



Faculteit Toegepaste Wetenschappen
Vakgroep Telecommunicatie en Informatieverwerking
Voorzitter: Prof. dr. ir. H. Bruneel

Vergelijkend onderzoek van textuurparameters in echografiebeelden

door Bruno Huysmans

Promotor: prof. dr. ir. W. Philips
Thesisbegeleider: lic. E. Vansteenkiste

Afstudeerwerk ingediend tot het behalen van de graad van burgerlijk
elektrotechnisch ingenieur, optie: communicatietechniek

Academiejaar 2002-2003

Toelating tot bruikleen

De auteur geeft de toelating dit afstudeerwerk voor consultatie beschikbaar te stellen en delen van het afstudeerwerk te kopiëren voor persoonlijk gebruik. Elk ander gebruik valt onder de beperkingen van het auteursrecht, in het bijzonder met betrekking tot de verplichting de bron uitdrukkelijk te vermelden bij het aanhalen van resultaten uit dit afstudeerwerk.

Bruno Huysmans

2 juni 2003

Dankwoord

Het schrijven van een thesis is een werk van lange adem. Ik ben dan ook heel dankbaar voor de vele hulp en steun die ik gekregen heb tijdens het schrijven.

In het bijzonder zou ik mijn promotor, prof. dr. ir. W. Philips, willen danken voor de mogelijkheid dit onderzoek uit te voeren. Mijn oprechte dank gaat ook uit naar Ewout Vansteenkiste, in wiens handen de dagelijkse begeleiding lag.

Tevens wil ik ook mijn ouders bedanken omdat zij het mogelijk maakten mijn studies en deze thesis tot een goed einde te brengen.

Tenslotte zou ik nog Maaïke Lootens en Micheline Vertriest willen danken voor het nalezen en corrigeren van de kladversies.

Vergelijkend onderzoek van textuurparameters in echografiebeelden

door

Bruno Huysmans

Afstudeerwerk ingediend tot het behalen van de graad van burgerlijk elektrotechnisch ingenieur, optie: communicatietechniek

Academiejaar 2002-2003

Universiteit Gent

Faculteit Toegepaste Wetenschappen

Promotor: prof. dr. ir. W. Philips

Samenvatting

In de geneeskunde wordt dikwijls gebruik gemaakt van echografiebeelden om diagnoses te stellen. Door de lage kwaliteit van deze beelden is het voor de arts vaak moeilijk de beelden correct te interpreteren. Een hulpmiddel voor de arts is de analyse van de beelden aan de hand van textuurparameters. In deze thesis wordt leukomalacie bestudeerd, een hersenziekte die vooral bij premature baby's voorkomt. De textuurparameters berekend uit de coöccurrentiematrix worden onderzocht. Ook worden enkele andere textuurmethodes bestudeerd: som- en verschilhistogrammen, de gewijzigde coöccurrentiematrix, Amelungs statistische parameters, de looplengtematrix, Laws' filter maskers en gaborfilters.

Tenslotte wordt ook nog de invloed onderzocht van de grootte van het venster waarin we de textuur bestuderen, de invloed van de hoek en de invloed van het compensatie-algoritme. Het compensatie-algoritme is een algoritme dat de invloed van de instellingen van het echografietoestel op de beelden minimaliseert.

Trefwoorden: echografie, textuur, leukomalacie.

Inhoudsopgave

1	Inleiding	1
2	Echografiebeelden	3
2.1	Algemene principes van echografie	3
2.1.1	Geluidsgolven	3
2.1.2	Reflectie en interferentie	3
2.1.3	Voor- en nadelen	4
2.2	Echografie en leukomalacie	5
2.2.1	Structuur van de hersenen	5
2.2.2	Periventriculaire leukomalacie	5
2.2.3	Echografiebeelden van leukomalaciepatiënten	6
2.2.4	Het compensatie-algoritme	8
3	Textuurparameters	12
3.1	Wat is textuur?	12
3.2	Textuur bij echografiebeelden	13
3.2.1	Hoekafhankelijkheid	13
3.2.2	Grootte van de ROI	14
3.2.3	Gebruikte methodes om de textuur te beschrijven	15
4	Principale Componenten Analyse	16
4.1	Inleiding	16
4.2	Theoretische achtergrond	16
4.3	Toepassing voor dimensionaliteitsreductie	19
5	Classificatiemethode	21
5.1	Inleiding	21
5.2	Bayesiaanse classificatie	22
5.2.1	De meerdimensionale normale verdeling	22
5.2.2	Bayesiaanse classificatie	23

5.2.3	Beslissingsgrenzen	24
6	Coöccurrentiematrix	26
6.1	Definitie	26
6.2	Voorbeeld	27
6.3	Eerste orde statistieken	28
6.4	Textuurparameters - Tweede orde statistieken	28
6.5	Toepassing op echografiebeelden	32
6.5.1	Implementatie	32
6.5.2	Overzicht	32
6.6	Gecompenseerde beelden	33
6.6.1	Coronale beelden	33
6.6.2	Sagittale beelden	37
6.6.3	Combinatie van coronale en sagittale beelden	38
6.7	Niet gecompenseerde beelden	41
6.7.1	Coronale beelden	41
6.7.2	Sagittale beelden	43
6.7.3	Combinatie van coronale en sagittale beelden	44
6.8	Invloed van de afstand d	45
6.8.1	Coronale beelden	45
6.8.2	Sagittale beelden	46
6.9	Overzicht van de bekomen resultaten	46
7	Varianten op de coöccurrentiematrix	54
7.1	Uners som- en verschilhistogrammen	54
7.1.1	Definitie	54
7.1.2	Voorbeeld	55
7.1.3	Parameters	55
7.1.4	Toepassing op echografiebeelden	56
7.1.5	Coronale beelden	57
7.1.6	Sagittale beelden	58
7.1.7	Coronale en sagittale beelden	59
7.1.8	Som- en verschilhistogrammen: algemeen besluit	60
7.2	Coöccurrentie met het gemiddelde van de omliggende pixels	64
7.2.1	Definitie en voorbeeld	64
7.2.2	Parameters	64
7.2.3	Toepassing op echografiebeelden: implementatie	65

7.2.4	Coronale beelden	65
7.2.5	Sagittale beelden	67
7.2.6	Coronale en sagittale beelden	67
7.2.7	Gewijzigde coöccurrentiematrix: algemeen besluit	68
7.3	Amelungs statistische parameters	70
7.3.1	Definitie	70
7.3.2	Parameters	70
7.3.3	Coronale beelden	70
7.3.4	Sagittale beelden	72
7.3.5	Coronale en sagittale beelden	72
7.3.6	Statistische parameters: algemeen besluit	73
8	Galloway's looplengtematrix	75
8.1	Definitie	75
8.2	Voorbeeld	75
8.3	Parameters	76
8.4	Toepassing op echografiebeelden	77
8.4.1	Laagdoorlaatfilteren	77
8.4.2	Grijswaardereductie	77
8.4.3	Coronale beelden	79
8.4.4	Sagittale beelden	81
8.4.5	Coronale en sagittale beelden	81
8.4.6	Besluit	82
9	Laws' filter maskers	88
9.1	Definitie	88
9.2	Toepassing op echografiebeelden	89
9.2.1	Implementatie	89
9.2.2	Resultaten	90
10	Gaborfilters	91
10.1	Definitie	91
10.1.1	Het Gabor prefilter	91
10.1.2	Het Gaussiaans postfilter	93
10.2	Toepassing	93
11	Besluit	97

Hoofdstuk 1

Inleiding

In de medische wereld wordt tegenwoordig vaak gebruik gemaakt van echografiebeelden. Iedereen kent natuurlijk de beelden van een foetus, maar echografieën hebben nog vele andere toepassingen. Men gebruikt ze namelijk bij het onderzoek van de meeste zachte weefsels, zoals de nieren, het hart, de lever, . . . In deze scriptie wordt *periventriculaire leukomalacie* bestudeerd, een hersenziekte die vooral voorkomt bij premature baby's met een laag geboortegewicht (minder dan 1600 gram).

Een nadeel van echografiebeelden is dat het beeld, door het typische specklepatroon, vaak van een lage kwaliteit is en veel ruis bevat. De diagnose is dan ook afhankelijk van de interpretatie van de arts, en bijgevolg subjectief. We kunnen de artsen bij hun diagnose bijstaan door een kwantitatieve analyse van de beelden uit te voeren, op basis van de *textuurparameters*. Tot nu toe gebruikte men steeds de methode van de *coöccurrentiematrix* om deze parameters te berekenen, er werd echter nog geen kwantitatief onderzoek uitgevoerd met deze methode. Een eerste doel van deze scriptie is dan ook om dit kwantitatief onderzoek uit te voeren.

Naast de methode van de coöccurrentiematrix bestaan nog talloze andere methoden om de textuur te beschrijven. Een tweede doel is om enkele van deze andere methoden te implementeren en de performantie ervan te vergelijken. Deze performantie wordt vooral bepaald door de mate waarin ziek en gezond hersenweefsel kunnen worden onderscheiden.

Een bijkomend probleem, naast de slechte kwaliteit van de beelden, is dat de parameters die bekomen worden afhankelijk zijn van de instellingen van de scanner. Deze instellingen kunnen verschillen van beeld tot beeld omdat de artsen die instellingen kiezen die voor hen het duidelijkste beeld opleveren. Voor het toestel waarmee onze beelden gemaakt zijn, is echter een algoritme ontwikkeld dat de invloed van de instellingen ongedaan maakt. Een derde doel is om te kijken welke parameters in welke mate afhankelijk zijn van deze instellingen.

Tenslotte wordt nog nagegaan of er betrouwbare classificatie van ziek en gezond weefsel

mogelijk is, en welke daarbij de optimale parameters zijn.

Voor deze classificatie tussen ziek en gezond weefsel hebben we gebruik gemaakt van een dertigtal beelden waarvan op voorhand geweten was tot welke categorie ze behoorden. Deze voorkennis maakt dat we van een *gesuperviseerd* probleem kunnen spreken. Dit is uiteraard veel eenvoudiger dan het *niet gesuperviseerd* probleem, waarbij op voorhand niet geweten is welke beelden ziek of gezond weefsel voorstellen.

In hoofdstukken 2 en 3 wordt kort ingegaan op de begrippen *echografie* en *textuur*. In hoofdstuk 4 wordt *Principale Componenten Analyse (PCA)* besproken, een techniek die we hier gebruikt hebben om de dimensie van een parameter ruimte te reduceren. Hoofdstuk 5 beschrijft het algoritme op basis waarvan ziek en gezond weefsel geïdentificeerd worden.

In hoofdstuk 6 wordt de coöccurrentiematrix uitgebreid besproken. De verschillende parameters berekend met deze methode worden kwantitatief onderzocht, en daaruit worden dan de optimale parametercombinaties bepaald (met het oog op classificatie tussen ziek en gezond weefsel).

In hoofdstuk 7 worden enkele varianten van de coöccurrentiematrix besproken (Uners som- en verschilhistogrammen, een gewijzigde coöccurrentiematrix en een statistische aanpak). Hoofdstuk 8 behandelt de techniek van de looplengtematrix. In hoofdstukken 9 en 10 wordt ingegaan op twee filtertechnieken: Laws' filter maskers en het gebruik van gaborfilters. De besluiten tenslotte komen in hoofdstuk 11 te staan.

Hoofdstuk 2

Echografiebeelden

2.1 Algemene principes van echografie

2.1.1 Geluidsgolven

Bij echografie wordt gebruik gemaakt van hoogfrequente geluidsgolven met een frequentie van 2 tot 20 MHz. Deze frequenties zijn voor de mens onhoorbaar. Door het gebruik van geluidsgolven spreekt men ook wel van *ultrasound*. Hoe lager de frequentie van de geluidsgolven, hoe dieper ze het lichaam binnendringen. Wil men weefsels bestuderen die dicht tegen de oppervlakte liggen dan gebruikt men dus golven met hogere frequentie, hoe dieper het weefsel, hoe lager de gebruikte frequentie.

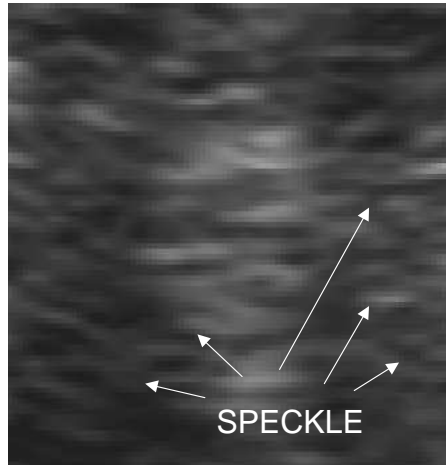
De golven worden in het lichaam gestuurd door een *transducer* op de huid te plaatsen. Een transducer is een klein toestel waarin zich piëzo-elektrische kristallen bevinden. Deze kristallen worden aan het trillen gebracht door elektrische signalen, waardoor golven gegenereerd worden. In de praktijk brengt men tussen de huid en de transducer nog een gel aan, omdat deze gel voor de golven een veel betere geleider is dan lucht.

2.1.2 Reflectie en interferentie

De geluidsgolven worden gereflecteerd als ze op de grens komen van twee gebieden met een verschillende *akoustische impedantie*. De akoustische impedantie van een weefsel is het product van de dichtheid van het weefsel en de snelheid waarmee de golf zich voortplant in dat weefsel. De snelheid mag constant ondersteld worden, zodat de impedantie enkel afhangt van de dichtheid.

De gereflecteerde golven komen terug samen aan de transducer en daar zullen ze door de piëzo-elektrische kristallen omgezet worden in een elektrisch signaal. Dit elektrisch signaal tenslotte wordt vertaald in een grijswaardebeeld.

Aan de probe van de transducer zal er steeds interferentie optreden. Bij constructieve interferentie hebben we een buik, wat zich op het beeld uit als een witte vlek. Bij



Figuur 2.1: Het typische specklepatroon van echografiebeelden.

destructieve interferentie hebben we een knoop, wat resulteert in een donkere vlek. Dit typische patroon van heldere en meer donkere vlekken wordt *speckle* genoemd, zie Fig. 2.1. Een nadeel van echografiebeelden is dat we vaak niet weten welke speckles nu relevante informatie bevatten en welke speckles ruis zijn.

Het specklepatroon hangt niet enkel af van het beschouwde weefsel, maar is ook in hoge mate afhankelijk van de apparatuur, vooral van de transducer. Het specklepatroon beschrijft dus niet alleen het weefsel, maar eerder het geheel van weefsel en gebruikte apparatuur.

2.1.3 Voor- en nadelen

De belangrijkste voordelen van echografie zijn:

- Goedkoop en snel.
- Niet invasief, dit wil zeggen dat de apparatuur niet in het lichaam komt. Er zijn achteraf ook geen nadelige gevolgen voor het lichaam¹.
- Realtime evaluatie mogelijk.
- Geen metaal-artefacten: in tegenstelling tot bijvoorbeeld een CT-scan of MR-scan zullen metalen in het lichaam bij echografie geen verstoring van het beeld opleveren.
- Draagbare apparatuur: de patiënt moet niet verplaatst worden, het onderzoek kan aan het ziekbed gebeuren.

¹Er wordt algemeen aangenomen dat de techniek volledig veilig is voor de mens, maar dit is tot op heden nog steeds niet aangetoond.

Er zijn natuurlijk ook nadelen:

- Reflectie door bot, lucht en metaal. Beenderen en lucht zullen de golven bijna volledig weerkaatsen zodat men geen delen kan bestuderen die zich achter beenderen bevinden. De reflectie door lucht wordt grotendeels verholpen door het gebruik van de geleidende gel.
- Door de ruis is het dikwijls moeilijk voor de arts om de beelden te interpreteren, de diagnose is dan ook subjectief. De ruis is ook nadelig voor verdere beeldverwerking.
- Het beeld is afhankelijk van de instellingen van het apparaat, waardoor kwantitatieve analyse gehinderd kan worden.

2.2 Echografie en leukomalacie

2.2.1 Structuur van de hersenen

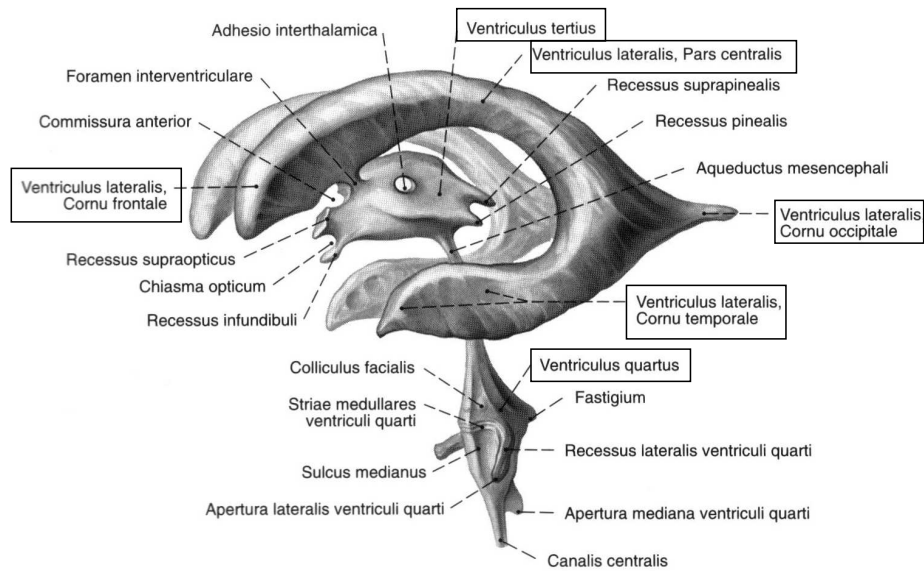
De hersenen van de mens bestaan uit drie delen: de grote hersenen, de kleine hersenen en de hersenstam. In de hersenen bevinden zich vier *hersenventrikels* of *hersenkamers*, zie Fig. 2.2. Twee zijventrikels (*ventriculus lateralis*) bevinden zich in de halfronden van de eindhersenen, een derde ventrikel (*ventriculus tertius*) verbindt deze halfronden en het vierde ventrikel (*ventriculus quartus*) bevindt zich in de hersenstam en loopt door naar het ruggenmerg. De ventrikels zijn gevuld met vocht, het *cerebrospinaal vocht* of *hersenvocht*. De functies van het hersenvocht zijn de volgende:

- Beschermen van de hersenen tegen schokken en stoten (een soort bumpereffect).
- Vervoeren van hormonen uit bepaalde hersendelen.
- Afvoeren van afvalproducten uit de hersenen.

Het hersenvocht wordt gevormd door ingestulpte bloedvatnetwerken, de *plexus choroideus*, in de twee zijventrikels. Het zijn vooral de twee zijventrikels die in dit onderzoek van belang zijn. Bij leukomalacie, de bestudeerde ziekte, wordt namelijk het weefsel rond deze ventrikels beschadigd.

2.2.2 Periventriculaire leukomalacie

Periventriculaire leukomalacie (PVL), kortweg leukomalacie genoemd, is een ziekte waarbij de witte hersenstof nabij (Grieks: *peri*) de ventrikels beschadigd wordt. Men denkt dat leukomalacie veroorzaakt wordt door een tekort aan bloed in de hersenen tijdens de zwangerschap, bij de geboorte, of tijdens de eerste dagen na de geboorte. Aangezien bij



Figuur 2.2: De vier hersenventrikels [11].

leukomalacie delen van de hersenen beschadigd worden, of zelfs geheel afsterven, hebben baby's met de ziekte een heel grote kans om mentaal en lichamelijk gehandicapt te zijn.

2.2.3 Echografiebeelden van leukomalaciepatienten

Alhoewel de geluidsgolven quasi volledig weerkaatst worden door beenderen, kunnen we voor hersenonderzoek van pasgeboren baby's toch echografie toepassen. De reden hiervoor is dat de schedel nog niet volledig toegesloten is, de nog aanwezige opening wordt *fontanel* genoemd.

De gebruikte dataset

We beschikken over twee soorten beelden, namelijk *sagittale* en *coronale* doorsneden van de hersenen, zie Fig. 2.3. Bij de studie van de textuurparameters moeten we rekening houden met deze twee soorten doorsneden, omdat de texturen verschillend kunnen zijn. Er zijn 35 beelden, de verdeling tussen ziek-gezond en sagittaal-coronaal is in Tabel 2.1 terug te vinden.

In Fig. 2.4 zien we links een echografiebeeld van de hersenen van een gezonde baby, rechts van een zieke baby. Wanneer het hersenweefsel rond de ventrikels afsterft wordt het zachter, en dit leidt tot de witte vlekken op de beelden van zieke patiënten.

Tabel 2.1: Verdeling van de beelden tussen ziek-gezond en sagittaal-coronaal

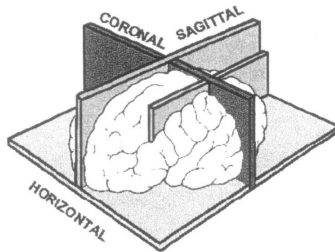
	ziek	gezond
sagittaal	5	9
coronaal	14	7

Tabel 2.2: Aantal testgebieden per beeld voor de verschillende categorieën beelden: ziek-gezond en sagittaal-coronaal

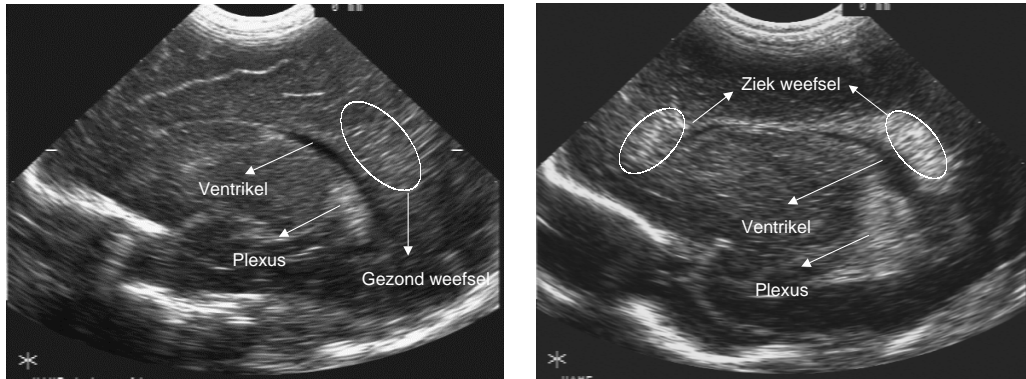
	ziek	gezond
sagittaal	4	2
coronaal	2	3

Beperkingen van de dataset

- Omdat we te weinig beelden hebben om statistisch significante data te verkrijgen zullen we uit elk beeld meerdere testgebieden halen, zie Tabel 2.2. Deze testgebieden zijn vierkantjes van 30x30, 40x40, 50x50 of 60x60 pixels. Het testgebied dat bestudeerd wordt zullen we vanaf nu *ROI (Region Of Interest)* noemen. Het gebruik van meerdere ROI's per beeld verhoogt het aantal waarnemingen, maar uiteraard is dit geen ideale oplossing. Het nadeel is dat de gebieden meestal sterk zullen overlappen, en daardoor grotendeels dezelfde informatie zullen bevatten. Hierdoor zullen de waarden van de parameters bekomen uit eenzelfde beeld dicht bij elkaar liggen, er zal dus clustering optreden. Deze clustering wordt geïllustreerd in Fig. 2.5, waar de gemiddelde grijswaarde en de variantie op de grijswaarde voor de zieke en gezonde sagittale beelden tegenover elkaar uitgezet zijn. We zien duidelijk dat bij de zieke beelden vier ROI's per beeld werden genomen, en bij de gezonde beelden twee ROI's



Figuur 2.3: Sagittale, coronale en horizontale doorsnedes van de hersenen.



Figuur 2.4: Links zien we de sagittale doorsnede van een beeld met gezond weefsel. Rechts zien we ziek weefsel. Het zieke weefsel uit zich op de figuur door heldere, witte vlekken. De witte vlek rechts, onder het ventrikel, is geen ziek weefsel maar stelt de *plexus choroideus* voor.

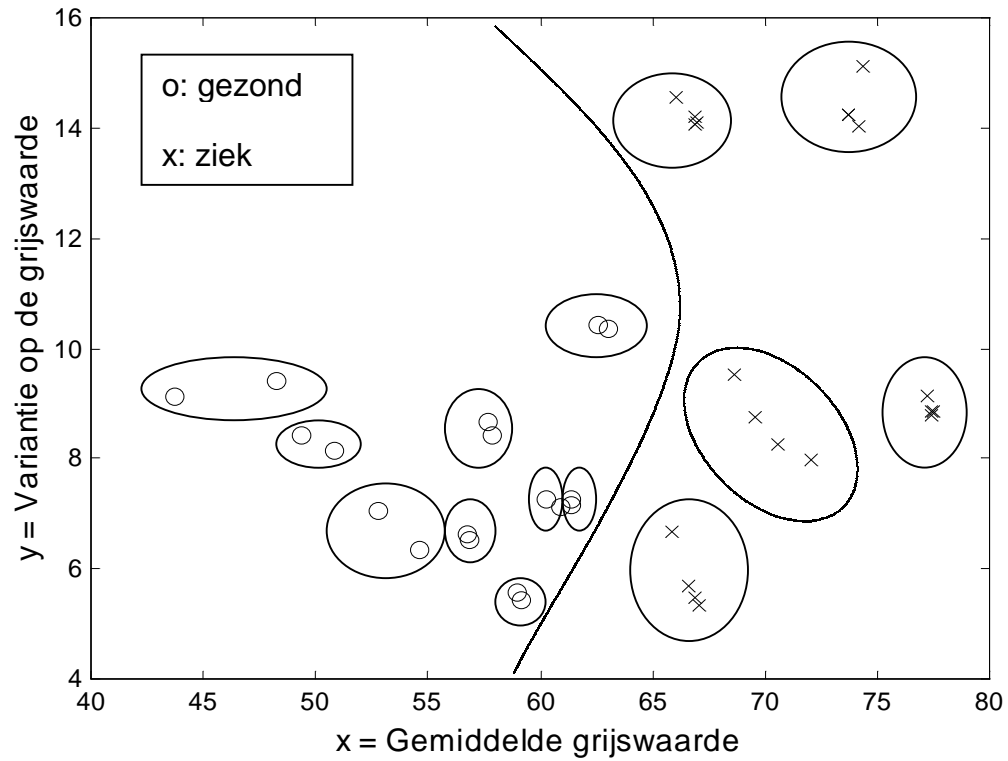
per beeld.

- Een tweede beperking van onze dataset, naast het tekort aan beelden, is dat de invalshoek van de golven niet overal dezelfde is. Daardoor zal bij sommige beelden het ventrikel er niet opstaan, bij andere wel, zie Fig. 2.6. Aangezien verschillende stukken weefsel verschillende textuureigenschappen kunnen hebben is dit ook een beperking voor onze testen. Ideaal zou zijn mocht de invalshoek overal dezelfde zijn. In de praktijk is dit echter moeilijk te realiseren, omdat we met (kwetsbare) premature baby's te maken hebben, en de plaatsing van de probe dus niet eindeloos kan gewijzigd worden.
- Een derde tekortkoming van de gebruikte beelden is dat ze niet allemaal in hetzelfde stadium van de ziekte werden genomen. Sommige beelden werden genomen bij het prille begin van de ziekte, andere beelden werden dan weer genomen bij patiënten met vergevorderde leukomalacie.

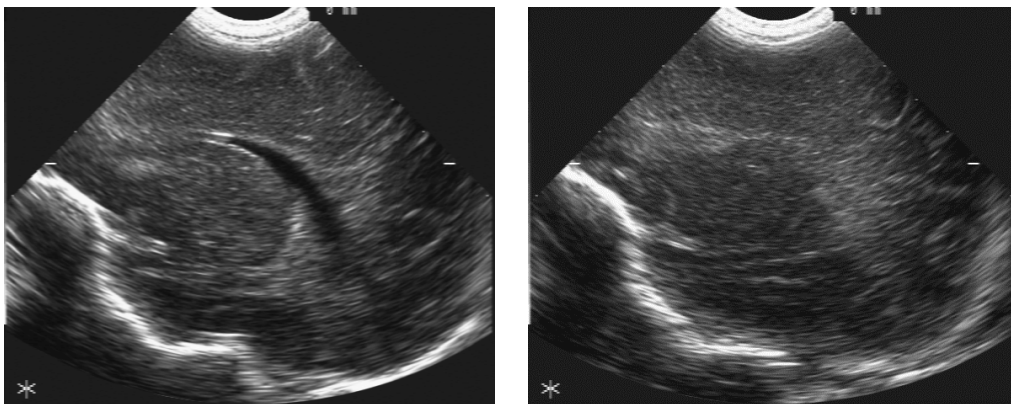
2.2.4 Het compensatie-algoritme

Wanneer de artsen echografiebeelden maken, kiezen ze de instellingen van het toestel zodanig, dat het beeld voor hen zo duidelijk mogelijk wordt. Instellingen die gewijzigd kunnen worden zijn bijvoorbeeld de amplitude en de versterking van de geluidsgolf. Afhankelijk van de situatie, de arts, ... zullen de beelden met andere instellingen genomen zijn. Dit kan uiteraard een probleem vormen voor kwantitatieve analyse van de beelden. Voor het toestel waarmee onze beelden gemaakt zijn, is echter een compensatie-algoritme ontwikkeld dat de invloed van de instellingen tracht te minimaliseren. Een voorbeeld van weefsel

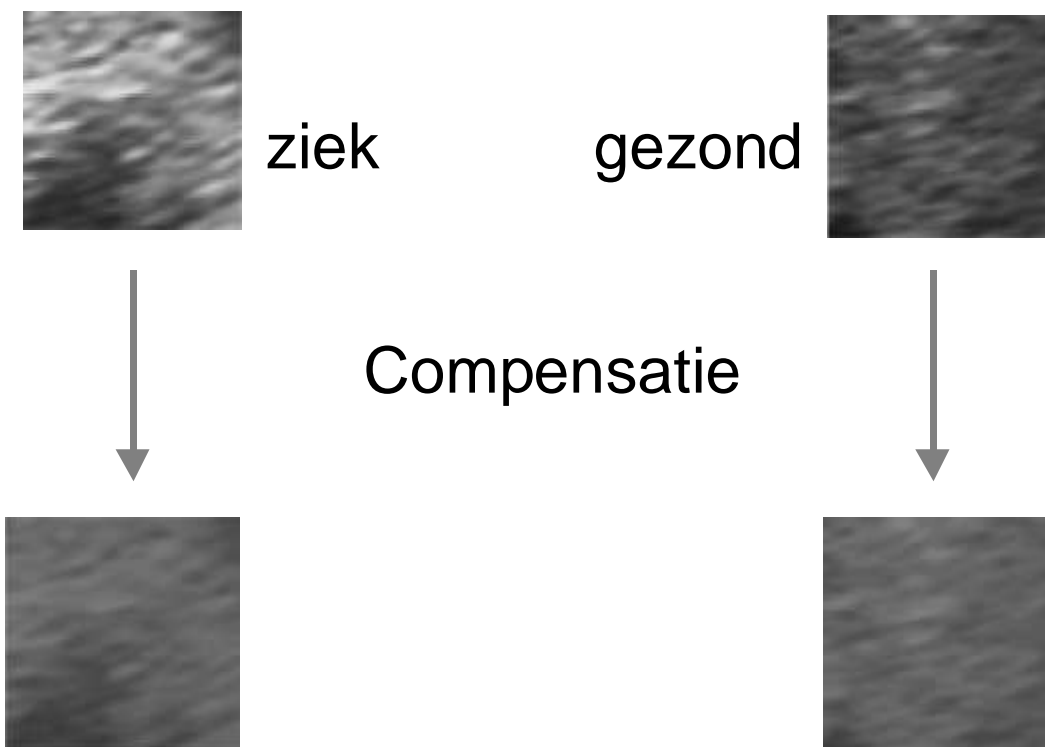
voor en na compensatie is in Fig. 2.7 gegeven. We zullen al onze experimenten zowel met gecompenseerde als met niet gecompenseerde beelden uitvoeren.



