Signaler, information & bilder, föreläsning 16

Michael Felsberg och Maria Magnusson
Computer Vision Laboratory (Datorseende)
Department of Electrical Engineering (ISY)
michael.felsberg@liu.se,
maria.magnusson@liu.se

LINKÖPINGS UNIVERSITET

Översikt

- Mer om Histogram och tröskelsättning
  - Histogramutjämning
  - Lokal tröskelsättning
  - Tröskelsättning med hysteres
- Medianfilter
- Segmentering och etikettering
  - Watershed algoritmen
- Korrelation (2D)
  - Vanlig
  - Normerad
  - Utan DC-nivå
- Vidareledande kurser och profil
- Aktuell forskning
- Teori: Kap. 2.2, 5.1, 5.2, 6.1, 6.2, 6.5, 7
- Bygger på Maria Magnussons föreläsningar

LINKÖPINGS UNIVERSITET

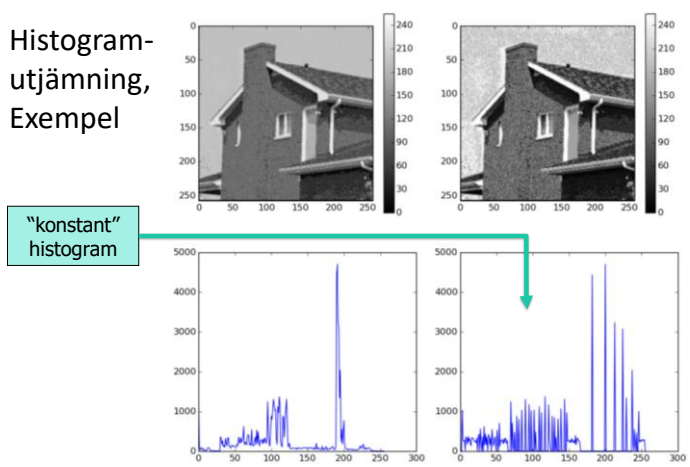
Histogramutjämning

- Om varje pixels gråvärde transformeras genom histogrammets integral (diskret: kumulativa summa), blir den nya bildens histogram konstant (!) (I diskreta fallet blir det dock inte perfekt konstant.)
- Phyton-kod:

```
P = p.cumsum()                  # p ej normaliserat histogram
P = 255*P / P[-1]               # skala om till rätt interval
f2 = np.interp(f.flatten(),np.arange(256),P)
                                # beräkna nya gråvärden
f2 = f2.reshape(f.shape)        # gör om till 2D bild
```

