

# Normálvektorok optimális becslése affin transzformációkból

Hajder Levente<sup>1,2</sup>, Baráth Dániel<sup>1,2</sup>, Molnár József<sup>1</sup>

<sup>1</sup> Szegedi Tudományegyetem

Képfeldolgozás és Számítógépes Grafika Tanszék

<sup>2</sup> MTA SZTAKI, Budapest

{hajder.levente,barath.daniel}@sztaki.mta.hu

**Kivonat** A tanulmány célja, hogy megmutassuk, kalibrált kamerák esetén a normálvektort meg lehet határozni affin transzformációk ismerete esetén. Ehhez először levezetjük, hogy egy felületdarab két vetülete közötti affin transzformáció hogyan függ össze a felület normálvektorával. Az összefüggés segítségével új normálvektorbecslő algoritmusokat vezetünk be. Azt is levezetjük, hogy ha a felületdarabka projektív mélységét ismerjük, legkisebb négyzetes értelemben optimális becslő is készíthető, ellenkező esetben iteráció segítségével juthatunk el a megoldásig. A módszereket szintetikus és valós adatokon egyaránt teszteljük.

## 1. Bevezetés

Bár a számítógépes látás több évtizede intenzíven kutatott terület, még mindig rengeteg megoldatlan problémára keresik a kutatók a választ. Ez a tanulmány egy még nyitott problémával foglalkozik: megmutatjuk, hogy sztereo képeken hogyan lehet affin transzformációból a megfelelő felületi darabka normálvektorát optimálisan becsülni, ha a kameráink kalibráltak.

A szakirodalmat áttekintve, azt találjuk, hogy normálvektorok becslésére a leggyakrabban alkalmazott eljárás a fotometrikus sztereo, amit körülbelül 20 évvel ezelőtt javasolt Woodham[1]. A módszer hatékony, azonban laboratóriumi körülményeket igényel, hiszen mesterségesen kell a rekonstruálandó tárgyakat több irányból megvilágítani. Általában párhuzamos fényforrásokat szokás alkalmazni [1], azonban pontszerű fényforrásokkal [2] is elvégezhető a felületi normálvektorok számítása.

Normálvektorok becslése pusztán képek segítségével is lehetséges. Ebben az esetben a két képen kivágott minták között homográfiát szokás becsülni, és a homográfia felbontásával kapjuk meg a normálvektort [3,4]. Ez csak akkor lehetséges, ha a kameránk belső paramétereit ismerjük. A külső paraméterek a homográfia felbontásával szintén meghatározhatóak. A módszer egyik problémája, hogy a felbontás nem egyértelmű, ahogyan azt Liu disszertációjában megmutatta [5].

A mi tanulmányunk azt mutatja meg, hogy nem szükséges a felületdarab képei közötti homográfia ahhoz, hogy a normálvektorokat kiszámítsuk: elég az affin

transzformációkat meghatározni. Ismereteink szerint ezzel a konkrét problémával még nem foglalkoztak a szakirodalomban. Munkánkhoz legközelebbi tanulmányt Habbecke és Kobbelt publikációjában [6,7] találtunk. A szerzőpáros fotokonvizisztencia alapon becsüli meg a térbeli síkdarabot. A sík paraméterei között a normálvektor is szerepel.

Egy másik hasonló megoldás Megyesi és munkatársai dolgozata [8], amiben megmutatták, hogy rektifikált képpárok esetén a két képpár közötti homográfia hogyan befolyásolja a normálvektort. (Megyesiék dolgozatukban sűrű illesztéssel foglalkoznak, azonban a normálvektorok meghatározása részeredményként szerepel a publikációban.) Számunkra ez a módszer azért nem ideális, mert egyrészt affin transzformációk helyett homográfiát becsülnek, másrészt a rektifikálással törvényszerűen hibát is visznek a rendszerbe, mi pedig az optimális megoldás megtalálására törekszünk.

Ismereteink szerint itt közöljük az első olyan módszert, amelyik affin transzformációkból legkisebb négyzetes értelemben véve optimális becslést ad a felületi normálvektorra. Tanulmányunknak két fő mondanivalója van:

- Megmutatjuk a kameraparaméterek és a felületi normálvektorok közötti általános összefüggést tetszőleges kameraparaméterekre.
- Többféle becslő eljárást javasolunk normálvektorok becslésére, melyek között az optimális is megtalálható, amely az ismert affin transzformációra minimalizálja a becslési hibát.

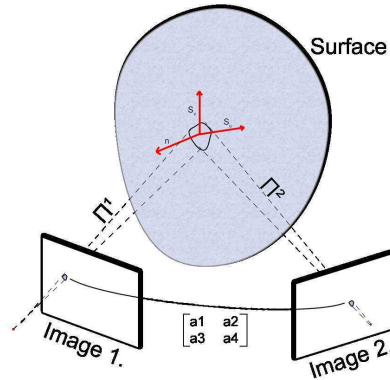
## 2.. Elméleti alapok

Adott egy térbeli lapka, amely két képen látszik. Lokális közelítésben a lapka tekinthető síknak. A feladat ábrázolása az 1. képen látható. A lapkát nem ismerjük térben, azonban a két képen a vetületekhez tartozó pixeleket kétdimenziós képfeldolgozási módszerekkel meg tudjuk becsülni. A cél a felületi lapka  $n$  normálvektorának meghatározása.

