

Képi alapú többszintű környezetelemzés

MTA doktori értekezés tézisei

Dr. Benedek Csaba, PhD



Magyar Tudományos Akadémia
Számítástechnikai és Automatizálási Kutatóintézet

Budapest, 2019

1. Az értekezés tárgya, előzmények, célkitűzések

Az értekezés a gépi érzékelés különböző problémáival foglalkozik, célul tűzve ki a megfigyelt statikus vagy dinamikus környezetünk automatikus elemzését különféle képi jellegű mérések alapján. A kérdéskörhöz kapcsolódó kutatási feladatok napjainkban lényegesen túlmutatnak a hagyományos optikai leképzéssel nyert fényképeken vagy videókon működő képfelismerő eljárások tervezésén. Egy adott probléma megoldása során gyakran felhasználhatjuk korszerű szenzorok, például többkamerás rendszerek, multispektrális, mélység- és hőkamerák, radarok valamint lézerszkennerek méréseit, akár együttesen több különböző forrás fuzionálásával is [57]. Egy hatékony megoldás kidolgozása tehát már nem csak a legjobb mintafelismerő algoritmus kitalálását jelenti, hanem a feladathoz illeszkedő legjobb hardver-szoftver konfiguráció kiválasztását is. Az új mérésmodalitások megjelenésén túl számolunk kell a már létező technológiák gyors minőségbeli fejlődésével is: növekszik a szenzorok térbeli illetve időbeli felbontása, csökken a zajszintjük, illetve új spektrumcsatornák bevezetésével és többszörös jelviszszaverődés érzékelésével anyagi tulajdonságjellemzők válnak közvetlenül mérhetővé.

A szenzorfejlődés különféle alkalmazási területeken teremt lehetőséget a környezeti méréseket elmező eljárások funkcióinak kibővítésére, megbízhatóságuk növelésére. Ugyanakkor a feldolgozó algoritmusok oldaláról számos új kihívás is jelentkezik, különösképp az automatikus jellemzőkinyerés, alakzat- és eseményfelismerés, gépi tanulás, indexelés és tartalom alapú visszakeresés területein. Az értekezésben három különböző szempontot tartunk szem előtt. *Elsőként* a szenzorfüggetlenség elvét, tehát biztosítani kívánjuk, hogy a magasabb szintű funkciókat ellátó megoldásaink különféle adatforrásokból is tudjanak kezelni. Így lehetővé válik a rendelkezésre álló szenzorok méréseinek rugalmas fuzionálása, és az elavult szenzorok lecserélése meghatározott időközönként újabb, jobb minőségű adatforrásokra a létező adatfeldolgozó algoritmusaink teljes újratervezése nélkül. *Második* szempont a környezeti modell térbeli skálázhatóságának a biztosítása. A mérések növekvő térbeli felbontásának eredményeként ugyanis egy adott mérésszegmensen gyakran egyszerre tudunk elemezni különböző jelenségeket különböző skálafaktor mellett. Igény mutatkozik tehát olyan felismerő algoritmusok létrehozására, amelyek képesek a képi tartalom többszintű hierarchikus elemzésére. Példaként említve egy nagyfelbontású légi fénykép vizsgálata során makro szinten elválaszthatjuk a városi és a mezőgazdasági területeket, majd a városon belül elkülöníthetjük a különböző jellegű városrészeket, megtalálhatjuk az utakat, kinyerhetjük és osztályozhatjuk az épületeket, vagy akár irányíthatjuk a figyelmünket kisebb objektumokra is, például járművekre, vagy lámpaoszlopokra [58, 59]. *Harmadik* szempontként szintén hatékonyan kell felhasználnunk az idődimenzió különböző léptékű mintavételezéséből származtatható információt. Míg a biztonsági kamerákkal felvett videókon a mozgó alakzatok helyzete kinyerhető az egymást követő képkockák pixelszintű összehasonlításával, addig az ugyanarról a területről több hónapos vagy akár több éves időkülönbséggel készült képi mérések összehasonlítása már magasabb szintű modellezési megközelítést igényel. A kutatómunkáknak így egy olyan komplex rendszer elérését kell célul kitűzniük, ahol a különböző adatforrások által szolgáltatott információs csomagok egy egységes hierarchikus helyszínmodellbe rendezhetők, lehetővé téve az entitások multimodális reprezentációját, felismerését és összehasonlítását a megfigyelésekből és előzetes ismereteinkből származó információk együttes felhasználásával.

Funkcionális oldalról az értekezésben bemutatott új módszerek részben alacsony szintű előfeldolgozó lépéseket valósítanak meg különböző elemzési funkciók megoldásához, részben magasabb szintű objektumalapú környezetelemző modulok működéséhez járulnak hozzá. Az első esetben a modellek a szenzorok méréseiből közvetlenül kinyerhető jellemzőkön alapulnak, felhasználva például a pixelek színértékeinek térbeli vagy időbeli statisztikai eloszlását, illetve lokális képrészleteken számolt különböző textúraleírókat. Az alacsony szintű eljárások kimenete az aktuális megfigyelés osztályozása - más néven szegmentálása - ami a mért adatok szemantikus címkézéseként írható le. Egy adott helyszínről készített képpáron például elkülöníthetjük a megváltozott és a változatlan

területeket, a városokról légi Lidarral készített pontfelhőkön pedig megkülönböztethetjük a tetők és az utak régióit. Bár az osztályozás elsősorban a megfigyelésfüggő lokális képi jellemzőkön alapul, kiegészítésként felhasználunk előzetes ismereteinken és feladatfüggő feltevéseinken alapuló úgynevezett *prior* kényszereket is a különböző osztályok jellemzőtérbeli átlapolódásának feloldására, és a mérészaj hatásainak csökkentésére. Az egyik legegyszerűbb, mégis gyakran felhasznált prior feltevés a megoldásként kapott címkekép összefüggőségének a biztosítása, hiszen számos szegmentálási probléma esetén feltételezhetjük, hogy az eredménykép homogén régiókat tartalmaz, tehát a szomszédos pixelek (vagy voxelek) többségükben ugyanahhoz a szemantikus osztályhoz tartoznak.

A képi osztályozás Markov véletlen mező (MVM) alapú modellje Geman és Geman 1984-es cikkének [60] megjelenése óta alapvető valószínűségi megközelítésnek számít, ugyanakkor témakör nyitott kérdései a 2000-es és 2010-es években újra a kutatások előterébe kerültek, amikor a szenzortechnológia fejlődésével szükségessé vált különböző forrásokból származó adatok hatékonyabb fuzionálása, és különféle adatfüggő és prior következtetési szabályok együttes megvalósítása. A közelmúltban számos, részben magyar vonatkozású módszer is született a jellemzőfúzió implementálására. Például Kató és munkatársai multinomiális jellemzőeloszlásokat használtak fel a MVM-k adattagjaiban [61], majd a megközelítés kibővítéseként többrétegű MVM modelleket vezettek be [62, 63], míg Szirányi és Shadaydeh kidolgoztak egy új fúziós MVM modellt [64]. Kevesebb létező megoldást találunk ugyanakkor különféle prior információ- és szabályhalmazok hatékony modellezésére a markovi keretrendszerben. Munkám vonatkozó szakaszában ezért elsősorban ezzel az utóbbi kérdéskörrel foglalkoztam: kialakítottam egy új címkefúzió alapuló többrétegű MVM struktúrát [12], valamint a korábbi következtetési modelleket kibővítve bevezettem a dinamikus markovi gráfok alkalmazását a képszegmentációs módszertanba [13], alapul véve az eredetileg génszekvenciák modellezésére javasolt *kevert Markov modell* [65] megközelítést. Az új többrétegű modellek bevezetése mellett szükségessé vált hatékony optimalizációs algoritmusaik kidolgozása, illetve az új és a korábban létező többrétegű MVM struktúrák részletes módszertani és kísérleti összehasonlítása, amit szintén elvégeztünk [4].

Nagyfelbontású méréseken a képszegmentálási eljárások régiószintű kimenete általában nem ad elég információt a képi tartalom részletes szemantikus elemzéséhez, így szükségessé válik a geometriai tartalom és az entitások közötti kölcsönhatások közvetlen modellezése. A környezet magasabb szintű értelmezésére ad lehetőséget a mért helyszínek objektumszintű leírason alapuló vizsgálata, ami központi kérdés számos gépi látáshoz köthető alkalmazásban a távérzékeléstől kezdve, optikai gyártásellenőrzési rendszereken keresztül a videomegfigyelésig.

A szakirodalomban található objektumdetekciós módszerek a tervezési információ belső áramlásának irányától függően a direkt (alulról felfele), illetve az inverz (felülről lefele) megközelítést követhetik. A direkt megoldások [66] az objektumokat a képeken megtalált különböző primitívekből állítják össze, például foltszerű régiókból, élrészletekből, vagy sarokpontokból. Bár a direkt módszerek tipikusan gyorsan futó, tehát kis számítási igényű implementációt tesznek lehetővé, érzékenyen reagálnak olyan helyzetekre, amikor a képi zaj, vagy a tartalom összetettségének következtében a primitíveket nem tudjuk megbízhatóan kinyerni. Klasszikus megközelítésként alkalmazhatunk itt Hough transzformáció vagy matematikai morfológia alapú módszereket, azonban ezek az eljárások csak korlátozott hatékonysággal alkalmazhatók sűrű objektumpopulációk megfigyelésére, különösen amikor a képek sok egymáshoz közeli vagy egymással érintkező alakzatot tartalmaznak. Ugyancsak hiányosság, hogy a direkt eljárások - illetve ezek determinisztikus hipotézis generálás-elfogadás alapú kiterjesztései [67, 68] - különálló objektumokat keresnek a teljes objektumkonfiguráció modellezése helyett, figyelmen kívül hagyva a populáció szintjén definiálható jellemzőket, például a szomszéd objektumok átlapolódásának büntetését, relatív pozíciójuk és orientációjuk előírását, színeik hasonlóságát, vagy az objektumok között várható térbeli távolságot [69].

Az említett hátrányok jelentős részben kiküszöbölhetők az úgynevezett *inverz* módszerek alkalmazásával, melyek egy fitness függvényt definiálnak a lehetséges objektumkonfigurációk terében, majd egy optimalizációs folyamat segítségével kísérelik meg elérni a legnagyobb fitness értékkel rendelkező populációt. Ebben a megközelítésben nagy szabadsági fokkal tudunk összetett objektummegjelenési modelleket tervezni, közvetlenül építhetünk prior geometriai feltevéseken alapuló (szoft) feltételeket a környezeti modellünkbe, valamint figyelembe vehetjük a szomszédos objektumok kölcsönhatásait. Ugyanakkor az inverz megközelítés szükségessé teszi egy számításintenzív keresési művelet végrehajtását a tipikusan nagy dimenziós populációterben, ahol a többnyire konkáv fitness függvény lokális maximumai megnehezítik az optimum megtalálását.

A jelölt pontfolyamat modellek (JPM) [70] hatékony inverz módszertant nyújtanak alakzatkonfigurációk robusztus kinyerése valószínűségi leírás alapján. A MVM megközelítést kiterjesztve a JPM gráfmodelljeiben pixelek helyett geometriai objektumok a csúcspontok, figyelembe véve, hogy egy adott képen tetszőleges (és előzetesen ismeretlen) számú objektum lehet látható. A Markov mezőhöz hasonlóan itt is adatfüggő tagok és prior kényszerek építhetők be egy közös energiamodellbe, valamint hatékony elméleti algoritmusok állnak rendelkezésre paraméterbecsléshez [71, 72, 73] és energiaoptimalizációra [74, 75].

Eredményes JPM alapú eljárások születtek a közelmúltban különböző populációszámlálási feladatokra, ahol nagy számú, de alakban és méretben kis varianciát mutató objektum megtalálása a cél [81]. Jól modellezhetők így például az épületek [82], fák [83, 84], madarak [71, 72, 78], vagy csónakok [74] távérzékelte képeken, sejtmagok biológiai felvételeken [85], galaxisok űrkutatásbeli alkalmazásokban [77] vagy személyek videomegfigyelési környezetben [86]. Mivel a JPM-k globális optimalizációjának számításigénye valós idejű alkalmazások esetén problémát jelenthet, hatékony közelítő megoldások születtek az energiaminimalizáció folyamatának felgyorsítására, például a többszörös születés-halálozás algoritmus [71] vagy a párhuzamos Monte Carlo Markov lánc alapú mintavételezés (RJCMC) [73] bevezetésével.

Bár a fenti szakirodalmi módszerek a JPM alapú megoldások egyértelmű gyakorlati alkalmazhatóságát mutatják, nem foglalkoznak sem a mérések idődimenziójának modellezésével, sem a helyszínek hierarchikus térbeli struktúráival és a különböző hierarchiaszinteken lévő entitások viszonyainak leírásával. Munkám során ezért új megoldásokat adtam a klasszikus JPM keretrendszer idő- és térdimenzióval történő kiterjesztésére, részletesen elemezve a létező eljárások hiányosságait, a megoldások lehetséges irányvonalait, valamint valós problémákon szemléltettem a fejlesztések eredményeként előálló módszerek hatékonyságát.