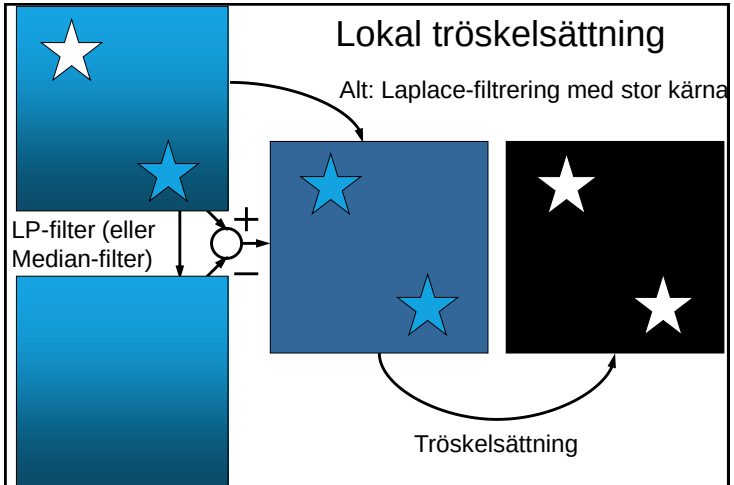
LINKÖPINGS UNIVERSITET

Histogram-utjämning, Exempel



LINKÖPINGS UNIVERSITET

Lokal tröskelsättning



Lokal tröskelsättning

Fig. 5.8

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

−

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

/25

=

-1	-1	-1	-1	-1
-1	-1	-1	-1	-1
-1	-1	24	-1	-1
-1	-1	-1	-1	-1
-1	-1	-1	-1	-1

/25

δ(x,y): Diskret centrerad dirac: Ger ingen effekt

M(x,y): Ska LP-filtrera och sudda ut

L(x,y): stor Laplace-kärna

$$f(x,y) * \delta(x,y) - f(x,y) * M(x,y) = f(x,y) * L(x,y)$$

OBS! I verkligheten behövs ännu större kärnor

Tröskelsättning med hysteres

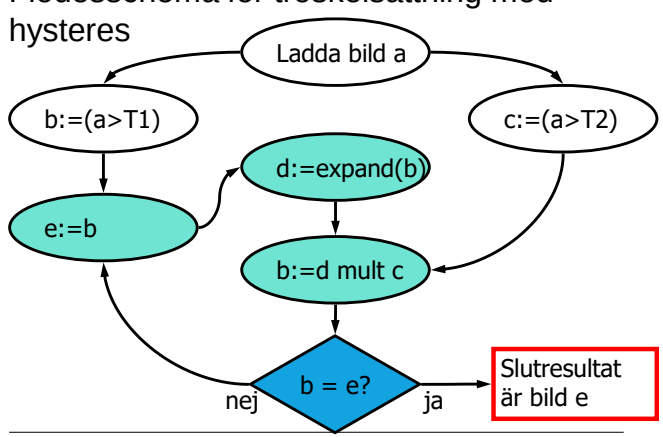
- Tröskel T1 tar bort delar av objektet
- Tröskel T2 bevarar delar av bakgrunden

a: Gråskale-bild | b: Trösklad på T1 | c: Trösklad på T2 | d: Slut-resultat

Bild b expanderas med bivillkor att expan-sionen ej får ske utanför objekten i bild c.

≈*Fig. 5.10*

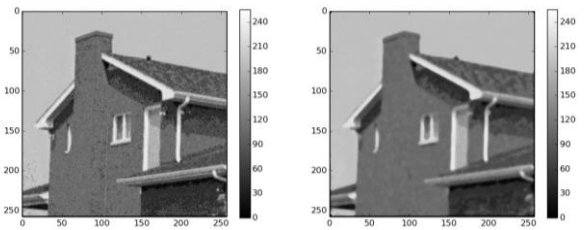
Flödesschema för tröskelsättning med hysteres



Medianfilter

- Medianen M är det gråvärde sådant att antalet gråvärden större än M är lika med antalet gråvärden mindre än M
- Ett medianfilter ersätter på arbetspunkten gråvärdet med medianen i filtermasken.

Im3 = signal.medfilt2d(Im,5)



LINKÖPINGS UNIVERSITET

Median-filtrering, teknik

- Median-filtret glider över in-bilden på liknande sätt som vid faltning.
- Värdena under filtret noteras. Exempel:

  1 1 9 2 3 3 3 3 3
- De sorteras till:

  1 1 2 3 3 3 3 3 9
- Median-värdet blir här 3. Det sätts i ut-bilden.

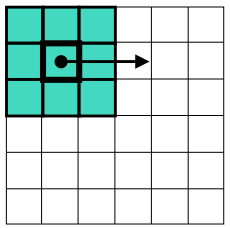


Bild med glidande median filter

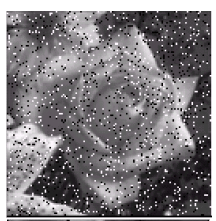
Sortering är beräkningstungt!

LINKÖPINGS UNIVERSITET

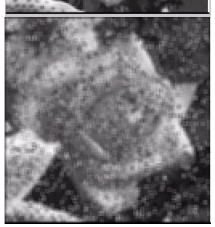
Jämförelse: Median- och LP-filter

Original-bild…

…med salt- och peppar brus



Efter LP-filtrering med 3x3-filter

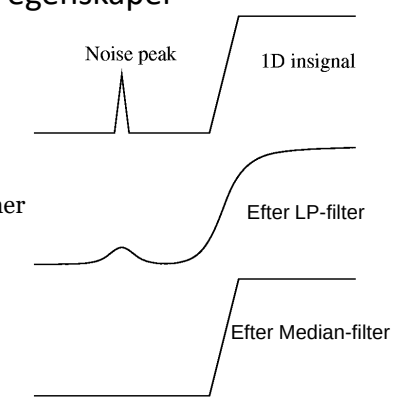


Efter Median-filtrering med 3x3-filter

LINKÖPINGS UNIVERSITET

Median-filtrering, egenskaper

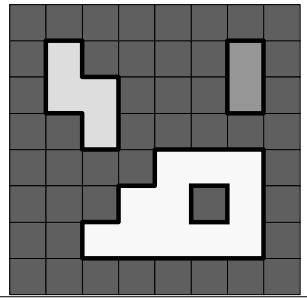
- Kanter bevaras
- Brus undertrycks
- Tunna linjer förstörs
- Jämna ytor uppkommer



