Kamera-kalibráció felületi normálisok felhasználásával

Eichhardt Iván[1,2], Hajder Levente[1,2]

[1] MTA SZTAKI, Budapest
{eichhardt.ivan,hajder.levente}@sztaki.mta.hu
[2] Eötvös Loránd Tudományegyetem, Budapest

Kivonat Klasszikus kamera-kalibrációs algoritmusok kizárólag pontok közötti összefüggéseket használnak, azonban a kalibrációs képek több információt tartalmaznak, mint néhány 2D koordinátapár. Ezen megközelítés alapján bemutatunk egy eljárást, amely képi vetületek közötti lokális affin összefüggéseket, illetve a kalibrációs objektum felületi normálisait képes felhasználni a kalibrációs probléma zárt alakú megoldására. A kezdetben becsült paramétereket numerikusan is finomítjuk. Az algoritmusaink pontosságának validálása szintetikus adatokon történik. Újszerű megközelítésünk valós alkalmazhatóságát egy 3D strukturált fény alapú szkenner kalibrációján keresztül mutatjuk be.

1. Bevezetés

A többnézetű geometria számos témája [5] között a kamera-kalibrációt gyakran tekintik megoldott problémának. Mérnökök és kutatók előszeretettel alkalmazzák a jól ismert, sok éve publikált, Zhang-féle kalibrációs eljárást [10]. Számos kereskedelmi és ingyenes számítógépes látással kapcsolatos függvénykönyvtár alkalmazza ezt a hatékony és egyszerű kalibrációs eljárást. Alternatívaként használhatunk egy 3D kalibrációs objektumot, ha ismertek az objektum és a kamerakép közötti pontszerű 3D→2D összefüggések. Ebben az esetben az úgynevezett Direct Linear Transformation [5] (DLT) segítségével egy kellően pontos becslést tehetünk a perspektív vetítés paramétereire, melyet numerikusan tovább finomíthatunk (pl. a Levenberg-Marquardt eljárással [7,9]).

A hagyományos kamera-kalibrációs eljárások kizárólag pontpárok közötti összefüggéseket alkalmaznak, kameraképek [10], vagy egy ismert geometriájú térbeli objektum és annak kamera-képe között [5]. Cikkünkben szemléltetjük, hogy utóbbi esetben a felületi normális további hasznos információt szolgáltat a kalibráció számára, *mely segítségével pontosabb becslést kaphatunk a kérdéses paraméterekre.*

A legjobb tudásunk szerint újszerű ötlet a felületi normálisokat alkalmazni belső és külső paraméterek becslésére. A külső paraméterek kiszámíthatóak

Levente Hajder, Ivan Eichhardt, Improvement of Camera Calibration using Surface Normals, ICPR2016 képeken szereplő síkok közötti homográfiák felbontásából [8,4] a felületi normális ismerete nélkül. A felületi normális is megbecsülhető, de a homográfia-dekompozíció egyik jelentős hátránya, hogy nem egyértelmű [6]. Egy másik hátulütője, hogy a homográfia alapú kamera-kalibráció kifejezetten zajérzékeny [11].

Munkánkat főként Barath et al. [1] tanulmánya inspirálta, ahol bemutatták, hogy az egymásnak megfelelő képfoltok közötti affin transzformáció leírható a sztereo-rendszer paraméterei és a hozzájuk tartozó felületi normális ismeretével. Bemutatott eljárásainkkal célunk, hogy pontos, zárt-alakú becslést adjunk a kameraparaméterekre a rendszer második kamerája esetén, ahol az első kamerát kalibráltnak tekintjük, továbbá egy felületről képfoltok ismertek 3D felületi normálisokkal és pozíciókkal.

Cikkünk a témához való fő hozzájárulása egy újszerű elmélet, mely bemutatja, hogyan becsülhetőek pontos kameraparaméterek, ismert képfoltokhoz rendelt normálisok és közöttük lévő affin transzformációk alapján. Állításainkat valós és szintetikus tesztekkel támasztjuk alá.

Algoritmusaink bemenete (i) a felületi normálisok a megfigyelt helyeken, (ii) a hozzájuk tartozó 3D pozíciók és (iii) vetületeik 2D helyzetei és közöttük lévő affin transzformációk. Ennek következtében, az algoritmus elsősorban speciális kalibrálási tárgyak használata mellett alkalmazható.

2. Jelölések

Általános megjegyzések. Ebben a munkában a mátrixokat félkövér nagybetűkkel, vektorokat pedig – a gradiensektől eltekintve – aláhúzással jelöljük. Egy felső index a legtöbbször a képi nézet sorszáma, kivéve, ha az zárójelek között van, ekkor az a megfigyelt képfolt számozása.

Legyen $\mathbf{B}$ egy $\mathbb{R}^{a \times b}$ mátrix, $i \leq j$, $m \leq n$ ahol $i, j \in \{1, \dots, a\}$ és $m, n \in \{1, \dots, b\}$. Az alábbi kifejezés $\mathbf{B}$ egy *al-mátrixát* jelöli ami a soraiból $\{i,\, i+1, \dots,\, j\}$ és oszlopaiból $\{m,\, m+1, \dots,\, n\}$ vett közös elemekből áll:

$$\mathbf{B}|_{(i:j,\, m:n)} \in \mathbb{R}^{(j-i+1)\times(n-m+1)} \tag{1}$$

Az egyszerűség kedvéért, akárcsak a Matlab-ban, az alábbi jelöléseket használjuk, ha a kérdéses mátrix összes sorát vagy oszlopát vesszük:

$$\mathbf{B}|_{(:,\, m:n)} = \mathbf{B}|_{(1:a,\, m:n)}\,,$$

és

$$\mathbf{B}|_{(i:j,\, :)} = \mathbf{B}|_{(i:j,\, 1:b)}\,.$$

Legyen $\mathtt{vec}\,(\mathbf{A})$ az $\mathbf{A}$ mátrix elemeinek oszloponkénti felsorolása:

$$\mathtt{vec}\,(\mathbf{A}) \coloneqq \left[\, \mathbf{A}|^T_{(:,\,1)}\ \mathbf{A}|^T_{(:,\,2)}\ \dots\ \mathbf{A}|^T_{(:,\,m)} \right]^T.$$

