

Pázmány Péter Katolikus Egyetem
Roska Tamás Műszaki és Természettudományi Doktori Iskola

Tér-időbeli vizuális jelenségek vizsgálata biológiai alapú algoritmusokkal

Fülöp Tamás

PhD Disszertáció Tézisei

Témavezető: Zarándy Ákos, DSc



Budapest

2018

1 Bevezetés

A világ, amely körülvesz minket tele van olyan jelenségekkel, melyek egy részét közvetlenül képesek vagyunk érzékelni, míg másokat csak közvetetten. Egy jelenség érzékelése az első lépcsője a megismerés, egészen konkrétan az észlelés folyamatának. Az érzékelés minőségén múlik, hogy milyen következtetéseket lehet levonni egy-egy jelenség megtörténtéből.

Az észlelés a biológiai rendszereknél a túlélés egyik nagyon fontos alapkövetelménye. Egy történés észlelésére adott válasz segíthet elkerülni sérülést, ütközést, támadást, vagy éppen segít a navigációban, táplálék felkutatásában. A percepció feltétlen és feltételes reflexet egyaránt eredményezhet.

Nincs ez másként az épített rendszerek esetében sem, hiszen valamilyen történést, mint inputot, a rendszer elemez, feldolgoz és rá valamilyen eredményt ad. Egy egyszerű rendszer esetén lehet egyszerűen csak leállítja a további működést (pl. ne okozzon a továbbműködés kárt), míg bonyolultabbak esetében mindenkori cél lehet, hogy a rendszer valamilyen biztonságos, kontrollált állapotban maradjon és adott esetben, ha korlátozásokkal is, de továbbra is maradjon működőképes részben, vagy egészben (pl. vasúti rendszerek, repülő berendezései). A cél tehát, hogy az észlelést követően megtörténjen egy tevékenység elindítása, folytatása, megváltoztatása vagy leállítása, valamint hibák, működési feltételek megváltozásának jelzése, vagy akár további adatok kérése.

Az észlelés mindenkor az érzékeléssel kezdődik. A biológiai rendszerekben a bemeneti ingerek különböző információkat hordoznak, melyek egy része azonnal feldolgozható jelentéssel bír. Sok olyan inger-válasz reakció van, melly a különböző biológiai rendszerek túléléshez szükséges alapvető viselkedési mintáit tartalmazza. Ezek lényegében mind feltétlen reflexek és ez az “utasításkészlet” alkalmas arra, hogy az egyedfejlődés során tapasztalt tipikus veszélyhelyzetekre, szituációkra átlagosan jó válaszokat adjon. Az észlelés fontos pontjának tekinthető, mikor az is realizálódik, hogy milyen veszélyhelyzetre történt valamilyen konkrét reakció.

Napjainkban az egyik legszámítás intenzívebb feladat a képfeldolgozási feladatok megoldása. Számos látórendszerhez köthető vizuális jelenségre ad az emberi

szervezet azonnali választ veszélyek, veszélyhelyzetek elkerülésére. Kutatásom arra irányult, hogy a tipikus képfeldolgozási láncban történő feladat megoldáshoz képest, a természetben is előforduló alapvető túlélési problémákra a látórendszerben milyen megoldás található, az mennyire hatékony, mennyire paraméterevezhető és hogyan használható fel. Egy ilyen tipikus veszélyt jelentő helyzet a közeledő objektumok érzékelése, mely a menekülésben és a navigációban egyaránt fontos szerepet játszik.

Kutatásom során azzal is foglalkoztam, hogy milyen már létező számítógépes architektúrán hogyan lehet hatékonyan implementálni egy ilyen biológia motiválta algoritmust.

2 Kutatás módszere

Kutatásom Roska Botond svájci kutatócsoportjának munkáját és eredményeit megismerve indult el 2008 végén, melyben a Celluláris Neurális Hálózatok (CNN) egyik komoly alkalmazási lehetőségét láttam meg. Először a kutatócsoport által megtalált retinacsatorna egyszerű modellezésével és a modell vizsgálatával foglalkoztam, a modellezési alapkérdéseket 2009-es diplomatervemben kezdtem el feloldozni, melyben felvázoltam már a lehetséges kutatási irányokat. 2009-ben, a Natureben megjelent cikkük adta meg a modell építésének konkrét sarokköveit [1] és a modell behatóbb vizsgálatának lehetőségét, valamint ez alapján indulhatott meg a modell validációja és kalibrációja.

A modellt elemző munka alapvető fontosságú volt, mely során először a retinacsatorna működésének mélyebb megismerése, megértése volt az elsődleges, valamint az egyes paraméterek meghatározásának kérdései. Ez alapján lehetett megkezdeni a modell validálását, valamint ehhez kapcsolódóan a modell kalibrálását, mely során vizsgáltam a modell működési tartományát és a kapcsolódó időzítési viszonyait.

A modell felépítése során rendelkezésre álló kutatási eszközök lehetőséget adtak arra, hogy a modell alkotáshoz CNN chipen történő implemetáció készüljön, így megvizsgáltam az algoritmus ezen való megvalósíthatóságát. Az algoritmust az Anafocus Eye-RIS v1.2 architektúráján készítettem el először, valamint portoltam az architektúra frissített változatára is a v1.3-ra.

