

Képi alapú többszintű környezetelemzés
(Image based multi-level environment analysis)

Válasz Prof. Szirmay-Kalos Lászlónak

az MTA doktori tézis opponensének

Benedek Csaba

2020. június 9.

Mindenekelőtt köszönöm Prof. Szirmay-Kalos László szakértő, gondos és alapos bírálatát. A bírálat megjegyzéseire és kérdéseire az alábbiakban adok részletes választ.

- 1. A Markov-i megközelítés a gépi látás egyik robusztus és hatékony technikája, amelyet számos előnye mellett a relatíve nagyobb számításigény jellemzi. Ezért kérem a Jelöltet, hogy helyezze el az eredményeit a feldolgozási idő, hardver igény skáláján is, és mutassa be, hogy mely algoritmus alkalmas valós-idejű, interaktív, illetve off-line feldolgozásra.*

A disszertációban bemutatott munka során kizárólag (korszerű) asztali számítógépeken illetve laptopokon dolgoztam. Többnyire C++ implementációjú prototípus jellegű, illetve demonstrációs programok készültek, egyes modulok Matlab-ban, illetve a mélytanulósos feladatoknál Python-ban készültek el.

Az egyes módszerek futási idejét érintő tesztelések eltérő mélységben és részletezettséggel készültek, így a következő áttekintés a pontos összehasonlításukra nem ad alapot, de úgy gondolom, hogy megfelelő közelítéssel választ ad a Bíráló kérdésére. Doktorandusz hallgatóimnak köszönhetően a 4.3 altézisben szereplő több módszernek is elkészült párhuzamos implementációja többmagos CPU vagy GPU platformokra, míg a többi tézispontnál közölt futásiidő-eredmények kizárólag soros implementációkhoz tartoznak

- 1.1 tézis L^3 MRF modell kameramozgástól független objektum-elmozdulások régióinak kinyerésére:**
A módszert 320×240 méretű képrészletekre teszteltük. Körülbelül 10 másodpercet számolt egy képpáron, a futási idő felét a képregisztráció és a jellemzőkinyerés műveleteire fordítottuk, a másik felét a háromrétegű Markov véletlen mező energiájának optimalizációja igényelte. A módszer ebben a formában és implementációban nem alkalmas valós-idejű objektumdetekcióra és –követésre, azonban használható forgalomsűrűség becslésére, vagy a sztereo-rekonstrukciót megelőző lépésben a dinamikus, „outlier”-ként megjelenő régiók kiszűrésére, akár interaktív akár offline alkalmazásokban.
- 1.2 tézis CXM modell lényeges változások kinyerésére nagy időkülönbséggel készült légi fényképek összehasonlító vizsgálata során:** A [4] cikkben közölt méréseink alapján az eljárás 25-30 másodpercig számolt 952×640 pixel méretű képpárokon, Intel(R) Core(TM) i7 3.20 GHz CPU processzorral ellátott asztali számítógépen. Ebből a jellemzőkinyerés körülbelül 5 másodpercet vett igénybe, a fennmaradó időt a négyrétegű MRF energiafüggvény optimalizációjára fordítottuk. A módszer jelenlegi implementációjában offline feldolgozásra alkalmazható, türelmes operátor esetén interaktív felhasználás is szóba jöhet. Valós idejű működés igénye ennél az alkalmazásnál ritkán merülhet fel, hiszen több éves különbséggel készült, előzetesen regisztrált képekkel dolgozunk.

