Representation and recognition methods on parallel and unconventional machines

Attila Stubendek

Thesis of Ph.D. Dissertation

Supervisors: Dr. Péter Szolgay, dr. Kristóf Karacs

Pázmány Péter Catholic University

Faculty of Information Technology and Bionics

Roska Tamás Doctoral School of Sciences and Technology

Budapest, 2018.

1. Introduction

Designing and building a machine that helps people, is the most important challenge and achievement of our civilization. For thousands of years by the development of tools used for moving and handling, one had to bear less and less burden on its shoulder and could devote more time to art and thinking. As a result, new, better machines were created giving aid to people in the field of arithmetic, logic, and even making inferences. Nowadays the primary indicator of the development of our civilization is not just the capability of the human family, but the capabilities of the machines made and taught by the people.

Today state-of-the-art engineering aims to understand and reproduce the process of the human thought, recognition, and learning – to mimic everything that has put us to the top of the evolution. Meanwhile, the scientific community will meet again and again with the recognition that life and patterns of organizations in nature, the excellent work of God is the perfect inspiration for creating the simplest and most complex devices and alike.

Sight is the most fundamental way of human perception. Therefore, any machine that aims to interact with humans and the human world is required not only to be capable of sensing physical light but also to be able to interpret the image.

Artificial vision or computer vision is employed in a number of applications that make the daily life easier for millions. Artificial vision provides more accurate and faster healthcare, automated industry and quality checking, safe unmanned vehicles, and exploration of places where humans cannot enter.

The quality of life of visually impaired can be improved by orders of magnitude by devices using reliable artificial vision.

Automated image understanding utilizes methods of several different disciplines. Image or image flow is first preprocessed, and mathematical descriptions are generated, composing an object representation. If given a specific task regarding the image, descriptions are used for analysis, modeling or comparison. Thus, the key of an effective image understanding is encapsulated in the choice of the features: the *representation* must highlight those *modalities* of the image that are *discriminative* for the given task and neglect other *irrelevant* aspects.

In a mathematical interpretation, object recognition is a *mapping* from an image to a few dimensional output vector. The output is an answer to the actual visual query regarding the presence, the position, the category or a numerical feature of the object. To give an answer, a *decision machine* is required that has the appropriate *a priori knowledge* gained by learning from already known examples.

Methods and architectures employed in computer vision applications are consciously getting closer to the known mechanisms of the human vision system and perception. Classical architectures and concepts are universal, although in several special cases dedicated architectures are significantly faster and provide lower power consumption by employing multiple processing cores and utilizing the *precedence of locality*. For that very reason, the computational schemes of the available devices should be considered during the design of any algorithm.

In my research, I have been striving to create general methods. However, a particular area of usage, the design and the implementation of algorithms employed in the Bionic Eyeglass was the primary source of my motivation. The Bionic Eyeglass is an application collection for visually impaired, realized as a part of a research project of the Hungarian Bionic Vision Center. Functions asked by the potential users of the applications and the limitations of hardware available defined the requirements, aims, and constraints concerning the algorithms. Objects to be recognized – pictograms, banknotes – are all two-dimensional and rigid patches, thereby in my research, I focused on algorithms that *describe and recognize shapes* this kind of nature.

The shape is the most meaningful visual aspect of an object, more than the color, material or texture. Usually, the two-dimensional image is a projection of a three-dimensional object. The two-dimensional projection of a flat patch depends only on the mutual position of the camera and the subject, and the patches from different angular viewports can be transformed to each other. In contrast, the projection of 3D objects depends on the actual positioning of the object – a sculpture or a mug has a completely different shape viewed from different sides, not necesseraly transformable to each other.

In my thesis I focus on flat, rigid shape patches obtained from an image taken by a standard camera or a cell phone, that has an extent only in 2D, such as pictograms, road signs, patches of banknotes. The challenge in recognition of such shapes lies in the high variety of objects, resulted by various lighting conditions, unambiguous segmentation and uncontrolled camera handling.

The basic mathematical features of a shape, as perimeter, area, eccentricity consist of only one or few scalar values, though by which they express the feature of the shape in a very compressed way. Complex shape descriptors belong to two groups. Edge-based features describe the contour of the shape; the region-based descriptors describe the shape based on every point of the shape. In my thesis I describe the most important shape descriptors.

Shapes of objects seen in our everyday life are compound, Thereby no unified aspect can be found that can categorize all the shapes unambiguously and clear. In my research, I aimed to create a compound shape descriptor, in which several modalities of the shape are described, and are suitable for fast, robust and reliable *recognition*.

Learning is essential to an *intelligent agent* as the capability of improving its behavior based on experience. In the *supervised learning* approach, the goal is to *predict* the decision mapping from *inputs* to output *labels*, when a certain number of elements of the mapping are given. The set of these a priori known mappings is called *training set*.

Depending on the output type, the label can be a qualitative or a quantitative variable. According to the naming convention, predictions of continuous, quantitative outputs are called *regression*, and the inference is called *classification* if the output label is discrete qualitative variable.

If the class labels take more than two values, we speak about multiclass classification. In specific tasks the number of the classes is not necessarily limited. In this case, we distinguish *relevant* and *not relevant* classes from the aspect of the actual task. The classification process is different if it is known that the input belongs to some of the relevant classes (i.e., boy or a girl), or if the input may be the member of any other classes (voices of birds are recognized, but other noises also may appear on the input record). Real-world objects and their representation often belong to an unbounded number of classes. Depending on the task, out of these classes, the number of relevant ones may be orders of magnitude smaller than the number of irrelevant classes, thus representing each irrelevant class is not efficient, if at all feasible.

The most important distinction of classifiers is based on the relationship between the input data size and the number of the prediction model parameters. If the model size changes with the amount of training data, we speak about *non-parametric models*, and the models having fixed number of parameters for a fixed task, independently from the training set, are called *parametric models*. In my thesis I will show some examples of both type of models, that are later mentioned or compared.

In the second part of my thesis I propose an extension to the standard nearest neighborhood method that makes the model capable to handle classification problems with non-relevant classes as well, and by a two-level decision model provide faster classification.

The efficient implementation on *dedicated architecture* was one of the main motivations during the design of both the descriptor and the classifier. Due to the requirements of the applications I aimed to design algorithms that are able to run on *kiloprocessor architectures* with low power consumption, and at the same time, on standard architectures of a computer or phone, provide reasonable processing time.

