

Kameramozgás becslése járműre szerelt kamerák esetén függőleges és vízszintes síkok használatával

Gál István Gergő¹, Baráth Dániel², Hajder Levente¹

¹ Eötvös Loránd Tudományegyetem, Budapest
{gallv5, hajder}@inf.elte.hu

² VRG, Department of Cybernetics, Czech Technical University in Prague; MPLab, SZTAKI, Budapest; and Department of Computer Science, ETH Zurich
barath.daniel@sztaki.mta.hu

Absztrakt. A cikk mozgó járműre szerelt kalibrált kamerák közötti relatív elmozdulás becslésére mutat be új módszert, csupán egyetlen affin megfeleltetés felhasználásával. A módszerek speciális, a gyakorlatban sűrűn előforduló síkokat feltételeznek, melyekhez a kapcsolódó homográfiák is speciálisak. A bemutatott algoritmusokat négy csoportba soroljuk. Az első három csoport esetén azt feltételezzük, hogy a keresett sík merőleges a kamera egyik tengelyére. A negyedik csoport esetében a sík a talajra merőleges, de ismeretlen normálvektorral rendelkezik. Ilyen eset például az épületek homlokzata. Valamennyi módszert lineáris rendszerrel oldunk meg, kis együtthatómátrixszal, így rendkívül gyorsak és hatékonyak. A minimális és a túlhatározott eseteket egyaránt megoldhatók a javasolt módszerekkel. Kiértékelésünket szintetikus adatokon és nyilvánosan elérhető valós adathalmazokon végeztük. Az összes tesztelt algoritmust magunk implementáltuk. A vizsgálataink azt mutatják, hogy az új eljárások pontosabbak, ugyanakkor gyorsabbak is a konkurens módszerekhez képest, ha a legkorszerűbb robusztus becslő keretrendszerekbe foglalják őket. A forráskód nyilvánosan elérhető [18].

1. Bevezetés

A sík és sík közötti egyenestartó megfeleltetések (azaz homográfiáinak) becslése képpárok között alapvető feladat a térbeli jelenet (színtér) geometriájának helyreállításában. Napjainkban a legkorszerűbb (state of the art - SOTA) módszerek a struktúrát a mozgásból (Structure from Motion) [36, 35, 34] vagy a szimultán lokalizáció és leképezés (Simultaneous Localization and Mapping - SLAM) [14, 2] algoritmusok robusztusak lesznek olyan esetekben, amikor a mozgás közel síkmozgás vagy csak forgást tartalmaz. A robusztusságot az epipoláris geometria és a homográfiai becslés kombinálásával érik el. Ebben a cikkben nem hagyományos bemeneti adatokat (affin megfeleltetéseket)

*A cikk a 2021 IEEE International Conference on Robotics and Automation (2021 ICRA) konferenciára került elfogadásra Pose Estimation for Vehicle-mounted Cameras via Horizontal and Vertical Planes néven. A magyar verzió tartalmában megegyező az elfogadott cikkel, azonban néhány helyen bővebb kifejtést tartalmaz.

használunk, és arra az esetre összpontosítunk, amikor a kamerát a mozgó járműre szerelik, és előzetes ismeretek állnak rendelkezésre a keresett síkról, például, hogy képpáron megtalálható a talaj (úttest) vagy függőleges síkok. Ilyen függőleges síkok lehetnek az épületek falai vagy a nagyobb járművek oldala.

Egy affin megfeleltetés (affine correspondence – AC) egy pontpárt és a hozzá tartozó 2×2 -es lokális affin transzformációt tartalmaz. Az affin transzformációk olyan transzformációk, amelyek az egyik képen a képpont körül található végtelenül kicsi környezetet a másik képen a képpontnak megfelelő végtelenül kicsi környezetbe transzformálják. Az 2D-s affin transzformációk négy komponense a forgatás, a két irány menti skálázás, valamint a nyírás. Fontos tulajdonsága az affin transzformációnak, hogy a pontokon átfutó egyenesek irányát helyesen transzformálja. A kétdimenzió azt jelenti, hogy a képtérben van értelmezve. Azért csak 2×2 -es transzformációval dolgozunk, mert az eltolás paramétert az affin transzformációból nem használjuk. Napjainkban számos algoritmus létezik [28, 20, 10, 6, 31, 30, 8, 4, 29, 5], amely affin megfeleltetéseket használ geometriai entitások becslésére, mint például homográfia, felületi normális vagy epipoláris geometria.

Ezeket a módszereket részletesen tárgyalják cikkünkben Baráth és mtsai. [7]. Bento-lila és Francos [10] munkájukban javasoltak egy módszert fundamentális mátrix becslésére három affin jellemzőből. Raposo és mtsai. [31, 30], valamint Barath és mtsai. [5] megmutatták, hogy két megfeleltetés elegendő a relatív kamera mozgás becslésére.

Továbbá kettő jellemző pár elegendő a részben kalibrált kamera esetének megoldására, például amikor a cél az esszenciális mátrix és egy közös, de ismeretlen fókusz távolság meghatározása [8].

Ezenfelül homográfia megbecsülhető két affin megfeleltetésből [20], valamint ismert epipoláris geometria esetén egyetlen megfeleltetés elegendő [4]. A lokális affin transzformációk és a felület normálisai megfeleltethetők egymásnak kalibrált kamerák esetén [20, 6]. Pritts és mtsai. [29] megmutatták, hogy a lencse torzítás paraméterei szintén kinyerhetők.

Az affin jellemzők magasabb fokú információt hordoznak magukban a jelenet geometriájáról, ezért az ezeket felhasználó algoritmusok a becslési feladatokat kevesebb megfeleltetésekből tudják elvégezni, mint a hagyományos pontmegfeleltetés alapú módszerek. A felhasznált megfeleltetések számának csökkentése kiemelt fontosságú a robosztus becslések során, például a RANSAC [15] módszer esetében, ahol a futási idő *exponenciálisan* függ a szükséges pontok számától.

Az algoritmusokhoz szükséges affin megfeleltetések könnyen kinyerhetők a képpárokból hagyományos affin-kovariáns jellemző detektorok (affine-covariant feature detectors) segítségével, melyeket részletesen tárgyalnak cikkünkben Mikolajczyk és mtsai. [23] Ilyen detektorok például a jól ismert MSER [22], Hessian-Affine, Harris-Affine, stb. Másik módszer lehet az affin jellemzők előállítására a nézet-szintetizálás (view - synthesizing), amelyet az Affine-SIFT [26] vagy MODS [24] eljárás használ. Ezek az affin teret, valamint az affin módon transzformálódó bemeneti képeket mintavételezik. Továbbá léteznek modern tanulás alapú megközelítések is, például a Hes-Aff-Net [25], ami az affin régiókat CNN-alapú Hessian kulcspontokon történő regresszió futtatásával nyeri ki.

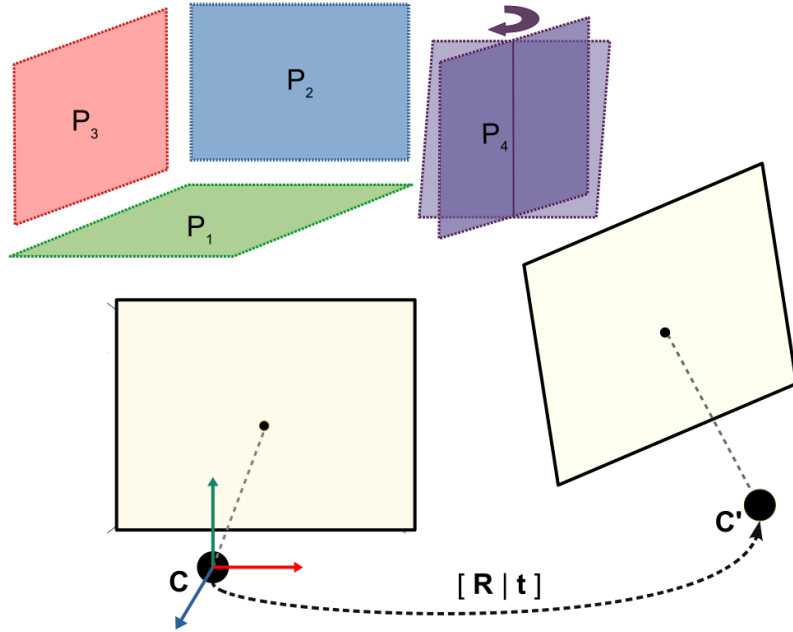
Az utóbbi időben az önvezető autózás egyre nagyobb figyelmet kap, így kifejezetten fontossá válik olyan algoritmusok létrehozása, melyek magukban foglalják az önvezető autó mozgását meghatározó megkötéseket és ezért jobb eredményt adnak az általános megoldó módszerekénél. Jól ismert megközelítés annak a figyelembevételére, hogy a kamerák egy síkon mozognak, például egy járműre vannak felszerelve. Így megközelítve a mozgás szabadsági foka háromra csökkenthető, ezért nagyban felgyorsítható a robotusztus becslés. Érdemes megjegyezni azonban, hogy ez a megkötés akkor tehető meg, ha ismert a függőleges irány, például IMU szenzor adataiból.

