

1. Na koji način je moguće detektovati ivice na slici koja poseduje šum korišćenjem najmanje dva operatora?

Primenom box filtera za uklanjanje šuma i sobel filtera za ekstrakciju ivica ili Gausovog filtera za uklanjanje šuma i sobel filtera za ekstrakciju ivica.

2. Koja je razlika između Gausovog šuma i šuma „so i biber“ i koji filtri se koriste za njihovo uklanjanje iz slika?

Gausov šum predstavlja šum na slici koji prati normalnu (Gausovu) distribuciju i odgovara višim frekvencijama na slici, dok šum „so i biber“ predstavlja oštre promene signala koje se manifestuju kao crni ili beli pikseli na slici.

Za uklanjanje Gausovog šuma se može koristiti box filter ili Gausov filter, dok se za uklanjanje „so i biber“ šuma koristi median filter.

3. Koji filter se koristi za pooštavanje slika i da li je bitno da li se koristi konvolucija ili korelacija?

Koristi se filter sa slike koji akcentira razlike u odnosu na lokalnu srednju vrednost. Nije bitno da li se koristi korelacija ili konvolucija zato što su to identične operacije kada je filter simetričan.

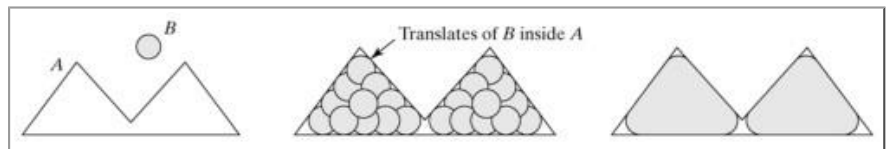
0	0	0
0	2	0
0	0	0

$$- \frac{1}{9}$$

1	1	1
1	1	1
1	1	1

4. Na koji način se morfološke operacije otvaranja i zatvaranja definišu preko dilacije i erozije? Grafički ilustrovati.

Otvoravanje je kombinovana operacija erozije nakon koje se vrši dilacija, pri čemu se koristi isti strukturni element. Otvoravanje A elementom B predstavlja uniju svih položaja B koji ulaze u A.

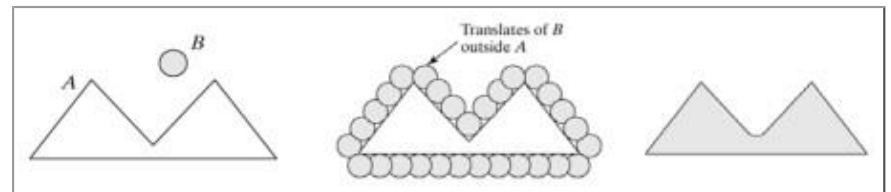


Zatvaranje je kombinovana operacija dilacije nakon koje se vrši erozija, pri čemu se koristi isti strukturni element.

Zatvaranje A elementom B

predstavlja komplement unije svih translacija B tako da se ne preklapa sa A.

Otvoravanje i zatvaranje su idempotentne operacije - njihova višestruka primena nema efekta.

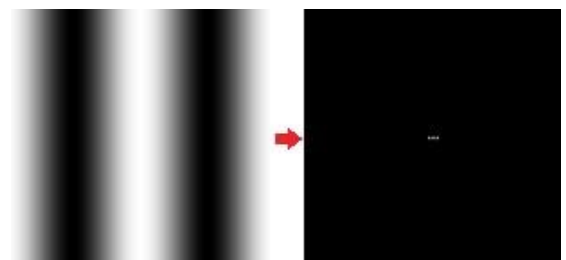


5. Koji se efekat postiže korišćenjem većeg strukturnog elementa kada se obavlja morfološka operacija zatvaranja u odnosu na manji? Grafički ilustrovati.

Sve praznine na slici veće od manjeg strukturnog elementa i svi oštri uglovi se popunjavaju pri korišćenju većeg strukturnog elementa. Grafička ilustracija - slika u pitanju iznad sa većim strukturnim elementom bi imala mnogo veći prevoj ukoliko bi se koristio veći krug kao strukturni element.

6. Kako izgleda izlaz Furijeove transformacije kada slika poseduje horizontalne sinusoidalne promene intenziteta? Skicirati i objasniti.

Postoje horizontalno raspoređene tačke od kojih svaka odgovara amplitudi određene frekvencije na slici. S obzirom na to da su promene sinusoidalne, tada one sadrže samo nekoliko niskih frekvencija, što se grafički prikazuje kao nekoliko tačaka na horizontalnoj osi. Tačke su bele jer su amplitude svih frekvencija maksimalne (jednake 1).

