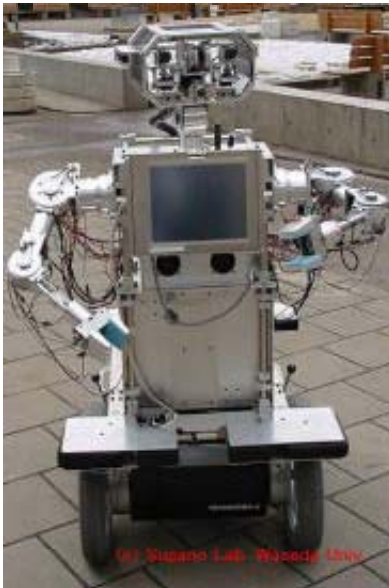


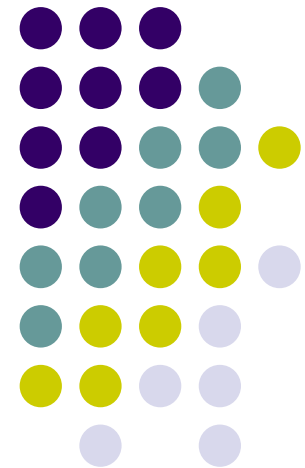
# Lekcija 5: *Robotska vizija*



Prof.dr.sc. Jasmin Velagić  
Elektrotehnički fakultet Sarajevo

Kolegij: Mobilna robotika

2012/2013



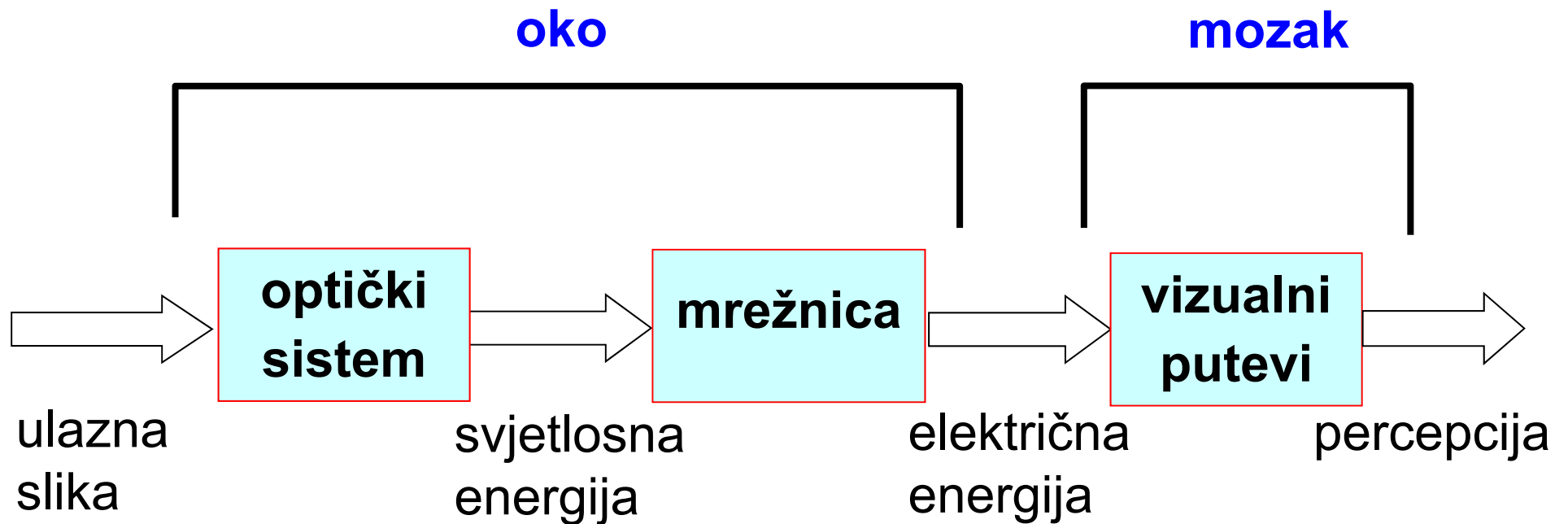
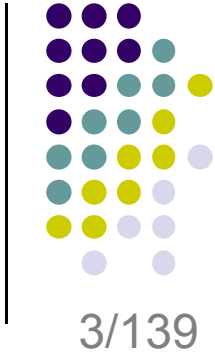
## 5.1. Uvod

- Ljudski vizualni sistem prihvaća informacije u obliku kodiranih svjetlosnih signala i transformira ih u elektrohemijske signale koji se dalje prenose neuronima u čovjekovom mozgu.
- **U biološkim znanostima se proučava kako se ove neuronske informacije obrađuju i na temelju njih grade interni prikazi okruženja i kako ljudi zaključuju o svojoj okolini na temelju ovih prikaza.**
- Sa stajališta robotike, **područja računarske ili robotske vizije istražuju oblasti računarskog predstavljanja okoline na temelju svjetlosnih signala i proučavaju tehnike umjetne inteligencije koje obavljaju zadatke zaključivanja ili planiranja na temelju dobivenog prikaza okoline.**



# Uvod

- Mehanizam ljudskog percipiranja okoline, odnosno objekata u njoj.
- Model ljudskog vizualnog sistema.



# Uvod

- **Glavni cilj sistema robotske vizije je oponašati sposobnosti i značajke ljudskog vizualnog sistema.**
- Kako je vizija najmoćniji način opažanja okoline slijedi da vizualni robotski sistemi osiguravaju veliki broj informacija o robotskom okruženju na temelju kojih se može ostvariti inteligentna interakcija u dinamičkoj sredini.
- Pri tome treba voditi računa da je **obrada podataka i gradnja mape okruženja na temelju vizualnih informacija značajno složenija od istih sa nevizualnim sensorima.**
- Temeljni procesi u robotskoj viziji: **obrada i analiza slike.**



# Uvod



- **Obrada i analiza slike** predstavljaju važne procese kojima se transformira slika u prihvatljiv oblik i izdvajaju, odnosno prepoznaju odgovarajuće značajke slike koje se mogu koristiti u različitim procesima unutar navigacijskog sistema mobilnog robota.
- **Digitalna obrada slike** je proces podvrgavanja numeričkih reprezentacija objekata u slici seriji operacija s ciljem postizanja željenog rezultata.
- Ulaz i izlaz u procesu obrade slike predstavljaju sliku, dok se sam proces obrade sastoji od **poboljšavanja** i **obnavljanja** slike.



# Uvod

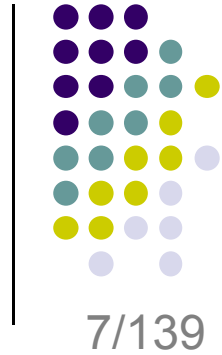
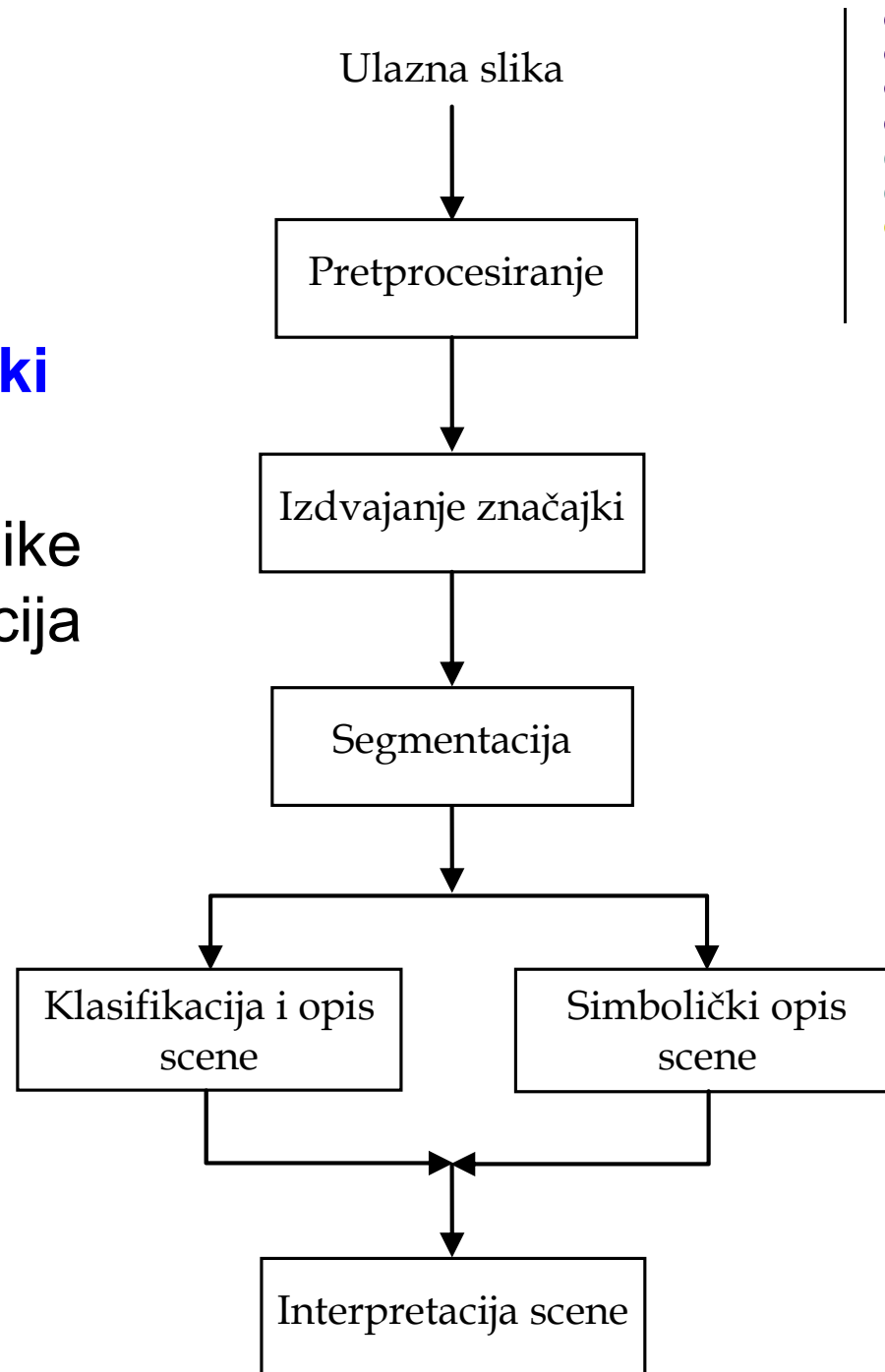
## Digitalna obrada slike obuhvaća sljedeće procese:

- **Operacije na slici** (konvolucija, Fourierova transformacija, Z transformacija,...).
- **Transformacije slike** (linearne, ortogonalne, diskretne).
- **Metode poboljšanja slike** (operacije na tački, prostorne operacije, upotreba transformacija, tehnike pseudokoloriranja, filtriranje – linearno i homomorfno,...).
- **Obnavljanje slike** (modeli degradacije slike, Inverzni i Wienerov filter, pseudoinverzija matrica, ...).
- **Rekonstrukciju slike iz projekcija** (Radonova transformacija, operator povratne projekcije, projekcijski teorem, inverzna Radonova transformacija, Hilbertova transformacija, Fourierova metoda rekonstrukcije, ...).
- **Kompresiju slike** koja pokušava smanjiti broj bitova potrebnih za pohranu slike bez ili sa gubitkom informacija.



# Uvod

- **Izlaz iz sistema za analizu nije slika, već numerički ili simbolički opis njenog sadržaja.**
- U sistemu za analizu slike postoje tri grupe operacija ***izdvajanje značajki***, ***segmentacija*** i ***opis scene***.



# Uvod

## Postupci u procesu analize slike:

- **Raspoznavanje uzoraka** (pravilo najbližeg susjeda, Bayesovo pravilo, algoritmi za grupiranje,...).
- **Izdvajanje (ekstrakcija) značajki** slike (prostorne osobine, osobine u domeni transformirane slike, detekcija rubova, detekcija granica).
- **Segmentacija slike** (amplitudna, obilježavanje komponenti, granice objekata, unutarjost objekta, grupiranje, ekspertni sistemi, neuronske mreže,... ).
- **Matematička morfologija** (binarna i siva).
- **Analiza tekstura, oblika i pokreta.**
- **Registracija slike** (klasifikacija metoda za registraciju, pregled geometrijskih transformacija i algoritama).
- **Stereovizija** (kalibracija kamere, stereovizija, problem korespondentnih tačaka).





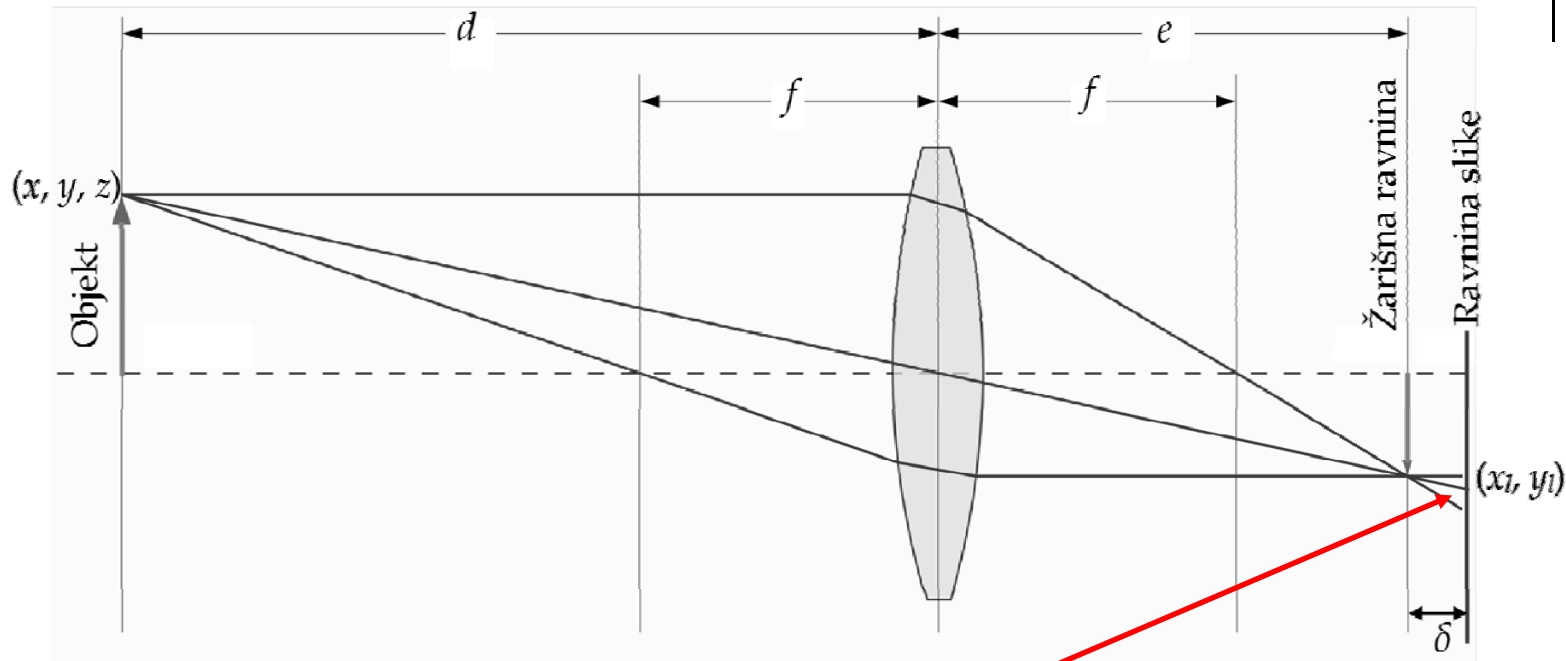
## 5.2. Proces formiranja slike

- Većina vizualnih robotskih sistema koriste kamere za opažanje prostora.
- Opažanje područja je izuzetno važno u mobilnoj robotici za izbjegavanje prepreka pri obavljanju nekog zadatka.
- Međutim, **kod vizualnog sistema (kamera) izražen je problem gubitka informacija o dubini.**
- Ako se mogu predvidjeti informacije o dimenzijama objekta, ili pak o njegovoj boji i refleksiji, tada se mogu izvući informacije o dubini.
- **Svojstva slike ne ovise samo o promjenama u sceni već i o parametrima kamere.**



# Proces formiranja slike

- Veza između parametara kamere i svojstava slike.



Slika je rezultat refleksije svjetlosti od objekta u smjeru kamere

- Udaljenost leće do objekta  $d$  i razmak između leće i žarišta  $e$  povezani su žarišnom duljinom leće  $f$ :

$$\frac{1}{f} = \frac{1}{d} + \frac{1}{e}$$



# Proces formiranja slike

- **Slika je rezultat refleksije svjetlosti od objekta u smjeru kamere.**
- **Ako je udaljenost između leće i ravnine slike jednaka e tada se nakon refleksije na ravnini slike dobiva tačka.**
- **Kada ravnina slike nije na udaljenosti e od leće svjetlost će se od objekta reflektirati na ravnini slike u obliku zamagljenog kruga.**
- Prva aproksimacija polazi od toga da je svjetlost homogeno distribuirana duž tog kruga i njegov promjer iznosi:

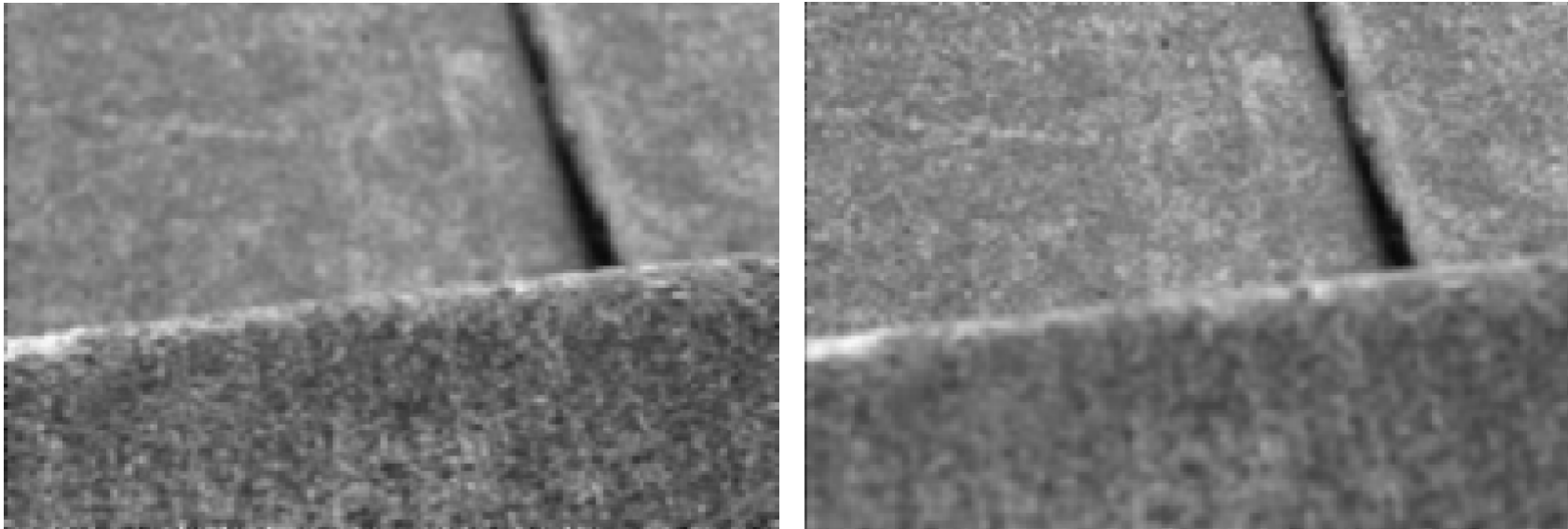
$$R = \frac{L\delta}{2e}$$

$L$  – promjer leće,  $\delta$  – razmak između ravnine slike i žarišne ravnine.



# Proces formiranja slike

- **Različit  $\delta$  sugerira različite pozicije fokusiranja za istu scenu.**
- Na slikama se zapaža različitost u oštrom rubova tekstone (bliži rubovi imaju veću oštrinu) – slike su dobivene za dvije različite vrijednosti  $\delta$ .



- **Iz niza slika iste scene, dobivenih sa različitim geometrijama kamere, može se rekonstruirati dubina slike.**



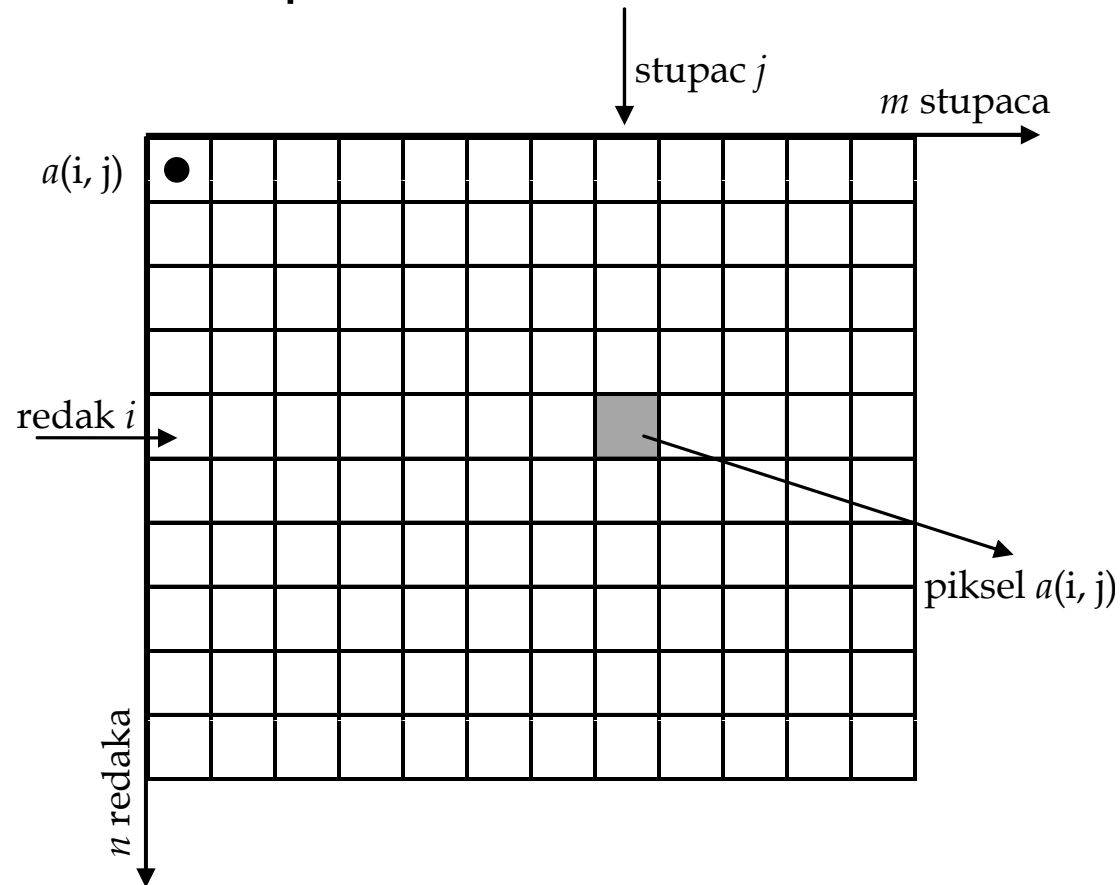
# Proces formiranja slike

- Stvarna slika je uzorkovana konačnim brojem tačaka (piksela).
- **Rezolucija uzorkovanja**, odnosno **rezolucija slike** predstavlja broj piksela digitalne slike.
- Tipične rezolucije su  $640 \times 480$  piksela,  $320 \times 240$  piksela, itd.
- **Kod rezolucije je također važno koliko često se mjerenja kamere prosljeđuju računaru.**
- Kod CCD kamera je standardni intenzitet mjerenja 30 slika/sekundi.
- Svaki piksel predstavlja pojedini uzorak na slici.
- **On je prikazan cjelobrojnom vrijednošću koja označava intenzitet (osvjetljenost) slike u toj tački.**



# Proces formiranja slike

- Prikaz slike u obliku dvodimenzionalnog polja.
- Indeksi piksela  $[i,j]$  predstavljaju cjelobrojne vrijednosti koje označavaju redak, odnosno stupac u kome se nalazi piksel.



