

Új markovi változásdetekciós modellek a képfeldolgozásban

Ph.D. disszertáció tézisei

Benedek Csaba
mérnök-informatikus

Tudományos vezető:
Dr. Szirányi Tamás
az MTA doktora



Pázmány Péter Katolikus
Egyetem
Információs Technológiai Kar



MTA
Számítástechnikai és
Automatizálási Kutatóintézet

Budapest, 2008

1. Bevezetés

Napjainkban a képsorozatokot kezelő, rögzítő illetve megjelenítő számítógépes rendszerek jelentősége és ezáltal száma is gyorsan növekszik, ami elengedhetlenné teszi az általuk szolgáltatott nagy mennyiségű információ automatikus feldolgozását. Csak néhány alkalmazást felsorolva, gondolhatunk biztonsági videomegfigyelésre, forgalomfigyelésre, városi forgalomirányításra, légi felderítésre, erdő-tűz észlelésére, vagy földrajzi területek vegetációs változásainak felmérésére.

A változások észlelése fontos kezdeti elemét képezi a látvány értelmezésének. Egy biztonsági kamera képe megváltozik, ha megjelenik egy új személy a helyszínen, egy táskát felejtnek a teremben vagy – hosszabb megfigyelés esetén – a monitorozott területen felépítenek egy házat. A kép azonban akkor is megváltozik, ha a kamerát elmozdítjuk, vagy ha a megvilágítás módosul, ami sok esetben nem tartalmaz lényeges információt, csupán zajként jelenik meg a feldolgozó rendszerek számára. Fontos megjegyezni, hogy mindig az adott feladat dönti el, hogy milyen típusú változások észlelésére van szükség, és melyek azok a képi különbségek, amiket figyelmen kívül kell hagyni.

Az előbbi példákat is tekintve sok esetben érvényes feltevés, hogy a képsorozatok egymást követő kockáit összehasonlítva az „érdekes módon” megváltozott képrészletek kijelölik a további vizsgálatra érdemes területeket. Ennél azonban gyakran több is igaz: a változott régiók alaki, pozíciós, és méretbeli jellemzői közvetlenül felhasználhatók objektumfelismerés és eseményanalízis különböző feladataiban. Más oldalról megközelítve, a szakirodalomban fellelhető számos magasabb szintű képfeldolgozási művelet épül az előzőleg kinyert pontos változásmaszokra, hatékonyságuk így nagymértékben függ az előfeldolgozási lépés minőségétől.

Mint azt az alkalmazások széles skálája is mutatja, a változásdetekció tág gyűjtőfogalom. Értekezésemben a témába illő három konkrét feladattal foglalkozom, melyek mind a pontos célokat, mind a körülményeket tekintve jelentősen különböznek.

Az első feladat mozgó alakzatok régióinak elkülönítése a *háttértől*

biztonsági videofelvételeken. Feltételezzük, hogy a képsorozatok álló kamerák vették fel, és a rögzítés folyamatosan, ám alacsony vagy ingadozó képkockasebességgel történt. Fő kihívás a mozgó árnyékok eltüntetése az alakzatok sziluettmaszkjáról, valamint a színük alapján háttérbe olvadó objektumrészletek azonosítása.

A második alkalmazás célja szintén objektum elmozdulások észlelése, ebben az esetben azonban mozgó légi járművekről készített képpárok képezik az eljárás bemenetét. A kameramozgás miatt a képeket először egymásra kell illeszteni. Az így kapott, ún. *regisztrált* képek azonban nem csak a mozgó alakzatok helyén térnek el, mivel az illesztési hibák és különféle 3D geometriai torzítások további hamis különbségeket okoznak. A legnagyobb kihívást az utóbbi zajhatások csökkentése jelenti.

A harmadik feladat kapcsán jelentős időkülönbséggel (több hónap vagy év) készített légi felvételeket vetünk össze. Mivel a megváltozott fényhatások, és évszakokra jellemző eltérések miatt az összehasonlítható területek megjelenése nagyon különböző lehet, ebben az esetben pontosan definiálni kell, milyen típusú változásokat akarunk jelezni. A dolgozatban az új beépült területek észlelésével foglalkozunk.