Az értekezésben bemutatásra kerülő JPM alapú módszerekben az *idődimenzió* két különböző alkalmazási környezetben jelentkezik. *Elsőként* egy többidejű JPM struktúrát vezetek be objektumszintű változásdetekció megvalósítására képpárokon, ötvözve az alakzatkinyerés és változásfelismerés alacsony szintű megközelítéseit geometriai objektummodellezéssel. Az alakzatokat tipikusan jelentős időkülönbséggel készült képeken vizsgáljuk, és eredményként egy olyan objektumpopulációhoz jutunk, melynek tagjaihoz a változásokat jelölő címkéket is rendelünk. A *második* megoldás célja karakterisztikus geometriájú objektumok JPM alapú követése alacsony minőségű képi méréseken, alkalmazási példaként bemutatva célpontelemzés megvalósítását radarképsorozatokon. A feladat megoldásához bevezettem egy többkeretes JPM keretrendszert, amely együttesen biztosítja a megfigyelt képi jellemzők és az illesztett modellek kölcsönös megfelelőségét az egyes képkockákon, valamint a realisztikus mozgásból eredő kényszerek teljesülését az egymásutáni időkeretek között.

A következő fő vizsgálati területünk a képi tartalom *térbeli hierarchikus modellezése*. A klasszikus JPM alapú képelemző eljárások [70, 71] kizárólag egy adott részletezettségi szinten reprezentálják az objektumokat. Bár itt is definiálhatunk különböző prior interakciós kényszereket az objektum-párok között (például átfedésmentességet, vagy párhuzamos elrendeződést), a globális környezetet

leíró magas szintű strukturális információt csak nehézkesen és erősen korlátozott mértékben tudunk így figyelembe venni. Ugyanakkor számos különböző alkalmazásban lényeges feladat az objektumok csoportosulási mintáinak az elemzése, illetve az objektumok és részobjektumok együttes kezelése. Ezért bevezettem egy hierarchikus JPM struktúrát *beágyazott jelölt pontfolyamat* néven, ami az objektumok és részobjektumok hierarchiáját szülő-gyermek kapcsolatok útján írja le, és az alakzatokat különböző objektumcsoportokhoz rendeli a populáció Bayes-i szegmentálásán keresztül.

A dinamikus környezetünk gépi érzékelése és automatikus elemzése kiemelt kutatási és mérnöki terület napjainkban, különösképpen az autonóm vezetéshez, a biztonságtechnikai rendszerekhez, és az okos városokhoz köthető alkalmazások gyors fejlődésének köszönhetően. Míg a hagyományos optikai elven működő kamerák továbbra is fontos eszközei a vizuális megfigyelésnek, a közelmúltban elterjedt lézer alapú Lidar mélységszenzorok hatékony alternatívát kínálnak számos érzékelési feladat megoldásához, közvetlen méréseket szolgáltatva a háromdimenziós környezet geometriájáról. A gyűjtött 3D téradatokat kiegészítve az idő dimenziójával *4D mérés-halmazokhoz* jutunk. A ma elérhető Lidar szenzorok fontos korlátja azonban, hogy kompromisszumot kell kötnünk a mérések térbeli és időbeli felbontása között, ami megnehezíti kis méretű alakzatok és jelenségek valós időben történő észlelését és elemzését. Fontos kutatási kérdés ezért azoknak a hasznos feladatoknak a felderítése, amelyek az új típusú 4D adatok felhasználásával válnak hatékonyabban megoldhatóvá. Szintén nyitott kérdések, hogy az ismert képelemző, volumetrikus környezetleíró, és az érzékeléshez köthető gépi tanulás alapú módszerek milyen mértékben adaptálhatók a Lidar adatokhoz, valamint milyen módon tudjuk a különböző Lidar és optikai szenzorok méréseit hatékonyan fuzionálni egy magasabb szintű környezetinterpretáció eléréséhez. Az értekezésben a *4D környezetérzékelés* három választott problémacsaládjával foglalkozunk. *Elsőként* egy új Bayes-i módszert mutatunk be személyek helyzetének a meghatározására és magasságuk becslésére többkamerás környezetben. *Másodikként* egy személyek megfigyelésére alkalmazható biztonsági rendszert és ehhez tartozó algoritmikus módszertant ismertetünk forgó többsugaras (FT) Lidar szenzorok méréseire támaszkodva, kitérve a mozgásérzékelés, mozgó alakzatok elkülönítése, követése, és a járás alapú biometrikus személyazonosítás feladataira. Végül a *harmadik* problémacsoportban vázoljuk egy újszerű rendszer folyamatmodelljét, és ebbe illeszkedően bemutatunk számos új algoritmust városi környezetekben történő folyamatok érzékelésére, felhasználva egy mozgó járműre szerelt FT Lidar szenzor méréseit és egy mobil lézerszkenneléssel előzetesen elkészített részletes 3D referenciaterképet.

2. Alkalmazott eszközök és módszerek

A kidolgozott módszerek hátterét főként a statisztika, valószínűségszámítás, geometriai modellezés, matematikai optimalizálás valamint kép- és videofeldolgozás, pontfelhő alapú 3D modellezés, adatfúzió és 3D/4D gépi látás területeiről összegyűjtött eredmények adják. Az elvégzett kutatómunka fő célja a számítógépes látás és mintafelismerés területein létező funkciókönyvtárak kiterjesztése, új összetett és minél általánosabban alkalmazható megoldásokkal. Valós életből származó különböző érzékelési problémákhoz igazodva modelleket dolgoztam ki helyszínek és események időbeli és térbeli többszintű dekompozíciójára.

Az általam bevezetett módszerek többsége matematikai értelemben az előző fejezetben ismertetett *Markov véletlen mezők* [60], *kevert Markov modellek* [65], *jelölt pontfolyamat modellek* (JPM [70]), és korszerű gépi tanulási eljárások különböző megvalósításainak tekinthetők. A munkámat így kidolgozott elméleti háttérre alapozhattam, ami biztosította a megoldásaim stabilitását, és egyszerűbbé tette az új módszerek alkalmazhatóságának, érvényességi körének és limitációinak a meghatározását is.

Mivel a disszertációban bemutatott főbb tudományos hozzájárulások különböző alkalmazha-

tó modellek és algoritmusok kidolgozásához köthetők, a részletekre kiterjedő kísérleti kiértékelés kritikus lépés volt az új eredmények jelentőségének bizonyításához a különböző alkalmazási környezetekben. Ezért a kísérletek során kiemelt figyelmet fordítottam releváns tesztalmozok és referenciaadatok (ground truth) használatára, részben nyilvánosan elérhető adatbázisok, részben általunk mért illetve létrehozott adatminták felhasználásával. Az MTA SZTAKI Gépi Érzékelés Kutatólaboratóriumában rendelkezésemre álltak különböző korszerű szenzorok, többek között nagyfelbontású optikai kamerák, valamint járművekre szerelhető Velodyne HDL 64-E és VLP16 forgó többsugaras Lidar szkennerrek. A távérzékelési feladatokhoz partnereink szolgáltattak megfelelő adatokat: légi és műholdképekhez valamint légi Lidar mérésekhez jutottunk az *Astrium Defense and Space Magyarországtól*, radar (ISAR) képsorozatokhoz a *Pisai Egyetemtől*, mobil lézerszkenneléssel nyert adatokhoz a *Budapest Közút Zrt-től*, illetve további légi fotókat vásároltunk a Földmérési és Távérzékelési Intézettől (FÖMI).

A kiértékeléshez készített referenciaadatok jelentős részét kutatócsoportunk honlapján nyilvánosan elérhetővé tettük (a disszertációban felsorolt elérhetőségekkel), így lehetőséget biztosítottunk a nemzetközi tudományos közösségnek is az adatok különböző célú felhasználásaira, illetve később megjelenő módszerek összehasonlítására az általunk elért eredményekkel.

Az új eljárások implementációjához és teszteléséhez főként C/C++ nyelvű fejlesztőeszközöket használtam, míg a prototípustervezés egyes lépéseit Matlab környezetben végeztük. A kép- és pontfelhőfeldolgozási algoritmusok C++ nyelvű implementációját a publikusan elérhető OpenCV [88] és PCL [89] könyvtár-csomagok könnyítették meg. A disszertáció, a szerző kapcsolódó publikációi és számos prezentációs anyag (előadásfóliák és posztterek) $\text{\LaTeX}$ szövegszerkeztővel készültek.

3. Az új tudományos eredmények összefoglalása tézisekben

Az értekezés eredményeit négy téziscsoportba soroljuk. Az *első téziscsoportban* távérzékeléssel készült képek pixelszintű összehasonlítására mutatunk be új módszereket - két egymástól jelentősen eltérő alkalmazási környezetben - többretegű, címkefúzió alapuló Markov mezőkkel. A *második téziscsoport* tárgya képi változások objektumszintű elemzése újszerű tér-időbeli jelölt pontfolyamat modellekkel. A *harmadik téziscsoportban* hierarchikus objektumstruktúrák kinyerését célozzuk meg digitális képeken egy új beágyazott jelölt pontfolyamat modellel. Végül a *negyedik téziscsoportban* különböző környezetelemzési feladatokra adunk megoldásokat új típusú szenzorokat felhasználva, megvalósítva a videofelügyeleti alkalmazásokban kritikus személyek lokalizációja és biometrikus felismerése funkciókat, valamint dinamikus városi környezetek mozgó járművekről történő automatikus elemzését korszerű térinformatikai adatbázis háttérrel.

Az *első három téziscsoportban* ismertetett problémákon témafelelős kutatóként dolgoztam, így feladatom volt a pontos kutatási célok definiálása, az elméleti modellek kidolgozása, az irodalomkutatás, valamint az implementációs és tesztelési feladatok főbb részeinek az elvégzése. Társ szerzőim a modellek végleges kialakításához és prezentációjához járultak hozzá tapasztalataikkal valószínűségi alapú modellezés, információfúzió és képi változásdetekciós témákban, illetve a problémák felállításához adtak fontos tanácsokat alkalmazásspecifikus ismereteiket felhasználva távérzékelés, radar képalkotás és ipari gyártástechnológia területeiről. A negyedik téziscsoportban köthető eredményeket részben posztoktor munkatársaimmal, illetve az általam felügyelt doktori és egyetemi hallgatókkal közösen értem el, itt a hozzájárulások pontos megosztását az egyes altézis pontoknál részletezem.



(a) 1.1 altézis. Független objektumelmozdulások kinyerése kameramozgás kompenzációjával



(b) 1.2 altézis. Hosszútávú változásdetekció légi képeken

1. ábra. Az 1. téziscsoportban bemutatott többretegű címkefúziós eljárások eredményeinek illusztrációja

1. téziscsoport: Többretegű, címkefúzió alapuló Bayesi eljárások

Tézis: Többretegű, címkefúzió alapuló Bayesi eljárásokat javasoltam különböző időpontokban, távérzékelési technológiákkal készített képek összehasonlítására és meghatározott típusú lényeges változások automatikus észlelésére. Hatékony optimalizációs algoritmusokat vezettem be a kifejlesztett modellekhez a módosított Metropolis dinamika relaxációs módszer továbbfejlesztésével. Kísérleti úton kiértékeltem a modellek hatékonyságát valós távérzékelési alkalmazásokhoz kapcsolódóan, és korábbi módszerekkel összevetve a kapott eredményeket bemutattam az új eljárások előnyeit.

1.1. altézis: Háromretegű Markov véletlen mező alapú modellt (L^3 MRF) vezettem be a kameramozgástól független objektum-elmozdulások régióinak kinyerésére mozgó légi platformról készített nagy felbontású képpárokon. Kísérletileg megmutattam a javasolt megközelítés előnyeit tisztán 2D regisztráció alapú, valamint lokális parallaxiscsökkentést alkalmazó korábbi modellek, illetve különböző információfúziós stratégiák 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 objektumelmozdulások detekciójához olyan képpárokon, melyeket előzőleg egy automatikusan számítható globális projektív transzformáció segítségével regisztrálunk.

Az előbbieken kinyert jellemzők integrálására bevezettem egy új címkefúzió alapuló háromretegű 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.

Az eljárás kiértékelését három különböző adathalmazon végeztem el, összesen 83 valódi légi kép-pár és hozzájuk tartozó kézzel szerkesztett pixel szintű referenciamaszk felhasználásával (1(a) ábra). Részletes kvantitatív kiértékelést végeztem, amely az új megoldás előnyeit igazolta a problémára javasolt öt különböző referenciamódszerrel szemben. A fentiekén túl az új címkefúziós megközelítés jelentőségét bemutattam egy kifejezetten módszertani vizsgálaton keresztül is, amely során ugyanazokra a képi jellemzőkre alapozva összehasonlítottam a L^3 MRF modell eredményeit alternatív jellemzőfúziós és döntésfúziós sémájú modellek kimeneteivel.

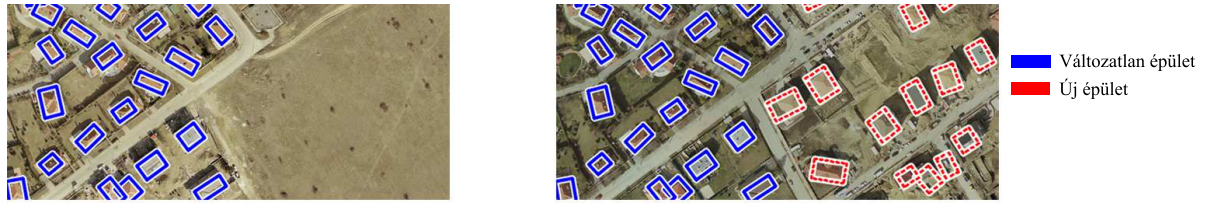
Az L^3 MRF modellt az IEEE TRANS. IMAGE PROCESSING folyóiratban publikáltuk 2009-ben [12], míg előzetes részeredményeket bemutattam korábbi publikációkban [46, 56]. Bár a modell főbb részeinek bemutatása szerepelt a Ph.D. disszertációmban is [55], a később megjelenő [12] közleményben jelentős újabb hozzájárulások szerepelnek a módszer minőségének és stabilitásának analízisével kapcsolatban, valamint a vizsgált referenciamódszerek számát és az összehasonlítások szempontjait is kibővítettem.

