



UNIVERSITETET I OSLO
Institutt for informatikk

**Lokalisering av
multistråle
sonarfotavtrykk
med
todimensjonale
bildebehandlings-
metoder**

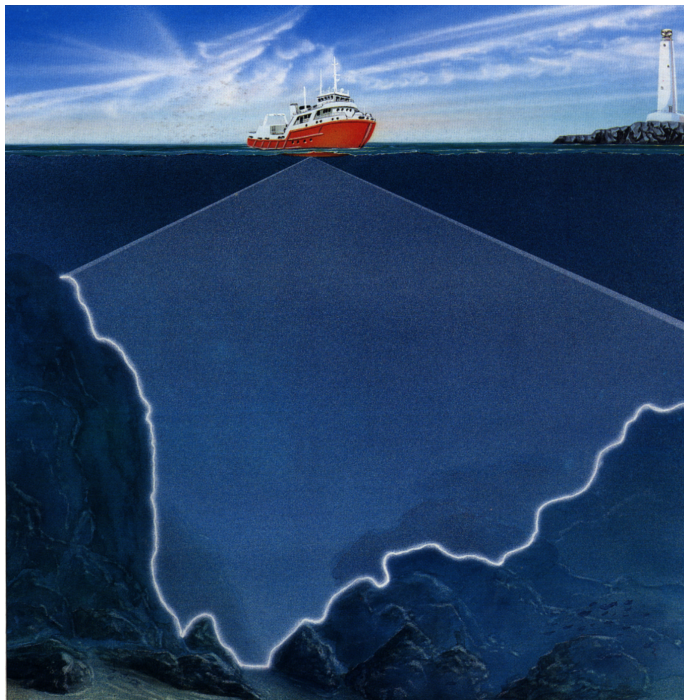
Jens Toivo Berger
Thielemann

Hovedfagsoppgave

Mai 1999



Lokalisering av multistråle sonarfotavtrykk med todimensjonale bildebehandlingsmetoder



Jens T. Berger Thielemann

Hovedfagsoppgave

Institutt for Informatikk

Universitetet i Oslo



iv

Forsidebildet gitt av Kongsberg SIMRAD.

Sammendrag

Multistrålesonarer er i dag viktige verktøy ved kartlegging av havbunnen. Kartleggingen gjøres ved å sende lydimpulser mot havbunnen, og registrere refleksjonen av disse lydimpulsene i hvert mottakerelement. Etter analog/digital-konvertering av elementsignalene brukes disse signalene til stråleforming i mange vinkler. Det tidsrommet innenfor hver stråle som inneholder refleksjonen fra havbunnen kalles strålens fotavtrykk. Samlingen av alle strålenes fotavtrykk kalles sonarfotavtrykket. En god lokalisering av sonarfotavtrykket er viktig for senere deteksjon av bunnekk og påfølgende kartproduksjon.

Eksisterende metoder for lokalisering av fotavtrykk er enkle, men gir et upresist resultat. Dette arbeidet har studert tersklingsmetoder som kan gi mer presis lokalisering av fotavtrykket.

Fem forskjellige lokalt adaptive todimensjonale tersklingsmetoder har blitt evaluert. Disse fem metodene kom best ut i en omfattende evaluering av tersklingsmetoder. Metodene ble testet eksperimentelt på et testsett bestående av femten sonarbilder fra en SIMRAD EM3000 multistrålesonar. Sonarbildene var fra tre forskjellige typer havbunn.

Bernsen's (1986) metode med ikke-kvadratiske vinduer ga det beste resultatet. Gradientbasert postprosessering og glatting av det tersklede bildet forbedret resultatet ytterligere. Mindre enn 5% av sonarfotavtrykket ble feilklassifisert som bakgrunnsstøy og ca 10% av bakgrunnsstøyen ble feilklassifisert som fotavtrykk. Feilen til metoden brukt som referansem metode var fire ganger så stor. Den nye metoden benyttet imidlertid ti ganger mer regnetid enn referansemetoden. Total regnetid per bilde overstiger likevel ikke to sekunder på en SUN SPARCstation 10.

Flere av de lokalt adaptive todimensjonale tersklingsmetodene var mer presise enn referansemetoden. Implementasjon av en eller flere av disse metodene i fremtidige sonarer vil trolig medføre betydelig mer nøyaktige havbunnskart.

Til Marit

Innhold

1	Introduksjon	1
2	Multistrålesonarer	5
2.1	Sonararrayets geometri og signaler	5
2.1.1	Sending og mottak av signal	5
2.2	Stråleforming	8
2.2.1	Forsink-og-summér	9
2.2.2	Fourier transformasjonen	9
2.3	Bunndeteksjon	10
2.3.1	Amplitudebaserte metoder	11
2.3.2	Fasebaserte metoder	13
2.4	Karttegning	15
2.4.1	Batymetriske kart	15
2.4.2	Refleksjonsbilder	15
3	Sonarfotoavtrykk	17
3.1	Sonarfotoavtrykkets egenskaper	17
3.1.1	Signalene innenfor fotoavtrykket	17
3.1.2	Sekundært ekko — uønsket reverberasjon	18
3.2	Bruk av sonarfotoavtrykk	18
3.3	Etablerte metoder for lokalisering av sonarfotoavtrykk	19
3.3.1	Ikke-kontekstuell global terskling av amplitudebildet	19
3.3.2	Enkel endimensjonal terskling	20
4	Nye metoder for lokalisering av sonarfotoavtrykk	21
4.1	Lokal adaptiv terskling	21
4.1.1	Bernsens metode	22
4.1.2	Eikvil, Taxt & Moens metode	23
4.1.3	Nibblacks metode	25
4.1.4	Yanowitz & Brucksteins metode	25
4.1.5	Forbedret White & Rohrer	26
4.2	Fjerning av falske objekter	27
4.2.1	Fjerning av initiell støy	27
4.2.2	Fjerning av objekter i homogene områder	28
4.3	Glating av tersklet bilde	29
5	Data og eksperimenter	31
5.1	Data	31
5.1.1	Stråleforming	32
5.1.2	Preprosessering	32
5.2	Evalueringsmetoder	36
5.2.1	Etablerte evalueringsmetoder	37
5.2.2	Evaluering av segmenteringsresultat	37
5.2.3	Evaluering av regnetid	38
5.2.4	Referansebilder	38

6 Eksperimentelle resultater	41
6.1 Segmenteringsresultat	41
6.1.1 Objektiv vurdering	41
6.1.2 Subjektiv vurdering	42
6.1.3 Parameterverdier	48
7 Diskusjon og konklusjon	53
A Eksperimentenes utførelse	57
A.1 Parallell kommandoutføring	57

Figurer

2.1	Skisse av en multistrålesonar.	5
2.2	Flytdiagram for kvadraturdemodulering	7
2.3	Rådata fra sonarmottakerene	7
2.4	Skisse av et primitivt mottakerarray	8
2.5	Stråleforming med FFT av rådata fra mottakerene	10
2.6	Multipel refleksjon innenfor en stråle som følge av ujevn havbunn.	11
2.7	Estimering av ankomstvinkel fra amplitudebilde ved tid t	12
2.8	Faseinterferometri ved hjelp av to subarrayer	13
2.9	Resultat av interferometri med flere subarrayer	14
2.10	Eksempel på deteksjon av tilnærmet flat bunn	15
2.11	Batymetrisk kart over havbunnen i Halifax havn i Canada	16
2.12	Refleksjonsbilde over et skipsvrak	16
3.1	Illustrasjon av en stråles fotavtrykk	17
3.2	Resultat av forskjellige signalbehandlingsmetoder for sonarsignaler fra flat bunn	18
3.3	Sekundært ekko — reverberasjon	19
4.1	Stegvis illustrasjon av Bernsen's (1986) metode.	23
4.2	Prinsipp for metoden til Eikvil, Taxt & Moen (1991)	24
4.3	Fjerning av «spøkelsesobjekter»	28
5.1	Eksempel på amplitudebilde og referansebilde til et ping fra hver testserie	33
5.2	Histogram til et amplitudebilde.	34
5.3	Plot av Rice-fordelingen.	35
5.4	Fordelingen for $\log[R + \epsilon]$ hvor R er Rice-fordelt	35
5.5	Resultat etter logaritmisk transform av amplitudebilde med påfølgende gjennomsnittsglatting, (3×3) vindu	36
5.6	Forenkling av manuell terskling ved hjelp av Photoshop	39
6.1	Resultat etter postprosessering	43
6.2	Resultatbilder etter postprosessering	44
6.3	Resultatbilder etter segmentering (før postprosessering)	46
6.4	Type I og Type II feil til segmenteringsresultatet som funksjon av utvalgte parametre	50

Tabeller

3.1	Parameterverdier for endimensjonal terskling.	20
4.1	Parameterverdier for Bernsens metode.	23
4.2	Parameterverdier for Eikvil, Taxt og Moens metode.	25
4.3	Parameterverdier for Niblack's metode.	25
4.4	Parameterverdier for Yanowitz & Brucksteins metode.	26
4.5	Parameterverdier for forbedret White & metode	27
4.6	Parameterverdier for fjerning av objekter i homogene områder	29
4.7	Parameterverdier for morfologiske operasjoner på tersklet bilde	29
4.8	Parameterverdier for medianglatting av tersklet bilde.	29
5.1	Noen viktige parametere for en SIMRAD EM3000 sonar.	32
5.2	Pikselverdier for tegning av referansebilder	39

Kapittel 1

Introduksjon

En god kartlegging av havbunnen gjør det mulig å bedre utnytte det mangfold av ressurser som havet besitter — olje, gass og mineraler. Å kunne kartlegge havbunnen godt er óg viktig for at skip og undervannsfartøy skal kunne seile trygt og uten forsinkelser.

Sonarer er dagens hovedverktøy for å kartlegge havbunnen. Kartleggingen gjøres ved å sende lydimpulser mot havbunnen, og måle refleksjonen av disse lydimpulsene. Disse målingene brukes til å danne kart over havbunnen. To typer «kart» benyttes: Batymetriske kart som kartlegger havdybden, og refleksjonsbilder som viser hvor godt havbunnen reflekterer lydimpulsen.

I dag finnes to typer sonarer som kartlegger havbunnen: Batymetriske side-seende sonarer og multistrålesonarar. Side-seende sonarer benytter to parallelle lineære arrays av kombinerte sendere og mottakere («transducere»). Transducerene som benyttes omformer akustisk energi til elektrisk energi og vice versa. Interferometriske metoder (de Moustier 1993, sec. 3.1) benyttes for å finne havbunnen. Imidlertid har side-seende sonarer problemer med ujevnt terreng.

Multistrålesonarar har ikke disse terrengavhengige problemene i samme grad. De benytter separate sende- og mottakerarrays som står normalt på hverandre. For å analysere signalet dannes stråler i mange vinkler ved hjelp av stråleformingsmetoder som for eksempel FFT eller forsink-og-sommer (Johnson & Dudgeon 1993). Stråleforming gjør det mulig å beregne signalets amplitude i forskjellige vinkelsektorer («stråler»). Ut fra disse beregningene kan man med amplitudebaserte eller interferometriske metoder finne mottakstidspunktet for bunnekket innenfor hver stråle.

Prosessen som lokaliserer bunnekket kalles bunndeteksjon og er essensiell for tegning av batymetriske kart og refleksjonsbilder. Det tidsrommet innenfor hver stråle som inneholder refleksjonen fra havbunnen kalles strålens fotavtrykk. Fotavtrykket er det tidsrommet der det mottatte signalet har en relativt høy amplitude innenfor strålen. Lengden på dette intervallet varierer som følge av strålens vinkel i forhold til bunnen. Samlingen av alle strålens fotavtrykk kalles sonarens fotavtrykk. En god lokalisering av fotavtrykket for hver stråle er viktig for en senere bunndeteksjon. En god bestemmelse av fotavtrykket gjør det også mulig å karakterisere havbunnen pålitelig.

I dag er metodene for å finne fotavtrykket for multistrålesonarar ganske enkle. De baserer seg på en-dimensjonal terskling og likefremme antagelser om bunngeometrien (de Moustier 1993, Lingsch & Robinson 1992, Satriano & Smith 1991). Ytelsen er derfor nokså begrenset. Dette kompliserer senere bunndeteksjon, og gjør bunndeteksjonen upresis.

Utviklingen innenfor hardwareteknologi og signal- og bildebehandlingsmetoder gjør det mulig å bruke todimensjonale signal- og bildebehandl-

ingsmetoder. Det finnes i dag mange tilgjengelige todimensjonale bildeanalysemetoder for segmentering av bilder, og det er skrevet atskillige oversiktsartikler og bøker om emnet (f.eks Pratt 1991, Zhang 1997, Trier & Taxt 1995*b*, Pal & Pal 1993).

Denne hovedfagsoppgaven forsøker å finne metoder som kan forbedre presisjonen ved lokalisering av sonarfotavtrykket. Jeg forsøker å forbedre presisjonen i tre trinn:

1. *Adaptive todimensjonale metoder.* Eksisterende metoder for lokalisering av fotavtrykk arbeider i beste fall endimensjonalt med dataene. Felles for lokalt adaptive bildetersklingsmetoder er at de arbeider todimensjonalt med dataene. Dermed kan mer lokal informasjon benyttes til å gjenkjenne fotavtrykket, noe som kan bedre presisjonen ved lokalisering.

I dette studiet har jeg testet fem lokalt adaptive bildetersklingsmetoder (Bernsen 1986, Eikvil et al. 1991, Niblack 1986, Yanowitz & Bruckstein 1989, Trier & Taxt 1995*a*). Disse fem metodene kom best ut i en omfattende evaluering av tersklingsmetoder for dokumentbilder (Trier & Taxt 1995*b*). Samtlige metoder er allerede implementert i bildebehandlingsverktøyet XITE (Bøe, Lønnestad & Milvang 1998, Bøe 1998*b*, Bøe 1998*a*).

2. *Fjerning av falske objekter.* Lokalisering av fotavtrykket ved hjelp av bildetersklingsmetoder gir flere disjunkte regioner som muligens tilhører fotavtrykket. Et problem er ofte at enkelte av disse regionene ikke tilhører fotavtrykket. For eksempel vil en del tersklingsmetoder (også eksisterende lokaliseringmetoder) klassifisere reverberasjoner som en del av fotavtrykket.

Metoder basert på gradientverdier (Yanowitz & Bruckstein 1989, Trier & Taxt 1995*b*) eller amplitudeverdier (Yang 1998*a*) testes for å fjerne slike falske regioner.

3. *Glatting av endelig resultat.* Når fotavtrykket er lokalisert, kan det være usammenhengende og med ujevne kanter. Ved å glatte det lokaliserte fotavtrykkets kanter og fylle inn hull i fotavtrykket kan resultatet trolig forbedres (Pratt 1991).

Mangfoldet av tersklingsmetoder skyldes at det ikke finnes noen metode som fungerer best for alle tersklingsproblemer. Derfor er det nødvendig å sammenligne resultatene fra mange lovende metoder. Evalueringen ble automatisert ved bruk av objektive mål dannet på grunnlag av enkle betraktninger om segmenterte bilder (Zhang 1996). For å skape en objektiv evaluering, tegnet jeg referansebilder ved å segmentere sonarbildene manuelt. Sonarfotavtrykket i disse referansebildene og sonarfotavtrykket i de automatisk segmenterte bildene ble sammenlignet piksel for piksel.

I dette studiet rangerte jeg de utvalgte metodene ved å teste mange metode-/parameterkombinasjoner. Cirka 18000 forskjellige kombinasjoner av metodesett og parametre ble undersøkt på et testsett bestående av femten amplitudebilder fra multistrålesonarar. Disse femten bildene er deler av tre forskjellige serier med sonardata fra havområder med ulike bunnforhold. Totalt har cirka 300000 tersklingsresultater blitt generert og evaluert. Genereringen og evalueringen skjedde parallelt i et PVM-nettverk (Geist, Beguelin, Dongarra, Jiang, Manchek & Sunderam 1994). Prosessen ble styrt av en spesialtilpasset make (Stallman & McGrath 1998, Zimmermann 1997).

Resultatene viste at Bernsen's (1986) metode var best egnet for å lokalisere fotavtrykk. Postprosesseringssteget til Yanowitz & Bruckstein's (1989) metode som fjerner falske objekter forbedret resultatet ytterligere (Trier & Taxt 1995*b*). Glatting av det segmenterte bildet med medianoperatoren eller en morfologisk operator ga en ytterligere forbedring av resultatet.

I det beste endelige resultatet ble mindre enn 5% av sonarfotavtrykket feilklassifisert som bakgrunnsstøy og ca 10% av bakgrunnsstøyen ble feilklassifisert som fotavtrykk. Feilen til referansemetoden var fire ganger så stor. Den nye metoden benyttet imidlertid ti ganger mer regnetid enn referansemetoden. Total regnetid per bilde overstiger likevel ikke to sekunder på en SUN SPARCstation 10.

Resten av denne hovedfagsoppgaven omfatter sju kapitler. I de to første kapitlene presenterer jeg multistrålesonarens oppbygning og virkemåte, samt egenskapene til sonardataene. Sonarfotavtrykkets egenskaper og bruksområde omtales spesielt.

Kapittel fire beskriver bildebehandlingsmetodene jeg benyttet, med vekt på lokalt adaptive tersklingsmetoder.

Kapittel fem omtaler først de data metodene ble testet på, og hvordan dataenes egenskaper ble forbedret før senere lokalisering. Eksperimentdelen av kapitlet diskuterer evalueringsmetoder for objektiv vurdering av resultatene.

Det påfølgende kapittel seks gir de eksperimentelle resultatene. De nye lokalt adaptive tersklingsmetodene og resultatene diskuteres videre i kapittel sju med tanke på relevans for fremtidige sonarer.

Den faktiske gjennomføringen av eksperimentene var svært regnekrevende. Tillegg A gir en oversikt over hvordan jeg paralleliserte beregningene og dermed reduserte tidsforbruket drastisk.

Takk

Først og fremst takker jeg mine to veiledere, Torfinn Taxt og Luren Yang, for uvurderlige kommentarer ved skrivingen av denne oppgaven. Jeg takker Fritz Albregtsen som velvillig leste og kommenterte et siste utkast av oppgaven. Jeg takker Are Rønhovde for nyttige kommentarer til resultatene og inspirerende diskusjoner underveis. Jeg takker Svein Bøe og driftsgruppen ved instituttet, som lånte meg 1GB disk til å utføre eksperimenter på. Jeg takker Roger Hansen og Sigurd W. Larsen for mye inspirasjon over en kopp te.

Kapittel 2

Multistrålesonarar

2.1 Sonararrayets geometri og signaler

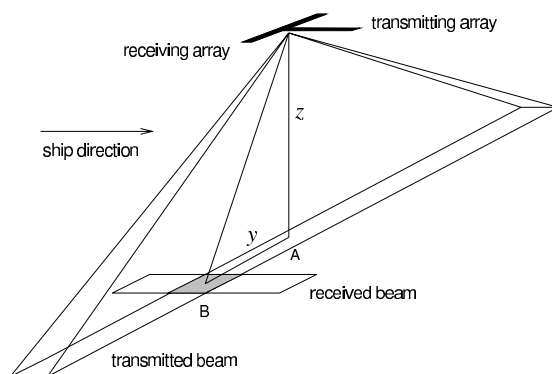
Multistrålesonarar benytter to arrayer for å henholdsvis sende og motta lyd-pulser. Hvert array består av 80-128 sendere/mottakere. Senderne monteres langs en linje parallell med skipets fartsretning, og med lik avstand mellom senderne. Mottakerene festes på tilsvarende måte, men linjen de monteres langs står vinkelrett på sendearrayet (Figur 2.1). Ved å montere arrayene slik, er det mulig å oppnå god oppløsning ved kartleggingen av havbunnen (Johnson & Dudgeon 1993, avsn. 3.3), (de Moustier 1993, avsn. 4). Dette kapitlet gir en oversikt over multistrålesonarens oppbygning og virkemåte.

2.1.1 Sending og mottak av signal

Sendearrayet sender med ujevne tidsintervaller en lyd-puls mot sjøbunnen. Etter at en puls har blitt sendt, venter man med å sende en ny puls inntil hele den sist sendte pulsen har blitt mottatt og behandlet. Tiden det tar før ny puls sendes ut er altså avhengig av havdybden (Augustin, Le Suave, Lurton, Voisset, Dugelay & Satra 1996, de Moustier 1993).

På grunn av begrensninger i senere prosessering (Fourier transformasjonen, avsnitt 2.2.2), betraktes kun fjernfeltet til signalet. Strålen som sendes ut vil i fjernfeltet være formet som en vifte (Figur 2.1).

For en typisk multistrålesonar vil strålen ha en utstrekning på ca 120° i



Figur 2.1: Skisse av en multistrålesonar.

retningen normalt på skipets fartsretning, og ca 1.5° i skipets fartsretning (Kongsberg SIMRAD 1994). Denne vifteformen skyldes at senderne ligger på en linje som er parallell med skipets fartsretning. Som beskrevet i Johnson & Dudgeon (1993) gir et linjeformet array høy oppløsning i linjens retning, og tilsvarende lav oppløsning normalt på linjens retning.

Strålen som mottas har i fjernfeltet også en form som en vifte. Denne viften er vridd 90° i forhold til sendestrålen. For en typisk multistrålesonar vil denne viften ha en utstrekning på ca 1.5° i retningen normalt på skipets fartsretning. Den mottatte strålen kan styres («stråleformes») i ønsket retning. Avsnitt 2.2 omhandler stråleforming av signalet og dannelsen av sonarbildet der bunnekkene lokaliseres. Stråleforming gjør det mulig å studere amplituderresponsen til små sektorer i det vifteformede mottatte signalet, og åpner dermed for identifikasjon av små regioner på havbunnen. Fysiske faktorer begrenser likevel den endelige oppløsningen.

Begrensninger i oppløsning

En vanlig måte å bedre oppløsningen på i det resulterende sonarbildet (altså å minske størrelsen på de tidligere nevnte enkeltregionene) er å øke lengden på arrayet. Et lengre array medfører en smalere hovedlobe og dermed bedre romlig oppløsning (Johnson & Dudgeon 1993, s. 144). Imidlertid begrenser fysiske faktorer bruken av lange arrayer, typiske sonarer har i dag 80–128 elementer i arrayene.

I tillegg til arraylengden er oppløsningen også avhengig av en godt tilpasset signalfrekvens. Høyfrekvente signaler vil i de fleste sammenhenger (Johnson & Dudgeon 1993) gi en bedre oppløsning. Høyfrekvente signaler dempes imidlertid mest av sjøvannet. Dette medfører at det må gjøres en avveining mellom demping og oppløsning, avhengig av den aktuelle havdybden.

