

LABORATORIJSKA VAJA 6 – ŠIVANJE PANORAMSKIH SLIK

IDEJA VAJE

Namen vaje je seznaniti študente z znim od področij uporabe slikovnih tehnologij. V okviru te vaje se študenti srečajo s postopki šivanja slik, s katerimi je mogoče sestavljati večje panoramske slike iz večjega števila slik iste scene. Študenti s v okviru vaje seznanijo s karakteristikami postopkov šivanja slik in primer takšnega postopa tudi udejanijo. Pri tem študenti podrobneje spoznajo

- Lokalne deskriptorje SIFT oz. SURF,
- Postopke za iskanje ujemanja lokalnih deskriptorjev,
- Postopek RANSAC,
- Perspektivne preslikave ter pristope k njihovi določitvi, in
- Geometrijske transformacije slik

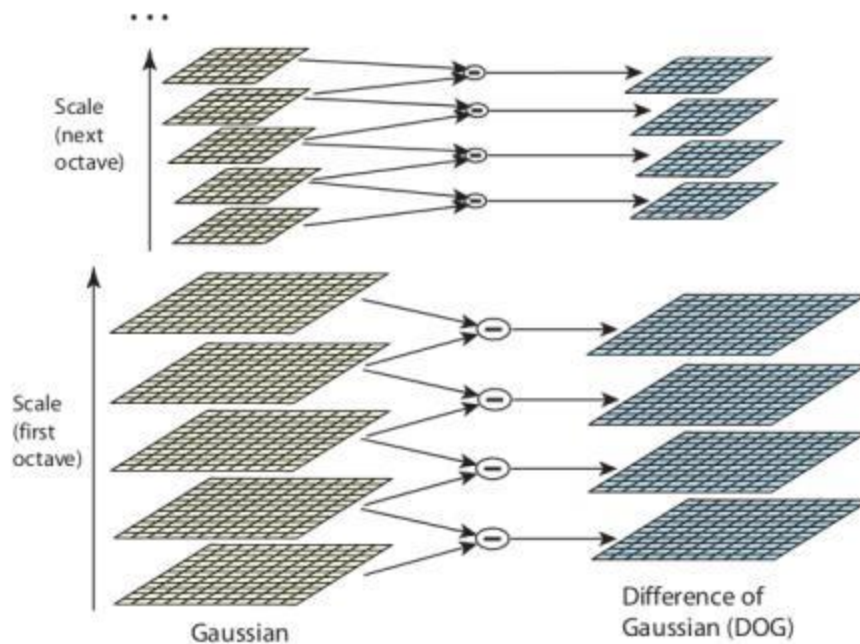
Za izvedbo vaje pridejo v poštev funkcije, ki so del knjižnice openCV. Ker je bil v zadnjih letih postopek za izračun SIFT deskriptorjev v ZDA zaščiten s patentom, za izračun teh deskriptorjev potrebujemo starejšo verzijo knjižnice openCV, ki se namesti z ukazom

```
pip uninstall opencv-python  
pip install -U opencv-contrib-python==3.4.2.16
```