Az  $[X, Y, Z]^T$  háromdimenziós vektorral megadott felületi pontok kétdimenziós koordinátáit a  $\Pi$  vetítőfüggvénnyel számoljuk a térbeli pontból, a felület háromdimenziós pontját pedig parametrikus alakban írjuk le:

$$\begin{aligned} x &= \Pi_x(X, Y, Z) & y &= \Pi_y(X, Y, Z) \\ X &= X(u, v), & Y &= Y(u, v), & Z &= Z(u, v). \end{aligned}$$

Differenciális geometriából [9] jól ismert tény, hogy az érintővektorok felírhatóak a paraméteres alakban megadott felület parciális deriváltjaiból, a normálvektor



1. ábra. Térbeli lapka perspektíven vetítve egy képpárra.

pedig a két érintővektor vektoriális szorzatából kapható meg:

$$S_u = \begin{bmatrix} \frac{\partial X(u,v)}{\partial u} \\ \frac{\partial Y(u,v)}{\partial u} \\ \frac{\partial Z(u,v)}{\partial u} \end{bmatrix}$$

$$S_v = \begin{bmatrix} \frac{\partial X(u,v)}{\partial v} \\ \frac{\partial Y(u,v)}{\partial v} \\ \frac{\partial Z(u,v)}{\partial v} \end{bmatrix}$$

$$n = S_u \times S_v$$

Az  $[X, Y, Z]^T$  térbeli pont és az  $S_u$  és  $S_v$  érintővektorok a felület adott pontjában lévő érintősíkját is meghatározzák. Lokálisan a felület közelíthető ezzel az érintősíkkal. Feltételezzük ugyebár, hogy a felületről két kép készült. A felületdarabka vetülete a lokális környezetében az elsőrendű Taylor sor segítségével jól közelíthető:

$$\begin{bmatrix} x + \Delta x \\ y + \Delta y \end{bmatrix} \approx \begin{bmatrix} \Pi_x(X, Y, Z) \\ \Pi_y(X, Y, Z) \end{bmatrix} + \begin{bmatrix} \frac{\partial \Pi_x(X, Y, Z)}{\partial u} & \frac{\partial \Pi_x(X, Y, Z)}{\partial v} \\ \frac{\partial \Pi_y(X, Y, Z)}{\partial u} & \frac{\partial \Pi_y(X, Y, Z)}{\partial v} \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix}$$

Nézzük meg ezek után, hogy a parciális deriváltak segítségével a térbeli érintősíkon és az egyik képen levő minta hogyan feleltethető meg affin transzformáció segítségével:

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \approx A \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix}$$

$$A = \begin{bmatrix} \frac{\partial \Pi_x(X,Y,Z)}{\partial u} & \frac{\partial \Pi_x(X,Y,Z)}{\partial v} \\ \frac{\partial \Pi_y(X,Y,Z)}{\partial u} & \frac{\partial \Pi_y(X,Y,Z)}{\partial v} \end{bmatrix}$$

A parciális deriváltak a láncszabály segítségével módosíthatóak. Például:

$$\begin{aligned} \frac{\partial \Pi_x(X,Y,Z)}{\partial u} &= \frac{\partial \Pi_x(X,Y,Z)}{\partial X} \frac{X}{\partial u} + \frac{\partial \Pi_x(X,Y,Z)}{\partial Y} \frac{Y}{\partial u} \\ &\quad + \frac{\partial \Pi_x(X,Y,Z)}{\partial Z} \frac{Z}{\partial u} = \nabla \Pi_x^T S_u, \end{aligned}$$

ahol  $\nabla \Pi_x$  a vetítőfüggvény gradiense  $X$ ,  $Y$  és  $Z$  felületi koordináták szerint. Hasonlóan:

$$\frac{\partial \Pi_x}{\partial v} = \nabla \Pi_x^T S_v \quad \frac{\partial \Pi_y}{\partial u} = \nabla \Pi_y^T S_u \quad \frac{\partial \Pi_y}{\partial v} = \nabla \Pi_y^T S_v.$$

Ebből következik, hogy magát az affin transzformációt így is fel tudjuk írni:

$$A = \begin{bmatrix} \nabla \Pi_x^T \\ \nabla \Pi_y^T \end{bmatrix} [S_u \ S_v].$$

Mivel sztereo képpárunk van, hiszen a térbeli alakzatunkról két képet készítetünk, az affin transzformáció a két képen látható ugyanazon minta között felírható az  $A_1$  transzformáció inverzének és az  $A_2$ -es transzformációnak a sorzatával. (Az előbbi az első képen levő minta és a térbeli minta közötti kapcsolatot írja le, az utóbbi a térbeli és a második képen levő minta kapcsolatát.) Formálisan felírhatjuk az alábbi összefüggést:

$$[\Delta x_2 \ \Delta y_2]^T = A_2 A_1^{-1} [\Delta x_1 \ \Delta y_1]^T$$

A két kép közötti affin transzformációt felírhatjuk tehát  $A_2 A_1^{-1}$  alakban. A további átalakításhoz nézzük meg, hogy az  $A$  mátrix inverze hogyan írható fel:

