

Szövegfelismerés, események észlelése és a celluláris hullámszámítógép-architektúrák néhány elméleti aspektusa

Celluláris hullámalgoritmusok szemantikus beágyazással és az
invertálhatóság vizsgálata a celluláris bináris dinamikában egy új
gráfrepresentáció segítségével

Ph.D. disszertáció tézisei

Karacs Kristóf

Témavezető:

Dr. Roska Tamás
az MTA rendes tagja

Konzulens:

Dr. Prószéky Gábor



Pázmány Péter Katolikus Egyetem
Információs Technológiai Kar
Interdiszciplináris Műszaki
Tudományok Doktori Iskola
(PPKE-ITK)



Magyar Tudományos Akadémia
Számítástechnikai és Automatizálási
Kutatóintézet
Analogikai és Neurális Számítások
Kutatólaboratórium
(MTA-SZTAKI)

„KérjeteK és adatik nékteK; keressetek és találtoK; zörgessetek és megnyittatik nékteK.”

Mt 7:7.

Tartalomjegyzék

Bevezetés, célkitűzések.....	4
Alkalmazott módszerek.....	6
Új tudományos eredmények.....	7
I. Tézis Analogikai algoritmusok térbeli szemantikus beágyazással kézzel írt szöveg felismerésére	7
II. Tézis Dinamikus események észlelése és megadott kétdimenziós mintázatok detektálása szakkadikus és zajos vizuális folyamatokban mozgó, vak platformon	10
III. Tézis Egydimenziós celluláris hullámszámítógépek invertálhatósága.....	13
Az eredmények alkalmazási területei.....	15
Köszönetnyilvánítás	16
Publikációk.....	18
Hivatkozások.....	20

Bevezetés, célkitűzések

Minden látó ember kisgyermekkorától természetesnek veszi, hogy képes felismerni a környezetünkben előforduló tárgyakat és mintázatokat. Ez olyannyira természetes mindannyiunk számára, hogy alig vesszük észre, mennyire összetett is agyunk vizuális mintaillesztő képessége. Az elmúlt időszakban a technológiai fejlődésre épülve az emberi látás kutatása egyre nagyobb figyelmet nyert a biológusok körében, de még mindig sok mindent meg kell értenünk azzal kapcsolatban, hogy mi zajlik le elménkben, amikor embereket és tárgyakat látva képesek vagyunk felismerni azokat. Ezzel párhuzamosan a mérnökök is megpróbálták modellezni a vizuális felismerési problémákat, azonban roppant nagy nehézségekkel kellett szembenéznük, és csak igen szűk alkalmazási körben születtek megoldások, amelyek általában nagy teljesítményfelvétellel járnak. Alapvetően két fő problémát különböztethetünk meg ezekkel a mintaillesztési feladatokkal kapcsolatban. Mindkettő ahhoz kapcsolódik, hogy még a legegyszerűbb és leghétköznapiabb feladatok között is nehezen találhatunk olyat, amely során az emberi agy ne építene hatalmas és rendkívül asszociatív szemantikus tudáshálójára. Az egyik probléma az, hogy a jelenlegi technológiai feltételek mellett nincsenek olyan modellek, amelyek ezeknek a hálózatoknak a működését minden szempontból megfelelően írják le. A másik fő nehézség abban rejti, hogy számos tudásszintet különböztethetünk meg ezekben a hálózatokban, amelyeket az agy teljesen párhuzamosan vesz igénybe. Olyannyira, hogy az érzékelés és a feldolgozás folyamata nem is igazán szétválasztható. Ez a párhuzamos modell nagyon távol áll azoktól a hagyományos szekvenciális modellektől, amelyek a különböző szintű információkat egymás után következő, egymástól független lépésekben dolgozzák fel, és némi visszacsatolást tartalmaznak az alapmodellből fakadó problémák kezelésére és a stabilitás biztosítására.

Amikor a kutatásaimat kezdtem meg voltam győződve róla, hogy a plaszticitást és a háttérinformációk párhuzamos kiértékelését be kell építenünk a modelljeinkbe. A felismerési feladatoknál a beérkező információ mennyisége hatalmas, amellyel csak úgy tudunk megbirkózni, ha megfelelő jellemzőket nyerünk ki. Ha a háttérinformációkat csak a jellemzők kinyerését követően használjuk fel, akkor számos olyan részletet túl hamar eldobunk, amely lehetővé teszi, hogy felismerjük a megfigyelt célt. Kutatásaim során két

érdekes és összetett mintaillesztési problémával foglalkoztam: Kézzel, folyóírással írt szövegek offline felismerésével valamint tömegközlekedési járművek járatszámának megtalálásával és felismerésével. Mindkét feladat mély szemantikai összefüggésekbe van beágyazva.

A kézírásfelismerés az emberi olvasási folyamat gépi megfelelője. Céлом az volt, hogy egy olyan rendszert tervezek, amely az emberi olvasást utánozva lehetővé teszi a nyelvi tudás azonnali alkalmazását már a betű- és a szószintű jellemzőknél anélkül, hogy a betűket vagy a szavakat előbb fel kellene ismernünk.

A táblák megtalálását és felismerését megvalósító algoritmusokat a „Bionikus Szemüveg” projekt motiválta. Ennek keretében egy hordozható eszköz elkészítését tűztük ki célul, amely segítségül szolgálhat vak, illetve látássérült embereknek olyan mindennapi élethelyzetekben, ahol a látás nélkülözhetetlen. Ezen funkciók megvalósításához a bekövetkező események környezetének alapos megértése szükséges, hasonlóan ahhoz, ahogy egy betűsorozat elolvasása a nyelvi környezetbe ágyazódik.