TEORETIČNO OZADJE

Lokalni deskriptorji

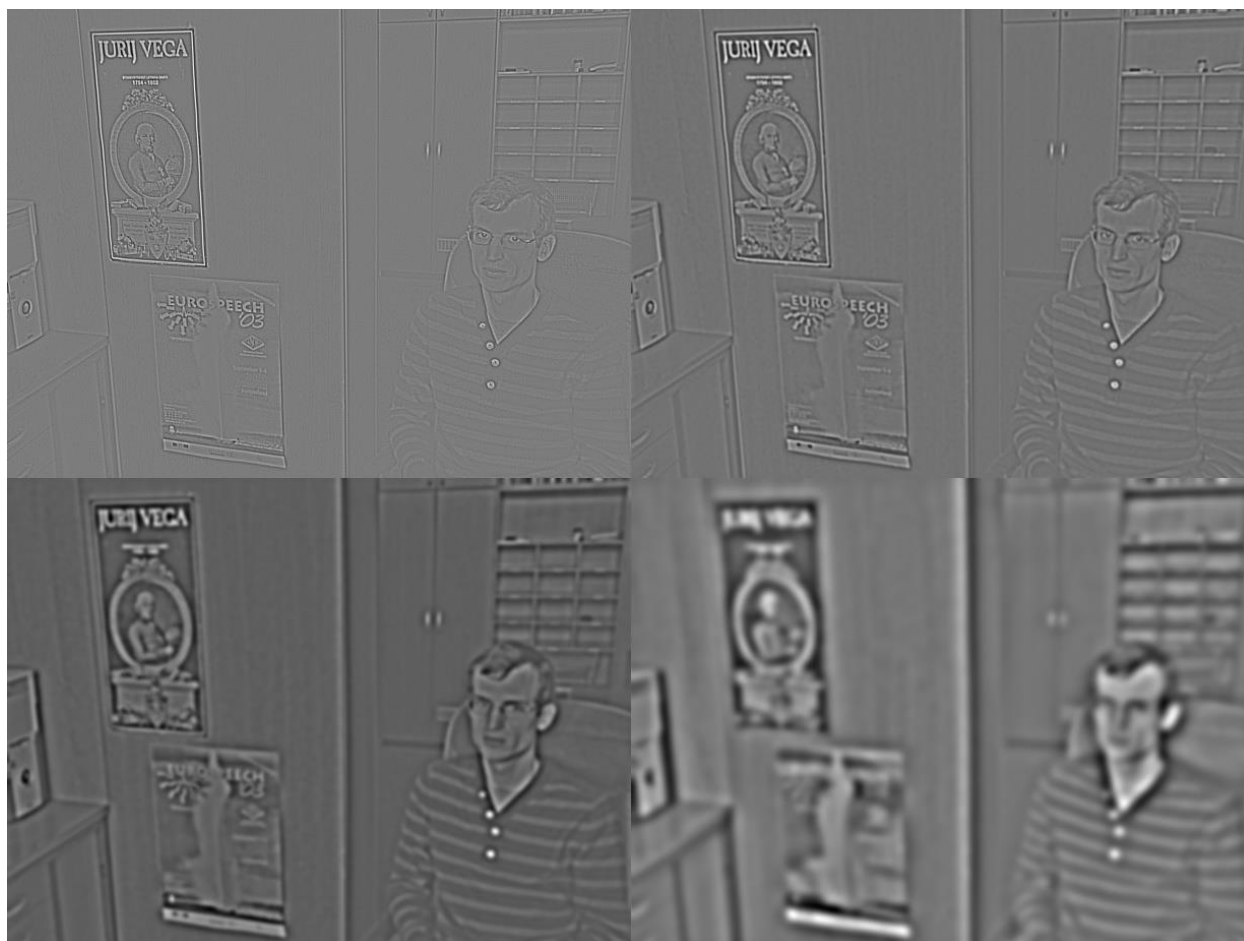
Prvi korak za izračun lokalnih deskriptorjev sestoji iz pretvorbe vhodne slike v t.i. prostor ločljivosti (angl. *scale space*). Dvodimenzionalno (npr. sivinsko) sliko $I(x, y)$ tu pretvorimo v trirazsežno predstavitev $I(x, y, \sigma)$, ki nam poda vrednost piksla s koordinatami (x, y) slike, dobljene s filtriranjem originalne slike z Gaussovim filtrom s standardno deviacijo enako σ . Po pretvorbi v prostor ločljivosti lahko definiramo sliko razlike Gaussov (angl. *Difference of Gaussians*, DoG), kot $DoG(\sigma_1, \sigma_2) = I(x, y, \sigma_1) - I(x, y, \sigma_2)$. Postopek ponavljamo in sliko rekurzivno zmanjšujemo, tako, da na ta način pridelamo piramido predstavitev, kot prikazuje spodnja slika:



Postopek lahko prikažemo grafično s pretvorbo spodnje slike v prostor ločljivosti:



Sliko pretvorimo v prostor skal s $\sigma_1, \dots, \sigma_5 = 0.5, 1, 2, 4, 8$, ter z odštevanjem sosednjih ločljivosti dobimo sledeče razlike Gaussov:



Iz slik je vidno, da postopek izpeljave DoG predstavitev deluje kot pasovno prepustni frekvenčni filter, kjer z nastavljanjem zgornje in spodnje vrednosti σ izbiramo frekvenčni pas, ki ga želimo iz slike izolirati.

