Csirik János

Képfeldolgozás és mesterséges intelligencia

Képfeldolgozás

A számítógépes képfeldolgozás alapvető feladata képi jellegű (leggyakrabban 2 vagy 3 dimenziós) adatok feldolgozása, elemzése. Egyebek mellett feladatai közé tartozik a képszegmentálás (a képen ábrázolt objektumok régióinak elkülönítése, körülhatárolása), a képregisztráció (a térben és/vagy időben elkülönülten felvett képek közötti térbeli/geometriai megfeleltetés), a képeken ábrázolt objektumok tulajdonságainak meghatározása, az objektumok és/vagy az ábrázolt színtér jellemzése. Képfeldolgozással ma már gyakorlatilag az élet szinte minden területén találkozhatunk az orvosi diagnosztikai rendszerektől az ipari gyártási és minőségbiztosítási folyamatokon át a biztonságtechnikai rendszerekig, az autonóm robotoktól a kamerákkal felvértezett okostelefonokig.

A képfeldolgozás szegedi története majdnem fél évszázados múltra tekint vissza, és egy orvosi diagnosztikai rendszer alapjainak lefektetésével indult.

A SEGAMS nukleáris medicinai rendszer

1969-ben Csernay László kérésére Kalmár László professzor Csirik Jánost ajánlotta segítségül radioaktív izotóp segítségével nyert képek számítógépes feldolgozásához. A feldolgozás kezdetben nagyon nehézkes volt: a szkenner által nyomtatott beütésszámokat először lyukszalagra kellett lyukasztani, a lyukszalag tartalmát mágnesszalagra írni, majd az így összegyűjtött anyagot a MINSZK-22 számítógépen lehetett feldolgozni. Az így nyert képeket a sornyomtatón kinyomtatva lehetett láthatóvá tenni.

1974-ben a Gamma Művek hozzákezdett egy orvosi diagnosztikai célú számítógépes gammakamera-rendszer hardverének kidolgozásához, a szoftver tervezésével és kivitelezésével a SzOTE Központi Izotópdiagnosztikai Laboratóriumában alakult munkacsoportot bízta meg Csernay vezetésével, aki a Kibernetikai Laboratóriumban frissen megalakult Képfeldolgozási kutatócsoport bevonásával vállalta a feladatot. Még nem volt kész a TPAi számítógépen működő rendszer (SEGAMS) első változata, amikor a Nemzet-

közi Atomenergia Ügynökség által meghirdetett versenyen Los Angelesben a SEGAMS-szal világviszonylatban harmadik helyezést értünk el.

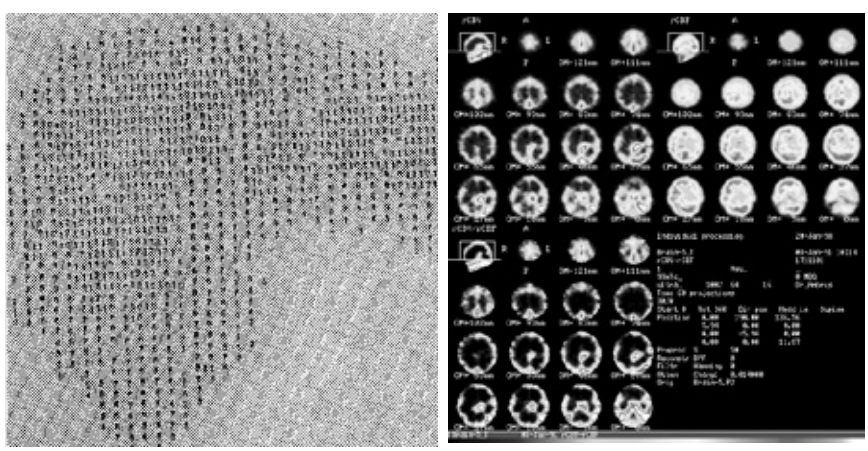
A Segams-80 volt az első olyan nukleáris medicinai rendszer a világon, amely lehetőséget adott arra, hogy a felhasználó anyanyelvén „társaloghasson" a rendszerrel. Csak a rendszer szövegkönyvtárának lefordítása volt szükséges egy újabb nyelvi változat elkészítéséhez. A standardizált rutinvizsgálatok előre definiált felvételi módként beépíthetők a rendszerbe, így elvégzésük rövidített beszélgetéssel kezdeményezhető. A vizsgálatok kiértékeléséhez „orvos barát" komplex programozási lehetőséget is biztosítottunk.

A TPAi alapú rendszerünk legfejlettebb változata a SUPER-SEGAMS már FORTRAN programozási lehetőséget is biztosított, lehetővé téve a rendszer valamennyi funkciójának egyszerű aktiválását. A SUPER-SEGAMS segítségével az EKG-val vezérelt szívvizsgálatok is elvégezhetők voltak, és alkalmas környezetett biztosított az emmissziós tomográfiás rendszer kísérleti megvalósítására is (Csernay és mtsai., 1982).

1987-ben kezdtük meg az Amiga alapú microSEGAMS fejesztését. A rendszer biztosította a nukleáris medicinai vizsgálatok – többek között tomográfiás vizsgálatok – elvégzését és komplett klinikai programokkal való kiértékelését. A SEGAMS változataiból több mint 300 darabot értékesített a GAMMA Művek a különböző országokban.

A munka vezetője, koordinátora prof. Csernay László (SzOTE Központi Izotópdiagnosztikai Laboratórium).

A hardver a Gamma Művek Nukleáris Osztályán készült, vezető tervező: dr. Billing Ádám és dr. Kári Béla.



1. ábra: Bal oldalon egy MINSZK-22-n, jobb oldalon egy microSEGAMS-szal végzett feldolgozás eredménye látható.