1.2. altézis: Új négyrétegű feltételes kevert Markov modellt (Conditional Mixed Markov model, CXM) dolgoztam ki a kevert Markov modell séma és a feltételesen független véletlen jellemzők kombinációjával, melyet lényeges változások kinyerésére használtam fel ugyanarról a területről nagy időkülönbséggel készült légi fényképek összehasonlító vizsgálata során. Kísérletileg kiértékeltem az új modellt, és megmutattam az előnyeit korábbi szakirodalmi megoldásokkal és különböző alternatív többrétegű Markov mezős megoldásokkal összehasonlítva a kapott eredményeket.

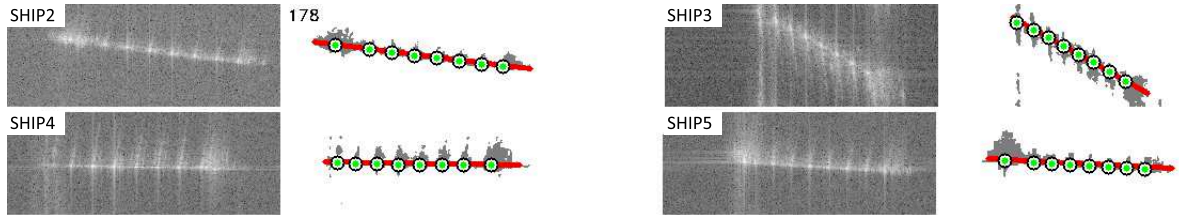
A jelentős - több hónapos vagy akár több éves - időkülönbséggel készült légi fényképek automatikus összevetése nehéz kihívásokat tartalmazó mintafelismerési feladat, mivel a változó megvilágítási körülményeknek, a növényzet szezonális változásainak és a felvételeket készítő kamerák eltérő paramétereinek köszönhetően a kinyerhető alacsony szintű képi leírók (például színérték, vagy textúra) akár a *változatlan* területeken is jelentősen eltérőek lehetnek. A különböző jellemzők alapján szegmentált rétegek címke szintű fuzionálása itt is kézenfekvő megoldást kínál a problémára, azonban kísérleteink azt mutatták, hogy a lokálisan domináns jellemzők kiválasztás folyamata ebben az esetben hatékonyabbá tehető kontextusfüggő következtetési lépések felhasználásával. Bár a statikus pixel-szomszédossági gráfokon definiált Markov véletlen mezők nem tudnak modellezni ilyen kontextusfüggő fuzionálási szabályokat, a közelmúltban bevezetett kevert Markov modellek [65] - megengedve adatfüggő kapcsolatokat a feldolgozó csomópontok között - lehetővé teszik a szükséges konfigurálható adatstruktúrák létrehozását a jellemzők integrációja során.

Új többrétegű modellstruktúrát javasoltam feltételes kevert Markov model (CXM) néven, ami az 1.1 tézisben bemutatott többrétegű címkefúziós keretmodellt kiterjeszti a kevert Markovi megközelítés alkalmazásával, létrehozva adatfüggő dinamikus kapcsolódási lehetőségeket a csomópontok között. Az egyes szegmentált rétegekhez megbízhatósági térképet rendeltem a képi jellemzők lokális statisztikai becslésére alapozva, melyeket a fúzió során felhasználtam.

Az új CXM modell hatékonyságát a légi képek hosszútávú változásdetekciós feladatán keresztül mutattam be (1(b) ábra). A szegmentálás alapjául szolgáló jellemzőként felhasználtam az összehasonlított képek együttes intenzitáseloszlását és a képek azonos helyzetű blokkjai között számított korrelációt, míg a lokális jellemzőválasztás a képek adott pixeleinek környezetében számolt kontasztertékek alapján történt. Kifejlesztettem egy szimulált lehűtésen alapuló energiaoptimalizációs eljárást, ami egyszerre biztosította az egyes rétegekhez rendelt képi megfigyelések szerinti osztályo-



(a) 2.1 altézis. Többidejű jelölt pontfolyamatmodell épületváltozások modellezésére



(b) 2.2 altézis. Többkeretes pontfolyamat modell mozgó objektumok analízisére radarképeken

2. ábra. A 2. téziscsoportban ismertetett tér-időbeli jelölt pontfolyamat modellek példa eredményei

zást, a megfelelő jellemzőválasztást, valamint az eredményül kapott változasmuszk összefüggőségét.

A kiértékelést légi- és műholdképek három különböző tesztalmazán végeztem, a Földmérési és Távérzékelési Intézet (FÖMI) valamint a Google Earth adatait felhasználva. A tesztalmaz 13 – kézzel is kiértékelte – fénykép-párt tartalmaz, melyek összesen 17 km² területet fednek le, a képek között néhány évtől két évtizedig terjedő időkülönbségekkel. A módszer hatékonyságát és előnyeit négy, a szakterület vezető folyóirataiban közölt korábbi szakirodalmi módszerrel összehasonlítva mutattuk be, valamint demonstráltuk az eljárás robusztusságát különböző modellparaméterek perturbációinak hatását vizsgálva.

A CXM módszert az IEEE TRANS. GEOSCIENCE AND REMOTE SENSING folyóirat közölte 2009-ben [13], míg részeredmények bemutatásra kerültek az ICPR 2008 konferencián [45]. 2015-ben egy áttekintő cikket közzeltünk az *ISPRS Journal of Photogrammetry and Remote Sensing* [4] folyóiratban, ahol összehasonlítottunk különböző kurrens többrétegű Markov mezős modellstruktúrákat a változásdetekciós feladatra, demonstrálva itt is a CXM több előnyös tulajdonságát. Az említett áttekintő cikk [4] 2017-ben alapjául szolgált egy vezetéssel benyújtásra került sikeres projektpályázathoz a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal „Jelentős nemzetközi hatású, kiemelkedő eredményeket elért kutatócsoportok támogatására” kiírt felhívására.

2. téziscsoport: tér-időbeli jelölt pontfolyamat modellek

Tézis: Újszerű tér-időbeli jelölt pontfolyamat modell (JPM) alapú eljárásokat javasoltam képi méreissorozatok objektumszintű elemzésére, a hagyományos JPM modellek időbeli dimenzióval történő kiegészítésével. Megoldásokat dolgoztam ki objektum alapú változásdetekció valamint mozgó célpont követése feladatokra távérzékelési alkalmazásokhoz, melyeket részletesen kiértékeltem kísérleti úton és bemutattam az új megközelítések előnyeit.

A jelölt pontfolyamat modellek (JPM) hatékony szakirodalmi megoldásokat nyújtanak objektumpopulációk számlálási feladataira. Azonban míg a korábban közölt eljárások statikus környezetet vizsgálnak, az alkalmazási oldalról igény jelentkezik napjainkban objektumszintű változások elemzésére, valamint mozgó célpontok észlelésére és követésére a JPM által nyújtott geometriai modellezés lehetőségeit és előnyeit kihasználva.

2.1. altézis: Új valószínűségi eljárást dolgoztam ki többidejű jelölt pontfolyamat (tJPM) néven objektumváltozások észlelésére, integrálva a geometriai modell alapú objektumkinyerést és az alacsony szintű változásdetekciós megközelítést egy közös keretrendszerben. A modellt épületek és épületváltozások együttes detekciójára alkalmaztam távérzékelt képeken, és az előnyeit bemutattam a feladat fellelhető szakirodalmi megoldásaihoz képest.

A beépített területek változásainak nyomon követése légi- és műholdképeken kulcsfeladat város-felügyeleti és környezetvédelmi alkalmazásokban. A feladat megoldására javasoltam egy új valószínűségi modell alapú eljárást (2(a) ábra), amely három részletében hozott módszertani újdonságot a szakirodalmi megoldásokhoz képest:

- (i) Új objektum alapú változásdetekciós megközelítést dolgoztam ki többidejű jelölt pontfolyamat modell (tJPM) néven, amely párhuzamosan használ fel az időrétegek között kinyerhető alacsony szintű változásjellemzőket, valamint az épített objektumokat leíró geometriai modelleket a képeken látható épületek automatikus megtalálására és az időbeli változásaik felismerésére.
- (ii) Mivel a létező légi- és műholdkép-adatbázisok nagy mértékben heterogén adathalmazokat tartalmaznak, modulárisan felépülő és rugalmasan tanítható keretmodellt javasoltam az objektumok leírására, amely különböző minőségű képeken megjelenő és különböző struktúrájú épülettípusokat is képes együttesen kezelni.
- (iii) Annak érdekében, hogy nagy feldolgozandó adatmennyiség mellett is biztosítani tudjuk alkalmazásainknak az optimalizáció konvergenciájára, az eredmény minőségére és a folyamat számítási igényeire vonatkozó praktikus követelmények teljesülését, a hatékony *többszörös születés és halál* optimalizációs eljárást illesztettem a változásdetekciós problémához, valamint kidolgoztam egy új inhomogén sztochasztikus objektumszületési stratégiát, amely alacsony szintű képi jellemzők alapján generál nagyobb valószínűséggel magas fitness értékű objektumjelölteket.

Az implementált tJPM modellt nyolc, jelentősen különböző légi- és műholdképeket tartalmazó adathalmazon teszteltem, összesen 662 kézzel is bejelölt épületet vizsgálva. Kvantitatív kiértékelést végeztem mind a megtalált objektumok számát, mind a geometriai illesztés pixel szintű pontosságát illetően, és bemutattam, hogy eredményesség szempontjából a módszerünk felülmúlja a létező szakirodalmi eljárásokat, míg az eljárásunk számítási időigénye is versenyképes a referenciamódszerekével. A javasolt egyesített objektum- és változásdetekciós megközelítés hatékonyságát összehasonlítottam egy kétlépéses, előbb az épületek kinyerését majd a találatok utólagos összehasonlítását végző eljárásával, ez a teszt szintén az új tJPM megoldás előnyeit igazolta. Végül kísérletekkel demonstráltam a képi jellemzőkön alapuló inhomogén születési folyamat hatékonyságát, valamint részletesen analizáltam a különböző paraméterek változtatásainak hatását.

A tJPM módszert bemutató publikáció az IEEE TRANS. PATTERN ANALYSIS AND MACHINE INTELLIGENCE folyóirat 2012 januári kiadványának kiemelt cikkeként jelent meg [10], míg kapcsolódó eredményeket különböző nemzetközi konferenciákon mutattunk be [40, 42, 43, 44].

2.2. altézis: Bevezettem egy rövidebb időbeli szekvenciák vizsgálatára alkalmas többkeretes pontfolyamat modell sémát, amely két szinten, objektumok és karakterisztikus pontok szintjén együttesen képes kinyerni és követni egy mozgó célpont geometriai paramétereit. Bemutattam a modell hatékonyságát rendkívül zajos távérzékelt adatok (radarképek) analízisére ahol a képalkotó eljárás fókuszhibái és zajszintje miatt az egyes képkockákról kinyert információ nem bizonyul megbízha-

tónak, ugyanakkor a képkockák közötti prior geometriai kapcsolatok (pozícióbecslés és alaki állandóság) felhasználásával robusztus módszer nyerhető.

Mozgó objektumok - például teherhajók - észlelése és mozgásanalízise inverz szintetikus apertúra radar (ISAR) képeken kulcsfeladat radar alapú automatikus célpontfelismerési alkalmazásokban. A távérzékeléssel nyert ISAR képek olyan helyzetekben is értékes információkat szolgáltatnak célpontok osztályozásához és azonosításához, amikor a hagyományos SAR képalkotás nem képes használni a távérzékeléssel nyert ISAR képeken a karakterisztikus célpontjellemzők robusztus kinyerése és követése tipikusan nehéz feladat a képek nagy zajszintje, és a megjelenített célpont struktúrájának alacsony részletettségé miatt.

Bevezettem egy vonalszakaszokat és pontthalmazokat együttesen kezelő *többkeretes pontfolyamat modell sémát (kJPM)*, hajók és repülőik struktúráinak automatikus észlelésére és követésére távérzékelte ISAR képsorozatokon (2(b) ábra). Robusztus megoldást dolgoztam ki az objektumok fő tengelyeinek kinyerésére, jellemző pontok detekciójára és követésére. A jelviszszaverő felületek által okozott karakterisztikus pontok fluktuációjának és az ISAR képeken jelentkező gyakori tüskezaj jelenségeknek a kezelésére az eljárás eredményeként létrejövő célpontszekvenciát iteratív sztochasztikus optimalizációval hozzuk létre, ami együttesen írja elő a megfigyelt képi adat és a becsült modellek illeszkedését, valamint különféle prior geometriai kényszerek teljesülését a célpont geometriájának egymást követő képkockákon mért paramétereinek között.

Elvégeztem a módszer kvantitatív kiértékelését nyolc különböző, teherhajókat és repülőgépeket tartalmazó valódi ISAR képszekvencián, 545 kézzel annotált képkockából álló adatbázis segítségével. Kísérletileg bemutattam, hogy zajos szekvenciákon a kJPM modell alkalmazása szignifikánsan javítja az egyes képkockákon külön-külön elvégzett detekció eredményeit.

