

Affin transzformációt becslő eljárások kvantitatív kiértékelése drónfelvételek által

Zoltán Pusztai^{1,2}, Levente Hajder²

¹ Gépi Érzékelés Laboratórium, MTA SZTAKI, Budapest, Magyarország
zoltan.pusztai@sztaki.mta.hu

² Algoritmusok és Alkalmazásaik Tanszék, ELTE IK, Budapest, Magyarország
hajder@inf.elte.hu

Kivonat A jellegzetes pont detektáló algoritmusokat gyakran használják a számítógépes látás területein. Mostanában, az affin transzformációt is becslő módszerek lettek népszerűek. Az affin transzformációkkal lehetséges az epipoláris geometria alapvető problémáinak megoldása kevesebb megfeleltetést használva. Ebben a cikkben az affin transzformációt is detektáló módszereket hasonlítjuk össze kvantitatíven, drón felvételek által. A valós (GT) transzformációk a kamera kötött mozgásából kiszámíthatóak. Az így létrejött adatbázissal a detektált pontok és az affin transzformációk pontossága is kiértékelhető. Az adatbázis segítségével a legnépszerűbb affin transzformációt becslő algoritmusok kvantitatíven összehasonlításra kerülnek.¹

1. Bevezetés

A jellegzetes pontkövető algoritmusokat már a számítógépes látás kezdete óta vizsgálják. Kezdetben csak a pontokat detektáló módszereket használták a kamera mozgás és epipoláris geometria becslésére. A lokális jellegzetes pontok a kép csupán kis részét írják le, ezért lehetséges őket hatékonyan használni pontmegfeleltetések keresésére, még nagy kamera nézetbeli és fényviszonybeli változás esetében is. Az affin detektorok ezzel szemben képesek a jellegzetes pontok körüli környezet torzulását is mérni. Az ebből a torzulásból nyert affin transzformációk által lehetséges az epipoláris geometria alapvető feladatainak megoldása (mint például a kamera mozgás meghatározása, objektum detektálása vagy 3D rekonstrukció), kevesebb megfeleltetést felhasználva. Ebben a cikkben az affin detektorok kvantitatív összehasonlítását mutatjuk be, ehhez valós környezetben felvett drónfelvételeket felhasználva.

A jellegzetes pont detektorokat már régóta tanulmányozzák a számítógépes látás területén. Az ismert Harris [6] és Shi-Tomasi sarokdetektor [13] több, mint két évtizede ismert. Azóta, új jellegzetes pont detektáló algoritmusok láttak napvilágot, mint például a SIFT [9], SURF [4], KAZE [1] vagy a BRISK [8]. A képeken külön-külön detektált pontokat meg kell egymásnak feleltetni, azok úgynevezett leíró vektorát felhasználva. Ez a leíró a jellegzetes pontok lokális környezetét

¹ Kuba Attila díjra beadott cikk.

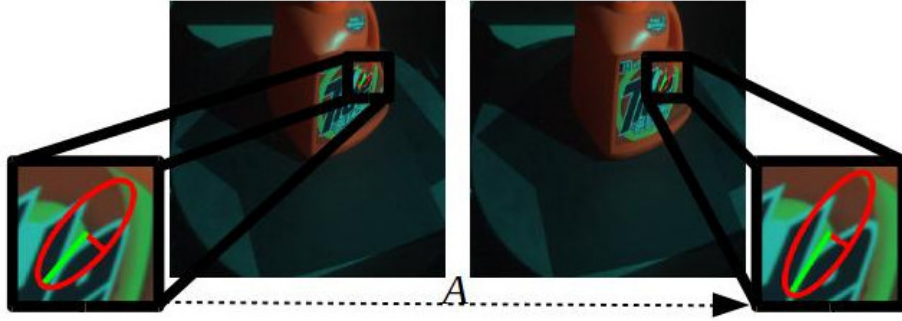
jellemzi egy kompakt vektor által. A pontmegfeleltetések keresése során, ezen leíró vektorok távolságát vesszük figyelembe. Azok a jellegzetes pontok, melyek leíró vektorai közel esnek egymáshoz, potenciálisan egy megfeleltetést alkotnak.

A jellegzetes pont kereső módszerek által általában csak pont megfeleltetésekre tehetünk szert. Viszont az affin transzformációt is detektáló módszerek képesek a lokális kép torzulását is mérni a pontok körül. Az affin transzformáció nem más, mint a kamera által leírt vetítő-függvény Taylor sorbafejtésének elsőrendű közelítése. Vagyis, ez egy 2×2 mátrix, amely az első képen található jellegzetes pont körüli lokális környezetet transzformálja át a neki megfelelő jellegzetes pont környezetébe a második képen. Az affin transzformáció elegendő információt tartalmaz a neki megfelelő térbeli ponthoz tartozó érintősík normálvektorának felírásához. A normálvektorra tett megkötés segítségével lehetséges a fundamentális mátrix, kamera mozgás vagy egyéb epipoláris geometriai tulajdonság becslése, kevesebb megfeleltetést felhasználva, mint pontok esetében.

Az összehasonlítások irodalma nem kielégítő. Talán a legjelentősebb összehasonlítást Mikolajczyk és társai által lett publikálva [11]. Ebben, különböző affin detektorok lettek összehasonlítva, valós környezetben felvett képek által. Azonban, ezek képek vagy csak sík objektumokról készültek, vagy pedig nem változott a kamera helyzete a felvételek között. Az affin transzformációk hibáját a általuk definiált ellipszisek átfedése adta. Habár a szerzők különböző zaj használatával gazdagították az összehasonlítást (elmosódás, JPEG tömörítés, fényviszony változás), a fix pozícióban tartott kamera és a sík objektumok túlságosan nagy megkötést jelentenek. Ezen felül az ellipszisek átfedésének mértéke nem bizonyul jó metrikának, hiszen ez függ azok méretétől. Tuytelaars és társai [14] szintén megjelentettek egy összehasonlítást, azonban ebben csak szintetikus előállított képeket használtak. Korábban, Pusztai és Hajder [12] egy 3D szkenneren alapuló összehasonlítást végeztek mozgó objektumokról. Azonban az összehasonlításban, csak kis méretű tárgyakat használtak.

A kvantitatív összehasonlítások irodalma nem teljes, annak ellenére, hogy mennyire fontos téma. A legtöbb detektáló algoritmust csak pár képen lett tesztelve, ezek paraméterei úgy lettek beállítva, hogy a legjobb eredményt adják ezen a néhány tesztképeken. A valós körülmények eltérhetnek a laboratóriumi körülményektől, ezért lehetséges, hogy más paraméterek használata szükséges. Illetve az is előfordulhat, hogy más algoritmus bizonyul jobbnak a gyakorlati felhasználás során. Ezért az összehasonlításnak valós kamera felvételeken kell alapulnia, miközben a kamera mozgást végez.