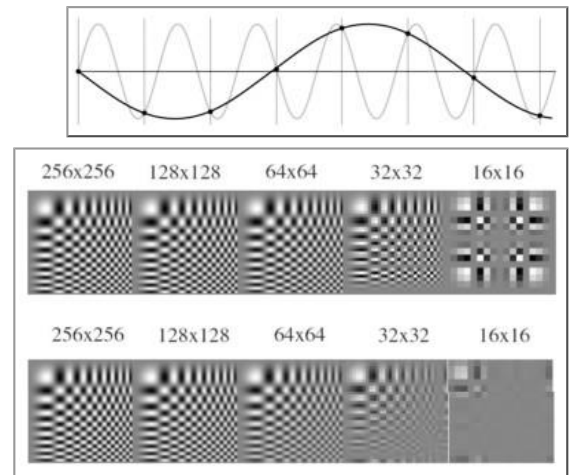


7. Zbog čega konvolucija slike sa Gausovim filtrom smanjuje visoke frekvencije?

Zbog toga što je konvolucija u prostornom domenu ekvivalentna množenju u frekventnom domenu, a Gausov filter u frekventnom domenu znatno smanjuje ili uklanja one vrednosti koje odgovaraju višim frekvencijama.

8. Kako se manifestuje „aliasing“ i šta ga uzrokuje? Grafički ilustrovati. Koja tehnika se primenjuje da bi se izbegao ovaj efekat?

Aliasing se manifestuje iskrzanim ivicama, deformisanim teksturama ili pojavama kao što su prividno okretanje točkova u suprotnom smeru ili čudan izgled prugastih košulja na televiziji. Aliasing je uzrokovan smplovanjem signala sa frekvencijom manjom od $2 \cdot f_{max}$, tako da rezultujući signal nema isti izgled kada se rekonstruiše. Da bi se izbegao aliasing možemo ili smplovati sa frekvencijom većom od $2 \cdot f_{max}$, ili odbaciti frekvencije signala više od polovine frekvencije smplovanja, što dovodi do gubitka informacija (za ovo se koristi low-pass filter). Aliasing se najčešće vidi pri smanjivanju slika ili smplovanju videa, bilo u prostornom (smanjenjem rezolucije) ili vremenskom domenu (smanjenjem broja frejmova u sekundi).



9. Opisati proces detekcije ivica korišćenjem Canny detektora?

Canny detektor ivica radi u četiri koraka:

- 1) Filtriranje slike izvodima Gausove funkcije - Nalaze se grube (višepikselne) ivice slike.
- 2) Nalaženje jačine i orijentacije gradijenata.
- 3) Potiskivanje ne-maksimuma - Tanje se višepikselne ivice sve dok se ne dođe do debljine od jednog piksela. Vršiti se proveru da li je piksel lokalni maksimum duž pravca gradijenta. Zove se još i tanjenje (thinning).
- 4) Povezivanje i pragovanje (histerezis): Definišu se viši i niži prag. Viši se koristi da se započnu, a niži da se nastave ivice.

10. Koji filter (funkciju) je moguće iskoristiti za izdvajanje ivica na slici?

Moguće je koristiti parcijalne izvode slike, sobel operator, izvod 2D Gausovog filtera ili laplasijan Gausove funkcije.

11. Standardna Hough-ova transformacija se zasniva na glasanju za određene parametarske oblike. Zbog čega se primenjuje ova tehnika i zašto daje rezultate?

Hough-ova transformacija se primenjuje zbog toga što je suviše zahtevno ispitati sve moguće modele ili kombinacije fičera poklapajući model sa svakim mogućim podskupom. Ovakav pristup daje rezultate zato što je doprinos šuma i clutter-a u glasanja beznačajan, tj. njihovi glasovi nisu konzistentni sa većinom „pravih“ glasova. Takođe, ukoliko neki fičeri nisu detektovani, model se izgrađuje na osnovu ostalih fičera.

12. Ukoliko je potrebno da se nađu loptice za bilijar na slici bilijarskog stola, da li je prikladnije koristiti Hough-ovu transformaciju ili RANSAC algoritam? Obrazložiti odgovor.

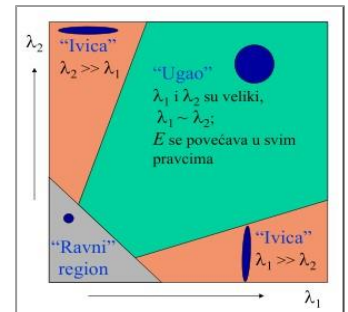
Prikladnije je koristiti Hough-ovu transformaciju, s obzirom na to da su bilijarske loptice krugovi istog poluprečnika koji se mogu predstaviti istim modelom. Hough-ova transformacija omogućava detektovanje više instanci istog modela u jednom prolazu, što je čini efikasnijom od RANSAC algoritma.

- 13. Kada kažemo da deskriptori ključnih tačaka trebaju da budu i „invarijantni“ i „prepoznatljivi“, na šta se odnosi „invarijantnost“ i zašto je ona neophodna?**

Invarijantnost se odnosi na nezavisnost prepoznatljivosti fičera od geometrijskih (skaliranje/rotacija) i fotometrijskih (promena osvetljenja) transformacija i neophodna je kako bi se isti fičer mogao prepoznati na više različitih slika, bez obzira na njegov položaj ili nivo osvetljenosti.

