

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 1556

**REKONSTRUKCIJA TRODIMENZIJSKIH OBJEKATA IZ STEREO
SLIKE**

Goran Jakovljević

Zagreb, rujan 2005.

SADRŽAJ

1	UVOD.....	3
1.1	OPIS PROBLEMA	4
2	OSNOVE STEREO VIDA.....	5
2.1	MODEL KAMERE.....	5
2.2	PROMJENA KOORDINATNOG SUSTAVA.....	7
2.2.1	<i>Promjena koordinatnog sustava ravnine slike</i>	7
2.2.2	<i>Promjena koordinatnog sustava scene</i>	12
2.3	ODSTUPANJE OD MODELA KAMERE	14
2.4	GEOMETRIJSKA INTERPRETACIJA REDAKA PROJEKCIJSKE MATRICE.....	16
3	KORESPODENTNE TOČKE	17
3.1	EPIPOLOVI I EIPOLARNE LINIJE.....	17
3.2	IZRAČUNAVNJE EPIPOLOVA	19
3.3	ISPRAVLJANJE EIPOLARNIH LINIJA	21
3.4	ODREĐIVANJE KORESPODENTNIH TOČAKA	23
3.5	METODE PRETRAŽIVANJE PODRUČJA SLIKE	26
3.6	PRETRAŽIVANJE POMOĆU PROZORA	27
4	3D REKONSTRUKCIJA.....	29
5	PROGRAMSKA IMPLEMENTACIJA	30
5.1	OPIS SUSTAVA	30
5.2	KALIBRACIJA KAMERA	31
5.3	UKLANJANJE RADIJALNE DISTORZIJE.....	33
5.4	ISPRAVLJANJE EIPOLARNIH LINIJA	34
5.5	PREPOZNAVANJE KORESPODENTNIH TOČAKA.....	35
6	ZAKLJUČAK	39
7	LITERATURA.....	40

1 Uvod

U posljednjih nekoliko godina vizualizacija 3D objekata pomoću računala doživjela je golem napredak. Mnoge tehnike su sazrijele i sada se obavljaju sklopovski. Ono što je prije nekoliko godina zahtijevalo računalo čija se cijena izražavala u milijunima, sada se može postići i s prosječnim računalom koje se koristi samo za igru.

Posljedica ove evolucije je zahtjev za kompleksnijim i realističnijim 3D modelima. Iako su alati za trodimenzionalno modeliranje sve moćniji, sintetiziranje realističnih modela je složeno i vremenski zahtjevno, pa time i skupo. Mnogi virtualni objekti inspirirani su stvarnim objektima, pa bi bilo interesantno kada bi ih se na neki način moglo 'prebaciti' na računalo.

Istraživači su proučavali metode za dobivanje 3D informacija o objektima i okolini mnogo godina i stvoreni su brojni sustavi koji na različite načine prikupljaju podatke o veličini objekata, izgledu i njihovom položaju u prostoru. U prošlosti glavna primjena bila je navođenje robota pri čemu je bilo dovoljno određivanje koordinata manjeg broja točaka. Danas su ti sustavi izašli iz laboratorija sa strogo kontroliranim uvjetima i sve češće nalaze primjenu u raznim granama industrije i gospodarstva. Iako ima dosta rješenja koja koriste takve sustave za autonomno navođenje vozila, provjeru kvalitete proizvodnje, digitalizaciju arheoloških iskopina i sl., ipak najviše primjera upotrebe možemo naći u filmskoj industriji. Korištenje sustava koji omogućuju reprodukciju objekata iz okoline donijelo je revoluciju u izradi animiranih filmova. Likovi se sada iz stvarnih 3D modela koji su izrađeni od gline ili sličnih materijala mogu prebaciti u neki program za 3D oblikovanje gdje se koriste u izradi animacija ili umjetnog okoliša, čime je umjetnicima vraćena sloboda stvaranja koja je na neki način bila sputana mogućnostima korištenog programa.

Potreba za što realističnijom rekonstrukcijom razlog je zbog kojeg su postojeći sustavi izgrađeni oko specijaliziranog sklopovlja (npr. projektori, laseri,...) što je uzrok njihove vrlo velike cijene. Za mnoge primjene, naprotiv, potreban je robustniji i jeftiniji sustav.

Jedan takav sustav za rekonstrukciju 3d objekta koji koristi dvije digitalne kamere bit će opisan u nastavku teksta.

1.1 Opis problema

Digitalna kamera, zbog svoje pristupačnosti, osnova je velikog broja sustava za prikupljanje podataka o trodimenzionalnom objektu. Postoje mnogi radovi u kojima se koristi jedna kamera i neka pomoćna sredstva kao npr. ogledala, laseri, projektori da bi se mogla odrediti prostorna koordinata promatrane točke na objektu. Dodavanjem još jedne kamere u takav sustav (sa ili bez pomagala) razvila se metoda nazvana stereovid (engl. stereovision).

“Stereo vid” je metoda dobivanja trodimenzionalnih podataka uz pomoć slika snimljenih s dvije kamere. Metoda je nastala oponašanjem načina na koji ljudi i životinje promatraju svijet i određuju udaljenosti i međusobne prostorne odnose između promatranih objekata. Sustav koji se koristi sastoji se od dvije kamere, postavljene jedna pored druge i pomoću njega dobijamo dvije slike istog objekta, koje se malo razlikuju zbog različite pozicije i orijentacije kamere (slika 1.1).



Slika 1.1 Slike istih objekata snimljene s dvije kamere postavljene jedna pored druge

Ovaj postupak zahtjeva poznavanje parametara kamera i njihov međusobni položaj (položaj i orijentacija jedne kamere u odnosu na položaj i orijentaciju druge). Parametri i položaj kamera određuju se kalibracijama kamera.

Kalibracijama kamera saznali smo na koji su način stvorene lijeva i desna slike, pa ako izaberemo iste točke snimljenog objekta na jednoj i na drugoj slici pomoću postupka obrnutog od nastanka slike moguće je odrediti točku u prostoru koja odgovara izabranim točkama sa slika.

Taj postupak je teoretski jednostavno provesti, ali zbog nesavršenosti opreme i nemogućnosti određivanja apsolutno točnih parametara kamera moramo se koristiti raznim trikovima kako bi ostvarili što točnije rezultate.

Iako je stereovid već duže vrijeme predmet proučavanja i iako postoje brojni radovi na tu temu, još uvijek nedostaju usporedbe učinkovitosti različitih postupaka što otežava izbor načina ostvarenja ovakvog sustava.

U ovom radu pokušat ćemo rekonstruirati neki objekt snimljen s dvije kamere, a pri tom ćemo opisati sustav i neke algoritme potrebna za taj zadatak.

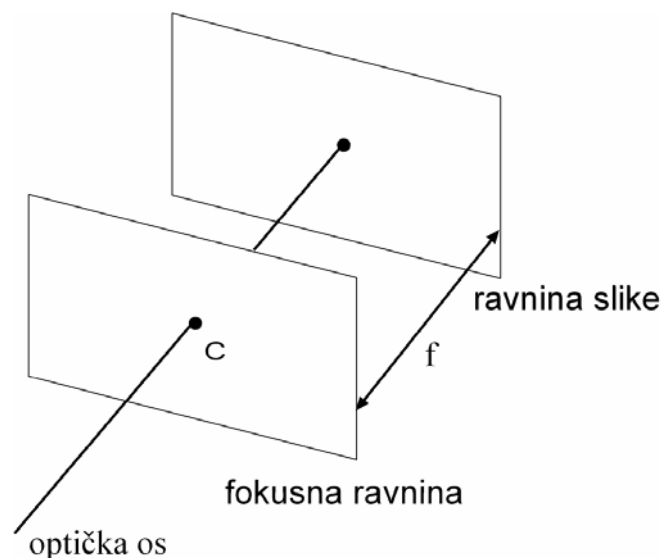
2 Osnove stereo vida

2.1 Model kamere

Da bi shvatili kako nastaje slika pomoću kamere, te da bi iz dvije slike mogli reproducirati trodimenzionalni oblik objekta, potrebno je odrediti model kamere.

Najjednostavniji geometrijski model kamere, prikazan na slici 2.1 sastoji se od dvije paralelne ravnine - **fokusne ravnine** i **ravnine slike**. Udaljenost između te dvije ravnine je **fokusna duljina** i označava se sa **f**. Na fokusnoj ravnini nalazi se točka C koja se naziva **optički centar**. Kroz optički centar prolazi **optička os** - pravac koji je okomit na ravninu slike.

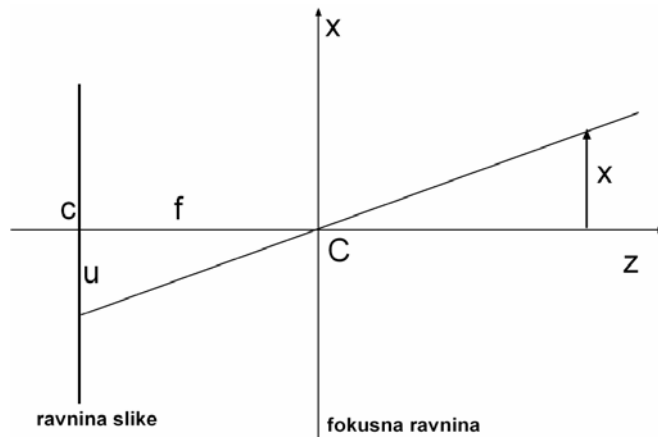
Slika neke točke M u prostoru formira se u ravnini slike u točki u kojoj pravac koji prolazi kroz točku M i optički centar siječe ravninu slike. To preslikavanje iz trodimenzionalnog u dvodimenzionalni prostor određeno je jednačbom perspektivne projekcije.



Slika 2.1 Prikaz optičke osi, fokusne ravnine i ravnine slike

Uvođenjem koordinatnog sustava prostora (naziva se još koordinatni sustav svijeta ili koordinatni sustav scene) (C, x, y, z) i koordinatnog sustava slike (c, u, v) moguće je odrediti jednačbu perspektivne projekcije. Da bi pojednostavili određivanje te jednačbe pretpostavit ćemo da vrijede slijedeće pretpostavke:

- optički centar (još se naziva i centar projekcije) podudara se s ishodištem koordinatnog sustava svijeta
- os kamere (optička os) poravnata je s z osi svijeta



Slika 2.2 Geometrijski prikaz nastanka slike

Nakon što smo uveli te pretpostavke iz sličnosti trokuta (slika 2.2) lako možemo odrediti

$$\frac{x}{z} = \frac{u}{f} \quad \text{i} \quad \frac{y}{z} = \frac{v}{f} \quad \Rightarrow \quad \frac{f}{z} = \frac{u}{x} = \frac{v}{y}$$

a to možemo izraziti i ovako

$$u = f \frac{x}{z} \quad v = f \frac{y}{z} \quad (2.1)$$

U homogenim koordinatama perspektivnu projekciju zapisujemo pomoću matrice, tako da s lijeve strane jednačbe imamo točku na slici, a s desne strane točku u prostoru pomnoženu s 3x4 matricom

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

(negativni predznak nam govori da će dobivena slika biti invertirana)

Gornja jednačba je definirana 'do na konstantu' (engl. up to scale) pa ju možemo zapisati i pomoću projekcijskih koordinata (X,Y,Z,T) točke M

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix} \quad (2.2)$$

Jednadžba 2.3 izražava činjenicu da je odnos između koordinata slike i koordinata u prostoru linearan u projekcijskim koordinatama i da se može zapisati u matričnom obliku

$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{M}} \quad (2.3)$$

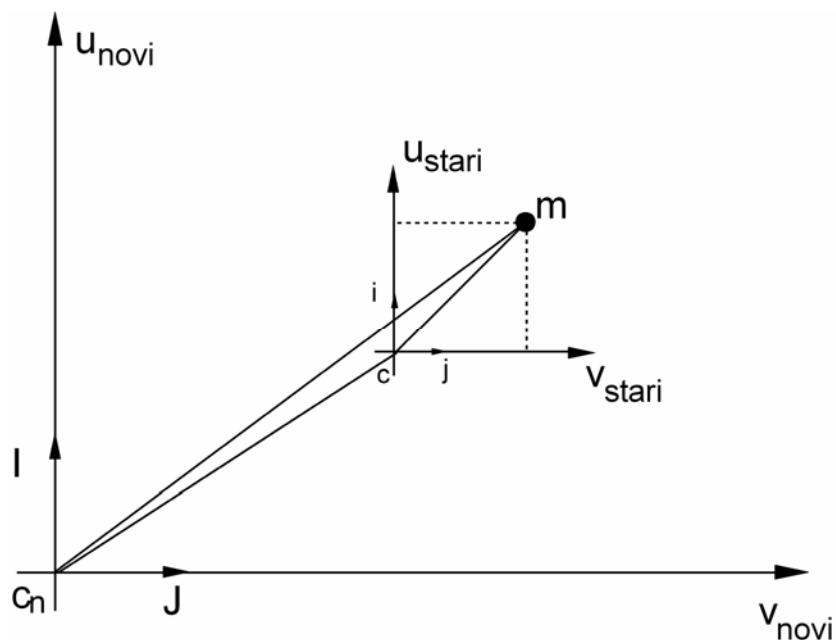
gdje je $\tilde{\mathbf{m}} = [U, S, V]^T$, a $\tilde{\mathbf{M}} = [X, Y, Z, T]^T$
 (homogeni ili projekcijski vektori i matrice u tekstu će se označavati valovitom crtom)

2.2 Promjena koordinatnog sustava

U nastavku ćemo opisati kako se matrica $\tilde{\mathbf{P}}$ mijenja kada mijenjamo koordinatni sustav slike i koordinatni sustav svijeta. Te promjene su važne jer će nam omogućiti definiranje unutrašnjih i vanjskih parametara kamere.