Ortin és Montiel [27] bebizonyította, hogy síkmozgás esetén az epipoláris geometria megbecsülhető két pontmegfeleltetés felhasználásával. Ez a modell, amit cikkünkben mi is feltételezünk. Ortin és Montiel munkája óta számos megoldó módszerre tettek javaslatot, amely két pontmegfeleltetés felhasználásával képes a relatív kamera póz becslésére [13, 12]. Továbbá Scaramuzza [33] javaslatot tett egy módszerre, amelyhez elegendő egyetlen pontpár felhasználása olyan esetekben, ahol a kamerára egy speciális, úgynevezett nem-holonomikus kényszer teljesül.

A cikkünk célja relatív kamera póz becslése speciális függőleges és vízszintes síkok felhasználásával síkmozgás feltételezése mellett. A bemutatott módszerekhez legjobban kapcsolódó munka Saurer és mtsai. [32] publikációja. Ők a homográfia mátrixokat becsülik meg pontmegfeleltetések felhasználásával azzal a feltételezéssel élve, hogy előre ismert a keresett síkok normális vektora. Azokat az eseteket vizsgálják, például amikor a normális vektor merőleges vagy párhuzamos azzal a síkkal amelyen a jármű, tipikusan egy autó vagy kvadrokopter mozog. Ehhez képest ebben cikkben az affin megfeleltetésekből származó többletinformáció is felhasználásra kerül.

Relatív póz becslésére más megközelítések is megtalálhatók. Banglei Guan és mtsai. [17] a cikkükben bemutatnak egy kettő pontnál kevesebbet használó algoritmust, abban az esetben, ha ismert a gravitációs vektor és rendelkezésre állnak a kamerától távol elhelyezkedő pontmegfeleltetések. Cikkünkötől eltérően alapvetően más megközelítést alkalmaznak és szétválasztják a forgatás és az eltolás vektor becslését két lépésbe. Egy távoli pontmegfeleltetés x koordinátáját (melyet csak kamerák közötti forgás határoz meg) és a gravitációs vektor ismeretét felhasználva első lépésként megbecsülik a forgatás vektor szabad paraméterét. Majd ezt követően egy újabb lépésben 1.5 pontmegfeleltetést felhasználva az eltolás vektor 3 paraméterét. Ettől eltérően mi egyetlen affin megfeleltetést használva és síkmozgást, valamint speciális síkok meglétét feltételezve határozzuk meg az ismeretlen forgatást és eltolást egy lépésben. Továbbá lehetőség van abszolút és relatív póz becslésére több kamera rendszerek esetén 2 és 3D egyenes párok, valamint ismert függőleges irány felhasználásával. [1]

Hozzájárulásaink. A cikkben javaslatot teszünk speciális homográfia becslésére *egyetlen affin megfeleltetés* felhasználásával. A tárgyalt homográfiahoz tartozó speciális síkok az 1. ábrán láthatóak. Az első típusú megoldók esetén azt feltételezzük, hogy a sík merőleges a kamera egyik tengelyére. A második típus esetén a síkok függőlegesek (mint például épületek falai) és egy ismeretlen normál vektorral rendelkeznek, mely egyetlen szöggel reprezentálható. Valamennyi módszert lineáris rendszerrel oldunk meg, kis együtthatómátrixszal, így rendkívül gyorsak, nagyságrendileg például 5–10 μ s futási idővel rendelkeznek. Mind a minimális és a túlhatározott eseteket egyaránt megoldhatók a javasolt módszerekkel. Kiértékelésünk szintetikus adatokon és nyilvánosan elér-



1. ábra: Az ábrán látható C és C' kamerák közötti kapcsolat leírható egy vertikális forgatás mátrixsal (R) és egy eltolás vektorral ($t = [t_x, 0, t_z]^T$), melynek második (vertikális) koordinátája nulla. Négy különböző eset látható, amikor a pontok (1) vízszintes P_1 , (2) szemben lévő függőleges P_2 , (3) oldalsó függőleges P_3 és (4) általános függőleges P_4 síkokból származnak.

hető valós adathalmazokon végeztük. A vizsgálataink az mutatják, hogy az új eljárások pontosabbak, vagy legalábbis összemérhető pontosságúak a hagyományos algoritmusokkal, és gyorsabbak, ha a mi algoritmusainkat a legkorszerűbb (State of The Art – SOTA) robusztus becslőkbe illesztjük.

2. Elméleti alapok

Legyen adott két kalibrált kamera és legyenek a kamerák kalibrációs mátrixai K és K' . Figyeljünk meg egy síkobjektumot! Fixáljuk le a világ koordinátarendszerét az első kamerához. A projekciós mátrixok legyenek $P = K[I | 0]$, $P' = K'[R | t]$, ahol az R mátrix, valamint a t eltolás vektor a nézetek között lévő forgatásmátrix és eltolás vektor, azaz a sztereó póz.

Ha a két kép között egymásnak megfeleltethető pontok találhatók, melyek homogén koordinátákban $u = [x \ y \ 1]$ és $u' = [x' \ y' \ 1]$ alakban adóttak, akkor a pontok koordinátái közötti kapcsolat lineáris jellegű a projektív térben. Ez a kapcsolat kifejezhető egy H homográfiával úgy, hogy $u' \sim Hu$, ahol a $\sim$ operátor jelöli az egyenlőséget egy általában ismeretlen skalár szorzó erejéig.

Kalibrált kamerák esetén, a kétdimenziós koordináták normalizálhatók a kamerák belső paramétereit tartalmazó mátrixok inverzével. A cikk további részében normalizált koordinátákat használunk: $\mathbf{u} \leftarrow \mathbf{K}^{-1}\mathbf{u}$ és $\mathbf{u}' \leftarrow \mathbf{K}'^{-1}\mathbf{u}'$.

A homográfia paraméterei kifejezhetők a kamerák közötti relatív póz segítségével [19]. A forgatás mátrixszal és eltolás vektorral a homográfia a következőképpen írható fel:

$$\mathbf{H} \propto \mathbf{R} - \frac{1}{d} \mathbf{t} \mathbf{n}^T, \quad (1)$$

ahol a skaláris d fejezi ki a távolságot az első kamera és a megfigyelt sík között, az $\mathbf{n}$ a sík normál vektora.

2.1. Síkmozgás

Tegyük fel, hogy adott egy kalibrált képpár egy közös XZ síkkal ($Y = 0$), ahol az Y tengely párhuzamos a képsíkok függőleges irányával. Az autonóm autóra szerelt kamera egy triviális példa ilyen megkötésre, ha jól rögzítik a kamerát. Ebben az esetben a kamera Y tengelye merőleges a talajra. Érdemes megjegyezni, hogy ez a megszorítás egyszerűen kielégíthető abban az esetben, ha a függőleges irány ismert. Tipikusan IMU szenzor segítségével szokták előállítani a függőleges, azaz gravitációs vektort. A kamera mozgás becsléséhez első lépésként leírjuk a probléma paraméterezését.

