

dc_494_12

MTA DOKTORI ÉRTEKEZÉS TÉZISEI

**Régiók és alakzatok:
Vizuális objektumok kinyerése és
illesztése**

(From Regions to Shapes:
The Extraction and Alignment of Visual Objects)

Kató Zoltán

*Képfeldolgozás és Számítógépes Grafika Tanszék,
Informatikai Tanszékcsoporth,
Szegedi Tudományegyetem*

2013

Tartalomjegyzék

1. Bevezetés	1
2. Előzmények	2
2.1. Koherens régiók kinyerése	2
2.2. Vizuális objektumok illesztése	2
3. Kutatási módszer	3
4. Az értekezés tézisei	5
1. Több felbontású és hierarchikus képszegmentáló Markov modellek	5
2. Valószínűségi modellek több-szemponjú képszegmentáláshoz és a 'gas of circles' alakzatmodell	6
3. 2D és 3D alakzatok illesztése megfeleltetések nélkül	7
5. Az eredmények hasznosítása	8
A szerző publikációi	10
További hivatkozások	14

1. Bevezetés

A digitális képek számítógépes elemzése, interpretációja alapvető fontosságú számos olyan területen, ahol a vizuális információ kiemelt szerepet játszik: például földterületek, ültetvények monitorozása műholdképek, légi felvételek segítségével; különböző orvosi képalkotó berendezésekhez kapcsolódó diagnosztikai feladatok, vagy termékek vizuális minőségellenőrzése a gyártási folyamat során. A számítógépes képelemző eljárásokkal szemben általános elvárás, hogy a bemeneti képeket lehetőleg emberi beavatkozás nélkül dolgozzák fel. A számítógépek alkalmazása sokrétű lehet: sok esetben az algoritmusok nem váltják ki teljesen az emberi képelemzést, hanem azt megkönnyítik például új vizuális tartalom előállításával, vagy a képi tartalom előkészítésével. Ez leginkább olyan kritikus alkalmazásoknál jellemző, mint pl. orvosi diagnosztika, ahol egy-egy diagnosztikai eljárás rutin feladati mellett mindig szükség van egyedi elemzésre is, amelyet szükségszerűen szakembernek kell megtennie. Más esetekben azonban a számítógép teljesen automatikusan elemzi a képet és nyeri ki a vizuális információt, emberi beavatkozásra nincs lehetőség. Ilyen tipikusan az ipari minőségellenőrzés, ahol egy nagy pontosságú kép alapú mérés, hibadetektálás emberi felügyeletére nincs mód, részben a rendkívül rövid feldolgozási idők miatt (egy gyártósoron ez akár néhány tized másodperc is lehet!), másrészt az emberi látás közvetlenül nem is alkalmas a látott kép ilyen pontosságú elemzésére.

Mindezekből jól látszik, hogy a számítógépek elsősorban a kontrollált környezetben, aprólékos részletek elemzésére kiterjedő feladatokban jobbak, míg az ember elsősorban az előre nem látható, intuíciót igénylő, átfogó jellegű vizuális információ kinyerésében jobbak. Ez döntően az ember és számítógép képi látásában mutatkozó fő különbségekre vezethető vissza:

- A számítógép alapvetően pixel-halmazként érzékeli a képeket, azok feldolgozását pixelenként végzi. Ebben a párhuzamos architektúra önmagában nem hoz lényegi változást, hiszen a műveleteket ekkor is pixel-szinten végezzük, ám ekkor ezek a műveletek párhuzamosan végezhető, melynek elsősorban a számítási komplexitásra van hatása. Ezzel szemben az emberi látás az egészet érzékeli. Ugyan a retinakép érzékelése elemi érzékelők (pálcikák, csapok) segítségével történik, az emberi érzékelésben ez a "diszkretizálás" közvetlenül nem jelenik meg.
- Míg az ember rengeteg előismeretet használ fel egy-egy kép elemzése során (ismeri a kontextust, felismeri a tárgyakat, következtet az alakjukra, képes kompenzálni a rossz látási viszonyokat, zavaró hatásokat), addig a számítógép "ismeretei" alapvetően a bemeneti kép pixelértékeire korlátozódnak. Míg az ember alapvetően objektumok érzékel, addig a számítógép pixeleket "lát", pontszerű objektumokban "gondolkodik" és ezekből próbálja meg összerakni a komplex vizuális objektumokat.

Munkám során olyan képfeldolgozó módszereket vezettem be, amelyek a fenti két korlátot próbálják meg kiküszöbölni. A kifejlesztett eljárások a gépi látás két alapproblémájára keresik a megoldást: vizuális objektumok (régiók, alakzatok) kinyerése illetve az ilyen objektumok több nézetből vagy több szenzor által készített képeinek illesztése, fúziója.

2. Előzmények

2.1. Koherens régiók kinyerése

Amikor megfigyelünk egy képet nem csak a kép intenzitásértékeit, színeit, textúráit nézzük és döntjük el mit ábrázol a kép, hanem jelentős mennyiségű való világból vett tapasztalattal is rendelkezünk, ami megkönnyíti a feladatot. Egy tárgy felismerésénél a tárgy alakját és azt is nézzük, milyen környezetben van. Az első képszegmentáló módszerek még csak a képek intenzitását használták, és így nem lehettek hatékonyak például erősen zajos képeken. Manapság a kutatások afelé haladnak, hogy a képen található információk (intenzitás, szín, textúra, stb.) mellett használjunk olyan előzetes információkat is, mint például a régiók simasága [53] vagy a keresett objektum alakja [63].

A korábbi aktív kontúros modellek egy előtér-háttér szegmentálásra koncentrálnak [59], adatmodellnek a kép gradiensét használják és egy előre definiált alak modellt [49, 48, 65, 52]. A kontúr mozgását a gradiens nagysága és a kontúr alakja határozza meg. Lehetséges lenne több információt is felhasználni az adatmodellhez, de korántsem triviális például, hogy egy aktív kontúr esetén hogyan használjuk fel a kép textúráját. A másik oldalon pedig ott van a Markov modell [53, 30], mely egy valószínűségi modell és könnyen kezeli a kép többi jellemzőit is, viszont nagyon kevés munka irányult arra, hogy előzetes alak információt építsenek be ilyen modellekbe.

