

hajder_274_24

MTA doktori értekezés tézisei

**Geometriai alapú háromdimenziós számítógépes
látás digitális kamerák és LiDAR érzékelők
segítségével**

Hajder Levente

**Eötvös Loránd Tudományegyetem
Informatikai Kar**

Algoritmusok és Alkalmazásaik Tanszék

Budapest, 2025

hajder_274_24

Tartalomjegyzék

1. Bevezetés és előzmények	2
2. Elméleti alapok	3
2.1. Visszavetítés és felületmetszés	3
2.2. Képalapú háromdimenziós rekonstrukció	4
2.2.1. Kétképes (sztereo) rekonstrukció	4
3. I. tétel: Elméleti affin eredmények	7
4. II. tételcsoport: Affin megoldók	8
5. III. tételcsoport: Kötegelt behangolás újszerű alkalmazása nem hagyományos problémaosztályokra	11
6. IV. tételcsoport: Új módszerek LiDAR mérőeszközök és perspektív kamerák kalibrációjára	13

1. Bevezetés és előzmények

A háromdimenziós érzékelés a műszaki tudományok körülbelül ötven éve művelt ága. A kezdetekben digitalizált kamerák képeinek feldolgozásából foglalkoztak a kutatók. Kezdetekben komoly nehézséget okozott, hogy a gépek számítási teljesítménye sok-sok nagyságrenddel a ma korszerűnek mondható architektúrák alatt volt. A számítógépes látás őskorában ezért csak a szomszédos pixelek közötti kapcsolatra tudtak építeni, akkor alkották meg a mai napig jól működő éldetektorokat [SHB07], később sarokdetektorokat [Har88, Bir97, BM02], alapvető mintaillesztő eljárásokat.

A 80-as évek második felétől már több képet is fel tudtak dolgozni, és ez lehetővé tette térbeli információ kinyerését a képekből. Az ezredforduló tájékán megjelentek azok az algoritmusok, amelyek már a mélységet is képesek voltak hardver segítségével egy nézőpontból meghatározni.

Jómagam a 2000-es évek eleje óta foglalkozom a területtel. Előbb karakterfelismeréssel kezdtem foglalkozni, majd doktoranduszi tanulmányaim során áttértem a háromdimenziós látásra, és meg is maradtam ezen a részterületen mind a mai napig. Különösen az ragadott meg, hogy az algebra és a geometria határterületén megtapasztalhattam, a gyakorlatban használható módszerek esetében milyen fontos, hogy elméletileg is meg legyenek alapozva.

Korábbi témavezetőm, a tavaly sajnálatosan elhunyt Csertverikov Dmitrij 2010 után ismerttetett meg rendkívül kreatív tanítványával, Molnár Józseffel, aki egy SZTAKI-SZTE projekt keretében elkezdett foglalkozni az affin transzformációk alkalmazásával [MC14]. Bár ő hamar felhagyott a területtel, de számomra és saját doktoranduszaim számára ez a terület egy teljesen új világot nyitott meg, és alapvetően meghatározta az évtized kutatási irányait. Az elméleti alapok kidolgozása után saját algoritmusokat kezdtünk készíteni a tanítványaimmal. Az első két téziscsoport ezzel a területtel foglalkozik, az első csoport az elért elméleti eredményeket tartalmazza, a második az elméletre épülő becslési eljárásokat, megoldókat ismerteti.

Időközben egy másik problémakörrel is foglalkoztam: a háromdimenziós rekonstrukciós eljárásokra kifejlesztett numerikus algoritmus, a "kötegelt behangolás" kiterjesztését végeztük el újabb problémaosztályokra. Megmutattuk, hogy sok paraméter esetén is lehetséges numerikus eljárással a becslést finomítani, amennyiben a mért értékek és a paraméterek között lazább a kapcsolat, és a Jakobi-mátrixban sok zérus érték található. Felületrekonstrukciós feladatokra adaptáltuk az eredeti módszert. Ezekből az eredményekből születtek a harmadik csoport tézisei.

A negyedik terület, amelyről a doktori eljárás során a tudományos tézisfüzetben említést teszek, LiDAR eszközök és perspektív kamerák kalibrációja. Amikor 2018-ban az ELTE Informatikai Karán elindult az autonómrendszer-informatikus (MSc) képzés, a járműre rögzített érzékelők feldolgozásával kezdtünk kollégáimmal foglalkozni. A kamerák mellett önvezető járművek esetén a LiDAR lézeres letapogatóeszközök is nagyon hasznosak/népszerűek. Azt tapasztaltuk, hogy kamerák és LiDAR-ok közös kalibrációjához a jelenlegi megoldások nem elég pontosak, ezért több új eljárást is javasoltunk az elmúlt években.

2. Elméleti alapok

A számítógépes látás az esetek többségében középpontos vetítést megvalósító kamerákat alkalmaz, ezen a kategórián belül is a leggyakrabban lyukkamerát szokás alkalmazni [HZ03]. Ezen leképző eszközök esetén a vetítés geometriai modellje egyszerű: a térbeli pontot megfigyelő pixel helyét úgy határozzuk meg, hogy a térbeli pontot összekötjük a kamera fókuszpontjával, így kapunk egy vetítő egyenest. Ennek az egyenesnek és a képsíknak a metszéspontja jelöli ki azt a pixelt, amelyen a térbeli pont vetülete látszik.

A digitális kamerák szinte kivétel nélkül középpontos vetítést valósítanak meg, de lehetőség van sík helyett más felületekre is vetíteni. Az érdeklődő Olvasók Geyer és Daniilidis munkájában [GD00] találhatnak kiváló összefoglalást a lehetséges kameramodellekre. Munkám során én elsősorban síkra vetítést megvalósító lyukkamerával foglalkoztam. A hétköznapiakban nemperspektív kamerákat is szoktak használni, de ezeket kiegyenesítés¹ algoritmussal perspektívokká lehet alakítani [Zha00, SMS06].

A képsíkra vetítés összefüggése így írható fel:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \mathbf{K} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} f_u X + u_0 Z \\ f_v Y + v_0 Z \\ Z \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

ahol f_u és f_v az optika fókusz távolságának és a szenzoron a (vízszintes és függőleges) pixelméretnek a szorzata. Ezt a két paramétert a szakmában fókusz távolságnak szokás nevezni, ebben az alakban a mértékegysége pixelben értendő. Ha a pixel alakja négyzet, ami az esetek döntő többségében igaz, akkor f_u és f_v egyenlőek. Az $[u_0 \ v_0]^T$ pixelpozíció az úgynevezett dőféspont, amit geometriai megközelítésben a képsík és az optikai tengely metszéspontjaként kaphatunk meg. Ez a pont a perspektíva középpontja a képen. Az összefüggésben a $\sim$ operátor a skálázás erejéig egyenlőséget jelenti, egyenlőséget a homogén osztás elvégzésével kaphatunk.

Ha a koordináta-rendszer "valahol a világban" van, egy $\mathbf{R}$ elforgatásmátrix és egy $\mathbf{t}$ eltolásvektor segítségével előbb a kamerába kell transzformálni a pontokat². A formula ekkor így alakul át:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \mathbf{K} [\mathbf{R}|\mathbf{t}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \quad (2)$$

Ebben az esetben már a térbeli pontot is homogén koordinátával kell leírni.

