

Városi forgalomfelügyelet kétszintű jelölt pontfolyamat modellel légi LiDAR felvételeken

Börcs Attila^{1,2}, Benedek Csaba¹

¹ Elosztott Események Elemzése Kutatólaboratórium, Magyar Tudományos Akadémia,
Számítástechnikai és Automatizálási Kutatóintézet

² Irányítástechnika és Informatika Tanszék, Budapesti Műszaki és Gazdaságtudományi
Egyetem

{borcs.attila,benedek.csaba}@sztaki.mta.hu

Absztrakt. Cikkünkben egy új objektum alapú hierarchikus valószínűségi modellel mutatunk be, melynek célja távérzékelte városi LiDAR pontfelhőkben lévő járművek észlelése és a forgalmi szempontból összetartozó járműcsoportok, *forgalmi szegmensek*, kinyerése. Első lépésként a háromdimenziós pontfelhőket szegmentáljuk, megkülönböztetve a *növényzet*, *járműjelölt*, *épületek tetőszerkezetei*, illetve *ritka pontfelhő* osztályokat. Ezután az egyes pontokhoz tartozó osztálycímkeket és a LiDAR eszköz által mért intenzitás (visszaverődés erősség) értékeket a talaj síkjára vetítjük. Az így kapott 2D címke- és intenzitásképen a felülnézetből látszódó járműveket téglalapokkal közelítjük. Mivel feladatunk egy időben a járművek elhelyezkedését és dimenzióit leíró téglalap populáció megtalálása, valamint az objektumok csoportosítása forgalmi szegmensekbe, egy hierarchikus, kétszintű jelölt pontfolyamat modellel (L^2 MPP - Two-Level Marked Point Process) dolgoztunk ki a probléma megoldására. Az optimális jármű és forgalmi szegmens konfigurációt iteratív sztochasztikus algoritmusmal határozzuk meg. A módszert valódi, összesen 471 járművet tartalmazó légi LiDAR adatakon teszteltük, kvantitatív módon kiértékeljük, és eredményességét két szakirodalmi módszerrel összehasonlítva igazoltuk.

1. Bevezetés

¹Napjainkban az automatikus közlekedéselemzési feladatok központi szerepet töltenek be forgalomirányítási, biztonságtechnikai, civil és katonai városfelügyeleti rendszerekben, céljaik balesetmegelőzéstől kezdve, környezetvédelmi szempontokon érvényesítésén keresztül, a közlekedési lámpák optimális összehangolásáig terjedhetnek. A járműforgalom hatékony automatikus elemzése azonban komplex, hierarchikus modellezési megközelítést kíván. Rendszerünknek az érzékelés szintjén képesnek kell lennie detektálni és elkülöníteni az egyes járműveket, míg magasabb szinten azonosítani kell különböző forgalmi helyzeteket, közlekedési statisztikákat gyűjteni és megjeleníteni, szükség esetén riasztás küldve az operátornak vagy a beavatkozó moduloknak. Ennek a

¹ A cikkben közölt eredmények eredetileg angol nyelven, az *ISPRS Congress 2012* [1] és *ICPR 2012* [2] konferenciák kiadványaiban jelentek meg.

folyamatnak az egyik alapvető lépése az összetartozó járműcsoportok, úgynevezett *forgalmi szegmensek* megtalálása, ami az itt bemutatásra kerülő munkánk központi eleme. Megközelítésünkben egy-egy forgalmi szegmensbe tartoznak például egy parkolóban, vagy az úttest szélén párhuzamosan várakozó járművek, egy közlekedési lámpa előtt felsorakozott autósor tagjai, vagy a többsávos utakon párhuzamosan közlekedő gépkocsik. Munkánk során egy kétszintű algoritmikus megoldást dolgoztunk ki, ami egyszerre képes a jármű objektumok vizuális információ alapján történő pontos lokalizációjára, ami önmagában is komoly kihívást jelentő feladat, és az aktuális forgalmi helyzet *szegmentálására*, tehát az objektumcsoportok azonosítására. A két feladatot azonban nem szekvenciálisan végezzük, hanem egy iteratív keretben biztosítjuk a rétegek kölcsönhatását. Ezáltal lehetővé válik csoport szintű információk kinyerése és visszavezetése a járműdetekciós szintre, hozzájárulva az alsóbb szintű felismerés finomításához.

Bár a forgalomfigyelést ma még a legtöbb esetben földi szenzorokkal, például videokamerákkal vagy indukciós hurkokkal végzik, a légi vagy űrből származó mérési adatok szerepe szintén egyre jelentősebb, mivel nagyobb látószögben adnak információt a megfigyelt helyszínről, és fentről tekintve az utcai objektumok is kevésbé takarják egymást. A szakirodalom több kapcsolódó módszere (optikai) légi fotókon vagy videoszekvenciákon végez járműfelismerést. Azonban pusztán optikai bemenetre alapozva rendkívül nehéz általánosan alkalmazható megoldást találni, mivel a mért képi adat több szempontból is heterogén: a képminőségre nagy hatással vannak a kameraszenzorok különböző paraméterei valamint a változó időjárás, évszaki és napszaki megvilágítási körülmények, és kihívást jelentenek járművek változatos megjelenési formái az optikai képeken [3].

A LiDAR (Light Detection and Ranging, “fény érzékelés és távmérés”) lézer alapú technológia jól kihasználható előnyöket biztosít a fenti problémák kezelésére mivel egyszerre szolgáltat nagy pontosságú közvetlen geometriai információt, úgynevezett pontfelhőt, a helyszínről, és kiegészítő adatokat ad az egyes strukturális elemek felületi (lézer)fényvisszaverő képességéről (intenzitás) és a felületek tömörségéről (adott irányból történő visszaverődések száma). A LiDAR mérések e mellett nagy mértékben függetlenek az időjárás tényezőktől és külső fényviszonyok változásaitól.

