

Nagy pontosságú szenzor-kalibráció geometria alapú számítógépes látáshoz

Pusztai Zoltán

Doktori Értekezés Tézisei

Témavezető:

Hajder Levente, PhD (ELTE IK)

Algoritmusok és Alkalmazásaik Tanszék
Informatikai Kar
Eötvös Loránd Tudományegyetem

Budapest, 2019.

Bevezetés

A szenzorkalibráció célja az eszközökhöz választott matematikai modell paramétereinek becslése. A disszertáció új eljárásokat mutat be a számítógépes látásban gyakran használt eszközök kalibrációjához, mint például a strukturált fényes szkener, kamerák és a LiDAR (Light Detection and Ranging) szenzor.

A kalibrációs lépés nélkülözhetetlen számos eljárás számára a számítógépes látás területén. Az egyes paramétereiből szerzett előzetes információ (pl. a fókusz távolság vagy a kamerák elhelyezkedése) által a rá épülő algoritmusok futási ideje csökken, míg pontosságuk nő. Néhány esetben ezeket a paramétereket online, az eszközök használata közben is meg lehet becsülni. Azonban az online kalibrációs eljárások általában pontatlanok, és csak abban az esetben használhatók, ha az eszközt nem lehet fizikailag megközelíteni, és a manuális kalibráció nem végrehajtható. A disszertációban taglalt eljárások célja a lehető legnagyobb pontosság elérése, szemben az automatikus vagy objektum nélküli kalibrációval. A bemutatott új eljárások sok esetben az ismert objektumok geometriáját használják fel és megfeleltetéseket alkotnak a különböző nézetek között.

Néha speciális eszközök esetében is szükség lehet kalibrációra. Amíg a szakirodalomban bőségesen fellelhetőek a kamera, LiDAR, GPS vagy IMU egységek kalibrációjára kidolgozott algoritmusok, a forgóasztal kalibrációjával csak kevés módszer foglalkozik. Forgóasztalt gyakran használnak 3D szkennerekhez, azonban a legtöbb szkennerekre kidolgozott kalibrációs eljárás

nem foglalkozik ezzel az eszközzel. A forgóasztal kalibrációja során az asztal forgástengelyét kell megkeresni a kamera koordinátarendszerében. Ennek ismeretében lehetséges a tárgyak 360 fokos szkennelése, és a különböző nézetekből rekonstruált pontfelhők pontos összeillesztése.

A 3D szkennerek pontos kalibrációja által lehetséges GT (Ground Truth) adatot generálni képi pontmegfeleltető algoritmusok kvantitatív összehasonlításhoz. A legtöbb jelenlegi összehasonlító adatbázis homográfiát használ a képpárok közötti megfeleltetésre. Azonban a pontos kalibráció és a 3D szkennerek által rekonstruált nagy felbontású térbeli pontfelhők segítségével valós, térbeli, mozgó objektumokat is használhatunk az összehasonlításhoz. Ezáltal a képi jellemzőket detektáló algoritmusok kvantitatív kiértékelése pontosabb képet ad azok működéséről. Ezenfelül az összehasonlító adatbázist kiegészítettük drónra szerelt kamera által készített kültéri képekkel, ezzel elhagyva a laboratóriumi körülményeket. Ebben az esetben a GT adat a drón kötött mozgásából nyerhető ki.

A kalibrációs lépés akkor jelenti a legnagyobb kihívást, amikor különböző modalitású eszközökről van szó. Ilyen eszközök például a kamerák és LiDAR szenzorok. Az előbbi nagy felbontású, színes képeket készít, míg az utóbbi ritka pontfelhőt generál. Még az általuk visszaadott kimenet dimenziója és felbontása is jelentősen eltérő, ezért a kalibráció során különös figyelmet kell fordítani arra, hogy pontos megfeleltetések nyerjünk a két eszköz között.

A több kamerából és LiDAR szenzorokból felépülő rendszereket előszeretettel használják napjainkban az önvezető járműveknél. Ezen eszközök

külső kalibrációja során azok egymáshoz viszonyított helyzetét kell meghatározni. A pontos kalibráció lehetővé teszi, hogy az önvezető jármű környezete feltérképezhető legyen még akkor is, ha az eszközök látómezeje közötti átfedés csekély.

