

# Usporedba metoda za kalibraciju kamera

---

Šarlija, Dino

Undergraduate thesis / Završni rad

2020

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Department of Mathematics / Sveučilište Josipa Jurja Strossmayera u Osijeku, Odjel za matematiku**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:126:447780>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-07-12**



Repository / Repozitorij:

[Repository of School of Applied Mathematics and Computer Science](#)



Sveučilište J.J. Strossmayera u Osijeku  
Odjel za matematiku  
Sveučilišni preddiplomski studij matematike

**Dino Šarlija**

**Usporedba metoda za kalibraciju kamera**

Završni rad

Osijek, 2020.

Sveučilište J.J. Strossmayera u Osijeku  
Odjel za matematiku  
Sveučilišni preddiplomski studij matematike

**Dino Šarlija**

## **Usporedba metoda za kalibraciju kamera**

Završni rad

Mentor: izv. prof. dr. sc. Domagoj Matijević  
Komentor: dr. sc. Ivica Skokić (Opservatorij Hvar,  
Geodetski fakultet, Sveučilište u Zagrebu)

Osijek, 2020.

## Sažetak

Problem kalibracije tj. određivanje parametara kamere iz slika gotovo je nezaobilazan korak u primjeni računalnog vida. U ovome radu prikazane su 3 najčešće korištene različite metode kalibracije kamere, a to su direktna linearna transformacija (DLT), Tsai-eva te Zhangova metoda. Svaka od metoda je detaljno opisana u okviru projekcijskog modela kamere uz parametre potrebne za kalibraciju. Metode su testirane na skupu slika pripremljenih za test, a rezultati su međusobno uspoređeni. Rezultati pokazuju da je Tsai-eva metoda najbolja kod slika sa značajnom radijalnom distorzijom, dok je Zhangova metoda najtočnija kod upotrebe više slika, a DLT metoda je općenito najbrža. Sve tri metode daju solidne rezultate, a izbor metode koja će se koristiti u određenoj primjeni ovisi o broju dostupnih slika za kalibraciju te potrebnoj brzini izvođenja.

## Ključne riječi

direktna linearna transformacija (DLT), Tsai-eva metoda, Zhang-ova metoda, kamera, slika, kalibracija kamere, koordinate, žarišna udaljenost, intrinzični parametri, ekstrinzični parametri, radijalna distorzija

## Summary

The problem of calibration i.e. determining the camera parameters from the images, is an almost unavoidable step in the application of computer vision. In this paper, 3 most commonly used different methods for camera calibration are presented, namely direct linear transformation (DLT), Tsai's and Zhang's method. Each of the methods is described in detail within the projection model of the camera with the parameters required for calibration. The methods were tested on a set of images prepared for the test, and the results were compared with each other. The results show that the Tsai method is best for images with significant radial distortion, while the Zhang method is most accurate when using multiple images, and the DLT method is generally the fastest. All three methods give solid results and the choice of method which will be used in a particular application depends on the number of images available to calibrate and the required execution speed.

## Key words

direct linear transformation (DLT), Tsai's method, Zhang's method, camera, image, camera calibration, coordinates, focal length, intrinsic parameters, extrinsic parameters, radial distortion

# Sadržaj

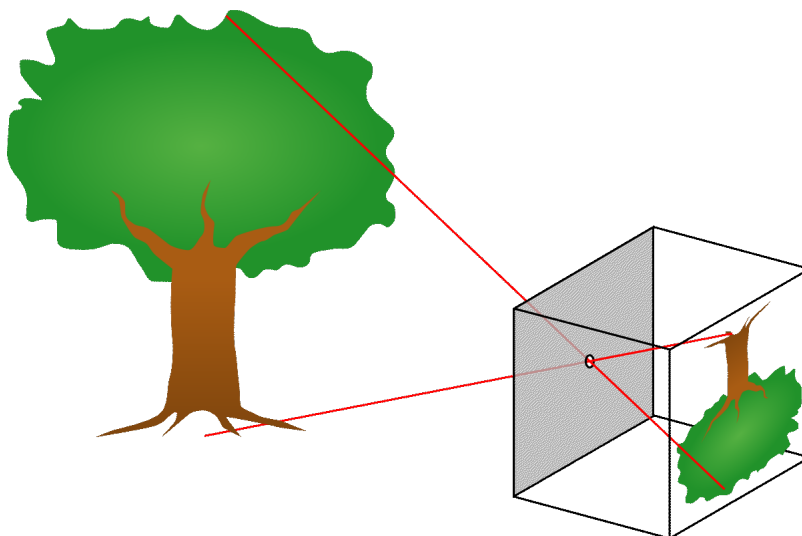
Uvod	i
<b>1 Projekcijski model kamere</b>	<b>1</b>
<b>2 Direktna Linearna Transformacija (DLT)</b>	<b>2</b>
2.1 Eksplicitna kalibracija kamere . . . . .	2
2.2 Procjena linearnih parametara . . . . .	3
<b>3 Tsai-eva metoda</b>	<b>5</b>
3.1 Unutarnja orijentacija - Kamera prema slici . . . . .	5
3.2 Vanjska orijentacija - Slika prema kameri . . . . .	5
3.3 Procjena rotacije i dijela translacije . . . . .	6
<b>4 Zhang-ova metoda</b>	<b>8</b>
4.1 Notacija . . . . .	8
4.2 Homografija između ravnine modela i njegove slike . . . . .	8
4.3 Rješavanje problema kalibracije kamere . . . . .	9
<b>5 Usporedba rezultata testiranja</b>	<b>11</b>
<b>6 Zaključak</b>	<b>16</b>
<b>Literatura</b>	<b>17</b>

## Uvod

Uspoređujemo tri bitno različita pristupa [1, 3, 4] za problem dobivanja značajki projekcije kamere tako da se snimi par slika kalibracijskog objekta koji se sastoji od planarnih uzoraka. Prva metoda s kojom se susrećemo će biti direktna linearna transformacija (eng. direct linear transformation - DLT) u kojoj ćemo objasniti intrinzične i ekstrinzične parametre te vidjeti kako dobiti DLT matricu. Druga metoda je Tsai-eva, a kod nje ćemo opisati njezine parametre za rotaciju i translaciju. Zadnja metoda koju koristimo je Zhang-ova kod koje ćemo promatrati odnos između ravnine modela i njegove slike te saznati o rješavanju problema kalibracije kamere. Ova postava je česta u računalnom vidu (eng. computer vision) kao i predstavljena rješenja. Cilj ovog rada je procjena točnosti navedenih algoritama koji omogućuje praktičnu obradu koeficijenta i korisnosti svake tehnike. Određeni skup parametara za naš model u određenom procesu naziva se model kamere. Sadrži informacije o vanjskoj orijentaciji (eng. outer orientation) kamere, značajkama unutarnje projekcije opisanim s intrinzičkim parametrima koji se definiraju kao unutarnja orijentacija (eng. inner orientation). Prvenstveno nas zanima unutarnja orijentacija, budući da intrinzični model služi kao skup ulaznih parametara za postupke 3D modeliranja.

