

A normalizált, 8-pontos algoritmus perturbációanalízise egy újabb szemszögből

Bódis-Szomorú András, Dabóczi Tamás

Budapesti Műszaki és Gazdaságtudományi Egyetem
Méréstechnika és Információs Rendszerek Tanszék
1117 Budapest, Magyar tudósok körútja 2.
{bodis,daboczi}@mit.bme.hu,
<http://www.mit.bme.hu/~{bodis,daboczi}>

Absztrakt. A két, kalibrálatlan nézet projektív viszonyát teljes mértékben leíró fundamentális mátrix meghatározása lényeges a nézetek közötti, robosztus pont-regisztrációtól, a sűrű sztereón és a mozgás-kompenzáción keresztül a 3D-rekonstrukciós alkalmazásokig. Az egyik, széles körben használt, ma már alapvetőnek számító módszer a fundamentális mátrix becslésére a 8-pontos algoritmus, amelynek direkt alkalmazása közismerten rosszul kondicionált problémára vezet, míg a kondíció megfelelő transzformációkkal javítható. Ugyan az alapvető forrásmunkák – szemléltető jelleggel – kitérnek rá, a helyes normalizáló transzformációk megválasztása, valamint ezek összefüggése a probléma kondicionáltságával azonban, paradox módon, matematikailag nem teljesen tisztázott vagy sokszor elnagyolt. Cikkünkben elsősorban ezekre a kérdésekre szeretnénk rávilágítani.

1. Bevezetés

A kétnézetes konfigurációra érvényes (projektív) geometriai kényszereket az ún. epipoláris geometria, algebrailag pedig az esszenciális mátrix illetve a fundamentális mátrix írja le. Meghatározásuk kulcsfontosságú probléma a gépi látásban.

Az esszenciális mátrix Longuet-Higgins [1] munkájában jelent meg először, míg azt, hogy az esszenciális mátrix projektív kényszereket ír le, és kalibrálatlan esetre is általánosítható, Hartley [2] és Faugeras [3] is felismerte. A fundamentális mátrix kiszámítási módjait Luong és Faugeras [4] és Zhang [5] tekintik át és hasonlítják össze az elsők között, de [6] egy újabb, kiváló összefoglaló. A *fundamentális mátrix* elnevezést [4] javasolja.

A számítások bizonytalanságának jellemzésére [5] bevezeti a perturbált epipoláris egyenesek nyalábját közrefogó kvadratikus görbét, [4] pedig stabilitási analízist végez linearizációval és munkapont körüli szimulált perturbációval. A kritikus, degenerált eseteket, amelyek mellett a számítás (a priori információk hiányában) nem végezhető el, például [4] és [7] vizsgálja.

A számítási módszerek rendszerint a két kép között megfelelően pontpárokra, mint bemenő adatokra illesztik a kétnézetes elrendezés projektív modelljét

valamilyen kritériumfüggvény alapján. A módszerek alapvetően a kritérium megfogalmazásában, ezen belül a hibatagok súlyozásában térnek el, és leginkább analitikus (máshol "lineáris"), iteratív és robusztus csoportokba szokás őket sorolni. A 8-pontos algoritmus és variánsai az analitikus, lineáris LS-módszerek közé sorolható. Ezek valamilyen – geometriai jelentéssel nem bíró – algebrai hibát minimalizálnak, de egyszerűek és gyorsak. Lineárisra visszavezethető, nemlineáris módszer pl. a virtuális parallaxis és a lineáris altér módszere [5]. Az iteratív módszerek rendszerint valamilyen – paraméterekben nemlineáris – geometriai hibát minimalizálnak nemlineáris optimalizálási módszerekkel (pl. Gauss-Newton, Levenberg-Marquardt). Ide sorolható a gradiens alapú módszer, a párosított pontok és az epipoláris egyenesek távolságának minimalizálása, valamint a Maximum Likelihood (ML) módszerek is (pl. Hartley optimális módszere [7]; HEIV, FNS és továbbfejlesztései [8]). Ezek a lineáris módszereknél jobb becslőket eredményeznek, de jóval bonyolultabbak, futási idejük nagyobb, az inicializáláshoz pedig lineáris módszer szükséges. Az iniciális becslő minőségétől nagy mértékben is függhet a megtalált lokális optimumhely és a konvergencia sebessége [7],[8]. Mind a lineáris, mind az iteratív módszerek érzékenyek a mérések, már kis számú, kiugró hibáira is. A robusztus módszerek ezek hatását igyekeznek csökkenteni (M-becslők módszere) vagy megszüntetni (LMedS, valamint RANSAC és válfajai). Ezek általában szintén nagy futási idejűek, és a detektált, téves adatok kizárásával célszerű a modellt – pl. lineáris módszerrel – újraindítani a becslő minőségének javításához. Mindmáig fontos szerepe, egyszerűsége és gyorsasága miatt cikkünk a 8-pontos algoritmusra fókuszál.

A 8-pontos algoritmus egy jól ismert, LS-regresszió alapuló, lineáris módszer, amellyel a fundamentális mátrix számítható, ha a képek között 8 vagy több, megfeleltetett pontpár adott. Kiemelkedő egyszerűsége és gyorsasága ellenére sokan támadták, mert a komplexebb módszerekhez képest – a bemenő adatokra nézve – lényeges instabilitást, valamint nagy illesztési hibát produkált (pl. [5]). Ezek az eredmények ellenben leginkább az algoritmus „nem körültekintő” alkalmazásának tulajdoníthatók [9]. Hartley a bemenő, képi adatok előzetes normalizáló transzformációját, majd a végeredmény inverz transzformációját javasolja a probléma numerikus kondíciójának javítására [9]. A transzformációk a pontrendszerek első és másodrendű nyomatékát változtatják meg. Noha a kondíció ezzel *átlagosan* valóban javítható, a számított fundamentális mátrix nem invariáns a koordináta-rendszerek transzformációjára [9],[10]. Torr és Fitzgibbon [10] rámutat arra, hogy emiatt Hartley módszerénél az epipoláris geometriai kényszer alapján nem szűrhető, téves pont-párosítások a normalizáló transzformációkat hibával terhelik¹, és ezért [10] olyan, ugyancsak egyszerű módszert mutat be, ami invariáns a koordináta-rendszer hasonlósági transzformációira. A javasolt módszer nem váltja ki teljesen a Hartley-algoritmust, mert kizárja az affin nézetek esetét. A Torr-Fitzgibbon-féle módszerrel ezért jelen munkánkban a továbbiakban nem foglalkozunk.

Hartley [9] arra is szemléltető iránymutatást ad, hogy a javasolt normalizáló transzformáció hogyan függ össze a probléma kondíciójával. Felkeltette a fi-

¹ hatnak a pontrendszer súlypontjára és tehetetlenségi nyomatékára

gyelmünket azonban, hogy a kondíciós szám egyrészt csak egy felső korlát, amelyre a magyarázat alapján alsó becslést kapunk (gyenge kritérium), másrészt a kondíciós szám inhomogén lineáris egyenletrendszerek érzékenységéről árulkodik, míg ezzel szemben a 8-pontos algoritmus egy másik típusú problémát, egy *homogén, túlhatározott, lineáris egyenletrendszert* old meg, ami egy sajátérték-sajátvektor problémára vezethető vissza [9],[11]. A „hagyományos” kondíciós szám helyett a sajátvektorok perturbációanalízise szolgáltatja az itt releváns felső korlátot [12]. Emellett a sajátvektorok perturbációjára egyszerű elsőrendű közelítés is létezik [13],[14]. Ezért meglepő, hogy az algoritmus statisztikai vizsgálataiban legtöbbször nem térnek ki erre a kérdésre [4],[5],[11], noha [14] egy ellenpélda. [14] a probléma igényes statisztikai vizsgálatát adja, de ennek során csak az egyik képen feltételez pont-lokalizációs hibát, ami nem reális. Ugyan Hartley igazolja [9], hogy a *normalizált* 8-pontos algoritmus eredménye *átlagosan* megközelíti a jóval komplexebb technikákét is, mégis úgy érezzük, hogy a magyarázat pontosításra és kiegészítésre szorul. Elsősorban a perturbációanalízis módszertanára hívnánk fel a figyelmet, másodsorban a Hartley által javasolt normalizáló transzformáció nem bizonyítottan „optimális”, így erre a kérdésre is kitérünk. Harmadsorban pedig a két lépéses, Hartley-féle módszer második, rang-korrekciós lépésénél alkalmazott eljárást is megvizsgáljuk.