A szoftvert a JATE Kibernetikai Laboratórium Képfeldolgozási munkacsoport készítette, vezető programfejlesztők: dr. Csirik János, dr. Makay Árpád†, dr. Máté Eörs, dr. Kuba Attila†.

A rendszer 1977-ben és 1983-ban BNV nagydíjat nyert, 1981-ben az MTA Akadémiai Díjban részesült prof. Csernay László, dr. Csirik János, dr. Makay Árpád† és dr. Máté Eörs.

A rendszer fejlesztése alapján 1 MTA doktori, 5 kandidátusi és 1 egyetemi doktori értekezés született. A rendszer használatához ezres nagyságrendű tudományos publikáció köthető.

A SZOTE-PACS

A FEFA III. és IV. (1994–1995) pályázatok elnyerésével közös fejlesztés indult Csernay László professzor vezetésével a SzOTE Központi Izotópdiagnosztikai Laboratórium, az Orvosi informatikai intézet, a Radiológiai Klinika és a Kuba Attila igazgatásával, önálló egységként működő JATE Alkalmazott Informatikai Tanszék munkatársainak közreműködésével. A digitális képalkotó technikák térhódításával szükségessé vált ugyanis egy olyan rendszer kifejlesztése, amely a különféle képalkotó berendezések által előállított vizsgálati képanyagot egységes formában tárolja és lehetővé teszi alkalmazásukat oktatási, kutatási és diagnosztikai célokra. A kifejlesztett SZOTE-PACS (Kuba és mtsai., 1996) volt az ország egyik első képarchiváló és továbbító rendszere, amely a nagy tárolókapacitású központi kiszolgálót, valamint megjelenítő munkaállomásokat kötött össze szabványos adatformátumok és kommunikációs protokollok alkalmazásával az egyetemi számítógép-hálózaton keresztül. A szerveren Oracle adatbázis segítette az archívumban történő gyors keresést, a tárolási fájlformátum és az eszközök kommunikációjának alapját pedig az akkor még csak születő DICOM nemzetközi szabvány biztosította. Ez utóbbi azóta a PACS rendszerek alapvető szabványává vált. A hazánkban a maga korában egyedülálló SZOTE-PACS 2005-ig üzemelt. Feladatát egy évtizeddel később a Radiológiai Klinikán egy későbbi, „gyári" PACS rendszer vette át.

3D orvosi képek szegmentálása és regisztrációja

A szegedi universitas megalakulása után a képfeldolgozó csapat az SZTE Képfeldolgozás és Számítógépes Grafika Tanszéken belül tevékenykedett tovább. Ekkor kezdődött az a gyümölcsöző kutatás-fejlesztési együttműködés, melyben a tanszék kutatói, a Radiológiai Klinika orvosai, valamint a GE Medical

Systems klinikai szoftvereket fejlesztő csapata fogott össze. A feladat szervek szegmentációja 3 dimenziós CT-képeken, azaz az egyes szerveket ábrázoló képi régiók meghatározása. Algoritmusokat dolgoztunk ki több hasi és kismedencei szerv (vese, máj, lép, prosztata, húgyhólyag), a gerinccsatorna, valamint a látószervi objektumrendszer komponenseinek szegmentálására CT-felvételekből, elsődlegesen sugárterápia-tervezési alkalmazáshoz.

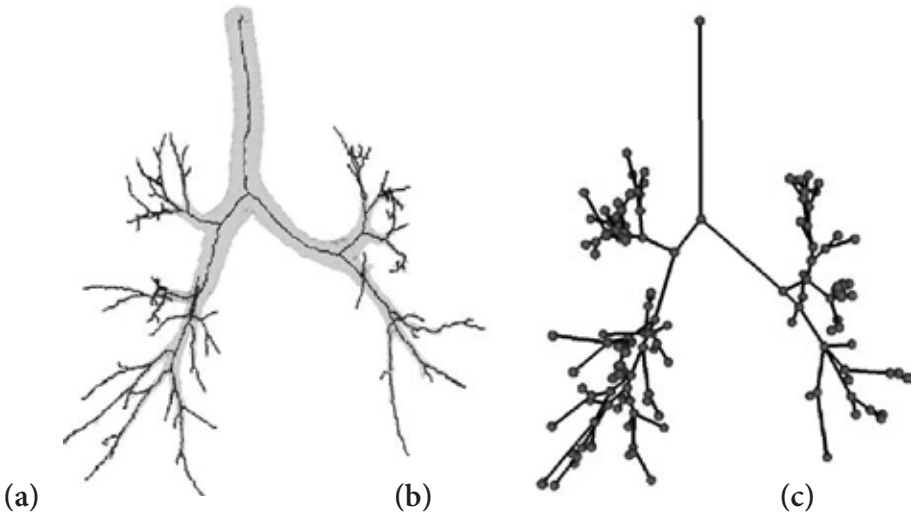
A modell alapú szegmentálási technikák, valamint az időbeli elemzések elengedhetetlen feltétele, hogy ismert legyen a különböző felvételek közötti térbeli megfelelőség. Ennek a meghatározásához automatikus képregisztrációs módszereket dolgoztunk ki, elsősorban a kismedencei régióhoz.

A szegmentációs és regisztrációs módszereink, melyekből több nemzetközi szabadalom is született (Fidrich és mtsai., 2008; Tanács és mtsai., 2007), beépítésre kerültek a GE termékeinek szoftverrendszereibe. A kutatás-fejlesztési munkákba egyetemi és PhD-hallgatók is bekapcsolódtak, és később erre a bázisra építve alakult meg a GE HealthCare szegedi szoftverfejlesztő részlege.

