

Opponensi vélemény

MTA TTK Árszéna Doktori Tanács TTK Árszéna	
Ikt./Értk.	2016 NOV 10. n.
Előadó	Ol
Értk.	
Megjegyzés	
2011	

Hajdu András (2016) *Diszkrét geometriai és fúziós módszerek objektumok detektálásához és a döntéstámogatáshoz*

A benyújtott disszertáció témája a digitális képfeldolgozás területén belül két fő problémakörre fókuszál. Az időrendben is első terület a képekről kinyert objektumok, alakzatok mint bináris képek elemzése elsősorban digitális geometriai eszközökkel. Míg ezen a területen elsősorban elméleti eredményeket ért el a szerző, addig a másik terület erősen alkalmazás-orientált: retinaképek elemzése orvosdiagnosztikai döntéstámogatáshoz. A szerző mindkét területen jelentős új eredményeket ért el. Tekintettel a két terület eltérő jellegére, a disszertáció jól mutatja, hogy a jelölt otthonosan mozog a képfeldolgozás elméleti és alkalmazott kutatási területein is.

Az értekezés felépítése, stílusa, kidolgozottsága

A 174 oldalas szépen szerkesztett, szemléletes ábrákkal illusztrált értekezés struktúrája és angolsága jó. A dolgozat 8 számozott fejezetből áll, a 2-7 fejezetek egy-egy altézishez kapcsolódó tudományos eredményt ismertetnek. Ezek követhető, jól felépített módon mutatják be a kutatás problémáját, a célkitűzéseket, az alapgondolatot és a kapcsolódó irodalmat, valamint a javasolt módszer gyakorlati tesztelésének eredményeit. Az eredmények tézisszerű összefoglalására az utolsó, *Summary of new scientific results* című fejezetben kerül sor, ahol a téziseket alátámasztó publikációk is tétélesen említésre kerülnek. A két részletesebb matematikai bizonyítás a mellékletben kapott helyet. Az 1. fejezetben, mint egy szövegben foglalja össze a szerző a dolgozatban több helyen is felhasznált fogalmakat, összefüggéseket, algoritmusokat és adatbázisokat. Kétségtelenül hasznos egy ilyen összefoglaló, de az olvasó számára nem igazán követhetőek ezek a mozaikszerűen összeszerkesztett, tematikailag azonban nem összetartozó fogalmaknak az ismertetése, különösen így a dolgozat elején. Én inkább egy függelékben helyeztem volna el ezt az anyagot. Sajnos a dolgozat több részében is hiányoznak a hivatkozások fontosabb felhasznált algoritmusokra, összefüggésekre (legyen az szakirodalmi hivatkozás vagy az 1. fejezet releváns része). A teljesség igénye nélkül:

1. Section 3.2: Eulerian curve, Floodfill8 algorithm

2. Section 5.6.1: (5.59) – (5.62) mértékekhez egyetlen hivatkozás sincs megadva
3. Section 6.1: Simulated annealing
4. Section 7.3.2: Egyetlen osztályozó algoritmushoz sincs megadva referencia

Ez némiképp megnehezíti a gondolatmenetek követését.

Tudományos eredmények

A szerző tudományos eredményeit 3 tézisben foglalja össze. A tézisekben bemutatott eredmények több magas színvonalú nemzetközi folyóirat- illetve konferenciákban lettek publikálva, az egyes tézispontok megfelelő számú és minőségű publikációval vannak lefedve. Ezek alapján tehát megállapítható, hogy a szerző az MTA doktora cím megszerzéséhez megfelelő minőségű tudományos munkát végzett. A továbbiakban az egyes tézispontok eredményeit értékelem.

1. Objektumok optimális mintavételezése és tömörítése

Az 1. altézisben, melynek bemutatására a 2. fejezetben kerül sor, a szerző a központi Voronoi felbontás elméletének kiterjesztésével bevezet egy új mintavételező eljárást folytonos- és diszkrét térbeli objektumok egyszerűsítésére. A szerző igazolta, hogy ez a technika optimális az ún. chamfer-illesztési módszerekben. Az eredményeket elméletileg igazolja a szerző, valamint gyakorlati tesztekben összehasonlítást végez klasszikus mintavételezési eljárásokkal két alkalmazási területen: emberi jelenlét detekció (Section 2.3.2) és retina képekben erek detektálására (Section 2.3.3). Mindkét tesztesettel az a fő probléma, hogy a szerző nem hasonlítja össze az általa javasolt Chamfer-illesztésen alapuló detektáló módszert más ember- illetve érrendszer-detektáló algoritmusokkal. Ezért a tesztek csak annyit igazolnak, hogy a mintavételezésre vonatkozó elméleti eredmények a gyakorlatban is kézzelfoghatóak. Azonban a tesztek alapján nem jelenthető ki, hogy a módszer az emberdetekció területén valójában mennyire jelentős. Ezért az altézis eredményei elsősorban a megfogalmazott elméleti eredmények és nem azok alkalmazása.