A második fejezetben bemutatjuk a Hartley-féle, normalizált 8-pontos algoritmust, és a harmadik fejezetben vizsgáljuk meg, hogy az algoritmus első fázisában mi mondható a fundamentális mátrix érzékenységéről. A negyedik fejezetben pedig a rang-korrekciónál jelentkező perturbációt és annak mérési módját vizsgáljuk. Az ötödik fejezet a normalizáló transzformációkat tárgyalja, végül pedig a hatodik fejezetben leszűrjük a tanulságokat.

2. Hartley-féle 8-pontos algoritmus

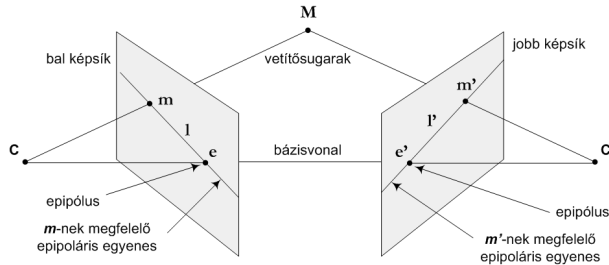
Ha adottak egy \mathbf{M} térbeli pont $\mathbf{m} = (x, y, 1)^T$ és $\mathbf{m}' = (x', y', 1)^T$ centrális vetületei, akkor az \mathbf{m} ponton keresztülhaladó vetítésugár képe a másik nézetben egy $\mathbf{I}' = (l'_1, l'_2, l'_3)^T$ ún. epipoláris egyenes (1. ábra). Az $\mathbf{m} \rightarrow \mathbf{I}' = \mathbf{F}\mathbf{m}$ leképezést az $\mathbf{F} = [f_{ij}]$ 3-adrendű fundamentális mátrix határozza meg. Az epipoláris kényszer szerint az \mathbf{m}' pont rajta van az \mathbf{I}' egyenesen, ezért $\mathbf{m}'^T \mathbf{I}' = \mathbf{m}'^T \mathbf{F}\mathbf{m} = 0$. A szimmetria miatt $\mathbf{m}' \mapsto \mathbf{l} = \mathbf{F}^T \mathbf{m}'$.

Mivel \mathbf{F} egyrészt homogén mátrix, vagyis skalárral szorozva ugyanazt reprezentálja, másrészt rangja 2, tehát szinguláris ($\det \mathbf{F} = 0$), ezért szabadságfoka 7. Belátható, hogy a fundamentális mátrix minden projektív információt magába foglal a kétnézetes elrendezésről, pl. az $\mathbf{e} \leftrightarrow \mathbf{e}'$ epipólusok az $\mathbf{F}\mathbf{e} = \mathbf{0}$ és az $\mathbf{F}^T \mathbf{e}' = \mathbf{0}$ kényszerekből számíthatók.

A 8-pontos algoritmus célja \mathbf{F} számítása elegendő számú $\mathbf{m}_i = (x_i, y_i, 1)^T \leftrightarrow \mathbf{m}'_i = (x'_i, y'_i, 1)^T$ pontpárból ($i = 1, 2, \dots, n$). A módszer az

$$\mathbf{m}'_i{}^T \mathbf{F}\mathbf{m}_i = \mathbf{a}_i^T \mathbf{f} = 0, \quad \mathbf{a}_i^T = (x'_i x_i, x'_i y_i, x'_i, y'_i x_i, y'_i y_i, y'_i, x_i, y_i, 1) \quad (1)$$

epipoláris kényszerből indul ki, ahol \mathbf{f} az \mathbf{F} mátrix elemeit tartalmazza soronként. Az összes kényszert, kompakt módon az $\mathbf{A}\mathbf{f} = \mathbf{0}_n$ homogén, lineáris egyenle-



1. ábra: A kétnézetes elrendezés geometriája (epipoláris geometria).

trendszer foglalja magába, ahol $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_n]^T$. Ennek akkor van nemtriviális megoldása, ha \mathbf{A} rangja 9-nél kevesebb. A degenerált esetektől eltekintve $\text{rg } \mathbf{A} = 8$, így a nulltér dimenziója 1, vagyis \mathbf{f} egy skálatényező erejéig meghatározható 8 vagy több pontpárból. Mindez ideális, zajmentes méréseket feltételez. A gyakorlatban azonban az $\mathbf{m}_i \leftrightarrow \mathbf{m}'_i$ párok lokalizációs (és párosítási) hibával terhelték. Ilyen esetben $\text{rg } \mathbf{A} = 9$ és az (1) egyenletnek nincs megoldása. Helyette célszerű azt az \mathbf{f} vektort megkeresni, amelyre az $\|\mathbf{A}\mathbf{f}\|^2 = \mathbf{f}^T \mathbf{A}^T \mathbf{A} \mathbf{f}$ norma minimális. \mathbf{F} szabadon hagyott skálatényezője túlparaméterezetté teszi a feladatot. Ezt és a triviális megoldást a 8-pontos algoritmus az $\|\mathbf{f}\| = \mathbf{f}^T \mathbf{f} = 1$ járulékos kényszerrel kerüli el. A Lagrange-multiplikátorok módszere szerint ez a feltételes minimalizálási probléma visszavezethető a

$$\mathcal{C}(\mathbf{f}, \lambda) = \mathbf{f}^T \mathbf{A}^T \mathbf{A} \mathbf{f} + \lambda(1 - \mathbf{f}^T \mathbf{f}) \tag{2}$$

költségfüggvény feltétel nélküli minimalizálására a λ segédváltozó és \mathbf{f} szerint. A $\partial \mathcal{C} / \partial \mathbf{f} = \mathbf{0}^T$ feltételből az $\mathbf{A}^T \mathbf{A} \mathbf{f}_{min} = \lambda \mathbf{f}_{min}$ sajátérték-sajátvektor problémára jutunk, ahol $\mathbf{A}^T \mathbf{A}$ szimmetrikus, pozitív szemidefinit mátrix, λ pedig ennek egy sajátértéke. Visszahelyettesítve $\mathcal{C}(\mathbf{f}_{min}, \lambda) = \lambda$, vagyis $\lambda = \lambda_{min}$ az $\mathbf{A}^T \mathbf{A}$ mátrix legkisebb sajátértéke, a keresett \mathbf{f}_{min} pedig az ehhez tartozó sajátvektor. Ezt a gyakorlatban az \mathbf{A} mátrix SVD-felbontásával határozzuk meg [7].

Eddig nem vettük figyelembe a $\det \mathbf{F} = 0$ nemlineáris szingularitási kényszert, ezért egyszerű módszert kaptunk, de nem garantált, hogy \mathbf{F}_{min} rangja 2, így az epipólusok nem számíthatók belőle (az epipoláris egyenesek nem metszik egymást az epipólusokban). A korrekció során az $\mathbf{F}_{min} = \mathbf{U}_f \text{diag}(\mu_1, \mu_2, \mu_3) \mathbf{V}_f^T$ SVD-felbontású mátrix helyett tekintjük az $\mathbf{F} = \mathbf{U}_f \text{diag}(\mu_1, \mu_2, 0) \mathbf{V}_f^T$ mátrixot ($\mu_1 \geq \mu_2 \geq \mu_3$). Belátható [5],[14], hogy ez $\|\mathbf{F} - \mathbf{F}_{min}\|_F$ értelemben az \mathbf{F}_{min} mátrixhoz lehető legközelebb álló 2-rangú mátrix, ahol $\|\cdot\|_F$ a Frobenius-norma.

