

Új környezetelemzési módszerek Lidar-adatok alapján

Ph.D. disszertáció tézisei

H. Zováthi Örkény Ádám

Témavezető:
Dr. Benedek Csaba, az MTA doktora



Roska Tamás Műszaki és Természettudományi Doktori Iskola
Pázmány Péter Katolikus Egyetem

Budapest, 2023

1. Bevezetés

A háromdimenziós (3D) adatgyűjtés új érzékelőinek felhasználása kulcsfontosságú lépés a környezetünkben található térbeli információ hatékony értelmezéséhez. Napjainkban széles körben elterjedtek a Lidar (Light Detection and Ranging) technológián alapuló szenzorplatformok, melyek nagy adatgyűjtési sebességgel pontos 3D méréseket biztosítanak [6]. Az elérhető lézerszkennerek időbeli és térbeli felbontása között egyelőre kompromisszumot kell kötni, ami miatt az előállított pontfelhők minősége és sűrűsége jelentősen különbözik [7], korlátozva a pontfelhőkben tárolt információ automatikus értelmezését végző általános módszerek használhatóságát és az egyes szenzorokhoz kifejlesztett, egyedi funkcionális követelményekkel rendelkező algoritmusok általánosíthatóságát.

Egyrészt az autonóm járművek és a mobil eseménymegfigyelési alkalmazások (például forgalomelemzés vagy tömegmegfigyelés) valós idejű, mobil platformon futó 3D adatgyűjtési és feldolgozási technikákat igényelnek. Dinamikus környezetérzékelési és -felismerési feladatokhoz, mint például jelenetelemzés és -értelmezés, periodikusan ismétlődő, jellemzően forgó többsugaras (RMB) Lidar-szenzorokat (pl. Ouster OS1 vagy Velodyne Puck modellek) [1] alkalmaznak. Az RMB Lidarok valós idejű 3D pontfelhőfolyamokat képesek előállítani (300 ezer-2 millió pont/mp), azonban a mérések térbeli sűrűsége alacsony, valamint a szenzorlátómező (FoV) kitöltöttsége a teljes szkennelési folyamat során állandó: A függőleges felbontást a lézersugarak száma határozza meg (16-128), míg a vízszintes felbontás a szenzor forgási frekvenciájának (5-20 Hz) függvényében változik.

Az RMB Lidarok alternatívájaként az újonnan elterjedt, lényegesen alacsonyabb költségű, nem periodikus szkennelésű (NRCS) Lidar-szenzorok is alkalmazhatóak valós idejű környezetelemzési feladatokhoz. Az NRCS Lidarok (pl. a Livox AVIA szenzor) – ellentétben az RMB Lidarokkal – képesek nagy területek sűrű feltérképezésére egy adott szkennelési pozícióból is, köszönhetően annak, hogy speciális szkennelési technológiájuk nem periodikus, pl. rozettamintát követ. Esetükben a fő kihívás a rögzített 3D adatok térbeli és időbeli felbontása közötti hatékony egyensúly megteremtése egy megfelelő

integrációs ablak használatával [2, 8].

Másrészt városüzemeltetési alkalmazásoknak, például városfejlesztés és -tervezés, közterület-felügyelet vagy útkarbantartás, részletes és pontos 3D várostérképekre van szükségük, amelyeket offline szkennelési technológiákkal készítenek el, majd új generációs térinformatikai rendszerekben (GIS) tárolnak és tartanak karban. Ezekhez az alkalmazásokhoz gyakran használnak korszerű földi (TLS) és mobil lézerszkennelő (MLS) platformokat, mert – időben szinkronizált Lidar-érzékelőkkel, kamerákkal és navigációs egységekkel felszerelve – sűrű, pontos, részletgazdag és georeferált pontfelhőket készítenek a környezetükről.

Disszertációm a Lidar-szenzorok környezetelemzésével kapcsolatos két fő feladattal foglalkozik. Az első részben megvizsgálom, hogy egy korszerű, valós idejű RMB Lidar szenzor érzékelése hogyan javítható egy MLS rendszer által offline rögzített, szemantikailag kiértékelt 3D pontfelhőtérképekből kinyert előzetes helyszíni információ alapján. A második részben azt vizsgálom, hogy a korszerű valós idejű NRCS Lidarok érzékelése hogyan javítható külső információ felhasználása nélkül, mélytanulási módszerekkel, a szenzormérések térbeli és időbeli karakterisztikái alapján.

