# Kapitel 2

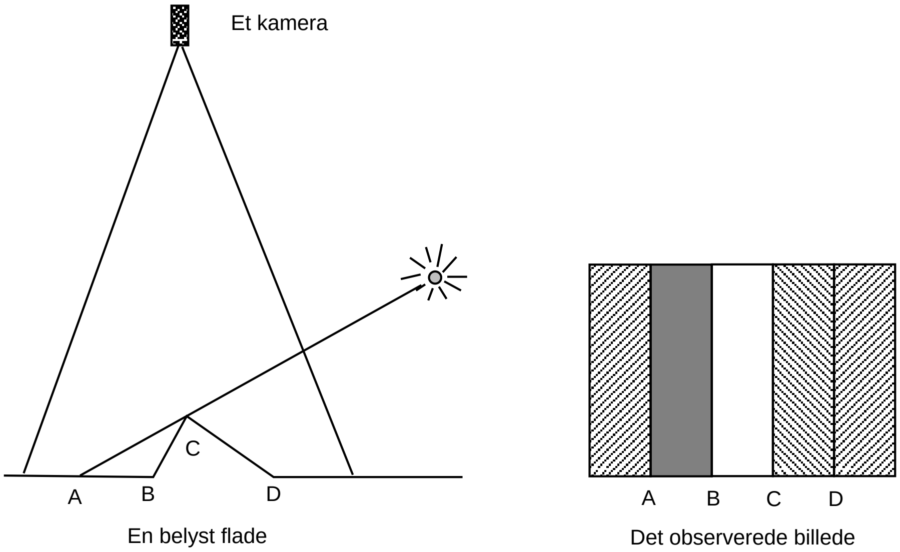
# Detektion af kanter i billeder

Detektion af kanter i et billede er en meget hyppig operation i mange industielle billedbehandlingsopgaver såvel som i metoder fra datamatsyn. Uformelt defineres billedkanterne ved en mængde af billedpunkter hvor kontrasten er stor, dvs. hvor intensitetsfunktionen markant ændrer værdi. Kanterne angiver således hvor i billedet "der sker noget". Disse er derfor (i visse sammenhænge) en god beskrivelse af billedet.

Kanterpunkterne i et billede vil typisk svare til projektioner af scenepunkter hvor scenematerialet ændrer reflektionsegenskaber, hvor scenefladen markant ændrer orientering, eller hvor illuminationen af scenen varierer (skygger). Billedkanter kan overordnet klasssificeres som forårsaget af topografiske eller fotometriske ændringer. Det er ikke umiddelbart muligt at klassificere de observerede kanter i et billede efter de "begivenheder" der har forårsaget kanterne.

Detektion af kanter er et modelproblem i billedbehandling. Metoder til kantdetektion er blevet udviklet over de sidste 30 år, og nye metoder kommer stadig til. Kantdetektion er en svær opgave fordi det ikke er veldefineret hvad en kant er, samt fordi man ofte ønsker uforenelige egenskaber ved en kantdetektor. I afsnit 2.1 diskuteres nærmere hvilke modeller af kanter der typisk anvendes. I afsnit 2.2 introduceres problemerne knyttet til støj i billeder, herunder hvorfor kantdetektion er en ustabil operation. I afsnit 2.3 og 2.4 introduceres nogle fundamentale metoder i kantdetektion. I afsnit 2.5 beskrives nogle metoder til markering af kantpunkter ud fra estimater af de partielle afledede af intensitetsfladen. Desuden gives en ultrakort indføring i differentialgeometri. Asnit 2.6 beskriver en såkaldt skelettransformation hvorved initielt detekterede "fede" kanter kan udtyndes. I afsnit 2.7 til 2.12 beskrives nogle mere avancerede metoder til kantdetektion, herunder en ultrakort introduktion til variationsregning. Afsnit 2.13 beskriver nogle metoder til kodning og repræsentation af kanter. I afsnit 2.14 diskuteres nogle metoder til detektion af rette liniestykker. Endelig indeholder afsnit 2.15 en beskrivelse af Hough-transformationen.

Ofte vil man kræve af en kantdetektor at den kan markere positioner, der intuitivt opfattes som kanter, at kun disse positioner markeres, samt at lokaliseringen er nøjagtig. Detektoren ønskes altså såvel robust som nøjagtig. Det kan vises at disse ønsker ikke kan opfyldes (optimalt) samtidig. Som vi senere skal se er det imidlertid muligt at konstruere detektorer hvor afvejningen mellem robusthed og nøjagtighed kan specificeres direkte ved




Figur 2.1: Til venstre ses en flade hvor punkterne B og D markerer materialeændringer. Punktet A markerer en skyggekant. Punkteterne B, C, og D markerer alle en ændring i fladeorienteringen. Til højre er der illustreret et billed af fladen.

en parameter.

## 2.1 Kantmodeller

Når vi taler om kanter i et billede, og vil udvikle en metode til detektion af disse, er det vigtigt eksplicit at formulere sin model af en kant. Typisk er modeller af kanter 1-dimensionale. Dette svarer til at betragte et snit af intensitetsfladen i intensitetsgradientens retning. Formålet med antagelsen er at simplificere modellen. Antagelsen implicerer at kanter i et billede er lokalt retlinede. Dette er naturligvis ikke altid tilfældet. Mange kantdetektorer har derfor også problemer med at lokalisere kanter når disse krummer (som ved et hjørne). I det følgende antages at kanterne er lokalt retlinede. Vi vil derfor kun formulere modeller af kanter for signaler. Problemet er således at formulere en enkel men tilstrækkelig generel model for hvorledes intensitetsværdien ændrer sig i en omegn af kanten. Antag at en en-dimensionel kant med amplitude $A$ er lokaliseret i $x_0$. Intensitetsfladens opførsel i en lille omegn om $x_0$ kan da modelleres ved:

$$E(x) \;=\; Au(x - x_0) \;+\; B \tag{2.1}$$

hvor funktionen $u(x)$ angiver kantmodellen, translateret således at kanten er lokaliseret i $x = 0$ og skaleret til intensitetsintervallet [0;1]. Den mest brugte model svarer til antagelsen at ændringen i intensitetsværdi sker diskontinuert som ved et trappetrin. Modellen kaldes derfor for trappetrinsmodellen

$$u(x) = \begin{cases} 0 & \text{hvis } x < 0 \\ 1 & \text{hvis } x \geq 0 \end{cases} \tag{2.2}$$

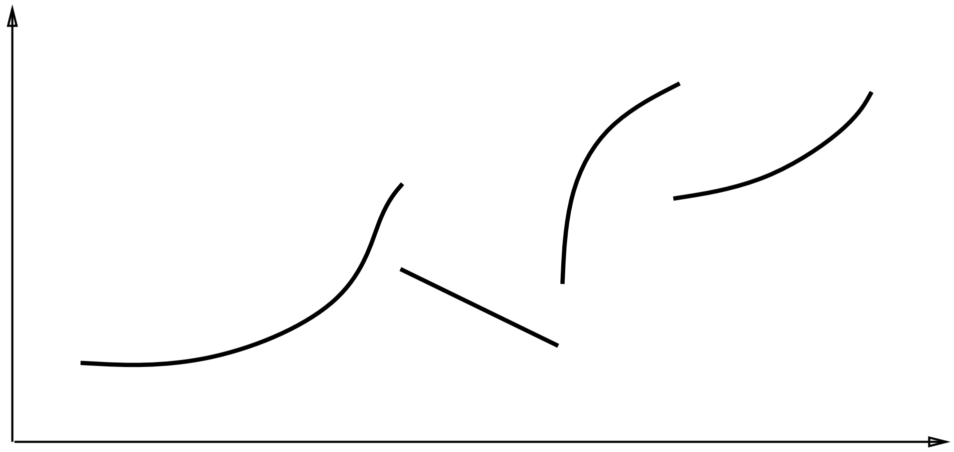
Typisk antages (som i ovenstående ligning) endvidere at intensitetsværdierne på hver side af kanten er lokalt konstante. Den primære fordel ved modellen er at den er simpel. Modellen har ingen parametre. Dette har betydning ved en teoretisk udledning af en optimal kantdetektor. I en anden kantmodel antages, at intensitetsændringen over kanten kan beskrives ved en rampe.

$$u(x) = \begin{cases} 0 & \text{hvis } x \leq \text{-t} \\ \frac{1}{2}(1+\frac{x}{t}) & \text{hvis } -t < x < t \\ 1 & \text{hvis } x \geq \text{t} \end{cases} \tag{2.3}$$

Denne kantmodel er mere kompliceret idet dennes udstrækning (2t) (og dermed ramppens hældningen $\frac{1}{2t}$) har betydning ved en detektion. En tredie kantmodel kan være en parametriseret arcus-tangens funktion. I denne model angiver parameteren $\tau$ hældningen af kantprofilen i kantpunktet.

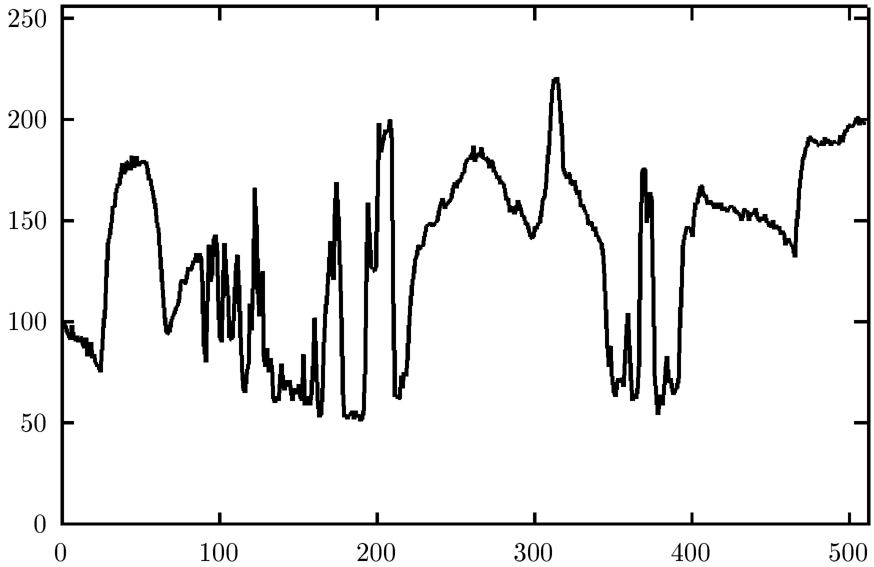
$$u(x) = \arctan(\tau x) \tag{2.4}$$

Som eksempel på en mere generel model af en kant er en metode at modellere intensitetsfladen på "højre" hhv. "venstre" side af kanten separat f.eks. ved brug af anden eller tredie ordens polynomier. I sidstnævnte tilfælde vil modellen af en kant således involvere 8 parametre, og vil kunne modellere kanter som i nedenstående illustration.



Figur 2.2: Illustration af nogle af de mere komplekse kanter, som kan forekomme i billeder.

Hvilken kantmodel man vælger må naturligvis afhænge af hvorledes kantprofilerne i de billeder man ønsker at behandle ser ud, samt af formålet med kantdetektionen. Generelt vil et billede ikke indeholde kanter med kun én profilform. Man står sig dog ofte ved at vælge en forholdsvis generel model, der ikke indeholder for mange parametre. Problemet er, at hvis man opererer med flere kantmodeller, kan det være svært at afgøre hvilken model, der i en gives situation skal anvendes. Spørgsmålet om valg af kantmodel er såvel teoretisk uafklaret, såvel som det er uafklaret hvilke kantprofiler der faktisk forekommer i billeder optagne under forskellige forhold. Nedenstående billede viser intensitetsvariationen i skanlinie nummer 300 i billedet af *Lenna*. Som det ses, kan antagelsen om en bestemt af de ovennævnte kantmodeller diskuteres.




Figur 2.3: En skanlinie taget fra et portrætbillede

## 2.2 Støjproblemet

Nogle teoretisk specificerede kantoperatorer gør ikke klart hvorvidt operatoren forudsættes anvendt på støjfri billeder, eller hvilke antagelser der gøres om en eventuel støj. Sædvanligvis indeholder de fleste billeder støj. Denne kan være tilført billedet under optagelsen eller ved transmition af billedet. Det er også muligt at støjen kan være tilført inden billedet er optaget, f.eks. ved at lysstrålerne har passeret gennem et inhohogent medie før de registreres af sensoren.

Uformelt beskriver kanterne i et billede de punkter hvor intensitetsfunktionen markant ændrer værdi. I disse punkter må intensitetsgradienten derfor være numerisk stor. Dette indikerer at detektion af kanter til en vis grad afhænger af målinger af de partielle afledede af intensitetsfunktionen.

Betragt et (kontinuert) signal $f(x)$, og antag at det observerede signal $g(x)$ fremkommer af $f(x)$ ved addition af et støjled $\eta(x)$, der her approximeres ved en højfrekvent sinusbølge med lille amplitude.

$$g(x) = f(x) + \beta sin(\omega x) \tag{2.5}$$

Ved differentiation af $g(x)$ ses, at selv om støjens amplitude er lille, vil støjledet have en dominerende indflydelse på den afledede, fordi $\omega$ er stor.

$$g'(x) = f'(x) + \beta\omega cos(\omega x) \tag{2.6}$$

*Differentiation er således en ustabil operation.* Ved højere ordens afledede øges graden af ustabilitet. Da kantdetektion jvf. den foregående diskussion er knyttet til måling af partielle afledede ses, at også kantdetektion er en ustabil operation. Det er ofte nødvendigt at foretage en støjfiltrering før (eller samtidig med) anvendelsen af differentialoperatorer.

Sædvanligvis antages at støjen er ukorreleret. I modsat fald vil det være nødvendigt at foretage en støjfiltrering, der er konstrueret til at fjerne netop den observerede støj. Det antages som regel også at støjen har middelværdi nul. Fordelingsfunktionen for støjen er afhængig af den process hvorved støjen er tilført billedet. En rimelig og generel antagelse er at støjen er normalfordelt. Dette svarer til at støjen er tilført billedet ved addition af mange små uafhængige støjbidrag, hvis individuelle fordelinger er vilkårlig.

I forbindelse med differentiation skal man yderligere være opmærksom på at der ved diskretiseringen af billedet er elimineret høje frekvenser i fourierspektret for det "oprindelige" kontinuerte billede. Yderligere er billedet kvantificeret til typisk 256 forskellige intensitetsværdier. Sidstnævnte forhold gør det problematisk at anvende en streng matematisk definition af en kant som en diskontinuitet (nemlig at grænseværdierne fra højre hhv. venstre er forskellige). Endelig kan der være tale om egentlige kvantificeringsfejl.

## 2.3 Numeriske metoder til måling af afledede

I det følgende skal vi betragte nogle simple metoder til måling af afledede for en diskrete endimensional funktione $f$. Lad $h$ være et positivt heltal. De forlæns, baglæns og centrale differens approximationer med støtte $h$ til den første ordens afledede af *f(x)* er defineret ved:

$$f_x^{for}(x) = \frac{f(x+h)-f(x)}{h} \tag{2.7}$$

$$f_x^{bag}(x) = \frac{f(x)-f(x-h)}{h} \tag{2.8}$$

$$f_x^{cen}(x) = \frac{f(x+h)-f(x-h)}{2h} \tag{2.9}$$

hvilket fås af Taylor-rækkeudviklingerne for den til *f(x)* svarende kontinuerte funktion *u(x)*:

$$u(x+h) = u(x) + hu'(x) + \frac{h^2u''(x)}{2} + \frac{h^3u'''(x)}{6} + \cdots \tag{2.10}$$

$$u(x-h) = u(x) - hu'(x) + \frac{h^2u''(x)}{2} - \frac{h^3u'''(x)}{6} + \cdots \tag{2.11}$$

$$u(x+h) - u(x-h) = 2hu'(x) + 2\frac{h^3u'''(x)}{6} + O(h^4) \tag{2.12}$$

Det ses at forlæns og baglæns approximationerne tilnærmer de kontinuerte afledede med en fejl af orden $O(h)$ og at fejlen ved den centrale approximation er af orden $O(h^2)$. Ligeledes ses at den centrale approximation er lig middelværdien af den baglæns og den forlæns approximation. Sædvanligt benyttes den centrale approximation, fordi denne giver symmetrisk svar ved anvendelse på en ideel trappetrins-kant. Nøjagtigheden af approximationen er afhængig af støttestørrelsen $h$. Var *f(x)* kontinuert ville approximationen konvergere mod den afledede for $h$ gående mod 0. I det diskrete tilfælde kan vi ikke vælge $h$ mindre end 1. Dette vil være det optimale valg under den antagelse at vi faktisk ønsker af måle den afledede af det observerede signal *f(x)*.

I praksis er kontrastovergangen ved en kant sjældet diskontinuert. Ofte kan denne beskrives ved foldningen af en (ukendt) punktspredningsfunktion med en ideel trappertinskant (eksemplificeret ved et billede optaget med dårlig fokusering på cenen). Desuden vil signalet/billedet ofte indeholde støj. Endelig er billedet diskretiseret og billedværdierne kvantificeret. Det er derfor nødvendigt at forholde sig til den *skala* hvorover intensitetsændringen sker, samt til mængen af støj. Begge forhold relaterer til opløsningen i gråtoneværdier.

Vi er som regel interesseret i at måle den afledede af det "oprindelige" signal, og ikke af det observerede signal. Ved direkte anvendelse af ovenstående differensapproximationer på signalet kan bestemmelsen blive unøjagtig eller helt forkert. En simpel metode til at øge robustheden (men ikke nøjagtigheden) af målingen er at øge støttestørrelsen $h$ (Overvej dette). En anden metode er at støjfiltrere signalet før eller evt. samtidig med anvendelsen af differensapproximationen. Anvendes eksempelvis den centrale differensapproximation med $h$ lig 1, på en boxfiltrering med støtte 3 af signalet, fås den samlede approximation til:

$$f_x^{boxC} = \frac{f(x+2) + f(x+1) - f(x-1) - f(x-2)}{6} \tag{2.13}$$

Generelt vil filtret blive mere robust jo kraftigere de høje frekvenser i *g(x)* bliver dæmpet ved lavpassfiltreringen. Dette sker på bekostning af at det lokale maksima (i filtersvaret for en kant) bliver mere afrundet (dårligere lokaliseret) og vil få mindre amplitude. Generelt vil det ikke være muligt at konstruere noget filter til måling af afledede, som både er robust og i stand til at lokalisere kanter nøjagtigt. Vi skal senere se nærmere på dette forhold. Det vel nok sværeste problem i kantdetektion er at bestemme det støjfilter, der er optimalt mht. samtidig at øge robustheden og lokaliseringsevnen. Udledningen af et sådant filter afhænger naturligvis af mere præcise definitioner for robusthed og lokaliseringsevne. I afsnit 2.9 skal vi se et eksempel på et sådant optimalt filter.

## 2.4 Beaudet operatorer

I dette afsnit beskrives en mere generel metode til udledning af filtre for differensapproximationer til de partielle afledede af et billede. Filtrene er todimensionale og relativt enkle at udlede. I modsætning til differensapproksimationerne er det muligt at angive et mål (støttestørrelsen) for robustheden af filtrene.