Vékonyítás, vázkijelölés, strukturális jellemzők

A képszegmentáció és képregisztráció valójában csak eszközök, és az igazi cél a képeken ábrázolt objektumok, szervek, régiók jellemzése, diagnosztikai célú leírása. Az egyszerű térfogati mérések mellett nagyon fontosak az alaki jellemzők is, mint pl. a csöves szerkezetű objektumok középvonala, mely fontos bizonyos mérésekhez (pl. lumen meghatározására), de ugyanúgy elengedhetetlen bizonyos megjelenítési technikákhoz is (pl. virtuális endoszkópia).

A vékonyítás (mint iteratív redukció) alkalmas bináris objektumok középvonalának meghatározására. Az általunk kifejlesztett, bizonyítottan topológia-megőrző 3-dimenziós vékonyító algoritmusokat (Palágyi és mtsai, 2012) főleg orvosi képekből szegmentált objektumok elemzésére alkalmazták. Sikeres együttműködést alakítottunk ki pl. a Medical University of Graz (Ausztria) és a The University of Iowa (USA) kutatóival (Sorantin és mtsai., 2002; Palágyi és mtsai., 2006).



2. ábra: (**a**) Légcsőszűkület elemzés. A CT-vizsgálatból szegmentált légcső átmérőjét a vékonyítással kapott középvonalra merőleges 2D képszeleteken becsüljük. (**b**-**c**) Légútfák kvantitatív kiértékelése. (**b**) A CT-vizsgálatokból szegmentált légútfa és középvonala, valamint (**c**) a középvonal alapján kapott formális struktúra, mely valamennyi ágra tartalmazza annak térfogatát, felszínét, hosszát és átmérőjét. A formális struktúrák illesztése lehetővé teszi az anatómiai címkézést, a légzés elemzését és számos tüdőbetegség automatikus diagnosztizálását.

Diszkrét tomográfia

A *tomográfia* a görög „tomos" (szelet) és „grapho" (írni) szavakból származik, és általánosságban szeleteken alapuló képalkotást értünk alatta. Egy háromdimenziós objektum szeleteit úgy képezhetjük, hogy azt egy tetszőleges kétdimenziós síkkal elmetsszük, ezáltal egy kétdimenziós keresztmetszetet kapunk. Gyakorlati szempontból a szeletek elkészítését nemroncsoló módon végezzük, leggyakrabban röntgensugarak segítségével. A vizsgált objektumon ezen sugarak áthatolnak, miközben a sugárnyaláb fotonjainak egy része elnyelődik. A fotonsugár kezdeti és az áthaladás után mért gyengült intenzitásából következtetni lehet arra, hogy a sugár milyen anyagon és milyen hosszan haladt át. Az egy irányból nyert intenzitásgyengüléseket leíró adatot az objektum egy *vetületének* nevezzük. A tomográfia feladata az, hogy ezen vetületekből összeállítsuk (rekonstruáljuk) az eredeti objektum leírását, azaz minden pontjában meghatározzuk az objektum anyagát.

A számítógépes tomográfia (CT) legrégibb és mai napig leggyakoribb alkalmazási területe az orvosi diagnosztika. A képalkotáshoz jelenleg legelterjedtebben

alkalmazott eljárás az úgy nevezett *szűrt visszavetítés*, mely több száz vetületet igényel a megfelelő minőségű kép előállításához.

Újabban az ipari roncsolásmentes vizsgálathoz is használnak CT-berendezéseket. A módszer segítségével felderíthetők a tárgyak belsejében keletkezett légbuborékok, repedések, vizsgálhatók a korrózió hatásai, de akár lerakódások vagy idegen anyag jelenléte is. A vetületképzés egyik korlátja az ipari objektum nagy kiterjedése lehet, azaz ha nem minden irányból helyezhető el a tárgy a CT-szkennerben, így nem minden irányból készíthető róla vetület. Hasonló a helyzet akkor is, ha a tárgy olyan anyagból van, amely elnyeli a röntgensugárzást. Ebben az esetben neutronsugárzás segítségével még mindig átvilágítható az objektum, az ilyen sugarak előállítása azonban költséges, amely szintén gátja lehet a megfelelő minőségű képalkotáshoz szükséges számú vetület kinyerésének.

Az ipari vizsgálatok esetében tehát sokszor csak kevés vetület áll rendelkezésünkre, vagy azok nem egyenletesen minden irányból adottak. A szűrt visszavetítés ilyen esetekben ritkán alkalmazható sikeresen. Mégis van remény jó minőségű kép előállítására, ha az előállítandó képről rendelkezünk előzetes információkkal. Ez lehet például a vizsgált tárgy megközelítő alakja, belső szerkezetének strukturáltsága (például műanyaghabok vagy golyókat tartalmazó tárgyak vizsgálata esetén), vagy az azt felépítő általában kevés számú anyag ismerete. Orvosi alkalmazások esetében is elvárás, hogy a pácienst lehetőleg minél kevesebb sugárzás érje, így itt is jelentős igény van olyan matematikai modellek és számítógépes módszerek kidolgozására, melyek minél kevesebb vetületből az előzetes információ felhasználásával tudnak a diagnózis felállításához még megfelelő minőségű képet előállítani. Ezen modellek és számítógépes algoritmusok tervezésével és elemzésével foglalkozik a *diszkrét tomográfia*.