A modell vizsgálatának következő kérdésköre a kalibráció tisztázása volt, a modell matematikai alátámasztásával. A modell kalibrálásával kapcsolatos megállapítások 2012-ben jelentek meg referált folyóirat publikáció keretében. Ezen publikáció már az egér retina alapú modell más biológiai alapú modellekkal való összevetését is tartalmazta.

A modell vizsgálata során több, kisebb jelentőségű hatás is felismerésre került, például a közeledési irány meghatározhatóságának kérdésköre, valamint a kalibrációs paraméterek megváltoztatásai során észlelt másodlagos hatások és ezek

elemzése, mely alapján később az impulzusüzemű fényforrások felismerésének lehetőségével is foglalkoztam.

A disszertáció 1. tézisének alapját képező publikáció:

[1] T. Münch, R. Silveira, S. Siebert, T. Viney, G. Awatramani és B. Roska, „Approach sensitivity in the retina processed by a multifunctional neural circuit,” *Nature Neuroscience, Vol 12.*, pp. 1308-1316, 2009.

3 Új tudományos eredmények

Munkám új tudományos eredményeit két fő téziscsoportba soroltam. Az első fő tézis-csoportba azon téziseim kerültek, melyek közvetlenül az egérben talált Pvalb-5 ganglionsejt csoporthoz kapcsolódó modell kutatási eredményei alapján születtek meg. A második téziscsoportba a modellvizsgálat során talált egy mellékhatás kapcsán született eredmények kerültek. A disszertációhoz kapcsolódó tudományos eredmények folyóiratban [J], könyvfejezetben [B] és referált konferencia publikációkban [C] jelentek meg.

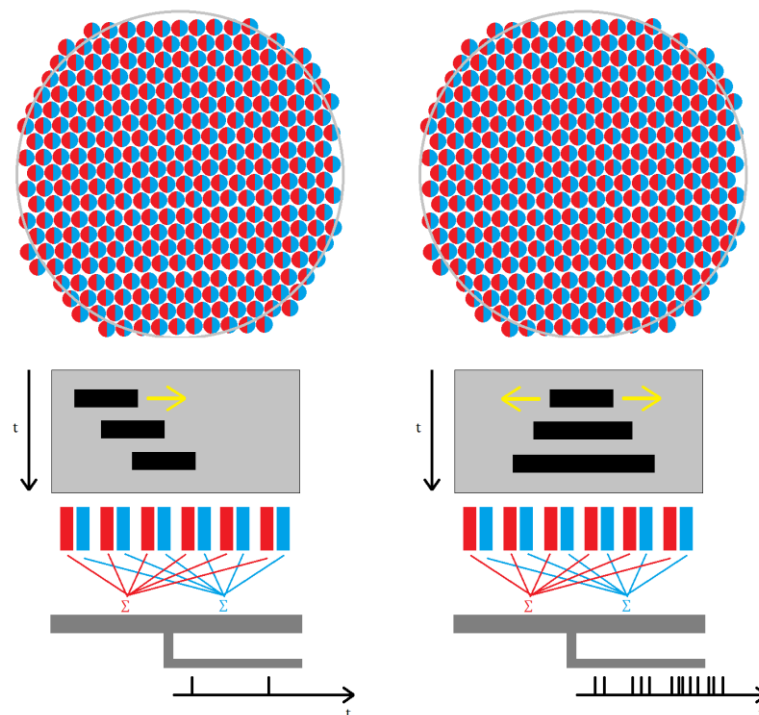
I. Téziscsoport:

1. Igazoltam, hogy az egerek retinájában felfedezett Pvalb-5 ganglionsejtek modellje képes az OFF receptív mezőn sötét tárgyak közeledését detektálni. Elemeztem a modell hatékonyságát implementálási szempontból. A modell viselkedési analízise kapcsán megmutattam, hogy az objektumok sebessége és iránya becsülhető.

Téziscsoporthoz kapcsolódó publikációm: [B], [J], [C1], [C2], [C3]

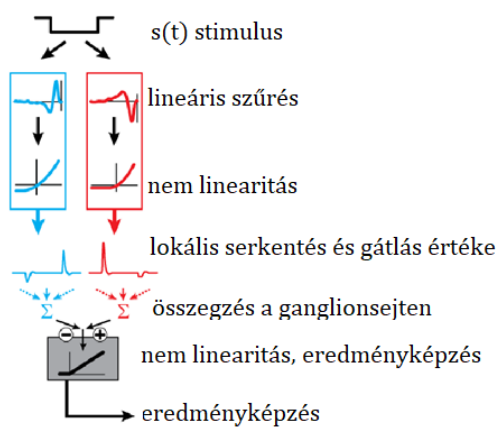
Az egér retinájában felfedezett Pvalb-5 ganglionsejt olyan OFF típusú sejt, mely érzékeny az objektumok közeledő mozgására. Egy Pvalb-5 ganglionsejt a látómező mintegy 10%-ának receptívmezőjét alkotó csap receptorsejtekből nyeri bemeneti jeleit. Egy receptív mező esetében a világos háttér előtt a retina felé közeledő sötétebb objektum sziluett növekedése nyomán előállt változás képes aktiválni a ganglionsejtet. A kísérletek bizonyították azt is, hogy a Pvalb-5 sejt ON irányban erősen gátló választ ad. Ennek jelentősége abban áll, hogy ezzel a sejt a laterális mozgásokra így teljesen érzéketlen, vagyis neuronválaszt nem ad. Az elkészített modellem ezt a működést bizonyította.

