

Bírálat

Dr. Kató Zoltán „Régiók és alakzatok: Vizuális objektumok kinyerése és illesztése” című MTA doktori értekezéséről

1. Témaválasztás. Kató Zoltán tudományos eredményei elsősorban a gépi látás, illetve képfeldolgozás területén születtek. Mind a képek szegmentációja, mind a különböző képi alakzatok összeillesztése nagyon fontos és nehéz problématerület. Mindkét feladat lényege, hogy „intelligens” vizuális információt nyerjünk ki elemi, zajos adatokból, amelyeket nagyméretű pixelhalmazok hordoznak. A nemzetközi képfeldolgozó közösség számos technikát és algoritmust dolgozott ki ezen feladatok megoldására az elmúlt évtizedek során. A tudományterület fejlődését sok új eljárás kifejlesztése fémjelzi, amelyek jobb minőséget, nagyobb megbízhatóságot és számítási hatékonyságot biztosítanak, valamint folyamatosan bővítik a lehetséges alkalmazási területek körét. Véleményem szerint Kató Zoltán eredményei hozzájárultak a tudományterület fejlődéséhez.

(Ezúton szeretném megjegyezni, hogy képfeldolgozás nem része szűkebb értelemben vett tudományos kutatási területemnek (3D geometriai modellezés), ennek ellenére szívesen fogadtam el a bírálói felkérést, mert úgy gondolom, hogy meg tudom ítélni Kató Zoltán tudományos munkásságának jelentőségét.)

2. A dolgozat felépítése és a fejezetek rövid ismertetése. Az értekezés angol nyelven íródott, hossza 183 oldal, a bevezetés és az eredmények összefoglalása között hat technikai fejezet található. Az értekezés több mint 80 szemléltető ábrát és táblázatot, valamint egy több mint 200 elemet tartalmazó referencia listát tartalmaz, ebből 51 a szerző saját publikációinak listája. Az értekezés tartalmi, szerkesztési és nyomdai szempontból egyaránt elismerésre méltó, professzionális munka. A hat technikai fejezet egységes szerkezetet követ: a probléma ismertetése és a feladat megoldásának motivációja után egy részletes matematikai tárgyalás következik, amelyet az alkalmazott numerikus módszerek ismertetése követ. A fejezetek végén képekkel szemléltetett alkalmazási példák találhatók.

A *szegmentálás* célja, hogy a képet egymástól elkülönülő, jól meghatározott tartományokra bontsuk. Az 1. fejezetben bemutatott szegmentáló eljárások valószínűségi modelleken alapszanak. A Markov véletlen mezőkön (MRF) alapuló módszerek feltételes valószínűségeket rendelnek az egyes pixelekhez (illetve környezetükhöz) attól függően, hogy az adott pixel feltehetően melyik szegmenshez tartozik. Ez alapján a teljes képre egy energiafüggvény definiálható, amelynek globális minimuma adja a legjobb szegmentációt. A minimum meghatározásához a leggyakrabban alkalmazott numerikus eljárás a szimulált hűtés (simulated annealing). Az 1. fejezetben leírt eljárás a képet egyidejűleg több felbontásban, hierarchikusan reprezentálja. Ezzel összhangban a szimulált hűtés hőmérsékletének beállításához egy új stratégia kerül bevezetésre, mely szerint a durvább felbontásokhoz magasabb, a finomabb felbontásokhoz alacsonyabb hőmérséklet tartományok rendelődnek. A Markov modell paramétereinek meghatározásához egy adaptív szimulált hűtés algoritmust javasol a szerző.

A 2. fejezetben tárgyalt szegmentálás több képi jellemzőt (szín, textúra, mozgás) együttesen kezel, feltéve, hogy ezek hasonló valószínűségi eloszlással jellemezhetők. A javasolt eljárás többváltozós Gauss sűrűségfüggvényeket használ, a Markov modell paramétereinek becslésére az ún. Reversible Jump MCMC eljárást alkalmazza. Az eljárás segítségével színes képeket lehet automatikusan szegmentálni. A 2.4. alfejezetben egy többrétegű Markov modell kerül

bevezetésre, amely különböző valószínűségi eloszlással rendelkező képi jellemzők párhuzamos kiértékelésén alapul. A módszert sikeresen tesztelték mozgás-szegmentálási és változás-detektálási feladatok során.

A 3. fejezet egy speciális szegmentálási feladattal foglalkozik, ahol a *detektálandó tartományok körszerűek*, hozzávetőlegesen egyenlő sugárral (GOC = gas of circles), a 4. fejezet pedig kiterjeszti ezt a feladatot olyan esetekre, amikor a körtartományok érintkeznek, illetve átlapolódnak. (Megjegyzés: jó lenne a 'gas of circles' magyar megfelelőjét megtalálni.) Ez a kutatás az ún. aktív kontúrokkal történő szegmentáció egy részfejezete; a tradicionális aktív kontúrok folytonos görbéket reprezentálnak, amelyek egy energia-minimalizálási eljárás során több lépésben húzódnak rá az azonosítandó tartományok határvonalára. A szerző egy újszerű Markov modell bevezetését javasolja, amely során az eredeti aktív kontúr algoritmus átforgalmazódik egy valószínűségeken és összefüggő tartományokon alapuló eljárássá, majd a szegmensek szimulált hűtés alapján állnak elő a korábban tárgyalt módszerekkel összhangban. A 4. fejezetben a fenti modell kiterjesztésére kerül sor, ahol az érintkező és átfedő köröket a képpel asszociált különböző rétegekben reprezentálja. A módszert sikeresen alkalmazták fakoronák azonosítására légi felvételek alapján, valamint mikroszkópiai képelemzések során.