Figuur 2.5: De gemiddelde grijswaarde en de variantie op de grijswaarde voor de zieke en gezonde sagittale beelden. Uit de figuur blijkt duidelijk de clustering van parameters bekomen uit hetzelfde beeld. We zien dat er uit elk ziek beeld vier, en uit elk gezond beeld twee, ROI's werden gehaald.



Figuur 2.6: Invloed van de invalshoek. Op de linkse figuur is het ventrikel duidelijk te zien, op de rechtse figuur staat het ventrikel er niet op, wel het weefsel dat zich voor het ventrikel bevindt.



Figuur 2.7: Ziek en gezond weefsel voor en na compensatie.

Hoofdstuk 3

Textuurparameters

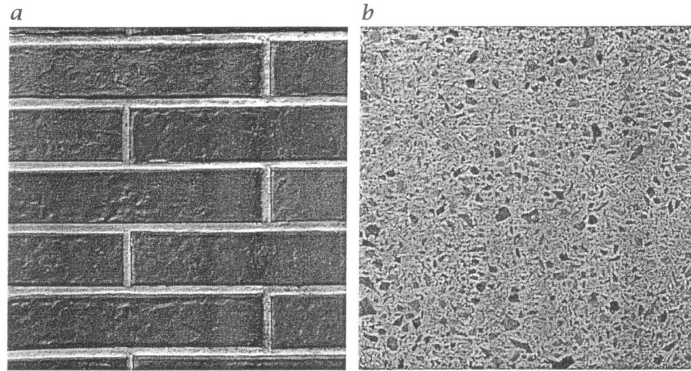
3.1 Wat is textuur?

In de patroonherkenning is het belangrijk dat men verschillende objecten (patronen) kan onderscheiden. Dit wordt gedaan door gebruik te maken van kleur, vorm en andere eigenschappen. Eén van die andere eigenschappen is *textuur*. Het is moeilijk om een exacte definitie van textuur te geven. Met textuur bedoelt men meestal een ‘*ruimtelijke verdeling van kleuren of intensiteiten, met een zekere regelmaat*’. Het begrip textuur is slechts zinvol voor een verzameling punten, men kan dus niet spreken over de ‘textuur van een punt’.

Textuur kan *deterministisch* en *periodiek* zijn, zoals de textuur van een schaakbord. Textuur kan ook een *stochastisch karakter* hebben. Een voorbeeld hiervan is de textuur van een grasveld: er is geen periodiek patroon, maar sommige statistische eigenschappen zullen hetzelfde zijn (bijvoorbeeld het gemiddeld aantal grassprietten per m^2 , de vorm van de grassprietten, ...). Twee andere voorbeelden van textuur worden getoond in Fig. 3.1.

Een belangrijk aspect is de *schaal* waarop men textuur beschouwt. We zullen dit duidelijk maken aan de hand van een voorbeeld: wanneer we een voldoende groot venster nemen, zal de textuur van een bakstenen muur bepaald worden door het regelmatige patroon dat bestaat uit de afwisseling van bakstenen en voegen. Wanneer het venster kleiner wordt dan een baksteen zullen we de eerder stochastische textuur van de baksteen zelf waarnemen.

De mens kan verschillende texturen onderscheiden. Dit doet hij op basis van verschillen in oriëntatie, schaal, kleur, ... Stochastische texturen kunnen onderscheiden worden door verschillen in de gemiddelde grijswaarde en in de variantie op de grijswaarde. Wanneer we met de computer texturen willen onderscheiden maken we gebruik van verschillende *textuurparameters* (features), die de textuur op een zo intelligent mogelijke manier trachten te beschrijven. Afhankelijk van de textuur zullen sommige parameters goede of slechte resultaten geven.



Figuur 3.1: Twee voorbeelden van textuur. Links zien we een muur met deterministische, periodieke textuur. Rechts is een voorbeeld van een stochastische textuur gegeven.

3.2 Textuur bij echografiebeelden

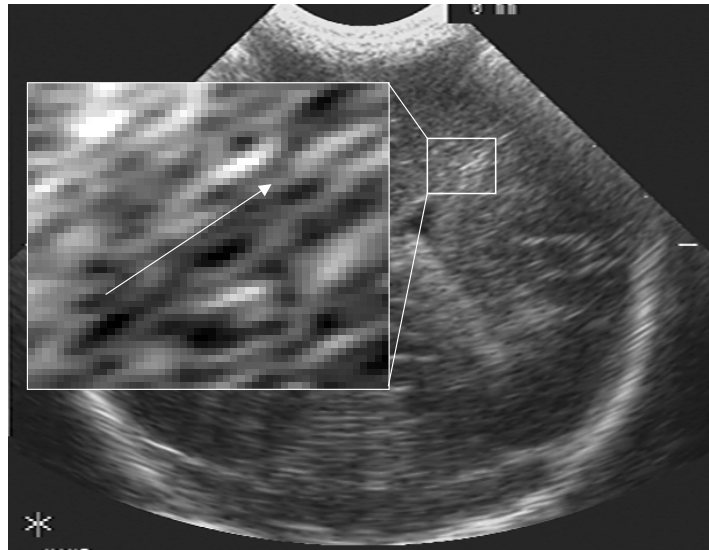
Zoals hoger al gezegd zullen sommige parameters de ene textuur goed beschrijven, terwijl ze voor een andere textuur minder geschikt zullen zijn. Het voornaamste doel van dit werk is om uit te zoeken welke parameters bij de echografiebeelden van de hersenen geschikt zijn om ziek en gezond weefsel te onderscheiden.

De (stochastische) textuur van echografiebeelden wordt gevormd door de afwisseling van heldere en donkere vlekjes (speckles). In hoofdstuk 2 werd reeds vermeld dat het specklepatroon niet enkel afhangt van het beschouwde weefsel, maar ook in hoge mate afhankelijk is van de apparatuur, vooral van de transducer. Het specklepatroon beschrijft dus niet alleen het weefsel, maar eerder het geheel van weefsel en gebruikte apparatuur. Een bepaalde textuurparameter kan het weefsel goed beschrijven voor het ene apparaat, terwijl dezelfde parameter voor hetzelfde weefsel maar andere apparatuur slechte resultaten kan geven [4].

3.2.1 Hoekafhankelijkheid

Coronale beelden Aan de rechterkant van het beeld heeft de textuur van het weefsel rond het ventrikel een duidelijke oriëntatie in een richting van 45 graden. Dit wordt geïllustreerd in Fig. 3.2. Aan de linkerkant is er een oriëntatie in een richting van 135 graden. We hebben de ROI steeds in het rechtergebied gekozen, zodat we voor alle ROI's dezelfde voorkeursrichting hebben.

Sagittale beelden Net als bij de coronale beelden loopt de textuur aan de rechterkant van het beeld weer in een richting van 45 graden. Naarmate we meer naar de linkerkant van het beeld gaan, nadert de voorkeursrichting naar 135 graden, zie Fig. 3.3. De verschillende



Figuur 3.2: Deze figuur toont de voorkeursrichting van 45 graden in het rechterdeel van de coronale beelden.

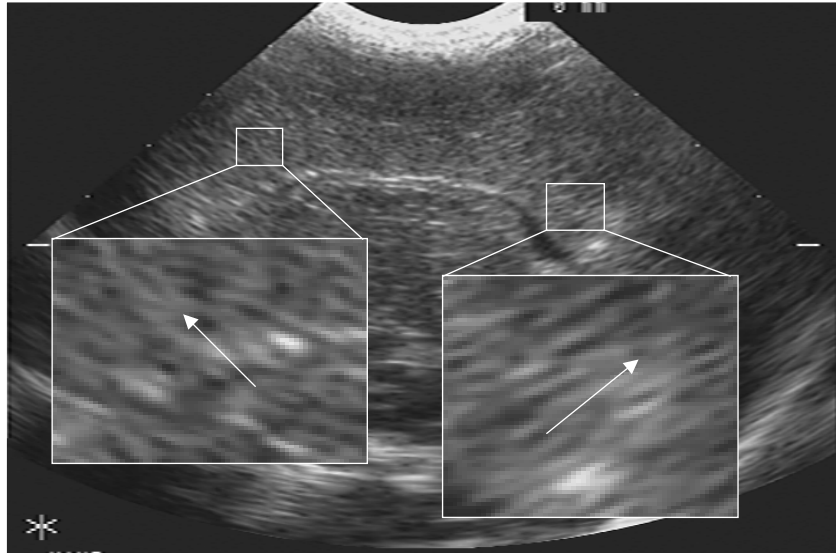
ROI's werden zowel uit het linker- als rechtergedeelte van de beelden gehaald, omdat er te weinig sagittale beelden zijn om steeds enkel rechts of enkel links ROI's te kiezen en ook nog voldoende ROI's te bekommen.

3.2.2 Grootte van de ROI

We hebben daarnet al vermeld dat de grootte van het venster waarin de textuur bekeken wordt belangrijk is. In de volgende hoofdstukken gebruiken we telkens vensters van 30x30, 40x40, 50x50 en 60x60 pixels. We onderzoeken de invloed van de venstergrootte op de textuurparameters. Er werden geen vensters kleiner dan 30x30 pixels gebruikt, omdat de schaal waarop de textuur beschreven wordt dan te klein zou geworden zijn, en de parameters teveel afhankelijk zouden worden van welk stukje weefsel nu precies geselecteerd werd. Vensters groter dan 60x60 pixels werden ook niet beschouwd, omdat het dan moeilijk wordt ervoor te zorgen dat het venster enkel relevant weefsel bevat.

De voordelen en nadelen van een *kleiner* ten opzichte van een *groter* venster zijn:

- Hoe groter het venster, hoe langer de benodigde tijd om de textuurparameters te berekenen.
- Een kleiner venster zal meestal niet buiten het bestudeerde weefsel vallen, en dus enkel relevant weefsel bevatten. Grote vensters kunnen soms een beetje buiten het weefsel vallen, en dit heeft uiteraard een nadelige invloed op de significantie van de berekende textuurparameters.



Figuur 3.3: Deze figuur toont de voorkeursrichting bij de sagittale beelden. Deze voorkeursrichting varieert van 45 graden (rechts) tot 135 graden (links).

- Bij een kleiner venster zullen ruis en foute informatie meer doorwegen, aangezien er minder wordt uitgemiddeld. Deze nadelen worden verminderd bij een groot venster.

3.2.3 Gebruikte methodes om de textuur te beschrijven

De gebruikte technieken om de textuur te bestuderen vallen uiteen in twee categorieën:

1. Tellen van grijswaarden en combinaties van grijswaarden De eerste soort technieken telt het voorkomen van bepaalde grijswaarden of bepaalde grijswaardecombinaties. Tot deze methodes behoren de gewone en gewijzigde coöccurrentiematrix, de som- en verschilhistogrammen, de statistische parameters en de looplengtematrix.

2. Filteren De tweede soort technieken maakt gebruik van filters. We zullen de beelden door een of meerdere filters sturen en uit de gefilterde beelden berekenen we dan de parameters. De gebruikte filtertechnieken zijn: Laws' filter maskers en gaborfilters.

Hoofdstuk 4

Principale Componenten Analyse

4.1 Inleiding

Bij de meeste methodes worden meerdere textuurparameters berekend. Sommige hiervan zullen beter in staat zijn om ziek en gezond weefsel te onderscheiden dan andere. Er kan ook steeds correlatie tussen de verschillende parameters zijn. Hoe sterker de correlatie tussen twee parameters, hoe meer de informatie overlapt. Het is duidelijk dat voor classificatie niet alle parameters nodig zijn, en dat de informatie dus compacter kan worden voorgesteld. De parameters die totaal geen onderscheid maken tussen ziek en gezond weefsel kunnen we laten vallen, sterk gecorreleerde parameters kunnen we vervangen door nieuwe parameters die de informatie als het ware ‘samenvatten’.

De *Principale Componenten Analyse (PCA)* [1] wordt gebruikt om een dergelijke parameter ruimte met hoge dimensie te reduceren tot een lagere dimensie. Hierbij moet natuurlijk zoveel mogelijk informatie behouden worden, en moet dus een optimale transformatie gezocht worden. Een belangrijke eigenschap van de PCA is dat de getransformeerde parameters ongecorreleerd zullen zijn.

We zullen ons hier beperken tot een herleiding naar twee of drie dimensies, omdat het resultaat dan eenvoudig grafisch voor te stellen is.

4.2 Theoretische achtergrond

We beschouwen de $N \times 1$ vector \mathbf{v} , waarvan elk element een parameter voorstelt,

$$\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \quad (4.1)$$

Er zijn dus N parameters. De vector μ bevat de gemiddelde waarden van de parameters,

$$\mu = E[\mathbf{v}] = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{pmatrix} \quad (4.2)$$

De reductie van N naar M parameters komt neer op een *projectie* van de N -dimensionale ruimte op een M -dimensionale subruimte. Deze projectie moet zo gekozen worden dat classificatie zo optimaal mogelijk kan gebeuren. Hiervoor maken we gebruik van de *covariantiematrix* \mathbf{K} ,

$$\mathbf{K} = E[(\mathbf{v} - \mu)(\mathbf{v} - \mu)^T] \quad (4.3)$$

Een eigenschap van de covariantiematrix is dat de eigenvector die correspondeert met de grootste eigenwaarde de richting aangeeft waarin de grootste variantie optreedt. De grootste eigenwaarde geeft dan de variantie in die richting. De eigenvector horend bij de tweede grootste eigenwaarde geeft dan de richting met de grootste variantie loodrecht op de eerste eigenvector, die ervoor zorgt dat de parameters volgens de nieuwe richtingen gedecorreleerd zijn. Ook nu weer wordt de variantie gegeven door de eigenwaarde. Het voorgaande kunnen we veralgemenen tot de k -de eigenvector, die dan de richting met de grootste variantie geeft loodrecht op de eerste $(k - 1)$ eigenvectoren, zo dat de parameters ongecorrleerd zullen zijn. De eigenwaarde is dan de variantie in die richting. De eigenwaarden van de covariantiematrix worden ook wel de *principale componenten* genoemd, dit verklaart meteen ook de naam van deze techniek.

De covariantiematrix \mathbf{K} kan ontbonden worden in het product

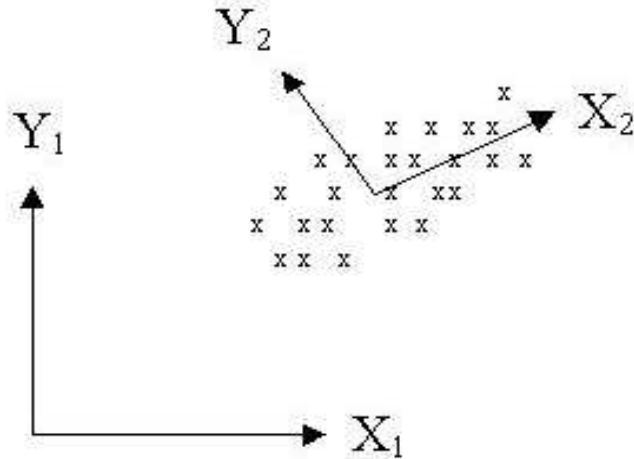
$$\mathbf{K} = \mathbf{BDB}^T \quad (4.4)$$

van matrices \mathbf{B} en \mathbf{D} , waarbij

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{pmatrix} \quad (4.5)$$

de diagonaalmatrix is met de eigenwaarden, geordend van groot naar klein ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$) en

$$\mathbf{B} = (\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_N) \quad (4.6)$$



Figuur 4.1: Deze figuur toont de toepassing van PCA in twee dimensies. Het oorspronkelijke assenstelsel (X_1, Y_1) wordt eerst naar het zwaartepunt van de wolk verschoven, vervolgens is er een rotatie die ervoor zorgt dat de nieuwe assen (X_2, Y_2) volgens de richtingen met de grootste varianties komen te liggen en de parameters ongecorreleerd worden.

de matrix is met de eigenvectoren b_i horend bij λ_i . Deze eigenvectoren zijn orthonormaal, waaruit volgt dat:

$$\mathbf{B}^T \mathbf{B} = \mathbf{B} \mathbf{B}^T = \mathbf{I} \quad (4.7)$$

We gaan nu over naar een nieuw coördinatenstelsel door eerst een translatie over vector μ uit te voeren, gevolgd door een rotatie volgens de eigenvectoren van de covariantiematrix \mathbf{K} . De getransformeerde vector \mathbf{w} wordt dan berekend uit vector \mathbf{v} door

$$\mathbf{w} = \mathbf{B}^T (\mathbf{v} - \mu) \quad (4.8)$$

Men kan aantonen dat de getransformeerde parametervector \mathbf{w} een verwachtingswaarde nul heeft en de nieuwe parameters ongecorreleerd zijn, met andere woorden: $E[\mathbf{w}] = 0$ en $\text{cov}[\mathbf{w}, \mathbf{w}] = \mathbf{D}$.

Een voorbeeld van de PCA transformatie in twee dimensies is te vinden in Fig. 4.1. Bij de PCA transformatie wordt de dimensie niet gereduceerd. Veel voordeel levert de transformatie hier dan ook niet op. Men behoudt dezelfde parameterruimte, maar bekijkt ze enkel vanuit een ander standpunt. De dimensie kan worden gereduceerd van N tot M door in plaats van de eigenvectormatrix \mathbf{B} een getrunceerde matrix \mathbf{B}_M te gebruiken. Deze matrix bestaat uit de M dominante eigenvectoren, dus de eigenvectoren horend bij de M hoogste eigenvectoren.

$$\mathbf{B}_M = (\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_M) \quad M \leq N \quad (4.9)$$

De transformatie met de getrunceerde matrix wordt nu

$$\mathbf{w}_M = \mathbf{B}_M^T(\mathbf{v} - \mu) \quad (4.10)$$

Men kan aantonen dat ook nu $E[\mathbf{w}_M] = 0$ en $\text{cov}[\mathbf{w}_M, \mathbf{w}_M] = \mathbf{D}_M$, met \mathbf{D}_M de diagonaalmatrix met de M grootste eigenwaarden.

4.3 Toepassing voor dimensionaliteitsreductie

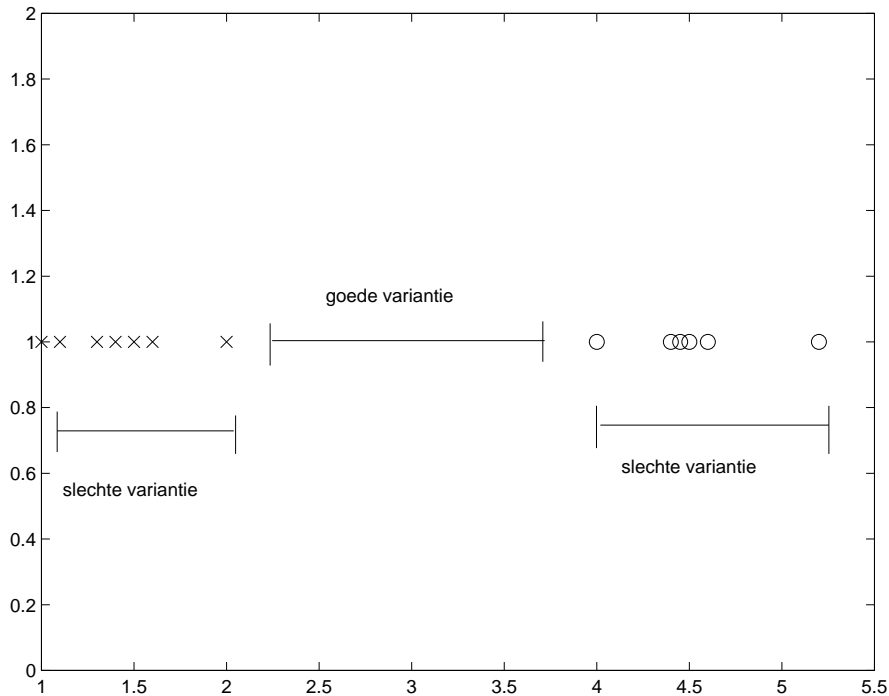
Twee of drie dimensies? We gebruiken PCA om de parameterruimte naar twee of drie dimensies te reduceren. We kiezen voor twee of drie dimensies omdat de resultaten dan eenvoudig grafisch voorstelbaar zijn. De eigenwaarden geven de variantie in een bepaalde richting aan. Stel dat er twee grote eigenwaarden zijn, en de rest verwaarloosbaar klein is, bijvoorbeeld: $[5 \ 4 \ 0.002 \ 0.0015]$. Herleiding naar drie dimensies heeft dan weinig voordeel ten opzichte van herleiding naar twee dimensies, omdat de variantie in de richting van de derde dimensie reeds heel klein is. Wanneer de eigenwaarden $[5 \ 4 \ 4 \ 4 \ 0.002]$ zijn kan het nuttig zijn om in vier dimensies te werken, omdat de variantie in de richting van de vierde dimensie nog steeds groot is.

Normalisatie Vooraleer PCA wordt toegepast moeten de textuurparameters *genormaliseerd* worden:

$$x_n = \frac{x - \mu}{\sigma} \quad (4.11)$$

Hierbij is μ de gemiddelde waarde en σ de standaardafwijking van de parameter. De reden voor deze normalisatie is duidelijk: parameters met grotere waarden zullen automatisch een grotere variantie vertonen, en dus zullen de assen eerder in de richting van die parameters gekozen worden. Na de normalisatie hebben alle parameters een verwachtingswaarde nul en een variantie één.

De normalisatie wordt eerst uitgevoerd op enkel de zieke of enkel de gezonde beelden. Vervolgens wordt **dezelfde** transformatie uitgevoerd op de andere set beelden. Wanneer we bijvoorbeeld eerst de zieke beelden normaliseren zullen we de gezonde beelden normaliseren met de μ en σ waarden van de zieke beelden. De reden hiervoor is de volgende: we splitsen de variantie in een ‘goede’ en een ‘slechte’ variantie, zie Fig. 4.2. De goede variantie wordt veroorzaakt door het feit dat gezond en ziek ver uit elkaar liggen. De slechte variantie wordt veroorzaakt door de spreiding op de zieke en gezonde waarnemingen zelf. Wanneer we eerst de zieke beelden normaliseren zullen we enkel de slechte variantie reduceren tot 1. Vervolgens gebruiken we dezelfde transformatie voor de gezonde beelden. Wanneer we de zieke en gezonde beelden tegelijk zouden normaliseren, is de variantie veel



Figuur 4.2: Voorbeeld van 'goede' en 'slechte' variantie.

groter, en zal de goede variantie 'bestraft' worden, zodat de afstand tussen ziek en gezond verkleint.

Parameters waarvoor ziek en gezond door elkaar liggen zullen na de normalisatie een variantie van ongeveer 1 hebben. De parameters waarbij ziek en gezond gescheiden zijn zullen een hogere variantie hebben, en dus bevoordeeld worden bij PCA.

De eigenlijke PCA Nadat alle parameters genormaliseerd zijn kunnen we PCA toepassen. We passen dit toe op de volledige testruimte, dus op de zieke en gezonde beelden tegelijkertijd. Parameters met een slecht onderscheidend vermogen, waar gezond en ziek dus door elkaar liggen, zullen een variantie van ongeveer 1 hebben. De parameters met een hoog onderscheidend vermogen zullen een grotere variantie hebben, en dus bij PCA bevoordeeld worden.

We merken tenslotte nog op dat de PCA techniek niet specifiek voor classificatie is ontworpen. Enkel de eerste en tweede orde momenten van het proces \mathbf{v} worden in rekening gebracht, er wordt nergens met classificatie rekening gehouden. Daarom geeft PCA niet altijd de beste resultaten voor classificatie. In veel gevallen geeft een combinatie van twee parameters uit een parameterset betere resultaten dan PCA reductie van de parameterset naar twee parameters.

Hoofdstuk 5

Classificatiemethode

5.1 Inleiding

Eens we de textuurparameters berekend hebben moeten we ze natuurlijk gebruiken om een onderscheid te maken tussen ziek en gezond weefsel. Een eenvoudige manier van classificatie is *grafisch*: we tekenen gewoon 1, 2 of 3 parameters en kijken hoe goed ziek en gezond weefsel gescheiden worden. De eenvoud wordt echter betaald met enkele nadelen:

- Vooreerst is het niet mogelijk om in meer dan drie dimensies te werken. De combinatie van meer dan drie parameters kan dus niet getest worden.
- Een ander nadeel is het feit dat het moeilijk is om de optimale grenzen tussen ziek en gezond te bekomen.
- Een derde nadeel is dat onderzoek van de parameters volgens deze methode heel arbeidsintensief is. Het is bijna onbegonnen werk om voor alle parameters alle mogelijke combinaties te bekijken.

Omwille van de nadelen verbonden met grafische classificatie zullen we onze toevlucht nemen tot een iets geavanceerder classificatie-algoritme, meer bepaald *Bayesiaanse classificatie* [2].

De Bayesiaanse classifier schat de waarschijnlijkheid dat een bepaalde waarneming tot een bepaalde klasse behoort. In ons geval zijn er maar twee klassen, namelijk ‘ziek’ en ‘gezond’. Het gaat hier dus om *harde classificatie* (hard labeling), omdat een sample steeds 100% tot de ene (gezond) of de andere (ziek) klasse behoort. Naast harde classificatie bestaat er ook nog *zachte classificatie* (soft labeling, fuzzy labeling). Hierbij wordt nog rekening gehouden met een tussenzone, een klasse ‘onbepaald’, waar men dus niet kan opmaken of de beelden ziek of gezond weefsel voorstellen. In dit onderzoek werd echter geen zachte classificatie gebruikt.

Wat evenmin gedaan werd is het gebruik van een techniek voor ‘outlier-removal’, dit is het verwerpen van samples die te afwijkende resultaten geven. In de praktijk wordt een dergelijke techniek dikwijls gebruikt, maar wegens het kleine aantal samples werd dit hier niet toegepast.

Bij Bayesiaanse classificatie wordt aangenomen dat de algemene vorm van de waarschijnlijkheidsdichtheidsfunctie van de parameters voor elke klasse (ziek en gezond) gekend is. De waarden van de parameters van de dichtheidsfunctie, zoals het gemiddelde of de variantie, zijn echter niet op voorhand geweten. Voor we de dichtheidsfuncties kunnen gebruiken moeten we deze parameters dus eerst schatten. Omdat we te weinig testbeelden hebben om de exacte dichtheidsfunctie te bepalen zullen we onderstellen dat de parameters steeds *normaal verdeeld* zijn, zodat we enkel de gemiddelden en de varianties moeten schatten. Als er meerdere parameters beschouwd worden zullen we een *meerdimensionale normale verdeling* onderstellen.

De onbekende parameters worden geschat uit de *trainingsset*. Dit zijn in ons geval de beelden waarvan we al weten of ze van een ziek of gezond kind komen. Om de efficiëntie van de classificatie te bepalen moeten we gebruik maken van een *testset*. In ons geval zullen de testset en de trainingsset uit dezelfde beelden bestaan. We zullen steeds 1 beeld als testbeeld gebruiken en de andere beelden als trainingsbeelden (*leave-one-out principe*). Vervolgens wordt nagegaan of het ene beeld juist wordt geclassificeerd. Wanneer we alle beelden eens als testbeeld gebruiken kunnen we het foutpercentage van de classificatie bepalen:

$$\text{Classificatiefout [\%]} = 100 \times \frac{\text{aantal verkeerd geclassificeerde beelden}}{\text{totaal aantal beelden}} \quad (5.1)$$

De eigenlijke validatie van alle resultaten zal echter moeten gebeuren met een *onafhankelijke* nieuwe testset, die momenteel echter onbeschikbaar is.

5.2 Bayesiaanse classificatie

5.2.1 De meerdimensionale normale verdeling

De gekende 1-dimensionale normale verdeling met gemiddelde waarde μ en standaardafwijking σ ,

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (5.2)$$

kan worden uitgebreid naar meerdere dimensies, we spreken dan van een *meerdimensionale normale verdeling*. Hierbij is het handig om gebruik te maken van matrixnotatie. De normaal verdeelde veranderlijken x_1, \dots, x_d worden gegroepeerd in de d-dimensionale

kolomvector \mathbf{x} .

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} \quad (5.3)$$

Verder definiëren we de gemiddelde waarde vector μ als

$$\mu = E[\mathbf{x}] = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_d \end{pmatrix} \quad (5.4)$$

Tenslotte voeren we nog de symmetrische covariantiematrix Σ in:

$$\Sigma = E[(\mathbf{x} - \mu)(\mathbf{x} - \mu)^T] = \begin{pmatrix} \sigma_{11} & \sigma_{21} & \dots & \sigma_{d1} \\ \sigma_{12} & \sigma_{22} & \dots & \sigma_{d2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1d} & \sigma_{2d} & \dots & \sigma_{dd} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1d} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1d} & \sigma_{2d} & \dots & \sigma_d^2 \end{pmatrix} \quad (5.5)$$

Hierbij is σ_{ij} de covariantie tussen x_i en x_j . De d-dimensionale normale dichtheidsfunctie kan nu, in matrixnotatie, worden geschreven als

$$p(\mathbf{x}) = \frac{1}{\sqrt{|\Sigma|} (2\pi)^d} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)} \quad (5.6)$$

5.2.2 Bayesiaanse classificatie

Bij Bayesiaanse classificatie berekent men voor een parametervector \mathbf{x} de conditionele waarschijnlijkheid $P(C_i | \mathbf{x})$ dat hij tot klasse C_i behoort. De klasse waarvoor deze waarschijnlijkheid het grootste is, is dan de klasse waarin \mathbf{x} wordt geclassificeerd. Om de waarschijnlijkheid $P(C_i | \mathbf{x})$ te berekenen maken we gebruik van de *regel van Bayes*:

$$P(C_i | \mathbf{x}) = \frac{P(C_i)P(\mathbf{x} | C_i)}{P(\mathbf{x})} = \frac{P(C_i)P(\mathbf{x} | C_i)}{\sum_{j=1}^k P(C_j)P(\mathbf{x} | C_j)} \quad (5.7)$$

waarbij k het aantal klassen voorstelt, $k = 2$ in ons geval.

