

Körbelátó kamerák
(oktatási segédanyag)

Kató Zoltán

2014 október

Tartalomjegyzék

1. A körbelátó kamera geometriai modellezése	7
1.1. A körbelátó optikákról	7
1.2. A katadioptrikus optika	9
1.3. A halszem optika	13
2. A körbelátó kamera kalibrációja	15
2.1. Körbelátó kamerák egységes Taylor modellje	16
2.2. Scaramuzza kalibrációs algoritmus	17

Az oktatási segédanyag az Európai Unió és Magyarország támogatásával, az Európai Szociális Alap társfinanszírozásával a TÁMOP 4.2.4.A/2-11-1-2012-0001 azonosító számú „Nemzeti Kiválóság Program” című kiemelt projekt keretei között valósult meg.

Előszó

A számítógépes látás az emberi látás azon funkcióit valósítja meg, amelyek a retinai kép elemzését végzik. A szakirodalomban ezen feladatok megoldására elsősorban hagyományos perspektív kamerákat használnak. Az utóbbi időben azonban széleskörűen elterjedtek a körbelátó panoramikus (vagy omnidirekcionális) kamerák. Ma már nem csak speciális ipari célra alkalmazzák ezeket a kamerákat, hanem az autóipar vezetésköz támogatási rendszerei is előszeretettel használják ezeket a vizuális szenzorokat főként navigációs célokra. A körbelátó kamerák változatos optikai konstrukcióban készülnek, de közös jellemzőjük, hogy a kamera látószöge jellemzően $180^\circ \times 360^\circ$. Ez óriási előny navigációs feladatok esetében, ugyanis a jármű egyetlen kamera által egy időben minden irányba képes „látni”. Bár a körbelátó kamerák geometriai modellje lényegesen eltér a perspektív kamerákétól, mégis hasonló módon, speciális kalibrációs sémákat használva megoldható a kamera pontos kalibrációja, mely a legfontosabb kezdő lépés a kamerák valós alkalmazásához. Jelen segédanyag célja a körbelátó kamerák általános geometriai modellezésének bemutatása illetve ezen modellek alapján a kamera belső paramétereinek meghatározása, vagyis a kamera kalibrációja. A tematika összeállításában nagy segítséget jelentett Christopher Geyer, Tomas Pajdla és Kostas Daniilidis összefoglalója [1] valamint Davide Scaramuzza PhD disszertációja [2]. A munka elkészítéséhez nagyban hozzájárult a témában folytatott közös kutatómunka Tamás Leventével (Kolozsvári Műszaki Egyetem) és Fröhlich Róberttel (Szegedi Tudományegyetem) [3,4]. A jelen munkában bemutatott tematika szerves részét képezi a Szegedi Tudományegyetemen tartott *Ipari Képfeldolgozás* MSc kurzusom anyagának.

Tudomásunk szerint jegyzetünk az első magyar nyelvű munka ebben a témakörben, ezért igyekeztünk az angol szakkifejezések minél kifejezőbb magyar megfelelőit is megtalálni. Reméljük, hogy jegyzetünket nem csak egyetemi vagy főiskolai hallgatók olvassák majd, hanem a műszaki fejlesztésekben részt vevő mérnökök, programozók is hasznosnak találják a bemutatott ismereteket.

Szeged, 2014 október 28.

Kató Zoltán *Szegedi Tudományegyetem*

1. fejezet

A körbelátó kamera geometriai modellezése

Egy közönséges digitális kamera felfogható egy olyan eszközként, amely a látóterében lévő 3D valós *látványról* egy 2D *képet* készít, vagyis matematikai értelemben nem más, mint egy 3D→2D leképezés. A legegyszerűbb ilyen kamera a lyukkamera (latinul *camera obscura*), amely egy zárt dobozból (az un. sötétkamrából) áll, melynek egyik oldalán egy kis méretű lyuk található. A lyukon keresztül a fénysugarak bejutnak a dobozba és a lyukkal átellenes oldalra vetítik a kamera által látott *látvány* fordított képét (ld. 1.1 ábra). A lyukkamera matematikai modelljét (1.2 ábra) használják a közönséges digitális fényképezőgépek és kamerák geometriai modellezésére [5]. Az ilyen kamerák látószöge általában még *nagylátószögű optika* alkalmazása esetén sem haladja meg a 100° -t.

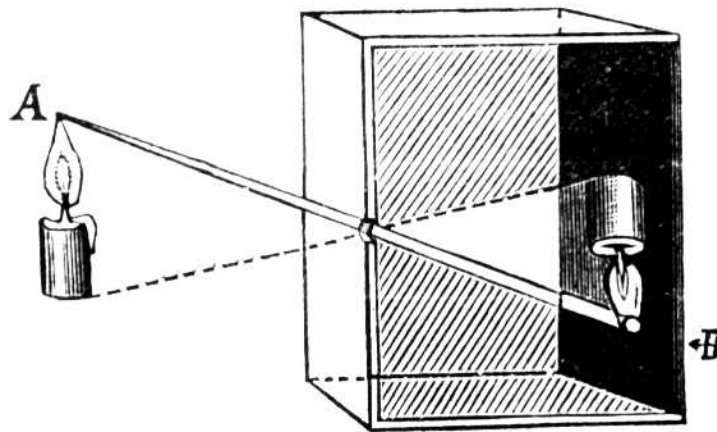
Jegyzetünkben olyan kamerákkal foglalkozunk, ahol ez a leképezés továbbra is *középpontos vetítés* formájában valósul meg, de a kamera látótere a szokványos perspektív kamerákat jelentősen meghaladó $270^\circ - 360^\circ$ tartományba esik és a leképezés megvalósításához speciális optikákat alkalmaznak. Ezeket a kamerákat *körbelátó* vagy *panoramikus* kameráknak is nevezik (angolul az *omnidirectional camera* kifejezés terjedt el) [6].