- 2.1 tézis *tJPM többidejű jelölt pontfolyamat modell épületek és épületváltozások detekciójára*: a [10] cikkben közölt méréseink alapján 0.5-1M pixelből álló képek feldolgozása 30-70 másodpercig tartott, ebben a tekintetben versenyképes volt a cikkben összehasonlított referenciamódszerekkel is. Az alkalmazás jellegéből adódóan itt is az offline felhasználásra nyílik tipikusan igény, de a módszerünknek egy interaktív vizsgálat során történő lefuttatása is elfogadható késleltetéssel jár.
- 2.2 tézis *többkeretes pontfolyamat modell ISAR (radar) képsorozatok analízisére*: a [7] cikkben közölt méréseink alapján 2fps és 5fps közötti sebességgel tudtunk 256×96 méretű képkockákból álló képsorozatokot feldolgozni (a sebesség itt függött az objektum méretétől, a kezdeti detekció minőségétől és a detektálható jellemzőpontok számától). A feldolgozás mindenképpen összegyűjtött keretsomagokon (*frame batch*-eken) történik, tehát nem végezhető a módszerrel valósidejű mozgáskövetés, azonban kis késleltetéssel lehetővé teszi például egy partközeli úszó ellenséges teherhajó észlelését és beazonosítását.
- 3 téziscsoport *BJPM beágyazott jelölt pontfolyamat modell*: a [2] cikkben közölt kísérleteink alapján tesztképpenként 10-30 másodperc közötti futási időt mértünk, ami 20-30%-os többletet jelentett hasonló paraméterű, de rétegenként külön futtatott modellek időigényéhez képest. A bemutatott statikus jellegű távérzékelési és optikai minőségellenőrzési feladatok offline feldolgozást, esetleg interaktív vizsgálatokat igényelnek, amit a módszer lehetővé tesz.
- 4.1 tézis: *3D jelölt pontfolyamat modell személyek többkamerás lokalizációjára*: ez a megoldásunk képkockánként több másodpercig számolt, főleg a jellemzőkinyerés bonyolultsága miatt, így valósidejű alkalmazásra nem volt használható. A modellt egy ismert benchmark adatbázison (PETS2009) definiált lokalizációs problémára hegyeztük ki, ahol a verseny fő szempontja a 3D detekció pontossága volt, ezért kevésbé fókuszáltunk a futási idő optimalizációjára.
- 4.2 tézis *videofelügyeleti eljárások forgó többsugaras (FT) Lidar szenzor méréseire alapozva*: a módszer mozgásdetekciós és alakzatkövető modulja valósidejűben működik, ezt több kutatók éjszakája rendezvényen be is mutattuk a közönségnek, valamint a demonstrációnk a *Velodyne Europe GmbH* meghívására a szenzorgyártó cég standján két hétig futott a 2017-es Frankfurti Autószalon (IAA) kiállításon. A személyazonosítás és a cselekvésfelismerés szintén működik valósidejűben, azzal a megkötéssel, hogy mivel itt az algoritmus futtatásához meghatározott számú képkocka összegyűjtésére van szükség a mozgásról, a detekciós eredmény mindenképpen kicsit késleltetve jelenik meg. (A disszertáció 6.21 ábrája ezt a késleltetést mutatja idődiagramon.)
- A bírálói kérdéshez kapcsolódóan azt külön kiemelném, hogy az első mozgásdetekciós lépés a pontfelhő mélységkép reprezentációjában Markov mezős optimalizációt is tartalmaz, ez azonban egy determinisztikus gráfágas alapú optimalizációs algoritmussal nagyon hatékonyan elvégezhető, így a [6] cikkben közölt méréseim alapján a Markov mező alapú háttérkivonás 15-16 fps sebességgel fut.
- 4.3 tézis *mozgó járművekre rögzített FT Lidar méréseinek elemzése*: az úttest detekciója és az objektumok elkülönítése valósidejűben működik (többszálú, CPU-n futó implementációval), ezek az eljárások később ipari rendszerekbe is beépítésre kerültek egy projekt kapcsán. A mélytanuláson alapuló objektumfelismeréshez nem készült teljes C++ implementáció, de megfelelő kódoptimalizálással ez is valósidejűben működhet. A multimodális pontfelhőregisztráció a dolgozatban is szereplő E.2 táblázat alapján tipikusan 1-3 másodpercet vett igénybe Velodyne képkockánként. Itt mindenképpen a valósidejű működés a kitűzött cél, aminek az elérésén jelenleg is dolgozunk (például [16] valósidejű közelítő megoldást ad a regisztrációra). A mobil lézerszenzéssel nyert 3D pontfelhő szemantikus osztályozása tipikusan offline feladat. Itt egy-egy kiválasztott pontfelhő-kocka neurális hálózat