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} \Pi_x^T S_u & -\Pi_y^T S_u \\ -\Pi_x^T S_v & \Pi_y^T S_v \end{bmatrix},$$

ahol  $\det(A) = \Pi_x^T S_u \Pi_y^T S_v - \Pi_x^T S_v \Pi_y^T S_u$ . Ha figyelembe vesszük, hogy  $S_v S_u^T - S_u S_v^T = [N]_\times$ , akkor egyszerű átalakításokkal a következő alakra hozható az affin transzformáció:

$$A_1^{-1} A_2 = \frac{1}{\Pi_x^{1T} [N]_\times \Pi_y^1} \begin{bmatrix} \Pi_x^{2T} [N]_\times \Pi_y^1 & \Pi_x^{1T} [N]_\times \Pi_x^2 \\ \Pi_y^{2T} [N]_\times \Pi_y^1 & \Pi_x^{1T} [N]_\times \Pi_y^2 \end{bmatrix}$$

Fontos megjegyezni, hogy skálázásra érzéketlen a formula, hiszen mind a determináns, mind a mátrix elemei  $[N]_{\times}$ -el meg lettek szorozva. Az  $a^T [N]_{\times} b$  kifejezést szokás skaláris hármasszorzatnak is hívni. Ha figyelembe vesszük, hogy  $a^T [n]_{\times} b$  egyenlő  $n^T (b \times a)$ -vel, az affin transzformáció végleges formáját így kapjuk meg:

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = A_1^{-1} A_2 = \frac{1}{n^T w_5} \begin{bmatrix} n^T w_1 & n^T w_2 \\ n^T w_3 & n^T w_4 \end{bmatrix}, \quad (1)$$

ahol  $w_1 = \nabla \Pi_y^1 \times \nabla \Pi_x^2$ ,  $w_2 = \nabla \Pi_x^2 \times \nabla \Pi_x^1$ ,  $w_3 = \nabla \Pi_y^1 \times \nabla \Pi_y^2$ ,  $w_4 = \nabla \Pi_y^2 \times \nabla \Pi_x^1$  és  $w_5 = \nabla \Pi_y^1 \times \nabla \Pi_x^1$ .

Ez az levezetés egy nagyon fontos összefüggéshez vezetett, hiszen az 1. egyenlet minden kameramodell esetén igaz. Mindössze a vetítő  $\Pi$  függvényeket kell megadni és a gradiensüket kiszámolni.

### 2.1.. Perspektív kamera

Ebben a fejezet részben megmutatjuk, hogy az általános egyenletet perspektív kamerára hogyan lehet alkalmazni. Projektív kamera esetén a vetítési egyenlet alakja a következő:

$$[x, y, 1]^T = \frac{1}{s} P_{persp} [X, Y, Z, 1]^T, \quad (2)$$

ahol  $[x, y]$  jelöli a vetített koordinátákat,  $s$  a projektív mélység,  $P_{persp}$  a  $3 \times 4$ -es perspektív kamera. Ha a kamera sorait rendre  $p_1^T$ ,  $p_2^T$  és  $p_3^T$ -mal jelöljük, a vetítő függvényeket így kapjuk meg:

$$\begin{aligned} \Pi_x &= \frac{p_1^T [X, Y, Z, 1]^T}{s} & \Pi_y &= \frac{p_2^T [X, Y, Z, 1]^T}{s}, \\ \nabla \Pi_x &= \frac{1}{s} \begin{bmatrix} P_{11} + xP_{31} \\ P_{12} + xP_{32} \\ P_{13} + xP_{33} \end{bmatrix}, \\ \nabla \Pi_y &= \frac{1}{s} \begin{bmatrix} P_{21} + yP_{31} \\ P_{22} + yP_{32} \\ P_{23} + yP_{33} \end{bmatrix}, \end{aligned}$$

ahol  $P_{ij}$  a projektív mátrix  $i$ -edik sorában a  $j$ -edik elem. A projektív mélység az  $s = p_3^T [X, Y, Z, 1]^T$  összefüggéssel számítható. Az affin transzformáció alakja a projektív kamerára így írható fel:

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = \frac{1}{\alpha n^T w_5} \begin{bmatrix} n^T w_1 & n^T w_2 \\ n^T w_3 & n^T w_4 \end{bmatrix}, \quad (3)$$

ahol  $\alpha = s^1/s^2$  projektív mélységek aránya a két kép között. Továbbá bevezettük az alábbi két egyszerűsítést:  $w_1 = s^1 s^2 (\nabla \Pi_y^1 \times \nabla \Pi_x^2)$ ,  $w_2 = s^1 s^2 (\nabla \Pi_x^2 \times \nabla \Pi_x^1)$ ,  $w_3 = s^1 s^2 (\nabla \Pi_y^1 \times \nabla \Pi_y^2)$ ,  $w_4 = s^1 s^2 (\nabla \Pi_y^2 \times \nabla \Pi_x^1)$ , és  $w_5 = s^2 s^2 (\nabla \Pi_x^1 \times \nabla \Pi_y^1)$ .

Fontos megjegyezni, hogy ha az  $s_i$  projektív mélységet nem ismerjük, de a két vetítő mátrix ( $P_1$  és  $P_2$ ) bal oldali  $3 \times 3$ -mas részmátrixait igen, akkor a gradiens vektorokat egy skálázás erejéig tudjuk meghatározni. (Ez az  $s_i$  skála reciprok.)

Szintén érdekes, hogy a  $w_1 \dots w_4$  vektorok  $s^1 s^2$ -vel, a  $w_5$  vektor pedig  $s^2 s^2$ -vel vannak skálázva.

Az összefüggés jól mutatja azt a tényt, hogy a normálvektor független a kamerák helyzetétől, hiszen a vetítő mátrixok utolsó oszlopa nem szerepel az összefüggésben.