A csapok a modell szempontjából egyszerre viselkednek serkentő és gátló sejtként, függően attól, hogy OFF vagy ON típusú eseményt érzékelnek. Az egyes receptorsejteken beérkező stimulus, függően attól, hogy serkentőként, vagy gátlóként viselkedik az előtte történtekekre, a Pvalb-5 ganglionsejten összegződik, mely neurális sejtválaszt ad. A feldolgozás biológiai útját szemlélteti az 1. ábra.



1. ábra Pvalb-5 ganglionsejt által lefedett receptívmező. Bal oldalon laterális mozgás esete, jobb oldalon közeledő mozgás (sziluett növekedés esete), melynél a serkentő és a gátló válaszok a ganglionsejten összegződnek.

Az elkészült modell sematikus információ áramlását a 2. ábra mutatja be, melyen az látható, hogy valamely receptorsejten beérkező $s(t)$ időbeli stimulusra (1) szerint számolható a serkentés és (2) szerint a gátló időbeli konvolúció. A serkentés és a gátlás esetén egyaránt lineáris szűrést kell végezni a (3) és (4) szerint számolható módon. Ennek jelentősége abban van, hogy a természetben csak “pozitív” sejtválasz létezik, ez a lépés ezt készíti elő. Emiatt van szerepe a gátló csatorna hatásnak.



2. ábra Pvalb-5 ganglionsejt számítási modellje valamely beérkező $s(t)$ stimulusra.

$$e_{i,j}(t) = \sum_{n=0}^{g-1} s_{i,j}(t-n)w_n^e \quad (1)$$

$$i_{i,j}(t) = \sum_{n=0}^{g-1} s_{i,j}(t-n)w_n^i \quad (2)$$

ahol:

- $s_{i,j}$ a fotóreceptort ért fényintenzitás értéke, a képkocka i,j pontján, a modell bemenete,
- g a diszkrét képfelvételek száma, mely az időbeli konvolúcióban résztvevő képkockák száma,
- w_n^e a serkentő csatorna időbeli konvolúciójához tartozó súlytényező vektor,
- w_n^i a gátló csatorna időbeli konvolúciójához tartozó súlytényező vektor,
- $e_{i,j}(t)$ adja meg a serkentő csatorna t időpillanatban vett időbeli konvolúciójának eredményét az i,j képkockára (a szűrés serkentő kimenete),
- $i_{i,j}(t)$ adja meg a gátló csatorna t időpillanatban vett időbeli konvolúciójának eredményét az i,j képkockára (a szűrés gátló kimenete).

$$h_e(x) = \begin{cases} (x+o_e), & \text{ha } (x+o_e) > 0 \\ 0 & \text{egyébként} \end{cases} \quad (3)$$

$$h_i(x) = \begin{cases} (x+o_i), & \text{ha } (x+o_i) > 0 \\ 0 & \text{egyébként} \end{cases} \quad (4)$$

ahol:

- o_e a serkentő csatorna eltolási értéke,
- o_i a gátló csatorna eltolási értéke,
- h_e a gátló csatorna átviteli függvénye,
- h_i a serkentő csatorna átviteli függvénye.

$$g_{k,l} = r\left(\sum_{s,t \in N_r(k,l)} (h_e(e_{s,t}) - h_i(e_{s,t}))\right) \quad (5)$$

ahol:

$N_r(k,l)$ a ganglionsejthez tartozó (k,l) méretű receptívmező,

$r(x)$ a sejtválaszt adó függvény, mely $r(x) = \frac{x(\text{sign}(x)+1)}{2}$ -ként írható fel.

A kialakult lokális eredmények a Pvalb-5 sejt által lefedett receptívmező összegzi és thresholdolja, mely egyben megadja a becsült választ a közeledés tényére, melyet a (5) ír le.

A modellt megvizsgálva látható volt, hogy a közeledést vizsgáló esetekben a receptívmező változásainál pozitív előjellel vett változás történik (serkentés), ahol nincs történés, ott semleges. A laterális mozgásnál a serkentő változással egyidőben gátló változások is megjelennek.

Tézis 1.1: Megmutattam, hogy az egerek retinájában felfedezett, Pvalb-5 típusú ganglionsejtnél leírt működés párhuzamos architektúrákon megvalósítható és hatékonyan implementálható.

Tézishez kapcsolódó publikációk: [B], [J], [C1]

A modellt leíró (1)-(5) képleteknél az (1)-(4) képletek egy pixelre elvégzett számítások, melyek egyszerű szorzások és összeadások, valamint egy szignum függvény alkalmazása. Az (5) képlet a térbeli összegzést valósítja meg. Mint a képletekből látható, elemi műveletekkel egyszerűen leírható a modell és a pixel szintű műveletek hatékonyan párhuzamosíthatók SIMD architektúrára. A térbeli összegzés a pixel számának és a rendelkezésre álló processzormagok számának, valamint a processzor architektúrális felépítésének függvénye.