2.2.1 Promjena koordinatnog sustava ravnine slike

Prvo ćemo promotriti kako promjena središta koordinatnog sustava slike te promjene jedinica na u i v osi utječu na matricu perspektivne projekcije $\tilde{\mathbf{P}}$. Prije nego što navedemo postupak, uvest ćemo pomalo nestandardnu notaciju vektora, preuzetu iz literature [1]. Tako ćemo vektor koji povezuje točku \mathbf{a} i točku \mathbf{b} (gdje su \mathbf{a} i \mathbf{b} radij vektori dvodimenzionalnih točaka) označiti sa \mathbf{ab} . To u ovom slučaju nije oznaka operacije množenja ova dva vektora, već vektor između te dvije točke. Iz toga slijedi slijedeće: ako oznaka \mathbf{a} označava ishodište nekog koordinatnog sustava, tada oznaka \mathbf{ab} označava radij vektor točke \mathbf{b} u tom koordinatnom sustavu.



Slika 2.3 Promjena koordinatnog sustava ravnine slike

Na slici 2.3 prikazana je promjena koordinatnog sustava slike. Ishodište starog koordinatnog sustava je u sjecištu \mathbf{c} optičke osi i ravnine slike i ima iste jedinice na obje osi. Iz starog koordinatnog sustava prelazimo u novi kojem je ishodište u točki \mathbf{c}_n slike (obično u jednom od kuteva) i koji će ponekad imati različite jedinice na koordinatnim osima.

Za piksel \mathbf{m} vrijedi

$$\mathbf{c}_n \mathbf{m} = \mathbf{c}_n \mathbf{c} + \mathbf{c} \mathbf{m} \quad (2.4)$$

gdje je $\mathbf{c}_n \mathbf{m}$ radij vektor točke \mathbf{m} u novom koordinatnom sustavu.

Primjer.

Kao primjer navest ćemo promjenu koordinatnog sustava slike s ishodištem u donjem lijevom kutu. Mjerenjem smo odredili da je centar radijalne distorzije (objašnjenje pojma navedeno je kasnije u tekstu) u točki \mathbf{c} s koordinatama $[500, 500]$. Budući da je svejedno koji koordinatni sustav uzimamo da je novi, a koji stari, uzmimo da je ishodište starog koordinatnog sustava u točki \mathbf{c} (takva pretpostavka uzeta je zbog modela kamere prema kojem je središnja koordinata slike $[0, 0]$), a slika koju imamo dana je u novom sustavu. Prebacivanje slike u stari sustav radimo pomoću dane formule

$$\text{koord. u starom sustavu} = \text{koord. u novom} - \text{ishodište starog sustava}$$

Ako uzmemo neku točku naše slike \mathbf{m} s koordinatama $[100, 200]$ dobit ćemo točku $[-400, -300]$ u starom koordinatnom sustavu.

Osim translacije koordinatnog sustava, u praksi često moramo promijeniti jedinice na koordinatnim osima. To ćemo postići promjenom veličina (smanjivanjem ili povećanjem) jediničnih vektora \mathbf{i} i \mathbf{j} koji određuju koordinatne osi. Skaliranje je dano matricom \mathbf{s}

$$\mathbf{s} = \begin{bmatrix} k_u & 0 \\ 0 & k_v \end{bmatrix} \quad (2.5)$$

Ako zapišemo $\mathbf{c} \mathbf{m}$ u starom koordinatnom sustavu $\mathbf{c} \mathbf{m} = u_{\text{stari}} \mathbf{i} + v_{\text{stari}} \mathbf{j}$ i uvedemo skaliranje iz starog koordinatnog sustava određenog s jediničnim vektorima (\mathbf{i}, \mathbf{j}) u novi (\mathbf{I}, \mathbf{J}) vrijedi

$$\mathbf{I} = s \mathbf{i} \quad \mathbf{J} = s \mathbf{j}$$

Primjer

Uzmimo koordinatni sustav određen jediničnim vektorima $\mathbf{i}=[1 \ 0]^T$ i $\mathbf{j}=[0 \ 1]^T$ i točku u tom sustavu \mathbf{m} s koordinatama $[100, 10]$. Ona se može zapisati pomoću vektora $\mathbf{m} = 100\mathbf{i} + 10\mathbf{j}$.

Ako želimo dobiti točku u koordinatnom sustavu s jediničnim vektorima \mathbf{I} i \mathbf{J} , tako da vrijedi $\mathbf{I} = \mathbf{i}$, $\mathbf{J} = -\mathbf{j}$ dobit ćemo je množenjem s matricom \mathbf{s}

$$\mathbf{s} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\mathbf{m}_{NOVA} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} * \mathbf{m}_{STAR} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} * \begin{bmatrix} 100 \\ 10 \end{bmatrix} = \begin{bmatrix} 100 \\ -10 \end{bmatrix}$$

Uz $\mathbf{c}_n \mathbf{c} = \mathbf{t}$ u novom koordinatnom sustavu dobivamo jednadžbu točke \mathbf{m}_{nova} u projekcijskim koordinatama

$$\tilde{\mathbf{m}}_{nova} = \tilde{\mathbf{H}} \tilde{\mathbf{m}}_{stara} \quad (2.6)$$

gdje je $\tilde{\mathbf{H}}$ 3x3 matrica

$$\tilde{\mathbf{H}} = \begin{bmatrix} k_u & 0 & t_x \\ 0 & k_v & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Iz navedenih jednadžbi možemo dobiti novu matricu perspektivne projekcije

$$\begin{aligned} \tilde{\mathbf{m}}_{stara} &= \tilde{\mathbf{P}}_{stara} \tilde{\mathbf{M}} \\ \tilde{\mathbf{m}}_{nova} &= \tilde{\mathbf{H}} \tilde{\mathbf{P}}_{stara} \tilde{\mathbf{M}} \\ \tilde{\mathbf{P}}_{nova} &= \tilde{\mathbf{H}} \tilde{\mathbf{P}}_{stara} \end{aligned} \quad (2.7)$$

Ako s u_0 i v_0 označimo koordinate vektora \mathbf{t} tada se matrica $\tilde{\mathbf{P}}$ može zapisati

$$\tilde{\mathbf{P}} = \begin{bmatrix} -fk_u & 0 & u_0 & 0 \\ 0 & -fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.8)$$

U rijetkim slučajevima kada u i v osi nisu međusobno okomite u model kamere (tj. u matricu perspektivne projekcije) može se dodati i smik koordinatnih osi pa matrica $\tilde{\mathbf{P}}$ poprima slijedeći oblik

$$\tilde{\mathbf{P}} = \begin{bmatrix} -fk_u & fk_u \cot \theta & u_0 & 0 \\ 0 & -\frac{fk_v}{\sin \theta} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.9)$$

(kut θ je kut između koordinatnih osi, ova matrica je navedena radi potpunosti opisa sustava, izvod se nalazi u literaturi [1])

Smik se pojavljuje kada radimo sliku slike i optička os nije okomita na ravninu slike. Posmak također može biti različit od nule kada kamera snima slike preko nekog optičkog pomagala npr. zrcala.

Budući da je posmak uglavnom jednak nuli mi ćemo ga u našem modelu zanemariti.

Neka je $\alpha_u = -f k_u$ $\alpha_v = -f k_v$. Parametri α_u , α_v , u_0 , v_0 ne ovise o poziciji i orijentaciji kamere u prostoru pa se stoga nazivaju **intrinzični** ili **unutarnji parametri kamere**.

Napomena

Parametar α_u može se procijeniti tako da podijelimo fokusnu duljinu leće kamere s širinom slike na ravnini slike (za digitalne kamere to je CCD polje) i rezultat pomnožimo s širinom dobivene slike izraženom u pikselima. Analogno možemo izračunati α_v .

Nakon što smo odredili kako promjene koordinatnog sustava slike utječu na matricu perspektivne projekcije definirat ćemo poseban koordinatni sustav koji nam dopušta da normiramo koordinate slike. Taj koordinatni sustav naziva se **normirani koordinatni sustav kamere** i on se vrlo često koristi u aplikacijama stereo vida jer nam dozvoljava da zanemarimo specifične karakteristike kamere i da radimo s idealnim sustavom. Matrica perspektivne projekcije tog sustava je

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.10)$$

Prema jednadžbi za prelazak između sustava s različitim matricama perspektivne projekcije $\tilde{\mathbf{P}}_{nova} = \tilde{\mathbf{H}}\tilde{\mathbf{P}}_{stara}$ matrica $\tilde{\mathbf{H}}$ ima slijedeću vrijednost

$$\tilde{\mathbf{H}} = \begin{bmatrix} 1 & 0 & -\frac{u_0}{\alpha_u} \\ \alpha_u & 1 & -\frac{v_0}{\alpha_v} \\ 0 & \alpha_v & \alpha_v \\ 0 & 0 & 1 \end{bmatrix}$$

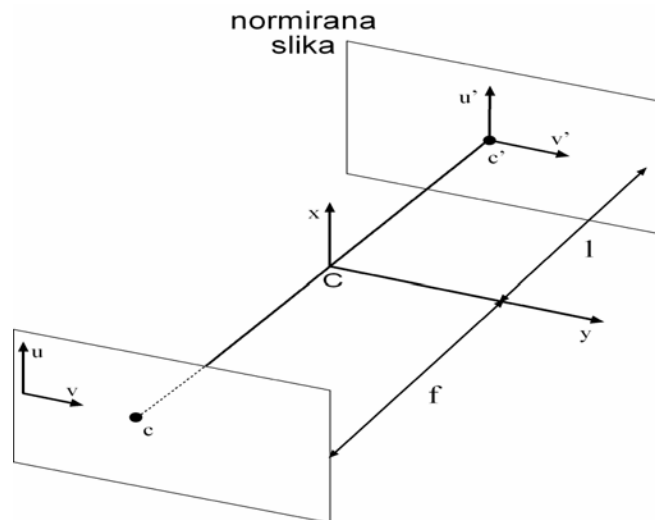
Tu matricu možemo dobiti ako u jednadžbu $\tilde{\mathbf{P}}_{nova} = \tilde{\mathbf{H}}\tilde{\mathbf{P}}_{stara}$ uvrstimo

$$\tilde{\mathbf{P}}_{stara} = \begin{bmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \tilde{\mathbf{P}}_{nova} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

i riješimo taj sustav jednadžbi.

Nove koordinate ravnine slike (kao što se može i uočiti na slici 2.4) dane su sa sljedećim jednadžbama

$$u' = \frac{u - u_0}{\alpha_u} \quad v' = \frac{v - v_0}{\alpha_v} \quad (2.11)$$



Slika 2.4 Normirani koordinatni sustav kamere

Iz dosadašnjih razmatranja matrice perspektivne projekcije mogu se odrediti jedinice u kojima je pojedina veličina izražena.

Ako uzmemo već navedenu jednadžbu

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

i izrazimo veličine x , y , z , f u jedinicama duljine (npr. metrima), a u i v u pikselima tada jednadžbe

$$u = \frac{U}{S} = -fk_u \frac{x}{z} + u_0 \quad \text{i} \quad v = \frac{V}{S} = -fk_v \frac{y}{z} + v_0$$

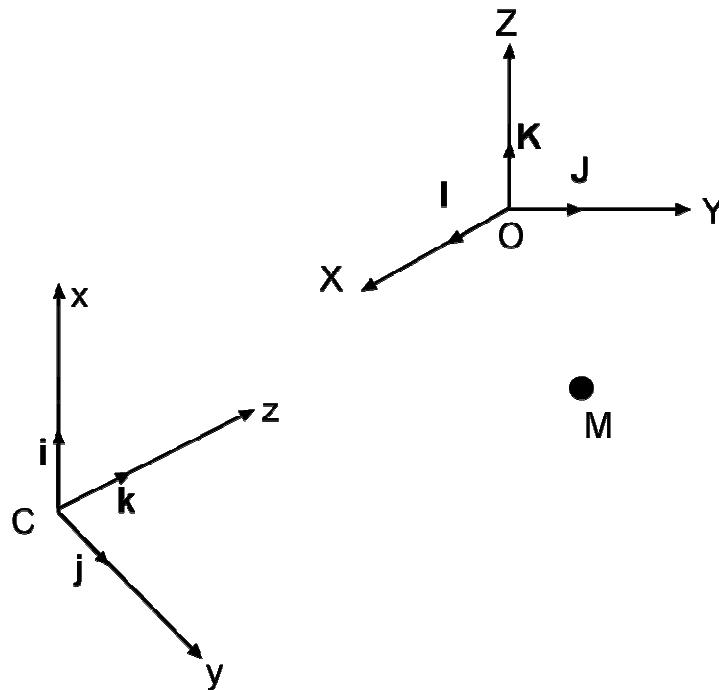
pokazuju da su k_u i k_v izraženi u jedinicama piksel/m, a u_0 i v_0 u pikselima.

