Kapitel 1

Filtrering af billeder

Ved filtrering af et billede ønsker man sædvanligvis af forbedre billedkvaliteten . Det givne billede kan eksempelvis være uskarpt, og filtreringens formål at øge kontrasten i billedet. Alternativt kan billedet være fremkommet ud fra et ideelt billede (der ikke nødvændigvis har eksisteret konkret) ved addition af støj. Filtreringens formål er her at fjerne støjen eller anderledes sagt at frembringe et billede der er tættest muligt på det ideelle.

I visse tilfælde, hvor billedforringningsprocessen er kendt, er det muligt at angive (matematisk set) optimale filtre. Ofte er denne process imidlertid ikke kendt eksakt. Dette betyder at vurderingen af en filtrering er en subjektiv sag. Dette er i særlig grad aktuelt hvor mennesker skal vurdere resultatet. En sådan vurdering er bl.a. afhængig af træning. Endelig er resultatet af en matematisk set optimal filtrering ikke nødvendigvis den der forekommer øjet optimal.

Ved valget af filtreringsmetode er det vigtigt at holde sig filtreringens formål for øje. Hvad er problemet med billederne? Hvordan er billederne karakteriseret? Kan der opstilles en model for billedforringelsen? Hvad skal de filtrerede billeder benyttes til?

Den simplest tænkelige filtrering af et billede er at modificere billedelementerne et for et. Dette er f.eks. tilfældet ved skalering af intensitetsværdierne i et billede. Som introduktion til dette "simpleste tilfælde" beskrives i afsnit 2 en meget effektiv skaleringsmetode, kaldet *histogram linearisering*.

I visse tilfælde har den forvrængning, som billederne har været udsat for, berørt billedkoordinatsystemet mere end intensitetsværdierne. I afsnit 3 beskrives nogle metoder til *opretning* af sådanne billeder. Herunder beskrives nogle metoder til interpolation af billedværdier.

I afsnit 4 beskrives *Fouriertransformationen*. Denne transformation er grundlaget for en lang række filtre. Desuden giver transformationen et beskrivelsesværktøj, der ofte benyttes til at karakterisere signaler og billeder.

I afsnit 5 udledes en simpel model for billedforringelse. I afsnit 6 beskrives en række lineære positionsinvariante filtre. I afsnit 7 diskuteres nogle ofte benyttede ikke-lineære filtre. I afsnit 8 beskrives en relativ ny type af (ikke-lineære) filtreringer, der bygger på *matematisk morfologi* (formlære).

1.1 Introduktion til filtrering

At filtrere et billede betyder i virkeligheden kun at modificere billedelementerne efter et i forvejen fastlagt skema. Signalet hhv. billedet betragtes ofte som en kontinuert eller samplet funktion i en eller to dimensioner, og filtreringen består ofte i en *foldning* af denne funktion med en filterfunktion. Foldningsoperatoren beskrives i et senere underafsnit. I denne klassiske metode til filtrering er det muligt at samtlige værdier i inddatabilledet indgår i beregningen af samtlige uddataværdier. Hvis dette er tilfældet kaldes filteret for *globalt* . Der tales også om IIR (Infinite Impulse Response) -filtrering .

Indgår kun en begrænset delmængde af inddataværdierne siges filteret at være *lokalt* , eller at have begrænset *støtte* . Der tales her om FIR (Finite Impulse Response)-filtrering . Ofte kan filtreringen beskrives ved for hver pixel at lægge en maske sådan at pixelen kommer i centrum af masken, og derefter at beregne filterværdien (for den givne pixel) som et vægtet gennemsnit af billedværdierne i masken.

Hvis værdierne af billedkoordinaterne indgår i beregningen af filterværdien kaldes filtret *positionsvariant* . I modsat fald kaldes filtret *positionsinvariant.* Hvis beregningsmåden for filterværdien afhænger af inddataværdierne siges filteret at være *adaptivt* . I modsat fald kaldes filteret *konstant* . Eksempelvis er foldninger positionsinvariante og konstante.

Eksempler

Nedenstående ligninger angiver beregningsmåder for fem filtreringer af et signal $I(x)$. Filterresultatet betegnes $O_i(x)$.

$$\begin{aligned} O_1(x) &= 2 * I(x) - 5 && (1.1)\\ O_2(x) &= [I(x-1) + I(x) + I(x+1)]/3 && (1.2)\\ O_3(x) &= O_3(x-1) - 2 * O_3(x-2) + I(x) && (1.3)\\ O_4(x) &= \text{if } I(x) \geq T \text{ then } 1 \text{ else } 0 && (1.4)\\ O_5(x) &= [I(x+1) - I(x-1)]/2 && (1.5)\\ O_6(x) &= x * I(x) && (1.6) \end{aligned}$$

I ovenstående eksempler har alle bortset fra det tredie filter begrænset støtte. Dette er et såkaldt IIR-filter, og er endvidere *kausalt* idet beregningen af filterværdien i x kun afhænger "tidligere beregnede" filterværdier, dvs. værdierne $I(x), I(x-1), I(x-2), \ldots$. Filtrene 2 hhv. 5 har en støtte på tre hhv. to pixels hvorimod det første og det fjerde filter har en støtte på 1 pixel. Alle filtrene, bortset fra det sjette, er positionsinvariante. Kun det fjerde filter kan med lidt god vilje kaldes adaptivt, idet konstanterne "0" og "1" kan opfattes som separate beregningsudtryk.

En simpel filtreringsmåde er at foretage en afbilding af intensitetsværdierne:

$$O(x) = T(I(x)) \tag{1.7}$$

hvor T eksempelvis er en afbildning $T : \Re \rightarrow \{0, 1\}$, eller $T : [0..255] \rightarrow [0..255]$. En simpel modifikation kunne være at negere billedet:

$$T(t) \quad = \quad 255 - t \tag{1.8}$$

Typisk foretages sådanne simple modifikationer ud fra en global statistisk analyse af fordelingen af intensitetsværdier i billedet. En lineær skalering af værdierne kunne være en sådan modifikation.

$$O(x, y) \quad = \quad \frac{I(x, y) - I_{min}}{I_{max} - I_{min}} M \tag{1.9}$$

hvor M er den største ønskede intensitetsværdi. Filtreringen består her af to passager. I første passage beregnes I_{max} og I_{min}. I anden passage foretages ovenstående korrektion. I dette tilfælde er filtreringen global (og positionsinvariant).

Simple lineære skaleringer (som ovennævnte) kan forbedre billedkvaliteten markant hvis billedet kun udnytter et bånd af de mulige intensitetsværdier. I mange situationer er denne transformation dog for simpel. Typisk kan filtreringen forbedres ved udnytte mere global information (end blot minimum og maksimum) og ved at udnytte lokal information (dvs. hvorledes I varierer i en omegn om det punkt hvor filtrerværdien beregnes). Dette er eksemplificeret ved ovenstående filter O_2, hvor der beregnes et lokalt gennemsnit af værdierne. Dette vil ofte have en støjdæmpende effekt.

En oplagt filtreringsmetode er at beregne filterværdien som et vægtet gennemsnit af inddataværdierne i omegnen.

$$O(x, y) \quad = \quad \sum_{(i,j) \in S} w_{ij} I(x - i, y - j) \tag{1.10}$$

hvor vægtene (*filterkoefficienterne*) w_{ij} bestemmer filterfunktionen, og S angiver støtten for filteret. For en typisk 3×3 støtte vil $S = [-1, 0, 1] \times [-1, 0, 1]$. Ovenstående eksempel svarer til filtrering ved foldning (se nedenfor). Andre metoder, f.eks. at benytte logiske operatorer i stedet for sum-operatoren, er også mulige (se afsnit om matematisk morfologi).

1.1.1 Foldninger

Foldningen af to en-dimensionale reelle funktioner $f(x)$ og $g(x)$ er defineret ved integralet:

$$f \star g(x) \quad = \quad \int_{-\infty}^{\infty} f(\alpha) g(x - \alpha) d\alpha \tag{1.11}$$

Definitionen udvides let til funktioner defineret på $\Re^n$.

Foldningsbegrebet er centralt i billedbehandling fordi alle lineære og positionsinvariante filtreringer kan beskrives som foldning af det oprindelige signal/billede med en filterfunktion. Som vi skal se senere stiller den klassiske (kontinuerte) funktionsteori et overordentlig nyttigt værktøj til rådighed ved beskrivelse af mange billedbehandlingsoperationer. Foldninger og fouriertransformationer er de operationer, der er mest benyttede. Om foldningsoperationen gælder følgende regneregler:

$$f \star g \quad = \quad g \star f \tag{1.12}$$

$$f \star (g + h) \;=\; f \star g \;+\; f \star h \tag{1.13}$$
$$f \star (g \star h) \;=\; (f \star g) \star h \tag{1.14}$$

Yderligere gælder at resultatet af en foldning er begænset og differentiabel (med begrænset afledet) når blot en af funktionerne er det. I dette tilfælde gælder desuden at den afledede af foldningen er lig foldningen af en af funktionerne med den afledede af den anden:

$$D(f \star g) \;=\; Df \star g \;=\; f \star Dg \tag{1.15}$$

hvor D angiver differentialoperatoren.

Det ville have været overordentlig nyttigt om foldningsoperationen var forsynet med et 1-element, dvs. at der fandtes en invers foldning. Dette er desværre **ikke** tilfældet. I billedbehandling er det ofte tilfældet, at det aktuelle optagne billede O er forvrænget på en måde der kan beskrives ved en foldning $O = h \star I$, hvor I angiver det ideelle billede. Selv om er h er kendt eksakt er det sjældent muligt at rekonstruere I ud fra observationen O, fordi det ikke er muligt at *udfolde* billedet (dvs. at folde baglængs).

Som det ses af definitionen på foldningen består denne for hver værdi af x i følgende elementer:

1) En spejling: $g(\alpha) \rightarrow g(-\alpha)$.
2) En translation: $g(-\alpha) \rightarrow g(-\alpha + x)$.
3) Konstruktion af produktfunktionen $h(\alpha) \;=\; g(-\alpha + x)f(\alpha)$.
4) Integration af produktfunktionen $h(\alpha)$.

Eksempel

Betragt følgende to funktioner:

$$f(\alpha) \;=\; \begin{cases} 1 & \text{hvis } 0 \leq \alpha \leq 1 \\ 0 & \text{ellers} \end{cases} \tag{1.16}$$

$$g(\alpha) \;=\; \begin{cases} 1/2 & \text{hvis } 0 \leq \alpha \leq 1 \\ 0 & \text{ellers} \end{cases} \tag{1.17}$$

Ved spejling af g og translation med $x \leq 0$ vil de områder af de to funktioner, der er forskellig fra nul, ikke lappe ind over hinanden. Produktfunktionen vil derfor være nulfunktionen, og integral-værdien være lig 0. Det samme argument holder når der translateres med en x-værdi større end 2. For $0 \leq x \leq 2$ vil de to funktioner $f(x)$ og $g(x - \alpha)$ overlappe i større eller mindre grad. Overlapningen er maksimal for $x = 1$. I dette tilfælde får produktfunktionen værdien $\frac{1}{2}$ for $0 \leq \alpha \leq 1$ og ellers nul. Integralværdien er således lig $\frac{1}{2}$. Det er relativt let at udregne foldningsresultatet:

$$f \star g(x) = \begin{cases} x/2 & \text{hvis } 0 \leq x \leq 1 \\ 1 - x/2 & \text{hvis } 1 \leq x \leq 2 \\ 0 & \text{ellers} \end{cases} \tag{1.18}$$

Foldningen af to "firkantfunktioner" er således en "trekant".

Impulsfunktionen

Impulsfunktionen $\delta(x)$ er defineret som den uegentlige funktion, der har værdi lig 0 uden for $x = 0$, og som har integral lig 1. En impulsfunktion $\delta(x - x_0)$ i x_0 defineres tilsvarende:

$$\int_{-\infty}^{\infty} \delta(x - x_0)dx = \int_{-x_0^-}^{x_0^+} \delta(x - x_0)dx = 1 \tag{1.19}$$

En impulsfunktion tegnes som en pil rettet opad fra x-aksen. Foldningen af en funktion f med en impulsfunktion i x_0 er givet ved:

$$f \star \delta(x_0) = \int_{-\infty}^{\infty} f(x)\delta(x_0 - x)dx = f(x_0) \tag{1.20}$$

idet kun værdien i $x = x_0$ vil bidrage til integrationen. Hvis vi har givet en kontinuert funktion $f(x)$ kan vi altså *sample* denne i x_0 ved ovennævnte foldning. Et endeligt/uendeligt *impulstog* er defineret ved en endelig/uendelig følge af impulser:

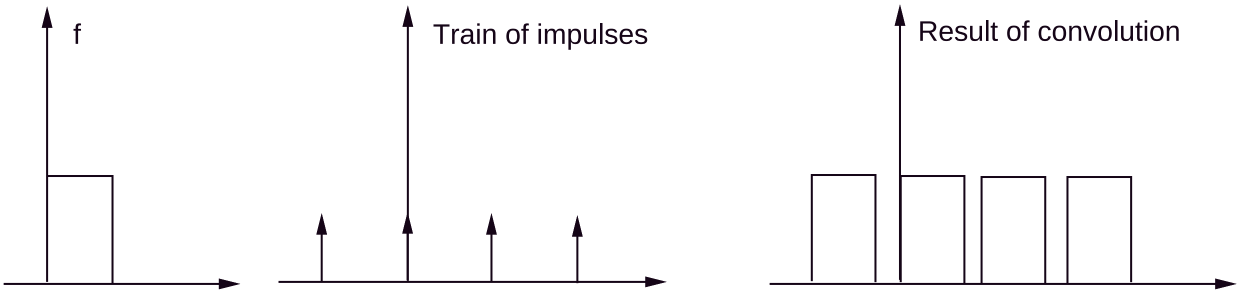
$$\sum_{n \in \Sigma} \delta(x - (nT + x_0)) \tag{1.21}$$

hvor Σ er en delmængde af de hele tal, og T og x_0 er reelle tal. Et impulstog (eller del af et impulstog) tegnes som et antal pile der udgår fra $x_0 - nT, \ldots, x_0 - T, x_0, x_0 + T, \ldots x_0 + nT$. Foldningen af en funktion med et impulstog kan betragtes som en kopiering. Tag f.eks. definitionen af $f(\alpha)$ i ovenstående eksempel (1.16), og antag at $x_0 = 0$ og at afstanden T mellem impulserne i ovenstående tog er større end 1. Da vil resultatet af foldningen være en funktion der har værdien 1 i alle punker x således at: $(x \bmod T) \leq 1$ og ellers værdien 0. Dette er illustreret i Figur 1.1. Vi skal i senere afsnit om *foldningssætningen* hhv. *samplingsætningen* se nærmere på anvendelsen af foldninger og impulstog.

Den diskrete foldning

Antag at f hhv. g er givet ved et endelig sæt af værdier $(f(0), f(1), f(2), \ldots, f(A-1))$ og $(g(0), g(1), g(2), \ldots, g(B-1))$. Som vi skal se i afsnittet om fouriertransformationen vil det ofte være en fordel at betragte hvert sæt af værdier som tilhørende en periode af en periodisk funktion. For at undgå at de første hhv. sidste værdier i en periode influerer på beregningen af de sidste hhv. første værdier af foldningen er det nødvendigt at udvide funktionerne f og g ved:

$$f_e(x) = \left\{ \begin{array}{ll} f(x) & \text{hvis } 0 \leq x \leq A-1 \\ 0 & \text{hvis } A \leq x \leq M-1 \end{array} \right\} \tag{1.22}$$



Figur 1.1: Foldning af en funktion med et tog af impulsfunktioner

og

$$g_e(x) = \left\{ \begin{array}{ll} g(x) & \text{hvis } 0 \leq x \leq B-1 \\ 0 & \text{hvis } B \leq x \leq M-1 \end{array} \right\} \tag{1.23}$$

hvor M er den fælles periode af de udvidede signaler, og også bliver perioden af foldningsresultatet. Det kan vises at M skal vælges således at:

$$M \geq A + B - 1 \tag{1.24}$$

Hvis M vælges mindre end denne størrelse vil de enkelte perioder kunne overlappe (når den ene funktion spejles og translateres). Dette vil ofte give anledning til andre værdier end de forventede. Fænomenet kaldes *warparound error*. Med ovenstående afklaringer er den diskrete to-dimensionale foldning defineret ved:

$$f_e(x,y) \star g_e(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_e(m,n) g_e(x-n, y-m) \tag{1.25}$$

hvor M og N er valgt så de opfylder ovennævnte krav. Som tidligere nævnt kan ovennævnte foldning betragtes som et vægtet gennemsnit.

1.2 Histogram linearisering

I det følgende beskrives en transformationsmetode hvorved intensiteterne ændres således at tæthedsfunktionen for det resulterende billede vil være uniform. Anderledes sagt, vil alle intensitetsværdier udnyttes lige godt (i teorien). Metoden kaldes histogram linearisering (eng. histogram equalization). I det følgende antages at intensitetsværdierne er skaleret til intervallet $[0:1]$. For at udnytte resultaterne fra sandsynlighedsregningen formuleres metoden i det kontinuerte tilfæde, dvs. at alle intensitetsværdier $r \in [0:1]$ er mulige. Intensitetsværdien r i et givet billedelement opfattes som en stokastisk variabel. Vi kan derfor tale om sandsynlighedsfordelingen $p_r(r)$. Transformationen kan skrives:

$$s = T(r) \tag{1.26}$$

Filterværdien fremkommer ved multiplikation af s med en skaleringsfaktor (typisk 255.0). Funktionen T antages at være monotont voksende, differentiabel med kontinuert afledede.

Desuden antages at T afbilder intervallet $[0:1]$ på sig selv. Dette betyder at afbildningen bevarer ordningen fra "mørke" intensiteter til "lyse", samt at vi kan tale om den inverse transformation T^{-1}. I det følgende vil vi få brug for følgende sætning fra sandsynligheds-regningen:

Sætning

Lad X være en kontinuert stokastisk variabel med tæthed f og definitionsmængde $[0:1]$. Lad ϕ være en bijektiv funktion af $[0:1]$ på $[0:1]$. Antag at ϕ' eksisterer og er kontinuert. Lad $\psi = \phi^{-1}$. Da er tætheden g for $Y = \phi(X)$ givet ved:

$$g(y) = f(\psi(y))\,|\psi'(y)| \quad \text{for} \quad y \in [0:1] \tag{1.27}$$

Beviset er i to dele: Først antages at ϕ er voksende. Da er:

$$G(y) = P(Y \leq y) = P(\phi(X) \leq y) = P(X \leq \psi(y)) = F(\psi(y)) \tag{1.28}$$

Af kædereglen for differentiation fås:

$$g(y) = G'(y) = F'(\psi(y))\psi'(y) = f(\psi(y))\,|\psi'(y)| \tag{1.29}$$

Antag nu at ϕ er aftagende. Da får vi:

$$G(y) = P(\phi(X) \leq y) = P(X \geq \psi(y)) = 1 - F(\psi(y)) \tag{1.30}$$

Derfor har vi:

$$g(y) = G'(y) = -F'(\psi(y))\psi'(y) = f(\psi(y))\,|\psi'(y))| \tag{1.31}$$

Hermed er sætningen bevist.

Givet et billede med tæthedsfunktion $p_r(r)$ samt transformationen $s = T(r)$, er tætheden $p_s(s)$ for s (jvf. ovenstående sætning) givet ved:

$$p_s(s) \;=\; \left[p_r(r)\frac{dr}{ds}\right]_{r=T^{-1}(s)} \tag{1.32}$$

Betragt nu transformationen givet ved den (cummulative) fordelingsfunktion:

$$s = T(r) = \int_0^r p_r(w)dw \quad \text{for} \quad 0 \leq r \leq 1 \tag{1.33}$$

For denne transformation er $ds/dr = p_r(r)$. Ved substitution i (1.32), ses at $p_s(s) = 1$. For den angivne transformation vil samtlige intensiteter i det transformerede billede altså være lige sandsynlige. Transformationen vil som regel kunne forbedre billedkvaliteten betydeligt. Dette kan især være tilfældet ved under- eller overbelyste billeder.

Eksempel

Lad $p_r(r) = \frac{\pi}{2} cos(\frac{\pi}{2}r)$. For $s = T(r) = \int_0^r p_r(w)dw = sin(\frac{\pi}{2}r)$ fås at: $r = \frac{2}{\pi}\arcsin(s)$. Vi kan nu indsætte i formel (1.32):

$$\begin{aligned} p_s(s) &= \left[p_r(r)\frac{dr}{ds}\right]_{r=T^{-1}(s)} \\ &= \left[\frac{\pi}{2}\cos(\frac{\pi}{2}r)\frac{dr}{ds}\right]_{r=\frac{2}{\pi}\arcsin(s)} \\ &= \frac{\pi}{2}\cos(\arcsin(s))\;\frac{2}{\pi}\frac{1}{\sqrt{1-s^2}} \\ &= 1 \end{aligned}$$

I praksis er intensiteterne i et digitalt billede ikke kontinuert fordelt. Dette nødvendiggør, efter beregning af de nye værdier, at tilnærme disse til de mulige intensitetsværdier. Herved kan det ske, at antal forskellige intensitetsværdier i det filtrerede billede bliver mindre end antal forskellige intensitetsværdier i det oprindelige billede. Væsentlig er også at bemærke, at alle pixels med samme intensitet før filtreringen naturligvis også vil have samme intensitet efter filtreringen.

Eksempel

Antag et billede med 8 mulige intensitetsværdier. Lad hyppighederne af de enkelte værdier være givet som angivet i tabellen:

intensitet	*hyppighed*	*accumuleret*	*interval*
0	0.00	0.00	0.000
1	0.09	0.09	0.125
2	0.01	0.10	0.250
3	0.30	0.40	0.375
4	0.29	0.69	0.500
5	0.10	0.79	0.625
6	0.11	0.90	0.750
7	0.10	1.00	0.875

I tabellen er også angivet den accumulerede sandsynlighed, samt den nedre intervalgrænse for tildelingen af den diskrete intensitetsværdi. Ved histogramlineariseringen ses at pixels med oprindelig intensitet 0, 1, og 2 alle får intensiteten 0 efter filtreringen. Afbildningen bliver:

$$\begin{array}{cccc} 0 \rightarrow 0 & 2 \rightarrow 0 & 4 \rightarrow 5 & 6 \rightarrow 7 \\ 1 \rightarrow 0 & 3 \rightarrow 3 & 5 \rightarrow 6 & 7 \rightarrow 7 \end{array}$$

I det filtrerede billede bliver intensitetsværdierne 1, 2, og 4 altså ikke udnyttet. Desuden vil man kunne se at histogrammet over de transformerede værdier kun meget approksimativt er ligefordelt. Dette skyldes tillempningen af den kontinuerte teori til det diskrete tilfælde.

Direkte histogramspecifikation

I visse tilfælde er formålet med filtreringen at fremhæve bestemte udsnit af et billede, eller generelt at forbedre kvaliteten, dog således at særlige detalier fremstår klarere. I sådanne tilfælde er den lineære histogrammodifikation som regel utilstrækkelig. Hvad man ofte ønsker er, interaktivt at kunne specificere et ønsket histogram. Dette er som regel muligt med moderne visualiseringsprogrammer, men er ikke hensigtsmæssigt når mange billeder skal filtreres således at deres intensitetsfordeling er ens. Problemet er her at vælge den resulterende fordeling. Det er overordentlig svært at angive et sådant "ideelt" histogram, selv med godt kendskab til det oprindelige billede.

Antag at formålet med filtreringen er at opnå et billede med tæthedsfunktion $p_t(t)$, f.eks. en normalfordeling med middelværdi lig 128 og med spredning lig 32. Igen vil vi kun betragte det kontinuerte tilfælde. Lad $v = G(t)$ være den transformation, der lineariserer fordelingen af v:

$$v \;=\; G(t) \;=\; \int_0^t p_t(w)dw \tag{1.34}$$

Forelagt et billede, med intensitetsfordeling $p_r(r)$ kan dette transformeres på samme måde. Lad $v = T(r)$ være den transformation, der lineariserer fordelingen af v. Vi kan derfor beskrive transformationen fra r til t ved transformationssammensætningen:

$$t \;=\; G^{-1}\left[T(r)\right] \tag{1.35}$$

I det kontinuerte tilfælde kan det være vanskeligt at finde et analytisk udtryk for G^{-1}. Dette problem er ikke til stede i det diskrete tilfælde, hvor vi kan beregne og foretage transformationen ved tabelopslag.

Det kan bemærkes at maskiner til billedbehandling ofte (i matriellet) har indbygget sådanne opslagstabeller (kaldet LUT's for "look up tables", eller farvetabeller). Transformationen af et billede sker her (i sand tid) ved ændring af tabelværdierne.

Lokal histogram linearisering

Lokal (eller adaptiv) histogram linearisering er en metode hvor billedet opdeles i en række områder, og lineariseringen foretages seperat for hvert område. Ved en opdeling af billedet i disjunkte områder vil man ofte kunne iagttage en "skakbord-effekt". Dette er så generende at områderne i praksis altid vælges overlappende. Som regel er områderne kvadratiske, og overlapper således at modifikationen for hvert område kun berører billedelementet i midten af området. Områdernes størrelse er meget betydningsfulde for filtrets virkemåde. Ved valg af små områder vil enhver intensitetsændring blive forstærket. Dette kan virke som en egentlig støjfremhævning. Afhængig af formålet med filtreringen, fås den bedste virkning typisk ved valg af områdestørrelsen større end 100-200 pixels.