A SZTE Képfeldolgozás és Számítógépes Grafika Tanszékén a tudományterület születésével lényegében egy időben, az 1980-as évek közepén kezdődött meg a diszkrét tomográfia elméletének és gyakorlati alkalmazhatóságainak vizsgálata. Tanszékünkről az első cikk Kuba Attila nevéhez kötődik, mely egy speciális geometriai tulajdonság, a *hv-konvexitás* feltételezésével ad hatékony rekonstrukciós technikát abban a szélsőséges esetben, amikor csak két vetület áll rendelkezésre, és az előállítandó kép bináris (Kuba, 1984). Kuba Attila vezetésével a szegedi diszkrét tomográfiai iskola az ezredforduló környékére a terület nemzetközileg is elismert csapatává nőtte ki magát. A szakterület két alapvető összefoglaló irodalmán túl (Herman és Kuba, 1999; Herman és Kuba, 2007) számos nemzetközi folyóirat- és konferenciapublikációnk jelent meg. Az általunk legintenzívebben kutatott területek a hv-konvex bináris képek

elmélete, a diszkrét tomográfia alkalmazása az ipari nemroncsoló tesztelésben, valamint a különböző irányú vetületek hasznosságának, információtartalmának vizsgálata (Varga és mtsai., 2014).

Tanszékünkön a diszkrét tomográfia területén jelenleg egy fiatal kutatókból és PhD-hallgatókból álló lelkes csapat dolgozik az újonnan felmerülő elméleti és gyakorlati problémák megoldásán. A kutatási területen az elmúlt években több doktori disszertáció készült, a cikkeinkre kapott független hivatkozások pedig azt mutatják, hogy eredményeink nemzetközileg is ismertek és elismertek.

Kollaboratív mobil vizuális számítások

Az utóbbi években a *kamerával felszerelt okostelefonok* robbanásszerű elterjedésének lehetünk szemtanúi, amelyek új lehetőségeket nyitnak a számítógépes látás tudományterület kutatásában és gyakorlati alkalmazásában is. Ezek az eszközök egyre jobb képalkotó képességekkel, folyamatosan növekvő számítási kapacitással, lokalizációs szenzorokkal, hálózati kapcsolódási lehetőségekkel, valamint felhasználók milliárdjaival rendelkeznek. Figyelembe véve ezeket a képességeket, az egy területen jelenlévő felhasználók és készülékeik egy *ad-hoc kamerahálózat* tagjainak tekinthetők. A készülékek közötti kommunikációval eloszthatók az elkészült képeken alapuló számítások, valamint új vizuális információk állíthatók elő, amelyek aztán megosztásra kerülhetnek a résztvevők között.

A számítógépes látás számos klasszikus és új problémája átültethető ebbe az új környezetbe. Ezek egyike a *térbeli rekonstrukció*, amelynek feladata fotók alapján a felhasználót körülvevő színtér térbeli modelljének előállítása. Alapját a mobil eszközökkel a színtérről készült fotók és szenzoradatok alkotják. Mivel a készülékek – kisebb-nagyobb hibával – képesek földrajzi pozíciójuk és földfelszínhez viszonyított orientációjuk kiszámítására, egy *láthatóság gráf algoritmus* (Kovács, 2013) segítségével meghatározhatóvá válik, mely kamerák azok, amelyek átfedő látómezővel rendelkeznek. Vizuális számítások elvégzéséhez rendszerint szükségünk van annak pontos ismeretére, hogy az adott készülék hogyan képezi le a 3D világot a vetítési síkra (*belső kalibráció*), valamint hogy a készülékek a térben hogyan helyezkednek el egymáshoz képest (kamerahálózat *külső kalibrációja*).

Az általunk javasolt rekonstrukciós megoldás nagyobb kiterjedésű, sík felületek felhasználását igényli, amely igen gyakori városi környezetben, épületekről készült fotók esetén. A megoldáshoz a kalibrációs paraméterek mellett

szükségünk van a sík felületek közötti megfeleltetés-keresésekre is a képeken (Domokos és mtsai., 2012). A rekonstrukció alapja, hogy ugyanazt a sík felületrészt a kamerák más nézőpontból látják, így az perspektív torzulást szenved. A torzulás ismeretében meghatározható a sík felület orientációja és kamerától vett távolsága.

3. ábra: Példa 3D rekonstrukcióra. A két eredeti kép (felső sor), valamint a rekonstrukció eredménye két nézőpontból (alsó sor). Az első képen vörös körvonal jelzi a kijelölt sík területek határát.

Irodalom:

- Csernay László, Makay Árpád, Csirik János, Máté Eörs: SUPER-SEGAMS kézikönyv, I–II kötet (magyar, német, angol, orosz nyelven), Budapest, 1982, p. 1–256.
- Kuba Attila, Alexin Zoltán, Nagy Antal, Nyúl László, Palágyi Kálmán, Nagy Marianna, Csernay László, Almási László: DICOM based PACS and its application in the education. In: Proc. of the 14th EuroPACS Meeting, Heraklion, Crete, pp. 46–49, 1996.

