

Bildbehandling i frekvensdomänen

Erik Vidholm
erik@cb.uu.se

9 december 2002

Sammanfattning

Detta arbete beskriver hur en bild kan tolkas som en tvådimensionell digital signal, hur denna signal Fouriertransformeras och hur transformen sedan kan visualiseras och tolkas. Filtrering av bilder i frekvensdomänen beskrivs, och resultat för olika typer av filter ges. Det ges även exempel på hur rörelseoskärpa eller dålig fokus kan korrigeras med invers filtrering.

Innehåll

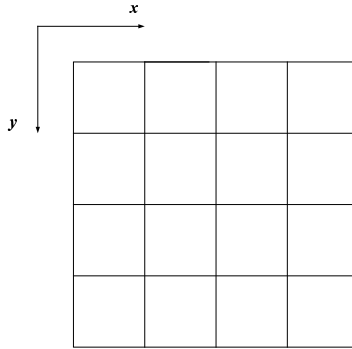
1	Introduktion	2
2	Digital representation av bilder	2
3	Diskret Fouriertransform i 2D	2
4	Visualisering	3
5	Filtrering av bilder	4
5.1	Invers filtrering	5
5.2	Vanliga filtertyper	6
5.3	Realisering av filter	6
6	Implementation i Matlab	8
7	Sammanfattning	9
	Referenser	10

1 Introduktion

Inom bildbehandling är filtrering ofta använt. De vanligaste filtren gör bilden mjukare eller hittar kanter i bilden. Sådana filter representeras ofta av en s.k. filterkärna av storlek 3x3, 5x5, 7x7 etc. Filtringen av bilden utförs som en tvådimensionell faltning med denna kärna. En faltning i spatialdomänen representeras som bekant av en multiplikation i Fourierdomänen. För stora kärnor innebär detta att antalet beräkningar kan reduceras avsevärt genom användandet av den snabba Fouriertransformen. Detta är ett sätt att arbeta med bilder i frekvensdomänen, men man har också stora möjligheter att skapa filter direkt i frekvensdomänen för att reducera eller framhäva vissa frekvenser i bilden. Att använda invers filtrering (avfaltning) är ett sätt att skapa filter i frekvensdomänen för att t.ex. korrigera dålig fokusering eller rörelseoskärpa.

2 Digital representation av bilder

En tvådimensionell digital bild kan beskrivas av en diskret funktion $f(x, y)$, där x och y är pixelpositionerna och $f(x, y)$ är intensiteten i punkten.



Figur 1: En bild som diskret funktion.

En grånivåbild representeras ofta av 8-bitars heltal. Det betyder att $2^8 = 256$ grånivåer kan visas, där 0 betyder svart och 255 vitt. Färgbilder representeras på ett liknande sätt, men där används flera lager för att representera olika spektralband (t.ex. RGB-modellen).

3 Diskret Fouriertransform i 2D

Den diskreta Fouriertransformen (DFT) av en bild $f(x, y)$ samplad i $M \times N$ punkter är definierad som

$$\mathcal{F}[f(x, y)] = F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2\pi i \left(\frac{ux}{M} + \frac{vy}{N} \right)}, \quad (1)$$

där $u = 0, \dots, M-1$ och $v = 0, \dots, N-1$ är diskreta frekvensvariabler. Frekvensen här har enheten 1/avstånd till skillnad från t.ex. ljudsignaler där enheten är 1/tid.

Originalbilden fås från den inverterade DFT'n:

$$\mathcal{F}^{-1}[F(u, v)] = f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi i \left(\frac{ux}{M} + \frac{vy}{N} \right)}, \quad (2)$$

för $x = 0, \dots, M-1$ och $y = 0, \dots, N-1$.