2. Módszerek, eszközök

A kidolgozott módszerek háttérét főként a valószínűségi számítás, geometriai modellezés és képfeldolgozás területeiről összegyűjtött eredmények adják. A modellek matematikai értelemben a *Markov véletlen mezők* [17] megvalósításainak tekinthetők. A feldolgozás kimenete egy szegmentált kép (például bináris mozgásmaszk), mely globális energiaoptimalizáció eredménye:

$$\arg \max_{\underline{\omega}} P(\underline{\omega} | \mathcal{O}) = \arg \min_{\underline{\omega}} \left(-\log P(\mathcal{O} | \underline{\omega}) + \sum_{C \in \mathcal{C}} V_C(\underline{\omega}) \right), \quad (1)$$

ahol \mathcal{O} megfigyelt képi jellemzőket jelöl, $\underline{\omega}$ egy lehetséges szegmentálást, \mathcal{C} az úgynevezett klikkek halmaza, ahol klikk alatt páronként kölcsönhatásban álló képpontok csoportját értjük, P valószínűség,

V_C pedig a klikkeken értelmezett potenciálfüggvény.

A dolgozat újdonságai a szegmentációs osztályok jellemzőinek kiválasztásában (\mathcal{O} származtatása), az osztályok valószínűségi leírásában ($P(\mathcal{O}|\underline{\omega})$ meghatározása), valamint a modellek strukturális felépítésében (\mathcal{C} , V_C definiálása) jelentkeznek.

A módszerek hatékonyságának mérése valós alkalmazási környezetből származó tesztvideókon illetve képsorozatokon történt. Az első feladat tesztkörnyezete a PPKEyes nevű digitális videomegfigyelő rendszer, mely a Pázmány Péter Katolikus Egyetem Információs Technológiai Karán működik. A megvalósított modell validációjához egy nyilvánosan elérhető tesztadatbázist is felhasználtunk. A második és harmadik feladat megoldása során használt légi képeket az ALFA NKFP projekt terhére vásároltuk, részben a Földmérési és Távérzékelési Intézettől (FÖMI).

A modellek tervezéséhez és teszteléséhez Matlab és Visual Studio .Net környezeteket használtunk. A képfeldolgozási algoritmusok C++ nyelvű implementációját az Intel OpenCV könyvtár csomagja könnyítette meg. A tézis és a kapcsolódó publikációk \LaTeX szövegszerkesztővel készültek.

3. Új tudományos eredmények

1. tézis: Új tér-időbeli statisztikát használó Markov véletlen mező alapú modellt dolgoztam ki videójelenetek előtér-háttér szegmentációjára és a mozgó alakzatok vetülő árnyékainak detekciójára. Kísérletileg igazoltam, hogy a módszer hatékonyabb korábbi hasonló célú és hasonló környezeti feltevéseket használó modelleknél.

A szerző kapcsolódó publikációi [1][2][4][5][14]

A disszertáció készítőjének társszerzős publikációi, melyekben az ismertetett modellnek, illetve előzményeinek az alkalmazásai szerepelnek: [8][9][10][11][12]

A bemutatott módszer célja az előtér, háttér és árnyék osztályok hatékony elkülönítése *valódi* biztonsági vagy forgalomfigyelő alkal-



1. ábra. Előtér- (fehér) árnyék- (szürke) és háttérregiók (fekete) elkülönítése biztonsági videofelvételeken [1. tézis]

mazások által szolgáltatott videókon. A modell feltételezi, hogy a felvételeket álló kamerák rögzítették, ám a videók rossz minőségűek és alacsony képkockasebességűek is lehetnek. A tervezés során figyelembe vettem a kameraajt, a megvilágítás időbeni változásait, valamint inhomogén fényvisszaverő felületek jelenlétét a helyszínen. Az (1) képletben felírt energiát az alábbi alakot ölti:

$$\sum_{s \in S} -\log P(o(s)|\omega(s)) + \sum_{\{r,s\} \in \mathcal{C}} \Theta(\omega(s), \omega(r)), \quad (2)$$

ahol $P(o(s)|\omega(s))$ annak a valószínűsége, hogy az s jelű pixelnél lokálisan mért $o(s)$ jellemzőt a pixel $\omega(s)$ címkéjével jelölt osztály (előtér, háttér vagy árnyék) generálta. Vizsgálataim hatékony jellemzők származtatására, és az osztályok megfelelő valószínűségi leírására irányultak. A szegmentálás simaságáért felelős $\Theta(.,.)$ tag a rács szomszédos képpontjain értelmezett, és bünteti ha a vizsgált pont-pár címkéi különbözőek.

