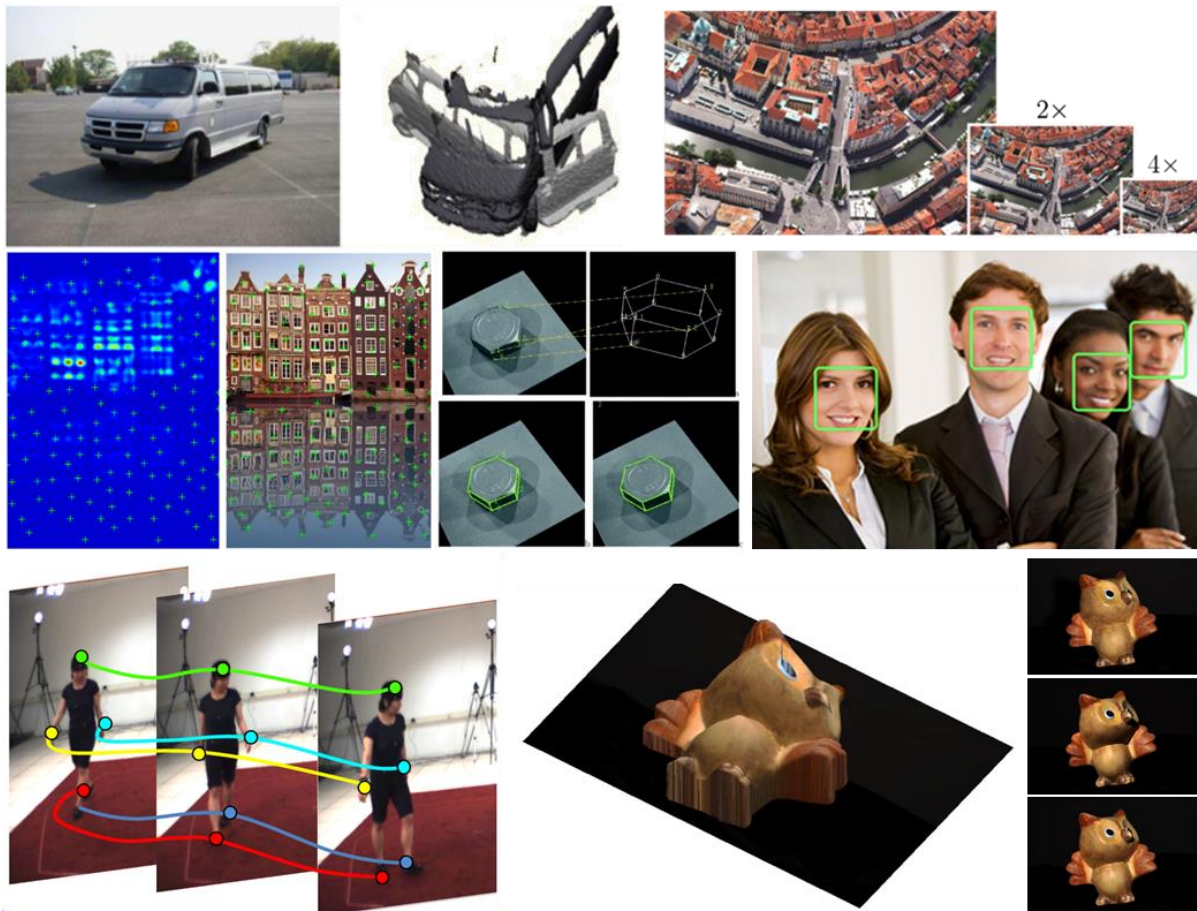




Robotski vid

Predavanja



Univerza v Ljubljani
Fakulteta *za elektrotehniko*



Žiga Špiclin

Robotski vid

Predavanja

Ljubljana, 2019

Predgovor

Pričujoča zbirka prosojnic predavanj predstavlja dopolnilno študijsko gradivo pri predmetu Robotski na Univerzitetnem študiju elektrotehnike 2. stopnje, smer Robotika. Nastala je iz gradiv za izvedbo predavanj pri tem predmetu v preteklih študijskih letih.

Namen gradiva je seznaniti študente z vsebinami predmeta in podati smernice za kasnejšo izvedbo laboratorijskih vaj, ki se tesno navezujejo na predavanja. Zbirka predavanj obsega 10 predavanj, ki študente seznanijo z osnovami vizualnega zaznavanja, s tehnikami in modeli zajema slik ter formati zapisa digitalnih slik, s postopki in parametri opredeljevanja kakovosti slik in slikovnih sistemov, s sivinskimi preslikavami in osnovnimi postopki obdelave slik kot so diskretna konvolucija za namen glajenja in ostrenja, ter povečava in pomanjšava slik z interpolacijo in decimacijo, s postopki zaznavanja objektov, od osnovnih detektorjev oslonilnih točk, robov in parametričnih oblik do postopkov strojnega učenja z globokimi nevronskimi mrežami, geometrijskimi preslikavami in poravnavo slik in oblik, s postopki določanja poze objekta iz 2D slik, s tehnikami in postopki rekonstrukcije 3D oblik, s postopki sledenja objektom v videih in osnovnimi koncepti načrtovanja in vrednotenja sistemov odločanja na podlagi značilnic izločenih iz slik.

Avtor se zahvaljuje vsem sodelavcem Laboratorija slikovne tehnologije na Fakulteti za elektrotehniko, Univeze v Ljubljani, ki so kakorkoli pripomogli k nastanku te zbirke.

Ljubljana, December 2019

Žiga Špiclin

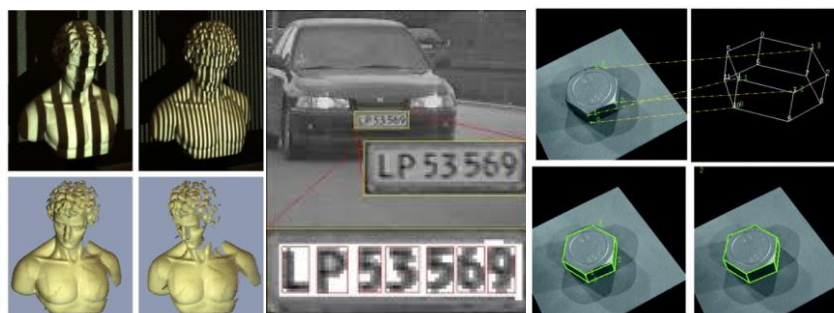
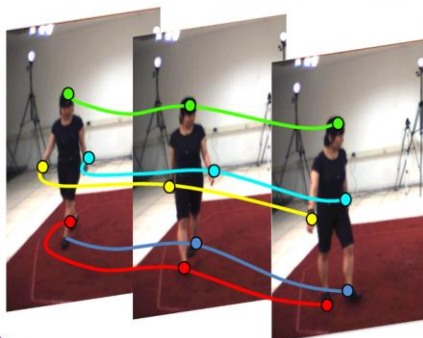
Kazalo

1. Uvod v vizualno zaznavanje	5
2. Zajem slik in slikovni formati	21
3. Parametri kakovosti slik in sivinske preslikave	40
4. Osnovna obdelava slik	63
5. Zaznavanje objektov v slikah	87
6. Geometrijska poravnava slik in oblik	121
7. Prileganje 3D modelov na 2D slike	139
8. Rekonstrukcija 3D oblik	165
9. Vizualna navigacija	178
10. Odločanje na podlagi slik	195

1. Uvod



Robotski vid



1. Uvod 2

Vsebina predmeta

Predavanja

- Uvod v vizualno zaznavanje
- Zajem slik in slikovni formati
- Kakovost slik in sivinske preslikave
- Osnovna obdelava digitalnih slik
- Zaznavanje objektov v slikah
- Poravnava slik in oblik

Laboratorijske vaje

- Uvod v programski jezik
- Upravljanje in prikazovanje slik
- Sivinske preslikave in obnova slik
- Osnovna obdelava slik
- Robustno iskanje 2D objektov
- Geometrijske preslikave slik in oblik

Osnovna znanja

Cilj je pridobiti **splošno uporabna znanja o zajemu in obdelavi digitalnih slik**

- Kalibracija optičnih sistemov
- Določanje poze objektov v prostoru
- Rekonstrukcija 3D oblik
- Vizualna navigacija
- Odločanje na osnovi slike

- Kalibracija optičnih sistemov
- Prileganje 3D modelov na 2D slike
- Rekonstrukcija 3D oblik
- Sledenje objektov
- Vizualna kontrola kakovosti

Specifična znanja

Cilj je pridobiti **specifična znanja za uporabo v aplikacijah robotskega vida**

Univerza v Ljubljani Robotski vid 1. letnik, Smer Robotika
 Fakulteta za Elektrotehniko doc. dr. Žiga Štešič 2. stopnja UN študija Elektrotehniko

1. Uvod 3

Človeški vid

- Vid je zelo pomembno človeško čutilo
- Človeško oko je logaritemsko občutljivo na intenziteto svetlobe
- Občutljivost očesa se spreminja glede na intenziteto svetlobe
- Odnos med zaznano in dejansko svetlobo je zelo kompleksen

Univerza v Ljubljani Robotski vid 1. letnik, Smer Robotika
 Fakulteta za Elektrotehniko doc. dr. Žiga Štešič 2. stopnja UN študija Elektrotehniko

1. Uvod 4

Človeški vid

Opazovani objekt

fizični prostor

Preslikava

oko in živčni sistem

Biološki sistem

možgani

Kaj pomeni videti?

„Vision is the act of knowing what is where by looking.“
– Aristotel

„Vision is the act of unconscious inference.“
– Helmholtz

Kaj omogoča?

Preživetje!

- zaznavanje
- orientacija
- lov
- delo
- žabava
- ...

Pomen čutov?

- vid
- otip
- okus
- sluh
- vonj
- ...

Univerza v Ljubljani Robotski vid 1. letnik, Smer Robotika
 Fakulteta za Elektrotehniko doc. dr. Žiga Štešič 2. stopnja UN študija Elektrotehniko

1. Uvod 5

Robotski vid

Opazovani objekt

fizični prostor

Preslikava
slikovne tehnike

Slika
prostor slike

Kaj pomeni videti?
Kako iz slik ugotoviti, kaj se nahaja v prostoru, kje so objekti in kakšne so njihove lastnosti, kaj se v sliki dogaja.

Robotski vid?
Na podlagi vidne informacije izvesti fizično akcijo.

Kaj omogočajo?

Namen uporabe?

- arhiviranje
- prenašanje
- nadzor
- merjenje
- vodenje
- ...

Aplikacije

- industrija
- medicina
- informiranje
- raziskave
- varnost
- vojska
- promet
- kmetijstvo
- zabava
- ...

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Štešič

1. letnik, Smerni Robotika
2. stopnja UN študija Elektrotehnika

1. Uvod 6

Kaj vidite na sliki?

ČESTITAM!
Ravnokar ste naredili nekaj,
kar je matematično
nemogoče opisati.

KAKO?
Vaša opažanja so temeljila na
predpostavkah o predhodnem
znanju / izkušnjah iz realnega sveta.

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Štešič

1. letnik, Smerni Robotika
2. stopnja UN študija Elektrotehnika

1. Uvod 7

Optične iluzije in kompleksnost vida

(a)

(b)

(c)

(d)

Univerza v Ljubljani
Fakulteta za Elektrotehniko

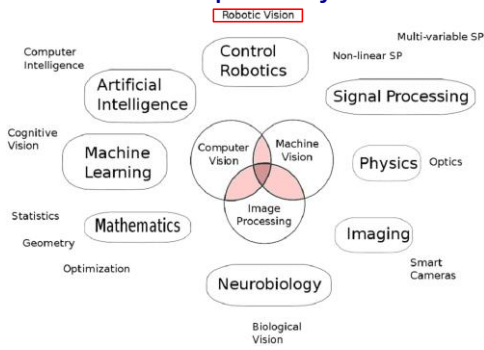
Robotski vid
doc. dr. Žiga Štešič

1. letnik, Smerni Robotika
2. stopnja UN študija Elektrotehnika

Različna področja in izrazi

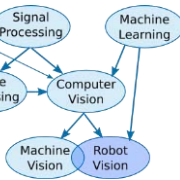
- RAČUNALNIŠKI VID (Computer Vision)
- STROJNI VID (Machine Vision)
- ROBOTSKI VID (Robot Vision)
- UMETNO ZAZNAVANJE (Artificial Perception/Cognition) } slike!
- ANALIZA SLIK (Image Analysis)
- OBDELAVA SLIK (Image Processing)
- RAZUMEVANJE SLIK (Image Understanding)

Povezave med področji



Hierarhija področij

- Tehnike in algoritmi na področju Robotskega vida vključujejo aspekte robotike, kot so kinematika, referenčni okvir, kalibracija in interakcija z okoljem
- Področja ločimo glede na vhod-izhod



Tehnika	Vhod	Izhod
Obdelava signalov	Električni signal	Električni signal
Obdelava slik	Slika	Slika
Računalniški vid	Slika	Informacije/značilnice
Razpoznavanje vzorcev/Strojno učenje	Informacije/značilnice	Informacije
Strojni vid	Slika	Informacije
Robotski vid	Slika	Fizična akcija

Računalniški vid

je področje, ki vključuje (se ukvarja s) **postopki za zajem, obdelavo, analizo in razumevanjem slik** (večdimenzionalnih podatkov) realnih prizorov, z namenom **oblikovanja numerične ali simbolične informacije**, npr. v obliki odločitev.

Ena izmed raziskovalnih tem tega področja je (bil) poskus posnemanja človeškega vida v obliki elektronskega zaznavanja in razumevanja slike. Na razumevanje slike lahko gledamo kot na odkrivanje simbolične informacije vsebovane v sliki z uporabo modelov, ki so zgrajeni s pomočjo geometrije, fizike, statistike ter teorije učenja. Na računalniški vid lahko gledamo kot na zahteven proces avtomatizacije in integracije širokega spektra procesov in predstavitev, ki vodijo k vizualni zaznavi.*

*na internetu poiščite še kakšno drugo definicijo računalniškega vida

Računalniški vid

- Naučiti računalnike, da bi razumeli vsebino slik in videa in znali odgovoriti na različna vprašanja.



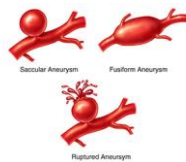
- Kakšen prizor prikazuje slika?
- Kje je oznaka za kraj?
- Kje je cesta?
- Kako daleč je prva hiša?

Računalniški vid

- Naučiti računalnike, da bi razumeli vsebino slik in videa in znali odgovoriti na različna vprašanja.



- Kje so anevrizme (vrečaste razširitve žilne stene)?
- Katera anevrizma lahko kmalu počí?
- Kakšna je optimalna trajektorija do anevrizme?



Računalniški vid

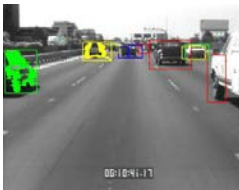
- Naučiti računalnike, da bi razumeli vsebino slik in videa in znali odgovoriti na različna vprašanja.



- Katere tablete / kapsule so poškodovane?
- Katere vrste poškodb se pojavljajo?
- Ali so prisotne rumene tablete?

Računalniški vid

- Naučiti računalnike, da bi razumeli vsebino slik in videa in znali odgovoriti na različna vprašanja.



- Kaj se dogaja v okolici mojega avta?
- Ali so na cestišču pešci ali druge ovire?
- Kako daleč je avto pred mano?

Računalniški vid

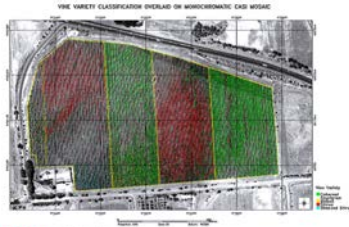
- Naučiti računalnike, da bi razumeli vsebino slik in videa in znali odgovoriti na različna vprašanja.



- Kje so stični spoji?
- Kakšna je kvaliteta zvara?

Računalniški vid

- Naučiti računalnike, da bi razumeli vsebino slik in videa in znali odgovoriti na različna vprašanja.



- Kje so vinogradi?
- Kakšna sorta trte je zasajena?
- Ali je trta obolela in kje ?

Problemi in izzivi

- Nekateri kategorije objektov imajo veliko variacijo pojavnosti



- Nekateri objekti enostavno *nočejo* biti vidni



Zelo kratka zgodovina računalniškega vida

- **1966:** M. Minsky (MIT) uvede računalniški vid kot poletni dodiplomski projekt



Navodilo študentu:
 „spend the summer linking a camera to a computer and getting the computer to describe what it saw“

→ zdaj vemo, da je to precej zahteven problem

1. Uvod 20

Zelo kratka zgodovina računalniškega vida

- **Šestdeseta:** interpretacija sintetične scene
- **Sedemdeseta:** interpretacija izbranih slik
- **Osemdeseta:** umetna inteligenca; premik proti geometriji in matematični strogosti
- **Devetdeseta:** razpoznavanje obrazov, v modi statistična analiza

Guzman '68

Turk & Pentland '91

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

1. Uvod 21

Zelo kratka zgodovina računalniškega vida

- **Dvatisoča:** širše razpoznavanje; velike označene baze slik; začetki obdelave videa
- **Dvatisočdeseta:** strojno učenje, umetne nevronske mreže

- **Veliko podjetij,** ki razvijajo produkte z računalniškim vidom
- **Nepopoln seznam:**
<http://www.cs.ubc.ca/~lowe/vision.html>

feature extraction classification

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

1. Uvod 22

Aplikacije računalniškega vida so številne

- Ena od aplikacij (uporab) računalniškega vida je **strojni vid**, ki združuje tehnologije in postopke v delujoč sistem za avtomatsko vizualno analizo surovin, polizdelkov in izdelkov, za kontrolo procesov ter njihovo vodenje. Če gre za vodenje s fizično akcijo je to aplikacija **robotskega vida**.

→ **Primeri aplikacij:**
<http://courses.cs.washington.edu/courses/cse455/09wi/Lects/lect1.pdf>

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Strojni vid nadomešča človeški vid

- Predvsem pri kontroli kakovosti v intenzivnih proizvodnih procesih

PREGLEDOVANJE TEKSTILA

PREGLEDOVANJE STEKLENIC

PREGLEDOVANJE TABLET IN KAPSUL

Univerza v Ljubljani, Fakulteta za Elektrotehniko

Robotski vid, doc. dr. Ziga Spacín

1. letnik, Smer Robotika, 2. stopnja UN študija Elektrotehnika

Robotski vid nadomešča človeški vid

- Predvsem v kontroliranih okoljih kot so proizvodni procesi, pa tudi drugod

MANIPULACIJA IZDELKOV NA TEKOČEM TRAKU

AVTONOMNA VOŽNJA

Univerza v Ljubljani, Fakulteta za Elektrotehniko

Robotski vid, doc. dr. Ziga Spacín

1. letnik, Smer Robotika, 2. stopnja UN študija Elektrotehnika

Slikovne tehnike

390 – 700 nm

Magnetna resonanca, Termalno slikanje, Svetlobne tehnike, Mikroskopija, Rentgensko slikanje, Računalniška tomografija, Nuklearna medicina

Infrardece slikanje, Fluorescenčna mikroskopija, DTI MRI, Elektronska mikroskopija

Wavelength (metri): Radio, Microwave, Infrared, Visible, Ultraviolet, X-ray, Gamma Ray

Radio: 10^3 , Microwave: 10^2 , Infrared: 10^5 , Visible: 5×10^6 , Ultraviolet: 10^8 , X-ray: 10^{10} , Gamma Ray: 10^{11}

About the size of: [Icons representing various scales from human to atomic]

Robotski vid, doc. dr. Ziga Spacín

1. letnik, Smer Robotika, 2. stopnja UN študija Elektrotehnika

1. Uvod 26

Problem – kompleksna struktura ambientalne osvetlitve

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Ziga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

1. Uvod 27

Problem – kompleksna struktura ambientalne osvetlitve

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Ziga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

1. Uvod 28

Robotski vid zato...

- **združuje** osvetlitev, zajem, obdelavo in analizo slik, ter tehnologije za manipulacijo z izdelki
- **da zagotovi** brezkontaktno lokalizacijo, opis in manipulacijo mirujočih in gibajočih se izdelkov

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Ziga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Zasnova in gradnja sistemov z robotskim vidom zahteva znanja o

- osvetlitvi
- optiki
- senzorjih (kamere, ...)
- mehanski manipulaciji
- algoritmih za obdelavo slik
- arhitekturi računalnikov
- uporabniških vmesnikov
- zagotavljanju kakovosti

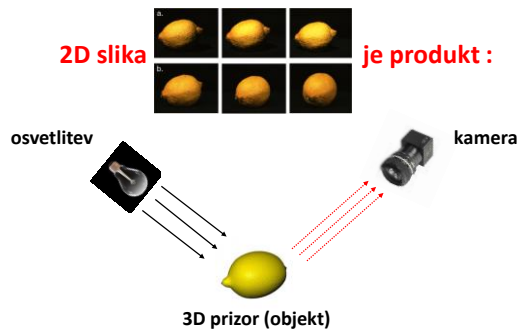


Sistemi z robotskim vidom

- Pogosto delajo v omejenih in kontroliranih okoljih in okoliščinah,
 - v katerih lahko vplivamo na oz. kontroliramo osvetlitev, kontrast in resolucijo slike ter položaj in orientacijo objektov in/ali poznamo lastnosti objektov, ki jih pregledujemo.



Zajem slike



Osvetlitev



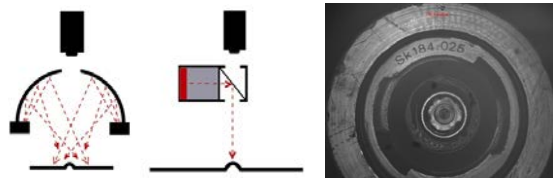
- Z osvetlitvijo od zadaj (**Back Lighting**) dobimo temen (črn) obris objekta na belem ozadju
 - Takšne slike olajšajo odkrivanje prisotnosti/odsotnosti objekta (delov objekta); določanje njegovega položaja in orientacije, dolžin in površin; manipulacijo in nadaljnjo obdelavo; ...



Osvetlitev



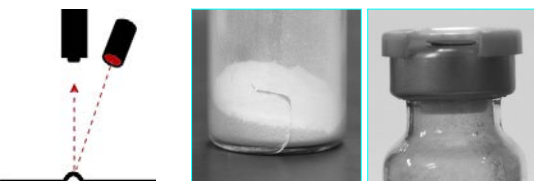
- Difuzna osvetlitev (**Diffuse Lighting or Full Bright Field Lighting**) je zelo učinkovita za osvetljevanje ukrivljenih in bleščečih površin



Osvetlitev



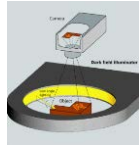
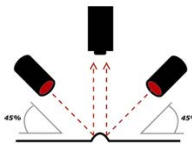
- **Direktno osvetlitev (Partial Bright Field or Directional Lighting)** uporabljamo za generiranje senc in s tem kontrasta ter za poudarjanje topografskih podrobnosti



Osvetlitev



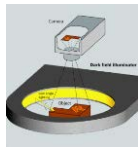
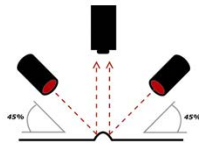
- **Osvetlitev pod nizkim kotom (Dark Field Lighting)** je verjetno najmanj razumljiva tehnika, četudi jo vsakodnevno uporabljamo
 - *Primer:* avtomobilski žarometi osvetljujejo cesto pod nizkim kotom. To omogoča zaznavanje poškodbe cestišča in objektov na njem



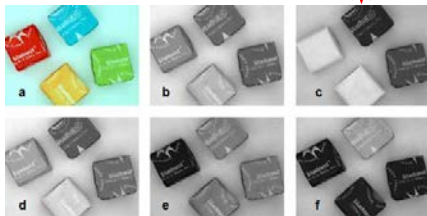
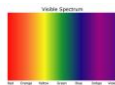
Osvetlitev pod nizkim kotom



Primeri:

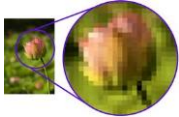
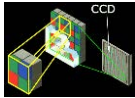


Barva osvetlitve



Slike bombonov. a) bela svetloba in barvna kamera b) bela svetloba in črno-bela kamera c) rdeča svetloba posvetli rdeče in rumene objekte in potemni modre in zelene d) rdeča & zelena svetloba posvetli rumene objekte bolj kot rdeče, e) zelena svetloba posvetli zelene in modre ter potemni rdeče objekte f) modra svetloba posvetli modre ter potemni druge.

Kamere: 2D (array scan)



Kamere: 1D (line scan)



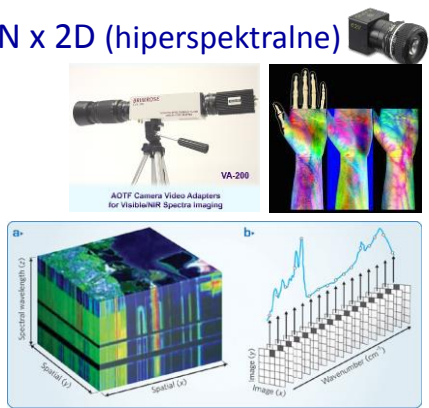
Kamere: infrardeče

- Zaznavajo svetlobo v infrardečem območju od 700 nm do 1 mm



Kamere: N x 2D (hiperspektralne)

- Slike pri različnih valovnih dolžinah svetlobe



Robotski vid

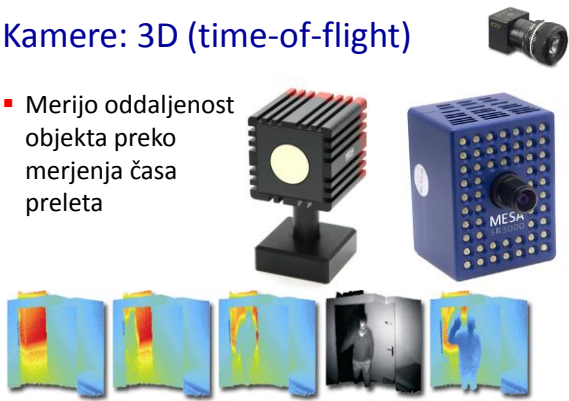
doc. dr. Ziga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Kamere: 3D (time-of-flight)

- Merijo oddaljenost objekta preko merjenja časa preleta



Robotski vid

doc. dr. Ziga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Kamere: pametne

- Integriran računalnik
- Programabilne
- Knjižnica za analizo slik



Robotski vid

doc. dr. Ziga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Povzetek

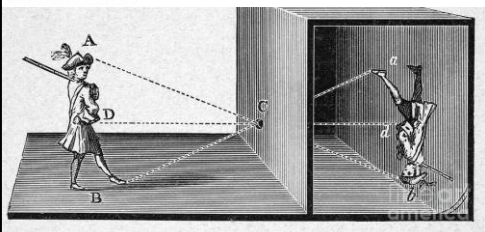
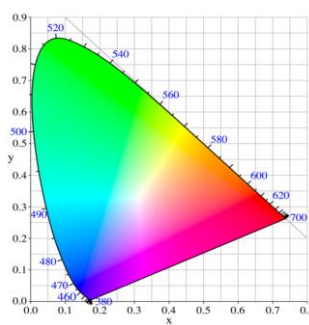
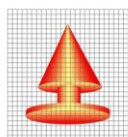
- Potek, vsebina, gradiva dostopna na <https://e.fe.uni-lj.si>
- Predavanja in vaje se tesno navezujejo, zato naj študij poteka redno in sprotno
- Robotski vid je interdisciplinarno področje, ki izhaja iz računalniškega vida, umetnega zaznavanja, obdelave in analize slik
- Združuje osvetlitev, zajem, obdelavo in analizo slik ter tehnologije za manipulacijo z izdelki za brezkontaktno lokalizacijo, opis in manipulacijo mirujočih in gibajočih se izdelkov



2. Zajem slik in slikovni formati



Robotski vid

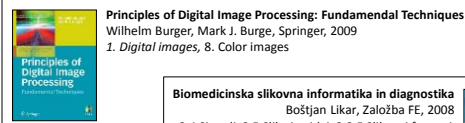


Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Špiclin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Študijska literatura



Principles of Digital Image Processing: Fundamental Techniques
Wilhelm Burger, Mark J. Burge, Springer, 2009
1. *Digital images*, 8. *Color images*

Biomedicinska slikovna informatika in diagnostika
Boštjan Likar, Založba FE, 2008
2.4 *Signali*, 2.5 *Slike in videi*, 3.3.5 *Slikovni formati*,
zalogovniki in standardi zgoščevanja



Computer vision: algorithms and applications
Richard Szeliski, Springer, 2010
Knjiga je dostopna na spletu
2.3.2 *Color*

Zajem slike

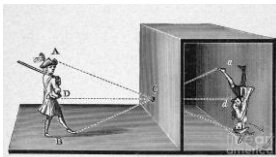


3D prizor

2D slika

- Postopek (proces), ki **preslika 3D prizor v 2D sliko** je spremenljiv in zahteven
- Večina postopkov obdelave in analize slik uporablja *model optične kamere z odprtino* kot geometrijski model zajema slike

Model kamere z odprtino („pinhole“)



http://sl.wikipedia.org/wiki/Camera_obscura



Fotografija Abelarda Morella, 1991

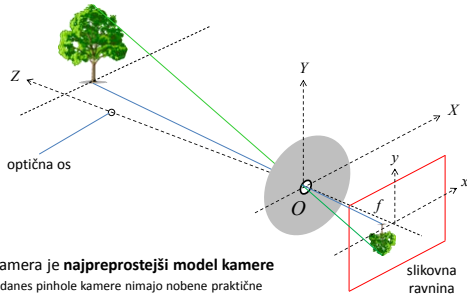
- **Camera obscura** (latinsko: temna soba), znana že v 13. stoletju, je predhodnica sodobnega fotoaparata
- V osnovi škatla, v katero skozi **majhno odprtino** (pinhole) **vdirajo svetlobni žarki**, na njeni zadnji steni **oblikujejo pomanjšano in obrnjeno sliko predmeta pred odprtino**
- Načelo delovanja te naprave še dandanes posnemajo fotoaparati in kamere → beseda kamera izhaja prav iz izraza camera obscura

Primer sobe z odprtino

- Prizor v sobi
- Pogled iz sobe

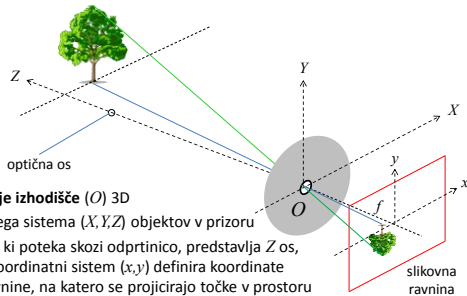


Model pinhole kamere



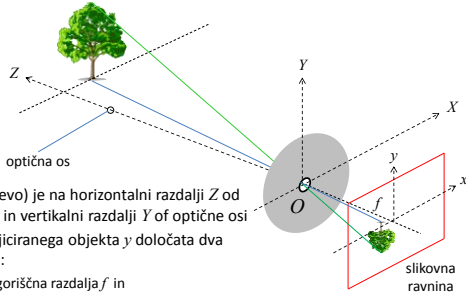
- Pinhole kamera je **najpreprostejši model kamere**
 - Četudi danes pinhole kamere nimajo nobene praktične vrednosti, so uporaben model za razumevanje osnovnih optičnih komponent kamere

Model pinhole kamere



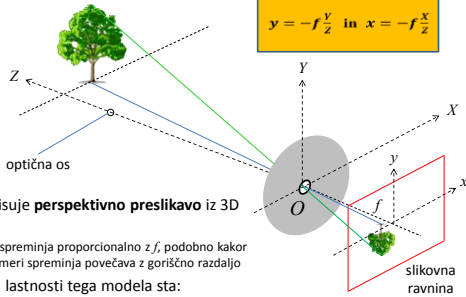
- Odprtina je izhodišče (O) 3D** koordinatnega sistema (X, Y, Z) objektov v prizoru
- Optična os**, ki poteka skozi odprtino, predstavlja Z os, ločeni 2D koordinatni sistem (x, y) definira koordinate slikovne ravnine, na katero se projicirajo točke v prostoru
- Razdalja f (goriščna razdalja)** med odprtino in slikovno ravnino določa skalo (velikost) projekcije

Model pinhole kamere



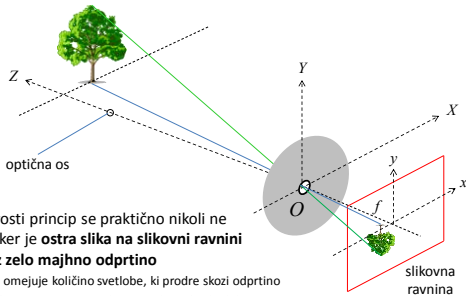
- Objekt (drevo) je na horizontalni razdalji Z od odprtine in vertikalni razdalji Y of optične osi
- Višino projiciranega objekta y določata dva parametra:
 - fiksna goriščna razdalja f in
 - oddaljenost Z objekta od koordinatnega izhodišča O
- S primerjanjem podobnih trikotnikov dobimo $y = -f \frac{Y}{Z}$ in $x = -f \frac{X}{Z}$

Model pinhole kamere



- Enačba opisuje **perspektivno preslikavo** iz 3D v 2D
 - Skala se spreminja proporcionalno z f , podobno kakor se pri kameri spreminja povečava z goriščno razdaljo
- Pomembni lastnosti tega modela sta:
 - premice v 3D se preslikajo v premice v 2D
 - krogi se preslikajo v elipse

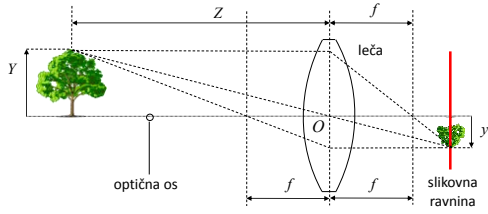
Model pinhole kamere



- Sicer preprosti princip se praktično nikoli ne uporablja, ker je **ostra slika na slikovni ravnini pogojena z zelo majhno odprtino**
 - ta precej omejuje količino svetlobe, ki prodre skozi odprtino
 - zajem slike po tem principu bi zato zahteval zelo dolge ekspozicijske čase

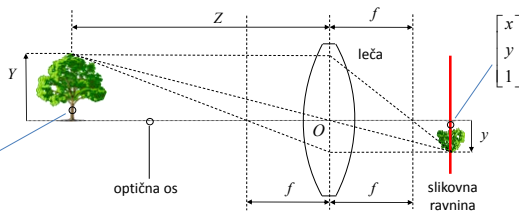
Model tanke leče

$$y = -f \frac{y'}{Z} \text{ in } x = -f \frac{x'}{Z}$$



- Pri tem modelu predpostavljamo, da je leča simetrična in neskončno tanka, tako da se vsi svetlobni žarki odklonijo na virtualni ravnini sredi leče
 - slikovna geometrija je praktično enaka kot pri *pinhole* modelu
 - ne odraža vseh fizikalnih lastnosti sistema leč, kot so geometrijska popačenja ter medsebojno različni odkloni svetlobe različnih valovnih dolžin

Perspektivna projekcija

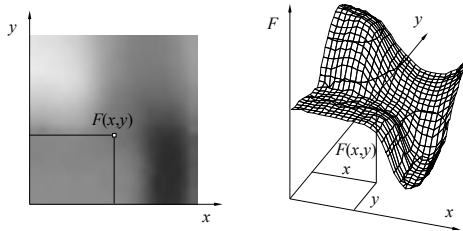


homogene koordinate

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/f & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ -Z/f \\ 1 \end{bmatrix} \Rightarrow [-f \frac{X}{Z}, -f \frac{Y}{Z}, 1] = [x, y, 1]$$

Slika

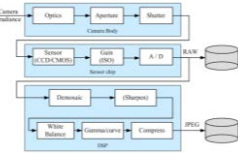
- Sistem leč projicira na senzor kamere dvodimenzionalno (2D), časovno spremenljivo, zvezno porazdelitev svetlobne energije



Digitalna slika

▪ Digitalni posnetek zvezno spreminjajoče se porazdelitve svetlobe dobimo v treh korakih:

1. Zvezno porazdelitev svetlobe moramo *prostorsko vzorčiti*
2. Dobljeno diskretno funkcijo moramo potem *vzorčiti v časovnem prostoru*, da dobimo eno sliko
3. Dobljene vrednosti moramo *kvantizirati* v omejeno množico številskih vrednosti



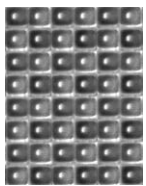
Realna slika Vzorčenje Kvantizacija Digitalna slika



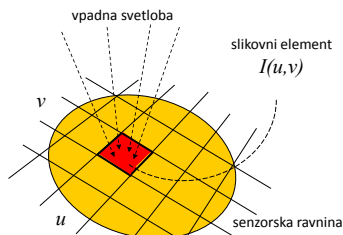
Korak 1: Prostorsko vzorčenje

▪ Prostorsko vzorčenje slike je preslikava zveznega signala v diskretno predstavitev

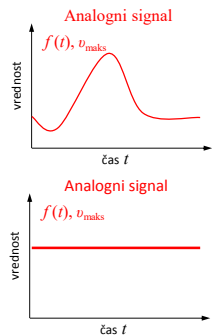
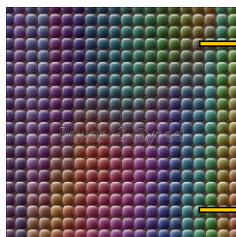
- Zavisí od geometrije elementov senzorna kamere - ponavadi pravokotno polje elementov na senzorski ravnini



Príkaz enakomerne razporeditve elementov v 2D rastru, kjer vsak element "meri" lokalno svetlobo, ki pada nanj.

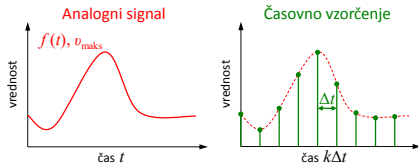


Korak 1: Prostorsko vzorčenje



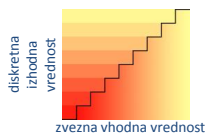
Korak 2: Časovno vzorčenje

- Časovno vzorčenje temelji na merjenju količine svetlobe, ki pade na senzor v določenem časovnem intervalu
 - pri CCD in CMOS senzorjih je časovno vzorčenje izvedeno s proženjem zbiranja električnega naboja, ki ga povzroča zvezni tok fotonov in merjenjem količine naboja, ki se je med ekspozicijo nabral na posameznem elementu senzorja

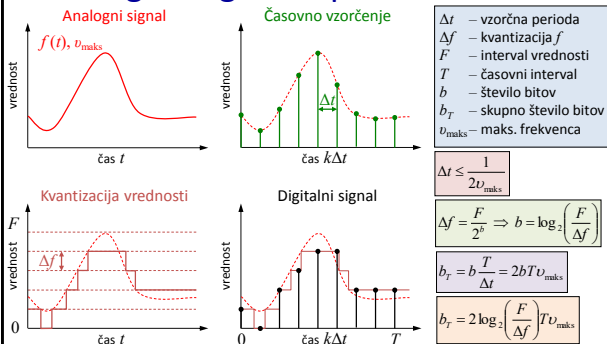


Korak 3: Kvantizacija vrednosti

- Da bi lahko shranili in obdelali sliko, ki je po korakih 1. in 2. opisana z vrednostmi, ki ustrezajo količini naboja na posameznih elementih, je potrebno te vrednosti pretvoriti v razpon celih števil
 - naprimer $256 = 2^8$ ali $4096 = 2^{12}$
- Pretvorba se izvede s pomočjo **analogno digitalnega pretvornika**, ki je ponavadi vgrajen v senzorsko elektroniko, ali pa s pomočjo posebnega vmesnika



Analogno digitalna pretvorba



Δt – vzorčna perioda
 Δf – kvantizacija f
 F – interval vrednosti
 T – časovni interval
 b – število bitov
 b_T – skupno število bitov
 v_{maks} – maks. frekvenca

$$\Delta t \leq \frac{1}{2v_{maks}}$$

$$\Delta f = \frac{F}{2^b} \Rightarrow b = \log_2 \left(\frac{F}{\Delta f} \right)$$

$$b_T = b \frac{T}{\Delta t} = 2bTv_{maks}$$

$$b_T = 2 \log_2 \left(\frac{F}{\Delta f} \right) Tv_{maks}$$

Analogno digitalna pretvorba

Vzorčni periodi (Δx in Δy)

→ Nyquistov teorem - brez izgube informacije

$$\Delta x = \Delta y = \frac{1}{2v_{\text{maks}}}$$

Kvantizacija vrednosti (Δf in F)

→ zanesljivost merjenja (šum, natančnost)
→ namen uporabe in variabilnost podatkov

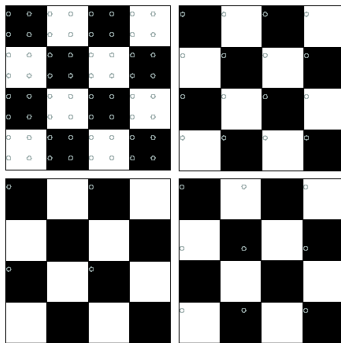
$$\Delta f = \frac{F}{2^b} \Rightarrow b = \log_2 \left(\frac{F}{\Delta f} \right)$$

Δx – vzorčna perioda po x osi
 Δy – vzorčna perioda po y osi
 X – velikost slike po x osi
 Y – velikost slike po y osi
 Δf – kvantizacija vrednosti $f(x,y)$
 F – dinamično območje $f(x,y)$
 b – število bitov za en piksel
 b_{XY} – skupno število bitov za sliko
 v_{maks} – maksimalna frekvenca slike

Skupno število bitov

$$b_{XY} = b \frac{X}{\Delta x} \frac{Y}{\Delta y} = 4bXYv_{\text{maks}}^2 = 4 \log_2 \left(\frac{F}{\Delta f} \right) XYv_{\text{maks}}^2$$

Nyquistov teorem – 2D primer



← Ustrezna vzorčna frekvenca

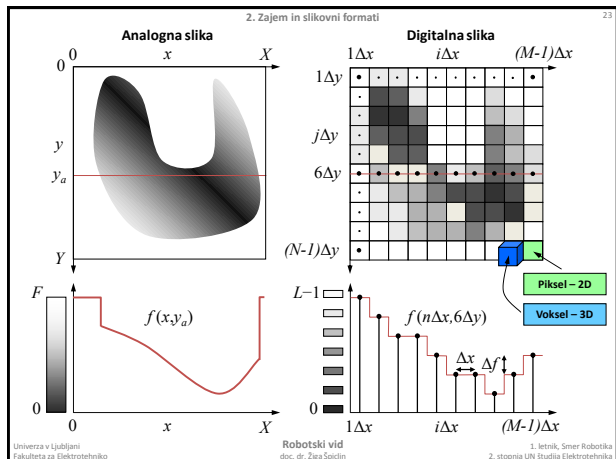
← Neustrezna vzorčna frekvenca
(rekonstrukcija originalne slike iz zajetih vzorcev ni možna)

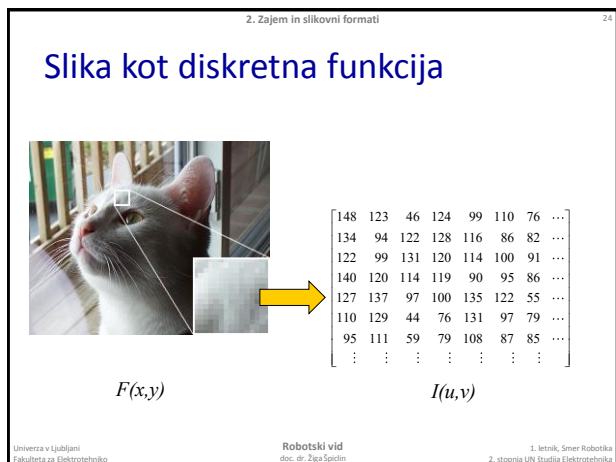
Slika kot diskretna funkcija

- Rezultat omenjenih treh korakov je opis slike v obliki urejene **dvodimenzionalne matrike celih števil**
- Formalno je digitalna slika I definirana kot:

$$I(u, v) \in F \quad \text{in} \quad u, v \in N$$

- Slika $I(u, v)$ lahko prenesemo v računalnik, jo shranimo, zgostimo, obdelamo, prikažemo ...
 - slika je le še dvodimenzionalno polje števil
 - za nadaljno obdelavo ni nujno več pomembno, kako je bila zajeta





2. Zajem in slikovni formati 25

Velikost in ločljivost slike

- Ločljivost = **resolution**
- Če predpostavimo, da je slika kvadratna, je njena velikost določena s širino M (število stolpcev) in višino N (število vrstic) slikovne matrice I
- Ločljivost je običajno podana s številom slikovnih elementov na enoto (kilometer, milimeter, mikrometer, inch, itd.), ki opisuje velikost prizora na sliki
 - naprimer **dots per inch** (dpi) ali **lines per inch** (lpi) pri tisku, pixels per micron pri mikroskopskih posnetkih, pixels per kilometer pri satelitskih posnetkih, itd.
- Ločljivost v navpični in vodoravni smeri je ponavadi enaka, kar pomeni, da so slikovni elementi kvadratne oblike

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Srečič

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

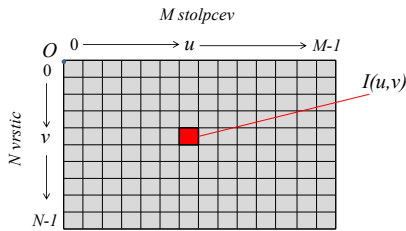
Velikost in prostorska ločljivost slike

- Prostorska ločljivost ne vpliva oz. ni pomembna za večino postopkov obdelave slik
 - vendar je pomembna, kadar želimo meriti absolutne razdalje med objekti na sliki ali pa površine objektov

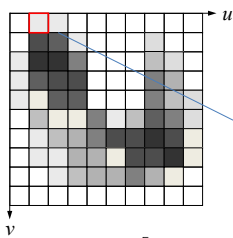


Koordinatni sistem slike

- Lokacija slikovnega elementa zavisi od koordinatnega sistema slike – izhodišče in osi



Matrični zapis 2D digitalne slike

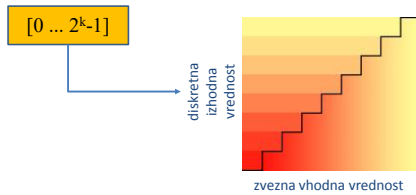


indeksiranje piksla $u=0, v=1$

$$I(x,y) \approx I(u,v) = \begin{bmatrix} I(0,0) & I(1,0) & \dots & I(M-1,0) \\ I(0,1) & I(1,1) & \dots & I(M-1,1) \\ \vdots & \vdots & \ddots & \vdots \\ I(0,N-1) & I(1,N-1) & \dots & I(M-1,N-1) \end{bmatrix}$$

Vrednosti slikovnih elementov (pikslov)

- Vrednosti sivinskih slik, ki so zajete z enim kanalom, predstavljajo intenziteto (intensity), svetlost (brightness) ali gostoto (density) slike
- Vrednosti so cela števila iz množice



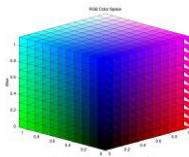
Vrednosti slikovnih elementov (pikslov)

- Zaloga vrednosti pikslov je zaradi dvojiškega zapisa informacije v pomnilniku iz nabora 2^k (npr. 64, 128, 256, 512, 1024) različnih vrednosti, kjer je k „bitna globlina“ (depth) slike

Sivinske (intenzitetne) slike	Št. kanalov	Bitov/Pixel	Razpon	Uporaba
	1	1	0 ... 1	<i>Binarne slike:</i> dokument, ilustracija, fax
	1	8	0 ... 255	<i>Univerzalna:</i> foto, sken, print
	1	12	0 ... 4095	<i>Visoka kvaliteta:</i> foto, sken, print
	1	14	0 ... 16383	<i>Profesionalna:</i> foto, sken, print
	1	16	0 ... 65535	<i>Najvišja kvaliteta:</i> medicina, astronomija

Vrednosti slikovnih elementov (pikslov)

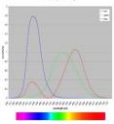
- Večina **barvnih (RGB) slik** temelji na primarnih barvah
 - rdeča (R-red), zelena (G-green) in modra (B-blue), kjer je vsaka od RGB komponent 8-bitna



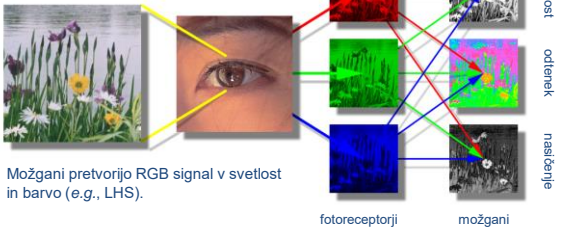
Barvne slike	Št. kanalov	Bitov/Pixel	Razpon	Uporaba
	3	24	$[0 \dots 255]^3$	<i>RGB splošno:</i> foto, sken, print
	3	36	$[0 \dots 4095]^3$	<i>RGB visoka kvaliteta:</i> foto, sken, print
	3	42	$[0 \dots 16383]^3$	<i>RGB profesionalna:</i> foto, sken, print

Zaznavanje barve pri človeku

spektralna občutljivost fotoreceptorjev →



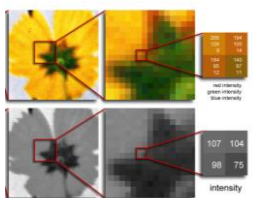
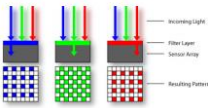
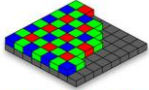
Oko ima tri tipe fotoreceptorjev, ki so občutljivi na svetlobo v rdečem, zelenem in modrem delu spektra vidne svetlobe.