Transformen $F(u, v)$ kallas för funktionen f 's spektrum, och är i det generella fallet komplex:

$$F(u, v) = \operatorname{Re}[F(u, v)] + i \operatorname{Im}[F(u, v)].$$

I polära koordinater:

$$F(u, v) = |F(u, v)| e^{i\phi(u, v)},$$

där

$$|F(u, v)| = \sqrt{\operatorname{Re}[F(u, v)]^2 + \operatorname{Im}[F(u, v)]^2}$$

och

$$\phi(u, v) = \arctan \left(\frac{\operatorname{Im}[F(u, v)]}{\operatorname{Re}[F(u, v)]} \right).$$

$|F(u, v)|$ kallas för *amplitudspektrum* och $\phi(u, v)$ kallas för *fasspektrum*.

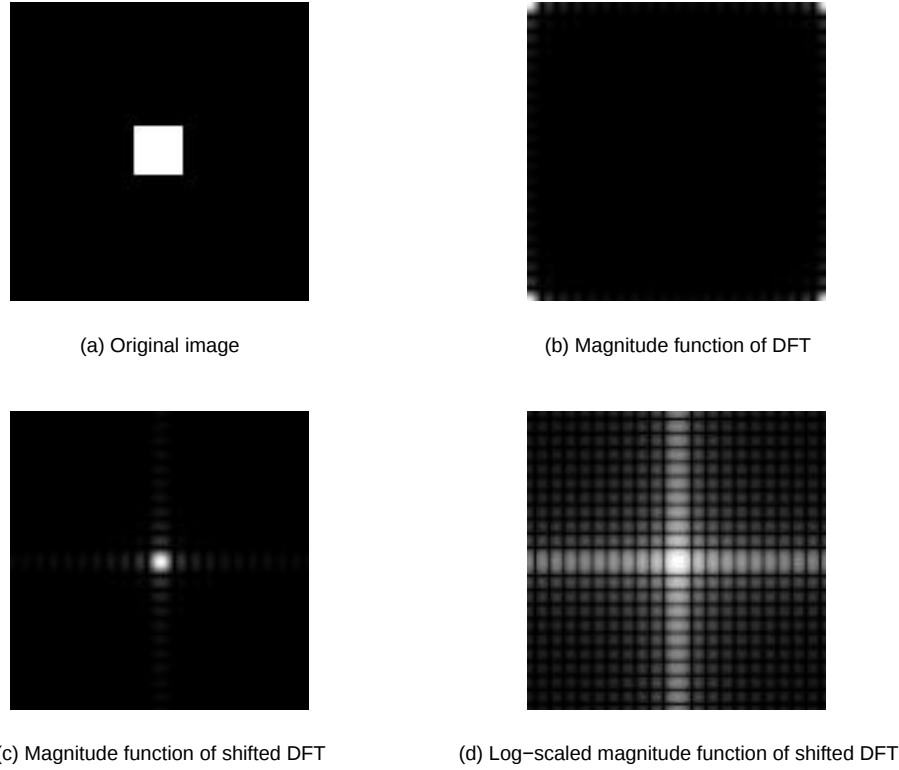
4 Visualisering

För att visualisera amplitudspektrum av en funktion som en bild används ofta funktionen

$$D(u, v) = c \log(1 + |F(u, v)|).$$

Där c är en skalfaktor för att passa in funktionsvärdena till gränivåskalan.

Man brukar också göra ett skift för att få origo i centrum, d.v.s. se till att låga frekvenser hamnar i mitten av bilden. Man beräknar alltså $F(u - M/2, v - N/2)$. I figur 2 visas hur amplitudspektrum kan visualiseras.



Figur 2: *Exempel på hur amplitudspektrum kan visualiseras.*

När det gäller fasspektrumet så är det inte intuitivt hur det ska visualiseras. Det finns modeller som utnyttjar en periodisk färgskala för att representera fasen, se [4].