A normálvektor becslése szempontjából két esetet különböztetünk meg:

1. Mindkét vetítő mátrix ( $P^1$  és  $P^2$ ) ismert. (Más szóval: a kamerák teljesen kalibráltak.)
2. Csak a  $3 \times 3$ -mas vetítő almátrixok ismertek. Ebben az esetben a térbeli pont projektív mélysége nem ismert, ezért a gradiensvektoroknak az irányát lehet csak meghatározni, a nagyságát nem.

Az összefüggés azt is mutatja, hogy ha  $w_5 = 0$ , akkor a normálvektort nem lehet megbecsülni. Ez az eset csak akkor állhatna fent, ha  $\nabla \Pi_x^1$  és  $\nabla \Pi_y^1$  párhuzamosak volnának. Perspektív kamera esetén ez azt jelenti, hogy a projektív mátrix első és második sora (a negyedik elemeket elhagyva) párhuzamos. Szerencsére valós kameráknál ez az eset nem fordulhat elő.

### 3. A normálvektor becslése

Ebben a fejezetben több normálvektor-becslőt javasolunk. Lesznek gyorsabb, de kevésbé pontos, illetve lassabb, de precízebb módszerek is. Megmutatjuk, hogy teljesen kalibrált kamerák esetén optimális módszert is lehet készíteni.

#### 3.1.. Gyors normálvektor-becslő (Fast Normal Estimation - FNE)

A 3. összefüggés összesen 4 egyenletet tartalmaz. Ha ezek közül kettőt kiválasztunk, és a hányadosukat vesszük, akkor kapunk egy egyenletet. Ugyanezt a másik kettőre elvégezve kapunk egy újabb egyenletet. Például:

$$\frac{w_1^T n}{w_2^T n} = \frac{a_1}{a_2} \quad (4)$$

$$\frac{w_3^T n}{w_4^T n} = \frac{a_3}{a_4} \quad (5)$$

A nevezőkkel szorzás után:

$$\begin{aligned} (a_2 w_1^T - a_1 w_2^T) n &= 0 \\ (a_4 w_3^T - a_3 w_4^T) n &= 0 \end{aligned}$$

A kapott összefüggésekből következik, hogy  $n$  egyaránt merőleges  $(a_2 w_1^T - a_1 w_2^T)$ -ra és  $(a_4 w_3^T - a_3 w_4^T)$ -ra. Ezért a két vektor keresztszorzata megadja a normálvektor irányát:

$$n = (a_2 w_1^T - a_1 w_2^T) \times (a_4 w_3^T - a_3 w_4^T). \quad (6)$$

### 3.2.. Optimális normálvektor becslése ismert projektív mélység esetén (OPT)

A kapott vektort természetesen normálni kell ahhoz, hogy egységvektor legyen. Ennek a becslőnek nagyon kellemes tulajdonsága, hogy a  $w_1 \dots w_4$  vektorok skálájára teljesen érzéketlen, hiszen a skála normáláskor eltűnik.

Ha az affin transzformációból kapott értékeket máshogy párosítjuk, a normálvektorra másik két becslést is tudunk adni:

$$n = (a_3 w_1^T - a_1 w_3^T) \times (a_4 w_2^T - a_2 w_4^T) \quad (7)$$

vagy

$$n = (a_4 w_1^T - a_1 w_4^T) \times (a_3 w_2^T - a_2 w_3^T) \quad (8)$$

### 3.2.. Optimális normálvektor becslése ismert projektív mélység esetén (OPT)

Ebben a szakaszban legkisebb négyzetes értelemben véve optimális becslő eljárást adunk. A cél a 3. egyenlet hibájának minimalizálása: az affin transzformáció négy értékére számoljuk a négyzetes hibák összegének minimumát. A feladat megegyezik a következővel:

$$\arg \min_n \sum_{k=1}^4 \left( \frac{n^T w_k}{n^T w_5} - a_k \right)^2 \quad (9)$$

Ez az optimalizálási feladat optimálisan megoldható, ahogyan az a függelékben le van írva. ( $\alpha = 1$  paraméterbeállítással kell a függelékben leírt módszert alkalmazni.)

### 3.3.. Normálvektor becslése ismeretlen projektív mélység esetén (ALT)

Amennyiben a projektív mélység nem ismert, a 9. összefüggésben definiált függvény nem optimalizálható egyszerűen, hiszen az  $\alpha = s_1/s_2$  paraméter sajnos nem ismert. Ezért a hibafüggvényt az alábbiak szerint módosítani kell:

$$\arg \min_n \sum_{k=1}^4 \left( \frac{n^T w_k}{\alpha n^T w_5} - a_k \right)^2 \quad (10)$$

Sajnálatos módon ez a feladat jelenlegi ismereteink szerint optimálisan nem oldható meg. Ezért egy alternáló módszert javasolunk, amely két egymás utáni lépést ismételt a konvergencia eléréséig:

1. EstimateAlpha: A költségfüggvény (10. összefüggés)  $1/\alpha$  szerint lineáris, hiszen  $A \frac{1}{\alpha} = b$  írható fel, ahol  $A = \begin{bmatrix} n^T w_1 \\ n^T w_5 \end{bmatrix}, \dots, \begin{bmatrix} n^T w_4 \\ n^T w_5 \end{bmatrix}^T$  és  $b = [a_1, \dots, a_4]^T$ . Az optimális megoldás ebben az esetben a becslélméletben jól ismert formula segítségével adódik:

$$\frac{1}{\alpha} = \frac{n^T w_5}{\sum_j (n^T w_j)^2} \sum_j n^T w_j a_j \quad (11)$$

2. EstimateNormal: A normálvektor becslése az optimális módszerhez hasonlóan adódik, csak  $\alpha$  értékét nem ismerjük egzaktul, hanem az előző lépésben kiszámított  $\alpha$ -t helyettesítjük be. Ezután a függelékben ismertetett módszerrel, a negyedfokú polinom gyökei közül a legjobbat kiválasztva kapunk becslést  $n$ -re.