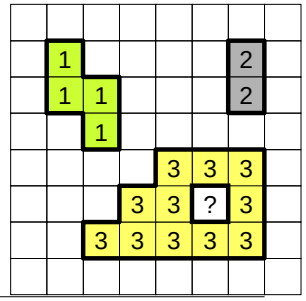
LINKÖPINGS UNIVERSITET

Segmentering skiljer ut och etiketterar sammanhängande objekt

Gråskalebild



Segmenterad och etiketterad bild



Exempel på Segmenteringsalgoritmer

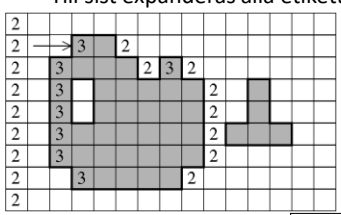
- Tröskelsättning
- Watershed, MSER (liknar watershed)
- mm
- Ett gemensamt slutsteg på en segmentering är ofta etikettering (labeling)

Exempel på Etikettering (labelling)

- Flood-fill (används ofta i datorgrafik)
- RB-algoritmen (Raster-scan Borderfollow)
- mm

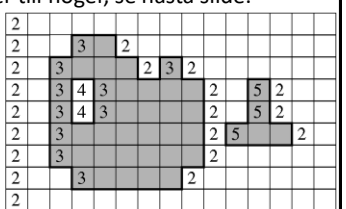
RB-algoritmen (Raster-scan Border-follow)

- Bilden rasterscannas vänster-höger uppifrån-nedåt.
- När scanningen passerar en kant (0->1 eller 1->0) avbryts skanningen och konturföljning sker.
- Under konturföljningen sätts etiketter till höger om kant.
- Till sist expanderas alla etiketter till höger, se nästa slide.



Efter kontur-följning av det första objektet.

Fig. 6.21

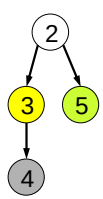


Efter konturföljning av båda objekten och ett hål.

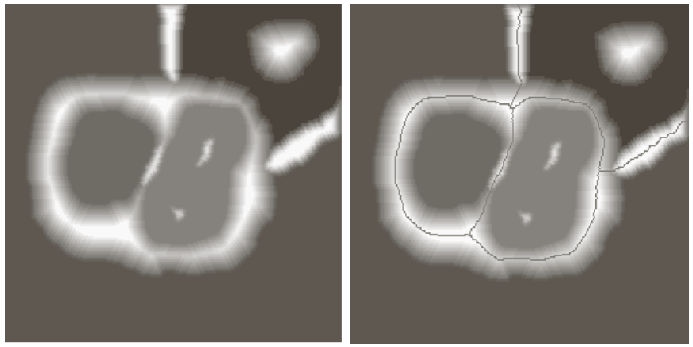
RB-algoritmen, slutresultat

2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	3	3	2	2	2	2	2	2	2	2	2	2
2	2	3	3	3	3	2	3	2	2	2	2	2	2	2
2	2	3	4	3	3	3	3	3	2	2	5	2	2	2
2	2	3	4	3	3	3	3	3	2	2	5	2	2	2
2	2	3	3	3	3	3	3	3	2	5	5	5	2	2
2	2	3	3	3	3	3	3	3	2	2	2	2	2	2
2	2	2	3	3	3	3	3	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Topo-logisk graf:



Watershed segmentering, introduktion



a) h)