A buszok és villamosok azonosítása az egyik legfontosabb azon tucatnyi funkció közül, amelyeket a lehetséges felhasználók képviselői gyűjtöttek össze. Kifejlesztettem egy szemantikus keretrendszert a különböző érzékszervi információk feldolgozására. A szemantikus beágyazási elvet alkalmazva új algoritmusokat fejlesztettem ki úgy, hogy a tipikus táblák és kijelzők szemantikus leírását celluláris hullámszámítógépeken futtatható eljárásokká transzformáltam.

A bináris celluláris nemlineáris hálózatok – más néven sejtautomaták – elméletével szintén foglalkoztam. Megadtam egy eljárást, amely lehetővé teszi egy egydimenziós, egyszomszédságú sejtautomata összes pályájának vizsgálatát arra nézve, hogy létezik-e olyan pont, ahol trajektóriák egybeolvadnak, és bebizonyítottam, hogy ha nem létezik ilyen pont, akkor „Édenkert”¹ állapot sem létezik, illetve az összes pálya „Édensziget”². Ez a téma a lokálisan összekötött architektúrák területén belül az elméleti témakörökhöz tartozik.

¹ Egy sejtautomata egy állapotát akkor és csak akkor hívjuk Édenkertnek, ha nincs olyan másik állapot, amely erre képződik le, azaz a globális leképezés alatt nincs előképe.

² Egy sejtautomata egy állapotát akkor és csak akkor hívjuk Édenszigetnek, ha kizárólag saját magát vetíti az n -szer iterált globális leképezés önmagára.

Alkalmazott módszerek

Kutatásaim jelentős része közvetlenül kapcsolódik a Celluláris Nemlineáris/Neurális Hálózatokhoz (CNN), a CNN Univerzális Géphez (CNN-UM) és a Celluláris Hullámszámítás paradigmájához. A kifejlesztett CNN template-eket az irodalomban publikált analitikus módszerek ([18],[19]) segítségével terveztem. A celluláris hullámalgoritmusok tervezésénél kihasználtam, hogy a CNN-UM egy Univerzális Gép a Folyamokon (UMF) [20].

A saját magam által tervezett template-eken kívül számos standard template-osztályt használtam [12], többek között morfológiai és egyéb képfeldolgozási művelteket megvalósító template-eket. Egy fontos módszer, amit kifejlesztettem, a szemantikus beágyazás. Ezt térbeli és multimodális téridőbeli észlelési (detektálási) feladatokban alkalmaztam topografikus jellemzők megtalálására oly módon, hogy megadtam a bementi kép- és videófolyam-minták szemantikus leírását strukturális képanalízis segítségével. A jellemzőket detektáló algoritmusokat standard és saját készítésű tesztalmazokon ellenőriztem.

A kézzel írt szövegek beolvasásához hagyományos síkágvas lapolvasókat használtam. A kétdimenziós videófolyamokon végzett kísérletekhez használt felvételek a Vakok által Rögzített Videófolyam Adatbázisból származnak, amelyeket kereskedelmi forgalomban kapható, beépített kamerával rendelkező mobiltelefonokkal illetve kompakt digitális fényképezőgépekkel rögzítettek vak személyek valóságos élethelyzetekben.

Algoritmusfejlesztéshez az Aladdin nevű programcsomagot használtam az Analogic Computers Kft-től [21], az ACE-16k celluláris vizuális mikroprocesszort [22] és a Matlab programozási környezetet [23] a MatCNN toolbox-szal együtt [53].

Kifejezésszintű nyelvmodellezésre szókettősöket használtam, mint egyszerű, nem-szemantikus modellt, és statisztikai módszerekre hagyatkoztam a valószínűségi becsléseknél.

A sejtautomaták területéhez kapcsolódóan a nemlineáris dinamika elméletének, a gráfelméletnek, a szimbolikus dinamikának, és különösen L. O. Chuának a bináris Celluláris Nemlineáris Hálózatok kvalitatív elméletében elért új eredményeire [50] támaszkodtam.

Új tudományos eredmények

1. Tézis: Analogikai algoritmusok térbeli szemantikus beágyazással, kézzel írt szöveg felismerésére

A kézírás stílusok nagy változatossága a kézzel írt szövegek általános felismerését igen nehézé teszi. Létrehoztam egy kézírás-felismerő rendszert, amely az emberi olvasási folyamatot veszi mintául azáltal, hogy celluláris hullámalgoritmusok felhasználásával modelleztem az észlelési folyamatot és integráltam nyelvi tudás felhasználását a felismerési folyamatba. A rendszer lexikon-szűkítést valósít meg nagyon magas szűkítési aránnyal (>99.9%) és 80% feletti fedéssel, ami összemérhető a az irodalomban szereplő korábbi eredményekkel.

1.1. Kifejlesztettem egy alakkódolásnak elnevezett eljárást kétdimenziós morfológiai alakjellemzők felismerésének szemantikus környezetbe való beágyazására

A nyelvi információ felhasználása jelentősen javíthatja a felismerő rendszerek teljesítményét, de a hagyományos módon történő utófeldolgozás révén nem érhető el kielégítő eredmény. Az alakkódolás lehetővé teszi a nyelvi tudás beágyazását a felismerési folyamatba anélkül, hogy a betűket ténylegesen fel kellene ismerni.

