PhD PROCEEDINGS

ANNUAL ISSUES OF THE DOCTORAL SCHOOL • 2014

FACULTY OF INFORMATION TECHNOLOGY & BIONICS

PhD PROCEEDINGS

ANNUAL ISSUES OF THE DOCTORAL SCHOOL
FACULTY OF INFORMATION TECHNOLOGY & BIONICS
PÁZMÁNY PÉTER CATHOLIC UNIVERSITY

PhD PROCEEDINGS

ANNUAL ISSUES OF THE DOCTORAL SCHOOL • 2014
FACULTY OF INFORMATION TECHNOLOGY & BIONICS

Pázmány University ePress
Budapest, 2014

HU ISSN 2064-7271

Kiadja a Pázmány Egyetem eKiadó
Budapest, 2014

Felelős kiadó
Ft. Dr. Szuromi Szabolcs Anzelm O. Praem.
a Pázmány Péter Katolikus Egyetem rektora

Készült a
TÁMOP-4.2.1.B-11/2/KMR-2011-0002 és a TÉT_10-1-2011-0058
projektek keretében, és az Új Széchenyi Terv támogatásával

Cover image by **Ágnes Polyák**, Fractalkine/Cx3CR1 system mediates macrophage (Mϕ) accumulation into adipose tissues during the development of obesity. Impaired fractalkine signaling prevents inflammation and macrophage accumulation into brown adipose tissue (BAT) and white adipose tissue (WAT), therefore macrophages do not express high levels of NPY, which enhances lipogenesis and decreases lipolysis, hence fractalkine receptor deficient mice do not gain excess body weight.

A borítón **Polyák Ágnes** ábrája látható: Az elhízás során a fraktalkin/fraktalkin receptor (Cx3CR1) rendszer szerepet játszik a makrofágok zsírszövetbe vándorlásában. A hibás fraktalkin jelátvitel csökkenti a gyulladás kialakulását, és a makrofágok fehér- és barna zsírszövetbe áramlását. A makrofágok kevesebb NPY-t termelnek, ami serkenti a lipogenezist, és csökkenti a lipolízist, így a fraktalkin receptor hiányos egerek nem híznak el.

Contents

Introduction 7

Program 1 • Bionics, Bio-inspired Wave Computers, Neuromorphic Models 9

Bálint Balázs • Symmetric light-sheet microscope for subcellular imaging 11

Erzsébet Farkas • Structural and functional characterization of the retrograde signaling system in the hypothalamic paraventricular nucleus 15

Katharina Hofer • A data analysis program for studying synchronous population activity in the human neocortex in vitro 19

Attila Gyula Jády • Metabolic changes during differentiation of neural stem cells 23

Balázs Knakker • Electrophysiological correlates of visual word form familiarity and crowding 27

András József Laki • Microvesicle fractionation from human blood using deterministic lateral displacement effect 31

Balázs Ligeti • Ranking drug combinations via logistic regression 35

Zselyke Magyari • In silico study on MAO enzymes: From structure to inhibitor specificity 39

Tamás Molnár • Movement related cortical potentials and oscillations on corticography 43

Dénes Pálfi • Spatially and temporally clustered inputs generate dendritic calcium spikes and local membrane oscillations with ripple frequency and in parvalbumin interneurons 47

Tamás Pardy • Chemical heating for non-instrumented nucleic acid amplification on lab-on-a-chip devices 51

Ágnes Polyák • Fractalkine/Cx3CR1 signaling in brown adipose tissue and diet induced obesity 55

Zsófia Sztyéhlikné Bérces • Neural cell response to nanostructured biosensor surfaces 59

Eszter Tóth • Diagnostic methods for microfluidics 63

Program 2 • Computer Technology Based on Many-core Processor Chips, Virtual Cellular Computers, Sensory and Motoric Analog Computers 67

Bernadett Ács • Computing linearly conjugate weakly reversible kinetic structures using graph theory and optimization 69

Vamsi Kiran Adhikarla • Fast and efficient data reduction approach for multi-camera light field display telepresence systems 73

Dóra Bihary • Modeling bacterial quorum sensing in open and closed environments 79

Bence József Borbély • A test platform for real-time measurement and analysis of myoelectric signals 83

Sándor Földi • Framework for examination the changes of arterial pulse signal waveform in time and space 87

Anna Gelencsér-Horváth • Using contour geometry as a merging cue in oversegmented images 91

László Kozák • Design of brain neuron signal recording system on UMC 0.18µ technology 95

Máté Németh • Compact, low-power, FPGA-based multi-camera vision system for real-time image processing 99

Norbert Sárkány • A mobile bio-signal measuring system on multiple limbs for rehabilitation purposes 103

Zoltán András Tuza • Analysis based parameter estimation of an in vitro transcriptional-translational system 107

Ádám Váły • Analyzing muscle synergies in cycling movements 111

Tamás Wilheim • Retrograde signalling in the GnRH network 115

Program 3 • Feasibility of Electronic and Optical Devices, Molecular and Nanotechnologies, Nano-architectures, Nanobionic Diagnostic and Therapeutic Tools 119

Imre Benedek Juhász • Simulation-based study of the spatial profile of fluorescence and photobleaching in two-photon-excited samples 121

Ádám Papp • Spin-wave based computing devices 127

Program 4 • Human Language Technologies, Artificial Understanding, Telepresence, Communication 131

Balázs Indig • Towards a psycholinguistically motivated performance-based parsing model 133

István Endrédy • Corpus driven research: ideas and attempts 137

Mátyás Jani • Experimental investigation of transitions for mixed speech and music playlist generation 141

Attila Novák • Vocabulary extension by paradigm prediction 145

Borbála Siklósi • Identifcation of abbreviations and multiword terms in clinical reports based on their contextual behavior 149

Zijian Győző Yang • Quality estimation of machine translated output without reference translation 153

Appendix 159

Introduction

It is our pleasure to publish this Annual Proceedings again to demonstrate the genuine multidisciplinary research done at the Jedlik Research Laboratories by young talents working in the Multidisciplinary Doctoral School of the Faculty of Information Technology and Bionics at Pázmány Péter Catholic University. The scientific results of our PhD students show the main recent research directions in which our faculty is engaged. Thanks are also due to the supervisors and consultants, as well as to the five collaborating National Research Laboratories of the Hungarian Academy of Sciences, the Semmelweis Medical School and the University of Pannonia. The collaborative work with the partner universities, especially, Katolieke Universiteit Leuven, Politecnico di Torino, Technische Universität München, University of California at Berkeley, University of Notre Dame, Universidad de Sevilla, Universita di Catania is gratefully acknowledged.

As an important development of this special collaboration, we—jointly with the Semmelweis Medical School—were able to accredit a new undergraduate curriculum on Molecular Bionics, the first of this kind in Europe.

We acknowledge the many sponsors of the research reported here. Namely,

- the Hungarian National Research Fund (OTKA),
- the Hungarian Academy of Sciences (MTA),
- the National Development Agency (NFÜ),
- the Gedeon Richter Co. Budapest, Hungary,
- the Office of Naval Research (ONR) of the US,
- NVIDIA Ltd.,
- Eutecus Inc., Berkeley, CA,
- MorphoLogic Ltd., Budapest, Hungary
- Analogic Computers Ltd., Budapest, Hungary
- AnaFocus Ltd., Seville, Spain

and some other companies and individuals.

Needless to say, the resources and support of the Pázmány Péter Catholic University is gratefully acknowledged.

Budapest, June 2014

Tamás Roska	Gábor Prószéky	Péter Szolgay
Head of the Jedlik Laboratory	Chairman of the Board of the Doctoral School	Head of the Doctoral School

Program 1

Bionics, Bio-inspired Wave Computers, Neuromorphic Models

Heads: Tamás Roska, Tamás Freund, György Karmos,
Zsolt Liposits, Sándor Pongor

Symmetric light-sheet microscope for subcellular imaging

Bálint Balázs
(Supervisors: Lars Hufnagel, Balázs Rózsa)
balint.balazs@embl.de

***Abstract*—To image subcellular processes in the context of a developing organism, high spatial and temporal resolution as well as low-light sensitivity is needed. Selective-plane illumination microscopy (SPIM) has proven to be a powerful imaging technique due to its unsurpassed acquisition speed and gentle optical sectioning. To increase resolution, most of these microscopes use multi-view image acquisition by rotating the sample, however rotation is not always possible (e.g. mouse embryos or cell cultures).**

To address this issue, we propose a symmetric SPIM setup using high NA objective lenses, where both objectives can be used for illumination and for detection, thus eliminating the need for sample rotation. The use of high NA objectives lenses, however, poses a geometrical problem, which we solve by implementing a non 90 degrees arrangement of the lenses and then optically tilting the light-sheet to restore focal plane illumination. Simulations show that this imaging system is capable of reaching close to isotropic resolution with 3-fold increase in axial resolution. This, combined with the high imaging speed also opens up the possibility of 3D imaging of large samples.

***Keywords*-light-sheet microscopy; multi-view deconvolution; isotropic 3D imaging**

I. Introduction

Development in early embryos is a highly dynamic process spanning over large scales Fundamental processes such as symmetry breaking in the early mouse embryo is still not fully understood despite [1]. To answer such questions, high resolution 3D imaging is required over different timescales: high imaging speed is necessary to capture subcellular processes, while imaging over long periods of time is needed to understand processes on the organismal scale.

Although confocal laser-scanning microscopy provides the necessary spatial resolution (at least in the lateral direction), it has many drawbacks for long-term live-imaging. A high power focused laser beam is scanned through the sample, imaging it point by point, which results in high phototoxicity and high photobleaching of the fluorophores. Acqusition speed is also a limiting factor, since the speed of many subcellular processes are to the scanning process.

Single-plane illumination microscopy [2] uses a thin light-sheet to illuminate the sample and capture a whole plane at once. This makes it two orders of magnitude faster than confocal microscopy, while also being gentle to the sample, sparing it from photodamage and bleaching. Using multiple views it's even possible to to image such an opaque specimen as a *Drosophila m.* embryo with sufficient quality to track all the individual nuclei through the different stages of development [3].

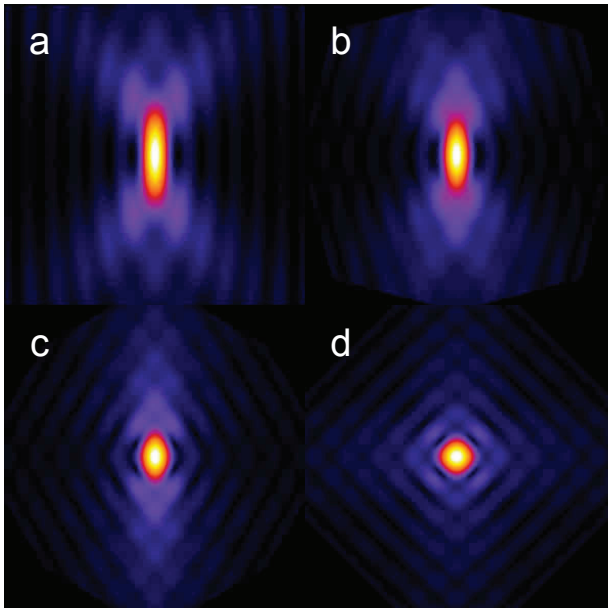


Fig. 1. Simulated point spread functions for different objective alignments. a) Simulatied PSF for a single objective. b)–d) Simulated compound PSF of two objectives in b) 30, c) 60 and d) 90 degrees to each other. Parameters used for calculation: NA=1.1, $\lambda_{ex} = 488$nm, $\lambda_{det} = 510$nm, $n = 1.333$ for water immersion.

II. Multi-view imaging

Even though selective plane illumination microscopy provides optical sectioning, and the illumination and detection are decoupled, it still suffers from significantly lower axial resolution compared to the lateral resolution. Fig. 1 a) shows the axial cross-section of a simulated PSF computed by using a scalar diffraction-limited model [4] with NA=1.1, $\lambda_{ex} = 488$nm, $\lambda_{det} = 510$nm for GFP, and n=1.333 for water. The ratio of axial and lateral resolution is 3.797, which results in a much worse resolution in the axial direction than in the lateral.

To circumvent this problem, it is possible to rotate the sample and image it from multiple directions, this way gaining more information from the different angles. In the case of a 90° rotation for example, the axial direction of the rotated view will correspond to the lateral direction of the original view, while the lateral direction of the rotated view will

B. Balázs, "Symmetric light-sheet microscope for subcellular imaging,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 11-14.

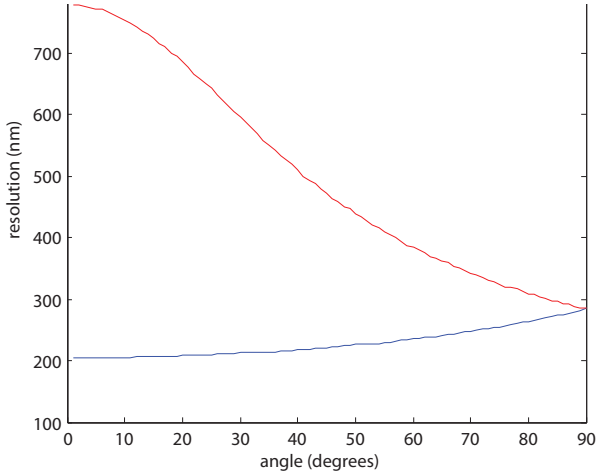


Fig. 2. Theoretical resolution depending on objective alignment

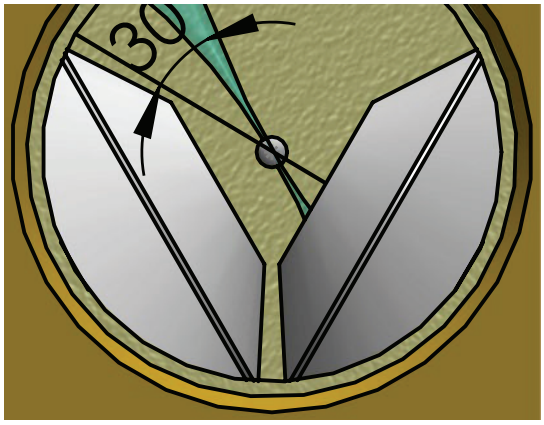


Fig. 3. Objective alignment. The focus of the two identical objectives coincide, while their optical axes are in 120 degrees. Light-sheet imaging in this configuration is still possible with a 30 degree tilted light-sheet.

correspond to the axial direction of the original view. This way we managed to gather high resolution information from our sample in all 3 dimensions. Fig. 1 b-d) shows the combined point spread function of two views with rotations of 30, 60 and 90 degrees.

The axial and lateral resolution defined as the full width half maximum of the simulated PSF was calculated for rotations between 0 and 90 degrees, and is depicted on Fig. 2. Above 60 degrees the difference between the axial and lateral resolution quickly diminishes, and at 90 degrees they become equal, resulting in a nearly isotropic image, and a 2.7-fold resolution gain in the axial direction.

Image processing steps

In order to actually gain from the rotation, several image processing steps are necessary before obtaining a single high quality image [5]. The two views need to be transformed to a common coordinate system (registered) and then they can be fused using several different methods. To be able to perform the image registration step accurately, the correct affine transformation matrix has to be found that will transform one of the views in the coordinate system of the other view. Ideally a simple rotation would be enough, but in practice translation, shearing and even scaling can occur.

A common method to find the appropriate transformation is to image fluorescent beads, which are much easier to align than the actual images themselves. Once the beads are aligned, the same transformation can be applied to the original data to perform the image registration.

After the registration there are several possibilities to perform the image fusion, but usually the Lucy–Richardson iterative deconvolution [6], [7] yields the best results. The iteration loop for the Lucy–Richardson algorithm can be summarized as

$$p_{k+1} = p_k \cdot h * \frac{p_0}{h * p_k} \tag{1}$$

where p_k is the estimate at iteration k, h is the PSF and $*$ denotes the convolution operator. This algorithm can be further improved by applying a regularizing step based on the total variation of the estimate, which will prevent the formation of ring artifacts [8]. Using this method simultaneously for the two views yields a single fused image which will have a superior resolution to either source images.

III. Symmetric light-sheet microscopy

To eliminate the need of rotation while still having multi-view capabilities, a symmetric light-sheet microscope can be used [9], [10]. In this case, instead of having two different objectives for illumination and detection, two identical objectives are used, and both of these can be used either for illumination or detection.

A microscope like this has numerous advantages over a conventional SPIM setup. Since the objectives are orthogonal, even without sample rotation it's possible to acquire orthogonal views from the sample, which then can be fused using the method mentioned earlier, resulting in a close to isotropic 3D resolution. Moreover, samples don't need to be mounted in an agarose cylinder, which results in easier sample handling, and also opens up the possibility to image specimens which are not compatible with the gel embedding procedure, such as mouse embryos, yeast, or cell culture.

The highest theoretical numerical aperture of such a system is

$$NA = n \cdot \sin\alpha = 1.333 \cdot 90/2 = 0.942 \tag{2}$$

which results in a lateral resolution of

$$d = 0.61 \cdot \frac{\lambda}{NA} = 0.61 \cdot \frac{510 \text{ nm}}{0.942} = 330.05 \text{ nm} \tag{3}$$

However, the objectives closest to this that are commercially available, have a numerical aperture of 0.8, which means the lateral resolution ideally is

$$d = 0.61 \cdot \frac{\lambda}{NA} = 0.61 \cdot \frac{510 \text{ nm}}{0.8} = 388.88 \text{ nm} \quad (4)$$

To image subcellular processes and maximize photon collection efficiency, higher NA is needed, since the light collection scales by the square of the numerical aperture. Using a higher NA objective however poses a geometrical problem, since arranging these orthogonally is not possible. Because of this reason, we propose a 120° objective alignment with a 30° tilted light-sheet (Fig. 3). Even though this will not result in an exactly isotropic resolution, with $d_{lat} = 235.5$ nm and $d_{axial} = 384.2$ nm lateral and axial resolution, and doubled light collection, this system can outperform the ones previously mentioned.

IV. Optical properties

To achieve the highest possible imaging speed we will use a cylindrical lens to generate the light-sheet, together with two Nikon Apo LWD 25X/1.1W objectives. The effective focal length of this objective is

$$f_O = \frac{f_{TL}}{M} = \frac{200 \text{ mm}}{25} = 8 \text{ mm} \quad (5)$$

and the back aperture diameter is 17.6 mm.

To generate the tilted light-sheet as shown on Fig. 3, the illumination beam will need to be displaced by

$$\delta = f_O \cdot \tan(30°) = 4.62 \text{ mm} \quad (6)$$

Since the Gaussian beam is not uniform, only a smaller portion of it can be used to maintain even illumination (Fig. 4 [10]). Because the size of an early mouse embryo is around 80 μm, we specify the length and the height of the light-sheet to be at least 100 μm.

The length and thickness of the light-sheet

The length of the light-sheet is determined by the Rayleigh-range of the beam in the zy plane. Since $l_{fov} = 2 \cdot z_{0,y} = 100\,\mu$m

$$z_{0,y} = 50\,\mu\text{m} \quad (7)$$

Since the Rayleigh range and beam width are coupled, the light-sheet thickness is:

$$2 \cdot W_{0,y} = 2 \cdot \sqrt{\frac{z_0 \cdot \lambda}{\pi}} = 5,57\,\mu\text{m} \quad (8)$$

when $\lambda = 488$ nm for GFP excitation.

From this, the divergence angle is

$$\theta_0 = \frac{\lambda}{\pi W_0, y} = 55.74 \text{ mrad} \quad (9)$$

This means, the numerical aperture needed to produce this light-sheet is:

$$NA_{ls} = n \cdot \sin(\theta_0) = 0.0743 \quad (10)$$

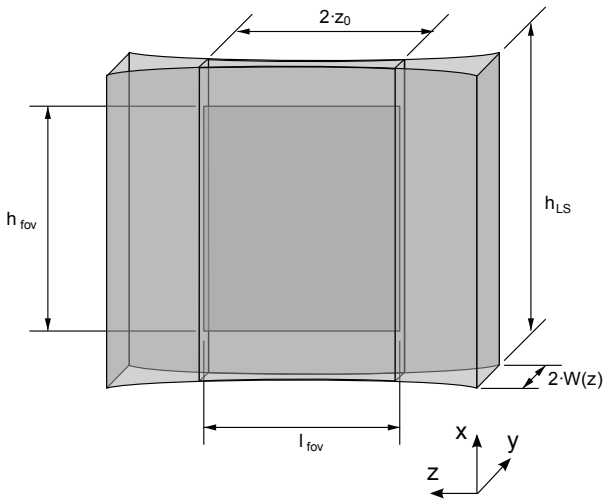


Fig. 4. Light-sheet dimensions. Light propagates in direction z, is focused in th zy plane but collimated in the zx plane. Length (l_{fov}) is determined by the Rayleigh-range in the zy plane ($z_{0,y}$), while height (h_{fov}) is determined by the beam waist in the plane zx ($W_{0,x}$). Adapted from [10].

Since $NA = 1.1$, and the diameter of the back aperture is $d = 17.6$ mm and $\theta_0 \ll 1$, using parallax approximation, the necessary beam width at the back focal plane in the y direction is

$$b_y = d \cdot \frac{NA_{ls}}{NA} = 1.19 \text{ mm} \quad (11)$$

Also, because the cylindrical lens focuses the beam only along x, the beam diameter before the cylindrical lens needs to be

$$b = b_y = 1.19 \text{ mm} \quad (12)$$

The height of the light-sheet

The height of the light-sheet is determined by the Gaussian intensity profile. To achieve an even illumination, the intensity at the sides of the FOV shouldn't be less than 80% of the intensity in the center. Thus,

$$I_s = I_0 \cdot exp\left(\frac{-2x^2}{W_{0,x}^2}\right) \quad (13)$$

where I_s is the intensity at the side of the FOV, I_0 is the intensity at the center, x is the distance from the center and $W_{0,x}$ is the beam waist radius in direction x. Since $I_s = 0.8 \cdot I_0$, and $x = 100\,\mu\text{m}/2 = 50\,\mu$m the necessary beam diameter at the focus is

$$b'_x = 2 \cdot W_{0,x} = 299.38\,\mu\text{m} \quad (14)$$

This corresponds to a (de)magnification of $m = b/b'_x = 0.252$, which means, the focal length of the cylindrical lens needs to be at most

$$f_{cyl} = \frac{f_O}{m} = 31.8 \text{ mm} \quad (15)$$

If the cylindrical focal length is longer than this, the height of the light-sheet will be smaller, resulting in a smaller field of

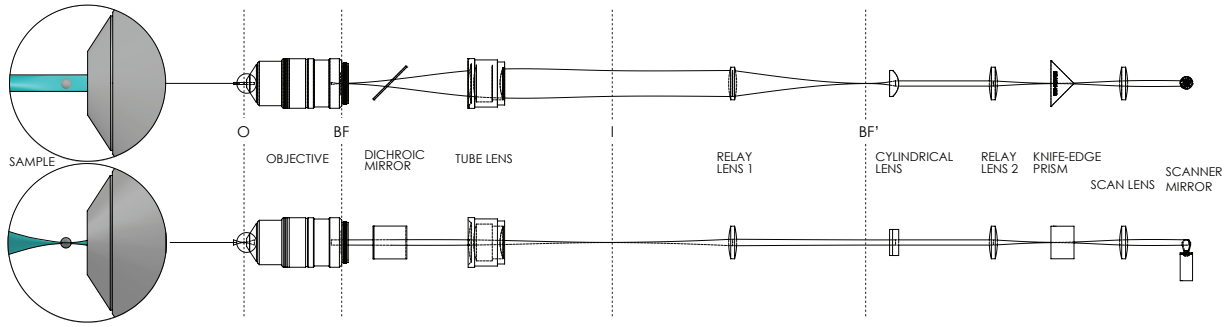


Fig. 5. Optics layout. Top: top view, bottom: side view. The light-sheet is generated using a cylindrical lens which focuses the illumination beam on the 1:1 conjugate plane of the objective back focal plane. A galvanometric (scanner) mirror is used to steer the illumination beam to the correct imaging branch and to quickly pivot the light-sheet to attenuate shadowing artifacts. O - object plane, BF - back focal plane, I - image plane, BF' - conjugate back focal plane

view. On the other hand, if it's smaller, the light-sheet height increases resulting in a more even illumination on the sample but also increasing the "wasted" laser power outside of it.

Practical considerations

A common artifact in light-sheet microscopy is the appearance of shadows, or stripes, due to the fact that the illumination beam can be absorbed by certain features in the sample. This can be greatly reduced by quickly pivoting the light-sheet, so the sample is illuminated from more directions during the exposure time [11]. The pivoting is performed by a galvanometric (scanner) mirror (Fig. 5) placed in the focus of the cylindrical lens.

To improve the flexibility of our setup, we introduced a relay lens system between the cylindrical lens and the scanner, and placed a knife-edge prism between the lenses. This way the same scanner can be used for destriping, and for switching between the two illumination arms.

Since the focal length of the cylindrical lens has to be relatively short [see equation (15)], we also relayed the back focal plane of the objective using two lenses, so it's easily accessible.

The components will be placed in a way that the light-sheet is already tilted, but in case if adjustment is needed, the tilt degree can be fine-tuned by translating the dichroic mirror with a micrometric screw.

V. Conclusion

In this paper we showed the importance of imaging live biological specimens with highest possible spatial resolution in all 3 dimensions, while collecting as much of the fluorescent emission as possible. High imaging speed is also imperative, since a majority of subcellular processes occur on a very short time scale, however, we also can not compromise on long term imaging, so we can observe organism-scale processes as well.

These requirements call for symmetric light-sheet imaging with high, 1.1 NA objectives, which will provide close to isotropic images at $d_{lat} = 235.5$ nm lateral and $d_{axial} = 384.2$ nm axial resolutions, which can be further improved using the Lucy–Richardson deconvolution method. The spatial constraints however, dictate to arrange the objectives in 120° relative to each other, which will require a tilted light-sheet illumination. Even though this will not result in achieving truly isotropic resolution, the high 3D resolution, light sensitivity and high-speed imaging capabilities of this microscope make it a prime candidate for future investigations in developmental biology.

Acknowledgment

I would like to thank my supervisor Lars Hufnagel, and Gustavo Q. G. de Medeiros and Nils Norlin for their help with my project.

References

[1] S. Wennekamp, S. Mesecke, F. Nédélec, and T. Hiiragi, "A self-organization framework for symmetry breaking in the mammalian embryo," *Nat Rev Mol Cell Biol*, vol. 14, no. 7, pp. 452–459, Jul. 2013.

[2] J. Huisken, J. Swoger, F. Del Bene, J. Wittbrodt, and E. H. K. Stelzer, "Optical sectioning deep inside live embryos by selective plane illumination microscopy," *Science*, vol. 305, no. 5686, pp. 1007–1009, 2004.

[3] U. Krzic, S. Gunther, T. E. Saunders, S. J. Streichan, and L. Hufnagel, "Multiview light-sheet microscope for rapid in toto imaging," *Nat Meth*, vol. 9, no. 7, pp. 730–733, Jul. 2012.

[4] P. A. Stokseth, "Properties of a defocused optical system," *J. Opt. Soc. Am.*, vol. 59, no. 10, pp. 1314–1321, Oct. 1969.

[5] U. Kržič, "Multiple-view microscopy with light-sheet based fluorescence microscope," Ph.D. dissertation, Ruperto-Carola University of Heidelberg, Germany, 2009.

[6] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *The Astronomical Journal*, vol. 79, p. 745, Jun. 1974.

[7] W. H. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Am.*, vol. 62, no. 1, pp. 55–59, Jan. 1972.

[8] M. Temerinac-Ott, O. Ronneberger, P. Ochs, W. Driever, T. Brox, and H. Burkhardt, "Multiview deblurring for 3-d images from light-sheet-based fluorescence microscopy," *Image Processing, IEEE Transactions on*, vol. 21, no. 4, pp. 1863 –1873, Apr. 2012.

[9] Y. Wu, P. Wawrzusin, J. Senseney, R. S. Fischer, R. Christensen, A. Santella, A. G. York, P. W. Winter, C. M. Waterman, Z. Bao, D. A. Colón-Ramos, M. McAuliffe, and H. Shroff, "Spatially isotropic four-dimensional imaging with dual-view plane illumination microscopy," *Nat Biotech*, vol. 31, no. 11, pp. 1032–1038, Nov. 2013.

[10] B. Balázs, "Development of a symmetric single-plane illumination microscope," Master's Thesis, Pázmány Péter Catholic University, Hungary, 2013.

[11] J. Huisken and D. Y. R. Stainier, "Even fluorescence excitation by multidirectional selective plane illumination microscopy (mSPIM)," *Opt. Lett.*, vol. 32, no. 17, pp. 2608–2610, 2007.

Structural and functional characterization of the retrograde signaling system in the hypothalamic paraventricular nucleus

Erzsébet Farkas
Supervisor: dr. Fekete Csaba
farkas.erzsebet@koki.mta.hu; fekete.csaba@koki.mta.hu

Abstract-Neuropeptide Y (NPY), the product of the orexigenic cell group of the arcuate nucleus (ARC) increases food intake and decreases energy expenditure. One of the main sites where NPY exerts its effect is the hypothalamic paraventricular nucleus (PVN). Results of our laboratory demonstrated that in addition to the regulation of gene expression of parvocellular neurons, NPY also influences the synaptic input of these cells via the activation of the endocannabinoid and nitric oxide (NO) retrograde signaling systems.

To understand the anatomical basis of this regulation, we have performed electron microscopic immunocytochemistry and quadruple-labeling immunofluorescence. Neuronal nitric oxide synthase (nNOS) was present in both pre- and postsynaptic sides of synapses in the PVN. The type 1 cannabinoid receptor (CB1) was found only in presynaptic terminals. The elements of both retrograde transmitter systems were present in a population of excitatory and inhibitory synaptic associations in the PVN.

Using quadruple-labeling immunofluorescence, we have observed that both vesicular glutamate transporter type 2 (VGLUT2) and NPY-immunoreactive (IR) axons formed frequently close association with the dendrites of parvocellular neurons in the PVN. The VGLUT2- and NPY-IR axons formed separate populations. The α1 subunit of the soluble guanylate cyclase (GCα1), an important component of the NO signaling, was observed in a population of VGLUT2-IR axon varicosities. NPY-, VGLUT2- and GCα1-containing inputs innervated the parvocellular neurons of the PVN. VGLUT2- and GCα1-containing boutons were located in the close proximity of NPY-IR axons on the surface of parvocellular neurons of PVN. Colocalization of NPY and GCα1 was not observed. Our data support, that NPY can regulate the synaptic input of parvocellular neurons via the endocannabinoid and NO retrograde-signaling systems.

Keywords-nitric oxide system; paraventricular nucleus; electron microscopy, confocal microscopy, endocannabinoid system

I. Introduction

The hypothalamic arcuate nucleus (ARC) plays a critical role in the regulation of energy homeostasis. Two main feeding-related neuronal groups are located in the ARC, the orexigenic neuron group that synthesizes neuropeptide Y (NPY) and agouti-related protein (AGRP) and the anorexigenic neuron population that synthesizes α-melanocyte stimulate hormone (α-MSH) and cocaine- and amphetamine- regulated transcript (CART). Both neuronal groups regulate the energy homeostasis and have antagonistic effects on the regulation of food intake and energy expenditure.

NPY, the product of the orexigenic cell group, increases food intake and decreases energy expenditure. One of the main sites where NPY exerts its effect is the hypothalamic paraventricular nucleus (PVN). Focal administration of NPY into the PVN results in marked increase of food intake and decrease of energy expenditure. In addition, the parvocellular neurons of the PVN are very densely innervated by NPY-containing axons of the ARC.

One effect of NPY on the neurons of the PVN is the regulation of the gene expression by the inhibition of cAMP synthesis. For example, NPY inhibits the synthesis of thyrotropin-releasing hormone (TRH) and corticotropin-releasing hormone (CRH) in these neurons.

Our group, however, recently raised the possibility that NPY may also regulate the parvocellular neurons of the PVN by modulating the synaptic input of these cells. This hypothesis was supported by patch clamp electrophysiology recordings showing that NPY can inhibit both the glutamatergic and GABA-ergic inputs of the PVN neurons via retrograde messenger systems and that the endocannabinoid [1] and nitric oxide systems [2] are involved in these processes.

Therefore, the goal of our current study was to understand the anatomical basis of these regulatory processes using multiple-labeling light and electron microscopic methods.

II. Material and Methodes

The experiments were carried out on ten adult, male, CD1 mice, weighing 30–35 g, housed under standard environmental conditions (light between 06:00 and 18:00 h, temperature 22±1 °C, mouse chow and water ad libitum). All experimental protocols were reviewed and approved by the Animal Welfare Committee at the Institute

E. Farkas, "Structural and functional characterization of the retrograde signaling system in the hypothalamic paraventricular nucleus," in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds. Faculty of Information Technology and Bionics, Pázmány Péter Catholic University. Budapest, Hungary: Pázmány University ePress, 2014, pp. 15-18.

of Experimental Medicine of the Hungarian Academy of Sciences.

Animals were deeply anesthetized with ketamine/xylazine (ketamine 50 mg/kg, xylazine 10 mg/kg body weight, ip). Five minutes later, the animals were perfused transcardially with 10 ml 0.01 M phosphate-buffered saline (PBS), pH 7.4, followed sequentially by 10 ml of 4% paraformaldehyde in Na-acetate buffer, pH 6.0, and then by 50 ml of 4% paraformaldehyde in Borax buffer, pH 8.5. The brains were rapidly removed and stored in 4% paraformaldehyde in 0.1 M phosphate buffer (PB), pH 7.4, for 24 h at 4 °C.

A. Quadruple-labeling immunofluorescence for NPY, VGLUT2, GCα1, MAP2

The sections were pretreated with 0.5% Triton X-100 and 0.5% H2O2 in 0.01M PBS for 15 min. Nonspecific antibody binding was blocked with 10% normal horse serum (NHS) in PBS for 15 min. The sections were then processed for quadruple labeling immunocytochemistry as described below. The sections were incubated in a mixture of rabbit anti-sGC α1 serum (1: 4,000), sheep anti-NPY serum (1:8000), mouse anti-MAP2 (1:500), as a neuronal marker, and guinea-pig anti VGLUT2 (1:500), as marker of glutamatergic neurons, diluted in 2% normal horse serum (NHS) in PBS for two days at 4°C. Then, the sections were incubated in the following mixture of secondary antibodies: Alexa 555-conjugated donkey anti-rabbit IgG (1:500), Alexa 488-conjugated donkey anti-mouse IgG (1:500), cy5 conjugated donkey anti sheep (1:500) and dylight 405 conjugated anti guinea-pig IgG (1:250) diluted at NHS in PBS at room temperature for 2 hours, followed by incubation. The sections were mounted onto glass slides and coverslipped with Vectashield (Vector) mounting medium. Images were taken using a Biorad Radiance 2000 confocal laser-scanning microscope (Bio-Rad Laboratories, Hemel Hempstead, UK) with 60X objectives using LaserSharp program. Image analysis was performed using ImageJ software. 488 nm for Alexa 488, 543 nm for Alexa 555, 637 nm for CY5, 405 nm for dylight 405 and dichroic/emission filters, 560 nm/500–530 nm for Alexa 488 and 570–590 nm for Alexa 555, 660 nm long pass filter for CY5, 420-480 nm for dylight 405. All images shown represent a single optical section (less than 0.8 μm thick) that were captured through 60x oil lens.

B. Immuno-electronmicroscopy

Serial 25μm thick coronal sections were cut on a Leica VT 1000S vibratome (Leica Microsystems, Wetzlar, Germany) throughout the part of hypothalamuc containing the PVN. The sections were treated with 0.5% H2O2 in PBS for 15 min. The sections were cryoprotected in 15% sucrose in PBS for 15 min at room temperature and in 30% sucrose in PBS overnight at 4°C and then, quickly frozen over liquid nitrogen and thawed three times to improve antibody penetration into the tissue. To detect the nNOS-immunoreactivity and CB1-immunoreactivity, pretreated sections were then placed in a mixture of rabbit anti-nNOS serum (1:200) and sheep anti-CB1 serum (1:800) for 4 days at 4 °C and after rinsing in PBS and 0.1% cold water fish gelatin/1% bovine serum albumin (BSA) in PBS, incubated in donkey anti-rabbit IgG conjugated with 0.8 nm colloidal gold (Electron Microscopy Sciences, Fort Washington, PA) diluted at 1:100 and biotinylated donkey anti-sheep IgG diluted at 1:500 in PBS containing 0.1% cold water fish gelatin and 1% BSA. After rinsing in 0.2 M sodium citrate, pH 7.5, the gold particles were silver intensified with the Aurion R-Gent SE-LM Kit followed by treatment in avidin-biotin-peroxidase complex (ABC Elite 1:1000). The CB1-immunoreactivity was detected in 0.05% DAB/0.15%Ni-ammonium-sulfate/0.005% H2O2 in 0.05 M Tris buffer, pH 7.6. Sections were osmicated, and then treated with 2% uranyl acetate in 70% ethanol for 30 min. Following dehydration in an ascending series of ethanol and propylene oxide, the sections were flat embedded in Durcupan ACM epoxy resin (Fluka) on liquid release agent (Electron Microscopy Sciences)-coated slides, and polymerized at 56 °C for 2 days.

After immunostaining, the sections were embedded in Durcupan ACM epoxy resin (Fluka) and then 60–70 nm thick utlrasections were cut with Leica ultracut UCT ultramicrotome (Leica Microsystems, Wetzlar, Germany). The ultrathin sections were mounted onto Formvar-coated single slot grids, contrasted with 2% lead citrate and examined with a Jeol-100 C transmission electron microscope.

C. Specificity of antisera

The specificity of nNOS, sGC α1 antisera was reported previously (Szabadits et al., J Neurosci, 2007). The specificity of the goat rabbit CB1 antiserum was described elsewhere (Makara JK et al., J Neurosci 2007). The specificity of NPY antisera was reported previously (Wittman et al., Neurosci Lett 2003).

TABLE 1. LIST OF ANTISERA

Primary antisera	Dilution
rabbit antiserum against nNOS (Zymed Laboratories, San Francisco, CA)	1:200
rabbit antiserum against soluble guanylyl cyclase α1 (catalog number G4280, lot number 011K4888; Sigma)	1:4000
mouse antiserum against MAP2 (mouse monoclonal antibody, cat. no #MAB3418, lot no #2062312, Millipore)	1:500
sheep antiserum against CB1 (C31; Watanabe, Kazanawa Uni, Japan)	1:800

guinea pig antiserum against VGLUT2 (C34aa; Watanabe, Kazanawa Uni, Japan)	1:500
sheep antiserum against NPY (FJL no. 14/3A; Merchenthaler, Wyeth Research, Collegeville, PA)	1:8000

III. Results

Using quadruple-labeling immunofluorescence detecting VGLUT2 as a marker of glutamatergic terminals, NPY, sGCα1 and MAP2 as a neuronal marker, we have observed that both VGLUT2- and NPY-IR axons frequently formed close association with the dendrites of MAP2-IR elements. The VGLUT2- and NPY-IR axons formed separate populations. Occasionally, sGCα1 and VGLUT2-immunoreactivity is observed in the same axon varicosities (Fig. 1). VGLUT2-, NPY-, sGCα1-IR axon varicosities were found in juxtaposition to the parvocellular neurons of the PVN (Fig. 1). NPY-, VGLUT2- and GCα1-containting inputs innervated the parvocellular part of the PVN. VGLUT2- and sGCα1-containting boutons were located in the proximity of NPY-IR varicosities on the surface of parvocellular neurons of PVN (Fig. 1). Coloclization of NPY- and sGCα1- or NPY and VGLUT2-immunoreactivity was not observed. sGCα1-immunoreactivity was also observed in dendrites and numerous cell bodies.

Ultrastructural examination of double-labeled preparations demonstrated that nNOS-immunoreactivity was abundantly present in neuronal perikarya (Fig. 2 A) and dendrites (Fig. 2 A-C) and also in axon varicosities (Fig. 2 C) in the parvocellular part of the PVN. In perikarya and dendrites, nNOS-immunoreactivity was widely distributed in the cytoplasm, primarily associated with the endoplasmatic reticulum. nNOS-immunoreactivity was also found to be associated with the perikaryal plasma membrane in close proximity to both symmetric and asymmetric synapses (Fig. 2 A, C), as well as within axon varicosities forming both symmetric and asymmetric type synapses (Fig. 2 C).

The cannabinoid receptor type 1 (CB1) was found in axon varicosities forming synapses on perikarya and dendrites of parvocellular neurons. CB1 was found in both excitatory (forming asymmetric synapse) and inhibitory (forming symmetric synapse) axon terminals (Fig. 2 A-C). In many instances nNOS-immunoreactivity (nNOS-IR) was observed in the close proximity of the postsynaptic side of the synapses (both symmetric and asymmetric type) formed by CB1-IR terminals.

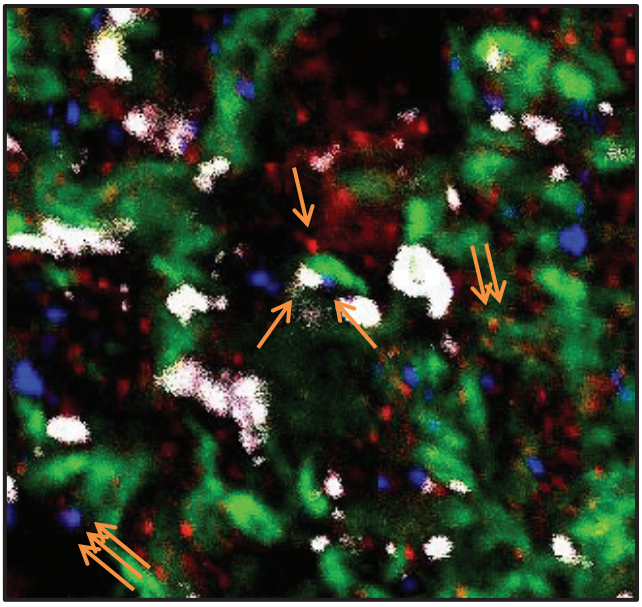


Figure 1. High-magnification confocal image demonstrate the dense VGLUT2- (blue) and NPY-IR (white) innervation of the dendrites of the parvocellular neurons of the PVN that was labelled using MAP2 as a neuronal marker (green), and the GCα1 (red) content of these dendrites and the axons terminating these neuronal profiles. NPY-, VGLUT2- and GCα1-containting boutons were located near to each other and juxtaposed to the parvocellular dendrites, (indicated by arrowheads). VGLUT2 is observed in axon varicosities and cocontained with GCα1 (indicated by three arrows). GCα1is observed mainly in axon varicosities, occasionally in dendrites and also is present in numerous cell bodies (indicated by two arrows) Scale bars=10µm

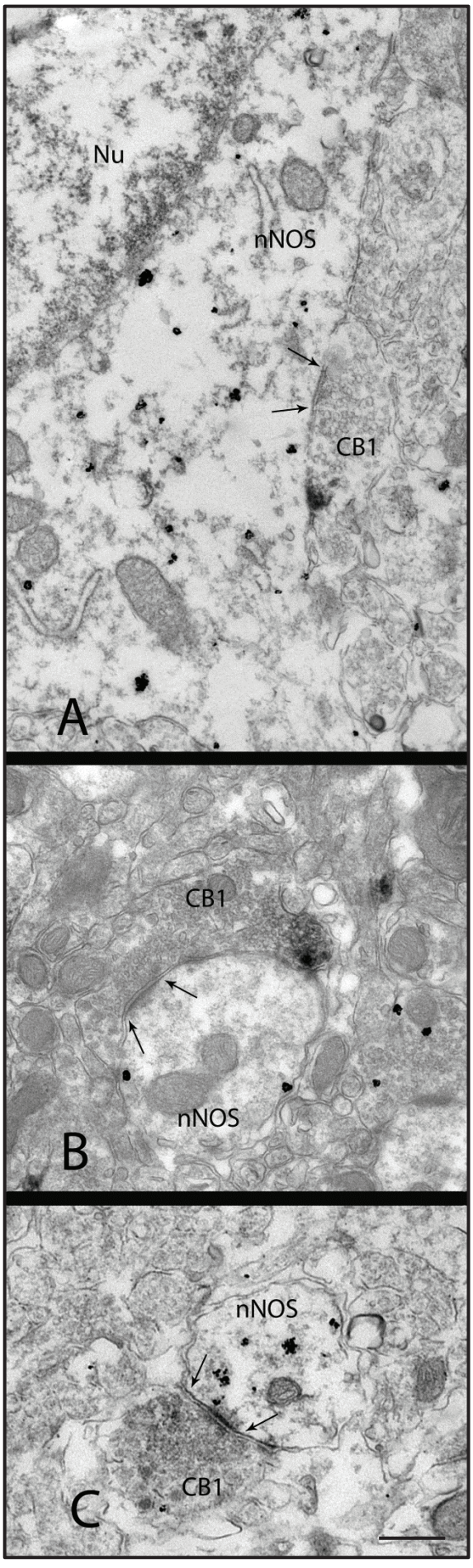


Figure 2. Electron micrographs illustrate the localization of nNOS- (highly electron dense gold-silver particles) and CB1-immunoreactivity (electron dense NiDAB precipitate) in the parvocellular division of the hypothalamic paraventricular nucleus (PVN). nNOS immunoreactivity is widely distributed in the cytoplasm of a parvocellular neurons (A) and primarily associated with the endoplasmatic reticulum. CB1-immunoreactive (IR) axon forms a symmetric type synapse on the surface of the nNOS-IR perikaryon. Silver grain denoting nNOS-immunoreactivity is located in the close proximity of the postsynaptic side of the synapse of CB1-IR terminal (arrowhead). nNOS-immunoreactivity (arrowhead) can also be observed in dendrites in the proximity of the postsynaptic site of both symmetric (A) and asymmetric (B, C) type synapses of CB1-IR terminals. Arrows point to edge of synapses. Scale bars=0,5µm in (A-C). Nu= nucleus

IV. Conclusion

nNOS-immunoreactivity is present in both the pre- and postsynaptic sites of inhibitory (symmetric) and excitatory (asymmetric) synapses as well.

CB1 immunreactivity is present in axon terminals forming both inhibitory and excitatory synapses.

Both the nitric oxide and the endocannabinoid systems are present in a population of axon terminals forming synapse on the parvocellular neurons.

Near to the NPY inhibitory inputs of the parvorcellular part of the PVN located excitatory inputs including VGLUT2-containing inputs and nearly located also the element of the nitric oxide system, CGα1.

Our data suggest that nitric oxide system and endocannabinoid system can be work simultaneously together in the same synapses in the parvocellular part of the PVN in mice and they can be mediate the inhibition effect of NPY on the extitatory and inhibitory inputs of the PVN neurons.

V. Acknowledgments

The authors are grateful to Dr. István Merchenthaler for providing the NPY antiserum for our studies.

VI. References

[1] Quarta C et al. (2010). CB(1) signaling in forebrain and sympathetic neurons is a key determinant of endocannabinoid actions on energy balance. Cell Metabolism, 273-85.

[2] Jin-Dong Ding, A. B. (2004). Distribution of Soluble Guanylyl Cyclase in the Rat Brain. The journal of comparative neurolog , 437-448.

A data analysis program for studying synchronous population activity in the human neocortex in vitro.

Katharina Hofer
(Supervisor: Ulbert István)
hofer.katharina@ttk.mta.hu

***Abstract*—In vitro human studies of cortical network activity employ tissue that was surgically removed from epileptic patients or brain tumour patients for treatment reasons. Similar to interictal spikes seen on scalp EEG recordings of epileptic patients, spontaneous population activity (SPA) was also observed in vitro in human neocortical slice preparations.**

Cellular and network properties of SPA were recorded from neocortical slices using a 24 channel laminar microelectrode. SPA could be observed in epileptic as well as in non-epileptic neocortical tissue. To better analyse the local field potential gradient (LFPg), multi-unit activity (MUA) and clustered single neuron activity, a program was written in C++.

Further refinements of the analysis will improve our understanding of the subtle border between physiological (non-epileptic) and pathological (epileptic) neuronal population activity.

***Keywords*-neuronal population activity; data analysis; laminar multielectrode; neocortical slice preparation; neuroscience; human**

I. INTRODUCTION

Cortical neural network activity and its pathology are extensively studied in vitro and in vivo in various animal models [1-4]. Epilepsy is thought to be related to hyperactivity of neuronal circuits. Pharmacological treatment is often effective, but a significant number of patients resists pharmacotherapy. Surgical tissue removal in these patients offers a remarkable possibility to study living human tissue known to be intimately involved in the generation of this neurological disorder. In this study, neocortical tissue obtained from resection surgery on brain tumour patients served as non-epileptic control tissue.

Similar to interictal spikes recorded on the scalp EEG [5], spontaneous population activity (SPA) could be observed in vitro in human epileptic neocortical [6] and hippocampal [7-10] slice preparations in a physiological perfusion solution.The preliminary data suggest that SPA was not only generated in neocortical slices of epileptic patients but also in patients with tumour but without epilepsy. To study subtle differences in the occurrence and appearance of SPA, a program was developed in C++ which is able to analyse large amounts of data unbiasedly and more accurately and quicker than if performed by the experimenter.

Supported by OKTA K81354, OKTA PD77864, ANR-TÉT Neurogen, ANR-TÉT Multisca, TÁMOP-4.2.1.B-11/2/KMR-2011-0002, Bolyai Research Fellowship (toLucia Wittner), PhD student grant by the Hungarian government (to KH)

II. METHODS

A. Tissue preparation

Postoperative neocortical human tissue was obtained from epileptic or tumour patients during brain surgery and immediately transferred into ice cold, oxygenated cutting solution (256mM sucrose, 10mM D-glucose, 25mM NaHCO3, 1mM KCl, 1mM CaCl2, 10mM MgCl2, phenol red, saturated with carbogen gas: 95% O_2, 5% CO_2). The tissue was cut into slices (500µm), perpendicular to the cortical layers. The slices were transferred to an interface chamber, where they were exposed to artificial cerebrospinal fluid (ACSF; 124mM NaCl, 10mM D-glucose, 25mM $NaHCO_3$, 3.5mM KCl, 1mM $CaCl_2$, 1mM $MgCl_2$; saturated with 95% O2, 5% CO2) at 33°C. The slices were equilibrated for 1h before recording.

B. Recording

For extracellular recordings, a 24 channel laminar multi-electrode was placed on the surface of the neocortical slice. Each of the contacts of the electrode served as the reference for the next contact, resulting in a spatial array of 23 channels, the local field potential gradient (LFPg). A custom made voltage gradient amplifier of pass-band 0.01Hz to 10kHz was used [11-14]. The data were recorded at 20 kHz sampling rate.

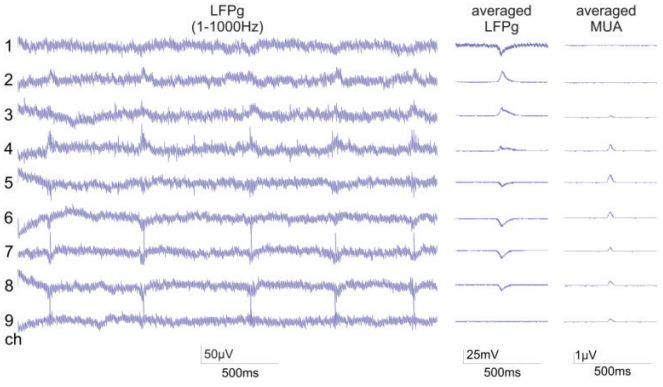


Figure 1. LFPg excerpt and averages reflecting SPA

C. Data analysis

SPA was detected on the optimal channel using routines written for MatLab (The MathWorks, Natick, MA, USA) [12-14], using an amplitude threshold of 3x the standard deviation after applying Hamming-window spatial smoothing and filtering from 3-30Hz. The occurrences of the SPAs were saved

K. Hofer, "A data analysis program for studying synchronous population activity in the human neocortex in vitro,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 19-22.

in an event file which was subsequently used to average the detected SPAs using the NeuroScan Edit 4.3 program (Compumedics NeuroScan, Charlotte, NC, USA). For the MUA analysis, the averaging was performed on high pass filtered (500Hz) and rectified data. The left panel of Figure 1shows a short (2.7 s) excerpt from a longer recording from a human neocortical slice showing recurring SPA. The middle and right panel shows the LFPg and MUA averaged for SPAs over the whole recording.

Neuronal firing was detected and clustered into single cell activity using routines written for MatLab. The occurrences of the cell spikes were also saved in an event file and averaged using NeuroScan to retrieve the averaged action potential.

III. The Analysis Program

Pre-existing code from Daren Weber (2000) was used for loading the average files into an STL container within the C++ program. The SPA, MUA and action potential amplitudes and widths were measured and the respective event files analysed separately as well as set in relation to each other.

A. SPA and MUA peak detection

After loading the SPA average file into an STL container, the SPA peak was searched across all 23 channels. Note that in case of the SPA this peak could have a negative or positive amplitude, which is why this will be referred to as an extreme search rather than a maximum search. However, depending on the number of SPA events averaged, the average could be quite noisy and the point with the largest amplitude was not always a good fit for the detection of the peak of the SPA. Thus, the extreme was searched for the sum of all data points within a 2.5ms shifting window (seeFigure 2).

After the sum of the amplitudes for the points was calculated for a given window position, the window was shifted by one data point and the sum of points was calculated again. Figure 2 shows three example windows (blue horizontal lines) on top of an example peak from one channel of an SPA average file. Note that the number of data points was reduced for the purpose of this illustration. As the acquiring frequency was 20 kHz, 50 data points were taken into consideration for each 2.5ms time window. A linear maximum search for the absolute value was employed to find the channel, time point and amplitude of the largest SPA peak (positive or negative) in the average file. The time point of the SPA peak was defined by the middle of the time window, whereas the amplitude was defined as the median amplitude of the data points in the time window.In Figure 2 the window achieving the largest sum of amplitudes is shown in dark blue, with the resulting amplitude and time point of the peak are indicated by a small black cross. The inset shows a larger portion of the same SPA average, indicating the amplitude and time point of the peak with a black cross.

The measurement of the MUA amplitude worked analogous to the SPA measurement with the exception that only positive peaks were possible due to the rectification of the data after the application of the 500Hz high pass filter.

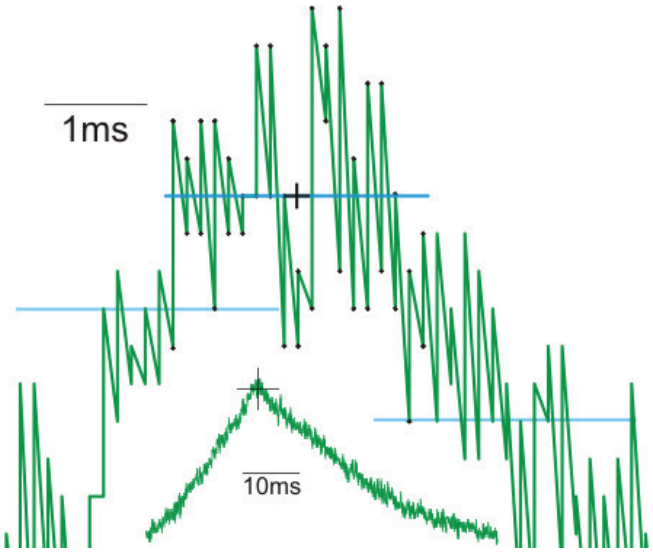


Figure 2. SPA and MUA peak detection

B. SPA and MUA peak width

The width of the SPA or MUA was calculated at various heights of the peak on the channel on which the largest peak was detected. To measure the width, the height at which the width was measured was defined as a fraction of the amplitude of the peak. Figure 3 shows the measurement at 50% of the peak amplitude (half maximal amplitude) relative to the baseline.

Starting from the peak (vertical dashed line in Figure 3), the data points were sampled outwards until a stopping threshold was reached. This threshold (blue horizontal line) was defined at 33% of the width measuring height (orange horizontal line). While the data points were sampled, each intersection with the width measuring height was stored (black dots). The data point used for calculating the width wais the one directly outside the median intersection (median intersection: yellow dot with black surrounding; data point used: black dot with blue surrounding.Note that the amplitude of the black doth with blue surrounding was irrelevant).

After the width was calculated for both sides of the peak, the width of the SPA was defined as the sum of the two sides, whereas the asymmetry of the SPA was defined as the ratio of the right side to the left side.

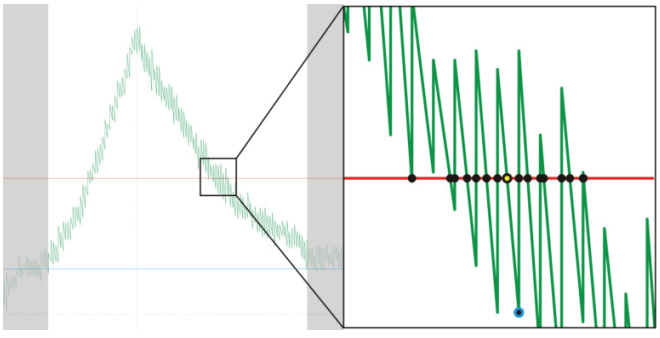


Figure 3. SPA width measurement at half maximal amplitude.

C. Action potential measurements

In case of measuring the features of an action potential, peaks were detected using extreme searches without using shifting time widows as described for the SPA and MUA. While the baseline was the reference for measuring the amplitude in case of SPA and MUA waves, itwas just one of the references used for measuring the shape of an action potential (Figure 4A, B and C). In addition, the minimum left for a positive action potential (maximum left for a negative action potential)was also taken as a reference (Figure 4 D).

In addition, the minimum right of the action potential (or maximum, depending on the direction of the action potential), called the afterhyperpolarisation, was also assessed (delay and amplitude: Figure 4E and F).

As the acquiring frequency was 20 kHz, the measuring points are 0.05ms apart from each other. While this is negligible for calculating SPA or MUA widths, it is a relevant difference when measuring the widths of action potentials. Therefore, the intersections (black dots with blue surrounding in Figure 5) of the action potential with the defined height (orange and blue horizontal line inFigure 5)were intrapolated between the points before and after the intersections (black dots in Figure 5). The width was calculated separately for the left and right side of the peak and used to calculate the width as well as the asymmetry of the action potential at the respective height.

The measurements of width and asymmetry at various heights of the action potential help to distinguish between the two major cell types in the neocortex: principal cells and interneurons. This permits conclusions about the balance between excitation and inhibition within the neuronal network, which is a major interest in the study of epileptic processes.

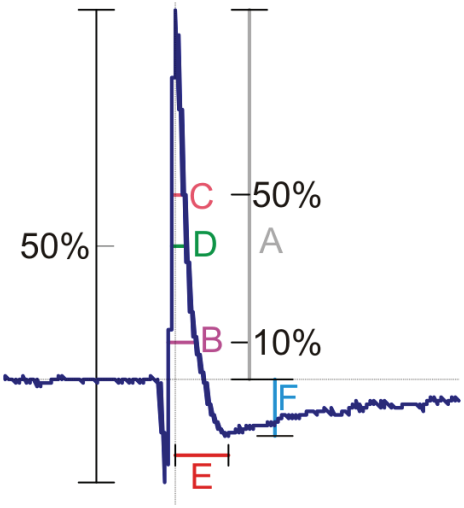


Figure 4. Action potential measurements

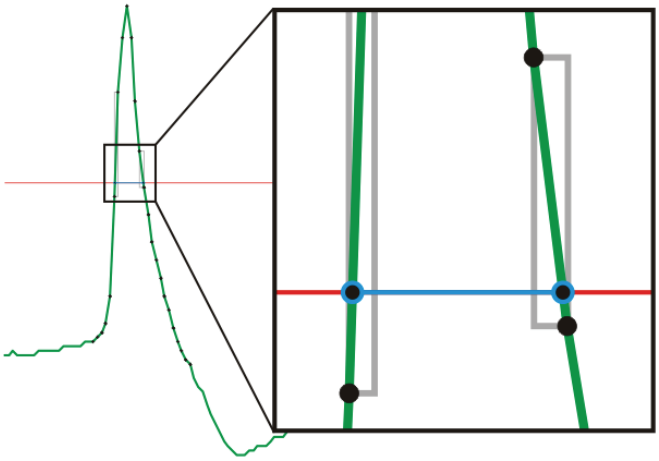


Figure 5. Action potential width intrapolation

D. Occurrence parameters

Event files contain information on when the SPAs and cell spikes occurred. From them, the average firing frequency, the median interspike interval and a histogram forthe interspike intervals were computed. For event files describing the firing properties of cells, an algorithm calculated the number of spikes contained within firing bursts, a burst being defined as a sequence of at least 3 action potentials with each interspike interval below 8ms [15]. For the purpose of excluding cells firing tonically at high frequencies, an upper limit of 15 spikes was added to the definition of a burst.

E. Correlation of cell firing and population activity

It is especially interesting to compare clustered cell activity to the occurrence pattern of SPAs. The program generated histograms for each SPA to each cell clustered from the same recording. The sweeps of the histogram were defined by the occurrence of the SPA, the cell spikes were sorted into the histogram bins. However, SPAs from different recordings can have different shapes. Thus, to be able to compare cell firing relative to the phase of the SPA, the bin size of the histogram was variable and adjusted for each SPA.

For this purpose, the widths left and right of the SPA peak were calculated at 33% height (see section III.B). They were then multiplied by 4 to define the area covered by the histogram (see Figure 6). The left and right sideswere each divided into 50 bins.

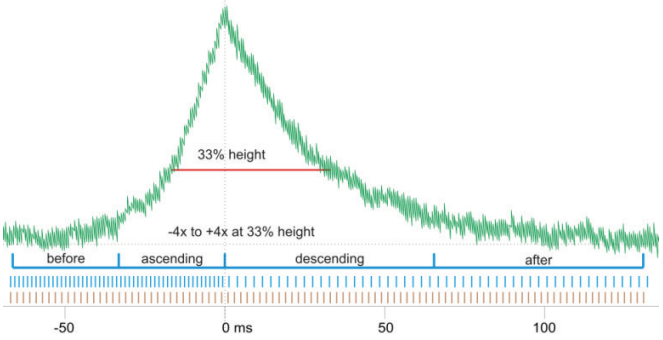


Figure 6. Maximal firing histogram: sweep and bin size definition

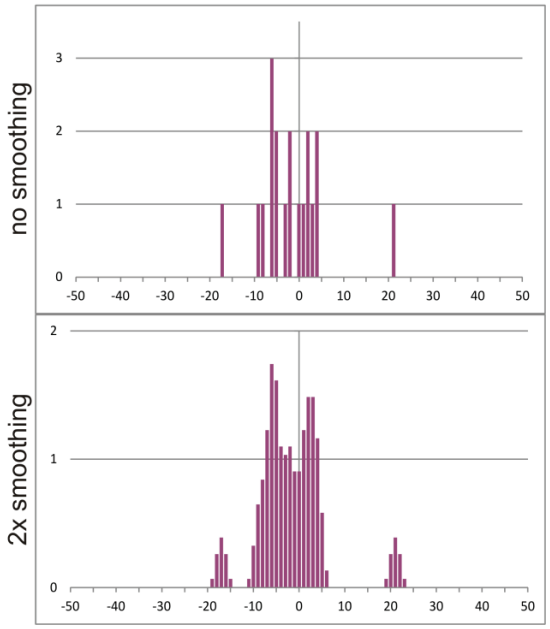


Figure 7. Maximal firing histogram: smoothing

The next step was to sort the cell spikes into the bins of the histogram. Note, that the sizes of the bins differed between the left and the right side of the peak (blue bins in Figure 6) and between histograms for different SPAs.However, this setup of the histogram allows for the comparison of histograms derived from different recordings. A similar histogram was generated using bins that werenot depending on the shape of the SPA. Those bins had a fixed size of 2ms (brown bins in Figure 6). As most of the SPAs were asymmetric, this means that the numbers of bins on the two sides of the peak were not equal.

As small changes in the exact position or width of the bins can change the appearance of a histogram substantially, at least 2 smoothing steps were performed to obtain a reproducible representation of the cells' firing pattern. During each smoothing step, half of the value of each bin was distributed over its neighbours. Thus, 3 neighbouring bins of the values 0, 1 and 0 thus resulted in 0.25, 0.5 and 0.25 after one smoothing step. At least two such smoothing steps were performed (see Figure 7). Afterwards, the program searched for the bin with the highest value. If there was no clear maximum, additional smoothing steps were performed until a unique maximum was found. In the case of the histogram with bin sizes relative to the SPA shape, this time point of maximal firing of a cell can be compared between cells, in relation to different SPAs and across recordings.

IV. Conclusion

The described program offers a possibility to analyse various features of SPA and clustered single neuron activity. It provides a way of correlating cell firing to population activity and proved useful in preliminary data analysis. It will be further developed to investigate more aspects of this type of neuronal population activity.

Acknowledgment

The author wishes to acknowledge Dr. István Ulbert and Dr. Lucia Wittner for excellent supervision as well as Dr. Kinga Tóth and Dr. Dániel Fabó for collaboration and advice as well as the neurosurgeons Dr. Attila Bagó, Dr. Loránd Erőss and Dr. László Entz from the National Institute of Clinical Neurosciences (OKITI) for their collaboration.

References

[1] V. Bouilleret, F. Loup, T. Kiener, C. Marescaux and J.M. Fritschy, "Early loss of interneurons and delayed subunit-specific changes in GABA(A)-receptor expression in a mouse model of mesial temporal lobe epilepsy," Hippocampus 2000, 10: pp.305-24.

[2] Z. Maglóczky and T.F. Freund, "Selective neuronal death in the contralateral hippocampus following unilateral kainate injections into the CA3 subfield," Neuroscience 1993, 56: pp.317-35.

[3] J.O. McNamara, M.C. Byrne, R.M. Dasheiff and J.G. Fitz, "The kindling model of epilepsy: a review," Prog Neurobiol 1980, 15: pp.139-59.

[4] L. Turski, E.A. Cavalhiero, M. Sieklucka-Dziuba, C. Ikonomidou-Turski, S.J. Czuczwar and W.A. Turski, "Seizures produced by pilocarpine: Neuropathological sequelae and activity of glutamate decarboxylase in the rat forebrain," Brain Res. 1986, 398: pp.37-48.

[5] M. de Curtis and G. Avanzini, "Interictal spikes in focal epileptogenesis," Prog Neurobiol 2001, 63: pp.541-67.

[6] R. Kohling, A. Lucke, H. Straub, E.J. Speckmann, I. Tuxhorn, P. Wolf et al, "Spontaneous sharp waves in human neocortical slices excised from epileptic patients," Brain 1998, 121 (Pt 6): pp.1073-87.

[7] I. Cohen, V. Navarro, S. Clemenceau, M. Baulac and R. Miles, "On the origin of interictal activity in human temporal lobe epilepsy in vitro," Science 2002, 298: pp.1418-21.

[8] G. Huberfeld, L. Wittner, S. Clemenceau, M. Baulac, K. Kaila, R. Miles et al, "Perturbed chloride homeostasis and GABAergic signaling in human temporal lobe epilepsy," J Neurosci 2007, 27: pp.9866-73.

[9] L. Wittner, G. Huberfeld, S. Clemenceau, L. Erőss, E. Dezamis, L. Entz et al, "The epileptic human hippocampal cornu ammonis 2 region generates spontaneous interictal-like activity in vitro," Brain 2009, 132: pp.3032-46.

[10] C. Wozny, A. Knopp, T.N. Lehmann, U. Heinemann and J. Behr, "The subiculum: a potential site of ictogenesis in human temporal lobe epilepsy," Epilepsia 2005, 46 Suppl 5: pp.17-21.

[11] I. Ulbert, E. Halgren, G. Heit and G. Karmos, "Multiple microelectrode-recording system for human intracortical applications," J Neurosci Methods 2001, 106: pp.69-79.

[12] I. Ulbert, G. Heit, J. Madsen, G. Karmos and Halgren E, "Laminar analysis of human neocortical interictal spike generation and propagation: current source density and multiunit analysis in vivo," Epilepsia 2004, 45 Suppl 4: pp.48-56

[13] I. Ulbert, Z. Maglóczky, L. Erőss, S. Czirják, J. Vajda, L. Bognár, S. Tóth, Z. Szabó, P. Halász, D. Fabó, E. Halgren, T.F. Freund and G. Karmos, "In vivo laminar electrophysiology co-registered with histology in the hippocampus of patients with temporal lobe epilepsy," Exp Neurol 2004, 187: pp.310-318.

[14] D. Fabó, Z. Maglóczky, L. Wittner, A. Pék, L. Erőss, S. Czirják, J. Vajda, A. Sólyom, G. Rásonyi, A. Szűcs, A. Kelemen, V. Juhos, L. Grand, B. Dombovári, P. Halász, T.F. Freund, E. Halgren, G. Karmos, I. Ulbert, "Properties of in vivo interictal spike generation in the human subiculum," Brain 2008, 131: pp.485-99.

[15] Ranck, J.B.,Jr.,"Studies on single neurons in dorsal hippocampal formation and septum in unrestrained rats: I. Behavioral correlates and firing repertoires." Experimental Neurology1973, 41, pp. 461-531.

Metabolic changes during differentiation of neural stem cells

Attila Jády
(Supervisor: EmíliaMadarász)
jah@net-home.eu

***Abstract*—According to previous results [1], neural stem cells survive at much lower oxygen supply than neurons, both in vivo and in vitro. In order to understand the diverse O_2-demand, metabolic analyses were carried out on one-cell derived populations of neural stem cells representing progenitors of the neural plate /early neural tube (NE-4C) [2] and the adult neurogenic zones (HC_A and SVZ_M) [3] and also on primary neuronal cultures isolated from the embryonic mouse forebrain. Depending on origin and developmental stages, different stem cells displayed different responses to "starvation" and to supplementing the medium with single metabolites (glucose, lactate, β-OH-butyrate, amino acids). The data indicated that the basic metabolism shifts with the advancement of neural differentiation, and the metabolic profile reflects the origin and stage of differentiation of neural cells**

***Keywords*: neural stem cell; metabolism; neuronal differentiation; oxygen consumption**

I. Introduction

Previous data indicated that neural stem cells and their differentiating progenies require significantly different environment for survival. Besides the needs for growth factors, adhesive surfaces and cell activation patterns, the changes in metabolism play important roles in decision on integration or decay of young neural cells in the course of development, regeneration and physiological neuron-replacement.

Under hypoxic ($[O_2] \leq 1$ (v/v)%) conditions, neural stem cells survive and proliferate but can not differentiate; under hypoxic conditions, committed neural precursors and maturing neurons die [1]. The composition of mitochondrial membranes is also changing during the formation of neurons as it was shown by the presence of TSPO 18 kDa (PBR; peripheral benzodiazepine receptor) in stem cells and early neuronal progenitors [4], but not in mature neurons.

Biochemical reasons underlying developmentally regulated changes of the metabolic machinery are not understood.

In order to explore biochemical processes behind differentiation-dependent metabolic changes, the O_2-consumption of one-cell derived neural stem cell populations representing progenitors of early brain vesicles (NE-4C) [2], adult neurogenic zones (HC_A and SVZ_M) [3], and primer neurons were investigated. In vitro induced neural differentiation of these cells provided models to investigate some metabolic characteristics of developing neural cells at defined stages of differentiation.

Tematic Grant of Richter Gedeon Pharmaceutical Inc.,Grant No.: 4700148815

II. Characterization of non-induced and partially differentiated neural stem/progenitor cells. Schedule of in vitro differentiation.

A. NE-4C, embryonic neural stem cells

The NE-4C neural stem cell linewas derived from the forebrain of a 9-day old, $p53^{-/-}$ mouse embryo [2].In non-induced state, the cell divide continuously and display epithel-like morphology. Treatment with 10^{-8}-10^{-6}M retinoic acid (RA) initiates neural differentiation of NE-4C cells resulting the formation of mature neurons approximately by the 7th , and astrocytes after the 14th days of induction.Neuron formation proceeds through well characterizedsteps [2, 5, 6, 7]: starts with aggregationofinduced cells (Day1-3)followed by migration out of the aggregates (Days 4-5) and formation of loose neuronal networks (Days 7-) on top of monolayer of substrate-attached non-neuronal cells. (Fig. 1, 2)

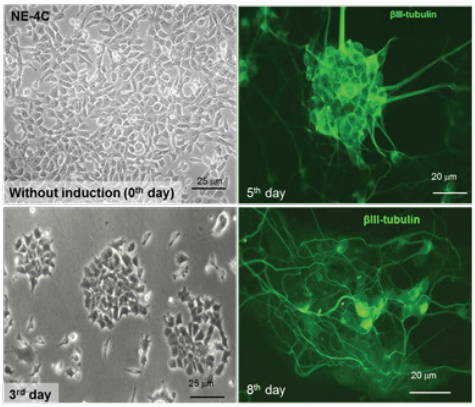


Figure 1. NE-4C cells, derived from the anterior brain vesicles of p53-deficient mouse embryo (E9), proliferate as non-differentiated epithel-like cells in maintaining cultures (RA0) , but give rise to neurons (5th and 8th days) if induced by all-trans retinoic acid (RA).

A. Jády, "Metabolic changes during differentiation of neural stem cells,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 23-26.

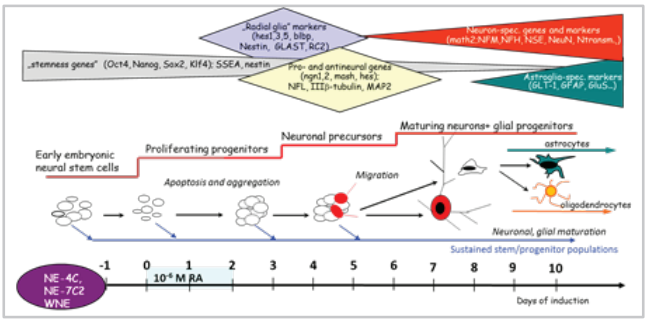


Figure 2. The differentiation proceeds through reproducible stages which had been morphologically, biochemically and physiologically characterized [5, 6, 7].

B. *SVZ_M and HC_A, adult neural stem cells*

The SVZ_M neural progenitor line was derived from the subventriculare zone and the HC_A line was cloned from the hippocampus of adult (P62) mice. The cells were grown on AK-c(RGDfC)-coated surfaces, in serum-free conditions with EGF (epidermal growth factor) supplementation [3]. These cells display characteristics of radial glial (RGl) cells and give rise to neurons upon EGF withdrawal. Neurons develop by the 5th day of induction. [3, 8] (Fig. 3, 4)

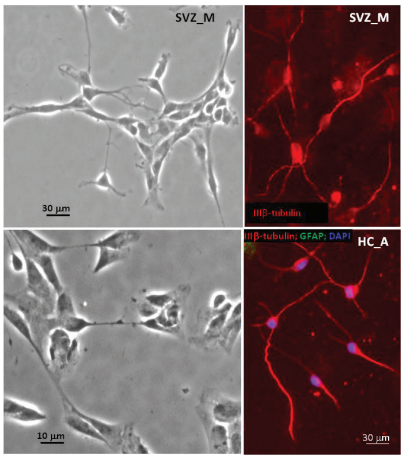


Figure 3. SVZ_M and HC_A cells, derived from the adult mouse subventriculare zone and hippocampus, respectively, proliferate as non-differentiated epithel-like cells in maintaining cultures, but give rise to neurons if induced by EGF withdrawal.

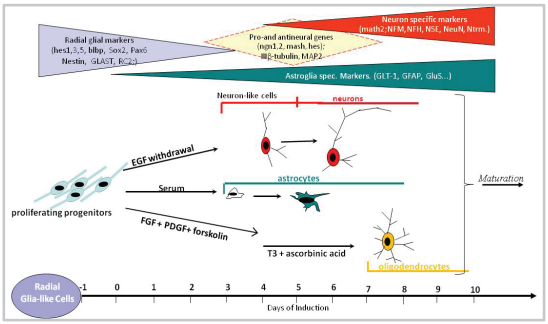


Figure 4. The differentiation proceeds through reproducible stages which had been morphologically, biochemically and physiologically characterized [3, 8].

C. *Primary neuron-enriched cultures*

Embryonic (E15) mouse forebrain cells were seeded on PLL-coated surfaces in media favoring the survival of neurons with the presence of negligible number of astroglial cells. By the 7th day in vitro, the cultures were composed of 75-80% neurons (Fig.5).

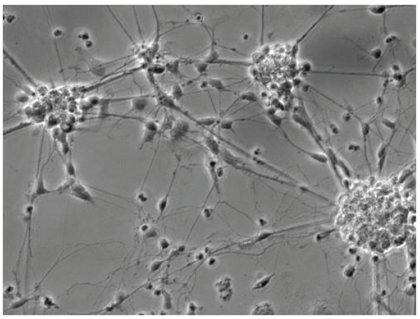


Figure 5. Primary neuron-enriched culture on the 7th day after plating

The primary neuron cultures were used to monitor metabolic features of brain-derived neurons and to compare to those of nerve cells generated by neural stem cells, in vitro.

III. Experimental Setup

A. *Cell cultures*

The cells were maintained in the appropriate media (table 1 and 2) in water-saturated air atmosphere containing 5% CO_2, at 37 °C. The culture media were changed on every 2nd day. NE-4C and RGlcells were serially split using 0,05 (w/v) % tripsin with 1mM EDTA [Invitrogen (Gibco)].

Cells were seeded into 96-well Seahorse plates (1-3 x 10^4 cells/well) coated with appropriate adhesive peptides and were maintained as non-differentiated stem cells or were induced with appropriate treatment (table 1) for neural differentiation.

TABLE 1. Cell culture data

Cells	**Embryonic cell line**	**Adult cell lines**		**Embryonic primary neurons**
	NE-4C	***HC_A***	***SVZ_M***	
Derived from	9 days old mouse embryo forebrain	adult mouse hippocampus	adult mouse subventriculare zone	E 15 mouse forebrain
Medium	5% FCS MEM	high glc DMEM + F12 + B27 + EGF		MEM-ITS
Induction	10^{-6} M retionic acid	EGF-withdrawal		
Substrate	poly-L-lysine	AK-cyclo[RGDfC]		poly-L-lysine

TABLE 2. Tissue culture media

5% FCS MEM	**Serum-free, DMEM-ITS**	**High glc DMEM + F12 + B27 (+ EGF)**
•MEM – minimum essential medium [Sigma], •5% heat-inactivated FCS – foetal calf serum [PAA], •0.4 mM L-glutamin [Sigma] •0.04 mg/ml Gentamicin [Chinoin]	•50% DMEM – Dulbecco's modified Eagle medium [Sigma], •50% F12 HAM [Sigma], •1% ITS – Insulin–Transferin–Selenium [Gibco], •0.4 mM L-glutamin [Sigma,] •0.04 mg/ml Gentamicin [Chinoin]	•50% DMEM – Dulbecco's modified Eagle medium [Sigma], •50% F12 HAM [Sigma], •2% B27 (with retinal) •(40 ng/ml EGF)

The ACSF (artificial cerebrospinal fluid) solution contains 45 mMNaCl [Reanal], 3 mMKCl [Reanal], 2 mM $CaCl_2$ [Sigma], 1 mM $MgCl_2$ [Sigma], 10 mM HEPES [Sigma]. The pH was 7.2. For metabolic assays, the base media were supplemented with one of the following metabolites: 5 mM D-glucose (glc) [Reanal], 5 mM Na-lactate (lac), 5 mM D,L-β-hydroxi-butyrate (βOHB) or5 mM non-essential amino acid mixture (aa) [Gibco].

B. Determinationof O_2 consumption with Seahorse Cell Metabolism Analyzer

The Seahorse XF96 Extracellular Flux Analyzer (Seahorse Bioscience)was employed to determine the impact of various metabolic fuel substrates on the bioenergetic processes of non-induced and differentiating NE-4C, SVZ_M andHC_A cells, and primary neurons. Fluorimetric sensors enabled sensitive in situ measurement of O_2 consumption rate (OCR) and the rate of the extracellular pH drift (ECAR) in a 2.28 µl fluid volume above the cells. (Fig. 6)

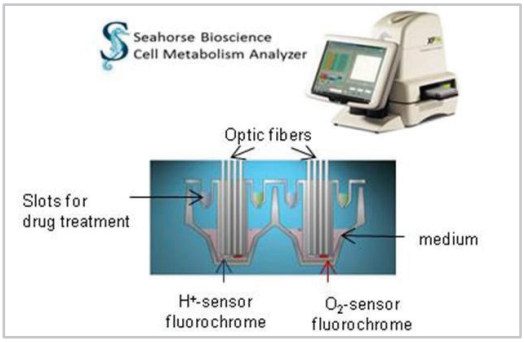


Figure 6. The measuring devices

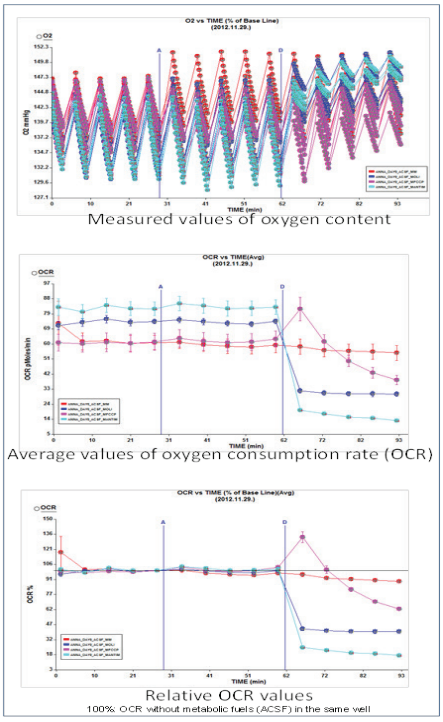


Figure 7. The measured data

In assay-mode, the device reduces the sensing volume to 2.28 µl fluid volume above the cells and produces a gas-tight measuring well. The oxygen content in the cell-covering fluid decreases with time in parallel with the O_2-consumption of cells. The device records the oxygen content in every 15 seconds for 3 minutes, then introduces atmospheric O_2 into the cell covering media by opening up and mixing the wells for 3 minutes. The oxygenation and assay steps alternate. Through ports in the assay-plate, solutions can be introduced to the assay space with a 5-min mixing period. Acidification (pH) of the extracellular medium was measured in parallel with oxygen content in each well. The data are plotted as OCR (oxygen consumption rate: pmole O_2 consumption/min) and ECAR (extracellular acidification rate mpH/min) as a function of time. For comparing reactions of defined wells to added material, OCR and ECAR values were related to those in non-treated control state and plotted as relative OCR and ECAR values. (Fig. 7)

The oxygen consumption rate (OCR) indicates the respiration activity, because the electron transport chain consumes oxygen.The extracellular acidification rate (ECAR)demonstrates mainly the intensity of glycolytic activity, because its lactate production acidifies the environment.

The metabolic state of cells was tested at the end of metabolic assays by monitoring mitochondrial responses to respiration blocking drugs.

Oligomycin blocks ATP synthase (Fig.8) resulting in reduced hydrogen ions consumption and accumulation of hydrogen ions in the intermembrane space of mitochondria. As a consequence, the electron transport chain will be blocked and the oxygen consumption decreased.

FCCP(fluoro3-carbonil cianide-methoxy-phenylhydrazone)opens free routes for hydrogen ionsthrough the inner mitochondrial membrane (Fig. 8)resulting in heavyincrease inthe oxygen consumption.

Cellsresponding accurately to the above drugs possessed functional mitochondria, thus were regarded “healthy”.