2.1. Visszavetítés és felületmetszés

A 2. összefüggés invertálásával a visszavetítést is jól meg tudjuk határozni:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \alpha \mathbf{K}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}, \quad (3)$$

¹ Angolul: rektifikáció

² Az $\mathbf{R}$ forgatás és a $\mathbf{t}$ eltolás segítségével az elképzelhető összes egybevágósági transzformációt le lehet írni a térben, ha $\mathbf{R}$ ortonormált, azaz $\mathbf{R}^T \mathbf{R} = \mathbf{I}$.

ahol α értéke tetszőleges lehet, és a skálázás erejéig egyenlőség operátort váltja ki a 2. összefüggésben. Ez az egyenlőség geometriailag úgy értelmezhető, hogy az adott pixelhez vetítés előtt végtelen sok pont tartozik, ezen pontok alkotják a vetítőegyenest. Az egyenesen az α paraméter változtatásával tudunk végigmenni. A gyakorlatban ez a paraméter a Z koordináta, amit szokás projektív mélységnek is nevezni.

Amikor egy térbeli pontot kamerával megvizsgálunk, a pixel, amin a pont látszik, egy vetítőegyenest határoz meg. Ha ismerjük azt a felületet, amin a térbeli pont elhelyezkedik, akkor a rekonstrukciót a vetítőegyenest és a felület metszéspontjaként lehet kiszámítani, ahogyan arra egy példát látunk a disszertációm IV.2. tézisében.

2.2. Képalapú háromdimenziós rekonstrukció

Ebben a fejezetben röviden összefoglalom a két- és sokképes rekonstrukció lépéseit, kiemelve a legfontosabb módszereket.

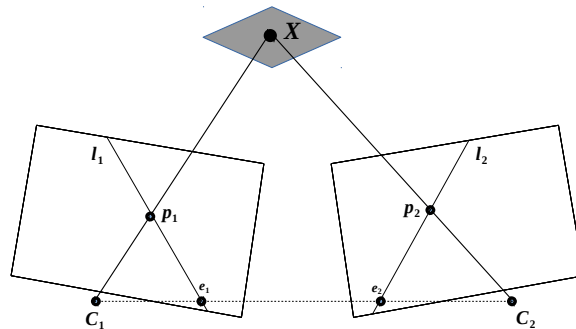
2.2.1. Kétképes (sztereó) rekonstrukció

A sztereó rekonstrukció problémáját az 1. ábrán láthatjuk. Egy térbeli felületdarabka X vektorral jellemzett pontját levetítjük a két kamerára. Középpontos vetítést alkalmazunk, a két kamera vetítési középpontjait, melyeket fókuszpontnak hívnak, C_1 -el és C_2 -vel jelöljük. A vetítés legegyszerűbb modellje a síkra vetítés, egy átlagos kamera igen jó közelítéssel megfelel ennek a modellnek. Ahogyan azt fentebb már leírtam, a térbeli X pont képét úgy kapjuk meg, hogy összekötjük a pontot a fókuszpontokkal, így vetítősugarakat kapunk, és a két kamerásíkon az elmetszett pixelek adják a képen a pont helyét.

A számítógépes látás feladata algoritmikusan visszaállítani a térbeli X pont helyét, ha a két képen ismerjük a vetületi pontokat. A rekonstrukciós folyamatokat az alábbi lépésekre szokták osztani:

1. kamera kalibráció,
2. epipoláris geometria becslése és
3. háromszögelés (trianguláció).

Kamera kalibráció. Amennyiben ismerjük a kamerának a belső paramétereit, azaz a fókusz táv(ok)at és a döféspontot, a rekonstrukciós egyenletek egyszerűsödnek, és az eredmények numerikusan is pontosabbak lesznek. Sakktábla segítségével az alkalmazott kamerákat általában előzetesen be lehet kalibrálni. Sakktáblát azért használnak széleskörűen [Zha00], mert szabályos a mintázata, és a mezők alakját (élek, sarokpontok) nagyon pontosan meg lehet határozni. Amennyiben a belső paramétereket ismerjük, a kamerákat *kalibráltaknak* nevezik.



1. ábra. Sztereó látás alapproblémája. A térbeli objektum X pontját a két képsíkra középpontosan vetítjük. A kamerákat (vetítést) a képsíkok, a C_1 , C_2 fókuszpontok és a fókusz távolság határozzák meg.

Epipoláris geometria becslése. Az epipoláris geometria a projektív geometria kétképes esete. Kalibrált kamerák esetén egyszerűbb a feladat, ekkor az úgynevezett lényegi (esszenciális) mátrix reprezentálja az epipoláris geometriát. Ismeretlen kameraparaméterek esetén csak az alapvető (fundamentális) mátrix meghatározása lehetséges.

Fundamentális mátrix becslése két kép között. Két általános perspektív kamera között megadható egy 3×3 -as mátrix, amely a képek között levő geometria kapcsolatát általános esetben is képes leírni. Ezt a mátrixot hívják alapvető (fundamentális) mátrixnak, és a szakirodalomban általában F -fel jelölik. A mátrixnak ugyan kilenc eleme van, de valójában hét paraméter segítségével leírható, mert két megkötést is fel lehet tenni:

1. Egyrészt a mátrix szinguláris, azaz a determinánsa zérus, ebből egy megkötés származik.
2. Másrészt skálázásra invariáns, azaz ugyanazt a geometriai kapcsolatot írja le a mátrix, ha egyazon valós számmal szorozzuk meg az alapvető mátrix minden elemét. Ez a tény még egy szabadságfokot csökkent.

Kalibrált kamerák esetén az alapvető mátrixban levő ismeretlenek számát tovább lehet csökkenteni, mivel csupán a kamerák külső paraméterei határozzák meg azt. Ebben az esetben kapjuk a lényegi mátrixot. A forgatás – melyet az $\mathbf{R}$ ortonormált mátrixszal szokás reprezentálni – három paraméterrel írható le, a $\mathbf{t} = [t_x \ t_y \ t_z]^T$ eltolásvektort további hárommal, de képalapú térbeli látás esetén az eltolás nagysága nem rekonstruálható, csupán az iránya, ezért csak két paramétert lehet az eltolásból meghatározni. Így összesen $3 + 2 = 5$ paramétert kell számítani, ebből a kettőből lehet a lényegi mátrixot előállítani az alábbi módon:

$$\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R}, \quad (4)$$

ahol $[\mathbf{t}]_{\times}$ a $\mathbf{t}$ eltolásvektorból készített, vektoriális szorzást reprezentáló 3×3 -as mátrix. Az F alapvető és E lényegi mátrixokat kalibrált kamerák esetén könnyen át lehet számítani az $\mathbf{E} = \mathbf{K}_2^T \mathbf{F} \mathbf{K}_1$ összefüggéssel, melyben $\mathbf{K}_1$ és $\mathbf{K}_2$ az első és a második kamera belső paramétereit leíró felső háromszög mátrix, melyet az 1. összefüggésben adtam meg.

