Válasz opponensi bírálatra

Opponens:

Dr. Kató Zoltán, MTA doktora, egyetemi tanár

MTA doktori értekezés címe:

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

Szerző:

Dr. Hajdu András

Tisztelt Prof. Dr. Kató Zoltán!

Mindenek előtt szeretnék köszönetet mondani az értekezésem bírálatára szánt idejéért és annak nagyon alapos véleményezésért. Az alábbiakban idézem az opponensi kérdéseket, illetve az általam megválaszolandónak érzett megjegyzéseket, majd az erre adott válaszaimat.

1) "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"

Válasz:

Az opponensi észrevétel jogos, mivel az értekezésben – részben terjedelmi okok miatt is – igyekeztem az általam közismertebbnek feltételezett fogalmak, eredmények hivatkozását mellőzni. Ennek megítélése természetesen részben szubjektív, illetve fennállt annak a hibának a lehetősége, hogy egyes hivatkozások kimaradnak. Az opponens által említett hiányosságokkal kapcsolatban:

3.2: Eulerian curve, Floodfill8 algorithm

A szakirodalom explicit módon nem ismeri az *Eulerian curve* (Euler-görbe) fogalmát, azaz itt hivatkozást megadni nem igazán lehetséges. Az értekezésben felvázolt és hivatkozott módon felhasználom a gráfelméletben használt Euler-tulajdonságot, beidézem az *abstract curve graph* (absztrakt görbegráf) fogalmát, amelyeket Euler-görbeként kapcsolok össze. Mivel a fogalom így nem elterjedt, valószínűleg hasznos lett volna ezt az eljárást egy magyarázó mondattal még tisztázni az értekezésben.

A Floodfill8 algoritmus hivatkozása az első előfordulásnál (3.1.1. szakasz, 47. oldal) az értekezésben az alábbiak szerint szerepel: "… we can use the recursive Floodfill8 algorithm [124]"

• 5.6.1: (5.59) – (5.62) mértékek

A mértékek definíciójának ismertetésekor valóban nem szerepel az explicit irodalmi hivatkozás, viszont az (5.59) egyenlet előtti bevezető bekezdésben megadásra kerültek a megfelelő

szakirodalmak: "In the current literature (see e.g. [166, 174]) the following diversity measures are suggested..:"

• 6.1: Simulated annealing

A Simulated annealing (szimulált hűtés) algoritmus hivatkozása az első előfordulásnál (6.1 szakasz, 117. oldal) az értekezésben az alábbiak szerint szerepel: "Our search framework is based on simulated annealing [201], which is a stochastic search algorithm"

• 7.3.2: Osztályozó algoritmusok

Az értekezésben szereplő osztályozók tesztelésére a Weka programcsomag szolgált, így ennek hivatkozása célszerű lett volna: Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

Az egyes osztályozókhoz (Alternating Decision Tree, kNN, AdaBoost, Multilayer Perceptron, Naive Bayes, Random Forest, SVM) tartozó szakirodalom:

Freund, Y., Mason, L.: The alternating decision tree learning algorithm. In: Proceeding of theSixteenth International Conference on Machine Learning, Bled, Slovenia, 124-133, 1999.

Fix, E., Hodges, J.L. Discriminatory analysis, nonparametric discrimination: Consistency properties. Technical Report 4, USAF School of Aviation Medicine, Randolph Field, Texas, 1951.

Freund, Y.; Schapire, R.E. (1997), "A decision-theoretic generalization of on-line learning and an application to boosting", Journal of Computer and System Sciences, Elsevier, 55 (1): 119–139.

Paul Werbos (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. PhD thesis, Harvard University.

Russell, Stuart; Norvig, Peter (2003) [1995]. Artificial Intelligence: A Modern Approach (2nd ed.). Prentice Hall.

Ho, Tin Kam (1995). Random Decision Forests (PDF). Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282.

Cortes, C.; Vapnik, V. (1995). "Support-vector networks". Machine Learning. 20 (3): 273-297.

2) "Mindkét tesztesettel az a fő probléma, hogy a szerző nem hasonlítja össze az általa javasolt Chamferillesztésen 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."

<u>Válasz:</u>

Egyetértve az opponensi véleménnyel, ennek a munkának a fő motivációja egy új, chamfer-illesztést tekintő eljárásokban legkisebb hibával járó mintavételezés kidolgozása volt a sablonillesztés gyorsításaként. A vizsgálatok kiterjedtek arra is, hogy a mintavételezésből származó hiba a mintavételezés fokának függvényében milyen módon nő, így becsülhető a sablonillesztés pontosságának visszaesése. Így annak a kérdésnek a megválaszolása, hogy a javasolt módszer mennyire alkalmas például az emberi alak detektálására elsősorban arra vezethető vissza, hogy a chamfer-illesztés alapú módszerek mennyire

alkalmasak erre feladatra, mivel a javasolt mintavételező eljárás ezt egy adott mértékben becsülhető módon rontja. A chamfer-illesztésen alapuló detektáló módszerek elsősorban olyan alkalmazásokban voltak népszerűek, ahol kisebb számítási kapacitás állt rendelkezésre (például gépjármű fedélzeti kamerájának valós idejű feldolgozása gyalogosok detektálására), így a pontosság mellett a futási idő is kiemelten fontos volt. Munkámat alapvetően Dariu M. Gavrila módszerei és eredményei inspirálták, aki 1997-2013 között a Daimler cég számára fejlesztett gyalogos-érzékelő eljárásokat, és a szakterületen a mai napig aktív (lásd www.gavrila.net) folyamatos, kifejezetten nívós publikációs tevékenységgel a módszerek valódi alkalmazása mellett. Legelemibb vizsgálódásaimkor még Gavrila és szerzőtársa 1999-es [1] munkájára támaszkodtam. Az azóta eltel időben az látszott, hogy a chamfer-illesztésen alapuló módszer a mai napig megmaradt a fejlesztett rendszerekben, csak folyamatosan kiegészült további hatékony és gyorsan számolható sajátságokkal. Mindezen leírók egy összetett (ensemble-based) keretrendszerben kerülnek aggregálásra; lásd például [2]: "Our Mixture-of-Experts framework [24] for pedestrian classification combines four modalities (shape, intensity, depth, and motion) and three features (Chamfer distance, HOG, and LBP)." A chamfer-távolság, HOG és LBP sajátságokról az is kiderül ebben a munkában, hogy meglehetősen diverz viselkedésűek, ami az összetett rendszerek esetében kifejezetten jótékony tulajdonság. A közleményben szintén szerepel egy összehasonlítás, ami a fals pozitívok számának csökkenését igazolja az összetett rendszer használatakor egy korábbi (csak HOG sajátságokat és lineáris SVM osztályozót) tekintő eljárás esetén, mely [3] alapján kifejezetten versenyképes a területen; az elemzésekben használt adatbázist a szerzők publikussá tették (lásd http://www.gavrila.net/Datasets/datasets.html).

A fentiek alapján jelenleg nagyon nehéz lenne egy korszerű összetett gyalogos-detektáló rendszer pontosságával kapcsolatban nyilatkozni arról, hogy a mintavételezés alkalmazása mekkora pontosságvesztést eredményezne a teljes rendszer szintjén. Mindenesetre, az összehasonlító eredmények azt sugallják, hogy a módszernek helye van a szakterületen, egy összetett rendszer egy lehetséges komponenseként mindenképpen. A mintavételezés viselkedésének valós hatását – az opponenssel egyetértve – kifejezetten érdekes lehet a jövőben megvizsgálni az elérhető adatbázisokon, valószínűsíthető azonban, hogy önmagában a chamfer-illesztés nem lenne már versenyképes. Az egyszerű empirikus tesztelés mellett izgalmas elméleti kérdéseket vethetne fel itt az, hogy a sablonok hierarchikus klaszterezéssel fába való rendszerezését milyen mértékben módosítja a mintavételezés.

A mintavételező eljárásnak az érhálózat detektálásában, mint egy másik lehetséges alkalmazásban, való felhasználása már lényegesen messzebb esik a klasszikus chamfer-illesztéstől. Itt az alapgondolat az volt, hogy a képpont-szintű leírók meghatározásakor az egyes osztályok (érpont/háttérpont) jobb/egyenletesebb reprezentációjához célszerű lehet irányított módon végezni a mintavételezést, mert például a határon lényegesen változatosabb sajátságvektorok jelenhetnek meg, így a tanulásokhoz relatíve több információ szükséges. Magát a mintavételezést pedig az indokolja, hogy ellenkező esetben nagyon magas lenne a képpontszám a tanulóeljárások számára. Az osztályok egyenletesebb reprezentációja természetesen a sajátságtérben is analizálható lenne a tanulóvektorok segítségével, viszont ennek a vizsgálatnak éppen az volt a motivációja, hogy hasonló eredmény a képtartományban is elérhető-e geometriai megfontolások alapján. A képpont-szintű érhálózat-detektorok versenyképesek a területen (lásd például Gyorgy Kovacs, Andras Hajdu: A Self-Calibrating Approach for the Segmentation of

Retinal Vessels by Template Matching and Contour Reconstruction, *Medical Image Analysis* **29** (2016 Apr), 24-46).

[1] D. Gavrila and V. Philomin, "Real-time object detection for smart vehicles," in IEEE International Conference on Computer Vision (ICCV), vol. 1, 1999, pp. 87–93.

[2] M. Enzweiler and D. M. Gavrila, "A Multilevel Mixture-of-Experts Framework for Pedestrian Classification," in IEEE Transactions on Image Processing, vol. 20, no. 10, pp. 2967-2979, Oct. 2011.

[3] Enzweiler M, Gavrila DM, Monocular pedestrian detection: survey and experiments, IEEE Trans Pattern Anal Mach Intell. 2009 Dec;31(12):2179-95.

3) "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)."

Válasz:

A bemutatott görbetömörítő eljárás valóban veszteséges, ezért a JBEAM [1] tömörítő eljárással való összevetés is úgy volt korrekt, hogy a JBEAM veszteséges tömörítést adó paraméterezésével lett hasonlítva (a JBEAM veszteségmentesen is tud tömöríteni, de a szerzők javaslata a veszteséges tömörítés).