Síkmozgás feltételezve a forgatás és eltolás transzformációk három paraméterrel kifejezhetők: egy kétdimenziós eltolással és a forgatás szögével. Formálisan:

$$\mathbf{R} = \begin{bmatrix} \cos \alpha & 0 & -\sin \alpha \\ 0 & 1 & 0 \\ \sin \alpha & 0 & \cos \alpha \end{bmatrix}, \quad \mathbf{t} = \rho \begin{bmatrix} \cos \beta \\ 0 \\ \sin \beta \end{bmatrix}. \quad (2)$$

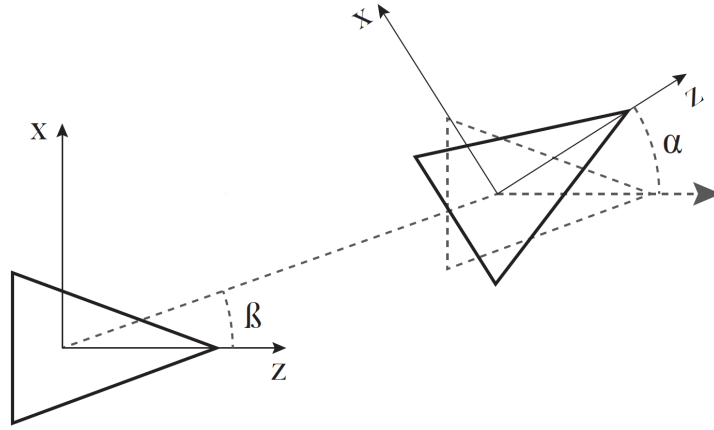
Az eltolás vektor kifejezhető egy $\beta \in [0, 2\pi)$ szöggel és a $\rho \in \mathbf{R}^+$ hosszával. Az $\alpha \in [0, 2\pi)$ az Y tengely körüli forgatás szöge. A mozgás szemléltetése a 2. ábrán látható.

2.2. Homográfia becslés

Ebben a cikkben a homográfia és a lokális affin transzformációk közötti kapcsolatot használjuk ki. Egy $\mathbf{H}$ homográfia reprezentálható egy 3×3 -as mátrixszal, míg egy affin transzformáció egy 2×2 -essel. A mátrixokat a továbbiakban a következőképpen indexeljük:

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}. \quad (3)$$

Egy képpáron két egymásnak megfeleltetett síkfelületen található területek között a $\mathbf{H}$ homográfia reprezentálja a projektív transzformációt. Az affin transzformáció egy első rendű közelítése a két kép egymásba való transzformálásának [6]. A homográfiaikat általában pontmegfeleltetések felhasználásával becsülik meg. Ha egy pont helyzetét



2. ábra: Síkmozgás paraméterei. A kamera mozgás leírható az Y tengely körüli forgatás α szögével és a $[\cos(\beta), 0, \sin(\beta)]^T$ eltolás vektorral. Az eltolás hossza nincs figyelembe véve, mivel abszolút távolságok nem nyerhetők ki a képpárokból, csak arányok.

$[x \ y]^T$ és $[x' \ y']^T$ koordinátákkal jelöljük az első és második képen, a koordinátáik közötti összefüggés [19] a két nézet között a következőképpen fejezhető ki:

$$\begin{aligned} x' (h_7 x + h_8 y + h_9) &= h_1 x + h_2 y + h_3, \\ y' (h_7 x + h_8 y + h_9) &= h_4 x + h_5 y + h_6. \end{aligned} \quad (4)$$

Következésképpen minden egyes pontmegfeleltetés (angolul Point Correspondence – PC) két darab egyenletet ad a homográfia becslésére.

Barath és Hajder [4] bizonyította cikkükben, hogy az affin része egy affin megfeleltetésnek további négy egyenletet ad, melyek a következők:

$$\begin{aligned} h_1 - (x' + a_1 x) h_7 - a_1 y h_8 - a_1 h_9 &= 0, \\ h_2 - (x' + a_2 y) h_8 - a_2 x h_7 - a_2 h_9 &= 0, \\ h_4 - (y' + a_3 x) h_7 - a_3 y h_8 - a_3 h_9 &= 0, \\ h_5 - (y' + a_4 y) h_8 - a_4 x h_7 - a_4 h_9 &= 0. \end{aligned} \quad (5)$$

Összesítve egy affin megfeleltetés (AC), ami a pontmegfeleltetéseket is magában foglalja, hat független megszorítás ad. Így egy AC és egy további PC elegendő egy általános nyolc szabadsági fokkal rendelkező homográfia becslésére.

3. Javasolt módszerek

Cikkünkben a következő probléma csoportokat tárgyaljuk: homográfai becslése (i) talajsíknál, (ii,iii) speciális függőleges síkok és (iv) általános függőleges síkok esetén. A fő célunk a kamera póz becslése a homográfákat felhasználva.

3.1. Talajsík

A talaj normál vektora $\mathbf{n} = [0 \ 1 \ 0]^T$. Síkmozgás esetén az 2. egyenletet és az $\mathbf{n}$ normál vektort az 1. egyenletbe helyettesítve a homográfia mátrix a következő alakban írható fel:

$$\gamma \mathbf{H} = \begin{bmatrix} \cos \alpha & 0 & -\sin \alpha \\ 0 & 1 & 0 \\ \sin \alpha & 0 & \cos \alpha \end{bmatrix} - \rho \begin{bmatrix} \cos \beta \\ 0 \\ \sin \beta \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}^T, \quad (6)$$

ahol a γ paraméter az ismeretlen faktor. A homográfia kilenc elemét kifejezve a következőt kapjuk:

$$\gamma \mathbf{H} = \begin{bmatrix} \cos \alpha & p & -\sin \alpha \\ 0 & 1 & 0 \\ \sin \alpha & q & \cos \alpha \end{bmatrix}, \quad \mathbf{H} \sim \begin{bmatrix} h_1 & h_2 & h_3 \\ 0 & h_5 & 0 \\ h_7 & h_8 & h_9 \end{bmatrix}, \quad (7)$$

ahol $p = -\rho \cos \beta$ és $q = -\rho \sin \beta$, valamint a $\mathbf{H}$ mátrix elemei a következők: $h_1 = h_9 = \cos \alpha$, $h_2 = p$, $h_3 = -h_7 = -\sin \alpha$ és $h_8 = q$, $h_5 = 1$. Ennek megfelelően három szabadsági fok becslésére van szükség, melyek az ismeretlen forgás α szöge és a 2D eltolás, melyet a p és q koordináták reprezentálnak.

Ha a homográfia, valamint a pont és affin paraméterek közötti kapcsolatot figyelembe vesszük, a 4. és 5. egyenletek alapján a becslési feladat linearizálható:

$$\mathbf{A}_{gd} \begin{bmatrix} \cos \alpha \\ \sin \alpha \\ p \\ q \end{bmatrix} = [0 \ -y \ 0 \ 0 \ 0 \ -1]^T, \quad (8)$$

ahol

$$\mathbf{A}_{gd} = \begin{bmatrix} x - x' & -x'x - 1 & y & -x'y \\ -y' & -y'x & 0 & -y'y \\ 1 - a_1 & -x' - a_1x & 0 & -a_1y \\ -a_2 & -a_2x & 1 - x' - a_2y & \\ -a_3 & -y' - a_3x & 0 & -a_3y \\ -a_4 & -a_4x & 0 & -y' - a_4y \end{bmatrix}. \quad (9)$$

A pont és affin paraméterek hat egyenletből álló $\mathbf{A}_{gd} \mathbf{h}_{gd} = \mathbf{b}_{gd}$ formájú lineáris egyenletrendszert adnak. A megoldásvektor $\mathbf{x} = [x_1 \ x_2 \ x_3 \ x_4]^T$ alakú, ahol $x_1 = \cos \alpha$, $x_2 = \sin \alpha$, $x_3 = p$, valamint $x_4 = q$. Az egyenletrendszer megoldására két megközelítés alkalmazható.