A legtöbb alakleíró modell egy vagy több minta alakzatot és azok változatait használja a modellezésre. Ezek az eljárások nem elég hatékonyak abban az esetben, amikor az objektumok száma ismeretlen vagy az objektumok kölcsönhatásban vannak egymással. Minden különálló objektum külön tárolást és számítást igényel, kiegészítve a geometriai transzformációkat tároló változókkal. Minden egymással kölcsönható objektumpár extra számítással jár. A magasabb rendű aktív kontúr (HOAC) modellek egy másfajta megközelítést használnak [63]. A HOAC modellek le tudják írni az alakjellemzőket sablonok nélkül, helyettük a régiókat megadó határpontok közötti nagy hatósugarú függőségeket ad meg, ezek eredendően forgatás invariánsak. Amikor a HOAC modelleket fázis-mező modellekkel reprezentáljuk, felhasználhatók több kölcsönható objektumot tartalmazó konfigurációk modellezésére [64].

2.2. Vizuális objektumok illesztése

A képregisztráció számos képfeldolgozással kapcsolatos alkalmazás során fontos szerepet játszik, ahol olyan képeket kell összevetni, melyek különböző időpontban vagy különböző modalitással esetleg eltérő nézőpontból készültek (pl. képmozaikozás, alakfelismerés, orvosi képfeldolgozás, stb.). Általánosan a feladat úgy fogalmazható meg, hogy egy adott *minta* kép és ennek torzult változata (a *megfigyelés*) között kell megtalálni azt a transzformációt, melyet alkalmazva a *minta* képre a *megfigyelést* kapjuk. A regisztrációs módszerek [67, 61] általában tartalmazzák a következő komponenseket: A *tulajdonságpontok* olyan könnyen detektálható pontok (pl. sarokpont, szakaszok metszéspontjai, stb.), melyeket a képek közötti *pontmegfeleltetések* meghatározására használnak fel. Az illesztés jóságának méréséhez szükség van egy

hasonlósági mértékre. A keresési tér tartalmazza a lehetséges transzformációkat. A cél megtalálni azt a transzformációt, melyet a *minta* képre alkalmazva a lehető legkisebb eltérést kapjuk a *megfigyeléshez* képest a definiált *hasonlósági mérték* szerint.

A regisztrációs módszerek egyik csoportjába tartoznak a *pont alapú módszerek* [67, 54]. Az ilyen típusú módszerek olyan geometriai jellemzőket használnak, melyek a transzformáció során nem változnak (pl. sarokpontok, szakaszok metszéspontjai, stb.), viszont ezek meghatározása általában nem könnyű feladat. Egy népszerű és gyakran alkalmazott pont alapú módszer például az *iterative closest point* (ICP) algoritmus [47]. Egy másik megközelítés, a *terület alapú módszerek* [62, 56] esetében nem használnak tulajdonságpontokat. A feladatot ilyenkor globális leírók [51] vagy objektum invariánsok [57] segítségével oldják meg.

Szürkeárnyalatos képek affin regisztrációjára létezik hatékony módszer [55], ahol a feladatot egy lineáris egyenletrendszer megoldására vezetik vissza. Így elkerülhető mindenféle optimalizációs vagy pontmegfeleltetéssel kapcsolatos lépés. Ez a módszer viszont csak szürkeárnyalatos képeken működik, bináris képek esetén nem alkalmazható. Másrészt feltételezi, hogy a két kép intenzitásértékei pontosan megfelelnek egymásnak, avagy a radiometriai torzítás egyszerűen modellezhető. Sajnos ezek a feltételek a legtöbb gyakorlati problémánál nem teljesülnek [58, 66], így a módszer gyakorlati alkalmazhatósága korlátozott.

3. Kutatási módszer

Munkám során az elsődleges cél olyan új hatékony módszerek bevezetése volt, amelyek elméletileg megalapozottak, és általánosan alkalmazhatóak. A megfelelő matematikai modellek alkalmazása döntő fontosságú, hiszen csak így lesz módunk az algoritmusok helyességét, számítási komplexitását, vagy az esetleges modellparaméterek becslését egzakt módon vizsgálni. Ez a fajta megközelítés a 80-as évektől kezdve meghatározóvá vált, melynek eredményeként áttörő eredmények születtek a gépi látás és képfeldolgozás területén, és jól érzékelhető elmozdulás történt a csupán pixel alapú, ún. *bottom-up* technikáktól a komplexebb képjellemzőket és különböző bonyolultságú *a priori* ismereteket alkalmazó módszerek felé. A következő fejezetben azokat az eredményeimet foglalom össze tézisszerűen, amelyekkel hozzájárultam ehhez a folyamathoz.

A modell alapú képszegmentálás valószínűségi irányzatának kidolgozásában meghatározó volt a S. Geman és D. Geman [53] által bevezetett Markov véletlen mezőkön (MRF) alapuló modellezés és az ehhez szorosan kapcsolódó szimulált hűtés algoritmusának képfeldolgozásbeli alkalmazása egy új mintavételező, a *Gibbs sampler* segítségével. Az eredeti cikkben [53] képjavításra dolgoztak ki valószínűségi modellt és elemezték annak hatékonyságát, illetve elméletileg bizonyították a Gibbs mintavételezőn alapuló szimulált hűtés konvergenciáját. Ezzel megnyílt az út az olyan összetett valószínűségi modellek alkalmazása előtt, amelyek már nem csak az egyes pixelértékek statisztikáival dolgoznak, hanem képesek egy pixel környezetére is adott statisztikai eloszlást előírni, reprezentálni. Az ilyen modellek tipikusan egy összetett, a teljes képtől függő ún. *energiafüggvényt* definiálnak, amely lényegét tekintve egy költségfüggvény, melynek globális minimumát megkeresve megoldható az adott

képfeldolgozási probléma. Mivel ezek az energiafüggvények nem konvexek, ezért előtérbe kerültek a hatékony mintavételezőkön alapuló szimulált hűtés algoritmusok, amelyek valószínűségi konvergenciája is bizonyítható (például a *Gibbs sampler*).