A bevezetett modell új elemeit az altézisekben részletezem.

1.1. Új statisztikus és adaptív színmodellt adtam mozgó árnyékok jellemzésére és észlelésére. Bemutattam, hogy az eljárás hatékonyabb a korábbi megközelítéseknél, ha fényvisszaverő környezet nem ideálisan Lamberti.

A szakirodalomban fellelhető árnyékmodellek mesterségesen leegyszerűsített környezetek modellezésén alapulnak (például tisztán *Lamberti* fényvisszaverő felületek jelenlétét feltételezik). Hatékonyságuk drasztikusan csökken olyan valódi élethelyzetekhez köthető helyszíneken, ahol a feltevések nem állják meg helyüket.

Új árnyékmodellt vezettem be, mely parametrikusan hangolható különböző megvilágítási körülményekhez és a környezetben megfigyelhető fényvisszaverő hatásokhoz, oly módon, hogy a megfigyelt helyszín és egy tisztán *Lamberti* visszaverő környezet különbségeit írja le valószínűségi alapon. A módszer ahelyett, hogy konstans fényvisszaverő tulajdonságokat feltételezne a teljes helyszínen, a lokális különbségeket sztochasztikus úton modellezi. A képpontok aktuálisan megfigyelt és háttérben mért színét felhasználva az árnyékok esetén szignifikáns értékkel rendelkező leírót származtattam, és ebben a jellemző térben az árnyékok által lefedett tartományt valószínűségi sűrűségfüggvényekkel írtam le. A paraméter adaptációs algoritmus az árnyék tartományának időbeni követésén alapul. Számos valódi környezetben végzett kísérlettel igazoltam, hogy a modell által detektált árnyék-régiók lényegesen pontosabbak, mint a tisztán *Lamberti* modell felhasználása esetén.

1.2. A szomszédos pixelek színértékeinek térbeli statisztikáját felhasználva az előtér osztály új valószínűségi leírását adtam meg. Megmutattam, hogy a megközelítés javítja a háttér vagy árnyék színű objektum részek detekcióját a korábban közölt modellekkel szemben, akkor is, ha a videó alacsony vagy bizonytalan képkockasebességű.

Az előtér régiókat több korábbi módszer pusztán a háttér és árnyékmodellhez nem illeszkedő képrészletekként azonosította, feltételezve, hogy az előtér objektumai bármely szint azonos valószínűséggel vehetnek fel. Ez a megközelítés sok esetben a háttér vagy árnyék színű objektumrészek hamis osztályozásához vezet. Más megoldások olyan képkockasebességre érzékeny jellemzőket származtattak (például alakzatok követése révén), melyek sok valódi alkalmazásban a képátvitelre fordítható korlátos sávszélesség miatt nem állnak rendelkezésre.

Multimodális eloszlással jellemzett színmodellt vezettem be az előtérre, amely a szükséges szín statisztikát nem a képkocka sebességtől nagyban függő időbeni, hanem az e tekintetben robusztus térbeli információból nyeri ki. Azt a feltevést használtam, hogy a különböző objektum régiók jellemző szín és/vagy textúra mintázatokkal rendelkeznek, ami alapján az előtérben előforduló színek gyakoriságát lokálisan jellemeztem az egyes pixelek környezetében. A tesztek igazolták, hogy módszeremmel számos olyan előtér részlet detektálhatóvá válik, melyeket az egyenletes eloszlást használó előtérmodellek hibásan figyelmen kívül hagytak.

1.3. Valószínűségi leírást adtam a háttérben illetve árnyékban mért tetszőleges mikrostruktúra jellemzők értékeire. Olyan textúra alapú kiegészítést vezettem be a korábbi altézisekben javasolt Markov mezős színmodellhez, mely a háttér lokális jellemzőihez illeszkedő kernelválasztási stratégiát használ. Mesterségesen generált és valódi tesztképeken is megmutattam, hogy a modell javítja a szegmentáció minőségét a puszta szín alapú, vagy egyetlen textúra jellemzőt használó megközelítésekhez képest.