1.1. A körbelátó optikákról

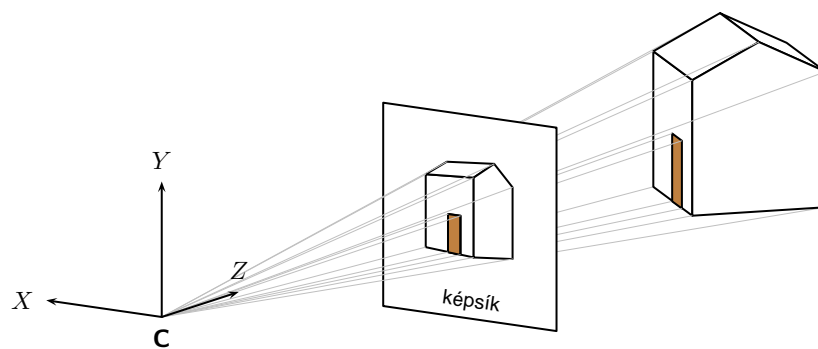
A körbelátó kamerák szerteágazó konstrukciós elvek szerint készülnek. Konstrukció szerint két nagyobb csoportra bonthatók a jelenleg elterjedt panoramikus kamerák:

1. egyetlen kamerából és egy speciális optikából álló rendszer
2. több hagyományos kamerából álló, rögzített elrendezésben összeszerelt kamera-rendszer (1.3 ábra)

A 2) esetre mutat példát az 1.3 ábra. Ezeknek a kameráknak nagy előnye, hogy szokványos perspektív kamerák képeiből állítják elő a panorámát, így a kamerák felbontása



1.1. ábra. Camera obscura (Fizyka, 1910)



1.2. ábra. A lyukkamera geometriai modellje: A képsíkot a **C** fókuszpont elé helyezve oldalhelyes képet kapunk.



1.3. ábra. PointGrey LADYBUG5 körbelátó kamera: 30 MegaPixel felbontás, a teljes gömb 90%/t lefedő kép (Forrás: <http://ww2.ptgrey.com/ld5ms/>).

maximálisan kihasználható és az egyes kameraképek torzításmentesek maradnak, ráadásul a hagyományos perspektív kamerákra alkalmazható módszerek változtatás nélkül használhatók. A konstrukció hátránya, hogy – mivel több kamerát is be kell építeni – drága és nagy méretű.

Jegyzetünkben az 1) esettel foglalkozunk részletesen. Ezeknek a kameráknak nagy előnye, hogy egyetlen kameraképben áll elő a panorámakép, így olcsó és kompakt eszközök állíthatók elő. Hátránya azonban, hogy az így kapott kép egy az optika leképezésétől függő nemlineáris torzítással áll elő, tehát ennek elemzése és feldolgozása új algoritmusokat igényel. A torzítás miatt a pixelek a valóságban más és más méretű felületelemnek felelnek meg, így a kamera felbontása sem egyenletes, ami megnehezíti például a pontos méréseket. Erre látunk példákat különböző optikák esetében az 1.4 ábrán. Ezeket a kamerákat elsősorban földi és légi mobil robotokon alkalmazzák vizuális navigációhoz valamint odometriához.

1.2. A katadioptrikus optika

A katadioptrikus optika egy speciális tükör és egy hagyományos perspektív kamerából álló képalkotó eszköz. A képalkotás során a tükör felszínén kialakult képet fényképezük le a projektív kamerával (ld. 1.5 ábra). Ez a leképező rendszer általános esetben nem centrális vetítést valósít meg. Belátható azonban, hogy kizárólag akkor lesz középpontos a vetítés, ha

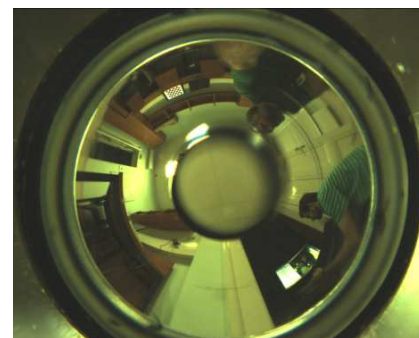
- a tükör keresztmetszete kúpszelet
- a tükörfelület a kúpszelet szimmetriatengelye körüli forgatásaként áll elő
- a perspektív kamera a tükör egyik fókuszában van