Beaudet operatorene af N'te orden med kvadratisk støtte $(2M+1)\times(2M+1)$ er defineret som de lineære positionsinvariante operatorer der, foldet med billedet, svarer til det bedste fit (i mindste kvadraters betydning) til intensitetsfladen med et n'te ordens polynomie. Anderledes sagt: For hver punkt $P$ i billedet indlægger vi et *lokalt koordinatsystem* $(P, x, y)$, og tilnærmer intensitetesfladen $E(x,y)$ med et polynomie *z(x,y)* defineret ved:

$$\begin{aligned} z(x,y) &= \sum_{i,j=0}^{\infty} \frac{\partial^{i+j}}{\partial x^i \partial y^j} E(x,y) x^i y^j \\ &= a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2 + \cdots \end{aligned} \tag{2.14}$$

Den reelle værdi $a_0$ tilnærmer således intensiteten i $P$. Til den *i'te* koefficient $a_i$ svarer en Beaudet operator, der foldet med intensitetsfunktionen i en omegn om $P$ giver værdien $a_i$. Opgaven er at bestemme $a$'erne i Taylor rækkeudviklingen for E(x,y). Hvert $a_i$ beskriver et filter $\mathbf{A_i}$. Bestemmelsen af $a_0$ vil eksempelvis give os et filter til estimation af intensitetsværdien i $(x,y) = (0,0)$

For N'te ordens operatorer afskæres rækken efter det N'te ordens led. Vi begår således en trunkeringsfejl. I det følgende ignoreres denne fejl. De søgte koefficienter bestemmes ved at betragte kvadratafvigelsen (kvadratfejlen) $e^2$ givet ved:

$$e^2 = \sum_{x=-M}^{M} \sum_{y=-M}^{M} [E(x,y) - z(x,y)]^2 \tag{2.15}$$

Opgaven, at finde de $a$'er, der minimerer kvadratfejlen $e^2$, løses lettest ved at opskrive udtrykket på matriceform. For at lette notationen i opstillingerne nummereres hvert punkt i støtteområdet med et tal $\in [1:n]$ hvor $n = [2M+1]^2$. For såvel koordinatværdier og intensiteter angiver indices en reference til et punkt. Nummererringen foretages søjle for søjle i støtteområdet (fra lille koordinatværdi mod større).

$$\mathbf{E} = \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_{n-1} \\ E_n \end{pmatrix} \qquad \mathbf{A} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{coeff(N)} \end{pmatrix} \tag{2.16}$$

hvor $coeff(N) = \frac{N(N+3)}{2} + 1$ er antallet af koefficienter i et N'te ordens polynomium. Søjlematricen $\mathbf{E}$ indeholder alle intensitetsværdierne inden for støtteområdet.

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 & y_1 & x_1 y_1 & x_1^2 & y_1^2 & \cdots \\ 1 & x_2 & y_2 & x_2 y_2 & x_2^2 & y_2^2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n-1} & y_{n-1} & x_{n-1} y_{n-1} & x_{n-1}^2 & y_{n-1}^2 & \cdots \\ 1 & x_n & y_n & x_n y_n & x_n^2 & y_n^2 & \cdots \end{pmatrix} \tag{2.17}$$

Med ovenstående definitioner kan vi nu skrive formel (2.15) som en matrixligning:

$$e^2 = \|\mathbf{E} - \mathbf{X}\mathbf{A}\|^2 \tag{2.18}$$

Den vektor $\mathbf{A}$ der minimerer $e^2$ fås ved differentiation:

$$\frac{\partial e^2}{\partial \mathbf{A}} = -2\mathbf{X}^t\left(\mathbf{E} - \mathbf{X}\mathbf{A}\right) = 0 \tag{2.19}$$

hvor t angiver transponering. Videre fås:

$$\mathbf{X}^t\mathbf{E} = \mathbf{X}^t\mathbf{X}\mathbf{A} \tag{2.20}$$

og dermed

$$\mathbf{A} = \mathbf{X}^\sharp\mathbf{E} = \left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{E} \tag{2.21}$$

hvor matricen $\mathbf{X}^\sharp$ kalden den *pseudoinverse* til matricen $\mathbf{X}$ (eller "den generaliserede Penrose-More inverse"). Bemærk at den pseudoinverse kun indeholder talværdier. Opskrives resultatet på sædvanlig form, ses at de aktuelle værdier af $\mathbf{A}$ fås ved foldning af funktionen E(x,y) med den (de) funktion(er) hvis matrice(r) er givet ved rækkerne i $\mathbf{X}^\sharp$.

### Eksempel

Som eksempel betragtes her udledningen af Beaudet operatorene af første orden for M = 1. I dette tilfælde er modellen givet ved: $z(x,y) = a_0 + a_1x + a_2y$. Matricen $\mathbf{X}$ får udseendet:

$$\mathbf{X} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \tag{2.22}$$

hvor anden og tredie søjle angiver $x$ hhv. $y$ -koordinatværdier. Videre fås:

$$\mathbf{X}^t\mathbf{X} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} \tag{2.23}$$

Ved matriceinvertering -og multiplikation fås at:

$$\mathbf{X}^\sharp = \frac{1}{18}\begin{pmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ -3 & 0 & 3 & -3 & 0 & 3 & -3 & 0 & 3 \\ -3 & -3 & -3 & 0 & 0 & 0 & 3 & 3 & 3 \end{pmatrix} \tag{2.24}$$

Dekomponeres denne matrice svarende til $a_0, a_1$ *og* $a_2$ får vi:

$$\mathbf{A_0} = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad \mathbf{A_1} = \frac{1}{6}\begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \qquad \mathbf{A_2} = \frac{1}{6}\begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \tag{2.25}$$

Jo større støtten for Beaudetoperatorene vælges, jo mere robuste vil disse angive de partielle afledede (ved foldning med intensitetsfladen). Tilsvarende vil nøjagtigheden af estimaterne blive mindre. Bemærk endvidere at Beaudetoperatorer (med samme støtte) vil være forskellige afhængigt af ordenen af det approximerende polynomie. Således er masken svarende til middelværdi-operatoren $A_0$ for de 2. ordens operatorer med støtte $3 \times 3$ givet ved:

$$\mathbf{A_0} = \frac{1}{9}\begin{pmatrix} -1 & 2 & -1 \\ 2 & 5 & 2 \\ -1 & 2 & -1 \end{pmatrix} \tag{2.26}$$

Læseren opordres til at eftervise nedenstående 2. ordens operatorer med støtte $5 \times 5$ (brug et "matematikprogram" som f.eks. **maple**, **mathlab** eller **matematica**).

$$\mathbf{A_0} = \frac{1}{175}\begin{pmatrix} -13 & 2 & 7 & 2 & -13 \\ 2 & 17 & 22 & 17 & 2 \\ 7 & 22 & 27 & 22 & 7 \\ 2 & 17 & 22 & 17 & 2 \\ -13 & 2 & 7 & 2 & -13 \end{pmatrix} \qquad \mathbf{A_1} = \frac{1}{50}\begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \end{pmatrix} \tag{2.27}$$

samt

$$\mathbf{A_2} = \frac{1}{50}\begin{pmatrix} -2 & -2 & -2 & -2 & -2 \\ -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 \end{pmatrix} \qquad \mathbf{A_4} = \frac{1}{35}\begin{pmatrix} 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ 2 & -1 & -2 & -1 & 2 \end{pmatrix} \tag{2.28}$$

og

$$\mathbf{A_5} = \frac{1}{35}\begin{pmatrix} 2 & 2 & 2 & 2 & 2 \\ -1 & -1 & -1 & -1 & -1 \\ -2 & -2 & -2 & -2 & -2 \\ -1 & -1 & -1 & -1 & -1 \\ 2 & 2 & 2 & 2 & 2 \end{pmatrix} \qquad \mathbf{A_3} = \frac{1}{100}\begin{pmatrix} 4 & 2 & 0 & -2 & -4 \\ 2 & 1 & 0 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -1 & 0 & 1 & 2 \\ -4 & -2 & 0 & 2 & 4 \end{pmatrix} \tag{2.29}$$

Beaudetoperatorerne er beskrevet i *P. R. Beaudet: Rotationally Invariant Image Operators; Int. joint Conf. on Pattern Recognition, 1978.*

## 2.5 Detektionsmetoder

Som nævnt i et foregående afsnit er detektion af kanter knyttet til måling af de partielle afledede af intensitetsfunktionen. Lad $\nabla E = (E_x, E_y)$ være intensitetsgradienten, $\rho = ||\nabla E||$ denne vektors størrelse, og lad $\phi = arctan(E_y/E_x)$ være intensitetsgradientens retning. Bemærk at intensitetsfladens stejlhed i en vilkårlig retning $v$ er bestemt ved:

$$E_v \; = \; E_x \cos(v) \; + \; E_y \sin(v) \tag{2.30}$$

Den simplest tænkelige detektionsmetode er at definere kantpunkterne i et billede ved de punkter (x,y) hvor $\rho$ overstiger en vis (foruddefineret) værdi T. Dette kaldes for detektion ved grænseværdiafskæring.

$$edge(x,y) = \begin{cases} 0 & \text{hvis } \rho < T \\ 1 & \text{hvis } \rho \geq T \end{cases} \tag{2.31}$$

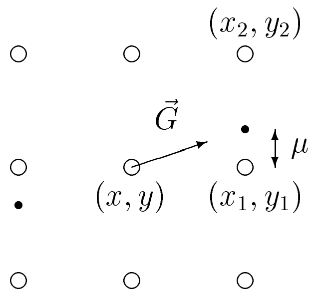
Metoden vil markere alle punkter hvor intensitetsfladen er stejlere end T. Disse punkter vil ofte udgøre fede kanter (afhængig af valget af T), men ofte vil billedpunkter, der intuitivt klart ses som kanter ikke markeres. Kvaliteten af metoden er derfor meget afhængig af et godt valg af T. Ofte vil det ikke være muligt at vælge T sådan at alle intuitivt observerede kanter markeres, og sådan at de markerede kanter er acceptabelt tynde.

Bemærk at vi ikke lagde os fast på nogen bestemt norm da vi definerede $\rho$. Sædvanligvis benyttes den euclidiske norm, fordi denne er rotationssymmetrisk. I situationer hvor horisontale og vertikale kanter har særlig preference kunne man i stedet bruge den beregningsmæssigt hurtigere maksimumsnorm. I det følgende antages at den euclidiske norm benyttes.

I forhold til ovennævnte detektionsmetode er en mere avanceret metode (kun) at markere kanter ved de punkter P hvor $\rho$ lokalt antager maksimal værdi (eng. non maximum suppression). Dette gøres ved at sammenligne $\rho$-værdien i P med $\rho$-værdierne i to punkter i en vis afstand (f.eks. 1 pixel) fra P. De to punkter skal vælges i intensitetsgradientens positive og negative retning. P markeres som kantpunkt hvis $\rho$ i P er større end $\rho$ i begge de andre punkter. Nedenstående C-programstump, der tænkes udført for alle billedpunkter illustrerer denne metode. Her angiver G gradientstørrelsen $\rho$. Gx hhv. Gy angiver estimerede mål af de partielle afledede af intensiteten.

```
/* do non maximum suppression in (x,y) */
#define round  (1.0 - sqrt(2.0)/2.0)
if ((m = G[x][y]) > THreshold)
{
   gx = Gx[x][y]/m;
   gy = Gy[x][y]/m;

   /*point on axis*/
   x1 = (gx > 0.0) ? (int)(gx + round) : (int)(gx - round);
   y1 = (gy > 0.0) ? (int)(gy + round) : (int)(gy - round);

   /*point on diagonal*/
   x2 = (gx > 0.0) ? 1 : -1;
```

```
   y2 = (gy > 0.0) ? 1 : -1;

   gxm = fabs(gx);
   gym = fabs(gy);
   my  = (gym < gxm) ? gym/gxm : ((gym > gxm) ? gxm/gym : 0.0);

   m1 = (1.0-my)*G[x+x1][y+y1]  + my*G[x+x2][y+y2];
   m2 = (1.0-my)*G[x-x1][y-y1]  + my*G[x-x2][y-y2];
   if ((m > m1) && (m > m2)) ''MARK LOCAL MAXIMA''
}
```




Figur 2.4: Interpolation ved detektion af lokale maksima

Bemærk at i ovenstående skitse angiver værdierne m1 og m2 interpolerede størrelser, der "måles" på randen af kvadratnettet hvor billedpunkterne er knuder. Dette er illustreret i Figur 2.4.

I ovennævnte metode benyttes en simpel grænseværdiafskæring i afgørelsen om et kantpunkt er tilstrækkeligt markant til at skulle medtages. Denne metode resulterer ofte i at en kant bliver opsplittet i mange korte kantstykker. I en alternativ metode, kaldet *grænseværdiafskæring med hysterese* er ideen at registrere alle kantpunkter hvor gradientstørrelsen enten er over en vis nedre værdi eller hvis kantpunktet ligger på en kant hvor gradientstørrelsen i mindst et kantpunkt er større en end øvre værdi. Metoden er overlegen i forhold til metoder hvor der kun benyttes en grænseværdi, fordi metoden bevarer kanter hvor gradientstørrelsen varierer meget. Ulemper ved metoden er, at den kræver angivelse af to værdier, og at den forudsætter en segmentering af kanterne i billedet. Segmenteringen kan foretages løbende, sådan at sammenhængende kanterpunkter gennemløbes og markeres som besøgte. Under gennemløbet "slettes" punkter hvor gradientstørrelsen er mindre end den nedre grænseværdi, og positionerne af de øvrige kantpunkter registreres. Desuden registreres den største gradientstørrelse. Ved afslutning af gennemløbet undersøges om denne værdi er større end den øvre grænse. I bekræftende fald markeres alle registrerede kantpunkter som accepterede. I modsat fald markeres de som forkastede.

En anden metode til detektion af kantpunkter er at betragte de anden ordens partielle afledede af E. For et signal vil ekstrema i den førsteafledede svare til en nulpunktsgennemskæring (eng. zero crossing, ofte forkortet zc) i den andenafledede. Problemet i det todimensionale tilfælde er primært hvilken kombination af de andenafledede (for E) der er fornuftigt at betragte. Mest simpelt kan kantpunkter defineres ved nulpunktsgennemskæringer af laplaceoperatoren:

$$\nabla^2 E = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)E = M = E_{xx} + E_{yy} \tag{2.32}$$

Laplaceoperatoren er rotationssymetrisk og lineær (i de afledede), hvilket betyder at dens værdi på et billede kan beregnes ved foldning af dette med eet filter. Man kunne også forestilles sig en ikke-lineære kombination af de andenafledede, f.eks. ved måling af Gausskrumningen af intensitetsfladen E:

$$K = E_{xx}E_{yy} - E_{xy}^2 \tag{2.33}$$

En tredie mulighed kunne være at definere kantpunkter som nulpunktsgennemskæringer af den andenafledede i retning givet ved intensitetsgradienten.

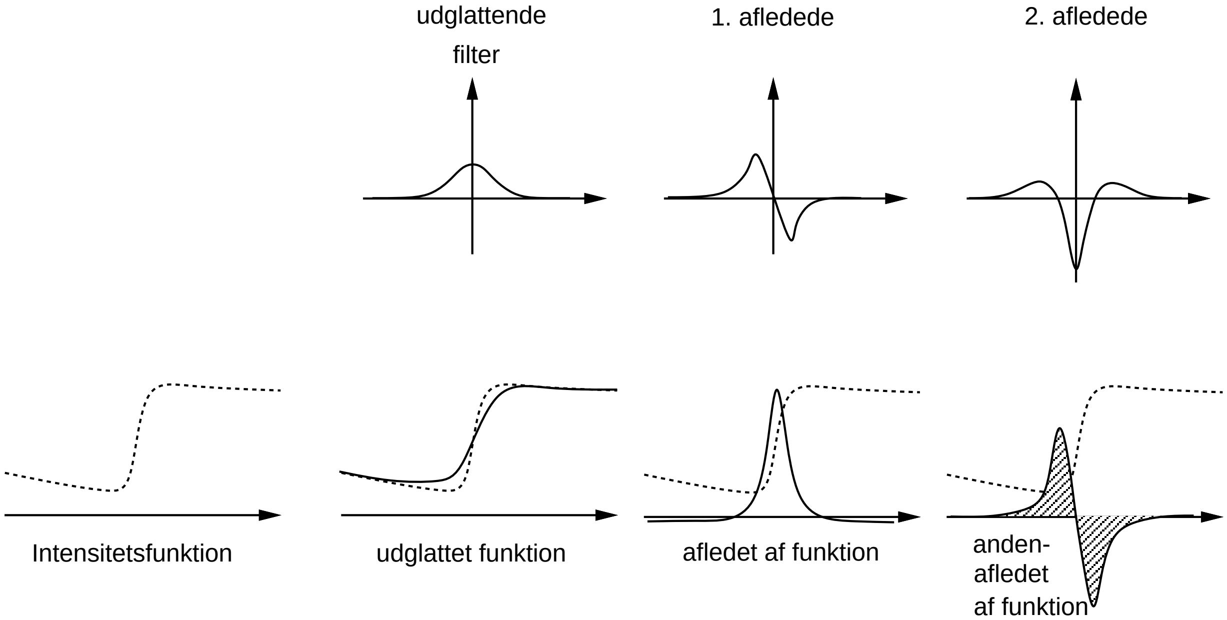
Detektion af nulpunktsgennemskæringer er relativt lette. Nulpunktsgennemskæringer i f(x,y) kan eksempelvis markeres hvis:

- f(x-1,y) $\cdot$ f(x+1,y) $<$ 0
- f(x,y) = 0 og [ f(x-1,y) $\neq$ 0 eller f(x+1,y) $\neq$ 0 ]
- f(x,y-1) $\cdot$ f(x,y+1) $<$ 0
- f(x,y) = 0 og [ f(x,y-1) $\neq$ 0 eller f(x,y+1) $\neq$ 0 ]

Valget mellen første- og anden ordens differentialoperatorer er tildels et spørgsmål om smag. Fordelen ved måling af førsteafledede er at disse kan bestemmes mere stabilt end andenafledede (jvf. afsnit 2.2). Fordelen ved anvendelse af andenafledede er at nulpunktsgennemskæringer er lettere at detektere end lokale ekstrema. Der findes særdeles udmærkede kantdetektorer af begge typer.

For detektorer baseret på de førsteafledede benyttes filterværdien i kantpunktet som indikation af detektionsstyrken. For detektorer baseret på de andenafledede kan stejlheden i nulpunktsgennemskæringen benyttes som mål for detektionsnøgagtigheden. Denne størrelse er klart relateret til den tredieafledede af intensiteten. Sikkerheden af detektionen (robustheden) er derimod knyttet til arealet af den numeriske værdi af det filtrerede signal (i en omegn om kantpunktet). Dette er illustreret i Figur 2.5. Begge størrelser skal naturligvis fortolkes ud fra viden om størrelsen af det udglattende filter. En mere detalieret diskussion af brugbarheden af disse mål falder uden for rammerne af disse noter.

Der findes kantdetektorer der ikke eksplicit involverer målinger af partielle afledede såvel som detektorer der involverer målinger af afledede af højere orden end 2. Vi skal ikke i disse noter beskrive sådanne metoder.




Figur 2.5: Detektion af kanter ved maxima hhv. nulpunktsgennemskæring af den første hhv. andenafledede af funktionen foldet med et udglattende filter.

### 2.5.1 Ekskurs i differentialgeometri

I ovennævnte afsnit blev to anden ordens differentialoperatorer, nemlig *Gausskrumningen* $K$ og Laplaceoperatoren eller *Middelkrumningen* $M$ fremhævet. Disse størrelser er ikke valgt tilfældigt, men er centrale i teorien for "kurver og flader i rummet". Enhver flade kan i ethvert punkt $P$ lokalt approksimeres med en paraboloide (den oskulerende paraboloide).