Bár különböző képszegmentációs feladatok kapcsán a szín és textúrajellemzők integrálása számos korábbi munkában megjelent, a textúrakomponensek csak akkor adnak valódi plusz információt a régiók elkülönítéséhez, ha az aktuálisan jelenlévő objektumok vagy a helyszín textúrája illeszkedik a kiválasztott jellemzőkhöz. Valós életből származó felvételek esetén általában nem találunk olyan textúra leírót, mely a teljes helyszínt tekintve megfelelő. Ugyanakkor az irreleváns jellemzők a szegmentáció segítése helyett a zajt növelik.

Valószínűségi modellt adtam a háttérben és árnyékban megfigyelhető mikrostrukturális textúrajellemzők értékeire, feltéve, hogy a leírókat tetszőleges 3×3 -as kernel generálja. A különböző pixeleken különböző kernelek használhatók, melyeket egy adaptív választási stratégia jelöl ki, figyelembe véve a háttér lokális textúráis jellemzőit. Megmutattam, hogy a bemutatott árnyékmodellhez is illeszthető a megközelítés, és a közelítő eloszlásparaméterek analitikusan számít-

hatók. Kísérletileg megmutattam, hogy az eljárás pontosabban jelöli ki az objektum régiókat, mint a pusztán szín alapú szegmentáció, illetve az egyetlen kernelt használó fúziós technikák.

*1.4. Kísérleti úton megmutattam, hogy az elterjedt színterek közül a CIE $L^*u^*v^*$ színmodell a leghatékonyabb az árnyék detekciójára mind az árnyék-leírótérben, a kézenfekvőnek bizonyult elliptikus elválasztó felülettel végzett pixel szintű osztályozás, mind a javasolt Markovi optimalizáció több szintérre történő általánosítása esetén.*

Nyitott kérdés a szakirodalomban, hogy melyik szintérben érdekes az árnyékdetekiós vizsgálatokat elvégezni, és mekkora a helyes választás jelentősége.

Két kísérletsorozattal megmutattam, hogy a megfelelő szintér kiválasztásával a szegmentálás minősége lényegesen javítható olyan esetekben, amikor az árnyékmodell paramétereinek számát praktikus okokból alacsonyan kell tartani.

Az első kísérletsorozat elvégzéséhez terveztem egy tisztán pixel szinten mért leírókon alapuló, több szintérral is együttműködni képes előtér-árnyék osztályozót. Mivel pixel szinten maradván az előtérpontok időbeni színeloszlása általában nehezen és pontatlanul becsülhető, az árnyék pontok leírói által lefedett tartományt az egyosztályos klasszifikáció eleve alapján írtam le a jellemzőtérben. Az elkülönítéséhez használt elliptikus elválasztó felület általános relevanciáját kísérletekkel és a szakirodalom részletes áttekintésével és elemzésével támasztottam alá. Az osztályozó modell és a korábban is használt teszt adatbázis segítségével a színtereket kísérleti úton összehasonlítottam.

Második lépésben bemutattam a korábbi altézisekben ismertetett Markov mezős szegmentációs modell szintér független kiterjesztését, és a színterek összehasonlítását ebben a keretben is elvégeztem.

Mindkét kísérleti megközelítés a CIE $L^*u^*v^*$ szintér egyértelmű dominanciáját hozta.

Fontos megjegyezni, hogy mivel az első kísérletsorozat az árnyék-leírókon kívül nem használ fel más járulékos információt, eredményei objektívebben és általánosabban jellemzik szintérválasztás közvetlen



2. ábra. Objektum elmozdulások detekciója mozgó légi járműről készült képpáron. Az ábrán az eljárás bemenetét képező, mozgó kamerával rögzített képek és az elmozdulások maszkja látható [2.1 tézis]

hatását. A Markov mezős kísérletek ugyanakkor azt bizonyítják, hogy a megfelelő színmodell kiválasztása az összetett – szomszédságot, textúrát és környezeti színtesztisztikát is figyelembe vevő – modellben is mérhető teljesítmény javulást eredményez.

2. tézis: Új háromrétegű, címkefúzióan alapuló Markov mezős modelleket adtam a kameramozgástól független objektum elmozdulások detekciójára, valamint több év időközönbséggel készített légi képpárokon megfigyelhető beépített területek változásainak észlelésére. Elvégeztem a modellek kísérleti validációját.