Ebben a cikkben bemutatjuk, hogy az affin kovariáns jellegzetes pont detektorokat kvantitatíven össze lehet hasonlítani valós környezetben készült drónfelvételeket használva. *A cikk fő újdonságai: (1) Bemutatásra kerül a pontos GT affin transzformációk és pont mozgások kiszámítása, különböző speciális kameramozgások esetében. Legjobb tudomásunk szerint, ez az első kvantitatív összehasonlítás, amely a GT adatot drónfelvételek által állítja elő. (2) A GT adatok alapján az ismertebb affin kovariáns detektorok összehasonlításra kerülnek. Mind a jellegzetes pontok pozíciójának detektálása, mind pedig az affin transzformációk pontosságára kiértékelésre kerül.*



1. ábra. A képen ellipszisek jelölik a detektált affin régiókat. A becsült affin transzformáció \mathbf{A} , az első képen látható affin régiót transzformálja át a második képen láthatóba.

2. Az affin kovariáns detektorok rövid áttekintése

Ebben a fejezetben az affin transzformációk és az affin kovariáns detektorok kerülnek röviden bemutatásra. A detektorok célja, hogy a jól megkülönböztethető jellegzetes pontokat megtalálják a képeken, külön-külön. Ha egy ilyen jellegzetes pontot megtaláltak, akkor lehetséges a pont körüli az affin régió megkeresése, amelyet egy ellipszissel szokás szemléltetni. Az 1. ábrán látható két, egymásnak helyesen megfeleltetett jellegzetes pont körüli affin régió. A különböző detektorok más és más eljárások által keresik meg a jellegzetes pontokat, és az azok körüli affin alakot. A továbbiakban ezen detektorok működése kerül bemutatásra.

Harris-Laplace, Harris-Affine. A Mikolajczyk [10] által publikált Harris-Laplace és Harris-Affine az ismert Harris pont detektoron alapul [6]. A Harris az úgynevezett második momentum mátrixot használja a jellegzetes pontok megtalálásához, amely a következőképpen néz ki:

$$\mathbf{M}(\mathbf{x}) = \sigma_D^2 G(\sigma_I) * \begin{bmatrix} f_x^2(\mathbf{x}, \sigma_D) & f_x(\mathbf{x}, \sigma_D)f_y(\mathbf{x}, \sigma_D) \\ f_x(\mathbf{x}, \sigma_D)f_y(\mathbf{x}, \sigma_D) & f_y^2(\mathbf{x}, \sigma_D) \end{bmatrix},$$

ahol

$$G(\sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|\mathbf{x}|^2}{2\sigma^2}\right), \quad f_x(\mathbf{x}, \sigma_D) = \frac{\partial}{\partial \mathbf{x}} G(\sigma_D) * f(\mathbf{x}).$$

Az \mathbf{M} mátrix a gradiens eloszlást tartalmazza az \mathbf{x} pont környezetében, σ_D , σ_I pedig a méret és integrációs skála. A jellegzetes pontot ott találjuk meg, ahol a $\det(\mathbf{M}) - \lambda \text{trace}(\mathbf{M})$ nagyobb, mint egy előre beállított küszöb érték. Ennek a jelentése az, hogy az \mathbf{M} sajátértékei nagyok, így sarokpontot találtunk.

Miután a jellegzetes pontot megkerestük, az affin alakzat meghatározása következik. Először a karakterisztikus skálát kell meghatározni a kör alakú Laplace operátor segítségével. Az a jellegzetes mérték kerül kiválasztásra, amelyre a pont környezete és a skálázott operátor hasonlósága a legnagyobb. Az affine kovariáns, Harris-Affine algoritmus esetében egy második skála is meghatározásra kerül, amelyet a következő iteratív algoritmus becsül meg:

1. A kezdeti középpont és a karakterisztikus skála meghatározása.
2. Az affin alakzat becslése \mathbf{M} másodrendű mátrix alapján.
3. Az affine régió körre normalizálása $\mathbf{M}^{1/2}$ -t felhasználva.
4. Új pozíció és skála kiszámolása a normalizált képen.
5. Goto (2), ha \mathbf{M} sajátértékei nem egyenlők.

Az iteráció minden esetben konvergál. A végleges pozíciót és hozzá tartozó ellipszis alakú régiót pedig a konvergálás után adja az algoritmus.

A második skála és alakzat kiválasztására leírt, 5 lépésből álló algoritmus tetszőleges kezdőpont esetében alkalmazható. A szintén Mikolajczyk által bemutatott **Hessian-Laplace** és **Hessian-Affine** detektorok a jellegzetes pontok megtalálására a Hesse mátrixot használják. Ez a következőképpen néz ki:

$$H(\mathbf{x}) = \begin{bmatrix} f_{xx}(\mathbf{x}, \sigma_D) & f_{xy}(\mathbf{x}, \sigma_D) \\ f_{xy}(\mathbf{x}, \sigma_D) & f_{yy}(\mathbf{x}, \sigma_D) \end{bmatrix}, \quad (1)$$

ahol $f_{xx}(\mathbf{x}, \sigma_D)$ a Gauss szűrő által elsimított képen számított másodrendű deriváltak. A Hesse mátrix folszerű képi jellemzők detektálására használható.

EBR. A Tuytelaars és társai által bemutatott EBR [15] szintén a sarokpontok körüli affine kovariáns régiók kinyerésére használható. A Harris sarokdetektáló algoritmust és a Canny éldetektort együttesen alkalmazzák. Az affine régiókat ott lehet kinyerni, ahol egy sarokpontban két él találkozik. Ha a sarokpontot (\mathbf{p}) és egy-egy pontot veszünk az éleken (\mathbf{p}_1 és \mathbf{p}_2), akkor azok egy paralelogrammát feszítenek ki. A végső affin régiót ott találja meg az algoritmus, ahol a parÖsszeselogramma a következő függvény maximumát adja:

$$f(\Omega) = abs \left(\frac{|(\mathbf{p} - \mathbf{p}_g)(\mathbf{q} - \mathbf{p}_g)|}{|(\mathbf{p} - \mathbf{p}_1)(\mathbf{p} - \mathbf{p}_2)|} \right) \times \frac{M_{00}^1}{\sqrt{M_{00}^2 M_{00}^0 - (M_{00}^1)^2}}, \quad (2)$$

ahol

$$M_{pq}^n = \int_{\Omega} I^n(x, y) x^p y^q dx dy, \quad \mathbf{p}_g = \left(\frac{M_{00}^1}{M_{00}^1}, \frac{M_{01}^1}{M_{00}^1} \right),$$

és \mathbf{p}_g pedig a paralelogramma tömegközéppontja. A paralelogrammára végül ellipszist illesztnek.

IBR. Amíg az EBR a sarokpontokban és az azokhoz kapcsoló éleken keresi az affin régiókat, az IBR az intenzitást veszi figyelembe. Először a képet elmoszák, majd a non-maximum suppression-t alkalmazzák. A kapott pontok pozícióját nem lehet pontosan meghatározni, azonban robusztusak az intenzitás változásra. Sugarakat lőnek ki a jellegzetes pontból minden irányban, majd pedig a következő függvényt értékelik ki az egyes sugarakon külön-külön:

$$f_I(t) = \frac{abs(I(t) - I_0)}{\max \left(\frac{\int_0^t abs(I(t) - I_0) dt}{t}, d \right)}, \quad (3)$$

ahol, t , $I(t)$, I_0 és d jelentései rendre az ívhossz, az intenzitás az adott pontban, az intenzitás maximuma és egy konstans, hogy elkerüljük a nullával való osztást.