Tézis 1.2: Kísérletileg bizonyítottam, hogy az egerek retinájában felfedezett, Pvalb-5 típusú ganglionsejtnél leírt működés topografikus processzor architektúrára elemi CNN műveletekkel megvalósítható, és analóg processzor tömbön hatékonyan implementálható sebesség-energia-terület viszonylatban.

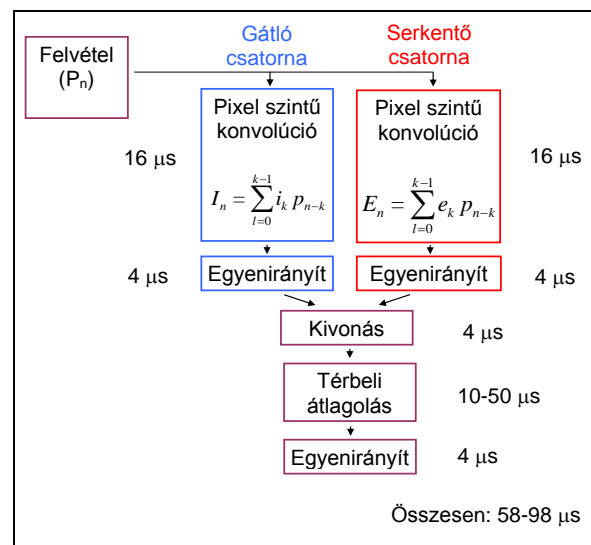
Tézishez kapcsolódó publikációk: [B], [J], [C1]

A Pvalb-5 ganglionsejt a látótérből 10° -ot fed le, ami nagyszámú csap stimulusának feldolgozását jelenti. Az egér látómezeje 120° - 130° körüli, ezt százezres nagyságrendű csap fed le, a különböző típusú csapok eloszlása a látómezőn nem egyenletes, kb. 15% egységnyi területen a sűrűségben az eltérés.

A modellezéshez az Eye-RIS v1.2 CNN-UM, analóg érzékelő számítógépre esett a választásom, mely a kutatás idején az egyik legkorszerűbb fókusz síkbeli analóg

processzortömbbel rendelkező érzékelő-számítógép. A rendszer egy Q-Eye szenzor-processzor chipet tartalmaz, mely szürkeárnyalatos felvételek készítésére akár 10000k fps sebességgel képes 176x144-es képméretben. A CNN műveletek végrehajtását alapértelmezés szerint saját FPP nevű funkcionális programozási nyelvvel utasításszinten támogatja. Egy parancs lefuttatása során kizárólag az analóg végrehajtási idővel kell számolni.

Az implementáció a ganglionsejt kimenetét az Eye-RIS-on 58-98 μ s alatt képes meghatározni a képkészítést követően. A 3. ábra bemutatja az egyes műveletek részidejeit.



3. ábra Az implementált modell folyamatábrája, az analóg rendszer számára szükséges műveletek idejeivel

A ganglionsejt lényegében egy morfológiai összegzést végez, ez a legdrágább művelet az egész folyamat során. Ez az összegzés lényegében egy diffúzió, melyet az Eye-RIS a képtől és egyéb beállításoktól függően 10-50 μ s alatt tud végrehajtani. A digitális architektúrákhoz képest itt az időben számottevő nyereséget lehet elérni. Az analóg diffúzió ideje függ a háttértől, zajoktól és a beállításoktól. Ez a diffúz összegzés valósítja meg a (6)-os képlet szerint leírtakat.

$$g_{k,l}(t) = h_g \left(\sum_{i,j \in N_r(k,l)} \left(h_e \left(\sum_{n=0}^S I_{i,j}(t - n\Delta t) w_n^e \right) - h_i \left(\sum_{n=0}^S I_{i,j}(t - n\Delta t) w_n^i \right) \right) \right) \quad (6)$$

Tézis 1.3: Megmutattam, hogy több Pvalb-5 típusú ganglionsejt a retina különböző térrészein különböző módon reagál ugyanazon gyorsan közeledő objektumra, s a legerősebben a közeledés centrális közepén reagál.

Tézishez kapcsolódó publikációk: [B], [C3]

Roska Botond kutatócsoportja bemutatta [1], hogy a ganglionsejtek által lefedett területek átfedésben vannak egymással, átlapolással fedik le a receptívmezőt. A modell módosításával átfedő hálózatot hoztam létre és megmutattam, hogy a receptívmező által látott méretét meg nem haladó látótérnövekedésben a középső receptívmezőre hat a leginkább és az átfedő részekre kevésbé, viszont a receptívmezőt kinőve a szomszédos mezők még továbbra is reagáltak, így nőtt az aktivitás mértéke.