De probabiliteiten $P(C_i)$ uit (5.7) worden de *a priori* probabiliteiten genoemd. De reden hiervoor is dat ze de waarschijnlijkheid van de klassen geven voordat we de waarden van de parameters kennen. Deze waarden worden berekend uit de trainingsset, of intuïtief toegekend. In ons geval hebben we de beide waarschijnlijkheden gelijk genomen aan de fractie van de beelden die tot een bepaalde klasse behoren.

De probabiliteiten $P(C_i | \mathbf{x})$ worden de *a posteriori* probabiliteiten genoemd, omdat ze de waarschijnlijkheid van de klassen geven na waarneming van de parameters. Om deze a posteriori waarschijnlijkheden te berekenen moeten we enkel nog de waarden $P(\mathbf{x} | C_i)$ kunnen berekenen. Hiertoe veronderstellen we dat de parameters berekend uit de beelden van het zieke en gezonde weefsel normaal verdeeld zijn. Ze worden dus beschreven door meerdimensionale normale dichtheidsfuncties. De waarschijnlijkheid $P(\mathbf{x} | C_i)$ wordt dan berekend door in (5.6) de juiste waarden van \mathbf{x} , μ en Σ in te vullen. Hierbij worden μ en Σ berekend uit de trainingsset.

5.2.3 Beslissingsgrenzen

Met de methode uit de vorige sectie zijn we in staat om voor alle klassen de waarschijnlijkheid $P(C_i | \mathbf{x})$ te berekenen. De waarneming behoort dan tot de klasse waarvoor deze waarschijnlijkheid het grootst is. Meestal is men niet geïnteresseerd in de precieze waarden van $P(C_i | \mathbf{x})$, maar enkel in de meest waarschijnlijke klasse. Hiervoor voert men *beslissingsgrenzen* in. Deze grenzen verdelen de parameterruimte in verschillende gebieden. Elk gebied correspondeert dan met een bepaalde klasse. In ons geval zijn er maar twee klassen, namelijk ziek en gezond. We beperken ons hier verder ook tot twee parameters, zodat de beslissingsgrenzen eenvoudig grafisch voorgesteld kunnen worden.

De optimale grens tussen de twee klassen correspondeert met de parameterwaarden (x, y) waarvoor:

$$P(C_1 | x, y) = P(C_2 | x, y) \quad (5.8)$$

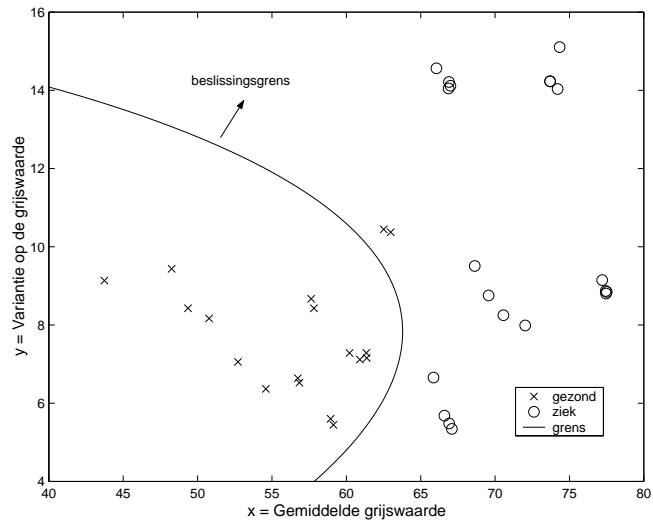
Toepassen van de regel van Bayes hierop leidt tot:

$$P(C_1)P(x, y | C_1) = P(C_2)P(x, y | C_2) \quad (5.9)$$

De a priori probabiliteiten $P(C_i)$ zijn gekend. Als we onderstellen dat de parameters binormaal verdeeld zijn kunnen we de $P(x, y | C_i)$ termen vervangen door de binormale dichtheidsfunctie, berekend uit de testset. De gelijkheid (5.9) wordt dan verder nog vereenvoudigd door van beide leden het natuurlijke logaritme te nemen. Hierdoor is het verband tussen x en y gekend, en kunnen we de grenzen berekenen. Voor normaal verdeelde parameters x en y is de algemene vorm van de beslissingsgrenzen:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (5.10)$$

Afhankelijk van de waarde $\Delta = B^2 - 4AC$ is de beslissingsgrens een ellips ($\Delta < 0$), een parabool ($\Delta = 0$) of een hyperbool ($\Delta > 0$). Een voorbeeld van een beslissingsgrens is gegeven in Fig. 5.1.



Figuur 5.1: De gemiddelde grijswaarde t.o.v. de variantie op de grijswaarde voor sagittale beelden, met een venster van 60x60 pixels. De zwarte lijn is de beslissingsgrens. We zien dat niet alle beelden juist zullen geïdentificeerd worden.

Hoofdstuk 6

Coöccurrentiematrix

De methode van de coöccurrentiematrix, ontwikkeld door Haralick [8] in 1973, is een van de bekendste technieken om textuur te beschrijven. Het is deze methode die tot op heden wordt gebruikt om de textuureigenschappen van echografiebeelden te karakteriseren. Er is een apart hoofdstuk aan deze techniek gewijd omdat veel van de hier geïntroduceerde ideeën later nog terugkomen. Ook wordt hier wat dieper ingegaan op de fysische betekenis van sommige parameters, en de gebruikte methode voor parametersselectie.

6.1 Definitie

De coöccurrentiematrix is een tweedimensionaal histogram dat het voorkomen van pixelparen op een *afstand* d en onder een *hoek* θ beschrijft. In de praktijk zijn de waarden voor θ beperkt tot 0, 45, 90 en 135 graden. Om de matrix op te stellen vertrekken we van een nulmatrix P met dimensie $G \times G$, waarbij G het aantal grijswaarden voorstelt. In de praktijk is G meestal 256, maar soms wordt het aantal grijswaarden gereduceerd om de berekeningen te versnellen.

Vervolgens nemen we de pixel op de eerste rij en de eerste kolom van het beeld, met grijswaarde i . We kijken over een afstand d en hoek θ verder en belanden bij een pixel met grijswaarde j . Een dergelijke overgang van grijswaarde i naar j betekent dat $P(i, j) = P(i, j) + 1$ in de coöccurrentiematrix. Zo overlopen we alle pixels van het beeld.

Wanneer de tweede pixel buiten het beschouwde beeld valt zullen we het pixelpaar gewoon niet in rekening brengen. Wanneer bijvoorbeeld $d = 1$ en $\theta = 0$, is dit het geval bij de laatste pixel van elke rij. Een gevolg hiervan is dat het aantal effectief in rekening gebrachte pixelparen R niet gelijk is aan het aantal pixels N van het beeld.

We kunnen de coöccurrentiematrix P nog *normaliseren* door te delen door het aantal mogelijke pixelparen R :

$$p(i, j) = \frac{P(i, j)}{R} \tag{6.1}$$

Merk op dat R ook gelijk is aan de som van de elementen van de coöccurrentiematrix:

$$R = \sum_{i=1}^G \sum_{j=1}^G P(i, j) \quad (6.2)$$

Verder kunnen we de *hoekafhankelijkheid* wegwerken door de coöccurrentiematrix te berekenen voor 0, 45, 90 en 135 graden, en het gemiddelde te nemen van deze vier matrices.

Tenslotte kunnen we de coöccurrentiematrix nog *symmetrisch* maken door zijn getransponeerde erbij op te tellen:

$$P_S(i, j) = P(i, j) + P^T(i, j) \quad (6.3)$$

Dit komt neer op het optellen van de matrix berekend voor een hoek θ met de matrix berekend voor een hoek $-\theta$.

6.2 Voorbeeld

In Fig. 6.1 wordt een voorbeeld gegeven van een gebied met 4 verschillende grijswaarden. Aangezien er vier grijswaarden zijn, zal de coöccurrentiematrix P uit vier rijen en vier kolommen bestaan. De corresponderende niet genormaliseerde coöccurrentiematrix P , met $d = 1$ en $\theta = 0$, wordt weergegeven in Fig. 6.2.

Om de coöccurrentiematrix te berekenen starten we met de pixel linksboven, met grijswaarde '1'. We vergelijken deze pixel met de pixel die $d = 1$ pixels verder ligt, onder een hoek van 0 graden. Dit is dus de pixel rechts ervan. Deze heeft ook grijswaarde '1'. We hebben een '1-1' overgang. Op rij 1 en kolom 1 van de matrix moet 1 bijgeteld worden en komt er nu dus een 1. De tweede grijswaarde is '1', rechts daarvan hebben we een grijswaarde '2'. Nu wordt op rij 1 en kolom 2 van de matrix 1 opgeteld. Als we het volledige gebied met grijswaarden doorlopen hebben, bekomen we uiteindelijk de bijhorende coöccurrentiematrix P .

1	1	2	2
1	1	2	2
1	3	3	3
3	3	4	4

Figuur 6.1: De grijswaarden van de pixels in een gebied

$$\begin{pmatrix} 2 & 2 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Figuur 6.2: Coöccurrentiematrix P met afstand $d = 1$ en hoek $\theta = 0$

6.3 Eerste orde statistieken

Gemiddelde grijswaarde

De gemiddelde grijswaarde is een *eerste orde statistiek*. De eerste orde statistieken worden rechtstreeks uit het grijswaardehistogram van het beeld berekend, en houden dus geen rekening met de spatiale eigenschappen van het beeld. Het zijn eigenlijk geen echte textuurparameters, maar ze kunnen toch nuttig zijn bij de classificatie. Wanneer N het aantal pixels in de ROI is en $g(x, y)$ de grijswaarde van pixel (x, y) dan hebben we:

$$f_1 = \mu = \frac{1}{N} \sum_{(x,y) \in ROI} g(x, y) \quad (6.4)$$

Variantie van de grijswaarde

Zoals de naam al zegt, geeft deze parameter de variantie op de grijswaarde:

$$f_2 = \sigma^2 = \frac{1}{N-1} \sum_{(x,y) \in ROI} (g(x, y) - \mu)^2 \quad (6.5)$$

Signal Noise Ratio

Deze parameter wordt gedefinieerd als de verhouding van de gemiddelde grijswaarde en de standaardafwijking op de grijswaarde:

$$f_3 = SNR = \frac{\mu}{\sigma} \quad (6.6)$$

6.4 Textuurparameters - Tweede orde statistieken

Eens we de coöccurrentiematrix berekend hebben, kunnen we hieruit verschillende tweede orde textuurparameters berekenen. Tweede orde statistieken beschrijven meestal de ruimtelijke eigenschappen van de textuur, zoals in ons geval de specklegrootte. De parameters worden hieronder opgesomd, en eventueel wat verduidelijkt. Waar mogelijk wordt ook ingegaan op de fysische betekenis van de parameters. Meer informatie over deze fysische betekenis kan gevonden worden in [3]. De parameters worden steeds uit de genormaliseerde coöccurrentiematrix berekend.

Angulair Tweede Moment (ATM)

$$f_4 = \sum_{i=1}^G \sum_{j=1}^G (p(i, j))^2 \quad (6.7)$$

Deze parameter wordt ook wel *uniformiteit*, *energie* of *homogeniteit* genoemd. Hoe meer verschillende grijswaarde-overgangen er zijn, hoe kleiner het ATM zal zijn. Zijn de grijswaarde-overgangen steeds dezelfde (bijvoorbeeld steeds van 90 naar 100), dan neemt het ATM toe. Deze parameter is dus een maat voor de *homogeniteit* van het beeld. Bij een homogeen beeld zal het ATM zijn maximale waarde 1 bereiken (wanneer we met de genormaliseerde coöccurrentiematrix werken uiteraard).

Contrast

$$f_5 = \sum_{i=1}^G \sum_{j=1}^G (i - j)^2 p(i, j) \quad (6.8)$$

Deze parameter¹ penaliseert kleine grijswaarde-overgangen en bevoordeelt overgangen van een kleine naar een grote grijswaarde. Het is dus een maat voor het contrast, zoals wij het kennen uit het dagelijkse leven. De diagonaalelementen van de coöccurrentiematrix, dit zijn overgangen tussen gelijke grijswaarden, tellen niet mee (want $i - j = 0$), overgangen tussen sterk verschillende grijswaarden daarentegen zullen een grote invloed hebben. In een homogeen beeld zal het contrast klein zijn.

Correlatie

$$f_6 = \frac{1}{\sigma_x \sigma_y} \left(\sum_{i=1}^G \sum_{j=1}^G (ij) p(i, j) - \mu_x \mu_y \right) \quad (6.9)$$

Hierbij zijn μ_x , μ_y , σ_x en σ_y de gemiddelde waarden en standaardafwijking van $p_x(i)$ en $p_y(j)$. Hierbij wordt $p_x(i)$ bekomen door de rijen van $p(i, j)$ op te tellen, $p_y(j)$ wordt bekomen door de kolommen op de tellen:

$$p_x(i) = \sum_{j=1}^G p(i, j)$$

$$p_y(j) = \sum_{i=1}^G p(i, j)$$

¹In het oorspronkelijke artikel van Haralick [8] werd deze parameter gedefinieerd als: $\sum_{g=0}^{G-1} g^2 \left[\sum_{i=1}^G \sum_{j=1}^G p(i, j) \right]_{|i-j|=g}$, wij gebruiken echter een eenvoudiger definitie.

De correlatieparameter² is een maat voor de lineaire afhankelijkheid tussen de grijswaarden van naburige pixels.

Variantie

$$f_7 = \sum_{i=1}^G \sum_{j=1}^G (i - \mu)^2 p(i, j) \quad (6.10)$$

De variantie is een maat voor de *heterogeniteit* en is sterk gecorreleerd met eerste orde parameters zoals de variantie van de grijswaarden. De variantie stijgt naarmate de grijswaarden meer afwijken van het gemiddelde. In tegenstelling tot wat men misschien op het eerste zicht zou denken is deze parameter weinig gecorreleerd met het contrast.

Invers Verschilmoment

$$f_8 = \sum_{i=1}^G \sum_{j=1}^G \frac{1}{1 + (i - j)^2} p(i, j) \quad (6.11)$$

Deze parameter is eigenlijk zowat het omgekeerde van contrast. De invloed van hoge grijswaarde-overgangen wordt nu sterk onderdrukt, de invloed van kleine overgangen wordt belangrijk.

Entropie

$$f_9 = - \sum_{i=1}^G \sum_{j=1}^G p(i, j) \log(p(i, j)) \quad (6.12)$$

De entropie is een maat voor de wanorde. Wanneer de grijswaarden sterk verspreid zijn, zullen veel elementen van de coöccurrentiematrix een kleine waarde aannemen en zal de entropie groot zijn. Wanneer we bijvoorbeeld een beeld beschouwen met random grijswaarden (witte ruis), dan benadert het grijswaardehistogram een constante functie en nadert de entropie zijn maximum. Entropie is sterk gecorreleerd met het ATM, maar dan wel invers: wanneer de ene parameter groot is, is de andere klein. Een voordeel van het gebruik van het ATM in plaats van de entropie is dat het ATM genormaliseerd is (tussen 0 en 1).

²Een andere veelgebruikte definitie van de correlatieparameter is: $\frac{1}{\sigma^2} \sum_{i=1}^G \sum_{j=1}^G (i - \mu)(j - \mu)p(i, j)$

Som Gemiddelde

$$f_{10} = \sum_{i=2}^{2G} i p_{x+y}(i) \quad (6.13)$$

waarbij

$$p_{x+y}(g) = \left[\sum_{i=1}^G \sum_{j=1}^G p(i, j) \right]_{i+j=g}$$

Deze parameter wordt ook berekend bij de methode van Unser, op basis van som- en verschilhistogrammen. Daarom hebben we deze parameter niet geïmplementeerd in deze methode. Dezelfde opmerking geldt voor som variantie, som entropie, verschil variantie en verschil entropie.

Som Variantie

$$f_{11} = \sum_{i=2}^{2G} (i - f_{10})^2 p_{x+y}(i) \quad (6.14)$$

Som Entropie

$$f_{12} = - \sum_{i=2}^{2G} p_{x+y}(i) \log(p_{x+y}(i)) \quad (6.15)$$

Verschil Variantie en Verschil Entropie

Deze parameters worden op een analoge manier berekend als *som variantie* en *som entropie*, enkel gebruiken we hier p_{x-y} in plaats van p_{x+y} :

$$p_{x-y}(g) = \left[\sum_{i=1}^G \sum_{j=1}^G p(i, j) \right]_{|i-j|=g}$$

Informatiematen i.v.m. de correlatie

$$f_{13} = \frac{HXY - HXY1}{\max(HX, HY)} \quad (6.16)$$

$$f_{14} = (1 - \exp[-2|HXY2 - HXY|])^{1/2} \quad (6.17)$$

Hierbij is HXY de entropie, dus parameter f_9 , zijn HX en HY de entropie van respectievelijk p_x en p_y en worden $HXY1$ en $HXY2$ gegeven door:

$$HXY1 = - \sum_{i=1}^G \sum_{j=1}^G p(i, j) \log(p_x(i)p_y(j))$$

$$HXY2 = - \sum_{i=1}^G \sum_{j=1}^G p_x(i)p_y(j)\log(p_x(i)p_y(j))$$

6.5 Toepassing op echografiebeelden

6.5.1 Implementatie

Om de coöccurrentiematrix en de bijhorende parameters te berekenen werd uitgegaan van een bestaande implementatie [6], waaraan volgende parameters werden toegevoegd: f_7 , f_{13} en f_{14} . De parameters f_{10} , f_{11} en f_{12} werden, zoals reeds vermeld, niet berekend, aangezien ze in Hoofdstuk 7.1 op een efficiëntere manier worden bekomen. Alle parameters werden berekend uit de *genormaliseerde, symmetrische* coöccurrentiematrix.

6.5.2 Overzicht

In de volgende paragrafen bespreken we:

- De optimale textuurparameters berekend uit de coöccurrentiematrix, met betrekking tot classificatie tussen ziek en gezond weefsel. We doen dit zowel voor sagittale en coronale beelden afzonderlijk, als voor sagittale en coronale beelden samen.
- De invloed van de grootte van de ROI.
- De invloed van de hoek waarvoor we de coöccurrentiematrix hebben berekend.
- De invloed van het compensatie-algoritme.
- De invloed van de afstand d waarvoor we de coöccurrentiematrix hebben berekend.

Eerst worden de gecompenseerde beelden beschouwd. We maken hierbij een onderscheid tussen coronale beelden, sagittale beelden en de combinatie van beide. Daarna worden de experimenten herhaald voor niet gecompenseerde beelden. Aanvankelijk beschouwen we enkel een afstand $d = 1$, omdat dit in de literatuur als optimale waarde wordt omschreven. In de laatste paragraaf bespreken we dan nog kort de invloed van de afstand d .

De performantie van de parameters werd eerst voor elke parameter afzonderlijk getest, door de classificatiefout na een Bayesiaanse classificatie te berekenen. Hieruit kunnen we ongeveer de helft van de parameters, ongeacht de grootte van de ROI en de hoek, al verwerpen voor classificatie.

6.6 Gecompenseerde beelden

In deze paragraaf worden de *gecompenseerde* beelden besproken. Het compensatie-algoritme zorgt ervoor dat de invloed van de instellingen van het echografietoestel geminimaliseerd wordt.

6.6.1 Coronale beelden

Verworpen parameters De elf parameters zijn uiteraard niet allemaal geschikt om de zieke en gezonde beelden te onderscheiden. Om het kaf van het koren te scheiden hebben we eerst de efficiëntie van elke parameter afzonderlijk getest. Dit hebben we gedaan door met elke parameter een Bayesiaanse classificatie uit te voeren, en dit voor alle venstergroottes (30x30, 40x40, 50x50 en 60x60 pixels) en alle hoeken (0, 45, 90 en 135 graden en uitgemiddeld over de hoeken). We behouden in eerste instantie enkel de parameters die tijdens deze classificatie minstens éénmaal een classificatiefout van minder dan 30% opleverden. We hebben voor een waarde van 30% gekozen omdat dit empirisch goede resultaten opleverde. De acht parameters die we na deze eerste selectie overhouden zijn:

- f_1 =Gemiddelde grijswaarde
- f_2 =Variantie op de grijswaarde
- f_4 =Angulair tweede moment (ATM)
- f_6 =Correlatie
- f_7 =Variantie
- f_9 =Entropie
- f_{13} =Informatiemaat i.v.m. correlatie 1 (IC_1)
- f_{14} =Informatiemaat i.v.m. correlatie 2 (IC_2)

De correlatie tussen de entropie en het ATM is (in absolute waarde) steeds groter dan 0.98, ongeacht de venstergrootte of de hoek. Beide parameters bevatten dus zo goed als identieke informatie, en één van de twee kunnen we dan ook laten vallen. We behouden het ATM, omdat deze parameter al genormaliseerd is. Eenzelfde opmerking geldt voor de variantie en de gemiddelde grijswaarde. Hier behouden we de gemiddelde grijswaarde. Er zijn dus nog zes parameters over.

We moeten natuurlijk zeker zijn dat geen goede parameters onterecht overboord werden gegooid. Sommige parameters kunnen namelijk heel slechte resultaten geven wanneer

men ze alleen beschouwt, maar in combinatie met andere parameters juist heel nuttig zijn. Om dergelijke parameters op te sporen hebben we de classificatie herhaald, en het foutpercentage berekend voor alle mogelijke combinaties van 1 behouden en 1 verworpen parameter. Combinaties die dan lage fouten gaven (lager dan die van de behouden parameter afzonderlijk) werden gedetecteerd, en de bijhorende verworpen parameter werd terug opgevist, en behouden. Op die manier werd de SNR aan de behouden parameters toegevoegd.

Invloed van de hoek De hoekafhankelijkheid is geïllustreerd in Fig. 6.7 en Fig. 6.8, voor vensters van 30x30 en 60x60 pixels. De twee andere venstergroottes geven analoge resultaten. Op deze figuur is duidelijk te zien dat de SNR afzonderlijk heel slechte resultaten oplevert.

De gemiddelde grijswaarde, de variantie op de grijswaarde en de SNR zijn uiteraard hoekonafhankelijk. De correlatie en de verwante parameters IC_1 en IC_2 geven de beste resultaten voor een hoek van 45 graden. Dit resultaat was ergens te verwachten, aangezien de textuur duidelijk in een hoek van 45 graden loopt en het lineaire verband in die richting dus het grootst zal zijn.

Het ATM geeft, behalve bij een venster van 50x50 pixels, de beste resultaten bij een hoek van 0 graden. Voor een venster van 50x50 pixels levert een hoek van 45 graden echter een verrassend lage classificatiefout van 20% op, waar deze fout elders meestal tussen 26%-35% blijft. Deze uitschieter is waarschijnlijk te wijten aan de keuze van de ROI's. We moeten de ROI voldoende groot kiezen, om de textuur op een voldoende grote schaal te kunnen beschrijven. Wanneer we het venster echter te groot kiezen, omvat het niet enkel weefsel meer, maar ook andere delen, waarin we uiteraard niet geïnteresseerd zijn. Door deze niet relevante gebieden zal de significantie van de bekomen textuurparameters natuurlijk verminderen en zal er soms een uitschieter optreden.

De classificatiefouten voor het ATM liggen voor 0 en 45 graden vrij dicht bij elkaar, zodat we voor de coronale beelden 45 graden als optimale hoek kunnen beschouwen.

Invloed van de grootte van de ROI We beschouwen enkel nog de parameters voor een hoek van 45 graden. Deze resultaten vinden we terug in Fig. 6.9. We zien dat het onderscheidend vermogen van de parameters slechts weinig varieert met de venstergrootte. Er zijn enkele uitschieters, zoals bijvoorbeeld de 33% bij de variantie op de grijswaarde en de 20% bij het ATM. Deze wispelturigheid is opnieuw te wijten aan de keuzes van de ROI. Vanaf hier werken we verder met een venster van 60x60 pixels, omdat ruis dan het best wordt uitgemiddeld. Bij de keuze van een venster van 60x60 pixels moet men er wel

op letten dat het gebied zoveel mogelijk ziek of gezond weefsel bevat, zodat de informatie in het venster relevant is.

Combineren van de parameters Nu we de optimale parameters en de optimale hoek bepaald hebben, moeten we de parameters nog op zo'n manier combineren dat de classificatiefout minimaal wordt. Wanneer we alle zeven de parameters gebruiken als input voor de Bayesiaanse classificatie (we onderstellen dus dat de parameters zevendimensionaal normaal verdeeld zijn), wordt maar 1 beeld verkeerd beoordeeld, wat overeenkomt met een classificatiefout van 2%. De reductie van zeven naar drie of twee parameters wordt hieronder kort beschreven. De correlatiematrix ρ van de zeven resterende parameters is:

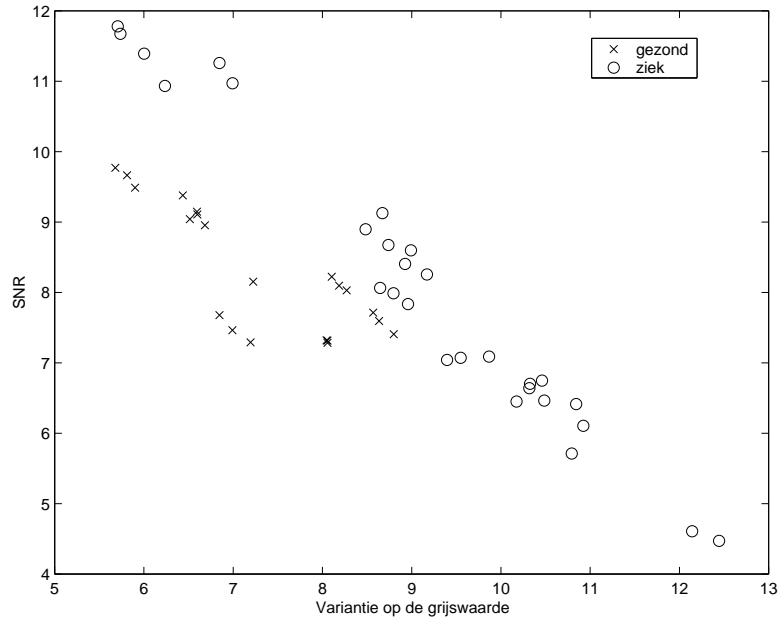
$$\rho = \begin{pmatrix} 1 & 0.23 & -0.30 & 0.49 & -0.55 & 0.44 & 0.25 \\ 0.23 & 1 & -\mathbf{0.86} & 0.55 & -0.51 & 0.70 & -\mathbf{0.86} \\ -0.30 & -\mathbf{0.86} & 1 & -0.29 & 0.17 & -0.44 & \mathbf{0.79} \\ 0.49 & 0.55 & -0.29 & 1 & -\mathbf{0.96} & \mathbf{0.97} & -0.25 \\ -0.35 & -0.51 & 0.17 & -\mathbf{0.96} & 1 & -\mathbf{0.94} & 0.25 \\ 0.44 & 0.70 & -0.44 & \mathbf{0.97} & -\mathbf{0.94} & 1 & -0.44 \\ 0.25 & -\mathbf{0.86} & \mathbf{0.79} & -0.25 & 0.25 & -0.44 & 1 \end{pmatrix}$$

Hierbij is $\rho(i, j)$ de correlatie tussen de i -de en de j -de parameter. Uiteraard is $\rho(i, i) = 1$. De parameters in de matrix zijn respectievelijk:

1. f_1 =Gemiddelde grijswaarde
2. f_2 =Variantie op de grijswaarde
3. f_4 =ATM
4. f_6 =Correlatie
5. f_{13} = IC_1
6. f_{14} = IC_2
7. f_3 =SNR

We zien dat de correlatie, IC_1 en IC_2 sterk ($|\rho| > 0.94$) gecorreleerd zijn en dat het ATM, de variantie op de grijswaarde en de SNR ook een hoge correlatie vertonen ($\rho = -0.86$ en $\rho = 0.79$ respectievelijk). We hebben dus drie 'sets' van parameters:

- De drie correlatieparameters



Figuur 6.3: De variantie op de grijswaarde en de SNR voor een venster van 60x60 pixels, bij coronale beelden.

- De gemiddelde grijswaarde
- Het ATM, de variantie op de grijswaarde en de SNR.

Classificatie op basis van een driedimensionele input, met 1 parameter uit elke set geeft een minimaal foutpercentage van 10.2% (bij correlatie, gemiddelde grijswaarde en variantie op grijswaarde).

Wanneer we echter de variantie op de grijswaarde en de SNR combineren krijgen we een classificatiefout van 0%, dus een perfecte classificatie, zie Fig. 6.3. Dit is eigenaardig omdat beide parameters vrij sterk gecorreleerd zijn en beide afzonderlijk geen goede resultaten geven. De hoge correlatie is duidelijk zichtbaar op de figuur. We zien op de figuur ook dat de zieke en gezonde waarnemingen heel dicht bij elkaar liggen. De meeste beelden liggen heel dicht tegen de beslissingsgrens en de variantie op de parameters moet bijgevolg heel klein blijven. De waarde van deze oplossing moet nog verder getest worden, door bijvoorbeeld een grotere dataset te gebruiken. Deze parametercombinatie is dus waarschijnlijk geen ideale oplossing, en daarom zullen we toch nog verder op zoek gaan naar andere combinaties.

Wanneer we de zeven parameters normaliseren en dan PCA toepassen om de parameter ruimte tot twee dimensies te reduceren bekomen we opnieuw een fout van 10,2%. Reductie tot een driedimensionale ruimte heeft hier geen voordeel ten opzichte van de tweedimensionale ruimte, omdat de derde eigenwaarde reeds heel klein is.

Besluit Bij de coronale beelden halen we de beste resultaten bij een hoek van 45 graden en nemen we het grootste venster, om de ruis zoveel mogelijk uit te middelen. Perfecte classificatie is mogelijk, op basis van de variantie op de grijswaarde en de SNR. De waarden voor ziek en gezond liggen echter heel dicht tegen elkaar. Wanneer we ons niet tot twee of drie dimensies beperken is een classificatiefout van 2% mogelijk.

6.6.2 Sagittale beelden

Verworpen parameters Net als bij de coronale beelden kunnen we hier op voorhand al enkele parameters verwerpen, omdat het onderscheidend vermogen ervan te slecht is. Wanneer we dezelfde selectiecriteria toepassen als bij de coronale beelden blijven de volgende parameters over:

- f_1 =Gemiddelde grijswaarde
- f_2 =Variantie op de grijswaarde
- f_4 =ATM
- f_{14} = IC_2

Correlatie is ook een goede textuurparameter, maar is sterk ($\rho = 0.997$) gecorreleerd met IC_2 . Deze laatste geeft net iets betere resultaten zodat we correlatie verwerpen. Om dezelfde reden hebben we ook f_7 verworpen, die sterk (ook $\rho = 0.997$) gecorreleerd is met de gemiddelde grijswaarde. Bij de sagittale beelden levert de combinatie van de SNR met andere parameters nergens goede waarden op. Deze parameter wordt hier dan ook niet behouden.