Antag at koordinatsystemet er lagt med begyndelsespunkt i $P$ således at $x-y$-planet er indeholdt i tangentplanet til fladen (i $P$). I dette system kan fladen fremstilles ved:

$$z \;=\; f(x,y) \tag{2.34}$$

Enhver plan indeholdende $z$-aksen vil skære fladen i et såkaldt *normalsnit*. Normalsnittes skæring med $x-y$-planet danner en ret linie $(x,y) \;=\; (\lambda\cos\theta, \lambda\sin\theta)$, hvor $\theta$ er vinklen mellem $x$-aksen og linien. Normalsnittet kan derfor fremstilles ved ligningen:

$$z \;=\; G(\lambda) \;=\; f(\lambda\cos\theta, \lambda\sin\theta) \tag{2.35}$$

For at finde krumningen af fladen langs normalsnittet dannes $G'(\lambda)$ og $G''(\lambda)$:

$$G'(\lambda) \;=\; f'_x(\lambda\cos\theta, \lambda\sin\theta)\cos\theta \;+\; f'_y(\lambda\cos\theta, \lambda\sin\theta)\sin\theta \tag{2.36}$$

$$\begin{aligned} G''(\lambda) \;=\;& f''_{xx}(\lambda\cos\theta, \lambda\sin\theta)\cos^2\theta \;+\; 2f''_{xy}(\lambda\cos\theta, \lambda\sin\theta)\cos\theta\sin\theta \\ &+\; f''_{yy}(\lambda\cos\theta, \lambda\sin\theta)\sin^2\theta \end{aligned} \tag{2.37}$$

Skrives

$$r \;=\; f''_{xx}(0,0)\;; \qquad s \;=\; f''_{xy}(0,0)\;; \qquad t \;=\; f''_{yy}(0,0) \tag{2.38}$$

kan vi opskrive normalsnittets krumning i punktet $P$:

$$\kappa_\theta \;=\; G''(0) \;=\; r\cos^2\theta \;+\; 2s\cos\theta\sin\theta \;+\; tsin^2\theta \tag{2.39}$$

Des ses let at krumningen langs alle normalsnit vil stemme overens med krumningen af den oskulerende paraboloide:

$$z \;=\; \frac{1}{2}[rx^2 \;+\; 2sxy \;+\; ty^2] \tag{2.40}$$

Ved at rotere $x - y$-koordinatsystemet omkring $z$-aksen kan man opnå at produktleddet $2sxy$ forsvinder. Herved opstår *Eulers formel*:

$$\kappa_\theta \;=\; \kappa_1\cos^2\theta \;+\; \kappa_2\sin^2\theta \tag{2.41}$$

hvor $\theta$ nu angiver vinklen mellem den ny $x$-akse og normalsnittet. Størrelserne $\kappa_1$ og $\kappa_2$ kaldes *hovedkrumningerne* i $P$, og er krumningerne for de til $\theta = 0$ og $\theta = \frac{\pi}{2}$ svarende normalsnit. Væsenligt er at tangentretningerne svarende til hovedkrumningerne er vinkelrette. Hovedkrumningerne bestemmes som rødderne i reduktionsdeterminanten:

$$\begin{vmatrix} r-\lambda & s \\ s & t-\lambda \end{vmatrix} \tag{2.42}$$

og tilfredsstiller ligningerne:

$$\kappa_1 \;+\; \kappa_2 \;=\; M == \; r+t\;, \quad og \quad \kappa_1\kappa_2 \;=\; K \;=\; rt-s^2 \tag{2.43}$$

Heraf ses at størrelserne $M$ og $K$ er rotationsinvatiante, dvs. at målene er uafhængige af hvorledes det benyttede koordinatsystem er lagt. Dette kan have meget stor betydning for fortolkningen af intensitetsfladens lokale opførsel.

Antag at $\kappa_1 \geq \kappa_2$ (ellers omdøb de to størelser). Af Eulers formel fås:

$$\kappa_\theta \;=\; \kappa_2 \;+\; (\kappa_1 - \kappa_2)\cos^2\theta \tag{2.44}$$

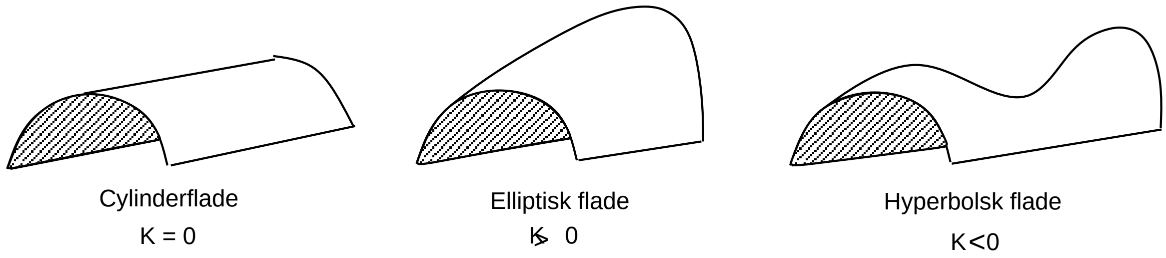
Heraf ses at der for alle $\theta$ gælder at: $\kappa_1 \geq \kappa_\theta \geq \kappa_2$. I det specialle tilfælde hvor $\kappa_1 = \kappa_2$ ses at $r = t$ og $s = 0$, hvilket svarer til at fladen lokal har kugleform. Mere generelt kan vi, som illustreret i Figur 2.6, klassificere fladeformen efter gausskrumningen $K$.

**Ellliptisk flade:** $K > 0$

Hovedkrumningerne $\kappa_1$ og $\kappa_2$ har samme fortegn. Den oskulerende paraboloide er derfor elliptisk, hvorfor punktet $P$ kaldes elliptisk. Fladen ligger, ligesom paraboloiden, helt på den ene side af tangentplanet.

**Hyperbolsk flade:** $K < 0$

Hovedkrumningerne $\kappa_1$ og $\kappa_2$ har modsat fortegn. Den oskulerende paraboloide er hyperbolsk, hvorfor punktet $P$ kaldes hyperbolsk. Fladen ligger på begge sider af tangentplanet. Paraboloiden skærer tangentplanet langs asymptoteretningerne for paraboloiden.




Figur 2.6: Klassifikation af flader i 3D.

**Parabolsk flade:** $K = 0$

En af hovedkrumningerne (eller begge) er lig nul. Hvis $\kappa_1 = \kappa_2 = 0$ er paraboloiden udartet til et plan (tangentplanet). Hvis kun en af hovedkrumningerne er lig nul beskriver den oskulerende paraboloide en cylinderflade. I dette tilfælde kaldes punktet $P$ for parabolsk.

En egenskab ved kanter defineret ved nulpunktsgennemskæringerne af gausskrumningen $K$ er at kanterne vil adskilde områder i intensitetsfladen, der er elliptiske ($K > 0$ hhv. hyperbolske ($K < 0$). Kanterne selv kaldes de parabolske kurver. Det har været diskuteret meget om denne klassifikation (segmentering) af intensitetsfladen er brugbar i datamatyn såvel som i en forståelse af det menneskelige syn. Diskussionen er ikke endt i noget klart svar.

### 2.5.2 Retningsspecifikke eller rotationssymetriske operatorer

Det har været diskuteret (i litteraturen) hvorvidt retningsspecifikke operatorer var at foretrække frem for rotationssymetriske. Diskussionen har været baseret på den tilgængelige viden om det menneskelige syn, idet antagelsen var at dette formodentlig anvendte den bedste metode. Fra neurofysiologiske studier af abehjerner vides, at disse (i den visuelle del af cortex) indeholder celler, der er meget følsomme over for retningsbestemte kanter. Se *Hubel and Wiesel: Brain Mechanism of Vision; Scientific American, september 1979* samt *Marr and Hildreth: Theory of Edge Detection; Proc. R. Soc. London B 207 1980.*

Fra et biologisk synspunkt er problemet, at hvis retningsopløsningen skal være god, da må detektionsmetoden involvere mange retningsbestemte filtre, hvilket kan være “dyrt i hardware”. Endvidere kan det være svært at kombinere inpulssvarene for disse filtre. Fra et datalogisk synspunkt er problemet enklere idet gradientretningen kan beregnes ud fra to filtre. Det er ikke nødvendigt at implementere de mange retningsspecifikke filtre.

Anvendelse af en rotationssymetrisk operator har den fordel, at kun et filter benyttes. Hvis dette er seperabelt vil en implementation af detektoren nok være hurtigere end en baseret på retningsspecifikke filtre. Endvidere kan det argumenteres at ingen kantorienteringer har speciel preference ved en rotationssymetrisk operator. I praksis er dette argument dog af mindre betydning. Mere væsenligt er, at rotationssymetriske operatorer ofte har en dårligere lokalisering af kanter der krummer meget (f.eks ved hjørner) i forhold til retningsspecifikke operatorer. Vi skal senere se nærmere på såvel rotationssymetriske som retningsspecifikke kantoperatorer.

## 2.6 Udtynding af kanter

Som tidligere nævnt er 1 pixels tynde kanter ofte at foretrække. Hvis den aktuelle anvendelse kræver dette, men kantdetektoren kun kan producere "fede" kanter, er ideen at udtynde eller bort-erodere alt andet end *skelettet* af kanterne. Denne operation kan oplagt udføres ved anvendelse af matematisk morfologi. Her skal vi imidlertid betragte en anden metode. Et krav til kantdetektoren (eller til billedet) er, at to initielt detekterede kanter ikke er "smeltet sammen" til en ekstra "fed" kant. Metoden, der omtales videre i kapitel 3, er opstået ud fra et ønske om at repræsentere *segmenter* i et billede ved *skelettet* af segmenterne. Metoden betegnes derfor ofte som *skelettransformationen* (eng. Medial Axis Transform). Aktuelt betragtes en initielt detekteret fed kant som et segment. I det følgende beskrives en (blandt mange) algoritmer til udtynding eller skelettransformation. Metoden er oprindeligt beskrevet i *Zhang and Suen: A Fast Parallel Algorithm for Thinning Digital Patterns, Comm. ACM vol 27 no3, 1984*.

Antag at kantpunkter har værdi 1, og at ikke-kantpunkter (baggrunden) har værdi 0. Metoden består i et antal gennemløb af billedet hvor hvert gennemløb har 2 faser. I hver fase *markeres* et antal kantpunkter som "ikke skeletpunkter". Efter hver fase slettes angivelsen af at disse punkter er kantpunkter. Algoritmen terminerer når ingen punkter er blever markeret i nogen af de to faser. For hvert billedpunkt i hvert fase foregår afgørelsen om markering på basis af en undersøgelse af antal og konfiguration af de tilbageblevne kantpunkter blandt de 8 nabopunkter.

Lad $(x, y)$ være det givne punkt i betragtning. Lad $N$ være antal nabopunkter til $(x, y)$, der stadig er angivet som kantpunkter. Lad $S$ være antal overgange fra ikke-kantpunkt til kantpunkt, når vi foretager en cirkulær vandretur omkring punktet, dvs. turen gennem: $(x + 1, y)$, $(x + 1, y + 1)$, $(x, y + 1)$, $(x - 1, y + 1)$, $(x - 1, y)$, $(x - 1, y - 1)$, $(x, y - 1)$, $(x + 1, y - 1)$, tilbage til $(x + 1, y)$.

For begge faser gælder, at et krav til en mulig markering af $(x, y)$ som et ikke-skeletpunkt er at:

$$2 \leq N \leq 6 \tag{2.45}$$

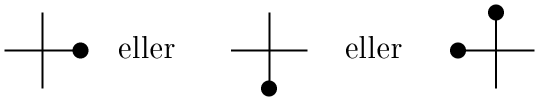
$$S = 1 \tag{2.46}$$

Begrundelsen for disse krav er, at hvis $N = 0$, da er punktet isoleret, og bør ikke slettes; Hvis $N = 1$ er punktet enden af en kant og bør ikke slettes heller. Hvis $N = 7$ eller $N = 8$ kan vi endnu ikke afgøre om punktet er et skeletpunkt (andre punkter må borteroderes først). Hvis $S \neq 1$ er punktet ikke lokaliseret på ydersiden af en fed kant, og skal derfor ikke slettes. Bemærk at for en 1-pixels tynd kant gælder at $S = 2$.

Kravet til at et punkt kan markeres under Fase 1 er yderligere begrænset af betingelserne:

$$E(x,y-1)\cdot E(x+1,y)\cdot E(x,y+1) = 0 \tag{2.47}$$
$$E(x+1,y)\cdot E(x,y+1)\cdot E(x-1,y) = 0 \tag{2.48}$$

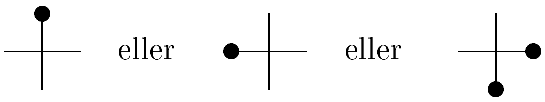
svarende til konfigurationerne:




hvor $E(x,y)$ som tidligere nævnt er lig 1 hvis $(x,y)$ er et kantpunkt og ellers 0. Hvis $E(x+1,y)=0$ eller $E(x,y+1)=0$ er punktet et “østligt” eller “sydligt” punkt på randen af en fed kant og kan slettes. Hvis både $E(x,y-1)$ og $E(x-1,y)$ er lig 0, er $(x,y)$ et “nordvestligt” hjørne af en fed kant og kan ligeledes slettes. Dette kræver en del overvejelser at indse. Læseren opfordres til at optegne situationerne og overbevise sig om rimeligheden i kravene. Kravet til at et punkt kan markeres under 2. fase er tilsvarende begrænset af:

$$E(x,y-1)\cdot E(x+1,y)\cdot E(x-1,y) = 0 \tag{2.49}$$
$$E(x,y-1)\cdot E(x,y+1)\cdot E(x-1,y) = 0 \tag{2.50}$$

svarende til konfigurationerne:




Som ovenfor nævnt er betingelserne svære at indse “matematisk”, men relativt lette at forstå intuitivt ved eksemplificering. Effektiviteten af algoritmen (og kvaliteten af de producerede tynde kanter) er ganske gode under førnævnte forudsætning af at kanterne er tilstrækkeligt separeret. Antallet af gennemløb er ofte ret stort (10 - 20). De fleste ikke-skelet kantpunkter fjernes i de første gennemløb. En ulempe ved algoritmen er at “enderne” af kanterne bliver noget afkortede. Afkortning vil omtrentlig svare til halvdelen af tykkelsen af den initielt detekterede kant.

## 2.7 Laplace-Gauss-filtret

Laplace-Gauss filtret er formodentlig et af de mest kendt og benyttede filtre i anvendelser hvor 1 pixel tynde kanter af god kvalitet kræves. Filtret er givet ved laplaceoperatoren på en gaussfunktion. Kantpunkterne defineres som nulpunktsgennemskæringerne i foldningen af filtret med intensitetsbilledet.

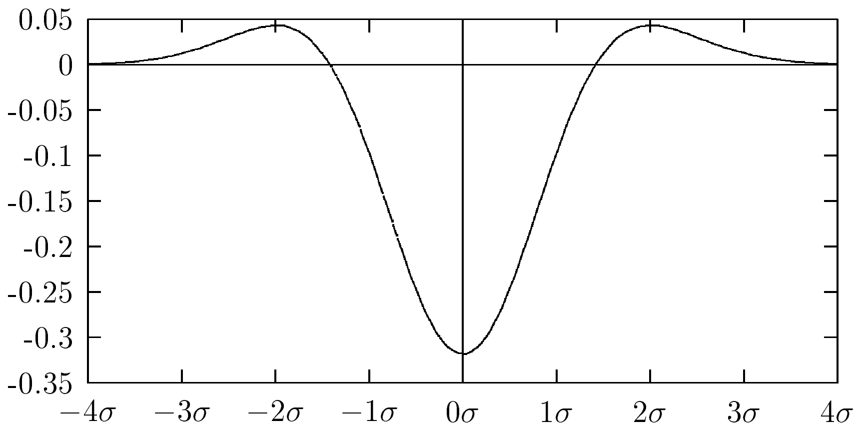
$$LG(x,y) = \nabla^2(G_\sigma * E(x,y)) = (\nabla^2 G_\sigma) * E(x,y) \tag{2.51}$$
$$= (\frac{\partial^2}{\partial x^2}G_\sigma + \frac{\partial^2}{\partial y^2}G_\sigma) * E(x,y) \tag{2.52}$$
$$= \frac{1}{2\pi\sigma^2}\left[\frac{x^2+y^2-2\sigma^2}{\sigma^4}\right] e^{-\frac{x^2+y^2}{2\sigma^2}} * E(x,y) \tag{2.53}$$

Filtret ses at være rotationssymmetrisk, og bestemt af parametren $\sigma$, der er spredningen af gaussfiltret, men ofte kaldes *skalaparametren.* Den fouriertransformerede af filtret vil også være rotationssymmetrisk. Skrives $s^2 = u^2 + v^2$ får vi den fouriertransformerede:

$$F\left\{\nabla^2 G_\sigma\right\}(u,v) = H(s) = -s^2 e^{-\frac{\sigma^2 s^2}{2}} \tag{2.54}$$

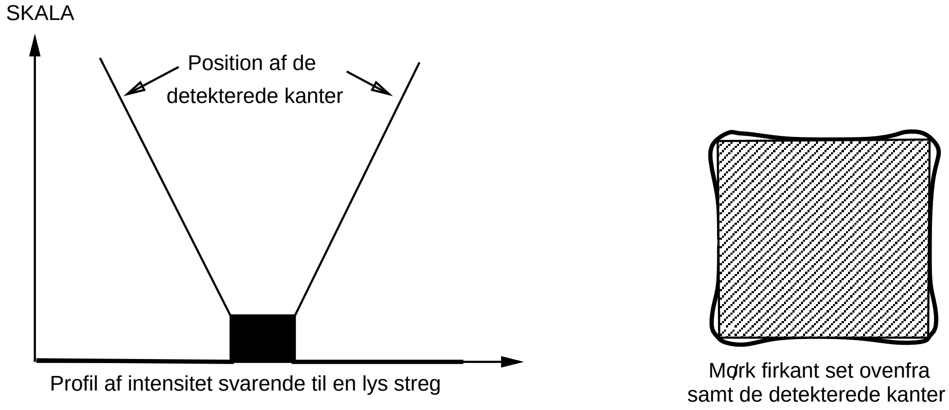
Det ses at filtret er et båndpasfilter med ekstrema for $s = \pm 2/\sigma^2$. Jo større skalaparametren $\sigma$ vælges, jo lavere frekvenser baseres kandetektionen på. Mindre variationer i intensitetsfunktionen vil altså ikke give anledning til nogen kanter. Vælges skalaparametren meget lille baseres detektionen i stedet på de høje frekvenser i frekvensspektret. Små og meget lokale intensitetsændringer vil i dette tilfælde resultere i nulpunktsgennemskæringer i resultatet af foldningen.

Et tværsnit af Laplace-Gaussfiltret har en form, der kan minde om en meksikansk hat (se Figur 2.7), hvorfor filtret ofte går under dette navn. Den centrale (negative) område




Figur 2.7: Et tværsnit af det to-dimensionale Laplacian-Gaussian filter

af filtret har diameter $w = 2\sqrt{2}\sigma$. Filtret implementeres ofte ved en kvadratisk støtte hvis sidelængde $l = kw$ hvor k er et lille tal mellem 3 og 5. Størrelsen af filtrets støtte er således afhængig af $\sigma$. Filtret giver som regel en god lokalisering af kanterne i et billede hvis disse er lineære og adskilder to intensitetsflader af konstant intensitet (trappetrinsmodellen).




Figur 2.8: Til venstre ses at afstanden mellen de detekterede kanter af en mørk/lys streg på en lys/mørk baggrund er lineært afhængig af $\sigma$. Til højre ses at Laplace-Gauss operatoren afrunder højner.

Ved krumme kanter (f.eks. hjørner) kan afvigelsen mellem den detekterede og den korrekte lokalisering dog blive stor (op mod $\frac{1}{2}\sigma$). Ukorrekt lokalisering forekommer også ved tynde aflange områder hvis intensitet afviger fra omgivelsernes intensitet. Hvis $w$ er større end områdets bredde n, vil afvigelsen være givet ved $(w-n)/2$. Disse tilfælde på fejllokalisering er illustreret i Figur 2.8. Videre diskussion kan f.eks. findes i *V. Berzins: Accuracy of Laplacian Edge Detectors, CVGIP vol 27, 1984.*

Bemærk at Laplace-Gaussfiltret ikke er seperabelt. Filtret kan imidlertid skrives som en sum af to seperable filtre:

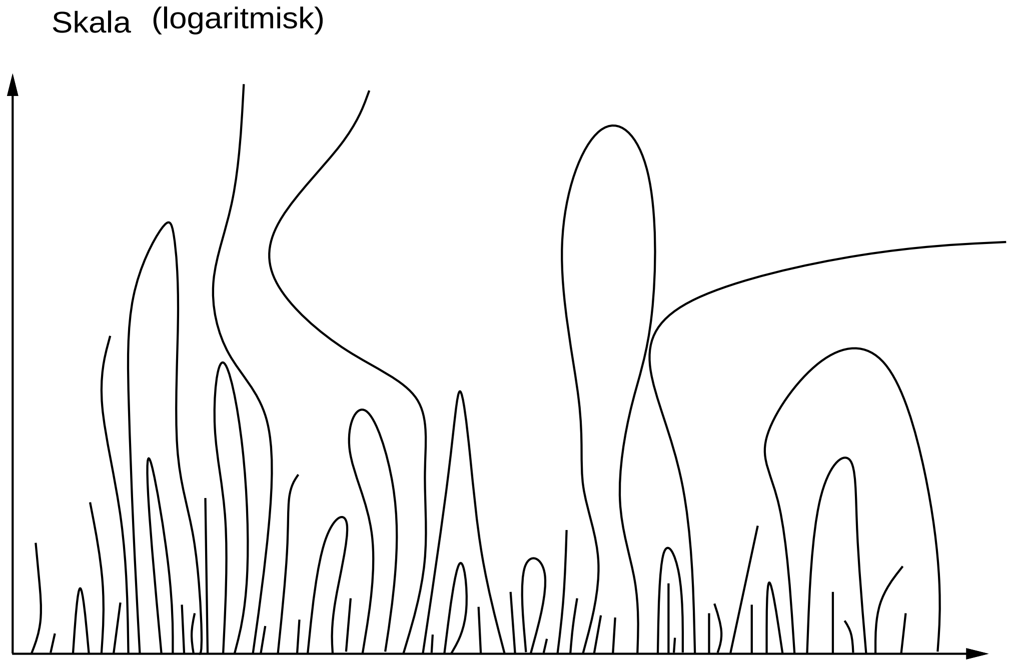
$$h_1(t) = \sqrt{K}\left[1-\frac{t^2}{\sigma^2}\right]e^{-\frac{t^2}{2\sigma^2}} \tag{2.55}$$

$$h_2(t) = \sqrt{K}\,e^{-\frac{t^2}{2\sigma^2}} \tag{2.56}$$

$$\nabla^2 G_\sigma(x,y) = h_1(x)h_2(y)+h_2(x)h_1(y) \tag{2.57}$$

hvor konstanten K for det uskalerede filter er lig $\frac{1}{2\pi\sigma^4}$. Idet filtret ofte kun benyttes som forbehandling til en detektion af nulpunktsgennemskæringer kan konstanten K imidlertid ofte vælges vilkårligt.

