



Eötvös Loránd Tudományegyetem

Informatikai Kar

Algoritmusok és Alkalmazásaik Tanszék

---

# Síkdetektálás sztereó képpárokon affin transzformációk segítségével

Dr. Hajder Levente  
egyetemi docens, PhD

Gál István Gergő  
Programtervező Informatikus MSc

Budapest  
2019

# Tartalomjegyzék

<b>1. Bevezetés</b>	<b>1</b>
<b>2. Elméleti háttér</b>	<b>3</b>
2.1. Perspektivikus vetítés . . . . .	3
2.2. Sztereó geometria alapfogalmai . . . . .	4
2.3. Az esszenciális mátrix bevezetése. . . . .	6
2.4. A fundamentális és esszenciális mátrix becslése 8 pontos algoritmus segítségével . . . . .	7
2.5. Síkhomográfia és becslése . . . . .	9
2.6. Affin transzformációk használata homográfia becslésére . . . . .	12
2.7. Homográfia becslése affin megfeleltetéséből és fundamentális mátrixból	15
2.8. Robosztus módszerek használata HAF algoritmussal . . . . .	17
<b>3. Síkdetektálás</b>	<b>19</b>
3.1. Síkok detektálása homográfia segítségével . . . . .	20
3.2. Normalizált HAF algoritmus használata síkok detektálására . . . . .	23
<b>4. Homográfia becslés affin megfeleltetésekkel</b>	<b>25</b>
4.1. Homográfia becslése ismert normálvektorú vízszintes sík esetén . . . .	26
4.2. Homográfia becslése ismert normálvektorú függőleges sík esetén . . .	29
4.3. Homográfia becslése ismertlen normálvektorú függőleges sík esetén . .	31
<b>5. Homográfia becslés síkmozgás esetén</b>	<b>34</b>
5.1. Általános orientációjú sík esete . . . . .	35
5.2. Vízszintes sík esete . . . . .	37
5.3. Függőleges sík esete . . . . .	40
<b>6. Tesztelés</b>	<b>43</b>
6.1. Szintetikus tesztek . . . . .	43
6.1.1. Metsző síkok . . . . .	43
6.1.2. Párhuzamos síkok . . . . .	45
6.2. Valós tesztek . . . . .	47



# Ábrák jegyzéke

2.1. Különböző koordináta rendszerek perspektív leképezés esetén . . . . .	4
2.2. Epipoláris geometria . . . . .	5
2.3. Áttérés másik kamera koordináta rendszerére. . . . .	6
2.4. Sík homográfia becslése . . . . .	11
6.1. Kamerával szemben lévő hegyes szöget bezáró metsző síkok szimulációs elrendezése és a kapott eredmény. . . . .	44
6.2. Kamerával szemben lévő tompa szöget bezáró metsző síkok szimulációs elrendezése és a kapott eredmény. . . . .	45
6.3. A kamera előtt elhelyezkedő kamera síkjával párhuzamos síkok szimulációs elrendezése és a kapott eredmények. . . . .	46
6.4. A kamera előtt elhelyezkedő kamera síkjára merőleges síkok szimulációs elrendezése és a kapott eredmény. . . . .	47
6.5. A legnagyobb kiterjedésű sík meghatározása beltéri képpárok esetén .	48
6.6. A legnagyobb kiterjedésű sík meghatározása kültéri képpárok esetén .	48
6.7. Kvadropteres affin pontmegfeleltetések. . . . .	49
6.8. Különböző RANSAC paraméterek esetén detektált síkok. A felső sorban látható a <i>4 pontos</i> algoritmus. Az alsó sorban a 4.1-ben definiált algoritmus. Egy oszlopon belül azonosak a paraméterek. . . . .	49
6.9. 1. Kertvárosi kvadropteres felvétel. A felső 4.1-ben definiált algoritmus eredménye. Az alsó pedig a <i>4 pontos</i> algoritmusé. . . . .	50
6.10. 2. Kertvárosi kvadropteres felvétel. A felső 4.1-ben definiált algoritmus eredménye. Az alsó pedig a <i>4 pontos</i> algoritmusé. . . . .	51

# 1. fejezet

## Bevezetés

Az önvezető autózás napjaink egyik igen intenzíven kutatott kutatás-fejlesztési területe, melyben a számítógépes látás legújabb eredményeit alkalmazzák. A terület egyik alapproblémája hatékony algoritmusok kidolgozása síkok detektálására. A dolgozat célja sztereó képeken pontmegfeleltetések és hozzájuk tartozó affin transzformációk segítségével síkok detektálása. A munkám során egyaránt foglalkoztam síkdetektálással kalibrálatlan kamerák által felvett valós adatokon és kalibrált kamerás rendszerek szimulációs eredményein. Megvizsgáltam, hogy metsző síkok esetén a síkok detektálhatóságát hogyan befolyásolja az általuk bezárt szög, valamint párhuzamos síkok esetén milyen távolságra kell lenniük egymástól, hogy meg lehessen különböztetni őket. További vizsgálataim közé tartozik, hogy városi környezetben hogyan alkalmazhatók a dolgozatban javasolt módszerek síkok detektálására.

A dolgozat a 2. fejezetben a sztereó képpárokra történő síkok detektálásához szükséges elméleti háttér áttekintésével kezdődik, ahol ismertetem a szakirodalom e témakörhöz tartozó legfontosabb alapfogalmait és algoritmusait, különös figyelmet fordítva a homográfiák becslésének elméletére affin transzformációk segítségével<sup>[1],[2]</sup>. Ezt követően a 3. fejezetben ismertetem, hogy hogyan lehet a *RANSAC*<sup>[3]</sup> robusztus algoritmust használni sztereó képpárokra affin pontmegfeleltetésekkel síkok detektálására. Továbbá bemutatok egy a *HAF*<sup>[2]</sup> algoritmusra építkező síkdetektálási módszert. A 4. fejezetben megmutatom, hogy speciális esetekben a hagyományos pontmegfeleltetések helyett affin megfeleltetéseket használva a szakirodalomban ismert<sup>[4]</sup> módszereknél kevesebb megfeleltetés használatával megadható síkból síkba transzformációk homográfia mátrixa. Az 5. fejezetben megvizsgálom, hogyan befolyásolja a homográfia becslést, ha a kamerák sík mentén mozognak. Ezt követően az 6. fejezetben részletesen elemzem szimulációk segítségével sztereó képpárokra történő síkok detektálhatóságát, különös figyelmet fordítva a valós életben úttestnek, járdának, valamint épület falaknak megfelelő esetekre. Végül valós életbeli kvadrokopteres felvételeken demonstrálok a javasolt síkdetektálási módszert.

**Tudományos újdonságok a dolgozatban** a 4. és 5. fejezetben tárgyalt módszerek. Levezetem, hogy speciális, felülnézeti kamerás elrendezésekben, valamint síkmozgás esetén, hagyományos pontmegfeleltetések helyett affin pontmegfeleltetéseket használva kevesebb megfeleltetésből becsülhető a homográfia mátrixa.

## 2. fejezet

# Elméleti háttér

Ebben a fejezetben tárgyalom a későbbi fejezetek megértéséhez szükséges alapfogalmakat. Először bevezetem a képfeldolgozás során használt különböző koordináta-rendszereket és a köztük fennálló kapcsolatokat. Ezt követően leírom a sztereó geometria alapfogalmait, valamint definiálom a fundamentális és esszenciális mátrixokat, és azok becslését. Az alapvető sztereó geometriai fogalmak ismertetése után leírom az affin transzformációk fogalmát, valamint, hogy hogyan lehet pontmegfeleltetések helyett affin megfeleltetéseket használni. Közben bevezetem a homográfiák fogalmát és kapcsolatukat a síkok detektálásához.

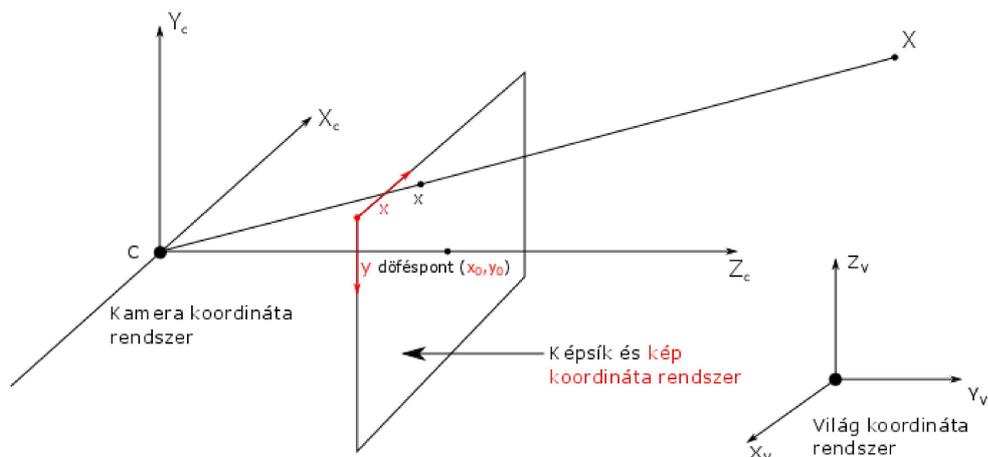
### 2.1. Perspektivikus vetítés

Ha adott egy kameránk  $\mathbf{C}$  kamera középponttal és egy  $\mathbf{X}$  térbeli pont, akkor homogén koordinátákat használva a kamera képsíkjának koordináta-rendszerében az  $\mathbf{X}$  pont képe megadható (2.1 ábra). Első lépésként a kamera koordináta-rendszeréből a kép koordináta rendszerébe vetítjük a pontot:

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} fk_x & 0 & x_0 \\ 0 & -fk_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}. \quad (2.1)$$

Ahol  $\mathbf{K}$  a kamera kalibrációs mátrixa, melyben  $f$  a kamera fókusztávolsága, ami a kamera origója és dőféspontja közötti távolság. A dőféspont az a pont, ahol a kamera koordináta-rendszerének  $Z$  tengelye metszi a kamera képsíkját.  $x_0, y_0$  a kamera dőféspontjának koordinátái a képsík koordináta-rendszerében.  $k_x, k_y$  arányossági tényezők, hogy mekkora távolság felel meg a kamera egy pixelének.

Ha a világ koordináta-rendszerében megadott  $\mathbf{X}$  vektort szeretnénk kifejezni a képsík koordináta-rendszerében, vetítés előtt ki kell fejeznünk a koordinátákat a kamera



2.1. ábra. Különböző koordináta rendszerek perspektív leképezés esetén

koordináta-rendszerében. Ez megtehető egy  $\mathbf{R}$  forgatás mátrix, valamint egy  $\mathbf{t}$  eltolás vektor segítségével:

$$\mathbf{X}_c = \mathbf{R}\mathbf{X} + \mathbf{t}. \quad (2.2)$$

$\mathbf{X}_c$ -t visszahelyettesítve a 2.1 egyenletbe:

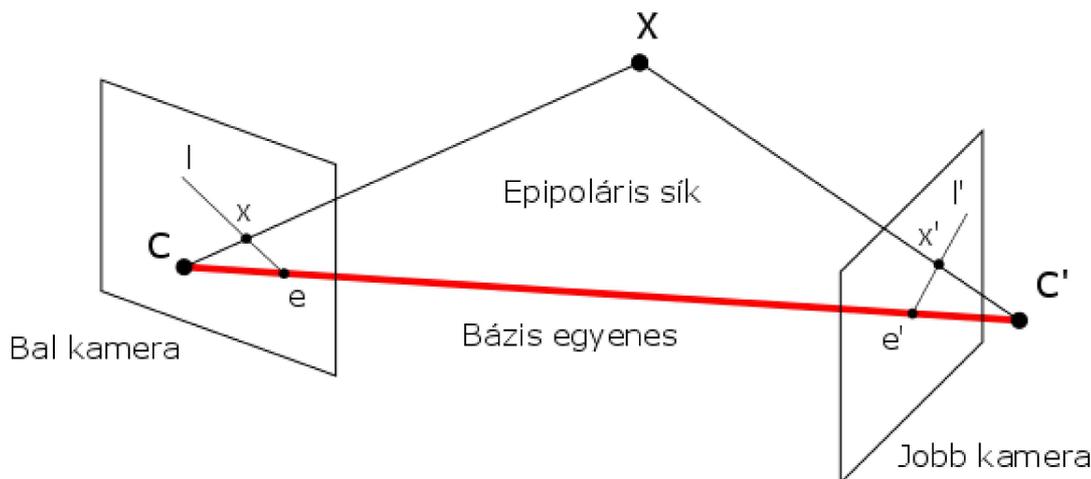
$$\mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{C}[\mathbf{R}|\mathbf{t}]\mathbf{X} = \mathbf{P}\mathbf{X}. \quad (2.3)$$

Ahol  $\mathbf{P}$  a kamera 3x4-es projekciós mátrixszá.

## 2.2. Sztereó geometria alapfogalmai

A számítógépes látás során sztereó képpárokról beszélünk akkor, ha ugyanazon 3D látványt leképező projektív kamerák vetítési középpontja nem esik egybe. Az ilyen kamerák által készített képpárokat sztereó képpároknak nevezzük. Az általános sztereó leképezés speciális esete, amihez a kamerák, valamint a képsíkok egy síkba esnek, és csak  $x$  irányú eltolás van köztük.

A kamerák által készített képek közötti kapcsolatot leíró geometriát epipoláris geometriának nevezzük. A két kamera középpontját összekötő egyenes a bázisegyenes, mely a képsíkokat az epipólusokban dőfi át. A kamerák középpontjai ( $\mathbf{C}, \mathbf{C}'$ ) a va-



2.2. ábra. Epipoláris geometria

lós térbeli  $\mathbf{X}$  pont és a képekre vetített pontjai  $(\mathbf{x}, \mathbf{x}')$  egy síkban vannak, melyet epipoláris síknak nevezünk. A képsíkok és az epipoláris sík metszetei az epipoláris egyenesek  $(\mathbf{l}, \mathbf{l}')$ , melyek az egyik kamera vetítősugarának képei a másik kamerában. Mindezeket az 2.2 ábrán láthatjuk.

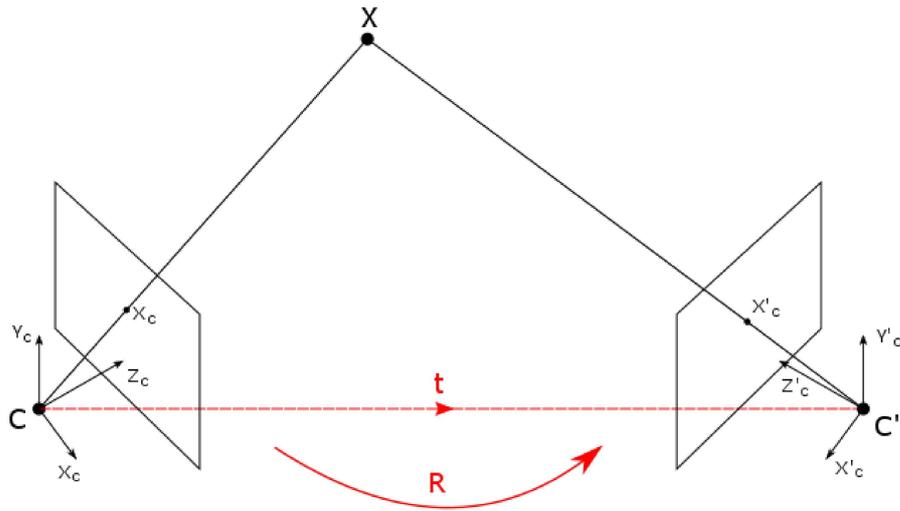
Érdeemes megjegyezni, hogy az epipoláris egyenesek és az epipoláris sík helyzete az  $\mathbf{X}$  három dimenziós ponttól függ. Minden  $\mathbf{X}$ -hez tartozik egy epipoláris sík és ezen epipoláris síkok egy közös egyenesben, a bázisegyenesben metszik egymást. Továbbá a különböző  $\mathbf{X}$  3D pontokhoz tartozó epipoláris egyenesek egy közös pontban metszik egymást, a kép epipólusában.

Ezekből következik, hogy a  $\mathbf{C}$  kamera középpont és az  $\mathbf{X}$  3D pont által meghatározott egyenesnek a másik  $\mathbf{C}'$  kamerán az  $\mathbf{l}'$  epipoláris egyenes felel meg. Így az  $\mathbf{X}$  pont  $\mathbf{x}'$  vetített képe a második kamerán az  $\mathbf{l}'$  egyenesen lesz. Összefoglalva tehát az epipoláris kapcsolat a két kamera között megad egy  $\mathbf{x}$  pontból  $\mathbf{l}'$  egyenesbe való leképezést, melyet algebrailag az  $\mathbf{F}$  fundamentális mátrix ad meg:

$$\mathbf{l}' = \mathbf{F}\mathbf{x}. \quad (2.4)$$

A fundamentális mátrix egy  $3 \times 3$ -as 2-es ranggal és 7 szabadsági fokkal rendelkező mátrix. Segítségével felírható a kapcsolat egy sztereó képpár  $\mathbf{x}$  és  $\mathbf{x}'$  pontjai között. Az  $\mathbf{x}'$  pont az  $\mathbf{l}'$  egyenesen található:

$$\mathbf{x}'^T \mathbf{l}' = 0. \quad (2.5)$$



2.3. ábra. Áttérés másik kamera koordináta rendszerére.

melybe  $\mathbf{l}'$ -t behelyettesítve kapjuk:

$$\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0. \quad (2.6)$$

Ez az eredmény egy megszorítást ad a fundamentális mátrixra, mely segítségével pontmegfeleltetések használatával a fundamentális mátrix becslhető lesz.