A függvény szélsőértékét ott veszi fel, ahol az intenzitás hirtelen változik. A sugarakon megtalált pontok általában egy amorf alakot adnak, így utolsó lépésben erre az alakra ellipszist illesztnek.

TBMR. A Tree-Based Morse Regions (TBMR) algoritmust a [16] cikkben publikálták. A detektort alapját a Morse elmélet adja. A jellegzetes pontok keresése Minimum és Maximum fákkal történik. Felfogható az MSER egy fajtájaként is, de invariáns a megvilágítás változásra és kevesebb paramétert igényel.

SURF. A SURF [4] a SIFT algoritmus gyorsított változata. A Hesse mátrixot doboz szűrőkkel becsli, így gyorsabban képes képi jellemzőket megtalálni.

3. GT adat generálás

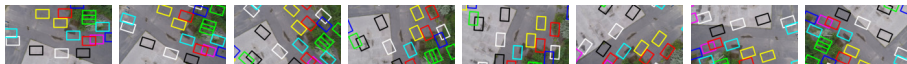
Ahhoz, hogy az affin detektorok pontosságát és az affin transzformációk megbízhatóságát kvantitatíven össze tudjuk hasonlítani, a valós (GT) transzformáció meghatározása szükséges. Más hasonló adatbázisok [11,5,17] esetében az összehasonlítás alapja a homográfia, amely csak síkszerű objektumok esetében használható. Ezzel szemben, a mi összehasonlításunk valós drónfelvételeken alapul és a GT affin transzformáció annak kötött mozgásából számítható. Az összehasonlítás motivációja, hogy az affin paraméterek sokkal pontosabban becsülhetők a kötött mozgás esetében, mint homográfia által. Sőt, ha a kamera mozgásának paraméterei is kiszámíthatóak, akkor a jellegzetes pontok képbeli mozgását is ki lehet számítani. Így lehetséges, hogy nem csak az affin transzformációkat, hanem a jellegzetes pontok detekciójának pontosságát is összehasonlíthatni.

3.1. Forgás

A forgás esetében a drón nem változtatja helyzetét, csak a vertikális tengelye körül forog. Ezért, a mozgás eltolás vektora $\mathbf{0}$. A készített képeket egy forgatás köti össze, amely kiszámítható ha annak középpontja és a szög ismertében. A 2. ábra első sorában a forgás során felvett képek láthatóak. A második sorban színezett téglalapok jelölik az azonos régiókat. A színes téglalapokat a kiszámolt forgatás alapján transzformáltunk a képek között.



(a) Példaképek a tesztesetről.



(b) A színezett téglalapok a forgatás paraméterei által lettek transzformálva. Látható, hogy ugyanazt a régiót tartalmazzák. A képek megegyeznek az előző sorban lévőekkel.

2. ábra. A képek készítése során a drón a vertikális tengelye körül forgott.

Legyen α_i a forgatás szöge az i . képen, ekkor a forgatási mátrix a következőképpen írható fel:

$$\mathbf{R}_i = \begin{bmatrix} \cos \alpha_i & -\sin \alpha_i \\ \sin \alpha_i & \cos \alpha_i \end{bmatrix}. \quad (4)$$

Ez a mátrix írja le a megfeleltethető affine régiók közötti transzformációt a forgatás esetében. Tegyük fel, hogy adottak a képeken az egymásnak megfeleltetett pontok, akkor az ezen pontok közötti transzformációt a követő egyenlettel lehet kifejezni:

$$\mathbf{p}_i^f = \mathbf{v} + \mathbf{R}_f (\mathbf{p}_i^1 - \mathbf{v}), \quad (5)$$

ahol \mathbf{p}_p^f , α_i , \mathbf{v} rendre a p -edik pont pozíciója a f -edik képen, a forgatás foka az első és a f -edik kép között és a forgatás középpontja. Utóbbi konstansnak tekinthető ezen mozgás alkalmazása során.

Ahhoz, hogy megbecsüljük a forgatás szögét és középpontját, a kiválasztott és becslt pontok közötti Euklideszi távolságot kell minimalizálni. A hibafüggvény a következőképpen írható fel:

$$\sum_{f=2}^F \sum_{i=1}^P \left\| \mathbf{p}_i^f - \mathbf{R}_f (\mathbf{p}_i^1 - \mathbf{v}) - \mathbf{v} \right\|_2^2, \quad (6)$$

ahol F és P a képek, illetve a kiválasztott pontok száma. A hibafüggvény csökkentése egy alternációs algoritmussal érhető el, amely két lépést tartalmaz. Először a forgatás középpontját, majd pedig a forgatás szögét finomítjuk. Mindezt iteratívan végrehajtva, amíg el nem érjük a hibafüggvény minimumát.

A forgás középpontjának becslése A forgás középpontjának becslése felírható $\mathbf{A}\mathbf{v} = \mathbf{b}$ alakban, ahol

$$\mathbf{A} = \begin{bmatrix} \mathbf{R}_2 - \mathbf{I} \\ \vdots \\ \mathbf{R}_F - \mathbf{I} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} \mathbf{R}_2 \mathbf{p}_1^1 - \mathbf{p}_1^2 \\ \vdots \\ \mathbf{R}_F \mathbf{p}_P^1 - \mathbf{p}_P^F \end{bmatrix}. \quad (7)$$

A legkisebb négyzetes értelemben vett optimális megoldás az \mathbf{A} mátrix pseudo-inverse által számítható: $\mathbf{v} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$.

A forgatás szögeinek becslése A forgatás szöge képpáronként külön-külön becsülhető. A becslés $\mathbf{C}\mathbf{x} = \mathbf{d}$ formába hozható, $\mathbf{x}^T \mathbf{x} = 1$ feltétellel, azaz:

$$\begin{bmatrix} x_1^1 - u & v - y_1^1 \\ y_1^1 - v & x_1^1 - u \\ \vdots & \vdots \\ x_P^1 - u & v - y_P^1 \\ y_P^1 - v & x_P^1 - u \end{bmatrix} \begin{bmatrix} \cos \alpha_i \\ \sin \alpha_i \end{bmatrix} = \begin{bmatrix} x_1^f - u \\ y_1^f - v \\ \vdots \\ x_P^f - u \\ y_P^f - v \end{bmatrix}, \quad (8)$$

ahol $[x_i^f, y_i^f]^T = \mathbf{p}_i^f$ és $[u, v]^T = \mathbf{v}$. Az optimális megoldást, azaz a forgatás szöveget egy negyedfokú polinom egyik gyöke adja. A számolás további részletei megtalálhatóak a Függelékben.

Konvergencia. A fenti két alternációs lépés egymásután iteratívan ismételjük a konvergencia eléréséig. A konvergencia sebessége nem jelentős az algoritmus felhasználása miatt. Azonban, tapasztalataink alapján a konvergencia 10-20 iterációs után beáll, ha a kép középpontját adjuk meg a forgatás középpontjának kiinduló értékeként.

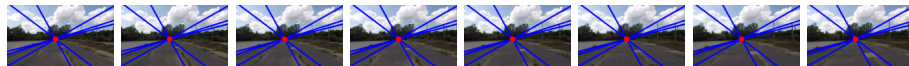
GT affin transzformációk A valós affin transzformációt az affine régiók között a forgatás mátrix adja.

