

Opponensi vélemény

Kató Zoltán: From Regions to Shapes: The Extraction and Alignment of Visual Objects

(Régiók és alakzatok:Vizuális objektumok kinyerése és illesztése)

c. MTA doktori (DSc) értekezéséről

Témaválasztás - problémafelvetés

Kató Zoltán értekezését érintő kutatása során a gépi látás két fontos kérdésére keresi a választ,

(1) miként lehetséges képekből hasonló tulajdonságú pixelek besorolásával összetett, vizuális objektumokat (régiók, alakzatok) kinyerni, azaz a képeket szegmentálni,

(2) objektumok több nézetből vagy több szenzor által készített képeit miként lehet egymásnak megfeleltetni, illeszteni. Ez, a regisztrációként is említett feladat rendkívül fontos a vizuális ellenőrzésben, az orvosi képek elemzésében, és a videó alapú tárgykövetésben.

Kiemelendő, hogy a jelelőlt mindkét feladat kapcsán olyan algoritmusok kidolgozását tűzte ki célul, melyek minél kevesebb emberi beavatkozást és felügyeletet igényelnek, és a korábbi megoldásoknál kevesebb számításigénynek köszönhetően új alkalmazási területeken is hasznosnak bizonyulhatnak.

Az első feladat megoldása kapcsán a legfontosabb kérdés, hogy a képek szegmentálásához a klasszikus információkon túl (pl. intenzitás, szín) miként lehet az előzetes ismereteket felhasználni, illetve mi az a magasabb szintű reprezentáció amely alkalmazásával a meglévő algoritmusok hatékonysága növelhető.

A második – megfeleltetési – feladat megoldásnak kulcsa szintén a magasabb szintű reprezentáció alkalmazásában rejlik. A megfeleltetési feladat lényege, hogy megtaláljuk azt a transzformációt, melyet a mintaképre alkalmazva az összehasonlítható képet (megfigyelést) kapjuk. A „klasszikus” pont és terület alapú módszerek alkalmazhatósága megköveteli az összehasonlítható objektumok pontos (manuális) definiálását. A csupán intenzitásértékeken alapuló technikák kizárólag akkor alkalmazhatók sikeresen, amikor a kapcsolódó objektumok intenzitásértékei pontosan megfelelnek egymásnak. A multimodális képek hatékony elemzése új megközelítésmódot igényelt. A megoldás kulcsa, hogy szegmentált objektumokra redukált képek kerülnek összevetésre.

A témaválasztás – a kutatás időpontjában – rendkívül időszerű volt. A vizsgált problémakör nagyon széles. A felvetett kérdések jelentőségét a publikációk és alkalmazások szép száma is bizonyítja.

A megközelítésmód

A (hierarchikus) modell alapú képszegmentálásra és a szegmentált objektumok alapján történő képösszevetésre kidolgozott eljárások mögött meghúzódó gondolatok rendkívül ötletesek.

A szegmentálás alapja a szomszédos pixelek közötti összefüggések feltárása. Markov mezőkkel modellezhetők. A Markov modell paramétereinek meghatározása összetett feladat. Multigrid módszereket széles körben alkalmaznak a konvergencia gyorsítására. Az elsőként ismertetett eljárásban e megközelítésmód alapján a hierarchia a bemeneti kép különböző felbontású változatain értelmezett.

A többváltozós Gauss keverék modellek – ahogy a számítástudomány sok más területén – itt is kiválóan alkalmazhatóknak bizonyulnak. A Jelölt által kifejlesztett többretegű modell a szegmentálás során e rétegekben definiált normális eloszlás keverék modellek megfelelő kombinációjával olyan információk feltárására alkalmas, melyek az egyes rétegekből (pl. szín, textúra) külön-külön nem tárhatók fel. Az EM algoritmuson túl a szimulált hűtés alkalmazása szintén széles körben alkalmazott technika.

A megközelítésmód nagyon fontos pozitív eleme, hogy az alkalmazott technikákhoz a jelölt kiemelkedő matematikai igényességgel nyúl, alaposan megvizsgálja a modellek strukturális ekvivalenciáit (pl. 3.5 fejezet), elemezi az algoritmusok fontosabb tulajdonságait, például bizonyítja az általa létrehozott többhőmérsékletes szimulált hűtés konvergenciáját.

Az elméleti megalapozottságon túl kiemelendő a kísérleti munka példaértékű alapossága. A fontosabb paraméterek hatását érthető és látványos példák illusztrálják, a fejezetek a létrehozott metódusok alkalmazhatóságát igazoló részletes, jól dokumentált esettanulmányokkal zárulnak. Az eredmények reprodukálhatóságát az interneten közzétett saját fejlesztésű szoftverek is biztosítják.

A dolgozat felépítése

A 183 oldalas rendkívül szépen szerkesztett, szemléletes ábrákkal illusztrált értekezés struktúrája mintaszerű, angolsága kiváló. A dolgozat a tématerületen kevésbé jártas olvasó számára is teljes mértékben követhető; jól felépített módon ismerteti a kutatás problémáját, a célkitűzéseket, az alapgondolatot és a kapcsolódó irodalmat.