A disszertáció az elméleti eredményeken kívül könnyen használható, új eljárásokat nyújt különböző modalitású eszközök kalibrációjára. Minden bemutatott módszer megbízhatóan működik valós körülmények között is, egyes módszerek az ipari partnerek figyelmét is felkeltették.

Irodalmi áttekintés

A disszertáció három témakör esetében nyújt újdonságot a számítógépes látás területéhez. Ebben a fejezetben a témakörök irodalmának áttekintése és a jelenlegi state-of-the-art módszerek leírása található.

Strukturált fényes szkennelés forgóasztallal

A strukturált fényt felhasználó 3D szkennelési technika régóta közkedvelt kutatási terület a számítógépes látás területén. Ezek a szkennerek kamerával és projektorral vannak felszerelve, és strukturált fényt használnak a tárgyak alakjának meghatározására. A strukturált fény egy képsorozat, amely minden egyes projektor pixelt egyedileg elkódol. Az objektumok alakja úgynevezett háromszögeléssel nyerhető ki. Egy előzetes kalibrációs lépés is szükséges, hogy meghatározzuk az eszközök külső és belső paramétereit. A kalibrációt jellemzően egy ismert geometriájú tárggyal végzik el, például sakktáblával. A sakktáblával a kamera belső paramétereit és a sakktábla kamerához viszonyított helyzetét lehet meghatározni. Ugyanez a kalibrációs eljárás használható a projektor esetében is, azonban ehhez a sakktábla sarokpontjainak vetületét ismerni kell a projektor síkján. A vetületek a strukturált fény visszafejtése által számolhatóak ki. A korábbi kalibrációs eljárások csak a nyersen, a megfelelő kamera pixelekből visszafejtett koordinátákat használták a kamera és projektor pixelek közötti megfeleltetések létrehozásához. Azonban ezek a megoldások csak pixel pontosságot képesek elérni, és nem robusztusak a dekódolás során bekövetkezett hibákra. Ezért az első tézis erre a problémára nyújt megoldást homográfia használatával. A homográfia egy térbeli sík két vetülete közötti transzformációt írja le, amelyet robusztusan becsülünk a sakktábla összes pontját felhasználva. Így lehetséges a homográfia

segítségével a zajosan visszafejtett projektor pixelek észlelése és a szubpixel pontosság elérése, ami javítja a kamera-projektor kalibráció pontosságát.

Az általunk használt szkener egy forgóasztallal is fel van szerelve, így lehetséges a tárgyak 360 fokos szkennelése. A szkener különböző nézőpontokból rekonstruálja az objektumok egyes részeit. Ahhoz, hogy ezeket a pontfelhőket összesítsük, a forgóasztal forgástengelyét kell meghatározni. Az ebben a témakörben publikált korábbi eljárások többsége az ICP (Iterative-Closest-Points) algoritmust használja, amely csak a nyers pontfelhőkkel és becsült kezdeti értékekkel dolgozik. Egyéb eljárások forgó sakktáblát használnak, majd a kamera sakktáblához viszonyított külső paramétereit alapján kört illesztenek a térben, amely középpontja a forgástengelyen helyezkedik el. Az imént említett technikák egyike sem ad kielégítő eredményt. A különböző nézetből összesített pontfelhők illesztési hibákat tartalmaznak, ezért egy teljesen új módszer került kidolgozásra a forgóasztal tengelyének kalibrációjához. Ez különösen nehéz feladat, hiszen a tengely becslése során már a kisebb pontatlanságok esetében is látható illesztési hibák lépnek fel az összeillesztett pontfelhőben. Az új módszer nem csak a forgástengelyt képes megbecsülni, hanem a kamera és projektor külső paramétereit is finomítja. Ezáltal még pontosabb kalibrációt kapunk.