Možgani pretvorijo RGB signal v svetlost in barvo (e.g., LHS).

Zaznavanje barve v barvnih digitalnih kamerah

Bayer vzorec

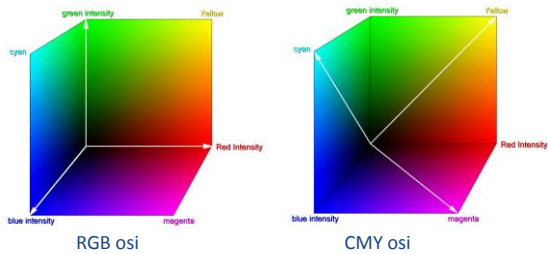


Barvne slike imajo 3 vrednosti na piksel, sivinske pa le eno

Manjkajoče vrednosti določa na podlagi vrednosti v sosednjih pikslih (interpolacija, demosaicing)

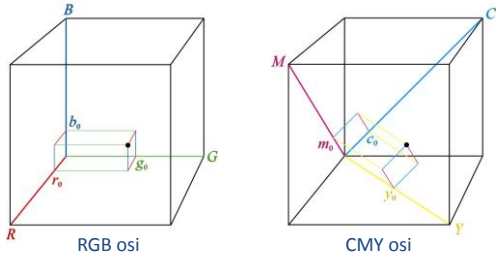
RGB barvni prostor

- Določen z **energijo svetlobe** po posamezni osnovni valovni dolžini
- Predstavitev barvnega prostora:



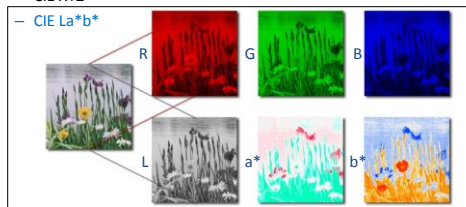
Predstavitev barve v različnih barvnih prostorih

- Ista barva ima različne koordinate



Standardni barvni prostori

- Vsak uporablja standarden nabor osnovnih barv oz. funkcij preslikave barve
- Linearni barvni prostori
 - RGB
 - CIE XYZ
- Nelinearni barvni prostori
 - HSV

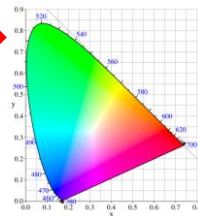
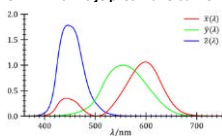


CIE XYZ 1931 barvni prostor

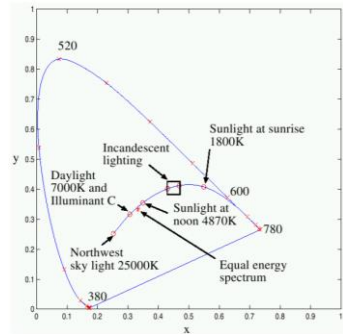
- CIE: Commission International d'Eclairage (CIE), 1931
- Določen s projekcijo na funkcije preslikave barve:

$$\begin{cases}
 X = \int_{380}^{780} L_{60,0,1}(\lambda) \bar{x}(\lambda) d\lambda \\
 Y = \int_{380}^{780} L_{60,0,1}(\lambda) \bar{y}(\lambda) d\lambda \\
 Z = \int_{380}^{780} L_{60,0,1}(\lambda) \bar{z}(\lambda) d\lambda
 \end{cases}
 \rightarrow
 \begin{cases}
 x = \frac{X}{X+Y+Z} \\
 y = \frac{Y}{X+Y+Z}
 \end{cases}$$

CIE XYZ funkcije preslikave barve

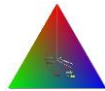


Predstavitev barv v xy prostoru

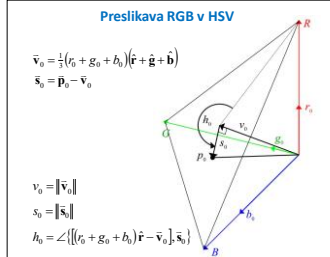
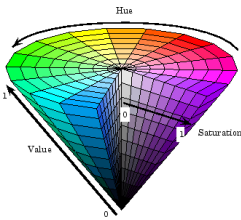


- Spektralna krivulja od 380 do 780 nm
 - barve monokromatske svetlobe
- Krivulja barv črnih teles
 - barve črnih teles glede na njihovo temperaturo
- Barva nekaterih standardnih svetil

HSV barvni prostor

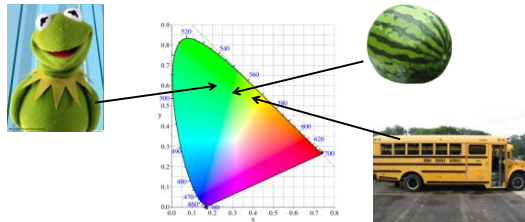


- Odtенок (hue), nasičenje (saturation), svetlost (value)
- Nelinearen – odraža topologijo barv tako, da kodira odtенок s kotom $0 - 360^\circ$



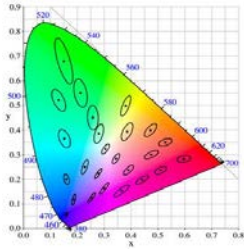
Razdalje v barvnem prostoru

- Ali so razdalje med točkami v barvnem prostoru smiselne glede na naše zaznavanje razlik med barvami?



Razdalje v barvnem prostoru

- ...niso nujno smiselne: CIE XYZ 1931 ni enakomeren v smislu barve → magnitude sprememb so slab indikator razlike v barvi



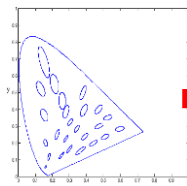
McAdamove elipse označujejo področja komaj zaznavne spremembe barve

- CIE 1960
- CIE 1964
- CIE 1976

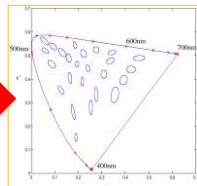
Enakomerni barvni prostori

- Z ustrežno preslikavo osi barvnega prostora lahko približamo koncept razdalje v barvnem prostoru dejanski zaznavni razliki
- Primera barvnih prostorov:

- CIE $u'v'$
- CIE La^*b^*



CIE XYZ



CIE $u'v'$

Slikovni formati, zalogovniki in standardi zgoščevanja

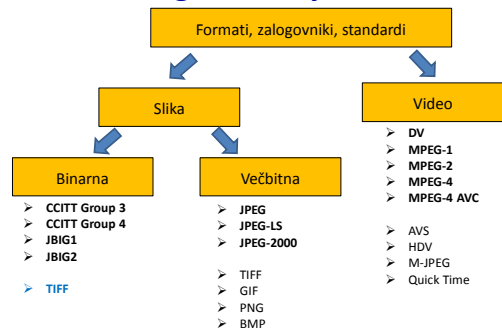
- **Slikovni datotečni format** (image file format)
 - standardni način organizacije in shranjevanja podatkov, ki določa urejenost in morebitno zgoščevanje podatkov
- **Slikovni zalogovnik** (image container)
 - podoben datotečnemu formatu, vendar lahko vsebuje več vrst podatkov
- **Standardi zgoščevanja** (compression standards)
 - določajo načine (postopke) zgoščevanja slik – predstavitev slikovne vsebine s čim manjšo količino podatkov



Slikovni formati, zalogovniki in standardi zgoščevanja

- Obstaja veliko standardiziranih datotečnih formatov
 - uporaba standardiziranih formatov omogoča lažjo izmenjavo slik in poveča verjetnost, da bo slika mogoče prebrati z drugo programsko opremo
- Izbira primernega formata je zavisi od:
 - vrste slike: črno bela, sivinska, sken dokumenta, barvna, barvna grafika, medicinska, satelitska
 - velikosti in kompresije: ali shranjevanje slike predstavlja problem, kakšen je postopek kompresije, ...
 - kompatibilnost: kako pomembna je izmenjava podatkov, kako dolgo naj bodo podatki v arhivu berljivi
 - področje aplikacije: na katerem področju bodo slike pretežno uporabljene, ali so namenjene tiskanju, medicini, grafiki, ...

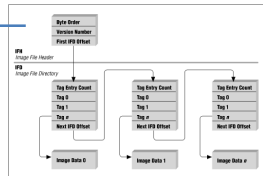
Slikovni formati, zalogovniki in standardi zgoščevanja



TIFF – Tagged Image File Format

- Široko uporaben, fleksibilen in univerzalen format
 - razvilo ga je podjetje Aldus, potem pa nadgradila Microsoft in Adobe
 - podpira sivinske, indeksirane in barvne slike, lahko vsebuje več slik z različnimi značilnostmi
 - dovoljuje različne postopke kompresije (LZW, ZIP, JPEG) in barvne prostore
 - v eni TIFF datoteki je možno shraniti različne primere iste slike – v različnih velikostih in predstavitvah

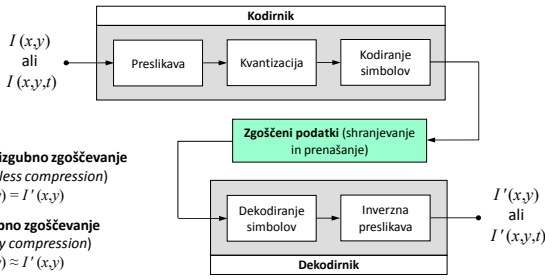
Length	Value
2	Byte Order (0x49, 0x49) for little endian ('L', 'L') (0x4D, 0x4D) for big-endian ('M', 'M')
2	TIFF ID Always (0x00, 0x2A)
4	Offset To 0th IFD From the start of the TIFF header



Ime Organizacija(e)	Opis
Videi	
AVS MII	Standard AVS (ang. <i>Audio-Video Standard</i>), razvit na Kitajskem, je podoben H.264, vendar uporablja Golombovo kodiranje.
HDV Konzorcij podjetij	Standard HDV (ang. <i>High Definition Video</i>) je razširitev DV standarda za HDTV, ki uporablja zgoščevanje podobno MPEG-2, vključno z odstranjevanjem časovne redundance z napovedovanjem razlik.
M-JPEG Različna podjetja	Standard M-JPEG (ang. <i>Motion JPEG</i>) uporablja neodvisno JPEG zgoščevanje posameznik slik.
QuickTime Apple Computer	Zalogovnik medijev, ki podpira DV, H.261, H.262, H.264, MPEG-1, MPEG-1, MPEG-4 in ostale formate oz. standarde za zgoščevanje videov.
VC-1 oz. WMV9 SMPTE	Najbolj uporabljen video format na internetu. Privojen za HD in Blue-ray HD DVD-je. Podoben je H.264/AVC. Uporablja celoštevilsko DCT s spremenljivo velikostjo blokov in vsebini prilagojene kode s spremenljivo dolžino, vendar brez napovedovanja med zaporednimi slikami.
Microsoft	

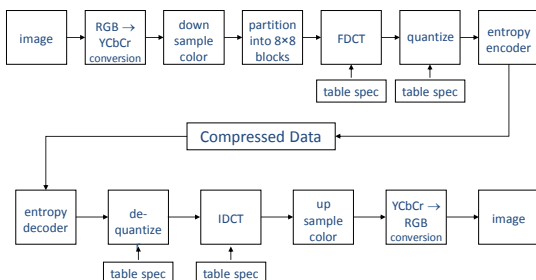
Sistem zgoščevanja slik

- Zaporedje dajlic
- Bločna predstavitev
- Predhodne in sledeče slike
- Prostor in čas. redundanca
- Povraten postopek
- Objektive ocene kakovosti
- Povprečno bitno razmerje (bit/s) → spremenljiva kakovost videa
- Nepomembna informacija
- **Nepovraten postopek**
- Konstantna ali spremenljiva koda
- Uploštevanje verjetnosti pojavljanja sivin
- Redundanca kodiranja
- Povraten postopek



- **Brezizgubno zgoščevanje** (lossless compression)
 $I(x,y) = I'(x,y)$
- **Izgubno zgoščevanje** (lossy compression)
 $I(x,y) \approx I'(x,y)$

Primer - sistem zgoščevanja JPEG



JPEG izgubna kompresija slik



Originalna slika 5244 x 4716
1200 ppi, velikost 127 MBajtov



Povzetek

- Preslikavo slike realnega objekta na slikovno ravnino opišemo s perspektivno projekcijo
- Digitalno sliko dobimo s prostorskim in časovnim vzorčenjem ter kvantizacijo vrednosti
 - zapis sivinske v obliki 2D matrike vrednosti
 - barvna slika s tremi vrednostmi v pikslu – vrednosti zavisijo od barvnega prostora
- Digitalne slike in videe lahko zapišemo v različnih standardnih formatih
 - razlike glede na tip slike (binarna, večbitna oz. sivinska, barvna)
 - možnosti zapisa večih slik, animacij, metapodatkov o sliki (EXIF)
 - načinom zgoščevanja, npr. izgubno ali brezizgubno



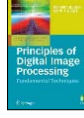
3. Parametri kakovosti slik in sivinske preslikave




Robotski vid




Študijska literatura



Principles of Digital Image Processing: Fundamental Techniques
 Wilhelm Burger, Mark J. Burge, Springer, 2009
 3. Histograms, 4. Point operations



Biomedicinska slikovna informatika in diagnostika
 Boštjan Likar, Založba FE, 2008
 2.4 Signali, 2.5 Slike in videl



Computer vision: algorithms and applications
 Richard Szeliski, Springer, 2010
 Knjiga je dostopna na spletu
 3. Image processing


Kakovost slik

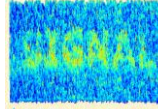
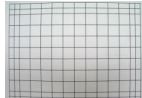
- Kompleksna večparametrična lastnost
- Odvisna od kakovosti preslikave iz fizičnega prostora v prostor slike → slikovne tehnike
- Določa uporabno vrednost slik

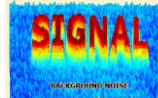


Parametri kakovosti slik

- Ločljivost oz. resolucija
- Kontrast
- Razmerje signal/šum
- Nepravilnosti oz. artefakti
- Točnost in natančnost

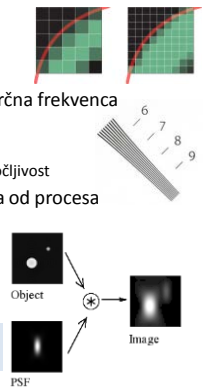




Prostorska ločljivost

- Pogosto napačno interpretirana kot vzorčna frekvenca
 - izražena kot število točk na mm oz. inčo (*DPI – dots per inch*)
 - vzorčna frekvenca je zgolj zgornja meja za ločljivost
- Dejanska prostorska ločljivost je odvisna od procesa zajemanja slik
 - fenomen zameglitve ostrih točk
 - opišemo ga s funkcijo razširitve točke (*PSF – point spread function*) →



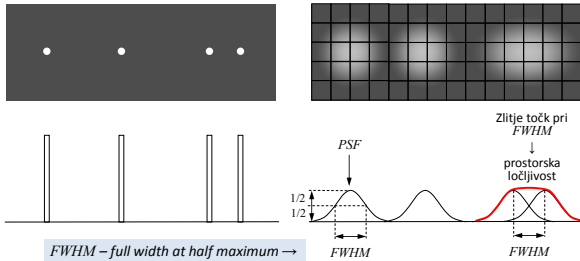
PSF opisuje odziv sistema za zajemanje slik na točkovni izvor svetlobe ali točkovni objekt.

Prostorska ločljivost

Proces zajemanja slik

Idealna slika

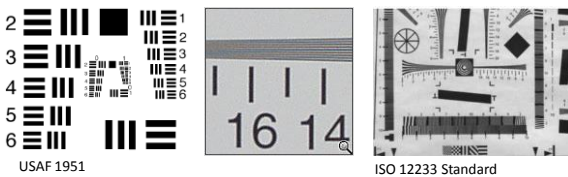
Realna slika



FWHM – full width at half maximum →

Prostorska ločljivost

- Uporabljajo se kalibri z različnimi gostotami linij
 - ločljivost izrazimo kot največje število linijskih parov na milimeter (lp/mm), ki jih na zajeti sliki še lahko razločimo

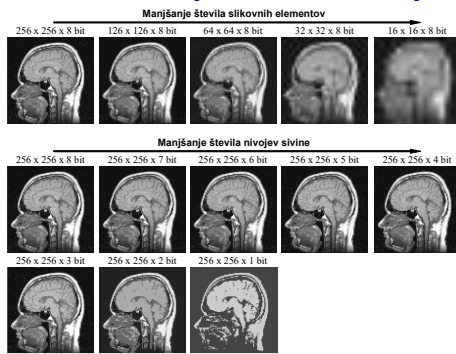


Sivinska ločljivost



- Sivinska ločljivost je najmanjša še razločljiva sprememba sivine
- Zelo subjektivno opravilo
 - človeški vid je bistveno bolj občutljiv na število slikovnih elementov, kot pa na število diskretnih nivojev sivine
- Zato namesto sivinske ločljivosti podajamo število kvantiziranih vrednosti sivine v bitih
 - 16-bitne, 12-bitne, 8-bitne, 4-bitne, 1-bitne (binarne oz. črno-bele) slike

Vpliv kvantizacije na zaznavanje



Kontrast

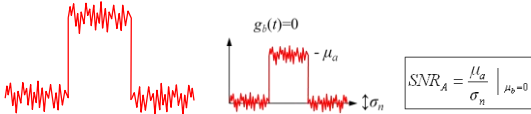


- Razlika sivinskih vrednosti med sosednjimi področji na sliki (lokalno-globalna lastnost)
- Odvisen je od procesa zajemanja slik
- Tudi od načina prikazovanja slik:
 - velik vpliv osvetljenosti okolja
 - omogočiti moramo spreminjanje kontrasta

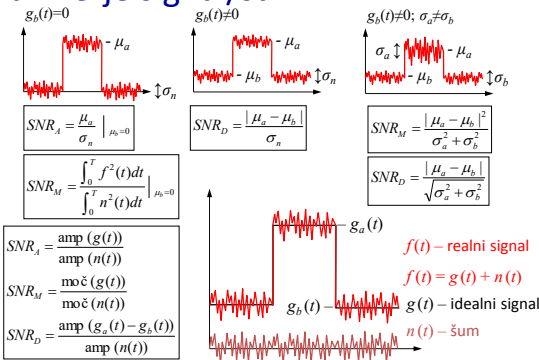


Razmerje signal/šum (SNR)

- Izvor, vrsta in velikost šuma je odvisna od procesa zajemanja slik
 - Najbolj znana sta Gaussov in Poissonov šum
- Kaj je razmerje signal/šum? (*signal-to-noise ratio - SNR*)
 - Merilo relativne jakosti signala glede na njegov šum
 - Merilo sposobnosti zaznavanja fizikalnih veličin danega signala



Razmerje signal/šum



Razmerje signal/šum

- Razmerje signal/šum ni standardizirano
- Izberemo ga glede na:
 - vrsto signala in šuma ter namen uporabe
 - tako da ocenjuje želeno in predstavlljivo lastnost signala
- Včasih ga podajamo v decibelih (dB):

Za amp. in dif. razmerje:

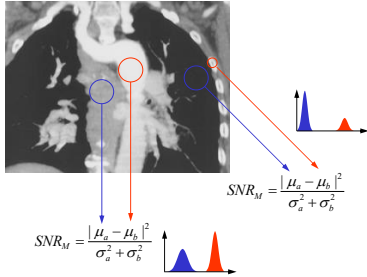
$SNR_{(dB)} = 20 \log_{10} SNR$

Za močnostno razmerje:

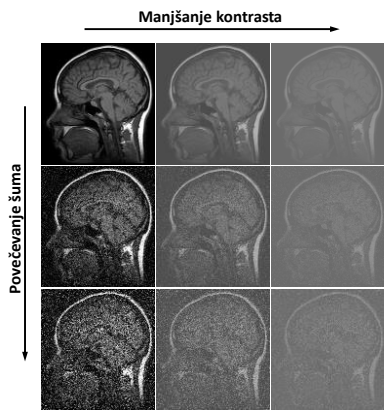
$SNR_{(dB)} = 10 \log_{10} SNR$

Razmerje signal/šum

- Kako določimo razmerje signal/šum?



Kontrast in šum



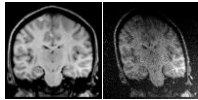
Nepravilnosti oz. artefakti

- Lastnosti, ki ne predstavljajo realnih objektov
- Izvirajo iz procesa zajemanja slik, naprimer:
 - premikanje opazovanih objektov
 - nehomogenost energije (svetloba, magnetno polje)
 - sivinske nehomogenosti
 - nepopolnost optike
 - optične aberacije: distorzije, barvne aberacije,...
 - kovinski predmeti
 - zvezdni artefakti (CT), izničen signal (MRI)
 - napake pri obdelavi slik
 - ostrenje robov, prevzorčenje, filtriranje slik, ...



Problemi zaradi nepravilnosti

- Lahko prikrijejo dejanske lastnosti objektov
- Lahko jih napačno interpretiramo kot veljavne lastnosti slikanih objektov
- Vplivajo na kvantitativno vrednotenje slik:
 - moramo jih čim bolje odpraviti ali
 - vsaj čim bolje razumeti in ovrednotiti njihove potencialne vplive na uporabo slik

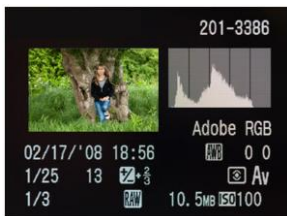


Razprava – kakovost slik

Parametri	Izvori / vzroki / odvisnosti	Rešitve strojna in/ali programska oprema
ostrina in ločljivost	optika (OTF, aberacije), CCD	boljša optika in CCD, ostrenje in obnova
globinska ostrina	kakovost optike, zaslonka	boljša optika, zaprta zaslonka
geometrijske deformacije	kakovost optike, goriščnica	boljša optika, kalibracija
nehomogenosti svetlosti	osvetlitev, optika, CCD	homogena osvetlitev, kalibracija, retrospektivna obnova
kontrast	osvetlitev, optika, CCD	spekter osvetlitve, boljša optika in CCD
šum	osvetlitev, CCD	več svetlobe, boljši CCD, filtriranje, večkratno zajemanje slik
občutljivost, hitrost zaslone, frekvenca zajemanja	lastnosti CCD in elektronike	boljši CCD in elektronika
...

Histogram

- Histogram podaja statistiko pojavljanja sivin (barv) v sliki
 - najosnovnejše orodje, ki omogoča preprosto in pregledno grafično prikazovanje frekvenčnih porazdelitev vrednosti slikovnih elementov



Histogram

- Histogram h diskretnih vrednosti meritev na območju $[0, K-1]$ je diskretna funkcija:

$$h(r_k) = n_k$$

- kjer r_k k -ta diskretna vrednost; $k = 0, 1, \dots, K-1$; n_k pa število vzorcev z diskretno vrednostjo r_k .

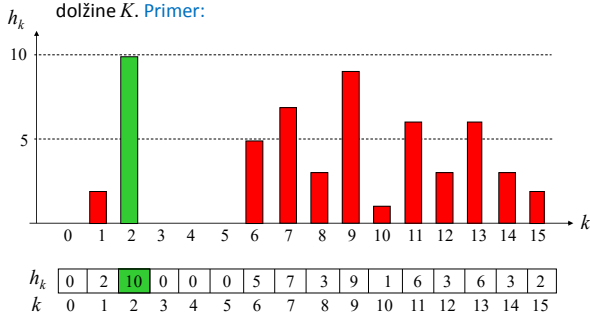
- Histogram 8-bitne ($K=256$) sivinske slike I pa je:

$$h(k) = n_k = \text{št. pixlov v sliki } I \text{ z vrednostjo } k$$



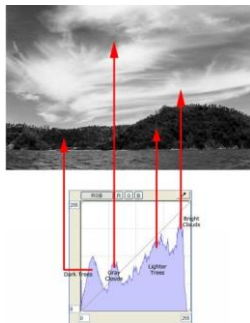
Histogram

- Rezultat računanja histograma je enodimenzionalni vektor h dolžine K . **Primer:**



Histogram

- Histogram sivinske slike \rightarrow
- Histogram **ne vsebuje prostorske informacije** – informacije o tem, kje v sliki so piksli z določeno vrednostjo
- Iz slike lahko zgradimo histogram, iz histograma pa ne moremo enoznačno zgraditi slike



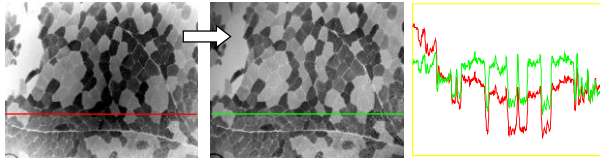
Lastnosti slike iz histograma

- S pomočjo histograma lahko na grobo hitro ocenimo nekatere lastnosti slike, npr. ali je bila slika primerno osvetljena



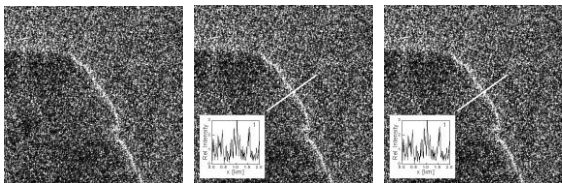
Prerez

- Prezi omogočajo prikazovanje vrednosti slikovnih elementov v prostorih nižjih dimenzij od originalnega
- Prikazovanje slik v obliki **1D prostorskih signalov** omogoča lažjo **detekcijo globalnih in lokalnih sprememb vrednosti slike**

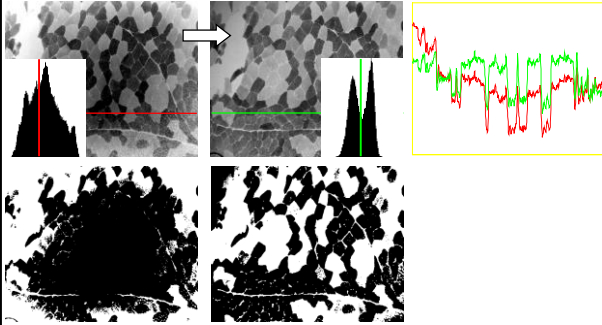


Prerez

- Prerez 2D slike $F(x,y)$ dobimo tako, da za poljubno konstantno vrednost x ali y koordinate, x_c ali y_c , vrednosti slike prikažemo le kot funkcijo druge koordinate, torej kot prostorski signal (1D prerez) vzdolž osi x ; $F(x_c, y)$, ali vzdolž osi y ; $F(x, y_c)$
- Če izrazimo eno koordinato kot funkcijo druge, npr. kot $y=g(x)$, lahko vrednosti slike prikažemo kot prostorski signal $f(x, g(x))$ vzdolž poljubno izbrane funkcije $g(x)$, ki je lahko premica ali pa krivulja v 2D

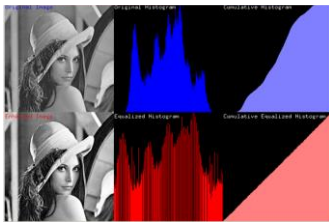


Histogram in prerez

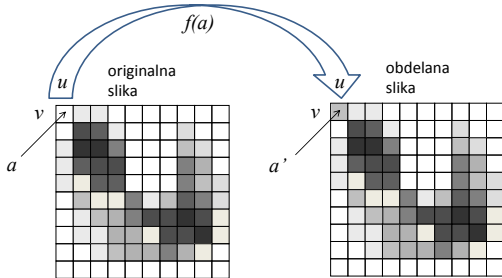


Točkovne operacije (Point operations)

- Spremenijo vrednosti slikovnih elementov
 - spremenijo histogram, ne pa velikosti, geometrije in lokalne strukture slike



Točkovne operacije



Nova vrednost piksla $a' = f(a)$ na lokaciji (u, v) je odvisna samo od prejšnje (originalne) vrednosti $a = I(u, v)$ na isti lokaciji in je neodvisna od vrednosti kateregakoli drugega piksla in še posebej od piksllov v soseščini (filtri).

Točkovne operacije

- Funkcija $f(a)$ preslika prvotne vrednosti v slikovnih elementih v nove vrednosti na vsaki lokaciji (u,v)

$$a' \leftarrow f(a)$$

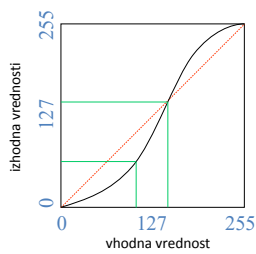
$$I'(u,v) \leftarrow f(I(u,v))$$

- Točkovne operacije se uporabljajo za
 - spreminjanje svetlosti ali kontrasta slike
 - poljubne preslikave intenzitet
 - kvantizacijo slik
 - globalno upravljanje
 - gama korekcijo
 - barvne preslikave



Točkovne operacije

- Preslikavo vrednosti lahko predstavimo kot *look-up tabela*



npr.	indeks	vrednost
...
	101	64
	102	68
	103	69
	104	70
	105	70
	106	71
...
	vhod	izhod

Homogene točkovne operacije

- Če je funkcija $f(a)$ neodvisna od koordinat slike (enaka povsod na sliki) je operacija *globalna* oz. *homogena*

- povečanje/pomanjšanje kontrasta

$$f_{k+}(a) = a \cdot 1,5$$

$$f_{k-}(a) = a \cdot 0,5$$

- osvetlitev/potemnitev

$$f_{o+}(a) = a + 10$$

$$f_{o-}(a) = a - 10$$

- invertiranje slike

$$f_{inv}(a) = -a + a_{max}$$

Kako te operacije predstavimo z look-up tabelo?

Aritmetične operacije

▪ Osnovne aritmetične operacije:

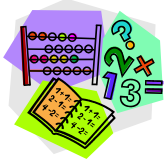
- **seštevanje, odštevanje, množenje in deljenje**, izvajamo med sivinskimi vrednostmi vseh istoležnih slikovnih elementov
- lahko jih kombiniramo in izvajamo med poljubnim številom slik istih dimenzij in enakih velikosti

$$s(x, y) = f(x, y) + g(x, y)$$

$$o(x, y) = f(x, y) - g(x, y)$$

$$m(x, y) = f(x, y) \times g(x, y)$$

$$d(x, y) = f(x, y) \div g(x, y)$$



Seštevanje slik – povprečenje

- Model nastanka šuma

$$v(x, y, z) = f(x, y, z) + n(x, y, z)$$

- Zajem več slik iste scene ter izračun **povprečne slike**:

$$\mu_n(x, y, z) = \frac{1}{K} \sum_{n=1}^K v_n(x, y, z)$$

- Potem je **pričakovana vrednost povprečne slike enaka sliki brez šuma**:

$$E[\mu_n(x, y, z)] = f(x, y, z)$$

- Standardna deviacija šuma povprečne slike pa je za **koren števila zajetih slik manjša** od standardne deviacije šuma ene same slike:

$$\sigma_{\mu_n(x, y, z)} = \frac{\sigma_{n(x, y, z)}}{\sqrt{K}} = \frac{\sigma_{v(x, y, z)}}{\sqrt{K}}$$



Seštevanje slik – povprečenje

Originalna slika s šumom



Povprečje 25 slik s šumom



Seštevanje slik – povprečenje

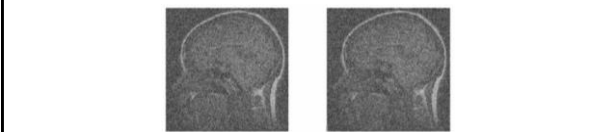
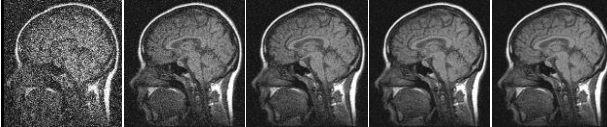
Povprečje 4 slik, 20 s

Povprečje 46 slik, 20 s



Seštevanje slik – povprečenje

$\nu(x,y)$ $\mu_s(x,y), K=10$ $\mu_s(x,y), K=20$ $\mu_s(x,y), K=30$ $\mu_s(x,y), K=40$



Odštevanje slik – angiografija

- Uporabljammo ko želimo **poudariti razlike med slikami**:

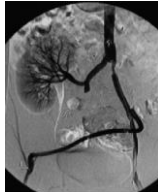
$$o(x,y) = f(x,y) - k(x,y)$$



Retina

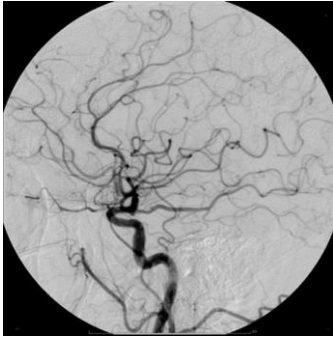


Ledvica



- Razlike slik lahko delamo v krajših časovnih intervalih – študije gibanja
- Potrebna je standardizacija geometrije zajemanja slik ali poravnava slik

Odštevanje slik – angiografija



Možgani

Odštevanje slik – radiografija

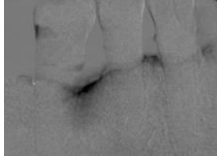
Referenčna slika



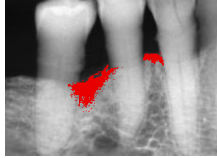
Kontrolna slika



Slika razik



Referenčna slika z razliko



Odštevanje ozadja – varnost



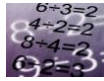
Deljenje – obnova homogenosti

- Uporabljamo npr. pri **obnovi sivinske homogenosti** slik
- Model nastanka nehomogene slike:

$$m(x, y, z) = f(x, y, z) \times g(x, y, z)$$

- Homogeno sliko lahko dobimo s pomočjo inverznega modela, če ocenimo polje nehomogenosti (kalibracija, retrospektivna obnova):

$$f(x, y, z) \approx \tilde{f}(x, y, z) = \frac{m(x, y, z)}{\tilde{g}(x, y, z)}$$



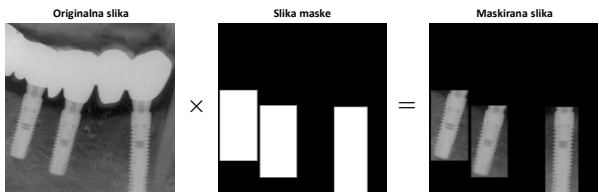
Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Ziga Srećin

1. letnik, smer Robotika
2. stopnja UNI Studija Elektrotehnika

Množenje – maskiranje

- Pogosto uporabljamo tudi pri **maskiranju področij zanimanja** (ang. regions of interest – ROI) na slikah



Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Ziga Srećin

1. letnik, smer Robotika
2. stopnja UNI Studija Elektrotehnika

Povzetek

- Kakovost slik je večparametrična lastnost
 - kontrast, razmerje signal/šum, prostorska ločljivost, šum
 - s parametri opišemo in medsebojno primerjamo kakovost sistemov in podsistemov za zajem slik
- Histogram slike predstavlja frekvenčno porazdelitev sivinskih vrednosti
 - z njim ocenimo lastnosti slike (osvetljenost, kontrast)
 - histogram barvne slike izračunamo za vsak kanal posebej
- Točkovne operacije vplivajo na vrednost v posameznem slikovnem elementu, ne pa na geometrijo in lokalno strukturo slike
 - primeri so linearna in gama preslikava, izenačevanje in prilaganje histograma, aritmetične operacije med slikami



Univerza v Ljubljani
Fakulteta za Elektrotehniko

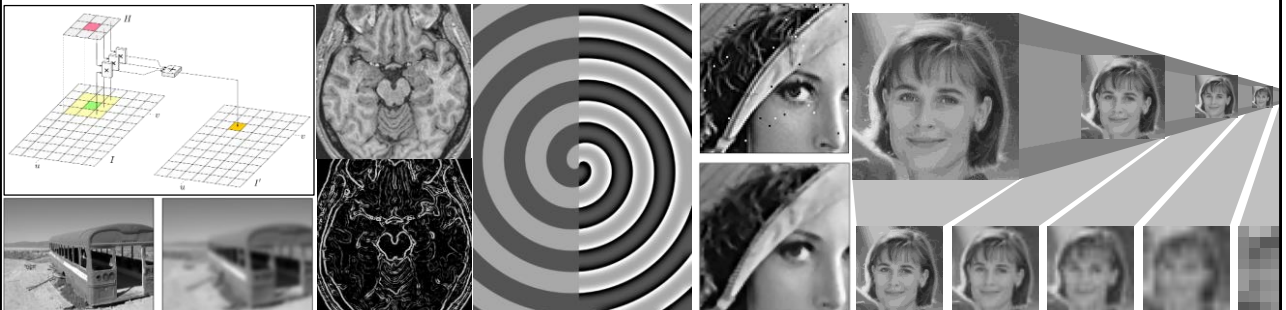
Robotski vid
doc. dr. Ziga Srećin

1. letnik, smer Robotika
2. stopnja UNI Studija Elektrotehnika

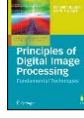
4. Osnovna obdelava slik




Robotski vid




Študijska literatura



Principles of Digital Image Processing: Fundamental Techniques
Wilhelm Burger, Mark J. Burge, Springer, 2009
5. Filters, 6. Edges and contours, 7. Morphological filters



Biomedicinska slikovna informatika in diagnostika
Boštjan Likar, Založba FE, 2008
3.2.1 Interpolacija slik, 3.2.2 Decimacija slik
4.2 Prostorsko filtriranje



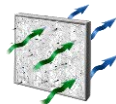
Computer vision: algorithms and applications
Richard Szeliski, Springer, 2010
Knjiga je dostopna na spletu
3. Image processing

Vsebina



- Filtriranje v prostoru slike
 - Osnove filtriranja
 - Glajenje in ostrenje slik
 - Statistično filtriranje
 - Prilagodljivo filtriranje
 - Morfološko filtriranje
- Interpolacija in decimacija

Prostorsko filtriranje



- Filtriranje v prostoru slike je eno od najpomembnejših orodij pri obdelavi slik:
 - uporabno za zmanjševanja šuma, ostrenje, decimacijo,...
- Ime filtriranje izhaja iz klasičnega filtriranja v frekvenčnem prostoru
 - prepuščanje določenih frekvenc
- Podobne operacije lahko izvajamo tudi v prostoru slike:
 - transformacija v frekvenčni prostor ni potrebna
 - bolj enostavno, pregledno in lažje izvedljivo
 - nudi še številne dodatne možnosti, kot je denimo nelinearno ali statistično filtriranje v prostoru slike

Prostorsko filtriranje



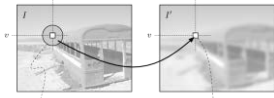
Razlika med filtri in točkovnimi operacijami

- točkovne operacije z funkcijo $g \rightarrow$ preslikajo sivinsko vrednost v vsakem slikovnem elementu

$$a' \leftarrow g(a, u, v)$$

$$I'(u, v) \leftarrow g(I(u, v), u, v)$$

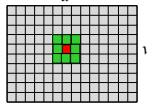
- filtri uporabljajo vrednosti večjega števila slikovnih elementov originalne slike $I(u, v)$, da določijo vrednost piksla nove slike $I'(u, v)$



- primer filtra:** vrednost v vsaki točki slike nadomestimo s povprečno vrednostjo vrednosti okoliških točk

$$I'(u, v) \leftarrow \frac{1}{9} \cdot [I(u-1, v-1) + I(u, v-1) + I(u+1, v-1) + I(u-1, v) + I(u, v) + I(u+1, v) + I(u-1, v+1) + I(u, v+1) + I(u+1, v+1)]$$