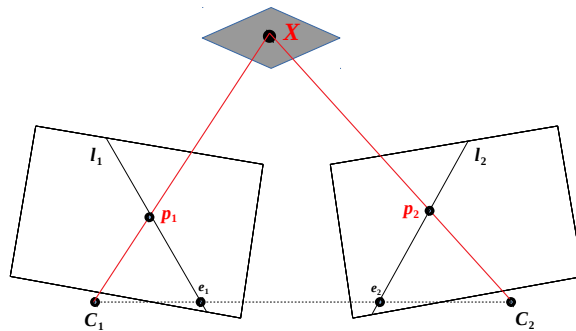
Mind az alapvető, mind a lényegi mátrixok meghatározására a sztereó képpárokon pontmegfeleléseket határoznak meg. Minden pontpár egy egyenletet ad, melyet az alábbi alakban lehet leírni:

$$\mathbf{p}_2^T \mathbf{F} \mathbf{p}_1 = \mathbf{p}_2^T \mathbf{K}_2^{-T} \mathbf{E} \mathbf{K}_1^{-1} \mathbf{p}_1^T = 0. \quad (5)$$

Az esszenciális mátrixhoz öt [Nis03], a fundamentális mátrixhoz hét pontpárra van szükség [HZ03], miután minden pontpár egy egyenletet ad. Praktikus okokból a fundamentális mátrix becslésére a nyolcpontos algoritmus [Har95] is igen népszerű, bár eggyel kevesebb pontból is meghatározható a feladat.

Háromszögelés (trianguláció)

Ha két képen ismerjük az egymásnak megfelelő pontokat, a pont térbeli helyzete is meghatározható. A képeken a két pont egy-egy vetítőegyenest határoz meg, az eredeti térbeli pont helye közel lesz ezekhez a vetítő egyeneshez. A feladathoz tartozó magyarázó rajzot a 2. ábra mutatja meg. Pirossal jelöltem a vetítőegyeneseket, a hozzájuk legközelebbi pont adja a becsült térbeli pont helyét.



2. ábra. A triangulációs probléma: a két képen $\mathbf{p}_1$ és $\mathbf{p}_2$ egymásnak megfelelő pontok meghatároznak egy-egy vetítőegyenest. Ideális esetben ezek metszéspontja adja a vizsgált térbeli pont helyét. A zaj miatt a gyakorlatban az egyenesek kitérőek, ezért a térbeli pont helyét becsülni kell.

Többféle eljárás létezik háromszögeléshez. Van, amelyik algebrai hibát [HZ03] minimalizál, másik megoldás a vetítőegyeneselekhez legközelebbi pontot találja meg. Létezik optimális trianguláció [HS97] is kétképes estre, amely attól optimális, hogy a képtérben minimalizálja a hibát, és a képek közötti geometriai megkötést is rögzíti, melyet az alapvető mátrix határoz meg.

3. I. t  zis: Elm  leti affin eredm  nyek

Az I. t  ziscsoport az affin transzform  ci  k elm  let  vel foglalkozik. Megmutattam a disszert  ci  ban, hogy milyen t  rv  nyszer  s  gek vezethet  k le lyukkamer  kra a lok  lis affin transzform  ci  k felhaszn  l  s  val.   sszesen h  rom t  rv  nyt siker  lt kimondani. Az els   k  t t  rv  nyt ugyan j  magam is publik  ltam, de nem tekintem saját eredm  nynek, ellent  tben az utols  val, amelyik   sszef  gg  st nekem siker  lt el  sz  r fel  rnom. A t  rv  nyek megfogalmaz  s  t ugyanakkor eredm  nynek tekintem, hiszen a kutat  csoportom munk  j  nak jobb meg  rt  s  t szolgálja.

Ezek alapj  n a t  ziscsoporton bel  l k  t alt  zist mondok ki:

I.1. t  zis. Affin transzform  ci  k t  rv  nyszer  s  gei [5,17,20].

Kutat  munk  nk   s a minket meg  l  z   szakemberek munk  ss  ga alapj  n h  rom t  rv  nyben fogalmaztam meg az affin transzform  ci  kkal kapcsolatos   sszef  gg  seket.

1. Az els   t  rv  ny a fel  leti norm  lvektor, a kameraparam  terek, a vet  t  si f  ggv  ny   s a vet  leti k  z  ppontok k  z  tti   sszef  gg  st   rja le:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \frac{1}{\nabla \Pi_u^1 T [\mathbf{n}]_\times \nabla \Pi_v^1} \begin{bmatrix} \nabla \Pi_u^2 T [\mathbf{n}]_\times \nabla \Pi_v^1 & \nabla \Pi_u^1 T [\mathbf{n}]_\times \nabla \Pi_u^2 \\ \nabla \Pi_v^2 T [\mathbf{n}]_\times \nabla \Pi_v^1 & \nabla \Pi_u^1 T [\mathbf{n}]_\times \nabla \Pi_v^2 \end{bmatrix}, \quad (6)$$

ahol a gradiens oper  tor (∇) jel  li a vet  t  si f  ggv  ny gradiens  t a t  rbeli koordin  ta szerint³, $\mathbf{n}$ pedig a fel  leti norm  lvektor. A kapott formula v  lem  nyem szerinti rendk  v  l nagy   rt  ke, hogy b  rmilyen vet  t  f  ggv  ny esetén m  k  dik, nem korl  toz  dik egyetlen kamerat  pusra sem. A fenti alapegyenlet tehát egy   ltal  nos   sszef  gg  st ad, az alkalmazott kamerat  pusok k  z  tti k  l  nbs  g a Π vet  t  f  ggv  nyekben jelenik meg.

2. A m  sodik t  rv  ny megmutatja, hogy milyen   sszef  gg  s k  pezhet   lyukkamera esetén a $\mathbf{H}$ homogr  fia   s a k  t k  pen egym  s  nak megfelel   $[u_1 \ v_1]^T$   s $[u_2 \ v_2]^T$ pixelkoordin  t  k k  z  tti, ha a lok  lis fel  let egy s  kr  l sz  rmazik. A parci  lis deriv  ltak alapj  n   sszesen n  gy   sszef  gg  st lehet fel  rni:

$$\begin{aligned} a_{11} &= \frac{\partial u_2}{\partial u_1} = \frac{H_{11}-H_{31}u_2}{s}, & a_{12} &= \frac{\partial u_2}{\partial v_1} = \frac{H_{12}-H_{32}u_2}{s}, \\ a_{21} &= \frac{\partial v_2}{\partial u_1} = \frac{H_{21}-H_{31}v_2}{s}, & a_{22} &= \frac{\partial v_2}{\partial v_1} = \frac{H_{22}-H_{32}v_2}{s}. \end{aligned} \quad (7)$$

3. A harmadik t  rv  ny a k  vetkez  , I.2. alt  zisben megfogalmazott   sszef  gg  s a fundament  lis (alapvet  ) m  trix   s a lok  lis affin transzform  ci   elemei k  z  tti.

I.2. t  zis. Line  ris kapcsolat az epipol  ris geometria   s az affin transzform  ci  k k  z  tti [5,17,20].