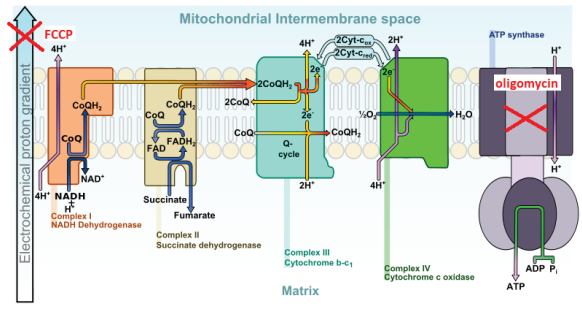


Figure 8. The effects of the used drugs

C. The metabolic treatment

For assaying metabolic characteristics, two different approacheswere used. As a common firststep, all cells were thoroughly washed with metabolite-free ACSF to remove metabolic components of the maintaining media. (Fig. 9)

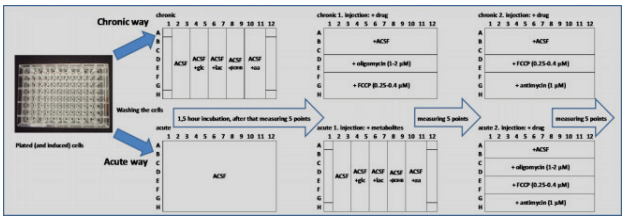


Figure 9. The flow chart of the metabolic treatments

In *chronic assays*, 180 µl ACSFsupplemented with the required metabolite was added to each well. According to the probed metabolites, there were 5 treatment-groups on each plate: without metabolite (starvation), with glucose, lactate, β-hydroxi-butyrate (keton-body) or aminoacid mixture.After 1.5 hour incubation, 5 OCR and ECAR data (60 – 60 datapoints) were recorded. After recording the metabolite effects, mitochondrial drugs were injected one after other, and 5 data points were measured from each treatment.

In *acute assays*, 180 µl ACSFwas added without any metabolite supplementation. (So the cells were starving.) After 1.5 hour incubation, 5 OCR and ECAR data (60 – 60 datapoints) were recorded as a baseline. A supplementary metabolite was then added through the injection port and further 5 data points were measured. At the end of metabolite testing, mitochondrial drugs were injected and 5 data points were measured.

IV. Brief results

During differentiation,the oxygen consumption and metabolite requirements of developing neural cells change significantly. Results obtained on viability, O_2-consumption and extracellular acidification of distinct neural stem cell lines demonstrated that different neural stem/progenitor populations and also their differentiating progenies display specific, cell-type and developmental stage-dependent demandsfor survival.

Acknowledgment

I would like to thank Prof. LászlóTretter,TündeKovács, Susan Van-Wert andKatalinGaál for their theoretical and technical help.

I also want to thank the help and guidance to my supervisior, Prof. EmíliaMadarász.

This work has been supported by the Tematic Grant of Richter Gedeon Pharmaceutical Inc., Grant No.: 4700148815

References

[1] Anita Zádori, Viktor Antal Ágoston, Kornél Demeter, Nóra Hádinger, Linda Várady, Tímea Kőhídi, Anna Göbl, Zoltán Nagy, Emília Madarász, “Survival and differentiation of neuroectodermal cells with stem cell properties at different oxygen levels”; Experimental Neurology, vol. 227, pp. 136–148,2011

[2] Katalin Schlett and Emília Madarász, “Retinoic Acid Induced Neural Differentiation in a Neuroectodermal Cell Line Immortalized by p53 Deficiency”; Journal of Neuroscience Research, vol. 47, pp. 405–415, 1997

[3] Károly Markó, Tímea Kőhídi, Nóra Hádinger,Márta Jelitai, Gábor Mező, Emília Madarász, “Isolation of radial glia-like neural stem cells from fetal and adult mouse forebrain via selective adhesion to a novel adhesive peptide-conjugate”; PLoS One (6) e28538, 2011

[4] Varga B, Markó K, Hádinger N, Jelitai M, Demeter K, Tihanyi K, Vas A, Madarász E.,“Translocator protein (TSPO 18kDa) is expressed by neural stem and neuronal precursor cells”; Neurosci.Letts., vol. 462, pp. 257-262, 2009

[5] Herberth B, Pataki A, Jelitai M, Schlett K, Deák F, Spät A, Madarász E.,“Changes of KCl sensitivity of proliferating neural progenitors during in vitro neurogenesis”; J Neurosci Res., vol. 67, pp. 574-582, 2002

[6] Jelitai M, Schlett K, Varju P, Eisel U, Madarász E., “Regulated appearance of NMDA receptor subunits and channel functions during in vitro neuronal differentiation”**;** J Neurobiol., vol. 51, pp. 54-65, 2002

[7] Jelitai M, Anderová M, Chvátal A, Madarász E., “Electrophysiological characterization of neural stem/progenitor cells during in vitro differentiation: study with an immortalized neuroectodermal cell line”; J Neurosci Res., vol. 85, pp. 1606-1617, 2007

[8] Madarasz E., “Diversity of neural stem/progenitor populations. In: Neural Stem Cells”; InTech Book Series; Ed. L.Bonfantini; in press, 2013

Electrophysiological Correlates of Visual Word Form Familiarity and Crowding

Balazs Knakker
(Supervisor: Zoltan Vidnyanszky)
knakker.balazs@itk.ppke.hu

***Abstract*—Reading is a complex ability that is built upon the plasticity and effectiveness visual object recognition. Visual word recognition is most effective for words in familiar format; reading altered-format words might require less effective alternative strategies. One example for format alteration is increasing or decreasing letter spacing. Letter spacing, apart from familiarity, is related to the degree of crowding, which occurs when the visual system counterproductively groups together similar objects that are close to each other, impairing their processing as separate items. In this study, we sought to explore how these two factors (crowding and visual familiarity) affect early visual processing. To this end, we presented normal, double-spaced and minimally spaced words and measured brain activity with electroencephalography. The analysis of event-related potentials revealed that the right N1 is modulated by crowding as expected based on previous findings using non-orthographic stimuli. In turn, the bilateral P2-N2 (and possibly the left N1) shows effects of orthographic familiarity. These results raise new possibilities for probing in what ways early visual processing of orthographic stimuli might be impaired in dyslexics.**

***Keywords*-visual word perception, reading, letter spacing, crowding, event-related potentials**

I. Introduction

The human ventral occipito-temporal cortex evolved to process complex visual stimuli both effectively and flexibly. Effectiveness is achieved through developing specialized circuitries for stimuli that are frequent and/or behaviorally relevant. The classical example of specialization in the visual system is the Fusiform Face Area [1]: faces are socially highly relevant, very frequently encountered stimuli, whose great variability and subtle information content is analyzed by a neural network involving this area. However, for infrequent, visually unfamiliar stimuli, the effectiveness gained by specialization is traded off, given that the brain has limited capacities – thus, most of the time, the processing of these stimuli is slower and might involve alternative, more general mechanisms. For instance, in the case of faces, seeing inverted faces causes smaller activity in this area than looking at upright ones, corresponding to our difficulty recognizing inverted faces compared to the ease of dealing with upright ones[3].

Like inverting faces, manipulating the format of printed words may invoke slower, alternative recognition and reading strategies. For example, increasing interletter spacing probably shifts perception towards a more serial, letter-by-letter approach, disrupting spatial relations that the fast, parallel system relies on. Interestingly, it has been shown that dyslexics, as opposed to typical readers, can benefit from increased letter spacing [4]. Increasing spacing might help dyslexics by diverting their perceptual processing from the typical reading pathway that is impaired in them. In particular, it has been shown that dyslexics are more sensitive to crowding [5]. Crowding occurs when the visual system – counterproductively – groups together similar objects that are close to each other in space, impairing their processing as separate items [6].

Perea et al. [7] inferred that the neural processes underlying the benefit of increased letter spacing most probably occurs the earliest perceptual stages. The effect of letter spacing has been studied with fMRI [2], but not with electromagnetic methods, which are more sensitive to early effects because of their superior temporal resolution. Therefore, we decided to explore how the letter spacing affects visual processing of orthographic stimuli, as indexed by event-related potentials. To this end, besides normally spaced words, we presented our subjects with words in which spacing was either doubled or decreased to a possible minimum. We expected two meaningful patterns to arise. First, processing stages reflecting specialized visual processing of familiar orthographic formats would make responses to normal stimuli differ from both manipulated formats with the same sign. In contrast, sensitivity to crowding would give rise to a pattern of amplitudes with monotonous differences parallel with changes in inter-letter spacing.

II. Materials and Methods

A. Subjects

12 healthy right-handed young adults participated in this study. All of them were native speakers of Hungarian, had typical reading skills and normal or corrected-to-normal vision. None of them had any history of neurological or psychiatric diseases. All participants gave informed consent before the beginning of the measurements.

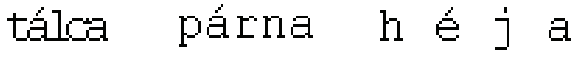


Fig. 1. Stimulus sample

B. Knakker, "Electrophysiological Correlates of Visual Word Form Familiarity and Crowding,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 27-30.

B. Stimuli and Procedure

The stimuli were 4 and 5 letter Hungarian nouns from two semantic categories (living and non-living), presented centrally on a TFT screen using a monospace font (Courier New). Random flanker words were also presented on both sides of the central word, to mimic the visual context during normal reading. There were three presentation conditions: for normal spacing, the built-in spacing parameter of the font was kept. For the two other conditions, spacing was doubled (double spacing), or decreased by a factor of 0.707 (minimal spacing), which results in the letters touching each other but no considerable overlapping (see 1). Normal words subtended approximately 2 degrees of visual field in the vertical dimension. A small blue fixation dot was always present in the center of the screen. The background was middle grey. The subjects were seated in a dark room, their head supported by a chin rest in a distance of 50 centimeters from the screen. The experiments were conducted in 6 runs, each lasting cca. 8 minutes, with some minutes of rest in between. Condition order was counterbalanced across subjects. Within runs, letter spacing of the words was constant. In half of the trials, words were presented without flanking words – these trials are not included in this report.

In each trial, a word was presented for 800 milliseconds. Subjects were told to respond after stimulus offset with a mouse button press, indicating which category the word (living or non-living) they had seen belonged to. The response interval was maximized in 2 seconds. The length of the inter-trial-interval (ITI, starting from the time of the response or from the end of the response interval) was chosen from a uniform probability distribution between 1250 and 1750 milliseconds. After every third trial an additional 650 ms of pause was added with the fixation dot turning red, and the subjects were asked to try to blink only during this period. The frequency and length of these blink windows were sometimes adjusted to individual differences of the given subject's propensity to blink. Stimulus presentation and subject response registration was implemented in MATLAB using PsychToolbox version 3 [8], [9].

C. Electrophysiological and Behavioral Measurements

EEG was acquired using 64 electrodes (Brain Products ActiCap; amplifier: BrainAmp MR) mounted on an elastic cap according to the extended 10/20 system. The sampling rate was 500 Hz and the signal was digitized using an external D/A converter supplied by Brain Products and recorded by the Brain Vision Recorder software. Eye movements were recorded using IView X Hi-Speed (SensoMotoric Instruments) at a sampling rate of 1250 Hz.

D. Data Analysis

Preprocessing of the EEG signal was done in Brain Vision Analyzer. The signal was bandpass filtered (Butterworth zero-phase filter, 0.1Hz-30Hz, 12 dB/octave) and segmented. Segments containing artefacts were marked using amplitude, amplitude difference and voltage step thresholds and by visual inspection; these segments were not used in further analyses.

Data were imported to Matlab, and surface Laplacian approximations of the scalp current density was calculated using the CSD Toolbox[8] (spline flexibility m=4, smoothing $\lambda = 10^{-5}$). Artefact-free segments were baseline-corrected and averages for conditions of interest were computed for each subject. Channels PO9 and PO10 were used in the ERP analysis. On the subject averages, P1 and N1 component peaks were detected semi-automatically. The P2-N2 complex displayed great inter-subject variability and could not be detected over both hemispheres for every subject. However, for subjects in whom peaks (or clearly visible inflections) were present, their timing was fairly consistent. Thus, we used a time-window based approach to assess how letter spacing modulates these two components. The mean amplitude in the [200 ms; 220 ms] and [240 ms; 260 ms] post-stimulus onset time windows in each condition were used to quantify the P2 and N2 components, respectively.

The amplitudes of the P1 and N1 component were entered into separate repeated-measures ANOVAs using within-subject factors spacing (3 levels: minimal, normal and double) and hemisphere (2 levels: PO9 amplitudes for Left and PO10 amplitudes for Right). The Greenhouse-Geisser correction for violation of sphericity was applied where necessary. Post-hoc comparisons were conducted using Tukey's Honestly Significant Differences test. Data from the P2 and N2 time windows was analysed in a common ANOVA, adding a factor of time, to assess whether they are differentially modulated or there was a common effect with a more wide time window.

III. Results

The P1 component was not affected by letter spacing (main effect of spacing: F(2,22)=0.32, p=0.73). For P1, the main effect of hemisphere was significant (F(1,11)=6.70, p=0.025), indicating larger amplitude over the right hemisphere (Fig. 3, first row).

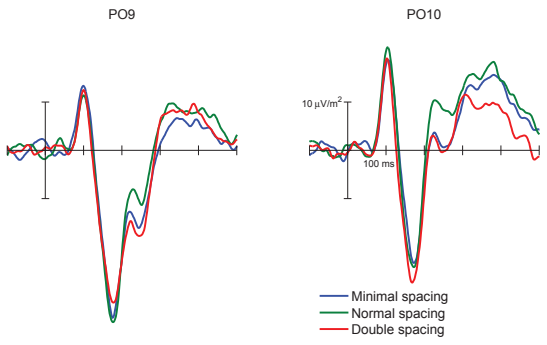


Fig. 2. Grand average event-related potentials from electrodes PO9 and PO10

The N1 component was significantly modulated by letter spacing over the right, but not over the left hemisphere (F(2,22)=6.71, p=0.006, ε=0.96, Fig. 3, second row). In particular, the right-hemisphere N1 became more negative as

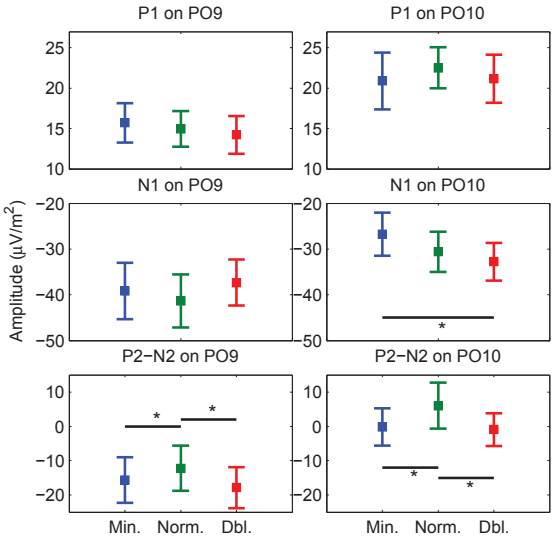


Fig. 3. Component amplitudes on PO9 and PO10. Boxes represent means across subjects, error bars indicate the standard error of the mean. Significant differences are indicated by lines between relevant comparisons and asterisks.

spacing increased (post hoc p=0.012 for minimal vs. double spacing, PO10, all p>0.17 for PO9; see also on Fig. 2). The left N1 displayed a pattern (of non-significant differences) consistent with that its amplitude might be dependent on visual familiarity, i.e. it was smaller in both the double-spacing and the minimal-spacing conditions than for normally spaced print. To put it differently, *decreasing* spacing from the normal level causes bilateral N1 decrement (only significant over the right hemisphere), but *increasing* spacing causes the N1 to diminish over the left, but strengthen over the right hemisphere. To assess this trend, post-hoc interaction tests were conducted, which showed that the dense-normal difference is similar on both sides (dense-normal × hemisphere interaction: F(1,11)=0.61, p=0.45, uncorrected), whereas for the sparse-normal comparison, an interaction is found with hemisphere (sparse-normal × hemisphere interaction: F(1,11)=6.21, p=0.030, uncorrected).

The P2 and N2 components were analyzed using a time-window based approach. However, neither the main effect of time (F(1,11)=0.27, p=0.62), nor any of the interactions (F(2,22)<1.60, p>0.2) were significant, indicating that the two components behaved very similarly (data pooled accross time windows is displayed on Fig. 3, in the third row). Amplitude in the P2-N2 time window was smaller (less negative) in the normal print than both of the altered-format conditions, which were, in turn, identical in amplitude (main effect of spacing: F(2,22)=9.70, p=0.0012, ε=0.95; post hoc comparisons: D<N p=0.010; S<N p=0.0011). Being more extended in time than the N1 effects, the pattern is also visible on the grand average event-related potential (Fig. 2). Again, the overall amplitude was more negative over the left hemisphere (main effect of hemisphere: F(1,11)=17.46, p=0.0015) .

IV. Discussion

Changing letter spacing can affect the perceptual processes of normal reading several ways. First, increasing spacing might decrease the effect of crowding and uncertainty in letter position coding[7]. Second, deviation from normal spacing renders the stimulus unfamiliar and might be unsuited for the specialized neural systems to process efficiently [7]. In typical readers, the costs of larger letter spacing seem to outweigh the benefits: both of our altered spacing stimuli were previously shown to be suboptimal for the typical skilled reader [9]. However, this does not exclude the possibility that both effects of letter spacing might be reflected in early event-related components.

Indeed, we have found that the N1 component over the right hemisphere displays a pattern compatible with a crowding modulation, whereas in the later P2-N2 time window, the normal format differs similarly from both altered-format conditions, yielding a familiarity pattern. The left N1 shows similar, although non-significant pattern of differences.

The right-lateralized crowding effect - the greater degree of crowding made the amplitude of the right N1 smaller - is compatible with recent findings of Chicherov et al [10]. They manipulated the similarity of peripheral flankers to foveal targets, and found that N1 amplitudes decreased in a right-lateralized fashion as crowding increased, and localized the effect most prominently to the right lateral occipital cortex. Their stimuli being non-orthographic, high level object knowledge did not influence their results, which solely reflected the effects of crowding on both the neural and behavioral level.

The P2-N2 difference clearly reflects that normally formatted and altered words might involve different mechanisms in this time window that is also potentially present in the earlier N1 time window over the left hemisphere. Traditionally, the N1 component is primarily thought to reflect the accumulation of category-specific visual information in the ventral occipitotemporal cortex, and the 200-300 ms time window is associated with sublexical orthographic and phonological processing [11]. However, it has more recently been suggested that processes that were assumed to be slow and sequential actually happen fast and almost simultaneously in the first 250 ms, followed by a slower reanalysis phase corresponding to most of the classical effects in the literature [12]. In this account, our P2-N2 modulations would rather belong to a second phase, reflecting the recruitment of additional attentional resources required for alternative processing strategies, possibly involving letter-by-letter processing.

V. Conclusion

Early event-related potentials reflect both the unfamiliarity of altered letter spacing and the effects of crowding. We have shown that crowding modulates the N1 component in a way that is consistent with effects known for non-orthographic stimuli. The left N1 is relatively robust to these changes; however, it shows trends that might be stronger in dyslexics, in whom crowding is known to cause more interference in reading processes. In the P2-N2 time window, the visual

cortex was shown to be sensitive to orthographic familiarity, possibly reflecting additional attentional effort required to process unfamiliar printed word stimuli. These results raise new possibilities for probing in what ways early visual processing of orthographic stimuli might be impaired in dyslexia.

REFERENCES

[1] N. Kanwisher, J. McDermott, and M. M. Chun, "The fusiform face area: A module in human extrastriate cortex specialized for face perception," *The Journal of Neuroscience*, vol. 17, no. 11, pp. 4302 –4311, Jun. 1997. [Online]. Available: http://www.jneurosci.org/content/17/11/4302.abstract

[2] L. Cohen, S. Dehaene, F. Vinckier, A. Jobert, and A. Montavont, "Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways," *NeuroImage*, vol. 40, no. 1, pp. 353–366, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811907010762

[3] G. Yovel and N. Kanwisher, "The neural basis of the behavioral face-inversion effect," *Current Biology*, vol. 15, no. 24, pp. 2256–2262, Dec. 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0960982205014107

[4] M. Zorzi, C. Barbiero, A. Facoetti, I. Lonciari, M. Carrozzi, M. Montico, L. Bravar, F. George, C. Pech-Georgel, and J. C. Ziegler, "Extra-large letter spacing improves reading in dyslexia," *Proceedings of the National Academy of Sciences*, p. 201205566, Jun. 2012, PMID: 22665803. [Online]. Available: http://www.pnas.org/content/early/2012/05/29/1205566109

[5] M. Martelli, G. D. Filippo, D. Spinelli, and P. Zoccolotti, "Crowding, reading, and developmental dyslexia," *Journal of Vision*, vol. 9, no. 4, p. 14, Apr. 2009, PMID: 19757923. [Online]. Available: http://www.journalofvision.org/content/9/4/14

[6] D. Whitney and D. M. Levi, "Visual crowding: a fundamental limit on conscious perception and object recognition," *Trends in cognitive sciences*, vol. 15, no. 4, pp. 160–168, Apr. 2011, PMID: 21420894 PMCID: PMC3070834. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3070834/

[7] M. Perea, C. Moret-Tatay, and P. Gómez, "The effects of interletter spacing in visual-word recognition," *Acta Psychologica*, vol. 137, no. 3, pp. 345–351, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0001691811000758

[8] J. Kayser and C. E. Tenke, "Principal components analysis of laplacian waveforms as a generic method for identifying ERP generator patterns: I. evaluation with auditory oddball tasks," *Clinical Neurophysiology*, vol. 117, no. 2, pp. 348–368, Feb. 2006. [Online]. Available: http://www.clinph-journal.com/article/S1388-2457(05)00369-X/abstract

[9] D. Yu, S.-H. Cheung, G. E. Legge, and S. T. L. Chung, "Effect of letter spacing on visual span and reading speed," *Journal of Vision*, vol. 7, no. 2, p. 2, Jan. 2007, PMID: 18217817. [Online]. Available: http://www.journalofvision.org/content/7/2/2

[10] V. Chicherov, G. Plomp, and M. H. Herzog, "Neural correlates of visual crowding," *NeuroImage*, vol. 93, Part 1, pp. 23–31, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811914001207

[11] J. Grainger and P. J. Holcomb, "Watching the word go by: On the time-course of component processes in visual word recognition," *Language and linguistics compass*, vol. 3, no. 1, pp. 128–156, Jan. 2009, PMID: 19750025 PMCID: PMC2740997. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2740997/

[12] F. Pulvermüller, Y. Shtyrov, and O. Hauk, "Understanding in an instant: Neurophysiological evidence for mechanistic language circuits in the brain," *Brain and Language*, vol. 110, no. 2, pp. 81–94, Aug. 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0093934X08001673

Microvesicle fractionation from human blood using deterministic lateral displacement effect

András J. Laki
Supervisors: Kristóf Iván Ph.D., Pierluigi Civera Ph.D.
laki.andras@itk.ppke.hu
June 2014

***Abstract*—Label-free fractionation of tumor-derived extracellular vesicles are essential in cancer research and in many diagnostic and therapeutic methods. Recently, there has been an interest in methods that avoid the use of biochemical labels, intrinsic biomarkers or electrical polarizability to identify microvesicles from human blood. Here, we report a microfluidic device to separate circulating extracellular vesicles from blood samples using the deterministic lateral displacement principle. The device continuously fractionates label-free extracellular vesicles and cells according to size and membrane flexibility by displacing them perpendicularly to the fluid flow direction in a micro-fabricated array of posts. Experimental data and computational fluid dynamic simulations are presented to create a compelling argument that microvesicles from human blood could be fractionated by deterministic lateral displacement arrays. Direct separation of different size micro- and nanospheres were demonstrated using a multi-stage separation strategy thus offering a potential route for novel cancer diagnostic approaches where microvesicles can be targeted and intercepted during cell to cell communication.**

I. Introduction

The extracellular space of multicellular organisms contains solutions of metabolites, ions, proteins and polysaccharides and a large number of mobile membrane-bounded vesicles, called extracellular vesicles such as exosomes (Exs), microvesicles (MVs) and apoptotic bodies (ABs) [1]. Recent advances in the study of tumor-derived microvesicles reveal new insights into the cellular basis of disease progression and the potential to translate this knowledge into innovative approaches for cancer diagnostics and personalized therapy [2]. A key step in cancer diagnostics and molecular biological observations is to separate cells, functionalized microbeads, extracellular vesicles, or other particles from a solution which may contain other undesirable elements [3]. Even though a number of microfluidic techniques have been developed to enhance on-chip blood fractionation [4], classification of membrane vesicles, protocols of their isolation and detection, molecular details of vesicular release, clearance and biological functions are still under intense investigation. The most frequently used methods to purify microvesicles and exosomes from cell culture supernatants or body fluids involve a series of centrifugation and filtration steps to remove cells, apoptotic bodies and other cellular contaminants by a final high-speed ultracentrifugation to pellet small extracellular vesicles [5]. These procedures require long preparation time, ultracentrifuge equipment and yield a relatively low amount of extracellular vesicles [6], making it difficult for application in clinical practice.

Due to the importance of the isolation of extracellular vesicles, several microfluidic devices have been developed for such purpose but these techniques mostly work in batch separation mode [7]. In our approach, we developed a device for continuous and label-free separation of extracellular vesicles across functional laminar streamlines in pressure-driven microfluidic flow using an asymmetric micropost array. By flowing human blood into and through functional streamlines, different size components can be fractionated for further biomedical diagnostic processes using a deterministic lateral displacement (DLD) array. The DLD technique, a size-based particle sorting method, has shown extremely high size selectivity, adaptability to sorting multiple particle sizes [8], and a broad range of operating conditions, sorting particles from 100 nm to 30 μm [9].

Being a continuous flow method, DLD has all the advantages that such methods have to offer. Furthermore, clinical implementation of DLD devices may open new perspectives in translational medicine both in diagnostics and therapy.

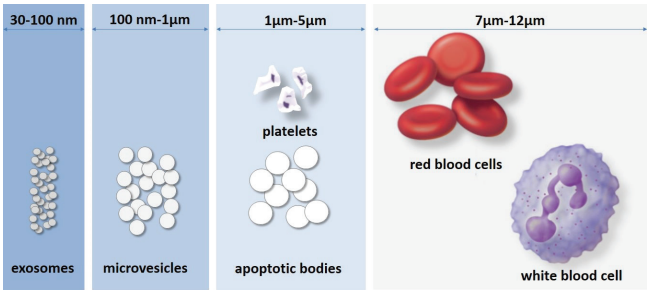


Fig. 1. Size ranges of major blood components. While exosomes share size distribution with viruses, microvesicles overlap in size with bacteria and protein aggregates (e.g. immune complexes), apoptotic bodies and thrombocytes fall into the size range of $1-5$ μm, whereas diameter of erythrocytes is around $6-8$ μm and the size of lymphocytes is from 7 μm up to 12 μm.

II. Device principles

A continuous-flow, label-free separation procedure reported by Huang et al. [8] and known as deterministic lateral displacement is able to separate micrometer-sized particles enhancing on-chip blood fractionation with an uncertainty of 10 nm. This technique, which shows a marked improvement over existing

A. J. Laki, "Microvesicle fractionation from human blood using deterministic lateral displacement effect," in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds. Faculty of Information Technology and Bionics, Pázmány Péter Catholic University. Budapest, Hungary: Pázmány University ePress, 2014, pp. 31-34.

methods, has been shown capable of separating erythrocytes (RBCs), white blood cells (WBCs), thrombocytes, plasma, circulating tumor cells, and nematodes/infections [10] from whole blood based on their size. The anticoagulant-treated blood is forced to flow through an array of circular posts under low Reynolds number conditions, in which each row of posts is slightly offset laterally with respect to the previous row (Fig. 2). Particles below a critical hydrodynamic diameter D_c are able to follow one such stream through the array (zigzagging mode) whereas bigger particles are forced through interactions with posts, to change streams many times, always in the same general direction, becoming laterally displaced (displacement mode) in each section (n). The DLD phenomenon is based on the column shift fraction (ϵ_n):

$$\epsilon_n = \frac{\Delta\lambda_n}{\lambda}, \tag{1}$$

which is the ratio of vertical (tangential to the flow) distance that each subsequent column is shifted ($\Delta\lambda$) to the vertical array period (λ_n). The critical diameter for separation is then described by the following equation [11]:

$$D_{c,n} = 2g\epsilon_n\eta_n \tag{2}$$

where g is the gap distance between two pillars in the same row ($g = D_{post} - \lambda$) and η_n is a unit-less parameter taking into account the parabolic flow profile between the pillars in the array, a consequence of pressure driven flow [11].

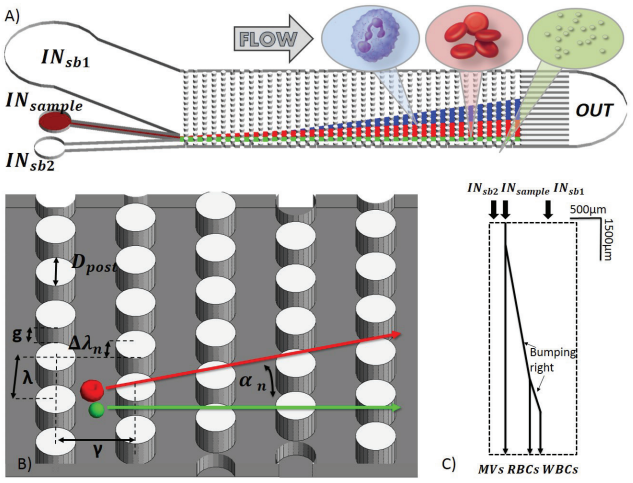


Fig. 2. A) Schematic illustration of the separation by deterministic lateral displacement in an array of microposts. The sample fluid (IN_{sample}) is hydrodynamically focused by shield buffers (IN_{sb1}, IN_{sb2}) and pushed through the array structure. B) Definition of the device geometry. The vertical (tangential to the flow) distance between two sets of columns is γ, the vertical array period is λ, each subsequent column is shifted by $\Delta\lambda_n$ and diameter of pillars is D_{post}. Microvesicles (green particles) remain within the flow stream and cells (red particles) are displaced by α_n angle at each obstacle. C) Diagram of average particle paths in our device.

A serial connection of different critical size DLD structures have also been implemented to retrieve size fractions of mixed particles. Our developed DLD array was designed with pillars of 20 μm diameter (D_{post}), the gap between adjacent pillars in each column (g) is 10 μm, the horizontal (perpendicular to the flow) array period (λ) is 30 μm and the vertical (tangential to the flow) array period (γ) is 40 μm. The column shift ratio (ϵ_n), which ranges from 0.1 up to 0.33 with steps of 1/60, describes 15 column sections (n) following each other thus the $D_{c,n}$ is between 3.9 μm and 7.7 μm in such an array sequence.

The DLD effect, unlike many separation technologies, relies on a deterministic process instead of a stochastic process such as diffusion. There are dimensionless numbers that describe the qualitative behavior of solute particles in a continuous single-phase liquid flow by using the particle Reynolds number Re_p, the Péclet number Pe and Stokes number St, which are expressed by the following equations [12]:

$$Re_p = Re\frac{d^2}{D_H^2} = \frac{v_{max}d^2}{\nu D_H} \tag{3}$$

Where Re is the channel Reynolds number, d is the particle diameter, v_{max} is the maximum flow velocity, ν is the kinematic viscosity (in our case blood at $3,33 \cdot 10^{-06}$ Pas) and D_H is the hydraulic diameter of a rectangular channel:

$$D_H = \frac{2wh}{w+h}, \tag{4}$$

where w is the width ($w = g_n = 10$ μm) and h is the height (20 μm) of channel thus the hydraulic diameter between neighbour columns is $1.33 \cdot 10^{-5}$ μm, where the Reynolds number has the highest value ($D_H = 13.3$ μm, $Re_p \approx 4{\cdot}10^{-9}$) and correlates with the flow velocity (Fig. **??**). If Re_p<1, the inertial force becomes a dominant parameter for driving the lateral migration of particles transverse to fluid streamlines, whereas if Re_p>1 the particle behavior is strongly promoted to follow the flow pattern by a viscous drag force acting on the particle surface. In the consideration of mass transport, the particle motion can be estimated by the Péclet number:

$$Pe = \frac{vD_H}{D}, \tag{5}$$

where v is the mean velocity of the fluid (in our case, $v \approx 0.001$ m/s) and D is the diffusion coefficient of particles. The diffusion coefficient of the particles is described by the Stokes-Einstein equation:

$$D = \frac{k_B T}{6\pi\mu r}, \tag{6}$$

where k_B is Boltzmann's constant ($1.38 \cdot 10^{-23}$ J/K), T is the absolute temperature (in our case, 298 K), and r is the radius of the spherical particle (in our case from 400 nm up to 10 μm). If Pe>1 the advection rate is bigger than the diffusion rate between rows of posts thus the small particles ($D_{particle}$>$D_{c,n}$) are basically confined to streamlines. When a particle encounters an accelerating flow in a nonlinear channel, the Stokes number estimates the particle behavior as the ratio of the relaxation time (τ_p) of the particle to the characteristic time (τ_f) of the flow:

$$St = \frac{\tau_p}{\tau_f} = \frac{\rho_p d^2/18\mu}{D_H/v_{max}} = \frac{\rho_p}{18\rho_f}Re_p, \tag{7}$$

where ρ_p and ρ_f are the particle density and fluid density. If $St>1$, a particle will continue in its original moving direction instead of following the fluid streamline when the flow turns suddenly by the channel geometry. Due to the low particle Reynolds number ($Re_p<1$), the high Peclét number ($Pe>1$) and low Stokes number ($St<1$), the observed blood particles are remain on the desired streamlines.

III. Results and discussion

Purified blood components, such as erythrocytes (Fig. 3.A), lymphocytes (Fig. 3.B) and microvesicles (Fig. 3.C), are loaded into the center inlet (IN_{Sample}) following each other at the same boundary conditions, whereas the sheathed buffers are introduced at the from ports on the left (IN_{sb1}) and right (IN_{sb2}) sides of the specimen port focusing samples to the desired width (Fig. 2.A). The concentration of erythrocytes was about $5 \cdot 10^6$ per μL, PBMCs were about $7 \cdot 10^3$ per μL and microvesicles were about $8 \cdot 10^4$ per μL.

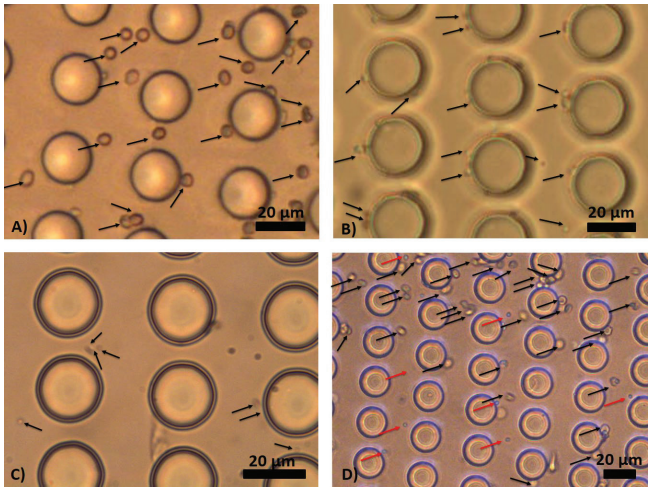


Fig. 3. Images from observed main blood components which were forced through the deterministic lateral displacement array A) erythrocytes B) peripheral blood mononuclear cells C) microvesicles D) platelet-free blood sample (prepared by 2500 g centifugation), where black arrows indicate erythrocytes whereas red arrows point to microvesicles.

To optically detect the blood elements, the biomedical samples are driven through the device at $0.001\ ml/h$ flow rate, which provides a suitable rate of cells for counting and a suitable residence time in front of the camera to be imaged (Fig. 3). We record the lateral position of particles at three different position along the device ($n = 1$, $n = 7$ and $n = 15$) in the deterministic lateral displacement array and bin these particles for the histograms which are shown on Fig. 4.

The microvesicles, which are below any critical hydrodynamic diameter $D_{c,n}$ are able to follow a given stream through the array in zigzagging mode whereas erythrocytes and lymphocytes become laterally displaced by every interaction with posts. The further displacement of PBMCs comes after section $n = 7$, when the diameter of RBCs becomes equal with the actual critical diameter ($D_{c,n}, n>7$) and RBCs enter in zigzagging mode meanwhile PBMCs are forced to adopt orientations that give them a greater displacement along the device.

Although the behavior of blood particles can be considered as deformable and non-spherical particles, which in a DLD array is significantly more complex than that of hard spheres. Shear forces, which result from gradients in the fluid velocity around a particle, may induce complex motions including rotation, tumbling and shape change [13]. This suggests that such blood cells appear to modify shape and diameters as they pass through the DLD device which can occur lower separation efficiency. The behavior of blood components in the developed DLD array results in smooth histograms (Fig. 4).

The overview of our DLD structure is shown Fig. 4.A, where results were bin for the histograms at the initial section (Fig. 4.I, $n = 1$), in the middle of the DLD array (Fig. 4.II, $n = 7$) and the end of the DLD array (Fig. 4.III, $n = 15$). The dispersion of microvesicles, which is shown of Fig. **??**.C, remains mainly the initial along the device due to the dimensionless numbers of fluid dynamics ($Re_p<1$, $Pe>1$ and $St<1$). The average laterally displacement of erythrocytes, which is represented on Fig. **??**.B, is around $100 - 120\ \mu m$ between the initial and the terminal section. Although the lymphocytes are displaced by $140 - 160\ \mu m$ from the initial position, which shown on Fig. **??**.A. The obtained and reported efficiency of fractionation can be increased by a longer device and the throughput by parallelized microfluidic devices.

Secondly, anticoagulant-treated human blood was forced through the DLD structure thus the lymphocytes does not or less coagulate to the surface of the device. Due to the overlapping size ranges of thrombocytes and vesicles, the thrombocytes were extracted from blood sample and concentration of microvesicles was increased by additional harvested sample. When the anticoagulant-treated human blood enters the device (typically at about $0.001\ ml/h$), as we have observed previously we get similar results thus the erythrocytes and lymphocytes are laterally displaced by the interactions with posts, whereas the extracellular vesicles follow the initial central stream through the array structure in zigzagging mode. The displacement of RBCs, PBMCs and microvesicles are observed at the terminal section by CCD camera-based image recording system and them distributions are corresponded with the sum of previous parallel cases.

IV. Conclusions and outlook

In this paper we have shown evidence that label-free fractionation of microvesicle from human blood can be delivered by using a deterministic lateral displacement array. This suggests that our DLD device may be able to provide rapid diagnostic information about the haemostatic condition of a blood sample, to explore cell to cell communication or to fractionate blood sample efficiently for clinical tests without the use of an activation specific label or marker. Our developed device may be useful in a wide range of medical procedures and diagnostics which involve extracellular vesicles. Several DLD devices could be joined in series to allow separation

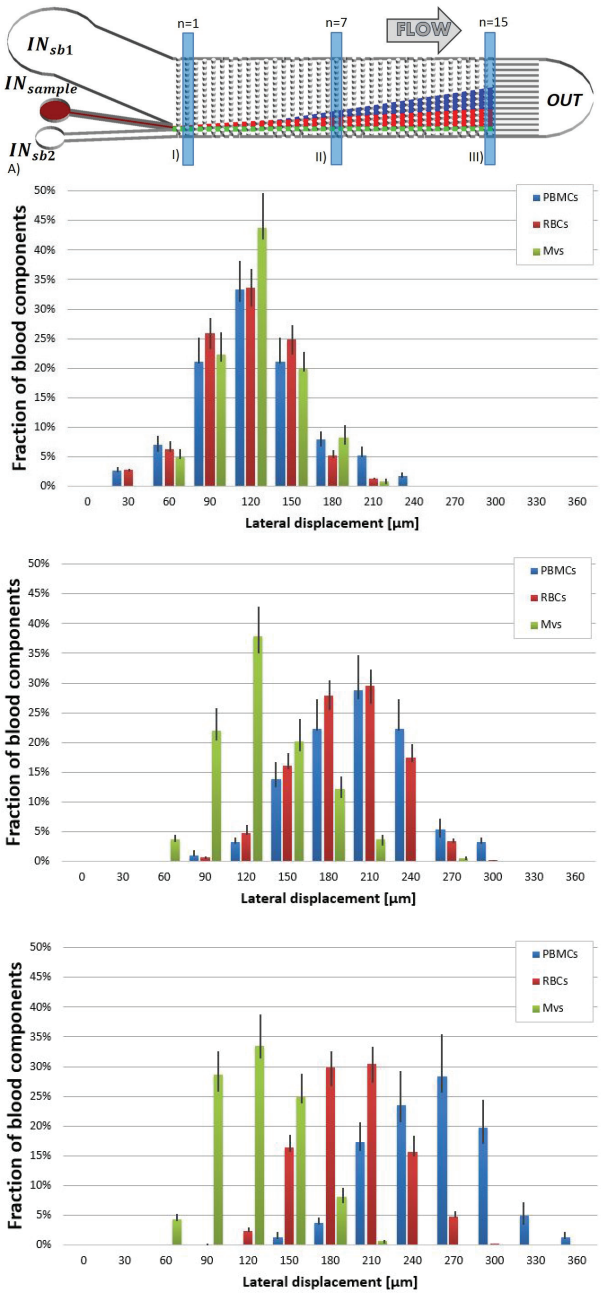


Fig. 4. A) Overview of our deterministic lateral displacement device. The greater the diameter of the particles the further they are laterally displaced. I) the distribution of blood components at the initial section ($n = 1$), II) in the middle of the DLD array ($n = 7$) and III) the end of the DLD array ($n = 15$). Vertical error bars are based on the average of 10 parallel measurements.

of several different droplet sizes or in parallel to increase throughput.

V. Acknowledgment

We would like to extend our gratitude to Péter Fürjes (MEMS Laboratory, Research Center for Natural Sciences, Hungarian Academy of Sciences, Budapest) for his kind help. The support of grants TÁMOP-4.2.1.B-11/2/KMR-2011-0002 and TÁMOP-4.2.2/B-10/1-2010-0014, OTKA 84043 and FP7-PEOPLE-2011-ITN-PITN-GA-2011-289033 DYNANO are gratefully acknowledged.

References

[1] B. György, T. G. Szabó, M. Pásztói, Z. Pál, P. Misják, B. Aradi, V. László, E. Pállinger, E. Pap, A. Kittel, G. Nagy, A. Falus, and E. I. Buzás, "Membrane vesicles, current state-of-the-art: emerging role of extracellular vesicles," *Cellular and molecular life sciences: CMLS*, vol. 68, no. 16, pp. 2667–2688, Aug. 2011, PMID: 21560073.

[2] C. D'Souza-Schorey and J. W. Clancy, "Tumor-derived microvesicles: shedding light on novel microenvironment modulators and prospective cancer biomarkers," *Genes & Development*, vol. 26, no. 12, pp. 1287–1299, Jun. 2012, PMID: 22713869. [Online]. Available: http://genesdev.cshlp.org/content/26/12/1287

[3] N. Pamme, "Continuous flow separations in microfluidic devices," *Lab on a Chip*, vol. 7, no. 12, pp. 1644–1659, Nov. 2007. [Online]. Available: http://pubs.rsc.org/en/content/articlelanding/2007/lc/b712784g

[4] H. W. Hou, A. A. S. Bhagat, W. C. Lee, S. Huang, J. Han, and C. T. Lim, "Microfluidic devices for blood fractionation," *Micromachines*, vol. 2, no. 3, pp. 319–343, Jul. 2011. [Online]. Available: http://www.mdpi.com/2072-666X/2/3/319

[5] C. Théry, S. Amigorena, G. Raposo, and A. Clayton, "Isolation and characterization of exosomes from cell culture supernatants and biological fluids," in *Current Protocols in Cell Biology*, J. S. Bonifacino, M. Dasso, J. B. Harford, J. Lippincott-Schwartz, and K. M. Yamada, Eds. Hoboken, NJ, USA: John Wiley & Sons, Inc., Apr. 2006. [Online]. Available: http://www.currentprotocols.com/WileyCDA/CPUnit/refId-cb0322.html

[6] H. G. Lamparski, A. Metha-Damani, J.-Y. Yao, S. Patel, D.-H. Hsu, C. Ruegg, and J.-B. Le Pecq, "Production and characterization of clinical grade exosomes derived from dendritic cells," *Journal of Immunological Methods*, vol. 270, no. 2, pp. 211–226, Dec. 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022175902003307

[7] C. Chen, J. Skog, C.-H. Hsu, R. T. Lessard, L. Balaj, T. Wurdinger, B. S. Carter, X. O. Breakefield, M. Toner, and D. Irimia, "Microfluidic isolation and transcriptome analysis of serum microvesicles," *Lab on a Chip*, vol. 10, no. 4, pp. 505–511, Feb. 2010. [Online]. Available: http://pubs.rsc.org/en/content/articlelanding/2010/lc/b916199f

[8] L. R. Huang, E. C. Cox, R. H. Austin, and J. C. Sturm, "Continuous particle separation through deterministic lateral displacement," *Science*, vol. 304, no. 5673, pp. 987–990, May 2004, PMID: 15143275. [Online]. Available: http://www.sciencemag.org/content/304/5673/987

[9] J. A. Davis, D. W. Inglis, K. J. Morton, D. A. Lawrence, L. R. Huang, S. Y. Chou, J. C. Sturm, and R. H. Austin, "Deterministic hydrodynamics: taking blood apart," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 40, pp. 14 779–14 784, Oct. 2006, PMID: 17001005.

[10] S. H. Holm, J. P. Beech, M. P. Barrett, and J. O. Tegenfeldt, "Separation of parasites from human blood using deterministic lateral displacement," *Lab on a chip*, vol. 11, no. 7, pp. 1326–1332, Apr. 2011, PMID: 21331436.

[11] D. W. Inglis, J. A. Davis, R. H. Austin, and J. C. Sturm, "Critical particle size for fractionation by deterministic lateral displacement," *Lab on a Chip*, vol. 6, no. 5, pp. 655–658, May 2006. [Online]. Available: http://pubs.rsc.org/en/content/articlelanding/2006/lc/b515371a

[12] C. Bonadonna, G. G. J. Ernst, and R. S. J. Sparks, "Thickness variations and volume estimates of tephra fall deposits: the importance of particle reynolds number," *Journal of Volcanology and Geothermal Research*, vol. 81, no. 34, pp. 173–187, May 1998. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0377027398000079

[13] J. P. Beech, S. H. Holm, K. Adolfsson, and J. O. Tegenfeldt, "Sorting cells by size, shape and deformability," *Lab on a chip*, vol. 12, no. 6, pp. 1048–1051, Mar. 2012, PMID: 22327631.

Ranking Drug Combinations via Logistic Regression

Balázs Ligeti
(Supervisor: Dr. Sándor Pongor)
ligeti.balazs@itk.ppke.hu

***Abstract*—Drug combinations offer a better treatment option for complex diseases - such as cancer - compared to monotherapy. However, designing such combinations is a challenging task due to the underlying complexity of the diseases. In this paper we will present a novel method that can efficiently prioritize cancer drug combinations based on the the assumption that those drugs can form efficient combinations that are linked to a large number of common perturbed proteins and share some therapeutic properties. It considers not only the network phenomena such as crosstalks, feedback and feed forward loops (identified via perturbation analysis of the constituent drugs), but also therapeutic and functional similarities between the components identified by analyzing the network of gene ontology data and therapeutic information. We compared our predictions with the outcome of recently finished clinical trials (carried out on trastuzumab, a well known and widely used cancer drug). The aggregated scores of the combinations containing trastuzumab and different cytotoxic drugs show significant correlation with the outcome of clinical trials, both with the objective response and the progression free survival.**

***Index Terms*—drug combination;drug interaction;**

I. Introduction

In the past few decades the number of novel marketed drugs have fallen much below the expectations despite the growing resources invested in this area [1], [2], [3]. Biological pathways, and in particular those involved in cancer like the EGFR pipeline are tough against drug effects since they have rich regulatory loops which can be utilized to compensate various perturbations [3], [4]. Multitarget drugs or drug combinations have been proposed as a general strategy to circumvent this phenomenon [5], [6], [7], [8]. In addition drug combinations have lower toxicity and therapeutic selectivity [9]. Nevertheless, the number of approved drug combinations is increasing, even though most of them were established by experience and intuition [10], [11]. Several experimental methods, even high throughput methods, have been developed for measuring the efficiency of drug combinations, such as Bliss independence or Loewe additivity [12], [13], but this kind of exhaustive search is impractical. Wong et al. used a stochastic search algorithm [14] while Calzoari et al. used a sequential decoding algorithms for finding the best combinations [15]. Yang et al. use differential equations to find a perturbation pattern that can revert the system from a disease state to a normal state [16]. Jin et al. employed a Petri net based model to microarray data in order to predict the synergism of drug pairs [17]. The common in these computational methods is that they require a large number of experiments or deep knowledge of the kinetic parameters of the pathways even if the search space is small. Others use data mining methods to integrate pharmacological and network data [18], [19], [20]. Li et al. used the concept of network centrality and disease similarity to prioritize drug combinations [21]. Wu et al. used the microarray profile of the individual drugs for the predictions [22], while others use the concept of synthetic lethality and the available gene interaction data [23], [24]. In this paper we present a novel drug combination prediction algorithm which is partly based on the assumption that the perturbations generated by the drugs propagate through the possible interactions between proteins. I also assume that the components of the combination have to share some therapeutic and functional properties that can be measured by using the Anatomical Therapeutic Chemical (ATC) Classification System and the Gene Ontology [28] annotations.

II. Methods

Figure 1 shows the general workflow of the method. Three different drug-drug interaction strength measures were used. One is based on the analysis of the perturbation made by the drugs individually (DIFF), the other two measure the functional (GO) and therapeutic similarities (ATC) between the components. Each measure can be seen as a feature that describes one aspect of the drug - drug interaction. Then a logistic regression modell was trained using different features of known and random combinations. Finally the trained model ranked the candidate combinations.

A. Network based drug-drug interaction measure (DIFF)

The effect of the perturbation generated by a drug is modelled by a diffusion process initiated from the drug target proteins. We define the subnetwork affected by a drug as the set of proteins significantly perturbed by the drug. These proteins are the nodes of the subnetwork. We assume that those drug combinations are strong that share many drug-affected proteins, in other words, their subnetworks substantially overlap. This overlap can be measured by the Jaccard measure (similarity measure between sets), where the elements of the sets are the nodes of the significantly perturbed subnetworks. The significance levels were computed with Monte Carlo simulations.

The network is a graph $G(V, E)$ where V, E are the set of nodes and edges, respectively. In this case the nodes represent genes or proteins, and the edges are the associations between

B. Ligeti, "Ranking Drug Combinations via Logistic Regression,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 35-38.

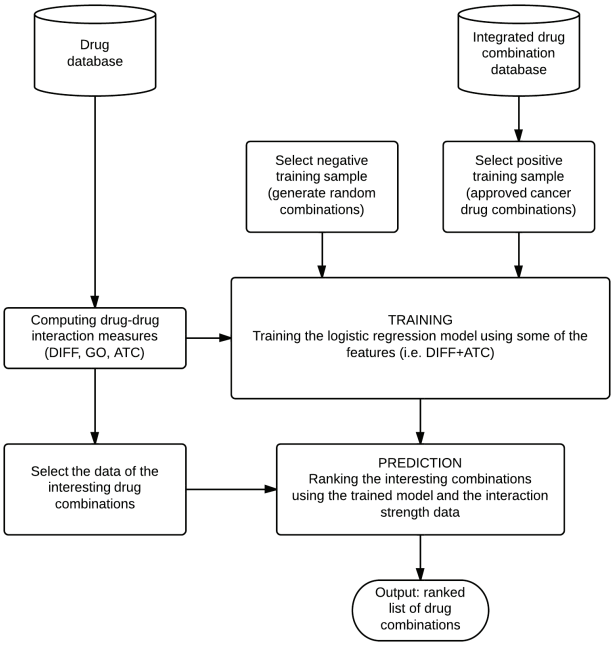


Fig. 1: *The input is a list of candidate combinations (i.e. combinations selected for clinical trials) and the set of known combinations (i.e. previously approved cancer combinations). The first step is to compute the drug combination measure (DIFF) and the drug interaction measures (GO, ATC) for all possible drug combinations. The database consists of the random generated drugs and of the components of the candidate and the known combinations. After the selection of the training sample (both the positive - known cancer combinations - and the negative one - random combinations) a logistic regression was trained using the previously computed DIFF and similarity values. In the next step the trained model is used for ranking a set of candidate combinations. The output is the ranked list of the drug combinations.*

them. The edges may have a weight, which can be interpreted as an association strength. Let A be the adjacency matrix of the graph. The element a_{ij} is the weight of the edge between node i and j, if there is no edge then it is 0. The drug affected proteins can be determined by using the Regularized Laplacian Exponential Diffusion Kernel ($K_{\mu,\alpha}$) [25] (DIFF). The formula of that kernel is:

$$K_{\mu,\alpha} = \sum_{k=1}^{\infty} \frac{\alpha^k}{k!}(-L_\mu)^k = e^{-\alpha L_\mu}$$

where L_μ is the regularized Laplacian of the graph:

$$L_\mu = \mu G - A$$

Where G is a diagonal matrix, where $g_i = \sum_{j=1}^{|V|} A_{ij}$. The ith drug (D_i) perturbation can be expressed with vector:

$$S_{DIFF}(D_i) = K_{\mu,\alpha} p_0 \tag{1}$$

where p_0

$$p_0 = \begin{cases} 1, & \text{if the protein } i \text{ is drug target} \\ 0, & \text{otherwise} \end{cases}$$

The jth element of $S(D_i)$ measures the disruption effect of D_i on protein j. In this application the diffusion paramater α was 0.005 and the regularization parameter μ was chosen as 0.1.

B. Randomizations and the drug affected proteins (DAP)

The statistical significance of each protein was estimated by a randomization procedure (10000 times) [26]. After that one can define the set of drug affected proteins (DAPs) as follows:

$$DAP = \{v_j | v_j \in V, p_j < 0.05\}$$

Where p_j is the p-value of protein j. I assumed that the sets of DAPs of the interacting drugs largely overlap, which is measured by the Jaccard coefficient, thus the DIFF drug-drug interaction strength is:

$$S_{DIFF}(D_i, D_j) = \frac{|DAP_i \cap DAP_j|}{|DAP_i \cup DAP_j|}, \tag{2}$$

C. Gene ontology

For each drug a GO vector (g_i) was built, where each entry of the vector represents the presence or the absence of a GO term annotated to the drug targets. The ith entry is 1 if the ith term is annotated to the target protein, 0 otherwise. Then the cosine similarities between drugs can be computed.

$$S_{GO}(D_i, D_j) = 1 - \frac{g_i^T g_j}{\|g_i\| \|g_j\|} \tag{3}$$

D. ATC

The Anatomical Therapeutic Chemical Classification (ATC) System classifies drugs into groups at five levels in a hierarchical way. Thus the classification system can be seen as a simple ontology, more specifically a forest (disjoint union of trees). The roots of the individual trees are the first level characters/classes of the ATC system and leaves are the full ATC codes (7 character). One could quantify the similarity between any two ATC codes with the Resnik measure [27], that is one of the most commonly used semantic similarity measure. The similarity is based on the common ancestors's information content (IC), that quantify the specificity and the informativity of an ATC code level. The IC of an ATC code level (c) is the negative log likelihood of the probability of the code ($p(c)$) being used. This can be estimated from the annotation frequency.

$$IC(c) = -\log p(c) \tag{4}$$

The Resnik similarity is the most informative common ancestor (MICA) of two ATC codes:

$$sim(ATC_i, ATC_j) = IC(MICA(ATC_i, ATC_j)) \tag{5}$$

Then one can compute the similarities between drugs by taking the maximum value of all possible pairwise Resnik similarities

between each set of ATC codes annotated to the two drugs. The ATC similarity is 0 if the drugs do not share any ATC codes at any level, 1.0 if they have the exactly same ATC code annotation.

E. Logistic regression model

The logistic regression is able to predict how successful an unknown drug combination will be. Briefly, for a series $m_1, m_2, \ldots, m_n$ measures to be combined, the logical regression model will calculate a combined measure S as:

$$S(D_i, D_j) = \frac{1}{1 + e^{\beta_0 + \sum_i \beta_i m_i}} \tag{6}$$

Where the regression parameters β_i were estimated by the glmfit MATLAB function.

F. Score of drug regimens

All the drug-drug interaction measures are only applicable in pairwise cases, where the combinations have only two components. In the multicomponent cases one could simply aggregate the score of all possible pairwise interactions in the interaction. Let the combination DC have n components $DC = \{D_1, D_2, \ldots, D_n\}$ then the score of DC is:

$$S(DC) = \frac{\sum_{(D_k, D_n) \in DC \times DC} S(D_k, D_n)}{n} \tag{7}$$

III. Description of the experiments

All the algorithms were implemented in MATLAB 2013b. The used network was STRING [29]. The drug related data (drug targets, synonyms, aliases, ATC codes) were downloaded from the Drugbank [30], Stitch [31] and TTD databases [32]. The drug combination data were taken from the DCDB [33], TTD database [32].

A. Clinical trial data

This dataset contains information about the recently conducted clinical studies that include trastuzumab. The dataset was created by the group of Balázs Győrffy. The measured response variables were the objective response (OR) and the progression free survival (PFS). The combination is identified by its components, thus the different combinations studied in the same trial were treated as independent entities, but no distinction has been made if the combinations had different dose or time schedule (these combinations have different OR, PFS values).

1) Performance measures: The performance of the numerical indices and index combinations was characterized by their ability to rank positive and negative examples. The ranking performance was calculated by ROC analysis [34], using a leave-one out cross-validation process. The trivial combinations and the combinations being under investigation in clinical trial were excluded from the training set. The drugs lack of ATC code were also excluded as in the case of interacting drugs. In the filtered training set the number of known cancer combinations is quite low (9), what may result in an unstable learning. To overcome that limitation the number of negative combinations (consisting of random generated drugs) were five fold more than that of the positive ones. This procedure was repeated 1000 times to obtain the final average AUC. The predictive performance of the various indices and index combinations was estimated from a correlation with clinical trial data as well.

IV. Results

TABLE I: *The performance of the algorithm on different sets of drug combinations.*

Method[1]	AUC[2]	OR[3]		median PFS[4]	
		r	p-val.	r	p-val.
DIFF+ATC	0.8603 ± 0.0017	0.71	0.0003	0.49	0.1341
DIFF	0.8386 ± 0.0011	0.71	0.0003	0.31	0.2520
DIFF+GO+ATC	0.8183 ± 0.0018	0.71	0.0003	0.88	0.0042
DIFF+GO	0.8141 ± 0.0014	0.74	0.0002	0.67	0.0510
ATC	0.8125 ± 0.0062	0.45	0.0256	0.80	0.0150
GO+ATC	0.7683 ± 0.0027	0.56	0.0063	0.56	0.0962
GO	0.6345 ± 0.0016	0.59	0.0036	0.56	0.0962

[1] The method indicates which drug-drug interaction measures were used for predicting the scores.
[2] AUC was calculated from a ranked list of known cancer and random drug combination by leave-one-out cross-validation as described in the text. The procedure was repeated 1000 times to obtain the final average AUC.
[3] Overall Response.
[4] Median Progression Free Survival.
[5] Spearman's rank correlation coefficient.
[6] P-values for Spearman's rank correlation coefficient.

Table I shows the average test AUC-s. The best prediction performance with leave one-out crossvalidation settings resulted in AUC of 0.86 ± 0.02. The predictor was the diffusion based overlap measure (DIFF) combined with the ATC code information. Predictor combinations containing DIFF are also outstanding. GO based prediction is weak but not completely inadequate, because both GO and GO+ATC show significant correlation with clinical outcome measures. **Table I** presents all the correlation values and the corresponding significance between the aggregated score (**equation 7**) and the outcome of the clinical studies. The most coherent predictor, DIFF+ATC shows strong correlation with the Overall Response (OR) ($r = 0.71$, $p = 0.0003$). In the case of median Progression Free Survival (PFS) the strongest correlation is presented by the DIFF+GO+ATC with $r = 0.88$, $p = 0.0042$. Predictors using GO information had lower performance on the known dataset compared to the diffusion based method. However, on the clinical trial data set they also showed strong correlation with the clinical outcomes, so the GO based predictors might be useful in several cases.

V. Conclusions

In this paper we presented a novel method that can successfully prioritize candidate drug combinations based on the hypothesis that those drugs can make efficient combinations that a) share therapeutic and functional properties and b) share a large number of perturbed proteins that can be simply measured by the Jaccard coefficient. I also showed that the

integration of different kind of drug-drug interaction measures improved the performance compared to the individual classifiers. However, the method has some limitations; for example, the exact nature of the interaction (e.g. synergetic or antagonistic) can not be predicted due to the scantiness of information about protein interactions. On the other hand, the predicted ranking of the candidate combinations showed correlation with the outcome of clinical studies.

ACKNOWLEDGMENT

BL is PhD student at the Multidisciplinary Doctoral School at the Faculty of Information Technology and Bionics of the Pázmány Péter Catholic University. The work was partially supported by the grants TÁMOP-4.2.1./B-11/2/KMR-2011-0002, TÁMOP-4.2.2./B-10/1-2010-0014 and TÉT 10-1-2011-0058. BGy was supported by the OTKA PD 83154 grant.

REFERENCES

[1] A. L. Hopkins, "Network pharmacology: the next paradigm in drug discovery," Nature Chemical Biology, vol. 4, 682–690, 2008.

[2] P. Imming, C. Sinning, and A. Meyer, "Drugs, their targets and nature and number of drug targets," Nature Reviews Drug Discovery, vol. 5, p. 821–834, 2006.

[3] K. Hiroaki, "A robustness-based approach to systems-oriented drug design," Nature Reviews Drug Discovery, vol. 5, p. 202–210, 2007.

[4] K. Hiroaki, "Biological Robustness," Nature Reviews Genetics, vol. 5, p. 826–837, 2007.

[5] S. I. Berger, and R. lyengar, "Network analyses in systems pharmacology," Bioinformatics, vol. 25, no. 19, 2466–2472, 2008.

[6] V. Ágoston, P. Csermely, and S. Pongor, "Multiple weak hits confuse complex systems: A transcriptional regulatory network as an example," Physical Review E, vol. 71, no. 5, http://link.aps.org/doi/10.1103/PhysRevE.71.051909, 2005.

[7] P. Csermely, V. Ágoston, S. Pongor, "The efficiency of multi-target drugs: the network approach might help drug design," Trends in Pharmacological Sciences, vol. 26, p. 178–182, 2005.

[8] T. Korcsmáros, M. S Szalay, Cs. Böde, I. A Kovács, P. Csermely, "How to design multi-target drugs: Target search options in cellular networks," Expert Opinion on Drug Discovery, vol. 2, p. 1–10, 2007.

[9] J. Lehár, A. S. Krueger, W. Avery, A. M. Heilbut, L. M. Johansen, E. R. Price, R. J. Rickles, G. F. Short III, J. E. Staunton, X. Jin, M. S. Lee, G. R. Zimmermann, and A. A. Borisy, "Synergestic drug combinations tend to improve therapeutically relevant selectivity," Nature Biotechnology, vol. 27, p. 659–666, 2009.

[10] C. T. Keith, A. A. Borisy, and B. R. Stockwell, "Multicomponent therapeutics for networked systems," Nature Reviews Drug Discovery, vol. 4, p. 71–78, 2005.

[11] G. R. Zimmermann, J. Lehár, and C. T. Keith, "Multi-target therapeutics: when the whole is greater than the sum of the parts," Drug Discovery Today, vol. 12, no. 1-2, p. 34–42, 2006.

[12] W. R. Greco, G. Bravo, and J. C. Parsons, "The search for synergy: a critical review from a response surface perspective," Pharmacological Reviews, vol. 47, no. 2, p. 331–385, 1995.

[13] A. A. Borisy, P. J. Elliott, N. W. Hurst, M. S. Lee, J. Lehar, E. R. Price, G. Serbedzija, G. R. Zimmermann, M. A. Foley, B. R. Stockwell, and C. T. Keith, "Systematic discovery of multicomponent therapeutics," Proceedings of the National Academy of Sciences, vol. 100, no. 13, p. 7977–82, 2003.

[14] P. K. Wong, F. Yu, A. Shahangian, G. Cheng, R. Sun, and C. M. Ho, "Closed loop control of cellular functions using combinatory drugs guided by a stochastic search algorithm," Proceedings of the National Academy of Sciences of the United States of America, vol. 105, no. 13, p. 5105–10, 2008.

[15] D. Calzolari, S. Bruschi, L. Coquin, J. Schofield, J. D. Feala, J. C. Reed, A. D. McCulloch, and G. Paternostro, "Search algorithm as a framework for the optimization of drug combinations," PLoS Computational Biology, vol. 4, no. 12, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2590660/, 2008.

[16] K. Yang, H. Bai, Q. Ouyang, L. Lai, and C. Tang, "Finding multiple target optimal intervention in disease-related molecular network," Molecular Systems Biology, vol. 4:228, Epub 2008 Nov 4.

[17] G. Jin, H. Zhao, X. Zhou, and S. T. Wong, "An enhanced Petri-net model to predict synergistic effects of pairwise drug combinations from gene microarray data," Bioinformatics, vol. 27, no. 13, p. 310–316, 2011.

[18] X. M. Zhao, M. Iskar, G. Zeller, M. Kuhn, V. Noort, and P. Bork, "Prediction of drug combinations by integrating molecular and Pharmacological data," PLoS Computational Biology, vol. 7, no. 12:e1002323, Epub 2011 Dec 29.

[19] H. Fu-Yan, S. Jiangning, and Z. Xing-Ming, "Exploring Drug Combinations in a Drug-Cocktail Network, " IEEE International Conference on Systems Biology (ISB), Zhuhai, China, September 2 - 4, 2011, p. 382–387.

[20] X. Ke-Jia, S. Jiangning and Z. Xing-Ming, "A network biology approach to understand combination of drugs," The Fourth International Conference on Computational Systems Biology (ISB2010); 2010, p. 347–354.

[21] S. Li, B. Zhang, and N. Zhang, "Network target for screening synergistic drug combinations with application to traditional Chinese medicine," BMC Systems Biology vol. 5, Suppl 1:S10, 2011.

[22] Z. Wu, X. M. Zhao, and L. Chen, "A system biology approach to identify effective cocktail drugs," BMC Systems Biology, vol. 4, Suppl 2:S7, 2010.

[23] M. Cokol, H. N. Chua, M. Tasan, B. Mutlu, Z. B. Weinstein, Y. Suzuki, M. E. Nergiz, M. Costanzo, A. Baryshnikova, G. Giaever, C. Nislow, C. L. Myers, B. J. Andrews, C. Boone, and F. P. Roth, "Systematic exploration of synergistic drug pairs," Molecular Systems Biology, vol. 7:544. doi: 10.1038/msb.2011.71, 2011.

[24] J. Xiong, J. Liu, S. Rayner, Z. Tian, Y. Li, and S. Chen, "Pre-Clinical Drug Prioritization via Prognosis-guided genetic interaction networks," PLoS One, vol. 5, no. 11:e13937, 2010.

[25] T. Ito, M. Shimbo, T. Kudo, Y. Matsumoto. "Application of kernels to link analysis.," International Conference on Knowledge Discovery and Data Mining archive Proceedings of the eleventh ACM SIGKDD Pages: 586 - 592.–, 2005.

[26] P. N. Westfall, and S. S. Young, *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*, Wiley, New York, 1993.

[27] P. Resnik, "Using information content to evaluate semantic similarity in a taxonomy," Proceedings of the 14th international joint conference on Artificial intelligence, vol. 1. p. 448–453, 1995

[28] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, and G. Sherlock, "Gene Ontology: tool for the unification of biology, " Nature Genetics, vol. 25, no. 1, p. 25–29, 2000.

[29] D. Szklarczyk, A. Franceschini, M. Kuhn, M. Simonovic, A. Roth, P. Minguez, T. Doerks, M. Stark, J. Muller, P. Bork, L. J. Jensen, and C. von Mering, "The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored," Nucleic Acids Research, vol. 39 (Database issue), p. 561–8, 2010.

[30] C. Knox, V. Law, T. Jewison, P. Liu, S. Ly, A. Frolkis, A. Pon, K. Banco, C. Mak, V. Neveu, Y. Djoumbou, R Eisner, AC. Guo, DS. Wishart., "DrugBank 3.0: a comprehensive resource for 'omics' research on drugs.," Nucleic Acids Research, vol. 39 (Database issue), p. 1035–8, 2011.

[31] M. Kuhn, D. Szklarczyk, A. Franceschini, C. von Mering, LJ. Jensen, P. Bork., "STITCH 3: zooming in on protein-chemical interactions. ," Nucleic Acids Research, vol. 40 (Database issue), p. 876–8, 2012.

[32] F. Zhu, Z. Shi, C. Qin, L. Tao, X. Liu, F. Xu, L. Zhang, Y. Song, X. Liu, J. Zhang, B. Han, P. Zhang, Y. Chen., "Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery," Nucleic Acids Research, vol. 40 (Database issue), p. D1128–8, 2012.

[33] Y. Liu, B. Hu, C. Fu, and X. Chen, "DCDB: Drug combination database " . Bioinformatics, vol. 26, no. 4, p. 587–588, 2010.

[34] P. Sonego, A. Kocsor, and S. Pongor, "ROC analysis: applications to the classification of biological sequences and 3D structures," Briefings in Bioinformatics, vol. 9, no. 3, p. 198–209, 2008.

In silico study on MAO enzymes: From structure to inhibitor specificity.

Zselyke Magyari
(Supervisor: Peter Matyus)
magyari.zselyke@pharma.semmelweis-univ.hu

***Abstract*—Members of the amine oxidase enzyme family, such as Semicarbazide-Sensitive Amine Oxidase (SSAO), Monoamine-Oxidase-A and -B (MAO-A, -B) are structurally diverse, different cofactor containing catabolizing enzymes, located in various types of tissues. A lot of MAO-A and MAOB inhibitors are available drugs in the market, as MAO-A isoform plays an important role in mental illnesses e.g. schizophrenia, depression, and MAO-B in several neurodegenerative diseases, e.g. Parkinson's and Alzheimer's diseases, all of them with selectivity and specificity problems causing serious off target effects. In the last twenty years SSAO enzymes have been related to inflammatory diseases and cancer types, but still no approved drugs against them. Investigations on these enzymes' substrates show a significant similarity. In the first semester our goal was to develop a complete, inhibitors (MAO−A, MAO−B, SSAO inhibitors) containing accurate database, molecules separated by their selectivity profile, with comparable IC_{50} values in it, and to explore the structural facts of ligand for further *in silico* molecular modeling.**

***Keywords*-monoamine oxidase; ChEMBL database; activity; assay type;**

I. Introduction

Semicarbazide-Sensitive Amine Oxidases exist in two isoforms, a membrane bounded form (different name: VAP-1: Vascular Adhesion Protein-1) which is a type II transmembrane protein, and a soluble form in the plasma. VAP-1 with SSAO activity is an adhesion molecule mediating leukocyte trafficking. SSAO is mainly represented in endothelial cells, smooth muscle cells and adipocytes, although it has been found in the lung, gut, gall bladder, kidney, cartilage, liver, retina, placenta, pancreas and plasma. In the brain it is absent from the glial cells and nerves, but present in the microvessels and may therefore contribute to the blood-brain barrier. The enzyme catalyzes the oxidative deamination of primary aliphatic and aromatic amines, such as tyramine, dopamine and 5-HT. Significantly higher SSAO activity was observed in various human disorders, including diabetes, congestive heart failure, liver cirrhosis, Alzheimer's disease and several inflammatory diseases, although the underlying causes are often unknown.
The enzymes's substrate specificity overlap partly with the MAOs, it has been studied whether SSAO may partially compensate for MAO oxidative deamination in cases when the MAO is dysfunctional [1][2].

Monoamine oxidase enzymes are localized in the mitochondrial outer membrane, and are expressed in most tissues. These enzymes metabolize a wide range of primary, secondary, and tertiary monoamines. Monoamine oxidases exist in two isoforms, MAO-A and MAO-B, encoded by two separate genes localized to the X chromosome, although they share a 70 % amino-acid sequence identity. MAO-A shows grates activity towards 5-hydroxy-tryptamine (5-HT) and noradrenaline (NA), mostly expressed in peripheral tissues e.g. liver, heart but also in some part of the brain e.g. perikarya of the locus coeruleus. MAO-B expression level is significantly higher in the brain e.g. perikarya of the serotonergic neurons, glial cells, ependyma and this isoform shows higher activity towards benzylamine (BZ) and 2-phenylethylamine (PE).
The substrate and inhibitor selectivity comes from the differences between MAO-A and MAO-B active sites. MAO-A binding pocket is a single chamber close to the membrane surface, but MAO-B combining site consists of two, an entrance cavity and an inner combining site. If a suitable substrate binds to MAO-B's entrance cavity an isoleucine (ileu199) rotates, due to this deformation the two chamber fuse, and the inner, substrate binding site becomes available (Fig.1.) [3][4].

Parkinson's disease, discovered by James Parkinson in 1817 is the second most frequent age related neurodegenerative disease affecting 1% of the population over the age 50. In Parkinsons disease decreasing level of dopamine caused by the accelerating degradation of nigrostriatal dopaminergic neurons effect the most relevant symptoms, so unfortunately clinical therapy mostly means dopamine replacement. Inhibition of MAO-B eventuates the elimination of the remaining amount of dopamine, and contributes to lower the level of reactive oxygen species in functional cells resulting in prolonged progression of the disease. Developing new neuroprotective MAO-B inhibitors could result in longer survival time and/or better quality of life to those people suffering from Parkinsons disease [5]. *In silico* molecular modelling methods e.g. structure-activity relationship and pharmacophore models are intended to be built to find molecules with MAO-B selective inhibitory activity applying database screening or to construct a new selective MAO-B inhibitor with occasionally additive activity towards other amine oxidase enzymes e.g. SSAO.

The overall reaction of MAO and SSAO enzymes [1][3]:

$$R - CH_2NHO_2 + O_2 + H_2O = R - CHO + NH_3 + H_2O_2$$

Zs. Magyari, "In silico study on MAO enzymes: From structure to inhibitor specificity,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 39-41.

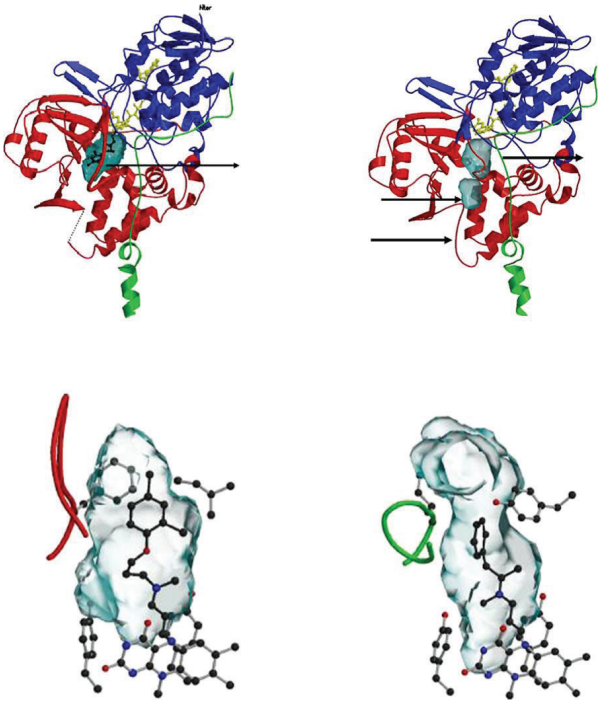


Fig. 1: MAO-A (left) and MAO-B (right) enzymes and binding pockets. FAD (Flavin Adenine Dinucleotide) cofactor is observable in both binding pockets, and in MAO-A the covalently binding clorgyline inhibitor, in MAO-B the covalent inhibitor, selegiline [3].

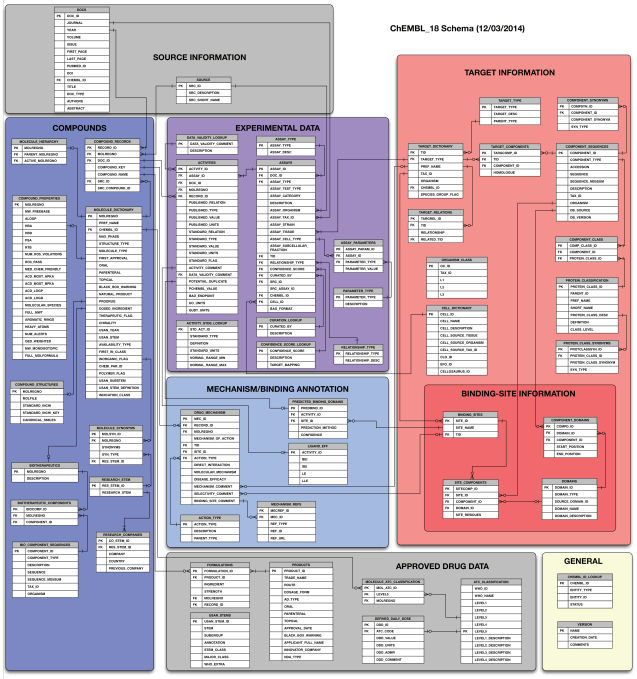


Fig. 2: Overall structure of ChEMBL database, data (different types) collected into separate tables, in all cases a unique primary key provides tables′ connection.

This study shows the workflow of how we constructed an accurate MAO and SSAO inhibitors containing database to be the base of our further modelling study, and also some details of the quantity and quality of the database.

II. Materials and Methods

First we downloaded the ChEMBL whole open source database. ChEMBL is a manually curated chemical database of bioactive molecules with drug-like properties [6]. It is maintained by the European Bioinformatics Institute (EBI), based on the Wellcome Trust Genome Campus, Hinxton, UK. The database, originally known as StARlite, was developed by a biotechnology company, Inpharmatica Ltd. latterly acquired by Galapagos NV. The data was acquired for EMBL in 2008 with an award from The Wellcome Trust, resulting in the creation of the ChEMBL chemogenomics group at EMBL-EBI, led by John Overington [6][7][8]. After applying a proper python script this database was automatically loaded to MySQL. After running a MySQL query we got inhibitor sets with all needed descriptors e.g. activity value, details of the papers they were publicated in, etc (Fig.2.).

Inhibitor selection criteria:

- activity values measured against human recombinant enzymes.
- existing IC_{50} value as activity value.
- publicated in well-known papers.

After our relevant inhibitors containing database was ready for analysis, the IC_{50} measurement assay type identification started. Since a description (a short description of the work detailed in the paper) was available as a property of all molecules we used a python script to find specific words which was expected to identify the assay types. For most molecules no assay circumstances were found in the description, so the structures were finally grouped by the assay type after reading the materials and methods parts for all papers. Two major types (marked with 1 and 2, data reloaded to MySQL with this additional property) were found and some questionable measurements, those inhibitors with no exact assay type were omitted.

III. Results

According to ChEMBL 1603 activity measurements were done with MAO-A 1065 with MAO-B and 105 with SSAO (Fig.3.). For 704 molecules both MAO-A and MAO-B, for 23 both MAO-B and SSAO and for 10 both MAO-A and SSAO activity values were registered. 9 molecules were tested against all three enzymes (Fig.4.). For the above mentioned 9 molecules measured activity values are shown in figure 4. IC_{50} values showing MAO-A inhibitory activity are high, considering none of the 9 molecules active on MAO-A, but it is obvious that 2 structures (CHEMBL2029540, CHEMBL489079) share high inhibitory activity on both MAO-B and SSAO enzymes.

We are planning to discover the minimal common substructure to learn unique structure components leading multiple enzyme (in this case MAO-B and SSAO) binding property, based on our databes which contains comparable activity values.

After all measurement circumstances were registered for MAO-A and MAO-B two well described assay types were identified (some molecules were needed to be omitted because the IC_{50} measurement was not well-defined):

1) Assay type 1 (612 molecules) - Amplex Red MAO Assay
   - enzyme: human recombinant MAO-A, -B (infected insect cells with baclovirus)
   - substrate: p-tyramine
   - measured: H_2O_2 cc. $->$ plus a specific reagent $->$ fluorescent signal

   This kit can be ordered from an American firm.

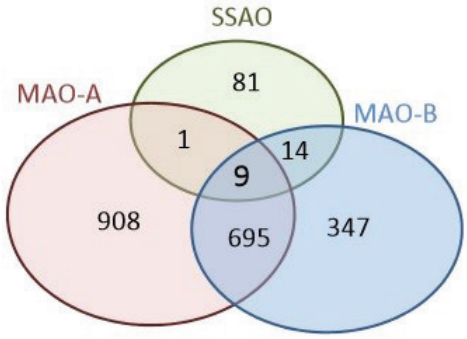
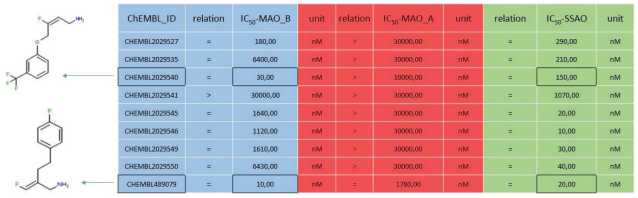


Fig. 3: Number of IC_{50} value measurements on MAO-A, MAO-B and SSAO enzymes according to ChEMBL database.



ChEMBL_ID	relation	IC_{50}-MAO_B	unit	relation	IC_{50}-MAO_A	unit	relation	IC_{50}-SSAO	unit
CHEMBL2029527	=	180,00	nM	>	30000,00	nM	=	290,00	nM
CHEMBL2029535	=	6400,00	nM	>	30000,00	nM	=	210,00	nM
CHEMBL2029540	=	30,00	nM	>	10000,00	nM	=	150,00	nM
CHEMBL2029541	>	30000,00	nM	>	30000,00	nM	=	1070,00	nM
CHEMBL2029545	=	1640,00	nM	>	30000,00	nM	=	20,00	nM
CHEMBL2029546	=	1120,00	nM	>	30000,00	nM	=	10,00	nM
CHEMBL2029549	=	1610,00	nM	>	30000,00	nM	=	30,00	nM
CHEMBL2029550	=	6430,00	nM	>	30000,00	nM	=	40,00	nM
CHEMBL489079	=	10,00	nM	=	1780,00	nM	=	20,00	nM

Fig. 4: 9 molecules wich have measured IC_{50} values on all 3 enzymes. On the left side of the picture two molecules are shown which inhibit both MAO-B and SSAO enzymes with high activity values.

2) Assay type 2 (542 molecules) - Kynuramine assay
   - enzyme: human recombinant MAO-A, -B (infected cells)
   - substrate: kynuramine
     kynuramine + MAO = fluorescent 4-hidroxyquinoline $->$ fluorescent signal
     (This assay has been publicated by Novaroli et al. Bioorg.Med.Chem., 2005, 13, 6212-6217.)

After clustering molecules not just by their enzyme (plus isoform) selectivity but their activity assay type it was clear that 37 distinct molecules were measure with both assay types (1 and 2). In some cases one molecule´s activity value with one assay was measured several times, and as the diagrams show nor the average values nor the minimum activity values differ significantly for all molecules (Fig.5.).

IV. Conclusion

The ChEMBL open source databaes were used to create an accurate MAO-A, MAO-B and SSAO inhibitors containing database. Joining tables in MySQL and applying the necessary queryes we got the molecules clustered by their measured enzyme inhibitory activity. Considering that these activities could be measured with different types of IC_{50} measurement assays leading to non-comparable activity values we clustered the molecules also by their assay types. According to the results we showed that ChEMBL-structures containing large clusters are apropriate to be the base for our further pharmacophore and structure-activity relationship modelling, and also that the measured activity values whith the above mentioned

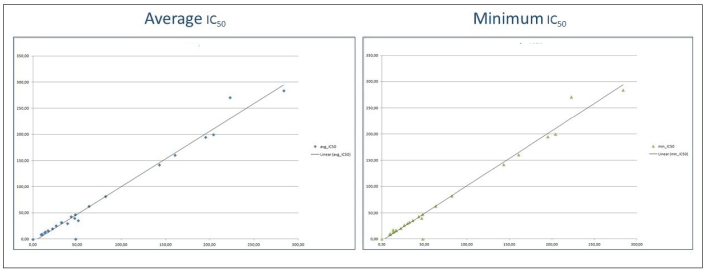


Fig. 5: Measured IC_{50} values with two distinct assays, x axis: applying assay 1, y axis: applying assay 2. For some molecules there were more tham one measurement with the same assay, so in one case (left diagram) we calculated the average measured activity value for one molecule, and the other case (right diagram) we found the minimum values to compare.

two main assay types are completely comparable.