Det skal bemærkes at beregningen af histogrammet (og transformationen) ikke behøver at ske på ny hver gang området flyttes een pixel. Her vil jo blot ske det, at en "søjle" forsvinder ud af området mens en ny kommer til syne. I praksis implementeres metoden ved løbende at modificere histogrammet for intensiteterne i området.

1.3 Geometrisk opretning af billeder

I visse tilfælde har den forvrængning, som billedet (under optagelsen) har været udsat for, bevirket at billedkoordinatsystemet ikke svarer til det, som vi ønsker at betragte eller analysere billedet i. I tilfælde hvor et *referencebillede* eksisterer på forhånd, er den geometriske opretning eller *registrering* relativt enkel. Denne situation opstår eksempelvis ved analyser af fly- og satellitbilleder, hvor man ofte ønsker at sammenholde et nyt billede med et ældre. Opgaven er, at transformere det ny billede, således at koordinatværdierne for en række udvalgte billedpunkter bliver identiske med koordinatværdierne for de samme billedpunkter i referencebilledet. Anderledes sagt ønsker man at bestemme den koordinattransformation der afbilder det nye billede over i referencebilledet. Det er sædvanligt, for hver koordinat, at benytte polynomier af lav grad som transformationer. Afhængig af billederne og den måde hvorpå de er optaget, kan transformationerne dog være vilkårligt komplicerede.

Lad (x_0, y_0) være koordinater i det observerede billede I_0 og lad (x, y) være koordinater i referencebilledet I således at $I(x, y)$ og $I_0(x_0, y_0)$ afbilder samme punkt (område) i scenen. En 2.ordens polynomiel koordinattransformation kan da beskrives ved:

$$\begin{aligned} x_0 &= a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 xy + a_5 y^2 \\ y_0 &= b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 xy + b_5 y^2 \end{aligned} \tag{1.36}$$

En 1.ordens transformation vil tillade en vilkårlig affin afbildning, omfattende translation, rotation, skalering og skævvridning (således er det ny koordinatsystem ikke nødvendigvis retvinklet). Ovennævnte koordinattransformation kræver bestemmelse af 12 parametre. Hvis en 1.ordens transformation kan beskrive forvrængningen behøves kun 6 parametre. Det er sjældent nødvendigt at gå over 3.ordens transformationer, der ialt kræver 20 parametre. I praksis estimeres parametrene ved at bestemme et stort antal (20 - 50 eller flere) "samhørende" koordinatværdier (*paspunkter*) (x, y) og (x_0, y_0) i de to billeder. Heved fremkommer et lineært ligningssystem, der kan "løses" ved *mindste kvadraters metode* .

Antag at der benyttes N paspunkter. Dette giver $2N$ ligninger. Lad $\mathbf{C_x}$ hhv. $\mathbf{C_y}$ være søjlevektorer bestående af de N x-koordinatværdier hhv. de N y-koordinatværdier i I_0-billedet. Lad $\mathbf{A}$ hhv. $\mathbf{B}$ være søjlevektorer af de 12 ukendte koefficienter. Lad endvidere $\mathbf{X}$ være en $N \times 6$ -matrice bestående af koordinatværdierne i I-billedet.

$$\mathbf{C_x} = \begin{pmatrix} x_0(1) \\ x_0(2) \\ \vdots \\ x_0(N) \end{pmatrix} \qquad \mathbf{C_y} = \begin{pmatrix} y_0(1) \\ y_0(2) \\ \vdots \\ y_0(N) \end{pmatrix} \qquad \mathbf{A} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_5 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_5 \end{pmatrix} \tag{1.37}$$

$$\mathbf{X} = \begin{pmatrix} 1 & x(1) & y(1) & x^2(1) & x(1)y(1) & y^2(1) \\ 1 & x(2) & y(2) & x^2(2) & x(2)y(2) & y^2(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x(N) & y(N) & x^2(N) & x(N)y(N) & y^2(N) \end{pmatrix} \tag{1.38}$$

Vi kan nu skrive ligningerne (1.36) som:

$$\mathbf{C_x} = \mathbf{XA} \quad \text{og} \quad \mathbf{C_y} = \mathbf{XB} \tag{1.39}$$

Parametrene $\mathbf{A}$ og $\mathbf{B}$ kan altså findes separat. I de følgende vil vi finde $\mathbf{A}$ ved anvendelse af *mindste kvadraters metode*. Bemærk at $\mathbf{C_x}$ omdøbes til $\mathbf{C}$.

For givet valg af $\mathbf{A}$ kan vi udregne kvadratfejlen:

$$e^2 = (\mathbf{C} - \mathbf{XA})^t(\mathbf{C} - \mathbf{XA}) = \|\mathbf{C} - \mathbf{XA}\|^2 \tag{1.40}$$

hvor t angiver transponering. Vi ønsker at vælge $\mathbf{A}$ således at e^2 minimeres. Dette gøres på sædvanlig måde, ved at differentiere, og sætte resultatet lig nul.

$$\frac{\partial e^2}{\partial \mathbf{A}} = -2\mathbf{X}^t(\mathbf{C} - \mathbf{XA}) = \vec{0} \tag{1.41}$$

Videre fås:

$$\mathbf{X}^t\mathbf{C} = \mathbf{X}^t\mathbf{XA} \tag{1.42}$$

og dermed

$$\mathbf{A} = \mathbf{X}^\sharp\mathbf{C} = \left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{C} \tag{1.43}$$

hvor matricen $\mathbf{X}^\sharp$ kalden den *pseudoinverse* til matricen $\mathbf{X}$. Bemærk at den pseudoinverse kun indeholder talværdier (nemlig, de målte koordinatværdier for paspunkterne). Estimationen kræver således invertering af to 6×6 matricer. Vi skal ikke her diskutere de numeriske problemer, som der kan være forbundet hermed.

Selve transformationen foregår ved for hver billedpunkt (x, y) i referencekoordinatsystemet, at bestemme den tilsvarende koordinat (x_0, y_0) i koordinatsystemet for det observerede billede I_0, og derefter at “kopiere” intensitetsværdien:

$$I^{new}(x, y) := I_0(x_0(x, y), y_0(x, y)) \tag{1.44}$$

Herved opstår der det problem at de estimerede koordinater x_0 og y_0 vil være reelle størrelse. Da intensiteten I_0 kun kendes i billedelementerne, dvs. for heltallige koordinatvædier, er det nødvendig at foretage en interpolation af intensitetsfunktionen.

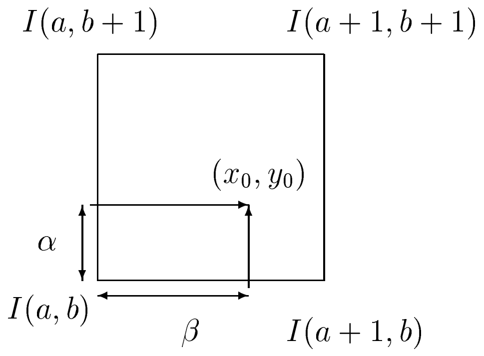
Interpolation af værdier ud fra diskrete stikprøver kan foretages på mange måder. Her skal nævnes tre metoder: *Nærmeste nabo interpolation* , *Bilineær interpolation* og *Kubisk interpolation* . Nærmeste nabo -metoden er den simpleste og giver ofte et dårligt resultat. Kubisk interpolation er mest kompliceret og giver kun sjældent bedre resultat end ved bilineær interpolation. Væsentlig i denne sammenhæng er, at interpolationspunkter aldrig ligger længere fra et målepunkt end 1/2 pixel, samt at vi ikke her ønsker at indbygge nogen form for støjfiltrering. Dette forhold gør mere avancerede interpolationsmetoder mindre relevente.

Nærmeste nabo interpolation Bestem $I_0(x_0, y_0)$ hvor x_0 og y_0 er relle ved $I_0(x_1, y_1)$, således at x_1 og y_1 er heltallige, og således at $|x_0 - x_1| \leq 1/2$ og $|y_0 - y_1| \leq 1/2$.

Bilineær interpolation Bestem $I_0(x_0, y_0)$ ud fra de fire omliggende naboer ved:

$$\begin{aligned} I_0(x_0, y_0) &= (1-\alpha)(1-\beta)I_0(a,b) + (1-\alpha)\beta I_0(a,b+1) + \\ &\quad \alpha(1-\beta)I_0(a+1,b) + \alpha\beta I_0(a+1,b+1) \end{aligned} \tag{1.45}$$

hvor $a = \lfloor x_0 \rfloor$, $b = \lfloor y_0 \rfloor$, og $\alpha = x_0 - a$ samt $\beta = y_0 - b$.



Figur 1.2: Skema til illustration af bilineær interpolation

Kubisk interpolation Ved kubisk interpolation bestemmes værdien ud fra 16 omgivende stikprøveværdier. Interpolationsformlen gives her i det endimensionelle tilfælde. Lad a og α være givet som ved den bilineære interpolation. Den interpolerede værdi kan da bestemmes ved:

$$\begin{aligned} I_0(x_0, y_0) &= \alpha(-1+2\alpha-\alpha^2)I_0(a-1) + (1-2\alpha^2+\alpha^3)I_0(a) + \\ &\quad \alpha(1+\alpha-\alpha^2)I_0(a+1) + \alpha^2(-1+\alpha)I_0(a+2) \end{aligned} \tag{1.46}$$

Det overlades til læseren at verificere formlen samt at udvide denne til det 2-dimensionelle tilfælde (Se også opgave 1.10).

1.4 Fouriertransformationen

Fouriertransformationen er klassisk inden for al signal- og billedbehandling. Dette skyldes dels historiske forhold, dels inspiration fra fysikken, og dels at en række foldninger kan beregnes væsentlig hurtigere ved brug af fouriertransformationen. Den fouriertransformerede af et signal beskriver frekvensgangen i signalet. Hvis signalet er en simpel sinusbølge vil den fouriertransformerede være to (symmetrisk positionerede) impulsfunktioner. Således giver fouriertransformationen mulighed for visuelt at fortolke signalet/billedet på en ny måde og for at specificere visse filtre lettere. Matematisk set er fouriertransformationen blot en ud af mange andre baseskiftende transformationer. De billedbehandlingsmæssigt mest interessante af disse er beskrevet i kapitel 4 af noterne. Som beskrevet senere kan det vises at der gælder en meget nyttig relation (foldningssætningen) mellem foldningen af to funktioner og deres fouriertransformerede.

1.4.1 Den kontinuerte fouriertransformation

Lad $f(x)$ være en kontinuert reel funktion af en reel variabel x. Antag endvidere at f er numerisk integrabel, dvs at $\int_{-\infty}^{\infty} |f(x)| dx$ er begrænset. Da er den fouriertransformerede af f defineret ved.

$$F\{f\} = \mathbf{f}(u) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i u x} dx \tag{1.47}$$

hvor $i = \sqrt{-1}$, og u er *frekvensvariablen* . Transformationen sender altså en funktion over i en anden funktion. Ofte skrives:

$$f \bowtie \mathbf{f} \tag{1.48}$$

hvor $\bowtie$ betyder "har den fouriertransformerede". Den fouriertransformerede er en kompleks funktion da integranden $f(x)e^{-2\pi i u x}$ involverer den komplekse eksponentialfunktion. Givet den fouriertransformerede $\mathbf{f}(u)$ kan $f(x)$ beregnes ved den *inverse fouriertransformation* :

$$F^{-1}\{\mathbf{f}\} = f(x) = \int_{-\infty}^{\infty} \mathbf{f}(u) e^{2\pi i u x} du \tag{1.49}$$

Da $\mathbf{f}(u)$ er kompleks kan den skrives som en sum af en realdel og en imaginærdel:

$$\mathbf{f}(u) = R(u) + iI(u) \tag{1.50}$$

Ofte beskrives $\mathbf{f}(u)$ dog på formen:

$$\mathbf{f}(u) = |\mathbf{f}(u)| e^{i\phi(u)} \tag{1.51}$$

hvor

$$|\mathbf{f}(u)| = \sqrt{R^2(u) + I^2(u)} \tag{1.52}$$

betegnes *fourierspektret*. Kvadratet på denne størrelse kaldes for *styrkespektret* for f.

$$\phi(u) \quad = \quad \arctan\{\frac{I(u)}{R(u)}\} \tag{1.53}$$

kaldes *fasevinklen* (eller blot fasen). At u kaldes for en frekvensvariabel er inspireret af *Eulers* formel :

$$e^{2\pi iux} \quad = \quad \cos(2\pi ux) \ + \ i\sin(2\pi ux) \tag{1.54}$$

Den to-dimensionale kontinuerte fouriertransformerede er defineret på tilsvarede måde. Således gælder:

$$F\{f\} \quad = \quad \mathbf{f}(u,v) \quad = \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-2\pi i(ux\,+\,vy)}dxdy \tag{1.55}$$

Eksempel

Lad:

$$f(x) \quad = \quad \begin{cases} H & \text{hvis } A \leq x \leq B \\ 0 & \text{ellers} \end{cases} \tag{1.56}$$

f er klart numerisk integrabel. Den fouriertransformerede er derfor:

$$\mathbf{f}(u) \quad = \quad \int_{-\infty}^{\infty} f(x)e^{-2\pi iux}dx \tag{1.57}$$

$$= \quad H\int_{A}^{B} e^{-2\pi iux}dx \tag{1.58}$$

$$= \quad \frac{-H}{2\pi iu}[e^{-2\pi iux}]_A^B \tag{1.59}$$

$$= \quad \frac{-H}{2\pi iu}(e^{-2\pi iuB} \ - \ e^{-2\pi iuA}) \tag{1.60}$$

I tilfældet $A = 0$ får vi:

$$\mathbf{f}(u) \quad = \quad \frac{-H}{2\pi iu}(e^{-2\pi iuB} \ - \ 1) \tag{1.61}$$

$$= \quad \frac{-H}{2\pi iu}[e^{-\pi iuB} \ - \ e^{\pi iuB}]e^{-\pi iuB} \tag{1.62}$$

$$= \quad \frac{H}{\pi u}\sin(\pi Bu)e^{-\pi iuB} \tag{1.63}$$

Den transformerede ses at være kompleks. Styrkespektret for f er bestemt ved:

$$|\mathbf{f}(u)| \quad = \quad HB\left|\frac{\sin(\pi Bu)}{\pi Bu}\right| \tag{1.64}$$

I tilfældet at $A = -B$ får vi:

$$\mathbf{f}(u) = \frac{-H}{2\pi iu}(e^{-2\pi iuB} - e^{2\pi iuB}) \quad (1.65)$$

$$= 2BH\frac{\sin(2\pi Bu)}{2\pi Bu} \quad (1.66)$$

Her ser vi at den transformerede er reel. Det kan vises at en hver lige funktion har en fouriertransformeret med imaginærkomponent lig nul. Omvendt gælder at enhver reel funktion har en fouriertransformeret, der er lige. Er f ulige ses tilsvarende at $\mathbf{f}$ vil være rent imaginær.

1.4.2 Den diskrete fouriertransformation

Antag at der er givet N stikprøver (samples) $(f(0), f(1), f(2), \ldots, f(N-1))$ udtaget med ækvidistant afstand fra en en-dimensional funktion f. Hver af de samplede værdier kan være reelle eller komplekse. Den *diskrete* fouriertransformation og dens inverse er da givet ved:

$$\mathbf{f}(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)e^{-2\pi iux/N} \quad (1.67)$$

$$f(x) = \sum_{u=0}^{N-1} \mathbf{f}(u)e^{2\pi iux/N} \quad (1.68)$$

for u hhv. x lig $0, 1, 2, \ldots, N-1$. Tilsvarende formler gælder i det to-dimensionale tilfælde:

$$\mathbf{f}(u,v) = \frac{1}{NM}\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f(x,y)e^{-2\pi i(ux/N\,+\,vy/M)} \quad (1.69)$$

$$f(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{M-1} \mathbf{f}(u,v)e^{2\pi i(ux/N\,+\,vy/M)} \quad (1.70)$$

for u og $x \in [0 \ldots N-1]$, og for v of $y \in [0 \ldots M-1]$.

Beviser for at ovenstående ligninger på fornuftig måde definerer den diskrete fouriertransformation ligger uden for rammerne af disse noter. Det bør bemærkes at der ikke i det diskrete tilfælde stilles krav om integrabilitet. Den diskrete fouriertransformation er altid defineret.

I billedbehandling benyttes der ofte kvadrariske billeder (N=M). I dette tilfælde er det sædvane at gøre den *forlæns* og den *baglæns* fouriertransformation symmetriske ved at splitte normaliseringskonstanten ($\frac{1}{N^2}$) op i to:

$$\mathbf{f}(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)e^{-2\pi i(ux\,+\,vy)/N} \quad (1.71)$$

$$f(x,y) = \frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} \mathbf{f}(u,v)e^{2i\pi(ux\,+\,vy)/N} \quad (1.72)$$

Som vi senere skal se eksisterer der en hurtig algoritme (FFT-algoritmen) til beregning af den fouriertransformerede. Denne metode kræver at såvel N som M er potenser af 2.

En grundlæggende antagelse (der dog kun har fortolkningsmæssig betydning) er at de N samples $(f(0), f(1), f(2), \ldots, f(N-1))$ er udtaget med konstant afstand Δx . Tilsvarende beskriver de N værdier af den fouriertransformerede en fortolkning hvor frekvensvariablen er samplet med en konstant afstand Δu. Som vi skal se senere i afsnittet om *samplingsætningen* gælder der følgende relation:

$$\Delta u = \frac{1}{N \Delta x} \tag{1.73}$$

For fast N ses at jo mindre samplingafstanden Δx er, jo større bliver Δu. Da der ialt er N samples af frevensvariablen u betyder dette, at højere frekvenser af signalet bliver repræsenteret. For fast Δx betyder færre/flere samples N blot færre/flere samples af frekvensvariablen, men ikke at det repræsenterede frekvensområde ændres.

1.4.3 Egenskaber ved fouriertransformationen

Fouriertransformationen har en lang række af egenskaber, der er vigtige at kende ved praktisk anvendelse. Af definitionen på transformationen følger direkte at denne er lineær:

$$F\{f+g\} = F\{f\} + F\{g\} \tag{1.74}$$
$$F\{\alpha f\} = \alpha F\{f\} \tag{1.75}$$

Desuden gælder der at:

$$f(\alpha x) \bowtie \frac{1}{|\alpha|}\mathbf{f}(u/\alpha) \tag{1.76}$$

Fouriertransformationen er separabel , idet der gælder:

$$F(u,v) = \frac{1}{NM} \sum_{x=0}^{N-1} e^{-2\pi iux/N} \left[\sum_{y=0}^{M-1} f(x,y) e^{-2\pi ivy/M} \right] \tag{1.77}$$

Dette betyder at en implementation af den to-dimensionale fouriertransformation blot kræver at vi kan foretage en en-dimensional fouriertransformation. For et kvadratisk billede (og ved brug af normaliseringskonstanten 1/N jvf. (1.72)) gælder således:

$$\mathbf{f}(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \hat{f}(x,v) e^{-2\pi iux/N} \tag{1.78}$$
$$\hat{f}(x,v) = N \left[\frac{1}{N} \sum_{y=0}^{N-1} f(x,y) e^{-2\pi ivy/N} \right] \tag{1.79}$$

En anden egenskab ved fouriertransformationen er at den komplekse eksponentialfunktion $e^{-2\pi iux/N}$ er periodisk med periode N . Dette betyder at:

$$\mathbf{f}(u,v) = \mathbf{f}(u+N,v) = \mathbf{f}(u,v+N) = \mathbf{f}(u+N,v+N) \tag{1.80}$$

Det er desuden let at vise at hvis f er reel da er den fouriertransformerede konjugeret symmetrisk, dvs.:

$$\mathbf{f}(u,v) = \mathbf{f}^{\star}(-u,-v) \tag{1.81}$$

Dette betyder som tidligere nævnt at:

$$|\mathbf{f}(u,v)| = |\mathbf{f}(-u,-v)| \tag{1.82}$$

Endelig skal det nævnes at den fouriertransformerede af et billede der er roteret θ grader i forhold til en reference, også vil være roteret θ grader i samme retning. Udtrykt i polære koordinater gælder altså:

$$f(r,\psi+\theta) \bowtie \mathbf{f}(\omega,\phi+\theta) \tag{1.83}$$

hvor (r,ψ) er tidsvariable og (ω,ϕ) er frekvensvariable.

Lokalisering og værdi i Origo for fourierspektret

Et praktisk problem ved visuel inspektion af den fouriertransformerede af et billede er at Origo for den betragtede periode vil være lokaliseret i såvel (0,0), (0,N), (N,0) og (N,N), hvor N (et multipla af 2) er størrelsen af det kvadratiske billede. Visuelt vil det være at foretrække at have Origo centreret i (N/2,N/2). Generelt er translationsegenskaberne for den fouriertransformerede givet ved:

$$f(x,y)e^{2\pi i(u_0x+v_0y)/N} \bowtie F(u-u_0,v-v_0) \tag{1.84}$$

hvor $\bowtie$ betyder "har den fouriertransformerede". Idet vi lader $u_0 = v_0 = N/2$, og bemærker at:

$$e^{2\pi i(xN/2+yN/2)/N} = e^{\pi i(x+y)} = (-1)^{x+y} \tag{1.85}$$

ses at translationen fremkommer ved fouriertransformation af $(-1)^{x+y}f(x,y)$.

Ved visuel inspektion af fourierspektret er vi ofte kun interesserede i at se styrkespektret $|F(u,v)|$. I praksis er en så stor del af energien i forurierspektret samlet i Origo, at ved en lineær skalering vil vi kun se Origo som en lys prik på en sort baggrund. Ofte benyttes derfor skaleringen:

$$V(u,v) = log(1+|\mathbf{F}(u,v)|) \tag{1.86}$$

hvorpå $V(u,v)$ skaleres lineært eller histogramlineariseres, sådan at alle intensitetsværdier udnyttes.

Den faktiske værdi af den fouriertransformerede i Origo, er relateret til middelintensiteten i billedet. Dette følger direkte af definitionen på den diskret fouriertransformerede:

$$\mathbf{F}(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)e^{-2\pi i(ux+vy)/N} \tag{1.87}$$

ved substitution af (u,v) med (0,0). Den fundne relation bliver:

$$\mathbf{F}(0,0) = N\bar{f}(x,y) \tag{1.88}$$

Som det ses giver formlen mulighed for direkte at angive den ønskede middelintensitet i et filtreret billede, blot ved at sætte værdien af den fouriertransformerede i (0,0).

Afledede

Som vi skal diskutere nøjere i kapitlet om kantdetektion, er det ofte man har brug for målinger af de partielle afledede af en intensitetsfunktion. Vi skal ikke her komme ind på de problemer der opstår hvis billedet indeholder støj etc., men betrage det idelle kontinuerte tilfælde hvor billedet antages C^∞ og reel. Da gælder regnereglen:

$$F\{\frac{\partial^n I}{\partial x^n}\} = (2\pi i)^n u^n F\{I\} \tag{1.89}$$

hvor $F\{I\}$ her angiver fouriertransformation af I. Ønsker vi eksempelvis at anvende laplaceoperatoren $(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2})$ på et billede I kan dette foregå ved transformationen: $F^{-1}\{\mathbf{h}F\{I\}\}$ hvor frekvensfiltret $\mathbf{h}$ er givet ved:

$$\mathbf{h}(u,v) = -4\pi^2(u^2+v^2) \tag{1.90}$$

Tilsvarende kan andre differentialoperatorer defineres i frekvensområdet. Det skal bemærkes at i praksis vil denne metode være problematisk idet billeder ofte indeholder højfrekvent støj, og at denne fremhæves markant ved, som i ovenstående eksempel, at multiplicere frekvensbilledet med en paraboloide.

Foldningssætningen

Som tidligere beskrevet kan en meget stor del af de klassiske (og nyere) metoder til billedfiltrering beskrives ved foldningen af billedet med et filter. En af de væsenligste grunde til at der er ofret så meget tid på fouriertransformationen er, at fouriertransformationen af en foldning er lig produktet af de fouriertransformerede:

$$f(x)\star g(x) \;\bowtie\; \mathbf{f}(u)\mathbf{g}(u) \tag{1.91}$$

og ækvivalent:

$$f(x)g(x) \;\bowtie\; \mathbf{f}(u)\star\mathbf{g}(u) \tag{1.92}$$

Dette resultat går under navnet *foldningssætningen*. Et bevis for sætningen ligger uden for rammerne af disse noter. I det diskrete tilfælde gælder tilsvarende sætninger. Ved anvendelse af foldningssætningen er det nødvendigt at sikre sig mod *warparound error* ved at udvide signalerne/billederne passende. I praksis er så meget af energien i et billede samlet i og umiddelbart omkring Origo, at udvidelserne ikke foretages. Dette vil resultere i en fejl, der dog ofte er så lille at den kan ignoreres.

Foldningssætningen siger altså, at vi kan formulere filtreringen i tidsområdet eller i frekvensområdet efter behag. Dette skal vi udnytte i de følgende afsnit. En anden fordel som sætningen giver er, at visse filtreringer kan beregnes mere effektivt i frekvensområdet. Dette bygger på at FFT-algoritmen (se senere) har en kompleksitet på $O(N \log N)$ for et signal på N elementer. Implementeres foldningen af de to signaler direkte fås en tidskompleksitet på $O(N^2)$. Ved først af fouriertransformere de to funktioner, derefter at multiplicere resultaterne element for element, og sluttelig invers fouriertransformere produktet fås igen en kompleksitet på $O(N \log N)$. For store værdier af N giver en implementation v.h.a. foldningssætningen altså mulighed for en "speed-up". Denne kan være betragtelig, især hvis den fouriertransformerede af filteret kan beregnes på forhånd.