The speed of the digital computing blocks used in processors is approaching their physical borders. Following Moore's Law about the development of architectures, the innovation of multi-core processors using parallel architectures began in the last decade. Meanwhile, the interest has increased for alternative technologies as well, breaking out of the Neumann's paradigm. One of these principal branches of research is the development of *non-Boolean*-based, non-digital computing units implemented in a network of interacting nanomagnets called Spin Torque Oscillators. The basic principle of the architecture is to use current to excite the magnets that, depending on the current and the topology, synchronize with a pattern of phases on each oscillator.

In my work, I defined the terms of data and program on oscillatory networks and showed how to use them as computing units. I conducted experiments aiming to find the class of problems that are effectively computed on oscillatory networks.

2. Methods

The designed algorithms, tests and experiments were mostly implemented in Matlab[1] environment. On lower levels of visual processing (matrix operations, convolution, morphology - propgen, torque computing) I employed the built-in functions of Matlab, but on higher levels I created my own implementations for better understanding and to allow higher flexibility to the methods.

The WEKA[2] and RapidMiner[3] machine learning implementation collections helped me a lot in exploring the basic applicability of methods. Based on these results I made my own implementations of the nearest neighborhood and neural network model, and also implemented a framework for genetic algorithms.

I tested the banknote recognition application of the Bionic Eyeglass with people with visual impairments. At the beginning of the tests, the participants were given detailed information about the operation of the device. After that, they had to identify banknotes without any external help, relying on the device only. During the tests, we stored every processed frame with the corresponding decisions, so the experiments can be completely reproduced later in a simulation environment. With these steps, the team was able to track the development of the algorithms. During the live tests we have written records and notes to collect all remarks according to the behavior of the algorithms, the usability of the device or to other environmental circumstances. Test sets presented in this thesis are compound of these live tests.

[1] http://www.mathworks.com/products/matlab/
[2] http://www.cs.waikato.ac.nz/ml/weka/
[3] https://rapidminer.com/

3. New Scientific Results

Thesis I.

I designed a new general shape descriptor called the Global Statistical and Projected Principal Edge Distribution descriptor that is based on different modalities. It combines the advantages of the global statistical and local edge-based descriptors. I implemented the descriptor and verified its efficiency through experiments.

Related publications: [A1][A3][A4][A7]

The descriptor consists of global mathematical shape features, and an edge based local feature set. The method of assembling more shape descriptors into one aims to describe different modalities independently from each other, and to speed up the recognition process.

The first part of the descriptor contains statistical shape features that describe the shape as whole. Employing these features resulted in a higher cover ratio, and accelerated decision accuracy in case of a comparison-based classification.

The first property is the eccentricity of the shape, which characterizes the elongation of the shape with one scalar value. The second property is the ratio of the shape area and the area of its bounding box. The following eight features are the first four statistical moments of the vertical and horizontal histograms.

The edge-based shape description part (Extended Projected Principal Shape Edge Distribution, EPPSED) characterizes the local properties of the shape. The algorithm was inspired by the PPED image descriptor, which selects and projects principal edge values. (see Figure 1.)

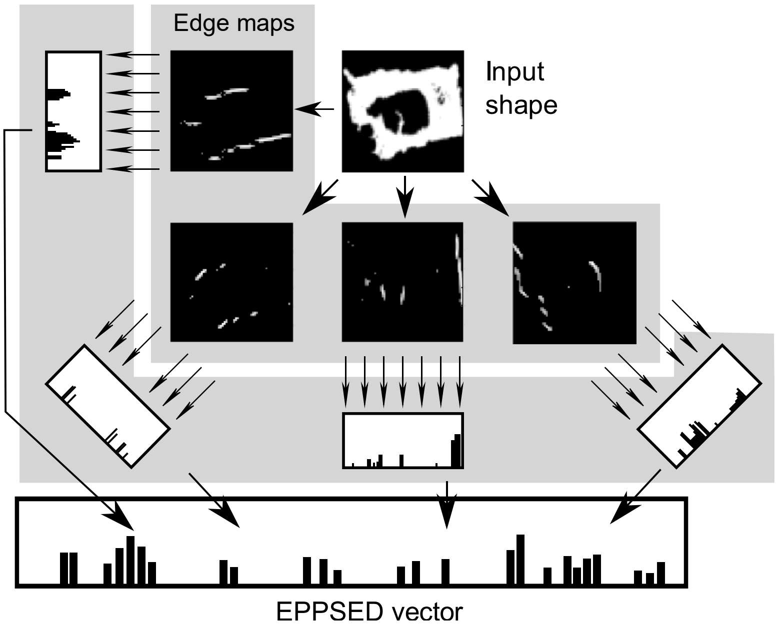

Figure 1. Construction of the EPPSED feature vector. Edges are detected in four directions; then thresholding and maxima selection are applied; finally, projections are concatenated and normalized.

Edges are detected in four directions, resulting in four edge maps. For every pixel location in the four edge maps, the principal value is highlighted, and the other values are weakened or totally neglected. Highlighting is performed by a soft-thresholding function, which decreases the effect of overfitting and provides the ability to use the algorithm on different architectures with different number representations. I measured the efficiency of the soft-thresholding, and it improves the performance of the shape description.

The rotation invariance can be ensured at two points of the recognition process. The first approach incorporates the invariance in the description directly or by a preprocessing including an angular normalization. The second point is at the classification phase, on the employing an invariant metric or by storing (all) the different rotations of one instance in the training set.

Descriptors of the family of the PPED are not rotation invariant. Therefore the rotation invariance is ensured by rotationally redundant training set. Although this is closer to the human sense, it would result in a vast database. From this consideration, I normalize the shape angularly, based on the orientation of the shape. The blob is rotated to have its major axis parallel to the horizontal axis. Finally, the shape is rotated by 180 degrees if its center of gravity is on its right side.

The scale-invariance is ensured by normalization, by resizing the blob to fit into a 64×64 pixel window. The position-invariance in the vertical dimension is ensured by moving the shape in the middle, thus in the other dimension, the shape touches both sides of the frame edge.

The shape descriptor algorithm I designed is feasible for implementing and using on dedicated VLSI architecture and on a Cellular Wave Computer.