$1/k_u$ i $1/k_v$ mogu se protumačiti kao veličina u metrima vodoravnih i okomitih piksela, a u_0 i v_0 kao veličina fokusne duljine u horizontalnim i vertikalnim pikselima.

2.2.2 Promjena koordinatnog sustava scene

U sljedećem dijelu promotrit ćemo kako promjena koordinatnog sustava scene utječe na matricu perspektivne projekcije. Pojedine veličine označavat ćemo na analogan način kao pri razmatranju promjene koordinatnog sustava slike.

Promijenit ćemo stari koordinatni sustav s optičkim centrom C s novim koordinatnim sustavom s optičkim centrom O i to pomoću rotacije R i translacije $T = CO$, kao što se vidi na slici 2.5.



Slika 2.5 Prikaz dva trodimenzionalna koordinatna sustava

Ako koristimo oznake sa slike, možemo zapisati

$$CM = CO + OM$$

OM možemo izraziti u novom koordinatnom sustavu

$$OM = x_{novi} \mathbf{I} + y_{novi} \mathbf{J} + z_{novi} \mathbf{K}$$

Uvodeći rotacijsku matricu R koja rotira stari koordinatni sustav u novi pišemo

$$\mathbf{I} = R * \mathbf{i} \qquad \mathbf{J} = R * \mathbf{j} \qquad \mathbf{K} = R * \mathbf{k}$$

Matrica R je 3×3 matrica.

Konačno, ako označimo vektor \mathbf{CO} u starom koordinatnom sustavu s $\mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$ dobivamo

$$\mathbf{CM} = \mathbf{t} + \mathbf{R} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (2.12)$$

U projekcijskim koordinatama to daje

$$\tilde{\mathbf{M}}_{stari} = \tilde{\mathbf{K}}\tilde{\mathbf{M}}_{nova} \quad (2.13)$$

gdje je $\tilde{\mathbf{K}}$ 4x4 matrica (oznaka $\tilde{\mathbf{K}}$ je odabrana zbog podudarnosti s literaturom i ne odnosi se na jedinični vektor \mathbf{K} koordinatne osi \mathbf{Z})

$$\tilde{\mathbf{K}} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

ili ako koristimo skraćenu notaciju

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix}$$

Ako izrazimo točku na slici u projekcijskim koordinatama s

$$\tilde{\mathbf{m}}_{stara} = \tilde{\mathbf{P}}_{stara} \tilde{\mathbf{M}}_{stara}$$

$$\tilde{\mathbf{m}}_{stara} = \tilde{\mathbf{P}}_{stara} \tilde{\mathbf{K}}\tilde{\mathbf{M}}_{nova}$$

napokon možemo zapisati novu projekcijsku matricu

$$\tilde{\mathbf{P}}_{nova} = \tilde{\mathbf{P}}_{stara} \tilde{\mathbf{K}}$$

Matrica \mathbf{R} i \mathbf{t} opisuju poziciju i orijentaciju kamere u odnosu na novi koordinatni sustav scene. Ti parametri nazivaju se ekstrinzični ili vanjski parametri kamere.

Sada možemo zapisati projekcijsku matricu kao funkciju unutarnjih i vanjskih parametara kamere:

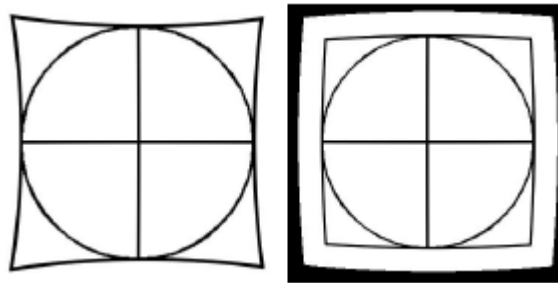
$$\tilde{\mathbf{P}} = \begin{bmatrix} \alpha_u \mathbf{r}_1 + u_0 \mathbf{r}_3 & \alpha_u t + u_0 t_z \\ \alpha_v \mathbf{r}_2 + v_0 \mathbf{r}_3 & \alpha_v t_y + v_0 t_z \\ \mathbf{r}_3 & t_z \end{bmatrix} \quad (2.14)$$

2.3 Odstupanje od modela kamere

Perspektivni model kamere opisuje proces nastanka slike dovoljno precizno za većinu kamera. Ipak, kada je potrebna vrlo velika preciznost ili kada se koriste jeftinije kamere moraju se u obzir uzeti i dodatni parametri.

Tada treba uzeti u obzir nemogućnost optičkog sustava da svaku zraku svjetlosti od jedne točke objekta projicira u jednu točku slike ili pripisanu geometrijsku poziciju. Ta pogreška optičkog sustava naziva se aberacija. Postoje mnogi tipovi aberacije (astigmatizam, kromatska aberacija, sferna aberacija, distorzijska aberacija).

Mnoge od tih pogrešaka su zanemarive u normalnim uvjetima. Radijalna distorzija, naprotiv, može se uočiti pri kraćim fokusnim dužinama. Radijalna distorzija je linearan pomak točaka slike radijalno prema ili od centra slike kao što je vidljivo na slici 2.6. Uzrok joj je činjenica da objekti na različitim kutnim duljinama od osi leće podliježu različitim povećanjima.



Slika 2.6 Radijalna distorzija

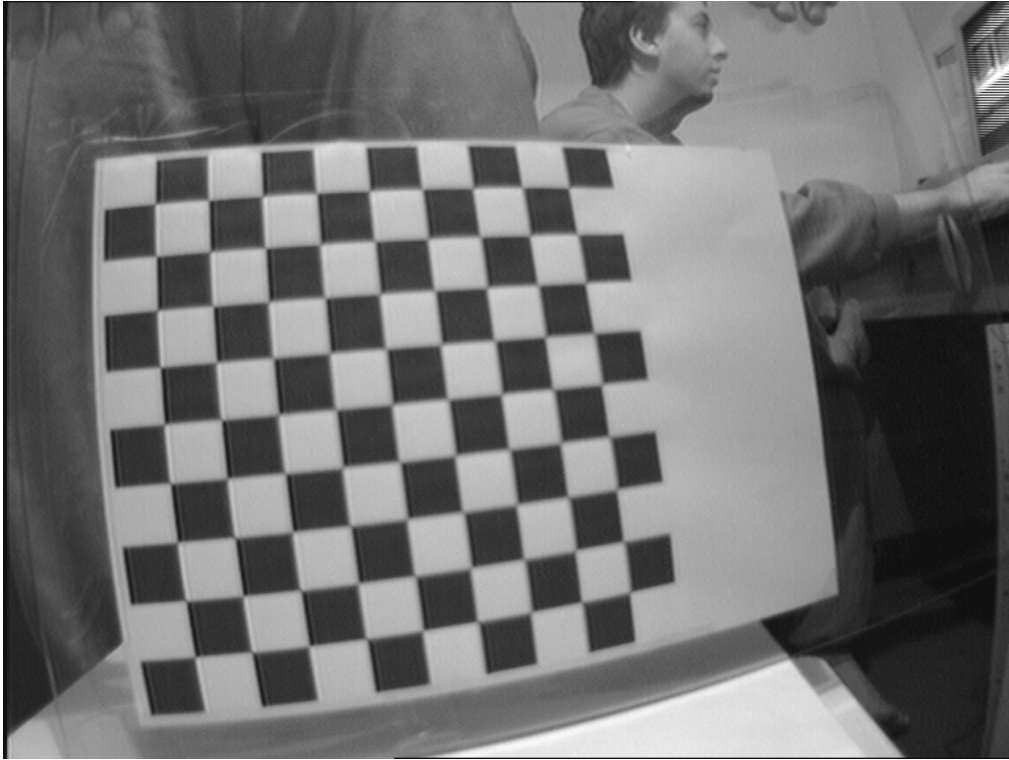
Moguće je ukloniti većinu tih efekata ‘savijanjem’ slike. Nedistorzirane koordinate mogu se pomoću sljedećih jednažbi dobiti iz distorzirane slike:

$$\begin{aligned}u &= u_0 * (1 + K_1 * r^2 + K_2 * r^4 + \dots) \\v &= v_0 * (1 + K_1 * r^2 + K_2 * r^4 + \dots)\end{aligned}\tag{2.15}$$

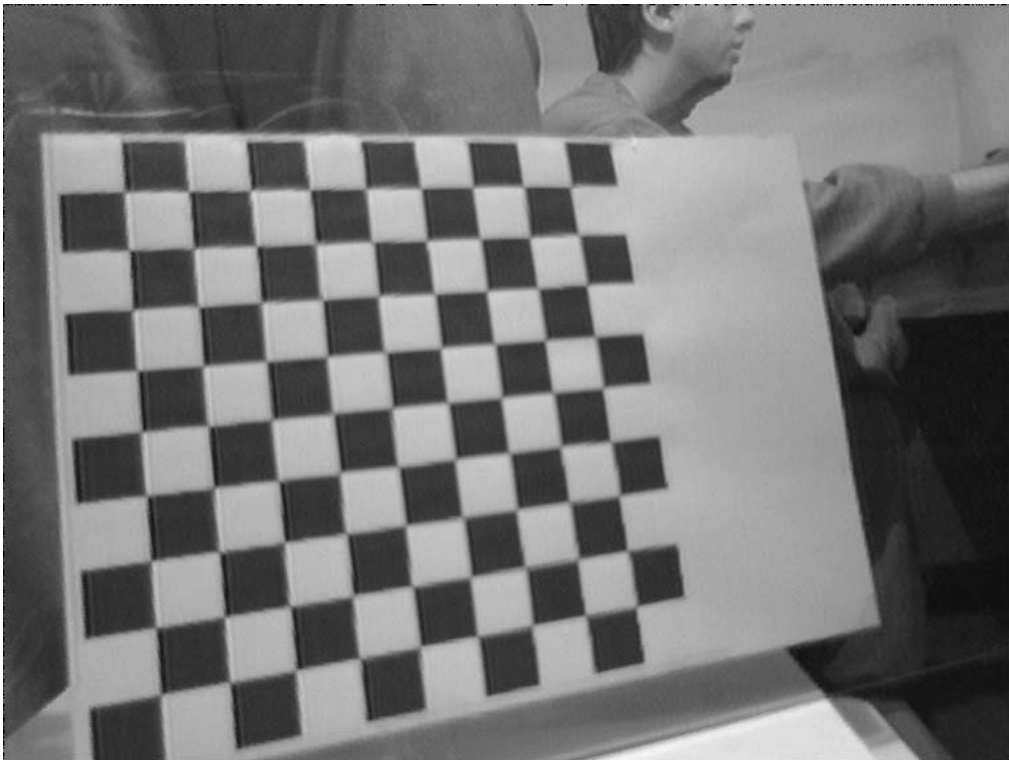
gdje su K_1 i K_2 prvi i drugi parametar radijalne distorzije, a r je

$$r^2 = u_0^2 + v_0^2$$

Pri tom valja naglasiti da se radi s normaliziranim koordinatama, pa pri tom nisu bitni unutrašnji parametri kamere.



Slika 2.7 *Slika s radijalnom distorzijom*



Slika 2.8 *Ispravljena slika*

2.4 Geometrijska interpretacija redaka projekcijske matrice

Geometrijsku interpretaciju redaka projekcijske matrice možemo dati ako ju zapišemo na slijedeći način:

$$\tilde{\mathbf{P}} = \begin{bmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \\ q_{31} & q_{32} & q_{33} & q_{34} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{Q}}_1^T \\ \tilde{\mathbf{Q}}_2^T \\ \tilde{\mathbf{Q}}_3^T \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^T & q_{14} \\ \mathbf{q}_2^T & q_{24} \\ \mathbf{q}_3^T & q_{34} \end{bmatrix}$$

gdje su $\tilde{\mathbf{Q}}_i$ za $i=1,2,3$ 4x1 vektori, a $\tilde{\mathbf{Q}}_i^T = [\mathbf{q}_i^T, q_{i4}]$.

Svaki vektor predstavlja jednu ravninu. $\tilde{\mathbf{Q}}_3^T \tilde{\mathbf{M}} = 0$ odgovara točkama u ravnini slike kojima je $S = 0$, tj točkama u beskonačnosti. Iz toga se zaključuje da je to **fokusna ravnina**. Ravnine $\tilde{\mathbf{Q}}_1^T \tilde{\mathbf{M}} = 0$ i $\tilde{\mathbf{Q}}_2^T \tilde{\mathbf{M}} = 0$ odgovaraju točkama u ravnini slike tako da je $U=0$, tj. $V=0$. Presjecište tih ravnina je linija koja prolazi kroz optički centar kamere i ishodište u ravnini slike. Valja naglasiti da ta linija nije optička os jer nije okomita na fokusnu ravninu.