Laplace-Gauss-filtret er enestående i den forstand, at det (ligesom gaussfiltret) opfylder nogle nydelige matematiske egenskaber. *Witkin: Scale Space Filtering: Proj IJCAI 1983* introducerede ideen at betragte filtrets respons som en kontinuert funktion af $\sigma$. Lad $f(x)$ være et signal og lad $g(x,\sigma) = G_\sigma * f(x)$ være *skalarumsudvidelsen* af $f$. Betragt for hvert fast $\sigma$ kanterne (defineret som nulpunktsgennemskæringerne i den andenafledede) i $g$. For lille $\sigma$ vil kanterne ligger forholdsvist tæt . Når $\sigma$ øges vil kantpunkterne tilnærmer sig hinanden parvis og forsvinde. Hvis man betragter et enkelt kantpunkts position som funktion af $\sigma$ vil positionen beskrive et glat kurvestykke, der ikke krydser nogen anden kurve. Et sådant plot af kanterne for et signal kaldes *fingeraftrykket* for signalet. Det er




Figur 2.9: Fingeraftryk for et signal sådan som det ofte detekteres. Bemærk at visse områder (positivt hhv. negativt fortegn) i fingeraftrykket ikke er afgrænset af en kurve. Dette illustrerer de numeriske problemer ved detektion af nulpunktsskæringer i signaler med numerisk meget lille værdi.

vist i *Yuille and Poggio: Scaling Theorems for Zero-crossings; MIT AI memo 722, 1983* og *Yuille og Poggio: Fingerprints Therorems for Zero-crossings; MIT AI memo 730, 1983* at:

1. Laplace-Gaussfiltret er den eneste lineære differentialoperator med pæne skaleringsegenskaber.

2. Der findes intet filter, baseret på en ikke-lineær retningsspecifik differentialoperator, der har pæne skaleringsegenskaber.

3. Nulpunkterne i $\nabla^2 G_\sigma * E$ (betragtet som funktion af position og skala) bestemmer E fuldstændigt bortset fra en konstant og en harmonisk funktion.

hvor der med *pæne skaleringsegenskaber* menes, at der ikke introduceres "nye" kantpunkter for voksende skala. Ud over det intuitivt smukke ved at kanterne gradvist forsvinder for voksende $\sigma$, betyder denne egenskab at kanterne (detekteret ved Laplace-Gauss-filtret) er attraktive byggesten i en skalamæssig hierakisk beskrivelse af et signal/billede.

Da hver kurve i fingeraftrykket adskiller områder i skalarummet med positiv hhv. negativ fortegn ses at fingeraftrykskurverne må være lukkede (evt. afgænset nedaf mod x-aksen). Som illustreret i Figur 2.9 bliver det filtrerede signals numeriske værdi ofte så numerisk lille at det kan være vanskeligt at detektere nulpunktsgennemskæringen.

Det er betydelig vanskeligere at beskrive hvorledes mængden af kantpunkter (i skalarummet) opfører sig for et billede. I det simpleste tilfælde vil mængden af kantpunkter (i skalarummet) udgøre et hierakisk sæt af "omvendte skåle". Det er relativt let at forestille sig situationer hvor en lukket kant (i billedplanet) på ét skalaniveau splitter op i to lukkede kanter på et lidt højere niveau. Vi vil ikke her gå dybere i dette emne.

At intensitetsfunktionen kan rekonstrueres (beviset er ikke konstruktivt) ud fra information om skalarumskanterne indikerer at "kodningen" ved disse kantpunkter er stærk. Uformelt beskriver kanterne de steder i signalet/billedet hvor der "sker noget". Det er faktisk også muligt at rekonstruere et billede alene ud fra kantinformationen (Se: Curtis, Oppenheim and Lim *Signal Reconstruction from Fourier Transform Sign Information*; IEEE trans. Acoustics, Speech and Signal Processing, ASSP-33 no 3, 1985).

Laplace-Gauss kantoperatoren er diskuteret overordentlig meget i literaturen. Operatoren er let at implementere og relativ hurtig i forhold til kvaliteten af de producerede kanter. Som vi skal se i det følgende findes der imidlertid andre detektorer, der på trods af Laplace-Gaussfiltrets enestående matamatiske egenskaber, ofte er bedre mht. detektion af kanter i billeder.

## 2.8 Canny's kantoperator

Canny's kantoperator er enestående i den forstand, at den er særdeles nydeligt udledt og at den danner grundlag for Deriche's kantoperator, som nok er den bedste kantoperator på markedet i 1988. Canny's udledning er matematisk, og leder til et udtryk som Canny ikke kan håndtere. Canny beslutter derfor at udnytte en approximation til hans iøvrigt optimale filter. Canny's approximation er bedre end de fleste andre filtre til kantdetektion, men ringere end filteret udledt af Deriche. I dette afsnit vil vi gengive Canny's udledning og kort kommentere Canny's kantoperator. I næste afsnit beskrives Deriche's videreudvikling af operatoren.

Canny ønsker at bestemme det lineære positionsinvariante en-dimensionale filter $f(x)$, der maksimerer:

1. Signal/Støjforholdet.

2. Lokaliseringen af en ideel trappetrins-kant.

3. Entydigheden af kantmarkeringen

I det følgende skal vi opstille udtryk for signal/støjforholdet og for lokaliseringsevnen. Dette leder til en funktional, hvis maksimum definerer filtret $f(x)$. Ved analyse af denne funktion skal vi se at kravet om entydighed er nødvendigt. Dette krav kan uformelt formuleres som ønsket om at hver kant kun markeres een gang. Væsenligt i Canny's analyse er:

- Modellen af de kanter hvor operatoren ønskes at fungere optimalt (trappetrinsmodellen).

- Modellen af den støj signalet antages behæftet med (normalfordelt hvid støj).
- Beslutningen om at kanter detekteres som lokale ekstrema i impulssvaret for foldning af $f(x)$ med det signal hvori kanterne ønskes detekteret.

### 2.8.1 Mål for robusthed og lokalisering

Lad $I(x)$ være en ideel trappetrinsfunktion med højde $A$ hvortil er adderet uafhængig hvid støj $\eta(x)$.

$$I(x) = Au(x) + \eta(x) \tag{2.58}$$

$$u(x) = \begin{cases} 0 & \text{hvis } x < 0 \\ 1 & \text{hvis } x \geq 0 \end{cases} \tag{2.59}$$

Svaret ved anvendelse af det endnu ukendte filter $f(x)$ er givet ved:

$$O(x_0) = f * I(x_0) = \int_{-\infty}^{\infty} I(x) f(x_0 - x) dx \tag{2.60}$$

Ved udnyttelse af definitionerne fås at:

$$O(x_0 = 0) = A \int_{-\infty}^{0} f(x) dx + \int_{-\infty}^{\infty} \eta(x) f(-x) dx \tag{2.61}$$

I udledningen af signal/støjforholdet S.N.R. antages at dette kan skrives som kvotienten af bidraget til $O(x)$ oprindende fra den ideelle kant og bidraget fra støjen. Tælleren i dette forhold er netop givet ved det første led i højresiden af ovenstående formel. I det følgende udledes en approximation til støjbidraget. Betragt den statistisk forventede værdi af kvadratet på bidraget fra støjen:

$$E\left\{\left[\int_{-\infty}^{\infty} f(x)\eta(-x) dx\right]^2\right\} = n_0^2 \int_{-\infty}^{\infty} f^2(x) dx \tag{2.62}$$

hvor $n_0^2$ angiver variansen af støjen. Udledningen kan begrundes ved grænseværdibetragtninger over en endelig diskret approximation af integralet med en summation. Centralt i denne udledning er følgende skitse:

$$E\left\{\left[\sum_{x=-N}^{N} f(x)\eta(-x)\right]^2\right\} = \tag{2.63}$$

$$E\left\{[f(-N)\eta(N) + \cdots\cdots + f(N)\eta(-N)]^2\right\} = \tag{2.64}$$

$$E\left\{\sum_{x=-N}^{N} f^2(x)\eta^2(-x) + \sum_{i=-N, i\neq j}^{N} \sum_{j=-N}^{N} f(i)f(j)\eta(-i)\eta(-j)\right\} = \tag{2.65}$$

$$\sum_{x=-N}^{N} f^2(x) E\left\{\eta^2(-x)\right\} + \sum_{i=-N, i\neq j}^{N} \sum_{j=-N}^{N} f(i) f(j) E\left\{\eta(-i)\eta(-j)\right\} = \tag{2.66}$$

$$E\left\{\eta^2(-x)\right\} \sum_{x=-N}^{N} f^2(x) = \tag{2.67}$$

$$n_0^2 \sum_{x=-N}^{N} f^2(x) \tag{2.68}$$

hvor det næstsidste lighedstegn kommer idet $\eta(i)$ og $\eta(j)$ er uafhængige og den forventede værdi af produktet derfor er lig nul, dvs. at $n_0^2 = E\left\{\eta^2\right\} - E^2\left\{\eta\right\} = E\left\{\eta^2\right\}$. Bidraget til filtersvaret O(x) oprindende fra støjen er derfor givet ved kvadratroden af højresiden af udtrykket i formel (2.62), og vi kan skrive målet for filtrets signal/støjforhold som:

$$S.N.R. = \frac{A\int_{-\infty}^{0} f(x)dx}{n_0\sqrt{\int_{-\infty}^{\infty} f^2(x)dx}} = \frac{A}{n_0}\Sigma \tag{2.69}$$

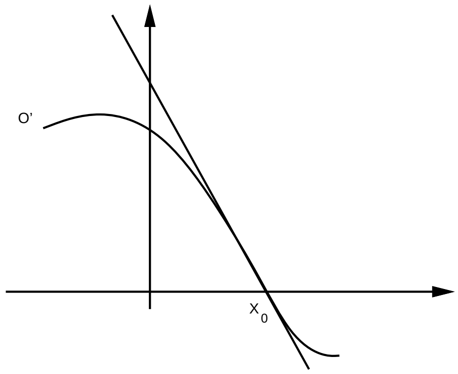
Trappetrinskanten er lokaliseret i $x = 0$, og lokaliseringsevnen for filtret kan derfor måles som den absolutte værdi af af maksimumspunktet i O(x). Antag at dette maksimum antages i $x_0$, dvs. at den førsteafledede $O'$ har en nulpunktsgennemskæring i $x_0$. Ønsket om at bestemme $f(x)$ så lokaliseringsevnen maksimeres, kan da formuleres som et ønske om at maksimere $x_0^{-1}$. Af regnereglerne for differentiation af en foldning fås:

$$\begin{aligned}
O'(x_0) &= (\frac{\partial}{\partial x} f * I)(x_0) && (2.70)\\
&= \int_{-\infty}^{\infty} f'(x) I(x_0 - x)dx && (2.71)\\
&= A\int_{-\infty}^{x_0} f'(x)dx + \int_{-\infty}^{\infty} f'(x)\eta(x_0 - x)dx && (2.72)\\
&= Af(x_0) + \int_{-\infty}^{\infty} f'(x)\eta(x_0 - x)dx && (2.73)\\
&= 0 && (2.74)
\end{aligned}$$

På samme måde som i udledningen af S.N.R. forholdet kan vi opsplitte i et bidrag oprindende fra signalet og et bidrag oprindende fra støjen. Signalbidraget kan approximeres ved:

$$O'_s(x_0) = Af(x_0) \approx Ax_0 f'(0) \tag{2.75}$$

Her antages altså at $f'(x)$ er omtrentlig lineær i en lille omegn omfattende $x_0$ og 0. I vurderingen af støjbidraget betragtes (som før) forventede værdier. Vi får:




Figur 2.10: Antagelse om linearitet i omegn om lokalt maximum af $O(x)$.

$$E\left\{O'^2_n(x)\right\} = n_0^2 \int_{-\infty}^{\infty} f'^2(x)dx \tag{2.76}$$

Vi kan nu skrive

$$\begin{aligned} O'(x_0) &= O'_s(x_0) + O'_n(x_0) = 0 \ \Rightarrow && (2.77) \\ O'_s(x_0) &= -O'_n(x_0) \ \Rightarrow && (2.78) \\ E\left\{O'^2_s(x_0)\right\} &= E\left\{O'^2_n(x_0)\right\} && (2.79) \end{aligned}$$

Højresiden af formel (2.79) er givet ved (2.76). Ved indsættelse af definitionerne, samt approksimationen i (2.75), kan venstresiden videreudvikles til:

$$E\left\{O'^2_s(x_0)\right\} \approx E\left\{A^2 x_0^2 f'^2(0)\right\} = A^2 f'^2(0) E\left\{x_0^2\right\} \tag{2.80}$$

Ved sammenstilling af de udledte formler, fås et udtryk for den forventede værdi af kvadratafvigelsen i lokalisering af kanten:

$$E\left\{x_0^2\right\} = \frac{n_0^2 \int_{-\infty}^{\infty} f'^2(x)dx}{A^2 f'^2(0)} \tag{2.81}$$

Vi ønsker at minimere ovenstående størrelse (eller ækvivalent kvadratroden af størrelsen). Dette svarer til ønsket om at maksimere $\left[E\left\{x_0^2\right\}\right]^{-1/2}$. Lokaliseringsevnen for filtret $f(x)$ kan derfor måles ved:

$$Lokalisering = \frac{A}{n_0} \frac{|f'(0)|}{\sqrt{\int_{-\infty}^{\infty} f'^2(x)dx}} = \frac{A}{n_0}\Lambda \tag{2.82}$$

### 2.8.2 Et performance mål for en kantoperator

Vi ønsker at optimere signal-støj-forholdet og lokaliseringsevnen samtidig. Dette kan gøres på mange måder. Mest oplagt er (som Canny vælger) at betragte produktet af de to udtryk. Det søgte filter $f(x)$ fremkommer da, som den funktion, der maksimerer produktfunktionalen:

$$\Theta(f) = \left(\frac{A}{n_0}\Sigma\right) \cdot \left(\frac{A}{n_0}\Lambda\right) = \frac{A^2}{n_0^2}\Sigma\Lambda \tag{2.83}$$

I ovenstående formel er størrelsen $A^2/n_0^2$ uafhængig af filtret $f(x)$. I det følgende vil vi derfor kun betragte produktet $\Sigma\Lambda$. Denne størrelse kaldes *performance målet* for en kantoperator. Før vi går videre med udledningen af det filter, der optimerer $\Theta$, skal vi kort bemærke at en skalering $f(x/w)$ ingen indflydelse har på performance målet $\Theta$.

$$\Theta(x/w) = \Sigma(x/w)\Lambda(x/w) = \left[\sqrt{w}\Sigma(f)\right] \cdot \left[\frac{1}{\sqrt{w}}\Lambda(f)\right] = \Sigma(f)\Lambda(f) = \Theta(f) \tag{2.84}$$

Heraf ses at hvis $f(x)$ optimerer $\Theta$, da vil også $f(x/w)$ optimere $\Theta$. Der eksisterer, med andre ord, en hel klasse af optimale filtre. Disse er karakteriseret ved deres robusthed (S.N.R.-forholdet) og ved lokaliseringsevnen. Vælges $w$ stor øges robustheden hvorimod lokaliseringsevnen forringes. Vi kan således opnå vilkårlig god robusthed og lokaliseringsevne, blot ikke samtidig. Denne observation er af fundamental betydning, ikke kun for Canny's kantoperator, men langt mere generelt.

### 2.8.3 Ekskurs i variationsregning

Variationsregning er et værktøj til bestemmelse af den/de funktioner, der ekstremiserer en funktional. Dette er (i en vis forstand) en pendent til det værktøj vi benytter når vi ønsker at finde stationære punkter for en funktion. I sidstnævnte tilfælde differentierer vi funktionen mht. hver variabel og sætter hvert udtryk lig nul. Denne metode kan vi ikke bruge når vi ønsker at optimere over et rum af funktioner. I det følgende introduceres meget overfladisk til variationsregning. Alle beviser udelades. For en grundigere gennemgang henvises til et utal af lærebøger i matematik. En udmærket gennemgang findes på dansk i: *Erik Hansen, Variationsregning, Polyteknisk Forlag, Lyngby 1976.*

I variationsregning betragter vi en funktional $\Phi(f) : (\Omega \rightarrow \Re) \rightarrow \Re$, hvor $f$ altså antages at afbilde $\Omega$ på de reelle tal, og hvor $\Phi$ er et *godhedsmål* på sådanne funktioner, dvs. et reelt tal. Ofte vælges $\Phi$ således at $\Phi(f_1) < \Phi(f_2)$ hvis $f_1$ er bedre end $f_2$. Lad os i det følgende antage at $f$ er defineret på en delmængde $\Omega$ af $\Re$. I dette tilfælde antages at godhedsmålet har formen:

$$\Phi(f) \;=\; \int\int_{\Omega} F(x, f, f_x, f_{xx}, \ldots)dx \tag{2.85}$$

Det antages altså at godhedsmålet er bestemt ved funktionen $F$ af stedkoordinaten $x$, af funktionen $f$ selv, samt af de afledede af $f$. I de fleste tilfælde (fra billedanalyse) indgår hverken $x$ eller afledede af orden større end 2. Til gengæld er det ofte at funktionen $f$ er defineret på en delmængde af $\Re^2$. I dette tilfælde fås:

$$\Phi(f) \;=\; \int\int_{\Omega\subseteq\Re^2} F(f, f_x, f_y, f_{xx}, f_{xy}, f_{yy})dxdy \tag{2.86}$$

I dette tilfælde har $F$ således seks parametre. I det følgende betragtes mængden af *tilladelige* funktioner , der er $2N$ gange differentiable, hvor $N$ er den højeste orden afledede af $f$ som optræder i $F$. Desuden skal der gælde, at alle afledede af $f$ af orden op til $N$ enten har værdi lig nul på randen af definitionsområdet $\Omega$ eller opfylder et sæt på forhånd givne *randbetingelser*. Det kan vises at en nødvendig betingelse for at en tilladelig funktion $f$ ekstremiserer $\Phi$ er at funktionen tilfredsstiller *Euler-Lagrange ligningen* for funktionalen. For en funktion $f$ af én variabel er Euler-Lagrange ligningen givet ved:

$$F_f \;+\; \sum_{n=1}^{N} (-1)^n \frac{\partial^n}{\partial x^n} F_{f^{(n)}} \;=\; 0 \tag{2.87}$$

Euler-Lagrange ligningen for funktioner af flere variable er analoge. F.eks. er Euler-Lagrange ligningen svarende til ovenstående eksempel lig:

$$F_f \;-\; \frac{\partial}{\partial x} F_{f_x} \;-\; \frac{\partial}{\partial y} F_{f_y} \;+\; \frac{\partial^2}{\partial x^2} F_{f_{xx}} \;+\; \frac{\partial^2}{\partial x \partial y} F_{f_{xy}} \;+\; \frac{\partial^2}{\partial y^2} F_{f_{yy}} \;=\; 0 \tag{2.88}$$

I ovenstående ligning skal faktoren $F_{f_{xx}}$ eksempelvis læses "den afledede af funktionen $F$ med hensyn til den fjerde parameter $f_{xx}$". Størrelserne $f_x$, $f_{xx}$, etc. betragtes altså her som symboler, og ikke som funktionsudtryk.