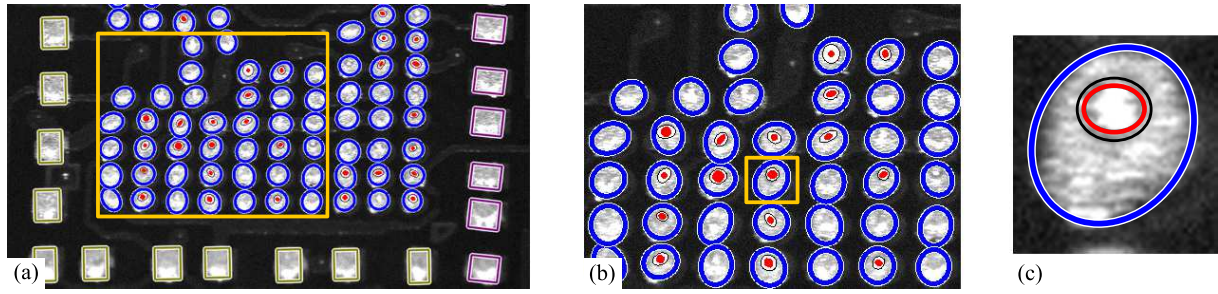
A többkeretes pontfolyamat modell megoldást az IEEE TRANS. GEOSCIENCE AND REMOTE SENSING folyóiratban publikáltuk 2014-ben [7], korábban a módszer egyes részeit az IET Radar konferencián [33] valamint az IEEE IGARSS távérzékelési konferencián mutattuk be [38].

3. téziscsoport: beágyazott jelölt pontfolyamat modell

Tézis: Bevezettem egy háromrétegű beágyazott jelölt pontfolyamat modell (BJPM) keretrendszeret összetett hierarchikus objektumstruktúrák kinyerésére különböző digitális képeken. A módszer hatékonyságát három különböző alkalmazási területen mutattam meg: lakott területek távérzékelte képeken történő analízise során, nyomtatott áramkörök hibáinak vizsgálatakor, és utcai járműforgalom felügyeleténél légi Lidar pontfelhőkön.

A korábban bevezetett jelölt pontfolyamat modell (JPM) alapú szakirodalmi módszerek a képi tartalom objektumszintű elemzését célozzák meg, azonban kevésbé alkalmasak hierarchikus mintafelismerési problémák megoldására. Ennek a hiányosságnak a kezelésére létrehoztam egy új beágyazott jelölt pontfolyamat modell (BJPM) keretrendszert, amelyben a klasszikus JPM módszereket két fő ponton kiegészítve a populáció tartalmazhat objektum-részobjektum struktúrákat, valamint az összetartozó objektumok különböző objektumcsoportokhoz csatlakozhatnak. A fenti két módszertani kiterjesztés motivációját konkrét alkalmazásokból merítettem nyomtatott áramkörök elemzése (3.1 tézispont) és távérzékelés alapú forgalomfelügyeleti problémák [5] vizsgálata során. A feladat-specifikus kutatómunkák elvégzését követően kidolgoztam és implementáltam egy általános háromrétegű modell keretrendszert (3.2 tézispont), amit végül három, jelentősen különböző alkalmazási területeken teszteltem és validáltam.

3.1. altézis: Kidolgoztam egy Bayes-i alapú automatizált képi ellenőrző eljárást nyomtatott áramkörök minőségvizsgálatához, ami képes egyszerre felismerni kü-



3. ábra. A 3. téziscsoportban ismertetett beágyazott jelölt pontfolyamat modell demonstrációja áramköri elemek optikai vizsgálata

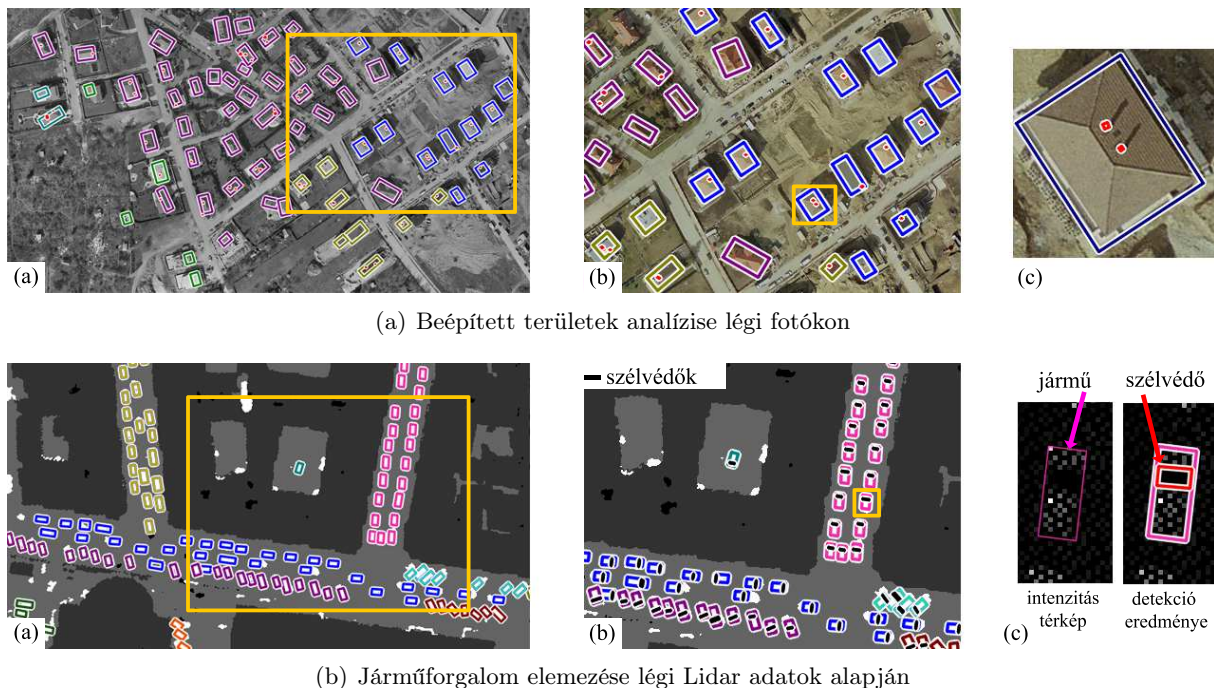
lőnböző alakú és méretű áramköri elemeket és azok hibás régióit, a JPM alapú sémát kiterjesztve objektum-részobjektum kapcsolatok modellezésével. A megközelítés hatékonyságát speciális hegesztési eljárással létrehozott áramköri elemek üregesedéseinek észlelését és területük becslét végző alkalmazáson keresztül mutattam be, amely fontos feladat a nyomtatott áramkörök minőségének biztosításához. Kidolgoztam egy Bayes-i alapú automatizált képi ellenőrző eljárást nyomtatott áramkörök minőségvizsgálatához, ami képes egyszerre felismerni különböző alakú és méretű áramköri elemeket és azok hibás régióit, a JPM alapsémát kiterjesztve objektum-részobjektum kapcsolatok modellezésével. A megközelítés hatékonyságát speciális hegesztési eljárással létrehozott áramköri elemek üregesedéseinek észlelését és területük becslét végző alkalmazáson keresztül mutattam be, amely fontos feladat a nyomtatott áramkörök megfelelő minőségének biztosításához.

Az automatikus optikai áramkörellenőrző technológiák rendkívül nagy felbontású ($10\mu\text{m}$) képeken dolgoznak, így a felismerő eljárásokban szükség van az áramköri struktúra hierarchikus modellezésére, kinyerve az áramköri régiókat és az egyes áramköri elemeket (ÁE), miközben modellezzük az ÁE-k elrendezését meghatározó interakciókat, és felismerünk jellegzetes mintázatokat az ÁE-ken belül, például a geometriailag modellezhető üregesedési hibákat (3. ábra).

Optikai ellenőrző eljárást készítettem nyomtatott áramkörök számára, amely az áramköri elemek és az üregesedési hibáik összetett modelljét a jelölt pontfolyamat keretrendszerbe ágyazott hierarchikus kapcsolati relációval írja le. Figyelembe véve eltérő alakú áramköri elemek együttes előfordulását egy adott áramköri lapon, a hierarchikus alakzatkinyerésre alkalmas modellben különböző geometriai objektumok detekciójára adunk egyszerre lehetőséget. Mivel ezek a módszerki-terjesztések a JPM megoldási tér dimenzióját lényegesen megnövelik, fontos hogy az ÁE populáció terét hatékonyan mintavételezzük. Ezért a 2.1. tézispontnál bemutatott képi jellemző alapú inhomogén születési folyamat megközelítést továbbfejlesztettem, és kidolgoztam egy alulról építkező (bottom-up) sztochasztikus objektumgenerálási stratégiát alacsony szintű statisztikai képleírók és prior struktúrajellemzők kombinálásával, ami az optimalizáció számítási komplexitását praktikus alkalmazásokban is kivitelezhető mértékűre szorította le.

A javasolt megoldás hatékonyságát áramköri ellenőrzéshez kapcsolódó valós adathalmazon értékeltük ki, amely 125 áramkör képét, és ezen belül több mint 10000 áramköri elemet tartalmaz. Az eredmény minőségét egy hagyományos morfológiai megoldással vetettük össze, és az új megoldás előnyeit egyértelműen kimutattuk.

A javasolt képi mintafelismerő eljárás központi elemeit egyszerűs cikkben publikáltam a *Pattern Recognition Letters* [11] folyóiratban (2011), és bemutattam az IEEE ICIP 2011 konferencián [37]. A teljes módszert az IEEE TRANS. INDUSTRIAL ELECTRONICS folyóiratban ismertettük



4. ábra. A 3. téziscsoportban ismertetett beágyazott jelölt pontfolyamat modell demonstrációja két távérzékelési alkalmazásra

2013-ban [8], ahol a választott optikai felismerési feladat relevanciáját és technológiai háttérét tár-szerzőim, a Budapesti Műszaki és Gazdaságtudományi Egyetem elektronikai technológiai szakértői biztosították.

3.2. altézis: Definiáltam az általános háromrétegű beágyazott jelölt pontfolyamat modellt (BJPM) és egy ehhez illeszkedő sztochasztikus energiaoptimalizációs algoritmust, amely egyszerre képes objektumcsoportok, objektumok és objektumrészek kinyerésére nagyfelbontású digitális képekről. Rugalmas tervezési lehetőségeket biztosítottam a modellben az adatfüggő és prior kényszerek felállításakor. Bemutattam, hogy a BJPM megközelítés számos különböző valós életbeli hierarchikus mintafelismerési probléma megoldásához illeszthető, és a modell eredményességét három különböző alkalmazási példán keresztül igazoltam.

Nagyfelbontású képek többszintű tartalmi elemzésére különböző szakirodalmi eljárások szület-tek, melyek az általam is ismertetett objektum alapú JPM megoldás [5, 8] mellett, régió alapú [58], vagy hibrid [59] megközelítéseket követtek. Ezeket a modelleket specifikus alkalmazási környezetbe, és meghatározott bemeneti adatmodalitásokhoz tervezték, például távérzékelte optikai képek [58, 59] és Lidar pontfelhők [5] elemzéséhez, vagy - a 3.1 altézisben bemutatott példában - nyomtatott áramkörök automatikus optikai vizsgálatára μm felbontású képeken [8]. A gyakorlati tapasztalatok ugyanakkor azt mutatják, hogy időigényes és nehéz feladat ezeket az összetett modelleket az eredeti-től eltérő alkalmazási környezethez igazítani, ami többnyire jelentős modellezői és implementációs munkát is igényel.

Munkám során ezért összegyűjtöttem hasonló problémákat a korábban vizsgált alkalmazási területeimről, majd kidolgoztam egy általános beágyazott jelölt pontfolyamat modell struktúrát (BJPM), amely konkrét alkalmazások széles köréhez illeszthető. A modell strukturális elemeit és a hierarchikusan felépülő energiafüggvény optimalizációs algoritmusát absztrakt szinten definiáltam és

implementáltam, míg a különböző alkalmazások felé egyszerű interfészeket biztosítottam, lehetővé téve a célfeladat rugalmas definícióját és az általánosítást újabb felmerülő problémákra.

A BJPM modell két fő újonságot vezet be a hagyományos *egyszintű* pontfolyamat megoldásokkal szemben:

- (i) Az objektumok és objektumrészek hierarchikus kapcsolatát a JPM keretrendszerbe beágyazott szülő-gyermek relációval írjuk le. A gyermek megjelenési modelljét közvetlenül befolyásolja a szülő entitás, például geometriai és radiometriai megjelenés alapú kényszereken keresztül.
- (ii) A (szülő) objektumok populációját partícionáljuk, létrehozva objektumcsoportokat, más néven konfigurációszegegmenteket. A szekvenciális megközelítésekkel szemben modellünkben az objektumokat egyidejűleg nyerjük ki az optimális szegementekkel egy közös energiaminimalizációs eljárás segítségével. A szegementenként különbözőképpen definiálható interakciók megengedésével adaptív objektumszomszédosságot hozunk létre.