Lavfrekvente signaler trenger ned i havbunnen, og påvirkes av grunnforholdene under bunnen. Dette setter også en nedre grense for valg av frekvens. Slik gjennomtrengning vil skape uønsket støy i signalene, og forlenge den temporale utstrekningen av bunnekket. Senere bunndeteksjonsmetoder (avsnitt 2.3) forsøker derimot å redusere den temporale utstrekningen av bunnekket. Gjennomtrengning av bunnen er derfor ikke ønskelig her.

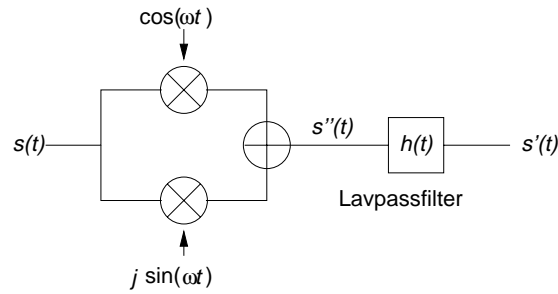
Signalfrekvensen som benyttes avtar derfor med økende dybde av havbunnen. Lavfrekvente signaler, typisk ned til 6 kHz, benyttes for dypvannssystemer (de Moustier 1993), mens frekvenser opp til 300 kHz benyttes for sonarer beregnet på grunt vann.

Kvadraturdemodulering

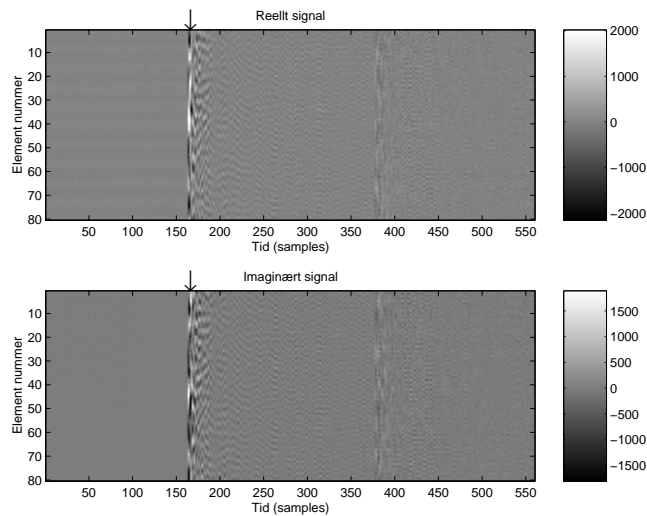
Før digital stråleforming kan gjøres, må det analoge signalet fra hver mottaker samples. Den faktiske digitaliseringsprosessen gjennomføres ved å sample signalet samtidig fra alle mottakerene med jevne tidsintervall. Da et smalbandet (6–8 kHz) høyfrekvent (300 kHz) signal benyttes (Kongsberg SIMRAD 1994), er det lite praktisk å sample signalet direkte. I stedet demoduleres signalet først analogt ved hjelp av kvadraturdemodulering. Dermed kan samplingsfrekvensen reduseres betydelig.

Kvadraturdemodulering («IQ-demodulering») er en vanlig teknikk innenfor signalbehandling av smalband radar- og sonarsignaler (Churchill, Ogar & Thompson 1981). Smalbandsignaler kan representeres ved $s(t) = A(t) \cos(\omega t + \phi(t))$, der bærefrekvensen ω er kjent og konstant. $A(t)$ og $\phi(t)$ er funksjoner som varierer langsomt, mens $\cos(\omega t + \phi(t))$ varierer raskt. Dermed ligger $A(t)$ og $\cos(\omega t + \phi(t))$ i forskjellige frekvensbånd.

Kvadraturdemodulering gjør det mulig å sample det lavfrekvente komplekse signalet $s'(t) = A(t)e^{j\phi(t)}$ i stedet for å måtte sample det reelle signa-



Figur 2.2: Flytdiagram for kvadraturdemodulering. $s(t) = A(t) \cos(\omega t + \phi(t))$ er innkommende reelt signal, $s'(t) = A(t)e^{j\phi(t)}$ er utgående kompleks signal.



Figur 2.3: Rådata fra sonarmottakerene. Pilene angir når første refleksjonen fra sjøbunnen mottas.

let $s(t)$. For en typisk sonar med sendefrekvens 300 kHz ville sampling av $s(t)$ i følge Nyquists samplingsteorem kreve en samplingsfrekvens på over 600 kHz (Oppenheim & Schaffer 1989, s. 86–87). Etter kvadraturdemodulering kan $s'(t)$ samples med en samplingsfrekvens tilpasset $s'(t)$, typisk rundt 8–16 kHz for sonarer.

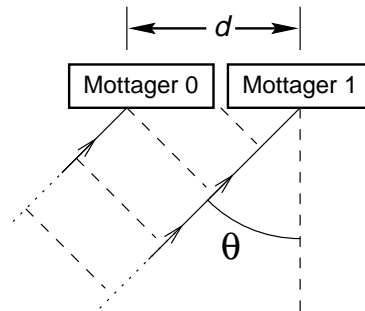
Kvadraturdemodulering genererer først fra det opprinnelige signalet $s(t)$ to reelle signaler $I(t) = 2 \cdot s(t) \cos \omega t$ («In-phase») og $Q(t) = -2 \cdot s(t) \sin \omega t$ («Quadrature»). Ved hjelp av enkle trigonometriske identiteter reduseres disse uttrykkene til

$$I(t) = A(t) \cos[\phi(t)] + A(t) \cos[2\omega t + \phi(t)] \quad (2.1a)$$

$$Q(t) = A(t) \sin[\phi(t)] - A(t) \sin[2\omega t + \phi(t)] \quad (2.1b)$$

Deretter formes det komplekse signalet $s''(t) = I(t) + jQ(t)$. En passende lavpassfiltrering av $s''(t)$ fjerner det siste leddet i (2.1a) og (2.1b), noe som gir det ønskede komplekse signalet $s'(t)$.

Det er enkelt å implementere denne metoden i praksis, da både multiplikasjonen med cosinus-/sinussignalene og lavpassfiltreringen kan gjøres med analog logikk. Dette er skissert i Figur 2.2. Figur 2.3 gir et eksempel på demodulerte og komplekse data fra en faktisk sonar.



Figur 2.4: Skisse av et primitivt mottakerarray. Lydbølgene ankommer mottakerene med en strålevinkel θ . Det er en viss tidsforskjell mellom når bølgefronten ankommer mottaker null og mottaker en. Ved stråleforming utnyttes at denne tidsforskjellen er avhengig av strålevinkelen.

2.2 Stråleforming

Som nevnt i avsnitt 2.1, er det nødvendig å oppdele det mottatte signalet i mange mindre stråler som peker i ulike vinkler. For en typisk multistrålesonar styres strålen i 128 retninger, med en grads mellomrom mellom strålene. Hver av disse strålene har for en slik sonar en utstrekning på ca 1.5° i retningen normalt på skipet (Kongsberg SIMRAD 1994).

De vifteformede sende- og mottaksstrålene står normalt på hverandre, og overlapper hverandre. Denne konfigurasjonen muliggjør måling av refleksjonen fra de små sektorene der strålene overlapper hverandre. Disse sektorene har et omfang på ca $(1.5^\circ \times 1.5^\circ)$ i typiske multistrålesonar. Oppdelingen av signalet i så smale stråler muliggjør en presis kartlegging av enkeltregioner på havbunnen.

Prinsippet som benyttes i stråleformingmetoder, er at et signal som er vinklet i forhold til mottakerene ikke vil ankomme samtidig til alle mottakerene. Dette skyldes at bølgefronten bruker tid på å forplante seg i vannet, og at avstanden mellom kilden og mottakerene er ulik.

En skisse av dette er gitt i Figur 2.4. Her ankommer bølgefronten fra venstre, vinklet $\theta = 45^\circ$ i forhold til mottakerene. Da signalet er i fjernfeltet, er bølgefronten plan (Johnson & Dudgeon 1993, avsn. 2.2). θ angir vinkelen mellom signalets retning og en linje som står normalt på mottakerarrayet. Som det fremgår av figuren ankommer bølgefronten først mottakeren til venstre, og deretter senere mottakeren til høyre.

I dette tilfellet vil forsinkelsen δ_1 mellom signalets ankomst til mottaker «1» i forhold til mottaker «0» være gitt ved (Johnson & Dudgeon 1993, kap. 4)

$$\delta_1 = \frac{d \sin \theta}{c} \quad (2.2a)$$

hvor d angir avstanden mellom de to mottakerene og c er bølgehastigheten i vann. En generalisering av dette uttrykket gir at for N mottakere er forsinkelsen mellom mottaker «0» og mottaker i gitt ved

$$\delta_i = \frac{id \sin \theta}{c} \quad (2.2b)$$

Dette fenomenet kan utnyttes til å forsterke signalet fra en spesiell vinkel, og dempe signaler som kommer fra andre retninger. De to følgende

avsnitt presenterer to standard måter å gjøre slik stråleforming på: Forsink-og-summér og Fourier transformasjonen.

I stedet for en av disse to standard stråleformingsmetodene, kan en høyoppløselig («adaptiv») stråleformingsmetode benyttes. Adaptive stråleformingsmetoder forsøker å minske bredden til hovedloben, for dermed å oppnå bedre oppløsning (Steinhardt & van Veen 1989). Disse metodene gir også bedre signal-støy forhold enn standardmetodene som presenteres her. Til gjengjeld er høyoppløselige metoder mer regnetunge og mer sårbar ved uventede avvik i dataene (Johnson & Dudgeon 1993, kap. 4, 7). En oversikt over bruk av høyoppløselige metoder for multistrålesonar er gitt i Rantzartzis, Moustier & Alexandrou (1993).

2.2.1 Forsink-og-summér

Dette er en svært enkel måte å gjøre stråleforming på (Johnson & Dudgeon 1993, avsn. 4.1). Som nevnt i forrige avsnitt vil signalet $s_i(t)$ for mottaker i være gitt ved

$$s_i[t] = s_1[t - \delta_i],$$

hvor vi plasserer arrayets senter i mottaker 1. Hvis signalet fra mottaker i forsinkes med en passende valgt $\Delta_i = -\delta_i$, vil denne forsinkelsen kansellere forsinkelsen som oppstod som følge av bølgespredningen (Figur 2.4). En påfølgende summering av de forsinkede signalene forsterker signalet i retning θ . Dette gir for et signal $s(t)$ som kommer i retning θ resultatanten $S[T]$;

$$S[t] = \sum_{i=0}^N s_i[t - \Delta_i] = \sum_{i=0}^N s[t - \Delta_i - \delta_i] = \sum_{i=0}^N s[t] \quad (2.3)$$

og den ønskede forsterkning oppnås. Signaler som kommer fra andre retninger vil forskyves i tid, og vil ofte kansellere hverandre.

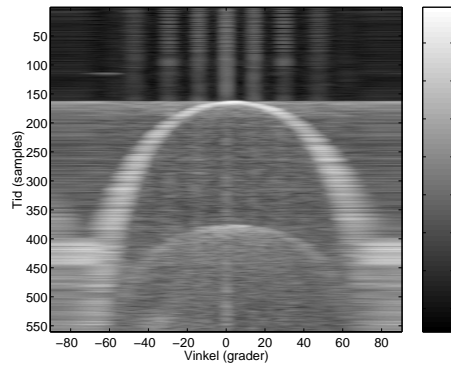
Fordelen til forsink-og-summér metoden er at den fungerer på alle typer signaler. Metoden fungerer for alle smalbandede og bredbandede signaler, uavhengig av om signalene kommer fra nær- eller fjernfeltet. Metoden er dessuten uavhengig av array-geometri, og kan derfor også benyttes på arrayer med ulik avstand mellom mottakerene.

Ulempen er at metoden er regnetung. For hver strålevinkel, må N tall summeres. Med M strålevinkler blir algoritmekompleksiteten da $\mathcal{O}(NM)$ (Weiss 1995). Den brukes derfor ikke i dagens sonarer.

2.2.2 Fourier transformasjonen

Fourier transformasjonen er en av de mest brukte transformasjonene innen signalbehandling. En egen versjon av transformasjonen finnes for diskrete data («Discrete Fourier Transform»). Denne kan implementeres svært raskt som «Fast Fourier Transform», oftest kalt «FFT». Da algoritmekompleksiteten til FFT er $\mathcal{O}(N \log N)$ (Oppenheim & Schafer 1989, Weiss 1995) eksekverer den svært raskt på dagens maskiner, så lenge N ikke er for stor. Fourier transformasjonen er derfor den metoden som benyttes til stråleforming i dagens sonarer.

Vanligvis benyttes Fourier transformasjonen til å transformere $x(t)$, en funksjon av tid t , til $X(\omega)$ som representerer signalet som en funksjon av frekvens ω . Ved stråleforming er det ønskelig å heller transformere fra mottakerposisjon i til strålevinkel θ . I stedet for strålevinkel, benyttes bølgenummer til å angi ønsket stråleretning. Bølgenummer er definert som $\vec{k} = \xi f/c$ hvor $\vec{\xi} = [\sin \theta, \cos \theta]$ er en enhetsvektor som peker i signalrets retning og c og f er henholdsvis bølgehastighet og signalfrekvens.



Figur 2.5: Stråleforming med FFT av rådata fra mottakerene. Bildet viser absoluttverdien (amplituden) til det komplekse signalet. Flat havbunn. Det lyse området formet som en «U» er det mottatte signalet («sonarfotavtrykket»), «U»-formen skyldes at geometrien er gitt i polarkoordinater. Det omliggende mørke området er bakgrunnsstøy. Logaritmisk transformasjon av pikselverdiene i amplitudebildet.

Den diskrete Fourier transformasjonen er for arrayer gitt ved (Johnson & Dudgeon 1993, avsn. 3.2)

$$X[\vec{k}, t] = \sum_{i=0}^N s_i[t] e^{j\vec{k}di},$$

hvor \vec{k} angir bølgenummeret, t er tidspunktet eller samplenummeret, d er avstanden mellom mottakerene og i er mottakernummer. Ved stråleforming reduseres uttrykket til

$$X[\theta, t] = \sum_{i=0}^N s_i[t] e^{jfd/cid \sin \theta} \quad (2.4)$$

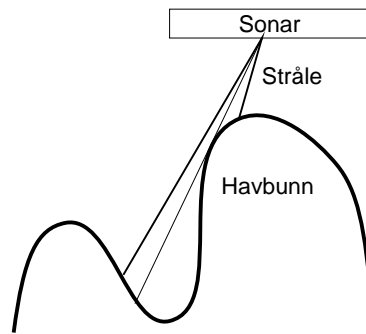
hvor en kjenner igjen leddet $\frac{fd \sin \theta}{c}$ fra ligning (2.2b).

For alle samtidige samples fra mottakerene beregnes $X[\theta_j, t]$ for vinklene $\theta_j = \theta_{\min} + j\Delta\theta, j = 1, \dots, M$. Figur 2.5 gir et eksempel på resultatet av denne transformasjonen. Her er $X[\theta, t]$ beregnet for $-90^\circ < \theta < 90^\circ$, og prosessen er gjentatt for $0 \leq t \leq 550$, hvor t angir samplenummer. Det hvite området som er formet som en «U» angir sonarens fotavtrykk. Fotavtrykkets form skyldes at bildets geometri er gitt i polare koordinater. Det er innenfor fotavtrykket bunnrefleksjonen er sterkest. Det mørke området rundt er hovedsaklig bakgrunnsstøy.

Fourier transformasjonen har enkelte klare begrensninger. Bredbåndssignaler som har frekvenskomponenter $|k| > \pi/d$ forårsaker aliasing. Metoden fungerer derfor bare for smalbåndssignaler fra fjernfeltet. Dessuten må signalet observeres med arrayer der avstanden er lik mellom alle mottakerene.

2.3 Bunndeteksjon

For å tegne batymetriske kart (altså dybdekart) er det nødvendig å finne havbunnen i det mottatte signalet fra sonaren. Tidsrommet innenfor hver



Figur 2.6: Multiplereksjon innenfor en stråle som følge av ujevn havbunn.

stråle som inneholder refleksjonen fra havbunnen kalles for strålens fotavtrykk (mer om fotavtrykk i kapittel 3). Samlingen av alle strålens fotavtrykk kalles for sonarens fotavtrykk.

Ideelt sett bør havbunnen identifiseres innenfor sonarfotavtrykket som en romlig kurve. Strålene som benyttes har imidlertid både romlig og temporal utstrekning. Bunndeteksjonsmetoder er derfor nødvendig for å redusere fotavtrykket til en tynn linje som representerer bunnen. Denne linjen gir informasjon om bunnforholdene rett under skipet, og kan benyttes videre til karttegning (avsnitt 2.4). Først gis en sammenfatning over noen av de viktigste metodene som benyttes til bunndeteksjon.

2.3.1 Amplitudebaserte metoder

Felles for de amplitudebaserte metodene er at de først estimerer fotavtrykket fra amplitudesignalet (avsnitt 3.3). Deretter estimeres havbunnens lokalisasjon ut fra fotavtrykkestimaten.

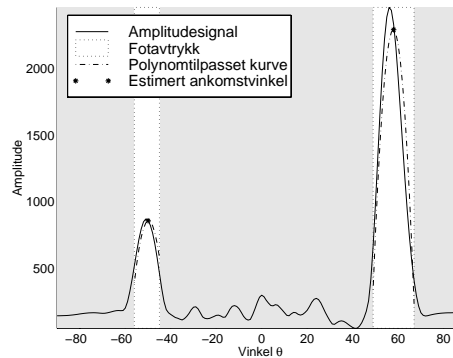
Det er vanlig å dele eksisterende bunndeteksjonsmetoder som benytter amplitudedata inn i to hovedgrupper (de Moustier 1993). Gruppe 1 forsøker å estimere ekkoets ankomsttid innenfor en stråle (Farr 1980, de Moustier & Alexandrou 1991). Gruppe 2 forsøker å isolere stråler som inneholder et ekko for et spesifikt tidspunkt (Satriano & Smith 1991). Begge gruppene arbeider med dataene utelukkende i en dimensjon om gangen. Sonardata er imidlertid todimensjonale (Tid \times Vinkel), noe som gir rom for forbedringer. Dette er den sentrale ideen i hovedfagsoppgaven.

Metoder for å finne ankomsttid

Metoder som finner ankomsttiden («TOA», Time-Of-Arrival) mottar først alle samples, og analyserer deretter en stråle om gangen. Ankomsttiden innenfor fotavtrykket estimeres separat for hver mottatt stråle. En vanlig brukt forenkling er å anta at hver stråle bare inneholder ett ekko fra havbunnen. Dette er riktig i de fleste tilfeller, men det er mulig å motta to eller flere bunnekkøer innen samme stråle (Figur 2.6).

Fotavtrykket for strålen defineres som et tidsintervall eller segment av strålen (avsnitt 3.3). En metode (Farr 1980) for å finne ankomsttidspunktet T_H innenfor fotavtrykket er å beregne tyngdepunktet til alle tidsindeksene innenfor fotavtrykket, altså

$$T_H = \sum_{t \in \mathcal{T}_F} I_A[\theta, t] \cdot t \quad (2.5)$$



Figur 2.7: Estimering av ankomstvinkel fra amplitudebilde ved tid t . Fotavtrykket estimeres, noe som gir to disjunkte regioner. Et polynom tilpasses til dataene innenfor hver region. Ankomstvinkelen estimeres så til å være polynomets toppunkt.

hvor \mathcal{T}_f er tidsindeksene til fotavtrykket, θ er strålevinkelen og $I_A[\theta, t]$ er amplitudebildet. En lignende metode (de Moustier & Alexandrou 1991) estimerer ankomsttidspunktet til midtpunktet av den kumulative summen av amplitudeverdiene, altså den T_H som oppfyller

$$\sum_{t=\min \mathcal{T}_f}^{T_H} I_A[\theta, t] = \frac{1}{2} \sum_{t \in \mathcal{T}_f} I_A[\theta, t] \quad (2.6)$$

Begge metoder kan gi et skjevt estimat. Dette kan skyldes et sterkt ekko i en nærliggende stråle. Da vil den nærliggende strålens sidelobeekko kunne adderes til refleksjonsekkoet i den aktuelle strålen.

I følge de Moustier (1993) er derfor TOA-metoder best egnet i tilfeller der ekkoet er sterkt og sentrert rundt strålen. Dette gjør at disse metodene har liten nytte utenfor det området som ligger rett under skipet.

Metoder for å finne ankomstvinkel

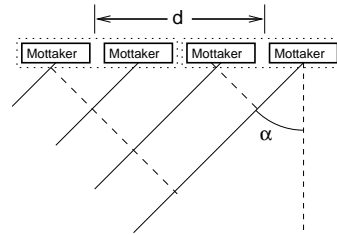
Den andre gruppen av amplitudebaserte metoder finner bunnekkets ankomstvinkeler («DOA», Direction-Of-Arrival) for hvert tidspunkt (Satriano & Smith 1991, de Moustier 1993). Når man ser på et tidspunkt om gangen, må man ta hensyn til at flere bunnekkere kan ankomme samtidig. Dette er vanlig — en flat bunn vil gi to samtidige refleksjoner (med strålevinkel $\pm \alpha$).

Etter at sonarfotavtrykket er estimert for tiden t (avsnitt 3.3), gjenstår som oftest flere disjunkte regioner i mottakerretningen som utgjør fotavtrykket. Ankomstvinkelen estimeres separat innenfor hver av disse regionene.

En enkel estimeringsmetode er å simpelthen velge det punktet innenfor hver region som har høyest amplitude (Satriano & Smith 1991), eller et vektet medium tilsvarende ligning (2.5) (de Moustier 1993).

En annen metode for å estimere ankomstvinklene er å benytte andre ordens polynomisk kurvetilpassing på dataene innenfor hver region, (Figur 2.7). Etter at en parabel har blitt tilpasset data innenfor hver region av fotavtrykket, estimeres ankomstvinkelene til parabelens toppunkt (Satriano & Smith 1991).

TOA-metodene og DOA-metodene utfyller hverandre. DOA-metodene fungerer best utenfor det området som er rett under båten; mens TOA-metodene er best egnet for området som er rett under båten.



Figur 2.8: Faseinterferometri ved hjelp av to subarrayer. Lydpulsen ankommer mottakerarrayet med en vinkel α . Prikkede linjer antyder gruppering av mottakere inn i to subarrayer. Stiplede linjer antyder bølgefront.