Katadioptrikus optika

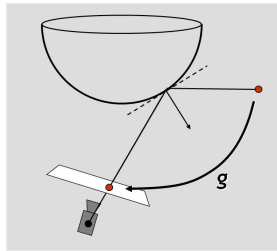


Halszem optika

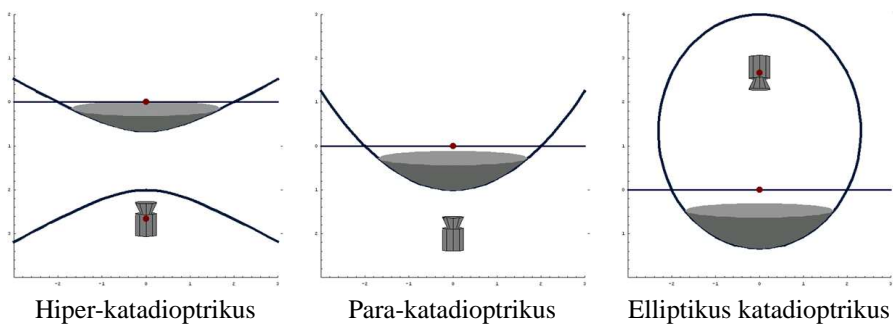


PAL optika

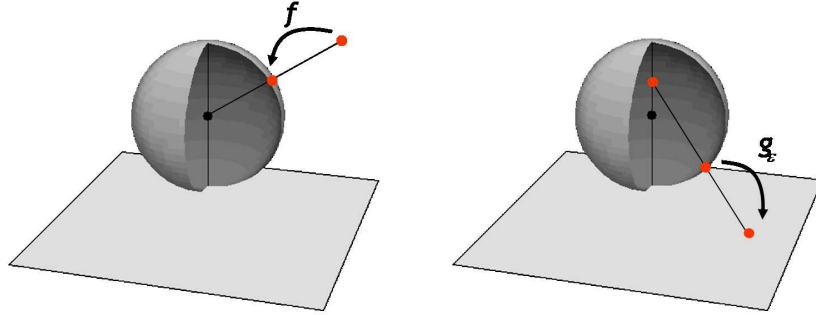
1.4. ábra. Különböző körbelátó optikák és az általuk készített kép.



1.5. ábra. Centrális katadioptrikus optika leképezése: A vetítésugarak a tükörről visszaverődve jutnak a kamerába, amely összességében a g -vel jelölt nemlineáris leképezést valósítja meg.



1.6. ábra. Centrális vetítést megvalósító katadioptrikus optikák.



1.7. ábra. Centrális katadioptrikus kamerák egységes szferikus modellje [7].

Ekkor a tükör másik fókuszpontja lesz a katadioptrikus kamera effektív vetítési középpontja [7]. Ilyen konstrukciókat mutat be az 1.6 ábra. Ezek közül a gyakorlatban leginkább a *hiper-katadioptrikus* (hiperbolikus tükör és perspektív kamera kombinációja) illetve a *para-katadioptrikus* (parabolikus tükör és telecentrikus kamera kombinációja) konstrukciók terjedtek el.

A centrális katadioptrikus és hagyományos perspektív kamerák egységes vetítési modelljét Geyer és Daniilidis dolgozta ki [7]. Az elmélet szerint minden katadioptrikus (para-, hiper-, elliptikus-) és standard projektív vetítés izomorf egy olyan gömb \rightarrow sík leképezéssel, ahol a gömb középpontja az effektív vetítési középpontban van, a síkvetítés középpontja pedig a gömb középpontján átmenő síkra merőleges tengelyen (ld. 1.7 ábra). Maga a leképezés pedig két vetítés kompozíciójaként definiálható:

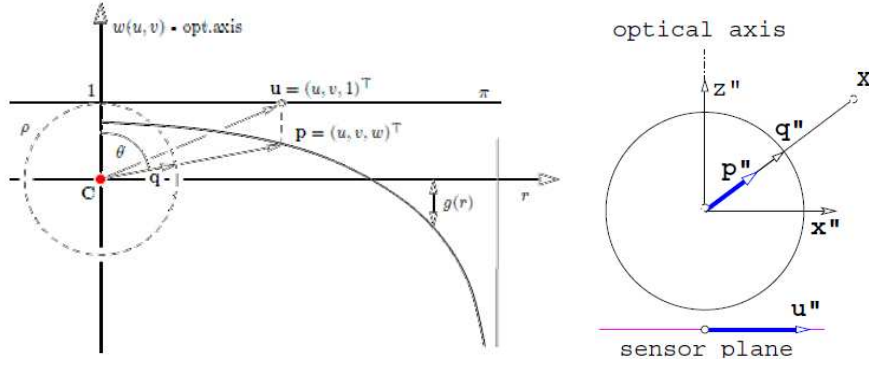
f az egységgömb felszínére vetít. A vetítési középpont a gömb középpontja, a vetítés eredménye egy olyan homogén pont, melynek 4. koordinátája a vetített 3D \mathbf{X} pont távolsága lesz. A megfelelő inhomogén vektor tehát a vetített pont irányába mutató egységvektornak felel meg:

$$f(\mathbf{X}) = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \sqrt{X_1^2 + X_2^2 + X_3^2} \end{pmatrix} \quad (1.1)$$

g_ϵ középpontos vetítés a gömb tengelyének egy (az ϵ excentricitás által meghatározott) pontjából a gömb tengelyére merőleges képsíkra:

$$g_\epsilon = \begin{pmatrix} 2\epsilon & & & \\ & 2\epsilon & & \\ & & -(1+\epsilon)^2 & \\ & & & -2\epsilon \end{pmatrix} \quad (1.2)$$

Vegyük észre, hogy az f által adott pontot az alsó félgömbre tükrözi! A \mathbf{C}_ϵ vetítési középpontra teljesül a $g_\epsilon \mathbf{C}_\epsilon = 0$ összefüggés, amiből \mathbf{C}_ϵ -t az alábbi



1.8. ábra. Centrális halszem és katadioptrikus kamerák egysége szferikus modellje [8, 9].

formában kapjuk:

$$\mathbf{C}_\epsilon = \begin{pmatrix} 0 \\ 0 \\ \frac{2\epsilon}{1+\epsilon^2} \\ 1 \end{pmatrix} \quad (1.3)$$