- 14. Harisov detektor uglova koristi matricu momenta drugog reda. Grafički ilustrovati i objasniti kakve sopstvene vrednosti te matrice (λ_1 i λ_2) ukazuju na dobar fičer za praćenje.**

Na dobar fičer za praćenje ukazuju velike vrednosti λ_1 i λ_2 koje su međusobno slične, $\lambda_1 \sim \lambda_2$. Iz ovih vrednosti se može izvesti „uglovitost“ - mera dobrote fičera. Uglovitost se izražava kao $C = \lambda_1 \cdot \lambda_2 - \alpha \cdot (\lambda_1 + \lambda_2)^2$. Fičer je dobar (ugao) ukoliko je $C > 0$, ivica ukoliko je $C < 0$, a ravan za male vrednosti $|C|$.



- 15. Na koji način je moguće fičer detektor (npr. SIFT) učiniti invarijantnim na promene u osvetljenju?**

SIFT detektor se čini invarijantnim na promene u osvetljenju tako što se sve vrednosti u vektoru normalizuju na magnitudu 1.0, zatim se odseku svi gradijenti sa magnitudom većom od 0.2 radi postizanja robustnosti na „nelinearne efekte osvetljenja“, a rezultujuće vrednosti se ponovo normalizuju.

- 16. Da li su Harisovi uglovi invarijantni na rotaciju? Objasni odgovor. Kakva je situacija po ovom pitanju kod SIFT-a?**

Harisovi uglovi su invarijantni na rotaciju jer se pri rotaciji slike rotiraju i elipse matrice momenta drugog reda, ali zadržavaju isti oblik i samim tim iste sopstvene vrednosti (na osnovu kojih se vrši odlučivanje).

SIFT je invarijantan ne samo na rotaciju, već i na translaciju, skaliranje i druge fotometrijske transformacije.

- 17. Na koji način se vrši formiranje vektora koji predstavlja SIFT deskriptor fičera?**

Vektor se računa na osnovu rotirane i skalirane verzije prozora, a na osnovu orijentacije i razmere tačke. Koriste se težinski gradijenti na bazi Gausove funkcije koja opada ka krajevima. Za svakih 4x4 gradijenata u vektoru se formira 8-elementni histogram njihovih orijentacija, pa se 4x4 niz histograma orijentacija gradijenata može predstaviti kao 4x4 piksela. Rezultujući vektor ima ukupno 128 dimenzija (4x4 histograma sa po 8 orijentacija). Zatim se vrši umekšavanje - magnitude se težinski skaliraju na osnovu udaljenosti od centra (Gausova funkcija). Na kraju se vrši redukcija efekta osvetljenja, tako što se vrednosti u vektoru normalizuju na magnitudu 1.0, zatim se odseku svi gradijenti sa magnitudom većom od 0.2 radi postizanja robustnosti na „nelinearne efekte osvetljenja“, a rezultujuće vrednosti se ponovo normalizuju.

- 18. Na koji način se RANSAC algoritam koristi za spajanje slika u panoramu? Skicirati i opisati postupak.**

Potrebno je naći homografiju pomoću RANSAC algoritma da bi se slike mogle spojiti:

- 1) Nasumično se biraju 4 para tačaka (p_i).
- 2) Izračunava se homografija H_k .
- 3) Određuju se konzistentni parovi tako da je $SSD(p_i', H_k \cdot p_i) < \epsilon$.
- 4) Sačuva se H_k ako je C_k najveći skup konzistentnih parova. Postupak se ponavlja više puta (vraćamo se na korak 1)).
- 5) Izračunava se H primenom metoda minimizacije kvadrata rastojanja koristeći uzorke iz skupa C_k .
- 6) Pronađena homografija se primenjuje na prvu sliku, čime se ona poklapa sa drugom.

19. Pri redukciji dimenzija korišćenjem PCA, prvo se vrši oduzimanje srednje vrednosti tačaka. Objasniti zašto?

Zato što je potrebno naći jedinični vektor \vec{u} u R^d prostoru takav da je varijansa po tom pravcu maksimalna. Srednja vrednost tačaka se oduzima jer se pri smanjenju dimenzija tačke karakterišu samo jednom od koordinata i srednjom vrednošću.

20. Na koji način je moguće iskoristiti PCA za detekciju lica na slici?

PCA se za detekciju lica koristi tako što se nalaze k vektora koji predstavljaju k pravaca maksimalne varijanse, pri čemu je pretpostavka da se većina slika lica nalazi u niskodimenzionalnom podprostoru određenom pomoću prvih k pravaca maksimalne varijanse, tzv. „prostoru lica“. Tada se sve moguće slike lica mogu predstaviti kao linearna kombinacija ovih vektora.

21. Objasniti razliku između generativnih i diskriminacionih metoda za klasifikaciju?

Generativni metod predstavlja korišćenje trening podataka da bi se izgradili reprezentativni modeli verovatnoća za svaku klasu ponaosob, dok diskriminacioni metod umesto nezavisnih modela klasa modeluje granicu odluke između klasa.

22. Opisati na koji način se vrši detekcija objekata na slici metodom klizajućeg prozora?

Prvo je potrebno obučiti klasifikator da bi se detekcija uopšte mogla vršiti:

- 1) Obezbediti trening podatke.
- 2) Definirati fičere.
- 3) Obučiti klasifikator.