References

[1] @Articlepmid23263543, Author="Tabi, T. and Szok?, E. and Merey, A. and Toth, V. and Matyus, P. and Gyires, K. ", Title="Study on SSAO enzyme activity and anti-inflammatory effect of SSAO inhibitors in animal model of inflammation", Journal="J Neural Transm", Year="2013", Volume="120", Number="6", Pages="963–967", Month="Jun"

[2] @Articlepmid18691041, Author="Dunkel, P. and Gelain, A. and Barlocco, D. and Haider, N. and Gyires, K. and Sperlagh, B. and Magyar, K. and Maccioni, E. and Fadda, A. and Matyus, P. ", Title="Semicarbazide-sensitive amine oxidase/vascular adhesion protein 1: recent developments concerning substrates and inhibitors of a promising therapeutic target", Journal="Curr. Med. Chem.", Year="2008", Volume="15", Number="18", Pages="1827–1839"

[3] @Articlepmid24607445, Author="Finberg, J. P. ", Title="Update on the pharmacology of selective inhibitors of MAO-A and MAO-B: Focus on modulation of CNS monoamine neurotransmitter release", Journal="Pharmacol. Ther.", Year="2014", Pages=" ", Month="Mar"

[4] @Articlepmid17573034, Author="Edmondson, D. E. and Binda, C. and Mattevi, A. ", Title="Structural insights into the mechanism of amine oxidation by monoamine oxidases A and B", Journal="Arch. Biochem. Biophys.", Year="2007", Volume="464", Number="2", Pages="269–276", Month="Aug"

[5] @Articlepmid24109179, Author="Ma, L. and Wei, L. and Wu, F. and Hu, Z. and Liu, Z. and Yuan, W. ", Title="Advances with microRNAs in Parkinson's disease research", Journal="Drug Des Devel Ther", Year="2013", Volume="7", Pages="1103–1113"

[6] @Articlepmid21948594, Author="Gaulton, A. and Bellis, L. J. and Bento, A. P. and Chambers, J. and Davies, M. and Hersey, A. and Light, Y. and McGlinchey, S. and Michalovich, D. and Al-Lazikani, B. and Overington, J. P. ", Title="ChEMBL: a large-scale bioactivity database for drug discovery", Journal="Nucleic Acids Res.", Year="2012", Volume="40", Number="Database issue", Pages="D1100–1107", Month="Jan"

[7] @Articlepmid21948594, Author="Andreas Bender.", Title="Databases: Compound bioactivities go public", Journal="Nature Chemical Biology", Year="2010", Volume="6", Pages="309",

[8] @Articlepmid19194660, Author="Overington, J. ", Title="ChEMBL. An interview with John Overington, team leader, chemogenomics at the European Bioinformatics Institute Outstation of the European Molecular Biology Laboratory (EMBL-EBI). Interview by Wendy A. Warr", Journal="J. Comput. Aided Mol. Des.", Year="2009", Volume="23", Number="4", Pages="195–198", Month="Apr"

Movement related cortical potentials and oscillations on corticography

Tamás Molnár
(Supervisor: LorándErőss MD, PhD)
tamas.molnar@mail.oiti.hu

Abstract: **Understandingthe mechanisms behind sensorimotor integration is essential todesign brain computer interfaces (BCI). Movement related cortical potentials (MRCP) have been studied for long, but their intracranial sources have not been evaluated in details.Eight patients with epilepsy implanted with subdural macro-electrodes arrays (MAE)over the central region were included in this study. Finger tapping(FT) and n. medianus somatosensory evoked potentials (SSEP)were recorded.Movement related cortical potentials were averaged to the button press event, and cortical surface voltage maps were visualized on the MAE array. Time–frequency analysis of the FT event was used to assess frequency characteristics in a low (8-30 Hz) and high (>60 Hz) frequencyband. Low amplitude motor potential (MP)dipole was identified from-10ms to 50ms with peak negativity postcentrally, positivity precentrally. The surface dipole configuration resembled the N20 component of intracranial SSEP. Time-frequency analysis revealed a decrease in spectral power in the lowfrequency band and an increase in the high frequencyband measured on the somatosensory cortex. The low amplitude MP negativity seems to be localized to the postcentral region despite previous data in the literature. Further aims are to investigate the sensory-motor integration, and try to compare it to preoperative high-density EEG results and additional time-frequency analysis.**

Keywords: **brain-computer interface, event related potential, corticography**

Abbreviations: **brain-computer interface (BCI), event related potential (ERP), movement related cortical potential (MRCP), motor potential (MP), subdural macro electrode (MAE), finger tapping (FT),somatosensory evoked potential (SSEP), event-related desynchronization (ERD), event-related synchronization (ERS), electrocorticography (ECoG), electroencephalography (EEG)**

I. INTRODUCTION

The motor system may be damaged selectively at the level of the primary motor cortex,the spinal cord, the perifericnerves or the musclesresulting the otherwise intact brain to be unable to control body functions and movements.Such a damage causes one's quality of life severely decreased, especially in terms of communication and mobility[1].Amyotrophic lateral sclerosis (ALS, Lou Gehrig's disease) is a well-knownexample of such a selective motor damage. Brain-computer interfaces (BCI) can help in these cases to replace the missing link between the brain and the environment. The principle idea of aBCI (*Figure1*) is thatintact brain areas are capable of producing signals which can be translated as inputs into a computer that can, on the other hand, drive a communication [2] or an external prosthetic device [3] to restore the patient's communication or movements abilities. Typicallyslow cortical potentials, sensorimotor rhythms, the P300 evoked potential, and cortical neuronal (unit) activities can be used as such BCI input signals[4].Two main types of BCI devices can be distinguished: 1) invasive and 2) non-invasive. Invasive BCI requires intracranially implanted electrodes, while the non-invasive ones use EEG signals recorded from the scalp[5]. There are input signals such as unit activities that can be detected only invasively. Not only the signal but also the cost, the easiness of the application, or the stability of action differs between these approaches.

The most common and best-confirmed input signals for BCI are scalp recorded event related potentials. Intracranial MRCPs have not been assessed in details so far, however these signals have higher functional resolution, and potentially better signal to noise ratio. A deeper understanding of these potentials is essential to utilize them for BCI application.

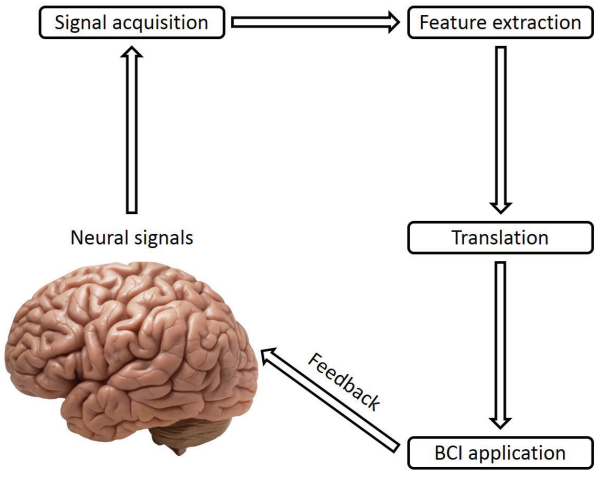


Figure 1. Simplified diagram of a BCI system

T. Molnár, "Movement related cortical potentials and oscillations on corticography,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 43-46.

II. METHODS

A. *Patient management, electrode implantation, functional mapping*

Eight patients (P1-P8) with therapy resistant focal epilepsy were implanted with subdural strip or grid electrode arraysover the sensory-motor region as part of their evaluation for epilepsy surgery in order to determine the exact location of epileptic focus. 5-7 days continuous video-EEG observation was conducted with the implanted electrodes to record epileptic seizures. During these days the dosage of the patients' antiepileptic drugs were typically lowered in order to induce more seizures. After this period the patients underwent a second surgery where the electrodes were removed and the determined epileptic focus was removed.

During video-EEG monitoring various tests were performed in order to localize the functional areas on the cortex. Functional cortical mapping includedhigh frequency electrical stimulation (bipolar 2-10 mA, 50 Hz, 0.5ms, 3-4 s trains) of the cortical areas which can induceneuropsychologic or neurologic alterations such as speech arrest, naming difficulties, movement changesaccording to the function of the stimulated area.

B. *Finger tapping and SSEP*

Intracranial MRCPs were evoked by self-paced (around 1 Hz) and/or auditory cue (on cue in every 4.5-7.5 s) paced finger tapping of the contralateral hand relative to the implantation side. SSEP were evoked stimulating the contralateral
n. medianustranscutaneously (8-12mA, 2Hz, 0.2ms, 200 stim).

C. *Analysis*

SSEP, and FT button press averages were calculated. MRCPswere detected and surface voltage maps at selected time lagswere visualized on the MAE array. Time-frequency analysis was performed using the event related spectral perturbation method. Changes in a low (8-30 Hz, alpha and beta) and in a high (>60 Hz) frequency band were assessed. The analyses were performed using Matlab (version 2012a, Mathworks, Nattick, MA, USA) based nswiew (version 4, Fabo, NICNS, Hungary) and eeglab (version 13, Delorme and Makeig, SCCN, Univerity of California, USA) toolboxes and Neuroscan (version 4.5, Compumedics, Charlotte, NC, USA).

III. RESULTS

A. *EEG averaging*

The averaged EEG of FT and SSEP can be seen on *Figure 2*. MRCP consist a low amplitude dipole that lasted from -10 ms to 50 ms (*Figure 4*). The peak negativity can be detected in the postcentral region while the peak positivity in the precentral region. The surface dipole configuration resembled the N20 component (short latency cortical response) of intracranial SSEP.

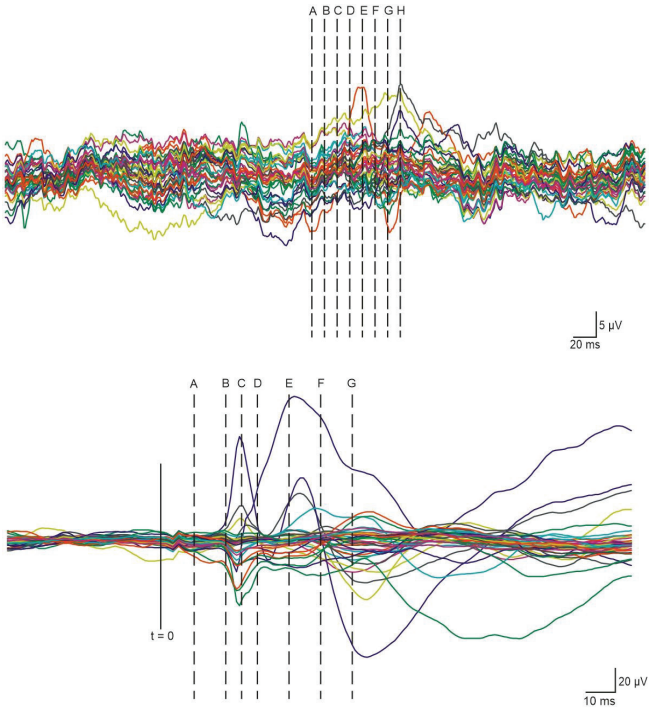


Figure 2. Top: Averaged electroencephalogram (EEG) around FT event (t=0). Bottom: Averaged EEG around SSEP. Different colors of curves represent different electrodes.Snapshotsof electrode potentials were taken at t=-10 ms, 0 ms, 10 ms, 20 ms, 30 ms, 40 ms, 50 ms, 60 ms (FT, timepoint A to H, *Figure 4*) and at t=10 ms, 20 ms, 25 ms, 30ms, 40 ms, 50 ms, 60 ms(SSEP, timepoint A to G, *Figure 4*) and were plotted on a grid array(*Figure 4*).

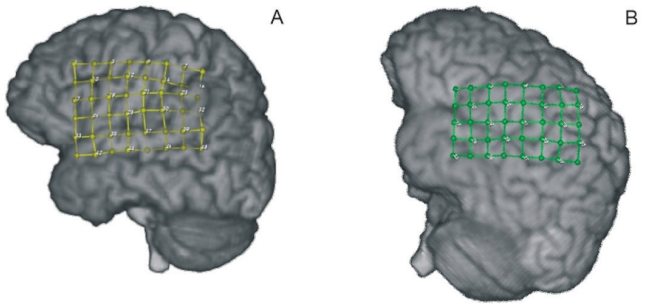
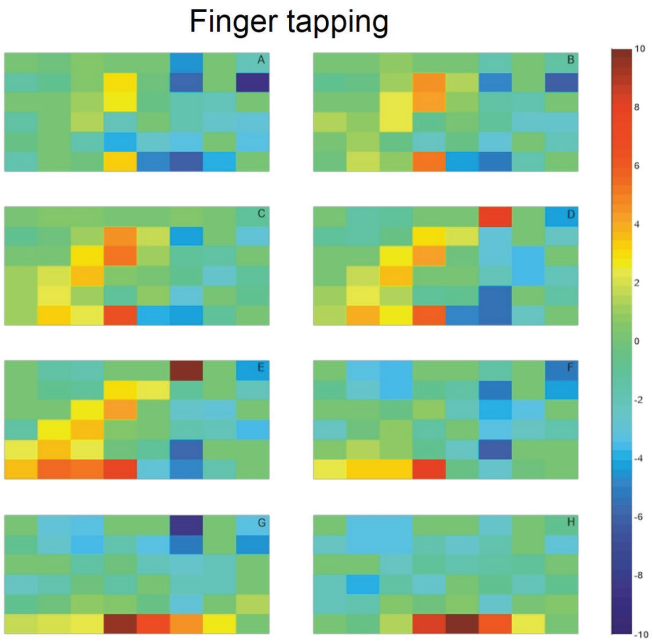


Figure 3. The subdural macro electrodes over the left central region of P5 (to the left, A) and also over the left central region of P8 (to the right, B). A: Electrode labeled '1' is the first upper left rectangle on the arrays on *Figure 4* (FT), while '48' is the last lower right. B: Electrode labeled '1' is the first upper left rectangle on the arrays on *Figure 4* (SSEP), while '40' is the last lower right.



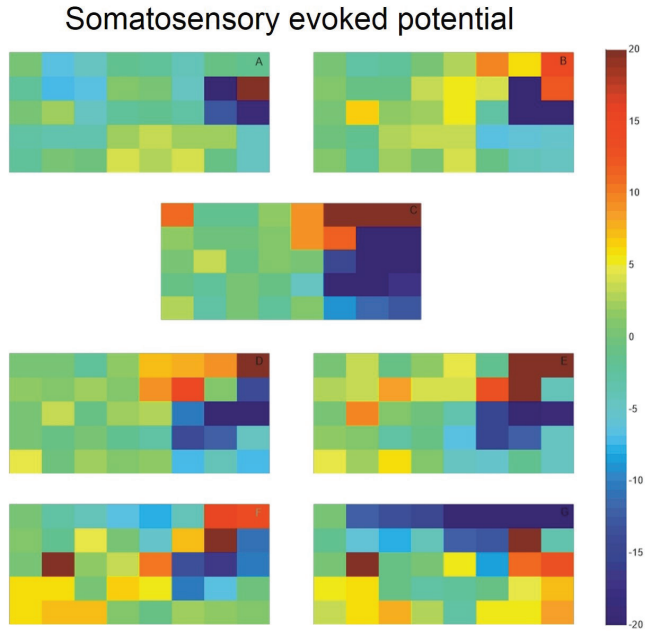


Figure 4. Snapshots of electrode potentials of the MAE at the time points shown in*Figure 2*. Blue colors correspond to negativity and red colors correspond to positivity.

B. Time frequency analysis

Decrease in spectral power was observed in the lower frequency band (8-30 Hz). This event related desynchronization (ERD) was detected across all examined patients (event: finger tapping, *Table 1*).The ERD typically began slightly before the FT event. Additionally, in a few cases this ERD was associated with an event related synchronization (ERS) in the high frequency band (*Figures5 and 6*). This particular phenomenonwas recorded on a single electrode while low frequency ERD was recorded on two electrodes on the cortex of patient P2. In general, the detected ERDs were spatially broad (observed at many electrodes) while the ERSs were spatially focal but both of them were restricted to the postcentral region (*Figure 5*).

Patient ID	ERD	ERS
1	X	X
2	X	X
3	X	
4	X	
5	X	
6	X	
7	X	
8	X	X

Table 1. Summary of patients showing low frequency ERD and high frequency ERS in the finger tapping experiment.

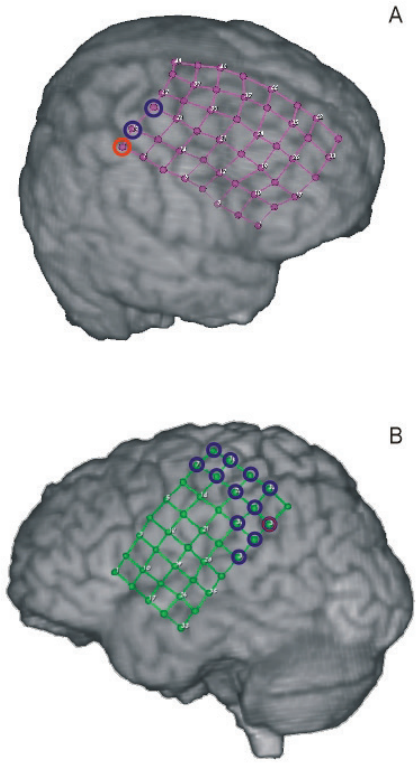


Figure 5. The MAE over the right hemisphere of P2 (A) and the left hemisphere of P8 (B). A: red and blue circlesrepresent ERS and ERD, respectively. The ERS is shown in the time-frequency domain in *Figure 6A*. B: ERD on many electrodes, time-frequency domain representaion of the recording at electrode 15 (the 7th electrode in the 2nd row of the grid) is in *Figure 6B*. The electrode marked with a red and blue circle corresponds to *Figure 6C*, where ERD and ERS take place simultaneously.

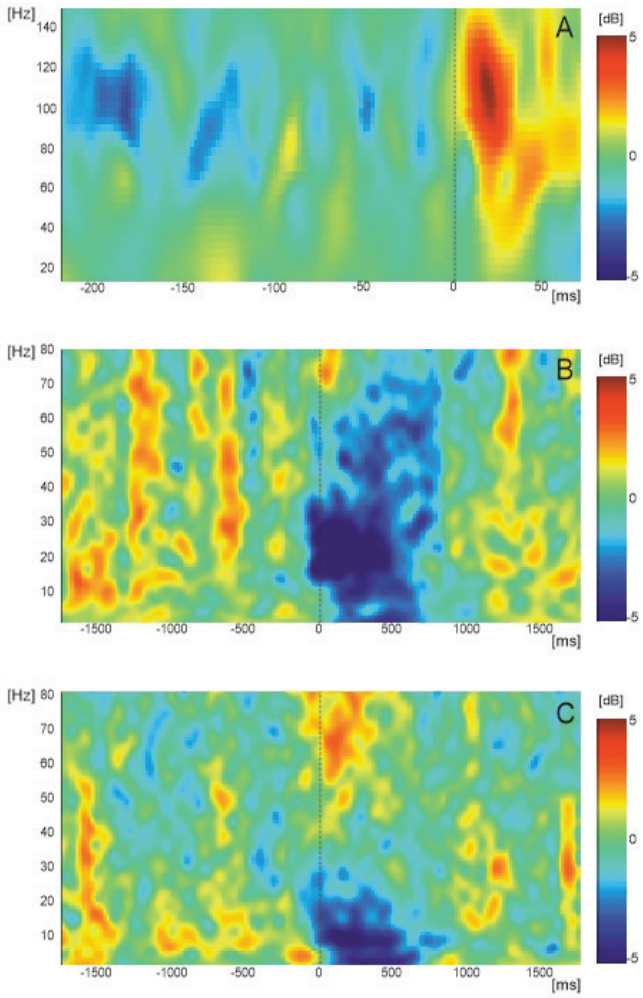


Figure 6. Time-frequency domain representation of recordings at various electrodes shown in *Figure 5*. A:Immediate ERS in the high frequncy range following FT event (dashed line: time 0). See *Figure 5A*, electrode labeled with red circle. B: ERD beginning slightly before the FT event (electrode 15 in *Figure 5B*). C: Simoltaneous low frequency ERD and high frequency ERS (electrode 39 in *Figure 5B*).

IV. Discussion

In contrast to the vast majority of motor related potential studies found in the literature our study was done usinginvasive intracranial, subdural electrodes that provide much higher spatial resolution.

The low amplitude MP negativity seemedto localize postcentrally. Time frequency analysis further confirmed this finding, both the detected ERDs and ERSs were localized to the same region. However, the spatial distribution of the synchronizations and desynchronizationsdiffered significantly. In almost all cases the ERDs were much more distributed along the somatosensory cortex than the ERSs, that was in line with a former study using intracranial electrodes [6], however, Pfurtschellerdescribedfocal ERD surrounded by ERS[7]. The clarification of the spatial properties of ERD and ERS patterns requires more recordings and analyses.

The detailed understanding of MRCPs can lead to better BCI applications. However we analyzed only effective movements, motor imaginary might be enough to create similarelectrical responses that have been evaluated in this paper. Patients without the ability to move their limbs can still imagine movementsand the electrical responses can be used in a BCI system to control e.g. an external robotic arm. In order to achieve the execution of fine movements with such a device, the detailed information of intracranial signals may be essential. We need further investigations to understandthe sensorimotor integrationand plan such precise BCI devices.

Our planned next step is to deduct quantified observations from the time-frequency analyses.

Acknowledgment

I would like to thank Dr. LorándErőss and Dr. DánielFabó for their continuous support in this research.

The research was supported by OTKA PD101754 grant.

References

[1] A. B. Jackson, M. Dijkers, M. J. DeVivo, and R. B. Poczatek, “A demographic profile of new traumatic spinal cord injuries: Change and stability over 30 years,” Archives of Physical Medicine and Rehabilitation, vol. 85, no. 11, pp. 1740-1748, 11//, 2004.

[2] N. G. N. Birbaumer, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub & H. Flor, “A spelling device for the paralysed,” Nature, vol. 398, no. 6725, pp. 297, 03/25/, 1999.

[3] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. van der Smagt, and J. P. Donoghue, “Reach and grasp by people with tetraplegia using a neurally controlled robotic arm,” Nature, vol. 485, no. 7398, pp. 372-375, 05/17/print, 2012.

[4] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain–computer interfaces for communication and control,” Clinical Neurophysiology, vol. 113, no. 6, pp. 767-791, 6//, 2002.

[5] M. A. Lebedev, and M. A. Nicolelis, “Brain–machine interfaces: past, present and future,” TRENDS in Neurosciences, vol. 29, no. 9, pp. 536-546, 2006.

[6] K. J. Miller, E. C. Leuthardt, G. Schalk, R. P. Rao, N. R. Anderson, D. W. Moran, J. W. Miller, and J. G. Ojemann, “Spectral changes in cortical surface potentials during motor movement,” The Journal of neuroscience, vol. 27, no. 9, pp. 2424-2432, 2007.

[7] G. Pfurtscheller, “The cortical activation model (CAM),” Progress in brain research, vol. 159, pp. 19-27, 2006.

Spatially and temporally clustered inputs generate dendritic calcium spikes and local membrane oscillations with ripple frequency and in parvalbumin interneurons

Dénes Pálfi
(Supervisor: Dr. Balázs Rózsa)
palfi.denes@itk.ppke.hu

Abstract— **Parvalbumin expressing, fast spiking interneurons (PV-FS-INs) are one of the key elements in feedforward and feedback inhibition networks which play an important role in different brain oscillations such as sharp wave ripple activity (SPW-R). Using two-photon glutamate uncaging in spatiotemporally clustered input patterns we could reproduce spontaneous SPW-R associated Ca^{2+} responses. We characterized the input-output function by plotting the simultaneously-recorded somatic membrane potential and Ca^{2+} responses as a function of the progressively increasing input number. At a given active input number (first threshold), the uncaging-evoked Ca^{2+} signal showed a supralinear increase and reproduced the spontaneous SPW-R associated local dendritic Ca^{2+} responses (dendritic hot spots). The first threshold was followed by a separate second threshold where the uncaging-evoked Ca^{2+} signal showed a sigmoid-like increase outside the zone of inputs (lateral dendritic region). Although Ca^{2+} transient slightly decreased as a function of distance from the border of the hot-spot region, it reproduced the large propagating Ca^{2+} waves we measured during SPW-Rs. These propagating Ca^{2+} waves were mediated dominantly by L-type voltage gated Ca^{2+} channels. In addition, these Ca^{2+} signals were well related with somatic membrane potential oscillations which were mediated voltage-gated Na^{+} channels. These oscillations had the same frequency as the SPW-R-associated field potential oscillations. We found that the integration mode of PV-FS-INs has changed, instead of the fast EPSP-AP coupling the APs outputs are coupled to the membrane ripple oscillations now. Our results indicate that short dendritic segments of parvalbumin interneuron generate dendritic calcium spike and oscillations with ripple frequency.**

***Keywords*-two photon microscopy; glutamate uncaging; sharp wave ripples; parvalbumin expessing interneurons**

I. INTRODUCTION

In the hippocampus, sharp wave ripples (SPW-Rs) are transient oscillatory events which reflect synchronized population discharges corresponding to the reactivation of previously established cell assemblies, and play a crucial role in establishing long-term memory traces in the neocortex[1]. According to the current model of hippocampal SPW activity, first a transient increase in pyramidal cell firing is generated autonomously in CA3, and this drives a widespread depolarization in both pyramidal cells and interneurons in CA1, leading to the generation of network ripple oscillations[2]. The precisely timed and anatomically structured input activity of CA3 neurons is nonlinearly transformed to neuronal output by the somatic and dendritic compartments of downstream neurons[3].Nonlinear dendritic processing is achieved mainly by voltage-gated ion channels, which interact through locally propagating and attenuating membrane potential fluctuations; in this way, dendritic signal integration can be clustered in small (~10 μm) dendritic computational subunits ('hot-spots'). When more synaptic inputs are activated in synchrony, voltage-gated ion channels can also induce more global signals, i.e. regenerative dendritic spikes[4]. The relationship between network ripple oscillations, SPWs, and dendritic hot-spot activity has not yet been investigated; nevertheless, several facts suggest that SPW-R-associated cell assemblies can activate dendritic hot-spots. First, synchronized cell assemblies have been shown to activate dendritic hot-spots[5], and during an SPW-R event up to 10% of the total neuronal population discharges in the hippocampus, making SPW-Rs the most synchronized cell assembly pattern in the cortex[1]. Second, synaptic inputs in dendritic hot-spots have been reported to be locally synchronized for an interval of around 60 ms[6], which matches the average length of individual SPW-R events[2]. However, whether and how these SPW-R-associated cell assemblies activate dendritic hot-spots, and if this activation changes the dendritic computation and action potential output of individual neurons have not been investigated yet.

Although many different types of cells are recruited during SPW-R events, hippocampal FS-PV INs are substantially more active than other neuron types, and their firing is strongly phase-locked to network ripple oscillations[7]. It has been suggested that FS-PV INs play key roles in the generation of SPW-Rs and during other synchronized cell assembly activities[2]. According to the generally accepted view, FS-PV INs act in cortical circuits as fast and, essentially, passive integrators of synaptic inputs[8]. Under conditions of low network activity, when the incoming synaptic activity is low, the passive and active characteristics of FS-PV INs result in accelerated kinetics of excitatory postsynaptic potentials (EPSPs), a reduced, sub-millisecond temporal window for

D. Pálfi, "Spatially and temporally clustered inputs generate dendritic calcium spikes and local membrane oscillations with ripple frequency and in parvalbumin interneurons," in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds., Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 47-50.

dendritic integration, and precise and fast coupling between EPSPs and action potential (AP) outputs[8]. Ca^{2+} dynamics in the aspiny dendrites of FS-PV INs has also been found to be fast and confined to small, approximately 1 μm long, dendritic microdomains[9]. Moreover, it has not been possible to evoke regenerative dendritic spikes in these cells, and back-propagating action potentials are severely attenuated. However, under physiological high-activity conditions such as SPW-Rs, when neurons are bombarded with precisely timed dendritic inputs, the rules of dendritic integration and EPSP-AP coupling can be different [10].

Here, using two-photon scanning in combination with electrophysiological recordings, and employing a novel caged glutamate compound, we show that the classical view of the operation and functional role of FS-PV INs in cortical circuits is incorrect when these cells are activated by large synchronous cell assemblies. We demonstrate SPW-Rs can generate a transition in FS-PV INs from the well-documented ground state of passive integration to an active state, where dendritic processing and the input-output transformation of these neurons are fundamentally different. Moreover, our data suggest a completely novel mechanism for the generation of SPW-Rs by demonstrating that the smallest functional units that can generate ripple-frequency oscillations in the brain are short (~ 20 μm) segments of FS-PV IN dendrites.

II. Spikes are mediated by L-type Ca2+ channels and ripples by Na+ channels

In order to investigate the functional role of different ion channels in the mechanism underlying dendritic ripples and spikes, we activated spatiotemporally clustered patterns which were above the second threshold, but below the threshold for somatic AP generation (43.8±2.9 active inputs). Long dendritic segments (Figures 1A, 1D) were selected, and inputs were activated at only one end of the imaged (Fluo4 and Alexa 594) dendritic segments in order to better separate the central and lateral dendritic regions in the pharmacological experiments. For a precise quantification of the pharmacological effects, we had to take into account the saturation and nonlinear response of the Ca^{2+} dye, and we therefore transformed the relative fluorescence data into $[Ca^{2+}]$.

AMPA and NMDA receptors play a dominant role in triggering dendritic Ca^{2+} response: the combined application of AMPA and NMDA receptor blockers (CNQX and AP5, respectively) reduced the Ca^{2+} responses to almost zero (Figure 1G) in both regions. In the lateral dendritic region, VGCC blockers evoked a drastic reduction in the Ca^{2+} signal (Figures 1A-1C, 1G). Therefore, our data indicate that VGCCs are mainly responsible for the Ca^{2+} influx in the lateral dendritic region, and are thus responsible for the dendritic Ca^{2+} spike. Hippocampal interneurons express P/Q-, R-, L-, N-, and T-type VGCCs, but we found that the L-type VGCC blocker nimodipine had the greatest effect on the Ca^{2+} responses of FS-PV INs, both in the present study and in our earlier work[10]. In the central hot-spot region, the source of Ca^{2+} influx was more complex as it was simultaneously mediated by NMDA, calcium-permeable AMPA receptors, and VGCCs, and further amplified by Na^{+} channels (Figure 1). In line with other observations, we noted that Ca^{2+}-permeable AMPA receptors had a larger effect on the postsynaptic Ca^{2+} influx than NMDA receptors[9]. The amplitude and area of the simultaneously recorded uncaging-evoked EPSPs was significantly decreased by all the ion-channel blockers we tested (Figures 1H and 1I). This change in the somatic EPSP likely reflects a similar change in the local dendritic voltage response, which may by itself affect the generation of membrane ripple oscillations. To test this possibility, we compensated for the EPSP amplitude drop by increasing the uncaging laser intensity when oscillations disappeared until the amplitude of the uncaging-evoked EPSPs reached the control value again, or interneuronal ripple oscillations reappeared. Under these circumstances, the oscillations recovered (or remained stable) for all pharmacons except during TTX application, which completely eliminated interneuronal ripple oscillations, indicating that they are dependent on Na^{+} channel activation (Figures 1I and 1J). In contrast, the oscillation frequency did not change significantly in the presence of AP5, nimodipine, IEM-1460, or the cocktail of VGCC blockers (Figure 1J). Our results show that the propagating dendritic Ca^{2+} spikes are predominantly mediated by L-type Ca^{2+} channels, while the related interneuronal ripple oscillations are determined by voltage-gated Na^{+} channels. In summary, we can conclude that dendritic spikes exist in FS-PV INs.

To further validate the dendritic origin of interneuronal ripple oscillations, we locally injected TTX (10 μM) onto the axosomatic region (Figure 2A). The local TTX application eliminated all APs induced by somatic current injection, but it did not block interneuronal ripple oscillations (control 212.83±24.18 Hz; TTX puff 182.72±18.72 Hz, paired t-test, p=0.156; Figures 2B-2D). This confirms their dendritic origin. In order to detect local membrane potential oscillations more directly, in a different set of experiments we combined whole-cell recordings with two-photon guided dendritic juxtacellular recordings (Figures 2E-2G) at distal dendritic locations (266.37±67.05 μm, mean±s.d.). Oscillations were induced by using spatiotemporally clustered input patterns, as above. Interneuronal ripple oscillations could be simultaneously detected in the somatic membrane potential and in juxtacellular signals (Figure 2F) at distal dendritic locations, where the bAP-induced juxtacellular signal was below the detection threshold (Figure 2G). The oscillation amplitude in the LFP signal decreased as the dendritic recording pipette was gradually moved away from the activated dendritic segment, but remained high enough to be detected at distances up to about 40 μm (Figures 2H-2I, spatial decay constant 15.6±4.9 μm, n=4). Finally we showed that, in contrast to somatic AP, the threshold of interneuronal ripple oscillations changed in proportion to the somatic membrane potential, which reflects the weak control over the oscillations by the somatic membrane potential. These data, together with the local uncaging experiments and propagation measurements strongly support the dendritic origin of the interneuronal ripple oscillations.

III. Discussion

In this study we propose a novel, dendritic hot-spot-related mechanism to be integrated into the currently accepted network model of SPW-R activity[2]. According to our results, membrane oscillations at ripple frequencies can be generated following strong depolarizing events in the dendrites of FS-PV INs. In intact hippocampal circuits, this depolarization can be provided predominantly by the synchronized firing of CA3 cell assemblies, which are directly responsible for the envelope of the SPW events[2]. Smaller cell assemblies can also provide the required depolarization, as individual CA3 subfields, and CA1 (or CA3) minislices have also been shown to be capable of generating SPW-Rs[11]. Moreover, local application of KCl to the dendritic layer, with a complete blockade of GABAergic and glutamatergic synaptic transmission, reproduced SPW events and was also capable of generating the associated network ripple oscillations in CA1 minisclices. These data also suggest that a single depolarization event in dendrites without any internal pattern is capable of activating intrinsic membrane mechanisms which then generate the ripple-oscillations. In line with this prediction, we showed that a single activation of clustered glutamatergic inputs in the distal dendrites of FS-PV INs, which generate a depolarizing hump and reproduce the hot-spots associated with spontaneous SPW-R events, is also capable of generating secondary membrane oscillations in the ripple frequency range. In summary, we can say that the phase-locked firing during SPW-Rs is not a simple reflection of the discharge pattern of presynaptic cell assemblies, but oscillations can be formed actively and intrinsically by the dendritic membrane.

Although smaller and smaller parts of the hippocampal formation have recently been shown to be possible sources of ripple oscillations, ripples have been exclusively considered as a network phenomenon. Here we demonstrate that the smallest functional unit which can generate fast oscillations in the ripple frequency range when activated by approximately 30 coincident active inputs is a short (~20 μm) segment of a dendrite, a dendritic hot-spot. Although the duration of the interneuronal ripple oscillations increased upon increasing active input number, and the onset latency decreased, the oscillation frequency remained stable, suggesting that these dendritic ripple generators are the integrated circuit elements which provide the stable ripple frequency of the network oscillation during SPWs. Interneuronal ripple oscillations could be detected in the local field potential at distances up to a few micrometers from the activated dendritic segments, but whether, and how, these independent dendritic oscillators interact within the dendritic arbor of the same neuron, and throughout the gap-junction-connected dendritic network of interneurons, and how they finally give rise to the local field potential, remains to be investigated.

References

[1] Buzsaki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. Neuron 68, 362-385.

[2] Buzsaki, G., and Silva, F.L. (2012). High frequency oscillations in the intact brain. Progress in neurobiology 98, 241-249.

[3] Katona, G., Kaszas, A., Turi, G.F., Hajos, N., Tamas, G., Vizi, E.S., and Rozsa, B. (2011). Roller Coaster Scanning reveals spontaneous triggering of dendritic spikes in CA1 interneurons. Proc Natl Acad Sci U S A 108, 2148-2153.

[4] Larkum, M.E., Nevian, T., Sandler, M., Polsky, A., and Schiller, J. (2009). Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle. Science 325, 756-760.

[5] Kleindienst, T., Winnubst, J., Roth-Alpermann, C., Bonhoeffer, T., and Lohmann, C. (2011). Activity-dependent clustering of functional synaptic inputs on developing hippocampal dendrites. Neuron 72, 1012-1024.

[6] Takahashi, N., Kitamura, K., Matsuo, N., Mayford, M., Kano, M., Matsuki, N., and Ikegaya, Y. (2012). Locally synchronized synaptic inputs. Science 335, 353-356.

[7] Klausberger, T., Magill, P.J., Marton, L.F., Roberts, J.D., Cobden, P.M., Buzsaki, G., and Somogyi, P. (2003). Brain-state- and cell-type-specific firing of hippocampal interneurons in vivo. Nature 421, 844-848.

[8] Pouille, F., and Scanziani, M. (2004). Routing of spike series by dynamic circuits in the hippocampus. Nature 429, 717-723.

[9] Topolnik, L. (2012). Dendritic calcium mechanisms and long-term potentiation in cortical inhibitory interneurons. Eur J Neurosci 35, 496-506.

[10] Chiovini, B., Turi, G.F., Katona, G., Kaszas, A., Erdelyi, F., Szabo, G., Monyer, H., Csakanyi, A., Vizi, E.S., and Rozsa, B. (2010). Enhanced dendritic action potential backpropagation in parvalbumin-positive basket cells during sharp wave activity. Neurochem Res 35, 2086-2095.

[11] Maier, N., Nimmrich, V., and Draguhn, A. (2003). Cellular and network mechanisms underlying spontaneous sharp wave-ripple complexes in mouse hippocampal slices. J Physiol 550, 873-887.

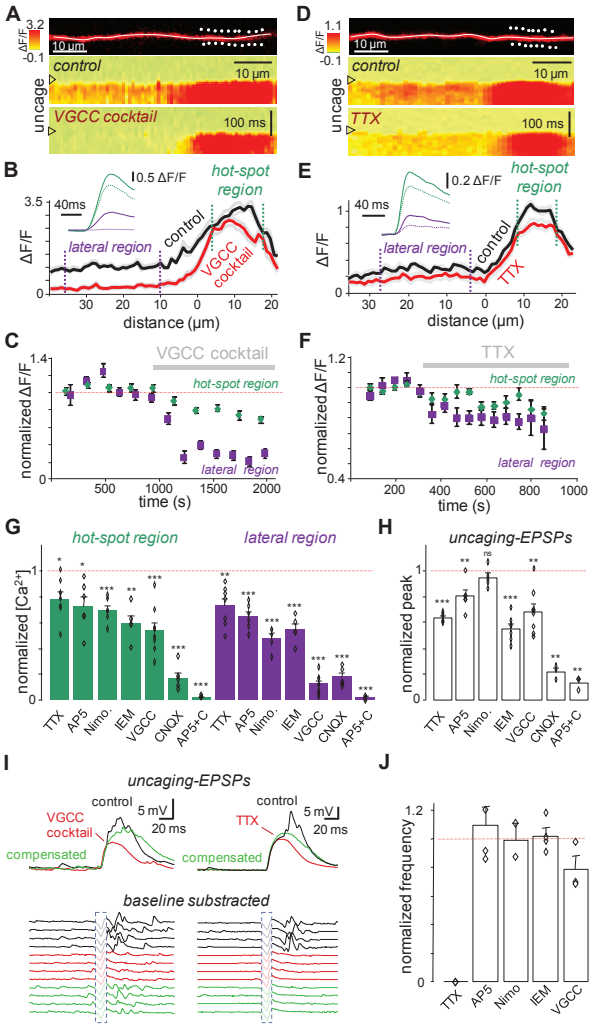


Figure 1. Dendritic Ca2+ spikes are mediated predominantly by L-type Ca2+ channels; the related interneuronal ripple oscillations are generated by Na+ channels. (A-C) Effect of VGCC blockers on uncaging-evoked Ca2+ responses. (A) Maximum intensity z-projection image of a distal dendritic segment of an FS-PV IN. Average uncaging-evoked Ca2+ responses in control conditions (middle), and in the presence of a cocktail of VGCC blockers (bottom). White points are active input locations used for DNI-Glu•TFA uncaging (top). (B) Spatial distribution of the peak dendritic Ca2+ response (mean ± s.e.m.) measured along the white line in A under control conditions (black) and in the presence of VGCC blockers (red). Inset: mean Ca2+ transients derived from the hot-spot (green) and lateral dendritic (magenta) regions before (solid line) and after (dashed line) application of the VGCC cocktail. (C) Time-course of the effect of the VGCC cocktail on Ca2+ responses in the hot-spot (green) and lateral dendritic (magenta) regions. (D-F) The same as A-C, respectively, but for TTX. (G) Effect of different ion channel blockers on the peak amplitude of [Ca2+]. Nimo. and AP5+C indicate nimodipine and AP5+CNQX, respectively. (H) The same as G, but for simultaneously recorded EPSPs. (I) Subthreshold EPSPs showing interneuronal ripple oscillations with (bottom) and without (top) baseline subtraction in control conditions (black), but not in the presence of TTX or the VGCC cocktail (red traces left and right respectively). When the uncaging laser intensity was increased (compensated), interneuronal ripple oscillations were restored in the presence of VGCC blockers, but not when TTX was present. (J) The effect of ion-channel blockers on interneuronal ripple oscillations. Ripple oscillations were only abolished by TTX.

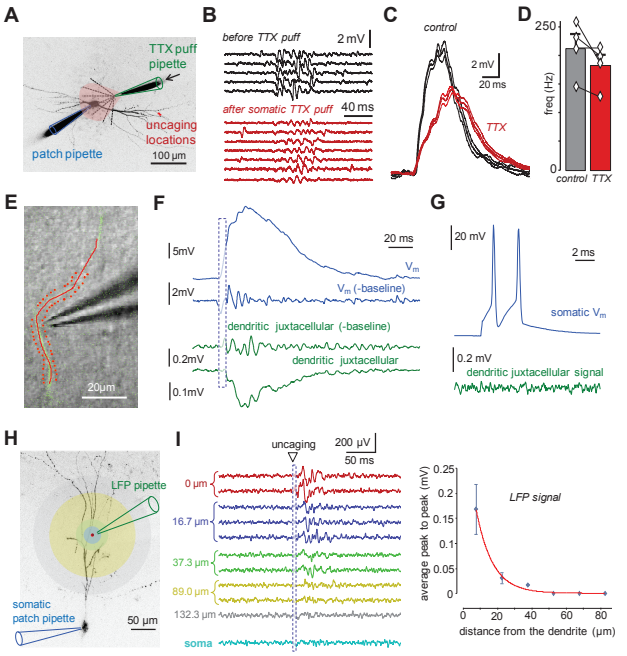


Figure 2. Interneuronal ripple oscillations can be detected by juxtacelular recording and are preserved following somatic TTX injection. (A) An inverted image of an FS-PV IN with the somatic recording and local TTX injection pipettes. (B) Representative somatic EPSPs after baseline subtraction show interneuronal ripple oscillations in control conditions (black) and after local TTX injection (red). Oscillations were induced by using spatiotemporally clustered input patterns in the distal dendritic location indicated in A. (C) Average of the EPSPs (mean ± s.e.m.). (D) Oscillation frequencies in the control case and after local TTX injection were not significantly different (n=4 cells). (E) Overlaid transmitted gradient and two-photon images of a dendritic segment from an FS-PV IN. (F) The juxtacellular signal (green) from the dendritic location in E, and the simultaneously recorded somatic membrane potential (blue) are shown with and without baseline substraction. Interneuronal ripple oscillations were induced by uncaging at the red points in E. (G) Somatically evoked bAPs were below the detection threshold in the juxtacellular signal (green) in the dendritic location in E. (H) The second patch pipette was used to record the LFP signal at variable distances from the dendrite. The same inputs, in the same dendritic region, were activated as shown in E to induce interneuronal ripple oscillations. (I) Individual LFP signals and amplitude of the oscillations (mean ± s.e.m., n=4) as a function of distance from the activated interneuronal dendritic segment. Red line is an exponential fit.

Chemical heating for non-instrumented nucleic acid amplification on Lab-on-a-Chip devices

Pardy Tamás
(Supervisor: dr. Iván Kristóf)
parta@digitus.itk.ppke.hu

***Abstract*—Nucleic acid amplification and detection has a wide variety of applications, ranging from the detection of infectious diseases, cancer, suspect identification in forensics to auditing in the food industry. However, due to the cost of the equipment and operating laboratory staff, these assays are not available to a significant proportion of people in need of them. Loop-mediated isothermal amplification (LAMP) is a simple, rapid, cost-effective nucleic acid amplification method, which requires a constant 65$^\circ C$ sample temperature during the process, and thus is well-suited for non-instrumented nucleic acid amplification assays, including disposable Lab-on-a-Chip (LOC) based rapid tests. In this paper, possible implementations of a heater that could be fully integrated into LOC devices are reviewed. In these devices, various highly exothermic reactions provide the necessary heat output. However, due to the rapid and intense nature of these reactions, heat output must be regulated, to which end phase-change materials (PCMs) are specifically engineered to melt in a well-defined temperature range. In this paper, various aspects of the available chemical heaters are reviewed. In addition, possibilities for numerical simulation via finite element analysis of chemical heating are evaluated.**

***Keywords*-NINA, LOC, LAMP, Lab-on-a-chip, Non-instrumented nucleic acid amplification, Loop-mediated isothermal amplification**

I. Introduction

Nucleic acid amplification, as of today, is a well-known and widespread method to identify the presence or absence of various living organisms in a wide variety of samples, ranging from food samples, or human bodily fluids to evidence collected from crime scenes. However, the tried and true method of polymerase chain reaction (PCR) requires a thermal cycler and trained personnel to use for amplification purposes, which limits its applicability due to its high cost, especially in a low-resource environment. Loop-mediated isothermal amplification (LAMP) is an isothermal nucleic acid amplification method developed by Eiken Chemical Co., Ltd., which is a rapid, reliable and cost-effective alternative to PCR. It can amplify a DNA sample $10^9 - 10^{10}$ times in 15-60 minutes, and it only requires a constant $65^\circ C$ temperature to be maintained during the amplification process [1], which removes the need for a thermal cycler, and this criterion can be met using simpler equipment, such as an electric hot plate with adequate temperature control. Therefore LAMP is suitable for application in so-called Non-instrumented nucleic acid amplification tests (NAAT) that do not require sophisticated equipment to function and are therefore cost-efficient and portable. Heating in NINA devices can be performed using highly exothermic reactions. Most of these are two-component oxidative reactions with a high energy output, which has to be regulated in order to keep the appropriate temperature range. This regulation is done via the inclusion of a specifically engineered phase change material (PCM) that melts in a certain temperature range, and through its melting, its heat conducting properties change, enabling it to carry away excess heat. In addition, phase change materials can dampen variations in ambient temperatures.

II. Background of device components

A. Exothermic reactions

The investigated exothermic reactions can be ordered into several subcategories. One category is that of bi-or multi-component oxidative or corrosive reactions, which provide a considerable amount of heat. One prominent example of these is the reaction of calcium oxide and water ($CaO(s) + H_2O(l) \rightleftharpoons Ca(OH)_2(aq) + heat$). However, in the case of this reaction the volume change of CaO as it hydrates has to be considered in the device geometry. Although CaO is cheap and easy to acquire, its quality and heat output may vary. Another very prominent example is that of the field ration heaters of the US Military (MRE), which are based on the reaction of water with a magnesium-iron alloy ($Mg + 2H_2O \rightarrow Mg(OH)_2 + H_2 + heat$), supplemented by salt and activated carbon to catalyze the galvanic corrosion of iron. The heater also contains anti-foaming agents, and are stored in a paper package which aids uniform hydration. The MgFe alloy is an engineered material, commercially available, and therefore has a consistent heat output, and reliable quality. This reaction has a much higher heat output than calcium oxide hydration, but the reaction results in the release of hydrogen gas, which has to be accounted for in the design (0.8 liters of gas per grams of MgFe fully reacted). The reaction of iron and air ($4Fe(s) + 3O_2(g) \rightarrow 2Fe_2O_3(s)$) is very exothermic with a heat output higher than the reactions listed before, but its heat output is hard to control [2].

The reactions described so far involved the hydration of solids. However, the reaction could be fully liquid-based. Gujit et al [3] proposed a device in which the reaction of sulfuric acid with water would provide the heat output and the evaporation of acetone could be used to cool the reactor. The application of sulfuric acid might, however, raise some safety concerns and therefore safe storage and disposability must be design considerations.

T. Párdy, "Chemical heating for non-instrumented nucleic acid amplification on Lab-on-a-Chip devices,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 51-54.

Manufacturer	Product name	Melting temperature [$^\circ C$]	Latent heat [kJ/kg]
Rubitherm	RT 60	60	144
	RT 62	62	146
	RT 65	65	152
PureTemp	PureTemp 60	61	230
	PureTemp 63	63	199
PlusICE	A60	60	145
	A62	62	145

TABLE I
COMMERCIALLY AVAILABLE PHASE CHANGE MATERIALS

Crystallization reactions were also considered, such as the crystallization of a supersaturated solution of sodium acetate, which releases heat upon crystallization. This reaction enable reusing the heater as the crystals can be dissolved in boiling water in a few minutes, however, the reaction has a low heat output of $58^\circ C$, which makes it inappropriate for LAMP. In theory, acid-base neutralization reactions could also be used to provide heat output.

B. Phase change materials

Phase change materials (PCM) are engineered to melt in a certain well-defined temperature range, and melting changes their heat conducting properties. Latent heat of melting [kJ/kg] is the energy absorbed or released when this phase change occurs. A well-chosen PCM can store the excess energy output of the exothermic reactions and in addition, dampen the variations in ambient temperature by melting and solidifying as temperatures fluctuate. Some of the currently market-available PCMs that melt in the 60-65 $^\circ C$ temperature range appropriate for LAMP are listed in Table 1.

PCMs are recyclable and nontoxic, which makes them suitable for use in disposable devices. The PATH group [2] proposed a method to calculate PCM volume requirements for a specific ambient temperature range:

$$V_{PCM} = \frac{(T_H - T_L)t}{R_{sys} L_{PCM} \rho_{PCM}} \tag{1}$$

where V_{PCM} is the volume of a specific PCM required to maintain assay temperatures over a given assay time t [s] in a specific ambient range (T_H [K] the high ambient limit, T_L [K] the low ambient limit). L_{PCM} is the latent heat of fusion of the PCM [J/kg], R_{sys} is the overall thermal resistance of the assay device [$(m^2 \cdot K)/W$] and $\rho[kg/m^3]$ is the density of the PCM. However, in addition to the phase change materials, device insulation is necessary to consider in the design. The reactor volume is too small to compensate for the heat dissipation into the ambient, therefore unless proper additional insulation is included, the reactor will cool down very quickly.

III. METHODS FOR FINITE ELEMENT ANALYSIS

For the introduction of the methods for numerical analysis, several basic assumptions are made. Firstly, the LAMP reaction requires a constant 60-65$^\circ C$ temperature in the amplification area. This temperature is produced by an exothermic reaction. Secondly, we assume that the heated are contains a pure DNA solution mixed with the proper primers to start the reaction, and that over the time frame of the reaction flow rate is sufficiently low to consider the liquid in the amplification are static. Space-dependence in the model is defined by a Cartesian coordinate system with 2 axes x [mm],y [mm], z [mm], whereas time-dependence is described by time t [s]. Every equation described in the model is meant to be interpreted for points in this coordinate system, therefore such dependencies are not mentioned later on.

To model chemical heating, a two-component exothermic reaction is considered, which is valid in most cases. The generation and transfer of heat is modeled by 3 separate sets of equations: reaction heat is generated by a reaction model, which uses the concentrations from a convection-diffusion equation and the heat of which is transferred by a heat transfer equation.

The heat transfer equation, at zero velocity is the following:

$$\rho C_p \frac{\partial T}{\partial t} + \nabla \cdot (-k \nabla T) = Q \tag{2}$$

where $\rho[kg/m^3]$ is the fluid density, C_p [J/kg K] the specific heat capacity at constant pressure, T [K] the absolute temperature in the system, k [W/(mK)] the thermal conductivity, and Q [W/m^3] is the heat source, which in this case is calculated from the reaction model.

Transport of species in the model is described by a convection-diffusion equation:

$$\nabla \cdot (-D_i \nabla c_i) + \underline{u} \cdot \nabla c_i = R_i \tag{3}$$

where D_i [m^2/s] is the diffusion coefficient, c_i [mol/m^3] is the species concentration, u [m/s] is the flow velocity of the flow field, and R_i [$mol/(m^3 s)$] is the reaction rate where i is the index of the chemical species. The assumption is made that diffusion through the walls is zero: $\underline{n} \cdot (-D_i \nabla c_i) = 0$, where $\underline{n}$ is a vector normal to the wall.

The reaction model in the heater is based on the Arrhenius equations to calculate rate constants (k), which in turn are used to calculate reaction rates (r) based on the concentration of species at a given spatial coordinate at a given time point. Heat output (Q) is calculated by multiplying the reaction rate with the standard enthalpy of formation (H). The hydration reaction of Calcium oxide is a second-order reaction, which is irreversible under the examined conditions and can be modeled as follows:

$$A + B \rightarrow C \tag{4}$$

Where A (Calcium oxide) and B (water) are the reagent species, and C is the product species. The Arrhenius equation is the following:

$$k = A e^{-E_a/(RT)} \tag{5}$$

Where k is the rate constant of the aforementioned reaction, E_a the activation energy [J], R the Universal gas constant (8.314 $[J/(mol \cdot K)]$), and T the temperature [K] of the system at the given spatial coordinate and time. Reaction rates are calculated as follows:

$$r = k \cdot c_A c_B \tag{6}$$

where r $[mol/(m^3 s)]$ is the reaction rate, by which the concentration $[mol/m^3]$ of species A and B decreases, and that of species C increases. The reaction heat, Q $[W/m^3]$ is calculated by multiplying the reaction rate with the enthalpy of formation (H [J/mol]) of the product species.

IV. Results to date

Singleton et al. from the PATH group demonstrated the first successful implementation of a NAAT device for malaria, based on the hydration reaction of Calcium oxide, the heat output of which was regulated by an engineered phase change material (EPCM). They managed to produce and maintain a temperature range of $62 - 65°C$ over 45 minutes using 77g of CaO, 17 g of water and a PCM that melted at $65°C$ in stainless steel tubes [4]. Hatano et al. successfully performed a LAMP reaction for the detection of anthrax using a system which consisted of a Styrofoam^{TM} cup, in which the active substance from disposable pocket warmers provided the necessary heat output, which was regulated by a chamber of PCM around the sample tubes [5]. Bau et al. developed a device that is more compact and smaller than the devices described so far, and most closely resembles a Lab-on-a-Chip setup (see Figure 1). In this particular device, the heating material of the Meal, Ready-to-Eat (MRE) packets of the US military are used, namely the Mg-Fe alloy mentioned earlier, which, upon coming into contact with water produces a significant amount of heat. The wetting of the reaction material is done with a strip of filter paper, which reaches into a water reservoir that is filled by the user. This setup enables the control of heating intensity and duration via the width of the filter paper strip. Additionally, the PCM surrounding the amplification chamber absorbs not only internal excess heat, but also external, meaning that the device can tolerate a wide range of ambient temperatures (22-40 $°C$). 0.36g of Mg-Fe alloy is used to provide the necessary heat output, which is an amount small enough to fit in Lab-on-a-Chip devices. The water reservoir is sized to contain 1.5 ml of water, but 0.8 ml is enough to fully wet the alloy. The device was demonstrated to work by the detection of E. coli DNA [6]. The PATH group proposed a paper-based single used disposable (SUD) system for NAAT, using the aforementioned Mg-Fe alloy and a PCM that melts at $58°C$, to produce and maintain the appropriate temperature range. Their system was designed so that it accepted a sample paper membrane and was activated by piercing a package of saline, which then came into contact with the Mg-Fe alloy, thus starting the reaction. A thick insulating layer protected against changes in ambient temperature [2].

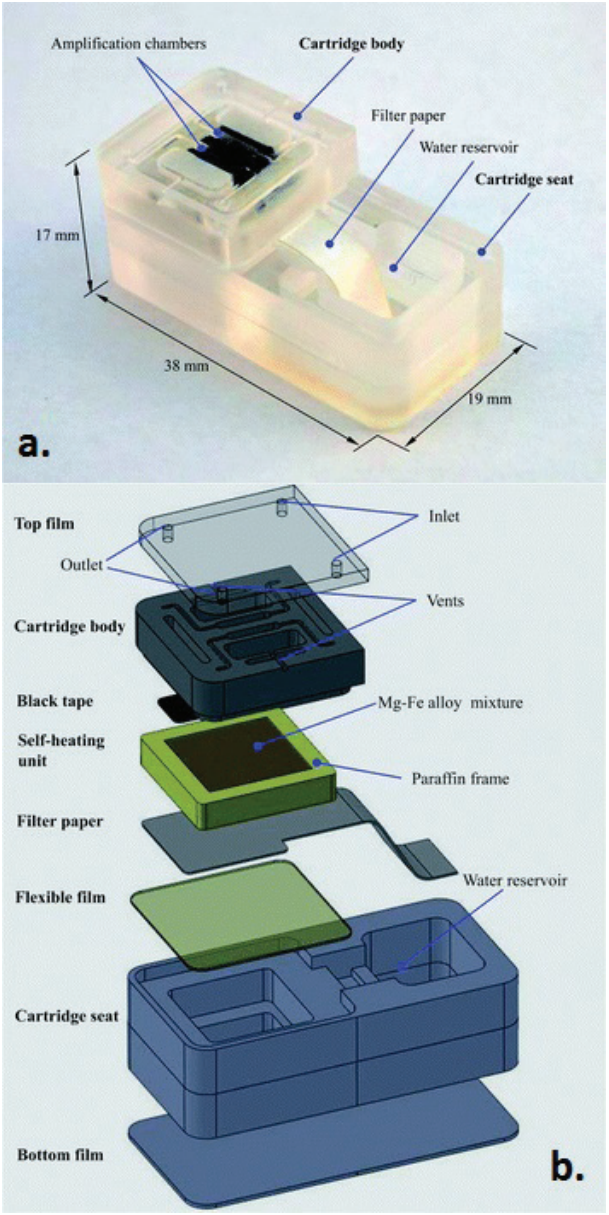


Fig. 1. A successful implementation of a Lab-on-a-chip device that employs chemical heating to perform non-instrumented nucleic acid amplification. Assembled device in isometric view (a.), and the exploded view of the device (b.) are shown.[7]

V. Conclusion

Non-instrumented isothermal nucleic acid amplification has been successfully performed using chemical heaters that employ highly exothermic reactions as the heat source, and phase change materials to regulate the heat output. Various setups were demonstrated to function with high amounts of reagents, in portable devices, the majority of which were still too large to meet the size criteria of Lab-on-a-Chip devices. The scaling down of these devices is made hard by the volume requirements of the chemical heater, as consistent, high heat output and sufficient insulation are key aspects and can prevent further miniaturization. However, a successful implementation

that can be considered a Lab-on-a-Chip device already exists and was demonstrated to work, opening new avenues for future research.

REFERENCES

[1] K. Nagamine, T. Hase, and T. Notomi, "Accelerated reaction by loop-mediated isothermal amplification using loop primers," Molecular and Cellular Probes, vol. 16, no. 3, pp. 223–229, Jun. 2002.

[2] J. Singleton, C. Zentner, J. Buser, P. Yager, P. LaBarre, and B. H. Weigl, "Instrument-free exothermic heating with phase change temperature control for paper microfluidic devices," 2013, p. 86150R.

[3] Guijt, R. M., Dodge, A., van Dedem, G. W., de Rooij, N. F., & Verpoorte, E., "Chemical and physical processes for integrated temperature control in microfluidic devices.", Lab on a Chip, 2003

[4] P. Labarre, J. Gerlach, J. Wilmoth, A. Beddoe, J. Singleton, and B. Weigl, "Non-instrumented nucleic acid amplification (NINA): instrument-free molecular malaria diagnostics for low-resource settings.," Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, vol. 2010, pp. 1097–9, Jan. 2010.

[5] Hatano, Ben, et al. "LAMP using a disposable pocket warmer for anthrax detection, a highly mobile and reliable method for anti-bioterrorism." Jpn J Infect Dis 63.1 (2010): 36-40.

[6] H. H. Bau, C. Li, M. G. Mauk, R. W. Hart, and X. Qiu, "Moisture-Activated Self-Heating Analysis Device," US20120315638 A12012.

[7] L., Changchun and Mauk, M. Hart, R. and Q., Xianbo and Bau, Haim H., "A self-heating cartridge for molecular diagnostics", Lab on a Chip, vol. 11., 2011

Fractalkine/Cx3CR1 signaling in brown adipose tissue and diet induced obesity

Ágnes Polyák
(Supervisor: Dr. Krisztina J. Kovács)
polyak.agnes@itk.ppke.hu

***Abstract*—Brown adipose tissue (BAT) plays an essential role in metabolic and thermoregulation using fatty acids for heat production via uncoupling of mitochondrial oxydative phosphorylation. Because BAT uses fuels stored in white adipose tissue (WAT), there is an incresing interest on the relationship between BAT and obesity. Both BAT and WAT contains significant population of immune cells, tissue specific macrophages, which have an impact on body metabolic and inflammatory functions.**

Fractalkine (Cx3CL1) is a chemokine that attracts macrophages into several tissues, including adipose tissues. Fractalkine receptor (Cx3CR1) is expressed mainly by macrophages. The aim of this study was to reveal the role of macrophages in brown adipose tissue (BAT) function in obesity. For this purpose we used transgenic mice wherein Cx3CR1 was replaced by gfp reporter gene. +/gfp mice have functioning Cx3CR1 while gfp/gfp mice do not express Cx3CR1. Mice were fed with fat enriched diet (FatED) for 10 weeks.

+/gfp FatED - but not gfp/gfp FatED group gained more weight and had higher relative BAT weight than normal diet (ND) fed mice. Lipid droplets of brown adipose cells were larger in FatED than ND groups. Rt-PCR analysis showed elevated Cx3CR1, MCP-1, GFP, IL-1b, TNFa, Arg1, UCP1, PPARg, NPY and nesfatin mRNA levels in BAT in response to FatED. Among these all (except UCP1 and PPARg) mRNA levels were lower in gfp/gfp FatED mice than +/gfp FatED mice.

These results reveal increased number of macrophages in BAT of obese mice and fractalkine signaling plays a role in the traffic and accumulation of these cells. Similarly with our previous results, mRNA levels of inflammatory citokines tend to be lower in gfp/gfp FatED mice than +/gfp FatED mice. We revealed that BAT is capable to express NPY, which might affect polarization of tissue specific macrophages and decrease lipolysis.

***Index Terms*—obesity, fractalkine, Cx3CR1, brown adipose tissue, inflammation, macrophage, NPY**

I. Introduction

Obesity and type2 diabetes are worldwide epidemics driven by the disruption in energy balance [1] characterized by excessive fat accumulation and low grade inflammation in the white adipose tissue (WAT) and decrease in energy expenditure. Brown adipose tissue (BAT) accounts for a significant proportion of energy expenditure via thermogeneration by uncoupling ATP synthesis. In humans, it has been estimated that activated BAT could contribute as much as 15% of energy expenditure [2], [3]. BAT activity correlates with body weight: metabolically active BAT seems to be particularly low in patients with obesity or type2 diabetes [3], [4].

BAT activation is under sympathetic nervous system (SNS) control, although other humoral and hormonal factors contribute to BAT regulation [3]. SNS activation results in the release of catecholamines (noradrenaline) in BAT and WAT. Noradrenaline - acting via the beta-3 adrenergic receptors - induces lipolysis in white adipocytes, whereas it stimulates the expression of thermogenic genes (e.g. uncoupling protein 1 (Ucp1)) in brown adipocytes [5]. Central delivery of the inflammatory cytokine, interleukin 1, or tumor necrosis factor alpha potently activates BAT thermogenesis. Several tissues (e.g. adipose tissues, liver, etc.) recruit macrophages and other immune cells and express inflammatory markers in obese subjects, which is associated with a shift from an anti-inflammatory to a pro-inflammatory profile [6] contributing to a chronic state of low-grade inflammation [7]. Macrophages are also essential for heat production in brown fat [8]. Exposure to cold temperature rapidly promote alternative activation of adipose tissue macrophages, which secrete catecholamines to induce thermogenic gene expression in brown adipose tissue and lipolysis in white adipose tissue [5].

Fractalkine (Cx3CL1, neurotactin) is a chemokine expressed in endothelial cells, vascular smooth muscle cells, hepatocytes, adipocytes and neurons as a transmembrane protein and involved in recruitment and capturing various leukocytes (monocytes, macrophages, microglia) expressing its cognate receptor, Cx3CR1. Fractalkine can be released from the cell surface by proteolytic cleavage and might act in paracrine and endocrine manner. Fractalkine is an important regulatory factor of microglia activity in the central nervous system mediating neuroinflammation. However its role in metabolic inflammation in general, and in connecting metabolic- and neuroinflammation in particular, remains to be elucidated [9]. The aim of this study was to reveal the role of fractalkine signaling in recruitment and polarization of tissue macrophages in BAT during diet-induced obesity.

II. Materials and methods

A. Animals and diet

Experiments were performed in male Cx3CR1 +/gfp, and Cx3CR1 gfp/gfp mice [10]. The C57Bl/6J strain has been shown to be genetically vulnerable to diet-induced obesity [11]. In these mice, the cx3cr1 gene was replaced by a gfp reporter gene such that heterozygote Cx3CR1+/gfp mice express GFP in cells that retain receptor function, whereas cells in homozygote Cx3CR1gfp/gfp (Cx3CR1/, knockout (KO)) mice are labeled with GFP and also lack functional Cx3CR1. Genotype of the animals has been verified by PCR using combination of three different primers as described by

Á. Polyák, "Fractalkine/Cx3CR1 signaling in brown adipose tissue and diet induced obesity,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 55-58.

Jung et al. [10].
Animals were housed in groups of 4-5/cage at the minimal disease (MD) level of the Medical Gene Technology Unit of the Institute, had free access to food and water and were maintained under controlled conditions: temperature, $21\,^{\circ}\mathrm{C} \pm 1\,^{\circ}\mathrm{C}$; humidity, 65%; light-dark cycle, 12-h light/12-h dark cycle, lights on at 07:00. At 35 days of age, both Cx3CR1 +/gfp (n=8) and Cx3CR1 gfp/gfp (n=10) mice were randomly distributed into two equal groups. The first group, normal diet (ND), received standard chow [VRF1 (P), Special Diets Services (SDS), Witham, Essex, UK calory content: 14,24 KJ/g]. The second group received fat-enriched diet (FatED), by providing a 2:1 mixture of standard chow and lard (Spar Budget, Budapest, Hungary). The calorie content of this mixture is 22,02 KJ/g). All procedures were conducted in accordance with the guidelines set by the European Communities Council Directive (86/609 EEC) and approved by the Institutional Animal Care and Use Committee of the Institute of Experimental Medicine (permit number: 22.1/3347/003/2007).

B. Experimental design

Mice were fed with normal diet (ND) or fat enriched diet (FatED) for 10 weeks, body weight and food consumption was measured weekly. Mice were then decapitated, trunk blood was collected on EDTA, and the plasma stored at $-20\,^{\circ}\mathrm{C}$ until assay. Brain, liver, visceral- and, subcutaneous white adipose tissue pads and interscapular brown adipose tissue were collected, sampled and stored at $-70\,^{\circ}\mathrm{C}$ for RT-PCR, or fixed in 4% buffered paraformaldehyde for histology

C. Histology

Tissues were fixed by immersion in 4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4 (PB) for 3 days. Subsequently, they were stored in 1% paraformaldehyde in 0.1 M PB at $4\,^{\circ}\mathrm{C}$. Tissues were paraffin-embedded, sectioned and stained with hematoxylin-eosin (H&E). Slides were digitalized with Pannoramic Digital Slide Scanner (3DHISTECH Kft., Hungary). Lipid droplet areas of brown adipose cells were counted under 40x magnification with ImageJ software (NIH, USA).

D. Gene expression analysis by quantitative real-time PCR

Total RNA was isolated from brown adipose tissue samples with QIAGEN RNeasyMiniKit (Qiagen, Valencia, CA, USA) according the manufacturers instruction. To eliminate genomic DNA contamination, DNase I (Fermentas) treatment was used. Sample quality control and the quantitative analysis were carried out by NanoDrop (Thermo Scientific). Amplification was not detected in the RT-minus controls. cDNA synthesis was performed with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA). The designed primers (Invitrogen) were used in the real-time PCR reaction with Power SYBR Green PCR master mix (Applied Biosystems, Foster City, CA, USA) on ABI StepOne instrument. The gene expression was analyzed by ABI StepOne2.0program. The amplicon was tested by Melt Curve Analysis on ABI StepOne instrument. Experiments were normalized to Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) expression.

E. Primer design

Primers used for the comparative CT (threshold cycle) experiments were designed by the Primer Express 3.0 program. Primer sequences were the following:
GAPDH:f: TGA CGT GCC GCC TGG AGA AA
r: AGT GTA GCC CAA GAT GCC CTT CAG
IL1a: f: CCA TAA CCC ATG ATC TGG AAG AG
r: GCT TCA TCA GTT TGT ATC TCA AAT CAC
IL1b: f: CTC GTG GTG TCG GAC CCA TAT GA
r: TGA GGC CCA AGG CCA CAG GT
IL6: f: TCC GGA GAG GAG ACT TCA CA
r: TGC AAG TGC ATC ATC GTT GT
TNFa: f: CAG CCG ATG GGT TGT ACC TT
r: GGC AGC CTT GTG CCT TGA
ARG1: f: GTC TGG CAG TTG GAA GCA TCT
r: GCA TCC ACC CAA ATG ACA CA
MCP-1:f: CCA GCA CCA GCA CCA GCC AA
r: TGG ATG CTC CAG CCG GCA AC
FKN: f: CCG CGT TCT TCC ATT TGT GT
r: GGT CAT CTT GTC GCA CAT GATT
GFP: f: GGA CGA CGG CAA CTA CAA GA
r: AAG TCG ATG CCC TTC AGC TC
UCP1: f: GGT CAA GAT CTT CTC AGC CG
r: AGG CAG ACC GCT GTA CAG TT
NPY: f: CAG ATA CTA CTC CGC TCT GCG ACA CTACAT
r: TTC CTT CAT TAA GAG GTC TGA AAT CAG TGT
PPARGv2 f: CTC CTG TTG ACC CAG AGC AT
r: TGG TAA TTT CTT GTG AAG TGC TCA
TH f: TCT CAG AGC AGG ATA CCA AGC A
r: GCA TCC TCG ATG AGA CTC TGC

F. Statistical analysis

The results are shown as means + SEM. Statistical analysis was performed by factorial ANOVA with NewmanKeuls post-hoc test in Statistica 11 (StatSoft Inc.). In all cases $p<0.05$ was considered significant. * $p<0.05$ vs. ND, # $p<0.05$ vs. +/gfp; 2 symbols $p<0.01$, 3 symbols $p<0.001$

III. Results

10 weeks after FatED relative BAT weight was higher in +/gfp FatED mice than in gfp/gfp FatED mice and in animals kept on control diet. Both treatment (diet) and genotype effect was significant ($F(1,14) = 11.50$, $p<0.01$; $F(1,14) = 166.27$, $p<0.05$, respectively), post-hoc analysis showed that only +/gfp FatED group differed from the others (Fig.1A). In FatED mice lipid droplet areas in the BAT shifted to the larger sizes, less droplets were under 15 μm^2 and more over 135 μm^2 ($F(1,14) = 8.62$, $p<0.05$; $F(1,14) = 16.76$, $p<0.01$, respectively), according to post-hoc analysis only in +/gfp FatED group were less lipid droplets under 15 μm^2 (Fig.1C).

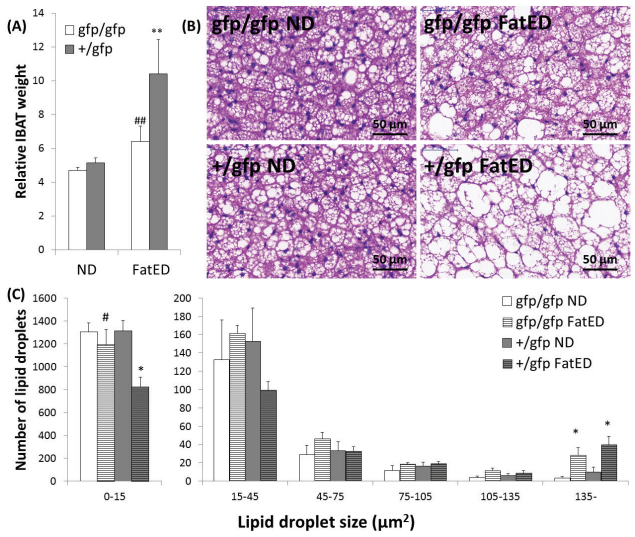


Figure 1. (A) Relative BAT weight (B) Histological sections of BAT, stained with H&E. (C) Frequency distribution of lipid droplets in BAT. * $p<0.05$, ** $p<0.01$ vs. ND; # $p<0.05$, ## $p<0.01$ vs. +/gfp.

Relative mRNA levels of monocyte attracting proteins were elevated in FatED mice (MCP1: genotype effect: $F(1,12) = 10.66$, $p<0.01$; treatment effect: $F(1,12) = 10.55$, $p<0.01$; genotype * treatment: $F(1,12) = 10.11$, $p<0.01$. Fractalkine: genotype effect: $F(1,12) = 10.61$, $p<0.01$; treatment effect: $F(1,12) = 115.17$, $p<0.001$). In case of MCP1 post-hoc analysis showed significant difference between +/gfp FatED and gfp/gfp groups. Elevated GFP mRNA levels in FatED mice refer to larger numbers of macrophages (treatment effect: $F(1,12) = 6.80$, $p<0.05$; genotype effect: $F(1,12) = 9.64$, $p<0.01$). Among the pro-inflammatory citokines Il-1b and TNFa was higher in FatED groups (IL-1b: treatment effect: $F(1,12) = 15.0$, $p<0.01$; genotype effect: $F(1,12) = 5.790$, $p<0.05$; TNFa: treatment effect: $F(1,12) = 9.60$, $p<0.01$; genotype effect: $F(1,12) = 10.33$, $p<0.01$), moreover post-hoc analysis showed that both IL-1b and TNFa level was higher in +/gfp than gfp/gfp mice. The anti-inflammatory Arg-1 expression was also higher in FatED mice (treatment effect: $F(1,12) = 9.79$, $p<0.01$; genotype effect: $F(1,12) = 9.52$, $p<0.01$). The mRNA levels of both UCP1 - the key protein in non-shivering thermogenesis - and PPARg, which increases UCP1 gene expression, was elevated in response to FatED (UCP1: treatment effect: $F(1,11) = 60.18$, $p<0.001$; PPARg treatment effect: $F(1,12) = 98.51$, $p<0.001$, genotype effect: $F(1,12) = 5.81$, $p<0.05$). BAT express elevated levels of NPY in FatED groups (treatment effect: $F(1,11) = 32.43$, $p<0.001$; genotype effect: $F(1,11) = 24.45$, $p<0.001$; treatment * genotype: $F(1,11) = 20.41$, $p<0.001$), with significant difference between +/gfp FatED and gfp/gfp FatED mice according to post-hoc analysis. There was no difference in tyrosine hydroxylase (TH) mRNA expression.

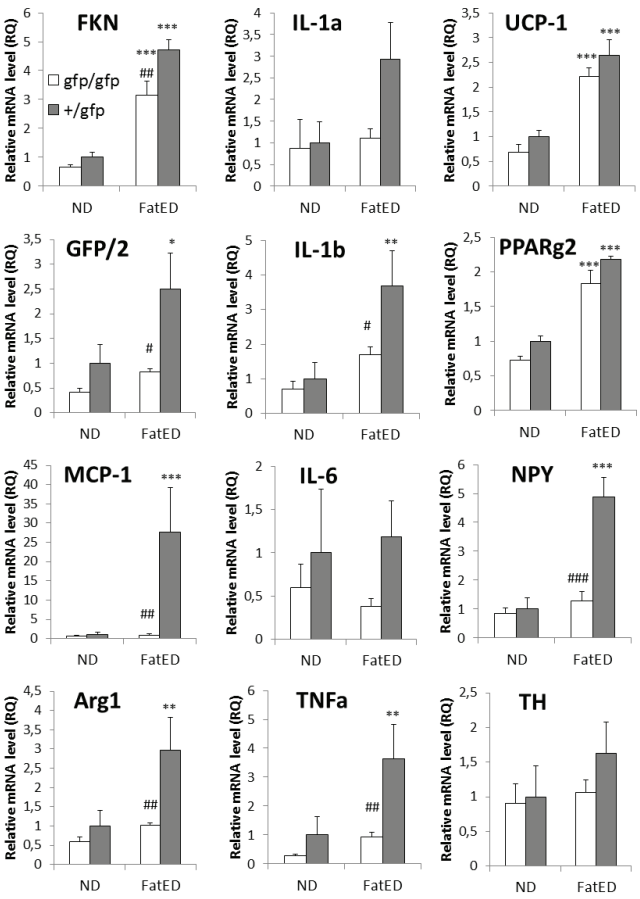


Figure 2. Relative mRNA levels of monocyte-attracting chemokines, GFP, anti- and proinflammatory cytokines, UCP1, PPARg2, NPY and TH. * $p<0.05$ vs. ND, # $p<0.05$ vs. +/gfp, 2 symbols $p<0.01$, 3 symbols $p<0.001$.

IV. Discussion

We showed that FatED induces fractalkine expression in BAT. Furthermore in BAT fractalkine mRNA level is 2-fold higher compared to mRNA levels in WAT, which we measured previously [9] in response to FatED. This means that BAT is an important source of fractalkine in obese individuals. MCP-1 (the main chemokine responsible for the recruitment of monocytes to sites of active inflammation) is also elevated in response to FatED. Accordingly the estimated number of GFP positive cells was higher in FatED groups. Both pro- and anti-inflammatory cytokine expressions were upregulated in FatED fed mice. Lack of fractalkine signaling prevented BAT inflammation in response to FatED. There was no significant difference between ND fed and gfp/gfp FatED fed mice in GFP positive cell accumulation into BAT and pro- and anti-inflammatory cytokine expression. We previously showed that Cx3CR1-deficient (gfp/gfp) mice gain significantly less weight on fat-enriched diet than heterozygote mice [9]. We hypothesized that the difference between body weights can be due to differential activation of BAT. In BAT inner mitochondrial membrane UCP1 generates heat rather than ATP, and for that it needs long-chain fatty acids (LCFA), as it is an LCFA anion/H+ symporter [12]. FatED upregulated both UCP-1 and its regulator PPARg mRNA expression in

both gfp/gfp and +/gfp mice, thus the difference may be elsewhere. Alternatively activated macrophages can secrete catecholamines, which induce lipolysis in WAT and UCP-1 expression in BAT. A defect in alternative macrophage activation was associated with impaired release of free fatty acids [5] but in our experiment Arg1, a marker of alternative macrophage activation, was upregulated in FatED +/gfp mice, suggesting an elevated number of alternatively activated macrophages in BAT. Even so there was no elevation in TH mRNA expression, which catalyses the initial, rate-limiting step in the catecholamine biosynthetic pathway [13]. In contrast to others, who found difference in NPY expression only in EWAT but not in BAT between ND and HFD fed mice after 16 weeks of HFD [14], we showed that BAT express NPY in response to FatED. The sources of NPY in adipose tissue are macrophages and NPY has autocrine, paracrine and systemic effects. HFD, obesity and inflammation increases NPY production in ATMs which dampens M1 ATM activity, while on pre-adipocytes and adipocytes NPY contribute to decreased lipolysis and it promotes adipogenesis [14]. We found that gfp/gfp FatED mice express less NPY than +/gfp FatED mice and therefore it may be that in +/gfp mice NPY causes a shift toward fat accumulation. It is known that obese individuals have downregulated adrenoceptor expression [3], which also contribute to decreased lipolysis in obese individuals as catecholamines induce lipolysis [5]. Therefore the difference between gfp/gfp and +/gfp FatED groups may be a result of decreased lipolysis, caused by elevated NPY expression in macrophages, and maybe decreased adrenoceptor expression. In conclusion, our results suggest that fractalkine signaling has a role in macrophage accumulation into BAT. Impaired fractalkine signaling prevents inflammation and macrophage accumulation into BAT. Therefore ATMs do not express high levels of NPY, which could decrease lipolysis and promote fat accumulation, hence fractalkine receptor deficient mice do not gain excess bodyweight. These results make fractalkine Cx3CR1 system a therapeutic target for obesity and metabolic diseases.

References

[1] B. Zafrir, "Brown adipose tissue: research milestones of a potential player in human energy balance and obesity," *Horm Metab Res*, vol. 45, no. 11, pp. 774–85, 2013.

[2] T. Yoneshiro, S. Aita, M. Matsushita, T. Kameya, K. Nakada, Y. Kawai, and M. Saito, "Brown adipose tissue, whole-body energy expenditure, and thermogenesis in healthy adult men," *Obesity (Silver Spring)*, vol. 19, no. 1, pp. 13–6, 2011.

[3] S. H. Lockie, A. Stefanidis, B. J. Oldfield, and D. Perez-Tilve, "Brown adipose tissue thermogenesis in the resistance to and reversal of obesity: A potential new mechanism contributing to the metabolic benefits of proglucagon-derived peptides," *Adipocyte*, vol. 2, no. 4, pp. 196–200, 2013.

[4] A. Bartelt and J. Heeren, "Adipose tissue browning and metabolic health," *Nat Rev Endocrinol*, vol. 10, no. 1, pp. 24–36, 2014.

[5] K. D. Nguyen, Y. Qiu, X. Cui, Y. P. Goh, J. Mwangi, T. David, L. Mukundan, F. Brombacher, R. M. Locksley, and A. Chawla, "Alternatively activated macrophages produce catecholamines to sustain adaptive thermogenesis," *Nature*, vol. 480, no. 7375, pp. 104–8, 2011.

[6] N. Esser, S. Legrand-Poels, J. Piette, A. J. Scheen, and N. Paquot, "Inflammation as a link between obesity, metabolic syndrome and type 2 diabetes," *Diabetes Research and Clinical Practice*, 2014.

[7] A. Taube, R. Schlich, H. Sell, K. Eckardt, and J. Eckel, "Inflammation and metabolic dysfunction: links to cardiovascular diseases," *Am J Physiol Heart Circ Physiol*, vol. 302, no. 11, pp. H2148–65, 2012.

[8] A. J. Whittle and A. Vidal-Puig, "Physiology immune cells fuel the fire," *Nature*, vol. 480, no. 7375, pp. 46–47, 2011.

[9] A. Polyak, S. Ferenczi, A. Denes, Z. Winkler, R. Kriszt, B. Pinter-Kubler, and K. J. Kovacs, "The fractalkine/cx3cr1 system is implicated in the development of metabolic visceral adipose tissue inflammation in obesity," *Brain Behav Immun*, vol. 38, pp. 25–35, 2014.

[10] S. Jung, J. Aliberti, P. Graemmel, M. J. Sunshine, G. W. Kreutzberg, A. Sher, and D. R. Littman, "Analysis of fractalkine receptor cx(3)cr1 function by targeted deletion and green fluorescent protein reporter gene insertion," *Molecular and Cellular Biology*, vol. 20, no. 11, pp. 4106–4114, 2000.

[11] S. Collins, T. L. Martin, R. S. Surwit, and J. Robidoux, "Genetic vulnerability to diet-induced obesity in the c57bl/6j mouse: physiological and molecular characteristics," *Physiol Behav*, vol. 81, no. 2, pp. 243–8, 2004.

[12] A. Fedorenko, P. V. Lishko, and Y. Kirichok, "Mechanism of fatty-acid-dependent ucp1 uncoupling in brown fat mitochondria," *Cell*, vol. 151, no. 2, pp. 400–13, 2012.

[13] Q. Y. Zhou, C. J. Quaife, and R. D. Palmiter, "Targeted disruption of the tyrosine-hydroxylase gene reveals that catecholamines are required for mouse fetal development," *Nature*, vol. 374, no. 6523, pp. 640–643, 1995.