Eksempel

Antag et billede af dimension $M \times M$, hvor M er en potens af 2, og et filter med støtte $N \times N$. Implementeres foldningen direkte er det totale antal multiplikationer lig N^2M^2. Benyttes fouriertransformationen (FFT-algoritmen) først på billedet samt det udvidede filter, og til slut på resultatet af den elementvise multiplikation af de to transformerede billeder, fås ialt $M^2(6\ \log(M) + 1)$ multiplikationer (Overvej dette).

Korrelationssætningen

Korrelation af to kontinuerte relle funktioner (med begrænset numerisk integral) er defineret ved:

$$f(x) \circ g(x) \;=\; \int_{-\infty}^{\infty} f^{\star}(\alpha) g(x \;+\; \alpha) d\alpha \tag{1.93}$$

Som det ses ligner foldninger og korrelationer hinanden meget. Forskellen er kun, at i korrelationen er funktionen f komplekst konjugeret og minusset i translationen ændret til et plus. De fleste af sætningerne om foldningen og sammenhæng mellem foldning og fouriertransformation kan gentages her. Således gælder korrelationssætningen:

$$f(x) \circ g(x) \bowtie \mathbf{f}^{\star}(u)\mathbf{g}(u) \tag{1.94}$$

og ækvivalent:

$$f^{\star}(x)g(x) \bowtie \mathbf{f}(u) \circ \mathbf{g}(u) \tag{1.95}$$

Korrelationer kan, som foldninger, i visse tilfælde implementeres effektivt ved brug af fouriertransformationen.

Korrelationer anvendes meget i billedbehandling til at finde "match" mellem en given prototype (eller 'template") og et forelagt billede. I det klassiske eksempel er prototypen et billede af en tank. Typisk er prototype-billedet lille (f.eks 32×32). Vi ønsker nu, forelagt et (stort) billede at finde positionerne af de muligt afbildede tank's i dette. Dette gøres ved at detektere de lokale maksimum i resultatet af korrelationen mellem prototypen og billedet (begge udvidet passende, således som det tidligere er beskrevet).

Hvis prototypen (som i ovenstående eksempel) er reel, da adskiller korrelationen sig kun fra foldningen ved den manglende spejling. I eksemplet fra afnit 1.1.1 om foldninger er det let at se, at maksimum i korrelationsfunktionen (korrelationen mellem f og g) vil antages i $x = 0$, og ikke som ved foldningen i $x = 1$. Da funktionerne fra dette eksempel er symmetriske vil formen af korrelationsfunktionen være ens med resultatet af foldningen, dvs. en "trekantsfunktion".

1.4.4 Samplingsætningen

Samplingsætningen omhandler hvor mange samples (stikprøver) af en kontinuert funktion vi skal tage for at sikre at der ikke går information tabt i samplingen. Eksempelvis kræver en konstant funktion kun én sampel. I det følgende beskrives (vha. tegneserier) hvad der sker ved sampling af en en-dimensional funktion. Resultaterne for det to-dimensionale tilfælde er analoge.

Lad $f(x)$ være en funktion, der antages defineret på hele den relle akse. $f(x)$ kaldes for *båndbegrænset* med båndbredde w , hvis den fouriertransformerede $\mathbf{f}(u)$ er konstant lig 0 uden for frekvensintervallet $[-w, w]$. Figur 1.3 illustrerer i billede a) og b) en båndbegrænset funktion. Bemærk at der i figuren benyttes store bogstaver, f.eks. $F(u)$ om den fouriertransformerede $\mathbf{f}(u)$.

Den fouriertransformerede af et impulstog

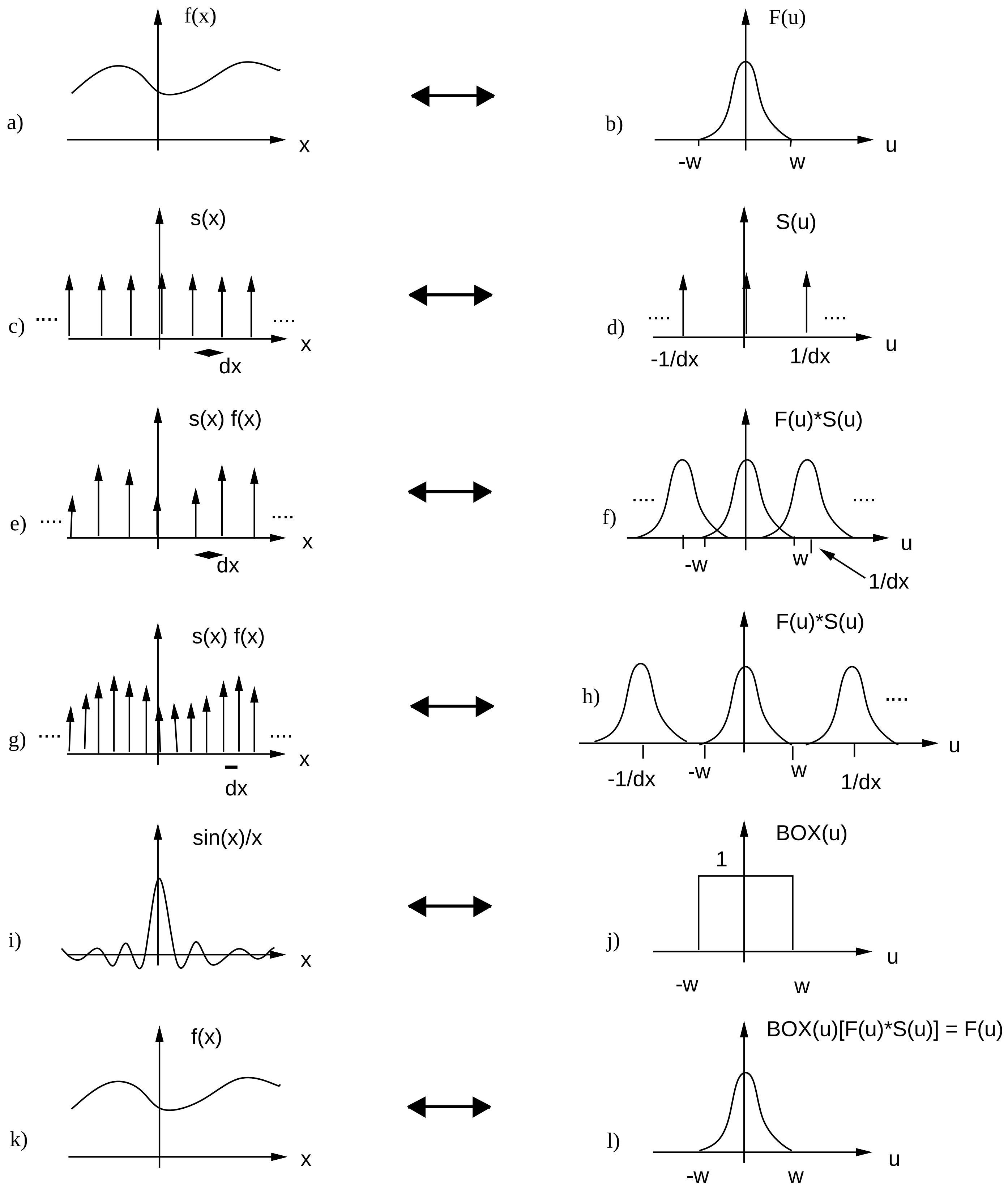
I Figur 1.3 billede c) er vist et uendeligt impulstog $s(x)$. Ved multiplikation af $f(x)$ med dette tog fås samplingen vist i billede e). Billederne på højresiden af figuren viser de fouriertransformerede af funktionerne vist i billederne på venstresiden af figuren. Billede c) og d) illustrerer at den fouriertransformerede $S(u)$ af et uendeligt impulstog $s(x)$ også er et impulstog med impulsafstand, der omvendt proportional med impulsafstanden i $s(x)$, dvs:

$$s(t) \;=\; \sum_{n=-\infty}^{\infty} \delta(t - nT) \quad \bowtie \quad S(u) \;=\; \frac{1}{T} \sum_{n=-\infty}^{\infty} \delta(u - \frac{n}{T}) \tag{1.96}$$

Dette resultat er ikke oplagt. Et bevis falder uden for rammerne af disse noter. Dog kan følgende løse argumenter sandsynliggøre resultatet: Betragt funktionen S_1 bestående af tre impulsfunktioner lokaliseret i $-1/T$, i 0 og i $1/T$. Af definitionen på fouriertransformationen følger direkte at den fouriertransformerede s_1 af S_1 vil være summen af en konstant (fra den midterste impulsfunktion) og en cosinus (fra de to symmetrisk positionerede impulsfunktioner).

$$s_1(u) \;=\; 1 + 2\cos(2\pi u/T) \tag{1.97}$$

Denne funktion har lokale maksima for $u = pT$, hvor p er et heltal. Konstruer nu S_2, S_3 etc. ved at tilføge impulsfunktioner i $-2T$ og $2T$, i $-3T$ og $3T$ etc. Herved adderer vi til den



Figur 1.3: Grafisk illustration af effekten af en uendelig sampling

fouriertransformerede nye cosinusfunktioner med dobbelt, tredobbelt etc. frekvens. I maksimumspunkterne har cosinusfunktionerne med dobbelt, tredobbelt etc. frekvens også lokalt maksimum. I intervallerne mellem de lokale maksima vil funktionerne svinge hurtigere og hurtigere, og vil have en tendens til af ophæve hinanden. I grænsen fås:

$$s_\infty(u) \;=\; \lim_{K\to\infty}\left[1 \;+\; 2\sum_{n=1}^{K}\cos(2\pi nu/T)\right] \;=\; \sum_{n=-\infty}^{\infty}\delta(t-nT) \tag{1.98}$$

Funktionerne s_3, s_7 og s_{11} er vist nedenfor i Figur 1.4.

Vi kan nu fortsætte med beskrivelsen af samplingsætningen. Multiplikationen af $f(x)$ med impulstoget s svarer i fourierdomænet til en foldning af $F(u)$ med $S(u)$. Denne foldning kan, som tidligere beskrevet, betragtes som en kopiering. Herved fremkommer billede f). Som det ses er afstanden dx mellem samples $s(x)$ valgt stor. Dette betyder af afstanden mellem de tilsvarende impulsfunktioner af den fouriertransformerede er lille. Herved kommer de enkelte perioder af foldningsresultatet til at overlappe. Dette fænomen kaldes *aliasing* . Aliasing vil umuliggøre en eksakt rekonstruktion af den kontinuerte funktion $f(x)$. For at undgå aliasing må der oplagt kræves (overvej dette) at:

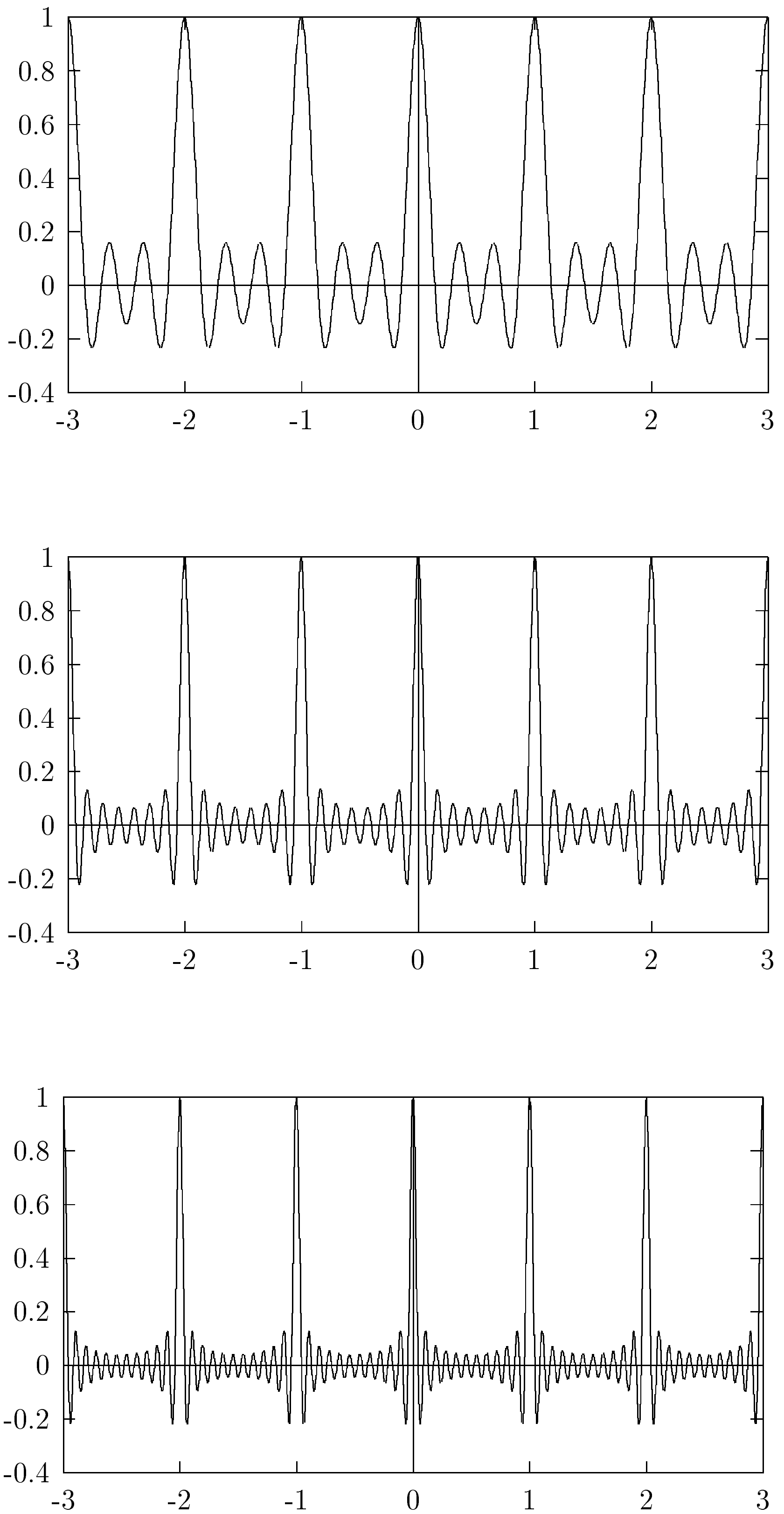
$$w \;\leq\; \frac{1}{2dx} \tag{1.99}$$

En situation hvor dx er valgt sådan at ovenstående ulighed er opfyldt er vist i billede g) og h). Her er de enkelte perioder af foldningsresultatet vel separeret. Billede j) og l) viser hvorledes det ved multiplikation af foldningsresultatet med et box-filter er muligt at ekstrahere præcis en periode. Ved invers fouriertransformation af denne funktion fås den oprindelige *kontinuerte* funktion $f(x)$. Vi kan derfor opsummere:

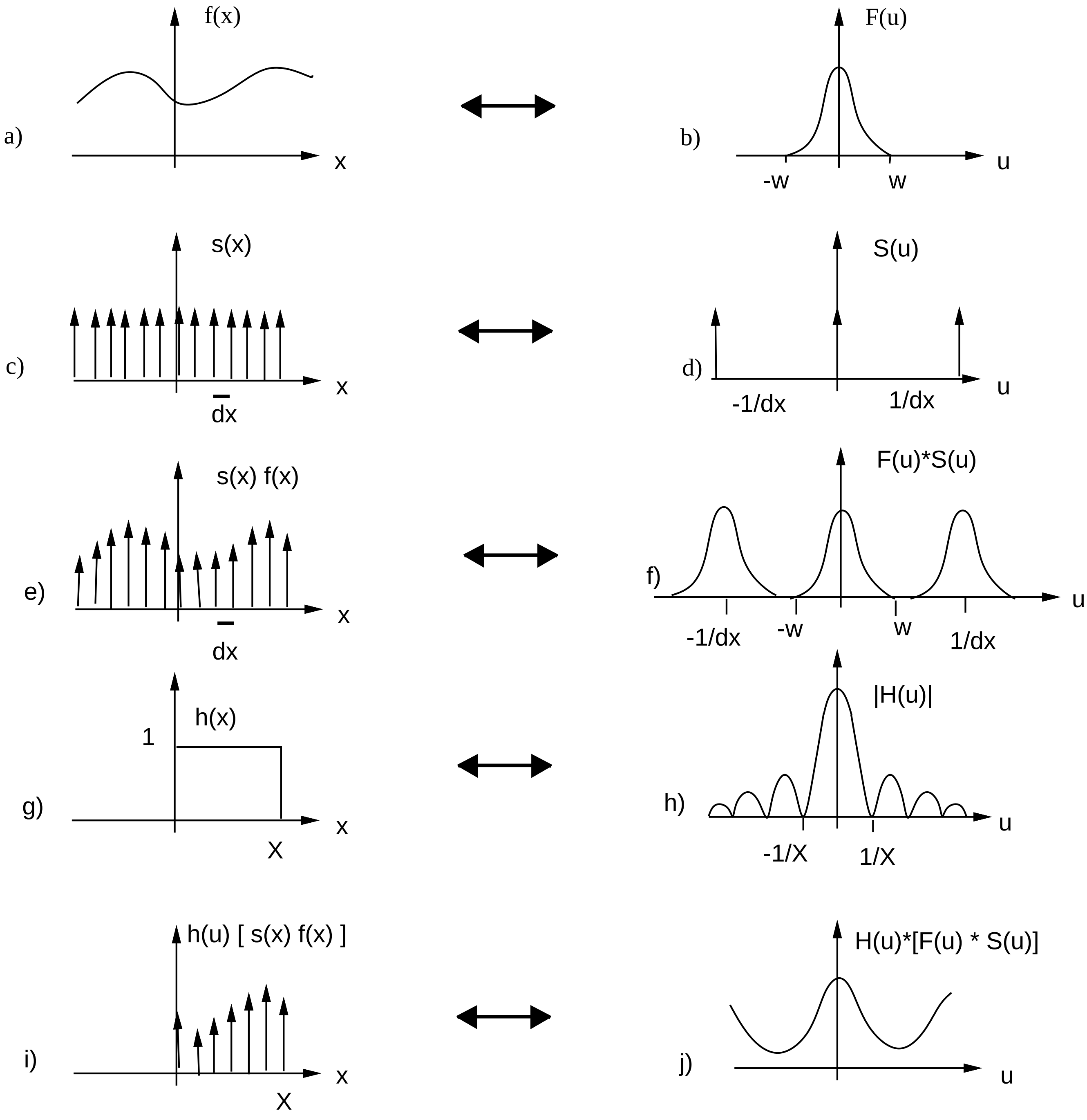
> Antag at den kontinuerte funktion $f(x)$ er defineret for alle reelle x og er båndbegrænset med båndbredde w. Lad $g(x)$ være en sampling af f med samplingafstand $dx \leq 1/(2w)$. Da kan den kontinuerte funktion f rekonstrueres eksakt ud fra de diskrete samples ved at invers fouriertransformere produktet af et box-filter med støtte $[-w, w]$ og den fouriertransformerede af g.

Sampling af endelige funktioner

Antag at $f(x)$ er defineret på hele den reelle akse, og at $f(x)$ er båndbegrænset med båndbredde w, og betragt en sampling således at samplingsætningen er opfyldt. Dette er illustreret i Figur 1.5 i billede a) til f). I billede g) og h) af Figur 1.5 ses hvorledes et endeligt antal samples kan ekstraheres ved multiplikation med et vindue. Som tidligere beskrevet svarer dette (i fourierdomænet) til en foldning med den fouriertransformerede af vinduesfunktionen. Billede j) viser resultetet af denne foldning. Da $H(u)$ har frekvensværdier der fortsætter mod uendelig forvrænges frekvensspektet i forhold til det viste i b). Selv om samplingen opfylder betingelserne for samplingsætningen er det derfor ikke længere muligt at rekonstuere f, fordi det oprindelige fourierspektrum ikke kan rekonstrueres. Bemærk at



Figur 1.4: Funktionerne $v_3(u)$, $v_7(u)$, og $v_{11}(u)$



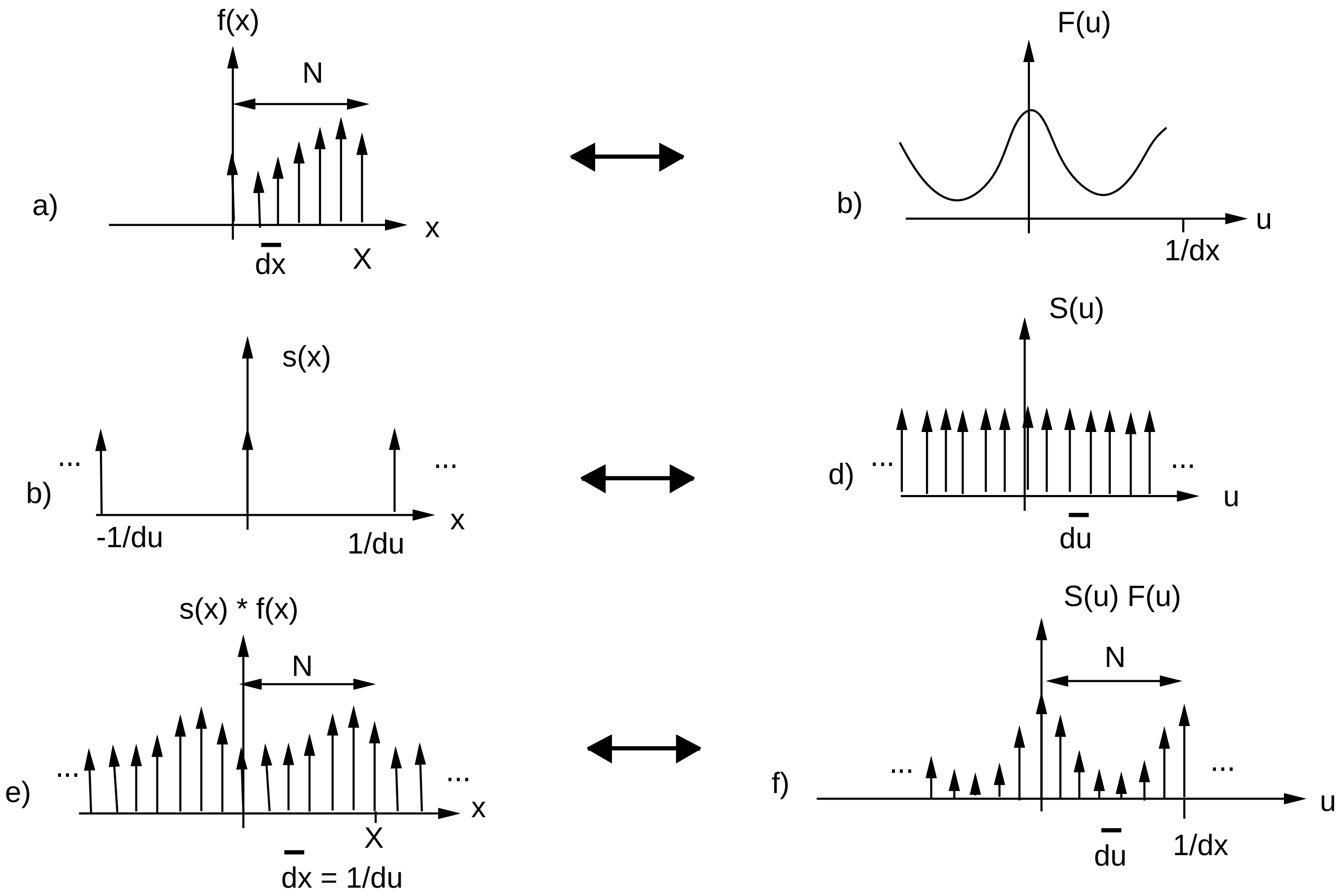
Figur 1.5: Grafisk illustration af effekten af en endelig sampling

ved invers fouriertransformation fås en funktion, der er defineret på hele den reelle akse og som *ikke* er lig nul uden for det område hvori f oprindeligt er samplet. Således gælder der at:

> Der eksisterer ingen funktioner med endelig støtte som er båndbegrænsede . Omvendt gælder at enhver båndbegænset funktion har uendelig støtte.

Periodicitet af den diskrete fouriertransformation

I den foregående diskussion har alle funktioner i frekvensområdet været kontinuerte. For at opnå en diskret repræsentation af en funktion i frekvensområdet multiplicerer vi blot denne med et impulstog. Figur 1.6 illustrerer (som en fortsættelse af Figur 1.5) hvorledes denne sampling foregår. I x-domænet svarer multiplikationen til en foldning med et impul-



Figur 1.6: Grafisk illustration af den diskrete fouriertransformation

stog med impulsafstand på $1/du$. Derved bliver foldningen vist i billede e) periodisk med periode $1/du$. Dette forudsætter naturligvis at du er valgt tilstrækkelig lille (jvf. samplingsætningen). Funktionen f såvel som dens diskret fouriertransformerede er repræsenteret ved

N samples. Disse er udtaget ækvidistant således at hver periode er dækket. Dette betyder at $Ndx = X$ i x-domænet og tisvarende at $Ndu = 1/dx$ i frekvensdomænet . Der gælder altså at:

$$du = \frac{1}{Ndx} \tag{1.100}$$

1.4.5 FFT-algoritmen

Betragt den diskrete en-dimensionale fouriertransformation:

$$\mathbf{f}(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)W_N^{ux} \tag{1.101}$$

hvor N er antal samples, hvor:

$$W_N = e^{-2\pi i/N} \tag{1.102}$$

og hvor faktoren W_N^{ux} skal læses "W_N ophævet til potensen ux". For hvert af de N værdier af u omfatter transformationen N komplekse multiplikationer af $f(x)$ med $W_N^{ux} = e^{-2\pi ixu/N}$ og $N-1$ additioner. Multiplikationer mv. til udregning af de N værdier af den komplekse eksponentialfunktion medregnes ikke, da disse kan tabelleres på forhånd. Den endimensionale fouriertransformation har derfor en tidskompleksitet på $O(N^2)$.