2. Új tudományos eredmények

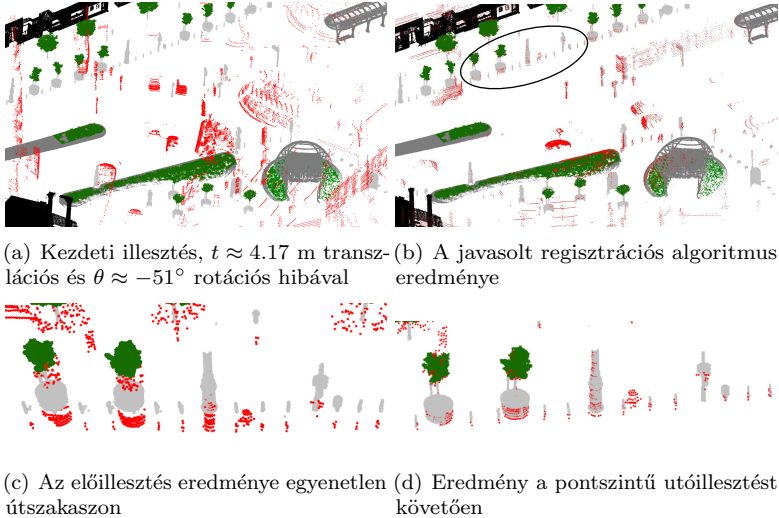
1. tézis: Javasoltam egy új változásdetekciós eljárást, amely képes 3D pontfelhők regisztrációjára és pontszintű változások szegmentálására jelentősen eltérő sűrűségkarakterisztikájú Lidar-pontfelhők között. Létrehoztam egy új városi adathalmazt egy korszerű forgó többsugaras (RMB) Lidar-szenzorral (tipikusan 50-500 pont/m²) és egy modern mobil lézerszkennelő (MLS) rendszerrel (több, mint 5000 pont/m²) felvett mérések alapján. Az új adathalmaz segítségével kvantitatív módon demonstráltam a javasolt algoritmus előnyeit a legkorszerűbb referenciamódszerekkel szemben.

A szerző kapcsolódó publikációi: [1][3][4][5]

A tézis három egymásra épülő részfeladattal foglalkozik: Először pontfelhőregisztrációt hajtok végre a ritkás RMB- és sűrű MLS-adatok pontos összeillesztéséhez, és bemutatok ennek az algoritmusnak egy továbbfejlesztett változatát, amely képes a mozgó jármű pózának valós idejű követésére még jelentősen kitakart környezetben is. Másodszer, egy változásdetektációs algoritmust javaslok, amely képes a jármű Lidar-mérésein pontszintű változások szegmentálására. Harmadszor bemutatok egy olyan módszert, amely a változásdetektáció kimenetét használja fel dinamikus objektumok detekciójának javítására.

1.1. Javasoltam egy új megközelítést kezdetben egymáshoz pontatlanul illeszkedő ritka RMB Lidar- és sűrű MLS-pontfelhők regisztrálására. A módszer először statikus objektumok geometriai jellemzőinek párosításával egy durva transzformációt végez, majd az illesztést a párosított objektumok lokális pontfelhőszegmensei alapján finomítja a számítási igény csökkentése érdekében. A javasolt algoritmus előnyeit különböző korszerű referenciamódszerekkel szemben városi helyszíneken demonstráltam a transzformációs mátrixok és a Ground Truth (GT) dekompozíciójából számolt translációs és rotációs hibák összehasonlításával, valamint a regisztrált pontfelhők közötti ponttávolságok vizsgálatával. Bemutattam az algoritmus egy hatékony alkalmazását pózkövetésre, felhasználva a bemutatott objektumpárosítási eljárás eredményét síkbeli pózbecsléshez, majd integrálva a becsült pózokat egy konstans sebességű járműdinamikai modellt feltételező Kálmán-szűrőbe.

A szakirodalomban számos pontfelhőregisztrációs algoritmus létezik, amelyek manuális [9, 10, 11] vagy tanult [12, 13] kulcspontok, szegmensek [14, 15] vagy pontok [16, 17, 18] jellemzői között keresnek megfeleltetéseket. A megcélzott RMB-MLS szcenárióban a feladat fő kihívását az adja, hogy az RMB Lidar pontfelhői túl ritkák geometriailag értelmezhető 3D kulcspontok kinyeréséhez (amelyek egyébként sűrű MLS-pontfelhők között jól definiálhatók), míg az MLS-pontfelhők lokálisan gyakran akár 100-1000-szer sűrűbbek, mint a nekik megfelelő RMB-régiók, ami félrevezeti a szegmensalapú



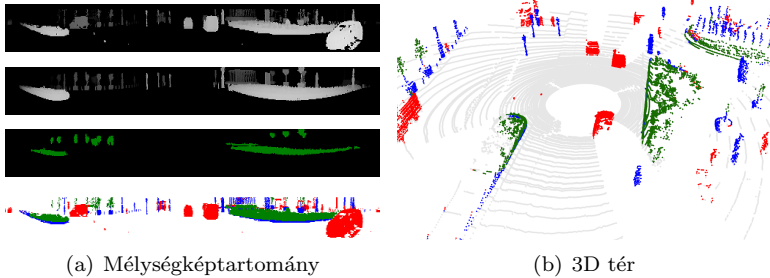
1. ábra. A javasolt pontfelhőregisztrációs algoritmus eredményei. (c) és (d) alábbiak a (b) alábbián feketével jelölt területet ábrázolják. Színkódok: Az RMB Lidar pontjai pirossal, az MLS-pontfelhő pedig a szemantikus osztályoknak megfelelő színekkel van jelölve.

illesztési módszereket. Egy eltérő megközelítést követve, első kontribúcióként az eredeti pontfelhők illesztése helyett az RMB Lidar és az MLS-térkép pontfelhőiből statikus objektumokat detektáltam és párosítottam egymással. Ebben a lépésben akár sok téves statikus objektumjelölt is megjelenhet (pl. forgalmi résztvevők, részben kitért objektumok), ami nagyszámú hibás párosítást eredményez. Ennek kezelésére az általánosított Hough-transzformáció szavazási sémáját alkalmaztam. Második kontribúcióként az előillesztési lépést a pontszintű Iterative Closest Point algoritmus [19] inicializálására használtam fel, amelyet csak a korábban összeillesztett objektumpárokhoz tartozó pontfelhőszegmensekre futtattam le. A javasolt megközelítéssel hat különböző pontfelhőregisztrációs módszerhez [13, 20, 16, 21, 22, 23] képest a ponttávolságok mediánértéke 1-2 nagyságrenddel csökkent.

