

**ANALIZA PROCESA 3D REKONSTRUKCIJE SCENE POMOĆU STEREO VIZIJE****AN ANALYSIS OF 3D SCENE RECONSTRUCTION PROCESS USING STEREO VISION**Aleksandra Ujfaluši, Željens Trpovski, *Fakultet tehničkih nauka, Novi Sad***Oblast – ANIMACIJA U INŽENJERSTVU**

**Kratak sadržaj** – Precizna 3D percepcija iz para slika je osnovna tema u kompjuterskoj viziji i robotici. Ovaj rad sadrži analizu epipolarne geometrije stereo sistema i analitički opis alata za računanje 3D oblaka tačaka iz uparenih piksela. Parametri stereo kamere iz procesa kalibracije kamere i mapa dispariteta su dva važna parametra za postizanje preciznih rezultata 3D rekonstrukcije. Proces 3D rekonstrukcije generiše koordinate globalnih tačaka (oblak tačaka).

**Ključne reči:** Kalibracija kamere, Fundamentalna matrica, Esencijalna matrica, 3D rekonstrukcija, Oblak tačaka

**Abstract** – Accurate 3D perception from image pair is a core subject in computer vision and robotics. This paper contains analysis of epipolar geometry of the stereo system and an analytical description of tools to compute a 3D point cloud from matched pixels. Stereo camera parameters from the camera calibration process and disparity map are two important parameters to obtain an accurate result for the 3D reconstruction. 3D reconstruction process generates the coordinates of world points (point cloud).

**Keywords:** Camera Calibration, Fundamental matrix, Essential matrix, 3D reconstruction, Point cloud

**1. UVOD**

Kompjuterska vizija je oblast koja se bavi kreiranjem novih informacija na osnovu obrade slika. Primena generisanja strukture iz niza slika ili videa je široka. Ovaj proces primenjuje se u raznim oblastima poput robotike, arheologije, geo-lokalizacije, računarske grafike itd.

Rezultati ovakvih 3D algoritama moraju se na neki način prikazati, a jedna od biblioteka koja se razvijala u tom pravcu jeste *OpenCV* biblioteka [1]. *OpenCV* biblioteka sadržala je uvek veoma pogodan temelj na osnovu kojeg su se razvijali algoritmi i programi za primenu kompjuterske vizije.

**2. MODEL KAMERE**

Za mnoge primene potrebno je znati parametre kamere kako bi se ona mogla efikasno koristiti kao vizualni senzor. Ovi parametri mogu se koristiti za korekciju distorzije sočiva, merenje objekata u realnom svetu,

pronalaženje lokacije kamere u prostoru i sl. Proces računanja parametara kamere naziva se kalibracija kamere. U toku procesa prikupljaju se sve informacije o kameri (parametri ili koeficijenti kamere) pomoću kojih se precizno opisuje odnos između 3D tačke u realnom svetu i njoj korespondentne 2D projekcije ili preciznije piksela u slici.

**2.1. Kalibracija kamere**

Tipično, kalibracija kamere podrazumeva računanje dve vrste parametara. Eksterni (eng. *extrinsics*) i interni (eng. *intrinsics*) parametri služe za opisivanje načina mapiranja 3D tačke u prostoru na 2D sliku. Eksterni parametri sadrže opis položaja kamere u prostoru i pravac u kom je kamera usmerena, a interni jesu unutrašnji parametri kamere koji zavise od modela kamere i utiču na način mapiranja 3D tačke na ravan kamere.

Prilikom mapiranja tačaka na 2D ravan prelazi se iz jednog u drugi koordinatni sistem. Tokom celog procesa javljaju se četiri koordinatna sistema sa odgovarajućim koordinatama:

1. Globalni koordinatni sistem
2. Koordinatni sistem kamere
3. Koordinatni sistem slike
4. Koordinatni sistem senzora

U globalnom koordinatnom sistemu definisani su svi objekti na sceni, što uključuje i koordinate položaja kamere. U koordinatnom sistemu kamere opisan je način na koji kamera vidi scenu tj. svi objekti opisani su u odnosu na centar projekcije kamere kada se koristi *pinhole*<sup>1</sup> model kamere.