3.2. Uniform Mozgás: Elülső nézet

Ebben a tesztesetben, a drón egy egyenes vonal mentén halad, miközben forgó mozgást nem végez. Ezért a kamera pozíciók közötti forgatás mátrix az egységmátrixszal egyenlő. A kamera a mozgás irányába néz, így az úgynevezett Focus of Expansion (FOE) pont kiszámítható. A FOE nem más, mint a mozgás irányvektorának a végtelenben vett vetülete a kameraképre. Ez a pont egyben egy epipólus is, ezért az egyenes mozgás során minden képen ugyanott található és minden epipoláris egyenes itt metszi egymást. Az epipoláris egyenesek meghatározhatóak, ha ismertek ugyanannak a térbeli pontnak a vetületei az egymást követő kameraképeken. A 3. ábra első sorában láthatóak példa képek erre a mozgásra. A második sorban pedig a piros pont a FOE-t, a kék vonalak pedig néhány epipoláris egyenest jelölnek.



(a) Néhány kép a tesztesetből.



(b) A piros pont FOE-t, a kék vonalak pedig néhány epipoláris egyenest jelölnek. A képek megegyeznek az első sorban láthatóakkal.

3. ábra. A képek készítése során a drón egyenesen előre haladt, párhuzamosan a talajjal.

A FOE-t numerikusan meg lehet határozni, az azonos térbeli pontok kameraképre vett vetületeinek az epipoláris vonalaktól számított négyzetes távolságának minimalizálása által [7]. Ezért, a minimalizálandó hibafüggvény a kijelölt pontok távolságát tartalmazza a hozzájuk tartozó epipoláris egyenestől:

$$\sum_{f=1}^F \sum_{i=1}^P \left((\mathbf{p}_i^f - \mathbf{m})^T \begin{bmatrix} -\sin \beta_i \\ \cos \beta_i \end{bmatrix} \right)^2, \quad (9)$$

ahol \mathbf{m} és β_i rendre a FOE, és az X tengellyel bezárt szöge az i -dik epipoláris egyenesnek. Vegyük észre, hogy a $[-\sin \beta_i, \cos \beta_i]^T$ nem más, mint az i -dik epipoláris egyenes normálvektora. Ez a hibafüggvény minimalizálható szintén egy alternációs algoritmus által, két lépésben. Először a FOE-t, majd pedig az epipoláris egyenesek X tengellyel bezárt szögeit pontosítva. Az alternáció során a kép középpontját használtuk fel a FOE kiindulási értékének. Az epipoláris egyenesek és az X tengely szögét pedig a következőképp számíthatjuk ki.

Epipoláris egyenesek becslése Az epipoláris egyenesek a FOE pontban metszik egymást, mert egyenes vonalú mozgást végez a drón. Sőt, ugyanannak a térbeli pont vetületeinek az epipoláris egyenesei megegyeznek a kameraképeken. Ezért minden egyenes epipoláris egyenes és a vízszintes tengely között bezárt szög felírható egy $\mathbf{Ax} = \mathbf{0}$ homogén egyenletrendszer által, a $\mathbf{x}^T \mathbf{x} = 1$ feltétellel:

$$\mathbf{A} = \begin{bmatrix} y_i^1 - m_y & m_x - x_i^1 \\ \vdots & \vdots \\ y_i^F - m_y & m_x - x_i^F \end{bmatrix}, \mathbf{x} = \begin{bmatrix} \cos \beta_i \\ \sin \beta_i \end{bmatrix}, \quad (10)$$

ahol $[m_x, m_y] = \mathbf{m}$ and $[x_i^f, y_i^f] = \mathbf{p}_i^f$ jelöli rendre a FOE és a kijelölt pontokat.

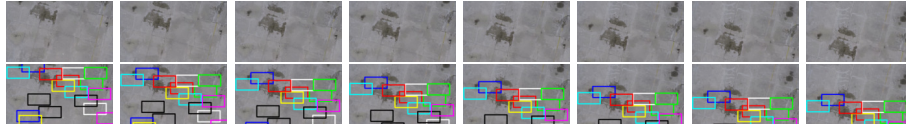
A megoldás, amely minimalizálja a hibafüggvényt, az $\mathbf{A}^T \mathbf{A}$ mátrix legkisebb sajátértékéhez tartozó sajátvektor adja. Az epipoláris egyenes és X tengely által bezárt szöget pedig ebből $\beta_i = \text{atan2}(v_y, v_x)$ alakban kapjuk.

A Focus of Expansion pont becslése A FOE az epipoláris egyenesek metszéspontjában található. Ezért, a FOE becslése egy lineáris egyenletrendszerrel $\mathbf{Am} = \mathbf{b}$ felírható, azaz

$$\begin{bmatrix} -\sin \beta_1 & \cos \beta_1 \\ -\sin \beta_1 & \cos \beta_1 \\ \vdots & \vdots \\ -\sin \beta_1 & \cos \beta_1 \\ -\sin \beta_2 & \cos \beta_2 \\ \vdots & \vdots \\ \vdots & \vdots \\ -\sin \beta_N & \cos \beta_N \end{bmatrix} \begin{bmatrix} m_x \\ m_y \end{bmatrix} = \begin{bmatrix} y_1^1 * \cos \beta_1 - x_1^1 \sin \beta_1 \\ y_1^2 * \cos \beta_1 - x_1^2 \sin \beta_1 \\ \vdots \\ y_1^F * \cos \beta_1 - x_1^F \sin \beta_1 \\ y_2^1 * \cos \beta_2 - x_2^1 \sin \beta_2 \\ \vdots \\ \vdots \\ y_N^F * \cos \beta_N - x_N^F \sin \beta_N \end{bmatrix}. \quad (11)$$

Az egyenletrendszer megoldását, a FOE pont helyzetét pedig az \mathbf{A} mátrix Pseudo inverze adja, azaz $\mathbf{m} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{Ab}$.

A Fundamentális mátrix A fenti alternációs lépések a konvergencia beállásáig ismétlődnek. A GT affin transzformációt lehetetlen kiszámítani ismeretlen környezetben, hiszen az a térbeli pont normálvektorjától függ [3], amely nem ismertek. Viszont egyéb megkötéseket lehet tenni az affin transzformációkra, ha a



4. ábra. A képek a drón a talajjal párhuzamos mozgása közben készültek, a kamera a talaj felé néz. **Első sor:** példaképek az Uniform Mozgás: Alulsó nézet tesztesetre. **Második sor:** a téglalapok képek közötti transzformációját a mozgás paraméterei adják.

fundamentális mátrix ismert. A fundamentális mátrix írja le az epipoláris egyenesek képek közötti transzformációját statikus környezet esetében. A fundamentális mátrix kiszámolható a kamera mátrix és a kameramozgás paramétereinek ismeretében, a következőképpen:

$$\mathbf{F} = \mathbf{K}^{-T} \mathbf{R} [\mathbf{t}]_x \mathbf{K}^{-1}, \quad (12)$$

ahol \mathbf{K} , \mathbf{R} és $[\mathbf{t}]_x$ rendre a kamera mátrix, a forgatás mátrix és a vektoriális szorzat mátrix-sza. A fundamentális mátrixot a FOE ismeretében is meg lehet határozni, $\mathbf{F} = [\mathbf{v}]_x$, ha a kameramozgás csak eltolást tartalmaz. Ha a fundamentális mátrix ismert két kép között, akkor legközelebbi, az epipoláris geometriának megfelelő affin transzformáció kiszámítható. Ezt fogjuk GT affin transzformációnak tekinteni ebben a speciális mozgás esetében. A fundamentális mátrixszon alapuló összehasonlítás részleteit a 4.2. Fejezetben ismertetjük.