Invloed van de ROI Bij een venster van 60x60 pixels levert de gemiddelde grijswaarde een classificatiefout van 0% op. Op basis van deze eerste orde parameter is dus perfecte classificatie mogelijk. Omdat de gemiddelde grijswaarde bij alle beelden (zowel sagittaal als coronaal) de beste parameter gebleken is, werken we vanaf nu verder met een venster van 60x60, zodat de ruis op de gemiddelde grijswaarde maximaal uitgemiddeld wordt. Ook voor de andere parameters (op de variantie op de grijswaarde na) geeft een venster van 60x60 het beste resultaat. Dit is geïllustreerd in Fig. 6.10, voor een hoek van 45 graden. Voor de andere hoeken bekomen we analoge resultaten.

Invloed van de hoek We beschouwen enkel nog een venster van 60x60 pixels. Het gemiddelde van de grijswaarde en de variantie erop zijn uiteraard weer onafhankelijk van de hoek waarvoor de coöccurrentiematrix werd berekend. In tegenstelling tot bij de coronale

beelden hebben we hier geen duidelijk voordeel bij een hoek van 45 graden. Hoeken van 45, 90 en 135 graden geven ongeveer gelijke resultaten, zie Fig. 6.11.

Combineren van de parameters Aangezien de gemiddelde grijswaarde een fout van 0% geeft, kunnen we eventueel classificeren op basis van enkel deze parameter. Toch kan het nuttig zijn om ook de andere parameters nog te gebruiken. Bij een grotere dataset zou het namelijk kunnen dat enkel de gemiddelde grijswaarde niet meer volstaat en we de andere parameters dus toch nodig hebben. Door naar twee of meerdere dimensies over te gaan kunnen we de euclidische afstand tussen de zieke en gezonde wolk trachten te vergroten.

Overgang naar meerdere dimensies kan echter ook nadelig zijn voor de Bayesiaanse classificatie. Wanneer we bijvoorbeeld de gemiddelde grijswaarde en de variantie op de grijswaarde combineren krijgen we een classificatiefout van 7.8%. Dit is geïllustreerd in Fig. 6.4. We zien op deze figuur dat de gemiddelde grijswaarde een perfect onderscheid mogelijk maakt. De twee gezonde beelden die buiten de beslissingsgrens vallen zullen echter dichterbij de ‘zieke wolk’ dan bij de ‘gezonde wolk’ aanleunen, en dus toch verkeerd geclassificeerd worden.

We kunnen op deze figuur de clustering, in clusters van vier, bij de zieke beelden duidelijk opmerken. De oorzaak hiervan is het kiezen van meerdere (hier dus vier) ROI's op een beeld, waarbij deze gebieden uiteraard (kunnen) overlappen.

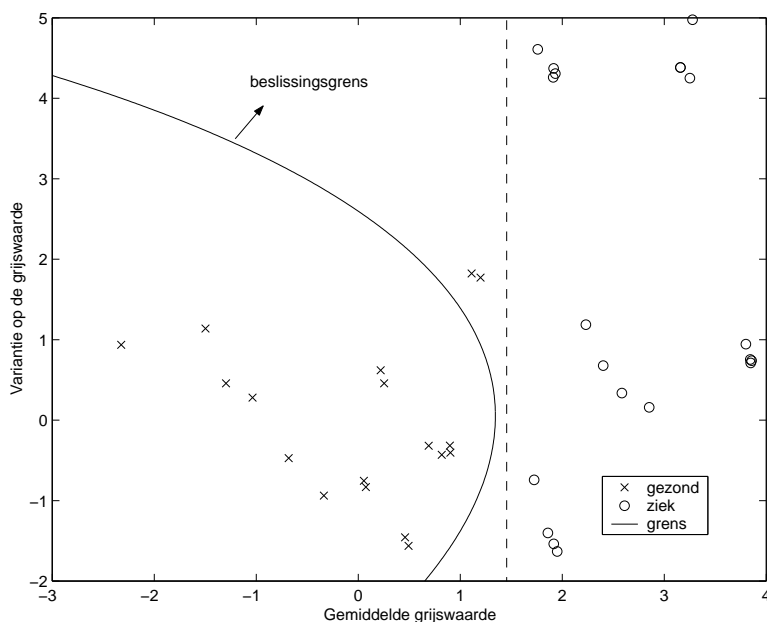
Wanneer we ATM, gemiddelde grijswaarde en variantie op de grijswaarde gebruiken blijft de fout 0%, ook de combinatie van het ATM en de gemiddelde grijswaarde maakt perfecte classificatie mogelijk, zie Fig. 6.5.

Besluit Met de gemiddelde grijswaarde is perfecte classificatie mogelijk bij een venster van 60x60 pixels. Het kan toch nuttig zijn om de andere parameters ook te gebruiken, om de afstand tussen de zieke en gezonde beelden te vergroten. In tegenstelling tot bij de coronale beelden is er hier geen beter resultaat bij een hoek van 45 graden.

6.6.3 Combinatie van coronale en sagittale beelden

Het is ook nuttig om de sagittale en de coronale beelden tesamen te bestuderen. Bepaalde parameters kunnen misschien een even goede classificatie opleveren wanneer we de beelden niet scheiden. In dat geval kunnen we de beelden classificeren zonder steeds het onderscheid te moeten maken.

Verworpen parameters Zoals te verwachten was, geven de textuurparameters die bij de afzonderlijke beelden optimaal waren ook hier de beste resultaten. We gebruiken dezelf-



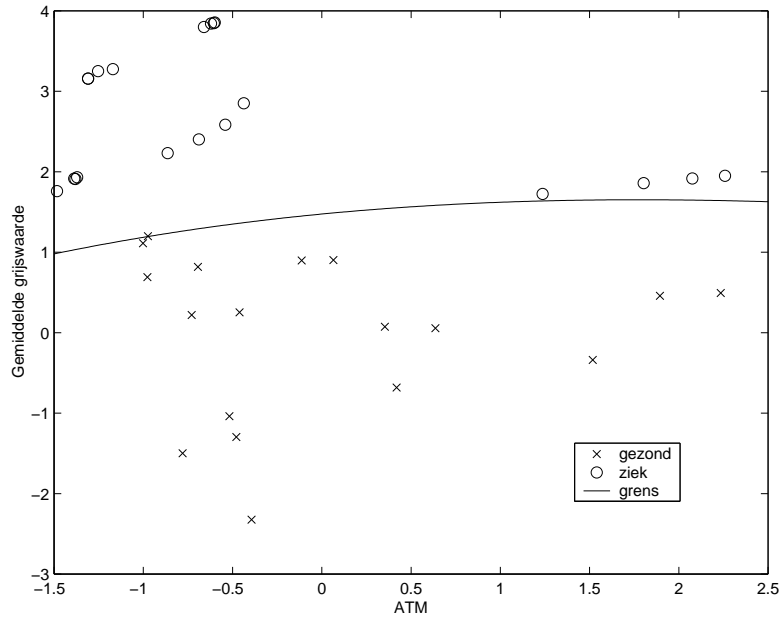
Figuur 6.4: De gemiddelde grijswaarde en de variantie op de grijswaarde, voor een venster van 60x60 pixels, bij sagittale beelden. De volle lijn is de beslissingsgrens, de streeplijn toont aan dat op basis van enkel de gemiddelde grijswaarde inderdaad perfecte classificatie mogelijk is.

de selectiecriteria als voor de coronale en sagittale beelden. Dit levert, op IC_1 na, dezelfde parameters op als bij de coronale beelden:

- f_1 =Gemiddelde grijswaarde
- f_2 =Variantie op de grijswaarde
- f_3 =SNR
- f_4 =ATM
- f_6 =Correlatie
- f_{13} = IC_2

Invloed van de ROI en de hoek Net als bij de sagittale beelden is er niet echt een hoek die opvallend betere resultaten geeft. De foutpercentages variëren weinig met de hoek. Bij voldoende grote vensters geven de correlatieparameters opvallend slechtere resultaten voor een hoek van 0 graden. We kiezen hier voor een hoek van 45 graden, omdat die net iets lagere classificatiefouten oplevert dan de andere hoeken.

We kiezen verder voor een ROI van 60x60 pixels. De reden voor deze keuze is dat het ATM betere resultaten geeft bij grotere vensters, en dat de gemiddelde grijswaarde, zoals



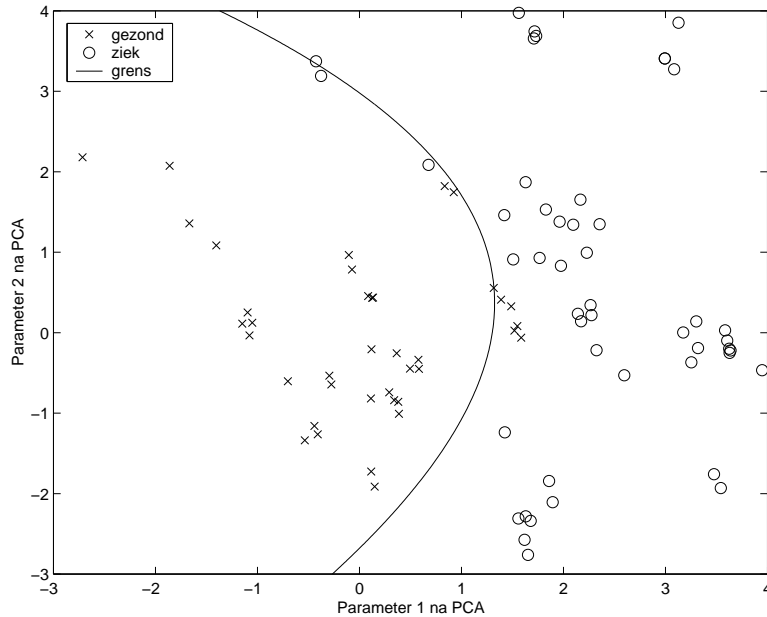
Figuur 6.5: Het ATM en de gemiddelde grijswaarde voor een venster van 60x60 pixels en een hoek van 45 graden, bij sagittale beelden.

vroeger reeds vermeld, dan ook beter werkt.

Combineren van de parameters We combineren de parameters weer en zoeken de combinatie die de laagste fout geeft. Wanneer we een Bayesiaanse classificatie toepassen op alle zes de parameters (we gebruiken dus een zesdimensionale normale distributie) krijgen we een classificatiefout van 8%. Om de resultaten grafisch te kunnen voorstellen reduceren we de parameterruimte naar twee of drie parameters. Hiervoor kunnen we PCA gebruiken: de grootste drie eigenwaarden van de covariantiematrix zijn: $[-63,4 \ -5.2 \ -0.0021]$. Aangezien de derde eigenwaarde reeds heel klein is reduceren we de ruimte naar twee dimensies, zie Fig. 6.6. De classificatiefout wordt dan 10%.

We kunnen ook naar een lagere dimensie overgaan door gewoon twee of drie parameters te combineren en de rest te laten vallen. De beste resultaten worden hieronder opgesomd:

- De combinatie ATM, gemiddelde grijswaarde en variantie op de grijswaarde geeft een fout van 3.4%.
- De combinatie van variantie op de grijswaarde met SNR leidt eveneens tot een foutpercentage van 3.4%. De twee klassen liggen echter weer heel dicht bij elkaar (cfr. Fig. 6.3), en we kunnen hier dezelfde opmerkingen maken als bij de coronale beelden.
- Combineren van de gemiddelde grijswaarde en de variantie op de grijswaarde resul-



Figuur 6.6: De twee parameters na een PCA reductie bij een venster van 60x60 pixels en een hoek van 45 graden, voor de coronale en sagittale beelden tesamen.

teert in een classificatiefout van 10%.

Besluit De parameters die bij de coronale of sagittale beelden goed werkten, geven ook bij de combinatie ervan de betere resultaten. Er is niet echt een ‘beste hoek’, ook de venstergrootte geeft geen enorme verschillen in performantie. Wanneer we classificeren op basis van de zes behouden parameters is een foutpercentage van 8% mogelijk. Wanneer we de juiste parameters kiezen (ATM, gemiddelde grijswaarde en variantie op de grijswaarde) is een classificatiefout van 3.4% mogelijk.

6.7 Niet gecompenseerde beelden

6.7.1 Coronale beelden

Verworpen parameters We gebruiken dezelfde selectiecriteria als bij de gecompenseerde beelden. De hieronder opgesomde parameters blijven over na de selectie. Dit zijn dezelfde parameters als bij de wel gecompenseerde coronale beelden:

- f_1 =Gemiddelde grijswaarde
- f_2 =Variantie op de grijswaarde
- f_3 =SNR

- $f_4=ATM$
- $f_6=Correlatie$
- $f_{13}=IC_1$
- $f_{14}=IC_2$

Invloed van hoek en venster De textuurparameters werden berekend voor een venster van 60x60 pixels. Een hoek van 45 graden geeft ook hier weer iets betere resultaten, vooral voor de correlatieparameters. We werken dan ook verder met deze hoek.

Vergelijking met de parameters van gecompenseerde beelden De tweede orde textuurparameters van de niet gecompenseerde beelden geven iets betere resultaten dan die van de wel gecompenseerde beelden. De eerste orde statistieken gemiddelde grijswaarde en variantie van de grijswaarde daarentegen geven opvallend betere resultaten bij de gecompenseerde beelden. Dit is geïllustreerd in Fig. 6.12, voor een venster van 60x60 pixels en een hoek van 45 graden. Voor de andere combinaties van venstergrootte en hoek werden analoge resultaten bekomen.

Het was te verwachten dat de eerste orde parameters betere resultaten zouden geven na compensatie: de grijswaarden van de niet gecompenseerde beelden hangen in hoge mate af van de instellingen van het echografietoestel. Afhankelijk van de sterkte van de invallende geluidsgolven zullen we grotere of kleinere grijswaarden krijgen. Wanneer we een geluidsgolf met een grote amplitude laten invallen op gezond weefsel kan de gemiddelde grijswaarde hoger zijn dan wanneer we een geluidsgolf met kleine amplitude laten invallen op ziek weefsel. Hierdoor kan het zieke weefsel (dat in het algemeen gekenmerkt wordt door hogere grijswaarden) verkeerdelijk als gezond weefsel beoordeeld worden, en kan het gezonde weefsel (dat in het algemeen gekenmerkt wordt door lagere grijswaarden) verkeerdelijk als ziek weefsel beoordeeld worden. Wanneer we het compensatie-algoritme toepassen, zal de invloed van de instellingen echter geminimaliseerd worden, en zal de classificatie op basis van de eerste orde statistieken dus beter zijn omdat alle beelden naar eenzelfde referentie teruggerekend zijn.

Aangezien de eerste orde parameters heel belangrijk zijn voor een goede classificatie kunnen we vermoeden dat de gecompenseerde beelden tot betere classificatie zullen leiden dan de niet gecompenseerde beelden. Dit vermoeden zal verder inderdaad nog enkele malen bevestigd worden.

Combineren van de parameters Wanneer we alle 7 de parameters gebruiken als input voor classificatie bekomen we een fout van 10.2%. Wanneer we echter enkel de gemiddelde grijswaarde, de variantie op de grijswaarde en de SNR (dus de drie eerste orde parameters) gebruiken als input krijgen we een classificatiefout van 2%. Wanneer in twee dimensies wordt gewerkt resulteren de gemiddelde grijswaarde en de SNR in een foutpercentage van 10.2%. Toepassen van PCA geeft geen verbetering van de resultaten: reductie tot drie dimensies geeft een classificatiefout van 18%, reductie tot twee dimensies van 16%.

Besluit Zonder compensatie kunnen we een classificatiefout van 2% halen, wanneer men de drie eerste orde parameters combineert. Met gecompenseerde beelden is perfecte classificatie mogelijk met dezelfde parameters. Compensatie is dus voordelig voor coronale beelden.

6.7.2 Sagittale beelden

Verworpen parameters Na de selectie houden we de volgende parameters over:

- f_1 =Gemiddelde grijswaarde
- f_2 =Variantie op de grijswaarde
- f_4 =ATM
- f_6 =Contrast
- f_{14} =IC₂

Net als bij de niet gecompenseerde beelden is SNR hier niet nuttig. In tegenstelling tot bij de gecompenseerde beelden geeft het contrast hier wel goede resultaten.

Invloed van de hoek en vergelijking met gecompenseerde beelden Net als bij de gecompenseerde sagittale beelden is er weinig hoekafhankelijkheid, we bekomen een figuur analoog aan Fig. 6.10. Ook hier zijn de tweede orde parameters beter bij niet gecompenseerde beelden. De eerste orde parameters gemiddelde grijswaarde en SNR werken dan weer beter bij de gecompenseerde beelden. Dit wordt geïllustreerd in Fig. 6.13 voor een hoek van 45 graden.

Combineren van de parameters Wanneer we alle vijf de parameters combineren krijgen we een classificatiefout van 2%. Combinatie van gemiddelde grijswaarde, ATM en IC₂ resulteert in een fout van 0%.

Ook nu weer geeft PCA geen betere resultaten, reductie naar 3 dimensies leidt tot een classificatiefout van 5.3%, reductie naar 2 dimensies leidt tot een foutpercentage van 23%.

Besluit Ook bij de niet gecompenseerde beelden is een perfecte classificatie mogelijk. Voor sagittale beelden is compensatie dus niet noodzakelijk.

6.7.3 Combinatie van coronale en sagittale beelden

Verworpen parameters Na selectie blijven dezelfde parameters over als bij de gecompenseerde beelden:

- f_1 =Gemiddelde grijswaarde
- f_2 =Variantie op de grijswaarde
- f_3 =SNR
- f_4 =ATM
- f_6 =Correlatie
- f_{12} =IC₁
- f_{13} =IC₂

We krijgen weer iets mindere resultaten voor een hoek van 0 graden, een hoek van 90 graden geeft voor de afzonderlijke parameters iets betere resultaten dan een hoek van 45 graden. Wanneer we de parameters combineren geeft een hoek van 45 graden echter even goede classificatie. De resultaten in de volgende paragraaf zijn bekomen met een hoek van 45 graden.

Combineren van de parameters Wanneer we de zeven parameters als invoer voor classificatie gebruiken, bekomen we een classificatiefout van 9%. PCA reductie naar twee dimensies leidt tot een classificatiefout van 12,6%. De beste resultaten worden weer bekomen door gewoon de juiste parameters te combineren:

- In drie dimensies: de combinatie van het ATM, IC₂ en de gemiddelde grijswaarde geeft een fout van 10%.
- In twee dimensies: combinatie van het ATM en de gemiddelde grijswaarde resulteert in een foutpercentage van 12.6%.

- Combineren van de gemiddelde grijswaarde en de SNR geeft hier een vrij hoge classificatiefout van 21%, daar waar dit bij gecompenseerde beelden de optimale combinatie was.

Besluit Er is een minimale classificatiefout van 9% mogelijk, door alle zeven de parameters te combineren. Bij de gecompenseerde beelden is met dezelfde parameters een fout van 3,4% mogelijk. Wanneer coronale en sagittale beelden samen worden beschouwd is compensatie voordeliger dan wanneer we de beelden apart beschouwen. Dit is als volgt te verklaren: omdat de texturen van de sagittale en coronale beelden te verschillend zijn, zullen de textuurparameters minder goed werken. De belangrijkste parameters zijn dus de eerste orde parameters, die bij compensatie veel beter werken.

6.8 Invloed van de afstand d

Om te besluiten wordt nog kort de invloed van de *afstand* d op de parameters onderzocht. Enkel afstanden van 1 tot 5 werden onderzocht. Uiteraard zullen de eerste orde parameters niet afhankelijk zijn van de afstand d .

6.8.1 Coronale beelden

Gecompenseerde beelden Bij de gecompenseerde coronale beelden is de classificatiefout bij de afzonderlijke parameters slechts weinig afhankelijk van de waarde van d . In Fig. 6.14 wordt dit geïllustreerd voor de tweede orde parameters.

Van de parameters die we bij $d = 1$ konden verwerpen, omdat hun classificatiefout te hoog lag of omdat ze te sterk gecorreleerd waren met een andere parameter, levert enkel het invers verschilmoment f_8 opmerkelijk betere resultaten op bij een grotere d . De 4 tweede orde parameters die bij $d = 1$ werden behouden, geven bij $d = 4$ iets lagere classificatiefouten.

We zoeken nu de optimale parametercombinaties bij $d = 4$, op basis van de parameters die bij $d = 1$ werden behouden, en het invers verschilmoment. Geen van de combinaties die de laagste classificatiefout opleveren, maakt gebruik van het invers verschilmoment, zodat de lage foutwaarde van deze parameter bij $d = 4$ onbelangrijk is. We halen een foutpercentage van 2% door het ATM, de entropie en de gemiddelde grijswaarde te combineren. Deze waarde is dezelfde als bij $d = 1$ haalbaar was, we hebben dus geen voordeel wanneer we $d = 4$ gebruiken.

Niet gecompenseerde beelden Ook bij de niet gecompenseerde beelden is de classificatiefout bij de afzonderlijke parameters weinig afhankelijk van d , hier hebben we een

heel klein voordeel bij een afstand $d = 1$, zie Fig. 6.15. Met $d = 1$ was door de eerste orde parameters te combineren een classificatiefout van 2% mogelijk. Bij de andere waarden voor d bekomen we hetzelfde resultaat: steeds een minimale classificatiefout van 2%, en steeds door combinatie van de eerste orde parameters.

Besluit Bij de gecompenseerde beelden geeft $d = 4$ iets lagere classificatiefouten voor de afzonderlijke parameters, bij de niet gecompenseerde beelden hebben we een licht voordeel bij $d = 1$. De significantie hiervan moet uiteraard wel nog gecontroleerd worden met een grotere dataset. Wanneer we de parameters combineren, verkrijgen we echter voor de meeste waarden van d dezelfde foutpercentages.

We kunnen dus besluiten dat een afstand $d = 1$ bij de coronale beelden een goede keuze is. De andere afstanden d geven niet noodzakelijk slechtere resultaten, maar ook geen betere.

6.8.2 Sagittale beelden

Gecompenseerde beelden De parameters die optimaal waren bij $d = 1$, zijn ook bij de andere waarden van d de beste. Van de parameters die bij $d = 1$ werden verworpen, zijn er geen die bij grotere d -waarden toch goede resultaten geven. Bij de gecompenseerde sagittale beelden is $d = 1$ dus een goede keuze.

Niet gecompenseerde beelden Bij de niet gecompenseerde sagittale beelden gelden dezelfde opmerkingen als bij de gecompenseerde sagittale beelden. Voor grotere afstanden hebben we wel opvallend lagere fouten bij het invers verschilmoment, zie Fig. 6.16. Net als bij de coronale beelden kunnen we dit voordeel echter niet uitbuiten om na combinatie van de parameters lagere classificatiefouten te behalen.

Besluit Ook bij de sagittale beelden is het gebruik van een afstand $d = 1$ een goede keuze. Met andere afstanden zijn even goede resultaten mogelijk, maar geen betere.

Aangezien $d = 1$ bij zowel de coronale als sagittale beelden de beste resultaten geeft, zal dit ook bij de combinatie van coronale en sagittale beelden zo zal zijn. Daarom gaan we op dit geval niet verder in.

6.9 Overzicht van de bekomen resultaten

De resultaten worden samengevat in Tabel 6.1. Voor elke combinatie van coronaal-sagittaal en gecompenseerd-niet gecompenseerd staan de optimale parameters aangeduid. Tenslotte

wordt bij elke combinatie het laagst haalbare foutpercentage gegeven. De parameters die voor deze laagste fout nodig zijn staan weergegeven als een dubbele 'x'.

Tabel 6.1: Samenvatting van de optimale parameters voor de verschillende combinaties sagittaal-coronaal en gecompenseerd-niet gecompenseerd. Een dubbele 'x' geeft aan dat de parameter nodig is om het laagste foutpercentage te behalen. g = gecompenseerd, ng = niet gecompenseerd

Parameter	Coronaal		Sagittaal		Beide	
	g	ng	g	ng	g	ng
Gemiddelde grijswaarde	xx	xx	xx	xx	xx	xx
Variantie grijswaarde	xx	xx	x	x	xx	xx
SNR	xx	xx			x	xx
ATM	x	x	x	xx	xx	xx
Contrast				x		
Correlatie	x	x			x	xx
IC ₁	x	x				xx
IC ₂	x	x	x	xx	x	xx
Minimale fout (%)	0	2	0	0	3.4	9

Invloed van de afstand d Een afstand $d = 1$ is een goede keuze. Andere afstanden werken niet noodzakelijk slechter, maar ook niet beter. De resultaten in Tabel 6.1 zijn dan ook allemaal voor een afstand $d = 1$.

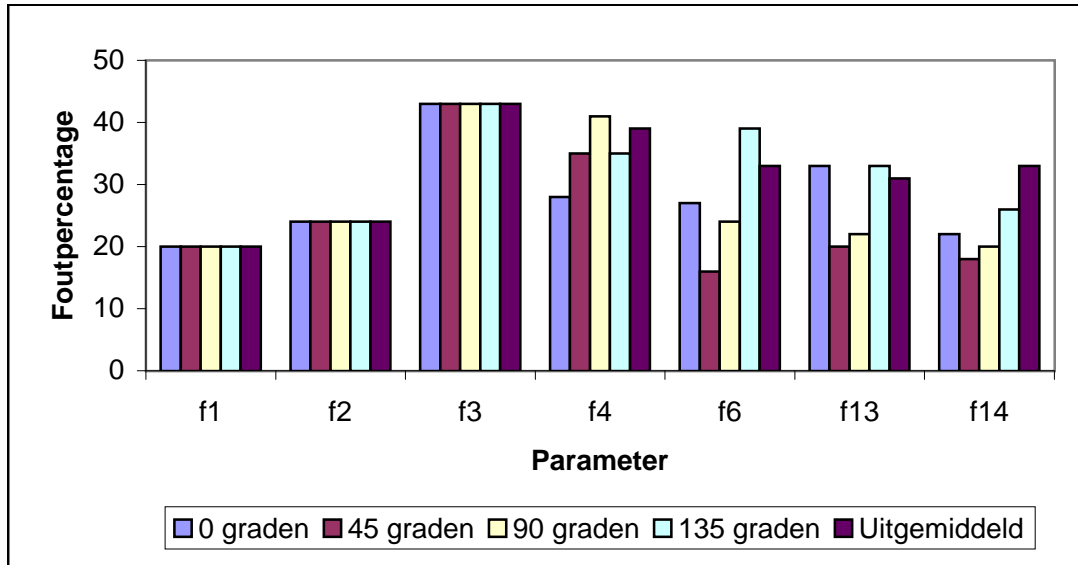
Hoekafhankelijkheid Voor de coronale beelden is het voordelig om een hoek van 45 graden te gebruiken. Dit was te verwachten omdat de ROI's steeds uit de rechterkant van het beeld gekozen werden, zodat de textuur in een richting van 45 graden loopt. Voor de sagittale beelden is de kwaliteit van de classificatie echter weinig afhankelijk van de hoek, omdat de textuur in de verschillende ROI's in verschillende richtingen loopt. De reden hiervoor is dat het aantal sagittale beelden beperkt is, en we uit elk beeld meerdere ROI's moeten kiezen, en we niet voor alle ROI's dezelfde hoek kunnen nemen (zie Hoofdstuk 3). Wanneer we alle beelden samen beschouwen, hebben we ook iets betere resultaten bij een hoek van 45 graden. We kunnen dus besluiten dat 45 graden de optimale hoek is.

Invloed van grootte ROI We hebben gekozen voor een venster van 60x60 pixels. De reden hiervoor is dat de andere venstergroottes geen betere resultaten gaven, en dat bij grotere vensters de ruis meer uitgemiddeld wordt. Bij de keuze van het venster moet men er wel op letten dat het zoveel mogelijk relevant weefsel bevat.

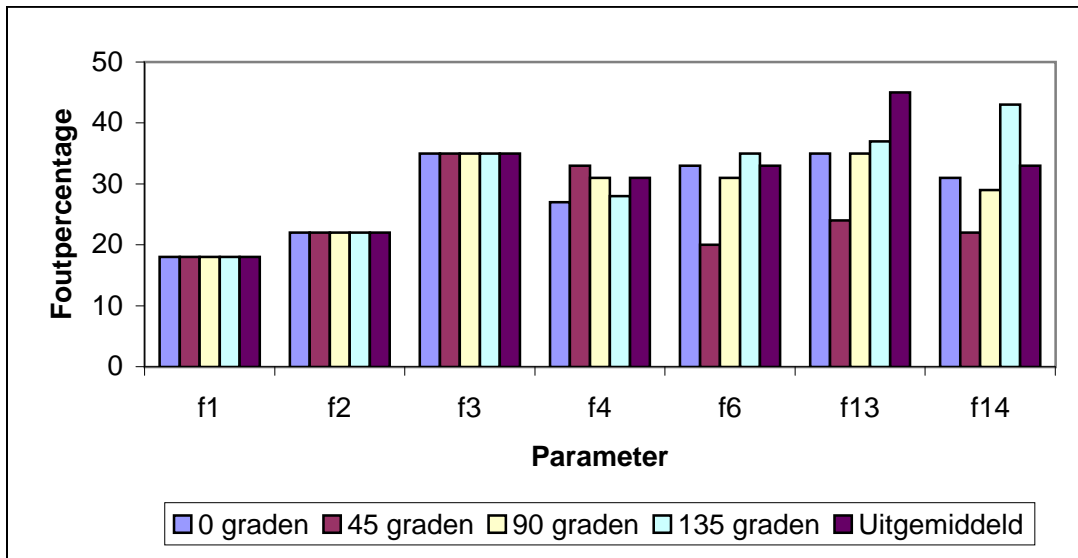
Invloed compensatie Het compensatie-algoritme zorgt er vooral voor dat de eerste orde statistieken voor een beter onderscheid kunnen zorgen. De eigenlijke textuurparameters berekend uit de coöccurrentiematrix daarentegen, werken meestal beter bij de niet gecompenseerde beelden. De eerste orde parameters zijn efficiënter voor classificatie, zodat compensatie globaal gezien iets betere resultaten zal geven.

Splitsen coronale en sagittale beelden We bekomen lagere foutpercentages wanneer we de coronale en sagittale beelden apart behandelen. We kunnen hieruit dus besluiten dat de texturen te verschillend zijn om de beelden samen te beschouwen.