2.3.2 Fasebaserte metoder

Istedenfor å benytte amplitudeinformasjonen, kan faseinformasjonen brukes til å lokalisere havbunnen (Augustin, Edy, Savoye & Le Drezen 1994). En etablert metode (Hammerstad, Pøhner, Parthiot & Bennett 1991) og en nylig publisert metode (Yang & Taxt 1997) for å lokalisere havbunnen presenteres her.

Aperturereparasjonsmetoden

I noen av dagens sonarer, f.eks. SIMRAD EM3000 og EM12, benyttes aperturereparasjonsmetoden for å estimere vinkelen til det reflekterte signalet. Denne metoden deler mottakerarrayet inn i to like store subarrayer (Hammerstad et al. 1991, Yang & Taxt 1997).

La α angi vinkelen til den reflekterte pulsen, og d være avstanden mellom subarrayenes sentre (Figur 2.8). Anta at bølgefronten er plan, slik at faseforskjellen mellom de to subarrayene er gitt ved

$$\Delta\Phi = 2\pi f \frac{d}{c} \cos \alpha, \quad (2.7)$$

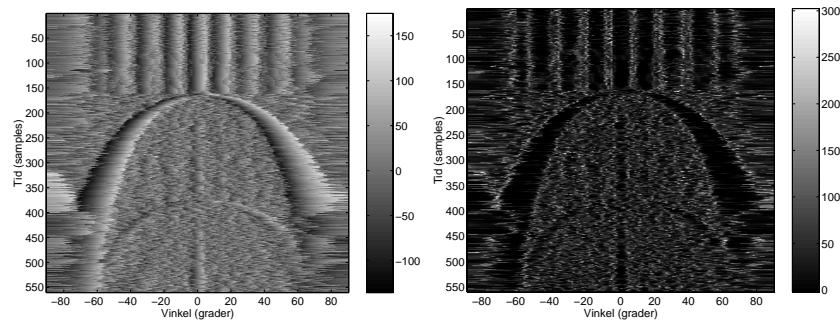
hvor f er signalfrekvensen og c er lydhastigheten i vann. Ut fra dette ser man at α finnes ved

$$\alpha = \arccos \left(\frac{c\Delta\Phi}{2\pi f d} \right) \quad (2.8)$$

Vanligvis gjøres stråleformingen innenfor hvert subarray med FFT (se avsnitt 2.2.2), og mottakerindeksen i justeres i ligning (2.4) slik at $\Delta\Phi$ er null når $\theta = \alpha$, hvor θ er strålevinkelen ved stråleforming. Denne justeringen gjøres ved å gi hver mottaker et indeksnummer, og la mottakeren beholde samme indeks uavhengig av hvilket subarray mottakeren er del av.

Innenfor fotavtrykket vil faseforskjellen $\Delta\Phi$ være en lineær funksjon av tiden t , med nullkrysning når ekkoet kommer i stråleretningen (Augustin et al. 1994). Utenfor fotavtrykket vil faseforskjellen opptre tilfeldig (Yang & Taxt 1997). Dette fenomenet kan benyttes ved lokalisering av havbunnen. Metoden forutsetter imidlertid at fotavtrykket er stort nok, slik at den lineære delen skiller seg godt nok ut fra den bakenforliggende støy.

En enkel måte å få et stort nok fotavtrykk, er å simpelthen redusere størrelsen på subarrayene. En slik reduksjon har imidlertid uønskede konsekvenser som at data kastes (Yang & Taxt 1997) og at signal-støy forholdet minsker (Johnson & Dudgeon 1993, avsn. 4.5). Støyen setter dermed en nedre grense for subarrayenes størrelse.



(a) Fasekurvebilde

(b) Feilfunksjon for fasekurvebildet

Figur 2.9: Resultat av interferometri med flere subarrayer. Fotavtrykket sees klart i bilde (a) som et kvasi-lineært område. I bilde (b) er fotavtrykket et område med svært liten intensitet. Fotavtrykket er formet som en «U». Geometrien er gitt i polarkoordinater.

Interferometri med flere subarrayer

En ny måte å benytte faseinformasjon på er presentert av Yang & Taxt (1997). Denne metoden splitter mottakerarrayet opp i mange overlappende subarrayer, mot tidligere kun to. På samme måte som i forrige avsnitt gjøres så stråleforming med FFT på de individuelle subarrayer.

Ved nærmere studie av ligning (2.7), ser man at faseforskjellen reelt sett er en lineær funksjon av avstanden d mellom mottakerene. Anta at signalet stråleformes i retning θ . La signalfasen til subarray i være gitt ved $\sigma_i[\theta, t]$, og la subarrayets posisjon være gitt ved x_i . For flere subarrayer kan ligning (2.7) generaliseres til

$$\sigma_i[\theta, t] = k[\theta, t]x_i + b[\theta, t] \quad (2.9)$$

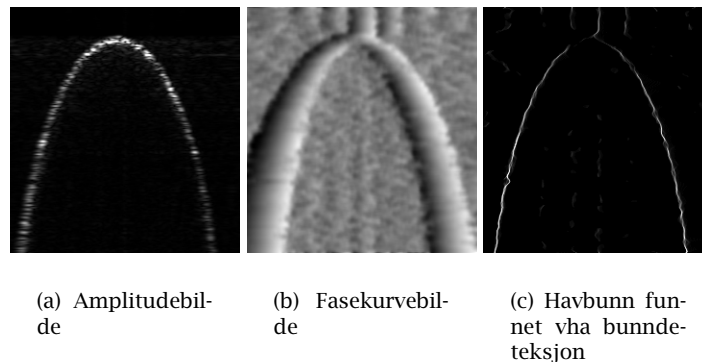
hvor $k[\theta, t]$ kalles for fasekurvekonstanten og $b[\theta, t]$ er en konstant.

Stråleforming kan nå gjøres på N subarrayer, noe som gir et datasett $s_i[\theta, t]$. Ved hjelp av minste kvadraters metode er det nå mulig å finne linjeparametrene $k[\theta, t]$ og $b[\theta, t]$ i ligning (2.9). Som ledd i denne linjetilpasningen finner man også estimeringsfeilen $e[\theta, t]$. Ved å beregne $k[\theta, t]$ og $e[\theta, t]$, er det mulig å danne bilder over disse variablene som funksjon av θ og t .

Et eksempel på et fasekurvebilde, altså $k[\theta, t]$, er vist i Figur 2.9(a). Som det går klart fram av figuren, er $k[\theta, t]$ i likhet med $\Delta\Phi$ en kvasi-lineær funksjon av tid innenfor fotavtrykket. $k[\theta, t]$ oppfører seg også som støy utenfor fotavtrykket (Yang & Taxt 1997).

Feilfunksjonen $e[\theta, t]$ har ingen klar parallell i den interferometriske metoden diskutert i forrige avsnitt. Et eksempel på et feilfunksjonsbilde er gitt i Figur 2.9(b). Fotavtrykket sees her klart som et område med svært liten intensitet.

Figur 2.10 viser bunndeteksjon med denne metoden. Amplitude- og fasekurvebildet sees i henholdsvis Figur 2.10(a) og 2.10(b). Deretter benyttes «Bottom Image Transform» (som baserer seg på både amplitude- og fasekurvebildet) til bunndeteksjon (Yang, Taxt & Albrechtsen 1997). Havbunnen sees som en tynn krum linje (Figur 2.10(c)).



Figur 2.10: Eksempel på deteksjon av tilnærmet flat bunn. (a) viser amplitudebildet; (b) viser fasekurvebildet. Bunndeteksjon gjøres vha «Bottom Image Transform», dette gir bilde (c). Bunnen sees som en tynn lys og krumme kurve mot en mørk bakgrunn. Bildene er tatt fra (Yang & Taxt 1997).

2.4 Karttegning

De foregående avsnittene har tatt for seg metodene som leder til de data som jeg bruker i denne oppgaven til å lokalisere sonarfotavtrykket. Resten av kapitlet viser to eksempler på videre databehandling etter at sonarfotavtrykket og havbunnen har blitt lokalisert.

2.4.1 Batymetriske kart

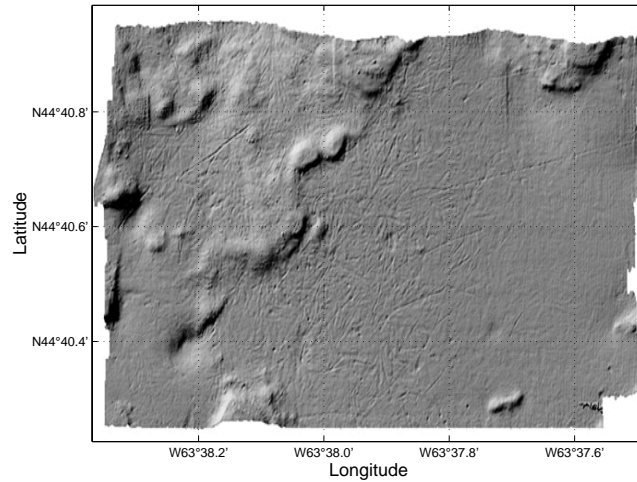
Produksjon av batymetriske kart eller dybdekart er den mest direkte bruk av havbunnsprofilene. Etter at bunndeteksjon har blitt gjort (avsnitt 2.3), finner man bunnen i et område på f.eks. ($1.5^\circ \times 128^\circ$) under skipet.

Ved å la skipet seile over et område (typisk frem og tilbake i parallelle linjer), får man en grov skisse av havbunnen. Denne skissen inneholder dybdeinformasjon for spredte punkter over havbunnen. Etter interpolasjon mellom disse punktene er det mulig å tegne kart over havbunnen (Augustin et al. 1994). Dybdekartet i Figur 2.11 viser en skyggelagt 3D-projeksjon av havbunnen ved Halifax havn i Canada. I tillegg til havbunnens topografi, er det også mulig å skimte også mindre detaljer, slik som skår i havbunnen som følge av ankring.

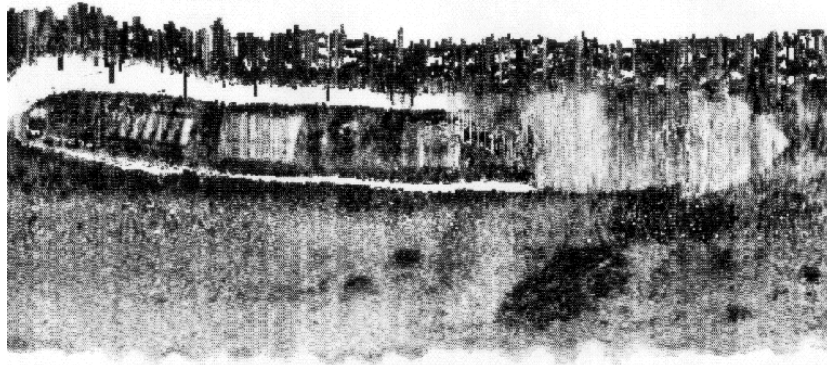
2.4.2 Refleksjonsbilder

Refleksjonsbildet gir informasjon om sedimentforholdene på havbunnen. Bildet viser hvor godt havbunnen reflekterer lydimpulsen. Automatiske bildebehandlingssystemer benytter refleksjonsbildet til å senere klassifisere havbunnen (Carmichael 1998, Augustin, Dugelay, Lurton & Voisset 1997). Ved hjelp av slike bilder har blant annet undersjøiske lavastrømmer blitt funnet (Somers 1993).

Refleksjonsbilder ble først lagd av sideseende sonarer, derav deres engelske navn: «Sidescan images». Det er mer regnekrevende å lage disse med multistrålesonarar, da mange flere stråler må prosesseres. Det gjennomsnittlig amplitudenivået til hver stråles fotavtrykk plottes (Mitchell 1996). Dette amplitudenivået må imidlertid korrigeres for avvik. Avviket skyldes



Figur 2.11: Batymetrisk kart over havbunnen i Halifax havn i Canada. Kartlagt med SIMRADs EM3000 sonar. Kartet er vist som et skyggelagt dybdebilde. Gitt av Kongsberg SIMRAD.



Figur 2.12: Refleksjonsbilde over et skipsvrak. Hvite områder angir sterk refleksjon, mørke områder svak eller ingen refleksjon. Aksene er kartesiske og angir geografisk posisjon. Gitt av Kongsberg SIMRAD.

at stråler som kommer fra rett under båten vil returnere vesentlig mer energi enn stråler som kommer fra siden av båten (Lingsch & Robinson 1992). Dessuten er det også en del tilfeldige amplitudevariasjoner over tid som må korrigeres (Fusillo & Satriano 1993).

Et eksempel på et refleksjonsbilde er vist i Figur 2.12. Omrisset av et skipsvrak vises tydelig mot en nokså uniform bakgrunn. Det er stor forskjell på hvor godt de enkelte deler av skipet reflekterer lydimpulsen. Metallskroget reflekterer lydimpulsen svært godt, og vises derfor som hvitt i refleksjonsbildet. Dekket er tydeligvis av et annet materiale, og reflekterer lydimpulsen dårligere — dette gjør at dekket vises med en mørkere farge.

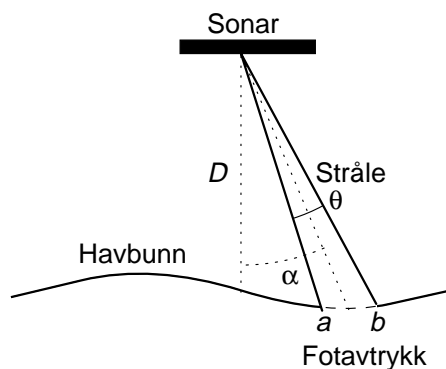
Kapittel 3

Sonarfotavtrykk

Innenfor en stråle kalles tidsrommet som inneholder refleksjonen fra havbunnen for strålens fotavtrykk. Samlingen av alle strålens fotavtrykk kalles for sonarens fotavtrykk. Figur 3.1 viser en stråle med vinkel α og utstrekning θ som gir et fotavtrykk på havbunnen fra punkt a til b . Da strålen har romlig utstrekning, er områdetets størrelse på havbunnen avhengig av α og θ . Strålen lager altså et «fotavtrykk» på havbunnen.

Ved behandling av sonarsignaler er det bare dataene innenfor sonarfotavtrykket som er interessante. Det omliggende området består bare av bakgrunnsstøy, og kan derfor ignoreres (Yang & Taxt 1997).

Resten av dette kapitlet omhandler sonarfotavtrykkets egenskaper, og fotavtrykkets bruk i batymetri og sonaravbildning. Til slutt gis en oversikt over eksisterende metoder for å finne sonarfotavtrykk.

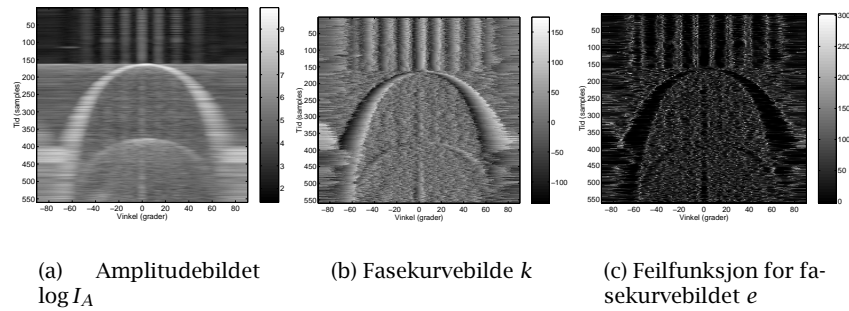


Figur 3.1: Illustrasjon av en strålens fotavtrykk. En stråle med vinkel α og utstrekning θ formes ut fra signalet som mottas fra havbunnen. Fotavtrykket til strålen (området mellom a og b) er vist som en stiplet linje langs havbunnen.

3.1 Sonarfotavtrykkets egenskaper

3.1.1 Signalene innenfor fotavtrykket

Signalene innenfor fotavtrykket har to vesentlige egenskaper: Stor amplitude og (ved flat havbunn) lineær faseforskjell mellom mottakerene (Yang & Taxt 1997). Så lenge man ikke mottar noe ekko fra havbunnen, har signalene som mottas liten amplitude. Som man ser av amplitudebildet i Fi-



Figur 3.2: Resultat av forskjellige signalbehandlingsmetoder for sonarsignaler fra flat bunn. Sonarfotavtrykket sees som et «U»-formet område i alle tre bilder. Formen skyldes bruk av polare koordinater.

Figur 3.2(a) medfører dette at fotavtrykket fremstår som et lyst område (dvs. stor amplitude) mot en mørk bakgrunn.

Innenfor fotavtrykket er faseforskjellen mellom to mottakere en lineær funksjon av avstanden mellom mottakerene (avsnitt 2.3.2). Faseforskjellen er også en omtrent lineær funksjon av tid. Dette er ikke tilfelle utenfor fotavtrykket.

Metoden for beregning av fasekurvebilde (Yang & Taxt 1997) gir også fotavtrykket særtrekk. For fasekurvebildet $k[\theta, t]$ (avsnitt 2.3.2) vil $k[\theta, t]$ være en kvasi-lineær funksjon av både θ og t innenfor fotavtrykket. Dette sees tydelig av Figur 3.2(b). På grunn av den lineære faseforskjellen innenfor feilbildet vil feilfunksjonsbildet $e[\theta, t]$ indikere svært liten estimeringsfeil innenfor fotavtrykket. Den svarte regionen i Figur 3.2(c) viser dette tydelig — fotavtrykket fremstår som et lavintensitetsområde mot en lysere bakgrunn.

3.1.2 Sekundært ekko — uønsket reverberasjon

Et vanlig problem ved identifikasjon av fotavtrykk oppstår som følge av sekundære ekkoer eller reverberasjoner. Figur 3.3(a) viser denne uønskede effekten. Lydpulsen sendes først fra sonaren. Noe av lydimpulsen treffer havoverflaten og reflekteres mot havbunnen på ny. Deretter returnerer lydimpulsen til sonaren.

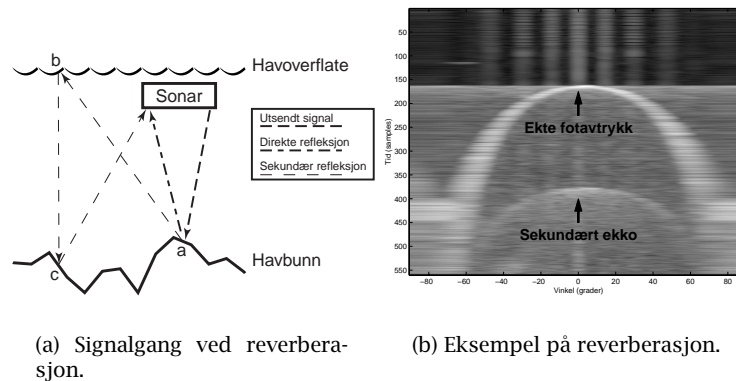
Disse reverberasjonene skaper områder i strålene med tilnærmet de samme egenskapene som det reelle fotavtrykket: Høyere amplitude og lineær faseforskjell mellom mottakerene. Amplituden innenfor det sekundære ekkoet er imidlertid som oftest lavere enn innenfor det reelle fotavtrykket.

Figur 3.3(b) viser et eksempel på hvordan slike reverberasjoner ser ut. Pilene øverst og nederst peker på henholdsvis reelt fotavtrykk og det sekundære ekkoet. Det sekundære ekkoet sees som en svak lokal økning i amplituden.

3.2 Bruk av sonarfotavtrykk

Refleksjonen fra havbunnen vil alltid finnes innenfor fotavtrykket (Yang & Taxt 1997). Et godt lokalisert fotavtrykk gir altså klare yttergrenser for hvor

3.3. ETABLERTE METODER FOR LOKALISERING AV SONARFOTAVTRYKK19



Figur 3.3: Sekundært ekko — reverberasjon. (a) Signalet sendes ut av sonaren, treffer punkt a på havbunnen, treffer havoverflaten i b , havbunnen i c og returnerer til sonaren. Dette skaper reverberasjon som vanskeliggjør presis identifikasjon av sonarfotavtrykk. (b) Pilene peker på det ekte fotavtrykket og den første av de uønskede reverberasjonene.

bunnen kan befinne seg. Dette begrenser avvik i bunndeteksjonsprosessen, og øker dermed presisjonen ved lokalisering av havbunnen.

En annen viktig gevinst ved et godt estimert fotavtrykk er bedre refleksjonsbilder (avsnitt 2.4.2). Slike bilder genereres ut fra amplitudeverdien til dataene innenfor fotavtrykket. Data som feilaktig hentes fra området utenfor fotavtrykket forringer estimatet av bunnrefleksjonsstyrken. Et presist lokalisert fotavtrykk vil redusere bruk av slike feilaktige data.

3.3 Etablerte metoder for lokalisering av sonarfotavtrykk

De to eksisterende metodene for lokalisering av sonarfotavtrykk baserer seg på at amplituden innenfor sonarfotavtrykket er høyere enn utenfor fotavtrykket. Metodene har ikke så god presisjon. Dette kan skyldes at metodene arbeider enten ikke-kontekstueelt eller bare er kontekstuelle i en dimensjon om gangen. Dessuten gjør ingen av metodene noe forsøk på å fjerne støy.

3.3.1 Ikke-kontekstuell global terskling av amplitudebildet

de Moustier (1993) nevner ikke-kontekstuell global terskling som en metode for lokalisering av fotavtrykk. Metoden er svært enkel og ikke regnetung.

Fotavtrykket estimeres ved å la $I_A[\theta, t]$ være amplitudebildet for tid t og strålevinkel θ . Fotavtrykket lokaliseres så til å være de $[\theta, t]$ hvor $I_A[\theta, t] > T$, hvor T er en forhåndsvalgt global terskel. Terskelen T velges manuelt på forhånd slik at terskelverdien er større enn amplituden til støy og sideløber. Det eksisterer også metoder for å automatisk finne T ut fra egenskaper til bildet (f.eks. Otsu 1979, Kapur, Sahoo & Wong 1985, Abulateb 1989, Kittler & Illingworth 1986).

Det viste seg at resultatet ved global terskling var for dårlig også når terskelverdien T ble valgt manuelt, særlig for data fra sonarer som ser rett

mot bunnen. Dette skyldes at metoden ikke kompenserer for lokale amplitudevariasjoner. I områder med variert bunnsammensetning, kan områder som reflekterer signalet dårlig bli ignorert. Videre gir metoden et lite sammenhengende og ujevnt fotavtrykk. Jeg valgte derfor å ikke studere globale metoder videre.