A korábbi LiDAR alapú járműdetekciós módszerek többsége két különböző szemléletmód egyikét követi: *cella alapú*, vagy *pontfelhő alapú* modellt valósít meg [4]. Az első csoportba tartozó eljárások [5, 6] a LiDAR pontfelhőből először egy 2.5 dimenziós adatrepresentációt, úgynevezett digitális magasságtérképet (Digital Elevation Model, DEM) állítanak elő, majd képfeldolgozási technikákat alkalmazva keresik meg a járművek 2D felülnézeti határoló pontjait. A második megoldási mód [7] esetén a járműfelismeréshez használt jellemzőket közvetlenül 3D-s pontfelhőből nyerik ki, és itt határozzák meg az egyes autók befoglaló térrészeket, elkerülve így a síkra vetítésből illetve a hiányzó adatok interpolációjából adódó információvesztést. Ugyanakkor az ilyen típusú módszerek erőforrásigénye magasabb, mint a cella alapú eljárásoké, és robosztusság szempontjából is kevésbé kiforrott technikákat eredményeznek. Az általunk javasolt módszer e téren egy hibrid megoldásnak tekinthető. A pontfelhő régiói háromdimenziós jellemzők alapján kerülnek osztályozásra, viszont ezután az osztálycímkéket és intenzitásértékeket a talaj síkjára vetítjük, így az optimális objektumkonfigurációt már egy 2D pixelrácson határozzuk meg. Hogy biztosítsuk megközelítésünk robosz-

tusságát, a járműforgalmat jelölt pontfolyamat (Marked Point Process, MPP) [8] modell segítségével írjuk le, amely egy hatékony Bayesi módszer családjának tagja, ami többek között objektumpopulációk jellemzésére használható digitális képeken. A MPP megközelítésben egyszerre tudunk figyelembe venni különböző adatfüggő objektummodelleket, továbbá prior geometriai megkötések segítségével úgynevezett *gyenge kényszereket* (soft-constraint) állíthatunk az objektumok közötti kapcsolatokra, azaz interakciókra is. Az irodalomban eddig közölt MPP modellek azonban csak korlátozottan alkalmasak hierarchikus környezetértelmezésre, mivel az interakciós tagjaik többnyire páros objektum kapcsolatokat kezelnek előre rögzített szimmetrikus objektum szomszédság értelmezésével. Valós forgalmi helyzetekben például gyakran találunk orientáció tekintetében rendezett járműcsoportosulásokat, ami egy szomszédosságon belüli *orientáció-simító* interakciós adattaggal modellezhető. Azonban problémát jelent, hogy a csoportokon belül gyakran előfordulnak szabálytalanul, a többiekétől eltérő szögben parkoló autók, melyeket nem kívánatos prior kényszerek alapján a csoport tagjaihoz illeszkedően “beforgatni”, hanem éppen kiugró, potenciálisan problémát okozó objektumként kellene jelezni (4. ábra). Ugyanitt megfigyelhető, hogy a járműcsoportoknak gyakran vékony és hosszúkas alakja van, ahol a centrálisan szimmetrikus szomszédosságra alapuló paramétersimítás nem bizonyul hatékony megoldásnak, az ugyanis épp a vékony kiugró részleteket hajlamos eltüntetni. A fenti modell-limitációk kiküszöbölésére egy újszerű, kétszintű jelölt pontfolyamat (Two-Level Marked Point Process, L²MPP) modellt dolgoztunk ki, ami felső szinten *forgalmi szegmensekre* bontja a populációt, alsó szinten pedig minden szegmensre meghatározza a hozzá tartozó járművek pozícióját és elhelyezkedését. A forgalmi szegmensek és a járművek kinyerése egy közös energia minimalizációs folyamat segítségével egy időben történik. Másik fontos újdonság, hogy *adaptívan alakítható objektum-szomszédosságot* megvalósítva, különbözőképpen kezeljük az azonos szegmenseken belül, és a különböző szegmensekhez tartozó objektumok interakcióit, kiküszöbölve a fix szomszédság alapvető korlátait.

A módszer tesztelését valós légi LiDAR adatbázisokon végeztük, ami összesen 471 járművet tartalmaz. Eljárásunk eredményességét két szakirodalmi módszerrel [5, 9] vetettük össze, és hatékonyságát igazoltuk.

2. Pontfelhők szegmentálása

Eljárásunk bemenete egy LiDAR pontfelhő, \mathcal{L} , amely l pontot tartalmaz $\mathcal{L} = \{p_1, \dots, p_l\}$. Valamennyi $p \in \mathcal{L}$ pont rendelkezik pozíció, intenzitás és visszaverődési sorszám paraméterekkel, amit az 1. táblázatban részletezünk. Egy adott pont ϵ -szomszédosságát $\mathcal{V}_\epsilon(p)$ -vel jelöljük:

$$\mathcal{V}_\epsilon(p) = \{q \in \mathcal{L} : \|q - p\| < \epsilon\},$$

ahol $\|r - p\|$ az r és p pontok térbeli pozíciójának euklideszi távolságát jelenti. $|\mathcal{V}_\epsilon(p)|$ -vel jelöljük a szomszédosság elemszámát. A térben közeli pontok hatékony meghatározásához első lépésben particionáltuk a pontfelhőt a k -d fa adatstruktúrát felhasználva.

A pontfelhő kezdeti szegmentálását egy 3D térben definiált energiaminimalizációs módszerrel végezzük, ami különböző háromdimenziós leírókat használ fel az egyes pontosztályok azonosítására. Modellünkben *talaj*, *növényzet*, *háztető*, *jármű* és *ritka régió* területeket különítünk el és $\xi(p)$ -val jelöljük az egyes pontok osztálycímekjét.

Paraméter	értékkészlet	Leírás
x_p, y_p, z_p	\mathbb{R}^3	a p pont 3-D geometriai koordinátái a világkoordináta rendszerben
g_p	[0,255]	a p ponthoz rendelt intenzitásérték (lézervisszaverődés erőssége)
n_p	{1, 2, 3, 4}	lézervisszaverődések (ekhók) száma a p pont irányából
r_p	{1, 2, 3, 4}	a p ponthoz rendelt visszaverődés sorszáma az egy irányból érkező ekhók között ($r_p \leq n_p$)

1. táblázat: Az egyes p pontokhoz rendelt leíró adatok a bemeneti \mathcal{L} pontfelhőben

Az osztályozás elvégzéséhez valamennyi ξ osztályhoz definiálunk egy $\mu_\xi(p) \in [0, 1]$ inverz tagsági függvényt, ami az írja le, hogy a $p \in \mathcal{L}$ pont mennyire illeszkedik a ξ szegmentációs osztály modelljéhez, kisebb értékek jelzik a jobb illeszkedést. A tagsági függvények származtatásához ζ szigmoid függvényeket használunk, amik fel foghatók gyenge küszöbfüggvényeknek (*soft threshold functions*):

$$\zeta(x, \tau, m) = \frac{1}{1 + \exp(-m \cdot (x - \tau))}.$$

ahol $x \in \mathbb{R}$ egy skalár értékű fitnessjellemző, τ az x -hez tartozó gyenge elfogadási küszöb, m pedig a normalizáláshoz használt meredekség.

A *talaj* pontosztály leírásához feltételezzük, hogy a talajrészletek az egyes terület-szegmenseken belül jól közelíthetők sík felületekkel. Ezért minden pontfelhőrészletre egy RANSAC alapú eljárás segítségével megbecsüljük a domináns T talajsíkot, majd ettől a síktól mért $d_p^T = \text{dist}(p, T)$ távolság függvényében értékeljük ki a pontokat:

$$\mu_{\text{talaj}}(p) = \zeta(d_p^T, \tau_{\text{talaj}}, m_{\text{talaj}}),$$

ahol a τ_{talaj} magasság-küszöbértéket és a m_{talaj} normalizáló paramétert felügyelt módon, tanító régiók segítségével határozzuk meg, mivel ezek nagyban függenek a bemeneti pontfelhő zajsztintjától és pontsűrűségétől.

