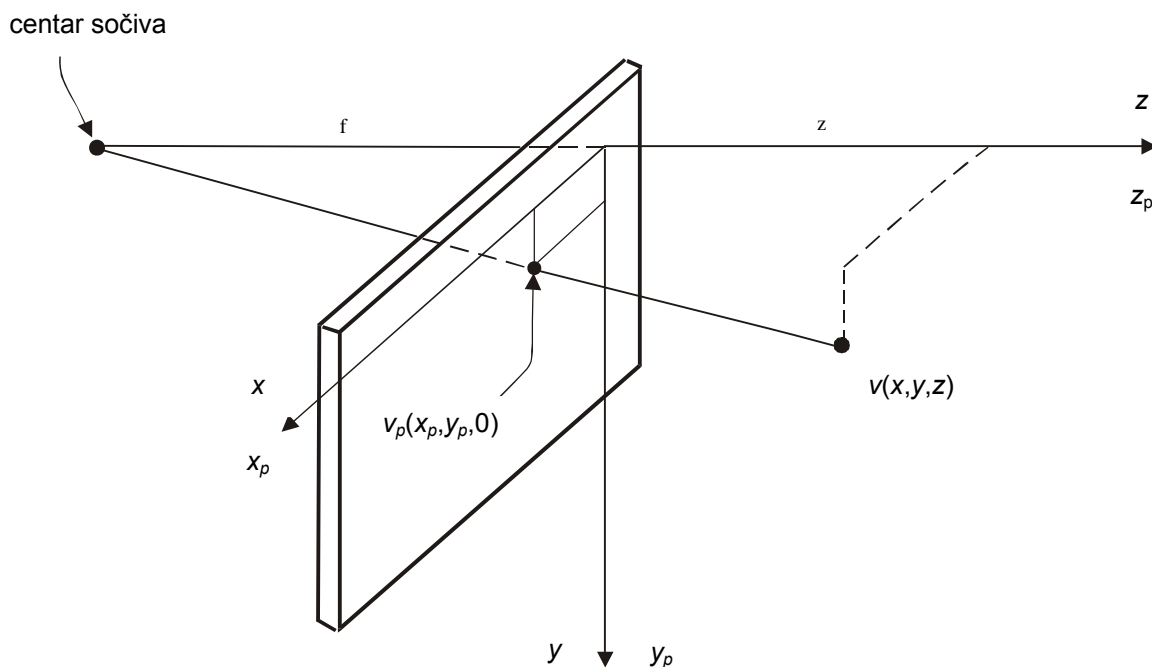


SISTEMI PREPOZNAVANJA KOD ROBOTA

Kompjutersko gledanje, odnosno prepoznavanje kod robota, se odvija po sledećim fazama:

- **formiranje slike,**
- **procesiranje slike,**
- **prepoznavanje,**
- **akcije robota.**

Formiranje slike i struktura sistema prepoznavanja



Centralna projekcija projektivnog zraka

Polazi se od kartezijanskih koordinata tačke v_p , koje se na osnovu modela prikazanog na slici mogu definisati na sledeći način:

$$v_p = \begin{bmatrix} x_p \\ y_p \\ 0 \end{bmatrix}; \quad x_p = \frac{xf}{(f+z)}; \quad y_p = \frac{yf}{(f+z)}; \quad z_p = 0, \quad (1)$$

Ako se primeni *direktna perspektivna transformacija* kao homogena matična transformacija, za model prikazan na slici, sledi da su homogene koordinate vektora \tilde{v}_p date preko sledećeg izraza:

$$\tilde{v}_p = T_p \cdot \tilde{v} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{f} & 1 \end{bmatrix} \cdot \begin{bmatrix} WX \\ WY \\ WZ \\ W \end{bmatrix} = \begin{bmatrix} WX \\ WY \\ WZ \\ \frac{WZ}{f} + W \end{bmatrix} \quad (2)$$

Vektor \tilde{v} u jednačini (2) je dat u homogenim koordinatama i u odnosu na aktuelne kartezijanske koordinate tačke $v = [x, y, z]^T$ proširen je i pomnožen koeficijentom skaliranja W . U istoj jednačini matrica T_p predstavlja homogenu transformaciju.