Az eredeti, 8-pontos algoritmus tehát első fázisban az $r_i = \mathbf{m}_i^T \mathbf{F} \mathbf{m}_i$ maradó hibák $\sum_i r_i^2 = \|\mathbf{A}^T \mathbf{f}\|^2$ négyzetösszegét minimalizálja az $\|\mathbf{f}\| = \|\mathbf{F}\|_F = 1$ kényszer mellett, majd a második fázisban az eredményre alkalmazza rangkényszert. Az eredmény azonban nem invariáns a bemeneti ponthalmazok síkbeli transzformációira, így függ attól is, hogy milyen képi koordináta-rendszereket választottunk a két nézetben. Amennyiben az $\mathbf{m} \leftrightarrow \mathbf{m}'$ általános pár helyett az $\tilde{\mathbf{m}} = \mathbf{H}\mathbf{m} \leftrightarrow \tilde{\mathbf{m}}' = \mathbf{H}'\mathbf{m}$ transzformált párt tekintjük, ahol \mathbf{H} és \mathbf{H}' független

homográfiaák, akkor $\mathbf{m}'^T \mathbf{F} \mathbf{m} = \tilde{\mathbf{m}}^T \mathbf{H}'^{-T} \mathbf{F} \mathbf{H}^{-1} \tilde{\mathbf{m}} = 0$ miatt az új koordináta-rendszerben az epipoláris kényszer $\tilde{\mathbf{m}}^T \tilde{\mathbf{F}} \tilde{\mathbf{m}} = 0$, ahol $\tilde{\mathbf{F}} = \mathbf{H}'^{-T} \mathbf{F} \mathbf{H}^{-1}$ a transzformált pontokra érvényes fundamentális mátrix. Zajos $\mathbf{m}_i \leftrightarrow \mathbf{m}'_i$ mérések esetén ugyan $\tilde{r}_i = \tilde{\mathbf{m}}^T \tilde{\mathbf{F}} \tilde{\mathbf{m}} = \mathbf{m}'^T \mathbf{F} \mathbf{m} = r_i$, vagyis transzformációval és anélkül ugyanazt az algebrai hibát minimalizáljuk, de $\|\mathbf{F}\|_F = \|\mathbf{H}'^T \tilde{\mathbf{F}} \mathbf{H}\| \neq \|\tilde{\mathbf{F}}\|_F$ miatt nem ugyanazon kényszer mellett, így az eredmény függ \mathbf{H} és \mathbf{H}' megválasztásától.

Felmerül a kérdés, hogy hogyan célszerű megválasztani a koordináta-rendszert. Mivel a 8-pontos algoritmus rossz numerikus kondicionáltsága ismeretes, Hartley a kondíciószám javítását tűzi ki célul. Javaslatára

$$\mathbf{H} = \begin{bmatrix} 1/s & 0 & -x_g/s \\ 0 & 1/s & -y_g/s \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{m}_g = \frac{1}{n} \sum_{i=1}^n \mathbf{m}_i, \quad s = \left(\frac{1}{2n} \sum_{i=1}^n \|\mathbf{m}_i - \mathbf{m}_g\|^2 \right)^{\frac{1}{2}}, \quad (3)$$

ahol $\mathbf{m}_g = (x_g, y_g, 1)^T$ az eredeti pontrendszer súlypontja, s pedig egy skálázási tényező. Ezzel a súlypontot az origóba helyezi, és a pontokat izotróp módon skálázza úgy, hogy a pontok RMS²-távolsága a súlyponttól $\sqrt{2}$ legyen. Megjegyzendő, hogy [9] és [15] még az *átlagos távolságról*, míg [7, 282.o.] már az *RMS-távolságról* nyilatkozik. \mathbf{H}' megválasztása hasonlóképpen történik. Hartley [9] tehát elvégzi az $\mathbf{m}_i \leftrightarrow \mathbf{m}'_i$ pontpárok normalizációját külön a két nézetben, az így kapott $\tilde{\mathbf{m}}_i \leftrightarrow \tilde{\mathbf{m}}'_i$ párokra alkalmazza a 8-pontos algoritmust (rang-kényszerrel), majd az eredményezett $\tilde{\mathbf{F}}$ fundamentális mátrixot visszatranszformálja az eredeti koordináta rendszerbe.

Hartley [15] az anizotróp skálázást is vizsgálja, amelynek során a pontfelhők főirányai mentén külön is normalizál, valamint azt is megvizsgálja, hogy mi a különbség, ha csak a 8-pontos algoritmus első, illetve csak a második lépése (rang-kényszer) előtt normalizál és utána transzformál vissza. Tapasztalata, hogy az izotróp és az anizotróp skálázás eredménye között elenyésző a különbség, míg a normalizáció-denormalizáció a teljes, két lépéses, 8-pontos algoritmus előtt és után elvégezve adja a legjobb eredményt.

A normalizált algoritmus – az illesztés *átlagos* maradó hibája tekintetében – kiemelkedően jobb eredményt hoz, mint a klasszikus változat, emelet megközelíti a legjobb iteratív algoritmusok illesztési hibáit [9],[5],[15]. Az *átlag* azonban nem alkalmas arra, hogy egy illesztés *stabilitására* következtessünk, egy konkrét realizáció lehetséges „kilengését” ugyanis nem jellemzi. Erre szolgálhat a szórások számítása [6, 1. táblázat], vagy a perturbációanalízis, amellyel általában alsó és felső korlátokat számítunk a statisztikai paraméterek helyett.

3. Az első fázis numerikus kondicionáltsága

3.1. Hartley értelmezése

Legyen $\mathbf{A} = [a_{ij}] = \mathbf{U} \mathbf{S} \mathbf{V}^T$ az $n \times 9$ méretű \mathbf{A} adatmátrix SVD-felbontása. Ekkor $\mathbf{A}^T \mathbf{A} = \mathbf{V} \mathbf{D} \mathbf{V}^T$ az $\mathbf{A}^T \mathbf{A}$ mátrix sajátérték-felbontása, ahol $\mathbf{V} = [\mathbf{v}_1 \dots \mathbf{v}_9]$, $\mathbf{D} = \mathbf{S}^T \mathbf{S} = \text{diag}(\lambda_1, \dots, \lambda_9)$ és $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_9 \geq 0$. Zaj esetén $\lambda_9 \neq 0$.

² Root-Mean-Square, négyzetes közép, az elemek négyzetes átlagának a négyzetgyöke

Hartley [9] összefüggésbe hozza a probléma kondícióját a $\kappa_H = \lambda_1/\lambda_8$ aránnyal, majd becslést ad arra, hogy a $\mathbf{B} = [b_{ij}] = \mathbf{A}^T \mathbf{A}$ mátrix szerkezete illetve annak megváltoztatása a javasolt normalizáló transzformációkkal hogyan befolyásolja (előnyösen) ezt az arányt. A sajátértékek összefűzési tétele alapján [12, 396. o.] belátható [9], hogy $\kappa_H > \max_k b_{kk}/(b_{22} + b_{11})$, ahol $b_{kk} = \sum_{i=1}^n a_{ik}^2$. Eszerint, ha a legnagyobb, és a két utolsó diagonálemet azonos nagyságrendűvé tesszük, akkor κ_H alsó korlátja csökkenthető. A \mathbf{B} mátrix diagonálemerei (1) alapján

$$\{\sum (x'_i x_i)^2, \sum (x'_i y_i)^2, \sum x_i'^2, \sum (y'_i x_i)^2, \sum (y'_i y_i)^2, \sum y_i'^2, \sum x_i^2, \sum y_i^2, n\} \quad (4)$$

ahol az összegzés $i = 1, 2, \dots, n$ szerint értendő. Hartley szemléltető magyarázata szerint egy átlagos homogén pont koordinátái alapesetben $\mathcal{O}(10^2, 10^2, 10^0)$, mivel a kép méretei tipikusan néhány száz pixeles nagyságrendűek. Az előbbi gondolatmenet alapján, $\kappa_H > 10^8/(10^4 + 10^1)$, tehát a κ_H szám legalább 10^4 nagyságrendű. Ha a koordinátákat a (3) transzformációval normalizáljuk, akkor a diagonálemek kedvező irányban változnak, ezért κ_H alsó küszöbe nagyságrendekkel csökkenthető [9].