E megközelítés előnyei kétrétűek. Egyrésztől megkerüli azt a problémát, hogy a betűk szétválasztása és felismerése kölcsönösen függ egymástól (Sayre-paradoxon [31]). Másrésztől a nyelvi tudás befolyásolhatja a geometriai jellemzők releváns voltát és fontosságát.

1.2. Meghatároztam hat holisztikus jellemzőt, amelyek CNN algoritmusokkal detektálhatók, és implementáltam számos jellemző-kinyerési és jellemző-osztályozási analogikai algoritmust

A holisztikus jellemzőket a betűk primitív részei alkotják, és a szóképeken detektálom őket anélkül, hogy azt betűkre kellene bontani. A hat jellemző a következő: lyukak, felnyúlók, lenyúlók, kereszteződési pontok, hegyek és völgyek. Néhány jellemzőt több osztályba sorolunk méretük, alakjuk illetve elhelyezkedésük alapján. Az 1. ábrán néhány tipikus jellemző-térkép látható.



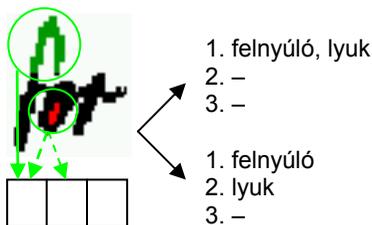
1. ábra (a) Lyukak illetve (b) felnyúlók és lenyúlók különböző osztályai

I.3. Az írásképre általánosan jellemző paramétereket határoztam meg, és eljárásokat adtam azok adaptív kiszámítására

Az azonosított paraméterek közé tartozik az írás alapvonalainak távolsága, a küszöbérték a felnyúlók és a lenyúlók minimális magasságára, méretintervallumok a lyukak osztályozásához és az átlagos betűszélesség. A kifejlesztett eljárások a saját kísérleteimen alapulnak és celluláris hullámalgoritmussal valósítottam meg őket.

I.4. Megalkottam egy eljárást a megtalált topografikus jellemzőknek a lehetséges szószintű betűpozíciókhoz való rendelésére

Egy topografikus jellemző a szón belül több betűpozícióhoz is tartozhat, így a leképezés nem egyértelmű. A 2. ábra egy egyszerű leképezése esetet mutat.



2. ábra A topografikus jellemzők leképezése nem egyértelmű, az 'o' betűben található lyuk két különböző betűpozícióhoz is tartozhat, megduplázva ezáltal a lehetséges leképezések számát.

Az eljárás megadja az összes lehetséges leképezést, amely megfelel a következő geometriai feltételeknek: az írás dőlésszöge, a jellemzők vízszintes sorrendisége és függőleges összekapcsolódása.

1.5. Megalkottam egy eljárást egy kézzel írt szóképen található karakterek számának becslésére a vízszintesen kapcsolódó komponenseket mint kvázi-független mértéket használva

Kísérletileg igazoltam, hogy egy folyóírással írt szóképen a vízszintesen kapcsolódó komponensek száma és a szóban található betűk száma erős korrelációt mutat, és a hányadosuk független az írás stílusától. Az új eljárás pontosabban meg tudja becsülni a betűk számát egy szóban, mint egy a szó pixeleiben mért szélességére alapuló becslés, amellet, hogy a bemenetét egyetlen celluláris hullámutasítással, a **HCCD** template-tel ki lehet számolni.

1.6. Kifejlesztettem egy új eljárást a felső alapvonal megtalálására, amely akkor is megbízható eredményt ad, ha a szóban található felnyúlók száma magas

A felső alapvonal minél pontosabb azonosítása nagyon fontos ahhoz, hogy a jellemzőket is pontosan lehessen detektálni. Celluláris hullámalgoritmust adtam a felső alapvonal megtalálására, amely a kapcsolódó komponenseket használva egy vízszintes alapvonalat határoz meg egy dőléskorrigált szóképen. A számított felső alapvonalat egy példán a 3. ábra mutatja. Megmutattam, hogy a jellemzőkinyerés pontossága nem csökken ha vízszintes felső alapvonalat használunk, és hogy a vízszintes alapvonalat megbízhatóbban lehet számítani, mint a két szabad paraméterrel rendelkező alapvonalat.



3. ábra A számított felső alapvonal egy példaként választott szóképen

II. Tézis: Dinamikus események észlelése és megadott kétdimenziós mintázatok detektálása egy mozgó, vak platformon, szakkadikus és zajos vizuális folyamatokban

Egy mozgó kamerával valóságos élethelyzetekben rögzített videófolyamokban táblákat megtalálni és elolvasni rendkívül nehéz feladat, amelyet az motivál, ahogy az emberek képesek ezt megtenni és ez alapján tájékozódni. Algoritmusokat adtam táblák megtalálására néhány megadott helyzetben, amelyek leírhatók egy előre definiált szemantikával, és a táblákon található számok felismerésére. Az algoritmusokat a Vakok által Rögzített Videófolyam Adatbázisból származó felvételeken teszteltem és ellenőriztem, amely több mint egy órányi felvételt tartalmaz száznál több dinamikus eseménnyel.

II.1. Definiáltam egy általános szemantikus keretrendszert a különböző érzékszervi információk beágyazott szemantikával történő hierarchikus feldolgozására

Megalkottam egy keretrendszert, amely lehetővé teszi a különböző modalitású, topografikus érzékelőktől származó információk hierarchikus feldolgozását és az érzékelés-feldolgozás folyamatát irányító szemantika integrálását. A hierarchiában a következő absztrakciós szinteket definiáltam: jellemzők, szimptomák, események, helyzetek. Az események az Eseménytárban rögzítődnek, amely – az aktív szituációkkal és az előzetes célokkal együtt – befolyásolja a figyelmi mechanizmust. A Figyelemirányító vezérli az aktuátorokat, amelyekeken keresztül a rendszer kapcsolatot tart a környezetével, és az érzékelőket azért, hogy elősegítse a jobb érzékelést az aktuális szituációk és események szempontjából.