Annak érdekében, hogy a módszer robusztusan működjön akkor is, ha nincs megfelelő számú párosítható objektum, az objektumillesztési algoritmus eredményéből kiszámoltam a mozgó jármű síkbeli (3DoF) pózát (síkbeli pozíció és függőleges tengely menti orientáció), és a becsült pózparamétereket egy konstans sebességű járműdinamikai modellet feltételező Kálmán-szűrőbe integráltam. A javasolt pózkövető algoritmus pontatlan, 5-10 méteres hibájú GPS-alapú helymeghatározásból kiindulva képes a jármű pozicionálási hibáját egy nagyságrenddel csökkenteni, és a függőleges tengely menti orientációs hibát 1° körül tartani a teljes trajektória alatt, jelentős drift nélkül, valós időben (20-25 Hz).

1.2. Javasoltam egy új, Markov véletlen mező alapú megközelítést (RangeMRF), amely képes regisztrált RMB- és MLS-pontfelhők közötti változások kinyerésére és többosztályos klasszifikációjára (dinamikus, szezonális vagy nincs változás) 2D mélységkép-reprezentációk felhasználásával. A javasolt algoritmus előnyeit különböző korszerű referenciamódszerekkel szemben kvalitatív és kvantitatív kiértékelésekkel igazoltam.

A háromdimenziós változásdetektáció a szakirodalomban aktívan kutatott téma [24], azonban a meglévő pont- [25, 26], szegmens- [27] vagy voxel- [28] alapú módszerek nem kezelik hatékonyan, ha a két összehasonlítandó pontfelhő karakterisztikájában jelentősen különbözik, és a megcélzott RMB-MLS szcenárióban vagy számos hamis pozitív (pl. zaj, növényzeti változások), vagy sok hamis negatív (pl. részletek elvesztése) találatot eredményeznek. Első kontribúcióként a javasolt RangeMRF módszer a pontfelhőkből generált 2D mélységképek között keres változásokat. A javasolt módszer a kompakt mélységkép-reprezentációnak köszönhetően kifejezetten gyors, eközben robusztusan kezeli a két pontfelhő közötti jelentősen eltérő sűrűségkarakterisztikáját is azáltal, hogy a sűrű MLS-adatoknak csak a releváns részeit tartalmazza. Másodszor, egy Markov véletlen mező alapú modellt alkalmaztam, amely robusztus a zajos mérési adatokra. Harmadszor, a képszegmentáló modellben három osztályt különböztettem meg: növényzeti régiók szezonális változásait, mozgó forgalmi résztvevők vagy a megváltozott/újrapozicionált statikus



2. ábra. A változásszegmentáció eredménye valós mérési adatokon a mélységképtartományban és a 3D térben. Az (a) alábra első két sora mélységképeket tartalmaz, ahol a világosabb pixelek kisebb távolságot jelölnek, a fekete pixelek pedig nem tartalmaznak méréseket. A harmadik sor zöld színnel jelöli azokat a pixeleket, amelyek az MLS-adat növényzetosztályához tartoznak. Az utolsó sor az RMB-adatok szegmentálásának kimenetét jelöli. A (b) alábra ugyanezt az eredményt jeleníti meg a 3D térben. A statikus háttér pixeljei/pontjai kékkel, a dinamikus változások pirossal, a szezonális változások pedig zölddel vannak jelölve.

utcai elemek által okozott dinamikus változásokat és változatlan, statikus területeket. A növényzeti régiók speciális érzékenységi figyelembevételével a javasolt módszer számos téves találatot képes kiszűrni növényzetet tartalmazó területeken, miközben a bemeneti pontfelhők között még kisebb dinamikus változások (pl. állomások vagy falszegmensek közelében álló gyalogosok) felismerésére is képes marad. Összehasonlítva négy referenciatechnikával [27, 24, 28, 26], a javasolt módszer vagy F1-metrika szerint (körülbelül 10-25%-kal), vagy a számítási komplexitásában (10-1000-szer gyorsabb) bizonyult jobbnak.

1.3. Javasoltam egy új módszert a bemutatott változásdetekciós megközelítés felhasználására a kizárólag Lidar-alapú dinamikus objektumdetekciós algoritmusok teljesítményének javítása érdekében. A legkorszerűbb objektumdetektálási módszert nagy

forgalmú városi útszakaszokon vizsgálva igazoltam, hogy a javasolt megközelítés képes hatékony egyensúlyt létrehozni a pontosság (precision) és felidézés (recall) értékei között, valamint egy jelentős általános javulást elérni mind járművek, mind gyalogosok esetében.