Zatim se vrši detekcija metodom klizajućeg prozora:

- 1) Generišu se kandidati klizajućim prozorom.
- 2) Ocenjuju se kandidati pomoću klasifikatora.

23. Na koji način se kaskadni klasifikator koristi za detekciju lica na slici? Koja je njegova glavna prednost?

Ideja kaskadnog klasifikatora je u što bržoj detekciji da prozor nije lice. Formira se kaskada sa malom greškom u detekciji negativnih primera na početku. Prednost je u brzini - primenjuju se manje precizni, ali veoma brzi klasifikatori koji odmah eliminišu prozore koji očigledno ne sadrže lice.

24. Šta predstavljaju integralne slike i na koji način se one koriste kod detekcije lica na slici?

Integralna slika je slika kod koje vrednost u tački (x, y) predstavlja sumu svih piksela iznad i levo od te tačke. Zahvaljujući ovoj osobini, moguće je izračunati svaku sumu u konstantnom vremenu, a suma u okviru proizvoljnog pravougaonika na slici se računa pomoću samo tri operacije sabiranja/oduzimanja. Integralne slike se koriste za predstavljanje lokalne teksture preko fičera koji se brzo računaju. Koriste se pravougaoni filtri kod kojih se izlaz računa kao razlika između susednih regiona.

25. Na koji način je moguće iskoristiti SVM kad podaci nisu linearno separabilni?

Ako podaci nisu linearno separabilni, vrši se mapiranje podataka na prostor sa više dimenzija gde je trening skup separabilan. U originalnom prostoru se obezbeđuje nelinearna granica preko kernel funkcija umesto eksplicitnog računanja transformacije. Primeri kernel funkcija su:

linearna funkcija - $K(x_i, x_j) = x_i^T x_j$, Gausova RBF - $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2 \cdot \sigma^2}\right)$ i presek

histograma - $K(x_i, x_j) = \sum_k \min(x_i(k), x_j(k))$.

26. Opisati kako se metod histograma orijentisanih gradijenata (HOG) u kombinaciji sa linearnim SVM-om koristi za detekciju osoba na slici.

Svaka ćelija u ulaznom prozoru se mapira na histogram gradijenata po orijentaciji (HOG). Obučava se linearni SVM korišćenjem trening skupa sa osobama i ne-osobama.

27. Na koji način se metod „vreća vizuelnih reči“ koristi za pretragu slika?

Primenjuju se sledeći koraci:

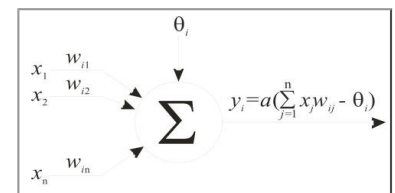
- 1) Nalaženje vizuelnih reči u slici za pretragu.
- 2) Nalaženje relevantnih slika iz baze korišćenjem inverznog indeksa.
- 3) Uređivanje rezultata po broju poklopljenih reči.

28. Objasniti upotrebu inverznog indeksa na vizuelne reči?

Vizuelne reči se čuvaju u rečniku (bazi). Za svaku vizuelnu reč se čuva niz referenci na slike koje je sadrže. Obrada nove slike se sastoji od nalaženja vizuelnih reči koje ona poseduje, a onda se na osnovu referenci pristupa slikama za svaku vizuelnu reč koju slika poseduje.

29. Opisati model veštačkog neurona. Kako izgleda ReLU aktivaciona funkcija? Zbog čega je ovaj tip aktivacione funkcije veoma zastupljen u dubokim neuronskim mrežama?

Veštački neuron predstavlja jednostavni procesni element koji obavlja relativno jednostavnu matematičku funkciju. ReLU aktivaciona funkcija je $f(x) = \max(0, x)$. Prednost ReLU je jednostavnost izračunavanja, što znači značajno manje vreme za trening i evaluaciju.



30. Šta predstavlja overfitting problem kod obučavanja veštačkih neuronskih mreža i kojetehnike se koriste kako bi se izbegao?

Overfitting znači da mreža „previše dobro“ modeluje trening podatke, tj. da su šumovi zapamćeni u modelu i negativno utiču na svojstvo modela da generalizuje. Tehnike koje se koriste su glasanje i dropout.

Regularizacija pokušava da reši problem overfitting-a, čak i kada imamo previše neurona u mreži. Regularizacija polazi od toga da velike težine proizvode velike promene u ponašanju za male promene na ulazima, pa je potrebno povećati cenu za velike težine.

Kod glasanja se vrši dodela slučajnih težina u više neuronskih mreža, tako da se konačan izlaz određuje glasanjem nakon obučavanja svih mreža. Dropout tehnika u svakom koraku obučavanja nasumično ignoriše određeni broj neurona u mreži, čime se postiže manja zavisnost od prethodnih izlaza i forsira se mreža da uči robustnije fičere.

31. Zbog čega su se duboke konvolucione neuronske mreže pokazale jako efikasnim u rešavanju različitih problema računarskog vida?