# 1 Projekcijski model kamere

Projekcijski model kamere (eng. pinhole camera model) temelji se na principu gdje se svaka točka u objektnom prostoru projicira ravnom crtom kroz projekcijski centar u ravninu slike. Središte koordinatnog sustava kamere nalazi se u projekcijskom centru na lokaciji  $(X_0, Y_0, Z_0)$  s obzirom na koordinatni sustav objekta, a z-os okvira kamere okomita je na ravninu slike. Rotacija je predstavljena pomoću Euler-ovih kutova  $\omega, \phi$  i  $\kappa$  koji definiraju niz od tri elementarne rotacije oko x, y, z-osi. Rotacije se izvode u smjeru kazaljke na satu, prvo oko x-osi, zatim y-osi koja je već jednom zakretana, i na kraju oko z-osi koja je dva puta zakretana tijekom prethodnih faza.



Slika 1: Primjer projekcijskog model kamere



## 2 Direktna Linearna Transformacija (DLT)

Direktnu linearnu transformaciju (DLT) prvotno su razvili Abdel-Aziz i Karara [6]. Kasnije je prerađena u nekoliko publikacija, npr. [2] i [11]. Za ovu metodu koristit ćemo pinhole model kamere.

### 2.1 Eksplicitna kalibracija kamere

Parametri fizičke kamere se obično dijele na ekstrinzične i intrinzične parametre. Ekstrinzični parametri su potrebni za transformaciju koordinata objekta u koordinatni okvir centriran u kameri. U sustavima s više kamera ekstrinzični parametri opisuju i odnos između kamera.

Kako bismo iskazali proizvoljnu objektnu točku P na lokaciji  $(X_i, Y_i, Z_i)$  u koordinatama slike, prvo je trebamo transformirati u koordinate kamere  $(x_i, y_i, z_i)$ . Ova transformacija se sastoji od translacije i rotacije, a može se izvesti pomoću sljedeće matrične jednadžbe:

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \quad (1)$$

gdje su

$$\begin{aligned} m_{12} &= \sin \omega \sin \phi \cos \kappa - \cos \omega \sin \kappa & m_{11} &= \cos \phi \cos \kappa \\ m_{22} &= \sin \omega \sin \phi \sin \kappa + \cos \omega \cos \kappa & m_{21} &= \cos \phi \sin \kappa \\ m_{13} &= \cos \omega \sin \phi \cos \kappa + \sin \omega \sin \kappa & m_{31} &= -\sin \phi \\ m_{23} &= \cos \omega \sin \phi \sin \kappa - \sin \omega \cos \kappa & m_{32} &= \sin \omega \cos \phi \\ & & m_{33} &= \cos \omega \cos \phi \end{aligned}$$

Intrinzični parametri kamere obično uključuju žarišnu duljinu  $f$ , faktor iskrivljenosti  $s_u$ , a centar slike  $(u_0, v_0)$  također se naziva glavna točka. Jedinica koordinata slike su pikseli, te su stoga potrebni koeficijenti  $D_u$  i  $D_v$  za promjenu mjernih jedinica u piksele. Korištenjem modela pinhole projekcija točke  $(x_i, y_i, z_i)$  do ravnine slike izražena je kao

$$\begin{bmatrix} \tilde{u}_i \\ \tilde{v}_i \end{bmatrix} = \frac{f}{z_i} \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (2)$$

Odgovarajuće koordinate slike  $(u'_i, v'_i)$  u pikselima dobivaju se iz projekcije  $(\tilde{u}_i, \tilde{v}_i)$  primjenom sljedeće transformacije

$$\begin{bmatrix} u'_i \\ v'_i \end{bmatrix} = \frac{f}{z_i} \begin{bmatrix} D_u s_u \tilde{u}_i \\ D_v \tilde{v}_i \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \quad (3)$$

Pinhole model je samo aproksimacija prave projekcije kamere. To je koristan model koji omogućuje jednostavnu matematičku formulaciju za odnos između objekta i koordinata slike. Međutim, to nije valjano kada je potrebna visoka točnost i stoga se mora koristiti sveobuhvatniji model kamere. Obično, pinhole model je osnova koja se proteže s nekim korekcijama za sustavno iskrivljene koordinate slike.

## 2.2 Procjena linearnih parametara

Metoda DLT temelji se na modelu pinhole kamere, a zanemaruje nelinearne radijalne i tangencijalne komponente distorzije. Postupak kalibracije sastoji se od dva koraka. U prvom koraku rješava se linearna transformacija od koordinata objekta  $(X_i, Y_i, Z_i)$  do koordinata slike  $(u_i, v_i)$ . Korištenjem homogenog prikaza matrice 3 x 4 za matricu A može se napisati sljedeća jednadžba:

$$\begin{bmatrix} u_i w_i \\ v_i w_i \\ w_i \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} \quad (4)$$

Možemo izračunati parametre  $a_{11}, \dots, a_{34}$  od DLT matrice uklanjanjem  $w_i$ . Definirajmo

$$L = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -X_1 u_1 & -Y_1 u_1 & -Z_1 u_1 & -u_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -X_1 v_1 & -Y_1 v_1 & -Z_1 v_1 & -v_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -X_i u_i & -Y_i u_i & -Z_i u_i & -u_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -X_i v_i & -Y_i v_i & -Z_i v_i & -v_i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_N & Y_N & Z_N & N & 0 & 0 & 0 & 0 & -X_N u_N & -Y_N u_N & -Z_N u_N & -u_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -X_N v_N & -Y_N v_N & -Z_N v_N & -v_N \end{bmatrix}$$

$$\mathbf{a} = [a_{11}, a_{12}, a_{13}, a_{14}, a_{21}, a_{22}, a_{23}, a_{24}, a_{31}, a_{32}, a_{33}, a_{34}]^T$$

$$L\mathbf{a} = 0 \quad (5)$$

gdje je N broj točaka. Zamjenom pravilnih točaka  $(u_i, v_i)$  od slike sa promatranim vrijednostima  $(U_i, V_i)$  možemo procijeniti parametre  $a_{11}, \dots, a_{34}$  na način najmanjih kvadrata. Kako bi se izbjeglo trivijalno rješenje  $a_{11}, \dots, a_{34} = 0$ , mora se primijeniti pravilna normalizacija. Abdel-Aziz i Karara [6] koristili su ograničenje  $a_{34} = 1$ . Tada, jednadžba može biti riješena pseudoinverznom tehnikom. Problem s normalizacijom je u tome što se uvodi singularnost, ako je ispravna vrijednost  $a_{34}$  blizu nule. Umjesto  $a_{34} = 1$  Faugeras i Toscani [13] predlažu ograničenje koje je singularity free.