Sada ćemo navesti kako možemo iz matrice perspektivne projekcije dobiti neke korisne informacije o našem sustavu.

Optički centar kamere definiran je kao presjek tri ravnine $\tilde{\mathbf{Q}}_1^T \tilde{\mathbf{M}} = 0$, $\tilde{\mathbf{Q}}_2^T \tilde{\mathbf{M}} = 0$, i $\tilde{\mathbf{Q}}_3^T \tilde{\mathbf{M}} = 0$ i može se dobiti rješavanjem sustava od tri linearne jednadžbe

$$\tilde{\mathbf{P}} \begin{bmatrix} \mathbf{C} \\ 1 \end{bmatrix} = 0 \quad (2.16)$$

Ako matricu $\tilde{\mathbf{P}}$ zapišemo kao $[\mathbf{P} \quad \tilde{\mathbf{p}}]$ gdje je \mathbf{P} matrica 3x3, a $\tilde{\mathbf{p}}$ vektor 3x1 \mathbf{C} se može izračunati na slijedeći način

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} \mathbf{C} \\ 1 \end{bmatrix} = 0$$

$$\mathbf{P}\mathbf{C} + \tilde{\mathbf{p}} = 0$$

$$\mathbf{C} = -\mathbf{P}^{-1}\tilde{\mathbf{p}} \quad (2.17)$$

Iz matrice perspektivne projekcije možemo naći jednadžbu 3D linije (\mathbf{C}, m) određene pikselom m i optičkim centrom \mathbf{C} . Ta linija naziva se **optička zraka** definirana sa m . Iz prethodne analize znamo izraziti optički centar. Druga točka što nam je potrebna je točka u beskonačnosti, $\tilde{\mathbf{D}}$, čije su projekcijske koordinate dane s $[\mathbf{D}^T, 0]^T$. Ta je točka dana jednadžbom $\mathbf{D} = \mathbf{P}^{-1}\tilde{\mathbf{m}}$. Budući da je \mathbf{D} paralelan s vektorom $\mathbf{C}m$ (vektorom koji spaja točke \mathbf{C} i m) točka \mathbf{M} na liniji dana je s $\mathbf{M} = \mathbf{P}^{-1}(-\tilde{\mathbf{p}} + \lambda\tilde{\mathbf{m}})$ gdje λ poprima vrijednosti u intervalu $(-\infty, +\infty)$.

3 Korespondentne točke

Najvažniji i najteži problem za rekonstrukciju 3d objekata iz stereo slika je određivanje korespondentnih točaka – točaka na ravninama slika koje su nastale od iste fizičke točke (neke točke na objektu kojeg promatramo) . Pošto je vrlo teško odrediti koje su dvije točke na slikama nastale od iste 3d točke, pogotovo na objektima koji nemaju mnogo značajki (npr. jednolike teksture) potrebno je uočiti i iskoristiti ograničenja da bi smanjili područje pretrage.

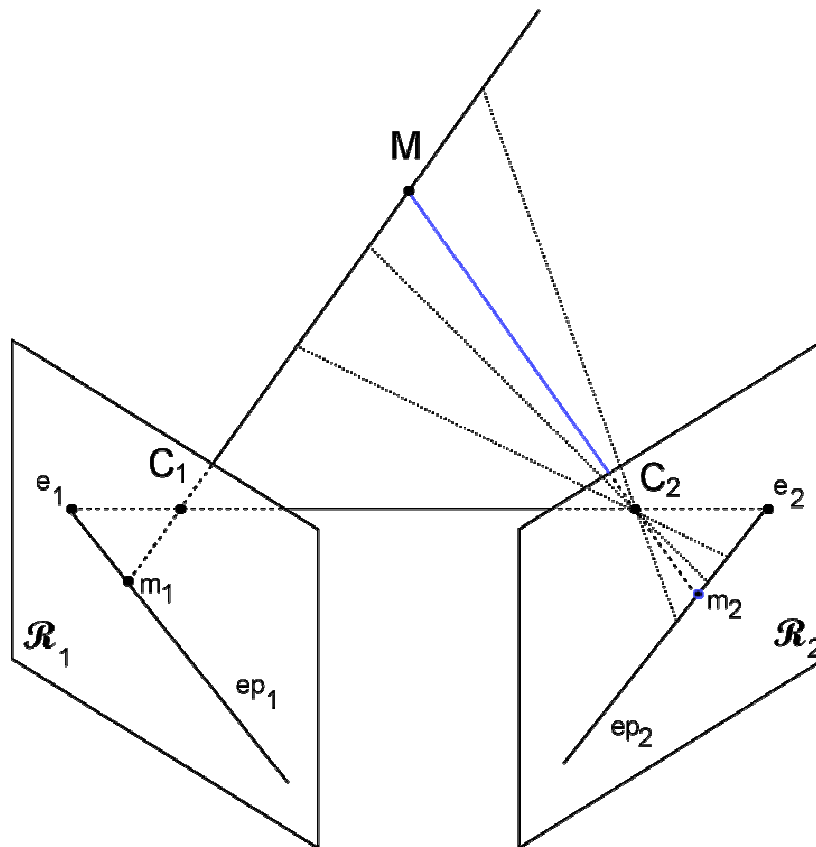
Ta ograničenja se mogu podijeliti na:

1. Geometrijska ograničenja nametnuta stereo sustavom . Najvažnije takvo ograničenje je epipolarno ograničenje zahvaljujući kojem možemo dvodimenzionalnu potragu pretvoriti u jednodimenzionalnu.
2. Geometrijska ograničenja koja potječu od objekata koje promatramo. Možemo pretpostaviti, na primjer, da njihova udaljenost od stereo sustava varira polako na cijeloj površini objekta.
3. Fizička ograničenja kao ona koja potječu od načina kako se osvjtljenje odražava na objektu.

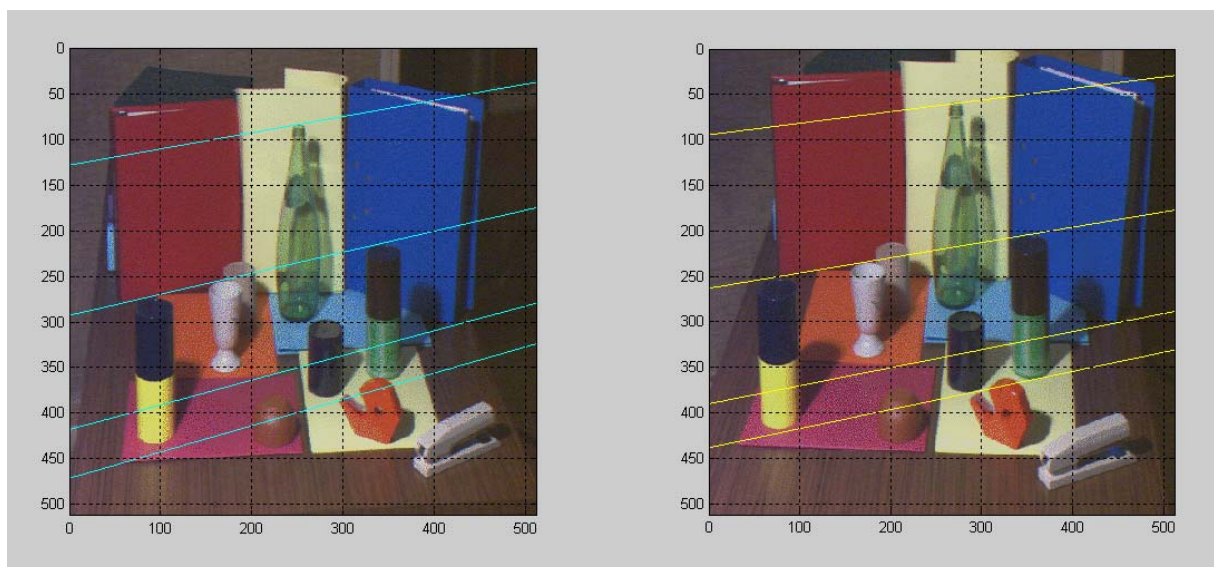
3.1 Epipolovi i epipolarne linije

Epipolarne linije su najvažnije ograničenje koje nam pojednostavljuje traženje korespondentnih točaka na stereo slikama. Pojmove epipol i epipolarna linija objasniti ćemo pomoću slike 3.1.

Ako izaberemo neku točku m_1 u ravnini slike R_1 , sve fizičke točke M koje su mogle 'proizvesti' tu točku nalaze na liniji koja prolazi kroz optički centar C_1 i točku m_1 . Sve točke u ravnini slike R_2 , koje mogu odgovarati točki m_1 nalaze se na slici te linije u ravnini R_2 . Ta slika je linija ep_2 i naziva se **epipolarna linija**.

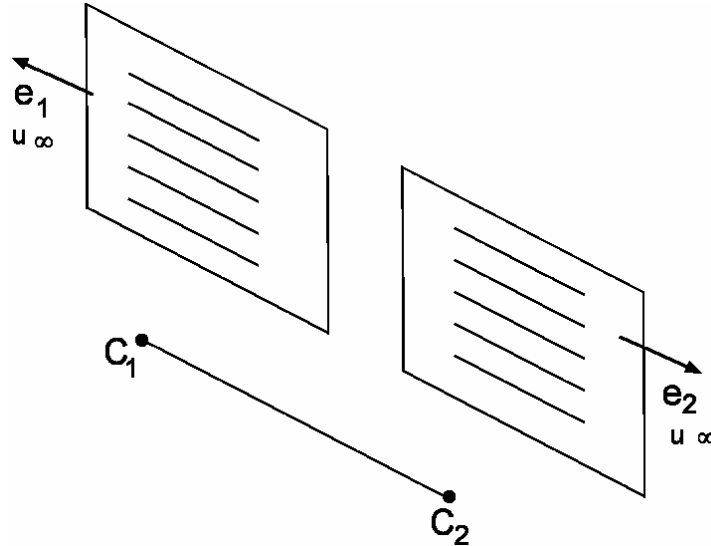


Slika 3.1 Epipolovi i epipolarne linije



Slika 3.2 Prikaz epipolarnih linija na stvarnim slikama

Epipol je slika (točka u ravnini slike) optičkog centra jedne kamere na ravnini slike druge kamere. Na slici su epipolovi označeni s e_1 i e_2 . U slučaju kada su ravnine slike R_1 i R_2 paralelne epipolovi su u beskonačnosti. Epipolarne linije su tada paralelne. To svojstvo sustava omogućava nam da dvodimenzionalni problem traženja korespondentnih točaka pretvorimo u jednodimenzionalan problem.



Slika 3.3 Ispravljene epipolarne linije

3.2 Izračunavnje epipolova

Koordinate dvaju optičkih centara C_i ($i = 1, 2$) možemo izračunati pomoću dva sustava linearnih jednadžbi (sustavi su dani formulom 2.15)

$$\tilde{\mathbf{P}}_i \tilde{\mathbf{M}} = 0 \quad \text{za } i = 1, 2$$

Budući da je svaki epipol e_i slika dobivena kamerom i svih optičkih centara C_j (s tim da je $j \neq i$), koordinate epipola e_i mogu se dobiti pomoću matrice $\tilde{\mathbf{P}}_i$ i vektora $\tilde{\mathbf{C}}_i$.

Nakon što smo odredili epipol, možemo odrediti i epipolarnu liniju.

Da bi odredili epipolarnu liniju potrebne su nam dvije točke. Jedna od njih je epipol e_2 koji je dan jednadžbom

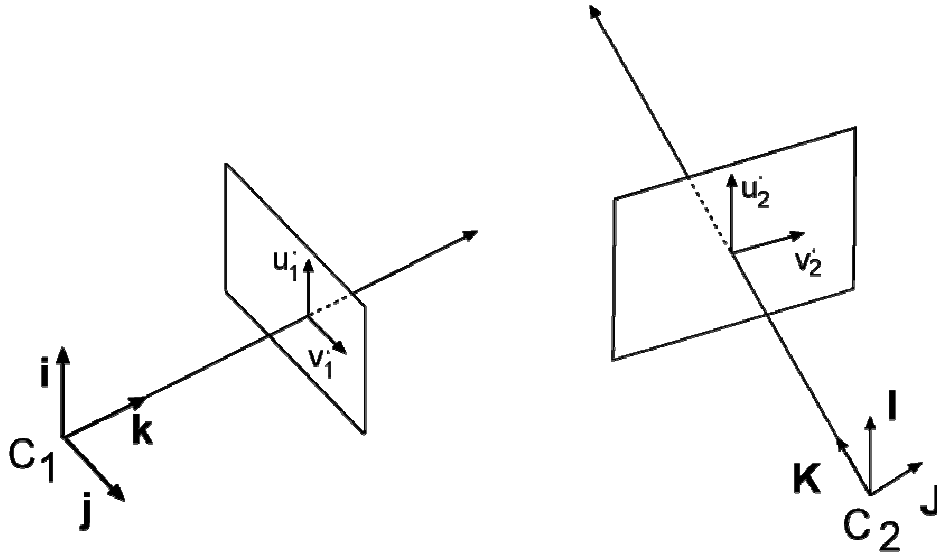
$$\tilde{\mathbf{e}}_2 = \tilde{\mathbf{P}}_2 \begin{bmatrix} -\mathbf{P}_1^{-1} \tilde{\mathbf{p}}_1 \\ 1 \end{bmatrix}$$

Druga točka je točka u beskonačnosti optičke zrake (C_1, m_1). Slika m_2 te točke u drugoj ravnini slike može se izračunati

$$\tilde{\mathbf{m}}_2 = \mathbf{P}_2 \mathbf{P}_1^{-1} \tilde{\mathbf{m}}_1$$

Budući da se linija koja prolazi kroz dvije točke predstavlja kao vektorski produkt te dvije točke epipolarna linija dana je s $\tilde{\mathbf{e}}_2 \times \tilde{\mathbf{m}}_2$

Ako pretpostavimo da koristimo normirani model kamere i ako je referentni koordinatni sustav C_1xyz iz slike 3.4 možemo izvesti dodatne zaključke o epipolarnim linijama.