3. Elméleti háttér

Munkánk elméleti háttere Barath et al. [1] cikkén alapszik, ahol bemutatták, hogy a 3D→2D vetítőfüggvények a térkoordinátákra vett gradiensei és a felületi normális közötti összefüggés a következőképpen írható le:

$$\begin{bmatrix} |\underline{n}\,\nabla y_1\,\nabla x_2| & |\underline{n}\,\nabla x_2\,\nabla x_1| \\ |\underline{n}\,\nabla y_1\,\nabla y_2| & |\underline{n}\,\nabla y_2\,\nabla x_1| \end{bmatrix} = c\mathbf{A}, \tag{2}$$

ahol $\nabla\underline{u}_2 = \begin{bmatrix} \nabla x_2 \\ \nabla y_2 \end{bmatrix}$ a vetítőfüggvények gradienseit jelöli,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

a két kamera képén vett képfolt közötti affin transzformáció. $c = |\nabla x_1\,\underline{n}\,\nabla y_1|$ pedig egy rövidítés, ahol $|...|$ felületi normálisok és vetítőfüggvények 2D koordinátakomponenseire vett gradiens-vektorai közötti vegyes szorzat.

A vegyes szorzat tulajdonságait felhasználva a következő lineáris egyenletrendszer hozható létre a (2)-es egyenlet egy kompaktabb alakjaként:

$$\nabla\underline{u}_2 \cdot \mathbf{N} = c\mathbf{A} \tag{3}$$

ahol

$$\mathbf{N} = \begin{bmatrix} \underline{n} \times \nabla y_1 \mid -\underline{n} \times \nabla x_1 \end{bmatrix} \in \mathbb{R}^{3\times 2}. \tag{4}$$

A (3)-ik egyenlet mutatja, hogy az első kamera paraméterei (∇x_1-ben és ∇y_1-ben) elválaszthatóak a második kameráétól ($\nabla\underline{u}_2$-ben). Továbbá ez a viszony linearizálható a koordináták gradienseit tekintve.

4. A projekciós mátrix becslése

A kalibrációs probléma bemeneteként a következőek adottak:

Kameramátrix: P^1;

$(\forall i \in \{1 \dots S\})$:

Minta-pontok: $\underline{u}_1^{(i)}$, $\underline{u}_2^{(i)}$ képpont-párok, $\underline{X}^{(i)}$ 3D pontok;
Felületi normálisok: $\underline{n}^{(i)}$.
Affin transzformációk: $\mathbf{A}^{(i)}$.

Ebben a részben a második kamera becslését taglaljuk. A kamerát egy 3×4-es projekciós mátrixszal reprezentáljuk.

A perspektív kamera koordinátafüggvényeinek gradiensét az A Függelékben írjuk le. Ez alapján az i-ik mintára ($i \in \{1 \dots S\}$) a következő linearizált alakot kapjuk:

$$\left(\mathbf{P}^2\big|_{(1:2,\,1:3)} - \underline{u}_2^{(i)} \cdot \mathbf{P}^2\big|_{(3,\,1:3)} \right) \cdot \mathbf{N}^{(i)} = s_2^{(i)} c^{(i)} \mathbf{A}^{(i)} \tag{5}$$

ahol $s_2^{(i)}$ a projektív mélység a második nézetből és $\mathbf{P}^2$ a kamera 3×4-es perspektív vetítési mátrixa.

A felső ($i \in \{1 \dots S\}$)-ös egyenlet ekvivalens a következővel:

$$E^{(i)}\underline{\mathrm{p}} = \underline{\mathrm{h}}^{(i)} \tag{6}$$

ahol

$$E^{(i)} = \begin{bmatrix} \mathbf{N}^{(i)}\big|_{(:,1)}^{T} & \underline{0}^T & -x_2^{(i)} \cdot \mathbf{N}^{(i)}\big|_{(:,1)}^{T} \\ \underline{0}^T & \mathbf{N}^{(i)}\big|_{(:,1)}^{T} & -y_2^{(i)} \cdot \mathbf{N}^{(i)}\big|_{(:,1)}^{T} \\ \mathbf{N}^{(i)}\big|_{(:,2)}^{T} & \underline{0} & -x_2^{(i)} \cdot \mathbf{N}^{(i)}\big|_{(:,2)}^{T} \\ \underline{0}^T & \mathbf{N}^{(i)}\big|_{(:,2)}^{T} & -y_2^{(i)} \cdot \mathbf{N}^{(i)}\big|_{(:,2)}^{T} \end{bmatrix} \tag{7}$$

$$\underline{\mathrm{h}}^{(i)} = s_2^{(i)} c^{(i)} \mathtt{vec}\left(\mathbf{A}^{(i)}\right) \tag{8}$$

Az összes mintát összekapcsolva az alábbi túlhatározott egyenletrendszert kapjuk:

$$\mathbf{E} \cdot \underline{\mathrm{p}} = \underline{\mathrm{H}} \tag{9}$$

ahol

$$\mathbf{E} = \begin{bmatrix} E^{(1)} \\ E^{(2)} \\ \vdots \\ E^{(S)} \end{bmatrix}, \quad \underline{\mathrm{H}} = \begin{bmatrix} \underline{\mathrm{h}}^{(1)} \\ \underline{\mathrm{h}}^{(2)} \\ \vdots \\ \underline{\mathrm{h}}^{(S)} \end{bmatrix}.$$

A következő alszakaszokban a (9)-es egyenletben foglaltak különböző megoldási lehetőségeit taglaljuk. Fontos megjegyezni, hogy ezen egyenletek alapján kizárólag a projekciós mátrix bal $3{\times}3$-as al-mátrixának elemeit ($\underline{\mathrm{p}} = \mathtt{vec}\left(\mathbf{P}^2\big|_{(1:3,1:3)}\right)$) becsülhetjük. Az utolsó, negyedik oszlop a megszokott DLT módszerrel [5] becsülhető. Ezen okból kifolyólag kizárólag a relatív elforgatás és a belső kameraparamétereket tudjuk első megközelítésben megbecsülni, a kamera-középpont relatív eltolásával nem tudunk számolni.