Az alternáló eljárások egyik nagy hátránya, hogy az algoritmusnak kezdeti értéket kell biztosítani. A mi esetünkben akár az FNE módszer, akár a következő részben ismertetett lineáris módszer (LNE-UPD) alkalmas a kezdőérték meghatározására.

### 3.4.. Normálvektor lineáris becslése (Linear Normal Estimation -LNE)

Az alap mátrixos egyenlet (3. összefüggés) sajnos nem lineáris, azonban a nevezővel való felszorzás segítségével lineárisra lehet tenni. A minimalizálandó függvény így változik:

$$\arg \min_n \sum_{k=1}^4 (n^T w_k - \alpha a_k n^T w_5)^2 \quad (12)$$

A nevezővel való felszorzás ismert trükk, sajnos ezzel a zajt torzítjuk, így az optimalitás elvész. A trükknek nagy előnye, hogy lineáris rendszereket sokkal könnyebben tudunk megoldani.

**Lineáris becslés ismert projektív mélység esetén (Linear Normal Estimation for Known Projective Depth – LNE-KPD).** Ha a projektív mélység ismert,  $\alpha = 1$  állítható be. A nevezővel való szorzás után a problémánk  $An = 0$  alakra hozható  $n^T n = 1$  megkötéssel, ahol

$$A = \begin{bmatrix} w_1 - a_1 w_5 \\ w_2 - a_2 w_5 \\ w_3 - a_3 w_5 \\ w_4 - a_4 w_5 \end{bmatrix}. \quad (13)$$

Ez egy homogén túlhatározott lineáris egyenletrendszer, melynek optimális megoldása az  $A^T A$  mátrix legkisebb sajátértékéhez tartozó sajátvektora.

**Lineáris becslés ismeretlen projektív mélység esetén (Linear Normal Estimation for Unknown Projective Depth – LNE-UPD).** Ha a projektív mélységet nem ismerjük, a minimalizálandó függvény (12. képlet) szintén homogén túlhatározott egyenletrendszerre vezet, melyet  $Bb = 0$  alakra hozhatunk. Az együttthatók és az ismeretleneket tartalmazó vektor a következőképpen



alakul:

$$B^T = \begin{bmatrix} w_1^T, -a_{11} \\ w_2^T, -a_{12} \\ w_3^T, -a_{21} \\ w_4^T, -a_{22} \end{bmatrix},$$

$$b = \begin{bmatrix} n \\ \alpha w_5^T n \end{bmatrix}.$$

A megoldás a  $B^T B$  legkisebb sajátértékéhez tartozó sajátvektor. Érdeemes megjegyezni, hogy az ismeretleneket tartalmazó vektorban az  $\alpha = s_1/s_2$  relatív mélység is szerepel. Másik fontos megjegyzés, hogy a normálvektorok csak az irányát kapjuk meg, a nagyságát nem, az eredményt ezért normálni kell.

#### 4.. Tesztelési eredmények

A javasolt módszert szintetikus és valós adatokon egyaránt teszteltük.

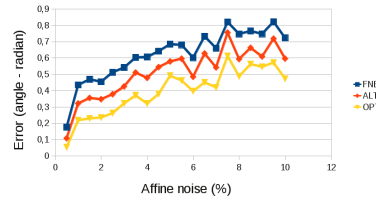
##### 4.1.. Vizsgálat szintetikus adatokon

A szintetikus tesztek alatt arra koncentráltunk, hogy különböző irányú normálvektorokat rekonstruáljunk. Ezért vettünk egy gömbfelületet, melyet gömbi koordináták segítségével mintavételeztünk. Összesen 72 különböző irányú normálvektor vizsgáltunk meg minden egyes tesztesetben. A két kamera a gömb középpontjától megadott távolságra áll, és a gömb középpontját nézi.

Az affin paramétereket szintén mesterségesen ki lehet számítani. A gömb pontjainak érintősíkjait ismerjük, a projekciós mátrixokat is, hiszen azokat is mesterségesen állítottuk elő, ezért a térbeli érintősík és a kameraképek közötti transzformációt ismerjük, ebből a két kép közötti affin transzformáció is megkapható. Tesztjeinkben hibaértéknek a valódi (ground truth) és a becsült értékek különbségvektorának hosszát használjuk. Minden tesztesetben a gömb összes normálvektorára elvégeztük a vizsgálatot, és összesen ötvenszer vettük a gömböt. Azaz eredményeink  $50 \cdot 72 = 3600$  normálvektor-becslés átlagaként jöttek ki.

**Tesztelés zajos affin mátrixokkal.** Amennyiben az ismert affin mátrixok értékéhez zajt adunk, össze tudjuk hasonlítani a becsülő eljárásaink zajérzékenységet. Az eredményt a 2. ábra bal oldalán láthatjuk, ahol az FNE, ALT és OPT módszereket hasonlítottuk össze. Természetesen mindig az optimális módszer adja a legjobb eredményt. A gyors (FNE) eljárás a legkevésbé hatékony, hiszen ott a gyors számításra koncentráltunk.