3.3.2 Enkel endimensjonal terskling

En noe mer adaptiv metode nevnes av Satriano & Smith (1991) og Lingsch & Robinson (1992). Denne metoden benytter en-dimensjonal terskling for å finne fotavtrykket. Metoden finner automatisk en ny terskel $T[t]$ for hvert tidpunkt t i amplitudebildet $I_A[\theta, t]$. Terskelen $T[t]$ for amplitudebildet $I_A[t, \theta]$ beregnes ved

$$T[t] = \frac{C}{N} \sum_{\theta=\theta_{\min}}^{\theta_{\max}} I_A[t, \theta] \quad (3.1)$$

hvor N er antall stråler i θ -retningen og C er en skaleringskonstant. Sonarfotavtrykket estimeres så til de $[t, \theta]$ hvor $I_A[t, \theta] \geq T[t]$. Selv om resultatet er noe bedre enn for global terskling, har også denne metoden de samme problemene med sekundære ekkoer.

Jeg valgte å benytte denne metoden som referansemetode.

Valg av parameterverdier

Denne metoden trenger en parameter for skalering av gjennomsnittsverdien. Satriano & Smith (1991) benytter $C = 2.0$. Jeg også forsøkt andre valg av parameterverdier, for å se om dette forbedret resultatet.

Skaleringskonstant C 2, 4, ..., 70

Tabell 3.1: Parameterverdier for endimensjonal terskling.

Kapittel 4

Nye metoder for lokalisering av sonarfotavtrykk

Dagens sonarer benytter ganske enkle endimensjonale klassifikasjonsmetoder («tersklingsmetoder») for å lokalisere sonarfotavtrykk (de Moustier 1993, Lingsch & Robinson 1992, Satriano & Smith 1991). Samtidig finnes det et mangfold av todimensjonale globale og lokalt adaptive tersklingsmetoder for lokalisering av bilderegioner med spesifikke egenskaper. Bildeterskling er i det enkleste tilfellet å klassifisere eller terskle hvert piksel i et bilde som enten forgrunn eller bakgrunn. Her tilsvarende forgrunnen (eller «objektet») fotavtrykket, og bakgrunnen bakgrunnstøyen.

En rekke globale og lokalt adaptive tersklingsmetoder er publisert (f.eks Pratt 1991, Zhang 1997, Pal & Pal 1993, Trier & Taxt 1995*b*). Mangfoldet av slike metoder gjør det nødvendig å begrense metodeutvalget. I oppgaven har jeg valgt å evaluere de fem metodene som kom best ut i en omfattende sammenligning av tersklingsmetoder (Trier & Taxt 1995*b*).

Resten av kapitlet er delt opp i to deler: Tersklingsmetoder og postprosesseringsmetoder. Avsnittene er organisert som følger:

Lokalt adaptiv terskling (avsnitt 4.1) presenterer de fem utvalgte tersklingsmetodene som gir den initiale klassifikasjon av bildet i forgrunns- og bakgrunnsobjekter.

Postprosesseringsmetoder (avsnitt 4.2–4.3) omhandler metoder som fjerner regioner som feilaktig har blitt klassifisert som del av fotavtrykket.

I metodebeskrivelsene har jeg valgt å opprettholde vanlig bildebehandlingsnotasjon. Innenfor bildebehandling er det vanlig å benytte (x, y) for å identifisere posisjonen til pikselene. Brukt på sonardata representerer x θ -aksen og y tidsaksen. Både x og y er her diskrete variabler og lineære funksjoner av θ og t .

4.1 Lokal adaptiv terskling

I motsetning til global terskling (avsnitt 3.3.1), prøver lokalt adaptive metoder å tilpasse tersklingsverdien individuelt til hver piksel. Dette gjøres ved å lage en funksjon $T(x, y)$ som angir tersklingsverdien til pikselen (x, y) i innbildet $I(x, y)$. Funksjonen $T(x, y)$ kalles ofte for bildets tersklingsflate. Etter at flaten $T(x, y)$ er blitt beregnet, gjøres selve tersklingen ved

$$B(x, y) = \begin{cases} 1 & T(x, y) \leq I(x, y) \\ 0 & T(x, y) > I(x, y) \end{cases} \quad (4.1)$$

hvor $B(x, y)$ er det segmenterte bildet.

Valg av vindusstørrelse

Lokalt adaptive metoder benytter ofte små lokale vinduer som bearbejdes mer eller mindre uavhengig. Tersklingsresultatet er ofte avhengig av valg av vindusstørrelse. Med mitt valg av stråleformingsparametre (avsnitt 5.1.1), hadde fotavtrykket en bredde på ca 3 pikseler. Ut fra dette valgte jeg vindusstørrelsen for slike lokalt adaptive metoder ut fra følgende resonnement:

- En vanlig tommelfingerregel er at vindusbredden bør være omtrent det dobbelte av objektets (fotavtrykkets) utstrekning. En slik vindusstørrelse sikrer at det i nærheten av fotavtrykket vil være omtrent like mye forgrunn som bakgrunn innenfor vinduet. Dette hjelper metoder som forsøker å estimere fordelingsparametre ut fra det lokale histogram.

Innledende forsøk antydte imidlertid at større vinduer i noen tilfeller ga bedre resultat. Jeg har derfor forsøkt med en rekke vindusbredder, fra en bredde litt under det dobbelte av fotavtrykkets bredde til en vindusbredde lik bildets bredde.

- Den optimale vinduslengden er noe lettere å argumentere rundt. Den bør være nokså liten. Dette skyldes at signalstyrken til ekkoet varierer vesentlig over tid. Det er ønskelig at filteret raskt tilpasser seg endringer i signalstyrken. En viss høyde vil samtidig annen side medføre at det blir noe glatting ved valg av tersklingsverdi. Der det var mulig, har jeg derfor forsøkt med vindushøyder mellom 3–9 pikseler.

4.1.1 Bernsens metode

Bernsen's (1986) metode benytter minimal- og maksimalverdiene innenfor et lokalt vindu for danne tersklingsflaten. Innenfor et område R rundt pikselen (x, y) i innbildet $I(x, y)$ beregnes tersklingsflaten $T(x, y)$ ved

$$T(x, y) = 0.5 \min_{(i,j) \in R} I(i, j) + 0.5 \max_{(i,j) \in R} I(i, j) \quad (4.2)$$

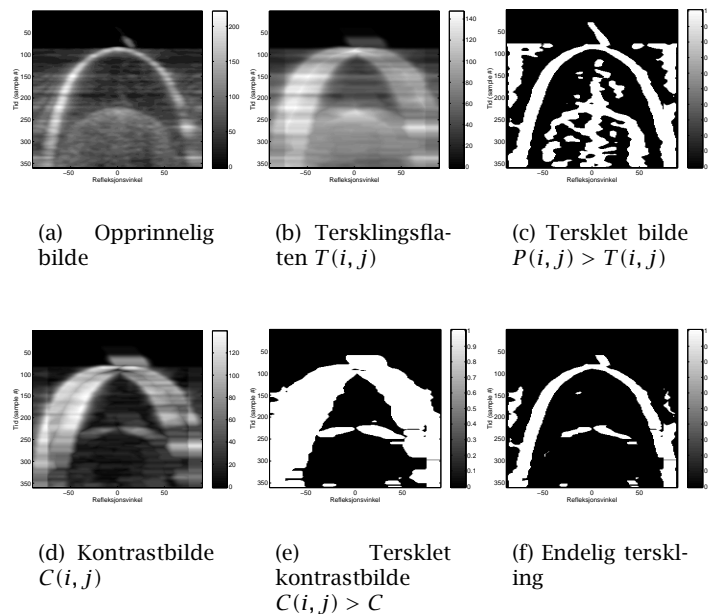
Metoden får imidlertid problemer hvis noen regioner er homogene og bare inneholder bakgrunn, noe som ofte er tilfelle i vårt bilde. Bernsen (1986) foreslo derfor å også se på hvor stor kontrast det er innenfor vinduet. Den lokale kontrasten $C(x, y)$ estimeres ved

$$C(x, y) = \max_{(i,j) \in R} I(i, j) - \min_{(i,j) \in R} I(i, j) \quad (4.3)$$

Hvis $C(x, y) > C$, hvor C er en valgt global terskel, er tersklingsflaten $T(x, y)$ sannsynligvis velegnet til å terskle denne regionen. Hvis ikke, må en annen metode benyttes. I dette tilfellet velger jeg å klassifisere slike lavkontrastområder som bakgrunn. Tersklingen skjer altså ved følgende ligning

$$B(x, y) = \begin{cases} 1 & C(x, y) > C \wedge I(x, y) > T(x, y) \\ 0 & \text{ellers} \end{cases} \quad (4.4)$$

Behovet for kontrastfiltrering sees tydelig fra Figur 4.1, hvor metoden har blitt benyttet på det glattede testbildet i Figur 4.1(a). Etter at tersklingsflaten $T(x, y)$ (Figur 4.1(b)) har blitt beregnet, terskles testbildet ved hjelp av ligning (4.1), noe som gir bildet vist i Figur 4.1(c).



Figur 4.1: Stegvis illustrasjon av Bernsen's (1986) metode.

Det er ennå store mengder med falske positive piksler i bildet. Denne feilklassifiseringen oppstår hovedsaklig i de områdene i bildet hvor det kun er støy, og ikke noe signal.

Kontrastbildet (Figur 4.1(d)) terskles derfor, noe som resulterer i Figur 4.1(e). Dette er et svært grovt estimat på fotavtrykket. Dette grove estimatet kan nå benyttes som en maske som fjerner store deler av områdene som kun inneholder støy. Den endelige tersklingen, vist i Figur 4.1(f) er vesentlig forbedret.

Valg av parameterverdier

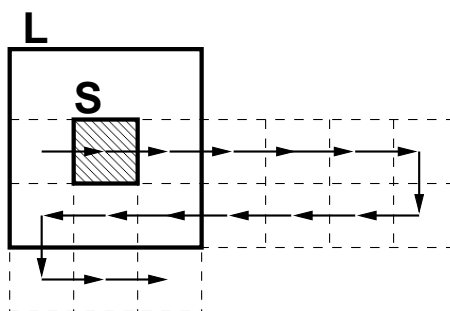
Tabell 4.1 gir parameterverdier for Bernsens metode. Vindusstørrelsen ble valgt ut fra betraktningene i avsnitt 4.1. Implementasjonen tillot bruk av rektangulære vinduer, vindusbredde og -lengde varierte derfor uavhengig. Tersklingsverdien for kontrastbildet ble valgt ut innledende eksperimenter der jeg beregnet kontrasten for noen utvalgte bilder.

Horisontal vindusstørrelse	3, 5, 7, 15, 30
Vertikal vindusstørrelse	3, 5, 7, 9
Kontrastterskel C	10, 25, 40, 55

Tabell 4.1: Parameterverdier for Bernsens metode.

4.1.2 Eikvil, Taxt & Moens metode

Denne metoden (Eikvil et al. 1991) er en forbedring av Taxt, Flynn & Jain's (1989) metode. I motsetning til den opprinnelige metoden, som bare benytter ett vindu, benytter denne metoden to vinduer som glir over bildet. Regelsettet som benyttes for å generere tersklingsflaten er også forbedret.



Figur 4.2: Prinsipp for metoden til Eikvil et al. (1991). Det store vinduet L glir over bildet med steg tilsvarende størrelsen på det lille vinduet S . S holdes sentrert i L i løpet av prosessen. Pikselene i S terskles på basis av egenskapene til pikselene i L .

Tersklingen av det lille vinduet S gjøres på grunnlag av egenskapene til et større vindu L som ligger konsentrisk rundt S . Disse to vinduene glir simultant over vinduet med steg lik størrelsen på S , slik som vist i Figur 4.2.

L terskles ved hjelp av en eksisterende global tersklingsmetode (f.eks. Otsu 1979, Kittler & Illingworth 1986) inn i forgrunn og bakgrunn, som finner en terskel T . Gjennomsnittlig pikselverdi for forgrunnen μ_0 og bakgrunnen μ_1 beregnes.

Underveis benyttes μ_1 og μ_2 også til å oppdatere den kumulerte gjennomsnittlige intensiteten til forgrunnen μ_1^i og bakgrunnen μ_2^i . Dette skjer ved hjelp av oppdateringsligningene

$$\mu_0^{i+1} = w \cdot \mu_0 + (1 - w)\mu_0^i \quad (4.5a)$$

$$\mu_1^{i+1} = w \cdot \mu_1 + (1 - w)\mu_1^i \quad (4.5b)$$

hvor w er en valgt konstant og i er antall vinduer S som har blitt behandlet. μ_0^0 og μ_1^0 beregnes ved å gjøre en global terskling av bildet som helhet, og beregne gjennomsnittlig intensitet til forgrunn og bakgrunn.

I likhet med Bernsens metode (avsnitt 4.1.1) testes det på hvorvidt L er homogen — altså om L inneholder bare bakgrunn eller forgrunn. Dette gjøres ved å beregne $\Delta\mu = |\mu_1 - \mu_2|$. Hvis $\Delta\mu > l$, hvor l er en global kontrastterskel, anses scenen som inhomogen. T benyttes da til å terskle S .

Hvis $\Delta\mu < l$ anses vinduet som homogent. Metoden forsøker da å klassifisere hele vinduet som enten forgrunn eller bakgrunn. μ_1^i og μ_2^i benyttes da til å klassifisere pikselene i S , og tersklingsresultatet beregnes ved

$$B(x, y) = \begin{cases} 1 & |T - \mu_0^i| \leq |T - \mu_1^i| \\ 0 & |T - \mu_0^i| > |T - \mu_1^i| \end{cases} \quad (4.6)$$

Alle pikselene i S tilordnes altså den klassen som har den nærmest liggende kumulerte gjennomsnittsverdien.

Valg av parameterverdier

Metoden benytter to sett med parametre: Et sett med parametre for vindusstørrelser, og et sett med parametre for justering av tersklingsflaten.

For å velge vindusstørrelse støttet jeg meg til diskusjonen i avsnitt 4.1. Da S vinduet må ligge innenfor L , valgte jeg å redusere størrelsen på det lille vinduet noe. Begrensninger i implementasjonen hindret bruk av annet enn kvadratiske vinduer.

De resterende parameterverdiene ble valgt ut fra et ønske om jevn fordeling av mulige parametre. Jeg valgte å benytte et større spenn ved valg av kontrastgrense l enn i Bernsens metode, da kontrastestimeringen ikke er like forutsigbar. Tabell 4.2 oppsummerer de utprøvde parameterverdiene.

Regionstørrelse L	$(5 \times 5), (7 \times 7), (15 \times 15), (30 \times 30)$
Vindusstørrelse S	$(3 \times 3), (5 \times 5), (9 \times 9)$
Kontrastgrense l	10, 30, 50, 70, 100
Vekt w	0.1, 0.4, 0.6, 0.9

Tabell 4.2: Parameterverdier for Eikvil, Taxt og Moens metode.

4.1.3 Niblack's metode

I Trier & Taxt's (1995b) sammenligning endte Niblack's (1986) metode (med postprosesseringssteget beskrevet i avsnitt 4.2.2) opp som den beste metoden for tersklingskart. Det var derfor interessant å se om metoden var like velegnet for disse dataene.

Metoden beregner tersklingsflaten $T(x, y)$ til pikselen $I(x, y)$ i innbildet I ved

$$T(x, y) = a \cdot \sigma(x, y) + \mu(x, y) \quad (4.7)$$

hvor $\sigma(x, y)$ og $\mu(x, y)$ er henholdsvis standardavviket og gjennomsnittsverdien i et område rundt pikselen. Denne tersklingsflaten benyttes så for å terskle bildet.

Valg av parameterverdier

Tabell 4.3 gir parameterverdier for Niblack's metode. Min implementasjon av Niblack's metode forutsatte kvadratiske vinduer. Dette begrenset valg av vindusstørrelser noe (jæmfør avsnitt 4.1 for valg av vindusstørrelser).

Trier (1995) anbefalte å gi standardavviket en vekt $a \approx 0.2$, som var egnet for hvite forgrunnsobjekter. Innledende eksperimenter antydte imidlertid at høyere verdier også ga gode resultater. Jeg valgte derfor å prøve noe høyere verdier også.

Vindusstørrelse	$(3 \times 3), (5 \times 5), (7 \times 7), (15 \times 15), (30 \times 30)$
Vekt a	0.1, 0.15, 0.2, 0.25, 0.3, 0.5, 0.6, 0.7, 0.8

Tabell 4.3: Parameterverdier for Niblack's metode.

4.1.4 Yanowitz & Brucksteins metode

Yanowitz & Bruckstein's (1989) metode lager tersklingsflaten basert på gradientinformasjon i bildet. Et gradientmagnitusbilde $G(x, y)$ beregnes ut fra innbildet $I(x, y)$ ved hjelp av Sobels kantdetektor (Pratt 1991). $G(x, y)$ terskles med en terskel T_G , noe som gir et bilde der kantene er markert. Disse kantene er oftest mer enn en piksel tynne. Ved hjelp av morfologiske operasjoner tynnes disse kantene. Dette gir et nytt bilde $M(x, y)$ som markerer bildet $I(x, y)$ sine kanter ved hjelp av en-piksel tynne linjer.

Ved hjelp av en iterativ interpoleringsmetode kan tersklingsflaten nå regnes ut slik at den går gjennom samtlige kantpikseler i bildet. Dette gjør

res ved å beregne tersklingsflaten $T(x, y)$ ved

$$T_{n+1}(x, y) = T_n(x, y) + \frac{\beta \cdot R_n(x, y)}{4} \quad (4.8a)$$

$$\begin{aligned} R_n(x, y) = & T_n(x - 1, y) + T_n(x + 1, y) \\ & + T_n(x, y + 1) + T_n(x, y - 1) \\ & - 4T_n(x, y) \end{aligned} \quad (4.8b)$$

hvor $R_n(x, y)$ er en skalert Laplace-filtrering av $T_n(x, y)$ (Pratt 1991, avsn. 16.3.1), og n angir iterasjonsnummer. De pikselene (x, y) som er kantpikselene fastholdes gjennom iterasjonene. Denne iterasjonen fortsetter inntil $\max_{(x,y)} R_n(x, y) < R_{\max}$ eller $n > N$, hvor R_{\max} og N er parametre til metoden. Yanowitz & Bruckstein (1989) angir $R_{\max} = 0.1$, dette har også blitt benyttet i min uttesting. Prosessen i ligning (4.8a) og (4.8b) kalles relaksjon, og er en lineær lavpassfiltrering av bildet.

Yanowitz & Bruckstein (1989) beskriver også et siste postprosesseringssteg. I likhet med Trier & Taxt (1995b) har jeg skilt ut dette steget til et separat steg. Dette steget er beskrevet i avsnitt 4.2.2.

Valg av parameterverdier

Verdier for tre parametere må velges: For interpolasjonskonstanten β , for gradientterskelen T_G og for maksimalt antall iterasjoner i . Yanowitz & Bruckstein (1989) anbefaler $1.0 \leq \beta \leq 2.0$, mens Trier (1995) anbefaler $\beta \leq 1.0$. Jeg valgte derfor å la β spre seg noenlunde jevnt mellom 0.5 og 1.5.

Jeg valgte antall iterasjoner ut fra et ønske om å kunne vekte kjøretid mot resultat. På den ene siden var det interessant å se hvor godt resultat man kunne få ved få iterasjoner; samtidig var det interessant å vite hva det beste resultatet ble. Jeg valgte derfor et nokså vidt spenn (5–150) med iterasjoner.

For å velge gradientterskelen, benyttet jeg en Sobel-operator på noen amplitudebilder, og valgte manuelt terskelen for disse bildene for å finne en terskel slik at mest mulig av kanten til fotavtrykket ble med, mens minst mulig annet ble inkludert (Tabell 4.4).

β	0.5, 0.9, 0.95, 1.0, 1.05, 1.1, 1.25, 1.5
Max iterasjoner N	5, 25, 50, 150
Gradientterskel T_G	5, 10, 20, 30, 40, 50

Tabell 4.4: Parameterverdier for Yanowitz & Brucksteins metode.

4.1.5 Forbedret White & Rohrer

En forbedring av White & Rohrer's (1983) integrerte funksjons algoritme har blitt utviklet av Trier & Taxt (1995a). I likhet med den opprinnelige metoden danner ikke denne metoden heller noen tersklingsflate.

Som i den opprinnelige metoden dannes en tre-nivå overflate ved å beregne en gradientlignende operator $A(x, y)$ («aktivitetsnivået») for innbildet $I(x, y)$. $A(x, y)$ terskles så med en global terskling T_A . Pikselene (x, y) i $A(x, y)$ gis merkelapp «0» hvis pikselverdien er lavere enn T_A .

For å klassifisere de resterende pikselene, benyttes Laplace-operatoren gitt ved

$$\nabla^2 I(x, y) = \frac{\delta^2 I(x, y)}{\delta x^2} + \frac{\delta^2 I(x, y)}{\delta y^2} \quad (4.9)$$

eller i det diskrete tilfellet (Pratt 1991, avsn. 16.3.1) ved for eksempel

$$\begin{aligned} \nabla^2 I(x, y) = I(x, y) - \frac{1}{4}(I(x-1, y) - I(x+1, y)) \\ - \frac{1}{4}(I(x, y-1) - I(x, y+1)) \end{aligned} \quad (4.10)$$

Hvis $\nabla^2 I(x, y) > 0$ merkes pikselen som «+», ellers som «-». «+»-piksel dominerer langs innsiden av objektkantene, mens «-»-piksel dominerer langs utsiden av objektkantene.

For å redusere effekten støy har på tersklingsresultatet, omgjøres små «+»-regioner med færre enn S piksel om til «0»-regioner. S er her en metodeparameter.

Dette trenivå bildet benyttes så for videre klassifikasjon. Metoden markerer først alle «+»-piksel som objektpiksel og alle «-»-piksel som bakgrunn. «0»-piksel klassifiseres så ved å se på deres åtteraboer. Hvis en «0»-piksel har overvekt av «+»-piksel som naboer klassifiseres den som objektpiksel. Videre markeres den som en «+»-piksel. Hvis «-»-piksel er i overvekt skjer det motsatte; pikselen klassifiseres som bakgrunn og markeres som en «-»-piksel.

Valg av parameterverdier

Trier & Taxt (1995a) spesifiserer også et pre- og postprosesseringsledd. Jeg velger her å ignorere disse leddene. Glatting gjøres allerede av dataene, og postprosesseringsleddet metoden spesifiserer omhandles senere i avsnitt 4.2.2.

Dette levner kun to parametre — terskelverdien for aktivitetsnivået og terskelverdien for minimal objektstørrelse. Aktivitetsgrensen valgte jeg ved samme metodikk som for Yanowitz & Bruckstein's (1989) Rohrsers metode. Et kort studie av aktivitetsnivåer til noen amplitudebilder grunngir verdi-valget.