A szerző kapcsolódó publikációi [3][6][13]

2.1. Új statisztikai modellt adtam kameramozgástól független objektum elmozdulások detekciójára légi járművekről készített képpárokon. Kísérletileg megmutattam a javasolt megközelítés előnyeit tisztán 2D regisztráció alapú, valamint lokális parallaxiscsökkentést alkalmazó korábbi modellek használatával szemben.

Eltérő kamerapozícióból készített fotók összehasonlítása a képek automatikus egymásra illesztésével kezdődik. Ez a regisztrációnak nevezett művelet azonban – a jelenleg elérhető eszközökkel – gyakran nem oldható meg pontosan, mivel a 3D környezet teljes és hiba

nélküli rekonstrukcióját igényelné. Módszerem ezért a „tökéletes regisztráció” feltételezése nélkül ad statisztikai becslést az előtér régióira. Az eljárás sűrű parallaxis hatás jelenlétével számol, azonban feltételezi, hogy az ebből eredő térbeli torzítás korlátos mértékű. Értekezésemben megmutattam, hogy az összetartozó pixelek szürkeségi értékének különbsége és a lokális korrelációs számítás egymást hatékonyan kiegészítő jellemzőket nyújt objektumelmozdulás detekcióhoz olyan képpárokon, melyeket előzőleg egy automatikusan számítható globális projektív transzformáció segítségével regisztrálunk.

Az előbbieken kinyert jellemzők integrálására bevezettem egy új háromrétegű Markov mezős struktúrát. A két szélső réteg szegmentálása a különböző jellemzőkön alapul, míg a középső réteg a végső változásmaszkot jeleníti meg anélkül, hogy a mérésekkel közvetlen összeköttetésben lenne. Rétegen belüli kapcsolatok biztosítják a szegmentált képek simaságát, míg rétegek közötti linkek felelősek a középső réteg szemantikailag helyes címkézéséért.

A markovi energiatag (1. egyenlet) az alábbi módon számítható:

$$\sum_s -\log P_s^d + \sum_s -\log P_s^c + \sum_{i,\{r,s\}} \Theta(\omega(r^i), \omega(s^i)) + \sum_s \zeta_s,$$

ahol P_s^d és P_s^c a megfigyelt jellemzők és a hozzájuk rendelt rétegek címkézésének konzisztenciáját írják le, a 2. képletben alkalmazott valószínűségi módon. A $\Theta(.,.)$ függvény felelős a sima szegmentációkért az egyes (i -vel indexelt) rétegeken, ζ_s értéke pedig $\pm\rho$ konstans, attól függően, hogy az s pixelhez tartozó címkék a három rétegen megfelelnek-e az előírt címke-fúziós szabályoknak.

A modell kísérleti validációja a megközelítés előnyeit mutatta ugyanerre a gyakorlati feladatra adott korábbi megoldásokhoz képest.

2.2. Markovi modellt terveztem, mely új beépített területek detekciójára alkalmas jelentős időkülönbséggel készült légi felvételeken. Megmutattam, hogy az összehasonlított képek szegmentációi között pixelszintű összeköttetéseket létesítve a zajos osztályleírókra robosztus régió szintű változásdetekciós eljárás nyerhető.

Jelentős időkülönbséggel készített felvételek összehasonlítása esetén a pixelszintű eljárások helyett gyakran a robosztusabb, régió szintű megközelítéseket alkalmazzák. Az összevetett képek korábbi módszerekkel történő régiókra bontása azonban a zajos, illetve szemantikailag hiányos osztályleírók hatása miatt gyakran nem teljesen egyforma, így a régiók összevetése számos hamisan megjelenő különbséget is tartalmazhat.