Szintén összehasonlítottuk a lineáris és nemlineáris módszereket. Ismert mélység esetén az OPT és a LIN-KPD módszerek versenyeztek egymással, ismeretlen mélység esetében az ALT és a LIN-UPD. Az eredmény a 3. grafikonon olvasható le. Érdekes, hogy az optimális módszer lényegesen jobb eredményt ad lineáris



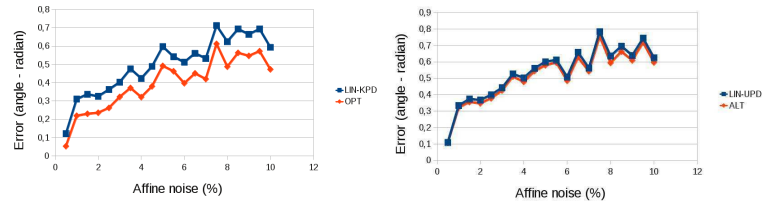
2. ábra. Javasolt módszerek összehasonlítása zajos adatok esetén.

társánál, az alternáló ellenben alig jobb a LIN-UPD-nél, ráadásul a sok iteráció miatt lényegesen lassabb is. Ezért az OPT illetve a LIN-UPD módszerek használatát javasoljuk, attól függően, hogy ismerjük-e a projektív mélységet.

Egy másik tesztben az eredmények szórását is összehasonlítottuk (1. táblázat). Kiemelnénk az eredmények közül két elvárható tulajdonságot: (1) az optimális módszer elsősege megkérdőjelezhetetlen (2) a projektív mélység ismerete jelentősen javít az eredményen.

1. táblázat. Hibavektorok hosszának szórása különböző becslések esetén.

FNE	LIN-UPD	ALT	LIN-KPD	OPT
0.55	0.449	0.433	0.352	<b>0.2919</b>



3. ábra. A lineáris és a megfelelő nemlineáris módszerek összehasonlítása.

#### 4.2.. Teszt valódi képeken

##### Kalibrált képek

Az itt bemutatott normálvektor-becslőket valódi képeken is teszteltük. Az oxfordi Visual Geometry Group honlapjáról<sup>1</sup> letöltött képekhez kalibrációs adatok és pontkövetések is tartoznak. Az affin transzformációkat saját nyers erő (brute force) megoldást alkalmazó képillesztő algoritmussal becsültük meg. A kivágott, illesztett minták általában  $60 \times 60$ -as méretűek voltak. Az eredmények a 4. ábrán láthatóak.

További eredményeket a 4. képen láthatunk. A folyosó esetében a bázistávolság kisebb, ezért a becslés értelemszerűen rosszabb.



4. ábra. Képpár a becsült normálvektorokkal (**Library** képsorozatból).



5. ábra. Becsült normálvektorok **House** (balra) és **Corridor** (középen) sorozatokra, illetve térbeli rekonstrukció (jobbra).

<sup>1</sup> <http://www.robots.ox.ac.uk/~vgg/data/>

A háromdimenziós felületeket a normálvektorok és a pontok segítségével is lehet rekonstruálni. Kipróbáltuk a MeshLab <sup>2</sup> szoftverbe beépített APSS módszert, a rekonstrukció eredménye a 4.2. ábrán látható.

#### Normálvektor becslése síkok esetén

A javasolt optimális (OPT) módszert épületek rekonstrukciójára szintén kipróbáltuk, ahogyan az a 6. ábrán látható. Épületek általában síkokból állnak, ezeket a síkokat pedig homográfia segítségével meg lehet becsülni. Tesztjeinkhez a SZTE tesztsorozatait [10] használtuk. A kapott homográfia elsőrendű parciális deriváltjaiból az affin transzformáció megkapható, ezek alapján elvégeztük a normálvektorok becslését különböző pontokban, melyeket a képekre fekete vagy fehér vonalakkal rá is rajzoltunk. (A pontok helyét háromszögeléssel kaptuk meg) Helyes homográfia esetén kivétel nélkül mindig jó irányba néző normálvektorokat kaptunk.

## 5.. Összefoglalás

Tanulmányunkban megmutattuk, hogyan lehetséges két kép közötti affin transzformáció segítségével a felület megfelelő pontjának normálvektorát megbecsülni. Több becslő eljárást is javasoltunk, az egyik közülük legkisebb négyzetes értelemben optimális megoldást ad, ha a kamerák külső és belső paramétereit egyaránt kalibrálták.

A teszteredmények alapján nyilvánvaló, hogy a normálvektor becslése érzékeny az affin transzformáció hibájára. Ezért a jövőben jelentős energiát szeretnénk a képalapú affin transzformáció becslésére fordítani. Másik ígéretes kutatási irány olyan térbeli felületek létrehozása, amelyek nem csak a rekonstruált pontfelhőket, hanem az affin transzformációból számolt normálvektorokat is figyelembe veszik. Reményeink szerint így pontosabb, valóságosabb térbeli modelleket kaphatók.

## Köszönetnyilvánítás

A kutatás az Európai Unió és Magyarország valamint az Európai Szociális Alap társfinanszírozásában a FuturICT.hu (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0013) projekt keretében valósult meg.