A valós idejű dinamikus objektumdetekció ritka 3D pontfelhőkben az autonóm vezetés egyik legaktívabban kutatott területe, amire a szakirodalomban számos geometriai [29] és mélytanuláson alapuló [30, 31, 32, 33, 34] algoritmus létezik. Ezeknek a megközelítéseknek azonban több limitációjuk is van: Egyrészt hamis pozitív találatok jelennek meg olyan statikus pontfelhőterületeken, amelyek a dinamikus objektumokhoz hasonló megjelenéssel és kontextusparaméterekkel rendelkeznek, másrészt hamis negatív találatokat okozhat, ha a dinamikus objektumok pontfelhőszegmensei összeolvadnak statikus utcai elemekkel, vagy részlegesen kitakarják őket más objektumok. Ezért egy új megközelítést javasoltam, amely a sűrű MLS-térképeket használja arra, hogy párhuzamosan csökkentse az objektumdetektáló algoritmusok hamis negatív és hamis pozitív találatait is. A javasolt megközelítés tartalmaz egy térkép alapú objektumvalidációt, a korábban bemutatott változásdetekciós eljárást és egy objektum-szintű változáselemzési lépést a regisztrált RMB- és MLS-pontfelhők között. Az összehasonlítások alapjául a *PointPillars* [34] korszerű objektumdetekciós módszert választottam, amelyhez képest a javasolt módszerrel sikerült pontosságban (precision) 5,96%-os, felidézésben (recall) 9,21%-os, F1-metrika szerint pedig 7,93%-os javulást elérni a saját adathalmazunkon.

2. tézis: Javasoltam egy új mélységképiegészítő eljárást egy nem periodikus szkennelésű (NRCS) Lidar ritkás egymás utáni mérései alapján egy mélytanulás alapú modell (ST-DepthNet) felhasználásával. Létrehoztam egy új városi adathalmazt, amely különböző szimulált és valós NRCS Lidar-adatokat tartalmaz. Az új adathalmaz segítségével kvalitatív és kvantitatív módon igazoltam a javasolt módszer előnyeit egy, a nyers szenzormérésekből készített sűrített mélységtérképhez, valamint két egymástól független, kizáró-

lag Lidar-alapú korszerű mélységképkiegészítő algoritmus-hoz képest.

A szerző kapcsolódó publikációja: [2]

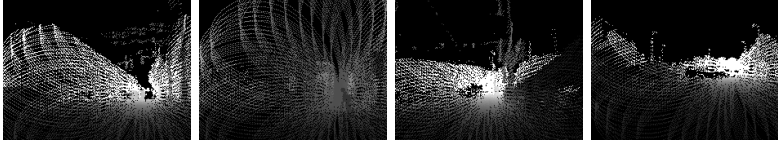
A tézis az NRCS Lidar-adatok hatékony adatvezérelt kiegészítésével foglalkozik. A nem periodikus szkennelési technológiájuknak köszönhetően az NRCS Lidarok képesek a látómező különböző területeit egy adott szkennelési pozícióból egymást követő időpillanatokban feltérképezni [8]. Az így keletkezett pontfelhő-szekvenciák elemzésének fő kihívása az adatok térbeli és időbeli felbontása közötti hatékony egyensúly megteremtése egy megfelelő integrációs ablak segítségével. Nagyobb integrációs idő (pl. 1 mp) esetén nagy a mérések térbeli felbontása, azonban számos kihívás megjelenik, amelyek megnehezítik a dinamikus események elemzését, például elmosódnak a megfigyelt járművek, gyalogosok vagy épületek. Kis időablak (pl. 0,2 mp) esetén megmarad a térbeli pontosság, viszont a mérések térbeli felbontása alacsony marad, ami a térbeli részletek jelentős elvesztését eredményezi. Az NRCS Lidar alapú érzékelés ezen tér-időbeli kompromisszumának feloldására egy új, mélytanuláson alapuló megközelítést javasoltam, amely képes a ritka NRCS Lidar-mérések sűrítésére, miközben megtartja a nagy időbeli felbontást és a térbeli pontosságot.

2.1. Javasoltam egy új tanítási keretrendszert az NRCS Lidarok 3D adatainak sűrítésére azáltal, hogy az egymást követő pontméréseket ritkás, 2 tizedmásodperc alatt integrált, 40% látómező-kitöltöttségű 2D mélységképekké alakítom. Létrehoztam egy új szintetikus adathalmazt, amely az NRCS Lidar működésének szimulálásával készített ritkás mélységképeket, valamint minden egyes ritka mintához sűrű mélységképeket is tartalmaz, kihasználva a virtuális világ teljes térbeli információját. Az adatokat kibővítettem ugyanezt a mélységkép-reprezentációt használó ritkás valós mintákkal is, és mindkét adathalmazt nyilvánosan elérhetővé tettem. A javasolt keretrendszer lehetővé teszi a mélységképkiegészítő algoritmusok tanítását és tesztelését szintetikus adatokon, valamint megbízhatóságuk validálását valós szituációkban is.