### 2.3. Az esszenciális mátrix bevezetése.

A 2.3 ábrán látható módon adott két kamera esetén a kamerák koordináta-rendszerei között egy  $\mathbf{R}$  forgatás mátrix és egy  $\mathbf{t}$  eltolás vektorral adható meg a kapcsolat:

$$\mathbf{X}'_C = \mathbf{R} \mathbf{X}_C + \mathbf{t}. \quad (2.7)$$

Ahol  $\mathbf{X}'_C$ ,  $\mathbf{X}_C$  jelöli az egyes kamera koordináta-rendszerekben azt a pontot, ahol a vetítő sugarak metszik a kamerák képsíkját. Ha az egyenletet balról vektoriálisan szorozzuk a  $\mathbf{t}$  eltolás vektorral ezt követően pedig skalárisan szorozzuk az  $\mathbf{X}'_C$ -vel:

$$\mathbf{X}'_C \cdot (\mathbf{t} \times \mathbf{R} \mathbf{X}_C) = 0. \quad (2.8)$$

Ami azt fejezi ki, hogy a  $\mathbf{C}\mathbf{X}_C$ ,  $\mathbf{C}'\mathbf{X}'_C$ ,  $\mathbf{C}\mathbf{C}'$  egy síkon fekszenek. Ha átírjuk a  $\mathbf{t}$  vektorral való vektoriális szorzást mátrixos alakra és bevezetjük az  $\mathbf{E} = [\mathbf{t}]_{\times}\mathbf{R}$  jelölést:

$$\mathbf{X}'_C[\mathbf{t}]_{\times}\mathbf{R}\mathbf{X}_C = \mathbf{X}'_C\mathbf{E}\mathbf{X}_C = 0. \quad (2.9)$$

Ahol a  $\mathbf{t}$  vektorral való vektoriális szorzás mátrixos alakban felírva általánosan is megadható. Legyen adott  $\mathbf{t}$  és  $\mathbf{x}$  vektorunk és fel szeretnénk írni vektoriális szorzatukat akkor:

$$\mathbf{t} \times \mathbf{x} = [\mathbf{t}]_x \mathbf{x}, \quad (2.10)$$

ahol:

$$[\mathbf{t}]_x = \begin{bmatrix} 0 & -t_z & -t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}. \quad (2.11)$$

Így a 2.9 egyenlet megadja az algebrai összefüggést a két kamera koordináta rendszerében vett képpontok között. Felhasználva a kamerák kalibrációs mátrixát ( $\mathbf{K}$ ) az egyenlet a következő alakban írható:

$$\mathbf{x}'^{-T}\mathbf{K}'^{-T}\mathbf{E}\mathbf{K}^{-1}\mathbf{x} = \mathbf{x}'^T\mathbf{F}\mathbf{x} = 0. \quad (2.12)$$

Ahol  $\mathbf{F}$  a már bevezetett fundamentális mátrix. Így felírható a kapcsolat a fundamentális és az esszenciális mátrix között:

$$\mathbf{F} = \mathbf{K}'^{-T}\mathbf{E}\mathbf{K}^{-1}. \quad (2.13)$$

Melyre szükségünk lesz a későbbiek során, mivel az esszenciális mátrix tartalmazza a külső kamera paramétereit így a fundamentális mátrix becsléséből, valamint a kamera kalibrációs mátrixából ki tudjuk majd számolni ezen paramétereit.

## 2.4. A fundamentális és esszenciális mátrix becslése 8 pontos algoritmus segítségével

Ha adott egy sztereó képpárunk és adottak ezen képpárokhoz tartozó olyan pontpárok  $(\mathbf{x}, \mathbf{x}')$ , melyek teljesítik az  $\mathbf{x}'^T\mathbf{F}\mathbf{x} = 0$  feltételt, azaz ugyanazon  $\mathbf{X}$  térbeli ponthoz tartoznak, akkor a fundamentális mátrix megbecsülhető. Homogén koordinátákat használva minden egyes pontmegfeleltetés egy feltételt ad a fundamentális mátrixra:

$$\begin{bmatrix} x'_i & y'_i & 1 \end{bmatrix} \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = 0. \quad (2.14)$$

Az egyenletben szereplő szorzásokat elvégezve, minden (n db) pontpárra felírva és átrendezve a következő  $\mathbf{A}\mathbf{f} = 0$  alakú lineáris egyenletrendszert kapjuk a fundamentális mátrix elemeire:

$$\begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots \\ x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} = 0. \quad (2.15)$$

A 2.15 egyenlet alapján a fundamentális mátrix legalább nyolc pont megfeleltetés alapján meghatározható egy skálázási tényező erejéig. Érdekes megjegyezni, hogy  $\mathbf{F}$  szabadsági foka 7 így 7 pontpár is elég lenne a meghatározásához, de ilyen esetben szükségessé válna  $\mathbf{F}$  szingularitásának felhasználása is és a megoldási módszer numerikusan instabil volna. Ebből kifolyólag általános esetben 8 vagy több pontmegfeleltetés alapján becsülhető a fundamentális mátrix.

Ha adott több mint 8 (zajjal terhelt) pontmegfeleltetésünk az egyenletrendszer túlhatározott alakú és a megoldása a következő módon kapható<sup>[5]</sup>:

$$\min_f \|\mathbf{A}\mathbf{f}\|^2. \quad (2.16)$$

feltéve, hogy  $\|\mathbf{f}\| = 1$ . Ez egy általános lineáris algebrai probléma, melynek megoldását az  $\mathbf{A}^T\mathbf{A}$  mátrix legkisebb sajátértékéhez tartozó sajátvektora adja. A sajátvektor megkapható az  $\mathbf{F}$  szinguláris érték dekompozíciójának (röviden SVD) felhasználásával:

$$\mathbf{F} = \mathbf{U}\mathbf{D}\mathbf{V}^T. \quad (2.17)$$

Ahol  $\mathbf{U}, \mathbf{V}$  az  $\mathbf{F}$  bal illetve jobb oldali sajátvektoraiból álló mátrix,  $\mathbf{D}$  pedig a sajátértékeket balról lefelé csökkenő sorrendben tartalmazó diagonális mátrix. A legkisebb sajátérték a  $\mathbf{D}$  utolsó eleme, melyhez tartozó sajátvektor a  $\mathbf{V}^T$  utolsó oszlopa.

Fontos megjegyezni, hogy az  $\mathbf{A}$  mátrix az egyes számot, zajjal terhelt pixelben mért koordinátákat, valamint azok szorzatát tartalmazza. Ennek következtében az elemek között akár több nagyságrendi különbség is lehet. Numerikus megoldás esetén ez instabilitáshoz vezethet valós adatok használatakor. Ezért érdemes a pixel térből egy transzformáció segítségével olyan reprezentációra áttérni, ahol az elemek közel azonos nagyságrendűek. A numerikus instabilitás elkerülhető, ha az  $(\mathbf{x}, \mathbf{x}')$  pontokat úgy transzformáljuk, hogy azok súlypontja az origóba essen valamit a súlypontjuktól

vett átlagos távolságuk  $\sqrt{2}$ -re normáljuk. Ezen transzformációk az  $(\mathbf{x})$  ponthalmazra a következőképpen adhatók meg:

$$\mathbf{T}_{\text{origo}} = \begin{bmatrix} 1 & 0 & -\mu_x \\ 0 & 1 & -\mu_y \\ 0 & 0 & 1 \end{bmatrix}; \quad \mathbf{T}_{\mathbf{d}} = \begin{bmatrix} \frac{\sqrt{2}}{d_{\text{atlag}}} & 0 & 0 \\ 0 & \frac{\sqrt{2}}{d_{\text{atlag}}} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.18)$$

Ahol  $\mu_{x,y}$  az  $(\mathbf{x})$  pontok x illetve y koordinátáinak átlaga,  $d_{\text{atlag}}$  pedig a súlypontjuktól vett átlagos távolsága. A két transzformáció egymás utáni végrehajtásával megkapjuk a normalizáló transzformációt:

$$\mathbf{T}_{\text{norm}} = \mathbf{T}_{\mathbf{d}} \mathbf{T}_{\text{origo}} = \begin{bmatrix} \frac{\sqrt{2}}{d_{\text{atlag}}} & 0 & -\mu_x \\ 0 & \frac{\sqrt{2}}{d_{\text{atlag}}} & -\mu_y \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.19)$$

Hasonlóan adható meg a  $\mathbf{T}'_{\text{norm}}$  transzformáció az  $(\mathbf{x}')$  pontokhoz. Ezeket felhasználva minden pontpár normalizálható, melyekkel a 8 pontos algoritmus numerikusan stabil. Az így kapott  $\hat{\mathbf{F}}$  fundamentális mátrix normalizált koordináták szerint lesz megadva. A  $\mathbf{F}$  pixeltérbeli koordináták szerint felírt fundamentális mátrix megadásához denormalizálásra van szükségünk. Normalizált koordinátás alakba átírva a 2.6 egyenletet az  $\hat{\mathbf{F}}$  fundamentális mátrixra a következőt kapjuk:

$$\mathbf{x}'^T \mathbf{T}'_{\text{norm}}{}^T \hat{\mathbf{F}} \mathbf{X}_{\text{norm}} \mathbf{x} = 0. \quad (2.20)$$

Melyet összehasonlítva a 2.6 egyenlettel megadható a kapcsolat  $\mathbf{F}$  és  $\hat{\mathbf{F}}$  között:

$$\mathbf{F} = \mathbf{T}'_{\text{norm}}{}^T \hat{\mathbf{F}} \mathbf{X}_{\text{norm}}, \quad (2.21)$$

valamint:

$$\hat{\mathbf{F}} = \mathbf{T}'_{\text{norm}}{}^{-T} \mathbf{F} \mathbf{X}_{\text{norm}}^{-1}. \quad (2.22)$$

A későbbiekben látni fogjuk, hogy a fundamentális mátrix becslése kevesebb pontból is megadható, ha pontmegfeleltetések helyett affin megfeleltetéseket használunk. Fontos megjegyezni továbbá, hogy ismert kamera paraméterek esetén az esszenciális mátrix is becsülhető a bemutatott *8 pontos* módszerrel a 2.13 egyenletet felhasználva.

## 2.5. Síkhomográfia és becslése

A későbbi fejezetekben látni fogjuk, hogy hogyan lehet a homográfiát használni síkok elkülönítésére és azok normális vektorának meghatározására ezért tekintsük a

2.4 ábrán látható elrendezést. Adott két kamerán és mindkét kamera látóterében egy  $\Pi$  sík. Szeretnénk megadni azt a transzformációt, amely az egyik kamera képterében látott  $\pi$  síkot áttranszformálja a másik kamara képterébe  $\pi'$  síkként. Tehát egy olyan transzformációt szeretnénk, ami minden  $\mathbf{x} \in \pi$ -hez megadja az  $\mathbf{x}' \in \pi'$  pontokat:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{H}\mathbf{x}. \quad (2.23)$$

Ahol  $\mathbf{H}$ -t a sík két képe között fennálló homográfia mátrixának nevezzük a kiszámítására pedig a jól ismert DLT (Direct Linear Transformation)<sup>[6]</sup> módszert fogjuk használni. Fontos megjegyezni, hogy az egyenlőség egy skála erejéig érvényes. Kiírva a mátrix szorzást és elvégezve a homogén osztást:

$$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}; \quad y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}. \quad (2.24)$$

A nevezővel beszorozva:

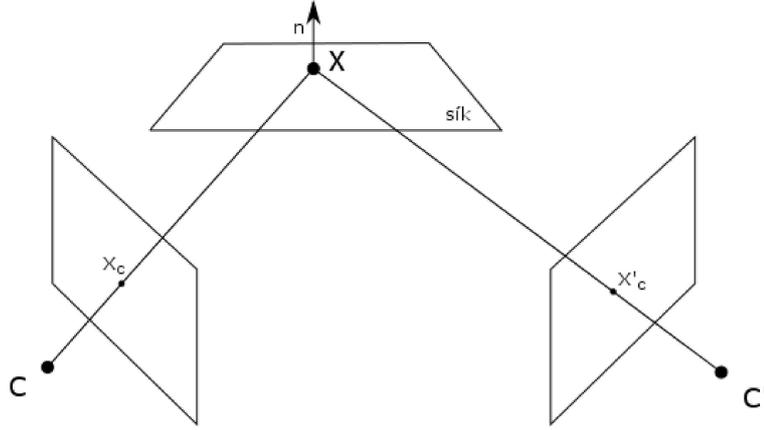
$$\begin{aligned} x'(h_{31}x + h_{32}y + h_{33}) &= h_{11}x + h_{12}y + h_{13}; \\ y'(h_{31}x + h_{32}y + h_{33}) &= h_{21}x + h_{22}y + h_{23}. \end{aligned} \quad (2.25)$$

Ebből következik, hogy minden pontmegfeleltetés  $(\mathbf{x}, \mathbf{x}')$  két megszorítást ad a  $\mathbf{H}$  mátrixra így,  $n \geq 4$  pont elegendő a homográfia becsléséhez.

A megoldásra két lehetőségünk van. Az első, hogy rögzítjük a homográfia mátrix egyik elemét, például a  $h_{33}$ -at. Ekkor  $n=4$  esetben egy lineáris egyenletrendszer kell megoldanunk,  $n>4$  esetben pedig a megoldást pseudo-inverz segítségével kaphatjuk. Ezzel a megközelítéssel az a probléma, hogy a valós megoldásban lehet  $h_{33} = 0$ , mely esetben ez nem elérhető ilyen megoldási módszerrel.

Másik megközelítés, hogy a homográfiát egy skála erejéig határozzuk meg. Más alakban írva a 2.25 egyenletet:

$$\begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -x'x & -x'y & -x' \\ 0 & 0 & 0 & x & y & 1 & -y'x & -y'y & -y' \end{bmatrix} \mathbf{h} = 0, \quad (2.26)$$



2.4. ábra. Sík homográfia becslése

ahol  $\mathbf{H} = (h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}, h_{33})^T$  a  $\mathbf{H}$  mátrix vektor alakban írva, így 4 pont esetén:

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1x_1 & -x'_1y_1 & -x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1x_1 & -y'_1y_1 & -y'_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x'_2x_2 & -x'_2y_2 & -x'_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y'_2x_2 & -y'_2y_2 & -y'_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x'_3x_3 & -x'_3y_3 & -x'_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -y'_3x_3 & -y'_3y_3 & -y'_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x'_4x_4 & -x'_4y_4 & -x'_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -y'_4x_4 & -y'_4y_4 & -y'_4 \end{bmatrix} \mathbf{h} = 0. \quad (2.27)$$

Az egyenlet  $\mathbf{A}\mathbf{h} = 0$  alakú, ahol az  $\mathbf{A}$  egy  $8 \times 9$ -es mátrix. A megoldás vektor  $\mathbf{h}$  így az  $\mathbf{A}$  mátrix egy dimenziós null tere. Ha  $n > 4$   $\mathbf{A}$  egy  $2n \times 9$ -es mátrix, így nem kapható egy  $\mathbf{A}\mathbf{h} = 0$  alakú megoldás. Ebben az esetben a korábbiakhoz hasonlóan megmutatható hogy a megoldásvektor az a vektor, mely a  $\mathbf{h}^T \mathbf{A}^T \mathbf{A} \mathbf{h}$ -t minimalizálja feltéve, hogy  $\|\mathbf{h}\| = 1$ , ami az  $\mathbf{A}^T \mathbf{A}$  legkisebb sajátértékhez tartozó sajátvektora lesz. Ha  $n=4$ , akkor ez a sajátérték 0, így a sajátvektor az  $\mathbf{A}$  mátrix nulltere.

Fontos továbbá, hogy a becsült homográfából kiszámíthatók a kamera külső paramétereire és a sík normális vektora. A homográfia egy síkból síkba leképezés, ezért a valós térben lévő sík az egyes kamerák képsíkjára való leképezése is felírható homográfia alakjában:

$$\mathbf{x} = \mathbf{H}_1 \mathbf{X}; \quad \mathbf{x}' = \mathbf{H}_2 \mathbf{X}. \quad (2.28)$$

Behelyettesítve kapható:

$$\mathbf{x}' = \mathbf{H}_2 \mathbf{X} = \mathbf{H}_2 \mathbf{H}_1^{-1} \mathbf{x} = \mathbf{H} \mathbf{x}. \quad (2.29)$$

Továbbá ha ismert a kamera kalibrációs mátrixa:

$$\mathbf{H} = \mathbf{K}'(\mathbf{R} + \mathbf{t} \cdot \mathbf{n}/d)\mathbf{K}^{-1}. \quad (2.30)$$

A későbbiekben ez lehetőséget ad a becsült homográfiákból kapott normálvektorok használatára síkdetektálás során.

## 2.6. Affin transzformációk használata homográfia becslésére

Eddig  $(\mathbf{x}, \mathbf{x}')$  pontmegfeleltetéseket használtunk a fundamentális, esszenciális és a homográfia becslésére. A pontmegfeleltetések csak a térbeli elhelyezkedés információját hordozzák magukban, felmerül a kérdés, hogy ha más információ is rendelkezésre áll, például az információ, hogy pontmegfeleltetésekhez tartozó lokális környezet hogyan transzformálódik tudunk-e pontosabb becslést adni, vagy kevesebb pontra lesz-e szükségünk?

Legyen adott a korábbiakhoz hasonlóan  $(\mathbf{x}, \mathbf{x}')$  pontmegfeleltetés és legyen továbbá adott ehhez a megfeleltetéshez tartozó lokális környezetek transzformációját leíró  $\mathbf{A}$  affin transzformáció:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & t_{13} \\ a_{21} & a_{22} & t_{23} \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.31)$$

Az affin transzformációkra nézve invariánsak a párhuzamos egyenesek; a területek aránya; az egy síkban vagy párhuzamos egyeneseken mért távolságok valamint a vektorok lineáris kombinációja. Két részből áll, az első rész, későbbiekben csak ezt fogjuk affin transzformációnak nevezni  $\mathbf{A} = a_{i,j}$ , ahol  $i,j = 1,2$ . A négy paraméter a hosszanti, függőleges skálázásért, forgatásért és nyírásért felel. A második része adja az affin transzformáció eltolását.