Parametri  $a_{11}, \dots, a_{34}$  nemaju nikakvo fizičko značenje pa se tako prvi korak u kojem se procjenjuju njihove vrijednosti može smatrati i kao implicitna faza kalibracije kamere. Postoje tehnike za izdvajanje nekih od fizičkih parametara kamere iz DLT matrice, ali neke tehnike nisu u mogućnosti riješiti sve od njih. Melen [2] je predložio metodu temeljenu na RQ dekompoziciji gdje se iz DLT matrice izdvaja skup od jedanaest parametara iz kamere. Dekompozicija je oblika:

$$A = \lambda V^{-1} B^{-1} F M T \quad (6)$$

gdje je  $\lambda$  ukupni faktor skaliranja i matrice M i T definiraju rotaciju i translaciju iz koordinatnog sustava objekta na koordinatni sustav kamere (vidi jednadžbu (1)). Matrice V, B i F sadrže žarišnu duljinu f, glavnu točku  $(u_0, v_0)$  i koeficijente za linearno iskrivljenje  $(b_1, b_2)$ :

$$V = \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 + b_1 & b_2 & 0 \\ b_2 & 1 - b_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad F = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Korekcija linearnog iskrivljenja koristi se kako bi se nadoknadile ortogonalne pogreške koordinatnih osi slike. U ovom postupku, faktor iskrivljenosti  $s_u$  se pretpostavlja da je 1. U slučaju strukture koplanarne kontrolne točke,  $3 \times 4$  DLT matrica postaje singularna. Tako se matrica  $3 \times 3$  s devet nepoznatih parametara mora koristiti. Melen [2] je također predložio metodu dekompozicije matrice  $3 \times 3$ , ali se samo može procijeniti podskup parametara fizičkih kamera.

### 3 Tsai-eva metoda

Metoda koju je predložio R.Y. Tsai [3] linearizira ogroman dio računanja tako da ograničava efekte izobličenja leće na radijalnoj distorziji. Tsai-eva tehnika kalibracije od dvije faze je brza i nije nam potrebna početna pretpostavka parametara kalibracije.

Tsai-eva metoda prvo pokušava dobiti procjene što više parametara pomoću linearne metode najmanjih kvadrata. To je zgodan i brz proces budući da se takvi problemi mogu riješiti pomoću pseudo-inverzne matrice. U početnom koraku ne provode se ograničenja između parametara (kao što je ortonormiranost matrice rotacije), a ono što je minimizirano nije pogreška u ravnini slike, već količina parametara koja pojednostavi analizu i dovodi do linearnih jednadžbi. To, međutim, ne utječe na konačni rezultat, budući da se te procijenjene vrijednosti parametara koriste samo kao početne vrijednosti za konačnu optimizaciju.

U sljedećem koraku, ostatak parametara dobiva se pomoću nelinearnih metoda optimizacije koje pronalaze najbolje poklapanje između promatranih točaka slike i onih predviđenih iz ciljanog modela. Parametri procijenjeni u prvom koraku su obrađeni u procesu.

Rješenje je općenito osmišljeno za jedno-prikaznu kalibraciju, ali je također primjenjivo za kalibraciju položaja s više prikaza.

#### 3.1 Unutarnja orijentacija - Kamera prema slici

Unutarnja orijentacija je odnos između centralnih koordinata kamere i koordinata slike. Koordinatni sustav kamere ima svoje ishodište u središtu projekcije, svoju z os duž optičke osi, te njegove x i y osi su paralelne s osima x i y od slike. Koordinate kamere i koordinate slike povezane su jednadžbama perspektivne projekcije:

$$\frac{x_1 - x_0}{f} = \frac{x_C}{z_C} \quad i \quad \frac{y_1 - y_0}{f} = \frac{y_C}{z_C}$$

gdje je  $f$  principna udaljenost (udaljenost od središta projekcije do ravnine slike), a  $(x_0, y_0)$  je principna točka (podnožje okomito od središte projekcije do ravnine slike). To jest, centar projekcije je na  $(x_0, y_0, f)^T$ , po mjerenjima u koordinatnom sustavu slike.

#### 3.2 Vanjska orijentacija - Slika prema kameri

Vanjska orijentacija je odnos između koordinatnog sustava sa scenom u centru i koordinatnog sustava sa kamerom u centru. Transformacija iz scene u kameru sastoji se od rotacije i translacije.

Koordinatni sustav scene može biti bilo koji sustav prikladan za određeni dizajn cilja. U slučaju planarnog cilja, z os se bira okomito na ravninu, a  $z = 0$  u ciljanom ravnini.

Ako su  $r_S$  koordinate od točke izmjerene u koordinatnom sustavu scene i koordinate  $r_C$  izmjerene u koordinatnom sustavu kamere, tada

$$r_C = R(r_S) + t$$

gdje je  $t$  translacija, a  $R(\dots)$  rotacija.

Ako smo za trenutak odabrali da koristimo ortonormiranu matricu za predstavljanje rotacije, onda to možemo napisati u komponentnom obliku:

$$\begin{pmatrix} x_C \\ y_C \\ z_C \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} x_S \\ y_S \\ z_S \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (7)$$

gdje je  $r_C = (x_C, y_C, z_C)^T$ ,  $r_S = (x_S, y_S, z_S)^T$  i  $t = (t_x, t_y, t_z)^T$ . Nepoznanice koje treba odrediti u problemu vanjske orijentacije su translacija vektora  $t$  i rotacija  $R(\dots)$ .