En annen mulighet jeg utforsket var å automatisk estimere aktivitetsgrensen for hvert bilde ved hjelp av Kittler & Illingworth's (1986) metode (markert med E i Tabell 4.5).

Som ledd av metoden gjøres små «+»-regioner om til «0»-regioner, altså regioner med færre enn S piksel. Minimal størrelse på slike «+»-regioner ble valgt ut fra et kort forhåndsstudie av noen utvalgte bilder.

Aktivitetsgrense T_A	$E, 40, 70, 100, 130$
Minimal regionstørrelse S	$15, 30, 50$

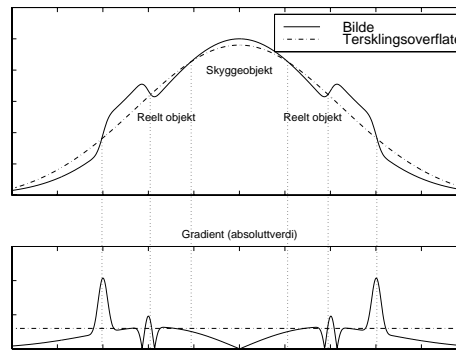
Tabell 4.5: Parameterverdier for forbedret White & metode. E angir at aktivitetsgrensen ble estimert med Kittler & Illingworth's (1986) metode.

4.2 Fjerning av falske objekter

Mesteparten av de tersklede bildene inneholder regioner som ikke tilhører fotavtrykket. Jeg prøvde derfor ut noen enkle metoder som forsøker å fjerne slike feilklassifiserte objekter.

4.2.1 Fjerning av initiell støy

Det tar noe tid fra vi sender ut lydimpulsen til den første refleksjonen ankommer mottakerene. Data i dette tidsrommet inneholder dermed bare bakgrunnsstøy.



Figur 4.3: Fjerning av «spøkelsesobjekter». Den genererte tersklingsflaten lager et falskt objekt midt i bildet. Dette objektet kan fjernes ved å benytte gradientinformasjon langs kanten på objektene (vertikale linjer). Ved å terskle gjennomsnittlig gradient (horisontal linje nederste graf) kan falske positive fjernes.

Dette støyfeltet forvirrer tersklingsmetodene, og skaper en del objekter feilklassifisert som fotavtrykk. En svært rask metode er utviklet av Yang (1998a) for å raskt fjerne disse falske fotavtrykksdelene.

Metoden estimerer tidspunktet for når første ekko blir mottatt. Estimaten gjøres ved å se på den maksimale amplitudeverdien innenfor hver tidslinje. For et amplitudebilde $A(x, y)$ estimeres tidspunktet y_e for første ekko ved

$$y_e = \min\{y \mid \max_x A(x, y) > T\} \quad (4.11)$$

hvor T er en valgt terskel. Alle pikselene i det tersklede bildet $B(x, y)$ hvor $0 \leq y \leq y_e$ markeres så som bakgrunns pikseler.

Imidlertid inneholder de 10–15 første tidslinjene i hvert sonarbilde ofte en del sterk støy. Denne sterke innledende støyen forvirrer metoden. Det første ekkoet ankommer ikke så tidlig. Metoden overser derfor disse tidslinjene når den estimerer y_e .

Valg av parameterverdier

Jeg valgte å studere noen bilder fra hver serie, og velge terskelen T ut fra dette. Etter noen forsøk fant jeg at resultatet stort sett var uavhengig av valgt terskel. For å spare tid valgte jeg å kun benytte en terskel, $T = 5$.

4.2.2 Fjerning av objekter i homogene områder

Ved terskling av bilder oppstår det ofte en viss mengde regioner feilklassifisert som deler av fotavtrykket. En mulig grunn til dette er en feilaktig tersklingsflate, som illustrert i Figur 4.3. Her er objektet på toppen feilaktig tatt med som forgrunnsobjekt grunnet en feilaktig tersklingsflate.

Som man ser, er imidlertid gradienten ved kanten til de faktiske objektene langt større enn ved kanten til «skygge»-objektet. Yanowitz & Bruckstein (1989) utnyttet dette til å fjerne slike objekter. I likhet med Trier & Taxt (1995b) valgte jeg å benytte denne delen av Yanowitz & Bruckstein's (1989) metode som en generell postprosesseringsmetode.

Metoden fungerer ved å for hvert forgrunnsobjekt beregner gjennomsnittlig gradient \bar{G}_i langs kanten til objekt nummer i . En terskel T_G benyttes så for å fjerne objekter der $\bar{G}_i < T_G$. Dette levner som oftest kun de «ekte» objektene.

Valg av parameterverdier

Det forutsettes et godt valg av terskelverdi T_G for at metoden skal lykkes. Jeg forsøkte å benytte denne metoden manuelt på noen bilder, og ut fra dette forsøkte jeg å velge parameterverdier. De innledende forsøkene antydte at verdier i området 8–20 var egnet. Dette er reflektert i parameterverdiene i Tabell 4.6.

Gradientterskel 8, 11, 14, 17, 20

Tabell 4.6: Parameterverdier for fjerning av objekter i homogene områder

4.3 Glatting av tersklet bilde

Etter terskling er ofte kantene til det lokaliserte fotavtrykkets ujevne. Det lokaliserte fotavtrykket har også små hull. For å fjerne disse defektene benyttet jeg tre standardmetoder for prosessering av binære bilder: Morfologisk åpning, morfologisk lukking og medianfiltrering av det binære bildet.

Lukkingsoperasjonen vil oftest øke objektenes omfang, og minske eventuelle hull i objektene. Åpningsoperasjonen gjør det motsatte — minsker objektomfang og forstørker hull i objektene. Medianfiltrering havner midt i mellom disse to metodene, og endrer objektenes omfang etter den lokale kontekst.

Valg av parameterverdier — morfologi

Morfologisk åpning setter en klar grense for hvor stort strukturelementet kan være. Hvis strukturelementet er større enn en del av objektet, vil denne delen forsvinne.

Morfologisk lukning setter en tilsvarende begrensning. Hvis avstanden mellom to egentlig disjunkte objekter er mindre enn strukturelementets størrelse, vil objektene bli feilaktig forenet.

Fotavtrykket har en bredde som for mine data var rundt 3–7 pikseler. Jeg valgte derfor strukturelementet til å være et sirkulært objekt med størrelse som vist i Tabell 4.7.

Strukturradius 3, 5, 7

Tabell 4.7: Parameterverdier for morfologiske operasjoner på tersklet bilde. Sirkulært strukturelement.

Valg av parameterverdier — medianfiltrering

Medianglatting trenger en parameter: Vindusstørrelsen. Fotavtrykket har en bredde på 3–5 pikseler. For at ikke fotavtrykket skal bli fjernet som helhet, må fotavtrykket minst kunne fylle halvparten av vinduet. Dette innebærer at vinduet ikke kan være større enn 6–10 pikseler. Dette danner grunnlag for valget av parameterverdier i Tabell 4.8.

Vindusstørrelse (3 × 3), (5 × 5), (7 × 7)

Tabell 4.8: Parameterverdier for medianglatting av tersklet bilde.

Kapittel 5

Data og eksperimenter

Ved hjelp av eksperimentene ønsket jeg å finne to typer resultater:

1. *Hvilken metode(-kombinasjon) er best egnet.* Det finnes mange tersklingsmetoder som kan benyttes til lokalisering av sonarfotavtrykk. Kapittel 3 presenterte to etablerte metoder, mens kapittel 4 presenterte fem lokalt adaptive tersklingsmetoder. Metodene i kapittel 4 ble valgt ut da de kom best ut i en omfattende evaluering av tersklingsmetoder for dokumentbilder (Trier & Taxt 1995*b*). Blant disse metodene ønsket jeg å finne den metoden som var best egnet for lokalisering av fotavtrykk.
2. *Hvilke parameterverdier er best egnet.* Innledende eksperimenter antydte at resultatet til en metode er svært avhengig av hvilke parameterverdier som benyttes. Uten å teste hver metode med forskjellige parametre er det ikke realistisk å kunne sammenligne metodene.

Metodene ble testet på reelle sonardata med ulike egenskaper. Første del av dette kapitlet omhandler utvalg av data for eksperimentene. Deretter følger en presentasjon og diskusjon av evalueringsmetoder for å kunne rangere metodene. Den faktiske gjennomføring av eksperimentene er omtalt i Tillegg A.

5.1 Data

Jeg benyttet femten forskjellige ping eller bilder til uttesting av metodene. Bildene kom fra tre forskjellige dataserier med sekvensielle ping. Dataene ble innsamlet fra forskjellige steder med ulike bunnforhold i sjøen utenfor Horten.

Jeg valgte å benytte to serier fra en skråstilt multistrålesonar og en serie fra en multistrålesonar montert vertikalt. Dette valget begrunnes i at skråstilte sonarer er vanligere, og at SIMRAD også anbefaler at multistrålesonaren monteres skråstilt. Følgende serier ble benyttet:

1. *Flat havbunn, skrå sonar.* Dette gir et sammenhengende fotavtrykk, som er nesten en rett linje (Figur 5.1(a1)). Dette skyldes at sonaren her er montert med en vinkel på 45° på skipets vertikallakse. Dette er muligens det «enkleste» fotavtrykket.
2. *Lite objekt på havbunnen, skrå sonar.* I denne serien er det et objekt på havbunnen som stikker opp. Dette objektet skygger igjen for havbunnen. Dermed blir fotavtrykket noe mer usammenhengende. Også

her er sonaren montert med en vinkel på 45° på skipets vertikallakse. Dette gir et fotavtrykk som også er tilnærmet lik en rett linje (Figur 5.1(a2)).

3. *Flat havbunn, vertikal sonar.* Sonaren er montert i vertikal retning. Havbunnen er nokså flat under båten. Dette gir et fotavtrykk som har en «U»-lignende form (Figur 5.1(a3)).

Samtlige testdata ble gitt av Kongsberg SIMRAD og kommer fra en SIMRAD EM3000 multistrålesonar. De viktigste parameterne for denne sonaren er gitt i Tabell 5.1 (Kongsberg SIMRAD 1994).

Antall mottakere:	80
Avstand mellom hver mottaker:	2.55 mm
Frekvens:	300 kHz
Båndbredde:	4-8 kHz

Tabell 5.1: Noen viktige parametere for en SIMRAD EM3000 sonar.

5.1.1 Stråleforming

Ved stråleforming benyttet jeg sju subarrayer med seks elementer mellom hvert subarray. Subarrayene hadde en lengde på 32 elementer, og jeg benyttet 32 stråler som dekket en spredning på 180° . Dette valg av stråler og subarrayer vurderes som aktuelt for et fremtidig sonarbehandlingsystem (Yang 1998b). En annen fordel ved dette verdivalg for stråleformingsparameterene var at bildene ble små og opptok lite lagringsplass.

Etter stråleforming ble bildene konvertert fra 64-bits flyttall til 8-bits heltall. Dette var nødvendig, da de lokalt adaptive tersklingsmetodene kun var implementert for 8-bits heltall i XITE. Dataene ble lineært skalert slik at det fulle spennet av pikselverdier ble benyttet i 8-bitsbildet. Skaleringen ble gjort for å minimere datatapet.

5.1.2 Preprosessering

Glatting av støy

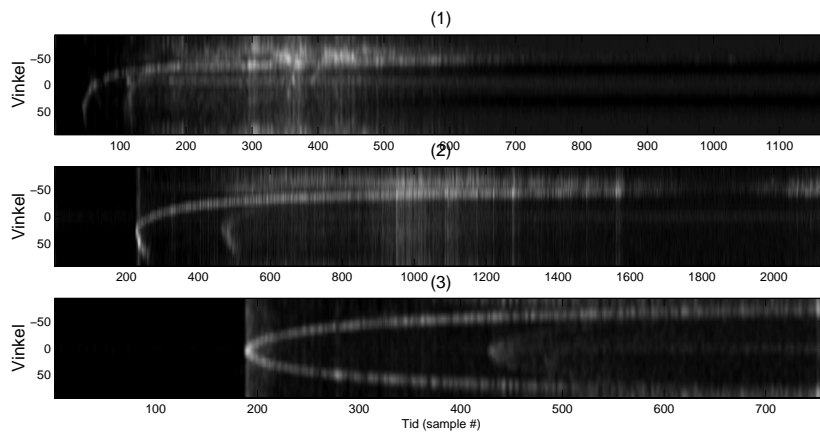
Støy kan forringe sonarbildet og gi en mer upresis lokalisering av fotavtrykket. Fjerning (eller glatting) av denne støyen gir et bedre resultat. Ut fra bildedannelsen antok jeg at bildene inneholdt additiv normalfordelt støy uten romlig korrelasjon. Bildene ble derfor glattet med et (3×3) gjennomsnittsfilter.

Histogramforbedring

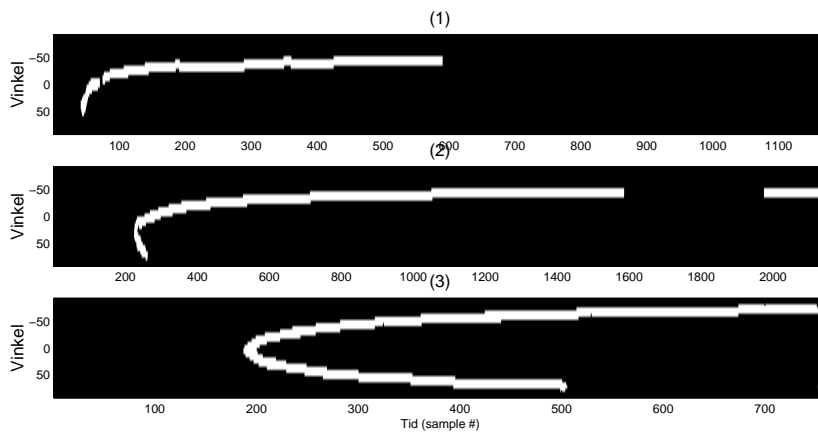
En del tersklingsmetoder (f.eks Eikvil et al. 1991, Kittler & Illingworth 1986) gjør en antagelse om at histogrammet til innbildet kan modelleres som en sum av to Gauss-fordelinger. Dette er ikke tilfelle for amplitudebildets histogram (Figur 5.2). Det er nødvendig å «normalisere» amplitudebildets histogram for å oppnå gode resultater.

Anta at bakgrunnsstøyen er normalfordelt og like sterk i alle vinkler. Ved stråleforming blir støy utenfor den aktuelle strålen filtrert bort. Bakgrunnsstøyen $f(x)$ kan dermed modelleres ved (Thomas 1969, s. 155-167)

$$f(x) = E(x) \cos[kx + \phi(x)] \quad (5.1)$$

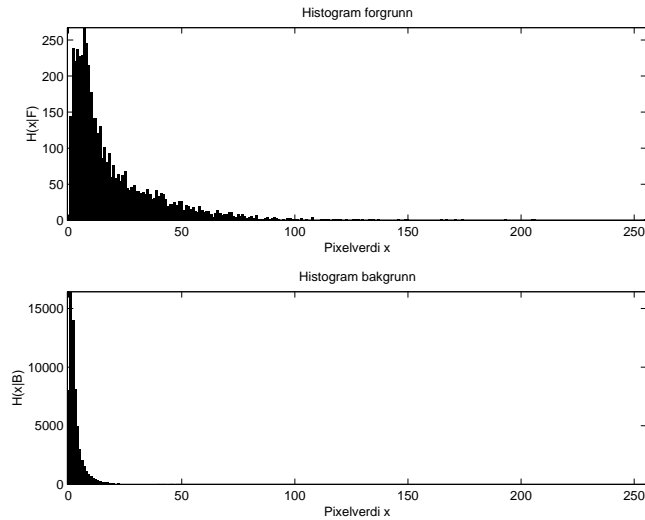


(a) Amplitudebilde.



(b) Referansebilde. Hvitt: Fotavtrykk; Svart: Bakgrunnsstøy.

Figur 5.1: Eksempel på amplitudebilde og referansebilde til et ping fra hver testserie. Serie (1): Flat havbunn; sonar montert 45° på vertikal retning; Serie (2): Havbunn med skyggende objekt; sonar montert 45° på vertikal retning; Serie (3) Ujevn havbunn; sonar montert i vertikal retning. Geometrien er gitt i polarkoordinater.



Figur 5.2: Histogram til et amplitudebilde.

hvor $E(x)$ er en stokastisk omhylningskurve, $\phi(x)$ er en stokastisk fasefunksjon, x er mottakerposisjon og k er bølgetallet for sentervinkelen i strålen (avsnitt 2.2).

Ekkoet av sonarens lydimpuls mottas fra havbunnen, og dette ekkoet adderes til bakgrunnsstøyen. Signalet som mottas kan dermed modelleres som et sinussignal addert til global hvit støy. Det mottatte signalet $g(x)$ blir dermed

$$g(x) = f(x) + A \cos[kx + \Psi] \quad (5.2)$$

hvor $\Psi \in [0, 2\pi]$ angir fase og A amplitude til havbunnsekkoet. Dette signalet modelleres så som

$$g(x) = R(x) \cos[kx + \theta(x)] \quad (5.3)$$

hvor $R(x)$ er en stokastisk omhylningskurve og $\theta(x)$ er en stokastisk fasefunksjon. R representerer den stokastiske fordelingen til amplituden i sonarbildet etter stråleforming.

Det kan vises (Thomas 1969, avsn. 4.7-4.8) at siden $f(x)$ er normalfordelt så er R Rice-fordelt, altså

$$P(R) = \frac{R}{\sigma^2} e^{-(R^2+A^2)/2\sigma^2} I_0\left(\frac{AR}{\sigma^2}\right) \quad (5.4)$$

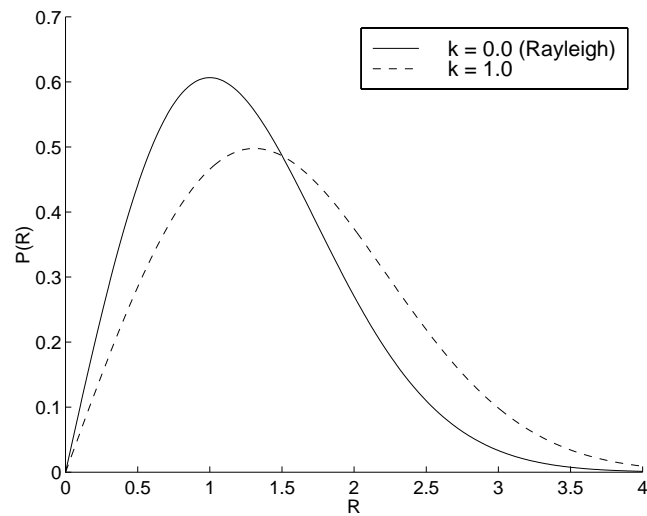
hvor I_0 er den modifiserte Bessel-funksjonen av første type og orden null. Fordelingen er vist i Figur 5.3.

Histogrammet i Figur 5.2 er altså en Rice-fordeling. Jeg ønsker å transformere amplitudesignalet R med en transform $u(\cdot)$ slik at $P(u(R))$ blir normalfordelt. Dette kan gjøres analytisk ved å benytte de kumulative fordelingene til normalfordelingen og Rice-fordelingen (Gonzalez & Woods 1992, s. 180-182). Dette er imidlertid en regnetung operasjon.

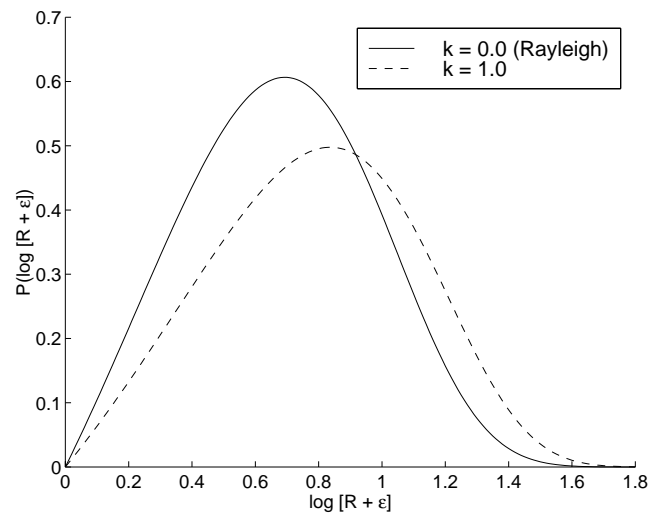
Det viser seg at $u(R) = \log(R + \varepsilon)$ gir gode resultater. Fordelingen $\log(R + \varepsilon)$ tilnærmer en normalfordeling og er tilnærmet symmetrisk (Figur 5.4).

Jeg danner derfor utbildet $U(x, y)$ ved å transformere amplitudebildet $I_A(x, y)$ ved

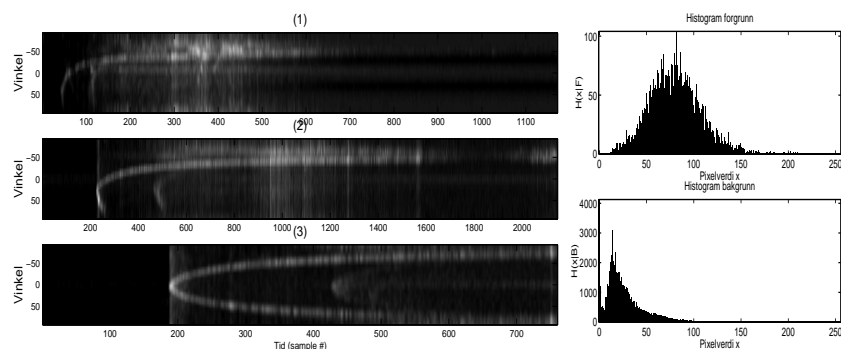
$$U(x, y) = \log[I_A(x, y) + \varepsilon] \quad (5.5)$$



Figur 5.3: Plot av Rice-fordelingen $f(x) = x \exp[-(x^2 + a^2)/2]I_0(ax)$. Amplitudesignalet har en slik distribusjon. For $a = 0$ tilsvarende Rice-fordelingen Rayleigh-fordelingen.



Figur 5.4: Fordelingen for $\log[R + \epsilon]$ hvor R er Rice-fordelt. $\epsilon = 1$. Logaritmetransformasjonen gir en fordeling som tilnærmer en normalfordeling.