alapú osztályozása viszonylag gyorsan elvégezhető (10^{-4} s egy $23 \times 23 \times 23$ -as voxelkörnyezetre), azonban a nagy kültéri pontfelhőhelyszínek szekvenciális végigolvasása időigényes feladat. A neurális hálók tanítása GPU-s implementációval történt, a disszertáció benyújtását követően publikált SZTAKI-CityMLS adatbázison (Nagy, Benedek, 2019), 36 órát vett igénybe Nvidia Geforce GTX 1080 GPU-n. A tanított hálót felhasználva egy kiválasztott $56 \text{ m} \times 111 \text{ m}$ alapterületű, 19M pontot tartalmazó 0.1m-es voxelfelbontású helyszín szemantikus felcímkézése 3 percet vett igénybe.

(Nagy, Benedek, 2019) B. Nagy and Cs. Benedek: "3D CNN Based Semantic Labeling Approach for Mobile Laser Scanning Data," IEEE Sensors Journal, vol. 19, no. 21, pp. 10034 - 10045, 2019

2. *[2. fejezet] A bevezető röviden kitér a mélytanulás alapú megközelítésekre és elválasztja a saját módszereket azoktól. Jó lett volna röviden a jellemző alapú (pl. SIFT) megközelítéseket is hasonló módon elhatárolni.*

Hely szűkében a bevezető és az alapokat bemutató fejezetekben csak a disszertáció új eredményeinek alapját adó tudományos módszerekről adtam áttekintést. Az új eredmények nagy része (1, 2 és 3. téziscsoportok, valamint a 4.1 altézis) a Markovi energiaminimalizációs modellek családjához tartozik, míg az utolsó téziscsoport 4.2 és 4.3 altézisei különböző gépi tanulási módszerek alkalmazásaira épülnek: ezért szerepel ez a két téma a bevezetésben és a *Foundamentals* fejezetben is. A teljesség kedvéért természetesen valóban érdekes lehetett volna több alternatív módszer-csoportot megemlíteni a dolgozatban szereplő módszerekkel összevetésben.

3. *[3. fejezet] A bemutatott módszerek általánosak, nem csak ezekre az alkalmazásokra jók. Bár azt jónak tartom, hogy a Jelölt az általános módszert az elejétől kezdve ezen konkrét alkalmazásokon keresztül mutatta be, mert így könnyű a döntéseket megérteni, de szerencsés lenne kiemelni azokat az általánosítható fogásokat és alkalmazási kört, amit a Jelölt ezzel a módszerrel javasol megoldani.*

Mind az L^3MRF , mind a CXM olyan jellemzőfúzióra alkalmazható modellstruktúrát valósít meg, ahol a fúzió alapja nem egy nagydimenziós jellemzőtérben végzett automatikus dimenziócsökkentés, vagy többdimenziós sűrűségfüggvénnyel történő modellezés, hanem prior, szakértői ismeretek következtetési szabályok formában való megfogalmazása, és annak beépítése a Markovi energiamodellbe.

Az L^3MRF esetén két olyan jellemzőt találtunk, melyek közül

- egyik sem okoz számottevő mennyiségű hamis negatív találatot (a mozgó alakzatokat mindketten nagy biztonsággal jelzik)
- egyenként gyakran, de egyszerre ritkán okoznak hamis pozitív találatot

A fenti megfigyelés miatt a fúziós szabály a két jellemző által javasolt mozgásmaszk ÉS kapcsolatban történő figyelembevétele, amit a simaság fenntartásához Potts-féle szomszédossági tagokkal ötvöztünk. Ez a séma tetszőleges logikai függvényre általánosítható (VAGY, XOR stb), az viszont feladat- és jellemzőfüggő, hogy éppen melyik fúziós szabály építhető be hatékonyan a markovi energiafüggvénybe.

Egy korábbi cikkünkben [SPPRA07] az L^3MRF -hez implementációs szinten nagyon hasonló struktúrájú, de a jellemzők szavazatai között XOR kapcsolatot használó módszert alkalmaztunk az 1.2 tézisben is megcélzott hosszútávú változásdetekciós feladatra. Ez a módszer – ami még a PhD fokozatom megszerzése előtt készült - nem szerepel a disszertációban, mert ugyanerre a feladatra CXM végül általánosabb és hatékonyabb megoldásnak bizonyult, azonban a példa mutathatja a megközelítés általánosíthatóságát.