Optimális megoldó. A cél egy inhomogén lineáris egyenletrendszer megoldása az $x_1^2 + x_2^2 = 1$ megkötéssel, ahol $x_1 = \cos \alpha$ és $x_2 = \sin \alpha$ az első két koordinátája az $\mathbf{x}$ vektornak. Az algebrai probléma optimálisan megoldható legkisebb négyzetek értelemben két kúp metszetének felhasználásával.

Gyors megoldó. A probléma megoldható továbbá homogén lineáris egyenletrendszerként is:

$$[\mathbf{A}_{gd} \mid -\mathbf{b}_{gd}] [\mathbf{x}^T \ 1]^T = \mathbf{0}. \quad (10)$$

Az $[\mathbf{A}_{gd} | -\mathbf{b}_{gd}]$ mátrix nullvektora megadja a szuboptimális megoldást. Az $x_1^2 + x_2^2 = 1$ megkötést a kapott megoldásvektor utolsó komponensének leosztásával garantálhatjuk. Ezt követően a szög megkapható:

$$\alpha = \text{atan2}(x_2, x_1). \quad (11)$$

3.2. Speciális függőleges síkok

Városi felvételeken gyakran előfordulnak olyan síkok (épületek falai), melyek merőlegesek a mozgó jármű haladási irányára. Ezekben az esetekben a síkok normál vektora $[1 \ 0 \ 0]^T$ vagy $[0 \ 0 \ 1]^T$. A hozzájuk kapcsolódó homográfiák a következők:

$$\mathbf{H}_{s_1} \sim \begin{bmatrix} \cos \alpha - p & 0 & -\sin \alpha \\ 0 & 1 & 0 \\ \sin \alpha - q & 0 & \cos \alpha \end{bmatrix}, \mathbf{H}_{s_2} \sim \begin{bmatrix} \cos \alpha & 0 & -\sin \alpha - p \\ 0 & 1 & 0 \\ \sin \alpha & 0 & \cos \alpha - q \end{bmatrix}. \quad (12)$$

Annak ellenére, hogy a homográfiák nem teljesen megegyezők az 7. egyenletben találhatóval, a probléma lineáris és a keresett α , p és q paraméterek is ugyanazok. Ezért a probléma is hasonlóan megoldható. Az algebrai alak a következő:

$$\begin{aligned} \mathbf{A}_{v_1} [\cos \alpha, \sin \alpha, p, q]^T &= \mathbf{b}_{v_1}, \\ \mathbf{A}_{v_2} [\cos \alpha, \sin \alpha, p, q]^T &= \mathbf{b}_{v_2}, \end{aligned} \quad (13)$$

ahol $\mathbf{A}_{v_1}$ és a $\mathbf{A}_{v_2}$ az új együtttható mátrixok és az inhomogén egyenletek jobb oldala teljesen megegyező a 8 egyenlet jobb oldalával. Következésképpen $\mathbf{b}_{v_1} = \mathbf{b}_{v_2} = \mathbf{b}_{gd}$. Valamint az együtttható mátrixok a következő alakúak:

$$\mathbf{A}_{v1} = \begin{bmatrix} x - x' & -x'x - 1 & -x & x'x \\ -y' & -y'x & 0 & y'x \\ 1 - a_1 - x' & -a_1x - 1 & a_1x + x' \\ -a_2 & -a_2x & 0 & a_2x \\ -a_3 & -y' - a_3x & 0 & a_3x + y' \\ -a_4 & -a_4x & 0 & a_4x \end{bmatrix}, \quad (14)$$

$$\mathbf{A}_{v2} = \begin{bmatrix} x - x' & -x'x - 1 & -1 & x' \\ -y' & -y'x & 0 & y' \\ 1 - a_1 - x' & -a_1x & 0 & a_1 \\ -a_2 & -a_2x & 0 & a_2 \\ -a_3 & -y' - a_3x & 0 & a_3 \\ -a_4 & -a_4x & 0 & a_4 \end{bmatrix}. \quad (15)$$

3.3. Általános függőleges síkok

Szintén fontos feladat az autonóm vezetés szempontjából az olyan esetek vizsgálata, ahol a sík függőleges, de ismeretlen orientációjú. Ilyen síkok lehetnek például más autók, táblák, valamint az autóhoz viszonyítva szögben álló épületek falai. Egy általános

függőleges sík $\mathbf{n} = [n_x \ 0 \ n_z]^T$ normálvektorral rendelkeznek. A felületi normál kifejezhető egy δ szöggel, mellyel az $\mathbf{n} = [\cos \delta \ 0 \ \sin \delta]^T$. Az ebből következő homográfia:

$$\mathbf{H} = \begin{bmatrix} h_1 & 0 & h_3 \\ 0 & h_5 & 0 \\ h_7 & 0 & h_9 \end{bmatrix}, \quad (16)$$

ahol $h_1 = (\cos \alpha - p \cos \delta)/\gamma$, $h_3 = (\sin \alpha - p \sin \delta)/\gamma$, $h_5 = 1/\gamma$, $h_7 = (-\sin \alpha - q \cos \delta)/\gamma$, és $h_9 = (\cos \alpha - q \sin \delta)/\gamma$. Így a probléma öt szabadsági fokú, az ismeretlen paraméterek α , δ , γ , p és q . A hat lineáris egyenletrendszer a 4. és a 5. egyenletekből a következő hat sorból álló rendszert alkotja:

$$\mathbf{A}_{vert} \mathbf{h} = \begin{bmatrix} 1 & 0 & 0 & -(x' + a_1 x) & -a_1 \\ 0 & 0 & 0 & -a_2 x & -a_2 \\ 0 & 0 & 0 & -(y' + a_3 x) & -a_3 \\ 0 & 0 & 1 & -a_4 x & -a_4 \\ x & 1 & 0 & -xx' & -x' \\ 0 & 0 & y & -xy' & -y' \end{bmatrix} \begin{bmatrix} h_1 \\ h_3 \\ h_5 \\ h_7 \\ h_9 \end{bmatrix} = \mathbf{0}. \quad (17)$$

Megoldó. A homográfia mátrix elemei megbecsülhetők az $\mathbf{A}_{vert}$ mátrix null mátrixából. A γ paraméter eliminálható a homográfia mátrix $h_5 = 1$ átskálázásával.

A fennmaradó négy paraméter kinyerhető a skálázott homográfia mátrixból, melyek a következő alakban írhatóak fel:

$$\begin{aligned} h_1 &= \cos \alpha - p \cos \delta, \\ h_3 &= \sin \alpha - p \sin \delta, \\ h_7 &= -\sin \alpha - q \cos \delta, \\ h_9 &= \cos \alpha - q \sin \delta. \end{aligned} \quad (18)$$

Az első és harmadik egyenletből p és q kifejezhető:

$$p = \frac{\cos \alpha - h_1}{\cos \delta}, \quad q = -\frac{h_7 + \sin \alpha}{\cos \delta}. \quad (19)$$

Ezeket a második és negyedik egyenletbe visszahelyettesítve, majd elvégezve az alapvető egyenlet rendezéseket a következő két egyenlet kapható:

$$\begin{aligned} h_3 \cos \delta &= h_1 \sin \delta + \sin(\alpha - \delta), \\ h_9 \cos \delta &= h_7 \sin \delta + \cos(\alpha - \delta). \end{aligned} \quad (20)$$

Mindez egy mátrix-vektor szorzatként írható fel:

$$\begin{bmatrix} h_9 - h_7 \\ h_3 - h_1 \end{bmatrix} \begin{bmatrix} \cos \delta \\ \sin \delta \end{bmatrix} = \begin{bmatrix} \cos(\alpha - \delta) \\ \sin(\alpha - \delta) \end{bmatrix}. \quad (21)$$