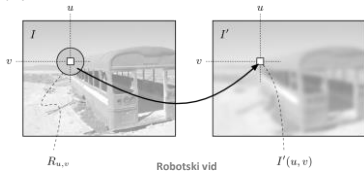
$$I'(u, v) \leftarrow \frac{1}{9} \cdot \sum_{j=-1}^1 \sum_{i=-1}^1 I(u+i, v+j)$$



Prostorsko filtriranje

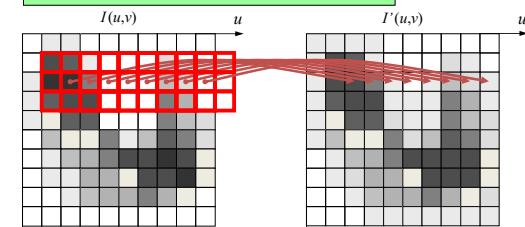
Pomembni parametri

- velikost okolice** določa koliko originalnih pikselov bo udeleženi pri določitvi vrednosti piksla v filtrirani sliki
- oblika okolice** – ta ni nujno kvadratna
- utež pikselov v okolici** – npr. bližnjim pikslom damo večjo utež
- linearne ali nelinearne karakteristike** – ali vrednosti pikselov v okolici nastopajo v linearnem ali nelinearnem izrazu



Prikaz izvedbe filtriranja

Splošni lokalni operator $I'(u, v) = T(I(u, v))$

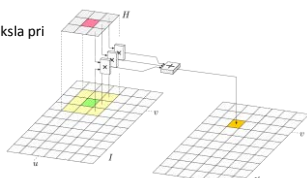
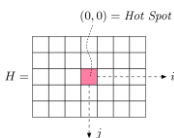


Linearni operator $I'(u, v) = \sum_{i=-a}^a \sum_{j=-b}^b I(u+i, v+j) \cdot H(i, j)$

Linearni filtri

- Določeni z lokalnim operatorjem $H(i, j)$

- tudi jedro ali maska ali matrika filtra, ki je podobno kot slika diskretna, dvodimenzionalna funkcija realnih vrednosti
- filter ima svoj koordinatni sistem, katerega izhodišče je ponavadi v središču
- vsak element $H(i, j)$ določa utež piksla pri seštevaju



$$I'(u, v) \leftarrow \sum_{i=-1}^{i=1} \sum_{j=-1}^{j=1} I(u+i, v+j) \cdot H(i, j) \quad \text{za vse piksele } (u, v)$$

Lastnosti linearnih filtrov

- Za diskretni 2D funkciji $I(u, v)$ in $H(i, j)$ je operacija konvolucije definirana kot:

$$I'(u, v) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} I(u-i, v-j) \cdot H(i, j)$$

$$I'(u, v) = \sum_{(i,j) \in R_H} I(u-i, v-j) \cdot H(i, j)$$

$$= \sum_{(i,j) \in R_H} I(u+i, v+j) \cdot H(-i, -j)$$

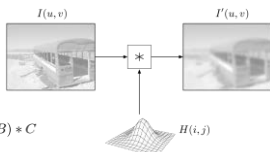
$$= \sum_{(i,j) \in R_H} I(u+i, v+j) \cdot H^*(i, j)$$

- Podobno izrazu za filtriranje – linearni korelaciji = korelacija z zrcaljeno matriko:

$$I'(u, v) \leftarrow \sum_{(i,j) \in R_H} I(u+i, v+j) \cdot H(i, j)$$

Lastnosti linearnih filtrov

- Matematični koncept linearnega filtriranja je konvolucija



- Lastnosti konvolucije

- komutativnost: $I * H = H * I$
- asociativnost: $A * (B * C) = (A * B) * C$
- linearnost: $(s \cdot I) * H = I * (s \cdot H) = s \cdot (I * H)$
 $(I_1 + I_2) * H = (I_1 * H) + (I_2 * H)$
 $(b + I) * H \neq b + (I * H)$

- Uporaba lastnosti

- če: $H = H_1 * H_2 * \dots * H_n$
- potem: $I * H = I * (H_1 * H_2 * \dots * H_n)$
 $= (\dots((I * H_1) * H_2) * \dots * H_n)$

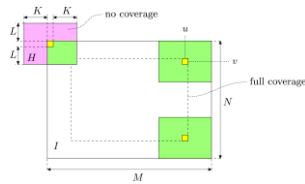
Ločljivost filtra:

$H_x = [1 \ 1 \ 1 \ 1 \ 1]$ in $H_y = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$I' = (I * H_x) * H_y = I * (H_x * H_y)$

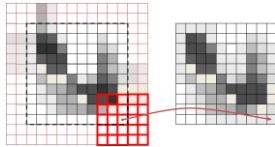
$H_{xy} = H_x * H_y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$

Problem filtriranja na robu slike



$$I'(u, v) = \sum_{i=-1}^{K-1} \sum_{j=-1}^{K-1} I(u+i, v+j) \cdot H(i, j)$$

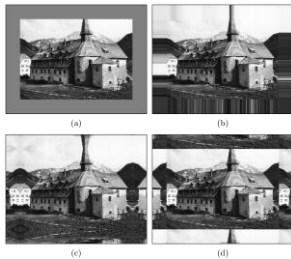
- Lokalni operator gleda čez rob slike, kjer vrednosti niso določene
- Z ekstrapolacijo sivinskih vrednosti ustrezno razširimo sliko



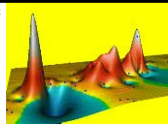
Problem filtriranja na robu slike

- Načini ekstrapolacije sivinskih vrednosti

- postavi neobdelane piksele na konstantno vrednost (npr. 0 – črna)
 - problem: slika je na robu popažena oz. se zmanjša
- razširi mejne piksele čez rob
- zrcali piksele na robu čez rob
- sliko periodično pomnigoteri v horizontalni in vertikalni smeri



Glajenje slik



- Filtriranje se najpogosteje uporablja za **glajenje in zmanjševanje šuma**
- Uporaba glajenja:
 - odstranjevanje manjših nepravilnosti npr. pred razgradnjo večjih objektov
 - povezovanje manjših prekinitev linij in krivulj
 - decimacija slik
- Šum odpravljamo z linearnim in nelinearnim filtriranjem

Glajenje s povprečenjem

Aritmetično povprečje:

- za odpravljanje manjših nepravilnosti
- nizkopasovno filtriranje

$$I'(x, y) = \frac{1}{9} \sum_{i=-1}^1 \sum_{j=-1}^1 I(u-i, v-j)$$

Uteženo povprečje

- navadno bolj utežimo središče filtra

$$I'(u, v) = \frac{\sum_{i=-a}^a \sum_{j=-b}^b I(u-i, v-j) \cdot H(i, j)}{\sum_{i=-a}^a \sum_{j=-b}^b H(i, j)}$$

Aritmetično povprečje

$$\frac{1}{9} \times$$

1	1	1
1	1	1
1	1	1

Uteženo povprečje

$$\frac{1}{16} \times$$

1	2	1
2	4	2
1	2	1

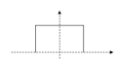
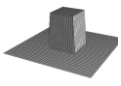
Gaussov filter

0,01	0,08	0,01
0,08	0,64	0,08
0,01	0,08	0,01

Glajenje s povprečenjem

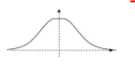
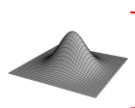
Aritmetično povprečje

(box filter)



0	0	0	0	0
0	1	1	1	0
0	1	1	1	0
0	1	1	1	0
0	0	0	0	0

Uteženo povprečje z Gaussovo funkcijo

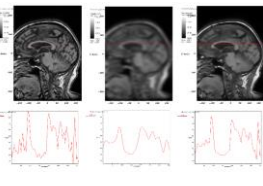


0	1	2	1	0
1	3	5	3	1
2	5	9	5	2
1	3	5	3	1
0	1	2	1	0

→ Gaussov filter

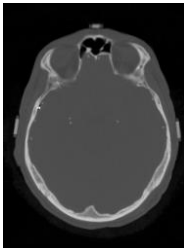
$$G_{\sigma}(r) = e^{-\frac{r^2}{2\sigma^2}} \text{ ali } G_{\sigma}(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}}$$

→ Parameter σ določa velikost matrike in stopnjo glajenja:

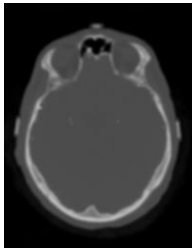


Glajenje s povprečenjem

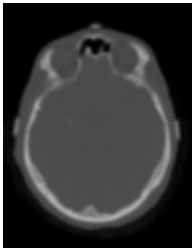
Originalna slika



Gaussov filter 3 × 3

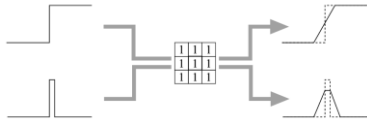


Gaussov filter 7 × 7



Slabosti linearnih filtrov

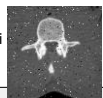
- Pri glajenju za namen odstranjevanja šuma bomo zgladili vse točke, robove in linije
 - zato so linearni filtri običajno neprimerni za odstranjevanje šuma



- Za hkratno odstranitev šuma in ohranitev robov uporabimo nelinearne filtre, kjer vrednosti pikslov v maski povezuje neka nelinearna funkcija

Glajenje s povprečenjem

Impulzni šum →



Geometrično povprečje:

- MN -ti koren iz produkta vseh MN sivinskih vrednosti znotraj domene filtra
- nelinearen operator

$$g(x, y) = \left(\prod_{i=-a}^a \prod_{j=-b}^b f(x-i, y-j) \right)^{\frac{1}{MN}}$$

Harmonično povprečje:

- učinkovito za svetli impulzni šum (sol) in ostale vrste šumov
- neučinkovito za temni impulzni šum (popper)
- nelinearen operator

$$g(x, y) = \frac{MN}{\sum_{i=-a}^a \sum_{j=-b}^b \frac{1}{f(x-i, y-j)}}$$

Nasprotno harmonično povprečje:

- Q = stopnja filtra: $Q > 0$ učinkovito za temni impulzni šum (popper), sicer za svetlega (sol)
- za $Q = 0$ dobimo aritmetično povprečje
- za $Q = -1$ pa harmonično povprečje

$$g(x, y) = \frac{\sum_{i=-a}^a \sum_{j=-b}^b (f(x-i, y-j))^{Q+1}}{\sum_{i=-a}^a \sum_{j=-b}^b (f(x-i, y-j))^Q}$$

Glajenje?



Ostrenje slik



- Glavni namen ostrenja (*sharpening*)
 - poudariti prehode med različno svetlimi področji
- Povprečenje je analogno prostorski integraciji
- Ostrenje je analogno prostorski diferenciaciji oz. odvajanju sivinskih vrednosti
 - poudarja robove in druge velike spremembe sivin
 - poudari tudi šum!
 - duži področja s konstantnimi in počasi spreminjajočimi se vrednostmi

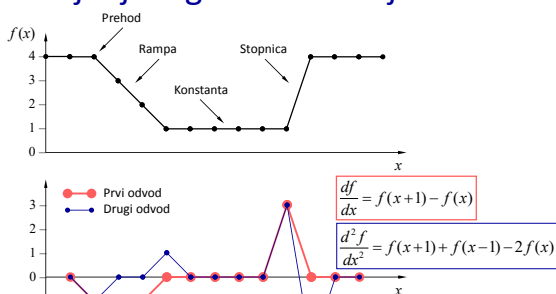


Ostrenje slik



- Odvodi digitalne funkcije so definirani z diferencami
 - za prvi odvod funkcije zahtevamo:
 - da ima vrednost nič na področjih s konstantno vrednostjo ter
 - da ima pozitivno oz. negativno vrednost na področjih, kjer se vrednost funkcije povečuje oz. zmanjšuje
 - za drugi odvod funkcije zahtevamo:
 - da ima vrednost nič na področjih s konstantno vrednostjo in na področjih ramp ter
 - da ima pozitivno oz. negativno vrednost na področjih, kjer se vrednost funkcije neenakomerno povečuje oz. zmanjšuje
- Digitalne funkcije imajo končne vrednosti
 - zato so tudi največje možne spremembe njihovih vrednosti končne
 - najkrajša razdalja na kateri se sprememba lahko zgodi je določena z vzorčno periodo

Odvajanje digitalne funkcije



Prvi odvod 2D slike – gradient

- Za 2D funkcijo $I(u,v)$ ga zapišemo v obliki stolpnega vektorja (linearna operacija):

$$\nabla I(u,v) = \text{grad}[I(u,v)] = \mathbf{g}(u,v) = \begin{bmatrix} g_u(u,v) \\ g_v(u,v) \end{bmatrix} = \begin{bmatrix} \frac{\partial I(u,v)}{\partial u} \\ \frac{\partial I(u,v)}{\partial v} \end{bmatrix}$$

- Dobimo **vektorsko sliko** gradienta $\mathbf{g}(u,v)$, ki kaže v smeri največje spremembe funkcije $I(u,v)$
- Velikost oz. **sliko gradienta** izračunamo kot (nelinearna operacija):

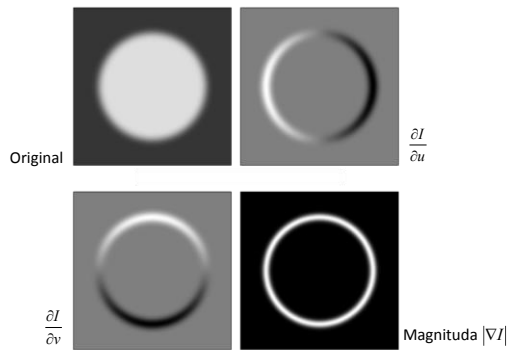
$$G(u,v) = \text{amp}[\nabla I(u,v)] = \sqrt{g_u^2(u,v) + g_v^2(u,v)}$$

- Sobelov operator** →
 - poudarjanje robov
 - razgradnja
 - kvantitativna analiza

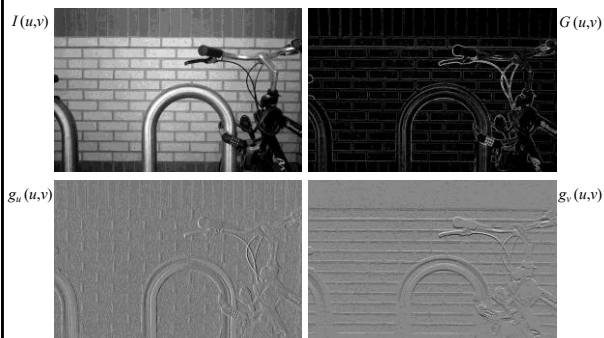
$$g_u = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$g_v = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Prvi odvod 2D slike – gradient

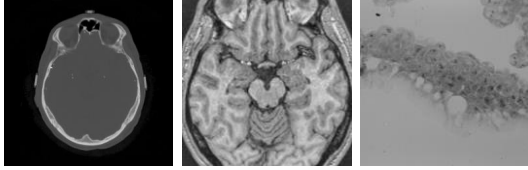


Sobelov operator – primeri

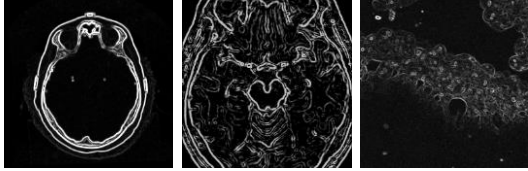


Sobelov operator – primeri

$I(u,v)$



$G(u,v)$



Drugi odvod 2D slike

- Najenostavnejši operator za drugi odvod je **Laplaceov operator**:

$$\nabla^2 I = \frac{\partial^2 I}{\partial u^2} + \frac{\partial^2 I}{\partial v^2}$$

- Za digitalno 2D funkcijo ga izrazimo z diferenciali:

$$\frac{\partial^2 I}{\partial u^2} = I(u+1, v) - 2I(u, v) + I(u-1, v)$$

$$\frac{\partial^2 I}{\partial v^2} = I(u, v+1) - 2I(u, v) + I(u, v-1)$$

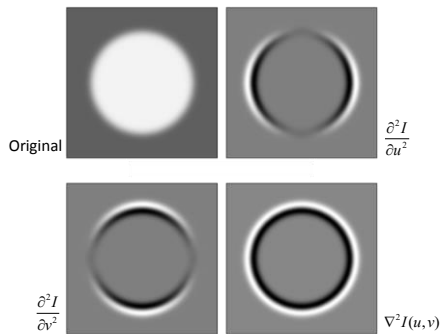
$$\nabla^2 I(u, v) = I(u+1, v) + I(u-1, v) + I(u, v+1) + I(u, v-1) - 4I(u, v)$$

0	1	0
1	-4	1
0	1	0

Navadno dodamo še diagonalne člene za 45° izotropnost →

1	1	1
1	-8	1
1	1	1

Drugi odvod 2D slike



Ostrenje z drugim odvodom

- Vhodni sliki $I(u,v)$ odštejemo uteženo sliko drugega odvoda:

$$I'(u,v) = I(u,v) - c [\nabla^2 I(u,v)]$$

- Konstanta c določa stopnjo ostrenja



Ostrenje z maskiranjem

- Masko dobimo tako, da vhodni sliki $I(u,v)$ odštejemo njeno zglajeno verzijo (*unsharp masking*):

$$m(u,v) = I(u,v) - F(I(u,v))$$

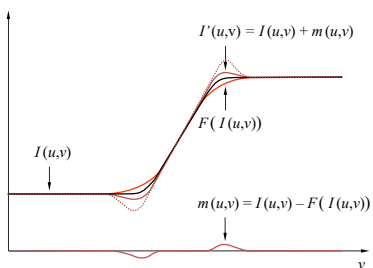
- Sliko maske $m(u,v)$ dodamo vhodni sliki:

$$I'(u,v) = I(u,v) + c \cdot m(u,v)$$

- Stopnjo ostrenja določa konstanta c
- Digitalni filter F pa določa območje ostrenja:
→ lokalno oz. globalno ostrenje (odvisno npr. od σ)



Maskiranje neostrih področij

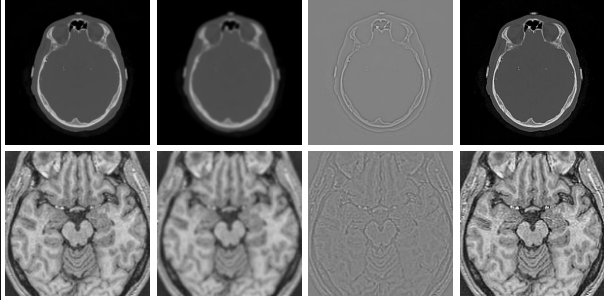


Maskiranje neostrih področij

$I(u,v)$

$F(I(u,v))$

$m(u,v) = I(u,v) - F(I(u,v))$ $I'(u,v) = I(u,v) + m(u,v)$



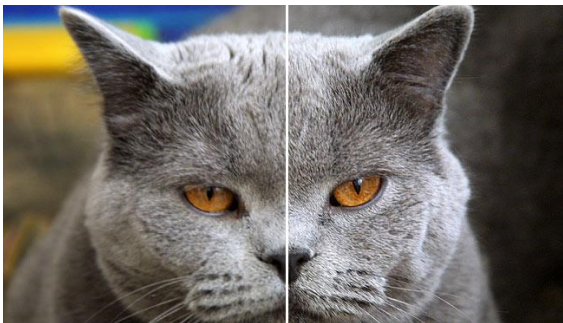
Maskiranje neostrih področij

$I(u,v)$



$I'(u,v)$

Maskiranje neostrih področij



Statistično filtriranje

- Sivinske vrednosti pikselov znotraj filtra **uredimo po velikosti** (od največje do najmanjše)
- **Urejeno statistično filtriranje je nelinearno in robustno:**
 - neobčutljivo na nepravilnosti slik (*outliers*)

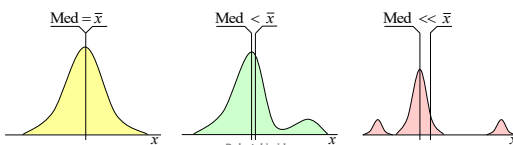


Statistično filtriranje – mediana

- Mediana je sredinska vrednost v urejenem nizu n vrednosti: z_1, z_2, \dots, z_n , polovica vrednosti je manjših, polovica pa večjih

$$\text{Med} = \begin{cases} z_{(n+1)/2} & \text{ko je } n \text{ liho število} \\ \frac{z_{n/2} + z_{n/2+1}}{2} & \text{ko je } n \text{ sodo število} \end{cases} \quad I(u, v) = \text{Med}_{i=a\dots u, j=b\dots v} [I(u-i, v-j)]$$

- Mediana je bolj robustna kot povprečje
- Primernejša za nesimetrične porazdelitve



Statistično filtriranje – mediana

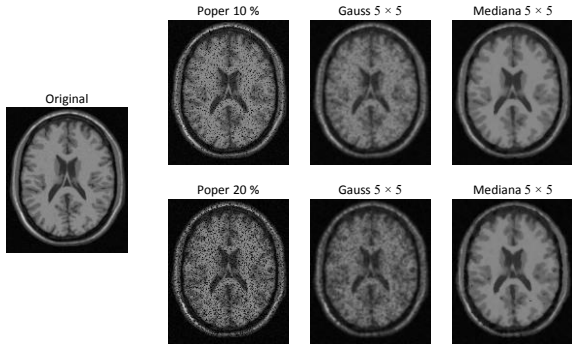
- Osnovna različica in nadgradnje

▪ **Mediana**

$$I(u, v) \leftarrow \text{median} \{ I(u+i, v+j) \mid (i, j) \in R \}$$

▪ **Utežena mediana**
– uteži položaje znotraj področja filtra

Primerjava Gauss – mediana



Primerjava box filter – mediana



Druge statistike

- Filter **maksimalne** in **minimalne** vrednosti:

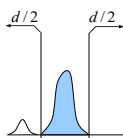
$$I(u, v) = \text{Maks}_{i=-a..a, j=-b..b} [I(u-i, v-j)] \quad I(u, v) = \text{Min}_{i=-a..a, j=-b..b} [I(u-i, v-j)]$$

- Filter **srednje vrednosti** (midpoint filter):

$$I(u, v) = \frac{\text{Maks}_{i=-a..a, j=-b..b} [I(u-i, v-j)] + \text{Min}_{i=-a..a, j=-b..b} [I(u-i, v-j)]}{2}$$

- Filter **popravnih obreznih vrednosti** (alpha-trimmed mean filter)

$$I(u, v) = \frac{1}{MN-d} \sum_{i=-a}^a \sum_{j=-b}^b I_p(u-i, v-j)$$



- Odstranimo ustrezen delež skrajnih vrednosti:
 - za $d = 0 \rightarrow$ dobimo aritmetično povprečje
 - za $d = MN-1 \rightarrow$ dobimo medianin filter

Vpliv Min in Maks filtra



Originalna slika

Slika glajena z
Min filtromSlika glajena z
Maks filtrom

Prilagodljivo filtriranje



- Spreminjamo lastnosti filtrov glede na statistične lastnosti slike (*adaptive filtering*)
- Izboljšamo lastnosti filtrov, poveča pa se kompleksnost filtriranja
- Npr. če poznamo standardno deviacijo šuma σ_n lahko določimo prilagodljivi filter kot:

$$I'(u, v) = I(u, v) - \left(\frac{\sigma_n}{\sigma_f} \right)^2 [I(u, v) - \mu_f]$$

kjer je μ_f povprečje in σ_f standardna deviacija slike znotraj filtra

- Bolj zgladi sliko na homogenih področjih $\rightarrow \sigma_f \approx \sigma_n$
- Manj pa na prehodih med homogenimi področji $\rightarrow \sigma_f > \sigma_n$
- Številni drugi načini prilagodljivega filtriranja:
 - npr. s spreminjanjem velikosti filtra glede na sliko

Morfološko filtriranje

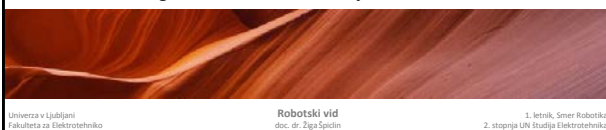
- Vrednost piksla izhodne slike določimo na podlagi pravila, ki ga določa binarni **strukturni element**

$b(u, v)$ s svojo velikost in obliko:

1	1	1	0	1	0	1	1	1	1
1	1	1	1	1	1	1			
1	1	1	0	1	0	1			

0	0	1	0	0
0	1	1	1	0
1	1	1	1	1
0	1	1	1	0
0	0	1	0	0

- Velikost in obliko strukturnega elementa določimo glede na namen filtriranja



Morfološko filtriranje

- Osnovni operaciji morfološkega filtriranja sta **erozija** in **dilacija**, ki ju za vhodno sliko $I(u,v)$ in binarni strukturni element $b(u,v)$ lahko določimo kot:

Erozija E oz. \ominus $(I \ominus b)(u,v) = \min_{i,j} \{I(u-j, v-f) \cdot b(i,j)\}$

Dilacija D oz. \oplus $(I \oplus b)(u,v) = \max_{i,j} \{I(u-j, v-f) \cdot b(i,j)\}$

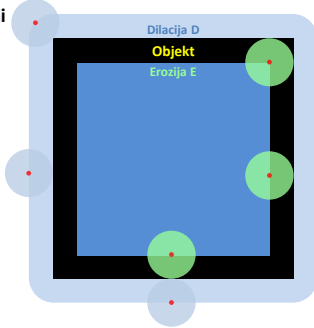
- S kombinacijo erozije in dilacije dobimo operaciji morfološkega **odpiranja** in **zapiranja** kot:

Odpiranje O oz. \circ $I \circ b = (I \ominus b) \oplus b$

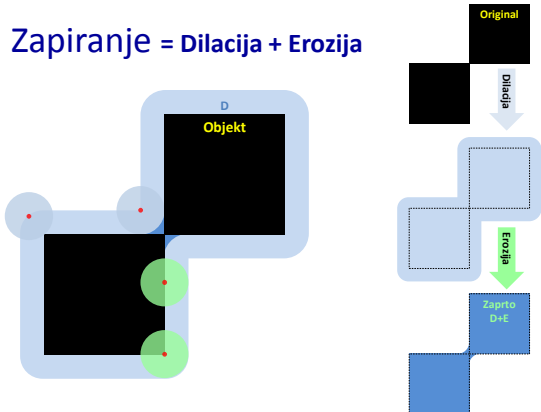
Zapiranje Z oz. \bullet $I \bullet b = (I \oplus b) \ominus b$

Erozija in dilacija

Strukturni element



Zapiranje = Dilacija + Erozija



4. Osnovna obdelava slik 47

Odpiranje = Erozija + Dilacija

Original

Erozija

Dilacija

Odpirto E+D

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

4. Osnovna obdelava slik 48

Zapiranje in odpiranje

Original

Zapiranje

Odpiranje

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

4. Osnovna obdelava slik 49

Morfološko binarno filtriranje

Dilacija D

Original

Erozija E

Zapiranje D+E

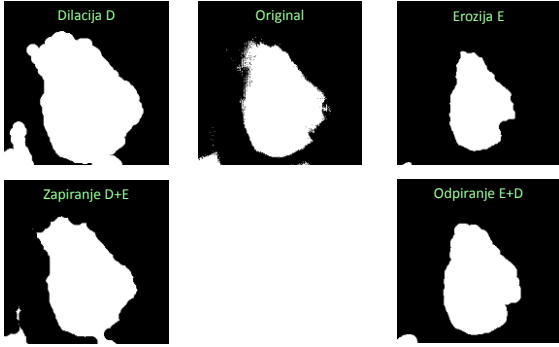
Odpiranje E+D

Univerza v Ljubljani
Fakulteta za Elektrotehniko

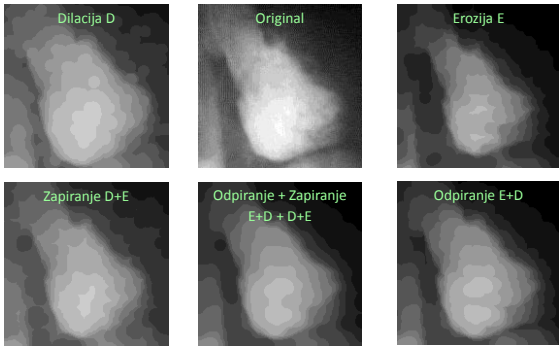
Robotski vid
doc. dr. Žiga Spacin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Morfološko binarno filtriranje



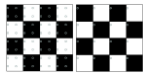
Morfološko sivinsko filtriranje



Morfološko binarno filtriranje



Interpolacija in decimacija



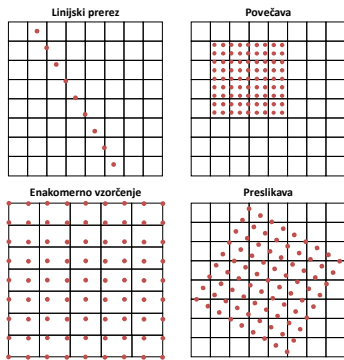
Omogočata spreminjanje vzorčne frekvence

- Z interpolacijo povečamo vzorčno frekvenco
 - omogočimo izvedbo povečav
 - geometrijskih preslikav
 - prerezov, projekcij in upodobitev slik
- Z decimacijo zmanjšamo vzorčno frekvenco
 - zmanjšamo velikost digitalnih slik
 - pohitrimo postopke obdelave in analize slik

Uporabnost interpolacije

Določevanje vrednosti v točkah, ki ne sovpadajo z diskretno mrežo slike:

- tvorjenje prerezov, projekcij in upodobitev
- povečevanje slik = dvig vzorčne frekvence
- enakomerno vzorčenje slik z neenakomernim vzorčenjem
- geometrijske preslikave



Optična in digitalna povečava

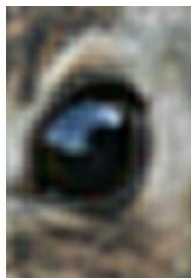
Original



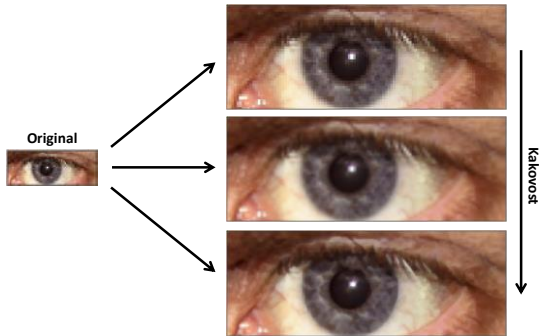
10 x optični zoom



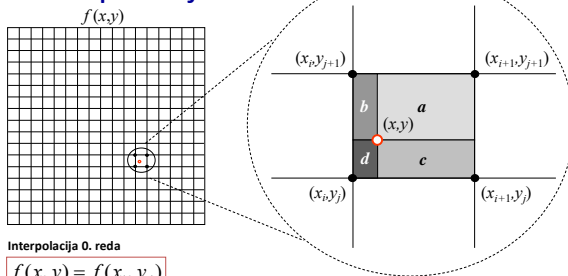
10 x digitalni zoom



Kakovost digitalne povečave



Interpolacija



Interpolacija 0. reda

$$f(x, y) = f(x_i, y_j)$$

Bilinearna interpolacija 1. reda

$$f(x, y) = a \cdot f(x_i, y_j) + b \cdot f(x_{i+1}, y_j) + c \cdot f(x_i, y_{j+1}) + d \cdot f(x_{i+1}, y_{j+1})$$

Interpolacije različnih redov

<p>Diskretne sivinske vrednosti slike</p>	<p>Interpolacija 0. reda</p>	<p>Bilinearna interpolacija 1. reda</p>	<p>Interpolacija z zlepkami 3. reda</p>
--	-------------------------------------	--	--

3D interpolacije:

- 0. reda – najbližji sosed – hitra
- Trilinearna – utežena z volumni
- Trikubična – sistem enačb s 64 neznankami
- Interpolacija z zlepkami

$$f(x, y, z) = \sum_{n=0}^3 \sum_{m=0}^3 \sum_{r=0}^3 a_{nmr} x^n y^m z^r$$

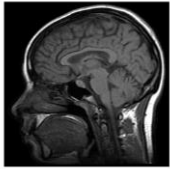
Bikubična interpolacija 3. reda

16 sosednjih točk – sistem enačb

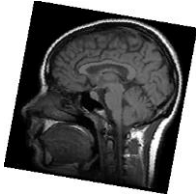
$$f(x, y) = \sum_{n=0}^3 \sum_{m=0}^3 a_{nm} x^n y^m$$

Preslikava z interpolacijo

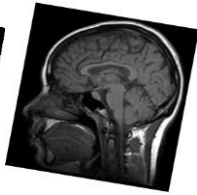
Prvotna slika



Zavrtena slika z interpolacijo 0. reda

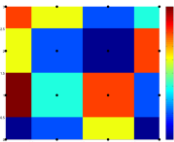


Zavrtena slika z interpolacijo 1. reda

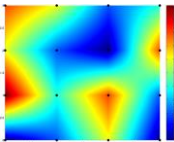


Primerjava interpolacij

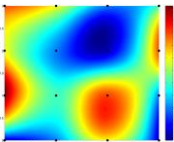
Najbližji sosed - 0. red



Bilinearna interpolacija - 1. red



Bikubična interpolacija - 3. red



Optična in digitalna povečava

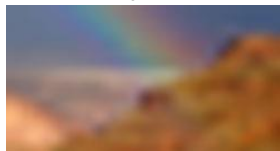
Original



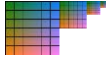
10 x optični zoom



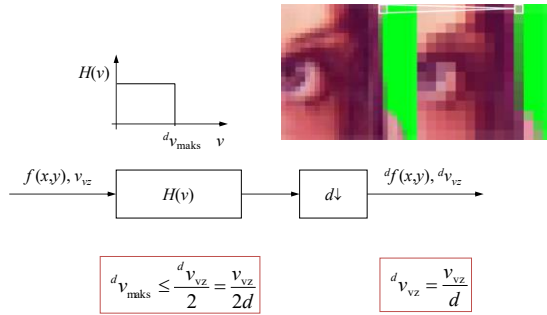
10 x digitalni zoom



Decimacija slik



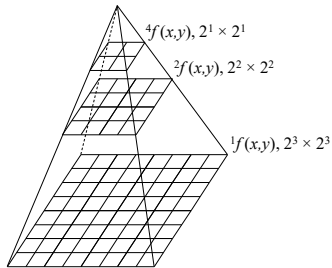
Decimacija slik



Piramidna decimacija



- Večstopenjska decimacija
- Vhodna slika ter decimirane kopije
- Velikost slik pada s potenco števila dva
- Zaporedno filtriranje in opuščanje vsakega drugega vzorca



$$2^n f(x, y) = \sum_{i=-a}^a \sum_{j=-a}^a c(i, j) \cdot f(2x-i, 2y-j)$$

Piramidna decimacija



– Pomembna je izbira koeficientov $c(i,j)$ digitalnega filtra:

- pozitivni,
- unimodalni
- simetrični in
- normalizirani

– Prispevek vseh slikovnih elementov slike na naslednji nivo mora biti enak

$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$
$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$
$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

$\frac{1}{400}$	$\frac{1}{80}$	$\frac{1}{50}$	$\frac{1}{80}$	$\frac{1}{400}$
$\frac{1}{80}$	$\frac{1}{16}$	$\frac{1}{10}$	$\frac{1}{16}$	$\frac{1}{80}$
$\frac{1}{50}$	$\frac{1}{10}$	$\frac{4}{25}$	$\frac{1}{10}$	$\frac{1}{50}$
$\frac{1}{80}$	$\frac{1}{16}$	$\frac{1}{10}$	$\frac{1}{16}$	$\frac{1}{80}$
$\frac{1}{400}$	$\frac{1}{80}$	$\frac{1}{50}$	$\frac{1}{80}$	$\frac{1}{400}$

$$2^n f(x, y) = \sum_{i=-a}^a \sum_{j=-a}^a c(i, j) \cdot f(2x-i, 2y-j)$$

Gaussova piramida $M=5$



128 x 128

64 x 64

32 x 32

16 x 16

8 x 8

Uporaba decimacije

- Velikost slike in s tem računska zahtevnost postopkov digitalne obdelave hitro padata:
 - $z \cdot d^2$ pri 2D in $z \cdot d^3$ pri 3D slikah
- Možna je postopna obdelava slik na različnih nivojih prostorske ločljivosti
- Zmanjševanje vzorčne frekvenca slik:
 - ko je frekvenčna širina slik zelo majhna
 - ko visokofrekvenčna informacija slik ni pomembna (velikokrat predstavlja šum)
- Prihranek pomnilniškega prostora, hitrejši prenos ter obdelava in analiza digitalnih slik

Povzetek

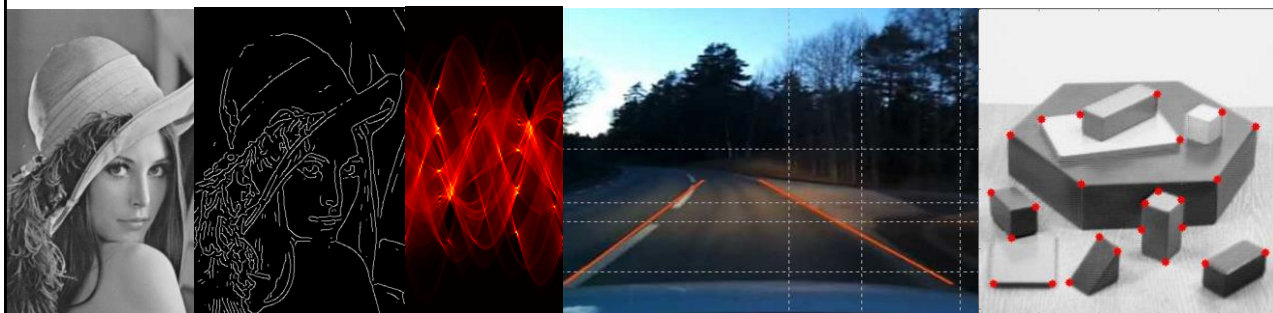
- Filtriranje v prostoru slike je eno od najpomembnejših orodij pri obdelavi slik, uporabno za zmanjševanje šuma, glajenje, zaznavo robov, ostrenje, decimacijo, itd.
 - linearno filtriranje temelji na diskretni konvoluciji z diskretnim jedrom
 - nelinearno filtriranje temelji na nelinearnih operatorjih, kot naprimer mediana, minimum in maksimum
- Morfološko filtriranje temelji na uporabi binarnega strukturnega elementa
 - običajno za filtriranje binarnih (lahko pa tudi sivinskih) slik s postopki erozije, dilacije, odpiranja, zapiranja
- Z interpolacijo in decimacijo spreminjamo vzorčno frekvenco slik (upoštevaj Nyquistov teorem!)
 - interpolacijo uporabljamo za digitalno povečavo slik, preslikave in prereze
 - decimacijo uporabljamo za digitalno pomanjšavo slik



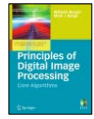
5. Zaznavanje objektov v slikah



Robotski vid



Študijska literatura



Principles of Digital Image Processing: Core Algorithms
 Wilhelm Burger, Mark J. Burge, Springer, 2009
 3. Detecting simple curves, 4. Corner detection

Biomedicinska slikovna informatika in diagnostika
 Boštjan Likar, Založba FE, 2008
 5.1.2 Razgradnja na podlagi odvodov,
 5.1.3.3 Globalno povezovanje robov
 s Houghovo transformacijo



Computer vision: algorithms and applications
 Richard Szeliski, Springer, 2010
 4. Feature detection and matching

Concise Computer Vision
 Reinhard Klette, Springer, 2014
 9.1 Invariance, Features, and Sets of Features,
 9.2 Examples of Features



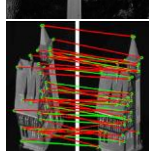
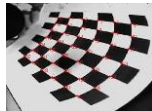
Vsebina



- Zaznavanje oslonilnih točk
- Zaznavanje in povezovanje robov
 - Cannyev detektor
- Zaznavanje parametričnih in neparam. oblik
 - Houghova preslikava
- Zaznavanje objektov s strojnim učenjem
 - Konvolucijske nevronske mreže

Zaznavanje oslonilnih točk

- Oslonilne točke so „robustne“ lokacije v sliki, ki jih lahko uporabljamo za
 - sledenje objektov,
 - poravnavo slik,
 - geometrijska merjenja,
 - kalibracijo sistemov za vizualno kontrolo kakovosti, itd.
- Primer oslonilnih točk so oglišča
 - v oglišču se intenziteta razlikuje od intenzitete sosednjih slikovnih elementov v več kot eni smeri
 - oglišča lahko hitro zaznamo, avtomatsko iskanje pa ni enostavno



Zaznavanje oglišč

- Iz gradienta v točki (u, v)

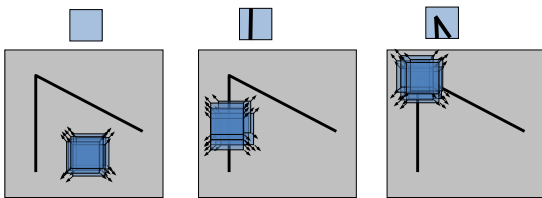
$$|\nabla I|(u, v) = \sqrt{\left(\frac{\partial I}{\partial u}(u, v)\right)^2 + \left(\frac{\partial I}{\partial v}(u, v)\right)^2}$$

ne moremo sklepati ali je točka oglišče ali ne

- potrebujemo statistiko gradientov znotraj majhnega okna
- Glede na statistiko ločimo
 - konstanto**: majhne spremembe intenzitet
 - rob**: velike spremembe intenzitet v eni smeri
 - oglišče**: velike spremembe intenzitet v dveh ortogonalnih smereh



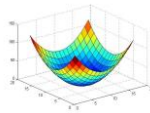
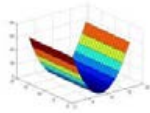
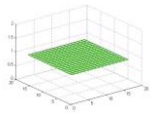
Harrisov detektor oglišč: Osnovna ideja



"konstantno" področje:

"rob":

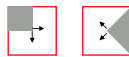
"oglišče":



Harrisov detektor oglišč

- Po intuiciji bosta oba parcialna odvoda

$$I_x(u, v) = \frac{\partial I}{\partial x}(u, v) \quad \text{in} \quad I_y(u, v) = \frac{\partial I}{\partial y}(u, v)$$



v oknu, kjer je oglišče, velika

- Za vsako točko (u, v) na sliki izračunamo vrednosti

$$\begin{aligned} A(u, v) &= I_x^2(u, v) \\ B(u, v) &= I_x I_y(u, v) \\ C(u, v) &= I_x(u, v) I_y(u, v) \end{aligned} \quad \text{matrika } M(u, v) \quad \rightarrow \quad M = \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

- Lastne vrednosti matrike $M(u, v)$ bodo sorazmerne spremembi intenzitete v dveh medsebojno ortogonalnih smereh slike

Harrisov detektor oglišč

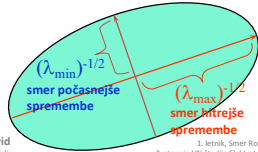
- Okno določa jedro, s katerim gladimo vrednosti $A(u, v)$, $B(u, v)$ in $C(u, v)$
 - če uporabimo Gaussov filter $H^{G, \sigma}$, potem σ določa velikost okna

$$\bar{M} = \begin{pmatrix} A * H^{G, \sigma} & C * H^{G, \sigma} \\ C * H^{G, \sigma} & B * H^{G, \sigma} \end{pmatrix} = \begin{pmatrix} \bar{A} & \bar{C} \\ \bar{C} & \bar{B} \end{pmatrix}$$

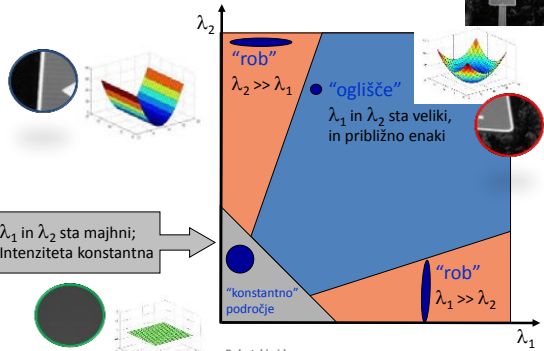
- Nato določimo lastni vrednosti matrike

$$g(\lambda) = |M - \lambda I| = \begin{vmatrix} \bar{A} - \lambda & \bar{C} \\ \bar{C} & \bar{B} - \lambda \end{vmatrix} = \lambda^2 - (\bar{A} + \bar{B})\lambda - \bar{C}^2 + \bar{A}\bar{B} = 0$$

$$\lambda_{1,2} = \frac{(\bar{A} + \bar{B}) \pm \sqrt{(\bar{A} - \bar{B})^2 + 4\bar{C}^2}}{2}$$



Harrisov detektor oglišč



Detektorji oglišč na osnovi λ_1 in λ_2

- Glede na minimalno velikost λ_1 in λ_2

$$Q(u, v) = \min(\lambda_1, \lambda_2)$$

- Kot kvocient produkta in vsote

$$Q(u, v) = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} = \frac{\det M}{\text{trace } M}$$

- Harrisov detektor oglišč

$$Q(u, v) = \det M - \kappa(\text{trace } M)^2 = \lambda_1 \lambda_2 - \kappa(\lambda_1 + \lambda_2)^2$$

Harrisov detektor oglišč

- Z analizo lastnih vrednosti določimo lastnosti okna

$$\lambda_{1,2} = \frac{(\bar{A} + \bar{B}) \pm \sqrt{(\bar{A} - \bar{B})^2 + 4\bar{C}^2}}{2}$$

- Ker je *bila leta 1988* operacija korenjenja „potratna“ je Harris predlagal računsko preprostejšo **kriterijsko funkcijo za merjenje izrazitosti oglišča**

$$Q(u, v) = (\bar{A}\bar{B} - \bar{C}^2) - \kappa(\bar{A} + \bar{B})^2 \quad \kappa \in [0 - 0,25]$$

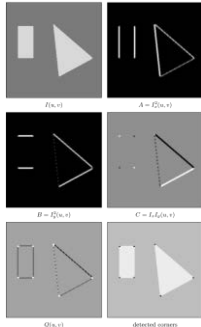
- tipično je κ okoli 0,05
- večji kot je κ , manj občutljiv je detektor in manj oglišč bo odkril

- Oglišča najdemo kot uprakovljene lokalne maksimume

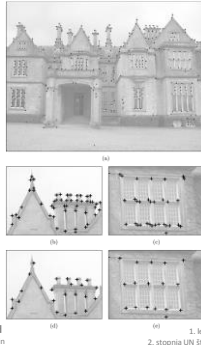
$$Q(u, v) > t_H, \quad t_H \in [10.000 - 1.000.000]$$

Harrisov detektor oglišč – primeri

- Sintetična slika



- Realna slika



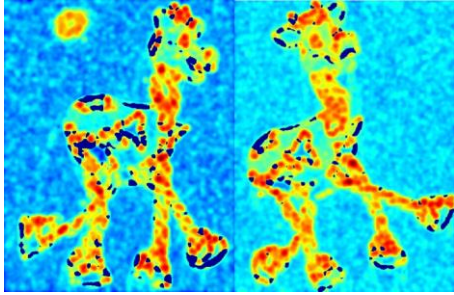
Harrisov detektor oglišč – primeri

- Primer iskanja soležnih točk (npr. za namen poravnave objektov)



Harrisov detektor oglišč – primeri

- Primer iskanja soležnih točk (npr. za namen poravnave objektov)



$Q(u,v)$

Harrisov detektor oglišč – primeri

- Primer iskanja soležnih točk (npr. za namen poravnave objektov)



$Q(u,v) > t_H$


Harrisov detektor oglišč – primeri

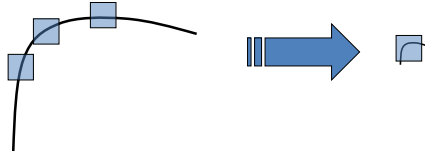
- Primer iskanja soležnih točk (npr. za namen poravnave objektov)



Lokalne
maksimalne
vrednosti
 $Q(u,v)$

Harrisov detektor oglišč – lastnosti

- Rotacijska invarianca → DA 
- Invarianca na velikost (skalo) → NE

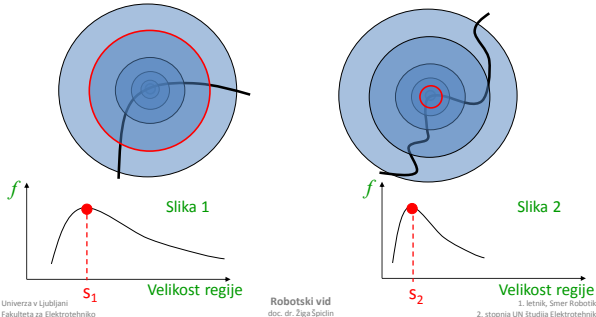


Vse točke bodo
zaznane kot robovi

Čeprav „na drugačni
skali“ izgleda kot oglišče!