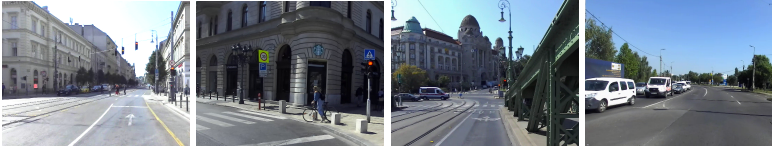
A 2D reprezentációhoz a Livox AVIA NRCS Lidar-szenzor által rögzített ritka pontfelhőket a Descartes-féle (x, y, z) koordinátarendszerből a gömbi (r, θ, ϕ) polárkoordináta-rendszerbe konvertáltam, majd 2 tizedmásodperces integrációs időablakot használva a vízszintes és függőleges látómezőket egy 400×400 pixelrácsra kvantáltam az egymást követő mérések összegyűjtéséhez. Ezzel a megközelítéssel a mélységkép pixeljeinek körülbelül 60%-a ugyan nem tartalmaz távolságértéket, viszont az elmosódás nem jelentős, emellett a képi reprezentáció lehetővé teszi kétdimenziós konvolúciós neurális hálózatok használatát a hiányzó mélységinformáció kitöltésére. Mivel a nagyobb integrációs idő elmosódott sziluettek eredményezne a környező dinamikus objektumok, valamint a platform független mozgása miatt, valós adatok esetén kihívást jelent a sűrű és nagy térbeli pontosságú GT mélységinformáció előállítás. Ezért a valós mérések mellett egy realiztikus virtuális világból készítettem egy szintetikus mélységkép-adathalmazt is a CARLA szimulátor [35] segítségével, ahol a Livox AVIA NRCS Lidar-szenzor viselkedését implementáltam. A virtuális világ lehetővé teszi a sűrű, nagy térbeli pontosságú mélységinformációk kinyerését, amelyek GT-ként használhatóak a Lidar ritkás méréseihez.

2.2. Javasoltam egy új, a klasszikus U-Net architektúrát egy tér-időbeli leskalázó ággal kibővítő mélységképkiegészítő mély neurális hálózatot (ST-DepthNet), amely képes nagy sűrűségű és pontos mélységképadatok valós idejű előállítására. A javasolt algoritmus előnyeit a legújabb referenciatechnikákkal szemben szintetikus és valós szituációkban is igazoltam.

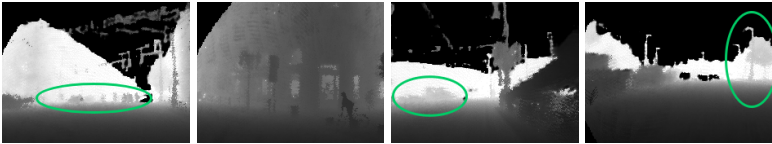
Első kontribúcióként azt használtam ki, hogy az alkalmazott Livox AVIA szenzor használatával egy 1 másodperces időintervallum elegendő, közel teljes lefedettségű mélységinformációt tartalmaz egy adott helyszínről. Ezért öt egymást követő ritka mélységképet – amelyek mindegyike 2 tizedmásodperces integrálási idővel készült – adtam a javasolt hálózat bemenetének, amely így elegendő információt tartalmaz a teljes látómezőről. Az egyetlen kimeneti mélységkép pontos előállításához egy kép-kép U-Net [36] architektúrát alkalmaztam, és a bemeneti képsorozat jellemzői közötti időbeli kapcsolatok



(a) Ritkás bemeneti adat 2 tizedmásodperces időablakkal integrálva



(b) RGB-referenciakép, kizárólag vizuális verifikációhoz



(c) A javasolt ST-DepthNet mély neurális hálózat kimenete

3. ábra. Eredmények valós mérési adatokon Budapesten. A javasolt módszer által pontosan és részletesen kiegészített objektumstruktúrák zöld ellipszissel vannak jelölve.

figyelembevételéhez a U-Net hálózat leskálázó részét kiterjesztettem Conv2DLSTM-rétegek [37] felhasználásával. Másodszer, a javasolt hálózat felskálázó ágát tisztán kétdimenziósként megtartva az egyes szinteken a közvetlen (skip) kapcsolatokat rekurrens összevonással végeztem, felhasználva a Conv2DLSTM-rétegek utolsó kimenetét, amely az utolsó 2 tizedmásodperces mérés jellemzőit tartalmazza. Harmadszor, az utolsó bemeneti képet közvetlenül is összekapcsoltam a hálózat kimenetével, kihasználva, hogy az utolsó bemeneti kép a térben aktuálisan pontos méréseket tartalmazza, és ezáltal a hálózatnak csak a mélységkép hiányzó régióit kell megtanulnia. A modellt az ismertetett szintetikus adathalmazon tanítottam, és kvantitatív módon igazoltam, hogy a javasolt megközelítés a két legkorszerűbb referenciamódszerrel [38, 39] szemben szintetikus ada-

tokon a 2D mélységképtartományban számolt átlagos négyzetes hiba négyzetgyöke (RMSE) alapján több mint 1 méterrel, a 3D tartományban pedig a normalizált Chamfer és medián távolságok szerint körülbelül fél méterrel pontosabb. A módszerek hatékonyságát valós NRCS mérési adatokon egy kérdőív segítségével hasonlítottam össze, melyben 20 gépilátás-szakértőt kértem fel a generált mélységképek osztályozására, akik a javasolt módszer kimenetét értékelték a legjobbnak.