Na osnovu (2), deljenjem svakog od prva tri člana vektora \tilde{v}_p četvrtim članom, mogu se definisati kartezijanske koordinate tačke v_p na ravni slike:

$$v_p = \begin{bmatrix} \frac{fx}{f+z} \\ \frac{fy}{f+z} \\ \frac{fz}{f+z} \end{bmatrix} \quad (3)$$

Prema datoj slici očigledno je da koordinata z_p (treći član vektora v_p) mora biti jednaka nuli, s obzirom da je slika u ravni.

Ako se sada primeni *inverzna perspektivna transformacija*, slično prethodnom, moguće je dobiti homogene koordinate vektora projektivnog zraka:

$$\tilde{v} = T_p^{-1} \cdot \tilde{v}_p = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{f} & 1 \end{bmatrix} \cdot \begin{bmatrix} WX_p \\ WY_p \\ WZ_p \\ W \end{bmatrix} = \begin{bmatrix} WX_p \\ WY_p \\ WZ_p \\ W - \frac{WZ_p}{f} \end{bmatrix} \quad (4)$$

Slično određivanju v_p , deljenjem svakog od prva tri člana vektora \tilde{v} četvrtim članom, dobija se da je:

$$v = \begin{bmatrix} \frac{fx_p}{f-z_p} \\ \frac{fy_p}{f-z_p} \\ \frac{fz_p}{f-z_p} \end{bmatrix}; \quad x = \frac{fx_p}{f-z_p}; \quad y = \frac{fy_p}{f-z_p}; \quad z = \frac{fz_p}{f-z_p} \quad (5)$$

Eliminacijom z_p iz (5), posle izračunavanja i sređivanja, sledi da je:

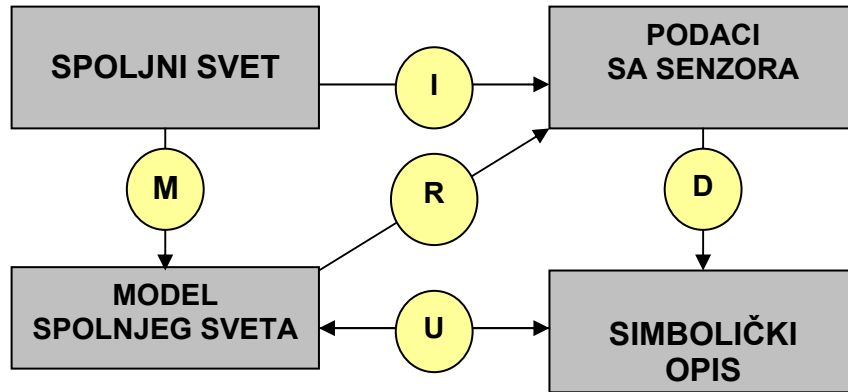
$$x = \frac{x_p}{y_p} y = \frac{x_p(f+z)}{f} \Rightarrow y_p = \frac{yf}{(f+z)}, \quad (6)$$

što je saglasno jednačini (1)

Vidi se da je moguće, korišćenjem homogenih koordinata i homogenih transformacija, odrediti kartezijanske koordinate tačke na ravni slike, kao i jednačinu projektivnog zraka koji prolazi kroz datu tačku na ravni slike i kroz centar sočiva.

Opšta struktura sistema prepoznavanja

Struktura opšteg modela sistema prepoznavanja je pokazana na slici.

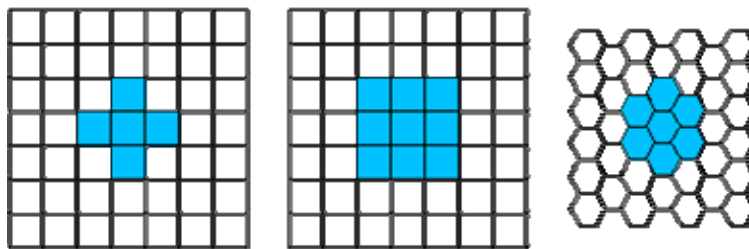


Struktura opšteg modela sistema prepoznavanja

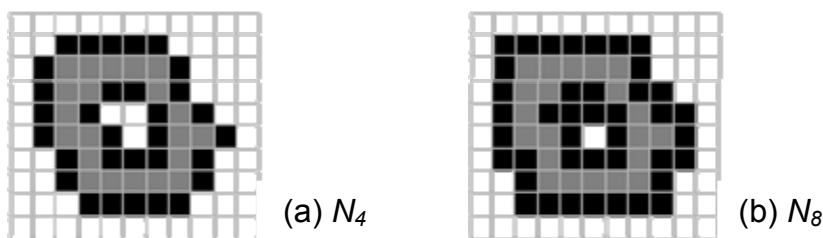
Opštu strukturu sistema prepoznavanja čini realan svet, domen digitalnih podataka senzora, domen modela i međudomen simboličkog opisa scene odnosno domen karakteristika, a njihovo povezivanje obuhvata pet osnovnih procesa:

- **proces formiranja digitalne slike (I),**
- **proces opisa – obrade podataka sa senzora (D),**
- **proces modeliranja realnih objekata (M),**
- **proces prepoznavanja – razumevanja (U),**
- **proces formiranja prikaza – sintetičkih senzorskih podataka (R).**

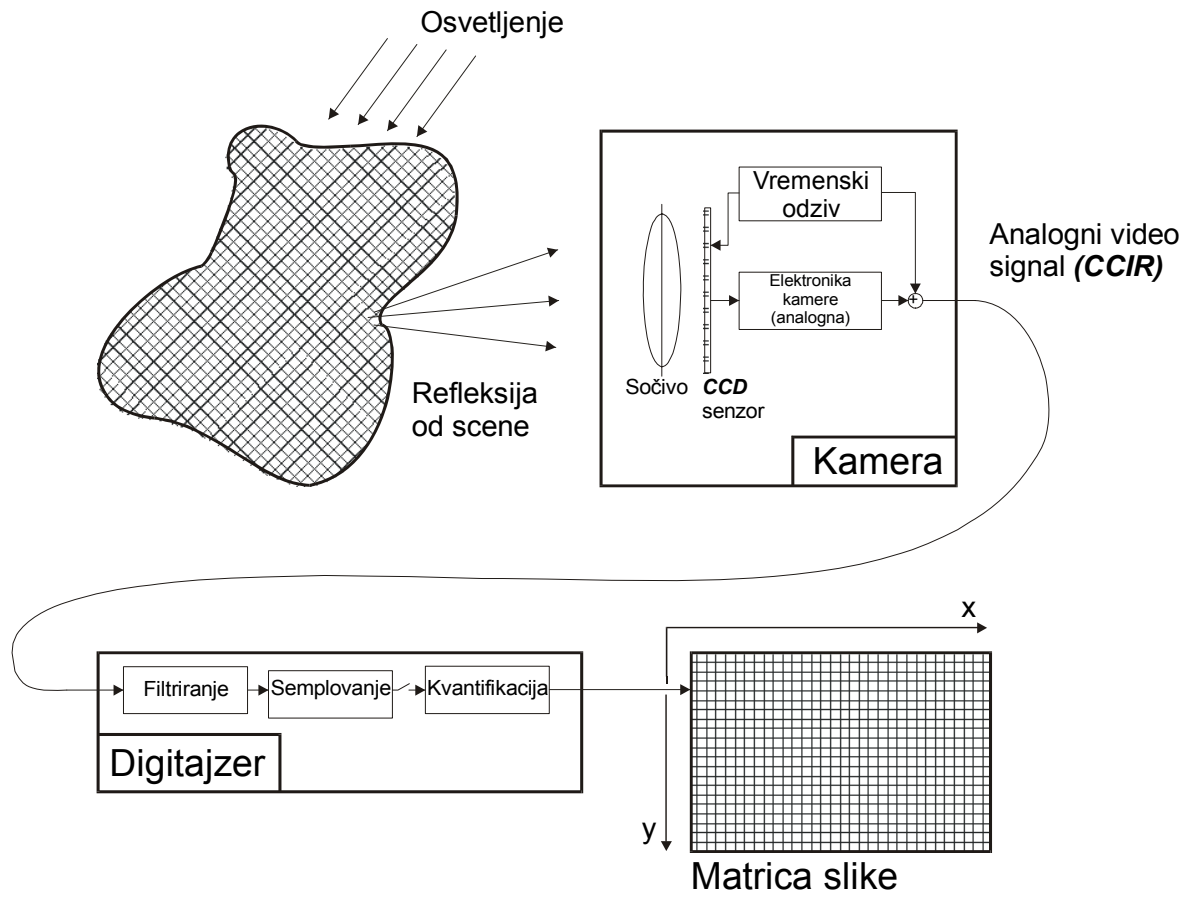
Karakteristike procesa formiranja digitalne slike