A homográfia és az affin transzformáció kapcsolata a következő módon kapható<sup>[1]</sup>: A tömörebb írásmód érdekében írjuk fel a  $\mathbf{H}$  mátrixot a következő alakban:

$$\mathbf{H} = \begin{bmatrix} h_1^T \\ h_2^T \\ h_3^T \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}. \quad (2.32)$$

Így az  $\mathbf{x}'$  vektor  $x'$  és  $y'$  komponense:

$$x' = \frac{h_1^T[x, y, 1]^T}{h_3^T[x, y, 1]^T} \quad y' = \frac{h_2^T[x, y, 1]^T}{h_3^T[x, y, 1]^T}. \quad (2.33)$$

Az affin komponensek a parciális deriváltjai a perspektív sík-sík transzformációnak, így:

$$\begin{aligned} a_{11} &= \frac{\partial x'}{\partial x} = \frac{h_{11}h_3^T[x, y, 1]^T - h_{31}h_1^T[x, y, 1]^T}{(h_3^T[x, y, 1]^T)^2} = \frac{h_{11} - h_{31}x'}{s}, \\ a_{12} &= \frac{\partial x'}{\partial y} = \frac{h_{12} - h_{32}x'}{s}, \\ a_{21} &= \frac{\partial y'}{\partial x} = \frac{h_{21} - h_{31}y'}{s}, \\ a_{22} &= \frac{\partial y'}{\partial y} = \frac{h_{22} - h_{32}y'}{s}, \end{aligned} \quad (2.34)$$

ahol:

$$s = h_3^T[x, y, 1]^T. \quad (2.35)$$

Felszorozva  $s$ -el és átrendezve 4 darab lineárisan független<sup>[7]</sup> egyenletet kapunk  $\mathbf{H}$  elemeire:

$$\begin{aligned} h_{11} - h_{31}(x' + a_{11}x) - h_{32}a_{11}y - h_{33}a_{11} &= 0, \\ h_{12} - h_{32}(x' + a_{12}y) - h_{31}a_{12}x - h_{33}a_{12} &= 0, \\ h_{21} - h_{31}(y' + a_{21}x) - h_{32}a_{21}y - h_{33}a_{21} &= 0, \\ h_{22} - h_{32}(y' + a_{22}y) - h_{31}a_{22}x - h_{33}a_{22} &= 0. \end{aligned} \quad (2.36)$$

Az egyenletrendszer nem tartalmazza azonban  $\mathbf{H}$  összes elemét, a  $h_{13}$  és a  $h_{23}$  hiányzik belőle. A hiányzó két elem reprezentálja az affin transzformáció eltolását, melyet a korábban megismert DLT algoritmusból megkaphatunk:

$$\begin{aligned} h_{11}x + h_{12}y + h_{13} - h_{31}xx' - h_{32}yx' - h_{33}x' &= 0, \\ h_{21}x + h_{22}y + h_{23} - h_{31}xy' - h_{32}yy' - h_{33}y' &= 0. \end{aligned} \quad (2.37)$$

A 2.36 és 2.37 egyenletek  $\mathbf{H}$ -ra kapcsolatot teremtenek a pontok koordinátái, az affin transzformáció elemei és a homográfia mátrixa között. Hat egyenlet kapható így egy affin pontmegfeleltetésből (AC-Affin correspondance), így összesen két AC-ből megbecsülhető  $\mathbf{H}$ , ami jelentős különbség az egyszerű DLT-hez képest, ahol négy pont megfeleltetésre volt szükségünk. A pontmegfeleltetések száma robusztus például RANSAC<sup>[3]</sup> algoritmusok használatakor igen jelentős, ahol erősen függ az elvégzendő iterációk száma a szükséges pontok számától.

Homogén lineáris egyenlet rendszer formájában ( $\mathbf{B}h = \mathbf{0}$ ) felírva:

$$Bh = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -(x' + a_{11}x') & -a_{11}y & -a_{11} \\ 0 & 1 & 0 & 0 & 0 & 0 & -a_{12}x & -(x' + a_{12}y') & -a_{12} \\ 0 & 0 & 0 & 1 & 0 & 0 & -(y' + a_{21}x) & -a_{21}y & -a_{21} \\ 0 & 0 & 0 & 0 & 1 & 0 & -a_{11}x & -(y' + a_{22}y) & -a_{22} \\ x & y & 1 & 0 & 0 & 0 & -xx' & -yx' & -x' \\ 0 & 0 & 0 & x & y & 1 & -xy' & -yy' & -y' \end{bmatrix} h = 0. \quad (2.38)$$

Tehát az optimális megoldást *HA (Homogphy from Affine transformation)*<sup>[1]</sup> eljárás során a korábbikhoz hasonlóan a  $\|\mathbf{h}\| = 1$  feltétel mellett a  $\mathbf{B}^T\mathbf{B}$  legkisebb sajátértékező tartozó sajátvektora adja.

A hagyományos négy ponton alapuló homográfia becsléshez hasonlóan itt is szükséges a felhasznált adatok (pontmegfeleltetések és affin transzformációk) normalizálása. Továbbá a numerikus stabilitás érdekében az affin transzformációk normalizálása is szükséges. A normalizálás elvégezhető a pontmegfeleltetések normalizálásakor bevezetett transzformációs mátrixok segítségével. <sup>[2]</sup>

A normalizált affin transzformációk kiszámítása nem triviális feladat, hiszen az affin transzformációk 2.34 levezetésénél láthattuk, hogy függenek a még ki nem számított normalizált  $\hat{\mathbf{H}}$  homográfia mátrixtól. A 2.34 egyenleteket átrendezve kapjuk:

$$\begin{aligned} (h_{31}x + h_{32}y + h_{33})a_{11} &= h_{11} - h_{31}x', \\ (h_{31}x + h_{32}y + h_{33})a_{12} &= h_{12} - h_{32}x', \\ (h_{31}x + h_{31}y + h_{33})a_{21} &= h_{21} - h_{31}y', \\ (h_{31}x + h_{32}y + h_{33})a_{22} &= h_{22} - h_{32}y'. \end{aligned} \quad (2.39)$$

Mely egyenleteket a 2.19-ben definiált normalizáló transzformáció a következőképpen módosítja:

$$\begin{aligned} (h_{31}x + h_{32}y + h_{33})\hat{a}_{11} &= \frac{l'_x}{l_x}h_{11} - \frac{l'_x}{l_x}h_{31}x', \\ (h_{31}x + h_{32}y + h_{33})\hat{a}_{12} &= \frac{l'_x}{l_y}h_{12} - \frac{l'_x}{l_y}h_{32}x', \\ (h_{31}x + h_{31}y + h_{33})\hat{a}_{21} &= \frac{l'_y}{l_x}h_{21} - \frac{l'_y}{l_x}h_{31}y', \\ (h_{31}x + h_{32}y + h_{33})\hat{a}_{22} &= \frac{l'_y}{l_y}h_{22} - \frac{l'_y}{l_y}h_{32}y'. \end{aligned} \quad (2.40)$$

Ahol  $l_x = T_{norm,11}$ ,  $l_y = T_{norm,22}$ ,  $l'_x = T'_{norm,11}$ ,  $l'_y = T'_{norm,22}$  a melyek a normalizáció transzformációk vízszintes és függőleges skálái. A 2.39 és 2.40 egyenleteket összevetve megadható kapcsolat a normalizált és normalizálatlan affin transzformá-

ciók komponensei között:

$$\begin{aligned} \hat{a}_{11} &= \frac{l'_x}{l_x} a_{11}, & \hat{a}_{12} &= \frac{l'_x}{l_y} a_{12}, \\ \hat{a}_{21} &= \frac{l'_y}{l_x} a_{21}, & \hat{a}_{22} &= \frac{l'_y}{l_y} a_{22}. \end{aligned} \quad (2.41)$$

A kapott  $\hat{\mathbf{H}}$  homográfia mátrix a normalizált koordináták és affin transzformációk szerint adott így  $\mathbf{H}$  megkapásához denormalizációra van szükségünk:

$$\mathbf{H} = \mathbf{T}'^{-1} \hat{\mathbf{H}} \mathbf{T}_{\text{norm}}. \quad (2.42)$$

Összegezve tehát sikerült megadnunk egy módszert homográfia becslésére affin pontmegfeleltetések használatával. Ennek nagy előnye, hogy 4 hagyományos pontmegfeleltetés helyett elegendő 2 affin megfeleltetés.

## 2.7. Homográfia becslése affin megfeleltetéséből és fundamentális mátrixból

Ebben a fejezetben bemutatom a kapcsolatot a perspektivikus homográfia és a fundamentális mátrix között. Az epipoláris geometriában jól ismert, hogy a  $\mathbf{F}$  fundamentális mátrix és a  $\mathbf{H}$  homográfia mátrix közötti kapcsolat felírható<sup>[6]</sup>:

$$[\mathbf{e}']_x \mathbf{H} = \lambda \mathbf{F}, \quad (2.43)$$

ahol  $\mathbf{e}' = [e'_x, e'_y, 1]^T$  a második kép epipólusa, valamint  $\lambda$  egy arányossági tényező (skalár). Az  $[\mathbf{e}']_x$  a 2.11 egyenletben bevezetett asszimmetrikus vektoriális szorzat mátrixa az  $\mathbf{e}'$  vektornak. Általában ez az alak arra használható, hogy a homográfia három pontmegfeleltetésből becsülhető legyen<sup>[6]</sup>. A következőkben megmutatom, hogyan lehet a homográfiát becsülni egy affin megfeleltetésből és a fundamentális mátrixból<sup>[2]</sup>.

Mivel az  $[\mathbf{e}']_x$  rangja kettő, ezért a 2.43 egyenletrendszer harmadik sora elhagyható, így:

$$\begin{bmatrix} 0 & -1 & e'_y \\ 1 & 0 & -e'_x \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} = \lambda \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \end{bmatrix}. \quad (2.44)$$

Ennek következményeként a szabadsági fokok száma háromra redukálódik. A  $\mathbf{H}$  homográfia mátrixnak első két sora megadható a harmadik sorának és a fundamentális

mátrix komponenseinek felhasználásával:

$$\begin{aligned} h_{11} &= e'_x h_{31} + \lambda f_{21}, & h_{21} &= e'_y h_{31} - \lambda f_{11}, \\ h_{12} &= e'_x h_{32} + \lambda f_{22}, & h_{22} &= e'_y h_{32} - \lambda f_{12}, \\ h_{13} &= e'_x h_{33} + \lambda f_{23}, & h_{23} &= e'_y h_{33} - \lambda f_{13}. \end{aligned} \quad (2.45)$$

Mind a fundamentális és a homográfia mátrix csak egy skála erejéig határozható meg. Ennek következtében egy skálát tetszőleges értékűre választhatunk, ezért legyen  $\lambda = 1$ .

A 2.45 egyenletet a 2.29 egyenletbe helyettesítve inhomogén lineáris egyenletrendszert kapunk  $\mathbf{H}$  elemeire:

$$\begin{aligned} (xe'_x - xx')h_{31} + (ye'_x - yy')h_{32} + (e'_x - x')h_{33} &= -xf_{21} - yf_{22} - f_{23}, \\ (xe'_y - xy')h_{31} + (ye'_y - yy')h_{32} + (e'_y - y')h_{33} &= xf_{11} + yf_{12} + f_{13}. \end{aligned} \quad (2.46)$$

Azonban az epipoláris kényszer miatt a pontpároknak a megfelelő epipoláris egyenesen kell elhelyezkedniük, ezért az egyenletek közül csak az egyik hordoz új információt. Ennek ellenére érdemes lehet mindkét egyenletet felhasználni a homográfia becslése során a zaj hatásának csökkentése érdekében. Ezeket az egyenleteket felhasználva egy a *DLT*-hez hasonló megoldás módszer adható meg. Mivel három pontmegfeleltetés elegendő a homográfia mátrix becsléséhez, ezért ez az algoritmus *3 pontos (3PT)* algoritmusnak nevezzük.

A szükséges megfeleltetések számát tovább lehet csökkenteni abban az esetben, ha pontmegfeleltetések helyett affin megfeleltetéseket használunk. A 2.45 egyenletet a 2.36 egyenletbe helyettesítve:

$$\begin{aligned} e'_x h_{31} + f_{21} - h_{31}(x' + a_{11}x) - h_{32}a_{11}y - h_{33}a_{11} &= 0, \\ e'_x h_{32} + f_{22} - h_{32}(x' + a_{12}y) - h_{31}a_{12}x - h_{33}a_{12} &= 0, \\ e'_y h_{31} - f_{11} - h_{31}(y' + a_{21}x) - h_{32}a_{21}y - h_{33}a_{21} &= 0, \\ e'_y h_{32} - f_{12} - h_{32}(y' + a_{22}y) - h_{31}a_{22}x - h_{33}a_{22} &= 0. \end{aligned} \quad (2.47)$$

További 4 lineárisan független egyenletet kapunk  $\mathbf{H}$  mátrix utolsó sorára. Lineáris, inhomogén egyenletrendszer mátrixos alakjába rendezve:

$$\begin{bmatrix} e'_x - x' - a_{11}x & -a_{11}y & -a_{11} \\ -a_{12}x & e'_x - x' - a_{12}y & -a_{12} \\ e'_y - y' - a_{21}x & -a_{21}y & -a_{21} \\ -a_{22}x & e'_y - y' - a_{22}y & -a_{22} \end{bmatrix} \begin{bmatrix} h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} = \begin{bmatrix} -f_{21} \\ -f_{22} \\ f_{11} \\ f_{12} \end{bmatrix}. \quad (2.48)$$

Az egyenletrendszerhez hozzávéve a 2.46 mátrixos alakját  $\mathbf{C}\mathbf{y} = \mathbf{d}$  alakú egyenletrendszert kapunk  $\mathbf{H}$  utolsó sorának elemeire:

$$\mathbf{C}\mathbf{y} = \mathbf{d}$$

$$\begin{bmatrix} e'_x - x' - a_{11}x & -a_{11}y & -a_{11} \\ -a_{12}x & e'_x - x' - a_{12}y & -a_{12} \\ e'_y - y' - a_{21}x & -a_{21}y & -a_{21} \\ -a_{22}x & e'_y - y' - a_{22}y & -a_{22} \\ xe'_x - xx' & ye'_x - yx' & e'_x - x' \\ xe'_y - xy' & ye'_y - yy' & e'_y - y' \end{bmatrix} \begin{bmatrix} h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} = \begin{bmatrix} -f_{21} \\ -f_{22} \\ f_{11} \\ f_{12} \\ -xf_{21} - yf_{22} - f_{23} \\ xf_{11} + yf_{12} + f_{13} \end{bmatrix}. \quad (2.49)$$

Az optimális megoldás legkisebb négyzetes értelemben megadható:

$$\mathbf{y} = \mathbf{C}^\dagger \mathbf{d}, \quad (2.50)$$

ahol  $\mathbf{C}^\dagger$  a  $\mathbf{C}$  mátrix Moore-Penrose pszeudo inverze. Mivel az  $\mathbf{y}$  megoldásvektor a  $\mathbf{H}$  mátrix utolsó sora, így 2.45 segítségével a homográfia összes eleme megadható. Egy egyenletünk már adott az utolsó sorra a pontmegfeleltetésekből (2.46), így a 2.47 egyenletekből csupán kettőre lenne szükségünk. Érdemes azonban megtartani a 2.47-ben szereplő mind a négy lineárisan független egyenletet a zaj hatásának csökkentése érdekében.

Összefoglalva tehát, affin megfeleltetéseket használva, a fundamentális mátrix ismeretében csupán egy megfeleltetésből megadható a homográfia mátrixa.

A most bemutatott *HAF* homográfia becselő módszer a korábbi módszerekhez hasonlóan pont és affin megfeleltetéseket használ bemenetként. Lehetőség van más, összetettebb objektumok használatára is, mint például egyenesek<sup>[6]</sup>, régiók<sup>[8]</sup>, kontúrok<sup>[9]</sup>, valamint kúpszeletek<sup>[10],[11]</sup>. Azonban ilyen objektumok detektálása komplexebb feladat, így nehezebb valós életbeli adatok esetén bemenetként használni őket.

## 2.8. Robosztus módszerek használata HAF algorit-mussal

Az előző alfejezetben bemutatott *HAF* algoritmus előnye, hogy a szükséges megfeleltetések számának minimuma egy. Ebből következik, hogy a globálisan optimális homográfia mátrix megtalálható, egy egyszerű lineáris keresés segítségével a homog-

ráfia mátrixok között:

$$\mathbf{H}_{\text{opt}} = \arg \min_{\mathbf{H} \in \mathcal{H}} \sum_{i=1}^n I(\mathbf{H}, \epsilon) \quad (2.51)$$

ahol  $\mathcal{H}$  a homográfiák halmaza,  $I$  az *inlier*-eket megszámloló függvény és  $\epsilon$  a hibahatár. Fontos megjegyezni, hogy ez a *RANSAC*-hoz hasonlóan más robusztus módszerek esetén is alkalmazható, mint például *MLESAC*<sup>[12]</sup>, *LMeDS*<sup>[13]</sup>, *J-Linkage*<sup>[14]</sup> és *T-Linkage*<sup>[15]</sup>.

A szakirodalom áttekintése után a következő fejezetekben az eddigi elméleti háttérre építkezve definiálni fogok egy síkdetektor algoritmust, ami a homográfia becslésre építkezve fog szeparálni síkokat sztereó képpárokban. A homográfia mátrixának becslése így különös fontosságot kap, ezért a 4 valamint 5 fejezetekben megmutatom, hogy bizonyos speciális esetekben a homográfia mátrixa a fundamentális mátrix ismerete nélkül is megadható két pont megfeleltetés helyett egy affin megfeleltetésből. Továbbá tesztelni fogom szintetikus és valós adatokon a homográfia mátrix becslésére támaszkodó algoritmus használhatóságát.