II.2. Analogikai algoritmusokat adtam tömegközlekedési járművek járatszámot tartalmazó tábláinak megtalálására kétdimenziós videófolyamban, és a lehetséges táblajelöltek ellenőrzésére

A mobil eszközök által rögzített videófolyamok minősége alacsony, és a felhasználónak megfelelően kell tartania a kamerát, hogy a tábla felé nézzen. Az általam készített algoritmus

azonosítja a lehetséges táblajelöltek helyét, és az alapján ellenőrzi őket, hogy található-e rajtuk nagy méretű szám. A detekciós algoritmus UMF diagramja a 6. ábrán látható. A kiszámított táblapozíciót egy példán a 4. ábra mutatja. A tényleges (validált) táblapozíciókat a rendszer követi a képkockákon, és a belőlük kinyert táblaképeket egymásra szuperponálva egy jobb minőségű táblaképet készít, amely jobb felismerést tesz lehetővé.



(a)

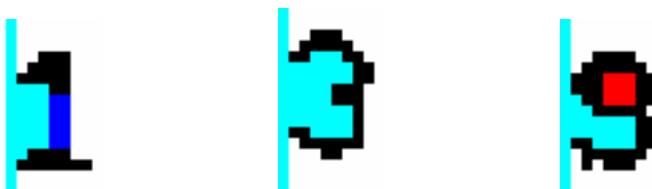


(b)

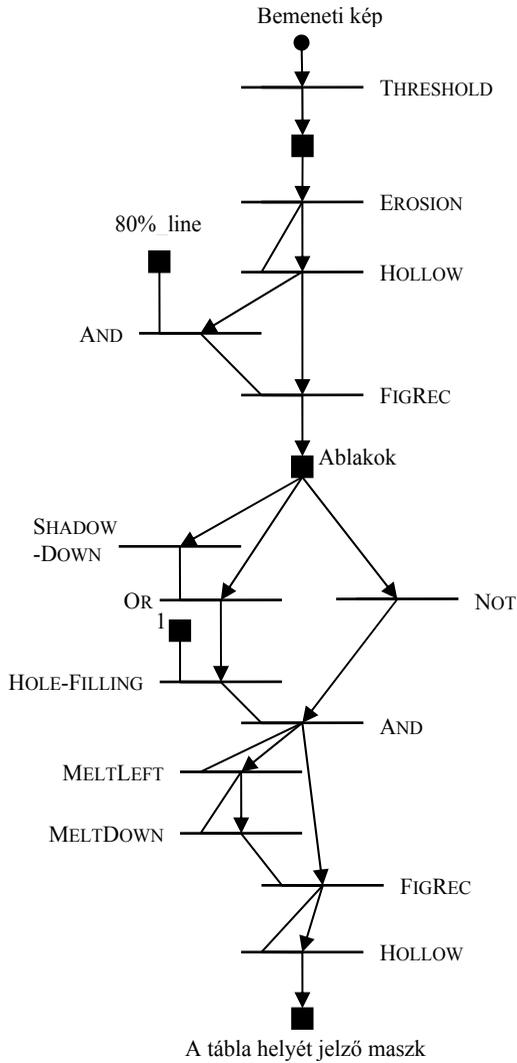
4. ábra A tábla helyzete egy villamoson. (a) Az eredeti képkocka
(b) A tábla helye a bináris képen

II.3. Eljárást adtam a táblákon megtalált számok felismerésére

A felismerés topografikus jellemzőkön alapul, amelyeket celluláris hullámalgorithmusokkal detektálok. Lyukakat és egyenes vonalakat használok jellemzőként, valamint egy doboz-modellt alkalmazok a számjegyek oldalán található félig nyitott lyukak megtalálására. Az 5. ábra néhány példát mutat topografikus jellemző-térképekre.



5. ábra Néhány jellemző-térkép. A jobbról nyitott lyukakat világoskék, a középső függőleges vonalakat sötétkék, a felső kerek lyukakat pedig piros szín jelzi.



6. ábra A fehér háttérű táblákat megtaláló algoritmus UMF diagramja

III. Tézis: Egydimenziós celluláris hullámszámító-gépek invertálhatósága

A dinamikai rendszerek invertálhatóságának vizsgálata nagy jelentőséggel bír. A L. O. Chua által bevezetett „Édensziget” fogalmára építve kifejlesztettem egy új típusú gráfot, amellyel felfedezhető, ha egy bináris Celluláris Nemlineáris Hálózat (sejtautomata) trajektóriái egybeolvadnak. Az *Édensziget* egy olyan pálya, amelynek minden állapotának kizárólag önmaga az előképe az n -szer iterált globális leképezés alatt. Bebizonyítottam egy általános tételt, amely az irányított gráf köreinek egy jól definiált halmazát az *Édenszigetek*hez köti, és mind a 256 elemi sejtautomatát megvizsgáltam nem-triviális invertálhatóság szempontjából.