[14] K. Singer, D. L. Morris, K. E. Oatmen, T. Wang, J. DelProposto, T. Mergian, K. W. Cho, and C. N. Lumeng, "Neuropeptide y is produced by adipose tissue macrophages and regulates obesity-induced inflammation," *PLoS One*, vol. 8, no. 3, p. e57929, 2013.

Neural cell response to nanostructured biosensor surfaces

Zsófia Bérces
(Supervisors: Anita Pongrácz PhD, Kristóf Iván PhD)
berces.zsofia@itk.ppke.hu

***Abstract*—Surface quality of implantable biosensors for recording electrical signals in the central nervous system is of great importance with respect to long-term use in chronic animal experiments. In our work the interaction of cell cultures and nanotextured surfaces was investigated as a model of implanted device surface and living tissue interaction. We have developed a robust, maskless nanostructuring method, which can be integrated into our neural biosensor fabrication process. Morphology of the fabricated nanograss has been characterised with and without adhesion protein coatings using scanning electron microscopy. The nanopillars are between 520-800 nm in height and their density is 18-70/μm^2. The specific surface area is 30 times larger compared to the reference. The nanostructured and platinized surfaces will be systematically characterised by electrochemical impedance spectroscopy, cyclic voltammetry and contact angle measurements and the chips will be tested in viability and adhesion assays using neural cell cultures.**

***Index Terms*—neural implant; nanostructuring; biocompatibility**

I. Introduction

Cells in their natural environment interact with nanoscale structures like the extracellular matrix and its proteins. Based on biomimetic consideration, creating nanopatterned implant surfaces promises better cell adhesion and therefore better biocompatibility. In case of fibroblasts and osteoblasts it's already demonstrated that surfaces with specific surface roughness parameters show better cell adhesive properties [2]. On the other hand, there are only a few results in the case of neural and glial cells. In 1997, Turner et al. showed better neuron adhesion on nanostructured samples compared to smooth ones using immortalised cell line. In the same article, the authors presented that primer neural and glial cells react the other way around. [3] Later on many groups investigated neural cell adhesion on nanostructured biosurfaces. Most of them used porous silicon [4]–[6] or etched Si surfaces with nanometer scale structures [7]–[11]. In other cases, they used different materials, such as GaP [12] or polymers [13]. These experiments aimed to enhance neural implant efficiency, and make the implants capable of conducting chronic in vivo experiments. Other groups have concluded that surface nanostructuring can be used to examine self-developing neural circuits [9] and it is also capable of neurite directioning [5], [9]. As a result, this substrate preparation approach is suitable for modelling neural circuits.

When cells adhere to artificial or natural surfaces they produce so-called adhesins which are essential for cell-substrate adhesion. In common cell culturing methods additional adhesive proteins are also used. Based on literature data, it seems to be possible to develop such surface that can be used for cell culturing without these adhesive proteins [8], [10], [14]. Omitting the use of adhesive proteins is preferred not only because of the easier culturing method, but the artificially added adhesive protein coating would also create an electrical insulating layer between the cells and the surface, which results in worse signal-to-noise ratio in case of electrode surfaces [16].

By modulating the specific surface area, wetting and nano-pattern regularity of the nanostructured samples, several groups published better neural cell adhesion and viability on nanostructured surfaces compared to the smooth references in the past few years [8], [11], [12], [15]. In 5 days long experiments, neural cells showed surface preference based on the work of Y.W. Fan et al. The cells were observed to migrate to nanostructured parts of the sample [10]. In the case of implants in living tissue, the reaction of the surrounding glial cells is as important as the adhesion of the neurons. The astrocytes and microglias are key factors in neuron metabolism and function. They play an emphatic role as the central nervous system's protector in diseases and injuries. In case of physical injury such as the microelectrode implantation they generate an aggressive neuroprotective reaction called gliosis [17]. As a result, the so-called glial scar remains around the injury and the foreign body which hinders electrical recordings. As a consequence in neural implant surface development it's also important to investigate the neuroprotective reaction of the glial cells to the proposed surface. Since these cells work in a close cooperation with neurons and with each other, it's also important to examine them in tissue-like co-cultures.

My work aims to synthesize bioimplant surfaces with nanoscale patterns using novel combination of micro- and nanomachining techniques. The proposed maskless nanopatterning method can easily be integrated into the fabrication process of neural microelectrodes. The expected results are envisioned to minimize the immune response of the neural tissue to the surface of the implanted microelectrodes and thus enable efficient functionality in long-term experiments.

II. Materials and Methods

A. Sample design and fabrication

My investigations are based on two types of biological measurements. Cell adhesion and viability is tested

Zs. Bérces, "Neural cell response to nanostructured biosensor surfaces,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 59-62.

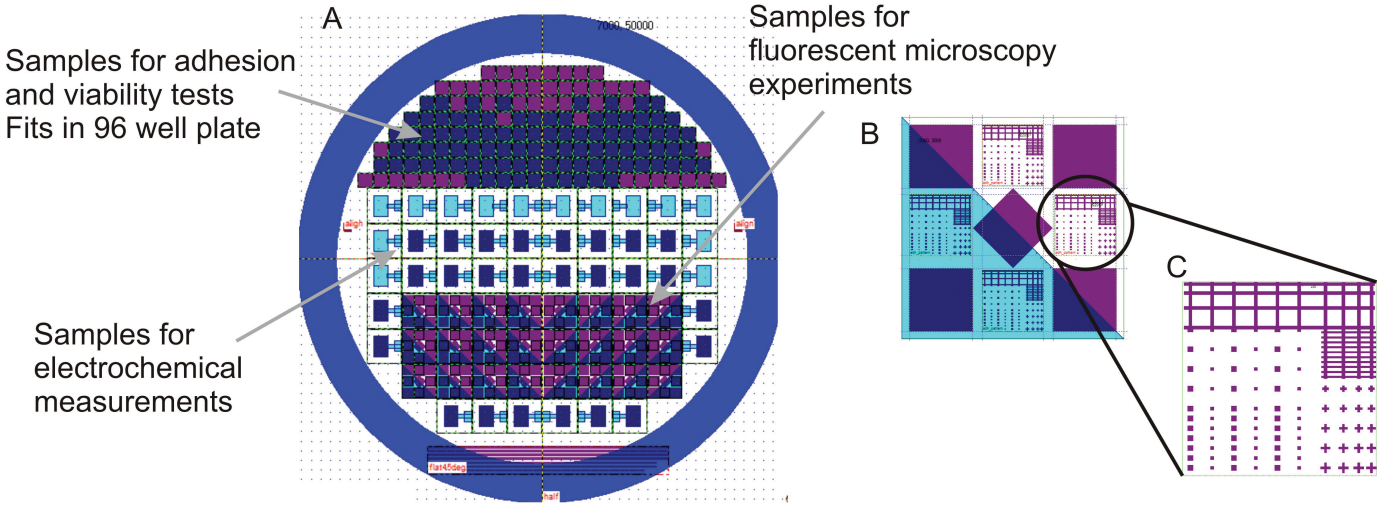


Fig. 1. A) Layout of the pattern to be transferred on the Si wafer. Light blue shows platinum, purple represents black silicon covered areas without platinum coating and dark blue areas represent platinum coated black silicon. The samples for different experiments can be seen. B) Sample for fluorescent measurements. C) Part of the sample shown on B). Different microstructures are designed based on literature or inspired by our microelectrode (contact site sizes and spacing).

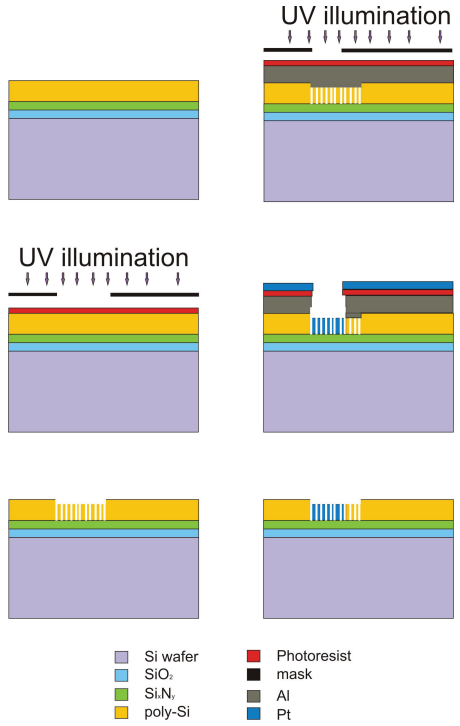


Fig. 2. Schematic process flow of the fabrication method

by methylene blue and MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide) assays in 96 well plates. Each experiment requires different uniformly nanostructured squares of 4.2 mm x 4.2 mm (chip), and smooth samples as reference. To measure the preferential behaviour of cells on nanostructured and smooth reference surfaces, different chips are designed. Applied patterns are in the range of the contact site of a microelectrode array usually used for neural recording. Chips carrying cultured fluorescent cells will be tested by fluorescent microscopy techniques. For electrochemical measurements, several identical test samples were designed. Figure 1 on the left side shows the micro pattern of the Si wafer (A). Different type of chips are marked with arrows. The right side of Figure 1 (B,C) shows the map of the chip for the fluorescent microscopy measurements.

During the fabrication process first, 500 nm thick thermal oxide is grown on a 4" (100) oriented Si wafer. Then 1000 nm poly-Si is deposited in a Tempress LPCVD equipment. The micropatterning of the black-Si is performed by photolitography, and microstructured in poly-etchant. Nanopattern formation is carried out by deep reactive ion etching (DRIE) at cryogenic temperature. After black-Si formation platinum deposition is performed and followed by a second patterning step using lift-off process. Figure 2 shows the schematic process flow of the fabrication method.

The advantages of the process are that the nanostructuring is maskless and it can fully be integrated into an implantable Si microelectrode fabrication process.

B. Surface characterisation

The effect of the different etching parameters and platinum layer thicknesses is investigated using SEM. From the images the pillar height and density parameters of the fabricated samples are also extracted.

1) Electrochemical measurements: The specific surface area of the nanostructured and platinised samples were recovered from the cyclic voltammetry measurements. Electrochemical impedance spectroscopy was carried out to show the impedance reduction, which also refers to surface area enhancement.

C. Adhesion enhancing protein coatings

In common cellular assays adhesion proteins are used. Before cell culturing on our chips we investigated the effect of these proteins on our samples. This way, misinterpretation of in vitro results due to possible planarization effects is minimized. The impedance of the protein coated samples was measured and results were compared with the results obtained using the protein free ones. Cell adhesion proteins coating was carried out in the Research Institute for Experimental Medicine, HAS (MTA KOKI). The proteins were PLL (Poly –L-lysine), which is an artificial protein for neural cell culturing, AK-c(RGDfC), which is a synthetic adhesive polypeptide [18] especially for neural progenitors, and bovine serum as the most common cell culturing medium.

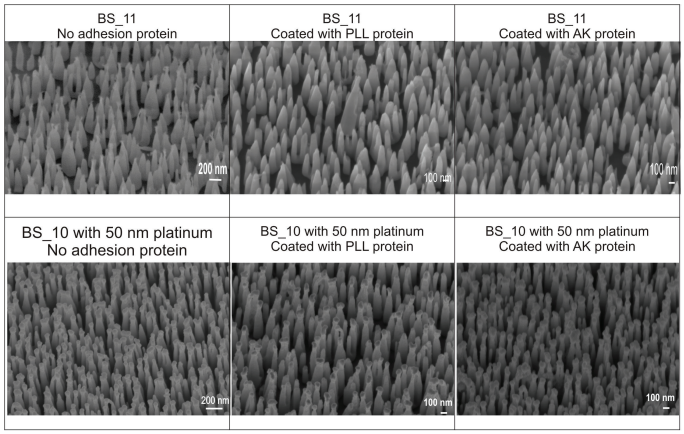


Fig. 3. Representative SEM images of protein coated nanostructures and non-coated reference samples. In the case of PLL coatings pillars became bulkier, in the case of AK only a thin layer can be seen.

III. Results

The morphological parameters of the samples (pillar density, pillar height) were derived from scanning electron micrographs. The nanopillars are between 520-800 nm in height, and their density is 18-70/μm^2 depending on the fabrication parameters of the DRIE process. Samples with surface adhesive proteins were examined. The AK-c(RGDfC)protein is supposed to cover the pillars in a monomolecular layer while PLL covers the surface in a thicker layer. Representative SEM pictures can be seen on Figure 3.

Impedance yield in case of both uncoated and protein coated samples was also measured. In the case of nanostructured samples a 30 times larger specific surface area can be calculated based on the cyclic voltammetry measurements (Figure 4 A). The expected impedance reduction can be seen on the results of the EIS measurements. (See Figure 4 B).

The comparison between EIS curves measured on non-coated and protein coated samples can be seen on Figure 5). While the AK monolayer coating shows no change in impedance, a significant impedance increasing can be seen in the case of PLL and serum.

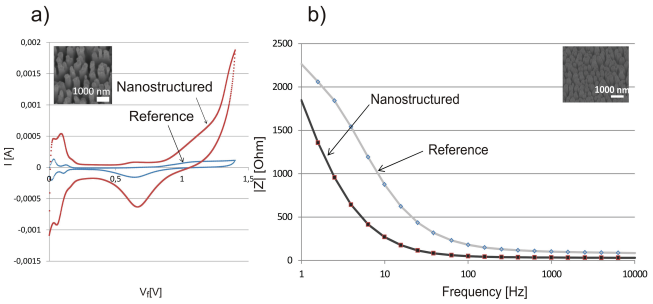


Fig. 4. Representative curves of electrochemical measurements a) Cyclic voltammetry, b) Electrochemical impedance spectroscopy. Specific surface area growth and impedance reduction can be seen.

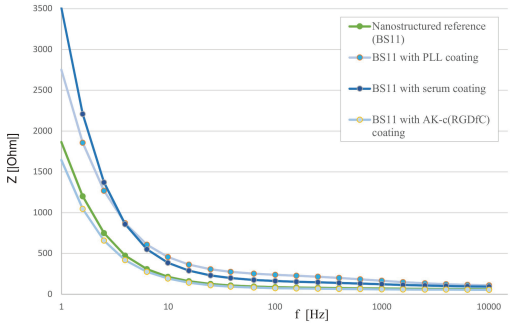


Fig. 5. AK monolayer coating shows no significant difference in impedance. A significant impedance increasing can be seen in the case of PLL and serum

IV. Conclusion and Future Plan

Based on our preliminary studies, the nanopattern morphology of our chips are tuneable. The surface area growth is significant compared to the Pt surface currently used as electrode contact site. The specific surface area growth can be measured by impedance reduction which is an additional advantage. The adhesion proteins do not planarise nanopatterns, however, impedance growth is detected as the samples have been coated. Samples will be investigated with and without these proteins in cellular assays. In the future, designed samples with 8 different parameters will be fabricated: 4 different b-Si nanopattern each with two type of Pt thickness. The samples will be tested in methylene blue and MTT assays using neural progenitor cell line in cooperation with MTA KOKI. Fluorescent cells will also be seeded on samples to examine cells surface selectivity by fluorescent microscope and SEM. Electrochemical measurements will be executed using the corresponding chips. Since the surface wetting seems to be a key factor in cell adhesion [19], mainly with respect to the precise electrochemical measurements, the wetting property of the samples will be characterised by contact angle measurements. Based on quantitative results of the cellular assays and the electrochemical measurements of the same type of samples, the surface with the best electrical properties and biocompatibility will be selected. The chosen nanopatterns will then be integrated into our Si microprobe as a seed layer of the electrical contact sites and will be tested in-vivo.

ACKNOWLEDGMENT

I would like to thank to my supervisors dr. Anita Pongrácz and dr. Kristóf Iván for their help and support. I would especially like to thank dr. Zoltán Fekete for all his work in black-Si method from fabrication to measurements and evaluations. Thanks are also due to Ágoston Horváth for his work in the electrochemical measurements. I also would like to thank dr. Emília Madarász and Attila Jády from MTA KOKI for their cooperation.

REFERENCES

[1] J.S. Perlmutter, J.W. Mink, 'Deep Brain Stimulation', Annual Reviewof Neuroscience 29 p. 229-57, 2006.

[2] Dalby, M.J.; Pasqui, D.; Affrossman, S.: 'Cell response to nano-islands produced by polymer demixing: a brief review', IEE Proceedings - Nanobiotechnology, 151, (2), p. 53-61, 2004.

[3] Turner, S. and Kam, L. and Isaacson, M. and Craighead, H. G. and Shain, W. and Turner, J.,'Cell attachment on silicon nanostructures' Journal of Vacuum Science and Technology B, 15, 2848-2854, 1997.

[4] Moxon, K.A.; Kalkhoran, Nader M.; Markert, M.; Sambito, M.A.; McKenzie, J.L.; Webster, J.T., 'Nanostructured surface modification of ceramic-based microelectrodes to enhance biocompatibility for a direct brain-machine interface', Biomedical Engineering, IEEE Transactions on , vol.51, no.6, pp.881,889, June 2004

[5] Fredrik Johansson, Martin Kanje, Cecilia Eriksson andLars Wallman,'Guidance of neurons on porous patterned silicon: is pore size important? ', physica status solidi (c) Volume 2, Issue 9, pages 3258–3262, June 2005

[6] Andrei V. Sapelkin, Susan C. Bayliss, Bayram Unal, Amanda Charalambou, 'Interaction of B50 rat hippocampal cells with stain-etched porous silicon', Biomaterials, Volume 27, Issue 6, Pages 842-846, February 2006.

[7] J. Ma, F. Z. Cui, B. F. Liu, Q. Y. Xu, 'Atomic force and confocal microscopy for the study of cortical cells cultured on silicon wafers', Journal of Materials Science: Materials in Medicine, Volume 18, Issue 5, pp 851-856, May 2007.

[8] Saida P. Khan, Gregory G. Auner, Golam M. Newaz, Influence of nanoscale surface roughness on neural cell attachment on silicon, Nanomedicine: Nanotechnology, Biology and Medicine, Volume 1, Issue 2, Pages 125-129, June 2005.

[9] N M Dowell-Mesfin, M-A Abdul-Karim, A M P Turner, S Schanz, H G Craighead, B Roysam, J N Turner and W Shain, Topographically modified surfaces affect orientation and growth of hippocampal neurons, J Neural Eng. ;1(2):78-90., Jun. 2004.

[10] Y.W. Fan, F.Z. Cui, S.P. Hou, Q.Y. Xu, L.N. Chen, I.-S. Lee, Culture of neural cells on silicon wafers with nano-scale surface topography, Journal of Neuroscience Methods, Volume 120, Issue 1, Pages 17-23, 15 October 2002.

[11] A. M. P. Turner, N. Dowell, S. W. P. Turner, L. Kam, M. Isaacson, J. N. Turner, H. G. Craighead and W. Shain, 'Attachment of astroglial cells to microfabricated pillar arrays of different geometries', Journal of Biomedical Materials Research Volume 51, Issue 3, pages 430–441, 5 September 2000

[12] Gaëlle Piret, Maria-Thereza Perez, Christelle N. Prinz, 'Neurite outgrowth and synaptophysin expression of postnatal CNS neurons on GaP nanowire arrays in long-term retinal cell culture', Biomaterials, Volume 34, Issue 4, Pages 875-887,January 2013.

[13] Samuel Bechara, Lucas Wadman, Ketul C. Popat, 'Electroconductive polymeric nanowire templates facilitates in vitro C17.2 neural stem cell line adhesion, proliferation and differentiation', *Acta Biomaterialia*, Volume 7, Issue 7, Pages 2892-2901, July 2011.

[14] S.C Bayliss, L.D Buckberry, I Fletcher, M.J Tobin, The culture of neurons on silicon, Sensors and Actuators A: Physical, Volume 74, Issues 1–3, Pages 139-142, 20 April 1999.

[15] YW Fan, FZ Cui, LN Chen, Y Zhai, QY Xu, IS Lee, Adhesion of neural cells on silicon wafer with nano-topographic surface, Applied surface science 187 (3), 313-318, 2002.

[16] YW Fan, FZ Cui, LN Chen, Y Zhai, QY Xu, Improvement of neural cell adherence to silicon surface by hydroxyl ion implantation, Surface and Coatings Technology 131 (1), 355-359, 2000.

[17] Alexei Verkhratsky (Editor), Arthur Morgan Butt (Editor), Glial Physiology and Pathophysiology, Wiley-Blackwell, ISBN: 978-0-470-97853-5, March 2013.

[18] Károly Markó, Tímea Kőhidi, Nóra Hádinger, Márta Jelitai, Gábor Mező,Emília Madarász, 'Isolation of Radial Glia-Like Neural Stem Cells from Fetal and Adult Mouse Forebrain via Selective Adhesion to a Novel Adhesive Peptide-Conjugate', PLoS ONE 6(12): e28538. , 2011.

[19] A. Ranella, M. Barberoglou, S. Bakogianni, C. Fotakis, E. Stratakis, Tuning cell adhesion by controlling the roughness and wettability of 3D micro/nano silicon structures, Acta Biomaterialia 6, 2711–2720, 2010.

Diagnostic methods for microfluidics

Eszter Tóth
(Supervisor: Kristóf Iván Ph.D, Péter Fürjes Ph.D)
totes@digitus.itk.ppke.hu

***Abstract*—Portable diagnostic devices have received a high attention in the medical field. Downscaling of the diagnostic procedures requires the development of various fluid manipulation methods in micrometer-scale channels. With new applications of microfluidic devices, a wide range of reactions have become possible that could not be carried out on macroscale.**

A fast emerging field of microfluidics is the development of Lab-On-a-Chip tools and Micro Total Analysis Systems (μTAS) for cheap and mass producable diagnostic devices.

Present work provides a brief overview of clinical diagnostic methods for microfluidic platforms. Recent developments and actual challenges in microfluidics-based diagnostic research are described and examples are given for state-of-the-art medical and biotechnological applications.

***Keywords*-microfluidics; Lab-On-a-Chip; μTAS; diagnostics**

I. Introduction

Microfluidics has offered a new range of applications in medical diagnostics with Lab-On-a-Chip (LOC) technology. Besides medical laboratories with high needs of human resources, costs and infrastructure, LOC devices provide a cheap, easy-to-use and mass producable way of diagnostics. Microfluidics aims to provide low-cost Point-of-Care devices for continous health monitoring and screening at the the bedside. Micro Total Analysis Systems (μTAS) require low sample volumes that largely simplifies the sampling and examination procedures.

LOC technology requires downscaling of traditional diagnostic methods that holds challenges in both the design and fabrication of the devices. New physical background has to be developed to solve basic tasks in diagnostics like fluid mixing, cell detection and separation as well as flow and reaction control. On the other hand, a new range of chemical reactions and physical phenomena have become possible on microscale. Finding new ways to solve the challenges of microfluidics and utilize the new possibilities for better and smaller diagnostic devices is an active field of research. In this work, a brief overview of recent developments is provided with examples of promising principles and devices of microscale medical diagnostics.

II. Diagnostic principles

A. Cytometry

1) Flow Cytometer: Flow cytometer is a laser-based technology for cell counting, cell sorting, biomarker detection and protein engineering. It supends the cells of interest in a stream. The laser beam of the sensor hits the cells. The forward scatter is proportional to the cell diameter, the side scatter is proportional to the cell granularity. [1]

2) Fluorescence-activated cell sorting FACS: (Fig. 1) Fluorescent-activated cell sorting (FACS) is a specialized flow cytometry method (Fig. 1). The different cells are labeled by specific fluorescent labels. The sensor observes the fluorescent charecter of the cell, then the stream of cells is broke into individual droplets by a vibrating mechanism. A charge is then used to trap the droplets. [2]

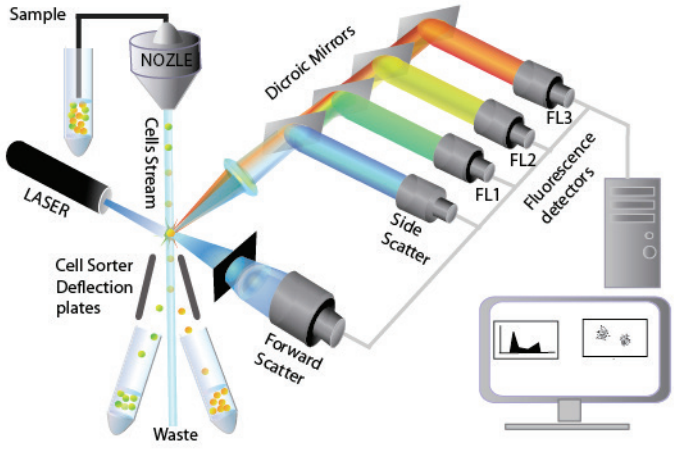


Fig. 1. Fluorescent activated cell sorting is a specialized flow cytometry method. The different cells are labeled and captured at the end of the channel. [3]

B. Immunoassays

1) ELISA: Enzyme-linked immunosorbent assay (ELISA) is one of the most comon immunological methods used for medical diagnostic and research applications to identify and detect proteins based on their binding to immobilized antibodies. In Direct ELISA the plate is coated with the antigenes from the sample, conjugate antibodies are used for amplification and detection. Sandwich ELISA (Fig. 2) coats the plate with antibodies specific for the test antigene. Target protein is captured by the antibody, a detection antibody is then added. The formed complex converts the substrate to a detectable form.

2) MELISA: Memory Lymphocyte Immunostimulation Assay (MELISA) detects Type-IV allergy to environmental toxins, molds, metals and chemicals from blood sample. The test requires live memory lymphocytes which are isolated and incubated. The blood sample is divided into three portion. The first portion remains unexposed to allergenes serving as a negative control. The second part is exposed to an universal allergen, the thirs is exposed to the suspected allergen. The

E. Tóth, "Diagnostic methods for microfluidics,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 63-66.

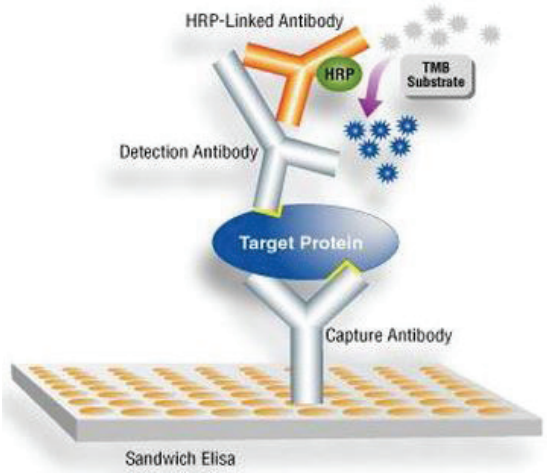


Fig. 2. Sandwich ELISA method. The target protein is captured by the antibody immobilized on the plate. Detection and HPR-Linked antibody is then added and binded to the protein. A substrate is added and is converted by the enzyme to a detectable form. [4]

level of lymphocite reactivity to the allergene is the Stimulation Index (SI). [5]

C. DNA amplification

1) PCR: Polymerase chain reaction (PCR) (Fig. 3) amplifies the DNA of interest cycle to cycle. The sample DNA is denaturated by high temperature so the specific primers can access the single-stranded DNA. The device is then cooled down to enable primer extension by DNA polymerase. [6]

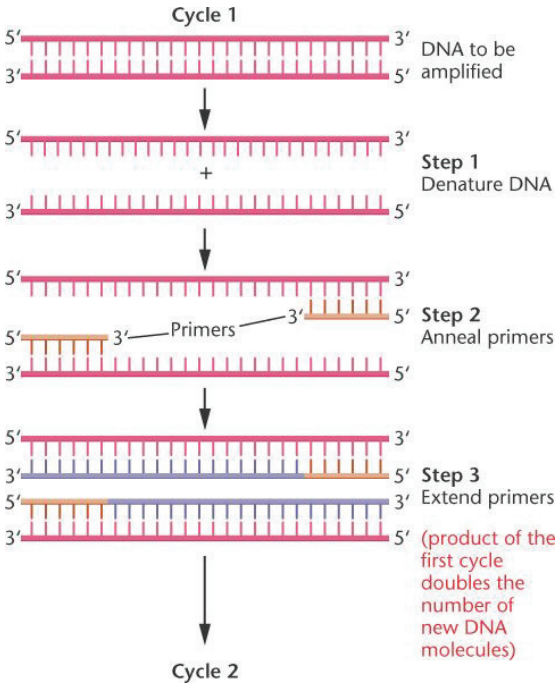


Fig. 3. Polymerase chain reaction (PCR) method. The DNA of the sample is denaturated by high temperature. The primers then annealed to the DNA strands. When the sample is cooled the polimerase extends the primers to a complementary DNA strand. [7]

2) RT-PCR: Reverse transcription polymerase chain reaction (RT-PCR) is commonly used in biology to detect mRNA expression levels. The reverse transkriptase creates a complementary DNA strand from the mRNA of interest. The chain reaction then amplifies the DNA. [8]

3) LAMP PCR: Loop-mediated isothermal amplification (LAMP) is a specialized nucleic acid amplification. The reaction is carried out at a single temperature making it more applicable for microfluidic aplications.

D. Spectroscopy

1) Surface-enhanced Raman Spectroscopy (SERS): Surface-enhanced Raman spectroscopy is a surface-sensitive tool that absorbes moleculed on metal surfaces enhancing Raman Scattering. The sensitivity of the principle makes it applicable to detect low-abundance biomolecules such as proteins in body fluids. [9]

2) Quartz Crystal Microbalance (QCM): Quartz Crystal Microbalance uses a quartz crystal resonator to measure mass. The change in resonation frequency of the crystal is proportional to the mass change. QCM can be used in vacuum, gas phase and in liquid as well. It is highly effective at determining the affinity of molecules to the functionalized crystal surface. It can also measure only a few molecules. [10]

III. Implemented Microfluidic Diagnostic Devices

Several new microfluidics-based diagnostic device has been reported in the past few years. In this section we review microfluidic diagnostic prototype devices by the analyte.

A. Biomarker detection

Cancer Biomarker Detection

1) T-cell lymphoma detection: The system extracts the DNA from whole blood, amplifies the target sequences of the T-cell receptor gene and detects a signature consistent with monoclonality by electrophoresis. [11]

2) On-chip NMR-aided detection: Lee et al. reported a miniaturized diagnostic magnetic resonance (DMR) system for multiplexed marker analysis. The sensor strategy is based on a self-amplifying proximity assay using magnetic nanoparticles. Binding of nanoparticles to their molecular target leads to the formations of nanoclusters decreasing bulk spin-spin relaxation time (T_2) (Fig. 4). [12]

3) Mass-based method: Zhang et al. reported a method similar to quartz crystal microbalance (QCM) to measure small changes in weight using a microfluidic system with integrated surface modified piezoelectric sensor. [13]

Cardiac Biomarker Detection

Cardiac Biomarkers are popular amongst Point-of-care device targets. They give information about recent cardiac events such as acute myocardial infarction (cardiac troponin I (cTnI), Creatine kinase), inflamatory processes (C Reactive Protein (CRP)) and heart failure (B-type natriuretic peptide (BNP), D-dymer).

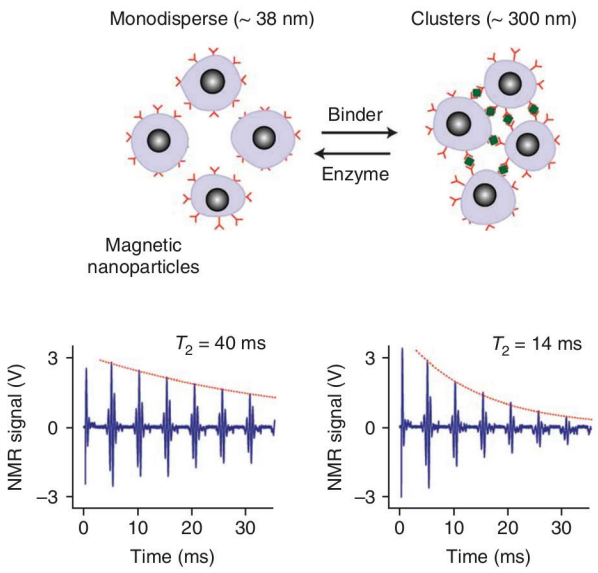


Fig. 4. Principle of proximity assay using magnetic particles. When monodisperse magnetic nanoparticles cluster upon binding to targets, the self-assembled clusters become more efficient at dephasing nuclear spins of many surrounding water protons, leading to a decrease in spin-spin relaxation time T_2 [12].

4) MEMS-based device for multiple cardiac marker detection: The device detects four different cardiac markers (vis. myoglobin, CRP, cTnI and BNP by Au nanoparticle-based fluorescence detection. [14]

5) Chemiluminescence-based immunoassay system: An ELISA-on-a-chip (EOC) based microfluidic system was created by Cho et al. for detection of CRP including an on-board detection module. [15]

6) Surface Acoustic Wave-based platform: Mitsakakis and Gizeli have developed an integrated microfluidic surface acoustic wave (SAW) platform enabling multianalyte detection without conatmination. The observed phase-change is proportional to the concentration. The devices was able to detect creatine kinase, CRP and D-dimer. [16]

B. Viral Detection

HIV detection

1) POCKET for quantifying anti-HIV-1 antibodies: Portable and cost-effective (POCKET) diagnostic device was made by Sie et al. for quantifying anti-HIV-1 antibodies in the sera of HIV-1 infected patients. The application however could not make a correlation of the output data with HIV disease states. [17]

2) RT-PCR-based POC diagnostic chip: Lee et al. developed a reverse transcription polymerase chain reaction (RT-PCR) based POC diagnostic chip for HIV. They used HIV markers p24 and gp120 for diagnostics. [18]

3) Two-stage microfluidics CD4+ T-cell capturing device: Cheng et al. reported a POC microfluidic CD4+ T-cell counting device. The device capture CD4+ cell after depletion of monocytes from whole blood. The microfluidic device is disposable and cheap, however the method requires expensive optical microscopes to count the captured cells. [19]

4) Lensless portable CCD-based microfluidic platform: The captured CD4+ T-lymphocytes are detected by a charge coupled device (CCD) sensor using lensless shadow imaging techniques and counted using automatic cell counting software in a few seconds. [20]

5) Non optical method of counting CD4+ T-cells: A non optical approach is cell counting by measuring changes in conductivity of the sorrounding medium due to ions released from the surface-immobilized cells within a microfluidic channel. [21]

Other Viral Detection

6) Detection of Dengue virus: On-chip surface enhanced Raman spectroskopy (SERS)-based biomolecular device was made by Huh et al. Fluid actuation is governed by electrokinetic forces. The device had a 30pM detection limit. [22]

7) Hepatitis B virus detection: Lee at al. reported a Lab-on-a-disc centrifugal microfluidics-based portable ELISA system for the detection of Hepatitis B virus. The limit of detection for the antigene was 0.51 ng/mL. [23]

8) Influenza virus detection: Monoclonal antibody (mAb)-conjugated immunomagnetic beads are introduced and mixed with the analyte. Virus-bound complexes are fluorescently labeled. External optical device is needed to analyze the intensity. The limit of detection is better comparing to conventional flow cytometry systems. [24]

9) H1N1 virus detection with MIMED: Ferguson et al. presented the Magnetic Integrated Microfluidic Electrochemical Detector (MIMED) for detection of H1N1 influenza virus from throat swab samples. They combined immunomagnetic target caption, concentration and purification, amplification by reverse-transcriptase polymerase chain reaction (RT-PCR) and sequence-specific electrochemical DNA detection. The detection limit was 4 order of magnitude below the classical clinical method. The device offers a general approach for multi-target diagnostics as well. [25]

C. Bacterial Detection

Optical detection methods

Optical sensing methods are often preferred for their high selectivity and sensitivity. Several microfluidic applications have been developed for bacteria using optical means.

1) Detection of Escherichia coli: Xiang et al. detected Esherichia coli using laser-optical fiber fluorescence detection. The detection limit of the system was 0.3 ng/μL thus an order of magnitude lower than that achieved for conventional fluorescence microscope. [26]

2) Bacterial antibody detection: Gao et al. captured labeled bacterial antibodies by pre-patterned antigens on the bottom wall of the microchannels in the electrokinetically controlled microfluidic system. E. coli and Helicobacter pylori antibodies could be simultaneously detected from human blood serum. [27]

Electrical and electrochemical methods

The main advantage of the electrical and electrochemical method is the ease of fabricating microelectrodes in the microchannel by litograpy and the absence of labeling steps. [28]

3) E.coli detection by impedance masurement: Boehm et al. reported a microfluidic sensor based on impedance measurement. Bacteria were recognized by antibodies and selectively immobilized on the functionalized glass surface increasing the measured impedance within the chamber. [29]

4) Detection of urinary pathogens by electrochemical sensor: A microfabricated electochemical sensor consisting of an array of 16 three-layered gold electrodes was able to detect relevant bacterial urinary pathogenes. The library contained Escherichia coli, Proteus mirabilis, Pseudomonas aeruginosa, Enterocococcus spp., and the Klebsiella-Enterobacter group. [30]

IV. Conclusion

Several diagnostic method in the medical and biotechnology field are applicable and adaptable on microscale bringing together the diagnostic power with the advantages of microfluidics such as low sample and reagents volume and small portable sizes.

The most recent microfluidic diagnostic devices are designed for a specific disease or biomarker detection. Although a large interest is shown towards these focused devices, further research aims to develop multifunctional devices to screen multiple diseases thus providing complete Lab-On-a-Chip functionality.

References

[1] H. M. Shapiro, *Practical Flow Cytometry*.John Wiley & Sons, 2005.

[2] W. A. Bonner, H. R. Hulett, R. G. Sweet, and L. A. Herzenberg, "Fluorescence Activated Cell Sorting," *Review of Scientific Instruments*, vol. 43, no. 3, pp. 404–409, Mar. 1972.

[3] FACS Flow Cytometry - ENEA UT-BIORAD Technical Unit of Radiation Biology and Human Health (online), http://utbiorad.casaccia.enea.it/facs-flow-cytometry.html (Accessed: May 2014)

[4] ELISA: Direct and Indirect "Sandwich" ELISA General Protocol, EIAab, Jul. 2012. , Available: http://www.eiaab.com/info/detail/456

[5] E. Valentine-Thon, K. Müller, G. Guzzi, S. Kreisel, P. Ohnsorge, and M. Sandkamp, "LTT-MELISA® is clinically relevant for detecting and monitoring metal sensitivity," *Neuro endocrinology letters*, vol. 27, pp. 17–24, 2006.

[6] H. A. Erlich, "Polymerase chain reaction," *J Clin Immunol*, vol. 9, no. 6, pp. 437–447, Nov. 1989.

[7] PCR (online), http://bio3400.nicerweb.com/Locked/media/ch19/19-17-PCR.jpg (Accessed: May 2014)

[8] W. Liu and D. A. Saint, "A New Quantitative Method of Real Time Reverse Transcription Polymerase Chain Reaction Assay Based on Simulation of Polymerase Chain Reaction Kinetics," *Analytical Biochemistry*, vol. 302, no. 1, pp. 52–59, Mar. 2002.

[9] S. Schlücker, *Surface Enhanced Raman Spectroscopy: Analytical, Biophysical and Life Science Applications*. John Wiley & Sons, 2013.

[10] S. Bruckenstein and M. Shay, "Experimental aspects of use of the quartz crystal microbalance in solution," *Electrochimica Acta*, vol. 30, no. 10, pp. 1295–1300, Oct. 1985.

[11] L. A. Legendre, C. J. Morris, J. M. Bienvenue, A. Barron, R. McClure, and J. P. Landers, "Toward a Simplified Microfluidic Device for Ultra-Fast Genetic Analysis with Sample-In/Answer-Out Capability: Application to T-Cell Lymphoma Diagnosis," *Journal of Laboratory Automation*, vol. 13, no. 6, pp. 351–360, Dec. 2008.

[12] Hakho Lee, E. Sun, D. Ham, and R. Weissleder, "Chip–NMR biosensor for detection and molecular analysis of cells," *Nature Medicine*, vol. 14, no. 8, pp. 869–874, Aug. 2008.

[13] K. Zhang, L.-B. Zhao, S.-S. Guo, B.-X. Shi, T.-L. Lam, Y.-C. Leung, Y. Chen, X.-Z. Zhao, H. L. W. Chan, and Y. Wang, "A microfluidic system with surface modified piezoelectric sensor for trapping and detection of cancer cells," *Biosensors and Bioelectronics*, vol. 26, no. 2, pp. 935–939, Oct. 2010.

[14] B. Hong, J. Kai, Y. Ren, J. Han, Z. Zou, C. H. Ahn, and K. A. Kang, "Highly Sensitive Rapid, Reliable, and Automatic Cardiovascular Disease Diagnosis with Nanoparticle Fluorescence Enhancer and Mems," in *Oxygen Transport to Tissue XXIX*, K. A. Kang, D. K. Harrison, and D. F. Bruley, Eds. Springer US, 2008, pp. 265–273.

[15] I.-H. Cho, E.-H. Paek, Y.-K. Kim, J.-H. Kim, and S.-H. Paek, "Chemiluminometric enzyme-linked immunosorbent assays (ELISA)-on-a-chip biosensor based on cross-flow chromatography," *Analytica Chimica Acta*, vol. 632, no. 2, pp. 247–255, Jan. 2009.

[16] K. Mitsakakis and E. Gizeli, "Detection of multiple cardiac markers with an integrated acoustic platform for cardiovascular risk assessment," *Analytica Chimica Acta*, vol. 699, no. 1, pp. 1–5, Aug. 2011.

[17] W. G. Lee, Y.-G. Kim, B. G. Chung, U. Demirci, and A. Khademhosseini, "Nano/Microfluidics for diagnosis of infectious diseases in developing countries," *Advanced Drug Delivery Reviews*, vol. 62, no. 4–5, pp. 449–457, Mar. 2010.

[18] S. H. Lee, S.-W. Kim, J. Y. Kang, and C. H. Ahn, "A polymer lab-on-a-chip for reverse transcription (RT)-PCR based point-of-care clinical diagnostics," *Lab Chip*, vol. 8, no. 12, pp. 2121–2127, Dec. 2008.

[19] X. Cheng, A. Gupta, C. Chen, R. G. Tompkins, W. Rodriguez, and M. Toner, "Enhancing the performance of a point-of-care CD4+ T-cell counting microchip through monocyte depletion for HIV/AIDS diagnostics," *Lab Chip*, vol. 9, no. 10, pp. 1357–1364, May 2009.

[20] S. Moon, H. O. Keles, A. Ozcan, A. Khademhosseini, E. Hæggstrom, D. Kuritzkes, and U. Demirci, "Integrating microfluidics and lensless imaging for point-of-care testing," *Biosensors and Bioelectronics*, vol. 24, no. 11, pp. 3208–3214, Jul. 2009.

[21] X. Cheng, Y. Liu, D. Irimia, U. Demirci, L. Yang, L. Zamir, W. R. Rodríguez, M. Toner, and R. Bashir, "Cell detection and counting through cell lysate impedance spectroscopy in microfluidic devices," *Lab Chip*, vol. 7, no. 6, pp. 746–755, May 2007.

[22] Y. S. Huh, A. J. Chung, B. Cordovez, and D. Erickson, "Enhanced on-chip SERS based biomolecular detection using electrokinetically active microwells," *Lab Chip*, vol. 9, no. 3, pp. 433–439, Feb. 2009.

[23] B. S. Lee, J.-N. Lee, J.-M. Park, J.-G. Lee, S. Kim, Y.-K. Cho, and C. Ko, "A fully automated immunoassay from whole blood on a disc," *Lab Chip*, vol. 9, no. 11, pp. 1548–1555, Jun. 2009.

[24] K.-Y. Lien, L.-Y. Hung, T.-B. Huang, Y.-C. Tsai, H.-Y. Lei, and G.-B. Lee, "Rapid detection of influenza A virus infection utilizing an immunomagnetic bead-based microfluidic system," *Biosensors and Bioelectronics*, vol. 26, no. 9, pp. 3900–3907, Maj. 2011.

[25] B. S. Ferguson, S. F. Buchsbaum, T.-T. Wu, K. Hsieh, Y. Xiao, R. Sun, and H. T. Soh, "Genetic Analysis of H1N1 Influenza Virus from Throat Swab Samples in a Microfluidic System for Point-of-Care Diagnostics," *J Am Chem Soc*, vol. 133, no. 23, pp. 9129–9135, Jun. 2011.

[26] Q. Xiang, G. Hu, Y. Gao, and D. Li, "Miniaturized immunoassay microfluidic system with electrokinetic control," *Biosensors and Bioelectronics*, vol. 21, no. 10, pp. 2006–2009, Apr 2006.

[27] Y. Gao, P. M. Sherman, Y. Sun, and D. Li, "Multiplexed high-throughput electrokinetically-controlled immunoassay for the detection of specific bacterial antibodies in human serum," *Analytica Chimica Acta*, vol. 606, no. 1, pp. 98–107, Jan. 2008.

[28] Jinseok Heo and S. Z. Hua, "An Overview of Recent Strategies in Pathogen Sensing," *Sensors*(14248220), vol. 9, no. 6, pp. 4483–4502, Jun. 2009.

[29] D. A. Boehm, P. A. Gottlieb, and S. Z. Hua, "On-chip microfluidic biosensor for bacterial detection and identification," *Sensors and Actuators B: Chemical*, vol. 126, no. 2, pp. 508–514, Oct. 2007.

[30] J. C. Liao, M. Mastali, V. Gau, M. A. Suchard, A. K. Møller, D. A. Bruckner, J. T. Babbitt, Y. Li, J. Gornbein, E. M. Landaw, E. R. B. McCabe, B. M. Churchill, and D. A. Haake, *"Use of Electrochemical DNA Biosensors for Rapid Molecular Identification of Uropathogens in Clinical Urine Specimens,"* J. Clin. Microbiol., vol. 44, no. 2, pp. 561–570, Feb. 2006.

Program 2

Computer Technology Based on Many-core Processor Chips, Virtual Cellular Computers, Sensory and Motoric Analog Computers

Head: Péter Szolgay

Computing linearly conjugate weakly reversible kinetic structures using graph theory and optimization

Bernadett Ács
(Supervisor: Gábor Szederkényi)
acs.bernadett@itk.ppke.hu

***Abstract*—An algorithm is given in this paper for computing dense weakly reversible linearly conjugate realizations of chemical reaction networks on a fixed set of complexes. The algorithm also decides whether such realization exists or not. The algorithm uses linear programming and graph theory.**

***Keywords*-chemical reaction network, weak reversibility, directed graph representation, optimization**

I. Introduction and basic notions

Chemical reaction networks (CRN) are used not only for modelling biochemical processes, but all kinds of nonlinear qualitative dynamical phenomena, for example describing the operation of mechanical or electrical systems. In this paper a special type of networks will be examined whose dynamics are governed by the mass action law.

Chemical reaction networks are usually described by the following three sets.

1) A set of *species*: $\mathcal{S} = \{X_i | i = 1, \dots, n\}$
2) A set of *complexes*: $\mathcal{C} = \{C_j | j = 1, \dots, m\}$,
   where $C_j = \sum_{i=1}^{n} \alpha_{ji} X_i \quad \forall j = 1, \dots, m$
   and $\alpha_{ji} \in \mathbb{N} \quad \forall j = 1, \dots, m, i = 1, \dots, n$.
   $\mathcal{C}$ is a finite set of the linear combinations of the species with nonnegative integer coefficients, that are called *stoichiometric coefficients*.
3) A set of *reactions*: $\mathcal{R} \subseteq \{(C_i, C_j) | C_i, C_j \in \mathcal{C}\}$,
   generally a real subset of all ordered pairs consisting of complexes. The ordered pair (C_i, C_j) corresponds to the reaction $C_i \to C_j$.

To each reaction in $\mathcal{R}$ there belongs a nonnegative real number k_{ij} called the *reaction rate coefficient*. The reaction (C_i, C_j) takes place only if the coefficient k_{ij} belonging to it is positive.

Assuming mass-action kinetics the following dynamical equations will be used to describe the species' concentrations depending on time:

$$\dot{x} = Y \cdot A_k \cdot \psi(x) \tag{1}$$

where x denotes the concentrations of the species, so it must be in the open positive orthant $\mathbb{R}^n_+$.
Y is the *complex composition matrix*, the element $[Y]_{ij}$ equals the coefficient α_{ji}. The columns of Y corresponds to the complexes, while the rows to the species.

The matrix A_k is a *Kirchhoff-matrix* or *column conservation matrix*, if

$$[A_k]_{ij} = \begin{cases} -\sum_{l=1, l\neq i}^{m} k_{il} & \text{if } i = j \\ k_{ji} & \text{if } i \neq j \end{cases} \tag{2}$$

and the function $\psi : \mathbb{R}^n \to \mathbb{R}^m$ is defined by the following coordinate functions:

$$\psi_j(x) = \prod_{i=1}^{n} x_i^{\alpha_{ji}} \quad j = 1, \dots, m \tag{3}$$

The dynamics of a reaction network are determined only by the matrices Y and A_k, but reaction networks with different structures and species sets can be described by the same differential equations. The reaction networks described by the matrix pairs (Y, A_k) and (Y', A'_k) are *dynamically equivalent*, if

$$\dot{x} = Y \cdot A_k \cdot \psi(x) = Y' \cdot A'_k \cdot \psi'(x) = f(x) \quad \forall x \in \mathbb{R}^n_+ \tag{4}$$

In this case these are *realizations* of the dynamical system determined by $\dot{x} = f(x)$.

If the set of complexes is fixed, the above definition can be extended. The reaction networks given by (Y, A_k) and (Y, A'_k) are *linearly conjugate* if there is a positive diagonal matrix T so that $Y \cdot A_k = T \cdot Y \cdot A'_k$.

II. Graph representation

A reaction network can also be be described by a weighted directed graph $G(V, E)$ called *Feinberg-Horn-Jackson graph*, or shorter *reaction graph*. The sets and properties of the reaction network are represented as follows:

1) the *vertices* correspond to the complexes, $V(G) = \mathcal{C}$,
2) the *directed edges* describe the reactions,
   there is a directed edge from the vertex C_i to C_j if and only if the reaction $C_i \to C_j$ takes place,
3) the *weights* of the edges are the reaction rate coefficients, the weight of the edge representing the reaction $C_i \to C_j$ is k_{ij} for all $(C_i, C_j) \in \mathcal{R}$.

In the reaction graph loops and multiple edges are not allowed.

A reaction network is called *reversible*, if for all reactions $C_i \to C_j$ the reverse reaction $C_j \to C_i$ takes place as well. It

B. Ács, "Computing linearly conjugate weakly reversible kinetic structures using graph theory and optimization,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 69-71.

means that if in the reaction graph there is an edge between two vertices, then there is one in both directions.

The previously mentioned property is rarely satisfied in general, therefore the following weaker one is preferable. It also guarantees the stability of the system. A reaction network is called *weakly reversible* if it holds for all $C_i, C_j \in \mathcal{C}$ that if the complex C_j is reachable from complex C_i after one or more reactions, then C_i is reachable from C_j as well. (It is not obviously true for any reaction network.)

A directed graph is called *strongly connected* if all the vertices are reachable on a directed path from all other vertices. A *strong component* of a directed graph is a maximal set of vertices where all the vertices are reachable from each other. If a strong component contains only one vertex, then it is called *trivial strong component*. The vertex set of every directed graph can be partitioned into strong components, at worst all the vertices are trivial strong components.

Lemma 1. *A reaction network is weakly reversible if and only if there are no edges between the strong components of the reaction graph belonging to it.*

Proof: If there are no edges between the strong components, and vertex v_1 is reachable from vertex v_2, then they must be in the same strong component, so v_2 is reachable from v_1 as well.

Assume that G_1 and G_2 are two disjoint strong components of the reaction graph, $v_1 \in V(G_1)$ and $v_2 \in V(G_2)$ are vertices, and there is a directed edge $\overrightarrow{v_1 v_2}$ connecting them. As the reaction network is weakly reversible, there must be a directed path from v_2 to v_1. But in this case there is a path between the vertices in G_2 and v_1 in both directions. Therefore v_1 must be in G_2, but it is a contradiction, for G_2 is maximal. ■

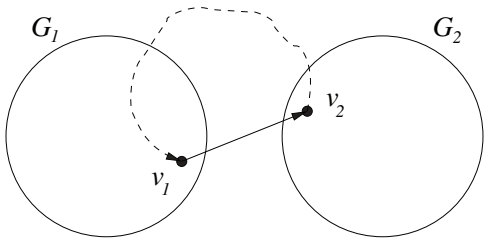


Figure 1. No edges between strong components

In Lemma 1 paths between strong components should not be mentioned, because all the vertices are in some strong component, even the interior points of the paths. Therefore if in the reaction graph of a weakly reversible realization there is a trivial strong component, then it must be an isolated vertex.

A realization of a CRN is called *dense realization* if its reaction graph has maximal number of edges, and it is called *sparse realization* if it has the minimal number of edges.

It was proven by Johnston et al. in [3] that the dense realizations are super-structures amongst dynamically equivalent and linearly conjugate realizations respectively. It means that the reaction graphs of all the possible dynamically equivalent realizations on the same vertex set not considering weights of the edges are subgraphs of the reaction graph of the dense realization. (From now a subgraph of a directed graph will always mean subgraph without edge weights.) The same holds for linearly conjugate realizations, all the linearly conjugate realizations belong to subgraphs of the reaction graph of the dense linearly conjugate realization. It is clear that in both cases the dense realization is unique, because there can not be two different graphs that are subgraphs of each other.

III. Algorithm

The algorithm demonstrated in this section was originally published by Szederkényi et al. in [1]. It computes the dense dynamically equivalent weakly reversible realization for given dynamics on a fixed set of complexes.

This algorithm can be applied in exactly the same way to find linearly conjugate weakly reversible realizations. The only exception is that we have to look for dense linearly conjugate, not dynamically equivalent realizations.

The dynamics of the original system is given by the following equation:

$$\dot{x} = Y \cdot A_k \cdot \psi(x) = M \cdot \psi(x) \tag{5}$$

If a reaction network determined by (Y, A'_k) is linearly conjugate to it, then for some positive definite matrix T it fulfils the following equality as well:

$$M \cdot \psi(x) = T \cdot Y \cdot A'_k \cdot \psi(x) \tag{6}$$

The function $\psi(x)$ is a monomial-type vector mapping, therefore a realization given by the pair (Y, A'_k) is linearly conjugate to the given system if and only if $Y \cdot A'_k = T^{-1} \cdot M$.

Since the complex set is given, Y and $\psi(x)$ are constants, and the coordinates of the matrix A_k and T are the variables. The elements of A_k determine the edges of the reaction graph – if $[A_k]_{ij} > 0$, then $C_j C_i$ is a directed edge. We can add constraints resulting the deletion of certain edges. The idea of this method is that the edges that can not be in the reaction graph of a weakly reversible realization can be recognized.

There are two procedures applied repeatedly during the algorithm:

- The dense linearly conjugate realization of the dynamical system with some subgraph of G as reaction graph can be determined using a polynomial time algorithm created by J. Rudan in [2]. It will be called FindDense(M, Y, G).
- The strongly connected components of the graph can be determined for example by breadth first search, run from all the vertices. It produces all the pairs of vertices that can be reached from each other on a directed path. By checking all the edges one can find those that have endpoints in different strong components. It takes two steps for each edge, so with the BFS it is also a polynomial time method. This procedure will be called FindCrossedges(G) and will return the set of edges between the strong components.

In the algorithm (T, A_k) always represents a realization, $G(T, A_k)$ the reaction graph belonging to it, $E(G)$ the edge set of graph G, and K_n the complete directed graph on n vertices, edges directed in both directions.

procedure STRUCTURE(M, Y)
  (T, A_k) :=FindDense(M, Y, K_n)
  $G := G(T, A_k)$
  while FindCrossedges(G) $\neq \emptyset$ **do**
    $E(G) := E(G) \setminus$ FindCrossedges(G)
    $(T, A_k) :=$ FindDense(M, Y, G)
    $G := G(T, A_k)$
  end while
  if $E(G) = \emptyset$ **then**
    There is no weakly reversible
    linearly conjugate realization.
  else
    (T, A_k) is a weakly reversible
    linearly conjugate realization.
  end if
end procedure

Figure 2. Algorithm for computing weakly reversible linearly conjugate realizations

In each step of the algorithm the actual reaction graph with maximal number of edges is examined. According to Lemma 1 there can not be any edge in a weakly reversible realization with endpoints in different strong components. So if the procedure FindCrossedges($G(T, A_k)$) returns a nonempty set S, then (T, A_k) is not a weakly reversible realization.

The realization (T, A_k) in each step is a dense realization with the restriction of not containing any forbidden edges. This is a linear constraint in the optimization model, so according to Lemma 2 there also must be a super-structure amongst realizations with this property. Since the actual $G(T, A_k)$ is dense under the given restrictions, it is enough to search for realizations with reaction graphs subgraphs of it.

Lemma 2. *For any CRN, there is a realization determining a super-structure amongst all the realizations linearly conjugate to the given CRN and fulfilling a finite set of linear constraints.*

Proof: The realizations are linearly conjugate to the CRN if and only if the following equation holds for them:

$$M = T \cdot Y \cdot A_k \tag{7}$$

Since T is a positive diagonal matrix, T^{-1} exists and it is also a positive diagonal matrix. So the equation can be transformed into a linear relation between the entries of T^{-1} and A_k, with M and Y being constants.

$$0 = Y \cdot A_k - T^{-1} \cdot M \tag{8}$$

Every realization can be represented as a point in $\mathbb{R}^{m(m+1)}$. As the sign of each coordinate is known, all points have to be in the same orthant. All constraints are linear, so the possible solutions are in a polyhedron, and the convex combinations of any two points are in it as well. Let us assume that there is a point P with maximal number of positive coordinates. Since all coordinates in T^{-1} are positive, and all in the main diagonal of A_k are negative, P has to have the maximal possible number of positive off-diagonal elements of A_k. Therefore the reaction graph belonging to P has maximal number of edges.

Let Q represent another realization. If there is a coordinate i so that $P_i = 0$ but $Q_i > 0$ (it must represent an offdiagonal element in A_k), then consider an interior point of the interval $(P, Q) = \{R \in \mathbb{R}^{m(m+1)} | R = c \cdot P + (1 - c) \cdot Q, c \in (0, 1)\}$.

All the coordinates that are positive in P or Q have to be positive in R as well. Therefore R has more positive coordinates than P which is a contradiction.

The points in the polyhedron can have positive coordinates only where P does, so the reaction graphs are subgraphs of the one describing P. ■

IV. Optimization model

Using optimization the following constraints should be fulfilled for a linearly conjugate realization:

$$Y \cdot A_k = T^{-1} \cdot M \tag{9}$$

$$\sum_{i=1}^{m} [A_k]_{ij} = 0 \quad \forall j \in \{1, 2, \dots m\} \tag{10}$$

$$[A_k]_{ij} \geq 0 \text{ if } i \neq j \quad \forall i, j \in \{1, 2, \dots m\} \tag{11}$$

To forbid a set S of reactions the following must hold:

$$[A_k]_{i,j} = 0 \quad (C_i, C_j) \in S \tag{12}$$

For dense realizations the number of positive offdiagonal elements in A_k should be maximized.

V. Conclusion

If there is a weakly reversible realization linearly conjugate (or as special case dynamically equivalent) to the given CRN, the algorithm presented above can find it.

In each step a dense realization for a set of additional constraints is constructed, therefore the result is a dense weakly reversible realization.

This algorithm is a more general version of the algorithm presented in [1], and runs in polynomial time.

Acknowledgement

This project was developed within the PhD program of the Multidisciplinary Doctoral School, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest.

Thanks are due to my supervisor, Gábor Szederkényi for the interesting topic and the fruitful discussions.

References

[1] G. Szederkényi, K. M. Hangos, Z. Tuza: Finding weakly reversible realizations of chemical reaction networks using optimization, MATCH Commun. Math. Comput. Chem. 67: 193-212, 2012

[2] J. Rudan, G. Szederkényi, K. M. Hangos, T. Péni: Polynomial time algorithms to determine weakly reversible realizations of chemical reaction networks, J Math Chem (2014) 52:13861404

[3] M. D. Johnston, D. Siegel, G. Szederkényi: Dynamical equivalence and linear conjugacy of chemical reaction networks: new results and methods, MATCH Commun. Math. Comput. Chem. (2012) Vol. 68, pp. 443-468.

Fast and Efficient Data Reduction Approach for Multi-Camera Light Field Display Telepresence Systems

Vamsi Kiran Adhikarla
(Supervisor: Péter Szolgay)
v.kiran@holografika.com

***Abstract*— Virtual presence is an immensely growing field of study since past few decades with many applications including video conferencing. Although the conception took shape from video telephony during early 19th century, the first working version was commercially made available during 1980 - 1990. However, fully functional telepresence system that stimulates the senses to give an illusion of 3D presence in real time is not available yet. Existing 2D video conferencing systems fail to provide sufficient depth cues and it is not possible to maintain eye contact in a group conferencing scenario. With the advent of light field display technology, it became possible to create an extreme sense of reality. Displays using this technology recreate the light field structure of the scene by shooting the light rays from different positions analogous to real world. Porting this technology to telepresence systems would greatly enhance the user experience. The main problem in using the light field displays arises from the size of the data needed to actuate the display. Light field realization needs multiple views at every time instant which can be captured by incorporating multiple cameras. Nonetheless, depending on the field of view of the display, using the existing networking solutions, it may not be possible to transfer the information over longer distances in real time. In this paper, we propose a method that makes efficient use of available bandwidth by exploiting existing compression schemes and light field display technology. We show that the method is simple to implement and results in high data compression.**

***Keywords*-3D-TV; image processing; 3-D video transmission; light field; HoloVizio; multi-view; telepresence; collaborative virtual environments.**

I. INTRODUCTION

Over the recent years, to cope with the necessity of modern communication world, we have been witnessing a firm increase in the demand of remote collaboration systems. Large high-resolution displays are becoming more and more popular in many fields and introducing these displays in to collaboration space gives a realistic impression of virtual telepresence. Light field displays [1], that are based on multiple projection mechanism are of very high resolution and capable of providing substantial depth cues for 3D perception. These displays provide sufficiently wide field of view and support multiple users simultaneously. In addition, smooth and continuous parallax enhances the ability to follow eye gaze and user gestures. Thus, it is highly advantageous to have such a display in the collaboration space.

Modern day telepresence systems [7, 9] incorporate multiple cameras locally to capture different perspectives. Inevitably they generate a huge payload collectively which needs to be stored, processed, transmitted and reproduced at the remote site. If the remote site is equipped with a light field display, the amount of data flow increases considerably because of the enlarged field of view. Under such conditions, transmission of data in real time is an enormous challenge. Thus, schemes that provide sufficient data reduction for real-time transmission and at the same time are flexible in adapting the bit-rate for transmission are very much needed.

In this paper we propose a simple, yet efficient data reduction approach for transmitting the light field information. The use of classical 2D video compression methods on various camera image streams is an alternative, but may not yield sufficient reduction in many cases. The main contributions of the work include reducing the data in the first step by carefully considering the light field reconstruction mechanism and then rearranging the reduced data to support better compression using existing data compression technologies. Note that the aim of this work is not to investigate a new encoding and decoding algorithms for data compression, rather to manipulate in the image space that leads to best compression.

The rest of the paper is organized as follows: in section 2, we present a short literature overview, section 3 presents the data reduction approach followed by results and conclusions in sections 4 and 5 respectively.

II. RELATED WORKS

In [11], the authors introduced dynamic camera selection process to reduce the number of recording cameras at the acquisition site. The idea here is to track the user position at the remote space then use this information to make a decision on the number and position of recording cameras. They also investigated the usage of selecting dynamic frustum in case of failure in camera selection. Lamboray et al., in [4], introduced a way of dividing the data stream into several parts namely: bulk data, sporadic-event data, and real-time streaming data. They explored and presented various aspects concerning image and geometry based reconstruction systems. They also rely on the positional data of the user in calculating the image update and then use this information for selective transmission. They implemented the system in a controlled

V. K. Adhikarla, "Fast and Efficient Data Reduction Approach for Multi-Camera Light Field Display Telepresence Systems,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 73-77.

network where both the sender and receiver sites are close to each other.

In [5], Lien et al. proposed model-driven data compression. The authors, in [8], present a concise study that suggests a method of data reduction which adapts to the changes in the scene. Concerning data transmission, Hutanu and Nimmagadda, in [2], have shown that using multiple UDP streams to transmit uncompressed video to a receiver can lead to congestion, even if there is sufficient bandwidth. They suggested is to shape the traffic by packetizing to avoid congestion and transmission bursts. However, in a multi user telepresence scenario, it is more favorable to have a method that does not depend on user position data.

In [6], Magnor et al. proposed light field compression schemes. They incorporated applied vector quantization, DCT coding and transform coding using spherical functions for light field compression and achieved interactive rendering. With wide FOV light field displays, compressing raw data may not yield sufficient reduction in bandwidth consumption.

III. BANDWIDTH REDUCTION SCHEME

In the current work, the light field information that needs to be reproduced at a remote site is captured using a set of horizontally aligned cameras. The light field capturing, transmission and displaying process chain is shown in Figure 1. Information from multiple cameras is acquired by a single acquisition computer, that runs as many processes as the number of cameras. The captured multiple image data streams are sent to a remote site via a Gigabit Ethernet switch with each stream consisting of continuous JPEG images which are made available on a distinct ports of the capturing node. Thus, it is possible to establish a TCP connection to a port from the receiver side to get hold of a specific camera image stream. To support control operations such as monitoring the streams and acquiring the calibration data, an application node is used at the capturing side.

Holographic light field displays use multiple projection modules to project light rays that hit a holographic screen which then re-directs the light rays for 3D visualization. To reconstruct the light field, the input images to the optical modules, which then emit this information in the form of light rays, must be properly rectified taking in to account the Holographic screen optical properties, origin of projection, camera calibration information and the region of interest in a given scene. These multiple optical modules are driven by a cluster of computers which take care of the required image rectification for light field rendering. For more information on light field rendering, please refer to [1] & [10].

Figure 2 shows the amount of information generated by the all the cameras at different time instances in an experimental capture, hereafter will be referred to as raw 3D frames. Each camera generates JPEG compressed RGB images of resolution 960x720 and the average amount of information is generated per raw 3D frame is 3.636 MB. If 25 frames are generated per second, the average amount of information to be transmitted in a second is approximately 91 MB. Currently, a single computer is capturing images from all the cameras and the effective 3D frame rate achieved is 10fps. In this case, the amount of information to be transmitted in a second would be 36.36 MB. The main aim of the current work is to reduce the size of the information being transmitted per second even further.

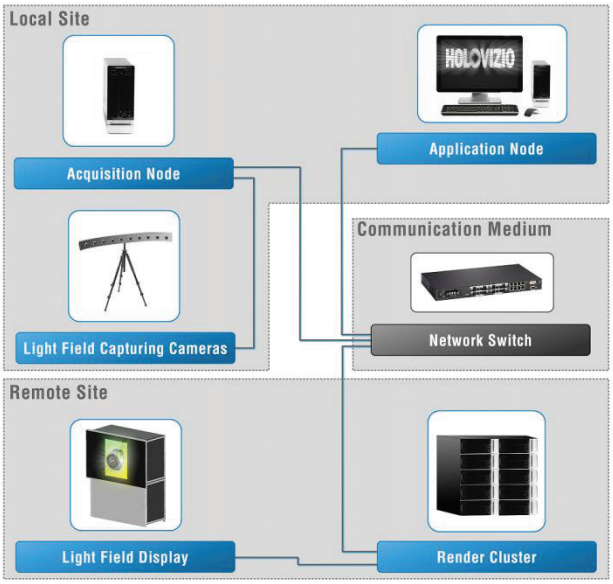


Figure 1. Light field capture, transmission and rendering pipeline.

Figure 2 shows the amount of information generated by the all the cameras at different time instances in an experimental capture, hereafter will be referred to as raw 3D frames. Each camera generates JPEG compressed RGB images of resolution 960x720 and the average amount of information is generated per raw 3D frame is 3.636 MB. If 25 frames are generated per second, the average amount of information to be transmitted in a second is approximately 91 MB. Currently, a single computer is capturing images from all the cameras and the effective 3D frame rate achieved is 10fps. In this case, the amount of information to be transmitted in a second would be 36.36 MB. The main aim of the current work is to reduce the size of the information being transmitted per second even further.

A. Experimental Setup

We assume a telepresence scenario in which the information is captured locally and sent to a remote site on a wired Gigabit Ethernet connection. Revisiting the Figure 1, at the transmission side, we have the multi-camera setup together with the acquisition and application node. The camera rig consists of 27 identical USB 2.0 cameras. The remote site is equipped with the light field display and the rendering cluster. In the current experiment, we use Holografika's HV80WLT light field display. This is the first 180 degree FOV light field display with a projection system incorporating virtual projection modules. In practice, the information from 27 cameras is not sufficient to account 180 degree FOV, but for experimental purpose we do not take this in to consideration in the current work. Figure 3 shows a sample capturing setup. Holographic displays incorporate horizontal only parallax and thus it is sufficient to arrange cameras in a 1D spatial topology. In the current experiment, we consider horizontally cameras. Estimating the capturing geometry involves

calculation of camera viewport matrix, projection matrix and view matrix. These matrices are defined collectively by camera calibration information.

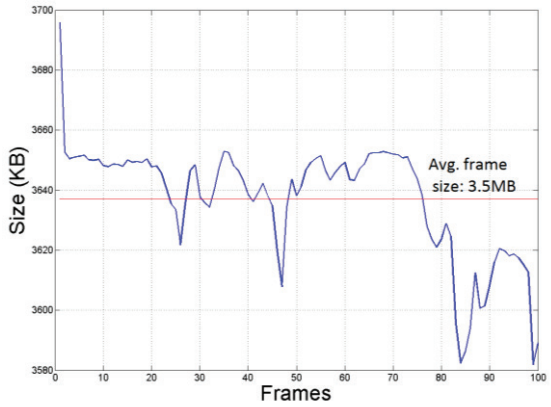


Figure 2. Amount of information produced by all the cameras per frame (JPEG compressed). in a sample capture

B. Data Reduction

As mentioned in the beginning of this section, part of light field reconstruction involves generating the projection module images from the incoming camera images. This inherently refers re-ordering the camera image pixels in order to suit the display 3D projection geometry. Depending the projection module composition group, each node in the render cluster efficiently implements this pixel re-ordering on GPU in real-time. As each rendering node drives a set of projection modules, depending on the location of this projection module group, the node makes use of pixels from many cameras. Thus for a single node, it is not necessary to have all the data from all the cameras. As the cameras are horizontally aligned they capture information from slightly different perspectives and thus it is more likely to have redundant information which might not be used during the process of light field reconstruction. The first step in our approach is to locate these unwanted pixels

Given the camera calibration and display geometry pixel re-ordering fashion remains identical temporally. This makes it possible to estimate the unwanted data during light field reconstruction (for more information on this, please refer to [3]). To facilitate data reduction, the pixel light ray mapping calculations are made locally at the capture side before the start of transmission and we safely store the information on appropriate pixel co-ordinates in the form of masks. Figure 4 shows the whole process.

At each time instant, from the incoming camera images, we extract the useful information carefully using the pre calculated masks. Although this itself reduces the size of data to be transmitted considerably, it is still possible to achieve significant data reduction by incorporating video compression schemes. From the masks it can be observed that the amount of information needed in vertical direction is more than that of horizontal and moreover it also remains same for all the camera images. This forms the basis for the second step in our approach. We integrate these patches and make a single 2D frame constituting multiple camera images (hereafter will be referred to as raw 3D frame). To speed up the memory access (rows of an image are stored in consecutive memory locations), these images are first transposed and integrated in to a raw 3D frame vertically. The idea behind the integration of patches is to explore and make use of the potential of legacy 2D video compression standards in transmitting raw 3D frames. The final step in our approach is to associate the H.264 codec in to the system. The raw 3D frames at each time instant are packetized using H.264 codec and sent to the receiver side. The reason behind choosing this codec is that it is an open source software that supports zero latency encoding and decoding and is currently one of the most commonly used formats for the recording, compression, and distribution of high definition video.

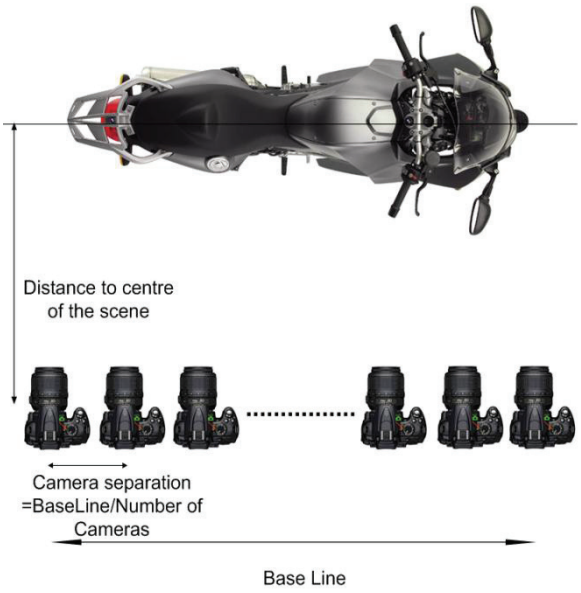


Figure 3. Sample multi-camera scene capture.

The encoded packets are broadcasted to all the nodes in rendering cluster on the receiving side. Each node decodes the packets it receives using corresponding decoder and uses the decoded raw 3D frame to produce a set of projection module images that it drives. Note that the bounding box information on useful pixels from each camera image is made available at the remote site beforehand to get back the original pixel co-ordinate system.

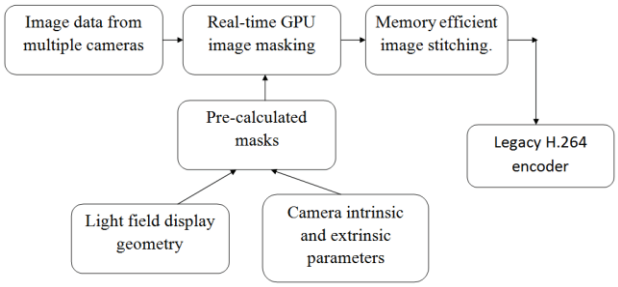


Figure 4. Data reduction approach

IV. Experimental Results

Figure 5 presents the data sizes after the initial and final stages of the presented approach. The curve in blue shows the size per raw 3D frame sent in the form of multiple streams on multiple ports. Comparing this with Fig [2], it can be clearly stated that there is approximately 70% of bandwidth saving at the first step.

With patch integration and single stream H.264 encoding, the average size per frame is further reduced. The codec settings were adjusted after trial and error to preserve the quality of decoded image. We found that the compression ratio of 9.81:1 would be enough to maintain the visual quality of light field rendering. Using this compression ratio, the average size of a raw 3D frame is calculated to be 42 KB and thus bandwidth saving is further increased by approximately 28%. Note that the 3D resolutions of the display and camera system are not identical. The codec settings were adjusted to emit one intra frame for every 25 frames which is the reason for the periodic peaks in the H.264 single stream curve. Results show that using the presented approach, 25fps raw 3D video can be comfortably streamed on a 10 Mbps line. The average connected speed needed is calculated to be 8.2 Mbps.

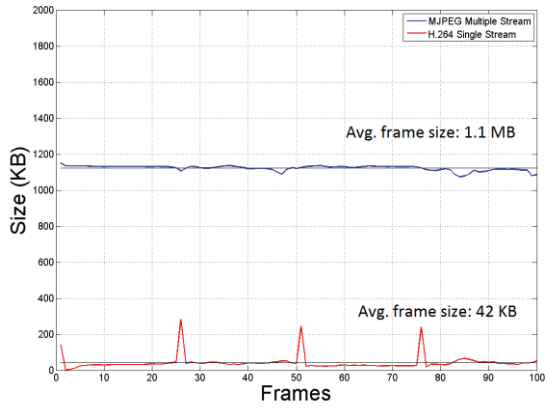


Figure 5. Data reduction approach results.

Figure 6 shows the PSNR in dB of the decoded raw 3D frames. Using the presented approach we could achieve an average PSNR of 36.38 dB.

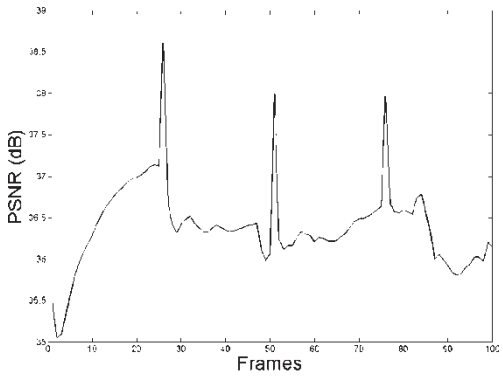


Figure 6. PSNR of decoded frames.

V. Conclusions and Future Work

In this paper we presented a dynamic bandwidth reduction approach for transmission of light field data in a telepresence scenario. The presented system is first of its kind in exploring and re-ordering the 3D information to fit into the legacy 2D compression methods. The captured light field information from multiple cameras is efficiently compressed using two stage data reduction process. We proposed simple image processing steps to extract and retrieve only required data which lead to fast and efficient compression. Results show that the approach is very useful and the amount of data being transmitted is reduced up to 98%. The decoded image stream has sufficient quality for light field reconstruction.

Although the required bit-rate for light field transmission is very much reduced by this approach, in practice it may not be sufficient for streaming the data over internet in real-time because of the additional overhead in transmission. Thus further investigation in the data reduction possibilities is required. In the current working system, from each camera image, pixels used by all the nodes on remote side are extracted. Thus all nodes receive slightly additional information that they do not require. Another possibility is to multicast different streams for different nodes rather than broadcasting the same information to all the nodes.

Acknowledgement

The research leading to these results has received funding from the DIVA Marie Curie Action of the People programme of the European Unions Seventh Framework Programme FP7/2007- 2013/ under REA grant agreement 290227.

References

[1] Tibor Agocs, Tibor Balogh, Tamas Forgacs, Fabio Bettio, Enrico Gobbetti, Gianluigi Zanetti, and Eric Bouvier, “A large scale interactive holographic display,” *IEEE Virtual Reality Conference (VR 2006)*, page 57, Washington, DC, USA, 2006.

[2] Hutanu Andrei and Nimmagadda Sandeep, “Using Traffic Shaping to Improve Performance for Raw Data Transfer and Video Streaming,” *Communication Networks and Services Research Conference (CNSR),* pages 49-55, May 2011.

[3] Vamsi Kiran Adhikarkla *et. al.,* “Fast and efficient data reduction approach for multi-camera light field display telepresence systems,” *3DTV-Conference: The True Vision-Capture, Transmission and Dispaly of 3D Video (3DTV-CON), 2013* , vol., no., pp.1,4, 7-8 Oct. 2013.

[4] Edouard Lamboray, Stephan Wurmlin, and Markus Gross, “Data Streaming in Telepresence Environments,” *IEEE Transactions on Visualization and Computer Graphics* Vol. 11, November 2005, pages 637-348.

[5] Jyh-Ming Lien, Gregorij Kurillo, and Ruzena Bajcsy. 2009, “Multi-camera tele-immersion system with real-time model driven data compression: A new model-based compression method for massive dynamic point data,” *The Visual Computer: International Journal of Computer Graphics*, Vol. 26, November 2009, pages 3-15.

[6] Magnor, M.; Girod, B., “Data compression for light-field rendering,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol.10, no.3, pp.338,343, Apr 2000.

[7] Maimone, A.; Fuchs, H., “Encumbrance-free telepresence system with real-time 3D capture and display using commodity depth cameras,” 2011 10th *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp.137,146, 26-29 Oct. 2011.

[8] Maimone, A. and H. Fuchs., “A First Look at a Telepresence System with Room-Sized Real-Time 3D Capture and Large Tracked Display,” *21st International Conference on Artificial Reality and Telexistence (ICAT),* Osaka, Japan, November 28-30, 2011

[9] Benjamin Petit and Jean-Denis Lesage and Clément Menier and Jérémie Allard and Jean-Sébastien Franco and Bruno Raffin and Edmond Boyer and François Faure, "*Multicamera Real-Time 3D Modeling for Telepresence and Remote Collaboration,*" International Journal of Digital Media Broadcasting, vol. 2010, 2009, pages = {247108--12}

[10] Balogh Tibor and Peter Tamas Kovacs , "Real-time 3D Lightfield Transmisison," *Proc. SPIE*, 7724:772406–772406–7, 2010

[11] Malte Willert, Stephan Ohl, and Oliver Staadt, "Reducing bandwidth consumption in parallel networked telepresence environments," In *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry* (VRCAI '12). ACM, New York, NY, USA, 247-254.

Modeling bacterial quorum sensing in open and closed environments

Dóra Bihary
(Supervisor: Dr. Sándor Pongor)
bihary.dora@itk.ppke.hu

***Abstract*—Quorum sensing (QS) is a process of bacterial communication and cooperation mediated by the release of jointly exploited signals and public goods into the environment. There are conflicting reports on the behavior of mutants deficient in the release of these materials. Namely, mutants that appear perfectly viable and capable of outgrowing wild type cells in a closed model system such as a culture flask, may not be viable or invasive on open surfaces such as agar plates. Here we show via agent based computational simulations that this apparent discrepancy is due to the difference between open and closed systems. We suggest that the experimental difference is due to the fact that wild type cells can easily saturate a well-mixed culture flask with signals and public goods so QS will be not necessary after a certain time point. As a consequence, QS deficient mutants can continue to grow even after the wild type population vanished. This phenomenon is not likely to occur in open environments including open surfaces and agar plate models. In other words, even if QS is required for survival, QS deficient mutants may grow initially faster in short term lab experiments or computer simulations, only WT cells appear stable on longer time scales, especially when adaptation to changing environments is important.**

***Keywords*-Bacterial communication; Agent-based model; Quorum sensing; Cooperation; Cheating**

I. Introduction

Quorum sensing (QS) is a cell-cell communication process in which bacteria release and detect molecular signals called autoinducers, which enables them to monitor cell population density and make a variety of coordinated responses [1]. Acyl homoserine lactones (AHLs), the major class of autoinducer signals used by Gram-negative bacteria, consist of a conserved homoserine lactone ring coupled to an acyl side chain, which may vary from 3 to 18 carbons in length. All AHLs are believed to freely diffuse across the cell envelope; however efflux pumps may actively export some longer chain AHLs [2]. In a typical AHL-QS circuit, the AHL signals are synthesized by a LuxI-type protein. At a critical concentration, the AHL binds to a LuxR-type protein, and the LuxR-AHL complex acts on target genes that affect a variety of cellular processes [3] including the production of exoenzymes and exopolysaccharides. As these QS-regulated factors are released into the environment and are accessible to other cells, they are often referred to as public goods i.e. openly accessible means of intercellular cooperation. For instance, cells of the ubiquitous opportunistic pathogen *Pseudomonas aeruginosa* has two AHL-QS systems , the so called *las* and the *rhl* system (see, e.g. [4] [5]). The *las* system activates the production and release of an elastase enzyme which will digest proteinaceous nutrients in the environment into amino acids that the bacteria can directly utilize, while the rhamnolipids released by the *rhl* system facilitate the movement of swarming cells. The two AHL systems of *P. aeruginosa* are interlinked which can give rise to complex cellular responses. For instance, mutants containing a single deletion in the *las* system affected several pathways. This is all the more remarkable since some Gram-negative bacteria use the same type of AHL signals. Sharing of signals and other public goods can then lead to a coordinated behavior of multispecies consortia in which different species communicate and cooperate.

The role of the *luxI* and *luxR* genes can be conveniently studied by deletion mutants [6]. Cells in which the *luxI* gene is deleted will not produce the signal, so they are termed "signal negative" or "SN". Even though these cells cannot produce the signal above a very low, baseline level, they can respond to signal molecules produced by other bacteria, for instance by producing public goods. In other words, SN cells cannot communicate but can cooperate. Cells in which the *luxR* gene is deleted cannot respond to the signal so they are termed "signal blind" or "SB". SB cells are not able to upregulate their signal production, nor do they produce public goods. In other terms SB cells neither communicate, nor cooperate.