4.1. Iteratív megoldás

Ebben az alszakaszban egy iteratív (alternáló) megoldó működését írjuk le, ahol a kameraparaméterek és a projektív mélységek becslése felváltva finomodik (lásd: 1. algoritmus). A két lépés részleteiben kifejtve a következőképpen alakul.

1. **A 3×3-al al-mátrix becslése**: A (9)-es egyenlet egy ekvivalens alakjaként a következőt vesszük:

$$[\mathbf{E},\, -\underline{\mathrm{H}}] \cdot \begin{bmatrix} \underline{\mathrm{p}} \\ w \end{bmatrix} = \underline{0} \tag{10}$$

mely egy homogén lineáris egyenletrendszer, ahol $w = 1$. Ennek nem-triviális megoldása legyen $\begin{bmatrix} \hat{\underline{\mathrm{p}}}' \\ \hat{w} \end{bmatrix}$ (illetve $\hat{\underline{\mathrm{p}}} = \frac{\hat{\underline{\mathrm{p}}}'}{\hat{w}}$), a $[\mathbf{E},\, -\underline{\mathrm{H}}]^T\, [\mathbf{E},\, -\underline{\mathrm{H}}]$ mátrix legkisebb sajátértékéhez tartozó sajátvektora.

2. **Projektív mélységek becslése**: Minden mintához $s_2^{(i)}$ a következő módon kiszámítható:

$$s_2^{(i)} = vec^+ \left(c^{(i)} \mathbf{A}^{(i)} \right) \cdot vec \left(\nabla \underline{u}_2^{(i)} \cdot \mathbf{N}^{(i)} \right) \tag{11}$$

ahol $vec^+ \left(c^{(i)} \mathbf{A}^{(i)} \right)$ a $c^{(i)} \mathbf{A}^{(i)}$ mátrix a Moore-Penrose pszeutodinverze. Ezt a megoldási módot alkalmazhatjuk, mert feltételezzük, hogy a második lépés problémája lineáris $s_2^{(i)}$-re vonatkozóan.

Ezen megoldó kezdeti paramétereit a DLT algoritmus [5] segítésével nyerjük ki. Mivel az alternáló lépések egyazon nemnegatív költségfüggvényt minimalizálják, az iteráció konvergál egy lokális vagy globális minimumhoz.

Algorithm 1 Alternáló iteratív megoldó.

1. $\forall i \in \{1 \dots S\}$: `init` $\left(s_2^{(i)} \right)$;
2. Ciklus a konvergencia eléréséig:
   (a) (9) megoldása $\underline{p}$-re a minden egyes mintára becsült $s_2^{(i)}$-vel, felhasználva a (10)-es egyenletet.
   (b) $\forall i \in \{1 \dots S\}$: számold újra $s_2^{(i)}$ értékét felhasználva a (11)-ik egyenletet.

4.2. Zárt alakú megoldás

A becslés megfogalmazható zárt alakban is, ha feltételezzük, hogy minden $s_2^{(i)}$ és minden egyéb paraméter ismeretlene egyazon lineáris egyenletrendszernek. Ez alapján a (9)-es egyenletet egy újabb módon fogalmazzuk meg, mellyel egy homogén lineáris egyenletrendszert kapunk:

$$E_2 \cdot \underline{x} = \underline{0} \tag{12}$$

ahol

$$\underline{x} = \left[\underline{p}^T \; s_2^{(1)} \; s_2^{(2)} \; \dots \; s_2^{(S)} \right]^T \in \mathbb{R}^{9+S} \tag{13}$$

és

$$E_2 = \begin{bmatrix} E_2^{(1)} \\ E_2^{(2)} \\ \vdots \\ E_2^{(S)} \end{bmatrix} . \tag{14}$$

Továbbá, $E_2^{(i)}$ a következő módon kapható meg

$$E_2^{(i)} = \left[E^{(i)} \; \dots \; \underline{0} \; -\underline{\mu}^{(i)} \; \underline{0} \; \dots \right] . \tag{15}$$

ahol $-\underline{\mu}^{(i)}$ az $(i+9)$-ik oszlopa $E_2^{(i)}$-nek és

$$\underline{\mu}^{(i)} = \frac{1}{s_2^{(i)}} \underline{\mathrm{h}}^{(i)} = c^{(i)} vec\left(\mathbf{A}^{(i)}\right). \tag{16}$$

A becslés problémáját a következő módon fogalmazzuk meg:

$$\min_{\underline{\mathrm{x}}} \{E_2 \cdot \underline{\mathrm{x}}\} \text{ s.t. } \|\underline{\mathrm{x}}\| = 1. \tag{17}$$

Jól ismert, hogy a fenti probléma megoldása az $E_2^T E_2$ mátrix legkisebb sajátértékéhez tartozó sajátvektora [3].

4.3. Hibrid módszer

A javasolt módszerben rejlő újdonság, hogy az affin transzformációkban rendelkezésre álló információt, ha részlegesen is, de felhasználja. Ezek alatt a nyírás, forgatás tengelyek menti skálázások transzformációját értjük. Ez az eljárás nem használja a két eltolásra vonatkozó paramétert. Ezekkel a paraméterekkel a klasszikus DLT [5] eljárás foglalkozik, segítségükkel becsüli meg a kameraparamétereket. Szerencsére a DLT eljárás egyenletei hozzáfűzhetőek a mi egyleteinkhez, hiszen ugyan azon ismeretleneket tartalmazzák (lásd a (13) -as $\underline{\mathrm{x}}$ vektorát). Ezáltal a DLT eljárás további két sort ad az általunk megfogalmazott lineáris egyenletrendszerhez:

$$\left(\mathbf{P}^2\big|_{(3,:)} \cdot \tilde{\mathbf{X}}^{(i)}\right) \cdot \mathbf{u}_2^{(i)} = \mathbf{P}^2\big|_{(1:2,:)} \cdot \tilde{\mathbf{X}}^{(i)} \tag{18}$$

mivel a becslés paramétere – az $\underline{\mathrm{x}}$ vektor – tartalmazza a $\mathbf{P}^2$ elemeit. A probléma továbbra is homogén lineáris egyenletrendszer marad, az optimális $\underline{\mathrm{x}}$ érték pedig továbbra is egy $\mathbf{G}^T\mathbf{G}$ mátrix legkisebb sajátértékéhez tartozó sajátvektora, ahol a $\mathbf{G}$ korábbiakhoz képest egy némileg módosított együtthatómátrix.