Slika 3.4 Dva normirana koordinatna sustava

Prema prethodnom razmatranju promjene koordinatnog sustava scene uzмимо slijedeće jednadžbe (koje vrijede za sustav na slici 3.4):

$$\mathbf{C}_1\mathbf{C}_2 = \mathbf{t}$$

$$\mathbf{R}[\mathbf{I}, \mathbf{J}, \mathbf{K}] = [\mathbf{i}, \mathbf{j}, \mathbf{k}]$$

$$\tilde{\mathbf{K}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix}$$

Matrica $\tilde{\mathbf{P}}_1$ je jednaka $[\mathbf{I}_3 \ \mathbf{0}]$, gdje je \mathbf{I}_3 jedinična matrica 3x3. Matrica $\tilde{\mathbf{P}}_2$, iskazana u koordinatnom sustavu C_2XYZ jednaka je matrici $\tilde{\mathbf{P}}_1$, pa je zbog toga u koordinatnom sustavu C_1xyz jednaka

$$\tilde{\mathbf{P}}_1\mathbf{K}^{-1} = [\mathbf{R}^T \quad -\mathbf{R}^T\mathbf{t}]$$

Pomoću tih projekcijskih matrica možemo izračunati epipolove E_1 i E_2

$$\tilde{\mathbf{e}}_1 = \tilde{\mathbf{P}}_1 \begin{bmatrix} \mathbf{C}_1\mathbf{C}_2 \\ 1 \end{bmatrix} = \tilde{\mathbf{P}}_1 \begin{bmatrix} \mathbf{t} \\ 1 \end{bmatrix} = \mathbf{t}$$

$$\tilde{\mathbf{e}}_2 = \tilde{\mathbf{P}}_2 \begin{bmatrix} \mathbf{O}_3 \\ 1 \end{bmatrix} = -\mathbf{R}^T\mathbf{t}$$

Ako izaberemo piksel m_1 predstavljen s $\tilde{\mathbf{m}}_1 = [u'_1, v'_1, 1]^T$ njegova epipolarna linija je

$$\tilde{\mathbf{e}}_2 \times \mathbf{P}_2 \mathbf{P}_1^{-1} \tilde{\mathbf{m}}_1 = -\mathbf{R}^T \mathbf{t} \times \mathbf{R}^T \tilde{\mathbf{m}}_1$$

Isto možemo izračunati za m_2

$$\tilde{\mathbf{e}}_1 \times \mathbf{P}_1 \mathbf{P}_2^{-1} \tilde{\mathbf{m}}_2 = \mathbf{t} \times \mathbf{R}^T \tilde{\mathbf{m}}_2$$

3.3 Ispravljanje epipolarnih linija

Da bi ispravili epipolarne linije, tj. da bi one bile u paralelne na obje slike, moramo poznavati matrice perspektivne projekcije. Budući da smo kalibrirali sustav, poznate su nam vrijednosti unutrašnjih i vanjskih parametara kamera, a iz tih parametara, kao što smo ranije pokazali, možemo izračunati te matrice.

Osnovna ideja postupka je pronalaženje (izračunavanje) novih matrica $\tilde{\mathbf{P}}_{n1}$ i $\tilde{\mathbf{P}}_{n2}$, tj. pronalaženje pozicije i orijentacije kamera za koje bi epipolovi bili u beskonačnosti, a epipolarne linije paralelne. Pošto imamo dva para matrica perspektivne projekcije, možemo govoriti o dva sustava kamera, kamerama prije transformacija i novim kamerama, tj. kamerama poslije transformacija.

Da bi osigurali paralelnost epipolarnih linija, potrebno je zarotirati svaku kameru oko njezina optičkog centra tako da im fokusne ravnine budu jednake (a time će fokusne ravnine sadržavati i liniju koja spaja optičke centre). Da bi dobili vodoravne epipolarne linije , pravac koji spaja optičke centre mora biti paralelan s novom X osi kamera. Iste y vertikalne koordinate istih točaka na obje slike možemo osigurati s zahtjevom da nove kamere imaju iste unutarnje parametre.

Ukratko, pozicije kamera (optičkih centara) jednake su i kod novih kamera, orijentacija (ista kod obje kamere) razlikuje se za odgovarajuću rotaciju; unutrašnji parametri isti su za obje kamere.

Nove matrice perspektivne projekcije možemo zapisati tako da iz njih možemo pročitati unutarnje i vanjske parametre.

$$\tilde{\mathbf{P}}_{n1} = \mathbf{K} * [\mathbf{R} \quad -\mathbf{R}\mathbf{c}_1]$$

$$\tilde{\mathbf{P}}_{n2} = \mathbf{K} * [\mathbf{R} \quad -\mathbf{R}\mathbf{c}_2]$$

Unutarnji parametri \mathbf{K} isti su za obje kamere i proizvoljni su. Matrica orijentacija \mathbf{R} također je ista za obje kamere, a konstruirat ćemo je od vektora koji će predstavljati X, Y i Z osi koordinatnog sustava kamere izraženog u koordinatnom sustavu scene.

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \mathbf{r}_3^T \end{bmatrix}$$

U skladu s prethodnim razmatranjima o svojstvima novih kamere, retci rotacijske matrice su

$$\mathbf{r}_1 = \frac{\mathbf{c}_1 - \mathbf{c}_2}{\|\mathbf{c}_1 - \mathbf{c}_2\|}$$

(nova X os paralelna s pravcem koji spaja optičke centre)

$$\mathbf{r}_2 = \mathbf{k} \times \mathbf{r}_1$$

(nova Y os okomita je na X os i na \mathbf{k} , gdje je \mathbf{k} proizvoljni jedinični vektor kojim potpuno određujemo Y os u ravnini okomitoj na X. U našem slučaju koristit ćemo jedinični vektor u smjeru Z osi lijeve kamere)

$$\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$$

(nova Z os mora biti okomita na X i Y)

Da bi ispravili sliku potrebno je odrediti transformaciju koja preslikava ravninu slike stare kamere u ravninu slike nove kamere. Tu transformaciju možemo dobit na slijedeći način:

- za svaku 3D točku vrijedi

$$\begin{aligned}\tilde{\mathbf{m}}_{stara} &= \tilde{\mathbf{P}}_{stara} \tilde{\mathbf{M}} \\ \tilde{\mathbf{m}}_{nova} &= \tilde{\mathbf{P}}_{nova} \tilde{\mathbf{M}}\end{aligned}$$

Budući da ne mijenjamo optičke centre, jednadžbe optičkih zraka su:

$$\mathbf{M} = \mathbf{c}_1 + \lambda_0 \mathbf{P}_{stara}^{-1} \tilde{\mathbf{m}}_{stara}$$

$$\mathbf{M} = \mathbf{c}_1 + \lambda_1 \mathbf{P}_{nova}^{-1} \tilde{\mathbf{m}}_{nova}$$

iz toga proizlazi jednadžba za točku u ravnini slike nove kamere

$$\tilde{\mathbf{m}}_{nova} = \lambda \mathbf{P}_{nova} \mathbf{P}_{stara}^{-1} \tilde{\mathbf{m}}_{stara}$$

Pomoću te jednadžbe možemo odrediti dvije 3x3 matrice, \mathbf{T}_{LIJEVA} i \mathbf{T}_{DESNA} kojima množimo piksele lijeve i desne slike kako bi dobili nove, ispravljene slike.

$$\mathbf{T}_{LIJEVA} = \mathbf{P}_{nova1} \mathbf{P}_{stara1}^{-1} \quad (3.1)$$

$$\mathbf{T}_{DESNA} = \mathbf{P}_{nova2} \mathbf{P}_{stara2}^{-1} \quad (3.2)$$

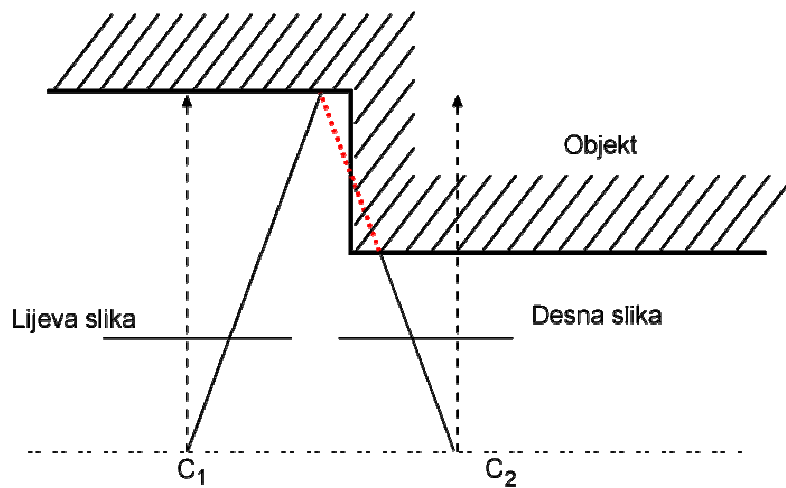
3.4 Određivanje korespodentnih točaka

Ispravljanjem epipolarnih linija, smanjili smo dimenziju pretraživanja korespodentnih točaka. Budući da je pretraživanje po boji jednog piksela na liniji besmisleno, razvijeno je mnogo metoda za rješavanje ovog problema. Najviše korištene metode možemo podijeliti u dvije grupe: u prvoj grupi nalaze se metode koje pretražuju sliku po nekim značajkama (npr. rubovima, linijama, krivuljama, ...) dok su u drugoj grupi metode koje uspoređuju područja slike samo na temelju boja (ili nijansi sive).

Princip rada metoda iz druge grupe temelji se na uspoređivanju područja oko neke točke na lijevoj slici s područjem oko točke na istoj liniji pretraživanja na desnoj strani. U nastavku teksta koristit ćemo se nazivom 'prozor' jer je to područje u pravilu pravokutnog oblika.

Prije nego što objasnimo način rada metoda za pronalaženje korespodentnih točaka, potrebno je obratiti pažnju na poteškoće koje se mogu pojaviti i koje će nam otežat zadatak.

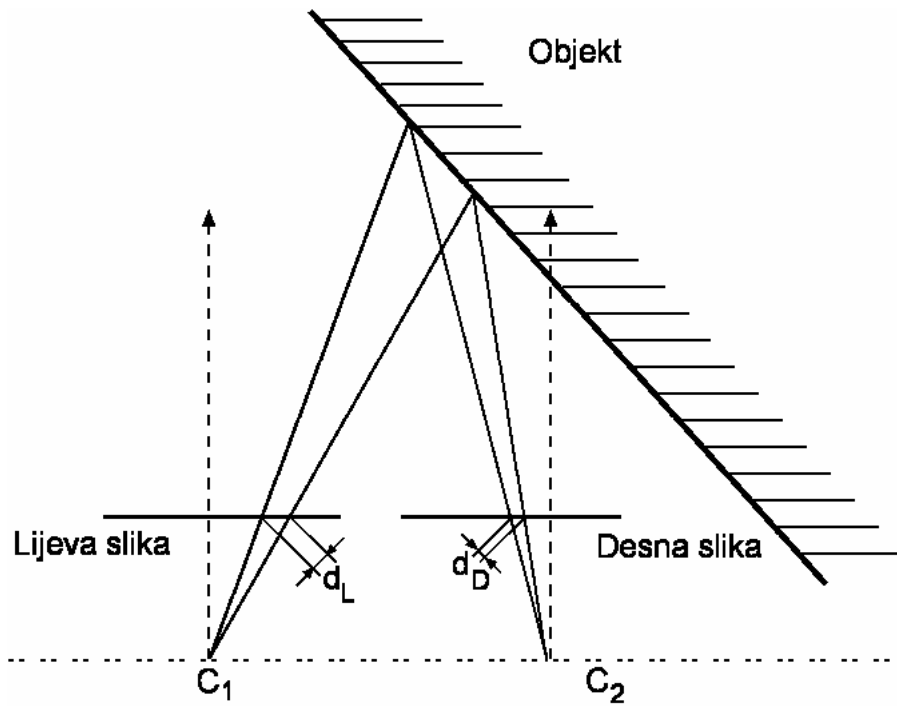
1. **Zaklanjanje** (engl. occlusion) – točke na jednoj slici ne odgovaraju niti jednoj točki na drugoj slici, npr. snimani objekt zaklanja 'pogled' jednoj kameri, a drugoj ne (slika 3.5 i slika 3.6)
2. **Šum** – na jednoj slici može postojati šum koji nije prisutan na drugoj
3. **Fotometrična distorzija** (engl. photometric distorsion) – boja (ili svjetlina) piksela projicirane trodimenzionalne točke različita je na dvije slike (dogđa se zbog različitog kuta refleksije svjetla s objekta) (slika 3.6)
4. **Projekcijska distorzija** – projicirani oblici nisu identični na obje slike
 - (1) kako se površina udaljuje od posmatrača, smanjuje se pa s time i tekstura na njoj postaje iskrivljena
 - (2) ako je površina nagnuta, tekstura na projiciranoj slici postaje stisnuta u smjeru nagiba (slika 3.7 i slika 3.8)
5. **Područje bez teksture** – u područjima gdje nema razlika u intenzitetu imat ćemo iste mjere sličnosti u cijelom području. *Višestruke korespodentne točke* možemo naći kada ne postoji tekstura u području većem od veličine prozora.