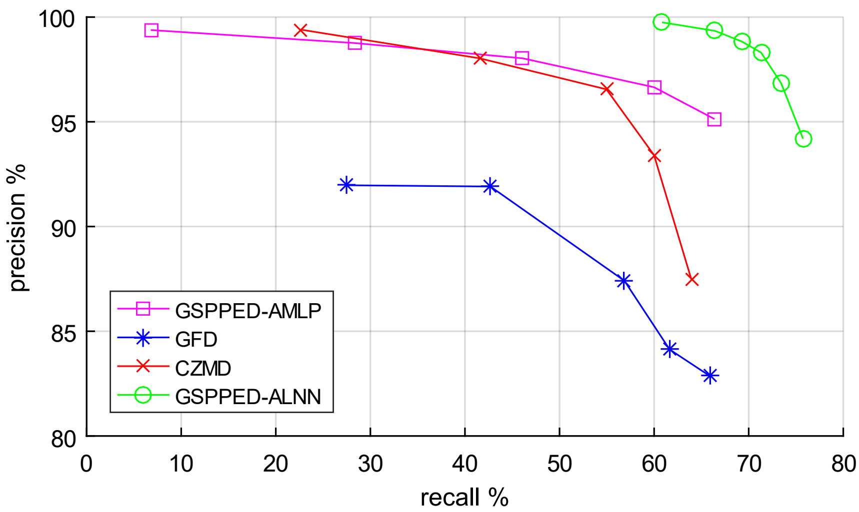

Figure 2. The GSPPED shape description compared to other shape features

I compared the GSPPED shape descriptor with the most widely used shape descriptors on a shape set containing patches from the Hungarian Forint Banknotes. Results show that the performance of the GSPPED is higher than the performance of the Complex Zernike Moments and the Generic Fourier Descriptor. (Figure 2.)

Thesis II.

I designed and implemented a reliable and robust, two-level, parametric, memory-based classification method. The classifier is able to handle multiple classes, it has the ability to reject distant inputs, and can handle compound features.

Related publications: [A1][A3][A4]

The two-level decision adapts to the semantics of the descriptor and to the discriminative power of each descriptor parts. First, the mathematical and statistical shape-descriptors are being compared for fast filtering, then by employing the Adaptive Limited Nearest Neighborhood algorithm, the edge-based properties are being compared.

The essence of the method is a multi-level application of the compound descriptor that speeds up the decision and increases the cover of the recognition. When comparing with representative instances, first only simple mathematical properties are analyzed to examine the presence of a basic similarity. Although the final decision cannot be expected from this primary comparison, it is proven to be efficient to filter and reject elements that obviously differ from the instances of the representative set.

II.1 I demonstrated that in a two-level, comparison-based classification, filtering based on simple but highly expressive features can both speed up the classification and can significantly increase the recall, if the goal of the second level is to maximize the precision.

In case of a comparison-based classification, the complexity of the decision increases with the size of the training set. Thus, if the difference exceeds a certain threshold level during the comparison, the input can be rejected without continuing the evaluation. The first properties of the GSPPED descriptor contain highly expressive features, that are suitable for rejection.

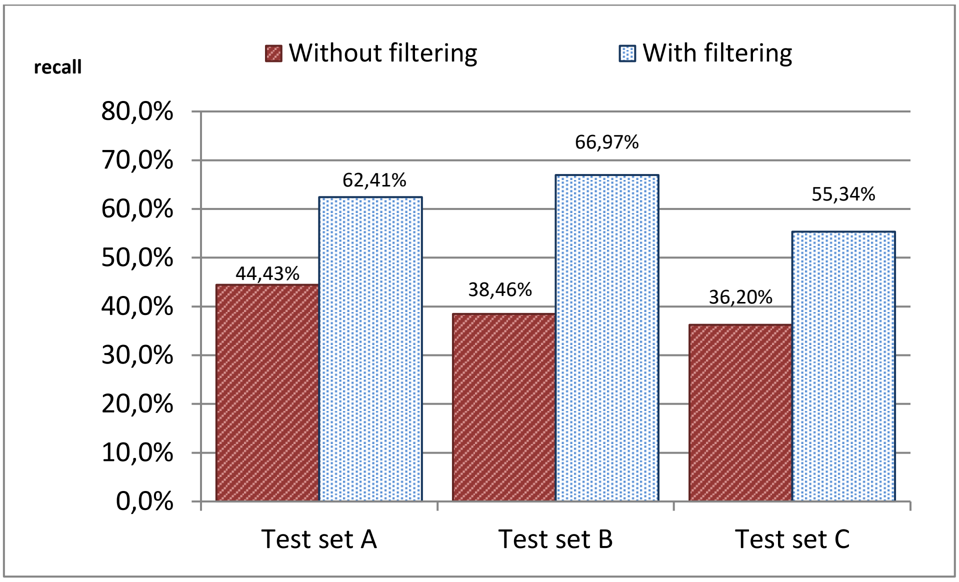

Figure 3. Recall depending on the usage of the filtering preprocessor.

Applying the proposed filtering, the average decision time decreases to 12% compared to the decision time without filtering. Besides the significant acceleration, filtering has a secondary effect, namely that the cover of the recognition increases by 17% in average in case when filtering is applied. This seemingly paradox phenomenon can be explained as follows: features applied for filtering are orthogonal to those applied in the second phase of comparison in information theoretical sense. To maximize the precision in the second phase, the acceptance radius is set as a function of the distance of the closest element of other classes. As a result, acceptance regions increase since instances closer to the representative elements might be rejected. The explanation is verified through measurements, where acceptance regions are 25% larger with filtering. Since the second phase focuses on maximizing the precision, the precision remains independent of the filtering. (Figure 3.)

II.2 I designed and implemented the Adaptive Limited Nearest-Neighbor (AL-NN) classifier. The presented method is capable of rejecting non-relevant (zero class) instances, that is one of the main benefits against standard KNN methods.

Nearest-neighbor classifiers compare inputs to already known, labeled instances of representative sets. To overcome the one of the main drawbacks of the traditional nearest-neighbor classifiers - namely the lack of rejecting non-relevant elements - I extended the algorithm by introducing limits to each stored instance.