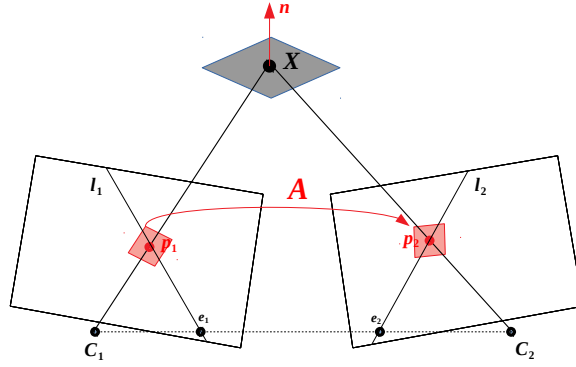
Amennyiben adott egy sz  nt  rr  l k  t, lyukkamer  val felvett k  p,   s a k  t k  p k  z  tti az $\mathbf{F}$ fundament  lis m  trix teremt kapcsolatot, akkor az affin transzform  ci     s a fundament  lis m  trix k  z  tti az al  bb vektoros (k  tdimenzi  s) egyenletet lehet megadni:

$$\mathbf{A}^T \hat{\mathbf{F}} \begin{bmatrix} \mathbf{p}_1 \\ 1 \end{bmatrix} = -\tilde{\mathbf{F}}^T \begin{bmatrix} \mathbf{p}_2 \\ 1 \end{bmatrix}, \quad (8)$$

³ Az als   index a v  zszintes/f  gg  leges   r  ny, a fels   index a k  p sz  m  t jel  li.

ahol $\tilde{\mathbf{F}}$ és $\hat{\mathbf{F}}$ a fundamentális mátrix "csonkított" változatai, melyeket úgy kapunk, hogy az eredeti $\mathbf{F}$ mátrixból az utolsó sort, illetve az utolsó oszlopot töröljük. $\mathbf{p}_1$ és $\mathbf{p}_2$ a két képen az egymásnak megfelelő pontok homogén koordinátákkal megadva, ezen pontok körül értelmezzük az $\mathbf{A}$ affin transzformációt.

Fontos megjegyzés. Az affin transzformációk alkalmazása esetén a háromdimenziós sztereó látás nemcsak térbeli helyeket (pontokat), hanem a felületi normálisokat is rekonstruálni képes, ahogyan az a 3. ábrán is látható. A rekonstrukció kimenete így nemcsak háromdimenziós pontfelhő, hanem *irányított pontfelhő* lesz. Irányított pontfelhő esetén minden egyes ponthoz a merőleges normálirány is adott.



3. ábra. Az affin transzformáció ismerete esetén a pirossal jelzett felületi normálvektor is meghatározható, amennyiben a kamera kalibrált, azaz ismerjük a vetítőfüggvényt.

4. II. téziscsoport: Affin megoldók

Ebben a téziscsoportban az I. csoport elméleti eredményeit alkalmazva kilenc algoritmus és algoritmuscsalád elkészítését ismertetem. Ezek az alábbi pontokban foglalhatóak össze:

II.1. tézis. Normálvektor optimális becslése affin transzformációból, kalibrált kamerapár esetén [14,16,7,28,31].

Az I.1 tézis szerinti első törvény alapján Baráth Dániel doktoranduszommal közösen az alábbiak szerint definiáltuk az optimális normálvektorbecslés problémáját legkisebb négyzetes értelemben:

$$\mathbf{n} = \arg_{\mathbf{n}} \min \sum_{i=1}^2 \sum_{j=1}^2 \left(\frac{\mathbf{n}^T \mathbf{w}_{ij}}{\mathbf{n}^T \mathbf{w}_c} - a_{ij} \right)^2, \quad (9)$$

ahol

$$\begin{aligned} \mathbf{w}_{11} &= (\nabla \Pi_v^1 \times \nabla \Pi_u^2), & \mathbf{w}_{12} &= (\nabla \Pi_u^2 \times \nabla \Pi_u^1), \\ \mathbf{w}_{21} &= (\nabla \Pi_v^1 \times \nabla \Pi_v^2), & \mathbf{w}_{22} &= (\nabla \Pi_v^2 \times \nabla \Pi_u^1), \\ \mathbf{w}_c &= (\nabla \Pi_u^1 \times \nabla \Pi_v^1). \end{aligned} \quad (10)$$

Megmutattuk, hogy ez a feladat legkisebb négyzetes értelemben optimálisan megoldható. Három eljárást is kidolgoztunk, a második és a harmadik eljárás esetén kutatócsoportunk kollégáinak segítségét is igénybe vettünk:

- Negyedfokú polinom segítségével.
- Harmadfokú polinom segítségével.⁴
- Lineáris összefüggések hányadosaként.⁵

Az újabb módszerek egyre gyorsabbak, a futási idő csökkentése motiválta az újabb eljárások kifejlesztését.

II.2. tézis. Lyukkamera vetítő mátrixának becslése affin transzformációból [15].

Eichhardt Iván kollégámmal, kutatócsoportunk oszlopos tagjával közösen megmutattuk, hogy az I.1 tézis szerinti első törvénye alapján – amelyet a 6. összefüggés ír le –, ha ismerjük az affin transzformációt, a felületi pontokat és normálvektorokat, továbbá a vetületeket a képeken, akkor lyukkamerára a projektív mátrix paramétereit meg tudjuk határozni az egyik kamerára, ha a másik kamerának belső paramétereit ismerjük. Erre alapozva kidolgoztunk egy algoritmust projektív mátrix becslésére.

A kalibrációs algoritmus működőképességét kamera-projektor páros kalibrációján keresztül mutattuk meg. A projektor maga inverz kamerának fogható fel, amely nem képet (leképzést) készít a háromdimenziós világról, hanem a síkbeli mintát vetíti ki a térbe.

II.3. tézis. Homográfia becslése affin transzformációkból [4,16].

Az I.1. altézis második törvényszerűsége négy egyenletet ír fel, amely összeköti a homográfiát és a lokális affin transzformációkat lyukkamera modell alkalmazása esetén. Az összefüggésekből Baráth Dániel doktorandusszal két becslő algoritmust készítettünk:

1. **HA algoritmus.** Általános esetben, ha a kamerák paramétereit teljesen ismeretlenek, két affin megfelelésből ki tudjuk számolni a kapcsolódó homográfiát; szemben a hagyományos, kizárólag pontmegfeleléseket alkalmazó módszerrel [HZ03], amelynek legalább négy megfelelésre van szüksége.
2. **HAF algoritmus.** Ha a két kép között az alapvető (fundamentális) mátrix ismert, egyetlen affin transzformáció is elég a becsléshez.

A javasolt módszerek univerzálisak abban az értelemben, hogy mind a minimális, mind a túlhatározott esetekben alkalmazhatóak.

II.4. tézis. Síkok klaszterezése lokális affin transzformációk segítségével, kalibrált kamerák esetén [18].