### 3.3 Procjena rotacije i dijela translacije

U početku pretpostavljamo da imamo razumnu procjenu položaja principne točke  $(x_0, y_0)$ . Ova točka je uobičajeno u blizini sredine CCD ili CMOS senzora. Koordinate upućujemo na ovu točku koristeći

$$x'_I = x_I - x_0 \quad \text{i} \quad y'_I = y_I - y_0$$

tako da je

$$\frac{x'_I}{f} = s \frac{x_C}{z_C} \quad \text{i} \quad \frac{y'_I}{f} = s \frac{y_C}{z_C} \quad (8)$$

Dalje, smatramo da je samo smjer točke na slici izmjeren od principne točke. To daje rezultat koji je neovisan o nepoznatoj principalnoj udaljenosti  $f$ . Rezultat je također neovisan o radijalnoj distorziji.

$$\frac{x'_I}{y'_I} = s \frac{x_C}{y_C} \quad (9)$$

Korištenjem proširenja u smislu komponenti rotacijske matrice  $R$  dobivamo:

$$\frac{x'_I}{y'_I} = s \frac{r_{11}x_S + r_{12}y_S + r_{13}z_S + t_x}{r_{21}x_S + r_{22}y_S + r_{23}z_S + t_y}$$

koje nakon sređivanja izraza dobivamo jednadžbu:

$$s(r_{11}x_S + r_{12}y_S + r_{13}z_S + t_x)y'_I - (r_{21}x_S + r_{22}y_S + r_{23}z_S + t_y)x'_I = 0 \quad (10)$$

ili

$$(x_S y'_I) s r_{11} + (y_S y'_I) s r_{12} + (z_S y'_I) s r_{13} + y'_I s t_x - (x_S x'_I) r_{21} - (y_S x'_I) r_{22} - (z_S x'_I) r_{23} + x'_I t_y = 0 \quad (11)$$

koju možemo gledati kao linearnu homogenu jednadžbu sa osam nepoznanica

$$s r_{11}, s r_{12}, s r_{13}, r_{21}, r_{22}, r_{23}, s t_x, t_y$$

Koeficijenti u jednadžbi su proizvodi komponenti odgovarajućih koordinata scene i koordinata slike.

Dobivamo jednu takvu jednadžbu za svaku korespondenciju između kalibracije ciljne točke  $(x_{S_i}, y_{S_i}, z_{S_i})^T$  i točke slike  $(x_{I_i}, y_{I_i})^T$ .

Imajmo na umu da je ovdje nepoznat faktor skaliranja jer su ove jednadžbe homogene. To jest, ako imamo rješenje za osam nepoznanica, onda je bilo koji višekratnik od tog rješenja također rješenje. Kako bismo dobili rješenje, homogenu jednadžbu možemo pretvoriti u nehomogenu jednadžbu tako da proizvoljno postavimo jednu nepoznanicu (recimo  $t_y$ ) na jedan.

Tada imamo sedam nepoznanica koje možemo riješiti ako imamo sedam korespondencija između ciljnih koordinata i koordinata slike. Ako imamo više od sedam korespondencija, onda možemo minimizirati zbroj kvadrata pogrešaka pomoću pseudo-inverzne metode.

Kada dobijemo rješenje, moramo se sjetiti da se osam nepoznanica (sedam za koje smo riješili plus ona koju smo postavili na jedan) može skalirati proizvoljnim faktorom. Pretpostavimo da je najbolje prikladno rješenje skupa jednadžbi

$$sr'_1, sr'_2, sr'_3, st'_x, r'_2, r'_2, r'_3, t'_y = 1$$

Možemo procijeniti točan faktor skaliranja ako primjetimo da bi redci matrice rotacije trebali biti normirani, to jest

$$r_{11}^2 + r_{12}^2 + r_{13}^2 = 1 \quad \text{i} \quad r_{21}^2 + r_{22}^2 + r_{23}^2 = 1$$

Moramo pronaći faktor skaliranja  $c$  koji će se primijeniti na rješenje kako bismo zadovoljili te jednakosti. Lako je vidjeti da

$$c = 1/\sqrt{r_{21}'^2 + r_{22}'^2 + r_{23}'^2} \quad (12)$$

i

$$c/s = 1/\sqrt{(sr'_{11})^2 + (sr'_{12})^2 + (sr'_{13})^2} \quad (13)$$

Ove jednadžbe omogućuju nam oporavak faktora  $c$  kao i omjera  $s$  od horizontalnog razmaka piksela do okomitog razmaka piksela.

Imajmo na umu da nismo provodili ortogonalnost na prva dva reda matrice rotacije. Možemo poboljšati stvari tako da ih prilagodimo kako bi bili ortogonalni.

## 4 Zhang-ova metoda

Metoda kalibracije kamere koja koristi tradicionalne tehnike kalibracije (poznate točke kalibracije) i tehnike samo-kalibracije (korespondencija između točaka kalibracije kada su u različitim položajima). Za izvođenje potpune kalibracije putem Zhang metode potrebne su najmanje tri različite slike od cilja/mjerila kalibracije, bilo pomicanjem mjerila ili same kamere. Ako su neki od intrinzičkih parametara dani kao podaci (ortogonalnost slike ili koordinate optičkog centra) broj potrebnih slika može se smanjiti na dvije [4], [12].

### 4.1 Notacija

Točka sa dvije dimenzije je zapisana sa  $m = [u, v]^T$ . Točka sa 3 dimenzije je zapisana sa  $M = [X, Y, Z]^T$ . Koristimo  $\tilde{x}$  za označavanje proširenog vektora dodavanjem 1 kao posljednjeg elementa:  $\tilde{m} = [u, v, 1]^T$  i  $\tilde{M} = [X, Y, Z, 1]^T$ . Kamera je modelirana po uobičajenoj pinhole kameri. Odnos između 3D točke M i njezine projekcije slike m je dan

$$s\tilde{m} = A[Rt]\tilde{M} \quad (14)$$

gdje je s proizvoljni faktor skaliranja, (R, t), koji se naziva ekstrinzični parametar, je rotacija i translacija koja povezuje globalni koordinatni sustav sa koordinatnim sustavom kamere, i A, pod nazivom intrinzična matrica kamere, dana izrazom

$$A = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

s  $(u_0, v_0)$  koordinatama principne točke,  $\alpha$  i  $\beta$  su faktori skaliranja osi u i v na slici, a  $\gamma$  parametar koji opisuje iskrivljenost dviju osi od slike.