Tézis 1.4. Karakterizáltam és kalibráltam a Pvalb-5 típusú ganglionsejt működését leíró modellt sebesség-méret-távolság relációkban

Tézishez kapcsolódó publikációk: [B], [J]

A karakterizálás [B] négy fő paraméter megadását jelenti:

1. Konvolúciós ablakok nagysága (a rendszer memóriája, reagálási képessége a közeledés sebességére)
2. A serkentő és gátló csatornák küszöbértékei (kis változások és zajok eliminálásának képessége, a méretváltozásra való érzékenység)
3. Receptív mező mérete (a közeledő objektum méretéhez, távolságához képest, a lehetséges sejtválasz nagyságát befolyásolja)
4. Ganglionsejt küszöbértékének meghatározása (a modell jelzési küszöbe, az előző paraméterek függvényében)

Az átfedő receptívmezők és méretezésük kapcsán megvizsgáltam, hogy a receptívmező mérete, a közeledő objektum mérete és sebessége, valamint távolsága, irányultsága milyen összefüggésben befolyásolja a Pvalb-5 ganglionsejt választát.

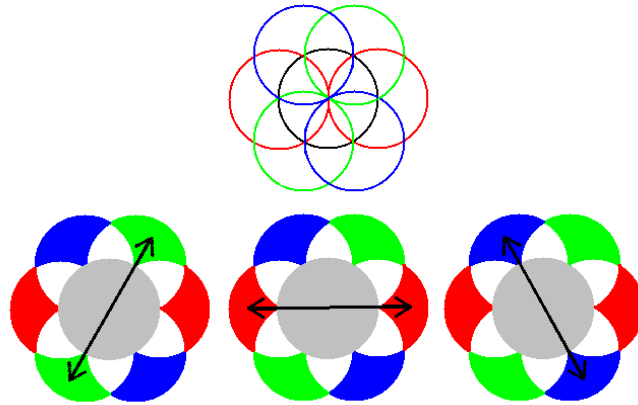
Közeledő, távolodó, valamint laterális mozgások és ezek kombinációival modellkísérleteket végeztem, valamint felírtam a modellre jellemző alapösszefüggéseket, valamint elemeztem a modell paraméterezését ezek függvényében.

A modell számítását plotterasztal segítségével validáltam, különböző sebességű és irányultságú mozgások vizsgálatán keresztül.

Tézis 1.5: Megmutattam, hogy a Pvalb-5 típusú ganglionsejt működése alapján, egy közeledő objektum iránya becsülhető, valamint a világos háttéren sötét objektumok laterális elmozdulásának irányát is képes megmutatni.

Tézishez kapcsolódó publikációk: [B], [C3]

Mivel a közeledő objektum az átfedő ganglionsejteket különböző mértékben aktíválja, így ezen aktivitási mintázat a ganglionsejt válaszána szignumfüggvényével vett küszöbölését követően, meghatározható a közeledés irányának vektora, az aktivitás mértékének függvényében, az átfedő receptormezők pozíciójához ké-



pest.

4. ábra Átfedő receptívmezők között számolt potenciálkülönbségek.

Ezen vektor lényegét tekintve egy receptívmező és annak szomszédos átfedő receptívmezőinek szuperpozíciójából adódik, mely a (7) képlettel számolható.

$$\vec{v}_e = \sum_{i \in n} a_i \cdot \vec{v}_i \quad (7)$$

ahol:

- a_i a potenciálkülönbség nagysága és előjeles iránya két szomszéd között,
- \vec{v}_i két egymásmelletti receptívmező középpontjai közötti normálvektor

Ez a 4. ábra esetében látható átfedés esetén a következő:

$$\begin{pmatrix} x_{size} \\ y_{size} \end{pmatrix}_{center} = \begin{pmatrix} (rR - rL) + \sqrt{(bR - bL) + \sqrt{(gR - gL)}} \\ \sqrt{(bR - bL) + \sqrt{(gR - gL)}} \end{pmatrix} \quad (8)$$

Ahol r : piros, g : zöld, b : kék receptívmezők értékei a középponthez képest balra (L), vagy jobbra (R).

A modellről megmutattam, hogy képes jelezni a laterális mozgásokat is.

II. Téziscsoport:

2. Megmutattam, hogy a periódikusan változó fényintenzitást kibocsájtó impulzus üzemű LED-es lámpák időbeli alul-mintavételezéssel detektálhatóak HSV színtartományal történő szegmentációt követően, mozgó- illetve álló platformos képfelvétel esetén.

Tézishez kapcsolódó publikációm: [C4]

A közeledésre érzékeny retinacsatorna vizsgálata során kiderült, hogy a periodikusan változó intenzitású LED-es fényforrások alacsony képrögzítési sebesség mellett a receptív mező modell számolása során a kimenet végeredményét befolyásolják. Ennek oka, hogy folyamatos ON-OFF impulzus éri egy adott területen a receptívmezőt és ezzel a pixel szintű kezdeti konvolúciót befolyásolja. Megfelelő színsegmentációt követően a frekvencia alapján egy jelzőlámpa felismerhető.

A szegmentáció során fontos szempont, hogy minél szűkebb, de az összes lehetséges találatot tartalmazó tartomány ki legyen jelölve. A LED-ek látható fényhullámhossza szűk tartományban van, így szegmentációjuk egyszerűbb, jelzőlámpák esetén pedig szabványok is vonatkoznak a lámpák színére és a fény erősségére. A LED-es lámpáknál az emittált fény színe konstans és csak a fényereje, valamint szaturációja változik.