LINKÖPINGS UNIVERSITET

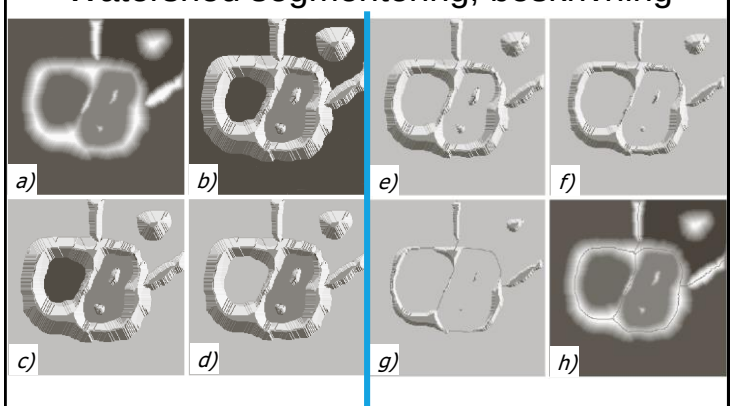
Gonzalez&Woods: Fig. 10.54

Watershed segmentation, introduction

- Fig. 10.54 a) The watershed algorithm regards a gray scale image as a landscape, where the lakes are preferably dark pixels and the ridges and mountains are preferably bright pixels.
- A watershed = a ridge between individual lakes and rivers.
- A catchments basin = a geographic area whose water flows to a certain lake.
- Fig. 10.54 h) The watershed algorithm can find the watersheds in the image. They form closed curves and the area inside each closed curve, the catchments basin, can be regarded as a segmented, individual object.

LINKÖPINGS UNIVERSITET

Watershed segmentering, beskrivning



LINKÖPINGS UNIVERSITET

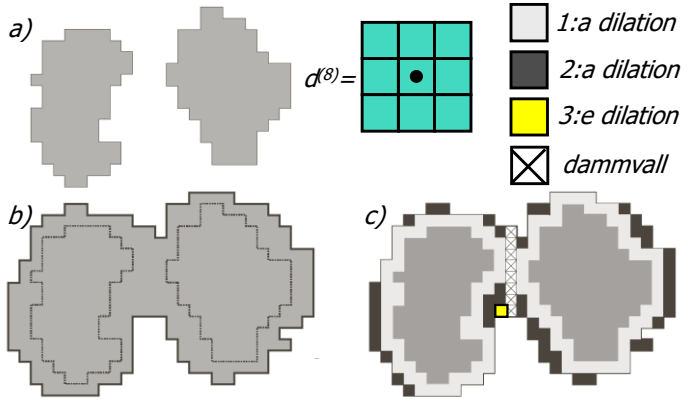
Gonzalez&Woods: Fig. 10.54

Watershed segmentation, description

- Fig. 10.54 b) is a topographic view of 10.54 a). It is intended to give a 3D-feeling.
- In the beginning, the landscape in a) and b) is totally dry. Local minima are detected and holes are punched in each of them.
- Then the entire landscape is flooded from below by letting water rise through the holes at a uniform rate.
- In c), there is water (light gray) in the background.
- In d), there is also water in the left lake.
- In e), there is water in both lakes.
- In f), there is the first sign on overflow between two catchments basins. Therefore, a short dam has been built between them.
- In g), there are rather long dams.
- Finally, the whole landscape is over flooded with water, separated by high dams, watersheds.
- In h), these watersheds are shown overlaid on the original image.

