Elosztott Események Elemzése Kutatólabor

Számítástechnikai és Automatizálási Kutató Intézet
Magyar Tudományos Akadémia
H-1111 Budapest, Kende u. 13-17.
http://web.eee.sztaki.hu

SZIRÁNYI Tamás, MTA doktora
Tudományos tanácsadó
Egyetemi tanár
Kutatási laborvezető

☎: 279-6106 📠 279-6292
💻: sziranyi.tamas@sztaki.mta.hu
Web: http://www.sztaki.hu/~sziranyi

Bírálat

Hajdu András

Diszkrét geometriai és fúziós módszerek objektumok detektálásához és a döntéstámogatáshoz

Discrete Geometric and Fusion Based Techniques for Object Detection and Decision Support

c. MTA Doktori Értekezéséről

Hajdu András folyamatosan magas szintű tudományos eredményt nyújt a számítógépes képfeldolgozás és alakfelismerés területén. Különösen értéke eredményei vannak a biológia és orvosi célú képek kiértékelésével kapcsolatban.
A disszertáció 174 oldalas, jól tagolt, a fejezetek mindenütt bevezető összefoglalóval és tartalomjegyzékkel ellátottak. A disszertáció megírásában komoly kihívást jelentett, hogy egyszerre szerepelnek igényes matematikai eszközök és matematikai okfejtések, együtt a mérnöki kísérleti munka elemeivel és kiértékelési követelményével. A jelölt többnyire eredményesen kezelte ezt az összetett - egyébként kivételesen becsülendő - követelményt; ugyanakkor, a tézisek megfogalmazása, a saját eredmények egyértelműsítése nehezen birkózik meg az összetett igényekkel.

A bevezető után az alapeljárásokat bemutató fejezet következik, majd tematikus 6 db. fejezet jön, melyeket az eredmények összefoglalója követi kétfajta bontásban is (*Related results* és *Scientific results*), végül a hivatkozáslista 219 tétele, melynek első 55 eleme a saját mű. Fontos megjegyezni, hogy a jelölt valamennyi tézisét kiemelkedő nemzetközi publikációi is alátámasztják.

Az első fejezet a gráfelmélet és statisztikai mérések egyfajta összefoglalója, de érdekes módon idekerültek a retina 4. fejezetbeli mérésének hivatkozott rutinjai is, valamint a képi adatbázisok leírása.
A második fejezet a gyors minta-illesztés új módszereiről szól. Itt a körvonal és a területalapú egyszerűsítés, amolyan adaptív újra-mintavételezés a cél. Ebbe beletartozik a szkeletonizálás optimalizálása is. Az eljárások alapvetően a bonyolultabb későbbi lépések (felismerés, szemantikai elemzés) megalapozó lépései, elsőrendű célja a gyorsítás.
Kérdés 1. Jó lenne látni, hogy a gyorsítás milyen szerepet játszik egy összetettebb feladat komplexitásában?

A példák között szerepel a hőképen érzékelt emberi alak felismerése. A hőkép jelentősen eltér az optikai kameraképektől, ezért speciális algoritmusokat/jellemzőket kíván. Az emberi alak felismerésében szívesen vennék egy összehasonlítást más emberi-alak detekciós eljárásokkal, melyek között szintén vannak gyorsított változatok: wavelet vagy hierarchikus mintavételezés alapon. A Voronoi megoldás biztosan alak-adaptívebb, kérdés a hatékonysága. Az általam fontosnak ismert, gyalogos-detekcióra vonatkozó publikációk:

- Dollárék mérvadó eredménye az ICF eljárás:
    - Dollar P, Tu Z, Perona P, Belongie S (2009a) Integral channel features. In: BMVC, vol 2, p 5
- Deformálható modellekre vinatkozóan pedig a HOG alapú eljárás egy változata hozott áttörést:
    - Felzenszwalb PF, Girshick RB, McAllester D, Ramanan D (2010) Object detection with discriminatively trained part-based models, IEEE T. Pattern Anal Mach Intell 32(9):1627–1645

A hőképes ember-detekcióra vonatkozóan itt is található összehasonlítás:

- Varga D, Sziranyi T, Kiss A, Sporas L, Havasi L (2015) A multi-view pedestrian tracking method in an uncalibrated camera network, ICCV 2016, IEEE, pp 37–44

Kérdés 2. Az ismertetett minta-kivételi eljárás hatékonysága hogyan viszonyul más eljárásokhoz?