A *növényzet* kiszűrése érdekében megvizsgáltuk a pontokhoz tartozó visszaverődési számot (ekhót). Ahogy az 1. táblázatban részletezzük, a LiDAR lézerekken a pontok háromdimenziós koordinátái mellett képes eltárolni az adott p pont irányában kibocsátott lézervisszaverődéshez tartozó összes visszaverődés számát (n_p) és a p ponthoz tartozó visszaverődés sorszáma is (r_p). Tipikusan a növényzettel borított régiókról többszörösen verődik vissza a lézervfény ($r_p < n_p$, azaz $n_p - r_p \geq 1$), ezt vizsgálva becsülhetők a pontfelhőben lévő növényzettel borított területek:

$$\mu_{\text{növény}}(p) = 1 - \zeta(n_p - r_p, 0.5, m_n)$$

A *háztető* pontosztályt illetően feltételezzük, hogy a pontok d_p^T magasságértéke adott $\tau_{\text{tető}}$ küszöböt meghalad, valamint a pontok sűrű régiókat alkotnak, azaz $|\mathcal{V}_\epsilon(p)| > \tau_V$. A kapcsolódó adattag így:

$$\mu_{\text{tető}}(p) = \left(1 - \zeta(d_p^T, \tau_{\text{tető}}, m_{\text{tető}})\right) \cdot \left(1 - \zeta(|\mathcal{V}_\epsilon(p)|, \tau_V, m_V)\right)$$

A ritka régiókban az előző esettel ellentétben kevés szomszédos pontot várunk:

$$\mu_{ritka}(p) = \zeta(|\mathcal{V}_\epsilon(p)|, \tau_\nu, m_\nu)$$

Végül a járművekhez tartozó pontok esetén azt várjuk, hogy a talajsíktól számított távolságuk egy minimális (τ_{jmin}) és egy maximális (τ_{jmax}) magasságérték közé essen, valamint az utolsó visszaverődést adják a hozzájuk tartozó irányból:

$$\mu_{jármű}(p) = \zeta(d_p^T, \tau_{jmax}, m_{jármű}) \cdot \left(1 - \zeta(d_p^T, \tau_{jmin}, m_{jármű})\right) \cdot \zeta(n_p - r_p, 0.5, m_n)$$

A tagsági függvények meghatározása után a lehetséges globális pontfelhő-címkézések terén értelmezünk egy E energiafüggvényt, ami a szomszédossági kapcsolatok leírásához a Potts modellt használja [10]-hoz hasonlóan:

$$E(\{\xi(p)|p \in \mathcal{L}\}) = \sum_{p \in \mathcal{L}} \mu_{\xi(p)}(p) + \sum_{p \in \mathcal{L}} \sum_{r \in \mathcal{V}_\epsilon(p)} \kappa \cdot \mathbf{1}\{\xi(p) \neq \xi(r)\} \quad (1)$$

ahol $\kappa > 0$ az interakciós tag súlytényezője, $\mathbf{1}\{\cdot\}$ indikátor függvény: $\mathbf{1}\{\text{igaz}\} = 1$, $\mathbf{1}\{\text{hamis}\} = 0$.

A (1) energiát minimalizálására jó közelítő megoldást adható gráf vágás alapú optimalizációs módszerekkel, amit [11] implementációját használva mi is teszteltünk. Ugyanakkor azt is tapasztaltuk, hogy a pontonkénti (azaz a Potts simítótagot figyelmen kívül hagyó) szegmentáció eredményénél már a gyors Iterated Conditional Modes (ICM) optimalizáció is számottevő javulást eredményez, amit az 1. ábra is szemléltet.

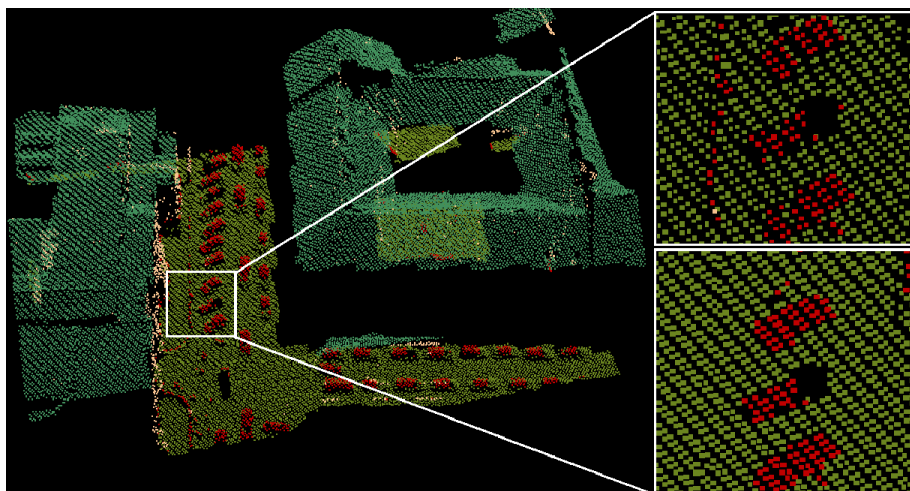
A 3-D térben végzett pontfelhő szegmentáció után egy diszkrét 2-D pixelrácstot feszítünk a talaj síkjára, ahol $s \in S$ jelöl egy tetszőleges pixelt. Ezután minden olyan LiDAR pontot levetítünk erre a rácsra, aminek a címkéje *talaj* (világoszöld az 1. ábrán), *jármű* (piros) vagy *háztető* (sötétzöld). A projekció eredménye egy 2-D címketérkép, ahol az egy rácspontra eső többszörös pontprojekciókat egy kiválasztó algoritmussal kezeltük, ami nagyobb precedenciát ad a jármű osztályhoz tartozó pontoknak. Másrésztől, a lokálisan több helyen ritka ponthalmaz vetítése egy szabályos pixelrácstra számos kimaradó pixel eredményez, ahol a címkeérték definiálatlan. Több korábbi megoldással [5, 6] ellentétben, algoritmusunk nem interpolálja ezeket a hiányzó adatokat, hanem egy külön „definiálatlan” címkeosztályt rendel hozzájuk (fekete pixelek a 2. (a) és (e) ábrákon). Így elkerüljük az interpolációból adódó információvesztést.