- Fidrich Márta, Makay Géza, Máté Eörs, Balogh Emese, Kuba Attila, Nyúl László G., Kanyó Judit: Systems and methods for segmenting an organ in a plurality of images. US Patent US7388973 B2, 2008.
- Tanács Attila, Máté Eörs, Kuba Attila: Method and system for automatically transforming CT studies to a common reference frame. US Patent US7259762 B2, 2007.
- Palágyi Kálmán, Németh Gábor, Kardos Péter: Topology Preserving Parallel 3D Thinning Algorithms. In: Brimkov V.E., Barneva R.P. (szerk.): Digital Geometry Algorithms. Theoretical Foundations and Applications to Computational Imaging, Springer, pp. 165–188, 2012.
- Kálmán Palágyi, Juerg Tschirren, Eric A. Hoffman, Milan Sonka: Quantitative analysis of pulmonary airway tree structures. Computers in Biology and Medicine 36, pp. 974–996, 2006.
- Erich Sorantin, Csongor Halmai, Balázs Erdőhelyi, Kálmán Palágyi, László G.Nyúl, Krisztián Ollé, Bernhard Geiger, Franz Lindbichler, Gerhard Friedrich, Karl Kiesler: Spiral-CT-based assessment of tracheal stenoses using 3-D-skeletonization. IEEE Transactions on Medical Imaging 21, pp. 263–273, 2002.
- Gabor T. Herman, Attila Kuba (szerk.): Discrete Tomography: Foundations, Algorithms, and Applications, Applied and Numerical Harmonic Analysis. Birkhauser, 1999.
- Gabor T. Herman, Attila Kuba (szerk.): Advances in Discrete Tomography and Its Applications, Applied and Numerical Harmonic Analysis. Birkhauser, 2007.
- Kuba Attila: The reconstruction of two-directionally connected binary patterns from their two orthogonal projections. Computer Vision, Graphics, and Image Processing, 27, pp. 249–265, 1984.
- Varga László G., Nyúl László G., Nagy Antal, Balázs Péter: Local and global uncertainty in binary tomographic reconstruction. Computer Vision and Image Understanding, megjelenés alatt, 2014.
- Domokos Csaba, Nemeth Jozsef, Kato Zoltan: Nonlinear Shape Registration without Correspondences. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34, pp. 943–958, 2012.
- Kovács Levente: Processing Geotagged Image Sets for Collaborative Compositing and View Reconstruction. In: Proc. of ICCV Workshop on Computer Vision for Converging perspectives, pp. 460–467, 2013.

Mesterséges intelligencia

Teljesen elosztott adatbányászat

A személyi számítógépek és az internet elterjedésével a 90-es évek végén egy forradalom zajlott le, amelynek során előtérbe kerültek azok az informatikai alkalmazások és fejlesztések, amelyek a személyi számítógépeket kihasználva nagyszámú felhasználóra épülő alkalmazásokat hoztak létre. Először központosított (elsősorban webes), később a közvetlen kommunikációra épülő ún. peer-to-peer (tipikusan fájlcserélő vagy internetes telefon) alkalmazások jelentek meg. A 2000-es évek második felétől egy második forradalom vette kezdetét, ami a személyi számítógépek új generációjára, az egyre terjedő tabletekre és okostelefonokra épül. Ezek rengeteg adat gyűjtését teszik lehetővé a felhasználójukról és annak környezetéről, ami a korábbinál sokkal gazdagabb élményt és több alkalmazást tesz lehetővé. Ezek az adatok származhatnak az eszközökben található számos érzékelőből, vagy keletkezhetnek különböző alkalmazások használata során, mint amilyenek pl. a böngészők vagy kapcsolattartásra szolgáló alkalmazások. A jelenlegi gyakorlat szerint ezeket az adatokat az egyes alkalmazások központi helyen tárolják (a felhőben), és itt végzik el a statisztikai elemzést és adatbányászatot. Ez a gyakorlat azonban számos problémát vet fel. Ide tartozik, hogy a központi adattárolás nagyon költségessé válik egy bizonyos adatmennyiség felett, ami kizárja a közcélú vagy kísérletező jellegű nonprofit alkalmazásokat. Ezen kívül jelenleg szinte lehetetlen több különböző alkalmazásból származó adatokat egyesíteni és közös modellben elemezni, holott ezek magán az eszközön együtt vannak jelen, míg a központi tárhelyen már elválnak.

Kutatásaink során azt vizsgáljuk, hogy a legfontosabbnak tekinthető, tipikusan felhőben végzett, számítási és adatfeldolgozási funkciókat hogyan lehetne a központi komponensek kikerülésével, vagy nagyon korlátozott bevonásával, megvalósítani, tisztán az egyes eszközök közötti közvetlen kommunikációra építve. Ez a ma már hagyományosnak tekinthető asztali gépek platformján is túlmutat az ott ismert alkalmazásokon, amik gyakran nem jelentenek számítási feladatokat, csak tárolási vagy kommunikációs problémákat oldanak meg. Emellett természetesen a modern okostelefonok platformján rengeteg új alkalmazást tenne lehetővé a jelenleg elérhető technológiánál sokkal olcsóbban és rugalmasabban.

Az általános adatbányászati problémák vizsgálatától eltekintve, az elosztott probléma vizsgálata is immár évtizedes múlttal rendelkezik a kutatócsoportban.

A kezdeti eredmények Jelasity Márk nevéhez fűződnek, aki pletykaalapú algoritmusokat javasolt néhány egyszerű statisztikai feladat elosztott megoldására, mint amilyen pl. az átlagszámítás (Jelasity és mtsai., 2004). Az ilyen algoritmusok alapötlete, hogy az egyes eszközök közösen egy diffúzióra emlékeztető folyamatot valósítanak meg, amelynek során kiegyenlítődnek az egyes eszközök közelítései.

Az évek során sokkal általánosabb problémákat is sikerült megoldanunk, megtartva a pletyka-algoritmusok hatékonyságát és robosztusságát. Itt az alapötletünk az ún. sztochasztikus gradiens módszerek adaptálása, amelynek során az adatokat leíró statisztikai modell (pl. legegyszerűbb esetben az egy átlag, bonyolultabb esetben akár egy nemlineáris döntési felület) véletlen sétát végez az eszközök felett, és minden eszközön az ott található adatokat felhasználva egy gradiens lépést hajt végre (Ormándi és mtsai., 2013). Természetesen az alapötletnek és magának a problémának számos vonatkozását vizsgáljuk, kezdve az okostelefonok használatának, halózati kapcsolatának, energiaszintjének stb., modellezésétől egészen olyan kurrens algoritmusok megvalósításáig mint amilyen pl. az ún. boosting algoritmus vagy az ún. többkarú bandita algoritmusok. Számos kapcsolódó témát is vizsgáltunk, mint amilyenek a magánjellegű adatok védelme, ami egyre fontosabb kutatási terület, és az algoritmusaink adaptivitása abban az esetben, ha dinamikusan változik az adathalmaz, amin dolgozunk.