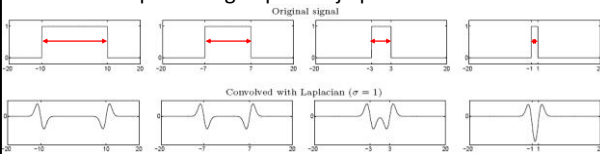
Avtomska izbira skale – intuicija

- Poišči skalo pri lokalno maksimalnem odzivu neke funkcije f glede na položaj in skalo

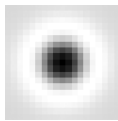
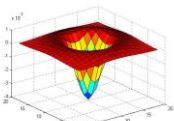


Avtomska izbira skale – Laplace

- Odziv Laplaceovega operatorja pri $\sigma = 1$



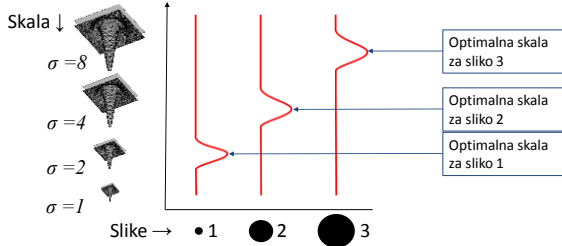
- Primer 2D jedra Laplaceovega operatorja



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

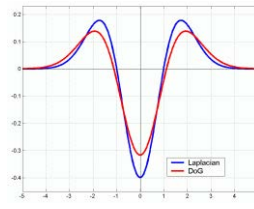
Avtomatska izbira skale – LoG

- Laplace-of-Gaussian (LoG) operator
 - Filtriranje z Laplacevim operatorjem, ki je glajen z različnimi Gaussovimi jedri glede na σ



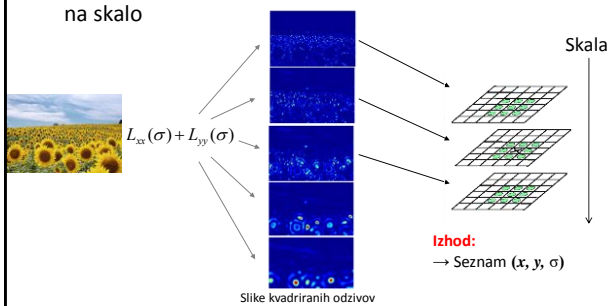
Avtomatska izbira skale – LoG/DoG

- LoG
 - $L = \sigma^2 (L_{xx}(x, y, \sigma) + L_{yy}(x, y, \sigma))$
(Laplacian-of-Gaussian)
- DoG
 - $DoG = G(x, y, k\sigma) - G(x, y, \sigma)$
(Difference-of-Gaussian)
 - Izračun z zaporednim filtriranjem



Od skale neodvisne oslonilne točke

- Lokalni maksimumi tako glede na položaj v sliki kot na skalo



Od skale neodvisne oslonilne točke

▪ Primer



Skala → ○ ○ ○ ○

Primerjava lastnosti detektorjev oslonilnih točk

Table 7.1 Overview of feature detectors.

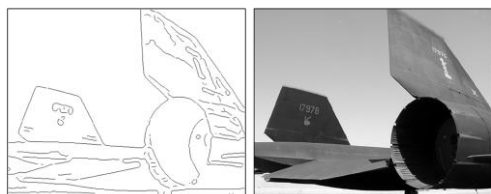
Feature Detector	Corner	Blob	Region	Rotation invariant	Scale invariant	Affine invariant	Repeatability	Localization accuracy	Robustness	Efficiency
Harris	✓			✓			+++	+++	+++	++
Hessian	✓	✓		✓			++	++	++	+
SUSAN				✓			++	++	++	+++
Harris-Laplace	✓	(✓)		✓	✓		+++	+++	++	+
Hessian-Laplace	(✓)	✓		✓	✓		+++	+++	+++	+
DOG	(✓)	✓		✓	✓		++	++	++	++
SURF	(✓)	✓		✓	✓		++	++	++	+++
Harris-Affine	✓	(✓)		✓	✓	✓	+++	+++	++	++
Hessian-Affine	(✓)	✓		✓	✓	(✓)	+++	+++	+++	++
Solent Regions	(✓)	✓		✓	✓		+	+	++	+
Edge-based	(✓)	✓		✓	✓		+++	+++	++	+
MISER		✓		✓	✓	✓	+++	+++	++	+++
Intensity-based		✓		✓	✓	✓	++	++	++	++
Superpixels		(✓)		✓	(✓)	✓	+	+	+	+

- Vsi detektorji so precej robustni
- Izbira zavisi od konkretne aplikacije

[Tinne Tuytelaars and Krystian Mikolajczyk \(2008\), "Local Invariant Feature Detectors: A Survey", Foundations and Trends® in Computer Graphics and Vision: Vol. 3: No. 3, pp 177-280.](#)

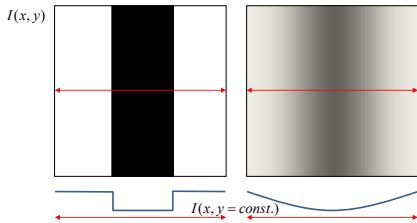
Zaznavanje robov

- Robovi igrajo pomembno vlogo pri človekovem vidnem zaznavanju
 - hitro opazni, z njimi lahko opišemo ali rekonstruiramo objekt



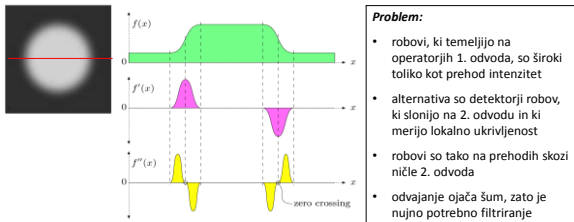
Zaznavanje robov

- Rob na sliki sestavlja skupina točk, v katerih se lokalna intenziteta močno spremeni vzdolž neke smeri
 - močnejša je sprememba intenzitete, močnejši je rob



Zaznavanje robov

- Velikost spremembe funkcije določa njen prvi odvod – na tem konceptu temeljijo tudi detektorji robov



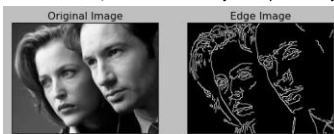
Problem:

- robovi, ki temeljijo na operatorjih 1. odvoda, so široki toliko kot prehod intenzitet
- alternativa so detektorji robov, ki slonijo na 2. odvodu in ki merijo lokalno ukrivljenost
- robovi so tako na prehodih skozi ničle 2. odvoda
- odvajanje ojača šum, zato je nujno potrebno filtriranje

Cannyev detektor robov



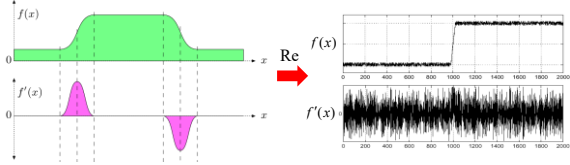
- Enega najbolj uporabljanih detektorjev robov je leta 1983 v okviru magistrske naloge na MIT razvil **John Canny**
- Pri razvoju detektorja robov je zasledoval tri cilje:
 - malo napak:** detektor naj bi našel večino robnih točk in malo takih točk, ki ne predstavljajo roba
 - dobra lokalizacija:** robne točke naj bi bile čim bližje dejanskim robovom
 - ena točka / rob:** vsak rob naj bi bil predstavljen z (širok) eno točko



J. Canny, A Computational Approach To Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8:6 79-714, 1986.

Cannyev detektor robov – gradniki

- Vzemimo 1D stopničast rob, ki je popačen z Gausovim šumom

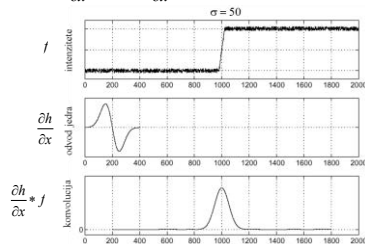


Kje je rob?

- **Rešitev:**
 - najprej zgladi sliko, potem odvajaj
 - rob je tam, kjer so vrhovi funkcije: $\frac{d}{dx}(h * f)$

Cannyev detektor robov – gradniki

- Združimo glajenje in odvajanje z upoštevanjem asociativnosti konvolucije: $\frac{\partial}{\partial x}(h * f) = \frac{\partial h}{\partial x} * f$

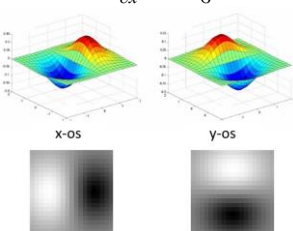


Cannyev detektor robov – gradniki

- Dobra aproksimacija detektorja 1D stopničastih robov, ki so popačeni z Gausovim šumom, je prvi odvod Gausove funkcije

$$h(x) = G(x) = e^{-\frac{x^2}{2\sigma^2}} \rightarrow \frac{\partial}{\partial x} e^{-\frac{x^2}{2\sigma^2}} = -\frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$$

- Razširitev na 2D
 - dobimo konvolucijski jedri \rightarrow
 - alternativno lahko uporabimo Gaussovo glajenje in filtriranje z operatorjem odvoda kot npr. Roberts, Prewitt, Sobel, itd.

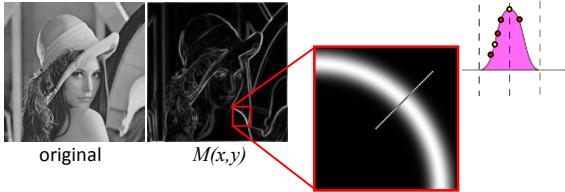


Cannyev detektor robov – gradniki

- Določimo magnitudo in orientacijo 1. odvoda v vsaki točki slike

$$|\nabla I(x, y)| = M(x, y) = \sqrt{g_x^2 + g_y^2} \quad \Phi(x, y) = \arctg \left[\frac{g_y}{g_x} \right]$$

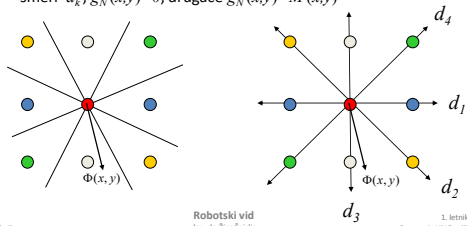
- Nato odstranimo **nemaksimalne vrednosti** v smeri odvoda



Cannyev detektor robov – gradniki

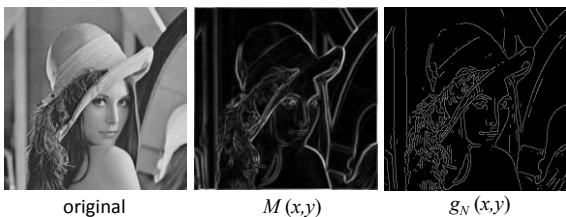
- Diskretna implementacija odstranjevanja nemaksimalnih vrednosti v smeri $\Phi(x, y)$

- Poišči smer d_k , ki je najbližja smeri $\Phi(x, y)$
- Če je vrednost $M(x, y)$ manjša od najmanj ene sosednje vrednosti v smeri d_k , $g_N(x, y) = 0$, drugače $g_N(x, y) = M(x, y)$



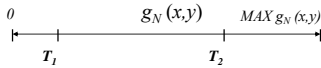
Cannyev detektor robov – gradniki

- Odstranjevanje nemaksimalnih vrednosti – primer:



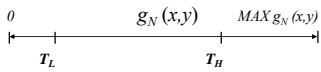
Cannyev detektor robov – gradniki

- Odstranjevanje napačnih robnih točk z upravljanjem $M(x, y)$ v točkah nemaksimalnih vrednosti $g_N(x, y)$



T_1 : nismo odstranili vseh napačnih robnih točk!
 T_2 : odstranili smo tudi prave robne točke!

- Upragovljanje s histerezo



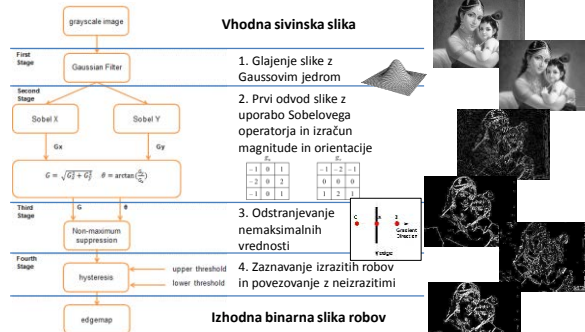
Canny: $\frac{T_L}{T_H} = \frac{1}{2}$ ali $\frac{1}{3}$

Poiščemo izrazite (NH) in neizrazite robove (NL):
 $g_{NH}(x, y) = g_N(x, y) \geq T_H$
 $g_{NL}(x, y) = g_N(x, y) \geq T_L$
 $g_M(x, y) = g_{NL}(x, y) - g_{NH}(x, y)$

Cannyev detektor robov – gradniki

- Povezovanje robov

Cannyev detektor robov – algoritem



Cannyev detektor robov – primeri

- Vpliv stopnje glajenja σ na zaznavanje robov



Original

Canny $\sigma = 1$

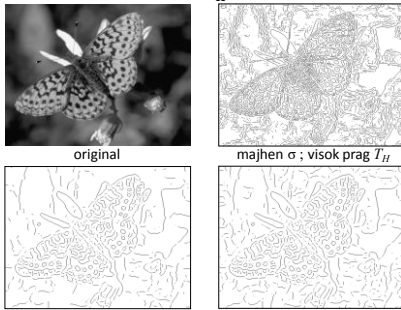
Canny $\sigma = 2$

Večji $\sigma \rightarrow$ detekcija izrazitejših robov (manj robov)

Manjši $\sigma \rightarrow$ detekcija tudi manj izrazitih robov (več robov)

Cannyev detektor robov – primeri

- Vpliv izbire σ in praga T_H



original

majhen σ ; visok prag T_H

velik σ ; visok prag T_H

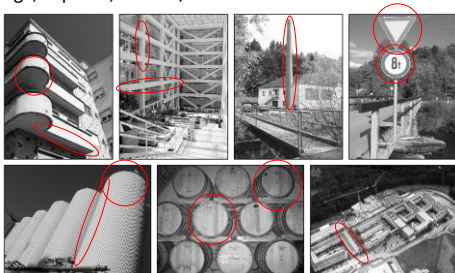
velik σ ; nizek prag T_H

Ali lahko na podlagi slike robov odgovorimo na naslednji vprašanji:

- Kje je metulj?
- Kaj je na sliki?

Zaznavanje umetnih objektov

- Motivacija:** umetni objekti, ki jih je naredil človek so ravni, okrogli, elipsasti, trikotni, kvadratni...



Zaznavanje parametričnih objektov

- **Problem:** kako na osnovi binarne slike robov poiskati umetne objekte določenih oblik – npr. parametrične krivulje?



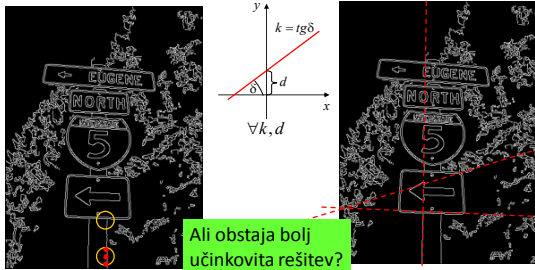
Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Ziga Spacín

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Zaznavanje parametričnih objektov

- **Rešitev 1:** lokalno sledenje konture objekta – izberi točko na binarni sliki robov, preglej njeno sosesčino in dodaj točke v sosesčini
- **Rešitev 2:** globalno izčrpano iskanje krivulj – npr. iščemo premice $y = kx + d \quad \forall k, d$, na katerih leži največ robnih točk



Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Ziga Spacín

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Zaznavanje parametričnih objektov

- Izberi **parametrični model objekta**, ki predstavlja skupek robov
- Pripadnost roba oceni globalno
 - nemogoče je ugotoviti „ali točka pripada modelu“ na podlagi opazovanja le te točke
- Ustreznost modela ne moremo preveriti za vse kombinacije robov → **uporabimo pristop z glasovanjem**
 - preletimo vse robove in pripišemo glasove ustreznim parametrom modela
 - izločimo parametre modela, ki dobi zadostno število glasov
- Z glasovanjem ublažimo negativen vpliv šuma in prekrivanja
 - robovi kot posledica šuma ali prekrivajočih objektov bodo neskladni z večino robov, ki pripadajo umetnim objektom
 - objekt zaznamo, tudi če je le delno viden

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Ziga Spacín

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Houghova preslikava

- **Robove preslika v parametrični prostor**
 - za enostavne krivulje kot so premice, krogi, elipse
 - posplošena različica za objekte poljubnih oblik
- **Kratka zgodovina:**
 - 1962: Paul Hough – US Patent 3,069,654
 - 1972: algoritem opisan v knjigi



Pattern Classification
Richard O. Duda, Peter E. Hart
Wiley 1973 (2nd edition 2000)



Richard Duda Peter Hart

Houghova preslikava – premice

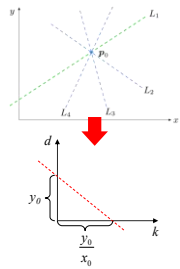
- Vsaka **premica** $L_j = \langle k_j, d_j \rangle$, ki poteka skozi točko $p_0 = (x_0, y_0)$, je opisana z enačbo: $L_j : y_0 = k_j x_0 + d_j$

$$L_j : y_0 = k_j x_0 + d_j$$



$$d_j = -x_0 k_j + y_0$$

$k_j, d_j \rightarrow$ spremenljivki
 $x_0, y_0 \rightarrow$ konstantna parametra



Houghova preslikava – premice

- Za poljubno točko $p_i = (x_i, y_i)$ enačba

$$M_i : d = -x_i k + y_i$$

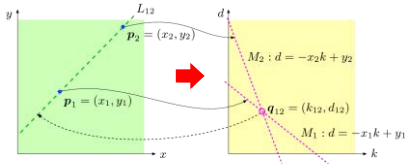
opisuje premico s parametroma x_i, y_i v **parametričnem prostoru**

- Temu prostoru pravimo **Houghov prostor** in je določen s koordinatama k in d

Prostor slike (x, y)		Parametrični prostor (k, d)	
točka	$p_i = (x_i, y_i)$	$M_i : d = -x_i k + y_i$	premica
premica	$L_j : y = k_j x + d_j$	$q_j = (k_j, d_j)$	točka

Houghova preslikava – premice

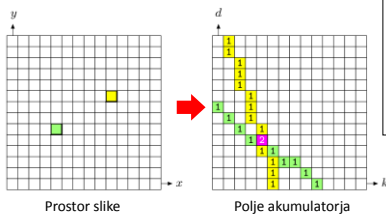
- V *parametričnem* prostoru nas zanimajo točke, kjer se dve ali več premic seka



- Če se v neki točki (k^*, d^*) *parametričnega* prostora seka N premic, potem N točk v prostoru slike leži na premici $y = k^*x + d^*$

Houghova preslikava – premice

- Parametrični prostor diskretiziramo v 2D polje
 - ovrednotimo lahko le končno mnogo modelov!
 - pravimo mu tudi akumulator**



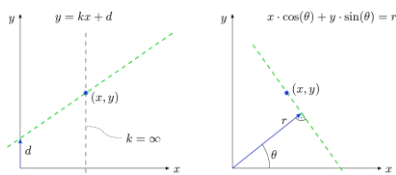
Problem: premica v parametrični obliki
 $L_j : y = k_j x + d_j$
 ima parametra k in d v neskončnih intervalih!

Houghova preslikava – premice

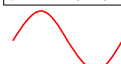
- Namesto eksplicitne oblike enačbe premice $y = kx + d$ je boljša *normalna ali Hessejeva* oblika enačbe premice

$$x \cdot \cos(\theta) + y \cdot \sin(\theta) = r \quad \text{kjer velja} \quad \begin{aligned} 0 \leq \theta \leq 2\pi \\ 0 \leq r \leq r_{\max} \end{aligned}$$

- Ta zapis omogoča linearno kvantizacijo kota θ in polmera r



Točka (x_j, y_j) predstavlja sinusoido v parametričnem prostoru (r, θ)



Houghova preslikava – premice

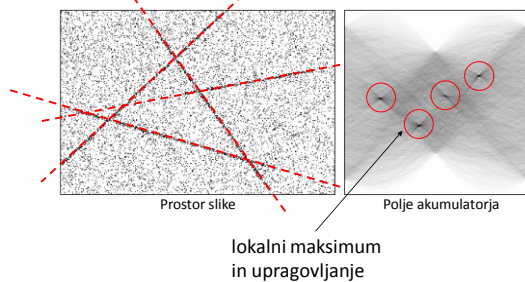
Algoritem:

```

1: HOUGHLINES(I)
   Returns the list of parameters  $\langle \theta_i, r_i \rangle$  corresponding to the strongest
   lines found in the binary image I.
2: Set up a two-dimensional array  $Acc[\theta, r]$  of counters, initialize to 0.
3: Let  $(u_c, v_c)$  be the center coordinates of the image I
4: for all image coordinates  $(u, v)$  do
5:   if I(u, v) is an edge point then
6:     Get coordinate relative to the image center  $(u_c, v_c)$ :
7:      $(x, y) \leftarrow (u - u_c, v - v_c)$ 
8:     for  $\theta_i = 0 \dots \pi$  do
9:        $r_i = x \cdot \cos(\theta_i) + y \cdot \sin(\theta_i)$ 
10:      Increment  $Acc[\theta_i, r_i]$ 
11: Return the list of parameter pairs  $\langle \theta_j, r_j \rangle$  for K strongest lines:
12: MaxLines ← FINDMAXLINES(Acc, K)
13: return MaxLines.
  
```

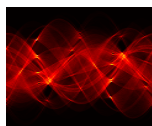
Houghova preslikava – premice

Sintetični primer:



Houghova preslikava – premice

Realni primer:



Parametrični
prostor

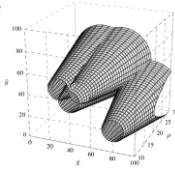
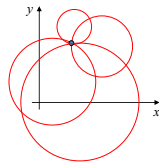
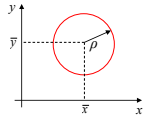
Kako določimo
položaj krajišč daljce?

Houghova preslikava – krožnice

- Enačba krožnice v kartezičnih koordinatah

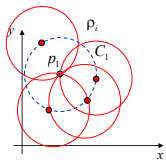
$$(x - \bar{x})^2 + (y - \bar{y})^2 = \rho^2$$

- Houghova preslikava krožnic zahteva **tridimenzionalni parametrični prostor** (\bar{x}, \bar{y}, ρ) , v katerem iščemo središče in polmer krožnice

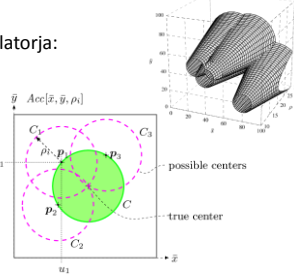


Houghova preslikava – krožnice

- Primer izgradnje polja akumulatorja:



$$(x - \bar{x})^2 + (y - \bar{y})^2 = \rho^2$$



prerez skozi 3D akumulator pri radiju ρ

Houghova preslikava - primer

- Zaznavanje premic in krogov

<https://www.youtube.com/watch?v=UUMvx9bMWRY>



Houghova preslikava - primer

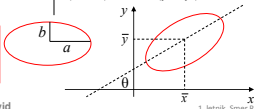
- Uporaba: zaznavanje voznega pasu
<https://www.youtube.com/watch?v=IE2SlzDdFM0>



Houghova preslikava za analitične krivulje

Analična oblika	Parametri	Enačba
Premica	θ, r	$x \cdot \cos(\theta) + y \cdot \sin(\theta) = r$
Krogla	\bar{x}, \bar{y}, ρ	$(x - \bar{x})^2 + (y - \bar{y})^2 = \rho^2$
Parabola	$\bar{x}, \bar{y}, \rho, \theta$	$(y - \bar{y})^2 = 4\rho(x - \bar{x})$
Elipsa	$\bar{x}, \bar{y}, a, b, \theta$	$(x - \bar{x})^2 / a^2 + (y - \bar{y})^2 / b^2 = 1$

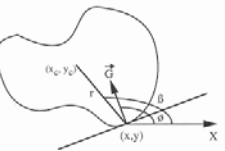
Z naraščanjem števila parametrov postane Houghova preslikava računsko in spominsko zelo zahtevna!



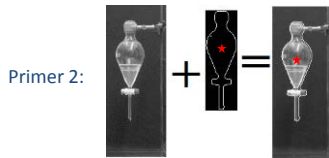
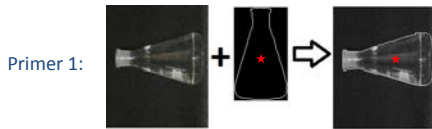
Posplošena Houghova preslikava

- Omoogoča zaznavanje poljubnih oblik
- Celotno obliko je potrebno zakodirati v R-tabelo
 - izberemo referenčno točko (x_c, y_c)
 - določimo vektor od (x_c, y_c) do roba objekta
 - shranimo informacijo o dolžini in orientaciji r, β vektorja v R-tabelo pri indeksu ϕ , ki predstavlja orientacijo gradienta G v točki roba
 - za posamezen indeks je v R-tabeli lahko več vnosov

$\phi_1 = 0$	$(r, \beta)_{1k_1}$	$(r, \beta)_{1k_2}$	\dots	$(r, \beta)_{1k_{n_1}}$
\dots	\dots	\dots	\dots	\dots
ϕ_j	$(r, \beta)_{jk_1}$	$(r, \beta)_{jk_2}$	\dots	$(r, \beta)_{jk_{n_j}}$
\dots	\dots	\dots	\dots	\dots
$\phi_k = \pi$	$(r, \beta)_{kk_1}$	$(r, \beta)_{kk_2}$	\dots	$(r, \beta)_{kk_{n_k}}$



R-tabela kodira preslikavo iz roba objekta v njegov center



Posplošena Houghova preslikava

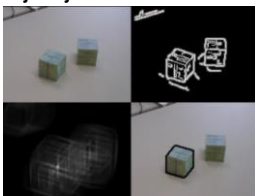
- **Algoritem** za iskanje centra objekta (x_c, y_c) pri danih robovih (x_i, y_i, ϕ_i)
 1. Ustvari polje akumulatorja $A(x_c, y_c)$ in ponastavi kot $A(x_c, y_c) = 0, \forall (x_c, y_c)$
 2. Za vsako točko roba (x_i, y_i, ϕ_i)
 - Za vsak vnos r_k^i v R-tabeli izračunaj

$$x_c = x_i + r_k^i \cos \beta_k^i$$

$$y_c = y_i + r_k^i \sin \beta_k^i$$
 - in povečaj vrednost akumulatorja $A(x_c, y_c) = A(x_c, y_c) + 1$
 3. Poišči lokalne maksimume v $A(x_c, y_c)$

Posplošena Houghova preslikava

- **Velikost in orientacijo oblike** lahko zaznamo z manipulacijo vnosov v R-tabeli
 - za skaliranje s faktorjem s pomnoži vektorje v R-tabeli s s
 - za rotacijo s kotom θ , rotiraj vektorje v R-tabeli za kot θ
- **Primer zaznavanja objektov:**

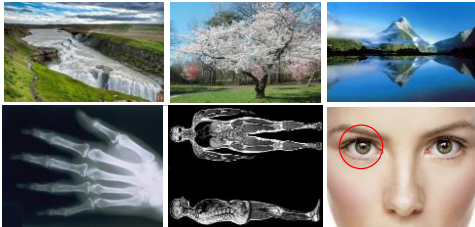


Houghova preslikava – povzetek

- **Splošne prednosti Houghove preslikave so, da:**
 - omogoča zaznavanje objektov,
 - je robustna na delno prekrivanje iskane oblike,
 - odporna na šum v sliki robov in
 - lahko zazna več primerkov oblike v sliki v enem samem izračunu
- **Glavni slabosti pa sta računska zahtevnost in zahteva po velikem pomenu za akumulator**
 - dimenzija narašča s številom parametrov/elementov oblike

V primerjavi z umetnimi so naravni objekti bolj kompleksni

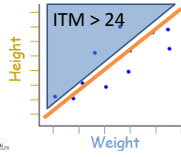
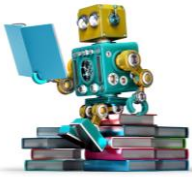
- V naravi redko najdemo ravne črte, krožnice, elipse,...
- Objekti so določeni tudi z barvo, teksturo



Zaznavanje objektov – ključna vprašanja

- Katere značilnice uporabiti?
 - sivine, robove, barvo, teksturo,... vse?
- Katere so optimalne značilnice?
 - omogočajo visoko občutljivost in specifičnost zaznave
- Ali se lahko algoritem razpoznavanja sam nauči, katere značilnice so optimalne?
 - če podam primere, tj. zbirko slik z oznakami objektov
- Kako dobro poslušuje znanje tak algoritem?
 - ali deluje, če dam novo, še nevideno sliko?
- Kateri so ti algoritmi? → **strojno učenje**

Strojno učenje ang. machine learning



- Stroj se učijo sposobnosti izvajanja nalog iz opažanj
- V osnovi so to statistični principi, postopki in algoritmi za učenje in napovedovanje na osnovi preteklih opažanj (podatkov)
- Primer: **regresija in razvrščanje**

Algoritmi strojnega učenja



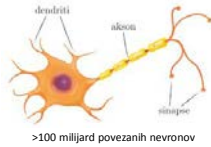
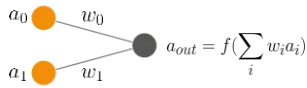
Globoko učenje ang. machine learning

- Podpodročje strojnega učenja
 - osnovano na nevronskih mrežah
- Preboj v letu 2012 pri razpoznavi kategorije objektov na barvnih slikah javne zbirke ImageNet
 - 1000 kategorij, 1 milijon slik
 - napaka enaka človeški (zdaj že manjša, <2%)



Umetne nevrnske mreže

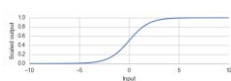
- Sestojijo iz povezanih enot a



- Povezanost enot določajo uteži w
- Aktivacija enote a_{out} zavisi od
 - aktivacije vhodnih povezanih enot a_0, a_1 ,
 - vrednosti uteži w_0, w_1 in
 - aktivacijske funkcije

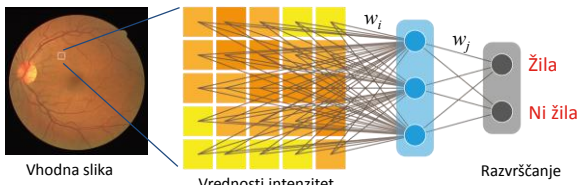
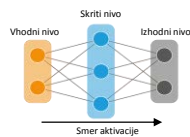
sigmoidna funkcija

$$f(x) = \frac{1}{1 + e^{-x}}$$



Večplastni perceptron

- Enote so povezane v več nivojih
 - vhodni nivo vsebuje značilnice vzorca
 - skriti nivo kombinira značilnice
 - izhodni nivo določa izhodne vrednosti
- Z „učenjem“ določimo vrednosti uteži w posameznih povezav

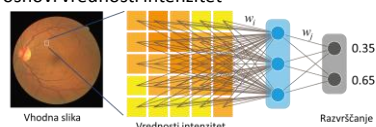


Tradicionalno vs. globoko strojno učenje

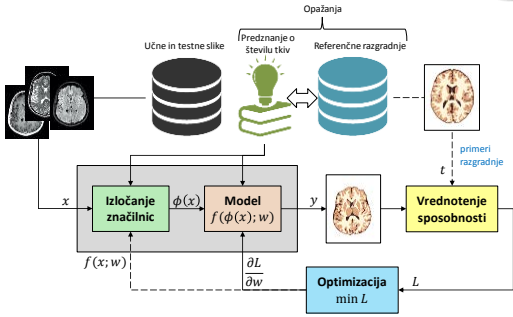
- Tradicionalno:** razvrščanje z ročno določenimi značilnicami za vsak košček slike



- Globoko:** Razvrščevalnik se nauči najbolj smiselno predstavitev le na osnovi vrednosti intenzitet

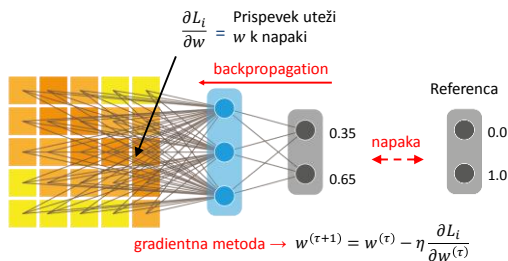


Učenje



Učenje nevronske mreže

- Optimiziramo funkcijo napake L_i



gradientna metoda $\rightarrow w^{(\tau+1)} = w^{(\tau)} - \eta \frac{\partial L_i}{\partial w^{(\tau)}}$

Funkcije napake L_i

- Za vzorec i in razvrstitev p_i ter referenčno oznako t_i naj bo
 - za **regresijo** - zvezne izhodne vrednosti
 - srednja kvadratna napaka

$$L_i = (t_i - p_i)^2$$

- za **razvrščanje** – diskretne izhodne vrednosti
 - natančnost (za vrednotenje)

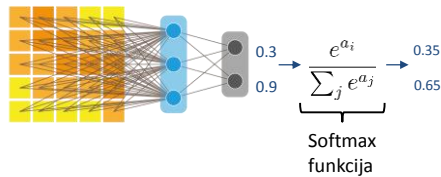
$$L_i = \mathbb{I}(t_i = \operatorname{argmax}_j p_{i,j})$$

- križna-entropija (za učenje)

$$L_i = - \sum_j t_{i,j} \log p_{i,j}$$

Normalizacija izhoda Softmax

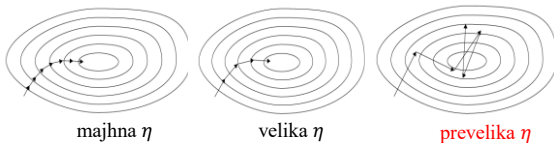
- Pri uporabi **križne entropije** moramo zagotoviti, da bo vsota izhodnih vrednosti enaka 1



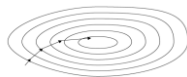
Gradientna metoda ali postopek najstrmejšega spusta

$$w^{(\tau+1)} = w^{(\tau)} - \eta \frac{\partial L_i}{\partial w^{(\tau)}}$$

- Na smiseln način se premikamo po „prostoru uteži“ do optimalnih uteži w
 - uporabimo vse podatke (**batch**) → računsko prezahtevno
 - uporabimo po en vzorec (**sequential**) → slaba konvergenca
 - uporabimo majhno število vzorcev (**minibatch**) → izboljšana konvergenca in računsko obvladljivo
- Vpliv koraka η (učna konstanta – **learning rate**)



Adaptivni učni korak η

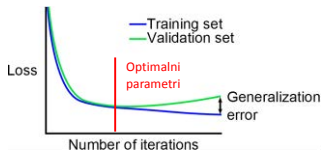


- Intuitivno naj bi η zmanjševali, ko se približujemo optimumu (v začetki daljši koraki, nato krajši)
- Ta potek sovпада z magnitudo gradienta $\frac{\partial L_i}{\partial w}$, kar izkoriščajo nekateri postopki
 - RMSP**rop in **AdaDelta** uporabljata tekoče povprečje magnitude gradienta
 - AdaGrad** podobno, vendar vsiljuje monotono padajoče vrednosti η
 - Omenjeni postopki stabilizirajo magnitudo popravkov uteži
 - Adam** upošteva moment (pospešek), kar stabilizira tako magnitudo kot smer popravkov uteži
 - trenutno najbolj priporočen postopek

Delitev podatkov pri učenju

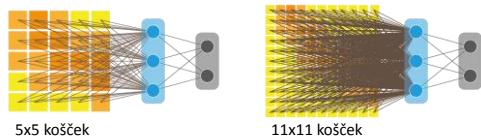
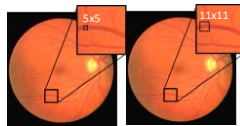


- **Učenje** na delu učne množice (vzorci+referenčne oznake)
- **Validacija** na delu učne množice (za optimizacijo učenja)
- **Testiranje** na testni množici (vzorci+referenčne oznake, nikoli uporabljena za učenje ali validacijo)



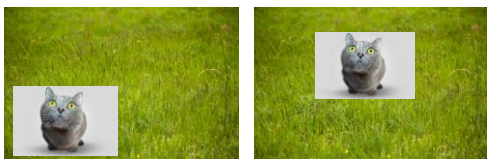
Problemi in omejitve

- Koliko informacije „vidi“ mreža med procesom razpoznavanja?
- Z večanjem števila povezanih vhodnih pikselov hitro narašča število parametrov
 - večja verjetnost nasičenja učenja in posledično
 - manjša zmožnost posploševanja na novih vzorcih



Problemi in omejitve

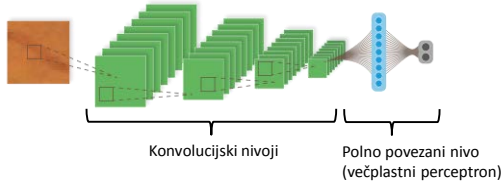
- V mnogih aplikacijah se lahko strukture in objekti pojavijo na različnih lokacijah v sliki
 - težko ustrezno naučiti večplastni perceptron
 - potreba po prostorski invarianci izločanja značilnic



Konvolucijske nevronske mreže

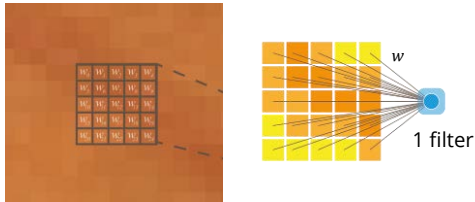
ang. convolutional neural networks - CNN

- **Konvolucijski nivo** izvaja konvolucijo s sliko tako, da uporablja majhna jedra → naučimo se več filtrov
 - vrednosti v jedrih niso predoločene, ampak se jih naučimo
 - zaporedno zlaganje konvolucijskih nivojev omogoča izločanje kompleksnejših značilnic (*globina!*)



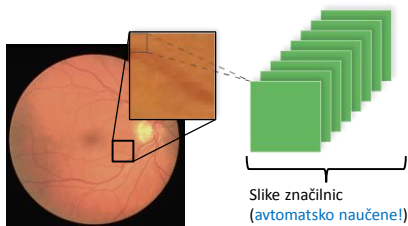
Konvolucijski nivo

- Podoben večplastnemu perceptronu
 - vrednosti filtra so uteži w
 - uteži je manj, ker je sprejemno polje lokalno in
 - uteži so preko uporabe konvolucije translacijsko invariantne (*manj parametrov!*)



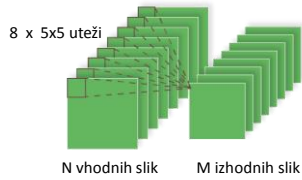
Konvolucijski nivo

- Filter se premika preko koščka slike ali pa celotne slike
 - kjer je veljavna konvolucija (za vse voksle moramo razširiti sliko)
 - eden izmed parametrov je korak premika (ang. stride)

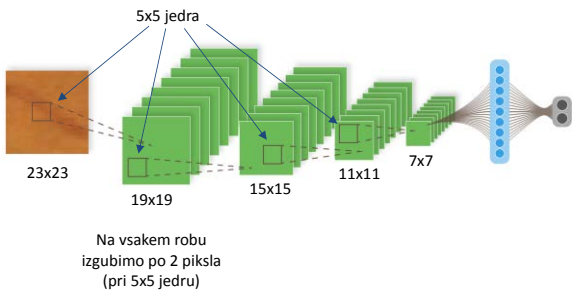


Konvolucijski nivo

- Vhodnih slik je lahko več (npr. RGB kanali barvne slike)
 - in jih preko nivojev kombiniramo v značilnice

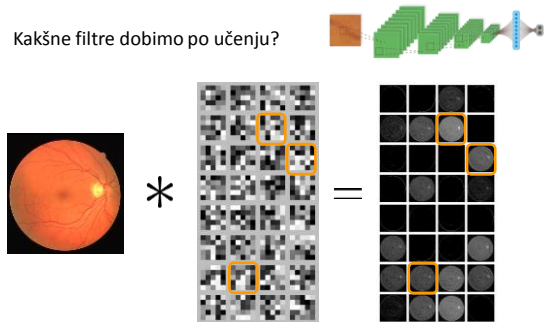


Veljavnost konvolucije in velikosti polj značilnic

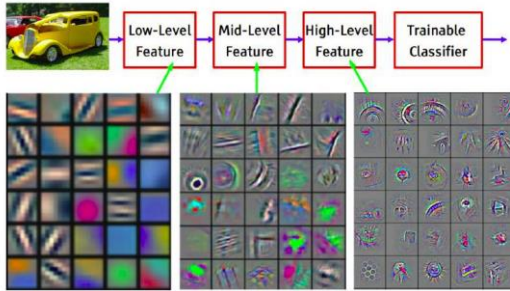


Primeri izločenih značilnic

- Kakšne filtre dobimo po učenju?