Affin kovariáns jellegzetes pontkövető algoritmusok kvantitatív kiértékelése

A jellegzetes képpontok, azaz feature pontok használata rendkívül elterjedt a számítógépes látás minden területén. Napjainkban ezeket kiegészítik az úgynevezett affin transzformációkkal is. Az affin transzformáció a kép lokális torzulását becsli a pontjellemzők körül. Ezért azon epipoláris geometriában

jelentkező problémák, amelyek megoldhatóak pontmegfeleltetések használatával, szintén megoldhatóak az affin transzformációkkal, jóval kevesebb megfeleltetést felhasználva. Annak ellenére, hogy rengeteg összehasonlító adatbázis létezik pontmegfeleltetések tesztelésére, kevés az affin transzformációkat is figyelembe vevő szakirodalom és kvantitatív összehasonlítás. Ezek közül a legismertebbek homográfiát használnak az összehasonlításhoz. A képek közötti homográfia használata során két probléma merül fel: egyrészt nem foglalkozik a képi jellemzők párosításának a problémájával, illetve csak sík objektumok vagy pedig statikus kamera esetében használható.

A hasonló kiértékelő rendszereknél használt mérték, melyet repeatability-nek neveznek, szintén nem megfelelő több szempontból sem. A repeatability egy arányszám a kölcsönösen megtalált pontok és a két képen talált jellegzetes pontok számának minimuma között. Egy jellegzetes pontpárt akkor tekintünk kölcsönösnek, ha a hozzájuk tartozó affin régiók metszete nagyobb, mint egy előre meghatározott érték. Azonban a nagyobb affin régióval rendelkező pontpárok könnyebben teljesítik ezt a feltételt, mint a kisebbek. Ezenkívül a képpárokon talált jellegzetes pontok számának minimuma miatt ez a mérték rosszul jellemzi az algoritmusok viselkedését, ha az egyik képen csupán kevés jellegzetes pontot sikerült találni.

Azért, hogy kijavítsuk a fentebb említett összehasonlítások hátrányait, két teljesen új módszert mutatunk be GT adat generálására és a kiértékelésre. Az első esetben a pontosan kalibrált strukturált fényes szkennert használjuk a GT előállítására, ahol a megfigyelt tárgyak forgó mozgást végeznek a képek között. A

második esetben a kiértékeléshez használt képeket kültéren, egy drónra szerelt kamera által vettük fel, és a GT a drón kötött mozgásából számolható. A repeatability mérték helyett pedig az affin transzformációk Frobenius normáját és a pontmegfeleltetések euklideszi normáját használjuk.

Multi-LiDAR Multi-Kamera Rendszerek Pontos Kalibrációja

Az autonóm vezetés napjaink egyik központi tématerülete a kutatói közösségek és az ipari partnerek körében is. Ahhoz, hogy az önvezető rendszerek környezetét feltérképezzük, több szenzor egyidejű használata szükséges. Azért, hogy ezek a szenzorok hatékonyan együtt tudjanak dolgozni, az egymáshoz viszonyított pozíciójuk és orientációjukat kell ismernünk, tehát az eszközök külső kalibrációját kell elvégezni.

Kamera és LiDAR szenzorok egyidejű használatával kompenzálhatjuk azok hátrányait, ezért előszeretettel alkalmazzák őket több különböző területen is. Ezen szenzorok külső kalibrációja manapság aktív kutatási területnek számít. Megkülönböztethetünk online és offline kalibrációs eljárásokat. Az online módszerek az eszközök használata közben működnek, és általában nem használnak külön kalibrációs tárgyat. Az offline módszerek esetében szükség van egy vagy akár több kalibrációs tárgyra. Általánosságban elmondható, hogy az offline módszerek pontosabb eredményt adnak.

A legtöbb korábban publikált kalibrációs eljárás sakktáblákat vagy pedig egyéb síkszerű tárgyat használ a kalibrációhoz. Ennek indoka az, hogy a sakktáblákat automatikusan és nagy pontossággal lehet használni kamera

kalibrálásra. Az összes hasonló módszer esetében problémát okoz a tábla széleinek pontos megtalálása a LiDAR pontfelhőben. Számos megoldás született ennek a problémának a megoldására, viszont egyik sem bizonyult elegendően pontosnak. Továbbá a sakktábla mintája növeli a LiDAR pontfelhő zajosságát, ami rontja a kalibráció pontosságát is.

Más módszerek a színtér egy részét rekonstruálják több kameraképet felhasználva, majd az így nyert pontfelhőt próbálják meg összeilleszteni a LiDAR-ral szkennelt pontfelhővel. Néhány algoritmus az ICP-t használja a pontfelhők összeillesztésére, azonban ennek az algoritmusnak jó kiindulási paraméterekre van szüksége, illetve a pontossága sem mindig megfelelő. Kifinomultabb módszerek közös síkok detektálása után keresnek megfeleltetéseket a síkok között. A síkok megfeleltetése szintén nehéz feladat, és több megoldás is lehetséges.