4.4. Numerikus finomítás a Levenberg-Marquardt [9] eljárással

A standard kalibrációs eljárások [5,10] a kezdeti becsléseket tovább finomítják, minimalizálva egy valamilyen geometriailag értelmezhető metrikát, mint például a visszavetítési vagy a Sampson-hibát. Ez esetünkben is lehetséges. Az optimalizáció DLT-vel rokonítható oldalához ((18)-as egyenlet) a visszavetítési hiba használatát javasoljuk. Az affin oldalhoz kapcsolódó hiba a következőképpen fejezhető ki a Frobenius norma segítségével:

$$\left\| \frac{1}{s_2^{(i)} c^{(i)}} \left(\mathbf{P}^2\big|_{(1:2,\,1:3)} - \underline{\mathrm{u}}_2^{(i)} \cdot \mathbf{P}^2\big|_{(3,\,1:3)}\right) \cdot \mathbf{N}^{(i)} - \mathbf{A}^{(i)} \right\|_F^2 .$$

Tanulmányukban Bentolila és Francos [2] bizonyította, hogy a (19)-es egyenletben írottak geometriailag értelmezhetőek, hasonlóan a visszavetítési hibához.

A visszavetítési hiba és az affin komponensek tapasztalati úton súlyozhatóak egymás között. Tesztjeinkben minden súlyt 1-nek választottunk meg, de mindez adaptívan változtatható lehet.

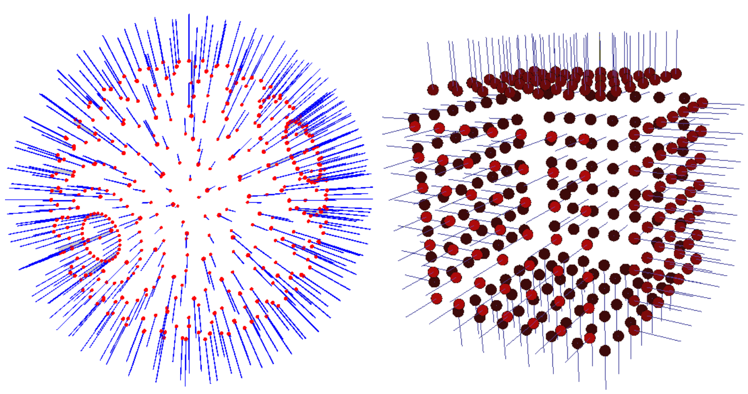
Megjegyezzük, hogy Bentolila és Francos [2] munkája egy kifinomultabb metrikát mutat be az affin hiba mérésére, mely transzformált képfoltok távolságát méri, ám eljárásuk nem integrálható munkánkba.

5. Tesztek

A DLT algoritmussal [5] való összehasonlítás céljából a javasolt kalibrációs módszereket szintetikus adatokon teszteltük. A zárt alakú megoldó valós körülmények közötti alkalmazhatóságát egy kamera-projektor rendszer kalibrálásán keresztül igazoltuk.

Választásunk azért esett a DLT-re, mint riválís metódus, mivel ez is egy hasonlóan linearizált alakja a projektív egyenletnek, akárcsak a mi megközelítésünk a ((2))-es egyenlet esetén. Mivel az összes becslőnk ezt a fajta algebrai hibát minimalizálja, kimenetük további numerikus finomításra szorul.

5.1. Tesztek szintetikus adatokon



1. ábra. Gömb (left) and cube (right) with surface normals. They were used for synthesized tests.

A szintetikus tesztek során a megközelítésünk a következő: Egy adott geometriai alakzat felületén kvázi egyenletes mintavételezéssel mintapontokat és hozzájuk tartozó felületi normálisokat állítottunk elő, majd ezek „levetítésével” képpontokat és affin transzformációkat (((2))-es egyenlet) kaptunk. A vetítéseket egy szintetikus stereo képpárra végeztük el, a kamerákat pedig egy-egy 3×4-es projekciós mátrix írta le. Az alábbi 3D geometriai alakzatokon végeztük a mintavételezést:

1. *Gömb*: A szintetikus gömb alakzat a felületi normálisokkal együtt az 1. ábrán látható, melyet 72 mintavételi pont segítségével állítottunk elő. Minden tesztlépésben véletlenszerűsítettük a mintavételezést.
2. *Kocka*: A szintetikus kocka alakzat a felületi normálisokkal együtt az 1. ábrán látható. A kocka minden oldalát egyenletes mintavételezéssel $7 \times 7 = 49$ helyen vizsgáltuk, így összesen $49 \cdot 6 = 294$ mintavételi ponttal.

A sztereo pár közötti affin paraméterek kiszámítása a következőképpen zajlott: (i) A mintapontban vett érintősík meghatározása (gömb ill. kocka esetén), (ii) majd ezt követte a képpárra való vetítés. (iii) A sík vetületei egy-egy homográfiát határoznak meg a képekre vonatkoztatva. (iv) Az affin paraméterek a két kép közötti relatív homográfia elsőrendű közelítése a vetületi pontokban. (v) Végezetül zajt adtunk a 2D vetületi pontokhoz és affin transzformációkhoz.

A tesztek során 3 eljárást hasonlítottunk össze: a standard (normalizált) **DLT** algoritmus [5], a javasolt iteratív eljárás (**ALTER**), illetve a zárt alakú megoldó (**CLOSED**). A kalibráció minőségét a becsült projekciós mátrix és a ground truth különbségének Frobenius-normájával mértük. (Megjegyezzük, hogy kizárólag a bal 3×3-as al-mátrix elemeit vizsgáltuk. A mátrixokat normalizáltuk, mielőtt vizsgáltuk különbségüket.) Minden tesztesetet 20-szor ismételtük meg. Az ábrákról leolvasható az átlagos hiba és a medián.