5 Filtrering av bilder

Filtrering av bilder i spatialdomänen används ofta för att göra bilden mjukare eller för att hitta kanter i bilden. Ett utsmetande 3×3 medelvärdesfilter t.ex., beräknar medelvärdet av varje pixel och dess 8 grannar. En illustration finns i figur 3.

$$\frac{1}{9} \times \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

Figur 3: *Ett 3×3 medelvärdesfilter.*

Filtrering med medelvärdesfiltret ovan kan ses som en faltning av bilden $f(x, y)$ med en filterkärna $h(x, y)$ för att erhålla den resulterande bilden $g(x, y)$. Faltningen är:

$$g(x, y) = f(x, y) * h(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) h(x - m, y - n) \quad (3)$$

Som bekant representeras en faltning i spatialdomänen av en multiplikation i frekvensdomänen:

$$G(u, v) = F(u, v) H(u, v). \quad (4)$$

Detta ger möjligheten att beräkna DFT'n av $f(x, y)$ och $h(x, y)$, multiplicera transformerna och sedan inverstransformera för att erhålla resultatet $g(x, y)$. För stora filterkärnor är detta en effektiv algoritm om man kan använda den snabba Fouriertransformen (FFT).

En annan möjlighet är att skapa filtret $H(u, v)$ direkt i frekvensdomänen för att framhäva eller reducera vissa frekvenser.

5.1 Invers filtrering

En bild som är degraderad beroende på t.ex. rörelseoskärpa eller dålig fokus kan tecknas matematiskt med formeln

$$g(x, y) = f(x, y) * h(x, y) + n(x, y),$$

där h är någon olinjär funktion och n är brus. I frekvensdomänen fås:

$$G(u, v) = F(u, v)H(u, v) + N(u, v).$$

Genom att estimeras $H(u, v)$ med en matematisk modell kan ett filter skapas för att korrigera degraderingen. Om estimatet är $\tilde{H}(u, v) \approx H(u, v)$ kan filtret definieras som $\tilde{H}^{-1}(u, v)$ vilket resulterar i

$$F(u, v) \approx \tilde{H}^{-1}(u, v) (G(u, v) - N(u, v)).$$

Många gånger kan man bortse från bruset N , men ibland måste man göra ett estimat även av det.

För några specialfall kan degraderingsfunktionen H estimeras med ganska enkla matematiska modeller. Oskärpa beroende på en rörelse med konstant hastighet i x -riktningen kan beskrivas i frekvensdomänen med funktionen

$$\tilde{H}(u, v) = \frac{\sin(\pi v T u)}{\pi v u},$$

där T är exponeringstiden och v är hastigheten. En degradering som beror på dålig fokus beskrivs av funktionen

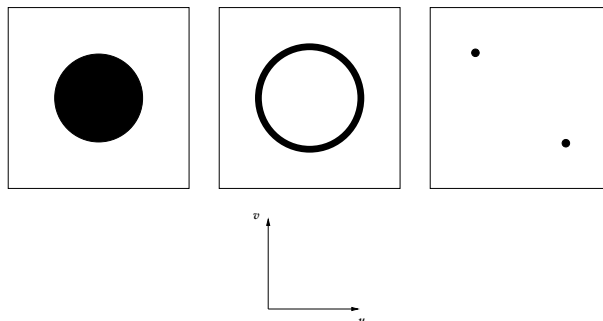
$$\tilde{H}(u, v) = \frac{J_1(a\rho)}{a\rho},$$

där J_1 är första ordningens Besselfunktion, $\rho^2 = u^2 + v^2$ och a är linsens förskjutning.

För dessa fall är det relativt lätt att hitta ett bra estimat till degraderingsfunktionen, men för en mer generell degradering är det ett svårt problem, se [2].

5.2 Vanliga filtertyper

Det finns olika typer av filter, men de vanligaste är lågpas- högpas-, bandpas- och bandstoppfilter. Även s.k. notchfilter används. Det vanligaste är att man använder filter som endast påverkar amplituden hos Fourierspektrumet, d.v.s. ingen fasvridning sker. Exempel på hur de olika typerna av filter ser ut i två dimensioner finns i figur 4.



Figur 4: Till vänster ett idealt högpasfilter, i mitten ett idealt bandstoppfilter, och till höger ett idealt notchfilter.