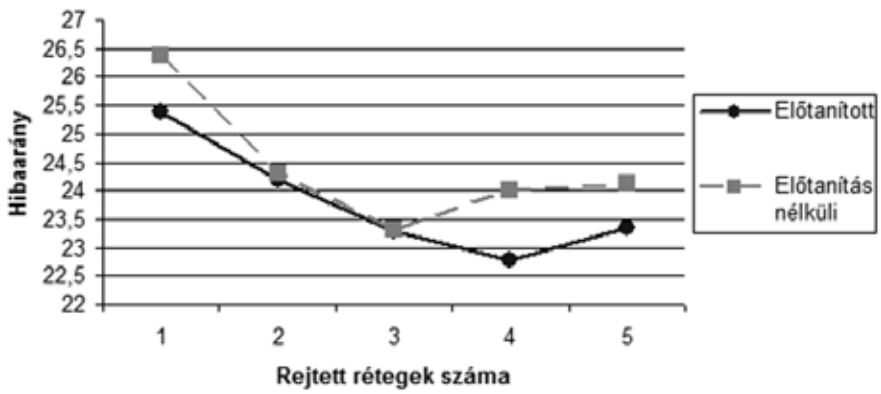
A témában számos nemzetközi kapcsolattal rendelkezünk, többek között egy EU FP7-es projektet (QLectives) is sikeresen lezártunk, ahol az algoritmusainkat peer-to-peer hálózatokban alkalmaztuk ajánlórendszerek és spamszűrés megvalósítására.

Beszédfelismerés mély neuronhálókkal

A gépi beszédfelismerés technológiájának matematikai alapjai 30 éve lényegében változatlanok. Ekkor vezették be a rejtett Markov-modelleket, amelyek egy egységes matematikai keretet adnak a sorozatok – így pl. a beszédjelek – statisztikai alapú felismeréséhez. Azóta természetesen rengeteg apró finomítást javasoltak a technológiához, és ezáltal az eredmények is sokat javultak, de általánosságban elmondható, hogy a beszédfelismerés területén mindeddig a kis lépésekben való előrehaladás volt jellemző. A 2006-os évben azonban megjelent egy új gépi tanulási algoritmus, melyet feltalálói „mély neuronhálónak” neveztek el. A módszert eleinte képi alakfelismerési feladatokon

tesztelték, az első beszédfelismerési kiértékelésére 2009-ben került sor. A mély neuronháló már ebben az első kísérletben minden korábbi eredményt megdöntött, nagy visszhangot váltva ki a kutatók körében.

Az első cikk óta a terület rohamosan fejlődik, a kiugróan jó eredmények miatt egyre többen fordulnak ehhez a technológiához, és így a megjelenő publikációk száma is exponenciálisan nő. Az MTA-SZTE Mesterséges Intelligencia Kutatócsoportnál már több mint egy évtizede foglalkozunk a beszédfelismerés neuronhálókat használó irányzatával, így szerencsés módon viszonylag gyorsan be tudtunk kapcsolódni a mély neuronhálók kutatásába. Legelső lépésként természetesen a külföldi eredményeket próbáltuk reprodukálni. A 4. ábra azt mutatja, hogy a neuronháló rétegszámának növelésével hogyan csökken a felismerő rendszer hibája:



4. ábra: Beszédfelismerő rendszer hibájának alakulása a mély neuronháló rétegszáma és az alkalmazott tanítási algoritmus függvényében.

Mint látható, a mély neuronhálók hatékonyságának egyik forrása maga a „mélység", azaz hogy sok rétegből állnak (ezzel szemben egy hagyományos neuronháló csak egyetlen rejtett réteget tartalmaz). A sok réteg azonban magában nem elég, hanem szükség van egy új tanító algoritmusra is, amelyet „előtanítás" névvel illet az irodalom. Amint az ábrán is látható, az előtanítás különösen sok réteg esetén fontos, ilyenkor az ily módon tanított hálózat lényegesen jobb eredményeket ad, mint a hagyományos, nem előtanított háló. Habár a Hinton és munkatársai által javasolt előtanítási algoritmus nagyon hasznosnak bizonyult, a gyakorlati alkalmazása nehézkes és időigényes. Ezért hamar megindult a kutatás a hasonló hatású, de egyszerűbb és gyorsabb megoldások iránt. Ebbe a kutatásba már csoportunk is be tudott kapcsolódni, így a beszédfelismerésben az elsők között próbáltuk ki az ún. „egyenirányított"

(rectified) neuronokból felépülő hálózatokat. Másokkal párhuzamosan azt kaptuk, hogy ez a hálótípus az előtanított hálózatokkal egyenértékű felismerési pontosságra képes, de az implementációja sokkal egyszerűbb, tanítása pedig gyorsabb (Tóth, 2013a).

A képi alakfelismerési algoritmusok hatékonyságát le szokta rontani, ha a felismerendő alak (például betű) nincsen a kép közepén, hanem valamelyik irányban elcsúszik. Ennek elkerülésére hagyományosan különféle „normalizálási" technikákat szoktak bevetni, de ennél hatékonyabb megoldás a neuronháló átalakítása oly módon, hogy a kép eltolt változatait is fel tudja ismerni. Ezt a technológiát konvolúciós modellezésnek hívják, és a képfelismerésben régóta kísérleteznek vele, de a beszédfelismerés eszköztárába csak mostanában került be. Mivel a képfeldolgozással ellentétben a beszédjel esetében a két ábrázolási tengelynek – idő és frekvencia – különböző szerepe van, ezért a konvolúciós technikát a két tengelyen eltérő módon kell alkalmazni. Mi elsőként az időtengely mentén való konvolúció technológiáját dolgoztuk ki, és megmutattuk, hogy a sztenderd mély neuronhálókhoz képest ezzel a módszerrel akár 10%-os hibacsökkenés is elérhető (Tóth, 2013b). Legújabban pedig kombináltuk módszerünket az Abdel-Hamid és mtsai. által alkalmazott, frekvenciatengely mentén működő konvolúcióval. Jelenleg ez a kombinált modell tartja a legkisebb felismerési hiba rekordját a széles körben vizsgált TIMIT adatbázison (Tóth, 2014).