I praksis beregnes fouriertransformationer ved anvendelse af *den hurtige fouriertransformation*: FFT-algoritmen. Det kan vises at antallet af (komplekse) multiplikationer hhv. additioner for denne metode er $\frac{N}{2}\log N$ hhv. $N\log N$. I det følgende skitseres kernen af den FFT-algoritme, der er baseret på den såkaldte "succesive dobling"-metode. Dette er den mest generelle og mest benyttede FFT-algoritme. En grundigere behandling beskrivelse findes f.eks. i *E. Oran Brigham: The Fast Fourier Transform and its Applications; Prentice-Hall International Editions, 1988.*

En begrænsende antagelse for FFT-algoritmen er at antallet af samples er en potens af 2, dvs: $N = 2^n$, hvor n er et positivt heltal. I dette tilfælde gælder at $N = 2M$. Jævnfør ovenstående definition af den diskrete fouriertransformation gælder der derfor:

$$\mathbf{f}(u) = \frac{1}{2M}\sum_{x=0}^{2M-1} f(x)W_{2M}^{ux} \tag{1.103}$$

$$= \frac{1}{2}\left[\frac{1}{M}\sum_{x=0}^{M-1} f(2x)W_{2M}^{u(2x)} + \frac{1}{M}\sum_{x=0}^{M-1} f(2x+1)W_{2M}^{u(2x+1)}\right] \tag{1.104}$$

$$= \frac{1}{2}\left[\frac{1}{M}\sum_{x=0}^{M-1} f(2x)W_{M}^{ux} + \frac{1}{M}\sum_{x=0}^{M-1} f(2x+1)W_{M}^{ux}W_{2M}^{u}\right] \tag{1.105}$$

hvor vi blot har opdelt summen i to led efter lige og ulige værdier af argumentet til f, og udnyttet at $W_{2M}^{2ux} = W_M^{ux}$. Det er nu oplagt at definere:

$$\mathbf{f}_{lige}(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(2x)W_M^{ux} \tag{1.106}$$

$$\mathbf{f}_{ulige}(u) \quad = \quad \frac{1}{M}\sum_{x=0}^{M-1} f(2x+1)W_M^{ux} \tag{1.107}$$

for $u = 0, 1, 2, \ldots, M-1$. Med disse forkortelser kan vi skrive:

$$\mathbf{f}(u) \quad = \quad \frac{1}{2}\left[\mathbf{f}_{lige}(u) \; + \; \mathbf{f}_{ulige}(u)W_{2M}^{u}\right] \tag{1.108}$$

$$\mathbf{f}(u+M) \quad = \quad \frac{1}{2}\left[\mathbf{f}_{lige}(u) \; - \; \mathbf{f}_{ulige}(u)W_{2M}^{u}\right] \tag{1.109}$$

hvor minusset i den sidste ligning fås fordi $W_{2M}^{M} = -1$. Formlerne beskriver en effektiv rekursiv metode til beregning af fouriertransformationen, idet en N-samples transformation opdeles i to grupper. I den første gruppe beregnes $\mathbf{f}(0)$ til $\mathbf{f}(N/2-1)$ ved evaluering af to N/2-samples transformationer. Værdierne af $\mathbf{f}(N/2)$ til $\mathbf{f}(N-1)$ fås direkte vha. ovenstående formel uden brug af yderligere transformationer.

Eksempel

En to-punkts fouriertransformation (dvs. $n = 1$) kræver beregning af $\mathbf{f}(0)$, dvs. $\mathbf{f}_{lige}(0)$ og $\mathbf{f}_{ulige}(0)$. Da begge disse er 1-punkts transformationer, og da fouriertransformationen af et enkelt værdi er lig værdien selv kræver bestemmelsen af disse størrelser ingen multiplikationer eller additioner. Derefter skal $\mathbf{f}_{ulige}(0)$ multipliceres med W_2^0. Heraf ses at det totale antal af multiplikationer hhv. additioner er $mult(n=1) = 1$ og $add(n=1) = 2$.

For $n = 2$ opdeles de fire værdier i to grupper som før. Den første gruppe kræver udregning af to 2-punkts transformationer. Dette kræver $2mult(1)$ hhv. $2add(1)$ multiplikationer hhv. additioner. Dertil kommer to multiplikationer og to additioner til at beregne $\mathbf{f}(0)$ og $\mathbf{f}(1)$. Endelig kræves to additioner til beregning af $\mathbf{f}(2)$ og $\mathbf{f}(3)$. Det totale antal operationer bliver således: $mult(2) = 2mult(1) + 2 = 4$ og $add(n=2) = 2add(1) + 4 = 8$.

Ved at fortsætte med $n = 3$ etc. er det let at indse at antallet at multiplikationer hhv. additioner for $N = 2^n$ (hvor $n \geq 1$) punkter bliver:

$$mult(n) \quad = \quad 2mult(n-1) \; + \; 2^{n-1} \tag{1.110}$$

$$add(n) \quad = \quad 2add(n-1) \; + \; 2^{n} \tag{1.111}$$

De tidligere nævnte størrelser for tidskompleksiteten af algoritmen følger nu let ved induktion.

Den inverse FFT

Den forlæns og baglæns en-dimensionale fouriertransformation adskilder sig (ud over en division med N) kun ved fortegnet på den komplekse eksponentialfunktion W_N. I overstående udledning af FFT-algoritmen antog vi at disse tal var tabbeleret på forhånd. For at kunne genbruge tabellerne i IFFT-algoritmen er det derfor oplagt at kompleks konjugere begge sider af ligningen (for definitionen af den inverse transformation). Ved tilsvarende at dividere med N fås:

$$\frac{1}{N} f^{\star}(x) = \frac{1}{N} \sum_{u=0}^{N-1} \mathbf{f}^{\star}(u) W_N^{ux} \tag{1.112}$$

Vi kan derfor benytte FFT-algoritmen til at beregne den inverse fouriertransformation. Blot skal vi før hhv. efter anvendelsen kompleks kongugere $\mathbf{f}(u)$ hhv. $f(x)$. Til slut skal resultatet multipliceres med N.

Det skal bemærkes at en helt tilsvarende måde kan benyttes i det todimensionale tilfælde. En forskel er dog, at der ikke her skal multipliceres med N til sidst. Implementeres den to-dimensionale transformation ved gentagende passager af den en-dimensionale transformation er det vigtigt kun at foretage kompleks konjugeringen før og efter *hele* transformationen, og altså ikke efter transformationen af hver række eller søjle. Yderligere skal det bemærkes at $f(x)$ ofte vides at være reel, hvorfor den afsluttende kompleks konjugering er overflødig.

Bit-spejlings-metoden

Antag at vi benytter FFT-algoritmen med succesiv dobling på et array med 8 punkter: $\{f(0), f(1), f(2), \ldots, f(7)\}$ Af beskrivelsen af FFT-algoritmen fremgår at der (for alle værdier af u) gælder at f-værdier med lige index indgår i beregningen af $\mathbf{f}_{lige}$ og at f-værdier med ulige index indgår i beregningen af $\mathbf{f}_{ulige}$. Inddatasættet kan altså opdeles i to grupper: $\{f(0), f(2), f(4), f(6)\}$ og $\{f(1), f(3), f(5), f(7)\}$. Tilsvarende gælder om hver af de 4-punkts transformationer der skal herefter skal benyttes. Dette giver opdelingen:

$$\{f(0), f(4)\}, \{f(2), f(6)\}, \{f(1), f(5)\}, \{f(3), f(7)\}$$

For at gøre FFT-algoritmen hurtig er det en fordel om array'et af inddataværdier kunne omordnes efter ovenstående skema, idet operationerne til indeksering herved kan spares. Heldigvis viser det sig, at der gælder en simpel regel (*the bit-reversal rule*) for hvorledes inddataværdierne skal omordnes. Dette er illustreret i Tabel 1.1. Det ses at for gyldigt index x fås det korrekte index i det omordnede array ved at udtrykke x binært og opskrive disse i omvendt rækkefølge. Yderligere diskussion af denne, og andre metoder, der er nyttige ved implementation af FFT-algoritmen findes mange andre steder f.eks. *E. Oran Brigham: The Fast Fourier Transform and its Applications; Prentice-Hall International Editions, 1988.*

1.5 En simpel model for billedforringelse

Ved opstilling af modeller for billedforringelse og for filtrering af billeder er det hensigtsmæssigt at formulere modellerne for kontinuerte billeder. Dette muliggør anvendelse af den sædvanlige mål- og integralteori. I det følgende antages derfor at det observerede billede I er en kontinuert todimensional reel funktion. Vi vil antage at I fremkommer ved forvrængning af en ideel intensitetsfunktion I_0 efterfulgt af en addition af en støjfunktion η:

Oprindeligt array	Oprindeligt binært indeks	Omordnet binært indeks	Omordnet array
f(0)	0 0 0	0 0 0	f(0)
f(1)	0 0 1	1 0 0	f(4)
f(2)	0 1 0	0 1 0	f(2)
f(3)	0 1 1	1 1 0	f(6)
f(4)	1 0 0	0 0 1	f(1)
f(5)	1 0 1	1 0 1	f(5)
f(6)	1 1 0	0 1 1	f(3)
f(7)	1 1 1	1 1 1	f(7)

Tabel 1.1: Omordning af array efter "bit-reversal" reglen

$$I = HI_0 + \eta \tag{1.113}$$

hvor H er en operator på mængden af kontinuerte 2-dimensionale reelle funktioner, og η er en endelig reel funktion. Det antages altså at støjen er uafhængig af billedet I_0 og af forvrængningen H.

Antag videre at operatoren H er lineær og positionsinvariant. Hvis støjen η et øjeblik antages lig nulfunktionen kan vi skrive:

$$I(x,y) = H \int\int I_0(i,j)\delta(x-i, y-j)didj \tag{1.114}$$

hvor δ er den sædvanlige deltafunktion (impulsfunktion). Da H er lineær og $I_0(i,j)$ er uafhængig af x og y fås:

$$I(x,y) = \int\int I_0(i,j)H\delta(x-i, y-j)didj \tag{1.115}$$

Fordi H er positionsinvariant kan vi skrive:

$$I(x,y) = \int\int I_0(i,j)h(x-i, y-j)didj \tag{1.116}$$

hvor $h(x,y) = H\delta(x,y)$. Funktionen h kaldes forvrængningens impulsrespons (alternative termer er impulssvar, eller punktspredningsfunktionen for forvrængningen). Det ses at hvis forvrængningen er positionsinvariant, da der denne fuldstændigt karakteriseret ved svaret på en impuls i punktet (0,0).

Af ovenstående ligning ses, at I kan skrives som foldningen af I_0 med h. I det generelle tilfælde hvor der optræder additiv støj får vi:

$$I = h * I_0 + \eta \tag{1.117}$$

Ovenstående model er, på trods af dens simple form, fundamentet for en lang række metoder til filtrering af billeder. Generelt er opgaven at finde en operator G således at $I_0 \approx GI$.

Dette problem benævnes gendannelsesproblemet (eng. restoration). I praksis kræves en vis viden om støjen η og om punktspredningsfunktionen h for at kunne foretage en approksimativ gendannelse. En eksakt gendannelse er sjældent mulig. F.eks. kan forringelsen af billedet (og kvantificeringen af billedværdier til heltal) have fjerne visse detalier helt. I det simple tilfælde hvor h er lig en impulsfunktion (dvs. $I = I_0 + \eta$), tales om støjfiltrering . I det følgende beskrives nogle få forvrængningstyper og nogle typiske eksempler på støj. I afsnittet om gendannelse behandles dette problem mere dybtgående.

1.5.1 Forvrængning

Et typisk eksempel på en forvrængning af et billede opstår ved fotografering af en scene gennem et røgfyldt eller tåget medium. Dette kan være tilfældet ved satellitbilleder, hvor et tyndt skylag bevirker en spredning af de reflekterede lysstråler. Et andet eksempel er et billede optaget med et kamera der ikke er fokuseret på scenen. I begge tilfælde vil det observerede billede være uskarpt (mere eller mindre udtværet) i forhold til det ideelle billede. I sådanne tilfælde vil forvrængningen ofte være rotationssymetrisk, og monotont aftagende. En prototype på en sådan punktspredningsfunktion er den to-dimensionale gaussfunktion.

I nogle tilfælde vil punktspredningsfunktionen ikke være rotationssymetrisk. Et eksempel er en afbilding af et objekt i bevægelse, f.eks. et billede af en bil i fart. Hvis eksponeringstiden er for stor, vil billedet af bilen være udtværet i retningen givet ved den projicerede bevægelse.

Generelt er det meget svært at bestemme punktspredningsfunktionen. Indeholder det ideelle billede et sort punkt på en hvid baggrund (eller omvendt) kan funktionen approximeres ved en skaleret udgave af det observerede punkt. Denne teknik kan også anvendes hvis I_0 indeholder en ideel trappetrinskant. Her kan vi dog kun måle forvrængningen vinkelret på kanten, og kun indirekte ved et integral af punktspredningsfunktionen (overvej dette). Som beskrevet tidligere er det imidlertid ikke let at gendanne det ideelle billede selv om punktspredningsfunktionen er kendt med stor nøjagtighed. I det følgende vil vi, med mindre andet nævnes, antage at det billede som vi vil filtrere ikke har været udsat for nogen forvrængning under optagelsen.

1.5.2 Støj

Overordnet antages, som tidligere beskrevet, at støjen er uafhængig af det ideelle billede og af en evt. forvrængning. Desuden antages støjen at være en stokastisk process hvor støjkomponenterne i to billedpunkter er uafhængige. Ofte er støjen additiv, men visse sensorer (f.eks. radar) tilfører billederne støj, der er multiplikativ. Hvis såvel middelværdien som variansen af støjen er konstant over billedet siges støjen at være *stationær*. Ofte antages at middelværdien af støjen er lig nul. I så tilfælde siges støjen at være hvid.

Væsentlige er også støjens statistiske egenskaber. Tre ofte anvendte støjtyper er:

1. Uniformt fordelt støj. Alle støjværdier er lige sandsynlige.

2. Gaussisk -eller normalfordelt støj. Vil typisk være rimelig ved støj introduceret i elektriske kredsløb, fordi denne kan beskrives som summen af mange små ensfordelte støjbidrag (jvf. de store tals lov i sandsynlighedsregningen vil summen være normalfordelt uafhængig af fordelingen af bidragene).

3. Salt og pebber støj. Støjen kan beskrives ved overskrivning af en vis procentdel af de oprindelige billedværdier med den mindst hhv. den størst mulige billedværdi. Denne støjtype kan f.eks. forekomme i slidte/gamle kameraer, hvor nogle celler er "faldet ud" og andre giver værdi som ved en overeksponering.

Det er meget vigtigt ved konstruktion af metoder til bortfiltrering af støj, og til billedbehandling generelt, at undersøge støjens natur, således at den korrekte model for forringelsen kan opstilles. Omvendt er det vigtigt, ved anvendelse af et filter, at gøre sig klart hvilke antagelser om støjens karakter filtret forudsætter, samt hvilken støj der aktuelt er til stede i billedet.

Mængden af støj i et signal eller billede måles ofte ved *signal-støjforholdet*, SNR . Dette er defineret ved energien i billedet divideret med energien i støjen:

$$SNR_{ms} = \frac{\int\int I_0^2 dxdy}{\int\int \eta^2 dxdy} \tag{1.118}$$

Ofte benyttes SNR_{rms} (eng. root mean square signal to noise ratio) defineret ved kvadratroden af SNR_{ms}-størrelsen. Under antagelsen om at støjen er hvid er det let at vise at variansen af støjen er lig middelenergien i støjen.

$$\sigma_\eta^2 = E(\eta^2) - E^2(\eta) = E(\eta^2) = \frac{1}{||A||}\int_A \eta^2 dxdy \tag{1.119}$$

Ligeledes kan man vise at styrkespektret for hvid støj (i teorien) er en konstant funktion hvis værdi netop er lig variansen.

Et andet mål er PSNR (*Peak signal to noise ratio*), defineret ved:

$$PSNR = 20\ \log_{10}(\frac{\mathrm{Max}}{SNR_{rms}}) \tag{1.120}$$

hvor Max er den størst mulige værdi (ofte 255.0). PSNR måles i decibel (dB). Bemærk at en forøgelse af PSNR på 20 dB svarer til en 10 gange mindre værdi af rms-værdien.

Der findes en række metoder til estimation af mængden af støj i billeder. Estimationen er vanskelig fordi det problematisk at skelne mellem egentlig støj og fin tekstur i billederne. Metoderne vil derfor ofte overestimere mængden af støj. En nærmere diskussion af metoder til støjestimation falder uden for rammerne af disse noter. En kort beskrivelse af nogle metoder samt en eksperimentel sammenligning af disse kan findes i *Olsen S. Estimation of Noise in Images: An Evaluation; DIKU-rapport 91/10.*

1.6 Lineære filtre

Lineære filtre er meget benyttede af to årsager. Dels er den matematiske teori bag lineære filtre veludviklet. Dette letter bevisførelsen for at et givet filter er optimalt i en given sammenhæng og muliggør teoretiske vurderinger af filtret. Dette gælder specielt vurderinger af filtrets egenskaber i frekvensområdet. Dels kan filtreringen beskrives ved foldning med punktspredningsfunktionen for filtret, hvilket kan udnyttes i specielt "foldnings-hardware". I nedenstående afsnit beskrives den lineære filtrering i frekvens- og intensitetsområdet.

1.6.1 Filtrering i frekvensområde

Det kan ofte være en fordel at specificere filtreringen i frekvensområdet , fordi visse støjtyper har en markant frekvenskarakteristik, eller fordi formålet med filtreringen netop er at fremhæve eller dæmpe visse frekvensområder. Det matematiske grundlag for frekvensfiltrering er foldningssætningen (se tidligere) . Frekvensfiltreringen foretages ved at øge eller dæmpe bestemte frekvenser svarende til værdierne af frekvensfilteret $\mathbf{h}$. I de følgende afsnit beskrives en række klassiske filtre $\mathbf{h}$ hvis formål er at dæmpe de høje frekvenser (lade de lave frekvenser passere), at dæmpe de lave frekvenser (lade de høje frekvenser passere), og at dæmpe udvalgte frekvenser (lade et bånd af frekvenser passere). Vi vil udelukkende betragte filtre der ikke berører fasen af billederne, men kun styrkespektret.

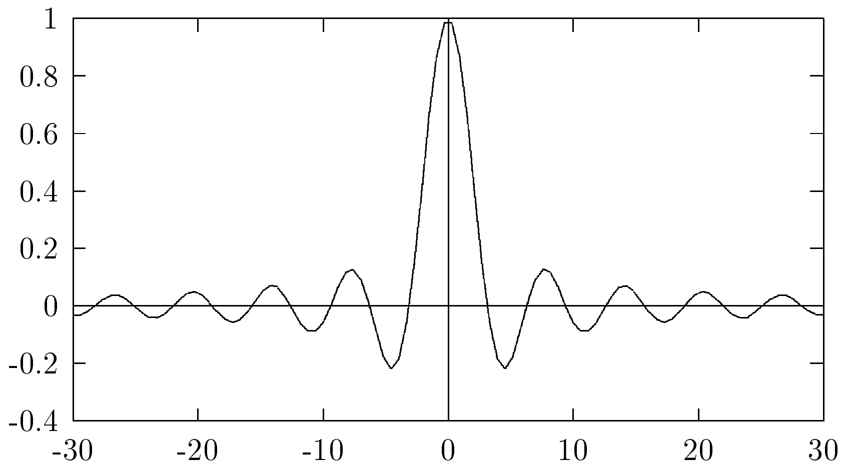
Lav-pass filtrering

Ved en lav-pass filtrering af et billede dæmpes eller undertrykkes de middelstore og høje frekvenser. Dette betyder at kontrasten i et billede forringes. Da mest støj er højfrekvent er det samtidig en metode til af filtrere denne støj bort . Grænsen mellem de frekvenser der skal passere og de der skal undertrykkes bestemmes ved "cut-off"-frekvensen D_0 , der defineres som den frekvens hvor filterværdien antager halvdelen af sin maksimale værdi.

Som regel vil næsten al energi i et billede være koncentreret i og omkring origo i den fouriertransformerede af billedet. Dette betyder at den meste energi bevares selv om "cut-off"-frekvensen vælges meget lille. Imidlertid er de høje frekvenser overordetlig bestemmende for det visuelle indtryk. Selv om blot få procent af energien i et billede afskæres ved at undertrykke de højeste frekvenser, vil billedet ofte have ringe lighed med det oprindelige billede. Dette viser at hvis formålet med filtreringen er at frembringe et billede, der skal benyttes af et menneske, da er det som regel en fordel at bevare en vis del af det højfrekvente indhold af billedet, selv om dette betyder at billedet er støjfyldt. I det følgende skal vi se på to elementære lav-pass filtre.

Idealfiltret Lad i det følgende $D = \sqrt{u^2 + v^2}$. Ideal lav-pass filtret, med "cut-off"-frekvens D_0 er givet ved:

$$\mathbf{h}(u, v) = \begin{cases} 1 & \text{hvis } D < D_0 \\ 0 & \text{ellers} \end{cases} \tag{1.121}$$



Figur 1.7: Funktionen $\sin(x)/x$

Idealfiltret er altså rotationssymmetrisk, og undertrykker alle frekvenser større end D_0 effektivt, hvorimod de lavere frekvenser passerer uændret.

Ved invers fouriertransformation af et idealfiltreret billede vil man som regel kunne observere en kraftig ringningseffekt . Denne optræder især ved kraftige kontraster i billedet. Årsagen til ringningseffekten kan forstås ud fra foldningssætningen. I det en-dimensionale tilfælde er den inverst fouriertransformerede af ideal-frekvensfiltret er givet ved:

$$h(x) = \frac{sin(2\pi D_0 x)}{\pi x} \tag{1.122}$$

Funktionen er skitseret i Figur 1.7: Jo større "cut-off"-frekvensen D_0 vælges, jo tættere vil de lokale ekstrema ligge, og jo større vil ekstremumsværdierne være. I dette tilfælde vil udtværingen af billedet være mindre, og ringningseffekten mindre synlig. Er D_0 lille vil der være et relativt lille antal ringe (lokale ekstrema) og disse vil derfor være mere synlige.

Årsagen til ringningseffekten er, at funktionen $\mathbf{h}$(x) er diskontinuert i D_0. Da foldningen kan betragtes som en kopiering, betyder dette at markante kontraster vil blive "kopieret" ud til positionerne af de lokale ekstrema i h(x). I praksis vil ringningseffekten være så generende, at idealfiltre sjældent kan benyttes. I stedet vil det være en fordel at benytte filtre hvor overgangen mellem de frekvenser der passerer og de der undertrykkes er mere glat.

Butterworthfilteret Et n'te ordens lav-pass butterworth filter med "cut-off"-frekvens D_0 er i frekvensområdet givet ved ligningen:

$$\mathbf{h}(u,v) = \frac{1}{1 + [D(u,v)/D_0]^{2n}} \tag{1.123}$$

Butterworthfiltret producerer (for lille n) kun i meget ringe grad nogen ringningseffekt. Årsagen er, at frekvensovergangen omkring "cut-off"-frekvensen D_0 er glat, og at halen af filtret tillader en vis portion af de høje frekvenser at passere. Dette betyder, at selv om højfrekvent støj dæmpes ved filtreringen, vil kanterne i det filtrerede billede stadig være bevaret i nogen grad (afhængig af valget af D_0 og n).

Filtre som butterworthfiltret kan være meget effektive til støjfiltrering, hvis denne er højfrekvent og har lille amplitude.

Hanning filtret

Hanning filtret er endnu et lav-pass filter. Dette er givet ved:

$$\mathbf{h}(u,v) = \frac{1}{2}\left[1 + \cos(\frac{\pi D}{2D_0})\right] \tag{1.124}$$

for $D \leq 2D_0$. Uden for denne cirkelskive er filterværdien nul.

Et grundlæggende problem ved alle de nævnte filtre er, at de undertrykker høje frekvenser relateret til kontraster i billederne på samme måde som høje frekvenser relateret til støj. Det er således ikke muligt at foretage en frekvensbaseret støjfiltrering uden samtidig at mindske kontrasten i det resulterende billede.

Høj-pass filtrering

Høj-pass filtrering foretages når formålet med filtreringen er at fremhæve kontrasterne i billedet (eller ved måling af afledede). Et høj-pass ideal-, hhv. første ordens butterworthfilter kan analogt til formlerne i foregående afsnit skrives:

$$\mathbf{h}(u,v) = \begin{cases} 0 & \text{hvis } D < D_0 \\ 1 & \text{ellers} \end{cases} \tag{1.125}$$

$$\mathbf{h}(u,v) = \frac{1}{1 + [D_0/D(u,v)]^2} \tag{1.126}$$

hvor D(u,v) $= \sqrt{u^2 + v^2}$, og "cut-off"-frekvensen D_0 som før er bestemt ved den frekvens hvor filtret antager havldelen af sin maksimale værdi.