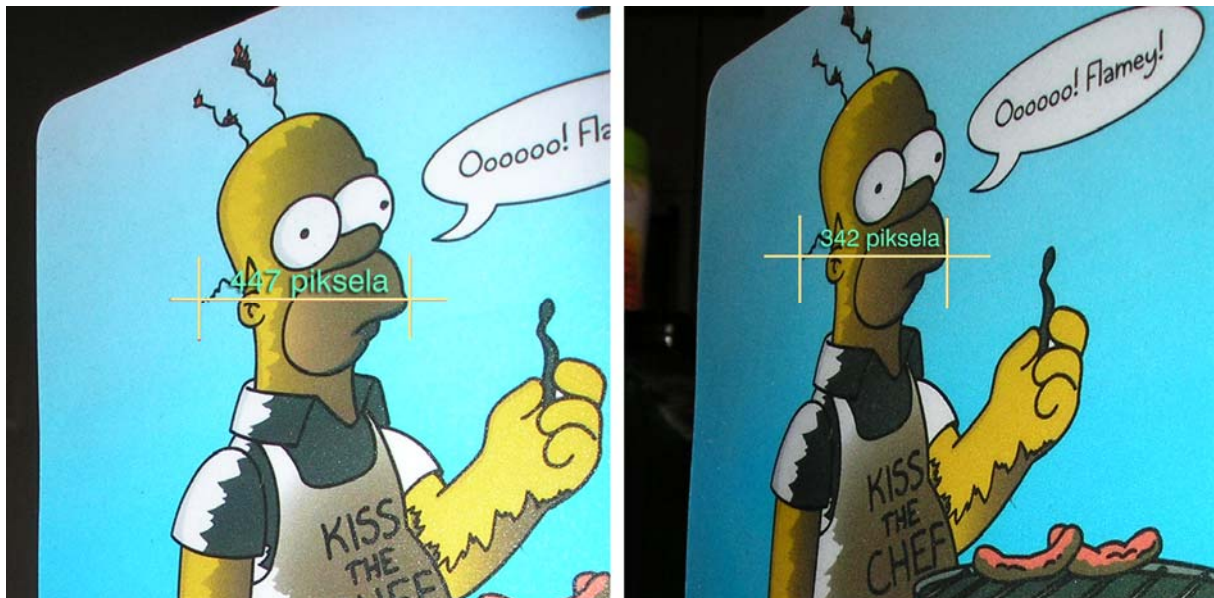
Slika 3.5 Grafički prikaz zaklanjanja



Slika 3.6 Zaklanjanje (lijevo je slika lijeve kamere, na toj se slici vidi tekst koji je zaklonjen objektom na desnoj slici)



Slika 3.7 Perspektivna distorzija (područje iste veličine projicira se u različite veličine d_L i d_R na različitim slikama)



Slika 3.8 Zbog nagiba površine Homerova je glava na desnoj slici stisnuta (lijeva slika snimljena je lijevom, a desna desnom kamerom)

3.5 Metode pretraživanje područja slike (engl. area-based methods)

Budući da ove metode uspoređuju područja na lijevoj i desnoj slici, potrebno je navesti načine na koji se dijelovi slike mogu uspoređivati.

Pri usporedbi u praksi se najčešće koristi suma kvadrata razlike (engl. SSD – sum of squared difference)

$$c(x, y, d) = \sum_{(h,w)} [\text{Im}_L(x+h, y+w) - \text{Im}_D(x+h+d, y+w)]^2 \quad (3.3)$$

gdje je Im_L lijeva slika, Im_D desna slika, $h = [-n, n]$ visina prozora, a $w = [-m, m]$ širina prozora, a d označava pomak.

Navedena formula daje ispravne rezultate samo kada su intenziteti (razine sive) piksela neke 3D točke jednaki na obje slike. Ako to nije slučaj, potrebno je koristiti normiranu inačicu ove jednadžbe:

$$c_{NORM}(x, y, d) = \sum_{(h,w)} \left[\frac{\text{Im}_L(x+h, y+w) - \overline{\text{Im}_L}}{\sqrt{\sum_{(h,w)} (\text{Im}_L(x+h, y+w) - \overline{\text{Im}_L})^2}} - \frac{\text{Im}_D(x+h+d, y+w) - \overline{\text{Im}_D}}{\sqrt{\sum_{(h,w)} (\text{Im}_D(x+h+d, y+w) - \overline{\text{Im}_D})^2}} \right]^2$$

gdje je $\overline{\text{Im}}$ srednja vrijednost intenziteta u prozoru.

Osim SSD može se koristiti i suma apsolutnih razlika (engl. SAD – sum of absolute differences). Opći oblik ove jednadžbe je

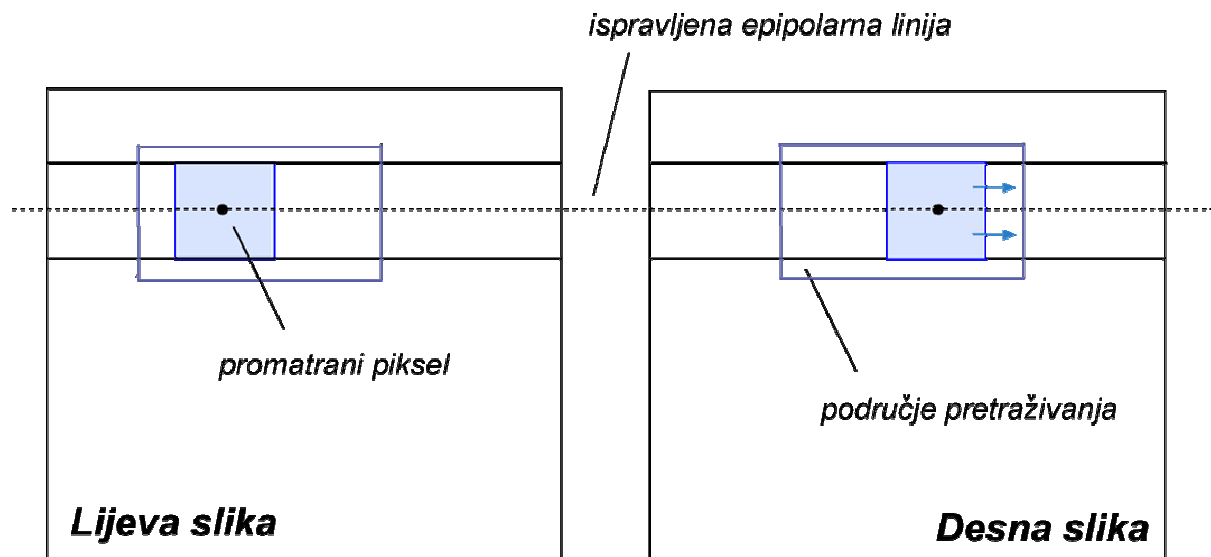
$$c(x, y, d) = \sum_{(h,w)} |\text{Im}_L(x+h, y+w) - \text{Im}_D(x+h+d, y+w)| \quad (3.4)$$

Ako želimo uspoređivati slike u boji, ovu jednadžbu možemo preformulirati u

$$c(x, y, d) = \sum_{(h,w)} \left[|R_L(x+h, y+w) - R_D(x+h+d, y+w)| + |G_L(x+h, y+w) - G_D(x+h+d, y+w)| + |B_L(x+h, y+w) - B_D(x+h+d, y+w)| \right]$$

3.6 Pretraživanje pomoću prozora

Osnova algoritma za pretraživanje (tj. za uparivanje) točaka na slikama je funkcija za usporedbu određenog broja piksela na nekom području. Neke od njih dane su u prethodnom poglavlju. Odabranom funkcijom računamo stupanj podudarnosti prozora na lijevoj slici s područjem na desnoj slici.



Slika 3.9 Traženje podudarnih područja

Prije nego što opišemo algoritam, potrebno je definirati neke pojmove (slika 3.9).

Promatrani piksel – piksel kojeg želimo pronaći na drugoj slici

Prozor usporedbe – pravokutnik oko promatranog piksela koji uspoređujemo s pravokutnikom istih dimenzija u drugoj slici

Područje pretraživanja – područje u kojem se pretražuje (veličina ovog područja zapravo određuje najveći pomak koji možemo naći)

Pretraživanje se vrši na slijedeći način:

Svaki piksel na lijevoj slici pokušavamo upariti s pikselom na desnoj slici. Pretpostavlja se da su slike rektificirane, tj. da su im epipolarne linije paralelne i da imaju iste vertikalne koordinate. Ulazni parametri algoritma su, osim dvije slike, i veličine prozora i područja pretraživanja. Valja naglasiti da je izbor ovih veličina važan za konačni rezultat. Ako npr. izaberemo premali prozor na rezultat će u velikoj mjeri utjecati šum na slikama, dok s druge strane, ako izaberemo preveliki moći ćemo rekonstruirati manji broj točaka (prozor će djelovati kao filter). Izbor prevelike širine područja pretraživanja rezultirat će s prevelikim brojem podudarnih točaka (pogotovo ako tekstura ima ponavljajući uzorak), dok će izbor premale veličine rezultirati s malim brojem pronađenih parova (pogotovo ako je razmak

između optičkih centara velik, tada će i razmak među točkama na slici biti velik pa uz premalu veličinu područja pretraživanja te točke neće biti pronađene).

Dakle, za svaku točku na epipolarnoj liniji lijeve slike odredimo prozor koji ju okružuje. Isto takav prozor odredimo na desnoj slici. Za razliku od prozora na lijevoj slici koji je fiksiran promatranom točkom, desni prozor se pomiče po području pretraživanja. Za svaku kombinaciju ta dva prozora računamo vrijednost funkcije koja određuje njihovu podudarnost. Najbolja podudarnost ostvarena je u minimumu te funkcije (razmak između točaka je razmak između pozicije prozora za koje je postignut minimum).

Funkcija za izračunavanje podudarnosti može imati više točaka minimuma, npr. u području bez teksture. Tada se ne može točno odrediti koje točka međusobno odgovaraju pa se zbog točnosti dobivenih podataka izračunati razmak odbacuje (uvijek je bolje imati rjeđe, ali točnije podatke).

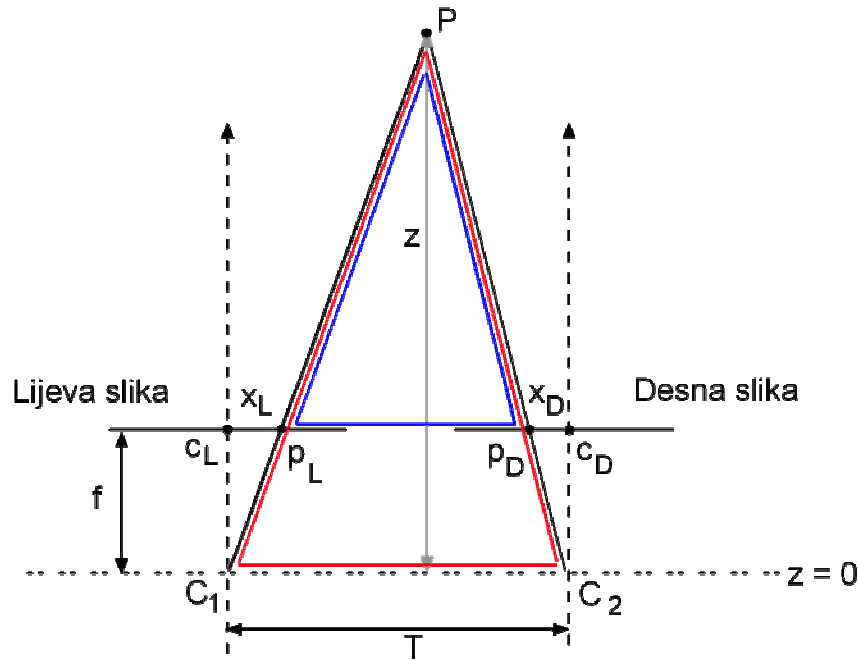
Dodatna preciznost postupka (tj. ne prihvaćanje nesigurnih rezultata) može se utvrditi slijedećim postupkom: nakon što se za neku lijevu točku odredi podudarna točka u desnoj slici postupak se može obrnuti, tako da za pronađenu točku na desnoj slici tražimo podudarnu točku na lijevoj. Iz svojstva stereo slike da svakoj točki na lijevoj slici mora odgovarati samo jedna točka na desnoj slici može se zaključiti da obrnutim postupkom moramo dobiti promatranu točku na lijevoj slici. U slučaju da nismo dobili identične piksele, zaključujemo da ta dva piksela nisu podudarna pa se njihov razmak izbacuje iz skupa rezultata.