LINKÖPINGS UNIVERSITET

Watershed segmentering, dammkonstruktion



LINKÖPINGS UNIVERSITET

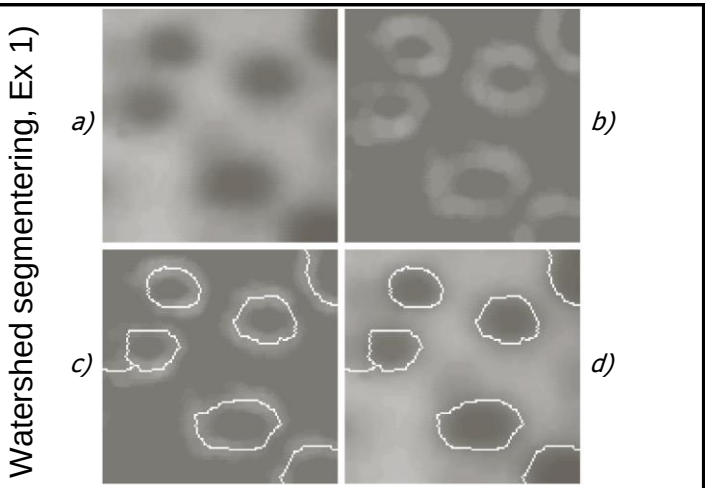
Gonzalez&Woods: Fig. 10.55

Watershed segmentation, dam construction

- Let $C_{n-1}(M_1)$ denote the pixels in lake M_1 at flooding step n-1. These are shown to the left in a).
- Let $C_{n-1}(M2)$ denote the pixels in lake M_2 at flooding step n-1. These are shown to the right in a).
- Flooding step n is shown in b).
- Let C(n-1) denote the union of $C_{n-1}(M_1)$ and $C_{n-1}(M_2)$. There are 2 objects in C(n-1) in a) and 1 object in C(n) in b). This indicates that something have happened in step n and that some action must be performed.
- Dilate the pixels in a) with the $d^{(8)}$-structuring element considering the following constraints:
  - Dilation is not allowed outside b).
  - Dilation is not allowed so that the pixels in M_1 and M_2 are united.
  - Instead, if the pixels in M_1 and M_2 are going to be united, denote: dam pixel.
- The first dilation in c) gives the light gray pixels.
- The second dilation in c) gives the dark gray pixels and the X-marked dam pixels.

LINKÖPINGS UNIVERSITET

Watershed segmentering, Ex 1)



LINKÖPINGS UNIVERSITET

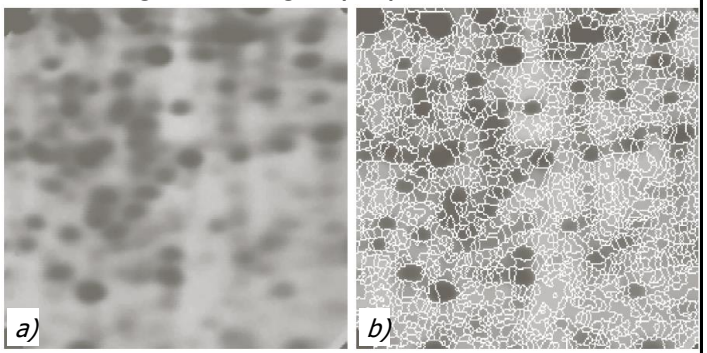
Gonzalez&Woods: Fig. 10.56

Watershed segmentation, Ex 1)

- a) Original image with dark objects that are going to be segmented.
- b) A usual pre-processing step is to calculate the magnitude of the gradient.
- c) The result of the watershed algorithm overlaid on the gradient image.
- d) The result of the watershed algorithm overlaid on the original image.

LINKÖPINGS UNIVERSITET

Watershed segmentering, Ex 2) översegmentering/superpixlar



LINKÖPINGS UNIVERSITET

Gonzalez&Woods: Fig. 10.57

2D Faltning och Korrelation

Faltning:

$$f * g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\cdot g(x-\alpha, y-\beta)\, d\alpha\, d\beta$$

Korrelation:

$$f \star g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\cdot g(x+\alpha, y+\beta)\, d\alpha\, d\beta$$

Ekv. 7.1

Korrelation är samma sak som faltning med vikt kärna (faltningsförfarandet utan vikning)

LINKÖPINGS UNIVERSITET

2D diskret Faltning och Korrelation

$$f * g(x,y) = \sum_{\alpha}\sum_{\beta} f(\alpha,\beta)\cdot g(x-\alpha, y-\beta)$$

$$f \star g(x,y) = \sum_{\alpha}\sum_{\beta} f(\alpha,\beta)\cdot g(x+\alpha, y+\beta)$$

Ekv. 7.2

0	-2	1
0	0	0
0	2	0

$f(x,y)$ ⋆

0	-2	1
0	0	0
0	2	0

$g(x,y)$ =

0	-4	2
0	0	0
-2	9	-2
0	0	0
2	-4	0

Notera symmetrin och att vi får max i origo!

0	4	-4	1
0	0	0	
0	-8	4	
0	0	0	
0	4	0	

Faltnings-resultat som jämförelse

LINKÖPINGS UNIVERSITET

Räknelagar för Faltning och Korrelation

Faltning kommuterar:

$$f * g(x,y) = g * f(x,y)$$

Korrelation kommuterar inte:

$$f \star g(x,y) = g \star f(-x,-y)$$

Faltning i Fourierdomänen:

$$\Im_2[f * g(x,y)] = F(u,v)\cdot G(u,v)$$

Korrelation i Fourierdomänen (f och g reella):

$$\Im_2[f \star g(x,y)] = F^*(u,v)\cdot G(u,v)$$

Bevisen är en trevlig övning!