Jelöljük $\chi(s) \subset \mathcal{L}$ -vel azon pontok halmazát, amik a s pixelre vetülnek. A pontok címkéitől ($\xi(p)$) való egyértelmű megkülönböztetés céljából az s pixelhez rendelt címkét $\nu(s)$ -sel jelöljük. A projekció után jármű, háttér és definiálatlan osztályokat különböztetünk meg, az alábbi módon:

$$\nu(s) = \begin{cases} \text{jármű} & \text{ha } \exists p \in \chi(s) : \xi(p) = \text{jármű} \\ \text{háttér} & \text{ha } \forall p \in \chi(s) : \begin{cases} \xi(p) = \text{háztető} \\ \text{VAGY} \\ \xi(p) = \text{talaj} \end{cases} \\ \text{definiálatlan} & \text{ha } \chi(s) = \emptyset. \end{cases}$$

Az osztálycímkék mellett minden s pixelhez egy $g(s)$ intenzitásértéket is rendelünk, ami 0, ha $\nu(s) = \text{definiálatlan}$, egyébként az s pixelre vetülő pontok intenzitásértékeinek az átlagát vesszük.

Módszerünk további lépéseiben kizárólag az itt kinyert címke- és intenzitásképeken dolgozunk. A felismerés főként a címketérképen alapul, de kiegészítő információként felhasználjuk az intenzitást is, figyelembe véve hogy számos jármű kiugró világos foltként jelenik meg a csillogó felülete miatt.

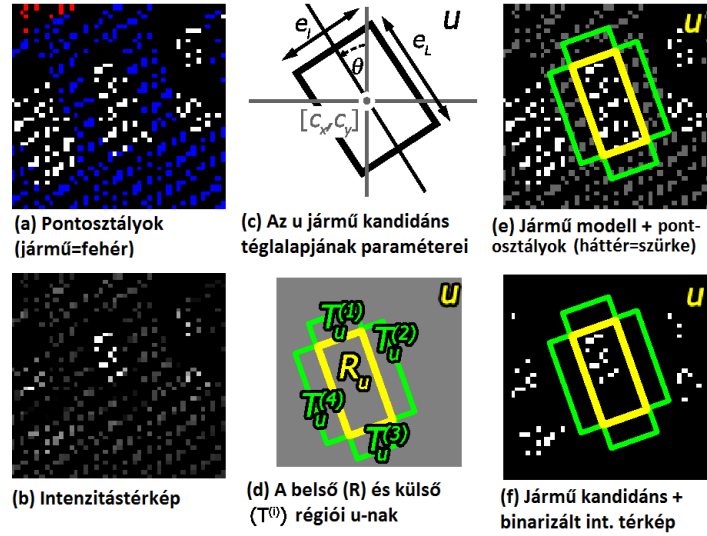


1. ábra: Pontfelhő szegmentációjának eredménye egy minta adatszeleten. Jobbra fent: a pontonkénti (azaz a Potts szomszédossági simítótágot figyelmen kívül hagyó) osztályozás eredménye. Jobbra lent: az (1) energiafüggvény ICM alapú közelítő optimalizációjával nyert zajmentes osztályozás.

3. Kétszintű jelölt pontfolyamat modell (L^2MPP)

A pontfelhő vetítését követően a forgalomelemzést egy kétdimenziós S pixelrácson végezzük, a rendelkezésünkre álló osztálycímke- és intenzitástérképekre támaszkodva. A járműfelismeréshez használt jellemzők többségét a címkékéről származtatjuk, viszont segédinformációként kihasználjuk azt az észrevételt is, hogy a csillogó felületük miatt a járművek többsége fényes foltként jelenik meg a LiDAR intenzitásképen. Modellünkben az egyes járművek sziluettjét felülnézetből téglalapokkal közelítjük.² Egy u járműjelöltet így öt paraméterrel írunk le: c_x és c_y a téglalap középpont koordinátái az S pixelrácson, e_L , e_l a téglalap oldalhosszai, illetve θ szög az orientáció (2. (c) ábra). A téglalapok \mathcal{H} -val jelölt objektumtérben definiálunk egy \sim szomszédossági kapcsolatot:

² Megjegyezzük, hogy a szakirodalomban egyes megoldások [4] a téglalapok helyett paralelogrammákat használnak a járművek pontfelhő projekcióinak leírására, mivel a pásztázó LiDAR technikák által szolgáltatott adatokon a mozgó járművek egy “nyíró effektust” követve torzulhatnak. MPP alapú modellünk szintén kiterjeszthető ebben az irányban, viszont az általunk használt tesztadatokon nem figyelhető meg ez a jelenség.



2. ábra: Az (a)-(b) bemenetek, (c) objektum befoglaló téglalapok és (d)-(f) adatagok kiszámítása.

$u, v \in \mathcal{H}$ esetén $u \sim v$ fennáll, amennyiben a téglalapok középpontjai közelebb vannak egymáshoz egy előre meghatározott küszöbértéknél.

A következő lépésben a járműforgalmat leíró újszerű, kétszintű pontfolyamat modellünket (L^2 MPP) definiáljuk. Jelöljön ω egy forgalmi konfigurációt, amely k forgalmi szegmens halmaza: $\omega = \{\psi_1, \dots, \psi_k\}$. Egy adott ψ_i ($i = 1 \dots k$) forgalmi szegmens n_i jármű tetszőleges konfigurációja: $\psi_i = \{u_1^i, \dots, u_{n_i}^i\} \in \mathcal{H}^{n_i}$. Előírjuk, hogy minden $i \neq j$ esetén $\psi_i \cap \psi_j = \emptyset$. A szegmensek k darabszáma és az egyes szegmensekhez tartozó járműhalmazok számossága n_1, \dots, n_k tetszőleges (kezdetben ismeretlen) nemnegatív egész értéket vehet fel. A későbbi egyszerűbb jelölés kedvéért bevezetjük a $u \prec \omega$ relációt is, ami fennáll, ha az u objektum a ω konfiguráció bármelyik szegmensének eleme: $\exists \psi_i \in \omega : u \in \psi_i$. Ω -val írjuk le a teljes forgalom-konfigurációs teret, azaz összes előfordulható ω forgalmi konfiguráció halmazát. Ezután, inverz modellezést követve, definiálunk egy $\Phi(\omega)$ energiatagot, amely alkalmas egy tetszőleges $\omega \in \Omega$ konfiguráció kiértékelésére, annak függvényében, hogy mennyire illeszkedik a megfigyelt adathoz, illetve egyes előre definiált prior geometriai kényszerekhez. Ezért az energiatagot egy adatfüggő (Φ_d) és egy prior energiatag (Φ_p) összegeként definiáljuk: $\Phi(\omega) = \Phi_d(\omega) + \Phi_p(\omega)$. Végül az ω_{opt} optimális forgalmi konfigurációt a $\Phi(\omega)$ függvény minimalizálásával közelítjük.