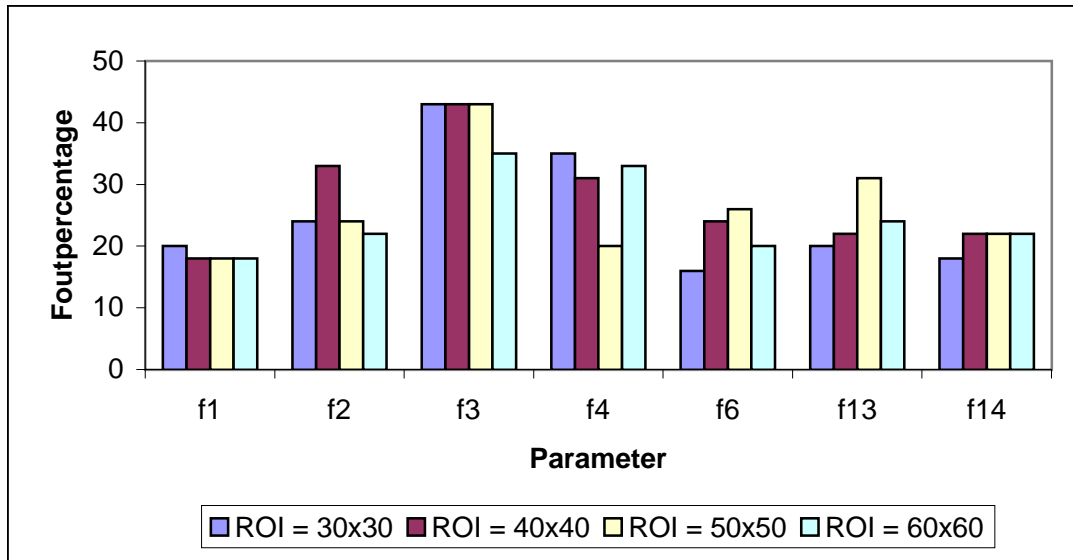
Parameterkeuze De beste resultaten worden bekomen wanneer gebruik gemaakt wordt van de eerste orde parameters gemiddelde grijswaarde en variantie op de grijswaarde, en deze parameters aangevuld worden met een of meerdere textuurparameters. Ook bij de methodes die verder besproken worden zal dit het geval zijn. Met enkel textuurparameters is geen goede classificatie mogelijk.



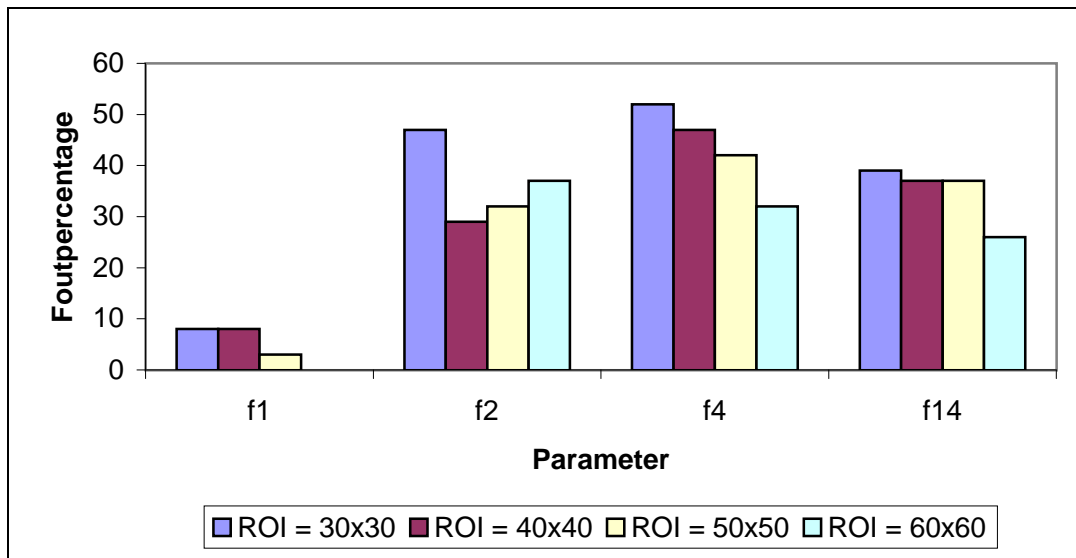
Figuur 6.7: Vergelijking van de classificatiefouten [%] bij de gecompenseerde coronale beelden, voor verschillende waarden van de hoek en een ROI van 30x30 pixels.



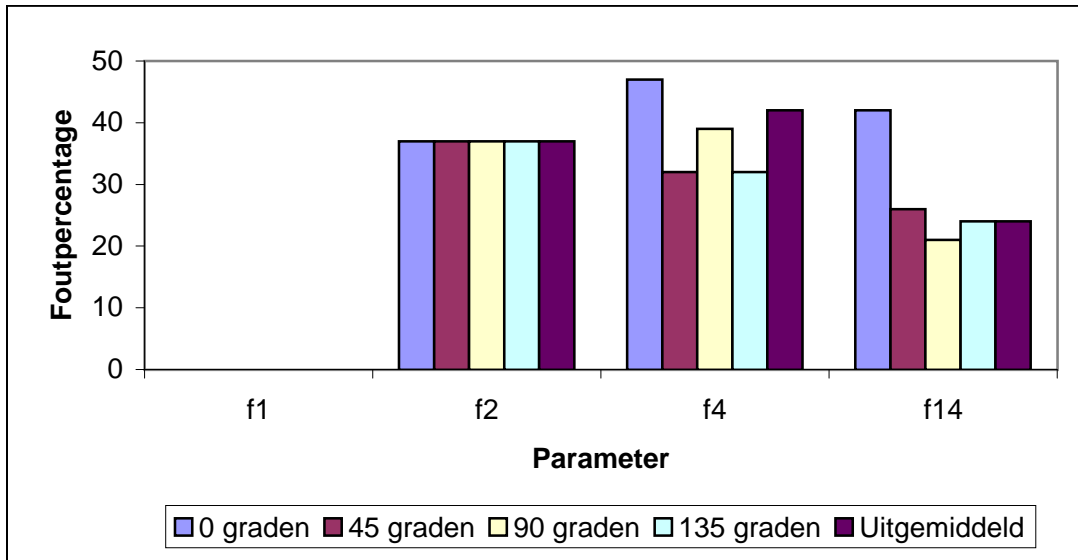
Figuur 6.8: Vergelijking van de classificatiefouten [%] bij de gecompenseerde coronale beelden, voor verschillende waarden van de hoek en een ROI van 60x60 pixels.



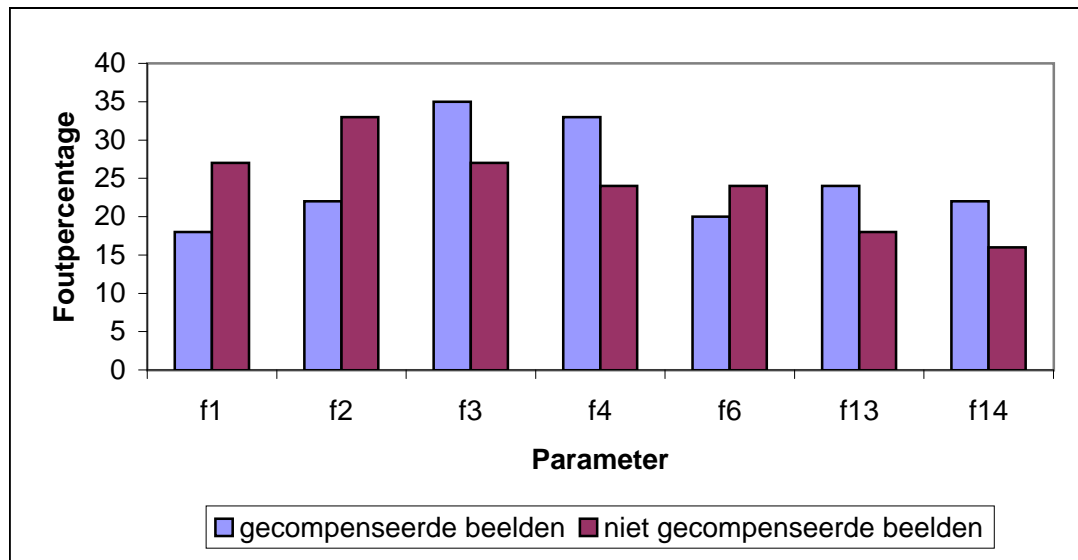
Figuur 6.9: Vergelijking van de classificatiefouten [%] bij de gecompenseerde coronale beelden, voor verschillende venstergroottes en een hoek van 45 graden.



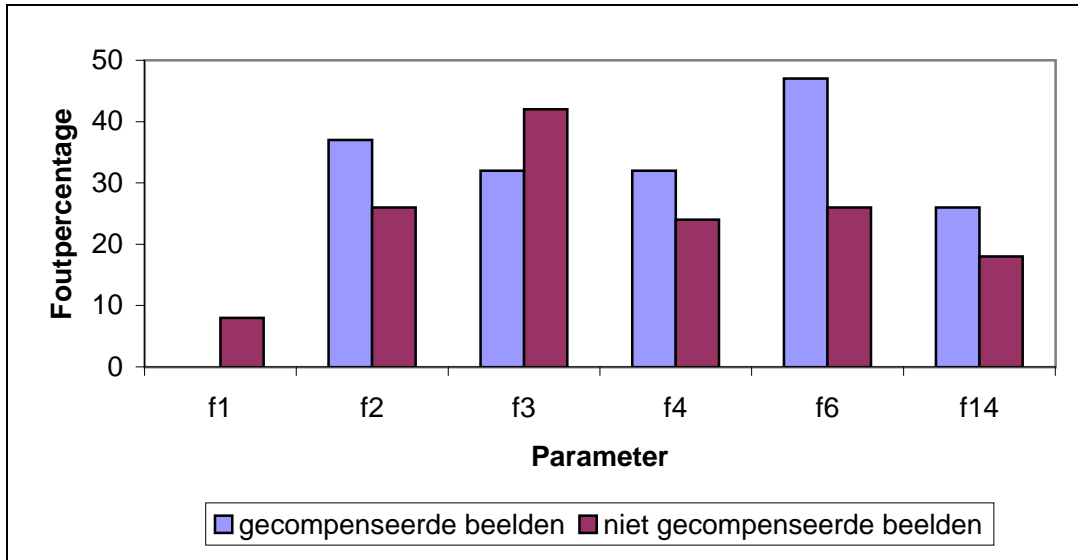
Figuur 6.10: Vergelijking van de classificatiefouten [%] bij de gecompenseerde sagittale beelden, voor verschillende venstergroottes en een hoek van 45 graden.



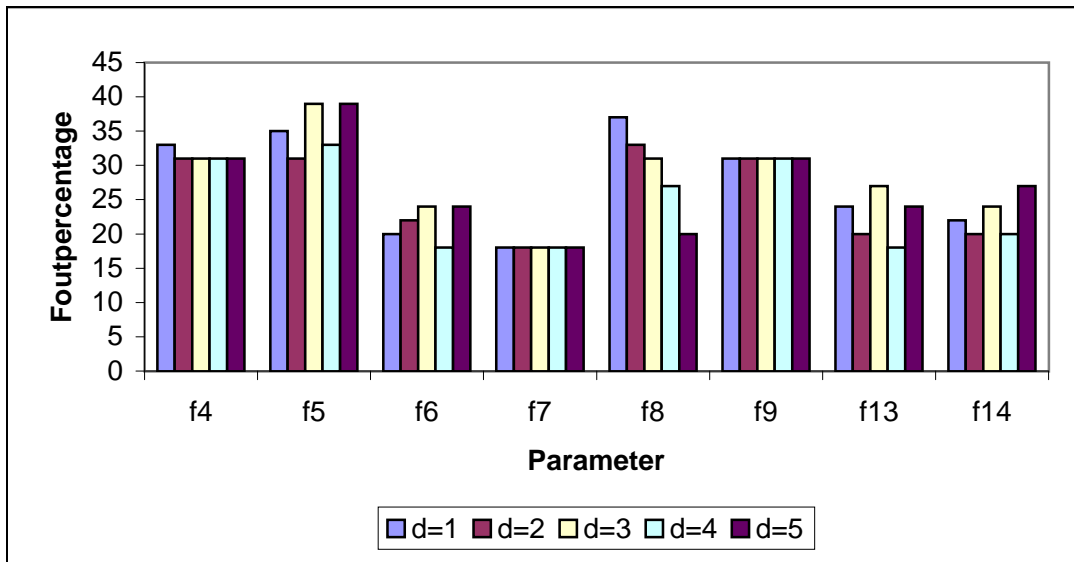
Figuur 6.11: Vergelijking van de classificatiefouten [%] bij de gecompenseerde sagittale beelden, voor verschillende waarden van de hoek en een ROI van 60x60 pixels.



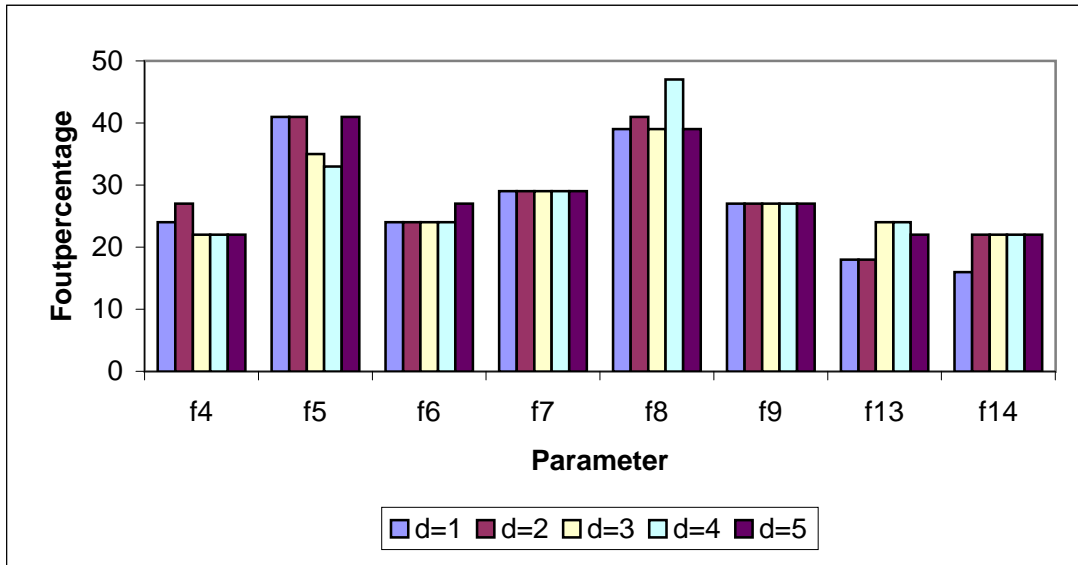
Figuur 6.12: Vergelijking van de classificatiefouten [%] bij de gecompenseerde en niet gecompenseerde coronale beelden, voor een hoek van 45 graden en een ROI van 60x60 pixels.



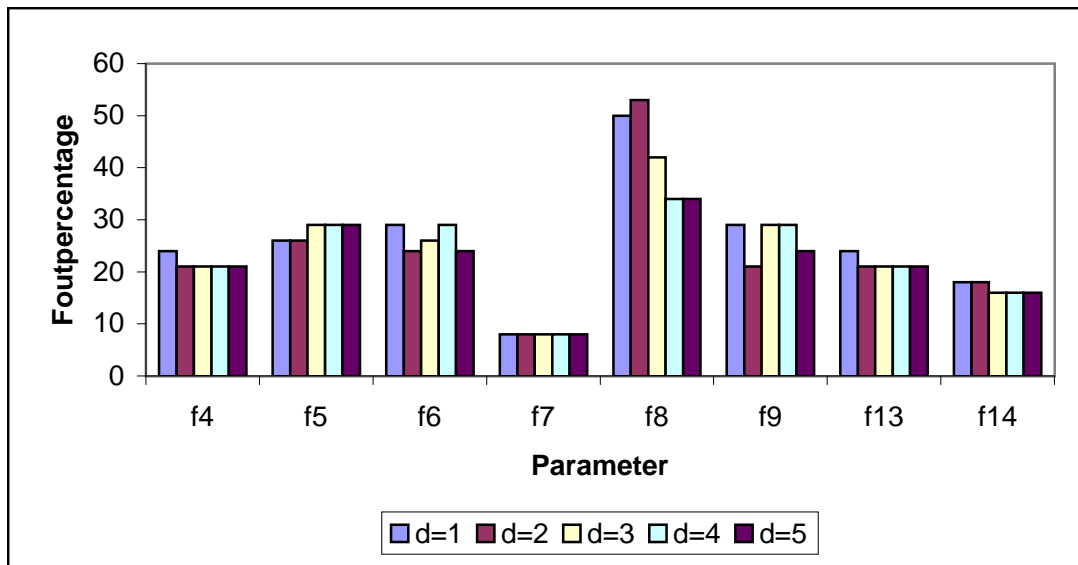
Figuur 6.13: Vergelijking van de classificatiefouten [%] bij de gecompenseerde en niet gecompenseerde sagittale beelden, voor een hoek van 45 graden en een ROI van 60x60 pixels.



Figuur 6.14: Vergelijking van de classificatiefouten [%] bij de gecompenseerde coronale beelden, voor afstanden d=1 tot d=5.



Figuur 6.15: Vergelijking van de classificatiefout [%] bij de niet gecompenseerde coronale beelden, voor afstanden d=1 tot d=5.



Figuur 6.16: Vergelijking van de classificatiefout [%] bij de niet gecompenseerde sagittale beelden, voor afstanden d=1 tot d=5.

Hoofdstuk 7

Varianten op de coöccurrentiematrix

In dit hoofdstuk worden drie methodes besproken die verwant zijn met de coöccurrentiematrix. Eerst worden Unsers som- en verschilhistogrammen besproken [12]. Deze methode is een vereenvoudiging van de coöccurrentiematrix, omdat met eendimensionale histogrammen in plaats van tweedimensionale matrices wordt gewerkt. Een tweede methode maakt gebruik van een gewijzigde coöccurrentiematrix [5]. Tenslotte wordt nog een statistische aanpak besproken, die gebaseerd is op grijswaarde- en gradiënthistogrammen.

7.1 Unsers som- en verschilhistogrammen

Unser ontwikkelde in 1986 een methode voor parameterextractie, die gebruik maakte van som- en verschilhistogrammen [12]. Zijn bedoeling was om een techniek te ontwikkelen die sneller was dan de coöccurrentiematrix en minder geheugen vereiste. De techniek is sneller omdat in plaats van tweedimensionale matrices eendimensionale histogrammen worden gebruikt. Deze snelheidswinst kan belangrijk zijn voor realtime toepassingen. In onze studie is de snelheid echter een minder cruciale factor. Een nadeel ten opzichte van de coöccurrentiematrix is dat door de vereenvoudigingen sommige parameters niet berekend kunnen worden.

7.1.1 Definitie

We beschouwen twee pixels y_1 en y_2 die door afstandsvector $\mathbf{d} = (d_1, d_2)$ gescheiden zijn:

$$\begin{cases} y_1 = y_{k,l} \\ y_2 = y_{k+d_1,l+d_2} \end{cases} \quad (7.1)$$

Zowel de som- als verschilhistogrammen zijn gebaseerd op de sommen $s_{k,l}$ en de verschillen $d_{k,l}$, die als volgt zijn gedefinieerd:

$$\begin{cases} s_{k,l} = y_{k,l} + y_{k+d_1,l+d_2} \\ d_{k,l} = y_{k,l} - y_{k+d_1,l+d_2} \end{cases} \quad (7.2)$$

De sommen $s_{k,l}$ kunnen waarden aannemen tussen 0 en $2G$, de verschillen tussen $-G$ en G , waarbij G de maximale grijswaarde voorstelt. We kunnen nu als volgt het somhistogram $P_S(i)$ en het verschilhistogram $P_D(j)$ definiëren:

$$\begin{aligned} P_s(i) &= h_s(i)/N; \quad (i = 0, \dots, 2G) \\ P_d(j) &= h_d(j)/N; \quad (j = -G, \dots, G) \end{aligned} \quad (7.3)$$

$$\text{waarbij} \begin{cases} h_s(i; d_1, d_2) = h_s(i) = \#\{(k, l) \in ROI, s_{k,l} = i\} \\ h_d(j; d_1, d_2) = h_d(j) = \#\{(k, l) \in ROI, d_{k,l} = j\} \\ N = \sum_{i=0}^{2G} h_s(i) = \sum_{j=-G}^G h_d(j) \end{cases} \quad (7.4)$$

7.1.2 Voorbeeld

Fig. 7.1 toont een beeldmatrix en de bijhorende som- en verschilhistogrammen. De histogrammen werden berekend onder een hoek van 90 graden en een afstand 1, dus $\mathbf{d}=(0,1)$. Om de histogrammen op te stellen beginnen we bij de pixel linksboven, met grijswaarde '0'. De pixel op afstand 1 en een hoek van 90 graden heeft ook '0' als grijswaarde . De som van de twee grijswaarden is dus '0'. In het somhistogram wordt het element dat overeenkomt met nul (dit is de eerste kolom) met 1 verhoogd. Nadat we op deze manier alle pixels overlopen hebben krijgen we het somhistogram. Op een analoge manier berekenen we het verschilhistogram. We delen beide histogrammen door 12, en niet door 16, omdat er maar 12 pixelparen in rekening werden gebracht.

7.1.3 Parameters

Men berekent voor zowel het som- als verschilhistogram de parameters opgesomd in Tabel 7.1. Sommige van deze parameters hebben weer een duidelijke fysische betekenis. Zo is gemakkelijk in te zien dat het gemiddelde van het somhistogram gelijk is aan twee maal de gemiddelde grijswaarde. Het contrast van het verschilhistogram is gelijk aan het contrast zoals berekend uit de coöccurrentiematrix. Ook nu is de entropie een maat voor de homogeniteit van het histogram en is ze maximaal voor uniforme histogrammen. Net als bij de coöccurrentiematrix zijn het angulaire tweede moment en de entropie sterk gecorreleerd.

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

$$P_s = \frac{1}{12} \begin{bmatrix} 3 & 0 & 4 & 2 & 1 & 2 & 0 \end{bmatrix}$$

$$P_d = \frac{1}{12} \begin{bmatrix} 0 & 2 & 4 & 6 & 0 & 0 & 0 \end{bmatrix}$$

Figuur 7.1: De grijswaarden van de pixels in een gebied en bijhorend somhistogram P_s en verschilhistogram P_d met afstand $d = 1$ en onder een hoek van 90 graden.

Tabel 7.1: Parameters berekend uit het som- en verschilhistogram

Parameter	Definitie
Gemiddelde	$f_1 = \sum_i iP(i)$
Angulair Tweede Moment	$f_2 = \sum_i (P(i))^2$
Contrast/Variantie	$f_3 = \sum_i (i - f_1)^2 P(i)$
Entropie	$f_4 = - \sum_i P(i) \log(P(i))$

7.1.4 Toepassing op echografiebeelden

Implementatie Het somhistogram wordt bekomen door de grijswaarden van het oorspronkelijk beeld, daar waar mogelijk, op te tellen met die van het over d_1 en d_2 verschoven beeld. Van dit nieuwe beeld wordt dan het histogram berekend. Op een analoge manier wordt het verschilhistogram bekomen.

De extractie van de parameters hebben we versneld door eerst de minimale grijswaarde G_{min} en de maximale grijswaarde G_{max} te bepalen. Het somhistogram bevat dan elementen tussen $2G_{min}$ en $2G_{max}$. We hebben enkel deze elementen beschouwd voor het berekenen van de parameters (in plaats van 0 tot 2×255). Eenzelfde redenering kan ook voor het verschilhistogram herhaald worden.

Overzicht We hebben de methode zowel op de sagittale en coronale beelden afzonderlijk, als op de combinatie van beide toegepast. De resultaten hiervan worden hieronder beschreven, waarbij kort wordt ingegaan op de invloed van de hoek en compensatie. Alle bekomen resultaten zijn voor een ROI van 60×60 pixels, waarbij deze venstergrootte ge-

kozen werd om dezelfde reden als bij de coöccurrentiematrix. Er werden geen parameters verworpen. Eerst werd de classificatiefout berekend wanneer alle parameters werden gecombineerd. Daarna werd de laagste fout gezocht die kon bekomen worden door twee of drie parameters te combineren. Tenslotte werd ook de classificatiefout na PCA reductie tot twee of drie dimensies berekend.

Wanneer we verder over parameters f_1 tot f_4 spreken, bedoelen we de parameters bekomen uit het somhistogram. Parameters f_5 tot f_8 zijn dan de parameters berekend uit het verschilhistogram.

7.1.5 Coronale beelden

Gecompenseerde beelden Figuur 7.2 geeft een overzicht van de foutpercentages voor de verschillende parameters en de verschillende hoeken, bij gecompenseerde coronale beelden. Uit de figuur volgt dat de parameters uit het somhistogram de beste resultaten geven. Toch hebben sommige parameters uit het verschilhistogram nog nut, in combinatie met de parameters uit het somhistogram. Zoals te verwachten was zijn het ATM en de entropie sterk gecorreleerd, zowel voor het som- als voor het verschilhistogram ($\rho = 0.99$ en $\rho = 0.7$ respectievelijk). Dit was ook zo bij de coöccurrentiematrix. Verder blijkt uit Fig. 7.2 nog dat, net als bij de coöccurrentiematrix, iets betere resultaten bekomen worden voor een hoek van 45 graden.

Wanneer, voor een hoek van 45 graden, alle parameters gebruikt worden als invoer van de classifier krijgen we een perfecte classificatie. Wanneer we ons beperken tot twee of drie dimensies geven volgende combinaties de beste classificatie:

- Combinatie van f_1 , f_2 en f_3 uit het somhistogram resulteert in een foutpercentage van 2%.
- Combinatie van f_1 en f_3 uit het somhistogram leidt eveneens tot een classificatiefout van 2%.
- PCA reductie naar drie dimensies geeft een foutpercentage van 8%, reductie naar twee dimensies van 12%.

Niet gecompenseerde beelden Bij de niet gecompenseerde beelden is het (kleine) voordeel van een hoek van 45 graden verdwenen, zoals blijkt uit Fig. 7.3. De gemiddelde grijswaarde f_1 en het contrast f_3 van het somhistogram werken beter bij gecompenseerde beelden, de andere textuurparameters zijn ongeveer even goed, of zelfs iets beter, bij de niet gecompenseerde beelden. De hieronder vermelde resultaten werden berekend met een hoek van 45 graden.

Wanneer alle parameters gebruikt worden als invoer voor classificatie wordt 8% verkeerd geïdentificeerd. De minimaal haalbare fout is 2%, deze wordt bereikt door parameters f_1 , f_3 en f_4 van het somhistogram te combineren.

Besluit Compensatie is bij de coronale beelden gewenst, aangezien de classificatiefout daardoor van 2% naar 0% kan worden teruggebracht. De fouten zijn dezelfde als bij de coöccurrentiematrix. Voor coronale beelden is deze techniek dus even efficiënt.

7.1.6 Sagittale beelden

Gecompenseerde beelden Ook hier zijn de parameters die uit het somhistogram bereikt worden de beste. Uit Fig. 7.4 blijkt dat een hoek van 45 graden hier geen betere resultaten geeft, en dat een hoek van 90 graden zelfs optimaal is. Dit was ook zo bij de coöccurrentiematrix. De reden is hier weer dat de ROI's bij de sagittale beelden zowel links, rechts als in het midden van het beeld werden gekozen, zodat verschillende voorkeursrichtingen samen werden beschouwd.

Op basis van de gemiddelde grijswaarde is perfecte classificatie mogelijk. Andere parameters kunnen ook hier weer nuttig zijn om de afstand tussen de twee wolken te vergroten.

Niet gecompenseerde beelden Bij de sagittale beelden geven de parameters bereikt uit de niet gecompenseerde beelden opvallend betere resultaten, op de gemiddelde grijswaarde na. Er is niet duidelijk een betere hoek, zie Fig. 7.5. Wanneer we de parameters combineren is er voor alle hoeken foutloze classificatie mogelijk. Onderstaande resultaten zijn bereikt met een hoek van 45 graden.

- Classificatie op basis van alle parameters geeft een classificatiefout van 5%.
- Na PCA reductie naar drie dimensies wordt 29% verkeerd beoordeeld, reductie naar twee dimensies leidt tot een fout van 32%.
- De combinatie van de gemiddelde grijswaarde f_1 en het ATM van het verschilhistogram f_6 leidt voor alle hoeken tot perfecte classificatie.

Besluit Compensatie is niet noodzakelijk bij de sagittale beelden. Net als bij de coöccurrentiematrix is zowel met als zonder compensatie perfecte classificatie mogelijk. Bij de niet gecompenseerde beelden moeten we echter meerdere parameters combineren, bij de gecompenseerde beelden volstaat het gebruik van de gemiddelde grijswaarde.

7.1.7 Coronale en sagittale beelden

Gecompenseerde beelden Er is, net als bij de coöccurrentiematrix, weinig afhankelijkheid van de hoek, zie Figuur 7.6. We werken verder met een hoek van 45 graden.

- Wanneer we alle parameters als input voor de classifier gebruiken wordt 8% verkeerd geclassificeerd.
- PCA reductie naar drie dimensies levert ook een foutpercentage van 8% op, reductie naar twee dimensies leidt tot een fout van 22%.
- Een classificatiefout van 2% wordt bekomen wanneer de gemiddelde grijswaarde f_1 , het contrast van het somhistogram f_3 en het ATM van het verschilhistogram f_6 worden gebruikt.

Niet gecompenseerde beelden Ook hier merken we op dat de gemiddelde grijswaarde bij niet gecompenseerde beelden minder goed werkt, terwijl de andere parameters beter zijn, zie Fig. 7.6 en Fig. 7.7. Er is weer geen hoek die opvallend betere resultaten geeft. De resultaten die hieronder worden vermeld, werden bekomen met een hoek van 45 graden. Er is een minimale classificatiefout van 11% haalbaar. Deze wordt bekomen door:

- Alle acht de parameters als input van de classifier te gebruiken.
- De gemiddelde grijswaarde, en de twee entropieën te combineren

Dit is veel slechter dan bij de gecompenseerde beelden, waar een fout van 2% werd gehaald.

Besluit Wanneer sagittale en coronale beelden niet gescheiden worden leidt compensatie tot veel betere resultaten. Dit is niet verwonderlijk: de textuurparameters werken minder goed wanneer coronale en sagittale beelden samen beschouwd worden, omdat de texturen te verschillend zijn. De belangrijkste parameters voor goede classificatie zijn dus de eerste orde parameters, die veel beter werken bij de gecompenseerde beelden (zie ook 6.7.1).