Po določitvi DoG predstavitve iščemo lokalne ekstreme v prostoru ločljivosti. Za vsak piksel v trirazsežni predstavitvi $I(x, y, \sigma)$ trdimo, da predstavlja lokalno ekstremno točko, če je njegova vrednost bodisi minimum, bodisi maksimum v lokalni $3 \times 3 \times 3$ okolici. Dobljena množica lokalnih ekstremov nam predstavlja kandidate za značilne točke na sliki. V nadaljnjih korakih z upravljanjem sivinske slike in DoG slik zmanjšamo množico kandidatov in ohranimo samo značilne točke na mestih z visokim lokalnim kontrastom ter v kotih objektov (zavračanje robov).

Po določitvi lokacij značilnih točk po opisanem postopku za vsako značilno točko določimo njeno ločljivost in orientacijo. Ločljivost je določena s položajem pripadajoče ekstremne točke v prostoru ločljivosti, orientacija pa glede na histogram smeri gradientov v lokalni okolici pripadajoče slike ustrezne ločljivosti. Okolico značilne točke velikosti 16×16 pikslov nato razdelimo v 16 področij velikosti 4×4 pikslov in za vsako področje določimo histogram orientacij gradientov, ki ga kvantiziramo na 8 nivojev. Vse orientacije nato normiramo na poprej določeno orientacijo značilne točke. Deskriptor značilne točke tako sestavimo iz vseh 16 kvantiziranih

histogramov, ki imajo skupaj 128 elementov. Po tem postopku vsaki značilni točki na sliki pripišemo deskriptor v obliki vektorja 128 števil.

Ujemanje lokalnih deskriptorjev

Za šivanje dveh slik v panorami potrebujemo poleg značilnih točk in njihovih deskriptorjev tudi informacijo o tem, kateri pari značilnih točk oz. deskriptorjev na slikah predstavljajo isti objekt. Pare iščemo z iskanjem najmanjših razdalj, pri čemer zahtevamo, da je razmerje med prvim in drugim najbližjim deskriptorjem večje od izbranega praga.

Določitev transformacijske matrike

Z izračunanimi pari značilnih točk lahko določimo transformacijsko matriko, ki nam določa perspektivno projekcijo druge slike glede na prvo. To lahko storimo npr. s postopkom RANSAC, tako, da večkrat zaporedoma vzorčimo ustrezno število parov točk (t.j. 4), na podlagi vzorčenih parov izračunamo transformacijo in preverimo, koliko ostalim parom točk ta transformacija ustreza.