3.1. Adatfüggő energiatagok

Az adatfüggő energiatagok az egyes $u = \{c_x, c_y, e_L, e, \theta\}$ járműjelöltek kiértékelését végzik az osztálycímké- és intenzitástérképek alapján, számításuk független azonban a forgalmi konfiguráció többi objektumától. A adatmodellezési folyamat két lépésből

áll. Először definiálunk különböző $f(u) : \mathcal{H} \rightarrow \mathbb{R}$ képi jellemzőket, amelyek egy-egy u járműhipotézis fitnessz értékét határozzák meg különböző szempontok alapján, úgy hogy ‘magas’ $f(u)$ érték tartozik a hatékonynak ítélt objektumokhoz. Második lépésben származtatunk egy $\varphi_d^f(u)$ energiatagot minden f jellemzőre, olyan módon, hogy $\varphi_d^f(u) < 0$ teljesüljön a valós objektumok esetén és $\varphi_d^f(u) > 0$ tartozzon a hamis jelöltekhez. A negatív értékű energiataggal rendelkező téglalapokat *attraktív objektumnak* nevezzük, mivel a járműpopulációhoz hozzáadva azokat, képesek lehetnek a konfigurációs energiát csökkenteni. Az energiatagok származtatásakor az f jellemzők különböző értékészleteit a $[-1, 1]$ intervallumra vetítjük le egy monoton csökkenő függvény segítségével: $\varphi_d^f(u) = \mathcal{Q}(f(u), d_0^f)$ [1, 8], ahol

$$\mathcal{Q}(x, d_0) = \begin{cases} \left(1 - \frac{x}{d_0}\right), & \text{ha } x < d_0 \\ \exp\left(-\frac{x-d_0}{0.1}\right) - 1, & \text{ha } x \geq d_0. \end{cases} \quad (2)$$

A fenti képletben láthatjuk, hogy a \mathcal{Q} függvény rendelkezik egy d_0^f paraméterrel, ami objektumelfogadási küszöbnek tekinthető az f jellemzőre vonatkozólag: u akkor és csak akkor *attraktív* $\varphi_d^f(u)$ alapján, amennyiben $f(u) > d_0^f$. Az egyes jellemzőkhöz tartozó elfogadási küszöbök minden esetben felügyelt módon, pozitív és negatív tanítóminták alapján határozzuk meg.

Modellünkben négy különböző adatfüggő jellemzőt használunk. Jelölje $R_u \subset S$ a képen a járműjelölt téglalapjának belsejébe eső pixelek halmazát, továbbá jelölje T_u^{up} , T_u^{bt} , T_u^{lt} , and T_u^{rg} az objektum felső, alsó, baloldali és jobboldali szomszédos régióját, amint azt a 2. (d) ábra szemlélteti. A felhasznált jellemzők definíciói a következők:

- I *Járműcímkék fedési aránya* $f^{\text{jf}}(u)$: megszámlaljuk az R_u téglalapon belüli, járműpontként klasszifikált pixeleket és a kapott értéket normalizáljuk az R_u téglalap területével.

$$f^{\text{jf}}(u) = \frac{1}{|R_u|} \sum_{s \in R_u} \mathbf{1}\{\nu(s) = \text{jármű}\},$$

ahol $|R_u|$ az R_u halmaz számosságát jelöli és $\mathbf{1}\{.\}$ itt is indikátor függvény.

- II *Külső háttérfedési tényező* $f^{\text{kh}}(u)$: meghatározzuk a jelölt téglalap szomszédosságában lévő négy külső régióban mért háttércímkék fedési arányát, és a második legkisebb értéket választjuk ki közülük:

$$f^{\text{kh}}(u) = \min_{i \in \{\text{up}, \text{bt}, \text{lt}, \text{rg}\}} \left(\frac{1}{|T_u^i|} \sum_{s \in T_u^i} \mathbf{1}\{\nu(s) = \text{háttér}\} \right)$$

így előírjuk, hogy az objektumot legalább három oldalról háttérként sűrűn klasszifikált pixelrégiók vegyék körül, viszont megengedjük, hogy legfeljebb az egyik oldala mentén megsértse ezt a feltételt, ami előfordulhat, ha az autó közvetlenül egy másik gépkocsi, vagy egy nagyobb ‘árnyék’ régió mellett helyezkedik el.

- III *Belső háttérfedési tényező* $f^{\text{bh}}(u)$: kiszámítjuk a háttérpixelek fedési arányát R_u -n belül. Valós objektumok belsejében csak kevés háttércímkéjű pixel jelenlétét engedélyezzük:

$$f^{\text{bh}}(u) = \frac{1}{|R_u|} \sum_{s \in R_u} 1 - \mathbf{1}\{\nu(s) = \text{háttér}\}.$$

A f^{if} , f^{kh} és f^{bh} jellemzők bemutatása a 2(e) ábrán látható.

IV *Intenzitás tényező*: kiszámítjuk, hogy a jármű-téglalap jelölt milyen arányban fed magas fényességi értékű területeket az intenzitástérképen:

$$f^{it}(u) = \frac{1}{|R_u|} \sum_{s \in R_u} \mathbf{1}\{g(s) > T_g\},$$

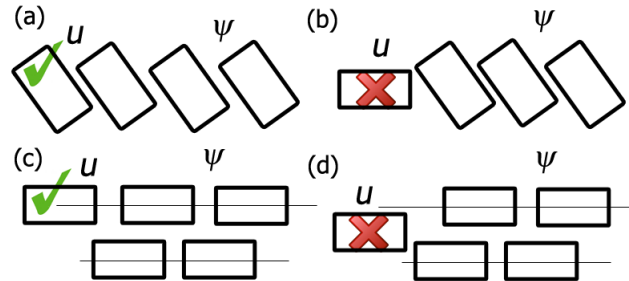
ahol T_g egy intenzitás küszöbérték.

A jellemzők definiálását követően a $\varphi_d^{if}(u)$, $\varphi_d^{kh}(u)$, $\varphi_d^{bh}(u)$, $\varphi_d^{it}(u)$ adattagok kiszámolhatók a \mathcal{Q} függvény segítségével. Az u objektum összesített adat-energiatagja az egyes jellemzőkhöz tartozó tagokból közvetlenül származtatható:

$$\varphi_d(u) = \max(\min(\varphi_d^{if}(u), \varphi_d^{it}(u)), \varphi_d^{kh}(u), \varphi_d^{bh}(u)).$$

Az itt alkalmazott min és max műveletek ekvivalensek a logikai “vagy”, illetve logikai “és” operátorokkal. általában az egyes jellemzők által támasztott kényszerek együttes teljesülését várjuk el, de nem írjuk elő a *járműcímke fedési és intenzitás* jellemzők egyidejű megköthetését, hiszen egyrészt nem minden jármű jelenik meg fényes foltként az intenzitás térképen, másrészt az előzetes pontfelhő-szegmentáció szintén tartalmazhat hibásan klasszifikált jármű pontokat. Az ω globális forgalmi konfiguráció adat-energiatagját az egyes önálló objektumok energiáinak összegeként számítjuk: $\Phi_d(\omega) = \sum_{u \prec \omega} \varphi_d(u)$.