Figur 5.5: Resultat etter logaritmisk transform av amplitudebilde med påfølgende gjennomsnittsglatting, (3×3) vindu. Det første bildet i hver av de tre dataseriene er vist. Histogram vist for det midterste bildet. Bunnekket er den lyse, bueformede strukturen i bildene. Geometri gitt i polarkoordinater.

hvor ε er en liten konstant. $U(x, y)$ brukes så for videre tersking. Etter gjennomsnittsglatting ligner histogrammet til fotavtrykket i $U(x, y)$ mye på en normalfordeling (Figur 5.5).

5.2 Evalueringmetoder

I eksperimentene genererte jeg svært mange bilder som skulle evalueres. Antallet parameterverdier og metodekombinasjoner gjorde en manuell vurdering av resultatet vanskelig. En menneskelig observatør vil dessuten benytte skjønn ved evaluering. Dette vanskeliggjør objektivitet i evalueringen, spesielt ved så mange bilder.

Jeg valgte å automatisere evalueringen av resultatene i så stor grad som mulig. Utvalget av publiserte evalueringmetoder er begrenset. Jeg har valgt å basere meg på de metodene som ble presentert i (Zhang 1996), som sammenligner og evaluerer kjente segmenteringsevalueringmetoder. Zhang (1996) grupperer automatiske evalueringmetoder inn i to grupper:

1. Metoder som ikke evaluerer ut fra et referanse bilde
2. Metoder som evaluerer ut fra et referanse bilde

Førstnevnte gruppe baserer seg på antagelser om sammenhenger mellom det segmenterte bildet og originalbildet. Disse antagelsene ligner ofte på de samme antagelsene som benyttes ved design av terskingsmetoder. Disse metodene kan ikke gi den nødvendige presisjonen ved evaluering. Jeg valgte derfor å ikke studere disse metodene videre.

Gruppe to forsøker å tallfeste forskjellen mellom det genererte bildet og et referansebilde, ut fra antagelser om hvordan øyet sammenligner to binære bilder. Bare metoder som faller innenfor denne gruppen kan gi en objektiv evaluering av segmenteringsresultatet (Zhang 1996). Videre er disse metodene overlegne til å finne det beste segmenteringsresultatet ut fra en samling med nær perfekte resultater. Problemet med metodene er at samtlige bilder må manuelt segmenteres. Dette kan være et tidkrevende arbeid.

5.2.1 Etablerte evalueringsmetoder

Segmenteringsfeilen kan kvantifiseres ved å telle antall feilsegmenterte pikseler. Hvis et bilde klassifiseres inn i N klasser, kan man sette opp en matrise C av størrelse $N \times N$, hvor $C_{i,j}$ angir antall j -klasse pikseler som blir klassifisert som i -klasse pikseler. To typer feilmål kan videre beregnes ut fra denne matrisen for en klasse k . Type I feil beregner hvor stor andel av det totale antall k -pikseler som ikke har blitt klassifisert som klasse k . Type II feil beregner hvor stor andel av det totale antall ikke- k pikseler som har blitt klassifisert som k -pikseler. Målene beregnes som følger:

$$M_I^k = 100 \times \frac{\left(\sum_{i=1}^N C_{i,k}\right) - C_{kk}}{\sum_{i=1}^N C_{ik}} \quad (5.6a)$$

$$M_{II}^k = 100 \times \frac{\left(\sum_{i=1}^N C_{k,i}\right) - C_{kk}}{\left(\sum_{i=1}^N \sum_{j=1}^N C_{ij}\right) - C_{kk}} \quad (5.6b)$$

Disse ligningene gir imidlertid mer enn et feilmål per segmenteringsresultat. I mitt to-klasse-tilfelle gir dette to forskjellige feilmål per bilde. Lee, Chung & Park (1990) foreslo derfor å kombinere målene ved å se på *sannsynligheten* for at en piksel er feilklassifisert. Denne feilsannsynligheten er gitt ved

$$P_E = P(F) \times P(B|F) + P(B) \times P(F|B) \quad (5.7)$$

hvor $P(B|F)$ er sannsynligheten for å klassifisere en objektpiksel som bakgrunn (vice versa for $P(F|B)$) og $P(F)$ og $P(B)$ er *a priori* sannsynlighet for henholdsvis objektpikseler og bakgrunns-pikseler.

I Zhang's (1996) sammenligning av evalueringsmetoder var feilklassifikasjonssannsynligheten det beste målet for å skille mellom nær perfekte segmenteringsresultater. Målet er også raskt å beregne — beregningen har kompleksitet lik $\mathcal{O}(N)$ hvor N er antall pikseler.

Å beregne avvik kun på grunnlag av antall forskjellige piksler samsvarer dårlig med hvordan det menneskelige øyet fungerer. Andre feilmål tar derfor også hensyn til posisjonen til de feilsegmenterte pikselene (Yasnoff, Mui & Bacus 1977, Strasters & Gerbrands 1991). Det finnes også enklere mål som ser på samsvar mellom objektregioner i stedet for enkeltstående piksler (Yasnoff et al. 1977, Strasters & Gerbrands 1991). Disse enkle målene er imidlertid uegnet til å skille mellom annet enn svært dårlige segmenteringsresultater (Zhang 1996, avsn. 6).

5.2.2 Evaluering av segmenteringsresultat

Et av målene for prosjektet var å finne en metode som kunne benyttes som et forprosesseringsteg i andre bunndeteksjonsalgoritmer. En presis identifikasjon av fotavtrykket medfører at dataene utenfor fotavtrykket kan ignoreres.

Det er derfor svært viktig at segmenteringsmetoden utelater svært lite av fotavtrykket, også om dette medfører at bakgrunnsstøy blir feilklassifisert som fotavtrykk. Samtidig er det ønskelig at minst mulig bakgrunnsstøy feilklassifiseres som fotavtrykk.

Det viser seg at ingen av de tidligere nevnte evalueringsmetodene er gode nok i seg selv. Følgende metoder er uaktuelle:

1. *Metoder som vektet forgrunn og bakgrunn likt.* Problemet med de fleste metodene omtalt i avsnitt 5.2.1 er at de bygger på en antagelse om at det er like viktig å klassifisere både forgrunn (fotavtrykk)

og bakgrunn riktig. Her er det imidlertid viktigst at hele forgrunnen (altså fotavtrykket) blir funnet — også om dette medfører at deler av bakgrunnen blir feilaktig klassifisert.

2. *Metoder som kun ser på forgrunn eller bakgrunn.* En mulighet er å kun se på hvor godt forgrunnen ble segmentert. Dette kan gjøres ved å se på Type I feilen M_I^F for forgrunnen (avsn. 5.2.1). Dette vil imidlertid ikke kunne skille en perfekt segmentering fra en segmentering som klassifiserer hele bildet som forgrunn. Et slikt mål blir derfor også uegnet.

Jeg valgte til slutt å kombinere Type I og Type II feil for forgrunnen. Det er viktig at $M_I^F < T_F$ hvor T_F er en valgt terskel. Dermed er andelen av fotavtrykket som mangler begrenset. Forøvrig er det ønskelig at å minimere sannsynligheten $M_{II}^F = M_I^B$ for at en bakgrunnspiksel er feilklassifisert. Ut fra disse to kriteriene valgte jeg å benytte følgende mål for å rangere resultatene:

$$PR = \begin{cases} 1 & M_I^F > T_F \\ M_{II}^F & M_I^F \leq T_F \end{cases} \quad (5.8)$$

Dette målet gjør at parameterverdier som ikke tar med hele fotavtrykket blir sortert som svært dårlige. Videre gjør målet det mulig å sortere på hvor mye bakgrunn som faktisk har blitt fjernet.

Ved å sammenligne faktiske tersklingsresultater med M_I^F fant jeg at $T_F = 0.05$ var en egnet grense. Jeg kunne da ikke se noen stor visuell forskjell mellom segmenteringsresultatet og det ideelle resultatet.

5.2.3 Evaluering av regnetid

For å kunne ha praktisk anvendelse, er det viktig at metoden ikke er for regnetung. Regnetunge metoder medfører at dyrere maskiner må benyttes, og kan i verste fall gjøre en metode uaktuell.

Her har jeg målt regnetiden som antall CPU-sekunder T som metoden benytter på en SUN SPARCstation 10. Denne tiden er imidlertid sterkt avhengig av sonarbildenes størrelse. For å normalisere resultatene, antar jeg at metodene har algoritmekompleksitet $\mathcal{O}(N)$ hvor N er antall piksler i bildet. Da kan tidsforbruket normaliseres ved

$$TM = \frac{T}{N}. \quad (5.9)$$

Dette målet benyttes til en rangering av metoderesultatene basert på regnetid.

5.2.4 Referansebilder

Tegning av referansebilder vanskeliggjøres ved at fotavtrykkets grenser er uklare. Langs kanten til fotavtrykket er det vanskelig å avgjøre hvorvidt en piksel er fotavtrykk eller ei. Jeg fant det derfor nyttig å introdusere «likegyldighetsregioner» i referansebildene. Likegyldighetsregioner er regioner som det er likegyldig om metoden inkluderer i fotavtrykket.

Det var nødvendig å modifisere evalueringsmetoden slik at den støttet likegyldighetsregioner. Dette ble gjort ved å tegne to bilder: Det ideelle segmenteringsresultatet S_I og et separat likegyldighetsbilde S_D som angir likegyldighetsregioner. Pikseldiene i hvert bilde ble valgt som i Tabell 5.2.

For et tersklingsresultat S kan det maskerte bildet S' beregnes ved

$$S' = S \cap S_D \quad (5.10)$$

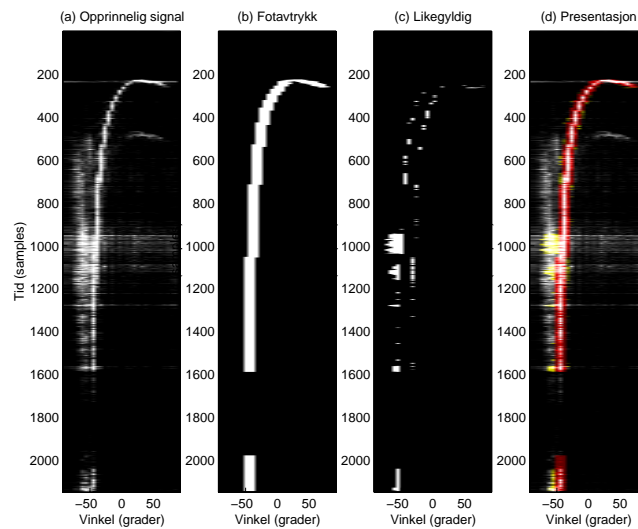
Pikseltype	S_I	S_D
Bakgrunn	0	0
Likegyldig	0	1
Forgrunn	1	0

Tabell 5.2: Pikselerverdier for tegning av referansebilder. S_I : Ideelt segmenteringsresultat. S_D : Likegyldighetsbilde.

hvor \bar{S}_D angir logisk negasjon. S_I og S' kan nå evalueres mot hverandre, uten at likegyldighetsregioner i S tas hensyn til.

Tegning av referansebildene

Jeg måtte tegne en del referansebilder som ledd i evalueringen. Dette ble gjort ved å ta utgangspunkt i et godt visuelt resultat fra en valgt segmenteringsmetode (i dette tilfellet Bernsens metode). Bildene ble så manuelt retusjert i Photoshop (Adobe 1998). Det var lett å segmentere bildet korrekt, da programmet la masken over sonarbildet (Figur 5.6).



Figur 5.6: Forenkling av manuell terskling ved hjelp av Photoshop. (a) Opprinnelig signal; (b) og (c) Maskene som skal redigeres; (d) Presentasjon av redigerbare masker over fastholdt signal i Photoshop. Rødt: Fotavtrykk. Gult: Likegyldig.

Kapittel 6

Eksperimentelle resultater

Prosesseringskjeden hadde to ledd: Terskling og postprosessering. Tersklingsleddet lokaliserte fotavtrykket og postprosesseringsleddet forbedret resultatet fra tersklingsleddet ytterligere. Postprosesseringen besto igjen av to trinn, det første trinnet fjernet falske objekter i de tersklede bildene det andre trinnet forsøkte å glatte objektenes kanter.

Resultatet til metodene i prosesseringskjeden var sterkt avhengig av hvilke parameterverdier som ble benyttet. For å kunne sammenligne metodene bedre, prøvde jeg ut hver metode med mange forskjellige parameterverdier. Innenfor hvert ledd i prosesseringskjeden ble så de dårligste resultatene ekskludert.

Resultatene ble evaluert ved å se på fotavtrykkets Type I og Type II feil. Feilmålene ble beregnet ved å sammenligne det automatisk segmenterte bildet med et referansebilde. Type I feil beregner hvor stor andel av fotavtrykket som ikke har blitt klassifisert som fotavtrykk. Type II feil beregner hvor stor andel av bakgrunnen som har blitt klassifisert som fotavtrykk. I evalueringen benyttet jeg femten testbilder fra tre forskjellige dataserier. En øvre grense ble satt for fotavtrykkets Type I feil M_I^F . Grensen ble satt til $M_I^F < 0.05$ (ligning (5.8)). Dermed er andelen av fotavtrykket som mangler begrenset. Fotavtrykkets Type II feil M_{II}^F ble deretter benyttet for rangering. Resultatene ble også vurdert subjektivt.

Gjennom hele kapitlet benyttes tre sonarbilder for å illustrere resultatene, et fra hver av dataseriene. De opprinnelige amplitudebildene er vist i Figur 5.1.

6.1 Segmenteringsresultat

Det er vanskelig å omtale resultatet før og etter postprosessering separat — resultatene endrer seg vesentlig ved postprosessering. Det beste resultatet etter segmentering gir ikke nødvendigvis det beste resultatet etter påfølgende postprosessering. Jeg velger derfor å omtale disse to stegene samlet.

6.1.1 Objektiv vurdering

Bernsens metode var klart beste metode til å lokalisere fotavtrykk. Metoden har både den laveste feilraten (Figur 6.1(a)) og rask eksekvering (Figur 6.1(b)). Selv uten påfølgende postprosessering var resultatet godt.

Eikvil et al.'s (1991) metode ga nest best resultat, og Niblack's (1986) metode tredje best. Regnetiden til begge metodene var omtrent lik (Figur 6.1). Yanowitz & Bruckstein's (1989) metode og Trier & Taxt's (1995a) metode ga dårligst resultat og benyttet lengst tid på å produsere det (Figur 6.1).

Den endimensjonale metoden for lokalisering av sonarfotavtrykk (referansemetoden) ga dårlige resultater. Segmenteringsfeilen var rundt tre ganger større enn for Bernsens metode. Referansemetoden var imidlertid ti ganger raskere enn Bernsens metode (Figur 6.1(b)). Bernsens metode brukte likevel under to sekunder per bilde.

Yanowitz & Bruckstein's (1989) gradientbaserte metode for fjerning av objekter i homogene områder (avsnitt 4.2.2) ga gode resultater for de vindusbaserte tersklingsmetodene, altså metodene til Bernsen (1986), Eikvil et al. (1991) og Niblack (1986) (Figur 6.1(a)). Metoden bedret ikke tersklingsresultatene for enkel endimensjonal terskling (Satriano & Smith 1991, Lingsch & Robinson 1992), Yanowitz & Bruckstein's (1989) metode og Trier & Taxt's (1995a) metode (Figur 6.1(a)). Metoden var imidlertid noe regnekrevende (Figur 6.1(b)).

Yang's (1998a) amplitudebaserte metode (avsnitt 4.2.1) forbedret tersklingsresultatet til Niblack's (1986) metode mye. For de resterende tersklingsmetodene forbedret denne postprosesseringsmetoden tersklingsresultatet lite.

Den gradientbaserte metoden ga altså gjennomgående bedre resultater enn den enklere amplitudebaserte metoden. Gradientbasert postprosessering tok imidlertid noe mer tid (Figur 6.1).

Etter falske objekter hadde blitt fjernet, ble segmenteringsresultatet glattet med morfologiske operatører eller medianoperatoren. Medianbasert postprosessering ga noe dårligere resultater enn morfologiske operatører (Figur 6.1(a)). Morfologiske operatører krever imidlertid mer regnekraft enn medianoperatoren (Figur 6.1(b)).

Kort oppsummert ble best resultat oppnådd ved bruk av Bernsen's (1986) metode etterfulgt av Yanowitz & Bruckstein's (1989) gradientbaserte postprosesseringssteg og morfologisk åpning av det segmenterte bildet.

6.1.2 Subjektiv vurdering

Bernsens metode (Figur 6.2(a)) ga et sammenhengende fotavtrykk. Fotavtrykkets kanter var nokså glatte, og reverberasjoner korrekt klassifisert som bakgrunnsstøy. Resultatet forringes riktignok noe av sideløber, disse sidelobene medfører at fotavtrykket er for bredt i enkelte områder. En subjektiv vurdering av resultatet til Bernsens metode bekrefter dermed de gode resultatene i avsnitt 6.1.1.

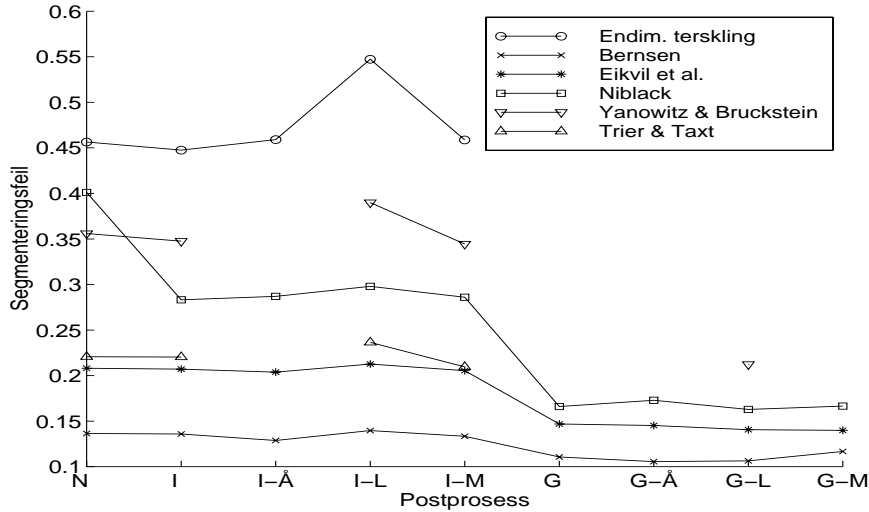
For Bernsens metode ga faktisk tersklingsresultatet vist i Figur 6.3(a) best resultat etter postprosessering. Dette bildet inneholder riktignok mye bakgrunnsstøy. Denne støyen fjernes imidlertid godt av postprosesseringsleddet til Yanowitz & Bruckstein's (1989) metode.

Eikvil et al.'s metode (Figur 6.2(b)) gir et godt resultat — faktisk nesten like godt som resultatet til Bernsens metode. Forskjellen består i at Eikvil et al.'s metode har noe større problemer med sideløber enn Bernsens metode. Dette medfører at fotavtrykket blir for bredt, og at kantene blir ujevne. Eikvil et al.'s metode feilklassifiserer også sekundære ekkoer (avsnitt 3.1.2) som fotavtrykk. Som i Bernsen's metode fjerner også her postprosesseringsleddet mye støy (jf. Figur 6.3(b)).

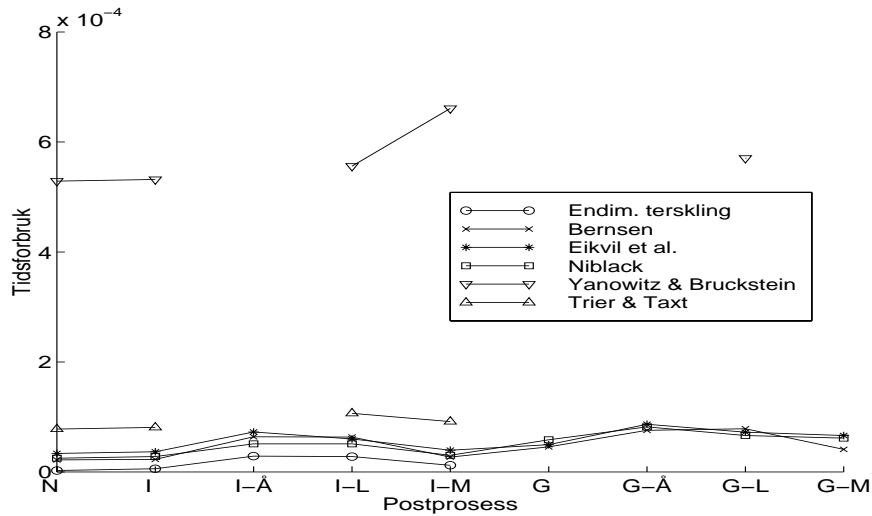
Niblack's metode (Figur 6.2(c) og 6.3(c)) har større problemer med sideløber. Problemene er så sterke at fotavtrykkets kanter ikke lenger er glatte, selv etter postprosessering.

Den eksisterende metoden for lokalisering av sonarfotavtrykk har spesielt store problemer med data fra sonarer som er montert vinklet (Figur 6.2(a1-2)). Metoden gir imidlertid usammenhengende fotavtrykk også for sonarer som ser rett mot bunnen (Figur 6.2(a3)).

Figur 6.2(e,f) viser også de dårlige resultatene til de resterende metodene.

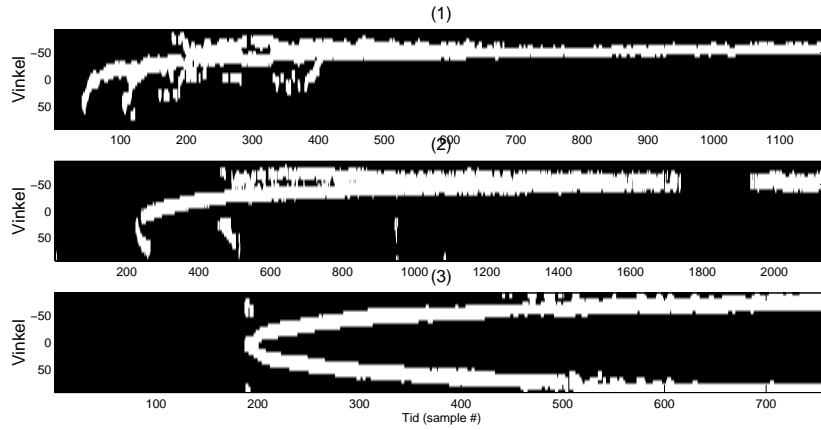


(a) Gjennomsnittlig segmenteringsfeil (ligning (5.8)) beregnet over alle femten bildene i treningssettet. Beste parameterkombinasjon valgt for hver metode.