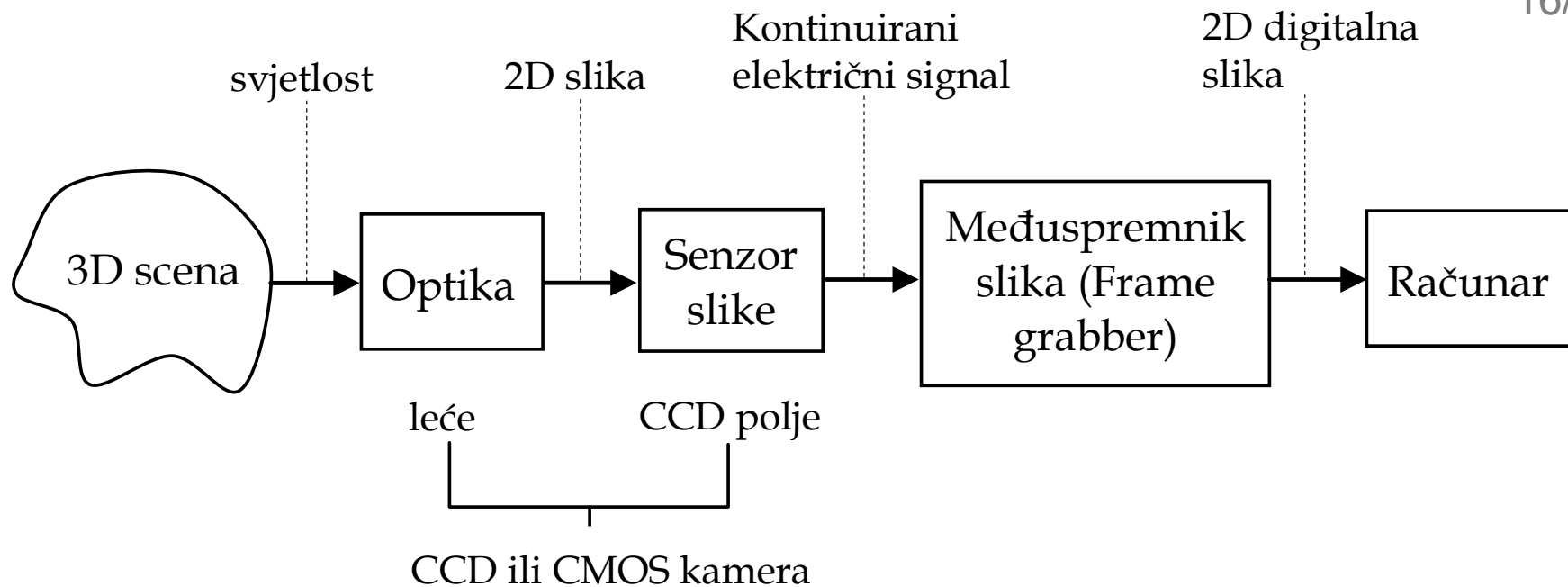
## Proces formiranja slike

- Osim rezolucije slike, važan parametar koji se odnosi na piksele je **kvantizacija slike**, odnosno razina intenziteta slike koja se koristi za prikaz piksela.
- Tipične razine kvantizacije slike iznose:  $2^8 = 256$ ,  $2^5 = 32$ , itd.
- Sistem vizije mobilnog robota obično se sastoji od dvije glavne komponente: **optičkog sistema** koji prikuplja svjetlost preko konačnog polja pogleda i fokusira ovu svjetlost na ravninu slike i **optoelektričkog sistema** koji preuzima ovu fokusiranu svjetlost i konvertira je za računar razumljive (čitljive) signale.



# Proces formiranja slike

- Sklopovlje sistema vizije prikazano je na slici.



- Proces kolekcije, odnosno prikupljanja svjetlosti i njenog pohranjivanja u računaru predstavlja složen proces.
- Zbog toga su i postupci koji se za njih koriste također složeni.



# Proces formiranja slike

- **Leće** predstavljaju zaseban modul optičkog sistema kamere.
- Važne specifikacije leće su: **žarišna udaljenost** (određuje širinu vidnog polja), **fokusiranje** (ručno/automatsko) i **distorzija leće** (problem korekcije distorzije).
- Slika trodimenzionalnog (3D) prostora dobivena kamerom oduzima prostoru jednu dimenziju, tako da slika postaje dvodimenzionalna (2D - **gubitak informacije o dubini**).
- Iz jedne slike prostora ne može se odrediti udaljenost od kamere do objekta u okruženju niti dimenzije objekta.



## Proces formiranja slike

- Za dobivanje ovih informacija mora se koristiti više slika.
- Pri tome treba naglasiti da rekonstrukcija 3D slike iz 2D slika nije jednoznačan problem.
- Rješavanje problema rekonstrukcije 3D slike zahtijeva:
  - **modeliranje procesa akvizicije slika (kalibracija kamere),**
  - **izdvajanje karakterističnih značajki iz slike,**
  - **identifikaciju odgovarajućih karakterističnih značajki iz dviju ili više slika.**



## 5.3. CCD i CMOS kamere

### CCD (Charged Coupled Device) kamera

- Danas najpopularniji sistem vizije.
- Osnovni element je CCD čip koji se sastoji od polja svjetlosno osjetljivih elemenata, ili piksela, obično između 20.000 i nekoliko miliona piksela (trenutno).
- Svaki piksel je dimenzija od  $5\mu\text{m} \times 5\mu\text{m}$  do  $25\mu\text{m} \times 25\mu\text{m}$  i ima mogućnost kapacitivnog punjenja, odnosno pražnjenja.
- Prvo se kondenzatori svih piksela u cijelosti napune, nakon čega započinje **proces integracije (generiranja)**.



## CCD i CMOS kamere

### CCD (Charged Coupled Device) kamera

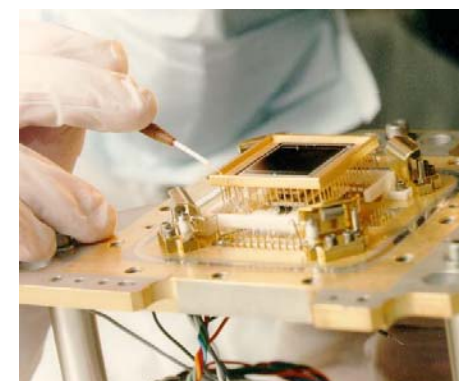
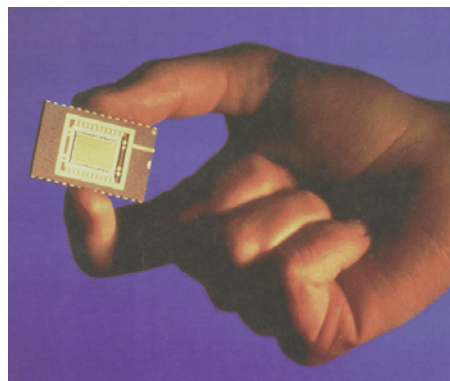
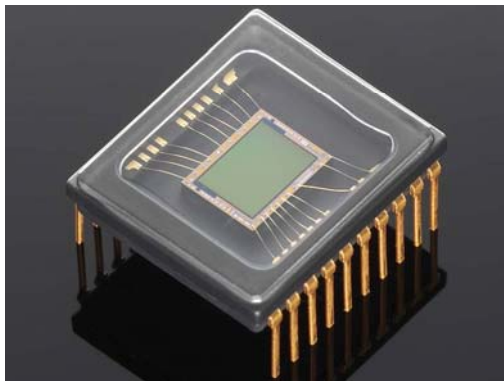
- Fotoni svjetlosti udaraju u svaki piksel, oslobađaju elektrone koji su obuhvaćeni električkim poljem i zaustavljaju se na pikselu (fotoelektrički efekt).
- Tokom vremena svaki piksel će akumulirati različit iznos energije na temelju ukupnog broja fotona koji su “udarili” u njega.
- Nakon što se period integracije kompletira, relativan iznos energije piksela se pohranjuje i čita – **proces kolekcije**.



# CCD i CMOS kamere

## CCD kamera - čip

- Dolazi u različitim oblicima i dimenzijama.
- Ukupan broj piksela -  $4008 \times 2672 = 11$  Mega Piksela.
- Veličina piksela – između 5 i 25  $\mu\text{m}$ .
- Veličina dijafragme – između 5 i 36 mm po osi.



- Fotodiode koje se koriste u CCD čipu nemaju jednaku osjetljivost na sve frekvencije svjetlosnog signala.



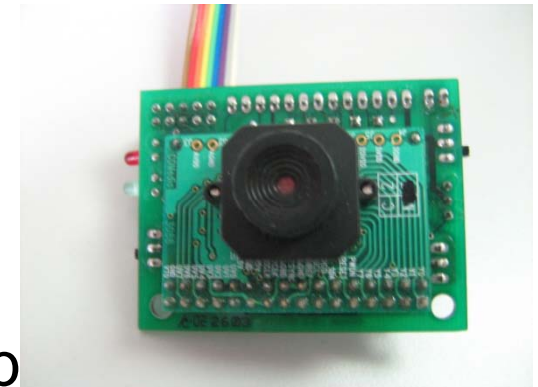
# CCD i CMOS kamere

## CMOS kamera

- CMOS čip se značajno razlikuje od CCD čipa.
- On također ima polje piksela, ali su oni locirani paralelno i svaki piksel koristi nekoliko tranzistora za svoje lociranje.
- **Proces integracije** (akumuliranje energije u pikselima) je isti kao kod CCD čipa.
- U **fazi kolekcije** CMOS koristi drukčiji pristup, gdje se svaki piksel mjeri i pojačava se njegov signal, pri čemu se to odvija paralelno za svaki piksel u polju.
- Rezultantne vrijednosti piksela se prenose do njihovih destinacija.



22/139



## CCD i CMOS kamere

- Obje tehnologije sistema vizije, CCD i CMOS, generiraju digitalne signale koji se mogu direktno iskoristiti u robotici.
- Čip u kameri posjeduje paralelne digitalne ulazne i izlazne pinove preko kojih se prenose diskretne vrijednosti sadržaja piksela.
- Neki moduli vizualnog sistema koriste ove digitalne signale i obrađuju ih u stvarnom vremenu.
- Za ovu svrhu koristi se čip koji predstavlja međuspremnik slika (frame grabber) između digitalnih izlaza kamere i digitalnih ulaza u računar. Ovakvi čipovi "hvataju" kompletnu sliku i omogućuju pristup pikselima, obično u pojedinačnom poretku.



# CCD i CMOS kamere

## Vizualni senzori za praćenje boja

- Važan aspekt opažanja vizualnog sistema je detektiranje i praćenje odgovarajuće boje (objekta) u okolini.
- Opažanje boje ima dvije važne prednosti:
  - Detekcija boje je funkcija pojedinačne slike.
  - Omogućuje kombinaciju sa drugim izvorima, recimo stereovizijom ili laserskim senzorom, kako bi se značajno poboljšale informacije.
- Danas postoje komercijalne verzije kamera sa dodatnom elektronikom koja omogućuje detektiranje i praćenje objekta na temelju njegove boje.





# CCD i CMOS kamere

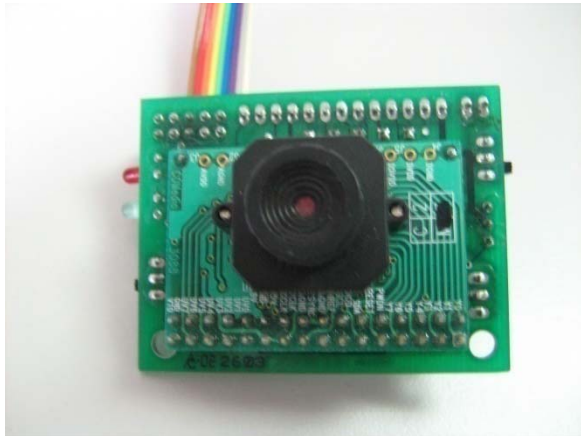
## Vizualni senzori za praćenje boja

- Vrlo je interesantan CMUcam vizualni senzor za robote.
- Ovaj vizualni sistem se sastoji od CMOS kamere, mikroprocesora koji radi na 50 MHz.
- Radi se o jeftinom inteligentnom vizualnom sistemu (smart kamera).
- Ovaj sistem osigurava visokokvalitetne informacije izvučene iz slike navedenom, vanjskom procesoru, koji obrađuje dobivene informacije i može upravljati, recimo, mobilnim robotom.
- Vanjski procesor konfigurira mod podataka koji se koristi za praćenje.

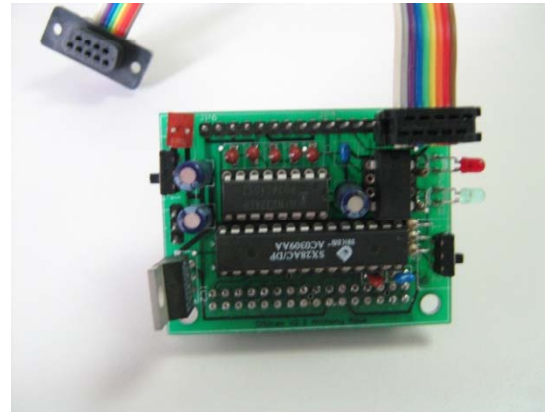


# CCD i CMOS kamere

## Vizualni senzori za praćenje boja

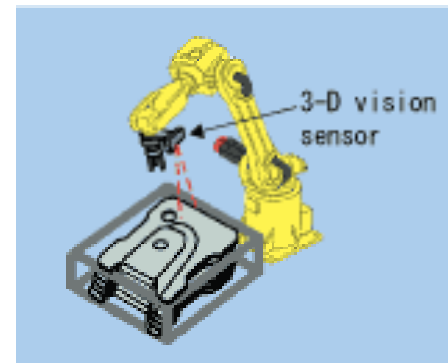


CMOS kamera



+ Mikroprocesorski bazirani sistem

## 3D senzor – primjena u robotici (za inspekciju)



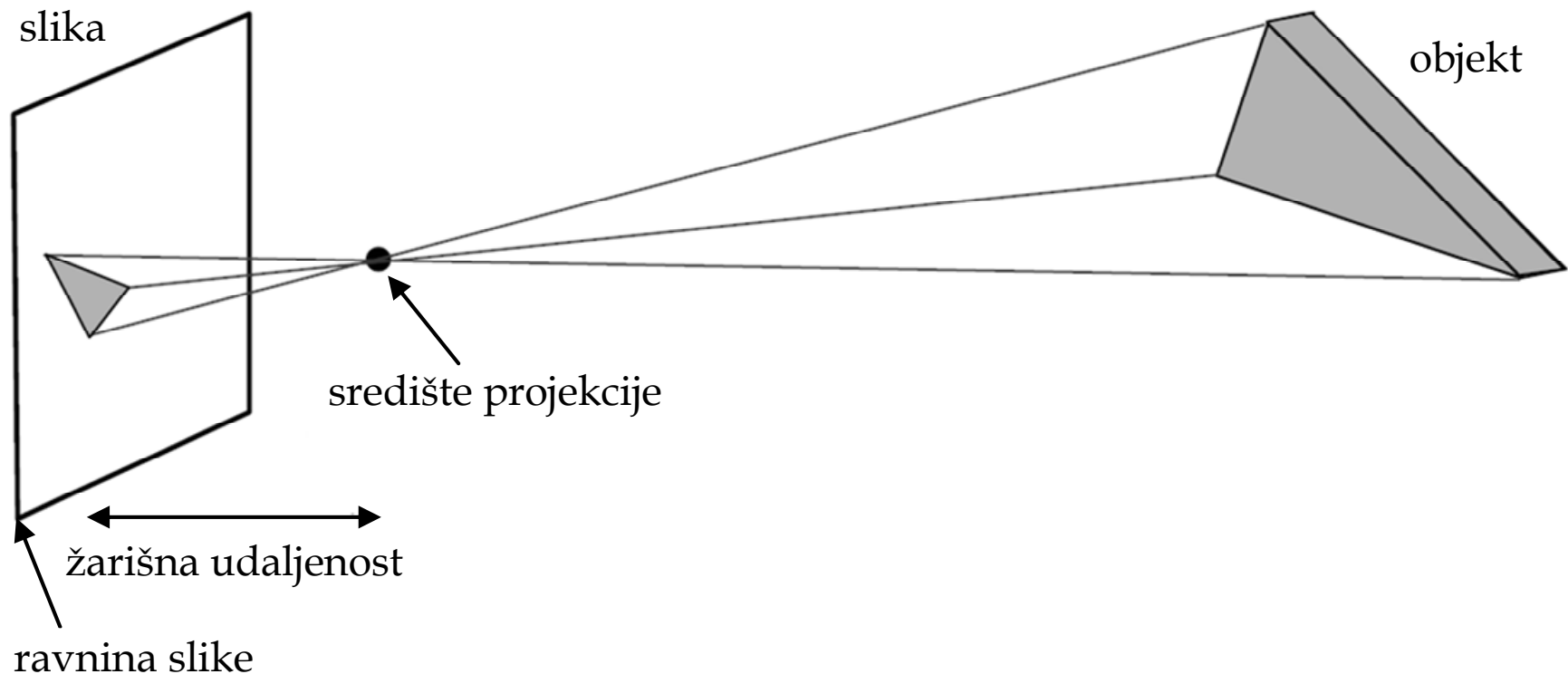


## 5.4. Kalibracija kamere

- **Redukcija trodimenzionalne pozicije tačke u njenu dvodimenzionalnu projekciju unutar kamere predstavlja fundamentalni problem korištenja vizije u mobilnoj robotici.**
- Za prevazilaženje ovog problema korišteni su mnogi mehanizmi koji **sužuju skup mogućih vrijednosti udaljenosti do objekta, upotrebu više kamera ili korištenje vremenske integracije.**
- Ključni koncept u transformaciji 3D slike u njenu 2D projekciju predstavlja ***perspektivna projekcija.***

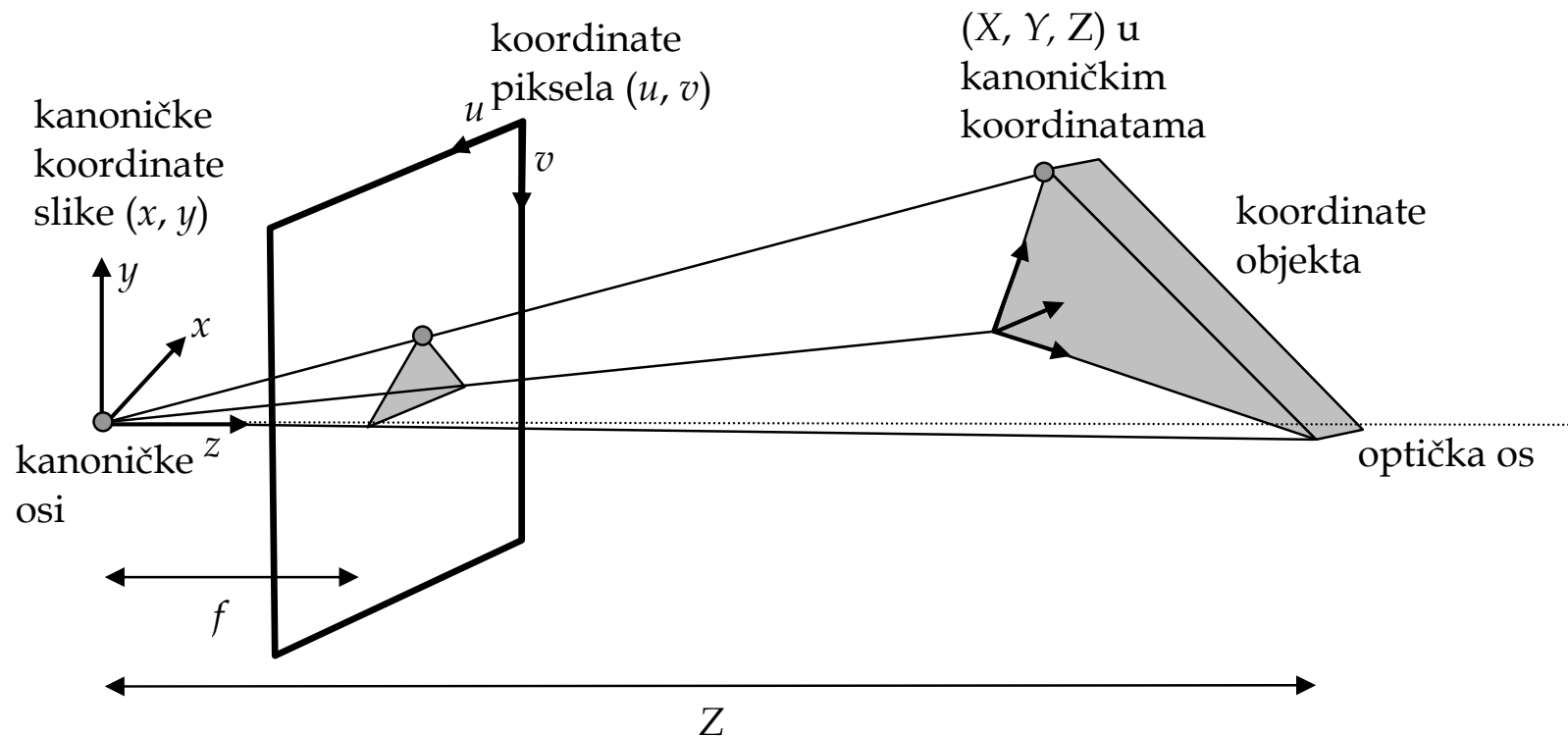
# Kalibracija kamere

- **Perspektivna projekcija** - svaka tačka slike se projicira po pravcu kroz središte projekcije, budući da se svjetlost prostire pravolinijski.
- Koncept perspektivne projekcije je ekvivalentan projekciji takozvane "pinhole" kamere.



# Kalibracija kamere

- Bolji geometrijski prikaz se dobiva ako se **slika pomakne ispred optičkog središta kamere** (središte projekcije).
- Tada se kanoničke osi slike postavljaju u središte projekcije.



## Kalibracija kamere

- Veza između kanoničkih koordinata slike  $x$  i  $y$  i koordinata objekta u prostoru  $X$ ,  $Y$  i  $Z$  dana je na sljedeći način:

$$x = \frac{fX}{Z}, \quad y = \frac{fY}{Z}$$

- **Povezanosti navedenih koordinata predstavljaju nelinearne transformacije.**
- Osnovni zadatak se sastoji u određivanju koordinata  $X$ ,  $Y$  i  $Z$  iz koordinata  $x$  i  $y$ , odnosno  $u$  i  $v$ .



## Kalibracija kamere

- Neka je zadana proizvoljna tačka u prostoru izražena u homogenim koordinatama  $X=[X_1 \ X_2 \ X_3 \ 1]^T$ , tada njene kanoničke koordinate u slici  $(x_1/x_3, x_2/x_3)$  iznose:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \tilde{P} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ 1 \end{bmatrix} \quad (*)$$

gdje je  $\tilde{P}$  **projekcijska matrica**.



# Kalibracija kamere

- Projekcijska matrica transformira 3D koordinatni sistem realnog svijeta u 3D koordinatni sistem koji je poravnat sa optičkim središtem (centrom) i idealnom ravninom slike kamere, a rezultat je perspektivne projekcije na idealni 2D koordinatni sistem ravnine slike.
- Neki parametri kamere ne ovise o vrsti upotrijebljene kamere i oni se nazivaju **vanjski parametri** kamere, dok se parametri koji se mijenjaju ovisno o vrsti kamere nazivaju **unutarnji parametri** kamere.
- Proces dobivanja, odnosno određivanja projekcijske matrice za određenu kameru naziva se **kalibracija kamere**.