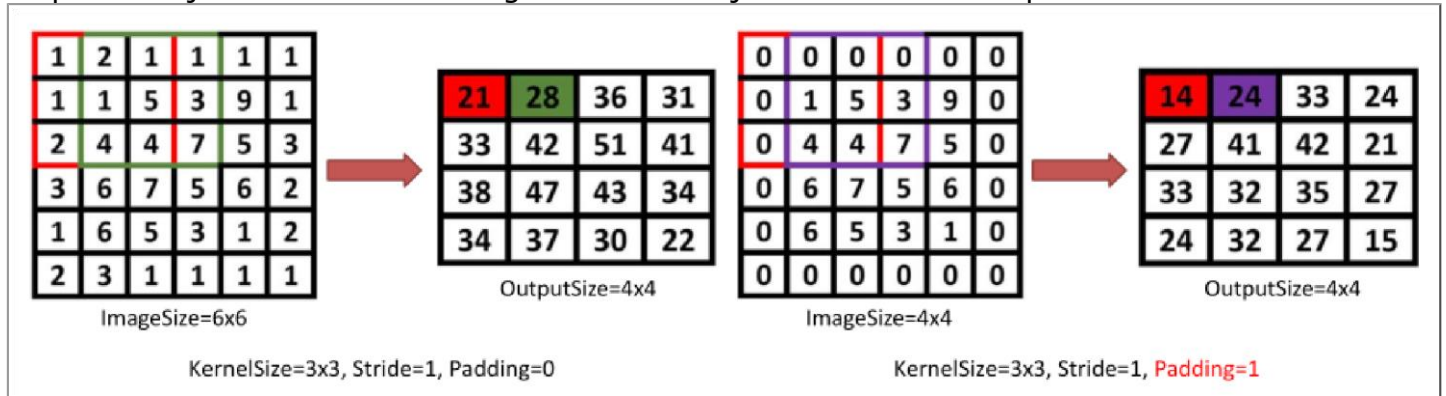
Duboke konvolucione neuronske mreže su vrlo efikasne zbog toga što je moguće paralelizovati ih korišćenjem grafičkih procesora, čime se znatno povećava brzina izračunavanja. Takođe, njihova konfiguracija je vrlo jednostavna, što znatno olakšava testiranje različitih struktura mreže, što je naročito bitno jer je potrebno napraviti neuronsku mrežu koja prepoznaje na hiljade različitih uzoraka, a nije uvek očigledna veza između strukture mreže i uzoraka koje ona prepoznaje.

32. Opisati strukturu tipične konvolucione neuronske mreže.

Konvolucionarna neuronska mreža (KNM) predstavlja višeslojnu veštačku neuronsku mrežu koja sadrži jedan ili više konvolucionih slojeva. Konvolucionni sloj se posmatra kao ekstraktor fičera iz ulazne slike. Iza konvolucionih slojeva se obično nalaze pooling slojevi koji vrše smanjivanje slike, čime se postiže invarijantnost na pomeranja i distorzije. Na kraju mreže se obično nalazi potpuno povezani sloj koji vrši klasifikaciju slike na osnovu prethodno pripremljenih fičera.

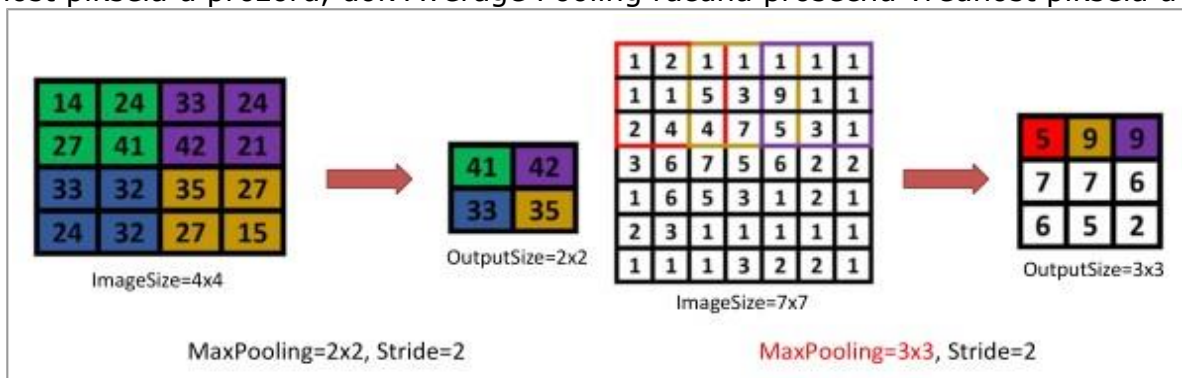
33. Koja je uloga i na koji način se formira izlaz konvolucionog sloja kod konvolucionih neuronskih mreža? Ilustrovati na primeru.

Konvolucioni sloj je najvažniji sloj konvolucione neuronske mreže (po njemu je dobila ime). Njegova uloga je ekstrakcija fičera kao što su ivice i uglovi iz ulazne slike. Izlaz se formira kao matrica dimenzije $outputSize = ((imageSize + 2 \cdot padding) - kernelSize) / stride + 1$, gde je $imageSize$ veličina ulazne slike, $padding$ broj nula koji se dodaje oko slike, $kernelSize$ veličina kernela (filtra) i $stride$ pomeraj kernela u svakom koraku. Svaki element izlazne matrice predstavlja sumu vrednosti originalne slike koje se nalaze unutar prozora.