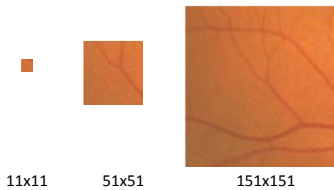
Primeri izločenih značilnic – ImageNet kategorizacija objektov



Sprejemno polje

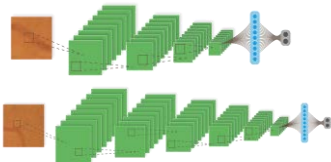
ang. receptive field

- Košček slike, ki ga mreža analizira za namen razvrščanja enega piksla (lahko tudi cela slika)
 - želimo čim večje, zato da uporabimo čim več informacije
 - z večjo globino mreže kodiramo bolj kompleksne značilnice



Problemi in omejitve

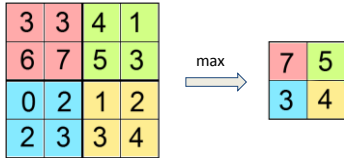
- Veliko sprejemno polje in večja globina mreže pomenita mnogo večje število parametrov za učenje
 - nevarnost nasičenja in manjša zmožnost posploševanja na nove vzorce
 - moderne arhitekture uporabljajo pretežno majhna 3x3 jedra in večje globine



Združevanje značilnic

ang. pooling

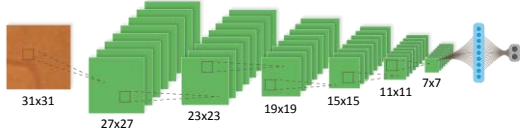
- Učinkovit način manjšanja števila parametrov je z združevanjem odzivov
 - izračunamo neko statistiko na podoknih slike, npr. maksimum, povprečje in v naslednji nivo izvozimo le vrednost statistike



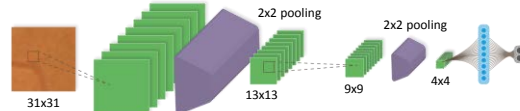
Združevanje značilnic

ang. pooling

- Originalna mreža: 8200 parametrov**



- Mreža z združevanjem: 1800 parametrov**



Izzivi in triki pri učenju

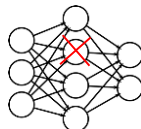
- Običajno malo vzorcev (slik) in preveč parametrov w
- Rešitve:
 - več podatkov (augmentation →)
 - zmanjšanje parametrov (pooling)
 - regularizacija učenja (adaptivna učna konstanta η , ↓batch normalization, dropout ↓)

$$x = \sum w_i \theta_i$$

$$\hat{x} = \frac{\hat{x} - \mu_B}{\sqrt{\sigma_B} + \epsilon}$$

$$y = \gamma \hat{x} + \beta$$

$$a_{out} = f(y)$$



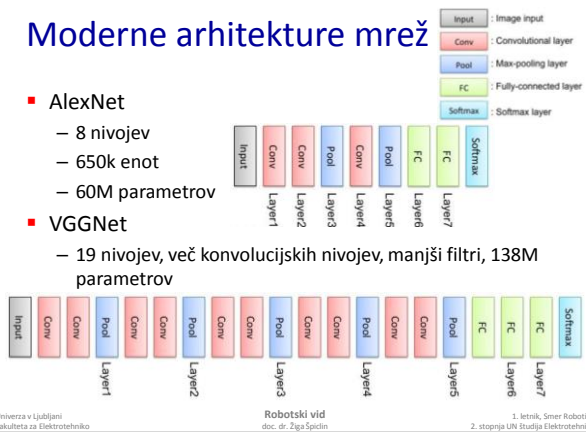
Moderne arhitekture mrež

▪ AlexNet

- 8 nivojev
- 650k enot
- 60M parametrov

▪ VGGNet

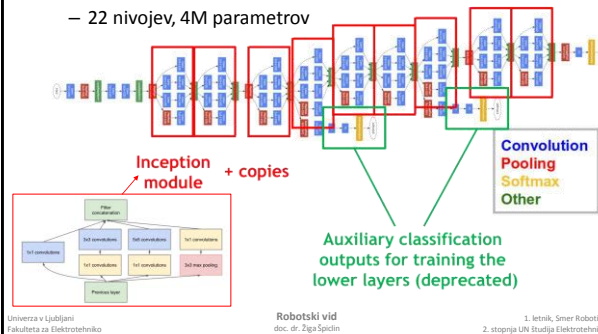
- 19 nivojev, več konvolucijskih nivojev, manjši filtri, 138M parametrov



Moderne arhitekture mrež

▪ GoogLeNet

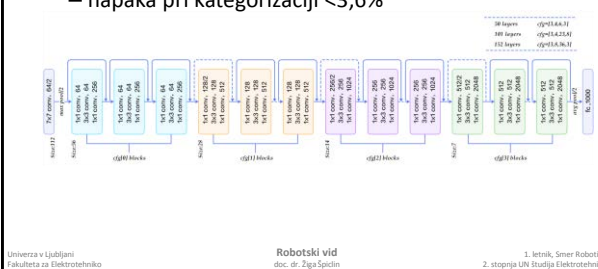
- 22 nivojev, 4M parametrov



Moderne arhitekture mrež

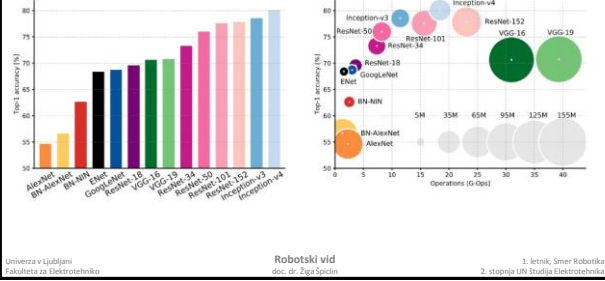
▪ ResNet

- 152 nivojev, „skip“ povezave
- napaka pri kategorizaciji <3,6%



Zmogljivosti arhitektur mrež

- Primerjava iz leta 2017



Primeri uporabe

- Razpoznavanje kategorije objektov



Primeri uporabe

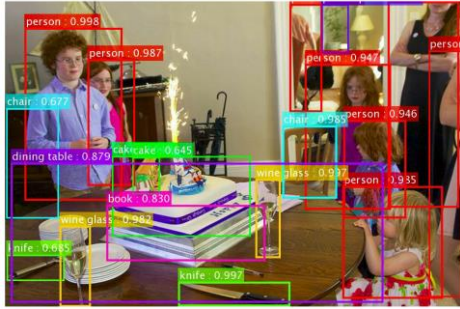
- Iskanje podobnih slik



Testne slike Najbolj podobne slike →

Primeri uporabe

- Lokalizacija in razpoznavanje objektov



Primeri uporabe

- Lokalizacija, razpoznavanje in obrisovanje objektov



Povzetek

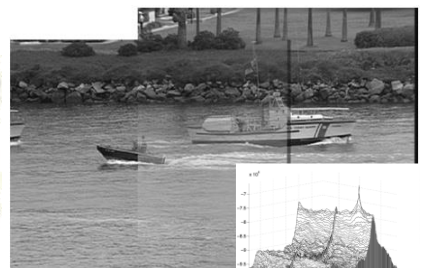
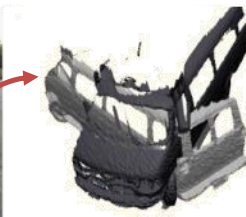
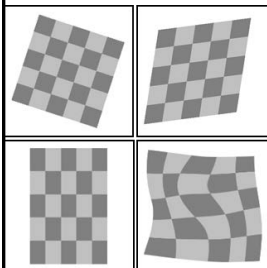
- Oglišča in robovi v sliki so področja visoke spremembe intenzitete in so pomembni za zaznavanje objektov
- Globalno povezovanje robov in zaznavanje objektov na osnovi robov lahko izvedemo s Houghovo preslikavo
 - podpira zaznavo parametričnih oblik (premice, krogi, itd.) in poljubnih neparametričnih oblik (R-tabela)
 - zaznamo več objektov hkrati, odporen na šum in prekrivanje
- Konvolucijske nevronske mreže
 - osnovni gradniki so enote, uteži-povezave, in aktivacijska funkcija, te enote so večnivojske
 - izločanje značilnic je samodejno preko učenja uteži-povezav v prostorsko invariantnih konvolucijskih jedrih
 - učenje je zahtevno zaradi velikega števila prostih parametrov



6. Geometrijska poravnava slik in oblik



Robotski vid




6. Geometrijska poravnava slik in oblik 2

Študijska literatura

Principles of Digital Image Processing: Core Algorithms
 Wilhelm Burger, Mark J. Burge, Springer, 2009
 10. Geometric operations, 4. Comparing images

Biomedicinska slikovna informatika in diagnostika
 Boštjan Likar, Založba FE, 2008
 3.2.4 Geometrijske preslikave,
 4.4 Geometrijska kalibracija in poravnava

Computer vision: algorithms and applications
 Richard Szeliski, Springer, 2010
 2.1 Geometric primitives and transformations
 3.6. Geometric transformations
 6.1 2D and 3D feature based alignment
 6.2 Pose estimation
 8.1 Translational alignment
 8.2 Parametric motion
 8.3 Spline-based motion




Univerza v Ljubljani Robotski vid 1. letnik, smer Robotika
 Fakulteta za Elektrotehniko doc. dr. Žiga Spičin 3. stopnja UN študijski Elektrotehniko

6. Geometrijska poravnava slik in oblik 3

Vsebina

- Definicija in delitev geometrijskih poravnav
 - 2D/3D, eno- ali več-modalna, linearna ali nelinearna
- Elementarne preslikave
 - linearna: translacija, rotacija, skaliranje, strig
 - nelinearne: projektivna, radialne funkcije, B-zlepki
- Mere podobnosti
 - med slikami, med geometrijskimi značilnicami
- Izvedba poravnave
 - eksplicitni izračun preslikave med kontrolnimi točkami in slikami
 - globalna ali lokalna numerična optimizacija




Univerza v Ljubljani Robotski vid 1. letnik, smer Robotika
 Fakulteta za Elektrotehniko doc. dr. Žiga Spičin 3. stopnja UN študijski Elektrotehniko

6. Geometrijska poravnava slik in oblik 4


Geometrijska poravnava

DEFINICIJA


- Iskanje optimalnih geometrijskih preslikav, ki bodo slike oz. oblike poravnale tako, da se bodo iste strukture na vseh slikah oz. oblikah nahajale na enakih položajih

$a(x,y)$

 Referenčni vzorec

T_1


$b_1(u,v)$

 Vzorec 1


T_2

$b_2(u,v)$

 Vzorec 2


\dots

T_N

$b_N(u,v)$

 Vzorec N


 Referenčni vzorec

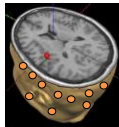
\dots


 Vzorec N

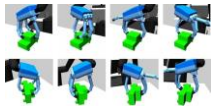
Univerza v Ljubljani Robotski vid 1. letnik, smer Robotika
 Fakulteta za Elektrotehniko doc. dr. Žiga Spičin 3. stopnja UN študijski Elektrotehniko

Uporaba geometrijskih poravnav

- **Kalibracija**
→ poravnava kalibracijskih slik oz. objektov
- **Integracija informacije**
→ številne aplikacije v medicini



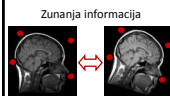
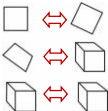
- **Določanje položaja in lege**
→ sledenje in vodenje



Delitev geometrijskih poravnav

GLEDE NA OBLIKO SLIKOVNE INFORMACIJE

- **Razsežnost slikovnega prostora:**
– 2D–2D, 2D–3D, 3D–3D
- **Osnova za poravnavo:**
– zunanja informacija (kontrolne točke)
– notranja informacija (mere podobnosti, kontrolne točke)
– neslikovna informacija (koordinatni sistemi)



Delitev geometrijskih poravnav

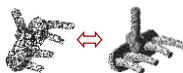
GLEDE NA OBLIKO SLIKOVNE INFORMACIJE

- **Slikovne tehnike oz. vrste slik:**
– enake (**enomodalna poravnava** – *monomodal registration*)
– različne (**večmodalna poravnava** – *multimodal registration*)
– poravnava slike oz. oblike in modela
– poravnava slike in oblike

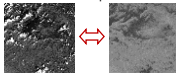
Enomodalna poravnava



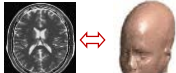
Poravnava oblika – model



Večmodalna poravnava



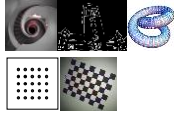
Poravnava slika – oblika



Delitev geometrijskih poravnav

GLEDE NA OBLIKO SLIKOVNE INFORMACIJE

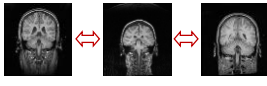
- Značilnice za poravnavo:
 - slikovna ali geometrijska informacija (sivine, barve, oglišča, robovi, obrisi, oblika, topologija...)
 - kalibracijski objekti
- Objekt poravnave:
 - slike oz. oblike istega objekta (*intra-object registration*)
 - slike oz. oblike različnih objektov (*inter-object registration*)



Poravnava slik oz. oblik istega subjekta



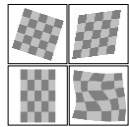
Poravnava slik različnih subjektov



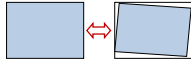
Delitev geometrijskih poravnav

GLEDE NA OBLIKO GEOMETRIJSKE PRESLIKAVE

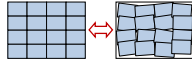
- Vrsta geometrijske preslikave:
 - toge in netoge
 - linearne in nelinearne poravnave
- Domena geometrijske preslikave:
 - globalne in lokalne poravnave



Globalna poravnava



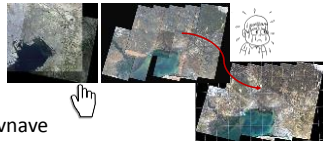
Lokalne poravnave



Delitev geometrijskih poravnav

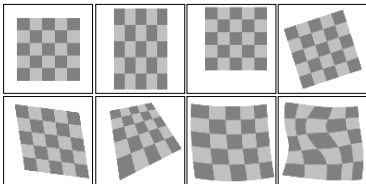
GLEDE NA NAČIN IZVEDBE

- Način delovanja:
 - ročne
 - polsamodejne
 - samodejne poravnave
- Določanje parametrov preslikav:
 - eksplicitni izračun
 - z optimizacijsko metodo



Geometrijske preslikave

- **Preslikava slik**
→ vse slikovne elemente preslikamo na nove lokacije
→ sivinske vrednosti se pri tem ne spremenijo
- **Preslikava oblik**
→ vse koordinate preslikamo na nove lokacije

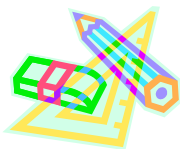


Geometrijske preslikave

- Formalni zapis geometrijske preslikave:

$$(u, v) = T(x, y)$$

$$(u, v, w) = T(x, y, z)$$



- Najbolj splošna linearna preslikava je afina preslikava, ki je sestavljena iz štirih elementarnih preslikav:



Afina preslikava

- Zapišemo kot produkt vektorja koordinat in homogene transformacijske matrike:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{T} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix} = \mathbf{T} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & t_x \\ a_{21} & a_{22} & a_{23} & t_y \\ a_{31} & a_{32} & a_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Afina preslikava

- Matrični zapis afine preslikave omogoča njen zapis z zaporednim množenjem matrik posameznih elementarnih preslikav v poljubnem vrstnem redu, npr.:

$$T_{\text{afna}} = T_{\text{sting}} T_{\text{rot}} T_{\text{trans}} T_{\text{skal}}$$

- Afina preslikava:
 - ohranja vzporednost med premicami
 - ne ohranja kotov med premicami
 - ne ohranja razdalj med poljubnimi točkami

Elementarne lin. preslikave

Skaliranje:

- ohranja vzporednost med premicami
- ohranja kote med premicami
- ohranja razdalje med poljubnimi točkami
- število parametrov v 2D
- število parametrov v 3D

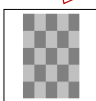
DA
DA
NE
2
3

 T_{2D}

$$\begin{bmatrix} k_x & 0 & 0 \\ 0 & k_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

 T_{3D}

$$\begin{bmatrix} k_x & 0 & 0 & 0 \\ 0 & k_y & 0 & 0 \\ 0 & 0 & k_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



Elementarne lin. preslikave

Translacija:

- ohranja vzporednost med premicami
- ohranja kote med premicami
- ohranja razdalje med poljubnimi točkami
- število parametrov v 2D
- število parametrov v 3D

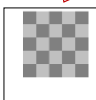
DA
DA
DA
2
3

 T_{2D}

$$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

 T_{3D}

$$\begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



Sestavljene lin. preslikave

Podobnostna = toga + enako skaliranje:

- ohranja vzporednost med premicami
- ohranja kote med premicami
- ohranja razdalje med poljubnimi točkami
- število parametrov v 2D
- število parametrov v 3D

DA
DA
NE
4
7

 T_{2D}

$$\begin{bmatrix} k c_{\alpha} & -k s_{\alpha} & t_x \\ k s_{\alpha} & k c_{\alpha} & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

 T_{3D}

$$\begin{bmatrix} k c_{\beta} c_{\gamma} & k(c_{\alpha} s_{\gamma} + s_{\alpha} s_{\beta} c_{\gamma}) & k(s_{\alpha} s_{\gamma} - c_{\alpha} s_{\beta} c_{\gamma}) & t_x \\ -k c_{\beta} s_{\gamma} & k(c_{\alpha} c_{\gamma} - s_{\alpha} s_{\beta} s_{\gamma}) & k(s_{\alpha} c_{\gamma} + c_{\alpha} s_{\beta} s_{\gamma}) & t_y \\ k s_{\beta} & -k s_{\alpha} c_{\beta} & k c_{\alpha} c_{\beta} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Sestavljene lin. preslikave

Afina preslikava:

- ohranja vzporednost med premicami
- ohranja kote med premicami
- ohranja razdalje med poljubnimi točkami
- število parametrov v 2D
- število parametrov v 3D

DA
NE
NE
6
12

 T_{2D}

$$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

 T_{3D}

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & t_x \\ a_{21} & a_{22} & a_{23} & t_y \\ a_{31} & a_{32} & a_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Nelinearne preslikave

Projektivna preslikava:

- ohranja vzporednost med premicami
- ohranja kote med premicami
- ohranja razdalje med poljubnimi točkami
- število parametrov v 2D
- število parametrov v 3D

NE
NE
NE
8
15

 T_{2D}

$$u = \frac{a_{11}x + a_{12}y + t_x}{p_x x + p_y y + 1}$$

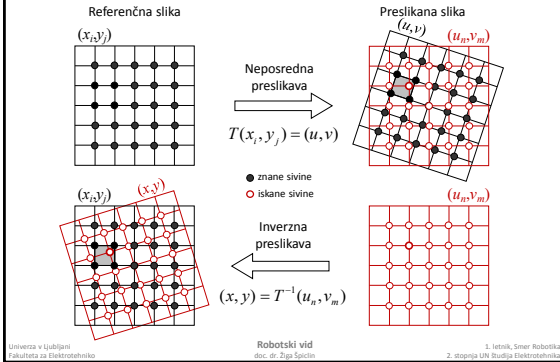
$$v = \frac{a_{21}x + a_{22}y + t_y}{p_x x + p_y y + 1}$$

$$u = \frac{a_{11}x + a_{12}y + a_{13}z + t_x}{p_x x + p_y y + p_z z + 1}$$

$$v = \frac{a_{21}x + a_{22}y + a_{23}z + t_y}{p_x x + p_y y + p_z z + 1}$$

$$w = \frac{a_{31}x + a_{32}y + a_{33}z + t_z}{p_x x + p_y y + p_z z + 1}$$

Interpolacija pri preslikavah



Mere podobnosti

DEFINICIJA IN LASTNOSTI

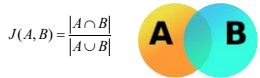
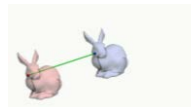
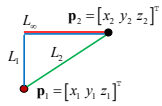
- **MP**: poljubna skalarna funkcija, določena nad vsemi istoležnimi elementi referenčne $a(x,y)$ in lebdeče slike oz. oblike $b(x,y)$
 - **čim bolj** občutljiva na geometrijska neskladja med slikama oz. oblikama
 - **čim manj** občutljiva na motilna neskladja (npr. šum, geometrijska in sivinska neskladja in popačenja)
 - smiselne, a ne nujne lastnosti so **zveznost**, **metričnost** in v nekaterih primerih **neobčutljivost na specifične preslikave** (geometrijske in/ali sivinske)
- Lastnosti metrike oz. mere razdalje **MR**
 - nenegativnost: $MR(a, b) \geq 0$
 - identiteta: $MR(a, b) = 0$ če in samo če $a = b$
 - simetričnost: $MR(a, b) = MR(b, a)$
 - trikotniška neenakost: $MR(a, c) \geq MR(a, b) + MR(b, c)$

MP ima lahko vse ali pa nobene od lastnosti metrike

Mere podobnosti

PODOBOST GEOMETRIJSKIH ZNAČILNIC

- L_p norma oz. razdalja
- Hausdorffova razdalja **HD** (Hausdorff distance)
 - razdalja med dvema množicama značilnic
- Jaccardov indeks J (Jaccard index)
 - podobnost med dvema množicama značilnic

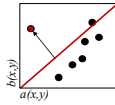


Mere podobnosti

PODOBNOŠT SLIKOVNIH ZNAČILNIC

- Srednja kvadratna napaka *MSE* (mean square error) → občutljiva zaradi kvadrata

$$MSE(a,b) = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J (a(x_i, y_j) - b(x_i, y_j))^2$$

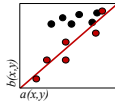


- Srednja absolutna napaka *MAE* (mean absolute error) → manj občutljiva kot MSE

$$MAE(a,b) = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J |a(x_i, y_j) - b(x_i, y_j)|$$

- Korelacijski koeficient *CC* (correlation coefficient) → neobčutljiv na linearne preslikave sivin

$$CC(a,b) = \frac{\sum_{i=1}^I \sum_{j=1}^J (a(x_i, y_j) - \bar{a}) \cdot (b(x_i, y_j) - \bar{b})}{\sqrt{\sum_{i=1}^I \sum_{j=1}^J (a(x_i, y_j) - \bar{a})^2 \cdot \sum_{i=1}^I \sum_{j=1}^J (b(x_i, y_j) - \bar{b})^2}}$$

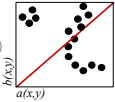


Mere podobnosti

PODOBNOŠT SLIKOVNIH ZNAČILNIC

- Pri večmodalni poravnavi slik, npr. poravnavi CT in MR slik, sivinske vrednosti niso funkcijsko, ampak le statistično odvisne → prej našete mere odpovedo
- Uveljavile so se informacijsko teoretične mere → še najbolj medsebojna informacija *MI* (mutual information)

$$MI(a,b) = H(a) + H(b) - H(a,b)$$



- Marginalni entropiji:

$$H(a) = - \sum_{s_a=0}^{L-1} p_a(s_a) \log p_a(s_a)$$

$$H(b) = - \sum_{s_b=0}^{L-1} p_b(s_b) \log p_b(s_b)$$

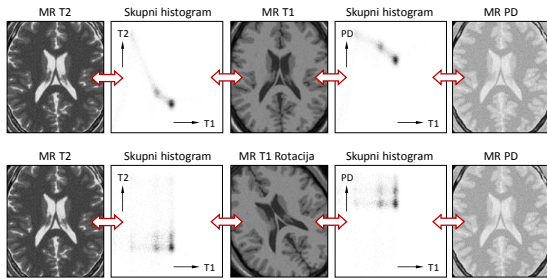
- Skupna entropija:

$$H(a,b) = - \sum_{s_a=0}^{L-1} \sum_{s_b=0}^{L-1} p_{ab}(s_a, s_b) \log p_{ab}(s_a, s_b)$$

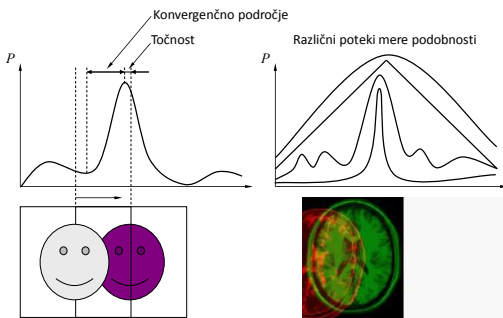
- Ocena skupne verjetnosti iz vezanega histograma sivinskih vrednosti:

$$p_{ab}(s_a, s_b) = \frac{h_{ab}(s_a, s_b)}{IJ}$$

Medsebojna informacija



Analiza mer podobnosti



Izvedba poravnave slik in oblik

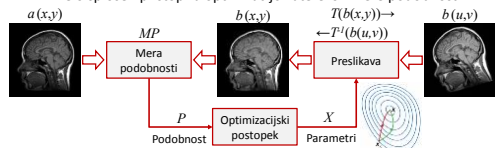
Poravnava geometrijskih značilnic
→ npr. kontrolnih točk
 $a(x,y)$ $b(u,v)$

Poravnava slikovnih značilnic
→ npr. sivinskih vrednosti
 $a(x,y)$ $b(u,v)$


Določanje parametrov preslikav

Z OPTIMIZACIJO MERE PODOBNOSTI

- Eksplisitni izračun**
 - za poravnavo kontrolnih točk z **znanimi pripadajočimi pari**
 - za poravnavo z nekaterimi merami podobnosti (npr. korelacijski koeficient)
- Z optimizacijskim postopkom**
 - za poravnavo kontrolnih točk **brez vnaprejšnjih pripadajočih parov**
 - zelo splošen pristop za optimizacijo katerekoli mere podobnosti



Poravnava kontrolnih točk

- Poznamo množico parov $(x_k, y_k) \leftrightarrow (u_k, v_k)$ **pridajajočih kontrolnih točk** na referenčni a in vhodni sliki oz. obliki b

- Pridajajoči pari točk (*corresponding points*) morajo predstavljati iste strukture na obeh slikah:
 - določimo jih lahko ročno → kliknemo na izrazite oz. značilne strukture, ki jih lahko zanesljivo določimo
 - npr. središča majhnih objektov, razcepišča, velike ukrivljenosti črt in robov
 - obstajajo tudi posebni samodejni in polsamodejni postopki za iskanje pripadajočih parov točk (ICP, RANSAC, SIFT, SURF, MSER,...)

Poravnava kontrolnih točk

- Iščemo preslikavo T , ki bo kontrolne točke čim boljše prekrila:

$$T(x_k, y_k) \leftrightarrow (u_k, v_k)$$

$$(x_k, y_k) \leftrightarrow T^{-1}(u_k, v_k)$$

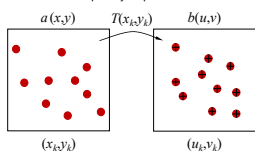
Metrika v prostoru vhodne slike (u, v)

→ lažja interpolacija pri preslikavi vhodne slike $T^{-1}(b(u, v))$

Metrika v prostoru referenčne slike (x, y)

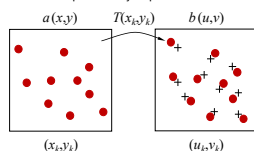
Popolno prekrivanje točk

Interpolacijska poravnava



Približno prekrivanje točk

Aproksimacijska poravnava



Afina interpolacija

• Afina preslikava:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = T \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

6 parametrov → potrebujemo 3 pare kontrolnih točk

• Sistem 6 enačb za 6 neznanih parametrov:

$$\begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix}$$

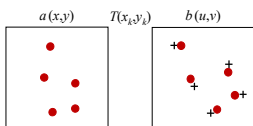
• Rešitev → matrika parametrov preslikave T :

$$T = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ 1 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix}$$

Afina aproksimacija

• Afina preslikava:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = T \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$



• Če imamo več kot 3 pare kontrolnih točk dobimo predoločen sistem enačb:

$$\begin{bmatrix} u_1 & u_2 & u_3 & \dots & u_K \\ v_1 & v_2 & v_3 & \dots & v_K \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_K \\ y_1 & y_2 & y_3 & \dots & y_K \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}$$

• Kontrolne točke z afino preslikavo ne moremo popolnoma poravnati → lahko ji le približno poravnamo

• Iščemo preslikavo T , ki bo **minimizirala povprečno kvadratno razdaljo** R^2 med vsemi pari:
 $R^2 = \frac{1}{K} \sum_{k=1}^K (T(x_k, y_k) - (u_k, v_k))^2 = \frac{1}{K} \sum_{k=1}^K ((a_{11}x_k + a_{12}y_k + t_x - u_k)^2 + (a_{21}x_k + a_{22}y_k + t_y - v_k)^2)$

Afina aproksimacija

$$R^2 = \frac{1}{K} \sum_{k=1}^K (T(x_k, y_k) - (u_k, v_k))^2 = \frac{1}{K} \sum_{k=1}^K ((a_{11}x_k + a_{12}y_k + t_x - u_k)^2 + (a_{21}x_k + a_{22}y_k + t_y - v_k)^2)$$

= prilaganje po postopku najmanjših kvadratov (least-squares fitting)

• Odvajamo R^2 po vseh parametrih in odvode postavimo na nič:

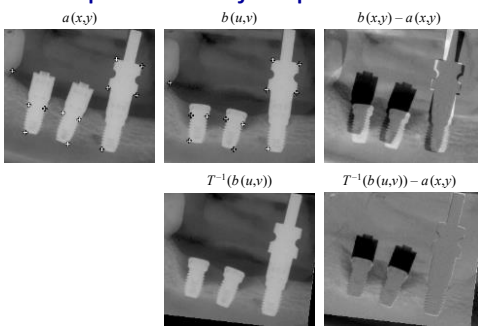
$$\frac{\partial R^2}{\partial a_{11}} = 0 \quad \frac{\partial R^2}{\partial a_{12}} = 0 \quad \frac{\partial R^2}{\partial t_x} = 0 \quad \frac{\partial R^2}{\partial a_{21}} = 0 \quad \frac{\partial R^2}{\partial a_{22}} = 0 \quad \frac{\partial R^2}{\partial t_y} = 0$$

• Dobimo sistem 6 enačb za 6 neznanih parametrov:

$$\begin{bmatrix} \bar{xx} & \bar{xy} & \bar{x} & 0 & 0 & 0 \\ \bar{xy} & \bar{yy} & \bar{y} & 0 & 0 & 0 \\ \bar{x} & \bar{y} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \bar{xx} & \bar{xy} & \bar{x} \\ 0 & 0 & 0 & \bar{xy} & \bar{yy} & \bar{y} \\ 0 & 0 & 0 & \bar{x} & \bar{y} & 1 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ t_x \\ a_{21} \\ a_{22} \\ t_y \end{bmatrix} = \begin{bmatrix} \bar{ux} \\ \bar{uy} \\ \bar{u} \\ \bar{vx} \\ \bar{vy} \\ \bar{v} \end{bmatrix}$$

$\bar{ux} = \frac{1}{K} \sum_{k=1}^K u_k x_k$
 • Rešitev: $\mathbf{P}_{xy} \cdot \mathbf{t} = \mathbf{p}_{uv}$ → parametri preslikave $\mathbf{t} = \mathbf{P}_{xy}^{-1} \cdot \mathbf{p}_{uv}$

Afina aproksimacija – primer



Toga interpolacija?

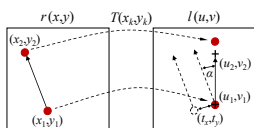


$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & t_x \\ \sin(\alpha) & \cos(\alpha) & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- 3 parametri → **interpolacija ni možna**:
- 1 par kontrolnih točk je premalo → 2 parametra
- 2 para kontrolnih točk je preveč → 4 parametre

- 2 para kontrolnih točk:

- s prvim določimo parametra translacije in točki interpoliramo
- z drugim določimo vektor za kot rotacije in točki aproksimiramo



Toga aproksimacija



- Če imamo več kot 1 par kontrolnih točk dobimo predoločen sistem enačb:

$$\begin{bmatrix} u_1 & u_2 & u_3 & \dots & u_K \\ v_1 & v_2 & v_3 & \dots & v_K \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & t_x \\ \sin(\alpha) & \cos(\alpha) & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_K \\ y_1 & y_2 & y_3 & \dots & y_K \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}$$

- Iščemo preslikavo T , ki bo **minimizirala povprečno kvadratno razdaljo** R^2 med vsemi pari:

$$R^2 = \frac{1}{K} \sum_{k=1}^K \left((x_k \cos(\alpha) - y_k \sin(\alpha) + t_x - u_k)^2 + (x_k \sin(\alpha) + y_k \cos(\alpha) + t_y - v_k)^2 \right)$$

- Odvajamo R^2 po vseh parametrih in odvode postavimo na nič:

$$\frac{\partial R^2}{\partial t_x} = 0 \quad \frac{\partial R^2}{\partial \alpha} = 0 \quad \frac{\partial R^2}{\partial t_y} = 0$$

$$\alpha = -\arctg \frac{\bar{u}y - \bar{v}x - \bar{u} \bar{y} + \bar{v} \bar{x}}{\bar{u}x + \bar{v}y - \bar{u} \bar{x} - \bar{v} \bar{y}}$$

- Rešitev sistema enačb: parametri preslikave →

$$t_x = \bar{u} - \bar{x} \cos \alpha + \bar{y} \sin \alpha$$

$$t_y = \bar{v} - \bar{x} \sin \alpha - \bar{y} \cos \alpha$$

Toga aproksimacija v 3D



- Če imamo več kot 2 para kontrolnih točk dobimo predoločen sistem enačb:

$$\begin{bmatrix} u_1 & u_2 & u_3 & \dots & u_K \\ v_1 & v_2 & v_3 & \dots & v_K \\ w_1 & w_2 & w_3 & \dots & w_K \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_K \\ y_1 & y_2 & y_3 & \dots & y_K \\ z_1 & z_2 & z_3 & \dots & z_K \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}$$

$$\mathbf{R} = \mathbf{R}_y \mathbf{R}_\rho \mathbf{R}_\alpha \quad (3 \times 3)$$

$$\mathbf{t} = [t_x \quad t_y \quad t_z]^T$$

- Iščemo rotacijsko matriko \mathbf{R} , ki bo **minimizirala povprečno kvadratno razdaljo** R^2 med usrediščeni kontrolnimi točkami:

$$\mathbf{d}_k = [u_k \quad v_k \quad w_k]^T \quad \mathbf{m}_k = [x_k \quad y_k \quad z_k]^T$$

$$R^2 = \frac{1}{K} \sum_{k=1}^K \|\hat{\mathbf{d}}_k - \mathbf{R} \hat{\mathbf{m}}_k\|^2 = \frac{1}{K} \sum_{k=1}^K (\hat{\mathbf{d}}_k^T \hat{\mathbf{d}}_k + \hat{\mathbf{m}}_k^T \hat{\mathbf{m}}_k - 2\hat{\mathbf{d}}_k^T \mathbf{R} \hat{\mathbf{m}}_k)$$

- Razdalja R^2 bo minimalna, ko bo člen $2\hat{\mathbf{d}}_k^T \mathbf{R} \hat{\mathbf{m}}_k$ **maksimalen**:

- definiramo korelacijsko matriko

$$\mathbf{M} = \sum_{k=1}^K \hat{\mathbf{m}}_k \hat{\mathbf{d}}_k^T \xrightarrow{\text{SVD}} \mathbf{M} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$$

- Parametri preslikave →

$$\mathbf{R} = \mathbf{V} \mathbf{U}^T$$

$$\mathbf{t} = \hat{\mathbf{d}} - \mathbf{R} \hat{\mathbf{m}}$$

Determinanta matrike $\mathbf{R}=1$,
sicer moramo rešitev zrcaliti!

Poravnava slik z maksimizacijo korelacijskega koeficienta

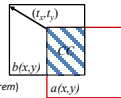


- Predpostavimo translacijsko preslikavo → vrednost CC zavisi od parametrov translacije t_x in t_y

$$CC(t_x, t_y) = \frac{\sum_{i=1}^J \sum_{j=1}^J (a(x_i, y_j) - \bar{a}) \cdot (b(x_i + t_x, y_j + t_y) - \bar{b})}{\sqrt{\sum_{i=1}^J \sum_{j=1}^J (a(x_i, y_j) - \bar{a})^2} \cdot \sqrt{\sum_{i=1}^J \sum_{j=1}^J (b(x_i + t_x, y_j + t_y) - \bar{b})^2}}$$

- Podobnost slik $a(x,y)$ in $b(x,y)$ za vsak par (t_x, t_y) vrednotimo v števcu s križno korelacijo:

$$\rho(t_x, t_y) = \sum_{i=1}^J \sum_{j=1}^J a(x_i, y_j) \cdot b(x_i + t_x, y_j + t_y)$$



- Izrabimo dve lastnosti Fourierjeve transformacije:

→ translacija signala vpliva le na fazo, ne pa magnitudo signala (*shift theorem*)

$$F\{b(x + t_x, y + t_y)\} = F\{b(x, y)\} \cdot e^{-2\pi i(t_x \lambda_x + t_y \lambda_y)}$$

→ konvolucija v prostoru slike se izraža kot množenje v prostoru frekvence (*convolution theorem*)

$$F\{a(x, y) * b(x, y)\} = F\{a(x, y)\} \cdot F\{b(x, y)\}$$

Poravnava slik z maksimizacijo korelacijskega koeficienta



- Križno korelacijo za vse pare (t_x, t_y) lahko naenkrat izračunamo v prostoru frekvence:

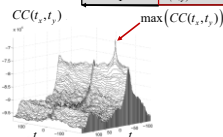
$$\rho(t_x, t_y) = F^{-1}\{F\{a(x, y)\} \cdot \overline{F\{b(x, y)\}}\}$$

Fourierjeva transformacija
predpostavlja periodične signale!

- Možna rešitev je dodajanje ničel
za velikost slike b (*zero padding*):



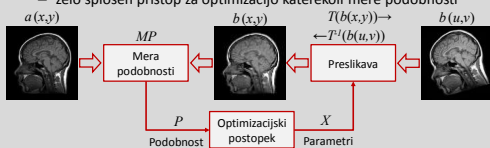
- Primer poravnave slik s korelacijskim koeficientom:



Določanje parametrov preslikav

Z OPTIMIZACIJO MERE PODOBNOSTI

- Eksplisitni izračun
 - za poravnavo kontrolnih točk z **znanimi pripadajočimi pari**
 - za poravnavo z nekaterimi merami podobnosti (npr. korelacijski koeficient)
- Z optimizacijskim postopkom
 - za poravnavo kontrolnih točk **brez vnaprej znanih pripadajočih parov**
 - zelo splošen pristop za optimizacijo katerekoli mere podobnosti

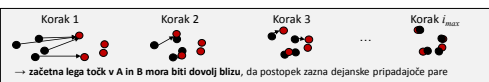


Iskanje dejanskih pripadajočih parov točk v množicah A in B

- Postopek **iterativno najbližje točke** (*iterative closest point - ICP*)
 1. Poišči vsaki točki iz množice A **najbližjo točko** iz B (pripadajoči par)
 2. Oцени parametre preslikave T_i s **trenutnimi pripadajočimi pari točk** z uporabo interpolacijske/aproksimacijske poravnave
 3. Preslikaj točke množice B na nove položaje $T_i(x_k, y_k) \leftrightarrow (u_k, v_k)$
 4. Ponavljaj korake 1-3 dokler je preslikava še smiselno velika $\|T_i\|_\infty < \varepsilon$ oziroma, ko ni doseženo maksimalno število korakov i_{max}
 - končno preslikavo T dobimo z združevanjem delnih preslikav T_i :

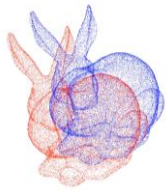
$$T = T_i(T_{i-1}(\dots(T_1))) \quad \text{oz. za linearne preslikave} \quad T = T_i T_{i-1} \dots T_1$$

Poravnava s postopkom iterativno najbližje točke



$$T = T_i T_{i-1} \dots T_1$$

- **Primer:**
 - poravnava dveh oblakov točk



Povzetek

- Pri poravnavi slik želimo najti **optimalno geometrijsko preslikavo** tako, da se po poravnavi dveh slik ali oblik iste strukture nahajajo na isti lokaciji
- Osnovni gradniki poravnave slik so **oblika geometrijske preslikave, mera podobnosti in način določanja preslikave**
 - pri poravnavi slik na osnovi sivinskih vrednosti moramo določiti še način interpolacije
- Mera podobnosti izberemo glede na obliko informacije, ki jo želimo primerjati
 - geometrijska vs. slikovna (npr. oslonilne točke ali sivinske vrednosti)
- Preslikavo določimo eksplicitno (analitično) ali preko numerične optimizacije mere podobnosti
 - eksplicitni izračun je običajno možen le v primeru linearnih geometrijskih preslikav



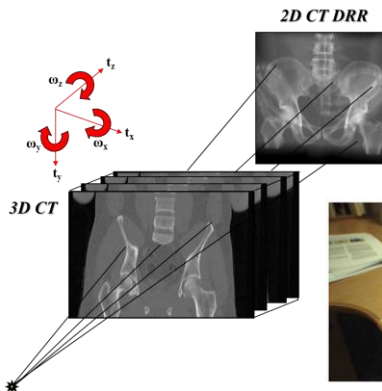
7. Prileganje 3D modelov na 2D slike



Robotski vid



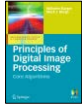
Univerza v Ljubljani
Fakulteta za Elektrotehniko



Robotski vid
doc. dr. Žiga Špiclin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Študijska literatura



Principles of Digital Image Processing: Core Algorithms
Wilhelm Burger, Mark J. Burge, Springer, 2009
10. Geometric operations, 4. Comparing images

Applied medical image processing: A basic course
Wolfgang Birkfellner, CRC Press, 2010
7. Spatial transforms
8. Rendering and surface models



Computer vision: algorithms and applications
Richard Szeliski, Springer, 2010
2.1 Geometric primitives and transformations
3.6. Geometric transformations
6.1 2D and 3D feature based alignment
6.2 Pose estimation

Vsebina predavanja

- Preslikava v slikovno ravnino in definicija problema
- Kalibracija kamere: intrinzični parametri
- Poravnava korespondenčnih točk
 - Linearna / nelinearna optimizacija
- Strategije določanja začetne poze
 - Zmanjšanje vpliva osamelcev: RANSAC
 - Uporaba slikovnih predlog kot začetnih poz
- Primeri in aplikacije
 - v navidezni resničnostni, industriji, robotiki
 - slikovno vodeni posegi v medicini

Preslikava v slikovno ravnino

Notranji (intrinzični) koordinatni sistem

Zunanji (ekstrinzični) koordinatni sistem

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \approx f_c \begin{bmatrix} \lambda f & sf & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Radialne distorzije ($n \geq 3$): x_c, y_c, k_1, \dots

Notranji (intrinzični) parametri ($n=5$): K

Zunanji (ekstrinzični) parametri ($n=6$): $[R | T]$

Projekcijska matrika ($n=11$)
→ rečemo ji homografija

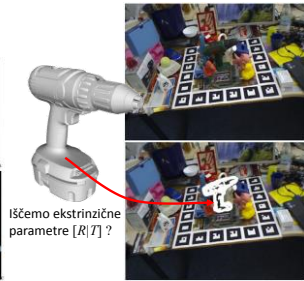
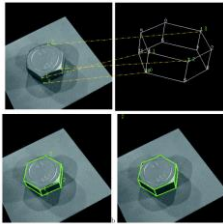
$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & 1 \end{bmatrix}$$

Robotski vid
doc. dr. Žiga Srebrnič

1. letnik, smer Robotika
3. stopnja UN Študij Elektrotehniko

Prileganje 3D modelov na 2D slike

- **Krajše 3D-2D poravnava** (ang. pose estimation, 3D-2D registration)



Iščemo ekstrinzične parametre $[R|T]$?