## 3. fejezet

# Síkdetektálás

Egy olyan detektor létrehozása, ami képes sztereó képpárokon csak a képek és kamerák információjára támaszkodva megkülönböztetni párhuzamos, de nem egybeeső síkokat különösen fontos, például az önvezető autózás során az úttest, valamint a járda ilyen síkok. Továbbá a megtalált síkok felhasználhatók a megfigyelt 3D-s világ síkokból álló reprezentálására. A kapott síkreprezentáció számos további alkalmazás bemeneteként szolgálhat. Lehetőségünk van a kamera képek és a síkreprezentáció használatával megfigyelt világ 3D rekonstrukciójára<sup>[16],[17]</sup>. Felhasználható továbbá a kamera belső paramétereinek meghatározására, segítségével kamera kalibráció végezhető<sup>[18],[19],[20]</sup>. A detektált síkokból alkotott reprezentáció felhasználható kiterjesztett valóság létrehozása során<sup>[21]</sup>. Mindezekon felül alkalmazható robotlátás során<sup>[22]</sup>, valamint beltéri navigáció esetén is<sup>[23]</sup>.

Az elméleti áttekintés fejezetben láttuk, hogy ha adott egy sztereó képpárunk, és a kamerák előtt egy sík található, akkor a két felvételen a sík vetületei között egy  $\mathbf{H}$  homográfia teremt kapcsolatot. Ez a homográfia felbontható egy  $\mathbf{R}$  forgatás mátrixra  $\mathbf{t}$  eltolás és a sík  $\mathbf{n}$  normális vektorára, továbbá tartalmazza még az első kamera és a sík közötti  $d$  távolságot. Tehát ha adott egy sztereó képpárunk és a képpáron különböző síkok találhatók, ezek alapján meg lehet különböztetni őket, mivel a sík normálisa tartalmazza a sík irányultságát. Ezen információ alapján meg lehet különböztetni nem párhuzamos síkokat. A sík irányultságán felül a homográfia tartalmazza még a  $d$  távolságot, ami alapján el lehet különíteni a párhuzamos, de nem egybeeső síkokat.

Ebben a fejezetben erre a problémára mutatok be két algoritmust, melyek közül az első csak a  $\mathbf{H}$  homográfia mátrixra támaszkodik, és ez alapján határozza meg a képpárokon található domináns síkokat. Majd ezt követően a második a homográfián felül még a képpontok  $\mathbf{x}, \mathbf{x}'$  valós térbe visszavetítését, és az így kapott  $\mathbf{X}$  vektort is felhasználja.

### 3.1. Síkok detektálása homográfia segítségével

*RANSAC*<sup>[3]</sup> algoritmust használva lehetőségünk van homográfia segítségével a legnagyobb síkot detektálni sztereó képpárokon. Legyenek adottak hagyományos pontmegfeleltetések (*PC*), melyek a két képen egymásnak megfelelő  $(\mathbf{x}, \mathbf{x}')$  pontokat tartalmaznak, vagy affin pontmegfeleltetések (*AC*), ahol a pont megfeleltetésekhez még egy  $\mathbf{A}$  affin transzformáció is tartozik. Ezek a bemenetként szolgáló affin megfeleltetések több módon is megkaphatók. Előállításukra használhatók a jól ismert *affin kovariáns* (affine covariant) detektorok, mint például az *MSER*, *Hessian-Affine* vagy a *Harris-Affine*<sup>[24]</sup> detektorok. Továbbá lehetőségünk van modernebb megközelítést használó *ASFIT*<sup>[25]</sup>, *ASURF*, *MODS*<sup>[26]</sup> módszerek használatára is, melyek a képpárok *jellemző részeit* (feature point) használják affin megfeleltetések keresésére. A megfeleltetésekből számított  $\mathbf{H}$  homográfia mátrixok segítségével, sztereó képpáron a legtöbb pontot tartalmazó sík az 1. algoritmus alapján megtalálható.

---

**Algoritmus 1:** Legnagyobb sík detektálása RANSAC használatával  
(LPDRANSAC - Largest Plane Detection with RANSAC)

---

**Bemenet:**

PC vagy AC megfeleltetések.  
 $maxIter$  - RANSAC algoritmus maximális iteráció száma,  
 $threshold$  - maximális reprojekciós hiba,  
 $confidence$  - konfidencia százalék.

**Kimenet:**

$S$  - A legdominánsabb síkhoz tartozó pontok.

**Algoritmus:**

1. Normalizáljuk a bemeneti  $PC - (P_1, P_2)$  vagy  $AC - (P_1, P_2, A)$  párokat.  
 $I :=$  üres,  
 $iterációszám := 0$ ,  
 $legjobbLefedettség := 0$ ,  
 $legnagyobbSík :=$  üres.
  2. Válasszunk véletlenszerűen DLT használata esetén 4, HA esetén 2 megfeleltetést. Legyenek ezekhez a megfeleltetésekhez tartozó pontok a  $randomSík$  pontjai.  
 $randomSík :=$  Kiválasztott megfeleltetések halmaza.
  3. Számítsuk ki a  $randomSík$ -hoz tartozó  $H$  homográfiát, DLT vagy HA algoritmussal.  
 $H := DLT(randomSík)$  PC esetén,  
 $H := HA(randomSík)$  AC esetén.
  4. A kapott  $H$  homográfia segítségével transzformáljuk át az első kép pontjait a másik képre:  
 $x'_H = Hx$  ahol  $x \in P_1$ .
  5. Vegyük a  $randomSík$ -kal egy síkba tartozóknak azokat a pontokat melyek transzformált pontjainak euklideszi távolsága a hozzájuk tartozó második képen lévő pontoktól kisebb mint a  $threshold$ .
  6. Legyen  $randomSíkLefedettsége$  a  $randomSík$  elemeinek száma osztva az összes megfeleltetés számával.
  7.  $iterációszám := iterációszám + 1$ .  
Ha a  $randomSíkLefedettsége > legjobbLefedettség$  akkor:  
 $legjobbLefedettség := randomSíkLefedettsége$ ,  
 $legnagyobbSík := randomSík$
  8. 2-7 ismétlése addig míg  $iterációszám < maxIter$  és  $legjobbLefedettség < confidence$  feltételek igazak.
-

A 1. algoritmussal a megfelelő bemeneti paraméterek használata esetén megtalálható egy domináns sík a képen. Ezt felhasználva a megtalált síkokkal iteratívan csökkentve a bemeneti megfeleltetések halmazát detektálható a legtöbb pontot tartalmazó  $n$  darab sík:

---

**Algoritmus 2:** Síkok detektálása affin pontmegfeleltetések alapján (Szekvenciális RANSAC<sup>[27]</sup> algoritmus affin transzformációkkal)

---

**Bemenet:**

$P_1, P_2$  - megfelelő pontok az első és második kamerán,  
 $A$  - a pontmegfeleltetésekhez tartozó lokális környezetet leíró affin transzformációk,  
 $n$  - a keresett síkok száma,  
maxIter - RANSAC algoritmus maximális iteráció száma,  
reProjThreshold - maximális reprojekciós hiba,  
confidence - konfidencia százalék.

**Kimenet:**

$S$  - lista mely síkok pontjait tartalmazó listákat tartalmaz.

**Algoritmus:**

1.  $P_1, P_2, A$  normalizálása a numerikus stabilitás érdekében.  
 $I :=$  üres
  2. A legmeghatározóbb sík meghatározása RANSAC algoritmussal.  
 $P_1^s, P_2^s, A^s :=$  LPDRANSAC(maxIter, reProjThreshold, confidence)  
 $iter := 0$ .
  3. A megtalált sík pontjainak kivétele a bemeneti adatokól.  
 $P_1 := P_1 - P_a^s 1$   
 $P_2 := P_2 - P_2^s$   
 $A := A - A^s$   
 $iter := iter + 1$ .  
A megtalált sík pontjainak hozzáadása  $S$ -hez:  
 $S := [S, (P_1, P_2)]$ .
  4. 2.-4. futtatása míg az  $iter$  el nem éri  $n$ -et vagy az összes pontot síkhoz nem rendeltük.
-

## 3.2. Normalizált HAF algoritmus használata síkok detektálására

Ebben az alfejezetben bemutatok egy a *normalizált HAF* algoritmusra épülő síkdetektálási módszert. Célunk az előző konfigurációhoz hasonló bemeneti adatok esetén síkok detektálása. Adott két kamera és az általuk készített két felvétel, melyek teljesítik az epipoláris kényszert. A képeken adottak egymásnak megfelelő  $\mathbf{P}_1, \mathbf{P}_2$  pontpárok és a hozzájuk tartozó  $\mathbf{A}$  affin transzformációk.

Eltérve az előző módszertől, ahol pusztán a homográfiát használtuk a síkok megkülönböztetésre, most kiszámítjuk a valós térbeli  $\mathbf{X}$  koordinátájukat, valamint a lokális környezetüket leíró érintő sík  $\mathbf{n}$  normális vektorát is.

A módszer két lépésből áll, először meghatározzuk a pontmegfeleltetésekhöz tartozó  $\mathbf{H}$  homográfia mátrixot, melyből megkapható a lokális érintősík normálvektora. Ezt követően a projekciós mátrixok segítségével visszavetítjük a pontmegfeleltetéseket a valós térbe, és kiszámítjuk a hozzájuk tartozó  $\mathbf{X}$  vektort.

### Lokális érintősík normálvektorának meghatározása HAF algoritmussal

A normalizált *HAF* algoritmus bemenete egy affin megfeleltetés. Kimeneteként megadja a megfeleltetésben szereplő pontok valós térbeli  $\mathbf{X}$  pontját tartalmazó lokális érintősíkhöz tartozó  $\mathbf{H}$  homográfia mátrixot. Az algoritmus használatához szükségünk van az képpárok közötti epipoláris kényszert kifejező  $\mathbf{F}$  fundamentális mátrixra.  $\mathbf{F}$  megkapható az elméleti részben bemutatott 8 pontos algoritmus segítségével. Továbbá normalizálnunk kell a bemeneti pont megfeleltetéseket és affin transzformációkat numerikus stabilitás érdekében.

Az algoritmust minden pontpárra lefuttatva megkapjuk a hozzájuk tartozó  $\mathbf{H}$  homográfia mátrixokat, melyek dekompozíciójából megkaphatók a pontpárokhoz tartozó  $\mathbf{n}$  normálvektorok. Továbbá eredményül kapjuk a kamerák koordinátarendszerei közötti  $\mathbf{R}$  mátrixot és  $\mathbf{t}$  eltolás vektort. Fontos megjegyezni, hogy a dekompozíció során  $\mathbf{R}$ -re  $\mathbf{t}$ -re és  $\mathbf{n}$ -re négy lehetséges megoldást kapunk. A megoldások közül kiválasztható a valós megoldás, ha megköveteljük a valós térbeli pont pozitív mélységét. A pozitív mélység megkövetelése azt jelenti, hogy a térbeli pont mindkét kamera előtt helyezkedik el.

### Valós térbeli koordináták meghatározása visszavetítéssel

Miután adottak a pontpárokhoz tartozó  $\mathbf{n}$  normál vektorok, a pontok síkokba történő klaszterezéséhez szükségünk van még a valós térbeli  $\mathbf{X}$  koordinátára. Ezek a koordináták megkaphatók a képpontok valós térbe történő visszavetítésével. A

visszavetítés megtehető a kamerák projekciós mátrixának segítségével, melyek:

$$\mathbf{Proj}_1 = [\mathbf{I}|\mathbf{0}], \quad \mathbf{Proj}_2 = [\mathbf{R}|\mathbf{t}], \quad (3.1)$$

ahol  $\mathbf{I}$  a 3x3-as egységmátrix,  $\mathbf{R}$  a kamerák közötti forgatás mátrix, valamint  $\mathbf{t}$  a kamerák közötti eltolás vektor. A forgatás mátrix, valamint az eltolásvektor megkaphatók az összes pont alapján kapott  $\mathbf{F}$  fundamentális mátrix felbontásából. Ezt követően a visszavetítést elvégezve megkapjuk a  $\mathbf{X}$  koordinátákat.

Miután adottak a valós térbeli koordináták és a hozzájuk tartozó normálvektorok, a pontokat síkokba tudjuk klaszterezni. Alkalmos klaszterezési eljárás lehet a *DBSCAN*<sup>[28]</sup> algoritmus módosított távolság metrikával. Legyen adott a távolság metrikánk a következő módon:

$$d = \begin{cases} \|\mathbf{X}_i - \mathbf{X}_j\|_2 & \|\mathbf{n}_i - \mathbf{n}_j\|_2 < threshold \\ \infty & \text{különben} \end{cases} \quad (3.2)$$

ahol  $\mathbf{X}_i, \mathbf{X}_j \in \mathcal{X}, i \neq j$ ,  $\mathcal{X}$  a visszavetítésből kapott valós térbeli pontok halmaza,  $\mathbf{n}_i, \mathbf{n}_j$  a hozzájuk tartozó normál vektorok, a *threshold* pedig egy általunk választott küszöbérték. A választott küszöbérték adja meg, hogy milyen eltérésű normál vektorokat tekintünk még egy síkhoz tartozóknak. Továbbá a  $\|\cdot\|_2$  jelöli a vektorok *Euklideszi* távolságát.

Ezek alapján a klaszterezési eljárást futtatva megfelelő bemeneti paraméterek esetén megkaphatók a képpárokban detektált síkok.

## 4. fejezet

# Homográfia becslés affin megfeleltetésekkel

Az előző fejezetben módszert vezetünk be homográfia alapján történő sík detektálásra, így kulcsfontosságú kérdés, hogy vannak-e olyan speciális esetek, melyekben kevesebb affin pontmegfeleltetésből kiszámítható a homográfia mátrixa. Ebben a fejezetben megmutatom, hogy abban az esetben, ha van egy közös irány a két egymást követő kamera képen, és van előzetes információnk a kamera által látott síkok normális vektoráról, akkor kevesebb megfeleltetés is elegendő a sík  $\mathbf{H}$  homográfia mátrixának kiszámítására. A közös irány meghatározható a kamera mellé szerelt egyéb szenzor (giroszkóp) segítségével, mely gyakran adott mobil eszközök, műholdak, repülőgépek estén. Továbbá ha nem áll rendelkezésre külső szenzor, a közös irány horizont detektálás<sup>[29]</sup> vagy eltűnő pontok segítségével<sup>[30]</sup> is meghatározható. A következő alfejezetekben olyan rendszert tekintünk, ahol a kamera merőleges ( $z$ ) tengelyét párhuzamosra állítottuk a közös iránnyal. Ezek a rendszerek igen fontos szerepet játszhatnak akkor, ha repülő, drón vagy műhold felvételeken szeretnénk felülnézeti városi környezetben síkokat például úttestet vagy épületfalakat detektálni. Megmutatható, hogy függőleges és vízszintes sík esetén melynek ismert a normálvektora, két hagyományos pontmegfeleltetés ( $PC$ ), függőleges ismeretlen normál vektorú sík esetén 2.5  $PC$  elegendő  $\mathbf{H}$  megadásához<sup>[4]</sup>. Ezzel szemben a következőkben megmutatom, hogy affin megfeleltetéseket használva elegendő egy megfeleltetés ezen esetek során a homográfia mátrixának becsléséhez.

## 4.1. Homográfia becslése ismert normálvektorú vízszintes sík esetén

Általános esetben egy homografikus kapcsolat egy 3D sík két képsíkra vetített pontjai között a korábbiakban megismert módon írható fel:

$$\mathbf{x}' = \mathbf{H}\mathbf{x}, \quad (4.1)$$

ahol  $\mathbf{x} = [x, y, 1]^T$  és  $\mathbf{x}' = [x', y', 1]^T$ , valamint:

$$\mathbf{H} = \mathbf{R} - \frac{1}{d}\mathbf{t}\mathbf{n}^T. \quad (4.2)$$

Tekintsük azt az esetet, hogy a kamerák előtt elhelyezkedő vízszintes sík normálvektora ismert, és a kamerák  $z$  tengelye párhuzamos ezzel az ismert normálvektorral. (A kamerák szembe néznek egy síkkal, például épületek tetejével vagy aszfalttal.) Tegyük fel továbbá, hogy a kamera belső paraméterei ismertek és a megfeleltetéseink normalizáltak. Ekkor az  $\mathbf{R}$  forgatás mátrix  $\mathbf{R}_z$  komponense marad ismeretlen, amely a  $z$  tengely körüli  $\theta$  szögű elforgatásnak felel meg:

$$\mathbf{H} = \mathbf{R}_z - \frac{1}{d}\mathbf{t}\mathbf{n}^T. \quad (4.3)$$

Mivel a 3D-s sík párhuzamos a képsíkkal és  $z$  tengelye párhuzamos  $\mathbf{n}$ -el ezért  $\mathbf{n} = [0 \ 0 \ 1]^T$  így:

$$\mathbf{H} = \begin{bmatrix} \cos(\Theta) & -\sin(\Theta) & 0 \\ \sin(\Theta) & \cos(\Theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{\mathbf{t}}{d} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}^T. \quad (4.4)$$

A  $d$  távolság ismeretlen, azonban a projektív transzformáció miatt  $\mathbf{t}$  csak egy skála erejéig határozható meg így a kamera-sík távolságot választhatjuk 1-nek és beleolvaszthatjuk  $\mathbf{t}$ -be:

$$\mathbf{H} = \begin{bmatrix} \cos(\Theta) & -\sin(\Theta) & 0 \\ \sin(\Theta) & \cos(\Theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}^T = \begin{bmatrix} \cos(\Theta) & -\sin(\Theta) & -t_x \\ \sin(\Theta) & \cos(\Theta) & -t_y \\ 0 & 0 & 1 - t_z \end{bmatrix}. \quad (4.5)$$

Mely azt jelenti, hogy a homográfiaiban 5 ismeretlenünk marad, és a következő alakban keressük:

$$\mathbf{H} = \begin{bmatrix} h_1 & -h_2 & h_3 \\ h_2 & h_1 & h_4 \\ 0 & 0 & h_5 \end{bmatrix}. \quad (4.6)$$

A 4.1 egyenlet felírható lineáris egyenletrendszer alakjában, ha vektoriálisan szorozzuk  $\mathbf{x}'$ -vel:

$$\mathbf{x}' \times \mathbf{H}\mathbf{x} = \mathbf{0}. \quad (4.7)$$

Azaz:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \times \begin{bmatrix} h_1 & -h_2 & h_3 \\ h_2 & h_1 & h_4 \\ 0 & 0 & h_5 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{0}. \quad (4.8)$$