## Hivatkozások

1. R. J. Woodham, „Photometric stereo: A reflectance map technique for determining surface orientation from image intensity,” in *Image Understanding Systems and Industrial Applications, Proc. SPIE*, vol. 155, 1978, pp. 136–143.
2. B. Fodor, C. Kazó, J. Zsolt, and L. Hajder, „Normal map recovery using bundle adjustment,” *IET Computer Vision*, vol. 8, pp. 66 – 75, 2014.

<sup>2</sup> [www.meshlab.org](http://www.meshlab.org)



6. ábra. Becsült normálvektorok síkokkal határolt épületeken.

3. O. Faugeras and F. Lustman, „Motion and structure from motion in a piecewise planar environment,” INRIA, Tech. Rep. RR-0856, 1988. [Online]. Available: <http://hal.inria.fr/inria-00075698>
4. E. Malis and M. Vargas, „Deeper understanding of the homography decomposition for vision-based control,” INRIA, Tech. Rep. RR-6303, 2007.
5. H. Liu, „Deeper Understanding on Solution Ambiguity in Estimating 3D Motion Parameters by Homography Decomposition and its Improvement,” Ph.D. dissertation, University of Fukui, 2012.
6. M. Habbecke and L. Kobbelt, „Iterative multi-view plane fitting,” in *In VMV'06*, 2006, pp. 73–80.
7. —, „A surface-growing approach to multi-view stereo reconstruction,” in *CVPR*, 2007.
8. G. Z. Megyesi and D. Chetverikov, „Dense 3d reconstruction from images by normal aided matching,” *Machine Graphics and Vision*, vol. 15, pp. 3–28, 2006.
9. E. Kreyszig, *Differential geometry*. Dover Publications, 1991.
10. A. Tanacs, A. Majdik, J. Molnar, A. Rai, and Z. Kato, „Establishing correspondences between planar image patches,” in *International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, 2014.

**Optimális normálvektor becslése** Az optimális normálvektort becsülő eljárás feladata a 10. összefüggést minimumát megtalálni az  $n$  normálvektor szerint. A normálvektor skálája nem számít, mindössze az irányát szeretnénk meghatározni. Az ilyesfajta feladatokat tipikusan Langrange multiplikátor segítségével szokás elkészíteni, azonban itt ez nem járható, mert a kifejezés túl bonyolult lesz, zárt alakú megoldást nem fogunk kapni. Ezért az alábbi trükköt alkalmazzuk: A normálvektor hosszát ne egységnyinek válasszuk, hanem alkalmazzunk egy másik megkötést: legyen a koordináták összege egy. Azaz a vektor három koordinátáját így írhatjuk fel:  $n = [n_x, n_y, 1 - n_x - n_y]^T$ . A feladat ezzel a jelöléssel az alábbiak szerint változik:

$$\arg \min_m \sum_{k=1}^4 \left( \frac{m^T q_k + w_{k,z}}{\alpha m^T q_5 + \alpha w_{5,z}} - a_k \right)^2,$$

ahol az  $m = [n_x, n_y]$  és  $q_i = [w_{i,x} - w_{i,z}, w_{i,y} - w_{i,z}]^T$  jelöléseket vezettük be. (Az  $x$ ,  $y$  és  $z$  indexek az első, második és harmadik koordinátákat jelölik.)

A szélsőértéke a függvénynek az  $m$  vektor szerinti deriválással kapható meg:

$$2 \sum_{k=1}^4 \beta_k \gamma_k = 0$$

ahol

$$\beta_k = \left( \frac{m^T q_k + w_{k,z}}{\alpha m^T q_5 + \alpha w_{5,z}} - a_k \right)$$

$$\gamma_k = \left( \alpha \frac{(m^T q_5 + w_{5,z})q_k - (m^T q_k + w_{k,z})q_5}{(\alpha m^T q_5 + \alpha w_{5,z})^2} \right)$$

Ha a törtet összevonjuk a legkisebb közös többszörös segítségével,  $\sum_{k=1}^4 \delta_k \kappa_k = 0$  alakra hozhatjuk a fenti egyenletet, ahol

$$\begin{aligned}\delta_k &= (m^T q_k + w_{k,z} - a_k \alpha m^T q_5 - a_k \alpha w_{5,z}) \\ \kappa_k &= ((m^T q_5 + w_{5,z}) q_k - (m^T q_k + w_{k,z}) q_5)\end{aligned}$$

Ez pedig  $\sum_{k=1}^4 e_k^1 e_k^2 = 0$  alakra egyszerűsíthető, ahol

$$\begin{aligned}e_k^1 &= (m^T (q_k - a_k \alpha q_5) + (w_{k,z} - a_k \alpha w_{5,z})) \\ e_k^2 &= ((m^T q_5) q_k - (m^T q_k) q_5 + w_{5,z} q_k - w_{k,z} q_5)\end{aligned}$$