A 3. ábrán látható, hogy a normálisokat alkalmazó eljárásaink (ALTER és CLOSED) egyaránt jelentős fölénnyel teljesítenek, a riválís DLT eljáráshoz képest. Ez igaz a gömbbel vagy a kockával végzett tesztekre is. A diagramokon látható, hogy az ALTER és CLOSED módszerek kimenetének minősége közel azonos. Ez elvárt viselkedés, hiszen azonos hibát próbálnak minimalizálni.

Az eljárások időigényét tekintve (4. ábra), a zárt alakú CLOSED algoritmus javasolt, hiszen az jelentősen gyorsabb az iteratív változatnál. A DLT algoritmus a többi eljárásnál gyorsabb, ám ez evidens, hiszen a megoldónál használatos együtthatómátrix kevesebb elemet tartalmaz, így annak felépítése vagy a sajátérték-felbontása is időben kevésbé költséges.

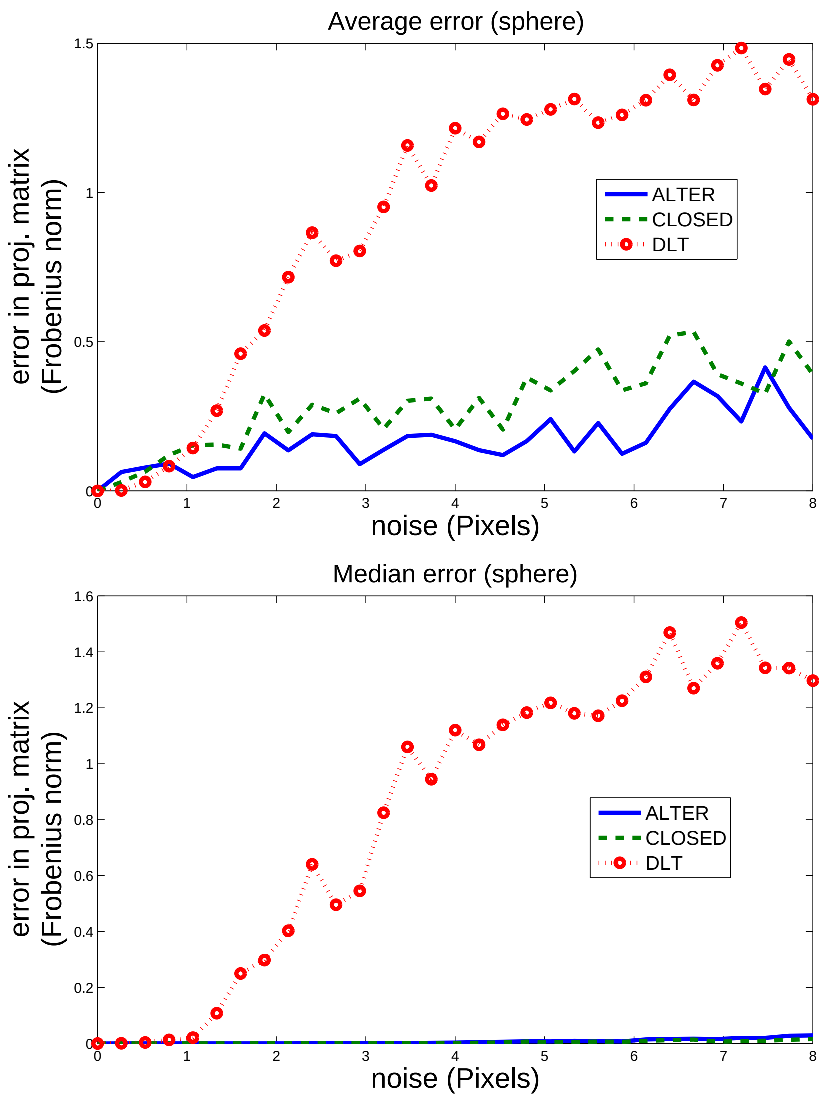
A következő szintetikus teszt a 4.3. részben bemutatott eljárást hasonlítja össze a DLT-vel. Mindkét eljárás kétféleképpen lett implementálva: (i) az eredeti zárt alakú becslő, illetve (ii) a Levenberg-Marquardt technikával numerikusan finomított változat, ahogyan az a 4.4. részben le van írva. Az eredmény a 5. ábrán látható.

A tesztből kiderül, hogy a hibrid verzió minden esetben jobban teljesít, mint az eredeti DLT eljárás, mivel előbbi több információt használ fel, így a túlhatározott egyenletrendszer is több egyenletet tartalmaz. Ez természetesen igaz a numerikus finomítást tartalmazó HybridLM és DirectLM esetére vetítve is.