Kérdés 3. A Snake alapú kontúrillesztés/detekció milyen körülmények között használható eredményesen? Mik a használhatóság határai?

Kérdés 4. A Fourier descriptor-t éppen hasonló célra használják, jó eredménnyel, de a disszertációban erről szintre semmi szó nem esik. Ha már megvan a körvonal, a FD nagyon gyors. Ezzel összevetették-e az eljárást?

Kérdés 5. A *false positive* és *true positive* esetekre jó lenne több példát mutatni. Eleve áll a disszertáció egészére, hogy bár korrekten dokumentált, de az olvasó jobban átlátná a problémákat több minta-kép esetén.

Ha a tárgy mozog a térben, és változik a nagysága (felbontása), a szkeletonizálásnak, Voronoi eljárásnak is hasonló eredményt kellene adnia.

Kérdés 6. Mit tud mondani a CVT eljárás skálázhatóságáról?

A ***harmadik fejezet*** a görbék kódolásáról szól, gráfelméleti megoldásokkal. A kidolgozott eljárások új és érdekes megoldások, és a JBEAM-hez képest jelentős javulást hoznak.

Az eljárásban szerepel a tracing/követés kifejezés, amit meg is valósít a legközelebbi irányok keresésével.

Kérdés 7. Érdekes lenne tudni, hogy a dinamikusabb offline aláírás-képekre hogyan működik?

A poligonokkal, illetve ügyesen bontott vonal-szegmensekre bontással történő görbe-leírás már régi trükk (*iterative endpoint fit*, Duda and Hart 1973), és hosszú a története. A szerző hatékonyabb, új eljárást dolgozott ki.

Kérdés 8. Mik a görbe-közelítés hatékonysága kritériumának a jellemző értékei: műveleti sebesség, kód-hossz, illeszkedési pontosság?

Kérdés 9. A vonal-szegmensek kódolásában a „run-length” vagy hasonló tömörítési elvek hogyan működnek? A különböző típusú görbék kódkönyvtára mennyire hasonló?

Kérdés 10. Az 56. Oldalon a CPP probléma 4 lépése megoldása saját eredmény?

Megjegyzés: a 3.11 ábrához tartozó szekvencia közepén: „…5, 8, 9, **11, 15**, 16, 17, 16 ….” A 11-15 élátmenet nem létezik, valójában 11 helyett 13-nak kell lenni.

Kérdés 11. A retinaképeken az erek vastagsága egy fontos jellemző, miért kellene eltüntetni? Pl. valamelyik jellemző pont (pl. OD közepe) körül vett körsávon áthaladó erek vonalkódot alkotnak, ami pl. a személyazonosításban használatos.

Megjegyzés: Itt is szívesen láttam volna részletesebb képi példákat is.

A ***negyedik fejezet*** a retina jellemző foltjai: az optikai diszk és makula meghatározásáról szólnak. Alapvetően többféle módszer (1. fejezet) alapján működő boosting eljárásokról van szó: gyenge leírókból súlyozással stb. erős jellemzőt generál. Hasonlók vannak pl. DeepLearning-el is:
Zilly et al: Boosting Convolutional Filters with Entropy Sampling for Optic Cup and Disc Image Segmentation from Fundus Images, 2015

Kérdés 12. Az együttes hatékonyság néha alig jobb, mint egy egyedi jellemzőé. Pl. a 4.1 táblázat OD4 vs. Ensemble, vagy 4.2 táblázat OD5 vs. Ensemble. Ezt hogyan értékeli?

A továbbiakban több értékes eredményt látunk, valahol szerepel „So, we examine only the applicability of strict theoretical approaches,” – hiányzik viszont egy átfogó gondolat(menet).

Kérdés 13. Mi az a fő elv, ami szerint az eljárást megítélhetnénk? Mi ennek a fejezetnek az elvi és a gyakorlati üzenete? Van-e tágabb jelentősége, mint a retina vizsgálata?

A 4.15-16 ábrák szemléletesen mutatják az eljárás hatékonyságát. Ugyanakkor, a megtalált foltok eleve eléggé egyértelműnek látszanak.

Kérdés 14. Kaphatnánk némi ízelítőt abból, hogy más eljárással összehasonlítva, mik a problémás esetek, és melyik eljárás hogyan érvényesül ilyen keretek között?