A javasolt BJPM módszertan eredményességét három különböző alkalmazási környezetben mutattam be (3. és 4. ábrák): beépített területek analízisére távérzékeléssel nyert képeken, járműforgalom elemzésére légi Lidar adatok alapján, valamint a 3.1 altézisben ismertetett modellt háromrétegűre kiterjesztve áramköri elemek optikai vizsgálatára. Az alkalmazások sikerességének demonstrációján túl részletes módszertani kiértékelést is végeztem. Kvantitatív tesztteredményekkel támasztottam alá a BJPM eljárás előnyeit egy hagyományosnak tekinthető szekvenciális megközelítéssel szemben, ahol az objektumpopulációt előbb egy egyrétegű JPM segítségével nyerjük ki, majd az objektumok csoportosítása utófeldolgozásban, egy elárasztásos eljárás alapján osztályozással történik. Ezen túl kísérleti úton mutattam meg a sztochasztikus optimalizáció végeredményének nagyfokú megismételhetőségét, valamint a módszer hatékonyságát a futási idő szempontjából is.

Az általános BJPM keretrendszert egyszerűs cikkben közöltem az IEEE TRANS. IMAGE PROCESSING folyóiratban 2017-ben [2], a modell különböző elemeit az IEEE ICASSP 2014 [25] és az ICIAR 2013 [28] konferenciákon mutattam be.

Korábban Börcs Attila doktoranduszommal egy kétrétegű alkalmazás-specifikus változatát is elkészítettük a modellnek L^2 JPM néven, amit járművek, és forgalmi szempontokból összetartozó járműcsoportok együttes kinyerésére használtunk. A L^2 JPM modellt az IEEE TRANS. GEOSCIENCE AND REMOTE SENSING folyóiratban jelentettük meg 2015-ben [5], míg egyes részeit az ICPR 2012 [32] és az ISPRS Congress 2012 [34] konferenciákon is bemutattuk. A L^2 JPM eljárás kidolgozása lényeges hozzájárulásokat tartalmazott Börcs Attila részéről is [87], különösképpen a pontfelhőszegementálás és a járműveket leíró jellemző alapú adatmodell megszerkesztésével kapcsolatban, míg magam itt elsősorban a prior adatmodell és az optimalizációs algoritmus kidolgozásához járultam hozzá.

4. téziscsoport: 4D környezetanalízis

Tézis: Új modelleket és algoritmusokat javasoltam videomegfigyelés és környezetérzékelés területén felmerülő különböző problémák megoldására, 4D téridőbeli méréseket szolgáltató korszerű szenortechnológiák felhasználásával. Kísérleti úton kiértékeltem a bevezetett új algoritmusok hatékonyságát valós méréseket tartalmazó reprezentatív adathalmazokon, és korábbi módszerekkel összevetve a kapott eredményeket bemutattam az új eljárások előnyeit.

A klasszikus mintafelismerési problémák, mint alakzatok lokalizációja, osztályozása, mozgáskövetés és változásdetektáció fontos megoldandó feladatként jelennek meg napjainkban is az intelligens környezetérzékelés különböző alkalmazásaiban, például biztonsági videofelügyeletnél, autonóm vezetésnél, vagy városmenedzsmentet érintő vizuális feladatokban. Ugyanakkor ezek az alkalmazások



5. ábra. A 4.1 altézisben bemutatott 3D jelölt pontfolyamatmodell eredményei egy választott kamera képére vetítve

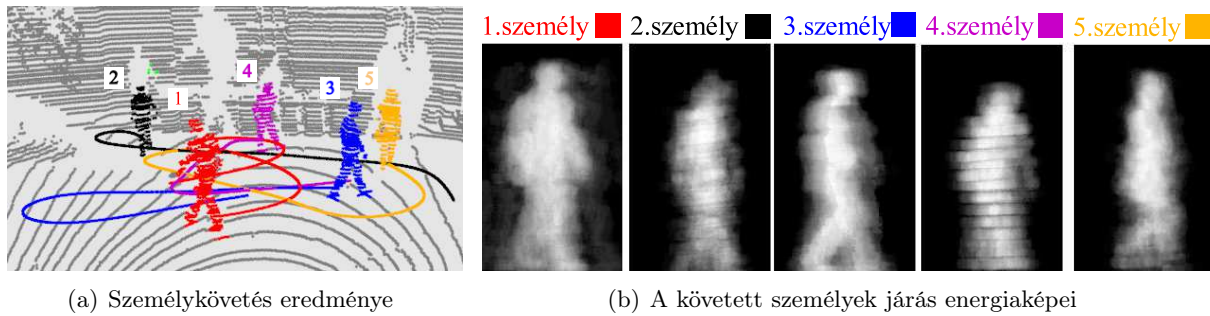
számos nehéz felismerési probléma megoldását is megkövetelik, különösen zsúfolt kültéri helyszíneken a változó megvilágítási tényezők, folyamatos háttérzaj és a helyszín mozgó és statikus objektumai között fennálló takarási jelenségek következtében. 3D/4D szenzorok méréseinek bevonása lényeges segítséget nyújthat a feladatok megoldásához, ugyanakkor az újfajta szenzorok használata számos új kihívást is támaszt a gépi látás algoritmusai számára. Eddigi munkám során három különböző problémakörrel foglalkoztam ezen a tématerületen, különböző és részben új típusú mérési hardver konfigurációkra támaszkodva.

4.1. altézis: Bevezettem egy henger objektumokat kezelő jelölt pontfolyamat modellt (JPM) többkamerás videofelügyeleti helyszíneken megfigyelhető, részben egymást takaró és egymással érintkező személyek csoportjainak jellemzésére. A kalibrált kamerák képeiből kinyerhető képi jellemzőket felhasználva megmutattuk, hogy a javasolt megközelítés hatékonyan alkalmazható egyszerre emberek 3D lokalizációjára és magasságuk becslésére, egy kiemelt szakirodalmi módszert meghaladó eredményességgel.

A kutatómunka során kidolgoztunk egy új valószínűségi alapú módszert személyek lokalizálására többkamerás környezetben. Első lépésként a társzerzőm által javasolt pixel szintű jellemzőket nyertünk ki a kamerarendszer méréseiből, amelyek a 2D optikai képalkotás fizikai modelljén alapulnak, információt szolgáltatva a személyek fejének és talpának a helyzetéről, megkülönböztetve álló és sétáló embereket is. Ezt követően a különböző kameranézetekből származtatott jellemzőket fuzionáltuk, becslést adva az egyes emberek talajsíkon meghatározott pozíciójára és a magasságára. Az így nyert jellemzők olyan esetben is használhatónak bizonyultak, amikor a mozgások észlelését csupán a helyszín egy meghatározott részletén kellett jelezniük, illetve a vetített ember-sziluettek és a háttérben történő zajszerű lényegtelen mozgásokból adódó környezeti változások területei jelentősen átlapolódtak.

Az előzőekben bemutatott jellemzőket felhasználva definiáltam egy háromdimenziós objektumkonfigurációs modellt a helyszín euklideszi koordináta-rendszerében, figyelembe véve prior geometriai kényszereket, és személyek közötti lehetséges interakciókra. Az emberek pozíciójának becsléséhez egy jelölt pontfolyamat modell realizációjaként 3D henger objektumok populációját használtam. A végső konfigurációt iteratív sztochasztikus energiaoptimalizációs eljárással nyertük ki.

A javasolt megoldást két publikusan elérhető tesztalmazon értékeltük ki (5. ábra), és összehasonlítottuk egy aktuális szakirodalmi referenciamódszerrel. A helyes számszerű kiértékeléshez referencia adatokat készítettünk a gyalogosok valódi pozíciójához, egy általunk erre a célra fejlesztett 3D annotáló szoftver segítségével [35]. Két eltérő hibametrika és különböző paraméterbeállítások



6. ábra. A 4.2 altézisben bemutatott Lidar alapú felügyeleti rendszer (a) személykövetés funkciója követés és (b) a kinyert biometrikus járásjellemzők.

mellett összegeztük a teszteredményeket, ami az új módszerünk jobb teljesítményét igazolta.

Az eljárást 2013-ban publikáltuk az IEEE TRANS. CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY folyóiratban [9], illetve bemutattuk a CVPR 2011 [39] élvonalbeli számítógépes látás konferencián (cikkelfogadási arány 26%), valamint a téma dedikált workshop-jain [35, 36, 41].

4.2. altézis: Új videofelügyeleti eljárásokat dolgoztam ki – megvalósítva mozgásdetekciót, személyek követését és meghatározott eseményfelismerési funkciókat – egyetlen forgó többsugaras (FT) Lidar szenzor méréseire alapozva, amely a dinamikus környezet egy rögzített térbeli pozícióból monitorozza. A javasolt megoldásokat kvantitatív módon kiértékeltek valós biztonságtechnikai környezetekben, bemutatva az új eljárások előnyeit a szakirodalmi módszerekhez képest.

Míg a hagyományos optikai illetve mélységszenzorok korlátozott látótérrel rendelkeznek, a forgó többsugaras (FT) Lidar - biztonsági kameraként használva - a teljes 360°-os látóteret leképezi, a szenzorok száma által meghatározott függőleges irányú felbontással, míg a vízszintes irányú felbontás a forgás sebességétől függően változtatható. Ugyanakkor a technológia inhomogén pontsűrűséget, jelentős zajjal terhelt méréseket, valamint a szekvenciális szkennelésből adódóan a rögzített pontfelhőben a mozgással összefüggő hibákat okoz, amelyek kezelése új algoritmikus megoldásokat igényel. A problémakörhöz a következőkben felsorolt főbb hozzájárulásokat tettem.

- (i) Hibrid 2D–3D módszert adtam az előtér és háttér régióinak elkülönítésére rögzített pozícióba helyezett FT Lidar szenzor mérőeszközatain. A javasolt megoldás a számítási igény szempontjából kritikus térbeli szűrést a 3D pontfelhőből származtatott 2D mélységképen oldja meg egy Markov véletlen mezős (MVM) megközelítéssel, ugyanakkor a pixelrácsra történő projekció kvantálási hibáiból adódó bizonytalanságot a valódi 3D pontpozíciók és a 2D címkék visszavetítésének együttes figyelembevételével oldja fel. A korábban a Ph.D. disszertációmban [55] is bemutatott térbeli előtérmodell [14, 15] mélységképekre történő kiterjesztésével szignifikánsan csökkentettem a lényegtelen háttérmozgások által okozott hibákat, melyeket elsősorban a mozgó fakoronák és bokrok régióiban eredményeztek hamis találatokat. A Lidarral rögzített jelenetek feldolgozásának kvantitatív pontszintű kiértékelésére kifejlesztettünk egy 3D pontfelhő annotáló alkalmazást, aminek a segítségével kézzel is kiértékeltek a tesztjeleneteket. A javasolt előtérkinyerő modult különböző referenciamódszerekkel hasonlítottuk össze, és az eredményességét demonstráltuk különböző biztonsági felügyeleti és forgalmi jeleneteken.
- (ii) Valós idejű módszert javasoltam mozgó személyek észlelésére és követésére FT Lidar méréseken több embert tartalmazó videofelügyeleti jelenetekben, a különböző időkereteken észlelt

alakzatok rövid távú és hosszú távú összerendelésével. Az egyes időkereteken belül a különböző személyekhez tartozó pontfelhőszegmenseket elkülönítettük egymástól, majd a rövid távú összerendelés (RTÖ) során az összetartozó személyek középpontjait párosítottuk az egymást követő kereteken keresztül. A hosszútávú összerendelés (HTÖ) az RTÖ hibái miatt megszakadt útpályák megfelelő összekötéséért felelős, az ismételten megjelenő személyek különböző egyedi jellemzők alapján történő újra-felismerésével.

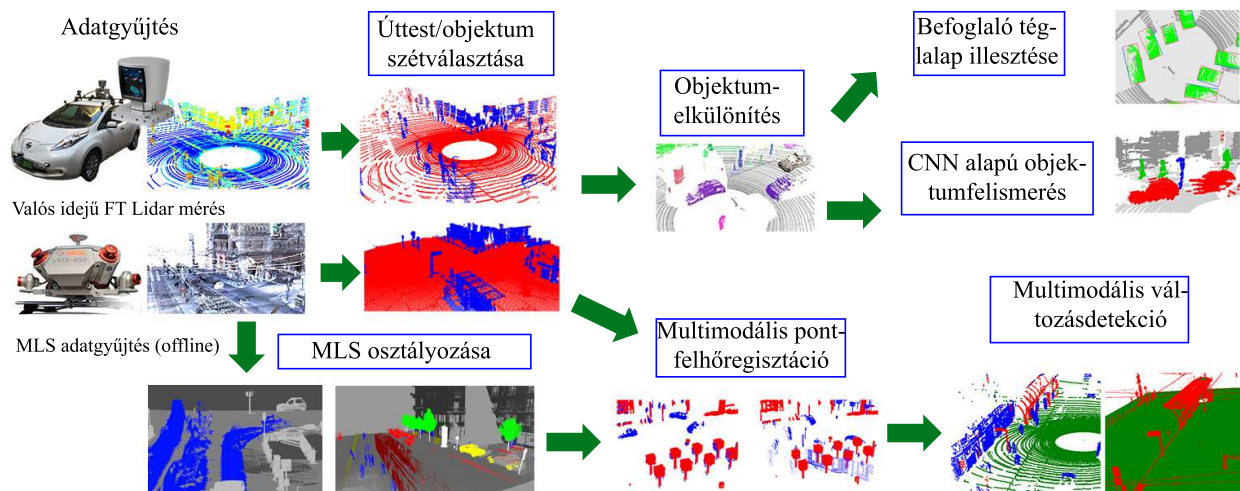
- (iii) Javaslatot tettem járás alapú biometrikus jellemzők használatára FT Lidar méréssorozatokon, a mozgó személyek követése során fent ismertetett hosszú távú összerendelés (HTÖ) lépés támogatására (6. ábra). Társzerzőimmel együttműködésben kinyertünk, teszteltünk és kvantitatívan összehasonlítottunk különböző alternatív képi leírókat, és bemutattuk az általunk javasolt Lidar alapú járásenergiakép (LJEK) megoldás hatékonyságát, előnyben részesítve a többi jellemzővel szemben. Szintén javaslatot tettünk a LJEK leíró kiterjesztésére különböző tipikus cselekvésminták felismerésére (például lehajolás, karóra leolvasása vagy integetés), amit sikeresen implementáltunk és kiértékelünk a kutatómunka során.