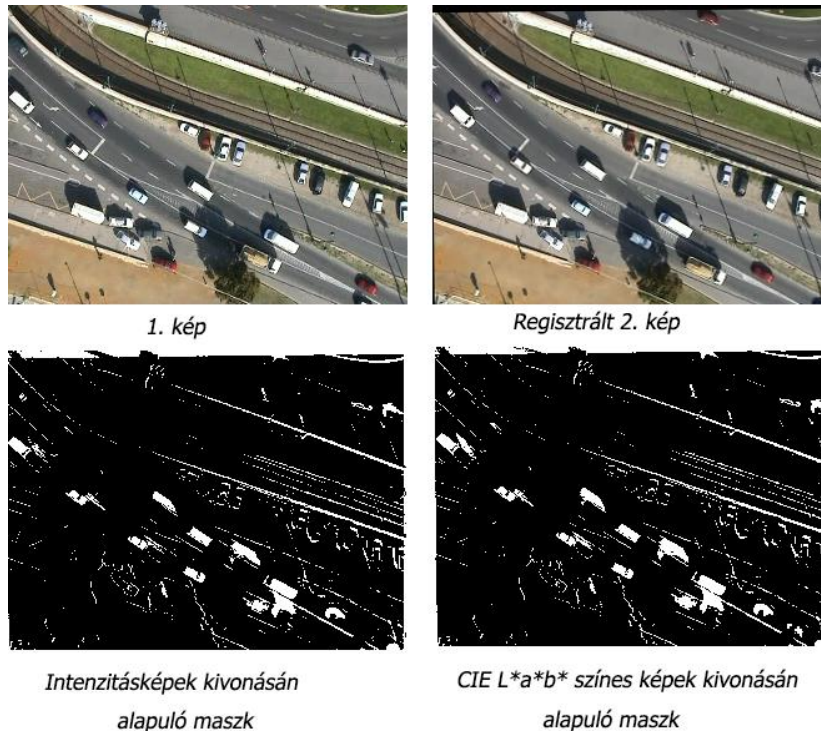
[SPPRA07] Cs. Benedek and T. Szirányi: "Markovian Framework for Structural Change Detection with Application on Detecting Built-in Changes in Airborne Images", IASTED International Conference on Signal Processing, Pattern Recognition and Applications (SPPRA), ACTA, pp. 68-73, Innsbruck, Austria, February 14-16, 2007 <https://eprints.sztaki.hu/4919/>

A CXM-hez kötődő hosszútávú változásdetekciós feladatban a fúziós szabály kicsit bonyolultabb, hiszen itt három jellemzőt fuzionálunk, melyek közül kettő az eredményül várt pixelcímkére szavaz (itt *változás* vagy *háttér*), míg a harmadik a jellemzők lokális megbízhatóságának becsléséhez járul hozzá. Ez a megoldás könnyen általánosítható N különböző jellemzőrétegre is, amelyeket a CXM-nél bemutatottakhoz hasonlóan egyetlen cím réteg (*address node layer*) köt össze. Meglátásom szerint a bemutatott séma receptként alkalmazható lehet más alkalmazásokra is, azonban mivel épp a prior tudás modellezésén alapul a fúziós szabály meghatározása, a jellemzőkiválasztást (*feature engineering*) minden esetben egyedien kell megoldani, vagy a lehetséges fúziós szabályok terében egy automatikus keresést végezni, és a legkisebb hibát adó fúziós szabályt alkalmazni

4. [3. fejezet] *A rövidtávú modell a képregisztráció után két jellemzőt fuzionál, az intenzitás különbséget és a környezetek keresztkorrelációját. Kérdésem, hogyan lehetne a színinformációt felhasználni, annak elhanyagolása ugyanis jelentős információvesztést jelent.*