Pretraživanjem cijele slike dobit ćemo dvodimenzionalnu matricu malo manju od veličine slika (zbog veličine prozora) s razmacima koji pripadaju pikselima na tim koordinatama. Iz tih razmaka, na način koji će biti opisan u slijedeće poglavlju, možemo izračunati z koordinatu pojedine točke. Zbog toga se 3d scena dobivena stereovidom često prikazuje slikom u kojoj je razmak kodiran sivim nijansama.

4 3D rekonstrukcija

Iz izračunatih pomaka piksela na desnoj slici u odnosu na položaj u lijevoj možemo vrlo jednostavno izračunati z koordinatu, tj izračunati koordinate 3D točke. Jedini uvijet je poznavanje udaljenosti optičkih centara dviju kamera.

Jednadžbu za izračunavanje z koordinate možemo izvesti ako naš sustav prikažemo kao na slici 4.1



Slika 4.1 Izračunavanje z koordinate iz pomaka

Uzmimo bazu crvenog trokuta i podijelimo ju s visinom istog trokuta. Rezultat je isti kao kad bazu plavog trokuta podijelimo s njegovom visinom. z koordinata točke P identična je visini crvenog trokuta.

Ako su x_L i x_D koordinate točaka p_L i p_D gledane od c_L odnosno c_D (po tome je x_L pozitivno, a x_D negativno) vrijedi:

$$\frac{T}{z} = \frac{T+d}{z-f} \quad \text{gdje je } d = x_D - x_L$$

$$zT - fT = (T+d)z$$

$$z(T - T + d) = fT$$

$$\boxed{z = f \frac{T}{d}} \quad (4.1)$$

5 Programska implementacija

5.1 Opis sustava

Korišteni sustav sastoji se od dva Olympus C-350 ZOOM digitalna fotoaparata, pričvršćena na podlogu jedan pored drugoga i orijentirani tako da su im optičke osi skoro paralelne.



Slika 5.1 *Stereo sustav*

5.2 Kalibracija kamera

Kalibracijom kamere određujemo njene intrinzične i ekstrinzične parametre te koeficijente radijalne distorzije.

Postoje više načina kalibracije kamere :

- pomoću 3d uzorka
- pomoću ravninskog uzorka
- 'autokalibracijom' ...

Kalibracija 2d uzorkom je najjednostavnija i najjeftinija. Potreban nam je samo laserskim pisačem iscrtani uzorak pričvršćen na tvrdu podlogu (npr. na karton ili staklo). Kao uzorak najčešće se koriste ili crno-bijeli kvadrati (šahovska ploča) ili krugovi.

Kalibracija pomoću ravninskog uzorka temelji se na prepoznavanju istih točaka (rubovi kvadrata ili težišta krugova) na više slika. Pomoću pozicija tih točaka na različitim slikama određuju se unutarnji parametri kamere i rotacija i pomak u odnosu na referentni koordinatni sustav. Prilikom fotografiranja važno je da stalno mijenjamo kutove uzorka prema kameri.

Za kalibraciju kamera koje su dio stereo sustava korišten je program *Matlab*. U nastavku ćemo opisati korake potrebne da se iz nekoliko fotografija ravninskog uzorka, pomoću *Matlaba*, odrede parametri kamere.

Učitavanje slika

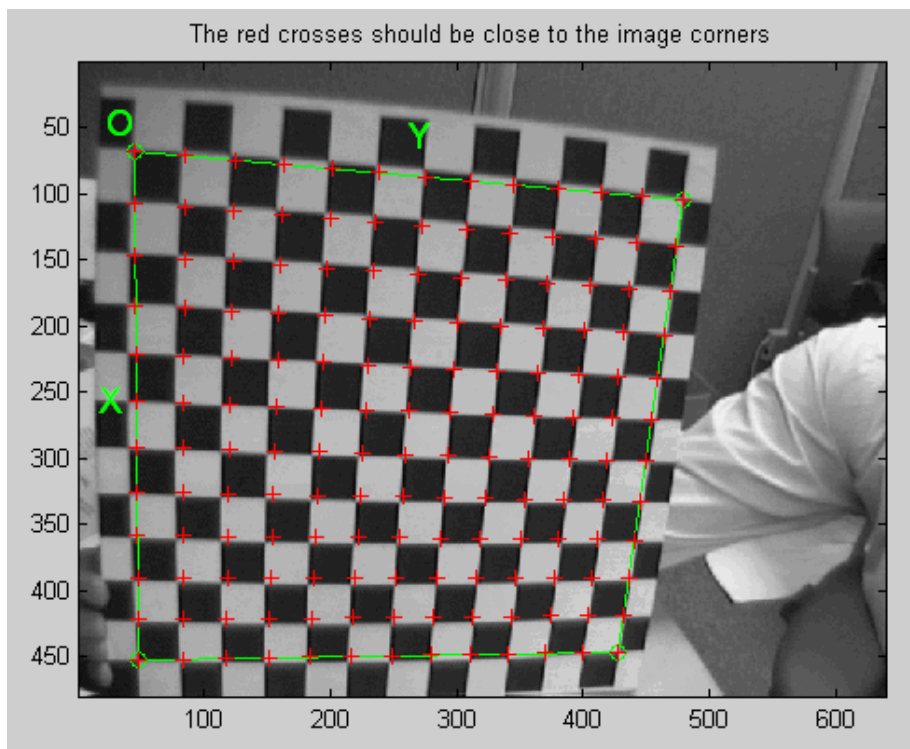
Slike ravninskog uzorka slikane pod različitim kutovima potrebno je snimiti u neki direktorij. Navođenjem prvog dijela imena koji je isti za sve slike i za načine zapisa, u memoriju se automatski učitavaju sve slike iz odabranog direktorija s tim parametrima (npr. ako upišemo *Slika* i odaberemo tip *tiff* učitavamo *Slika0.tiff*, *Slika1.tiff*, ...).

Odabir rubova

Za kalibraciju važno je da aplikacija točno odredi svaki rub kvadrata s slike. Potrebno je odrediti četiri rubne točke ručno, a aplikacija će automatski izabrati ostale točke. Ako želimo kalibrirati dvije kamere koje su dio stereo sustava pri odabiru rubnih točaka moramo paziti da je smjer odabira isti, jer bi se u protivnom sparile različite točke na slikama i ne bi bilo moguće odrediti prostorni odnos među kamerama.

Određivanje koeficijenata distorzije

Nakon što su odabrani kutevi svih kvadrata, *Matlab* na slici prikazuje poziciju svih točaka (slika 5.2). Zbog prisustva distorzije, ponekad postoje veća odstupanja između stvarnih rubova i onih koje je *Matlab* izabrao pa je potrebno upisati pretpostavljeni koeficijent distorzije. *Matlab* tada ponovo odabire rubove. Ako je odstupanje i dalje veliko, ponavljamo proces dok nismo zadovoljni.



Slika 5.2 Odabir rubova kalibracijskog uzorka

Računanje parametara kamere

Pritiskom na tipku Calibration izračunavaju se parametri kamere i prikazuju se u ovom obliku:

```

Focal Length:      fc = [ 661.67001  662.82858 ] ± [ 1.17913  1.26567 ]
Principal point:   cc = [ 306.09590  240.78987 ] ± [ 2.38443  2.17481 ]
Skew:              alpha_c = [ 0.00000 ] ± [ 0.00000 ]
Distortion:        kc = [ -0.26425  0.22645  0.00020  0.00023  0.00000 ]
Pixel error:       err = [ 0.45330  0.38916 ]

```

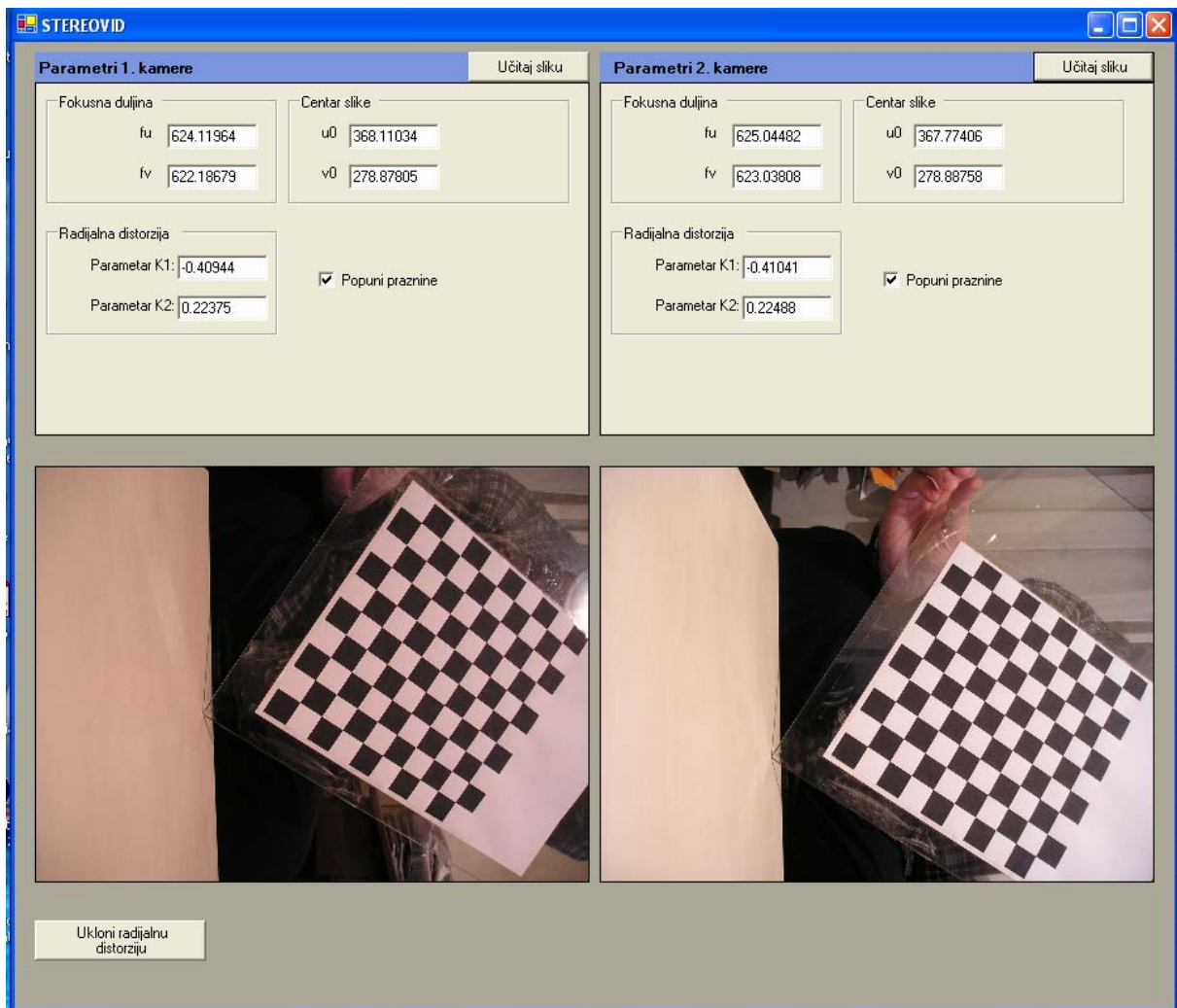
Slika 5.3 Unutrašnji parametri kamere

Vanjski parametri

Prilikom kalibriranja stereo sustava, osim unutarnjih parametara kamera program računa i vanjske parametre sustava, orijentaciju i položaj kamera (pri tom je jedna kamera u ishodištu koordinatnog sustava, a položaj druge dan je u odnosu na nju).

5.3 Uklanjanje radijalne distorzije

Za uklanjanje radijalne distorzije izrađen je program u programskom jeziku C++. Ulazni parametri programa su unutarnji parametri kamere, naziv ulazne slike i naziv ispravljene slike. Pored tih osnovnih parametara postoje još dva posebna ulazna parametra kojima biramo da li želimo prikazati ispravljenu sliku i da li želimo popuniti šupljine nastale zbog ispravljanja slike. Osim osnovnog programa kojem je jedina funkcija da na temelju parametara generira ispravljenu sliku, izrađen je dodatan program s pristupačnijim sučeljem.