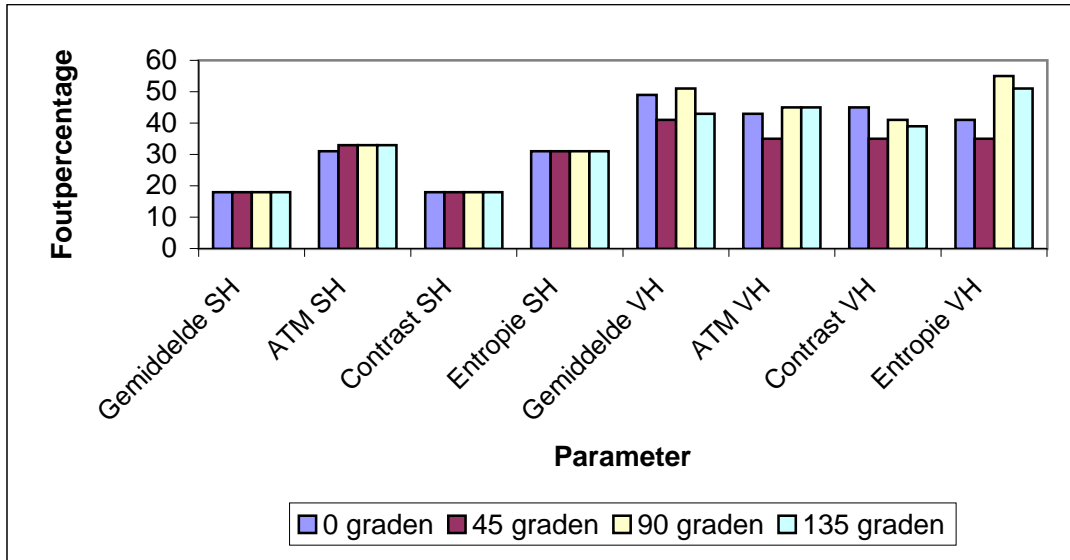
Bij de gecompenseerde beelden is met deze methode een lagere classificatiefout mogelijk dan met de coöccurrentiematrix (classificatiefout van 2.3% t.o.v. 3.4%). Door de vele beperkingen van onze testset (zie 2.2.3) moeten deze resultaten echter met een korreltje zout genomen worden: een verschil van 1% is verre van voldoende om te besluiten dat de ene methode beter werkt dan de andere!

7.1.8 Som- en verschilhistogrammen: algemeen besluit

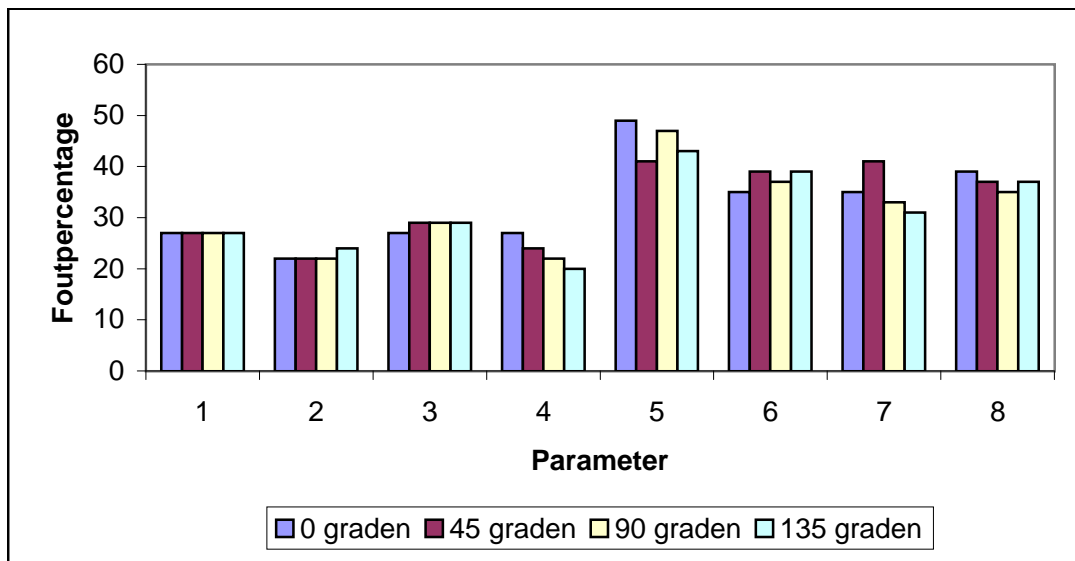
Er is even goede classificatie mogelijk met de som- en verschilhistogrammen als met de coöccurrentiematrix, zie Tabel 7.2. Het voordeel van de som- en verschilhistogrammen is dat ze sneller kunnen worden berekend. Net als bij de coöccurrentiematrix leidt het compensatie-algoritme tot betere resultaten, vooral wanneer men de coronale en sagittale beelden samen beschouwt. Enkel de gecompenseerde coronale beelden geven iets betere resultaten bij een hoek van 45 graden. Bij de andere beelden is de invloed van de hoek op de efficiëntie van de classificatie veel kleiner.

Tabel 7.2: Vergelijking van de minimale foutpercentages bekomen met de coöccurrentiematrix en met de som- en verschilhistogrammen. g=gecompenseerd, ng=niet gecompenseerd.

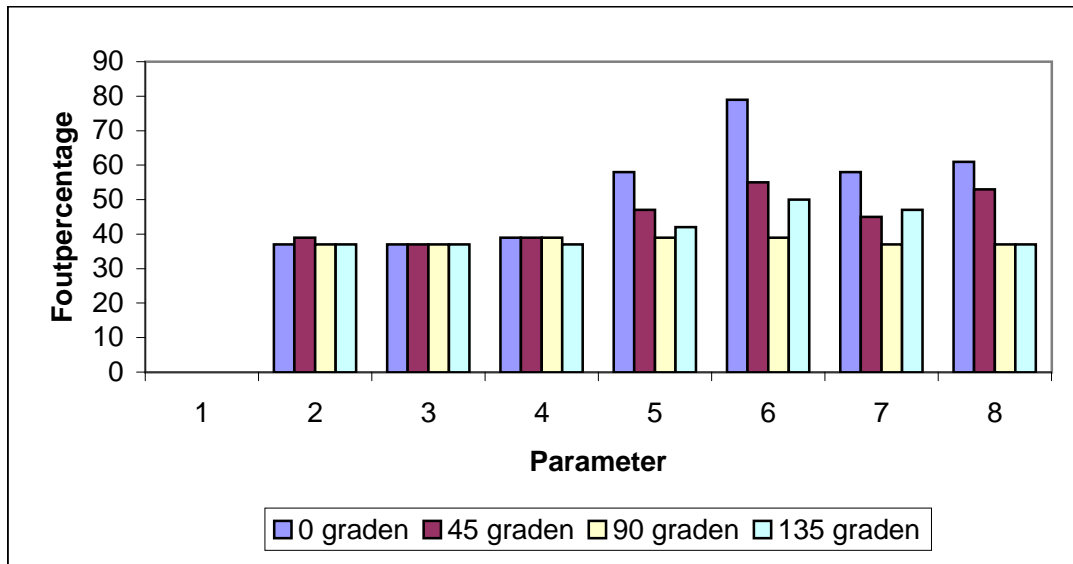
<i>Vergelijking</i> <i>Foutpercentages</i>	Coronaal		Sagittaal		Beide	
	g	ng	g	ng	g	ng
Coöccurrentiematrix	0	2	0	0	3	9
Som- en verschilhistogrammen	0	2	0	0	2	11



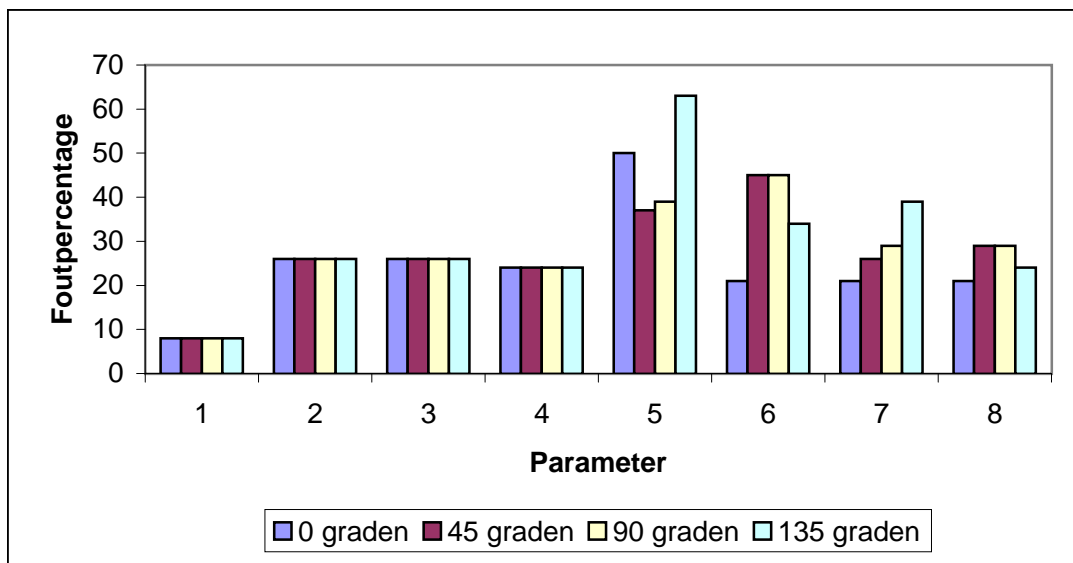
Figuur 7.2: Vergelijking van de foutpercentages bij de gecompenseerde coronale beelden, voor verschillende waarden van de hoek en een ROI van 60x60 pixels.



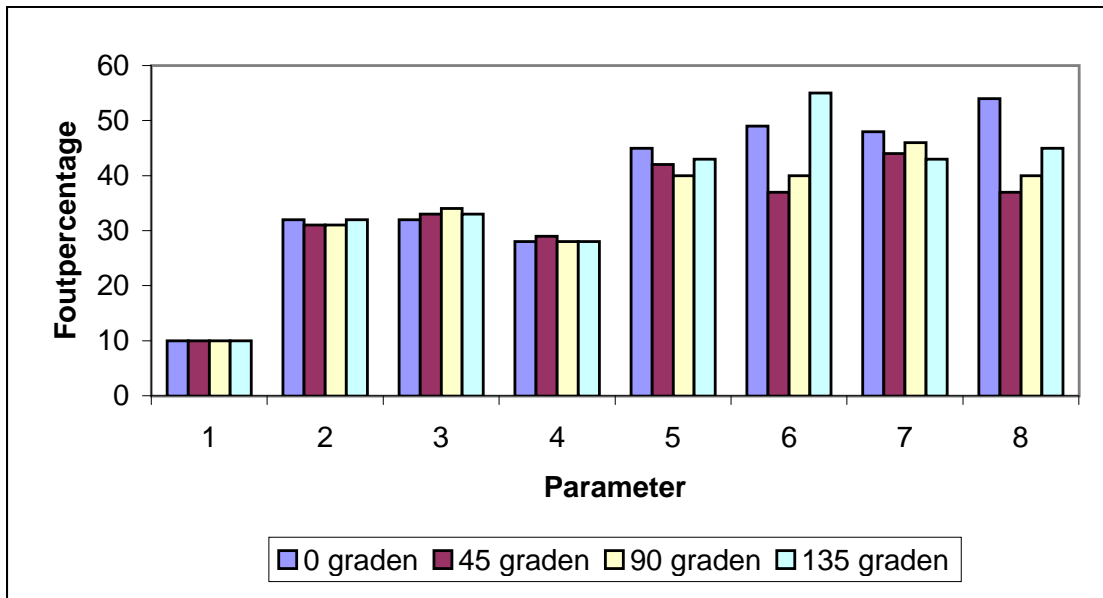
Figuur 7.3: Vergelijking van de foutpercentages bij de niet gecompenseerde coronale beelden, voor verschillende waarden van de hoek en een ROI van 60x60 pixels.



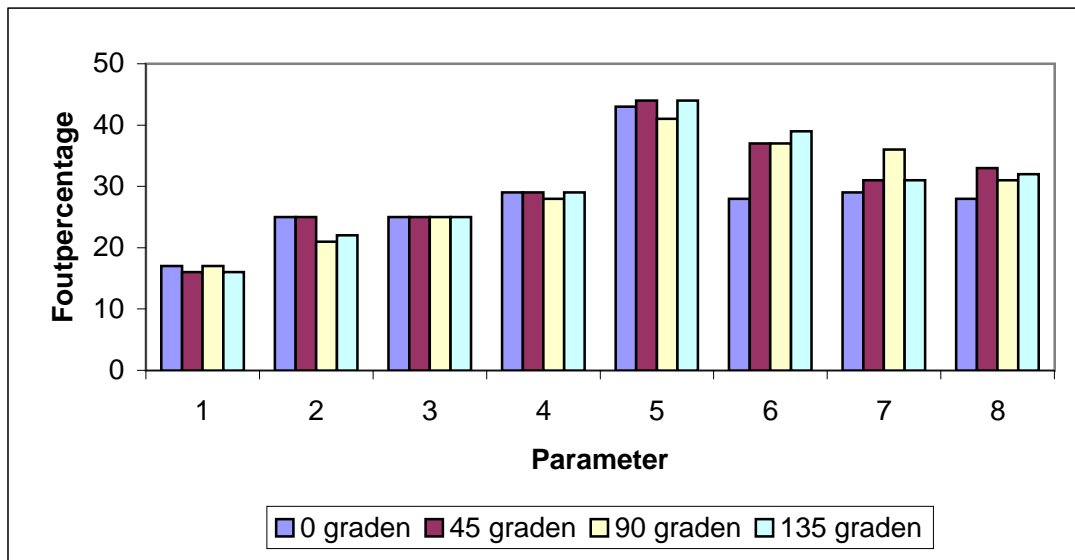
Figuur 7.4: Vergelijking van de foutpercentages bij de gecompenseerde sagittale beelden, voor verschillende waarden van de hoek en een ROI van 60x60 pixels.



Figuur 7.5: Vergelijking van de foutpercentages bij de niet gecompenseerde sagittale beelden, voor verschillende waarden van de hoek en een ROI van 60x60 pixels.



Figuur 7.6: Vergelijking van de foutpercentages bij de niet gecompenseerde coronale en sagittale beelden, voor verschillende waarden van de hoek en een ROI van 60x60 pixels.



Figuur 7.7: Vergelijking van de foutpercentages bij de gecompenseerde coronale en sagittale beelden, voor verschillende waarden van de hoek en een ROI van 60x60 pixels.

7.2 Coöccurrentie met het gemiddelde van de omliggende pixels

7.2.1 Definitie en voorbeeld

Dapeng en Zhongrong ontwikkelden in 1986 een *gewijzigde coöccurrentiematrix* $P(g,d)$ [5]. Deze matrix beschrijft de coöccurrentie tussen de grijswaarde g in een centrale pixel en een grijswaardeverschil d . Dit grijswaardeverschil is het verschil tussen de centrale grijswaarde en de gemiddelde grijswaarde van de acht aangrenzende pixels.

De methode wordt geïllustreerd in Fig. 7.8 en Fig. 7.9. De grijswaarden van het beeld bevinden zich tussen '0' en '3'. Er kunnen vier grijswaarden optreden, en dus hebben we vier rijen. De eerste rij telt de pixels met grijswaarde '0', de tweede telt de pixels met grijswaarde '1', enz.

Wanneer we bijvoorbeeld de pixel op de tweede rij en de tweede kolom nemen, zien we dat de gemiddelde grijswaarde van de burens '1' is. Het verschil tussen de grijswaarde van de centrale pixel en de gemiddelde grijswaarde van zijn burens bedraagt dus '-1'. Dit houdt in dat we het element op de rij die met grijswaarde '0' overeenkomt (de eerste rij) en de kolom die overeenkomt met een verschil '-1' de waarde met '1' verhogen. Voorgaande methode passen we toe op de vier middelste pixels, en daaruit volgt dan de gewijzigde coöccurrentiematrix. Uiteraard kunnen we voor de randpixels geen waarde berekenen. Dit vormt geen probleem wanneer we voldoende grote beelden bestuderen, wat bij ons het geval is (minstens 30x30 pixels).

7.2.2 Parameters

Uit de gewijzigde coöccurrentiematrix worden 13 textuurparameters berekend. Deze parameters worden opgesomd in Tabel 7.3. Hierbij is N het aantal effectief gebruikte pixels, G de maximale grijswaarde en D het maximale verschil.

2	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Figuur 7.8: De grijswaarden van de pixels in een gebied

		verschil					
			↓				
		...	-1	0	1	...	
grijswaarde	{	0	...	1	0	0	...
		1	...	0	1	0	...
		2	...	0	1	1	...
		3	...	0	0	0	...

Figuur 7.9: Gewijzigde coöcurrentiematrix

7.2.3 Toepassing op echografiebeelden: implementatie

Implementatie De benodigde rekestijd werd verminderd door eerst voor elke pixel de gemiddelde waarde van de burens te berekenen, door een tweedimensionale convolutie met het filter:

$$\frac{1}{8} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (7.5)$$

Om de rekestijd verder te reduceren werden de dimensies van de gewijzigde coöcurrentiematrix tot een minimum herleid. Dit werd gedaan door enkel aan de grijswaarden g tussen de minimale en maximale grijswaarde een rij in de coöcurrentiematrix toe te kennen, en door eerst het kleinste en grootste verschil te zoeken, zodat het aantal kolommen minimaal kon gemaakt worden.

7.2.4 Coronale beelden

Gecompenseerde beelden De gewijzigde coöcurrentiematrix is uiteraard hoekonafhankelijk. Bij tien van de dertien parameters gaf een venster van 60x60 pixels het laagste foutpercentage, zie Fig. 7.10. Dit is de reden waarom vanaf hier dan ook met dit venster wordt verdergewerkt.

Ook met deze methode is perfecte classificatie mogelijk, door het combineren van drie parameters.

- Parameters f_4 , f_5 , f_6 en f_8 zijn sterk gecorreleerd ($|\rho| > 0.93$). De combinatie van een van deze parameters met f_{12} en f_{13} leidt tot foutloze classificatie.
- Bij de vorige methodes werden steeds goede resultaten bereikt op basis van de gemiddelde grijswaarde en de variantie op de grijswaarde. Ook hier is dit zo: combineren van f_1 , f_2 en f_{11} geeft ook een foutpercentage van 0%.

Tabel 7.3: Parameters berekend uit de gewijzigde coöcurrentiematrix

Parameter	Definitie
Gemiddelde grijswaarde	$f_1 = \frac{1}{N} \sum_g g \sum_d P(g, d)$
Variantie van de grijswaarde	$f_2 = \frac{1}{N} \sum_g (g - f_1)^2 \sum_d P(g, d)$
Gemiddeld verschil	$f_3 = \frac{1}{N} \sum_d d \sum_g P(g, d)$
Variantie van het verschil	$f_4 = \frac{1}{N} \sum_d (d - f_3)^2 \sum_g P(g, d)$
Emphasis klein verschil	$f_5 = \frac{1}{N} \sum_g \sum_d \frac{P(g, d)}{d^2}$
Emphasis groot verschil	$f_6 = \frac{1}{N} \sum_g \sum_d d^2 P(g, d)$
Distributie van de grijswaarde	$f_7 = \frac{1}{N} \sum_g \left(\sum_d P(g, d) \right)^2$
Distributie van het verschil	$f_8 = \frac{1}{N} \sum_d \left(\sum_g P(g, d) \right)^2$
Angulair tweede moment	$f_9 = \frac{1}{N^2} \sum_g \sum_d (P(g, d))^2$
Entropie	$f_{10} = \frac{-1}{\log(\frac{1}{GD})} \sum_g \sum_d P(g, d) \log\left(\frac{P(g, d)}{N}\right)$
Contrast	$f_{11} = \frac{1}{N} \sum_g \sum_d (g - d)^2 P(g, d)$
Invers verschilmoment	$f_{12} = \frac{1}{N} \sum_g \sum_d \frac{1}{1+(g-d)^2} P(g, d)$
Covariantie	$f_{13} = \frac{1}{N f_2 f_4} \sum_g \sum_d (g - f_1)(d - f_3) P(g, d)$

- Wanneer in twee dimensies gewerkt wordt is een classificatiefout van 2% haalbaar: deze wordt bekomen na PCA reductie tot twee dimensies, of door combineren van f_1 en f_{12} .

Niet gecompenseerde beelden In Fig. 7.11 worden de foutpercentages van de parameters voor wel en niet gecompenseerde beelden met elkaar vergeleken. Uit deze figuur blijkt duidelijk dat de foutpercentages van de parameters bij deze methode enorm afhankelijk zijn van het compensatie-algoritme, veel meer dan bij de andere methodes.

Net als bij de voorgaande methodes is bij de niet gecompenseerde coronale beelden geen perfecte classificatie mogelijk. De laagste classificatiefout wordt bekomen door f_1 , f_2 en f_7 te gebruiken. In dat geval wordt 4% van de beelden verkeerd beoordeeld. Wanneer

alle parameters als invoer van de classifier worden gebruikt bedraagt de classificatiefout 6%. Deze resultaten zijn iets minder dan deze bij de andere methodes.

Besluit Bij de coronale beelden geeft een venster van 60x60 pixels de beste resultaten. Ook hier is compensatie voordelig. Bij de gecompenseerde beelden is perfecte classificatie mogelijk, bij de niet gecompenseerde beelden is het laagst haalbare foutpercentage 4%. Voor de coronale beelden is deze methode dus zeker niet beter dan de coöccurrentiematrix.

7.2.5 Sagittale beelden

Gecompenseerde beelden Ook bij de sagittale beelden worden de laagste foutpercentages behaald met vensters van 60x60 pixels. Bij de gecompenseerde beelden laat de gemiddelde grijswaarde weer foutloze classificatie toe.

Niet gecompenseerde beelden Ook bij de niet gecompenseerde beelden is perfect onderscheid mogelijk, namelijk door f_3 , f_5 , f_8 of f_{12} met f_{11} te combineren.

Besluit Aangezien zowel bij de gecompenseerde als de niet gecompenseerde beelden perfecte classificatie mogelijk is, kunnen we besluiten dat voor sagittale beelden de methode even efficiënt is als die van de coöccurrentiematrix.

7.2.6 Coronale en sagittale beelden

Gecompenseerde beelden Bij de gecompenseerde beelden is een foutpercentage van 2% haalbaar, door combinatie van f_1 , f_4 en f_{13} . In twee dimensies is een minimale fout van 8% mogelijk, door f_1 en f_{10} als ingang van de classifier te gebruiken. PCA reductie tot twee of drie dimensies leidt tot foutpercentages van 11.5%. We hebben dus meer baat bij het combineren van de juiste parameters.

Niet gecompenseerde beelden Bij de niet gecompenseerde beelden wordt in het beste geval een foutpercentage van 9% bekomen, door f_4 , f_9 en f_{11} te gebruiken. De laagst haalbare classificatiefout in twee dimensies bedraagt 13%, en wordt verkregen door f_4 en f_{11} te gebruiken. Net als bij de gecompenseerde beelden heeft PCA hier weinig voordeel, reductie tot twee of drie dimensies leidt tot een foute classificatie van 15% van de beelden.

Besluit Wanneer geen onderscheid werd gemaakt tussen sagittale en coronale beelden, was bij de voorgaande methodes het laagst haalbare foutpercentage steeds groter dan wanneer we de beelden afzonderlijk bekeken. Ook bij deze methode is dit het geval. Bij

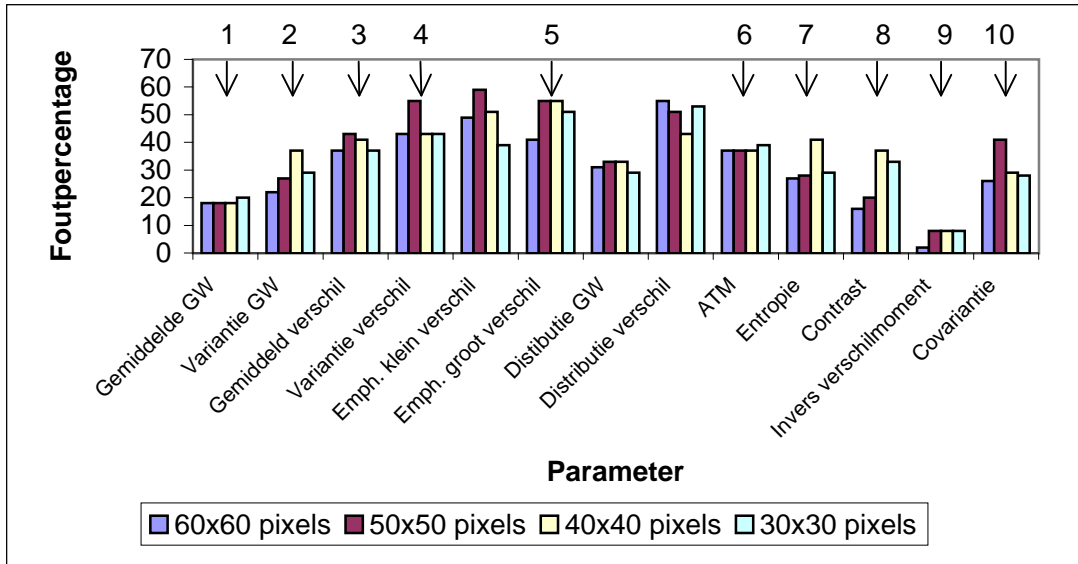
combinatie van de coronale en sagittale beelden zijn de foutpercentages ongeveer dezelfde als bij de coöccurrentiematrix.

7.2.7 Gewijzigde coöccurrentiematrix: algemeen besluit

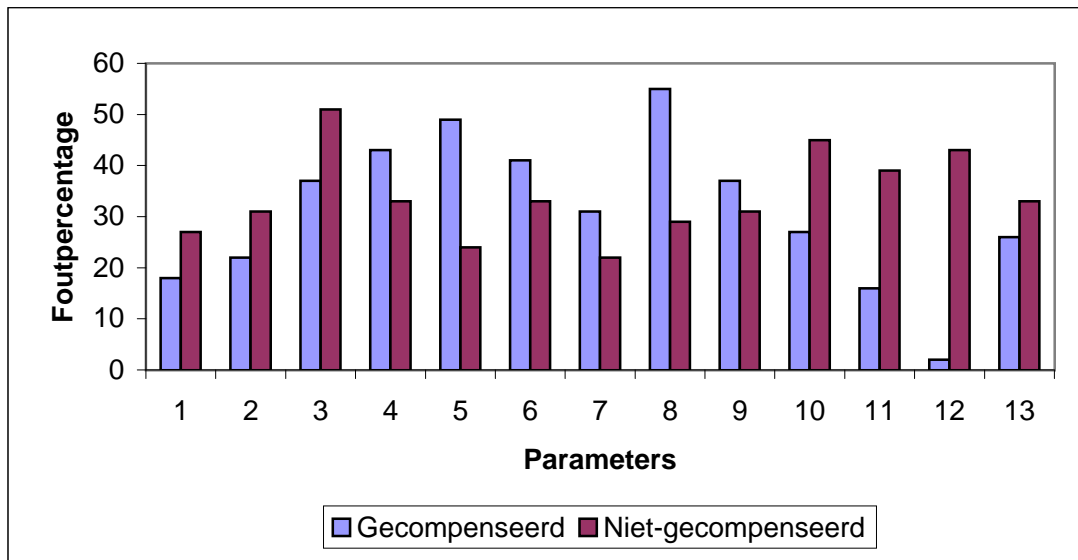
In Tabel 7.4 worden de laagst haalbare classificatiefouten voor de gewone en gewijzigde coöccurrentiematrix met elkaar vergeleken. Het blijkt duidelijk dat beide technieken aan elkaar gewaagd zijn. Wegens de beperktheid van onze dataset is het niet mogelijk om een ‘beste methode’ van de twee naar voor te schuiven.

Tabel 7.4: Vergelijking van de minimale foutpercentages bekomen met de gewone coöccurrentiematrix en met de gewijzigde coöccurrentiematrix. g=gecompenseerd, ng=niet gecompenseerd.

<i>Vergelijking</i> <i>Foutpercentages</i>	Coronaal		Sagittaal		Beide	
	g	ng	g	ng	g	ng
Gewone coöccurrentiematrix	0	2	0	0	3	9
Gewijzigde coöccurrentiematrix	0	4	0	0	2	9



Figuur 7.10: Vergelijking van de foutpercentages voor de vier verschillende venstergroottes, bij gecompenseerde coronale beelden.



Figuur 7.11: Vergelijking van de foutpercentages voor wel en niet gecompenseerde coronale beelden, ROI=60x60 pixels.

7.3 Amelungs statistische parameters

7.3.1 Definitie

De onderzoeker Amelung ontwikkelde een systeem, AST genoemd, om statistische parameters te berekenen uit zeven parametersets. Wij beschouwen hier enkel de parameters die afgeleid zijn van het *grijswaardehistogram* en de *gradiënthistogrammen*.

Er zijn twee van deze gradiënthistogrammen: één voor de gradiënt in de x-richting, en één voor de gradiënt in de y-richting. Vooraleer men de afgeleide in horizontale richting berekent zal men de ruis onderdrukken door het beeld in verticale richting te laagdoorlaatfilteren met het masker:

$$\frac{1}{4} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad (7.6)$$

Daarna wordt de afgeleide in horizontale richting berekend door convolutie met:

$$\frac{1}{2} \begin{pmatrix} -1 & 0 & 1 \end{pmatrix} \quad (7.7)$$

De combinatie van beide maskers komt overeen met het filteren met het volgende masker:

$$\frac{1}{8} \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad (7.8)$$

Van het gefilterde beeld zal met dan het histogram berekenen, en zo verkrijgt men het gradiënthistogram in de horizontale richting. Wanneer men de gradiënt in verticale richting wil berekenen gebruikt men de getransponeerde van (7.8). Deze twee filtermaskers worden ook wel *Sobeloperatoren* genoemd, zie ook [10], blz. 110 e.v.

7.3.2 Parameters

Eens de drie histogrammen opgesteld zijn, worden uit elk histogram 6 parameters berekend, deze zijn terug te vinden in Tabel 7.5. De meeste parameters berekend uit het grijswaardehistogram zijn terug te vinden bij de andere methodes. De reden om deze methode toch te onderzoeken is het feit dat de gradiënthistogrammen misschien goede classificatieresultaten kunnen opleveren.

7.3.3 Coronale beelden

Voor de afzonderlijke parameters worden weer net iets lagere classificatiefouten behaald met een venster van 60x60 pixels, daarom werken we vanaf hier met dit venster verder.

Tabel 7.5: Statistische parameters volgens Amelungs methode

Afkortingen	
N : aantal pixels	
G : aantal verschillende grijswaarden in het beeld	
Parameter	Definitie
Gemiddelde	$f_1 = \frac{1}{N} \sum_g P_H(g) \cdot g$
Variantie	$f_2 = \frac{1}{N} \sum_g P_H(g) \cdot (g - f_1)^2$
Derde moment	$f_3 = \frac{1}{N} \sum_g P_H(g) \cdot (g - f_1)^3$
Vierde moment	$f_4 = \frac{1}{N} \sum_g P_H(g) \cdot (g - f_1)^4$
Angular tweede moment	$f_5 = \frac{1}{N^2} \sum_g (P_H(g))^2$
Entropie	$f_6 = \frac{-1}{\log(\frac{1}{G})} \sum_g P_H(g) \cdot \log(\frac{P_H(g)}{N})$

Gecompenseerde beelden Een vergelijking van de classificatiefouten van de parameters bekomen uit de drie verschillende histogrammen is te vinden in Fig. 7.12. Uit deze figuur blijkt dat de gemiddelde grijswaarde en de variantie op de grijswaarde ook hier de laagste fouten geven, bovendien zijn deze parameters weinig gecorreleerd ($\rho = 0.18$). Classificatie op basis van deze twee parameters leidt tot een fout van 2%. Perfecte classificatie is mogelijk door deze parameters te combineren met een gepaste derde parameter. Zowel f_6 van het grijswaardehistogram, f_3 uit het gradiënthistogram in de x-richting, als f_5 uit het gradiënthistogram in de y-richting zijn geschikt als derde parameter.