A Lidar alapú biztonsági felügyeleti megoldásunk teljes munkafolyamatát a IEEE TRANS. CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY folyóiratban publikáltuk 2018-ban [1], fő hangsúlyt helyezve a személyek újra-felismerése és a cselekvésfelismerés funkciókban elért eredményeinkre. A személyek háromdimenziós megfigyelésének általam javasolt megközelítését elsőként egy egyszerű cikkben ismertettem a *Pattern Recognition Letters* folyóiratban 2014-ben [6]. Az előtér-háttér szegmentációs algoritmust a WDIA 2012 ICPR Workshop-on mutattam be először [31], míg a járásfelismeréssel kapcsolatos részeredményeinket a VISAPP 2017 [19], IWCIM 2015 [21] és EUSIPCO 2015 [22] konferenciákon is ismertettük.

4.3. altézis: Kidolgoztam egy új eljárásorozat folyamatmodelljét dinamikus városi környezet analízisére mozgó járművekről rögzített FT Lidar pontfelhősorozatok alapján, referenciaként felhasználva mobil lézerszkenneléssel (MLS) a helyszínről előzetesen rögzített sűrű pontfelhőket. Hozzájárulásokat tettem új algoritmusok kifejlesztéséhez és kiértékeléséhez objektumdetekció, osztályozás, a háttér szeman-tikus szegmentálása, multimodális pontfelhőregisztráció és változásdetekció feladattaira.

A lézerszkenneléses technológiák nagy precizitású háromdimenziós méréseket szolgáltatnak a környezetünkről. Ugyanakkor a napjainkban elérhető szenzorokat használva egyelőre kompromisszumot kell kötnünk a rögzített pontfelhő-sorozatok térbeli vagy időbeli felbontásával kapcsolatban. A járművekre szerelhető forgó többsugaras Lidar kamerák másodpercenként 15-30 időkeretes rögzítési sebességgel képesek pontfelhőket gyűjteni, ami lehetővé teszi dinamikus események analízisét, ugyanakkor a 4.2 altézisben is leírtak szerint a mérések térbeli felbontása alacsony és erősen inhomogén. Másrészről a mobil lézerszkennelő (MLS) rendszerek nagy felbontású statikus 3D pontfelhőket szolgáltathatnak nagy méretű városrészekről is, melyek nagy pontosságúak, viszont a valóban a helyszínhez tartozó statikus elemek mellett (például út, épülethomlokzatok, lámpaoszlopok) tartalmaznak mobilis (pl. parkoló autók) és időben változó (növényzet) objektumokat, valamint úgynevezett fantomrégiókat, amelyeket a szkennelést végző járművel együtt mozgó utcai alakzatok (mozgó járművek, gyalogosok) okoznak.

Kidolgoztam egy eljárásfolyamatot a járművek érzékelő platformjára szerelt FT Lidar szenzorok méréseinek és az előzetesen tárolt georeferált MLS pontfelhőket tartalmazó térbeli adatbázisok együttes felhasználásával instant környezetérzékelési és offline városfelügyeleti funkciók támogatásához (7. ábra). Ph.D. és egyetemi hallgatóimmal közösen új algoritmikus eszköztár bevezetésén



7. ábra. A 4.3 altézisben ismertetett folyamatmodell dinamikus városi környezet analízisére valós idejű FT Lidar pontfelhősorozatok és mobil lézerszkenneléssel nyert háttérmodell együttes felhasználásával

dolgoztunk, amely lehetővé teszi a térinformatikai (GIS) rendszerekben tárolt információk valós idejű felhasználását önjáró autók döntéstámogatásához, és lehetőséget kínál a meglévő GIS adatbázisok frissítésére a napi forgalomban közlekedő járművek méréseit figyelembe véve. A munka során több új modellt és algoritmust javasoltunk, melyeket a következőkben röviden ismertetek. Valós időben működő eljárást dolgoztunk ki az úttest és az akadályok elkülönítésére FT Lidar pontfelhőkön, megkülönböztetve az *alacsony előtér* (alacsony utcai objektumok régiói) és magas előtér (épületek homlokzatai és magas objektumok) szegmenseket, felkészülve az FT Lidar keretek inhomogén pontsűrűségének kezelésére, egyenetlen útfelkszínre, valamint a városi sűrű forgalmi helyzetekre [30]. A szegmentációra alapozva elkészítettünk egy kétrétegű rácson alapuló gyors objektumkinyerő algoritmust az alacsony illetve magas előtér régiókra, amely lehetővé teszi közeli objektumok valós idejű szétválasztását is [24], valamint 2D felülnézeti befoglaló téglalapok hatékony illesztését strukturális kényszerek alkalmazásával [23]. A következő lépésben az alacsony előtér objektumait négy szemantikus osztály egyikébe soroljuk – jármű, gyalogos, utcai háttérobjektum és alacsony fal – a pontfelhőszegmensek mélységkép reprezentációján alapuló (tehát 2D képi bemenettel dolgozó) konvolúciós neurális hálózat (CNN) segítségével [3].

A fentiekén túl javasoltunk egy 3D voxelmodell alapú CNN hálózatot MLS pontfelhők szemantikus szegmentációjára [18], ami lehetővé teszi a fantom régiók és a mozdítható objektumok eltüntetését a nyers MLS mérésekről, megjelölve a statikus – így tájékozódási pontnak is használható – utcai alakzatokat, amelyeket a járművek pontos pozíciójának és orientációjának a meghatározásához használunk a nagyfelbontású 3D térképen. Ezt követően multimodális pontfelhőregisztrációs algoritmust dolgoztunk ki, amelynek segítségével meghatározható az aktuális FT Lidar mérések pontos pozíciója és orientációja az MLS térkép háromdimenziós globális koordináta rendszerében [16, 20]. Végül egy Markov véletlen mező alapú változásdetekciós eljárást vezettünk be az előzőekben regisztrált multimodális pontfelhők összehasonlítására [17].

Az FT Lidar alapú objektumdetekciós és -osztályozó megoldásunkat az IEEE GEOSCIENCE AND REMOTE SENSING LETTERS [3] folyóiratban közzétettük, míg a további eredményeket különböző konferenciákon mutattuk be: az ISPRS VMC 2013 Workshop-on [30], az ECCV 2014 [24] ACCV 2014 [23] és ECCV 2018 [16] konferenciák autonóm vezetéssel foglalkozó Workshop-jain, valamint az ICPR 2016 [20], IJCNN 2017 [18] és ICIAR 2017 [17] IEEE, illetve LNCS kiadványokat közlő

konferenciákon.

Az ehhez az altézishez kapcsolódó munkát vezető kutatóként koordináltam, meghatározva célokat és a munkafolyamat lépéseit, hozzájárulva az új algoritmusok kidolgozásához, valamint felügyelve a kapcsolódó kutatási-fejlesztési tevékenységet. A specifikus technikai hozzájárulásokat megosztva hoztuk létre a felügyeletem alatt dolgozó *Ph.D.* és egyetemi hallgatókkal, kiemelve közülük társszerzőimet, Börcs Attilát (*Ph.D.*, 2018, [87]), Nagy Balázst (jelenleg *Ph.D.* hallgatóm), Gálai Bencét (*M.Sc.*, 2017) és Józsa Oszkárt (*B.Sc.*, 2013).

4. Az eredmények alkalmazási területei

A disszertációban szereplő új modellek és eljárások felhasználhatók lehetnek különböző korszerű, gépi látásfunkciókkal rendelkező rendszerekben, például környezetfelügyelet, biztonságtechnika, távérzékelés, ipari minőségellenőrzés, filmelőkészítés, robotika és autonóm vezetés alkalmazásaiban.

A javasolt eljárások közül több is közvetlenül kapcsolódik az MTA SZTAKI részvételével az elmúlt években folyó kutatási projektekhez. Tudományos hozzájárulásaim elsősorban a következőkben felsorolásra kerülő projekteket érintették, melyekben koordinátorként, vezető kutatóként, illetve (hazai) projektvezetőként vettem részt.

- (i) Az Európai Védelmi Ügynökség (EDA) támogatásával megvalósított *Array Passive ISAR adaptive processing* (APIS) projekt egy új nagy felbontású passzív radarrendszer kifejlesztését és funkcionalitásainak demonstrálását végezte. Az APIS tudomásunk szerint a világon az első olyan passzív rendszert valósította meg, amely az inverz szintetikus apertúra radar (ISAR) technológia felhasználásával képes a megfigyelt célpontok elemzésére alkalmas képeket generálni. A projektben az MTA SZTAKI munkáját koordináltam hazai témavezetőként, valamint kutatóként részt vettem a szakmai munkában különböző képfeldolgozási és mintafelismerési feladatok megoldásával. Mozgó célpontok ISAR képsorozatokon történő analizisével kapcsolatos hozzájárulásaim [7] (részletesen ismertetve a 2.2 altézisben) lényeges szerepet kaptak projektzáró dokumentumban és a projekt bemutatásra került demonstrátorában.
- (ii) Koordinátora voltam az MTA SZTAKI integrált 4D (i4D) című kutatási projektjének, amely során egy újszerű hardver-szoftver környezetet fejlesztettünk ki egy forgó többsugaras (FT) Lidar és egy 4D rekonstrukciós stúdió, mint két lényegesen különböző téridőbeli információforrás együttes felhasználásával. A kétféle adattípus integrálásának fő tudományos célja a környezetünk különböző szinteken történő vizuális mérése és reprezentációja volt. Megközelítésünkben a Lidar szenzor dinamikus kültéri helyszínek globális leírását adta időben változó pontfelhősorozatok formájában. A mozgó alakzatokat leválasztottuk a statikus háttértől, és elkészítettük a környezet háromdimenziós vektormodelljét. A 4D stúdióban részletgazdag dinamikus modelleket készítettünk mozgó szereplőkről, tipikusan személyekről. A két adatfajta újszerű együttes megjelenítésével képesek lettünk a helyszínmodell megváltoztatására, valamint a stúdióban készült mozgó avatárokkal történő benépesítésére. A projekt során különböző alkalmazási területeket céloztunk meg, például 4D városrekonstrukciót, városi környezet köztereinek felügyeletét, videomegfigyelést, kiterjesztett valóság-alkalmazásokat és telekommunikációt. Az i4D rendszer főbb tudományos és technikai újításait egy nemzetközi szabadalommal írtuk le [47], valamint referált cikkekben ismertettük az IEEE Coginfocom 2013 [26], IEEE CBMI 2013 [29] és ICVS (LNCS) 2013 [27] konferenciákon. Az i4D technológiát felhasználva kifejlesztettük egy 4D filmelőkészítő rendszer prototípusát, amit az FMX 2017 vizuális effektek témájú konferencián és kiállításon mutattunk be, és megoldásunkat beválogatták a „100 legérdekesebb magyar innováció” 2018-as kiadványába.

- (iii) Projektvezetője és résztvevő kutatója voltam az Európai Űrügynökség által finanszírozott DUSIREF (Dynamic Urban Scene Interpretation and REconstruction through remotely sensed data Fusion) projektnek. A munka fő célja városi környezet magas szintű analízise és a megfigyelhető időbeli változások kategorizálása volt heterogén távérzékelte adatokon, főként optikai és TerraSAR (radar) műholdfelvételeken és légi LIDAR méréseken. Új automatikus felismerési és vizualizációs módszereket fejlesztettünk ki négydimenziós mérések reprezentációjára, kezelve multimodális, többféle részletettségű szintű, idősorozat alapú adathalmazokat. A projektben felhasználtuk többek között az 1.2, 2.1 és 3.2 altézisekben ismertetett különféle szakmai hozzájárulásaimat.
- (iv) A jelen disszertáció benyújtásának idején vezető kutatója vagyok a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal két felfedező kutatási projektjének *Instant környezetmegfigyelés mobil járművekről újgenerációs térinformatikai adatbázis háttérrel* (K-16 kutatói kezdeményezésű témapályázat, 2016-2020) valamint *Változásdetekció és eseményfelismerés képi és Lidar mérések fúziójával* (KH-17, jelentős nemzetközi hatású, kiemelkedő eredményeket elért kutatócsoportok támogatására kiírt pályázat, 2017-2019) címmel, melyek tématerülete jelentős átfedésben van a 4.2 és 4.3 altézisekben bemutatott kutatómunkával.

További tudományos eredményeim kerültek alkalmazásra az MTA SZTAKI, a Pázmány Péter Katolikus Egyetem (PPKE) illetve a Budapesti Műszaki és Gazdaságtudományi Egyetem (BME) részvételével megvalósult különböző hazai és nemzetközi projektekben. A 4.1 altézisben ismertetett, videomegfigyelési módszereket a MEDUSA című EDA és a THIS EU projektben használtuk fel. Az áramkörvizsgálati technológiánk (3.1 altézis) kifejlesztése részben kötődött BME-n futó TÁMOP-4.2.1/B-09/1/KMR-2010-0002 projekthez. A légi képpárokon megfigyelhető objektumelmozdulások vizsgálata (1.1 altézis) az ALFA NKTH projekthez (NKFP 2/046 /04) kapcsolódott. 3D környezetelemző algoritmusainkat régészeti térinformatikai adatok elemzésére is alkalmazzuk a PPKE EFOP-3.6.2-16-2017-00013 projektjének keretein belül. A járműre szerelhető forgó Lidaros fejlesztéseink (4.3 altézis) autópári projektben kerültek alkalmazásra. A Lidar alapú személymegfigyelést végző eljárásunkat (4.2 altézis) egy valós idejű demonstrátorba építettük be, amit különböző hazai megjelenések mellett a 2017-es Frankfurti autószalonon (IAA) is bemutathattunk.