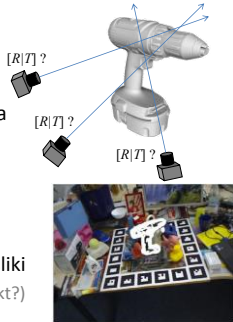
Iskanje 3D poze: definicija problema

- **Izhodišča:**

- Poznamo model objekta
- Dana je slika tega objekta
- *Poznamo intrinzične parametre kamere*

- **Cilj:**

- Poišči položaj kamere, ki odraža pozo objekta na sliki
(Kje je kamera glede na objekt?)



Osnovni gradniki 3D-2D poravnave

- Geometrijska preslikava

1. preslikava poze objekta zanimanja ali kamere
2. **preslikava za ujemanje dimenzij (projekcija)**

- Mera podobnosti

- interpolacija

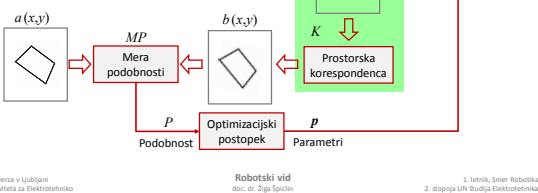
- Način določanja poravnave

- optimizacijski postopek

Izvedba 3D-2D poravnave

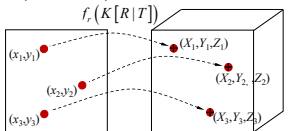
- Informacijo za poravnavo preslikamo v isti prostor

- preslikava med koordinatnimi sistemi slik $(X, Y, Z \leftrightarrow x, y)$
- dimenzijska korespondenca (npr. projekcija iz 3D v 2D)



Strategija iskanja 3D poze

- Poravnavamo pare korespondenčnih točk $x_i \leftrightarrow X_j$



- Poravnava običajno temelji na minimizaciji Evklidske razdalje med korespondenčnimi pari točk

$$E^2 = \min_{(R, \kappa)} \sum_j \| \underbrace{x_j}_{\text{dejanske točke v sliki}} - \underbrace{f_r(K[R, T]X_j; \kappa)}_{\text{projekcija točk na modelu s trenutno preslikavo}} \|^2 \rightarrow \hat{x}_j$$

Strategija iskanja 3D poze

- če ne poznamo korespondenčnih parov, jih poiščemo s posebnimi postopki
 - ICP (ang. iterative closest point), RANSAC (ang. random sample consensus), ipd.
- največja natančnost z iterativno optimizacijo nelinearne preslikave $f_r(K[R, T]; \kappa)$
 - za toge objekte so nelinearnosti v optičnem sistemu, karakteriziramo jih le enkrat s kalibracijo kamere
 - za konvergenco rabimo dober začetni približek parametrov
- približek običajno dobimo z linearno preslikavo $[R, T]$
 - manj natančno, a hitreje
 - več postopkov (PnP, DLT (=P6P), POSIT)
 - izkoriščajo linearno algebro in razcepe SVD, RQ, QR, Cholesky

Osnovni koncepti optimizacije

Linearna

- preslikava $[R, T]$ je linearna, zato velja

$$\Delta x = x - \hat{x} = J(x) p$$

- $J(x)$ predstavlja Jacobianovo matriko prvih odvodov glede na parametre preslikave

- zapišemo nastavke in izpeljemo

$$E^2 = \sum_i \|J(x_i) p - \Delta x_i\|^2 = p^T A p - 2p^T b + c$$

$$A = \sum_i J^T(x_i) J(x_i) \quad \text{Odvajamo } E^2 \text{ na } p \text{ in izenačimo z 0! Dobimo linearen sistem: } A p = b$$

$$b = \sum_i J^T(x_i) \Delta x_i$$

$$c = \sum_i \|\Delta x_i\|^2$$

Nelinearna

- preslikava $f_c(K[R, T]; x)$ je nelinearna, razvijemo jo v Taylorjevo vrsto glede na trenutne vrednosti parametrov p ,

$$p_{i+1} = p_i + \Delta p \quad \left. \begin{array}{l} \text{ITERATIVNA} \\ \text{METODA} \end{array} \right\}$$

- $J(x; p_i)$ predstavlja Jacobianovo matriko prvih odvodov glede na parametre preslikave p_i ,
- izpeljava podobna kot pri linearni, dobimo sistem

$$A \Delta p = b$$

- običajno rešujemo z Levenberg-Marquard

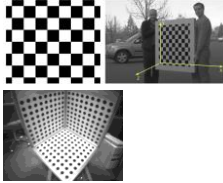
$$[A + \lambda \text{diag}(A)] \Delta p = b, \quad \lambda > 0$$

Kalibracija kamere

- Določanje intrinzičnih parametrov $(f, \lambda, s, x_0, y_0)$ in (radialnih) popačenj leč (x_c, y_c, K_1, \dots)

- Uporabimo množico slik „znanih“ objektov

- 2D (ravninski) objekti
- 3D objekti (ang. vanishing points, N-planes)
- samodejna kalibracija, npr. iz rotacijskega gibanja



Kalibracija kamere z 2D objekti

- Uporabimo visoko kontrastne vzorce
- znane oblike in znanih dimenzij



- Poiščemo korespondenčne točke na sliki kalibra $[U, V, 1]^T$ in modelu $[X, Y, Z]^T$



Kalibracija kamere z 2D objekti

- Točke $[X, Y, Z]^T$ ležijo v ravnini
 - zato je $Z = 0$



- zapišemo preslikavo **BREZ** (radialnih) popačenj:

Homografijo določimo iz vsaj 4ih pripadajočih parov točk $\Rightarrow P = K \begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{bmatrix}$ Matriko K poenostavimo pri množenju z $[X, Y, 0, 1]^T$

$$[h_1 \ h_2 \ h_3] = K [r_1 \ r_2 \ T]$$

Kalibracija kamere z 2D objekti

- Rotacijski del $[r_1, r_2, r_3]$ tvori ortonormalno bazo zato velja
 - $r_1^T r_2 = 0$ in $\|r_1\| = \|r_2\| = 1$
- Iz enačbe $[h_1 \ h_2 \ h_3] = K [r_1 \ r_2 \ t]$ veljata enačbi
 - $r_1 = K^{-1} h_1$ in $r_2 = K^{-1} h_2$
- Če upoštevamo $r_1^T r_2 = 0$ dobimo $h_1^T K^{-T} K^{-1} h_2 = 0$
- Če upoštevamo $\|r_1\| = \|r_2\| = 1$ dobimo $h_1^T K^{-T} K^{-1} h_1 - h_2^T K^{-T} K^{-1} h_2 = 0$
- Definiramo $B = K^{-T} K^{-1}$, ki je simetrična matrika

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} \Rightarrow \text{Unikatne parametre organiziramo v vektor } b = [b_{11} \ b_{12} \ b_{13} \ b_{22} \ b_{23} \ b_{33}]^T$$

Kalibracija kamere z 2D objekti

- Zapišemo sistem enačb $h_1^T B h_1 = 0$ in $h_1^T B h_1 - h_2^T B h_2 = 0$ tako, da uporabimo parametre $b = [b_{11} \ b_{12} \ b_{13} \ b_{22} \ b_{23} \ b_{33}]$
 - za vsak pogled kalibra dobimo dve enačbi
 - potrebujemo najmanj 3 poglede, ker imamo 6 neznank
 - potrebujemo vsaj 4 točke v vsaki ravnini, ker imamo 8 parametrov projektivne preslikave
- Sistem enačb uredimo v obliko $Vb = 0$
 - $V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}$, kjer je $v_{ij} = [h_{i1} h_{i2} h_{i3} h_{i1} h_{i2} + h_{i2} h_{i1}, \dots]$
- Poiščemo rešitev sistema $b^* = \arg \min Vb$
 - iz elementov b^* določimo elemente matrike K

Za N slik

$$\begin{bmatrix} v_{11}^T & v_{12}^T \\ v_{21}^T & v_{22}^T \\ \vdots & \vdots \\ v_{11}^T & v_{12}^T \\ v_{21}^T & v_{22}^T \end{bmatrix} b = 0$$

Slika 1
Slika N

Imamo intrinzične parametre kamerel $r_1 = K^{-1} h_1$
 \rightarrow pa tudi pozno kalibracijskega objekta v vsaki sliki $r_2 = K^{-1} h_2$
 $t = K^{-1} h_3$

Kalibracija kamere z 2D objekti

- Prikazana kalibracija ne upošteva (radialnih) popačenj
 - izračunane vrednosti K, R, t_c uporabimo kot začetni približek
- Optimiziramo nelinearno funkcijo

$$E^2 = \min_{K, R, t_c} \sum_y \left\| \mathbf{x}_y - f_p \left(K [R, T_c] \mathbf{X}_y; \kappa \right) \right\|^2$$

dejanske točke v sliki projekcija točk z modelom

- običajno Levenberg-Marquard, lahko tudi s simpleksno optimizacijo

Primer: kalibrirane slike

- Pred kalibracijo:

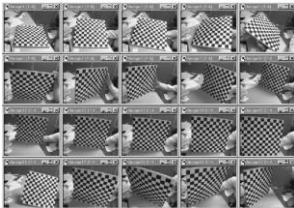


- Po kalibraciji:

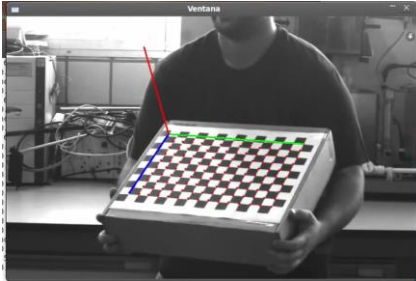


Praktična kalibracija z 2D objekti

- Uporabi npr. vzorec šahovnice
 - enostaven za razpoznavo oglišč (Harris corner detektor)
- Uporabi čim več slik šahovnice v različnih pozah



Primer: kalibrirane slike in poza



<https://www.youtube.com/watch?v=I6HOXIP3I4>

Direktna linearna transformacija (DLT)

- Ocenimo parametre linearne preslikave (brez popačenja)
 - linearna s stališča posplošenih parametrov h

$$\begin{bmatrix} x' \\ y' \\ z \end{bmatrix} = \begin{bmatrix} \lambda f & s & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \Rightarrow \begin{matrix} x = x'/z \\ y = y'/z \end{matrix}$$

$$\mathbf{x} = \mathbf{K} [\mathbf{R} | \mathbf{T}] \mathbf{X}$$

- Predpostavimo, da poznamo intrinzične parametre \mathbf{K}
 - preslikamo točke iz slikovne ravnine v normalizirane koordinate $\tilde{\mathbf{x}}$
 - rešujemo enačbo $\mathbf{K}^{-1}\mathbf{x} = \tilde{\mathbf{x}} = [\mathbf{R} | \mathbf{T}] \mathbf{X}$

Direktna linearna transformacija (DLT)

- Normalizirane $\tilde{\mathbf{x}}$ koordinate zapišemo kot

$$\tilde{\mathbf{x}} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \Rightarrow \begin{matrix} x = \frac{r_{11}X + r_{12}Y + r_{13}Z + t_x}{r_{31}X + r_{32}Y + r_{33}Z + t_z} \\ y = \frac{r_{21}X + r_{22}Y + r_{23}Z + t_y}{r_{31}X + r_{32}Y + r_{33}Z + t_z} \end{matrix}$$

- Pomnožimo z imenovalcem

$$\begin{aligned} r_{11}X + r_{12}Y + r_{13}Z + t_x - x(r_{31}X + r_{32}Y + r_{33}Z + t_z) &= 0 \\ r_{21}X + r_{22}Y + r_{23}Z + t_y - y(r_{31}X + r_{32}Y + r_{33}Z + t_z) &= 0 \end{aligned}$$

- Zapišemo v obliki linearnega sistema $\mathbf{A}\mathbf{p} = 0$

$$\text{dve vrstici za vsak par točk } \tilde{\mathbf{x}} \text{ in } \mathbf{X} \Rightarrow \begin{bmatrix} X & Y & Z & 0 & 0 & 0 & -xX & -xY & -xZ & 1 & 0 & -x \\ 0 & 0 & 0 & X & Y & Z & -yX & -yY & -yZ & 0 & 1 & -y \end{bmatrix} \begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ r_{21} \\ r_{22} \\ r_{23} \\ t_x \\ t_y \\ t_z \end{bmatrix} = 0$$

\mathbf{A} \mathbf{p}

Direktna linearna transformacija (DLT)

- Za vsak par točk $\tilde{\mathbf{x}}$ in \mathbf{X} dobimo dve enačbi
 - za 12 neznanek torej potrebujemo vsaj 6 parov točk
- Rešitev sistema $A\mathbf{p} = 0$ dobimo s SVD razcepom
 - rešitev \mathbf{p} je lastni vektor, ki ustreza **najmanjši lastni vrednosti** matrice $A^T A$
 - želimo minimizirati

$$\|A\mathbf{p}\|^2 = (A\mathbf{p})^T A\mathbf{p} = \mathbf{p}^T A^T A\mathbf{p} \quad \text{s pogojem} \quad \mathbf{p}^T \mathbf{p} = 1$$
 - z Lagrangeovim multiplikatorjem λ dobimo izraz za minimizacijo

$$L(\mathbf{p}) = \mathbf{p}^T A^T A\mathbf{p} - \lambda(\mathbf{p}^T \mathbf{p} - 1)$$
 - odvajamo glede na \mathbf{p} in izenačimo z 0

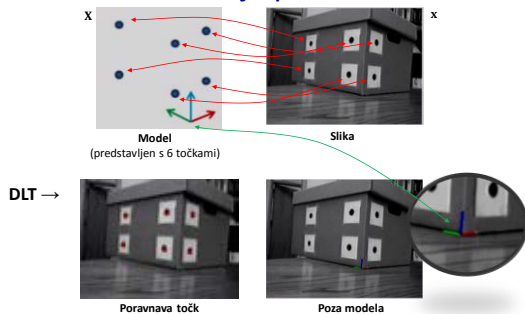
$$A^T A\mathbf{p} - \lambda\mathbf{p} = 0$$
 - iz česar sledi, da je λ lastna vrednost $A^T A$ s pripadajočo lastno vrednostjo $\mathbf{p} = \mathbf{e}_6$
 - funkcija $L(\mathbf{e}_6) = \lambda$, pa je minimalna pri $\lambda = 0$

Direktna linearna transformacija (DLT)

- Preverimo rotacijsko matriko R
 - imeti mora determinanto enako +1
 - če je negativna pomnožimo R z -1
 - če je absolutna vrednost različna od 1, skaliramo matriko tako, da signalarne vrednosti SVD razcepa matrice R postavimo na 1
- **Lastnosti DLT**
 - v primeru, da *ni kalibracije* uporabimo QR razcep, da dobimo matriki K in $[R|T]$

$$M = \begin{bmatrix} Q & R \end{bmatrix}^{-1} = \begin{bmatrix} R^{-1} & Q^{-1} \end{bmatrix}$$
 - hiter izračun parametrov, dober začetni približek
 - nato lahko uporabimo še nelinearno optimizacijo

Primer: Določanje poze 3D modela



Primer: Določanje poze 3D modela

Model
(predstavljen s 6 točkami)

Slika

Nelinearna optimizacija →

Poza modela
i=1 i=2 i=3 i=4 i=5

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

3. letnik, smer Robotika
3. stopnja UN Ljubljane, Elektrotehniko

Določanje poze 3D objekta na podlagi slikovne informacije

- Katero informacijo o modelu lahko uporabimo?
 - Običajno oglišča in robovi, to informacijo izluščimo tudi iz slik (Harris, Canny)

Slika **Canny rob** **Poravnane točke 3D modela**

Običajno veliko z objektom neskladnih ali manjkajočih robov (šum, prekrivanje!)

Težko je zanesljivo najti začetno pozol!

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

3. letnik, smer Robotika
3. stopnja UN Ljubljane, Elektrotehniko

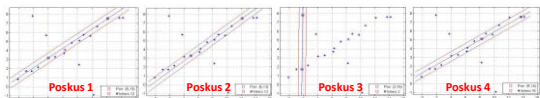
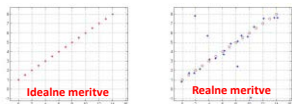
Strategije določanja začetne poze

- Naključno poizkušanje: **RANSAC**
- Uporaba slikovnih predlog
 - grobo vzorčene začetne poze
 - izkoristimo značilnice (npr. sivine, barve, teksture) za poravnavo / zaznavo objekta
 - najdemo najbolj podobno predlogo v bazi
- Uporaba informacije o obliki / oddaljenosti
 - uporaba rekonstrukcije oblike oz. slike oddaljenosti objekta (stereo vid, fotometrični stereo, aktivna osvetlitev, laserski skener,...)



RANSAC (ang. RANdom SAmple Consensus)

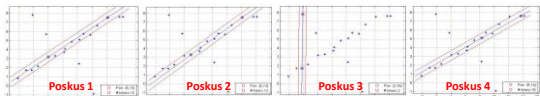
- **Ideja postopka:** Če je število osamelcev (ang. outliers) v množici meritev majhno, izberi podmnožico meritev. Obstaja visoka verjetnost, da bo podmnožica brez osamelcev.
- Primer s prilaganjem premice: vzročimo po dve točki



RANSAC (ang. RANdom SAmple Consensus)

Algoritem:

1. **Naključno izberi majhno podmnožico meritev** in prilagaj model z minimizacijo napake
2. Oцени napako modela glede na vse ostale meritve – **konzenzna množica** je tista množica meritev, kjer je napaka modela manjša od vnaprej določenega praga
3. Večkrat ponovi postopek in na koncu **izberi tisti model, ki ima največjo konzenzno množico**



Lastnosti postopka RANSAC

- Število ponovitev S postopka mora biti zadostno, da bo verjetnost naključne izbire podmnožice brez osamelcev visoka
- Število ponovitev S lahko določimo glede na
 - predvideni delež osamelcev $p = \text{št. „dobrih“ točk} / \text{št. vseh točk}$
 - verjetnost naključne izbire podmnožice brez osamelcev je p^k , kjer je k število točk v podmnožici

$$S = \frac{\log(1-P)}{\log(1-p^k)}$$

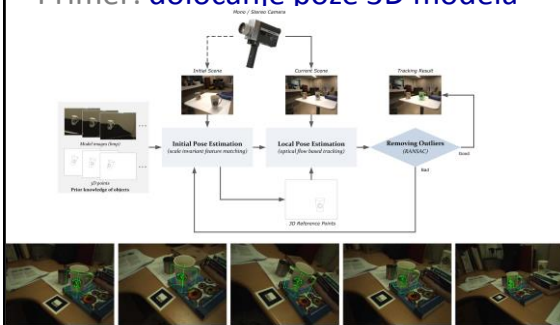
- Potrebno število ponovitev S za $P=99\%$ uspešnost postopka

k	p	S
3	0,5	35
6	0,6	97
6	0,5	293

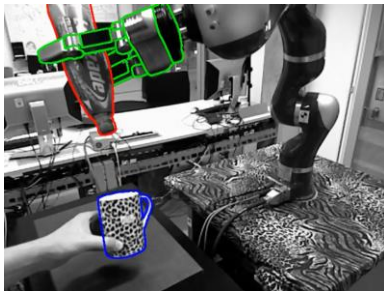
RANSAC in iskanje poze 3D objekta

- **Naključno določimo po 6 parov točk**
 - 6 točk na modelu, 6 točk iz slike robov
 - naključno določimo točke in korespondence med pari
- **Določimo pozo z DLT**
- **Preverimo velikost konsenzne množice**
 - Preslikamo vse točke modela na sliko
 - Določimo korespondence kot najbližje točke (ICP)
 - Preštejemo točke bližje kot npr. 3 piksele
- **Ponavljamo postopek**, na koncu izberemo model z največjo konsenzno množico
- **Po potrebi izboljšamo rešitev** z linearno / nelinearno optimizacijo v obliki postopka ICP
 - hkrati iščemo tudi korespondence

Primer: določanje poze 3D modela



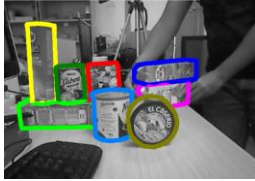
Primer: Sledenje poze 3D modela



<https://www.youtube.com/watch?v=5jhX8fXKPE>

Primer: Sledenje poze množici 3D modelov in različnih objektov na sliki

object detection / 3D pose estimation

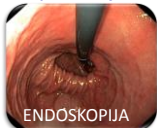


arbitrary view rendered with estimated 3D pose



<https://www.youtube.com/watch?v=H4JSU0M7Yc>

Poleg industrijskih aplikacij tudi mnogo aplikacij v medicini



ENDOSKOPIJA



RADIOTERAPIJA /
RADIOKIRURGIJA



INTERVENTNA
RADIOLOGIJA

- Nprestani vložki v razvoj
 - čim manj invazivnih,
 - čim bolj natančnih in
 - čim bolj uspešnih posegov
- Ena izmed novejših tehnologij, ki se pospešeno uvaja v medicinske posege je slikovno vodenje



KIRURGIJA

Roboti v medicini

- Brahiterapija
- Biopsija



Gabor Fichtinger



© DLR
www.dlr.de

7. Prileganje 3D modelov na 2D slike 38

Slikovno vodenje posegov v medicini

Načrt

„Živa“ slikovna informacija

Tarča

Robotski vid

Univerza v Ljubljani, Fakulteta za Elektrotehniko, 1. letnik, Smer Robotika, 3. stopnja UN študijski Elektrotehniko

7. Prileganje 3D modelov na 2D slike 39

Zakaj slikovno vodenje?

- Primer: stabilizacija hrbtenice z vstavljanjem pedikularnih vijakov

VEČJA NATANČNOST

MINIMALNA INVAZIVNOST

VEČJA USPEŠNOST

Robotski vid

Univerza v Ljubljani, Fakulteta za Elektrotehniko, 1. letnik, Smer Robotika, 3. stopnja UN študijski Elektrotehniko

7. Prileganje 3D modelov na 2D slike 40

Gradniki sistema za slikovno vodenje

- PRED-OPERATIVNE 3D SLIKE IN NAČRT POSEGA**
- SLEDENJE LEGE KIRURŠKIH ORODJI**
- PROSTORSKA PORAVNAVA MED BOLNIKOM IN 3D SLIKO**
- SUPERPOZICIJA SLIK, MODELOV, NAČRTOV IN MED-OPERATIVNIH 2D SLIK NA PRED-OPERATIVNE 3D SLIKE**
- ROBOTIKA**

Robotski vid

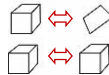
Univerza v Ljubljani, Fakulteta za Elektrotehniko, 1. letnik, Smer Robotika, 3. stopnja UN študijski Elektrotehniko

Prostorska poravnava med bolnikom in sliko

- **Cilj:** podatke zajete pred posegom (slike, modeli in načrti) združi s podatki zajetimi med posegom (bolnik, slike, orodja)



- Glede na dimenzionalnost teh podatkov je poravnava lahko
 - 3D-2D
 - 3D-3D



3D-3D: Določanje preslikave $T=(t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)$

- Uporabimo togo aproksimacijsko poravnavo

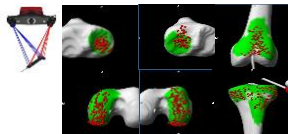
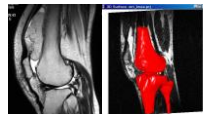
$$\begin{bmatrix} u_1 & u_2 & u_K \\ v_1 & v_2 & v_K \\ w_1 & w_2 & w_K \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_K \\ y_1 & y_2 & y_K \\ z_1 & z_2 & z_K \\ 1 & 1 & 1 \end{bmatrix} \quad \mathbf{R} = \mathbf{R}_x \mathbf{R}_y \mathbf{R}_z \quad (3 \times 3)$$

$$\mathbf{t} = [t_x \quad t_y \quad t_z]^T$$

- rešitev z minimizacijo srednje kvadratne napake med korespondenčnimi pari točk
- V kolikor ne poznamo korespondenčnih parov točk
 - uporabimo postopek iterativno najbližje točke (ICP)

Določanje preslikave $T=(t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)$

- **Slika:** površine dobimo z razmejčitvijo predoperativne 3D slike
- **Bolnik:** površine poravnamo z oblakom točk, ki jih dobimo z dotikanjem korespondenčne površine med posegom

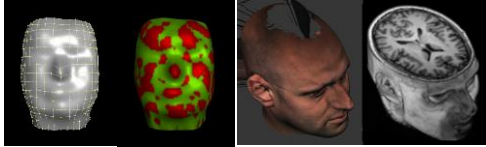


Določanje preslikave $T=(t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)$

- Korespondenčne površine na bolniku dobimo z uporabo laserja, strukturirane svetlobe ali stereo kamer



<http://groups.csail.mit.edu/>



Določanje preslikave $T=(t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)$

- V praksi najpogosteje pred posegom na bolnika pritrdimo markerje ali pa uporabimo anatomske oslonilne očke



vgrajeni v kost pritrjeni na kožo pritrjeni na okvir

- Bolnika slikamo z markerji



anatomske oslonilne točke

Določanje preslikave $T=(t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)$

- Med posegom se z instrumenti, ki jih sledimo v prostoru, dotikamo markerjev na bolniku in korespondenčnih markerjev na predoperativni sliki bolnika

- Tehnike sledenja markerjem:

<p>Mehansko STEREOSTAKTIČNI OKVIR → ROBOT</p>	<p>Optično → sledenje markerjem</p> <p>pasivno aktivno</p>	<p>$\{p_i\}$ $\{q_i\}$</p> <p>$i = 1, 2, \dots, N$</p>
<p>Elektromagnetno → oddajnik induktivni tok v trah med. pravok. navtičnih senzorja</p> <p>oddajnik senzor</p>	<p>Zaznavanje globine → RAZČUNALNIŠKI VID → LASER →</p>	

7. Prilaganje 3D modelov na 2D slike 47

3D-2D: Namesto markerjev uporabimo kar anatomsko informacijo v slikah

$T = (t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)$

S_W , $S_{MED,2}$, $S_{MED,1}$, S_{PRED} , t_z , ω_z , t_x , ω_x , t_y , ω_y

+ možnost avtomatske izvedbe
+ poravnava strukture zanimanja
- dimenzijska neskladja

Univerza v Ljubljani, Fakulteta za Elektrotehniko, Robotični vid, doc. dr. Žiga Savič, 1. letnik, Smerni Robotika, 3. stopnja UN študijski Elektrotehniko

7. Prilaganje 3D modelov na 2D slike 48

Poravnava slika - bolnik

T , S_W , $S_{MED,2}$, $S_{MED,1}$, S_{PRED} , t_z , ω_z , t_x , ω_x , t_y , ω_y

3D/2D poravnava

Univerza v Ljubljani, Fakulteta za Elektrotehniko, Robotični vid, doc. dr. Žiga Savič, 1. letnik, Smerni Robotika, 3. stopnja UN študijski Elektrotehniko

7. Prilaganje 3D modelov na 2D slike 49

3D-2D poravnava: različne dimenzije

S_{PRED} , CT, MR, S_W , $S_{MED,1}$, $S_{MED,2}$, $S_{MED,N}$

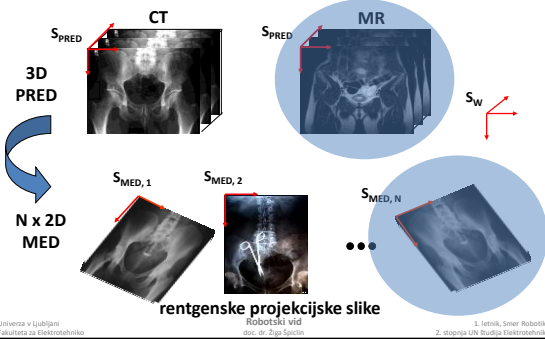
3D PRED

$N \times 2D$ MED

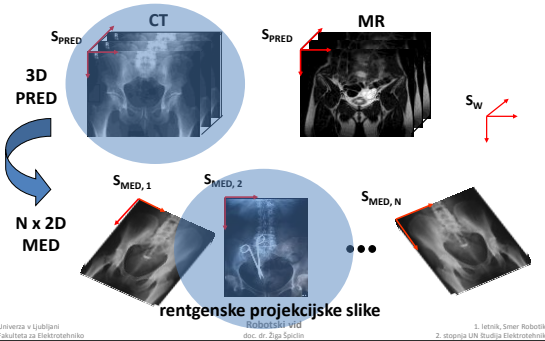
rentgenske projekcijske slike

Univerza v Ljubljani, Fakulteta za Elektrotehniko, Robotični vid, doc. dr. Žiga Savič, 1. letnik, Smerni Robotika, 3. stopnja UN študijski Elektrotehniko

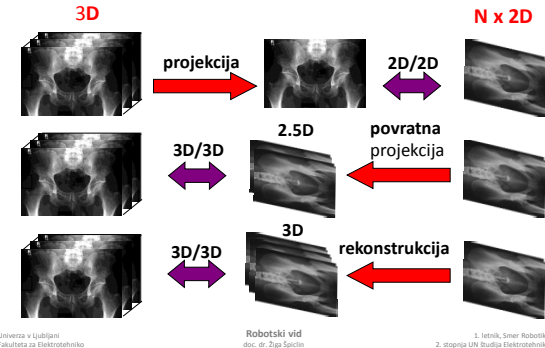
3D-2D poravnava: različne modalitete



3D-2D poravnava: neskladja



Ujemanje dimenzij za poravnavo



7. Prileganje 3D modelov na 2D slike 53

Postopki 3D-2D poravnave Preslikava dimenzije

Markelj et al.,
Media 2008

Projekcija **Povratna proj.** **Rekonstrukcija**

Osnova za poravnavo
Odvodi Intenziteta Značilnice

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, smer Robotika
2. stopnja UNB Študij Elektrotehnika

7. Prileganje 3D modelov na 2D slike 54

Ujemanje dimenzij

Projekcija **Vzratna proj.** **Rekonstr.**

Osnova za poravnavo
Značilnice Intenziteta Gradient

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, smer Robotika
2. stopnja UNB Študij Elektrotehnika

7. Prileganje 3D modelov na 2D slike 55

Postopki, ki temeljijo na značilnicah

- **Značilnice so geometrijske veličine**, kot so izolirane točke ali množica točk, ki predstavljajo krivuljo, obris ali površino
- **Značilnice se izločijo z razmejitvijo** (segmentacijo) ujemajočih se anatomskih struktur na 3D in 2D slikah

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, smer Robotika
2. stopnja UNB Študij Elektrotehnika

7. Prileganje 3D modelov na 2D slike 56

Značilnice / Projekcija (2D-2D) / Vzratna projekcija (3D-3D)

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, smer Robotika
3. stopnja UN Študij Elektrotehnika

7. Prileganje 3D modelov na 2D slike 57

Značilnice / Projekcija (2D-2D)

J. Feldmar et al. 3D-2D projective registration of free-form curves and surfaces. *Computer Vision Image Understanding* 65: 403-424, 1997.

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, smer Robotika
3. stopnja UN Študij Elektrotehnika

7. Prileganje 3D modelov na 2D slike 58

Značilnice / Povratna projekcija (3D-3D)

S. Lavallee et al. Recovering the position and orientation of free-form objects from image contours using 3D distance maps. *IEEE PAMI* 17: 378-390, 1995.

A. Guerez et al. Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot. *IEEE TMI* 17: 715-728, 1998.

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, smer Robotika
3. stopnja UN Študij Elektrotehnika

7. Prileganje 3D modelov na 2D slike 62

Intenziteta / Rekonstrukcija (3D-3D)

2D image 1, 2D image 2, ..., 2D image N

coarse reconstruction of a 3D image

3D image

D. Tomazevic et al. 3-D/2-D registration by integrating 2-D information in 3-D. IEEE TMI 25: 17-27, 2006.

Univerza v Ljubljani, Fakulteta za Elektrotehniko, Robotski vid, doc. dr. Žiga Savič, 3. letnik, Smerni Robotika, 3. stopnja UN Študija Elektrotehnika

7. Prileganje 3D modelov na 2D slike 63

Ujemanje dimenzij

Projekcija Vzratna proj. Rekonstr.

Osnova za poravnavo

Značilnice

Intenziteta

Gradient

Univerza v Ljubljani, Fakulteta za Elektrotehniko, Robotski vid, doc. dr. Žiga Savič, 3. letnik, Smerni Robotika, 3. stopnja UN Študija Elektrotehnika

7. Prileganje 3D modelov na 2D slike 64

Metode, ki slonijo na gradientih

- **Motivacija:** izraziti gradienti intenzitet na rentgenskih slikah so povezani z obrisi (volumenskimi gradienti) 3D anatomskih struktur
- Žarki, ki izhajajo iz izvora žarčenja in padajo na robove struktur na rentgenski sliki, so tangencialni na površine anatomskih struktur

Univerza v Ljubljani, Fakulteta za Elektrotehniko, Robotski vid, doc. dr. Žiga Savič, 3. letnik, Smerni Robotika, 3. stopnja UN Študija Elektrotehnika

Metode, ki slonijo na gradientih

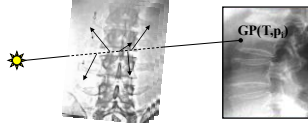
Lastnost projekcije gradienta:

- velikost gradienta v točki rentgenske slike je enaka integralu uteženih projekcij volumenskih gradientov na slikovno ravnino
- utež je enaka relativni razdalji med 3D točko in izvorom žarčenja

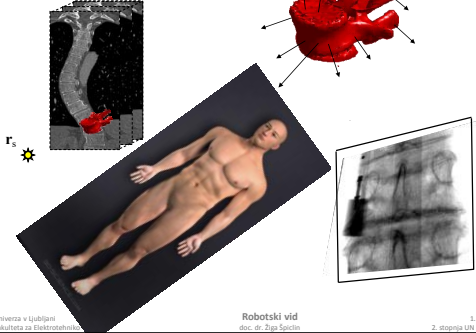
H. Lityatan et al. Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. *IEEE TMI* 22: 1395-1406, 2003.

D. Tomazevic et al. 3-D/2-D registration of CT and MR images to X-ray images. *IEEE TMI* 22: 1407-1416, 2003.

$$GP(T, p_i) \propto \left(\frac{h^T}{v^T} \right) \cdot T \int \nabla V(T^{-1}(p_i, \lambda)) d\lambda$$



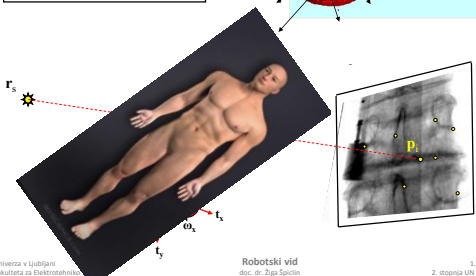
Gradient / Projekcija (2D-2D) / Povratna projekcija (3D-3D)



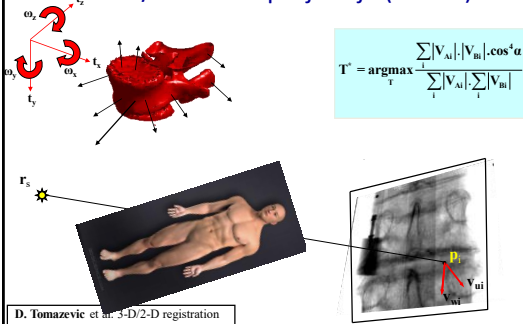
Gradient / Projekcija (2D-2D)

H. Lityatan et al. Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. *IEEE TMI* 22: 1395-1406, 2003.

$$\max_x \sum_i \|GP(T, p_i)\|$$



Gradient / Povratna projekcija (3D-3D)

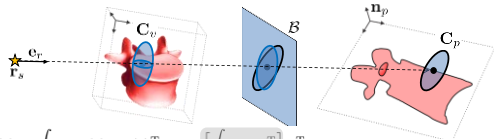


$$T^* = \operatorname{argmax}_T \frac{\sum_i |V_{A_i}| |V_{B_i}| \cdot \cos^4 \alpha}{\sum_i |V_{A_i}| \sum_i |V_{B_i}|}$$

D. Tomazevic et al. 3-D/2-D registration of CT and MR images to X-ray images. *IEEE TMI* 22: 1407-1416, 2003.

Gradient / Povratna projekcija (3D-3D)

- Lokalna kovariančna matrika kodira lokalno obliko anatomske strukture → analiza lastnih vrednosti



$$C_v(\mathbf{r}) = \int_{\Omega_v} \mathbf{g}_{bv}(\mathbf{r}) \mathbf{g}_{bv}(\mathbf{r})^T \approx \Gamma_B \left[\int_{\Omega_v} \mathbf{g}_v \mathbf{g}_v^T \right] \Gamma_B^T$$

$$C_p(\mathbf{r}) = \int_{\Omega_p} \mathbf{g}_{bp}(\mathbf{r}) \mathbf{g}_{bp}(\mathbf{r})^T \approx c^2 \Gamma_B^{-1} \left[\int_{\Omega_p} \mathbf{g}_p \mathbf{g}_p^T \right] \Gamma_B^{-1}$$

izračunana v lokalnem pravokotnem oknu
 $\Omega_v = \Omega_v(\mathbf{r}_v, h)$
 $\Omega_p = \Omega_p(\mathbf{r}_p, c \cdot h)$

Ž. Špičič et al. Fast and robust 3D to 2D image registration by backprojection of gradient covariances. *WBI*, 2014.

Gradient / Povratna projekcija (3D-3D)

- MP** je tenzorski skalarni produkt (TSP)

Peeters et al. *VPTF* 2009

$$s_{TSP}(C_v, C_p) = \sum_{i=1}^n \sum_{j=1}^n \lambda_{v,i} \lambda_{p,j} (\mathbf{e}_{v,i} \circ \mathbf{e}_{p,j})^2$$



– normaliziramo TSP (NTSP), da izenačimo skalo

- NTSP izračunamo za vseh K 3D oslonilnih točk

$$SM = \sum_{i=1}^K s_{NTSP}(C_v, C_p)$$



7. Prileganje 3D modelov na 2D slike 71

Gradient / Rekonstrukcija (3D-3D)

P. Markež et al. Robust gradient-based 3-D/2-D registration of CT and MR to X-ray images. *IEEE TMI* 27: 1704-1714, 2008.

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, Smer Robotika
3. stopnja UNB Študij Elektrotehnika

7. Prileganje 3D modelov na 2D slike 72

Gradient / Rekonstrukcija (3D-3D)

P. Markež et al. Robust gradient-based 3-D/2-D registration of CT and MR to X-ray images. *IEEE TMI* 27: 1704-1714, 2008.

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, Smer Robotika
3. stopnja UNB Študij Elektrotehnika

7. Prileganje 3D modelov na 2D slike 73

Rezultat 3D-2D poravnave se uporabi za ustvarjanje povratne vizualne informacije

- Superpozicija modelov, kirurških orodij, ali 2D rentgenskih slik na pred-operativne 3D slike

Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, Smer Robotika
3. stopnja UNB Študij Elektrotehnika

Povzetek

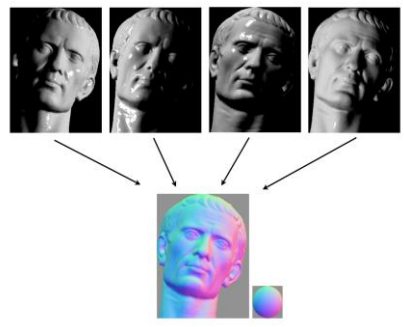
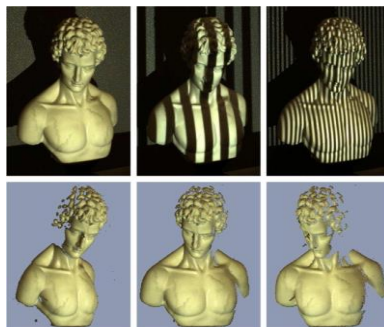
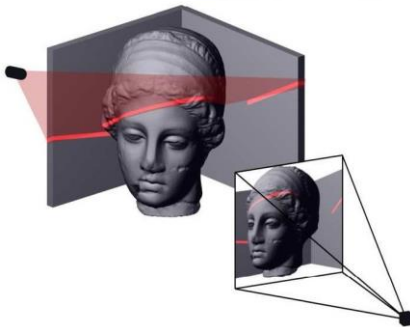
- Za določanje poze iz ene slike moramo predhodno določiti **intrinzične parametre kamere**
 - uporabimo kalibre znane oblike in dimenzije
- Če poznamo **vsaj 6 korespondenčnih parov točk**, potem pozo določimo z **direktno linearno transformacijo**
 - za začetni približek, izboljšamo z optimizacijo nelinearne preslikave
- V kompleksnih scenah (slaba osvetlitev, prekrivanje objektov) izločanje oglišč ali robov vrne veliko osamelcev
 - naredimo robustno z uporabo RANSACa
- **Aplikacija v medicini:** slikovno vodenje posegov
 - določanje poze 3D anatomije iz predoperativne slike v prostoru bolnika na podlagi žive 2D projekcijske slike



8. Rekonstrukcija 3D oblik



Robotski vid



Študijska literatura



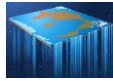
Concise Computer Vision
Reinhard Klette, Springer, 2014
7 3D Shape Reconstruction

Computer vision: algorithms and applications
Richard Szeliski, Springer, 2010
12 3D Reconstruction



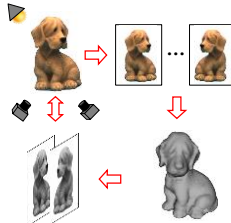
Rekonstrukcija 3D oblik

DEFINICIJA PROBLEMA



• Iskanje optimalne 3D oblike objekta zanimanja na osnovi ene ali več 2D slik, zajetih iz različnih pogledov in/ali pod različnimi osvetlitvami objekta tako, da bo 3D oblika čim boljše sovpadala z zajetimi 2D slikami.

- Vhodni podatki za rekonstrukcijo
 - zbirka 2D slik objekta zanimanja
 - kalibracijski parametri kamere in/ali svetlobnih virov
- Parametri rekonstrukcije
 - zapis 3D informacije oz. oblike
 - kriterij ujemanja 3D oblik in 2D slik (*photo-consistency function*)



Zapis 3D informacije

Z 2D slikami

Spete slike oz. panorama

Slika oddaljenosti

→ od 19. stoletja

→ zastranje/prekrivanje
→ zavisni od pogleda
→ omejena ločljivost

S 3D modeli

Oblak točk

3D slika

Obrisi

Mreža trikotnikov

manj prostorske informacije

več prostorske informacije

Zajem 3D oblik

Prek dotika merjenja

Nedestruktivno
→ robotska tipala

Destruktivno
→ razrez na rezine

Brez dotika merjenja

Z odbojem
→ računalniška tomografija (CT)

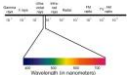
S presvetlitvijo

V območju vidne svetlobe
→ geofotometrija
→ interferometrija
→ čas preleta

Izven območja vidne svetlobe
→ mikrovalovni radar
→ ultrazvočni sonar

Zajem 3D oblik

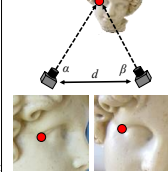
V OBMOČJU VIDNE SVETLOBE



Geofotometrija

Določanje položaja vidne točke s trikotniškimi pravili, npr. z merjenjem stranskih kotov do vidne točke z znano postavitvijo optičnega sistema (triangulacija).

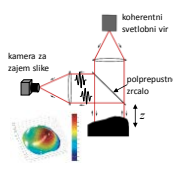
- pasivna
- aktivna



Interferometrija

Merjenje razdalje z detekcijo interference med referenčnim in od objekta odbitim koherentnim valovanjem.

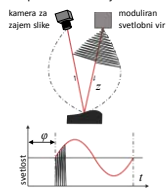
- multivalovna interferometrija
- pegasta interferometrija
- interferom. z belo svetlobo



Čas preleta

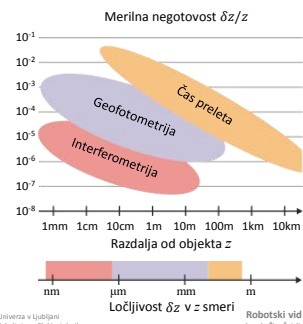
Merjenje razdalje z merjenjem faznega zamika med oddanim in prejetim, od objekta odbitim moduliranim valovanjem.