### 4.2 Homografija između ravnine modela i njegove slike

Pretpostavimo da je ravnina modela na  $Z = 0$  u globalnom koordinatnom sustavu. Označimo i-ti stupac rotacijske matrice R s  $r_i$ . Iz (14) imamo

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \begin{bmatrix} r_1 & r_2 & r_3 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = A \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

I dalje ćemo koristiti M za označavanje točke na ravnini modela, ali  $M = [X, Y]^T$  budući da je Z uvijek jednaka 0,  $\tilde{M} = [X, Y, 1]^T$ . Stoga su točka modela M i njezina slika m povezane s homografijom H:

$$s\tilde{m} = H\tilde{M} \quad , \quad H = A \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \quad (15)$$

S obzirom na sliku ravnine modela može se procijeniti homografija. Označimo je s  $H = [h_1 \ h_2 \ h_3]$ . Iz (15) imamo

$$[h_1 \ h_2 \ h_3] = \lambda A \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}$$

gdje je  $\lambda$  proizvoljan skalar. Koristeći informacije da su  $r_1$  i  $r_2$  ortonormirani, imamo

$$h_1^T A^{-T} A^{-1} h_2 = 0 \quad (16)$$

$$h_1^T A^{-T} A^{-1} h_1 = h_2^T A^{-T} A^{-1} h_2 \quad (17)$$

### 4.3 Rješavanje problema kalibracije kamere

Neka je

$$\begin{aligned} B = A^{-T} A^{-1} &\equiv \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\alpha^2} & -\frac{\gamma}{\alpha^2 \beta} & \frac{v_0 \gamma - u_0 \beta}{\alpha^2 \beta^2} \\ \frac{\gamma}{\alpha^2 \beta} & \frac{\gamma^2}{\alpha^2 \beta^2} + \frac{1}{\beta^2} & -\frac{(v_0 \gamma - u_0 \beta)^2}{\alpha^2 \beta^2} - \frac{v_0}{\beta^2} \\ \frac{v_0 \gamma - u_0 \beta}{\alpha^2 \beta} & -\frac{(v_0 \gamma - u_0 \beta)^2}{\alpha^2 \beta^2} - \frac{v_0}{\beta^2} & \frac{(v_0 \gamma - u_0 \beta)^2}{\alpha^2 \beta^2} + \frac{v_0^2}{\beta^2} + 1 \end{bmatrix} \end{aligned} \quad (18)$$

Imajmo na umu da je B simetrična matrica, definirana 6D vektorom

$$b = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T \quad (19)$$

Neka je  $i$ -ti stupac vektor od H zadan s  $h_i = [h_{i1}, h_{i2}, h_{i3}]^T$ . Tada imamo

$$h_i^T B h_j = v_{ij}^T b \quad (20)$$

pri čemu je

$$v_{ij} = [h_{i1} h_{j1}, h_{i1} h_{j2} + h_{i2} h_{j1}, h_{i2} h_{j2}, h_{i3} h_{j1} + h_{i1} h_{j3}, h_{i3} h_{j2} + h_{i2} h_{j3}, h_{i3} h_{j3}]^T$$

Stoga se dva temeljna ograničenja (16) i (17), iz dane homografije, mogu prepisati kao 2 homogene jednadžbe u b:

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0 \quad (21)$$

Ako se promatra  $n$  slika s ravnine modela, slaganjem  $n$  takvih jednadžbi kao što je (14) imamo

$$Vb = 0 \quad (22)$$

gdje je  $V$   $2n \times 6$  matrica. Ako je  $n \geq 3$ , imat ćemo općenito jedinstveno rješenje  $b$  definirano do faktora skaliranja. Ako je  $n = 2$ , možemo staviti ograničenje na iskrivljenost  $\gamma = 0$ , odnosno,  $[0, 1, 0, 0, 0, 0]b = 0$ , koje se dodaje kao dodatna jednadžba na (22). (Ako je  $n = 1$ , možemo riješiti intrinzične parametre od dvije kamere, npr.  $\alpha$  i  $\beta$ , pod pretpostavkom da su poznati  $u_0$  i  $v_0$  (npr. u centru slike) i  $\gamma = 0$ .) Rješenje za (22) dobro je poznato kao svojstveni vektor od  $V^T V$  povezan s najmanjom svojstvenom vrijednošću (ekvivalentno, desni singularni vektor od  $V$  povezan je s najmanjom singularnom vrijednošću).

Jednom kada se  $b$  procijeni, možemo izračunati intrinzičnu matricu  $A$  za sve kamere.

Nakon što je  $A$  poznat, ekstrinzični parametri za svaku sliku su lako izračunljivi. Iz (15), imamo



$$\begin{aligned}r_1 &= \lambda A^{-1}h_1 \\r_2 &= \lambda A^{-1}h_2 \\r_3 &= r_1 \times r_2 \\t &= \lambda A^{-1}h_3\end{aligned}$$

s  $\lambda = 1/\|A^{-1}h_1\| = 1/\|A^{-1}h_2\|$ . Naravno, zbog mogućnosti pogreške u podacima, tako izračunata matrica  $R = [r_1, r_2, r_3]$  općenito ne zadovoljava svojstva matrice rotacije.

## 5 Usporedba rezultata testiranja

Kreirali smo 3 uzorka slika po jedan za svaku metodu u besplatnom i eng. "open source" programskom paketu za stvaranje 3D okruženja Blender-u (verzija 2.82). Na svaki smo model stavili teksturu koja odgovara izgledu šahovske ploče. Osim kreiranja slika, u Blenderu dobivamo podatke o lokaciji, rotaciji i veličini slike, ali ujedno i postavljamo kameru na određenu lokaciju ( $x, y, z$ ) sa specifikacijama o rezoluciji slike ( $w$  - širina,  $h$  - visina), žarišnoj udaljenosti ( $f$ ), senzoru (sensor), parametar iskrivljenosti ( $s$ ) i Euler-ove kutove ( $\omega, \psi, \kappa$ ) u stupnjevima, vrijednosti tih parametara su:

$$\begin{aligned} x &= 7 \quad , \quad y = -7 \quad , \quad z = 5 \\ w &= 1920 \text{ px (piksela)} \quad , \quad h = 1080 \text{ px} \\ f &= 20 \text{ mm} \quad , \quad \text{sensor} = 36 \times 24 \text{ mm} \quad , \quad s = 0 \\ \omega &= 54.22431579^\circ \quad , \quad \psi = 39.05283433^\circ \quad , \quad \kappa = 24.41759732^\circ \end{aligned}$$

Metode koje smo koristili su prethodno navedene, a to su direktna linearna transformacija, Tsai-eva i Zhang-ova metoda. Sve tri metode su implementirane u programskom jeziku Python (verzija 3.7). Moramo uzeti u obzir da je svaka od metoda specifična stoga slike koje koristimo u testiranju ne sadrže iste oblike ili elemente.