3. Az eredmények alkalmazási területei

A kifejlesztett algoritmusok mindegyike mobil robotok és intelligens járművek RMB vagy NRCS Lidar-érzékelőkkel felszerelt modern érzékelési platformjain használható. Az első tézis városi környezetben alkalmazható RMB Lidarral és GPS-vevővel felszerelt járművek támogatására olyan helyszíneken, ahol elérhetőek részletes 3D pontfelhőtérképek. A javasolt módszerek térkép alapú, valós idejű környezetértelmezési feladatokhoz járulhatnak hozzá, például pontos lokalizáció és pózkövetés, változásalapú környezetsegmentálás és hatékony dinamikus objektumfelismerés. A második tézis NRCS Lidar-érzékelővel felszerelt robot- vagy járműplatformok számára alkalmazható, hogy pontos és sűrű mélységtérképeket készítsenek a környezetükről nagy időbeli felbontással, ami egy fontos köztes lépés lehet magasabb szintű környezetértelmezési vagy térképezési feladatokhoz. A Kooperatív Doktori Program részeként a javasolt algoritmusok nagy része közvetlenül is hozzájárult a SZTAKI és a PPKE részvételével zajló K+F projektekhez, emellett az eljárásokkal kapcsolatban két szabadalmi bejelentést is tettünk.

4. Adatok és implementációs részletek

Az RMB Lidar-mérések és az MLS-adatok közös felhasználására létrehoztam egy új adatbázist, amely egy Velodyne HDL 64E 64 sugaras RMB Lidar-szenzorral rögzített pontfelhőszekvenciákat tartalmaz Budapest különböző, olyan belvárosi területeiről, ahol egy

Riegl VMX450 MLS szkennertől rögzített, nagy sűrűségű, georeferált pontfelhőtérképek szintén rendelkezésre állnak. Az adatbázis három különböző mérési szekvenciát tartalmaz, amelyek mindegyike hozzávetőlegesen 300 méteres útszakaszokat fed le, szegmentált referencia MLS-adatokkal [40]. Releváns kvantitatív kiértékeléshez ellenőrzött Ground Truth (GT) adatok is rendelkezésre állnak, mind a regisztrációhoz, mind a pózkövetéshez, mind a változás- és objektumdetekcióhoz.

A második részben javasolt módszer tanításához és teszteléséhez egy realiztikus virtuális világból létrehoztam egy szintetikus mélységkép-adatbázist (*LivoxCARLA*) a CARLA szimulátor [35] segítségével, amelynek során a Livox AVIA NRCS Lidar-szenzor érzékelését szimuláltam. A *LivoxCARLA* adathalmaz 11726 véletlenszerűen mintavételezett bemeneti-kimeneti mélységképpárból áll, amelyek közül 10000 a tanításhoz, 500 a validáláshoz és 1226 a teszteléshez volt használva. A *LivoxCARLA* adathalmazon kívül valós mérési szekvenciákat is gyűjtöttem Budapestről. Ezekben a kísérletekben a Livox AVIA érzékelőt a tesztjármű tetejének elejére szerelve egy összesen 5,5 kilométeres útvonal lett bejárva, részben belvárosi szakaszokon, részben pedig gyorsforgalmi utakon.

A pontfelhők feldolgozásának fő platformját C++ nyelven, az OpenCV és PCL könyvtárakkal implementáltam, míg a neurális hálózatokat Python nyelven, a Pytorch vagy Keras keretrendszerekkel implementáltam és tanítottam.

5. Köszönetnyilvánítás

Szeretném kifejezni köszönetemet témavezetőmnek, Prof. Benedek Csabának a doktori tanulmányaim és munkám során nyújtott folyamatos támogatásáért, motivációjáért és türelméért. Külön köszönöm Prof. Szirányi Tamásnak, a Gépi Érzékelés Kutatólaboratórium vezetőjének, hogy támogatta kutatásomat a Számítástechnikai és Automatizálási Kutatóintézetben (SZTAKI). Köszönöm kollégáimnak munkájukat és tanácsaikat: Benedek Csabának, a kutatócsoport vezetőjének, Jankó Zsoltnek, KDP-kutatásom vállalati szakértőjének, valamint Nagy Balázsnak, Kovács Lórántnak, Pálffy Balázsnak,

Kövendi Józsefnek, Yahya Ibrahimnak és Tizedes Lászlónak. Köszönettel tartozom a Pázmány Péter Katolikus Egyetemnek (PPKE) is, elsősorban Prof. Szederkényi Gábornak a tanulmányi lehetőségért. Köszönöm a disszertáció bírálóinak munkájukat és észrevételeiket. Külön köszönet azoknak, akiket talán nem említettem név szerint, de közvetlenül vagy közvetve támogattak doktori kutatásom elvégzésében.