A homográfiával való szorzást elvégezve:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \times \begin{bmatrix} h_1x - h_2y + h_3 \\ h_2x + h_1y + h_4 \\ h_5 \end{bmatrix} = \mathbf{0}. \quad (4.9)$$

Az  $\mathbf{x}'$ -vel való vektoriális szorzást elvégezve:

$$\begin{bmatrix} y'h_5 - h_2x - h_1y - h_4 \\ h_1x - h_2y + h_3 - x'h_5 \\ x'h_2x + x'h_1y + x'h_4 - y'h_1x + y'h_2y - y'h_3 \end{bmatrix} = \mathbf{0}. \quad (4.10)$$

Melyek közül csak két egyenlet lineárisan független:

$$\begin{bmatrix} -yh_1 - xh_2 - h_4 + y'h_5 \\ xh_1 - yh_2 + h_3 - x'h_5 \end{bmatrix} = \mathbf{0}. \quad (4.11)$$

Ezt a következő alakban írhatjuk:

$$\begin{bmatrix} -y & -x & 0 & -1 & y' \\ x & -y & 1 & 0 & -x' \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \end{bmatrix} = \mathbf{0}. \quad (4.12)$$

Így egy pontmegfeleltetésből két megszorítást kaptunk a  $\mathbf{H}$ -ra, azonban rendelkezésünkre áll még a pontok lokális környezetének változását leíró  $\mathbf{A}$  affin transzformáció. A 4.6 homográfiát a 2.36 egyenletbe helyettesítve újabb négy egyenletet kapunk:

$$\begin{aligned} h_1 - a_{11}h_5 &= 0, \\ -h_2 - a_{12}h_5 &= 0, \\ h_2 - a_{21}h_5 &= 0, \\ h_1 - a_{22}h_5 &= 0. \end{aligned} \quad (4.13)$$

Mátrixos reprezentációban írva:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -a_{11} \\ 0 & -1 & 0 & 0 & -a_{12} \\ 0 & 1 & 0 & 0 & -a_{21} \\ 1 & 0 & 0 & 0 & -a_{22} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \end{bmatrix} = \mathbf{0}. \quad (4.14)$$

Így összesen 6 megszorítást tettünk az 5 ismeretlen paraméterrel rendelkező  $\mathbf{H}$ -ra, melyeket összefoglalva a következő alakban írhatjuk fel:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -a_{11} \\ 0 & -1 & 0 & 0 & -a_{12} \\ 0 & 1 & 0 & 0 & -a_{21} \\ 1 & 0 & 0 & 0 & -a_{22} \\ -y & -x & 0 & -1 & y' \\ x & -y & 1 & 0 & -x' \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \end{bmatrix} = \mathbf{0}. \quad (4.15)$$

Amely azt jelenti, hogy csupán egy affin megfeleltetésből  $\mathbf{A}\mathbf{h} = \mathbf{0}$  -alakú egyenletrendszer alakjában megkapható a megoldás  $\mathbf{H}$ -ra, ami lehetőséget ad, hogy nagyban javítsuk a robusztus módszerek hatékonyságát.

A megoldás  $\mathbf{h}$ -ra az  $\mathbf{A}^T\mathbf{A}$  mátrix legkisebb sajátértékéhez tartozó saját vektoraként kapható. A szabad skálázási tényezőt rögzítjük a  $\|\mathbf{h}\| = 1$  feltétellel. Ezt követően, hogy  $\mathbf{R}_z$  komponenseire valós forgatás paramétereit kapjunk,  $h_1$  és  $h_2$ -re trigonometrikus megkötést kell tennünk:

$$h_1^2 + h_2^2 = 1. \quad (4.16)$$

A 4.16 megkötés a  $\mathbf{h}$  megoldásvektor  $\pm\sqrt{h_1^2 + h_2^2}$ -vel való osztását vonja maga után. A megoldásvektorból kifejezhető  $\mathbf{R}$  és  $\mathbf{t}$  is:

$$\mathbf{t} = \begin{bmatrix} -h_3 \\ -h_4 \\ 1 - h_5 \end{bmatrix}, \quad (4.17)$$

$$\mathbf{R} = \begin{bmatrix} h_1 & -h_2 & 0 \\ h_2 & h_1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.18)$$

Fontos megjegyezni, hogy a trigonometrikus megkötés miatt két megoldást kapunk  $\mathbf{H}$ -ra,  $\mathbf{R}$ -re és  $\mathbf{t}$ -re. Jelen esetben nem tettünk megkötést  $\mathbf{t}$  vektorra, azonban lehetőség van a kettős megoldás megszüntetésére abban az esetben, ha a 4.5 egyenletben

$1 - t_z$  nem nulla<sup>[4]</sup>.

## 4.2. Homográfia becslése ismert normálvektorú függőleges sík esetén

A következő speciális eset a valós életbeli épületfalak detektálásának mintájára irányul. Legyen adott egy, a kamerával szemben lévő merőleges sík ismert  $\mathbf{n}$  normálvektorral. (Azaz felül nézetből tekintünk egy függőleges síkra, melynek normál vektora merőleges a kamera  $z$  tengelyére.) Ekkor a  $\mathbf{H}$  homográfia mátrix a függőleges síkra a

$$\mathbf{H} = \mathbf{R}_z - [t_x, t_y, t_z]^T [n_x, n_y, 0] \quad (4.19)$$

alakban írható. Az egyenletet kibontva a következőt kapjuk:

$$\mathbf{H} = \begin{bmatrix} \cos(\Theta) - n_x t_x & -\sin(\Theta) - n_y t_x & 0 \\ \sin(\Theta) - n_x t_y & \cos(\Theta) - n_y t_y & 0 \\ -n_x t_z & -n_y t_z & 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & 0 \\ h_3 & h_4 & 0 \\ h_5 & \frac{n_y}{n_x} h_5 & 1 \end{bmatrix}. \quad (4.20)$$

A 4.20-ben kapott  $\mathbf{H}$  homográfiát a 4.7 egyenletbe helyettesítve:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \times \begin{bmatrix} h_1 & h_2 & 0 \\ h_3 & h_4 & 0 \\ h_5 & \frac{n_y}{n_x} h_5 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{0}. \quad (4.21)$$

A homográfiával való szorzást elvégezve:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \times \begin{bmatrix} h_1 x + h_2 y \\ h_3 x + h_4 y \\ h_5 x + \frac{n_y}{n_x} h_5 y + 1 \end{bmatrix} = \mathbf{0}. \quad (4.22)$$

Végül a vektoriális szorzást elvégezve:

$$\begin{bmatrix} y' h_5 x + y' \frac{n_y}{n_x} h_5 y + y' - h_3 x - h_4 y \\ h_1 x + h_2 y - x' h_5 x - x' \frac{n_y}{n_x} h_5 y - x' \\ x' h_3 x + x' h_4 y - y' h_1 x - y' h_2 y \end{bmatrix} = \mathbf{0}. \quad (4.23)$$

Mely egyenletek közül kiválasztva két lineárisan független egyenletet és mátrixos alakba írva kapjuk:

$$\begin{bmatrix} 0 & 0 & -x & -y & y'x + y'y\frac{n_y}{n_x} \\ x & y & 0 & 0 & -x'x - x'y\frac{n_y}{n_x} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \end{bmatrix} = \begin{bmatrix} -y' \\ x' \end{bmatrix}. \quad (4.24)$$

Amely két megszorítást ad a  $\mathbf{H}$  homográfiára. Az előző alfejezethez hasonlóan a 4.20-et a 2.36 egyenletbe helyettesítve újabb feltételeket kapunk:

$$\begin{aligned} h_1 - (x' + a_{11}x)h_5 - a_{11}y\frac{n_y}{n_x}h_5 - a_{11} &= 0, \\ h_2 - (x' + a_{12}y)\frac{n_y}{n_x}h_5 - a_{12}xh_5 - a_{12} &= 0, \\ h_3 - (y' + a_{21}x)h_5 - a_{21}y\frac{n_y}{n_x}h_5 - a_{21} &= 0, \\ h_4 - (y' + a_{22}y)\frac{n_y}{n_x}h_5 - a_{22}xh_5 - a_{22} &= 0. \end{aligned} \quad (4.25)$$

Az eredményt mátrixos alakban írva:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -(x' + a_{11}x) - a_{11}y\frac{n_y}{n_x} \\ 0 & 1 & 0 & 0 & -(x' + a_{12}y)\frac{n_y}{n_x} - a_{12}x \\ 0 & 0 & 1 & 0 & -(y' + a_{21}x) - a_{21}y\frac{n_y}{n_x} \\ 0 & 0 & 0 & 1 & -(y' + a_{22}y)\frac{n_y}{n_x} - a_{22}x \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{bmatrix}. \quad (4.26)$$

Majd összevonva a 4.24 egyenlettel:

$$\begin{bmatrix} 0 & 0 & -x & -y & y'x + y'y\frac{n_y}{n_x} \\ x & y & 0 & 0 & -x'x - x'y\frac{n_y}{n_x} \\ 1 & 0 & 0 & 0 & -(x' + a_{11}x) - a_{11}y\frac{n_y}{n_x} \\ 0 & 1 & 0 & 0 & -(x' + a_{12}y)\frac{n_y}{n_x} - a_{12}x \\ 0 & 0 & 1 & 0 & -(y' + a_{21}x) - a_{21}y\frac{n_y}{n_x} \\ 0 & 0 & 0 & 1 & -(y' + a_{22}y)\frac{n_y}{n_x} - a_{22}x \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \end{bmatrix} = \begin{bmatrix} -y' \\ x' \\ a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{bmatrix}. \quad (4.27)$$

Mely azt jelenti, hogy 6 feltételt adtunk az 5 szabad paraméterrel rendelkező  $\mathbf{H}$  mátrixra, ami így meghatározható. A rendszerre két megkötés szabható, a trigonometrikus megkötés, valamint a normálvektor-megkötés:

$$\cos^2(\Theta) + \sin^2(\Theta) = 1; \quad n_x^2 + n_y^2 = 1. \quad (4.28)$$

Ezeket felhasználva a kapott  $\mathbf{h}$  megoldásból kiszámíthatók a kamerák közti transzformációk paraméterei:

$$\begin{aligned} t_z &= -\frac{h_5}{n_x}, \\ t_x &= n_x(h_4 - h_1) - n_y(h_2 + h_3), \\ t_y &= n_y(h_1 - h_4) - n_x(h_2 + h_3). \end{aligned} \quad (4.29)$$

Melyeket felhasználva megadható a forgatás paramétere is:

$$\cos(\Theta) = h_1 + n_x t_x. \quad (4.30)$$

Így megoldást tudunk nyújtani csupán egy  $AC$  felhasználásával ismert normálvektorú függőleges síkok homográfiájának becslésére.

### 4.3. Homográfia becslése ismertlen normálvektorú függőleges sík esetén

Az előző alfejezetben tárgyalt elrendezésű kamera rendszerben, ismeretlen  $\mathbf{n}$  normálvektor esetén a  $\mathbf{H}$  homográfia a következő alakban írható fel:

$$\mathbf{H} = \begin{bmatrix} \cos(\Theta) - n_x t_x & -\sin(\Theta) - n_y t_x & 0 \\ \sin(\Theta) - n_x t_y & \cos(\Theta) - n_y t_y & 0 \\ -n_x t_z & -n_y t_z & 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & 0 \\ h_3 & h_4 & 0 \\ h_5 & h_6 & 1 \end{bmatrix}. \quad (4.31)$$

Melyet a 4.7 egyenletbe helyettesítve és elvégezve a homográfiával való szorzást:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \times \begin{bmatrix} h_1 x + h_2 y \\ h_3 x + h_4 y \\ h_5 x + h_6 y + 1 \end{bmatrix} = \mathbf{0}. \quad (4.32)$$

Elvégezve a vektoriális szorzást:

$$\begin{bmatrix} y' h_5 x + y' h_6 y + y' - h_3 x - h_4 y \\ h_1 x + h_2 y - x' h_5 x - x' h_6 y - x' \\ x' h_3 x + x' h_4 y - y' h_1 x - y' h_2 y \end{bmatrix} = \mathbf{0}. \quad (4.33)$$

Kiválasztva két lineárisan független egyenletet és mátrixos alakban írva:

$$\begin{bmatrix} 0 & 0 & -x & -y & y'x & y'y \\ x & y & 0 & 0 & -x'x & -x'y \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \end{bmatrix} = \begin{bmatrix} -y' \\ x' \end{bmatrix}. \quad (4.34)$$

Továbbá 4.31-ot a 2.36 egyenletbe helyettesítve újabb 4 feltételt kapunk:

$$\begin{aligned} h_1 - (x' + a_{11}x)h_5 - a_{11}yh_6 - a_{11} &= 0, \\ h_2 - (x' + a_{12}y)h_6 - a_{12}xh_5 - a_{12} &= 0, \\ h_3 - (y' + a_{21}x)h_5 - a_{21}yh_6 - a_{21} &= 0, \\ h_4 - (y' + a_{22}y)h_6 - a_{22}xh_5 - a_{22} &= 0. \end{aligned} \quad (4.35)$$

Melyet mátrixos alakban írva és egybevonva 4.34 egyenlettel:

$$\begin{bmatrix} 0 & 0 & -x & -y & y'x & y'y \\ x & y & 0 & 0 & -x'x & -x'y \\ 1 & 0 & 0 & 0 & -(x' + a_{11}x) & -a_{11}y \\ 0 & 1 & 0 & 0 & -a_{12}x & -(x' + a_{12}y) \\ 0 & 0 & 1 & 0 & -(y' + a_{21}x) & -a_{21}y \\ 0 & 0 & 0 & 1 & -a_{22}x & -(y' + a_{22}y) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \end{bmatrix} = \begin{bmatrix} -y' \\ x' \\ a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{bmatrix}. \quad (4.36)$$

Így hat egyenletet kaptunk egy hat szabadsági fokú  $\mathbf{H}$  mátrix becslésére. Az egyenletrendszer megoldása után a kamerák közti transzformációk, valamint a sík normális vektora is megadható. A  $\mathbf{n}$  normális vektorra a következő feltétel szabható:

$$n_x^2 + n_y^2 = 1. \quad (4.37)$$

A 4.31 és a 4.37 egyenletet figyelembe véve  $t_z$ -re a következő két megoldás adható:

$$t_z = \pm \sqrt{h_5^2 + h_6^2}. \quad (4.38)$$

$t_z$ -t felhasználva  $n_x$  és  $n_y$  megadható:

$$n_x = -\frac{h_5}{t_z}, \quad (4.39)$$

$$n_y = -\frac{h_6}{t_z}. \quad (4.40)$$

A 4.39 és 4.40 egyenleteket felhasználva  $t_x$  és  $t_y$  is megadható:

$$\begin{aligned} t_x &= n_x(h_4 - h_1) - n_y(h_2 + h_3), \\ t_y &= n_y(h_1 - h_4) - n_x(h_2 + h_3). \end{aligned} \tag{4.41}$$

Végül  $t_x$  felhasználásával a forgatás paramétere is kiszámítható:

$$\cos(\Theta) = h_1 + n_x t_x. \tag{4.42}$$

Összefoglalva tehát: Sikertült megadnunk egy affin megfeleltetés felhasználásával ismert normál vektorú függőleges és vízszintes síkok homográfiáját, valamint ismeretlen normálvektorú függőleges sík homográfiáját is. Mely konfigurációk valós életbeli adatok esetén felülnézetből készült műholdas vagy repülőgépi adatoknak felelnek meg. Összehasonlítva a hagyományos módszerrel, ez lehetőséget ad, hogy az alkalmazott robusztus módszereink kevésbé érzékenyek legyenek, és kevesebb iteráció számmal lefussanak. A homográfiák becslése mellett sikertült továbbá megadnunk a kamerák közötti transzformációk paramétereit is arra az esetre, ha a módszert nem síkok detektálására, hanem a kamerák mozgásának követésére szeretnénk használni.

## 5. fejezet

# Homográfia becslés síkmozgás esetén

Az előző fejezetben tárgyaltam a homográfia mátrixának becslését felülnézeti rendszerek esetén. Ebben a fejezetben az önvezető autózás esetén előforduló rendszerek analógiájára síkbeli mozgást leíró sztereó képpárokon történő síkdetektálással foglalkozom. Abból a feltételezésből indulok ki, hogy a képpárok közötti elmozdulás a gravitáció irányában (függőlegesen) nulla. Levezetem, hogy ilyen esetekben affin megfeleltetéseket használva elegendő egy megfeleltetés a homográfia becsléséhez. Ezt követően megvizsgálom az előző fejezetben is látott speciális eseteket: nevezetesen a kamerákkal szemben lévő vízszintes és függőleges síkok esetét.