Tehát egy megszorított mátrix-vektor egyenletet kaptunk, mivel $\mathbf{B} \mathbf{v}_1 = \mathbf{v}_2$ s.t. $\mathbf{v}_1^T \mathbf{v}_1 = \mathbf{v}_2^T \mathbf{v}_2 = 1$. A $\mathbf{B}$ mátrix SVD felbontása a következő:

$$\mathbf{B} = \mathbf{R}_1 \text{diag}(\sigma_1^2, \sigma_2^2) \mathbf{R}_2, \quad (22)$$

ahol $\mathbf{R}_1$ és $\mathbf{R}_2$ ortonormált mátrixok. Az algebrai probléma így írható:

$$\mathbf{B}\mathbf{v}_1 = \mathbf{R}_1 \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \mathbf{R}_2 \mathbf{v}_1 = \mathbf{v}_2. \quad (23)$$

Mindkét oldalt $\mathbf{R}_1^T$ -el megszorozva:

$$\text{diag}(\sigma_1^2, \sigma_2^2) \mathbf{R}_2 \mathbf{v}_1 = \mathbf{R}_1^T \mathbf{v}_2, \quad (24)$$

ahol $\mathbf{v}'_1 = \mathbf{R}_2 \mathbf{v}_1$ és $\mathbf{v}'_2 = \mathbf{R}_1^T \mathbf{v}_2$. Végül a következő egyenletek kell megoldani:

$$\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \mathbf{v}'_1 = \mathbf{v}'_2. \quad (25)$$

Érdemes megjegyezni, hogy $\mathbf{v}'_1{}^T \mathbf{v}'_1 = \mathbf{v}'_2{}^T \mathbf{v}'_2 = 1$ mivel a forgatás nem változtatja meg a vektor hosszát. Következésképpen ez egy egyszerű geometriai probléma: egy origó középpontú ellipszis és egy origó középpontú egység sugarú kör található az egyenlet bal, illetve jobb oldalán.

A $\mathbf{v}'_1$ és $\mathbf{v}'_2$ vektorok felírhatók két szög α és β függvényeként:

$$\mathbf{v}'_1 = \begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix}, \quad \mathbf{v}'_2 = \begin{bmatrix} \cos \beta \\ \sin \beta \end{bmatrix}. \quad (26)$$

Mivel $\sin^2 \alpha + \cos^2 \alpha = \sin^2 \beta + \cos^2 \beta = 1$, a 25. egyenlet felírható a következő két egyenlet segítségével:

$$\sigma_1^2 \cos \alpha = \cos \beta, \quad \sigma_2^2 \sqrt{1 - \cos^2 \alpha} = \sqrt{1 - \cos^2 \beta}. \quad (27)$$

Négyzetre emelés után a következő lineáris egyenletrendszer kapható a szögek koszinuszainak négyzetére:

$$\begin{bmatrix} \sigma_1^4 & -1 \\ -\sigma_2^4 & 1 \end{bmatrix} \begin{bmatrix} \cos^2 \alpha \\ \cos^2 \beta \end{bmatrix} = \begin{bmatrix} 0 \\ 1 - \sigma_2^4 \end{bmatrix}. \quad (28)$$

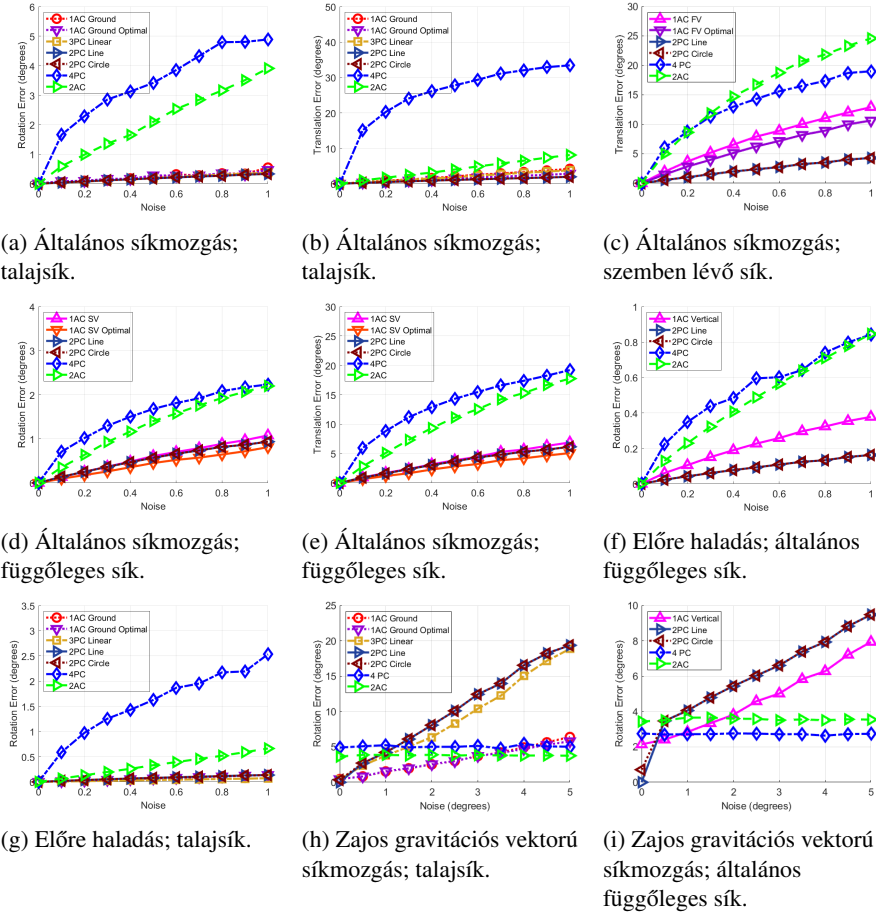
Így a végső alak:

$$\cos^2 \alpha = \frac{1 - \sigma_2^4}{\sigma_1^4 - \sigma_2^4}, \quad \cos^2 \beta = \frac{\sigma_1^4 - \sigma_1^4 \sigma_2^4}{\sigma_1^4 - \sigma_2^4}. \quad (29)$$

Az egyenletnek összesen négy megoldása van. A megoldások közül a helyes megoldás kiválasztható felhasználva, hogy minden 3D pontnak, amelyből a relatív póz meg lett becsülve, mindkét kamera előtt kell elhelyezkednie. A négy megoldás teljesen hasonló az általános sztereó rekonstrukció [19] során kapott négy megoldáshoz. A többértelműség oka, hogy a perspektív kamera modellje a kamera mögötti pontokat is leképezi a képsíkra, a láthatóságot nem veszi figyelembe.

4. Kiértékelés

A javasolt módszereket szintetikus adatokon és a Malaga adathalmazban [11] található valós képpárokon teszteltük. Az összes tesztelt algoritmust magunk implementáltuk.



3. ábra: Szintetikus kiértékelés. Az összehasonlított módszerek a következők: a javasolt talajsíkon alapuló megoldók a 1AC Ground (gyors megoldó) és a 1AC Ground Optimal (optimális megoldó); a szemben lévő falat felhasználó megoldók a 1AC FV (gyors) és a 1AC FV Optimal (optimális); oldalsó falat használók a 1AC SV (gyors) és a 1AC SV Optimal (optimális); az általános függőleges fal esetét megoldó 1AC Vertical; a két általános homográfiát becslő megoldó: normalizált DLT algoritmus [19] (4PC) és a 2AC algoritmus [4] valamint a 3PC Linear [12], 2PC Line [12], 2PC Circle [12] algoritmusok. A 2a-2g ábrákon három különböző zaj hozzáadás történik a bemeneti adatokhoz. Zajt adtunk hozzá a pontmegfeleltetésekhez, valamint az affin skálához és forgatáshoz. Az egyes zajok 0-tól 1-ig terjednek pixelben, százalékban és szögben.