Habár a publikálhatóság szempontjából az angol nyelvű tesztek a fontosak, természetesen igyekszünk az újonnan feltalált módszereket a magyar nyelv felismerésére is bevetni. Kísérleteink alapján úgy látjuk, hogy a mély neuronhálók a magyar nyelvű felismerésben is lényeges eredményjavulást tudnak hozni (Grósz és Tóth, 2014). Jelenleg is folyamatban vannak azok a Budapesti Műszaki Egyetemmel közösen futó kísérletek, amelyekben a mély neuronhálóinkat az ő nagyszótáras magyar nyelvű felismerőrendszerükbe építve értékeljük ki. A kezdeti eredmények azt mutatják, hogy mély neuronhálókkal akár 10-30%-os hibacsökkenés is várható a hagyományos felismerési technológiához képest.

Számítógépes nyelvészet

A mesterséges intelligencia egyik ága, a számítógépes nyelvészet vagy más néven a nyelvtechnológia az emberi nyelv számítógépes eszközökkel történő megértését és feldolgozását célzó tudományterület. A Szegedi Tudományegyetemen az 1990-es évek vége óta zajlanak nyelvtechnológiai kutatások

informatikus és nyelvész kollégák együttműködésében, továbbá a kutatómunkában részt vesznek egyetemi és PhD-hallgatók is. A folyó kutatások három fő pillér köré csoportosíthatók: nyelvi adatbázisok létrehozása, nyelvi elemző eszközök fejlesztése és információkinyerő alkalmazások fejlesztése. A műhely magyar és angol nyelvű szövegek feldolgozásával is foglalkozik, céljuk, hogy az egyes kifejlesztett technológiák nyelvfüggetlenek vagy legalább könnyen adaptálhatóak legyenek.

Gépi tanuló algoritmusok segítségével nagymértékben csökkenthető a szövegek feldolgozásához szükséges emberi munkaerő-ráfordítás. A gépi tanuláson alapuló módszerek alkalmazása azonban megköveteli nagyméretű, kézzel annotált szöveges adatbázisok, úgynevezett korpuszok meglétét. A csoport aktív közreműködője volt a magyar nyelvre kifejlesztett Szeged Korpusz és Treebanknek, amely 82 000 mondatot és 1,2 millió szövegszót tartalmaz, ezzel a legnagyobb méretű, kézzel annotált, magyar nyelvű adatbázis. A Szeged Korpusz és Treebankben megtaláljuk minden egyes szövegszó lehetséges morfológiai elemzését, az aktuális kontextusnak megfelelő morfológiai elemzését, valamint minden mondathoz szintaktikai elemzés is tartozik (Csendes és mtsai., 2005). A magyar wordnet lexikális adatbázis létrejöttében szintén fontos szerepet vállalt a kutatócsoport: az adatbázis a magyar szavakat a köztük húzódó jelentéstani kapcsolatok szerint rendezi csoportokba (Miháltz és mtsai., 2008). Mindemellett számos kisebb, adott feladatra kifejlesztett angol és magyar nyelvű korpuszt is létrehoztak a munkatársak.

A magasabb rendű nyelvtechnológiai alkalmazásoknak előfeltétele a szövegek előelemzése, azaz mondatra és szövegszóra bontása, valamint a szavak szófaji egyértelműsítése és a mondatok szintaktikai elemzése. A csoport által létrehozott *magyarlanc* nevű programcsomag magyar nyelvű szövegek előelemzését teszi lehetővé mondat- és szószegmentáló, szófaji elemző és függőségi (szintaktikai) elemző moduljaival (Zsibrita és mtsai., 2013). A csoport ezen túl rendelkezik tulajdonnév-felismerő és -kategorizáló, illetve jelentés-egyértelműsítő technológiákkal is mind angol, mind magyar nyelvre.

Az alap szintaktikai és szemantikai elemző szoftvercsomagokra építve számos nyelvtechnológiai alkalmazás került kifejlesztésre a csoportban (egy részük ipari megrendelésre). Ilyen alkalmazás például a vásárlási események kiemelése üzleti hírekből, az [origo] hírportál archívumának (400 ezer hír) automatikus címkézője, fehérje-interakciós információkinyerő biológiai publikációkból, betegségek és tüneteik azonosítása orvosi zárójelentések folyó szöveges részeiben, illetve önéletrajzok automatikus feldolgozása. A fentiek mellett a csoport munkatársainak érdeklődési területe többek között a többszavas kifejezések számítógépes kezelésére, a bizonytalanság és tagadás gépi

felismerésére, az automatikus kulcsszókinyerésre és a webes szövegek feldolgozására terjed ki.

A szegedi nyelvtechnológiai csoport 16 magyar pályázatban vett részt az elmúlt években, és a témában több mint 100 nemzetközi publikációval rendelkezik, valamint igen szép eredményeket ért el nemzetközi számítógépes nyelvészeti versenyeken (automatikus BNO-kódolás: 1. helyezés; orvosi rekordok anonimizálása: 1. helyezés; metonímiafeloldás: 1. helyezés, tweetek polaritásának elemzése: 1. helyezés). A csoport tagjai írták a 2007 nyarán megjelent *Szövegbányászat* című könyv információkinyeréssel foglalkozó fejezetét is. Kutatócsoportunk alapító tagja a magyarországi nyelv- és beszédtechnológiai műhelyeket tömörítő nyelv- és beszédtechnológiai platformnak, melynek keretében aktív szerepet vállal a tudományág eredményeinek minél szélesebb körben való közérthető ismertetésében és népszerűsítésében.