Valós alkalmazások esetében azonban a kép struktúrális információjára sokszor csak összetett jellemzőkkel írható le jól, tehát szükséges az ilyen komplex régiók, vizuális objektumok modellezése. Ráadásul az egyszerű simasági/homogenitási kritérium mellett célszerű a keresett objektumok alakjait is figyelembe venni a szegmentálási modellünkben. A kérdés tehát úgy vetődik fel, hogy lehetséges-e egy ilyen komplex modellt alkotni, ami matematikailag jól leírható, és ezért a szükséges stabilitási és konvergencia-kritériumok teljesülése igazolható is. Ugyanakkor az alkalmazás oldaláról legalább ilyen fontos az is, hogy a feladat minél kevesebb emberi beavatkozással, lényegében felügyelet nélkül, hatékony algoritmussal megoldható legyen. Ezekre a kérdésekre több szakaszban kerestem a választ és értem el fontos eredményeket. Egyrészt olyan Markov modelleket alkottam, amelyek nem egyszerűen a bemeneti kép pixelein definiáltak, hanem a kép több felbontású változatain, azok közötti interakcióval, vagy pedig a kép különböző pixeltulajdonságainak reprezentációin, és az egyes tulajdonságok közötti interakcióval. Ezek a modellek tipikusan az első esetben több felbontású, míg az utóbbi esetben több rétegű hierarchiába szervezett Markov véletlen mezők lesznek. Másrészt az aktív kontúrok területén alkalmazott alakzatmodelleket vizsgáltam és megállapítottam, hogy azok számos megszorítással élnek: tipikusan csak képi gradiens alapú adatmodellt és sablon-szerű alakmodellt alkalmaznak, ami csak egyetlen előfordulást képes leírni az egész képen. Lehetséges-e a Markov modellek hatékony adatmodellezését párosítani valamilyen alakzatrepresentációval? Itt a fő kihívást az jelenti, hogy ezek a modellek régiókat reprezentálnak, míg a legtöbb alakzat-leíró kontúrokkal dolgozik. Sikertült egy olyan modellt létrehozni, amely egyszerű geometriai primitíveket képes adott kiterjedésű kölcsönhatásokkal reprezentálni, és ezáltal egy régió alapú alakzatmodellt létrehozni. A korábbi modellekhez képest nagy előny, hogy egy képen több, akár átfedő objektum is megtalálható, ha a képi régiók fölött egy többrétegű objektum-modellt definiálunk, ezzel megőrizve a módszer számítási hatékonyságát.

A modellek által definiált energiafüggvény minimalizálása ugyan lehetséges a klasszikus *Gibbs sampler* segítségével is, azonban a modellek konstrukcióját kihasználó ún. több hőmérsékletes algoritmusok sokszor bizonyultak hatékonyabbnak. Természetesen szükséges volt az új algoritmus konvergenciájának elméleti igazolása is. A szükséges modellparaméterek becslésére szolgáló algoritmusok kidolgozása nagyban hozzájárult a módszerek gyakorlati alkalmazhatóságához, melyre elsősorban a távérzékelés és mikroszkópia területén került sor.

A kinyert régiók, alakzatok, vizuális objektumok közötti geometriai kapcsolat meghatározása, vagyis a *regisztráció* nagyon fontos lépése egy-egy képfeldolgozási feladat megoldásának. Ha leegyszerűsítjük a problémát két képre, adott egy sablon és egy megfigyelés kép és a köztük ható geometriai transzformációt akarjuk meghatározni. A gyakorlati alkalmazásokban sok kikötést tehetünk az ismeretlen transzformációra, mint a transzformáció erőssége, invertálhatósága. A képek között ható transzformáció meghatározásához ismernünk kell a transzformáció erősségét, ami lineáris esetben lehet például merev test vagy affin, míg nemlineáris esetben polinomiális vagy elasztikus transzformáció. A klasszikus eljárások sokfélék, de valamennyi

közös jellemzője, hogy vagy pontmegfeleltetések alapján, vagy pedig az egyes képrészletek (radiometriai) tartalma közötti korreláción alapulnak. Vajon lehetséges-e olyan hatékony algoritmust konstruálni, amely ezek nélkül az információk nélkül is működik? Eldobhatjuk-e a képi intenzitásokat csak a bináris, szegmentált objektumokra redukálva a képet? Meghatározható-e a transzformáció pontmegfeleltetések kinyerése nélkül? Az általam kidolgozott eljárások képesek erre, ami komoly előrelépést jelent például a multimodális (többféle szenzor által készített) képek regisztrációjában, de az olyan problémák esetében is, ahol nem áll rendelkezésre megfelelő gazdagságú intenzitás-mintázat (pl. jelek, táblák, vagy bármilyen szegmentált objektum). Megmutatom hogyan oldható meg ez a feladat egy nemlineáris egyenletrendszer megoldásaként úgy, hogy az algoritmus gyors és közvetlenül szolgáltatja a megoldást optimalizálás vagy pontmegfeleltetés használata nélkül. A módszert számos orvosi és ipari alkalmazási területen teszteltük, mindenütt alacsony számításigényű algoritmussal sikerült a korábbiaknál jobb eredményeket elérni.

4. Az értekezés tézisei

Az alábbiakban az általam elért új tudományos eredményeket foglalom össze három téziscsoportban. Ezek közül az első a PhD fokozat megszerzése előtti eredményeimet tartalmazza, míg a másik kettő a fokozatszerzés utáni munkámat foglalja össze. Minden tézispont végén megadom az eredményekről beszámoló publikációimat.

1. Több felbontású és hierarchikus képszegmentáló Markov modellek

A S. Geman és D. Geman [53] által javasolt módszert sikeresen alkalmaztam szűrkeárnyalatos SPOT műholdképek szegmentálására. A módszer hátránya, hogy nagyobb képek esetén az algoritmus számításideje jelentősen megnő, másrészt ha a négy legközelebbi szomszédnál nagyobb kiterjedésű kölcsönhatást szeretnénk alkalmazni a pixelek között, akkor a komplexitás hatványozottan növekszik.