A javasolt görbetömörítő tömörítési minőségének vizsgálatát illetően a legfontosabb szempont, hogy egy görbeszakasz feldolgozása egy online lineáris szakaszokra bontó algoritmussal történik, melynek kimeneteit Bresenham-szakaszok egy ábécéjéből helyettesítjük. Ez az eljárás bizonyítottan garantálja, hogy az eredeti és tömörített görbe távolsága mindenhol 1 képponton belül marad, mivel az eredeti felbontásával nyert egyenes szakaszok és azok Bresenham-reprezentációinak a végpontja megegyeznek. A JBEAM esetében a szerzők által javasolt tömörítési szinten nincsen garancia az eredeti görbétől való 1 képponton belüli távolságra, ahogyan ezt az alábbi példa is mutatja ([1]-ből idézve):





Figure 6: (a) Country borders. We use them as our test images. (b) Reconstruction of maps in the lossy beamlet coding.

Az MPEG-4 kontúrkódolással egyszerűbb a minőségi összevetés. Nevezetesen – ahogyan ez az értekezésben is szerepel (lásd 3.4.2. szakasz, 53. oldal) – az MPEG-4 poligonfelosztáson alapuló approximációs algoritmusa akkor eredményez lineáris görbedarabokat, ha a felosztás során használt távolságküszöb sqrt(2)/2. Az MPEG-4 a minőség rovására nagyobb tömörítési hatékonyságot ér el a küszöb növelésével, és a tesztek alapján a javasolt módszer tömörítési hatékonysága kb. az 1 küszöbértéknél egyezett meg az MPEG-4 hatásfokával (miközben a javasolt módszer betartotta a lineáris tulajdonságot, azaz jobb minőséget biztosított). Ahogyan ez szintén szerepel az értekezésben, az MPEG-4 szabvány –

pszichológiai tesztek alapján – 1,4-es küszöbértéket jelöl meg még vizuálisan elfogadhatóként, amiből az is következik, hogy a javasolt eljárás tömörítési minősége stabilan az MPEG-4 iránymutatása felett van.

[1] Huo, X., Chen, J., Donoho, D. L.: JBEAM: Coding Lines and Curves via Digital Beamlets, in IEEE DATA COMPRESSION CONFERENCE; 449-458.

4) "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."

Válasz:

Egyetértve a bírálói kritikával, valóban elmondható, hogy a 4.1.2 szakaszban bemutatott eljárás módszertanilag egyszerű. Mindenképpen figyelembe kell venni azonban, hogy időrendben ez volt a legelső fúziós modell, melynek gyakorlati sikeressége erősen motiválta a későbbi elmélyültebb vizsgálatokat a további javításokhoz az elméleti alapok lefektetésével együtt. A tárgyalt fúziós módszertanokban vizsgálataimban mindig fontos szerepet kapott, hogy a paraméterek beállítása lehetőleg ne kézzel, hanem valamilyen optimalizációs eljárás segítségével történjen. Az opponens által említett paraméterek a 4.1.2 szakaszban azonban kevésbé szabadon változtathatóak, mivel alapvetően humánbiológiai értékekhez kapcsolódnak. Nevezetesen, a vakfolt átmérőjeként használt d_OD paraméter várható értéke a klinikumban pontosan definiálható egy egészséges ember vakfoltjának átmérőjeként. A retinaképeken ez a paraméter a digitális képekre a látószög (FOV) alapján nagy biztonsággal kalkulálható. Ez a becslés empirikusan elégségesnek bizonyult. Hozzátehetjük azonban, hogy a FOV ismeretének hiányában is megoldható a paraméter precíz beállítása, mert például klinikai metrikus adat ismert a vakfolt alatt megfigyelhető legvastagabb kilépő ér átmérőjére, melynek detektálásával a vakfolt átmérője képpontban is megadhatóvá válik.

A másik paraméter a szavazategyezés esetében hozott döntési mechanizmusban kerül felhasználásra a kép egy erőteljes Gauss simításában. A paraméter beállítása itt valóban empirikusan történt, viszont értéke elég nagy skálán mozogható, mivel a cél lényegében egy erős elmosás.

A kérdésre adott választ összegezve tehát elmondható, hogy nem is annyira a paraméterek beállítása okozhat gondot más típusú képekre való alkalmazhatóság esetén, hanem a teljes fúziós modell adaptálása. Ez elsősorban a detektálni kívánt objektum megfelelő geometriai paraméterekkel való reprezentációjának feladatát hordozza magában.

5) "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."

Válasz:

Az opponensi felvetés jogos, hiszen a (4.8) formula nem veszi figyelembe a részt vevő algoritmusok várható pontosságát és függőségi viszonyait, szemben például a (4.7) eljárással. A nagyon egyszerű válasz a kérdésre természetesen az, hogy jelentős különbséget nem tapasztaltunk az egyszerűbb formula használatakor, különös tekintettel arra, hogy ez a hiba az objektum detektálásának helyességét már nem befolyásolja, csak annak pontos középpontja meghatározását, mely információ nem volt kritikus az

alkalmazásainkban. Mindazonáltal az alábbi szempontok lényegesen tovább árnyalják ezt az elsőre triviális választ:

 i) A (4.7) formula alapján történő számolás meglehetősen impraktikus, mivel a páronkénti függőségek (kovarianciák) számolása időigényes, különösen nagyobb számú algoritmus esetében. Valós alkalmazásban szintén nehézkes lehet az eljárás, mert például egy új algoritmus felvétele esetén a döntéshez mindig ismételni kell a teljes kiértékelést.

ii) Elméleti szempontból is kardinális problémát jelent, hogy bár a (4.7) eljárás szakirodalmi ajánlás, ez a lépés önmagában kevés a rendszer teljes függőségének feltárására, amihez a páronkénti függőségek megállapítása még nem elegendő. Ezért, mivel a függőségek feltárása a fúziós rendszerek gyakorlati használhatóságának szempontjából kritikus, ezt a kérdéskört az értekezésben is tárgyalt módon később lényegesen alaposabb vizsgálatnak vetettük alá (lásd 5.2 és 5.6 szakasz). Ez a vizsgálat a mai napig folyik, elsősorban hatékonyabb diverzitásmérők bevezetésével. Az opponens által hiányolt vizsgálat így reményeink szerint a függőségi viszonyok egy hatékonyabb feltárása mellett teljesedik ki.

6) "A 4.2. táblázatban közölt eredmények esetében nem világos, hogy OD5 miért ad konzisztensen jobb eredményt a javasolt halmaz-alapú módszernél. A 4.1 táblázat szerint pedig a javasolt módszer alig teljesít jobban, mint OD4. Mi ennek a magyarázata?"

Válasz:

A fúziós rendszerek elméleti eredményekkel is alátámasztható előnye, hogy az egyedi megközelítéseknél általában nagyobb stabilitást mutatnak, mely tulajdonságuk igen népszerűvé tette őket számos alkalmazás körében. A retinafelvételek elemzésénél ez a megállapítás szintén helytállónak bizonyult, melyben fontos szerepe lehet a kapcsolódó képi adatok szakterületen jól ismert meglehetősen heterogén voltának. Gyakorlati szempontból további ingadozást jelenthet az egyedi algoritmusok teljesítménye esetén, ha a fejlesztés során módosul egy olyan eljárás, amire az adott algoritmus ráépül. Így például, ha a feldolgozási idővel való takarékoskodás jegyében egyetlen érhálózat-detektort használunk az összetett rendszerben és arra vakfolt (OD) detektáló algoritmusok épülnek, akkor azok egyedi pontossága változhat az érhálózat-detektor cseréjével. A fúziós rendszerek előnye, hogy ezeket az anomáliákat is képesek kiegyenlíteni.

A 4.2. táblázatban az OD algoritmusok másodlagos hibája van összefoglalva, nevezetesen, hogy mekkora az általuk adott jelölt és a kézzel annotált középpont átlagos euklideszi távolsága. Ezt a másodlagos hibát azonban csak úgy korrekt számolni, ha a jelölt kellően közel volt az annotált középponthoz (azaz elsőrendű hiba szerint megtalálta az OD-t, vagyis beleesett a kézzel jelölt OD területbe), egyébként óriásivá nőhet. Más szóval, mivel nem szándékoztuk az elsődleges detektálás szempontjából pontosabb összetett rendszert ilyen módon előnybe hozni, minden egyedi algoritmusnál – így az OD5-nél is – az elsődlegesen hibás jelölteket kihagytuk a másodlagos hiba számításakor. A 4.1. táblázatból megfigyelhető, hogy az OD5 elsődleges detektálási pontossága 15-20%-kal volt gyengébb a vizsgált adatbázisokon az összetett rendszernél, így ennyivel kevesebb jelölt alapján került a másodlagos hibája is meghatározásra. Összességében tehát kijelenthető, hogy ha még a pontos középpont kijelölésénél a mért adatbázisokon árnyalatnyival nagyobb is a pontossága, az OD5 semmiképpen sem tekinthető hatékonyabb detektornak az összetett rendszernél, mivel pontossága jelentősen elmarad attól a sokkal lényegesebb elsődleges hibára nézve.

Az OD4 algoritmus 4.1. táblázatban közölt, az összetett rendszer pontosságát közelítő teljesítménye valósan mért adat, mellyel kapcsolatban jogosan merült fel az opponensi kérdés. Az értelmezéssel kapcsolatban azonban szeretnék visszautalni az összetett rendszerek nagyobb stabilitásával kapcsolatos korábbi fejtegetésre. A közölt adatbázisokon túl számos más – nagyobb heterogenitást mutató – adaton is folyamatosan teszteljük a rendszert, és míg az összetett rendszerek nagy megbízhatósággal dolgoznak, az egyedi algoritmusok teljesítménye meglehetősen nagy változatosságot mutat (például egy-egy befolyásoló, más típusú detektor módosításakor). Az ebben a szakaszban tárgyalt megközelítések kronológiájához hozzátartozik továbbá, hogy az azonos típusú (OD/MAC) detektorok független fúziós rendszereit azok geometriai alapú aggregálásával kívántuk a detektálás hatékonyságát tovább javítani. Ez a törekvés a 4.5. táblázat alapján sikeresnek is bizonyult, így az összetett rendszer és az OD4 algoritmus közötti pontosságkülönbség tovább nőtt.

7) "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."