# Kalibracija kamere

- Postoje različiti mehanizmi kalibracije pojedinačne kamere, ali je možda najjednostavniji i najviše korišteni postupak korištenja **kalibracijskog objekta sa poznatom 3D pozicijom i projekcijom**.
- Za zadanu skupinu kalibracijskih tačaka u koordinatnom sistemu realnog svijeta  $X_i$  (napravljeno više slika sa kalibracijskim objektom) odrede se njihove projekcije u ravnini slike  $(u_i, v_i)$ .
- Nakon toga se izraz (\*) može napisati kao:

$$\begin{bmatrix} u_i x_3 \\ v_i x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} \tilde{P}_{11} & \tilde{P}_{12} & \tilde{P}_{13} & \tilde{P}_{14} \\ \tilde{P}_{21} & \tilde{P}_{22} & \tilde{P}_{23} & \tilde{P}_{24} \\ \tilde{P}_{31} & \tilde{P}_{32} & \tilde{P}_{33} & \tilde{P}_{34} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ 1 \end{bmatrix} = \begin{bmatrix} \tilde{P}_{11} X_1 & \tilde{P}_{12} X_2 & \tilde{P}_{13} X_3 & \tilde{P}_{14} \\ \tilde{P}_{21} X_1 & \tilde{P}_{22} X_2 & \tilde{P}_{23} X_3 & \tilde{P}_{24} \\ \tilde{P}_{31} X_1 & \tilde{P}_{32} X_2 & \tilde{P}_{33} X_3 & \tilde{P}_{34} \end{bmatrix} = \begin{bmatrix} (\tilde{P}X_i)_1 \\ (\tilde{P}X_i)_2 \\ (\tilde{P}X_i)_3 \end{bmatrix}$$

gdje je  $u = x_1/x_3$  i  $v = x_2/x_3$



## Kalibracija kamere

- Da bi se preslikavanje  $\tilde{P}X_i$  obavilo sa najmanjom pogreškom u slici uzima se da je  $\tilde{P}_{3,4} = 1$  i primijenjuje se **metod najmanjih kvadrata** (LS metod – engl. least square) koji minimizira sljedeći izraz:

$$\sum \left( (\tilde{P}X_i)_1 - (\tilde{P}X_i)_3 u_i \right)^2 + \left( (\tilde{P}X_i)_2 - (\tilde{P}X_i)_3 v_i \right)^2$$

- Rješenje minimizacije se dobiva rješavanjem matrične jednadžbe  $AX = B$ , gdje je:

$$A = \begin{bmatrix} -X_i & -Y_i & -Z_i & -1 & 0 & 0 & 0 & 0 & u_i X_i & u_i Y_i & u_i Z_i \\ 0 & 0 & 0 & 0 & -X_i & -Y_i & -Z_i & -1 & v_i X_i & v_i Y_i & v_i Z_i \\ \dots & & & & & & & & & & \end{bmatrix}$$

$$X^T = [\tilde{P}_{1,1} \quad \tilde{P}_{1,2} \quad \tilde{P}_{1,3} \quad \tilde{P}_{1,4} \quad \tilde{P}_{2,1} \quad \tilde{P}_{2,2} \quad \tilde{P}_{2,3} \quad \tilde{P}_{2,4} \quad \tilde{P}_{3,1} \quad \tilde{P}_{3,2} \quad \tilde{P}_{3,3}]$$

# Kalibracija kamere



$$B = \begin{bmatrix} -u_i \\ -v_i \\ \dots \end{bmatrix}$$

- Rješenje LS metode je:

$$X = (A^T A)^{-1} A^T B$$

- Optičko središte kamere ( $C_x, C_y, C_z$ ) u koordinatnom sistemu prostora (objekta) definira se na sljedeći način:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \tilde{P} \begin{bmatrix} C_x \\ C_y \\ C_z \\ 1 \end{bmatrix}$$

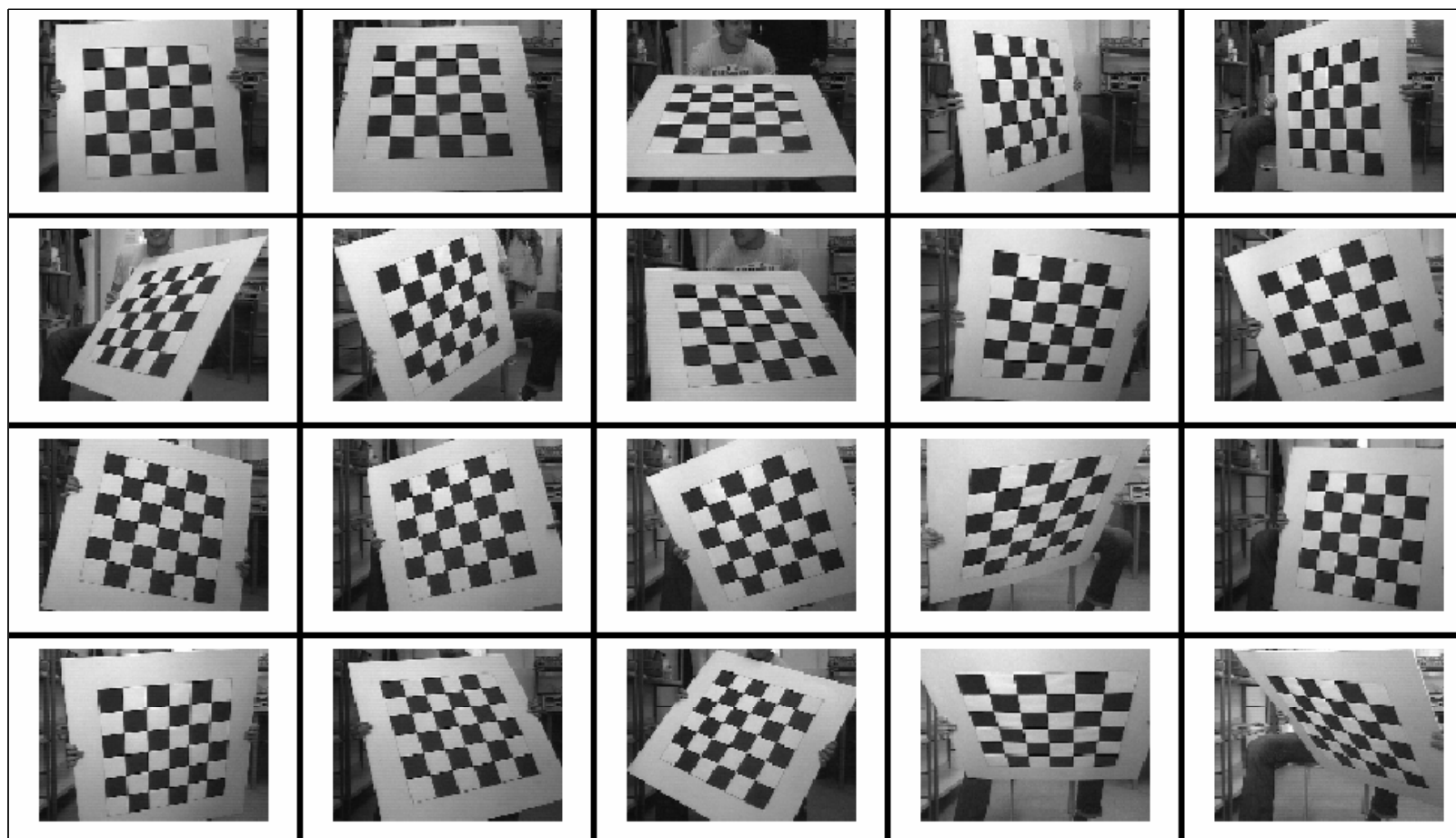
## Kalibracija kamere

- Svaki redak matrice  $\tilde{P}$  definira ravninu koja prolazi kroz tačku optičkog središta kamere.
- Presjecište  $\tilde{P}_1$  i  $\tilde{P}_2$  definira usmjerenost kamere, a presjecište  $\tilde{P}_2$  i  $\tilde{P}_3$  zakret kamere oko te osi.



# Kalibracija kamere

- **Primjer 1.** Kalibracija kamere korištenjem kalibracijskog objekta.
- Kao kalibracijski objekt koristi se šahovska ploča.



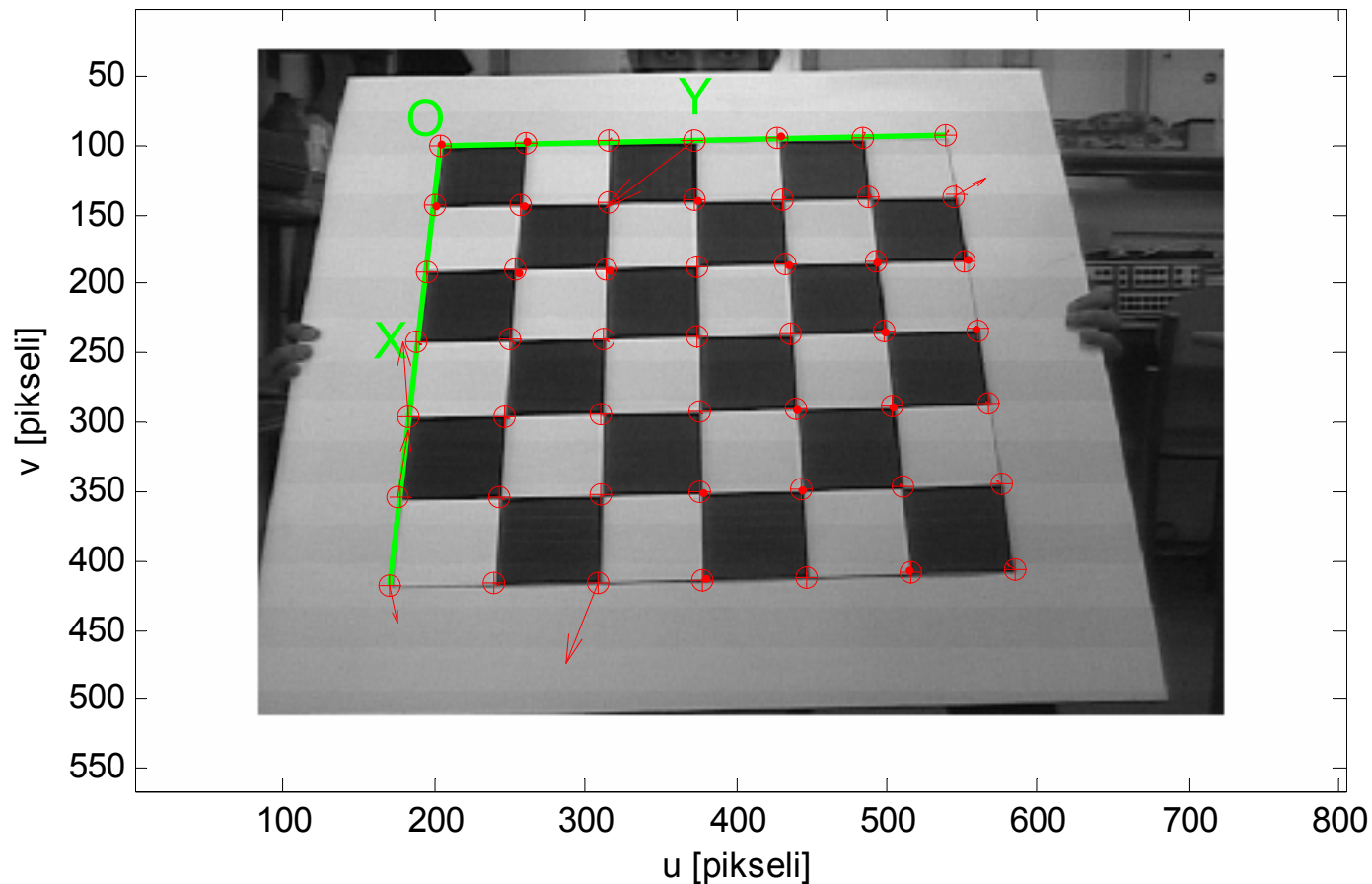
# Kalibracija kamere

- **Proces kalibracije** = određivanje unutarnjih (intrinzični) i vanjskih (ekstrinzičnih) parametara.
- U postupku kalibracije prvo se detektiraju vrhovi šahovskih polja na svim slikama te se nacрта mreža na slici šahovske ploče kako bi trebala izgledati bez distorzije.
- Ukoliko je neki ugao predaleko od stvarnog, onda se pomoću distorzijskih koeficijenata ručno podešavaju uglovi.
- Na sljedećoj slici prikazan je primjer detektiranja uglova za drugi uzorak sa prethodne slike.
- Ideničan postupak se provodi i za ostale uzorke.
- Stvarni uglovi na slici označeni su znakom +, kružići predstavljaju reprojicirane uglove, dok strelice predstavljaju gradijent pogreške.



# Kalibracija kamere

- Nakon toga se te informacije o položaju uglova koriste u procesu određivanja ekstrinzičnih i intrinzičnih parametara, odnosno minimizaciji srednje kvadratne pogreške.



Stvarni vrhovi slike (+) i detektirani vrhovi na projiciranoj mreži (o)

# Kalibracija kamere

## Dobiveni intrinzični parametri

- **Žarišna duljina:**

$$(f / \rho_w, f / \rho_h) = (913.49376, 921.60192) \pm (1.50531, 1.61318)$$

- **Centar slike:**  $(u_0, v_0) = (393.65084, 232.70717) \pm (2.91469, 2.77715)$

- **Iskošenost senzora:**  $\alpha = 0.00000 \pm 0.00000$

- **Distorzija:**

$$k_c = [-0.13214 \quad 0.39384 \quad -0.00938 \quad -0.00309 \quad 0.00000]^T$$

$$\pm [0.01276 \quad 0.08875 \quad 0.00087 \quad 0.00089 \quad 0.00000]^T$$

- **Pogreška piksela:**  $err = [0.19534 \quad 0.19198]^T$





# Kalibracija kamere

## Dobiveni ekstrinzični parametri

- Vektor translacije:

$$\mathbf{p}_t = [-269.983575 \quad -74.307502 \quad 1712.84948 \quad 0]^T$$

- Vektor rotacije:  $\mathbf{R}_v = [1.902719 \quad 1.757589 \quad -0.043652]^T$

- Matrica rotacije:

$$\mathbf{R} = \begin{bmatrix} 0.147023 & 0.931663 & 0.332249 \\ 0.914020 & 0.000430 & -0.405668 \\ -0.378089 & 0.363325 & -0.851495 \end{bmatrix}$$

- Pogreška piksela:  $\mathbf{err} = [0.27080 \quad 0.34822]^T$



# Kalibracija kamere

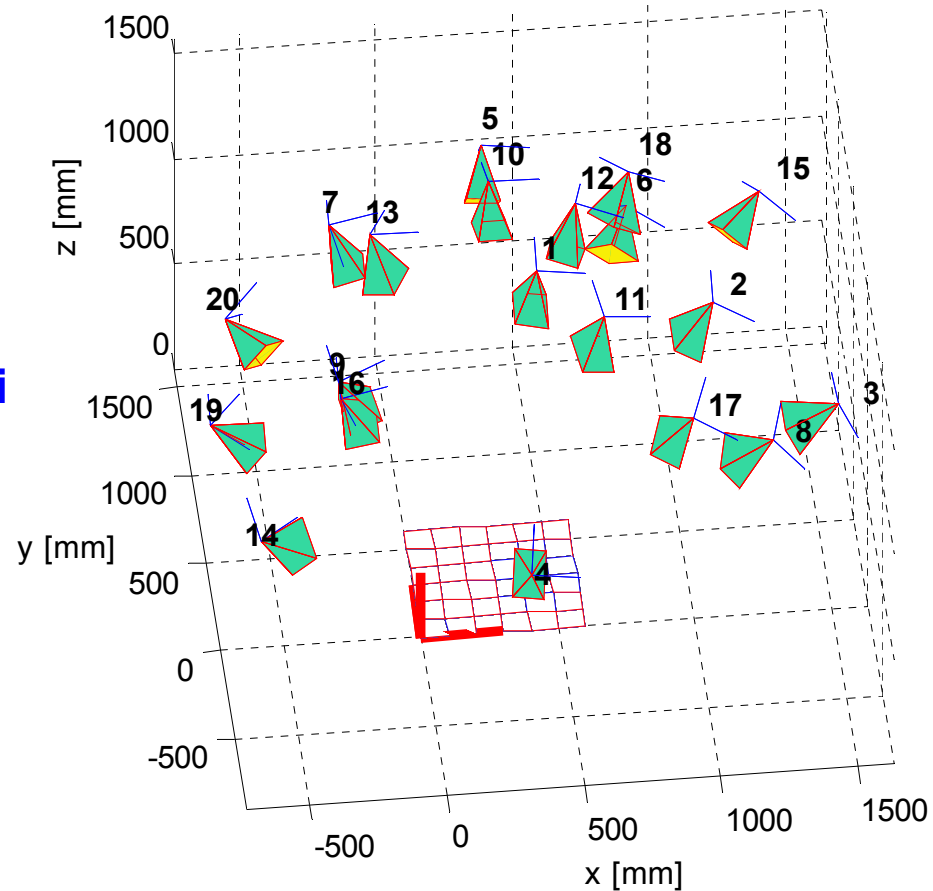
- Vektor  $k_c$  određuje distorziju kamere, odnosno definira odstupanje položaja piksela u stvarnoj slici u odnosu na istu dobivenu na idealnoj ravnini slike.
- Vektor rotacije  $R_v$  iskazuje smjer rotacije slike u 3D prostoru oko optičke osi (os isijavanja).
- Matrica  $R$  opisuje rotaciju slike u 3D prostoru u odnosu na koordinatni sistem kamere.
- Vektor translacije  $p_t$  definira pomak slike u odnosu na idealnu ravninu slike.
- Intervali pouzdanosti intrinzičnih parametara dobivaju se na temelju minimiziranja odstupanja piksela većeg broja uzoraka iste slike dobivenih iz različitih uglova snimanja.



# Kalibracija kamere

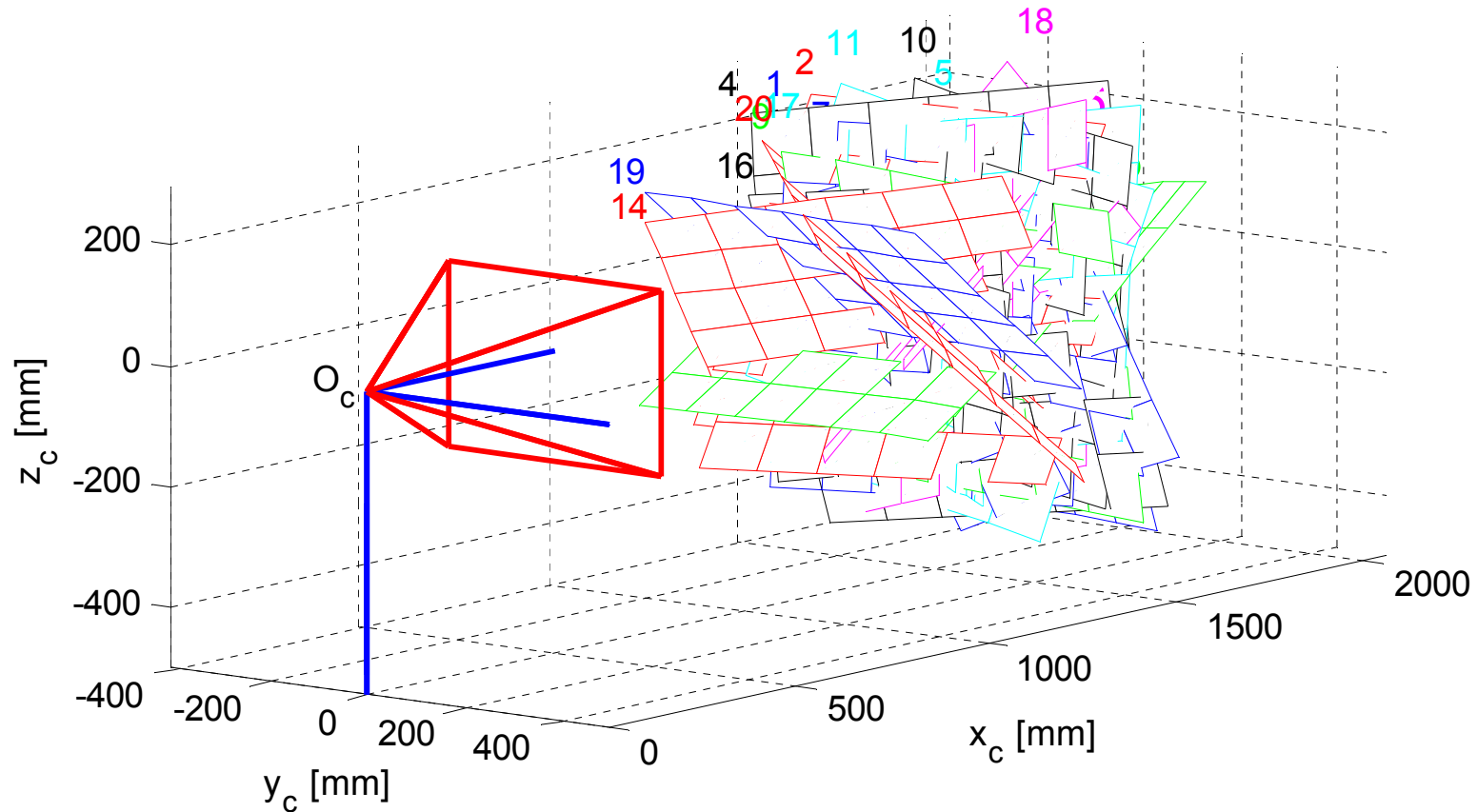
- Estimirane lokacije kamere u odnosu na šahovsku ploču za svih 20 uzoraka, odnosno pogleda kamere prema šahovskoj ploči.
- Estimirane lokacije određene su matricom rotacije i vektorom translacije.

**Ekstrinzični parametri (pozicija i orijentacija šahovske ploče u odnosu na koordinatni sistem kamere).**



# Kalibracija kamere

- 3D prikaz pozicije i orijentacije uzoraka šahovske ploče u koordinatnom sistemu kamere, gdje je središte kamere, odnosno ishodište pridruženog joj koordinatnog sistema označeno sa  $O_c$ .



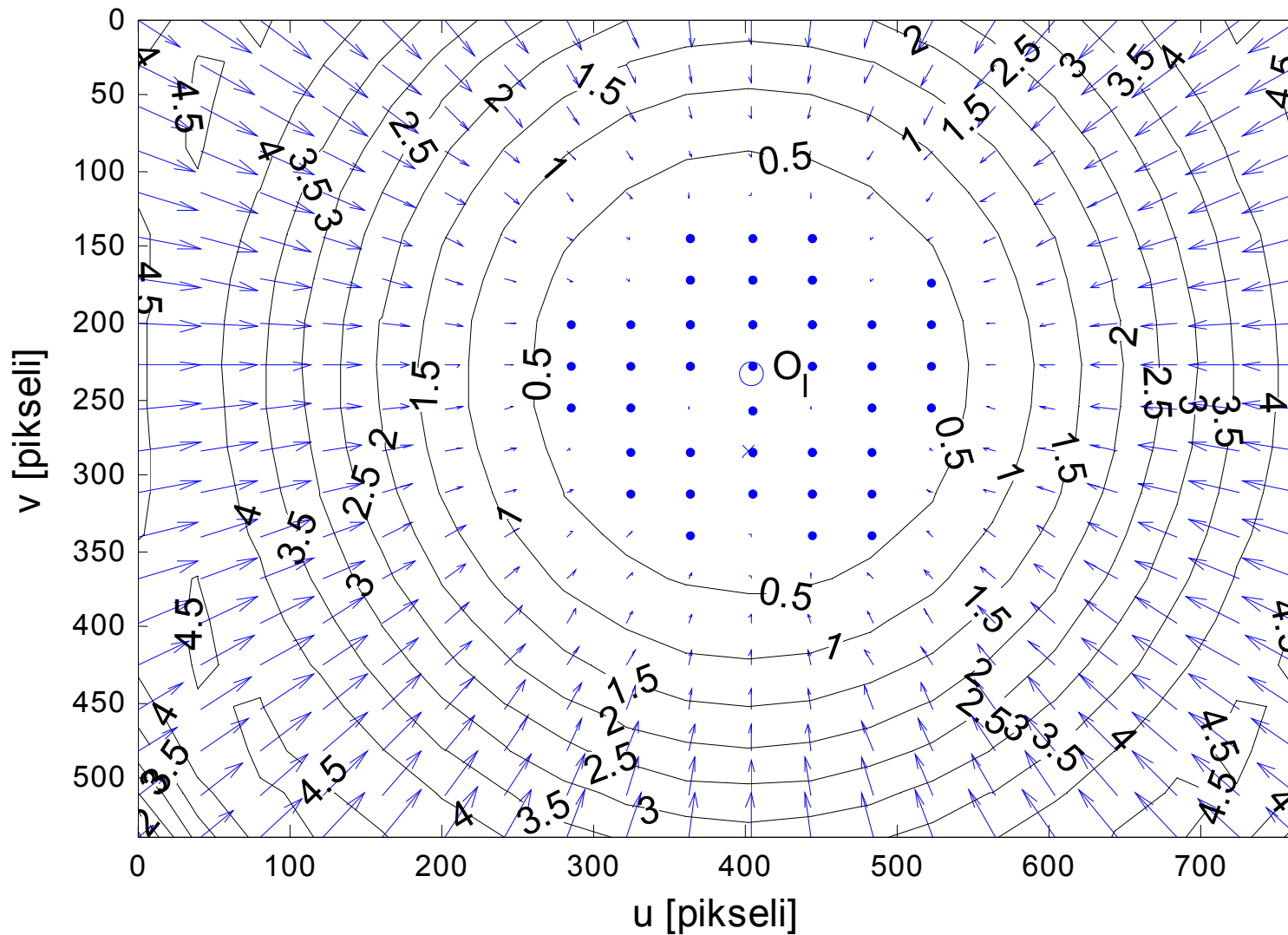
# Kalibracija kamere

- Na sljedećem slajdu prikazana radijalna distorzija leće kamere.
- Općenito, distorzija u optici leća označava odstupanje od idealnog modela projekcije gdje ravne linije u sceni ostaju ravne linije na projekcijskoj ravnini.
- Kod kamere su najviše izražene radijalna i tangencijalna distorzija.
- Radijalna distorzija pokazuje kako se pomjeraju tačke šahovske ploče u  $xyz$  koordinatnom sistemu prilikom akvizicije slike i koristi se za estimaciju stvarnih pozicija i orijentacija šahovske ploče.
- Ona je jednaka nuli u centru leće te se povećava kako se udaljavamo od centra prema periferiji.



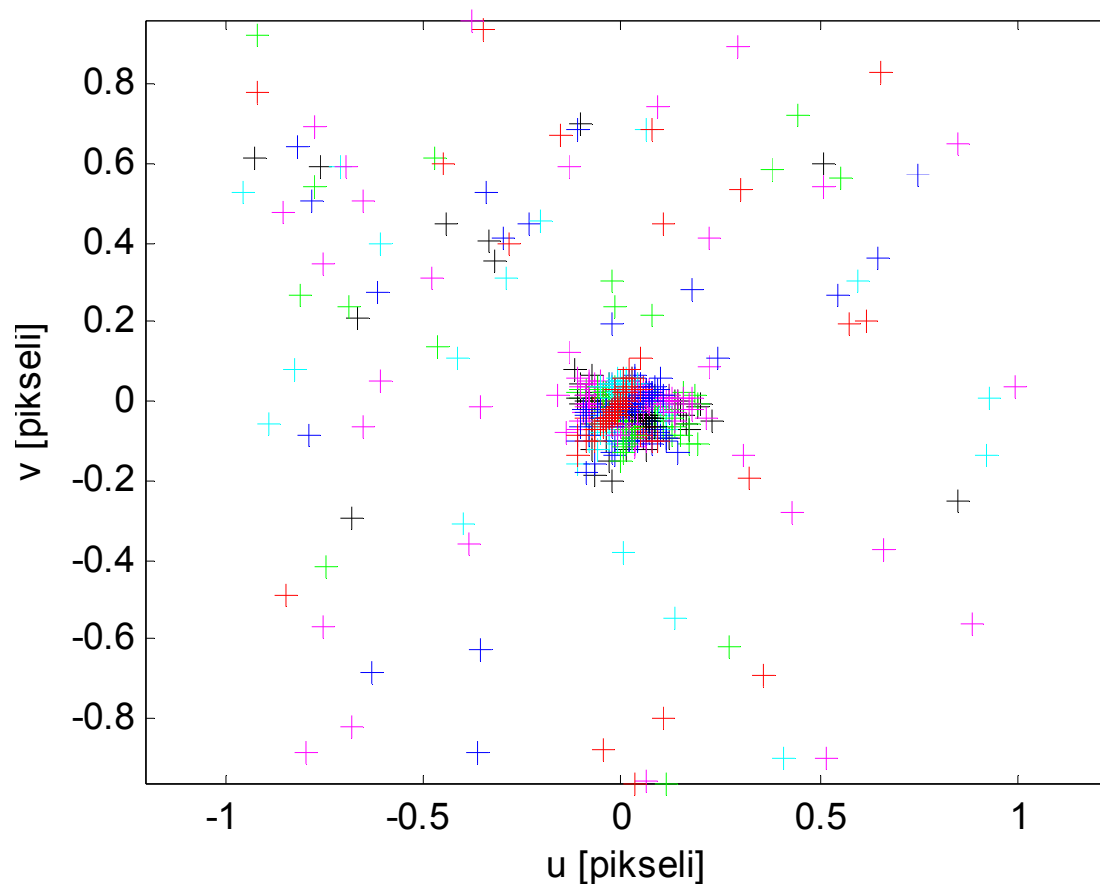
# Kalibracija kamere

- Radijalna distorzija leće kamere.