LINKÖPINGS UNIVERSITET

Mönsterdetektering med korrelation

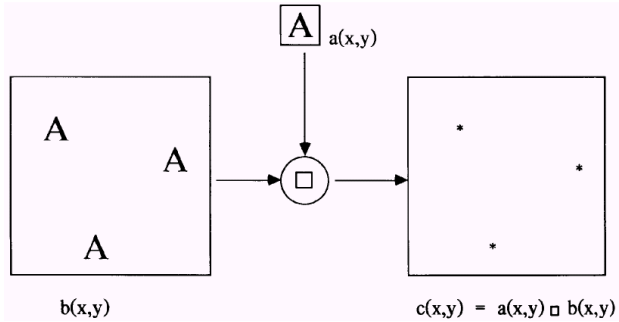
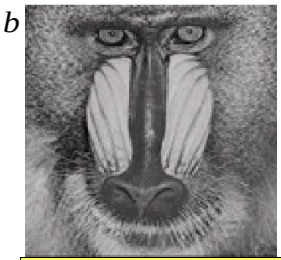


Fig. 7.1

Mönsterdetektering med vanlig korrelation

$$c(x,y)=\sum_{\alpha}\sum_{\beta}a(\alpha,\beta)\cdot b(\alpha+x,\beta+y)$$ *Ekv. 7.12*

b 

a 

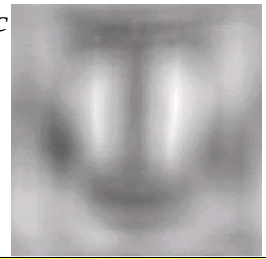
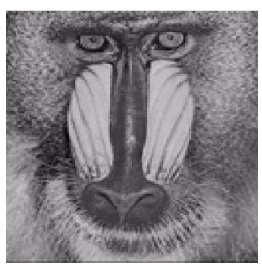
c 

Fig. 7.2

Hög positiv signalstyrka i bilden b(x,y) och positivt mönster kan ge högt korrelationsresultat även om mönstret och bilden ej överensstämmer.

Mönsterdetektering med normerad korrelation

$$a_n(x,y)=\frac{\sum_{\alpha,\beta}a(\alpha,\beta)\cdot b(\alpha+x,\beta+y)}{\sqrt{\sum_{\alpha,\beta}a^2(\alpha,\beta)\sum_{\alpha,\beta\in\Omega}b^2(\alpha+x,\beta+y)}}$$ *Ekv. 7.14*

b 

a 

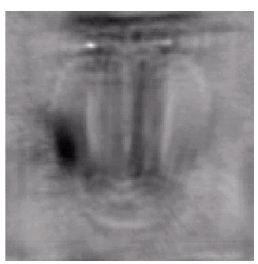
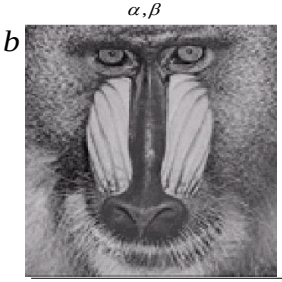
a_n 

Fig. 7.2

Mönsterdetektering med korrelation utan DC-nivå

$$d(x,y)=\sum_{\alpha,\beta}[a(\alpha,\beta)-\mu_a]\cdot[b(\alpha+x,\beta+y)-\mu_b]$$
$$=\sum_{\alpha,\beta}[a(\alpha,\beta)-\mu_a]\cdot[b(\alpha+x,\beta+y)]$$ *Ekv. 7.15*