Válasz:

A javasolt modell összetett rendszerek összeállítására valóban az adott alkalmazásra nézve került bemutatásra (mikroaneurizma detektálása). Mindazonáltal, ugyanezen szűrőrendszeren belül a módszer hatékonyságát sikerült más megjelenési formával rendelkező objektum (exudátum) detektálására is igazolni, ahogyan ez az értekezés 7.1.2 szakaszában röviden megjegyzésre is kerül. A használt megközelítést más – jelenleg kidolgozás alatt álló – klinikai szűrőrendszerekben is teszteljük és várhatóan alkalmazni fogjuk. A kapcsolódó elméleti igényű vizsgálatok viszont igen komoly kihívásnak ígérkeznek.

8) "Az eredményekkel kapcsolatban csupán annyi (inkább a prezentációra vonatkozó) kritikát fogalmaznék meg, hogy 113. oldalon tárgyalt szimulált hűtés algoritmusánál illett 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 elvész a konvergencia elméleti garanciája. Ugyanakkor a gyakorlatban elterjedt ún. exponeciális ütemező, melyet vélhetően a szerző is használ, csak akkor ad jó eredményt, ha a disszertációban q-val jelölt szorzófaktor közel 1!"

Válasz:

Jogos a kritika a szimulált hűtés valamivel mostohább kezelését illetően az értekezésben közölt eredményekben. Az adott munkában kisebb figyelmet szenteltünk az optimumkeresés hatékonyságának, lévén a fő eredményben ez a kérdéskör kisebb súlyt kapott. Számunkra itt a szimulált hűtés egy egyszerű eszköz volt, amit a szakterület nagyon jól ismer és széles körben alkalmaz. Választásunk azonban nem volt teljesen *ad hoc*, mivel a szimulált hűtést gyakran alkalmazzák diszkrét problémákra is, és a paraméterezésben követtük a minimális szakterületi – beleértve a globális konvergenciára vonatkozó – elvárásokat, így az általunk valóban használt exponenciális ütemező szorzófaktora 1-hez közeli (0,975) volt. Távolról sem biztos továbbá, hogy a szimulált hűtés az adott esetben a leghatékonyabb kereső heurisztika. Terveink között szerepel, hogy a jövőben jobban feltárjuk fúziós rendszereink energiafüggvényeinek viselkedését, és ehhez kapcsolódóan a szimulált hűtésnél hatékonyabb optimalizációs eljárásokat válasszunk ki a szakterület tapasztalt kollégáinak segítségével, akik szintén felhívták erre a lehetőségre a figyelmünket.

9) "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."

Válasz:

A 7) opponensi megjegyzésre adott válaszomhoz hasonlóan most is elmondható, hogy a javasolt fúziós modell valóban az adott alkalmazásra nézve került bemutatásra (diabéteszes retinopátia automatikus szűrése). A használt megközelítést más – jelenleg kidolgozás alatt álló – klinikai szűrőrendszerekben is teszteljük és várhatóan alkalmazni fogjuk értelemszerűen kiegészítve azt az azóta megjelent korszerű technikák (például mélytanulás) nyújtotta lehetőségekkel.

Az opponens joggal kifogásolja továbbá, hogy egyes kapcsolódó módszerek részletes bemutatása nem szerepel az értekezésben. Mindazonáltal, az elsőként említett módszer (morfológia-alapú sárgafolt detektáló) az értekezésben a MAC5 néven hivatkozott detektorként az 1.6.3 szakaszban tömören bemutatásra kerül (17.oldal), a második eljárás (rejtett Markov mezőn alapú érhálózat szegmentáló) az 1.7.1 szakasz "Quality assessment" bekezdésében leírt, míg a harmadik módszer (intenzitásprofilok vizsgálatán alapuló mikroaneurizma detektáló) az 1.6.4 szakaszban MA5 néven hivatkozott detektornak felel meg. Összességében azonban valóban elmondható, hogy terjedelmi okok miatt az automatikus szűrőrendszerben felhasznált saját fejlesztésű egyedi algoritmusok mélyebb bemutatását kénytelen voltam mellőzni az értekezésben. Igyekeztem egyúttal ilyen módon kevésbé megtörni a fúziós modellek vizsgálata által nyújtott kohéziót, mely modelleken alapuló rendszerekben a nevezett algoritmusok mégis fontos szerepet töltenek be.

Az opponensi véleményre reagálva az értekezésben leírtnál valamivel részletesebb módon most bemutatom a három említett módszert. Válaszom terjengősségét elkerülendő, a további részletek feltárását az eredményeket részletesen bemutató, az értekezésben is hivatkozott közlemények megjelölésével teszem, melyeket válaszomhoz függelékként csatolok is.

i) Morfológia-alapú sárgafolt detektáló [1]: Ebben a munkában bemutatásra kerül egy elsősorban orvosi képfeldolgozásban használt detektorok hatékonyságának növelését szolgáló megközelítés, ami az

algoritmusok paramétereinek finomhangolását megfelelő tesztadatbázisokon való kiértékeléssel végzi el. A javasolt eljárás egy magas dimenziós paraméterterek esetében is hatékony sztochasztikus keresésen alapul. Megközelítésünk hatékonyságát saját klinikai alkalmazásunk vizsgálatán keresztül végezzük. Nevezetesen, beszámolunk biztató eredményeinkről a sárgafolt automatikus detektálását illetően retinaképeken, ahol a javasolt módszer alkalmazásának segítségével detektorunk négy másik korszerű sárgafolt-detektornál bizonyult pontosabbnak. A sárgafolt detektálásához elsőként kiválasztjuk a bementi RGB kép zöld csatornáját, majd előfeldolgozásként ebből kivonjuk annak egy medián-szűrt változatát. Az így előállt háttérképet küszöböléssel binarizáljuk, majd kijelöljük a legnagyobb komponenst a sárgafolt végső jelöltjeként. Eljárásunk során számos további hagyományos morfológiai eljárást tekintünk kisebb korrekciókként. A detektor két szabad paramétere (medián-szűrő mérete és binarizálási küszöb) sztochasztikus optimalizációval kerül beállításra.

ii) Rejtett Markov mezőn alapú érhálózat szegmentáló [2]: A retina érhálózatának pontos szegmentálása jelenleg egy sokak által vizsgált terület. A feladat sikeres megoldása esetén ugyanis nagyot lépünk előre a retinaképek alapján működő automatikus szűrőrendszerek napi klinikai rutinba való bevezethetősége felé. Nevezetesen, az érhálózat helyének pontos ismeret nagy segítséget nyújt más kritikus anatómiai képletek (sárgafolt, vakfolt, fő érívek) és elváltozások (mikroaneurizma, vérzések) lokalizálásához, továbbá, az erek morfológiájának torzulása közvetlenül utal egyes betegségek jelenlétére (például magas vérnyomás). Ebben a munkánkban egy új, rejtett Markov mezőkön alapuló eljárást vezetünk be a retina érhálózatának szegmentálására. A rejtett Markov mező modellek optimalizálási problémáját a kép érintővektor-mezője alapján kiterjesztjük az elnyújtott elemekből felépülő érhálózat teljes összefüggőségének erősítéséhez. A valószínűségi becslés javításához a rejtett Markov probléma megoldásakor a hagyományosan használt Bayes becslés helyett az ú.n. Averaged One-Dependence becslést (AODE) használjuk, mivel az a sajátságok teljes függetlenségénél gyengébb feltételt tekint. Módszerünk empirikus, kvantitatív kiértékelését egy publikusan elérhető adatbázison végeztük el.

iii) Intenzitásprofilok vizsgálatán alapuló mikroaneurizma detektáló [3]: A retinaképek elemzésében diagnosztikai szempontból az egyik legfontosabb és legnagyobb kihívást hordozó terület a mikroaneurizmák (MAk) felismerése. Így a diabéteszes retinopátia (DR) feltárását célzó automatikus szűrőrendszerek is nagyban hagyatkoznak ezen elváltozások szegmentálására. Ebben a munkában eljárást adunk egy, az MAk jelenlétét mutató pontozókép meghatározására. A vélhető pozíciókat a magasabb valószínűségi csúcsok jelzik, amelyeket egy képponton áthaladó, különböző irányú intenzitásprofilokból származtatunk kihasználva az MA alapvetően körszerű jellegét. A végső MA jelöltek bináris küszöböléssel nyerhetők ki a pontozóképből. A legtöbb elérhető MA detektorral szemben ez a módszer nem igényel tanítást. Mindazonáltal, a gyakorlati tesztek alapján a javasolt eljárás versenyképes a területen, különösen az érhálózathoz közel eső mikroaneurizmák detektálásakor. Algoritmusunkat egy publikusan elérhető online versenyen is teszteltük.

[1] B. Antal and A. Hajdu, "A stochastic approach to improve macula detection in retinal images," Acta Cybernetica, vol. 20, pp. 5–15, 2011.

[2] G. Kovacs and A. Hajdu, "Extraction of vascular system in retina images using averaged onedependence estimators and orientation estimation in hidden markov random fields," in IEEE International Symposium on Biomedical Imaging (ISBI), 2011, pp. 693–696. [3] I. Lazar and A. Hajdu, "Microaneurysm detection in retinal images using a rotating crosssection based model," in IEEE International Symposium on Biomedical Imaging (ISBI), 2011, pp. 1405–1409.

Befejezésül, még egyszer köszönöm az opponens munkáját és az értekezés nyilvános védésére bocsátásának, valamint a fokozat odaítélésének támogatását.

Debrecen, 2016. november 30.

Tisztelettel,

.....

Dr. Hajdu András egyetemi docens

Függelék

9) opponensi megjegyzésre adott válaszhoz

A Stochastic Approach to Improve Macula Detection in Retinal Images

Bálint Antal and András Hajdu^{*}

Abstract

In this paper, we present an approach to improve detectors used in medical image processing by fine-tuning their parameters for a certain dataset. The proposed algorithm uses a stochastic search algorithm to deal with large search spaces. We investigate the effectiveness of this approach by evaluating it on an actual clinical application. Namely, we present promising results with outperforming four state-of-the-art algorithms used for the detection of the center of the sharp vision (macula) in digital fundus images.