Aangezien de combinatie f_1 , f_2 en f_6 uit het grijswaardehistogram foutloze classificatie mogelijk maakt, is de bijdrage van de gradiënthistogrammen hier miniem. Op basis van enkel de gradiënthistogrammen is een minimale fout van 12% mogelijk, door de twee gemiddeldes te combineren met het ATM van het gradiënthistogram in de y-richting.

Niet gecompenseerde beelden De gemiddelde grijswaarde en de variantie op de grijswaarde geven bij de niet gecompenseerde beelden weer hogere classificatiefouten dan bij de gecompenseerde beelden. Er is echter nog steeds perfecte classificatie mogelijk met de parameters berekend uit het grijswaardehistogram, meerbepaald door combinatie van of f_1 ,

f_3 en f_6 of f_1 , f_4 en f_5 . Met twee parameters, f_1 en f_4 , is een fout van 4% mogelijk. Ook bij de niet gecompenseerde beelden zijn de gradiënthistogrammen dus niet noodzakelijk om tot goede resultaten te komen.

Besluit Bij de coronale beelden hebben we geen voordeel bij het toepassen van compensatie, aangezien ook de niet gecompenseerde beelden foutloos kunnen geïdentificeerd worden. De gradiënthistogrammen zijn hier overbodig, aangezien de parameters uit het grijswaardehistogram volstaan om perfecte classificatie mogelijk te maken.

7.3.4 Sagittale beelden

Gecompenseerde beelden De gecompenseerde sagittale beelden kunnen weer perfect geïdentificeerd worden op basis van de gemiddelde grijswaarde.

Niet gecompenseerde beelden Ook bij de niet gecompenseerde beelden is een foutpercentage van 0% mogelijk, hiervoor hebben we echter de gradiënthistogrammen nodig (combinatie van gemiddelde grijswaarde en f_5 en f_6 uit het gradiënthistogram in de y-richting). Wanneer we enkel de parameters uit het grijswaardehistogram gebruiken bedraagt de laagst haalbare fout 8%, dit door combinatie van f_1 , f_3 en f_5 .

Besluit Bij de sagittale beelden is zowel met als zonder compensatie foutloze classificatie mogelijk. Het voordeel van compensatie is hier dat de gradiënthistogrammen dan niet moeten berekend worden.

7.3.5 Coronale en sagittale beelden

Gecompenseerde beelden Bij de gecompenseerde beelden is er een minimale fout mogelijk van 1%, door de gemiddelde grijswaarde te combineren met het ATM in de x-richting en het vierde moment in de y-richting. Wanneer we enkel parameters uit het grijswaardehistogram zouden gebruiken is er nog steeds een fout van slechts 3% mogelijk, door f_1 , f_2 en f_6 te combineren.

Niet gecompenseerde beelden De laagst haalbare classificatiefout ligt bij niet gecompenseerde beelden iets hoger dan bij de gecompenseerde beelden. Het laagste foutpercentage bedraagt hier 6%, en wordt verkregen door gebruik te maken van de gemiddelde grijswaarde, en het gemiddelde en de variantie van het gradiënthistogram in de y-richting. Wanneer enkel gebruik wordt gemaakt van parameters uit het grijswaardehistogram is er een classificatiefout van 14% haalbaar, door, net als bij de gecompenseerde beelden, f_1 , f_2 en f_6 te combineren.

Tabel 7.6: Vergelijking van de minimale foutpercentages bekomen met de coöccurrentiematrix en met Amelungs statistische methode. g=gecompenseerd, ng=niet gecompenseerd.

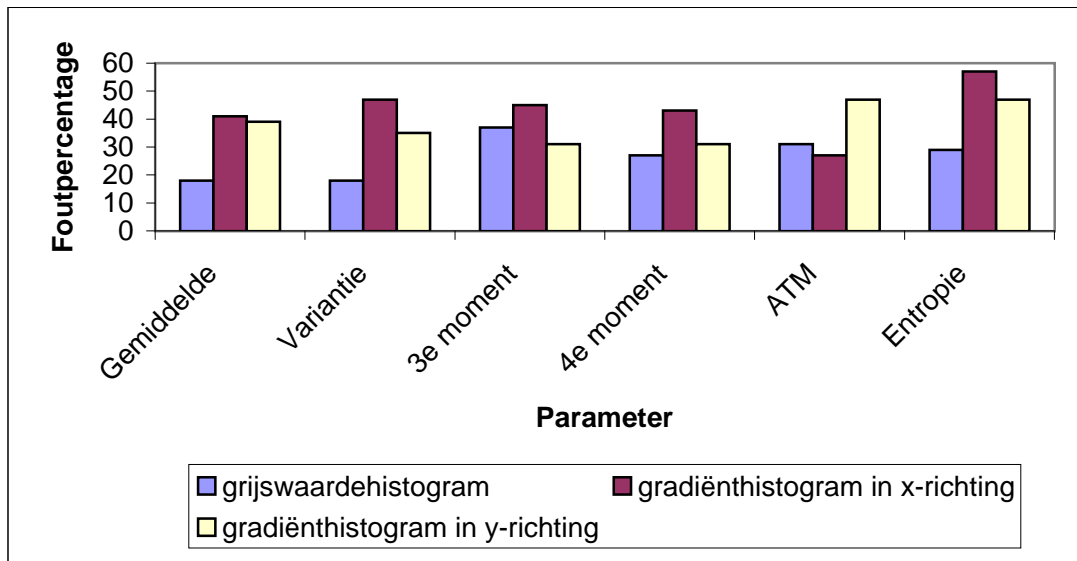
<i>Vergelijking</i> <i>Foutpercentages</i>	Coronaal		Sagittaal		Beide	
	g	ng	g	ng	g	ng
Coöccurrentiematrix	0	2	0	0	3	9
Statistische methode	0	0	0	0	1	6

Besluit Wanneer geen onderscheid wordt gemaakt tussen coronale en sagittale beelden, bekomen we de beste resultaten door parameters uit het grijswaardehistogram en de gradiënthistogrammen te combineren. In tegenstelling tot de situatie waarin de beelden apart bestudeerd werden zijn de gradiënthistogrammen hier dus wel nuttig.

7.3.6 Statistische parameters: algemeen besluit

Ook hier is het voordelig om de coronale en sagittale beelden apart te bestuderen. Wanneer men ze apart bestudeert volstaan de parameters uit het grijswaardehistogram. De parameters berekend uit de gradiënthistogrammen zijn echter nuttig om de classificatiefout naar beneden te halen wanneer coronale en sagittale beelden samen beschouwd worden. Net als bij de vorige technieken levert compensatie ook hier iets lagere foutpercentages op.

In Tabel 7.6 worden de laagst haalbare fouten vergeleken met die van de coöccurrentiematrix. We kunnen besluiten dat deze techniek gewaagd is aan de coöccurrentiematrix, en misschien zelfs beter is. Dit vermoeden kan enkel worden bevestigd door herhaling van de experimenten met een betere dataset.



Figuur 7.12: Vergelijking van de foutpercentages van de drie verschillende histogrammen, voor de gecompenseerde coronale beelden, ROI=60x60 pixels.

Hoofdstuk 8

Galloway's looplengtematrix

8.1 Definitie

Een vierde methode is die gebaseerd op de *looplengtematrix* (Engels: runlength matrix) [7]. De achterliggende idee is dat de lengte van bepaalde vormen kan gebruikt worden om de textuur te beschrijven. We spreken van een *looplengte r met grijswaarde g* als er r pixels met dezelfde grijswaarde g in een bepaalde richting een geconnecteerd pad vormen. Looplengtes zijn dus hoekafhankelijk, in de praktijk beperkt men zich weer tot hoeken van 0, 45, 90 en 135 graden.

De looplengtematrix P voor de richting ϕ wordt gedefinieerd als:

$$P_{\phi}(g, r) = (a_{g,r}) \quad (8.1)$$

Hierbij is $a_{g,r}$ het aantal geconnecteerde pixelintervallen van looplengte r in de richting ϕ waarvan alle pixels grijswaarde g hebben.

8.2 Voorbeeld

In Fig. 8.1 hebben we een beeldmatrix, in Fig. 8.2 de bijhorende looplengtematrix voor een hoek ϕ van nul graden. Deze bekomen we als volgt:

- De grijswaarden van het beeld bevinden zich tussen '0' en '3'. Er kunnen vier grijswaarden optreden, en dus hebben we vier rijen. De eerste rij telt de pixels met grijswaarde '0', de tweede telt de pixels met grijswaarde '1', enz.
- Een looplengte kan nooit groter dan vier zijn, omdat de grootste afmeting van het beeld vier is. Er zijn dus vier kolommen.
- We beginnen linksboven het beeld. Op de eerste rij staat er een pixel met grijswaarde '2'. De rechterbuur, dit is de pixel onder een hoek van nul graden, heeft een

2	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Figuur 8.1: De grijswaarden van de pixels in een gebied

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Figuur 8.2: Looplengetmatrix voor een hoek ϕ van nul graden

verschillende grijswaarde, we hebben dus een looplengte van 1. We verhogen het element op de derde rij (want grijswaarde ‘2’) en de eerste kolom (want looplengte ‘1’) met 1.

- De volgende pixel heeft grijswaarde ‘0’. Aangezien zijn rechterbuur een van ‘0’ verschillende grijswaarde heeft, hebben we weer een looplengte van ‘1’. We verhogen het element op de eerste rij (want grijswaarde ‘0’), en de eerste kolom (want looplengte 1) met 1.
- Tenslotte hebben we nog een looplengte van ‘2’, want we hebben twee enen naast elkaar. Bijgevolg moeten we het element op de tweede rij (want grijswaarde ‘1’) en de tweede kolom (want looplengte ‘2’) met 1 verhogen.
- Nu gaan we over naar de volgende rij, en overlopen op deze manier het volledige beeld.

8.3 Parameters

In de praktijk wordt de looplengetmatrix berekend voor hoeken van 0, 45, 90 en 135 graden. Om voldoende hoge looplengtes te bekomen wordt het aantal grijswaarden meestal gereduceerd.

De gebruikte parameters zijn terug te vinden in Tabel 8.1. Parameters f_1 tot f_5 komen uit het oorspronkelijke artikel van Galloway [7], een motivatie voor het gebruik van f_6 tot

f_{11} is te vinden in [13]¹

De *korte looplengte emphasis* is een parameter die de lage looplengtes bevoordeelt, en de hoge onderdrukt, de *lange looplengte emphasis* doet net het omgekeerde. De *korte looplengte emphasis* zal groot zijn als er veel kleine looplengtes met dezelfde intensiteit zijn, de *lange looplengte emphasis* zal groot zijn als er veel grote looplengtes met dezelfde grijswaarde zijn.

De *grijswaardedistributie* zal groot zijn als er maar weinig grijswaarden in het beeld voorkomen, of als een bepaalde grijswaarde dominant is. Hoe meer de grijswaarden verspreid zijn, hoe kleiner de grijswaardedistributie zal zijn. Dezelfde redenering kan gemaakt worden voor de looplengtedistributie, door 'grijswaarde' door 'looplengte' te vervangen.

De betekenis van *KGE*, of *Kleine Grijswaarde Emphasis*, is analoog aan die van *korte looplengte emphasis*, enkel worden hier kleine grijswaarden bevoordeeld i.p.v. kleine looplengtes. De *GLE*, of *Grote Grijswaarde Emphasis* zal kleine grijswaarden bestraffen en grote bevoordelen.

Parameters f_8 tot f_{11} zullen het bevoordelen van hoge/lage grijswaarden en lange/korte looplengtes combineren. Zo zal de *LLGGE*, de *Lange Looplengte en Grote Grijswaarde Emphasis*, een maat zijn voor het aantal lange looplengtes die bestaan uit hoge grijswaarden. Op eenzelfde manier kunnen f_9 tot f_{11} geïnterpreteerd worden.

8.4 Toepassing op echografiebeelden

8.4.1 Laagdoorlaatfilteren

De techniek van de looplengtematrix is gevoelig voor ruis. Omdat de echografiebeelden veel ruis bevatten, werden ze eerst door een laagdoorlaatfilter gestuurd, zodat de ruis gedeeltelijk uitgemiddeld werd. Hierna werd een grijswaardereductie gedaan, en dan pas hebben we de looplengtematrix berekend. Het uitmiddelen hebben we gedaan door convolutie met het masker:

$$filter = \frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (8.2)$$

8.4.2 Grijswaardereductie

De bestudeerde echografiebeelden hebben typisch ongeveer 50 verschillende grijswaarden, de grijswaarden van de gezonde beelden blijven meestal tussen 35-85, en die van de zieke

¹Bij sommige van deze extra parameters treedt een deling door nul op, namelijk bij $g = 0$. Dit hebben we vermeden door in onze implementatie g steeds te vervangen door $(g + 1)$.

Tabel 8.1: Parameters berekend uit de looppengtematrix

Afkortingen	
G : aantal grijswaarden in het beeld	
R : maximale looppengte in het beeld	
$N = \sum_{g=0}^{G-1} \sum_{r=1}^R P(g, r)$	
Parameter	Definitie
Korte looppengte emphasis	$f_1 = \frac{1}{N} \sum_{g=0}^{G-1} \sum_{r=1}^R \frac{P(g, r)}{r^2}$
Lange looppengte emphasis	$f_2 = \frac{1}{N} \sum_{g=0}^{G-1} \sum_{r=1}^R r^2 P(g, r)$
Grijswaardedistributie	$f_3 = \frac{1}{N} \sum_{g=0}^{G-1} \left(\sum_{r=1}^R P(g, r) \right)^2$
Looppengtedistributie	$f_4 = \frac{1}{N} \sum_{r=1}^R \left(\sum_{g=0}^{G-1} P(g, r) \right)^2$
Looppercentage	$f_5 = \frac{N}{N_{pix}}$
KGE	$f_6 = \frac{1}{N} \sum_{g=0}^{G-1} \sum_{r=1}^R \frac{P(g, r)}{g^2}$
GGE	$f_7 = \frac{1}{N} \sum_{g=0}^{G-1} \sum_{r=1}^R g^2 P(g, r)$
LLGGE	$f_8 = \frac{1}{N} \sum_{g=0}^{G-1} \sum_{r=1}^R g^2 r^2 P(g, r)$
LLKGE	$f_9 = \frac{1}{N} \sum_{g=0}^{G-1} \sum_{r=1}^R \frac{r^2}{g^2} P(g, r)$
KLHGE	$f_{10} = \frac{1}{N} \sum_{g=0}^{G-1} \sum_{r=1}^R \frac{g^2}{r^2} P(g, r)$
KLLGE	$f_{11} = \frac{1}{N} \sum_{g=0}^{G-1} \sum_{r=1}^R \frac{1}{g^2 r^2} P(g, r)$

tussen 40-90. Bij de oorspronkelijke beelden zijn de meeste looplengtes 1, 2 of 3 pixels lang, slechts zelden is er eens een looplengte van 5 of 6 pixels. Om het aandeel van grote looplengtes te verhogen zullen we het aantal grijswaarden reduceren. We mogen dit aantal echter niet te sterk reduceren omdat daardoor belangrijke informatie kan verloren gaan. In ons geval hebben we tot 8 of 16 grijswaarden gereduceerd. De grijswaardereductie hebben we op twee verschillende manieren geïmplementeerd:

1. Histogramegalisisatie Het aantal grijswaarden na reductie stellen we voor door G . Deze methode zal de oorspronkelijke grijswaarden g afbeelden op het interval $f(g) = [0G - 1]$, op zo een manier dat het histogram van het getransformeerde beeld zo uniform mogelijk is. De grijswaardetransformatie die we gebruiken hebben is (zie [10], blz. 42-44):

$$f(g) = \left\lfloor \frac{G}{N - h(0)} \sum_{n=1}^g h(n) \right\rfloor \quad (8.3)$$

Hierbij is $\lfloor x \rfloor$ het kleinste geheel getal dat kleiner is of gelijk aan x , $h(n)$ het histogram van het oorspronkelijke grijswaardebeeld en N het aantal pixels.

2. Schalen van de grijswaarden Hierbij verschuiven we de grijswaarden eerst zodat de minimale grijswaarde nul wordt, en gaan we vervolgens de grijswaarden schalen, zodat ze in het interval $[0 G-1]$ komen te liggen. Wanneer G_{min} en G_{max} respectievelijk de minimale en maximale grijswaarde voorstellen krijgen we:

$$f(g) = \text{round} \left[(G - 1) \times \frac{g - G_{min}}{G_{max} - G_{min}} \right] \quad (8.4)$$

8.4.3 Coronale beelden

Invloed van de ROI Zoals reeds eerder vermeld werd, is de looplengtematrix gevoelig aan ruis. Om de ruis zoveel mogelijk uit te middelen opteren we bij zowel de coronale als de sagittale beelden voor een venster van 60x60 pixels. De foutpercentages van de afzonderlijke parameters zijn voor de vier venstergroottes terug te vinden in Fig. 8.3. Uit deze figuur blijkt dat een ROI van 60x60 pixels bij de meeste parameters de laagste fout geeft.

Invloed van de hoek Aangezien de textuur van de coronale beelden duidelijk in een richting van 45 graden loopt, verwachten we de beste resultaten voor deze richting. Dat dit inderdaad zo is, blijkt uit Fig. 8.4. Deze figuur geeft de foutpercentages na grijswaardereductie tot 8 grijswaarden door histogramegalisisatie. Grijswaardereductie tot 16 grijswaarden, of met de schalingstechniek, levert analoge resultaten op.

Een hoek van 45 graden geeft enkel bij f_6 opvallend slechtere resultaten. Bij de verdere besprekingen zal deze parameter zich echter steeds als een buitenbeentje gedragen. Aangezien de parameter voor classificatie niet belangrijk is, is dit afwijkend gedrag hier niet erg.

We werken vanaf hier dus verder met een hoek van 45 graden.

Grijswaardereductie Een tweede vraag die zich stelt, is welke methode optimaal is voor grijswaardereductie: histogramegalitatie of schaling? Hierbij kunnen we ons verder nog de vraag stellen naar hoeveel grijswaarden we best reduceren, 8 of 16. De foutpercentages zonder grijswaardereductie worden in Fig. 8.5 vergeleken met deze bekomen na grijswaardereductie, voor een hoek van 45 graden. We zien dat reductie tot 8 grijswaarden door histogramegalitatie net iets betere resultaten geeft dan de andere methodes (met weer net het tegenovergestelde resultaat voor f_6).

Invloed van compensatie Tenslotte moet ook nog nagegaan worden of het compensatie-algoritme voordelig is bij deze methode. De vergelijking van de foutpercentages bij wel en niet gecompenseerde beelden wordt gegeven in Fig. 8.6. Uit deze figuur blijkt dat compensatie bij negen van de elf parameters gelijke of betere classificatie toelaat, hoewel de verschillen klein zijn. We kiezen dus voor de gecompenseerde beelden. De resultaten van de volgende paragraaf zullen nogmaals bevestigen dat compensatie beter werkt bij de coronale beelden.

Resultaten Uit de voorgaande beschouwingen volgt dus dat we het best werken met:

- Een venster van 60x60 pixels, omdat ruis dan maximaal wordt uitgemiddeld.
- Een hoek van 45 graden, omdat de textuur duidelijk in die richting loopt.
- Grijswaardereductie naar 8 grijswaarden door histogramegalitatie.
- Gecompenseerde beelden.

Parameters f_1 tot f_5 , dus de oorspronkelijk door Galloway gebruikte parameters, zijn in ons geval heel sterk gecorreleerd: de correlatie tussen f_1 en f_2 bedraagt 0.88, alle andere combinaties leveren correlatiewaarden hoger dan 0.93 op. Wanneer we enkel deze parameters zouden gebruiken is een minimale classificatiefout van slechts 16% mogelijk, door f_1 en f_4 te gebruiken. De extra parameters f_6 tot f_{11} zijn dus noodzakelijk om tot een betere classificatie te kunnen komen. Wanneer we de juiste parameters, namelijk f_3 , f_7 en f_9 , combineren is een foutpercentage van 8% mogelijk. Wanneer we ons tot

twee parameters beperken is de laagst haalbare fout 14% (door f_9 en f_{10} te combineren). PCA biedt hier weinig soelaas: reductie tot 3 dimensies geeft een fout van 26%, tot twee dimensies een fout van 24%.

Bij gebruik van de niet gecompenseerde beelden bedraagt de laagst haalbare classificatiefout 16%, door f_4 , f_6 en f_7 als invoer van de classifier te gebruiken. Dit is hoger dan bij de gecompenseerde beelden (waar 8% haalbaar was), dus is het inderdaad voordelig de gecompenseerde beelden te gebruiken.

8.4.4 Sagittale beelden

Invloed van de hoek Bij de voorgaande technieken was er meestal geen duidelijke voorkeursoek voor de sagittale beelden. Bij deze techniek leidt een hoek van 45 graden echter tot opvallend lagere classificatiefouten dan de andere hoeken, zie Fig. 8.7. We werken dus ook hier voort met een hoek van 45 graden.

Grijswaardereductie We bekomen analoge resultaten als bij de coronale beelden, en werken dus verder met grijswaardereductie tot 8 grijswaarden door histogramegalitatie.

Invloed van compensatie Bij de coronale beelden waren de foutpercentages bij de gecompenseerde beelden kleiner dan bij de niet gecompenseerde beelden. Bij de sagittale beelden zijn deze foutpercentages meer aan elkaar gewaagd. Dit wordt geïllustreerd in Fig. 8.8. Daarom zullen we verder werken met zowel de gecompenseerde als de niet gecompenseerde beelden.

Resultaten De parameters f_1 tot f_5 zijn weer sterk gecorreleerd, de correlatie tussen twee van deze parameters is steeds groter dan 0.95. Ook hier zullen de extra parameters dus nuttig zijn.

Bij de *gecompenseerde* beelden is een foutpercentage van 5% haalbaar, door f_1 , f_9 en f_{10} te combineren, of door f_7 en f_8 te gebruiken. Reductie van de parameterruimte door PCA leidt tot een minimale fout van 8% (reductie naar drie dimensies), PCA is hier dus ook niet voordelig.

Bij de *niet gecompenseerde* beelden wordt slechts 3% verkeerd beoordeeld wanneer we f_1 , f_9 en f_{10} gebruiken als invoer voor classificatie. Wanneer we in twee dimensies werken kunnen we een minimale fout van 5% halen door f_4 en f_5 te combineren.

8.4.5 Coronale en sagittale beelden

Aangezien de coronale en sagittale beelden apart de beste resultaten gaven met een hoek van 45 graden en grijswaardereductie door histogramegalitatie wordt hier ook met deze

instellingen gewerkt. Uit Fig. 8.9 blijkt dat de resultaten voor gecompenseerde en niet gecompenseerde beelden weer aan elkaar gewaagd zijn.

Bij de *gecompenseerde* beelden is een classificatiefout van 16% mogelijk, door f_9 en f_{10} te combineren, eventueel nog met f_1 erbij. PCA reductie leidt tot een minimaal haalbare fout van 23%.

Ook bij de *niet gecompenseerde* beelden resulteert de combinatie van f_9 en f_{10} in de laagste fout, hier 17%. Wanneer we alle beelden samen beschouwen heeft het compensatie-algoritme bij deze methode dus weinig voordeel.

8.4.6 Besluit

De beste resultaten worden bekomen voor:

- Een venster van 60x60 pixels.
- Een hoek van 45 graden.
- Grijswaardereductie naar 8 grijswaarden door histogramegalisisatie.

Het compensatie-algoritme is hier niet echt noodzakelijk, enkel bij de coronale beelden leidt het tot iets betere classificatie.

De oorspronkelijke vijf parameters f_1 tot f_5 volstaan hier niet om de textuur te beschrijven, omdat ze te sterk gecorreleerd zijn. De extra parameters f_6 tot f_{11} zijn veel minder gecorreleerd en zijn noodzakelijk voor een goede classificatie. Dat deze extra parameters ‘goed werken’ was ergens te verwachten: ze zijn, in tegenstelling tot de oorspronkelijke parameters, allemaal afhankelijk van de grijswaarden g (er staat steeds g^2 in de noemer of de teller). De grijswaarden van de zieke beelden liggen in het algemeen iets hoger dan die van de gezonde beelden, en de waarden van de parameters zullen hier dan ook voordelig (met oog op classificatie tussen ziek en gezond weefsel) door beïnvloed worden.

De vergelijking tussen de laagst haalbare foutpercentages bij de coöccurrentiematrix en bij de looplengtematrix is terug te vinden in Tabel 8.2.

We kunnen uit deze tabel besluiten dat de looplengtematrix geen bedreiging vormt voor de coöccurrentiematrix. Uit de tabel blijkt ook dat, louter op basis van textuureigenschappen, dus zonder gebruik te maken van de eerste orde statistieken, een vrij goede classificatie mogelijk is. De foutpercentages kunnen echter nog verlaagd worden door eerste orde statistieken zoals de gemiddelde grijswaarde op te nemen bij de parameters. De gemiddelde grijswaarde volgt gemakkelijk uit de looplengtematrix:

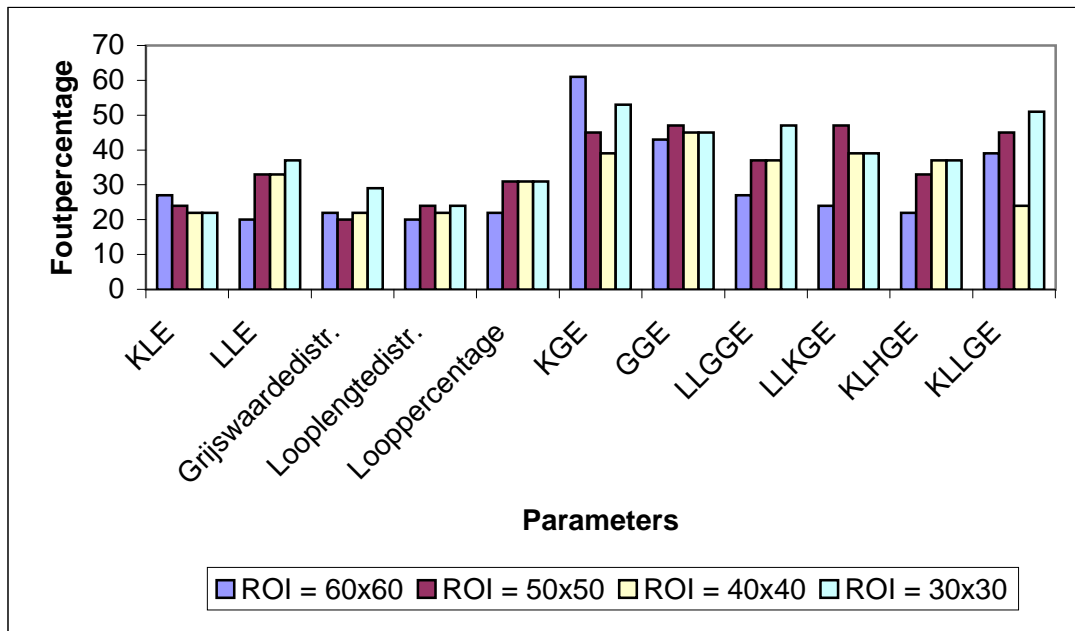
$$\mu = \frac{1}{N} \sum_{g=0}^{G-1} g \sum_{r=1}^R r P(g, r) \quad (8.5)$$

Tabel 8.2: Vergelijking van de minimale foutpercentages bekomen met de coöccurrentiematrix en de looplengtematrix, met en zonder gebruik van de gemiddelde grijswaarde als parameter. g=gecompenseerd, ng=niet gecompenseerd.

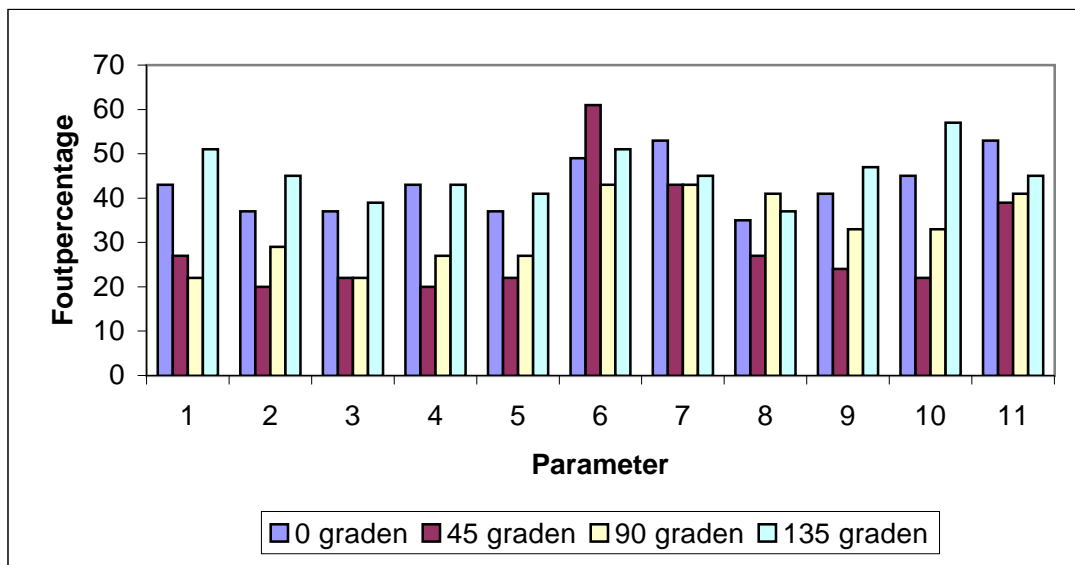
<i>Vergelijking foutpercentages</i>	Coronaal		Sagittaal		Beide	
	g	ng	g	ng	g	ng
Coöccurrentiematrix	0	2	0	0	3	9
Looplengtematrix (zonder gemiddelde grijswaarde)	8	16	5	3	16	17
Looplengtematrix (met gemiddelde grijswaarde)	4	10	0	0	4	5

Wanneer we de gemiddelde grijswaarde toevoegen aan de parameterset en de experimenten herhalen, merken we dat de classificatiefouten opvallend lager zijn dan zonder gemiddelde grijswaarde, zie Tabel 8.2. We gaan hier niet verder in op de ideale parametercombinaties, maar merken enkel op dat bij alle optimale parametercombinaties de gemiddelde grijswaarde betrokken was.