The behavior of AHL QS mutants of *Pseudomonas aeruginosa* has been studied both in laboratory models and with computational models. Diggle et al. [6] studied the experimental behavior of $\Delta lasI$ (SN^L) and $\Delta lasR$ (SB^L) deletion mutants in shaken cultures grown in defined medium. In this medium, SN^L and SB^L mutants grew slower than the wild type (WT). As adding exogenous signal molecules to the system restored the growth of SN^L to the WT levels, but not that of SB^L. If however the mutants were grown in pairwise competitions with the WT, both were able outgrow the wild type cells. This phenomenon was ascribed to the lower metabolic costs of deletion mutants, however one could also argue that the active *rhl* system may also have contributed to the growth of the mutants. For this reason, Venturi and coworkers [7] constructed Δlas ΔrhlI (SN^{LR}) and Δlas ΔrhlR (SB^{LR}) double deletion mutants, and studied their behavior on swarming agar plates. On these plates WT *Pseudomonas aeruginosa* cells formed branched colonies typical of swarming cells, but neither SN^{LR} nor SB^{LR} cells could swarm alone. Non-swarming colonies remained at the starting position without any visible sign of growth. In pairwise competitions, $WT + SN^{LR}$ cells formed branched colonies only slightly

D. Bihary, "Modeling bacterial quorum sensing in open and closed environments,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 79-82.

smaller than those of pure WT populations. $WT + SB^{LR}$ cells however could not swarm, but formed a colony slightly bigger than the starting population, as if the population starting to grow would suddenly collapse. Our question was why SB^L mutants can outgrow the WT cells in pairwise competitions in liquid culture while $WT+SB^{LR}$ populations collapse on agar plates. One of the immediate answers could be the difference between SB^L and SB^{LR} mutants, i.e. SB^L mutants may grow better because of their active *rhl* system, i.e. there may be a difference between single and double knockout mutants. This is however not the case, since the behavior of the single and double mutants (both SN and SB) was indistinguishable when tested on agar plates [7] This identical behavior lead us to speculate that the difference may be due to some aspect of the test system, probably the open or closed nature of the experimental system.

Here we test the hypothesis whether or not the different and seemingly contradictory results reported previously are due to the differences between closed and open systems. Namely, well stirred, shaken cultures can be quite rightly considered as closed systems where each cells of the populations can be in contact and can influence the behavior of any other cell in the population. Agar plates, on the other hand, mimic open surfaces inasmuch as the medium is not stirred so that cells interact only with their close neighbors. Even though the natural environments of bacteria can be considered only approximately closed or open, we have to note that QS signaling is profoundly affected by the closeness/openness of the environment. For instance, a single bacterial cell will turn on QS if placed in a very small volume, and similar conditions can exist when bacterial cells reach the microcapillaries of plants.

In this paper we use *in silico* modeling to study the behavior of WT *Pseudomonas aeruginosa* as well as its SN and SB mutants in open and closed environments and show that the previous, seemingly contradictory results are not artifacts but are in fact due to the differences between open and closed experimental systems.

II. Results and discussion

We carried out two types of experiments: competition and invasion. In competition experiments, the participating species were present in equal numbers. In invasion experiments, a small number of mutant cells were added to a large population of wild type agents. In both cases we used two types of environment, the so called open and closed models. The concept of the two models is shown in Figure 1.

Our open model is a longitudinal surface, bacterial agents start from the bottom of it and move upwards during the simulation (Fig. 1B). The surface is divided into cells in which the concentration of the solutes (signal, factor and nutrient) is the same. The simulations start with a finite amount of nutrient in each, which decreases as bacteria consume it. As a result, nutrient will diffuse to the square from areas with higher nutrient concentration. In a similar way, the signal and public goods produced by the cell also spread via diffusion.

A closed system is simpler, there is only one spatial unit ('square' - i.e. toroid surface) (Fig. 1A), so the concentration level of nutrient, signal and public goods is constant throughout the whole system. We introduced an infinite amount of nutrient otherwise all simulations ended with the bacteria (WT or mutant) perishing by starvation. In order to obtain realistic, sigmoidal growth curves we maximized the number of bacteria in a way that the agents consumed less energy as they approached the population limit, which is a standard approach in many areas of biological modeling [8]. All calculations were carried out with explicit representation of cells, diffusible signals and public goods as described in [9] [10] [11] and shown in Figure 1.

A. Competition experiments

First we carried out competition experiments, where we aimed to examine the behavior of competing species. The participating species (wild type and a mutant) were present in equal numbers. We tested four cases: WT + SN and WT + SB species, on open and closed surface (Figure 2).

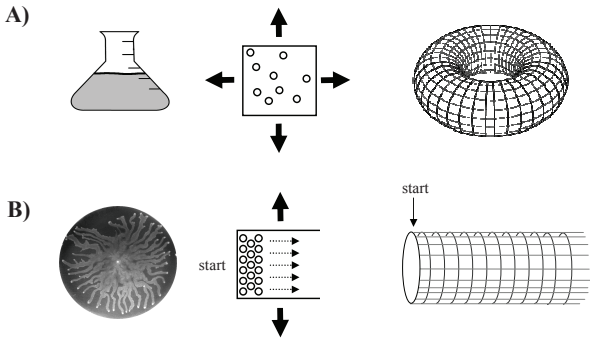


Fig. 1. Modelling closed and open systems in 2D. A) A closed system (culture flask) is represented as a square in which arrows indicate periodic boundary conditions on all sides (left), which corresponds to a toroidal surface (right). B) The open system (agar plate) has periodic boundary conditions on two sides only, indicated by black arrows (left), which corresponds to a cylindrical surface open on one end (right). Circles indicate cell agents. The stating population is randomly distributed all over the closed system. In the open system the cell agents are randomly positioned in the vicinity of the "start", and during the modeling experiment the (growing) community spontaneously proceeds as indicated by the dotted arrows.

TABLE I
Typical results for the position and density dependence of invasion experiments*

	Proximal	Distal
Uniform	"-"(1)	"+"(2)
Concentrated	"+"(3)**	"+"(4)

* 50 non-communicating SN mutants + 1950 wild type agents placed in different ways shown in Figure 3. The numbers in parenthesis refer to the positions indicated in Figure 3.
** stochastic coexistence occoured in 35% of cases (in the other three cases the observed outcome could be obtained in 100% of the simulations).

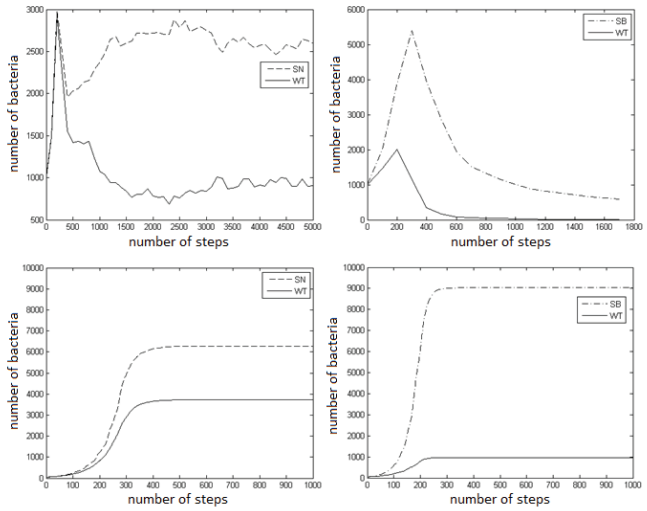


Fig. 2. Competition of WT and mutant bacterial models in open (top) and closed (bottom) systems. The experiments were set up with an equal number of WT (solid line) and mutant (SN indicated by dashed, or SB indicated by dash-dotted lines) cells and the population size was plotted as a function of time.

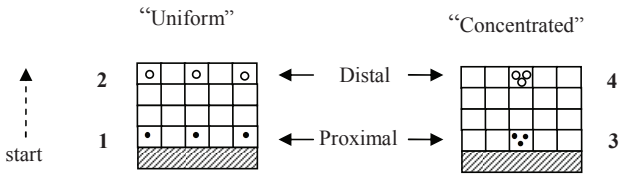


Fig. 3. Modeling position and density dependence in an open system. A) Uniform distribution means equidistant positioning of agents. B) Concentrated distribution means all agents are placed at a minimal allowed distance from each other. Proximal positioning means placing the agents close to the boundary of the system while distal means placing it further apart. Note that proximal/distal positioning is only possible in our open system as the closed system has no fixed boundaries.

SN mutants form a stable community with the WT species both in the open (top left) and closed system (bottom left), and mutants grow faster in both cases than WT bacteria. On the other hand, SB species can outgrow the WT only in the closed system (bottom right). In the open system we obtained a population collapse (top right). Each of our models have a single QS system so we use the notations SB and SN for the mutants, without the upper indices.

Our results are similar to the ones that were obtained by Venturi et al. [7] and Diggle et al. [6]. As mentioned above, Venturi and coworkers saw swarming populations in WT plus SN competitions, however collapsing, not swarming populations when adding SB mutants to WT agents, like in our open model. In Diggle and associates work both mutants could outgrow the WT agents in pairwise competitions similarly as agents behave in our closed model. These results suggest that the difference in the two experiments can be explained with the difference of the environments, and not with the difference of the mutant species used in the two studies (single knockout mutants in Diggle et al. [6] and double knockout mutants in Venturi et al. [7] work).

The difference between the two experimental systems can be further characterized by considering the nutrient supply and the response of bacterial populations. In open environments, a population constantly proceeds to pristine areas of the agar plate, however the local nutrient supply is exhaustible and must be competed for. This leads to the collapse of QS signaling in some cases, and there will be no growth even though nutrients are still available. In closed, well-mixed, submerged cultures, experimenters usually add a large amount of nutrients to the culture that is not exhausted during the time of the experiment which is typically less than 48 hours, and a growth inhibition effect will stop the growth of the populations before all nutrients are exhausted. But is it possible to reach collapse in a submerged culture? We believe not: as long as there are nutrients in the system, both competing populations will grow, and since the nutrients are evenly distributed, growth will be detected throughout the entire system. Moreover, it is a well known experimental fact that spent media contain large amounts of signals and public goods. We can simplify this situation by saying that a closed flask is easily saturated with signals and public goods so after a certain time QS will not be necessary for survival. This is not the case in realistic, open systems.

B. Invasion experiments

Invasion experiments are meant to determine whether or not a given mutant is able to spread in a WT population. Our goal was to determine the invadability of a mutant in open and closed systems. Exploratory experiments showed that there is a huge positional and density dependence in the system, the results whether the invasion is successful or not were remarkably dependent on these two factors. Invading mutants positioned far from each other and near the system boundary (position "1" in Figure 3) the success rate of the invasion was usually lower than in the opposite case (position "4" in Figure 3). Typical results are shown in Table I. It was also noted that the models behave in a stochastic way, i.e. the same number of mutants positioned in the same way invaded the community only in a fraction of the experiments. Large number of experiments were carried out in which the density of the invading mutants was systematically varied (Table II). We found, that both the non-communicating and the non-cooperating mutants can invade the wild type community by growing initially faster than the wild type.

Our simulations thus show that both SN and SB mutant species can invade the wild type cultures in a closed environment, regardless of their initial density (Table II). Simulations in open cultures show a stochastic behavior, and SN and SB can invade the WT population, however the outcome is either collapse (SB mutants) or a stable mixed community (SN mutants).

TABLE II
INVADABILITY IN OPEN AND CLOSED MODELS *

Type of invasion	Response in the closed space model	Response in the open space model
Non-communicating mutant SN		
1 SN : 1999 WT	"+" (100%)	"+" (0-65% coexistence) **
50 SN : 1950 WT	"+" (100%)	"+" (0-100% coexistence)
1000 SN : 1000 WT	"+" (100%)	"+" (100%)
Non-cooperating mutant SB		
1 SB : 1999 WT	"+" (100%)	"+" (0-55% collapse)
50 SB : 1950 WT	"+" (100%)	"+" (0-100% collapse)
1000 SB : 1000 WT	"+" (100%)	"+" (100% collapse)

* "+" means that mutant species could grow up, resulting a coexistence for SN, and a collapse for SB mutants. Values in parenthesis represent the % of cases that the "+" outcome was observed in 20 repetitions.
** stochastic, position and density dependent response

We mention that both the theoretical and the experimental aspects of quorum sensing are subject of intensive studies [12]. The most frequent goal is to discuss the long term evolutionary stability of cooperation of which bacterial QS is often considered to be a good model. Our general approach is different: i) We study the role of QS in short-term community formation; ii) We use agent-based i.e. particle-like models better known in physicochemistry and physics. So we can not make strong claims about the evolutionary fate of mutants, we can only suggest that mutants that die out in short times are probably not stable on the evolutionary time scales either. More importantly, we believe that QS is meant to facilitate bacterial survival in changing environments, so it may go lost only in environments where wide adaptation is no longer important. A shaken culture flask, saturated with signals and public goods is such an environment.

III. CONCLUSIONS

The behavior of QS bacteria is often characterized with laboratory experiments, especially in closed systems such as well stirred submerged cultures. *In silico* modeling experiments that we presented in this work show that QS deficient mutants that are not viable in an open environment may appear viable and capable of invading wild type colonies in closed systems. We suggest that this difference can be explained by the fact that wild type cells can easily saturate a culture flask with signals and public goods allowing QS deficient mutants to continue to grow even in the absence of WT cells, since they do not need an intact QS system - a condition rarely met in natural environments. As quorum sensing bacteria often shuttle between different environments, the differential behavior with respect to open and closed systems may need to be considered when describing the viability of mutants. Importantly, while QS mutants grow initially faster in short term lab experiments or computer simulations, only WT cells are deemed to be stable on longer time scales, where adaptation to fluctuating environments is important.

This work is in press [13].

IV. ACKNOWLEDGEMENT

This project was developed within the PhD program of the Multidisciplinary Doctoral School, Faculty of Information Technology, Pázmány Péter Catholic University, Budapest. Thanks are due to my supervisor, Prof. Sándor Pongor and to Ádám Kerényi, who is the major developer of the agent-based simulation program.

REFERENCES

[1] W C Fuqua, S C Winans, and E P Greenberg. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J Bacteriol*, 176(2):269–275, 1994.

[2] T Kohler, C van Delden, L K Curty, M M Hamzehpour, and J C Pechere. Overexpression of the MexEF-OprN multidrug efflux system affects cell-to-cell signaling in Pseudomonas aeruginosa. *J Bacteriol*, 183(18):5213–5222, 2001.

[3] K A Egland and E P Greenberg. Quorum sensing in Vibrio fischeri: elements of the luxI promoter. *Mol. Microbiol.*, 31:1197–1204, 1999.

[4] Z Gelencser, B Galbats, J F Gonzalez, K S Choudhary, S Hudaiberdiev, V Venturi, and S Pongor. Chromosomal Arrangement of AHL-Driven Quorum Sensing Circuits in Pseudomonas. *ISRN Microbiol*, 2012:484176, 2012.

[5] V Venturi. Regulation of quorum sensing in Pseudomonas. *FEMS Microbiol Rev*, 30(2):274–291, 2006.

[6] S P Diggle, A S Griffin, G S Campbell, and S A West. Cooperation and conflict in quorum-sensing bacterial populations. *Nature*, 450(7168):411–414, 2007.

[7] V Venturi, I Bertani, A Kerenyi, S Netotea, and S Pongor. Co-swarming and local collapse: quorum sensing conveys resilience to bacterial communities by localizing cheater mutants in Pseudomonas aeruginosa. *PLoS One*, 5(4):e9998, 2010.

[8] F. J. Richards. A Flexible Growth Function for Empirical Use. *Journal of Experimental Botany*, 10(2):290–301, 1959.

[9] S Netotea and S Pongor. Evolution of robust and efficient system topologies. *Cellular Immunology*, 122(244):80–83, 2006.

[10] V Venturi, A Kerenyi, B Reiz, D Bihary, and S Pongor. Locality versus globality in bacterial signalling: can local communication stabilize bacterial communities? *Biology Direct*, 5:30, 2010.

[11] A Kerenyi, D Bihary, V Venturi, and S Pongor. Stability of multispecies bacterial communities: signaling networks may stabilize microbiomes. *PLoS One*, 8(3):e57947, 2013.

[12] Daniel M Cornforth and Kevin R Foster. Competition sensing: the social side of bacterial stress responses. *Nature reviews. Microbiology*, 11(4):285–93, April 2013.

[13] D. Bihary, M. Tóth, Á. Kerényi, V. Venturi, and S Pongor. Modeling bacterial quorum sensing in open and closed environments: Potential discrepancies between agar plate and culture ask experiments. *Journal of Molecular Modelling*, in press.

A test platform for real-time measurement and analysis of myoelectric signals

Bence József Borbély
(Supervisor: Dr. Péter Szolgay)
borbely.bence@itk.ppke.hu

Based on conference papers written with the following co-authors: Zoltán Kincses, Zsolt Vörösházi, Zoltán Nagy and Péter Szolgay.

***Abstract*—A hardware platform for real-time measurement and analysis of human electromyographic signals (EMG) is proposed. The system is designed to have modular structure to facilitate further integration of various measurement, signal processing and actuator control tasks. To assess reachable system performance in the target application of EMG classification, two different classification methods are tested on the Xilinx Zynq-7000 All Programmable SoC using offline recorded multichannel lower arm muscle signals.**

I. Introduction

The field of intentionally driven electromechanical upper limb prostheses has evolved significantly in the last decade. As a result, prosthetic devices for almost all level of upper limb amputations are available today, providing a wide range of complexity and dexterity from simple grippers to the Modular Prosthetic Limb [1].

On the control side, the most widely used source is the myoelectric signal (electromyogram, EMG) that can be measured on the covering skin of muscles during muscle contraction. This signal reflects the summed motor unit activity around the recording area, therefore it can be related to muscle contraction and exerted force (however, in most cases this relationship is highly non-linear). The EMG signal has become relevant in upper limb prosthesis control because it can be used to analyze movement patterns at the level of muscle activations and tell specific movement intents in a non-invasive way even in the case when the actuated end-effector is missing from the system (for example consider a hand amputee patient who's hand is missing but the muscles responsible for wrist and finger movements are still present in his/her forearm). Different strategies have been developed for prosthesis control using the myoelectric signal ranging from the conventional method (mostly used for simple gripper actuators where the controller needs only two recording sites to drive one degree of freedom (DoF)) to more sophisticated pattern recognition techniques. These methods eliminate the need for isolated EMG signals and increase the amount of extracted information from the recordings (allowing the control of multiple degrees of freedom) [2]–[6].

The standard pattern recognition algorithms are optimized for real-time operation on low-power embedded processors, considering that the final system should fit into the fixing structure of the prosthesis (e.g. a forearm socket) and must be powered from a suitable battery. These requirements, while being application-centric, imply some limitations for the possible analysis algorithms using only conventional embedded processors.

In this study we propose a reconfigurable prototyping platform featuring a Field Programmable SoC to develop and test advanced prosthesis control methods. The system integrates both signal measurement and analysis while providing low power consumption and great flexibility for algorithm design because of the on-chip FPGA fabric. As test cases, two pattern recognition based classification algorithms were implemented and tested on the system.

II. System design and implementation

The main goal of this study was to compose a prototyping system that fulfills the following requirements: (1) provide a safe environment for measurement and analysis of biomedical signals, (2) enable the design, implementation and real-time operation of algorithms with higher computational needs, including parallel processing, (3) provide modularity and the possibility of reconfiguration in hardware, and (4) keep power consumption and total area of the system as low as possible.

A. Measurement Unit

To provide a complete platform for both measurement and analysis of myoelectric signals, we have started to develop a standalone measurement device that is capable of recording up to 8 (bipolar) EMG channels with 1 kHz sampling rate and 16 bit resolution. The unit is designed to be battery powered and carried by the subject to assure the less possible limitation in movement space during the measurement. To minimize noise corruption of the recorded signal, an active electrode design was chosen where signal amplification and digitization are performed on the electrode site. The recorded data can be stored on an SD card or streamed in real-time using the built-in Bluetooth interface. Using this approach we can assure that the measurement is completely safe for the subject while having access to an online stream of measured data. The block diagram of the Measurement Unit is depicted in Figure 1.

B. Digilent Zedboard

Considering the aforementioned key requirements we have chosen the Avnet/Digilent Zedboard as the central element of the system. The board is based on a Xilinx Zynq 7020 SoC architecture, which is built up from a dual-core ARM Cortex-A9 PS (Processing System) and a Xilinx Series-7 PL (Programmable Logic) fabric using 28 nm manufacturing process. The SoC tightly integrates the embedded microprocessor unit (MCU) with the FPGA fabric by using standard AMBA

B. J. Borbély, "A test platform for real-time measurement and analysis of myoelectric signals"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 83-86.

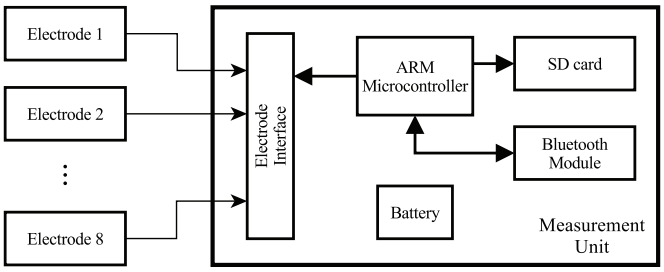


Fig. 1. Block diagram of the Measurement Unit. Sensor data from the active electrodes are captured via the electrode interface and stored on an SD card or streamed via Bluetooth connection.

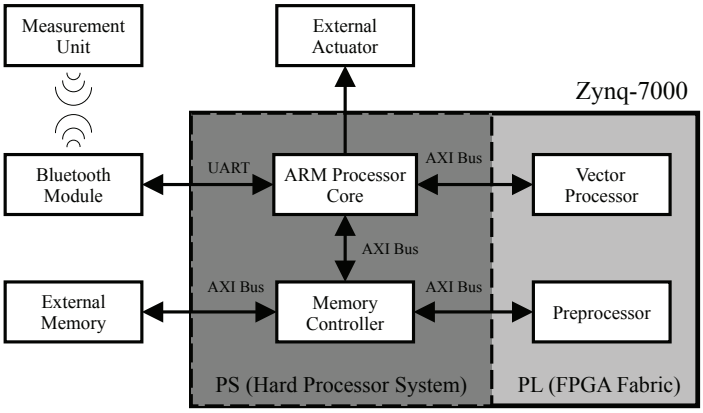


Fig. 2. Block diagram of the implemented system. Sensor data is received via Bluetooth connection from a measurement unit and processed with a distributed architecture on the Zynq-7020 processor (with computationally demanding blocks implemented in hardware on the PL side of the SoC). This implementation provides two operation modes for signal processing - an *offline* mode where previously measured data are processed from the SD card, and an *online* mode where a continuous stream of measured data is processed directly.

AXI4 bus interfaces. As in the case of traditional SoCs this platform also allows the so-called processor centric approach, where the startup sequence is partitioned into two steps: first, the Processing System is initialized then the Programmable Logic is configured. The interconnections between the PL (which is based on the Xilinx's Artix FPGA family) and the PS are implemented using several AXI4 buses. Four 32-bit wide interfaces (bridges) are dedicated to low latency access to the peripheral registers implemented in the PL. To handle fast transmission of large amounts of data between the PL and the different memories four 64-bit wide, high performance AXI4 buses are accessible. A specialized 64-bit wide coherent AXI4 bus is connected to the snoop protocol unit of the L2 cache, which is also accessible for tightly integrated coprocessors. As it can be seen, the Zynq-7020 architecture provides good basis for high performance embedded systems and algorithm implementation even at the hardware level.

In addition, the PS has wide variety of different I/O interfaces to connect the system to the outside world such as UART, I2C, SPI, USB and Gigabit Ethernet to name a few. These features, utilized on the Zedboard assure that a fully modular system can be built, where measurement data storage (e.g. on SD card) and analysis is performed on the board. Large variety of external measurement and actuator devices can be easily attached to the system using standard communication protocols. This modularity allows the utilization of the system in various bio-signal processing tasks, including but not limited to myoelectric control applications.

C. Test platform for MES applications

By extending the Zedboard with the Measurement Unit we propose a complete test system for MES algorithm prototyping depicted schematically in Figure 2. As the figure shows, the central element of the current implementation is one ARM processor core that manages data streaming (via Bluetooth) and storage (through the Memory Controller), accelerated data processing involving custom hardware implementations in the PL and a possible actuator control. It is important to note that this implementation is specific to the tested processing algorithms (more details in sections III. and IV.) and system architecture can be modified for further specific needs. This implementation provides two operation modes: an *offline* mode where previously measured data are processed from the SD card, and an *online* mode where a continuous stream of measured data can be processed directly. However the development of the Measurement Unit is still in progress, its behaviour can be simulated using a PC with Bluetooth connection, streaming offline measured test MES data to the system, which allows the testing of *online* operation.

III. Classification methods

To test the performance of the system in real signal processing applications, two pattern recognition based classification algorithms are being implemented. The first is the widely used "standard" pattern recognition method, while the second is a modified version that was recently proposed [7].

A. Standard pattern recognition - velocity control

The standard pattern recognition method is based on a supervised learning algorithm, where the two stages of operation (training and classification) are separated. In the training session, multichannel myoelectric signals are recorded during different predefined isometric contractions of muscles (different states) and specific features of these recordings are stored as separate state descriptors. Classification data based on all measured states is calculated offline (after the measurements). Following this stage an online stream of myoelectric data from the same electrodes is recorded and classified to categorize the actual signal into one of the trained classes. One important feature of the method to note is that it allows velocity control of the specified degrees of freedom, where usually two classified states are required to control one degree of freedom (e.g. hand opening/closing).

MES data is non-stationary and stochastic in nature therefore most of the related analyses apply processing windows to extract descriptive features of the signal. In the test implementation a 150 ms long processing window was used because it enables optimal classification performance [5]. The spatial selectivity (the number of separable movement classes) in the system is highly determined by the number of separate recording channels. Previous studies justified that in the case of lower arm recordings four channels of MES are suitable to classify online measured data into one of six separate

classes with high efficiency [4]. Based on these results we implemented a four-channel system as the basis of the test environment.

In real prosthetic applications overall latency and response time are critical factors of device acceptance which are determined by the processing window length and the amount of processing window shift (or sampling delay) during operation. Among these two factors window shift value can be varied to obtain different temporal resolutions, resulting that shorter shifts yield better response times at the cost of bigger computational overhead. In this study the algorithm was tested with window shifts ranging between 1-50 ms.

1) Signal features: To reduce data dimensionality while preserving signal characteristics at the same time, the standard four element time-domain feature set was calculated for each data window (150 ms) and channel. These features are the mean absolute value (MAV), number of zero crossings (NZC), number of slope sign changes (NSSC) and the waveform length (WL) as described in previous studies [2], [3]. It is important to note that these features give only estimations of specific signal properties (e.g. NZC $\sim$ frequency) but it has been shown that they provide as good basis as frequency-domain features for classification of stationary signals for less computational cost and induce lower latency in the system [4].

2) LDA classifier: To partition the feature space into six selected classes for pattern classification, linear discriminant analysis (LDA) was used as described in [8]. LDA is a statistical method that is able to sufficiently reduce feature space dimensionality considering the properties of the separate subspaces. More specifically it finds projection vectors in the complete feature space which best separate the individual classes when the dataset is projected to them. After the projection vectors are calculated (num. of projection vectors $\ll$ num. of feature space dimensions) data points from the original feature space are projected to these vectors to get a more separable set of target classes (dimension = num. of projection vectors).

During online classification the actual recorded data is first transformed into the feature space (by calculating its time-domain features) followed by the projection to the same vectors obtained with the LDA algorithm. The classification takes place when these projected values are compared to the stored projections of the target classes and class labels are assigned to the data based on a chosen distance metric (e.g. Euclidean or Mahalanobis distance) from the stored class values.

B. Modified pattern recognition - position control

The previous section showed how the standard LDA-based pattern recognition method can be used to classify static kinematics from EMG recordings of isometric contractions. This approach has the advantage of controlling multiple DoF of a prosthetic hand sequentially but without mode switching, which highly increases usability compared to conventional systems. On the other hand, because classifier training is performed on data sets of isometric contractions, the standard approach does not allow continuous estimation of the instantaneous kinematics, which could be the next step towards better prosthesis usability.

As a possible solution to this phenomenon, a proof of concept of a modified version of the standard pattern recognition method was recently proposed [7]. The published method uses the same time domain features (MAV, NZC, NSSC, WL), window length (150 ms) and window shift (50 ms) values as the implemented standard method. The difference lies in the classifier training method - the classifier is trained with classes identifying different angle ranges of only one selected DoF (wrist flexion) instead of referring to the extreme values only. This modification enables prediction of the actual kinematic state of the selected DoF, providing position control of the prosthesis. Using this approach in a real prosthetic application, each DoF would need a separate classifier to train with DoF-specific training data. In this study we implemented the modified classifier for only one DoF, because additional measurement data was not available at the time of writing.

IV. Algorithm implementation

A. Hard Processor System

The classification algorithms described in Section III. were implemented in ANSI-C on a laptop computer having an Intel Core i5-540M CPU running at 2.53 GHz. The extracted four time-domain features were *MAV, NZC, NSSC, WL*. All linear algebra calculations (mostly involved by the LDA) were performed using self-written implementations. Inverse matrix calculation was performed based on Gauss-Jordan elimination. Calculation of LDA projection vectors requires only eigenvectors to be calculated accurately, but not eigenvalues, therefore eigenvalues were only estimated using the QR iteration with limited number of steps. The eigenvectors were then accurately calculated applying the Inverse Iteration to the estimated eigenvalues.

The development system was running Ubuntu Linux 12.04 LTS operating system and the *gcc* compiler was used to generate executables. To compile the source code for the ARM cores of the Zynq processor, gcc's cross compiler version (arm-linux-gnueabi-gcc) was used. For optimal performance the -O3 compiler option was applied in both situations.

B. Architecture on FPGA

Our proposed architecture contains four main parts. The *ARM Processor Core* and the *Memory Controller* are implemented on the hard-processor system (PS) fabric of the Zynq SoC. The *Vector Processor* and the *Preprocessor* are located on the PL (FPGA fabric) part of the Zynq platform as can be seen on Figure 2. The *ARM Processor Core* executes the high-level steps of the algorithm as described in Section III, while the *Vector Processor* fulfills the vector operations that are required mostly in the training part of the algorithm. The time-domain features are calculated by the *Preprocessor*. In offline mode, the data received from the built-in UART peripheral of the PS and stored in the external DDR3 memory, while in online mode the data is fed directly to the *ARM Processor Core*. The *Memory Controller* is responsible for the control of the external memory and the load and store operations in offline mode. The implemented units communicate with each other using AXI-4 Interconnect buses.

As described before the *Preprocessor* is responsible for the calculation of the time-domain features. With these features, the dimension of the actual EMG data window can be reduced. This unit is built up from four main components using the

data of a 150 ms long processing window. The *MAV Unit* responsible for the calculation of the average of the summed-absolute values in the actual processing window. The *NZC Unit* specifies all of the possible zero-crossings on an incoming signal. If the difference between two values with opposite signs is larger than a pre-defined threshold, a zero-crossing point is found. To eliminate false zero-crossing detection caused by environmental noise the selection of an appropriate threshold value is important [3]. The *NSSC Unit* determines the number of direction changes. In these cases the first or the last changes among the three consecutive values are larger than a pre-defined limit [3]. This limiting factor is required for the filtering of external environmental noise. Finally, the function of the *WL Unit* is to compute the length of the waveform, which is a characteristic feature of signal complexity [2].

In the EMG processing system, the training phase of the algorithm requires double precision floating-point vector-, and matrix operations. Unfortunately, in one hand, the built-in Neon SIMD engine in the ARM Cortex-A9 Core does not support double precision vector floating-point operations. On the other hand, scalar floating-point computing performance is not high enough to perform the required operators at acceptable speed. Due to these restrictions the *Vector Processor* [9] was used to perform and speed up these operations. The *Vector Processor* is built-up from a scratch-pad memory, several vector registers, and a floating-point adder and multiplier (because the majority of the required operations are multiplications and additions). It is capable of computing simple addition, multiplication and multiply-addition operations. Moreover, an addition and a multiplication operation can be computed in parallel when separate result registers are used. Matrices are stored in the scratch-pad memory, where the high-speed memory access by the *ARM Processor Core* is critical. The length of the vectors is limited by the depth of the vector registers, and they can be configured on-the-fly to adapt to the requirements of the classification algorithm.

V. Test results

Classification accuracy of the test implementations were compared and validated with previous, self-written MATLAB versions of the pattern recognition methods. In case of the standard algorithm, the average classification accuracy (about 90%) was reached across several test trials using different training and test data in each execution. As confirmed by the comparison, our design on the prototyping system provided the same functionality and efficiency as the reference implementation. Concerning the running times of the standard method, results show that even in the case of the computationally intensive training phase, using the custom designed firmware elements in the PL (*Preprocessor* and *Vector Processor*) enables the platform to reduce execution times with one order of magnitude (ranging from $\sim$100 ms to 5 s, depending on window shift value) compared to the raw performance of an ARM core. Furthermore, running times of signal classification remained under 0.5 ms in each case.

Similarly, in case of the modified algorithm, classification accuracy was consistent between the prototype system and the reference implementation, at about 60%. The reason for this low value is that the algorithm provides acceptable prediction of the actual wrist angle, and some temporal differences were present in the output (that cannot be compensated by the accuracy determining algorithm). The running times of the training and the classification phases show similar performance speed-up presented by the standard method.

VI. Summary and future work

In this study we proposed a prototyping platform for real-time measurement and analysis of myoelectric signals. Because of its modular structure, the platform can be used with various measurement and actuator devices to develop and test novel bio-signal processing techniques. The used Xilinx Zynq architecture provides real-time operation of even more sophisticated algorithms by implementing the computationally intensive steps directly on the Programmable Logic section of the FPGA SoC. To validate the feasibility of our platform, two specific classification algorithms were implemented for myoelectric prosthesis control. In the future, some novel approaches for MES processing will be implemented on the system and the development of the Measurement Unit will be finished. In addition, development of further measurement modules are being prepared, including custom designed board extensions with increased number of channels (preparing the system for possible EEG applications).

Acknowledgment

This research project was supported by the OTKA Grant No. K84267. The publication/research was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program' and TÁMOP-4.2.1.B-11/2/KMR-2011-0002.

References

[1] M. Johannes and J. Bigelow, "An overview of the developmental process for the modular prosthetic limb," *Johns Hopkins APL Technical Digest*, vol. 30, no. 3, pp. 207–216, 2011.

[2] B. Hudgins, P. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control." *IEEE transactions on bio-medical engineering*, vol. 40, no. 1, pp. 82–94, Jan. 1993.

[3] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control." *IEEE transactions on bio-medical engineering*, vol. 50, no. 7, pp. 848–54, Jul. 2003.

[4] L. J. Hargrove, K. Englehart, and B. Hudgins, "A comparison of surface and intramuscular myoelectric signal classification." *IEEE transactions on bio-medical engineering*, vol. 54, no. 5, pp. 847–53, May 2007.

[5] L. H. Smith, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "Determining the optimal window length for pattern recognition-based myoelectric control: balancing the competing effects of classification error and controller delay." *IEEE transactions on neural systems and rehabilitation engineering*, vol. 19, no. 2, pp. 186–92, Apr. 2011.

[6] E. J. Scheme, K. B. Englehart, and B. S. Hudgins, "Selective classification for improved robustness of myoelectric control under nonideal conditions." *IEEE transactions on bio-medical engineering*, vol. 58, no. 6, pp. 1698–705, Jun. 2011.

[7] B. J. Borbély and P. Szolgay, "Estimating the instantaneous wrist flexion angle from multi-channel surface EMG of forearm muscles," in *IEEE Biomedical Circuits and Systems Conference, BioCAS 2013*, Rotterdam, 2013.

[8] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2000.

[9] Z. Nagy, A. Kiss, A. Zarándy, T. Zsedrovits, B. Vanek, and T. Péni, "Volume and power optimized high-performance system for UAV collision avoidance," in *IEEE International Symposium on Circuit and Systems, ISCAS 2012*, Seoul, 2012.

Framework for examination the changes of arterial pulse signal waveform in time and space

Sándor Földi
(Supervisors: Dr. György Cserey, Dr. Péter Sótonyi)
folsa@digitus.itk.ppke.hu

***Abstract*—Pulse diagnostics is a non-invasive and painless diagnostic method. It can diagnose cardiovascular and different inner organ diseases. Its roots are from the Traditional Chinese Pulse Diagnostics (TCPD), which has a great disadvantage by its subjectivity. Therefore an objective automatized realization is required. From the '90s it is getting more and more popular to create an automatized TCPD device. Most of these studies concentrate only on the principles of TCPD and does not question them, does not try to describe it quantitatively. Therefore there are still a lot of unanswered questions about the background of pulse diagnostics. One of these is how the different pulse signal waveforms are formed, what information is behind the differences. This paper summarizes some ideas how these differences could be studied.**

***Keywords*-pulse diagnostics, signal processing, wavelet filter, pulse waveform**

I. Introduction

The principle of pulse diagnostics is that the vascular system is in connection with every organs of the human body, they are interacting and this interaction appears in the pulse signal [1]. This principle has not been proved scientifically yet, but there are some promising results in case of several inner organ diseases, like Duodenal Bulb Ulcer, pancreatitis and appendicitis [2]. Most of these studies relies on the Traditional Chinese Pulse Diagnoses. It is applied by the Chinese physicians more than 2000 years [3]. It defines three positions on both wrists, where the physicians examine the patient's pulse signal. The examination takes place by palpation with three fingers simultaneously on the three positions. During the examinations the physician applies different depth of touch which gives more information according to TCPD. The scientific background of this is still unknown, but recently it is started to be studied [4].

The arterial blood pressure signal is created by the heart. The cardiac cycle consist of two phase: the systole, when the heart contracts and pumps the blood from the ventricle to the aorta and the diastole, when the ventricles fill with blood. After the blood pumped out from the ventricle, it propagates along the vascular system. The typical waveform of arterial blood pressure is changing away from the heart. This is represented by Fig. 1. This can be referred as the pulse waveform changes in space. These change in space is occurred because of the arterial wall and blood interactions and the propagating and reflected wave properties [6]. Fig. 2. shows a typical pulse signal waveform measured at the wrist. This is a so called TRIWave signal, which is one of the healthy pulse waveforms.

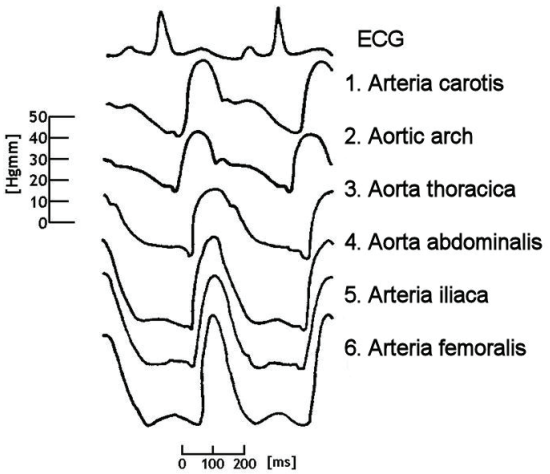


Fig. 1. This figure shows the typical blood pressure waveform on different points of the vascular system. This figure is edited from:[5]

It consists of three waves: the percussion wave, the tidal wave and the dicrotic wave. There is another frequent waveform, which is called DUOWave [7]. In case of DUOWaves, the tidal wave can not be seen, it merges with the percussion or the dicrotic wave.

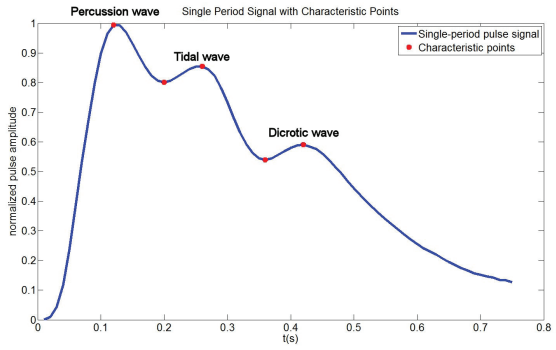


Fig. 2. The typical pulse signal waveform measured at the wrist. It consist of three waves: the percussion wave, which corresponds to the systolic peak, the dicrotic wave, which is corresponds to the closure of heart valves and the tidal wave, which is the reflected wave from the peripheral vascular system.

S. Földi, "Framework for examination the changes of arterial pulse signal waveform in time and space"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 87-90.

II. Method

A. Pulse signal measurement

To measure the pulse signal we use Optoforce 3D tactile force sensor [8]. This sensor is fixed over the radial artery on the wrist by velcro, as illustrated on Fig. 3. It can measure

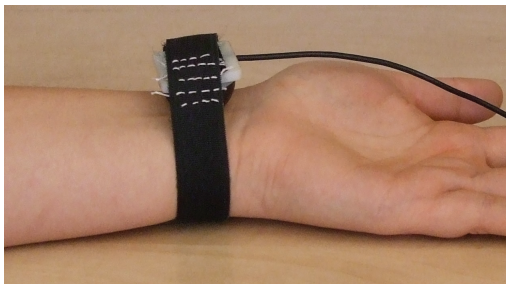


Fig. 3. The optoforce 3D tactile force sensor fixed on the wrist.

not only the force, but its direction too, which can help finding the right measuring position. It is extremely sensitive, it can measure really small forces, that is made it able to record the pulse signal at the wrist. Another advantage of this sensor is its material. It is not irritating, durable and easy to clean. During the examination the subject should not move, because that could not be filtered out yet. The length of examination is depend on the aim of it. To get the subject's typical pulse signal waveform only a few minutes (3-6) could be enough. To examine the time and space waveform changes, it requires much more time (at least 10 minutes, but 30 minutes would be more preferred), to be accurate enough. Therefore these examinations should take place in a calm, comfortable environment. It also requires more patience from the subject.

B. Signal processing

Signal processing consists of two main steps:

1) Signal filtering: The signal processing is realized with the Cascaded Adaptive Filter similar to the one presented in [9]. It can be separated into several steps. First the measured pulse signal is filtered by a Discrete Meyer wavelet filter. It can filter most of the noises caused by little hand movement or breath. After that the so called onset points must be found. These points are the start point of the systole, so in other words the onset point is the nearest minimum point left from the percussion peak. To find onset points, an adaptive windowing algorithm is applied, similar to presented in [10]. Using these onset points and the previously filtered signal, a spline estimation filter is used to remove the baseline wander drift.

2) Segmentation: After filtering the signal, it is segmented along its onset points. It creates the single-period pulse signals. These can be further examined. In case of diagnostic purpose, the similar single-period pulse signals are averaged. This averaged signal is suitable for feature extraction and after that for classification and diagnostic decision making. For examining the signal waveform changes in time or space, all of the single-period signals should be used. The signal processing method until the segmentation is summarized in Fig. 4.

C. Cross-correlation

For studying the changes in pulse signal waveform in time, cross correlation is used. This method can check the similarity of two different length time series, which is a great advantage in this case. Each single-period pulse signal is compared to the first measured signal. By this it can be examined, how permanent is the pulse signal waveform. Sometimes, even in healthy subjects, an outlier signal could occur, but in general this should not corrupt the results. The similarity measure between the first and the ith single-period signal is calculated as follows:

$$Similarity = max\Big(\frac{abs(xcorr(signal(1), signal(i)))}{norm(signal(1))norm(signal(i))}\Big), \quad (1)$$

where $signal(1)$ is the first single-period signal and $signal(i)$ is the ith.

III. Results

This subsection presents the results of pulse signal waveform similarity examinations. Fig. 5. shows the similarity values along a 30 minutes long measurement compared all single-period signals to the first single-period signal. According to

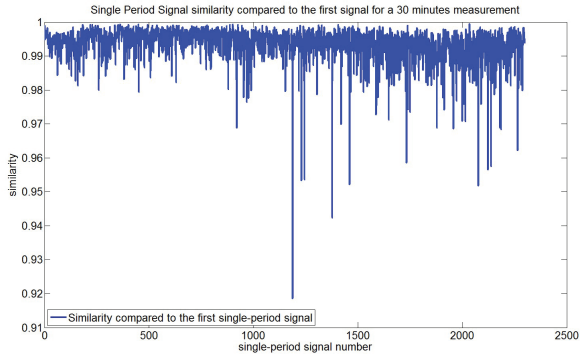


Fig. 5. Similarity between the first single-period pulse signal, and all the following ones during a 30 minutes long measurement.

this, the signals have high similarity. Fig. 6. shows the first, one with the highest similarity value, and the least similar single-period signals.

We have a database of 350 three minutes long pulse signals measured by the Optoforce 3D tactile force sensor [8]. Fig. 7. shows the maximum and the average similarity value for each measurement. Most of the cases the similarity value is high.

IV. Discussion

According to the results, we can state, that the pulse signal waveform does not change significantly in a given (maximum 30 minutes) time period. This allows us to assume that the averaged single-period pulse signal could be used for diagnostic purposes. But according to Fig. 6. this similarity value must be quiet high, at least around 0.98, because otherwise

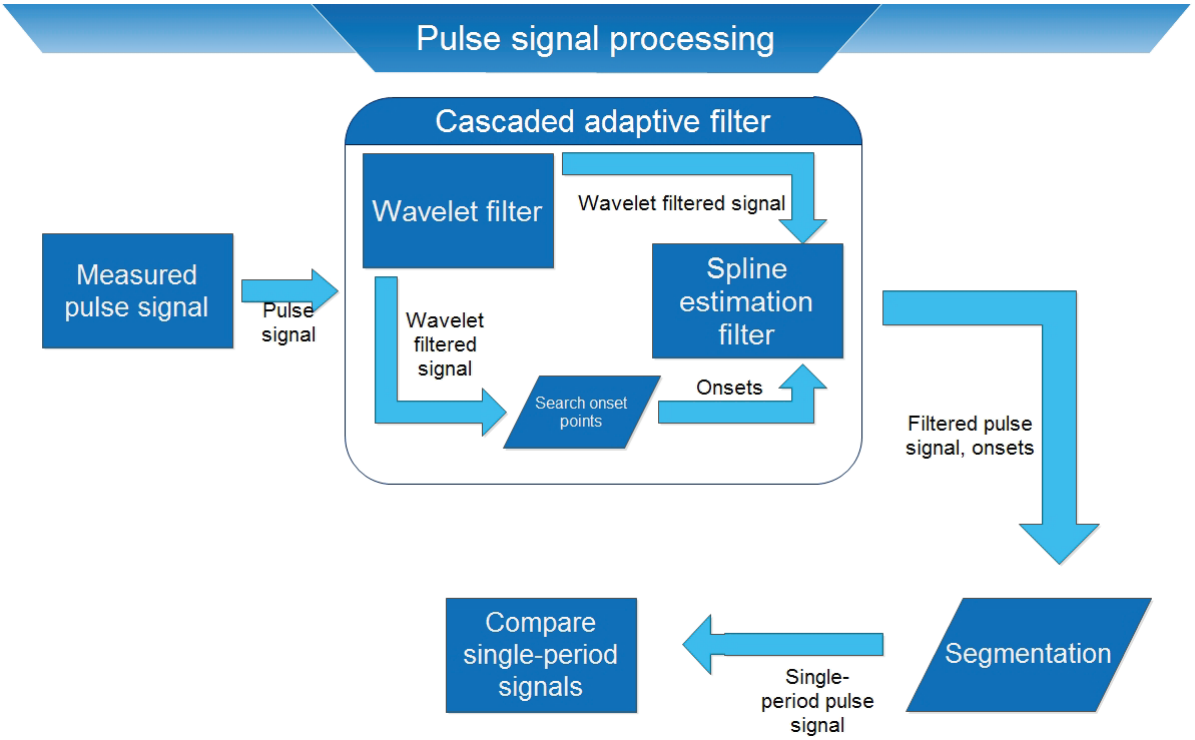


Fig. 4. Signal processing steps

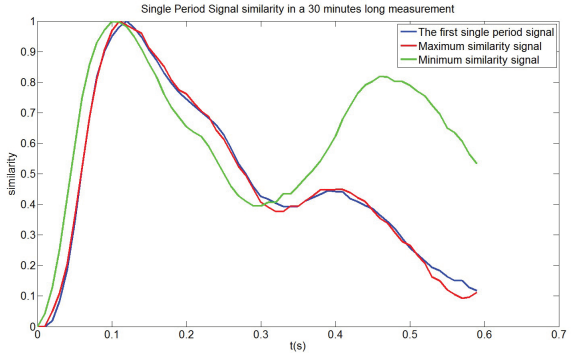


Fig. 6. The first single-period signal and the one with the highest similarity and another with the lowest during a 30 minutes long measurement.

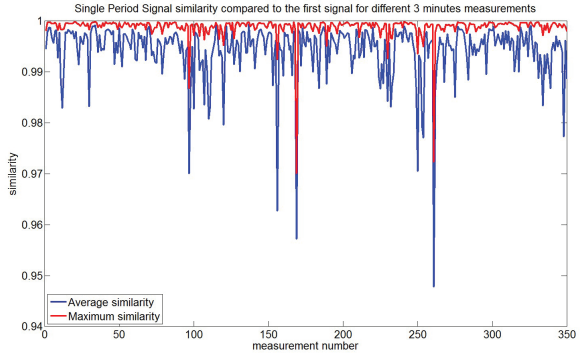


Fig. 7. Average and maximal similarity values in 350 three minutes long measurement.

the characteristic points could be really different which could lead to wrong conclusions, diagnosis.

Studying the 350 three minutes long measurements, the average similarity is above of 0.98 in 97.43 percent (341 out of 350), even 87.71 percent (307 out of 350) is above 0.99 average similarity value. The average standard deviation along the average similarities is only 0.0048. This means that in most cases the pulse signal waveform does not change much along the examination. That again confirm the assumption that the averaged single-period pulse signal could be used for feature extraction and therefore classification either. The reason of its importance is that the single-period signals usually not very smooth, which is a great problem in feature extraction. The averaged pulse signal is smooth in most of the cases, so feature extraction is easier and thanks to the similarity, it could be also more accurate.

V. Future plans

The similarity between two single-period pulse signals should be studied further. There need to be more long measurements on different subjects. There should also be some long examination on patients suffering in different kind of cardiovascular or inner organ diseases. There should be tested some other methods to calculate the similarity of different length time series, like dynamic time wrapping. The changes

of pulse signal waveform in time should be investigated also in site of diagnostic purposes. It could also be available to diagnose different kind of arrhythmias.

The differences between pulse signal characteristic in space will be also studied. The first positions will be at the wrist and the ankle. We have only one Optoforce sensor to take the pulse signals yet, so we could not make any comparison between the signals of these two positions, because it is important to record the signals simultaneously. So it will be further examined when we get the necessary equipment.

Acknowledgment

The author would like to thank Dr. György Cserey and Dr. Péter Sótonyi for their help and guide. The support of the Multidisciplinary Doctoral School at Faculty of Information Technology and Bionics (ITK - Research Faculty) at the Pázmány Péter Catholic University (PPCU - University of National Excellence) and the Városmajori Szív- és Érgyógyászati Klinika is also gratefully acknowledged. The technical background is acknowledged to Optoforce Ltd. The author would also like to thank Flóra Zieger for the support and help in the project.

References

[1] Bilton, Karen, Leon Hammer, and Chris Zaslawski, *Contemporary Chinese Pulse Diagnosis: A Modern Interpretation of an Ancient and Traditional Method*, Journal of Acupuncture and Meridian Studies, 2013.

[2] Sun, Yunlian, *Computerized wrist pulse signal diagnosis using KPCA*, Medical Biometrics, Springer Berlin Heidelberg, 2010. 334-343.

[3] B. Flaws, *The Secret of Chinese Pulse Diagnosis*, Blue Poppy Press, 1995.

[4] de Sá Ferreira, Arthur, and Agnaldo José Lopes, *Pulse waveform analysis as a bridge between pulse examination in Chinese medicine and cardiology*, Chinese journal of integrative medicine 19.4, 2013. 307-314.

[5] Emil Monos, *Hemodinamika: a vérkeringés dinamikája*, Semmelweis Egyetem, Egyetemi jegyzet, 2001.

[6] Nichols, Wilmer, Michael O'Rourke, and Charalambos Vlachopoulos, *McDonald's blood flow in arteries: theoretical, experimental and clinical principles*, CRC Press, 2011.

[7] Pei-Yong, Zhang, and Wang Hui-Yan, *A framework for automatic time-domain characteristic parameters extraction of human pulse signals*, Eurasip Journal on advances in signal Processing, 2008.

[8] Optoforce Home Page, *http://optoforce.com/*

[9] Xu, Lisheng, *Baseline wander correction in pulse waveforms using wavelet-based cascaded adaptive filter*, Computers in Biology and Medicine 37.5 (2007): 716-731.

[10] Chen, Yinghui, *Wrist pulse signal diagnosis using modified Gaussian models and Fuzzy C-Means classification*, Medical engineering and physics 31.10 (2009): 1283-1289.

Using contour geometry as a merging cue in oversegmented images

Gelencsér-Horváth Anna
(Supervisor: Dr. Karacs Kristóf)
horvath.anna@itk.ppke.hu

***Abstract*—Merging regions in over-segmented image using color similarity information is often not adequate if adjacent segment values are close to each other in the colorspace. We propose a method to improve the separation of distinct image components based on the geometry of the cluster contours. The chains built of adequate contour-components support the merging strategy by indicating object boundaries. In this article a proof-of-concept implementation is introduced.**

***Keywords*-over segmentation; merging; connect contour-components;**

I. Introduction and background

In image understanding segmentation is one of the most useful tools. On one hand background-foreground segmentation, or even distinguishing among foreground elements may often be necessary.

The output of segmentation algorithms is an oversegmented image, which means that we have more regions per object which should be merged. In this work we use over segmented images which were generated with mean shift (MS) algorithm. Mean shift is an iterative, non-parametric kernel based method [1]. As a first step a feature space is defined for the image. The MS kernel is convolved with the feature space elements and locates dense regions (mean points) as a result of convergence.

As a result of the output image is over-segmented, and requires further processing to enable us detecting and/or recognizing objects. There are several algorithms for merging, based on either statistical properties [3]–[5], similarity of adjacent regions [6], or graph properties [7]–[9].

However in some cases color information and similarity of the regions may be insufficient for region merging. For the picture in Figure 1 the trees between the roofs are segmented to regions with very similar color to the regions of the adjacent roof. Therefore during the merging it is coupled with the regions of the roof. In Figure 2, we may see a contour of the roof composed by contour component of the regions. The geometry of these contour components may indicate a different merging decision improving the final segmented output.

Fig. 1. Original image before segmentation. We focus on the image area containing the roofs and the trees between them. Due to similarity trees are merged to the same cluster as the roof of the building on the right.

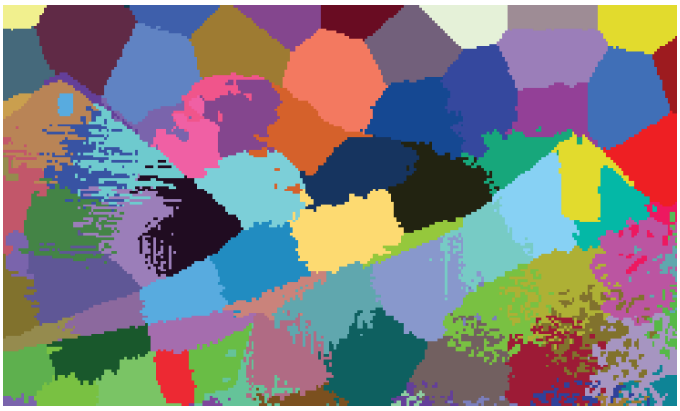


Fig. 2. Output of the segmentation algorithm. Each region of the oversegmented image has unique color corresponding to the cluster ID (index image).

II. Introduction to the proposed algorithm

The algorithm runs on an over segmented image. The aim is to detect the contours which are constructed from the region contours *well-defined*. We focus on superpixels contructed as 2x2 pixels, with exactly three unique indexes (cluster values), and define them as T-candidates. Edge between two T-candidates is defined as the pixel path between them, on the boundary of the region. A pixel path has pixels from a single cluster, which we define to be the cluster with the smaller index value. We define T-type connections among T-candidates by assigning a straightness value to each edge pair of the candidate based on the difference on their angle and 180°. We build *chains* from these connections (T-points) along the edges if possible using the assigned straightness values. The longer the chain is, the more probable it is to separate the region set on one side from the region set on the other side. The code is written in MATLAB.

A. Gelencsér-Horváth, "Using contour geometry as a merging cue in oversegmented images"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 91-94.

A. Methods

We detect T points (2x2 size subpixels of exactly three different index values) with a kernel - which can be adapted to parallel architecture easily. In the 3x3 size kernel lets define left-upper, right-upper, left-under, right-under superpixels as the 2x2 pixel size part, where all of them contains the middle element. If the middle pixel is part of a T-point, then there are exactly three indexes contained by the kernel window around it. Therefore we multiply the number of unique elements in the left/right upper/under superpixels: if the result of this four element multiplication is divisible with 3 but not divisible with 9 (exactly one of the 2x2-s have three colors, all others exactly two) then the middle element is set to 1 else 0 (see illustration in Figure**??**). A 3x3 kernel filters the upper left pixel for each T-point, these are inserted into an array containing point objects with coordinates (4 x 2), unique ID, index values corresponding to the coordinates.

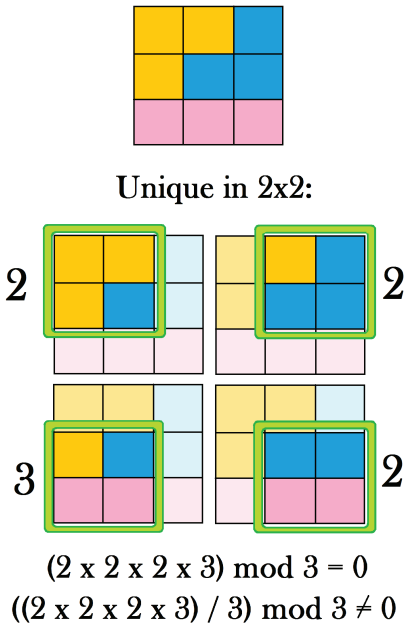


Fig. 3. For each 3x3 size kernel we count unique indexes for all 2x2 subkernels containing the middle element to detect T-candidates. The middle element is recognized as part of a candidate superpixel if exactly one of the subkernels have three, other exactly two unique indexes.

From each T-point we follow each region border until the next T point in the region with the smaller index value. If represented as a graph where points are vertices, these contour sections are the edges. Each edge is detected twice (starting from both ends), but not exactly with the same trajectory due to the possible negative and positive turns – as the step preference along the two-color border is up, left, right, down, up-left, up-right, down-right, down-left.

Each edge is inserted into an array of edge objects, where edge has the following properties: unique ID, ID of points of both sides, and the pixel route between these points.

We define a minimum edge length for the sake of efficiency (presently set to 3 but it may be defined in percentage of the neighbourhood size in angle estimation), and all sorter edges are filtered out. To remove duplicates from edge array we match the neighbourhood size (*N*) long middle parts of the routes in a loop and choose each contour section only once.

There are three main edge types as illustrated in Figure 4. Most edges are regular (type I.), still the proposed algorithm deal with type II. and III. as well.

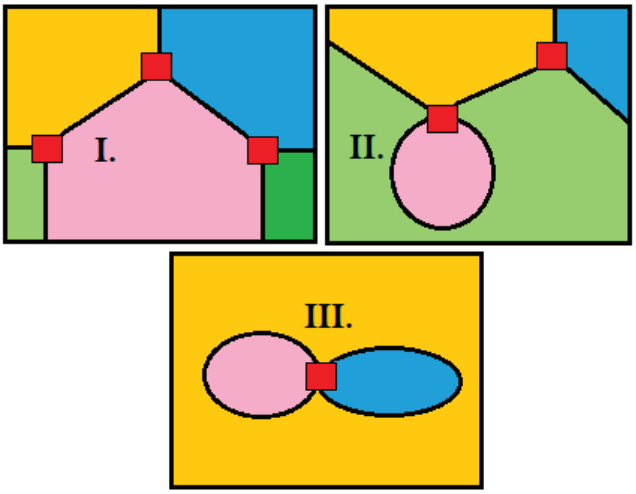


Fig. 4. Main types of edges in oversegmented image. Red squares denote T-point dedicates: at intersection of three regions of different index value. The I. type is regular, II. and III. are special cases.

We loop in the point array of T points, and set save the edge ID-s that *meet* at that point. For each edge we use the first *N* pixels to count its angle to the horizontal unit vector, than based on the magnitude and sign we calculate the three relative angles: each between two egde section. We save this data for each point object in the array.

For each angle value we calculate a straightness value that measures hos close it is to 180°. At a single T point two angles may be close 180° if the third is rather small (i.e. around 5°), therefore the straightness value f_{180} is defined zero for angles less then 120 and more then 240. Lets define δ (see Eq.1).

$$\delta = 0.01 \times (currentAngle - 180°) \tag{1}$$

The straightness value f_{180} is calculated according to equation 2. if δ < 0.03, and according to equation3. if $0.06 < \delta$ < 0.03 where *param* is a predefined parameter to set function gradient.

$$f_{180} = 1 - \delta^2 \tag{2}$$

$$f_{180} = (\delta - 0.06^2) \times param \tag{3}$$

We sort points based on the straightness value: the closer an angle is to 180°, the sooner the corresponding T-point is examined for chaining.

We build *chains* from the contour sections using edge-point-edge structures where also the straightness value of the angle

of the two edges at the T-point is known. Therefore each chain has an edge-point-edge-...-point-edge array with IDs of adjacent points and edges in the chain, and a straightness array where the value at index i corresponds to point ID at index $2 \times i$ in the edge-point-edge array.

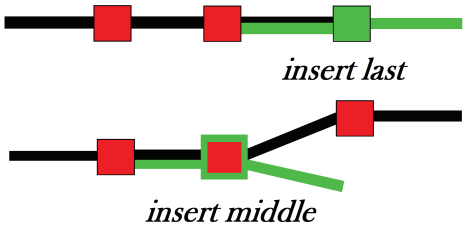


Fig. 5. Insert new element into chain array - cases. The new structure which will be inserted to the chain array in the current iteration is denoted with green (two edges and a point). Points already in a chain are denoted with red, edges are denoted with black.

The basis cases of chain element insertion are shown in Figure 5. The actual edge-point-edge structure (actual input for insertion) is denoted by green color. If neither edge ID is contained in any of the chains, then the structure is inserted to the chain array as a new chain. If at least one of the edges already exists in any of the chains, then based on the type we perform one of the following operations:

1) Insert last shall be applied if one edge is exactly the last edge on either side of a chain and the point is not included.
2) Insert middle is applied if either edge of the input structure and also the point is included in a chain. We insert a new chain to the array built from the copy of the common part of the chain appended to the point and the non-fitting edge of the input structure.
3) We perform concatenation when both edges are already present among the chain array elements. Either the same or different chains contain these edges, the containing chain(s) is removed from the array, and the appended one is inserted as a new one.

Chain size can be an indicator in merging decisions: a longer chain suggest a more reliable contour, which means merging regions along both sides, but not across may give a more proper region contour if color information indicates otherwise.

III. Results and Discussion

Since the length of the chains has an impact on the likelihood of merging we use a threshold on the chain size.

Chains with a length of at least three are shown on Figure 6 and 7. Figure 6 shows the relevant part of the original image, where the roof and the trees shall be separated. See Figure 7 for the difference in mergeing decision: the colored areas are constructed without using contour geometry, while chains indicate that the regions along the different sides shall be labelled different. This allows for the separation of the roof and the trees.

Fig. 6. Result on original image.

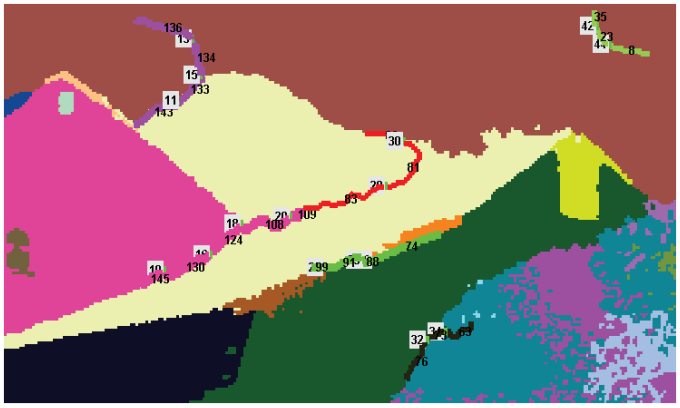


Fig. 7. Result of T-point based method on the merged image from color similarity.

However, the parameters and functions shall be revised, on this sample image we can see (by now only as a visual feedback) how the algorithm can improve the region clustering of the image. As a next step (in addition to the improvement of the code to enable more and longer chains if feasible), we will define evaluation method that can be measured. We will measure the quality of the proposed method in merging on a large testset. In the future we would like to embed our method into a merging algorithm that currently uses cluster neighbourhood information and color similarity [2]. Geometry along color information shall improve merging. We consider implementing learning the following: fuzzy function of angles, neighbourhood size of points, and thresholding on fuzzy value when chaining.

References

[1] D. Comaniciu and P. Meer. *Mean Shift: A Robust Approach Toward Feature Space Analysis* IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(5):603–619, 2002

[2] B. Varga and K. Karacs *Towards a balanced trade-off between speed and accuracy in unsupervised data-driven image segmentation* Machine vision and applications, 24(6), 1267-1294., 2013.

[3] R. Nock and F. Nielsen *Statistic region merging* IEEE Trans. on Pattern Analysis and Machine Intelligence, vol 26, pages 1452-1458, 2004.

[4] F. Calderero, F. Marques *Region merging techniques using information theory statistical measures* IEEE Transactions on Image Processing. Volume: 19 , Issue: 6. page(s): 1567-1586, 2010.

[5] F. Calderero, F. Marques *General region merging approaches based on information theory statistical measures* The 15th IEEE International Conference on Image Processing (ICIP). pp: 3016-3019, 2008.
[6] J. Ning, L. Zhang, D. Zhang and C. Wu *Interactive Image Segmentation by Maximal Similarity based Region Merging* Pattern Recognition, vol. 43, pp. 445-456, Feb, 2010
[7] P.F. Felzenszwalb and D.P. Huttenlocher *Efficient Graph-Based Image Segmentation* International Journal of Computer Vision. Vol. 59, Number 2, September 2004.
[8] B. Peng, L. Zhang and J. Yang *Iterated Graph Cuts for Image Segmentation* In Asian Conference on Computer Vision, 2009
[9] H. Liu, Q. Guo, M, Xu, I. Shen *Fast image segmentation using region merging with a k-Nearest Neighbor graph* IEEE Conference on Cybernetics and Intelligent Systems. Page(s): 179-184, 2008.

Design of Brain Neuron Signal Recording System on UMC 0.18μ Technology

László Tamás Kozák
(Supervisor: Péter Földesy)
kozla@digitus.itk.ppke.hu

***Abstract*—The ability to monitor the simultaneous electrical activity of multiple neurons in the brain enables a wide range of scientific and clinical endeavors. The main challenge of designing such a recording system is the limited power and area. This paper shows a possible alternative how the circuit size can be improved to reduce the device cost while the functionality remains the same.**

I. Introduction

In the last semester the addressed work was to determine the analog front-end schematic of a biomedical ASIC device. The task of this device is to process the signals of those brain neurons that are selected with a MEMS electrode array and transmit the preprocessed data wirelessly to the corresponding medical application. This document contains only the working and/or integrated analog parts that fulfill the targeted specification (Table I).

DC spec.		min	typ	max
T	Operation temperature	30^{o}C	36^{o}C	42^{o}C
VDDA	3.3V domain	2.7V	3.3V	3.65V
VDD	1.8V LDO output	1.75V	1.8V	1.85V
IDD	LNA on current	-	-	10μA
AC spec.				
Gain	LNA gain	60dB	-	-
CMRR	LNA CMRR	90dB	-	-
TR	ADC data-rate	15ksps	-	-

Table I: More important specifications

The selected technology is the 0.18μm mixed-mode twin-well process of the UMC that contains devices for 3.3V and 1.8V domains. The typical supply level was chosen as 3.3V but the largest part of the device works with stable 1.8V level that is generated by a Low Drop Out regulator (LDO). Under normal circumstances the body temperature cannot change between wide ranges that was exploited in case of those circuits where the accuracy is one of the most important factor. Since there is no layout and/or prototype the real behavior cannot be validated by measurement therefore the functionality is proven by only spice corner and/or monte-carlo simulations. The simulated corners that are based on the specification listed in Table II that means 64 different circumstances beside the nominal one.

II. Block Diagram

For reduced cost all the necessary blocks that are responsible for generating the proper digital and analogue control signals should be put onto the same die. These signals can have enable, reset functions or they can be different references. The Fig.1. shows the block diagram of the device that is one possible solution [1].

name	value
process	ss,ff,snfp,fnsp
resistor	min,max
capacitor	min,max
temperature	30^{o}C,42^{o}C
vdd3v	2.7V,3.6V

Table II: Table of 64 corners

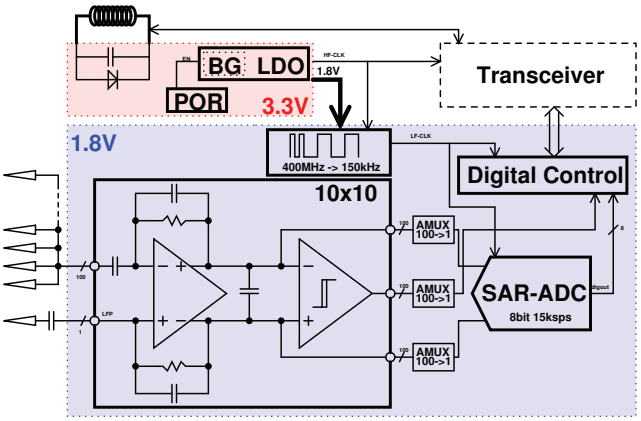


Figure 1: Block diagram of the device

Since the final device will be fed through transformer coupling and the user does not have the opportunity the switch on the device manually the device needs a power-on-reset block (POR) that enables the reference if the 3.3V rail level is within the specified range. The supply block is in the 3.3V domain and responsible for the bias and the stable 1.8V rail level. It contains a CMOS 2-way rectifier, a bandgap reference and an LDO. The amplifier array has the LNA itself and the comparator that has adjustable hysteresis. The 8bit successive approximation ADC (SAR-ADC) with 15ksps data rate converts those outputs of the amplifier array that are selected by the analog multiplexers. The necessary 150kHz clock signal for the ADC and the control is generated from the transformer feed-in frequency. In this way the built-in oscillator is not necessary that reduces the complexity and the cost.

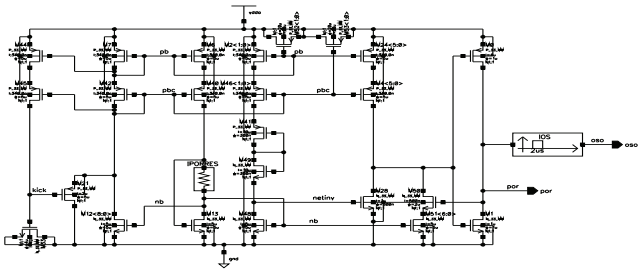
L. T. Kozák, “Design of Brain Neuron Signal Recording System on UMC 0.18µ Technology”
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 95-98.

III. Design

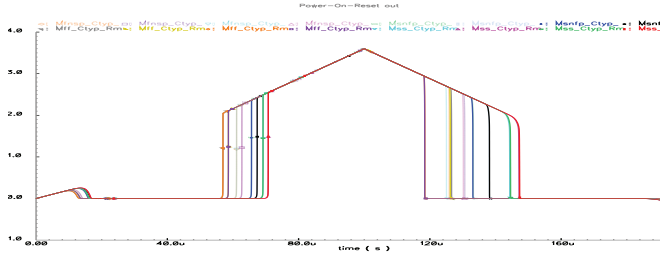
The largest part of the work was made with Cadence design tools that include the Virtuoso Schematic Editor L, while the simulation parameters were set up with Analog Design Environment where the spectre was the applied simulator.

A. POR

The ultimate goal is to make device work with wireless powering to reduce the risk of infection caused by open path into the inner parts of the body. Since there is no opportunity to use separate enable signal to switch the device on, a built-in POR block is necessary that determines if the supply level is between the specified range or not. The POR schematic (Fig.2.a) and the transient behavior of its output (Fig.2.b) can be seen below.



(a)



(b)

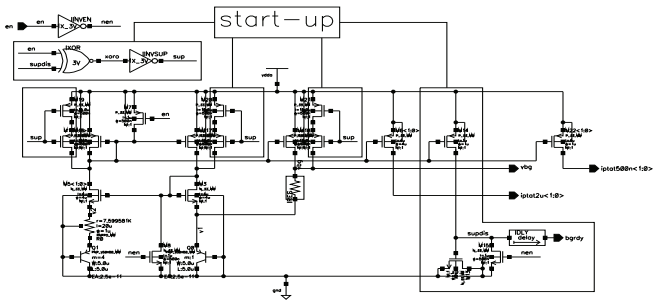
Figure 2: POR schematic (a) and its output (b)

The 3.3V rail is risen from 0V up to 3.6V in 100μs and than back to 0V in 100μs as well. The lowest supply level at the rising edge of the POR out is around 2V that is enough for the bandgap to be started up. The simulation result also shows that the POR output gets back to low level when the rail level is close to its specified lower bound. There is only one corner where the POR switches off the bandgap at around 2.8V.

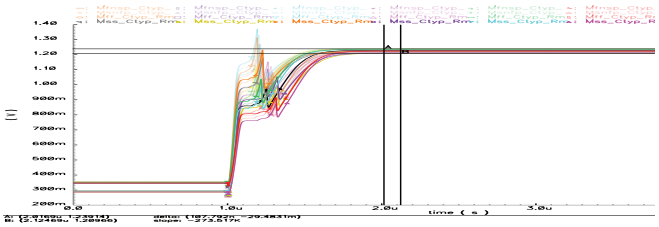
B. Bandgap, LDO

The supply is built up from a CMOS rectifier [4], a band-gap reference (BG) that generates stable 1.22V reference for the LDO and different PTAT currents. The BG schematic is in Fig.3.a and its transient output is in Fig.3.b. In Fig.3.b the macro is switched on at 1μs so the worst case start-up time is less than 1.12 μs and the output error is less than 30mV.

Fig.4.a shows LDO schematic and the transient behavior of its 1.8V level output can be seen in Fig.4.b where the error



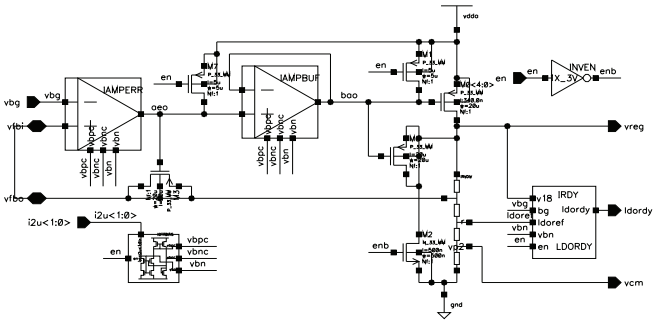
(a)



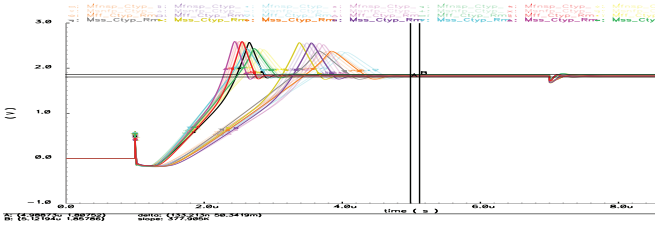
(b)

Figure 3: Bandgap schematic (a) and its output (b)

is around 50mV that is determined mainly by the BG output error. The enable input goes high at 1μs like in case of the BG that means less than 4μs start-up time maximum. Beside the 1.8V level output the LDO generates LDORDY output also that goes high when the LDO output reaches the operation level. Fig.4.b also shows the LDO output response for a 100μA load current step at 7μs where the worst case voltage drop is around 100mV and the LDO rise back its output within 200ns.



(a)



(b)

Figure 4: LDO schematic (a) and its output (b)

Those block whose current consumption exceeds this $100\mu A$ limit have successive starting method when their sub-blocks were started with different delays. In this way the size and power consumption of the LDO can be smaller and that few ns delay is not critical respect to the general few times 10us start-up time of the amplifiers that have large time constant due to their low frequency operation.

C. SAR-ADC

The ADC design is always a critical point of a project. This macro has more important parameters that has to be considered such like absolute and relative accuracy, data rate, power consumption and size. The type of the selected ADC is determined by the actual design. Because of the speed and accuracy requirement in [1] and [3] I chose successive approximation ADC type whose top schematic is in Fig.5. Since the capacitors in the LNA array take large area I used the R-2R resistor ladder ADC topology instead of capacitor type. Additionally the resistors belong to the front-end-of-line (FEOL) step and the frequency of the application is not too high so they can be placed below the MIM type capacitors without significant parasitic effects. One drawback of the resistor ladder is its noise and larger consumption.

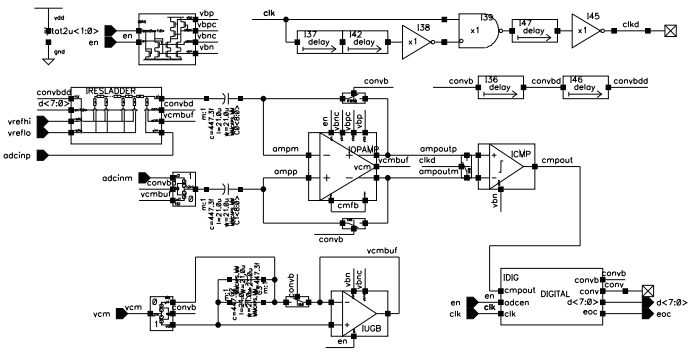
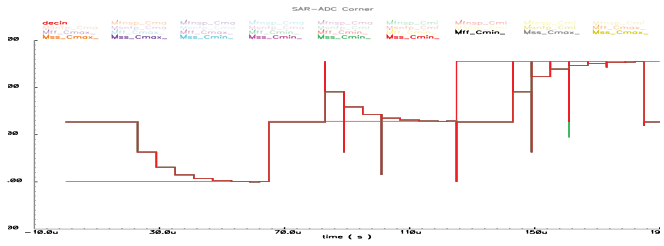
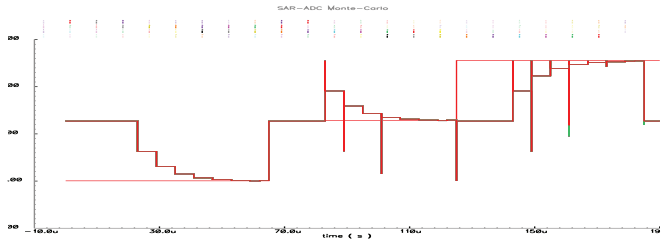


Figure 5: Schematic of the SAR-ADC

One conversion period takes 10 clock periods. This means 15ksps data rate with the 150kHz clock frequency that is generated by the frequency divider that makes 150kHz from the 400MHz reference frequency. The testing of the 2^8 possible input levels would be extremely long therefore the relative accuracy and the dynamic behavior cannot be checked by spice simulations. The schematics were tested with only 3 different input levels that include the min-max and middle cases. The Fig.6. shows the SAR algorithm of the ADC where the straight line represents the actual input and the other curves represent the ADC output decimally and how these corner dependent decimal values approximate the input successively. The corner simulation result with the 15ksps data rate and 8bit absolute accuracy can be seen in Fig.6.a while the Fig.6.b shows the mismatch simulation result of typical case where 3*30*3 runs did not show error.



(a)

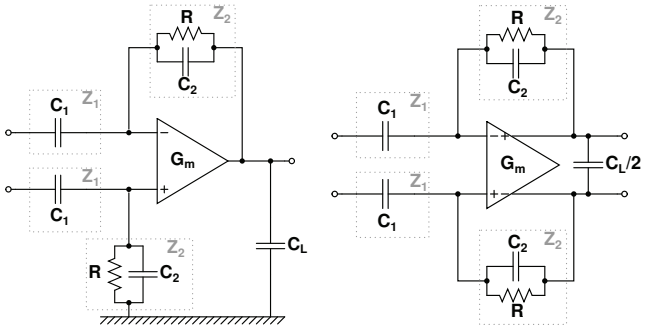


(b)

Figure 6: ADC digital output of the 64+1 corners (a) and the Monte-Carlo of 3*30*3 inputs (b)

D. LNA

The Fig.7.a shows the original topology of the LNA [1] [2] [3] that has asymmetric topology. The gain of this structure can be calculated as (1). The equivalent fully differential structure of the original asymmatric version can be seen in Fig.7.b.



(a) (b)

Figure 7: Schematic of the original (a) and the proposed (b) LNA

The asymmetric version needs large capacitor at both input terminals to cut of the DC level. In case of the proposed differential version the LFP inputs of all the amplifier can use the same C_1 capacitor. This cannot be done with the asymmetric topology since the parasitic C_{gs} of 100 NMOS input transistors would change (1).

$$V_o/(V_p - V_m) = Z_2/Z_1 \tag{1}$$

The transfer function of the amplifier can be expressed as (2) where the input parasitic capacitances are neglected.

$$H(s) = \frac{sC_1R}{1+sC_2R} \cdot \frac{sC_1C_2 + G_mC_2}{sC_1C_2 + G_mC_2 + sC_L(C_1+C_2)} \quad (2)$$

The other effect of the common C_1 capacitor is the lower first zero frequency that can be seen in Fig.7 of [1] that is shown here for your convenience

$$f_{zl} = \frac{C_2}{C_1 + C_2 + C_{in}} \cdot \frac{1}{2\pi RC_2} \quad (3)$$

where C_{in} is the input parasitic capacitance of the LNA. The (3) shows that if the parasitic of all the 100 amplifiers are connected in parallel the f_{zl} slides lower. Only one C_1 capacitor on LFP side means approximately 50% capacitor area saving. The lower zero frequency also reduces the 1/f noise at the output. The drawback of this solution is the additional CMFB part in the LNA whose size can be neglected beside of $100 \cdot C_1$.

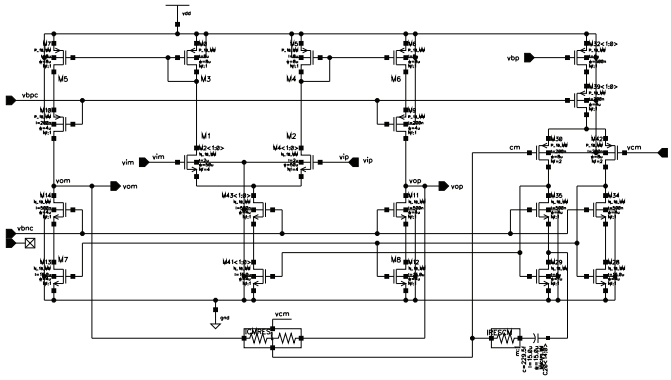


Figure 8: Schematic of the LNA

The Fig.8. shows the proposed LNA schematic with the CMFB part at the right and Fig.9. shows the AC corner result of the differential output where the midband gain is set as $A = C_1/C_2 = 100$.