Prelazak iz globalnog koordinatnog sistema u koordinatni sistem kamere obuhvata neki vid euklidske transformacije. Zatim, iz pogleda kamere mapiraju se 3D tačke na ravan kamere čime se prelazi u koordinatni sistem slike. Prelaskom u koordinatni sistem senzora dobijaju se individualni pikseli u slici za digitalne kamere. Prelasci kroz ove koordinatne sisteme mogu se zapisati kao transformacija:

$$\begin{bmatrix} {}^s x \\ {}^s y \\ 1 \end{bmatrix} = {}^s H_c {}^c H_k {}^k H_o \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

gde su:

**NAPOMENA:**

Ovaj rad proistekao je iz master rada čiji mentor je bio dr Željens Trpovski, vanr. prof.

<sup>1</sup> *Pinhole* model kamere jeste jednostavna kamera bez sočiva sa malim otvorom blende.

${}^kH_o$  – matrica transformacije iz globalnog koordinatnog sistema u koordinatni sistem kamere

${}^cH_k$  – matrica transformacije iz koordinatnog sistema kamere u koordinatni sistem slike

${}^sH_c$  – matrica transformacije iz koordinatnog sistema slike u koordinatni sistem senzora

$X, Y, Z$  – koordinate 3D tačke u globalnom sistemu

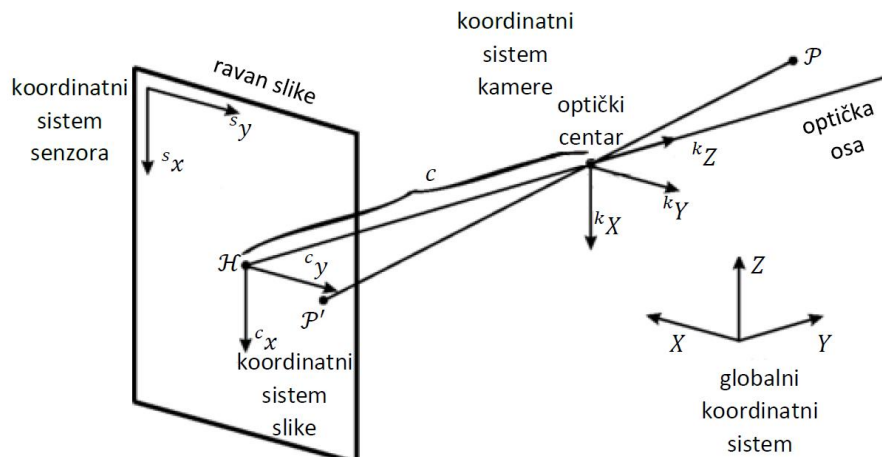
${}^s x, {}^s y$  – koordinate piksela u slici,

a međusobni položaj centara ovih koordinatnih sistema prikazan je na Slici 1.

Sa Slike 1 može se primetiti da se pravci koordinatnih osa ( $x$  i  $y$ ) u sistemu kamere i sistemu slike poklapaju.

Rastojanje između koordinatnog početka ova dva sistema naziva se konstanta kamere i obeležava se oznakom  $c$ . Veoma jednostavan odnos postoji i između sistema slike i sistema senzora. Koordinatni sistem senzora pomeren je jednostavnom translacijom od koordinatnog sistema slike.

U procesu projekcije javlja se podela na eksterne i interne parametre koji su prikazani na Slici 3.4. Eksterni parametri sadrže opis pozicije i rotacije kamere u globalnom koordinatnom sistemu, dok su interni parametri vezani za model kamere i ne menjaju se tokom akvizicije slike. Potrebno je naglasiti da se interni parametri menjaju sa promenom sočiva na kameri.



Slika 1. Međusobni odnos koordinatnih sistema

## 2.2. Projekcija iz 3D sistema u 2D sistem

Mapiranje iz 3D scene na ravan senzora pomoću pomenute affine transformacije naziva se direktna linearna transformacija. Takvo mapiranje može se zapisati kao

$$x = PX \quad (2)$$

gde je

$$P = KR[I_3 | -X_o] \quad (3)$$

a  $K$  je matrica kalibracije

$$\begin{bmatrix} c & cs & x_H \\ 0 & c(1+m) & y_H \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

Matrica kalibracije sadrži ukupno 5 parametara: konstanta kamere  $c$ , principijelna tačka  $x_H$ ,  $y_H$ , skaliranje  $m$  i smicanje  $s$ .