5. Egyetemi oktatói és hazai publikációs munkák

Az elsősorban MTA SZTAKI-hoz kötődő kutatói tevékenységem mellett 2015 óta részmunkaidős egyetemi docensként dolgozom a Pázmány Péter Katolikus Egyetem (PPKE) Információs Technológiai és Bionikai Karán, ahol jelenleg a *Számítógépes grafika alapjai*, valamint a *Képelemzés* kurzusok tárgyfelelős oktatója és előadója vagyok. Habilitációs pályázatomat 2017-ben védtem meg [48]. Egyetemi és doktorandusz hallgatóim számos díjat nyertek el, köztük négy *1. helyezést* Országos Tudományos Diákköri Konferenciákon, Gábor Dénes Ösztöndíjat, Kuba Attila Díjat, a Hírközlési és Informatikai Tudományos Egyesület Országos Szakdolgozat Pályázatát és különböző SZTAKI-s intézeti díjakat.

2013 óta elnökségi tagja, 2019 januárja óta elnöke vagyok a Neumann János Számítógéptudományi Társaság (NJSZT) Képfeldolgozók és Alakfelismerők Szakosztályának (KÉPAF). A disszertációhoz kötődő számos publikációm magyar nyelven is megjelent a KÉPAF konferenciák belső kiadványaiban, a többi között [49, 51, 52]. Többször voltam résztvevője és előadója az NJSZT Magyar Számítógépes Grafika és Geometria Konferenciának is.

6. Köszönetnyilvánítás

A disszertáció alapjául szolgáló kutatómunkám elvégzéséhez szükséges anyagi és szellemi feltételeket valamint az infrastruktúra jelentős részét a Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutatóintézete (MTA SZTAKI), és a Pázmány Péter Katolikus Egyetem Információs Technológiai és Bionikai Kara (PPKE ITK) biztosította.

Nagy köszönettel tartozom közvetlen munkatársaimnak az MTA SZTAKI Gépi Érzékelés (korábbi nevén: Elosztott Események Elemzése) Kutatólaboratóriumából, akik az évek során folyamatosan nagy hatással voltak a munkáimra és eredményeimre. Különös köszönet illeti a laboratórium vezetőjét és korábbi doktori témavezetőmet, prof. Szirányi Tamást, aki mentorom volt az elmúlt 15 évben.

A felügyeletemre bízott B.Sc., M.Sc. és Ph.D. hallgatók a PPKE ITK-ról, és a Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnöki és Informatikai Karáról (BME VIK) sokban hozzájárultak a született eredményekhez. Köszönöm valamennyiük munkáját, kiemelve közülük is a doktorandusz hallgatóimat, Börcs Attilát, Nagy Balázst és Yahya Ibrahimot, valamint a számos tehetséges egyetemi hallgatóm közül Gálai Bencét és Józsa Oszkárt, akikkel közösen több publikációm is megjelent.

Köszönök minden szakmai és emberi segítséget a Gépi Érzékelés laborban dolgozó senior és Ph.D. hallgató kollégáimnak: Baráth Dánielnek, Eichhardt Ivánnak, Hajder Leventének, Havasi Lászlónak, Jankó Zsoltnak, Keszler Anitának, Kiss Ákosnak, Kiss Attilának, Kovács Leventének, Kriston Andrásnak, Majdik Andrásnak, Manno-Kovács Andreának, Pusztai Zoltánnak, Rózsa Zoltánnak, Maha Shadaydeh-nek, Spórács Lászlónak, Szilávik Zoltánnak, Tizedes Lászlónak, Varga Domonkosnak és Utasi Ákosnak. Köszönet Barti Mónikának és Vágvolgyi Anikónak az adminisztrációs csapatból, és a labor valamennyi egyetemi hallgatójának és szoftverfejlesztőjének. Köszönöm Nagy Eszter segítségét a SZTAKI könyvtárából, akire utazásaim szervezése és a publikációs adatok számbavétele során bármikor számíthattam az elmúlt években.

Különböző kutatási feladatokban közvetlen segítséget nyújtottak a velünk együttműködő külföldi és hazai intézmények kutatói is: Josiane Zerubia és Xavier Descombes a franciaországi INRIA Sophia Antipolis-ból, Marco Martorella és Fabrizio Berizzi a Pisa-i Egyetemről, Kató Zoltán a Szegedi Tudományegyetemről, Jakab László és Krammer Olivér a BME Elektronikai Technológia Tanszékéről és Csetverikov Dmitrij az MTA SZTAKI-ból. Szintén köszönöm Anuj Srivastava-nak, hogy látogatóként fogadott a posztdoktori időszakom alatt a Floridai Állami Egyetemen.

Köszönöm a PPKE ITK dékánjainak, Nyékyné Gaizler Juditnak, Szolgay Péternek és Iván Kristófnak, valamint a BME VIK Elektronikai Technológia, illetve Irányítástechnika és Informatika tanszékek vezetőinek, Harsányi Gábornak, Szirmay-Kalos Lászlónak és Kiss Bálintnak, hogy lehetőséget biztosítottak számomra a SZTAKI-s főállásom mellett oktatási és témavezetői munkára az egyetemeken.

A disszertációban bemutatott munkát a befogadó intézeteken túl a Bolyai János Kutatási Ösztöndíj valamint a 4. fejezetben részletezett különböző projektek is támogatták.

Végtelenül hálás vagyok Lívinek és Lócinak, az egész családomnak és valamennyi barátomnak, akik minden lehetséges módon támogattak.

7. Publikációk

7.1. A szerző folyóiratpublikációi

- [1] **C. Benedek**, B. Gálai, B. Nagy, and Z. Jankó, „Lidar-based gait analysis and activity recognition in a 4D surveillance system,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 1, pp. 101–113, 2018. IF: 3.558*.
- [2] **C. Benedek**, „An embedded marked point process framework for three-level object population analysis,” *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4430–4445, 2017. IF: 5.071.
- [3] A. Börcs, B. Nagy, and **C. Benedek**, „Instant object detection in Lidar point clouds,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 992 – 996, 2017. IF: 2.892.
- [4] **C. Benedek**, M. Shadaydeh, Z. Kato, T. Szirányi, and J. Zerubia, „Multilayer Markov random field models for change detection in optical remote sensing images,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 107, pp. 22–37, 2015. Special Issue on Multitemporal Remote Sensing Change Detection, IF: 4.188.
- [5] A. Börcs and **C. Benedek**, „Extraction of vehicle groups in airborne lidar point clouds with two-level point processes,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 3, pp. 1475–1489, 2015. IF: 3.360.
- [6] **C. Benedek**, „3D people surveillance on range data sequences of a rotating Lidar,” *Pattern Recognition Letters*, vol. 50, pp. 149–158, 2014. Special Issue on Depth Image Analysis, IF: 1.551.
- [7] **C. Benedek** and M. Martorella, „Moving target analysis in ISAR image sequences with a multiframe Marked Point Process model,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 4, pp. 2234–2246, 2014. IF: 3.514.
- [8] **C. Benedek**, O. Krammer, M. Janóczki, and L. Jakab, „Solder paste scooping detection by multi-level visual inspection of printed circuit boards,” *IEEE Transactions on Industrial Electronics*, vol. 60, no. 6, 2013. IF: 6.500.
- [9] Á. Utasi and **C. Benedek**, „A Bayesian approach on people localization in multi-camera systems,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 1, pp. 105–115, 2013. IF: 2.259.
- [10] **C. Benedek**, X. Descombes, and J. Zerubia, „Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 33–50, 2012. IF: 4.795.
- [11] **C. Benedek**, „Detection of soldering defects in printed circuit boards with hierarchical Marked Point Processes,” *Pattern Recognition Letters*, vol. 32, no. 13, pp. 1535 – 1543, 2011. IF: 1.034.
- [12] **C. Benedek**, T. Szirányi, Z. Kato, and J. Zerubia, „Detection of object motion regions in aerial image pairs with a multi-layer Markovian model,” *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp. 2303–2315, 2009. IF: 2.848.

- [13] **C. Benedek** and T. Szirányi, „Change detection in optical aerial images by a multi-layer conditional mixed Markov model,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 10, pp. 3416–3430, 2009. IF: 2.234.
- [14] **C. 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. IF: 3.315.
- [15] **C. 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. Special Issue on Applied Color Image Processing, IF: 0.482.

7.2. A szerző válogatott nemzetközi konferenciapublikációi

- [16] B. Nagy and **C. Benedek**, „Real-time point cloud alignment for vehicle localization in a high resolution 3D map,” in *Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving at ECCV'18*, vol. 11129 of *Lecture Notes in Computer Science*, pp. 226–239, Munich, Germany: Springer, 2019.
- [17] B. Gálai and **C. Benedek**, „Change detection in urban streets by a real time Lidar scanner and MLS reference data,” in *International Conference on Image Analysis and Recognition (ICIAR)*, vol. 10317 of *Lecture Notes in Computer Science*, pp. 210–220, Montral, Canada: Springer, July 2017.
- [18] B. Nagy and **C. Benedek**, „3D CNN based phantom object removing from mobile laser scanning data,” in *International Joint Conference on Neural Networks (IJCNN)*, (Anchorage, Alaska, USA), pp. 4429–4435, May 2017.
- [19] B. Gálai and **C. Benedek**, „Gait recognition with compact lidar sensors,” in *International Conference on Computer Vision Theory and Applications (VISAPP)*, (Porto, Portugal), pp. 1–7, Feb 2017.
- [20] B. Gálai, B. Nagy, and **C. Benedek**, „Crossmodal point cloud registration in the Hough space for mobile laser scanning data,” in *International Conference on Pattern Recognition (ICPR)*, (Cancun, Mexico), pp. 3363–3368, IEEE, Dec 2016.
- [21] B. Gálai and **C. Benedek**, „Feature selection for lidar-based gait recognition,” in *International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM)*, (Prague, Czech Republic), pp. 1–5, IEEE, Oct 2015.
- [22] **C. Benedek**, B. Nagy, B. Gálai, and Z. Jankó, „Lidar-based gait analysis in people tracking and 4D visualization,” in *European Signal Processing Conference (EUSIPCO)*, (Nice, France), pp. 1138–1142, Aug 2015.
- [23] A. Börzs, B. Nagy, M. Baticz, and **C. Benedek**, „A model-based approach for fast vehicle detection in continuously streamed urban LIDAR point clouds,” in *Workshop on Scene Understanding for Autonomous Systems at ACCV'14*, vol. 9008 of *Lecture Notes in Computer Science*, pp. 413–425, Singapore: Springer, April 2015.
- [24] A. Börzs, B. Nagy, and **C. Benedek**, „Fast 3-D urban object detection on streaming point clouds,” in *Workshop on Computer Vision for Road Scene Understanding and Autonomous*

- Driving at ECCV'14*, vol. 8926 of *Lecture Notes in Computer Science*, pp. 628–639, Zürich, Switzerland: Springer, March 2015.
- [25] **C. Benedek**, „Hierarchical image content analysis with an embedded marked point process framework,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, (Florence, Italy), pp. 5147–5151, May 2014.
 - [26] A. Börcs, B. Nagy, and **C. Benedek**, „On board 3D object perception in dynamic urban scenes,” in *IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, (Budapest, Hungary), pp. 515–520, December 2013.
 - [27] **C. Benedek**, Z. Jankó, C. Horváth, D. Molnár, D. Chetverikov, and T. Szirányi, „An integrated 4D vision and visualisation system,” in *International Conference on Computer Vision Systems (ICVS)*, vol. 7963 of *Lecture Notes in Computer Science*, pp. 21–30, St. Petersburg, Russia: Springer, July 2013.
 - [28] **C. Benedek**, „A two-layer marked point process framework for multilevel object population analysis,” in *International Conference on Image Analysis and Recognition (ICIAR)*, vol. 7950 of *Lecture Notes in Computer Science*, pp. 160–169, Póvoa de Varzim, Portugal: Springer, June 2013.
 - [29] A. Börcs, O. Józsa, and **C. Benedek**, „Object extraction in urban environments from large-scale dynamic point cloud dataset,” in *IEEE International Workshop on Content-Based Multimedia Indexing (CBMI)*, (Veszprém, Hungary), pp. 191–194, June 2013.
 - [30] O. Józsa, A. Börcs, and **C. Benedek**, „Towards 4D virtual city reconstruction from Lidar point cloud sequences,” in *ISPRS Workshop on 3D Virtual City Modeling*, vol. II-3/W1 of *ISPRS Annals Photogram. Rem. Sens. and Spat. Inf. Sci.*, pp. 15–20, May 2013.
 - [31] **C. Benedek**, D. Molnár, and T. Szirányi, „A dynamic MRF model for foreground detection on range data sequences of rotating multi-beam Lidar,” in *International Workshop on Depth Image Analysis (WDIA)*, vol. 7854 of *Lecture Notes in Computer Science*, pp. 87–96, Tsukuba City, Japan: Springer, November 2012.
 - [32] A. Börcs and **C. Benedek**, „Urban traffic monitoring from aerial LIDAR data with a two-level marked point process model,” in *International Conference on Pattern Recognition (ICPR)*, (Tsukuba City, Japan), pp. 1379–1382, November 2012.
 - [33] **C. Benedek** and M. Martorella, „Ship structure extraction in ISAR image sequences by a Markovian approach,” in *IET International Conference on Radar Systems*, (Glasgow, UK), October 2012.
 - [34] A. Börcs and **C. Benedek**, „A marked point process model for vehicle detection in aerial LIDAR point clouds,” in *XXII. ISPRS Congress*, vol. I-3 of *ISPRS Annals Photogram. Rem. Sens. and Spat. Inf. Sci.*, pp. 93–98, August 2012.
 - [35] Á. Utasi and **C. Benedek**, „A multi-view annotation tool for people detection evaluation,” in *Proceedings of the First International Workshop on Visual Interfaces for Ground Truth Collection in Computer Vision Applications*, (Capri, Italy), May 2012.
 - [36] D. Baltieri, R. Vezzani, R. Cucchiara, Á. Utasi, **C. Benedek**, and T. Szirányi, „Multi-view people surveillance using 3D information,” in *Proceedings of The 11th International Workshop on Visual Surveillance*, (Barcelona, Spain), pp. 1817–1824, November 2011.