Keywords: biomedical image processing, simulated annealing, learning and adaptive systems

1 Introduction

Diabetic Retinopathy (DR) is the most common cause of blindness in the developed countries. Nowadays, the automatic screening of DR received much attention in the medical imaging community [1], [7], [9], since replacing a resource-demanding and expensive manual screening is a very challenging task. Automatic screening is based on the analysis of retinal images taken at eye hospitals. One class of the difficulties originates from the use of different kinds of retinal images, which leads to varying performance in the anatomy or lesion detection processes. Some detectors are based on machine learning, while others consider non-training approaches.

In this paper, we present a technique to improve a detection algorithm on retinal images via a learning-based approach. The idea behind this technique is to finetune the parameter setup for a certain detector. Since the selection of the optimal parameter setup usually traverses a large search space, we decided to use a stochastic approach, simulated annealing for this task. To demonstrate the effectiveness of this technique, we present a novel macula detector and show that the proposed framework improves detection performance. The contribution in this particular area is justified by the fact that the detection of macula involves the lowest number of reported works in the field of DR screening research [16]. A comparative analysis

DOI: 10.14232/actacyb.20.1.2011.2

^{*}University of Debrecen, Hungary. E-mail: {antal.balint, hajdu.andras}@inf.unideb.hu

reveals that our tuned algorithm outperforms other state-of-the-art algorithms in the field.

The main contributions of the paper are organized into the following sections:

- 1. A stochastic approach to improve detector performance is introduced. We also discuss the advantages of using simulated annealing over stochastic hill climbing (Section 2).
- 2. We show how to adopt the simulated annealing based search method to improve the performance of the proposed macula detector (Section 3).
- 3. A novel macula detector is proposed, which, in addition to its good performance, can be easily fine-tuned by a search algorithm (Section 3.1).
- 4. We define an error measure to efficiently characterize macula detection performance (Section 3.3).
- 5. We evaluate the performance of our macula detector using the proposed tuning also in comparison with four state-of-the-art algorithms (Section 4).

2 A stochastic approach to improve detector performance

In this section, we present our approach to select an optimal parameter setup for a detector algorithm. For this task, we have to prepare for a large search space, since these algorithms may operate with several parameters. In literature, stochastic hill climbing is often recommended [10] [13]. Stochastic hill climbing is based on the idea that using random jumps between the neighbouring elements of the search space converges faster to the extrema than using exhaustive enumerations. An element is accepted, if it provides better result than the current extremum. This approach is an effective solution for many problems, but it can get stuck in a local extrema in search spaces with many peaks.

To overcome this difficulty, we used a simulated annealing-based method. Simulated annealing [5] avoids getting stuck in local extrema by using a random acceptance function for rejected elements. That is, if an element does not provide a better result than the current one, it is still accepted if the acceptance function allows that. See Figure 1 for a visual comparison of hill climbing and simulated annealing.

The formal description of the algorithm can be given as follows:

Algorithm 1.: Parameter setup selection by simulated annealing.

Input:

- An initial temperature $T \in \mathbb{R}$.
- A minimal temperature $T_{min} \in \mathbb{R}$.



Figure 1: The path of the hill climbing and the simulated annealing algorithm is represented with gray and black colors, respectively. While hill climbing reaches only the local optimum, simulated annealing can continue towards the global optimum by using chaotic jumps.

- A temperature change $q \in \mathbb{R}$ with $(0 \le q \le 1)$.
- A search space $S \subset \mathbb{R}^n$ with $s \in S$ is a parameter setup.
- A function r(X), which chooses a random element x from a set X.
- A function $accept : \mathbb{R}^4 \to \{true, false\}$, which is defined as the follows:

$$accept (e, e_i, T, y) = \begin{cases} true, & if \exp\left(\frac{e-e_i}{T}\right) > y, \\ false, & otherwise. \end{cases}$$

• An energy function $E: S \to \mathbb{R}$.

Output: $s_{optimal} \in S$, where $E(s_{optimal}) = \min_{s \in S} E(s)$. That is, $s_{optimal}$ is the parameter setup minimizing the energy function E.

1. $s \leftarrow r(S)$ 2. $e \leftarrow E(s)$ 3. $S \leftarrow S - \{s\}$ 4. while $S \neq \emptyset$ or $T < T_{min}$ do 5. $s_i \leftarrow r(S)$ 6. $e_i \leftarrow E(s_i)$ 7. $S \leftarrow S - \{s_i\}$ 8. if $e_i < e$ then 9. $s \leftarrow s_i$

```
10.
            e \leftarrow e_i
            T \leftarrow T \cdot q
11.
12.
        else
            y \leftarrow r(\mathbb{R})
13.
            if accept(e, e_i, T, y) = true then
14.
               s \leftarrow s_i
15.
16.
               e \leftarrow e_i
               T \leftarrow T \cdot q
17.
18.
            end if
19.
        end if
20. end while
21. return s
```

3 Using the proposed approach: an example

In this section, we present an example to demonstrate the power of the proposed method. For this task, we chose a novel approach for macula detection in retinal images, which algorithm requires only two parameter to be optimized. Our proposed approach for obtaining the optimal parameter setup can be adapted to any similar problem, as well.

3.1 Macula detection

The macula is the central region of sharp vision in the human eye, with its center referred to as the fovea (see Figure 2). Any lesions (e.g. microaneurysms) which appear here can lead to severe loss of vision. Therefore, the efficient detection of the macula is essential in an automatic screening system for diabetic retinopathy.

3.2 A novel algorithm for macula detection

In this section, we present a novel approach to detect macula in a retinal image. As we can see later on, this algorithm outperforms state-of-the-art macula detectors with the use of the proposed framework for optimal parameter setup.

The proposed macula detection algorithm can be formulated as follows:

Algorithm 2.: A novel macula detector

Input:

- A digital retinal image I in 24 bit RGB format.
- A parameter $q \in \mathbb{R}$ with $0 \leq q \leq 1$ to adjust of the mask size in the median filtering step.
- A threshold $t \in [-255, \ldots, 255]$.

8



Figure 2: A sample fundus image with the main anatomical parts annotated.

Output: An image containing the macula region of the eye.

- 1. Extract the green intensity channel G from I.
- 2. Let $A = [Min(width(I), height(I)) \cdot q].$
- 3. Produce image M with the same size as G by applying median filtering [12] on G with a mask size $A \times A$.
- 4. Create the difference image D = G M.
- 5. Produce a binary image B by assigning all pixels with larger intensity than t in the D to the foreground, while the rest to its background.
- 6. Select the largest binary component to locate the macula.

The results after each step of the algorithm can also be observed in Figure 3.



Figure 3: Steps of the proposed macula detection: (a) The green channel of the input image. (b) The result of the median filtering. (c) The difference image. (d) The binary image after thresholding and largest component selection.

3.3 Error measurement of macula detectors

To select the optimal parameter setup for the above detector algorithm, we need a proper energy function to be minimized. An obvious choice for this task is to minimize the distance of the centroid of the macula found by the detector and the manually selected center of the macula for each image in a dataset. To avoid overtraining, we also take into account the distance from the optic disc (see Figure 2), as the macula and the optic disc are spatially constrained [11].

Thus, we define the following energy function for this problem:

$$E = \sum_{I \in DS} d\left(M_{alg}\left(I\right), M_{hm}\left(I\right)\right) + \sum_{I \in DS} \left|d\left(M_{alg}\left(I\right), O\right) - MO_{avg}\right|,$$

where

- DS is the dataset,
- d denotes the 2D Euclidean distance,
- M_{alg} is the centroid pixel of the detected macula,
- M_{hm} is the manually selected macula center,
- O is the manually selected optic disc center,
- MO_{avg} is the average Euclidean distance of the manually selected macula and optic disc center for the dataset DS.

4 Comparative results

We evaluate our method by comparing it with four other state-of-the-art macula detectors (Section 4.1) on different datasets (Section 4.2). As our results will show, the novel macula detector outperforms the others after finding its optimal parameter setup.

4.1 State-of-the-art macula detection algorithms

In this section, we list four macula detection algorithms, which are involved in our comparative analysis. The parameters of the algorithms were set according to the corresponding recommendations in literature.

4.1.1 Petsatodis et al. [8]

In [8] a region of interest (ROI) is defined to process macula detection. A Gaussian low-pass filter is applied to smooth the image. The statistical mean and standard deviation of the ROI area are used to compute a threshold for segmentation to get binary objects. The object that is located nearest to the center of the ROI is labelled as macula. Its center of mass is considered to be the center of the macula. However, we did some modification to this approach, because it is not mentioned how this ROI is defined; therefore we applied the smoothing to the whole image using a large kernel (70 × 70 pixels with $\sigma = 10$) so that vascular network and small patches do not interfere in detection. Then, an iterative thresholding process is launched to generate a set of binary images corresponding to different threshold values. In each binary image, the component satisfying the area and distance from the center constraints are identified, and the component found nearest to the center with minimum area is marked as macula.

4.1.2 Sekhar et al. [11]

In [11] a region of interest (ROI) for macula is defined regarding its spatial relationship to the optic disc. That is the portion of a sector subtended at the center of the optic disc by an angle of 30° above and below the line between this center and

the center of the retinal image disc. The macula is identified within this ROI by iteratively applying a threshold, and then applying morphological opening (erosion followed by dilation) on the resulting blob. The value of the threshold parameter is selected such that the area of the smoothed macula region is not more than 80% of that of the detected optic disc. The fovea is simply determined as the centroid of this blob.

4.1.3 Fleming et al. [2]

Fleming et al. [2] proposed to identify the macular region based on the information of the temporal arcade and OD center. First, the arcade is found by using semielliptical templates. Next, the optic disc is detected by using Hough transformation with circular templates having diameters from 0.7 to 1.25 OD diameter (DD). Finally, the fovea was detected by finding the maximum correlation coefficient between the image and a foveal model. The search was restricted to a circular region with diameter 1.6 DD centered on a point that is 2.4 DD from the optic disc and on a line between the detected optic disc and the center of the semi-ellipse fitted to the temporal arcades.

4.1.4 Zana et al. [17]

Zana et al. [17] presented a region merging algorithm based on watershed cell decomposition and morphological operations for macula recognition. After noise removal, morphological closing followed by opening is performed to remove the small dark holes and white spots. A watershed based decomposition of the gradient image into cells is done, and the cell with darkest gray level inside the macula is selected as the first step of a merging algorithm. A complex criterion based on the gray values and of edges of the filtered image is calculated to merge the cells of the macula, while rejecting perifoveal inter-capillary zones in order to produce the contour of the macula.