A intenzitásértékek különbségéként kapott $f_a(s) \in \mathbb{R}$ skálárjellemező helyett használhatjuk 3D színvektorok $\bar{f}_a(s) \in \mathbb{R}^3$ különbségvektorát, amit a statikus területeken háromdimenziós multinomiális Gaussi eloszlással modellezhetünk. Célszerű lehet perceptuálisan uniform – például CIE $L^*a^*b^*$ vagy CIE $L^*u^*v^*$ - színterekben dolgozni, ahol a 3D vektorok távolsága arányos a színek érzékelt távolságával, ekkor a különbségvektor Gauss függvénye is jól közelíthető diagonális kovariancia mátrixszal, így gyorsan számítható. A keresztkorreláció szintén felírható csatornánként is, de a modell azt is lehetővé teszi, hogy az eltérő rétegekben eltérő dimenziójú jellemzőket használjunk, hiszen a jellemzővektorokat nem fuzionáljuk közvetlenül, csak a kapcsolódó címkéiken keresztül vannak hatással egymásra.

Részben a kérdéshez kapcsolódik, hogy korábbi (a PhD fokozat előtti) munkáim során egy videofelügyeleti alkalmazás kapcsán részletesen tanulmányoztam a megfelelő szintér kiválasztásának számszerű jelentőségét mozgó árnyékok szegmentálására [15]. Itt egyes esetekben akár 5-10%-os javulást is tapasztaltam, amikor szürkeárnyaltos képről CIE $L^*a^*b^*$ vagy $L^*u^*v^*$ színreprezentációra váltottam. A légi platformhoz kötődő L^3MRF model kapcsán nem készült ilyen elemzés, de előzetes tesztek azt mutatták, hogy a módszer a rendelkezésre álló alacsony szaturációval rendelkező felvételeken nem javulna számottevően, ezért nem is jártuk ennél jobban körbe a kérdéskört. Az 1. ábra egy ehhez kapcsolódó kísérletet mutat: először a szürkeárnyaltossá konvertált képeket vontam ki egymásból, majd a különbségképet binarizáltam; másodszor pedig a CIE $L^*a^*b^*$ szintérbe konvertált színes képek különbségét binarizáltam közvetlenül: az eredmények nagyon hasonlóak lettek.



1. ábra szürkeárnyaltos illetve CIE $L^*a^*b^*$ színtartománybeli különbségképekészítés összehasonlítása.

[15] Cs. Benedek and T. Szirányi: "Study on Color Space Selection for Detecting Cast Shadows in Video Surveillance", International Journal of Imaging Systems and Technology, Special Issue on Applied Color Image Processing, vol. 17, no. 3, pp. 190-201, Wiley, 2007

5. [3. fejezet] A hosszú távú változásdetektáláshoz a Jelölt még egy újabb leíró alkalmaz, amely a többi leíró lokális megbízhatóságát mutatja. Ezt szép eredeti ötletnek tartom, az eredmények ismét jól alátámasztottak és meggyőzőek. A hosszú távú modellhez kapcsolódóan a regisztráció módjára kérdezek rá és arra, hogy annak pontossága mennyire veszélyezteti a javasolt algoritmus pontosságát

A regisztrációs folyamatra valóban nem tértem ki ennél a modellnél, ugyanis a Földmérési és Távérzékelési Intézet (FÖMI, ma a Lechner Tudásközpont része) által rendelkezésünkre bocsátott ortofotók pixelszinten lényegében tökéletesen regisztrált képekként érkeztek. Mivel itt nagy időtávú vizsgálatokról van szó, és az adatokat gyűjtő és rendszerező cégek mindenképpen elvégzik az ortofotóvá alakítást a további feldolgozás előtt, ez a feltevés sok esetben gyakorlati szempontból is indokolható.

A bírálói kérdéshez kapcsolódóan elvégeztem egy újabb kísérletet. A disszertáció 3.1(b) és 3.7 ábráin is szereplő Szada 2000 – 2005 képpár esetén különböző globális offset értékekkel regisztrációs hibákat szimuláltam, és az így eltolt képeken végeztem el a CXM alapú változásdetektációt. A kvalitatív eredményeket a következő oldalon szereplő 2. ábra mutatja. Azt tapasztalhatjuk, hogy y, és x irányokban 1-1 pixellel szándékosan eltolva az egyik bemeneti képet még közelítően jó eredményt kapunk (mivel a blokk-korrelációs jellemző kismértékű regisztrációs hibákat tud kezelni), ennél nagyobb offset-ekre viszont gyorsan romlik az algoritmus pontossága.