- [37] **C. Benedek**, „Analysis of solder paste scooping with hierarchical point processes,” in *Proc. IEEE International Conference on Image Processing (ICIP)*, (Brussels, Belgium), pp. 2121–2124, September 2011.
- [38] **C. Benedek** and M. Martorella, „ISAR image sequence based automatic target recognition by using a multi-frame marked point process model,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, (Vancouver, Canada), pp. 3791–3794, July 2011.
- [39] Á. Utasi and **C. Benedek**, „A 3-D marked point process model for multi-view people detection,” in *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, (Colorado Springs, USA), pp. 3385–3392, June 2011.
- [40] A. Kovács, **C. Benedek**, and T. Szirányi, „A joint approach of building localization and outline extraction,” in *IASTED International Conference on Signal Processing, Pattern Recognition and Applications (SPPRA)*, (Innsbruck, Austria), February 2011.
- [41] Á. Utasi and **C. Benedek**, „Multi-camera people localization and height estimation using multiple birth-and-death dynamics,” in *Workshop on Visual Surveillance*, (Queenstown, NZ), November 2010.
- [42] **C. Benedek**, X. Descombes, and J. Zerubia, „Building detection in a single remotely sensed image with a point process of rectangles,” in *International Conference on Pattern Recognition (ICPR)*, (Istanbul, Turkey), August 2010.
- [43] **C. Benedek**, „Efficient building change detection in sparsely populated areas using coupled marked point processes,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, (Honolulu, Hawaii, USA), July 2010.
- [44] **C. Benedek**, X. Descombes, and J. Zerubia, „Building extraction and change detection in multitemporal remotely sensed images with multiple birth and death dynamics,” in *IEEE Workshop on Applications of Computer Vision (WACV)*, (Snowbird, Utah, USA), pp. 100–105, December 2009.
- [45] **C. Benedek** and T. Szirányi, „A mixed markov model for change detection in aerial photos with large time differences,” in *Proc. International Conference on Pattern Recognition (ICPR)*, (Tampa, FL, USA), IAPR, December 2008.
- [46] **C. 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, September 2007.

7.3. A szerző egyéb válogatott publikációi

- [47] **C. Benedek**, Z. Jankó, T. Szirányi, D. Chetverikov, O. Józsa, A. Börcs, and I. Eichardt, „Method and system for generating a three-dimensional model.” US Patent No. 10,163,256 (PCT number: PCT-HU2014-000017), Dec. 2018.
- [48] **C. Benedek**, „Probabilistic Approaches on Environment Perception with 2D-3D Imaging Sensors.” Habilitation thesis, Péter Pázmány Catholic University, Faculty of Information Technology and Bionics, June 2017.

- [49] B. Gálai and **C. Benedek**, „Járás alapú személyazonosítás és cselekvésfelismerés LiDAR szenzorokkal,” in *Képfeldolgozók és Alakfelismerők Társaságának Országos Konferenciája (Képf)*, (Szováta), január 2017.
- [50] A. Börcs, B. Nagy, and **C. Benedek**, „Dynamic environment perception and 4D reconstruction using a mobile rotating multi-beam Lidar sensor,” in *Handling Uncertainty and Networked Structure in Robot Control*, Studies in Systems, Decision and Control, pp. 153–180, Springer, 2016.
- [51] A. Börcs, B. Nagy, and **C. Benedek**, „Valós idejű járműdetekció LIDAR pontfelhő sorozatokon,” in *Képfeldolgozók és Alakfelismerők Társaságának Országos Konferenciája (Képf)*, (Kerekegyháza), január 2015. **Kuba Attila Díjat nyert.**
- [52] **C. Benedek**, „Hierarchikus jelölt pontfolyamat modell objektumpopulációk többszintű elemzéséhez,” in *Képfeldolgozók és Alakfelismerők Társaságának Országos Konferenciája (Képf)*, (Kerekegyháza), január 2015.
- [53] **C. Benedek** and T. Szirányi, „Shadow detection in digital images and video,” in *Computational Photography: Methods and Applications*, Digital Imaging and Computer Vision, pp. 283–312, CRC Press, Taylor & Francis, 2010.
- [54] **C. Benedek**, X. Descombes, and J. Zerubia, „Building extraction and change detection in multitemporal aerial and satellite images in a joint stochastic approach,” Research Report 7143, INRIA, Sophia Antipolis, December 2009.
- [55] **C. Benedek**, *Novel Markovian Change Detection Models in Computer Vision*. PhD thesis, Péter Pázmány Catholic University, Faculty of Information Technology, June 2008.
- [56] **C. Benedek**, T. Szirányi, Z. Kato, and J. Zerubia, „A three-layer MRF model for object motion detection in airborne images,” Research Report 6208, INRIA Sophia Antipolis, France, June 2007.

7.4. A tézisfüzetben idézett külső hivatkozások

- [57] H. Zhu, K. Yuen, L. Mihaylova, and H. Leung, „Overview of environment perception for intelligent vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, pp. 2584–2601, Oct 2017.
- [58] G. Scarpa, R. Gaetano, M. Haindl, and J. Zerubia, „Hierarchical multiple Markov chain model for unsupervised texture segmentation,” *IEEE Trans. on Image Processing*, vol. 18, no. 8, pp. 1830–1843, 2009.
- [59] J. Porway, Q. Wang, and S. C. Zhu, „A hierarchical and contextual model for aerial image parsing,” *International Journal of Computer Vision*, vol. 88, no. 2, pp. 254–283, 2010.
- [60] S. Geman and D. Geman, „Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.
- [61] Z. Kato and T. C. Pong, „A Markov random field image segmentation model for color textured images,” *Image and Vision Computing*, vol. 24, no. 10, pp. 1103–1114, 2006.

- [62] Z. Kato, T. C. Pong, and G. Q. Song, „Multicue MRF image segmentation: Combining texture and color,” in *International Conference on Pattern Recognition*, (Quebec, Canada), pp. 660–663, Aug. 2002.
- [63] P. Singh, Z. Kato, and J. Zerubia, „A Multilayer Markovian Model for Change Detection in Aerial Image Pairs with Large Time Differences,” in *International Conference on Pattern Recognition*, (Stockholm, Sweden), May 2014.
- [64] T. Szirányi and M. Shadaydeh, „Segmentation of remote sensing images using similarity-measure-based fusion-MRF model,” *IEEE Geosci. Remote Sens. Lett.*, vol. 11, pp. 1544–1548, Sept 2014.
- [65] A. Fridman, „Mixed Markov models,” *Proc. National Academy of Sciences of USA*, vol. 100, no. 14, pp. 8092–8096, 2003.
- [66] S. Müller and D. Zaum, „Robust building detection in aerial images,” in *ISPRS Object Extraction for 3D City Models, Road Databases and Traffic Monitoring - Concepts, Algorithms and Evaluation*, (CMRT05), (Vienna, Austria), pp. 143–148, 2005.
- [67] B. Sirmaçek and C. Ünsalan, „Building detection from aerial imagery using invariant color features and shadow information,” in *Int. Symp. on Computer and Information Sciences (ISCIS)*, (Istanbul, Turkey), 2008.
- [68] Z. Song, C. Pan, Q. Yang, F. Li, and W. Li, „Building roof detection from a single high-resolution satellite image in dense urban area,” in *ISPRS Congress*, (Beijing, China), pp. 271–277, 2008.
- [69] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny, „Structural approach for building reconstruction from a single DSM,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 1, pp. 135–147, 2010.
- [70] X. Descombes and J. Zerubia, „Marked point processes in image analysis,” *IEEE Signal Processing Magazine*, vol. 19, no. 5, pp. 77–84, 2002.
- [71] X. Descombes, R. Minlos, and E. Zhizhina, „Object extraction using a stochastic birth-and-death dynamics in continuum,” *Journal of Mathematical Imaging and Vision*, vol. 33, pp. 347–359, 2009.
- [72] A. Gamal-Eldin, X. Descombes, and J. Zerubia, „Multiple birth and cut algorithm for point process optimization,” in *International Conference on Signal-Image Technology and Internet-Based Systems (SITIS)*, (Kuala Lumpur, Malaysia), pp. 35–42, 2010.
- [73] Y. Verdie and F. Lafarge, „Detecting parametric objects in large scenes by Monte Carlo sampling,” *International Journal of Computer Vision*, vol. 106, pp. 57–75, 2014.
- [74] S. Ben Hadj, F. Chatelain, X. Descombes, and J. Zerubia, „Parameter estimation for a marked point process within a framework of multidimensional shape extraction from remote sensing images,” in *ISPRS Technical Commission III Symposium on Photogrammetry Computer Vision and Image Analysis (PCV)*, (Paris, France), 2010.
- [75] F. Chatelain, X. Descombes, and J. Zerubia, „Parameter estimation for marked point processes. application to object extraction from remote sensing images,” in *Energy Minimization Methods*

- in *Comp. Vision and Pattern Recogn.*, vol. 5681 of *Lecture Notes in Computer Science*, pp. 221–234, Bonn, Germany: Springer, 2009.
- [76] Y. Li and J. Li, „Oil spill detection from SAR intensity imagery using a marked point process,” *Remote Sensing of Environment*, vol. 114, no. 7, pp. 1590 – 1601, 2010.
 - [77] F. Chatelain, A. Costard, and O. J. J. Michel, „A Bayesian marked point process for object detection. Application to muse hyperspectral data,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Prague, Czech Republic), pp. 3628–3631, 2011.
 - [78] A. Gamal-Eldin, X. Descombes, and J. Zerubia, „A novel algorithm for occlusions and perspective effects using a 3D object process,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Prague, Czech Republic), pp. 1569–1572, 2011.
 - [79] C. Mallet, F. Lafarge, M. Roux, U. Soergel, F. Bretar, and C. Heipke, „A marked point process for modeling Lidar waveforms,” *IEEE Trans. on Image Processing*, vol. 19, no. 12, pp. 3204–3221, 2010.
 - [80] B.-T. Vo and B.-N. Vo, „Labeled random finite sets and multi-object conjugate priors,” *IEEE Trans. on Signal Processing*, vol. 61, no. 13, pp. 3460–3475, 2013.
 - [81] X. Descombes, ed., *Stochastic geometry for image analysis*. Digital Signal and Image Processing, Wiley-ISTE, 2011.
 - [82] M. Bredif, O. Tournaire, B. Vallet, and N. Champion, „Extracting polygonal building footprints from digital surface models: A fully-automatic global optimization framework,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 77, no. 1, pp. 57–65, 2013.
 - [83] J. Zhou, C. Proisy, P. Couteron, X. Descombes, J. Zerubia, G. le Maire, and Y. Nouvellon, „Tree crown detection in high resolution optical images during the early growth stages of eucalyptus plantations in brazil,” in *Asian Conf. on Pattern Recognition*, pp. 623–627, 2011.
 - [84] Y. Yu, J. Li, H. Guan, C. Wang, and M. Cheng, „A marked point process for automated tree detection from mobile laser scanning point cloud data,” in *International Conference on Computer Vision in Remote Sensing (CVRIS)*, (Xiamen, China), pp. 140–145, IEEE, 2012.
 - [85] A. Veillard, S. Bressan, and D. Racoceanu, „SVM-based framework for the robust extraction of objects from histopathological images using color, texture, scale and geometry,” in *International Conference on Machine Learning and Applications (ICMLA)*, vol. 1, pp. 70–75, 2012.
 - [86] W. Ge and R. Collins, „Marked point processes for crowd counting,” in *IEEE Conference on Computer Vision and Pattern Recognition*, (Miami, FL, USA), pp. 2913–2920, 2009.
 - [87] A. Börcs, *Four-dimensional Analysis of Dynamic Urban Environments in Terrestrial and Airborne LiDAR Data*. PhD thesis, Budapest University of Technology and Economics, June 2018.
 - [88] R. Laganière, *OpenCV 2 Computer Vision Application Programming Cookbook*. Packt Pub Limited, 2011.
 - [89] R. B. Rusu and S. Cousins, „3D is here: Point Cloud Library (PCL),” in *IEEE International Conference on Robotics and Automation (ICRA)*, (Shanghai, China), 2011.