4.1. Szintetikus kiértékelés

A szintetikus adatokon való kiértékeléshez egy, a Saurer és munkatársaihoz [32] hasonló elrendezést használtunk. A jelenet tartalmaz egy talajsíkot, szemben és oldalt

lévő síkokat, valamint ismeretlen orientációjú síkokat. Az első kamerát a síkktól 10 egységnyi távolságra helyeztük el. A kamerák bázistávolsága 1 egység, a fókusztávolság 10 egység. Valamennyi algoritmus kiértékelését különböző hozzáadott zajok esetén végeztük el. Az összes algoritmust három különböző mozgás esetén vizsgáltuk: tisztán előre haladó (a Z tengely mentén), tisztán oldalra haladó (az X tengely mentén) és véletlenszerű síkmozgás esetén.

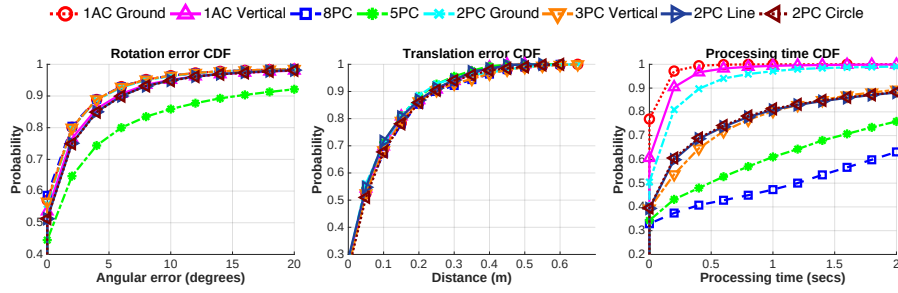
Az algoritmusok pontosságának kiértékeléséhez összehasonlítottuk a becsült relatív pózokat. A relatív forgásban lévő eltérés mérésére kiszámítottuk a szögeltérést a valós és a becsült forgatások között a $\xi_R = \cos^{-1}((\text{tr}(\mathbf{R}\hat{\mathbf{R}}^T) - 1)/2)$ összefüggés segítségével, ahol $\hat{\mathbf{R}}$ a valós, $\mathbf{R}$ pedig a becsült forgatás. Mivel az eltolás egy skála erejéig adható meg, ezért az eltolásban lévő hiba a valós eltolás vektor és a becsült eltolás vektorok által bezárt szög.

A javasolt módszereket a normalizált DLT [19] (4PC), 2AC [4], 2PC Line és Circle [12], valamint a 3PC Linear [12] megoldókkal hasonlítottuk össze. Minden tesztet 10000-szer futtattunk. Két típusú tesztet vizsgáltunk. Az első esetben az Y tengelye a kameráknak párhuzamos, ezek láthatóak a 3a - 3g ábrákon. Ebben az esetben három típusú zajt adtunk az adatokhoz. A pontmegfeleltetésekhez 0-tól 1-ig terjedő pixelzajt, az affin transzformációk forgás komponenséhez 0-tól 1-fokig, a skála komponenséhez pedig 0-tól 1 százalékgig terjedő zajt adtunk. A 3h - 3i ábrákon az algoritmusok érzékenységét teszteltük arra az esetre, ha a mozgás nem teljesen síkmozgás. Ezeknél a teszteknel a második kamera függőleges irányát forgattuk el az X és a Z tengely körül egy 0-tól 5 fokig terjedő szöggel. Emellett fix 1 pixeles pontzajt és 1 fokos (affin forgatás komponenshez), valamint 1 százalékos affin skála zajt adtunk az adatokhoz.

Talajsíkok. A javasolt 1AC Talaj (Ground) Gyors és Optimális (Optimal) megoldókat hasonlítottuk össze az általános 4PC és 2AC, valamint a 3PC Linear, 2PC Line és Circle algoritmusokkal (3a - 3b ábrák). A kamerák mozgása véletlenszerű síkmozgás. Eredményeink azt mutatják, hogy a javasolt algoritmusok a legjobban teljesítő megoldókéhoz hasonló alacsony hibát produkálnak.

Speciális függőleges síkok. Abban az esetben, amikor a megfigyelt sík a kamerákkal szemben vagy oldalt található, a 1AC Függőleges (Vertical) algoritmus speciális eseteit teszteltük. Fontos megjegyezni, hogy az ilyen elrendezésű síkok a 3PC Linear algoritmusnak degenerált esetei, ezért a 3PC algoritmust itt kihagytuk. A 3c. ábrán látható esetben véletlenszerű síkbeli mozgás és egy $[0 \ 0 \ 1]^T$ normálvektorú sík mellett teszteltük az algoritmusokat. Mind a pont és affin komponensekhez zajt adtunk. A javasolt módszerek jobb eredményt érnek el az általános megoldóktól, azonban a SOTA 2PC alapú módszerek kevésbé érzékenyek a zajra. Érdemes megjegyezni, hogy habár a zaj mind a pont, mind az affin komponenseket módosítja, az affin transzformációkon alapuló algoritmusok érzékenyebbek az affin zajra. A forgatás szögének becslése hasonló karakterisztikájú eredményt ad, ezért nem foglaltuk bele a cikkbe. A 3d - 3e ábrákon az $[1 \ 0 \ 0]^T$ normál vektorú sík esete látható véletlenszerű síkmozgás mellett. A javasolt 1AC Optimal (Optimális) megoldó vezet a legpontosabb eredményekre, mind a forgatás, mind az eltolás esetén.

Tisztán egyenes mozgás. A 3f ábrán vizsgáljuk azt az esetet, amikor tisztán egyenes előre haladó mozgás mellett általános függőleges síkok találhatóak a jelenetben. A javasolt módszerek jobb eredményhez vezetnek, mint az általános 2AC és 4PC módszerek,



4. ábra: A Malaga adathalmaz 15 jelenetének 9 064 képpárján elvégzett tesztek kumulatív eloszlásfüggvénye: a forgás hibájának (fokban), az eltolás hibájának (méterben) és a futási időnek (másodpercben). Minél jobban közelít egy adott algoritmus a bal felső sarokhoz, annál pontosabb. Robosztus becsléshez a GC-RANSAC-ot [3] használtuk. Az összehasonlított algoritmusok: 1AC Ground, 1AC Vertical megoldók; a nyolc pontos (8PC)- [19] és az öt pontos általános (5PC) [37]; a módszerek a [32] cikkből, 2PC Ground és 3PC Vertical, valamint a 2PC-alapú algoritmusok [12] alapján, 2PC Line és 2PC Circle.

azonban a Choi és mtsai. [12] által javasolt módszerek pontosabbak. A 3g ábrán látható tesztelésben tisztán egyenes mozgás és egy talajsík jelenléte mellett teszteltük az algoritmusokat. A javasolt módszerek a legjobban teljesítő eljárásokhoz hasonlóan pontos eredményhez vezettek.

Függőleges irány. A 3h - 3i ábrákon kis mértékben elrontjuk azt az előfeltételt, hogy a kamerák egy sík felületen mozognak az által, hogy elforgatjuk őket az X és Z tengelyeik mentén. Az elvárásoknak megfelelően a 4PC és a 2AC algoritmusokra nincs hatással ez a típusú zaj, hiszen általános mozgást feltételeznek. A javasolt módszerek jelentősen jobb eredményhez vezetnek a 2PC Line és Circle megoldóknál. A javasolt 1AC Ground és Optimal megoldók pontosabbak az általános 4PC és 2AC (általános) megoldóknál, ha a zaj körülbelül 3 fok alatti. A javasolt 1AC Vertical módszer pontosabb az általános módszereknél, ha a zaj 1 fok alatti. A tesztekéből az következik, hogy algoritmusaink jobban teljesítenek az összehasonlításban szereplőknél, mivel az függőleges irány általában 1 fok pontossáig meghatározható [16] egy egyszerű okostelefon (Nokia N900 vagy iPhone 4) segítségével. Napjainkban a modern autókban vagy okoseszközökben használt gyorsulásmérő szenzorok 0.06° , valamint a drágább jó minőségű társaik $< 0.02^\circ$ pontossáig képesek meghatározni a függőleges irányt [16].