Fremhævning af høje frekvenser

Ved høj-pass filtrering vil vi ofte observere at kanter fremhæves og at alle områder med glat varierende (oprindelig) intensitet vil få ens intensitet efter filtrering. Dette uafhængigt af at områderne før filtreringen havde markant forskellig (men glat varierende) intensitet. Årsagen er, at at de lave frekvenser er for effektivt undertrykt. En metode til at undgå dette er, at addere en konstant til høj-pass filtret (i frekvensområdet). Denne metode kaldes "at fremhæve de høje frekvenser" (eng. high frequency emphasis) . Er konstanten passende vil en anelse af de lavere frekvenser blive bevaret, og tonaliteten forbedret.

Et yderligere raffinement er at foretage en histogram linearisering af intensitetsværdierne i det inverst fouriertransformerede høj-pass filtrerede (frekvensfremhævede) billede. Dette vil ofte forbedre tonaliteten yderligere.

Ved tonaliteten af et billede forstås hvor godt intensitetsområdet udnyttes i forhold til detailrigdommen i billedet. Hvis billedet kun indeholder meget få forskellige gråtoner er tonaliteten dårlig. Det skal bemærkes at begrebet ikke er veldefineret, men at tonaliteten er en subjektiv bestemmelse.

Bånd-pass filtrering

Ved en bånd-pass filtrering tillades kun et bestemt frekvensbånd at passere. Passagen kan ske uhindret, som i et ideal bånd-pass filter eller dæmpes glat ved overgangen mellem de frekvenser der passerer og de der undertrykkes. Et ideal bånd-pass filter vil have formen:

$$\mathbf{h}(u,v) = \begin{cases} 1 & \text{hvis } D_0 - W < D < D_0 + W \\ 0 & \text{ellers} \end{cases} \tag{1.127}$$

hvor $D = \sqrt{x^2 + y^2}$, D_0 er "midterfrekvensen" af det frekvensbånd der skal passere, og $2W$ er bredden af frekvensbåndet.

Ofte benyttes et båndafvisningsfilter (eng. band reject filter), der i det ideale tilfælde fremkommer af ovenstående ligning ved ombytning af betingelserne i de to tilfælde.

Hvis der forekommer meget regelmæssig støj i et billede, f.eks en overlejret sinus bølge med en vis orientering og frekvens, vil denne støj optræde som punkter (u_0, v_0) og $(-u_0, -v_0)$ med numerisk stor værdi i styrkespektret for billedet. Det er ofte muligt at se sådanne punkter i et logaritmisk plot af styrkespektret. Interaktivt, er det derfor muligt at lægge to cirkelskiver omkring de to punkter, og at bestemme frekvensfiltret som det bånd-afvisnings filter, der tillader alle frekvenser undtagen dem i cirkelskiverne at passere uhindret. Eksempelvis kan et første ordens bånd-afvisnings butterworth filter til dæmpning af alle frekvenser i et cirkelbånd skrives som:

$$\mathbf{h}(u,v) = \frac{1}{1 + [\frac{D(u,v)W}{D^2(u,v) - C_0^2}]^2} \tag{1.128}$$

hvor C_0 angiver afstanden fra origo til centret af cirkelskiverne $\sqrt{u_0^2 + v_0^2}$.

1.6.2 Gendannelsesproblemet

Udgangspunktet for gendannelse af forringede billeder er den tidligere beskrevne simple model. Gendannelsesproblemet består i at finde en operator G der, anvendt på et forringet billede I, vil tilnærme det oprindelige ideelle billede I_0. Jævnfør den simple model $I = HI_0 + \eta$ bliver estimatet:

$$\hat{I} \;=\; GI \;=\; GHI_0 \;+\; G\eta \tag{1.129}$$

Gendannelsesfejlen bliver således:

$$e = ||I_0 - \hat{I}|| = ||(1 - GH)I_0 - G\eta|| \tag{1.130}$$

hvor "1" angiver identitetsoperatoren. Af ovenstående ligning fremgår at G dels bør vælges således at $G\eta$ er lille, dels som den (approksimativt) inverse til operatoren H. Disse to ønsker er ofte uforenelige. Desuden er det på ingen måde sikkert at den inverse til H eksisterer. Hvis f.eks. H er idealfilteret vil informationen knyttet til de eliminerede frekvenser ikke kunne genskabes. Selv for operatorer hvis punktspredningsfunktion har en strengt positiv fouriertransformeret vil rekonstruktionen ofte være utilfredsstillende. Dette skyldes at repræsentationen af flydende tal kun har en begrænset nøjagtighed. Endelig er det sjældent givet, for et givet billede I, hvilken forvrængning H billedet har været udsat for. Ligeledes er mængden og typen af støj i billedet som regel ukendt.

I det følgende beskrives en række klassiske metoder til gendannelse. Metoderne bygger alle på operationer i frekvensdomænet. I det simpleste tilfælde, hvor støjen η antages lig nul, vælges G som den (approximativt) inverse af H. Dette kaldes "invers filtrering" . Generelt er det nødvendigt at medtage støjkomponenten i modellen. Et eksempel på et sådant filter er "wienerfiltret" . De nævnte filtre kan udledes og beskrives ved brug af flere forskellige matematiske eller statistiske metoder. Vi skal her benytte den algebraiske tilgang.

Det inverse filter

Lad i det følgende g være et observeret billede fremkommet ud fra et idelt billede f ved forvrængning H og addition af uafhængig støj η. Som tidligere antages at H er lineær og positionsinvariant. Ved at nummererer alle billedpunkter (f.eks. række for række) kan funktionerne f, g, og η opskrives som tre søjlevektorer. Indeholder disse billeder N punkter ses, at operatoren H kan opskrives som en $N \times N$ matrice (da H er lineær). Den simple model for billedforringelse kan derfor skrives som et sæt af lineære ligninger:

$$\eta = g - Hf \tag{1.131}$$

Idet vi (på sædvanlig måde) definerer kvadrat-normen $||x||^2$ af en vektor x ved det indre produkt $x'x$, hvor x' angiver den transponerede af x, kan vi skrive:

$$||\eta||^2 = ||g - Hf||^2 \tag{1.132}$$

I udledningen af det inverse filter antages, at der ikke eksisterer nogen a priori viden om støjvektoren η. I stedet søges den løsning $\hat{f}$ der minimerer normen af støjleddet. Anderledes sagt søges $\hat{f}$ således at $H\hat{f}$ approksimerer g bedst muligt (i mindste kvadraters betydning) , dvs. som den funktion, der minimerer:

$$J(\hat{f}) = ||g - H\hat{f}||^2 \tag{1.133}$$

Den søgte løsning $\hat{f}$ findes på sædvanlig måde ved at differentiere J mht. $\hat{f}$, og ved at sætte resultatet lig nul:

$$\frac{\partial J(\hat{f})}{\partial \hat{f}} \;=\; -2H'[g \;-\; H\hat{f}] \;=\; 0 \tag{1.134}$$

hvor H' betegner den transponerede af matricen H. Videre fås:

$$\hat{f} \;=\; H^{\sharp}g \;=\; [H'H]^{-1}H'g \tag{1.135}$$

Filtret $H^{\sharp} \;=\; [H'H]^{-1}H'$ kaldes den *pseudoinverse* til H. Hvis det videre antages at H^{-1} eksisterer, dvs. at determinanten af matricen H er forskellig fra nul, reduceres ovennævnte filter til *det inverse filter* :

$$[H'H]^{-1}H' \;=\; H^{-1}(H')^{-1}H' \;=\; H^{-1} \tag{1.136}$$

Antag nu at H er en $N \times N$ matrice, og definer $N \times N$ matricen W ved:

$$W(j,k) \;=\; e^{\frac{2\pi i}{N}jk} \tag{1.137}$$

hvor $i = \sqrt{-1}$. Det kan vises at W^{-1} eksisterer og at $W^{-1}(j,k) \;=\; e^{-\frac{2\pi i}{N}jk}$. Desuden vil det gælde, at H kan skrives som:

$$H \;=\; WDW^{-1} \tag{1.138}$$

hvor D er en diagonalmatrice, hvor diagonalelementerne (egenværdierne for matricen H) netop er givet ved den fouriertransformerede $\mathbf{h}$ af punktspredningsfunktionen h for H. Det tekniske bevis for ovenstående er forholdsvist omstændeligt. Der udnyttes at h antages positionsinvariant, hvilket betyder at H er blokcirkulær. Interesserede lærese henvises til f.eks. *Gonzales and Woods: Digital Image Processing, Addison Wesley 1992* (kapitel 5 p. 257-268). Endelig kan det (let) vises at:

$$W^{-1}f \;=\; \mathbf{f} \tag{1.139}$$

hvor $\mathbf{f}$ er den diskret fouriertransformerede af f. Ved udnyttelse af ovennævnte resultater kan den inverse filtrering skrives:

$$\begin{aligned} \hat{f} \;&=\; H^{-1}g \\ &=\; (WDW^{-1})^{-1}g \\ &=\; WD^{-1}W^{-1}g \end{aligned} \tag{1.140}$$

og dermed:

$$W^{-1}\hat{f} \;=\; D^{-1}W^{-1}g \tag{1.141}$$

Da D er en diagonalmatrice med elementer $\mathbf{h}(u,v)$ betyder ovenstående ligning at den inverse filtrering kan formuleres i frekvensområdet ved:

$$\hat{\mathbf{f}}(u,v) \;=\; \frac{\mathbf{g}(u,v)}{\mathbf{h}(u,v)} \tag{1.142}$$

Denne meget ufuldstændige og dog omstændelige gennemgang bekræfter dermed hvad man intuitivt ville have forventet ud fra foldningssætningen.

Ud over antagelsen om at H er kendt, er der i udledningen af det inverse filter gjort to skrappe antagelser, nemlig at H^{-1} eksisterer og at støjen er lig nul. Hvis $\mathbf{h}(u,v)$ er forsvindende eller lig nul er divisionen i ovennævnte ligning ikke veldefineret. Hvis problemet kun opstår i et lille antal punkter i frekvensområdet er en brugbar metode at definere resultatet af divisionen til nul. Problemet med støjen er større. Hvis η ikke er lig nul fås:

$$\hat{\mathbf{f}}(u,v) \;=\; \mathbf{f}(u,v) \;-\; \frac{\boldsymbol{\eta}(u,v)}{\mathbf{h}(u,v)} \tag{1.143}$$

Typisk vil støjen være højfrekvent (dvs. at $\boldsymbol{\eta}(u,v)$ er ikke-forsvindende for store værdier af u og v), mens $\mathbf{h}(u,v)$ typisk aftager meget stærkt som funktion af u og v. Dette betyder at sidste led i ovennævnte ligning bliver dominerende. Anderledes sagt vil gendannelsen $\hat{f}$ kun have lidt til fælles med f. Af ovennævnte grunde er det inverse filter ikke anvendeligt i almindelighed. Smartere filtre, der i højere grad kan tage hensyn til eksistensen af støj, er nødvendige for en stabil gendannelse.

Wienerfilteret

Wienerfiltret hører til en klasse af betingede restaureringsfiltre. Før beskrivelsen af filtret skal vi kort gennemgå denne klasse. Som tidligere omtalt ønsker vi at finde en løsning $\hat{f}$ således at $(g - H\hat{f}) = \eta$. Denne matrix-ligning består af N ligninger med $2N$ ubekendte (N fra forvrængningen og N fra støjleddet). Matrix-ligningen har derfor mange løsninger. I tilfældet med det inverse filter løste vi problemet ved at antage at halvdelen af de ubekendte (dem fra støjleddet) var lig nul. I klassen af betingede restaureringsfiltre tilføjes N nye (lineære) ligninger til de N som de aktuelle data beskriver. Herefter kræves samtlige ligninger opfyldt bedst muligt. Denne metode går under navnet *regularisering* . I klassen af betingede rekonstruktioner formuleres de N nye ligninger ved: $Q\hat{f} = 0$, hvor Q er en lineær operator på f. En sammensat funktional $J(\hat{f})$ opbygges ved anvendelse af en såkaldt *Lagransk multiplikator* α:

$$J(\hat{f}) \;=\; ||Q\hat{f}||^2 \;+\; \alpha(||g - H\hat{f}||^2 - ||\eta||^2) \tag{1.144}$$

Løsningen til det sammensatte problem findes på sædvanlig måde ved at differentiere J og ved at sætte resultatet lig nul:

$$\frac{\partial J(\hat{f})}{\partial \hat{f}} \;=\; 2Q'Q\hat{f} \;-\; 2\alpha H'[g \;-\; H\hat{f}] \;=\; 0 \tag{1.145}$$

Løses denne ligning mht. $\hat{f}$ fås:

$$\hat{f} \;=\; (H'H \,+\, \gamma Q'Q)^{-1}H'g \tag{1.146}$$

hvor $\gamma = 1/\alpha$. Den endelige løsning findes ved at vælge γ således at kravet $||g - H\hat{f}||^2 \;=\; ||\eta||^2$ er opfyldt. Alternativt kan γ betragtes som en “fri parameter”. I denne situation vil betingelsen naturligvis ikke kunne blive opfyldt eksakt.

Det grundlæggende problem er at vælge en meningsfuld bibetingelse $Q\hat{f}$. I wienerfiltret er matricen Q valgt således at:

$$Q'Q \;=\; R_f^{-1}R_\eta \tag{1.147}$$

hvor

$$R_f^{-1} \;=\; E\{ff'\}^{-1} \tag{1.148}$$
$$R_\eta \;=\; E\{\eta\eta'\} \tag{1.149}$$

hvor $E\{\dot{}\}$ angiver forventet værdi, og f hhv. η er søjlematricer af billedværdier hhv. støj-værdier. R_f^{-1} hhv. R_η angiver således korrelationsmatricerne for f hhv. η. Typisk aftager korrelationen mellem billedelementerne meget stærkt som funktion af afstanden mellem disse. Dette betyder at matricerne R_f^{-1} og R_η langs hoveddiagonalen typisk vil indeholde et bånd af værdier der er forskellig fra nul. Uden for dette bånd vil matriceværdierne være lig nul. For sådanne matricer kan det vises (beviset falder langt uden for rammerne af disse noter) at matricerne kan approksimeres med blokcirkulære matricer. Dette betyder at vi kan skrive:

$$R_f^{-1} \;=\; WAW^{-1} \tag{1.150}$$
$$R_\eta \;=\; WBW^{-1} \tag{1.151}$$

hvor A hhv. B er diagonalmatricer indeholdende de fouriertransformerede af korrelations-elementerne i R_f^{-1} og R_η. Disse værdier, der i det følgende betegnes $\mathbf{S}_f$ hhv. $\mathbf{S}_\eta$, er igen lig værdierne af styrkespektret for f hhv. η. Idet vi udnytter at:

$$H \;=\; WDW^{-1} \tag{1.152}$$
$$H' \;=\; WD^\star W^{-1} \tag{1.153}$$

hvor $D^\star$ angiver den kompleks konjugerede af D, kan vi nu opskrive wienerfiltret:

$$\hat{f} \;=\; (H'H \,+\, \gamma R_f^{-1}R_\eta)^{-1}H'g \tag{1.154}$$
$$\;=\; (WD^\star DW^{-1} \,+\, \gamma WA^{-1}BW^{-1})^{-1}WD^\star W^{-1}g \tag{1.155}$$

Ved multiplikation af begge sider med W^{-1} og ved reduktion fås:

$$W^{-1}\hat{f} \;=\; (D^\star D \,+\, \gamma A^{-1}B)^{-1}D^\star W^{-1}g \tag{1.156}$$

Idet såvel D som A og B er diagonalmatricer kan vi derfor opskrive frekvensfilteret:

$$\hat{\mathbf{f}}(u,v) = \left[\frac{\mathbf{h}^{\star}(u,v)}{|\mathbf{h}(u,v)|^2 + \gamma\,[\mathbf{S}_\eta(u,v)/\mathbf{S}_f(u,v)]}\right]\mathbf{g}(u,v) \tag{1.157}$$

$$= \frac{1}{\mathbf{h}(u,v)}\left[\frac{|\mathbf{h}(u,v)|^2}{|\mathbf{h}(u,v)|^2 + \gamma\,[\mathbf{S}_\eta(u,v)/\mathbf{S}_f(u,v)]}\right]\mathbf{g}(u,v) \tag{1.158}$$

Det ægte wienerfilter fremkommer ved valg af $\gamma = 1$. Hvis γ betragtes som variabel tales om *det parametriske wienerfilter* . Bemærk, at hvis der ingen støj er tilstede, da vil styrkespektret $\mathbf{S}_\eta = 0$. I dette tilfælde reducerer filtret altså til det inverse filter. Bemærk også ved valget $\gamma = 1$ er der ikke længere nogen garanti for at betingelsen $||g - H\hat{f}||^2 = ||\eta||^2$ er opfyldt. Derimod kan det vises at udtrykket $E\{[f(x,y) - \hat{f}(x,y)]^2\}$ antager minimal værdi for dette valg.

For at kunne anvende wienerfiltret er det et krav at såvel styrkespektret $\mathbf{S}_f$ for det ideelle signal f er kendt og at styrkespektret $\mathbf{S}_\eta$ for støjen er kendt. Dette er ikke normalt tilfældet. Hvis støjen er hvid kan det vises at $E\{\mathbf{S}_\eta\} = \sigma_\eta^2$. Dette løser imidlertid ikke problemet omkring kendskabet til $\mathbf{S}_f$. I praksis ses nedenstående approksimation ofte anvendt:

$$\hat{\mathbf{f}}(u,v) \approx \frac{1}{\mathbf{h}(u,v)}\left[\frac{|\mathbf{h}(u,v)|^2}{|\mathbf{h}(u,v)|^2 + K}\right]\mathbf{g}(u,v) \tag{1.159}$$

hvor K^{-1} angiver signal-støjforholdet (for alle frekvenser). I praksis virker wienerfiltret meget bedre end det inverse filter hvis det oprindelige billede indeholder blot anelse støj. Af formel (1.157) fremgår, at frekvenserne i estimatet vægtes med det forventede støj/signal-forhold. Effekten af at minimere $||Q\hat{f}||$ bliver da, at betone de frekvenser i estimatet, hvor det forventede signal/støj-forhold er højt, og undertrykke de frekvenser, hvor forholdet er lavt. Denne effekt kan samtidig forklare, hvorfor wienerfiltret i praksis har en tendens til at opbløde kanter i billedet. Hvis støjen er hvid vil $\mathbf{S}_\eta$ (i teorien) være konstant. Derimod vil $\mathbf{S}_f$ aftage med afstanden fra Origo. Resultatet bliver at de lave frekvenser betones, mens de høje frekvenser undertrykkes. Filtret har altså lav-pass karakteristik.

Det geometriske middelværdifilter

Man kunne ønske at udvikle et filter, der lå "tæt ved" det inverse filter, men ikke døjede med singulariteterne i dette filter. En mulig parametrisering kunne være:

$$\hat{f} = (\text{invers filter})^{\alpha}\,(\text{parametrisk wiener filter})^{1-\alpha}\, g \tag{1.160}$$

hvor $0 \leq \alpha \leq 1$. I fourierområdet kan denne ide formuleres ved filteret:

$$\hat{\mathbf{f}}(u,v) = \left[\frac{\mathbf{h}^{\star}(u,v)}{|\mathbf{h}(u,v)|^2}\right]^{\alpha}\left[\frac{\mathbf{h}^{\star}(u,v)}{|\mathbf{h}(u,v)|^2 + \gamma\,[\mathbf{S}_\eta(u,v)/\mathbf{S}_f(u,v)]}\right]^{1-\alpha}\mathbf{g}(u,v) \tag{1.161}$$

Det geometriske middelværdifilter fremkommer nu ved valg af $\alpha = 1/2$. Hvis punktspredningsfunktionen er symmetrisk vil $\mathbf{h} = \mathbf{h}^{\star}$. I dette tilfælde får filteret formen:

$$\hat{\mathbf{f}}(u,v) = \left[\frac{\mathbf{S}_f(u,v)}{\mathbf{S}_f(u,v)\,|\mathbf{h}(u,v)|^2 + \gamma\mathbf{S}_\eta(u,v)}\right]^{1/2} \mathbf{g}(u,v) \tag{1.162}$$

I visse tilfælde virker dette filter bedre end både det inverse og wienerfilteret. Generelt er det vanskeligt at vælge en optimal værdi af α.

Bayes Restoration

En teoretisk anden måde at bestemme restaurationsfilteret på er ved at betragte billederne som stokastiske felter (markov random fields) hvor de enkelte billedelementer er stokastiske variable. Dette muliggør anvendelsen af *Bayes sætning*, idet man så kan tale om sandsynligheden for observation af bestemte billeder. Bayes sætning siger at:

$$p(f|g) = \frac{p(g|f)p(f)}{p(g)} \tag{1.163}$$

I det konkrete tilfælde kan ligningen læses: Sandsynligheden for det restaurerede billede f (vores model/fortolkning) givet det observerede billede g er lig sandsynligheden for observationerne givet modellen ganget med sandsynligheden for modellen og delt med sandsynligheden for observationerne. Givet g ønsker vi at bestemme det billede f, der er mest sandsynligt, dvs. vi ønsker at finde det f, der maksimerer den *a posteriori sandsynlighed* $p(f|g)$. Dette kaldes et *MAP estimat.* Da $p(g)$ er konstant kan dette udtryk ignoreres i optimeringen. Opgaven er derfor at bestemme f sådan at $p(g|f)p(f)$ maksimeres. Dette kræver at der opstilles analytiske udtryk for fordelingerne af vores prior $p(f)$ og af likelihoodfunktionen $p(g|f)$.

En rimelig og generel a priori antagelse er at intensiteten f er multinomialt fordelt med middelværdi $\bar{f}$ og covariansmatrice C_{ff}:

$$p(f) = \frac{1}{(2\pi)^{n/2}\,|C_{ff}|^{1/2}}\, e^{-\frac{1}{2}(f-\bar{f})^t C_{ff}^{-1}(f-\bar{f})} \tag{1.164}$$

hvor f betragtes som en søjlevektor af de $n = N \times M$ værdier i billedet (f.eks. opstillet række for række), og hvor:

$$\bar{f} = \mathcal{E}\{f\} \tag{1.165}$$

$$C_{ff} = \mathcal{E}\left\{(f-\bar{f})(f-\bar{f})^t\right\} \tag{1.166}$$

hvor $\mathcal{E}\{\cdot\}$ angiver statistisk forventet værdi. $\bar{f}$ er altså en søjlevektor med længde n af de statistisk forventede intensiteter. C_{ff} er en $n \times n$ matrice hvor det (i,j)'te element angiver covariansen mellem intensiteterne i positionerne i og j. Bemærk at diagonalelementerne af C_{ff} indeholder varianserne af de statistisk forventede intensiteter i hvert billedpunkt.

En typiskt observation i mange billedpunkter er at intensiteten er sammenlignelig med intensiteterne i en lille omegn om punktet. Denne observation angiver en mulig måde at estimere $\bar{f}$ på, og svarer til den typiske antagelse om at f er glat. Tilsvarende vil covariansern for punkter langt fra hinananden som regel være forsvindende. Dette betyder at C_{ff} vil have båndform med de største værdier liggende tæt omkring diagonalen. I det følgende skal vi antage den mest simple form af covariansmatricen, nemlig at C_{ff} er en diagonalmatrice med samme konstante værdi i alle diagonalelementer:

$$Cov(f_i, f_j) \;=\; 0 \qquad \text{og} \qquad Var(f_i) \;=\; \sigma_f^2 \tag{1.167}$$

Dette betyder at:

$$p(f) \;=\; \frac{1}{(2\pi)^{n/2}\;\sigma_f^n}\; e^{-\frac{1}{2\sigma_f^2}(f-\bar{f})^t(f-\bar{f})} \tag{1.168}$$

Det ses at σ_f specificerer vores a priori antagelse om hvor glat intensiteten f er. Jo større σ_f jo mindre glat. Problemet om specifikation af den a priori fordeling $p(f)$ er nu blevet reduceret til et spørgsmål om estimation af $\bar{f}$ og af σ_f. $\bar{f}$ kan mest simpelt estimeres ved middelværdien af intensiteterne i de umiddelbare nabopunkter. Dette betyder at $f - \bar{f}$ kan skrives som foldningen $c(x,y) \star f(x,y)$, hvor:

$$c(x,y) \;=\; \begin{cases} \;\;1 & \text{hvis} \;\; x = y = 0 \\ -\frac{1}{4} & \text{hvis} \;\; (x,y) \in \{(0,1),(0,-1),(1,0),(-1,0)\} \end{cases} \tag{1.169}$$

Alternativt kan vi skrive $f - \bar{f} \;=\; Cf$, hvor C er en $n \times n$ matrice:

$$C \;=\; \frac{1}{4}\begin{bmatrix} 4 & -1 & & & -1 & & & & & \\ -1 & 4 & -1 & & & -1 & & & & \\ & -1 & 4 & -1 & & & & & & \\ & & & 4 & & & \ddots & & & \\ & & & & & & & & & \\ -1 & & & & & \ddots & & & -1 & \\ & -1 & & & & & & & & -1 \\ & & & & & & & & & \\ & & & \ddots & & & & & & \\ & & & & & & & 4 & & \\ & & & & -1 & & & -1 & 4 & -1 \\ & & & & & -1 & & & -1 & 4 \end{bmatrix} \tag{1.170}$$

Den betingede sandsynlighed $p(g|f)$ beskriver fordelingen af støjen i billedet, idet vi antager at g er fremkommet ved forvrængning med punktspredningsfunktionen h og addition af støjen η: $g = h \star f + \eta$. Givet f kan vi derfor for enhver g udregne η. Vi vil i det følgende antage at støjen er normalfordelt med middelværdi 0 og med varians σ_η:

$$p(g|f) \;=\; \frac{1}{(2\pi)\;\sigma_\eta^2}\; e^{-\frac{1}{2\sigma_\eta^2}(g-Hf)^t(g-Hf)} \tag{1.171}$$

hvor H er den til punktspredningsfunktionen svarende matrice.