3.3. Uniform Mozgás: Alsó nézet

A mozgás hasonló az előző fejezetben leírthoz, azonban a kamera itt a föld felé néz, nem pedig előre. Mivel a mozgás párhuzamos a talaj síkjával, ezért ez az előző degeneratív esetének tekinthető, ahol a FOE a végtelenben található. Ebben az esetben a képi jellegzetes pontok közötti kapcsolat leírható egy eltolással. Azonos térbeli pontok vetületei az egymást követő kameraképeken egyeneseket alkotnak, akár csak az előző mozgás esetében, azonban ezek az epipoláris egyenesek párhuzamosak egymással és a drón mozgása által leírt vektorral is. A 4. ábra első sora mutat példaképet a mozgásra, a második sorában a színezett dobozok pedig ugyanazokat a régiókat jelölik, amelyek egymás eltolásából származnak.

Jelöljük az epipoláris egyenesek és az X tengely által bezárt szöget γ -val, és legyen l^i az i -edik epipoláris egyenesen fekvő pont. A hibafüggvény tartalmazza a kijelölt pontok négyzetes távolságát a hozzájuk tartozó epipoláris egyenestől:

$$\sum_{f=1}^F \sum_{i=1}^P \left((\mathbf{p}_i^f - \mathbf{l}^i)^T \begin{bmatrix} -\sin \gamma \\ \cos \gamma \end{bmatrix} \right)^2. \quad (13)$$

Az alternációs algoritmus, amely minimalizálja a fenti hibafüggvényt, először az epipoláris egyeneseket szögét (γ), majd pedig a hozzájuk tartozó pontokat (\mathbf{l}^i) finomítja. Az első alternációs lépés felírható egy $\mathbf{A}\mathbf{x} = 0$ homogén lineáris egyenletként, $\mathbf{x}^T \mathbf{x} = 1$ megkötéssel, hasonlóan a korábbi 10. Egyenlethez. Viszont a



5. ábra. A képek készítése közben a drón süllyedő mozgást végzett. **Első sor:** Példa képek a skálázás teszteléséhez. **Második sor:** A színezett téglalapok ugyanazt a régiókat jelölik. A képek között a a téglalapok a skálázás paramétereit által lettek transzformálva.

mostani tartalmazza az összes pontot az összes képen:

$$\begin{bmatrix} x_1^f - l_x^1 & y_1^f - l_y^1 \\ \vdots & \vdots \\ x_2^f - l_x^2 & y_2^f - l_y^2 \end{bmatrix} \begin{bmatrix} -\sin \gamma \\ \cos \gamma \end{bmatrix} = 0, \quad (14)$$

a megoldást az $\mathbf{A}^T \mathbf{A}$ mátrix legkisebb sajátértékéhez tartozó sajátvektor adja, majd pedig $\gamma = \text{atan2}(v_x, v_y)$.

A második alternációs lépés az epipoláris egyenesek pontjait finomítja. Ez felírható $\mathbf{A} \mathbf{l} = \mathbf{b}$ alakban, ahol

$$\begin{bmatrix} \cos \gamma & \sin \gamma \\ \vdots & \vdots \\ \cos \gamma & \sin \gamma \end{bmatrix} \begin{bmatrix} l_x^i \\ l_y^i \end{bmatrix} = \begin{bmatrix} x_i \cos \gamma + y_i \sin \gamma \\ \vdots \\ x_i \cos \gamma + y_i \sin \gamma \end{bmatrix}, \quad (15)$$

a pont az epipoláris egyenesen pedig az \mathbf{A} mátrix Pseudo inverzeként írható fel, azaz $\mathbf{l} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A} \mathbf{b}$.

GT Affin transzformáció A valós, GT affin transzformáció ezen mozgás esetében pedig nem más, mint az egységmátrix: $\mathbf{A} = \mathbf{I}$.

3.4. Skálázás

A skálázást a drón süllyedésével szimuláljuk, miközben a kamera a talaj felé néz. A mozgás iránya megközelítőleg merőleges a kamera képsíkjára, ezért ez a fajta mozgás felfogható a Uniform Mozgás: elülső nézet egy speciális esetének. Az egyetlen különbség, hogy a talaj síkszerű, ezért a mozgás paramétereit és az affin transzformációkat ki lehet számolni. A 5. ábrán első sorába láthatóak példaképek erre a mozgástípusra.

Mivel a mozgás iránya és a kamera síkja megközelítőleg merőleges egymásra, így a FOE és az epipoláris egyeneseket meg lehet határozni. Hasonlóan, mint a korábbi elülső nézetes mozgás esetében, a jellegzetes pontok az epipoláris egyenesek mentén mozognak. A FOE és az epipoláris egyenesek kiszámítását a 3.2. Fejezetben részleteztük.

Az egymásnak megfeleltethető jellegzetes pontok az egymást követő képeken felírhatók a skálázás paraméterét használva. Jelölje s ezt a skálázási paramétert, ekkor a jellegzetes pontok FOE-től való távolsága felírható a következőképpen:

$$(\mathbf{p}_i^2 - \mathbf{v}) = s_2 (\mathbf{p}_i^1 - \mathbf{v}) \quad i \in [1, P]. \quad (16)$$

Ezért, a skálázási paraméter kiszámítható, ha átlagoljuk a képeken található jellegzetes pontok távolságait, azaz:

$$s_2 = \frac{1}{P} \sum_{i=1}^P \frac{\|\mathbf{p}_i^2 - \mathbf{v}\|_2}{\|\mathbf{p}_i^1 - \mathbf{v}\|_2}. \quad (17)$$

GT affin transzformáció A GT affin transzformáció ebben az esetben a skálázás paramétere adja, azaz $\mathbf{A}_j = s\mathbf{I}$.

4. Kiértékelési Módszerek

A jellegzetes pont és affin detektorokat kétféleképpen hasonlítjuk össze. Az első összehasonlításban a jellegzetes pontok detektálásának és az affin transzformációk pontosságának hibáját mérjük. Ahhoz, hogy ezeket össze tudjuk kvantitatíven hasonlítani, a kamera mozgást kell megbecsülnünk. Ez kiszámítható, ha a kamera speciális mozgást végez, ennek részleteit az előző fejezet tartalmazza. Azonban, a Uniform Mozgás: Elülső nézet esetében ez nem tehető meg. Ezért egy második összehasonlítást is elvégeztünk, amelyben a fundamentális mátrixot használtuk.

4.1. Affin Összehasonlítás

Az affin összehasonlítás során a jellegzetes pontok detektálásának és az affin transzformáció pontosságának a hibáját mérjük.