**Eksempel**

Nedenstående anføres en række integrander $F$ til godhedsmåled $\Phi(f)$, samt de tilhørende Euler-Lagrange ligninger.

$$\begin{array}{lcl} F = f_x^2 + f_y^2 & : & f_{xx} + f_{yy} = 0 \\ F = f - f_x^2 + f_{xx}^2 & : & 1 + 2f_{xx} + 2f_{xxxx} = 0 \\ F = \sqrt{1 + f_x^2} & : & f_{xx} = 0 \end{array}$$

hvor der i sidste eksempel er udnyttet at kvadratrodsfunktionen er monotont voksende. Ekstremum for $\sqrt{g(x)}$ og for $g(x)$ antages derfor i de samme punkter.

Det er vigtigt ved anvendelse af variationsregning, at Euler-Lagrange ligningen kun angiver et nødvendigt krav til en funktion $f$ om at ekstremiserer $\Phi(f)$. Selv om Euler-Lagrange ligningen er opfyldt er vi ikke garanteret at $f$ er et globalt ekstremum. Det ikke let at opstille kriterier for en tilstrækkelig betingelse for ekstremum. Dog gælder, at hvis funktionen $F$ er konveks, da angiver Euler-Lagrange ligningen en tilstrækkelig betingelse.

**Eksempel**

Betragt variationsregningsproblemet:

$$\Phi(f) \;=\; \int_a^b \sqrt{1 \;+\; f_x^2(x)}dx \;\;,\quad f(a) \;=\; f_a \;\;,\quad f(b) \;=\; f_b$$

Her søger vi altså den korteste glatte kurve gennem punkterne $(a, f_a)$ og $(b, f_b)$. Klassen af tilladelige funktioner er mængden af 2 gange differentiable relle funktioner på delmængden [a:b]. Integranden $F$ er i dette tilfælde konveks. Euler-Lagrange ligningen er givet ved den anden ordens differentialligning $f_{xx} = 0$. Ved integration får vi derfor løsningen: $f(x) = Ax + B$, hvor $A$ og $B$ er reelle integrationskonstanter. Vi har altså vist at den korteste vej mellem to punkter er en ret linie. Konstanterne $A$ og $B$ skal vælges således at *randbetingelserne* $f(a) = f_a$ og $f(b) = f_b$. Herved bliver løsningen specificeret fuldstændigt.

Hvis randbetingelserne i et variationsregningsproblem udelades kan det vises at et krav for ekstremitet (ud over Euler-Lagrange ligningen) er opfyldelsen af de såkaldte *naturlige randbetingelser*. I tilfældet $F = F(x, f, f')$ vil de naturlige randbetingelser diktere at $F_{f'}$ er lig nul på randen (i intervalgrænserne) for definitionsområdet. Hvis vi, som eksempel, i ovenstående eksempel undlod randbetingelserne $f(a) = f_a$ og $f(b) = f_b$ ville de naturlige randbetingelser angive at $f_x(a) = f_x(b) = 0$, svarende til kravet $A = 0$. Løsningen ville derfor ikke blive én funktion, men klassen af funktioner med konstant værdi.

I mange variationsregningsopgaver er der, ud over godhedsmålet $F_0$, givet en antal bibetingelser :

$$
\begin{aligned}
\Phi_1(f) &= \int_\Omega F_1(\ldots) dxdy = c_1 \\
\Phi_2(f) &= \int_\Omega F_2(\ldots) dxdy = c_2 \\
&\vdots \\
\Phi_n(f) &= \int_\Omega F_n(\ldots) dxdy = c_n
\end{aligned}
$$

hvor alle integraler deler samme grænser. Det kan vises at løsningen til ovenstående variationsregningsproblem er den samme som løsningen til problemet:

$$
\Phi(f) = \int_\Omega \left[F_0 + \lambda_1 F_1 + \lambda_2 F_2 + \ldots + \lambda_n F_n\right] dxdy \tag{2.89}
$$

hvor $F_0$ er det oprindelige godhedsmål, og hvor sættet $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots \lambda_n) \neq (0, 0, \ldots, 0)$ vælges således at bibetingelserne er opfyldt. $\vec{\lambda}$ kaldes for et sæt af Lagranske multiplikatorer. Løsningen til Euler-Lagrange ligningen vil derfor være parametriseret ved $\vec{\lambda}$. Denne vektor bestemmes (efter at den generelle løsning er fundet) således at bibetingelserne bliver opfyldt.

### Eksempel

Lad klassen af tilladelige funktioner være lig mængden af to gange differentiable funktioner på intervallet [-w:w]; lad godhedsmålet $\Phi_0(f)$, og bibetingelsen $\Phi_1(f)$ være defineret ved nedenstående ligninger:

$$
\Phi_0(f) = \int_{-w}^{w} f_x^2(x) dx
$$

$$\Phi_1(f) \;=\; \int_{-w}^{w} f^2(x)dx \;=\; 1$$

Lad der endvidere gælde følgende randbetingelser: $f(0) = 0$, $f(w) = A$ og $f(w) = -A$. Først opstilles den sammensatte funktional:

$$\Phi(f) \;=\; \int_{\Re}[f_x^2(x) + \lambda f^2(x)]dx$$

Denne har Euler-Lagrange ligningen: $\lambda f - f_{xx} = 0$, der har løsningen: $f(x) = c_1 e^{mx} + c_2 e^{-mx}$, hvor $m = \sqrt{\lambda}$. Randbetingelsen $f(0) = 0$ giver nu at $c_1 + c_2 = 0$. Herved kan løsningen skrives: $f(x) = 2c_1 \sinh(mx)$. Konstanten $c_1$ kan nu bestemmes således at bibetingelsen bliver opfyldt. Endelig kan konstanten $m$ bestemmes således at de to sidste randbetingelser bliver opfyldt.

**Eksempel**

Lad os antage at vi har målt en todimensional funktion i et antal punkter $M$, og at vi ved at funktionen er glat i den forstand at funktionen kan modelleres ved en membran. Hvis dette er tilfældet kan godheden af funktionen $f$ måles ved den energi, som der skal til at bøje funktionen (membranen) således at den passerer tæt ved de givne målepunkter. Det kan vises, at integralet af kvadratet på gradientstørrelsen er et mål for den omtalte energi. Vi kan derfor opskrive kravet til $f$ som et godhedsmål $\Phi_0$ samt en bibetingelse $\Phi_1$:

$$\begin{aligned}
\Phi_0(f) &= \int_{\Re^2}[f_x^2(x,y) \;+\; f_y^2(x,y)]dxdy \\
\Phi_1(f) &= \int_{\Re^2}\delta_M(x,y)[f(x,y) \;-\; c(x,y)]^2 dxdy \;=\; 0
\end{aligned}$$

hvor

$$\delta_M(x,y) \;=\; \begin{cases} 1 & \text{hvis } (x,y) \in M \\ 0 & \text{ellers} \end{cases}$$

og hvor

$$c(x,y) \;=\; \begin{cases} \text{den målte værdi} & \text{hvis } (x,y) \in M \\ \text{udefineret} & \text{ellers} \end{cases}$$

Den sammensatte funktional bliver derfor:

$$\Phi(f) \;=\; \int_{\Re^2}[f_x^2(x,y) \;+\; f_y^2(x,y) \;+\; \lambda\delta_M(x,y)[f(x,y) - c(x,y)]^2]dxdy$$

Euler-Lagrange ligningen bliver:

$$\lambda\delta_M(x,y)[f(x,y) \;-\; c(x,y)] \;+\; f_{xx}(x,y) \;+\; f_{yy}(x,y) = 0$$

Vi har således transformeret det globale optimeringsproblem til et problem om at løse en differentialligning, der skal holde for alle $(x,y)$. Hvis vi kræver bibetingelsen overholdt eksakt (svarende til at vælge $\lambda$ lig $\infty$) er løsningen (ud over overensstemmelserne mellem

$f$ og $c$ på $M$) givet ved et sæt af koblede differentialligninger $f_{xx}(x,y) + f_{yy}(x,y) = 0$. Det er ikke muligt direkte at opskrive en analytisk løsning.

Antag i det følgende at vi ikke kræver at bibetingelsen overholdes eksakt, men kun at funktionen $f$ passerer tæt forbi de målte værdier $c$. Denne situation svarer til at vælge $\lambda < \infty$. Antag endvidere at såvel $f$ som $c$ er diskrete funktioner. En ofte benyttet diskret approksimation til laplaceoperatoren er:

$$\nabla^2 f(x,y) \;=\; f_{xx}(x,y) \;+\; f_{yy}(x,y) \;=\; 4\bar{f}(x,y) - 4f(x,y)$$

hvor

$$\bar{f}(x,y) \;=\; \frac{1}{4}[f(x-1,y) \;+\; f(x+1,y) \;+\; f(x,y-1) \;+\; f(x,y+1)]$$

Indsættes den diskrete approksimation, og løses med hensyn til funktionsværdien $f(x,y)$ fås følgende relation:

$$f(x,y) \;=\; \frac{4\bar{f}(x,y) \;+\; \lambda\delta_M(x,y)c(x,y)}{4 \;+\; \lambda\delta_M(x,y)}$$

der skal være opfyldt i alle $(x,y)$. I punkter hvori der ikke er nogen måling ses at værdien af $f$ er bestemt ved gennemsnittet af de omkringliggende værdier. I punkter $(x,y) \in M$ er $f$ bestemt ved et vægtet gennemsnit af denne værdi og af den målte værdi. Jo større $\lambda$ vælges jo mere vægt lægges der på overensstemmelse med de målte værdier i forhold til glathedskriteriet. Ovenstående udledning kan sammenlignes med det tilsvarende udtryk for regularisering af et en-dimensionalt signal (se afsnit 1.7.3).

### 2.8.4 En optimal kantoperator (første forsøg)

Den optimale filterform $f(x)$ findes ved sædvanlig variationsregning. Dog skal enkelte ting på plads først. Det er nødvendigt at integralerne i funktionalen har de samme grænser. Det er let at se at $f(x)$ må være en ulige funktion. Antag for et kort øjeblik at kanten er ulige, dvs. $u(x) = sign(x)$ (sammenlign med (2.59)). Enhver funktion $f$ kan skrives som summen af en lige $f_{lige}$ komponent og en ulige komponent $f_{ulige}$. Det ses at $f_{ulige}$ vil bidrage til både tæller og nævner af SNR-forholdet (2.69). $f_{lige}$ vil derimod kun bidrage til nævneres. Hvis $f(x)$ har en lige komponent vil SNR-forholdet derfor være dårligere end for en ren ulige funktion. For andre trappetrinskanter $u(x) = Asign(x) + B$ gælder tilsvarende observation fordi foldningen er lineær og fordi filtersvaret for en konstant funktion skal være lig nul. Under antagelsen at $f(x)$ er ulige kan vi nu skrive:

$$\int_{-\infty}^{\infty} f^2(x)dx \;=\; 2\int_{-\infty}^{0} f^2(x)dx \tag{2.90}$$

$$\int_{-\infty}^{\infty} f'^2(x)dx \;=\; 2\int_{-\infty}^{0} f'^2(x)dx \tag{2.91}$$

I praksis er Canny ikke interesserede i at støtten for $f(x)$ er uendelig stor. Antag derfor at $f(x)$ har endelig støtte [-W,W], og antag at $f(-W) = f(0) = 0$. Lad endvidere $|f'(0)| = c_1$.

Klassen af tilladelige funktioner er således lig mængden af endelige to gange differentiable reelle funktioner på støtteområdet. Opgaven er nu, at finde $f(x)$ i denne klasse sådan at funktionalen:

$$\frac{\int_{-W}^{0} f(x)dx}{\sqrt{\int_{-W}^{0} f^2(x)dx}\sqrt{\int_{-W}^{0} f'^2(x)dx}} \tag{2.92}$$

maksimeres under bibetingelsen $f'(0) = c_1$ og under randbetingelserne:

$$f(-W) = 0 \qquad f(0) = 0 \tag{2.93}$$

Der er ikke her behov for introduktion af nogen naturlige randbetingelser. Løsningsmetoden er den sædvanlige, idet Lagrange-multiplikatorer benyttes. Optimering af kvotienten $\frac{\Theta_2}{\Theta_1\Theta_3}$ under de givne randbetingelser giver samme løsning(er) som ved optimering af $\Theta_1$ under bibetingelserne $\Theta_2 = c_2$ og $\Theta_3 = c_3$ samt randbetingelserne. Vi skal altså finde ekstrema af et udtryk af formen $\int_a^b F(x, f, f')$ under bibetingelser af formen $\int_a^b G_i(x, f, f') = c_i$. Da alle integraler deler samme grænser, kan vi i stedet finde ekstremum for den sammensatte funktional:

$$\Psi(x, f, f') = F(x, f, f') + \lambda_1 G_1(x, f, f') + \lambda_2 G_2(x, f, f') \tag{2.94}$$

hvor

$$F(x, f, f') = f^2 \qquad G_1(x, f, f') = f \qquad G_2(x, f, f') = f'^2 \tag{2.95}$$

idet kvadratrodsfunktionen er nonotont voksende. Ekstrema for $\sqrt{h(x)}$ svarer derfor til ekstrema for $h(x)$ selv. Funktionalen har altså formen:

$$\Psi = f^2 + \lambda_1 f + \lambda_2 f'^2 \tag{2.96}$$

Et nødvendigt (men ikke garanteret tilstrækkeligt) krav til $f$ om at være ekstremum for funktionalen er, at $f$ er løsning til *Euler-Lagrange ligningen* for funktionalen. I dette simple tilfælde er Euler-Lagrange ligningen givet ved:

$$\Psi_f - \frac{\partial}{\partial x}\Psi_{f'} = 0 \tag{2.97}$$

Ved indsættelse og udregning fås Euler-Lagrange ligningen til:

$$2f + \lambda_1 - 2\lambda_2 f'' = 0 \tag{2.98}$$

Det søgte filter $f(x)$ er altså løsning til ovenstående anden ordens differentialligning. Løsninger er:

$$f(x) = \beta + a_1 e^{\alpha x} + a_2 e^{-\alpha x} \tag{2.99}$$

hvor $\alpha = \lambda_2^{-1/2}$, $\beta = -\lambda_1/2$, og $a_1$ og $a_2$ er bestemt af randbetingelserne $f(0) = f(-W) = 0$. Idet $f'(0) = \alpha(a_1 - a_2) = c_1$ ses at filtret er bestemt pånær en parameter (f.eks. $\alpha$). Filtret kan, for $x \in [-W; 0]$ omskrives til:

$$f(x) = \beta \left[ 1 - \frac{\cosh(\alpha(x + \frac{W}{2}))}{\cosh(\alpha \frac{W}{2})} \right] \tag{2.100}$$

Det fundne filter er afhængig af parametren $\alpha$. En måde at bestemme $\alpha$ på er, at undersøge performance-målet $\Theta$'s værdi for forskellige valg af $\alpha$. Parametren bør vælges sådan at målet maksimeres. Det kan (med nogen anstrengelse) vises at:

$$\lim_{\alpha \to \infty} \Sigma = \sqrt{2} \tag{2.101}$$

$$\lim_{\alpha \to \infty} \Lambda = \infty \tag{2.102}$$

$$\lim_{\alpha \to 0} \Sigma = \sqrt{\frac{5}{6}} \tag{2.103}$$

$$\lim_{\alpha \to 0} \Lambda = \sqrt{\frac{3}{4}} \tag{2.104}$$

hvoraf vi ser at:

$$\lim_{\alpha \to \infty} \Theta(f) = \infty \qquad \lim_{\alpha \to 0} \Theta(f) = \sqrt{\frac{5}{8}} \tag{2.105}$$

Vi får således optimal performance ved at vælge $\alpha$ meget stor. Det kan (også med noget besvær) vises at for $\alpha$ gående mod uendelig, vil $f(x)$ konvergere mod en differens af to box-filtre:

$$f_\infty(x) = \begin{cases} \beta & \text{for } -W < x < 0 \\ -\beta & \text{for } 0 < x < W \\ 0 & \text{ellers} \end{cases} \tag{2.106}$$

Resultatet af ovenstående (forholdsvist omstændelige) analyse er umiddelbart skuffende. Det optimale filter kan beskrives ved (samtidig) anvendelse af den centrale differens-approksimation til den førsteafledede (se formel (2.10)) og et boxfilter med støtte 2W. Den fouriertransformerede af $f_\infty(x)$ er rent imaginær og givet ved:

$$F\{f_\infty\} = (-2i\beta W)\frac{\cos 2\pi u W}{2\pi u W} \tag{2.107}$$

Som diskuteret i kapitel 1 er problemet ved anvendelse af boxfiltret, at dette har meget stor båndbredde. Hvis signalet I(x) indeholder støj, betyder dette, at filtret ofte vil give mange lokale maksima. Disse vil ofte ligge tæt, og det vil ikke være muligt at vælge et maksima (oprindende fra kanten) frem for andre (oprindende fra støjkomponenten). For at undgå dette problem (multiple lokale maksima) er det nødvendigt at tilføje endnu et led til funktionalen.

### 2.8.5 En optimal kantoperator (andet forsøg)

Vi ønsker at begrænse antallet at lokale maksima, sådan at der inden for bredden af impulssvaret for en ideel trappetrinskant kun forekommer et maksima. Lad $\xi$ være den gennemsnitlige afstand mellem lokale maksima i O(x), eller ækvivalent den gennemsnitlige afstand mellem nulpunktsgennemskæringer i O'(x). Vi ønsker nu at bestemme det filter $f(x)$, der ekstremiserer funktionalen:

$$\Theta(f) = \Sigma(f)\Lambda(f)\xi(f) \tag{2.108}$$

Canny benytter nu er resultat af Rice *Rice: Mathematical Analysis of Random Noise; Bell Sys. Tech. J. 24, 1945*, der siger at den gennemsnitlige afstand mellem nulpunktsgennemskæringer i svaret for en funktion $g(x)$ på gaussisk fordelt støj, er givet ved:

$$x_{ave} = \pi \left( \frac{-R(0)}{R''(0)} \right)^{\frac{1}{2}} \tag{2.109}$$

hvor $R(\tau)$ er autokorrelationsfunktionen af g, dvs.:

$$\begin{aligned} R(\tau) &= g \otimes g(\tau) \\ &= \int_{-\infty}^{\infty} g^*(x) g(x+\tau) dx \\ &= g \star g(-\tau). \end{aligned} \tag{2.110}$$

idet $g$ er reel. Heraf fås videre at:

$$\begin{aligned} R''(\tau) &= [g' \star g(-\tau)]' \\ &= -g' \star g'(-\tau) \end{aligned} \tag{2.111}$$

Derfor har vi at:

$$R(0) = \int_{-\infty}^{\infty} g^2(x) dx \qquad og \qquad R''(0) = -\int_{-\infty}^{\infty} g'^2(x) dx \tag{2.112}$$

I vores tilfælde er vi interesserede i den gennemsnitlige afstand mellem nulpunktsgennemskæringer i $f'(x)$. Idet $g = f'$, får vi resultatet:

$$x_{ave} = \pi \left( \frac{\int_{-\infty}^{\infty} f'^2(x) dx}{\int_{-\infty}^{\infty} f''^2(x) dx} \right)^{\frac{1}{2}} \tag{2.113}$$

Afstanden mellem nabomaksima i O(x) oprindende fra støjen vil blive det dobbelte af $x_{ave}$ fordi der mellem hvert lokalt maksima nødvændigvis må være et lokalt minima. Som mål for den søgte afstand vælger Canny forholdet mellem $2x_{ave}$ og filterstørrelsen W. Idet konstanten $\frac{2\pi}{W}$ ignoreres får vi derfor den ny (sammensatte) funktional:

$$\Theta(f) = \frac{\int_{-\infty}^{0} f(x) dx}{\sqrt{\int_{-\infty}^{\infty} f^2(x) dx}} \cdot \frac{|f'(0)|}{\sqrt{\int_{-\infty}^{\infty} f'^2(x) dx}} \cdot \frac{\sqrt{\int_{-\infty}^{\infty} f'^2(x) dx}}{\sqrt{\int_{-\infty}^{\infty} f''^2(x) dx}} \tag{2.114}$$

Ovenstående funktional er betydelig sværere at håndtere end den forrige. En årsag er, at der forekommer udtryk med kvadratet af den første afledede i både tæller og nævner i funktionalen. Vi skal ikke har gå nærmere i den (mere avancerede) variationsregning, der er nødvendig for at finde $f(x)$, men kun referere resultaterne. Svarende til ovenstående funktional fremkommer formen:

$$\Psi(x, f, f', f'') = f^2 + \lambda_1 f'^2 + \lambda_2 f''^2 + \lambda_3 f \tag{2.115}$$

Eulerligningen for denne højere ordens form er:

$$\Psi'_f - \frac{\partial}{\partial x}\Psi_{f'} + \frac{\partial^2}{\partial x^2}\Psi_{f''} = 0 \tag{2.116}$$

Hvoraf vi kan udlede at et nødvendigt krav til filtret f(x) er, at det er løsning til differentialligningen:

$$\frac{\lambda_3}{2} + f - \lambda_1 f'' + \lambda_2 f'''' = 0 \tag{2.117}$$

Idet Canny definerer $\alpha$ og $\omega$ ved:

$$\alpha^2 - \omega^2 = \frac{\lambda_1}{2\lambda_2} \qquad og \qquad 4\alpha^2\omega^2 = -\frac{\lambda_1^2 - 4\lambda_2}{4\lambda_2^2} \tag{2.118}$$

udleder han (udledningen er absolut ikke triviel, se *Canny: A Computational Approach to Edge Detection, IEEE PAMI vol. 8 no. 6 1986*), at det søgte filter har formlen:

$$f(x) = e^{\alpha x}\left[a_1 \sin \omega x + a_2 \cos \omega x\right] + e^{-\alpha x}\left[a_3 \sin \omega x + a_4 \cos \omega x\right] + c \tag{2.119}$$

hvor de fire konstanter $a_1$ til $a_4$ er bestemt af bibetingelsen $f'(0) = s$, og grænsebetingelserne:

$$f(0) = 0 \quad f(-W) = 0 \quad f'(-W) = 0 \tag{2.120}$$

Disse betingelser er nødvendige for at sikre at evalueringen af integralerne over hele den reelle akse svarer til evalueringen over det endelige interval (se Canny: *A Computational Approach to Edge Detection*, IEEE PAMI vol. 8 no. 6 1986, p 865 venstre spalte midt).