Limits are set individually to each element of the representative set based on the irrelevant and out-of-class instances included in the training set. For this very reason I introduced a combined training set including not only relevant instances but non-relevant ones as well. Although it is not discussed in the present thesis, the model is capable of making decisions in case of overlapping classes on the instance-level.

The model is built up as follows: The training set contains labeled instances, including non-relevant (also called as zero class) elements as well. To all relevant instances of the training set a limit is defined based on the surrounding instances. In case of having only non-relevant neighbors, the limit is set to the half of the distance between the closest neighbor and the current trainable element. When only an instance from a different class is within the current filter region, the limit is set to the distance of the nearest outlier and the current example.

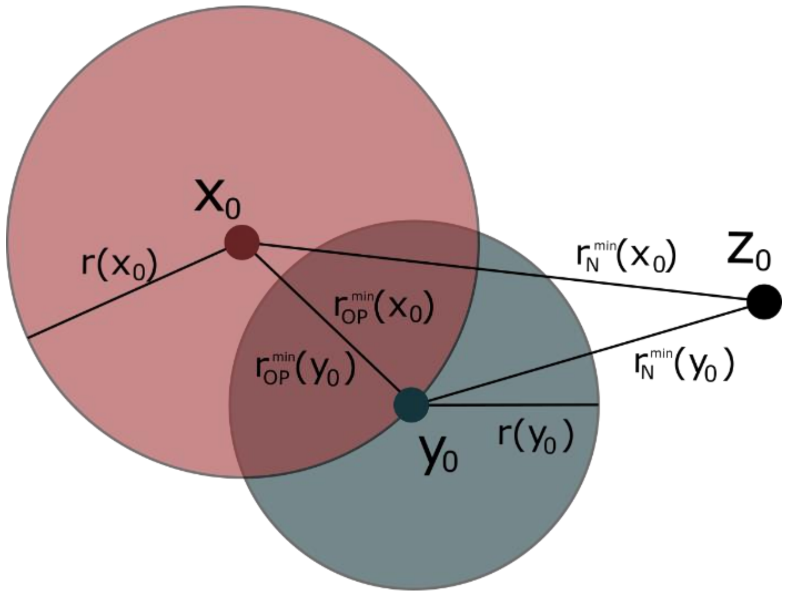

Figure 4. Definition of the acceptance range. x_0 and y_0 represent relevant elements from different classes, z_0 the closest irrelevant element. Acceptance threshold for y_0 ($r(y_0)$) is set to the half of the distance to the closest irrelevant element ($r_N^{min}(y_0)$). Acceptance threshold for x_0 ($r(x_0)$) is chosen as the distance to the closest element of another class ($r_{OP}^{min}(x_0)$). Since z_0 is an irrelevant template, it is not included in the representative set, thus no acceptance region is defined for it.

If all elements within the filter region are of the same class, the limit is defined as the distance of the farthest instance of the class. However, in the most cases more instances of more classes fall close to the example, thus the limit will be the minima of the limits defined above. (Figure 4.) Finally, if the corresponding filter region of the current training instance is empty, we reject the instance due to the lack of information about its surrounding. The result of the training is the representative set, containing only relevant class instances. The limit can be fine-tuned with a vigilance parameter, that can adjust the recall-precision relation as well.

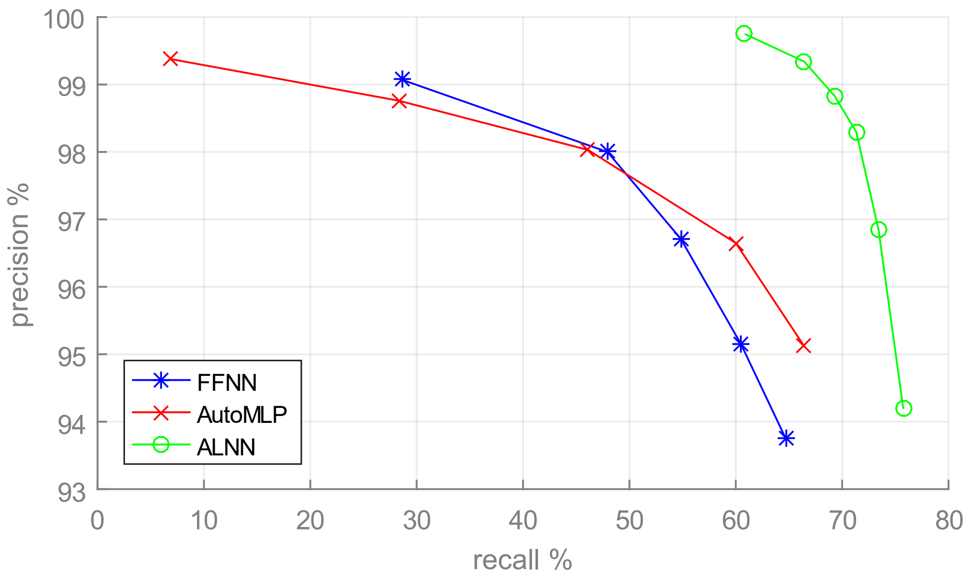

Figure 5. The performance of the ALNN algorithm compared to other classifiers.

I compared the classifier with other widely used methods, like KNNs, AutoMLPs, FFNNs and SVMs. The AL-NN shows higher performance compared to the other methods. (Figure 5.)

II.3 I designed a representative set optimizer and an adaptive online update method for the AL-NN classifier. By using the optimizer, the size of the representative set can be decreased without a significant decrease of the cover. The online update of the representative set provides the ability to include new instances from test sets automatically.

The usability of the decision methods is significantly better if the model can be automatically or semi-automatically updated and optimized during tests or measurements. A special scenario of such use cases is when the initial training set contains only a few instances at the beginning of the first test, and the model continuously increases by performing tests.