Az általunk fejlesztett módszer papírdobozokat használ kalibrációs tárgyként. Ezen dobozoknak három merőleges síkja van, amit nagy pontossággal lehet a LiDAR pontfelhőben detektálni. Ezen síkok metszései adják a doboz éleit, amikből végül a doboz sarokpontjait lehet meghatározni. Ezen sarokpontok vetületeit a kameraképen lehetséges nagyon pontosan megtalálni. Az így nyert 3D-2D pontmegfeleltetésekkel a külső kalibráció problémáját visszavezettük a PnP (Perspective-n-Point) problémára, amely megoldására már léteznek hatékony módszerek. Az eljárás képes LiDAR-LiDAR kalibrációra is. Ebben az esetben a külső paramétereket a térbeli pontok regisztrációja adja. A módszert könnyen ki lehet terjeszteni több LiDAR és több kamera együttes kalibrációjára is.

Tézisek

A disszertáció a következő három tézispontot tartalmazza.

1. Tézis – Forgóasztalos strukturált fény használó szkennerek kalibrációja [4;7;8]

Egy új módszert dolgoztunk ki forgóasztalos strukturált fényt használó szkennerek kalibrációjára, amely képes szubpixel pontos megfeleltetéseket kinyerni a kamera- és projektor pixelek között. A módszer erre a célra homográfiát használ, amelyet a sakktábla kamerából és projektorból egyaránt látszó felületén számít ki robusztusan. Ezt követően a homográfiával lehetséges a rosszul visszafejtett strukturált fény javítása, illetve a megfeleltetések szubpixel pontos finomítása. A forgóasztal kalibrációjához egy teljesen új eljárást mutattunk be a forgástengely kiszámítására a kamera koordinátarendszerében. A módszer iteratív módon finomítja a kamera és a projektor külső paramétereit. A valós környezetben végzett vizuális összehasonlításban a javasolt módszer pontosabb eredményt ad, mint a konkurens módszerek.

2. Tézis – GT adat generálás affin kovariáns jellegzetes pontokat detektáló algoritmusokhoz és kvantitatív kiértékelésük [1;3;6]

Új, valós adatot generáló algoritmusokat javasoltunk az affin invariáns jellegzetes pontokat detektáló algoritmusok kiértékelésére. A valós adat nem

csak pontmegfeleltetéseket tartalmaz, hanem a hozzájuk tartozó affin transzformációkat is. A létrehozott adatbázis felülmúlja a korábbiakat, ugyanis ez valós, térbeli, forgó mozgást végző tárgyakat is tartalmaz, míg a korábbi hasonló munkák csak síkszerű vagy pedig statikus kamerával felvett eseteket fedtek le. Az adatbázist kiegészítettük kültéri felvételekkel is, amelyek drónra felszerelt kamerával készültek. Több algoritmust is bemutatunk, amelyek a drón kötött mozgásából képesek a GT affin transzformáció kiszámítására.

3. Tézis – Kamera-LiDAR rendszerek külső kalibrációja

[2;5]

Új módszert javasoltunk a kamera-LiDAR szenzorok külső kalibrációjához. A jelenlegi state-of-the-art módszerekkel ellentétben a javasolt módszer térbeli kalibrációs tárgyakat használ síkszerűek helyett. Először a dobozok 3D koordinátáit számolja ki a LiDAR pontfelhőben nagy pontossággal, majd a kamera és LiDAR külső kalibráció problémáját visszavezeti a PnP (Perspective-n-Point) problémára 3D és 2D pontmegfeleltetéseket használva. LiDAR-LiDAR kalibráció szintén lehetséges ugyanazon térbeli sarokpontok megfeleltetése és pontregisztráció segítségével. A javasolt módszer a szintetikus és a valós tesztek során is nagyobb pontosságot ért el a konkurens algoritmusoknál. Ezenfelül a módszert könnyen ki lehet terjeszteni LiDAR és kamera rendszer kalibrációjára, azaz több egyidejűleg használt szenzor használata esetén is alkalmazható.