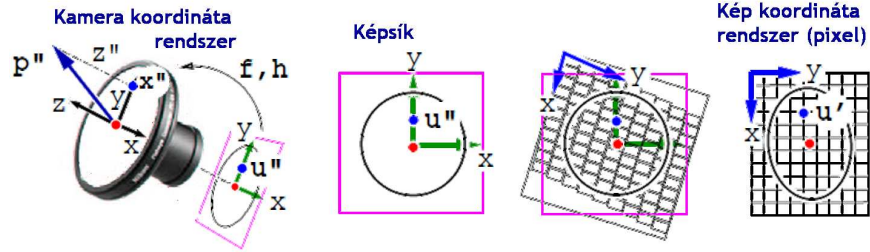
A katadioptrikus leképezés tehát $g_\epsilon \circ f$ alakban írható, melyben f nemlineáris de független a leképezés ϵ paraméterétől, míg g_ϵ ugyan függ ϵ -től, de lineáris:

$$\begin{aligned} g_\epsilon \circ f(\mathbf{X}) &= \begin{pmatrix} 2\epsilon & & & \\ & 2\epsilon & & \\ & & -(1+\epsilon)^2 & -2\epsilon \\ & & & -2\epsilon \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \sqrt{X_1^2 + X_2^2 + X_3^2} \end{pmatrix} \\ &= \begin{pmatrix} X_1 \\ X_2 \\ -(1+\epsilon)^2 X_3 - 2\epsilon \sqrt{X_1^2 + X_2^2 + X_3^2} \end{pmatrix} \quad (1.4) \end{aligned}$$

Vegyük észre, hogy $\epsilon = 0$ esetén a jól ismert perspektív kamera leképezését kapjuk (a képsík a vetítési középpont mögött van).

1.3. A halszem optika

A halszem optika tükör nélkül, speciális lencsetagokkal készül, amellyel a katadioptrikus optikákhoz hasonló látószöget lehet elérni. A leképezés egységes modelljét Micusik és Pajdla dolgozta ki [8, 9]. A képsík egy pontját reprezentálhatjuk mint a pontból húzott vetítősugar metszéspontját egy \mathcal{S} egységgömb retinán (ld. 1.8 ábra): $\mathbf{q} = \mathbf{P}\mathbf{X}$, ahol \mathbf{P} 3×4 vetítómátrix és \mathbf{X} valós pont. A \mathbf{q} szferikus pontnak megfelelő \mathbf{u} képpontot egy f nemlineáris leképezéssel kapjuk $\mathbf{q} \rightarrow \mathbf{p} \rightarrow \mathbf{u}$. Halszemoptikák esetén a leggyakoribb az ekvidisztáns ill. ekvianguláris leképezés ahol a képpont radiális távolsága arányos a vetítési szöggel.



1.9. ábra. A kamera- és a kép koordináta rendszerek közötti kapcsolat. [8].

Micusik és Pajdla modellje alkalmazható mind katadioptrikus mind halszemoptikákra, feltéve hogy a vetítés centrális és tengelyesen szimmetrikus, valamint a képsík merőleges a tengelyre. Egy \mathbf{X} 3D pontnak megfelelő szferikus \mathbf{q}'' képpontot, amelyet a vetítősugarának \mathbf{p}'' (nem feltétlenül egységnyi) irányvektorával reprezentálunk, egy \mathbf{u}'' pontba vetít a képsíkon:

$$\mathbf{p}'' = \begin{pmatrix} h(\mathbf{u}'')\mathbf{u}'' \\ f(\mathbf{u}'') \end{pmatrix} \tag{1.5}$$

ahol f és h forgásszimmetrikus függvények, azaz egy tetszőleges \mathbf{R} forgatásra $h(\mathbf{R}\mathbf{u}'') = h(\mathbf{u}'')$ és $f(\mathbf{R}\mathbf{u}'') = f(\mathbf{u}'')$. h a perspektív vetítést írja le, ami halszemoptikák esetén $h = 1$, azaz ortografikus vetítés, f pedig a tükör alakjától függ. Az alábbi két képlet egy katadioptrikus és egy halszem optika vetítési függvényére ad példát [8]:

Para-katadioptrikus optika	Nikon FC-E8 halszem optika
$\mathbf{p}'' = \begin{pmatrix} 1\mathbf{u}'' \\ \frac{a''^2 - \ \mathbf{u}''\ ^2}{2a''} \end{pmatrix}$	$\mathbf{p}'' = \begin{pmatrix} 1\mathbf{u}'' \\ \frac{\ \mathbf{u}''\ }{\tan \frac{a''\ \mathbf{u}''\ }{1+b''\ \mathbf{u}''\ }} \end{pmatrix}$

ahol a'', b'' a leképező rendszer fizikai paraméterei. Megjegyezzük, hogy ez a modell a perspektív vetítést is magában foglalja: $h = f = 1$.

Az eddig tárgyalt vetítés a kamera 3D koordináta rendszeréből képez le a kamera 2D képsíkjára. Ez azonban a gyártási pontatlanságok miatt a valóságban nem feltétlenül esik egybe a szenzor síkjával. Ezt kompenzálendő, Micusik és Pajdla modelljében [8, 9] a képsík koordináta rendszerből a szenzor pixel koordináta rendszerébe egy affin leképezés hat, vagyis