3.2. Prior energiatagok



3. ábra: Előnyben részesített (\checkmark), illetve büntetett (\times) jármű konfigurációk egy forgalmi szegmensben belül

A prior energiatagok szerepe, hogy az adott ω forgalmi konfigurációt a benne szereplő objektumok kölcsönhatásait figyelembe véve értékeljék ki, így különböző geometriai kényszerek teljesülését írják elő. Modellünkben a prior tag két fő részből áll:

$$\Phi_p(\omega) = \sum_{\substack{u, v \prec \omega \\ u \sim v}} I(u, v) + \sum_{u \prec \omega, \psi \in \omega} A(u, \psi) \quad (3)$$

Mivel különálló járműveket keresünk, a $I(u, v)$ átlapolódási koefficiens bünteti az ω konfiguráción belül két objektum befoglaló téglalapjának az átfedését hasonlóan korábbi MPP megoldásokhoz [12, 13]:

$$I(u, v) = \frac{\text{Area}\{R_u \cap R_v\}}{\text{Area}\{R_u \cup R_v\}}. \quad (4)$$

Modellünk módszertani újdonságát a két szemantikai réteg (jármű és járműcsoport) közötti kölcsönhatások bevezetése jelenti, amit a járművek és forgalmi szegmensek közötti $A(u, \psi)$ interakciós tagon keresztül valósítunk meg. Először definiálunk egy $d_\psi(u) \in [0, 1]$ távolságmértéket, ami az u jármű illeszkedését jellemzi a ψ forgalmi szegmenshez. $d_\psi(u)$ -t két adatag átlagaként számoljuk. Az *első* a normalizált szögeltérés az u objektum orientációja és a ψ forgalmi szegmens belüli objektumok átlagos orientációja között, így büntetjük, ha például egy párhuzamosan parkoló autócsoporttal egy szegmensbe kerül egy tőlük jelentősen eltérően orientációval rendelkező jármű (3. ábra (a)-(b)). A *második tag* meghatározásához a RANSAC algoritmus segítségével egy, vagy több párhuzamos egyenest fektetünk a ψ szegmens belüli objektumok középpontjaira, és kiszámítjuk a távolságot a legközelebbi egyenes és az u objektum középpontja között. Így büntetünk olyan, adott szegmenshez tartozó járműveket is, melyek bár párhuzamosak a csoport többi tagjával, elhelyezkedésükkel eltérnek a rendezett soroktól vagy oszlopoktól (3. ábra (c)-(d)).

Annak érdekében, hogy biztosítsuk térben összefüggő forgalmi szegmensek létrejöttét, egy konstans magas $d_\psi(u)$ távolságértéket használunk, ha az u objektumnak nincs szomszédja ψ szegmens belüli, a korábban definiált \sim reláció alapú objektum-szomszédosságot használva. Az így módosított távolságmértéket a következő módon számoljuk:

$$\hat{d}_\psi(u) = \begin{cases} 1 & \text{ha } \nexists v \in \psi \setminus \{u\} : u \sim v \\ d_\psi(u) & \text{különben} \end{cases}$$

A (3) formulában használt $A(u, \psi)$ interakciós tagot a $\hat{d}_\psi(u)$ távolságérték függvényeként definiáljuk. Bár lehetővé tesszük egy elemű szegmensek létrehozását (például egy szabálytalanul parkoló autó esetén), többségében a nagyobb csoportok létrejöttét részesítjük előnyben. Ezért enyhén büntetjük azokat a szegmenseket, amelyek csak egy járművet tartalmaznak: $A(u, \psi) = c$ akkor és csak akkor, ha $\psi = \{u\}$, ahol $0 < c \ll 1$ konstans. Egyébként, *büntetjük* a nagy $\hat{d}_\psi(u)$ távolság értéket egy adott ψ szegmens és egy általa tartalmazott $u \in \psi$ objektum között. Amennyiben viszont $u \notin \psi$, a kis \hat{d} távolság esetén számítunk fel büntetőtagot:

$$A(u, \psi) = \mathbf{1}_{u \in \psi} \cdot \hat{d}_\psi(u) + \mathbf{1}_{u \notin \psi} \cdot (1 - \hat{d}_\psi(u))$$

ahol $\mathbf{1}_E \in \{0, 1\}$ E esemény indikátor függvénye.

4. Optimalizáció

Az optimális objektum konfigurációt iteratív sztochasztikus eljárással közelítjük. MPP-energiaoptimalizálására hatékony megoldást kínál a többszörös születés és halál di-

namika (Multiple Birth and Death Dynamics, MBD) módszer [12–14]. Ezt az alapalgoritmust módosítottuk, hogy a kétszintű MPP modellünk energiafüggvényét is kezelni tudjuk. A módosított algoritmus lépései a következők:

Inicializáció: induljunk ki egy üres forgalmi konfigurációból, azaz $\omega = \emptyset$, inicializáljunk megfelelő egy b_0 születési gyakoriság értéket, kezdeti $\beta = \beta_0$ inverz hőmérséklet paramétert a $\delta = \delta_0$ diszkrétizációs lépcsőt [14]-ben leírt útmutatásokat követve.

Fő program: iteráljuk a következő három lépést:

- **Születés:** látogassuk meg egymásután az S pixelrács pixeleit. Minden s pixelben δb_0 valószínűséggel generáljunk egy s középponttal rendelkező u objektumot, és az e_L, e_l, θ paramétereit állítsuk be véletlenszerűen, az előforduló járműméretekről rendelkezésre álló prior statisztikákat követve. Az új járműveket vagy hozzávesszük egy meglévő szegmenshez, vagy új szegmenst hozunk létre számára a következők szerint: Minden egyes új u objektumhoz $p_u^0 = \mathbf{1}_{\omega=\emptyset} + \mathbf{1}_{\omega \neq \emptyset} \cdot \min_{\psi_j \in \omega} \hat{d}_{\psi_j}(u)$ valószínűséggel generáljunk egy üres ψ forgalmi szegmenst, és adjuk hozzá u objektumot. Végül a ψ szegmenst adjuk hozzá az ω globális objektum konfigurációhoz. Egyébként adjuk hozzá u objektumot egy létező $\psi_i \in \omega$ forgalmi szegmenshez az alábbi valószínűséggel: $p_u^i = (1 - \hat{d}_{\psi_i}(u)) / \sum_{\psi_j \in \omega} (1 - \hat{d}_{\psi_j}(u))$.