Detekciós hiba. A GT jellegzetes pont meghatározható, ha ismert annak pozíciója az előző képen, illetve a mozgás paraméterei. A különböző mozgásokat, és annak paramétereinek kiszámítását, az előző fejezetben részleteztük. A jellegzetes pontok detekciós hibáján, a pontok és a GT kiszámított pontok közötti Euklideszi távolságot értjük.

Affin hiba. Amíg a detekciós hiba az Euklideszi távolságon alapszik, az affin transzformációk hibáját a GT és a detektor által becsült affin mátrixok különbségének a Frobenius normája fogja adni. Formálisan, A Frobenius norma geometriai jelentéssel bír az affin transzformációk esetében, a részleteket a [2]-ben találhatók.

4.2. Fundamentális Mátrix Kiértékelés

A Uniform Mozgás: Elülső nézet esetében az affin transzformációk kiszámítása nem lehetséges. Azon objektumok, melyek közelebb helyezkednek a kamerához több pixelt mozdulnak el két szomszédos kép között, mint amelyek távol. Viszont a mozgás iránya behatárolható az epipoláris egyenesek által, melyek egymást a FOE-ben metszik. A fundamentális mátrix nagy pontossággal kiszámolható a FOE-ből, $\mathbf{F} = [\mathbf{v}]_x$, és felhasználható az affin transzformációk finomítására.

A finomított affin transzformációkat tekintjük GT adatnak, és ezek segítségével számoljuk az affin transzformációk hibáját. A fundamentális mátrix általi affin transzformáció finomítása[2] egy lineáris probléma, amely egy hatodfokú polinom gyökeinek megkeresésével ekvivalens. A hiba mértékét, a javított és az eredeti affin transzformációk különbségmátrixának a Frobenius normálja adja.

1. táblázat. Az oszlopok a módszereket, míg a sor hármasok a teszteseteket mutatják. Az **első sorokban** a jellegzetes pontok száma, a **második sorokban** a párosított pontok száma, a **harmadik sorokban** pedig a futási idők láthatóak.

		EBR	HARAFF	HARHES	HARLAP	HESAFF	HESLAP	IBR	MSER	SEDGELAP	SURF	TBMR
Skálázás	# Összes	217	18720	21198	19033	3866	3994	1277	336	20132	783	4122
	# Megfeleltetett	91	8369	9544	8534	1702	1757	478	235	10071	636	2259
	Futási idő (s)	20.43	3.27	2.32	2.06	1.19	0.97	7.59	0.41	3.29	0.65	0.68
Forgás	# Összes	114	12686	14266	12795	2655	2751	892	223	13173	508	3141
	# Megfeleltetett	33	5235	5943	5294	1070	1111	395	133	5872	390	1447
	Futási idő (s)	18.01	2.21	1.69	1.56	0.91	0.77	6.27	0.35	2.41	0.54	0.54
Alsó nézet1	# Összes	59	7514	8122	7629	1211	1216	441	92	7222	204	2688
	# Megfeleltetett	27	3450	3730	3503	522	524	226	64	3743	169	1049
	Futási idő (s)	18.09	1.69	1.37	1.30	0.80	0.73	5.52	0.38	1.85	0.51	0.67
Alsó nézet2	# Összes	26	12484	13020	12615	1119	1135	588	117	9461	139	3051
	# Megfeleltetett	8	4451	4681	4506	418	426	241	68	4094	107	1037
	Futási idő (s)	20.49	2.22	1.65	1.64	0.78	0.72	5.25	0.36	2.11	0.57	0.95
Elülő Nézet1	# Összes	78	10128	13639	10415	4708	5032	695	233	14869	803	2083
	# Megfeleltetett	36	4623	6476	4758	2444	2583	416	164	7650	653	1058
	Futási idő (s)	17.97	3.16	2.12	1.66	1.70	1.19	8.47	0.29	2.67	0.81	0.76
Elülő Nézet2	# Összes	59	9210	11193	9369	2108	2812	592	50	12345	511	1856
	# Megfeleltetett	30	4615	5764	4703	1192	1617	398	38	7275	447	1034
	Futási idő (s)	15.56	2.07	1.65	1.48	1.01	0.87	6.58	0.29	2.31	0.65	0.55

5. Összehasonlítás

Tizenegy affin jellegzetes pont detektort hasonlítottunk össze a teszt során. A detektorok eredeti, szerző által közzétett implementációját a Visual Geometry Group, oxfordi egyetem honlapjáról töltöttük le², kivétel a TBMR esetében amelyet a szerző saját honlapjáról értünk el³. A legtöbb detektort már ismertettük a 2. Fejezetben. Ezekhez képest a HARHES és SEDGELAP detektorokkal kiegészítettük az összehasonlítást. Előbbi a Harris-Affine és Hessian-Affine kombinációja, utóbbi pedig az élek mentén keresi az affin régiókat, a 2. Fejezetben ismertetett módszer alapján.

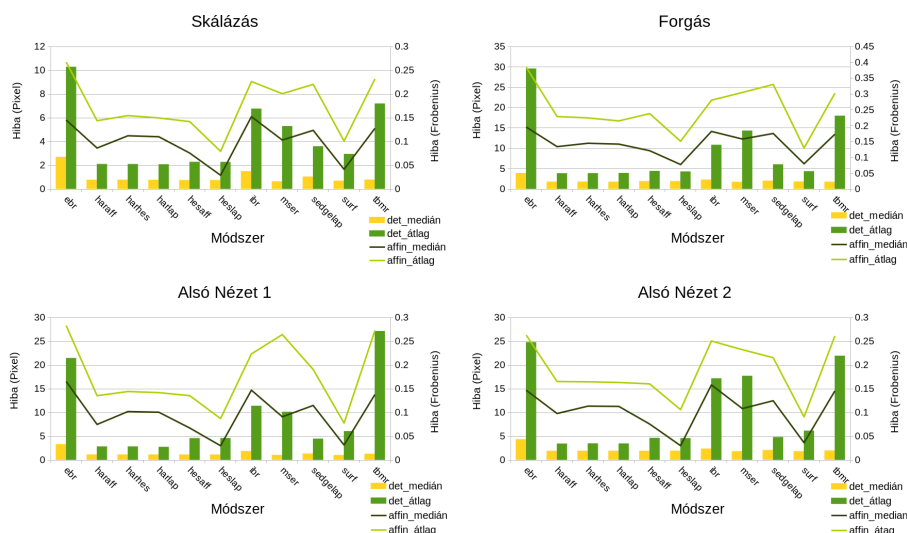
Miután a jellegzetes pontokat kinyertük a képekből, ezek megfeleltetése következik. Az outlier-ek kiszűrésére az ismert hányados tesztet [9] alkalmaztuk, amely a legközelebbi és a második legközelebbi leíró vektor távolsága alapján szűr. Az affin transzformációkat a szűrt jellegzetes pontok között határoztuk meg. A detektorok csak ellipszis alakú affin régiót képesek detektálni, azok orientációja nélkül. A hiányzó orientációt a SIFT leíró domináns iránya adja. Végül, az affin transzformáció két affin régió között a következőképp számolható:

$$\mathbf{A} = \mathbf{A}_2 \mathbf{R}_2 (\mathbf{A}_1 \mathbf{R}_1)^{-1}, \quad (18)$$

ahol $\mathbf{A}_i, \mathbf{R}_i$ $i \in [1, 2]$ rendre az affine régiók és a hozzájuk tartozó domináns irányok a SIFT leíróból kinyerve. Az 1. Táblázat tartalmazza a jellegzetes pontok és a megfeleltetett pont jellemzők számát, valamint a mért futási időket. Az SEDGELAP és HARHES detektorok találják a legtöbb jellemző pontot, viszont a sok pont bonyolítja a megfeleltetés folyamatát és növeli annak időigényét. Általánosságban, az EBR, IBR és MSER több száz, a Harris alapú módszerek (HARAFF, HARLAP) 10 – 20 ezer, míg a Hessian alapú módszerek (HESAF, HESLAP) néhány ezer jellegzetes pontot detektálnak. Az outlier szűrés [9] során előfordul, hogy egyes jellegzetes pontok nem kerülnek megfeleltetésre. A táblázat második soraiban található az outlier szűrés után kapott pontok száma. Vegyük