$$\mathbf{u}'' = \mathbf{A}\mathbf{u}' + \mathbf{t}, \tag{1.6}$$

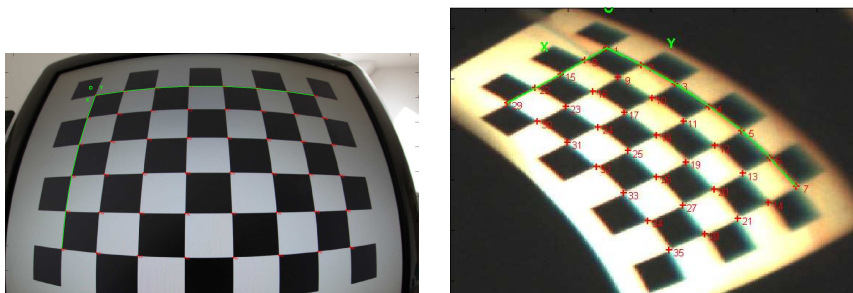
ahol \mathbf{u}' a pixel koordináta rendszerbeli pont, (\mathbf{A}, \mathbf{t}) pedig a pixel koordináta rendszerből a kép koordináta rendszerbe képző affin transzformáció. Tehát az 1.5 képlet az alábbi alakban írható:

$$\mathbf{p}'' = \begin{pmatrix} h(\mathbf{u}'')\mathbf{u}'' \\ f(\mathbf{u}'') \end{pmatrix} = \begin{pmatrix} h(\mathbf{A}\mathbf{u}' + \mathbf{t})(\mathbf{A}\mathbf{u}' + \mathbf{t}) \\ f(\mathbf{A}\mathbf{u}' + \mathbf{t}) \end{pmatrix} \tag{1.7}$$

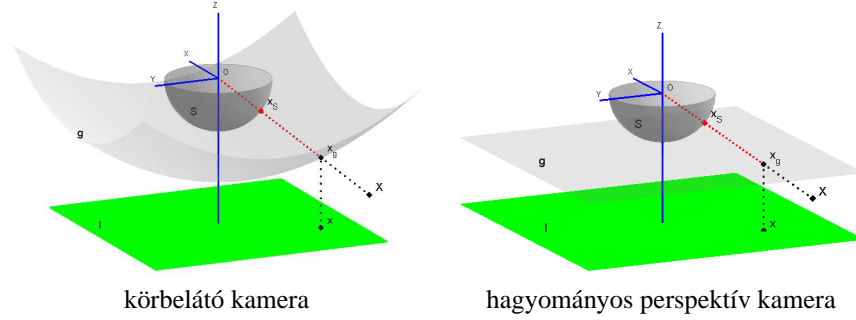
2. fejezet

A körbelátó kamera kalibrációja

Az előzőekben láttuk, hogy egy kamera a leképezési függvényével egyértelműen reprezentálható. Míg a perspektív kamerák esetében ez a kamera vetítési mátrixa [5], addig a körbelátó kamerák esetében többféle geometriai modell is lehetséges. Bármelyik modellt is használjuk azonban, ha ismert a kamera vetítő függvénye (azaz annak paraméterei), akkor egy tetszőleges pont képét elő tudjuk állítani, illetve bármely képpontot vissza is tudunk vetíteni. A gyakorlatban tehát fontos, hogy meg tudjuk határozni a kamera paramétereit. Ezt a folyamatot *kamera kalibrációnak* nevezzük. Ebben a fejezetben egy olyan kalibrációs eljárást fogunk bemutatni, ahol elegendő számú *pontmegfeleltetés* áll rendelkezésünkre a 3D valós pontok és azok 2D képei között. Ilyen pontmegfeleltetéseket a gyakorlatban úgy szoktunk előállítani, hogy egy ún. *kalibrációs mintát* helyezünk a kamera elé, melynek ismerjük a pontos 3D méreteit, és az erről készült képen azután könnyen megtalálhatjuk a mintának megfelelő képpontokat. Ezek a könnyen azonosítható pontok a mintázat sarokpontjai, ezért általában sakktábla-szerű kalibrációs mintákat szoktak használni. Egy ilyen kalibrációs mintát mutat a 2.1 ábra.



2.1. ábra. Tipikus kalibrációs minta képe halszemoptikával (balra) illetve PAL optikával (jobbra). A pontmegfeleltetéseket a négyzetek sarokpontjai szolgáltatják.



2.2. ábra. Központos vetítést megvalósító kamerák ekvivalens gömb-modellje

2.1. Körbelátó kamerák egységes Taylor modellje

A körbelátó kamera kalibrációjához elengedhetetlen a kamera geometriai leképezésének megfelelő modellezése. A Geyer és Daniilidis [7] által bemutatott modell, a középpontos kamerát egy képzeletbeli egységnyi sugarú gömb felületére való vetítésként írja le. Ezt a modellt többen is alkalmazták, a kamera belső leképezésének paramétereit különböző módon határozva meg. A Micusik és Pajdla által javasolt modell [8, 9] hasonlóan egy egységgömb segítségével modellezi a körbelátó kamerát, de Ők másképpen paraméterezték magát a vetítőfüggvényt. Ez utóbbi modell ugyan egységesen alkalmazható katadioptrikus és halszemoptikákra is, de a vetítési függvényt meghatározó f és h függvények parametrikus alakja a konkrét kamerától függően más és más lesz. Ez nyilvánvalóan nehézkesen használható egy általános célú kamera kalibrációs algoritmus fejlesztéséhez. Ezért David Scaramuzza egy új, kamera-független geometriai modellt dolgozott ki [2, 10, 11] (ld. 2.2 ábra). Az alábbiakban ezt tárgyaljuk.