- zvezno valovanje
- pulzna modulacija



Lastnosti tehnik zajema

V OBMOČJU VIDNE SVETLOBE

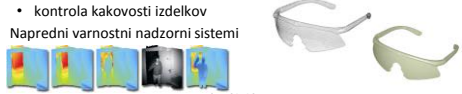


- Geofotometrija**
→ enostavna strojna oprema
→ natančnost zavisi od oblike/površine
→ prostorsko spremenljiva negotovost
- Interferometrija**
→ ločljivost reda $\lambda/100$
→ omejeno območje globine z zaradi preslikovanja faznih zamikov večjih od $\lambda/2$
- Čas preleta**
→ neposredna meritev razdalje
→ zahtevna meritev časa ($c \approx 6,7$ ps/m)
→ ločljivost je konstantna in neodvisna od razdalje



Uporaba rekonstrukcije 3D oblik

- 3D modeliranje
 - stavb in urbanih okolij
 - v medicini
- 3D igre in filmska industrija (avtarji)
- restavracija
- Vzratno inženirstvo (*reverse engineering*)
 - zasnova in razvoj izdelkov
 - kontrola kakovosti izdelkov
- Napredni varnostni nadzorni sistemi



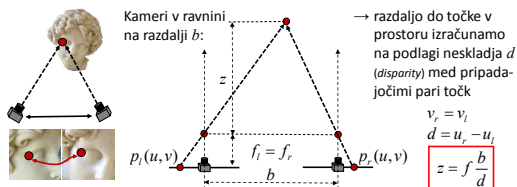
Postopki na osnovi geofotometrije

- Triangulacija
 - stereo vid (pas.)
 - strukturirana svetloba (akt.)
- Oblika iz obrisov
- Oblika iz fokusa
- Oblika iz senc
- Fotometrični stereo



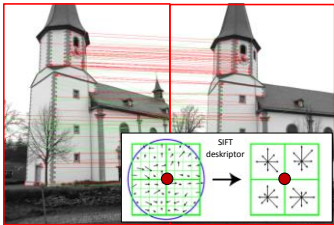
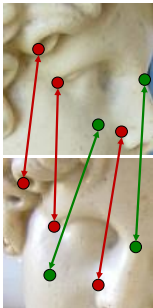
Stereo vid

- 3D stereoskopsko slikanje
 - posnemajo biološke sisteme
- Določanje razdalje od kamere do objekta
 - potrebujemo pripadajoče pare točk in kalibriran optični sistem



Stereo vid

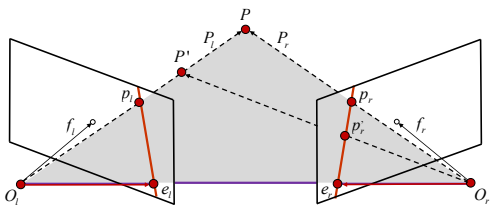
ZAZNAVANJE PRIPADAJOČIH PAROV TOČK



→ značilne točke: Harris detektor ogljšč, LoG, DoG
 → iskanje pripadajočih parov točk: deskriptorji okolice točk SIFT in SURF, RANdom Sample Consensus, epipolarna geometrija

Stereo vid

EPIPOLARNA GEOMETRIJA



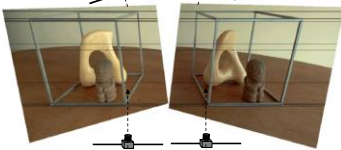
→ vsaki točki v eni sliki pripada točka na epipolarni črti v drugi sliki
 → na ta način omejimo prostor iskanja pripadajočih parov točk

Stereo vid

PRESLIKAVA SLIK V PRAVOKOTNO RAVNINO



→ slikovni ravnini preslikamo v ravnino tako, da sta optični osi kamer vzporedni
 → nato pripadajoči pari točk ležijo vzdolž horizontalnih premic



Stereo vid

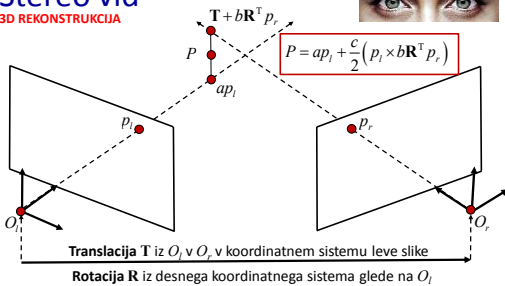
3D REKONSTRUKCIJA



- Za dane pripadajoče pare točk lahko določimo 3D položaj s triangulacijo
- Rekonstrukcija 3D položaja zavisi od „stopnje“ kalibracije optičnega sistema:
 - Poznamo notranje in zunanje parametre kamer
 - možna je metrična 3D rekonstrukcija
 - Poznamo le notranje parametre kamer
 - možna je 3D rekonstrukcija do konstante natančno (faktor skale)
 - Ne poznamo niti notranjih niti zunanjih parametrov
 - možna je 3D rekonstrukcija do projektivne transformacije natančno

Stereo vid

3D REKONSTRUKCIJA



→ Rešimo sistem enačb za skalarje a, b, c : $a p_l + c(p_l \times b R^T p_r) = T + b R^T p_r$

Stereo vid

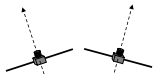
PRIMER REKONSTRUKCIJE



leva slika

desna slika

slika razdalje z

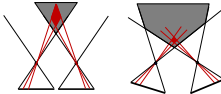


majhen z \leftrightarrow velik z

Stereo vid

PROBLEMI IN OMEJITVE

- Zanesljivost rekonstrukcije zavisi od zanesljivosti iskanja pripadajočih parov točk, le-ta pa zavisi od slikanega prizora in osvetlitve
 - Gladki predmeti brez teksture
 - Periodične teksture oz. vzorci
 - Visoko-odbojne ali prosojne površine
- Kaj storiti na področju, kjer ne najdemo pripadajočih parov točk?
- Prostorsko spremenljiva negotovost dž

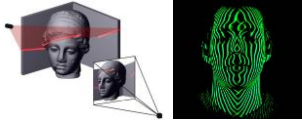


Robotski vid
doc. dr. Žiga Šturm

1. letnik, Smer Robotika
3. stopnja UN Škofja Loka/Elektrotehnika

Strukturirana svetloba

- Ideja:** geometrijska oblika površine popači vzorec projicirane strukturirane svetlobe



- Rekonstrukcija 3D oblike temelji na zajemu in analizi popačenih vzorcev strukturirane svetlobe
 - s primerjavo popačenih in projiciranih vzorcev lahko natančno določimo 3D obliko



Robotski vid
doc. dr. Žiga Šturm

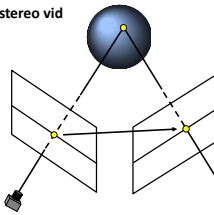
1. letnik, Smer Robotika
3. stopnja UN Škofja Loka/Elektrotehnika

Strukturirana svetloba

PREDNOSTI PRED PASIVNO TRIANGULACIJO

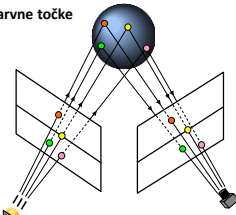
- Enostavna zaznava pripadajočih parov točk:

stereo vid



→ na podlagi slikovne informacije

barvne točke



→ na podlagi projiciranih vzorcev



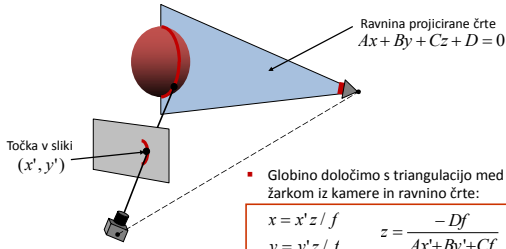
Robotski vid
doc. dr. Žiga Šturm

1. letnik, Smer Robotika
3. stopnja UN Škofja Loka/Elektrotehnika

Strukturirana svetloba

OSNOVNI PRINCIP

- Z laserjem projiciramo črto na objekt



- Globino določimo s triangulacijo med žarkom iz kamere in ravnino črte:

$$\begin{aligned} x &= x'z / f & z &= \frac{-Df}{Ax' + By' + Cf} \\ y &= y'z / f \end{aligned}$$

Strukturirana svetloba

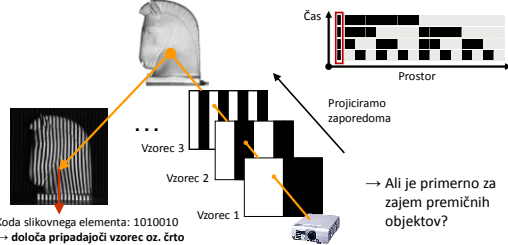
TEHNIKE ZAJEMA IN OSVETLITVE

- Zaporedno projiciranje**
 - več zajemov slik
 - binarna koda
 - Grayeva koda
 - fazni zamik
 - Grayeva koda in fazni zamik
- Naslavljanje z mrežo**
 - en sam zajem slike
 - psevdonaključne binarne točke
 - minivzorci s kodnimi besedami
 - mreža z barvno kodo
 - 2D polje barvnih točk
- Naslavljanje s črtami**
 - en sam zajem slike
 - barvno kodirane črte
 - prekinjene črte
 - črte z Grayevo kodo
 - De Bruijn zaporedje
- Zvezno spremenljiv vzorec**
 - en sam zajem slike
 - 3D zajem z vzorcem mavrice
 - zvezno spremenljiva barvna koda
- Mešane tehnike**

Strukturirana svetloba

ZAPOREDNO PROJICIRANJE BINARNE KODE

- Projiciramo več črt hkrati
 - manj zajemov: imamo $2^n - 1$ črt v n projiciranih slikah



Strukturirana svetloba

NEPOSREDNO NASLAVLJANJE



- Potrebujemo enega ali dva zajema
 - v realnem času, za dinamične ter kompleksne prizore



slika



slika s projiciranimi vzorci



rekonstrukcija oblike

- Kakovost rekonstrukcije zavisi od gostote vzorcev

ena črta
več zajemovveč črt
več zajemovkompleksen vzorec
en zajem

počasno, zanesljivo, natančno

Robotski vid
doc. dr. Žiga Savič

hitro, nezanesljivo, grobo

Univerza v Ljubljani
Fakulteta za Elektrotehniko

1. letnik, Smerni Robotika

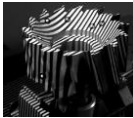
2. stopnja UNB Študija Elektrotehnika

Strukturirana svetloba

POVZETEK LASTNOSTI



- Potrebujemo aktivni vir osvetlitve
 - prostorsko-časovno-barvno spremenljivi vzorci
- Potrebujemo kalibracijo optičnega sistema
- Izogremo se problemu iskanja pripadajočih točk
- Izogremo se problemu izgleda slikanih površin
 - gladke površine brez texture in periodični vzorci
- Visoka natančnost do reda μm
- Primerni za uporabo v industriji
 - sestavljanje delov, kontrola kakovosti



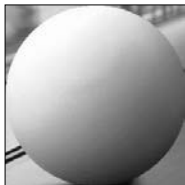
Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, Smerni Robotika

2. stopnja UNB Študija Elektrotehnika

3D oblika in svetlost slik



- Ali lahko sivine izkoristimo za rekonstrukcijo 3D oblik?
- Ali obstaja povezava med obliko in sivinami v sliki?

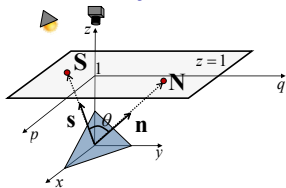
Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Savič

1. letnik, Smerni Robotika

2. stopnja UNB Študija Elektrotehnika

Interakcija svetlobni vir-površina



Enačba ravnine: $Ax + By + Cz + D = 0$

$$\text{oz. } \frac{A}{C}x + \frac{B}{C}y + z + \frac{D}{C} = 0$$

Označimo:

Vektor v smeri normale na površino:

$$-\frac{\partial z}{\partial x} = \frac{A}{C} = p \text{ in } -\frac{\partial z}{\partial y} = \frac{B}{C} = q \iff \mathbf{N} = \left(\frac{A}{C}, \frac{B}{C}, 1 \right) = (p, q, 1)$$

Definiramo **prostor odvodov**
 \rightarrow prostor pq oz. ravnina $z = 1$
 \rightarrow vsaka točka pripada eni normalni

Normala na površino

$$\mathbf{n} = \frac{\mathbf{N}}{\|\mathbf{N}\|} = \frac{(p, q, 1)}{\sqrt{p^2 + q^2 + 1}}$$

Smer svetlobnega vira

$$\mathbf{s} = \frac{\mathbf{S}}{\|\mathbf{S}\|} = \frac{(p_s, q_s, 1)}{\sqrt{p_s^2 + q_s^2 + 1}}$$

Kot med vektorjema

$$\cos \theta = \mathbf{n} \cdot \mathbf{s}$$

Model odbojnosti površine



- Povezuje svetlost slikovnih elementov $I(x, y)$ z orientacijo površine (p, q) za dano smer svetlobnega vira (p_s, q_s)

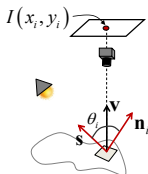
- Lambertov model difuzne površine:

k : svetlost svetlobnega vira

ρ : albedo oz. odbojnost površine

c : konstanta optičnega sistema

$$I(x, y_i) = \frac{\rho}{\pi} kc \cos \theta_i = \frac{\rho}{\pi} kc \mathbf{n}_i \cdot \mathbf{s}$$



- Če $\frac{\rho}{\pi} kc = 1$ potem:

$$I(x_i, y_i) = \cos \theta_i = \mathbf{n}_i \cdot \mathbf{s}$$

Model odbojnosti površine

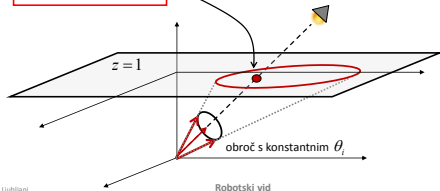


- Lambertov model:

$$I(x_i, y_i) = \cos \theta_i = \mathbf{n}_i \cdot \mathbf{s} = \frac{(p_i p_s + q_i q_s + 1)}{\sqrt{p_i^2 + q_i^2 + 1} \sqrt{p_s^2 + q_s^2 + 1}} = R(p_i, q_i)$$

Izočrta enake svetlosti

Model odbojnosti



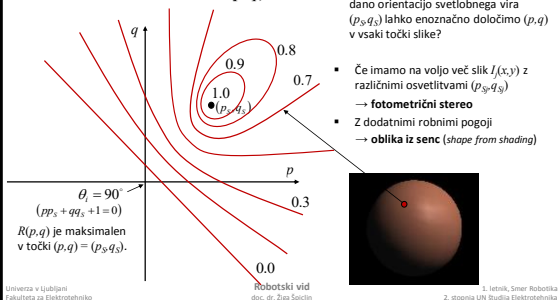
Model odbojnosti površine

PRIMER

Lambertov model $R(p, q)$

→ Ali za dan model odbojnosti $R(p, q)$ in dano orientacijo svetlobnega vira (p_s, q_s) lahko enoznačno določimo (p, q) v vsaki točki slike?

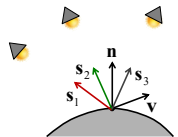
- Če imamo na voljo več slik $I(x, y)$ z različnimi osvetlitvami (p_s, q_s) → **fotometrični stereo**
- Z dodatnimi robnimi pogoji → **oblika iz senc** (*shape from shading*)



Fotometrični stereo

Lambertov model:

$$I = \frac{\rho}{\pi} kc \cos \theta = \rho \mathbf{n} \cdot \mathbf{s} \quad \left(\frac{kc}{\pi} = 1 \right)$$

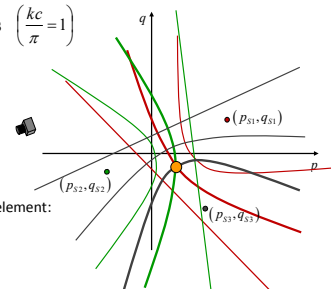


Sistem enačb za vsak slikovni element:

$$I_1 = \rho \mathbf{n} \cdot \mathbf{s}_1$$

$$I_2 = \rho \mathbf{n} \cdot \mathbf{s}_2$$

$$I_3 = \rho \mathbf{n} \cdot \mathbf{s}_3$$



1. letnik, Smer Robotika
2. stopnja UNI Beograd, Elektrotehnika

Fotometrični stereo

REŠEVANJE ENAČB

Enačbe zapišemo v matrični obliki:

$$\begin{cases} I_1 = \rho \mathbf{n} \cdot \mathbf{s}_1 \\ I_2 = \rho \mathbf{n} \cdot \mathbf{s}_2 \\ I_3 = \rho \mathbf{n} \cdot \mathbf{s}_3 \end{cases} \Leftrightarrow \underbrace{\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}}_{\mathbf{I}^{(3 \times 1)}} = \underbrace{\begin{bmatrix} \mathbf{s}_1^T \\ \mathbf{s}_2^T \\ \mathbf{s}_3^T \end{bmatrix}}_{\mathbf{S}^{(3 \times 3)}} \underbrace{\rho \mathbf{n}}_{\tilde{\mathbf{n}}^{(3 \times 1)}} \Leftrightarrow$$

Rešitev:

$$\tilde{\mathbf{n}} = \mathbf{S}^{-1} \mathbf{I}$$

$$\rho = \frac{\|\tilde{\mathbf{n}}\|}{\|\mathbf{n}\|}$$

$$\mathbf{n} = \frac{\tilde{\mathbf{n}}}{\|\tilde{\mathbf{n}}\|} = \frac{\tilde{\mathbf{n}}}{\rho}$$

Robustna ocena z N svetili in uteževanjem s svetlostjo

$$\begin{cases} I_1(I_1) = I_1(\rho \mathbf{n} \cdot \mathbf{s}_1) \\ \vdots \\ I_N(I_N) = I_N(\rho \mathbf{n} \cdot \mathbf{s}_N) \end{cases} \Leftrightarrow \underbrace{\begin{bmatrix} I_1^2 \\ \vdots \\ I_N^2 \end{bmatrix}}_{\mathbf{I}^{(N \times 1)}} = \underbrace{\begin{bmatrix} I_1 \mathbf{s}_1^T \\ \vdots \\ I_N \mathbf{s}_N^T \end{bmatrix}}_{\mathbf{S}^{(N \times 3)}} \underbrace{\rho \mathbf{n}}_{\tilde{\mathbf{n}}^{(3 \times 1)}} \Leftrightarrow$$

Rešitev:

$$\mathbf{I} = \mathbf{S} \tilde{\mathbf{n}}$$

$$\mathbf{S}^T \mathbf{I} = \mathbf{S}^T \mathbf{S} \tilde{\mathbf{n}}$$

$$\tilde{\mathbf{n}} = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \mathbf{I}$$

1. letnik, Smer Robotika
2. stopnja UNI Beograd, Elektrotehnika

Fotometrični stereo

DOLOČANJE GLOBINE



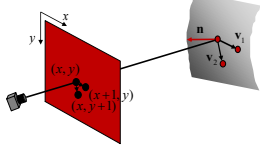
- Analiza odvodov površine

$$v_1 = (x+1, y, z(x+1, y)) - (x, y, z(x, y)) = (1, 0, z(x+1, y) - z(x, y))$$

$$v_2 = (x, y+1, z(x, y+1)) - (x, y, z(x, y)) = (0, 1, z(x, y+1) - z(x, y))$$

Velja:
 $n \cdot v_1 = 0, n \cdot v_2 = 0$

Rešitev:
 $Mz = v \rightarrow M$ je redka matrika
 $z = (M^T M)^{-1} M^T v$

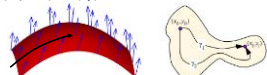


- Integriranje po krivulji:

→ z določajo prispevki normal n po poljubni poti $s(x, y)$

$$n = (n_x, n_y, n_z)^T$$

$$z(x, y) = \int_0^x n_x(u, y) du + \int_0^y n_y(x, v) dv$$

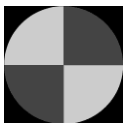
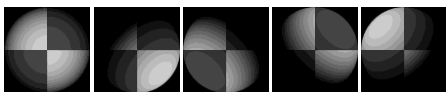


Fotometrični stereo

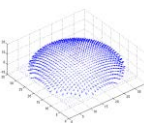
PRIMER REKONSTRUKCIJE 3D OBLIKE



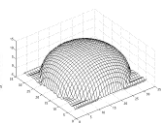
vhodne slike



rekonstruirana odbojnost površine



rekonstruirane normale površine



rekonstruirana 3D oblika

Fotometrični stereo

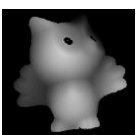
PRIMER REKONSTRUKCIJE 3D OBLIKE



vhodne slike



rekonstruirana odbojnost površine



rekonstruirana 3D oblika



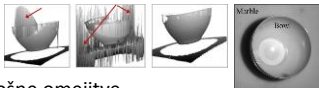
rekonstruirana 3D oblika s teksturo

Fotometrični stereo

PROBLEMI IN OMEJITVE



- Ne deluje za predmete z zrcalno površino ali prosojne predmete
- Motijo lastne sence in večkratni odboji svetlobe



- Splošne omejitve
 - Kamera in svetlobni viri morajo biti daleč od objekta
 - Potrebna je kalibracija sistema
 - kalibracija položaja in svetlosti svetlobnih virov
 - svetlostna kalibracija kamere

Povzetek

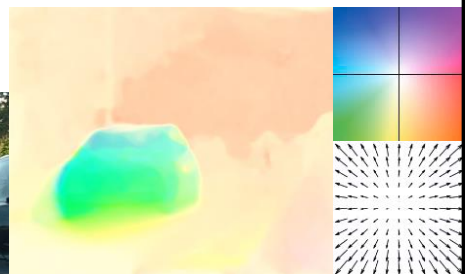
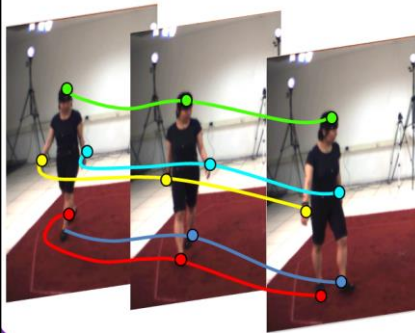
- Osnovni principi v vidnem delu svetlobe temeljijo na **geofotometriji, interferometriji in času preleta**
- S pomočjo geofotometrije 3D obliko rekonstruiramo
 - na osnovi ene ali več 2D slik, zajetih iz različnih pogledov in/ali
 - pod različnimi osvetlitvami objekta
 tako, da bo 3D oblika čim bolj sovpadala z zajetimi 2D slikami
- **Osnovne tehnike z uporabo kamere in svetil so**
 - triangulacija: stereo vid (pasivna-), strukturirana svetloba (aktivna osv.)
 - modeliranje odbojnosti: fotometrični stereo, oblika iz senc
 - obrezovanje s povratno projekcijo obrisov (podobno kot rekonstrukcija CT slik)
 - zamegljenost objektov (merjenje stopnje razširitve točke – PSF)



9. Vizualna navigacija



Robotski vid



Univerza v Ljubljani
Fakulteta za Elektrotehniko

Robotski vid
doc. dr. Žiga Špiclin

1. letnik, Smer Robotika
2. stopnja UN študija Elektrotehnika

Študijska literatura



Concise Computer Vision
Reinhard Klette, Springer, 2014
4 Dense motion analysis
9.3 Tracking and updating features

**Understanding the Basis of the Kalman Filter
Via a Simple and Intuitive Derivation**
Ramsey Faragher, *IEEE Signal Processing Magazine*, 29(5), 2012



Kalman filter for robot vision: a survey
S.Y. Chen, *IEEE T Industrial Electronics*, 59 (11), 2012



Eppur si muove (Galileo Galilei)



- Osnova za vizualno navigacijo je zaznavanje, razumevanje in predvidevanje oz. **analiza gibanja**

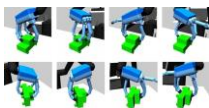


Vizualna navigacija

DEFINICIJA PROBLEMA



- Na podlagi zaporedja slik iz ene ali več kamer analiziramo gibanje objektov v 3D prostoru in s to informacijo načrtujemo povratno-zančno vodenje, manipulacijo, sledenje ter izvajamo rekonstrukcijo in detekcijo objektov.
- Sorodni problemi in rešitve
 - robustno zaznavanje objektov zanimanja
 - iskanje korespondenc med slikami in poravnava slik
 - kalibracija parametrov kamere in/ali svetlobnih virov
 - predstavitev objektov, modeli pojavnosti





Uporaba analize gibanja

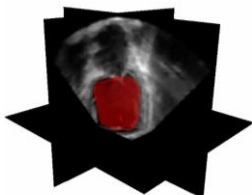
- Izboljšava videa oz. zaporedja slik (image and video enhancement)
 - stabilizacija, povečanje ločljivosti, zmanjšanje šuma





Uporaba analize gibanja

- Sledenje in razpoznavanje objektov
 - upravljanje vozila in pomoč vozniku, biomedicina





Uporaba analize gibanja

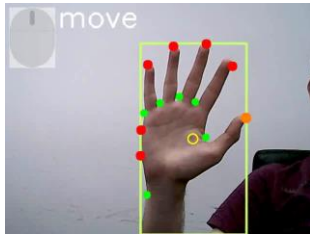
- Nadzorni (varnostni) sistemi
 - nadzor ulic in prometa, varovanje zgradb





Uporaba analize gibanja

- **Uporabniški vmesniki** (human computer interfaces - HCI)
→ razpoznavanje gibov, gest, načina hoje, sledenje očem





Izvori gibanja v zajetih slikah

- **Gibanje osvetlitve, objektov in/ali kamere**
- **Kombinacija teh izvorov določa uporabo, pogosto pa tudi spekter možnih rešitev**



Fiksna kamera, premični objekti
→ notranji nadzorni sistemi



Premična kamera, fiksni prizori
→ 3D rekonstrukcija, merjenje hitrosti



Premična kamera, premični prizori
→ sledenje objektu zanimanja: šport, filmi



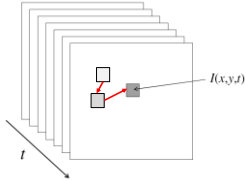
Fiksna kamera, premični objekti, sprem. osvetlitve
→ nadzor prometa, zunanji nadzorni sistemi

Analiza gibanja iz videa

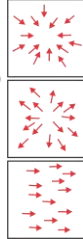
- **Video je zaporedje slik zajetih v času**
– slikovna informacija je funkcija prostora (x, y) in časa (t)

→ iskanje korepondenc

→ smeri oz. polje gibanja



Korepondence potrebujemo za določanje trajektorije tarče, konsistentno označevanje tarče, določanje lastnosti tarče (orientacija, področje v sliki, oblika)



Pomanjšava (-), oddaljevanje kamere ali oddaljevanje tarče

Povečava (+), približevanje kamere ali približevanje tarče

Premik kamere v levo ali gibanje tarče v desno

Izzivi pri analizi gibanja

- Izguba informacije pri preslikavi iz 3D prostora v 2D slikovno ravnino
- Spremembe jakosti in smeri osvetlitve, sence, odboji
- Šum v zajetih slikah, zameglitev zaradi gibanja kamere (motion blur)
- Obdelava slik v realnem-času (sproti, npr. 30 fps)
- Vizualno podobne tarče in interakcija med tarčami**
- Analiza zapletenega gibanja (**nezveznost, hitre spremembe**)
- Sprememba izgleda tarče**
- Deformacije in sestavljeni gibi
- Delno ali popolno zakrivanje**

Kompleksnost analize gibanja zavisi od količin in ciljev!

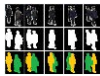


Predpostavke pri analizi gibanja

- Gibanje tarče:** zvezno brez naglih sprememb, konstantna hitrost ali konstanten pospešek
- Poznamo število, velikost, izgled in obliko tarče
- Geometrija tarče:** toga ali počasi spreminjajoča tarča
- Izgled tarče:** neprozorne površine, brez senc in popolnih odbojev (Lambertove površine)
- Osvetlitev:** fiksna ali pa se spreminja počasi
- Zajem slike:** brez zameglitve (no motion blur), malo šuma

Zaznavanje tarče

ODŠTEVANJE OZADJA



- Enostavna in pogosto uporabljena tehnika za zaznavanje premičnih tarč pri fiksni kameri



slika $I(x, y, t)$ ozadje $B(x, y, t)$ maska potencialnih tarč

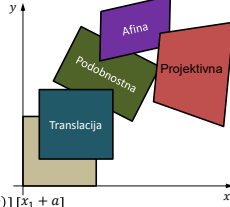
- Ozadje $B(x, y, t)$ ocenimo iz videa $I(x, y, t)$
 - kot prejšnjo sliko iz zaporedja $B(x, y, t) = I(x, y, t - 1)$
 - s povprečenjem zaporednih slik $B(x, y, t) = \frac{1}{N} \sum_{i=0}^{N-1} I(x, y, t - i)$
 - z mediano zaporednih slik $B(x, y, t) = \text{median}\{I(x, y, t - i)\}_{i=0, \dots, N-1}$

Sledenje s parametričnim modelom

- Preslikava: $(x_1, y_1) \rightarrow (x_2, y_2)$
 - (x_1, y_1) : točka v sliki 1
 - (x_2, y_2) : korespondenčna točka v sliki 2
- Globalni model preslikave
 - $(x_2, y_2) = f(x_1, y_1; \theta)$

Preslikave

- Translacija: $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + a \\ y_1 + b \end{bmatrix}$
- Podobnostna: $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = s \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} x_1 + a \\ y_1 + b \end{bmatrix}$
- Afina: $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} ax_1 + by_1 + c \\ dx_1 + ey_1 + f \end{bmatrix}$
- Projektivna: $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} ax_1 + by_1 + c \\ dx_1 + ey_1 + f \end{bmatrix}, z = gx_1 + hy_1 + i$



Mera podobnosti za sledenje

- Privzamemo, da so intenzitete konstantne

$$I_1(x, y) = I_2(x + u, y + v) + n \quad \leftarrow \text{šum}$$

$$n \sim N(0, \sigma^2)$$

- Podobnost med slikama

- Katere lastnosti slik se ne spremenijo?
 - Svetlost, gradienti, faza, druge značilnice (katere?)
- Mera razdalje (L2 ali bolj robustne funkcije)

$$E(u, v) = \sum_{x,y} (I_1(x, y) - I_2(x + u, y + v))^2$$

→ korelacija, korelacijski koeficient (NCC)

Sledenje tarči s param. modelom

POSTOPEK LUCAS-KANADE

- PROBLEM:** Za dani sliki $I_1(x, y)$ in $I_2(x, y)$ določi parametrično preslikavo, ki sliko I_1 preslika v sliko I_2
- Naj $\mathbf{x} = (x, y)^T$ označuje stolpni vektor s koordinatami pikslov
- Najpogostejši preslikavi

- Translacija: $W(\mathbf{x}; \mathbf{p}) = \begin{bmatrix} x + p_1 \\ y + p_2 \end{bmatrix}$

- Afina: $W(\mathbf{x}; \mathbf{p}) = \begin{bmatrix} p_1x + p_3y + p_5 \\ p_2x + p_4y + p_6 \end{bmatrix} = \begin{bmatrix} p_1 & p_3 & p_5 \\ p_2 & p_4 & p_6 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

- CILJ:** minimizacija mere podobnosti

$$\mathbf{p}^* = \arg \min_{\mathbf{p}} \sum_{\mathbf{x}} [I_2(W(\mathbf{x}; \mathbf{p})) - I_1(\mathbf{x})]^2$$

Matrika prvih odvodov (Jacobi)

POSTOPEK LUCAS-KANADE

- $\nabla I_2 = [I_x \ I_y]$ je prvi odvod I_2 v točki $W(\mathbf{x}; \mathbf{p})$

→ izračunaj odvode slike I_2 in jih preslikaj v I_1

- Afina transformacija: $\frac{\partial W}{\partial \mathbf{p}} = \begin{bmatrix} x & 0 & y & 0 & 1 & 0 \\ 0 & x & 0 & y & 0 & 1 \end{bmatrix}$
 $\nabla I_2 \frac{\partial W}{\partial \mathbf{p}} = [I_x x \quad I_y x \quad I_x y \quad I_y y \quad I_x \quad I_y]$

- Naj bo matrika $\mathbf{B} = [I_x X \ I_y X \ I_x Y \ I_y Y \ I_x \ I_y] \in \mathbb{R}^{n \times 6}$
 → I_x in X sta stolpna vektorja in produkt $I_x X$ je množenje po komponentah

Gauss-Newton optimizacija

POSTOPEK LUCAS-KANADE

- Mera podobnosti

$$\Delta \mathbf{p}^* = \arg \min_{\Delta \mathbf{p}} \sum_x \left[I_2(W(\mathbf{x}; \mathbf{p})) + \nabla I_2 \frac{\partial W}{\partial \mathbf{p}} \Delta \mathbf{p} - I_1(\mathbf{x}) \right]^2$$

Oziroma v vektorski obliki:

$$\Delta \mathbf{p}^* = \arg \min_{\Delta \mathbf{p}} (\mathbf{I}_t + \mathbf{B} \Delta \mathbf{p})^T (\mathbf{I}_t + \mathbf{B} \Delta \mathbf{p})$$

Kjer je $\mathbf{B} = [I_x X \ I_y X \ I_x Y \ I_y Y \ I_x \ I_y] \in \mathbb{R}^{n \times 6}$

$$\mathbf{I}_t = I_2(W(\mathbf{p})) - I_1$$

- REŠITEV:

$$\Delta \mathbf{p}^* = -(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{I}_t$$

aproximacija Hessianove matrike

Gauss-Newton optimizacija

POSTOPEK LUCAS-KANADE

- Podobno za translacijo

$$\text{Jacobijeva matrika: } \frac{\delta W}{\delta \mathbf{p}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\nabla I_2 \frac{\delta W}{\delta \mathbf{p}} = [I_x \ I_y]$$

- Naj bo matrika $\mathbf{B} = [I_x \ I_y] \in \mathbb{R}^{n \times 2}$

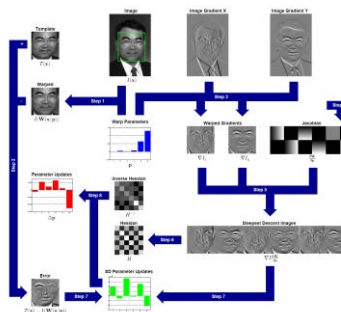
- REŠITEV:

$$\Delta \mathbf{p}^* = -(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{I}_t = - \begin{bmatrix} I_x^T I_x & I_x^T I_y \\ I_x^T I_y & I_y^T I_y \end{bmatrix}^{-1} \begin{bmatrix} I_x^T I_t \\ I_y^T I_t \end{bmatrix}$$

struktura matrika
(Harris detektor oglišč)

Kako deluje?

POSTOPEK LUCAS-KANADE

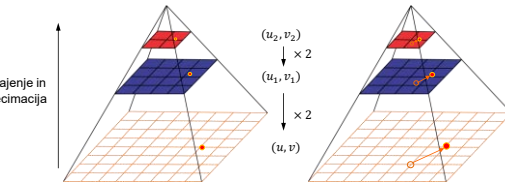


1. Izračunaj:
 $\mathbf{B} = \begin{bmatrix} \nabla I_2 & \frac{\partial W}{\partial \mathbf{p}} \end{bmatrix}$
2. Izračunaj inverz Hessian matrike:
 $(\mathbf{B}^T \mathbf{B})^{-1}$
3. Izračunaj:
 $\mathbf{B}^T \mathbf{I}_t$
4. Reši linearen sistem:
 $\Delta \mathbf{p}^* = -(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{I}_t$
5. Posodobí parametre:
 $\mathbf{p} \leftarrow \mathbf{p} + \Delta \mathbf{p}^*$

Piramidna shema za konvergenco

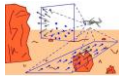
POSTOPEK LUCAS-KANADE

- Lucas-Kanade postopek konvergira v najbližji lokalni minimum
- Dober začetni približen je potreben za ustrezno sledenje
 - če so premiki med slikama $I_1(x, y)$ in $I_2(x, y)$ veliki, uporabi piramidno shemo
 - najprej ocenimo grobe premike pri nizki ločljivosti, nato fine pri visoki ločljivosti



Optični tok

POSTOPEK LUCAS-KANADE



- Korak za oceno optičnega toka (du, dv) v pikslu (x, y)

$$I_2(x + u + du, y + v + dv) - I_1(x, y)$$

$$= I_2(x + u, y + v) + I_x(x + u, y + v)du + I_y(x + u, y + v)dv - I_1(x, y)$$

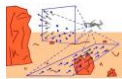
$$I_x du + I_y dv + I_t = 0$$
- Rešitev (du, dv) za $K \times K$ okno

$$\begin{bmatrix} du \\ dv \end{bmatrix} = - \begin{bmatrix} I_x^T I_x & I_x^T I_y \\ I_x^T I_y & I_y^T I_y \end{bmatrix}^{-1} \begin{bmatrix} I_x^T I_t \\ I_y^T I_t \end{bmatrix}$$

- Vektor optičnega toka posodobimo v vsakem pikslu
- Gladkost optičnega toka zagotovimo z medianim filtriranjem

Optični tok - primer

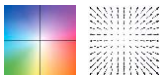
POSTOPEK LUCAS-KANADE



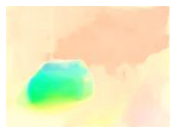
Vhodni sliki



LK s piramidno shemo



Vizualizacija optičnega toka



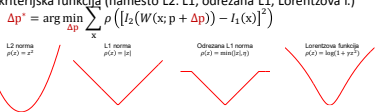
LK s piramidno shemo in medianim filtrom



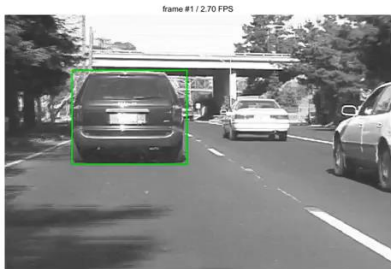
Lastnosti Lucas-Kanade



- Ocena parametričnega gibanja in optičnega toka (z medianim filtriranjem)
 - natančnost pod velikostjo slikovnega elementa
 - ocena gibanja je nezanesljiva na nestrukturiranih delih slike (homogene regije in ravni robovi, inverz Hessiana ne obstaja!)
 - pri daljšem sledenju se napaka ocene gibanja akumulira (**lezenje**)
- Nadgradnje osnovnega postopka:
 - groba ocena premika s poravnavo po blokih (MPEG: bloki 16 × 16)
 - sledenje zanesljivim tarčam in interpolacija premika v nezanesljivih pikslih
 - robustna kriterijska funkcija (namesto L2: L1, odrezana L1, Lorentzova f.)



Lezenje pri sledenju z Lucas-Kanade



<https://www.youtube.com/watch?v=4oy60OeOwPQ>

Upoštevanje dinamike tarče pri določanju premika

Uveljavljeni postopki

Kalmanov filter

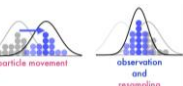
- za linearne dinamične sisteme, enomodalne verjetnostne porazdelitve

Razširjeni Kalmanov filter

- za nelinearne dinamične sisteme, enomodalne verjetnostne porazdelitve

Filter z delci

- večmodalne verjetnostne porazdelitve
- CONDENSATION - conditional density propagation



Kalmanov filter

MOTIVACIJA

- Iterativni postopek** za zmanjšanje pogoška pri meritvah iz različnih virov – pri sledenju filtra zaporedje pošumljenih meritev ocene premika
 - ocene optimalne v smislu srednje kvadratne napake med oceno stanja in dejanskim stanjem

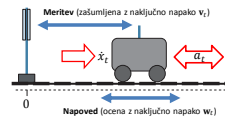
Primer:



Kalmanov filter: napoved + meritev → **boljša ocena premika**

Kalmanov filter

NAPOVEDOVANJE PREMIKA



Napoved položaja in hitrosti:

$$x_t = x_{t-1} + (\dot{x}_{t-1} \times \Delta t) + a_t \frac{(\Delta t)^2}{2}$$
$$\dot{x}_t = \dot{x}_{t-1} + a_t \Delta t$$

V matrični obliki:

$$\begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \dot{x}_{t-1} \end{bmatrix} + a_t \begin{bmatrix} \frac{(\Delta t)^2}{2} \\ \Delta t \end{bmatrix}$$

oziroma:

$$x_t = F_t x_{t-1} + w_t$$

Kovarianca suma:

$$Q_t = E[w_t w_t^T]$$

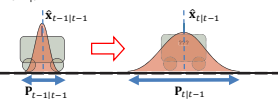
1. Napoved premika v času t na podlagi predhodne ocene

- imamo filtrirano oceno premika $\hat{x}_{t-1} = \hat{x}_{t-1|t-1}$
- in njeno negotovost → kovarianca $P_{t-1|t-1}$

$$\hat{x}_{t|t-1} = F_t \hat{x}_{t-1|t-1}$$

$$P_{t|t-1} = F_t P_{t-1|t-1} F_t^T + Q_t$$

→ izpeljava $P_{t|t-1} = E[(\tilde{x}_t - \hat{x}_{t|t-1})(\tilde{x}_t - \hat{x}_{t|t-1})^T]$



Kalmanov filter

KOMBINIRANJE NAPOVEDI IN MERITVE PREMIKA

Meritev glede na stanja modela:

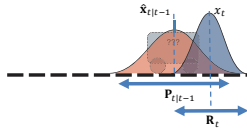
$$\begin{bmatrix} x_t \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} v_t \\ 0 \end{bmatrix}$$

v matrični obliki:

$$y_t = H_t x_t + v_t$$

Kovarianca šuma:

$$R_t = E(v_t v_t^T)$$



2. Nova ocena v času t s kombiniranjem napovedi in meritve

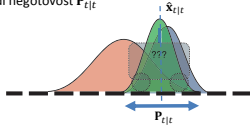
- posodobimo tako oceno $\hat{x}_{t|t}$ kot tudi negotovost $P_{t|t}$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (y_t - H_t \hat{x}_{t|t-1})$$

$$P_{t|t} = P_{t|t-1} - K_t H_t P_{t|t-1}$$

Kalmanovo ojačenje (Kalman gain):

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1}$$



Kombiniranje napovedi in meritve

- Imamo dve neodvisni oceni x_1, x_2 spremenljivke x

- prva ima varianco σ_1^2 , druga pa σ_2^2
- kakšna je njuna optimalna kombinacija, ki da oceno \hat{x} ?

$$\hat{x} = w_1 x_1 + w_2 x_2 \quad \text{pri čemer je} \quad w_1 + w_2 = 1$$

- optimalni vrednosti uteži naj minimizirata varianco σ^2 ocene \hat{x}

$$\sigma^2 = E[(\hat{x} - E[\hat{x}])^2] = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2$$

$$\rightarrow \text{naj bo } w_2 = w, \text{ potem } w_1 = 1 - w$$

$$\sigma^2 = (1 - w)^2 \sigma_1^2 + w^2 \sigma_2^2$$

$$\frac{d\sigma^2}{dw} = -2(1 - w)\sigma_1^2 + 2w\sigma_2^2 = 0$$

Optimalna kombinacija je produkt varianc (kovarianc)
→ enostavno v primeru, da je šum Gaussove oblike

Rešitev: $w = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$

iz katere sledi:

$$\hat{x} = \frac{\sigma_2^2 x_1 + \sigma_1^2 x_2}{\sigma_1^2 + \sigma_2^2} = x_1 + \frac{\sigma_1^2 (x_2 - x_1)}{\sigma_1^2 + \sigma_2^2}$$

$$\sigma^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} = \sigma_1^2 - \frac{\sigma_1^4}{\sigma_1^2 + \sigma_2^2}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (y_t - H_t \hat{x}_{t|t-1})$$

$$P_{t|t} = P_{t|t-1} - K_t H_t P_{t|t-1}$$

Kalmanovo ojačenje (Kalman gain):

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1}$$

$$K_t = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

Predpostavke Kalmanovega filtra

- Matrice procesa F_t, H_t so natančno poznane
- Procesni šum w_t in merilni šum v_t sta medsebojno nekorelirana bela (Gaussova) šuma z ničelno srednjo vrednostjo
 - poznamo njihni kovariančni matriki Q_t, R_t
- Podani sta začetni oceni vektorja stanj $\hat{x}_{0|0}$ in pripadajoča kovariančna matrika $P_{0|0}$

Izvajanje Kalmanovega filtra

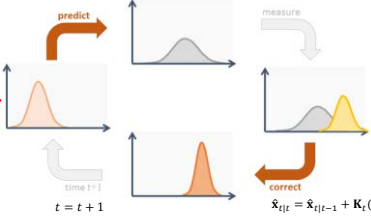
- Iterativno izvajanje v času t

$$\hat{x}_{t|t-1} = F_t \hat{x}_{t-1|t-1}$$

$$P_{t|t-1} = F_t P_{t-1|t-1} F_t^T + Q_t$$

$$y_t = H_t x_t + v_t$$

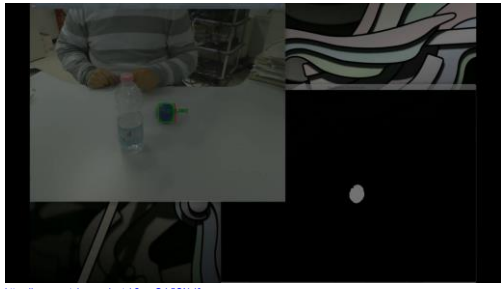
Začetne vrednosti:
 $\hat{x}_{0|0}, P_{0|0}$
 Q_t, R_t



$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (y_t - H_t \hat{x}_{t|t-1})$$

$$P_{t|t} = P_{t|t-1} - K_t H_t P_{t|t-1}$$

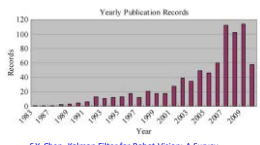
Kalmanov filter – primer



<https://www.youtube.com/watch?v=6-G-H5ONeJs>

Različice Kalmanovega filtra

- Kalman filter
- self-tuning Kalman filter
- steady-state Kalman filter
- ensemble Kalman filter
- adaptive Kalman filter
- switching Kalman filter
- fuzzy Kalman filter
- extended Kalman filter (EKF)
- motor EKF
- hybrid EKF
- augmented state EKF
- modified covariance EKF
- iterative adaptive EKF
- unscented Kalman filter (UKF)
- ...