# Kalibracija kamere

- Pogreška projekcije izražena brojem piksela.
- Na slici su prikazane pogreške za svih 20 kalibracijskih uzoraka, odnosno različitih pogleda na šahovsku ploču.



## 5.5. Uzorkovanje slike

- **Digitalna slika predstavlja kolekciju uzoraka vizualnog sadržaja okoline.**
- U principu, scena se sastoji od polja svjetlosnih zraka pridruženih svim mogućim smjerovima pogleda iz svih mogućih položaja.
- Ovakav opis scene naziva se ***lumigraf***.
- Iako slika prikazuje kontinuiranu površinu, ona se sastoji od **diskretnih tačaka uzorkovanih vremenski i prostorno**.
- **Diskretizacija uvodi fizička ograničenja na kvalitetu prikaza kontinuiranih površina, pri čemu se signali visokih frekvencija gube iz slike, a izobličuju se signali nižih frekvencija.**
- Ovaj fenomen se zove ***aliasing***.





## Uzorkovanje slike

- Minimalna frekvencija uzorkovanja koja se zahtijeva za kompletni prikaz kontinuiranog signala zove se ***Nyquistova frekvencija***.
- **Za prikaz kontinuiranog signala potrebno je da se signal uzorkuje sa frekvencijom koja je najmanje dvaput veća od najveće frekvencije u signalu (Shanonov teorem).**
- S tim u vezi neka se promatra jednodimenzionalni signal  $f(x)=\cos(\pi x)$ .
- Ako se signal  $f(x)$  uzorkuje u cjelobrojnim tačkama (to jest u 0, 1, 2, 3, itd.) tada se postižu vrijednosti 1, -1, 1, -1,.....



## Uzorkovanje slike

- Ako se  $f(x)$  uzorkuje u neparnim cjelobrojnim tačkama tada se dobivaju vrijednosti 1, 1, 1, 1, ....
- Izraženo pomoću ugaone frekvencije, frekvencija uzorkovanja mora biti  $\omega \geq \pm \pi$ , što znači da kontinuirani signal frekvencije iznad  $\pi$  i ispod  $-\pi$  ne može biti prikazan.
- Amplitudni spektar  $|\hat{f}(k)|$  kontinuiranog signala  $f(x)$ , za  $k, x \in \mathbb{R}^n$ , predstavlja se **Fourierovom transformacijom**:

$$\hat{f}(k) = \int_{-\infty}^{+\infty} f(x) e^{-ixk} dx$$





## Uzorkovanje slike

- Filtri ili operatori transformacije slike sa velikim iznosima amplitudnog spektra za  $k$  blizu  $\pm\pi$  (Nyquistova frekvencija) **pojačavaju visoke frekvencije u slici, odnosno izobličuju sliku.**
- Operacije nad slikom najčešće se mogu karakterizirati konvolucijskom sumom ( $y = x * h$ ):

$$y(n_1, n_2) = \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} x(k_1, k_2) h(n_1 - k_1, n_2 - k_2)$$

- Tako, naprimjer, izgladivanje slike  $x(k_1, k_2)$  može se provesti operatorom:

$$h(k_1, k_2) = \frac{1}{8} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (*)$$

## Uzorkovanje slike

- $(k_1, k_2) = (0, 0)$  predstavlja centralni element od  $h$ .
- Dvodimenzionalna Fourierova transformacija  $\hat{x}(\omega_1, \omega_2)$  ulaznog signala  $x(n_1, n_2)$  povezana je sa dvodimenzionalnom Fourierovom transformacijom  $\hat{y}(\omega_1, \omega_2)$  izlaznog signala  $y(n_1, n_2)$  relacijom:

$$\hat{y}(\omega_1, \omega_2) = \hat{h}(\omega_1, \omega_2) \hat{x}(\omega_1, \omega_2)$$

- Frekvencijska karakteristika se može koristiti za analizu odziva operatora (filtra) za proizvoljni ulazni signal.



# Uzorkovanje slike

- Za dvodimenzionalni operator  $h(x_1, x_2)$  predstavljen diskretnom konvolucijom,  $\hat{h}(k_1, k_2)$  je definiran izrazom:

$$\hat{h}(k_1, k_2) = \sum_{\alpha_1=-\infty}^{\alpha_1=+\infty} \sum_{\alpha_2=-\infty}^{\alpha_2=+\infty} h(\alpha_1, \alpha_2) e^{-i(\alpha_1 k_1 + \alpha_2 k_2)}$$

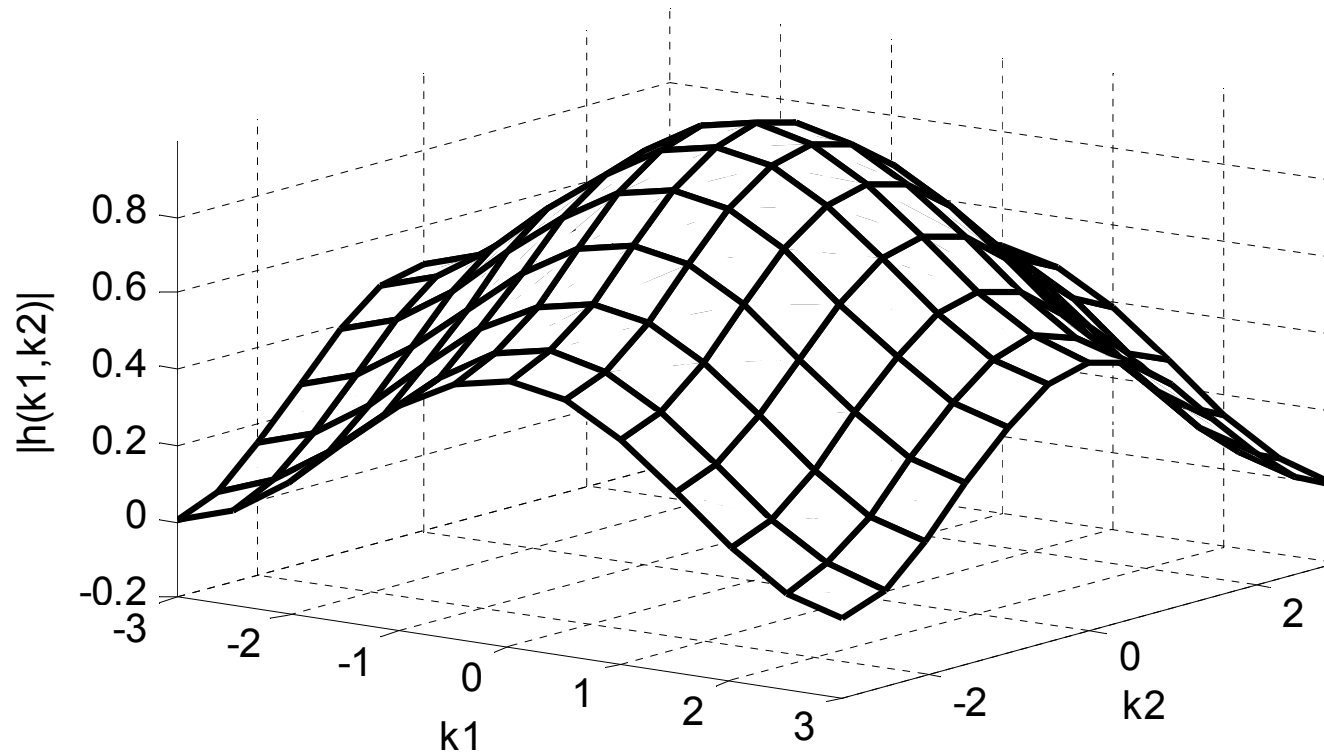
gdje su  $k_1$  i  $k_2$  kontinuirane vrijednosti prostorne frekvencije. Za operator iz izraza (\*) slijedi:

$$\begin{aligned} \hat{h}(k_1, k_2) &= \sum_{\alpha_1=-\infty}^{\alpha_1=+\infty} \sum_{\alpha_2=-\infty}^{\alpha_2=+\infty} h(\alpha_1, \alpha_2) e^{-i(\alpha_1 k_1 + \alpha_2 k_2)} \\ &= \frac{1}{2} + e^{ik_1} + e^{-ik_1} + e^{ik_2} + e^{-ik_2} \\ &= (4 + 2 \cos(k_1) + 2 \cos(k_2)) / 8 \end{aligned}$$



# Uzorkovanje slike

- Amplituda od  $\hat{h}$  prikazana je na slici.



- Sa slike se uočava veliko pojačanje za frekvencije blizu Nyquistove.
- Ovo znači da operator (\*) neće dobro operirati u prisustvu šuma.



## 5.6. Značajke slike

- Osnova mnogih vizualnih zadataka u robotskoj viziji je usporedba dvaju ili više pogleda istog objekta dobivenih sa dvije/više kamera ili dobivenih iz jedne kamere sa različitim pozicija, pri čemu je jedan od njih referentni pogled.
- Često je korisno izabrati takve apstrakcije slika koje pojednostavljaju usporedbu.
- Pojedinačni pikseli su samo indirektno povezani sa intenzitetom (boja) objekta u sceni.
- Iznosi piksela su rezultat složene interakcije između osvjetljenosti prostora, položaja kamere, prisustva drugih objekata u prostoru i refleksijskih svojstava objekata.



## Značajke slike

- Vrijednosti pojedinačnih piksela ne koriste se direktno u većini zadataka robotske vizije.
- Zbog toga je potrebno načiniti njihovo preobradbu da se istaknu korisne strukture u slici. U nastavku se obrađuju sljedeće značajke slike:
  - ***boja i sjena,***
  - ***ograničenja osvjetljenosti slike (optički tok),***
  - ***korelacija,***
  - ***izdavajanje značajki,***
  - ***izdvajanje rubova.***





## Boja i sjena

- Kamere u boji postižu intenzitet od skupa senzora osjetljivih na različite valne duljine svjetlosti (obično RGB).
- Nakon toga se slika predstavlja vektorom vrijednosti u svakom pikselu, radije nego samim intenzitetom.
- Jedna od najefikasnijih metoda za opis objekta u sceni je **histogram boja** kojim se objekt opisuje određenom kombinacijom boja.
- Slika se predstavlja vektorom  $(h_1, h_2, \dots, h_n)$  u  $n$ -dimenzionalnom vektorskom prostoru, gdje svaki element  $h_i$  označava broj piksela  $i$ -te boje u slici.



## Boja i sjena

- Sličnost slike  $H$  opisane histogramom  $(h_1, h_2, \dots, h_n)$  i slike  $K$  opisane sa  $(k_1, k_2, \dots, k_n)$  može se izračunati prema izrazu:

$$d(H, K) = \sum |h_i - k_i|$$

- U standardnim primjenama skup mogućih boja se reducira na manji broj (256 je tipičan broj boja) i tada se slika može predstaviti sa 256 vrijednosti.
- **Histogrami se normiraju tako da je mjera sličnosti neovisna o veličini slike.**



## Optički tok

- Iako vrijednosti pojedinačnih piksela nisu stabilne uz velike promjene u slici, mogu se smatrati stabilnim pri malim promjenama.
- Drugim riječima, intenzitet tačke se ne mijenja njezinim kretanjem.
- Ovo je iskazano **jednadžbom ograničenja osvjetljenosti**, odnosno jednadžbom ograničenja gradijenta:

$$f(x + dx, y + dy, t + dt) = f(x, y, t) \quad (**)$$

gdje je  $f(x, y, t)$  **funkcija intenziteta** slike.



# Optički tok

- Razvojem funkcije  $f(x, y, t)$  u Taylorov red dobiva se:

$$f(x + dx, y + dy, t + dt) = f(x, y, t) + \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial t} dt + O(x^2)$$

- Zanemarenjem članova visokog reda ( $O(x^2)$  – drugog i višeg reda po  $x$ -u) i uzimajući u obzir izraz (\*\*\*) slijedi:

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} \quad (***)$$

- Ova jednažba se naziva **jednažbom ograničenja optičkog toka**.



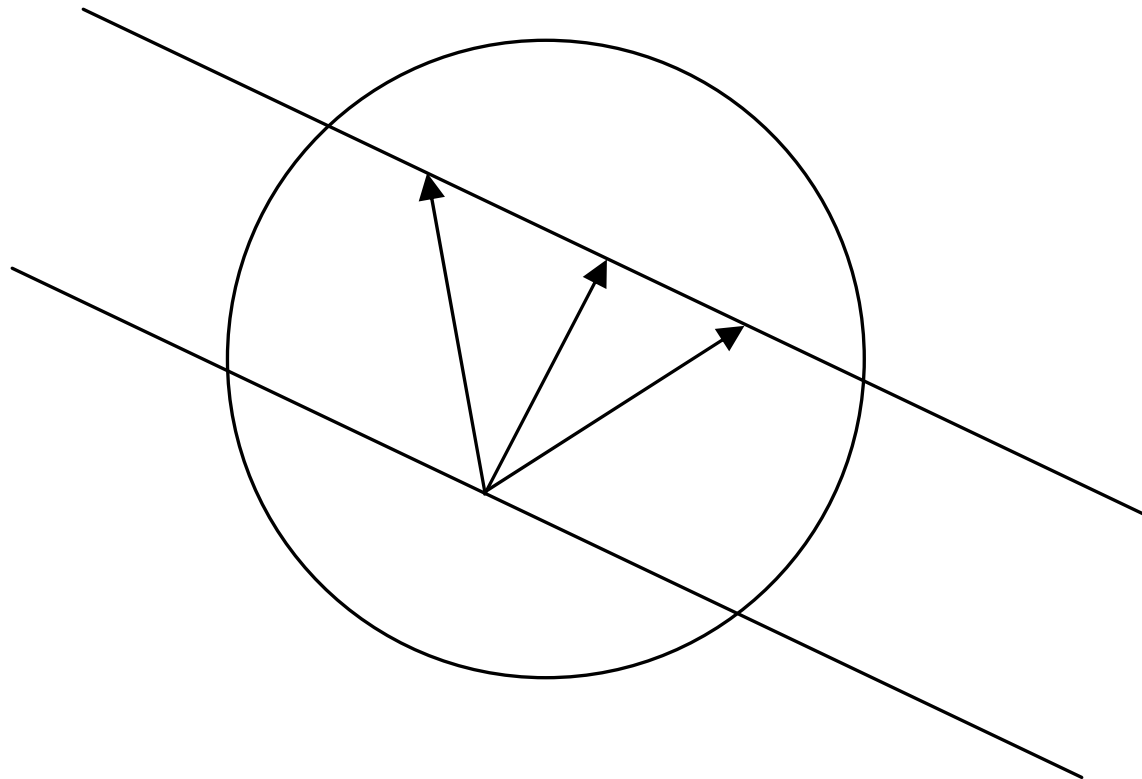
# Optički tok

- Jednadžba ograničenja optičkog toka postavlja ograničenje na brzinu slike  $(u, v) = (dx/dt, dy/dt)$  u ovisnosti o lokalnim derivacijama funkcije intenziteta slike  $(E_x, E_y, E_t) = (df/dx, df/dy, df/dt)$ .
- Jednadžba (\*\*\*) nije dovoljna za izračunavanje optičkog toka  $(u, v)$ , već samo linearnog ograničenja  $E_x u + E_y v + E_t = 0$ .
- Lokalne informacije sadržane u jednadžbi ograničenja optičkog toka daju linearna ograničenja na  $(u, v)$ , pri čemu optički tok predstavlja vektorsko polje  $(u, v)$ .



# Optički tok

- Ovo postavlja ograničenje na tok u smjeru gradijenta osvjetljenosti ali ne i na ograničenje u smjeru okomitom na smjer gradijenta osvjetljenosti (slika ispod).



## Optički tok

- Također je potrebno računati  $u$  i  $v$  za svaki piksel, što nije moguće navedenim postupkom jer jednažba ograničenja optičkog toka daje samo **jednu jednažbu po pikselu**.
- Da bi se dobilo jednoznačno polje optičkog toka neophodne su dodatne informacije.
- Nedostupnost lokalnih informacija optičkog toka u smjeru okomitom na smjer gradijenta osvjetljenosti poznato je pod imenom ***aperture problem*** (problem dijafragme kamere).





## Optički tok

- Rješavanje aperture problema se omogućuje uvođenjem dodatnih izraza koji izračunavaju **integral kvadrata amplitude gradijenta optičkog toka**:

$$e_s = \iint (u_x^2 + u_y^2 + v_x^2 + v_y^2) dx dy$$

- i **pogrešku u jednadžbi ograničenja optičkog toka**:

$$e_c = \iint (E_x u + E_y v + E_t)^2 dx dy$$



## Optički tok

- Oba izraza trebaju poprimiti što je moguće manje vrijednosti kada se želi minimizirati  $e_s + \lambda e_c$ , odnosno izbjeći aparture problem, gdje je  $\lambda$  parametar koji "otežava" pogrešku u jednadžbi kretanja scene (slike).
- Velika vrijednost ovog parametra se koristi ako je mjerenje osvjetljenosti tačno, a mala ako su mjerenja prekrivena šumom (zašumljena).
- **U praksi se parametar  $\lambda$  ručno podešava kako bi se postigle što je moguće bolje performanse.**



# Optički tok

- Primjena Eulerove jednačbe daje:

$$\nabla^2 u = \lambda(E_x u + E_y v + E_t) E_x$$

$$\nabla^2 v = \lambda(E_x u + E_y v + E_t) E_y$$

gdje je:

$$\nabla^2 = \frac{\partial^2}{dx^2} + \frac{\partial^2}{dy^2}$$

Laplaceov operator.



# Optički tok

- Gornje jednađbe formiraju par eliptičkih diferencijalnih jednađbi drugog reda.
- U slučaju pojave silueta (jedan objekat zatvara drugi) javlja se diskontinuitet u optičkom toku.
- Ovo naravno narušava ograničenje glatkoće.
- Jedna od mogućnosti prevazilaženja ovog problema je pokušati pronaći rubove koji nagovještavaju ovakva preklapanje objekata, a da se pri tome isključe pikseli u blizini takvih rubova na temelju računanja optičkog toka.
- Optički tok obećava da bude sastavni dio budućih algoritama vizije koji će kombinirati znakove korištenjem višestrukih algoritama.
- Međutim, algoritmi izbjegavanje prepreka i upravljački navigacijski sistemi koji isključivo koriste optički tok nisu još uvijek u širokoj upotrebi.



# Korelacija

- Iako se intenzitet strukture scene mijenja sa promjenom pogleda, može postojati slučaj da lokalna struktura slike blizu promatrane, interesantne tačke ostane konstantna.
- Jedan od najpopularnijih pristupa koji umjesto promatranja jednog piksela, kao kod optičkog toka, promatra **skup piksela** (prozor u slici) i uzorak iz jedne slike se korelira sa drugom slikom da se pronađe njegov položaj u drugoj slici.
- Ovaj pristup se naziva **korelacija**.
- Sličnost dviju slika  $f$  i  $g$  definirana je izrazom:

$$d^2(\delta x, \delta y) = \sum \sum (f(x, y) - g(x + \delta x, y + \delta y))^2$$





## Korelacija

- Gornja suma je načinjena preko područja dviju slika koje se uspoređuju.
- Mjestu gdje se slika  $g$  može naći u slici  $f$  pridružuje se mala vrijednost  $d^2$ , odnosno slike  $f$  i  $g$  su slične ako je  $d^2$  malo.
- Obavljanjem operacije kvadriranja dobiva se:

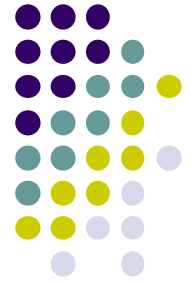
$$d^2(\delta x, \delta y) = \sum \sum f(x, y)^2 + g(x + \delta x, y + \delta y)^2 - 2f(x, y)g(x + \delta x, y + \delta y)$$

- Ako se pretpostavi da su  $f$  i  $g$  konstantni unutar cijelog prozora, tada je minimizacija izraza  $d^2(\delta x, \delta y)$  ekvivalentna maksimizaciji korelacije  $f$  i  $g$ -a, gdje je korelacija dana sa:

$$\text{corr}(\delta x, \delta y) = \sum \sum f(x, y)g(\delta x + x, \delta y + y)$$

# Korelacija

- U tipičnim primjenama je mali prozor u jednoj slici u korelaciji sa dijelovima prozora druge slike.
- Korelacija je relativno jeftina.
- Cijena je povezana sa veličinom prozora i područja koji se pretražuje.
- Također, jednostavna korelacijska funkcija dana gornjim izrazom je osjetljiva na lokalne osjetljivosti slika  $f$  i  $g$ .
- Ova funkcija je također osjetljiva na šum u ulazu i distorziju između dviju slika koje nisu rezultat jednostavnog pomjeranja (shiftovanje).



# Korelacija

- Normiranjem se dobiva sljedeća korelacijska funkcija:

$$\frac{\sum \sum W(x, y) f'(x + \delta x, y + \delta y) g'(x + \delta x, y + \delta y)}{|v_f(\delta x, \delta y) v_g(\delta x, \delta y)|^{\frac{1}{2}}}$$

- gdje su  $f'$  i  $g'$  korigirane verzije ulaza,  $v_f$  i  $v_g$  su lokalne varijance od  $f$  i  $g$  izračunate unutar funkcije prozora  $W$ .



## Izdvajanje značajki slike

- Umjesto razmatranja cjelokupne slike neki pristupi koriste heuristike za identificiranje pojedinih tačaka na slici koje su važne za određeni zadatak i kasnije korištenje tih tačaka za prikazivanje slike.
- Korneri, naprimjer, često su veoma važni za vizualne zadatke i za njihovo identificiranje se koriste ranije izgrađeni detektori kornera.
- Opća klasa takvih operatora je poznata pod imenom **detektori značajki** ili operatori od interesa.
- U praksi ovi detektori trebaju identificirati lokacije koje su stabilne unutar malih promjena u slici i malih promjena u tački pogleda.
- Oni također sažimaju sliku u obliku brojnih podesivih značajki.







## Izdvajanje značajki slike

- Jedan od najpopularnijih i najviše korištenih detektora značajki je ***Moravecov operator***.
- Ovaj operator generira potencijalne tačke od interesa na temelju mjerenja karakterističnih parametara lokalne strukture slike.
- Moravecov operator definira mjeru varijance u pikselu  $(x, y)$  kao:

$$\text{var}(x, y) = \left\{ \sum_{k, l \in S} [f(x, y) - f(x + k, y + l)]^2 \right\}^{\frac{1}{2}}$$

gdje je:

$$S = \{(0, a), (0, -a), (a, 0), (-a, 0)\}$$

$a$  je parametar.

## Izdvajanje značajki slike

- Varijanca se računa u okolini piksela  $(x, y)$  i minimalna varijanca se računa na sljedeći način:

$$\text{Moravec}(x, y) = \min_{(\delta x, \delta y) \in N \times N} \text{var}(x + \delta x, y + \delta y)$$

pri čemu se uzimaju samo minimalne lokalne vrijednosti operatora koje premašuju specificirani prag.

- Moravecov operator je dizajniran s ciljem identifikacije tačaka u slici koje imaju visoku varijancu, kao što su, naprimjer, korneri.



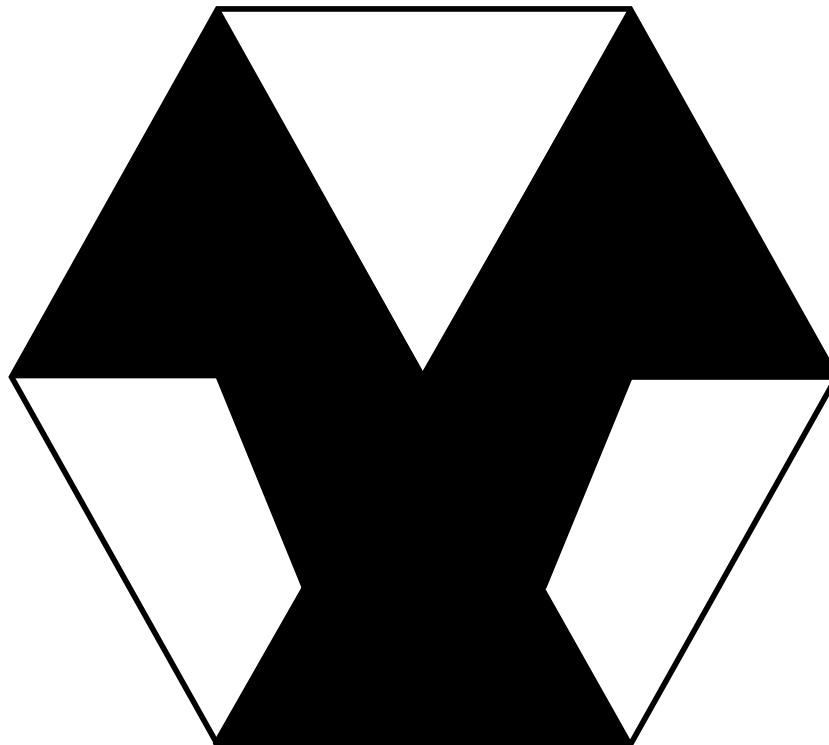
## Izdvajanje značajki slike

- Iako su tačke sa visokom vrijednošću varijance od interesa veoma je teško reći koji detektor će biti odgovoran.
- Nagel (Nagel, 1982) je razvio detektor koji veoma dobro reagira na uzorke intenziteta koji se pojavljuju u blizini kornera u slici.
- Nažalost, distribucija intenziteta slike identificiranih Nagelovim detektorom kornera samo indirektno odgovaraju kornerima na slici.