Síkbeli mozgás esetén a kamerák közötti  $\mathbf{R}$  forgatás mátrix  $x$  és  $z$  tengely körüli forgást leíró komponense nulla így  $\mathbf{R}$  felírható  $\mathbf{R}_y$  alakban:

$$\mathbf{R} = \mathbf{R}_y = \begin{bmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix}, \quad (5.1)$$

ahol  $\alpha$  az  $y$  tengely körüli forgatás szöge. Továbbá sztereó képpárok esetén a kamerák közötti  $\mathbf{t}$  elmozdulás vektor csak egy skála erejéig határozható meg. Az előző alfejezetekhez hasonlóan válasszuk meg úgy a skálát, hogy a detektálni kívánt sík  $d$  távolságát a  $\mathbf{t}$  elmozdulásvektorba olvasztva egységvektort kapjunk. Ilyen esetben az elmozdulásvektor síkmozgás esetén a következő alakban írható:

$$\mathbf{t} = \begin{bmatrix} \cos(\beta) \\ 0 \\ \sin(\beta) \end{bmatrix}. \quad (5.2)$$

## 5.1. Általános orientációjú sík esete

Általános esetben, ha nem teszünk semmilyen feltételezést a képpárokban található síkokról a 4.2 egyenlet a következő alakban írható:

$$\mathbf{H} = \mathbf{R}_y - \frac{1}{d} \mathbf{t} \mathbf{n}^T. \quad (5.3)$$

Ahol  $\mathbf{R}_y$  és  $\mathbf{t}$  az előzőekben (5.1, 5.2) bevezetett forgatás mátrix és eltolás vektor. Így a  $\mathbf{H}$  homográfia mátrixra a következőt kapjuk:

$$\mathbf{H} = \begin{bmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} - \begin{bmatrix} \cos(\beta) \\ 0 \\ \sin(\beta) \end{bmatrix} \begin{bmatrix} n_x & n_y & n_z \end{bmatrix} \quad (5.4)$$

A műveletek elvégzése után:

$$\mathbf{H} = \begin{bmatrix} \cos(\alpha) - \cos(\beta)n_x & -\cos(\beta)n_y & -\sin(\alpha) - \cos(\beta)n_z \\ 0 & 1 & 0 \\ \sin(\alpha) - \sin(\beta)n_x & -\sin(\beta)n_y & \cos(\alpha) - \sin(\beta)n_z \end{bmatrix} \quad (5.5)$$

Mely alapján a homográfia mátrix a következő alakban kereshető:

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ 0 & 1 & 0 \\ h_4 & h_5 & h_6 \end{bmatrix}. \quad (5.6)$$

A kapott  $\mathbf{H}$  mátrixban hat ismeretlen komponens van, melyek megadhatók lesznek egy affin megfeleltetésből. Négy lineárisan független egyenletet kaphatunk az affin transzformációkból 2.36 alapján:

$$\begin{aligned} h_1 - h_4(x' + a_{11}x) - h_5a_{11}y - h_6a_{11} &= 0, \\ h_2 - h_5(x' + a_{12}y) - h_4a_{12}x - h_6a_{12} &= 0, \\ -h_4(y' + a_{21}x) - h_5a_{21}y - h_6a_{21} &= 0, \\ 1 - h_5(y' + a_{22}y) - h_4a_{22}x - h_6a_{22} &= 0. \end{aligned} \quad (5.7)$$

Mátrixos alakban írva:

$$\begin{bmatrix} 1 & 0 & 0 & -(x' + a_{11}x) & -a_{11}y & -a_{11} \\ 0 & 1 & 0 & -a_{12}x & -(x' + a_{12}y) & -a_{12} \\ 0 & 0 & 0 & -(y' + a_{21}x) & a_{21}y & -a_{21} \\ 0 & 0 & 0 & a_{22}x & (y' + a_{22}y) & a_{22} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \quad (5.8)$$

Továbbá két egyenletet kapunk a pontmegfeleltetésekből 2.37 alapján:

$$\begin{aligned} h_1x + h_2y + h_3 - h_4xx' - h_5yx' - h_6x' &= 0, \\ y - h_4xy' - h_5yy' - h_6y' &= 0. \end{aligned} \quad (5.9)$$

Mely szintén mátrixos alakban írva:

$$\begin{bmatrix} x & y & 1 & -xx' & -yx' & -x' \\ 0 & 0 & 0 & xy' & yy' & y' \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (5.10)$$

A 5.8 és a 5.10 egyenletrendszereket együtt  $\mathbf{Ch} = \mathbf{d}$  alakban felírva:

$$\begin{bmatrix} 1 & 0 & 0 & -(x' + a_{11}x) & -a_{11}y & -a_{11} \\ 0 & 1 & 0 & -a_{12}x & -(x' + a_{12}y) & -a_{12} \\ 0 & 0 & 0 & -(y' + a_{21}x) & a_{21}y & -a_{21} \\ 0 & 0 & 0 & a_{22}x & (y' + a_{22}y) & a_{22} \\ x & y & 1 & -xx' & -yx' & -x' \\ 0 & 0 & 0 & xy' & yy' & y \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ y \end{bmatrix}. \quad (5.11)$$

Így  $\mathbf{h}$ -ra, valamint  $\mathbf{h}$  vektoron keresztül  $\mathbf{H}$ -ra megadható a megoldás:

$$\mathbf{h} = \mathbf{C}^{-1}\mathbf{d}, \quad (5.12)$$

Ahol  $\mathbf{C}^{-1}$  a  $\mathbf{C}$  mátrix inverze. Ezzel a módszerrel tehát síkmozgás esetén is megadható a homográfia mátrixa csupán egy affin megfeleltetés felhasználásával. A kapott  $\mathbf{h}$  vektorból 5.5 egyenletet 5.6-al összevetve  $\beta$  megadható:

$$\beta = \operatorname{atan2}\left(\frac{h_5}{h_2}\right). \quad (5.13)$$

$\beta$ -t felhasználva  $n_y$  megadható lenne:

$$n_y = -\frac{h_2}{\cos(\beta)}, \quad (5.14)$$

vagy:

$$n_y = -\frac{h_5}{\sin(\beta)}, \quad (5.15)$$

alakban, azonban a  $\beta$  lehet 0 vagy 90 fok is ezért érdemesebb  $\beta$  megadására kihasználni, hogy  $\cos^2(\beta) + \sin^2(\beta) = 1$ , így:

$$n_y = \pm \sqrt{h_2^2 + h_5^2}. \quad (5.16)$$

Az  $n_y$  komponensre két megoldást kapunk. A továbbiakban tekintsük úgy, hogy a képletekben mindkét megoldással számoltunk, azaz a két megoldás halmazunk lesz. A  $\beta$  szögparaméter ismeretében  $n_x$ -re és  $n_z$ -re két egyenlet írható fel:

$$\begin{aligned} -\cos(\beta)n_z - \sin(\beta)n_x &= h_3 + h_4, \\ -\cos(\beta)n_x + \sin(\beta)n_z &= h_1 - h_6. \end{aligned} \quad (5.17)$$

Ami  $\beta$  ismeretében egy kétismeretlenes inhomogén lineáris egyenletrendszer  $n_x$ -re és  $n_z$ -re nézve, így megoldható. A kapott  $n_x$  és  $n_z$  komponenseket felhasználva  $\alpha$  is számítható:

$$\alpha = \operatorname{atan2}\left(\frac{h_4 + \sin(\beta)n_x}{h_1 + \cos(\beta)n_x}\right), \quad (5.18)$$

alakban.

Összegezve ebben az alfejezetben megmutattam, hogyan lehet megadni síkbeli mozgás esetén a  $\mathbf{H}$  homográfia mátrix komponenseit affin megfeleltetések alapján. Továbbá megmutattam, hogyan kaphatók meg a síkbeli mozgást leíró szögparaméterek a kapott  $\mathbf{H}$  mátrix komponensei segítségével. A következő fejezetekben megvizsgálom, hogyan egyszerűsödik a homográfia megadása abban az esetben, ha előzetes információnk van a sík normálisvektorával kapcsolatban.

## 5.2. Vízszintes sík esete

Ebben az alfejezetben azzal az esettel foglalkozom, mikor a sztereó képpárokban egy ismert a kamera  $z$  tengelyével párhuzamos sík található. Ebben az esetben a sík  $\mathbf{n}$  normálvektora:

$$\mathbf{n} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \quad (5.19)$$

A mozgás során a kamerák koordináta rendszerei közötti kapcsolatot a 5.1 egyenletben bevezetett  $\mathbf{R}_y$  forgatás mátrix, és a 5.2-ben bevezetett  $\mathbf{t}$  eltolásvektor adja meg. A  $\mathbf{R}_y$  forgatás mátrixot, valamint  $\mathbf{t}$  és  $\mathbf{n}$  vektorokat a 4.2 egyenletbe helyettesítve a következőt kapjuk a  $\mathbf{H}$  homográfia mátrixra:

$$\mathbf{H} = \begin{bmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} - \begin{bmatrix} \cos(\beta) \\ 0 \\ \sin(\beta) \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \quad (5.20)$$

A műveletek elvégzése után:

$$\mathbf{H} = \begin{bmatrix} \cos(\alpha) & -\cos(\beta) & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & -\sin(\beta) & \cos(\alpha) \end{bmatrix} \quad (5.21)$$

Ez alapján a  $\mathbf{H}$  homográfia mátrix a következő alakban kereshető:

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & -h_3 \\ 0 & 1 & 0 \\ h_3 & h_4 & h_1 \end{bmatrix}. \quad (5.22)$$

A kapott  $\mathbf{H}$  homográfia mátrixban tehát csupán négy ismeretlen paraméter található így megadható lesz egy affin megfeleltetéssel. Az előző alfejezethez hasonlóan négy lineárisan független egyenletet kapható az affin transzformációkból 2.36 alapján:

$$\begin{aligned} h_1 - h_3(x' + a_{11}x) - h_4a_{11}y - h_1a_{11} &= 0, \\ h_2 - h_4(x' + a_{12}y) - h_3a_{12}x - h_1a_{12} &= 0, \\ -h_3(y' + a_{21}x) - h_4a_{21}y - h_1a_{21} &= 0, \\ 1 - h_4(y' + a_{22}y) - h_3a_{22}x - h_1a_{22} &= 0. \end{aligned} \quad (5.23)$$

Melyet mátrixos alakban írva:

$$\begin{bmatrix} 1 - a_{11} & 0 & -(x' + a_{11}x) & -a_{11}y \\ -a_{12} & 1 & -a_{12}x & -(x' + a_{12}y) \\ -a_{21} & 0 & -(y' + a_{21}x) & a_{21}y \\ a_{22} & 0 & a_{22}x & (y' + a_{22}y) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \quad (5.24)$$

Továbbá két egyenletet kapunk a pontmegfeleltetésekből 2.37 alapján:

$$\begin{aligned} h_1x + h_2y - h_3 - h_3xx' - h_4yx' - h_1x' &= 0, \\ y - h_3xy' - h_4yy' - h_1y' &= 0. \end{aligned} \quad (5.25)$$

Melyet szintén mátrixos alakba írva:

$$\begin{bmatrix} x - x' & y & -1 - xx' & -yx' \\ y' & 0 & xy' & yy' \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (5.26)$$

A 5.24 és a 5.26 egyenletrendszereket együtt  $\mathbf{Ch} = \mathbf{d}$  alakban felírva:

$$\begin{bmatrix} 1 - a_{11} & 0 & -(x' + a_{11}x) & -a_{11}y \\ -a_{12} & 1 & -a_{12}x & -(x' + a_{12}y) \\ -a_{21} & 0 & -(y' + a_{21}x) & a_{21}y \\ a_{22} & 0 & a_{22}x & (y' + a_{22}y) \\ x - x' & y & -1 - xx' & -yx' \\ y' & 0 & xy' & yy' \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ y \end{bmatrix}. \quad (5.27)$$

Fontos megjegyezni, hogy bár 4 ismeretlenünk van érdemes megtartani mind a hat egyenletet a kapott egyenletrendszerből, mivel így csökkenthető a zaj hatása.

A 5.27 egyenlet  $\mathbf{h}$  megoldásvektora megadható az előző alfejezethez hasonló módszerrel:

$$\mathbf{h} = \mathbf{C}^\dagger \mathbf{d}, \quad (5.28)$$

Ahol  $\mathbf{C}^\dagger$  a  $\mathbf{C}$  mátrix Moore-Penrose pszeudo inverze.

A bemutatott módszerrel megadható tehát egy affin megfeleltetés segítségével a vízszintes ( $\mathbf{n} = [0 \ 1 \ 0]^T$  normálvektorú) sík homográfia mátrixa. Ennek nagy a jelentősége, mivel a megoldási módszerünk tartalmazza a normálvektorra vonatkozó feltételt. Így az ilyen irányú normálvektorokra érzékenyebb módszert adtunk, amely a gyakorlatban azt jelenti, hogy jobban elkülöníthetők a vízszintes síkok ezzel a módszerrel számítva mint az előző alfejezetben bemutatott általános módszert használva.

A kapott  $\mathbf{h}$  vektorból 5.21 egyenletet 5.22-al összevetve  $\alpha$  és  $\beta$  megadható:

$$\beta = \text{atan2}\left(\frac{h_4}{h_2}\right), \quad (5.29)$$

valamint:

$$\alpha = \text{atan2}\left(\frac{h_3}{h_1}\right). \quad (5.30)$$

Összegezve tehát módszert adtam síkmozgás esetén vízszintes síkok detektálására egy affin megfeleltetés használatával és a kapott  $\mathbf{H}$  homográfia mátrix elemeiből megadtam a síkmozgást jellemző szögparamétereket. A következő alfejezetben megvizsgálom a kamerák előtt elhelyezkedő függőleges sík esetét.

### 5.3. Függőleges sík esete

Ebben az alfejezetben megvizsgálom azt az esetet, mikor a sztereó képpárokon egy függőleges sík található. Ez az eset felel meg a valós életbeli önvezető autózás során felmerülő épület falak detektálásának. A függőleges orientáció a következő megkötést vezet be a sík  $\mathbf{n}$  normálvektorára:

$$\mathbf{n} = \begin{bmatrix} n_x \\ 0 \\ n_z \end{bmatrix}. \quad (5.31)$$

A síkmozgás során bevezetett  $R_y$  forgatás mátrixot és  $t$  eltolás vektort (5.1, 5.2), valamint az  $\mathbf{n}$  normál vektort 5.31-ből a 4.2 egyenletbe helyettesítve, a következőt kapjuk a  $\mathbf{H}$  homográfia mátrixra:

$$\mathbf{H} = \begin{bmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} - \begin{bmatrix} \cos(\beta) \\ 0 \\ \sin(\beta) \end{bmatrix} \begin{bmatrix} n_x & 0 & n_z \end{bmatrix}. \quad (5.32)$$

A  $\mathbf{t}$  és  $\mathbf{n}$  vektorok diadikus szorzását elvégezve:

$$\mathbf{H} = \begin{bmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} - \begin{bmatrix} \cos(\beta)n_x & 0 & \cos(\beta)n_z \\ 0 & 0 & 0 \\ \sin(\beta)n_x & 0 & \sin(\beta)n_z \end{bmatrix}. \quad (5.33)$$

A mátrixok kivonását elvégezve  $\mathbf{H}$ -ra kapjuk:

$$\mathbf{H} = \begin{bmatrix} \cos(\alpha) - \cos(\beta)n_x & 0 & -\sin(\alpha) - \cos(\beta)n_z \\ 0 & 1 & 0 \\ \sin(\alpha) - \sin(\beta)n_x & 0 & \cos(\alpha) - \sin(\beta)n_z \end{bmatrix} \quad (5.34)$$

Ez alapján a  $\mathbf{H}$  homográfia mátrix a következő alakban kereshető:

$$\mathbf{H} = \begin{bmatrix} h_1 & 0 & h_2 \\ 0 & 1 & 0 \\ h_3 & 0 & h_4 \end{bmatrix}. \quad (5.35)$$

A kapott  $\mathbf{H}$  mátrixban függőleges sík esetén négy ismeretlen komponens van, melyek a korábbi alfejezetekhez hasonlóan megadhatók egy affin megfeleltetésből. Az előzőekben bemutatottakkal analóg módon négy lineárisan független egyenletet kapható

az affin transzformációkból 2.36 alapján:

$$\begin{aligned}
 h_1 - h_3(x' + a_{11}x) - h_4a_{11} &= 0, \\
 -h_3a_{12}x - h_4a_{12} &= 0, \\
 -h_3(y' + a_{21}x) - h_4a_{21} &= 0, \\
 1 - h_3a_{22}x - h_4a_{22} &= 0.
 \end{aligned} \tag{5.36}$$

Mátrixos alakban írva:

$$\begin{bmatrix} 1 & 0 & -(x' + a_{11}x) & -a_{11} \\ 0 & 0 & a_{12}x & a_{12} \\ 0 & 0 & (y' + a_{21}x) & a_{21} \\ 0 & 0 & a_{22}x & a_{22} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{5.37}$$

Továbbá két egyenletet kapunk a pontmegfeleltetésekből 2.37 alapján:

$$\begin{aligned}
 h_1x + h_2 - h_3xx' - h_4x' &= 0, \\
 y - h_3xy' - h_4y' &= 0.
 \end{aligned} \tag{5.38}$$

Mely mátrixos alakban írva:

$$\begin{bmatrix} x & 1 & -xx' & -x' \\ 0 & 0 & xy' & y' \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{5.39}$$

A 5.37 és a 5.39 egyenletrendszereket együtt  $\mathbf{Ch} = \mathbf{d}$  alakban felírva:

$$\begin{bmatrix} 1 & 0 & -(x' + a_{11}x) & -a_{11} \\ 0 & 0 & a_{12}x & a_{12} \\ 0 & 0 & (y' + a_{21}x) & a_{21} \\ 0 & 0 & a_{22}x & a_{22} \\ x & 1 & -xx' & -x' \\ 0 & 0 & xy' & y' \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ y \end{bmatrix}. \tag{5.40}$$

Így a négy ismeretlent tartalmazó  $\mathbf{h}$  megoldás vektorra hat egyenletet kaptunk. A korábbiakhoz hasonlóan itt is érdemes megtartani az összes egyenletet a zaj hatásának csökkentése érdekében. Az egyenletrendszer megoldása:

$$\mathbf{h} = \mathbf{C}^\dagger \mathbf{d}. \tag{5.41}$$

A  $\mathbf{h}$  megoldásvektor felhasználásával függőleges sík esetén nem adhatók meg a síkmozgást leíró  $\alpha$  és  $\beta$  szögparaméterek, valamint a sík normál vektorának  $n_x$  és  $n_z$  komponensei az előző alfejezetekhez hasonló egyszerű alakban. A négy ismeretlen paraméterre négy egyenlet adható 5.34 és 5.35 összevetésével, azonban a kapott egyenletrendszer nem lineáris egyenletrendszer formájában adódik:

$$\begin{aligned} \cos(\alpha) - \cos(\beta)n_x &= h_1, \\ -\sin(\alpha) - \cos(\beta)n_z &= h_2, \\ \sin(\alpha) - \sin(\beta)n_x &= h_3, \\ \cos(\alpha) - \sin(\beta)n_z &= h_4. \end{aligned} \tag{5.42}$$

Ezért ebben az esetben egyszerűbb a keresett paraméterek meghatározása a  $\mathbf{H}$  homográfia mátrix dekompozíciója alapján. A felbontás elvégezhető<sup>[31]</sup> numerikusan az *SVD*-re támaszkodó *Faugeras*<sup>[32]</sup> vagy *Zhang*<sup>[17]</sup> dekompozíciók felhasználásával.