Köszönet a Nemzeti Kutatási, Fejlesztési és Innovációs Alapnak, az ÚNKP-20-3 és ÚNKP-21-3 Új Nemzeti Kiválósági Program, valamint a KDP-2020 Kooperatív Doktori Program (KDP-977852) keretében nyújtott anyagi támogatásért. A SZTAKI-ban végzett kutatásom az Autonóm Rendszerek Nemzeti Laboratórium és a Mesterséges Intelligencia Nemzeti Laboratórium, valamint a TKP2021-NVA-01, illetve OTKA K-120233 és K-143274 projektek keretében is támogatva volt.

Végül, de nem utolsósorban, szeretnék köszönetet mondani a családomnak, amiért végig támogattak doktori tanulmányaim során és általában az életemben. Édesanyámnak, aki olyan feltétel nélküli szeretettel nevelt fel, amit soha nem fogok tudni meghálálni, még ha ezer évig is élek, édesapámnak, aki mindig arra ösztönzött, hogy merjek nagyot álmodni és megtanított arra, hogy a kemény munka mindig kifizetődik, testvéreimnek, Bendinek, Kendének, Csaninak és Dominak, valamint gyönyörű feleségemnek, Katának, aki a legnehezebb időkben is motivált és támogatott.

Nagyszüleimnek a véget nem érő áldásaikért és keresztanyámnak a folyamatos támogatásáért.

6. Publikációk

6.1. A szerző folyóiratpublikációi

- [1] **Ö. Zováthi**, B. Nagy, and C. Benedek, „Point cloud registration and change detection in urban environment using an onboard lidar sensor and MLS reference data,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 110, p. 102767, 2022. IF = 7.5, Scimago Q1/D1.
- [2] **Ö. Zováthi**, B. Pálffy, Z. Jankó, and C. Benedek, „ST-DepthNet: A spatio-temporal deep network for depth completion using a single non-repetitive circular scanning lidar,” *IEEE Robot. and Autom. Lett.*, vol. 8, no. 6, pp. 3270–3277, 2023. IF = 5.2*, Scimago Q1/D1.

6.2. A szerző nemzetközi konferenciapublikációi

- [3] **Ö. Zováthi**, B. Nagy, and C. Benedek, „Exploitation of dense MLS city maps for 3D object detection,” in *Int. Conf. Image Anal. Recognit.*, vol. 12131 of *Lecture Notes in Computer Science*, pp. 393–403, Springer, 2020.
- [4] **Ö. Zováthi**, B. Pálffy, and C. Benedek, „Real-time vehicle localization and pose tracking in high-resolution 3D maps,” in *European Signal Process. Conf.*, pp. 1786–1790, 2022.
- [5] **Ö. Zováthi**, L. Kovács, B. Nagy, and C. Benedek, „Multi-object detection in urban scenes utilizing 3D background maps and tracking,” in *Int. Conf. Control, Artif. Intell., Robot. Optim.*, pp. 231–236, 2019.

6.3. A disszertációhoz kapcsolódó válogatott publikációk

- [6] C. Benedek, A. Majdik, B. Nagy, Z. Rózsa, and T. Szirányi, „Positioning and perception in LIDAR point clouds,” *Digit. Signal Process.*, vol. 119, p. 103193, 2021.

- [7] J. Behley, V. Steinhage, and A. B. Cremers, „Performance of histogram descriptors for the classification of 3D laser range data in urban environments,” in *IEEE Int. Conf. Robot. Autom.*, (St. Paul, MN, USA), pp. 4391–4398, May 2012.
- [8] L. Kovács, M. Kégl, and C. Benedek, „Real-time foreground segmentation for surveillance applications in NRCS lidar sequences,” *ISPRS Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLIII-B1-2022, pp. 45–51, 05 2022.
- [9] R. Hänsch, T. Weber, and O. Hellwich, „Comparison of 3D interest point detectors and descriptors for point cloud fusion,” *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. II-3, pp. 57–64, 09 2014.
- [10] I. Sipiran and B. Bustos, „Harris 3D: A robust extension of the Harris operator for interest point detection on 3D meshes,” *The Vis. Comput.*, vol. 27, pp. 963–976, 11 2011.
- [11] Y. Zhong, „Intrinsic shape signatures: A shape descriptor for 3D object recognition,” in *Int. Conf. Comput. Vis. Workshops*, (Kyoto, Japan), pp. 689–696, 11 2009.
- [12] J. Li and G. H. Lee, „USIP: Unsupervised stable interest point detection from 3D point clouds,” in *IEEE/CVF Int. Conf. Comput. Vis.*, (Seoul, Korea), pp. 361–370, 2019.
- [13] F. Lu, G. Chen, Y. Liu, Z. Qu, and A. Knoll, „RSKDD-Net: Random sample-based keypoint detector and descriptor,” in *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 21297–21308, 2020.
- [14] B. Douillard, A. Quadros, P. Morton, J. P. Underwood, M. D. Deuge, S. Hugosson, M. Hallström, and T. Bailey, „Scan segments matching for pairwise 3D alignment,” in *IEEE Int. Conf. Robot. Autom.*, (St. Paul, MN, USA), pp. 3033–3040, May 2012.
- [15] R. Dubé, A. Cramariuc, D. Dugas, J. Nieto, R. Siegwart, and C. Cadena, „SegMap: 3D segment mapping using data-driven descriptors,” in *Robot.: Sci. Syst.*, 06 2018.