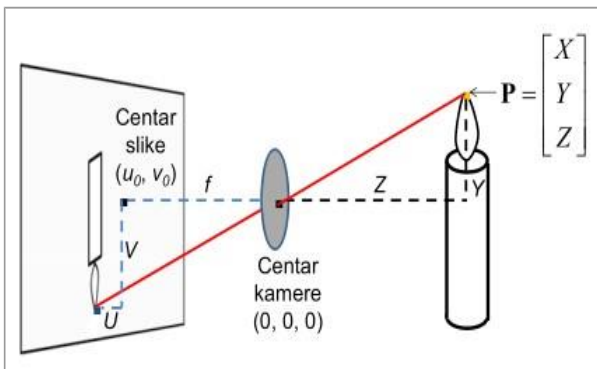


34. Koja je uloga i na koji način se formira izlaz pooling sloja kod konvolucionih neuronskih mreža? Ilustrovati na primeru.

Uloga pooling sloja konvolucione neuronske mreže je smanjivanje rezolucije slike (obično duplo), čime se postiže invarijantnost na pomeranja i distorzije, kao i detekcija fičera pri različitim razmerama. Prema načinu na koji se vrši smanjivanje, razlikujemo Max Pooling i Average Pooling slojeve. Max Pooling računa piksel rezultujuće slike kao maksimalnu vrednost piksela u prozoru, dok Average Pooling računa prosečnu vrednost piksela u prozoru.

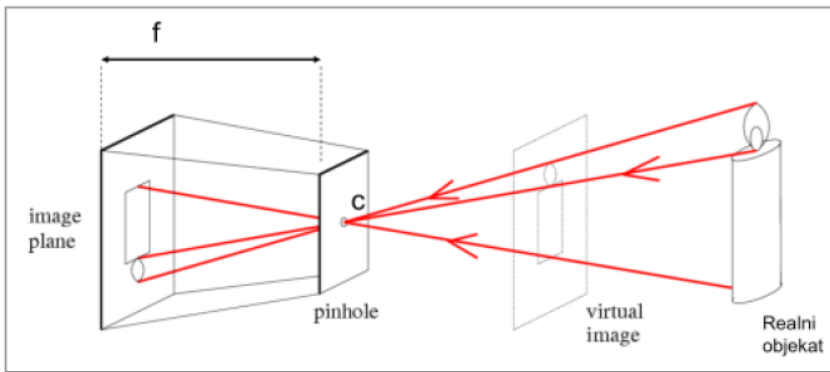


35. Gde se 3D tačka (x, y, z) u prostoru kamere prilikom perspektivne projekcije preslikava u prostor slike? Skicirati i napisati izraze za (u, v) .



Koordinate projekcije su $p = \begin{bmatrix} U \\ V \end{bmatrix}$, gde je $U = -X \cdot \frac{f}{Z}$ i $V = -Y \cdot \frac{f}{Z}$. f na slici predstavlja fokalnu dužinu kamere, tj. rastojanje centra kamere od ravni slike.

36. Grafički ilustrovati i opisati pinhole model kamere. Objasniti koji parametri određuju unutrašnju i spoljašnju matricu.



Pinhole kamera je jednostavna kamera bez objektivu, koja ima mali otvor koji propušta svetlost sa scene i na taj način se projektuje obrnuta slika na suprotnoj strani, što je poznato kao camera obscura efekat.

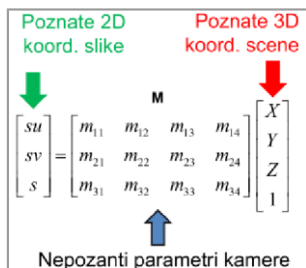
Koordinate na slici ($x = [u \ v \ 1]$) se dobijaju kao proizvod unutrašnje (intrinzične) matrice K , spoljašnje (ekstrinzične)

matrice i svetskih koordinata centra kamere ($X = [X \ Y \ Z \ 1]$). Unutrašnja matrica je 3×3 gornje trougaona matrica koja transformiše 3D koordinate kamere u 2D homogene koordinate slike. f_x i f_y predstavljaju fokalnu dužinu - rastojanje između pinhole-a i ravni slike. x_0 i y_0 predstavljaju offset „glavne“ tačke - „glavna“ osa kamere je normalna na ravan slike i prolazi kroz pinhole, njen presek sa ravni slike je „glavna“ tačka, tako da je offset rastojanje početka slike od ove tačke. Spoljašnja matrica ($[R \ t]$) lokaciju i pravac kamere u svetu (translaciju i rotaciju). R je 3×3 rotaciona matrica, čije su kolone pravci osa izraženi preko koordinata kamere, a t translacioni vektor, koji daje poziciju svetskog koordinatnog početka preko koordinata kamere.

37. Zbog čega paralelne linije konvergiraju u jednoj tački na slici?

Dužina, površina i uglovi se gube pri projektivnoj geometriji. U stvarnom svetu deluje da prave paralelne linije konvergiraju u linije nestajanja, tj. horizont, što se na slici predstavlja kao konvergencija ka jednoj tački.