Köszönetnyilvánítás

Hálásan köszönöm a következő projektek által nyújtott támogatást:



EMBERI ERŐFORRÁSOK
MINISZTERIUMA

Az Emberi Erőforrások Minisztériuma ÚNKP-18-3 kódszámú Új
Nemzeti Kiválóság Programjának támogatásával készült.

EFOP-3.6.3-VEKOP-16-2017-00001: Tehetséggondozás és kutatói utánpótlás fejlesztése autonóm járműirányítási technológiák területén – A projekt a Magyar Állam és az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

A szerző témában megjelent publikációi:

- [1] **Pusztai, Z;** Hajder, L. *Quantitative Affine Feature Detector Comparison based on Real-World Images Taken by a Quadcopter*. In: Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 5: VISAPP, 2019, ISBN:978-989-758-354-4, pages 704-715. DOI:10.5220/0007372907040715
- [2] **Pusztai, Z;** Eichhardt I; Hajder L. *Accurate Calibration of Multi-LiDAR-Multi-Camera Systems*. In: SENSORS (Basel). 2018; 18(7):2139. Published 2018 Jul 3. doi:10.3390/s18072139
- [3] **Pusztai, Z;** Hajder, L: *Affine Invariant Feature Tracker Evaluation based on Ground-Truth Data*, In: Szirmay-Kalos, László; Renner, Gábor, IX. magyar számítógépes grafika és geometria konferencia, GRAFGEO 2018, Budapest, Magyarország : Neumann János Számítógép-tudományi Társaság (NJSZT), (2018) pp. 164-173. , 10 p. ISBN : 978-963-313-282-1
- [4] **Pusztai, Z;** Hajder, L: *Ground-truth tracking data generation using rotating real-world objects* In: Braz, J; [et, al.] Computer Vision, Imaging and Computer Graphics Theory and Applications, 11th International Joint Conference, VISIGRAPP 2016, Revised Selected Papers Cham (Németország), Németország : Springer, (2017), pp. 395-417. ISBN: 978-3-319-64870-5
- [5] **Pusztai, Z;** Hajder, L: *Accurate Calibration of LiDAR-Camera Systems using Ordinary Boxes*. In: CVF 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), (2017) pp. 394-402. , 9 p. ISSN: 2473-9944, DOI: 10.1109/ICCVW.2017.53
- [6] **Pusztai, Z;** Hajder, L: *Quantitative Comparison of Affine Invariant Feature Matching* In: Imai, F; Tremeau, A; Braz, J: VISAPP 2017. Proceedings of

the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. Vol. 6, Setúbal, Portugália : SciTePress, (2017), pp. 515-522. , 8 p., DOI: 10.5220/0006263005150522

[7] **Pusztai, Z**; Hajder, L: *Quantitative Comparison of Feature Matchers Implemented in OpenCV3* In: Čehovin, L; Mandeljc, R; Štruc, V (szerk.), Proceedings of the 21st Computer Vision Winter Workshop. CVWW 2016, Ljubljana, Szlovénia : SDRV, (2016), pp. 1-9. ISBN 978-961-90901-7-6

[8] **Pusztai, Z**; Hajder, L: *A Turntable-based Approach for Ground Truth Tracking Data Generation* In: Magnenat-Thalmann, N et al. (szerk.), VISIGRAPP 2016. Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. Vol. 3. VISAPP 2016, Setúbal, Portugália : SciTePress, (2016), pp. 500-511., ISBN: 978-989-758-175-5, DOI:10.5220/0005719404980509

A szerző egyéb publikációi:

[9] Gábor, Rácz ; **Zoltán, Pusztai** ; Balázs, Kósa ; Attila, Kiss: *An improved Community-based Greedy algorithm for solving the influence maximization problem in social networks*, ANNALES MATHEMATICAE ET INFORMATICAЕ 44 pp. 141-150. , 10 p. (2015), ISSN 1787-5021

[10] Balázs, Kósa ; Márton, Balassi ; Péter, Englert ; Gábor, Rácz ; **Zoltán, Pusztai** ; Attila, Kiss: *A Basic Network Analytic Package for RapidMiner* In: Simon, Fischer; Ingo, Mierswa (szerk.) Proceedings of the 5th Rapidminer World 2014, Aachen, Németország : Shaker Verlag, (2014), pp. 47-59. , 13 p., ISBN-13: 978-3844029468