- [16] Z. Zhang, „Iterative point matching for registration of free-form curves and surfaces,” *Int. J. Comput. Vis.*, vol. 13, p. 119–152, Oct. 1994.
- [17] A. Gressin, C. Mallet, and N. David, „Improving 3D LIDAR Point Cloud Registration Using Optimal Neighborhood Knowledge,” *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. I-3, pp. 111–116, 08 2012.
- [18] A. Gressin, B. Cannelle, C. Mallet, and J.-P. Papelard, „Trajectory-Based Registration of 3D LIDAR Point Clouds Acquired with a Mobile Mapping System,” *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. I-3, pp. 117–122, 08 2012.
- [19] Z. Zhang, „Iterative point matching for registration of free-form curves and surfaces,” *Int. J. Comput. Vis.*, vol. 13, no. 2, pp. 119–152, 1994.
- [20] K. S. Arun, T. S. Huang, and S. D. Blostein, „Least-squares fitting of two 3-D point sets,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 5, pp. 698–700, 1987.
- [21] S. Choi, Q. Zhou, and V. Koltun, „Robust reconstruction of indoor scenes,” in *IEEE Conf. Comput. Vis. Pattern Recognit.*, (Boston, MA, USA), pp. 5556–5565, 2015.
- [22] H. Yang, J. Shi, and L. Carlone, „Teaser: Fast and certifiable point cloud registration,” *IEEE Trans. Robot.*, vol. 37, no. 2, pp. 314–333, 2021.
- [23] X. Huang, G. Mei, and J. Zhang, „Feature-metric registration: A fast semi-supervised approach for robust point cloud registration without correspondences,” in *IEEE Conf. Comput. Vis. Pattern Recognit.*, June 2020.
- [24] W. Xiao, B. Vallet, M. Brédif, and N. Paparoditis, „Street environment change detection from mobile laser scanning point clouds,” *ISPRS J. Photogramm. Remote Sens.*, vol. 107, pp. 38 – 49, 2015.

- [25] D. Liu, D. Li, M. Wang, and Z. Wang, „3D change detection using adaptive thresholds based on local point cloud density,” *ISPRS Int. J. Geo-Inf.*, vol. 10, no. 3, 2021.
- [26] D. Girardeau-Montaut, M. Roux, R. Marc, and G. Thibault, „Change detection on point cloud data acquired with a ground laser scanner,” *ISPRS Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 36, 01 2005.
- [27] A. Schlichting and C. Brenner, „Vehicle localization by lidar point correlation improved by change detection,” *ISPRS Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLI-B1, pp. 703–710, 06 2016.
- [28] K. Liu, J. Boehm, and C. Alis, „Change detection of mobile lidar data using cloud computing,” *ISPRS Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLI-B3, pp. 309–313, 06 2016.
- [29] A. Börcs, B. Nagy, and C. Benedek, „Instant object detection in lidar point clouds,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, pp. 992–996, July 2017.
- [30] K. Shin, Y. Kwon, and M. Tomizuka, „RoarNet: A robust 3D object detection based on region approximation refinement,” *IEEE Intelligent Vehicles Symposium*, pp. 2510–2515, 2018.
- [31] M. Simon, S. Milz, K. Amende, and H.-M. Groß, „Complex-YOLO: Real-time 3D object detection on point clouds,” *ArXiv*, vol. abs/1803.06199, 2018.
- [32] Y. Yan, Y. Mao, and B. Li, „SECOND: Sparsely embedded convolutional detection,” *Sensors*, vol. 18, p. 3337, 10 2018.
- [33] B. Yang, W. Luo, and R. Urtasun, „PIXOR: Real-time 3D object detection from point clouds,” in *IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 7652–7660, 2018.
- [34] A. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom, „PointPillars: Fast Encoders for Object Detection from Point

- Clouds,” in *IEEE Conf. Comput. Vis. Pattern Recognit.*, (Long Beach, CA, USA), pp. 12689–12697, 06 2019.
- [35] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, „CARLA: An open urban driving simulator,” in *Proc. Ann. Conf. Rob. Learn.*, pp. 1–16, 2017.
- [36] O. Ronneberger, P. Fischer, and T. Brox, „U-net: Convolutional networks for biomedical image segmentation,” in *Proc. Int. Conf. Med. Image Comput. Comp.-Ass. Interv.*, pp. 234–241, 2015.
- [37] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-k. Wong, and W.-c. Woo, „Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” in *Proc. Int. Conf. NIPS*, p. 802–810, 2015.
- [38] J. Ku, A. Harakeh, and S. L. Waslander, „In defense of classical image processing: Fast depth completion on the CPU,” *Conf. Comput. Rob. Vis.*, pp. 16–22, 2018.
- [39] M. F. F. Khan, N. D. Troncoso Aldas, A. Kumar, S. Advani, and V. Narayanan, „Sparse to dense depth completion using a generative adversarial network with intelligent sampling strategies,” in *Proc. ACM Int. Conf. Multimedia*, p. 5528–5536, 2021.
- [40] B. Nagy and C. Benedek, „3D CNN-based semantic labeling approach for mobile laser scanning data,” *IEEE Sens. J.*, vol. 19, pp. 10034–10045, Nov 2019.