Ground Truth változásmaszk



CXM eredmény regisztrációs hiba nélkül



CXM kimenet $\Delta y=1$, $\Delta x=1$ pixel regisztrációs hibával



CXM kimenet $\Delta y=2$, $\Delta x=2$ pixel regisztrációs hibával



CXM kimenet $\Delta y=3$, $\Delta x=3$ pixel regisztrációs hibával



CXM kimenet $\Delta y=4$, $\Delta x=4$ pixel regisztrációs hibával

2. ábra A hosszútávú változásdetektációt megvalósító CXM módszer kimenetének vizsgálata a képek között szimulált offset típusú regisztrációs hiba függvényében (Szada 2000 és Szada 2005 képekre)

6. *[4. fejezet] A Jelölt a konkrét feladatra megalapozott döntésekkel javasol megoldást. Kérdezem, hogy milyen szintig van lehetőség általános recepteket adni ezen problémák megoldására.*

A 2.1 altézisben bevezetett többidejű jelölt pontfolyamat (tJPM) modell általánosan alkalmazható objektumváltozások észlelésére olyan esetekben, amikor az egyes képeken a vizsgált objektumok JPM modellel hatékonyan kinyerhetők, illetve találunk a rétegek közötti változások jelzésére megfelelő alacsony szintű jellemzőket (az 1. téziscsoportban bemutatott L^3MRF vagy CXM modelleken alapuló eljárások például alkalmazhatók lehetnek erre a feladatra.)

Számos korábban publikált JPM megoldás adat-energiatagja építhető be a javasolt változásdetektációs modellünkbe, így felhasználható lehet épületdetektációra domborzati térképekről (például az [ISPRS13]-ben bemutatott egyrétegű modell integrálásával), fák számlálására és fakivágások észlelésére [ACPR11], [CVRS12], akna és csatornafedelekek detektációja [GRSL14], vagy úthálózatok változásainak jelzésére [PAMI05].

A 2.2 altézisben bemutatott többkeretes pontfolyamat modell bár szintén általános megfogalmazású, felhasználási köre vélhetően jóval szűkebb: egyszerű geometriájú alakzatok kifejezetten rossz minőségű képsorozatokból történő kinyerésére lehet hatékony, hasonlóan a tézisben is bemutatott passzív radaros hajódetektációs feladathoz.

A válaszban említett hivatkozások:

[ISPRS13] M. Bredif, O. Tournaire, B. Vallet, and N. Champion, "Extracting polygonal building footprints from digital surface models: A fully-automatic global optimization framework," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 77, no. 1, pp. 57–65, 2013.

[ACPR11] J. Zhou, C. Proisy, P. Couteron, X. Descombes, J. Zerubia, G. le Maire, and Y. Nouvellon, "Tree crown detection in high resolution optical images during the early growth stages of eucalyptus plantations in brazil," in Asian Conf. on Pattern Recognition, 2011, pp. 623–627.

[CVRS12] Y. Yu, J. Li, H. Guan, C. Wang, and M. Cheng, "A marked point process for automated tree detection from mobile laser scanning point cloud data," in International Conference on Computer Vision in Remote Sensing (CVRS), Xiamen, China, 2012, pp. 140–145.

[GRSL14] Y. Yu, J. Li, H. Guan, C. Wang, and J. Yu, "Automated detection of road manhole and sewer well covers from mobile LiDAR point clouds," IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 9, pp. 1549–1553, Sept 2014.

[PAMI05] C. Lacoste, X. Descombes, and J. Zerubia, "Point processes for unsupervised line network extraction in remote sensing," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no. 10, pp. 1568–1579, 2005.

7. *A dolgozatban sok helyen előjön az algoritmusok paramétereinek beállítása, amely alapvetően befolyásolhatja a konvergencia létét vagy sebességét. A javasolt módszer általában egy fajta tanítás, amikor is ismert képek alapján lőjük be automatikusan ezeket a paramétereket, amelyeket azután már az éles képekre alkalmazunk. Milyen más lehetőség volna még?*

A munka során valóban többnyire a felügyelt tanítású megközelítést alkalmaztuk a paraméterbeállításra, ennek mind a Markovi modelleknél (például maximum likelihood stratégiák,

illetve expectation maximization alapú becslés), mind a neurális hálózatoknál (backpropagation számos speciális kiterjesztéssel) jelentős szakirodalma van, amit felhasználtunk és alkalmaztunk.