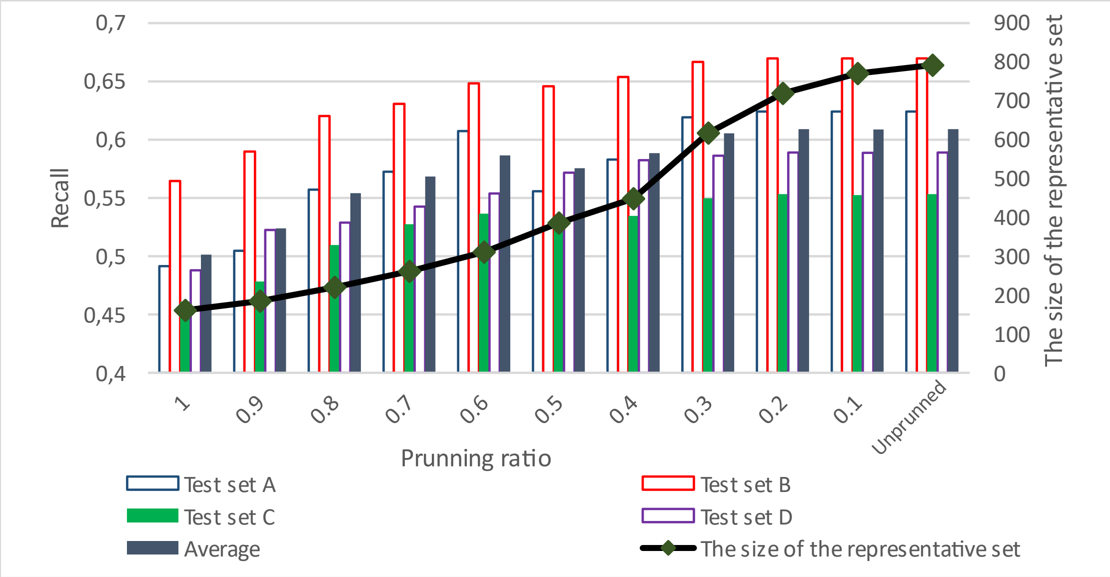

Figure 6. The recall and the size of the representative set depending ont he optimization.

During Adaptive Limited Nearest-Neighborhood classification not only the output class of each input can be determined but also the confidence of the output, based on the ratio of the measured distance and the acceptance radius. In case a new instance that is classified is located at the border of a class, it can be added to the representative set to extend the cover of the current class, or to increase the precision of the borderline between classes. To ensure that the representative set is extended with proper instances, I used the temporal order to verify the classification result.

By estimating the size of the intersections of acceptance regions of the instances in the representative set, redundant elements can be dropped. I suggested an iterative optimizer, where representative instances are ordered based on the number of other instances they cover, and in each step the instance with the larger cover is kept, and other instances covered by it are removed. By employing the optimizer, the representative set can be reduced to the 30-50% of the original size, while the cover decreases only by 5-15%. (Figure 6.)

Thesis III.

I formalized the network of weakly coupled oscillators as a computational unit, and I defined the concepts of a program: the data, the input and the output on the system. I provided an experimental proof that the oscillator networks are capable of resulting in a desired signature on response to a given input excitation, thus the unit is suitable to be used as a computational unit.

Related publications: [A2][A3][A8]

Two or more oscillators are interacting and form a coupled oscillator network. The Spin Torque Oscillators and the networks built from STOs are the subjects of recent researches of non-Boolean computational units on several levels.

The main emphasis today is on investigating the feasibility of using the STOs and other non-Boolean units by simulations. In my research an STO network model is used, where the interaction is determined only by the topology of the network. The atomic operation is considered as a synchronization, where the individual oscillators might be in a different phase. I formalized the program as the topology and the oscillator dynamics. The data is implemented as a phase-difference array, the input is the excitation, and the output is represented as the data of the output nodes.

I designed and implemented a method that creates an OCNN topology with a genetic algorithm, that, as an OCNN, transforms an input to a separable output space. The applicability of the method is proven through experiments.

In classical, two level object detection models the multidimensional input is first reduced in dimensionalities through special descriptors, and the final decisions are made through these compressed representations. Trainability and correctness of the decision highly depends on both compactness and discriminative power of the descriptors according to the output classes.

I proved the OCNN’s capability of performing the desired transformation with a learning example. I proposed a new classification method by inserting an OCNN array between the feature extractor and the classifier, thus the desired function of the array was to transform the feature to improve the final classification results. To design the proper topology of the OCNN array, I used a genetic algorithm that maximized the classification performance. I tested the proposed system with the classification of EPPSED data and verified the results with H-MAX vectors.

During the experiment I succeeded in improving classification performance on all the test sets. I also proved by measurement that the compactness of the classes in the feature space increases significantly.

I designed and implemented a method that based on input-output pairs, defines the topology of a pyramidal oscillator network built up of STOs, performing the desired classification. The applicability of the method for classification of spatio-temporal signals is proven through experiments.

I proved the feasibility of using a STO network as a computational unit by building a classifier. Classification encompasses not only a transformation but also separation, thus the task is more complex. I tested the classifier on static and spatiotemporal data as well.

I proposed a multi-layer topology for classification. The number of the oscillators in the top layer equals the dimension of the output. In the next layers the number of the oscillators is decreasing until the last layer with only two oscillators.

The topology had been tested for static and spatio-temporal inputs. A static sample performing a summation and a XOR was examined with a success. A moving signal function was used as a dynamic input with changing directions and noise. In each of the experiments the capability of teaching topology was proved.

4. Application of the results in practice

The first and second thesis was developed in the framework of the Bionic Eyeglass project, actively utilizing experiences collected from the live tests. By employing the shape description and classification I proposed, a complex classifier has been created that is suitable for the recognition of Hungarian Forint banknotes. The application identifies the banknote based on more classifiers, resulting in a robust, reliable device.

Beyond banknotes, the system might be used for other visual tasks as well, which require the classification of rigid, flat shapes. Within the Bionic Eyeglass project algorithms might be used for recognition of other banknotes and pictograms, reading of LED displays and identification of information boards.

The developed descriptor and classifier are already implemented on FPGA as well, providing faster processing compared to standard CPUs. (The implementation on FPGA was not the part of my research.)

In the future, oscillator networks may exchange the part of present-day classic CMOS components. Application areas are examined in many research institutes, this thesis is also part of this process.

5. Acknowledgement

First of all, I would like to thank my supervisor Tamás Roska that he involved me into the scientific community, and launched my scientific career, helped me with his knowledge, enthusiasm, and respect. I would like to thank him for allowing me to visit research institutes in Tokyo and in Portland.

I am especially grateful to Kristóf Karacs for being my patient mentor during my doctoral studies. I am grateful for his ideas and extremely detailed, constructive critiques.