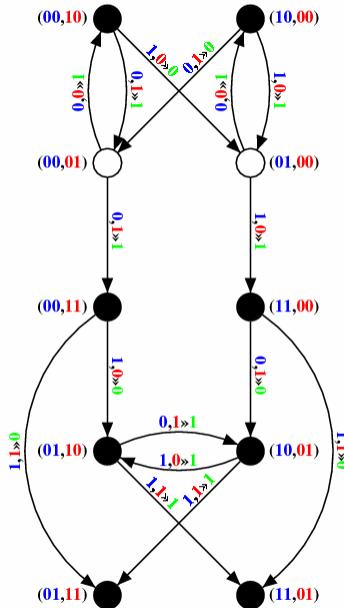
III.1. Definiáltam egy Édensziget gráfnak keresztelt irányított gráfot elemi sejtautomaták bemeneti mintázatpárjainak vizsgálatához, és konstruktív algoritmust adtam, amely bármely elemi sejtautomatának megadja az Édensziget gráffját. Bebizonyítottam egy általános tételt, amely kimondja, hogy akkor és csak akkor nincs L hosszú degenerált kör egy adott sejtautomata Édensziget gráffjában, ha minden pálya Édensziget, azaz pontosan akkor, amikor a sejtautomata invertálható.

Az *Édensziget* gráf kétbites bináris mintázatpárok feletti de Bruijn gráf, és az L hosszú körei két olyan L hosszú bináris karakterláncnak felel meg, amelyeket a sejtautomata ugyanarra a kimenetre képez le. Ha a mintázatok megegyeznek, akkor a kört degeneráltnak hívjuk. A nem degenerált köröknek különböző karakterláncok felelnek meg, így ezek a trajektóriának egy egybeolvadási pontját határozzák meg. A tétel bizonyításához néhány segédtelet alkalmazásával bebizonyítottam, hogy egy adott sejtautomata és adott hosszúságú mintázat estén a következő állítások ekvivalensek:

1. Minden kör degenerált
2. A globális leképezés szürjektív (nincs Édenkert állapot)
3. A globális leképezés injektív (nincs egybeolvadási pontja a trajektóriának)
4. Minden pálya *Édensziget*
5. A sejtautomata invertálható

III.2. Meghatároztam minden olyan elemi sejtautomatát, ami végtelen sok mintázat-hosszúságra invertálható, és bebizonyítottam, hogy a nem invertálható mintázat-hosszúságok periodikusan fordulnak elő

Numerikus szimulációkat végeztem az elmei sejtautomaták terén, és megvizsgáltam a globális leképezés szürjektivitását minden 20-nál kisebb mintázat-hosszúságra. Az eredmények alapján meghatároztam azokat a sejtautomatákat, amelyek végtelen sok mintázat hosszra invertálhatóak lehetnek. Megvizsgáltam ezen automaták Édensziget gráfjait, és bebizonyítottam, hogy azok a mintázat-hosszúságok, amelyekre nem invertálható periodikusan fordulnak elő (I. táblázat). A 7. ábrán a 45-ös szabály Édensziget gráfja látható.



7. ábra A 45-összabály Édensziget gráfja. Minden kör páros hosszú benne, mivel a köröket tartalmazó részgráfok mind párosak.

I. táblázat Azon sejtautomata szabályok, amelyek végtelen sok mintázat-hosszúságra invertálhatóak

A nem invertálható állapotok előfordulási periódusa	Szabály száma
2	45, 75, 89, 101 154, 166, 180, 210
3	105 150
∞	15, 85 51 170, 240 204

Az eredmények alkalmazási területei

Kutatásaim nagy részét közvetlenül az eredmények felhasználhatósága ösztönözte. A kézírás-felismerés területén általam kifejlesztett eljárások előnye, hogy általános célú algoritmusok, és nem korlátozódnak kisméretű szótárakra. A lehetséges alkalmazások körébe tartozik a személyes feljegyzések és a hivatalos formanyomtatványok feldolgozása is.

A szemantikus beágyazási keretrendszer széles körben felhasználható, bármilyen multimodális és multiszenzoros információ-feldolgozási feladat során. Ehhez természetesen az általános keretrendszert az adott probléma sajátosságaihoz kell igazítani, és tovább kell finomítani.

A járatszám-azonosító és -felismerő algoritmusok hordozható eszközökben alkalmazhatóak, mint például a "Bionikus Szemüveg", hogy busz- és villamosmegállókban segítsék a látássérülteket a megfelelő közlekedési eszköz kiválasztásában. Egyben alapul szolgálnak egy általánosabb, ún. Mobil Vaknavigációs keretrendszer kidolgozásához, amely képes tárgyak magas szintű azonosítására és felismerésére. Ez a keretrendszer nemcsak látássérültek által használható, hanem veszélyes környezetben való működésre kifejlesztett autonóm robotok automatikus navigációjának koordinálásához is hasznos lehet.

Az elemi sejtautomata invertálhatóságára vonatkozó eredmények az e területen folytatott további elméleti kutatások során

használhatóak, és alapjául szolgálhatnak a kvantumszámítógépet felhasználó új számítási modellek valamint a fizika és kiszámíthatóság kapcsolatával foglalkozó általánosabb kutatásoknak is. Az eljárás felhasználható továbbá az életben előforduló, könnyen előállítható, de nehezen gépi úton nehezen felismerhető valós minták felderítésére vonatkozó kutatások során is.

Köszönetnyilvánítás

Szeretnék köszönetet mondani témavezetőmnek, *Roska Tamás* *Professzor Úrnak*, hogy folyamatosan irányt mutatott számomra az évek során, köszönöm szüntelen lelkesedését, amellyel mindig a kiválóságra törekedett és engem is erre ösztönzött, és hogy minden törekvésemet támogatta. Köszönöm *Prószéky Gábornak*, hogy segített megérteni, hogy hogyan érdemes megközelíteni a természetes nyelveket, és megosztotta velem meglátásait, amelyek nagyon hasznosnak bizonyultak munkám során.