Tovább alakítva kétdimenziós vektoregyenlet kapható:

$$\sum_{k=1}^4 r \begin{pmatrix} m^T (q_5 q_{k,x} - q_k q_{5,x}) + w_{5,z} q_{k,x} - w_{k,z} q_{5,x} \\ m^T (q_5 q_{k,y} - q_k q_{5,y}) + w_{5,z} q_{k,y} - w_{k,z} q_{5,y} \end{pmatrix} = 0$$

ahol  $r = (m^T (q_k - a_k \alpha q_5) + (w_{k,z} - a_k \alpha w_{5,z}))$

Bevezetve az  $m = [x, y]^T$  jelölést, az alábbi formulát kapjuk:

$$\sum_{k=1}^4 (\Omega_k x + \Psi_k y + \Gamma_k) \begin{pmatrix} \Omega_k^1 x + \Psi_k^1 y + \Gamma_k^1 \\ \Omega_k^2 x + \Psi_k^2 y + \Gamma_k^2 \end{pmatrix} = 0$$

ahol

$$\begin{aligned}\Omega_k &= q_{k,x} - \alpha q_{5,x} a_k & \Psi_k &= q_{k,y} - \alpha q_{5,y} a_k \\ \Gamma_k &= w_{k,z} - a_k \alpha w_{5,z} & \Omega_k^1 &= 0 \\ \Psi_k^1 &= q_{5,y} q_{k,x} - q_{k,y} q_{5,x} & \Gamma_k^1 &= w_{5,z} q_{k,x} - w_{k,z} q_{5,x} \\ \Omega_k^2 &= q_{5,x} q_{k,y} - q_{k,x} q_{5,y} & \Psi_k^2 &= 0 \\ \Gamma_k^2 &= w_{5,z} q_{i,y} - w_{i,z} q_{5,y}\end{aligned}$$

Az egyenlet sorai speciális kvadratikus görbéket adnak meg. Ezeket felírhatjuk implicit egyenleteik segítségével:  $\sum_{k=1}^4 A_k^l x^2 + \sum_{k=1}^4 B_k^l y^2 + \sum_{k=1}^4 C_k^l xy + \sum_{k=1}^4 D_k^l x + \sum_{k=1}^4 E_k^l y + \sum_{k=1}^4 F_k^l = 0$ , ahol  $A_k^l = \Omega_k \Omega_k^l$ ,  $B_k^l = \Psi_k \Psi_k^l$ ,  $C_k^l = \Omega_k^l \Psi_k + \Psi_k^l \Omega_k$ ,  $D_k^l = \Omega_k^l \Gamma_k + \Gamma_k^l \Omega_k$ ,  $E_k^l = \Psi_k^l \Gamma_k + \Gamma_k^l \Psi_k$  and  $F_k^l = \Gamma_k \Gamma_k^l$ ,  $l \in 1, 2$ . Ezek azért speciálisak, mert  $A_k^1 = 0$  and  $B_k^2 = 0$ .

Az optimális megoldás két kvadratikus egyenlet megoldásából (kétdimenziós metszéspontból) számítható:

$$\begin{aligned}B_1 y^2 + C_1 x y + D_1 x + E_1 y + F_1 &= 0 \\ A_2 x^2 + C_2 x y + D_2 x + E_2 y + F_2 &= 0\end{aligned}$$

Az utóbbi egyenletből  $y$  kifejezhető:

$$y = -\frac{A_2 x^2 + D_2 x + F_2}{C_2 x + E_2}$$

$y$ -t az első egyenletbe behelyettesítve a következő kifejezés adódik:

$$\begin{aligned}
& B_1 \left( \frac{A_2 x^2 + D_2 x + F_2}{C_2 x + E_2} \right)^2 - \\
& C_1 x \frac{A_2 x^2 + D_2 x + F_2}{C_2 x + E_2} + D_1 x - \\
& E_1 \frac{A_2 x^2 + D_2 x + F_2}{C_2 x + E_2} + F_1 = 0
\end{aligned}$$

Ha mindkét oldalt megszorozzuk  $(C_2 x + E_2)^2$ -tel, az egyenletünk így módosul:

$$\begin{aligned}
& B_1 (A_2 x^2 + D_2 x + F_2)^2 - \\
& C_1 x (A_2 x^2 + D_2 x + F_2) (C_2 x + E_2) + \\
& D_1 x (C_2 x + E_2)^2 - \\
& E_1 (A_2 x^2 + D_2 x + F_2) (C_2 x + E_2) + F_1 (C_2 x + E_2)^2 = 0
\end{aligned}$$

Ez egy negyedfokú polinom, melynek az együtthatói a következők:

$$\begin{aligned}
x^4 : & \quad B_1 A_2^2 - C_1 A_2 C_2 \\
x^3 : & \quad 2B_1 A_2 D_2 - C_1 A_2 E_2 - \\
& \quad C_1 D_2 C_2 + D_1 C_2^2 - E_1 A_2 C_2 \\
& \quad B_1 D_2^2 + 2B_1 A_2 F_2 - \\
x^2 : & \quad C_1 D_2 E_2 - C_1 F_2 C_2 + 2D_1 C_2 E_2 - \\
& \quad E_1 A_2 E_2 - E_1 D_2 C_2 + F_1 C_2^2 \\
& \quad 2B_1 D_2 F_2 - C_1 F_2 E_2 + \\
x^1 : & \quad D_1 E_2^2 - E_1 D_2 E_2 - E_1 F_2 C_2 + \\
& \quad 2F_1 C_2 E_2 \\
x^0 : & \quad B_1 F_2^2 - E_1 F_2 E_2 + F_1 E_2^2
\end{aligned}$$

Azt a speciális esetet, amikor  $C_2 x + E_2 = 0$  szintén figyelembe kell venni. (Ebben az esetben könnyebb a dolgunk, hiszen az első egyenlet  $y$ -től nem függ,  $x$  lehetséges értékei könnyen kiszámíthatóak.  $y$  pedig az  $x$  értékeinek behelyettesével a második egyenletből jön ki.)