I am thankful to the leaders of the Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, especially to Péter Szolgay, Judit Nékyné Gaizler, Ágnes Bérczesné Novák for creating and managing such a professional and value-oriented research center.

I acknowledge the kind help of my supervisor in Tokyo, Tadashi Shibata, that he helped me to broaden my horizons and to make progress in my researches.

I am the most grateful to my family for teaching me to respect our values and for the inspiration to improve my knowledge continuously.

I am thankful to my bride Timi for her patience and love.

I am grateful to my colleagues and friends, András Horváth, Mihály Radványi, Miklós Koller and Tamás Zsedrovits and Tamás Fülöp for the discussions and support.

6. List of publications

[A1] A. Stubendek, K. Karacs, “Shape Recognition Based on Projected Edges and Global Statistical Features”, Mathematical Problems in Engineering, vol. 2018, Article ID 4763050, 18 p, 2018.

[A2] A. Horváth, M. Koller, A. Stubendek, T. Roska, “Spatial-temporal Event Detection via Frameless Cellular Wave Computing - a Review”, Nonlinear Theory and Its Applications, IEICE, 2014, vol 5 (3), pp. 391-408

[A3] A. Stubendek, K. Karacs, T. Roska, „Shape Description Based on Projected Edges and Global Statistical Features”, International Symposium on Nonlinear Theory and its Applications (NOLTA 2014), Luzern, Switzerland, 2014

[A4] K. Karacs, M. Radványi, A. Stubendek and B. Bezányi, “Learning Hierarchical Spatial Semantics for Visual Orientation Devices,” in Proc. of the IEEE/CAS-EMB 2014 Biomedical Circuits and Systems Conference (BIOCAS 2014), Lausanne, Switzerland, pp. 141–144, Oct. 2014

[A5] A. Horváth, A. Stubendek, T. Roska, “One dimensional cellular array of in-plane spin torque oscillators,” 13th IEEE International Conference on Nanotechnology, Beijing, China, 2013

[A6] Z. Solymar, A. Stubendek, M. Radvanyi, K. Karacs, "Banknote Recognition for Visually Impaired" in Proc. of the European Conference on Circuit Theory and Design (ECCTD’11), Linköping, Sweden, Aug 2011.

[A7] M. Radvanyi, Z. Solymar, A. Stubendek, K. Karacs, „Mobile Banknote Recognition – Topological Models in Scene Understanding”, in Proc. of the 4th International Symposium on Applied Sciences and Communication Technologies (ISABEL 2011), Barcelona, Spain, 2011.

[A8] T. Roska, A. Horvath, A. Stubendek, F. Corinto, Gy. Csaba, W. Porod, T. Shibata, G. Bourianoff, “An Associative Memory with Oscillatory CNN Arrays using Spin-Torque Oscillator Cells and Spin-Wave Interactions Architecture and End-to-end Simulator,” 13th IEEE international workshop on cellular nanoscale networks and applications, Turin, Italy, 2012.

7. References

[1] D. Poole, A. Mackworth, “Artificial Intelligence: Foundations of Computational Agents”, Cambridge University Press, ISBN 978-0-521-51900-7, 2010.

[2] K. P. Murphy, “Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series)”, The MIT Press, ISBN-13: 978-0262018029, 2006.

[3] T. Hastie, R. Tibshirani, J. Friedman, “The Elements of Statistical Learning – Data Mining, Inference, and Prediction”, Springer Series in Statistics, ISBN: 978-0-387-84857-0, 2011.

[4] Ch. M. Bishop, “Pattern Recognition and Machine Learning”, Springer Science + Business Media, LLC, ISBN: 978-0387-31073-2, 2009.

[5] S. Russel, P. Norwig, “Artificial Intelligence: A modern approach”, Pearson Education, ISBN 963545411-2, 2010.

[6] W. Ertel, “Introduction to Artificial Intelligence”, Springer Science + Business Media, LLC, ISBN: 978-085729-298-8, 2011.

[7] M. Kantrardzic, “Data Mining: Concepts, Models, Methods, and Algorithms”, Institute of Electrical and Electronic Engineers, 2nd ed., ISBN: 978-0-470-89045-5, 2011.

[8] I. Witten, E. Frank, “Data Mining: Practical Machine Learning Tools and Techniques”, 2nd ed., Morgan Kaufmann series in data management systems, ISBN: 0-12-088407-0, 2005.

[9] P. Flach, “Machine Learning: The Art and Science of Algorithms that Make Sense of Data”, Cambridge University Press, ISBN: 978-1-107-09639-4, 2012.

[10] S. Haykin, “Neural Networks and learning machines”, 3rd edition, Pearson Education Inc, ISBN: 978-0-13-129376-2, 2009

[11] M. L. Minsky, S. A. Papert, “Perceptrons: An Introduction to Computational Geometry”, MIT Press, Cambridge, ISBN: 0-262-63022-2, 1969.

[12] S. Marshland, “Machine Learning: An Algorithmic Perspective”, Chapman & Hall / CRC machine learning & pattern recognition series, ISBN: 978-1-4200-6718-7, 2009.

[13] J. W. Rudy, “The neurobiology of Learning and Memory”, Sinauer Associates, ISBN: 978-0-87893-669-4, 2008

[14] R. O. Duda, P. E. Hart, D. G. Stork, “Pattern Classification”, A Wiley-Interscience Publication, John Wiley & Sons, 2nd edition, ISBN: 0-471-05669-3, 2001.

[15] E. Alpaydin, “Introduction to Machine Learning”, MIT Press, Cambridge, ISBN: 9780262012119, 2004.

[16] D. E. Goldberg, “Genetic Algorithms in Search, Optimization and Machine Learning”, Addison-Wesley, Reading, ISBN: 0-201-15767-5, 1989.

[17] M. Mitchell, “An Introduction to Genetic Algorithms” Cambridge, MA: MIT Press, ISBN 9780585030944., 1996.

[18] A. Andreopoulos, John K. Tsotsos, “50 Years of object recognition: Directions forward”, Computer Vision and Image Understanding, vol. 117 (8), pp. 827-891, 2013.

[19] W. K. Pratt, “Digital Image Processing”, John Wiley & Sons, ISBN: 0-471-37407-5, 2001.