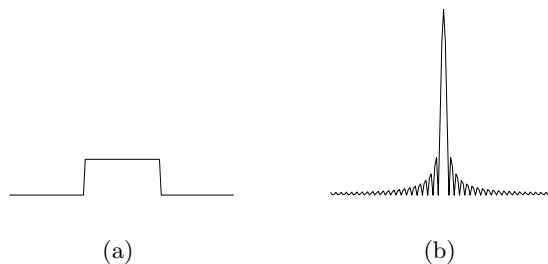
Ett lågpasfilter producerar utjämnade bilder, medan högpasfilter används för att ta fram snabba förändringar i bilden (kanter).

Bandstoppfilter dämpar frekvenser i ett band på ett givet avstånd från origo. Denna typ av filter används ofta för brusreducering när bruset består av kända frekvenser. Ett bandpassfilter kan användas t.ex. för att extrahera bruset för separat analys.

Notchfilter används för att dämpa frekvenser i en omgivning av centrumfrekvensen (u_0, v_0) .

5.3 Realisering av filter

I figur 4 är alla filter ideala. Ett idealt filter har värdet ett för frekvenser som ska släppas igenom, och noll för övrigt. I figur 5 ses ett endimensionellt idealt lågpasfilter och dess inversa Fouriertransform. Ideal filtrering ger ringningseffekter p.g.a. de skarpa övergångarna mellan frekvenserna.



Figur 5: (a) visar ett endimensionellt idealt lågpasfilter och (b) dess inversa Fouriertransform.

För att undvika ringningseffekter använder man approximationer av ideala filter istället. Vanligt är att använda förenklade versioner av Butterworthfilter som endast påverkar amplituden.

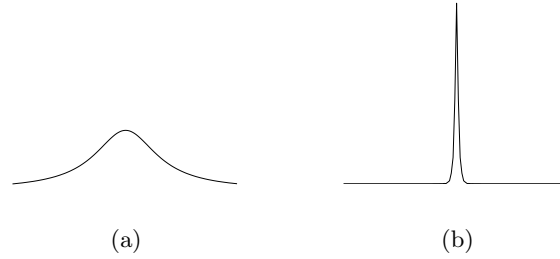
Ett Butterworth lågpasfilter av ordning n , och med gränsfrekvens på avståndet ρ_0 från origo definieras av funktionen

$$H_{\text{BWLP}}(\rho) = \frac{1}{1 + (\rho/\rho_0)^{2n}}, \quad (5)$$

där $\rho = \sqrt{u^2 + v^2}$. Ju högre grad n , desto mer liknar filtret ett idealt filter. Ett Butterworth högpasfilter definieras av funktionen $H_{\text{BWHP}}(\rho) = 1 - H_{\text{BWLP}}(\rho)$.

Den definition av Butterworthfilter som ges här är hämtad från [1] och skiljer sig från definitionen i [3]. Ekvation (5) beskriver egentligen den kvadrerade frekvensgången hos ett Butterworthfilter, men är det som används inom bildbehandlingen enligt [1].

I figur 6 syns ett endimensionellt Butterworth lågpasfilter och dess inversa Fouriertransform. Ringningarna är borta p.g.a. filtrets mjuka övergång mellan låga och höga frekvenser.



Figur 6: (a) visar ett endimensionellt Butterworth lågpasfilter och (b) dess inversa Fouriertransform.

Ett Butterworth bandpassfilter av ordning n och bandbredd W definieras av funktionen

$$H_{\text{BWBP}}(\rho) = \frac{1}{1 + \left(\frac{\rho^2 - \rho_0^2}{\rho W} \right)^{2n}}, \quad (6)$$

där ρ_0 är avståndet till bandets centrum. Motsvarande bandstoppfilter definieras av funktionen $H_{\text{BWBS}}(\rho) = 1 - H_{\text{BWBP}}(\rho)$.