38. Jedan način da se odrede parametri modela kamere je da se odrede parovi 3D tačaka usceni i 2D tačaka na slici. Koliko je minimalno tačaka potrebno za određivanje parametara kamere? Ukoliko imamo veći broj tačaka, na koji način se nalaze parametri kamere?



Treba naći tačke na slici koje odgovaraju 3D tačkama u sceni i matricu M koja rezultuje najmanjom srednjom kvadratnom greškom između dobijenih (p') i poznatih (p) 2D tačaka. Matrica M je dimenzija 3×4 , što daje 12 nepoznatih, a 11 stepeni slobode. Potrebna je jedna jednačina po nepoznatoj, tj. 5.5 uparenih tačaka. Ukoliko ih ima više, znači da ima više rešenja za M . U praksi se koristi više od 6 parova zbog robustnosti na greške u izmerenim koordinatnim tačkama.

39. Šta je afina transformacija i koliko parova tačaka je potrebno za njeno određivanje? Grafički ilustrovati.

Afina transformacija je kombinacija translacije, skaliranja, rotacije i smicanja. Matrica preslikavanja je dimenzija 2×3 , što daje 6 nepoznatih, a 5 stepeni slobode. Potrebna je jedna jednačina po nepoznatoj, tj. 2.5 uparenih tačaka. U praksi se koristi 3 ili više parova tačaka.

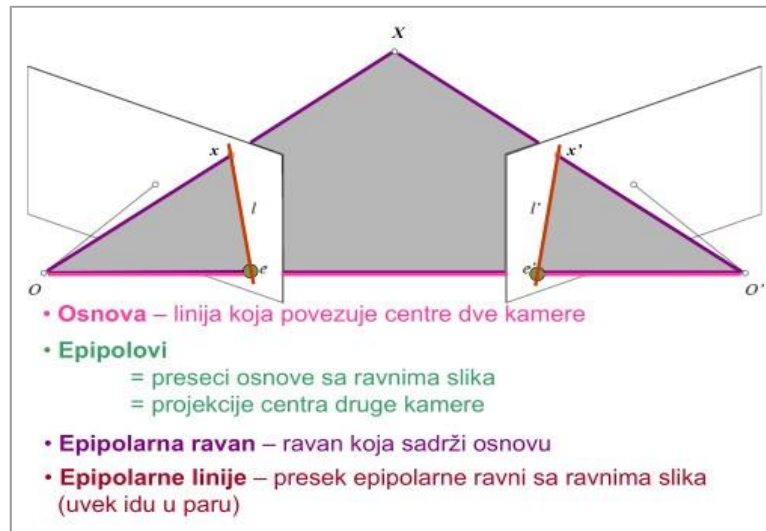
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

40. Šta je homografija i koliko parova tačaka je potrebno za njeno određivanje? Grafički ilustrovati.

Homografija je projektivna transformacija, tj. transformacija koja slika jedan projektivni prostor u drugi. Matrica homografije je dimenzija 3×3 , što daje 9 nepoznatih, a 8 stepeni slobode. Potrebna je jedna jednačina po nepoznatoj, tj. 4 uparenih tačaka, koliko se najčešće koristi u praksi.

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

41. Ukoliko imamo dve kamere i tačku X u prostoru, prikazati položaj epipolarnih linija naravnima slika.



42. Šta je epipol? Gde se nalaze epipolovi kod kretanja duž ose kamere, a gde kod kretanja upravno na osu kamere? Grafički ilustrovati.

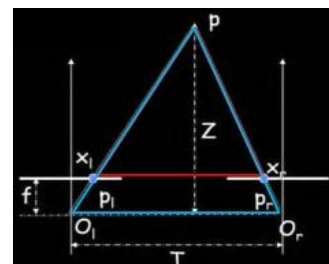
Epipol je projekcija centra jedne kamere na ravan slike druge. Posmatramo sliku iznad. Na slici su epipolovi obeleženi sa e i e' . U slučaju kretanja duž ose kamere, epipolovi ne menjaju svoje pozicije, dok se u slučaju kretanja upravno na osu kamere epipolovi šetaju po pravama l , odnosno l' .

43. Koja je razlika između esencijalne i fundamentalne matrice?

Esencijalna matrica sadrži informacije o translaciji i rotaciji, koje opisuju lokaciju druge kamere u odnosu na prvu u globalnim koordinatama. Fundamentalna matrica ima iste informacije, ali transformiše globalne koordinate u pixel koordinate koje su nam potrebne i omogućava da odredimo vezu između kamere za koje nisu poznati unutrašnji (intrinsic) parametri, tj. koje nisu kalibrisane. Drugim rečima, fundamentalna matrica mapira tačku u jednoj slici na epipolarnu liniju druge slike. Ovo se računa iz istih tačaka sa obe slike. Neophodno je najmanje 8 tačaka da bi se našla fundamentalna matrica, a poželjno je da ih ima više.

44. Koji je odnos između dispariteta piksela u kalibrisanom stereo sistemu i daljine odgovarajuće tačke od kamere? Grafički ilustrovati.