4.2 Datasets

We have tested our approach on 199 images from three publicly available data sources: DiaretDB0 [3], DiaretDB1 [4] and DRIVE [15]. The characteristic properties of these datasets can be seen in Table 1. We have selected the optimal parameter setup for each dataset using a separate training subset of a total of 60 images (20 images from each dataset). For each dataset, the ground truth are used only for parameter selection.

4.3 Results

Table 2 shows the selected optimal parameters for each dataset. The size parameter q and the threshold parameter t have been found by the proposed stochastic approach. Each dataset performed optimally using a different parameter setup. We have evaluated our approach in two aspects [6]: whether the detected macula

Dataset	Images	Normal	DR	FOV	Resolution
DiaretDB0	130	20	110	50	$1500 \mathrm{x} 1152$
DiaretDB1	89	5	84	50	$1500 \mathrm{x} 1152$
DRIVE	40	33	7	45	768x584

Table 1: Properties of the datasets.

Dataset	q	t
DiaretDB0	0.6	0
DiaretDB1	0.6	5
Drive	0.7	0

Table 2: Parameters selected by the proposed algorithm for macula detection.

center falls into the 0.5DD (Optic Disc Diameter) distance of the manually selected macula center and we also measured the Euclidean distance of them (calculated on normalized images). Table 3 and 4 contain the quantitative results using these measures, respectively. We disclose the results for each macula detector evaluated in all dataset. For the more straight-forward comparison, we also calculated the simple average of these performance values. In the terms of the first measure, the use of the proposed algorithm on the novel macula detector resulted in a 85% average accuracy, while the second best method only earned 77%. However, in the terms Euclidean error it is only third in the comparison, mainly because of its difficulties on the DRIVE database.

Dataset	Petsatodis	Sekhar	Fleming	Zana	Proposed
DiaretDB0	68%	72%	85%	63%	86%
DiaretDB1	62%	76%	79%	71%	92%
DRIVE	66%	76%	53%	82%	68%
Average	66%	74%	77%	69%	85%

Table 3: Percentage of detected macula centers falling in the correct region.

5 Conclusion

In this paper, we have presented an approach to improve detection algorithms by fine-tuning their parameters. For this task, we have used a simulated annealingbased search algorithm. As our experiments have proved, this approach is capable of improving a detector that outperforms state-of-the-art algorithms in the field of macula detectors. As a future work, the selection of different preprocessing methods for the dataset can further improve the detection of the macula. In addition, both simulated annealing [14] and the proposed detector could be implemented in parallel to reduce their computational needs.

Dataset	Petsatodis	Sekhar	Fleming	Zana	Proposed
DiaretDB0	26.59	26.85	37.82	24.11	24.02
DiaretDB1	26.32	27.45	35.67	24.77	25.72
DRIVE	18.15	26.20	37.29	20.85	30.25
Average	23.69	26.83	36.92	23.24	26.75

Table 4: Average euclidean distance of the detected macula centers from the manually selected ones.

Acknowledgement

This work was supported in part by the János Bolyai grant of the Hungarian Academy of Sciences, and by the TECH08-2 project DRSCREEN - Developing a computer based image processing system for diabetic retinopathy screening of the National Office for Research and Technology of Hungary (contract no.: OM-00194/2008, OM-00195/2008, OM-00196/2008). We also acknowledge the Moore-fields Eye Hospital, London for their clinical support. We are thankful to Brigitta Nagy and Ignác Csősz for their technical assistance.

References

- Abramoff, M. D., Reinhardt, J. M., Russell, S. R., Folk, J. C., Mahajan, V. B., Niemeijer, M., and Quellec, G. Automated early detection of diabetic retinopathy. *Ophthalmology*, 117(6):1147–1154, 2010.
- [2] Fleming, A. D., Philip, S., Goatman, K. A., Olson, J. A., and Sharp, P.F. Automated assessment of diabetic retinal image quality based on clarity and field definition. *Investigative Ophthalmology and Visual Science*, 47:1120–1125, 2006.
- [3] Kauppi, T., Kalesnykiene, V., Kamarainen, J.K., Lensu, L., Sorri, I., Uusitalo, H., Kalviainen, H., and Pietila, J. Diaretdb0: Evaluation database and methodology for diabetic retinopathy algorithms. Technical report, Lappeenranta University of Technology, Lappeenranta, Finland, 2006.
- [4] Kauppi, T., Kalesnykiene, V., Kmrinen, J.K., Lensu, L., Sorri, I., Raninen, A., Voutilainen, R., Uusitalo, H., Klviinen, H., and Pietil, J. Diaretdb1 diabetic retinopathy database and evaluation protocol. Proc. of the 11th Conf. on Medical Image Understanding and Analysis (MIUA2007), pages 61–65, 2007.
- [5] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. Optimization by simulated annealing. *Science*, 220:671–680, May 13, 1983.
- [6] Kovacs, L., Qureshi, R.J., Nagy, B., Harangi, B., and Hajdu, A. Graph based detection of optic disc and fovea in retinal images. In Soft Computing Applications (SOFA), 2010 4th International Workshop on, pages 143–148, 2010.

- [7] Niemeijer, M., van Ginneken, B., Cree, M.J., Mizutani, A., Quellec, G., Sanchez, C.I., Zhang, B., Hornero, R., Lamard, M., Muramatsu, C., Wu, X., Cazuguel, G., You, J., Mayo, A., Li, Q., Hatanaka, Y., Cochener, B., Roux, C., Karray, F., Garcia, M., Fujita, H., and Abramoff, M.D. Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs. *IEEE Transactions on Medical Imaging*, 29(1):185–195, 2010.
- [8] Petsatodis, T. S., Diamantis, A., and Syrcos, G. P. A complete algorithm for automatic human recognition based on retina vascular network characteristics. *Era1 International Scientific Conference, Peloponnese, Greece*, 16-17 September 2006.
- [9] Ravishankar, S., Jain, A., and Mittal, A. Automated feature extraction for early detection of diabetic retinopathy in fundus images. In *CVPR*, pages 210–217. IEEE, 2009.
- [10] Ruta, D. and Gabrys, B. Classifier selection for majority voting. Information Fusion, 6(1):63 – 81, 2005. Diversity in Multiple Classifier Systems.
- [11] Sekhar, S., Al-Nuaimy, W., and Nandi, A. K. Automated localization of optic disc and fovea in retinal fundus images. 16th European Signal Processing Conference, Lausanne, Switzerland, 2008.
- [12] Sinha, P. K. and Hong, Q. H. An improved median filter. *IEEE Transactions on Medical Imaging*, 9(3):345–346, 1990.
- [13] Sivanandam, S. N. and Deepa, S. N. Introduction to Genetic Algorithms. Springer, 2008.
- [14] So, O. and Özdamar, L. Parallel Simulated Annealing Algorithms in Global Optimization. Journal of Global Optimization, 19:27–50, 2001.
- [15] Staal, J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., and van Ginneken, B. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23:501–509, 2004.
- [16] Winder, R.J., Morrow, P.J., McRitchie, I.N., Bailie, J.R., and Hart, P.M. Algorithms for digital image processing in diabetic retinopathy. *Computerized Medical Imaging and Graphics*, 33(8):608–622, 2009.
- [17] Zana, F., Meunier, I., and Klein, J. C. A region merging algorithm using mathematical morphology: application to macula detection. In *ISMM '98: Proceedings of the fourth international symposium on Mathematical morphol*ogy and its applications to image and signal processing, pages 423–430, Norwell, MA, USA, 1998. Kluwer Academic Publishers.

EXTRACTION OF VASCULAR SYSTEM IN RETINA IMAGES USING AVERAGED ONE-DEPENDENCE ESTIMATORS AND ORIENTATION ESTIMATION IN HIDDEN MARKOV RANDOM FIELDS

György Kovács, András Hajdu

University of Debrecen Faculty of Informatics 4010, POB 12, Debrecen, Hungary

ABSTRACT

The proper segmentation of the vascular system of the retina currently attracts wide interest. As a precious outcome, a successful segmentation may lead to the improvement of automatic screening systems. Namely, the detection of the vessels helps the localization of other anatomical parts and lesions besides the vascular disorders. In this paper, we recommend a novel approach for the segmentation of the vascular system in retina images, based on Hidden Markov Random Fields (HMRF). We extend the optimization problem of HMRF models considering the tangent vector field of the image to enhance the connectivity of the vascular system consisting of elongated structures. To enhance the probability estimation during the solution of the Hidden Markov problem, the Averaged One-Dependence Estimator (AODE) is used instead of the commonly used naive Bayes estimators, since AODE uses a weaker assumption than total independence of features. The advantages of our method is discussed through a quantitative analysis on a publicly available database.

Index Terms— averaged one-dependence estimator, hidden markov fields, retina, vessel

1. INTRODUCTION

Algorithms for segmenting the vascular system of the retina have high importance in automatic systems for detecting diseases (such as diabetic retinopathy) based on digital fundus images. One of the most reliably detectable anatomical parts even for low quality images is the vascular system, because there are no similar lesions on the retina. Using an accurately segmented vascular system, we have better chances to locate other anatomical parts (e.g. optic disc and fovea) of the retina and explore the disorders of the vascular system itself. The extraction of the vascular system is a rather complex task, knowing only that the vascular system has lower intensity than its background and the widths of the vessels are varying. Motivated by its importance, several vessel segmentation approaches can be found in the literature including intensity edges [1], adaptive thresholding [2], region growing [3], mathematical morphology [4], machine learning [5], ensemble learning [6]. In this paper we propose a novel vasculature segmentation approach based on Hidden Markov Random Fields (HMRF). HMRF is a popular approach for the statistical segmentation of images. The segmentation process means the maximization of an energy functional defined on the set of all possible labelings of the image, where the global maximum of the energy functional belongs to the best segmentation according to the image model. HMRF-based vascular system extraction of the retina has already been suggested in [7]. This method considers angiographic retinal images as input, where the appearance of the vessel system is much more defined than in normal fundus images. However, the usage of angiography is not allowed for the state-of-the-art automatic screening systems and the method presented in [7] gives poor results fo normal retinal images. In the commonly used form of the HMRF based segmentation, the background and foreground probability estimations are performed by coninuous naive Bayesian estimation[8], based on the continuous model extracted from the training data. In the segmentation of color images this approach performs well, however using features extracted from intensity images, like the green channel of retinal images, the assumption that the distribution of features is Gaussian or Mixture of Gaussians and the coordinates of the feature vectors are independent, seems to be too strong, smoothing the boundaries of classes in the feature space. To overcome this problem, we have integrated the averaged one-dependence estimator (AODE) [9] into our segmentation model, which does not have the assumption of Gaussian feature distributions and total independence of feature coordinates. The use of AODE estimators has not yet been reported in HMRF segmentation models. To enhance the segmentation of thin vessels, we have trained the AODE model for three classes: background, thin vessels and thick vessels, but during the segmentation, the latter two classes are considered to be one class. Since HMRF is a general approach that does not consider orientation information at the pixels in its basic formulation, we integrate a corresponding new term into the model. However, instead of following the classic way with using the local gradients directly, first we determine the structure tensor field from them. Then, the tensor field is smoothed and weighted and the local tangents are estimated as the eigenvectors corresponding to the smaller eigenvalues of the tensors. This approach leads to a more reliable estimation of the vessel directions that helps to preserve the connectivity of the vascular system during segmentation. The rest of the paper is organized as follows. Section 2 describes the steps of our proposed method. In Section 3 we exhibit our experimental results on the publicly available database DRIVE [10] and compare the results to a state-of-the-art method. Finally in Section 4, some conclusions are drawn.