## Izdvajanje značajki slike

- Mnogo efikasniji je takozvani ***binarni akvizicijski predložak*** koji predstavlja posebno izrađen predložak poznate veličine koji se stavi u prostor i kome se lahko odrede pozicija i orijentacija.

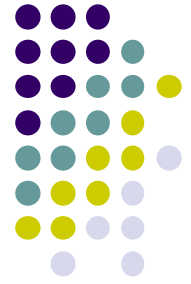


Binarni akvizicijski predložak



## Izdvajanje rubova

- Mnoge slike mogu se opisati u obliku strukture promjena u intenzitetu slike. **Rubovi** predstavljaju konture u slici na kojima se skokovito mijenja intenzitet.
- Budući da je slika opisana kontinuiranom funkcijom  $f(u, v)$  slijedi da se rubovi pojavljuju na mjestima gdje su **velike vrijednosti gradijenata funkcije  $f(u, v)$** .
- Gradijent funkcije  $f(u, v)$  je vektor  $\nabla f(u, v) = [\partial f / \partial u, \partial f / \partial v]$ .
- Trenutno postoji veliki broj diskretnih aproksimacija gradijenta.



## Izdvajanje rubova

- **Kod gradijentnih operatora moraju se izračunavati lokalni maksimumi i minimumi u slici**, što može biti vrlo teško ostvariti.
- Ovaj problem se ublažava primjenom gradijentnih operatora drugog reda, kod kojih se traže tačke prolaska druge derivacije funkcije  $f(u, v)$  kroz nulu.
- Ako su  $\Delta_1$  i  $\Delta_2$  ortogonalne usmjerene derivacije intenziteta slike, tada je amplituda gradijenta intenziteta slike dana sa  $\sqrt{\Delta_1^2 + \Delta_2^2}$  dok je smjer ovog gradijenta dan sa  $\tan^{-1} (\Delta_2/\Delta_1)$ .



# Izdvajanje rubova

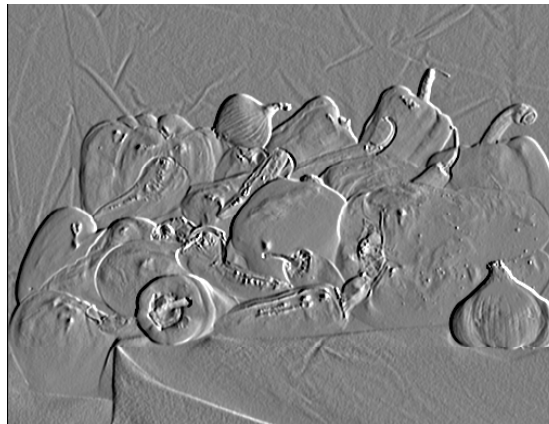
- Naprimjer, **Sobelov operator** se implementira pomoću korelacije slike sa operatorima  $\Delta_1$  i  $\Delta_2$ , gdje su oni dani sa:

$$\Delta_1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad \Delta_2 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

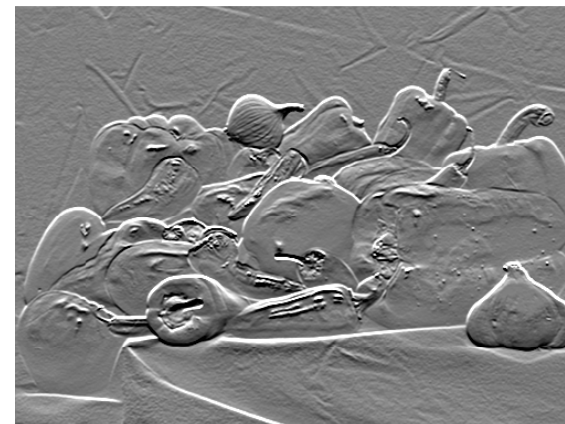
- Efekt djelovanja Sobelovog operatora prikazan je na slikama.



original ( $I$ )



$\Delta_1 I$



$\Delta_2 I$



## Izdvajanje rubova

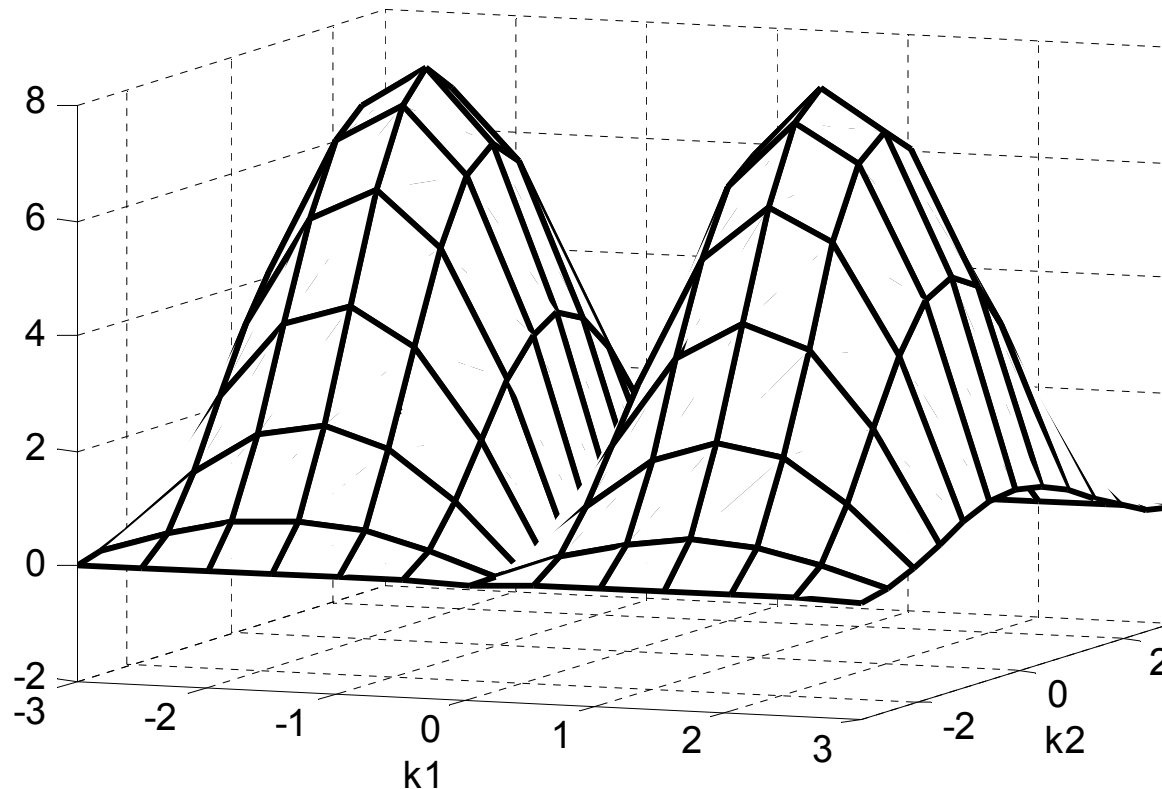
- **Nažalost, operatori koji koriste derivacije imaju tendenciju pojačavanja šuma u slici.**
- Zbog toga je potrebno obaviti filtriranje da bi se smanjio šum, pa tek onda primijeniti Sobelov operator.
- U vezi s tim je važno promatrati odziv svakog od Sobelovih filtera. Tako za  $\Delta_1$  prijenosna funkcija postaje:

$$\begin{aligned}\hat{\Delta}_1(k_1, k_2) &= \sum_{\alpha_1=-\infty}^{\alpha_1=+\infty} \sum_{\alpha_2=-\infty}^{\alpha_2=+\infty} h(\alpha_1, \alpha_2) e^{-i(\alpha_1 k_1 + \alpha_2 k_2)} \\ &= -2i(\sin(k_1 + k_2) + \sin(k_1 - k_2) + 2\sin(k_1))\end{aligned}$$



# Izdvajanje rubova

- Modul prijenosne funkcije  $\hat{\Delta}_1$  prikazan je na slici.



Iako  $\Delta_1$  ima veliku amplitudu na visokim frekvencijama može se očekivati njegovo slabo djelovanje u prisustvu šuma. Korištenje inicijalnog filtriranja s ciljem uklanjanja visokih frekvencija čini operatore slične Sobelovim mnogo stabilnijim.



## Izdvajanje rubova

- Drugi glavni problem sa gradijentnim operatorima je da se mora **pretraživati njihov izlaz kako bi se pronašli i identificirali lokalni maksimum i minimum (poznati kao šiljci) na slici.**
- Šiljci su predstavljeni značajkama koje je veoma teško pronaći u slici.
- **Upotreba gradijentnih operatora drugog reda reducira ovaj problem jer se pretraga šiljaka zamjenjuje sa pronalaženjem tačaka prolaska kroz nulu.**
- Postoji mnogo vrsta gradijentnih operatora drugog reda koji se ovdje mogu razmatrati, pri čemu je možda **Laplaceov operator** najpopularnijim.





## Izdvajanje rubova

- Ovaj operator je oblika:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

- Laplaceov operator je neusmjereni operator drugog reda.
- **Operatori drugog reda također pojačavaju visoke frekvencije, što ima za posljedicu njihovu osjetljivost na šum i potrebno je također prefiltriranje slike niskopropusnim filtrom, naprimjer, Gausovim filtrom oblika:**

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

## Izdvajanje rubova

- Laplaceov operator Gaussova filtra iznosi:

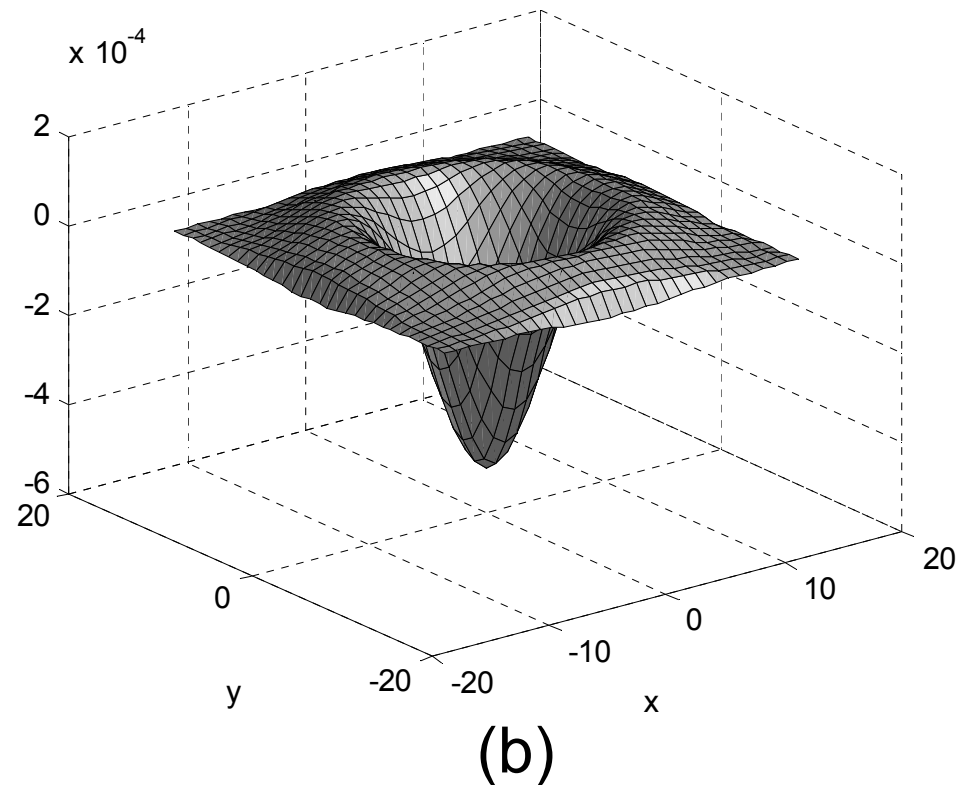
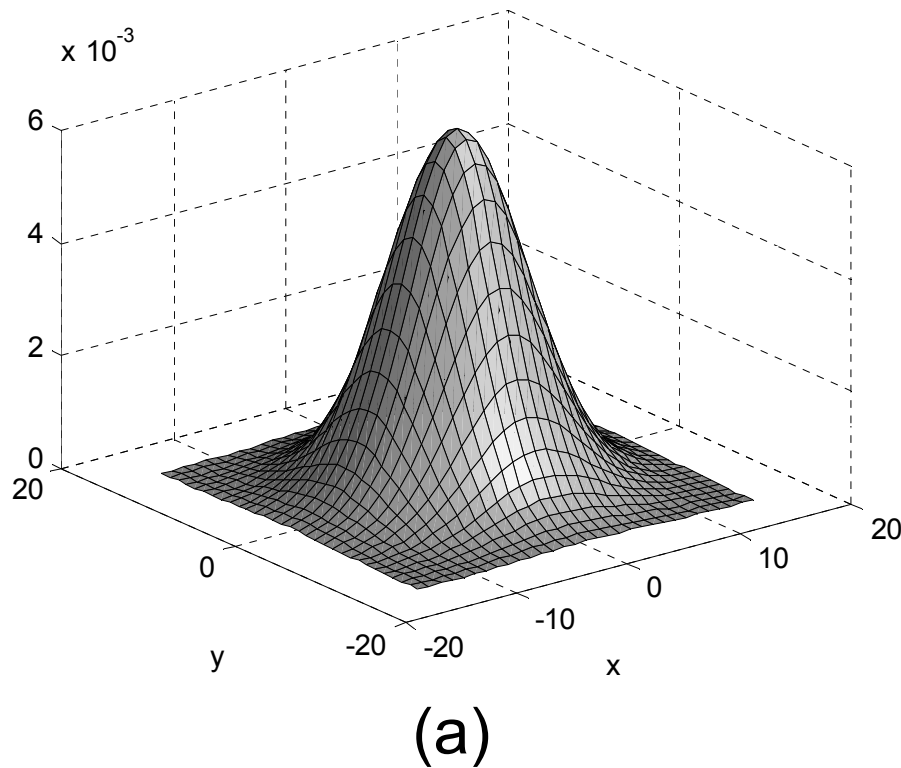
$$\nabla^2 (G * I) = (\nabla^2 G) * I$$

- **Ovaj se operator koristi za računanje prolaska kroz nulu druge derivacije radije nego za računanje šiljaka u prvoj derivaciji.**
- Operator  $\nabla^2 G$  je izotropni operator što znači da nije selektivan prema orijentaciji rubova.
- Odzivi Gaussova filtra i njegovog Laplaceovog operatora prikazani su na sljedećem slajdu.



# Izdvajanje rubova

- Odzivi Gausova filtra (a) i njegovog Laplaceovog operatora (b)

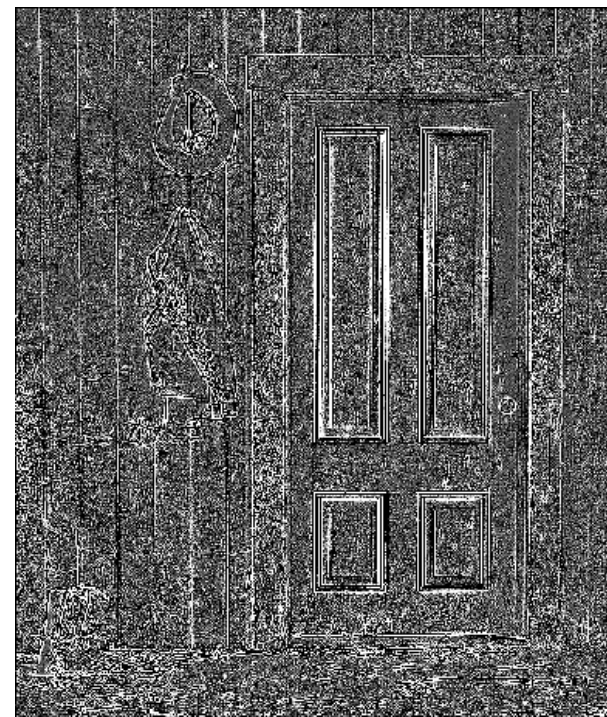


# Izdvajanje rubova

- Rezultati dobiveni Laplaceovim operatorom Gaussova filtra primijenjenim za istu sliku uz različite vrijednosti parametra  $\sigma$  prikazani su na sljedećim slikama.



original



$\sigma = 1$



# Izdvajanje rubova



$\sigma = 0.5$



$\sigma = 2$



## Izdvajanje rubova

- Sa slika je jasno da se **promjenom ovog parametra može prilagođavati rezolucija prikaza prostora.**
- Izdvojeni individualni rubni pikseli obično se povezuju u složenije strukture prije daljnje obrade.
- Naprimjer, skup tačaka koje predstavljaju linijski segment može se u njega i pretvoriti.
- Na taj se način smanjuje količina podataka kojom je prikazana slika.
- Ovaj postupak se zove ***vektorizacija slike***.





# Izdvajanje rubova

- **Cannyjev operator**
- Ova metoda koristi višerazinski algoritam za detekciju rubova u slici.
- Za razliku od ostalih metoda za detekciju i izdvajanje rubova, Cannyjeva metoda koristi dvije različite vrijednosti pragova za detektiranje jakih i slabih rubova, te uključuje slabe rubove samo u slučaju kada su povezani sa jakim rubovima.
- Na ovaj način omogućeno je kvalitetnije izdvajanje rubova iz slike u prisustvu šuma.
- Cannyjeva metoda predstavlja "optimalnu" metodu za detekciju, odnosno izdvajanje rubova.
- Optimalna funkcija Cannyjevog detektora opisana je zbrojem eksponencijalnih članova.

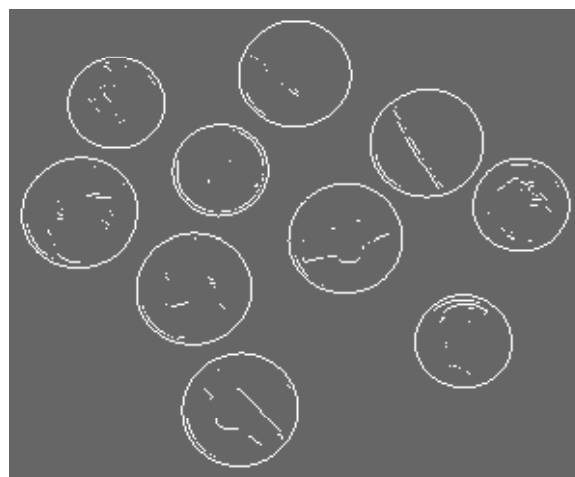


# Izdvajanje rubova

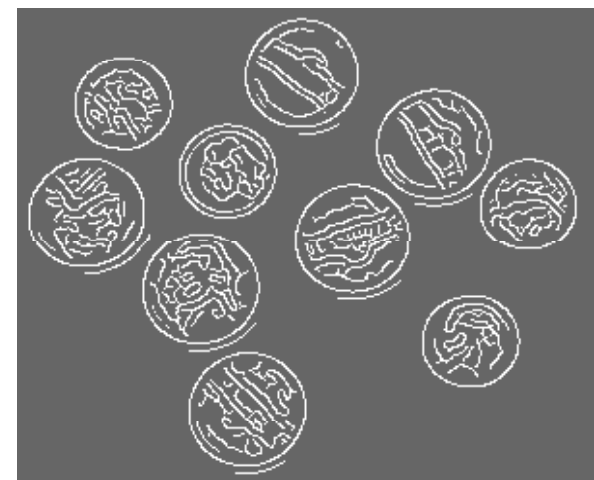
- Cannyjev algoritam može se razmatrati kao filter koji ulaznu crno-bijelu sliku obrađuje tako da na izlaznoj slici postoje samo rubovi ili ne-rubovi.
- Detekcija rubova odvija se u četiri faze: *izgladivanje (usrednjavanje) slike, određivanje gradijenta intenziteta slike, stanjivanje rubova i uporedba s pragom.*



originalna slika



Sobel



Canny

## 5.7. Segmentacija slike

- **Segmentacija slike bavi se dekompozicijom scene u njezine sastavne dijelove.**
- Osnovni cilj njenog korištenja je odvajanje piksela koji predstavljaju objekte od interesa od ostalih piksela u sceni.
- Proces segmentacije slike obuhvaća tri potprocesa: ***klasifikacija piksela, reprezentacija i opis.***
- Na temelju ovih procesa razvijene su različite metode segmentacije.
- U nastavku poglavlja se razmatraju sljedeći postupci segmentacije: ***binarna, segmentacija s pragom (amplitudna segmentacija), grupiranje, obilježavanje komponenti i metode koje koriste granice objekta.***



# Segmentacija slike

## Binarna klasifikacija s pragom

- Klasifikacija piksela predstavlja proces zaključivanja koji se primjenjuje na svaki piksel i pridružuje ga jednoj od  $C$  klasa,  $c \in \{0, \dots, C-1\}$ , gdje se klasa piksela  $c$  predstavlja cjelobrojnom vrijednošću.
- Kada je  $C = 2$  tada se govori o **binarnoj klasifikaciji**. U tom slučaju, pikseli se mogu klasificirati kao objekti ( $c = 1$ ) ili neobjekti ( $c = 0$ ) predstavljeni bijelom ili crnom bojom.
- Binarna klasifikacija:

$$c(u, v) = \begin{cases} 0 & f(u, v) < t \\ 1 & f(u, v) \geq t \end{cases} \quad \forall (u, v) \in f$$

gdje  $f$  predstavlja sliku i  $t$  prag



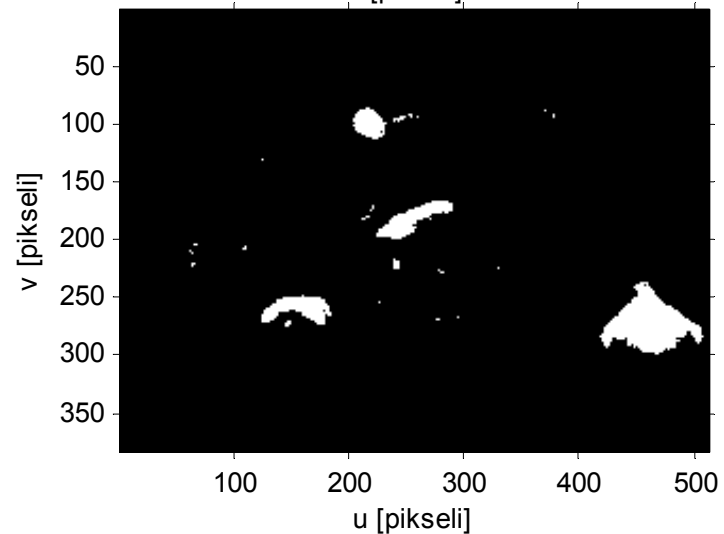
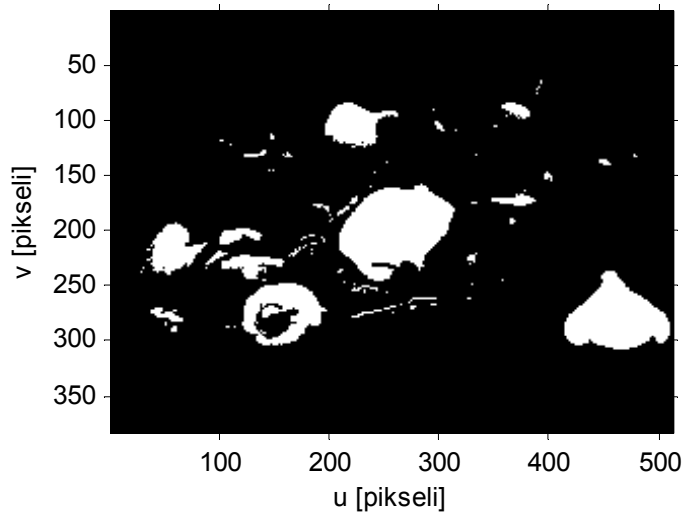
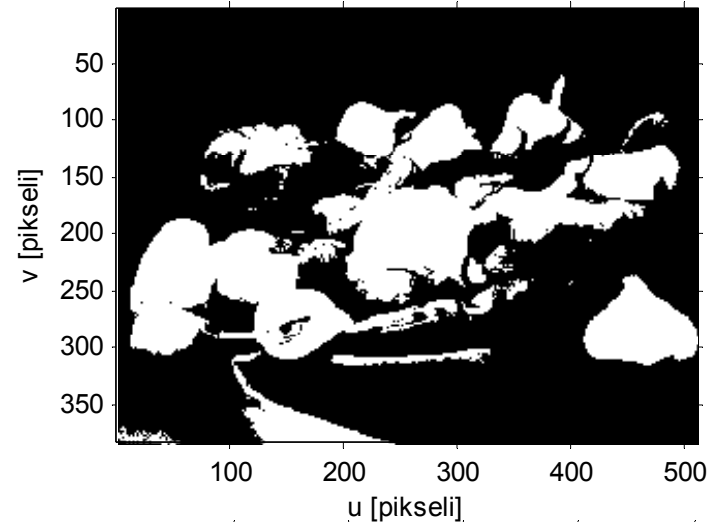
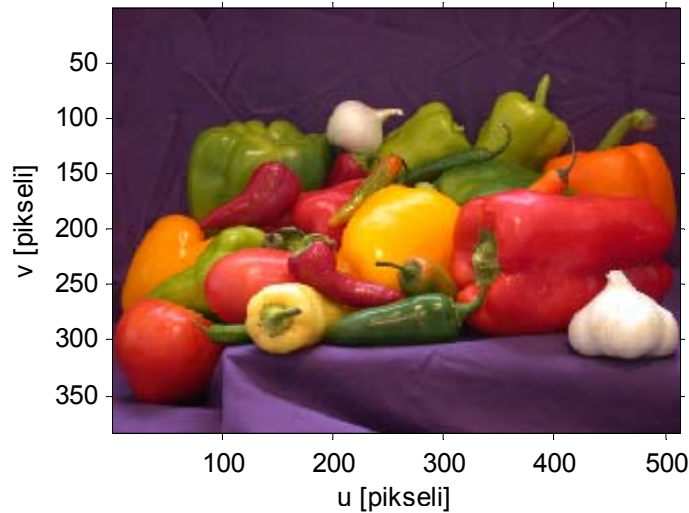
# Segmentacija slike

- Prag  $t$  može se odrediti korištenjem histograma prvog reda, odnosno računanjem minimuma u histogramu.
- Određuje se iznos praga tako da dio piksela ima svjetlinu nižu od praga ili odrediti prag na temelju histograma izračunatog za piksele u slici koji zadovoljavaju neki kriterij (velik iznos gradijenata).
- Kod mnogih slika vrijednost piksela pada u jednu od dvije skupine, tamne ili svijetle tačke.
- U takvim se slučajevima objekt može izdvojiti od pozadine na temelju vrijednosti amplitude.
- Kod izbora praga mogu nastati problemi usljed djelovanja šuma i smetnje, kada objekt i pozadina slike imaju širok raspon vrijednosti i u slučaju neuniformne pozadine.