Összefoglalva, ebben a fejezetben síkbeli mozgás során sztereó képpárokon történő síkhomográfia kiszámításával foglalkoztam. Az általános esettől haladva két speciális eset felé. Megmutattam, hogy a tárgyalt esetek során megadható a síkhomográfia mátrixa egy affin megfeleltetésből. Az első speciális eset, a képpárok elhelyezkedő vízszintes síkok esetét tárgyalta. Ebben és az általános esetben megadtam a síkmozgás szögparamétereit a kapott homográfia mátrix segítségével. A vízszintes sík esete felel meg az önvezető autózás során a gépjármű előtt elhelyezkedő úttest és járda esetének. A második speciális eset az ismeretlen normálvektorú függőleges orientációjú sík esete volt. A függőleges orientációjú sík esete megfelel az önvezető autózás során előforduló gépjármű előtt elhelyezkedő épületfalak esetének. Ilyen rendszerek esetén is kiszámíthatók a síkmozgás paraméterei numerikus módszerek felhasználásával. A kapott eredmény fontossága, hogy a bemutatott módszerek segítségével azokat robusztus algoritmusokba illesztve hatékonyabb síkdetektor állítható elő. Az iteratív robusztus módszerek esetén (pl. *RANSAC*) különösen fontos, hogy hány megfeleltetésből tudjuk számítani a homográfia mátrixát, mivel az iterációszámuk erősen nő a pontmegfeleltetések számával<sup>[4]</sup>. Továbbá, mivel beépítettük a detektálni kívánt sík paraméterét ( $\mathbf{n}$  normálvektorát), a megadott algoritmusunk érzékenyebb lesz az ilyen normálvektorú síkokra, így elősegítve azok megkülönböztethetőségét a többi síktól.

## 6. fejezet

# Tesztelés

Ebben a fejezetben megvizsgálom, hogy a homográfiára támaszkodó síkdetektálás mennyire hatékony. Először szintetikus generált adatokon tesztelem, hogy metsző, illetve párhuzamos síkok ideális zaj mentes környezetben mennyire különböztethetők meg egymástól. Majd ezt követően valós városi felvételeken keresek a 2. algoritmus segítségével.

### 6.1. Szintetikus tesztek

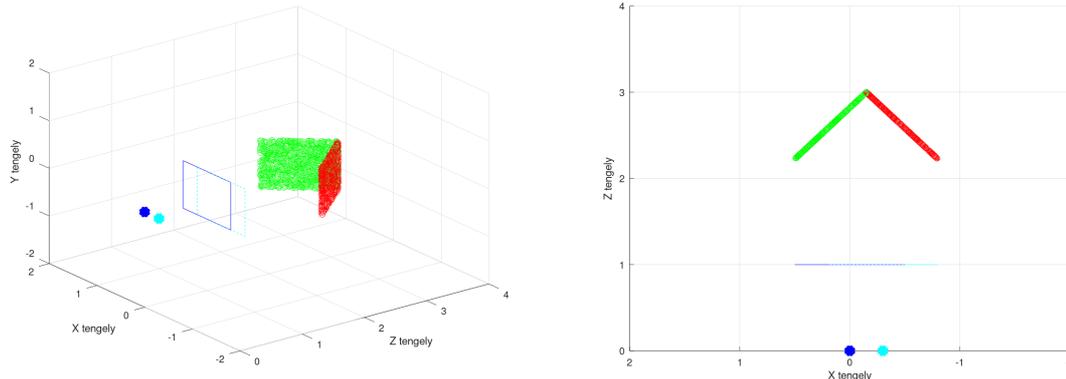
A szintetikus tesztek során standard sztereó képpárokat tekintettem. A távolságokat mértékegység nélküli egységekben mértem. Szimuláltam két kamerát, melyeket összekötő  $\mathbf{R}$  forgatás mátrix és  $\mathbf{t}$  eltolás vektor kapcsol össze:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}, \quad \mathbf{t} = 0.3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \quad (6.1)$$

Továbbá a kamerák  $f$  fókusztávolságát azonosnak és  $f = 1$ -nek vettem. A szimuláció során minden esetben a kamerák előtt elhelyezkedő két darab egységnyi területű síkrészlet pontjaiból számított homográfiákat vizsgáltam. Először metsző síkok esetére tértem ki, melyek a valós adatokban az épületfalakkal vonhatók kapcsolatba, majd az egymással párhuzamos síkok esetét vizsgáltam.

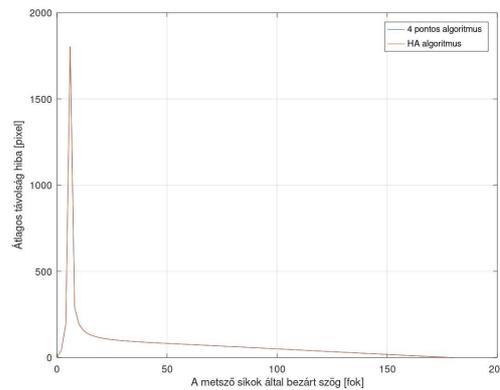
#### 6.1.1. Metsző síkok

Elsőként a 6.2 ábrán látható elrendezésben kamerák előtt elhelyezkedő metsző síkok esetét vizsgáltam. A metsző síkok esete kültéri vagy beltéri képek esetén párhuzamba vonható a kamerák előtt elhelyezkedő épület falainak találkozásával. A szimuláció során a síkok elhelyezkedése olyan, hogy az egyenes, melyben metszik egymást, a kamerák bázisegyenesének felezősíkján található. A szimulációt egybeeső síkok, azaz



(a) 3D nézet

(b) Felülnézet

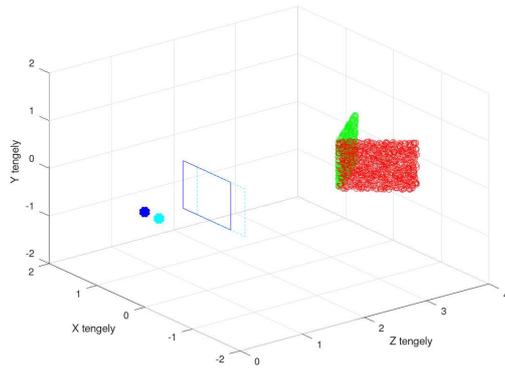


(c) Átlagos hiba

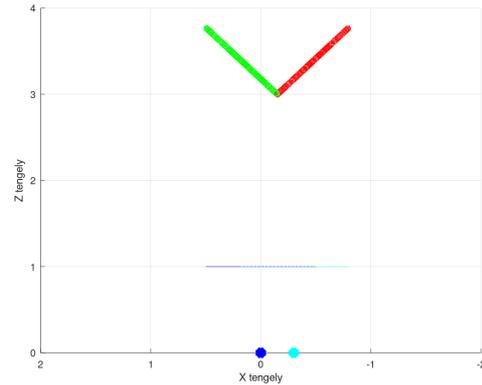
6.1. ábra. Kamerával szemben lévő hegyes szöget bezáró metsző síkok szimulációs elrendezése és a kapott eredmény.

nulla fokos  $\alpha$  szöget bezáró síkok helyzetéből kezdtem, majd lépésenként növeltem az általuk bezárt szöget, míg újra egybeesővé nem váltak ( $\alpha = 180^\circ$ ). Miden lépésben a *zöld* sík pontjai által becsültem  $\mathbf{H}_z$  homográfiát 4 pontos és *HA* algoritmussal 1000 megfeleltetésből. A becsült homográfiát felhasználva a *piros* sík első kamerán lévő  $\mathbf{x}_{\text{piros}}$  pontjait transzformáltam  $\mathbf{H}_z$ -val a második kamerára. Az így kapott  $\mathbf{x}_{\text{piros}}^{\mathbf{H}_z}$  pontok átlagos euklideszi távolságát mértem a valós, második kamerán található  $\mathbf{x}'_{\text{piros}}$  pontjuktól.

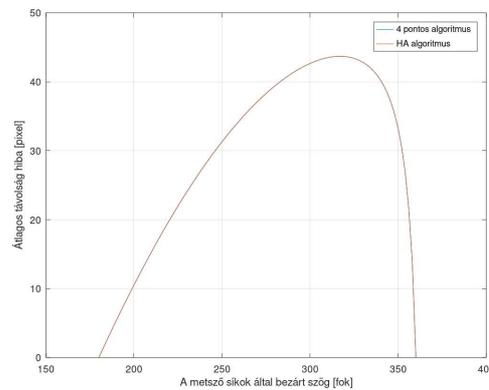
A kapott hibagörbéből látszik, hogy homográfia alapján elkülöníthető a két sík egymástól, ugyanis  $\alpha = 0^\circ$  és  $180^\circ$ , azaz egybeeső síkok esetén számítási pontosságon belül nullát kapunk, a többi esetben pedig nullától különböző a megkülönböztetéshez felhasználható számot. Fontos megjegyezni, hogy az ábrán látható egy divergenciát tartalmazó rész, mely azzal az esettel felel meg, mikor a kamerák pontosan a síkokon helyezkednek el, ilyenkor képeiken a síkok csupán egyenesként jelennek meg, így nem értelmezett a homográfia rajtuk. Az eredményből látszik, hogy bár a *4 pontos* és a *HA* algoritmus azonos eredményt ad 1000 megfeleltetésből. Általánosan a *HA* algoritmus robusztus módszerek alkalmazása esetén hasznosabb lesz, mivel kevesebb



(a) 3D nézet



(b) Felülnézet



(c) Átlagos hiba

6.2. ábra. Kamerával szemben lévő tompa szöget bezáró metsző síkok szimulációs elrendezése és a kapott eredmény.

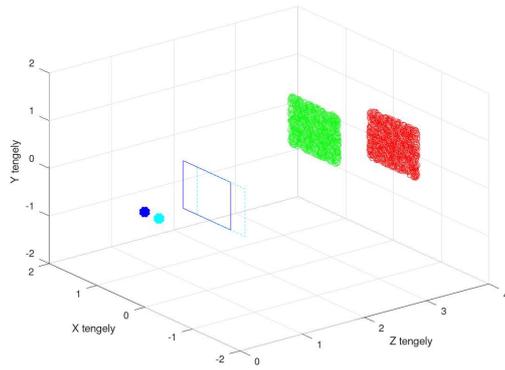
megfeleltetésből képes homográfiát becsülni mint a *4 pontos*.

A következő esetben azt vizsgáltam, hogyan változik a számított átlagos hiba, ha  $\alpha \in [180^\circ, 360^\circ]$ , mely a 6.1 ábrán látható. Az előző esethez hasonlóan az itt kapott eredmények is arra utalnak, hogy homográfiák alapján jól elkülöníthetők a síkok. Hasonlóan kapjuk továbbá azt is, hogy amint a két sík közeledik a párhuzamossághoz, az átlagos hiba nullához tart.

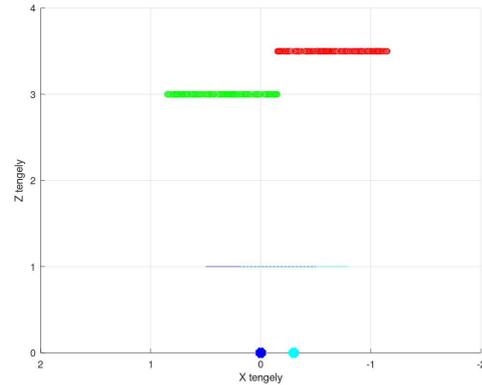
Összegezve, a szimulációs eredményekből látható tehát, hogy metsző síkok elkülönítésére lehetőségünk van homográfiák alapján. A következő alfejezetben a párhuzamos síkok esetét vizsgálom.

### 6.1.2. Párhuzamos síkok

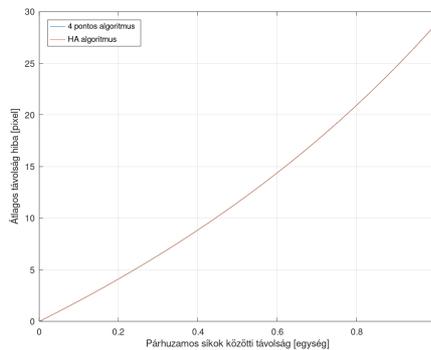
Az előző alfejezethez hasonló szimulációs környezetben vizsgáltam a párhuzamos síkok elkülöníthetőségét. Először szimuláltam a kamerákkal szemben lévő, képsíkjukkal párhuzamos síkok esetét, mely a valós életben függőleges szemben lévő párhuzamos épületfalaknak felel meg, vagy légi felvételek estén épületek és az utcaszint képé-



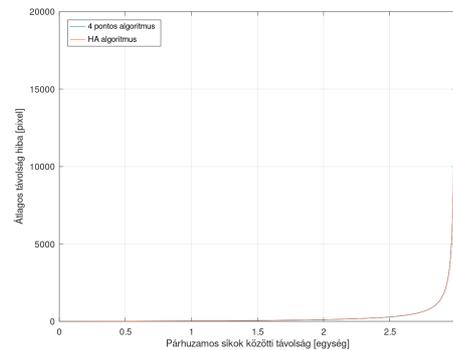
(a) 3D nézet



(b) Felülnézet



(c) Átlagos hiba 0-tól 1 egységig terjedő síktávolság esetén



(d) Átlagos hiba 0-tól 3 egységig terjedő síktávolság esetén

6.3. ábra. A kamera előtt elhelyezkedő kamera síkjával párhuzamos síkok szimulációs elrendezése és a kapott eredmények.

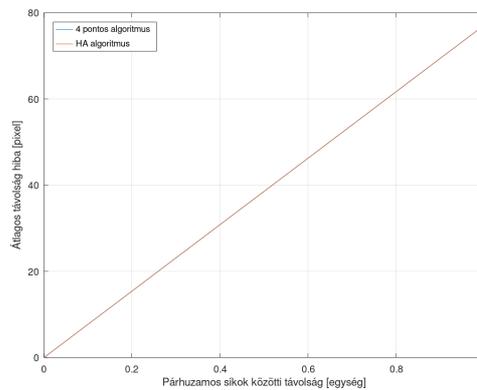
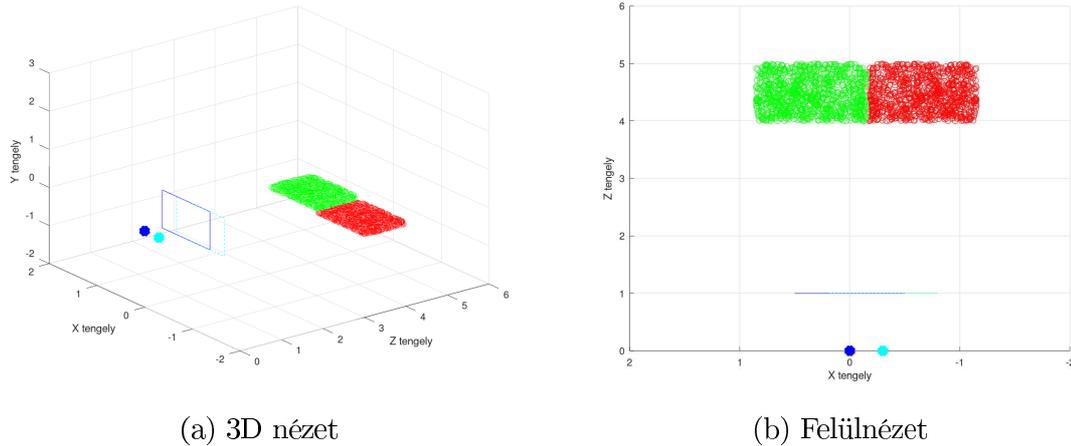
vel. Ezt követően a járdák és úttest esetét leíró szimulációt végeztem el, melyben a kamerával szemben található párhuzamos síkok merőlegesek a kamerák képsíkjára.

### Kamera síkjával párhuzamos síkok (épületfalak)

A szimuláció során kezdetben egybeeső síkokból kiindulva a kamerák  $z$  tengelye mentén szimulációs lépésenként eltoltam a síkokat. A 6.3 ábrán látható szimulációs eredményekből következik, hogy ebben az esetben is lehetőségünk van síkok szeparálására a becült homográfiák alapján.

### Kamera síkjára merőleges síkok (úttest és járda)

A kamera síkjára merőleges síkok esetén is egybeeső síkokból indultam ki, majd a kamerák  $y$  tengelye mentén szimulációs lépésenként eltoltam a síkokat. A 6.4 ábrán látható szimulációs eredményekből összevetve a korábbi eredményekkel (6.1, 6.2 és 6.3 ábrák) arra következtethetünk, hogy síkhomográfiák becslése lehetőséget ad különböző párhuzamos és metsző síkok elkülönítésére.



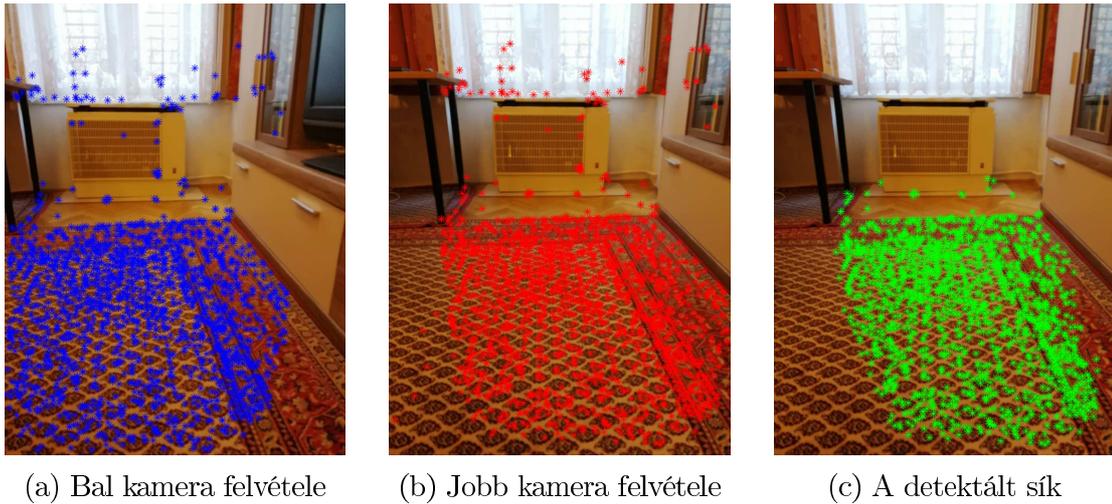
6.4. ábra. A kamera előtt elhelyezkedő kamera síkjára merőleges síkok szimulációs elrendezése és a kapott eredmény.