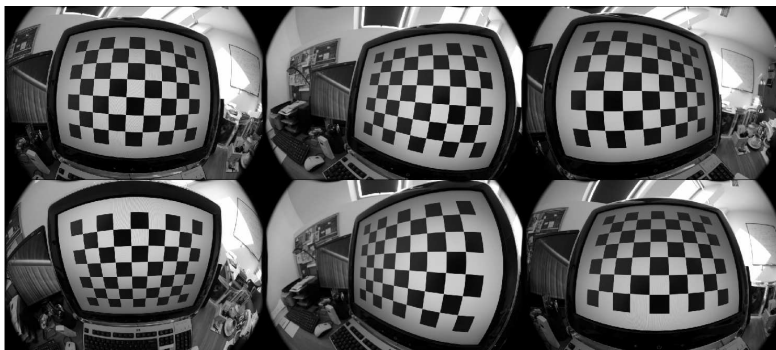
A modell lényegében az 1.5 képlet egy módosításának tekinthető, ahol a két függvény helyett egyetlen $g = f/h$ függvényt használunk:

$$\mathbf{p}'' = \begin{pmatrix} h(\mathbf{u}'')\mathbf{u}'' \\ f(\mathbf{u}'') \end{pmatrix} = \begin{pmatrix} \mathbf{u}'' \\ g(\mathbf{u}'') \end{pmatrix} \quad (2.1)$$

Továbbá a kamera-specifikus g függvény helyett annak egy általános n -ed fokú (Taylor) polinom közelítését használjuk:

$$g(\mathbf{u}'') = a_0 + a_1\|\mathbf{u}''\| + a_2\|\mathbf{u}''\|^2 + \dots + a_n\|\mathbf{u}''\|^n \quad (2.2)$$

Középpontos kamerák esetén a gyakorlatban $a_1 = 0$ [2, 10, 11], így a modell paramétereit maga az n , valamint a_0, a_2, \dots, a_n . A gyakorlatban egy negyedfokú ($n = 4$) polinom már elegendően pontos modellt ad függetlenül a kamera fizikai konstrukciójától. Vegyük észre, hogy ha $g \neq 0$ konstans, akkor a leképezés megfelel egy (a konstanstól függő fókusz távolságú) perspektív kamera leképezésének (ld. 2.2 ábra). Megjegyzendő még, hogy az 1.3 fejezetben tárgyaltak értelmében itt is egy, az 1.6 képlet szerinti affin transzformáció hat a képsík és a szenzor sík (vagyis a pixeles koordináta rendszer) között.



2.3. ábra. Kalibrációs minta képe halszemoptikával több nézőpontból.

A vetítési modell a következőképpen működik (ld. 2.2 ábra): A kamera koordináta rendszer egy \mathbf{X} 3D pontját középpontos vetítéssel vihetjük fel az egységnyi sugarú gömb felszínére a neki megfelelő \mathbf{x}_S gömbi pontba. Ez a vetítő sugár egy \mathbf{x}_g pontban metszi a 2.2 képlet által meghatározott g parametrikus görbe felszínét, ami nem más, mint a kamera belső leképezését modellező polinom függvény. Az \mathbf{x}_g pontot pedig ortografikus vetítéssel a képsíkra vetítve megkapjuk az \mathbf{X} -nek megfelelő \mathbf{x} képpontot a kamera képén. A két pont között tehát az alábbi összefüggés áll fenn:

$$\mathbf{x}_g = \left(a_0 + a_2 \|\mathbf{x}\|^2 + \dots + a_n \|\mathbf{x}\|^n \right) \mathbf{x}, \quad (2.3)$$

Vegyük észre, hogy az \mathbf{x}_S gömbi pontban találkozik az \mathbf{x}_g és a 3D \mathbf{X} pontok gömbre vontakozó középpontos vetítése. Továbbá ha \mathbf{X} egy tetszőleges világ koordináta rendszerben adott, akkor a kamera (\mathbf{R}, \mathbf{t}) külső paramétereit figyelembe véve az alábbi összefüggést kapjuk \mathbf{X} , \mathbf{x}_S és \mathbf{x}_g között:

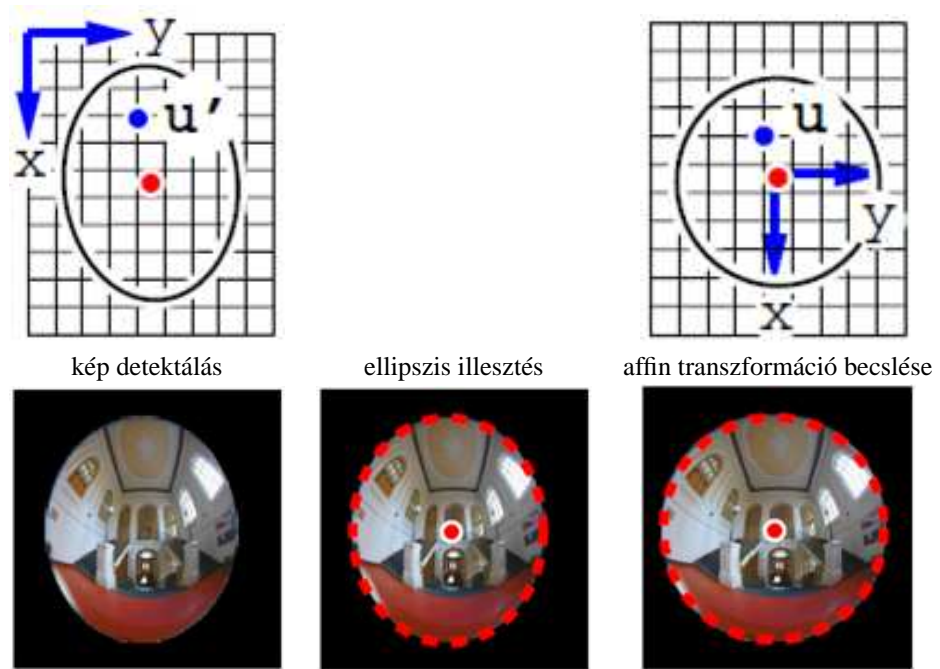
$$\frac{\mathbf{R}\mathbf{X} + \mathbf{t}}{\|\mathbf{R}\mathbf{X} + \mathbf{t}\|} = \mathbf{x}_S = \frac{\mathbf{x}_g}{\|\mathbf{x}_g\|} \quad (2.4)$$

2.2. Scaramuzza kalibrációs algoritmus

Ebben a fejezetben áttekintjük a Davide Scaramuzza [2] által kifejlesztett kalibrációs algoritmust, melynek Matlab implementációja is elérhető [12].