A II.3. altézisben kimondott HAF algoritmus segítségével kidolgoztunk Baráth Dániellel és Jiri Matas cseh kollégával közösen egy szegmentáló eljárást, amely az azonos síkhoz tartozó affin megfeleltetéseket csoportosítja, hiszen egyetlen megfelelésből a homográfiát

⁴A levezetés megalkotásában Nghia Le Minh BSc-s hallgató segítségét kell kiemelni, aki a formális bizonyítást végső alakját megalkotta.

⁵Itt Lóczi Lajos kollégám segítségét kell kiemelni, aki a Mathematica alkalmazását vetette be, és a leggyorsabb megoldót így találta meg.

meg tudja becsülni a HAF algoritmus, ha a két kép között az epipoláris geometriát leíró fundamentális mátrix ismert. A kiszámított homográfia alapján a hozzá tartozó térbeli síkot meg lehet határozni, és az azonos síkokat eredményező lokális affin transzformációk kiválasztása alapján a síkfelületet a képen szegmentálni lehet. Ez a módszer a HAF algoritmus kiterjesztésének tekinthető több (multi) modell illesztésére.

II.5. tézis. Lényegi és alapvető mátrix becslése lokális affin transzformáció alapján [5,20].

Az I.2-es tézis egy vektoros összefüggést ad, amelyik egy lokális affin transzformációt és a fundamentális mátrixot összeköti. Ennek az egyenletnek a felhasználásával doktorandusz halgatómmal közösen készítettünk egy homogén lineáris egyenletrendszert, amelyből legkisebb négyzetes értelemben optimális becselő eljárást alkottunk meg. Az affin összefüggésből két homogén egyenletet kapunk, ezt egészítettük ki a pontmegfelelésekből származó összefüggéssel [HZ03], ezért minden lokális affin transzformáció három egyenletet ad. Algoritmusunk ezért három megfelelésből, homogén lineáris egyenletrendszer megoldásán keresztül becsli meg a fundamentális mátrixot, szemben a hagyományos pontalapú algoritmussal [HZ03], amely legalább hét megfelelés meglétét feltételezi.

A javasolt algoritmusunk minimális és túlhatározott esetben egyaránt működőképes.

II.6. tézis. Félig kalibrált kamera kalibrációja két lokális affin transzformációból [20].

A II.5-ös altézis speciális esete, amikor a kamera félig kalibrált, azaz ismerjük a döféspontot, és csak a fókusztáv ismeretlen. Ekkor a kamera belső paraméterei közül csupán egy ismeretlen, így az alapvető mátrix összesen hat szabadságfokkal bír. Ezért két lokális affin transzformáció elégséges a megoldáshoz, hiszen ezek összesen hat egyenletet adnak. Ehhez javasoltunk egy algoritmust, amelyik a homogén lineáris egyenletrendszerből származó nulltérben (magtérben) geometriai és optikai törvényszerűségeket figyelembe véve megkeresi a feltételeknek legjobban megfelelő megoldást.

II.7. tézis. Síkban mozgás optimális becslése kalibrált kamerák esetén egyetlen lokális affin transzformációból [24].

Javasoltunk Baráth Dániellel doktorandusszal egy becselő algoritmust, amelyik síkban mozgás és teljesen kalibrált kamerák esetén a síkban mozgás két szabadságfokát képes megbecsülni. A két szabadságfok a függőleges elforgatás és a síkban elmozdulás iránya. Az elmozdulásnak a nagyságát a térbeli látás törvényszerűségei miatt nem lehet meghatározni, csak az irányára lehet becslést adni.

Az algoritmusunk egy algebrai hibát minimalizál, a pontmegfelelésekből egy, a lokális affin transzformációból további kettő egyenletet lehet felírni, így összesen három egyenlet áll rendelkezésre a két szabadságfokra, ezért a feladat már egy affin megfelelés teljes megléte esetén is túlhatározott. A lineáris egyenletrendszer megoldására legkisebb négyzetes értelemben optimális algoritmust javasolunk, amely tetszőleges számú affin megfelelés esetén működni képes. Az eljárás minimális ⁶ és a túlhatározott eseteket egyaránt kezelni képes.

II.8. tézis. Síkban mozgás becslése félig kalibrált kamerák esetén egyetlen lokális affin transzformációból [25].

⁶ Minimális esetet akkor kaphatunk, ha a három egyenletből az egyiket elhagyjuk, például a pontmegfeleléseket nem vesszük figyelembe, csak a két affin törvényszerűséget.

A II. 7. altézishez nagyon hasonló algoritmust dolgoztunk ki Baráth Dániellel közösen arra az esetre, ha a síkban mozgás esetén ugyanazzal a félig kalibrált kamerával veszünk fel két képet. Ekkor a két képen az egyetlen ismeretlen belső paraméter a fókusztávolság.

Megmutattuk, hogy az alapvető mátrixnak így három szabad paramétere van: a közös fókusztávolság, a két kép közötti függőleges elforgatás és az eltolás iránya. Javasoltunk egy lineáris megoldót, amely egy lokális affin transzformációból képes ezt a három paramétert megbecsülni. Az eljárás túlhatározott esetre is működik, amikor egynél több transzformációból nyerünk információt, de az optimalitása a módszernek ebben az esetben nem garantált.

II.9. tézis. Lokális affin transzformációk becslése egyenesekből [28].

Nghia Le Minh ELTE-s BSc hallgatómmal közös munkánkban azt mutattuk meg, hogy a képeken egymásnak megfeleltetett egyenesek irányait felhasználva is ki lehet nyerni az affin transzformációkat. Munkánkban megkülönböztettünk skálázott és skálázatlan egyeneseket. Utóbbi esetében az egyenesek iránya adott csupán, skálázott esetben a szakaszok hosszváltozatának lokális approximációja is ismert.

Ha az alapvető mátrix adott, akkor az eipoláris egyenesek normálvektoraira fel tudunk írni egy skálázott összefüggést, ahogyan azt például a 8. összefüggésben már meg is tettük, ezáltal egy inhomogén lineáris összefüggésünk van az affin transzformációra.

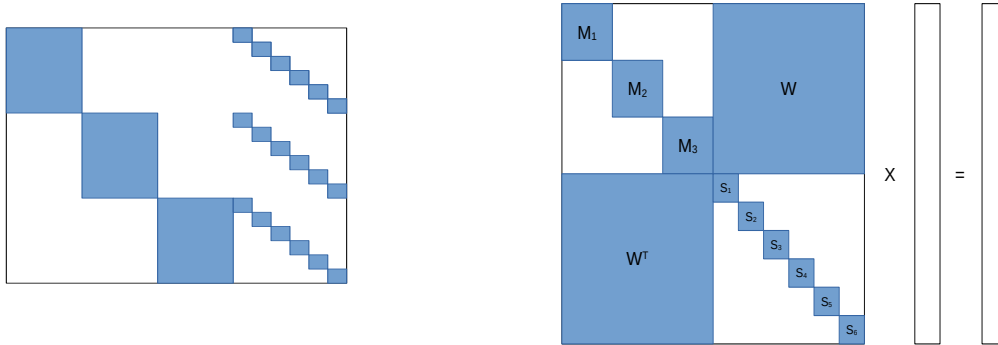
Ezeknek az affin transzformációknak a kinyeréséhez többféle algoritmust készítettünk. A becslési módszereink a lineáris részét képesek meghatározni az affin transzformációnak, az eltolást a pozíciók adhatják meg. Ha a skála nem ismert, a négy affin paraméterhez minden egyes ismeretlen skála egy becsültnivaló paramétert ad hozzá. A lineáris egyenletek ebben az esetben homogének az affin transzformációra nézve. Fontos feltétel, hogy legalább egy skálát ismerni kell, tehát egy inhomogén egyenletre szükség van a teljes becsléshez.