1. **Bevezettem** a többfelbontású hierarchikus Markov modelleket, amelyekben a bemeneti képet egyszerre több felbontásban reprezentáltam, az egyes felbontások közötti markovi interakcióval. Az eredményül kapott modell áttételesen képes volt a 4 szomszédnál lényegesen nagyobb területre kiterjeszteni a homogenitási kényszert úgy, hogy a modell energiáminimalizálásának időigénye nem emelkedett számottevően. *Kapcsolódó publikációk:* [12, 20, 18, 21, 19, 22, 30].
2. A számításidő szinten tartására **bevezettem** egy új többhőmérsékletes szimulált hűtést, melynek konvergenciáját a S. Geman és D. Geman [53] cikkben közölt tétel általánosításával **igazoltam**. Az így kapott képszegmentáló eljárást sikeresen lehetett alkalmazni SPOT műholdképeken a különböző vegetációjú parcellák kinyerésére. A módszer implementációját Franciaországban bejegyzett szoftver oltalom védi, és felhasználásra került a CNES-ben SPOT képek elemzésére. *Kapcsolódó publikációk:* [22, 30].

3. További feladatot jelentett azonban a Markov modellünk paramétereinek automatikus becslése, ami elengedhetetlenül szükséges az emberi interakció minimálisra csökkentéséhez. **Megmutattam**, hogy a modellparaméterek *maximum likelihood* alapú becslése előállítható egy adaptív szimulált hűtés algoritmussal. A így kapott automatikus képszegmentálási eljárás implementációját Franciaországban bejegyzett szoftver oltalom védi, és felhasználásra került a CNES-ben SPOT képek elemzésére. *Kapcsolódó publikációk*: [31, 23, 32, 33, 30].

2. Valószínűségi modellek több-szempontrú képszegmentáláshoz és a 'gas of circles' alakzatmodell

A különböző képelemzési feladatok során nem lehet mindig a pixelek intenzitásértéke alapján koherens régiókra bontani a képet, és így a keresett vizuális objektumok kinyerése megbízhatatlanná válik. Egy hatékony képszegmentáló eljárásnak tehát az egyszerű pixelértékek mellett olyan összetettebb jellemzőket is szükséges modelleznie, mint szín, textúra, mozgás.

1. Amennyiben a modellezni kívánt vizuális tulajdonságok hasonló módon jellemezhetők (pl. Gauss eloszlásokkal), akkor lehetséges egy közös, több változós valószínűségi eloszlás segítségével modellezni a jellemzőket. **Megmutattam**, hogy alkalmasan megválasztott színtér és textúra jellemzők esetén koherens színes textúrált régiók nyerhetők ki egy olyan Markov modellel, amely adatmodellként többváltozós Gauss sűrűségfüggvényt használ és a szükséges modellparaméterek az EM algoritmus alkalmazásával hatékonyan becsülhetők. **Módszert adtam** a többváltozós Gauss adatmodellt alkalmazó Markov modellek paraméterbecslésére a *Reversible Jump MCMC* mintavételező általánosításával. A módszert színes képek teljesen automatikus szegmentálására alkalmaztam, és a *Berkeley Segmentation Benchmark* keretében gyakorlati tesztekkel **igazoltam** az új módszer hatékonyságát. *Kapcsolódó publikációk*: [24, 26, 13, 14, 30].
2. **Bevezettem** egy új, többretegű Markov modellt képek komplex képjellemzők alapján történő szegmentálására (pl. szín, textúra, mozgás...). A modell alkalmas olyan jellemzők szimultán használatára is (pl. szín és mozgás), melyek eltérő viselkedésűek és ezért nem írhatók le jól egy közös valószínűségi eloszlással. A módszer hatékonyságát gyakorlati tesztekkel **igazoltam**. Több területen is sikeres volt a modell alkalmazása: mozgás-szegmentálás video képkockákon (pl. az MPEG kódolás egyik fontos lépése), illetve változás-detektálás légi felvételeken. *Kapcsolódó publikációk*: [28, 29, 25, 27, 1, 2, 30].
3. A magasabb rendű aktív kontúr (HOAC) modellek a szegmentálandó régiók alakzatára vonatkozó információt a távolabbi kontúrpontrókat interakciójaként tartalmazzák. Megmutattuk, hogy a HOAC modell paraméterei beállíthatók úgy, hogy a preferált régiók körszerűek (egy előre adott sugárral) – ezt neveztük 'gas of circles' HOAC modellnek. A modell egy ekvivalens, fázismező-alapú reprezentációjából kiindulva **bevezettem** egy ekvivalens Markov mo-

dellt, amely a fenti tulajdonságú régiók kinyerésére egy valószínűségi modellt ad – a 'gas of circles' MRF modellt. Gyakorlati úton **igazoltam**, hogy az így kapott modell energiája szimulált hűtéssel minimalizálható és a kívánt szegmentálást adja. A módszer sikeresen lett alkalmazva fakoronák kinyerésére légi felvételekről. *Kapcsolódó publikációk:* [11, 3].

4. Biológiai képek interpretációjánál a 'gas of circles' modellben egy jelentős korlátot jelent, hogy érintkező vagy átfedő köröket nem reprezentál. A probléma megoldására **bevezettem** egy alternatív alakzat-reprezentációt: a többrétegű 'gas of circles' modellt, amely a számítási hatékonyság megőrzése mellett képes érintkező és átfedő köröket, mint vizuális objektumokat reprezentálni a kép régiói fölött. A modell folytonos, fázismező-alapú, és diszkrét Markov mező alapú változata is kidolgozásra került, és mindkét módszer hatékonynak bizonyult a mikroszkópai képelemzés különböző szegmentálási feladatainak megoldásában. *Kapcsolódó publikációk:* [40, 37]

3. 2D és 3D alakzatok illesztése megfeleltetések nélkül

Egy objektum több nézőpontból, esetleg több szenzor által készült képe közötti geometria kapcsolat meghatározása, az ún. képregisztráció, gyakran felmerülő képelemzési probléma. **Bevezettem** egy általános keretrendszert, amely teljesen automatikusan, a korábbi módszerekhez szükséges pontmegfeleltetések illetve hasonlósági mértékek nélkül képes 2D és 3D alakzatok általános diffeomorf deformációinak meghatározására egy nemlineáris egyenletrendszer megoldása alapján.