Za DLT metodu u Blender-u smo napravili 3D oblik koji odgovara kocki (vidi sliku 2) visine 2m, širine 2m i dužine 2m (metri su osnovna mjerna jedinica u Blenderu pa ćemo ih takve ostaviti) koja ima središte u točki (0,0,0). Osim lokacije i dimenzije naše kocke potrebne su nam koordinate vrhova kocke, a one iznose:

$$(x, y, z) = (-1, -1, 1), (-1, -1, -1), (1, -1, -1), (1, 1, -1), (1, 1, 1), (-1, 1, 1), (1, -1, 1)$$

Nakon što smo napravili naš uzorak i izmjerili podatke o njemu prelazimo na implementacijski dio koji je odrađen u Python-u. DLT metoda je malo modificirana kako bi što više odgovarala za naš slučaj, a implementacija je preuzeta sa [13]. Naš program prvo pretvara Euler-ove kutove iz XYZ u ZYX oblik pa izračuna matricu translacije kamere iz Euler-ovih kutova. Nakon toga računamo projekcijsku, translacijsku, rotacijsku i intrinzičnu matricu te pomoću tih parametara i 3D koordinata vrhova naše kocke dobivamo 2D koordinate ( $u, v$ ) vrhova. Kada smo dobili sve te podatke pokrenemo DLT metodu na 3D ( $x, y, z$ ) i 2D ( $u, v$ ) koordinatama i dobivamo nove vrijednosti za translacijske vektore, ukupnu pogrešku, projekcijsku, rotacijsku i intrinzičnu matricu. Pomoću novo dobivene projekcijske matrice izračunamo 2D koordinate kako bi ih usporedili s originalnim 2D koordinatama.

**Primjer 1.** *Ovdje vidimo originalne ( $u, v$ ) koordinate i ( $u, v$ ) koordinate dobivene DLT metodom.*

*Originalne 2D ( $u, v$ ) koordinate:*

$$(606, 307), (633, 737), (960, 927), (1287, 737), (1314, 306), (960, 187), (960, 458)$$

*DLT 2D ( $u, v$ ) koordinate:*

$$(604.94309952, 306.27726548), (632.82928882, 736.24953362), (959.50000007, 926.83209601),$$

(1286.1707113, 736.24953358), (1314.05690055, 306.27726543), (959.50000002, 187.06391436),  
(959.50000004, 457.53223951)

**Primjer 2.** Imamo dvije projekcijske matrice od kojih prva (gore) sadrži originalne vrijednosti, a druga (dole) vrijednosti dobivene pomoću DLT metode.

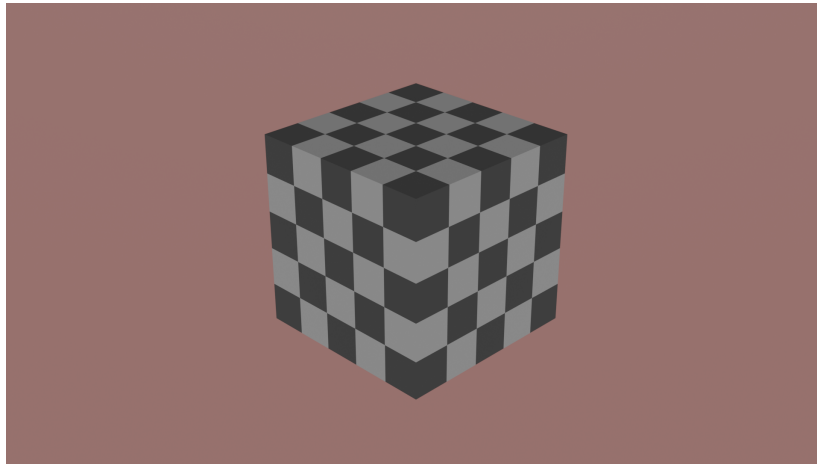
$$\begin{bmatrix} 115.513423 & 224.529619 & -39.2773255 & 959.500000 \\ 46.5397772 & -46.5397772 & -236.324156 & 530.063900 \\ -0.0568088568 & 0.0568088568 & -0.0409352011 & 1.00000000 \end{bmatrix}$$

$$\begin{bmatrix} 116.116819 & 223.302078 & -38.0847418 & 959.880077 \\ 46.7682458 & -47.3046857 & -235.888893 & 530.338291 \\ -0.0561338490 & 0.0555658322 & -0.0396917972 & 1.00000000 \end{bmatrix}$$

**Primjer 3.** Imamo dvije rotacijske matrice od kojih prva (gore) sadrži originalne vrijednosti, a druga (dole) vrijednosti dobivene pomoću DLT metode.

$$\begin{bmatrix} 0.707106781 & 0.707106781 & -6.81477652e - 11 \\ -0.321019761 & 0.321019761 & 0.891006524 \\ 0.630036755 & -0.630036755 & 0.453990500 \end{bmatrix}$$

$$\begin{bmatrix} 0.70445035 & 0.70974827 & -0.00266444 \\ -0.31701493 & 0.31800289 & 0.89351871 \\ 0.63502066 & -0.6285949 & 0.44901805 \end{bmatrix}$$



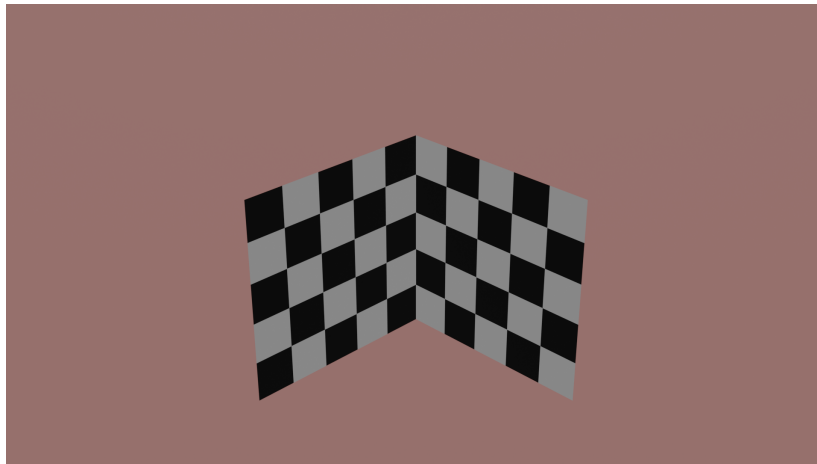
Slika 2: Primjer kocke s uzorkom šahovske ploče koju koristimo kao model za DLT metodu

Za Tsai-evu metodu u Blender-u smo napravili dvije ravnine međusobno okomite koje imaju jednu zajedničku stranicu (vidi sliku 3) visine 2m, širine 2m i dužine 2m. Prva ravnina ima središte u točki (0, -1, 0) dok druga ravnina ima središte u točki (1, 0, 0), opet su nam potrebne koordinate vrhova tih ravnina koje su dane:

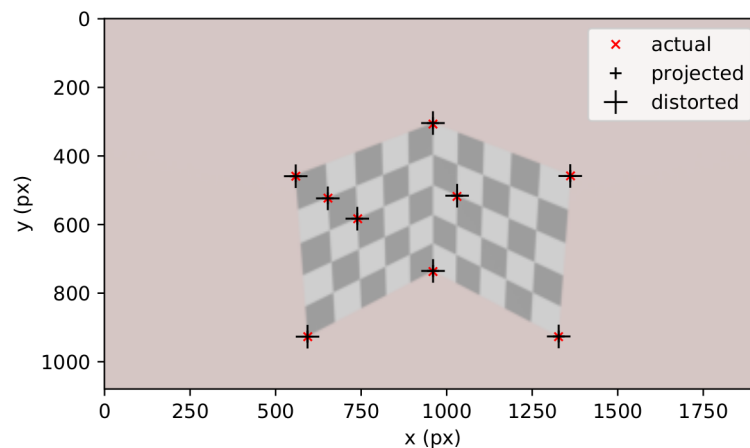
$$(x, y, z) = (0, -2, -1), (0, -2, 1), (0, 0, -1), (0, 0, 1), (2, 0, -1), (2, 0, 1)$$

Nakon što smo napravili naš uzorak i izmjerili podatke o njemu prelazimo na implementacijski dio koji je odrađen u Python-u. Tsai-eva metoda je također malo modificirana kako bi što

više odgovarala našem slučaju, a implementacija je preuzeta sa [14]. Naš program zahtjeva da unesemo podatke vezane za kameru koje smo naveli na početku ovog poglavlja kao i podatke o koordinatama svakog vrha u 3D i 2D prostoru. Prvo pretvaramo 2D koordinate uz pomoć žarišne udaljenosti, senzora, veličine piksela i rezolucije u 3D koordinate, s novo dobivenim 3D koordinatama uz još par operacija dobivamo nove 2D koordinate. Sad kad znamo nove 2D i 3D koordinate računamo translacijski vektor, parametar distorzije, projekcijsku i rotacijsku matricu.



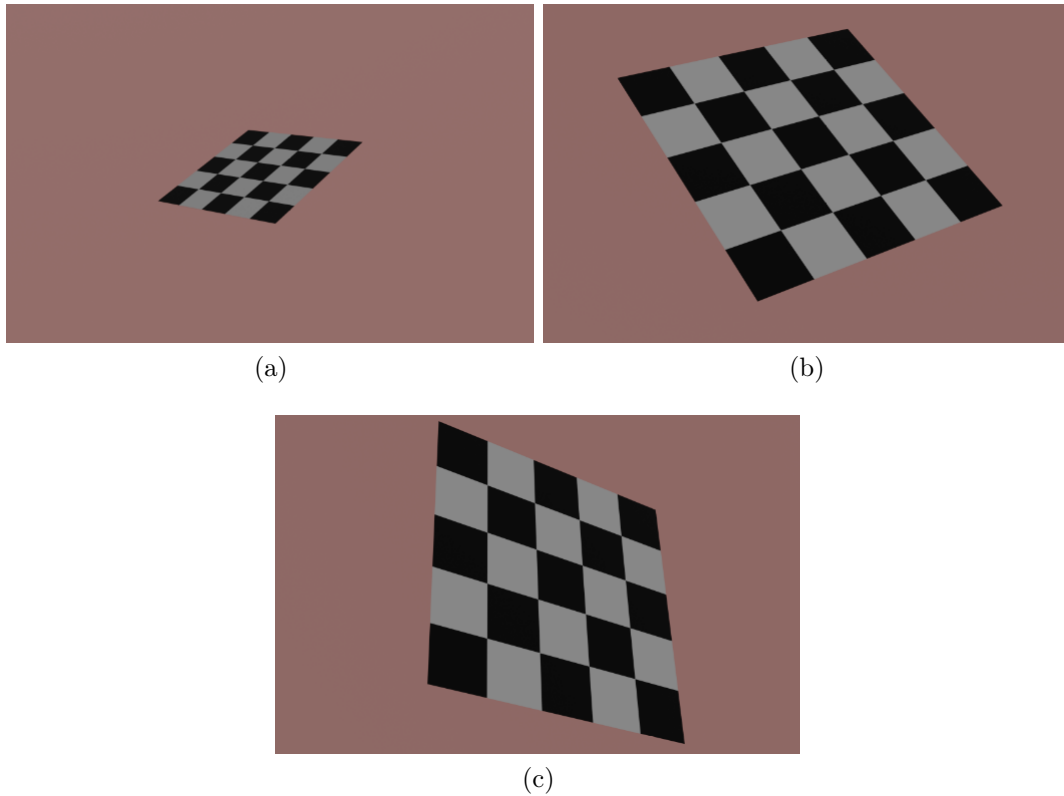
Slika 3: Prije dvije ravnine međusobno okomite koje koristimo kao model za Tsai-ovu metodu kalibracije



Slika 4: Primjer rezultata koje dobijemo izvođenjem Tsai-eve metode

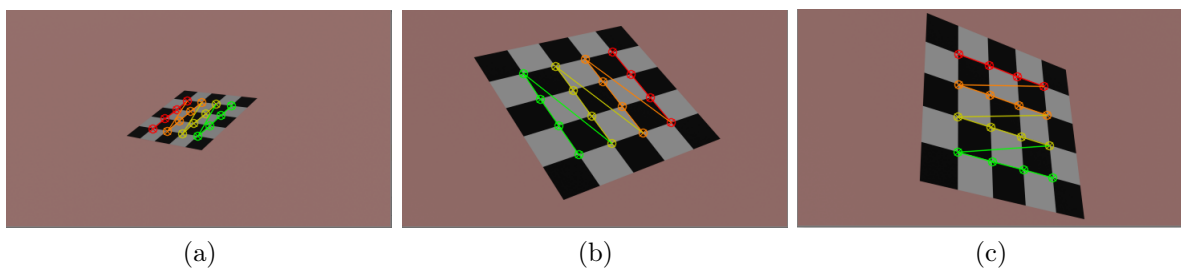
Za Zhang-ovu metodu odlučili smo koristiti 3 slike (vidi sliku 5) zato što nam je za Zhang-ovu metodu potrebno 3 ili više slika koje sadrže istu ravninu samo rotiranu u različitim pozicijama. Dimenzije naše ravine su 3m x 3m, a lokacije njezinih vrhova nisu nam potrebni jer ćemo provodeći Zhang-ovu metodu doći do 3D i 2D koordinata tekture koja podsjeća na šahovsku ploču.