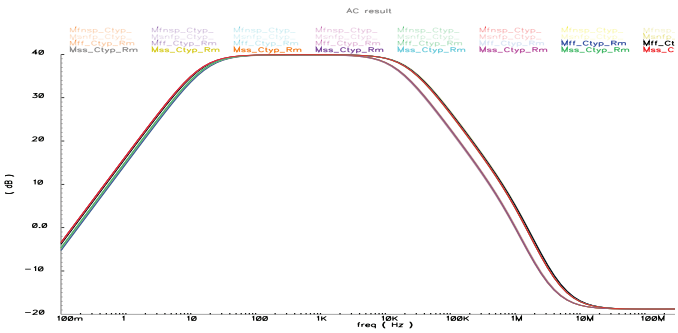


Figure 9: AC simulation result of the proposed topology

1) LNA noise: Since the amplitude of the neural signal is on the order of 100μV the noise analysis is very important. The input referred noise of the proposed LNA can be expressed in almost the same way like in [1]. The effect of the cascode transistors at the output can be neglected just like the noise that comes from the 1.8V rail through the LDO. The input referred noise of the amplifier can be expressed as

$$\overline{V_{nia}^2} = \left(\frac{C_1 + C_2 + C_{in}}{C_1}\right)^2 \frac{8kT}{3g_{m1}} \left(1 + 2\frac{g_{m3}}{g_{m1}} + \frac{g_{m7}}{g_{m1}}\right) \Delta f \quad (4)$$

The $\Delta f = (\pi/2) f_h$ is the brick wall frequency where $f_h = G_m/(2\pi C_L(C_1/C_2))$. The equation (13) of [1] expresses the area as the function of MIM capacitors per amplifier. In case of fully differential topology this equation is the following.

$$Area = \frac{2}{3}\frac{kT}{V_{ni,rms}^2 AC'} + (A+1)\frac{C_2}{C'} + \frac{C_2}{C'} \quad (5)$$

For varying midband gain (5) has a minimum at

$$A = \frac{1}{V_{ni,rms}}\sqrt{\frac{2}{3}\frac{kT}{C_2}} \quad (6)$$

that is the same as (14) in [1].

IV. Summary

Although the post layout simulation results can make modifications necessary on the schematic the largest part of the addressed analog work is done. The next setps can be the design of the wireless powering with the CMOS rectifier based on [4] and the transceiver block with the corresponding digital control and baseband. The layout design of the analog part also can be started simultaneously. Due to the significant size and complexity of this work the cooperation of a digital and more layout designers is not only possible but also necessary that can give an exciting topic and opportunity of their research.

References

[1] R.R. Harrison, "The Design of Integrated Circuits to Observe Brain Activity," *Proceedings of the IEEE*, Vol.96, pp.1203-1216 No. 7, July 2008.

[2] R.R. Harrison, "A Low-Power Integrated Circuit for a Wireless 100-Elecrode Neural Recording System," *IEEE Journal of Solid-State Circuit*, Vol.42, pp.123-133 No. 1, January 2007.

[3] R.R. Harrison, "A Low-Power Low-Noise CMOS Amplifier for Neural Recording Applications," *IEEE Journal of Solid-State Circuit*, Vol.38, pp.958-965 No. 6, June 2003.

[4] M. Ghovanloo, "Fully Integrated Wideband High-Current Rectifiers for Inductively Powered Deviced," *IEEE Journal of Solid-State Circuit*, Vol.39, pp.1976-1984 No. 11, November 2004.

[5] B. Razavi, "Design of Analog CMOS INtegrated Circuits," pp.377-404, McGraw Hill International Edition, 2001.

[6] J. Huijsing, "Operational Amplifiers: Theory and Design," in Fully Differential Operational Amplifiers pp.365-386, The Kluwer Academic Publishers, 2001.

Compact, low-power, FPGA-based multi-camera vision system for real-time image processing

Mate Nemeth
(Supervisor: Dr. Akos Zarandy)
nemeth.mate@itk.ppke.hu

***Abstract*—Autonomous maneuvering of UAVs (Unmanned Aerial Vehicles) in civil airspace is forbidden by flight authorities [1], [2], because of various safety shortcomings. One of the most important missing capability of a UAV is the lack of collision warning and avoidance system. This paper presents a vision based on-board development platform for UAVs (FPGA board, interface board, miniature camera system). The system is capable of receiving the video-stream of maximum five cameras, using LVDS (Low Voltage Differential Signaling) technology. Our goal is to develop a massively parallel, real-time image processing FPGA architecture based on this platform, which realizes the task of collision warning from visual sensor data.**

***Keywords*-real-time vision system, UAV, multi-camera, FPGA, multi-core processing, collision warning, LVDS**

I. Introduction

An ongoing project in MTA-SZTAKI is aiming at developing a compact UAV, capable of autonomous maneuvering, even in highly utilized airspace, due to its on-board non-cooperative collision warning and avoidance system [3]. According to Federal Aviation Administration (FAA) [1] and Eurocontrol [2], the integration of UAVs to utilized civil airspace requires that UAVs are capable of performing at an equivalent level of safety (ELOS) to that of manned aircraft [5], [6], to lower the probability of mid-air collisions. Collision avoidance has been identified as one of the most important issues of the above mentioned integration by international flight organizations.

Several solutions exist in the field of collision warning: GPS based active radio transponders (Traffic Alert and Collision Avoidance System, TCAS), radars, or vision based systems. It turns out from the literature [7], [8], that those systems using visual remote aircraft sensor for the detection of potentially dangerous intruder aircraft, are mainly using PCs for the calculations. To fit on a compact, low-weight UAV, the on-board systems have to meet various power consumption, weight and size considerations. As a result, a "miniature" embedded system should be used for the task, which seems to be a novelty according to the literature. The initial system [4] was presented at ICVS (International Conference on Computer Vision Systems), and an other publication [3], related to the vision system, is submitted to JRTIP (Journal of Real-Time Image Processing). Our goal is to develop a hardware platform, which can be applied on a suitable UAV. This could be used in real operating conditions, as a framework for the experimental validation of later architectures and algorithms for low-cost, low-power, small-form-factor, mobile image-processing applications.

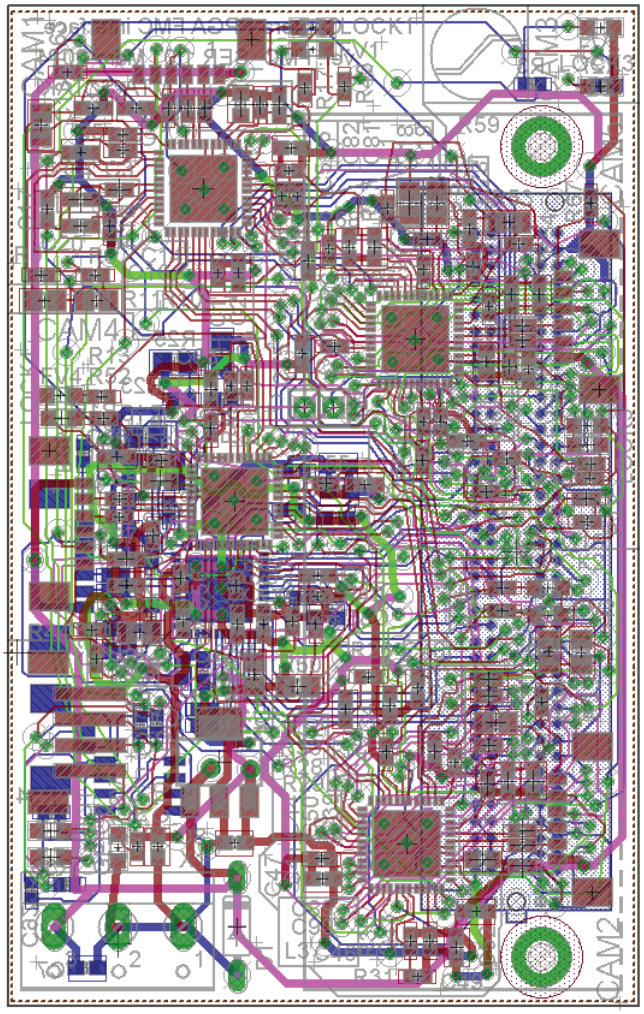


Fig. 1. Interface board layout, with five deserializers. (dimensions: $7.3cm \times 4.6cm$)

The detection of intruder aircraft requires a permanent monitoring of a sufficiently large field of view ($220° \times 70°$ in our case), with high resolution cameras, to satisfy the separation volume defined by flight safety requirements in any given situation. The following sections describe the hardware elements of the sense-and-avoid development platform.

M. Németh, "Compact, low-power, FPGA-based multi-camera vision system for real-time image processing" in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds. Faculty of Information Technology and Bionics, Pázmány Péter Catholic University. Budapest, Hungary: Pázmány University ePress, 2014, pp. 99-102.

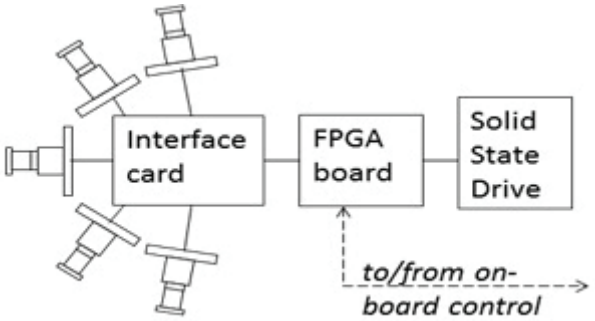


Fig. 2. Block diagram of the system.

II. System Architecture

As mentioned previously, an essential part of the project has been designed in the form of a development platform. It functions as a framework to the evaluation of the computational architectures to be designed on the reconfigurable fabric, even in a hardware-in-a-loop manner. The framework consists of four elements:

- **FPGA board**, with FMC-LPC connector
- **Camera modules** (up to five), with the Texas FPD-Link III protocol
- **Interface card** between the FPGA board and the cameras, which re-converts serial input data into parallel data lines
- **SSD drive**, to store the images

The block diagram of the system is shown in Figure 2. The following subsections describe the design considerations and features of each component.

A. FPGA Board

Instead of designing a new one, we decided to use an existing FPGA development board. However the development of a small-form-factor FPGA board is planned, which fits the later described interface board in size. Three optional choices emerged during the development process: Xilinx SP605, LX150T, and Digilent ZedBoard. The first two contains Xilinx series 6 Spartan FPGAs, while the latter comes with a mounted 7 series Zynq SoC, with two physical ARM Cortex-A9 microprocessors attached to the reconfigurable fabric, which allows more flexibility in the case of hardware-software partitioning. The interface card has been designed to be compatible with all three platforms. In the later stages of the development, the ZedBoard seems to be the preferable choice, because of the above mentioned flexibility, and the fact, that new generation Xilinx HLS [9] (High Level Synthesis) design flow supports only the 7 series FPGAs. The reason why the series 6 boards are considered as a real option, is that an image processing architecture has already been designed [3] in SZTAKI Computational Optical Sensing and Processing Laboratory, which we want to be compatible with.

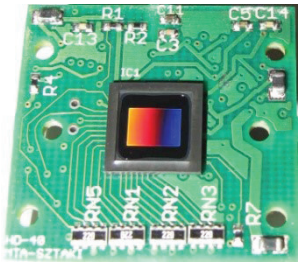


Fig. 3. 1.2 Mpixel camera module, with the MT9M021 sensor.

B. Camera [3]

The key component of a special purpose vision system is the camera. We decided to use multiple miniature cameras with S-mount (M12) lenses, instead of one high resolution camera with a low distortion ultra-wide angle optics, because of weight, size and especially financial considerations. We named two fundamental criteria against the cameras to be chosen, namely: it should have global shutter, and must be triggerable.

A rolling shutter camera does not provide geometrically coherent image on a vibrating platform, because different horizontal bands of the image are captured in different time instants, with different camera axes.

Triggered mode is essential, because we have to capture the images in known time instances. The remote aircraft positions derived from the camera images are relative to our UAV, and the Inertial Measurement Unit (IMU) contains the exact position of our UAV, hence the synchronization of these units is vital.

Formerly, we used off-the-shelf WVGA (752×480) camera modules: MBSV034M-FFC from Mobisens, but its resolution turned out to be insufficient, so we designed a new module based on 1280×960 sized Aptina sensor (MT9M021, Figure 3), with 3.66 mm focal length High Resolution Infinite Conjugate μ-Video^{TM} Imaging Lenses from Edmund Optics.

The advanced feature of our camera module is the applied DS90UB913Q serializer chip, which exploits the benefits of Texas Instruments FPD-Link III [10] serial protocol. FPD-Link is an application of the LVDS standard, designed specifically for high-speed digital video interfaces. The protocol embeds the clock, GPIO signals and even a bidirectional communication channel in the data signal and therefore uses only one differential pair to transmit all the necessary signals: the power, the digital video stream, the control signals, and the I2C bus, which is needed for setting up the sensor chip.

C. Interface Board

Interfacing of the serial input stream of five cameras to the FMC-LPC connector of the FPGA board is an essential task of the framework (Figure 1). It is performed through the applied DS90UB914Q deserializers. An important design constraint was to ensure, that the corresponding signals of one camera are wired to the same IO bank of the FPGA (on all the three FPGA boards), to minimize the delay due to signal propagation, and

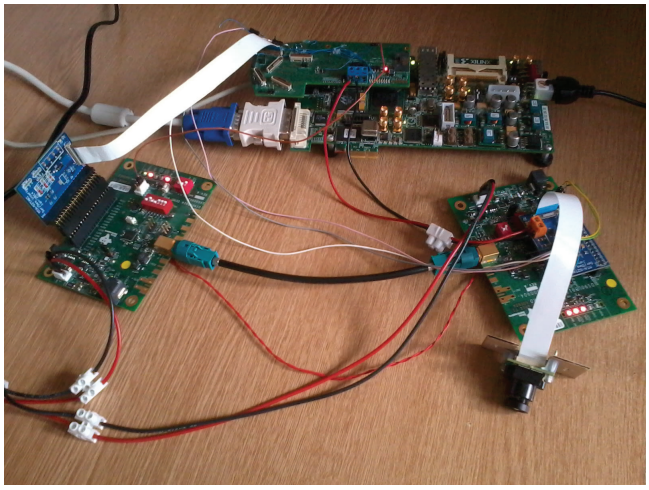


Fig. 4. Serializer/deserializer evaluation board test setup.

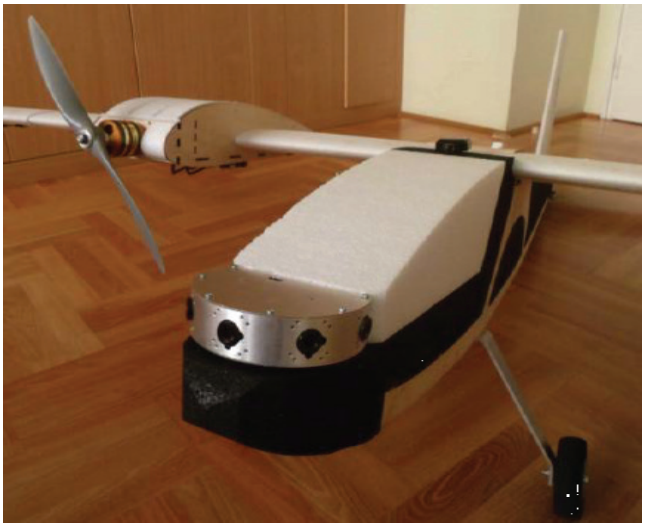


Fig. 5. The camera holder frame, mounted on the nose of the aircraft

to ensure that the IO buffers can be sychronised to the pixel clocks of each camera. The interface board is realized on a credit card sized ($7.3cm \times 4.6cm$) PCB, which is mechanically compatible with the above mentioned FPGA boards.

III. SERDES Evaluation Module Test

Since the interface board and the camera modules are still in production, the integration tests haven't been performed. However I managed to set up a hardware-in-the-loop test with one MBSV034M-FFC camera (with parallel data interface), and a serializer/deserializer evaluation module. The measurement setup contains an SP605 FPGA board, the SERDESUB-913 evaluation module (Figure 4), and two interface cards between the camera and ser/des cards. A Microblaze soft-processor (realized on the Spartan6 FPGA fabric) implements the control sequence of the ser/des ICs. The image data of the camera is streamed through the FPGA using AXI Stream protocol [11], to a monitor, so the video stream can be visualized in real-time. According to the previously mentioned measurement, the concept seems to be working, since the serial interface is capable of transmitting the video-stream at a maximum line rate of 1.4 Gbps. The power consumption of the components at a relevant 15 FPS (approx. 150 Mbps) transmission have been measured: Serializer: 360mW; Deserializer: 330mW; Image sensor: 270mW.

As a result, an estimation for the power consumption of the whole system (without the FPGA board) appears to be approximately 4.8W, which is significantly less than in the case of any general purpose computational platform, that provides the same processing bandwidth.

IV. Integration to the airframe

Physical system integration is always a key point of a complex embedded system. It is especially true for an airborne vision system with multiple cameras, where the fixed relative camera orientations, and cancelling any cross vibrations are critical. As described in [3], [4] an aluminum camera holder frame (Figure 5) has been mounted on the nose of the two engine polystyrene carrier aircraft, in a way that the optical axis of the front camera is aligned with the horizontal axis of the aircraft. The complete vision system with the FPGA board, the SSD drive, and the supply system is planned to be carried in the hull of the UAV.

V. Conclusion and Future Plans

As a conclusion of this paper, a low-cost, low-power, small-form-factor development platform has been designed for mobile image processing applications, which interfaces the image data to the FPGA, where various image processing architectures and algorithms can be evaluated and validated. The concept has been confirmed, with a hardware-in-a-loop test.

We named two research directions for the future:

- **Architectural**: study of scientific results related to hardware generation from high abstraction level languages (HLS - High Level Synthesis), hardware-software co-design - partitioning, adaptive partial reconfiguring.
- **Algorithmical**: the examination of the above mentioned architectural considerations in image processing point of view [12], especially architectural FPGA implementation issues.

Acknowledgment

The author wishes to acknowledge the guidance of Dr. Akos Zarandy, Dr. Zoltan Nagy, Dr. Andras Kiss, members of MTA-SZTAKI, Computational Optical Sensing and Processing Laboratory.

References

[1] Department of Defense, *Unmanned Aircraft System Airspace Integration Plan*, Tech. Rep. March, Department of Defense, 2011

[2] CARE, *Integration of UAVs into future air traffic management*, Cooperative Actions of R&D in Eurocontrol, 2001

[3] Akos Zarandy, Mate Nemeth, Zoltan Nagy, Andras Kiss, Levente Santha, Tams Zsedrovits, *A Real-Time Multi-Camera Vision System for UAV Collision Warning and Navigation*, submitted to JRTIP

[4] Akos Zarandy, Zoltan Nagy, Balint Vanek, Tamas Zsedrovits, Andras Kiss, Mate Nemeth, *A Five-Camera Vision System for UAV Visual Attitude Calculation and Collision Warning*, 9th International Conference, ICVS 2013, St. Petersburg, Russia, July 16-18, 2013. Proceedings

[5] David W. King, Allen Bertapelle, Chad Moses, *UAV Failure Rate Criteria for Equivalent Level of Safety*, International Helicopter Safety Symposium, Montreal, 2005

[6] Shane Degen, *Reactive Image-based Collision Avoidance System for Unmanned Aircraft Systems*, Thesis, Queensland University of Technology, 2011

[7] Dey, Debadeepta; Geyer, Christopher; Singh, Sanjiv; and Digioia, Matt, *Passive, long-range detection of Aircraft: Towards a field deployable Sense and Avoid System* (2009). Robotics Institute.Paper 962.

[8] Mejias, Luis and Ford, Jason J. and Lai, John S. (2010) *Towards the implementation of vision-based UAS sense-and-avoid* In: Proceedings of the 27th International Congress of the Aeronautical Sciences(ICAS 2010), 1924. September 2010, Acropolis Conference, Centre, Nice

[9] (2012, December 18). *Vivado Design Suite User Guide - High-Level Synthesis* [Online] http://www.xilinx.com/products/design-tools/vivado/integration/esl-design/

[10] (2013, April). *AN-1032 An Introduction to FPD-Link* [Online] Available: http://www.ti.com/lit/an/snla045b/snla045b.pdf

[11] (2011, March 7). *AXI Reference Guide* [Online] Available: http://www.xilinx.com/support/documentation/ip_documentation/ug761_axi_reference_guide.pdf

[12] Donald G. Bailey *Design for Embedded Image Processing on FPGA*, Massey University, New Zealand, JohnWiley & Sons (Asia) Pte Ltd., 2011

A mobile bio-signal measuring system on multiple limbs for rehabilitation purposes.

Norbert Sárkány, Attila Tihanyi, Péter Szolgay
Pázmány Péter Catholic University Faculty of Information Technology and Bionics Budapest, Hungary
{sarkany.norbert, tihanyi.attila, szolgay.peter}@itk.ppke.hu

***Abstract*—For any paralyzed or amputated patient rehabilitation is a necessary procedure. Rehabilitation for paralyzed patients is very important because of health protection and for amputated patients it helps to learn to control prosthetic. This paper presents the design of a mobile bio-signal measuring system. The goal of the project is to develop a robust mobile measuring system which can be used in several rehabilitation programs. The system is capable to measure ECG, EEG, EMG signals through multiple channels. At the moment the main scope is to measure EMG signals on the arm, and to optimize the signal condition and data compression through wireless data transmission. The current circuit and system designs are presented and a possible solutions for signal pre-processing is suggested.**

I. Introduction

Real-time motion capturing is a very frequent and important technique in biomedical signal processing, robotics and other applications based on human-computer interactions. The current state of electric and biological technologies led to a large variety of wearable systems capable of measuring and logging human body motion[1]. Such system is needed to measure and collect data from real life activity outside of a laboratory. Studies based on human motion analysis were carried out by many research laboratories[2]. After leaving the laboratory environment, this system can be used for rehabilitation purposes like relearning movement or for study gait, functional electrical stimulation, monitoring physical work processes and real life activity.

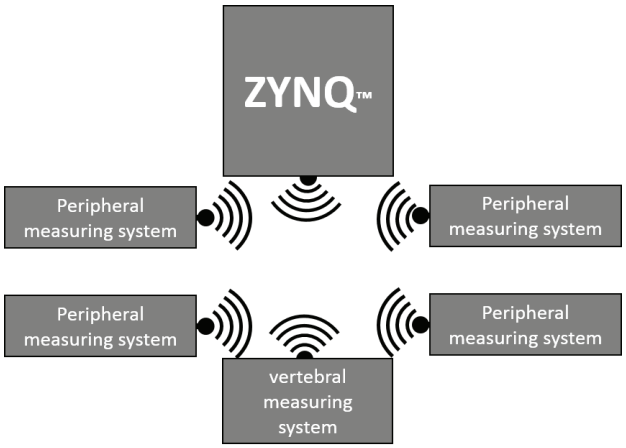
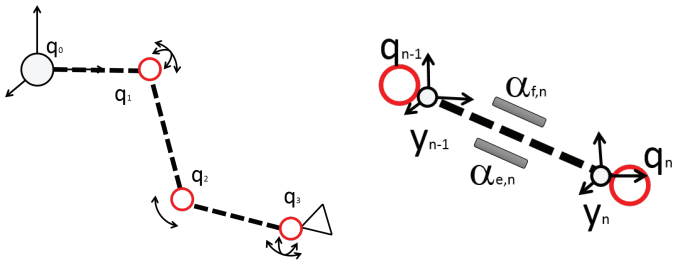


Fig. 1: Main processing system design

This paper presents a system design of a mobile multichannel bio-signal measuring system for rehabilitation purposes. After diseases like stroke or TIA (Transitoricus Ischaemias Attack) aftermath like paralysis or partial loss of ability to move are common and the patients require intensive rehabilitation. The rehabilitation dose not end after leaving the clinic it has to be continued also at home. This system achieves the simultaneous measurement of orientation and muscle activity of the specific limb, and can be used for relearning of the lost movement, to improve the training of athletes, to analyze the movement of healthy people which can be used in physiotherapy, occupational safety and health, in robotic applications and tele-operation.

At the moment the whole work focus on measuring of an arm's kinetics and muscle activity.

II. System Design



(a) The full arms kinematic model. (b) A limb segment.

Fig. 2: Fig.2a shows the kinematic model of the full arm, where q_i represents the joints. In Fig.2b a general limb segment is shown where y_i represents the 9-axis MEMS sensor and $\alpha_{i,n+1}$ the differential EMG electrodes for the flexor and extensor muscle parts.

This section discusses how to construct a system which is capable of simultaneously measuring the orientation and muscle activity of an arm, within it the individual components and their main parameters. Fig.1 shows the main system design, the core module being a Zynq-7000 architecture and the required peripheral device for measure arms, legs and back. The raw and preprocessed data is transferred wirelessly from peripheries to the core module which implements complex data processing. Every periphery implements basic Kalman

N. Sárkány, "A mobile bio-signal measuring system on multiple limbs for rehabilitation purposes"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 103-106.

filtering[3] on the data of the 9-axis Motion sensor and the EMG, which is shown in Fig.6

In order to fully understand what do we want to measure, we have to define the kinematic structure and the placement of the sensors. In Fig.2a we can see that the measuring is made on the whole arm from the shoulder(q_0) to the hand(q_3). Fig.2b shows the scenario of the placement of the sensor on a segment. At the end of a segment there is a 9-axis MEMS sensor(y_{n-1}, y_n) and a pair of differential EMG electrodes($\alpha_{i,n+1}$).

The 9-axis MEMS sensor stretches out a vector in the space which shows the orientation of the limb segment. The EMG electrodes $\alpha_{i,n+1}$ are for measuring of extensor and flexor muscle activity for the next actuated segment.

For an arm we need at last eight 9-axis MEMS sensors and six pairs of differential EMG electrodes. The communication between the peripheral devices micro-controller and the 9-axis MEMS sensors are realized through I^2C bus and with the EMG front end with SPI. The design of the peripheral measuring system is shown in Fig.3. If the research requires a more complex telemetry than the EMG sensor can be changed to an EEG or to an ECG front-end or we can daisy chain tham on the same SPI port.

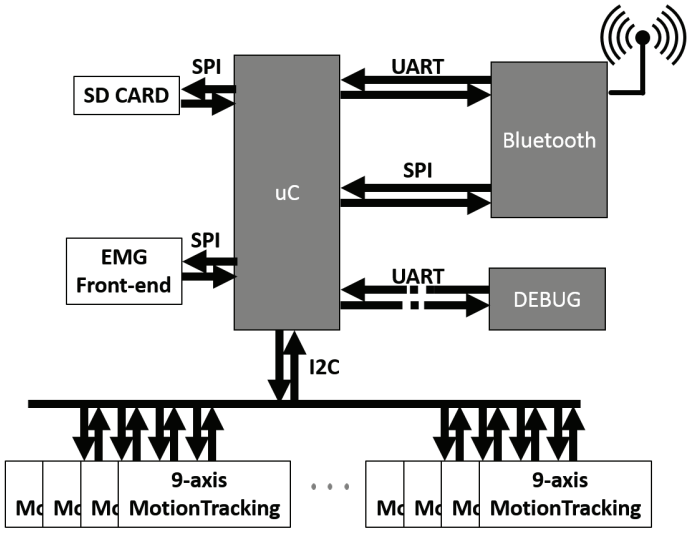


Fig. 3: Peripheral measuring system design.

A. Used parts in the peripheral module.

1) Micro controller: The peripheral devices computation is accomplished with a Microchip micro controller pic24FJ256GB108 [4], which is a 16-bit system with a maximal speed of 16 MIPS.

2) EMG Front-end: The EMG measurment are executed with TI's ADS1298 [5] device, which has 8 channels, 24-bit resolution, and a 32 KSPS maximal sample rate. If it necessary we can increase the amount of input of the front and and can change it to measure ECG and EEG signals to.

3) 9 Axis MEMS motion tracking module: is used, Invensense's MPU-9150 [6] module showen in Fig. 4 , which is a fully integrated system of 3-axis gyro, accelerometer and magnetometer all in one IC. The gyro's and the accelerometer's are sampled with 16-bit resolution and that of the magnetometer with 13-bit.

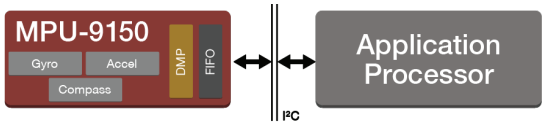


Fig. 4: 9 Axis MEMS motion tracking module.

The gyroscope measures the angular velocity which is linear to rate of rotation. It responds quickly and accurately and the rotation can be computed by time-integrating the gyroscope output. The accelerometer measure linear acceleration based on the acceleration of gravity.The problem with accelerometers is that they measure both acceleration due to the device's linear movement and acceleration due to earth's gravity, which is pointing toward the earth.

4) Communication and data storage: To get a flexible and convenient system, communication between the peripheries and the core module are achieved with Bluetooth.This allows to easily change or increase the number of devices. The data is saved to an SD Card on the devices.

B. Vertebral measuring system

For a complex measurement and to extend the telemetry at the same time measuring the movement of the limbs we can measure the muscle activity which allows the main spine movement. Because of the above-mentioned daisy chain we can add a front end to the vertebral measuring to measure ECG and the brain activity with EEG to obtain every data that is generated during their movement.

The measuring of the vertebral is similar to the limbs it contains measuring of the main muscle part which are responsible for its actuation and the movement of the spine to connect them to gather. The EMG electrodes are placed from bottom to top on the following muscel parts showen in Fig.5:

- Fascia thoracolumbar: the banding of the spine and fix the spine and the lumbar spine
- Latissimus dorsi: pulls the scapula away from the spine and keep them in the plane of the back
- Trapezius: moves up, down and towards each of the scapula

Simultaneously with the EMG the spatial movement of the spine and the upper body are measured with 9 Axis MEMS motion tracking module chain.

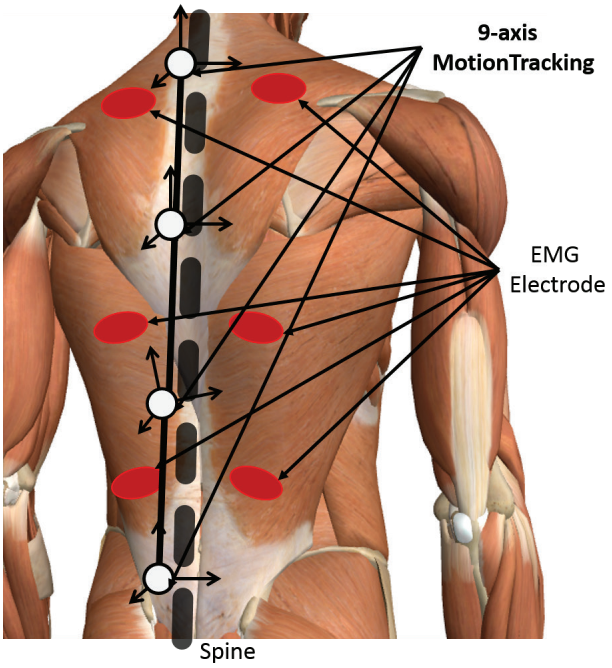


Fig. 5: The vbertebral measuring system design.

C. Core module of the motion analyzing system

The main processing to integrated the individual peripheral measurements are executed on Digilent's ZYBO development platform built around a Xilinx ZYNQ-7000 familly [7] the Z-7010. The big advantage is that this architecture tightly integrates an ARM Cortex-A9 processor with Xilinx 7-series Field Programmable Gate Array (FPGA).

The integration of the individual limb and vertebral measuring are realized with a cascade Kalman filter.

III. Kalman filter design to determine muscle activity for the actuation and the changing of the orientation of the limb.

The raw data not ready to use and they need to be calibrated. To calibrate these data it has to be scaled and the measuring bias has to be determined. Due to the reason that in our system data from many sensors are used and we want to reduce the error of the measurement and the integration of the data we implement a Kalman Filter for that purpose.In this section a Kalman filter model is introduced which which is based on the data of the peripheral measuring module.

The drifting rotation angle of the gyroscope signal is determined with trapezoid integration shown in equation (**??**).

$$\int_a^b f(x)\,\mathrm{d}x = (b-a)f(a) + (b-a)[\frac{f(b)-f(a)}{2}] \qquad (1)$$

Using the output of the accelerometer, a sensors rotation around the X -axis is the *Roll*, and the rotation around the Y -axis is the *Pitch* and can be calculated shown in equation (2) and (3) where acc_X, acc_Y and acc_Z are measurements in the X, Y and Z axes.

$$Roll = arctan(\frac{acc_Y}{(acc_Y)^2 + (acc_Z)^2}) \qquad (2)$$

$$Pitch = arctan(\frac{acc_X}{(acc_X)^2 + (acc_Z)^2}) \qquad (3)$$

Neither the accelerometer nor the gyroscope give accurate rotation measurements alone we have to combine the signals of the 3-axis gyroscope , the 3-axis accelerometer and the 3-axis magnetometer to determine the position of the endpoints show in Fig.2b (y_{n-1}, y_n) of a segment, and we combine the EMG signals($\alpha_{i,n+1}$) from the front end to reduce the error of the placement and to increase the signal quality.

The structure of the Kalman filter is shown in Fig.6. To achive the registration of data for the changing of a segment orientation in time and the necessary muscle activity, a third Kalman Filter will be implemented directly on the core modules FPGA.

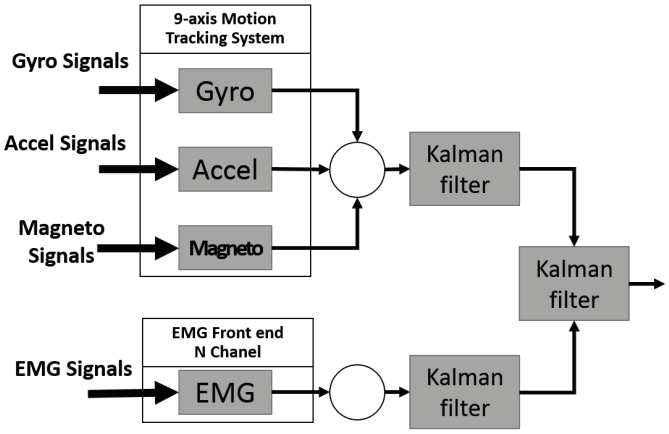


Fig. 6: The signal processing design.

IV. Conclusion

The system presented here is capable to measured limb orientation and muscle activity and to realize the goal of the project, complex motion analysis. With this measurement system we can follow down to the neuromuscular synapse the control signals from motor cortex and combine them with the movement of the actual limb.

Acknowledgment

This research project was supported by the University of National Excellence Program and the Research Faculty grant awarded to the Pázmány Péter Catholic University, Faculty of Information Technology and Bionics.

References

[1] K. Lorincz, B.-r. Chen, G. W. Challen, A. R. Chowdhury, S. Patel, P. Bonato, M. Welsh *et al.*, "Mercury: a wearable sensor network platform for high-fidelity motion analysis." in *SenSys*, vol. 9, 2009, pp. 183–196.

[2] S. Corazza, L. Muendermann, A. Chaudhari, T. Demattio, C. Cobelli, and T. Andriacchi, "A markerless motion capture system to study musculoskeletal biomechanics: Visual hull and simulated annealing approach," *Annals of biomedical engineering*, vol. 34, no. 6, pp. 1019–1029, 2006.

[3] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[4] M. T. Inc., accessed: 13.03.2014. [Online]. Available: http://ww1.microchip.com/downloads/en/DeviceDoc/39897c.pdf

[5] T. I. Incorporated, accessed: 13.03.2014. [Online]. Available: http://www.ti.com/lit/ds/symlink/ads1298.pdf

[6] I. Inc., accessed: 13.03.2014. [Online]. Available: http://www.invensense.com/mems/gyro/documents/PSMPU9150A00v43.pdf

[7] X. Inc., accessed: 13.03.2014. [Online]. Available: http://www.xilinx.com/products/silicon-devices/soc/zynq-7000/silicon-devices/index.htm

Analysis Based Parameter Estimation of an *in vitro* Transcriptional-Translational System

Zoltan A. Tuza
(Academic advisor: Gabor Szederkenyi)
tuza.zoltan@itk.ppke.hu

***Abstract*—Recent advances in mRNA measurement technology give us new insight into the dynamics of transcription activity in real time by measuring e.g. the Malachite Green Aptamers' fluorescent level. Thus, simultaneous measurements with fluorescent proteins in an easy-to-make cost effective *in vitro* transcriptional-translational system sets up a foundation for a molecular breadboard to test out biocircuits with various functions. In this paper we present a mass action type dynamic model for such an experiment and show the parameter estimation procedure, where the time series data contains information about both stages of gene expression. The identification process is supported by structural identifiability and sensitivity analysis.**

***Keywords*-Nonlinear systems, system identification, Reaction Networks**

I. Introduction

The amount and available types of data are often seriously limited in systems and synthetic biology experiments. In many cases, time-resolution or the sensitivity of the measurement technique is a serious obstacle for effective parameter estimation besides the structural non-identifiability of the mathematical model [1], [2], [3]. However, recent developments in mRNA measurement with Malachite Green RNA Aptamers along with florescent proteins provide a more reliable foundation for real time tracking of concentrations of the labeled species. Thus, we can measure transcription and translation simultaneously with sufficiently high frequency and specificity to directly use the obtained time series data for parameter estimation [4]. Furthermore, in synthetic biology we have the freedom to stitch together DNA parts (e.g. promoters, terminators, genes) in a matter of hours and test them in either *in vivo* or *in vitro*. Thus, we can put together many combination of DNA parts to achieve certain functions which may have or may have not existed before in nature. The set of these DNAs is often called biocircuits.

Conducting experiments *in vitro* has the benefit of shorter incubation time, ability to work with linear DNA and very good repeatability [5]. From the system identification point of view having the DNA and inducer initial concentrations as manipulable parameters that helps us to perturb the system dynamics to get sufficient data for parameter estimation.

Our aim is to develop and validate an ODE-based model, that can be used not only to fit the current data and show the dynamics of unmeasured states, but within appropriate limits, it can be used for predictive modeling. In our previous paper we worked out a possible mass action based model where we validated some of the parameters via wetlab experiments [6]. Recently, we conducted a wide-spread study, where several experimental conditions were tested [7]. In this paper we reiterate the model in terms of reaction speeds, time scales, structural identifiability and parameters sensitivity. The applied parameter estimation method is also implemented as part of the a modeling toolbox called transcription-translation or TXTL toolbox.

II. Experimental background

In this section, we briefly describe the experimental background of the modeled process. Figure 1. shows the main steps for the preparations of the cell extract. The *E. coli* cell's internal milieu cell is preserved through series of chemical processes and with additional resources this extract is capable of expressing non-genomic circular DNA. The external DNA was simply mixed with the cell extract and then the transcription and translation were followed simultaneously by measuring the fluorescence emission of the mixture at two wavelength. The transcription was monitored through Malachite Green Aptamer (MGApt) which is a short RNA segment with the capability of binding the Malachite green (triphenylmethane) dye. The MGApt:MG interaction enhances the fluorescence of the dye and thus enables the monitoring of the mRNA concentration in the solution. Translation was followed by using Green Fluorescent Protein (GFP), which exhibits a green fluorescence when exposed to blue or ultraviolet light. The separated emission profiles give us the opportunity to monitor the concentration dynamics of both stages of gene expression in real time throughout the life time of the cell extract.

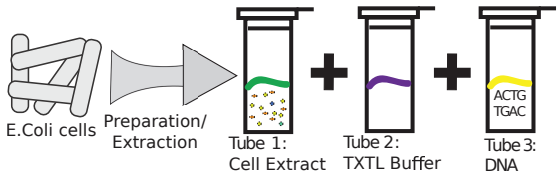


Figure 1. *in vitro* system overview: *E. coli* cell internal intent is processed through series of chemical procedure. The optimal cell condition and the missing resources are restored with the energy and buffer solution. The DNA tube represents our biocircuit which will be expressed in the *in vitro* system and the dynamics of certain parts are observed via fluorescent reporters. - taken from our CDC paper will be replaced in the camera ready version

A. In vitro transcriptional and translational system

All experiments were performed in a cell-free environment derived from *E. coli* crude extract. This extract contains all the endogenous system necessary for transcription and translation (e.g. ribosomes, RNA polymerase, translation initiation and elongation factors, etc.) but free from larger cellular compartments and from genomic DNA. For cell-free protein synthesis the cell extract needs to be supplemented with energy source, nuncleotides, amino

Z. A. Tuza, "Analysis Based Parameter Estimation of an in vitro Transcriptional-Translational System"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 107-110.

acids, tRNAs, and most importantly with the DNA of the desired gene. The detailed description of the system and the preparation steps of the crude extract can be found in [8]. Each cell extract can be made in large quantities, called batches, but the expression levels in each batches vary significantly; to avoid variation between extracts the same batch was used for all measurements.

B. Measurements

An experiment takes place in a 10μl reaction volume at $29\ ^\circ$C. Each experiment was run for 14 hours in a Biotek plate reader. The DNA construct contains the following elements: the promoter is a strong constitutive promoter drives the production of GFP-MGApt construct, the MGApt is placed at the 5' UTR, hence it will get not get translated. Fluorescence was measured in every 180 seconds for both MGApt (excitation: 610 nm; emission: 650 nm) and GFP (excitation: 485 nm; emission: 525 nm) channels. The sampling time was optimized to avoid sample photobleaching. Each experiment was repeated three times and standard deviation were calculated out of these experiments. In all the experiments the same DNA construct was added to the mixture, but with different concentrations. MG (triphenylmethane) dye has high autofluorescence and the free MG dye molecules degrades under high energy light for these reasons we have run extra wells ($10\mu l$ cell-extract + MG dye without DNA) to observe the dynamics of the dye itself. All measurements were corrected by these control wells. We changed the initial DNA concentration between 0.01nM and 20nM Figure 1. in [7], in the sequel we will use a subset of this data.

III. Process Modeling and Analysis

The mass-action type biochemical model and the corresponding kinetic ODEs are presented in this section. This kinetic ODE is further analyzed in terms of steady-state assumptions, structural identifiability and parameter sensitivity. This model will serve as base for parameter estimation.

A. Biomolecular Model

Among many possibilities to describe reaction networks we choose the mass action kinetics (MAK) framework, because the main criteria for MAK are met in the experimental conditions. Based on this fact, we can build a MAK based model for transcription and translation. The model accounts for resource consumption in order to cover resource limits and resource sharing effects. The details of the model can be found in [6]. Mass action based modeling has some advantages from the point of view of systems theory and parameter estimation. First of all, it is easy to apply stochastic solver to the model and investigate the dynamics of the non-deterministic regime [9]. Secondly, chemical reaction network theory (particularly for mass action systems) offers strong results for characterizing the qualitative dynamical properties of the studied system often without knowing precisely the model parameters [10], [11], [12].

1) Transcription Equations: The transcription in *E. coli* starts when the sigma factor (σ) activates the RNA polymerase (RNAP) and thus facilitates the binding of the promoter region on the DNA. Transcription occurs when RNAP builds the mRNA from the matching RNA nucleotides (NTP: ATP, GTP, CTP, UTP) that are paired with complementary DNA nucleotides of one DNA strand. We model transcription initiation and elongation as a one step reaction, where actual NTP consumption is calculated from the gene base pair length. Based on these considerations, the reaction scheme for transcription is the following.

$$RNAP+\sigma \underset{k_{\sigma r}}{\overset{k_{\sigma f}}{\rightleftharpoons}} RNAP\sigma \tag{1}$$

$$DNA+RNAP\sigma \underset{k_{DNAr}}{\overset{k_{DNAf}}{\rightleftharpoons}} DNA{:}RNAP\sigma \tag{2}$$

$$NTP+DNA{:}RNAP\sigma \underset{k_{NTPr}}{\overset{k_{NTPf}}{\rightleftharpoons}} NTP{:}DNA{:}RNAP\sigma \tag{3}$$

$$NTP{:}DNA{:}RNAP\sigma \xrightarrow{k_{TX}} DNA+RNAP\sigma+mRNA \tag{4}$$

2) Translation Equations: Just like in the previous step, the translation initiation and elongation is also modeled as a one step process. The ribosome (Ribo) binds first the free mRNA and then the codons of the mRNA pair with complementer anticodons of the tRNA. The tRNAs are loaded with the corresponding amino acids (AA) and hence the genetic code is translated to protein sequence. Since the tRNA and the AAs are supplemented externally to the crude extract in high quantity therefore we assume that charged tRNA is always abundant in the system and the tRNA charging dynamics is not modeled for that reason. When the translation is terminated the complex disassociates and the translated protein (GFP) is released.

$$mRNA+Ribo \underset{k_{Ribo\,r}}{\overset{k_{Ribo\,f}}{\rightleftharpoons}} mRNA{:}Ribo \tag{5}$$

$$AA+mRNA{:}Ribo \underset{k_{AAr}}{\overset{k_{AAf}}{\rightleftharpoons}} AAmRNA{:}Ribo \tag{6}$$

$$AAmRNA{:}Ribo \xrightarrow{k_{TL}} mRNA+Ribo+GFP \tag{7}$$

3) Protein Maturation: The translated GFP has a maturation time, which has been investigated independently in [13] and it puts the maturation between $5-7$ min. Note, that during the operation of the *in vitro* system the protein level does not achieve a steady state via balance in production and degradation but rather the system runs out of resources and the lack of protein degradation sets the final protein concentration.

4) mRNA degradation: The mRNA degradation enzymes are present in the *in vitro* system, therefore mRNA degradation should be included in the model. The total concentration of mRNA within the system can be calculated via

$$[mRNA]^{tot}=[mRNA{:}Ribo]+[AAmRNA{:}Ribo]+[mRNA] \tag{8}$$

The degradation enzymes attack the free mRNA and transform the functional mRNA into a nonfunctional one, this captured by a first order reaction where the mRNA is transformed into the zero complex.

5) Resource Degradation: The dynamics of *in vitro* systems is driven by the finite amount resources and the change of conditions (waste accumulation, pH change, etc.) during the incubation. On the modeling side there have been attempts to cover these effects, e.g. in [14], the resource degradation and enzyme activity decays over the course of the experiment as lumped terms.

B. State-space model

From the mass action type chemical reaction equations we can write a state space model serves as a basis for further analysis and parameter estimation. The state space model has 14 states and 15 parameters. Looking into the initial values, we see that out of 14 states only 5 of them has non-zero initial values, they belong to two groups. The first group what we have control over is constituted by the resources such as nucleotides NTP and amino acids AA and the plasmid or linear DNA concentration. The second set contains the enzymes facilitating the transcription and translation. The initial values of latter set were not measured in our *in vitro* system, but some literature values related to their orders of magnitude are available as a starting point for modeling and later on for parameter estimation.

Table I
TABLE SHOWS THE TWO GROUPS OF INITIAL VALUES THAT WE HAVE IN OUR MODEL, THE VALUES FOR THE FIRST GROUP IS GIVEN BY THE CRUDE-CELL EXTRACT PROTOCOL [8]. THE SECOND SET OF VALUES ARE TAKEN FROM THE LITERATURE BUT THEY ARE SUBJECT TO PARAMETER ESTIMATION.

NTP	1.2 mM	Resource	Protocol
AA	1.5 mM	Resource	Protocol
$RNAP$	100 nM	Enzyme	Literature
$Ribo$	1000 nM	Enzyme	Literature
$\sigma 70$	60 nM	Enzyme	Literature

1) Measured outputs: We measured two stages of gene expression simultaneously. The mRNA dynamics was tracked through MGApt and in our model the total concentration of mRNA is calculated as (8). We assume that the mRNA has the same fluorescence in all complexes, hence the measured output is simply the sum of the concentration of the three complexes. The GFP protein is a final product in the system, hence it exists only in one complex. To calculate concentartions scaling factors S_1 and S_2 were determined using an independent measurement with purified components. These values are highly machine specific numbers.

IV. PARAMETER ESTIMATION APPROACH

A. Prediction error minimization

After model reduction and taking into account the literature values we end up with 14 reaction constant and one initial value and one steady-sate value that are subject to parameter estimation.

The model has multiple outputs, $h_1(t)$, $h_2(t)$ for the mRNA level and for the protein level respectively and we are also performing curve fit with different DNA concentrations at once in order to incorporate more information about the parameters. To use the multiple outputs into the cost function we normalize each term with the two-norm of the corresponding time series data $y(\,\cdot\,)^k$ where k denotes the cases with differnt initial DNA concentrations.

$$C(\theta) = \sum_{k=1}^{m} \sum_{i=1}^{2} \frac{1}{\|y(\,\cdot\,)^k\|_2} \sum_{t=1}^{n} \frac{(y^k(t) - h_i^k(\theta, t))^2}{(\sigma_i^k)^2} \tag{9}$$

where $h_i^k(\theta, t)$ is the kth initial case of the model at time point t. The σ_i^k denotes the standard deviation of the kth case and the ith output.

The kinetic ODE model equations were solved by CVODES [15] because the model itself is still stiff. Most probably the NTP consumption happening on fast time scale, but this is the model crucial aspect, hence an efficient solver is needed since more than 90% percent of simulation based parameter estimation time is spent on integrating the differential equations.

Once we have our cost function we impose a set of bounds on the parameter values. A kinetic model requires positive parameter values, hence we should restrict the range of possible parameters onto the positive orthant. Also appropriate upper limits were imposed on all of the parameters. The Least Square optimization was performed with the Nomad solver, witch is a gradient-free global optimizer [16].

$$\underset{\theta \in \Theta}{\operatorname{argmin}}\, C(\theta) \tag{10}$$

$$0 \leq \theta \leq UB \tag{11}$$

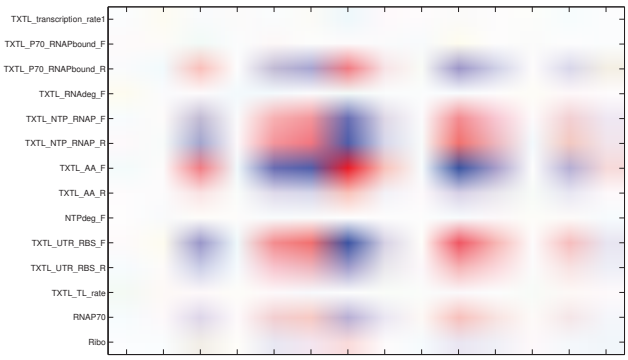


Figure 2. Estimated Covariance matrix of the paramteres. White means the parameters are not correlated, the red means positive covarinace and blue shows the negative covariance. The intensity of the color indicate the value of the covariance between the variables, higher intensity means higher correlation.

V. RESULTS AND DISCUSSION

Structural identifiability analysis tells us all of the parameters at least locally identifiable. That means that the value of each parameter contribute to the outputs, therefore the quality of the parameter estimation depends mostly on the quality of the data. The next step,sensitivity analysis helped us to show the importante of resource degradation in the model. After inspecting the reaction rates for transcription and translation on Figure 3. explains two characteristic feature of the *in vitro* system. First, we observed that after 2-3 hours of incubation the protein expression curve deviates from a strait line. This event also can be observed on the mRNA dynamics, peak of production occurs here. To explain that, we should look at the reaction rates for transcription, we can see that transcription capability constantly decays over time and around the 150 min (it's depending on the initial DNA concentration) crosses reaction rates for mRNA degradation, from this point the mRNA degradation dominates the systems's dynamics. The other feature of the *in vitro* system is that the translation stops much later

than the peak of the transcription. The right hand side of Figure 3 shows that the reaction rate for translation peaks around 150-200 min then decreases, which is in agreement with the mRNA dynamics shown on the left of 4, a clear sign of that the mRNA dynamics drives the protein production.

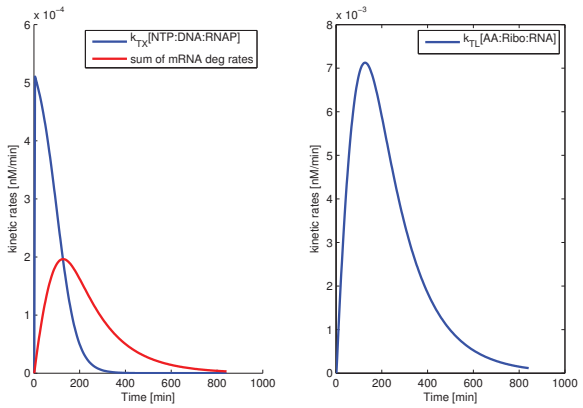


Figure 3. Reaction rates for the main contributors of the system dynamics. On the left the reaction rate for transcription decays over time and eventually crosses the rates of mRNA degradation. On the right the reaction rate for translation is shown for comparison with the mRNA dynamics

Figure 4. shows the measurements of both channels on the same time scale and converted to nM and μM for mRNA and GFP, respectively.

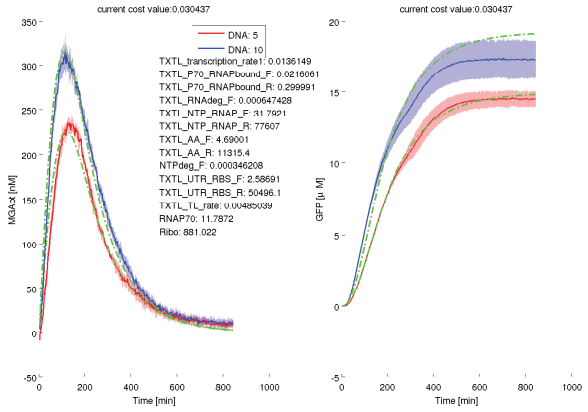


Figure 4. Result of parameter estimation for both channels with different initial concentration DNA in the system. The left figure shows the dynamics of MGApt, which is proportional to the mRNA concentration. Figure on right contains GFP dynamics. The fluorescent count for each channel has been converted with conversion factors.

The estimated covarance matrix shows where there a relationship between TL rates and mRNA degradation (in this case the mRNA degradation doesn't match with the measured one). Becouse of that we restricted the allows parameter values for mRNA degradation, the upper and lower bounds are based on independent measurements.

VI. Conclusion

In this paper we have shown how to utilize the rich measurement data and fit a mass action kinetic based model of gene expression. The model put emphasize on resource utilization, which is a crucial element of the cell-free systems since the resources are fixed no internal resource generation is available. Assessment of the model itself helped us to identify which parameters are identifiable in this model structure and also sensitivity analysis gives us how the parameters contribute to the dynamics. We have shown that how high resolution measurement techniques such as MGApt measurement can establish a foundation for parameter estimation in a cell-free environment. As we have seen that independent measurement of the mRNA play a crucial role, to push the estimation into right local minimum. Without that we would jump to a false conclusion about that system, just because the other minimum has a very large basin of attraction.

References

[1] O. Chis, J. R. Banga, and E. Balsa-Canto, "Structural identifiability of systems biology models: A critical comparison of methods," *PLoS ONE*, vol. 27, pp. 2610–2611, 2011.

[2] G. Craciun and C. Pantea, "Identifiability of chemical reaction networks," *Journal of Mathematical Chemistry*, vol. 44, pp. 244–259, 2008.

[3] G. Szederknyi, J. R. Banga, and A. A. Alonso, "Inference of complex biological networks: distinguishability issues and optimization-based solutions," *BMC Systems Biology*, vol. 5, p. 177, 2011.

[4] A. Raue, M. Schilling, J. Bachmann, A. Matteson, M. Schelke, D. Kaschek, S. Hug, C. Kreutz, B. D. Harms, F. J. Theis, U. Klingmller, and J. Timmer, "Lessons learned from quantitative dynamical modeling in systems biology," *PLOS ONE*, 2013.

[5] Z. Z. Sun, E. Yeung, C. A. Hayes, V. Noireaux, and R. M. Murray, "Linear DNA for Rapid Prototyping of Synthetic Biological Circuits in an Escherichia coli Based TX-TL Cell-Free System." *ACS Synth Biol*, 2013.

[6] Z. A. Tuza, V. Singhal, J. Kim, and R. M. Murray, "In silico modeling toolbox for rapid prototyping of circuits in a biomolecular "breadboard" system," in *Decision and Control (CDC), 2013 51th IEEE Conference on*, 2013.

[7] D. Siegal-Gaskins, Z. A. Tuza, J. Kim, V. Noireaux, and R. M. Murray, "Gene circuit performance characterization and resource usage in a cell-free breadboard info icon," *ACS Synth Biol*, 2014.

[8] Z. Z. Sun, C. A. Hayes, J. Shin, F. Caschera, R. M. Murray, and V. Noireaux, "Protocols for Implementing an Escherichia coli Based TX-TL Cell-Free Expression System for Synthetic Biology." *J Vis Exp*, vol. 79, p. e50762, 2013.

[9] S. Macnamara, A. M. Bersani, K. Burrage, and R. B. Sidje, "Stochastic chemical kinetics and the total quasi-steady-state assumption: application to the stochastic simulation algorithm and chemical master equation." *J Chem Phys*, vol. 129, no. 9, p. 095105, Sep 2008. [Online]. Available: http://dx.doi.org/10.1063/1.2971036

[10] M. Feinberg, "Chemical reaction network structure and the stability of complex isothermal reactors - I. The deficiency zero and deficiency one theorems," *Chemical Engineering Science*, vol. 42 (10), pp. 2229–2268, 1987.

[11] G. Craciun, Y. Tang, and M. Feinberg, "Understanding bistability in complex enzyme-driven reaction networks," *Proc. of the National Academy of Sciences of the USA*, vol. 103 (23), pp. 8697–8702, 2006.

[12] G. Shinar and M. Feinberg, "Structural sources of robustness in biochemical reaction networks," *Science*, vol. 327, pp. 1389–1391, 2010.

[13] J. Shin and V. Noireaux, "Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70," *J Biol Eng*, vol. 4, p. 8, 2010.

[14] T. Stgbauer, L. Windhager, R. Zimmerc, and J. O. Rdlerab, "Experiment and mathematical modeling of gene expression dynamics in a cell-free system," *Integr. Biol.*, vol. 4, pp. 494–501, 2012.

[15] C. the Sensitivity-Enabled ODE Solver in SUNDIALS, "R. serban and a. c. hindmarsh," *Proceedings of IDETC/CIE 2005*, 2005.

[16] S. Le Digabel, "Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm," *ACM Transactions on Mathematical Software*, vol. 37, no. 4, pp. 1–15, 2011.

Analyzing muscle synergies in cycling movements

Adam Valy
(Supervisors: Jozsef Laczko, Peter Szolgay)
valy.adam@itk.ppke.hu

***Abstract*—An experimental 3D-sEMG setup with a parallel processing software was created to study muscle synergies and the effect of resistance and speed on cycling movements. Experiments performed on 15 subjects yielded valuable results which will be a source for new results in biomechanic studies.**

***Keywords*- surface EMG, parallel processing, biomechanics**

I. INTRODUCTION

Individuals with lower limb lesions because of stroke or spinal cord injury need special methods (such as nerve stimulation) to maintain the fitness of the limb and prevent damage. A precise stimulation pattern for thigh muscles can be designed for different limb geometries to prevent unnecessary nerve or muscle dysfunction. This research is particularly important due to the increasing numbers of individuals with such lesions every year.

Non-invasive measurement the lower limbs during cycling execises is an efficient way of studying muscular electric signals in the presence of varying resistance at different speeds [1][2].

For the experiments, volunteers were asked to do exercieses on a SciFit ISO7000R recumbent bike, while their lower limb kinematics and myoelectric activity was measured by a Zebris CMS-HS ultrasonic 3D and sEMG system. The recorded data was then processed by parallel-capable National Instruments LabVIEW software.

II. SUBJECTS AND METHODS

Subjects of various qualities (age 20-28) were asked to do cycling exercises at two different speeds, and three resistances levels. These levels and speeds were chosen randomly for each participant but all exercises lasted about 20 seconds. The number of measurements for each subject is 6 * 12 (levels * markers), meaning a total of 1080 measurements.

Limb movements were recorded using seven 3D markers with a frequency of 100Hz, while the sEMG activity for 5 muslces was recorded at 900Hz.

For a 20 second exercise, 2000 3D data samples were obtained on 7 channels and 18000 sEMG samples for each of the 5 channels.

The seven adhesive kinematic markers were positioned as follows:

- marker 1: on the pedal axle to determine the crank angle
- marker 2: the ankle
- marker 3: on the fibula head
- marker 4: epicondylus (knee end)
- marker 5: close proximity to tuberculum maior
- markers 6 & 7: used as a horizontal reference

These markers, attached to fix points on the lower limb allowed monitoring knee and hip angles during exercises. Knee angle was monitored by determining the angle between vectors defined by marker pairs 2,3 and 4,5. Hip angle can be computed from vectors 4,5 and 6,7. Using a horizontal reference, there was no need for a marker on the shoulder, as it is almost impossible to find a steady location there.

The activity of five thigh muscles were monitored:

- Vastus Lateralis
- Rectus Femoris
- Vastus Medialis
- Semitendinosus
- Biceps Femoris

The first three are extensor muscles, while the last two are flexors.

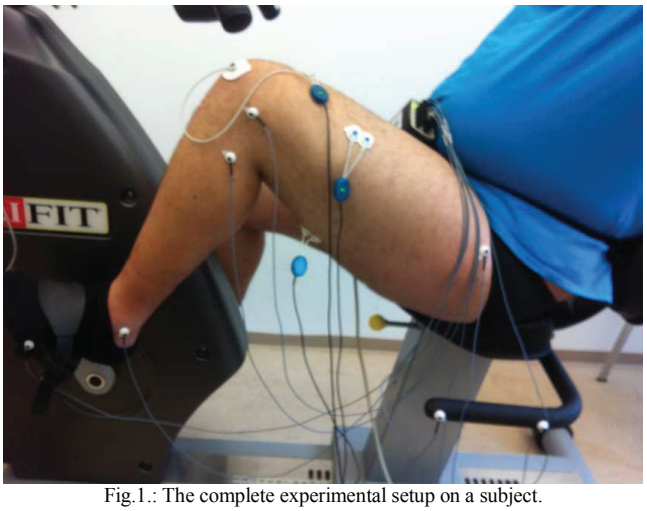


Fig.1.: The complete experimental setup on a subject.

III. COMPUTATIONS - EMG AMPLITUDES

Kinematic data (100Hz) and sEMG (900Hz) recordings were to be time-normalized. Before any further signal processing was performed, both 3D and sEMG measurements were error-corrected, rooting out outliers from multipath propagation (3D) and interruptions due to contact errors (3D, sEMG). For correct data analysis, 3D signals were transformed so that their sampling rate changed to 900Hz as

A. Valy, "Analyzing muscle synergies in cycling movements"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 111-114.

well. This was done by making 8 interpolated points between adjacent pairs of 3D data. Further processing and feature extraction was done with this new normalized data.

The following figures will demonstrate the normalization and interpolation process.

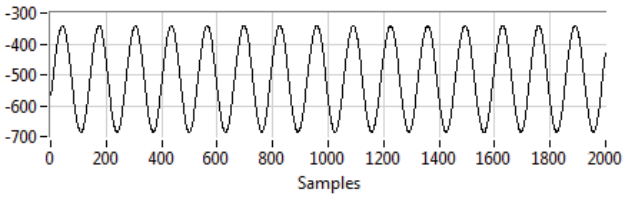


Fig.2.a.: Origial 100 Hz 3D data (2000 samples)

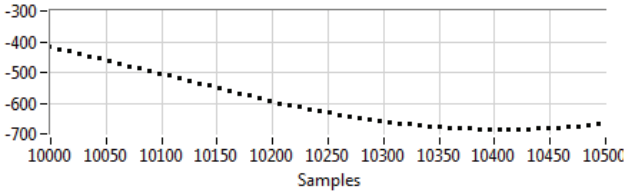


Fig.2.b.: Previous measurement now at 900Hz but not yet interpolated

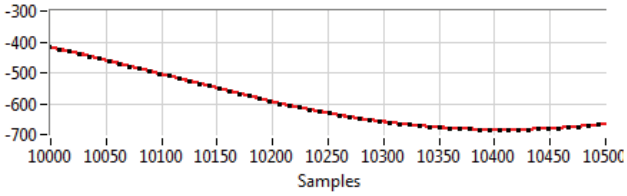


Fig.2.c.: Final result showing original 3D points in black, and new interpolated points in red (18000 samples). The Y axis in all cases is the z coordinate of the movement

sEMG signal filtering was performed with a fourth order Butterworth filter. The cutoff frequency domains were 25-48Hz and 52-200Hz which were determined by a spectrum analysis of the raw signal data. This filtered signal was then smoothed with a simple RMS filter using a 40 ms window size [3].

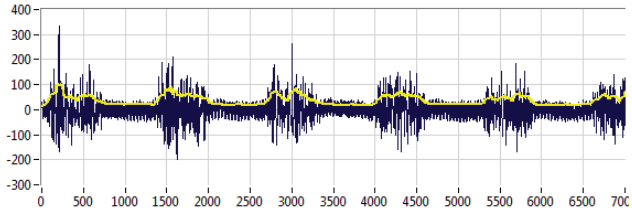


Fig.3.: Raw (dark blue) and filtered (yellow) sEMG signals

As mentioned before, a single marker was sufficient to determine crank angle after designating the reference up position as zero degrees. The correct sawtooth angle-time diagram was acquired as a result which was used to split the whole cycling into separate 360-degree periods for every participant.

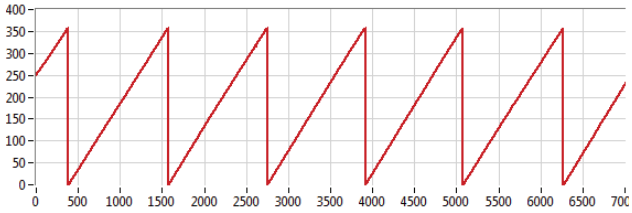


Fig.4.: Crank angle-time sawtooth diagram - each "tooth" is a complete cycle

To get feasible data from each participant, a 10 complete cycle average was calculated by acquiring the period start and endpoints throughout the complete exercise. Measurement data was exported into a 10 x m dimensional array, where 10 is the number of cycles to be averaged, and m is the measurement vector length. The elements of the result measurement vector are the averages of each column of the original 10 x m array.

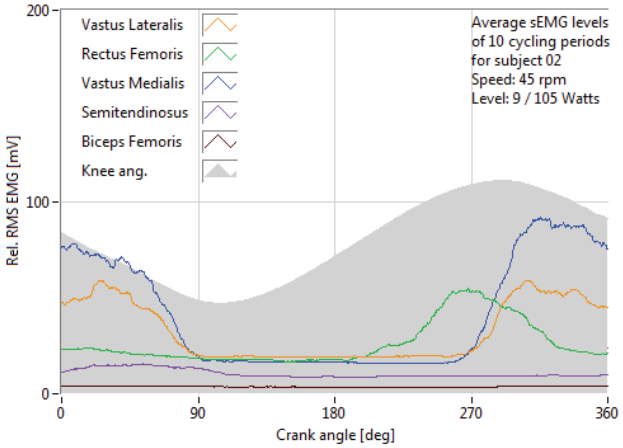


Fig.5.a: A 10-period average for a participant at low speed, high resistance. Knee angle is also shown.

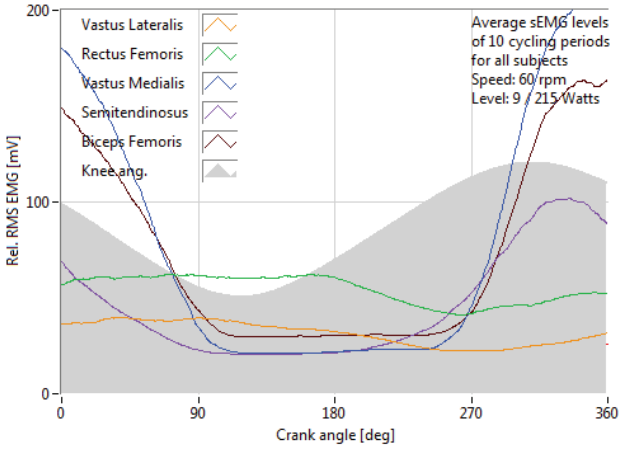


Fig.5.b.: Average of 10 trials averaged over all participants, at high speed and resistance

For all 6 levels, it is therefore possible to calculate an overall average sEMG pattern, which - with respect to crank

angle - can be used in an appropriate stimulator to help spinal cord injured or stroke patients maintain the fitness of the lower limb.

IV. SOFTWARE

Measurements by the Zebris CMS-HS unit are transmitted to a PC via parallel port. Zebris' WinData software is capable of saving this data into text files. These text files are later processed by NI LabVIEW software. Unlike text-based languages, LabVIEW uses a graphical method which allows fast prototyping, debugging and the easy creation of parallel tasks which can mean an advantage when working with large amounts of vector and matrix data.

Programs are designed in a way that the programmer can easily piece together elementary and advanced functions to get the desired results.

Elementary functions are subroutines that do basic tasks, such as reading a measurement file or displaying data. Advanced functions include sEMG filtering, crank angle calculation, etc. Apart from measurement control, display and interface, the most commonly used mathematics are also implemented.

The aim of this modular software system is to be able to process data from different experimental setups with the same efficiency, since developing a specialized software for every new test scheme takes time which could otherwise be spent more effectively.

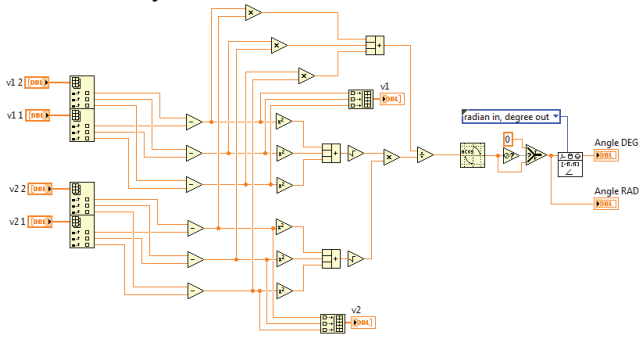


Fig.6.: Block diagram for angle calculation, input is 4 points (v11, v12, v21, v22) in 3D, output is an angle between vectors v1 and v2

The above elementary function is an example for building blocks of more complex algorithms. The calculation of the average activity across 10 trials is demonstrated by the scheme below.

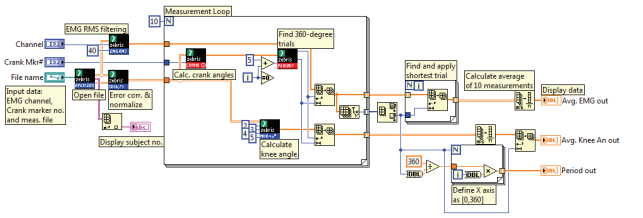


Fig.7.: LabVIEW block diagram of the 10 trial averaging algorithm. Note that blue lines represent integer values, while orange lines mean double variables.

The simplified flowchart of this algorithm can be seen on figure 8.

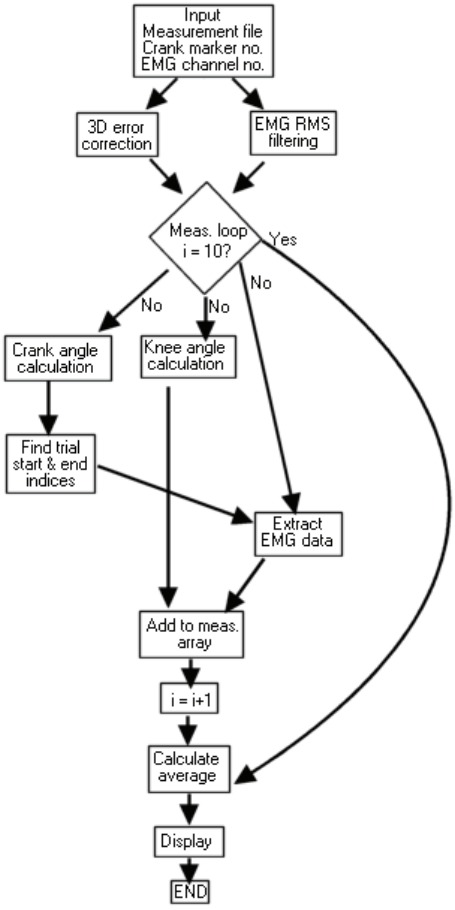


Fig.8.: Average sEMG amplitude calculation for one participant at one level

The novelty of the results compared to those of [1] and [4] lies in the fact that the stimulation patterns depend on the bike crank angle which was computed form marker coordinates recorded with a sampling rate of 100Hz instead of 25Hz as it was done earlier. This enables computation of crank angles at higher precision during cycling with variable speed. Another issue is that the new experimental setup enables to define stimulation patterns including 5 muscles. Thus, 3 extensor muscles are considered instead of 1 and this helps to design stimulation patterns based on muscle geometry and on co-activation among these parts of the quadriceps muscle group.

V. FUTURE WORK

It is desirable to take a statistical look at the measurements and computed averages. Is there a significant increase in sEMG amplitudes when cycling speed increases? What is the exact effect of resistance on the sEMG of one muscle [4]? This requires implementation of further functions, however, NI LabVIEW is an efficient tool.

Having a reasonable amount of data at hand, it is possible to further this research by examining activation times, determining the cooperative range of different muscles. The treshold to determine muscular electric activity is well defined [5].

This work is the continuation of what was started in January 2013 [6]. One year ago, a complex 3D-sEMG analysis and prediction system was proposed. The software created and used now is a potential part of that system, which would not only be used for research but for actual rehabilitation purposes as well.

ACKNOWLEDGMENTS

The author wishes to acknowledge the guidance of József Laczkó and Péter Szolgay. This work is supported by the Pázmány Péter Catholic University - University of National Excellence, the Faculty of Information Technology - Research Faculty, grants TÁMOP-4.2.1./B-11/2-KMR-2011-0002 and TÁMOP-4.2.2./B-10/1-2011-0014 by the European Union.

REFERENCES

[1] T. PILISSY, A. KLAUBER, G. FAZEKAS, J. LACZKO, J. SZECSI:
"Improving functional electrical stimulation driven cycling by proper synchronization of the muscles",
Clinical Neuroscience/Ideggyogy Szle.61(5-6)pp. 162-167

[2] P. KATONA, J. LACZKO:
"*Efficiency during cycling: the effect of altered pedaling speed on muscle activities"*
International Scientific Conference Motor Control 2012, Wisla, From Theories to Clinical Applications, Book of Abstracts, Ed. K. Slomka and G. Juras. p.37.

[3] G. KAMEN, D. A. GABRIEL:
„Essentials of Electromyography,"
Human Kinetics, 2010, ISBN(10): 0-7360-6712-4

[4] R. TIBOLD, G. FAZEKAS AND J. LACZKO:
Three-Dimensional Model to Predict Muscle Forces and Their Relation to Motor Variances in Reaching Arm Movements,
Journal of Applied Biomechanics, 2011, 27, 362-374

[5] OZGUNEN K. T., CELIK U., KURDAK S.:
"Determination of an optimal threshold value for muscle activity detection in EMG analysis"
Journal of Sport Science and Medicine,
Volume 9, Issue 4 (2010) pp. 620-628

[6] A.VALY:
"*A computer-aided setup for studying relations between EMG prediction, signals and muscular activity*,"
Proceedigns of the Interdisciplinary Doctoral School, 2012-2013 Academic Year, Faculty of Information Technology, Pázmány Péter Catholic University, Budapest Pázmány University ePress, pp. 71-75, HU ISSN 1788-9197

Retrograde signalling in the GnRH network

Wilheim Tamás
(Supervisor: Dr Kalló Imre)
wilheim.tamas@koki.mta.hu

***Abstract*—It was previously shown, that GnRH neurons mediate their glutamate- and GABAergic input through endocannabinoid signalling. Since these classical neurotransmitters are often co-synthesised with peptides or catecholamines, immunohistochemical triple labelling and 3 dimensional reconstruction, were carried out to identify the afferents of GnRH neurons which are target of endocannabinoid signalling i.e. immunoreactive for type 1 endocannabinoid (CB1) receptors. CB1 receptor-immunoreactivity was detected in kisspeptin-, neuropeptide Y- and TH-immunoreactive varicosities adjacent to GnRH cells. This suggests, that some of the peptidergic and catecholaminergic afferents of GnRH neurons are subject of CB1 retrograde signalling, via release of peptides/catecholamines and the classic neurotransmitters could be differentially regulated in the mouse brain.**

***Keywords*-GnRH; kisspeptin; retrograde signalling; afferents;**

I. Introduction

The GnRH neurons provide the main output of the hypothalamic circuitry controlling reproduction by secreting the GnRH hormone into the portal vessels of the hypophysis. The GnRH hormone is released at a pulsatile manner in both sexes, and during the female cycle, prior to ovulation, the pulsatile release is replaced by a surge release of GnRH hormone. The GABAergic network provides part of the estrogen feedback to the GnRH neurons. Due to the elevated chloride levels in the adult GnRH cells, the GABA transmission causes depolarization in GnRH neurons, thus GABA input is excitatory. The GnRH cells are capable to inhibit the input from GABAergic and glutamatergic afferents by releasing endocannabinoids [2,3]. The GnRH neurons receive one of their main peptidergic input from the kisspeptin secreting neural network. [8-13] There are two kisspeptin cell populations in rodents: the mainly glutamatergic neurons in the Arcuate nucleus (ARC), and the GABAergic population in the rostral periventricular area of the third ventricle (RP3V)[1,14]. By using immunohistochemical triple labelling it was investigated in this paper, whether kisspeptin, neuropeptide Y and TH expressing afferents of the GnRH cells are targets for the endocannabinoid - CB1 retrograde signalling pathway. To make sure, that the CB1 receptor is visible in the afferent terminal membrane, 3D reconstruction was used on the image stacks created with a confocal microscope.

II. Materials and methods

A. Animals

Adult female CD1 mice (2-3 months old, 2-30g b.w.; Charles River, Hungary) were housed under controlled lighting (12:12h light-dark cycle; lights on at 07:00h, and temperature (22±2 °C), with access to food and water ad libitum. All studies were carried out with permission from the Animal Welfare Committee of the Institute of Experimental Medicine (No. 2285/003) and in accordance with legal requirements of the European Community (Decree 86/609/EEC). Surgery was performed on animals under deep anaesthesia induced by an intraperitoneally injected cocktail of ketamine (25mg/kg b.w.), xylavet (5mg/kg b.w.) and pipolphen (2.5mg/kg b.w.) in saline.

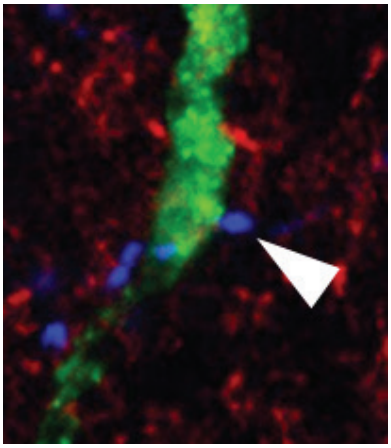


Fig. 1: 2D confocal image of CB1- and KP-IR fibers on a GnRH neuron. Green channel: GnRH, red channel: CB1, blue channel: kisspeptin. The white arrow shows the kisspeptin terminal on the GnRH cell, that contains CB1 receptor and was 3D reconstructed as shown in Fig.2

B. Tissue Preparation for Confocal Microscopy

The animals were ovariectomised (OVX) and 24h before they were sacrificed they were treated with estradiol(E2) subcuttan. The animals were perfused transcardially with phosphate-buffered saline (PBS; 0.1M) containing 4% paraformaldehyde (PFA), and 30% sucrose for cryoprotection. The brain was removed, and 30m thick coronal sections were cut on a freezing microtome on the day of perfusion. After the endogenous peroxidase activity had been quenched with 0.5% hydrogen peroxide (10 min), sections were permeabilised with 0.5% Triton X-100 (23,472-9, Sigma, 20min). Finally, 2% normal horse serum was applied (20min) to reduce non-specific antibody binding. Subsequent treatments and interim rinses in PBS (3×5min) were carried out at room temperature, except for incubation in the primary antibody or fluorochrome.

T. Wilheim, "Retrograde signalling in the GnRH network"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 115-118.

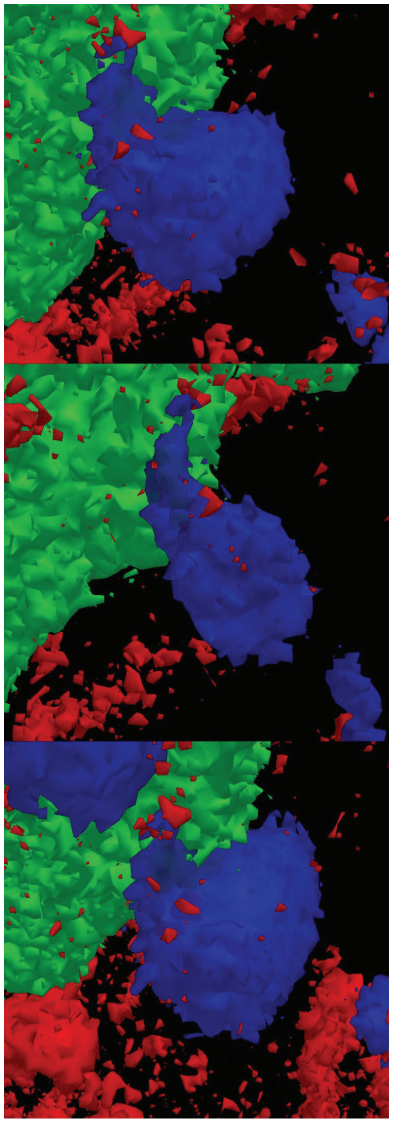


Fig. 2: 3D reconstruction of the confocal image shown in Fig. 1. Green channel: GnRH, red channel: CB1, blue channel: kisspeptin. The blue kisspeptin terminal is transparent, the red CB1-IR is visible on the terminal surface, and inside the terminal. The 3 images show the same KP-GnRH connection from different angles. There is no visible gap bewtween the two cells. By rotating the 3D object interactively, if the terminal is not on the GnRH cell, a gap is visible, that would not be easily discernable on a 2D optical slice.

C. Tripple Immunofluorescence for GnRH, CB1 and GnRH afferent marker Immunoreactivity

Sections from the RP3V and pre-optic region of the OVX+E2 animals were incubated (72h) in a cocktail of the guinea pig anti-GnRH (#1018, Hrabovszky, 1: 50,000) [6] and rabbit anti-KP (#566, Caraty, 1: 30,000) [7], goat anti CB1([5], 1:600) primary antibodies for the kisspeptin afferents, mouse anti TH (#MAB5280, Chemicon, 1:6000), guinea pig anti GnRH (#1018, Hrabovszky, 1: 50,000), goat

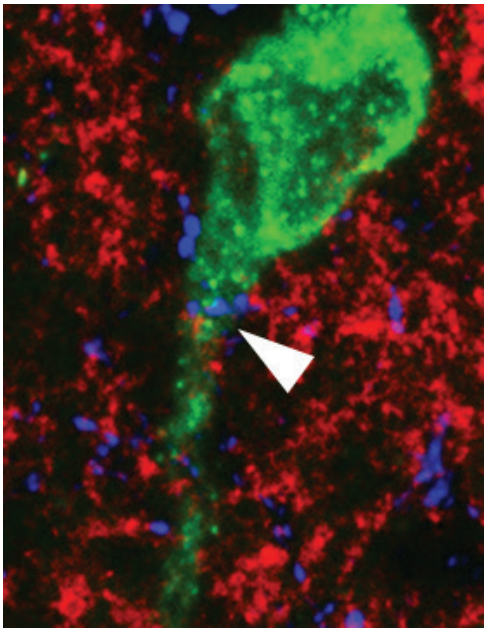


Fig. 3: 2D confocal image of CB1- and NPY-IR fibers on a GnRH neuron. Green channel: GnRH, red channel: CB1, blue channel: NPY. The white arrow shows the NPY terminal on the GnRH cell, that contains CB1 and was 3D reconstructed as shown in Fig.4

anti CB1([5], 1:600) for the TH afferents and rabbit anti NPY (Peninsula, #T-4070, 1:5000), guinea pig anti-GnRH (#1018, Hrabovszky, 1: 50,000) [6], goat anti CB1([5], 1:600) for the NPY afferents. GnRH immunoreactivity was visualised with FITC-conjugated donkey anti-guinea pig IgG (#706-095-148, Jackson ImmunoResearch Laboratories, 1: 1,000, 2h). To visualise the KP-IR structures, the sections were incubated in CY5-conjugated donkey anti-rabbit (#711-065-152, Jackson ImmunoResearch Laboratories, 1: 1,000, 2 h). To visualise the CB1-IR structures, the sections were incubated in CY3-conjugated donkey anti-goat (#706-165-147, Jackson ImmunoResearch Laboratories, 1: 1,000, 2 h). TH-IR was visualized with CY5 conjugated donkey anti mouse (#000, Jackson ImmunoResearch Laboratories, 1: 1,000, 2 h) and the NPY-IR was visualized with CY5-conjugated donkey anti-rabbit (#711-065-152, Jackson ImmunoResearch Laboratories, 1: 1,000, 2 h).