Ovakvo mapiranje sadrži ukupno 11 parametara od kojih su 6 eksterni parametri, a 5 interni parametri potrebni za kalibraciju.

## 2.3. Nelinearne greške

Do sada su razmatrane samo idealne situacije prilikom projekcije tačaka. U stvarnom svetu dolazi do nelinearnih grešaka usled nesavršenosti sočiva ili samog modela sočiva. Na primer, riblje oko je sočivo kod koga se

očigledno može primetiti da se ravne linije u stvarnom svetu projektuju u krive linije na slici. Takođe, greške se mogu javiti i usled neplanarnosti sočiva unutar kamere.

Nelinearne greške uvode se pomoću opšteg mapiranja. Takvo mapiranje zavisi je od koordinata piksela  $x$  u koordinatnom sistemu slike i od dodatnog parametra  $q$ .

$${}^a x = {}^s x + \Delta x(x, q) \quad (5)$$

$${}^a y = {}^s y + \Delta y(x, q) \quad (6)$$

Funkcija koja opisuje pomeraj koristi koordinate iz sistema slike, jer se u njemu nalazi principijelna tačka na koordinatama  $(0,0)$ .

Nelinearne greške obično su simetrične u odnosu na principijelnu tačku, pa je na ovaj način najlakše definisati funkciju koja opisuje takve greške.

Najčešće se dešava da lokacija principijelne tačke ostaje nepromenjena dok je ostale tačke na istoj udaljenosti od principijelne tačke potrebno približiti ili udaljiti u odnosu na istu.

Primer ovakve distorzije može se videti kod širokougaonih sočiva u kojoj se prave linije u stvarnosti projektuju u krive linije (Slika 2).



Slika 2. Uklanjanje distorzije kod široko-ugaonih sočiva

Modifikovanjem postojećih jednačina kako bi se uzele u obzir nelinearne greške, jednačine (2)-(4) postaju:

$${}^a x = {}^a P(x, q) X \quad (7)$$

gde je

$${}^a P(x, q) = {}^a K(x, q) R [I_3] - X_0 \quad (8)$$

a  $K$  je matrica kalibracije

$$\begin{bmatrix} c & cs & x_H + \Delta x(x, q) \\ 0 & c(1 + m) & y_H + \Delta y(x, q) \\ 0 & 0 & 1 \end{bmatrix} \quad (9)$$

## 2.4. Inverzija procesa mapiranja

Celokupno mapiranje može se predstaviti kao proces iz dva koraka: projekcija afinom kamerom i uvođenje nelinearnih grešaka. Kada se traži geometrija scene potrebno je ispratiti ceo proces u suprotnom smeru. Kreće se od piksela na slici  ${}^a x$ , otklanjaju se nelinearne greške i pokušavaju se pronaći 3D koordinate tačke  $X$ .

Prvi korak, otklanjanje nelinearnih grešaka podrazumeva prelazak iz koordinata senzora u koordinate slike, tj.  ${}^a x \rightarrow {}^s x$ . Po prirodi same transformacije potrebno je iterativno rešenje.

Transformacija zavisi od ulaznih vrednosti koordinata koje se u ovom koraku traže, tako da pronalazak inverzne transformacije nije trivijalan problem. Iz razloga što mnoge deformacije sočiva ne deformišu ekstremno sliku, iterativni proces vrlo brzo konvergira ka tačnom rešenju.

Nakon ovog koraka dobijene su 2D koordinate u sistemu senzora. Krajnji cilj jeste dobijanje 3D koordinata, međutim prilikom perspektivne projekcije dolazi do gubitka informacija. To znači da nije moguće sa sigurnošću rekonstruisati tačno 3D koordinate svake tačke, već je moguće samo rekonstruisati zrak koji prolazi kroz 3D tačku, tj. pravac projektovanja.