Az 5. és 6. fejezet *alakzatok regisztrációjával*, más szóval alakzatok összeillesztésével foglalkozik. Adott egy mintaalakzat (template), és adott egy másik kiértékelendő kép (observation), amelyen valamilyen formában megjelenik a mintaalakzat. A leképzést egy ismeretlen transzformáció írja le, célunk a lehető legjobb transzformáció meghatározása. A szakirodalomban alkalmazott tradicionális megoldás, hogy először jellegzetes alaksajátosságokat határoznak meg a két képen (angolul: feature points, feature curves), majd ezeket összeillesztve számítják ki az optimális transzformációt. Az 5. és 6. fejezetben javasolt megoldás viszont egy nagyon érdekes alternatív koncepción alapul, ahol nincs szükség alakzatsajátosságok kijelölésére és összeillesztésére. Az optimális transzformáció egy nemlineáris egyenletrendszer megoldásaként adódik, ahol az egyenleteket az objektumokon értelmezett függvények integráljaként határozzuk meg. Az 5. fejezetben a probléma átforgalmazott változata is megtalálható, ahol egy lineáris egyenletrendszert kell megoldani, így ez értelemyszerűen lényegesen hatékonyabb.

Az affin transzformációkon túllépve, a 6. fejezet nemlineáris összeillesztésekkel foglalkozik, különböző általánosabb deformációs modelleket feltételezve (pl. projektív leképezések, vagy a fizikai analógián alapuló vékonylemez deformációk.) Az összeillesztés nélküli regisztrációs módszerek hatékonyságát meggyőzően szemlélteti a 6.5. és 6.6. ábrásorozat, ahol a szerző – különböző mintákat választva – összehasonlítja saját eredményeit néhány korábban publikált módszerrel.

3. A tézisek értékelése. Bizonyítotttnak tekintem, hogy a jelölt nemzetközileg jelentős tudományos eredményekkel gyarapította a képfeldolgozás területét, és hozzájárult ezen tudományos ágazat fejlődéséhez. A bíráló dolga aránylag könnyű, ha az eredmények már korábban megjelentek magas impakt faktorral rendelkező szakfolyóiratokban, - többek között: IEEE Transactions on Image Processing, Pattern Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Image and Vision Computing, Medical Image Analysis - következőképp a nemzetközi tudományos

közösség már korábban jelentősnek minősítette ezeket az eredményeket. Megjegyzem, hogy a jelölt idézettségi indexe is igen magas.

Mindhárom tézis fontos eredményeket tartalmaz az alábbiak szerint:

Az első tézist – *'Több felbontású és hierarchikus képszegmentáló Markov modellek'* – elfogadom, mint új tudományos eredményt, beleértve az új reprezentációt és a szegmentáló algoritmust, valamint a többhőmérsékletes szimulált hűtés algoritmusát.

A második tézist – *'Valószínűségi modellek több-szempontú képszegmentáláshoz és a 'gas of circles' alakzatmodell'* – elfogadom, mint új tudományos eredményt, beleértve a többváltozós Markov modellt, az RJMCMC paraméterbecslés algoritmusát, valamint a többretegű Markov modellt, amely különböző eloszlású képjellemzőkön alapul. Elfogadom a körlapok szegmentációjára vonatkozó algoritmust, amely a tradicionális aktív kontúr szegmentációt egy Markov modell alapú eljárással helyettesíti, továbbá a többretegű megoldást is az érintkező/átfedő körök szegmentációjára.

A harmadik tézist – *'2D és 3D alakzatok illesztése megfeleltetések nélkül'* – elfogadom, mint új tudományos eredményt, beleértve magát az algoritmust, a nemlineáris egyenletrendszer felírását, a javasolt numerikus eljárást, és az alternatív lineáris változatot. Ugyancsak elfogadom az algoritmus kiterjesztését, amikor affin leképezések helyett az illesztés egy általánosabb deformációs modellre épül.

Megjegyzés. Szeretném felhívni a figyelmet a Doktori Szabályzat 1. paragrafusának (f) pontjára, miszerint a doktori cím adományozásának feltétele, hogy a jelölt

f) az általa művelt tudományágat és szakterületet a tudományos fokozat megszerzését követően jelentős eredeti tudományos eredménnyel gyarapította, amivel hozzájárult a tudomány továbbfejlődéséhez.

Jelölt az első téziscsoportot a PhD fokozat megszerzése előtti eredménynek nyilvánította, ez a fentiek alapján nem szerencsés megfogalmazás. A fenti ellentmondás feloldására két lehetőség van:

a) az első tézisinél csak azokat a publikációkat vesszük figyelembe, amelyek a 94-es PhD védést követően születtek, azaz a [22, 30, 33] számmal jelletteket; ezek igazolják, hogy a jelölt a PhD-ben ismertett kutatásait a továbbiakban is folytatta, és új eredményeket ért el.

b) Miután mind az első, mind a második téziscsoport valószínűségi modelleken alapuló szegmentációs technikákkal foglalkozik (Markov Random Fields), ezért a fenti probléma teljes mértékben eliminálódik, ha az első és második téziscsoportot összevonva, együttesen fogadjuk el.

A fenti megjegyzés természetesen nem befolyásolja azon véleményemet, hogy az értekezésben foglalt eredmények egyértelműen bizonyítják a jelölt jelentős tudományos munkásságát.

4. Összefoglalás. A fentiek alapján javaslom a nyilvános vita kitűzését és az MTA Doktora cím odaítélését.

Budapest, 2014. május 4.



Dr. Várady Tamás
az MTA doktora