² www.robots.ox.ac.uk/vgg/research/affine/index.html

³ <http://laurentnajman.org/index.php?page=tbmr>



6. ábra. Az affin összehasonlítás eredményei. A detektálás hibáját az oszlopdiagrammok mutatják a bal oldali tengelyen. Az affin hibát a vonaldiagrammok mutatják, skálája a jobb oldali oszlop.

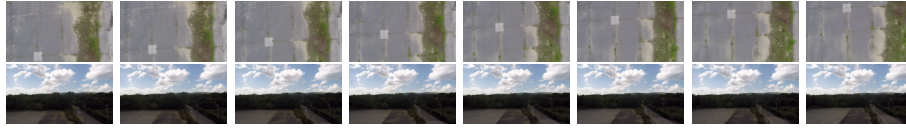
észre, hogy körülbelül a jellegzetes pontok 50%-a elvész a megfeleltetés során. A harmadik sorok a futási időket tartalmazzák, 5MP felbontású képeket esetében. Egyértelmű, hogy az MSER a leggyorsabb az összes közül, őt a SURF és a TBMR követi a sorban. A Hessian alapú módszerek 1, míg a Harris alapú módszerek 1.5-2 másodpercig futottak. A lelassabb, az IBR és az EBR. Felhívjuk a figyelmet, hogy párhuzamosított vagy GPU implementációt felhasználva valószínűleg más eredményt kaptunk volna.

5.1. Affin Kiértékelés

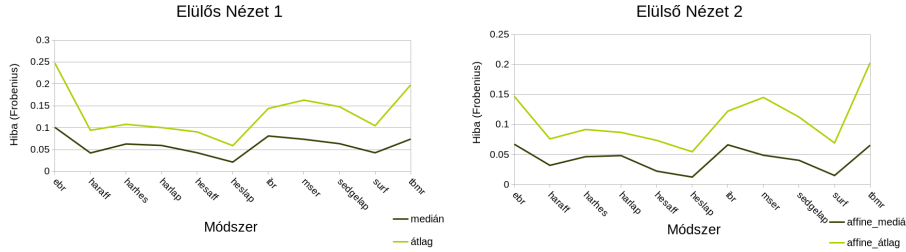
Az első kiértékelés becsült kamera mozgást és az affin transzformációkat használja, melyeket a 3. Fejezetben részleteztünk. A kvantitatív kiértékelés kettős, hiszen a jellegzetes pontok detektálásának és az affin transzformációk pontosságának hibáját is mérni tudjuk. Négy teszt esetet vizsgáltunk meg. Ezek közül az egy a skálázás, egy a forgás és kettő teszteset pedig az Uniform Mozgás: alsó nézetet képviseli. A 2. Ábra a forgást, és a 5. Ábra a skálázás, míg a 4 és 7. Ábrák első sora a Uniform Mozgás:Alsó Nézet tesztesetét szemléltetik.

A 6. Ábra foglalja össze a detektorok hibáját. A pont detektálás átlag és medián hibáját az oszlopdiagram szemlélteti, melynek mértéke – a valós és becsült pont közötti Euklideszi távolság – a bal oldali tengelyen látható. Az affin transzformációk átlag és medián hibáját pedig rendre a zöld és fekete vonaldiagrammok szemléltetik. Ennek mértéke a jobb oldali tengelyen látható, ami a valós és becsült affin transzformációk különbségének a Frobenius normálja.

A 6. Ábra diagramjai hasonló eredményeket mutatnak az összes tesztesetben. A detekciós hiba minden a Hessian és a Harris alapú módszerek esetében



7. ábra. A képeket a drón a talajjal párhuzamos mozgás közben végezte. **Első sor:** Alulnézet 2 teszt eset. A kamera a talaj nézi Alulnézet 2 teszt eset. **Második sor:** Elülső Nézet teszt eset. A kamera a mozgás irányába néz.



8. ábra. A fundamentális összehasonlítás hibája.

a legkisebb, ami azt jelenti, hogy ezek a sokkal pontosabban képesek a képi jellegzetes pontokat megtalálni, mint a többiek. Ez a hiba az EBR, IBR és MSER esetében a legnagyobb. Az átlag hiba 10 pixelnél is nagyobb lehet, kivéve a skálázás teszt esetében. Vegyük észre, hogy a medián értékek alacsonyabban, mint az átlagok. Ez valószínűleg abból adódik, hogy az outlier szűrés ellenére néhány hibás megfeleltetést nem sikerült kiszűrni, amik félrehúzzhatják az átlagot. Meglepő módon, az affin hiba a HESLAP és a SURF esetében a legkisebb, míg az EBR, IBR és TBMR esetében a legnagyobb.

5.2. Fundamentális Mátrix Kiértékelés

A fundamentális mátrixot használó kiértékelés során a drón előre mozog a talajjal párhuzamosan, míg a kamera szintén előre néz. A valós affin transzformáció és a pontok mozgás nem megállapítható. Ezért, a fundamentális mátrix segítségével finomítjuk az affin transzformációkat. A hibát, a mért és a finomított affin transzformációk különbségének a Frobenius normálja adja.

Két teszt esetet végeztünk. Az első példaképei a 3. Ábrán, a második tesztet tartozó példaképek a 7. Ábra második sorában láthatóak. Az első esetben a drón kb. két méterrel a talaj felett, a második esetben pedig a fák tetejének magasságában mozgott.

A 8. Ábra diagramjai mutatják az affin transzformációk hibáját a két teszt esetben. A HARRAFF, HARLAP és SURF affin transzformáció vétik a legkevesebb hibát, így ezek megbízhatónak bizonyulnak. A hibák karakterisztikája megegyezik az előző összehasonlításban látottakkal.

6. Konklúzió

Ebben a cikkben a legelterjedtebb affin transzformációt is detektáló algoritmusokat hasonlítottuk össze kvantitatíven. A fő újdonság, hogy az összehasonlítás va-

lós, drón által készített képeken alapszik. A tesztképeken komplexebb környezet figyelhető meg, mint a korábbi, csak homográfián alapuló összehasonlítások esetében. Mellékesen, a pont detektáló algoritmusok pontossága is összehasonlításra kerültek, hiszen ezek szükségesek az affin transzformációk meghatározásához is.

Legfontosabb meglátásunk, hogy az affin detektáló algoritmusok megbízhatósága nem függ szorosan a környezettől. A tesztek alapján a SURF, Harris, illetve Hessian alapú módszereket ajánjuk.