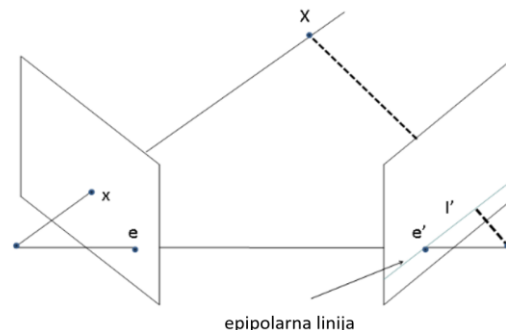
Ukoliko su poznate koordinate tačke nakon projekcije  ${}^s x$ , kao i eksterni parametri kamere koji se odnose na položaj kamere u prostoru  $X_0$ , rotaciju kamere  $R$  i interni parametri sačuvani u matrici  $K$ , moguće je naći pravac u 3D prostoru na kojem leži projektovana tačka.

$$\begin{aligned} X &= (KR)^{-1} KR X_0 + \lambda (KR)^{-1} {}^s x \\ &= X_0 + \lambda (KR)^{-1} {}^s x \end{aligned} \quad (10)$$

Jednačina 10 opisuje liniju koja počinje u tački  $X_0$ , a član  $\lambda (KR)^{-1} {}^s x$  opisuje pravac zraka iz centra koordinatnog sistema kamere  $X_0$  do 3D koordinata tačke  $X$ .

## 3. FUNDAMENTALNA I ESENCIJALNA MATRICA

Kada je poznat način na koji se tačka iz 3D sveta projektuje na ravan slike, potrebno je pronaći odnos između dve slike koje posmatraju istu scenu iz različitih uglova. Kada se dve kamere nalaze na horizontalnom odstojanju jedna od druge tada se takva postavka naziva stereo vizija.



Slika 3. Postavka kamera za stereo viziju

Na Slici 3 tačka  $X$  u globalnom koordinatnom sistemu projektuje se u tačku  $x$  na ravan slike. Projekcija tačke  $X$  nalaziće se uvek na istom mestu čak i kada bi se tačka pomerila duž ose projekcije u prostoru. To pokazuje da ukoliko je nužno pronaći projekciju tačke  $X$  na drugoj slici, potrebno je izvršiti pretragu dužinom projekcije ove linije na drugoj slici.

Ova zamišljena linija koja prolazi kroz tačku  $x$  naziva se epipolarna linija tačke  $x$ . Takva linija sa sobom povlači fundamentalno ograničenje koje pokazuje da se odgovarajući par tačke  $x$  iz prve slike na drugoj slici traži duž epipolarne linije tačke  $x$  u drugoj slici.

### 3.1. Fundamentalna matrica

Fundamentalnom matricom može se smatrati transformacija 2D tačke sa slike iz jednog ugla u epipolarnu liniju na slici iz drugog ugla. Fundamentalna matrica između para slika može se izračunati rešavanjem seta jednačina koje sadrže određeni broj poznatih uparenih tačaka između dve slike. Minimalan broj takvih parova je 7, a optimalan broj je 8. Zatim, za tačku u jednoj slici fundamentalna matrica daje jednačinu linije na kojoj bi se njena korespondentna tačka trebala nalaziti u drugoj slici. Ukoliko je tačka  $(x', y')$  korespondentna tačka tačke  $(x, y)$  i fundamentalna matrica između dve slike je matrica  $F$ , važi sledeća jednačina

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}^T F \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0 \quad (11)$$

Ova jednačina opisuje odnos između dve korespondentne tačke i poznata je kao epipolarno ograničenje.

### 3.2. Esencijalna matrica

Esencijalna matrica može se takođe nazvati specijalizacijom fundamentalne matrice, gde se matrica računa koristeći kalibrirane kamere što znači da je potrebno prethodno znanje o kamerama u sceni.

Iz tog razloga, da bi se izračunala esencijalna matrica, prethodno potrebna je matrica sa internim parametrima kamere (matrica koja sadrži optički centar i žižnu daljinu posmatrane kamere).

Proces pronalazjenja različitih parametara kamere jeste kalibracija kamere. Ideja je da se pred kamere postavi set tačaka za koje su poznate 3D pozicije u stvarnom svetu. Zatim se praćenjem projekcija tih tačaka na ravan slike, uz dovoljan broj 3D tačaka i pridruženih 2D tačaka, mogu izračunati tačni parametri kamere iz jednačine projekcije.