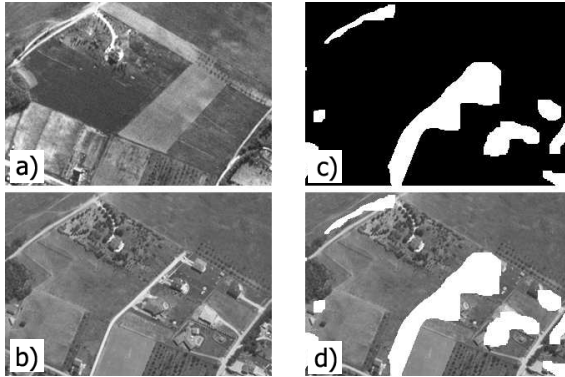
A disszertációban Markovi keretrendszert adtam strukturális területváltozások detekciójára a 2.1 tézispontban bemutatott háromrétegű modell segítségével. Esetünkben a két szélső réteg ugyanarról a földrajzi területről, de jelentős időkülönbséggel (több hónap vagy év) készített felvételek szegmentációját végzi, míg a középső réteg ismét a detektált változástérképet jeleníti meg. A módszer a képek együttes szegmentálását és a változasmask generálását egyetlen globális optimalizációként végzi el, az összetartozó pixelek között összeköttetéseket létesítve. Így a módszer sima és összefüggő régiók létrehozását segíti elő mind a szegmentált képeken, mind a változástérképen. A modellt beépített és természetes területek változásait detektáló alkalmazás esetén teszteltem. Kísérletileg megmutattam a többrétegű megközelítés előnyeit hagyományos egyrétegű Markov mezők kompozíciójával szemben.

4. Az eredmények alkalmazási területei

A dolgozatban bemutatott algoritmusok magasabb szintű képfeldolgozási eljárások előfeldolgozási lépéseként használhatók, főként videomegfigyelés és légi felderítés területein.

A módszerek közvetlenül kapcsolódnak az elmúlt években lezárult, illetve jelenleg is folyó kutatási projektekhez a Pázmány Péter Katolikus Egyetem, valamint az MTA-SZTAKI részvételével. A következőkben összegzem a legfontosabbakat.

A MUSCLE EU Project *Shape Modelling E-csoportja* alakzatok felismerését célozta meg képadatbázisok indexelése céljából. Fő feladatai objektumok alak alapú analízise és alak jellemzők tanulása, prior ismeretek alapján történő szegmentáció és alakzat alapján történő eseményfelismerés. Az említett feladatokban létfontosságú



3. ábra. Beépített területek változásának detekciója jelentős időkülönbséggel készült légi képpáron. a)-b) légi felvétel 2000 és 2005 évről c) a detektált változások maszkja (új házak és utak) d) a változásmaszk visszavetítve a második képre [2.2 tézis]

az alakzatok pontos sziluettjeinek a kinyerése, mint kezdeti lépés. Az *Autonóm Légi Felderítés és Navigáció* projekt (NKFP 2/046/04 NKTH) fő feladata egy olyan kompakt látórendszer létrehozása, amely a pilóta nélküli repülő járművek önálló vizuális felismerő és navigációs egysége lehet. A rendszernek érzékelnie és földi támogatás nélkül értelmeznie kell a vizuális információt (hosszú távú) navigációs döntések autonóm kialakításához. A fejlesztett rendszer polgári felhasználási célja jellegzetes terepformák és felszíni morfológiák, tereptárgyak, mozgó objektumok, illetve megváltozott környezeti viszonyok felismerése.

A *GVOP (3.1.1.-2004-05-0388/3.0)* projekt elsődlegesen a videoanyagok értékelésével, besorolásával, másrészt a számunkra érdekes részletek visszakeresésével, vagy a videofelügyeleti rendszerekben az érdelemes események azonosításával foglalkozik. Valamennyi feladat esetén az objektumelmozdulások detekciója fontos információt ad.

5. Köszönetnyilvánítás

Mindenekelőtt köszönettel tartozom témavezetőmnek, Szirányi Tamásnak, aki töretlen bizalommal támogatta munkám az elmúlt években, türelmesen bátorított a kezdeti akadályok leküzdésének idején és mértéktartásra tanított a sikeres időszakokban.

A doktori munkám elvégzéséhez szükséges anyagi és szellemi feltételeket valamint az infrastruktúrát a Pázmány Péter Katolikus Egyetem Információs Technológiai Kara (PPKE ITK) és a Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutatóintézete (MTA SZTAKI) biztosította. Köszönöm Roska Tamásnak, a PPKE ITK Multidiszciplináris Műszaki Tudományok Doktori Iskolájának és a SZTAKI Analogikai és Neurális Számítógépek Laboratórium vezetőjének, hogy lehetőséget biztosított számomra az itt folyó *Ph.D.* tanulmányaimra.