2. THE PROPOSED METHOD

The input of our algorithm is the RGB retina image, the output is a label image containing two labels: vessel, back-ground. For the training of the segmentation model, the DRIVE database is prepared by applying simple morphological operators to isolate thick and thin vessels. In Figure 1 a sample from the training database, the corresponding manually labeled mask and the extracted thick and thin vessels can be seen, respectively.



Fig. 1. Retina image from the training set (a), its standard manual labeling (b), the extracted thin (c) and thick (d) vessels

Let $I \in [0..255]^{m \times n}$ be an intensity image (in our case the green channel, since mainly the green channel contains information about the vascular system) and $L \in \{fg, bg\}^{m \times n}$ a possible labeling of the input image. The HMRF based segmentation means the maximization of the functional

$$E(I,L) = \sum_{i,j} P(L(i,j)|I(i,j)) + \sum_{i,j} C_D(L,i,j), \quad (1)$$

where P is the conditional probability of the (i, j) pixel belonging to the L(i, j) class and $C_D(L, i, j)$ measures the connectivity of the labeling in the (i, j) position. The maximization of the functional is performed by simulated annealing.

2.1. Probability term

The first term of the energy functional E estimates the probability of L being the most appropriate labeling of I. The more accurate probability estimation we use, the more accurate the

segmentation becomes. Since the distribution of the features we use to characterize the pixels and their neighborhood usually does not follow Gaussian distribution, we have chosen discrete classifiers. The drawback of the commonly used discrete naive Bayesian classifier is that it considers the correlation of one feature and the class label independently from other features, that is it assumes the total independence of features. The averaged one-dependence estimator (AODE), published in 2002 [9] considers the joint probabilities of every pair of features, as well. Therefore, the probability estimation becomes more accurate, but the computation demand is still a fraction of that of multivariate Bayesian classification.

Let Y be the set of class labels and $X \subset \mathbb{Z}^d$ be a set of d dimensional discrete feature vectors with

$$\{x_i | x \in X\} \bigcap \{x_j | x \in X\} = \emptyset, \quad 1 \le i, j \le d, i \ne j.$$
(2)

The training of the AODE classifier means the maintenance of a two dimensional $\hat{P}(y, x_i)$ and a three dimensional $\hat{P}(y, x_i, x_j)$ joint frequency table. The probability of the *d*-dimensional feature vector \hat{x} belonging to the class *y* can be estimated by the following formula:

$$\hat{P}(y|\hat{x}) = \frac{\sum_{i} \hat{P}(y, \hat{x}_{i}) \prod_{j} \hat{P}(\hat{x}_{j}|y, \hat{x}_{i})}{\sum_{y' \in Y} \sum_{i} \hat{P}(y', \hat{x}_{i}) \prod_{j} \hat{P}(\hat{x}_{j}|y', \hat{x}_{i})}.$$
(3)



Fig. 2. Test image (a), standard manual segmentation (b), orientation field (c), HMRF-segmentation (d)

2.2. Connectivity term

The connectivity term of the HMRF optimization problem tries to quantitatively characterize the connectivity of the labeling to avoid the fragmented segmentation results based on the probability term only. In the classic model of HMRF, 4connectivity is preferred:

$$C(L, i, j) = -(d(L_{i,j}, L_{i+1,j}) + d(L_{i,j}, L_{i-1,j}) + d(L_{i,j}, L_{i,j+1}) + d(L_{i,j}, L_{i,j-1})) * \beta,$$
(4)

where $d(l_1, l_2) = 1$ if the l_1 and l_2 labels are equivalent. Otherwise, $d(l_1, l_2) = 0$. β is a penalty parameter of the objective functional E, used to control the connectivity property. Since thin vessels have fine structure of 1 pixel width, this commonly used connectivity term is a rough constraint on the final segmentation. To enhance this component, we have developed the direction dependent connectivity term C_D , where the connectivity is examined and penalized only in the direction of the strongest tangents in the neighborhood of each pixel (i, j):

$$C_D(L, i, j) = -(d(L_{i,j}, L_{i+t_x(I,i,j),j+t_y(I,i,j)}) + d(L_{i,j}, L_{i-t_x(I,i,j),j-t_y(I,i,j)}) * \beta,$$
(5)

where $t_x(I, i, j)$ and $t_y(I, i, j)$ are the estimation of the x and y components of the strongest tangent in the local neighborhood of (i, j), respectively:

$$t_x(I, i, j) = \begin{cases} 1, \text{ if } A(I)_{i,j} < \frac{3*\pi}{8} \text{ or } A(I)_{i,j} > \frac{5*\pi}{8}, \\ 0, \text{ otherwise,} \end{cases}$$
(6)

$$t_y(I, i, j) = \begin{cases} 1, \text{ if } A(I)_{i,j} > \frac{\pi}{8} \text{ and } A(I)_{i,j} < \frac{7*\pi}{8}, \\ 0, \text{ otherwise,} \end{cases}$$
(7)

where A(I) is the estimation of the main tangent angels in each pixel (i, j) of I. For the estimation of A, we have computed the structure tensor field STF of the image I, assigning a 2×2 sized matrix to each pixel:

$$STF_{i,j} = \begin{bmatrix} G_{x_{i,j}}^2 & G_{x_{i,j}} * G_{y_{i,j}} \\ G_{x_{i,j}} * G_{y_{i,j}} & G_{y_{i,j}}^2 \end{bmatrix}, \quad (8)$$

where $G_{x_{i,j}}$ and $G_{y_{i,j}}$ are the respective x and y components of the gradient at pixel (i, j), estimated by the commonly used Sobel-operators. The tensor field is then weighted by the edge magnitudes and convolved by a Gaussian mask to provide smooth estimates:

$$SFT_{smoothed_{i,j}} = G(\sigma, k) * (T_{i,j}(G_{x_{i,j}}^2 + G_{y_{i,j}}^2)^{\frac{\mu}{2}}).$$
(9)

This step is similar to the step of smoothing the orientation information in the scale-invariant feature transform (SIFT). Using this smoothed tensor field the estimation of object tangent directions comes as

$$A(I)_{i,j} = \tan^{-1}\left(\frac{E_{x_{i,j}}}{E_{y_{i,j}}}\right),$$
(10)

where E is the vector field containing the eigenvector belonging to the smallest eigenvalue of SFT, elementwise. In Figure 2, one test image and the orientation field can be seen, the latter one represented by spikes.



Fig. 3. Convolution based MLSF (a), correlation based MLSF, matched Gabor filter magnitude (c), adaptive local contrast enhancement (d)

3. RESULTS

The proposed method is trained on the training set and tested on the testing set of the publicly available DRIVE database. In the training phase, 200 features (eccentricity, invariant image moments, central moments, matched line segment filters using convolution and correlation, matched Gabor filters, and the result of several contrast enhancement methods with various parametrizations) have been extracted for each pixel in the database and after forward feature subset selection, the final AODE classifier is trained on the features providing the best accuracy (in our case matched line segment filters using correlation and convolution, Gabor filters and the result of the adaptive local contrast enhancement with three parameterization). In Figure 3, four feature images of the six selected features can be seen. To enhance the recognition of thin structures, we have used three classes in the training process, but in the segmentation process the thin and thick vessel labels are considered to be equivalent. Using only the AODE classifier the accuracy became 93.7%, but the result is noisy. Keeping only the largest connected component, the accuracy decreases to 85.3%. See Figure 2. Using the HMRF-based segmentation with orientation estimation and directional connectivity term, the result becomes less noisy and after extracting the largest connected component, the accuracy increases to 91.2%. The specificity of the HMRF-based segmentation becomes 89.8% alghough several thin vessels are not labeled on the manually segmented images but labeled by our method. In Figure 4 one part of the same test image is highlighted, in which the thin and low contrast vessel found by our method is not labeled in the manual annotation available as part of the DRIVE database. The state-of-the-art method published by Oost et al. [6] reached 88.5% accuracy and 93.6% specificity, on the same database. See Figure 2 for the standard manual labeling and the HMRF segmentation of a test image.



Fig. 4. A really thin and low contrast vessel on the test image (a), its standard manual labeling (b), and the segmentation of the proposed method.

4. SUMMARY

The proposed method delegates several new approaches into the field of segmentation of vascular system in retinal images. In the method we utilize the powerful AODE classifier which does not use the assumption of Gaussian distribution of features like many other classifiers and also weakens the strong assumption of naive Bayesian classifiers, i.e. the features are independent from each other. To increase the connectivity of the final segmentation, the classic HMRF model is extended with a direction dependent connectivity term. To increase accuracy, in the training database the thin and thick vessels are distinguished, but in the segmentation process, when the connectivity is examined, both of them is considered belonging to the same class. The overall method is straightforward, increasing the variety of extracted features, the accuracy and specificity of the segmentation may be increased further.