Külön hálával és köszönettel tartozom feleségemnek, *Ágnesnek*, szüleimnek és egész családnak, akik szeretetükkel és bátorításukkal folyamatosan támogattak tanulmányaim alatt.

Köszönöm *Wolfgang Porod Professzor Úrnak*, hogy a Notre Dame-i Egyetemen tölthettem egy évet vendégkutatóként. Nagyon hálás vagyok *Csurgay Árpád Professzor Úrnak* és *Csurgay Ildikónak* támogatásukért, bölcs tanácsaikért és a velük folytatott ösztönző beszélgetésekért.

Szeretnék továbbá köszönetet mondani korábbi és jelenlegi kollégáimnak, különösen *Timár Gergelynek*, aki felkeltette érdeklődésemet a kézírás-felismerés iránt, és akivel együtt kezdtük el felfedezni e terület szépségét, *Török Leventének* és *Orzó Lászlónak* a laborban folytatott személyes beszélgetésekért és a matematikai, programozási és nyelvi nehézségek leküzdésében nyújtott segítségükért, *Cserey Györgynek* baráti támogatásáért és segítőkészségéért, *Bálya Dávidnak*, akihez tudományos kérdésekben fordulhattam, *Szuhaj Mihálynak*, *Wagner Róbertnek* és *Lázár Annának*, akikkel a „Bionikus Szemüveg” projekten dolgozunk együtt, *Pohl Gábornak* és *Hodász Gábornak* a nyelvi korpuszok kezelésében nyújtott segítségükért valamint *Rekeczky Csabának* és *Zarándy Ákosnak* a tőlük kapott segítségért és tanácsokért. Sokan kollégám járult még hozzá ahhoz a termékeny intellektuális és baráti légkörhöz, amely nélkülözhetetlen a kutatáshoz, többek között *Szolgay Péter*, *Radványi András*, *Szirányi Tamás*, *Kék László*,

Földesy Péter, Szatmári István, Szlávik Zoltán, Ahmed Ayoub, Petrás István, Gál Viktor, Binzberger Viktor, Hillier Dániel, Soós Gergely, Mózsaár András, Fodróczy Zoltán, Jónás Péter, Rakos Balázs, Rác Zoltán, Imre Alexandra, Iván Kristóf és Oláh András. Ezúton köszönöm a segítségüket.

Itt szeretném kifejezni köszönetemet korábbi tanárainknak is, különösen *Horváth Gábornak*, aki diplomamunkám során volt témavezetőm, és megismertetett a neurális hálózatok alapjaival, *Recski András Professzor Úrnak*, akinek lenyűgöző algebra előadásai kitörölhetetlen nyomot hagytak emlékezetemben, *Rónyai Lajos Professzor Úrnak*, *Friedl Katalinnak* és *Bach Ivánnak*, amiért bevezettek a számításelmélet alapjaiba, valamint *Hirka Antal OSB* és *Pintér Ambrus OSB* atyáknak, akik matematikára és fizikára oktattak a Pannonhalmi Bencés Gimnáziumban.

Különösen is köszönöm *Keserű Katalinnak* és *Kékné Ladányi Katalinnak* (MTA-SZTAKI) valamint *Csókási Annának* és *Adorján Líviának* (PPKE-ITK) az adminisztráció útvesztőiben nyújtott támogatást.

Publikációk

A szerző folyóiratcikkei

- [1] **K. Karacs**, G. Prósztéky, and T. Roska, “CNN algorithms with spatial semantic embedding for handwritten text recognition,” *International Journal of Circuit Theory and Applications*, megjelenés alatt
- [2] L. O. Chua, **K. Karacs**, V. I. Sbitnev, J. Guan, and J. Shin, “A Nonlinear Dynamics Perspective of Wolfram’s New Kind Of Science. Part VIII: More Isles of Eden,” *International Journal of Bifurcation and Chaos*, megjelenés alatt
- [3] G. Tímár, **K. Karacs**, and Cs. Rekeczky, “Analogic Preprocessing and Segmentation Algorithms for Off-line Handwriting Recognition,” *Journal of Circuits, Systems and Computers*, vol. 12, no. 6, pp. 783–804, Dec. 2003.

A szerző konferencia-kiadványban megjelent cikkei

- [4] **K. Karacs** and T. Roska, “Locating and Reading Color Displays with the Bionic Eyeglass,” in *Proc. of the 18th European Conference on Circuit Theory and Design (ECCTD 2007)*, Seville, Spain, Aug. 2007. pp. 515–518.
- [5] G. E. Paziienza and **K. Karacs**, “An Automatic Tool to Design CNN-UM Programs,” in *Proc. of the 18th European Conference on Circuit Theory and Design (ECCTD 2007)*, Seville, Spain, Aug. 2007. pp. 492–495.
- [6] **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.
- [7] **K. Karacs** and T. Roska, “Route Number Recognition of Public Transport Vehicles via the Bionic Eyeglass,” in *Proc. of the 10th IEEE International Workshop on Cellular Neural Networks and their Applications (CNNA 2006)*, Istanbul, Turkey, Aug. 2006, pp. 79–84.
- [8] T. Roska, D. Bálya, A. Lázár, **K. Karacs**, R. Wagner, and M. Szuhaj, “System aspects of a bionic eyeglass,” in *Proc. of the 2006 IEEE International Symposium on Circuits and Systems (ISCAS 2006)*, Island of Kos, Greece, May 21–24, 2006, pp. 161–164.
- [9] **K. Karacs** and T. Roska, “Holistic Feature Extraction from Handwritten Words on Wave Computers,” in *Proc. of the 8th IEEE International Workshop on Cellular Neural Networks and their*