# Segmentacija slike

- Binarno klasificirane slike sa različitim vrijednostima praga: 0.4, 0.6 i 0.8.



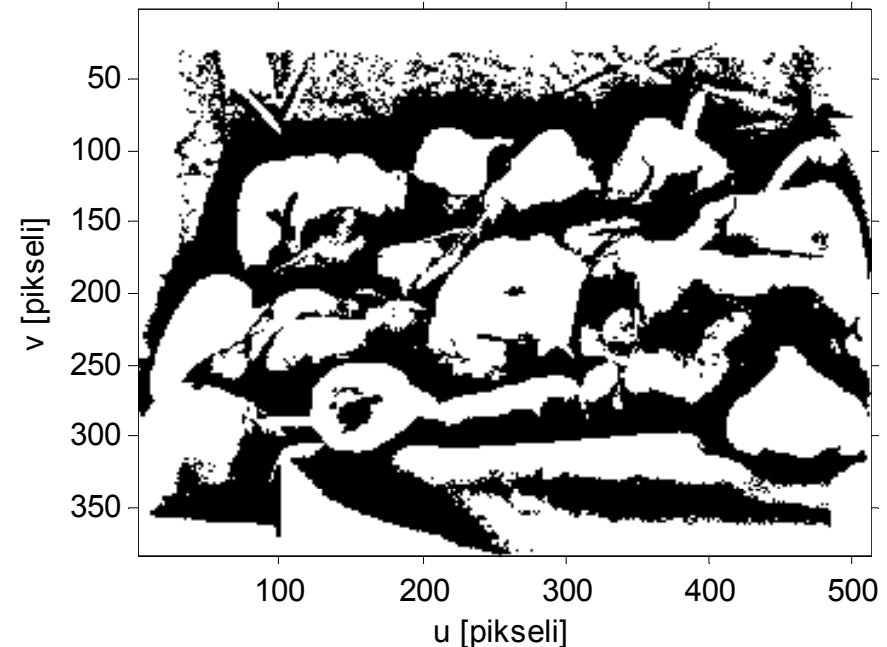
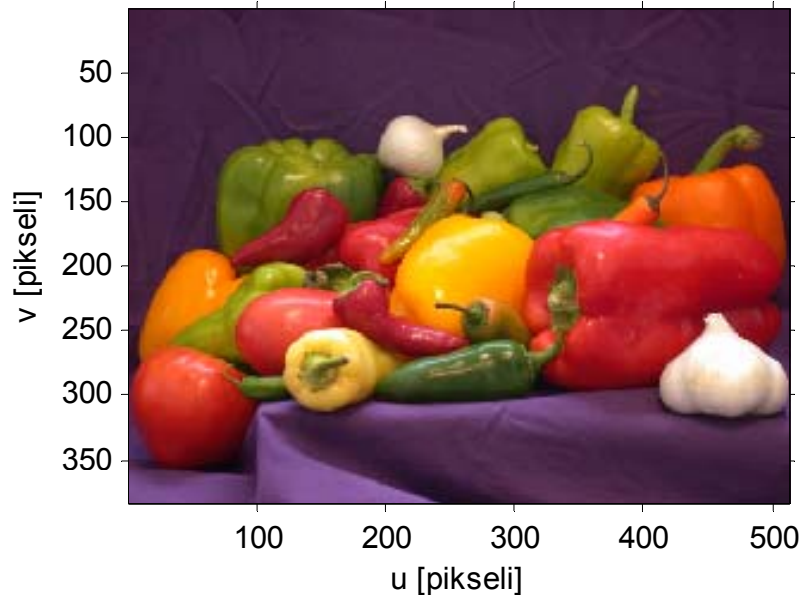
# Segmentacija slike

## Binarna klasifikacija sa Niblackovim algoritmom

- Niblackov algoritam koristi lokalni prag koji se računa na sljedeći način:

$$t(u, v) = \mu(W) + k\sigma(W)$$

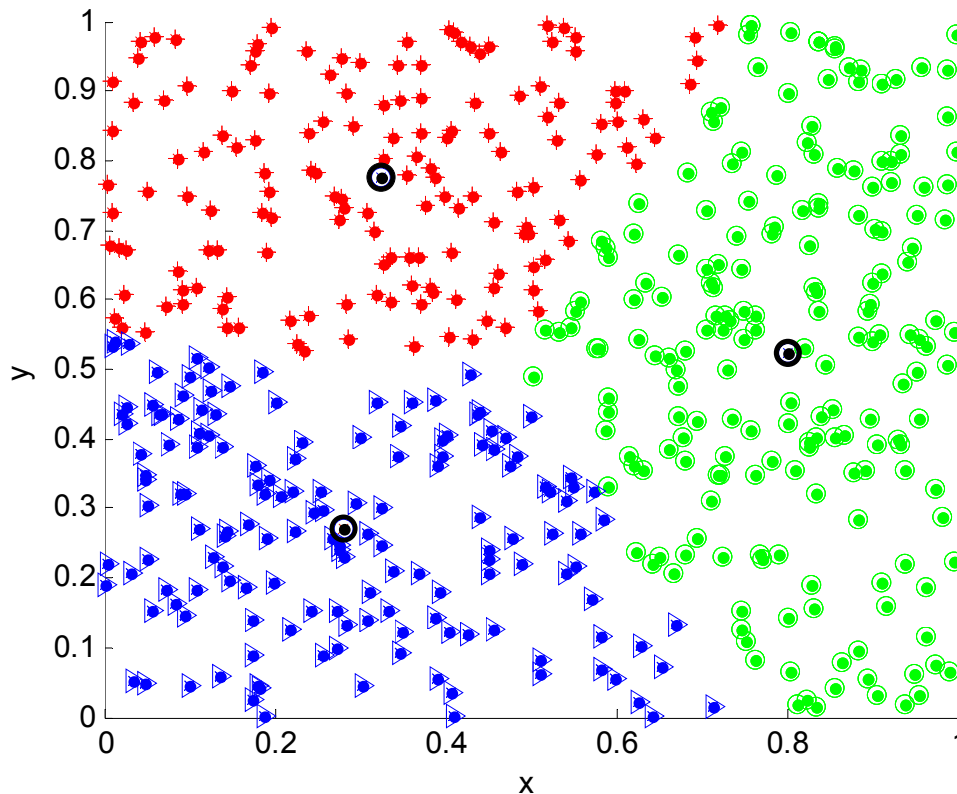
gdje je  $W$  veličina regija oko tačke  $(u, v)$ ,  $\mu(\cdot)$  srednja vrijednost i  $\sigma(\cdot)$  standardna devijacija.



# Segmentacija slike

## Grupiranje s K srednjih vrijednosti (engl. K-means clustering)

- Pronalazi grupe (engl. clusters) tačaka u ravnini, gdje svaka grupa odgovara skupini piksela sa istaknutom razlikovnom bojom.



**Tri grupe tačaka sa centrima, gdje su centri prikazani većim krugovima.**





# Segmentacija slike

- Ovdje se polazi od pretpostavke da pojedinačna scena sadrži četiri različita elementa u boji. Pikseli se grupiraju u četiri kromatske klase ( $C = 4$ ).
- Algoritam grupiranja s  $K$  srednjih vrijednosti je iterativni algoritam za grupiranje  $N$ -dimenzionalnih tačaka u  $K$  prostornih grupa.
- Svaka grupa definirana je centrom  $c_k$  koji predstavlja  $N$ -dimenzionalni vektor.
- Neka je  $x$   $N$ -dimenzionalni vektor značajki slike u tački  $(i, j)$ .
- Segmentacija treba grupirati vektore tako da značajke unutar jedne grupe budu uniformne.
- Nadalje, neka je poznat broj grupa  $K$  i neka je  $c_k(n)$  centar  $k$ -te grupe u  $n$ -toj iteraciji. Inicijalno se  $c_k(0)$  postave na bilo koju vrijednost.



# Segmentacija slike

- U  $n$ -toj iteraciji odabere se jedan vektor  $\mathbf{x}_i$  i dodijeli se grupi tačaka  $R_k$  čijem je centru najbliži:

$$\mathbf{x}_i \in R_k \Leftrightarrow d(\mathbf{x}_i, \mathbf{c}_k(n)) = \min_{j=1, \dots, K} \{d(\mathbf{x}_i, \mathbf{c}_j(n))\}$$

- Zatim se ponovo izračunaju centri grupa kao vektori koji minimiziraju udaljenosti za vektore iz pojedine grupe:

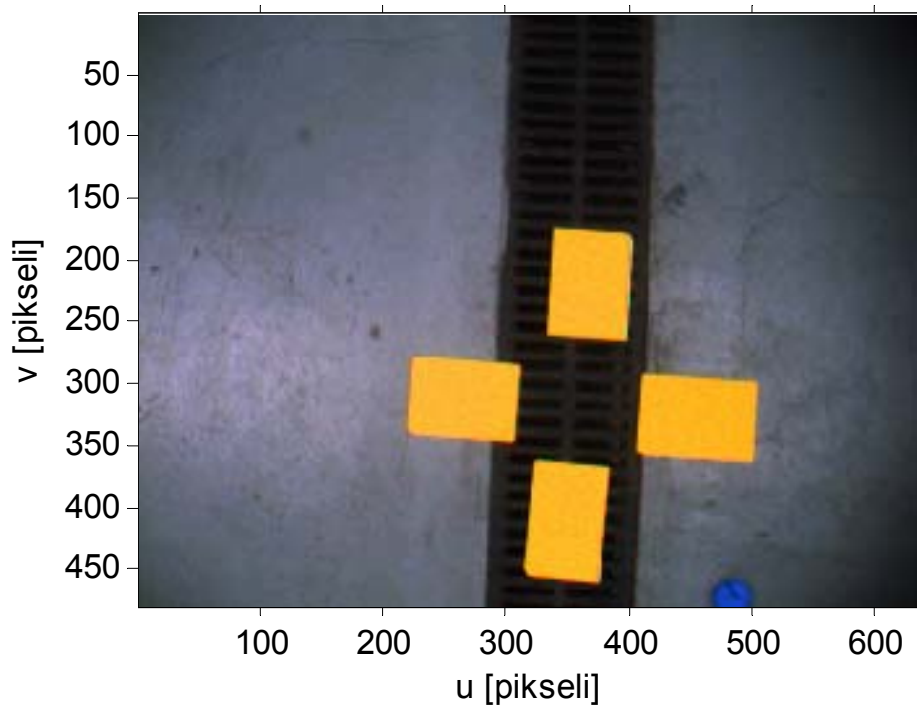
$$\mathbf{c}_k(n+1) : \sum_{\mathbf{x}_i \in R_k} d(\mathbf{x}_i, \mathbf{c}_k(n+1)) = \min_y \{d(\mathbf{x}_i, \mathbf{y})\}, \quad k = 1, \dots, K$$

- Postupak se ponavlja sve dok se položaj centara više ne mijenja, odnosno dok se sve tačke ne pridruže pojedinoj grupi.

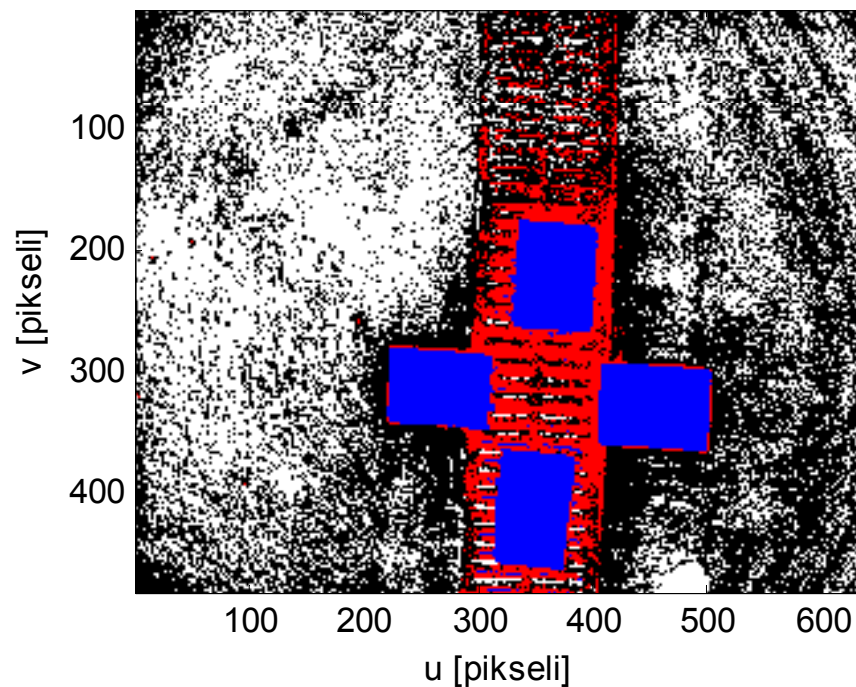


# Segmentacija slike

- **Primjer 2.** Segmentacija zasnovana na grupiranju s  $K$  srednjih vrijednosti

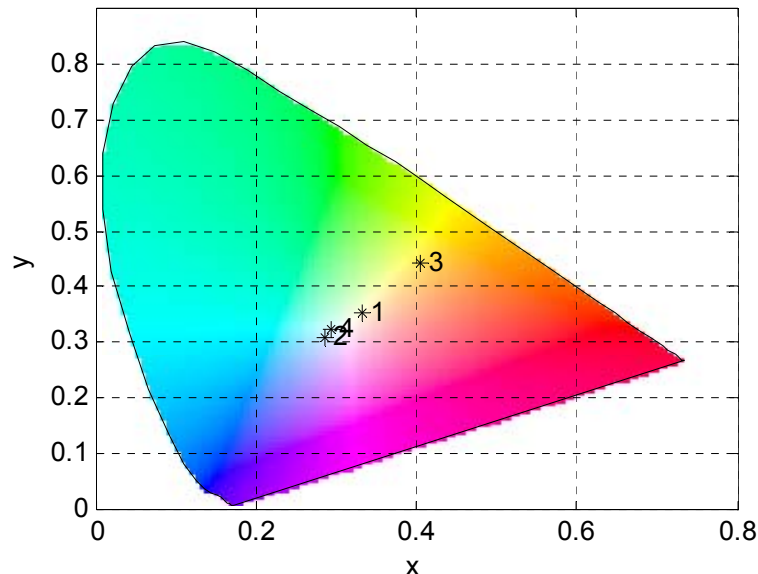


originalna slika

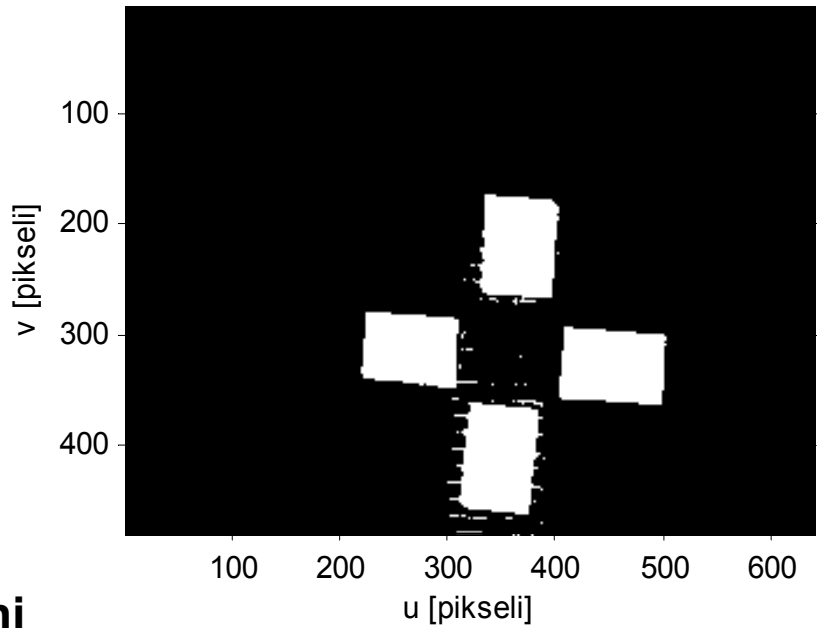


slika dobivena  
klasifikacijom piksela za  $C = 4$

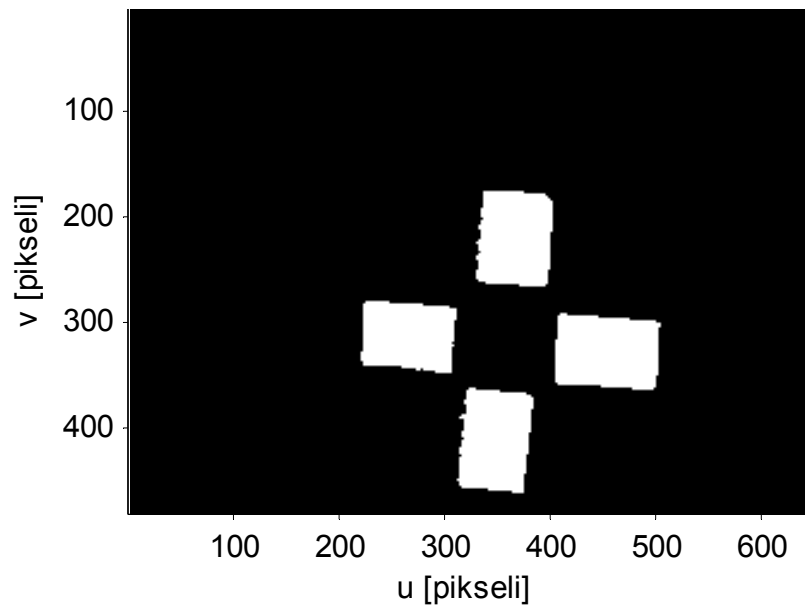
# Segmentacija slike



centri klasa u xy kromatskoj ravnini



pikseli klase 3.



slika nakon primjene morfološke operacije zatvaranja s kružnim strukturnim elementom polumjera  $r = 2$

# Segmentacija slike

## Segmentacija zasnovana na obilježavanju komponenti

- Jednostavna i efikasna metoda koja se zasniva na ispitivanju povezanosti piksela sa njegovim susjedima i obilježavanja povezanih skupova tačaka.
- **Algoritam:** Prvo se binarnom slikom prolazi liniju po liniju, slijeva na desno i odozgo prema dolje.
- Trenutni piksel  $X$  može biti objekt (vrijednost 1) ili pozadina (vrijednost 0).
- Ako je vrijednost piksela 1 tada se on obilježi ispitivanjem susjeda.
- Ako je neki od susjeda već obilježen da pripada nekom objektu onda i  $X$  dobiva istu oznaku.



# Segmentacija slike

- U slučaju da neki od susjeda imaju različite oznake tada se pripadni objekti stapaju (ekvivalentni objekti).
- Nova oznaka se stvara kod svakog prijelaza iz vrijednosti na izolirani piksel 1 (novi objekt).
- Rezultat pretraživanja je segmentacija slike u regije, to jest skupine objekata slike.



# Segmentacija slike

## Segmentacija zasnovana na granicama objekta

- Vrši segmentaciju objekata u slici na temelju granica objekata.
- Granice objekta mogu se izdvojiti sljedećim metodama: ***praćenje granice objekta***, ***interpolacija krivulje*** i ***Houghova transformacija***.
- **Metoda praćenja granice** (konture) daje kao izlaz niz koordinata tačaka koje se nalaze na granici objekta.
- Dobiveni niz određuje regiju slike koja pripada objektu.
- **Metoda interpolacije krivulje** povezuje tačke rasporeda rubova da bi se dobila zatvorena kontura koja definira regiju.
- Za to je potrebno razbiti konture u dijelove koji se interpoliraju i primijeniti neki od postupaka interpolacije.



# Segmentacija slike

- **Houghova transformacija**
- Houghova transformacija koristi se za detekciju linija u slici.
- Pravu liniju, odnosno pravac, potrebno je opisati u obliku minimalnog broja parametara.
- Standardni zapis  $v = au + b$  je problematičan u slučaju vertikalnog pravca kada je  $a = \infty$ .
- Umjesto toga može se koristiti  $(\rho, \theta)$  parametrizacija:

$$u \sin \theta + v \cos \theta = \rho$$

odnosno:

$$v = -u \tan \theta + \frac{\rho}{\cos \theta}$$





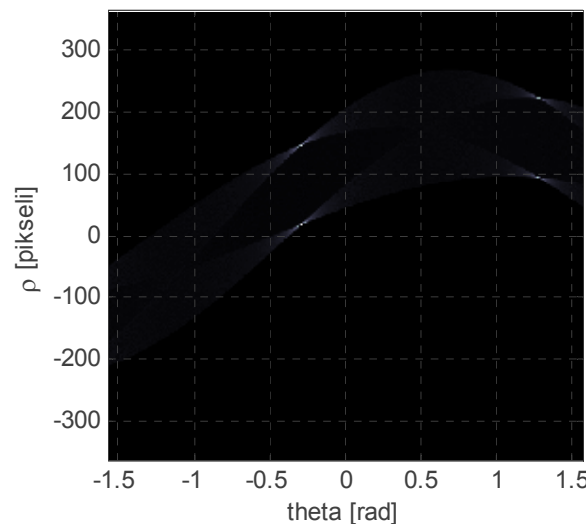
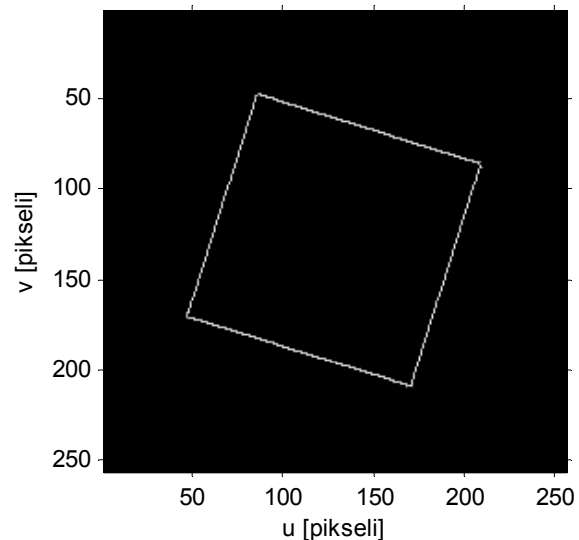
# Segmentacija slike

- $\theta \in [-\pi/2, \pi/2)$  je ugao nagiba i  $\rho$  udaljenost pravca od ishodišta, za koji vrijedi  $\rho \in [\rho_{\min}, \rho_{\max}]$ .
- Horizontalni pravac ima  $\theta = 0$  i vertikalni  $\theta = -\pi/2$ .
- Svaki pravac može se promatrati kao tačka  $(\rho, \theta)$  u dvodimenzionalnom prostoru svih mogućih linija.
- Houghova transformacija preslikava pravac u tačku u koordinatnom sistemu  $\rho\theta$ .
- Familija pravaca koja polazi kroz jednu tačku bi se preslikala u skup tačaka koje leže na sinusoidi.
- U praksi se razmatraju pravci iz konačnog skupa.
- U tom smislu se ravnina  $\theta\rho$  kvantizira i koristi se odgovarajuće polje  $A$  dimenzija  $N_\theta \times N_\rho$ .
- Ovo polje naziva se **Houghova ravnina**.