Ønsket om at maksimere $p(g|f)p(f)$ kan ækvivalent udtrykkes som ønsket om at minimere $J(f) = -2\log(p(g|f)p(f))$ (da logaritmefunktionen er monotont voksende).

$$J(f) \;=\; konstant \;+\; \frac{1}{\sigma_f^2}(Cf)^t(Cf) \;+\; \frac{1}{\sigma_\eta^2}(g-Hf)^t(g-Hf) \tag{1.172}$$

Ved at løse $\frac{dJ(f)}{df} \;=\; 0$ fås:

$$\frac{1}{\sigma_f^2}C^tCf \;+\; \frac{1}{\sigma_\eta^2}H^t(g-Hf) \;=\; 0 \tag{1.173}$$

På samme måde som tidligere kan ligningen omskrives til at gælde i fourierområdet (ligningen har - bortset fra andet valg af matricen C - helt samme form som i udledningen af wienerfilteret). Løsningen er derfor:

$$\hat{\mathbf{f}}(u,v) \;=\; \frac{\mathbf{h}^\star(u,v)}{|\mathbf{h}(u,v)|^2 \;+\; \gamma|\mathbf{C}(u,v)|^2}\,\mathbf{g}(u,v) \tag{1.174}$$

hvor $\mathbf{C}(u,v)$ er styrkespektret til funktionen $c(x,y)$, og hvor $\gamma \;=\; \sigma_\eta^2/\sigma_f^2$.

I visse tilfælde kan det være hensigtsmæssigt at bestemme $\bar{f}$ ved gennemsnittet af intensiteterne i en omegn, hvis form ændrer sig fra punkt til punkt. På denne måde vil man kunne undgå at glathedskravet forhindrer rekonstruktion af skarpe kontraster (overvej dette). I dette tilfælde kan "løsningen" mest hensigtsmæssigt skrives:

$$f \;=\; \bar{f} \;+\; \gamma^{-1}H^t(g \;-\; Hf) \tag{1.175}$$

Udtrykket indikerer en beregningsmetode ved iterativ estimation, hvor $f^{(0)}$ bestemmes ved g selv, eller - meget bedre - f.eks. ved invers eller wienerfiltrering. En god initialværdi er normalt afgørende for en hurtig konvergens. Derefter itereres (til konvergens):

$$f^{(i+1)} \;=\; \bar{f}^{(i)} \;+\; \gamma^{-1}H^t(g \;-\; Hf^{(i)}) \tag{1.176}$$

Sidste led i opdateringsformlen implementeres lettest i forurierområdet. Den iterative metode er naturligvis mere beregningskrævende, men tillader positionsvariante elementer hvilket afgørende kan forbedre kvaliteten af rekonstruktionen.

Som i de tidligere nævnte metoder til rekonstruktion benytter ovenstående metode sig også af kendskab til en række parametre. Det drejer sig om tallene σ_f^2, σ_η^2, samt om punktspredningsfunktionen h. Som tidligere beskrevet kan variansen af støjen estimeres rimeligt ud fra billedet g selv. Variansen af g vil ofte være et overtal for variansen af f (men det afhænger af f og af mængden af støj). Ved iterative metoder kan det være en fordel at starte med at vælge σ_f^2 lille (svarende til antagelse om stor glathed af f). Værdien kan så øges op mod variansen af $f^{(t)}$.

Iterative suboptimale Filtre

Der findes mange måder hvorpå wienerfiltret kan forbedres. En måde er at formulere bibetingelser $Q\hat{f} = 0$, der bedre er i stand til at bevare kanterne i rekonstruktionen end valget (1.147) er i stand til. En anden måde er iterativt at bestemme nogle af de parametre, der enten er ukendte eller dårligt estimerede. Som eksempel kræves i den generelle formulering at γ vælges i overensstemmelse med betingelsen $||g - H\hat{f}||^2 = ||\eta||^2$. Idet vi kan definere residualvektoren $r = (g - H\hat{f})$ fås:

$$r(\gamma) \;=\; g \;-\; H(H'H \;+\; \gamma Q'Q)^{-1}H'g \tag{1.177}$$

Det kan vises at $||r||^2$ vokser monotont som funktion af γ. Hvis vi kan finde et udtryk for $||\eta||^2$ er det derfor let at konstruere en procedure, hvorved γ iterativt justeres indtil $||r||^2 - ||\eta||^2$ antager minimal værdi. Det kan vises at der gælder:

$$||n||^2 \;=\; N[\sigma_\eta^2 \;+\; \bar{\eta}^2] \tag{1.178}$$

hvor $\bar{\eta}$ hhv. σ_η^2 er middelværdien og variansen af støjen, og N er det totale antal billedelementer. For nærmere diskussion af denne metode refereres til *Gonzales, Woods: Digital Image Processing. Addison Wesley 1992.* Se også *Andrews, Hunt: Digital Image Restoration. Prentice Hall 1977.*

1.6.3 Filtrering i intensitetsområde

Filtrering i intensitetsområdet foretages med fordel når støtten for filtret er lille eller når specielt foldningsmaskineri er til rådighed. Fordelen ved filtrering i intensitetsområdet bliver særlig stor hvis filtret er separabelt, dvs. når filtret $F(x, y)$ kan skrives:

$$F(x, y) \;\;=\;\; f_1(x)f_2(y) \tag{1.179}$$

Eksempelvis er kernen for boxfiltret og gaussfiltret (se nedenstående afsnit), samt kernen for fourierfiltret alle separable.

Antag i det følgende at filtreringens formål er at fjerne støjen og at denne er hvid. Ved lineær filtrering i intensitetsområdet foretages filtreringen ved beregning af et lokalt vægtet gennemsnit. Hvis støtten for filtret er stor vil filtrets egenskab mht. undertrykkelse af støj som regel være god, idet støjen i mange billedpunkter indgår i beregningen, og fordi den forventede værdi $E\{\eta\}$ af støjen er lig nul.

$$F * I = F * I_0 + F * \eta \approx E\{I_0\} + E\{\eta\} \approx E\{I_0\} \tag{1.180}$$

I områder hvor I_0 kun varierer lidt vil $I_0 - F * I$ defor nærme sig nul for voksende støtte. I områder hvor I_0 varierer meget vil $I_0 - F * I$ imidlertid vokse med voksende støtte. Der er således tale om en afvejning mellem et filters støjfiltreringsegenskab og dets egenskab mht. bevarelse af det "oprindelige" billede. Generelt er det nødvendigt at øge støtten jo ringere signal-støjforholdet i et billede er. I dette tilfælde er mere global information nødvendig.

Hvis filtreringens formål er at producere et billede til efterfølgende menneskelig inspektion vil det imidlertid ofte være en fordel at holde støtten lille. Mennesker er som regel i stand til at abstrahere fra ret store mængder af støj, men har svært ved at beskrive et billede hvor strukturene er udviskede. Filtrets støtte er naturligvis ikke enebestemmende for filtrets egenskaber. Filterformen er overordentlig vigtig. I de følgende afsnit beskrives to meget benyttede filtre samt deres egenskaber.

Box-filtre

Box-filtre har kvadratisk støtte og konstant værdi. Filtret antages at have integral 1.0 således at et $N \times N$ box-filter har værdier $1/N^2$. Box-filtret er ofte benyttet da det er separabelt og let at implementere. Den fouriertransformerede af et box-filter f er givet ved:

$$\mathbf{f}(u,v) = \frac{sin(2\pi Nu)}{\pi Nu}\frac{sin(2\pi Nv)}{\pi Nv} \tag{1.181}$$

Heraf ses, at nogle høje frekvenser (svarende til nulpunkterne i den fouriertransformerede) vil blive dæmpet totalt, hvorimod andre høje frekvenser (svarende til ekstrema) kun dæmpes moderat. Da amplituden af den fouriertransformerede aftager som $1/(\pi Nv)$ ses at støjdæmpningen vil være moderat med mindre N vælges meget stor. Til gengæld vil filtreringen så også berøre det lave og det mellemstore frekvensindhold. I praksis vælges N ofte lille (få pixel) da filtreringen så bliver meget hurtig.

Gauss-filtre

Gaussfiltret er vel nok det mest interessante af alle filtre. Vi skal møde gaussfilteret et stort antal gange, både her hvor det optræder som støjfilter, ved detektion af kanter i billeder samt ved kodning og repræsentation af billeder. Det todimensionale gaussfilter er givet ved:

$$D_\sigma(x,y) = \frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1.182}$$

Gaussfiltret er rotationssymetrisk, separabelt, parametriseret ved kun een parameter σ, der angiver spredningen af filtret. σ kaldes ofte for *skalaparameteren* . Den fouriertransformerede af en gaussfunktion med spredning σ er selv en gaussfunktion, og har spredning $1/2\pi\sigma$.

Gaussfunktionen er positiv men aftager så hurtigt, at den i beregningsmæssige sammenhænge ofte trunkeres til nul, for r = $\sqrt{x^2+y^2}$ større end 2σ eller 3σ. For store værdier af σ vil støtten for filtret (i intensitetsområdet) derfor blive stor, og selv om separabiliteten udnyttes vil foldningen tage forholdsvist lang tid. I dette tilfælde kan filtreringen ofte foretages hurtigere ved implementation i frekvensområdet.

Generelt vil der gælde at jo mere koncentreret (mindre støtte) en funktion er i intensitetsområdet, jo mindre koncentreret (mindre støtte) vil dens foruriertransformerede være i frekvensområdet. Således er den fouriertransformerede af en impulsfunktion jo en konstant. Graden af koncentration kan f.eks måles ved spredningen af funktionen, eller ved "cut-off"-frekvensen. Man kunne da spørge hvilket filter der var mest muligt koncentreret i såvel intensitetsområdet som i frekvensområdet. Det kan vises at dette filter netop er gaussfiltret. Denne egenskab er en af årsagerne til at gaussfiltret benyttes så ofte.

Som vi skal diskutere i kapitlet om kantdetektion har gaussfiltret andre enestående egenskaber, f.eks at det er løsning til varmeledningsligningen. Betragter vi et billede I som repræsenterende temperatur hvor stor hhv. lille intensitet repræsenterer varme hhv. kulde og betragter vi et kontinuum af foldninger $G_\sigma * I$ som funktion af spredningen σ (tiden), vil de varme hhv. kolde områder langsomt udlignes når tiden går (σ øges). Kolde områder vil blive varmere og varme områder koldere. For $\sigma \to \infty$ vil temperaturen (intensiteten) overalt tilnærme middeltemperaturen (middelintensiteten). Udvidelsen af signalet/billedet med den ekstra dimension σ (tiden) kaldes for *skalarummet* . Der forskes meget i hvorledes et signal/billede kan beskrives kompakt og robust ved "features" i skalarummet (dvs. målinger af hvorledes signalet/billedet ændrer sig som funktion af σ). Hvis en sådan repræsentation er mulig vil den beskrive signalet/billedet hierakisk, og ikke blot som en følge af ækvidistante målinger. Vi skal ikke her gå nærmere ind i dette (interessante) emne.

En tredie enestående egenskab ved gaussfiltret er at dette filter vil fremkomme ved gentagende filtrering af et vilkårligt filter med sig selv. Dette er en naturlig konsekvens af *de store tals lov* i sandsynlighedsregningen, og kan f.eks udnyttes hvis specielt hurtigt "hardware" til eksempelvis 3×3 -boxfiltrering er til rådighed. En acceptabel approximation til en gaussfiltrering kan da opnås ved gentagende box-filtrering. Ofte vil et lille antal iterationer være tilstrækkeligt. En anden konsekvens er, at foldningen af to gaussfiltre igen bliver et gaussfilter. Eksempelvis vil:

$$G_{\sigma_1} * G_{\sigma_2} = G_{\sqrt{\sigma_1^2+\sigma_2^2}} \tag{1.183}$$

1.7 Ikke-lineære filtre

Filtrering med ikke-lineære filtre kan naturligvis ikke foretages ved nogen foldning. Ikke-lineære filtre virker ofte bedre end lineære filtre fordi de bedre kan tage hensyn til abrudte ændringer i intensitetsfunktionen. Dette bliver eksemplificeret ved medianfiltret og Naga-ofiltret, der begge er udviklet til at foretage støjfiltrering uden at ødelægge kanter i intensitetsfunktionen. Anisotrop diffusion er en teknik med samme formål, men udviklet ved generalisation af skalarumsteknikken. Regularisering er egentlig en metode til lineær filtrering, men medtages her fordi metoden let kan (og ofte bliver) udvidet på en sådan måde at filtreringen bliver ikke-lineær. Ikke-lineær filtrering kan også foretages i frekvensområdet. Dette eksemplificeres ved homomorfisk filtrering, hvor ideen er at seperere filtreringen af belysnings- og reflektionskomponenten af intensiteten.

1.7.1 Filtrering i intensitetsområdet

Ved ikke-lineær filtrering i intensitetsområdet er filtersvaret en ikke-lineær funktion af billedværdierne i en (lille) omegn af punktet. Alternativt foretages den ikke-lineære operation på et sæt af værdier beregnet ved lineære filtreringer. Dette er tilfældet ved Nagaofiltret, der er retningsspecifikt. Ikke-lineære filtreringer foretages som regel i intensitetsområdet, fordi disse filtre er lettest at formulere i intensitetsområdet.

Medianfiltret

Filtersvaret for et $N \times M$ medianfilter er bestemt ved medianen af de NM intensitetsværdier i støtten. Filtret er overordentlig effektivt til at fjerne støj hvor intensiteten i en lille del af billedpunkterne er meget støjfyldt, som f.eks. ved "salt og peber"-støj. Abrudte ændringer i intensitetsfunktionen efterlades forholdsvist skarpe. Ofte kan filtret anvendes til egentlig kontrastfremhævning.

Den mest oplagte implementation af filtret består i en sortering af de NM værdier og udvælgelse af den midterste værdi. Denne implementation benyttes ikke i praksis da den er meget langsom. I stedet benyttes en glidende modifikation af et histogram over intensitetsværdierne.

Støjfiltreringsegenskaben øges med filterstørrelsen. Ofte vil strukturen af kanter i billederne imidlertid blive forringet (kanterne bliver flossede). Alternativt kan filtreringen itereres et lille antal gange. Det bør endvidere bemærkes at støtten for filtret ikke behøver at være rektangulært, men f.eks. kan have form af et kors e.lign.

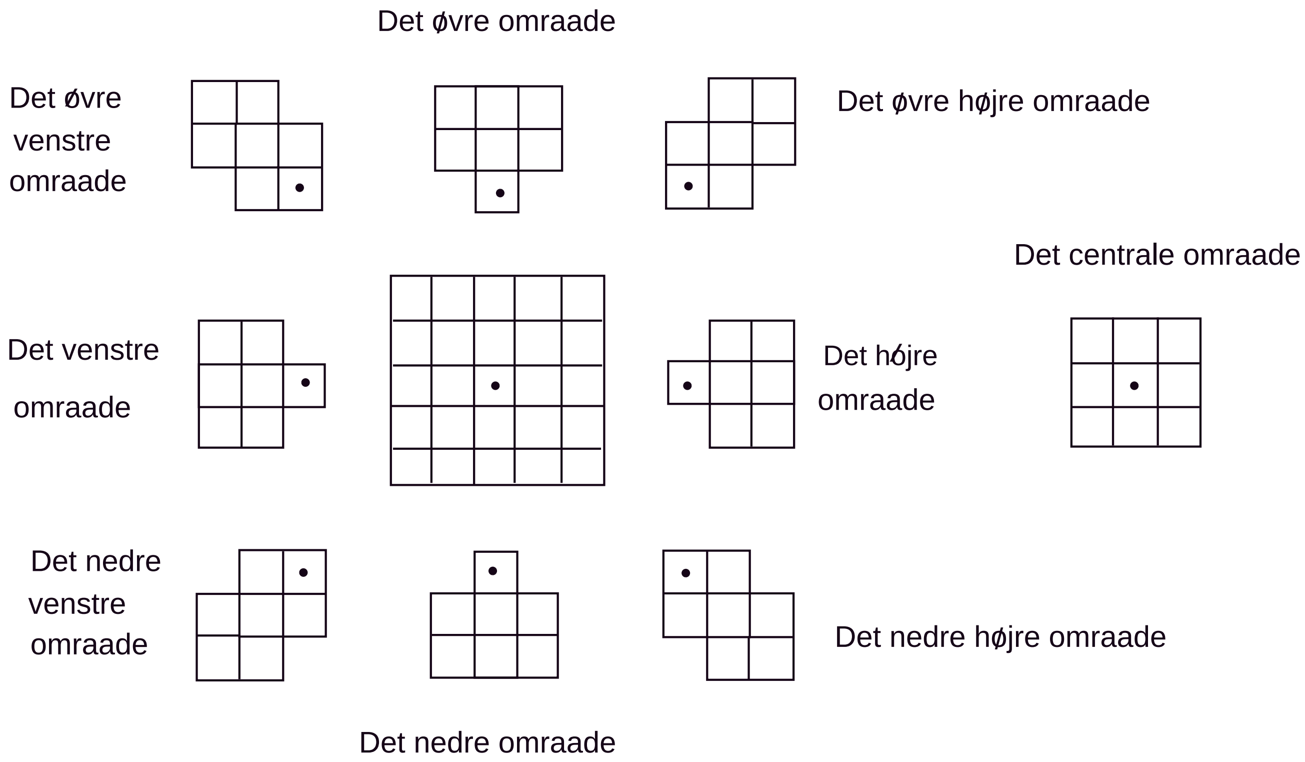
Nagaofiltret

Nagaofiltret er et retningsspecifikt filter, hvis formål er at foretage støjfiltrering uden at ødelægge kanter i intensitetsfunktionen. Filterstørrelsen er sædvanligvis 5×5. Ofte itereres filtreringen et lille antal gange. Støtten opdeles i 9 (overlappende) retningsspecifikke områder, der eksempelvis kan benævnes: det centrale, det øvre-venstre, det øvre-højre, det nedre-venstre, det nedre-højre, det øvre, det venstre, det højre og det nedre område. Disse er illustreret i nedenstående Figur 1.8, hvor billedelementet hvis filterværdi er under beregning er markeret ved en sort klat:

Metoden er for hvert område at bergne middelværdi og spredning af intensitetsværdierne for billedpunkterne i området. Filterværdien bestemmes da som middelværdien for det område, der har mindst spredning. Filtreringen kan således betragtes som en "retningsspecifik box-filtrering". Retningen vælges sådan at intensiteten i det givne punkt er mest i overensstemmelse med intensitetsværdierne i området svarende til retningen. Nagaofiltret er beskrevet i *N. Nagao and T Matsuyama: A Structural Analysis of Complex Aerial Photographs, 1980.*

1.7.2 Anisotrop diffusion

Som tidligere nævnt består en skalarumsudvidelse $I(x, y, t)$ af et billede $I(x, y)$ af en repræsentation af kontinuert mange foldninger af billedet med gaussfunktioner med voksende



Figur 1.8: De ni betragtede områder i nagaofilteret

varians t:

$$I(x, y, t) \;=\; G(0, t) \star I(x, y, 0) \tag{1.184}$$

hvor $I(x, y, 0) = I(x, y)$ og hvor $G(\eta, \sigma^2)$ er gaussfunktionen med middelværdi η og spredning σ. Ækvivalent kan vi betragte billedet $I(x, y, t)$ som en repræsentation af varme til tiden $t = 0$, og skalarummet vil angive varmefordelingen som funktion af tiden. Denne analogi er holdbar fordi gaussfunktionen er løsning til varmelednings- eller diffusionsligningen :

$$I_t \;=\; c[I_{xx} \;+\; I_{yy}] \tag{1.185}$$

hvor c er en konstant, kaldet varmeledningskoefficienten . Den grundlæggende idé ved anisotrop diffusion er at løsne på antagelsen at varmeledningskoefficienten er konstant. I dette tilfælde bliver diffusionsligningen:

$$\begin{aligned} I_t \;&=\; div(c(x, y, t)\nabla I) \;=\; c\Delta I \;+\; \nabla c \cdot \nabla I \\ &=\; c(x, y, t)[I_{xx}(x, y) + I_{yy}(x, y)] \;+\; c_x(x, y, t)I_x(x, y) \;+\; c_y(x, y, t)I_Y(x, y) \end{aligned} \tag{1.186}$$

Hvis c tillades at være lille hvor kontrasten i billedet er stor (og iøvrigt vælges fornuftigt) da vil resultatet af diffusionen blive at:

- Intensiteten, som funktion af tiden t, vil blive mere og mere glat mellem de store kontraster i billedet. Til slut vil billedet være stykvist konstant. Alt støj mellem kanterne vil være filtreret bort.

- Kontrasterne i billedet bliver skarpere (svarende til en udfoldning af en eventuel forvrængning).

Problemet er at finde et fornuftigt udtryk for $c(x, y, t)$. Det kan vises at ovenstående kan opfyldes hvis c vælges således at:

1. $c(x, y, t) \;=\; g(||\nabla I(x, y, t)||)$
2. $g(0) \;=\; 1$
3. $g(\cdot)$ er kontinuert, differentiabel, positiv, og aftagende

To eksempler på valg af g er:

$$g(s) \;=\; \frac{1}{1 \;+\; [\frac{s}{K}]^{1+\alpha}} \tag{1.187}$$

$$g(s) \;=\; e^{-(\frac{s}{K})^2} \tag{1.188}$$

hvor K og α er parametre. Antag at vi har valgt c som skitseret ovenfor, og at vi kun ser på den en-dimensionale diffusion (langs gradienten) over en kant. Lad $\phi(I_x) = cI_x = g(I_x)I_x$. Diffusionsligningen kan nu skrives:

$$I_t \;=\; \frac{\partial}{\partial x}\phi(I_x) \;=\; \phi'(I_x) \cdot I_{xx} \tag{1.189}$$

Antag videre at I er 3 gange differentiabe, og at $c > 0$. Tidsgradienten af intensitetshældningen kan nu udregnes (fordi vi kan ombytte rækkefølgen af differentiationer):

$$\begin{aligned} \frac{\partial}{\partial t}(I_x) \;&=\; \frac{\partial}{\partial x}(I_t) \\ &=\; \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\phi(I_x)\right) \\ &=\; \phi'' \cdot I_{xx}^2 \;+\; \phi' \cdot I_{xxx} \end{aligned} \tag{1.190}$$

Antag videre at kanten er orienteret således at $I_x > 0$. I vendetangent-punktet gælder at $I_{xx} = 0$ og at $I_{xxx} < 0$. Derfor vil $\frac{\partial}{\partial t}(I_x)$ have modsat fortegn af $\phi'(I_x)$ i en omegn omkring vendetangent-punktet. Hvis $\phi'(I_x) > 0$ vil hældningen af intensitetsfladen derfor mindskes med tiden t. Hvis $\phi'(I_x) < 0$ vil intensitetsgradienten i kantpunktet vokse som funktion af t. Der eksisterer derfor en grænseværdi T for hvilken kanter udglattes hhv. skærpes afhængig af størrelsen af intensitetsgradienten i det oprindelige billede $I(x, y, 0)$. Størrelsen T vil være parametriseret af K. Hvis $g(s)$ vælges som gaussfunktionen i formel (1.188) er det let at vise at T er givet ved $T = \sqrt{2}K/2$. Værdien af K er således afgørende for om en kontrast i billedet er tilstrækkelig stor til at forhindre udglatning over kontrasten. Typisk angiver brugeren værdien af K. Et godt valg er ofte afgørende for resultatet.