- **Halál:** tekintsük az aktuális ω forgalmi konfiguráció összes objektumát, és rendezzük őket sorba $\varphi_d(u) + A(u, \psi)|_{u \in \psi}$ alapján, csökkenő sorrendben. Az objektumokat vizsgáljuk meg ebben a sorrendben, és minden egyes u -ra számítsuk ki a $\Phi_{\mathcal{D}}(\omega)$ konfigurációs energia potenciális változását, amit u törlése okozna ω -ból:

$$\Delta\Phi_{\omega}(u) = \Phi_{\mathcal{D}}(\omega/\{u\}) - \Phi_{\mathcal{D}}(\omega)$$

Az energiaváltozás alapján kiszámítunk egy objektum-törlési valószínűséget:

$$d_{\omega}(u) = \Gamma(\Delta\Phi_{\omega}(u)) = \frac{\delta \exp(-\beta \cdot \Delta\Phi_{\omega}(u))}{1 + \delta \exp(-\beta \cdot \Delta\Phi_{\omega}(u))},$$

Ezután $d_{\omega}(u)$ valószínűséggel töröljük az u objektumot. Amennyiben u törlésével az őt eredetileg tartalmazó forgalmi szegmens is üressé válik, azt is eltávolítjuk az ω globális konfigurációból.

- **Járműcsoport újrendezése:** véletlenszerűen javasolunk járműcsoport összevonásokat, kettéválasztást és újracsoportosításokat. Minden egyes javasolt potenciális M művelethez kiszámoljuk a hozzájuk tartozó lehetséges $\Delta\Phi_{\omega}^M$ energiaváltozást (hasonlóan ahhoz, ahogy az előző lépésben a *törlés* művelethez tettük), ezután a javasolt műveletet $\Gamma(\Delta\Phi_{\omega}^M)$ valószínűséggel hajtjuk végre.

Konvergencia teszt: amíg a folyamat nem konvergál növeljük a β inverz hőmérsékletet, csökkentjük a diszkrétizációs lépcsőt δ geometriai séma szerint, és visszalépünk a születés lépésre. A konvergencia tényét akkor állapítjuk meg, ha az egymást követő ciklusok végén már nem változik a konfiguráció.

5. Kiértékelés

Módszerünk kiértékelését négy, Budapest különböző területei felett készített légi LiDAR adathalmazon végeztük el, amit az Infoterra Magyarország Kft.© bocsátott ren-

2. táblázat: A DEM-PCA (DP)[5], h-maXima (hX) [9] és a javasolt kétszintű L²MPP (²M) módszerek objektum és pixel szintű F-mérték értékei (%-ban), illetve az L²MPP modell helyes csoport-klasszifikációs arány (CSK) F-mérték értékei (%-ban).

Adathalmaz	JSZ*	Objektum szintű %			Pixel szintű %			Járműcsoport (CSK)
		DP	hX	² M	DP	hX	² M	² M
#1	78	78	68	96	64	46	89	94
#2	91	90	93	98	77	77	88	93
#3	132	70	74	83	61	46	66	86
#4	170	85	87	89	77	76	64	92
Teljes	471	83	82	91	70	61	80	91

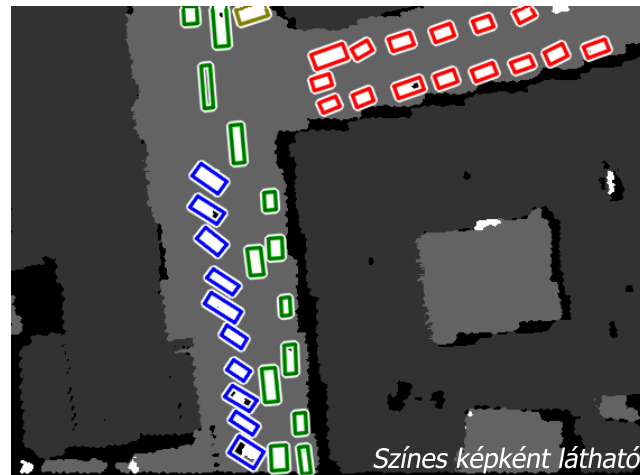
*JSZ = járművek valódi száma az adathalmazban

delkezésünkre. A tesztalmozok összesen 471 járművet tartalmaznak, különböző forgalmi helyzetekben: egy- vagy többsávos úton haladnak, forgalmi lámpánál állnak, parkolnak az úttesttel párhuzamosan, merőlegesen vagy ferden, stb. Ahhoz, hogy számszerűen tudjuk jellemezni a detekció pontosságát szükség, volt egy referencia adatbázis (*Ground Truth - GT*) létrehozására. Ezért elkészítettünk egy grafikus felhasználói felülettel rendelkező programot, ami lehetővé teszi, hogy az operátorok a címkevetítéssel nyert képen manuálisan vegyenek fel, pozícionáljanak és különböző csoportokba soroljanak jármű téglalapokat, elkészítve teljes GT forgalmi konfigurációkat.

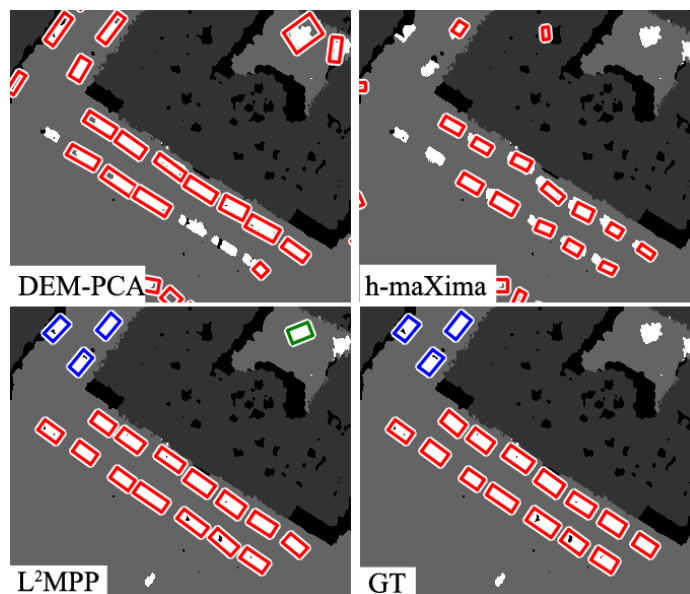
Módszerünket és a szakirodalmi referenciamódszereket az így kapott *GT* eredménnyel hasonlítottuk össze. A kvantitatív kiértékelés során figyelembe vettünk objektum- és pixelszintű metrikákat is. Objektum szinten meghatároztuk a helyesen detektált, a hamis pozitív illetve hamis negatív jármű találatok számát, ebből kiszámítottuk a detekció pontosságát (precision) és fedési arányát (recall rate), végül a két utóbbi arányérték harmonikus közepeként definiált F-mértékét. Pixel szinten a járművek sziluett maszkjait hasonlítottuk össze a referencia eredmény (*GT*) maszkjaival, és ugyancsak kiszámoltuk a detekció pixel szintű F-mértékét [13]. Meghatároztuk továbbá a helyes csoport-klasszifikáció arányát is, ami a módszer által szolgáltatott forgalmi szegmensbe sorolás és az operátor által megadott csoportosítás közötti hasonlóságot adja meg.