A dolgozat számozott fejezetei egy-egy altézishez kapcsolódó tudományos eredményt ismertetnek. Az eredmények tézisszerű összefoglalására az utolsó, Conclusion címet viselő fejezetben kerül sor, ahol a téziseket alátámasztó publikációk is tételenként említésre kerülnek. A három részletesebb matematikai bizonyítás a mellékletben kapott helyet.

Az eredmények összefoglalása

Az első tézist alátámasztó első fejezet a Geman és Geman által szürkeárnyaltos képekre kidolgozott szegmentálási technikát többfelbontású Markov mező modelleken alapuló fejlesztését ismerteti. A Jelölt a modell paramétereinek identifikálására új szimulált hűtésen alapuló optimalizációs eljárást dolgozott ki, mely konvergenciáját elméletileg is igazolta.

A 2.1 és 2.2. téziseket alátámasztó második fejezet többváltozós normál eloszlás keverék modellek színes képek Bayes-i szegmentálására való alkalmazhatóságát vizsgálja. A Reversible Jump MCMC mintavételező eljárás általánosításán alapuló paraméterbecslési eljárás hatékonyan kezeli azt a problémát, hogy a címkeszám előzetesen nem ismert.

2.3 tézist alátámasztó harmadik fejezet a magasabb rendű aktív kontúrok (HOAC) családjába tartozó, közel azonos sugarú körök halmazát leíró 'gas of circles' (GOC) modell alkalmazásával foglalkozik. A Markov mezőként reprezentált MOC modell előnyeit távérzékelésében, fakoronák szegmentálásában igazolta.

A 2.4 tézist alátámasztó negyedik fejezet a GOC modellek alkalmazhatóságát bővíti azzal, hogy egymáshoz kapcsolódó, egymással átlapoló köröket többretegű modell segítségével reprezentálja.

A harmadik tézis 2D és 3D alakzatok megfeleltetések nélküli illesztésére alkalmas eljárások fejlesztéséhez köthető eredményeket foglalja össze. A létrehozott keretrendszer szerint a megfeleltetéseket nemlineáris függvények integráljaiként kapjuk.

A 3.1 tézist alátámasztó ötödik fejezet egy olyan eljárást ismerteti, amely az illesztendő alakzatokat többváltozós normál eloszlásfüggvénnyel jellemzi, mely reprezentáció előnye, hogy ezek transzformációja lineáris egyenletrendszerre vezet.

A 3.2 tézist alátámasztó hatodik fejezet annak a feltételeit keresi, hogy mikor lehet az integrálás alól kivonni az ismeretleneket, lehetővé téve, hogy az integrálegyenlet helyett egyszerű nemlineáris egyenlettel dolgozhassunk. Fontos eredmény, hogy hatványfüggvényeken alapuló reprezentáció és bázisfüggvények lineáris kombinációjaként megadott deformációs modellek esetén lineáris időkomplexitású és egyszerűen megvalósítható, a megfeleltetések (manuális) meghatározása nélkül egyértelmű megoldást biztosító transzformáció adható.

Értékelés, megjegyzések

A tézisekhez köthető publikációk mennyisége (51 darab) és minősége (többségében IEEE konferenciák folyóirat publikációk), a hivatkozások 1200-at meghaladó száma (a Google Scholar szerint) kiemelkedő tudományos munkásságot indikál. A tudományos eredményeket kivétel nélkül sikeres gyakorlati alkalmazások igazolják.

Meglepő, hogy a 2013-ban elkészített értekezés 207 hivatkozása között 2010-nél frissebb dátumú alig szerepel (az önhivatkozások nélkül csak egy hivatkozás származik 2012-ből), a hivatkozások döntő többsége az 2000-es évekből való. Miért nincsenek friss irodalmi hivatkozások, az elmúlt időszak eredményei miként támaszkodtak a javasolt megközelítésmódokra, milyen más kapcsolódó eljárások váltak sikeressé és széles körben alkalmazottá?

Sztereoszkópikus képalkotás, szuperfelbontás területe rohamosan fejlődött az elmúlt évtizedben. Ezek a feladatok miként támogathatók a létrehozott eszközökkel?

A gauss keverék modellek végigkísérik a dolgozatot. A szerző is említi, hogy ezek identifikálása gyakran numerikus problémákkal terhelt (pl. 103. oldal). Volt-e szükség valamilyen „trükk” alkalmazására az ilyen esetekben (pl. a kovariancia mátrix kezelése kapcsán)?

Összefoglaló

A Jelölt tézisértékű új kutatási eredményeket ért el, ezeket elegáns módon publikálta és rájuk számos érdemi független hivatkozást kapott.

A tézisek kivétel nélkül új tudományos eredményeket fogalmaznak meg.

A benyújtott MTA doktori értekezés számomra példaértékű stílusú, rendkívüli kreativitásról, alaposságról, átgondoltságról és jó pedagógiai érzékről tesz tanúbizonyságot.

A fentiek alapján javaslom a munka nyilvános védésre bocsátását, sikeres védés esetén az MTA Doktora cím odaítélését.

Veszprém, 2014. árpilis 10.

Abonyi János
MTA doktora