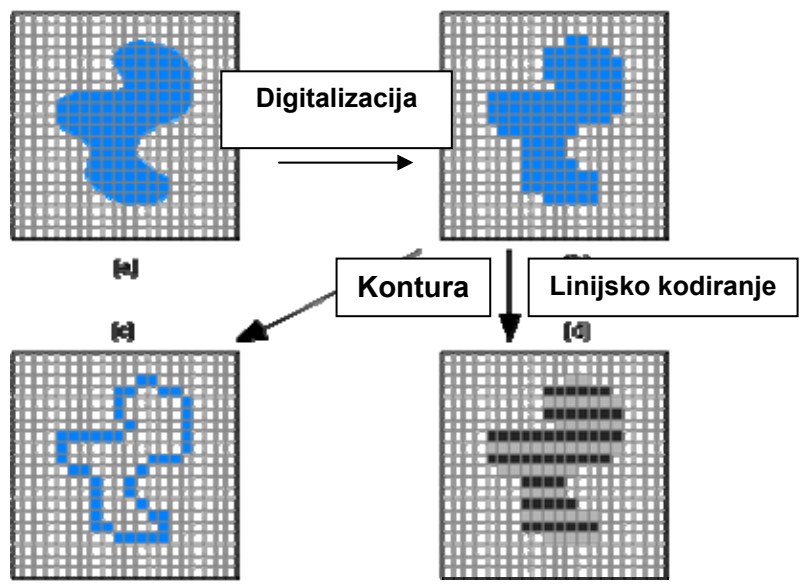
Spajanje elemenata slike



Ilustracija pojave dilatacije za N_4 i N_8



Formiranje digitalne slike



Konverzija kontinualne analogne slike u digitalni oblik

Procesiranje i analiza slike

Da bi se aktuelan objekat prepoznao, potrebno je da se prvo identifikuju regioni slike koji mogu da pripadaju tom objektu. Proces simplifikacije kompleksnog objekta podrazumeva grupisanje tih identifikovanih regiona u skupove po odgovarajućim svojstvima. Ovakav proces se naziva "**segmentacija**" i postoje dva osnovna pristupa ("tehnike"): **preko kontura** i **preko regiona**.

Proces segmentacije se, generalno posmatrano, realizuje u tri koraka:

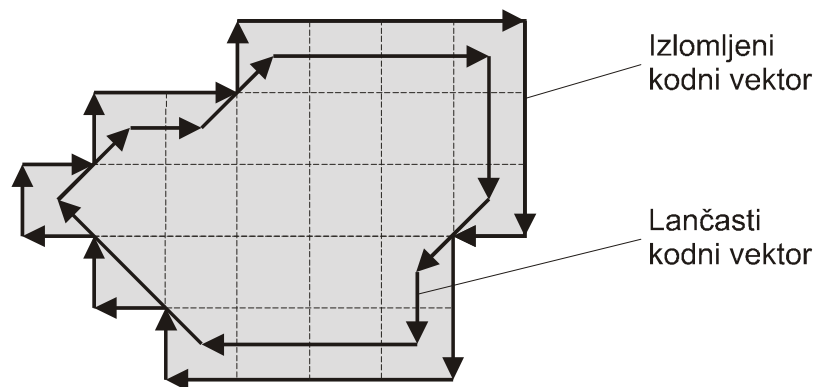
1. **Klasifikacija**, gde se pikseli klasifikuju u skupove prema intenzitetu osvetljenosti.
2. **Predstavljanje**, gde se prethodno klasifikovani skupovi prikazuju u formi koja je pogodna za dalji opis, bilo povezanih regiona bilo kontura.
3. **Opisivanje**, gde su skupovi piksela opisani u skalarnom ili vektorskom obliku.

Segmentacija preko kontura

Pri izdvajanju ivica, u toku procesiranja digitalne slike, koja može da bude i binarna, neophodne su sledeće informacije:

- lokalna orijentacija elemenata ivice,
- kontrast u osvetljenosti susednih regiona,
- širina pojasa elemenata ivice,
- određivanje položaja tačke ivice,
- intenzitet elemenata ivice, i
- intenzitet susednih regiona.