D. Confocal Laser Analysis and 3-D Reconstruction of GnRH-IR Afferents

The double-labelled sections were analysed using a Nikon A1R confocal microscope (Nikon, Japan). Multiple stacks of optical slices (1024×1024 pixels, z-steps 0.15 μm) were obtained by scanning all of the KP-IR neurones unilaterally in each of the selected coronal sections using a $60\times$ oil immersion objective. The FITC, CY3, and CY5 fluorochromes were detected with laser lines 488nm, 561 nm, and 641nm laser excitation. The separately recorded green, red and the artificially coloured far red channels were merged and displayed

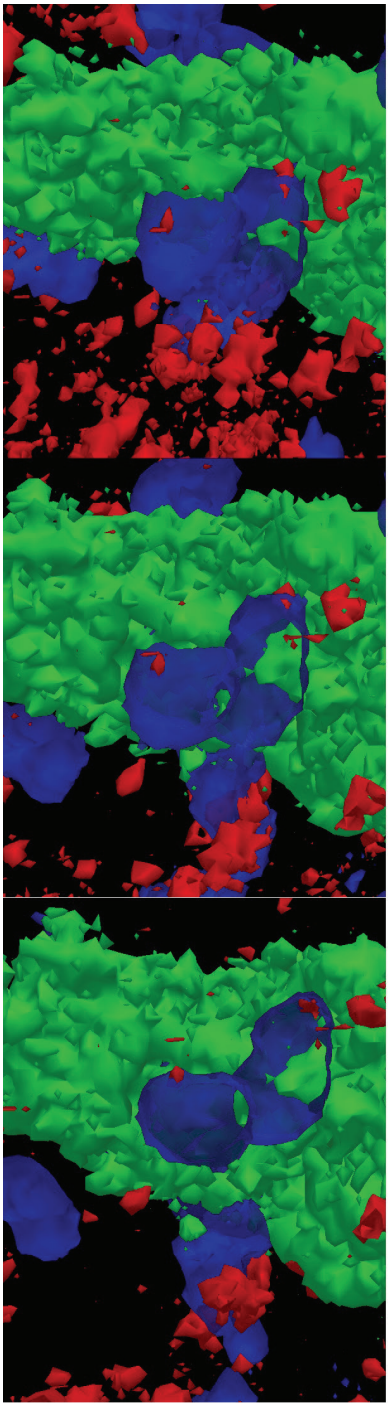


Fig. 4: 3D reconstruction of the confocal image shown in Fig. 3. Green channel: GnRH, red channel: CB1, blue channel: NPY. On the 3D reconstructed images, a CB1 positive NPY terminal is shown on a GnRH cell. The blue NPY terminal is transparent, the red CB1-IR is visible on the terminal surface, and inside the terminal. The 3 images show the same NPY-GnRH connection from different angles. There is no visible gap between the two cells.

with the ImageJ software [5] running on an IBM-compatible

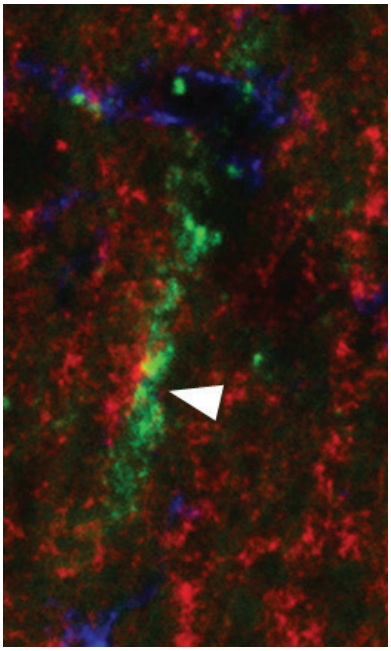


Fig. 5: 2D confocal image of CB1- and TH-IR fibers on a GnRH neuron. Green channel: GnRH, red channel: CB1, blue channel: TH. The white arrow shows the TH terminal below the GnRH cell with blue colour. The 3D reconstructed stack from this terminal is shown in Fig.6

personal computer. The images acquired with the confocal laser microscope were further investigated using three-dimensional (3D) analyses. The stack of optical slices were loaded into the visualisation software ImageSurfer 1.27 and the three channels containing images of consecutive optical slices were rendered in three dimensions with the Isosurface visualisation application. The isovalue was set individually for each image and colour channel to minimise any noise, while maintaining the proper cellular boundaries. The isosurfaces generated from the three channels in the same optical volume were visualised to check for cell-to-cell contacts, and the presence of CB1 receptor in the GnRH afferent terminals. This enabled verification of the findings from the two dimensional confocal image analyses.

III. Results

Visualizing membrane receptors is a hard task, because it is easy to wash out these molecules during the normal processing. To visualize the CB1 receptor, a weak fixation method was used, and the sections were processed on the day of perfusion to minimize CB1 signal loss. The fixation, sectioning, membrane permeabilization and transferring the sections into primary antibody was carried out on the same day to minimize signal loss. With this method it was possible to detect the fine grained CB1 signal in the preoptic area. This way, it was possible to show GnRH afferent terminals on GnRH cells that contain CB1 receptor. The presence of CB1 receptor in the afferent terminal is necessary for the retrograde signalling pathway that inhibits synaptic input via releasing

endocannabinoids that are detected by the CB1 receptors on the afferent terminal membrane. [2,3] To make sure that the CB1 receptor visible in the kisspeptin cell membrane is not in a neighbouring cell, the kisspeptin signal is visualized opaque. It can be seen, that the CB1 IR is present inside the kisspeptin terminal, and visible on the surface too. Further investigations are planned to make sure, that these receptors are not in a different terminal terminating on the afferent button. This investigation will be possible with the STORM technic due to its superior resolution, and possibility to correlate with conventional confocal microscopy.

Acknowledgment

The project was founded by OTKA K101326 grant.

References

[1] A. S. Kauffman, J. T. Smith: *Kisspeptin Signaling in Reproductive Biology* Springer Science+Business Media, ISBN 978-1-4614-6198-2

[2] Farkas I., Kalló I., Deli L., Vida B., Hrabovszky E., Fekete C., Moenter S. M., Watanabe M., Liposits Zs.: *Retrograde endocannabinoid signaling reduces GABAergic synaptic transmission to gonadotropin-releasing hormone neurons* Endocrinology. 151(12):5818-5829.

[3] Glanowska K. M., Moenter S. M.: *Endocannabinoids and prostaglandins both contribute to GnRH neuron-GABAergic afferent local feedback circuits* J Neurophysiol. 106(6):3073-81.

[4] Rasband, W.S., *ImageJ*, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2014.

[5] Makara J, Katona I, Nyíri G, Nemeth B, Ledent C, Watanabe M,de Vente J, Freund T, Hájos N: *Involvement of nitric oxide in depolarization-induced suppression of inhibition in hippocampal pyramidal cells during activation of cholinergic receptors.*, 2007 J Neurosci 27:1021110222

[6] Hrabovszky E, Molnár CS, Sipos MT, Vida B,Ciofi P, Borsay BA, Sarkadi L, Herczeg L,Bloom SR, Ghatei MA, Dhillo WS, Kalló I, Liposits Z: *Sexual dimorphism of kisspeptin and neurokinin B immunoreactive neurons in the infundibular nucleus of aged men and women.*, Front Endocrinol (Lausanne) 2011;2:80

[7] Franceschini I, Lomet D, Cateau M, Delsol G, Tillet Y, Caraty A: *Kisspeptin-immunoreactive cells of the ovine preoptic area and arcuate nucleus co-express estrogen receptor-α.*, Neurosci Lett 2006;401:225230

[8] Li XF, Kinsey-Jones JS, Cheng Y, Knox AM, Lin Y, Petrou NA, Roseweir A, Lightman SL, Milligan SR, Millar RP, OByrne KT: *Kisspeptin signalling in the hypothalamic arcuate nucleus regulates GnRH pulse generator frequency in the rat.*, PLoS One 2009; 4:e8334.

[9] Kinoshita M, Tsukamura H, Adachi S, Matsui H, Uenoyama Y, Iwata K, Yamada S, Inoue K, Ohtaki T, Matsumoto H, Maeda K: *Involvement of central metastin in the regulation of preovulatory luteinizing hormone surge and estrous cyclicity in female rats.*, Endocrinology 2005; 146: 44314436.

[10] Adachi S, Yamada S, Takatsu Y, Matsui H, Kinoshita M, Takase K, Sugiura H, Ohtaki T, Matsumoto H, Uenoyama Y, Tsukamura H, Inoue K, Maeda K: *Involvement of anteroventral periventricular metastin/kisspeptin neurons in estrogen positive feedback action on luteinizing hormone release in female rats.* J Reprod Dev 2007; 53: 367378.

[11] Clarkson J, dAnglemont de Tassigny X, Moreno AS, Colledge WH, Herbison AE: *Kisspeptin-GPR54 signaling is essential for preovulatory gonadotropin-releasing hormone neuron activation and the luteinizing hormone surge.* J Neurosci 2008; 28: 86918697.

[12] Roseweir AK, Kauffman AS, Smith JT, Guerriero KA, Morgan K, Pielecka-Fortuna J, Pineda R, Gottsch ML, Tena-Sempere M, Moenter SM, Terasawa E, Clarke IJ, Steiner RA, Millar RP: *Discovery of potent kisspeptin antagonists delineate physiological mechanisms of gonadotropin regulation.* J Neurosci 2009; 29: 39203929.

[13] Pineda R, Garcia-Galiano D, Roseweir A, Romero M, Sanchez-Garrido MA, Ruiz-Pino F, Morgan K, Pinilla L, Millar RP, Tena-Sempere M: *Critical roles of kisspeptins in female puberty and preovulatory gonadotropin surges as revealed by a novel antagonist.* Endocrinology 2010; 151: 722730.

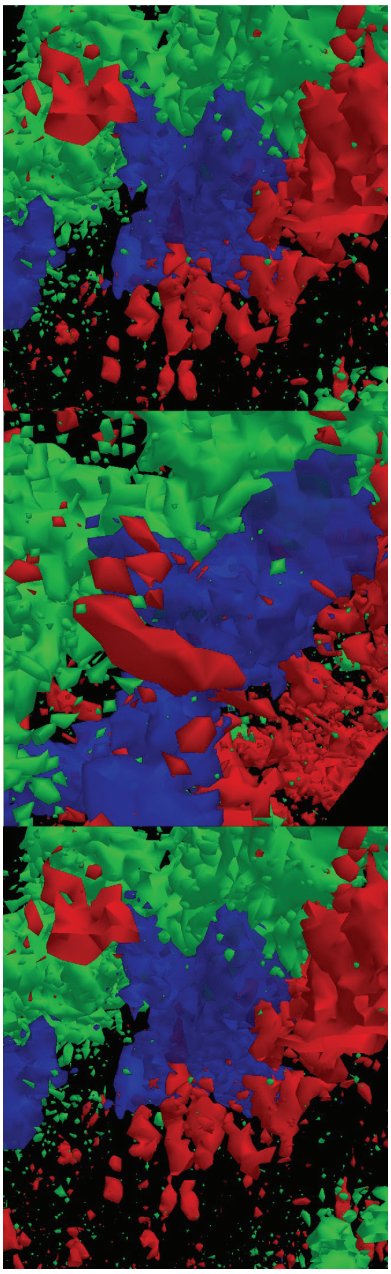


Fig. 6: 3D reconstruction of the confocal image shown in Fig. 5. Green channel: GnRH, red channel: CB1, blue channel: TH. On the 3D reconstructed images, a CB1 positive TH terminal is shown on a GnRH cell. The blue TH terminal is transparent, the red CB1-IR is visible on the terminal surface, and inside the terminal. The 3 images show the same TH-GnRH connection from different angles. There is no visible gap between the two cells.

[14] Kalló I, Vida B, Bardóczi Z, Szilvásy-Szabó A, Rabi F, Molnár T, Farkas I, Caraty A, Mikkelsen J, Coen CW, Hrabovszky E, Liposits Z.: *Gonadotropin-Releasing Hormone Neurones Innervate Kisspeptin Neurones in the Female Mouse Brain*, Neuroendocrinology 2013;98:281-289

Program 3

Feasibility of Electronic and Optical Devices, Molecular and Nanotechnologies, Nano-architectures, Nanobionic Diagnostic and Therapeutic Tools

Head: Árpád Csurgay

Simulation-based study of the spatial profile of fluorescence and photobleaching in two-photon-excited samples

Imre Benedek Juhász
(supervisor: Árpád Csurgay)
juhim@digitus.itk.ppke.hu

***Abstract*—We present a simulation-based study on the spatial distribution of fluorescence and photobleaching in two-photon-excited samples. We model the sample as a homogenous fluorophore solution in which the fluorophore molecules are described by a multi-level photodynamic model including also photobleaching. Diffusion of the molecules is also taken into consideration. Based on the model, we demonstrate the change of the three-dimensional profile of fluorescence over time: due to photobleaching, a "dark hole" evolves at the focus.**

Keywords-fluorescence; photobleaching; two-photon excitation; diffusion; photodynamic model; simulation

I. Introduction

Fluorescent techniques are widely used in present-day research, especially in biomedical sciences due to their high specificity and sensitivity [1]. In general in these techniques, the fluorescent molecules of the sample are excited by a laser beam, while the fluorescent photons emitted by the relaxing molecules are collected and counted. The intensity, spatial and temporal distribution, moreover the fluctuations of fluorescence refer to the number, spatial distribution, motion, and photodynamic properties of the fluorophore molecules. However, to acquire this information from the measurements, one has to know as precisely as possible where fluorescent photons originate.

In this paper, we summarize our work that was presented at the *SPIE Photonics Europe 2014* conference [2]: we demonstrate the simulation-based investigation of the aforementioned processes in the case of a homogenous fluorophore solution illuminated by a focused pulsed laser beam that causes two-photon excitation. In particular, we demonstrate the spatial distribution of fluorescence and its change over time due to photobleaching, i.e. photochemical reactions that destroy the fluorescent ability of the fluorophore molecule irreversibly.

In the first part of the paper, we delineate our simulation model, whereas in the second part, we present the simulation results.

II. Simulation model

The modeled arrangement is simple: a pulsed laser beam is focused on a homogenous fluorophore solution in which excitation, relaxation processes, photon emission, and photobleaching take place. The sample is divided into cubic cells among which diffusion occurs. In every volume cell, the number of the fluorophore molecules in each photodynamic state (see Section II-B) and the number of the emitted fluorescent photons are calculated in each time step.

The arrangement has a cylindrical symmetry with respect to the optical axis of the illuminating laser beam and thus a volume cell can be referred by two cylindrical coordinates, namely the distance r from the beam axis and the distance z measured from the focus along the beam axis; the origin of the coordinate system is the focus of the illuminating beam. Taking advantage of the cylindrical symmetry, computational complexity can be remarkably reduced: it is sufficient to simulate only one radial slice of the illuminated region [3], which reduces the number of volume cells from $O(n^3)$ to $O(n^2)$.

A. Model of the illumination

To provoke two-photon excitation, mode-locked lasers are used which emit short pulses with a duration in the femtosecond range. In this work, we describe the temporal shape of the pulses with a Gaussian function. The intensity in the pulsed Gaussian beam at point (r,z) at time instance t is calculated by the following formula given in [4]:

$$I(r,z,z) = A_0^2 \frac{w_0^2}{w^2(z)} \exp\left[-\left(t - \frac{r^2}{2cR(z)} - \frac{z}{c}\right)\right] \cdot \cdot \exp\left[-\frac{2r^2}{w^2(z)} + \frac{r^4}{w^4(z)}\left(\frac{\delta_0}{\omega_0}\right)^2\right], \quad (1)$$

where A_0^2 is the intensity at the origin at $t = 0$, c is the speed of light, t_0 is the $1/e$ pulse width, $\delta_0 = 1/t_0$ is the bandwidth of the spectrum of the pulse, ω_0 is the carrier frequency of the pulse,

$$w^2(z) = w_0^2\left(1 + \left(\frac{z}{z_0}\right)^2\right), \quad (2)$$

$$R(z) = z\left(1 + \left(\frac{z_0}{z}\right)^2\right), \quad (3)$$

$$z_0 = \frac{\pi w_0^2}{\lambda_0} \quad (4)$$

B. Juhász, "Simulation-based study of the spatial profile of fluorescence and photobleaching in two-photon-excited samples," in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds. Faculty of Information Technology and Bionics, Pázmány Péter Catholic University. Budapest, Hungary: Pázmány University ePress, 2014, pp. 121-125.

is the Rayleigh range, λ_0 is the wavelength at the center frequency,

$$w_0 = \sqrt{\frac{k}{k_0}} W_0 \tag{5}$$

is the waist size at the center frequency, k is the wavenumber, k_0 is the wavenumber at the center frequency. We determine W_0 as

$$W_0 = \frac{\lambda}{\pi\theta}, \tag{6}$$

where λ is the wavelength and θ is the beam divergence [5]. As it turns out from the above equations, the Rayleigh range and the waist size depend on the wavelength (and thus on the wavenumber). To calculate $I(r,z,t)$, we use only the center frequency, i.e. wavenumber $k = k_0$ and wavelength $\lambda = \lambda_0$, which results in

$$w_0 = W_0 = \frac{\lambda}{\pi\theta}. \tag{7}$$

However the formula

$$I_0 = \frac{2P}{\pi W_0^2} \tag{8}$$

for the light intensity I_0 in the focus was derived for a continuous wave Gaussian beam of power P [5], we suppose that it remains approximately valid also for our pulsed case and thus we can estimate A_0^2 in (1) as

$$A_0^2 \approx \frac{2P_{max}}{\pi W_0^2}, \tag{9}$$

where

$$P_{max} = P_{avg} 2\sqrt{\frac{\ln 2}{\pi}} \frac{T_{rep}}{\tau_{pulse}} \tag{10}$$

is the peak power of the pulses with Gaussian time profile, P_{avg} is the average power over a long time period, T_{rep} is the repetition time and τ_{pulse} is the full width at half maximum that characterizes the length of the pulses.

B. Photodynamic model of the fluorophore molecules

The states and state transitions are modeled by a three-level photodynamic model taken from the literature [6], [7], [8] with modifications, including four different photobleaching routes starting from the excited states. The model contains the following states (Fig. 1): ground state S_0, excited state S_1, triplet state T_1, and four photobleached states B_1, B_2, B_3, and B_4. Vibrational and rotational states are neglected. Thus, electronic states are considered to be discrete (sharp). In the model, the following transitions are possible:

- $S_0 \rightarrow S_1$, $S_1 \rightarrow B_2$, and $T_1 \rightarrow B_4$ via two-photon absorption,
- $S_1 \rightarrow B_1$ and $T_1 \rightarrow B_3$ via one-photon absorption,
- $S_1 \rightarrow S_0$ non-radiative relaxation (internal conversion),
- $S_1 \rightarrow S_0$ radiative relaxation followed by emission of a fluorescent photon,
- $S_1 \rightarrow T_1$ intersystem crossing,
- $T_1 \rightarrow S_0$ non-radiative relaxation.

In our model, we omit stimulated emission, as it is neglected also in [6], [7], [9]. The transition from state A to state B is described as follows: the time derivative of the number of molecules in state A in volume cell (r,z) at time instant t (supposing the change is due only to transition $A \rightarrow B$) is

$$\dot{A}(r,z,t) = -k_{AB}(r,z,t)A(r,z,t), \tag{11}$$

whereas the time derivative of the number of molecules in state B (supposing the change is due only to transition $A \rightarrow B$) is

$$\dot{B}(r,z,t) = k_{AB}(r,z,t)A(r,z,t), \tag{12}$$

where k_{AB} is the rate constant of the transition. If more than one transition affects a state, then the changes of the numbers of molecules deriving from each transition are added up.

The photodynamic processes in the model can be divided into two groups:

1) Transitions in the first group are related to photon absorption. Their rate constant depends on the illumination intensity either linearly or quadratically, and it is proportional to either the one-photon or two-photon absorption cross section (σ_{pb1}, σ_{pb3}; δ_{01}, δ_{pb2}, δ_{pb4}) which characterizes the process. The rate constant of such transitions can be calculated as

$$k_{AB}(r,z,t) = \sigma_{AB} \frac{\lambda_{exc}}{hc} I(r,z,t) \tag{13}$$

if the transition is induced by one-photon absorption, and as

$$k_{AB}(r,z,t) = \frac{1}{2} \delta_{AB} \left(\frac{\lambda_{exc}}{hc} \right)^2 I^2(r,z,t) \tag{14}$$

if it is coupled to two-photon absorption [7], where σ_{AB} and δ_{AB} are the one-photon and two-photon absorption cross sections at wavelength λ_{exc} respectively; h is Planck's constant, c is the velocity of light, whereas $I(r,z,t)$ is the illuminating light intensity at point (r,z) at time instance t in units of W/cm^2. Space and time-dependence of the rate constants arise from the space and time-dependence of the illumination intensity. In (14), the factor ½ expresses that two photons are required for the two-photon excitation process.

2) The rest of the transitions (relaxation transitions) are intensity-independent, and can be characterized by time constants (τ_{IC}, τ_f, τ_{ISC}, τ_T). The rate constant of such processes is the inverse of their time constant.

We suppose that the pulse duration is much shorter than the time constant of the relaxation processes (i.e. $\tau_{pulse} \ll \tau_{IC}, \tau_f, \tau_{ISC}, \tau_T$), furthermore the pulse length is much shorter than the pulse repetition time (i.e. $\tau_{pulse} \ll T_{rep}$). Thus, during the light pulse, the relaxation processes can be neglected and it can be assumed that a fluorophore molecule is excited at most once per pulse [6], [7], [8], [9]. Consequently, it is reasonable to construct two differential equation systems: one

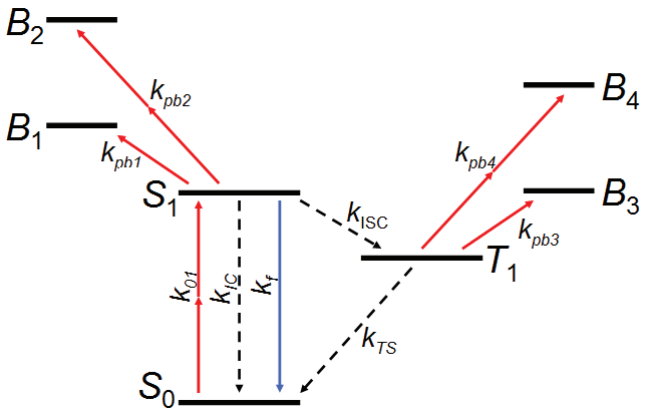


Figure 1: The Jablonski diagram of the fluorophore molecule represents the photodynamic states and possible state transitions of the model. S_0, S_1, and T_0 denote the ground state, excited state, and triplet state respectively; B_1, B_2, B_3, and B_4 denote the photobleached states. Solid lines refer to transitions that are related to either photon absorption or emission; dashed lines denote transitions without participation of photons.

that describes photon absorption processes taking place during light pulses (corresponding to group 1), and another one that describes relaxation transitions occurring during the interpulse periods (corresponding to group 2).

C. Model of the diffusion

The diffusion of the fluorophore molecules is modeled in the following way: in each laser cycle the probability that a molecule moves to the neighboring volume element along one of the coordinate axes during time T_{rep} is

$$p_{diff} = DT_{rep}/(\Delta x)^2, \tag{15}$$

where D is the diffusion coefficient and Δx is the size of the volume cells. It can be shown that after a large number of steps, the evolving distribution agrees with the distribution derived from Fick's law.

III. Simulations and results

A. Simulation parameters

In the presented simulation, we used the following parameters: wavelength of illuminating light $\lambda = 800\,\mathrm{nm}$, pulse length $\tau_{pulse} = 100\,\mathrm{fs}$, pulse repetition time $T_{rep} = 12.5\,\mathrm{ns}$ (which corresponds to pulse repetition rate $f_{rep} = 80\,\mathrm{MHz}$), beam divergence $\theta = 60°$. In order to get realistic results, fluorophore parameters were set to be in the same order of magnitude as values found in the literature [7], [8] for real fluorophores. We set the lifetime of fluorescent relaxation to $\tau_f = 100\,\mathrm{ns}$; however, we disabled internal conversion and intersystem crossing setting their lifetimes to $\tau_{IC} = \infty$ and $\tau_{ISC} = \infty$ respectively. Diffusion coefficient D was set to either zero or $2 \cdot 10^{-10}\mathrm{m}^2/\mathrm{s}$, two-photon absorption cross section δ_{01} to $2 \cdot 10^{-49}\mathrm{cm}^4\mathrm{s}$. For photobleaching cross section σ_{pb1}, we used $5 \cdot 10^{-23}\mathrm{cm}^2$, a value found in [8], while the three other photobleaching routes were disabled (their cross section was set to 0). Finally, 30 mW was used for the laser power. Under such illumination according to the model, the maximal local probability that a fluorophore molecule is excited during one pulse is 71.1%. The appropriate dimensions of the simulation volume and the cell size were determined based on simulations. We found that setting the radius and height of the cylindrical simulation volume to 5 μm and 10 μm, furthermore, the cell size to $\Delta x = 0.02$ μm is a reasonable trade-off between accuracy and computational time.

B. Simulation results

The simulation results are summerized in Fig. 2, which depicts the time-development of the spatial profile of fluorescence. This figure also compares the case without diffusion (left column) to that one when diffusion is present (right column).

In the first row [Fig. 2(a)-(d)], the number of emitted photons is depicted along axes r and z during the 1st, 10 000th, 20 000th, ..., 100 000th laser cycles. The darker the color of the curve, the larger the number of the preceding laser cycles is. If there is no diffusion in the sample and as it is hit by more and more laser pulses, the number of the fluorescent photons falls due to photobleaching, especially at the focus, where a "dark hole" evolves. Another noticeable feature of this scenario is the appearance of the small "hills" in the fluorescence profile next to the focus with a much smaller width than that of the initial fluorescent profile depicted by the uppermost grayish curve in each plot. However, if diffusion occurs in the sample, it can compensate photobleaching: the profile of fluorescence remains practically the same as the initial profile. Figs. 2(e) and 2(f) show the photon emission profiles in the 100 000th laser cycle as a three-dimensional plot.

Figs. 2(g) and 2(h) depict the number of emitted photons as a function of the distance of the photon source from the focus. Figs. 2(i) and 2(j) show characteristic distances in the case of different pulse numbers (i.e. different illumination durations): r_{50}, r_{75}, and r_{90} denote the minimal radii of spheres which are located at the focus and contain the volume cells that emit 50%, 75%, and 90% of the fluorescent photons. Along with the average distance of the location of the photon emission from the focus (denoted by the linked black squares), these values can be used to characterize the spatial resolution of the fluorescent method. The trends mentioned in the previous paragraph can be observed also in these diagrams: if there is no diffusion, the number of fluorescent photons decreases gradually in time, especially in the focal region, causing the increase of characteristic distances and thus the deterioration of the resolution.

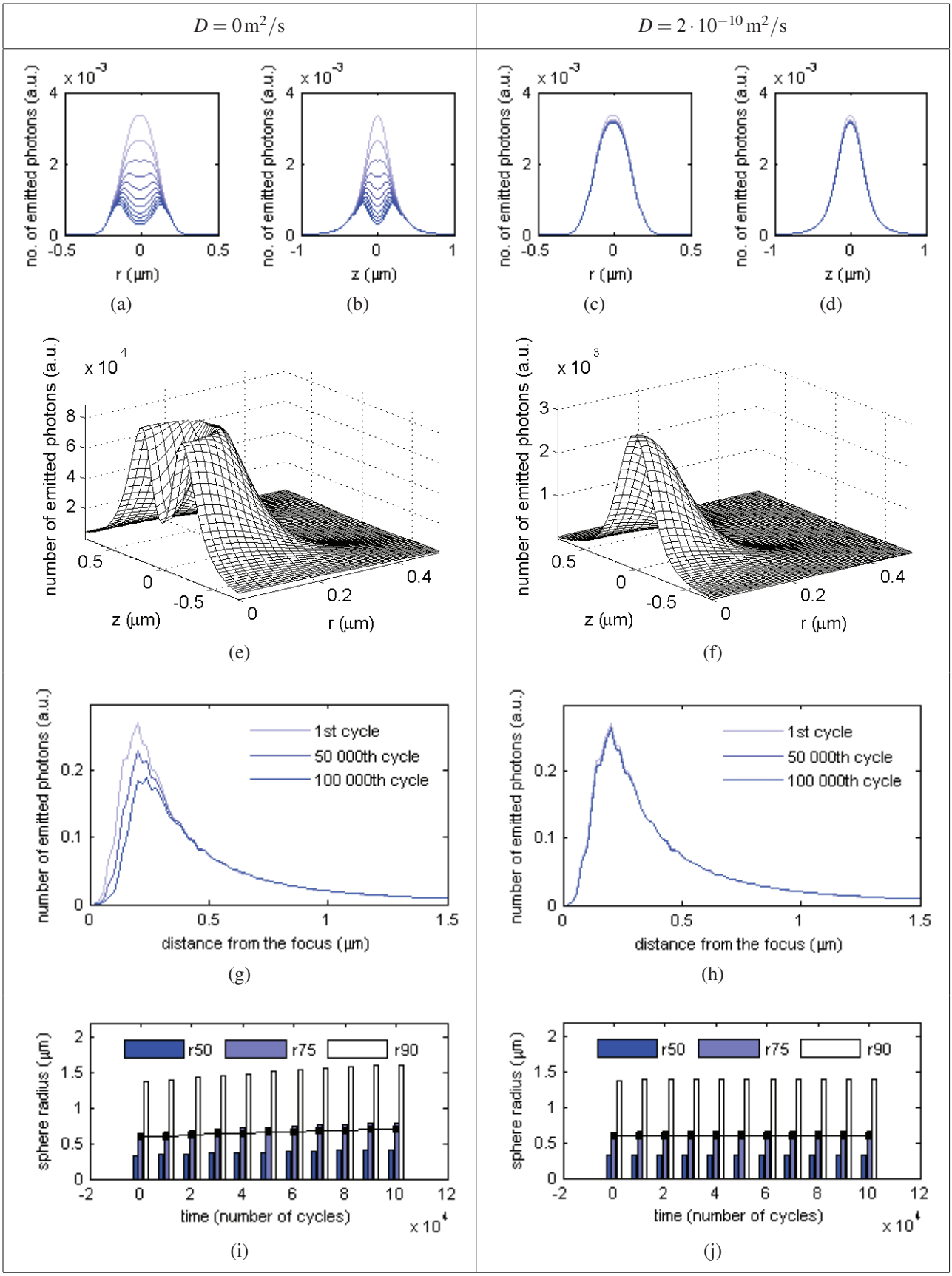


Figure 2: Spatial distribution of fluorescence and its time-development. (*Cont.*)

Figure 2: Diagrams (a)–(d) depict the number of emitted photons along axes r and z during the 1st, 10 000th, 20 000th, ..., 100 000th cycles. The darker the color of the curve, the latter the state is. Diagrams (e) and (f) illustrate the number of emitted photons per volume cell during the 100 000th laser cycle. Diagrams (g) and (h) shows the number of emitted photons as a function of the distance of the photon source from the focus, while (i) and (j) depict the characteristic distances of fluorescence over time: r_{50}, r_{75}, and r_{90} denote the minimal radii of spheres which are located at the focus and contain the volume cells that emit 50%, 75%, and 90% of the fluorescent photons. The average distance of the location of photon emission from the focus is denoted by the linked black squares. In the simulations, laser power was set to 30 mW and photobleaching cross section σ_{pb1} to $5 \cdot 10^{-23}\mathrm{cm}^2$. In the simulations showed in the left column, there was no diffusion in the sample ($D = 0\,\mathrm{m}^2/\mathrm{s}$), whilst diffusion occurred in the case of the right column ($D = 2 \cdot 10^{-10}\mathrm{m}^2/\mathrm{s}$). Note the difference between the scales of the horizontal axes of diagrams (a) and (b), as well as (c) and (d), furthermore between the vertical axes of diagrams (e) and (f).

IV. Conclusions

We presented a numerical model that allows the quantitative study of both the spatial and temporal characteristics of the photodynamic processes (including photobleaching) taking place in multi-photon-excited samples. Based on the model, we showed a simulation example that demonstrates quantitatively how photobleaching affects the spatial distribution of fluorescence over time. The results reveal that as photobleaching progresses, the most photons are not emitted from the focus but from an ellipsoid or, in more expressed cases, from a surface with an 8-shape cross section situated around the focus (not shown in this paper). In other words, a "dark hole" evolves at the focus. However, diffusion can inhibit this phenomenon. The results demonstrate that photobleaching not only decreases the number of emitted photons, but it also deteriorates the resolution of the fluorescent techniques.

There are several possible application fields of the model: taking into account the effects described above and using more accurate fluorescent profiles might facilitate the solution of the inverse problem (i.e. the reconstruction of the fluorophore distribution from the results of a two-photon microscope measurement) more effectively, which could result in a higher image quality. Moreover, the model might help to optimize the operating parameters of the microscope in different experimental setups. However, in these approaches, the lack of the photophysical parameters can be a limiting factor. Conversely, the simulator might be useful for the evaluation of the results of experiments in which the diffusion constant or photodynamic parameters are searched for. Finally, we note that the validation and calibration of the model require experimental measurements. Its adaptation to the particular techniques [e.g. two photon microscopy, super-resolution microscopy, fluorescence recovery after photobleaching (FRAP)] also needs further work.

Acknowledgment

The author would like to thank Professor Árpád Csurgay for his guidance. Grant from TÁMOP-4.2.1.B-11/2/KMR-2011-0002 is gratefully acknowledged.

References

[1] Barry R. Masters and Peter T. C. So, "Classical and quantum theory of one-photon and multiphoton fluorescence spectroscopy," in *Handbook of biomedical nonlinear optical microscopy* (Barry R. Masters and Peter T. C. So, eds.), pp. 91–152, New York: Oxford University Press, 2008.

[2] Imre B. Juhász and Árpád I. Csurgay, "Simulation-based investigation of the three-dimensional distribution of fluorescence and photobleaching in multi-photon excited samples," in *Proceedings of SPIE* (Frank Wyrowski, John T. Sheridan, Jani Tervo, and Youri Meuret, eds.), vol. 9131, (Brussels), 2014.

[3] A. Delon, Y. Usson, J. Derouard, T. Biben, and C. Souchier, "Photobleaching, mobility, and compartmentalisation: Inferences in fluorescence correlation spectroscopy," *Journal of Fluorescence*, vol. 14, no. 3, pp. 255–267, 2004.

[4] Z. Wang, Z. Zhang, Z. Xu, and Q. Lin, "Space-time profiles of an ultrashort pulsed gaussian beam," *IEEE Journal of Quantum Electronics*, vol. 33, no. 4, pp. 566–573, 1997.

[5] Bahaa E. A. Saleh and Malvin Carl Teich, "Beam optics," in *Fundamentals of Photonics*, pp. 74–101, Hoboken, New Jersey: John Wiley & Sons, Inc., second edition ed., 2007.

[6] J. Mertz, "Molecular photodynamics involved in multi-photon excitation fluorescence microscopy," *The European Physical Journal - D-Atomic, Molecular, Optical and Plasma Physics*, vol. 3, no. 1, pp. 53–66, 1998.

[7] C. Eggeling, A. Volkmer, and C. A. M. Seidel, "Molecular photobleaching kinetics of rhodamine 6G by one- and two-photon induced confocal fluorescence microscopy," *ChemPhysChem*, vol. 6, pp. 791–804, May 2005.

[8] R. Niesner, W. Roth, and K.-H. Gericke, "Photophysical aspects of single-molecule detection by two-photon excitation with consideration of sequential pulsed illumination," *ChemPhysChem*, vol. 5, pp. 678–687, May 2004.

[9] Martin Kauert, Patrick C. Stoller, Martin Frenz, and Jaro Rièka, "Absolute measurement of molecular two-photon absorption cross-sections using a fluorescence saturation technique," *Optics Express*, vol. 14, pp. 8434–8447, Sept. 2006.

Spin-Wave Based Computing Devices

Ádám Papp
(Supervisors: Dr. Árpád Csurgay, Dr. György Csaba (Notre Dame), Dr. Wolfgang Porod (Notre Dame))
papad@digitus.itk.ppke.hu

***Abstract*—We introduce the concept of spin-wave based wave-computing in magnetic thin films in analogy to the well established optical computing algorithms. Using micromagnetic simulations we demonstrate the basic building blocks of such a system and we show that these can realize more complex functionalities such as Fourier transformation or filtering of 1D vectors.**

***Keywords*-nanomagnet; spin-wave; Fourier transformation; non-Boolean computing; optical computing**

I. Introduction

In the last decade several spin-wave based computing devices has been proposed, most of them focused on developing novel switches [1] as an alternative to CMOS technology. We followed a different way by using spin-waves to represent analog information and borrowing the already established algorithms of optical wave-computing [2] [3] to realize complex functionalities.

Spin-waves and optical waves, however, are based on different physical phenomena and have major differences. Unlike optical waves spin-waves are not linear (although for small amplitude region they are close to linear) and need propagating media in which they have considerable attenuation due to the magnetic damping. A great advantage over the optical waves is the short wavelength (potentially $<$100 nm) which enables the on chip integration. We demonstrate that spin-waves display similar interference behavior as optical waves, and can be used in the linear regime to realize optical like devices. Spin-waves have 10^5 times smaller propagation speed than optical waves, however this still results only in a few nanosecond delay in a few micrometer distance.

Spin-waves can be generated in several different ways, e.g. spin-torque (oscillators), locally applied magnetic field, parametric excitations, strain [4] and Spin Hall Effect. For detection we can use giant magnetoresistance, magnetic induction, magnetic tunneling junction or inverse Spin Hall Effect. These technologies are compatible with CMOS, so the system could be integrated in a CMOS chip as well. Spin-wave devices are potentially low power and high speed, the most dissipation expected to occur by the electric-magnetic conversion.

In this paper we use thin permalloy films as propagation media, which is commonly used and widely available, however the damping is not low enough to allow spin-wave propagation over a few microns. There are materials with an order of magnitude lower damping (e.g. YIG), however these face other problems, like difficult shaping and low electric conductivity. We used spin torque to generate the spin waves in simulation, however in insulators we have to seek for different type of sources.

We designed the basic primitives of optics for spin-waves and demonstrated that they can serve as building blocks of computing devices that are designed to realize more complex functionalities. For verification of concepts we used the Object Oriented MicroMagnetic Framework (OOMMF) [5], with the Oxs_SpinXferEvolve module to include spin-transfer effects. We assumed $T = 0$ K for all simulations presented below and a 5 nm thick permalloy film with $M_s = 8.6 \cdot 10^5$ A/m, $A_{exch} = 1.3 \cdot 10^{-11}$ J/m and damping constant $\alpha = 0.008$.

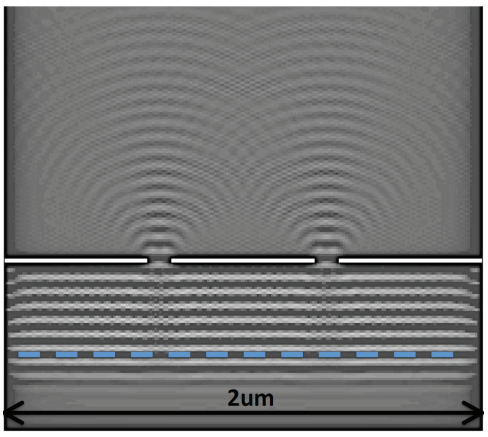


Fig. 1. Double slit experiment for spin-waves. The dashed blue line indicates the location the spin-wave sources creating a planar wave. The white walls are cutouts from the film and act as a mirror for spin-waves. [6]

II. Building blocks of spin-wave computing devices

A. Spin-wave sources

For spin-wave generation we used STO (Spin Torque Oscillator) structures. It is experimentally well established, that STOs generate narrow band spin waves which can propagate in the permalloy films to several micrometers, however the phase of the STOs can be easily changed by other waves inside the magnetic film, which is highly undesirable if the goal is to create plane waves or other type waves with fix phase pattern at the input nodes. To keep the phase constant we applied an $i = 1$ mA AC current at $f = 25$ GHz on the STOs instead of the usual DC. This forces the STO to an oscillation at the driving frequency and constant phase. For the input wave source we used a line of 20 nm diameter STOs, with individually controlling current amplitudes on the STOs

Á. Papp, "Spin-Wave Based Computing Devices,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 127-130.

according to the input vector analog values (same amplitude for plane wave).

B. Effective refractive index

Spin-waves show similar interference behavior as the optical waves for low and middle range amplitudes. Since the spin-waves are non-linear by nature, it is not exactly true, that they can cross each-other without interaction, however can be a good approximation.

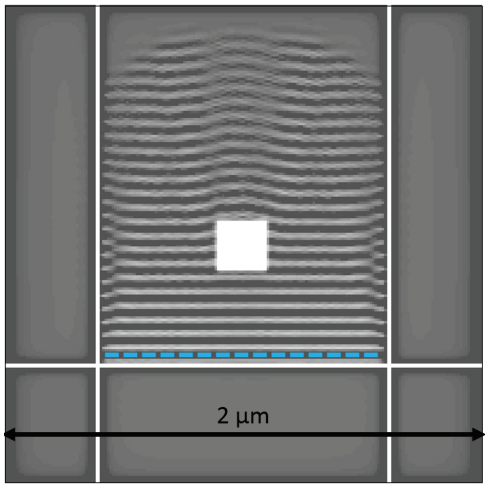


Fig. 2. Phase shifter realization in a permalloy film. In the white square region in the middle the applied external magnetic fields amplitude is 1.7 T. This shifts the wavefront by approximately 180°. At the top of the film we applied absorbing boundary conditions. [6]

One of the basic properties of optical waves is that they have different wavelength in materials with different refractive indexes. Using properly shaped geometries and multiple materials one can alter the way of the waves and create arbitrary interference patterns in specified locations/planes. We designed a method to alter the refractive index of the magnetic film for spin-waves by an applied magnetic field. Fig. 2 shows the result of a simulation where we measured the spin-wave wavelength in a magnetic film with different applied magnetic fields. The refractive index can be defined as $n_{21} = \frac{\lambda_1}{\lambda_2}$, where λ_1 is the wavelength in the first medium, λ_2 in the second medium.

C. Spin-wave mirrors and lenses

Spin-waves reflect back from the film boundaries or at the edge of other abrupt parameter changes. Thus the simplest way to create a mirror in the film is cutting out the proper shape.

To design a spin-wave lens we need to create a lens shaped external magnetic field over the film. For a simple approximation we can use the lensmaker's formula as for optical lenses, however creating homogeneouseous high amplitude magnetic fields over an arbitrary shaped region is not straightforward. Even if one can apply such a field in simulation, the demagnetization field of the film will change considerably, thus creating a space dependent effective magnetic field. The optimal external field distribution probably could be designed

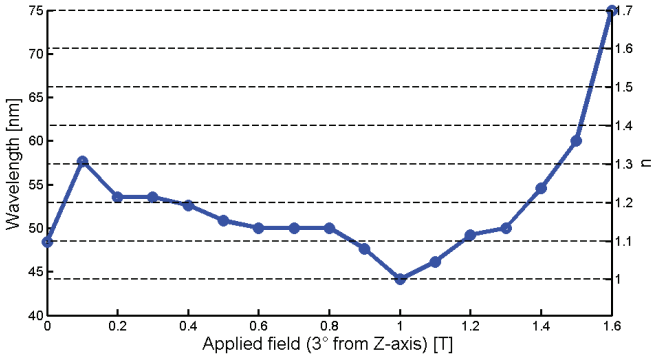


Fig. 3. The wavelength of the spin-waves in different applied external magnetic fields (OOMMF simulation). The refractive index on the right axis is calculated relative to wavelength in 1 T field. [2]

by proper calculations, we simply used homogeneous magnetic field as an approximation. Note, that in our design the focusing lens is concave as opposed to the usual optical focusing lenses, since the wavelength in the film outside the lens is smaller than inside the lens.

Although the lens is usually preferable to create one way signal propagation in a system, we can see, that the design and structure of a mirror can be much simpler, than of a lens.

D. Active damping

The Landau-Lifshitz-Gilbert-Slonczewski (LLGS) equation includes a spin torque term [7]:

$$\frac{d\vec{m}_f}{dt} = -\mu_0 \frac{|\gamma|}{1+\alpha^2}\vec{m}_f \times \vec{H}_{eff} - \alpha \frac{|\gamma|}{1+\alpha^2}\vec{m}_f \times \left(\vec{m}_f \times \vec{H}_{eff}\right)$$

$$+ \frac{g(\Theta_m)\mu_e I}{eM_s V}\vec{m}_f \times (\vec{m}_f \times \vec{m}_p)$$

If $\vec{m}_p$ is parallel to $\vec{H}_{eff}$ we can define:

$$\alpha_{eff} = \alpha \frac{|\gamma|\mu_0 H_{eff}}{1+\alpha^2} - \frac{g(\Theta_m)\mu_e I}{eM_s V}$$

Thus using negative DC current can increase the damping. We verified this by micromagnetic simulations. We estimate that to decrease the amplitude of a wave by an order of magnitude through a 20 nm wide region 3-5 mA current required, which is more, than what we use for generation (1 mA), but still a reasonable value.

III. Fourier transformation by lens or mirror

It is well known that an optical lens is able to produce the Fourier transformation of an image which is placed in the focal plane. We will demonstrate, that this is possible by using magnetic lenses (or parabolic mirrors) as well. Since the film only allows two dimensions for wave propagation, our device can only operate on one dimensional vectors.

A. Lens

Fig. 4 shows the design of a lens and the result of the Fourier transformation. An 1.6 mT magnetic field is applied in the lens shape, the input and output STOs are placed in the focal planes. The lens is quite thick, the range of refractive indexes is not wide enough to create a thin lens with short focal distance. Note, that in this case we did not allowed negative values in the input wave amplitude, which results in a DC component.

The operation of the system is simple, we can constantly apply the input signal in the input focal plane, and approximately 2 ns later read out the amplitude (and/or phase) at the output focal plane. This system has the potential to realize frameless computing, since the input can be changed continuously, however the difference between the propagation time of different paths in the film can set a minimal time between two consecutive input vectors.

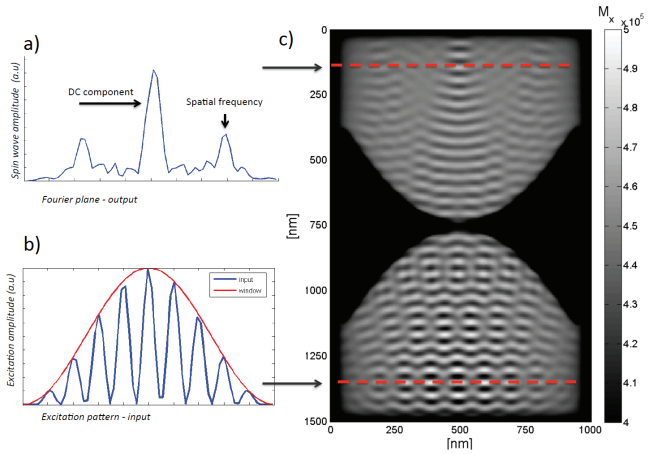


Fig. 4. Fourier transform of a 64 element vector by a lens. Part c) shows the magnetization in the film, the black concave shape is the lens, the red dashed lines show the two focal planes. The input shown in b) is applied in the bottom focal plane, the output wave amplitude at the top focal plane is shown at a). Note the two (symmetrically arranged) peaks that represent the frequency of the sinusoidal input. The middle peak is the DC component of the input (i.e. a focused plane wave). [2]

B. Mirror

Using a mirror has several advantages over using a lens, it is simple to design and produce, there are no losses due to reflections at the boundaries, and there is no limit for the focal distance. In exchange for these we lose the frameless behavior of the system, we need to use time multiplexing for the input and readout.

Fig. 5 shows the design of the system, the parabolic cutout at the top acts as a mirror. We applied a 0.9 T magnetic field almost perpendicular to the film, this sets the magnetization angle approximately to $45°$ from the plane. In the first cycle we apply the input wave in the focal plane for 1.5 ns: this is the time while the first waves reflect back to the focal plane and the readout can start for the same length of time.

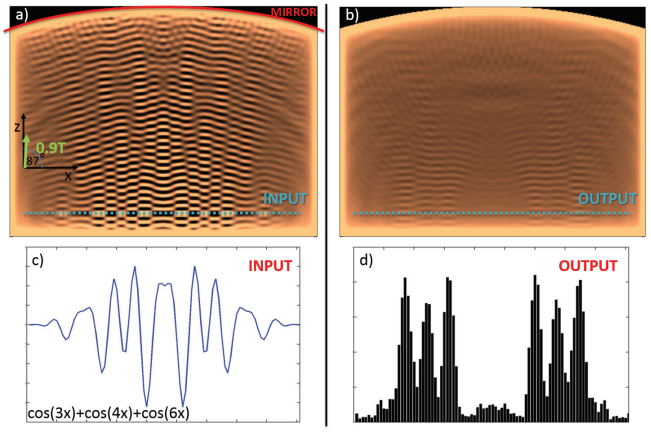


Fig. 5. Fourier transformation of a 99 element vector by a parabolic mirror. a) Applying the input b) readout c) input waveform (contains 3 sinusoid components) d) output wave amplitudes (3 peaks for the 3 frequency components). [8]

C. Double mirror

In order to realize more complex operations on the input data one need at least two Fourier transformations. There is no straightforward way to connect mirrors serially one after, as it is easy to do using lenses. In this design we attempt to do this by two facing mirrors with matched focal planes as it is shown in Fig. 6. This design can provide double amplitude output signal, however matching the two mirror can be challenging by physical realization (for 45 nm wavelength the mismatch should be under 10 nm).

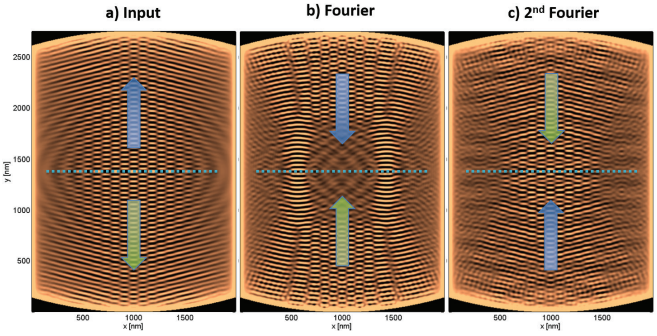


Fig. 6. Double mirror Fourier transformation. The blue dotted line in the center shows the focal plane, where the input and output takes place separated by time. The arrows show the wave propagation. After a) applying the input b) the first reflections form the Fourier transform of the input vector. c) The wave passes the focal plane and reflects back again, this is a second Fourier transform on the signal which is the original signal. [8]

Since in this system the waves should travel to the mirror and back at least twice, the damping can be a serious issue. We showed with simulations that geometrical losses (waves leaving at sidewalls) lower the total energy approximately by a factor of 4 in a cycle, and the permalloy damping causes losses in the same order of magnitude. In the design presented here we used the tenth of permalloy damping, which can be achieved by e.g. using YIG instead of permalloy.

The input is applied same way as in the one mirror case,

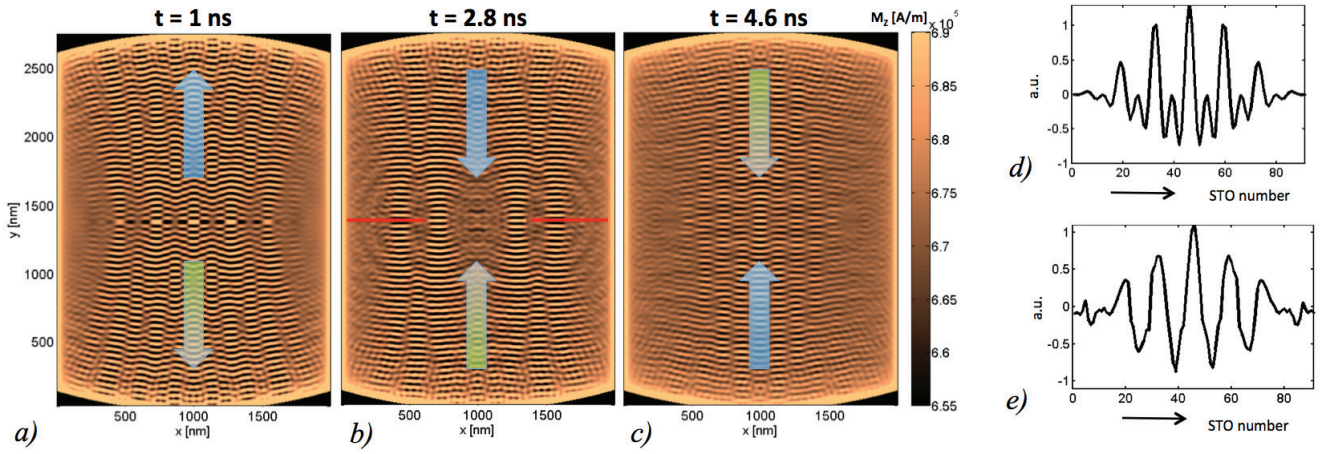


Fig. 7. Signal filtering by the double mirror system. The filter (indicated by red lines in b) forming a low-pass filter) is applied in the focal plane after the first reflections, i.e. while the Fourier transform is formed. d) the input signal e) the filtered signal (high frequency component removed). [9]

the waves propagate two directions symmetrically, and reflects back to form the Fourier transformation. After this they cross the focal plane, and propagate towards the opposite mirror just as they would be a new input. As a result, after the second reflection of the waves the second Fourier transformation of the input signal forms in the focal plane, which is the original signal. Fig. 8 shows the result of the second Fourier transformation compared to the input vector.

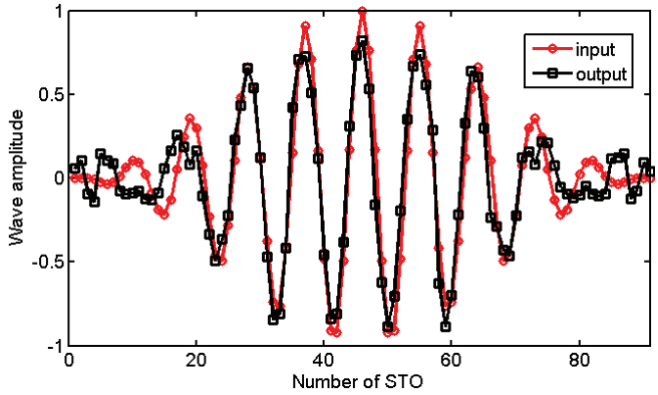


Fig. 8. The input waveform and the output wave amplitudes in the double mirror system after the second reflections. The good match indicates the systems capability for more complex signal processing. [8]

IV. Filtering a signal in Fourier space

As an example of a more complex operation realized by a spin-wave based system we demonstrate a filter device operating in a 99 element analog input vector. Using the double mirror arrangement one can apply active damping as a filter in the focal plane while the first Fourier transform is forming and thus absorb some components in the Fourier space. In Fig. 7b the red lines represent these filters. Everything else is the same as in the simple double mirror case, but here the output waveform is filtered as it can be seen in Fig. 7d,e.

V. Conclusion

We demonstrated basic primitives of an optical like spin-wave based computing system. We designed a simple filter as an example of the concept, and verified our designs by micromagnetic simulations. The designed system has potentially low power consumption while high operation frequency in a small chip area, and compatible with CMOS. It might have advantages over the present filter systems in areas where high throughput required but the operations can be divided to work on short (few hundred elements) and low precision vectors.

Acknowledgment

I would like to thank Dr. György Csaba and Dr. Wolfgang Porod for their help as my advisors at the University of Notre Dame, Dr. Árpád Csurgay for his support as my advisor at the Pázmány Péter Catholic University. I am also grateful for stimulating discussions with G. Bourianoff (Intel Corp.).

References

[1] T. Schneider *et al.*, "Realization of spin-wave logic gates," *Applied Physics Letters*, vol. 92, 2008.
[2] G. Csaba, A. Papp, and W. Porod, "Spin-wave based realization of optical computing primitives," *accepted to Journal of Applied Physics*, 2014.
[3] D. G. Feitelson, *Optical computing: a survey for computer scientists*. MIT Press, 1988.
[4] S. Cherepov *et al.*, "Electric-field-induced spin wave generation using multiferroic magnetoelectric cells," *Applied Physics Letters*, vol. 104, no. 8, Feb. 2014.
[5] M. J. Donahue and D. G. Porter, *OOMMF User's Guide, Version 1.2a3*. National Institute of Standards and Technology, Gaithersburg, MD, 1999, interagency Report NISTIR 6376. [Online]. Available: http://math.nist.gov/oommf/
[6] A. Papp, G. Csaba, and W. Porod, "Non-boolean computing using spin waves," in *16th International Workshop on Computational Electronics*, Nara, Japan, June 2013, pp. 78–79.
[7] S. E. Russek *et al.*, "Spin-transfer nano-oscillators," in *Handbook of nanophysics: Functional Nanomaterials*, K. D. Sattler, Ed. Addison Wesley, 2010, p. 10.
[8] A. Papp, G. Csaba, G. I. Bourianoff, and W. Porod, "Spin-wave-based computing devices," in *IEEE International Conference on Nanotechnology*, Toronto, Canada, 2014, accepted for publication.
[9] ——, "Signal processing using spin-wave interference," in *International Conference on Superlattices, Nanostructures and Nanodevices*, Savannah, USA, 2014, submitted for publication.

Program 4

Human Language Technologies, Artificial Understanding, Telepresence, Communication

Head: Gábor Prószéky

Towards a Psycholinguistically Motivated Performance-Based Parsing Model

Balázs Indig
(Supervisor: Gábor Prószéky)
indig.balazs@itk.ppke.hu

***Abstract*—In this paper I present a new paradigm and framework for syntactic and semantic analysis, which is based on the following principles: (1) *Psycholiguistically* motivated means, my model hold on to the inner workings of human language processing as much as possible. (2) As a *performance-based* system the model tries to process all natural language utterances (a few sentences long coherent texts) that occur in (written) texts. My focus is not the handling of theoretically existing but in practice rather rare phenomena. Instead, I consider and try to interpret any – no matter how badly formed or agrammatic – text that appears as a natural language utterance. (3) In my model the parser *processes the text strictly left-to-right incrementally* and does not utilize or reference the parts that succeed the current position. (4) The general architecture of the parser framework is naturally *parallel* from the beginning, in contrast to the traditional approach, where the analysis is generated at the end of a pipeline of modules. Here the program processes the actual word using parallel threads (a morphological analyser thread, a corpus statistics thread, etc.). These threads analyse each word together at the same time and communicate to correct each other's errors and to make a final decision in the analysis. (5) The framework's processing and representational units are not individual sentences, rather *utterances* consisting of one or more sentences. This enables the unified handling of intra- and intersentential anaphoric relations. (6) In accordance with the principles described so far, in order to be able to handle all the different phenomena at the same time the *representation* is not necessarily a tree, but a connected graph containing different types of edges. After describing its theoretical foundations, I present a *pilot implementation* of the framework. My pilot program illustrates the basic principles and performs some of the analysis steps conforming to the theory.**

***Keywords*-psycholinguistically-motivated; performance-based, incremental, semantic parsing**

I. Introduction

In this paper I present the foundations of a performance-based parser framework, that attempts to process real language data and tries to stay close to human language processing as possible. I process text simultaneously with the generation of the utterance (i.e., strictly left-to-right) and try to use all the information that is necessary for its interpretation, even if it is not well-formed in the traditional grammatical sense. The parser therefore is limited to *assume* some not yet available parts, based on those which are heard or read until the utterance has come to an end. This does not mean that there are no utterances which "circumvent" the most likely analysis, sometimes forcing even the human parser to backtrack, but in everyday communication humans seem to follow the maxims of Grice [1], and avoid these constructions. They typically occur in jokes or intentional misrepresentations[1]. I cannot use a traditional PoS-tagger which utilizes global information about all elements of the sentence when determining the tags of an incomplete segment. I'm using an n-gram model based on data acquired from a large corpus, which can provide probabilistic information about the possible tags of the current word based on the *preceding* words only. It also continuously yields information on how typical the current span of text is, and what kind of units are expected or required at the current position. I need larger parts of text as I attempt to uniformly handle ellipsis, anaphora/coreference and coordination which may or may not span beyond sentence boundaries. My basic unit is not the sentence but an utterance/paragraph (few sentences long coherent texts). Similarly to Discourse Representation Theory [2], in our group, we construct a unified semantic representation containing all entities referred to in the processed text but in contrast to discourse representation structures, our representations are "ontologically promiscuous" in the sense of [3] and contain reified abstract entities such as eventualities and propositions in order to handle logical complexity. I employ traditional tools/modules (morphological analyser, identification of verbal and other constructions, corpus frequencies, ontologies etc.) as so-called **resource threads** constantly running and working concurrently. This allows the different tools to communicate, complement or even override each other at every parsing step in order to correct each others errors. The architecture of the analyser therefore is parallel instead of the traditional pipeline design. To store information provided by words already processed, I have a second kind of threads, the so-called **structure threads**. These threads are initiated and closed by particular words, and they realize different "offers" and "demands". To operate this offer-demand system I need a certain description of the main phenomena of the language, namely such a grammar which enumerates all possible roles for linguistic units (e.g. a noun in nominative case or a comma). The output of my parser will contain syntactic and semantic information to identify both participants and events, building a representation of the whole utterance, and formulating statements about who does what, where and when. My offer-demand mechanism and Combinatory Categorial Grammar are both strongly lexicalist,

[1] In my group, we have begun to examine large corpora to identify these complicated constructions that are in the focus of modern grammatical theories and to collect information about the frequency of their occurrences to prove that they are quite rare in everyday life.

B. Indig, "Towards a Psycholinguistically Motivated Performance-Based Parsing Model,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 133-136.

proposing that offer-demand or functor-argument construction processes are governed by deep lexical information rather than by an extra set of syntactic rules. Hungarian being a configurational language, I must deal with arbitrarily ordered offer-demand pairs in the case of verb arguments, something that is dealt with in CCG for example by relaxing category definitions [4].

II. Background

Generative models do not provide an effective solution from the computer's perspective to analyse texts. This is partly due to the fact that these models are concerned primarily with ideal speakers-listeners and focus on sentences rather than larger text segments [5]. In addition, the notion of "efficient parsing" does not belong to competence, but to performance. The term *performance-based* means that everything that actually occurs as text needs to be processed. Constructions that could theoretically exist but in practice do not are less important. Therefore, I investigated the most important performance-based parsers [6] and those theoretical approaches which state that efficient parsing belongs to the area of competence (e.g. [7]). Several efficient dependency parsers exist, although many of them are based on theories which do not agree with our concepts: MaltParser [8], Stanford Parser [9], finite-state dependency parser [10], which are in fact aimed at the description of the relationship between different linguistic units, but are heavily tied to the separate processing of subsequent sentences. A dependency description for Hungarian was developed in the late 1980s for the DLT system [11]. Another important grammatical source for Hungarian, Szeged Treebank is also available in dependency format [12], from which a statistical dependency parser was also created [13]. I think that many phenomena of language can not be described properly with dependency relations. The MetaMorpho rule-based parser [14], the most thorough Hungarian parser to date is also at our group's disposal and I am currently using its verb frame constructions along with dependency descriptions of noun phrases. None of the mentioned parsers and resources handle the task of disambiguation properly and all have bad fault tolerance: a single out-of-vocabulary or unusual word may result in the failure of the entire analysis. I also found that all the presently known parsers do one-way, pipelined processing almost exclusively, i.e. there is no communication back and forth between the different processing levels, a tool begins to operate only after the previous tool completed its task, and takes the input over "as is" from the previous tool together with the possible errors. Therefore, in our parser we abandon the idea of the traditional pipeline architecture. One of the reasons is that errors of the lower level modules in a pipeline are carried over to higher levels without correction and get amplified later and weaken the quality of later modules. Typically, parsers try to address this problem by applying a simple frequency based filter, but this solution eliminates unusual analyses even when they would actually be the correct ones. As Prószéky points out [15], decisions taken during the human analysis of linguistic structures can override the lexicon. All the aforementioned parsers are using exclusively their preliminary knowledge in their decisions. In my parser, we intend to take an approach in which the unusual structure of the actual input does not conflict with either preliminary statistics or the often misleading output of mechanistic rules. My initial hypothesis is that in the listener's/reader's mind two systems exist: one which relies upon the learned structures and another that makes current decisions and is able to perform real-time processing even if the learned structures contradict each other in their grammaticality (e.g. contain mismatching features). Therefore, my parsing algorithm is looking for a kind of consensus between different pieces of linguistic knowledge [16]. Thus, as in human language processing, the analysing and interpreting modules work in parallel and in close cooperation in my parser [17].Current competence-based models do not assume any cooperation with non-linguistic information processing systems. I can say, however, that the performance can not be separated from other cognitive processes which have impact on language, therefore, from the very first step of analysis the parser relies on the simultaneous handling of some linguistic and non-linguistic modules (world knowledge, mood, etc.) on various levels, depending on how fine-grained the model is. In the grammar, various levels of processing (morphological analysis, identification of verbal constructions, corpus frequencies, world ontologies, etc.) work concurrently in separate resource threads in the background and can complement or even override each other at every parsing step.

III. Architecture

Basically, two language element-initiated thread types seem necessary. An *offer* thread provides information on the current element (e.g. an element is in nominative case), and a *demand* thread is looking for a required element with a specific property (e.g. a possession noun looks for its possessor, a postpositional particle needs a noun, a determiner seeks the NP head, a transitive verb needs its object etc.) To overcome the problem of rejecting non-frequent but valid constructions I propose an extralinguistic decision system, which considers relevant information from the corpus frequency thread on the one hand (how well the analysis corresponds to usual patterns), and rule-based analysis on the other hand. As I process the actual word I consider the frequency relationships between the word and the preceding words. For example, after the word *esik* (to fall) I expect *sz* (word) as subject because it is relatively frequent (15% of all possible subjects for *esik*). After this step, I expect a noun phrase with the case inflection suffix *-rl* because it is very frequent (appears in 9 cases out of 10 after *esik sz*), forming an idiom meaning something is talked about. Based on the processing efficiency of the human parser our framework is trying to avoid the combinatorial explosion, so it uses aggregated statistics as "preliminary knowledge" mined from corpora. Frequent constructions are used as whole units without real-time analysis and are inserted into the representation. This method is called *caching* in IT, but in psycholinguistics is also well-known

in the sense of human language processing, and are called "**Gestalts**". The desired output is a network of semantic relations built from the underlying text, which can answer questions, and can generate statements which are contained only implicitly in the original text. I favour the use of **non-tree-based** dependency graphs as coreferences and relative pronouns are not necessarily realized in the same sentence and marking them with an edge can ruin the tree structure producing a directed acyclic graph. To cope with **ellipsis phenomena** I allow the threads to run across sentences. I argue that analysis should not stop at the sentence level, as utterances are the natural units in human communication. The topics of subsequent sentences can be identical, and therefore in natural human communication can be and are in fact mostly omitted (ellipsis phenomena), which causes most parsers to lose track of the analysis. It is very important to identify the participants in the text and to determine their **coreference** relations (which entities in the text refer to the same entities in the real world). In other words, which participants are "new" at the point of their introduction in the text and which refer to already mentioned ones, and what kind of relationship holds among them. Essentially all nouns and nominative or accusative phonologically empty arguments inferred from verb suffixes can be considered as **event participants**. According to the principles of **neo-Davidsonian event semantics** [18] even events (verbs) may be participants, as I can refer to them as well. In order to aid the resolution of such references, all the participants in all the preceding sentences are stored in our system. Further threads utilize the rich descriptions of **lexical units** and **lexico-syntactic constructions** available from the MetaMorpho parsers databases. An "offer" type thread annotates units with semantic features (such as *animate, human, abstract* etc.), available for 118,000 words and multi-word expressions. On the other hand, a "demanding" type thread, by using 35,000 of MetaMorphos open construction rules proposes connections between verbs and their possible arguments and nouns or adjectives and their arguments (e.g. *hostility towards something, interested in something* etc.)

I am also experimenting with methods to aid the prediction of verb-argument connections based on verb-noun corpus co-occurence data from the Verb Argument Browser [19] and ontological information [20]. In our group, we are working on a method to map verb arguments from the corpus to Hungarian WordNet in order to identify generalized semantic classes that correspond to the verb arguments **semantic types**.

IV. Current Work

A. Practical problems

As a first step of the approach, I have identified and implemented some formalized operation types to be used during processing. First, I had to find those elements which predict what sort of elements can come after them. For example, a determiner predicts some nominal element at the end of the construction starting with the actual determiner. There are elements which fulfil an earlier prediction, e.g. a verbal argument fulfils a slot in the argument frame of a verb that occurred earlier in the input stream. If the verb itself comes later than one of its arguments, this argument construction can fulfil automatically the right slot in the verbal frame. There are other sorts of operations: conjunction structures, for example, can be identified only when a conjunctive element arises. It can be an 'and', an 'or' or a comma: they introduce the next element of the conjunctive construction. When the system identifies an element like this, it should modify the representation of the previous element, making that element the first one in a conjunctive structure. I treat a conjunction as one unit, without deciding whether it has a head or not. A working pilot implementation of our approach attempts to deal with frequent, fundamental phenomena that are not necessarily easy to handle in other frameworks, such as linking a separated verbal prefix and the verb stem, linking parts of possessive constructions, identifying enumerations/coordination as complex units, identifying the actual role of a comma (whether it triggers e.g. another clause, an enumeration, a parenthetical/interjection or an apposition), or identifying the scope of a negation. Currently in our group, we are working on the description and formalisation of further such important linguistic phenomena in Hungarian, taking into account the constraint of left-to-right incremental processing, concerning for example exocentric constructions where the phrase does not have a head, which is a challenge for dependency parsers.

B. Implementation

My prototype implementation proceeds left-to-right taking the elements (currently words) of the input text one by one. It processes the subsequent word taking all the information provided by the resource threads into account, then either (a) *stop*s or (b) *start*s some threads or (c) *leave*s some threads unchanged or (d) *creates an edge* in the representation graph. One kind of information I use are morphological patterns to identify the features of words that will be crucial in the syntax: suffixation information, parts-of-speech and word lemmas. The aforementioned operations are currently triggered by these morphological patterns, which the actual word matches. I allow multiple morphological patterns to match a single entry where necessary, as each word suffix can contain multiple valuable information (e.g. possession and case), then all of the corresponding structural threads must start.

Each element, that has – or introduces some later element that will have – reference is noted in the list of participants. In this list, each mention of a participant is linked with coreferring mentions. This includes the special cases of pronominal verb arguments that may or may not be present in Hungarian, since the inflection on the verb carries enough information to identify their number and person. For example, the 3rd person singular form of a finite verb introduces a participant that may or may not have surface realisation in the sentence, so at the time of processing the verb the parser introduces a new participant marked as "phonologically empty". If a surface realisation of this participant actually appears in the input at a later point, the parser links it to the formerly introduced participant.

Currently the parser builds a tree over the sentence units using the recognised dependency relations combined with verb frame constructions. In the future, we will add long distance dependencies and coreference relations that will turn the representation into a less strict graph structure that is no longer a tree.

My prototype has been tested on summaries of **news articles** taken from a Hungarian news portal's RSS feed (www.inforadio.hu). The 2-3 sentence long paragraphs describe single political or economic events. Their syntactic complexity is close to what I would like to model, thus they serve as an adequate input to our parser. First, the input text is lemmatised as a pre-processing step (playing the role of a simple lexical lookup in other, less inflecting languages).

V. Future work

Among the aformentioned future development goals, some of the presented tools are not production ready, they still need to be developed for integrating in the parser program. While the parser's internal representation is still not stable I'm working toward stabilizing it and keep the way open for possible extensions for other languages for example English.

VI. Conclusion

In this paper, I described ongoing research which aims at a psycholinguistically motivated, performance-based, strictly left-to-right, utterance-based parallel processing parser framework. In my view, currently existing parser models are not sufficient at modeling such aspects of human language processing. I laid down the fundamental principles of our linguistic theory and presented some details of a pilot parser implementation.

Acknowledgment

I would like to acknowledge my supervisor, Gábor Prószéky and my collegues Bálint Sass and Márton Mihály for his kind help and his knowledge on this multidisciplinary linguistic field. Also the support TMOP-4.2.1.B 11/2/KMR-20110002 and TMOP-4.2.2/B 10/120100014 is kindly acknowledged.

References

[1] H. P. Grice and G. Harman, *Logic and conversation*. Encino:Dickenson, 1975.
[2] H. Kamp and U. Reyle, *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*. Springer, 1993.
[3] J. R. Hobbs, "Ontological promiscuity," in *Proceedings of the 23rd annual meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 1985, pp. 60–69.
[4] M. Steedman and J. Baldridge, "Combinatory categorial grammar," in *Non-Transformational Syntax*, R. Borsley and K. Börjars, Eds. Blackwell, 2011, pp. 181–224.
[5] N. Chomsky, *Syntactic structures*. The Hague:Mouton, 1957.
[6] H. Bunt, P. Merlo, and J. Nivre, *Trends in Parsing Technology*. Dordrecht:Springer, 2010.
[7] B. L. Pritchett, *Grammatical competence and parsing performance*. University of Chicago Press, 1992.
[8] J. Nivre, *Inductive dependency parsing*. Springer, 2006.
[9] M.-C. De Marneffe, B. MacCartney, and C. D. Manning, "Generating typed dependency parses from phrase structure parses," in *Proceedings of LREC*, vol. 6, 2006, pp. 449–454.
[10] K. Oflazer, "Dependency parsing with an extended finite-state approach," *Computational Linguistics*, vol. 29, no. 4, pp. 515–544, 2003.
[11] G. Prószéky, I. Koutny, and B. Wacha, "A dependency syntax of Hungarian," *Metataxis in Practice (Dependency Syntax for Multilingual Machine Translation)*, pp. 151–181, 1989.
[12] V. Vincze, D. Szauter, A. Almási, G. Móra, Z. Alexin, and J. Csirik, "Hungarian dependency treebank." in *LREC*, 2010, pp. 1855–1862.
[13] J. Zsibrita, V. Vincze, and R. Farkas, "magyarlanc: A toolkit for morphological and dependency parsing of hungarian," in *Proceedings of RANLP*, Hissar, Bulgaria, 2013, pp. 763–771.
[14] G. Prószéky, L. Tihanyi, and G. Ugray, "Moose: A robust high-performance parser and generator," *Proceedings of the 9th Workshop of the European Association for Machine Translation*, pp. 138–142, 2004.
[15] G. Prószéky, "Számítógépes morfológia," in *Morfológia (Strukturális magyar nyelvtan III)*, F. Kiefer and Z. Bánréti, Eds. Akadémiai Kiadó, Budapest, 2000, vol. 3, pp. 151–1064.
[16] P. Csaba, *Mondatmegértés a magyar nyelvben*. Osiris Kiadó, Budapest, 1999.
[17] C. Valéria, "Az olvasó agy," *Akadémiai Kiadó, Budapest*, 2006.
[18] P. Terence, "Events in the semantics of English: A study in subatomic semantics," 1990.
[19] B. Sass, "The Verb Argument Browser," *11th International Conference on Text, Speech and Dialog (TSD)*, pp. 187–192, 2008.
[20] M. Mihály, C. Hatvani, J. Kuti, G. Szarvas, J. Csirik, G. Prószéky, and T. Váradi, "Methods and results of the Hungarian WordNet project," in *Proceedings of the Fourth Global WordNet Conference (GWC-2008)*, T. A., C. D., V. V., C. Fellbaum, and P. Vossen, Eds. Szeged, University of Szeged, 2008, pp. 311–321.

Corpus driven research: ideas and attempts

István Endrédy
(Supervisor: Dr. Gábor Prószéky)
endredy.istvan.gergely@itk.ppke.hu

***Abstract*—Many areas of human language technologies need large corpora. The bigger is the better, but it is a rather expensive work to build a huge corpus. On the other hand, web is free and its content is growing every second. If we build a good quality corpus based on the web, it gives the benefit of the even increasing size and the opportunity as well, to explore the frequent structures of the given language. This paper describes a method to detect noun phrases in sentences. The process is corpus driven: ideas are drawn from the corpus and they are checked also in it. This task is a part of a greater project, an overview will be given about its state.**

***Keywords*-corpus building; web crawler; noun phrases;**

I. The big picture

It would be a great step in the history of humanity if machines could understand natural language. It could change the usage of computers, the quality of machine translation and the search methods as well.
Perhaps this vision is far-fetched, but even a 8,000 mile journey consists of small steps. My research is a small step on this road: it deals with the automatic detection of frequent structures of a language.
A great corpus contains a lot of information about the world, the people and even the language itself. Most frequent noun pairs show for instance the famous or popular persons in the given age of the text.
Nowadays the greatest and free corpus can be the web. My research is about to create huge corpus from the web (1), detect the most frequent patterns of noun phrases and sentences (2), and prove the following hypothesis: the number of really used sentence structures is final.

I have worked in the area of language technology as a developer for 12 years. Projects from these years have given experience in data mining, stemming and searching in corpus. This research is organically connected to these.
Our research group (MTA-PPKE Hungarian Language Technology Research Group) works on a grammatical analyser with brand new approaches. One of the innovations of this project is strongly connected to frequent patterns: if a known structure is detected in the text then analyser should use the precalculated analysis of the structure, just like humans. The text processing of an adult and a three-year-old mainly differ in the number of the known patterns. The adult knows more one, hence (s)he can process text faster. This is the basic idea of our would-be analyser, it should work in a similar manner to humans.
This purpose will be helped with the corpus from the web, which is created by my software. The detected frequent structures will be added to the "world knowledge" of the analyser which will be described later.

This project may help to create a better grammatical analyser, and the experiences might improve the quality of other language processing tools tools: (i) HuMor morphological analyser [1], (ii) PurePos tagger[2], (iii) web crawler and (iv) Gold Miner [3] text extracting algorithm.

II. Why do we need a new corpus?

There are various large Hungarian corpora available for the researchers. One of the biggest made 10 years ago at BME MOKK [1] relies on 600 million words. The Hungarian National Corpus [2] with its 190 million words is somewhat smaller, but its consciously selected content is fully POS-tagged. Nowadays we feel that there would be a need for a large, comprehensive, annotated and updated huge database.

III. State of the crawler

In the first part of the project a crawler was written in c++, which is able to download many pages, and extract the main part of the pages. The extraction is called boilerplate removal, which means to eliminate header, footer, menus, ads from web page, only the valuable article is saved into the corpus.
In the second step was to develop a better boilerplate removal algorithm, which is able to make clean, duplicate free corpus, without any advertisements and other unwanted content. The name of the developed algorithm was Gold Miner, and it was introduced in Mexico [3].
The crawler was started in 2013, and it runs all the time. The size of its corpus is 53 million words.
Because of its size, the database of the crawler has to be migrated into MySQL database engine.

IV. Ideas

There is an open linguistic question: the number of the sentences' structures is finite or not. The bigger is a corpus, the more chance we have to answer it. The problem will be opened at the end of article, too, but we can get closer to the solution.
The variability of a sentence is mainly depends on the number of its noun phrases (NP).
The patterns of NPs and sentences are observed in a big corpus. If the frequency of a certain pattern is above a predetermined threshold, it will be regarded as a known pattern. Then, every known NP will be replaced as "NP" in the sentence,

I. Endrédy, "Corpus driven research: ideas and attempts,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 137-140.

therefore the sentence structure complexity will be decreased significantly (demonstrated in Table I).

TABLE I
NP DETECTION EXAMPLE

Egyedi elbírálást kér a kormánytól a károk enyhítésénél az árvízsújtotta Felsőzsolca önkormányzata.
(Egyedi elbírálást) kér (a kormánytól) (a károk enyhítésénél) (az árvízsújtotta Felsőzsolca önkormányzata).
NP-t kér NP-től NP-nél NP.

It helps the structure detection and processing of a sentence. Of course a rare pattern may occur in a text, and it most probably will give extra analysing task to the analyser. But in most cases the analyser will work faster and more accurately: it will know many patterns.

The first attempt was to extract minimal[4] NPs [5] [6]. In our attempt it was made by simple n-gram detection. Its results can be seen on Table II.

TABLE II
EXAMPLE N-GRAMS FROM CORPUS

Pattern: [noun] [noun] [adjective] [noun]
Matolcsy György nemzetgazdasági miniszter (694)
Barack Obama amerikai elnök (664)
Sólyom László köztársasági elnök (367)
Angela Merkel német kancellár (345)
Nicolas Sarkozy francia elnök (256)
Schmitt Pál köztársasági elnök (229)
Vlagyimir Putyin orosz elnök (186)
Pattern: [det] [noun] [verb]
A szóvivő elmondta (660)
A szakember elmondta (480)
A miniszter közölte (320)
A bíróság megállapította (29)
Pattern: [noun] [noun]
Orbán Viktor (8181)
New York (4085)
Wall Street (1358)
Harry Potter (691)
Johnny Depp (645)
Angelina Jolie (627)
Puskás Ferenc (380)

This simple n-gram searching approach showed us that corpus contains many information about the world: famous people and places, typical activities of a given role are all coded into the corpus. This fact inspirated to use this resource somehow, to make better analyser.
But of course every N-gram is not an NP, the most frequent NPs from the n-grams can be selected only by humans. Another approach was needed.

V. NOUN PHRASE DETECTION

A rule based maximum NP searching algorithm was built, which is works on part of speech tagged sentences. Some of its rules is demonstrated on Table III. It can detect great NP even with 20 (!) words.

TABLE III
NP DETECTION RULES

[Macro definitions] *# symbols on the right side will be replaced with the left one* *# "KOT:hogy" means where surface="hogy" AND category is "KOT"* THAT=KOT:hogy. BECAUSE=KOT:mert NAND=KOT:!és WHO=FN—NM:Ki FN=FN—NM MN=NM—NM SZ=SZ—NM POINT=PUNCT:.
[General pattern rules] *# NPStart: these categories start an NP (begining of a sentence is included).* NPStart=DET,FN,MN,SZN,NN_OK,!NN_BAD,!PUNCT *# NPEnd: these categories end an NP (end of a sentence is included)* NPEnd=THAT,KOT,NAND,WHO,POINT,NN_BAD,IGE, MN—NM,HA—NM,!_MIB,!_OKEP *# NPLastItem: these categories should be the last item of an NP* NPLastItem=FN,SZN,MN,!_ESSMOD

Three corpora (inforadio.hu, mno.hu and a book) were ran through NP detection algorithm, and an online web application was developed to be able search and list them as it is shown in Figure 1.

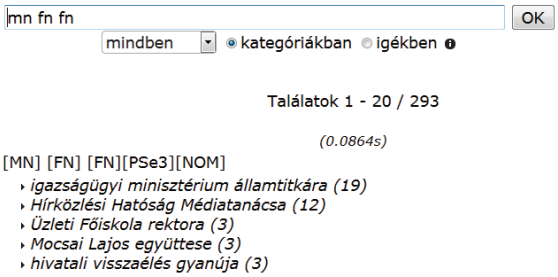


Fig. 1. Online NP search application

The application has two main features. On the one hand NP patterns can be listed based on input patterns. For example adj noun noun will list every NP which has this pattern. On the other hand sentences can be listed based on verb: it will give the simplified sentences, each NPs are reduced to a single "NP" string (Figure 2).
In other words you have also the possibility to search given structure of NPs or to list sentences with NP related to a verb (Figure 2). This option is very similar to Mazsola system [7], but this tool give whole sentences but Mazsola gives a verb frame [8].

The first feature gives the opportunity to redefine the categories of the word analyser: the errors of the NP results might suggest the missing or incorrect part of speech categories. For instance names or professions are nouns as well, but they play different role in a noun phrase. This may inspire to make new part of speech categories.

The precision of the algorithm was decreased by NPs which were splitted into two NP instead of one. Mainly it caused by *"és"*, as it is shown in Table IV.