A disszertáció alapjául szolgáló munkában témavezetőmön kívül közvetlen segítséget nyújtottak külföldi és hazai kollégáim: Josiane Zerubia, Xavier Descombes (mindketten INRIA Ariana) és Kató Zoltán (Szegedi Egyetem). Prof. Zerubia meghívására többször szakmai látogatást tehettem az általa vezetett Ariana kutatócsoportnál Franciaországban. Ugyanitt említeném Xavier Vilasis-Cardona-t, aki a Barcelonai Ramon Llull Egyetemre invitált meg szemináriumra és egy ezt követő szakmai vitára.

Köszönöm a disszertáció előbírálóinak és bírálóinak munkáját és felvetéseiket.

Tudományos pályám elindításáért külön köszönet illeti Vámos Tibort, aki egyetemi hallgatóként alkalmazott csoportjában az MTA-SZTAKI-ban.

Közvetlen munkatársaim –mai nevén– a SZTAKI EEE kutatócsoportjából számos tanáccsal, ötlettel láttak el, közleményeim ellenőrzésében, javításában segítettek: Szlávik Zoltán, Havasi László, Petrás István, Kovács Levente.

A Pázmány Egyetemen folyó oktatási munkába Vágó Zsuzsa, Szirányi Tamás és Ruttkay Zsófia segítségével kapcsolódhattam be, akik emellett önálló témafeldolgozást is vezettek számomra.

Köszönöm Péri Mártonnak angoltudásom fejlesztését, és a kézírata-

imba ennek ellenére becsúszó nyelvi hibák javítását.

Itthoni és külföldi doktorandusz évfolyamtársaim érdemei vitathatatlanok szakmai és baráti szempontból egyaránt: Harczos Tamás, Hegyi Barnabás, Bankó Éva, Soós Gergely, Gyimesi Gergely, Ercsey-Ravasz Mária, Szálka Zsolt, Zeffer Tamás, Horváth Péter és Giovanni Paziienza, Szolgay Dániel. Köszönöm Iván Kristófnak a disszertáció elkészítéséhez adott technikai segítséget. Köszönet ezenkívül valamennyi munkatársamnak a PPKE-ITK, MTA-SZTAKI és INRIA intézményekben.

Köszönöm családomnak, kedvesemnek és valamennyi barátomnak, akik minden lehetséges módon támogattak.

A kutatómunkához további anyagi segítséget nyújtottak magyarországi és Európai Unió K+F projektek: OTKA #49001, EU project MUSCLE (FP6-567752), ALFA Magyar K+F Projekt, GVOP (3.1.1.-2004-05-0388/3.0), és a MUSCLE Shape Modelling E-Team.

6. Publikációk

6.1. A szerző folyóiratpublikációi

- [1] **Cs. Benedek** and T. Szirányi, „Bayesian foreground and shadow detection in uncertain frame rate surveillance videos,” *IEEE Transactions on Image Processing*, vol. 17, no. 4, pp. 608–621, 2008.
- [2] **Cs. Benedek** and T. Szirányi, „Study on color space selection for detecting cast shadows in video surveillance,” *International Journal of Imaging Systems and Technology*, vol. 17, no. 3, pp. 190–201, 2007.

6.2. A szerző nemzetközi konferenciapublikációi

- [3] **Cs. Benedek**, T. Szirányi, Z. Kato, and J. Zerubia, „A multi-layer MRF model for object-motion detection in unregistered airborne image-pairs,” in *Proc. IEEE International Conference on Image Processing (ICIP)*, vol. VI, (San Antonio, Texas, USA), pp. 141–144, IEEE, Sept. 2007.
- [4] **Cs. Benedek** and T. Szirányi, „Markovian framework for foreground-background-shadow separation of real world video scenes,” in *Proc. Asian Conference on Computer Vision (ACCV), Lecture Notes in Computer Science (LNCS) 3851*, (Hyderabad, India), pp. 898–907, Springer, Jan. 2006.
- [5] **Cs. Benedek** and T. Szirányi, „Color models of shadow detection in video scenes,” in *Proc. International Conference on Computer Vision Theory and Applications (VISAPP)*, vol. IFP/IA, (Barcelona, Spain), pp. 225–232, INSTICC, March 2007.
- [6] **Cs. Benedek** and T. Szirányi, „Markovian framework for structural change detection with application on detecting built-in changes in airborne images,” in *Proc. IASTED International*

Conference on Signal Processing, Pattern Recognition and Applications (SPPRA), (Innsbruck, Austria), pp. 68–73, ACTA, February 2007.