Általánosságban vizsgáljuk meg mit jelent valójában a Hartley által használt kondíciós szám. A 2-mátrixnorma alapú kondíciós szám kvadratikusan, invertálható \mathbf{C} mátrixra definiált $\kappa_2 \triangleq \text{cond } \mathbf{C} \triangleq \|\mathbf{C}\|_2 \cdot \|\mathbf{C}^{-1}\|_2$ szerint. Jelölje $\sigma_k(\mathbf{X})$ egy tetszőleges \mathbf{X} mátrix k -edik legnagyobb szinguláris értékét, $\lambda_k(\mathbf{X})$ pedig sajátértékét. Mivel $\mathbf{B} = \mathbf{A}^T \mathbf{A}$ szimmetrikus, pozitív szemidefinit mátrix, ezért $\lambda_k(\mathbf{B}) = \sigma_k(\mathbf{B}) = \sigma_k^2(\mathbf{A})$. A $\lambda_k = \lambda_k(\mathbf{B})$ és $\sigma_k = \sigma_k(\mathbf{A})$ jelölésekkel élve, belátható, hogy $\kappa_2 = \sigma_1(\mathbf{B})/\sigma_9(\mathbf{B}) = \lambda_1/\lambda_9 = \sigma_1^2/\sigma_9^2$.

A κ_2 kondíciós szám a $\mathbf{C}\mathbf{x} = \mathbf{b}$ típusú egyenletrendszer érzékenységét jellemzi. Ha \mathbf{x} ennek egy megoldása, továbbá \mathbf{y} a $(\mathbf{C} + \Delta\mathbf{C})\mathbf{y} = (\mathbf{b} + \Delta\mathbf{b})$ perturbált egyenletrendszer megoldása, ahol a relatív perturbáció korlátos $\|\Delta\mathbf{C}\| < \epsilon\|\mathbf{C}\|$, $\|\Delta\mathbf{b}\| < \epsilon\|\mathbf{b}\|$ és $\epsilon < 1$ szerint, akkor $\mathbf{C} + \Delta\mathbf{C}$ is nonszinguláris és a megoldás $\|\mathbf{y}\|/\|\mathbf{x}\|$ relatív perturbációja nem nagyobb, mint $2\epsilon\kappa/(1 - \epsilon\kappa)$ [12, 83. o.]. Ha $\Delta\mathbf{b} = \mathbf{0}$, akkor $\epsilon\kappa/(1 - \epsilon\kappa)$ a felső küszöb.

A fundamentális mátrix meghatározásánál, ugyan $\mathbf{B} = \mathbf{A}^T \mathbf{A}$ kvadratikusan, de akár szinguláris is lehet, ráadásul nem egy $\mathbf{B}\mathbf{x} = \mathbf{b}$ típusú problémával állunk szemben, hanem egy $\mathbf{B}\mathbf{v} = \lambda\mathbf{v}$ alakú, szimmetrikus sajátérték-sajátvektor problémával. Ezen belül a λ_9 legkisebb sajátértékhez tartozó \mathbf{v}_9 sajátvektor érzékenységére vagyunk kíváncsiak valójában, hiszen ez adja a fundamentális mátrixot.

3.2. Felső korlát a sajátvektorok perturbációjára

A fundamentális mátrix meghatározása szempontjából – az eddigiek alapján – az a lényeges, hogy hogyan módosul a sajátvektor, ha az \mathbf{A} adatmátrix az $\mathbf{A} + \Delta\mathbf{A}$ mátrixra módosul a koordináták bizonytalansága miatt. Ekkor a szimmetrikus $\mathbf{B} = \mathbf{A}^T \mathbf{A}$ mátrix helyett a

$$\mathbf{B} + \Delta\mathbf{B} = (\mathbf{A} + \Delta\mathbf{A})^T (\mathbf{A} + \Delta\mathbf{A}) = \mathbf{A}^T \mathbf{A} + \mathbf{A}^T \Delta\mathbf{A} + \Delta\mathbf{A}^T \mathbf{A} + \Delta\mathbf{A}^T \Delta\mathbf{A} \quad (5)$$

mátrixot kapjuk. Látható, hogy a $\Delta\mathbf{B} = \mathbf{A}^T \Delta\mathbf{A} + \Delta\mathbf{A}^T \mathbf{A} + \Delta\mathbf{A}^T \Delta\mathbf{A}$ perturbáció is szimmetrikus. Legyen \mathbf{B} spektrálfelbontása $\mathbf{V}^T \mathbf{B} \mathbf{V} = \text{diag}(\lambda_1, \dots, \lambda_9)$, ahol $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_9$. Transzformáljuk a perturbált $\mathbf{B} + \Delta\mathbf{B}$ mátrixot is ugyanezzel a hasonlósági transzformációval, azaz

$$\mathbf{V}^T \mathbf{B} \mathbf{V} + \mathbf{V}^T \Delta\mathbf{B} \mathbf{V} = \text{diag}(\lambda_1, \dots, \lambda_n) + \mathbf{E}, \quad \mathbf{E} = \mathbf{V}^T \Delta\mathbf{B} \mathbf{V} = \begin{bmatrix} \mathbf{E}_{11} & \mathbf{e} \\ \mathbf{e}^T & e \end{bmatrix}. \quad (6)$$

Ha az \mathbf{E} mátrixban $\|\mathbf{e}\| = 0$, akkor \mathbf{B} utolsó sajátvektora sajátvektora a $\mathbf{B} + \Delta\mathbf{B}$ mátrixnak is, vagyis a fundamentális mátrix becslőjét jelentő sajátvektor a $\Delta\mathbf{B}$ perturbáció hatására egyáltalán nem mozdul, noha a többi sajátvektor és a sajátértékek változhatnak (de a cél szempontjából ezek nem lényegesek).

Ennél több is mondható, ha a perturbáció

$$\|\Delta\mathbf{B}\| = \|\mathbf{E}\| \leq \frac{\lambda_8 - \lambda_9}{5} \quad (7)$$

szerint korlátos, ahol $d = \lambda_8 - \lambda_9$ a többi sajátérték *szeparációja* a legkisebb sajátértéktől. Alkalmazható ugyanis a következő perturbációs tétel, ami a [12, 8.1.12 tétel] egy speciális esete. Particionáljuk a \mathbf{V} mátrixot $\mathbf{V} = [\mathbf{V}' \ \mathbf{v}_9]$ szerint. Belátható, hogy létezik olyan $\mathbf{p} \in \mathbb{R}^8$ vektor, amire $\|\mathbf{p}\| < 4\|\mathbf{e}\|/d$, ugyanakkor $\tilde{\mathbf{v}}_9 = (\mathbf{v}_9 + \mathbf{V}'\mathbf{p})/\sqrt{1 + \mathbf{p}^T \mathbf{p}}$ a perturbált $\mathbf{B} + \Delta\mathbf{B}$ mátrix egységnyi 2-normájú sajátvektora. A sajátvektor perturbációjának egy lehetséges mértéke a perturbált és perturbálatlan sajátvektor által kifeszített *alterek távolsága*, a fenti tétel folytatása szerint ugyanis

$$\text{dist}(\mathcal{S}, \tilde{\mathcal{S}}) \leq 4 \frac{\|\mathbf{e}\|}{\lambda_8 - \lambda_9}, \quad (8)$$

ahol $\mathcal{S} = \text{span}(\mathbf{v}_9)$ az eredeti és $\tilde{\mathcal{S}} = \text{span}(\tilde{\mathbf{v}}_9)$ a perturbált sajátvektor által kifeszített, 1-dimenziós altér. Az alterek távolságának értelmezéséhez az alábbi lemmát használhatjuk fel.