4.2. Valós tesztek

A javasolt módszereket a Malaga adathalmazon [11] teszteltük. Ez az adathalmaz városi környezetben közlekedő autóra szerelt szenzorok adatait tartalmazza. Köztük található például egy nagy felbontású sztereó kamera felvételei és öt lézer szkennerek adatai. Cikkünkben a nagy felbontású kamera képsorozatának minden tizedik képét használjuk fel. A javasolt módszereket futtattuk az így kapott képsorozat összes egymás követő

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	átlag
Idő (s)	1AC(G)	0.12	0.11	0.04	0.08	0.03	0.10	0.18	0.12	0.10	0.31	0.06	0.13	0.07	0.07	0.08	0.11
	1AC(V)	0.19	0.21	0.07	0.14	0.06	0.15	0.33	0.20	0.16	0.61	0.11	0.24	0.12	0.12	0.15	0.19
	8PC	4.57	3.82	0.96	2.55	0.79	4.56	5.71	3.36	2.49	4.54	1.72	3.64	2.63	1.92	1.86	3.01
	5PC	3.02	2.87	0.60	3.00	0.58	3.11	2.78	1.63	0.98	4.07	1.13	1.92	1.92	1.23	1.26	2.01
	2PC(G)	0.33	0.46	0.10	0.24	0.12	0.23	0.58	0.32	0.29	1.04	0.15	0.36	0.28	0.19	0.26	0.33
	3PC(V)	1.25	2.02	0.21	1.28	0.50	1.14	2.10	0.87	0.79	3.58	0.42	1.34	1.18	0.65	0.72	1.20
	2PC(L)	0.74	2.53	0.34	0.83	0.70	0.39	4.52	1.26	1.20	10.06	0.54	2.03	1.16	1.06	1.79	1.94
	2PC(C)	0.75	2.50	0.34	0.79	0.68	0.38	4.71	1.26	1.26	10.15	0.52	2.00	1.12	1.04	1.74	1.95
Szöghiba (°)	1AC(G)	2.40	2.08	1.01	1.50	3.64	4.15	4.63	1.39	3.16	5.15	0.99	2.52	4.87	1.37	1.96	2.72
	1AC(V)	2.54	2.98	2.60	1.94	3.69	4.25	4.89	2.56	3.21	5.52	1.94	2.98	5.50	2.18	3.30	3.34
	8PC	2.42	2.06	0.93	1.46	3.64	4.17	4.79	1.34	2.89	5.18	0.88	2.53	4.75	1.27	1.95	2.68
	5PC	3.67	7.41	6.05	10.27	4.09	6.25	7.13	2.16	12.54	6.87	9.26	6.16	12.72	8.98	7.48	7.40
	2PC(G)	2.37	2.23	1.13	1.52	3.65	4.14	4.73	1.52	3.12	5.18	1.09	2.57	4.95	1.54	2.09	2.79
	3PC(V)	2.40	2.09	1.03	1.52	3.64	4.16	4.76	1.37	3.08	5.09	1.05	2.54	4.88	1.47	1.93	2.73
	2PC(L)	2.39	2.80	2.32	2.04	3.71	4.34	4.73	2.15	3.30	6.23	1.92	2.76	6.29	2.73	3.19	3.39
	2PC(C)	2.36	3.09	2.17	1.99	3.70	4.23	4.92	2.08	3.66	7.21	1.88	2.84	6.18	2.74	3.81	3.52

1. táblázat: Az ábrán az átlagos futási idők (másodpercben) és a relatív póz becslésének forgatás hibája látható a Malaga adathalmaz 15 jelenetén (oszlopok). Robosztus becsléshez a GC-RANSAC [3] módszert használtuk. Az összehasonlított algoritmusok az öt pontos Stewenius és mtsai. (5PC) [37], a normalizált nyolc pontos megoldó (8PC) [19], Saurer és mtsai. [32] által bevezetett két pontos talajsíkot (2PC(G)) és három pontos függőleges síkot (3PC(V)) feltételező módszerek, Choi által bevezetett [12] vonalon (2PC(L)) és körön (2PC(C)) alapuló megoldók, valamint a javasolt két affin megfeleltetéseket használó módszer: a talajsíkot használó (1AC(G)) és a függőleges síkot használó (1AC(V)). Az eredményekhez tartozó kumulatív eloszlásfüggvények a 4. ábrán láthatók.

képpárjára. Referenciaként (ground truth) az adathalmazhoz tartozó GPS koordinátákból kapott trajektóriát használjuk.

Összesítve 9 064 képpárt használtunk fel a kiértékelésben. Az affin transzformációk kinyeréséhez [21] a VLFeat könyvtárat [38] használtuk, mely a Difference-of-Gaussians algoritmust használja az affin alakzatok hozzárendelésének összekombinálásával [9]. A tesztjeinkben az affin alakzatok hozzárendelése csupán egy kis ($\sim 10\%$) extra időt igényel az általános jellemző kinyeréshez képest. A megfeleltetéseket általános SNN ráta teszteléssel szűrjük [21]. Robosztus becsléshez a Graph-Cut RANSAC [3] (GC-RANSAC) algoritmust használjuk.

A GC-RANSAC (és más RANSAC-hoz hasonló módszerekben) két különböző megoldót használunk: (a) egyet a minimális minta illesztésére és (b) egyet a nem minimális minta illesztésére, amikor a modell finomítását végezzük az összes belső ponton vagy a lokális optimalizációs lépések során. Az (a) esetén a fő cél, hogy minél kevesebb adatpontot felhasználva oldjuk meg a problémát, hiszen a futási idő exponenciálisan függ a felhasznált pontok számától. A javasolt és konkurens módszerek ebben a részében szerepelnek a robusztus illesztésnek. Fontos megjegyezni, hogy a megfigyelt speciális síkok kevesebb belső pontot tartalmaznak az általános esetekhez képest, mivel lokalizáltan helyezkednek el a képen. Következésképpen a homográfia kiszámítása helyett a RANSAC ciklusban azonnal az esszenciális mátrix kerül kiszámítására a kinyert relatív pozíciókból, és nem a homográfia maga. A (b) részben az általános nyolc pontos relatív póz algoritlussal becsüljük meg az esszenciális mátrixot a nem minimális belső pontok halmazán.

Az összehasonlításban felhasználjuk az összes módszert a szintetikus tesztekben, valamint a 2PC Ground és 3PC Vertical [32], 5PC [37] megoldókat és a normalizált 8PC [19] algoritmust. A valós kiértékelés során leteszteltük, hogy az optimális és gyors megoldók közötti pontosságkülönbség a robusztus becselő által kiegyensúlyozott. Így a tesztek során a gyors megoldókat használtuk, hiszen nem vezetett jelentős pontosságvesztéshez, azonban sokkal gyorsabban kiszámítja az eredményeket.

A kumulatív eloszlásfüggvénye a forgatás (fokban) és eltolás hibáinak (méterben), valamint a futási időknak (másodpercben) a 4. ábrán látható. A hiba az esszenciális / homográfia mátrixok 3D forgatás és eltolássá való dekompozíciójából számítható. Minél pontosabb egy algoritmus, annál jobban közelít az ábrák bal felső sarkához. Az eltolásvektorok hibája megadható az adathalmazból vett GPS koordináták felhasználásával. Mind a két javasolt megoldó (1AC Ground és 1AC Vertical) a legpontosabbak között található.

Robosztus becslések teljes futási ideje a 4. ábrán látható. Az elvárásoknak megfelelően a javasolt 1AC Ground módszer, amely *messze a leggyorsabb* minden esetben maximum 0.4 – 0.5 másodperc alatt végez. Az esetek 80% százalékában a futási ideje 0.01 másodperc. A javasolt 1AC Vertical módszer a második leggyorsabb. Az átlagos forgatás hiba és futási idő a 1. táblázat tartalmazza. Jól látható, hogy minden jelenet esetén a 1AC Ground algoritmus a leggyorsabb amellet, hogy a második legpontosabb eredményre vezet.