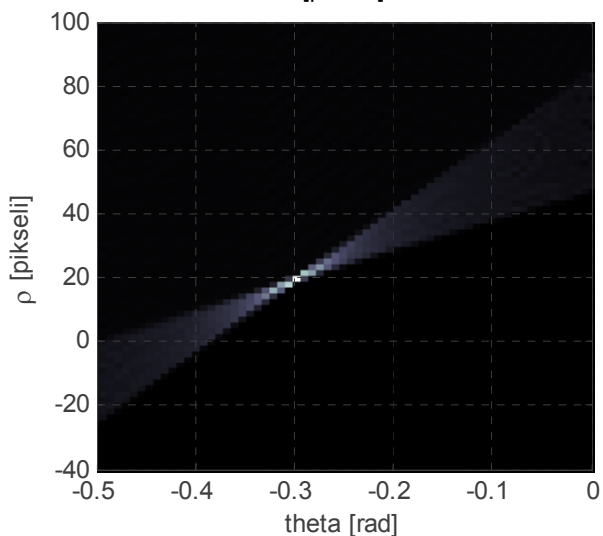


# Segmentacija slike

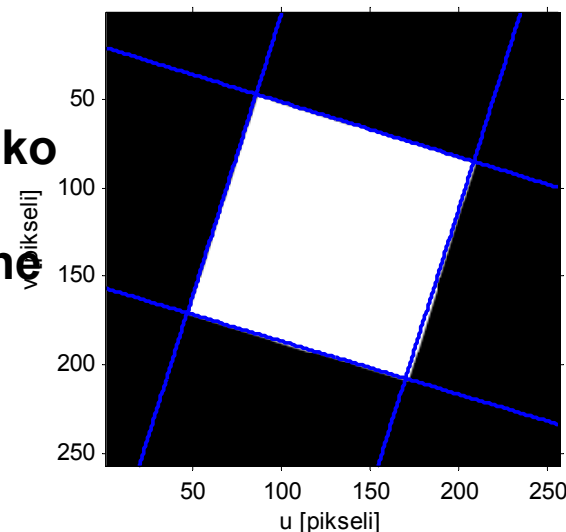
- **Primjer 3.** Promatra se pravokutni objekt zakrenut za određeni ugao u smjeru kretanja kazaljke na satu.



**Houghova ravnina**



**Prikaz oko jedne presječne tačke**



**Projekcija detektiranih linija (Houghovi pravci) na originalnu sliku**

# Segmentacija slike

## Segmentacija zasnovana na računanju momenta

- **Momenti** su efikasna i računarski nezahtjevna klasa svojstava slike koji mogu opisati veličinu regije, njenu lokaciju i oblik.
- Moment slike  $f$  je skalar opisan izrazom:

$$m_{pq} = \sum_{(u,v) \in f} u^p v^q f(u,v)$$

gdje je  $(p + q)$  red momenta.

- Za  $p = q = 0$  imamo moment nultog reda, odnosno:

$$m_{pq} = \sum_{(u,v) \in f} f(u,v)$$



# Segmentacija slike

- Za binarnu sliku u kojoj su vrijednosti piksela podloge jednaki nuli, moment predstavlja broj ne-nultih piksela (bijeli), odnosno veličinu regije izraženu brojem piksela.
- Neka je ukupna masa regije jednaka  $m_{00}$  tada je centar mase ili centroid regije:

$$u_c = \frac{m_{10}}{m_{00}}, \quad v_c = \frac{m_{01}}{m_{00}}$$

gdje su  $m_{10}$  i  $m_{01}$  momenti prvog reda.

- Centralni momenti  $\mu_{pq}$  računaju se u odnosu na centroid kao:

$$\mu_{pq} = \sum_{(u,v) \in f} (u - u_c)^p (v - v_c)^q f(u, v)$$



# Segmentacija slike

- Ovi momenti povezani su sa momentima  $m_{pq}$  na sljedeći način:

$$\mu_{10} = 0, \quad \mu_{01} = 0$$

$$\mu_{20} = m_{20} - \frac{m_{10}^2}{m_{00}}, \quad \mu_{02} = m_{02} - \frac{m_{01}^2}{m_{00}}, \quad \mu_{11} = m_{11} - \frac{m_{10}m_{01}}{m_{00}}$$

- Matrica inercije regije:

$$J = \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix}$$

sadrži momente oko osi koje su paralelne sa osima  $u$  i  $v$  i koje prolaze kroz centroid regije. Drugi centralni momenti  $\mu_{20}$  i  $\mu_{02}$  su momenti inercije i  $\mu_{11}$  je produkt inercije.



## Segmentacija slike

- Produkt inercije poprima ne-nultu vrijednost ako je oblik regije nesimetričan u odnosu na osi regije.
- Na ovaj način dobiva se **ekvivalentna elipsa** koja ima istu matricu inercije kao i regija.
- Svojstvene vrijednosti i svojstveni vektori matrice  $J$  povezani su sa radiusima elipse i njenom orijentacijom u odnosu na osi koordinatnog sistema.
- Radijusi elipse računaju sa kao:

$$a = 2\sqrt{\frac{\lambda_1}{m_{00}}}, \quad b = 2\sqrt{\frac{\lambda_2}{m_{00}}}$$

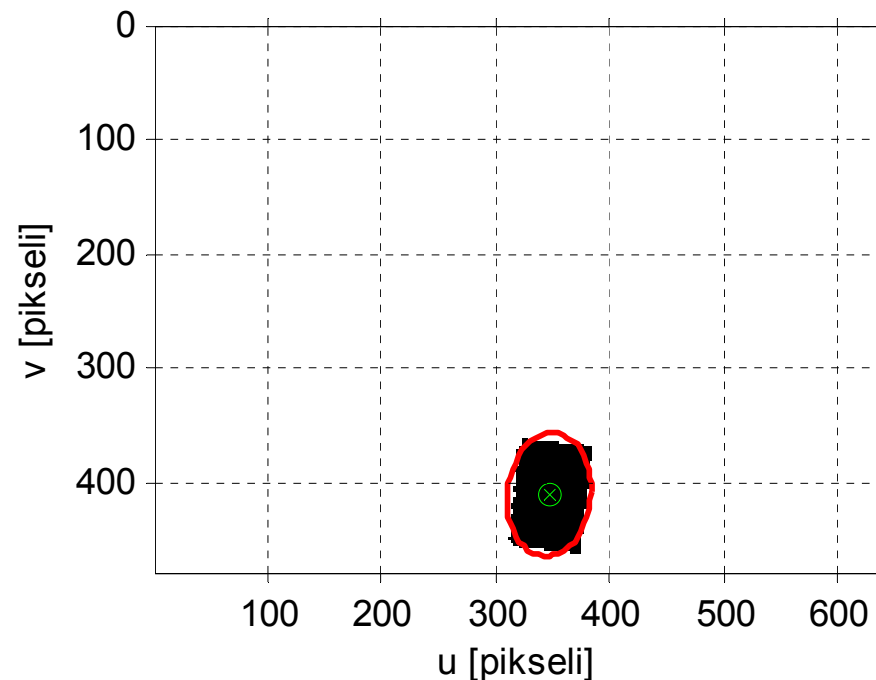
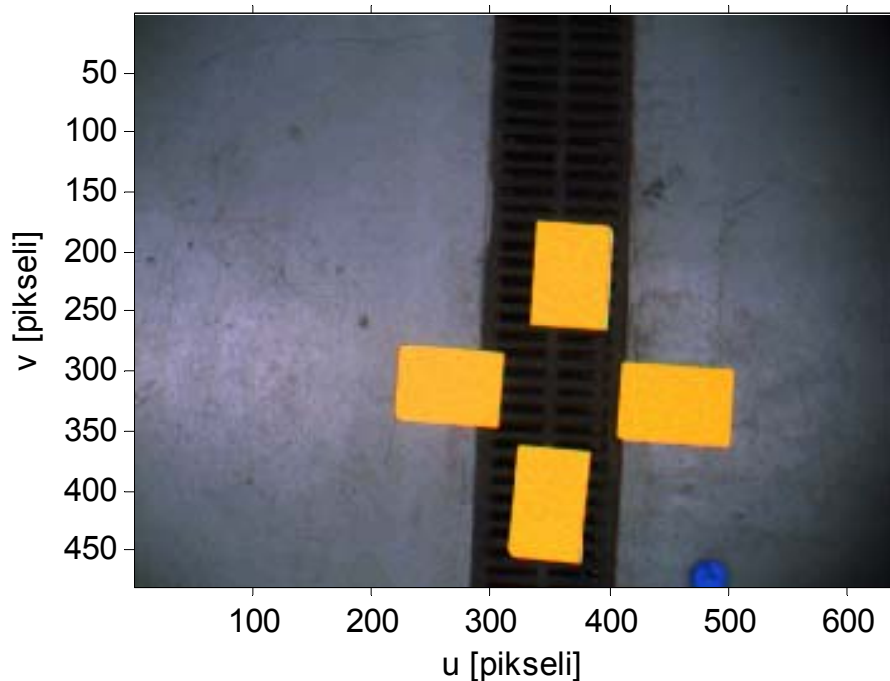
- Zadnji stupac matrice svojstvenih vektora ( $v_x$  i  $v_y$ ) određuje ugao koji ovaj vektor zatvara sa  $x$  osi:

$$\theta = \tan^{-1}\left(\frac{v_y}{v_x}\right)$$



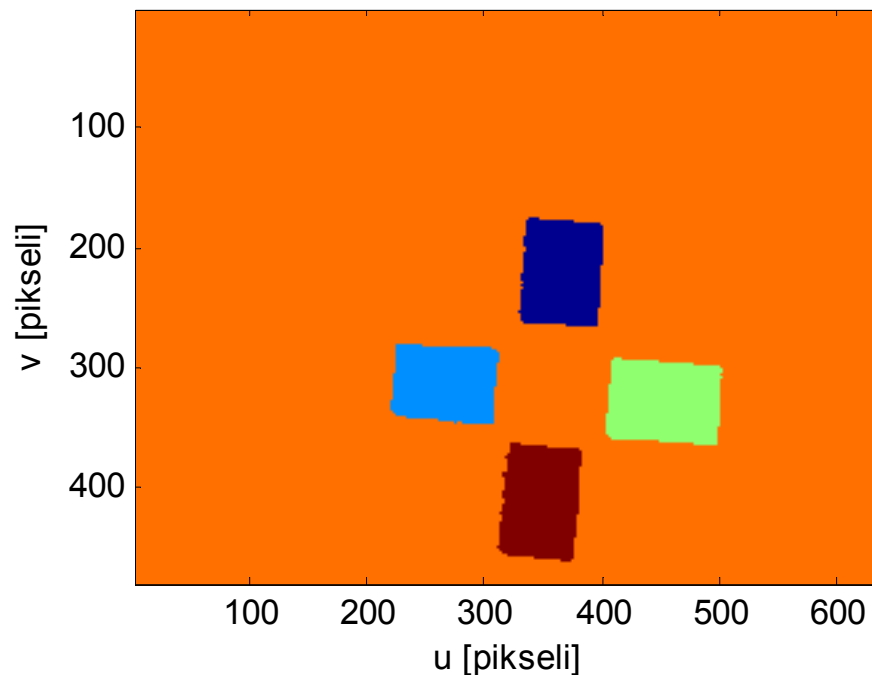
# Segmentacija slike

- **Primjer 4.** Segmentacija zasnovana na računanju momenta
- Ciljna regija označena je crnim pravokutnikom na drugoj slici. Za ovu regiju računaju se: momenti slike, centroid regije, matrica inercije regije (centralni momenti), radijusi elipse i ugao koji s horizontalnom osi zatvara zadnji vektor stupac matrice svojstvenih vektora.



# Segmentacija slike

- Regije sa prve slike izdvojene su postupkom segmentacije obilježavanja komponenti.



- Moment nultog reda iznosi  $m_{00} = 6004$ , što predstavlja veličinu regije izražene u pikselima.
- Koordinate centroida iznose  $u_c = 346.7539$  piksela i  $v_c = 410.4446$  piksela.





# Segmentacija slike



- Ciljna regija je ograničena sa  $[u_{min}, u_{max}] = [310, 382]$  piksela i  $[v_{min}, v_{max}] = [361, 460]$  piksela.
- Matrica inercije regije  $J$ , čiji su elementi centralni momenti regija, jednaka je:

$$J = \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix} = 10^6 \begin{bmatrix} 2.0436 & -0.2329 \\ -0.2329 & 4.4181 \end{bmatrix}$$

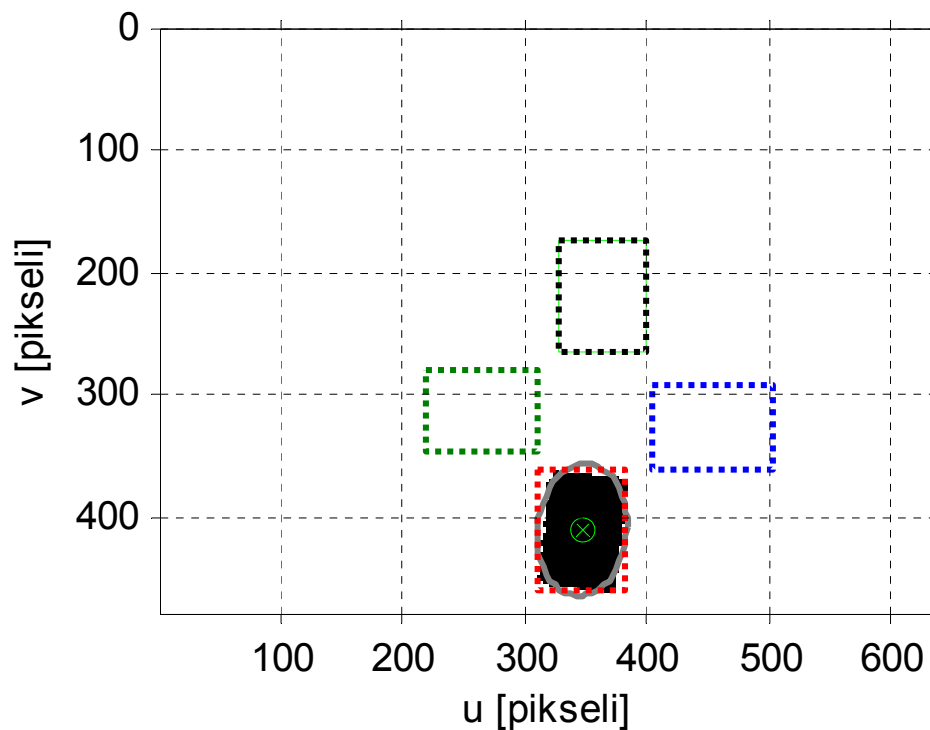
- Svojstvene vrijednosti matrice  $J$  i radijusi elipse:

$$\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = 10^6 \begin{bmatrix} 2.0210 \\ 4.4408 \end{bmatrix} \quad a = 2\sqrt{\frac{\lambda_1}{m_{00}}} = 36.6904, \quad b = 2\sqrt{\frac{\lambda_2}{m_{00}}} = 54.3880$$

- Omjer  $a/b$  iznosi 0.6746 i ugao koji zatvara vektor zadnjeg stupca matrice svojstvenih vektora  $v = [v_x \ v_y]^T = [-0.0962 \ 0.9954]^T$  sa horizontalnom osi iznosi  $95.5498^\circ$

# Segmentacija slike

- Omeđene kutije, odnosno pravokutnici, sadrže granične linije okvira unutar kojih su smještene četiri regije sa originalne slike (prva slika).
- Dimenzije ovih regija definirane su intervalima  $[u_{min}, u_{max}]$  i  $[v_{min}, v_{max}]$ .



## 5.8. Registracija slike

- **Problem registracije dviju slika sastoji se u određivanju geometrijske transformacije kojom se jedna slika preslikava u drugu.**
- Nakon registracije za svaku tačku u prvoj slici poznato je koja joj je tačka korespondentna u drugoj slici.
- Registracija je neophodna da bi se mogli upoređivati ili integrirati podaci dobiveni iz različitih mjernih sistema.
- Prema tome, registracija slike je važan problem analize slike kada imamo više slika jednog objekta dobivenih različitim načinima snimanja, kada imamo više slika jednog objekta snimljenih u različitim vremenskim trenucima, kada želimo uporediti slike više objekata i kada se želi uporediti slika s modelom.



# Registracija slike

- Proces registracije sastoji se od tri potprocesa: **transformacija slike**, **mjere sličnosti** kojom se određuje kvaliteta registracije i **funkcija optimiranja** kojom se maksimizira mjera sličnosti.
- Nakon što su dvije slike registrirane moguće je analizirati, odnosno detektirati razlike između dviju istovrsnih slika, naprimjer, snimljenih u dva različita vremenska trenutka i kombinirati informacije sadržane u dvjema raznovrsnim slikama, naprimjer, snimljenim različitim načinima snimanja.
- Što se tiče transformacije slike najčešće se koriste geometrijske transformacije, kao što su: **linearne** (uključuju translaciju, rotaciju, skaliranje i druge affine operacije), **elastične** i **rigidne** (neelastične).



# Registracija slike

- Geometrijska transformacija vrši se na prostoru u kojem se slika nalazi, a ne na intenzitetima slike, zbog toga se ona može matematički predočiti u općenitom obliku:

$$f_T(u, v) = f_S(T(u, v))$$

- Na ovaj način obavlja se registracija ulazne slike  $f_S$  na referentnu sliku  $f_R$  tako što se primjenjuje transformacija  $T$  na prostor u kojem se slika  $f_S$  nalazi.
- Sljedeći važan segment u registraciji slike je ***korespondencija slika***.
- Korespondencija više slika je problem pronalaženja koordinata piksela u više različitih slika koje odgovaraju istoj tački u stvarnom svijetu.
- Ove tačke mogu se detektirati različitim pogledima iste scene bez obzira na tačku pogleda ili svjetlosni ambijent.



# Registracija slike

- Pretpostavimo da imamo dvije slike iste scene dobivene dvjema različitim kamerama (veličina piksela, žarišna duljina, broj piksela po slici,...) postavljenim na različite položaje.
- Ukoliko scene nisu usklađene tada određivanje korespondencije između slika predstavlja značajan problem.
- U tom slučaju poseže se za nekim od detektora značajki, naprimjer, Harrisov detektor uglova koji odabire samo karakteristične, odnosno svojstvene tačke.
- **Harrisov detektor** računa se prema izrazu:

$$C_H(u, v) = \det(A) - k \cdot \text{tr}(A)$$



# Registracija slike



- Elementi matrice  $A$  računaju se na temelju gradijenata duž  $x$  i  $y$  osi slike ( $f_x$  i  $f_y$ ), bilo kvadriranjem ili množenjem, i izgladivanja pomoću težinske matrice  $W$ , na sljedeći način:

$$A = \begin{bmatrix} \sum W(i, j) f_x^2(u+i, v+j) & \sum W(i, j) f_x(u+i, v+j) f_y(u+i, v+j) \\ \sum W(i, j) f_x(u+i, v+j) f_y(u+i, v+j) & \sum W(i, j) f_y^2(u+i, v+j) \end{bmatrix}$$

- Matrica  $A$  može se napisati i u obliku:

$$A = \begin{bmatrix} \sum G(\sigma) \otimes f_x^2 & \sum G(\sigma) \otimes f_x f_y \\ \sum G(\sigma) \otimes f_x f_y & \sum G(\sigma) \otimes f_y^2 \end{bmatrix}$$

gdje je operacija konvolucije  $\otimes$  definirana kao:

$$G(\sigma_1) \otimes G(\sigma_2) = G\sqrt{\sigma_1^2 + \sigma_2^2}$$

$$G(\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{u^2+v^2}{2\sigma^2}}$$

# Registracija slike

- Na ovaj način reducira se šum u slici i poboljšava stabilnost i pouzdanost detektora. Prema tome, gradijenti slike  $f_x$  i  $f_y$  računaju se pomoću Gaussovog operatora sa mogućnošću podešavanja koeficijenta izgladivanja  $\sigma$ .
- U jednadžbi najčešće se uzima da je  $k = 0.04$ .
- Velike vrijednosti Harrisova detektora predstavljaju izoštrene i razlikovne uglove.
- Budući da je  $\det(A) = \lambda_1\lambda_2$  i  $\text{tr}(A) = \lambda_1 + \lambda_2$  slijedi da je ovaj detektor efikasan kada se izbjegava računanje svojstvenih vrijednosti matrice  $A$ , što ponekad može biti računarski izuzetno zahtjevno.



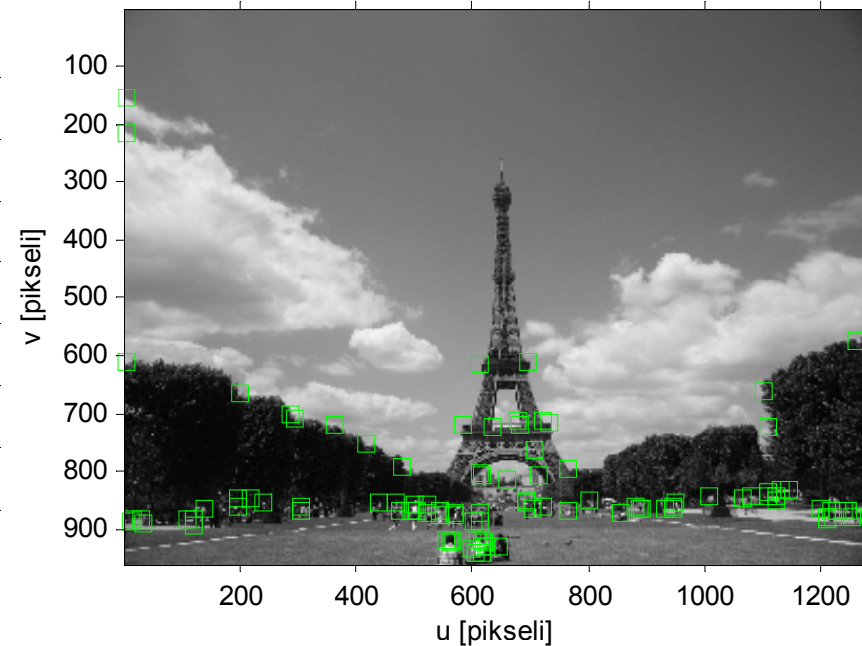
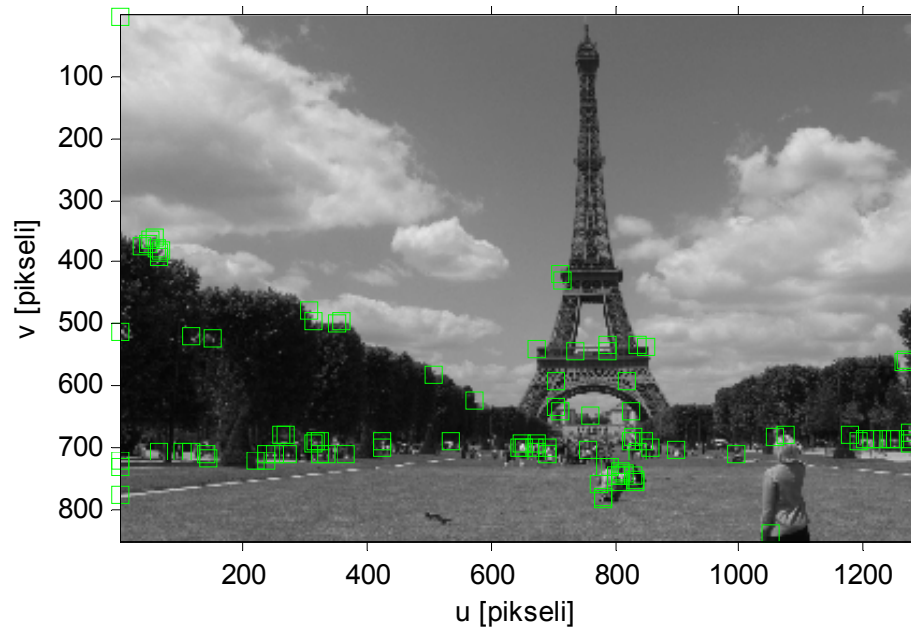


# Registracija slike

- **Primjer 5. Korespondencija slika.**
- Promatraju se dvije identične slike, dobivene dvjema kamerama različitog broja piksela (7 i 10 Mega piksela) i stupnja uvećanja slike



121/139



**Na slikama je primijenjen Harrisov operator kojim se detektira 100 najizraženijih uglova (kornera).**

# Registracija slike

- Za mjeru sličnosti dviju slika koje se registriraju zasnovanu na vrijednostima piksela mogu se koristiti različiti kriteriji, kao što su: kvadratna razlika intenziteta, korelacijske tehnike, uniformnost kvocijenta slike, kriteriji zasnovani na teoriji informacija (naprimjer, zajednička entropija dviju slika),...
- Ovdje se razmatra kriterij koji koristi korelacijske tehnike, odnosno računanje **korelacijskog koeficijenta**  $s$ :

$$s = \frac{\sum_{i=1}^M \sum_{j=1}^N (f_1(i, j) - \bar{f}_1) \cdot (f_2(i, j) - \bar{f}_2)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N (f_1(i, j) - \bar{f}_1)^2} \cdot \sqrt{\sum_{i=1}^M \sum_{j=1}^N (f_2(i, j) - \bar{f}_2)^2}}$$



# Registracija slike



- $N$  je dimenzija prostora,  $\bar{f}_1$  i  $\bar{f}_2$  predstavljaju prosječne vrijednosti od  $f_1$  i  $f_2$ , koje se računaju na sljedeći način:

$$\bar{f}_k = \frac{1}{NM} \sum_{i=1}^M \sum_{j=1}^N f_k(i, j), \quad k = 1, 2$$

- Za jednodimenzionalni slučaj dobiva se:

$$\begin{aligned} S &= \frac{\sum_{i=1}^N ((f_1(i) - \bar{f}_1) \cdot (f_2(i) - \bar{f}_2))}{\sqrt{\sum_{i=1}^N (f_1(i) - \bar{f}_1)^2 \cdot \sum_{i=1}^N (f_2(i) - \bar{f}_2)^2}} \\ &= \frac{\sum_{i=1}^N (f_1(i) - \bar{f}_1)}{\sqrt{\sum_{i=1}^N (f_1(i) - \bar{f}_1)^2}} \cdot \frac{\sum_{i=1}^N (f_2(i) - \bar{f}_2)}{\sqrt{\sum_{i=1}^N (f_2(i) - \bar{f}_2)^2}} \\ &= \underbrace{\frac{\sum_{i=1}^N (f_1(i) - \bar{f}_1)}{\sqrt{\sum_{i=1}^N (f_1(i) - \bar{f}_1)^2}}}_{d_1} \cdot \underbrace{\frac{\sum_{i=1}^N (f_2(i) - \bar{f}_2)}{\sqrt{\sum_{i=1}^N (f_2(i) - \bar{f}_2)^2}}}_{d_2} \end{aligned}$$

- Normirani vektori  $d_i$  mogu se koristiti kao deskriptori značajki.