Funktionen $f(x)$ er således parametriseret af $\alpha$, $\omega$, c og s. Disse parametre bør bestemmes således at $f(x)$ ekstremisere funktionalen (betragtet som funktion af disse fire parametre). Canny opgiver (fornuftigvis ?) at løse dette problem analytisk, og bestemmer i stedet det optimale filter ved stokastisk optimering.

### 2.8.6 Opsummering

Et væsenligt resultat ved Canny's analyse er, at der, som vi så tidligere, findes en hel klasse af optimale filtre. Disse er bestemt ved at have konstant produkt af lokaliseringsevne $\Lambda$ og robusthed $\Sigma$. Som før vil det være muligt at opnå vilkårlig god lokaliseringsevne eller vilkårlig god robusthed, kun kan dette ikke opnås samtidig. *Nøgletallene* $\Sigma\Lambda$ hhv. $\xi$ for et kantfilter vil ligge i intervallerne $[0:\infty]$ hhv. $[0:1]$. For Canny's optimale filter (bestemt ved stokastisk optimering) er nøgletallene:

$$\text{Canny:} \qquad \Sigma \cdot \Lambda = 1.12 \qquad og \qquad \xi = 0.58 \tag{2.121}$$

Canny observerer at det optimale filter har en form, der er meget lig formen af den førsteafledede af et gaussfilter.

$$G'(x) = -\frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \tag{2.122}$$

Canny vælger (på dette løse grundlag) at erstatte sit optimalt udledte filter med den førsteafledede af gaussfiltret, men bemærker dog at gaussfiltretes performance mål $\Lambda\Sigma$ og dets evne $\xi$ til seperering af nabomaksima er dårligere end for det optimale filter. Nøgletallene for gauss-approximationen er:

$$\text{G':} \qquad \Sigma \cdot \Lambda = 0.92 \qquad og \qquad \xi = 0.51 \tag{2.123}$$

Canny argumenterer (løst) at approximationens ca. 20% dårligere performance næppe vil kunne bemærkes ved anvendelse på ikke-syntetiske billeder. Som vi skal se i næste afsnit, er det muligt at konstruere et filter med endnu bedre nøgletal end Canny's optimale filter. Det skal også bemærkes, at det er let at beregne *nøgletallene* for en række andre mulige kantdetektorer, f.eks.:

$$\frac{\partial}{\partial x}\left(\frac{1}{1+\alpha x^2}\right) \qquad \Sigma \cdot \Lambda = 1.47 \qquad og \qquad \xi = 0.41 \tag{2.124}$$

$$\frac{\partial}{\partial x}\left(\frac{1}{\cosh(\alpha x)}\right) \qquad \Sigma \cdot \Lambda = 1.27 \qquad og \qquad \xi = 0.48 \tag{2.125}$$

Den sædevanlige implementation af Canny's kantoperator er at benytte filtret til beregning af de "førsteafledede" ved horisontal hhv. vertikal foldning af billedet med filtret. Ud fra disse størrelser beregnes "gradientstørrelse" -og orientering, og kantpunkter detekteres som lokale ekstrema af "gradientstørrelsen" i "gradientretningen" (Se afsnit 2.5 om non-maximum-suppression). Canny benytter derefter en grænseværdiafskæring med hysterese. Når der refereres til Cannys kantdetektor er det ofte uklart hvorvidt der er tale om det optimale filter eller "kun" om filtret (2.122). Canny's kantoperator hører til de absolut bedste på markedet, og benyttes meget hvor der kræves god kvalitet af de detekterede kanter. Operatoren benyttes derfor ofte som reference-operator når nyudviklede metoder skal sammenlignes med eksisterende.

De væsentlige formål med dette afsnit har være at præsentere en systematisk og elegant metode til udledning af "optimale" filtre samt en metode til vurdering af "performance" for disse, at præsentere en god og meget benyttet kantoperator, samt at lægge op til Deriche's videreudvikling af Canny's operator. Yderligere beskrivelse af Canny's detektor kan findes i *J. F. Canny: A Computational Approach to Edge Detection, IEEE PAMI vol. 8, no. 6, 1986*. Cannys tilgang kan anvendes til design af optimale operatorer for andre modeller end trappetrinskanten. I ovenstående reference beskriver Canny også udledningen i det generelle en-dimensionelle tilfælde. Andre forfattere har brugt Cannys systematik til udledning af optimale to-dimensionale filtre (f.eks. hjørnedetektorer).

## 2.9 Deriche's kantoperator

Deriche tager i sin videreudvikling af Canny's kantoperator udgangspunkt i filterformlen:

$$f(x) = e^{\alpha x}\left[a_1 \sin \omega x + a_2 \cos \omega x\right] + e^{-\alpha x}\left[a_3 \sin \omega x + a_4 \cos \omega x\right] + c \tag{2.126}$$

hvor de fire konstanter $a_1$ til $a_4$ er bestemt af bibetingelsen $f'(0) = s$, og grænsebetingelserne. Canny opgav at finde et analytisk udtryk for det optimale filter. Hovedideen i Deriche's arbejde er at erstatte visse af disse randbetingelserne svarende til antagelsen om et filter med uendelig støtte. Vi skal senere se hvorledes Deriche implementerer et sådant filter ved *rekursiv filtrering*. Primært antages randbetingelserne:

$$f(0) = 0 \qquad f(-\infty) = 0 \qquad f'(-\infty) = 0 \qquad |f'(0)| = s \tag{2.127}$$

Ved indsættelse kan det vises at disse betingelser (i hele det reelle område) leder til filtret:

$$f(x) = -ce^{-\alpha\cdot|x|} \cdot \sin \omega x \tag{2.128}$$

hvor c, $\alpha$ og $\omega$ er reelle positive konstanter. Af filtrets fouriertransformerede:

$$F(u) = \frac{i4c\alpha\omega u}{(\alpha^2+\omega^2) + u^4 + 2u^2(\alpha^2-\omega^2)} \tag{2.129}$$

ses at filtret er af bånd-pass type. Ved udregning af de integraler, som indgår i Canny's *performance index* fås:

$$\Lambda = (2\alpha)^{\frac{1}{2}} \tag{2.130}$$

$$\Sigma = \left[\frac{2\alpha}{\alpha^2+\omega^2}\right]^{\frac{1}{2}} \tag{2.131}$$

$$\xi = \left[\frac{\alpha^2+\omega^2}{5\alpha^2+\omega^2}\right]^{\frac{1}{2}} \tag{2.132}$$

$$\Lambda \cdot \Sigma = \frac{2\alpha}{\sqrt{\alpha^2+\omega^2}} \tag{2.133}$$

Lad nu $\alpha = m{\cdot}\omega$, og lad os undersøge performance for Deriche's kantoperator ved forskellige valg af $m$. Betragt først valget $m = 1$. I dette tilfælde fås nøgletallene:

$$m = 1: \qquad \Lambda\Sigma = \sqrt{2} \approx 1.414 \qquad \xi = 0.58 \tag{2.134}$$

Det ses at for $m = 1$ er separationen af lokale maksima den samme som for Canny's filter. Derimod har Deriche's filter en performance, som er ca. 25 % bedre end Canny's filter. Vælges $m \ll 1$ fås et filter, hvor separationen af lokale maksima er størst (nemlig 1.0). her er performancemålet $\Sigma\Lambda$ imidlertid dårligt (nemlig $2m$, der er lille da $m$ er lille). Den maksimale performance fås ved valg af $m$ meget større end 1:

$$m \gg 1: \qquad \Lambda\Sigma \approx 2.0 \qquad \xi = \frac{\sqrt{5}}{5} \approx 0.45 \tag{2.135}$$

Det ses at performance for dette valg af $m$ er omkring 79 % bedre end Canny's operator, hvorimod filtrets evne til separation af lokale maksima er lidt ringere end for Canny's operator. Når vi i det følgende refererer til Deriche's kantoperator forstås det filter hvor konstanten $m$ er valgt meget stor. Før vi går videre med at beskrive hvorledes filtret kan implementeres, skal vi dog først se en simplifikation af filtret. Bemærk at for stort $m$ vil $\sin \omega x \approx \omega x$. Dette leder til filtret:

$$g(x) = -cxe^{-\alpha|x|} \tag{2.136}$$

Det simplificerede filter har (naturligvis) samme performance som det forrige udledte filter (de to filtre er ikke til at skelne). Deriche's filter er derfor (essentielt) parametriseret ved kun en parameter $\alpha$. Denne parameter har omtrent samme funktion som $1/\sigma$ i (den førsteafledede af) gaussfiltret. Øges/mindskes $\alpha$ øges/mindskes lokaliseringsevnen og robustheden mindskes/øges. I det følgende skal vi ikke benytte det simplificerede filter, fordi filtret $f(x)$ i modsætning til $g(x)$ kan implementeres effektivt ved rekursiv filtrering.

### 2.9.1 Implementation af Deriche's kantoperator

Deriche's kantoperator kan naturligvis implementeres på sædvanlig måde ved to foldninger af billedet med de endimensionale filter-funktioner (svarende til x hhv. y -retningerne). Grundet støjproblemer ønsker vi imidlertid først at udvide filtrene til 2 dimensioner, ved at multiplicere disse med en orthogonal udglatningsfunktion. Deriche vælger her (forholdsvist oplagt) at at benytte integralet af funktion som projektions- eller udglatningsfunktion.

$$\int_0^t f(x)dx = \frac{\alpha \sin(\omega|t|) + \omega\cos(\omega|t|)}{\alpha^2 + \omega^2} e^{-\alpha|t|} \tag{2.137}$$

De to-dimrensionale filterfunktioner defineres således ved:

$$f^x = \frac{\left[-ce^{-\alpha|x|}\sin(\omega x)\right]\left[k(\alpha\sin(\omega|y|) + \omega\cos(\omega|y|))e^{-\alpha|y|}\right]}{\alpha^2 + \omega^2} \tag{2.138}$$

$$f^y = \frac{\left[-ce^{-\alpha|y|}\sin(\omega y)\right]\left[k(\alpha\sin(\omega|x|) + \omega\cos(\omega|x|))e^{-\alpha|x|}\right]}{\alpha^2 + \omega^2} \tag{2.139}$$

hvor konstanterne c og k kan bestemmes som normaliseringskonstanter (dvs. at c og k bestemmes således at integralet over den positive del af filteret bliver 1). Herved fås:

$$c = \frac{1 - 2e^{-\alpha}\cos(\omega) + e^{-2\alpha}}{e^{-\alpha}\cos(\omega)} \tag{2.140}$$

$$k = \frac{\left[1 - 2e^{-\alpha}\cos(\omega) + e^{-2\alpha}\right](\alpha^2 + \omega^2)}{2\alpha e^{-\alpha}\sin(\omega) + \omega - \omega e^{-2\alpha}} \tag{2.141}$$

Givet $\alpha$ og $\omega$ er vi nu i stand til at beregne de to to-dimensionale filtre. Vi ønsker ikke at implementere disse på sædvanlig måde ved foldning af filtrene (eller de filtre der fremkommer ved unyttelse af seperabiliteten) med billedet. årsagen er, at vores antagelse i udledningen af filtrene var, at disse havde uendelig støtte. I stedet benyttes IIR-filtrerings teknikker, der kort beskrives i det følgende.

### IIR-filtrering

Ved IIR-filtrering (Infinite Impulse Response filtering) benyttes *kausale filtre* . Vi opfatter hver række i et billede som en sekvens af målinger opsamlet en efter en, således at vi til et givet tidspunkt (en given pixel) kun har kendskab til de forudgående målinger (pixels med mindre koordinat). Ved en løbende (on-line) filtrering vil alle pixels med koordinat mindre end den aktuelle kunne indgå i beregningen. Anderledes sagt er filterkoefficienterne $f^+(n)$ lig 0 for $n > 0$. Ønsker vi at "simulere" filtrering med et ikke-kausalt filter $f(x)$, opdeles dette i to kausale filtre $f^+(n)$ og $f^-(n)$ sådan at deres sum giver $f(x)$.

Fordelen ved opdeling af et filter i kausale delfiltre er irrelevant hvis filtret vides at have lille støtte (idet en traditionel implementering da vil være hurtig). Hvis filterstøtten er stor er det i visse tilfælde muligt at implementere hvert af de kausale delfiltre sådan at filtreringen vil være meget hurtig. Vi skal ikke her beskrive kravene til at implementationen kan foretages (dette hænger primært på konvergensegenskaberne for Z-transformationen af de kausale filtre). Givet at betingelserne er opfyldt vil filtreringen kunne foretages efter følgende skema :

$$y^+(t) = a_{11} * x(t-1) + a_{21} * x(t-2) \cdots a_{n1} * x(t-n) \tag{2.142}$$

$$+b_{11} * y(t-1) + b_{21} * y(t-2) \cdots a_{m1} * y(t-m) \tag{2.143}$$

$$y^-(t) = a_{12} * x(t+1) + a_{22} * x(t+2) \cdots a_{n2} * x(t+n) \tag{2.144}$$

$$+b_{12} * y(t+1) + b_{22} * y(t+2) \cdots a_{m2} * y(t+m) \tag{2.145}$$

$$y(t) = y^+(t) + y^-(t) \tag{2.146}$$

hvor a'erne og b'erne er konstanter, udledt fra formen af de to kausale filtre, $n$ og $m$ er to små heltallige konstanter. $x(t)$ er inddata-signalet og $y(t)$ er uddata-signalet (ved foldning med $f(x)$). Generelt gælder at jo mere kompliceret filterformen er, jo større skal n og m være. For de filtre vi taler om er det som regel tilstrækkeligt at lade $n = m = 2$. Foretages filtreringen som i ovenstående skema kaldes denne for *rekursiv* . Af ovenstående rekursive filtrering ses, at IIR-filtreringen foretages med en filterstøtte på $n$ ($m$) pixels. Til gengæld udnyttes information om samtligt tidligere "besøgte" pixels ved reference til filtersvaret $y$ for $t-1$, $t-2$ osv. Det er oplagt at for små $n$ og $m$, vil en filtrering efter ovenstående skema være overordentlig hurtig.

Problemet er bestemmelsen af filterkonstanterne, dvs. a'erne og b'erne. Vi skal ikke i disse noter gå nærmere ind på disse bestemmelser, lige som vi ikke beskriver de teoretiske krav til at en filtrering kan foretages rekursivt.

Til illustration dels af rekursiv IIR-filtrering, dels af Deriche's kantoperator, er denne beskrevet (i sin helhed) ved nedenstående programstump. I programmet angiver I inddata-intensitetsbilledet, S1 og S2 er temporære billeder, og Gx hhv. Gy er uddata-billeder angivende filtrets svar ved anvendelse horisontalt hhv. vertikalt. Der foretages fire rekursive filtreringer: en fra venstre mod højre, en fra højre mod venstre, en fra toppen af billedet mod bunden, og en i den modsatte retning. Bemærk at $\omega$ sættes til $\alpha/1000$.

I praksis fungerer IIR-implementationen af Deriche's kantoperator betydeligt bedre end både Laplace-Gauss-kantoperatoren og Canny's kantoperator. En forskel er (jvf. nøgletallene) at Deriche's operator er mindre følsom over for støj i billederne. En anden forskel er at IIR-implementationen er meget hurtig. Metoden og implementationen er beskrevet i detalie i *Deriche: Using Canny's Criteria to derive a Recursively Implemented Optimal Edge Detector, Int. Jour. on Computer Vision 1987.*

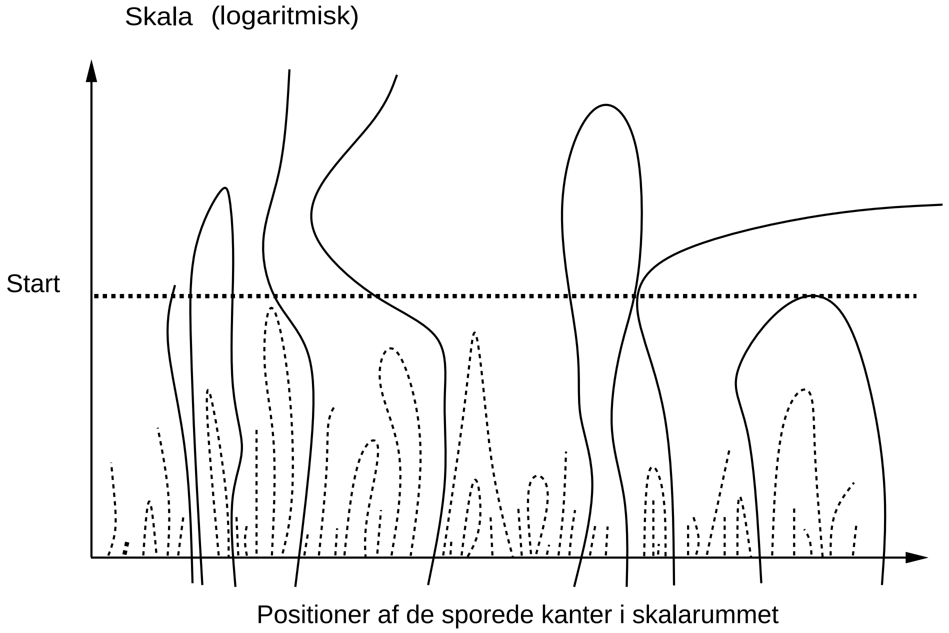
```
/* Declare images*/
grey  image  I;
float image  S1, S2;
float image  Gx, Gy;

/* Declare variables*/
register int x, y;
int          dimx, dimy;
float        alfa, omega;
float        a, a0, a1, a2, a3, b1, b2, c, c1, c2, k;
float        ea, e2a, co, so, ao;
float        t12, t13, t22, t23, t32, t33, t42, t43;
float        t1, t2, t3, t4;

alfa  = ''READ PARAMETER VALUE'';
omega = alfa/1000.0;

/*initialize*/
ea   = exp(-alfa);                  e2a  = exp(-2.0*alfa);
co   = cos(omega);                  so   = sin(omega);
ao   = alfa*alfa + omega*omega;     c    = 1.0 - 2.0*ea*co + e2a;
k    = c*ao;                        c   /= ea*so;
k   /= (2.0*alfa*ea*so + omega - omega*e2a);
c1   = k*alfa/ao;                   c2   = k*omega/ao;
b1   = -2.0*ea*co;                  b2   = e2a;
a    = -c*ea*so;                    a0   = c2;
a1   = (c1*so - c2*co)*ea;          a2   = a1 - c2*b1;
a3   = -c2*b2

/*do first filtering passage*/
t12 = t13 = t22 = t23 = t32 = t33 = t42 = t43 = 0.0;
for (x = 0; x < dimx; x++)
for (y = 2; y < dimy; y++)
{
```

```
      t1   = (float)I[x][y-1] - b1*t12 - b2*t13;
      t13  = t12;        t12  = t1;
      t2   = (float)I[y-1][x] - b1*t22 - b2*t23;
      t23  = t22;        t22  = t2;
      y1   = dimy - y;
      t3   = (float)I[x][y1] - b1*t32 - b2*t33;
      t33  = t32;        t32  = t3;
      t4   = (float)I[y1][x] - b1*t42 - b2*t43;
      t43  = t42;        t42  = t4;

      S1[x][y]    += a*t1;   S2[y][x]    += a*t2;
      S1[x][y1-1] -= a*t3;   S2[y1-1][x] -= a*t4;
   }

   /* Do second filtering passage*/
   t12 = t13 = t22 = t23 = t32 = t33 = t42 = t43 = 0.0;
   for (x = 0; x < dimx; x++)
   for (y = 2; y < dimy; y++)
   {
      t1  = a0*S1[y][x]   + a1*S1[y-1][x] - b1*t12 - b2*t13;
      t13 = t12;        t12 = t1;
      t2  = a0*S2[x][y]   + a1*S2[x][y-1] - b1*t22 - b2*t23;
      t23 = t22;        t22 = t2;
      y1  = dimy - y - 1;
      t3  = a2*S1[y1+1][x] + a3*S1[y1+2][x] - b1*t32 - b2*t33;
      t33 = t32;        t32 = t3;
      t4  = a2*S2[x][y1+1] + a3*S2[x][y1+2] - b1*t42 - b2*t43;
      t43 = t42;        t42 = t4;

      Gy[y][x]  += t1;   Gy[y1][x] += t3;
      Gx[x][y]  += t2;   Gx[x][y1] += t4;
   }
```

## 2.10 Kantfokusering i skalarummet

Kantdetektionerne ved såvel Laplace-Gauss-operatoren, Canny's operator samt Deriche's operator kræver angivelse af en *skalaparameter*. Spredningen $\sigma$ for Laplace-Gauss-operatoren og parameteren $\alpha$ for de andre to. Som diskuteret i afsnit 2.9, findes der en klasse af "lige gode filtre" i den forstand at de har samme *performance index*. Filtrene adskilder sig ved deres evne til korrekt lokalisering af kanterne samt deres robusthed over for støj. En optimal detektion af kanter i et forelagt billede kræver altså at vi (givet formålet med kantdetektionen) finder det optimale valg af skalaparameteren. De forudgående analyser giver ingen retningslinier for et sådant valg. I det følgende beskrives en metode, der i en vis forstand eliminerer kravet om angivelse af en skalaparameter.