Ett notchfilter av grad n centrerat i punkten (u_0, v_0) definieras av funktionen

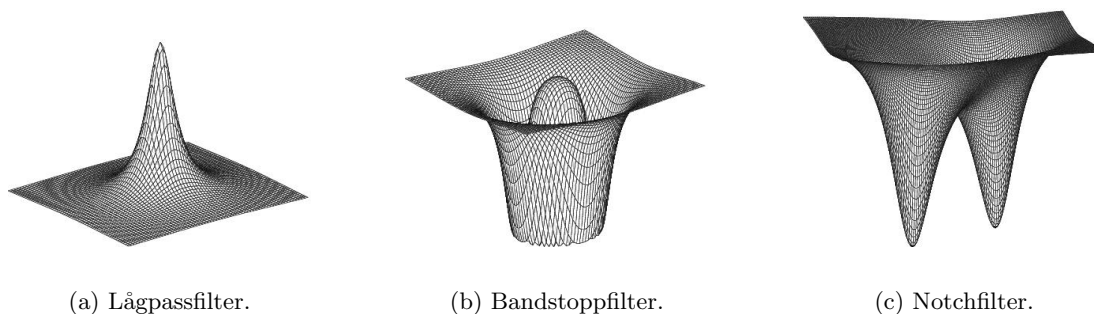
$$H_{\text{BWNP}}(u, v) = \frac{1}{1 + \left(\frac{D_1(u, v) D_2(u, v)}{D_0^2} \right)^n}, \quad (7)$$

där D_0 är den lokala gränsfrekvensen och

$$D_1(u, v) = \sqrt{(u - u_0)^2 + (v - v_0)^2},$$

$$D_2(u, v) = \sqrt{(u + u_0)^2 + (v + v_0)^2}.$$

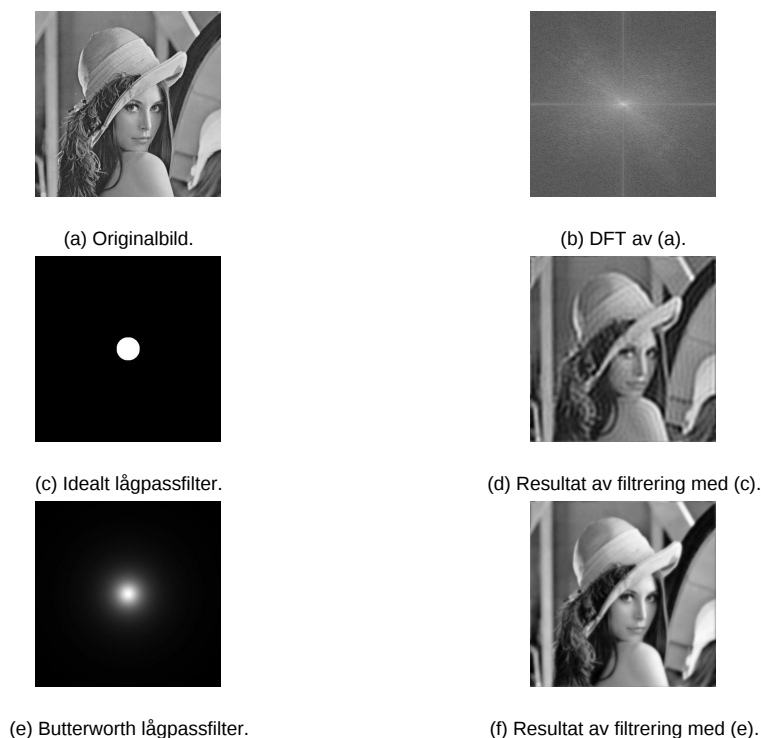
I figur 7 visas ytplottar på några av filtren ovan.



Figur 7: Ytplottar på Butterworthfilter.

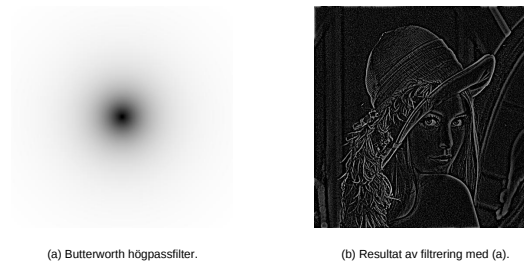
6 Implementation i Matlab

Några av filtren har skapats i Matlab och använts för att filtrera ett antal testbilder. Först jämfördes resultaten av ett idealt lågpassfilter och ett motsvarande Butterworthfilter. Originalbilden är av storlek 512×512 och gränsfrekvensen ligger på avståndet $\rho_0 = 512/16$ från origo. Se resultat i figur 8.