1. lemma ([12], 2.6.1. tétel). *Amennyiben $\mathbf{W} = [\mathbf{W}_1 \ \mathbf{W}_2]$ és $\mathbf{Z} = [\mathbf{Z}_1 \ \mathbf{Z}_2]$ n -edrendű ortonormált mátrixok, és $\mathcal{S}_1 = \text{span}(\mathbf{W}_1)$ valamint $\mathcal{S}_2 = \text{span}(\mathbf{Z}_1)$, akkor az \mathcal{S}_1 és \mathcal{S}_2 alterek távolsága $\text{dist}(\mathcal{S}_1, \mathcal{S}_2) = \|\mathbf{W}_1^T \mathbf{Z}_2\| = \|\mathbf{Z}_1^T \mathbf{W}_2\|$.*

Esetünkben a $\mathbf{V} = [\mathbf{V}' \ \mathbf{v}_9]$ eredeti és a $\tilde{\mathbf{V}} = [\tilde{\mathbf{V}} \ \tilde{\mathbf{v}}_9]$ perturbált modálmátrixok ortonormáltak. Ezekre alkalmazva a lemmát, egyrészt

$$\text{dist}(\mathcal{S}, \tilde{\mathcal{S}}) = \|\tilde{\mathbf{v}}_9^T \mathbf{V}'\| = \|\tilde{\mathbf{v}}_9^T [\mathbf{v}_1 \ \dots \ \mathbf{v}_8]\| = \left\{ \sum_{i=1}^8 \cos^2 \varphi(\tilde{\mathbf{v}}_9, \mathbf{v}_i) \right\}^{\frac{1}{2}} \quad (9)$$

ahol $\varphi(\tilde{\mathbf{v}}_9, \mathbf{v}_i)$ a \mathbf{v}_i zajmentes sajátvektor és a $\tilde{\mathbf{v}}_9$ perturbált sajátvektor által bezárt szög, és kihasználtuk, hogy $\tilde{\mathbf{v}}_9^T \mathbf{v}_i = \varphi(\tilde{\mathbf{v}}_9, \mathbf{v}_i)$. Másrészt, a lemma szerint (9) ugyanazt adja, ha a $\varphi(\tilde{\mathbf{v}}_9, \mathbf{v}_i)$ szögeket kicseréljük a $\varphi(\mathbf{v}_9, \tilde{\mathbf{v}}_i)$ szögekre.

A lemma következménye, hogy

$$0 \leq \text{dist}(\mathcal{S}, \tilde{\mathcal{S}}) \leq 1. \quad (10)$$

A távolság 0, ha $\mathcal{S} = \tilde{\mathcal{S}}$, azaz $\tilde{\mathbf{v}}_9 = \mathbf{v}_9$, és 1, ha $\mathcal{S} \cap \tilde{\mathcal{S}}^\perp \neq \{0\}$ [12, 77.o.].

Könnyen ellenőrizhető, hogy kis perturbációk esetén elég pontos a

$$\text{dist}(\mathcal{S}, \tilde{\mathcal{S}}) \approx \|\Delta \mathbf{v}_9\| = \|\tilde{\mathbf{v}}_9 - \xi \mathbf{v}_9\|, \quad (11)$$

közelítés, ahol $\xi = \pm 1$ aszerint, hogy melyikkel kisebb a $\|\Delta \mathbf{v}_9\|$ norma. Ez azt jelenti, hogy fenti tétel felső korlátot szab a valódi és perturbált fundamentális mátrixok közötti eltérés Frobenius-normájára, ha a különbség nem túl nagy és ha a mátrixok normalizálva vannak (a ξ előjelre is figyelve).

Visszatérve a (8) korlátra, látható, hogy a 8-pontos algoritmus első fázisának kondicionáltsága nem a $\kappa_2 = \lambda_1/\lambda_9$, de nem is a $\kappa_H = \lambda_1/\lambda_8$ aránytól függ, hanem a $\lambda_8 - \lambda_9$ szeparációtól, és az $\|\mathbf{e}\|$ normától, azaz a perturbáció szerkezetétől.

Összefoglalva, a fundamentális mátrix vektorizált változatát jelentő \mathbf{v}_9 sajátvektor változatlan, ha a $\Delta \mathbf{B}$ perturbáció transzformáltjában $\mathbf{e} = \mathbf{0}$. Ha ez nem teljesül, de teljesül a (7) szerinti $\|\mathbf{E}\| < d/4$ feltétel, akkor minél nagyobb a $d = \lambda_8 - \lambda_9$ szeparáció, annál szűkebb tartományban perturbálódhat a megoldás. Ebben az esetben cél a szeparáció növelése. Látható, hogy a Hartley által használt λ_1/λ_8 arány – legalábbis explicit módon – nem jelent meg a küszöb felírásában. A (7) feltétel esetünkben, a \mathbf{B} mátrix szerkezete miatt, nagyon szigorúnak bizonyul, de az $\|\mathbf{e}\|$ norma relevanciáját a perturbációanalízisben ez nem csökkenti.

3.3. A perturbáció lineáris közelítése

Az előbb levezetett felső korlát – mint a lehető legrosszabb eset – akár igen pesszimisztikus is lehet. Lehetőség van azonban a sajátvektor perturbációjának közelítésére is kellően kis mértékű perturbáció mellett. [13] a fundamentális mátrix egyik robusztus módszerénél használja ezt, míg [14] egy olyan sztochasztikus zajmodell felállításánál, amely csak az egyik nézetben feltételez zajt. Mi a vizsgálatinkban ilyen megkötést nem szeretnénk alkalmazni, mivel távol áll a valós helyzetektől.

A $\mathbf{B}\mathbf{v}_9 = \lambda_9\mathbf{v}_9$ sajátérték-sajátvektor-probléma kis mértékben perturbált változata $(\mathbf{B} + \delta\mathbf{B})\tilde{\mathbf{v}}_9 = (\lambda_9 + \delta\lambda_9)\tilde{\mathbf{v}}_9$. Írjuk fel a számunkra fontos $\tilde{\mathbf{v}}_9 = \mathbf{v}_9 + \delta\mathbf{v}_9$ perturbált sajátvektort az eredeti sajátvektorok lineáris kombinációjaként

$$\tilde{\mathbf{v}}_9 = \mathbf{v}_9 + \sum_{i=1}^8 \alpha_i \mathbf{v}_i \quad (12)$$

alakban. Ekkor $\|\mathbf{v}_i\| = 1$ esetén $\|\tilde{\mathbf{v}}_9\|$ nem feltétlenül 1. A perturbált egyenlet

$$(\mathbf{B} + \delta\mathbf{B}) \left(\mathbf{v}_9 + \sum_{i=1}^8 \alpha_i \mathbf{v}_i \right) = (\lambda_9 + \delta\lambda_9) \left(\mathbf{v}_9 + \sum_{i=1}^8 \alpha_i \mathbf{v}_i \right) \quad (13)$$

A szorzásokat elvégezve, és figyelembe véve, hogy $\mathbf{B}\mathbf{v}_i = \lambda_i\mathbf{v}_i$ ($\forall i$)

$$\sum_{i=1}^8 \alpha_i \lambda_i \mathbf{v}_i + \delta\mathbf{B} \left(\mathbf{v}_9 + \sum_{i=1}^8 \alpha_i \mathbf{v}_i \right) = \lambda_9 \sum_{i=1}^8 \alpha_i \mathbf{v}_i + \delta\lambda_9 \mathbf{v}_9 + \delta\lambda_9 \sum_{i=1}^8 \alpha_i \mathbf{v}_i. \quad (14)$$