A következő fejezetben tesztelem a síkok detektálását homográfia becslés alapján valós beltéri és kültéri adatok esetén.

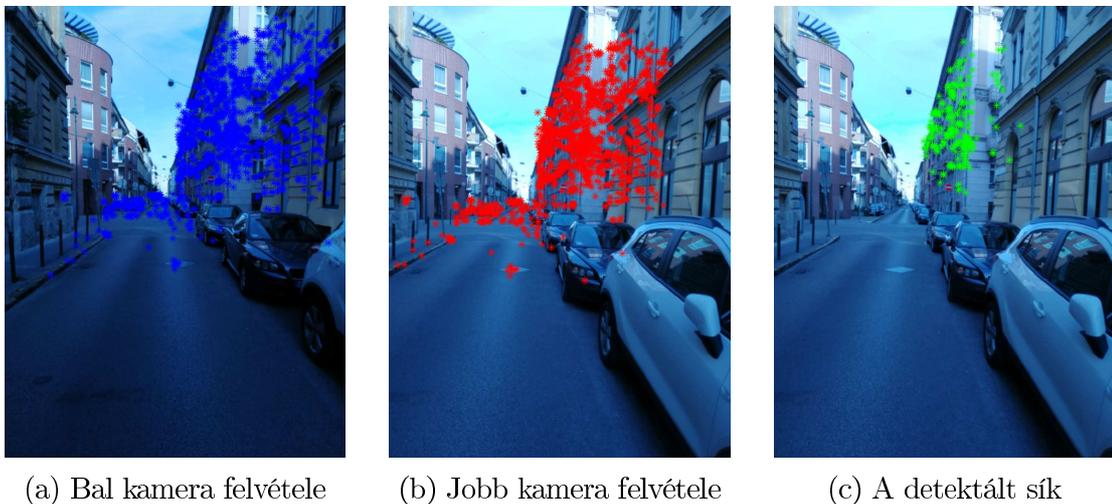
## 6.2. Valós tesztek

Ebben a fejezetben megmutatom, hogy milyen eredmény adható a harmadik fejezetben definiált (*Szekvenciális RANSAC*) algoritmussal erősen zajjal terhelt környezetben. A teszteléshez állítható fókusz távolsággal rendelkező telefon segítségével beltéri (szoba) és kültéri (úttest épületek között) képeket készítettem és megkerestem a képpárokban található legdominánsabb síkokat. Majd ezután összehasonlítom a 4.1 fejezetben definiált módszert a hagyományos *4 pontos* algoritmussal felülnézeti, kvadrokopteres képpárokban.

Első lépésként pontmegfeleltetések meghatározására volt szükségem a sztereó képpárokban. Ez valós életből származó adatok esetén nagy kihívást jelent, mivel a legelterjedtebb pontmegfeleltetéseket adó algoritmusok, mint például a *SIFT*<sup>[33]</sup> és az



6.5. ábra. A legnagyobb kiterjedésű sík meghatározása beltéri képpárok esetén

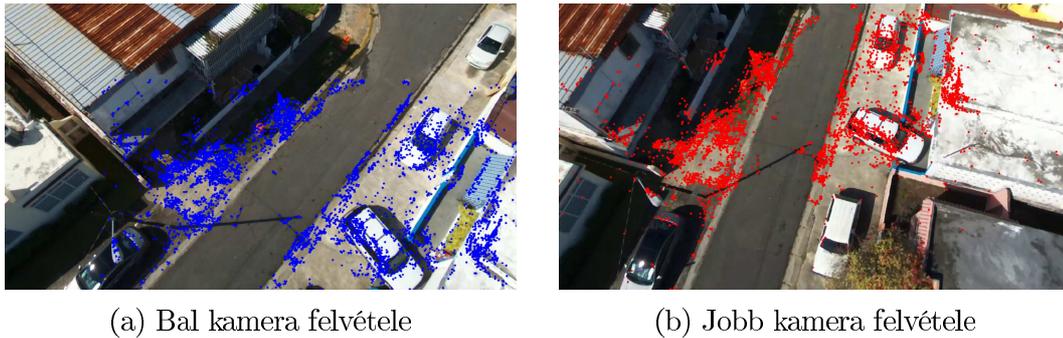


6.6. ábra. A legnagyobb kiterjedésű sík meghatározása kültéri képpárok esetén

*ASIFT*<sup>[25]</sup>, jellemző pontokra (*feature points*) támaszkodnak, amelyek a számomra detektálni kívánt objektumokon (síkok) tipikusan nagyon ritkák. Az általuk megtalált megfeleltetések általában síkok szélének közelében található nagy térbeli geometria változást tartalmazó részek (például ablakkeretek sarka, épület falának mintázata).

Ezt követően ezen zajjal erősen terhelt adatokon futtatva az algoritmust a 6.5 és 6.6 ábrán látható eredményeket kaptam a legmeghatározóbb síkokra. Fontos megjegyezni, hogy különös nehézséget jelent, hogy kültéri adatok esetén tipikusan sok megfeleltetést tartalmaztak az adatok a kamera felvételek horizontjának közelében, ami nem kívánatos, mivel a horizonton található pontok távolságuk miatt a kamera képeken egy síknak látszanak, ezért ezen adatokat figyelmen kívül hagytam a síkdetektálás során.

Érdeemes megemlíteni, hogy a megtalált síkok lehetőséget adnak a kamerákat össze-



6.7. ábra. Kvadrokopteres affin pontmegfeleltetések.

6.8. ábra. Különböző RANSAC paraméterek esetén detektált síkok. A felső sorban látható a  $4$  pontos algoritmus. Az alsó sorban a 4.1-ben definiált algoritmus. Egy oszlopon belül azonosak a paraméterek.

kötő transzformációk meghatározására, ha homográfia mátrixukat felbontjuk<sup>[31]</sup>. A dekompozíció elvégzéséhez azonban előbb ismernünk kell a kamera belső paramétereit, melyeket megkaphatunk például a jól ismert *sakktábla módszerrel*.<sup>[20][34]</sup> Összegezve tehát valós, igen zajos, és zömében a nem detektálni kívánt pontokat tartalmazó adatok esetén is használható az algoritmus domináns síkok detektálására.

Miután elvégeztem a városi és beltéri tesztelést, az 6.7- 6.10 ábrán látható kvadrokopteres felvételeken teszteltem a 4.1 alfejezetben bemutatott homográfia-becselő módszert. A kvadrokopteres felvételek jó közelítésben teljesítik az alfejezetben definiált megszorításokat. A felvételek felül nézetből készültek, továbbá a kamera  $z$  tengelyével közel párhuzamos normálvektorú síkok találhatók rajta, mint például úttest, járda és épületek teteje.

*Szekvenciális RANSAC* módszert használva síkok detektálására összehasonlítottam



6.9. ábra. 1. Kertvárosi kvadrokopteres felvétel. A felső 4.1-ben definiált algoritmus eredménye. Az alsó pedig a *4 pontos* algoritmusé.

a *4 pontos* homográfia becslést, valamint a 4.1 alfejezetben bemutatott módszert. A tesztelés során azt tapasztaltam, hogy az egy affin megfeleltetéssel dolgozó módszer érzékenyebb a *4 pontos* módszernél azonos *RANSAC* paraméter beállítások esetén. Ennek eredményeképpen jobban képes megkülönböztetni egymással párhuzamos síkokat, mint például az aszfalt és a tető síkját. A kapott eredmények két különböző paraméterezés esetén az 6.8- 6.10 ábrán láthatók. Megfigyelhető, hogy a *4 pontos* módszer az épület tetejéhez sorol néhány az utcán elhelyezkedő pontot is. Ezzel ellentétben az egy affin megfeleltetéssel dolgozó módszer képes külön választani az épületek tetejét az utca szintjétől. Fontos megjegyezni, hogy az affin transzformá-



6.10. ábra. 2. Kertvárosi kvadropteres felvétel. A felső 4.1-ben definiált algoritmus eredménye. Az alsó pedig a *4 pontos* algoritmusé.

ciókra támaszkodó algoritmus nem csupán kevesebb adatból volt képes pontosabb síkmeghatározásra, hanem pontatlanabb adatokból is. A bemeneti adatként használt becült affin transzformációk sokkal pontatlanabbak voltak, mint a pontmegfeleltetések.

További valós tesztelési lehetőség lett volna a függőleges síkok detektálása hasonló környezetben, azonban ennek elvégzéséhez nem álltak rendelkezésemre megfelelő adatok.

Összegezve a kapott eredményeket, sikerült megmutatnom, hogy valós, felülnéze-

ti kvardokpteres felvételeken jobb eredmény érhető el a 4.1 alfejezetben bevezetett módszerrel, mint a hagyományos *4 pontos* algoritmus segítségével.

## 7. fejezet

# Összefoglalás

A dolgozatban elsőként áttekintettem a sztereó látás alapfogalmait. Ezt követően a 3. fejezetben bemutatam, hogyan használható homográfiák segítségével síkdetektálásra a *szekvenciális RANSAC* algoritmus. A 4. fejezetben megmutattam, hogy speciális esetekben a hagyományos pontmegfeleltetések helyett affin megfeleltetéseket használva a szakirodalomban ismert módszereknél kevesebb megfeleltetés használatával megadható síkból síkba transzformációk homográfia mátrixa. Ennek nagy a jelentősége, hiszen homográfia mátrixok segítségével mind síkokat, mind a sztereó képpár kamerái közötti transzformációt meg lehet adni. Az ilyen módszerek esetén különösen fontos, hogy hány megfeleltetés szükséges a becsléshez, mivel a robusztus módszerek érzékenysége, valamint iterációk száma függ ezen megfeleltetések számától. Az 5. fejezetben megmutattam, hogy síkmozgás esetén affin megfeleltetéseket használva elegendő egy megfeleltetés a homográfia becsléséhez. Ezt követően az 6. fejezetben részletesen elemeztem szimulációk segítségével kalibrált sztereó képpáron történő síkok detektálhatóságát. Kielemeztem a valós életben járda és úttest megkülönböztethetőségének megfelelő szimulációs esetet, valamint épület falainak detektálhatóságát a bezárt szögük függvényében. Végül az 6.2 alfejezetben valós életből vett kvadrokopteres felülnézeti adatokon összehasonlítottam a 4.1 alfejezetben javasolt módszert a hagyományos *4 pontos* algoritmussal.

**Tudományos újdonságként a dolgozatban** a 4. fejezetben levezetem, hogy speciális, felülnézeti kamerás elrendezésekben hagyományos pontmegfeleltetések helyett affin pontmegfeleltetéseket használva kevesebb megfeleltetésből becsülhető a homográfia mátrixa. Továbbá megmutattam, hogy az 4.1 alfejezetben levezetett algoritmus *szekvenciális RANSAC* használatával azonos paraméterezés esetén jobban képes elkülöníteni párhuzamos síkokat a hagyományos *4 pontos* algoritmussal.

Továbbá az 5. fejezetben megmutattam, hogy síkmozgás esetén függőleges és vízszintes síkok homográfia mátrixa egy affin megfeleltetésből számítható.

# Köszönetnyilvánítás

Szeretnék köszönetet mondani a témavezetőmnek, Dr. Hajder Leventének az általa nyújtott szakmai segítségért.

EFOP-3.6.3-VEKOP-16-2017-00001: Tehetséggondozás és kutatói utánpótlás fejlesztése autonóm járműirányítási technológiák területén – A projekt a Magyar Állam és az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

# Irodalomjegyzék

- [1] D. Barath and L. Hajder, „Novel ways to estimate homography from local affine transformations,” 2016.
- [2] D. Barath and L. Hajder, „A theory of point-wise homography estimation,” *Pattern Recognition Letters*, vol. 94, pp. 7–14, 2017.
- [3] M. A. Fischler and R. C. Bolles, „Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM*, pp. 381–395, 1981.
- [4] O. Saurer, P. Vasseur, R. Bouteau, C. Demonceaux, M. Pollefeys, and F. Fraundorfer, „Homography based egomotion estimation with a common direction,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 2, pp. 327–341, 2017.
- [5] A. Bjorck, *Numerical methods for least squares problems*, vol. 51. Siam, 1996.
- [6] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [7] D. Barath, J. Matas, and L. Hajder, „Accurate closed-form estimation of local affine transformations consistent with the epipolar geometry,” 2016.
- [8] J. Nemeth, C. Domokos, and Z. Kato, „Recovering planar homographies between 2d shapes,” in *2009 IEEE 12th International Conference on Computer Vision*, pp. 2170–2176, IEEE, 2009.
- [9] C. Jawahar and P. Jain, „Homography estimation from planar contours,” in *Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06)*, pp. 877–884, IEEE, 2006.
- [10] P. K. Mudigonda, C. Jawahar, and P. Narayanan, „Geometric structure computation from conics,” in *ICVGIP*, pp. 9–14, 2004.

- [11] A. Sugimoto, „A linear algorithm for computing the homography from conics in correspondence,” *Journal of Mathematical Imaging and Vision*, vol. 13, no. 2, pp. 115–130, 2000.
- [12] S. Choi, T. Kim, and W. Yu, „Performance evaluation of ransac family,” *Journal of Computer Vision*, vol. 24, no. 3, pp. 271–300, 1997.
- [13] P. J. Rousseeuw, „Least median of squares regression,” *Journal of the American statistical association*, vol. 79, no. 388, pp. 871–880, 1984.
- [14] R. Toldo and A. Fusiello, „Robust multiple structures estimation with j-linkage,” in *European conference on computer vision*, pp. 537–547, Springer, 2008.
- [15] L. Magri and A. Fusiello, „T-linkage: A continuous relaxation of j-linkage for multi-model fitting,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3954–3961, 2014.
- [16] T. Werner and A. Zisserman, „New techniques for automated architectural reconstruction from photographs,” in *European conference on computer vision*, pp. 541–555, Springer, 2002.
- [17] Z. Zhang and A. R. Hanson, „3d reconstruction based on homography mapping,” *Proc. ARPA96*, pp. 1007–1012, 1996.
- [18] Z. Chuan, T. Da Long, Z. Feng, and D. Z. Li, „A planar homography estimation method for camera calibration,” in *Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation. Computational Intelligence in Robotics and Automation for the New Millennium (Cat. No. 03EX694)*, vol. 1, pp. 424–429, IEEE, 2003.
- [19] T. Ueshiba and F. Tomita, „Plane-based calibration algorithm for multi-camera systems via factorization of homography matrices,” in *null*, p. 966, IEEE, 2003.
- [20] Z. Zhang, „A flexible new technique for camera calibration,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, 2000.
- [21] S. J. Prince, K. Xu, and A. D. Cheok, „Augmented reality camera tracking with homographies,” *IEEE Computer graphics and Applications*, vol. 22, no. 6, pp. 39–45, 2002.
- [22] J. Zhou and B. Li, „Robust ground plane detection with normalized homography in monocular sequences from a robot platform,” in *2006 International Conference on Image Processing*, pp. 3017–3020, IEEE, 2006.

- [23] S. Mann, J. Huang, R. Janzen, R. Lo, V. Rampersad, A. Chen, and T. Doha, „Blind navigation with a wearable range camera and vibrotactile helmet,” in *Proceedings of the 19th ACM international conference on Multimedia*, pp. 1325–1328, ACM, 2011.
- [24] K. Mikolajczyk and C. Schmid, „An affine invariant interest point detector,” in *European conference on computer vision*, pp. 128–142, Springer, 2002.
- [25] G. Yu and J.-M. Morel, „ASIFT: An Algorithm for Fully Affine Invariant Comparison,” *Image Processing On Line*, vol. 1, pp. 11–38, 2011.
- [26] D. Mishkin, J. Matas, and M. Perdoch, „Mods: Fast and robust method for two-view matching,” *Computer Vision and Image Understanding*, vol. 141, pp. 81–93, 2015.
- [27] E. Vincent and R. Laganière, „Detecting planar homographies in an image pair,” in *ISPA 2001. Proceedings of the 2nd International Symposium on Image and Signal Processing and Analysis. In conjunction with 23rd International Conference on Information Technology Interfaces (IEEE Cat.*, pp. 182–187, IEEE, 2001.
- [28] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, *et al.*, „A density-based algorithm for discovering clusters in large spatial databases with noise,” in *Kdd*, vol. 96, pp. 226–231, 1996.
- [29] O. Oreifej, N. Lobo, and M. Shah, „Horizon constraint for unambiguous uav navigation in planar scenes,” in *2011 IEEE International Conference on Robotics and Automation*, pp. 1159–1165, IEEE, 2011.
- [30] J.-C. Bazin, H. Li, I. S. Kweon, C. Demonceaux, P. Vasseur, and K. Ikeuchi, „A branch-and-bound approach to correspondence and grouping problems,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 7, pp. 1565–1576, 2013.
- [31] E. Malis and M. Vargas, *Deeper understanding of the homography decomposition for vision-based control*. PhD thesis, INRIA, 2007.
- [32] O. D. Faugeras and F. Lustman, „Motion and structure from motion in a piecewise planar environment,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 2, no. 03, pp. 485–508, 1988.
- [33] T. Lindeberg, „Scale invariant feature transform,” 2012.
- [34] J.-Y. Bouguet, „Matlab calibration tool,” URL [http://www.vision.caltech.edu/bouguetj/calib\\_doc](http://www.vision.caltech.edu/bouguetj/calib_doc), 2015.