Pod pretpostavkom da su optičke ose paralelne i da su poznati parametri kamere (kamere su kalibrisane), iz sličnosti trouglova (p_l, P, p_r) i (O_l, P, O_r) sledi $\frac{T + x_l - x_r}{Y - f} = \frac{T}{Z}$, iz čega se dobija izraz $Z = f \cdot \frac{T}{x_l - x_r}$, gde je $x_l - x_r$ disparitet piksela.



45. Koje ograničenje se koristi za nalaženje uparenih piksela na stereo slikama? Grafički ilustrovati.

Ograničenje koje se nameće je da je determinanta fundamentalne matrice jednaka nuli. Koristimo RANSAC algoritam nasumično birajući po 8 parova tačaka, a proces ponavljamo sve dok se ne dobiju parametri modela sa visokim stepenom poverenja. Zadržavaju se samo parovi tačaka koji su „saglasni“ sa „najboljom“ fundamentalnom matricom.

46. Lukas-Kanade metod za nalaženje optičkog toka se zasniva na korišćenju gradijenata na statičkoj slici.

Sa kojim pretpostavkama je izveden ovaj metod i koje tehnike se primenjuju kako bi se metod primenio na situacije kada ove pretpostavke ne važe?

Lukas-Kanade metod se oslanja na dve pretpostavke:

- 1) Boja je konstantna - piksel $I(x, y, t)$ izgleda isto kao i piksel $I(x, y, t + 1)$.
- 2) Mali pokret - pomeraj između piksela je mali.

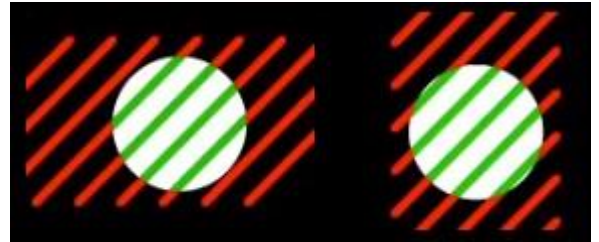
Ako nije zadovoljena prva pretpostavka, vrši se pretraga okoline normalizovane korelacije ili praćenje fičera. Ako nije zadovoljena druga pretpostavka, prvo se smanjuje rezolucija pa se tek onda primenjuje Lukas-Kanade metod.

47. Objasniti problem aperture kod određivanja pravca kretanja. Ilustrovati primerima.

Problem aperture je problem neodređenosti pravca kretanja. Ovaj problem se javlja jer senzor kretanja ima konačan prostor razmatranja "otvora". U okviru ovog otvora se različiti fizički pokreti ne razlikuju.

Ovde, na primer, niz linija koje se kreću s desna na levo rezultuju istim prostornim rasporedom kao da su se kretale odozdo naviše. Problem aperture podrazumeva da će

neuroni osetljivi na pokret uvek reagovati na konturu koja im je u prostoru razmatranja, nezavisno od njene istinske orijentacije, dok je smer kretanja u skladu sa preferiranim pravcem.



48. Koja je veza između Lukas-Kanade metoda za određivanje optičkog toka i Harisovog detektora uglova?

Parametri jednačine koji se dobija Lukas-Kanade metodom su kriterijumi za Harisonov detektor uglova. Ti parametri su magnituda gradijenata i odnos sopstvenih vrednosti matrica. Na osnovu tih parametara, odnosno Harisovog detektora, može se zaključiti da li se radi o regionu bez tekstone (mala magnituda gradijenata, male sopstvene vrednosti), ivici (veliki gradijenti, prva sopstvena vrednost velika a druga mala) ili visoko teksturisanim regionu (velika magnituda i gradijenti se razlikuju, prva sopstvena vrednost velika a druga mala).

49. Objasniti princip rada dubinske kamere kod Kinect-a.

Kinect poseduje senzor sa infracrvenim laserom, koji prikuplja video podatke u 3D-u pri bilo kakvom ambijentalnom svetlu. Dobija se dubinska mapa. Omogućeno je 3D snimanje pokreta na osnovu senzora prikačenih na ljudsko telo i određivanje poze na osnovu slika celog tela.

Svaki piksel poseduje informaciju o daljini od kamere umesto klasične RGB boje. Problem poze je preveden u problem klasifikacije (31 klasa za delove tela). Zbog velikih različitosti pri svakom objektu ili čoveku koji se snima, kreiraju se različiti modeli i varijacije.

Klasifikacija se pravi na osnovu stabla odluke.

50. Objasniti metod srednjeg pomeraja (mean shift). Na koji način se on koristi za segmentaciju delova tela kod Kinect-a?

Mean shift algoritam pravi predikciju položaja zglobova veoma brzo, tako što nalazi centre klastera na osnovu zadatog skupa tačaka. Postupak je sledeći:

- 1) Izabrati veličinu i funkciju kernela.
- 2) Za svaku tačku:
 - a) Postaviti prozor na tu tačku.
 - b) Sračunati srednju vrednost (centar mase) tačaka koje upadaju u prozor.
 - c) Pomeriti prozor pretrage na novoodređeni centar mase.
 - d) Ponavljati korake b) i c) sve dok ne dođe do konvergencije.
- 3) Dodeliti tačke koje dolaze do istih centara istim klasterima.