A 2. altézisben (3. fejezet) a szerző új gráfelméleti, hatékony tömörítést szolgáló módszereket ad tetszőleges topológiájú digitális görbék tömörítéséhez. A módszer lényege, hogy a görbék Euler-felbontással Euler gráfokra bontva bejárhatók, illetve a felbontás mellőzhető a kínai postás probléma megoldásának alkalmazásával, ahol a súlyokat a görbedarabok tömörítési igénye adja. A vázkijelöléssel nyert digitális görbék vázkijelölés

során deformálódó kereszteződéseinek detektálásához elméleti eredményeken alapuló módszert javasol, mely a kereszteződő szegmenseket sávok formájában kezeli, és azok vastagságának, valamint bezárt szögének függvényében határozza meg a deformáció nagyságát. A módszer összehasonlítására a JBEAM eljárással a 3.4.1 fejezetben kerül sor. A 3.1 táblázatban azonban csak a tömörítés mértéke szerepel, mint egyetlen paraméter. Mivel a tömörítés veszteséges, ezért legalább ilyen fontos lenne a *tömörítés minőségének* elemzése is. Ugyanez mondható el a 3.4.2 fejezetben tárgyalt MPEG-4 alakzat kódolásról is (3.2 táblázat).

2. Egy- és több példányban megjelenő objektumok detektálása

Az egyik leggyakoribb képfeldolgozási feladat adott objektumok jelenlétének és helyzetének minél pontosabb meghatározása. A szerző a területen új, térbeli szavazórendszerek modelljeit vezette be, melyek alkalmasságát elméleti eszközökkel bizonyította, valamint kísérletileg igazolta.

Az 1. altézisben (4. fejezet) a vak- és sárgafolt retinaképeken történő detektálásához javasol a szerző több lehetséges módszert a szakirodalomban elérhető detektáló algoritmusok kombinálására. A tézisfüzetben ugyan általánosabban kerül megfogalmazásra az eredmény, mint az altézist tárgyaló 4. fejezetben („...A képen egy példányban megjelenő, egy képponttal reprezentálható objektumok detektálásához fúziós módszereket vezettem be, amelyek különböző elveken alapuló detektorok kimeneteit egyesítik többségi- vagy súlyozott többségi szavazómodellek alapján...”), de erről az általánosabb alkalmazhatóságról kevés szó esik a disszertációban, így inkább a konkrét alkalmazás keretein belül tudom elfogadni az itt bemutatott eredményeket. A 4.1.2 alfejezetben tárgyalt módszer meglehetősen kezdetleges, sok a kézzel beállított és láthatóan képfüggő paraméter. Nem világos, hogy a módszer mennyire alkalmazható más típusú képekre. A gráf alapú módszer már inkább tekinthető elméletileg is megalapozott eredménynek. A 69. oldal alján a szerző javasol egy közelítő formulát (4.8), amelyet a tesztekben használ. Sajnos itt hiányzik annak elemzése, hogy ezen közelítés alkalmazása mennyire van hatással az eredményekre, a közelítés hatása pontosan miben jelentkezik az elméleti korrekt formulához képest. A 4.2. táblázatban közölt eredmények esetében nem világos, hogy OD_5 miért ad konzisztensen jobb eredményt a javasolt halmaz-alapú módszernél. A 4.1 táblázat szerint pedig a javasolt módszer alig teljesít jobban, mint OD_4 . Mi ennek a magyarázata? Véleményem szerint a 4.4.

fejezetben tárgyalt módszer van a legalaposabban kidolgozva. Mind elméleti mind gyakorlati oldalról meggyőző eredményekkel van alátámasztva a javasolt eljárás hatékonysága. Azt azonban itt is meg kell jegyezni, hogy a szakirodalomban javasolt releváns módszerekkel itt sincs kellő alaposítási összehasonlítás.