- [7] D. Szolgay, **Cs. Benedek**, and T. Szirányi, „Fast template matching for measuring visit frequencies of dynamic web advertisements,” in *Proc. International Conference on Computer Vision Theory and Applications (VISAPP)*, (Funchal, Madeira, Portugal), pp. 228–233, INSTICC, January 2008.
- [8] Z. Szlávik, L. Havasi, **Cs. Benedek**, and T. Szirányi, „Motion-based flexible camera registration,” in *Proc. IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*, (Como, Italy), pp. 439–444, Sept. 2005.
- [9] Z. Szlávik, T. Szirányi, L. Havasi, and **Cs. Benedek**, „Optimizing of searching co-motion point-pairs for statistical camera calibration,” in *Proc. IEEE International Conference on Image Processing*, vol. II, (Genoa, Italy), pp. 1178–1181, Sept. 2005.
- [10] Z. Szlávik, T. Szirányi, L. Havasi, and **Cs. Benedek**, „Random motion for camera calibration,” in *European Signal Processing Conference (EUSIPCO)*, (Antalya, Turkey), Sept. 2005.
- [11] L. Havasi, Z. Szlávik, **Cs. Benedek**, and T. Szirányi, „Learning human motion patterns from symmetries,” in *Proc. ICML Workshop on Machine Learning for Multimedia*, (Bonn, Germany), pp. 32–37, Aug. 2005.
- [12] L. Havasi, **Cs. Benedek**, Z. Szlávik, and T. Szirányi, „Extracting structural fragments from images showing overlapping pedestrians,” in *Proc. IASTED International Conference on Visualization, Imaging, and Image Processing (VIIP)*, (Marbella, Spain), pp. 943–948, Sept. 2004.

6.3. A szerző egyéb válogatott publikációi

- [13] **Cs. Benedek**, T. Szirányi, Z. Kato, and J. Zerubia, „A three-layer MRF model for object motion detection in airborne im-

ages,” Research Report 6208, INRIA Sophia Antipolis, France, June 2007.

- [14] **Cs. Benedek** and T. Szirányi, „A Markov random field model for foreground-background separation,” in *Proc. Joint Hungarian-Austrian Conference on Image Processing and Pattern Recognition (HACIPPR)*, (Veszprém, Hungary), May 2005.

6.4. A disszertációhoz kapcsolódó válogatott publikációk

- [15] S. Z. Li, *Markov random field modeling in computer vision*. London, UK: Springer-Verlag, 1995.
- [16] L. Havasi, Z. Szlávik, and T. Szirányi, „Detection of gait characteristics for scene registration in video surveillance system,” *IEEE Trans. Image Processing*, vol. 16, no. 2, pp. 503–510, 2007.
- [17] S. Geman and D. Geman, „Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, pp. 721–741, 1984.
- [18] L. Li and M. Leung, „Integrating intensity and texture differences for robust change detection,” *IEEE Trans. Image Processing*, vol. 11, no. 2, pp. 105–112, 2002.
- [19] I. Mikic, P. Cosman, G. Kogut, and M. M. Trivedi, „Moving shadow and object detection in traffic scenes,” in *Proc. International Conference on Pattern Recognition*, 2000.
- [20] A. Prati, I. Mikic, M. M. Trivedi, and R. Cucchiara, „Detecting moving shadows: algorithms and evaluation,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 918–923, 2003.

- [21] J. Rittscher, J. Kato, S. Joga, and A. Blake, „An HMM-based segmentation method for traffic monitoring,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1291–1296, 2002.
- [22] C. Stauffer and W. E. L. Grimson, „Learning patterns of activity using real-time tracking,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, 2000.
- [23] Y. Wang, K.-F. Loe, and J.-K. Wu, „A dynamic conditional random field model for foreground and shadow segmentation,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 279–289, 2006.
- [24] M. Irani and P. Anandan, „A unified approach to moving object detection in 2D and 3D scenes,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp. 577–589, 1998.
- [25] I. Miyagawa and K. Arakawa, „Motion and shape recovery based on iterative stabilization for modest deviation from planar motion,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 7, pp. 1176–1181, 2006.
- [26] H. Sawhney, Y. Guo, and R. Kumar, „Independent motion detection in 3D scenes,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1191–1199, 2000.