Za implementaciju smo koristili Python te OpenCV (eng. Open Source Computer Vision Library) iz kojega smo koristili već implementirane funkcije za kalibraciju kamere napravljene po uzoru na Zhang-ovu metodu. Prvo što trebamo napraviti je odrediti koliko kuteva (eng.



Slika 5: Prijer ravnine (šahovske ploče) u tri različita položaja koju koristimo kao model za Zhang-ovu metodu kalibracije

corners) ima na našoj ravnini ( $4 \times 4$ ), tako da znamo prosljediti dobar parametar funkciji koja ih traži na slici, ali prije nego što se pokrene ta funkcija trebali bi sliku prebaciti u sive tonove (eng. grayscale) te ako je velikih dimenzija kao u našem slučaju skalirati ih da bi što brže dobili rezultat. Ako nađemo kuteve tada želimo poboljšati koordinate kuteva tako da ih sada lociramo na izvornoj neskaliiranoj slici tako što drugoj funkciji damo približnu lokaciju dobivenu s male slike. Na slici 6 možemo vidjeti kako izgledaju naši kutevi kada su označeni. Koordinate kuteva predajemo funkciji za kalibraciju kao 2D koordinate zajedno s 3D koordinatama koje se nalaze u rasponu od -1 do 1 kojima je u svakoj točki koordinata  $z = 0$  (stavljamo ravninu u središte sustava  $(0, 0, 0)$  te skaliramo x i y koordinate da budu u zadanom rasponu).



Slika 6: Primjer ravnine (šahovske ploče) s kutovima dimenzije  $4 \times 4$ , zahvaljujući kojima dobivamo 2D koordinate

Nakon što smo ukratko opisali proces kalibracije kamere, pripremanja i dobivanja parametara za sve 3 metode u tablici 1 pokazt ćemo elemente intrinzičnih matrica kao što su žarišna udaljenost u pikselima  $a_x$  (ili  $\alpha$ ),  $a_y$  (ili  $\beta$ ), kordinate centra slike u pikselima  $u_o$ ,  $v_o$ , iskrivljenost senzora  $s$  (ili  $\gamma$ ).

Metoda	<i>Originalni podatci za DLT</i>	DLT	<i>Originalni podatci za Tsai</i>	Tsai	<i>Originalni podatci za Zhang</i>	Zhang
$a_x$ (px)	-2666.66667	-2719.42013	2666.667	2805.89	2666.67	2667.68
$a_y$ (px)	2666.66667	2722.27236	2666.667	2805.89	2666.67	2669.79
$u_o$ (px)	959.500000	947.213291	959.50	960	960	967
$v_o$ (px)	539.500000	525.856571	539.50	540	540	547
$s$	0	-1.92894159	0	0.997373	0	0

Tablica 1: Testni rezultati za sve tri metode kalibracijske metode

## 6 Zaključak

Usporedili smo tri različita pristupa za rješavanje problema kalibracije jedne kamere s planarnom kalibracijom cilja. Uzmimo u obzir da pri korištenju Zhang metode smo koristili 3 slike dok za druge dvije metode smo koristili samo jednu sliku. Prateći rezultate u tablici 1 lako je uočiti da Zhang-ova metoda daje najbolje rezultate, to možemo vidjeti na parametru iskrivljenosti ( $s$ ) koji je ostao 0 te po parametrima žarišne udaljenosti ( $a_x$  i  $a_y$ ) koji su najbliži originalnim vrijednostima. Jedini nedostatak Zhang-ove metode je brzina jer da bi dobili što točnije podatke treba nam više slika, a time i više vremena da ih obradimo. Ako nam je bitna brzina izvršavanja metode onda nam odgovaraju DLT (trajanje izvršavanja našeg koda iznosi 0.0089936 sekundi) i Tsai-eva (trajanje izvršavanja našeg koda iznosi 0.1449105 sekundi) metoda. Ako nam je bitnije da žarišna udaljenost bude što točnija originalnim vrijednostima tada bi trebali odabrati DLT metodu, a ako nam je bitnije da parametar iskrivljenosti bude bliži 0 tada bi trebali odabrati Tsai-evu metodu. Često će nam odabir kalibracijske metode ovisiti o tome kakve slike obrađujemo.

## Literatura

- [1] J. HEIKKILA AND O. SILVEN, *A four-step camera calibration procedure with implicit image correction*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97), San Juan, Puerto Rico, 1997. IEEE.
- [2] T. MELEN., *Geometrical modelling and calibration of video cameras for underwater navigation*, PhD thesis, Institutt for teknisk kybernetikk, Norges tekniske hogskole, 1994.
- [3] R. Y. TSAI, *An efficient and accurate camera calibration technique for 3d-machine vision*, In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 364–374, Miami Beach, Florida, 1986. IEEE.
- [4] Z. ZHANG, *A flexible new technique for camera calibration*, In IEEE Transactions on Pattern Analysis and Machine Intelligence, pages pp. 1330–1334, 2000.
- [5] H. ZOLLNER, *Camera calibration using pose estimation*, Technical Report 85, Pattern and Image Processing Group (PRIP), Institute of Computer Aided Automation, Computer Science Dept., Vienna University of Technology, Vienna, Austria, August 2003.
- [6] ABDEL-AZIZ, Y. I., KARARA, H. M, *Direct linear transformation into object space coordinates in close-range photo-grammetr*, In Proceedings Symposium on Close Range Photo-grammetry , Urbana, Illinois, 1971.
- [7] FAIG, W., *Calibration of close-range photogrammetric systems: Mathematical formulation*, Photogrammetric Engineering and Remote Sensing 41(12): 1479-1486.
- [8] TSAI, ROGER Y., *An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision*, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, 1986, pp. 364–374
- [9] HORN, B.K.P., H. M. HILDEN, S. NEGAHDARIPOUR, *Closed Form Solution of Absolute Orientation using Orthonormal Matrices*, Journal of the Optical Society A, Vol. 5, No. 7, pp. 1127–1135, July.
- [10] D. C. BROWN, *Close-range camera calibration*, Photogrammetric Engineering, 37(8):855–866, 1971.
- [11] FAUGERAS, O. D., TOSCANI, G., *Camera calibration for 3D computer vision*, Proc. International Workshop on Industrial Applications of Machine Vision and Machine Intelligence, Silken, Japan, 1987
- [12] P. STURM AND S. MAYBANK, *On plane-based camera calibration: a general algorithm, singularities, applications*, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Fort Collins, CO, USA, June 1999
- [13] <https://github.com/acvictor/DLT/blob/master/DLT.py>



[14] <https://github.com/bailus/tsai-calibration>