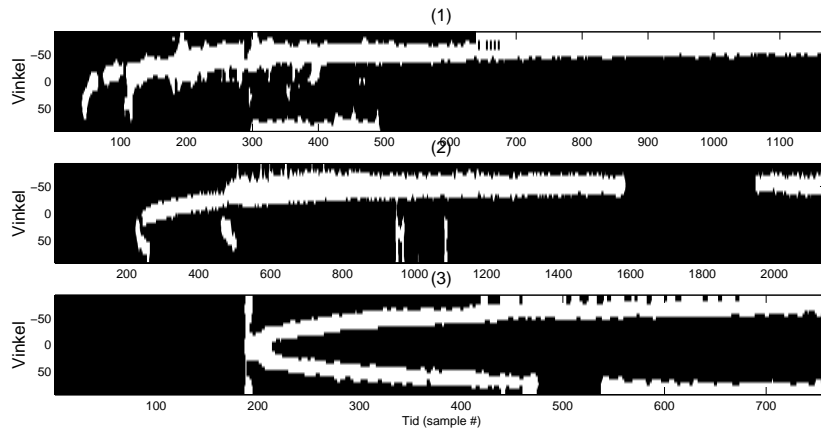


(b) Gjennomsnittlig tidsforbruk (ligning (5.9)) for tre av bildene i treningssettet (et i hver serie). Total CPU-tid (sekunder) per piksel på en SPARCstation 10.

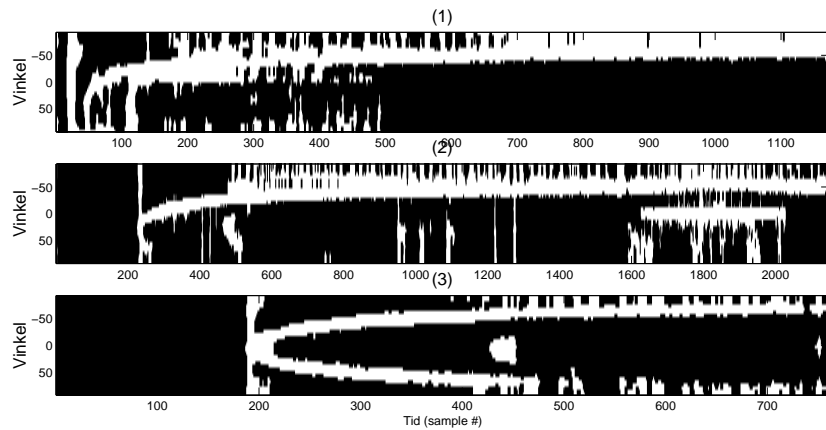
Figur 6.1: Resultat etter postprosessering. Ikke plottede punkter: Ingen utprøvd parameter medførte $M_I^F < 0.05$. G: Gradientbasert postprosessering (avsnitt 4.2.2); I: Amplitudebasert postprosessering (avsnitt 4.2.1); Å: Morfologisk åpning; L: Morfologisk lukking; M: Medianfiltrering av segmentert bilde; N: Ingen postprosessering.



(a) Bernsen's (1986) metode. Kontrastgrense $C = 10$. Vindusstørrelse $(dx \times dy) = (15 \times 3)$.

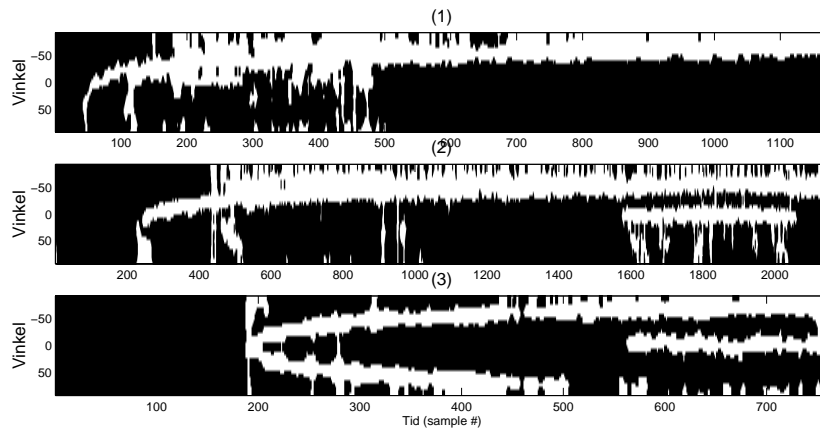


(b) Eikvil et al.'s (1991) metode. Størrelse stort vindu $L: (15 \times 15)$, størrelse lite vindu $S: (3 \times 3)$, kontrastgrense $\Delta\mu = 10$, oppdateringsvekt $w = 0.1$.

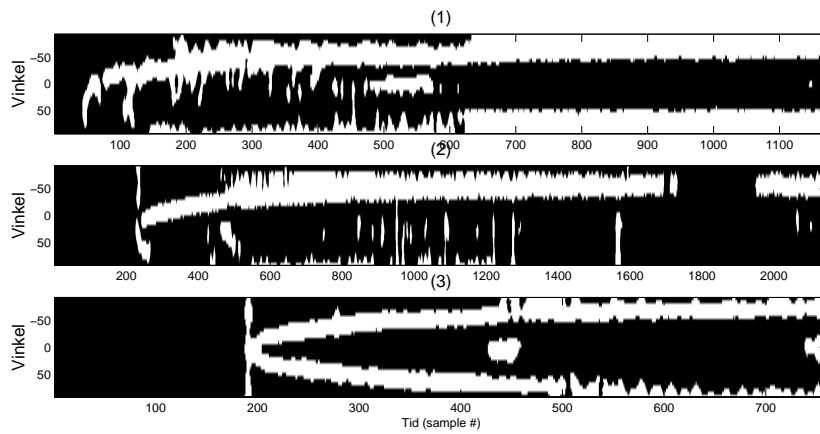


(c) Niblack's (1986) metode. Vindusstørrelse $s: (30 \times 30)$, standardavviksvekt $w = 0.1$.

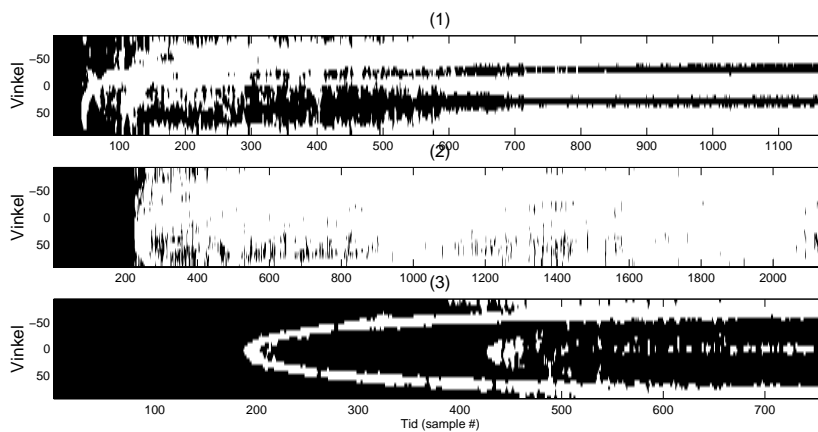
Figur 6.2: Resultatbilder etter postprosessering. Beste utfall vist for hver metode. Det første bildet i hver dataserie er vist.



(d) Yanowitz & Bruckstein's (1989) metode. Relaksjonskonstant $\beta = 1.0$, iterasjonsgrense $i = 150$, gradientterskel $T_G = 5$.

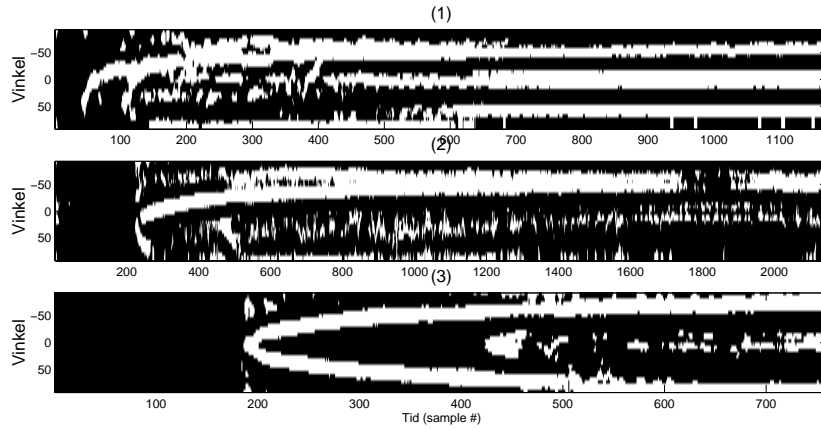


(e) Trier & Taxt's (1995a) metode. Aktivitetsgrense $T_A = 40$, minimal regionsstørrelse $S = 30$.

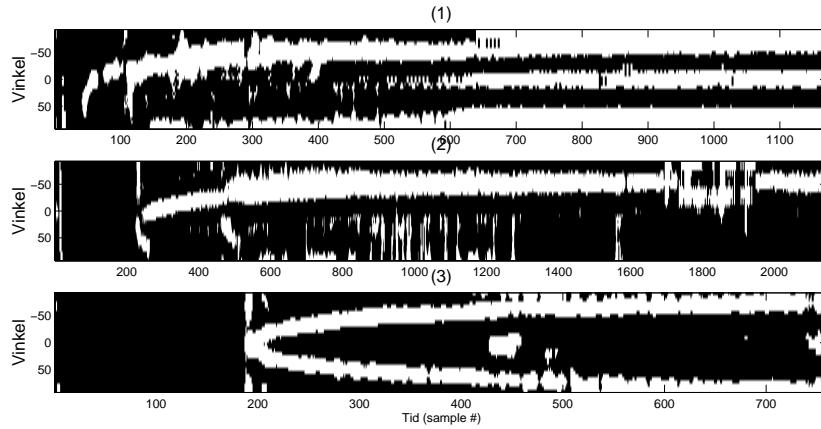


(f) Enkel endimensjonal terskling. Kontrastgrense $C = 36$.

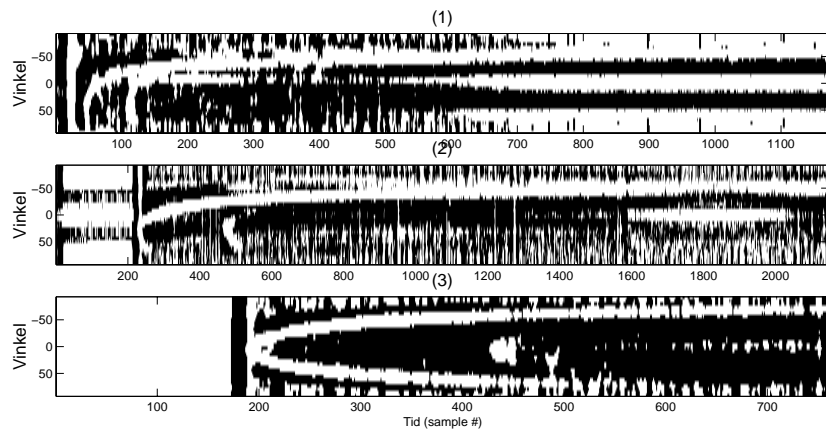
Figur 6.2: (forts.)



(a) Bernsen's (1986) metode.

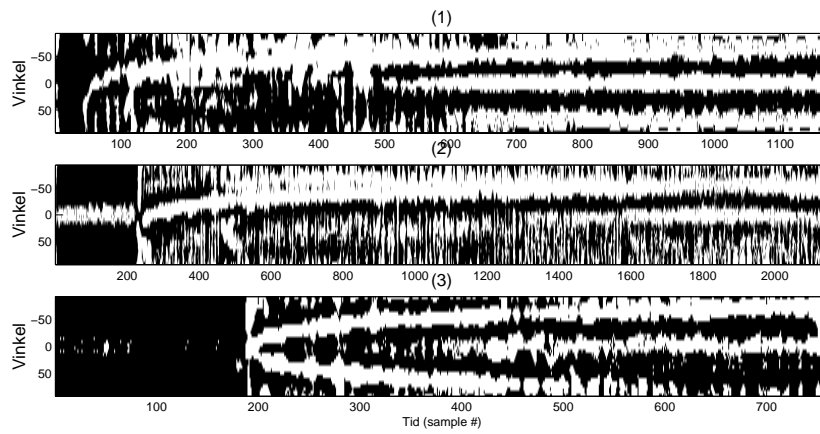


(b) Eikvil et al.'s (1991) metode

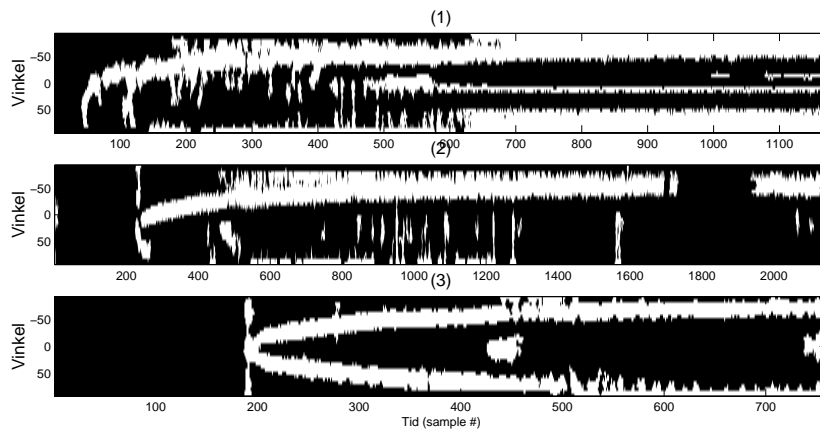


(c) Niblack's (1986) metode.

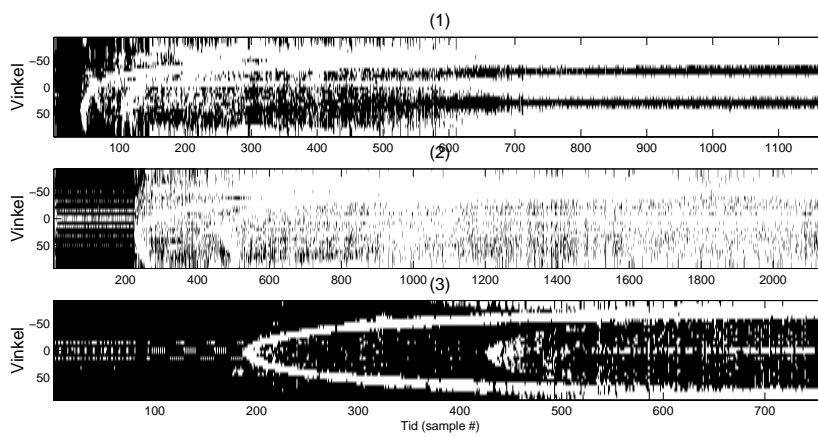
Figur 6.3: Resultatbilder etter segmentering (før postprosessering). De samme bildene er vist etter postprosessering i Figur 6.2.



(d) Yanowitz & Bruckstein's (1989) metode.



(e) Trier & Taxt's (1995a) metode.



(f) Enkel endimensjonal terskling.

Figur 6.3: (forts.)

6.1.3 Parameterverdier

Mange av parameterene tillater ikke noen entydig parameterverdi som samtidig minimerer M_I^F og M_{II}^F . Dette skyldes at det ofte er en samvariasjon mellom Type I feilen $M_I^F[\alpha]$ og Type II feilen $M_{II}^F[\alpha]$ for en metodeparameter α . Hvis α endres slik at $M_I^F[\alpha]$ minsker, medfører dette ofte at $M_{II}^F[\alpha]$ øker tilsvarende. α kan her representere vindushøyde, skaleringsvekt eller lignende parametre.

Det er vanskelig å finne det «beste» verdivalg for α når M_I^F og M_{II}^F avhenger av α på denne måten. Parametere som gir denne samvariasjonen mellom $M_I^F[\alpha]$ og $M_{II}^F[\alpha]$ påvirker bare hvor mye av bildet som klassifiseres som bakgrunn. Jeg gjengir i Figur 6.4 hovedsaklig parametre som viser invers samvariasjon.

Segmenteringsfeilen PR i ligning (5.8) bør være så liten som mulig, samtidig som $PR > 0$. De viktigste karakteristika for valg av parameterverdier er som følger:

Bernsens metode: Denne metoden styres av tre parametre: Vindusbredden, vinduslengden og en kontrastgrense. For å oppnå en akseptabel M_I^F må kontrastgrensen C velges liten.

Nokså brede vinduer er å foretrekke (Figur 6.4(a)), bare for brede vinduer blir M_I^F liten nok til at $PR < 1$. Samtidig er M_{II}^F stor for brede vinduer. En vindusbredde rundt 7–15 gir de beste resultatene.

Best resultater oppnås videre med små vinduslengder — både M_{II}^F og M_I^F minker med minkende vinduslengde (Figur 6.4(a)). Også kontrastgrensen C må velges liten for å oppnå en akseptabel M_I^F , og holdes her konstant. En liten C medfører imidlertid en større M_{II}^F .

Eikvil et al.'s metode: Denne metoden benytter to vinduer med størrelse $(L \times L)$ og $(S \times S)$, der $L > S$, samt en oppdateringsvekt w og en lokal kontrastgrense $\Delta\mu$.

Oppdateringsvekten w påvirker ikke resultatet noe særlig. Størrelsen S på det lille vinduet påvirker heller ikke resultatet nevneverdig. S og w holdes derfor konstante.

For å oppnå en akseptabel M_I^F må L være størst mulig (Figur 6.4(b)). En stor L øker samtidig ikke M_{II}^F vesentlig.

Kontrastgrensen $\Delta\mu$ må være liten for å oppnå en tilfredsstillende M_I^F (Figur 6.4(b)). Dette medfører at M_{II}^F øker betraktelig.

Niblack's metode: Denne metoden benytter et kvadratisk vindu med størrelse $(s \times s)$ og en skaleringskonstant w .

Både M_I^F og M_{II}^F minsker med økende vindusstørrelse (Figur 6.4(c)). Store vinduer egner seg altså best for denne metoden. w må videre velges liten for å få en akseptabel M_I^F (Figur 6.4(c)). Dette medfører at M_{II}^F øker tilsvarende.

Forbedret White & Rohrer's metode: To terskler styrer denne metoden, en terskel T_A for lokalt aktivitetsnivå («gradient») og en for minimal regionstørrelse S .

Aktivitetsnivåterskelen T_A må velges lav for at M_I^F skal være akseptabel (Figur 6.4(d)). Dette medfører at M_{II}^F blir tilsvarende høy.

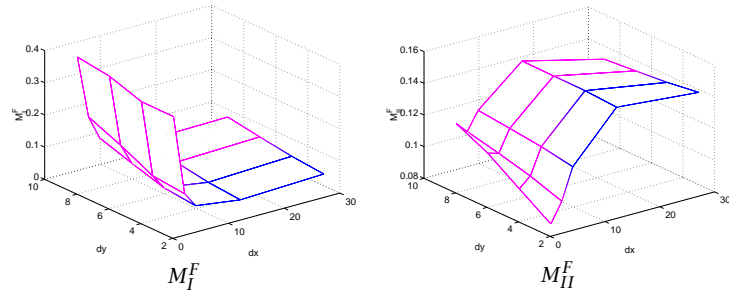
Minimal regionstørrelse S ser ikke ut til å påvirke resultatet nevneverdig når T_A er liten (Figur 6.4(d)). Konvertering av «+»-pikseler til «0»-pikseler hjelper ikke metoden vesentlig.

Yanowitz & Brucksteins metode: Tre parametre styrer denne metoden, en relaksjonskonstant β , en gradientterskel T_G og en iterasjonsgrense i . i påvirker ikke resultatet nevneverdig når $i > I, I \approx 10$, i holdes derfor konstant.

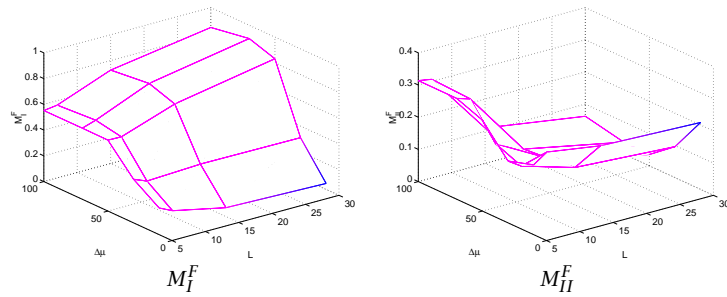
Hverken M_I^F eller M_{II}^F påvirkes nevneverdig av β så lenge $\beta < 1$. Feilen øker imidlertid sterkt straks $\beta \geq 1$ (Figur 6.4(e)).

M_I^F er ikke særlig påvirket av T_G , mens M_{II}^F minker med stigende T_G (Figur 6.4(e)). T_G må imidlertid velges svært liten for å passere den «magiske» grensen $M_I^F < 0.05$ i ligning (5.8). En større T_G gir en lavere M_{II}^F , men også et et fotavtrykk som er splittet opp i flere disjunkte regioner.

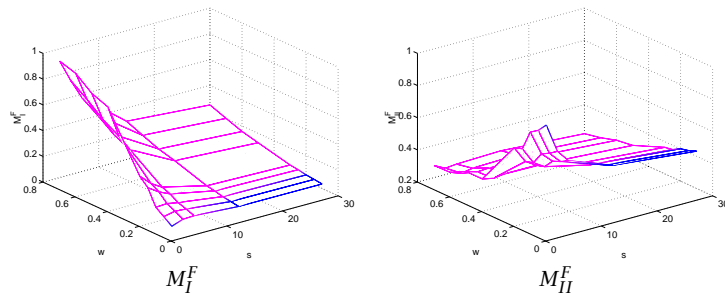
Enkel endimensjonal terskling: Denne metoden krever en skaleringskonstant C . $M_I^F[C]$ og $M_{II}^F[C]$ samvarierer som nevnt over, og $M_I^F[C]$ øker med økende C . Som tidligere nevnt gir dette ikke mulighet til særlig presise estimater.



(a) Bernsens metode. dx : Vindusbredde; dy : Vindushøyde. Kontrastgrense $C = 10$.

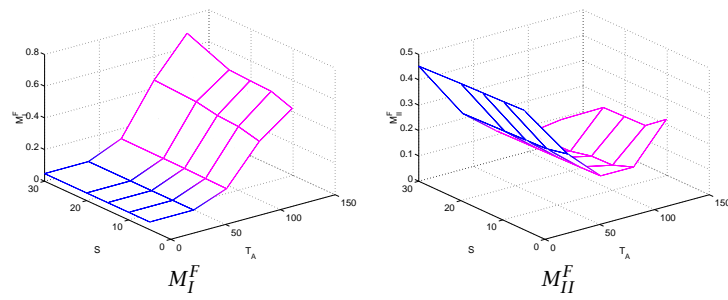


(b) Eikvil et als metode. L : Vindusstørrelse omliggende vindu; w : Oppdateringsvekt. Vindusstørrelsen S og oppdateringsvekten w holdes konstante, $S = 5$, $w = 0.9$.