Balról szorozva \mathbf{v}_k^T -tal ($k \neq 9$), majd figyelembe véve, hogy $\mathbf{v}_k^T \mathbf{v}_i = \delta_{kj}$, ahol δ_{kj} a Kronecker-delta, és átrendezve: $\alpha_k(\lambda_9 + \delta\lambda_9 - \lambda_k) = \mathbf{v}_k^T \delta\mathbf{B}\tilde{\mathbf{v}}_9$, ahonnan

$$\alpha_k = \frac{\mathbf{v}_k^T \delta\mathbf{B}\tilde{\mathbf{v}}_9}{\tilde{\lambda}_9 - \lambda_k} \approx \frac{\mathbf{v}_k^T \delta\mathbf{B}\mathbf{v}_9}{\lambda_9 - \lambda_k}, \quad k = 1, 2, \dots, 8. \quad (15)$$

A közelítés $\delta\mathbf{v}_9 \approx 0$ és $\delta\lambda_9 \approx 0$ környékén jó, és azért szükséges, hogy a keresett fundamentális mátrix elemeit tartalmazó, perturbált $\tilde{\mathbf{v}}_9$ megoldásvektort az eredeti (zajmentes) sajátvektorokkal fejezhessük ki (12) szerint. Az eredményből látható, hogy a sajátvektorok $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_9$ sorrendje miatt a legnagyobb súlyt itt is a $d^{-1} = (\lambda_8 - \lambda_9)^{-1}$ tényező (a szeparáció inverze) jelenti.

A jövőben meg fogjuk vizsgálni, hogy ha az $\|\mathbf{E}\|$ perturbáció meghaladja a (7) szerinti $d/4$ küszöböt (vagyis nem alkalmazható az előző alfejezetben ismertett küszöb), de a perturbáció a (15) közelítés jóságát nem rontja nagyon el, akkor $\delta\mathbf{B}$ és a perturbálatlan \mathbf{B} mátrix szerkezete alapján mi állítható a probléma numerikus kondícionáltságáról a segítségül hívott approximáción keresztül.

4. A rang-kényszer fázisában jelentkező perturbáció

A rang- illetve szingularitási kényszer fázisában azt az \mathbf{F} mátrixot határozzuk meg, ami $\text{rg } \mathbf{F} = 2$ illetve $\det \mathbf{F} = 0$ mellett minimalizálja az

$$\|\mathbf{F} - \mathbf{F}_{min}\|_F^2 = \|\mathbf{f} - \mathbf{f}_{min}\|^2 = \sum_{i=1}^9 (f_i - f_{min,i})^2 \quad (16)$$

normát, ha \mathbf{F}_{min} az első fázisban meghatározott mátrix. Ez a módosítás egy újabb perturbációként jelentkezik az epipoláris kényszerben, és növeli az $\|\mathbf{A}\mathbf{f}\|^2$ algebrai hibát, amit az első fázisban minimalizáltunk. Kiemelnénk, hogy ez a fajta „perturbáció” teljesen más jellegű, így megkülönböztetendő az előző fázisban analizált, adatokban jelentkező zajtól.

Fontos észrevétel [9], hogy a fundamentális mátrix elemeiben jelentkező változásnak csak másodlagos szerepe van. Általában az az érdekes, hogy ez mennyire perturbálja az epipoláris egyeneseket és az epipólusokat. A fundamentális mátrix és az egyenesek kapcsolata $\mathbf{l}'_i = \mathbf{F}\mathbf{m}_i$ illetve $\mathbf{l}_i = \mathbf{F}^T \mathbf{m}'_i$. Az előbbi részletesen

$$\mathbf{l}'_i = \begin{pmatrix} a'_i \\ b'_i \\ c'_i \end{pmatrix} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = \begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_i & y_i & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & x_i & y_i \end{bmatrix} \mathbf{f} = \mathbf{Z}_i \mathbf{f}. \quad (17)$$

Hartley a fundamentális mátrix módosításának relatív hibáit tartja fontosnak: mivel a pontok koordinátái $(x, y, 1) = \mathcal{O}(100, 100, 1)$ nagyságrendűek átlagosan, ezért a bal felső 2×2 elem abszolútértéke jellemzően nagyságrendekkel kisebb, mint a többié. A (16) célfüggvény az eltéréseket minden elemnél azonosan bünteti, ami értelemszerűen nagyobb relatív hibát okozhat a kisebb abszolútértékű elemekben. Az epipoláris egyenesek számításánál épp ezek az elemek találkoznak a pontok $x \gg 1$ és $y \gg 1$ koordinátaival, amelyek átlagosan 2 nagyságrenddel

nagyobbak az egységnyi homogén koordinátánál. Ez a rang-kényszer által behozott perturbációt különböző mértékben viszi át az egyenesek $\mathbf{l}'_i = (a'_i, b'_i, c'_i)^T$ paramétereire. Hartley úgy érvel, hogy a normalizáló transzformációk azonos nagyságrendűvé teszik a koordinátákat, így várhatóan \mathbf{F} elemeiben sem jelentkeznek a nagyságrendi eltérések, és az $\mathbf{F}\mathbf{m}_i$ szorzat is azonos mértékben viszi a perturbáció hatását át az egyenes paramétereire. Ez csak szemléltető magyarázat, az állítások teljesülésére nincs garancia: \mathbf{F} elemei akár nagyságrendileg is eltérhetnek, az egyenesek paramétereire pedig eltérő mértékben mozoghatnak.

Ugyanakkor az egyenes paramétereinek azonos mértékű relatív megváltozása nem ugyanolyan megváltozást okoz az egyenes – képen elfoglalt – helyzetében, hiszen $(a'_i, b'_i)^T$ az egyenes normálisa, c'_i pedig az eltolása. Számunkra az a cél, hogy az egyenes helyzete minél kisebb mértékben változzon a rang-kényszerítés hatására. Racionális ezt úgy megfogalmazni, hogy az első lépésben optimalizált

$$\|\mathbf{A}\mathbf{f}\|^2 = \sum_i r_i^2 = \sum_{i=1}^n (\mathbf{m}_i'^T \mathbf{l}'_i)^2 = \sum_{i=1}^n (\mathbf{m}_i'^T \mathbf{F}\mathbf{m}_i)^2 \quad (18)$$

algebrai hiba maradjon a lehetőségekhez képest kicsi. Még kifejezőbb, ha a detektált pontok és az epipoláris egyenesek távolságát, azaz a geometriai hibát mérjük, vagyis az \mathbf{m}_i pont és az \mathbf{l}_i egyenes illetve az \mathbf{m}'_i pont és \mathbf{l}'_i egyenes közötti távolságot mérjük és összesítjük:

$$\sum_{i=1}^n \{d^2(\mathbf{m}_i, \mathbf{l}_i) + d^2(\mathbf{m}'_i, \mathbf{l}'_i)\} = \sum_{i=1}^n \left\{ \frac{(\mathbf{m}_i^T \mathbf{l}_i)^2}{a_i^2 + b_i^2} + \frac{(\mathbf{m}'_i{}^T \mathbf{l}'_i)^2}{a_i'^2 + b_i'^2} \right\} = \sum_{i=1}^n w_i^2 r_i^2, \quad (19)$$

ahol $w_i^2(\mathbf{f}) = (a_i^2 + b_i^2)^{-1} + (a_i'^2 + b_i'^2)^{-1}$. Eszerint a geometriai hibatagokban megjelennek az algebrai hibatagok is [5]. Az összegzésben az összes rendelkezésre álló pontpár hibájának szerepelnie kell, nem csak azoké, amelyekre a modellt illesztettük. Eszerint nem az illesztési hibát, hanem a modell és az összes rendelkezésre álló adat konzisztenciáját mérjük.