b 

a 

d 

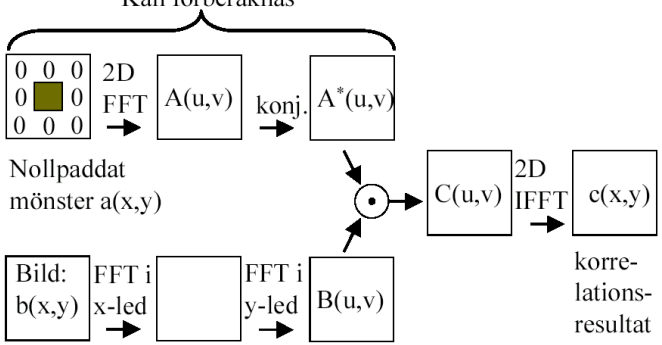
Fig. 7.2

Beräkning i Fourierdomänen

$$C(u,v) = A^*(u,v) \cdot B(u,v)$$

Ekv. 7.13

Kan förberäknas



LINKÖPINGS UNIVERSITET

Kovariansmatris

För den speciellt matematik-intresserade…

- Korrelation utan DC nivå blir kovarians
- Cyklisk kovarians av en 1D signal med sig själv motsvarar en symmetrisk, cyklisk matris (1)
- Egenvektorerna är kosinusfunktioner (2)
- Med hjälp av egenvektorerna *dekorreleras* signalen, d v s energin kan beräknas punktvis (3)
- (1)-(3) är relativt enkelt att bevisa
  - (1) stationäritet, variabelbyte
  - (2) fourierbasen som LTI egenvektorer + symmetrin
  - (3) ortogonalitet

LINKÖPINGS UNIVERSITET

Lab4: Bildkompression med JPEG

- JPEG är namnet på en (förstörande) bildkompressions-metod. Genom att använda denna teknik tar bilden mycket mindre plats att lagra.
- JPEG bygger på transformteknik. Man använder cosinus-transformen som är en nära släkting till fouriertransformen.
- Bilden delas upp i 8x8-rutor och varje ruta transformeras och kodas var för sig.
- Tanken är att de låga frekvenserna är viktigare än de höga.
  - Första tanken är att nollställa de högsta frekvenserna, särskilt i 8x8-rutor med liten variation.
  - Smartare är att kvantisera de höga frekvenserna hårdare.

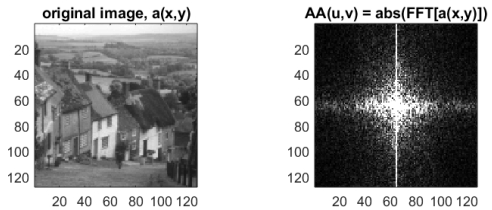
LINKÖPINGS UNIVERSITET

Bild, komprimerad med JPEG

Störningar syns från kodningen

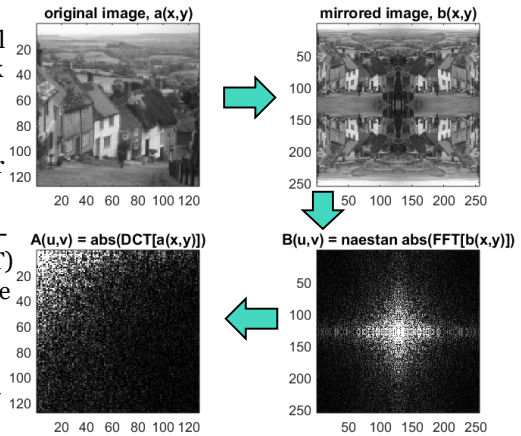
Diskret fouriertransform (DFT), beräknas med FFT

- DFT:n tycker att bilden sitter ihop i över- och underkant, samt i vänster- och högerkant.
  - Den stora skillnaden i över- och underkant ger upphov till den skarpa lodräta linjen i FFT:n.
  - vänster- och högerkant är mer lika och ger bara en svag horizontell kant.



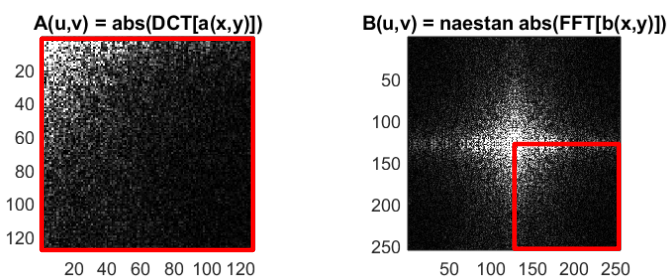
Spegling löser problemet!

- DFT:n (här B(u,v)) blir reell och symmetrisk (jämn).
- Den skarpa lodräta linjen är borta.
- Diskret cosinus-transform (DCT) motsvarar nedre högra delen av B(u,v).



Diskreta Cosinustransformen (DCT:n) och DFT:n

- DCT:n har de låga frekvenserna uppe till vänster.
- DFT:n har de låga frekvenserna i mitten.