- Applications* (CNNA 2004), Budapest, Hungary, July 2004, pp. 364–369.
- [10] **K. Karacs**, G. Prószéky, and T. Roska, “Intimate Integration of Shape Codes and Linguistic Framework in Handwriting Recognition via Wave Computers,” in *Proc. of the 16th European Conference on Circuits Theory and Design* (ECCTD 2003), Krakow, Poland, Sept. 2003, pp. 409–412.
- [11] G. Tímár, **K. Karacs**, and Cs. Rekeczky, “Analogic Preprocessing and Segmentation Algorithms for Off-Line Handwriting Recognition,” in *Proc. of the 7th IEEE International Workshop on Cellular Neural Networks and their Applications* (CNNA 2002), Frankfurt, Germany, July 2002, pp. 407–414.

A szerző egyéb publikációi

- [12] L. Kék, **K. Karacs**, and T. Roska, Eds., *Cellular Wave Computing Library (Templates, Algorithms and Programs) V. 2.1.* MTA-SZTAKI, Budapest, Hungary, 2007.

Hivatkozások

A CNN technológiához kapcsolódó hivatkozások

- [13] L. O. Chua and L. Yang, “Cellular Neural Networks: Theory,” *IEEE Trans. on Circuits and Systems*, vol. 35, pp. 1257–1272, Oct. 1988.
- [14] L. O. Chua and L. Yang, “Cellular Neural Networks: Applications,” *IEEE Trans. on Circuits and Systems*, vol. 35, pp. 1273–1290, Oct. 1988.
- [15] L. O. Chua and T. Roska, “The CNN Paradigm,” *IEEE Trans. on Circuits and Systems*, vol. 40, pp. 147–156, Mar. 1993.
- [16] T. Roska and L. O. Chua, “The CNN Universal Machine: An Analogic Array Computer,” *IEEE Trans. on Circuits and Systems – II: Analog and Digital Signal Processing*, vol. 40, pp. 163–173, 1993.
- [17] A. Rodríguez-Vázquez, S. Espejo, R. Dominguez-Castron, J. L. Huertas, and E. Sanchez-Sinencio, “Current mode techniques for the implementation of continuous- and discrete-time cellular neural networks,” *IEEE Trans. on Circuits and Systems – II*, vol. 40, no.3, pp. 132–146, Mar. 1993.
- [18] Á. Zarándy, “The Art of CNN Template Design,” *International Journal of Circuit Theory and Applications*, vol. 27, pp. 5–23, 1999.
- [19] L. O. Chua and T. Roska, *Cellular neural networks and visual computing, Foundations and applications*. Cambridge, UK & New York: Cambridge University Press, 2002.
- [20] T. Roska, “Computational and Computer Complexity of Analogic Cellular Wave Computers,” *Journal of Circuits, Systems, and Computers*, vol. 12, pp. 539–562, 2003.
- [21] Analogic Computers Ltd, *Aladdin Professional*, <http://www.analogic-computers.com>
- [22] A. Rodríguez-Vázquez, G. Liñán, L. Carranza, E. Roca, R. Carmona, F. Jiménez, R. Dominguez-Castro, and S. Espejo, “ACE16k: The Third Generation of Mixed-Signal SIMD-CNN ACE Chips Toward VSoCs,” *IEEE Trans. on Circuits and Systems*, vol. 51, no. 5, pp. 851–863, 2004.
- [23] MTA-SZTAKI, *MatCNN*, <http://lab.analogic.sztaki.hu/Candy/matcnn.html>

Kézírásfelismeréshez kapcsolódó hivatkozások

- [24] M. Côté, E. Lecolinet, M. Cheriet, and C. Y. Suen, “Automatic reading of cursive scripts using human knowledge,” in *Proc. of the International Conference on Document Analysis and Recognition*, 1997, pp. 107–111.
- [25] D. Guillevic, D. Nishiwaki, and K. Yamada, “Word lexicon reduction by character spotting,” in *Proc. of the 7th International Workshop on Frontiers in Handwriting Recognition*, Amsterdam, Netherlands, 2000, pp. 373–382.
- [26] A. L. Koerich, R. Sabourin, and C. Y. Suen, “Large vocabulary off-line handwriting recognition: A survey,” *Pattern Analysis & Applications*, vol. 6, no. 2, pp. 97–121, 2003.
- [27] G. Lorette, “Handwriting recognition or reading? What is the situation at the dawn of the 3rd millennium?” *International Journal on Document Analysis and Recognition*, vol. 2, pp. 2–12, 1999.
- [28] S. Madhvanath and V. Govindaraju, “Holistic lexicon reduction for handwritten word recognition,” in *Proc. of the 8th SPIE International Symposium On Electronic Imaging: Science and Technology*, vol. 2660, San Jose, Mar. 1996, pp. 224–234.
- [29] S. Madhvanath and V. Krpāsundar, “Pruning large lexicons using generalized word shape descriptors,” in *Proc. of the International Conference on Document Analysis and Recognition (ICDAR 1997)*, Ulm, Germany, Aug. 1997, pp. 552–555.
- [30] S. Madhvanath and V. Govindaraju, “The role of holistic paradigms in handwritten word recognition,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 149–164, 2001.
- [31] K. M. Sayre, “Machine recognition of handwritten words: A project report,” *Pattern Recognition*, vol. 5, pp. 213–228, 1973.
- [32] L. Schomaker and E. Segers, “Finding features used in the human reading of cursive handwriting,” *International Journal on Document Analysis and Recognition*, vol. 2, pp. 13–18, 1999.
- [33] W. Senior and A. J. Robinson, “An Off-line Cursive Handwriting Recognition System,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 309–321, 1998.
- [34] T. Steinherz, E. Rivlin, and N. Intrator, “Offline cursive script recognition – a survey,” *International Journal on Document Analysis and Recognition*, vol. 2, pp. 90–110, 1999.
- [35] G. Timár, “Applied High-Speed Analogic Algorithms for Multitarget Tracking and Offline Handwriting Segmentation,” PhD. dissertation, Analogical and Neural Computing Laboratory, Computer and Automation Institute, Hungarian Academy of Sciences, Budapest, Hungary, 2006.