Segmentacija preko kontura polazi od ocrtavanja regiona preko ivica koje mogu biti predstavljene pomoću pogodnih krivih i pravih linija, lančastim ili izlomljenim kodnim vektorom, što je prikazano na slici. Izlomljeni kodni vektor prikazuje ivicu kao niz horizontalnih i vertikalnih linijskih segmenata, koji prate isprekidanu konturu oivičavajući granicu skupa piksela, dok lančasti kodni vektor određuje ivicu pomoću povezanih centara graničnih piksela konture.

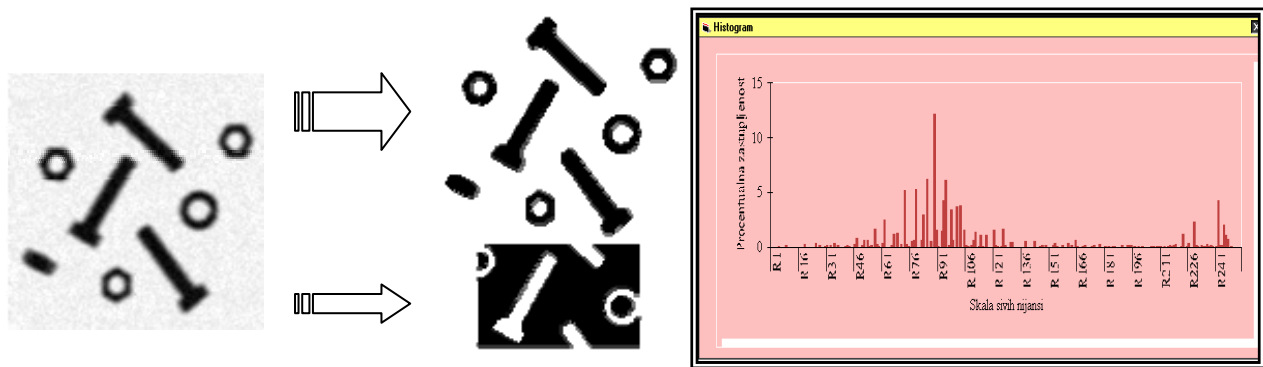


Predstavljanje konture regiona

Segmentacija preko regiona

Tehnika segmentacije preko regiona je bazirana na jednostavnom konceptu, koji se realizuje kroz grupisanje elemenata slike sa sličnim vrednostima osvetljenosti pomoću graničnog nivoa osvetljenosti-praga (*binarna segmentacija*).

Proces **binarne segmentacije** je postupak koji na najjednostavniji način vrši izdvajanje objekta od pozadine. Da bi se realizovao ovaj postupak neophodno je sprovesti nekoliko procesnih koraka, zasnovanih na uniformnim regionima. Posle transformacije slike, grupe piksela formiraju regione, odnosno segmente. Kada su regioni definisani, može se pristupiti utvrđivanju nivoa osvetljenosti. Za objekte na slici formira se **histogram**, koji daje informaciju kolika je procentualna zastupljenost formiranih regiona sa istom nijansom sivog, odnosno istim nivoom osvetljenosti.



Proces binarne segmentacije korišćenjem histograma osvetljenosti

Histogram h za digitalnu sliku I je definisan pomoću:

$$h(m) = \#\{(x,y) \mid I(x,y) = m\},$$

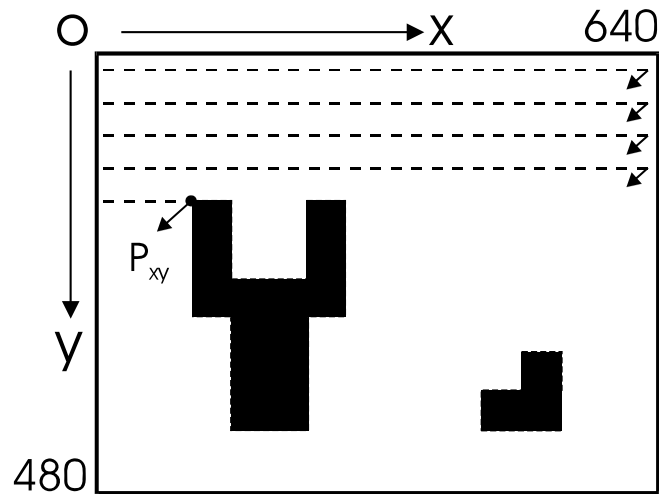
gde m obuhvata svaku vrednost nivoa sivog, a $\#$ je operator koji računa broj elemenata u skupu.