Parciális deriválással megállapítható, hogy a megfogalmazott kritériumfüggvényeket sem a fundamentális mátrix elemeinek, sem az egyenesek homogén koordinátáinak a relatív megváltozásai nem ugyanolyan mértékben befolyásolják. Mivel elsősorban a geometriai hibának van valódi, gyakorlati jelentősége, ezért az előbbi megállapítás szerint nem mérvadó, hogy azonos mértékű relatív megváltozást érzünk el az epipoláris egyenesek koordinátáiban vagy a fundamentális mátrix elemeiben.

Fontos megjegyezni, hogy a legtöbb, komplexebb módszer is úgy próbálja iteratíván minimalizálni valamelyik hibamértéket, hogy közben folyamatosan kikényszeríti a $\det \mathbf{F} = 0$ feltételt [5].

Az eredmény jóságát a (18) vagy (19) értelemben a \mathbf{H} és \mathbf{H}' transzformációk megválasztásával befolyásolhatjuk: megválaszthatjuk a koordináta-rendszert, amelybe transzformálva az első fázisban meghatározott \mathbf{F}_{min} mátrixot, a

$$\|\mathbf{H}'^{-T} \mathbf{F}_{min} \mathbf{H}^{-1} - \tilde{\mathbf{F}}\|_F \rightarrow \min_{\tilde{\mathbf{F}}}, \quad \det \tilde{\mathbf{F}} = 0 \quad (20)$$

minimalizálási problémát megoldjuk a 2. fejezetben leírt, SVD-alapú módszerrel. Az eredményezett fundamentális mátrixot vissza kell transzformálni az eredeti koordináta-rendszerbe $\mathbf{F} = \mathbf{H}'^T \tilde{\mathbf{F}} \mathbf{H}$ szerint. Ez a transzformáció a mátrix rangját nem rontja el, mert \mathbf{H}' és \mathbf{H} nonsinguláris mátrixok.

Összefoglalva, célszerű aszerint normalizálni a rang-kényszer fázisában, hogy a fundamentális mátrix egyes paraméterei minimális mértékben módosítsák az egyenes helyzetét, és ezzel együtt az első fázis célját, a már minimalizált algebrai hibát, vagy az egyenes eltávolodását geometriailag jobban kifejező geometriai hibák összegét. Mindezt ahelyett, hogy – Hartley érvelése szerint – az egyenes homogén koordinátáinak azonos mértékű perturbációjára törekednénk.

5. Normalizáló transzformációk

Az „optimális” normalizáció meghatározásánál azt a \mathbf{H}_1 és \mathbf{H}'_1 homográfiát keressük, amelyekkel az első fázisban maximális mértékben javítható a probléma numerikus kondicionáltsága, és ezzel együtt a sajátvektor-probléma stabilitása a pontok koordinátaiban jelentkező zajra nézve. Keressük azt a \mathbf{H}_2 és \mathbf{H}'_2 homográfiát is, amelyekkel a második fázisban jelentkező perturbáció hatása az első fázisban minimalizált algebrai hibát minimális mértékben növeli. Nem garantált, hogy $\mathbf{H}_1 = \mathbf{H}_2$ és $\mathbf{H}'_1 = \mathbf{H}'_2$, ezért a vizsgálatokban ésszerű külön kezelni a két fázist. Ez a két fázis között úgy értendő, hogy az első fázis végeredményét visszatranszformáljuk az eredeti koordináta-rendszerbe \mathbf{H}_1 és \mathbf{H}'_1 felhasználásával, majd alkalmazzuk a \mathbf{H}_2 és \mathbf{H}'_2 transzformációkat (20) szerint.

Az értelemszerű, hogy a transzformációkat a pontrendszerek jellemzőiből vezetjük le, de nem kizárható, hogy az optimális transzformációk az egyes esetekben függenek az epipoláris elrendezéstől és így a fundamentális mátrixtól is.

A keresett $\{\mathbf{H}_i, \mathbf{H}'_i\}$ transzformációk ($i = 1, 2$) alapesetben általános, 8 szabadságfokú homográfiák. Ahogy meg fogjuk mutatni, tetszőleges pontrendszer első- és másodrendű momentuma tetszés szerint befolyásolható már egy affin transzformációval is, noha ez nem garancia arra, hogy magasabb fokszámú modellel nem lehet jobb eredményt elérni. [15] már az általános, affin transzformáció, valamint az eltolás utáni izotróp skálázás között sem tapasztalt számottevő különbséget az eredményben, ezért első vizsgálatainkban mi sem alkalmazunk az affinnál magasabb szabadságfokú transzformációt. A transzformációkat ezért a

$$\mathbf{H}_A = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{S}^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}^T & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \quad (21)$$

alakban keressük, ahol $\mathbf{S} = \text{diag}(s_1, s_2)$ és $s_1 \geq s_2$. Belátható, hogy tetszőleges affin transzformációnak létezik ilyen dekompozíciója [7].

Az általános esethez tekintsük azt az ellipszist, amely egy origó középpontú, egységnyi tömegű, egyenletes tömegeloszlású, egységsugarú körből keletkezett a $\mathbf{T} = \mathbf{R}(\alpha)\mathbf{S}$ transzformációval. (Az ellipszis már nem egyenletes tömegeloszlású.)

Integrálással belátható, hogy az ellipszis súlypontja is az origóban van, másodrendű momentumainak 2×2 -es mátrixa pedig $\mathbf{J}^e = \frac{1}{2}\mathbf{R}\mathbf{S}^2\mathbf{R}^T$. Ha \mathbf{J}^e adott, akkor az ellipszis paraméterei kiolvashatók a \mathbf{J}^e pozitív szemidefinit mátrix $\mathbf{J}^e = \mathbf{U} \text{diag}(\mu_1, \mu_2)\mathbf{U}^T$ spektrálfelbontásából, ahol \mathbf{U} ortonormált mátrix. $\det \mathbf{U} = 1$ és $\mu_1 \geq \mu_2$ pedig elemi transzformációkkal biztosítható. Ezek után $\mathbf{R}(\alpha) \equiv \mathbf{U}$ és $\frac{1}{2}\mathbf{S}^2 \equiv \text{diag}(\mu_1, \mu_2)$, vagyis a α kiolvasható az $\mathbf{R}(\alpha)$ forgatómátrixból, az ellipszis főtengeley-irányú megnyúlásai pedig $s_i = \sqrt{2\mu_i}$ ($i = 1, 2$).

Tekintsük most a $\mathbf{p}_i = (x_i, y_i)^T$ pontokból álló pontrendszert ($i = 1, 2, \dots, n$). Ennek elsőrendű momentumuma és másodrendű momentumainak mátrixa rendre

$$\mathbf{p}_g = \frac{1}{n} \sum_{i=1}^n \mathbf{p}_i = \bar{\mathbf{p}}, \quad \mathbf{J} = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix} = \frac{1}{n} \sum_{i=1}^n (\mathbf{p}_i - \bar{\mathbf{p}})(\mathbf{p}_i - \bar{\mathbf{p}})^T. \quad (22)$$

Fizikai szemlélettel I_{xx} és I_{yy} a súlyponton keresztülhaladó, az x - illetve y - tengellyel párhuzamos tengelyekre vonatkoztatott tehetetlenségi nyomatékok. Helyettesítsük a pontrendszert egy olyan ellipszissel, amit az előbb tárgyalt módon származtatunk egy körből, de tömegközéppontja $\bar{\mathbf{p}}$, másodrendű momentumainak mátrixa pedig $\mathbf{J}^e \equiv \mathbf{J}$. Az \mathbf{S} és \mathbf{R} mátrixok \mathbf{J}^e -ből kiolvashatók az előbb tárgyalt módon. Ezt az ellipszist a ρ sugarú, origó középpontú körbe a $\tilde{\mathbf{p}} = \rho\mathbf{S}^{-1}\mathbf{R}^T(\mathbf{p} - \bar{\mathbf{p}})$ affin transzformáció viszi. Ugyanezt a transzformációt alkalmazva a \mathbf{p}_i pontokra, a transzformált $\tilde{\mathbf{p}}_i$ pontok elsőrendű momentumuma az origó, másodrendű momentumainak mátrixa, $\frac{\rho^2}{2}\mathbf{I}$, a súlypontra vonatkoztatott tehetetlenségi nyomaték pedig $\text{trace } \mathbf{J} = \rho^2$.