1. **Módszert adtam** 2D és 3D alakzatok megfeleltetések nélküli affin regisztrációjára, ahol az illesztő transzformációt egy nemlineáris egyenletrendszer megoldásaként kapjuk. Az egyenleteket az illesztendő objektumok fölött vett nemlineáris függvények integráltjaiként kapjuk. Ennek következtében tetszőleges számú egyenletet elő lehet állítani, hiszen ez csak a felhasznált nemlineáris függvények számától függ. Ezzel egy flexibilis módszert kapunk elegendő számú egyenlet előállítására. **Megfogalmaztam** azt az eredményt, hogy hatványfüggvényeket alkalmazva mindig polinom egyenletrendszert kapunk, melynek megoldása numerikusan könnyen megoldható. **Bevezettem** a módszer egy alternatív formáját, ahol az illesztendő alakzatokra előbb egy Gauss sűrűségfüggvényt illesztünk, majd a nemlineáris függvényeket ezekre alkalmazva vesszük az integráljaikat az alakzatok fölött. Ez egy lineáris egyenletrendszerre vezet, melynek megoldása lényegesen kisebb számításigényű, mint a nemlineáris rendszeré. A kifejlesztett módszerek sikeresen alkalmazhatók orvosi képelemzési problémák megoldására. *Kapcsolódó publikációk:* [15, 4, 9, 5, 7, 6, 44, 46, 45, 8].
2. Az affin regisztrációs módszer továbbfejlesztésével **módszert adtam** általános nemlineáris diffeomorf transzformációk megfeleltetések nélküli becslésére. Konkrétan kidolgozásra került a módszer síkhomográfiára, polinom és vékony fémlemez (*thin plate spline*) transzformációkra, de más parametrikus modellekre is könnyen alkalmazható az eljárás. **Megfogalmaztam** az elméletei

feltételeit annak, hogy mikor lehet az integrálás alól kivonni az ismeretleneket és ezzel egy integrál egyenlet helyett egy egyszerű nemlineáris egyenlettel dolgozni. Ehhez az szükséges, hogy hatvány függvényeket használjunk és maga a deformációs modellünk bázisfüggvények lineáris kombinációjaként legyen megadva. A módszert felhasználva többféle regisztrációs problémát is sikerült hatékonyan megoldani az alkalmazási területek széles körében, mint például orvosi képelemzés vagy vizuális minőségellenőrzés. *Kapcsolódó publikációk:* [39, 38, 10, 35, 34, 36, 42, 41, 16, 17, 43]

5. Az eredmények hasznosítása

Az általam kifejlesztett és a disszertációmban részletesen is bemutatott eljárások számos alkalmazási területen bizonyultak hasznosnak és általuk lehetségessé vált olyan képelemzési feladatok megoldása, amely korábban csak lényegesen nagyobb számításigényű és sokszor kevésbé eredményes algoritmusokkal volt lehetséges.

A kifejlesztett többfelbontású hierarchikus képszegmentáló modellem felhasználásával sikeresen oldottuk meg műholdképek szegmentálását, melyből a Francia Űrkutatási Intézet (CNES) részére két szoftver implementáció is készült, melyeket Franciaországban bejegyzett szoftver oltalom véd. Ezeket a modelleket számos azóta megjelent monográfia is részletesen bemutatja, mint a Markov modellezés egy-egy fontos eredményét [60, 50]. A többrétegű és komplex jellemzőket alkalmazó Markov modelleimet sikeresen alkalmaztuk színes textúrált képek szegmentálására, video képkockák mozgás alapú szegmentálására, illetve változásdetekcióra légi felvételeken. A keresett vizuális objektumok alakját is figyelembe vevő szegmentáló modellem pedig hasznosnak bizonyult faültetvények légi felvételeken történő elemzésére illetve többféle mikroszkópiai kép feldolgozásában, különösen az érintkező és átfedő struktúrák megbízható kinyerésére.

Ezen eredményeimről meghívott előadóként/kutatóként előadásokat tartottam több rangos kutatóintézet és egyetem szemináriumain, mint pl. INRIA Sophia Antipolis (Franciaország), CWI Amszterdam (Hollandia), SZTAKI Budapest, illetve Poitiers (Franciaország) és a Hong Kong-i műszaki egyetemen. Ezek az eredmények meghatározó részévé váltak több általam kidolgozott kurzus tematikájának, mint pl. a Szegedi Tudományegyetemen beindított *Energiaminimalizációs módszerek a képszegmentálásban* speciálkollégium, az Informatika Doktori iskolában több ízben megtartott *Markov mezők a képfeldolgozásban* PhD kurzusom, az SSIP (Summer School on Image Processing) nyári iskola *Markov Random Fields in Image Segmentation* előadásai mellyel 2008-ban Bécsben "Best Lecturer" díjat is nyertem), valamint a 2009-ben a Linz-i egyetemen vendégprofesszorként tartott *Random Fields and Active Contours in Image Segmentation* kurzusom (<http://www.ssw.uni-linz.ac.at/Teaching/Lectures/SpezialLVA/Kato/>).

A képregisztráció területén elért eredményeimnek számos alkalmazása került kidolgozásra már eddig is: röntgenképek fúziója, törött csontok automatikus helyreállítása (műtéti tervezés fontos lépése), multimodális MRI és ultrahang prosztata képek regisztrációja, kézzel írott karakterek illesztése, KRESZ táblák illesztése, valamint vizuális minőségellenőrzés az autópárhánban használt gumitömlők felületi jeleinek

ellenőrzésére. Számos transzformációs modellre készítettünk demo implementációt, melyet publikusan letölthetővé tettünk. Az utóbbi két évben több száz letöltést regisztráltunk. Az eredményeimről meghívott előadóként plenáris előadást tartottam az IEEE SITIS 2011 (Dijon, Franciaország), VISIGRAPP 2012 (Róma, Olaszország) és az IEEE IITM 2013 (Allahabad, India) konferenciákon. Több vezető kutatóintézet és egyetem is meghívott szemináriumi előadást tartani a témában: UTIA, Cseh Tudományos Akadémia (Prága), Prágai Műszaki Egyetem, Zágrábi Műszaki Egyetem, Szingapúri egyetem (NUS) Műszaki kara, valamint az IPAL-CNRS kutatólabor (Szingapúr). Az eredmények az általam tartott nemzetközi oktatási előadásokba is beépültek, mint pl. Erasmus Mundus MSc előadás a Burgundi Egyetemen (Franciaország) valamint az SSIP 2012 (Bécs) nyári iskolán.