A 2. altézis (5. fejezet) eredményei talán a legszebbek az egész disszertációban. A szerző mind elméleti mind gyakorlati oldalról kellő alaposítással tárgyalja a módszer viselkedését. A fejezet a klasszikus többségi és súlyozott többségi szavazás térbeli szavazómodellekké történő kiterjesztéséről szól, ahol az egyes szavazatok detektorok jelöltjeiként, képpontok formájában érkeznek. Ehhez egy olyan valószínűségi elemet vezet be, ami speciálisan megfigyelhető egy detektálni kívánt objektum alakjára vonatkozó feltételnek.

A 3. altézis (7. fejezet) megfogalmazása ismét általánosabb, mint ahogy az a disszertációban tárgyalásra kerül. Ezért itt is csak az alkalmazás erejéig tudom elfogadni az eredményeket, azok általánossága nincs kellőképpen alátámasztva. Ez azonban semmit nem von le az eredmények értékéből, ezt az altézist az előzővel együtt a dolgozat legértékesebb részének tekintem. Az eredményekkel kapcsolatban csupán annyi (inkább a prezentációra vonatkozó) kritikát fogalmaznék meg, hogy 113. oldalon tárgyalt szimulált hűtés algoritmusánál illetet volna megemlíteni, hogy ez a sokféle változat közül az ún. *Metropolis-Hastings* algoritmus. Továbbá nem szerencsés, hogy az algoritmus konvergenciája szempontjából kritikus *hőmérséklet ütemező* egyáltalán nincs a jelentőségének megfelelően kezelve. Nincs megemlítve sem, hogy elméletileg csak logaritmikus skálán szabad változtatni a hőmérsékletet, különben elvesz a konvergencia elméleti garanciája. Ugyanakkor a gyakorlatban elterjedt ún. *exponenciális ütemező*, melyet vélhetően a szerző is használ, csak akkor ad jó eredményt, ha a disszertációban q -val jelölt szorzófaktor közel 1!

3. Összetett rendszerek a döntéstámogatásban

A tézisek közül ezt tartom a legkevésbé kidolgozottnak (7. fejezet). Ugyan a tézist alátámasztó publikációk megfelelőek a tézis elfogadásához, mégis szerencsésebb lett volna inkább az előző tézissel egy egységben kezelni. Annál is inkább, mivel itt kerülnek alkalmazásra a 2. tézispont eredményei egy komplex diagnosztikai döntéstámogató rendszerben. Sajnos a tézis megfogalmazása itt is túlságosan általánosra sikerült, melyet megint csak a disszertáció 7. fejezetében konkrétan tárgyalt diabéteszes retinopátia automatikus szűrésének keretében tudok elfogadni. Abból a szempontból az eredmények

teljessége megfelelő tesztekkel van alátámasztva és a szükséges összehasonlító elemzések is megtörténtek a terület elfogadott módszereivel. A kiértékelés alapját a diagnosztikában megfogalmazott alapelvek adják, így minden bizonnyal a gyakorlatban is relevánsak az elért eredmények. Sajnos a tézis alábbi mondatában szereplő új módszereket nem tudtam azonosítani a disszertációban bemutatott módszerekkel, ezért ezeket csak úgy tudom elfogadni új eredményként, ha a szerző válaszában kitér ezekre a módszerekre:

„Matematikai morfológiai eszközökkel a sárgafolt, rejtett Markov mezőt használva a retina érhálózatának, míg intenzitásprofilok vizsgálatával mikroaneurizmák kinyerésére adtam eljárást.”

Összegzés

A Jelölt értékes kutatási eredményeket ért el, ezeket nemzetközi mércével is rangos folyóiratokban és konferenciákon publikálta, melyekre számos érdemi független hivatkozást kapott. A tézisek kivétel nélkül új tudományos eredményeket fogalmaznak meg, azokat elfogadom azzal a megjegyzéssel, hogy a tézisfüzetben a tézisek megfogalmazása több esetben általánosabb, mint ahogy az a disszertációban ténylegesen bemutatásra kerül. Az eredmények általánosabb alkalmazhatósága jövő kérdése. Továbbá a 3. tézis súlya kisebb, azt jelen formájában inkább a 2. tézis eredményeinek szép és gyakorlati szempontból is teljes értékű alkalmazásának lehet tekinteni.

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

Szeged, 2016 november 5.



Kató Zoltán