Diskret cosinustransform (DCT), alternativ beskrivning

- Tvinga fram symmetri:

1	-4	4	0	0	4	-4	1

- 1D-DFT ger då 1D-DCT (origo flyttas med ½)
- För bildregioner, kan DCT beräknas rad- och kolumnvis:

$$F_C(k,l) = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} f_d(n,m)\cos\left[\frac{\pi k}{N}(n+\tfrac{1}{2})\right]\cos\left[\frac{\pi l}{M}(m+\tfrac{1}{2})\right]$$

- med korrekt normalisering blir DCT:n ortogonal
  - DC komponenten (i varje led) multipliceras med: $1/\sqrt{2}$
  - hela matrisen med basfunktioner mult. med: $2/\sqrt{MN}$
  - Detta behövs dock egentligen inte i JPEG kompression

PSNR: Globalt bildavstånd i gråvärde

- Peak Signal to Noise Ratio (PSNR)
- Klassisk mått på hur bra t ex avbrusning fungerar.
- Vi använder det för att mäta bildkvalitet efter JPEG-kompression.
- Bygger på kvadratiska medelfelet mellan två bilder, Mean Square Error (MSE)

$$\text{MSE}=\sqrt{\text{mean}\left((\text{bild1}-\text{bild2})^2\right)}$$

- Mäts i dB (decibel)

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{255^2}{\text{MSE}}\right)$$

LINKÖPINGS UNIVERSITET

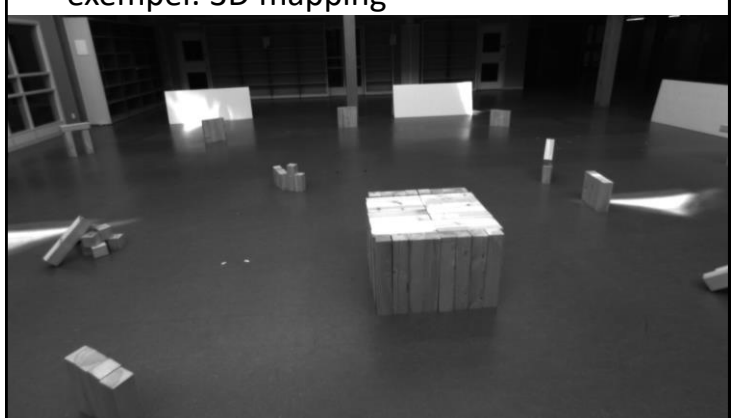
Bild-relaterade kurser: AI&ML profil

- TSBB06 Multidimensionell signalanalys
- TSBB08 Digital bildbehandling grundkurs
- TSBB09 Bildsensorer
- TBMI26 Neuronnät och lärande system
- TSBB11 Bilder och grafik, projektkurs, CDIO
- TSBB15 Datorseende
- TSBB17 Visuell detektion och igenkänning
- kurser inom bildkodning (TSBK): grafik, datorspel, m fl
- interna och externa exjobb

Deep Learning

LINKÖPINGS UNIVERSITET

exempel: 3D mapping



LINKÖPINGS UNIVERSITET

exempel: studentprojekt

LINKÖPINGS UNIVERSITET

exempel: a pose from a picture

LINKÖPINGS UNIVERSITET

exempel: studentprojekt

LINKÖPINGS UNIVERSITET

Alumni + Samarbete

- svenska företag
  - SAAB & Vricon
  - Spotscale
  - Autoliv & Zenuity
  - Scania
  - Termisk Systemteknik & Visage Technologies
- internationella företag
  - Daimler
  - SICK
  - Apple
  - Siemens

LINKÖPINGS UNIVERSITET

Ett lite äldre exempel på vår forskning: trafikskyltar och fourierdeskriptorer

- Bildens konturer erhålls från Watershed eller MSER algoritmerna.
- Konturen kan beskrivas som en komplex funktion: $c : [0, 1] \rightarrow \mathbb{C}, t \mapsto x(t) + jy(t)$
- 1D fourierserien C av denna funktion kallas fourierdeskriptor.
- Den har invariansegenskaper för:
  - position
  - orientering
  - storlek

LINKÖPINGS UNIVERSITET

Exempel: forskning

Aktuell forskning

- Computational imaging
  - rullande slutare
- Detection, tracking and recognition
  - 1:a, 2:a, och 3:e plats i VOT(-TIR) & OpenCV tävlingar
- 3D structure and pose estimation
  - under en tid 1:a plats i KITTI listan
- Robot vision, autonomous systems, AI
  - del av WASP (största svenska projektet i historien)

Exempel: forskning

Exempel: 3D struktur och rörliga objekt