Den grundlæggende ide er først at foretage kantdetektionen for stort valg af skalaparameteren. Herved findes de væsentlige kanter i billedet robust. Lokaliseringen af kanterne er dog ringe. Herefter spores de detekterede kanter (nu opfattet som kanter i skalarummet, dvs. rummet udspændt af billedet og skalaparameteren) mod mindre skalaværdi. Sporingen foretages iterativt ved langsom at mindske skalaværdien og detektere kanterne i en lille omegn om de groft fundne positioner (markeret for større skalaværdi). Der foretages altså kun een fuldstændig detektion af kanter i hele billedet. Dette gør at metoden ikke er væsentlig tungere beregningsmæssigt, end de tidligere beskrevne metoder (især hvis disse benytter grænseværdiafskæring med hysterese e.lign.).




Figur 2.11: Kantfokusering i skalarummet.

Fordelen ved metoden er at kanterne detekteres både robust og stabilt. Kanter forårsaget af støj vil kun optræde for lille valg af skalaparameteren og vil ikke blive detekteret. En yderligere fordel er, at metoden ikke kræver angivelse af en optimal skalaværdi. Dog kræves angivelse af en initiel skalaværdi. Denne er lettere at angive fordi de markante kanter i et billede ofte "overlever langt op" i skalarummet. Et optimalt valg af den initielle skalaværdi er derfor ikke essentiel for resultatets kvalitet.

Ulempen ved metoden er at det er uklart hvorvidt alle "væsentlige" kanter i et billede er tilstede på skalaniveauet angivet ved den initielle værdi, samt om mange "uvæsentlige" kanter ikke tilsammen kan manifestere en kant på dette niveau. Begge tilfælde er teoretisk mulige, omend betingelserne for disse tilfælde ikke er teoretisk fuldt afklaret. Praktisk anvendelse af metoden indikerer, at visse intuitivt oplagte kanter ikke detekteres, samt at visse detekterede kanter ikke intuitivt opfattes som sådanne. Selv om kantdetektionen således i mange situationer er bedre en detektion ved Canny's, Deriche's eller andre mere simple metoder, løser ovenstående metode ikke endegyldigt problemet omkring kantdetektion. For en nærmere beskrivelse af metoden refereres til : *F. Bergholm: Edge Focusing; IEEE trans. PAMI vol. 9, no. 6, 1987, p. 726.*

### 2.10.1 Kantdetektion ved Anisotrop diffusion

Jævnfør afsnit 1.7.2 i Kapitel 1, er det muligt at benytte anisotrop diffusion til at "konstruere" et skalarum, hvor positionen af kanterne er konstant over skala. Foretages en anisotrop diffusion ved itteration (simulering) af (den diskrete version af) diffusionsligningen vil det resulterende billede blive stykvist konstant. Detektionen af kanterne i dette billede er således meget simpel. Den eneste parameter i dektektionen er konstanten $K$ (se afsnit 1.7.2 for nærmere forklaring). Hvis denne ikke kan angives af brugeren er en idé at foretage en serie af diffusioner med forskellig værdi af $K$. For voksende værdie af $K$ vil kanterne gradvist forsvinde (som i det sædvanlige skalarum). Det tilsvarende fingeraftryk vil altså bestå af lodrette liniestykker.

## 2.11 Afsluttende kommentarer til kantdetektion

Vi har nu set en række eksempler på en- og to-dimensionale filtre til detektion af kantpunkter i billeder. Vi har også set at der eksisterer en hel klasse af filtre, der optimerer Canny's performancemål. Vi kan derfor vælge det "bedste" filter når blot vi har valgt hvilken skala, som vi ønsker at undersøge signalet/billedet på. Forudsætningerne for optimalitet var at:

1. At støjen er hvid (og normalfordelt).

2. At kanten er en isoleret trappetrinskant.

3. At kantoperatoren er lineær og positionsinvariant.

I mange tilfælde er den første antagelse rimelig. Mængden af støj i moderne CCD-kamera'er forholdsvis lille ($n_0^2 \approx 2$ til 4). For visse kamera er der endvidere et tydelig multiplikativt støjled (jo mere lys, jo mindre støj). Dette gør egenskaberne af de omtale filtre mht. filtrering af ren additiv støj noget problematisk. At kanten antages at have trappetrinsform er en grov antagelse, der næppe er opfyldt noget sted i praksis. Dette forhold problematiserer betragtningerne omkring optimalitet. At der ikke (inden for støtteområdet) befinder sig andre kanter er en virkelig grov antagelse. I områder med tætliggende kanter, eller kanter der mødes vil de fleste kantoperatorer enten fejllokalisere kanterne eller undlade at detektere disse. Dette er især tilfældet ved de såkaldte T- og Y-, og X- "junctions", samt ved hjørner. At kantoperatoren blev antaget lineær og positionsinvariant var blot for at kunne opskrive filtersvaret som en foldning, dvs. en beregningsteknisk forudsætning. Ofte er der grund til at lade støttestørrelsen afhænge af den lokale kanttæthed, dvs. at gøre filtret positionsvariant. Ligeledes kan det argumenteres at sumoperatoren i foldningen med fordel kunne erstattes af en ulineær operation (f.eks. medianen). Selv om at der er blevet arbejdet med detektion af kanter og andre lav-niveau features i billeder i over 25 år er der således endnu ingen endegyldig afklaring af hvorledes disse detekteres bedst muligt.

## 2.12 Kodning af kanter

I visse tilfælde vil det være hensigtsmæssigt at kode de fundne kanter i et billede. Ofte vil kantbilledet være af gråtone-type, selv om det reelt kun antager to værdier. Kodning af kanterne vil derfor betyde en meget stor datareduktion. Den mest banale kodning vil være at registrere koordinaterne for alle kantpunkter. Denne metode vil imidlertid ofte give en ringe kompressionsfaktor medmindre antallet af kantpunkter er lille i forhold til antal billedpunkter. I praksis vil man ofte bruge 2 gange 16 bit til angivelse af koordinaterne for et punkt. Hvis eksempelvis hvert 10'ende punkt er markeret som kantpunkt ses at vise at kompressionsfaktoren kun er 2.5. Kan koordinaterne kodes i 2 gange 9 bit øges den tilsvarende kompressionsfaktor til ca. 4.5.

I det følgende beskrives nogle mere avanceret metoder til kodning af kanter. En vigtig forskel mellem metoderne er om de koder kanterne eksakt eller ineksakt. Generelt ønsker man at kodningen er korrekt, men i mange anvendelser kan små fejl i de rekonstruerede kanter accepteres hvis kompressionsfaktoren samtidig kan øges væsenligt. I visse tilfælde, f.eks. hvor kantpositionerne er støjfyldte, kan en ineksakt kodning ligefrem være en fordel. I visse anvendelser benyttes kantkodningen direkte i forbindelse med genkendelse af objekter. Eksempelvis kan silhouetten af et objekt entydigt identificere dette. Ved kodning af silhouetkanten kan koden benyttes direkte i en sammenligning med et sæt af prototyper (se kapitel 5). Hvis objektet kan forekomme overalt i synsfeltet (billedet) ønskes at kodningen er invariant over for translation. Hvis objektet kan være roteret vilkårligt ønskes invarians over for rotationer. Hvis afstanden mellem objekt og kamera er ukendt, kan der være en skalaforskel mellem prototypen og observationen. Her ønskes at kodningen er skalainvariant. Endelig ønskes ofte at kodningen er invariant over for valg af startpunkt.

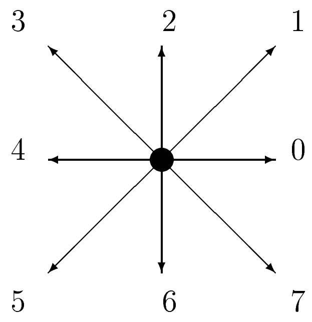
Ved de fleste former for kantkodning kræves at de detekterede kantpunkter grupperes sådan at hver gruppe udgør en kant. Ved kodning med polygoner kræves desuden at hver kant opdeles i kantsegmenter. Grupperingen af kantpunkter kan (mest oplagt) foretages løbende ved anvendelse af en "kantfølger"-algoritme. En (blandt mange) sådanne algoritmer beskrives i det følgende afsnit.

### 2.12.1 Kantfølgning

Kantfølgning er en basal operation i billedbehandling og i grafik. Vi skal ikke her forsøge at give en oversigt over metoder til kantfølgning, men kun introducere en metode. Denne er hentet fra *Pavlidis: Graphics and Image Processing; Springer 1982.* Alternative metoder hvor kantfølgning betragtes som hhv. søgning i en graf, hhv. som et problem fra dynamisk programmering kan eksempelvis findes i *Ballard and Brown: Computer Vision; Prentice Hall 1982.*

Man taler om 4-sammenhængelde eller 8-sammenhængende retningskoder, svarende til hvorvidt de fire nabopixel (langs koordinatakserne) eller samtlige 8 nabopixels defineres som nabokantpunkter. Valget mellem de to kodningsmetoder har eksempelvis betydning ved reproduktion af "skrå" linier ud fra kodningen. Algoritmen fra Pavlidis benytter 8-sammenhængende kanter. De otte retninger refereres ved tallene 0 til 7. Retningen 0 angiver billedelementet med første-koordinat der er een større end første-koordinaten i det aktuelle punkt. Retningerne svarende til de øvrige nabopunkter fås ved fortløbende nummerering modsat urets omløbsretning. Operationer på retninger foregår naturligvis modulo 8.

Kantpunkter antages at have værdi 1 og baggrunds punkter værdien 0. Isolerede 1-pixels kanter registreres ikke. Dette er årsagen til at inderste sløjfe gennemløbes højst tre gange. Det antages at funktionen *nabo* giver koordinaterne for nabopunktet svarende til retningsparametren. Funktionen *edge* angiver (ved værdierne 0, 1 eller 2) om en given posi-




Figur 2.12: 8-sammenhængende retningskoder

tion ikke angiver et kantpunkt, om det angiver et ubesøgt kantpunkt eller om kantpunktet allerede er del af en anden kontur. Parameteren 'or' angiver om konturen skal gennemløbes "højre-om" ('or' = 2) eller "venstre-om" ('or' = 6).

```
procedure tracer(startpoint, or) begin
   point       pos, new;
   boolean     first, found;
   integer     itt;

   pos   := startpoint;
   first := true;

   while ((pos != startpoint) or (first = true)) begin
       found := false;
       itt   := 3;
       while (( found = false) and (itt > 0)) begin
           if (edge(new := nabo(pos,(or-1) mod 8)) = 1) then begin
               pos   := new;
               or    := (or - 2) mod 8;
               found := true;
           end
           else if (edge(new := nabo(pos,or)) = 1) then begin
                pos   := new;
                found := true;
           end
           else if (edge(new := nabo(pos,(or + 1) mod 8)) = 1) then begin
                pos   := new;
                found := true;
           end
           else or := (or + 2) mod 8;
           itt   := itt - 1;
       end
       edge(pos) := edge(pos) + 1;
       first := false;
   end
end
```

Algoritmen må styres af et program, der bestemmer startpunkterne samt gennemløbsretningerne. En mulig styrealgoritme er at gennemløbe samtlige billedpunkter. Når et ikke markeret kantpunkt mødes kaldes proceduren *tracer* med billedpunktet som startpunkt to gange, nemlig med orientering med hhv. mod urets omløbsretning. Proceduren *tracer* modificeres således at den returnerer en liste af koordinatpar for kantpunkterne i konturen. Styrealgoritmen må da sammenkæde de to delkonturer. En mere avanceret styrealgoritme er beskrevet hos *Pavlidis*.

Ovenstående algoritme er særlig anvendelig hvis kanten af et segment skal “traces”. I tilfældet hvor vi på forhånd har givet at kanterne er én pixel tynde er det let at konstruere en hurtigere algoritme, idet vi kan udnytte at de fleste kanter er lokalt lineære. Dette kræver at vi vedligeholder en preferenceorientering “or”. Lad P være det aktuelle kantpunkt. En mulig implemtation kan da være:

```
procedure trace(P, or, List) begin
    point   p_m, p_0, p_p;
    integer N;

    p_m := nabo(P,(or-1) mod 8);
    p_0 := nabo(P,or);
    p_p := nabo(P,(or+1) mod 8);
    N   := edge(p_m) + edge(p_0) + edge(p_p);

    if (N = 1) begin
         if      (edge(p_0) then trace(p_0, or, cons(List, P));
         else if (edge(p_m) then trace(p_m, (or-1) mod 8, cons(List, P));
         else                    trace(p_p, (or+1) mod 8, cons(List, P));
    end else return(List);
end
```

Ovenstående skitse kan let forbedres således at registreres mere information om hvorledes kanten har artet sig. Dette kan f.eks. udnyttes i situationen N=0, til i visse tilfælde at se længere end blot én pixel frem. Ønsker vi at detektere rette liniestykker er det oplagt at afbryde sporingen når kanten “svinger” mere end usikkerheden på de allerede sporede punkter på kanten antyder er sandsynligt. Bemærk at ovenstående skitse ofte vil bryde sammenhængende kanter op i mindre stykker (Hvad kan der sker ved sporing af kanterne i en T-junction ?).

### 2.12.2 Retningskodning

Ideen ved retningskodning er at udnytte at kantpunkterne som regel er sammenhængende. I stedet for at kode samtlige koordinatpar, kodes kun koordinaterne for starten af kanten samt en følge af retningskoder, angivende den relative retning fra et kantpunkt til det næste. En væsentlig antagelse er at kanterne kun er 1 pixel tykke. En sammenhængende (1 pixel tyk) kant indeholdende N kantpunkter kan altså kodes med 2 heltal (start-koordinaterne) plus hhv. 2N hhv 3N bit for 4-sammenhængende hhv. 8-sammenhængende retningskoder. Retningskodning er eksakt, men i sin rå form hverken invariant over for startpositionen, rotation eller skalering. Fordelen ved metoden er primært aden er enkel. Bortset fra start-koordinaterne, er metoden uafhængig af translationer i billedplanet. Metoden kan let modificeres således at den bliver invariant mht. rotationer. Derimod er invarians over for skaleringer svært at opnå.

Det er muligt at optimere kodningen ved at erstatte en sekvens af ens retningskoder ved et antal samt retningskoden. Hvor meget denne optimering gavner afhænger naturligvis af hvor retlinede og hvor lange kanterne er.

### 2.12.3 Kodning ved Fourierkoefficienter

En anden kodningsmetode af en kant er at registrere samhørende x- hhv. y- koordinater for en kant i to endimensionale "arrays", at foruriertransformere disse og at registrere de N føste fourierkoefficienter for de to koordinatfunktioner. Alternativt kan koordinaterne kodes som et komplekst tal $z = x + iy$, resulterende i kun en funktion. Det skal bemærkes at de to metoder vil give forskellige kodninger (og rekonstruktioner).

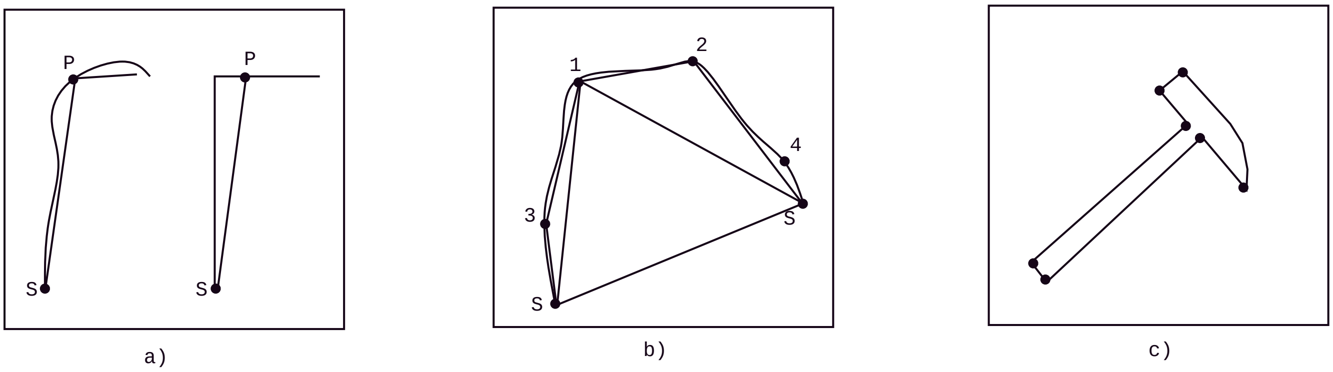
Kodningen er oplagt mere tidskrævende end retningskodning, men har en række fordele (eller ulemper, afhængigt af synspunkt). Ved registrering af samtlige fourierkoefficienter kan kanten rekonstrueres eksakt. Dette vil imidlertid ikke give mogen acceptabel kompressionsfaktor. Tværdigmod vil lagerforbruget stige, da vi for hver koordinatpar $(x, y)$ før skulle lagre 2 heltal (af begrænset størrelse) i stedet skal lagre et komplekst tal. Alternativt kan man eksempelvis nøjes med at registrere de 17 første koefficienter, eller det antal koefficienter der svarer til bevaring af (eksempelvis) 95% af energien i koordinatfunktionen. Ved alternativet kan kompressionsfaktoren øges. Dog kan funktionerne så ikke rekonstrueres eksakt. En væsentlig fordel ved kodning med fourierkoefficienter er at kodningen kan modificeres således at denne er invariant over for både translation, rotation og skalering. Skalering hhv. rotation af kanten vil bevirke multiplikation af samtlige fourierkoefficienter med en reel konstant hhv. en imaginær konstant $e^{i\theta}$. Der kan derfor normaliseres for skalering og rotation ved at dividere samtlige fourierkoefficienter med F(0). Normalisering mht startpunktet er noget vanskeligere (og skal ikke diskuteres her), Se evt *Gonzales og Woods: Digital Image Processing; 1992, p. 497-502.*

### 2.12.4 Kodning ved polygonale approximationer

I kodning ved polygonale approximationer opdeles hver kant i et sæt af kantsegmenter, der hver approksimeres ved et ret liniestykke. Mere avanceret kan man approksimere ved et polynomie. Segmenteres kanten i N segmenter og approksimeres med liniestykker kan kanten altså kodes ved 2(N+1) heltal (koordiatværdierne for endepunkterne). Sædvanligvis approksimeres i mindste kvadraters betydning.