Hasonlóan a perspektív kamera kalibrációhoz, a körbelátó kamerával egy kalibrációs mintáról (pl. 2.1 ábra) készítünk képeket. A képkészítés során ügyelni kell az alábbiakra:

- A mintát közelről fényképezzük le, különben túl kicsi lesz a képe, ami bizonytalaná teszi a sarokpontok detektálását.
- Lehetőleg minden oldalról készítsünk képet (ld. 2.3 ábra), hogy a leképező rendszer minden része benne legyen a képalkotásban (az esetleges tükrös/lencse hibák



2.4. ábra. A kép koordináta rendszer és szenzor (pixel) koordináta rendszer közötti affin transzformáció meghatározásának lépései.

csak így detektálhatóak). Ha a kalibrációs minta dobozként körbeveszi a kamerát, akkor ez automatikusan teljesül, egyébként a kamerával járjuk körbe a kalibrációs mintát.

- A kontrolpontok detektálása sarokdetektorral történik. Mivel egyenes képe nem egyenes, így a perspektív kameráknál gyakran alkalmazott egyenesítés itt nem alkalmazható.

A kép koordináta rendszer és a szenzor (pixel) koordináta rendszer közötti affin transzformáció egyszerűen meghatározható, ha a körbe látó kamera a szenzoron ellipszis alakú képet ad (a legtöbb kamera ilyen). Ekkor a szenzor által rögzített képen egyszerűen detektálhatjuk a kép határvonalát, hiszen a háttér fekete. Az így kapott kontúrról tudjuk, hogy ideálisan kör alakú, de a valóságban ellipszis alakú lesz. Tehát egy ellipszist illesztve a kontúrra, megkapjuk annak \mathbf{C} középpontját. \mathbf{C} -t eltolva a kép középpontjába megkapjuk a \mathbf{t}' eltolásvektort, míg az \mathbf{A}' affin mátrixot úgy kapjuk, hogy az ellipszist körre transzformáljuk. Ezzel tehát megkaptuk a két koordináta rendszer között ható $(\mathbf{A}', \mathbf{t}')$ transzformációt. A Scaramuzza által javasolt eljárás ennél bonyolultabb és működik akkor is, ha a kamera képe részben vagy egészében kitölti a szenzor teljes területét (ld. [2, 10, 11]).

A következő lépésben minden egyes képre meghatározzuk a kalibrációs minta síkjának relatív helyzetét a kamerához képest. Feltehetjük, hogy $\mathbf{A}' = \mathbf{I}$ és $\mathbf{t}' = \mathbf{0}$ (az

előző lépésben ezeket meghatároztuk és a képpontokat előállítottuk a pixel koordinátákból).

Jelölje a kalibrációs minta homogén pontjait $\mathbf{M}_i = (M_{i1}, M_{i2}, 0, 1)$, a nekik megfelelő képpontokat pedig $\mathbf{m}_i = (m_{i1}, m_{i2})$. Mivel a kalibrációs minta pontjai egy síkba esnek, feltehetjük, hogy $M_{i3} = 0$. Továbbá legyen $\mathbf{P} = [\mathbf{R}|\mathbf{t}]$ a kamera külső paramétereit (relatív helyzetét) leíró kamera mátrix, amely a 3D világ koordináta rendszer pontjait átviszi a 3D kamera koordináta rendszerbe:

$$\mathbf{P}\mathbf{M}_i = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{t}] \begin{pmatrix} M_{i1} \\ M_{i2} \\ 0 \\ 1 \end{pmatrix} = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{t}] \begin{pmatrix} M_{i1} \\ M_{i2} \\ 1 \end{pmatrix}, \quad (2.5)$$

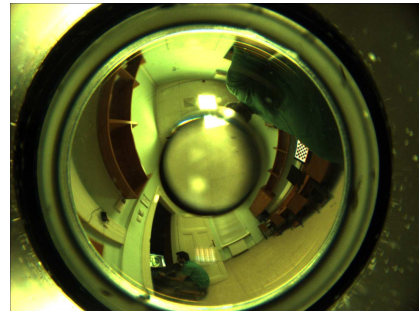
ahol \mathbf{r}_i az \mathbf{R} forgató mátrix i . oszlopa. Az \mathbf{m}_i képpontnak megfelelő g felszínen vett \mathbf{m}_{gi} pont a 2.4 képlet értelmében közös gömbi vetítőegyenesre esik az \mathbf{M}_i ponttal (ld. 2.2 ábra), tehát:

$$\mathbf{m}_{gi} \times \mathbf{P}\mathbf{M}_i = \begin{pmatrix} m_{i1} \\ m_{i2} \\ g(\mathbf{m}_i) \end{pmatrix} \times [\mathbf{r}_1, \mathbf{r}_2, \mathbf{t}] \begin{pmatrix} M_{i1} \\ M_{i2} \\ 1 \end{pmatrix} = 0 \quad (2.6)$$