Anisotrop diffusion implementeres lettest ved diskretisering af diffusionsligningen, isolering af udtrykket for $I(x, y, t)$, og konstruktion af et program der genberegner udtrykket for I ved iteration (og således forsøger at holde diffusionsligningen opfyldt). Den globale information om intensitensfladen vil typpisk sprede sig med en hastighed af 1 pixel per iteration. Diameteren af det største *segment* i billedet vil derfor bestemme det mindste antal iterationer der er nødvendige. Ofte spredes den lavfrekvente komponent af intensitetsfunktionen meget langsommere, hvorfor mange flere iterationer må foretages før konvergens. Hvis et subjektivt tilfredsstillende resultat er tilstrækkeligt kan man ofte nøjes med relativt få iterationer. En konkret implementation er beskrevet i: *P. Perona and J. Malik: Scale-Space and Edge Detection Using Anisotropic Diffusion, IEEE PAMI vol. 12, no. 7 1990.* Her er en lang række forhold omkring teknikken diskuteret med større grundighed. En simpel implementation er beskrevet herunder:

Betragt grafen med fem knuder svarende til punktet (x, y) og dets fire naboer. Lad $I^t_{x,y}$ angive intensiteten i punktet (x, y) til tiden t. Lad diffusionskoefficienten $c^{i,j}_{x,y}$ angive evnen til varmeledning mellem punkterne (x, y) og $(x + i, y + j)$. Lad endvidere:

$$\begin{aligned} D^{-1,0}_{x,y} &= I_{x-1,y} - I_{x,y} \\ D^{1,0}_{x,y} &= I_{x+1,y} - I_{x,y} \\ D^{0,-1}_{x,y} &= I_{x,y-1} - I_{x,y} \\ D^{0,1}_{x,y} &= I_{x,y+1} - I_{x,y} \end{aligned} \tag{1.191}$$

I det følgende approkimeres normen af gradienten af I med den absolutte værdi af projektionerne langs de fire kanter i grafen. Dette betyder at:

$$\begin{aligned} c^{-1,0}_{x,y} &= g(|D^{-1,0}_{x,y}|) \\ c^{1,0}_{x,y} &= g(|D^{1,0}_{x,y}|) \\ c^{0,-1}_{x,y} &= g(|D^{0,-1}_{x,y}|) \\ c^{0,1}_{x,y} &= g(|D^{0,1}_{x,y}|) \end{aligned} \tag{1.192}$$

Første led i diffusionsligningen (1.186) kan nu approksimeres ved:

$$\lambda[c^{-1,0}_{x,y} \cdot D^{-1,0}_{x,y} + c^{1,0}_{x,y} \cdot D^{1,0}_{x,y} + c^{0,-1}_{x,y} \cdot D^{0,-1}_{x,y} + c^{0,1}_{x,y} \cdot D^{0,1}_{x,y}] \tag{1.193}$$

hvor λ er en normaliseringskonstant. Perona og Malik ignorerer det andet led $\nabla c \cdot \nabla I$ i diffusionsligningen. Begrundelsen herfor (Perona og Malik giver ingen forklaringer) er at ∇I hurtigt bliver forsvindende i næsten alle punkter. Kun tæt på kanterne i billedet vil ∇I vedblive at være stor. Dette betyder at $c \approx 1$ næsten overalt. Derfor vil $\nabla c \approx \vec{0}$ næsten overalt. Tæt på en kant vil ∇c være parallel med ∇I. Kun her vil sidste term i diffusionsligningen derfor bidrage væsenligt. Argumentet er nu, at det område hvori $\nabla c \cdot \nabla I$ bidrager væsenligt meget hurtigt bliver forsvindende. Approksimationen til diffusionsligningen kan nu udtrykkes:

$$I^{t+1}_{x,y} = I^t_{x,y} + \lambda[c^{-1,0}_{x,y} \cdot D^{-1,0}_{x,y} + c^{1,0}_{x,y} \cdot D^{1,0}_{x,y} + c^{0,-1}_{x,y} \cdot D^{0,-1}_{x,y} + c^{0,1}_{x,y} \cdot D^{0,1}_{x,y}] \tag{1.194}$$

Diffusionen foregår ved iterativt at opdatere samtlige intensitetsværdier og diffusionskoefficienter. Det er ofte praktisk at lade diffusionskoefficienterne på randen af billedet være lig 0 (fast).

1.7.3 Regularisering

Regularisering er en teknik, der kan benyttes, hvis man ønsker at tilføje a priori viden (bibetingelser) til et minimeringsproblem. Hvis der eksempelvis er for få ligninger i forhold til antallet af ubekendte, kan den a priori viden formuleres som et sæt af yderligere ligninger, således at det samlede problem kan løses. Hvis der er ligninger nok, men løsningen er ustabil (f.eks. ved at konditionen af en matrixligning er dårlig) da kan tilføjelsen af et regulariseringsled stabilisere løsningen. I det følgende skal vi se et eksempel på anvendelse af regularisering til rekonstruktion af støjfyldte signaler.

Lad $c(x)$ være en diskret funktion af en variabel, og antag at $c(x)$ er fremkommet ud fra et ideelt signal $g(x)$ ved tilføjelse af støj $\eta(x)$. Vi ønsker nu at bestemme et signal $f(x)$ (løsningen), der dels ligger tæt på målingerne $c(x)$, og dels "opfører sig pænt". Sidstnævnte ønske formuleres tit som et krav til den afdelede $f'(x)$ at være numerisk lille. Sammenlagt ønsker vi at finde den funktion $f(x)$, der minimerer *energifunktionen* :

$$\Theta(f) \;=\; \int [\;||f(x)-c(x)||^2 \;+\; \lambda||f'(x)||^2]dx \tag{1.195}$$

hvor λ kaldes *regulariseringsparameteren* . Jo større λ vælges, jo mere vægt lægges der på opfyldelsen af sidste term, dvs. på glathedskravet. Jo mindre λ jo større vægt på overensstemmelse med det givne signal. Det skal bemærkes at differentialoperatoren i glathedsledet kan omfatte mere end den/de førsteafledede. I to dimensioner eller mere er valget af differentialoperator ikke triviel. Som vi skal se i Kapitel 2, kan ovenstående ligning løses ved brug af *variationsregning* . Her skal vi benytte en anden metode hvor vi først diskretiserer ligningen:

$$Min\;\sum\left\{[f(x)-c(x)]^2 \;+\; \lambda[f(x+1) \;-\; f(x)]^2\right\} \tag{1.196}$$

Ved nu at differentiere mht. $f(x)$, og derefter at sætte resultatet lig nul fås:

$$2[f(x)-c(x)] \;-\; 2\lambda[f(x+1) \;-\; f(x)] \;+\; 2\lambda[f(x) \;-\; f(x-1)] \;=\; 0 \tag{1.197}$$

For løsningen $f(x)$ til minimeringsproblemet skal ovenstående ligning være opfyldt for alle x. Dvs:

$$f(x)[1+2\lambda] = c(x) \;+\; \lambda[f(x+1) \;+\; f(x-1)] \tag{1.198}$$

Det er let at forestille sig hvorledes $f(x)$ kan bestemmes ved ud fra en startsituation $f^0(x) = c(x)$ at iterere nedenstående opdatering indtil der i alle punkter x gælder at $f^{t+1}(x) = f^t(x)$.

$$f^{t+1}(x) \;=\; \frac{c(x) \;+\; \lambda[f^t(x+1) \;+\; f^t(x-1)]}{1+2\lambda} \tag{1.199}$$

Denne *Jacobi* -opdatering svarer oplagt til løsning af et lineært ligningssystem $\mathbf{Xf} = \mathbf{c}$ med N ligninger og ubekendte (hvor N er antallet af oprindeligt givne data). Lad $\mathbf{c}$ hhv. $\mathbf{f}$ være søjlevektorer bestående af de givne data hhv. de ønskede funktionsværdier. Ved omskrivning af ovenstående ligninger er det let at se at matricen $\mathbf{X}$ får formen:

$$\mathbf{X} = \begin{pmatrix} 1+\lambda & -\lambda & & & & \\ -\lambda & 1+2\lambda & -\lambda & & & \\ & -\lambda & 1+2\lambda & -\lambda & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\lambda & 1+2\lambda & -\lambda \\ & & & & -\lambda & 1+\lambda \end{pmatrix} \tag{1.200}$$

hvor matriceværdierne uden for det viste bånd er lig nul. Det ses at matricen har båndform og er diagonaldominant. Dette garanterer konvergens af en Jacobi -eller en Gauss-Seidel opdatering.

I praksis giver regularisering, som beskrevet i ovennævnte eksempel, anledning til to problemer. Det første er at brugeren skal specificere en værdi af regulariseringsparameteren λ. Dette kan være vanskeligt og vil ofte kræve flere forsøg før et acceptabelt resultat er fundet. Det andet problem er mere fundamentalt: Diskontinuiteter i signalet vil blive udglattet på samme måde som støj. Denne observation er egentlig ikke overraskende fordi formuleringen af rekonstruktionen er lineær. Det er imidlertid ikke svært at modificere den basale regulariseringsligning (1.195), sådan at diskontinuiteter i signalet (til en vis grad) bliver bevaret, og ofte fremhævet. Årsagen til, at løsningen til regulariseringsligningen også kan formuleres som løsningen til et lineært ligningssystem er at alle termer i ligningen er kvadratiske. Ved *diskontinuert regularisering* er ideen at erstate det kvadratiske regulariseringsled med en mere kompliceret funktion af (f.eks.) den førsteafledede $f'(x)$.

$$Min \int [\ ||f(x) - c(x)||^2 \ + \ \lambda\ \rho(f'(x))]dx \tag{1.201}$$

Ved at vælge funktionen ρ passende er det muligt at bevare, og i visse tilfælde fremhæve, diskontinuiteter i den rekonstruerede funktion $f(x)$. Et eksempel på en sådan funktion er:

$$\rho(t) \ = \ \begin{cases} t^2 & \text{hvis } t \leq T \\ T^2 & \text{ellers} \end{cases} \tag{1.202}$$

hvor T er en ny parameter. Virkningen af denne funktion er, at store gradienter i funktionen f ikke straffes ud over konstanten T^2. Dette betyder at der ikke (i diskontinuitetspunkter) virker nogen kræfter på $f(x)$, som forsøger at udglatte funktionen. I stedet tilpasser funktionen $f(x)$ sig helt til de givne data $c(x)$ i sådanne punkter. Ovenstående model går i litteraturen under navnet *Den svage streng* , fordi funktionen (strengen) knækker når data "hiver for meget" i den. Problemet ved denne model er, at løsningsrummet ikke længere er konvekst. Der kan med andre ord findes flere funktioner f_i der lokalt minimerer den opstillede funktional. Derfor er det nødvendigt at benytte algoritmer som *simuleret udglødning* (Se Geman og Geman: *Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images, IEEE PAMI vol. 6 no. 6 pp. 721-741, 1984*) eller *Graduated Non-Convexity* -algoritmer (Se Blake og Zisserman: *Visual Reconstruction*, MIT-press, 1987) til at finde en løsning. Disse algoritmer er langsomme sammenlignet med lineære algoritmer og med metoder baseret på *Gradient descent* -metoden. En anden ρ-funktion er:

$$\rho(t) \;=\; \sigma \log(1 \;+\; t^2) \tag{1.203}$$

hvor σ er spredningen af støjen η. Ovenstående funktion går også under navnet *The Lorentzian estimator*. Det kan vises at funktionen er velegnet til diskontinuert regularisering, og at løsningsrummet er konveks hvis $\sigma \leq \sqrt{2/3}$. I denne situation kan *Gradient descent* metoden derfor benyttes. En dybere diskussion kan findes i: Mads Nielsen *Isotropic Regularization* DIKU-report 93/20.

Støjfiltrering og kontrastbevaring

Som vi har set i de foregående afsnit er høje frekvenser i fourierspektret for et billede både relateret til støj i billedet og til markante "spring" (kontraster) i intensitetsfunktionen. Hvis vi ønsker at bevare (eller evt. fremhæve) kontrasten samtidig med en bortfiltrering af støj, kan dette ikke ske ved nogen lineær filtrering. Medianfiltret og Nagaofiltret er eksempler på ikke-lineære filtre, der forsøger at opnå dette formål uden eksplicit at måle kontrasten og/eller støjen i biledet.

En anden mulighed er at tilpasse filtreringen efter målinger af den lokale mængde af kontrast og støj. Lad $I(x, y)$ være et billede og $m(x, y)$ være middelværdien af I beregnet ud fra en støttestørrelse på $N \times N$. En adaptiv filtreringsmetode kan da angives ved:

$$O(x, y) \;=\; \gamma(x, y)I(x, y) \;+\; (1 - \gamma(x, y))\, m(x, y) \tag{1.204}$$

hvor $\gamma(x, y)$ er en funktion, der vokser med øget lokal kontrast, og aftager med mængden af støj. Bemærk at γ ikke nødvendigvis behøver at være begrænset af $0 \leq \gamma \leq 1$. Eksempelvis kan γ vælges omvendt proportional med spredningen σ af intensitetsværdierne i støtteområdet. Alternative kan separate metoder til måling af støj (Se *Olsen S. Estimation of Noise in Images: An Evaluation; DIKU-rapport 91/10*) benyttes.

Et adaptivt kontrastfremhævende filter, der er opbygget lidt efter ovenstående pricip er *Wallis-filteret* :

$$O(x, y) \;=\; S\frac{I(x, y) - m^N(x, y)}{1 + \sigma^N(x, y)} \;+\; \alpha M \;+\; (1 - \alpha)I(x, y) \tag{1.205}$$

hvor $m^N(x, y)$ hhv. $\sigma^N(x, y)$ angiver middelværdi hhv. spredning af intensitetsfunktionen målt i et $N \times N$ vindue. Filteret er bestemt ud fra fire værdier:

1) N: Vinduesstørrelsen.
2) M: Den ønskede lokale middelværdi (typisk lig 128).
3) S: Den ønskede lokale spredning (typisk lig 32).
4) α: En vægtfaktor mellem 0 og 1.

Som det ses af formlen består udregningen af tre dele. De sidste to er en afvejning mellem den aktuelle værdi i det givne punkt og den ønskede middelværdi. Det første led kompensere værdien hvis denne adskiller sig væsenligt set i forhold til variationen af værdierne i omegnen. Filtreringsmetoden er udviklet til af bringe detalier med lille kontrast frem i et billede, der iøvrigt er domineret af områder med indbyrdes store kontraster. Filtret er ikke noget egenligt støjfilter. Dog ses det, at hvis σ er stor, da vil filterværdien tilnærme M for valg af α lig 1. Det er let af forestille sig en lang række varianter af denne metode.

1.7.4 Filtrering i frekvensområdedet: Homomorfisk filtrering

Homomorfisk filtrering er en ikke-lineær operation, der er baseret på en simpel antagelse om sammensætningen af frekvensspektret for et billede. Antagelsen er, at den observerede intensitet I kan beskrives ved produktet LR af en belysningskomponent L og en reflektionskomponent R. Logaritmen til I vil da være summen af logaritmerne til L og R. Om de fouriertransformerede af logaritmerne gælder da at:

$$F\{\log \mathrm{I}\} = \mathrm{F}\{\log \mathrm{L}\} + \mathrm{F}\{\log \mathrm{R}\} \tag{1.206}$$

Skrives ligningen $\mathbf{i} = \mathbf{l} + \mathbf{r}$, og betegnes filterfunktionen (i frekvensområdet) med $\mathbf{h}$ kan vi skrive:

$$I^{fil} = F^{-1}\{\mathbf{hi}\} = F^{-1}\{\mathbf{hl}\} + F^{-1}\{\mathbf{hr}\} = L^{fil} + R^{fil} \tag{1.207}$$

Tages nu eksponentialfunktionen af begge sider kan resultatet igen skrives som produktet:

$$I^* = L^* R^* \tag{1.208}$$

Som ovenfor beskrevet består homomorfisk filtrering i først fouriertransformere logaritmen til billedet. Den fouriertransformerede multipliceres med frekvensfiltret og resultatet invers fouriertransformeres. Endelig tages eksponentialfunktionen af den inverst fouriertransformerede.

Under antagelsen om ligningen I = L R, ses, at hvis de “interessante” frekvensområder for L og R ikke overlapper, da kan filtret **h** konstrueres således at det modificerer frekvensområderne forskelligt. I praksis vil L som regel være lavfrekvent, hvorimod R vil have ikke forsvindende høje frekvenser.

Homomorfisk filtrering er meget anvendelig hvis belysningen af den afbildede scene er ujævn. I så tilfælde skal filterværdierne vælges små for små frekvenser. Ønskes en samtidig kontrastfremhævning vælges filterværdierne store for større frekvenser. Den eksakte form af filtret kan variere meget. Et muligt parametriseret filter er:

$$\mathbf{h}(u,v) = \lambda - \frac{\lambda - \epsilon}{1 + \alpha(u^2 + v^2)} \tag{1.209}$$

eller

$$\mathbf{h}(u,v) = \lambda - (\lambda - \epsilon) e^{-\frac{u^2+v^2}{2\alpha^2}} \tag{1.210}$$

Her angiver λ og ϵ maksimums og minimums værdien af filtret, og α vil bestemme filterformen. I praksis er det relativt "let" at vælge fornuftige værdier af λ og ϵ (eksempelvis til 2.0 og 0.1). Det kan være betydelig sværere at vælge en optimal værdi af parameteren α. Igen skal det pointeres, at valget af parameterværdier afhænger af det konkrete billede samt af formålet med filtreringen.

1.8 Filtrering ved Matematisk morfologi

Morfologi er læren om form (især i biologi). Matematisk morfologi dækker over en række metoder til filtrering af signaler og billeder. Metoderne kan ofte beskrives intuitivt ved hvorledes de ændrer formen af et signal/billede. Matematisk morfologi bygger på mængdelære. I de næste par afsnit skal vi begrænse os til binær morfologi, dvs. morfologiske operatorer anvendt på binære billeder. Et sådant billede kan beskrives ved den delmængde $A \subset \mathcal{Z}^2$ af pixels der har værdi lig nul (er sorte). Ved anvendelse af en morfologisk operator dannes en ny mængde, der beskriver resultatet af filtreringen. Som vi skal se senere kan binær morfolgi let udvides til gråtone- (eller farve-) morfologi. Her betragtes blot mængder $A \subset \mathcal{Z}^p$. For gråtonemorfologi er $p = 3$ for farvemorfologi er $p = 5$. Det skal også bemærkes at operationerne let kan formuleres i det euclidiske rum E^p. Der er ingen krav om numerabilitet af elementerne i mængderne. Derimod er det et krav, at det rum hvori mængderne er defineret, er udstyret med et koordinatsystem. Således kan hvert element (punkt, pixel) i en mængde tilordnes en koordinat. Før de grundlæggende morfologiske operationer: Ekspansion (eng. dilation) og Erosion beskrives skal vi kort repetere nogle velkendte definitioner fra mængdelæren.

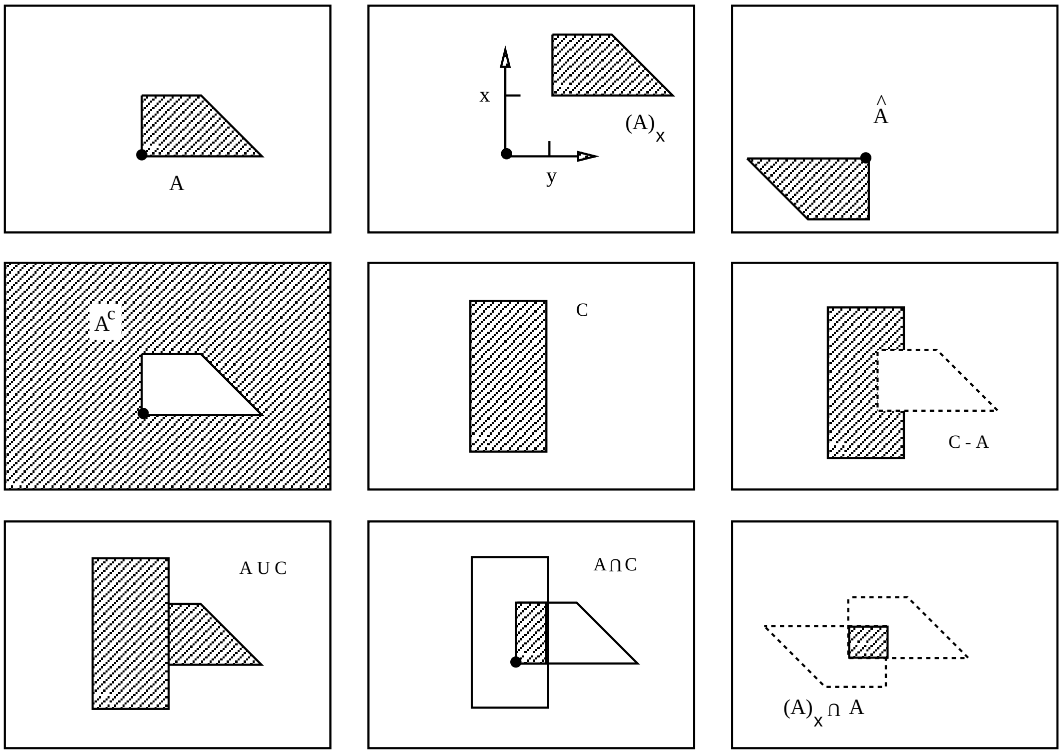
Foreningsmængde og fællesmængde: Foreningsmængden $A \cup B$ hhv. fællesmængden $A \cap B$ af to mængder A og B betegner de elementer (pixels) der tilhører en af de to mængder (ikke nødvendigvis dem begge) hhv. de elementer der netop tilhører begge mængder:

$$A \cup B = \{x \mid x \in A \text{ eller } x \in B\} \tag{1.211}$$

$$A \cap B = \{x \mid x \in A \text{ og } x \in B\} \tag{1.212}$$

Translation: Translationen $(A)_x$ af en mængde A med en værdi x er givet ved:

$$(A)_x = \{c \mid c = a + x, \text{ for a} \in A\} \tag{1.213}$$



Figur 1.9: De basele mængdeoperationer. Det sorte punkt "•" markerer Origo for koordinatsystemet.

Spejling: Spejlingen $\hat{B}$ af en mængde B er defineret ved:

$$\hat{B} \;=\; \{x \mid x = -b,\; b \in B\} \tag{1.214}$$

Komplement: Komplementet A^c til mængden A er givet ved:

$$A^c \;=\; \{x \mid x \notin A\} \tag{1.215}$$

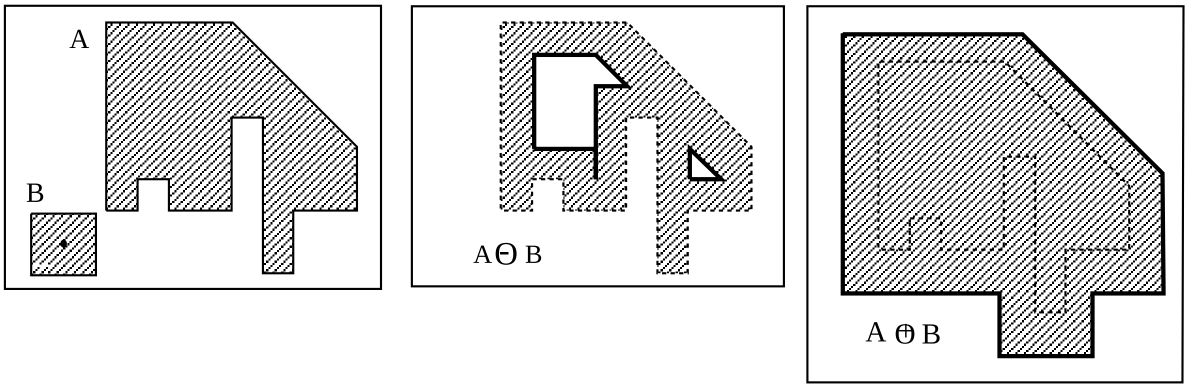
Difference: Differencen $A - B$ mellem to mængder A and B er givet ved:

$$A - B \;=\; A \cap B^c \;=\; \{x \mid x \in A, x \notin B\} \tag{1.216}$$

De basele mængdeoperationer er illustreret i Figur 1.9:

1.8.1 Ekspansion og Erosion

Ekspansion og erosion er de to vigtigste basisoperationer inden for matematisk morfologi. I det følgende vil A betegne billedmængden, dvs. mængden af "sorte" pixels i billedet.



Figur 1.10: Erosion og ekspansion

Mængden B betegnes *strukturelementet* , og svarer (i en vis forstand) til hvad vi normalt forstår ved et filter. Ekspansionen af mængden A med strukturelementet B er defineret ved:

$$A \oplus B = \left\{ x \mid (\hat{B})_x \cap A \neq \emptyset \right\} \tag{1.217}$$

$$= \left\{ x \mid (\hat{B})_x \cap A \subseteq A \right\} \tag{1.218}$$

Ekspansionen er således defineret ved mængden af alle forskydelser x således at $(\hat{B})_x$ og A overlapper med mindst ét element. Som det ses ligner definitionen meget udtrykket for en foldning. Erosion af mængden A med strukturelementet B er defineret ved:

$$A \ominus B = \{ x \mid (B)_x \subseteq A \} \tag{1.219}$$

Erosionen er således defineret ved de punkter x hvor B, translateret med x, er helt indeholdt i A. Det kan let vises at ekspansion og erosion er duale processer i den forstand at:

$$(A \ominus B)^c = A^c \oplus \hat{B} \tag{1.220}$$

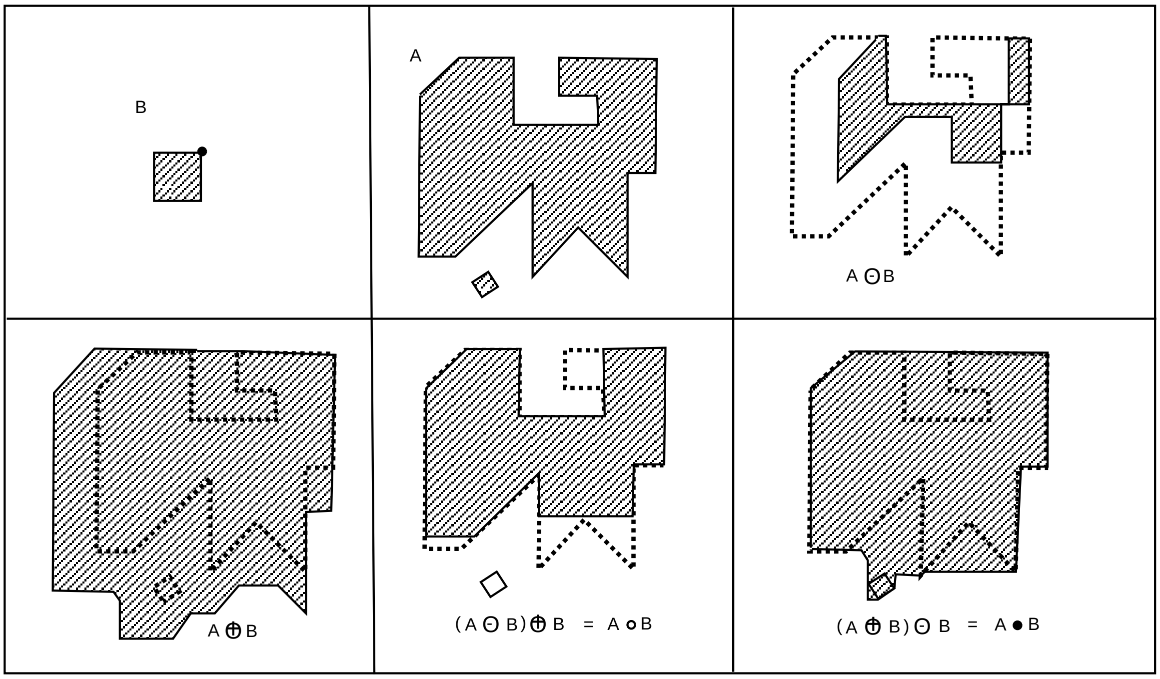
Resultatet af at erodere "objektet" A med strukturelementet B er således (bortset fra complementaritet) identisk med ekspansionen af "baggrunden" med det spejlede strukturelement. De to operationer er illustreret i Figur 1.10. Som det ses virker erosionen ved at "skrælle" et lag af "objekterne" i billedet. Herunder forsvinder mindre spidser og udbulinger. Ekspansionen, derimod, "lægger et lag på". Herunder opfyldes mindre huller og indhak.