- [36] A. Vinciarelli, “A survey on Off-Line Cursive Script Recognition,” *Pattern Recognition*, vol. 35, no. 7, pp. 1433–1446, 2002.
- [37] M. Zimmermann and J. Mao, “Lexicon Reduction Using Key Characters in Cursive Handwritten Words,” *Pattern Recognition Letters*, vol. 20, pp. 1297–1304, 1999.

Táblák detektálásához és felismeréséhez kapcsolódó hivatkozások

- [38] C.-H. Cheng, C.-Y. Wu, B. Sheu, L.-J. Lin, K.-H. Huang, H.-C. Jiang, W.-C. Yen, and C.-W. Hsiao, “In the blink of a silicon eye,” *IEEE Circuits and Devices Magazine*, vol. 17, no. 3, pp. 20–32, May 2001.
- [39] A. Dollberg, H. G. Graf, B. Höflinger, W. Nisch, J. D. S. Spuentrup, and K. Schumacher, “A Fully Testable Retinal Implant,” in *Proc. of International Conference on Biomedical Engineering (BioMED)*, Salzburg, Austria, 2003, pp. 255–260.
- [40] P. Silapachote, J. Weinman, A. Hanson, R. Weiss, and M. Mattar, “Automatic Sign Detection and Recognition in Natural Scenes,” in *Proc. of IEEE Workshop on Computer Vision Applications for the Visually Impaired (in conjunction with CVPR)*, San Diego, California, June 2005.
- [41] R. Wagner and M. Szuhaj, “Color Processing in Wearable Bionic Eyeglass,” in *Proc. of 10th IEEE International Workshop on Cellular Neural Networks and their Applications (CNNA 2006)*, Istanbul, Turkey, Aug. 28-30, 2006, pp. 85–90.
- [42] T. Yamaguchi, Y. Nakano, M. Maruyama, H. Miyao, and T. Hananoi, “Digit Classification on Signboards for Telephone Number Recognition,” in *Proc. of 7th International Conference on Document Analysis and Recognition*, Edinburgh, UK, Aug. 2003, pp. 359–363.

Természetes nyelvek feldolgozásához kapcsolódó hivatkozások

- [43] G. Prószéky, M. Naszódi, and B. Kis, “Recognition Assistance,” in *Proc. of COLING-2002*, vol. II, Taipei, Taiwan, 2002, pp. 1263–1267.
- [44] G. Prószéky, “Humor: a Morphological System for Corpus Analysis,” in *Proc. of First TELRI Seminar on Language Resources and Language Technology*, Tihany, Hungary, 1996, pp. 149–158.
- [45] G. Prószéky and B. Kis, “Agglutinative and Other (Highly) Inflectional Languages,” in *Proc. of the 37th Annual Meeting of the Association for Computational Linguistics*, College Park, Maryland, USA, 1999, pp. 261–268.

- [46] *The British National Corpus*, version 2 (BNC World). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk/>, 2001.

A sejtautomatákhoz kapcsolódó hivatkozások

- [47] S. Wolfram, *A New Kind of Science*. Wolfram Media, Champaign, IL, USA, 2002.
- [48] L. O. Chua, V. I. Sbitnev, and S. Yoon, “A non-linear dynamics perspective of Wolfram’s new kind of science. Part III: Predicting the unpredictable,” *International Journal of Bifurcation and Chaos* vol. 14, no. 11, pp. 3689–3820, 2004.
- [49] L. O. Chua, V. I. Sbitnev, and S. Yoon, “A non-linear dynamics perspective of Wolfram’s new kind of science. Part IV: From Bernoulli shift to $1/f$ spectrum,” *International Journal of Bifurcation and Chaos* vol. 15, no. 4, pp. 1045–1183, 2005.
- [50] L. O. Chua, J. Guan, V. I. Sbitnev, and J. Shin, “A non-linear dynamics perspective of Wolfram’s new kind of science. Part VII: Isles of Eden,” *International Journal of Bifurcation and Chaos*, megjelenés alatt
- [51] T. Toffoli, “Computation and Construction Universality of Reversible Cellular Automata,” *Journal of Computer and System Sciences*, vol. 15, pp. 213–231, 1977.
- [52] R. P. Feynman, “Simulating Physics with Computers,” *International Journal of Theoretical Physics*, vol. 21, no. 6-7, pp. 467–488, June 1982.

Egyéb hivatkozások

- [53] Mathworks, *MATLAB*, <http://www.mathworks.com>