[20] I. Biederman, G. Ju, “Surface versus edge-based determinants of visual recognition”, Cognitive Psychology, vol. 20, pp. 38–64, 1988.

[21] L. K. Samuelson, L. B. Smith, “They call it like they see it: Spontaneous naming and attention to shape”, Developmental Science, vol 8 (2), pp. 182–198, 2005.

[22] D. K. Prasad, “Survey of the problem of object detection in real images.“ International Journal of Image Processing (IJIP), vol. 6.6, pp. 441. 2012.

[23] S. L. Tanimoto, “An Interdisciplinary Introduction to Image Processing”, Massachusetts Institute of Technology, ISBN: 978-0-262-01716-9, 2012.

[24] Jähne, “Digital Image Processing: concepts, algorithms, and scientific applications”, Springer-Verlag Berlin Heidelberg, ISBN: 3-540-56941-3, 1991.

[25] J. W. H. Tangelder, R. C. Veltkamp, “A survey of content based 3D shape retrieval methods,” Proceedings Shape Modeling Applications, 2004., Genova, Italy, 2004, pp. 145-156., 2004.

[26] S. Zhang, “Recent progresses on real-time 3D shape measurement using digital fringe projection techniques.”, Optics and lasers in engineering, vol 48.2, pp.149-158, 2010.

[27] J. C. Russ, “The Image Processing Handbook”, CRC Press, 1995, ISBN: 0-8493-2516.

[28] S. A. Dudani, K. J. Breeding, R. B. McGhee, “Aircraft identification by moment invariants,” IEEE Transactions on Computers, vol. 26 (1), pp. 39–46, 1977.

[29] L. Gupta and M. D. Srinath, “Contour sequence moments for the classification of closed planar shapes,” Pattern Recognition, vol. 20, no. 3, pp. 267–272, 1987.

[30] C. Chang; D. J. Buehrer; S. M. Hwang, “A shape recognition scheme based on relative distances of feature points from the centroid”, Pattern Recognition, Vol: 24, Issue: 11, pp 1053-1063, 1991.

[31] R. Davies, “Machine Vision: Theory, Algorithms, Practicalities“, vol. 54, Academic Press, New York, NY, USA, 1991.

[32] P. J. van Otterloo, “A Contour-Oriented Approach to Shape Analysis“, Prentice-Hall International (UK) Ltd, New Jersey, NJ, USA, 1991.

[33] M. Sonka, V. Hlavac, and R. Boyle, “Image Processing, Analysis and Machine Vision”, Chapman and Hall Computing, Boca Raton, Fla, USA, ISBN: 978-0-412-45570-4, 1993.

[34] Q. M. Tieng and W. W. Boles, “Recognition of 2D object contours using the wavelet transform zero-crossing representation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 8, pp. 910–916, 1997.

[35] Freeman, “On the encoding of arbitrary geometric configurations,” IRE Transactions on Electronic Computers, vol. EC-10, no. 2, pp. 260–268, 1961

[36] D. Zhang, G. Lu, “Review of shape representation and description techniques”, Pattern Recognition 37, pp.1 – 19, 2004

[37] M. K. Hu, “Visual pattern recognition by moment invariant,” IRE Transactions on Information Theory, vol. 8, no. 2, pp. 179–187, 1962.

[38] A. Khotanzad and Y.H. Hong, “Invariant image recognition by zernike moments”, IEEE Trans. on Pattern Anal. and Machine Intell., 1990.

[39] R. B. Yadava, N. K. Nishchalb, A. K. Guptaa, V. K. Rastogi, “Retrieval and classification of objects using generic Fourier, Legendre moment, and wavelet Zernike moment descriptors and recognition using joint transform correlator”, Optics & Laser Technology, vol. 40 (3), pp. 517-527, 2008

[40] C.-H. Teh and R. T. Chin, “On Image Analysis by the Methods of Moments,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, pp. 496-513, 1988.

[41] P. Rosin, J. Zunic, “Handbook of Applied Algorithms: Solving Scientific, Engineering, and Practical Problems, 2D Shape Measures for Computer Vision”, Wiley-IEEE Press, ISBN: 978-0-470-04492-6, 2008.

[42] Flusser, T. Suk, B. Zitova, “Moments and Moment Invariants in Pattern Recognition”, John Wiley & Sons, ISBN: 978-0-470-69987-4, 2009.

[43] D. Zhang and G. Lu, “Shape-based image retrieval using generic Fourier descriptor,” Signal Processing: Image Communication, vol. 17, no. 10, pp. 825–848, 2002.

[44] D. Zhang, G. Lu. “Generic Fourier descriptor for shape-based image retrieval.” ICME (1). 2002.

[45] S. Li, M.-C. Lee, and C.-M. Pun. “Complex Zernike moments features for shape-based image retrieval”, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, vol 39. (1) , pp. 227-237., 2009.

[46] M. Vorobyov, “Shape classification using zernike moments”, Technical Report, iCamp-University of California Irvine, 2011.

[47] E. R. Kandel, J. H. Schwartz, T. M. Jessel, S. A. Siegelbaum, A. J. Hudspeth, (Ed.), “Principles of Neural Science”, E. R. Kandel (Ed), McGraw-Hill Companies, ISBN: 987-0-07-139011-8, 2013.

[48] M. Yagi, T. Shibata, “An Image Representation Algorithm Compatible with Neural-Associative-Processor-Based Hardware Recognition Systems,” IEEE Trans. Neural Networks, Vol. 14, No. 5, pp. 1144-1161, 2003.

[49] D. Geer, “Chip makers turn to multicore processors“, Computer, IEEE, vol 38. (5), pp 11-13. 2005.

[50] T. Roska, P. Giovanni, W. W. Chai. “Cellular wave computing via nanoscale chip architectures.“ International Journal of Circuit Theory and Applications 40.12, pp. 1187-1189. 2012.

[51] Y. Zhang, W. Zhao, J. O. Klein, W. Kang, D. Querlioz, Y. Zhang, “Spintronics for low-power computing.” Proceedings of the conference on Design, Automation & Test in Europe. European Design and Automation Association, pp. 303, 2014.

[52] U. Ingard, “Fundamentals of Waves & Oscillations”, Cambridge University Press, ISBN: 0-521-32734-2, 1988.