1.8.2 Åbning og Lukning

åbning og lukning af en mængde A med strukturelementet B er defineret ved:

$$\text{åbning:} \quad A \circ B = (A \ominus B) \oplus B \tag{1.221}$$

$$\text{Lukning:} \quad A \bullet B = (A \oplus B) \ominus B \tag{1.222}$$



Figur 1.11: åbning og lukning

Som det fremgår af definitionerne på erosion og ekspansion er virkningen af såvel åbning som lukning at udglatte konturer. åbninger vil herunder eliminere tynde udbulinger, hvor lukninger vil fylde små indhak. Anderledes sagt: åbninger fjerner små "positive" støjelementer hvor lukninger fjerner "negative" støjelementer. Dette er illustreret i Figur 1.11. Som for ekspansion og erosion gælder der at åbning og lukning er duale operationer mht. komplement og spejling, dvs.:

$$(A \bullet B)^c = (A^c \circ \hat{B}) \tag{1.223}$$

Desuden gælder der:

$$A \circ B \subseteq A \subseteq A \bullet B \tag{1.224}$$
$$(A \bullet B) \bullet B = A \bullet B \tag{1.225}$$
$$(A \circ B) \circ B = A \circ B \tag{1.226}$$
$$C \subseteq D \Rightarrow C \circ B \subseteq D \circ B \tag{1.227}$$
$$C \subseteq D \Rightarrow C \bullet B \subseteq D \bullet B \tag{1.228}$$

Der kan desuden vises en lang række sætninger, der beskriver relationer mellem udtryk involverende de morfologiske operationer og de basale mængdeoperationer. Interesserede henvises f.eks. til: *Haralick et.al. : "Image Analysis Using Mathematical Morphology", IEEE PAMI vol. 9 no. 4 1987.*

1.8.3 Morfologiske operationer

En lang række operationer på binære billeder kan konstrueres ud fra de ovennævnte basale morfologiske operatorer. Af de mere simple kan der nævnes:

1. Operationen $f - (f \ominus B)$ ekstraherer indersiden af objekterne i et binært billede.
2. Operationen $(f \oplus B) - f$ ekstraherer ydersiden af objekterne i et binært billede.
3. Operationen $(f \circ B) \bullet B$ angiver en (formindsket) udglatning af objekterne i et binært billede.
4. Operationen $(f \bullet B) \circ B$ angiver en (forstørret) udglatning af objekterne i et binært billede.
5. Operationen $(f \bullet B) - (f \circ B)$ er et mål for støjen, herunder iregulariteten, af et binært billede.

Andre (mere komplekse) operationer er:

- Regionsfyldning (eng. region filling).
- Detektion af objektsammenhæng (eng. connected component analysis).
- Bestemmelse af det konvekse hylster.
- Udtynding af objekter (eng. thinning), herunder skelettransformationer.

Ovennævnte fire operationer implementeres ofte ved iterativ anvendelse af de mere basale operationer. En god beskrivese af sådanne implementationer kan f.eks. findes i: *Gonzales and Woods: "Digital Image Processing", 1992.* En af de mere simple operationer *Hit-or-miss operationen* beskrives kort herunder.

Hit-or-Miss transformationen

Hit-or-Miss transformationen har til formål at bestemme "match" mellem en prototype B og forekomster af samme figur i et billede A. Det er klart at det er et krav til et objekt i billedet at B er indeholdt i objektet. Der vil således gælde at $M \subseteq (A \ominus B)$ er mængden af de mulige placeringer af prototypen i billedet. Hvis billedet indeholder et objekt, der er meget større end B vil M indeholde en delmængde af dette objekt, selv om der ikke er tale om noget match. For at udelukke disse situationer betragtes komplementærmængden A^c. Denne mængde må indeholde komplementærmængden til prototypen B. Lad W være et lille vindue, der omslutter B. W kan eventuelt være fremkommet ved ekspansion af B. Komplementærmængden til prototypen B approksimeres nu ved $W - B$. Jævnfør diskussionen skal der gælde at: $M \subseteq A^c \ominus (W - B)$. Hit-or-Miss transformationen kan nu defineres:

$$A \odot B \;=\; (A \ominus B) \cap [A^c \ominus (W - B)] \tag{1.229}$$

$A \odot B$ angiver således de punkter hvor B finder et match i A samtidig med at $(W - B)$ finder et match i A^c.

Væsenligt ved al matematisk morfologi er valget af det strukturerende element . Hvis dette er valgt uhensigtsmæssigt vil resultatet af filtreringen ofte være meget anderledes end forventet. Specielt er det vigtigt at kende skalaen for de "fænomener" i billederne som man ønsker at fremhæve eller bortfiltrere. Problemet med valg af strukturelement svarer helt til det tilsvarende problem ved valg af sædvanlige filterfunktioner ved lineær filtrering. Der er ingen lette løsninger. Dog vil det oftest være fortolkningsmæssigt fordelagtigt at vælge et strukturelement, der er symmetrisk om Origo, f.eks. et kvadrat.

1.8.4 Gråtone-morfologi

Gråtonemorfologi defineres let ud fra binær morfologi ved at udvide dimensionen af mængderne. Således kan de to-dimensionale gråtoneoperationer *ekspansion* og *erosion* defineres ved:

$$\begin{aligned} f \oplus b(s,t) &= MAX\,\{f(s-x,t-y) + b(x,y) \;\;|\;\; (s-x,t-y) \in Def_f, (x,y) \in Def_b\} \\ f \ominus b(s,t) &= MIN\,\{f(s+x,t+y) - b(x,y) \;\;|\;\; (s+x,t+y) \in Def_f, (x,y) \in Def_b\} \end{aligned} \tag{1.230}$$

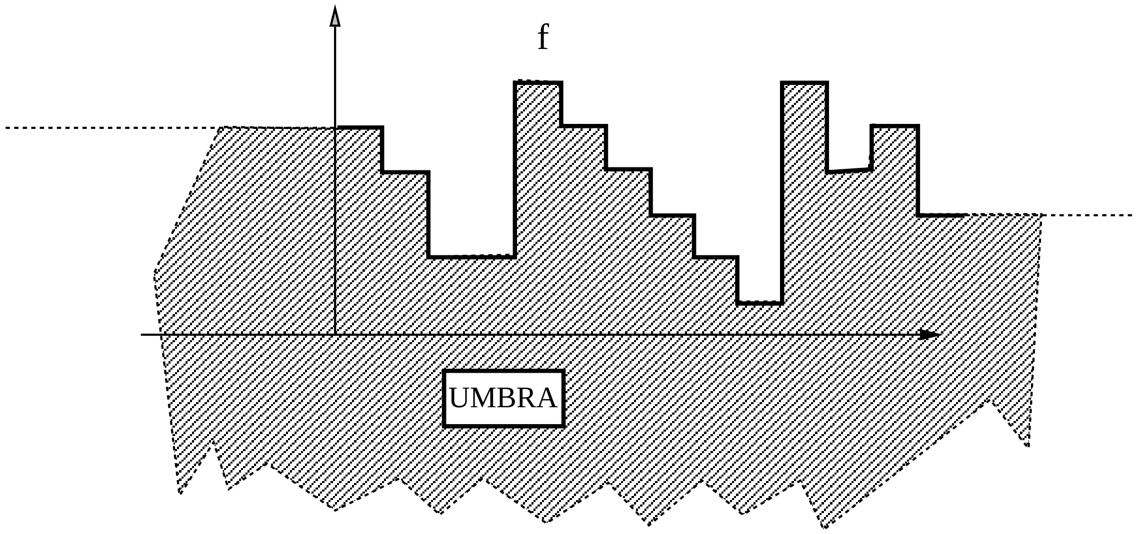
hvor Def_f hhv. Def_b er definitionsområderne for billedet f hhv. for strukturelementet b. Som det fremgår minder ovenstående udtryk meget om definitionerne af foldning og korrelation. Man kan forestille sig såvel billedet f som strukturelementet b som flader $(x, y, f(x, y))$ i 3D. Ekspansionen og erosionen fremkommer nu ved at lade fladeflisen (strukturelementet) b røre (glide hen over) den øvre hhv. nedre side af fladen f, og registrere de punktmængder som Origo for $\hat{b}$ hhv. b herved gennemløber. Det ses at erosionen generelt vil gøre billedet mørkere hvor ekspansionen vil gøre billedet lysere. Erosionen vil desuden fjerne relativt lyse pletter med bredde mindre end størrelsen af strukturelementet. Ekspansionen vil modsat fjerne mørke områder der ligeledes er mindre end strukturelementet.

En anden, måske lettere forståelig, fortolkning af de basale gråtonemorfologiske operationer baserer sig på definitionen af en **umbra** for en funktion. Lad $f(x)$ være en endimensional funktion med reelle eller heltallige værdier. Umbra'en for f er da defineret ved den punktmængde der ligger under funktionen. Dette er illustreret i Figur 1.12. De morfologiske operationer på gråtonefunktionen f kan ækvivalent defineres som de tilsvarende binære definitioner på umbra'en for f. Denne fortolkning kan umiddelbart udvides til flere dimensioner.

Definitionen af åbning og lukning af gråtone signaler eller billeder er helt identisk med definitionerne i det binære tilfælde, dvs.:

$$\text{åbning:} \qquad A \circ B = (A \ominus B) \oplus B \tag{1.231}$$

$$\text{Lukning:} \qquad A \bullet B = (A \oplus B) \ominus B \tag{1.232}$$



Figur 1.12: Umbra for en en-dimensional funktion

Idet vi udnytter definitionerne $f^c = -f(x,y)$ og $\hat{b} = b(-x,-y)$ kan man vise gråtone-dualitetsrelationen:

$$-(f \bullet b) \;=\; (-f \circ \hat{b}) \tag{1.233}$$

Hvis man antager at strukturelementet er kugleformet, kan man forestille sig en gråtone åbning som den (øvre) 3D-flade, som kuglen beskriver, når den bliver rullet hen over *undersiden* af fladen f. Tilsvarende kan lukningen beskrives ved den (nederste) flade som kuglen gennemøber når den rulles hen over *overfladen* af f.

1.8.5 Morfologiske gråtoneoperationer

Tre typiske morfologiske gråtoneoperationer er:

$$\text{Udglatning} \quad (f \circ b) \bullet b \tag{1.234}$$

$$\text{Gradient} \quad (f \oplus b) - (f \ominus b) \tag{1.235}$$

$$\text{Top-hat} \quad f - (f \circ b) \tag{1.236}$$

Det er ikke let at beskrive virkningen af de forskellige operationer i termer af frekvens-operationer. Visuelt har udglatnings -operationen lav-pas-karakteristik, mens gradient -operationen og top-hat -operationen har høj-pas-karakteristik. Sidstnævnte operator er velegnet til at fremhæve detalier. Mere avancerede operationer kan sammenstykkes ud fra de basale operationer ved iterativ anvendelse, ved kombination med de basale mængdeo-perationer, samt med kombination af simple histogramteknikker og tælleoperationer. En populær morfologisk metode til billedsegmentering (se kapitel 3 af noterne) er baseret på detektion af vandskel (eng. watershed). Det vil føre for vidt her at beskrive metoden nærmere.

1.9 Opgaver

1. Beskriv en hurtig implementation af den adaptive histogram-linearisering. Metoden skal **ikke** programmeres, men kun skitseres.

2. Beskriv en hurtig implementation af medianfiltret.

3. Bestem den største værdi af $v\sigma$ hvor en gaussfiltrering i intensitetsområdet (implementeret med støtte $4\sigma \times 4\sigma$) vil være hurtigere end den tilsvarende filtrering i frekvensområdet. Udnyt foldningssætningen. Dette betyder at der ingår to fouriertransformationer i frekvensfiltreringen. Benyt først at gaussfiltrer er separabelt, dernæst en implementation hvor separabiliteten ikke udnyttes. Antag et billede at størrelse $2^n \times 2^n$, og at den endimensionale fouriertransformation tager $N log(N)$ tidsenheder. (multiplikationer). Operationer til beregning af filterværdierne skal ikke medregnes.

4. Sammenlign eksperimentelt de fire filtreringer baserede på: 1) Boxfilteret; 2) Gaussfilteret; 3) Medianfilteret; og 4) Nagaofilteret. Vær omhyggelig med valget af billeder, samt med at beskrive de interessante forskelle som du iagttager. Vedlæg billeddokummentation.

5. Sammenlign eksperimentelt en lav-pass idealfiltrering med en lav-pass butterworthfiltrering. Vær omhyggelig med valget af billeder, samt med at beskrive de interessante forskelle som du iagttager. Vedlæg billeddokummentation.

6. Sammenlign eksperimentelt en høj-pass idealfiltrering med en høj-pass butterworthfiltrering. Vær omhyggelig med valget af billeder, samt med at beskrive de interessante forskelle som du iagttager. Vedlæg billeddokummentation.

7. Karakteriser eksperimentelt betydningen af valget af parametrene til den homomorfiske filtrering, sådan som det er beskrevet i noterne i formel (1.210).

8. Implementer en (vilkårlig) metode til støjfiltrering.

9. Skriv et program til rotation af et billede. Redegør for de problemer og de valg som du træffer i udformningen af programmet. Vis at programmet virker (for alle rotationsvinkler).

10. Formlen for den kubiske interpolation (afsnit 1.2) er taget fra *W. Nieblack: In Introduction to Digital Image Processing, Prentice Hall 1986*. Formlen svarer ikke til en sædvanlig interpolation med funktionen: $y = a_0 + a_1 x + a_2 x^2 + a_3 x^3$. Find koefficienterne til $I_0(a-1)$, $I_0(a)$ etc. for denne interpolation. Find interpolationsforskellen i $\alpha = 1/2$, for $I_0(a-1) = 0$, $I_0(a) = 2$, $I_0(a+1) = 2$, og $I_0(a+2) = 5$.

11. Udvid den kubiske interpolation til to dimensioner.

12. Angiv hvornår matricen $\mathbf{X}^t\mathbf{X}$, der indgår i den pseudoinverse (se afsnit 1.2) er invertibel.

13. Implementer en metode til geometrisk opretning af et billede ud fra en liste af "paspunkter". Benyt en lineær koordinattransformation.

14. Betragt Hanning (frekvens-) filtret: $\mathbf{h}(u) = \frac{1}{2}(1+\cos(u))$. Udled det tilsvarende filter i billedområdet. Benyt eventuelt foldningssætningen og opgave 27.

15. Beskriv en version af "Nagao-filteret" som benytter en støtte på 3x3 pixels. Implementer filteret, således at filtreringen kan itereres et vist antal gange. Sammenlign filteret eksperimentelt med det sædvanlige 5×5 "Nagao-filter".

16. Lad f være det eendimensionelle diskrete boxfilter med støtte 3. Fold filtret med sig selv 1,2,3 og 4 gange. Beskriv "udviklingen" af filterformen.

17. Sammenlign eksperimentelt den adaptive histogramlinearisering (for forskellig filterstøtte) med den globale histogram linearisering. Vær omhyggelig med valget af billeder, samt med at beskrive de interessante forskelle som du iagttager. Vedlæg billeddokummentation.

18. Lad r betegne (kontinuerte) intensitetsværdier i intervallet $[0:1]$, og lad tætheden $p_r(r)$ være givet ved: $p_r(r) = 2(1-r)$ for $0 \leq r \leq 1$. Beskriv, trin for trin, metoden for histogramlinearisering, og godtgør, at de resulterende intensiteter er uniformt fordelt i intervallet $[0:1]$.

19. Antag at vi har givet et 64x64 billede med 8 mulige intensitetsværdier, og at disse forekommer i billedet som angivet i tabellen.

0	1	2	3	4	5	6	7
201	843	902	103	32	1619	312	84

    Udled den transformation $s = T(r)$, der bevirker at s (teoretisk set) vil være uniformt fordelt. Beskriv den faktiske fordeling af s på de 8 intensitetsværdier.

20. Betragt et billede med intensitetsfordeling som i ovenstående opgave. Foretag en histogramtransformation, der resulterer i et billede med fordelingsfunktion $p_t(t) = 2t$.

21. Hvor godt kan i fjerne støjen i billederne *noisy.i* og *noisy2.i*. Der er frit valg mht. metode.

22. Lad

$$f(x) = \begin{cases} A & \text{hvis } 0 \leq x \leq 1 \\ 0 & \text{ellers} \end{cases}$$

$$g(x) = \begin{cases} B & \text{hvis } -1 \leq x \leq 1 \\ 0 & \text{ellers} \end{cases}$$

    Find foldningen $f(x) \star g(x)$. Find derefter korrelationen $f(x) \circ g(x)$.

23. Lad
$$f(x) = \begin{cases} x/2 & \text{hvis } -1 \le x \le 0 \\ 1 - x/2 & \text{hvis } 0 \le x \le 1 \\ 0 & \text{ellers} \end{cases} \tag{1.237}$$
og lad $g(x)$ være et impulstog med impulsafstand T. Find først foldningen $f(x) \star g(x)$ for T = 2, og dernæst for T = 1.5.

24. Lad
$$f(x) = \begin{cases} A & \text{hvis } -1 \le x \le 1 \\ 0 & \text{ellers} \end{cases}$$
$$g(x) = \begin{cases} -1 & \text{hvis } x < 0 \\ 1 & \text{hvis } 0 < x \\ 0 & \text{hvis } x = 0 \end{cases}$$
Find foldningen $f(x) \star g(x)$.

25. Lad
$$f(x) = \begin{cases} A & \text{hvis } -T \le x \le T \\ 0 & \text{ellers} \end{cases}$$
$$g(x) = \begin{cases} 0 & \text{hvis } x < -1 \\ -1 & \text{hvis } -1 \le x < 0 \\ 0 & \text{hvis } x = 0 \\ 1 & \text{hvis } 0 < x \le 1 \\ 0 & \text{hvis } x > 1 \end{cases}$$
Find foldningen $f(x) \star g(x)$.

26. Vis at den fouriertransformerede af en lige funktion er reel.

27. Bestem den fouriertransformerede af $\delta(x - x_0)$. Bestem den fouriertransformerede af $\cos(2\pi x u_0)$.

28. En reel funktion $f(x)$ kan dekomponeres i en sum af en lige og en ulige funktion. Vis først at $f_{lige} = \frac{1}{2}[f(x) + f(-x)]$ og $f_{ulige} = \frac{1}{2}[f(x) - f(-x)]$. Vis dernæst at $\mathbf{f_{lige}} = Re(\mathbf{f})$ og $\mathbf{f_{ulige}} = iIm(\mathbf{f})$.

29. Vis at fouriertransformationen af autokorrelationen $f(x) \circ f(x)$ af $f(x)$ er lig styrkespektret $|\mathbf{f}(u)|^2$.

30. Lad Γ være en lukket kurve i et digitalt billede, dvs. en mængde af pixels, der hver har præcis to naboer tilhørende Γ. Diskutér hvorledes en fourierrepræsentation af kurven kan dannes.

31. Lad $f(x)$ være en diskrete funktion defineret på intervallet $[0, 7]$ ved følgende værdier: $(0, 2, 3, 4, 4, 2, 2, 1)$. Beregn den diskret fouriertransformerede af f, samt styrkespektret for f.

32. Lad $f(x)$ være en følge af N reelle tal og lad $\mathbf{f}(u)$ være den diskret fouriertransformerede af f. Vis *Parsevals sætning for den diskrete fouriertransformation*:

$$\sum_{x=0}^{N-1} |f(x)|^2 = \frac{1}{N} \sum_{u=0}^{N-1} |\mathbf{f}(u)|^2 \tag{1.238}$$

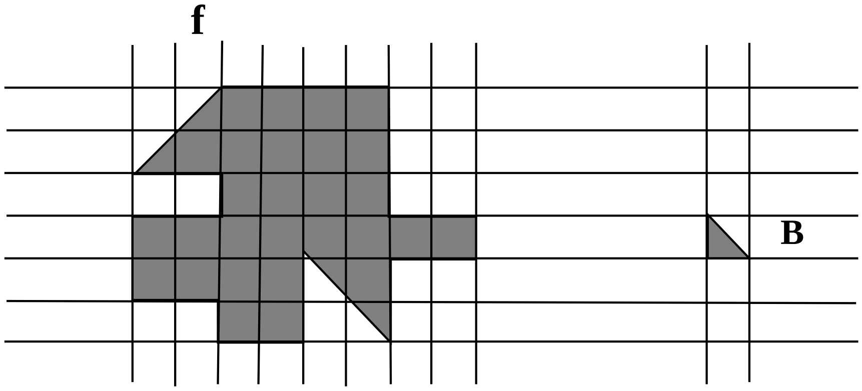
33. Skriv din egen implementation af FFT-algoritmen, og anvend den på et to-dimensionalt billede.

34. Skriv din egen implementation af *det inverse filter* og undersøg eksperimentelt dets styrker og svagheder.

35. Skriv din egen implementation af *wienerfilteret* og undersøg eksperimentelt dets styrker og svagheder.

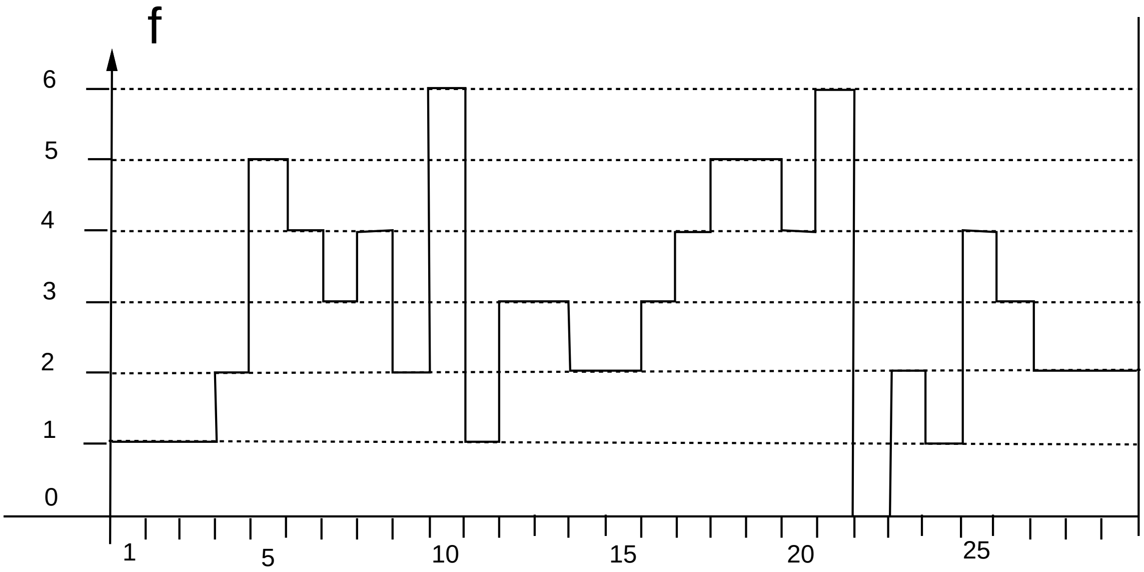
36. Skriv din egen implementation af *wallis-filtret* og undersøg eksperimentelt betydningen af de enkelte parametre.

37. Lad $f(x, y)$ være en binær figur som illustreret herunder. Lad B være et binært strukturelement med Origo i nederste venstre hjørne (ved den rette vinkel). Find nu:



a) $f \oplus B$ og $f \ominus B$
b) $f \bullet B$ og $f \circ B$
c) $(f \circ B) \bullet B$
d) $(f \oplus B) - (f \ominus B)$
Beskriv resultaterne af filtreringerne.

38. Lad $f(x)$ være ovenstående diskrete en-dimensionale funktion (med heltallige værdier): Lad filteret (stukturelementet) B bestemt ved:



$$B(x) \ = \ \text{if} \ \ x \ \in [-1, 0, 1] \ \ \text{then} \ \ 1 \ \ \text{else} \ \ 0$$

Filtrér funktionen f ved:
a) $f \oplus B$ og $f \ominus B$
b) $f \bullet B$ og $f \circ B$
c) $(f \oplus B) \ - \ (f \ominus B)$
d) $(f \bullet B) \ - \ (f \circ B)$
Beskriv resultaterne af filtreringerne.

39. Udled en Gauss-Seidel opdatering for løsning af det ligningssystem der fremkommer ved rekonstruktion vha. regularisering af en todimensional diskret funktion $f(x, y)$ ud fra et sæt af *spredte* målinger $c(x, y)$ for $(x, y) \in C$, og ved brug af glathedsfunktionalen: $G(f) \ = \ \int\int(f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2)dxdy$

40. Implementér den simple udgave af anisotrop diffusion, der er beskrevet i noterne. Sammenlign resultaterne ved anvendelse af metoden med en 1) Gauss-filtrering, 2) Nagao-filtrering og 3) Median-filtrering.