A [7] által javasolt (3) transzformáció ennek analógiájára egy eltolás és egy izotróp skálázás $\tilde{\mathbf{p}} = \frac{\rho}{r}(\mathbf{p} - \bar{\mathbf{p}})$ szerint, vagyis $\alpha = 0$ és $s_1 = s_2 = \rho/r$. Ezzel az ekvivalens, egyenletes tömegeloszlású, $\bar{\mathbf{p}}$ origójú, r^2 tehetetlenségi nyomatékú kört az origó középpontú, ρ sugarú körbe transzformáljuk, amelynek másodrendű momentumuma ρ^2 , hasonlóan a transzformált pontrendszer másodrendű momentumához. A transzformációban r a pontok RMS-távolsága a súlyponttól:

$$r = \left\{ \frac{1}{n} \sum_i \|\mathbf{p}_i - \bar{\mathbf{p}}\|^2 \right\}^{\frac{1}{2}}. \quad (23)$$

A pontrendszer RMS-távolságát az origótól [7] $\sqrt{2}$ -re állítja (az ekvivalens kör sugara $\rho = \sqrt{2}$), míg [9] a pontok átlagos távolságát normalizálja $\sqrt{2}$ -re, noha egyik megoldás optimális volta sem igazolt. Hartley [15] az anizotróp skálázást is vizsgálja, de a (21) paraméterezést nem használja. Az itt bemutatott, ekvivalens ellipszist használó módszer helyett egy Cholesky-dekompozíció alapú módszerrel határozza meg az affin transzformációs mátrix elemeit. Az általunk felírt paraméterezés előnye, hogy az elforgatás szöge explicit módon megjelenik, így közvetlen, geometriai jelentéssel bíró paramétereket használhatunk az ideális transzformációk megkeresésekor, amelyek így összefüggésbe hozhatók a pontfelhő ferdeségének szögével. A keresést a (18) vagy a (19) célfüggvény minimalizálásával érhetjük el a transzformációk paraméterei szerint.

[15] az izotróp és anizotróp skálázás közötti elenyésző különbséget egyetlen esetre igazolja, amelynél a pontfelhő alakjáról nincs információ. Így nem zárható

ki, hogy más elrendezések vagy a bemeneti pontfelhők nagyobb mértékű aszimmetriája esetén szerephez jut az affin normalizáló transzformáció is. A jövőben ezt is szeretnénk megvizsgálni.

6. Összefoglalás

A fundamentális mátrix kiszámítása kulcskérdés a számítógépes látásban, és egyik legalapvetőbb módszere a 8-pontos algoritmus. Ennek első fázisa egy homogén legkisebb-négyzetes probléma, ami egy sajátérték-sajátvektor problémára vezethető vissza, második fázisa pedig egy korrekció, amellyel a fundamentális mátrix szingularitását kényszerítjük ki. A második fázisban elvégzett korrekció tipikusan növeli az első fázisban minimalizált algebrai hibát.

A szakirodalom szerint a 8-pontos algoritmus alapesetben még nagy pontszám és kis lokalizációs hibák esetén is meglehetősen instabil eredményeket ad, ezért normalizáló transzformációk alkalmazása szükséges, amelyekkel *átlagosan* kisebb illesztési hibát érhetünk el [15]. Arra, hogy a normalizált algoritmus miért ad jobb eredményt, mint a normalizálatlan változat, Hartley [9],[15] adott egy magyarázatot.

Hartley egy olyan kondíciószámra hozza összefüggésbe az első fázis kondicionáltságát, ami matematikailag egy más típusú probléma, az inhomogén egyenletrendszer kondicionáltságát jellemzi. Ezért megmutattuk, hogy a Hartley-féle kondíciószám helyett a sajátérték-sajátvektor probléma perturbációanalízise milyen matematikai eszközöket kínál az első fázis numerikus kondicionáltságának jellemzésére. A jövőben numerikusan is meg fogjuk vizsgálni az egyes módszerek alkalmazhatósági körét bizonyos tipikus geometriai elrendezések mellett.

Emellett rámutattunk arra, hogy a második fázis érzékenységének magyarázatánál a fundamentális mátrix elemeinek azonos mértékű relatív perturbációja helyett célszerű az első fázisban minimalizált algebrai hibákat vagy az epipoláris egyenesek elmozdulását jobban jellemző geometriai hibákat minimalizálni a normalizáló transzformációk megválasztásánál. Mivel a két fázisban eltérő cél szerint is dolgozhatunk, más és más normalizáló transzformációk lehetnek optimálisak. Ezért a vizsgálatokban célszerű lehet az első fázisban kapott eredményt visszatranszformálni az eredeti koordináta-rendszerbe, majd a második fázis szempontjából előnyös transzformációval újra normalizálni.

A jövőben arra keressük a választ az itt bemutatott matematikai eszközök és mérési módszerek segítségével, hogy bizonyos esetekben és az egyes fázisokban mennyire és miért „jó” a Hartley által javasolt normalizáció, van-e ennél jobb, illetve mennyire és hogyan függ ez az elrendezéstől.

Irodalom

1. H. C. Longuet-Higgins, “A computer algorithm for reconstructing a scene from two projections,” *Nature*, 293:133–135, Szeptember 1981.
2. R. I. Hartley, “Estimation of relative camera positions for uncalibrated cameras,” in *Proc. European Conference on Computer Vision, LNCS 588*, pp. 579–587. Springer-Verlag, 1992.

3. O. D. Faugeras, "What can be seen in three dimensions with an uncalibrated stereo rig?," in *Proc. European Conference on Computer Vision, LNCS 588*, pp. 563–578. Springer-Verlag, 1992.
4. Q.-T. Luong and O. Faugeras, "The fundamental matrix: theory, algorithms, and stability analysis," *International Journal of Computer Vision*, 17:43–75, 1996.
5. Z. Zhang, "Determining the epipolar geometry and its uncertainty - a review," *International Journal of Computer Vision*, 27(2):161–195, March 1998.
6. X. Armangué and J. Salvi, "Overall view regarding fundamental matrix estimation," *Image and Vision Computing*, 21:205–220, 2003.
7. R. Hartley and A. Zissermann, *Multiple View Geometry in Computer Vision, Second Edition*, Cambridge University Press, 2006.
8. K. Kanatani and Y. Sugaya, "High accuracy fundamental matrix computation and its performance evaluation," in *Proc. 17th British Machine Vision Conf (BMVC 2006), vol.1*, pp. 217–226, 2006.
9. R. I. Hartley, "In defence of the 8-point algorithm," in *Proc. International Conference of Computer Vision*, pp. 1064–1070, 1995.
10. P. H. S. Torr and A. W. Fitzgibbon, "Invariant fitting of two view geometry," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(5):648–650, 2004.
11. W. Chojnacki, M. J. Brooks, A. van den Hengel, and D. Gawley, "Revisiting hartley's normalized eight-point algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(9):1172–1177, 2003.
12. G. H. Golub, , and C. F. van Loan, *Matrix Computations, Third Edition*, The John Hopkins University Press, 1996.
13. P. H. S. Torr, *Motion segmentation and outlier detection*, Ph.D. thesis, Department of Engineering Science, University of Oxford, 1995.
14. M. Müttlich and R. Mester, "The role of total least squares in motion analysis," in *Proc. European Conference of Computer Vision*, pp. 305–321, 1998.
15. R. I. Hartley, "In defense of the eight-point algorithm," *IEEE Trans. Pattern on Pattern Analysis and Machine Intelligence*, 19(6):580–593, 1997.