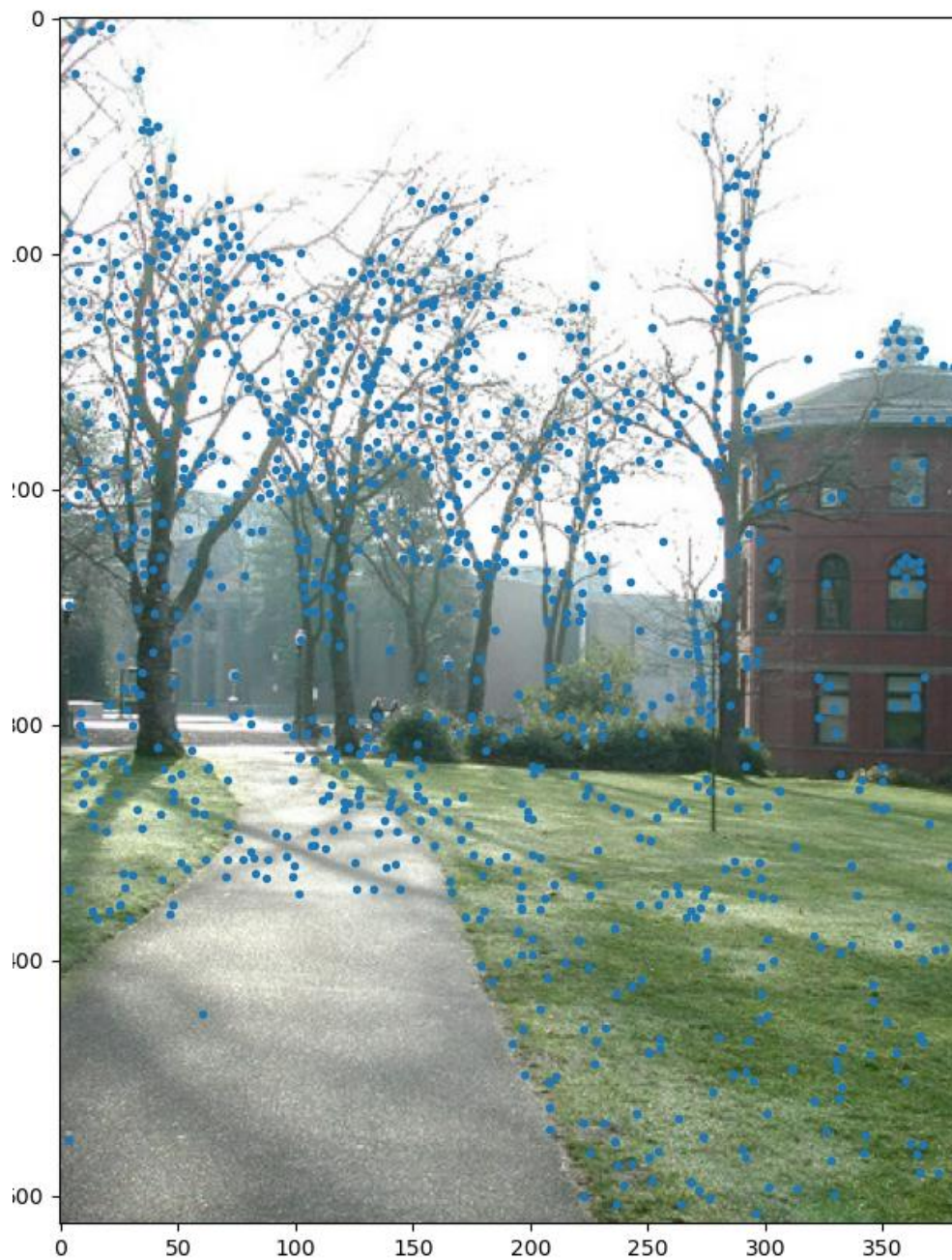
IZVEDBA

Naloga 1 (1 točka)

V mapi „slike“ se nahaja 9 parov slik, ki bi jih radi šivali. Poskušajte slike sešiti v enem izmed urejevalnikov slik (npr. GIMP, Photoshop), z uporabo osnovnih operacij za transformacijo slik, kot so rotacija, translacija in strižne deformacije.

Naloga 2 (1 točka)

Spišite funkcijo najdi_sifte, ki za dano sliko izračuna njene značilne točke in njihove deskriptorje. Za inicializacijo sift detektorja uporabite objekt `cv2.xfeatures2d.SIFT_create()`. Lokacije ključnih točk pretvorite v numpy array in prikažite grafično, kot je podano na spodnji sliki:



Naloga 3 (2 točki)

Spišite funkcijo `matchDescriptors`, ki najde pare ujemajočih deskriptorjev oziroma njihove indekse. Naj matriki $D_1 \in \mathbb{R}^{N_1 \times 64}$ in $D_2 \in \mathbb{R}^{N_2 \times 64}$ vsebujeta deskriptorje značilnih točk dveh slik, ki jih je N_1 za prvo sliko ter N_2 za drugo. Pare deskriptorjev poiščemo tako, da v prvem koraku izračunamo matriko evklidskih razdalj med vsakim možnim parom deskriptorjev v D_1 in D_2 , ter razdalje shranimo v matriko M . Nato za vsak deskriptor v D_1 njegov pripadajoč par določimo kot deskriptor v D_2 z najmanjšo razdaljo, vendar le, če je razmerje velikosti med njemu prvim in drugim najbližjim deskriptorjem manjše od 0.25. Vsak par nato podamo z indeksoma vrstic D_1 in D_2 , na katerih se nahajata deskriptorja danega para.

Najdene pare lahko preverite z grafičnim prikazom v funkciji `matchSift()`. Spodnja slika prikazuje primer pravilnega delovanja implementiranih funkcij:



Naloga 4 (1 točka)

V main delu programa zaporedoma naložite vsakega izmed parov slik, na njih poiščite pare značilnih točk, ter z uporabo funkcije `cv2.findHomography` določite pripadajočo transformacijsko matriko. Nato z uporabo funkcije `cv2.warpPerspective` izvedite ustrezno transformacijo ene izmed slik ter par slik sešijte v panoramo. Spodnji slike prikazujeta pravilno delovanje postopka:



Rezultati so na kakšnem izmed parov slik glede na izbrane parametre SIFT detektorja in funkcije `cv2.findHomography` lahko slabši. V poročilu opišite pomen parametrov ter poiščite vrednosti, s katerimi dobite sprejemljive rezultate na čim večih izmed parov slik.