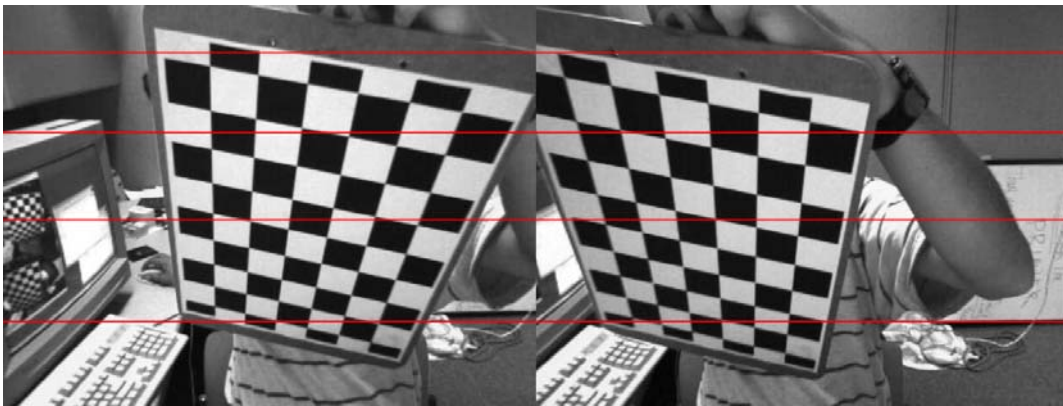
Slika 5.4 Program za uklanjanje radijalne distorzije

Ovim programom možemo jednostavno upisati parametre kamere, učitati dvije slike i ispraviti ih. Prilikom startanja programa parametri kamere se učitavaju iz xml datoteke što nam olakšava ispravljanje slika ako koristimo uvijek isti, kalibrirani, stereo sustav.

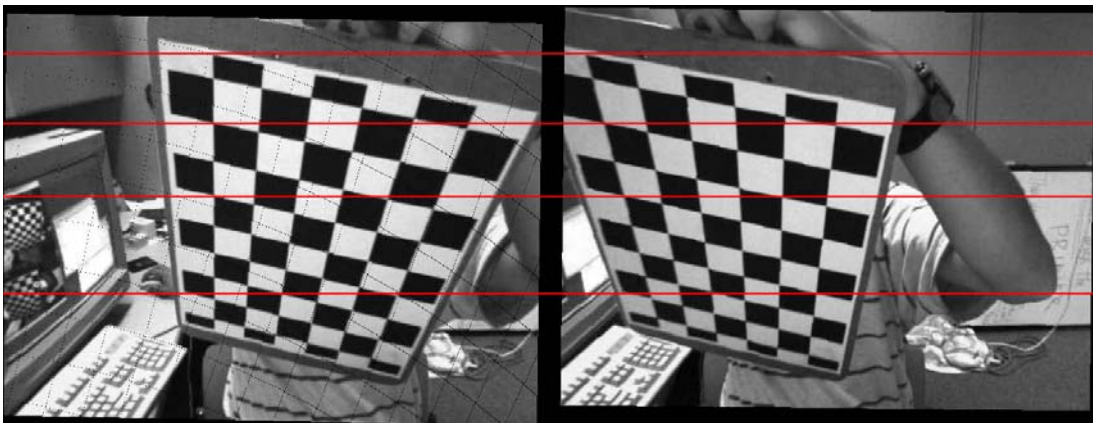
5.4 Ispravljanje epipolarnih linija

Program (ili bolje rečeno funkcije) za ispravljanje epipolarnih linija pisane su u programskom jeziku C (i prevedene da se mogu pozivati iz Matlab-a). Razlog programiranja u C-u je u načinu na koji se izvršava kod pisan u Matlab-u. Naime, Matlab svaku liniju koda prevodi u toku izvršavanja pa su zbog toga velike petlje (petlje s puno iteracija) pisane u Matlab-u jako spore.

Epipolarne linije su ispravljene na način opisan u poglavlju 3.3.



Slika 5.5 Slike prije ispravljanja



Slika 5.6 Slike nakon ispravljanja

5.5 Prepoznavanje korespondentnih točaka

Program za prepoznavanje korespondentnih točaka izrađen je pomoću Matlab-a. Zbog velike količine podataka koje je potrebno obraditi, napisana je funkcija u C-u koja je prevedena tako da se može pozivati iz Matlab-a.

Program se temelji na algoritmu danom u poglavlju 3.6. Ulazni parametri su lijeva i desna slika, veličina prozora i veličina područja na kojem se pretražuju korespondentne točke. Izlazni parametar je matrica čije vrijednosti predstavljaju pomak pojedine točke lijeve slike u odnosu na desnu (njen položaj u desnoj).

Iako je program napisan u C-u, problem u brzini (sporosti) izvođenja time nije riješen. Za manje slike i manja područja pretrage vrijeme izvođenja bilo je prihvatljivo (desetak sekundi bilo je potrebno da se obradi slika 255x255 s područjem pretrage 15x15 piksela), ali za veće površine (slike ili područja pretrage) brzina je naglo padala pa je za sliku dimenzija 1024x1024 trebalo i nekoliko sati. Razlog sporosti bio je izvođenje velikog broja operacija oduzimanja i velik broj poziva funkcije za određivanje apsolutne vrijednosti. Detaljnija analiza pokazala je da bi se broj operacije mogao smanjiti korištenjem podataka izračunatih u prethodnom koraku.

U optimiranoj verziji programa suma apsolutnih razlika izračunava se unaprijed za svaki prozor na lijevoj slici i za svaku udaljenost od tog prozora na desnoj. Za prvi redak se izračunavaju sve sume, ali za sve ostale retke koriste se već izračunate sume (suma prozora koji prelazi u redak ispod sebe razlikuje se samo za novi donji redak koji moramo dodati staroj sumi i za stari gornji redak kojeg trebamo oduzeti od te iste sume). Na taj način riješili smo se mnogo operacija oduzimanja.

Osim što smo ovim postupkom vidno ubrzali aplikacija, dobili smo i 'alat' kojim možemo ostvariti još jedno poboljšanje programa. Ovaj put poboljšanje se ne odnosi na brzinu, već na točnost postupka. Unaprijed izračunavanjem suma apsolutnih vrijednosti i pohranjivanjem izračunatih vrijednosti u matricu omogućili smo provedbu ograničenja koje kaže da piksel na jednoj slici može odgovarati samo jednom pikselu na drugoj slici. Ovo ograničenje se 'provodi' tako da nakon što nađemo minimum na nekoj udaljenosti od promatranog piksela koji se nalazi na lijevoj slici, obrnutim pretraživanjem (promatrani piksel je sada s desne strane) moramo dobiti promatrani piksel na lijevoj strani. Ako ne dobijemo iste vrijednosti velika vjerojatnost je da se jedan od tih piksela nalazi u zaklonjenom dijelu scene (slika 3.7)

Iz rezultata ovog programa (slike 5.7 do 5.12) može se primijetiti prilična neučinkovitost ovog postupka. Za sintetizirane stereo slike (slika 5.11) rezultati izgledaju dobro, ali u stvarnoj sceni, u prisustvu perspektivne i fotometrične distorzije algoritam je pokazao sve svoje nedostatke i slabosti.

U literaturi su dane modifikacije osnovnog algoritma koje rješavaju neke od navedenih problema. One uvode promjene u veličinu prozora i proširuju pretragu na više prozora, ali i povećavaju kompleksnost, a time i vrijeme izvođenja programa.



Slika 5.7 Lijeva i desna slika scene



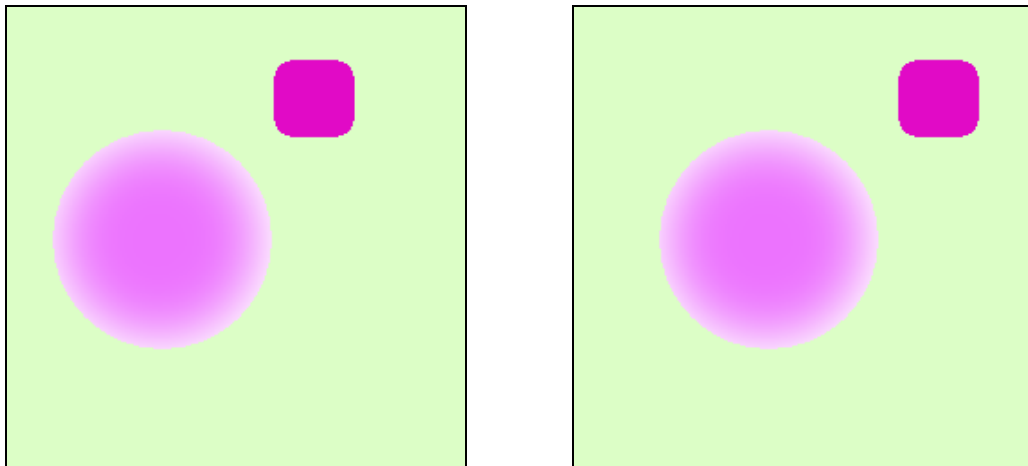
Slika 5.8 Prikaz pomaka u slici



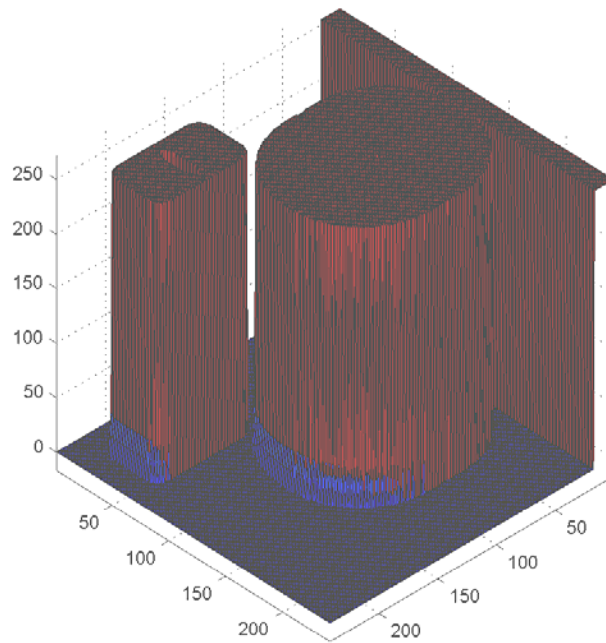
Slika 5.9 *Lijeva i desna slika scene*



Slika 5.10 *Prikaz pomaka u slici*



Slika 5.11 *Lijeva i desna slika scene*



Slika 5.12 *3d prikaz scene*

6 Zaključak

Stereo vid je vrlo kompleksna metoda dobivanja 3d modela. Iako se na prvi pogled sastoji od dijelova koje je jednostavno realizirati, skoro u svakom dijelu izrade aplikacije naišli smo na poteškoće zbog fizičkih ograničenja sustava (njegove nesavršenosti), procesorskih zahtjeva i zahtjeva za velikom količinom memorije.

U ovom radu pokušali smo opisati jedan stereo sustav – njegove fizičke značajke i algoritme koji se koriste u rekonstrukciji scene. Korištene metode nisu najbolje i najbrže, ali kao što se vidi iz rezultata, pomoću njih smo ipak dobili podatke o izgledu promatranih (snimljenih) objekata. Budući da smo čak i s vrlo ograničenim resursima dobili rezultate koji su dovoljno dobri za neke primjene, jasno je da ova metoda ima velike potencijale i da će se već u skoroj budućnosti naći u brojnim sustavima kojima je za rad potrebno znanje o izgledu svijeta (scene) koji ih okružuje ili s kojim su u interakciji.

7 Literatura

- [1] O. Faugeras, Three-Dimensional Computer Vision, A Geometric Viewpoint, 1993.g.
- [2] S. Birchfield, An Introduction to Projective Geometry (for computer vision), 1998.g.
- [3] A. Fusiello, E. Trucco, A.Verri, Rectification with Unconstrained Stereo Geometry, RESERCH MEMORANDUM,
- [4] C. Loop, Z.Zhang, Computing Rectifying Homographies for Stereo Vision, Technical report MSR-TR-99-21, 1999.g.
- [5] K.Muehlmann, D.Maier, J.Hesser, R.Maenner, Calculating Dense Disparity Maps from Color Stereo Images, an Efficient Implementation
- [6] I. J. Cox, S. L. Hingorani, S. B. Rao, A Maximum Likelihood Stereo Algorithm, Computer Vision and Image Understanding , 63:3, pp.542-567, May 1996
- [7] A. Fusiello, V.Roberto, E. Trucco, Efficient Stereo with Multiple Windowing, 1997.g.
- [8] J.P.Lewis, Fast Normalized Cross-Correlation
- [9] Di Stefano, M. Marchionna, S.Mattoccia, G.Neri, A Fast Area-Based Stereo Matching Algorithm
- [10] C. Sun, Fast Stereo Matching Using Subregioning and 3D Maximum-Surface Techniques, International Journal of Computer Vision, vol. 47, pp.99-117, May 2002
- [11] A.K. Jain, Fundamentals of Digital Image Processing, 1989