A fenti egyenletet kifejtve minden pontra kapunk 3 egyenletet, amiből egy lesz lineáris az (\mathbf{R}, \mathbf{t}) paraméterekben, konkrétan a $\mathbf{h} = (r_{11}, r_{12}, r_{21}, r_{22}, t_1, t_2)^T$ vektorbeli paraméterekre. Ezeket az egyenleteket minden pontra felírva és megfelelően rendezve kapunk egy lineáris túlhatározott egyenletrendszert $\mathbf{M}\mathbf{h} = \mathbf{0}$ alakban, ahol \mathbf{M} csak az ismert $(\mathbf{M}_i, \mathbf{m}_i)$ 3D-2D pontoktól függő mátrix. Az egyenletrendszert az \mathbf{M} mátrix SVD felbontásával oldjuk meg. Mivel \mathbf{R} ortonormált, ezért a \mathbf{h} -ből hiányzó r_{31}, r_{32} is egyértelműen meghatározható, t_3 meghatározása pedig a belső paraméterekkel együtt történik.

Az eddigiekben pontonként a 2.6 képlet egyetlen egyenletét használtuk, amellyel meghatároztuk $\mathbf{r}_1, \mathbf{r}_2, t_1, t_2$ értékeket. Ezeket pontonként behelyettesítve a maradék két egyenletbe, majd az összes kép összes pontpárjára kapott egyenleteket megfelelően rendezve felírhatunk egy újabb lineáris egyenletrendszert, melynek ismeretleni a képenkénti t_3 és a g paraméterei: a_0, a_2, \dots, a_n . A megoldást ismét SVD felbontással kapjuk. Ezzel a kamera belső paramétereit reprezentáló g valamennyi paraméterét meghatároztuk, valamint minden képre megkaptuk a megfelelő kalibrációs minta és a kamera relatív helyzetét leíró (\mathbf{R}, \mathbf{t}) paramétereket. Az így kapott megoldás azonban csak egy algebrailag optimális megoldást ad, amely a perspektív kameráknál is alkalmazott *visszavetítési hibával* nem-lineárisan tovább finomítható [2, 10, 11].

Mivel a körbelátó kamera leképezését a kamera koordináta rendszerből a kép koordináta rendszerbe teljes mértékben leírja a g függvény, így ennek ismeretében például visszavetíthetjük a kamera képét egy síkra, ezzel „kiterítve” a nemlineárisan torzult képet egy hagyományos perspektív képpé. Természetesen a körbelátó kamerák változó felbontása és nagy látószöge miatt a kiterített kép minősége nem éri el egy hagyományos perspektív kamera képminőségét, mégis sok alkalmazásban hasznos lehet egy ilyen virtuális kép előállítás. Erre mutat példát a 2.5 ábra.



Halszem optika

PAL optika

2.5. ábra. Kalibrált körbelátó kamera képeiből generált „kiterített” perspektív kép.

Irodalomjegyzék

- [1] C. Geyer, T. Pajdla, and K. Daniilidis. (2003) ICCV2003 short course on omnidirectional vision. [Online]. Available: <http://www.cis.upenn.edu/~kostas/omni/geyer03tutorial.ppt>
- [2] D. Scaramuzza, „Omnidirectional vision: from calibration to robot motion estimation,” Ph.D. dissertation, ETH Zurich, 2008.
- [3] R. Frohlich, L. Tamas, and Z. Kato, „Homography estimation between omnidirectional cameras without point correspondences,” in *Proceedings of ICRA Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras*, IEEE. Hong Kong: IEEE, June 2014, accepted.
- [4] L. Tamas, R. Frohlich, and Z. Kato, „Relative pose estimation and fusion of omnidirectional and lidar cameras,” in *Proceedings of the ECCV Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving*, ser. Lecture Notes in Computer Science. Zurich, Switzerland: Springer, Sept. 2014, accepted.
- [5] Z. Kató and L. Czúni, *Számítógépes látás*. Typotex kiadó, 2011.
- [6] L. Puig and J. J. Guerrero, *Omnidirectional Vision Systems: Calibration, Feature Extraction and 3D Information*. Springer, 2013.
- [7] C. Geyer and K. Daniilidis, „A unifying theory for central panoramic systems,” in *European Conference on Computer Vision (ECCV)*, 2000, pp. 445–462.
- [8] B. Micusik, „Two-view geometry of omnidirectional cameras,” Ph.D. dissertation, Center for Machine Perception, Czech Technical University in Prague, June 2004.
- [9] B. Micusik and T. Pajdla, „Structure from motion with wide circular field of view cameras,” *Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 7, pp. 1135–1149, July 2006.
- [10] D. Scaramuzza, A. Martinelli, and R. Siegwart, „A toolbox for easily calibrating omnidirectional cameras,” in *IEEE/RSJ International Conference on Intelligent Robots*. Beijing: IEEE, October 9–15 2006, pp. 5695–5701.

- [11] D. Scaramuzza, A. Martinelli, and R. Siegwart, „A flexible technique for accurate omnidirectional camera calibration and structure from motion,” in *Proceedings of the Fourth IEEE International Conference on Computer Vision Systems*, ser. ICVS-06. Washington, USA: IEEE Computer Society, 2006, pp. 45–51.
- [12] D. Scaramuzza. Ocamcalib: Omnidirectional camera calibration toolbox for matlab. [Online]. Available: <https://sites.google.com/site/scarabotix/ocamcalib-toolbox>