5. ACKNOWLEDGEMENTS

This work was supported in part by the Janos Bolyai grant of the Hungarian Academy of Sciences and by the TECH08-2 project DRSCREEN - Developing a computer based image processing system for diabetic retinopathy screening of the National Office for Research and Technology of Hungary (contract no.: OM-00194/2008, OM-00195/2008, OM-00196/2008).

6. REFERENCES

- A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms," *IEEE Trans. Inf. Technol. Biomed.*, vol. 3, pp. 125–138, February 1999.
- [2] X. Jiang and D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 25, pp. 131–137, January 2003.
- [3] M. E. Martinez-Prez, A. D. Hughes, A. V. Stanton, S. A. Thom, A. A. Bharath, and K. H. Parker, "Retinal blood vessel segmentation by means of scale-space analysis and region growing," in *Medical Image Computing and Cmoputer-Assisted Intervention MICCAI99*, 2003, vol. I, pp. 131–137.
- [4] F. Zana and J. C. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," *IEEE Trans. Image Proc.*, vol. 10, pp. 1010–1019, July 2001.
- [5] M. Niemeijer, J. J. Staal, B. van Ginneken, M. Loog, and M. D. Abramoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database," *SPIE Medical Imaging*, vol. 5370, pp. 648–656, 2004.
- [6] E. Oost, Akatsuka Y., A. Shimizu, H. Kobatake, D. Furukawa, and A. Katayama, "Vessel segmentation in eye fundus images using ensemble learning and curve fitting," in *Biomedical Imaging: From Nano to Macro*, 2010, IEEE International Symposium on. IEEE, 2010, pp. 676–679.
- [7] A. Simo and E. de Ves, "Segmentation of macular fluorescein angiographies: A statistical approach," *Pattern Recognition*, vol. 34, pp. 128–139, April 2001.
- [8] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts," ACM Transactions on Graphics, vol. 23, pp. 309–314, 2004.
- [9] Geoffrey I. Webb, Janice Boughton, and Zhihai Wang, "Averaged one-dependence estimators: Preliminary results," in *University of Technology Sydney*, 2002, pp. 65–73.
- [10] J. J Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imaging*, vol. 23, pp. 501–509, April 2004.

MICROANEURYSM DETECTION IN RETINAL IMAGES USING A ROTATING CROSS-SECTION BASED MODEL

Istvan Lazar, Andras Hajdu

lazar.istvan@inf.unideb.hu, hajdu.andras@inf.unideb.hu Faculty of Informatics, University of Debrecen, POB 12, 4010 Debrecen, Hungary

ABSTRACT

Retinal image analysis is currently a very vivid field in biomedical image analysis. One of the most challenging tasks is the reliable automatic detection of microaneurysms (MAs). Computer systems that aid the automatic detection of diabetic retinopathy (DR) greatly rely on MA detection. In this paper, we present a method to construct an MA score map, from which the final MAs can be extracted by simple thresholding for a binary output, or by considering all the regional maxima to obtain probability scores. In contrary to most of the currently available MA detectors, the proposed one does not use any supervised training and classification. However, it is still competitive in the field, with a prominent performance in the detection of MAs close to the vasculature, regarding the state-of-the-art methods. The algorithm has been evaluated in a publicly available online challenge.

Index Terms— Biomedical image processing, pattern recognition, microaneurysm detection, intensity profile



1. INTRODUCTION

Fig. 1. Retinal image with microaneurysm marked.

Automated analysis of retinal (fundus) images has become a quite active field among ophthalmologists and researchers in digital image processing. One of the most important motivations is the need for an effective computer-aided diagnostic (CAD) system to recognize diabetic retinopathy (DR). DR is the damage of the retina caused by diabetes. It is a sight-threatening disease that develops in most of the patients with long-standing illness. An automated DR screening system would be a great assist in the processes of diagnosing and progression tracking. Currently, these tasks are done by human graders who categorize each image manually according to the diagnosis protocol.

The presence of microaneurysms (MAs) in the retina is the earliest symptom of retinopathy, thus their reliable detection is essential in an automatic DR screening system. MAs are circular dark spots, whose diameters are less than the diameter of the major optic veins. Figure 1 shows a retinal image with an MA marked. In this paper, we present a method for retinal MA detection, which in contrast to most of the existing algorithms does not require any training. Still, it is able to precisely locate MAs with high sensitivity at low false positive rate. Moreover, it has an outstanding performance on lesions close to the vasculature regarding the state-of-the-art methods.

In section 2 we will briefly review the available algorithms for retinal MA detection, after which in section 3 the proposed method is described in details. In section 4, the results of our method in an open online challenge of MA detectors are presented.

2. STATE-OF-THE-ART MA DETECTORS

Most of the currently available methods divide MA detection into two consequent stages: candidate extraction and classification. Usually, the first step of candidate extraction is image preprocessing to reduce noise and improve contrast. After preprocessing, specific image segmentation is used to extract as much regions as possible that probably correspond to MAs. In the second step, the resulting candidates are labeled as true or false ones using a supervised learning based method. This classification requires a training set to establish the boundaries of the classes. The training set consists of pairs of feature vectors and class labels. Feature vectors are ordered sets of certain property values, mostly geometrical or color descriptors that may help to distinguish MAs from other objects. Such state of the art methods are proposed by e.g. Niemeijer et al. [1], Quellec et al. [2], and Zhang et al. [3].

Methods different from this two-phase approach have been also proposed. E.g. in [4] Giancdardo et al. presented a method based on Radon transformation that does not require explicit training.

3. THE PROPOSED METHOD

The method described in this paper is based on the two most characteristic attributes of MAs: diameter and circularity. We will show that it is possible to construct a score-map that assigns a score value to every pixel in the image, where the dark small circular objects will achieve the highest scores, while the vessel structure almost completely disappears.

3.1. Preprocessing

The only actual parameter of the proposed algorithm is the maximal MA diameter (d_{max}). We consider this value as 7 pixels in the case of a horizontal resolution of 760 pixels for the visible retinal part. Thus, input images in different spatial resolution need to be rescaled. When working with a retinal image, the usual approach is to use its green channel, since it has the most useful information content. Though the proposed method does not require a specific preprocessing, a certain amount of noise reduction is necessary in most fundus images. We suggest the use of a simple two dimensional Gaussian blurring with a standard deviation of 1.4. Also, the input images are inverted, thus MAs and vessels will appear as bright structures. From now on, when referring to a retinal image, it is meant as the smoothened inverted green channel I_{sg} of the original retinal image.

3.2. Cross-sectional profiles

Probably the most challenging task during automatic MA detection is the distinction between MAs and thin vessel fragments. Vessel segments are such structures in the retinal image whose points are local maxima in at least one direction. On the other hand, MAs are circular objects, thus they appear as local maxima from all directions.

To utilize this approach in practice, we need to test whether a pixel in the image is a local maximum in a particular direction. Therefore, we consider discrete line segments (scan lines) with different slopes, passing through the specific position being tested. The recordings of the pixel values along these scan lines result in a set of one dimensional intensity profiles, or in other words, in a set of cross-sectional profiles.

If we examine these profiles, we find that MAs show a definite peak for all directions. In the case of vessel points, only the profiles of scan lines perpendicular to the vessel show strong peaks, and as the direction of the scan line approximates the vessel segment, the peak on the profile becomes more and more elongated, until it almost completely disappears. Figure 2 shows the cross-section profiles of a vessel segment and an MA point at different directions, respectively.



respectively.

3.3 Scanning of the entire image

To exploit the observation about the cross-section profiles of MAs and vessel segments, every pixel of the image has to be accessed from every possible direction. For this reason we will use scan lines that pass through the entire image. Let us consider the *slope-intercept* equation $y = m \cdot x + b$ of the line. By modifying the *y-intercept* parameter *b* at a fixed slope *m*, every pixel of the image can be accessed for that given *m* direction, except the vertical case which has to be handled separately. This way, we define the scan line $SL_{\alpha,b}$, a point vector, whose i^{th} element $(SL_{\alpha,b}[i])$ is the i^{th} point of the discrete line defined by the $y = \tan \alpha \cdot x + b$ equation. That is,

$$SL_{\alpha,b}[i] = \begin{cases} -90^{\circ} < \alpha < -45^{\circ} : \left(\left|\frac{-i}{\tan\alpha}\right|, \lfloor b \rfloor - i\right), \\ -45^{\circ} \le \alpha \le 45^{\circ} : (i, \lfloor \tan\alpha \cdot i + b \rfloor), \\ 45^{\circ} < \alpha < 90^{\circ} : \left(\left|\frac{i}{\tan\alpha}\right|, \lfloor b \rfloor + i\right). \end{cases}$$

The intensity profile of $SL_{\alpha,b}$ is defined as an integer valued vector, whose i^{th} element is the intensity value of the image at the $SL_{\alpha,b}[i]$ position. That is, $IP_{\alpha,b}[i] = I_{gs}(SL_{\alpha,b}[i])$. To handle the intensity profile of vertical scan lines, we define the VIP_x vector, whose i^{th} element is the intensity value of the image in the i^{th} row and x^{th} column. That is, $VIP_x[i] = I_{sg}(x, i)$. As it can be seen from the definition of the scan lines, the possible directions cover 180°. Also, we specify that the scan directions have to be equidistantly sampled, i.e.:

$$\alpha \in \Phi = \{\varphi_1, \varphi_2 \dots\}: -90^{\circ} < \varphi_i \le 90^{\circ} \land \varphi_{i+1} - \varphi_i = \omega,$$

where ω denotes the scan step parameter. Theoretically, choosing a small ω would result in a finer scanning. However, our experiments showed that, choosing ω to be less than 3° will not make relevant difference in the final results. Figure 3 shows 4 examples of scan lines for different directions. The arrows represent the direction of the scanning.



Fig. 3. Examples of scan lines with different directions.

3.4. Peak localization and height measurement

Since MAs appear as peaks on the intensity profiles of scan lines, the next step is to detect these peaks. In this case we have a width restriction to the peaks, that is, only those peaks whose width is less than or equal to d_{max} can correspond to MAs. Besides the position of the peaks, the measure of their height is also required, since this information describes how much the object is separated from its surroundings in a particular direction. Such measure is needed that will consider both the peak height and its width, that is, peaks that are wide will receive a lower value, even if their height would suggest otherwise. This requirement is fulfilled by applying area opening, a special type of attribute opening [5] on the intensity profiles. By considering the difference of the result and the original profile (top-hat transformation), we may obtain a proper corrected peak height measure for every position in the profile.