Kérdés 15. A *Hidden Naıve Bayes (HNB)* model bizonyult a leghatékonyabbnak a 4.9 táblázat alapján. Ehhez mi volt a fő hozzájárulása, mi az újdonság?

Az ***ötödik fejezet*** a többségi szavazás hatékony képelemzési megoldásairól szól. Ez tulajdonképpen a 4. fejezetben bemutatott boosting-szerű eljárások folytatása, általánosabb szintre emelése. Példának ugyanakkor marad a retinakép. Az elv valóban fontos, hogy ne általános súlyozott átlagot képezzünk, hanem a jó döntéseket próbáljuk összeválogatni. Az együttes kibővítésének a hatása, a változó szavazóközösség elemzése értékes eredmény.

Kérdés 16. Ugyanakkor, ez a fejezet nekem nem állt össze egybe: nem látom, mi a pontos állítás, mennyire jók az eredmények, és hova vezetnek az elvi állítások? A fejezet végére kellett volna egy alapos összefoglaló.

A hatodik fejezet a microaneurysm (MA) detektálásáról szól. Ez elvileg a többes alakfelismerést célozza meg. A problémának széles szakirodalma van, az eljárások a morfológiai megoldásoktól az alakelemzésig terjednek. Jelen esetben a megoldást a szimulált hűtéssel iterált optimalizálás jelenti.

Kérdés 17. Jó lett volna, ha kapunk valamilyen áttekintést arról, hogy miről is van szó pontosan, ez miért probléma, és hogy néz ki, miért és mi a feladat?

Kérdés 18. Mi a választott elvi megoldás, és ez a feladatot tekintve miért optimális?

A ***hetedik fejezet*** szintén „ensemble" alapú elemzésről szól, de most a diabetikus retinopátia vizsgálatára. Érdekes, hogy ez már a hatodik fejezetben is szerepelt, pl. a 6.14 táblázatban. Lehetséges, hogy a 6. és a 7. fejezet együttes kezelése, a módszerek összehasonlítása értékesebb leírást adott volna.

A Tézisekről:

Valamennyi tézisre igaz, hogy bár komoly matematikai apparátus áll mögöttük, és alapos kísérleti elemzési munka támasztja alá az eredményeket, a tézisek megfogalmazása kevéssé hatékony. Megtisztelne a jelölt, ha a védésen a téziseket tömör és lényegre törő formában mutatná be, és azt is jelezné, hogy a disszertációban hol találhatók a saját eredmények (képletszám stb. szerint, pontosan).

1. Tézis: Objektumok optimális újra mintavételezése: mindkét altézis értékes új eredmény:
   i. A Voronoi alapú mintavételezés (2. fejezet)
   ii. Görbék gráfelméleti alapú tömörítése (3. fejezet)
2. Tézis: Objektumok detektálása: az altézisek együtt új eredménynek tekinthetők.
   i. Egyobjektumos felismerés (4. fejezet). Ez alatt az OD-t érti, ami azért nagyon speciális objektum.
   ii. Többségi szavazás kiterjesztése térbeli modellekre (5. fejezet)
   iii. Többes megjelenésű alakzatok detektálása (6. fejezet). Kidolgozatlansága miatt ezt a pontot gyengének érzem.
3. Tézis: Összetett rendszerek a döntéstámogatásban (7. fejezet): Túlzásnak tartom az utolsó fejezetre alapozva a két altézist. A két altézis együtt a 2. Tézis egy értékes kiegészítő altézise lehet:
   i. Új döntéstámogató modell
   ii. Sajátsági jellemzők és előfeldolgozók kidolgozása.

Összefoglaló: A tézisek és a hozzá tartozó disszertáció gondos és szép munka, nemzetközi szempontból is kiemelkedő eredményekkel.
A tézisek és a dolgozat gyengéje viszont, hogy az alapos munkát nem kíséri a saját eredmények egyértelmű kiemelése a disszertációban, és a tézisfüzetben megfogalmazott állításokból is hiányzik a célratörő megfogalmazás.
Fentiek nem kisebbítik azt a meggyőződésemet, hogy egy kiváló és szép munkáról van szó, kimagaslóan értékes publikációs eredményekkel. Különösen fontos, hogy a matematikai gondolkodást sikerült a mérnöki/orvosi problémák felé eredményesen alkalmazni.

Javaslom a *disszertáció nyilvános védésre bocsájtását, és az MTA doktora fokozat odaítélését.*

Budapest, 2016. Szeptember 5.

Szirányi Tamás