Der findes et utal af metoder til beregning af brydningspunkter. En metode er, at beregne det konvekse hyster af den figur der beskrives af kanten. Brydningspunkterne kan defineres som de kantpunkter hvis afstand til det konvekse hylster er maksimal (og evt. minimal). En anden metode er at transformere koordinatfunktionerne til polære koordinater, og definere brydningspunkterne som lokale ekstrema i argumentfunktionen samt ved diskontinuiteter i radiusfunktionen.

En alternativ metode er, givet et startpunkt på kanten, at betragte arealet mellem kanten og den rette linie defineret ved startpunktet og et (det næste) punkt på kanten. Er arealet mindre end en vis grænse gentages processen for næste kantkunkt. I modsat fald markeres punktet som brydningspunkt samt som nyt startpunkt. Bemærk at denne metode i særlig grad er afhængig af valget af startpunkt. Metoden kan modificeres i det




Figur 2.13: Bestemmelse af brydningspunkter. Billede a) viser to situationer hvor kriteriet er at arealet mellem forbindelseslinien og kurven skal være mindre end en grænseværdi. Billede b) illustrerer brugen af split-and-merge. Billede c) illustrerer kantpunkter med maksimal krumning.

uendelige. I stedet for at betragte arealet, kan lokale maksima i afvigelsen mellem kanten og den rette linie betragtes.

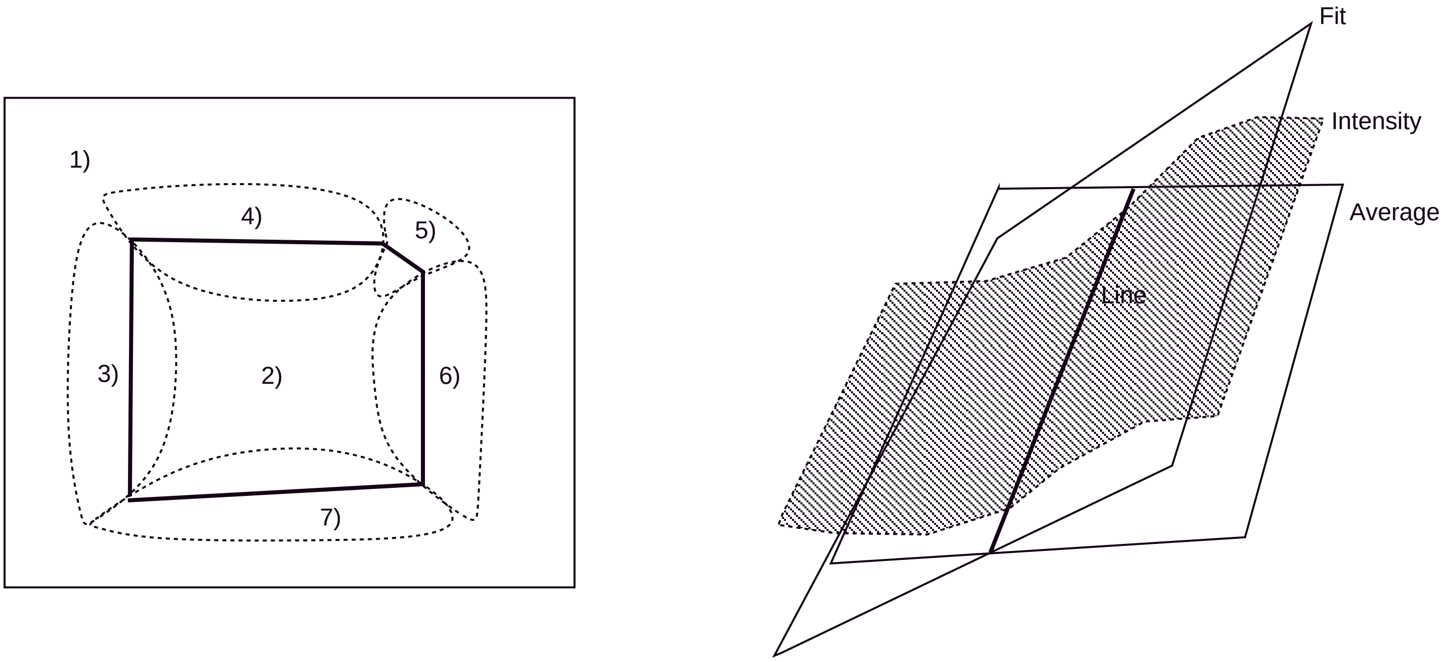
En væsentlig metode er anvendes af *split and merge* -metoder. Først betragtes hele kanten og et “bedste” brydningspunkt findes mht. et optimalitetsmål. Er målet større end en vis grænseværdi registreres brydningspunktet, og processen gentages rekursivt for de to delkanter. Når ingen delkant kan nedbrydes yderligere undersøges samtlige nabo-delkanter. Hvis målet for delkanten bestående af to sådanne nabokanter er mindre en grænseværdien slettes brydningspunktet der adskilder de to delkanter.

En stor klasse af metoder til bestemmelse af brydningspunkter tager udgangspunkt i en måling af krumningen af kanten (som funktion af kurvelængden). Ofte støjfiltreres denne funktion. Brugbare brydningspunkter kan herefter defineres som de kantpunkter hvor krumningen er lokalt maksimal, minimal, eller hvor kantkrumningen har vendetangent. Metoden er mere detaljeret beskrevet i *Asada and Brady: The Curvature Primal Sketch; IEEE PAMI vol. 8 no. 1, 1986.* Se også *Bengtson: Shape Representation by Multiscale Contour Approximation, Proc. SCIA 1989.* Væsenligt i denne metode er muligheden for at styre den detalieringsgrad hvounder kanten ønskes kodet/repræsenteret. Dette sker ved graden af (lav-pas) filtrering.

## 2.13 Liniedetektion

I visse situationer er man ikke interesseret i generelt at detektere kantpunkter, men kun i at detektere rette liniestykker. En oplagt mulighed er at først at detektere kantpunkterne, og derefter gruppere disse således at mængden af kantpunkter inden for hver gruppe ligger på en ret linie. Denne metoder er nok den mest benyttede, men har, jvf. ovenstående afsnit, en række problemer, herunder følsomhed over for startpunkt, diverse grænseværdier mv.

En anden mulighed er at detektere liniestykkerne direkte. Dette kan gøres ved først




Figur 2.14: Liniedetektion. Det venstre billede viser segmenteringen efter gradientorientering. Det højre billede illustrerer liniebestemmelsen i et segment.

at lavpasfiltrere billedet og måle gradientretningen af det filtrerede billede i hver punkt. Derefter segmentere billedet således at hvert segment indeholder punkter med ens (næsten ens) gradientretninger. Denne segmentering er ikke helt let, hvilket vi skal diskutere videre i Kapitel 3. Nu tilpasses en plan flade til intensiteten i hvert segment, hvor middel gradientstørrelen er over en vis grænseværdi. Desuden bestemmes middelintensiteten i segmentet. Endelig skæres det fundne plan med middelintensiteten, og den del af den resulterende linie, som ligger inden for segmentet registreres. En ulempe ved metoden er at den er ret følsom mht. måling af middelintensiteten og af segmenteringens kvalitet.

En tredie metode til detektion af rette linier i billeder er af foretage en såkaldt *Hough-transformation.*

## 2.14 Hough-transformationen

Hough-transformationen er egentlig ikke en metode til hverken detektion eller kodning af kanter, men en metode til lokalisering af parametriserede kurver. Metoden medtages her fordi den ofte benyttes efter en indledende kantdetektion, og fordi den i visse situationer med fordel kan indbygges i en kantdetektion. Forudsætningen for at benytte metoden er, at der søges efter kurver som er simpelt parametriseret. Det klassiske eksempel er søgning efter rette liniestykker. Ideen er at foretage en (ikke reversibel) transformation fra billedplanet til et parameterrum. Det antages at de enkelte parametre varierer diskret mellem en kendt minimal og maksimal værdi. For en ret linie vil man ofte benytte parametriseringen:

$$y = ax + b \tag{2.147}$$

Parameterrummet er to-dimensionalt og udspændt af $a$ og $b$. Hough-transformationen består i korthed af, for hvert detekteret kantpunkt $(x_0, y_0)$ at tælle *Hough-accumulatoren* $A[a][b]$ en op for samtlige par $(a, b)$ der opfylder overstående ligning, dvs:

$$A[a][b] := A[a][b] + 1 \tag{2.148}$$
$$\text{for } a_{min} < a < a_{max} \; og \tag{2.149}$$
$$b = y_0 - ax_0 \tag{2.150}$$

Når transformationen er foretaget kan linierne bestemmes ved de $(a, b)$ hvor accumulatoren $A$ har lokalt maksimum. Problemet ved ovenstående transformation er at grænserne for $a$ er $\pm\infty$ for lodrette linier. Problemet kan let løses ved at benytte en anden parametrisering:

$$x \cos\theta + y \sin\theta = \rho \tag{2.151}$$

Det skal bemærkes at detektionen kun vil angive lokalisering af linier, og ikke udstrækningen af liniestykkerne.

Hough-transformationen kan (eksempelvis) også anvendes ved detektion/lokalisering af cirkler :

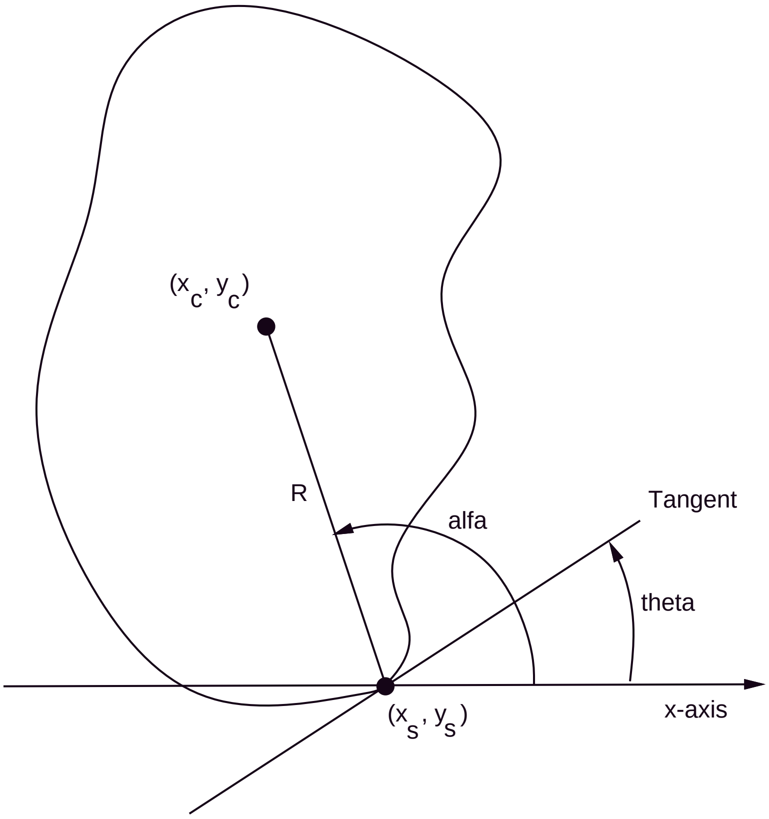
$$(x - a)^2 + (y - b)^2 = r^2 \tag{2.152}$$

Her er parameterrummet tre-dimensionalt. Eksemplet illustrerer at anvendeligheden af Hough-transformationen ligger i parameterrummets størrelse. Transformationen vil være beregningsmæssig kostbar, hvis antallet af parametre er stort, eller hvis antal diskrete værdier som hver parameter kan antage er stort. Hough-transformationen kan benyttes ved detektion af mere komplicerede parametriserede former. Et eksempel hvor ellipser anvendes er eksempelvis beskrevet i *Ballard and Brown: Computer Vision, Prentice Hall 1982.*

Det er også muligt at generalisere Hough-transformationen til geometriske figurer, der ikke er parametriserede men tabellerede. Tabelleringen af en silhouet af et kendt "objekt" foregår ved at vælge et referencepunkt $(x_c, y_c)$ inden for dette. Derefter beregnes afstanden $R$ for hvet punkt (x,y) på silhouetten til referencepunktet, samt tangentretningen $\theta$ til silhouetten i (x,y) og vinklen $\alpha$ mellem x-aksen og linien der forbinder $(x_c, y_c)$ og (x,y). Tabeller nu de to funktioner $\alpha(\theta)$ og $R(\theta)$. Antag nu at de to tabeller er givet og lad $(x_s, y_s)$ være et kantpunkt med tangentretning $\theta_s$. Da vil de mulige lokaliseringer af referencepunktet være givet ved:

$$x_c = x_s + R(\theta_s) \cos\alpha(\theta_s) \tag{2.153}$$
$$y_c = y_s + R(\theta_s) \sin\alpha(\theta_s) \tag{2.154}$$

Hough-akkumulatoren A tælles, som tidligere beskrevet, 1 op for alle punkter $(x_c, y_c)$ der opfylder ovenstående ligning, og lokaliseringen af referencepunkterne for figurer af den




Figur 2.15: Hough-transformationen for tabbelerede figurer.

tabellerede type findes ved bestemmelse af lokale maksima i akkumulatoren. Ovenstående skitse tager ikke højde for at en aktuel silhouet kan være roteret (eller skaleret) i forhold til prototypen. Det er dog relativt let at indføje dette ved at øge dimensione af akkumulatoren. Tilføjelse af en rotationsdimension $\phi$ og en skaleringsdimension $S$ vil ændre ovenstående ligninger til:

$$x_c = x_s + SR(\theta)\cos(\alpha(\theta) + \phi) \tag{2.155}$$

$$y_c = y_s + SR(\theta)\sin(\alpha(\theta) + \phi) \tag{2.156}$$

Ovenstående modifikation øger altså akkumulatorens dimension fra to til fire. Den beregningsmæssige omkostning ved transformationen stiger derfor svarende til at tre (mod før en) parametre skal gennemløbes for hvert kantpunkt. Den generelle Hough-transformation er beskrevet i *D. H. Ballard: Generalizing the Hough transform to arbitrary shapes; Pattern Recognition 13,2; 1981.*

Der eksisterer en lang række varianter af Hough-transformationen. Eksempelvis kan estimationen foretages adaptivt, således at kvantificeringen i starten er meget grov, men for hver iteration forfines omkring de tidligere fundne lokale maksima i akkumulatoren. Dette øger beregningshastigheden væsenligt. I andre varianter søges metodens påvirkelighed af støj begrænset ved mere avancerede opdateringsmetoder.

Hough-transformationen er særlig anvendelig hvis antallet af kantpunkter hørende til søgte strukturer er meget større end antallet af irrelevante kantpunkter. Hvis billedet, eksempelvis, er meget tekstureret vil det/de korrekte maksima i akkumulatoren ofte drukne i støj. Metoden skal derfor anvendes med varsomhed.

## 2.15 Opgaver

1. Udled formlen for den andenafledede af en kontinuert to gange differentiabel funktion E(x,y) i retning givet ved gradientretningen. Er udtrykket lineært ?

2. Beskriv forskellen mellem anvendelse af de forskellige differensapproksimationer til den første afledede af et signal. Hvad er effekten af at øge $h$? Hvad er effekten af at foretage en boxfiltrering før den diskrete diffentiation? Hvorledes varierer svaret (i en omegn om en kant) som funktion af boxfiltrets størrelse?

3. Udled differensapproksimationerne til de 2. og 3. ordens afledede af et signal.

4. Udled de seks anden ordens Beaudetoperatorer med støtte $4 \times 4$ pixels. Det kan anbefales at benytte et matematikprogram som f.eks. MATHLAB.

5. Diskuter problemerne relateret til valg af kantmodel, og optimalitet af filtreringsmetode. Tag udgangspunkt i *LeClerk and Zucker: The Local Structure of Discontinuities in One Dimension, IEEE PAMI vol. 9 no. 3, 1987.*

6. Skriv et program, der måler de førsteordens afledede vha. de centrale differensapproksimationer med støtte 3x1 hhv. 1x3. Beregn "gradientstørrelsen" og foretag grænseværdiafskæring ud fra en parameterangivet værdi.

7. Diskuter hvorledes de to parametre i en grænseværdiafskæring med hysterese kan bestemmes internt (og evt. adaptivt).

8. Implementer Laplace-Gaussfiltret som en sum af to seperable filtre, og detekter nulpunktsgennemskæringerne i resultatet af foldning af et billede med operatoren.

9. Sammenlign eksperimentelt resultatet af kantdetektionerne givet ved:
   a) Laplace-Gauss kantoperatoren (LGzc)
   b) Deriche kantoperatoren (deriche)
   c) Deriche kantoperatoren med hysterese (derihyst)
   d) 3x3 Beaudet-operator med grænseværdi (gradthres)
   e) Udtynding af kanter beregnet som i d) (gradient3x3 samt skeleton)

10. Beskriv (teoretisk) resultatet af en kantdetektion vha. Laplace-Gauss operatoren på et billedudsnit hvor tre områder, hver med konstant men forskellig intensitet, mødes.

11. Foretag en "håndkørsel" af algoritmen til udtynding af kanter på en akseparallel kantfigur af længde 12 pixels og bredde 5 pixels. Beskriv resultatet. Foretag en tilsvarende "håndkørsel" på en "skrå" kantfigur. Beskriv hvor stor en del af figuren, der fejlagtigt slettes som ikke-skeletpunkter.

12. Foretag en binarisering af et gråtonebillede (f.eks. ved simpel tærskelværdiafskæring), og anvend algoritmen til udtynding af kanter. Beskriv resultatet.

13. Konstruer et program til sporing af kanter i et binært kantbillede, f.eks. beregnet ved anvendelse af en Laplace-Gauss kantdetektion. Hvorledes er traceren afhængig af startpunkt, af forgreninger, støj, "dobbeltmarkeringer" mv.

14. Konstruer et program til detektion af rette liniestykker i en binært kantbillede. Uddata er en fil hvor hver kantstykke repræsenteres ved koordinaterne for endepunkterne af liniestykket.

15. Konstruer et program, der på basis af en kandetektion, koder kanterne ved "Chain"-koder. Diskuter problemerne og eksperimenter med kodningens effektivitet.

16. Konstruer et program, der givet en sporet kant i et billede beregner en fourierkoefficientkodning hvor kun p procent af energien er bevaret. Antag her at p er brugerbestemt. Konstruer også et program til optegning af den kodede kant.

17. Diskuter problemerne relateret til forholdet mellem kantdetektion og kantfølgning.

18. Implementer en detektion af linier ved Hough-transformation af resultatet af en kantdetektion.

19. Som forrige opgave, blot detektion af cirkler.

20. Foretag en eksperimentel analyse af i hvilken grad forskellige kantdetektorer er i stand til at detektere kantpunkter i nærheden af "T"-, "Y", og "X"-foreninger.

21. Diskutér hvorledes den tredieafledede kan benyttes ved detektion af kanter. Hvordan kan den lokale opførsel af den andenafledede, i en omegn om en nulpunktsgennemskæring) benyttes. Implementér metoderne og foretag en evaluering.

22. Implementer en billedfiltrering hvor den filtrerede intensitet lokalt bestemmes ved et robust lokalt fit vha. Hough-transformationen. Sammenlign med resultatet af en medianfiltrering. Diskutér om/hvorledes kanter i billedet kan bestemmes på basis af filtreringen.

23. Diskutér om/hvorledes en kantdetektion baseret på foldning med filtre med variabel støtte kan foretages. Problemet er, at ingen fast støttestørrelse vil være optimal. Hvis intensiteten er glat i en stor omegn bør filterstørrelsen være stor. Hvis billedet lokalt indeholder fin textur skal støtten være lille. I modsat fald vil nabokanter få indflydelse på detektionen.

24. Diskuter Cannys kriterier for optimalitet på basis af det faktum at filteret $f(x)$ implementeres diskret.

25. Diskuter eventuelle ulemper ved Deriche's filter.