5. III. téziscsoport: Kötegelt behangolás újszerű alkalmazása nem hagyományos problémaosztályokra

A kötegelt behangolás olyan problémák esetén alkalmazható, amikor nagy mennyiségű adat viszonylag kevés paramétertől függ, és az adatok kétdimenziós mátrixba szervezhetőek. A 4. ábra mutat egy példát a probléma algebrai alakjára. Eredeti verzióját sokképes rekonstrukcióra fejlesztették ki [Bil00]. Amennyiben van N darab kameránk és K darab pontunk, minden egyes kamerához minimum hat, maximum tizenegy paramétert kell meghatározunk. Vegyük az első esetet, ami kalibrált kamerák esetén lehetségesen, azaz a belső paraméterek ismertek, és csak a külső paramétereket kell becsülni: három forgatási és három eltolási paraméter számítandó ebben az esetben. Az M darab térbeli pont értelem-szerűen három szabad paraméterrel, a három térbeli koordinátával rendelkezik. Összesen tehát $6N + 3M$ paraméter becslését jelenti a rekonstrukciós folyamat.

Ha minden pont minden képen látszik – ami csak elméletileg lehetséges, a gyakorlati esetekben takarások is előfordulnak –, összesen $2NM$ koordinátaértéket tudunk detektálni. A kettes szorzó onnan jön, hogy a képeken vízszintes és függőleges koordinátákat egyaránt le tudunk olvasni. Nagy kép- és pontszám esetén $2N \cdot M \gg (6N + 3M)$, azaz a feladat jelentősen túlhatározott.

Ha felírunk valamifajta hibafüggvényt a rekonstrukcióra, az i -edik kép j -edik pontja csak az i -edik kamera paramétereitől és a j -edik pont térbeli koordinátáitól függ. Az összes többi kamera és a térbeli pozíció nem befolyásolja a vetített pont helyét. Ezért, ha a hibafüggvényt az ismeretlen paraméterek szerint deriváltjuk, a laza függőség miatt sok zérus értéket kapunk. Más szóval, a Jacobi mátrix elég ritka lesz, ahogyan azt a 4. ábra bal oldali képén látjuk három kamera és hat pont rekonstrukciója esetén. Numerikus algoritmusok alkalmazása során a linearizált normálegyenlet is ritka lesz, az ábrának a jobb oldali képe ezt a tényt mutatja meg. A normálegyenletet invertálni kell a megoldáshoz. Általános esetben az invertálandó mátrix mérete a paraméterek számával egyenlő. A normálegyenlet speciális formáján jól látható, hogy a főátló mentén kisebb blokkok találhatóak. A teljes mátrix inverziója a kisebb blokkok inverziójának segítségével megadható, ahogyan azt Laurakis [LA04] munkája nagyon alaposan elmagyarázza.



(a) A Jacobi mátrix felépítése három kamerás és (b) A bal oldali feladathoz tartozó normálegyenlet pontos rekonstrukció esetében.

4. ábra. Numerikus optimalizálás kötegelt behangolás segítségével. A nemnulla elemek világoskékek az ábrán. Észrevehető, hogy a normálegyenletben a főátló mentén kisebb blokkok találhatóak, ezek invertálása gyorsan elvégezhető.

Tudományos munkám során sikerült megmutatni, hogy a kötegelt behangolás algoritmus más problémaosztályok esetén is sikerrel alkalmazható. Az elért tudományos eredményeimet az alábbi tézispontokba foglalom össze:

III.1. tézis. Kötegelt behangolás alkalmazása fotometrikus sztereóra [3].

Fodor Bálint doktorandusz hallgatómmal és Jankó Zsolt SZTAKI-s kollégámmal közösen sikerült egy újszerű eljárást készíteni kötegelt behangolás alkalmazásával, amely a fotometrikus sztereó problémát oldja meg pontszerű fényforrás esetén. Tetszőleges számú fényforrást alkalmazni tudunk, a bemenet egy képsorozat, ahol a kamera helyét a rekonstruálandó tárgyhoz képest nem változtatjuk meg. Minden megvilágítás egy képet eredményez a sorozatban. A módszer laboratóriumi körülményeket igényel abból a szempontból, hogy külső megvilágításból nem szűrődhet be fény. Minden felvételnél egyetlen megvilágítás engedélyezett.

A numerikus optimalizációnál a kezdeti paraméterek meghatározásához párhuzamos fényforrás alkalmazását javasoljuk. A rekonstrukció felírható paraméter-minimalizálásként, ahol az optimalizálandó paraméterek a fények és a felületi normálvektorok. A kötegelt behangolás alkalmazásával tetszőleges fényforrás alkalmazható, mi pontszerű fényforrás feltételezésével mutattuk meg a módszer létjogosultságát.

III.2. tézis. Kötegelt behangolás alkalmazása irányított pontok rekonstrukciójára, affin transzformációk alkalmazásával [21].

Eichhardt Iván kollégámmal, kutatócsoportunk tagjával közösen megmutattuk 2017-ben, hogy a hagyományos kötegelt behangolás algoritmus kiegészíthető az affin transzformációkkal, ezért Structure from Motion (SfM) eljárások esetén pontos rekonstrukciós eredmény kapható numerikus optimalizálással annak ellenére, hogy a finomítandó paraméterek száma igen nagy. A hagyományos pontalapú rekonstrukcióhoz képest minden térbeli pont kiegészíthető a felületi normálvektorral is. A normálvektornak a nagysága nem tartalmaz érdemi információt, csak az iránya, ezért a térbeli normálvektor két további paramétert jelent, azaz az irányított pont összesen öt paraméterrel írható le. Így a becslés kiegészül pontonként további ismeretlenekkel, ugyanakkor az egyes képmintákhoz tartozó affin transzformációkat a képekből meg lehet határozni, azaz a képeken a 2D koordinátákon túl az affin transzformációk is megjelennek ismert adatként.

6. IV. téziscsoport: Új módszerek LiDAR mérőeszközök és perspektív kamerák kalibrációjára

Az utolsó kutatási terület, melyet jelenleg is nagy intenzitással művelek, különböző modalitású szenzorok közös kalibrációja. Az elmúlt években LiDAR és digitális kamerák összekalibrálására koncentráltunk kutatócsoportomban, tudományos eredményeimet is ezen a területen értem el.

Tézispontjaimat az alábbiakban foglalom össze:

IV.1. tézis. Kamera-LiDAR kalibráció téglatest (doboz) segítségével [6,22].