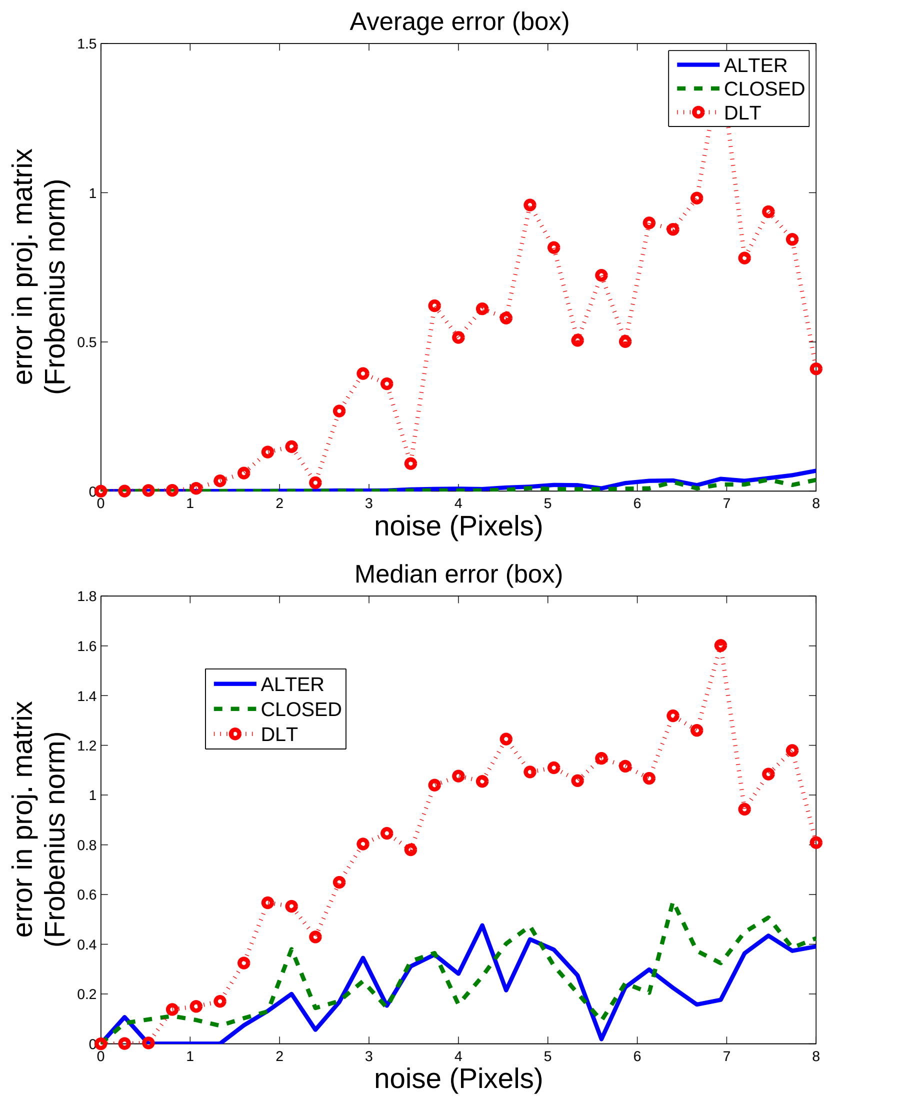
5.2. Tesztek valós adatokon

Kalibrációs eljárásunkat sikeresen alkalmaztuk egy forgóasztalos kamera-projektor strukturált fényes rendszer esetén. A felszerelésről egy kép a 6. ábrán látható.

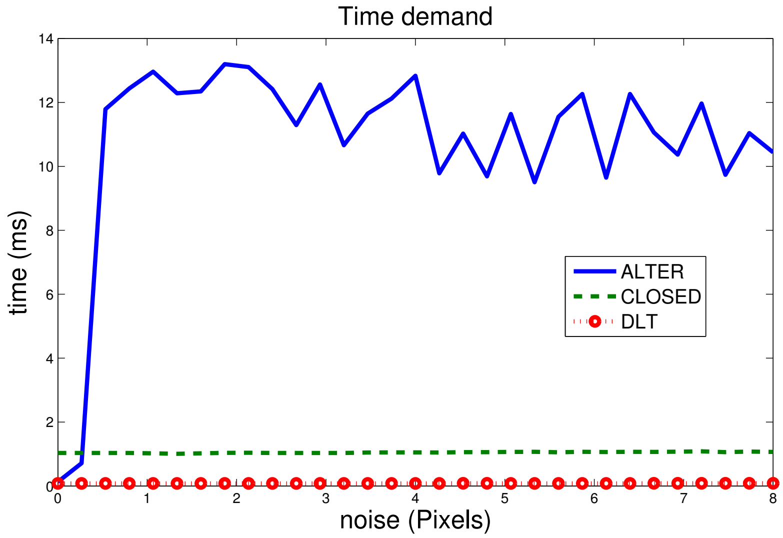
A javasolt módszerben többek között újdonság, hogy *a teljes eljárásnak elegendő egyetlen bemeneti kép a projektor által kivetített ábráról, a két eszköz pedig*



2. ábra. Gömb: A linearizált becslők szintetikus tesztjeinek eredménye. A javasolt eljárásaink jelentősen nagyobb pontosságot nyújtanak a DLT-vel szemben.



3. ábra. Kocka: A linearizált becslők szintetikus tesztjeinek eredménye. A javasolt eljárásaink jelentősen nagyobb pontosságot a javasolt DLT-vel szemben.



4. ábra. A riválisjmetódusok időigénye milliszekundumokban. A javasolt zárt alakú „CLOSED” eljárás jelentősen gyorsabb, mint az iterációt igénylő ALTER. Mindkét javasolt eljárás lassabb a DLT-nél: kompromisszumot kell kötni a minőség és a sebesség között.

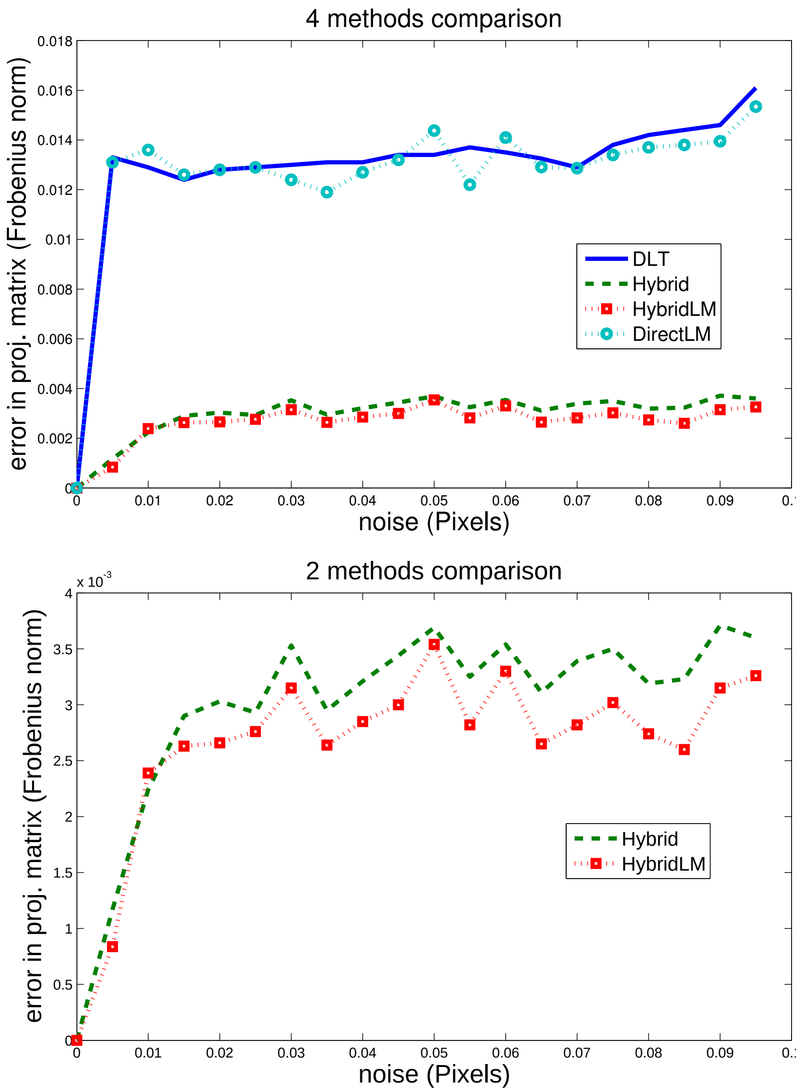
tetszőleges perspektív paraméterekkel rendelkezhet. A választott kalibrációs objektum egy kocka, melyet a projektorral egy speciális minta segítségével világítunk meg, ahogyan az a 6. ábra bal oldalán látható.