S.Y. Chen, Kalman Filter for Robot Vision: A Survey, IEEE T Industrial Electronics, 59 (11), 2012

Filter z delci (CONDENSATION algorithm)

LASTNOSTI V PRIMERJAVI S KALMANOVIM FILTROM

→ Razširjanje pogojne gostote verjetnosti (stanj) v času

- | | |
|--|---|
| <ul style="list-style-type: none"> ▪ KALMANOV FILTER • Predpostavlja enomodalno (Gaussovo) porazdelitev • Napove eno novo stanje za vsak objekt, ki mu sledimo • Posodobi stanje glede na napako med napovedanim stanjem in meritvami | <ul style="list-style-type: none"> ▪ FILTER Z DELCI • Podpira poljubne večmodalne porazdelitve • Napove več možnih stanj za vsak objekt, ki mu sledimo • Vsako stanje ima različno verjetnost • Verjetnosti napovedanih stanj temeljijo na meritvah |
|--|---|

Filter z delci – primer



<https://www.youtube.com/watch?v=wCMk-phzScE>

Povzetek

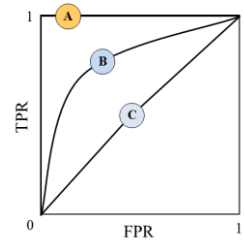
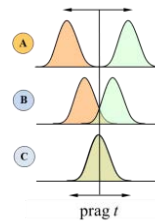
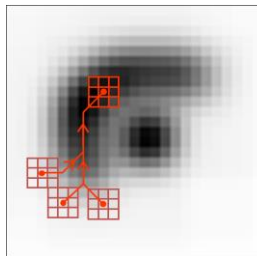
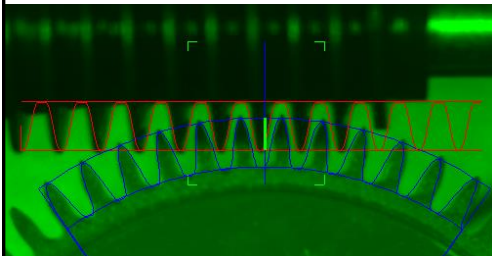
- S sledenjem objektu v zaporedju slik (videu) lahko analiziramo gibanje objektov (**tarče**) v 3D prostoru
 - je kot poravnava, ki čim bolj izkorišča dober začetni približek
- Osnovni gradniki sledenja tarči so
 - zaznavanje objektov (odštevanje ozadja, razgradnja, razvrščevalniki)
 - iskanje korespondenc med slikami (mere podobnosti, deskriptorji)
 - predstavitev tarče, modeliranje oblike/izgleda/dinamike
- Tarčo je lahko zapisana kot točka, geometrijski lik, obris, skelet
- Postopki sledenja tarči
 - sledenje točk: deterministični, statistični (Kalmanov filter, filter z delci)
 - sledenje predlogi: Lucas-Kanade, Mean-shift
 - sledenje obrisom: (netoga) poravnava oblik, prilagodljivi modeli (kače)
 - kombinacije omenjenih postopkov (npr. Lucas-Kanade in Kalmanov filter)



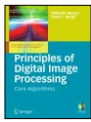
10. Odločanje na podlagi slike



Robotski vid



Študijska literatura

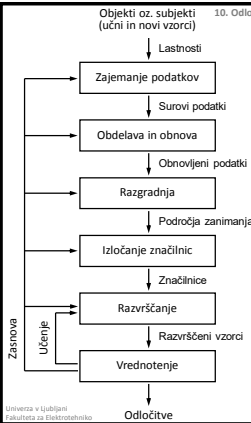


Principles of Digital Image Processing: Core Algorithms
Wilhelm Burger in Mark J. Burge, Springer, 2009
2 Regions in binary images

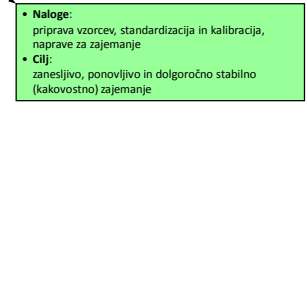


Concise Computer Vision: An Introduction to Theory and Algorithms
Reinhard Klette, Springer, 2014
3 Image analysis
5 Image segmentation
10 Object detection

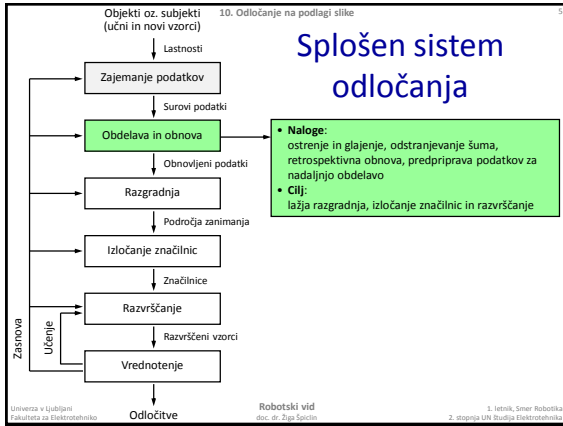
Splošen sistem odločanja

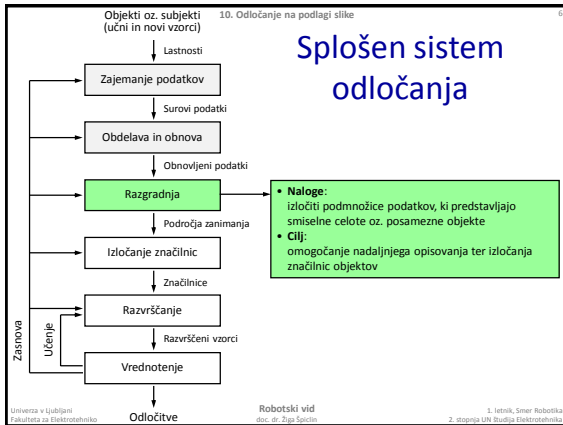


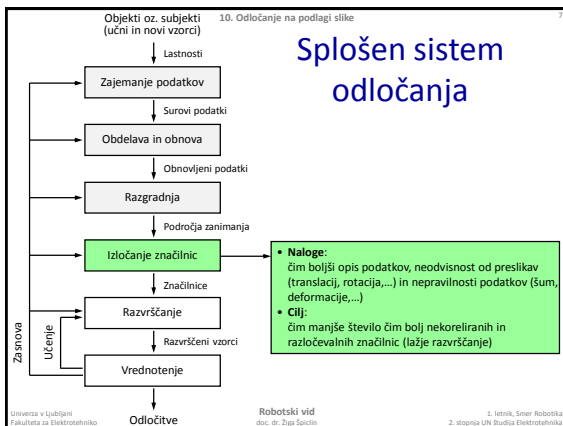
Splošen sistem odločanja



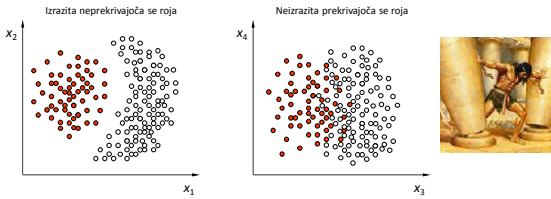
- **Naloga:** priprava vzorcev, standardizacija in kalibracija, naprave za zajemanje
- **Cilj:** zanesljivo, ponovljivo in dolgoročno stabilno (kakovostno) zajemanje



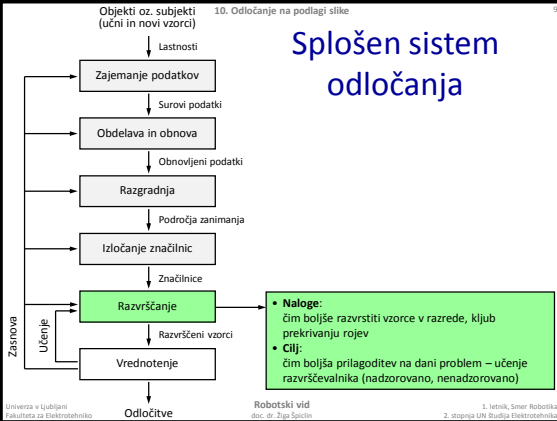




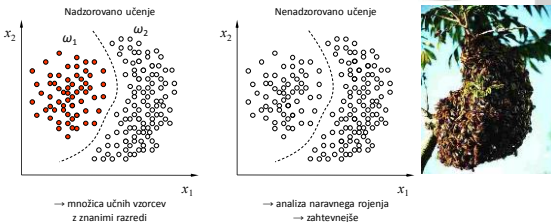
Razločevalnost značilnic

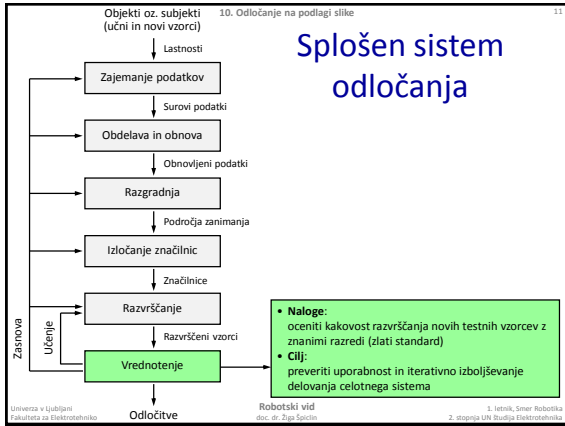


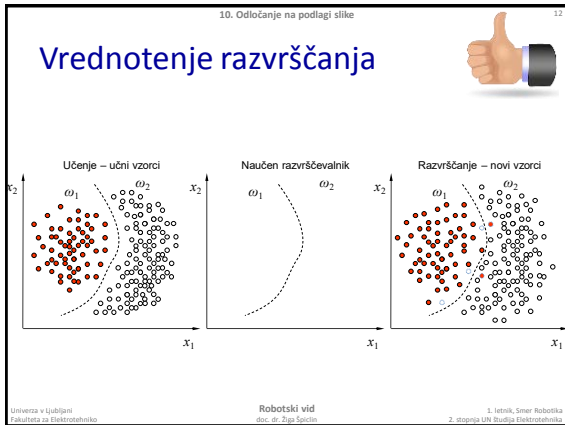
Splošen sistem odločanja

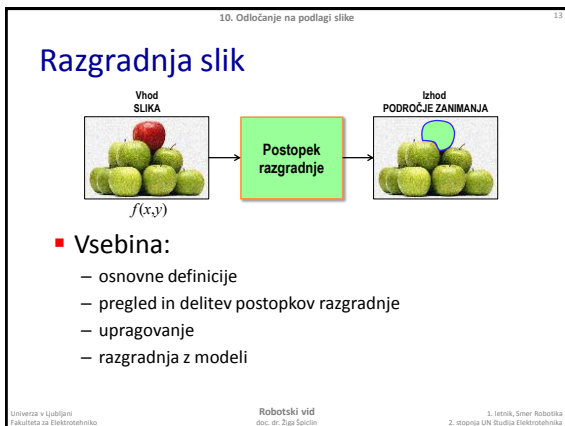


Učenje razvrščevalnika









Razgradnja slik



- Razgradnja ali segmentacija slik (*segmentation*):
 - razdeljevanje slik na osnovna področja oz. objekte
- Podrobnost razgradnje:
 - odvisna od namena uporabe
 - omejena z vsebino in kakovostjo slik
 - ločljivost, ostrina, sum, kontrast, sivinske nehomogenosti, ...
- Razgradnja je (najbolj?) zahteven postopek obdelave slik
- Zanesljivost in natančnost se dviguje zaradi:
 - razvoja slikovnih tehnik
 - dvigovanje kakovosti slik
 - razvoja sodobnih postopkov razgradnje
 - omogočajo uporabo predznanja o slikanih objektih

Razgradnja – definicije

- Bodí R domena slike, ki jo razgradimo na n področij; R_1, R_2, \dots, R_n , tako da zadostimo naslednjim petim pogojem:

1. Razgradnja mora biti popolna → vsak slikovni element mora biti razvrščen v področje:

$$R_1 \cup R_2 \cup \dots \cup R_n = R$$

2. Področja morajo biti medsebojno nezdržljiva:

$$R_i \cap R_j = \emptyset \text{ za vsak } i \text{ in } j; i \neq j$$

3. Slikovni elementi v področju morajo biti povezani, npr. z robovi ali z vogali:

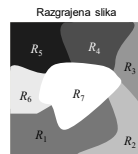
$$R_i \text{ je povezana množica slikovnih elementov; } i = 1, 2, \dots, n$$

4. Vse vrednosti v področju morajo biti enake ali pa dovolj podobne:

$$Q(R_i) = 1 \text{ za vsak } i = 1, 2, \dots, n$$

5. Sosednja področja morajo biti različna v smislu logične izjave Q :

$$Q(R_i \cup R_j) = 0 \text{ za vsaki dve sosednji področji } R_i \text{ in } R_j$$



Razgradnja – osnovni pristopi



- **Upragovanje**
 - velike spremembe sivin med področji
- **Razgradnja z modeli**
 - sliko geometrijsko poravnamo z nekim modelom, ki predstavlja področja slike
- **Razgradnja z razvrščanjem**
 - področja so glede na določen kriterij oz. značilnico dovolj homogena (sivine, std, tekstura, ...)

10. Odločanje na podlagi slike 20

Histogram in Upragovanje

VPLIV SPREMENB KONTRASTA IN SVETLOSTI

Kontrast ↑
↓

Svetlost ← →

Univerza v Ljubljani
Fakulteta za Elektrotehniko
Robotski vid
doc. dr. Žiga Srečan

1. letnik, Smerni Robotika
3. stopnja UNI Študija Elektrotehnika

10. Odločanje na podlagi slike 21

Histogram in Upragovanje

VPLIV ŠUMA

Slika →

↑
↓

Šum

Histogram →

Univerza v Ljubljani
Fakulteta za Elektrotehniko
Robotski vid
doc. dr. Žiga Srečan

1. letnik, Smerni Robotika
3. stopnja UNI Študija Elektrotehnika

10. Odločanje na podlagi slike 22

Histogram in Upragovanje

VPLIV SIVINSKIH NEHOMOGENOSTI

Slika a Polje b $c = a * b$

↑
↓

Histogram →

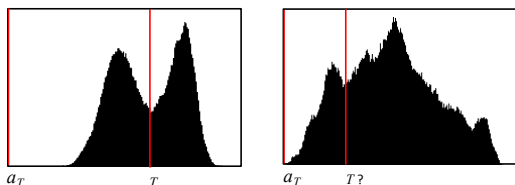
→ siviška kalibracija s slikanjem homogene scene (npr. bela podlaga brez objektov)

Univerza v Ljubljani
Fakulteta za Elektrotehniko
Robotski vid
doc. dr. Žiga Srečan

1. letnik, Smerni Robotika
3. stopnja UNI Študija Elektrotehnika

Upragovanje

- Kje je optimalni prag?



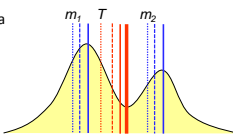
- obstaja več avtomatskih postopkov iskanja optimalnega praga
- npr. z minimizacijo variance med razgrajenima deloma histograma (Otsu postopek)

Iterativno določanje optimalnega praga

- Začetna ocena praga T , nato ponavljamo postopek:

1. Sliko razgradimo s pragom $T \rightarrow$ dobimo dve področji:
 - področje G_1 s svinami $> T$
 - področje G_2 s svinami $\leq T$
2. Izračunamo povprečni svinini obeh področij: m_1 in m_2
3. Izračunamo novo vrednost praga $T = (m_1 + m_2) / 2$
4. Ponavljamo korake od 1 do 3 dokler sprememba praga med zaporednimi iteracijami ne pade pod določeno stopnjo ΔT

- ΔT določa natančnost in hitrost postopka
- Postopek deluje dobro, če je med rojema v histogramu dovolj velika dolina
- Začetno vrednost praga T lahko postavimo kar na povprečno vrednost celotne slike oz. kamorkoli med oba roja



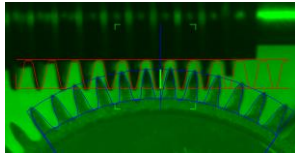
Razgradnja z modeli



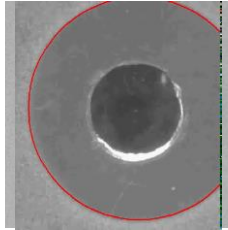
- Modeliranje objektov zanimanja:
 - opis predznanj o pričakovani obliki
- Matematični modeli:
 - geometrijski liki: premice, krivulje, kroge, elipse, ploskve, krogle, elipsoide, valje, cevi,...
- Fizikalni modeli:
 - upoštevajo tudi fizikalne lastnosti, denimo elastične lastnosti materiala
- Statistični modeli:
 - zgradimo jih na podlagi analize variabilnosti oblik večjega števila reprezentativnih objektov, npr. s statistično analizo velikosti oz. dolžine objekta zanimanja

Poravnava modela na sliko

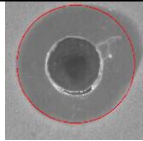
Merjenje ujema zobnikov



Merjenje premera con



Razgradnja z modeli



Model

- Opisuje attribute objektov zanimanja
- Naprimer obliko in izgled (pojavnost)

Prileganje

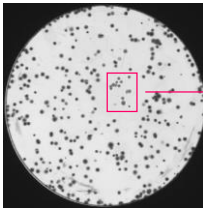
- **Zunanja energija:** iščemo pripadajoče pare točk, prekrivajočo slikovno informacijo
- **Notranja energija:** regulariziramo s prostostnimi stopnjami, možnimi deformacijami modela
- **Optimiziramo** mero kakovosti prileganja (ang. quality-of-fit; QoF)

Razgradnja

- Model z optimalnimi parametri določa razgradnjo oblike in/ali slike
- Poleg razgradnje tudi detektiramo objekte v sliki

Poravnava modela na sliko

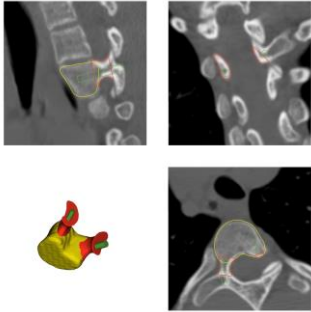
Štetje celičnih kolonij



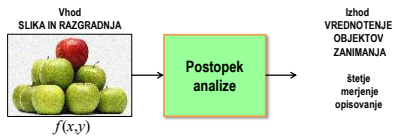
Štetje kovancev



Prileganje modela vretenca, pedikla in pedikularnega vijaka na CT sliko



Analiza objektov



■ Vsebina:

- označevanje objektov
- predstavitev objektov (**izločanje značilnic**)
- opisovanje mej in področij
- opisovanje z osnovnimi komponentami

Analiza objektov

- Vsakemu objektu priredimo svojo oznako
- Predstavitve so zgoščeni zapisi objektov:
 - omogočajo lažje **opisovanje** = **določanje značilnic**
- Predstavitve in opisi:
 - **zunanje lastnosti** (meje objektov)
 - **opisovanje oblike**
npr. obseg, usmeritev, največja dolžina, gladkost meje,...
 - **notranje lastnosti** (slikovni elementi objektov)
 - **opisovanje pojavnosti**
npr. homogenost, barva, tekstura, itn.
 - pogosto se uporabljata oba načina opisovanja:
 - komplementaren opis vseh lastnosti



Označevanje objektov

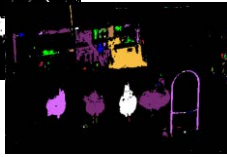
Sivinska slika z objekti



Upragovljena binarna slika

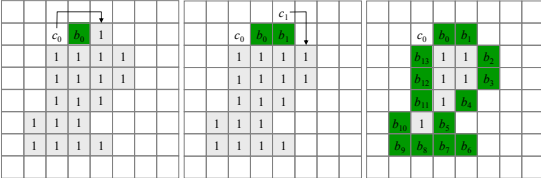


Barvne oznake



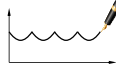
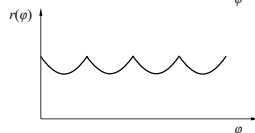
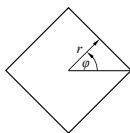
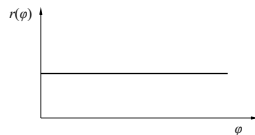
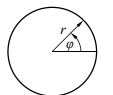
Sledenje mej

- Zapis zunanjih mej binarnih objektov z zaporedjem mejnih točk (v smeri urinega kazalca)
- Dobimo bolj zgoščen zapis, ki poenostavi opisovanje oblike



Podpisi mej

- Predstavitve 2D mej objekta z 1D funkcijami:
 - razdaljo med središčem in mejo zapišemo kot funkcijo kota





Podpisi mej

- Neodvisnost zapisa od translacije objekta

- Odvisnost od rotacije!

Rešitev?

- določimo isto začetno točko, ne glede na rotacijo

- izberemo najbolj oddaljeno točko od središča objekta

→ lahko jih je več!

- izberemo začetno točko, ki ustreza glavni osi objekta

→ nedoločenost $\pm 180^\circ$

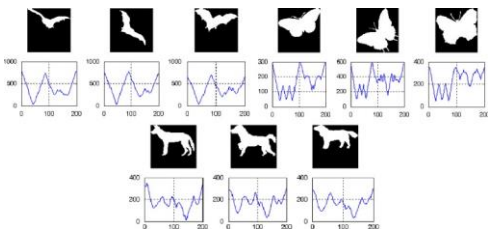
- Odvisnost od skaliranja!

Rešitev?

- funkcijo podpisa delimo z njegovim povprečjem ali standardno deviacijo



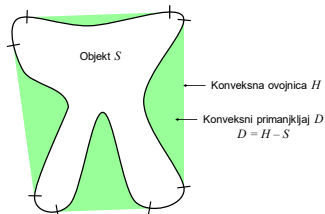
Podpisi mej





Odseki mej

- Učinkovito za izrazito konkavne meje objekta
→ pomembna informacija o obliki
- Mejo objekta razdelimo na odseke glede na njen prehod iz ali v konveksni primanjkljaj

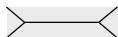


Odseki mej

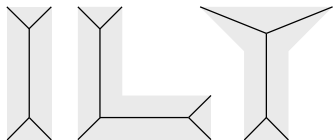


- Neodvisnost od velikosti in orientacije objekta
- Pred razdeljevanjem mej te navadno zgladimo:
 - tako se izognemo manjšim nepravilnostim
 - kako lahko zgladimo mejo oz. krivuljo?
- Predstavitev po odsekih omogoča opisovanje oblike objektov, npr.:
 - s površino objekta in konveksnega primanjkljaja
 - s številom odsekov v konveksnem primanjkljaju
 - z relativnimi položaji odsekov
 - drugi kvantitativni in statistični opisi odsekov

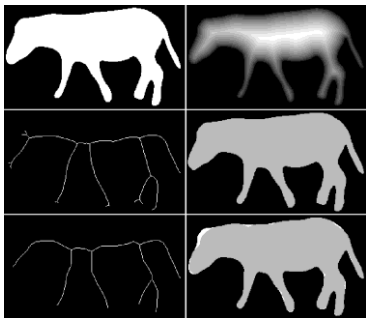
Skeleti



- Predstavitve strukturiranih objektov tako, da jih stanjšamo ("shujšamo") → dobimo samo skelete
- Skelet objekta določimo s transformacijo srednje osi (*medial axis transformation – MAT*)
- Točke skeleta so enako oddaljene od najmanj dveh točk meje



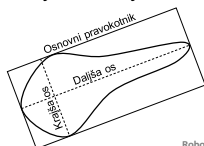
Skeleti



Osnovni opisi mej



- Dolžina meje = št. točk vzdolž meje:
 - verižna koda: št. elem. x in $y + \sqrt{2} \times$ št. diag. elem.
- Daljša os = največja razdalja med točkami
- Krajša os = pravokotna na daljšo os
- Osnovni pravokotnik = popolnoma objema mejo
- Ekscentričnost = razmerje obeh osi
- Ukrivljenost meje



Osnovni opisi področij



- Površina A → določena s številom vseh slikovnih elementov področja
- Obseg O → enak obsegu njegove meje
- Zgoščenost Z → opis oblike področja
 - neodvisen od velikosti, rotacije in translacije področja
 - razmerje površine in kvadrata obsega:

$$Z = \frac{A}{O^2}$$

- Krožno razmerje K_r → alternativen opis zgoščenosti
 - razmerje med dejansko površino področja A in površino kroga z enakim obsegom O kot področje
 - površina kroga z obsegom O je $O^2/4\pi$, torej je krožno razmerje:

$$K_r = 4\pi \frac{A}{O^2}$$

- Poleg opisov binarnih področjih, se uporabljajo tudi osnovne statistike sivinskih vrednosti področij:
 - povprečje, mediana, maks. in min. vrednost ter ostali percentili

Topološki opisi področij



- Globalno opisovanje področij
- Topologija odraža lastnosti podobe, ki se kljub njenim elastičnim deformacijam ohranjajo (brez trganja in združevanja)
- Najosnovnejši opis topologije je Eulerjevo število
 E = razlika med številom povezanih komponent C in številom lukenj H na sliki:

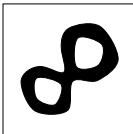
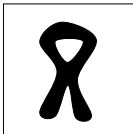
$$E = C - H$$

$$E = 1 - 0 = 1$$

$$E = 1 - 1 = 0$$

$$E = 1 - 2 = -1$$

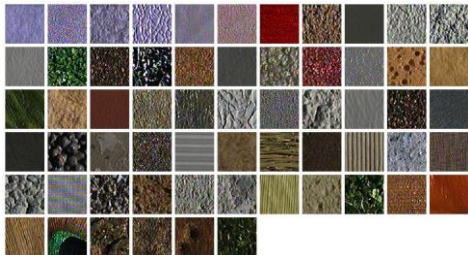
$$E = 2 - 4 = -2$$



Opisi področij s teksturo



- Pomembni kvantitativni opisi področij, brez formalne definicije, nanašajo pa se na lastnosti kot npr. gladkost, urejenost, usmerjenost, strukturiranost oz. znatost, itn.



Opisi področij s teksturo



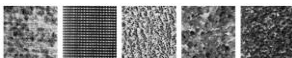
- Najbolj preprosto teksture opišemo s statističnimi momenti sivinskih vrednosti $\rightarrow p(z_i)$ = ocena verjetnostne porazdelitve

$$\mu_r(z) = \sum_{i=0}^{L-1} (z_i - m)^r p(z_i) \quad m = \sum_{i=0}^{L-1} z_i p(z_i)$$

- Moment $\mu_0=1, \mu_1=0, \mu_2=\sigma^2$ oz. varianca
 \rightarrow varianco lahko uporabimo za opis gladkosti področja G :

$$G = 1 - \frac{1}{1 + \sigma^2}$$

- Tretji moment μ_3 je merilo nesimetričnosti histograma
- Četrty moment μ_4 merilo njegove relativne ploskosti
- Momente višjih stopenj je težje povezati z obliko histograma



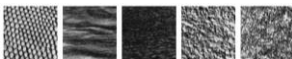
Opisi področij s teksturo



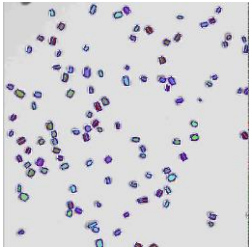
- Poleg momentov sta pomembni statistični značilnici teksture tudi homogenost H in entropija E , ki sta definirani kot:

$$H = \sum_{i=0}^{L-1} p^2(z_i) \quad E = - \sum_{i=0}^{L-1} p(z_i) \log p(z_i)$$

- Homogenost H ima največjo vrednost 1, ko je področje popolnoma homogeno, takrat pa ima entropija E najmanjšo vrednost 0
- Ostali načini opisovanja tekstur:
 - \rightarrow matrice so-pojavnosti sivinskih vrednosti področij
 - \rightarrow opisi strukture oz. sestave področij (*structural descriptors*)
 - \rightarrow opisi spektralnih oz. frekvenčnih lastnosti področij
npr. s Fourierovo transformacijo sivinskih vrednosti ter analizo spektra

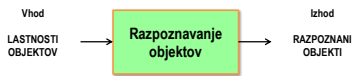


Primer: označevanje in opisovanje objektov



→ štetje in analiza celičnih kolonij

Razpoznavanje objektov



- Razpoznavanje objektov oz. vzorcev:
 - Osnovne definicije
 - Primer razpoznavanja za vizualno kontrolo kakovosti
 - Vrste razvrščevalnikov
- Vrednotenje binarnega razpoznavanja:
 - Predstavitev problema in zlati standard
 - Mere sposobnosti in ROC krivulja

Razpoznavanje vzorcev

- Računalniško podprto razpoznavanje:
 - razvoj umetnih sistemov za zanesljivo razpoznavanje vzorcev v različne namene:
 - razpoznavanja govora, prstnih odtisov, znakov, obrazov, dednih zapisov, zdravstvenega stanja, ...



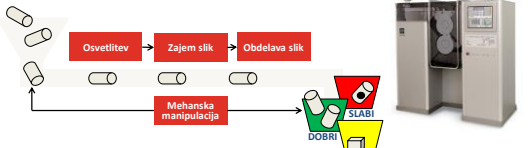
Primer: avtomatska vizualna kontrola kakovosti farmacevtskih izdelkov

- Bi zaužili tako tableto / kapsulo? Tveganja?



Primer: avtomatska vizualna kontrola kakovosti farmacevtskih izdelkov

- Koncept robotskega vida



- Obdelava slik

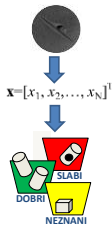


Definicije

- Vzorec je vektor x značilnic x_1, x_2, \dots, x_N opazovanega objekta:

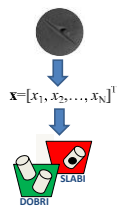
$$x = [x_1, x_2, \dots, x_N]^T$$

- Vsak vzorec želimo razvrstiti v nek razred vzorcev → podmnožico vzorcev s podobnimi lastnostmi
- Problem razpoznavanja vzorcev je problem razdeljevanja populacije vzorcev na podmnožice oz. razrede: $\omega_1, \omega_2, \dots, \omega_K$
- Razpoznavanje se prevede na razvrščanje vzorcev:
 - vsak vzorec x_i razvrstimo v pripadajoči razred; $x_i \rightarrow \omega_k$
 - in sicer glede na značilnice oziroma lastnosti opazovanega objekta

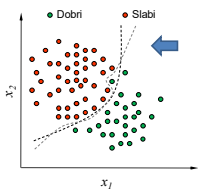


Kaj počne razvrščevalnik?

- Na podlagi porazdelitev značilnic določi mejo
- Razvrsti vzorec x glede na mejo



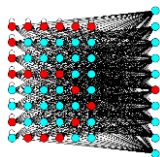
$$\mathbf{x} = [x_1, x_2, \dots, x_N]^T$$



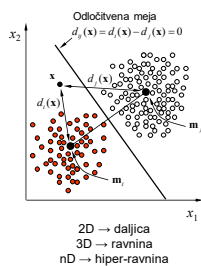
- Razločevalnost značilnic določa kompleksnost razvrščanja
→ vpliva na kakovost
- Kakovost razvrščanja zvisi tudi od algoritma razvrščanja, ki določi mejo med značilnicami

Vrste razvrščevalnikov

- Odločitvene funkcije:
 - linearne ali nelinearne odločitvene meje
- Razvrščanje glede na minimalno razdaljo:
 - razvrščanje glede na minimalno razdaljo, glede na najbližjega sosedo, iskanje rojev: *c-Means*, *ISODATA*,...
- Statistično razvrščanje:
 - Bayesov razvrščevalnik,...
- Mehko razvrščanje:
 - fuzzy c-Means*,...
- Neuronske mreže:
 - perceptron, backpropagation
- Odločitvena drevesa
 - CART, random forests,...
- ...



Odločitvene funkcije



Vzorec razvrstimo v razred z najmanjšo Evklidsko razdaljo D do središča roja

$$\mathbf{m}_i = \frac{1}{N_i} \sum_{\mathbf{x} \in \omega_i} \mathbf{x}; \quad i = 1, 2, \dots, K$$

$$D_i(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_i\| \quad \|\mathbf{a}\| = (\mathbf{a}^T \mathbf{a})^{1/2}$$

$$D_i^2(\mathbf{x}) = \mathbf{x}^T \mathbf{x} - 2\mathbf{x}^T \mathbf{m}_i + \mathbf{m}_i^T \mathbf{m}_i$$

Kar je ekvivalentno največji odločitveni funkciji:

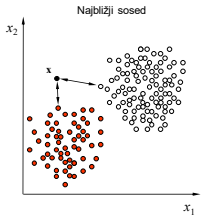
$$d_i(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_i - \frac{1}{2} \mathbf{m}_i^T \mathbf{m}_i; \quad i = 1, 2, \dots, K$$

ki da odločitveno mejo:

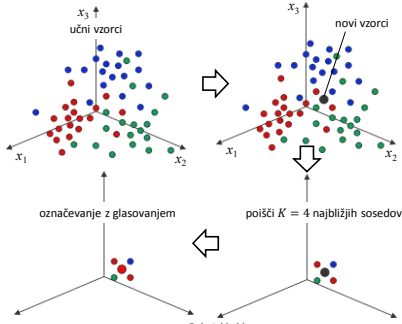
$$d_0(\mathbf{x}) = d_1(\mathbf{x}) - d_2(\mathbf{x}) = 0$$

$$d_0(\mathbf{x}) = \mathbf{x}^T (\mathbf{m}_1 - \mathbf{m}_2) - \frac{1}{2} (\mathbf{m}_1 - \mathbf{m}_2)^T (\mathbf{m}_1 + \mathbf{m}_2) = 0$$

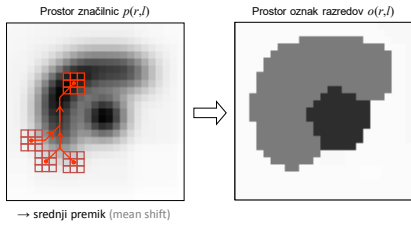
Minimalna razdalja



Razvrščanje s K najbližjih sosedov (KNN)

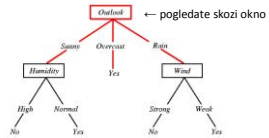


Neparametrično razvrščanje



Odločitvena drevesa

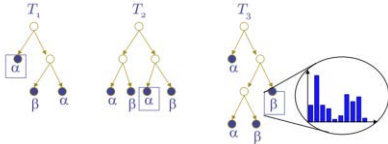
- Kako se odločite ali greste v soboto zjutraj na tek?



- Uporabna za različne tipe vhodnih vzorcev
 - zvezne in diskretne vrednosti (tudi take brez vrstnega reda)
- Enostavna interpretacija: če ... sicer pravila (ang. if-else rules)
- Odpornost na manjkajoče podatke (značilnice)

Naključni gozdovi

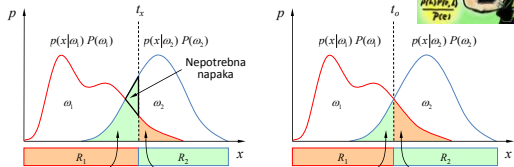
- Nauči večje število L naključnih dreves T



- Preko vseh dreves $l = 1, \dots, L$ in vseh listov η akumuliraj posteriori verjetnosti za namen razvrščanja

$$P(C|\mathbf{x}) = \frac{1}{L} \sum_{l=1}^L p_{l,\eta}(C|\mathbf{x})$$

Bayesov razvrščevalnik



$$P(E) = \int_{R_1} p(x|\omega_2) P(\omega_2) dx + \int_{R_2} p(x|\omega_1) P(\omega_1) dx$$

$$P(E) = \int_{R_1} p(x|\omega_2) P(\omega_2) dx + \int_{R_2} p(x|\omega_1) P(\omega_1) dx$$

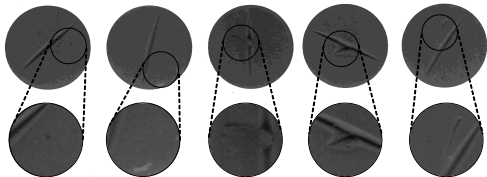
$x \rightarrow \omega_1$; če je $p(x|\omega_1)P(\omega_1) > p(x|\omega_2)P(\omega_2)$
 $x \rightarrow \omega_2$; če je $p(x|\omega_2)P(\omega_2) > p(x|\omega_1)P(\omega_1)$

→ optimalen v smislu najmanjše napake razvrščanja

Razvrščanje v več kategorij?

- Kombiniramo binarne razvrščevalnike ali
- učimo večkategorični razvrščevalnik (zavisi od algoritma)

D_1 – Dot D_2 – Spot D_3 – Emboss D_4 – Deboss D_5 – Crack



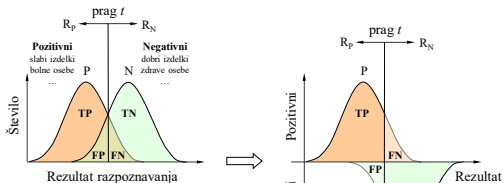
Vrednotenje razvrščanja



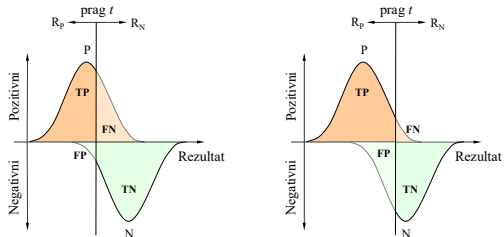
- Ocena pravilnosti razvrščanja
- Ocena **tveganja** (cena nepravilnega razvrščanja):
 - odvisna od pravilnosti razvrščanja in področja uporabe
 - odločitve imajo lahko resne posledice
- Npr. tablete s kritično napako ne smemo razpoznati kot dobre kakovosti
 - povečati moramo občutljivost na to napako:
 - odločitveno mejo premaknemo proti roju dobrih tablet
 - več dobrih tablet bomo razpoznali kot slabe:
 - izpad proizvodnje, več odpada za uničenje → večji stroški, ugled proizvajalca?
- Pomen objektivnega vrednotenja postopkov razpoznavanja oz. sprejemanja odločitev!



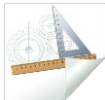
Vrednotenje razpoznavanja



Vpliv vrednosti praga

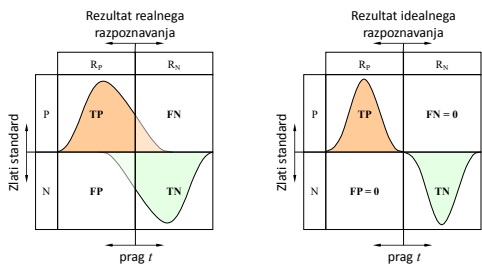


Kontingenčna tabela



		Rezultat pozitiven R_p		Rezultat negativen R_N			
Pozitivni	P	TP	FN	$P = TP + FN$ $N = FP + TN$		$R_p = TP + FP$ $R_N = FN + TN$	
	Negativni	N	FP				

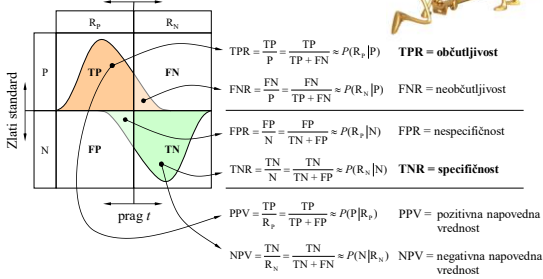
Kakovost razpoznavanja



Mere sposobnosti

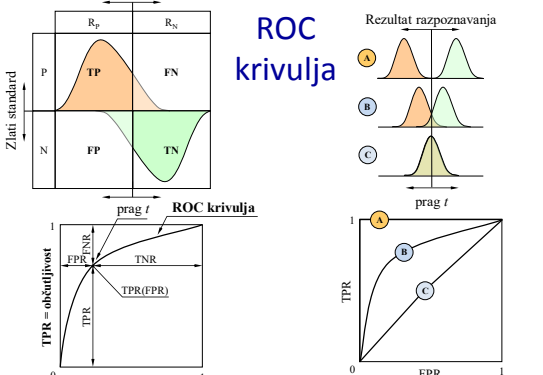


Rezultat razpoznavanja



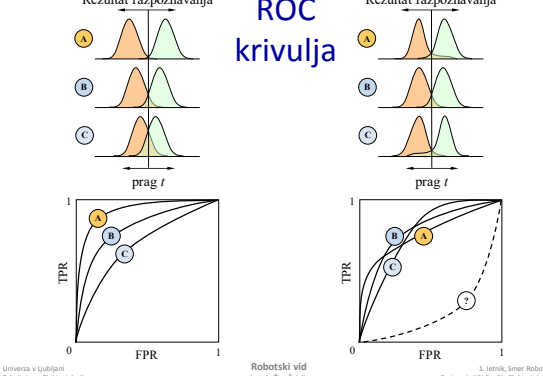
ROC krivulja

Rezultat razpoznavanja



ROC krivulja

Rezultat razpoznavanja

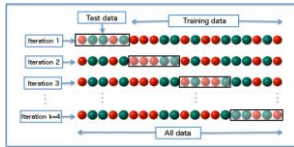


Izvedba validacije



- Ločevanje učnih in testnih slik
 - na učnih nastavimo optimalne parametre ali izvedemo učenje nadzorovanih postopkov razvrščanja
 - na testnih izračunamo vrednosti mer kakovosti
- k-koračna križna validacija (k-fold cross-validation)

1. razdelimo množico slik v k skupin
2. učimo na $k-1$ skupinah, testiramo na 1 skupini
3. ponovimo tako, da je vsaka skupina vsaj enkrat testna
4. izračunamo mere vrednotenja za vseh k eksperimentov in povprečimo



Povzetek

- Splošen sistem odločanja na podlagi slike sestoji iz
 - zajema, obdelave in obnove slik, razgradnje, analize objektov in
 - izločanja ter razvrščanja značilnic (po potrebi tudi vrednotenje)
- Z razgradnjo izločimo objekte zanimanja
 - osnovni principi so upragovanje, razgradnja z modeli, razvrščanjem, itd.
- Analiza izloči značilnice objektov
 - vsak objekt je opisan z vzorcem, ki predstavlja nabor značilnic tega objekta
 - značilnice kodirajo notranje in zunanje lastnosti objektov
- Algoritem razvrščanja iz porazdelitve značilnic določi mejo
 - meja je lahko linearna / nelinearna, zavisi od konkretnega algoritma
 - razvrščamo z eksplicitnimi odločitvenimi funkcijami, razdaljo do učnih vzorcev (KNN), odločitvenimi drevesi in gozdovi (hierarhično razvrščanje), Bayesov razvrščevalnikom,...
- Z vrednotenjem določimo lastnosti sistema odločanja, kot npr.
 - občutljivost, specifičnost, ROC krivuljo → kakovost, zanesljivost, uporabnost odločitev