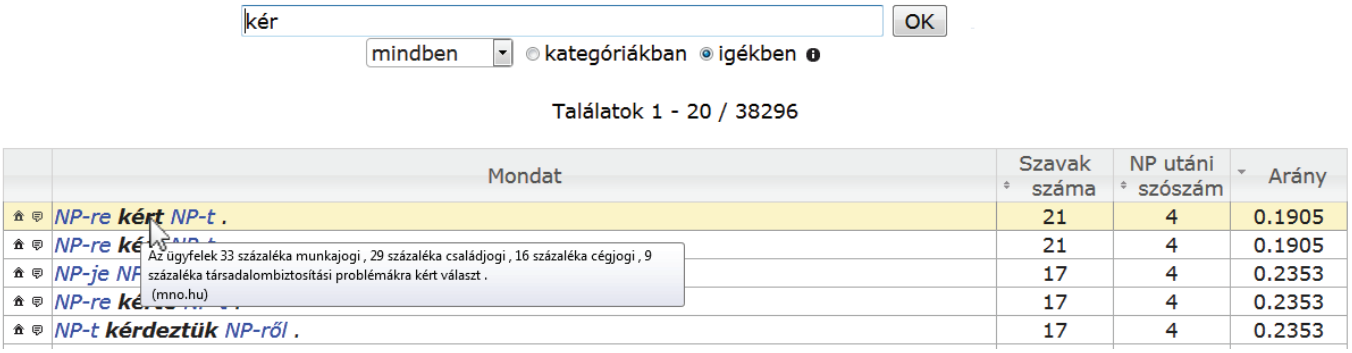


Fig. 2. Online NP search in sentences - without NPs

This problem can not be solved with only syntactic methods. Algorithm has to ”know” that the two NPs are related, and they form together a single NP. How can we implement and use semantic information to help NP detection?

TABLE IV
EXAMPLE FOR A TYPICAL NP DETECTION ERROR

Aláírják (a finanszírozási szerződést) (a Budapesti Közlekedési Központ igazgatósága) és (a Fővárosi Közgyűlés jóváhagyásával).
Aláírják NP-t NP és NP-vel. Aláírják NP-t NP-vel.

VI. SOLUTION: SEMANTIC INFORMATION FORM CORPUS

The algorithm would have to “understand” the sentence to be able to recognise that last two NPs are actually one NP. It has to know that the two entities are closely related. The idea was to extract semantic information from corpus itself. The following patterns are extracted from the corpus:

```
 NP1 és NP2
NP1 valamint NP2
NP1 blabla1, NP2 pedig blabla2
```

It is assumed that the heads of these NPs are in the same domain, in other words they are related. A small piece of text (75M) gave 5,500 word pairs which forms the basis for a small world knowledge. This pairs can be used in NP algorithm: if two NP is connected with és and their heads are pairs in this database, then we can assume that they are related (Table V), what is more, they compose one NP.
In other words the corpus helps to detect NP in a better way. The greater is the corpus the better will be the NP detection quality.

The related words extracted from corpus can be shown as a graph with the weight of their connection’s frequency. A small piece of the graph is demonstrated on Figure 3.

TABLE V
EXTRACT SEMANTIC INFORMATION FROM CORPUS

IMF	EU
fagylalt	jégkrém
bér	nyugdíj
munka	kenyér
délután	reggel
növekedés	foglalkoztatás
...	...

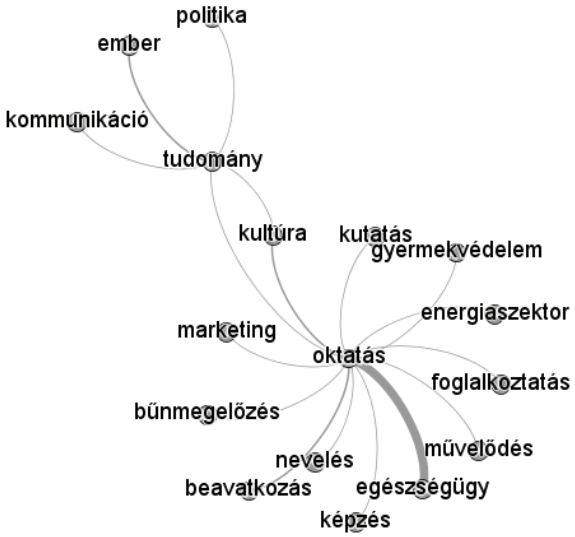


Fig. 3. Graph of related words extracted from corpus

VII. FUTURE PLANS

The next step in the project will be the extraction of huge number of related NP patterns V. This world knowledge will help to make better decision in NP detection in ambiguous cases (IV). We hope that the NP detection quality will be increased by this improvement, pursuing this further the structures of sentences will be more simple and cleaner.
This will help to get closer to the basic question: the number of really used sentence structures is final or not.

VIII. Conclusion

A method was introduced in this article which helps to detect maximum NP in a text. The recognition is ruled based, but it is improved by world knowledge extracted from corpora. It is language independent, it is based and driven by corpus, and made by automatically without any human intervention.

Acknowledgment

I would like to express my gratitude to Attila Novák and Gábor Prószéky for their constructive ideas, encouragement and patient suggestions during this research.

References

[1] G. Prószéky and B. Kis, "A unification-based approach to morpho-syntactic parsing of agglutinative and other (highly) inflectional languages." in *ACL*, R. Dale and K. W. Church, Eds. ACL, 1999. [Online]. Available: http://dblp.uni-trier.de/db/conf/acl/acl1999.html#ProszekyK99

[2] G. Orosz and A. Novák, "PurePos 2.0: a hybrid tool for morphological disambiguation," in *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2013)*. Hissar, Bulgaria: INCOMA Ltd. Shoumen, BULGARIA, 2013, p. 539–545. [Online]. Available: http://aclweb.org/anthology//R/R13/R13-1071.pdf

[3] I. Endrédy and A. Novák, "More effective boilerplate removal—the goldminer algorithm," *Polibits*, vol. 48, pp. 79–83, 2013. [Online]. Available: http://polibits.gelbukh.com/2013_48

[4] G. Recski and D. Varga, "A hungarian np chunker," *The Odd Yearbook*, 2009.

[5] A. Z. Gabor Recski, András Rung and A. Kornai, "Np alignment in bilingual corpora," in *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, B. M. J. M. J. O. S. P. M. R. D. T. Nicoletta Calzolari (Conference Chair); Khalid Choukri, Ed. Valletta, Malta: European Language Resources Association (ELRA), may 2010.

[6] G. A. Recski and D. Varga, "Magyar főnévi csoportok azonosítása," *Általános Nyelvészeti Tanulmányok*, pp. 81–95, 2012.

[7] B. Sass, "Extracting idiomatic hungarian verb frames," in *FinTAL*, ser. Lecture Notes in Computer Science, T. Salakoski, F. Ginter, S. Pyysalo, and T. Pahikkala, Eds., vol. 4139. Springer, 2006, pp. 303–309.

[8] B. Sas, "The verb argument browser," in *TSD*, ser. Lecture Notes in Computer Science, P. Sojka, A. Horák, I. Kopecek, and K. Pala, Eds., vol. 5246. Springer, 2008, pp. 187–192.

Experimental Investigation of Transitions for Mixed Speech and Music Playlist Generation

Mátyás Jani
(Supervisor: György Takács)
jani.matyas@itk.ppke.hu

***Abstract*—In the case of a costumizable, Internet-based radio program, beside a good selection of content, the acoustic and semantic quality of the transitions between pieces of audio material also play a significant role influencing the listening experience. Our goal is to automate radio program editing for individually customized (Internet) radio programmes, containing both pieces of speech and music.**

This paper deals with the acoustic features of speech to music transitions. A set of audio test data – pairs of speech and music – was selected and subjective opinion test on the quality of the transition was conducted. The analysis aimed at predicting the subjective acoustic quality of a speech to music transition based on automatically extractable features of speech and music. The attributes found to be significantly influencing the acoustic quality are related to the tempo and Mel spectrum values. The numeric quality of the match can be predicted with a cross-correlation of 0.55 and the digitised quality with a performance of 75%.

***Keywords*-playlist generation; speech; music; audio features; data mining**

I. Introduction

Current Internet technologies allow transmitting individually customized media content. The customisation can be based e.g., on users' general interests, their social networks and their current context. Rapidly spreading smartphones lower the access barriers and increase the significance of individually customized media content.

In case of audio media the customisation of the content is called playlist generation, the aim of which is to assemble a good playlist, i.e., a list of audio materials – currently exclusively pieces of music – that are meant to listen to sequentially. Playlist generation is a complex task and considers criteria on three levels: the selection of the pieces of music (similarly to recommender systems), the pairs of music that follow each other (e.g., smooth transitions) and the playlist as a whole (e.g., coherence, diversity).

Current approaches for playlist generation are limited to music. However, in customized online radios or innovative, voice-based systems, such as in audio social media [1] the handling of spoken content, e.g., news, reports, interviews, discussions (in the following simply referred to as speech) is also needed. For mixed speech and music playlist generation, a number of additional challenges arise. The most important one is that the acoustic quality must ensure a good listening experience, most importantly regarding consecutive pieces of speech and music. Additional issues also arise, e.g., the selection of single pieces of speech, the semantic relationships between different pieces of speeches or criteria concerning the whole playlists.

This paper focuses on the issue which we consider the most important one in mixed speech-music playlist generation: the matching of consecutive pieces of speech and music. The quality of the match is affected by the acoustics and the content of the speech and music parts. We expect the first, the acoustics to be more important, as it is more directly related to emotions, a central issue in playlists, and more fundamentally rooted in humans than facts or reasoning. Therefore, in this paper, we concentrate on the acoustics of speech to music transitions.

The goal of the paper is to predict the acoustic quality of speech to music transitions based on features that can be automatically extracted from the music and speech signals. For this, a set of audio material was prepared, the quality of the matches was measured with the help of a group of test persons, and acoustic features were extracted from the audio materials. The data collected from these sources – audio features of the pieces of music and speech vs. the quality of the match – was analysed by means of data mining methods.

In this paper the continuation of our previous works [2], [3] is presented. The results were published in the Internation Conference on Multimedia Retrieval [4].

II. Related Work

In the literature, some initial work can be found on sequential recommender systems that consider not only the set of selected items but also arrange them sequentially. In particular, [5] is concerned with general aspects of human reception, not specifically with pieces of audio material. In [6] a sequential recommender for video is introduced, however, the emphasis is on other aspects, such as implicit feedback and dealing with the cold-start problem, rather than on our focus, i.e. a smooth listening experience.

In audio and video media, automatically assembling the playlist is called playlist generation. The area has gained significant research and industrial interest over the last decade. A good entry point can be found e.g. in the dissertation [7] and tutorial [8] of Ben Fields (et. al.). In playlist generation, besides other aspects (such as the selection of pieces of music), also the matching of pieces following each other is an issue. Interesting approaches for playlist generation are described in [9] and [10]. However, current research in playlist generation focuses solely on music, ignoring speech material. Similarly, current customized internet radio stations, such as Pandora,

M. Jani, "Experimental Investigation of Transitions for Mixed Speech and Music Playlist Generation,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 141-144.

Last.fm and Spotify deal only with music. Stitcher doesn't have this focus on music, but automatic playlist generation is not of primary interest there.

Also, different methods and measures for the acoustic similarity of pieces of music were introduced [11], [12]. They aim at selecting pieces of music similar to those a user likes, rather than sequentially arranging two pieces of music for playlist generation. In general, no work, either general or focusing on the match of consequent pairs, is known on mixed music and speech playlist generation.

III. Data Collection and Preparation

A. Overview

Pairs of speech and music pieces from authentic radio recordings suitable for acoustic tests were selected. Our test set contained 25 original pairs from radio programs presumably edited by professional editors and 25 randomly selected pairs as well.

A test based on the listening experience was constructed to collect subjective opinion ratings for the speech and music pairs. The test was conducted using an online survey integrated with the audio test data. It involved ordinary listeners and the result was stored in a database.

The preparation of the radio recordings and the subjective opinion listening test are described in detail in our previous works [2], [3].

A set of potentially relevant basic acoustic features were selected for the speech and the music parts. These features were extracted from the recordings used in the test. The basic features were extended by comparison features (taking basic features for both the piece of speech and of music). In a next step, the most relevant features were selected, partly with attribute preselection methods known in data mining, partly manually.

The three types of the collected data – the original pairing matched by the editor, the subjective ratings from the listeners and the automatically extracted acoustic features – were analysed. We were looking for the relationships, patterns in the data with the final goal of predicting the subjective ratings using acoustic features.

B. Acoustic Feature Extraction

For the analysis covered in this paper, pieces of speech and music have to be characterized by distinctive features, the values of which can automatically be extracted from the audio data.

The features were selected based on our general experience in speech and music processing. Only relatively few and basic features were selected to avoid overfitting the limited amount of available data.

The chosen features for the speech were:

1) average of the fundamental frequency (*sf_f0_avg*),
2) standard deviation of the fundamental frequency (*sf_f0_stddev*),
3) speech tempo (*sf_tempo*),
4) dynamic range (*sf_dynrange*) and
5) spectral coefficients.

The following features were chosen for music:

1) music tempo (*mf_tempo*),
2) dynamic range (*mf_dynrange*) and
3) spectral coefficients.

The fundamental frequency extraction has been done using the WaveSurfer program [13]. The speech tempo was measured in vowels per second, ignoring possible silent parts at the beginning and at the ending. The music tempo was measured in quarter beats per second.

For the dynamic range we implemented the International Telecommunication Union recommendation of loudness unit measure [14] and the loudness range measure of European Broadcasting Union [15].

The Mel spectrum coefficients were extracted using the OpenSMILE toolkit [16]. First and second order polynomial (line and parabola) were fitted on the time-average of the coefficients. The coefficients of the polynomials were used as separate speech (*sf_melspec_p1_0*, *sf_melspec_p1_1*, *sf_melspec_p2_0*, *sf_melspec_p2_1*, *sf_melspec_p2_2*) and music (*mf_melspec_p1_0*, *mf_melspec_p1_1*, *mf_melspec_p2_0*, *mf_melspec_p2_1*, *mf_melspec_p2_2*) features.

C. Data Preprocessing

The amount of available data was limited. Collecting more similarly detailed data would have resulted in high costs, due to the manpower intensity of subjective opinion tests. Concerning the general approach, in order to squeeze the most out of the limited data, domain knowledge was intensively used. In order to avoid overfitting, only simple methods were used.

The set of the acoustic features extracted from the audio files had to be extended with additional, calculated attributes. Our main goal is to predict the quality of match of speech and music pairs. We took into consideration that most data mining algorithms cannot compare several attributes directly or they require large amounts of data for that. Therefore, we calculated some comparison attributes that might make sense in the application domain. Among these the notable ones were:

1) ratio of the speech tempo and music tempo (*tempo_ratio*);
2) ratio of the normalized speech tempo and music tempo (*temponorm_ratio*);
3) difference of the normalized speech and music tempo (*temponorm_diff*);
4) absolute value of the above (*temponorm_diffabs*);
5) relative tempo difference: the difference of the normalized speech and music tempo (s. point 2, above) divided by the speech tempo (*temponorm_diffrel*);
6) ratio of the normalized speech and music dynamic ranges (*dynrange_ratio*);
7) difference of the normalized speech and music dynamic ranges (*dynrange_diff*);
8) absolute value of the above (*dynrange_diffabs*);

9) a linear combination of the speech tempo and the music dynamic range (*sf_tempo_mf_dynrange*);
10) largest spectral coefficient for the speech part (*sf_mel_max*) and the largest spectral coefficient for the music part (*mf_mel_max*), as well as the sum of both (*mel_maxsum*).

This aggregation of the values was preceded by a visual investigation of all spectral coefficients. It was found that large values in the music and, somewhat less dominantly, in the speech for the higher Mel spectral coefficients indicate that the piece of music doesn't fit the piece of speech. Also, relatively lower values, but occurring both in the speech and in the music part of a pair also mean less good matches. As the highest values for music and speech are comparable, we could simply sum up the two values.

IV. Results

A. Overview

Test persons gave binary feedback for speech-music pairs, indicating that the pair was matching (yes) or random (no). These were aggregated for our investigations: the ratio of the yes-votes, noted as *yes-ratio*, a value in the [0, 1] interval was used.

The main task is to predict the subjective opinions by using the previously described attributes. Two groups of methods were tested: firstly, predicting the *yes-ratio* of the subjective opinion votes, a numeric value; secondly, predicting the discretized *yes-ratio*, a binary value.

B. Numeric Prediction

One approach is to predict the *yes-ratio* numeric values. The experiments were made with two simple classifiers: Linear Regression (LR) and linear Support Vector Regression (SVR).

The attribute selection plays a major role in the process. The different selections are concluded in Table I. The first test (a) was run using all the available attributes (basic acoustic features, comparison attributes). This produced a very poor result. Then using the built-in supervised attribute selection algorithms of Weka data mining software [17] (Correlation-based Feature Subset Selection with a greedy search: CfsSubsetEval, BestFirst -D 1 -N 5), a subset of these were selected automatically (b). The results slightly improved, but were still quite poor. By selecting those attributes which were selected by the Linear Regression classifier after training it on the whole data set (c), the results got better. After some further experiments the (d) attribute set was found to give the best result.

The tests were made using 3-fold cross validation. The reason for not using the generally favoured 10-fold is that it would leave too few instances in one fold. Also 5-fold cross validation was tested for the best attribute set, but there was no significant improvement over 3-fold. We tested each setup 10 times, with different seed numbers for the cross validation splits (from 0 to 9). The setups and the results are presented in Table II and Fig. 1.

TABLE I
Selections of attributes used for prediction

(a)	all attributes
(b)	*sf_tempo_mf_dynrange*, *temponorm_diffrel*, *mf_mel_max*, *dynrange_ratio*, *dynrange_diffabs*, *temponorm_ratio*, *temponorm_diff*, *tempo_ratio*, *mf_tempo*, *mf_dynrange*, *sf_dynrange*, *sf_melspec_p2_2*, *sf_melspec_p2_1*, *sf_melspec_p1_0*, *sf_tempo*, *sf_f0_stddev*
(c)	*mel_maxsum*, *tempo_ratio*, *mf_tempo*, *sf_melspec_p2_2*
(d)	*sf_tempo*, *sf_melspec_p1_0*, *mel_maxsum*

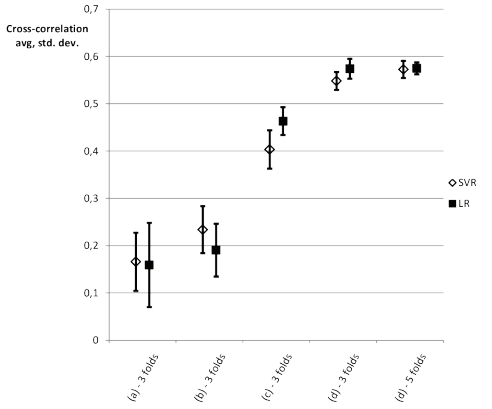


Fig. 1. Cross-correlation average and standard deviation for Support Vector Regression (SVR) and Linear Regression (LR)

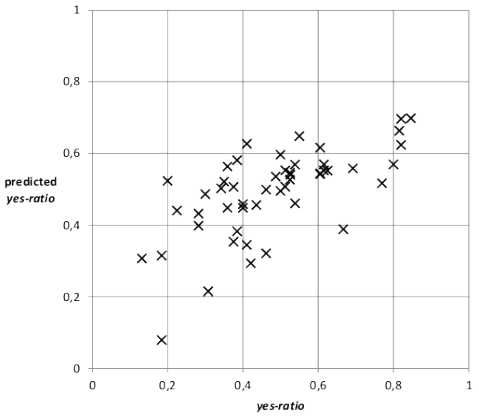


Fig. 2. The predicted *yes-ratios* in the function of the real *yes-ratios* by Linear Regression

In Fig. 2 each point represents an instance (speech-music pair). The x coordinate of the point is the *yes-ratio*, and the y coordinate is the predicted *yes-ratio* of the pair. This figure was created using Linear Regression algorithm on the whole dataset.

C. Binary Prediction

The binary prediction on this data was investigated in our previous article [3]. The best overall performance was improved using the findings in the numeric prediction. An increased accuracy of 75% was achieved using the attribute set (d) with the Linear Regression method with pre-discretized *yes-ratio*.

TABLE II
ESTIMATION PERFORMANCE WITH ATTRIBUTE SETS (A) TO (D) (TABLE I) FOR METHODS

Attribute set	(a)		(b)		(c)		(d)		(d)	
Folds	3		3		3		3		5	
Classifier	LR	SVR	LR	SVR	LR	SVR	LR	SVR	LR	SVR
Cross-corr. avg.	0.16	0.17	0.19	0.23	0.46	0.40	0.57	0.55	0.57	0.55
Cross-corr. stddev.	0.18	0.12	0.18	0.99	0.059	0.081	0.042	0.0038	0.025	0.036
Mean abs. err. avg.	0.17	0.18	0.16	0.16	0.13	0.13	0.12	0.12	0.12	0.12
Mean abs. err. stddev.	0.029	0.014	0.013	0.009	0.0077	0.0062	0.007	0.004	0.0035	0.0038

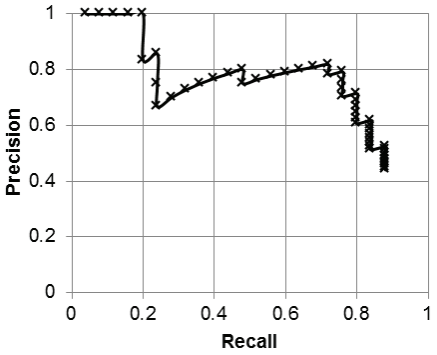


Fig. 3. Precision-recall curve of Linear Regression classifier

The precision-recall curve of the Linear Regression model is depicted on Fig. 3. For this, the speech-music pairs with a *yes-ratio* larger than 0.5 were considered to be hits.

V. CONCLUSION

In this paper, we described our investigations on the acoustic aspects of mixed speech and music playlist generation. We used a set of speech-music pairs, some of them being original matches from radio programmes, some of them random matches. Data collected by subjective opinion test corresponded to whether the match was an original match or a random match.

The major part of our investigations concentrated on finding acoustic features for speech and music and on predicting the quality of the match – the result of the subjective opinion test – on the basis of those acoustic features, by means of data mining methods.

The results are promising. Based on attributes for tempo and Mel spectrum coefficients, the match can be predicted with a cross- correlation of 0.55 with the Linear Regression method. When discretising the match, the Boolean value can be predicted with a performance of 75%. These predictions are not perfect, which suggests, that further factors may also play a role.

Future work is planned on the basis of larger data sets collected from traditional radio stations and audio social media systems. In addition, the semantic aspects (effect of the content) will also be investigated, using among others a large song lyrics database.

ACKNOWLEDGMENT

The support of the grant TÁMOP-4.2.1. B-11/2/KMR-2011-0002 is gratefully acknowledged.

REFERENCES

[1] Gergely Lukács, Beáta Pethesné Dávid, and Bea Madocsai, "Impact of personalized audio social media on social networks," in *XXXIII. Sunbelt Social Networks Conference of the International Network for Social Network Analysis Abstract Proceedings*. INSNA, 2013, p. 210.

[2] M. Jani, G. Takács, and G. Lukács, "Evaluation of speech music transitions in radio programs based on acoustic features," in *Proc. of 11th International Workshop on Content-Based Multimedia Indexing (CBMI)*, 2013, pp. 97–102.

[3] Gergely Lukács, Mátyás Jani, and György Takács, "Acoustic feature mining for mixed speech and music playlist generation," in *Proc. of 55th International Symposium ELMAR-2013*, 2013, p. 275–278.

[4] Mátyás Jani, Gergely Lukács, and György Takács, "Experimental investigation of transitions for mixed speech and music playlist generation," in *Proc. of ACM International Conference on Multimedia Retrieval*. ACM, 2014, pp. 392–398.

[5] Eliens A. and Wang Y., "EXPERT ADVICE AND REGRET FOR SERIAL RECOMMENDERS," in *Proc. EUROMEDIA 2007*. Delft, Netherlands: Eurosis, 2007, pp. 111–118.

[6] D. Zibriczky, B. Hidasi, Z. Petres, and D. Tikk, "Personalized recommendation of linear content on interactive TV platforms: beating the cold start and noisy implicit user feedback," in *Workshop and Poster Proceedings of the 20th Conference on User Modeling, Adaptation, and Personalization (UMAP), Montreal*, 2012.

[7] Ben Fields, "Contextualize your listening: The playlist as recommendation engine," PhD, Goldsmiths, University of London, London, 2011.

[8] Ben Fields and Paul Lamere, "Finding a path through the jukebox – the playlist tutorial, ISMIR," Utrecht, 2010.

[9] Arthur Flexer, Dominik Schnitzer, Martin Gasser, and Gerhard Widmer, "Playlist generation using start and end songs," in *ISMIR 2008, 9th International Conference on Music Information Retrieval, Drexel University, Philadelphia, PA, USA, September 14-18, 2008*, 2008, pp. 173–178.

[10] B. Logan, "Content-based playlist generation: Exploratory experiments," in *ISMIR 2002, 3rd International Conference on Music Information Retrieval, Paris, France, October 13-17, 2002, Proceedings*, 2002, pp. 295–296.

[11] J.-J. Aucouturier and F. Pachet, "Finding songs that sound the same," in *Proc.1st IEEE Benelux Workshop on Model based Processing and Coding of Audio (MPCA-2002)*, Nov. 2002, pp. 1–8.

[12] B. Logan and A. Salomon, "A music similarity function based on signal analysis," in *Proc. of IEEE International Conference on Multimedia and Expo, ICME*. IEEE, 2001, pp. 745–748.

[13] K. Sjölander and J. Beskow, "WaveSurfer," http://www.speech.kth.se/wavesurfer/, [Online; accessed 14-February-2013].

[14] ITU, "Recommendation ITU-R BS.1770-3 algorithms to measure audio programme loudness and true-peak audio level," 2012.

[15] EBU, "Tech doc 3342 loudness range: A measure to supplement loudness normalisation in accordance with EBU r 128," 2011.

[16] Florian Eyben, Martin Wöllmer, and Björn Schuller, "openSMILE - the munich versatile and fast open-source audio feature extractor," in *Proc. ACM Multimedia*. Florence, Italy: ACM, Oct. 2010, pp. 1459–1462.

[17] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *SIGKDD Explor. Newsl.*, vol. 11, no. 1, p. 10–18, Nov. 2009.

Vocabulary extension by paradigm prediction

Attila Novák
(Supervisor: Dr. Gábor Prószéky)
novak.attila@itk.ppke.hu

***Abstract*—Morphological analysis and generation are important tasks in natural language processing systems, especially in the case of morphologically complex languages. Computational morphologies often consist of a lexicon and some rule component, the creation of which requires various competences and considerable effort. Such a description, on the other hand, makes an easy extension of the morphology with new lexical items possible. Most freely available morphological resources, however, contain no rule component. They are usually based on just a morphological lexicon, containing base forms and some information (often just a paradigm ID) identifying the inflectional paradigm of the word, possibly augmented with some other morphosyntactic features. The aim of the research presented in this paper was to prepare an algorithm that makes the integration of new words into such resources similarly easy to the way a rule-based morphology can be extended. This is achieved by predicting the correct paradigm for words, which are not present in the lexicon. The supervised learning method described in this paper is based on longest matching suffixes and lexical frequency data, and is demonstrated and evaluated for Russian.**

***Keywords*-morphology; paradigm prediction; Russian**

I. Introduction

Morphological analysis is an important task in any natural language processing chain, preceding any further analysis of texts. It is also unavoidable in information retrieval, or indexing algorithms, where the lemma of words are to be used in order to have a robust representation of the information present in the documents.

Large-scale computational morphologies are usually created using a morphological grammar formalism that minimizes the amount of information necessary to include in the source lexicon about each lexical item by providing some rule-based method of formalization of the morphological behavior of words. This allows an easy extension of the morphology with new lexical items. This approach also gives the creator of the morphology complete control over the quality of the resource. Building rule-based morphological grammars, however, requires threefold competence: familiarity with the formalism, knowledge of the morphology, phonology and orthography of the language, and extensive lexical knowledge. Many morphological resources, on the other hand, contain no explicit rule component. Such resources are created by converting the information included in some morphological dictionary to some simple data structures representing the inflectional behavior of the lexical items included in the lexicon. The representation often only contains base forms and some information (often just a paradigm ID) identifying the inflectional paradigm of the word, possibly augmented with some other morphosyntactic features. With no rules, the extension of such resources with new lexical items is not such a straightforward task, as it is in the case of rule-based grammars. However, the application of machine learning methods may be able to make up for the lack of a rule component. In this paper, we intend to solve the problem of predicting the appropriate inflectional paradigm of out-of-vocabulary words, which are not included in the morphological lexicon. The method is based on a longest suffix matching model for paradigm identification, and it is showcased with and evaluated against an open-source Russian morphological lexicon.

Table I
The size of each test set

	rare	average	frequent
Number of words	3970	36917	9633

II. Training and test data

In the experiments described in this paper, we used the LGPL-licensed open-source Russian morphology available from www.aot.ru [1]. The core vocabulary of this morphology is based on Zaliznyak's morphological dictionary [2]. It contains 174 785 lexical entries, each of which are classified into one of 2 767 paradigms. For the evaluation of the performance of the paradigm assignment algorithm, we also used the frequency distribution of Russian lemmas, taken from Serge Sharoff's Russian internet frequency list.[1]

The morphological lexicon was then separated into training and test sets in three different settings based on lemma frequencies. First, rare words were separated from the lexicon. These are the ones occurring at most 10 times in the internet corpus. As the frequency list includes words with at least 8 occurrences, this was the lower limit. In the second setting, the middle range words were separated for testing, i.e. the ones that occur at most 100 times. In the third case, the most frequent words were considered, which correspond to a frequency value of at least 1000. The training set in each case was the complement of the training set with regard to the whole lexicon. The size of each set is shown in Table I.

III. Features affecting the paradigmatic behavior of Russian words

When attempting to predict the inflectional paradigm for Russian words, certain grammatical features of the lexical item need to be known in order to have a good chance of guessing right. Lemma and part of speech are obviously necessary features, although part of speech can be guessed

[1] http://corpus.leeds.ac.uk/frqc/internet-ru.num

A. Novák, "Vocabulary extension by paradigm prediction,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 145-148.

from the lemma for adjectives and verbs with rather good confidence. Nevertheless, we assumed these to be known, as these properties of words are present in any dictionary.

For nouns, a number of additional features (gender, countability and animacy) play a role in determining the morphosyntatctic feature combination slots which make up the paradigm of the given lemma. There are also nouns, which are undeclinable. Of these features, gender is indicated for each headword in any dictionary, and undeclinable nouns are also usually marked as such. Certain abstract, collective and mass nouns (and, in the aot resource, also many proper names) lack plural forms, while there are also pluralia tantum, which have no singular.

Animacy affects the nominal paradigm in a manner that does not influence the actual set of possible word forms. However, there is a case syncretism in Russian, which depends on animacy. For animate nouns, plural accusative coincides with genitive (for masculine nouns, the same applies also to singular). For inanimate nouns, on the other hand, the form of accusative matches that of the nominative. This difference is still present in the case of homonyms, where one of the senses of the word is animate, and another form is inanimate. Note, however, that the animacy feature, although it is present in the aot lexicon, is not generally made explicit in other dictionaries, because a human user can infer this information from the meaning of the word. We thus have not used this information. Similarly, the set of valid morphosyntactic feature combinations for verbs depends on verbal aspect and transitivity/reflexivity. Thus, these properties need to be known for verbs, and, indeed, they are listed in dictionaries. E.g. nontransitive verbs lack passive participles; verbs of perfective aspect lack present participle forms; and many verbs of imperfect aspect lack past participial (especially passive) forms. The adverbial participial forms a verb may assume also depend on aspect (and also on other idiosyncratic lexical features).

Defectivities of the adjectival paradigm, e.g. the lack of short predicative forms and synthetic comparative and superlative forms depends on semantic and other, seemingly idiosyncratic, features of the lexeme. E.g. relational adjectives usually lack these forms. Such properties, however, were not made explicit in the aot lexicon, neither are they present in normal dictionaries, so we did not use any lexical features for adjectives beside part of speech.

Thus, when defining the feature set for predicting inflectional paradigms of words, we assumed that the lemma and the lexical properties mentioned above: part of speech, gender, verb type, etc., are known. However, some morphological characteristics relevant from the aspect of inflection cannot be derived neither from a simple dictionary, nor from the surface form of a word.

The other set of features we used are n-character-long suffixes of the lemma for various lengths n. The maximum suffix length is a parameter of the algorithm. It was set to 10 in the experiments reported in this paper. In order to exploit this information, a suffix model is created based on the lexicon. An illustration of how this model including both the endings and the lexical features is generated is shown in Figure 1.

мумиё[N.n.*.-];prd:25	мумиё|n*[N.n-25]
остриё[N.n.-];sfx:ё;prd:1709	остри#ё|n[N.n-1709]
бабьё[N.n-];sfx:ё;prd:210	бабь#ё|ns[N.n-210]
дубьё[N.n-];sfx:ё;prd:210	дубь#ё|ns[N.n-210]
свежевьё[N.n-];sfx:ё;prd:210	свежевь#ё|ns[N.n-210]
цевьё[N.n.-];sfx:ьё;prd:1433	цев#ьё|n[N.n-1433]
жнивьё[N.n];sfx:ё;prd:1103	жнивь#ё|n[N.n-1103]
суровьё[N.n];sfx:ё;prd:210	суровь#ё|ns[N.n-210]
мостовьё[N.n];sfx:ё;prd:210	мостовь#ё|ns[N.n-210]

Figure 1. A portion of the suffix model. The format of the right column is: `lem#ma|lex-features[PosTag-paradigmID]`, where `ma` is a required ending of the lemma for all items in the paradigm identified by `paradigmID`.

IV. Creation of the suffix model

A suffix trie is built of words input to the training algorithm in the form shown in the right column of Figure 1. The lemma is decorated with the following features (from right to left):

- The tag in brackets consists of two parts: part of speech (and, in the example below: gender) is followed by the appropriate paradigm ID from the aot database; the two are separated by a hyphen. This is the information to be predicted by the algorithm for unknown words. After processing the training data, terminal nodes of the suffix trie link to a data structure representing the distribution (relative frequency) of tags for the given suffix.
- A suffix following a vertical bar is attached to the end of the lemma. This represents the available lexical knowledge about the lexical item in an encoded form.[2]
- Some paradigms are restricted to lemmas ending in a specific suffix. There is a hash mark at the beginning of the suffix of the lemma that is required by the given paradigm ID to be valid. The given paradigm ID is not applicable to words not having that ending. E.g. all lemmas in paradigm 1433 must end in *ъё*.

V. Ranking

The suffix-trie-based ranking algorithm that we used was inspired by the suffix guesser algorithm used in Brants' TnT tagger to estimate the lexical probability of out-of-vocabulary words ([3]). However, that model did not prove to perform well enough in this task. So we modified the model step-by-step until we arrived at a model that turned out to be simpler, yet to perform much better. The paradigms are predicted by assigning a score to each paradigm for each word. Then, the higher this score is for a paradigm tag for a certain word, the more probable it is that the word belongs to that paradigm. We select the top-ranked paradigm to be the predicted inflectional class.

The score for each paradigm in the case of a word is calculated for all suffixes of the word, including the lexical properties, from shortest to longest. More formally, for all tags, the rank is calculated iteratively according to Formula 1.

$$rank^{i+1}[tag] = sign \times len\_sfx \times rel\_freq + rank^{i}[tag] \quad (1)$$

[2]n: neuter noun, *: undeclinable, s: singular only

where

$sign$	is negative if the suffix is shorter than the minimal suffix required by the paradigm
len_sfx	is the length of suffix w/o lexical properties
rel_freq	is the relative frequency of tag for the suffix
$rank^i[tag]$	is divided by len_sfx if $len\_sfx > 1$ is negated if $sign > 0$ and $rank^i[tag] < 0$ before calculating $rank^{i+1}[tag]$

The applied ranking score clearly prefers the most frequent paradigm for the longest matching suffix.

VI. Evaluation

Evaluation of the ranking algorithm was performed for the four different test sets described in Section II. These are rare words (LT10), average words (LT100), and frequent words (MT1000). We used standard evaluation metrics for measuring the performance of our method. *First-best accuracy* measures the ratio of having the correct paradigm ranked at first place. This reflects the ability of the system to automatically classify new words to paradigms. In addition, the accuracy values for 2^{nd} to 9^{th} ranks were also calculated. *Recall* is the ratio of having the correct paradigm in the set of the first ten highest ranked candidates. Following the metrics used by [4], precision was calculated as *average precision at maximum recall*, i.e. $1/(1+n)$ for each word, where n is the rank of the correct paradigm. This measures the performance of the ranking algorithm. As it might be the case that paradigm prediction is used to aid human classification, this metric reflects the ratio of noise a human must face with when verifying the results. Finally, *f-measure* is the harmonic mean of precision and recall.

In order to measure the advances in the performance, two baselines were created. The first one uses Brants' suffix guesser model ([3]) instead of the longest suffix matching method. This model uses a θ factor to combine tag probability estimates for endings of different length in order to get a smoothed estimate. θ is set as the standard deviation of the probabilites of tags. First, the probability distribution for all suffixes is generated from the training set, then it is smoothed by successive abstraction according to Formula 2.

$$P(t|l_{n-i+1},...l_n) = \frac{\hat{P}(t|l_{n-i+1},...l_n) + \theta_i P(t|l_{n-i},...l_n)}{1+\theta_i} \quad (2)$$

for $i = m...0$, with the initial setting $P(t) = \hat{P}$, where

$\hat{P}$	are maximum likelihood estimates from the frequencies in the lexicon
θ_i	weights are the standard deviation of the unconditioned maximum likelihood probabilities of the tags in the training set for all i

The other baseline assigns the most frequent paradigm identifier to each word based on its part of speech. The results of these baselines compared to our system are shown in Table II. As expected, the second baseline, choosing the most frequent tag, has a very low accuracy, however, our longest suffix method outperforms the first baseline as well. A key difference between the two models is that Brants' model assigns more weight to unconditioned tag distributions and ones conditioned on shorter suffixes than those conditioned on longer ones. This is just the other way round in the longest suffix algorithm.

Table II
First-best accuracy of full tags achieved by the longest suffix match algorithm, Brants' model, and by assigning the most frequent paradigm tag

	Longest suffix	Brants' model	Most frequent tag
LT10	0.9274	0.6269	0.3433
LT100	0.9174	0.6148	0.3386
MT1000	0.8087	0.5687	0.3287

Table III
Evaluation of each test set for the ranked results

	LT10		LT100		MT1000	
	full	equip	full	equip	full	equip
#1	0.8924	**0.9274**	0.8750	**0.9174**	0.7416	**0.8087**
#2	0.0614	0.2322	0.0685	0.2278	0.0684	0.2371
#3	0.0168	0.2090	0.0223	0.2201	0.0314	0.2435
#4	0.0057	0.1518	0.0078	0.1452	0.0168	0.1900
#5	0.0035	0.1692	0.0037	0.1723	0.0090	0.2165
#6	0.0015	0.1884	0.0019	0.1683	0.0083	0.1697
#7	0.0000	0.1871	0.0012	0.1836	0.0032	0.1562
#8	0.0005	0.1400	0.0011	0.1496	0.0043	0.1418
#9	0.0010	0.1095	0.0007	0.1573	0.0017	0.1078
precision	0.9329	0.9538	0.92195	0.9481	0.8067	0.8550
recall	0.9841	0.9876	0.9832	0.9875	0.8872	0.9158
f-measure	0.9578	**0.9704**	0.9516	0.9674	0.8450	0.8843

The tags assigned to paradigms and syntactic features define a very sophisticated classification of words. However, some of the features that distinguish two different paradigms are not relevant from the aspect of their inflectional behavior, such as the subtype of a non-inflecting adverb. Also some paradigm differences are irrelevant from the point of view of a pure lemmatization task, because they do not affect the set of word forms in the paradigm. To see how the algorithms perform in that task, equivalence classes of paradigms were generated, and a prediction was considered correct if the set of inflected forms generated by the predicted paradigm was identical to the set of word forms generated by the correct paradigm. Of the 2 767 different paradigms, 921 non-unique paradigms could be collapsed into 283 equivalence classes. Table III shows the results for each setup, where columns 'full' and 'equip' correspond to full tag and equivalence class evaluations respectively. Note that the values in the 'equip' columns do not sum up to 1, since, in many cases, two or more paradigms on the list of top-ranked paradigm candidates would generate the same set of inflected word forms for a given lexical item.

As the numbers show, our system performs best on rare words, while it achieved the worst results on very frequent words. This is not very surprising, as irregular words tend

Table IV
FIRST-BEST ACCURACY OF FULL TAG PREDICTION IN THE CASE OF ALL TYPES OF WORDS, NOUNS, VERBS AND ADJECTIVES

	ALL	NOUNS	VERBS	ADJECTIVES
LT10	0.9166	0.9547	0.8158	0.8665
LT100	0.9038	0.9489	0.8114	0.8381
MT1000	0.7678	0.8594	0.6884	0.5991

to be frequent words, while rare words have regular inflectional behavior. Correctly predicting the exact paradigm of an unknown personal pronoun or an irregular verb is indeed a rather difficult task. Since our aim was to extend existing morphological lexicons, and such resources already contain the most frequent words of the language, the results obtained for rare words are the ones which are relevant for our task.

Also note that beside similar recall values, precision and first-best accuracy are significantly higher when equivalent paradigms are collapsed. The prediction algorithm works reasonably well for extending resources for tasks that do not require full morphological analysis such as indexing for information retrieval or dictionary lookup.

Table IV shows the first-best accuracy results for all words, nouns, verbs and adjectives separately. In this table, instead of full tag agreement, only the paradigm identifiers were considered. The exact paradigm of verbs and adjectives turned out to be more difficult to guess than that of nouns, due to semantic factors and stress variation as explained in the next section of this paper.

VII. ERROR ANALYSIS

The most frequent confusions of the longest suffix algorithm for infrequent words are due to failure to correctly predict

- whether an adjective has synthetic comparative forms
- whether a *-ние*-final abstract noun has an alternative *-нье* spelling
- whether a noun has a second genitive form (used in partitive constructions)
- stress in past passive participles of certain verb classes (this results in an *е*∼*ё* contrast not normally reflected in orthography)
- whether an adjective has synthetic superlative forms
- stress in short and comparative forms of certain adjectives (this results in an *е*∼*ё* contrast not normally reflected in orthography)
- whether a non-inflecting noun can be interpreted as plural
- whether an imperfective verb has past passive participle forms
- optional stress variation across the paradigm
- whether an adjective has short predicative forms

Except for stress-related issues and semantically motivated or idiosyncratic defectivity, incorrect forms are very rarely predicted by the algorithm. Humans would probably make similar mistakes for words they do not know, especially if they do not know the meaning of the word either. The system sometimes highlights inconsistencies in the original aot data that even the authors, who are not native or even advanced speakers of Russian can identify as errors, e.g. that while the name of the energy company *Кубаньэнерго* is categorized as lexically non-plural, the similarly formed *Сахалинэнерго* does not have this property.

When looking at errors the algorithm makes when applied to frequent words, we find that the types of errors are similar. Nevertheless, failure to predict superlatives, comparatives, second genitives or special locative forms is much more prevalent for this data, as a much higher proportion of very frequent words have these "irregular" forms.

The most frequent errors of Brants' original suffix guesser algorithm, on the other hand, include absurd errors that would not be made even by beginning learners of Russian. This is due to overemphasizing distributions conditioned on shorter suffixes over those on longer ones. The top-ranked candidate paradigm is often totally inapplicable to words having the ending the given lexical item has, such as the paradigm of *-ский*-final adjectives to *-ный*-final ones (the most frequent error of that algorithm for infrequent words).

VIII. CONCLUSION

In this article, we presented and evaluated a suffix-trie-based supervised learning algorithm capable of predicting inflectional paradigms for words based on the ending of their lemma and some basic lexical properties. The algorithm can be used to automatically extend the vocabulary of computational morphologies lacking an independent rule component. The experiments were demonstrated for Russian, however, with minimal adaptation the tool can be used for any language provided there is a morphological resource available. Moreover, we assumed that a dictionary with some lexical features is also available, thus such features could be used for disambiguating paradigm candidates. The results showed that our method performs above 90% in all the different setups, achieving the best performance on relatively rare words, which are good candidates of being absent in the original lexicon.

We found that assigning more weight to distributions conditioned on longer suffixes than on shorter ones yields much better prediction performance, not only in terms of the number of exact predicted paradigm matches, but especially when taking into account what sorts of errors the system makes. While the baseline suffix guesser algorithm often proposes paradigms inapplicable to the given lexical item, our algorithm makes errors that arise due to the lack of lexical semantic information. Humans would make similar errors in similar situations.

REFERENCES

[1] A. V. Sokirko, "Morphological modules at the site www.aot.ru," in *Dialog'2004*, 2004.

[2] A. A. Zaliznyak, *Russian grammatical dictionary – Inflection*. Moskva: Russkij Jazyk, 1980.

[3] T. Brants, "Tnt - a statistical part-of-speech tagger," in *Proceedings of the Sixth Applied Natural Language Processing (ANLP-2000)*, Seattle, WA, 2000.

[4] K. Linden, "Entry generation by analogy – encoding new words for morphological lexicons," in *Journal Northern European Journal of Language Technology*, 2009, pp. 1–25.

Identification of abbreviations and multiword terms in clinical reports based on their contextual behavior

Borbála Siklósi
(Supervisor: Dr. Gábor Prószéky)
siklosi.borbala@itk.ppke.hu

***Abstract*—The automatic processing of clinical documents created at clinical settings has become a focus of research in natural language processing. However, standard tools developed for general texts are not applicable or perform poorly on this type of documents. Moreover, several crucial tasks require lexical resources and relational thesauri or ontologies to identify relevant concepts and their connections. In the case of less-resourced languages, such as Hungarian, there are no such lexicons available. The construction of annotated data and their organization requires human expert work. In this paper we show how applying statistical methods can result in a preprocessed, semi-structured transformation of the raw documents that can be used to aid human work. The modules detect and resolve abbreviations, identify multiword terms and derive their similarity, all based on the corpus itself.**

***Keywords*-clinical text processing; abbreviations; multiword terms; distributional similarity; less-resourced languages**

I. Introduction

Clinical records are documents created at clinical settings with the purpose of documenting every-day clinical cases or treatments. The quality of this type of text stays far behind that of biomedical texts, which are also the object of several studies. Biomedical texts, mainly written in English, are the ones that are published in scientific journals, books, proceedings, etc. These are written in the standard language, in accordance with orthographical rules ([1], [2]). On the contrary, clinical records are created as unstructured texts without using any proofing tools, resulting in texts full of spelling errors and nonstandard use of word forms in a language that is usually a mixture of the local language (Hungarian in our case) and Latin ([3], [4]).

In order to access the hidden data, an efficient representation of the facts and statements recorded in the texts should be created. Several attempts have been made to apply general text processing tools to clinical notes, but their performance is much worse on these special texts, than on general, well-formed documents. Moreover, applications used for processing domain-specific texts are usually supported by some hand-made lexical resources, such as ontologies or vocabularies. In the case of less-resourced languages, there are very few such datasets and their construction needs quite an amount of human work. Furthermore, as facts and statements in clinical records are from a narrow domain, applications of the sublanguage theory ([5]) have been used in similar approaches, which also requires a domain-specific categorization of words of the specific sublanguage.

In order to be able to support the adaptation of existing tools, and the building of structured resources, we examined a corpus of Hungarian ophthalmology notes. In this study, statistical methods are applied to the corpus in order to capture as much information as possible based on the raw data. Even though the results of each module are not robust representations of the underlying information, these groups of semi-structured data can be used in the real construction process.

II. The corpus

In this research, anonymized clinical documents from the ophthalmology department of a Hungarian clinic were used. This corpus contains 334 546 tokens (34 432 sentences). The models were built using this set, and tested on another set of documents, which contained 5 599 tokens (693 sentences).

Compared to a general Hungarian corpus, there are reasonable differences between the two domains, which explains some of the difficulties that prevent tools developed for general texts working in the clinical domain. These differences are not only present in the semantics of the content, but in the syntax and even in the surface form of the texts and fall into three main categories discussed in the following subsections. The corpus used in the comparison as general text was the Szeged Corpus ([6]), containing 1 194 348 tokens (70 990 sentences) and the statistics related to this corpus was taken from [7].

A. Syntactic behaviour

The length of the sentences used in a language reflects the complexity of the syntactic behaviour of utterances. In the general corpus, the average length of the sentences is 16.82 tokens, while in the clinical corpus it is 9.7. Doctors tend to use shorter and rather incomplete and compact statements. This habit makes the creation of the notes faster, but being in lack of crucial grammatical constituents, most parsers fail when trying to process them.

Regarding the distribution of part-of-speech (pos) in the two domains, there are also significant differences. While in the general corpus, the three most frequent types are nouns, verbs and adjectives, in the clinical domain nouns are followed by adjectives and numbers in the frequency ranking, while the number of verbs in this corpus is just one third of the number of the latter two. Another significant difference is that in the clinical domain, determiners, conjunctions, and pronouns are

B. Siklósi, "Identifcation of abbreviations and multiword terms in clinical reports based on their contextual behavior,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 149-152.

also ranked lower in the frequency list. Furthermore, most of the numbers in the clinical corpus are numerical data.

B. Spelling errors

Clinical documents are usually created in a rush without proofreading. The medical records creation and archival tools used at most Hungarian hospitals provide no proofing or structuring tools. Thus, the number of spelling errors is very high and a wide variety of error types occur ([8]). These errors are not only due to the complexity of the Hungarian language and orthography, but also to characteristics typical of the medical domain and the situation in which the documents are created.

A common characteristic of these phenomena is that the prevailing errors vary with the doctor or assistant typing the text. Thus it can occur that a certain word is mistyped and should be corrected in one document while the same word is a specific abbreviation in another one, which does not correspond to the same concept as the corrected one. Latin medical terms usually have a standard form based on both Latin and Hungarian orthography, however what we find in the documents is often an inconsistent mixture of the two (e.g. tensio/tenzio/tensió/tenzió). Even though the spelling of these forms is standardized, doctors tend to develop their own customs which they use inconsistently.

Compared to the Szeged Corpus, the ratio of misspelled words was 0.27% in the general Hungarian texts, while 8.44% in the clinical notes. Moreover, the general corpus has several subcorpora, including one of primary school essays, which still had only an 0.87% error rate.

C. Abbreviations and word forms

The use of a kind of notational text is very common in clinical documents. This dense form of documentation contains a high ratio of standard or arbitrary abbreviations and symbols, some of which may be specific to a special domain or even to a doctor or administrator. These short forms might refer to clinically relevant concepts or to some common phrases that are very frequent in the specific domain. For the clinicians, the meaning of these common phrases is as trivial as the standard shortened forms of clinical concepts due to their expertise and familiarity with the context. The difference in the ratio of abbreviations in the general and clinical corpora is also significant, being 0.08% in the Szeged Corpus, while 7.15% in the clinical corpus, which means that the frequency of abbreviations is two order of magnitude larger in clinical documents than in general language.

III. Applied methods

A. Resolving abbreviations

The task of abbreviation resolution is often treated as word sense disambiguation (WSD). The best-performing approaches of WSD use supervised machine learning techniques. In the case of less-resourced languages, however, neither manually annotated data, nor an inventory of possible senses of abbreviations are available, which are prerequisites of supervised algorithms [9]. On the other hand, unsupervised WSD methods are composed of two phases: word sense induction (WSI) must precede the disambiguation process. Possible senses for words or abbreviations can be induced from a corpus based on contextual features. However, such methods require large corpora to work properly, especially if the ratio of ambiguous terms and abbreviations is as high as in the case of clinical texts. Due to confidentiality issues and quality problems, this approach is not promising either.

Thus, in this research, a corpus-based approach was applied for the resolution of abbreviations with using the very few lexical resources available in Hungarian. Even though the first approach was based on the corpus itself, it did not provide acceptable results, thus the construction of a domain-specific lexicon was unavoidable. But, instead of trying to create huge resources covering the whole field of medical expressions, it was shown in [10] that a small domain-specific lexicon is satisfactory and the abbreviations to be included can be derived from the corpus itself.

Having this lexicon and the abbreviated tokens detected, the resolution was based on series of abbrevitaions. Even though standalone abbreviated tokens are highly ambiguous, they more frequently occur as members of multiword abbreviated phrases, in which they are usually easier to interpret unambiguously. For example *o.* could stand for any word either in Hungarian or in Latin, starting with the letter *o*, even if limited to the medical domain. However, in the ophthalmology reports, *o.* is barely used by itself, but together with a laterality indicator, i.e. in forms such as *o. s., o. d.*, or *o. u.* meaning *oculus sinister* ‘left eye’, *oculus dexter* ‘right eye’, or *oculi utriusque* ‘both eyes’, respectively. In such contexts, the meaning of the abbreviated *o.* is unambiguous. It should be noted, that these are not the only representations for these abbreviated phrases, for example *oculus sinister* is also abbreviated as *o. sin., os, OS*, etc. Thus, when performing the resolution of abbreviations, we considered series of such shortened forms instead of single tokens. A series is defined as a continuous sequence of shortened forms without any unabbreviated word breaking the sequence.

Moreover, in order to save mixed phrases (when only some parts of a multiword phrase is abbreviated) and to keep the information relevant for the resolution of multiword abbreviations, the context of a certain length was attached to the detected series. Beside completing such mixed phrases, the context also plays a role in the process of disambiguation. The meaning (i.e. the resolution) of abbreviations of the same surface form might vary in different contexts.

These abbreviation series are then matched against the corpus, looking for resolution candidates, and only unresolved fragments are completed based on searching in the lexicon. The detalis of the algorithm and the results are published in [10] and [11]. It is shown there that having the corpus as the primary source is though insufficient, but provides more adequate resolutions in the actual domain.

B. Extracting multiword terms

In the clinical language (or in any other domain-specific or technical language), there are certain multiword terms that express a single concept. These are important to be recognized, because a disease, a treatment, a part of the body, or other relevant information can be in such a form. Moreover, these terms in the clinical reports could not be covered by a standard lexicon. For example, the word *eye* is a part of the body, but by itself it does not say too much about the actual case. Thus, in this domain the terms *left eye, right eye* or *both eyes* are single terms, referring to the exact target of the event the note is about. Moreover, the word *eye* seldom occurs in the corpus without a modifier. This would indicate the use of common methods based on mutual information or collocation. After a review of such methods, in [12] a c-value approach is described for multiword term extraction, emphasising the recognition of nested terms.

We used a modified version of this c-value algorithm. First, a linguistic filter is applied in order to ensure that the resulting list of terms contains only well-formed phrases. Phrases of the following forms were allowed:

$$(Noun|Adjective|PresentParticiple|Past(passive)Participle)^{+}Noun$$

This pattern ensures that only noun phrases are extracted and excludes fragments of frequent cooccurrences. It should be noted that other types of phrases, such as verb phrases, might be relevant as well, however, as described in Section II-A, the ratio of verbs is much lower in the clinical corpus, than in a general one. Thus, having only a relatively small corpus of this domain, statistical methods would be inefficient to build accurate models.

After collecting all n-grams matching this pattern, the corresponding c-value is calculated for each of them, which is an indicator of the termhood of a phrase. The c-value is based on four components: the frequency of the candidate phrase; the frequency of the candidate phrase as a subphrase of a longer one; the number of these longer phrases; and the length of the candidate phrase. These statistics are derived from the whole corpus of clinical notes. The details of the algorithm are found in [12].

C. Distributional semantic models

Creating groups of relevant terms in the corpus requires a similarity metric measuring the closeness of two terms. Instead of using an ontology for retrieving similarity relations between words, the unsupervised method of distributional semantics was applied. Thus, the similarity of terms is based on the way they are used in the specific corpus.

The theory behind distributional semantics is that semantically similar words tend to occur in similar contexts ([13]) i. e. the similarity of two concepts is determined by their shared contexts. The context of a word is represented as a set of features, each feature consisting of a relation (r) and the related word (w'). In other studies these relations are usually grammatical relations, however in the case of clinical texts, the grammatical analysis performs poorly, resulting in a rather noisy model. In this research a set of features based on the surface form and pos tag of the target word and those of the related ones were used. Each feature is associated with a frequency determined from the corpus. From these frequencies the amount of information contained in a tuple of (w,r,w') can be computed by using maximum likelihood estimation. This is equal to the mutual information between w and w'. Then, to determine the similarity between two words (w_1 and w_2) the similarity measure described in [14] was used, i.e.:

$$\frac{\sum_{(r,w)\in T(w_1)\bigcap T(w_2)}(I(w_1,r,w)+I(w_2,r,w))}{\sum_{(r,w)\in T(w_1)} I(w_1,r,w)+\sum_{(r,w)\in T(w_2)} I(w_2,r,w)}$$

where T(w) is the set of pairs (r,w') such that I(w,r,w') is positive.

Having this metric, the pairwise distributional similarity of any two terms can be counted. The similarity of multiword terms corresponds to the similarity of the last noun in the phrase.

IV. Results

The aim of this research was to create a transformation of clinical documents into a semi-structured form to aid the constrution of hand-made resources and the annotation of clinical texts. For a set of randomly selected documents taken one by one, the modules described above were applied. First, abbreviations were detected, collected and resolved. The resolutions were expanded with Latin and Hungarian variants as well. Then, multiword terms were identified and ranked for each document by the corresponding c-value. Finally, the pairwise similarity values for these terms were displayed in a heatmap and by listing the gropus of the most similar terms.

An example for a processed document is shown in Figure 1. The similarity of terms reveals that *tiszta törőközeg*, 'clean refractive media' and *békés elülső szegmentum*, 'calm anterior segment' behave very much alike, while they are different from *bal szem*, 'left eye'. This differentiation could indicate two types of annotation or two clusters of these three terms in this small example. These clusters can then be populated from more related terms extracted from other documents as well. Moreover, standalone words (nouns in our case) can be added to these clusters, when they are not part of any longer terms.

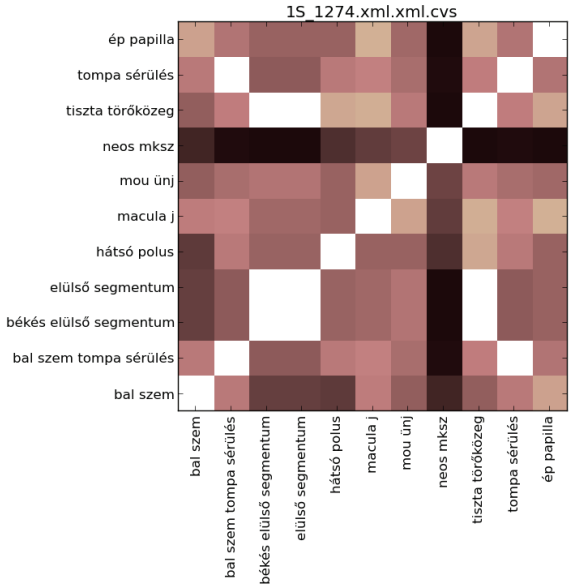
Other reasonable clusters that arise during the processing are ones that collect measurement units, such as *d sph* and *d cyl*. These terms appear as abbreviations as well in such documents, thus their resolution can also be linked to the cluster. Names of diseases are also grouped automatically, such as *asteroid hyalosis* and *cat. incip.* which are found in the same document. Note that the term *cat. incip.* is also an abbreviated form and is correctly recognized by both the multiword term extractor as a single term, and the abbreviation resolver, which generated the correct resolution as *cataracta inicipiens*. Another phenomenon that can be observed is the

Term		c-value
bal szem	'left eye'	2431.708
ép papilla	'intact papilla'	1172.0
tiszta törőközeg	'clean refractive media'	373.0
békés elülső szegmentum	'calm anterior segment'	160.08
hátsó polus	'posterior pole'	47.5
tompa sérülés	'faint damage'	12.0

(a) Multiword terms

Abbreviation	Resolutions	English translation
mydr	mydrum	mydrum
mksz	mindkét szem; oculi utriusque	both eyes
V	visus	visus
D	dioptria	dioptre
ünj	üveg nem javít	glasses do not help
St.	status	state
o. u	oculi utriusque; mindkét szem	both eyes
j.o.	jobb oldal	right side

(b) Abbreviations



(c) Heatmap of multiword term similarities. The lighter a square is, the more similar the two terms are.

Fig. 1. A fraction of the representation of a processed document with some examples of the automatically extracted multiword terms (a), some resolved abbreviations (b) and the heatmap of the similarities between multiword terms found in the actual document.

collection of word form or even phrase form variants, such as misspelled forms.

Being presented with such preprocessed, semi-structured transformation of the raw documents, human experts are more willing to complete the annotation. For example, if a group of semantically related terms are collected (such as diseases, treatments, etc.), the human annotator can just name the group and the label will be assigned to all items. Moreover, similarity values of looser relations can also be used as an initial setting when building a relational thesaurus.

V. Conclusion

Clinical documents represent a sublanguage regarding both the content and the language used to record them. However, one of the main characteristics of these texts is the high ratio of noise due to misspellings, abbreviations and incomplete syntactic structures. It has been shown that for a less-resources language, such as Hungarian, lexical resources are not available, which are used in similar studies to identify relevant concepts and relations. Thus, such lexicons should be built manually by human experts. However, an initial preprocessed transformation of the raw documents makes the task more efficient. Due to the availability of efficient implementations, statistical methods can be applied to a wide variety of text processing tasks. That is why, in this paper, we have shown that corpus-based approaches (augmented with some linguistic restrictions) perform well on abbreviation resolution, multiword term extraction and distributional similarity measures. Applying such methods can result in a semi-structured representation of clinical documents, appropriate for further human analyses.

References

[1] N. Sager, M. Lyman, C. Bucknall, N. Nhan, and L. J. Tick, "Natural language processing and the representation of clinical data," *Journal of the American Medical Informatics Association*, vol. 1, no. 2, Mar/Apr 1994.

[2] S. Meystre, G. Savova, K. Kipper-Schuler, and J. Hurdle, "Extracting information from textual documents in the electronic health record: a review of recent research," *Yearb Med Inform*, vol. 35, pp. 128–44, 2008.

[3] B. Siklósi, G. Orosz, A. Novák, and G. Prószéky, "Automatic structuring and correction suggestion system for Hungarian clinical records," in *8th SaLTMiL Workshop on Creation and use of basic lexical resources for less-resourced languages*, G. De Pauw, G.-M. De Schryver, M. Forcada, F. M. Tyers, and P. Waiganjo Wagacha, Eds., 2012, pp. 29.–34.

[4] B. Siklósi, A. Novák, and G. Prószéky, "Context-aware correction of spelling errors in Hungarian medical documents," in *Statistical Language and Speech Processing*, A.-H. Dediu, C. Martin-Vide, R. Mitkov, and B. Truthe, Eds., vol. LNAI 7978. Springer Verlag, 2013.

[5] Z. S. Harris, "The structure of science information," *J. of Biomedical Informatics*, vol. 35, no. 4, pp. 215–221, Aug. 2002.

[6] D. Csendes, J. Csirik, and T. Gyimóthy, "The Szeged Corpus: A POS Tagged and Syntactically Annotated Hungarian Natural Language Corpus." in *TSD*, ser. Lecture Notes in Computer Science, P. Sojka, I. Kopecek, and K. Pala, Eds., vol. 3206. Springer, 2004, pp. 41–48.

[7] V. Vincze, "Domnek kzti hasonlsgok s klnbsgek a szfajok s szintaktikai viszonyok eloszlsban," 2013, pp. 182–192.

[8] B. Siklósi, A. Novák, and G. Prószéky, *Context-Aware Correction of Spelling Errors in Hungarian Medical Documents*. Springer Berlin Heidelberg, 2013, no. Lecture Notes in Computer Science 7978, pp. 248–259.

[9] M. Nasiruddin, "A state of the art of word sense induction: A way towards word sense disambiguation for under-resourced languages," *CoRR*, vol. abs/1310.1425, 2013.

[10] G. P. Borbála Siklósi, Attila Novák, "Resolving abbreviations in clinical texts without pre-existing structured resources," *Fourth Workshop on Building and Evaluating Resources for Health and Biomedical Text Processing, LREC 2014.*, 2014.

[11] B. Siklósi and A. Novák, *Detection and Expansion of Abbreviations in Hungarian Clinical Notes*, ser. Lecture Notes in Artificial Intelligence. Heidelberg: Springer-Verlag, 2013, vol. 8265, pp. 318–328.

[12] K. Frantzi, S. Ananiadou, and H. Mima, "Automatic recognition of multi-word terms:. the c-value/nc-value method," *International Journal on Digital Libraries*, vol. 3, no. 2, pp. 115–130, August 2000.

[13] J. R. Firth, "A synopsis of linguistic theory 1930-55." vol. 1952-59, pp. 1–32, 1957.

[14] D. Lin, "Automatic retrieval and clustering of similar words," in *Proceedings of the 17th international conference on Computational linguistics - Volume 2*, ser. COLING '98. Stroudsburg, PA, USA: Association for Computational Linguistics, 1998, pp. 768–774.

Quality Estimation of Machine Translated Output Without Reference Translation

Zijian Győző Yang
(Supervisor: Dr. Gábor Prószéky)
yang.zijian.gyozo@itk.ppke.hu

***Abstract*—Nowadays machine translation is widely used. It is a great challenge to predict the quality of the machine translated output. There are many methods for evaluation of machine translation system. The traditional methods, which use reference translation for evaluation, have some critical problems. The first problem is that traditional methods cannot evaluate the output of machine translation systems in real-time. Another problem is that between human evaluation and traditional automatic evaluation, the correlation is very low. Last but not least, traditional evaluation systems use human translated reference translation, which is very expensive to produce. This research introduces a completely new approach. The purpose of this study is to build a quality estimation system for machine translation, which can calculate quality indicators in real time, does not use reference translation and the quality metrics are highly correlated to human evaluation. The new method is called Quality Estimation. This new method extracts different feature indicators from source and machine translated segments. Using these indicators and learning mechanism, the system builds a quality estimation model, which is trained on human annotations. As this system learns from human judgements, it can give more accurate quality values for the machine translated segments than the traditional methods. This new approach has never been used for the English-Hungarian language pair before.**

***Keywords*-evaluation; quality estimation; machine translation**

I. Introduction

As machine translation became popular among people, the measurement of translation output became an important task. There is an emerging need for quality indicators for machine translated sentences. The quality scores of machine translated segments can help human annotators in their post-edit tasks, or help machine translation systems and translation memory to find the best translation. Last but not least, quality indicators can filter out and inform about unreliable translations for everybody who uses a machine translation system. The traditional evaluation methods cannot perform this task well enough, because these need reference translation. It means that after machine translation, we also have to create a human translated sentence to compare to machine translated output. Creating human translation is very expensive and slow, and the quality is strongly dependent on the reference translation. Thus it needs a completely new approach to solve these problems. So we need a method, which can predict quality value in real-time and does not need reference translation. This research introduces a new real-time method without reference translation. This method, which is called Quality Estimation, can save considerable time for human annotators and researchers. Since it does not use reference translation, it can save a lot of money as well. This new method has never been researched for Hungarian.

In this paper, first I will shortly introduce the traditional methods and a quality estimation method that I used, then I show the benchmarking and finally I will present my experiments and results.

II. Evaluation Methods

There are two kinds of evaluation method. One uses of them is the methods which use reference translation. It means that these methods use human translated sentences and compare the machine translated sentences to these sentences, and measure the similarities or differences between them. The commonly used metrics with reference translation include BLUE [1], NIST [2], METEOR [3], TER [4], etc. One of the most popular metric is the BLEU (bilingual evaluation understudy) metric. The BLEU measures correspondence between machine translation and reference translation. It uses n-gram precision with weights. The NIST is based on BLEU method. It gives weights to particular n-grams, the rarer that n-gram is, the higher weight values will be given. The METEOR (metric for evaluation of translation with explicit ordering) counts weighted harmonic mean of unigram precision and unigram. The TER (translation error rate) measures the number of edits required to change a machine translated sentence into a reference translation.

Other methods do not use reference translation. This approach is an unsupervised evaluation method, which is called Quality Estimation of machine translation. The quality estimation addresses the problem by evaluating the quality of machine translated segments as a prediction task. For example, Gamon et al. [5] use the linear and nonlinear combinations of language model and SVM classifier to find the erroneous translations. Albrecht and Hwa [6] present a method which uses regression learning and a set of indicators of adequacy and fluency as references to evaluate the translation at sentence level. This study presents the QuEst [7] method, which was developed by Specia at al.

The QuEst system consists of two modules: a feature extraction module and a machine learning module. The feature extraction module (see Figure 1) can extract many different feature indicators (see Figure 2) from source and machine translated segments. From source sentences, complexity features can be extracted. From translated sentences the QuEst ex-

Z. Gy. Yang, "Quality Estimation of Machine Translated Output Without Reference Translation,"
in *PhD Proceedings Annual Issues of the Doctoral School - 2014*, T. Roska, G. Prószéky, P. Szolgay, Eds.
Faculty of Information Technology and Bionics, Pázmány Péter Catholic University.
Budapest, Hungary: Pázmány University ePress, 2014, pp. 153-156.

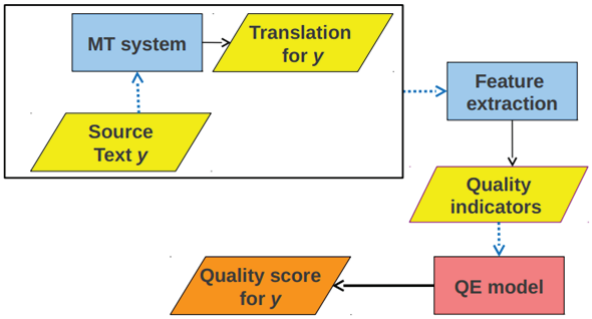


Fig. 1. Feature extraction module [8]

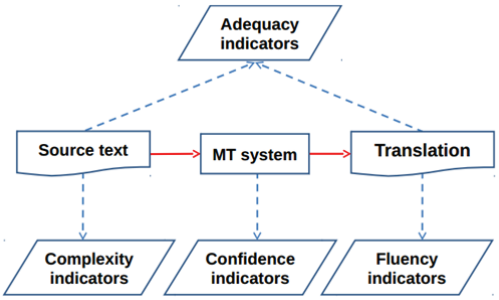


Fig. 2. Feature types [7]

tract fluency features. From the comparison between the source and the translated sentences, adequacy features are extracted. Finally, the system can extract features from information of the machine translation system, these are the confidence features. The QuEst with these features can extract quality indicators for building the quality estimation model.

The machine learning module (see Figure 3) uses scikit-learn library implemented in Python. Using the quality indicators and learning mechanism, the system builds the quality estimation model, which was trained on human annotations. The system can learn from this model and using the results of the model, it can predict the quality scores for new output segments. The learning mechanism uses SVR (support vector regression). Because the program learns from human judgements, the correlation is high between the estimated results and the human evaluation. The evaluation is real-time. The QuEst system can give more accurate quality indicators for the machine translated segments than the traditional methods. For the English-Spanish language pair, Beck et al. [9] for the feature extraction task, tried nearly two hundred features, but there were only 17 features that were really relevant to the result. Hence the task is to find the most relevant baseline feature set for the English-Hungarian language pair.

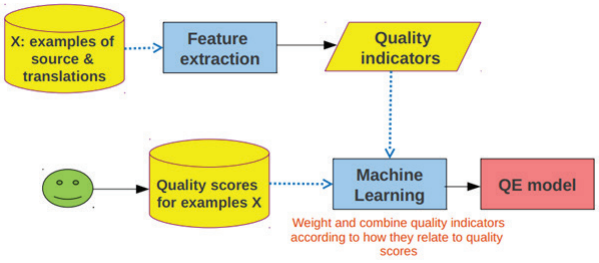


Fig. 3. Machine learning module [8]

TABLE I
BLEU AND NIST SCORES

BLEU	0.1511
NIST	4.8523

III. BENCHMARKING

The Hungarian translations were translated by MOSES [10]. The training set for MOSES was the Hunglish corpus [11] with 1 million sentences. The test set contained 3000 sentences, which were randomly selected from the Hunglish corpus. The source language was English. The target language was Hungarian.

For using QuEst system, the 3000 test set was split into 80%-20% training and test set for evaluation.

To evaluate the QuEst system, MAE (mean absolute error) and RMSE (root mean square error) metrics were used.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|H(s_i) - V(s_i)|$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(H(s_i) - V(s_i))^2}$$

where N is the number of test sentences, H is the score predicted by the quality estimation model and V is the actual score for the test sentence.

IV. EXPERIMENTS AND RESULTS

The first task was building the MOSES translation system for English-Hungarian. After building and training MOSES, I evaluated the system with BLEU and NIST (see Table I). Based on the BLEU score, the MOSES translation is $\sim$15% correct. Thereafter, I set up the QuEst system for English-Hungarian to evaluate the MOSES.

For building the QuEst system, we need feature indicators and human annotated translations. My first task was create the human annotated training set. I developed a website for human annotators to evaluate the translations. In this website we can see an English source sentence, and a machine translated Hungarian sentence. People can give quality scores from 1 to 5, from two points of view [12]: adequacy and

TABLE II
EVALUATION SCALE FOR HUMAN ANNOTATORS

Adequacy	Fluency
1 - none	1 - incomprehensible
2 - little meaning	2 - disfluent Hungarian
3 - much meaning	3 - non-native Hungarian
4 - most meaning	4 - good Hungarian
5 - all meaning	5 - flawless Hungarian

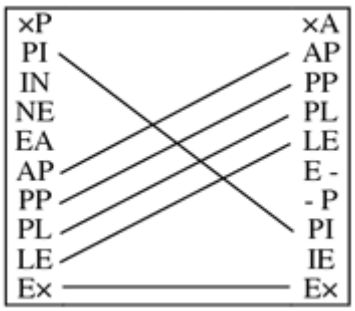


Fig. 4. OrthoBleu [13]

fluency (see Table II). 100 sentences were evaluated by 4 people. The problem was that 3000 sentences, should have been evaluated. Thus the human scores had to be replaced by an automatic evaluation method. For replacement, the OrthoBleu [13] method was chosen. The OrthoBleu method uses letter based n-gram model. It splits the source and the target sentences into n length letter sequences and counts f-measure from them.

I counted n=1 to n=10 and word-based OrthoBleu values, but none of these result had high correlation enough with human scores, hence using linear regression model and all the OrthoBleu scores trained on human scores, I created a predicted score, which was named predicted-orthobleu value. I gained 0,6 Pearson correlation, which was high enough for the further research.

The other main part of the research was the feature extraction task. Features may belong to two categories: ”blackbox” (features independent from the machine translation system) and ”glassbox” (features depend on the machine translation system). In my research I tried 67 blackbox features. First I evaluated the test set with the 67 features, and the 17 baseline feature set, which is optimized for English-Spanish. My task was to find the most relevant features from these 67 features for English-Hungarian and make the result better than that of the 67 and the 17 feature sets.

First I used the Random Lasso [14] method to filter the 67 features. After the filtering, 45 features remained. Then I sorted the features in descending order by the coefficient values used in the regression model. Thereafter I took the feature with the highest coefficient value and I evaluated the system with it. Then I added the second one and I evaluated again, then the third one and so on. With this method I

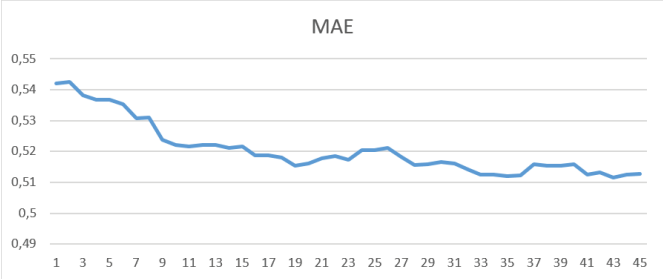


Fig. 5. MAE scores

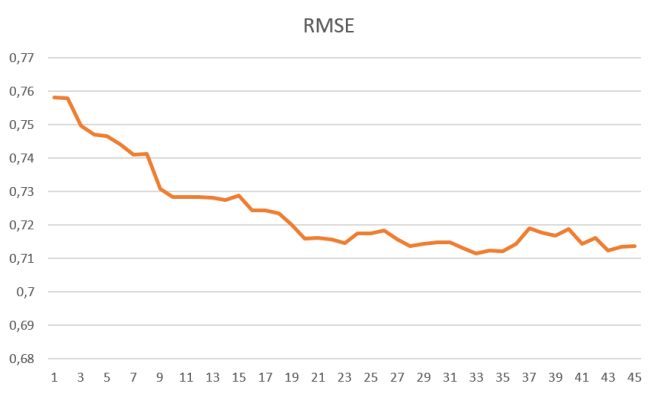


Fig. 6. RMSE scores

TABLE III
RESULTS

	23 baseline (en-hu)	17 baseline (en-es)	67 features (en-hu)
MAE	**0.5037**	0.5161	0.5184
RMSE	**0.7004**	0.7125	0.7172

could not achieved much better result. But as we can see in Figure 5 and Figure 6, during this process there were features which increased the result achieved by having only previous features, while others caused decreasement. I took the features, which increased the mae or rmse scores and I sorted them in descending order by the degree of increasement, and I restarted adding the features. With this method, the first 23 features produced the best result. Therefore I created the baseline feature set with this 23 features.

The final baseline set contains 23 features (see Table IV). The other features made the result worse. As we can see in the Table III, the result of the baseline set optimised for English-Hungarian language pair is really better than the feature set which contains 67 features, or than the baseline set (17 features) for English-Spanish language pair. The final MAE value is 0.5037, which means on average, the quality score predicted by QuEst differs ∼10% from the predicted-orthobleu values.

TABLE IV
23 BASELINE FEATURE SET FOR ENGLISH-HUNGARIAN

23 baseline set
Number of tokens in the source sentence.
Average source token length.
Percentage of numbers in the source.
Ratio of percentage of tokens a-z in the source and tokens a-z in the target.
Number of punctuation marks in the target sentence.
Number of punctuation marks in the source sentence.
Average number of translations per source word in the sentence (threshold in giza1: prob $>$ 0.5).
Source sentence LM perplexity.
Target sentence LM perplexity.
Percentage of distinct trigrams seen in the corpus (in all quartiles).
Absolute difference between number of tokens and source and target normalised by source length.
Average number of translations per source word in the sentence (threshold in giza: prob $>$ 0.5) weighted by the inverse frequency of each word in the source corpus.
Percentage of distinct bigrams seen in the corpus (in all quartiles).
Average unigram frequency in quartile 4 of frequency (lower frequency words) in the corpus of the source sentence.
Source sentence log probability.
Average number of translations per source word in the sentence (threshold in giza1: prob $>$ 0.5) weighted by the frequency of each word in the source corpus.
Number of occurrences of the target word within the target hypothesis (averaged for all words in the hypothesis - type/token ratio).
Average number of translations per source word in the sentence (threshold in giza: prob $>$ 0.01) weighted by the inverse frequency of each word in the source corpus.
Absolute difference between number of ! in source and target.
Average trigram frequency in quartile 4 of frequency (lower frequency words) in the corpus of the source sentence.
Percentage of tokens in the target which do not contain only a-z.
Percentage of distinct unigrams seen in the corpus (in all quartiles).
Average unigram frequency in quartile 3 of frequency (lower frequency words) in the corpus of the source sentence.

V. CONCLUSION AND FUTURE WORK

I built a statistical machine translation system (MOSES), for English-Hungarian language pair. Then I evaluated the machine translated sentences with BLEU and NIST methods, which use reference translation, then with QuEst methods, which does not use reference translation.

I built the QuEst system for English-Hungarian language pair. Then I proposed a baseline feature set, which contains 23 features and is optimized for Hungarian.

In the future, the first task, which I have already started, is create a large human annotated training set for the QuEst. If there is a large and useful training set, I can improve or replace the predicted-orthobleu values to gain higher correlation. Another task is to extract more features, for example glassbox features or linguistic features. I also would like to use semantic models for the evaluation. Last but not least, I would like to use the QuEst system for other language pairs, like Hungarian-Chinese.

ACKNOWLEDGMENT

I would like to say thank you to Laki László who gave me lots of good ideas and inspiration, and to Gábor Prószéky for their help.

REFERENCES

[1] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: A method for automatic evaluation of machine translation," in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ser. ACL '02. Stroudsburg, PA, USA: Association for Computational Linguistics, 2002, pp. 311–318. [Online]. Available: http://dx.doi.org/10.3115/1073083.1073135

[2] C.-Y. Lin and F. J. Och, "Automatic evaluation of machine translation quality using longest common subsequence and skip-bigram statistics," in *Proceedings of the 42Nd Annual Meeting on Association for Computational Linguistics*, ser. ACL '04. Stroudsburg, PA, USA: Association for Computational Linguistics, 2004. [Online]. Available: http://dx.doi.org/10.3115/1218955.1219032

[3] A. Lavie and A. Agarwal, "Meteor: An automatic metric for mt evaluation with improved correlation with human judgments," 2005, pp. 65–72.

[4] M. Snover, B. Dorr, R. Schwartz, L. Micciulla, and J. Makhoul, "A study of translation edit rate with targeted human annotation," in *In Proceedings of Association for Machine Translation in the Americas*, 2006, pp. 223–231.

[5] M. Gamon, A. Aue, and M. Smets, "Sentence-level mt evaluation without reference translations: Beyond language modeling," in *In European Association for Machine Translation (EAMT*, 2005.

[6] J. S. Albrecht and R. Hwa, "2007b. regression for sentence-level mt evaluation with pseudo references," in *In Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL*, pp. 296–303.

[7] L. Specia, K. Shah, J. G. de Souza, and T. Cohn, "Quest - a translation quality estimation framework," in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. Sofia, Bulgaria: Association for Computational Linguistics, August 2013, pp. 79–84. [Online]. Available: http://www.aclweb.org/anthology/P13-4014

[8] L. Specia. (2013, Oct.) Quest - an open source tool for translation quality estimation @ONLINE. [Online]. Available: http://staffwww.dcs.shef.ac.uk/people/L.Specia/projects/quest.html

[9] D. Beck, K. Shah, T. Cohn, and L. Specia, "Shef-lite: When less is more for translation quality estimation," in *Proceedings of the Workshop on Machine Translation (WMT)*, August 2013.

[10] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. Dyer, O. Bojar, A. Constantin, and E. Herbst, "Moses: Open source toolkit for statistical machine translation," in *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*, ser. ACL '07. Stroudsburg, PA, USA: Association for Computational Linguistics, 2007, pp. 177–180. [Online]. Available: http://dl.acm.org/citation.cfm?id=1557769.1557821

[11] D. Varga, H. P., A. Kornai, V. Nagy, L. Németh, and V. Trón, "Parallel corpora for medium density languages," in *Recent Advances in Natural Language Processing Conference*, Borovets, Bulgaria, September 2005, pp. 590–596.

[12] P. Koehn, *Statistical Machine Translation*, 1st ed. New York, NY, USA: Cambridge University Press, 2010.

[13] FTSK. (2014, May) Orthobleu mt evalution based on orthographic similarities @ONLINE. [Online]. Available: http://www.fask.uni-mainz.de/user/rapp/comtrans/d05orthobleu.html

[14] S. Wang, B. Nan, S. Rosset, and J. Zhu, "Random lasso," *The Annals of Applied Statistics*, vol. 5, no. 1, pp. 468–485, 03 2011. [Online]. Available: http://dx.doi.org/10.1214/10-AOAS377

Appendix

PhD-studies started in 2011-2012:
- István Endrédy
- Anna Gerencsér-Horváth
- Balázs Knakker
- László Kozák
- Balázs Ligeti
- Attila Novák
- Borbála Siklósi
- Zoltán András Tuza

PhD-studies started in 2012-2013:
- Vamsi Kiran Adhikarla
- Dóra Bihary
- Bence József Borbély
- Erzsébet Farkas
- Katharina Hofer
- Balázs Indig
- Attila Gyula Jády
- Mátyás Jani
- Imre Benedek Juhász
- András József Laki
- Dénes Pálfi
- Ágnes Polyák
- Norbert Sárkány
- Ádám Vály

PhD-studies started in 2013-2014:
- Bernadett Ács
- Bálint Balázs
- Sándor Földi
- Zselyke Magyari
- Tamás Molnár
- Máté Németh
- Ádám Papp
- Tamás Pardy
- Zsófia Sztyéhlikné Bérces
- Eszter Tóth
- Zijian Győző Yang
- Tamás Wilheim