Pusztai Zoltán doktoranduszommal közösen kidolgoztunk egy újszerű eljárást LiDAR letapogató eszköz és digitális kamera kalibrációjára, amely téglatestet (dobozt) alkalmaz kalibrációs objektumként, szemben a meglévő módszerekkel [GMCS12, VŠMH14], melyek a legtöbb esetben sakktábla segítségét veszik igénybe a kalibrációs feladat megoldásra.

A LiDAR pontfelhő feldolgozására javasoltunk egy félautomatikus eljárást, amely a mérőlegesség figyelembe vételével meghatározza a doboz csúcspontját és az éleit. Az eljárás először robusztus illesztéssel merőleges oldalakat keres, majd iteratív eljárással beforgatja és betolja a látható oldalakat a megfelelő pozícióba. Ismert dobozméret esetében így a doboz csúcsainak pozíciója egyértelműen megbecsülhető.

A kamera képén a doboz sarkait manuálisan kell meghatározni. A pontfelhőn becsült és a képen meghatározott pontok között egy Perspektív n-pont [LMNF09] (PnP) algoritmus adja meg a két eszköz közötti relatív eltolást és elforgatást.

Az eljárás könnyen kiterjeszthető tetszőleges számú kamera és LiDAR letapogató eszköz alkalmazására.

IV.2. tézis. Kamera-LiDAR kalibráció gömb segítségével [9,26].

Tóth Tekla doktoranduszom közreműködésével kidolgoztunk egy új módszert, amely gömb segítségével határozza meg digitális kamerák és LiDAR letapogató eszközök elhelyezkedését. Az eljárásunk teljesen automatikus.

Egy kamera képén egy gömb kontúrpointjai speciális ellipszist alkotnak, az ellipszis paramétereit meghatároztuk a gömb pozíciójának és sugarának függvényében. Megmutattuk, hogy a detektált ellipszisből hogyan lehet meghatározni a gömb középpontját a képen. Ismert sugár és kamera belső paraméterek esetében a gömb térbeli pozíciója is meghatározható.

A LiDAR pontfelhőn robusztus illesztés segítségével a középpont és a sugár szintén meghatározható, erre a feladatra a területen szabványosnak mondható eljárások robusztifikálását javasoljuk.

A kamerák és LiDAR-ok közötti elforgatás és eltolás megoldására pontregisztrációs eljárást alkalmazunk. Módszerünk tetszőleges számú kamerára és LiDAR eszközre működik, ha legalább négy gömbre el tudjuk a szükséges illesztési feladatokat végezni. Túlhatározott esetben, azaz négynél több felvétel esetén is működik az eljárásunk, a gömbök számának növelésével a kalibráció pontossága növekszik.

IV.3. tézis. Egy forgatási paraméterre redukált kamera-LiDAR kalibráció [30].

Tófalvi Tamás és Tóth Tekla doktorandusz hallgatókkal közösen egy harmadik eljárást is kidolgoztunk LiDAR-kamera kalibrációra, amely sakktáblát használ a külső paraméterek kiszámítására. A fő újdonság, hogy a kamerát és a LiDAR-t egy speciális, saját tervezésű és gyártású 3D-nyomtatott alkatrész segítségével egymáshoz rögzítjük ⁷, így a két eszköz relatív orientációja egyszerűsödik, mindössze egy forgatási szöget kell meghatározni.

Az általunk javasolt módszer sakktáblák síkjait vizsgálja, és ezek normálvektorait használja a helyes forgatási szög megbecslésére. A javasolt megoldó lineáris problémaként írható fel, a forgatásra így legkisebb négyzetes értelemben optimális megoldást tudunk adni, amelyik minimális és túlhatározott esetekben egyaránt működik.

A módszer egy képet igényli csak, ezért online kalibrálásra is használható, ha a jármű elhalad egy falra (síkra) rögzített sakktáblaminta mellett.

⁷ A rögzítés tervezése és 3D nyomtatása Kovács Bandó mérnökünk munkája.

A szerzőnek a PhD fokozat megszerzése óta megjelent fontosabb publikációi

Folyóiratcikkek

- [1] **Levente Hajder**, Ákos Pernek, Csaba Kazó. Weak-perspective structure from motion by fast alternation. *The Visual Computer*, 27(5), 387-399, 2011.
- [2] Ákos Pernek, **Levente Hajder**. Automatic focal length estimation as an eigenvalue problem. *Pattern Recognition Letters*, 34(9), 1108–1117, 2013.
- [3] Bálint Fodor, Csaba Kazó, Zsolt Jankó, **Levente Hajder**. Normal map recovery using bundle adjustment.. *IET Computer Vision*, 8(1), 66–75, 2014.
- [4] Dániel Baráth, **Levente Hajder**. A theory of point-wise homography estimation. *Pattern Recognition Letters*, 94, 7–14, 2017.
- [5] Dániel Baráth, **Levente Hajder**. Efficient Recovery of Essential Matrix From Two Affine Correspondences. *IEEE Transactions on Image Processing*, 27(11): 5328–5337, 2018.
- [6] Zoltán Pusztai, Ivan Eichhardt, **Levente Hajder**. Accurate Calibration of Multi-LiDAR-Multi-Camera Systems. *Sensors*, 18(7), 2139, 2018.
- [7] Dániel Baráth, Ivan Eichhardt, **Levente Hajder**. Optimal Multi-View Surface Normal Estimation Using Affine Correspondences. *IEEE Transactions on Image Processing*, 28(7), 3301-3311, 2019.
- [8] Gábor Baranyi, Bruno Carlos Dos Santos Melício, Zsófia Gaál, **Levente Hajder**, András Simonyi, Dániel Sindely, Joul Skaf, Ondrej Dusek, Tomás Nekkunda, András Lorincz. AI Technologies for Machine Supervision and Help in a Rehabilitation Scenario. In *Multimodal Technologies and Interaction*, 6(7): 48 (2022)
- [9] Tekla Tóth, **Levente Hajder**. A Minimal Solution for Image-Based Sphere Estimation. *International Journal of Computer Vision*, 131(6), 1428-1447, 2023.

Konferenciakiadványok

- [10] Ákos Pernek, **Levente Hajder**, Csaba Kazó. Metric Reconstruction with Missing Data under Weak Perspective. In *Proceedings of the British Machine Vision Conference*, 1–10, 2008.
- [11] Csaba Kazó, **Levente Hajder**. Rapid weak-perspective Structure from Motion with missing data. In *ICCV Workshops*, 491–498, 2011.
- [12] Ákos Pernek, **Levente Hajder**. Precise 3D Pose Estimation of Human Faces.. In *VISAPP (3)*, 618–625 , 2014.
- [13] Dániel Baráth, József Molnár, **Levente Hajder**. Optimal Surface Normal from Affine Transformation. In *VISAPP (3)*, 305–316, 2015.
- [14] Dániel Baráth, József Molnár, **Levente Hajder**. Novel Methods for Estimating Surface Normals from Affine Transformations. In *VISIGRAPP (Revised Selected Papers)*, Springer, 316–337, 2015.
- [15] Ivan Eichhardt, **Levente Hajder**. Improvement of camera calibration using surface normals. In *International Conference on Pattern Recognition*, 3745–3750, 2016.
- [16] Dániel Baráth, József Molnár, **Levente Hajder**. Novel Ways to Estimate Homography from Local Affine Transformations. In . *VISIGRAPP (3: VISAPP)*, 434–445, 2016.
- [17] Dániel Baráth, Jiri Matas, **Levente Hajder**. Accurate Closed-form Estimation of Local Affine Transformations Consistent with the Epipolar Geometry. In *British Machine Vision Conference*, 2016.
- [18] Dániel Baráth, Jiri Matas, **Levente Hajder**. Multi-H: Efficient recovery of tangent planes in stereo images.. In *British Machine Vision Conference*, 2016.
- [19] **Levente Hajder**. Weak-Perspective and Scaled-Orthographic Structure from Motion with Missing Data. In *VISIGRAPP (Revised Selected Papers)*, Springer, 128–153 , 2017.
- [20] Dániel Baráth, Tekla Tóth, **Levente Hajder**. A Minimal Solution for Two-View Focal-Length Estimation Using Two Affine Correspondences. In *IEEE International Conference on Computer Vision*, 2557-2565 , 2017.