# Registracija slike

- Bolji rezultati u slaganju slika dobivaju se korištenjem **SURF** (engl. Speeded Up Robust Feature) algoritma koji se zasniva na računanju maksimuma u aproksimiranoj Hessijan-Gaussovoj sekvenci.
- On računa vektorski deskriptor koji sadrži 64 elementa i opisuje svojstva piksela na način da je invarijantan na skaliranje i rotaciju i koristi piksele unutar regije od interesa.
- Segmentacija slike je jako efikasna za pronalaženje objekta u slici, praćenje objekta u nizu slika (film), kao i za određivanje orijentacije objekta na slici. Osim navedenih deskriptora, odnosno detektora značajki, veliku primjenu nalazi i **SIFT** (engl. scale-invariant feature transform).



## 5.9. Stereovizija

- Jedna kamera ne može lokalizirati objekat u 3D prostoru, ona daje informaciju **o smjeru**, a ne i o **udaljenosti** do njega.
- Korištenjem kamere, 3D slika se pretvara u 2D sliku.
- Postavlja se pitanja: **“Kako iz kanoničkih koordinata slike  $x_i$  i  $y_i$ , izvući informaciju o koordinatama objekta  $x, y$  i  $z$ ?”**
- Za određivanje udaljenosti do objekta nužno je koristiti više slika i identificirati značajke u njima.



# Stereovizija

- U mobilnoj robotici je važno izbjeći objekat, a ne imati informacije o njegovoj boji.
- U vezi s tim je zgodno izdvojiti rubne tačke (piksele) i povezati ih i na taj način dobiti linijski segment.
- Na ovaj način se smanjuje količina podataka sa kojima je prikazana slika.
- Problem izvlačenje informacija o dubini na temelju dvije slike koje oslikavaju scenu iz različitih perspektiva se rješava **stereovizijom**.
- Problem se može poopćiti i na više slika, odnosno više kamera. Korespondentne tačke moguće je matematički povezati **epipolarnom geometrijom**.



# Stereovizija

## Epipolarna geometrija

- Epipolarna geometrija opisuje odnose između slika dobivenih pomoću dvije kamere koje snimaju scenu iz dva različita ugla.

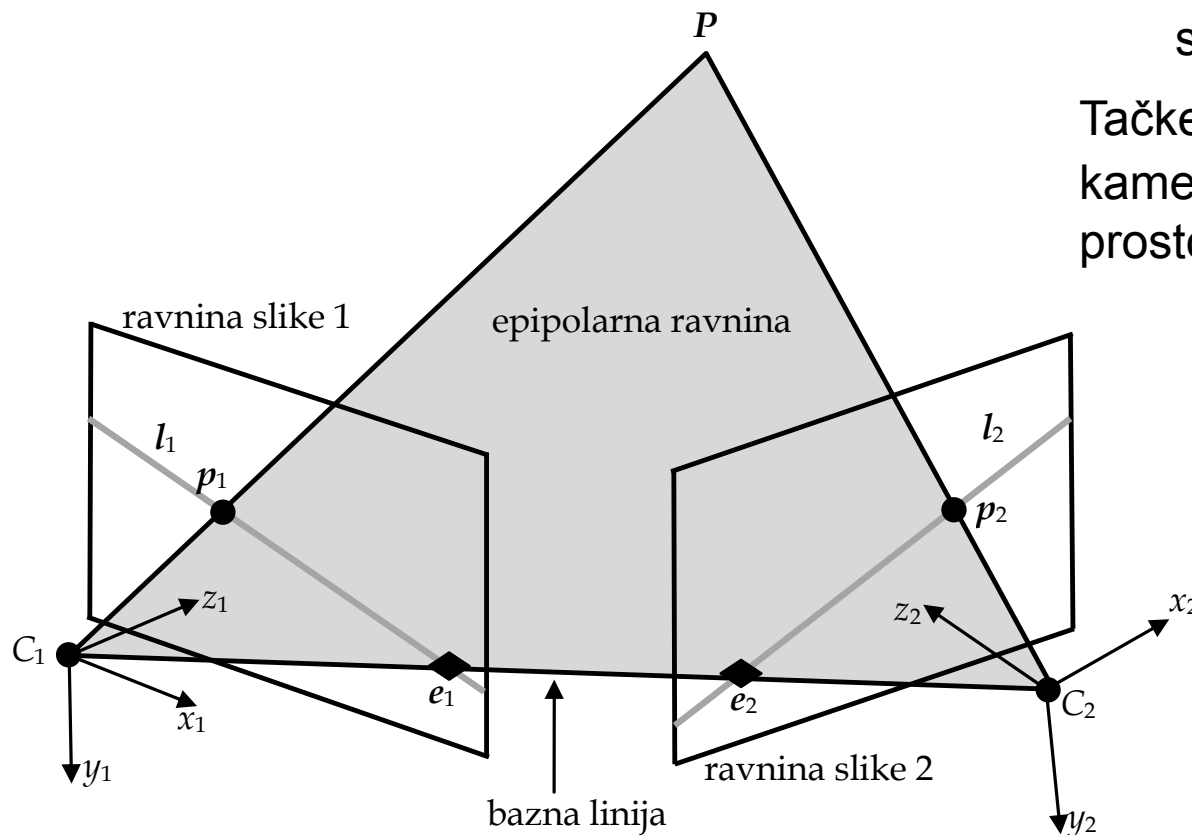


127/139

$P = (x, y, z)$  tačka u 3D prostoru scene

Tačke  $p_1$  i  $p_2$  definiraju optičke zrake kamere koje prolaze kroz tačku  $P$  u 3D prostoru.

Ravnina koja prolazi tačkom  $P$  i centrima projekcije obje kamere naziva se **epipolarna ravnina**. Sve tri tačke  $P$ ,  $p_1$  i  $p_2$  su koplanarne, odnosno leže u istoj epipolarnoj ravnini. Problem koji definira epipolarna geometrija jest određivanje odnosa između tačaka  $p_1$  i  $p_2$



# Stereovizija

- Linija na kojoj se nalaze obje kamere naziva se **bazna linija** (engl. baseline), a mjesto na kojem se sijeku bazna linija i ravnina slike naziva se **epipol** ili epipolarna tačka (označena sa  $e_1$  i  $e_2$ ).
- Ustvari, epipol predstavlja projekciju optičkog centra jedne kamere na ravninu slike druge kamere.
- Linije koje predstavljaju presjeke epipolarne ravnine s ravninama slika nazivaju se **epipolarne linije**, koje su označene sa  $l_1$  i  $l_2$ .
- Epipolarne linije se također mogu definirati kao linije na slici koje prolaze *epipolom* i projekcijom tačke  $P$  (to jest tačkom  $p_1$ , odnosno  $p_2$ ).





# Stereovizija

- Epipolarna linija može se promatrati kao projekcija optičke zrake prve kamere koja prolazi optičkim centrom prve kamere i tačkom  $P$  na sliku druge kamere.
- Budući da se tačka  $P$  može nalaziti samo na toj optičkoj zruci, to znači da se njezina projekcija na drugu sliku može nalaziti samo na razmatranoj epipolarnoj liniji.
- Ta činjenica naziva se ***epipolarno ograničenje***.
- Epipolarno ograničenje (uvjet) govori da se tačka može nalaziti samo na optičkoj osi gdje epipolarna ravnina siječe ravninu slike dobivene s druge kamere



# Stereovizija

- Kada se za neku tačku prve slike traži tačka na drugoj slici koja odgovara istoj 3D tački scene, tada je pretraživanje druge slike ograničeno samo na epipolarnu liniju koja odgovara tački  $p_1$ .
- Uz poznavanje intrinzičnih parametara obje kamera te njihovog međusobnog položaja, moguće je za svaku tačku prve slike odrediti njoj odgovarajuću epipolarnu liniju na drugoj slici i obratno.



# Stereovizija

## Fundamentalna matrica

- Fundamentalna matrica matematički opisuje epipolarnu geometriju.
- Tačka  $p_2$  mora uvijek ležati na epipolarnoj liniji  $l_2$  ravnine druge slike.
- Da bi se pronašla tačka  $p_2$  potrebno je samo pronaći preslikavanje tačke  $p_1$  na epipolarnu liniju  $l_2$ .
- Ovo se obavlja na sljedeći način.
- Neka se promatra pomoćna ravnina  $\pi$  koja ne siječe niti jedan od centara kamere.
- Mjesto gdje optička zraka prve kamere siječe ravninu  $\pi$  označeno je tačkom  $P$ .



# Stereovizija

- Projekcija ove tačke na ravninu druge slike odgovara tački  $p_2$  koja se nalazi na epipolarnoj liniji  $l_2$ .
- Budući da epipolarna tačka  $e_2$  i tačka  $p_2$  leže na epipolarnoj liniji, tada vrijedi:

$$l_2 = e_2 \times p_2$$

- Ukoliko bi tačku  $p_2$  prikazali kao umnožak vektora  $p_1$  i neke matrice  $H_\pi$  tada bi se prethodni izraz mogao zapisati kao:

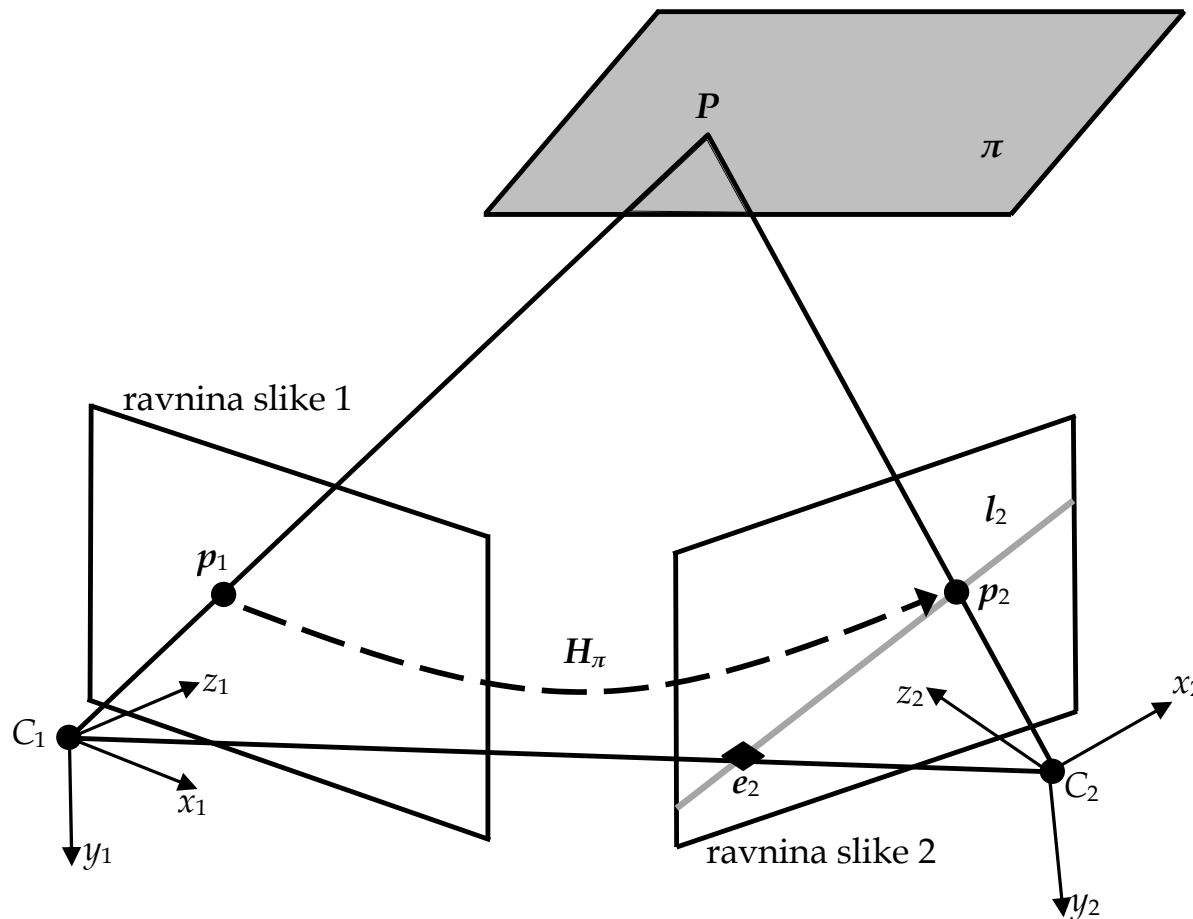
$$l_2 = e_2 \times p_2 = e_2 \times H_\pi p_1 = F p_1$$

gdje je  $F$  fundamentalna matrica i  $H_\pi$  2D homografska matrica koja se može definirati pošto su preslikavanja tačke  $p_1$  u tačku  $p_2$  projekcijski ekvivalentna u odnosu na tačku  $P$ .



# Stereovizija

- Važno je naglasiti da u slučaju kada međusobni položaj kamera nije poznat tada je potrebno estimirati fundamentalnu matricu.



# Stereovizija

- Važno je naglasiti da u slučaju kada međusobni položaj kamera nije poznat tada je potrebno estimirati fundamentalnu matricu.
- **Primjer . Određivanje korespondentnih tačaka i epipolarnih linija pomoću sistema stereovizije, odnosno sistema dviju kamera.**
- Na sljedećoj slici prikazane su slike dobivene lijevom i desnom kamerom, koje su poravnate i međusobno razmaknute za 30 cm, gdje je  $f = 5.2$  mm.
- Na temelju procesa slaganja prve i druge slike korištenjem SURF deskriptora pronađene su 132 korespondentne tačke.
- Nakon toga je primijenjena metoda *konsenzusa slučajnih uzoraka* (RANSAC - Random Sampling and Consensus) za estimaciju fundamentalne matrice.

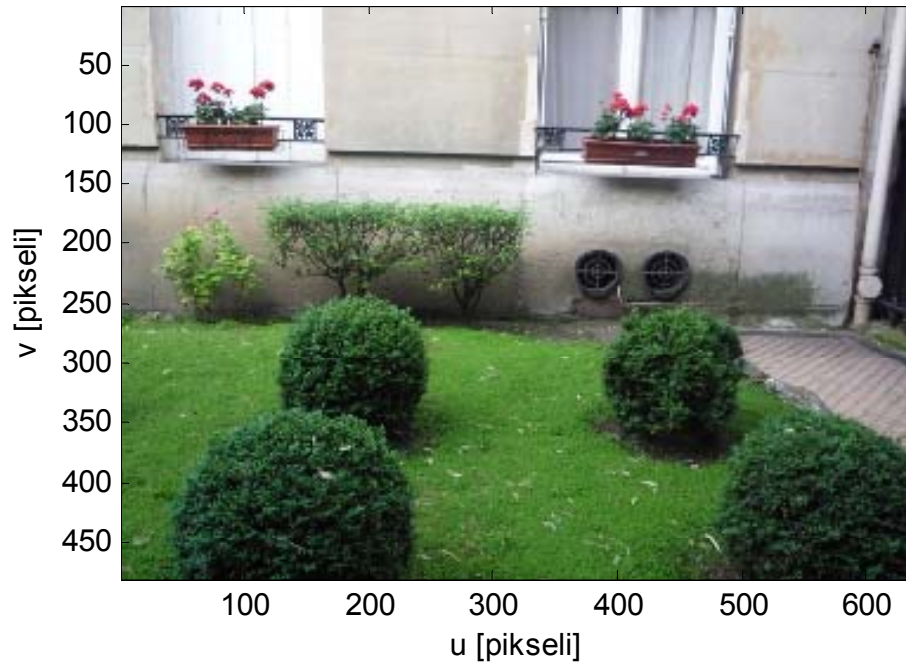


# Stereovizija

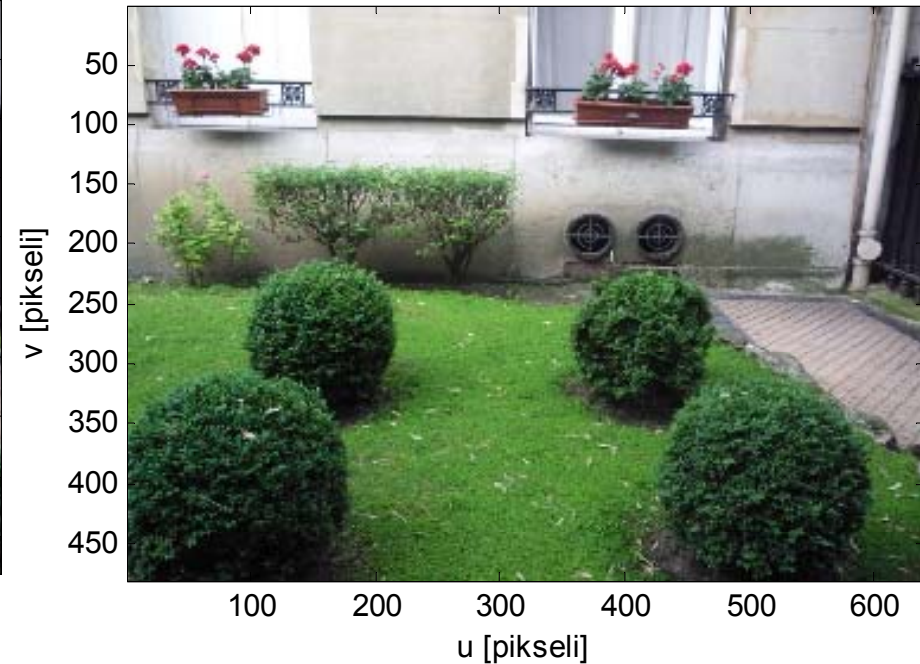
- Važno je naglasiti da u slučaju kada međusobni položaj kamera nije poznat tada je potrebno estimirati fundamentalnu matricu.



135/139



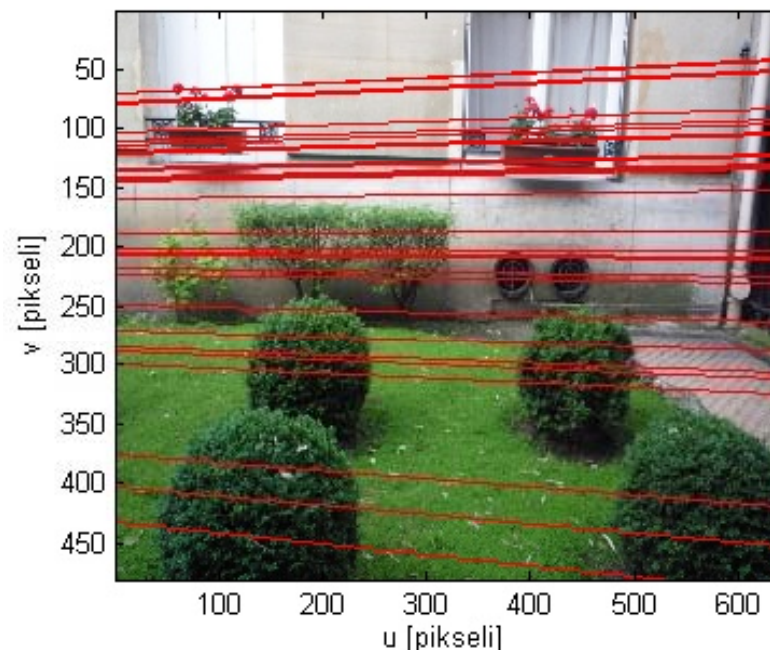
**Slika dobivena lijevom kamerom**



**Slika dobivena desnom kamerom**

# Stereovizija

- U ovom primjeru dobiveno je 53 podobnih (40.2%) i 79 nepodobnih (59.8%) tačaka.
- Položaj podobnih tačaka na sceni dobivenoj pomoću lijeve kamere prikazan je na lijevoj slici, dok su na desnoj slici predočene epipolarne linije.



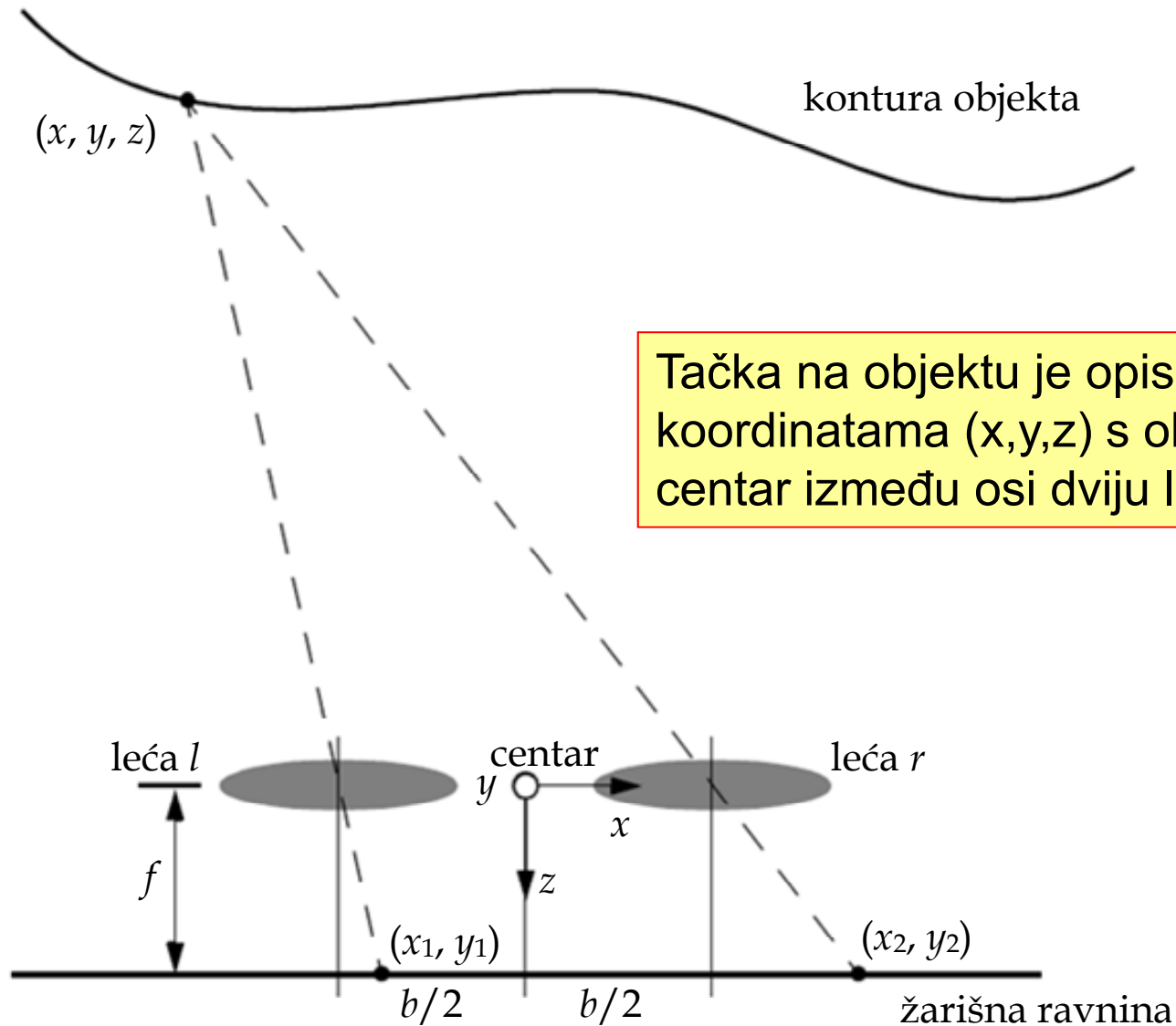


# Stereovizija

## Određivanje dubine slike pomoću triangulacije



137/139



# Stereovizija

- Iz prethodne slike slijedi:

$$\frac{x_1}{f} = \frac{x + b/2}{z}$$
$$\frac{x_2}{f} = \frac{x - b/2}{z}$$

odnosno,

$$\frac{y_1}{f} = \frac{y_2}{f} = \frac{y}{z}$$

Iz gornjih izraza dobiva se izraz za **disparitet** (razlika u koordinatama slike  $(x_1 - x_2)$ ):

$$\frac{x_1 - x_2}{f} = \frac{b}{z}$$



# Stereovizija

- Samo se mjerenjem dispariteta mogu izlučiti informacije o dubini.
- Na temelju gornjih jednadžbi dobivaju se izrazi za koordinate objekta:

$$x = b \frac{(x_1 + x_2) / 2}{x_1 - x_2}$$
$$y = b \frac{(y_1 + y_2) / 2}{x_1 - x_2}$$
$$z = b \frac{f}{x_1 - x_2}$$

- Udaljenost kamere do objekta je inverzno proporcionalna disparitetu.
- Disparitet je proporcionalan sa  $b$ ; za zadanu pogrešku dispariteta sa povećanjem  $b$ -a povećava se tačnost estimacije dubine.
- Kako  $b$  raste, zbog fizičke odvojenosti kamera, neki objekti se mogu pojaviti na jednoj kameri, a na drugoj ne.

Sva navedena priča vrijedi ako su kamere idealno poravnate, što u praksi nije slučaj – tada se koristi kalibracija kamere.