Postupak koji se najčešće koristi u toku analize slike podrazumeva poređenje slike "piksel po piksel" sa odabranim pragom T , tako da objekat može da bude "obojen" u crno a pozadina u belo, ili obrnuto, čime je izvršena **binarizacija slike**. Izbor praga T predstavlja centralni problem pri realizaciji binarne segmentacije.

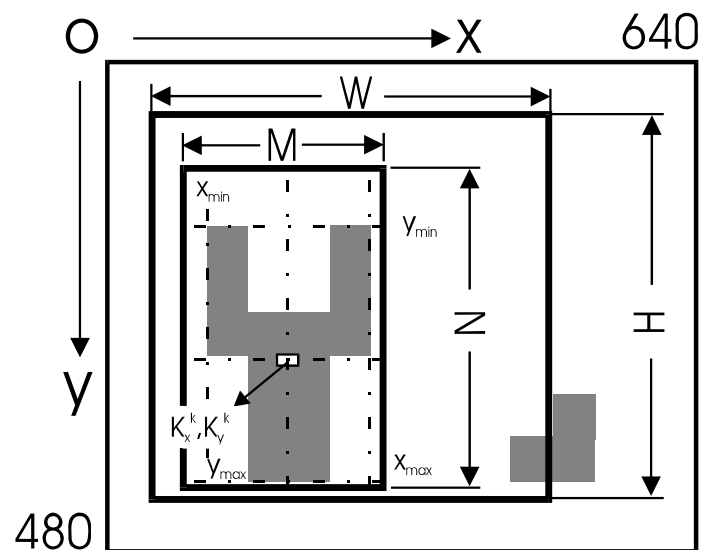
Binarna segmentacija je najjednostavniji vid klasifikacije i ostvaruje se primenom granične vrednosti osvetljenosti – praga T , tako da se za binarnu sliku može napisati:

$$f_{i,j} \in \begin{cases} S_b & \text{ako je } I_{ij} < T \\ S_f & \text{ako je } I_{ij} \geq T \end{cases}$$

gde je $f_{i,j}$ piksel (i,j) , a S_b i S_f su respektivno dati skupovi piksela koji određuju pozadinu scene (belo $\rightarrow S_b = 0$) i objekat (crno $\rightarrow S_f = 1$). Pomoću histograma je, jednostavnim procedurom, moguće odrediti vektor R , koji je dimenzionisan od 0 do MAX. Vrednost 0 je najmanja vrednost od mogućih vrednosti nivoa sivog, a MAX je najveća. Utvrđeno je, za primer sa slike gore, da je izabrani granični nivo osvetljenosti - prag T , sa najvećom zastupljenošću sivih nijansi, najčešće bio između 100 i 150 (MAX=256).



Skeniranje binarne slike



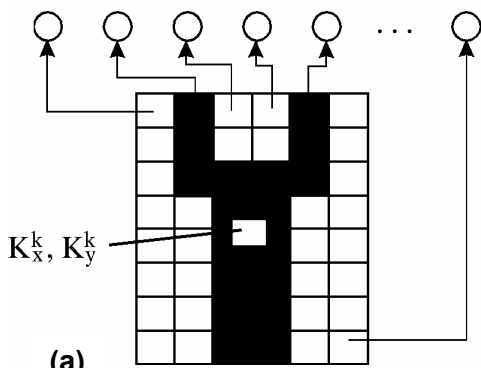
Izdvajanje regiona objekta i određivanje težišta

Kada je, korišćenjem analize povezanosti, izvršena segmentacija binarne slike na regione, moguće je izračunati karakteristike svakog labelisanog regiona, kao što su površina, težište, obim, itd. Značajno je da se uoči da je na slici izdvojena podoblast slike $M \times N$ u kojoj se nalazi labelisani region objekta, zbog uštede u memoriji računara.