Az általunk javasolt L²MPP módszert a kiértékelés során összehasonlítottuk egy cella alapú algoritmussal [5] (DEM-PCA), és egy közelmúltban publikált modellel [9], amely h-Maxima (hMax) transzformációt és *watershed* szegmentációt használ a járművek kinyeréséhez. Megjegyezzük, hogy ezeknek a módszereknek, hasonlóan a többi elérhető szakirodalmi referenciamegoldáshoz, nem célja a járművek csoportokba osztása, csak az egyes objektumok azonosítása és körvonaluk meghatározása.

A 4-5 ábrákon kvalitatív eredmények láthatók a módszerünk kimenetéről minta pontfelhőrszleteken, a kvantitatív kiértékelés eredményeit a 2. táblázat foglalja össze. Megfigyelhetjük, hogy az általunk javasolt kétszintű jelölt pontfolyamat (L²MPP) modellel objektum szinten és pixel szinten is felülmúlja az összehasonlításban használt módszereket.



4. ábra: Az L^2MPP módszerünk kimeneti eredménye egy pontfelhő-részleten. Az eljárás négy járműcsoportot azonosított, amiket különböző színnel jelenítettünk meg. Háttérben a címketérkép interpolált változata látható a szemléletesebb demonstráció céljából, a következő árnyalat kódokat használva, fehér: jármű osztály, világos szürke: úttest, sötét szürke: tető, fekete: definiálatlan osztály.



5. ábra: A javasolt L^2MPP módszerünk eredményének összehasonlítása a DEM-PCA, és h-maXima referenciamódszerek kimeneteivel, és a manuálisan rögzített GT-val

A módszerünk futási idő igénye is versenyképes a legtöbb szakirodalmi referenciamódszerrel, egy 200000 pontot tartalmazó pontfelhőrésztlet feldolgozása körülbelül 15 másodpercet vesz igénybe.

6. Összefoglalás és továbbfejlesztési lehetőségek

Cikkünkben egy új, kétszintű jelölt pontfolyamat (L^2MPP) modellt mutattunk be járművek és járműcsoportok együttes kinyerésére légi LiDAR pontfelhőkből. A módszer valós LiDAR méréseken teszteltük, és annak előnyeit két referencia módszerrel összehasonlításban is bemutattuk. Bár a jelenlegi implementációinkban a járművek csoportosítása elsősorban a hasonló orientációjú objektumok összefüggő csoportokba sorolásán keresztül valósul meg, azt tapasztaltuk, hogy a módszer kisebb görbülettel rendelkező ívelt utakon is működőképes. Jövőbeni célunk a modellben lévő prior energiatagok körének kibővítése, annak érdekében, hogy a módszer komplexebb járműcsoportosítási formákat is tudjon kezelni, például meredeken kanyarodó úttesten vagy körforgalmakban haladó összefüggő szegmenseket is megtaláljunk.

7. Köszönetnyilvánítás

A szerzők köszönetet mondanak az Infoterra Magyarország Kft[©]-nek, hogy biztosították a Budapest belvárosáról készített LiDAR mérési adatokat. A munkát részben az OTKA #101598 “Távérzékelte adatok átfogó elemzése” című projekt finanszírozta. A második szerző munkáját a Bolyai János Kutatási ösztöndíj is támogatta.

Irodalom

1. Börcs, A., Benedek, C.: A marked point process model for vehicle detection in aerial lidar point clouds. In: ISPRS Congress, Melbourne, Australia (2012)
2. Börcs, A., Benedek, C.: Urban traffic monitoring from aerial LIDAR data with a two-level marked point process model. In: International Conference on Pattern Recognition (ICPR), Tsukuba City, Japan (2012)
3. Tuermer, S., Leitloff, J., Reinartz, P., Stilla, U.: Automatic vehicle detection in aerial image sequences of urban areas using 3D HoG features. In: ISPRS Photogrammetric Computer Vision and Image Analysis, Paris, France (2010) B:50
4. Yao, W., Stilla, U.: Comparison of two methods for vehicle extraction from airborne lidar data toward motion analysis. *IEEE Geoscience and Remote Sensing Letters* **8** (2011) 607–611
5. Rakusz, Á., Lovas, T., Barsi, Á.: Lidar-based vehicle segmentation. *International Archives of Photogrammetry and Remote Sensing* **XXXV** (2004) 156–159
6. Yang, B., Sharma, P., Nevatia, R.: Vehicle detection from low quality aerial LIDAR data. In: *IEEE Workshop on Applications of Computer Vision (WACV)*. (2011) 541–548
7. Yao, W., Hinz, S., Stilla, U.: Extraction and motion estimation of vehicles in single-pass airborne lidar data towards urban traffic analysis. *ISPRS Journal of Photogrammetry and Remote Sensing* **66** (2011) 260–271
8. Lafarge, F., Gimel'farb, G., Descombes, X.: Geometric feature extraction by a multi-marked point process. *IEEE Trans. Pattern Anal. Mach. Intell.* **32** (2010) 1597–1609

9. Yao, W., Hinz, S., Stilla, U.: Automatic vehicle extraction from airborne lidar data of urban areas aided by geodesic morphology. *Pattern Recogn. Letters* **31** (2010) 1100 – 1108
10. Lafarge, F., Mallet, C.: Creating large-scale city models from 3D-point clouds: A robust approach with hybrid representation. *International Journal of Computer Vision* (2012) in press.
11. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26** (2004) 1124–1137
12. Utasi, Á., Benedek, C.: Személyek lokalizálása és magasságuk becslése többszörös születés és halál dinamikával többkamerás környezetben. In: *Képfeldolgozók és Alakfelismerők Társaságának Konferenciája, Szeged, Magyarország (2009)* 364–375
13. Benedek, C., Descombes, X., Zerubia, J.: Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics. *IEEE Trans. Pattern Anal. Mach. Intell.* **34** (2012) 33–50
14. Descombes, X., Minlos, R., Zhizhina, E.: Object extraction using a stochastic birth-and-death dynamics in continuum. *J. Mathematical Imaging and Vision* **33** (2009) 347–359