Az általam javasolt módszerek egy részének demo implementációja publikusan (*GNU General Public License* feltételeivel) letölthető a honlapomról (<http://www.inf.u-szeged.hu/~kato/software/>):

- *Supervised Image Segmentation Using Markov Random Fields* : MRF alapú algoritmus szürkeárnyaltos képek felügyelt és interaktív szegmentálására.
- *Supervised Color Image Segmentation in a Markovian Framework*: Színes képek felügyelt és interaktív szegmentálása Markov mezős modell alapján.
- *Affine Registration of Planar Shapes*: Alakzatok automatikus affin regisztrációja polinom egyenletrendszer direkt megoldásával (Windows alatt futó JAVA program).
- *Affine Registration of 3D Objects*: 3D objektumok automatikus affin illesztése túlhatározott egyenletrendszer megoldása alapján (Többszálás platformfüggetlen JAVA implementáció, ≈ 0.2 sec. futásidő több megavoxeles objektumokon).
- *Nonlinear Shape Registration without Correspondences*: Síkbeli alakzatok nemlineáris illesztése. A program projektív transzformáció (síkhomográfia) alapján illeszt, de más parametrikus nemlineáris modellre is könnyen kiterjeszthető.

A szerző publikációi

- [1] C. Benedek, T. Sziranyi, **Z. Kato**, and J. Zerubia, „A multi-layer MRF model for object-motion detection in uregistered airborne image-pairs,” in *Proceedings of International Conference on Image Processing*, (San Antonio, Texas, USA), pp. 141–144, IEEE, IEEE, Sep. 2007.
- [2] C. Benedek, T. Sziranyi, **Z. Kato**, and J. Zerubia, „Detection of object motion regions in aerial image pairs with a multilayer Markovian model,” *IEEE Transactions on Image Processing*, vol. 18, pp. 2303–2315, Oct. 2009.
- [3] T. Blaskovics, **Z. Kato**, and I. Jermyn, „A Markov random field model for extracting near-circular shapes,” in *Proceedings of International Conference on Image Processing*, (Cairo, Egypt), pp. 1073–1076, IEEE, IEEE, Nov. 2009.
- [4] C. Domokos and **Z. Kato**, „Binary image registration using covariant Gaussian densities,” in *International Conference on Image Analysis and Recognition*, (A. Campilho and M. Kamel, eds.), (Póvoa de Varzim, Portugal), pp. 455–464, Springer, June 2008.
- [5] C. Domokos and **Z. Kato**, „Affine alignment of compound objects: A direct approach,” in *Proceedings of International Conference on Image Processing*, (Cairo, Egypt), pp. 169–172, IEEE, IEEE, Nov. 2009.
- [6] C. Domokos and **Z. Kato**, „Affine puzzle: Realignment deformed object fragments without correspondences,” in *Proceedings of European Conference on Computer Vision*, (K. Daniilidis, P. Maragos, and N. Paragios, eds.), (Crete, Greece), pp. 777–790, Springer, Sep. 2010.
- [7] C. Domokos and **Z. Kato**, „Parametric estimation of affine deformations of planar shapes,” *Pattern Recognition*, vol. 43, pp. 569–578, March 2010.
- [8] C. Domokos and **Z. Kato**, „Simultaneous affine registration of multiple shapes,” in *Proceedings of International Conference on Pattern Recognition*, (Tsukuba Science City, Japan), pp. 9–12, IAPR, IEEE, Nov. 2012.
- [9] C. Domokos, **Z. Kato**, and J. M. Francos, „Parametric estimation of affine deformations of binary images,” in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, (Las Vegas, Nevada, USA), pp. 889–892, IEEE, IEEE, Apr. 2008.
- [10] C. Domokos, J. Nemeth, and **Z. Kato**, „Nonlinear shape registration without correspondences,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, pp. 943–958, May 2012.
- [11] P. Horvath, I. Jermyn, **Z. Kato**, and J. Zerubia, „A higher-order active contour model of a gas of circles’ and its application to tree crown extraction,” *Pattern Recognition*, vol. 42, pp. 699–709, May 2009.

- [12] **Z. Kato**, M. Berthod, and J. Zerubia, „Parallel image classification using multiscale Markov random fields,” in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, (Minneapolis), pp. 137–140, IEEE, IEEE, Apr. 1993.
- [13] **Z. Kato**, „Reversible jump Markov chain Monte Carlo for unsupervised MRF color image segmentation,” in *Proceedings of British Machine Vision Conference*, (A. Hoppe, S. Barman, and T. Ellis, eds.), (Kingston, UK), pp. 37–46, BMVA, Sep. 2004.
- [14] **Z. Kato**, „Segmentation of color images via reversible jump MCMC sampling,” *Image and Vision Computing*, vol. 26, pp. 361–371, March 2008.
- [15] **Z. Kato**, „A unifying framework for correspondence-less linear shape alignment,” in *International Conference on Image Analysis and Recognition*, (A. Campilho and M. Kamel, eds.), (Aveiro, Portugal), pp. 277–284, Springer, June 2012.
- [16] **Z. Kato**, „Linear and nonlinear shape alignment without correspondences,” in *Proceedings of International Joint Conference on Computer Vision, Imaging and Computer Graphics - Theory and Applications (Revised Selected Papers)*, (P. Richard, M. Kraus, R. S. Laramée, G. Csurka, and J. Braz, eds.), pp. 3–17, Rome, Italy: Springer, 2013. Keynote talk.
- [17] **Z. Kato**, „A unifying framework for correspondence-less shape alignment and its medical applications,” in *Proceedings of International Conference on Intelligent Interactive Technologies and Multimedia*, (A. Agrawal, R. Tripathi, and M. Tiwari, eds.), pp. 40–52, Allahabad, India: Springer, March 2013. Keynote talk.
- [18] **Z. Kato**, M. Berthod, and J. Zerubia, „A hierarchical Markov random field model and multi-temperature annealing for parallel image classification,” Research Report 1938, INRIA, Sophia Antipolis, France, Aug. 1993.
- [19] **Z. Kato**, M. Berthod, and J. Zerubia, „A hierarchical Markov random field model for image classification,” in *Proceedings of International Workshop on Image and Multidimensional Digital Signal Processing*, (Cannes, France), IEEE, Sep. 1993.
- [20] **Z. Kato**, M. Berthod, and J. Zerubia, „Multiscale Markov random field models for parallel image classification,” in *Proceedings of International Conference on Computer Vision*, (Berlin, Germany), pp. 253–257, IEEE, May 1993.
- [21] **Z. Kato**, M. Berthod, and J. Zerubia, „Parallel image classification using multiscale Markov random fields,” in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, (Minneapolis, USA), pp. 137–140, IEEE, Apr. 1993.