Karakteristike koje se najčešće koriste mogu se svrstati u dve grupe:

1. one koje su nezavisne od pozicije i orijentacije objekta, kao što su površina, obim, faktor oblika, broj rupa, itd.,
2. one koje su zavisne od pozicije i orijentacije objekta, kao što su težište, momenti inercije za težišne ose, pravac glavne ose, itd.

$$\vec{I}_k = \begin{bmatrix} 0100100100100 \\ 111100 \dots 1100 \end{bmatrix}$$



(a)

Vrsta	K o l o n a					
	0	1	2	3	4	5
0	0	1	0	0	1	0
1	0	1	0	0	1	0
2	0	1	1	1	1	0
3	0	0	1	1	0	0
4	0	0	1	1	0	0
5	0	0	1	1	0	0
6	0	0	1	1	0	0
7	0	0	1	1	0	0

(b)

Vrsta	Trojka
0	0 → [0,1,0][1,1,1][2,2,0][4,1,1][5,1,0]
1	5 → [0,1,0][1,1,1][2,2,0][4,1,1][5,1,0]
2	10 → [0,1,0][1,4,1][5,1,0]
3	13 → [0,2,0][2,2,1][4,2,0]
4	16 → [0,2,0][2,2,1][4,2,0]
5	19 → [0,2,0][2,2,1][4,2,0]
6	22 → [0,2,0][2,2,1][4,2,0]
7	25 → [0,2,0][2,2,1][4,2,0]

(c)

Linijsko kodiranje binarne slike.

(a) Slika objekta. (b) Raspored piksela. (c) Linijski kod.

Izdvajanje karakteristika iz binarne slike se najčešće izvršava istovremeno sa analizom povezanosti, pri čemu treba naglasiti da je uglavnom neophodno postprocesiranje slike posle označavanja. Geometrijske karakteristike za određeni region u labelisanoj slici su aproksimativne, jer je slika digitalna. Pripadnost elementa slike regionu je određena opisanim tehnikama, tako da se može definisati njegova vrednost kao:

$$B(x, y) = \begin{cases} 1, & \text{(za elemente slike u regionu)} \\ 0, & \text{(za elemente slike van regiona)} \end{cases}$$

Za svaki region nekog objekta je moguće izračunati sledeće karakteristike:

- površinu **A** slike objekta:

$$A = \sum_{x=l}^n \sum_{y=k}^m B(x, y)$$

- koordinate težišta slike objekta:

$$K_x^k = \frac{1}{A} \cdot \sum_{x=l}^n \sum_{y=k}^m x \cdot B(x, y)$$

$$K_y^k = \frac{1}{A} \cdot \sum_{x=l}^n \sum_{y=k}^m y \cdot B(x, y)$$

- momente inercije za težišne ose slike objekta:

$$I_x = \frac{1}{A} \cdot \sum_{x=l}^n \sum_{y=k}^m (y - K_y^k)^2 \cdot B(x, y)$$

$$I_y = \frac{1}{A} \cdot \sum_{x=l}^n \sum_{y=k}^m (x - K_x^k)^2 \cdot B(x, y)$$

$$I_{xy} = \frac{1}{A} \cdot \sum_{x=l}^n \sum_{y=k}^m (x - K_x^k)(y - K_y^k) \cdot B(x, y)$$

- pravac glavne ose:

$$\Theta = \frac{1}{2} \arctg \frac{2 \cdot I_{xy}}{I_y - I_x}$$

Skaliranjem pomoću faktora s_x i s_y moguće je odrediti realne veličine elementa slike (širinu i visinu), odnosno karakteristika koje su iz slike izdvojene. Dakle, moguće je raditi sa stvarnom površinom P i koristiti realne karakteristike pri generisanju jedinstvenog opisa objekta, pri čemu će se one porediti sa karakteristikama ostalih regiona u cilju prepoznavanja objekta i određivanja njegove pozicije i orijentacije. Jednačine koje se tada koriste su date u nastavku:

- površina P objekta

$$P = s_x s_y A$$

- koordinate težišta objekta

$$x_k = s_x \frac{\sum x}{A}$$

$$y_k = s_y \frac{\sum y}{A}$$

- glavni poluprečnici inercije objekta

$$i_{(1)} = \sqrt{\frac{a+b+e}{2f}}$$

$$i_{(2)} = \sqrt{\frac{a+b-e}{2f}}$$

- orijentacija objekta

$$\Re = \frac{1}{2} \arctg 2(2c, a-b)$$

gde su

$$a = \frac{4}{\pi} s_x^3 s_y \left(\sum x^2 - \frac{(\sum x)^2}{A} \right)$$

$$b = \frac{4}{\pi} s_x s_y^3 \left(\sum y^2 - \frac{(\sum y)^2}{A} \right)$$

$$c = \frac{4}{\pi} s_x^2 s_y^2 \left(\sum xy - \frac{\sum x \sum y}{A} \right)$$

$$e = \sqrt{(a-b)^2 + \frac{4}{\pi} c^2}$$

$$f = \sqrt[4]{ab - c^2}$$

Za primer sa slike i faktore skaliranja $s_x = s_y = 1$, a na osnovu gornjih izraza izdvojene su sledeće karakteristike:

- *granice podoblasti slike*
 $x_{\min} = 1$ $x_{\max} = 4$
 $y_{\min} = 0$ $y_{\max} = 7$
- *površina objekta*
 $P = A = 18$
- *koordinate težišta objekta*
 $x_k = 2.5$
 $y_k = 3.3$
- *glavni poluprečnici inercije objekta*

$$\left. \begin{array}{l} \sum x = \sum (1+4+1+4+\dots+2+3) = 45 \\ \sum y = \sum (0+0+1+1+\dots+7+7) = 60 \\ \sum xy = 150 \\ \sum x^2 = 129 \\ \sum y^2 = 288 \\ a \rightarrow 21.02 \\ b \rightarrow 112.1 \\ c \rightarrow 0.0 \\ e \rightarrow 91.08 \\ f \rightarrow 6.96 \end{array} \right\} \Rightarrow \begin{array}{l} i_{(1)} = 4.01 \\ i_{(2)} = 1.73 \end{array}$$

- *orijentacija objekta*
 $\mathfrak{R} = 0^\circ$

Za povezan, odnosno labelisan region, moguće je jednostavnim algoritmima izdvojiti i njegovu konturu, kako bi se definisao oblik regiona. Izdvojeni elementi slike, koji pripadaju konturi regiona, dalje se mogu predstaviti pomoću dvodimenzionalne matrice, redosledno memorisanim koordinatama svakog elementa konture ili kodnim lancem. Kodni lanac opisuje digitalnu liniju koja je predstavljena početnom tačkom i nizom kodova nagiba, a taj opis je značajan za kompaktno geometrijsko povezivanje.

Prepoznavanje objekta

Prepoznavanje objekta, poređenjem kompletne figure sa referentnim šablonom ("**template matching**"), je kompleksan postupak baziran na konvencionalnoj metodi. Problemi koji su prisutni kod konvencionalnih sistema su uglavnom vezani za dugo vreme sračunavanja, posebno kada se lik (objekat) koji se "traži" u slici pojavljuje u različitim orijentacijama.

Danas se primenjuju i napredne metode koje su bazirane na veštačkim neuronskim mrežama. Poređenje objekta za koji je formiran binarni vektor sa objektom koji se "traži", realizuje se korišćenjem ART-1 veštačke neuronske mreže, tako da se danas sve češće koristi metoda bazirana na toj veštačkoj neuronskoj mreži, kao i na hibridnom sistemu gde se ova neuronska mreža pojavljuje kao FUZAMP, odnosno *fuzzy* ARTMAP neuronska mreža. Procedura je slična konvencionalnom poređenju preko šablona, s tim što se sličnost među objektima određuje kroz kompetitivno obučavanje ART-1 veštačke neuronske mreže. Oni mogu da budu u proizvoljnoj orijentaciji, što predstavlja značajnu prednost ove nove procedure.

Pretpostavimo da je šablon za poređenje dat nizom elemenata slike S , a da je scena u kojoj se on nalazi data nizom elemenata slike W , i da je $W \supseteq S$. Za tačku x , koja se nalazi unutar slike I , što se može utvrditi poređenjem ako se translacijom šablona S u sliku I ona nalazi i unutar šablona S , dok u protivnom ta tačka pripada sceni W , odnosno okolini I^c , što se može matematički iskazati na sledeći način:

$$S_x \subseteq I ; (W - S_x) \subseteq I^c$$

Primena sistema prepoznavanja u upravljanju robotom

Klasifikacija upravljanja robotom na bazi sistema prepoznavanja se svodi na osnovna dva vida:

1. upravljanje na osnovu karakteristika koje su izdvojene iz slike i koje su korišćene za estimaciju (ocenjivanje) položaja cilja (objekta, ivice i sl.) u odnosu na kameru ("*position-based visual servo control*");
2. upravljanje na osnovu greške koja je u direktnoj korelaciji sa parametrima slike, što je u suprotnosti sa prethodnim vidom upravljanja koji definiše grešku u koordinatnom prostoru tehnološkog zadatka robota ("*image-based visual servoing*").