Megvizsgáltam a különböző színtereket a jelzőlámpa felismerés problémaköre kapcsán, s úgy találtam, hogy a legalkalmasabban a HSV színtér alkalmazható szegmentációs feladatra. A H (Hue) értéke jól meghatározottnak tekinthető a szabványok miatt, míg a másik két paraméter függ a környezettől és a képképző eljárástól.

A LED egyenletes ütemű, impulzus szerű villogása alacsony képsebesség mellett a jelenség alulmintavételezését jelenti, melynek frekvenciameghatározása a Nyquist-Shannon mintavételezési tétel miatt nem lehetséges. Kihhasználva viszont, hogy állandó a villogás frekvenciája, így a Goertzel algoritmussal, megfelelő mintaszám esetén még alulmintavételezés mellett is következtethetünk impulzusüzemű fényforrás, így jelzőlámpa felismerésére.

Az egyszerű színszegmentációs algoritmust követően látható volt, hogy az algoritmus képes volt a frekvencia alapján a felismerési feladatra.

13 000 képkockán 5 399 helyen volt piros jelzésekép. A piros jelzéseképeket 86,76%-ban ismerte fel az eljárásom és 8% körül alakult a hibás pozitív (fp) eredmények száma ebben az egyszerű modellben. Az alábbi táblázat megadja az eredményeket:

	Felismert Pozitív	Felismert Negatív
Pozitív	4693(tp)	716(fn)
Negatív	1400(fp)	17304(tn)

tp: helyesen talált, fp: helytelen találat,

tn: helyesen nem talált, fn: nem talált meg létezőt.

Ez a piros jelzésekép esetében az alábbi eredményeket jelenti az algoritmus minőségét tekintve: Precision= 0,77 Recall= 0,87

Tézis 2.1: Adható olyan analóg CNN processzortömbön is implementálható megoldás a frekvencia számolására, mely képes impulzus üzemű fényforrások detektálására.

A Goertzel algoritmus előnyös tulajdonsága, hogy egyszerűen számolható pixel szinten, s elegendő mennyiségű analóg memória mellett:

```

realW = 2*Cos(2*pi*k/N); //FPGA művelet cos miatt
imagW = Sin(2*pi*k/N); //FPGA művelet sin miatt

LAM_1 = 0; //CNN művelet
LAM_2 = 0; //CNN művelet
LAM_0 = 0; //CNN művelet
for (n=0; n<N; ++n) //ciklus FPGA közreműködéssel
{
    LAM_0 = LAM_x(n)+(realW*LAM_1)-LAM_2; //CNN művelet
    LAM_2 = LAM_1; //CNN művelet
    LAM_1 = LAM_0; //CNN művelet
}
outreal = 0.5*realW * LAM_1 - LAM_2; //CNN művelet
outimag = imagW * LAM_1; //CNN művelet

```

A Goertzel algoritmust megvalósító kódban a ciklusban található műveletek a LAM analóg memóriacellák között értelmezett műveletek, a ciklus hívások pedig az FPGA-ról kezdeményezhetők, ahogy a szinusz és koszinusz számolás is. Az Eye-RIS v1.3 rendszeren a fenti algoritmus nem implementálható, mivel az analóg processzortömb 6 analóg memóriacellája nem elegendő a számoláshoz.

4 A képfeldolgozási lánc és biológiai vonatkozásai, a kutatás eredményeinek felhasználása

A közlekedés-informatikában legyen az vasút, autóipar, vagy éppen repülés, nagy mennyiségű adat feldolgozására van igény. Mindhárom rendszerre igaz, hogy teljesen autonóm rendszer még jelenleg nem létezik, talán legelőrébb a vasút áll, mely zárt és jól karakterizált pálya esetén (CBTC alapú metrók) képes alacsony sebesség mellett akár tízméteres nagyságrendű követési távolságot betartva is egymás után szerelvényeket közlekedtetni radar, vagy más a távolságot aktívan mérő technológia nélkül, pusztán helyadatok és sebességmérés, valamint a saját állapotának ismeretében (sebesség, megfékezhettség). A “nagyvasúton”, mely nem zárt rendszer mint a metró, valós probléma az objektumok pályára kerülésének figyelése, akkor is, ha az ütközés nem elkerülhető, mert ilyenkor egy nagyobb tárgyal való ütközés előtti figyelmeztetésre a mozdonyvezetők a mozdonytér védettebb területére tudnak menekülni.

A repülőök autonóm vezetéséhez nagymennyiségű szenzort vetnek be, melyek egy része az objektumérzékelések kapcsán aktív (radarok), de az autonóm repülés esetében (UAV) számtalan kutatásban kerül elő a kamerák felhasználásának kérdése is az objektumészlelésnél.

A vezetést támogató rendszerek kapcsán a ma viszonylag drága és kevésbé praktikus, aktív eszközök, mint a radar és LIDAR megoldások mellett az elmúlt 15 évben számtalan cég foglalkozott kamerás rendszerekre épülő megoldásokkal. Az autonóm vezetésben a közeledő objektumok felismerése kritikus és állandóan jelenlévő probléma, melyre hatékony megoldást kell adni. Jelenleg a legtöbb megoldás a közeledő objektumok érzékelését egyedül kamerára még nem meri bízni, így a most piacon lévő egyik legfejlettebbnek tekinthető beágyazott EyeQ4 chip megoldás is tartalmazza a radarral együtt való alkalmazás lehetőségét, pedig maga a chip alkalmas biztonságkritikus számítási feladatok ellátására.