- [22] **Z. Kato**, M. Berthod, and J. Zerubia, „A hierarchical Markov random field model and multi-temperature annealing for parallel image classification,” *Computer Vision, Graphics and Image Processing: Graphical Models and Image Processing*, vol. 58, pp. 18–37, Jan. 1996.
- [23] **Z. Kato**, M. Berthod, J. Zerubia, and W. Pieczynski, „Unsupervised adaptive image segmentation,” in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, (Detroit, Michigan, USA), pp. 2399–2402, IEEE, May 1995.
- [24] **Z. Kato** and T. C. Pong, „A Markov random field image segmentation model using combined color and texture features,” in *Proceedings of International Conference on Computer Analysis of Images and Patterns*, (W. Skarbek, ed.), (Warsaw, Poland), pp. 547–554, Springer, Sep. 2001.
- [25] **Z. Kato** and T. C. Pong, „Video object segmentation using a multicue Markovian model,” in *Joint Hungarian-Austrian Conference on Image Processing and Pattern Recognition*, (D. Chetverikov, L. Czuni, and M. Vincze, eds.), (Veszprem, Hungary), pp. 111–118, KEPAF, OAGM/AAPR, Austrian Computer Society, May 2005.
- [26] **Z. Kato** and T. C. Pong, „A Markov random field image segmentation model for color textured images,” *Image and Vision Computing*, vol. 24, pp. 1103–1114, Oct. 2006.
- [27] **Z. Kato** and T. C. Pong, „A multi-layer MRF model for video object segmentation,” in *Proceedings of Asian Conference on Computer Vision*, (P. J. Narayanan, S. K. Nayar, and H.-Y. Shum, eds.), (Hyderabad, India), pp. 953–962, Springer, Jan. 2006.
- [28] **Z. Kato**, T. C. Pong, and G. Q. Song, „Multicue MRF image segmentation: Combining texture and color,” in *Proceedings of International Conference on Pattern Recognition*, (Quebec, Canada), pp. 660–663, IAPR, IEEE, Aug. 2002.
- [29] **Z. Kato**, T. C. Pong, and G. Q. Song, „Unsupervised segmentation of color textured images using a multi-layer MRF model,” in *Proceedings of International Conference on Image Processing*, (Barcelona, Spain), pp. 961–964, IEEE, Sep. 2003.
- [30] **Z. Kato** and J. Zerubia, *Markov random fields in image segmentation. Foundations and Trends in Signal Processing*, Now Publishers, Sep. 2012. 164 pages.
- [31] **Z. Kato**, J. Zerubia, and M. Berthod, „Unsupervised parallel image classification using a hierarchical Markovian model,” Research Report 2528, INRIA, Sophia Antipolis, France, Apr. 1995.
- [32] **Z. Kato**, J. Zerubia, and M. Berthod, „Unsupervised parallel image classification using a hierarchical Markovian model,” in *Proceedings of International*

Conference on Computer Vision, (Cambridge, MA, USA), pp. 169–174, IEEE, June 1995.

- [33] **Z. Kato**, J. Zerubia, and M. Berthod, „Unsupervised parallel image classification using Markovian models,” *Pattern Recognition*, vol. 32, pp. 591–604, Apr. 1999.
- [34] J. Mitra, **Z. Kato**, S. Ghose, D. Sidibe, R. Martí, X. Llado, A. Oliver, J. C. Vilanova, and F. Meriaudeau, „Spectral clustering to model deformations for fast multimodal prostate registration,” in *Proceedings of International Conference on Pattern Recognition*, (Tsukuba Science City, Japan), pp. 2622–2625, IAPR, IEEE, Nov. 2012.
- [35] J. Mitra, **Z. Kato**, R. Martí, A. Oliver, X. Llado, S. Ghose, J. C. Vilanova, and F. Meriaudeau, „A non-linear diffeomorphic framework for prostate multimodal registration,” in *Proceedings of International Conference on Digital Image Computing: Techniques and Applications*, (Noosa, Queensland, Australia), pp. 31–36, IEEE, Dec. 2011.
- [36] J. Mitra, **Z. Kato**, R. Martí, A. Oliver, X. Llado, D. Sidibe, S. Ghose, J. C. Vilanova, J. Comet, and F. Meriaudeau, „A spline-based non-linear diffeomorphism for multimodal prostate registration,” *Medical Image Analysis*, vol. 16, pp. 1259–1279, Aug. 2012.
- [37] C. Molnar, **Z. Kato**, and I. Jermyn, „A multi-layer phase field model for extracting multiple near-circular objects,” in *Proceedings of International Conference on Pattern Recognition*, (Tsukuba Science City, Japan), pp. 1427–1430, IAPR, IEEE, Nov. 2012.
- [38] J. Nemeth, C. Domokos, and **Z. Kato**, „Nonlinear registration of binary shapes,” in *Proceedings of International Conference on Image Processing*, (Cairo, Egypt), pp. 1001–1004, IEEE, November 2009.
- [39] J. Nemeth, C. Domokos, and **Z. Kato**, „Recovering planar homographies between 2D shapes,” in *Proceedings of International Conference on Computer Vision*, (Kyoto, Japan), pp. 2170–2176, IEEE, September 2009.
- [40] J. Nemeth, **Z. Kato**, and I. Jermyn, „A multi-layer ‘gas of circles’ Markov random field model for the extraction of overlapping near-circular objects,” in *Proceedings of Advanced Concepts for Intelligent Vision Systems*, (J. Blanc-Talon, W. Philips, D. Popescu, P. Scheunders, and R. Kleihorst, eds.), (Ghent, Belgium), pp. 171–182, Springer, Aug. 2011.
- [41] Z. Santa and **Z. Kato**, „Elastic registration of 3D deformable objects,” in *Proceedings of International Conference on Digital Image Computing: Techniques and Applications*, (Fremantle, Australia), pp. 1–7, IEEE, Dec. 2012.
- [42] Z. Santa and **Z. Kato**, „A unifying framework for non-linear registration of 3D objects,” in *Proceedings of International Conference on Cognitive Infocommunications*, (Kassa, Slovakia), pp. 547–552, IEEE, IEEE, Dec. 2012.