2003 óta minden évben Szegeden rendezik meg a Magyar Számítógépes Nyelvészeti Konferenciát, mely napjainkra a hazai nyelvtechnológiai kutatások legfőbb szakmai fórumává vált. A csoport nemzetközi konferenciák szervezéséből is kiveszi a részét: a csoport kutatóinak szervezésében valósult meg a CoNLL-2010 konferenciához kapcsolódó nemzetközi verseny, továbbá a 2008-as Global WordNet Conference-nek is Szeged adott otthont.

Fuzzy halmazok, többtényezős döntések és a genetikus algoritmusok

A fuzzy (elmosódott határú) halmaz a humán, „nem egzakt" fogalmak formális leírását valósítja meg, egyfajta bizonytalanság (uncertainty) leírása. A nem egzakt fogalmak matematikai leírására szolgáló halmazhoz tartozási függvény helyett a karakterisztikus függvény approximációjaként előálló elemi preferencia függvényt vezettük be, amivel az elvárási szint, döntési küszöb és a preferencia élessége is megadható. Megmutattuk, hogy a halmazhoz tartozási függvény mindig előállítható két függvény kompozíciójaként. Segítségével a szubjektív faktor leválasztható a fuzzy elméletről. A fuzzy elmélet kutatások másik célja a folytonos logikai operátorok vizsgálata.

A természetes nyelvű módosító szavak az unáris operátorokkal modellezhetők, itt megadtuk ezek általános alakját. A többtényezős döntéseknél meghatározó, hogy az egyes tényezők milyen fontossággal szerepelnek az értékelés során. Megmutattuk, hogy a súlyozási eljárás is operátorfüggő transzformáció és linearizálható a súlyokra nézve.

A fenti eredmények alapján létrehoztunk egy új folytonos logikai rendszert (rugalmas rendszert), ahol a döntési szint változtatható (Dombi 2013,

Dombi 2012, Dombi 2011). Így egy olyan logikát kaptunk, amelyik jól közelítik a klasszikus kétértékű logikát. A rugalmas rendszer segítségével a gépi tanulás neurális hálózatai leírhatók, illetve tovább általánosíthatók. Így a tanuló algoritmusok egy új osztályát hoztuk létre. A rugalmas rendszer segítségével adatbázisok természetes nyelvű leképezése valósítható meg.

Big Data projektekben a folytonos logikai operátorok segítségével diszkrét változók vizsgálata gyorsan elvégezhető. Optimalizálásnál a szigorú korlátok feloldása folytonos logikai operátorok segítségével kezelhető. Genetikus algoritmussal pedig nehéz feladatok optimalizálása végezhető el, ahol az optimumot fokozatosan közelítjük.

Új típusú tanuló algoritmusok: neurális hálózatok csomópontjainak folytonos logikai operátorokkal való helyettesítésével a robotikához kapcsolódó feladatok hatékonyan oldhatók meg.

Irodalom:

- Márk Jelasity, Alberto Montresor, Ozalp Babaoglu: Gossip-based aggregation in large dynamic networks. *ACM Transactions on Computer Systems*, 23 (3): 219–252, August 2005. (doi:10.1145/1082469.1082470)
- Ormándi Róbert, Hegedűs István, Jelasity Márk: Gossip learning with linear models on fully distributed data. *Concurrency and Computation: Practice and Experience*, 25(4):56–571, 2013. (doi:10.1002/cpe.2858)
- Tóth László: Phone Recognition with Deep Sparse Rectifier Neural Networks, Proc. ICASSP 2013, 6985–6989, 2013.
- Tóth László: Convolutional Deep Rectifier Neural Nets for Phone Recognition, Proc. Interspeech 2013, 1722–1726, 2013.
- Tóth László: Combining Time- and Frequency-Domain Convolution in Convolutional Neural Network Based Phone Recognition, Proc. ICASSP 2014, 190–194, 2014.
- Grósz Tamás, Tóth László: Új eredmények a mély neuronhálós magyar nyelvű beszédfelismerésben, *Proc. MSZNY 2014*, 3–13, 2014.
- Csendes Dóra, Csirik János, Gyimóthy Tibor, Kocsor András: The Szeged Treebank. In: *Proceedings of TSD 2005*, Karlovy Vary, 123–131, 2005.
- Miháltz Márton, Hatvani Csaba, Kuti Judit, Szarvas György, Csirik János, Prószéky Gábor, Váradi Tamás: Methods and Results of the Hungarian WordNet Project . In: *Proceedings of the Fourth Global WordNet Conference*, Szeged, 310–320, 2008.

- Zsibrita János, Vincze Veronika, Farkas Richárd: magyarlanc: A Toolkit for Morphological and Dependency Parsing of Hungarian. In: *Proceedings of RANLP 2013*, Hissar, 763–771, 2013.
- Dombi József: *On a certain class of aggregative operators* Information Sciences (2013) ISSN 0020-0255
- Dombi József: *Pliant operator system* In *Studies in computational intelligence*, Recent Advances in Intelligent Engineering Systems Ed. by J. Fodor, R. Klempous, C.P. Suárez Araujo: *Springer* 31–58, 2012.
- Dombi József: *Fuzziness measure in the Pliant system: The Vagueness measure* Acta Technica Jaurinensis; Series Intelligentia Computatorica Vol.4. No.1, 3–11, 2011