Kamera kalibrációja A kocka látható sarkait kézzel megjelöltük, majd a kamera ismeretlen paramétereit a DLT eljárással [5] becsültük meg. A becsült mátrixot belső, illetve külső (forgatás és eltolás) paraméterekre bontottuk.

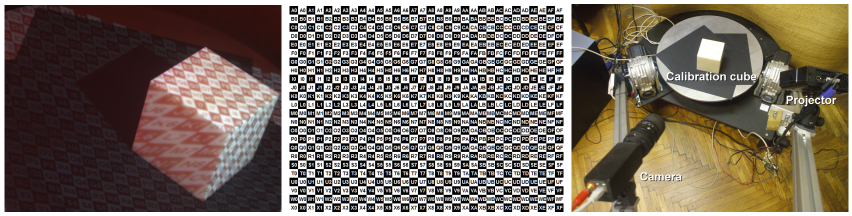
Projektor kalibrációja Az eljárásunkban a fő újdonság a projektor paramétereinek becslése. Az algoritmus képes a projektort reprezentáló perspektív vetületi mátrix bal felső 3×3-as al-mátrixát megbecsülni. Az eljárásnak szüksége van affin transzformációkra a mintavételi pontok helyén. A kamera és a projektor perspektivitása miatt homográfiával le lehet írni a kocka sík oldalaira kivetített minta és a kamera képe közötti transzformációt. Az affin paraméterek a homgográfia első deriváltjából nyerhetőek ki, ahogyan az a B. függelékben is olvasható.

Ha a kockának 7 sarka látható a kamera képén, akkor a 3 látható oldal síkjai és felületi normálisai is ismertek. Ezen normálisok és a számított affin transzformációk szolgálnak a kalibrációs eljárás bemeneteként.

A kivetített minta kódokat tartalmaz, ahogyan az a 6. ábra közepén is látható. Használatuk a sakktábla-sarkak helyes manuális kiválasztása miatt elenged-



5. ábra. Linearizált és numerikusan finomított kimenetek összehasonlítása. Vízszintes tengely: relatív zajszint. Függőleges tengely: A becsült projekciós mátrix a ground truthtól vett normalizált távolsága. DLT: a standard Direct Linear Transformation eljárás. Hybrid: DLT + CLOSED algoritmusok elegye. DirectLM: A DLT kimenetének Levenberq-Marquardt (LM) eljárással numerikusan finomított változata. HybridLM: LM finomítás a Hybrid kimenetén. Felső ábra: A rivális eljárások összehasonlítása. Alsó ábra: A hibrid és a finomított változatának összehasonlítása.



6. ábra. A javasolt eljárást valós körülmények között a strukturált fény alapú rendszerünk kalibrációján keresztül validáltuk. Bal: A speciális minta kivetetve a kalibrációs objektumra. Közép: A projektor által kivetített kalibrációs minta. Jobb: Fénykép a teljes felszerelésről és a kalibrációs objektumról.

hetetlen. Egy jövőbeni munkaként lehetséges egy automatikus eljárás kifejlesztése a sarkak detekciójára, ám ez a jelen tanulmány keretein túlmutat.

A végső lépések tömör összefoglalása a következő: a zárt alakú (CLOSED) eljárás megbecsüli a projektor perspektív vetületi mátrixát. A belső paramétereket RQ-felbontással nyerjük ki ebből. A becsült paraméterek a következőek: vízszintes és függőleges fókusztávolságok $f_u = 1136$, $f_v = 1068$; nyírás $s = 59$; döféspont $\underline{p}_0 = [478, 758]$. A becsült paraméterek érvényesek, hiszen a vízszintes és függőleges fókusztávolságok megközelítőleg azonosak, a nyírás pedig jelentősen kisebb náluk. A projektor felbontása 800×600, így a döféspont koordinátái is elfogadható helyen helyezkednek el.

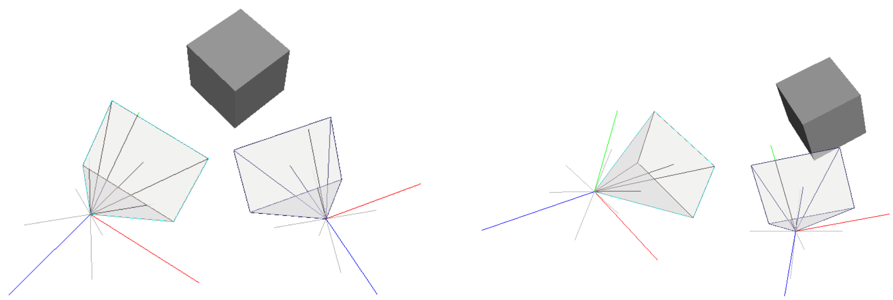
A projektort a teszttől függetlenül, egy standard sakktábla alapú eljárással [10] is kalibráltuk, illetve azt tovább finomítottuk numerikusan. Bemenetként *néhány tucat* képet használtunk. Az így becsült döféspont $[467, 764]$ és fókusztávolságok 1454 és 1461 közel esnek az algoritmusunk által becsülthez.

A kalibrált rendszer a 7. ábrán két nézetből. A gúlák a perspektív kameraparamétereket reprezentálják a kamera (bal) és a projektor (jobb) esetén. A kalibrációs objektum (kocka) a képek közepén látható.