- [21] **Levente Hajder**, Ivan Eichhardt Computer Vision Meets Geometric Modeling: Multi-view Reconstruction of Surface Points and Normals Using Affine Correspondences. In *ICCV Workshops*, 2427-2435, 2017.
- [22] Zoltán Pusztai, **Levente Hajder**. Accurate Calibration of LiDAR-Camera Systems Using Ordinary Boxes. In *ICCV Workshops*, 394-402, 2017.
- [23] Tekla Tóth, **Levente Hajder**. Robust Fitting of Geometric Primitives on LiDAR Data. In *VISAPP*, 622..629, 2019.
- [24] **Levente Hajder**, Dániel Baráth. Least-squares Optimal Relative Planar Motion for Vehicle-mounted Cameras. In *International Conference on Robotics and Automation*, 8644-8650 , 2020.
- [25] **Levente Hajder**, Dániel Baráth. Relative planar motion for vehicle-mounted cameras from a single affine correspondence. In *International Conference on Robotics and Automation*, 8651-8657 , 2020.
- [26] Tekla Tóth, Zoltán Pusztai, **Levente Hajder**. Automatic LiDAR-Camera Calibration of Extrinsic Parameters Using a Spherical Target. In *International Conference on Robotics and Automation*, 8580-8586 , 2020.
- [27] **Levente Hajder**, Tekla Tóth, Zoltán Pusztai:. Automatic Estimation of Sphere Centers from Images of Calibrated Cameras. In *VISAPP (4)*, 490–497, 2020.
- [28] Nghia Le Minh, **Levente Hajder**. Affine Transformation from Fundamental Matrix and Two Directions. In *VISAPP (4)*, 819–826, 2020.
- [29] István Gergő Gál, Dániel Baráth, **Levente Hajder**. Pose Estimation for Vehicle-mounted Cameras via Horizontal and Vertical Planes. In *International Conference on Robotics and Automation*, 7620-7626 , 2021.
- [30] Tamás Tófalvi, Bandó Kovács, **Levente Hajder**, Tekla Tóth. LiDAR-camera Calibration in an Uniaxial 1-DoF Sensor System. In *VISIGRAPP (4: VISAPP) 2022*: 730-738
- [31] **Levente Hajder**, Lajos Lóczi, Daniel Barath. Fast Globally Optimal Surface Normal from an Affine Correspondence. In *International Conference on Computer Vision 2023*: 3367-3378

Egyéb szakirodalmi hivatkozások

- [Bil00] Bill Triggs and Paul McLauchlan and Richard Hartley and Adrew Fitzgibbon. Bundle Adjustment – A Modern Synthesis. In W. Triggs, A. Zisserman, and R. Szeliski, editors, *Vision Algorithms: Theory and Practice*, LNCS, pages 298–375. Springer Verlag, 2000.
- [Bir97] S. Birchfield. Derivation of Kanade-Lucas-Tomasi Tracking Equation. <http://www.ces.clemson.edu/stb/klt/birchfield-klt-derivation.pdf>, 1997.
- [BM02] Simon Baker and Iain Matthews. Lucas-Kanade 20 Years On: A Unifying Framework: Part 1. Technical Report CMU-RI-TR-02-16, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, July 2002.
- [GD00] Christopher Geyer and Konstantinos Daniilidis. A unifying theory for central panoramic systems and practical applications. In David Vernon, editor, *Computer Vision - ECCV 2000, 6th European Conference on Computer Vision, Dublin, Ireland, June 26 - July 1, 2000, Proceedings, Part II*, volume 1843 of *Lecture Notes in Computer Science*, pages 445–461. Springer, 2000.
- [GMCS12] Andreas Geiger, Frank Moosmann, Omer Car, and Bernhard Schuster. Automatic camera and range sensor calibration using a single shot. In *IEEE International Conference on Robotics and Automation, ICRA 2012, 14-18 May, 2012, St. Paul, Minnesota, USA*, pages 3936–3943. IEEE, 2012.
- [Har88] Harris, C. and Stephens, M. A combined corner and edge detector. In *Fourth Alvey Vision Conference*, pages 147–151, 1988.
- [Har95] R. I. Hartley. In defence of the 8-point algorithm. *International Conference on Computer Vision*, pages 1064–1070, 1995.
- [HS97] R. I. Hartley and P. Sturm. Triangulation. *Computer Vision and Image Understanding: CVIU*, 68(2):146–157, 1997.
- [HZ03] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.
- [LA04] M. Lourakis and A. Argyros. The design and implementation of a generic sparse bundle adjustment software package based on the Levenberg–Marquardt algorithm. *Institute of Computer Science - FORTH, Tech. Rep. FORTH-ICS TR-340-2004*, 2004.
- [LMNF09] Vincent Lepetit, Francesc Moreno-Noguer, and Pascal Fua. Epnp: An accurate $\mathcal{O}(n)$ solution to the pnp problem. *International Journal Of Computer Vision*, 81:155–166, 2009.

- [MC14] József Molnár and Dmitry Chetverikov. Quadratic transformation for planar mapping of implicit surfaces. *J. Math. Imaging Vis.*, 48(1):176–184, 2014.
- [Nis03] D. Nister. An efficient solution to the five-point relative pose problem. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2:195, 2003.
- [SHB07] Milan Sonka, Vaclav Hlavac, and Roger Boyle. *Image Processing, Analysis, and Machine Vision*. CENGAGE-Engineering, third edition edition, 2007.
- [SMS06] Davide Scaramuzza, Agostino Martinelli, and Roland Siegwart. A flexible technique for accurate omnidirectional camera calibration and structure from motion. In *2006 IEEE International Conference on Computer Vision Systems, January 5-7, 2006, St. Johns University, Manhattan, New York City, New York, NY, USA, Proceedings, CDROM*, page 45. IEEE Computer Society, 2006.
- [VŠMH14] Martin Velás, Michal Španěl, Zdeněk Materna, and Adam Herout. Calibration of rgb camera with velodyne lidar. In *WSCG 2014 Communication Papers Proceedings*, pages 135–144, 2014.
- [Zha00] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.