Hivatkozások

1. P. F. Alcantarilla, A. Bartoli, and A. J. Davison. Kaze features. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, editors, *Computer Vision – ECCV 2012*, pages 214–227, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.
2. D. Barath, J. Matas, and L. Hajder. Accurate closed-form estimation of local affine transformations consistent with the epipolar geometry. In *Proceedings of the British Machine Vision Conference 2016, BMVC, York, UK 2016*, 2016.
3. D. Barath, J. Molnár, and L. Hajder. Optimal surface normal from affine transformation. In *VISAPP 2015 - Proceedings of the 10th International Conference on Computer Vision Theory and Applications, Volume 3, Berlin, Germany, 11-14 March, 2015.*, pages 305–316, 2015.
4. H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110(3):346–359, June 2008.
5. K. Cordes, B. Rosenhahn, and J. Ostermann. High-resolution feature evaluation benchmark. In R. Wilson, E. Hancock, A. Bors, and W. Smith, editors, *Computer Analysis of Images and Patterns*, pages 327–334. Springer Berlin Heidelberg, 2013.
6. C. Harris and M. Stephens. A combined corner and edge detector. In *In Proc. of Fourth Alvey Vision Conference*, pages 147–151, 1988.
7. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, New York, NY, USA, 2 edition, 2003.
8. S. Leutenegger, M. Chli, and R. Y. Siegwart. Brisk: Binary robust invariant scalable keypoints. In *Proceedings of the 2011 International Conference on Computer Vision, ICCV '11*, pages 2548–2555, Washington, DC, USA, 2011. IEEE.
9. D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, Nov. 2004.
10. K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors, *Computer Vision – ECCV 2002*, pages 128–142, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.
11. K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool. A comparison of affine region detectors. *International Journal of Computer Vision*, 65(1):43–72, Nov 2005.
12. Z. Pusztai and L. Hajder. Quantitative comparison of affine invariant feature matching. In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017) - Volume 6: VISAPP, Porto, Portugal, 2017.*, pages 515–522, 2017.
13. J. Shi and C. Tomasi. Good features to track, 1994.
14. T. Tuytelaars and K. Mikolajczyk. Local invariant feature detectors: A survey. *Found. Trends. Comput. Graph. Vis.*, 3(3):177–280, July 2008.
15. T. Tuytelaars and L. Van Gool. Matching widely separated views based on affine invariant regions. *International Journal of Computer Vision*, 59(1):61–85, 2004.

16. Y. Xu, P. Monasse, T. Géraud, and L. Najman. Tree-based morse regions: A topological approach to local feature detection. *IEEE Transactions on Image Processing*, 23(12):5612–5625, Dec 2014.
17. L. Zitnick and K. Ramnath. Edge foci interest points. International Conference on Computer Vision, January 2011.

FÜGGELÉK

Célunk, hogy megmutassuk, hogy az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer, $\mathbf{x}^T \mathbf{x} = 1$ feltétellel megoldható. A λ Lagrange multiplikátort használva az egyenletrendszer átírható a következőképpen:

$$J = (\mathbf{Ax} - \mathbf{b})^T (\mathbf{Ax} - \mathbf{b}) + \lambda \mathbf{x}^T \mathbf{x}.$$

Az optimális megoldást a költségfüggvény \mathbf{x} szerinti deriváltja adja:

$$\frac{\partial J}{\partial \mathbf{x}} = 2\mathbf{A}^T (\mathbf{Ax} - \mathbf{b}) + 2\lambda \mathbf{x} = 0.$$

Ezért az optimális megoldás felírható $\mathbf{x} = (\mathbf{A}^T \mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^T \mathbf{b}$ alakban.

Az egyszerűség kedvéért vezessük be a $\mathbf{v} = \mathbf{A}^T \mathbf{b}$ vektort és $\mathbf{C} = \mathbf{A}^T \mathbf{A}$ szimmetrikus mátrixot. Ekkor $\mathbf{x} = (\mathbf{C} + \lambda \mathbf{I})^{-1} \mathbf{v}$. Végül az $\mathbf{x}^T \mathbf{x} = 1$ megkötés figyelembevételével: $\mathbf{v}^T (\mathbf{C} + \lambda \mathbf{I})^{-T} (\mathbf{C} + \lambda \mathbf{I})^{-1} \mathbf{v} = 1$.

Definíció által tudjuk, hogy:

$$(\mathbf{C} + \lambda \mathbf{I})^{-1} = \frac{\text{adj}(\mathbf{C} + \lambda \mathbf{I})}{\det(\mathbf{C} + \lambda \mathbf{I})}.$$

Tehát, ha $\mathbf{C} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$, akkor $\mathbf{C} + \lambda \mathbf{I} = \begin{bmatrix} c_1 + \lambda & c_2 \\ c_3 & c_4 + \lambda \end{bmatrix}$.

A $\mathbf{C} + \lambda \mathbf{I}$ determináns és konjugált transzponált mátrix átírható:

$$\det(\mathbf{C} + \lambda \mathbf{I}) = (c_1 + \lambda)(c_4 + \lambda) - c_2 c_3, \quad \text{adj}(\mathbf{C} + \lambda \mathbf{I}) = \begin{bmatrix} c_4 + \lambda & -c_2 \\ -c_3 & c_1 + \lambda \end{bmatrix}$$

és

$$\begin{bmatrix} c_4 + \lambda & -c_2 \\ -c_3 & c_1 + \lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \lambda + c_4 v_1 - c_2 v_2 \\ v_2 \lambda + c_1 v_2 - c_3 v_1 \end{bmatrix}.$$

Mindazonáltal a $\mathbf{v}^T (\mathbf{C} + \lambda \mathbf{I})^{-T} (\mathbf{C} + \lambda \mathbf{I})^{-1} \mathbf{v} = 1$ kifejezés átírható:

$$\mathbf{v}^T \frac{\text{adj}^T(\mathbf{C} + \lambda \mathbf{I}) \text{adj}(\mathbf{C} + \lambda \mathbf{I})}{\det(\mathbf{C} + \lambda \mathbf{I}) \det(\mathbf{C} + \lambda \mathbf{I})} \mathbf{v} = 1,$$

$$\mathbf{v}^T \text{adj}^T(\mathbf{C} + \lambda \mathbf{I}) \text{adj}(\mathbf{C} + \lambda \mathbf{I}) \mathbf{v} = \det^2(\mathbf{C} + \lambda \mathbf{I}).$$

Az egyenlet mindét oldala polinomokat tartalmaz. A fokszámok rendre $2n - 2$ és $2n$. Ha a kifejezés egyik oldalát kivonjuk a másiktól, akkor egy $2n$ fokú polinomot kapunk. Figyeljük meg, hogy síkbeli mozgás esetén $n = 2$. Az optimális megoldást tehát e polinom valós gyökei adják. Az λ_i , $i \in 1, 2$ -nek megfelelő vektor felírható: $\mathbf{g}_i = (\mathbf{L} + \lambda_i \mathbf{I})^{-1} \mathbf{r}$. Majd a legkisebb $\|\mathbf{F} \mathbf{g}_i - \mathbf{h}\|$ normát adó fogja az optimálist megoldást adni.