- [43] Z. Santa and **Z. Kato**, „Correspondence-less non-rigid registration of triangular surface meshes,” in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, (Portland, Oregon, USA), IEEE, IEEE, June 2013. Accepted.
- [44] A. Tanács, C. Domokos, N. Sladoje, J. Lindblad, and **Z. Kato**, „Recovering affine deformations of fuzzy shapes,” in *Proceedings of Scandinavian Conferences on Image Analysis*, (A.-B. Salberg, J. Y. Hardeberg, and R. Jenssen, eds.), (Oslo, Norway), pp. 735–744, Springer, June 2009.
- [45] A. Tanacs and **Z. Kato**, „Fast linear registration of 3D objects segmented from medical images,” in *Proceedings of International Conference on BioMedical Engineering and Informatics*, (Shanghai, China), pp. 299–303, IEEE, Oct. 2011.
- [46] A. Tanács, N. Sladoje, J. Lindblad, and **Z. Kato**, „Estimation of linear deformations of 3D objects,” in *Proceedings of International Conference on Image Processing*, (Hong Kong, China), pp. 153–156, IEEE, IEEE, Sep. 2010.

További hivatkozások

- [47] P. J. Besl and N. D. McKay, „A method for registration of 3-D shapes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 239–256, Feb. 1992.
- [48] Y. Chen, H. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. Gopinath, R. Briggs, and E. Geiser, „Using prior shapes in geometric active contours in a variational framework,” *International Journal of Computer Vision*, vol. 50, no. 3, pp. 315–328, 2002.
- [49] D. Cremers, F. Tischhauser, J. Weickert, and C. Schnorr, „Diffusion snakes: Introducing statistical shape knowledge into the Mumford-Shah functional,” *International Journal of Computer Vision*, vol. 50, no. 3, pp. 295–313, 2002.
- [50] P. Fieguth, *Statistical Image Processing and Multidimensional Modeling*. Springer, 2011.
- [51] J. Flusser and T. Suk, „A moment-based approach to registration of images with affine geometric distortion,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, pp. 382–387, March 1994.
- [52] A. Foulonneau, P. Charbonnier, and F. Heitz, „Multi-reference shape priors for active contours,” *Int. Journal of Computer Vision*, vol. 81, pp. 68–81, January 2009.
- [53] S. Geman and D. Geman, „Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, 1984.

- [54] H. Guo, A. Rangarajan, S. Joshi, and L. Younes, „Non-rigid registration of shapes via diffeomorphic point matching,” in *Proceedings of International Symposium on Biomedical Imaging: From Nano to Macro*, (Arlington, VA, USA), pp. 924–927, IEEE, April 2004.
- [55] R. Hagege and J. M. Francos, „Linear estimation of sequences of multi-dimensional affine transformations,” in *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, (Toulouse, France), pp. 785–788, IEEE, May 2006.
- [56] M. S. Hansen, M. F. Hansen, and R. Larsen, „Diffeomorphic statistical deformation models,” in *Proceedings of International Conference on Computer Vision*, (Rio de Janeiro, Brazil), pp. 1–8, IEEE, October 2007.
- [57] J. Heikkilä, „Pattern matching with affine moment descriptors,” *Pattern Recognition*, vol. 37, pp. 1825–1834, September 2004.
- [58] S. Kaneko, Y. Satohb, and S. Igarashi, „Using selective correlation coefficient for robust image registration,” *Pattern Recognition*, vol. 36, pp. 1165–1173, May 2003.
- [59] M. Kass, A. Witkin, and D. Terzopoulos, „Snakes: Active contour models,” *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [60] S. Z. Li, *Markov Random Field Modeling in Image Analysis*. Springer, 3rd edition ed., 2009.
- [61] J. B. A. Maintz and M. A. Viergever, „A survey of medical image registration,” *Medical Image Analysis*, vol. 2, pp. 1–36, March 1998.
- [62] S. Mann and R. W. Picard, „Video orbits of the projective group a simple approach to featureless estimation of parameters,” *IEEE Transactions on Image Processing*, vol. 6, pp. 1281–1295, September 1997.
- [63] M. Rochery, I. H. Jermyn, and J. Zerubia, „Higher order active contours,” *International Journal of Computer Vision*, vol. 69, pp. 27–42, August 2006.
- [64] M. Rochery, I. H. Jermyn, and J. Zerubia, „Phase field models and higher-order active contours,” in *Proc. IEEE International Conference on Computer Vision (ICCV)*, (Beijing, China), October 2005.
- [65] M. Rousson and N. Paragios, „Shape priors for level set representations,” in *Proceedings of European Conference on Computer Vision*, (Copenhagen, Denmark), pp. 78–92, Springer, 2002.
- [66] K. M. Simonson, S. M. Drescher, and F. R. Tanner, „A statistics-based approach to binary image registration with uncertainty analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 112–125, January 2007.
- [67] B. Zitová and J. Flusser, „Image registration methods: A survey,” *Image and Vision Computing*, vol. 21, pp. 977–1000, October 2003.