[53] D. Landau, E. M. Lifshitz, “Theory of the dispersion of magnetic permeability in ferromagnetic bodies“, Phys. Z. Sowietunion, 8, 153, 1935.

[54] A. Slavin: “Microwave sources: Spin-torque oscillators get in phase”, Nature Nanotechnology 4, pp. 479-480. 2009.

[55] S. E. Russek, W. H. Rippard, T. Cecil, R. Heindl, “Spin-transfer nano-oscillators“, Handbook of Nanophysics: Functional Nanomaterials, 2010.

[56] G. Bourianoff, “Towards a Bayesian processor implemented with oscillatory nanoelectronic arrays“, Keynote lecture at IEEE International Workshop on Cellular Nanoscale Networks and Applications, Turin, Italy, 2012.

[57] Gy. Csaba, Á. Papp, W. Porod, “Perspectives of using spin waves for computing and signal processing“, Physics Letters A, vol. 381 (17), 2017.

[58] A. Horváth, F. Corinto, Gy. Csaba, W. Porod, T. Roska, “Synchronization in Cellular Spin Torque Oscillator Arrays,” 13th IEEE International Workshop on Cellular Nanoscale Networks and Applications, Turin, Italy, 2012.

[59] O. Chua and T. Roska, Cellular Neural Networks and visual computing. Cambridge, UK: Cambridge University Press, 2002.

[60] T. Roska and L. O. Chua, “The CNN universal machine: an analogic array computer,” IEEE Trans. Circuits Syst. II, vol. 40, pp. 163–173, Mar. 1993.

[61] Kék, K. Karacs, and T. Roska. (2007) Cellular wave computing library, version 2.1 (templates, algorithms and programs). [Online]. Available: http://cnn-technology.itk.ppke.hu/Library v2.1b.pdf visited on 02-12-2009.

[62] K. Karacs, A. Lázár, R. Wagner, D. Bálya, T. Roska, and M. Szuhaj, “Bionic Eyeglass: an Audio Guide for Visually Impaired,” in Proc. of the First IEEE Biomedical Circuits and Systems Conference (BIOCAS 2006), London, UK, Dec. 2006, pp. 190 – 193.

[63] K. Karacs, A. Lázár, R. Wagner, B. Bálint, T. Roska, and M. Szuhaj, “Bionic Eyeglass: The First Prototype, A Personal Navigation Device for Visually Impaired,” in Proc. of First Int’l Symp. on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2008), Aalborg, Denmark, 2008

[64] K. Karacs, Á. Kusnyerik, M. Szuhaj, T. Roska, “Situation-specific Scene Interpretation in a Bionic Navigation Device for Visually Impaired, International Conference on Vision in 3D Environments, Toronto, Canada, Jul, 2009.

[65] K. Karacs, M. Radványi, Á. Kusnyerik, T. Roska, “Basic Scene Understanding and Navigation with a Bionic Camera,” International Conference on The Eye and The Auto, Detroit, MI, Sept, 2009

[66] K. Karacs, Á. Kusnyerik, M. Radványi, T. Roska, M. Szuhaj, "Towards a Mobile Visual Navigation Device", CNNA2010, 2010.

[67] Radvanyi, G. E. Pazienza, and K. Karacs, "Crosswalk Recognition through CNNs for the Bionic Camera: Manual vs. Automatic Design," in Proc. of the European Conference on Circuit Theory and Design (ECCTD'09), Antalya, Turkey, Aug, 2009.

[68] K. Karacs, M. Radványi, M. Görög, T. Roska, "A Mobile Visual Navigation Device: New algorithms for crosswalk and pictogram recognition," in Proc. of 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2009), Bratislava, Slovakia, Nov, 2009.

[69] M. Radvanyi, K.Karacs, "Navigation through Crosswalks with the Bionic Eyeglass" in Proc. of 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010), Rome, Italy, 2010.

[70] M. Radványi, K. Karacs, "Autonomous Detection of Information Patterns Through Hierarchical Peeling," in Proc. of the 21st European Conference on Circuit Theory and Design (ECCTD 2013), Dresden, Germany, Sept. 2013.

[71] M. Fritz, B. Schiele, "Towards Integration of Different Paradigms in Modeling, Representation, and Learning of Visual Categories", Object Categorization: computer and human vision perspectives, edited by S. J. Dickinson, A. Leonardis, B. Schiele, M. J. Tarr, Cambridge University Press, ISBN: 978-0-521-88738-0, 2009.

[72] E. Pekalska, R. P. W. Duin, "The Dissimilarity Representation for Pattern Recognition: Foundations and Applications", World Scientific Publishing Co. Pte., ISBN: 981-256-530-2, 2005

[73] S. Khanam, S. Jang, W. Paik, "Shape Retrieval Combining Interior and Contour Descriptors", International Conference FGCN, Jeju Island, pp.120-128, 2011.

[74] T. G. Dietterich, "An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization", Article Machine Learning, vol. 40, (2), pp 139-157, 2000.

[75] T. G. Dietterich, "Multiple Classifier Systems", Chapter Ensemble Methods in Machine Learning, Volume 1857 of the series Lecture Notes in Computer Science pp 1-15, 2000.

[76] D. Opitz, R. Maclin, “Popular Ensemble Methods: An Empirical Study”, vol. 11, pp. 169-198, 1999.

[77] M. Yang, Kidiyo Kpalma, Joseph Ronsin. “A Survey of Shape Feature Extraction Techniques”, Peng-Yeng Yin. Pattern Recognition, IN-TECH, pp.43-90, 2008.

[78] G. Carneiro, A. B. Chan, P. J. Moreno, N. Vasconcelos, “Supervised Learning of Semantic Classes forImage Annotation and Retrieval”, IEEE T. on Pattern Analysis and Machine Intelligence, 2007.

[79] B. Kotsiantis, Sotiris, I. D. Zaharakis, P. E. Pintelas. “Supervised machine learning: A review of classification techniques.” vol. 3-24, 2007.

[80] R. C. Gonzalez, R. E. Woods. “Digital Image Processing”. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, ISBN: 978-0131687288, 2001.

[81] M. Riesenhuber, T. Poggio, “Hierarchical models of object recognition in cortex,” Nature Neuroscience, vol. 2, no. 11, pp. 1019-1025, 1999.