Ettől a megközelítéstől valamelyest eltér a 6.3.1.2 fejezetben ismertetett Lidar alapú háttérmodell konstruálása, amely egy felügyelet nélküli on-line k-means eljárással illeszt egy Gaussi keverékeloszlást a mélységkép adott pixelében különböző időpillanatokban mért mélységértékekre, majd a legnagyobb súlyú Gaussi tényező várható értékétől távol eső valamennyi mért értéket előtérhez tartozónak tekinti. Az eljárás tesztjeink szerint robusztusan működik addig, míg a megfigyelési időnek legalább az 50%-ban a háttérpont látszik a pixelben keresztül, vagy ha az előtér objektumok egymástól függetlenül, a szenzortól különböző távolságokban jelennek meg (tehát nem alakul ki egy másik domináns sűrűsödési pont az időbeli statisztikai elemzés során). Az ehhez a példához hasonló felügyelet nélküli klaszterező eljárások akkor működhetnek jól, ha az osztályozás szempontjából „inlier” régiók (például a hosszútávú változásdetekciónál a változatlan képrészletek) területe szignifikánsan nagyobb mint az outlier-eké (pl változásoké), és találunk olyan jellemzőteret, amiben az inlier-ek kompakt régiókat foglalnak el, melyekből az outlier-ek nagy valószínűséggel szignifikánsan kilógnak. Munkám során a problémák döntő többségénél azt tapasztaltam, hogy a vizsgált jellemzőterekben az egyes osztályok közelsége és gyakori átlapolódása miatt a felügyelt tanuláson alapuló paraméterbecslés szükséges volt a különböző kategóriák és objektumok hatékony elkülönítéshez, ezért az esetek többségében ezt az utat követtem. Tapasztalatom szerint a módszerek gyakorlati használatát ez a választás nem nehezíti meg lényegesen. A Markov-i modellek viszonylag kevés paraméterrel rendelkeznek (a neurális hálózatokhoz képest feltétlenül kevesebb paraméterrel), és a megfelelő paraméterek gyakran egy-egy jellemzőhisztogram felrajzolása után könnyen megadhatók, illetve a nehezebben vizsgálható paraméterek beállítására (például a simaságot biztosító δ tényezők, vagy a többszörös születés-halálozás algoritmus konvergenciasebességét meghatározó tényezők) széleskörben bevált gyakorlatok (*best practice*) léteznek, így ezeknek a paramétereknek a jelentős részét én is a szakirodalmi iránymutatások alapján határoztam meg. Távérzékelt adatok feldolgozásánál az automatikus vagy félautomatikus eljárásokat egy eredetileg többnyire manuálisan végzett munkafolyamatba kell beilleszteni, ahol így adott a kezdeti tanítóadat jelenléte, illetve lehetőség van az automatikus eredmények manuális javítására. A mobil lézekszenneres adatfeldolgozásban például már az is nagy segítséget jelentett a Budapest Közút munkatársainak, hogy eljárásainkkal a teljesen manuális annotáláshoz képest az operátor által végzett egérműveletek számát körülbelül a felére sikerült redukálni.

8. Kritikai megjegyzésként megemlítem, hogy a dolgozat több helyen túlságosan formális, az érthetőséget javította volna, ha képletek helyett néhány mondatban fogalmazta volna meg a mondanivalót.

Elfogadom a kritikát, az írás során én is sok esetben mérlegeltem, hogy a matematikai precizitás és a jelölések konzisztenciájának és következetességének a fenntartása, vagy az érthetőség és az olvashatósság miatt a kollokvialisabb stílus alkalmazása a megfelelőbb megoldás. Mivel tervezek a témáról szakkönyvet is írni a közeljövőben, a stilisztikai egyensúly fenntartására e tekintetben is különösképpen figyelni fogok.

Budapest, 2020. június 9-én

Benedek Csaba