A figyelmeztető rendszerek esetében fontos a hibás pozitív jelzések minimálisra szorítása, míg az autonóm rendszereknél mindemellett szükséges a nagy biztonságú felismerés is, mely csak meghatározottan tolerálható szintű hibázást tesz lehetővé.

A gyors és pontos felismeréshez mindenképpen fontos az, objektum közeledés tényét időben észlelni, hogy a figyelmeztetési vagy beavatkozási döntésre több idő juthasson.

A kutatásom óta eltelt időben egyértelmű tendencia volt a klasszikus, feladatorientált szegmentáció háttérbeszorulása és a mély tanuló algoritmusok, valamint hatékony feldolgozási hálózatok előretörése.

Ezekben az élő szervezetekhez hasonló megközelítésűek már, de mint látható, azal a kivétellel, hogy bizonyos fontos képfeldolgozási műveletek nem magas (agyi) szinten kerülnek felismerésre, hanem már alacsony szinten (retina), mint a közeledő objektum esetében.

A sáskában felfedezett, a Pvalb-5-nél jóval közismertebb és durvább felbontású sáska inspirálta LGMD1 modellel mind a mai napig többen foglalkoznak a robotikában utolsó pillanatos ütközésselkerülő megoldások kapcsán, ennek vélhető oka, hogy chip implementációja is létezik. Figyelemre méltó, hogy a precízebb felismerést támogató LGMD2 megoldás sem terjedt el a Pvalb-5-höz hasonlóan.

A Pvalb-5-öt számára tipikusan a jövőbeli alkalmazási terület lehet az érzékelő szenzor-processzortömbök további fejlődése és szélesebb körben való elterjedése, melyek a szintén biológiai környezet inspirálta ConvNetek számára is alkalmas futási környezetet jelenthetnek.

A lámpa jelzések felismerése kapcsán a frekvencia elemzés alapján történő felismerést nagysebességű felvételeken mások is kutatták a publikációmát követően. Az erre a felismerésre építő algoritmusok a felismerésben az elmúlt évben született ConvNet-es lámpa felismerési eljárásokhoz képest jobb eredményt mutatnak. Abban továbbra is egyedinek tekinthető a publikált eljárásom, hogy alulmintavételezéssel való megközelítéssel, hogy hagyományosabb képfelvévő eszközökkel és számítási apparátussal oldható meg vele a fényforrások felismerési feladata, ez pedig a számítási feladatokban is előnyösebbé teszi. A Goertzel algoritmus mindemellett alkalmas analóg CNN eszközön implementálható.

Eredményeim rámutatnak arra is, hogy a természetben számos problémára van megoldás, így utat kíván mutatni annak is, hogy a természettől mind több el lehet lesni és megvalósítani.

5 Publikációs lista

Könyvfejezet

[B] Á. Zarándy és T. Fülöp, „Implementation and Validation of a Looming Object Detector Model Derived from Mammalian Retinal Circuit,” in *Focal-Plane Sensor-Processor Chips*, Springer, 2011, pp. 245-260.

Folyóirat

[J] Á. Zarándy és T. Fülöp, „Approaching object detector mouse retina circuit model analysis and implementation on cellular sensor-processor array,” *INTERNATIONAL JOURNAL OF CIRCUIT THEORY AND APPLICATIONS*, Vol. 40., pp. 1249-1264., 2012.

Konferencia

[C1] T. Fülöp és Á. Zarándy, „Bio-inspired looming object detector algorithm on the Eye-RIS focal plane-processor system,” in *12th international workshop on cellular nanoscale networks and their applications - CNNA*, Berkeley, 2010.

[C2] Á. Zarándy és T. Fülöp, „Retinal approaching object detector model implementation and validation,” in *IEEE international symposium on circuits and systems - ISCAS*, Paris, 2010.

[C3] T. Fülöp és Á. Zarándy, „Bio-inspired looming direction detection method,” in *13th International Workshop on Cellular Nanoscale Networks and their Applications - CNNA*, Turin, 2012.

[C4] T. Fülöp és Á. Zarándy, „Chromatic Spatial-temporal Frequency Analysis for Traffic Light detection,” in *4th IEEE International Conference on Cognitive Infocommunications - CogInfoCom*, Budapest, 2013.

[C5] T. Fülöp és Á. Zarándy, „Real-time moving object segmentation algorithm implemented on the Eye-RIS focal plane sensor-processor system,” in *International symposium on nonlinear theory and its applications - NOLTA*, Budapest, 2008.

6 Köszönet

Sok türelmet és támogatást köszönhetek sokaknak. Azért, hogy eljutottam, eljuthattam idáig. Öröm, hogy ennyi jó ember vesz körül.

Ákosnak köszönöm a sok iránymutatást, javaslatot, szakmai és személyes beszélgetést, biztatást, munkáim átnézését, vagy éppen angol írásaim végigszenvedését, amivel 2007 óta feltétlenül támogatott. Közös munkánk kapcsán nem tudom nem hivatkozni *Lukács 15,11–32-t*. Köszönöm mindenkori és feltétlen támogatásodat.