Kao i kod fundamentalne matrice i kod esencijalne matrice postoji ograničenje u obliku jednačine

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}^T E \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0 \quad (12)$$

gde je  $E$  esencijalna matrica koja sadrži parametre rotacije i translacije. Esencijalna matrica je dimenzija  $3 \times 3$ , što znači da sadrži ukupno 9 parametara. Kako je iz ograničenja koje važi i za fundamentalnu matricu jasno da će množenjem esencijalne matrice skalarom izraz ostati nepromenjen iz toga sledi da esencijalna matrica ima 8 promenljivih koje je potrebno izračunati [2]. To znači da je za rešavanje matrice  $E$  potrebno ukupno osam jednačina tj. osam ili više korespondencija (podudarnih tačaka). Isto ograničenje je zapisano u dužem obliku glasi:

$$\begin{bmatrix} x' & y' & 1 \end{bmatrix} \begin{bmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0 \quad (13)$$

Zatim sve poznate vrednosti prebacuju se u jednu matricu, a nepoznate u drugu. Cilj je zapisati jednačinu u formi  $Ax = 0$  gde je  $x = (E_{11}, E_{12}, \dots, E_{33})$ . U praksi, matrica  $A$  će imati više redova (biće dimenzije  $n \times 9$ ), gde se svaki red matrice javlja kao jedna korespondencija. Matrica  $A$  mora imati najmanje osam redova, jer kao što je prethodno napomenuto, potrebno je najmanje osam korespondencija. Zatim, rešava se sistem homogenih jednačina  $Ax = 0$ . Ako se ignoriše trivijalno rešenje  $x = 0$ , moguće je pronaći jedinstveno rešenje za  $x$  koje daje rešenje najmanjeg kvadrata za

$$\sum \left( \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}^T E \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \right)^2 \quad (15)$$

Rezultat dobijen ovakvim postupkom može biti nestabilan usled loših brojevnih uslova. Mala količina šuma u slici može da ima veliki uticaj na grešku u rezultatu.

#### 4. 3D REKONSTRUKCIJA SCENE

U zavisnosti od izabrane funkcije za pronalazjenje obeležja u slici kao i od vrednosti parametara u funkciji dobijaju se različiti rezultati. Najbolji rezultati, prikazani na Slici 4, dobijeni su korišćenjem funkcije SURF.



Slika 4. Originalna scena i njena 3D rekonstrukcija

#### 5. ZAKLJUČAK

Prepoznavanje važnih obeležja u slici je težak zadatak u kome se izazovi javljaju zbog promene osvetljenja u sceni, zaklonjenosti obeležja nekim drugim predmetom ili zamućenja kao posledice pokreta kamere. Sa povećanjem broja kamera u sceni matematičke jednačine koje se kriju iza kalibracije postaju kompleksnije.

Kako bi se resili navedeni problemi, potrebno je obaviti dalje istraživanje na temu kalibracije, kako jedne kamere, tako i više kamera u sceni, kao i detaljnije testirati koji tipovi sočiva daju bolje rezultate kalibracije u određenim situacijama. Osim kamere postoje načini za unapređenje algoritama za uparivanje tačaka kao i za njihovo filtriranje. U oblasti percepcije sa vizijom vrše se ogromna istraživanja uz pomoć napretka u poljima mašinskog učenja i dubokog učenja.

#### 6. LITERATURA

- [1] A. Baksheev, „Viz - New 3D visualization module in the OpenCV library“, 28 March 2014. Dostupno na: <https://sudonull.com/post/113808-Viz-New-3D-Visualization-Module-in-OpenCV-Library-Intel-Blog>. [Poslednji pristup 28 April 2021].
- [2] W. Hoff. EGGN 512. Class Lecture, Topic: “Structure from Motion.” University of Mines and Technology, Golden, Colorado, Mar. 25, 2012.

#### Kratka biografija:



**Aleksandra Ujaluši** rođena je 1994. godine u Vrbasu. Osnovne akademske studije na Fakultetu tehničkih nauka završila je 2019. godine i iste godine upisala je master akademske studije na Fakultetu tehničkih nauka, smer Animacija u inženjerstvu.



**Željen Trpovski** rođen je u Rijeci 1957. godine. Doktorirao je na Fakultetu tehničkih nauka 1998. god. Oblast interesovanja su telekomunikacije i obrada signala.

#### Zahvalnica:

Izradu ovog rada pomogao je Fakultet tehničkih nauka u Novom Sadu, Departman za energetiku elektroniku i telekomunikacije, u okviru projekta pod nazivom: "Istraživanja u oblasti energetike, elektronike, telekomunikacija i primenjenih informacionih sistema u cilju modernizacije studijskih programa".