6. Következtetések

Bemutattuk, hogy a felületi normálisok ismerete egy kalibrációs objektum esetén felhasználható jobb minőségű kameraparaméterek becsléséhez. Két fő metódusvariánst mutattunk be a felületi normálist felhasználó kalibrációhoz. Taglaltuk az újfajta eljárások tulajdonságait és validáltuk azokat valós és szintetikus tesztek során.

A tesztek megmutatták, hogy a zárt alakú megoldót érdemes használni, gyorsasága és pontossága végett. Egy projektor-kamera rendszert is sikeresen kalibráltunk a bemutatott eljárások segítségével.



7. ábra. A kalbirált kamera-projektor rendszer két nézetből bemutatva. A kamerát és a projektort gúlákkal reprezentáljuk, illetve a képeken látható kocka a kalibrációs objektum.

Hivatkozások

1. Daniel Barath, Jozsef Molnar, and Levente Hajder. Novel methods for estimating surface normals from affine transformations. In *Computer Vision, Imaging and Computer Graphics Theory and Applications*, pages 316–337. Springer International Publishing, 2016.
2. Jacob Bentolila and Joseph M. Francos. Conic epipolar constraints from affine correspondences. *Computer Vision and Image Understanding*, 122:105–114, 2014.
3. Åke Björck. *Numerical Methods for Least Squares Problems*. Siam, 1996.
4. Olivier Faugeras and F. Lustman. Motion and structure from motion in a piecewise planar environment. Technical Report RR-0856, INRIA, 1988.
5. R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.
6. Liu He. *Deeper Understanding on Solution Ambiguity in Estimating 3D Motion Parameters by Homography Decomposition and its Improvement*. PhD thesis, University of Fukui, 2012.
7. K. Levenberg. A method for the solution of certain problems in least squares. *Quart. Appl. Math.*, 2:164–168, 1944.
8. Ezio Malis and Manuel Vargas. Deeper understanding of the homography decomposition for vision-based control. Technical Report RR-6303, INRIA, 2007.
9. D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math.*, 11:431–441, 1963.
10. Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.
11. Chuan Zhou, Da-Long Tan, Feng Zhu, and Zaili Dong. A planar homography estimation method for camera calibration. In *Proceedings of the IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pages 424–429, 2003.

A. Perspektív kamera parciális deriváltjai

A függelék ezen része bemutatja, hogy a vetítőfüggvény $[x, y]$ komponenseire hogyan számolhatóak ki a $[X, Y, Z]$ térkoordináták parciális deriváltjai.

Jelöljük a $\mathbf{P}$ perspektív mátrix i, j-ik komponensét a mint P^{ij} és a k-ik sort mint P^k.

$$\frac{\partial x}{\partial X} = \frac{\partial \frac{P^1[X,Y,Z,1]^T}{P^3[X,Y,Z,1]^T}}{\partial X} =$$

$$\frac{P^{11}s - P^{31}\left(P^1\left[X,Y,Z,1\right]^T\right)}{s^2} = \frac{1}{s}\left(P^{11} + xP^{31}\right)$$

ahol $s = P^3\left[X,Y,Z,1\right]^T$ a projektív mélység. Hasonlóan,

$$\begin{array}{ll} \frac{\partial x}{\partial Y} = \frac{1}{s}\left(P^{12} + xP^{32}\right), & \frac{\partial x}{\partial Z} = \frac{1}{s}\left(P^{13} + xP^{33}\right), \\ \frac{\partial y}{\partial X} = \frac{1}{s}\left(P^{11} + yP^{31}\right), & \frac{\partial y}{\partial Y} = \frac{1}{s}\left(P^{12} + yP^{32}\right), \\ \frac{\partial y}{\partial Z} = \frac{1}{s}\left(P^{13} + yP^{33}\right). & \end{array}$$

A cikk java részének jelöléseit felhasználva a vetítés térkoordinátákra vett Jacobi-ja a következő kompakt alakban fejezhető ki:

$$\nabla\underline{\mathrm{u}} = \frac{1}{s}\left(\mathbf{P}|_{(1:2,\,1:3)} - \underline{\mathrm{u}} \cdot \mathbf{P}|_{(3,\,1:3)}\right) \tag{19}$$

ahol $\underline{\mathrm{u}} = [x, y]^T$.

B. Affin paraméterek kinyerése a homográfiából

Ez a rész bemutatja, hogyan számíthatóak az affin paraméterek képpáron lévő egymásnak megfelelő $[u, v]^T$, $[u', v']^T$ pontok esetén, ha a síkok homográfiája ismert a képek között. Ezt a módszert használtuk valós tesztek esetén, az 5.2. részben.

Az affin transzformáció paramétereit a homográfiából nyerjük ki. Legyen adott a H homográfia, mint 3×3-as mátrix. Az említett pont-pár egymásnak megfelel. Ez a következőképpen fejezhető ki:

$$u' = \frac{h_1^T[u,v,1]^T}{h_3^T[u,v,1]^T}, \quad v' = \frac{h_2^T[u,v,1]^T}{h_3^T[u,v,1]^T}$$

ahol H a 3×3 homográfia-mátrix a következőképpen:

$$H = \begin{bmatrix} h_1^T \\ h_2^T \\ h_3^T \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}.$$

Az affin transzformáció paraméterei a síkok közötti perspektív transzformáció parciális deriváltjai. A bal felső eleme a_{11} az affin transzformációnak a következő:

$$a_{11} = \frac{\partial u'}{\partial u} = \frac{h_{11} h_3^T [u,v,1]^T - h_{31} h_1^T [u,v,1]^T}{\left(h_3^T [u,v,1]^T\right)^2} = \frac{h_{11} - h_{31} u'}{s}$$

ahol $s = h_3^T [u, v, 1]^T$. A többi komponense az affin transzformációnak hasonlóan számítható:

$$a_{12} = \frac{\partial u'}{\partial v} = \frac{h_{12} - h_{32} u'}{s}, \quad a_{21} = \frac{\partial v'}{\partial u} = \frac{h_{21} - h_{31} v'}{s}, \\ a_{22} = \frac{\partial v'}{\partial v} = \frac{h_{22} - h_{32} v'}{s}.$$