Ákos mellett Vető István Tanár úr az, akinek nagyon sokat köszönhetek mind szakmai, mind emberi oldalról egyaránt, a sok egyetem környékén töltött év alatt. Értékes beszélgetéseink nagyon hiányoznak. Embersége, szakmaisága, mérnökséghez való hozzáállása és mindenkor bátorító, jóindulatú és alázatos hozzáállása követendő és továbbadandó példaként áll számomra.

Roska Tamás professzor úrnak köszönhetem, hogy 2007-ben a SZTAKI-ban, majd 2009-ben már a doktori iskolában lehetőséget adott, hogy megismerjem a kutatói életutat, megtudjam, hogy mit jelent az elmélyült kutatás az alázatos munka, lelkesen hozzáállni egy-egy feladathoz. A disszertáció megírásakor ezt újra átélhettem, s ez most vissza hozta a kedvet, hogy ne szakadjak el teljesen a továbbiakban sem az egyetemi és kutatói körtől, s lássak benne perspektívát a valamikori jövőben.

Köszönöm Szolgay Péter és Vida Tivadarné folyamatos támogatását, mindenkori jelenlétét és rám szánt figyelmét, intó szavaikat. Ez tette lehetővé, hogy megembereljem magam, és végül a befejezés mellett döntsek.

Köszönöm, hogy Vágó Zsuzsanna megbuktatott Analízis II-ből a harmadik vizsgán is. “Én úgy látom, hogy maga ezt nem érti.” szavai máig fülemben csengenek. Akkor és ott helyén való volt: emlékeztetett újra Apum és általános iskolai kémia tanárom, Kaposiné Héger Ágnes szavaira, akik mindig arra tanítottak, hogy értsük meg a dolgokat, a világot, az összefüggéseket.

A Kalmárból Pataky Margit tanárnőnek köszönöm azokat a stabil és mély számítástechnikai szakmai alapokat, melyekkel megalapozta szakmai hozzáállásom.

Szombatfalvy Csillának, hogy megtanított írni, a komolyzene megszerettetésével együtt: Az orosz ötöktől, Esterházy Pál Harmonia caelestisén át Krzysztof Pendereckiig.

Köszönöm a sok év közös oktatói munkát, bizalmat és biztatást Bércesné Novák Ágnes tanárnőnek, valamint Lukács Gergelynek, akinél TDK-mal a kutatói utat először kipróbálhattam. Oláh Andrásnak is köszönetet mondok, akivel az egyetemi szakmai versenyélet és utánpótlásigazgatás elindításában dolgozhattam együtt. Köszönettel tartozom Tihanyi Attilának és Péri Mártonnak, akik álltak rendelkezésemre bármikor, amikor szakterületükről segítséget kértem.

Hálás vagyok mindenkori egyetemi kollégáimnak. Különösen, akik részt vettek és mindig társaim voltak doktoranduszként a mindennapokban, gyakorlatok tartásánál, akikkel beszélhettem szakmai és társadalmi kérdésekről, nehézségekről, vagy épp eredményekről, meglátásaimról: Radványi Mihály, Koller Miklós, Tornai Kálmán, Füredi László, Stubendek Attila, Tornai Gábor, Horváth András, Kiss András, Tisza Dávid, Tóth Emília, Kozák László. Kiemelném Zsedrovits Tamást, akinek számtalan közös gyakorlatvezetést, beszélgetést és angol lektorálást köszönhetek az elmúlt bő tíz évben.

Köszönöm, hogy belekóstolhattam az ipari kutatásba is a FETI-nél, a közös munkát, különösen Kováts Antalnak, Soós Balázs Gergőnek és Aoyagi-sannak.

Köszönöm jelenlegi munkahelyem, a Prolan Irányítástechnikai Zrt. támogatását, kollégáim türelmét és érdeklődését, hogy időt adtak arra, hogy mindennapi feladataim mellett nyugodtan tudjam megírni disszertációm is.

Külön köszönöm mindazok támogatását, akik végig mellettem voltak, biztattak, érdeklődtek. Legelőször és legjobban Orsinak, akivel életem akkor fonódott össze, amikor neki kezdtem ennek a kalandnak és akkor lett igazán egy, mikor épp a befejezéshez készülök már. Nem kevés türelmét köszönöm.

Szüleimnek. Mindhárom testvéremnek. Keresztszüleimnek.

Legjobb barátaimnak is köszönöm türelmüket és támogatásukat: Balázsnak, Csongornak, Ferinek, Gergőnek. Közeli barátaimnak is a folyamatos érdeklődést, biztatást: Bálintnak, Dávidnak, Eriknek, Fenyőfalvi Tamásnak, KV-nak, Bálint Zolinak, Mandácskó Zolinak és sógoromnak, Ciánnak.

A MAVÉ-s és EESTEC-es barátaimnak, külön kiemelve Franciskát, Ticz Ádámot, Krisztát, Krisztiánt, Renit, Csabit, akik mindannyian többször/sokszor kérdezték találkozónk alkalmával, hogy hogy állok velem.

A PhD disszertációt az Európai Unió támogatásával valósítottam meg, felhasználva az Európai Szociális Alap támogatását (EFOP -3.6.3-VEKOP-16-2017-00002)