5. Összefoglalás

Cikkünkben javaslatot tettünk mozgó járműre szerelt kalibrált kamerák mozgásának becslésére egyetlen affin megfeleltetés és speciális síkok felhasználásával. Ez a probléma kiemelt fontosságú önvezető autózás során. A javasolt módszerek általában egy egyszerű, 6×5 méretű lineáris egyenletrendszer megoldását igénylik. Ilyen kis méretű egyenletrendszerek megoldása rendkívül gyors, átlagosan $5\text{--}10\ \mu\text{s}$ futási időt igényelnek C++ implementáció esetén. Továbbá a kevés szükséges megfeleltetés miatt jelentősen felgyorsítják a legkorszerűbb robosztus becslő módszereket. Tesztjeink azt mutatják, hogy a talajsíkot használó megoldónk teljesített legjobban. Átlagosan a második legpontosabb eredményt éri el a Malaga adathalmaz megközelítőleg 9000 képpárján, amellett, hogy a leggyorsabb futási idővel rendelkezik.

A javasolt módszerek gyorsaságuk révén beilleszthetők valós időben működő robot és (fél-)autonóm járművekbe. Mivel egy ember által alkotott környezet gyakran tartalmaz vízszintes és függőleges síkokat ezek a módszerek valós körülmények között sokszor használhatóak.

Köszönetnyilvánítás

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.

Az Alkalmazási terület-specifikus nagy megbízhatóságú informatikai megoldások című projekt a Nemzeti Kutatási Fejlesztési és Innovációs Alapból biztosított támogatással, a Tématerületi kiválósági program (TKP2020-NKA-06, Nemzeti Kihívások Alprogram) finanszírozásában valósult meg.

Irodalom

1. Abdellali, H., Kato, Z.: Absolute and relative pose estimation of a multi-view camera system using 2d-3d line pairs and vertical direction. In: 2018 Digital Image Computing: Techniques and Applications (DICTA). pp. 1–8. IEEE (2018)
2. Bailey, T., Durrant-Whyte, H.: Simultaneous localization and mapping (slam): Part ii. IEEE robotics & automation magazine **13**(3), 108–117 (2006)
3. Baráth, D., Matas, J.: Graph-cut RANSAC. Conference on Computer Vision and Pattern Recognition pp. 6733–6741 (2018)
4. Barath, D., Hajder, L.: A theory of point-wise homography estimation. Pattern Recognition Letters **94**, 7–14 (2017)
5. Barath, D., Hajder, L.: Efficient recovery of essential matrix from two affine correspondences. IEEE Transactions on Image Processing **27**(11), 5328–5337 (2018)
6. Barath, D., Molnár, J., Hajder, L.: Optimal surface normal from affine transformation. In: International Conference on Computer Vision Theory and Applications. vol. 2, pp. 305–316. SciTePress (2015)
7. Barath, D., Polic, M., Förstner, W., Sattler, T., Pajdla, T., Kukulova, Z.: Making affine correspondences work in camera geometry computation. In: European Conference on Computer Vision (2020)

8. Barath, D., Toth, T., Hajder, L.: A minimal solution for two-view focal-length estimation using two affine correspondences. In: Conference on Computer Vision and Pattern Recognition. pp. 6003–6011 (2017)
9. Baumberg, A.: Reliable feature matching across widely separated views. In: Conference on Computer Vision and Pattern Recognition. pp. 774–781. IEEE (2000)
10. Bentolila, J., Francos, J.M.: Conic epipolar constraints from affine correspondences. *Computer Vision and Image Understanding* **122**, 105–114 (2014)
11. Blanco, J.L., Moreno, F.A., Gonzalez-Jimenez, J.: The málaga urban dataset: High-rate stereo and lidars in a realistic urban scenario. *International Journal of Robotics Research* **33**(2), 207–214 (2014)
12. Choi, S., Kim, J.: Fast and reliable minimal relative pose estimation under planar motion. *Image Vision Computing* **69**, 103–112 (2018)
13. Chou, C., Wang, C.: 2-point RANSAC for scene image matching under large viewpoint changes. In: International Conference on Robotics and Automation. pp. 3646–3651 (2015)
14. Durrant-Whyte, H., Bailey, T.: Simultaneous localization and mapping: part i. *IEEE Robotics & Automation Magazine* **13**(2), 99–110 (2006)
15. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* **24**(6), 381–395 (1981)
16. Fraundorfer, F., Tanskanen, P., Pollefeys, M.: A minimal case solution to the calibrated relative pose problem for the case of two known orientation angles. In: European Conference on Computer Vision. pp. 269–282. Springer (2010)
17. Guan, B., Vasseur, P., Demonceaux, C., Fraundorfer, F.: Visual odometry using a homography formulation with decoupled rotation and translation estimation using minimal solutions. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). pp. 2320–2327. IEEE (2018)
18. Gál, I.G.: Matlab implementation of the proposed algorithms (2021), <https://github.com/Elenadar/Pose-Estimation-for-Vehicle-mounted-Cameras-via-Horizontal-and-Vertical-Planes>
19. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge University Press (2003)
20. Köser, K.: Geometric estimation with local affine frames and free-form surfaces (2009), PhD. Thesis
21. Lowe, D.G.: Object recognition from local scale-invariant features. In: International Conference on Computer Vision. vol. 2, pp. 1150–1157. Ieee (1999)
22. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide baseline stereo from maximally stable extremal regions. In: British Machine Vision Conference (2002)
23. Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Van Gool, L.: A comparison of affine region detectors. *International Journal of Computer Vision* **65**(1-2), 43–72 (2005)
24. Mishkin, D., Matas, J., Perdoch, M.: MODS: Fast and robust method for two-view matching. *Computer Vision and Image Understanding* **141**, 81–93 (2015)
25. Mishkin, D., Radenovic, F., Matas, J.: Repeatability is not enough: Learning affine regions via discriminability. In: Conference on Computer Vision. pp. 284–300 (2018)
26. Morel, J.M., Yu, G.: ASIFT: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences* **2**(2), 438–469 (2009)
27. Ortín, D., Montiel, J.M.M.: Indoor robot motion based on monocular images. *Robotica* **19**, 331–342 (2001)
28. Perdoch, M., Matas, J., Chum, O.: Epipolar geometry from two correspondences. In: International Conference on Pattern Recognition. vol. 4, pp. 215–219 (2006)

29. Pritts, J., Kukulova, Z., Larsson, V., Chum, O.: Radially-distorted conjugate translations. Conference on Computer Vision and Pattern Recognition pp. 1993–2001 (2018)
30. Raposo, C., Barreto, J.: π match: Monocular vslam and piecewise planar reconstruction using fast plane correspondences. In: European Conference on Computer Vision. pp. 380–395 (2016)
31. Raposo, C., Barreto, J.P.: Theory and practice of structure-from-motion using affine correspondences. In: Computer Vision and Pattern Recognition. pp. 5470–5478 (2016)
32. Saurer, O., Vasseur, P., Boutteau, R., Demonceaux, C., Pollefeys, M., Fraundorfer, F.: Homography based egomotion estimation with a common direction. IEEE transactions on pattern analysis and machine intelligence **39**(2), 327–341 (2016)
33. Scaramuzza, D.: 1-point-ransac structure from motion for vehicle-mounted cameras by exploiting non-holonomic constraints. International Journal of Computer Vision **95**(1), 74–85 (2011)
34. Schonberger, J.L., Frahm, J.M.: Structure-from-motion revisited. In: Conference on Computer Vision and Pattern Recognition. pp. 4104–4113 (2016)
35. Snavely, N., Seitz, S.M., R: Modeling the world from internet photo collections. International Journal of Computer Vision **80**(2), 189–210 (2008)
36. Snavely, N., Seitz, S.M., Szeliski, R.: Photo tourism: exploring photo collections in 3d. In: ACM Transactions on Graphics. vol. 25, pp. 835–846. ACM (2006)
37. Stewenius, H., Engels, C., Nistér, D.: Recent developments on direct relative orientation. ISPRS Journal of Photogrammetry and Remote Sensing **60**(4), 284–294 (2006)
38. Vedaldi, A., Fulkerson, B.: VLFeat: An open and portable library of computer vision algorithms, <https://www.vlfeat.org/>