Attribute opening can be applied for both one and two dimensional data as well. It is based on the successive extension of all the regional maxima by progressively lowering the local threshold level, until a certain criterion becomes satisfied.

In the case of area opening, this halt condition is that the area of the extended region becomes larger than the given parameter. Let $\Gamma_{\lambda}^{a}(v)$ denote the result of area opening applied to vector v with parameter λ . This way, the area opening top-hat operator is defined as $\overline{\Gamma_{\lambda}^{a}}(v) = v - \Gamma_{\lambda}^{a}(v)$. Figure 4 shows the result of the attribute opening top-hat operator applied to an intensity profile with $\lambda = d_{max}$. It can be seen, that the top-hat values give a proper measure of the peak height values along the profile.



Fig. 4. Top-hat result of attribute opening on an intensity profile.

3.5. Assigning a directional height value to every pixel

To assign a height value for a given direction to every pixel, we define the directional height map function $DHM(x, y, \alpha)$, whose value is calculated as the corresponding value of the area opening top-hat operator applied to the intensity profile of the scan line that passes through the (x, y) point at an α angle. That is,

$$DHM(x, y, \alpha) = \begin{cases} -90^{\circ} < \alpha < -45^{\circ} : \overline{\Gamma_{d_{\max}}^{a}}(IP_{\alpha,b^{*}})[[b^{*}] - y], \\ -45^{\circ} \le \alpha \le 45^{\circ} : \overline{\Gamma_{d_{\max}}^{a}}(IP_{\alpha,b^{*}})[x], \\ 45^{\circ} < \alpha < 90^{\circ} : \overline{\Gamma_{d_{\max}}^{a}}(IP_{\alpha,b^{*}})[y - [b^{*}]], \\ \alpha = 90^{\circ} : \overline{\Gamma_{d_{\max}}^{a}}(VIP_{x})[y], \end{cases}$$

where $b^* = y - \tan \alpha \cdot x$. Let DHV(x, y) denote the vector containing the height values for a given position, i.e.: $DHV(x, y)[i] = DHM(x, y, \varphi_i)$. Considering the DHV of an MA and a vessel point, we find that only MAs achieve nearly equal height values for all the scan directions, while in the case of vessel segments, there is a gap at the directions when the scan line fits the vessel segment. The DHV of an MA and a vessel point can be seen in Figure 5.



Fig. 5. Plot of height values for an MA and a vessel point as the scan direction is rotating.

3.6. MA score map construction

Eventually, our goal is to create a score map, from which the MAs can be extracted by a simple thresholding for a binary output, or by considering all the regional maxima to obtain candidate points with the corresponding probability scores. We have several requirements regarding this map. It has to have high values for small circular objects (MAs), while the response to elongated structures (vessels) has to be minimal. Moreover, the response to MAs has to be proportionate to their strength.

We have seen that MAs have nearly equal height values for all scan directions. To obtain a measure of the proportion of directions in which the height value was larger than a given *h* level to the number of all scan directions, we define the peak ratio map function PRM_h , i.e.:

$$PRM_{h}(x, y) = \frac{|\alpha \in \Phi : DHM(x, y, \alpha) \ge h|}{|\Phi|}$$

 PRM_h has a value near to 1.0 in the case of MA points. However, usually it has high values for vessels, especially for vessel crossings, as well. Figure 6 shows PRM_6 , the peak ratio map for height level 6 for the fundus image in Figure 1.



Fig. 6. The peak ratio map at height level 6 (PRM_6) of the fundus image in Fig. 1.

By applying a modified version of diameter opening top-hat operation on PRM_h , vessels can be filtered out, and an MA probability image can be constructed for height level h. Diameter opening is a type of attribute opening, where the diameter of the extended region is tested. We apply two modifications. First, only those maximum regions have to be extended, where the PRM_h of the seed point is near to 1.0, since only these regions correspond to MAs. Theoretically, this value should be always 1.0 in the case of circular objects. However, we found that, in practice, it is recommended to lower the seed threshold to e.g. 0.9. This limitation for the seed regions result in a faster execution, since most of the regional maxima can be ignored. Second, in the resulting top-hat, the extended regions are multiplied with the *PRM_h* value of the corresponding seed. Let $\Gamma_{\lambda}^{d}(X)$ denote the result of this modified diameter opening tophat operation on image X, with parameter λ . This way, we obtain the MA probability map at height level h, denoted as P_h , by applying $\overline{\Gamma_{\lambda}^d}$ on PRM_h , i.e.:

$$P_h = \overline{\Gamma^d_{d_{max}}}(PRM_h).$$

To construct the final MA score map, which is independent from the height level parameter, we consider all possible P_h probability maps and combine them in a weighted pixelwise maximum selection. The value of each pixel in the final score map will be the product of the maximal probability value among the probability maps, and the corresponding height level, i.e.:

$$Score(x, y) = P_{h_{max}(x, y)}(x, y) \cdot h_{max}(x, y)$$
$$h_{max}(x, y) = argmax_i P_i(x, y),$$

where *Score* denotes the final MA score map. We found that it is sufficient to consider only the first 30 ($h = \{1 \dots 30\}$) probability maps. Moreover, the calculation of the P_h probability maps can be performed in parallel for the different h values, which gives a further boost to the method. The final MA score map of the fundus image in Figure 1 is shown in Figure 7. We can see that the highest score value was achieved by the strongest MA in the image.



Fig. 7. The final MA score map of the fundus image in Fig. 1.

4. EXPERIMENTAL RESULTS

To examine the performance of the proposed method in comparison with other state-of-the-art algorithms, we have submitted our results to the Retinopathy Online Challenge (ROC) [6]. To the best of our knowledge, currently ROC is the only possibility to compare MA detectors under the same conditions. This dataset consists of a training and a test set. Both sets contain 50 fundus images. The reference MAs are available for the training set only, which enables the participating teams to train their algorithms, and then submit the results obtained on the test set. The organizers evaluate the received data, and construct the FROC (free receiver operating characteristic) curve. The final score is calculated as the average sensitivity at seven fixed false positive (FP) rates (from 0.125 to 8 FPs per image). Figure 8 shows the FROC curve of our algorithm. At the time of the submission of this paper, our method (team Lazar et al.) is ranked at the 5th place among the 10 best participants. Besides the method proposed by team ISMV [4], our method is the only one that has achieved such a competitive score without using any supervised training or classification. Moreover, our method has achieved the highest score with respect to detecting MAs close to the vasculature (0.339). The currently first placed method (team DRSCREEN) is an ensemble-based detector that combines multiple individual algorithms, including this one, as well. The results of the individual teams can be seen in Table 1.



	Team name	Score
1.	DRSCREEN	0.434
2.	Niemeijer et al. [1]	0.395
3.	LaTIM [2]	0.381
4.	OKmedical [3]	0.357
5.	Lazar et al.	0.355
6.	GIB Valladolid	0.322
7.	Fujita Lab	0.310
8.	IRIA-Group	0.264
9.	ISMV [4]	0.256
10.	Waikato Retinal Imaging Group	0.206

Table 1. The participating teams in ROC with their achieved scores.

The proposed method was implemented entirely in Java SE 1.6. The average run time of an image from the ROC dataset was 15 seconds, using a PC with an Intel® Core[™]2 Quad Q8200 Processor and 2 GB RAM.

5. CONCLUSION

In this paper, we have presented a method that is capable of constructing a score map to a fundus image, in which the points corresponding to MAs will achieve high scores, while the response to the vasculature is minimal. This is accomplished without the usage of supervised learning based classification, and the only actual parameter of the method is the maximal MA diameter. The final MAs can be extracted by simple thresholding for a binary output, or by considering all the regional maxima to obtain probability scores. The proposed algorithm has been evaluated in a public online challenge aiming the comparison of MA detectors, where it has achieved a competitive score among the ten best participants; moreover, it has achieved the highest score with respect to the detection of MAs close to the vasculature. The method is also a component of an ensemble based algorithm, which at the time of the submission of this paper, holds the first place in the mentioned online competition.

ACKNOWLEDGMENTS

This work was supported in part by the Janos Bolyai grant of the Hungarian Academy of Sciences, and by the TECH08-2 project DRSCREEN- Developing a computer based image processing system for diabetic retinopathy screening of the National Office for Research and Technology of Hungary (contract no.: OM-00194/2008, OM-00195/2008, OM-00196/2008).

REFERENCES

[1] M. Niemeijer, J. Staal, M. D. Abramoff, M.S. A.Suttorp-Schulten, and B. van Ginneken, "Automatic detection of red lesions in digital color fundus photographs," *IEEE Transactions on Medical Imaging*, vol. 24, pp. 584–592, 2005.

[2] G. Quellec, M. Lamard, P. M. Josselin, G. Cazuguel, B. Cochener, and C. Roux, "Optimal wavelet transform for the detection of microaneurysms in retina photographs," *IEEE Trans. on Medical Imaging*, vol. 27(9), pp. 1230-1241, 2008.

[3] B. Zhang, X. Wu, J. You, Q. Li, and F. Karray, "Hierarchical detection of red lesions in retinal images by multiscale correlation filtering," *SPIE Medical Imaging 2009: Computer-Aided Diagnosis*, Vol. 7260, pp. 72601L, 2009.

[4] L. Giancardo, F. Meriaudeau, T. P. Karnowski, K. W. Tobin, Y. Li, and E. Chaum, "Microaneurysms Detection with the Radon Cliff Operator in Retinal Fundus Images," *SPIE Medical Imaging*, Vol. 7623, pp. 29, 2010.

[5] L. Vincent, "Morphological Area Openings and Closings for Greyscale Images," *Proc. NATO Shape in Picture Workshop*, Springer-Verlag, pp. 197-208, 1992.

[6] M. Niemeijer, B. van Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sanchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu, X. Wu, G. Cazuguel, J. You, A. Mayo, Q. Li, Y. Hatanaka, B. Cochener, C. Roux, F. Karray, M. Garcia, H. Fujita, M. D. Abramoff, "Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs," *IEEE Transaction on Medical Imaging*, 29(1), pp.185-95, 2010.