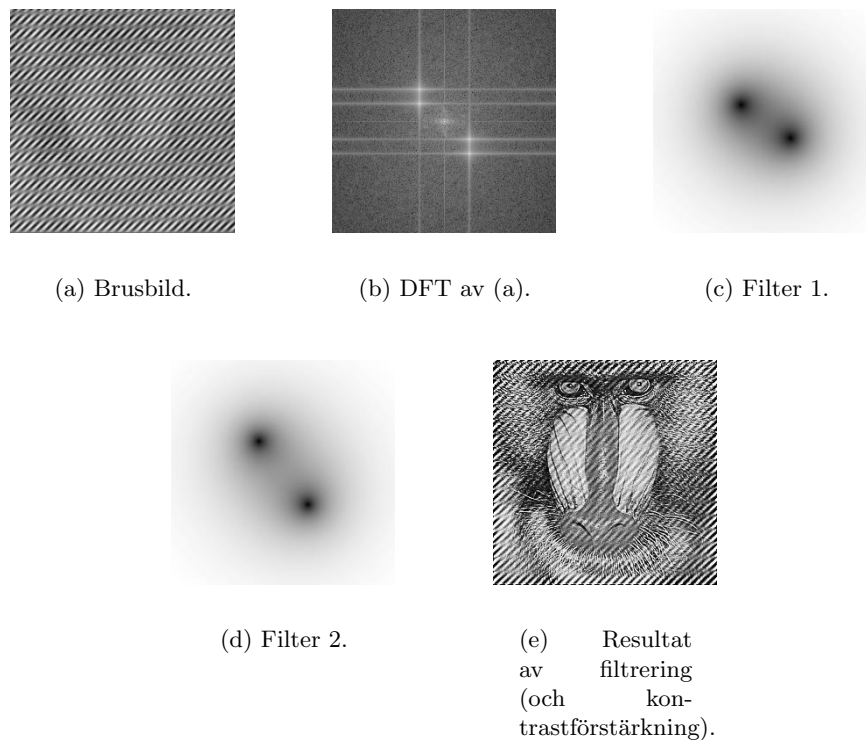
Figur 8: Jämförelse av lågpassfiltrering med ett idealt lågpassfilter och ett motsvarande Butterworthfilter. Ringningseffekterna är påtagliga i (d).

Ett Butterworth högpasfilter implementerades också för att testas på samma bild som i figur 8. Gränsfrekvensen ligger på avståndet $\rho_0 = 512/8$. Se resultat i figur 9.



Figur 9: *Högpasfiltrering med Butterworthfilter. Filtringen framhäver kanterna i bilden.*

För att undersöka hur brus kan reduceras genom filtrering i frekvensdomänen genererades en testbild med kraftigt brus bestående till största del av två frekvenser. Figur 10 visar hur sådant brus kan reduceras genom att använda två notchfilter. SNR i testbilden är ca -9 dB.



Figur 10: *Resultat av filtrering med Butterworth notchfilter.*

7 Sammanfattning

Arbetet har visat hur bilder kan behandlas i frekvensdomänen genom att använda den diskreta Fouriertransformen. Det har givits exempel på hur Fourierspektrumet kan visualiseras och hur man med ganska enkla filter kan få goda resultat vid filtrering i frekvensdomänen.

Referenser

- [1] R.C. Gonzalez, R.E. Woods. *Digital image processing*.
Addison Wesley publishing company, 2nd edition, 1992.
- [2] M. Sonka, V. Hlavac, R. Boyle. *Image processing, analysis and machine vision*.
PWS Publishing, 1999.
- [3] A. Svärdröm. *Signaler och system*.
Studentlitteratur, 1999.
- [4] K. Cowtan. *Kevin Cowtan's picture book of Fourier transforms*.
URL: <http://www.yorvic.york.ac.uk/~cowtan/fourier/fourier.html> (2002-12-08)