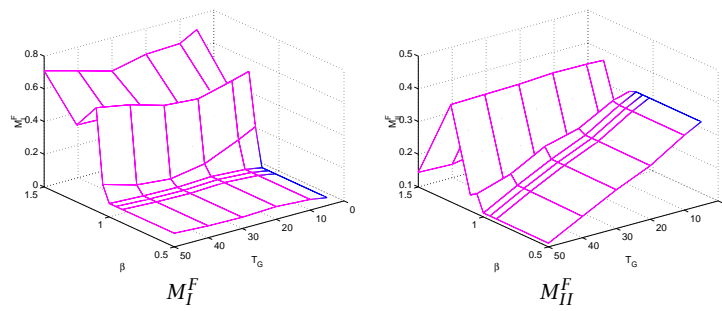


(c) Niblack's metode. w : Skaleringskonstant; s : Vindusstørrelse ($s \times s$)

Figur 6.4: Type I og Type II feil til segmenteringsresultatet som funksjon av utvalgte parametre. Magenta: $M_I^F > 0.05$; Blått: $M_I^F < 0.05$. Jf. ligning (5.8).



(d) Forbedret White & Rohrer's metode. T_A : Aktivitetsgrense; S : Minimal regionstørrelse



(e) Yanowitz & Bruckstein's metode. T_G : Gradientterskel; β : Relaksjonskonstant. Iterasjonsgrensen i holdes konstant, $i = 150$.

Figur 6.4: (forts.)

Kapittel 7

Diskusjon og konklusjon

Målet for denne hovedfagsoppgaven var å bedre lokaliseringen av sonarfotavtrykket ved hjelp av lokalt adaptive todimensjonale tersklingsmetoder. Som referansemetode brukte jeg en endimensjonal adaptiv metode (Satriano & Smith 1991, Lingsch & Robinson 1992). Jeg har evaluert de fem metodene som kom best ut i en omfattende sammenligning (Trier & Taxt 1995*b*). Metodene ble evaluert på et testsett bestående av femten sonarbilder, tatt fra tre forskjellige dataserier.

For å vurdere segmenteringsresultatene benyttet jeg to objektive evalueringsmål. De to målene beregnet Type I og Type II feil til sonarfotavtrykket. Type I feil beregner hvor stor andel av fotavtrykket som ikke ble klassifisert som fotavtrykk. Type II feil beregner hvor stor andel av bakgrunnen som ble klassifisert som fotavtrykk. En øvre grense ble satt for fotavtrykkets Type I feil, og fotavtrykkets Type II feil ble deretter benyttet for rangering. Resultat fra bilder som overskred grensen for Type I feil ble ekskludert fra de endelige resultatene.

Segmentering Resultatene viste at Bernsen's (1986) metode var best egnet til å lokalisere sonarfotavtrykk. Eikvil et al.'s (1991) metode var nest best og Niblack's (1986) metode var tredje best. Visuell inspeksjon av resultatbildene fra Bernsens metode bekreftet også de gode resultatene. Metod rangeringen samsvarer forøvrig med rangeringen til Trier & Taxt (1995*b*), med unntak av at Bernsens metode og Niblack's metode har byttet plass.

Sammenlignet med den eksisterende metoden for lokalisering av fotavtrykk (enkel endimensjonal terskling, (Satriano & Smith 1991, Lingsch & Robinson 1992)) kommer Bernsens metode svært godt ut. Selv med post-prosessering har den eksisterende metoden en feilrate som er tre ganger så stor som resultatene til Bernsens metode. Metoden har spesielt store problemer med data fra sonarer som er montert skråstilt.

Interessant nok er resultatet for referansemetoden for lokalisering av fotavtrykk så dårlig at postprosesseringssteget til Yanowitz & Bruckstein's (1989) metode ikke virker. Vanligvis er denne teknikken egnet for alle typer tersklingsmetoder (Trier & Taxt 1995*b*), og egnet seg også for de fleste tersklingsmetodene studert i denne oppgaven. Den endimensjonale metoden inkluderer imidlertid for mye av bakgrunnen i fotavtrykket, slik at kanten til det segmenterte objektet ikke lenger ligger langs kanten til fotavtrykket. Dette medfører at den gjennomsnittlige gradienten langs objektets kant blir for lav. Dermed fjernes fotavtrykket som helhet, i stedet for bare omliggende støy.

Parameterverdier Jeg studerte sammenhengen mellom parameterverdier og endelig resultat. For en del parametre samvarierte Type I feilen og

Type II feilen etter som parameterens verdi ble endret. Denne type samvariasjon er også tidligere kjent i litteraturen (Pratt 1991, s. 533).

Bernsen's (1986) metode og Eikvil et al.'s (1991) metode består av to uavhengige trinn. Trinn én gjør den initielle tersklingen av hver piksel. Trinn to finner homogene områder og klassifiserer disse områdene som bakgrunnsstøy. Begge metoder klassifiserer områder som homogene hvis den lokale kontrast mindre enn en viss terskel.

Begge metoder oppnådde best resultater når denne terskelen ble valgt liten. Dette kan tolkes som at metodene ikke selv bør klassifisere homogene områder som bakgrunnsstøy for denne type data. Slik klassifikasjon av homogene områder ga oftest et usammenhengende fotavtrykket. Homogene områdene klassifiseres mer presist som bakgrunnsstøy av postprosesseringssteget til Yanowitz & Bruckstein's (1989) metode.

Resultatene antydte at ikke-kvadratiske vinduer er best egnet når sonarbildene presenteres i polare koordinater. Bernsens metode ga bedre resultater med slike vinduer. Det er mulig at ytelsen til Eikvil et al.'s metode og Niblack's metode ble begrenset av at begge metodeimplementasjonene kun benytter kvadratiske vinduer. Denne begrensningen reduserte metodenes mulighet til å tilpasse seg raskt over tid, og kan forklare begge metodenes problemer med sidelober og suboptimale resultat.

De to lokalt adaptive todimensjonale tersklingsmetodene som ga dårligst resultat var Yanowitz & Bruckstein's metode og forbedringen av White & Rohrer's metode. Begge metodene benytter en global gradientterskel i stedet for små lokale vinduer. For å ikke utelate for store deler av fotavtrykket måtte denne terskelen velges svært lav. En så lav gradientterskel medførte at mye bakgrunnsstøy ble feilklassifisert som fotavtrykk. Feilklassifikasjonen skyldtes trolig at amplituden både innenfor og utenfor fotavtrykket varierer over tid. Global terskling av gradientbildet, i likhet med global terskling av amplitudebildet, er derfor lite egnet.

Postprosessering Tersklingsresultatet ble klart forbedret ved å benytte postprosesseringssteget til Yanowitz & Bruckstein's (1989) metode. Dette samsvarer med resultatene i Trier & Taxt (1995*b*). Metoden er imidlertid tidkrevende, og kan eventuelt utelates i en svært tidskritisk applikasjon. Å utelate slik postprosessering medførte bare 2-3% forskjell for Bernsens metode.

Den amplitudebaserte postprosesseringsmetoden (Yang 1998*a*) forbedret bildene lite. Den type støy som denne metoden fjerner, eliminerer Bernsens metode allerede godt ved hjelp av beregningen av lokal kontrast. Den amplitudebaserte metoden hadde en viss effekt for Niblack's (1986) metode, men ellers en marginal effekt for de resterende tersklingsmetodene.

Bruk av morfologiske operatører eller medianglatting forbedrer resultatet ytterligere. Forskjellen i resultatet mellom de to typer glatting var imidlertid for liten (1-1.5%) til å anbefale en spesifikk metode. Morfologiske operatører krevde imidlertid mer regnekraft.

Videre arbeid Segmenteringsresultatet fra Bernsens metode med postprosessering er som sagt svært godt. Sidelober skapte imidlertid noen problemer. Der sidelobene er svært sterke økes fotavtrykkets omfang feilaktig. Metodene jeg har benyttet har ikke hatt noen *a priori* kjennskap til fotavtrykkets form, og har derfor ikke evnet å redusere fotavtrykkets omfang tilsvarende. Det er mulig at segmenteringsresultatet kan forbedres ytterligere ved å postprosessere resultatet med metoder som tar hensyn til fotavtrykkets form.

Metodene til Eikvil et al. (1991) og Niblack (1986) bør prøves med ikke-kvadratiske vinduer. Siden Bernsens metode oppnådde bedre resultater

med slike vinduer, vil trolig samme forbedring skje for disse to metodene.

Et feilbilde og et fasekurvebilde kan også dannes fra rådataene (Yang & Taxt 1997). Verken feilbildet eller fasekurvebildet har blitt benyttet i dette studiet. Bruk av feilbildet sammen med amplitudebildet krever multispektral analyse, men dette kan gi bedret presisjon ved lokalisering av sonarfotavtrykk (Pratt 1991, avsn. 10.6). Det er også mulig at en teksturoperator anvendt på fasekurvebildet kan gi et egenskapbilde som kan forbedre resultatet ytterligere, spesielt ved klassifikasjon av de homogene områdene.

Konklusjon Dette studiet har lokalisert multistråle sonarfotavtrykk med todimensjonale bildebehandlingsmetoder. Fem lokalt adaptive tersklingsmetoder ble evaluert. Metodene ble testet på femten forskjellige sonarbilder tatt fra tre forskjellige serier med sonardata. Sonarbildene inneholdt flere store utfordringer for en lokaliseringsmetode — slik som ikke-sammenhengende fotavtrykk og varierende signalstyrke.

Bernsens metode med postprosessering ga gode resultater for alle bildene i testsettet. Metoden er derfor svært lovende for å lokalisere sonarfotavtrykk. Ved bruk av denne metoden kan fremtidens havkart trolig dannes enklere og nøyaktigere.

Tillegg A

Eksperimentenes utførelse

Jeg ønsket å teste ut et relativt stort antall metoder med mange forskjellige parameterverdier. Derfor utviklet jeg et system for å administrere datamengden og eksperimentene. Tersklingsmetodene jeg ønsket å teste var allerede implementert i bildebehandlingspakken XITE (Bøe et al. 1998, Bøe 1998*b*, Bøe 1998*a*).

Små shell-script ble skrevet for hver metode. Disse shell-scriptene var uavhengige av valgte parameterverdier. Parameterverdiene til metodene ble i stedet beregnet ut fra katalognavnet resultatfilen skulle legges i. Hvis katalognavnet var «bernsen/_10_15_3» ble Bernsens metode benyttet med terskel 10 og vindusstørrelse (15×3).

Likevel var det nødvendig å kjøre et stort antall kommandoer for å generere resultatene. For å lette arbeidet med å kontrollere hvilke kommandoer som skulle utføres, benyttet jeg GNU make (Stallman & McGrath 1998) for å kontrollere kjøringen. Reglene for kompilasjon og avhengigheter kunne imidlertid ikke formuleres enkelt ved hjelp av den innebygde mønstergjenkjenningensmekanismen i GNU make. Jeg skrev derfor programvare for å automatisk generere makefiler. Dette var svært arbeidsbesparende, samlet størrelse på makefilene var over 10 megabyte.

A.1 Parallell kommandoutføring

Svært mange kommandoer skulle utføres som del av eksperimentene. Hvis alle kommandoene skulle kjøres på en og samme maskin, ville dette ta uforholdsmessig lang tid.

GNU make er imidlertid bygget slik at brukeren kan skifte ut kommandotolkeren, altså programmet som kjører kommandoer. Videre er det mulig å be GNU make eksekvere N kommandoer samtidig. Å eksekvere flere kommandoer samtidig har imidlertid liten hensikt når maskinen bare har en prosessor. Det finnes imidlertid allerede programvare («rsh») som kan utføre kommandoer på en *annen* maskin.

Jeg forsøkte derfor å bygge en ny kommandotolker, «dish» («Distributed Shell»), som delegerte kommandoutføringen til et cluster av maskiner i instituttets nettverk. Dette fungerte relativt godt for et begrenset antall maskiner, til tross for en primitiv implementasjon i «bash». En fordel med denne måten å parallelisere arbeidet var at det var trivielt å balansere belastningen på hver maskin. «dish» sørget for at det kjørte maksimalt en kommando på hver maskin. Dette ga en god utnyttelse av maskinressurserne.

Ved hjelp av søkemaskiner fant jeg imidlertid en kraftigere implementasjon av samme idé (Zimmermann 1997), implementert via PVM (Parallell

Virtual Machine) (Geist et al. 1994). PVM er en teknologi for å få mange heterogene maskiner til å arbeide parallelt med et beregningsproblem. Dette programmet, «ppmake», gjorde det mulig å øke clusterstørrelsen vesentlig.

Jeg benyttet derfor et cluster bestående av to SPARCstation 10 kloner og 40-50 SUN Ultra 1 og Ultra 5 maskiner. Dette ga svært stor hastighetsøkning. Riktignok virket ikke hastighetsøkningen til å være lineært proporsjonal med antall maskiner, men det er trolig at nettverksbåndbredde kan ha vært en forsinkende faktor. Dessuten skalerer en master-slave arkitektur (som denne løsningen benytter) dårlig når det blir mange slaver.

Bibliografi

- Abulateb, A. S. (1989), 'Automatic thresholding of gray-level pictures using two-dimensional entropy', *Computer Vision, Graphics and Image Processing* **47**, 22-32.
- Adobe (1998), Adobe Photoshop. <http://www.adobe.com/prodindex/photoshop/main.html>.
- Augustin, J. M., Dugelay, S., Lurton, X. & Voisset, M. (1997), Applications of an image segmentation technique to multibeam echo-sounder data, in 'Proc. Oceans'97', Vol. 2, IEEE, IEEE, pp. 1365-1369.
- Augustin, J. M., Edy, C., Savoye, B. & Le Drezen, E. (1994), Sonar mosaic computation from multibeam echo sounder, in 'Proc. Oceans'94', IEEE, pp. 433-438.
- Augustin, J. M., Le Suave, R., Lurton, X., Voisset, M., Dugelay, S. & Satria, C. (1996), 'Contribution of the multibeam acoustic imagery to the exploration of the sea-bottom', *Mar. Geophys. Res.* **18**, 459-486.
- Bernsen, J. (1986), Dynamic thresholding of grey-level images, in 'Proc. 8th Int. Conf. Pattern Recogn.', pp. 1251-1255. Paris.
- Bøe, S., Lønnestad, T. & Milvang, O. (1998), XITE (X-based Image Processing Tools and Environment) — User's Manual, Technical Report 56, Image Processing Laboratory, Department of Informatics, University of Oslo. <http://www.ifi.uio.no/~blab/Software/Xite/UsersManual/>.
- Bøe, S. (1998a), XITE (X-based Image Processing Tools and Environment) — Programmer's Manual, Technical Report 92, Image Processing Laboratory, Department of Informatics, University of Oslo. <http://www.ifi.uio.no/~blab/Software/Xite/ProgrammersManual/>.
- Bøe, S. (1998b), XITE (X-based Image Processing Tools and Environment) — System Administrator's Manual, Technical Report 91, Image Processing Laboratory, Department of Informatics, University of Oslo. <http://www.ifi.uio.no/~blab/Software/Xite/SysAdmManual/>.
- Carmichael, D. (1998), Image processing techniques for the analysis of sidescan sonar survey data, in 'IEE Colloquium on Underwater Applications of Image Processing', IEE, IEE, pp. 3/1-6.
- Churchill, F. E., Ogar, G. W. & Thompson, B. J. (1981), 'The correction of I and Q errors in a coherent processor', *IEEE Trans. on Aerospace and Electronic Systems* **17**(1), 131-137.
- de Moustier, C. & Alexandrou, D. (1991), 'Angular dependence of 12 kHz seafloor acoustic backscatter', *J. Acoust. Soc. Am.* **90**(1), 522-531.

- de Moustier, C. (1993), Signal processing for swath bathymetry and concurrent seafloor acoustic imaging, *in* J. M. F. Moura & I. M. G. Lourtie, eds, 'Acoustic Signal Processing for Ocean Exploration', Kluwer Academic Publishing, pp. 329-354.
- Eikvil, L., Taxt, T. & Moen, K. (1991), A fast adaptive method for binarization of document images, *in* 'Proceedings of the First International Conference on Document Analysis and Recognition', Saint-Malo, France, pp. 435-443.
- Farr, H. K. (1980), 'Multibeam bathymetric sonar: Sea beam and hydrochart', *Marine Geodesy* 4(2), 77-93.
- Fusillo, L. J. & Satriano, J. H. (1993), Processing techniques for SASS multi-beam sidescan imagery, *in* 'Proc. Oceans'93', Vol. 2, IEEE, pp. 443-448.
- Geist, A., Beguelin, A., Dongarra, J., Jiang, W., Manchek, R. & Sunderam, V. (1994), *PVM: Parallel Virtual Machine*, MIT Press.
- Gonzalez, R. C. & Woods, R. E. (1992), *Digital Image Processing*, Addison-Wesley.
- Hammerstad, E., Pøhner, F., Parthiot, F. & Bennett, J. (1991), Field testing of a new deep water multibeam echo sounder, *in* 'Proc. OCEANS'91', Vol. 2, pp. 743-749.
- Johnson, D. H. & Dudgeon, D. E. (1993), *Array Signal Processing — Concepts and Techniques*, Prentice-Hall Signal Processing Series, Prentice-Hall.
- Kapur, J. N., Sahoo, P. K. & Wong, A. K. C. (1985), 'A new method for gray-level picture thresholding using the entropy of the histogram', *Computer Vision, Graphics and Image Processing* 29, 273-285.
- Kittler, J. & Illingworth, J. (1986), 'Minimum error thresholding', *Pattern Recognition* 19(1), 41-46.
- Kongsberg SIMRAD (1994), *SIMRAD EM3000 — New Portable High Resolution Multibeam Echo Sounder — Product specification*.
- Lee, S. U., Chung, S. Y. & Park, R. H. (1990), 'A comparative performance study of several global thresholding techniques for segmentation', *Computer Vision, Graphics and Image Processing* 52, 171-190.
- Lingsch, S. C. & Robinson, C. S. (1992), 'Acoustic imagery using a multibeam bathymetric system', *Mar. Geod.* 15, 81-95.
- Mitchell, N. C. (1996), 'Processing and analysis of SIMRAD multibeam sonar data', *Marine Geophysical Researches* 18, 729-739.
- Niblack, W. (1986), *An Introduction to Digital Image Processing*, Prentice Hall, pp. 115-116.
- Oppenheim, A. V. & Schaffer, R. W. (1989), *Discrete-time signal processing*, Prentice-Hall.
- Otsu, N. (1979), 'A threshold selection method from gray-level histograms', *IEEE transactions on Systems, Man and Cybernetics T-SMC* 9(1), 62-66.
- Pal, N. R. & Pal, S. K. (1993), 'A review on image segmentation techniques', *Pattern Recognition* 26(9), 1277-1294.
- Pratt, W. K. (1991), *Digital Image Processing*, second edn, John Wiley & Sons, New York.

- Rantzartzis, D., Moustier, C. & Alexandrou, D. (1993), Applications of high-resolution beamforming to multibeam swath bathymetry, *in* 'Proc. Oceans'93', Vol. 2, IEEE, pp. 77-82.
- Satriano, J. H. & Smith, L. C. (1991), Signal processing for wide swath bathymetric sonar, *in* 'Proc. Oceans'91', IEEE, pp. 558-561.
- Somers, M. L. (1993), Sonar imaging of the seabed: Techniques, performance, applications, *in* J. M. F. Moura & I. M. G. Lourtie, eds, 'Acoustic Signal Processing for Ocean Exploration', Kluwer Academic Publishing, pp. 355-369.
- Stallman, R. M. & McGrath, R. (1998), GNU Make Version 3.77. http://www.delorie.com/gnu/docs/make/make_toc.html.
- Steinhardt, A. O. & van Veen, B. D. (1989), 'Adaptive beamforming', *International Journal of Adaptive Control and Signal Processing* 3, 253-81.
- Strasters, K. & Gerbrands, J. J. (1991), 'Three-dimensional image segmentation using a split, merger and group approach', *Pattern Recognition Letters* 12, 307-325.
- Taxt, T., Flynn, P. J. & Jain, A. K. (1989), 'Segmentation of document images', *IEEE Trans. Pattern Analysis and Machine Intelligence* 11(12), 1322-1329.
- Thomas, J. B. (1969), *An Introduction to Statistical Communication Theory*, John Wiley & Sons.
- Trier, Ø. D. & Taxt, T. (1995a), 'Improvement of 'integrated function algorithm' for binarization of document images', *Pattern Recognition Letters* 16(3), 277-283.
- Trier, Ø. D. & Taxt, T. (1995b), 'Evaluation of binarization methods for document images', *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(3), 312-315.
- Trier, Ø. D. (1995), binarize - binarize a grey-scale image. http://www.ifi.uio.no/~blab/Software/Xite/ReferenceManual/binarize_1.ht%ml.
- Weiss, M. A. (1995), *Data structures and algorithm analysis*, Benjamin/Cummings.
- White, J. M. & Rohrer, G. D. (1983), 'Image thresholding for optical character recognition and other applications requiring character image extraction', *IBM Journal of Research and Development* 27(4), 400-411.
- Yang, L., Taxt, T. & Albregtsen, F. (1997), Bottom detection for multibeam sonars with active contours, *in* 'Proc. Oceans'97', IEEE, pp. 943-950.
- Yang, L. & Taxt, T. (1997), Multibeam sonar bottom detection using multiple subarrays, *in* 'Proc. Oceans'97', IEEE, pp. 932-938.
- Yang, L. (1998a), personal communication.
- Yang, L. (1998b), A new multibeam bottom detection system, Technical note, Kongsberg SIMRAD AS.
- Yanowitz, S. & Bruckstein, A. (1989), 'A new method for image segmentation', *Comput. Vision Graphics Image Proc.* 46(1), 82-95.
- Yasnoff, W. A., Mui, J. K. & Bacus, J. W. (1977), 'Error measures for scene segmentation', *Pattern Recognition* 9, 217-231.

- Zhang, Y. J. (1997), 'Evaluation and comparison of different segmentation algorithms', *Pattern Recognition Letters* **18**, 963-974.
- Zhang, Y. (1996), 'A survey on evaluation methods for image segmentation', *Pattern Recognition* **29**(8), 1335-1346.
- Zimmermann, S. (1997), PVM Parallel Make (ppmake).
<http://www3.informatik.tu-muenchen.de/~zimmerms/ppmake/>.