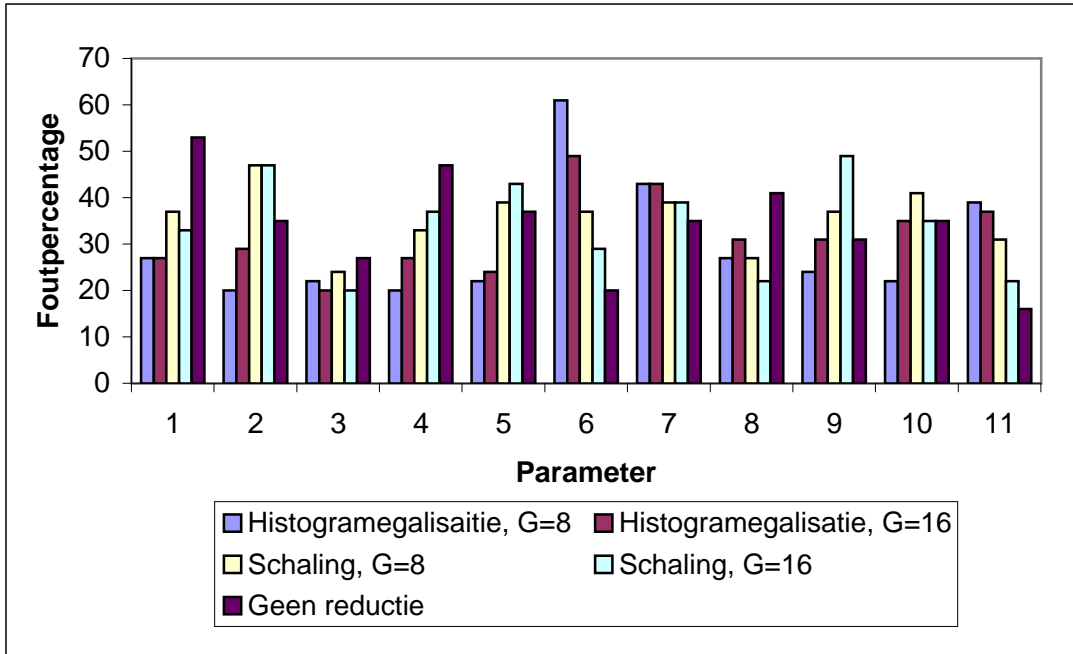
Om optimale classificatieresultaten te behalen zijn dus steeds eerste orde parameters nodig, vooral de gemiddelde grijswaarde is een belangrijke parameter. Wanneer enkel tweede orde textuurparameters gebruikt worden, is minder goede classificatie mogelijk. Dit kan betekenen dat de texturen van het zieke en gezonde weefsel niet genoeg verschillen om met enkel textuurparameters een optimale classificatie uit te voeren, ofwel dat de gebruikte technieken niet optimaal geschikt zijn om de textuur van het hersenweefsel te bestuderen.



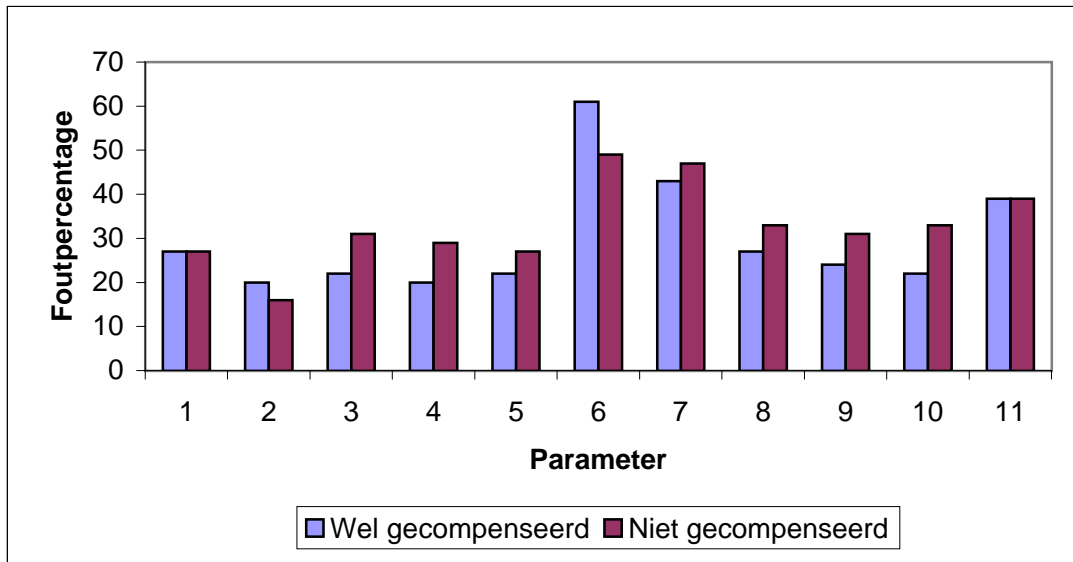
Figuur 8.3: Vergelijking van de foutpercentages voor de vier verschillende venstergroottes bij gecompenseerde coronale beelden. De beelden werden eerst door een laagdoorlaatfilter gestuurd, en het aantal grijswaarden werd door histogramegalisisatie tot 8 gereduceerd.



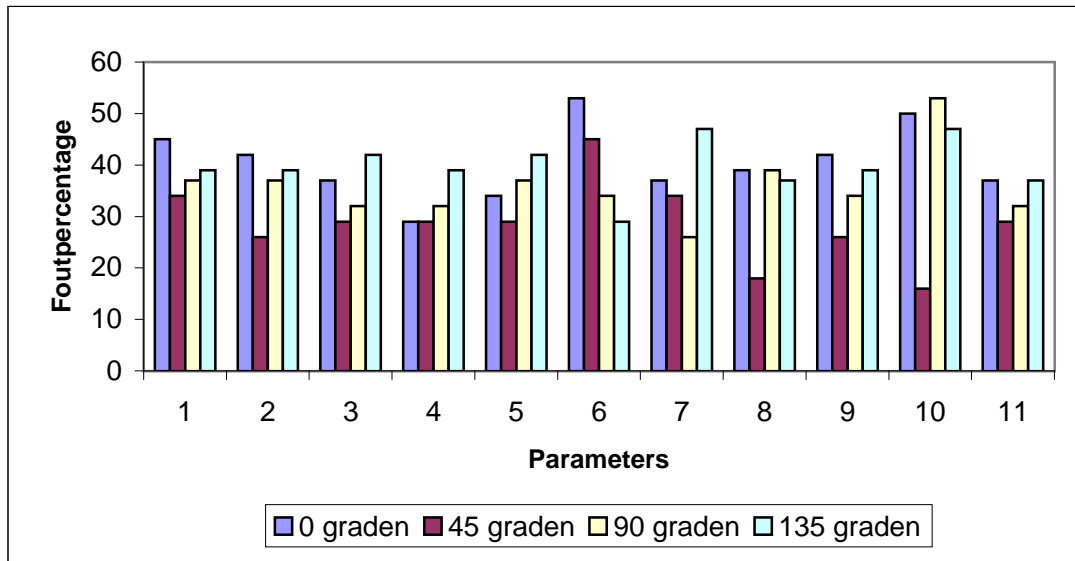
Figuur 8.4: Vergelijking van de foutpercentages voor de vier verschillende hoeken bij gecompenseerde coronale beelden. De beelden werden eerst door een laagdoorlaatfilter gestuurd, en het aantal grijswaarden werd door histogramegalisisatie tot 8 gereduceerd.



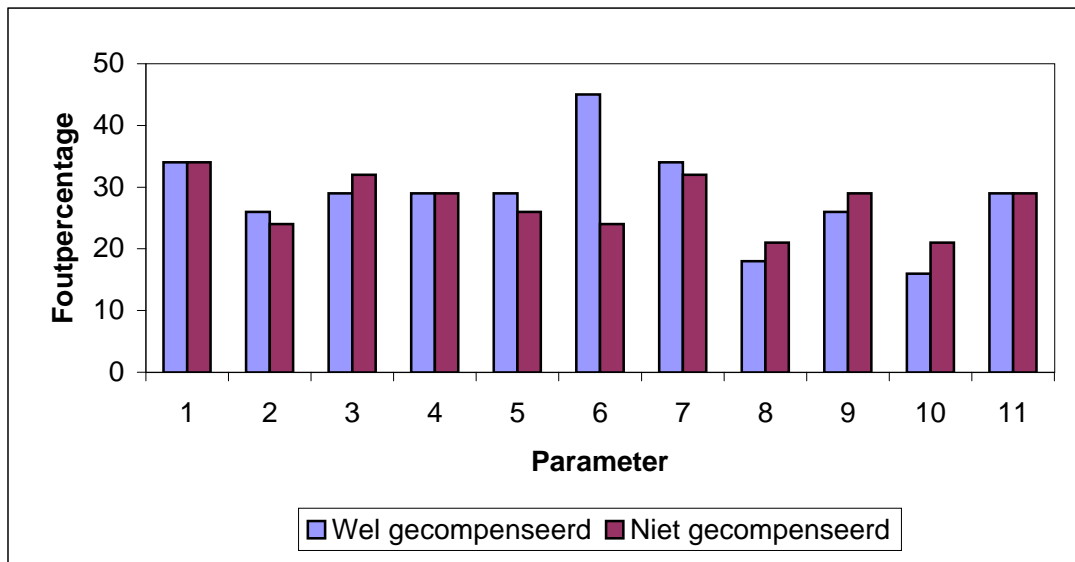
Figuur 8.5: Vergelijking van de foutpercentages zonder grijswaardereductie met de foutpercentages na grijswaardereductie tot 8 of 16 grijswaarden door histogramequalisatie en schaling, bij coronale beelden en een hoek van 45 graden.



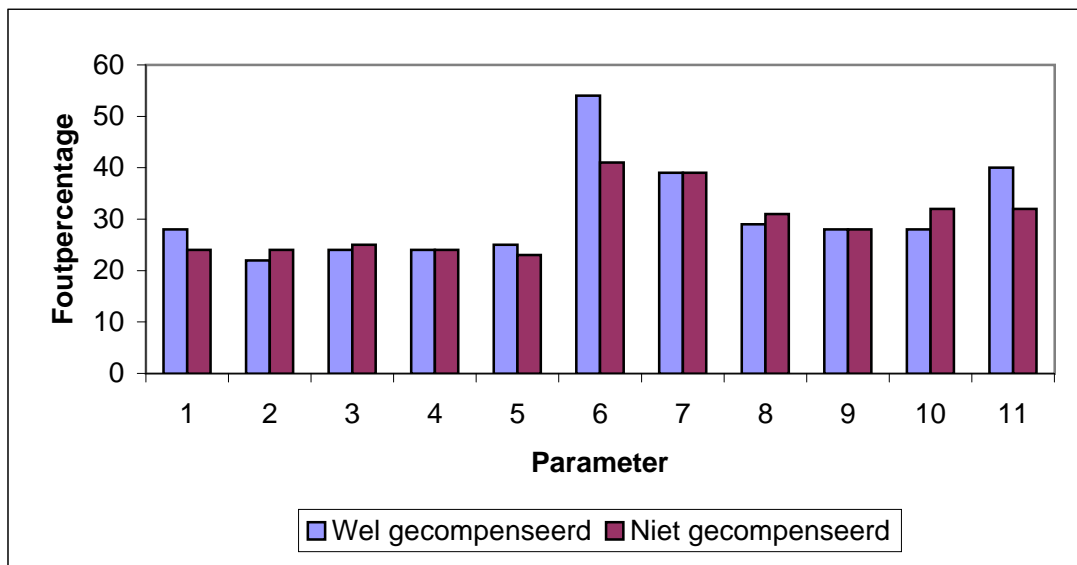
Figuur 8.6: Vergelijking van de foutpercentages bij wel en niet gecompenseerde coronale beelden, voor een hoek van 45 graden. De beelden werden eerst door een laagdoorlaatfilter gestuurd, en het aantal grijswaarden werd door histogramequalisatie tot 8 gereduceerd.



Figuur 8.7: Vergelijking van de foutpercentages voor de vier verschillende hoeken bij gecompenseerde sagittale beelden. De beelden werden eerst door een laagdoorlaatfilter gestuurd, en het aantal grijswaarden werd door histogramegalisatie tot 8 gereduceerd.



Figuur 8.8: Vergelijking van de foutpercentages bij wel en niet gecompenseerde sagittale beelden, voor een hoek van 45 graden. De beelden werden eerst door een laagdoorlaatfilter gestuurd, en het aantal grijswaarden werd door histogramegalisatie tot 8 gereduceerd.



Figuur 8.9: Vergelijking van de foutpercentages bij wel en niet gecompenseerde beelden, coronaal en sagittaal samen, voor een hoek van 45 graden. De beelden werden eerst door een laagdoorlaatfilter gestuurd, en het aantal grijswaarden werd door histogramegalisisatie tot 8 gereduceerd.

Hoofdstuk 9

Laws' filter maskers

De technieken die in de vorige hoofdstukken werden beschreven zijn allemaal gebaseerd op het tellen en vergelijken van grijswaarden. De technieken die in dit en het volgende hoofdstuk besproken worden zijn gebaseerd op het *filteren* van het beeld, wat overeenkomt met een tweedimensionale convolutie met een filtermasker. De textuureigenschappen volgen dan rechtstreeks uit het gefilterd beeld.

9.1 Definitie

Laws ontwikkelde 5 eendimensionale filtermaskers om textuur te beschrijven [9]. De maskers hebben de naam 'level', 'edge', 'spot', 'wave' en 'ripple', zie Tabel 9.1. De naam van de filtermaskers verklaart ook grotendeels de werking ervan. Zo zal het 'edge' masker grote grijswaarde-overgangen detecteren. Zo'n overgangen komen meestal overeen met randen, vandaar de naam 'edge'. De 'level' vector geeft een gewogen gemiddelde, de 'spot' vector is gevoelig voor heldere spots en de 'wave' en 'ripple' vectoren zullen golfvormen detecteren. Uit deze 5 eendimensionale maskers kunnen 25 tweedimensionale maskers geconstrueerd worden, bijvoorbeeld:

$$\mathbf{F}_{1e} = \mathbf{l}_5 \mathbf{e}_5^T = \begin{pmatrix} 1 \\ 4 \\ 6 \\ 4 \\ 1 \end{pmatrix} \begin{pmatrix} -1 & -2 & 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -2 & 0 & 2 & 1 \\ -4 & -8 & 0 & 8 & 4 \\ -6 & -12 & 0 & 12 & 6 \\ -4 & -8 & 0 & 8 & 4 \\ -1 & -2 & 0 & 2 & 1 \end{pmatrix} \quad (9.1)$$

Als textuurparameters gebruiken we dan de som van de grijswaarden na de inwerking van een 2-D masker op het beeld \mathbf{B} . Aanvankelijk zijn er dus 25 parameters g_{xy} :

$$g_{xy} = \sum_m \sum_n \left((\mathbf{x}_5 \mathbf{y}_5^T) * \mathbf{B} \mid \mathbf{x}_5, \mathbf{y}_5 \in \{\mathbf{l}_5, \mathbf{e}_5, \mathbf{s}_5, \mathbf{w}_5, \mathbf{r}_5\} \right) \quad (9.2)$$

Tabel 9.1: Definitie van Laws' 1-D filtermaskers

Laws filtermaskers				
$l_5 = \begin{pmatrix} 1 \\ 4 \\ 6 \\ 4 \\ 1 \end{pmatrix}$	$e_5 = \begin{pmatrix} -1 \\ -2 \\ 0 \\ 2 \\ 1 \end{pmatrix}$	$s_5 = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 0 \\ -1 \end{pmatrix}$	$w_5 = \begin{pmatrix} -1 \\ 2 \\ 0 \\ -2 \\ 1 \end{pmatrix}$	$r_5 = \begin{pmatrix} 1 \\ -4 \\ 6 \\ -4 \\ 1 \end{pmatrix}$

Dit aantal wordt echter gereduceerd om de richtingsafhankelijkheid te verminderen. Deze richtingsafhankelijkheid wordt gedeeltelijk weggewerkt door de ‘symmetrische’ parameters te combineren. Zo zal de textuurparameter na filteren met F_{l_e} opgeteld worden bij de parameter na filteren met $F_{e_l} = F_{l_e}^T$.

Dit herleidt het aantal parameters tot 15 (10 maal een combinatie van twee verschillende parameters, 5 maal een combinatie van dezelfde parameter). Tenslotte wordt nog een parameter, g_{ll} , gebruikt om de andere parameters te normaliseren. Dit resulteert in:

$$F_{xy} = \frac{g_{xy} + g_{yx}}{g_{ll}} \quad (9.3)$$

We hebben uiteindelijk dus 14 parameters.

9.2 Toepassing op echografiebeelden

9.2.1 Implementatie

Aangezien voor elk beeld 25 tweedimensionale convoluties berekend moeten worden, kan men de berekeningen versnellen door naar het frequentiedomein over te gaan. We hebben de berekeningen echter in het plaatsdomein uitgevoerd, aangezien de rekentijd ook in het plaatsdomein aanvaardbaar blijft, mede dankzij het feit dat we met scheidbare filters te maken hebben.

De normalisatieparameter g_{ll} wordt bekomen door inwerking van het F_{ll} filtermasker op het beeld. Dit is het enige 2-D masker met enkel positieve elementen, bij de overige 24 maskers is de som van alle elementen gelijk aan nul. De parameter g_{ll} zal dan ook veel groter zijn dan de andere parameters, in die mate zelfs dat na deling door g_{ll} de andere parameters te klein worden om nog nauwkeurig voorgesteld te worden in MATLAB. Om dit verlies van precisie te vermijden hebben we bij de normalisatie gedeeld door $(g_{ll} \times 10^{-9})$ in plaats van door g_{ll} .

Tabel 9.2: Vergelijking van de minimale foutpercentages bekomen met de coöccurrentiematrix en met Laws' filter maskers. g=gecompenseerd, ng=niet gecompenseerd.

<i>Vergelijking</i> <i>Foutpercentages</i>	Coronaal		Sagittaal		Beide	
	g	ng	g	ng	g	ng
Coöccurrentiematrix	0	2	0	0	3	9
Laws' filter maskers	33	33	11	13	29	29

9.2.2 Resultaten

Classificatie met Laws' techniek geeft veel hogere classificatiefouten dan classificatie met de coöccurrentiematrix of zijn varianten. De laagste foutpercentages bekomen met Laws' maskers worden in Tabel 9.2 vergeleken met die bekomen uit de coöccurrentiematrix. Hieruit blijkt duidelijk dat deze techniek veel minder geschikt is voor echografiebeelden. Daarom zullen we hier ook niet verder ingaan op de verschillende parameters en parametercombinaties. We merken enkel nog op dat het compensatie-algoritme hier geen lagere classificatiefouten toelaat, omdat de methode geen gebruik maakt van eerste orde parameters.

Hoofdstuk 10

Gaborfilters

De techniek die in dit hoofdstuk wordt besproken maakt gebruik van *Gaborfilters*. Gaborfilters worden vaak toegepast in de beeldverwerking, niet enkel voor segmentatie of classificatie van textuur, maar bijvoorbeeld ook voor documentanalyse, analyse van vingerafdrukken, beeldcompressie, . . .

Het gebruik van Gaborfilters in de beeldverwerking vindt zijn oorsprong in de biologische relevantie ervan: de werking van een Gaborfilter vertoont namelijk veel gelijkenissen met de werking van de lichtgevoelige receptoren in de visuele cortex.

De Gaborfilters kunnen hier voor twee zaken toegepast worden: enerzijds is er de classificatie tussen ziek en gezond weefsel, anderzijds kunnen we ook trachten door middel van Gaborfilters segmentatie uit te voeren tussen het relevante (zieke of gezonde) weefsel rond de ventrikels en de andere gebieden van de echografiebeelden.

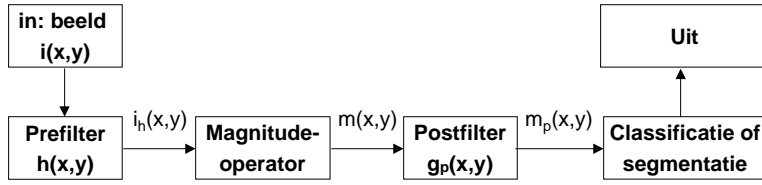
10.1 Definitie

Figuur 10.1 geeft een blokschema van de door ons gebruikte filters [14]. Het beeld $i(x, y)$ wordt eerst door een *Gabor prefilter* $h(x, y)$ gestuurd, het gefilterde beeld noemen we $i_h(x, y)$. Omdat het Gaborfilter een complex impulsantwoord bezit zal $i_h(x, y)$ ook complex zijn, en wordt bijgevolg de *magnitude* genomen vooraleer het beeld door een *Gaussiaans postfilter* $g(x, y)$ te sturen. De uitgang van dit filter wordt gebruikt voor segmentatie of classificatie. We zullen hieronder de afzonderlijke delen van het filter kort bespreken.

10.1.1 Het Gabor prefilter

Het beeld wordt dus eerst door een Gabor prefilter gestuurd. Dit filter heeft als impulsantwoord $h(x, y)$:

$$h(x, y) = g(x, y)e^{-j2\pi(Ux+Vy)}$$



Figuur 10.1: Blokschema van de gebruikte filters.

$$= \frac{1}{2\pi\sigma_g^2} e^{-\frac{(x^2+y^2)}{2\sigma_g^2}} e^{-j2\pi(Ux+Vy)} \quad (10.1)$$

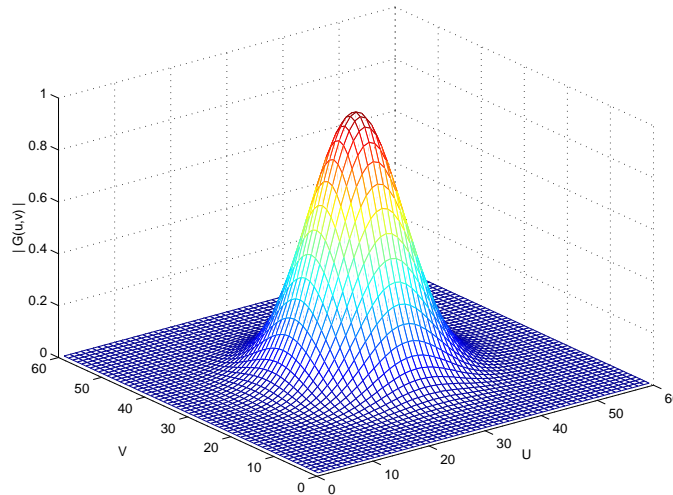
De 2-D Fouriergetransformeerde van de impulsrespons wordt gegeven door:

$$\begin{aligned} H(u, v) &= G(u - U, v - V) \\ &= e^{-2\pi^2\sigma_g^2((u-U)^2+(v-V)^2)} \end{aligned} \quad (10.2)$$

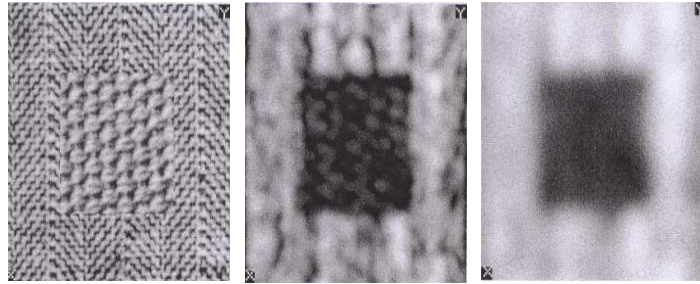
Deze Fouriergetransformeerde $H(u, v)$ is dus een in het frequentiedomein verschoven versie van $G(u, v)$, waarbij $G(u, v)$ de Fouriergetransformeerde is van de Gaussiaanse functie $g(x, y)$. $H(u, v)$ is dus een bandpassfilter met *centrale frequentie* (U, V) en een *bandbreedte* die bepaald wordt door σ_g . Een exemplaar van $H(u, v)$ wordt weergegeven in Fig. 10.2. We merken tenslotte nog op dat we voor de eenvoud een symmetrische Gaussiaanse functie $g(x, y)$ nemen. De uitgang van het prefilter $i_h(x, y)$ wordt gegeven door

$$i_h(x, y) = h(x, y) * i(x, y) \quad (10.3)$$

waarbij $*$ de tweedimensionale convolutie voorstelt. Aangezien de impulsrespons (10.1) complex is, zal de uitgang van het filter meestal ook complex zijn. Daarom wordt eerst de



Figuur 10.2: Voorstelling van het prefilter in het frequentiedomein.



Figuur 10.3: Voorbeeld van segmentatie m.b.v. Gaborfilters [14]. Links: ingangsbeeld, midden: ingang van postfilter, rechts: uitgang van postfilter.

magnitude genomen van het geprefilterde beeld, alvorens het door het Gaussiaans postfilter te sturen:

$$m(x, y) = | i_h(x, y) | = | h(x, y) * i(x, y) | \quad (10.4)$$

10.1.2 Het Gaussiaans postfilter

Het doel van het Gaussiaans postfilter $g_p(x, y)$ is het uitmiddelen van de variaties in de prefilter uitgang. Deze variaties kunnen bijvoorbeeld het gevolg zijn van fluctuaties in de textuur of ruis in het beeld. De uitgang van het postfilter $m_p(x, y)$ volgt uit:

$$\begin{aligned} m_p(x, y) &= m(x, y) * g_p(x, y) \\ &= m(x, y) * \frac{1}{2\pi\sigma_p^2} e^{-\frac{(x^2+y^2)}{2\sigma_p^2}} \end{aligned} \quad (10.5)$$

Het postfilter wordt volledig bepaald door de parameter σ_p , die een maat is voor de breedte van de Gaussiaanse functie. De uitgang van het postfilter wordt tenslotte gebruikt als ingang van het classificatie- of segmentatie-algoritme.

Als voorbeeld van segmentatie met Gaborfilters, en om het belang van het postfilteren te tonen, illustreren we in Fig. 10.3 de segmentatie van het linkse beeld. De grijswaarden die boven een bepaalde drempelwaarde liggen worden aan de ene textuur toegekend, grijswaarden onder de drempelwaarde worden aan de andere textuur toegekend. We zien dat bij de uitgang van het prefilter (middelste figuur) nog verschillende pixels verkeerd zullen worden beoordeeld (de verticale zwarte strepen), na uitmiddelen in het plaatsdomein (rechtse figuur) zal de efficiëntie al veel hoger zijn.

10.2 Toepassing

Filterbank en filterontwerp De grote uitdaging bij het gebruik van Gaborfilters is het vinden van goede waarden voor de parameters van het filter. Er zijn twee aanpakken

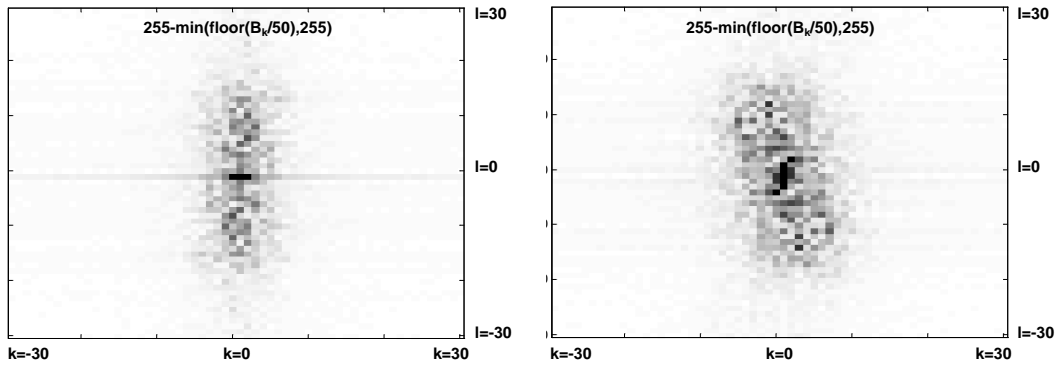
mogelijk: een *filterbank* aanpak of een *filterontwerp* (filterdesign) aanpak. Bij de filterbank aanpak gebruiken we een groot aantal filters in parallel, waarbij de parameters van de verschillende filters op voorhand vastliggen. Bij de filterontwerp aanpak zullen we uit de testbeelden de optimale waarden voor de parameters trachten te berekenen. De filterbank aanpak is eenvoudiger, maar heeft ook enkele nadelen:

- De filterparameters worden vooraf ad hoc gekozen, en zijn dus niet noodzakelijk optimaal voor de gekozen texturen.
- Dikwijls zijn er voor een efficiënte classificatie of segmentatie veel minder filters nodig.
- Vrij rekenintensief, wat bij sommige toepassingen niet gewenst is.
- We hebben parametervectoren met hoge dimensies, wat kan leiden tot gecompliceerde classificatie-algoritmes.

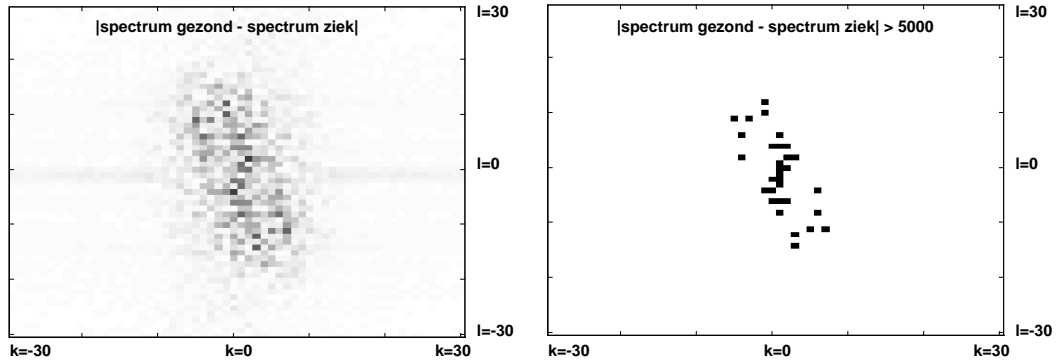
We hebben gekozen voor een filterontwerp aanpak. De methode bestaat erin voor de verschillende klassen de moduli van de discrete spectra te berekenen, te kijken voor welke frequenties deze spectra het meeste verschillen, en deze frequenties als centrale frequentie van het filter te kiezen.

Segmentatie De bedoeling van de segmentatie is om het weefsel rond de ventrikels te onderscheiden van het andere, niet-relevante weefsel. In Fig. 10.4 zien we links de modulus van het discrete spectrum van een sample van het weefsel rond de ventrikels (60x60 pixels), rechts zien we de modulus van het discrete spectrum van een sample van niet-relevant weefsel (60x60 pixels). Beide samples werden uit hetzelfde testbeeld gehaald. In Fig. 10.5 zien we links de absolute waarde van het verschil tussen deze twee spectra, rechts zien we de frequenties waarvoor dit verschil groter is dan een bepaalde grenswaarde, hier 5000.

Deze laatste frequenties zullen we gebruiken als centrale frequentie van het gaborfilter. De beste segmentatie voor het testbeeld werd bekomen met $(U, V) = (20, 24)$, $\sigma_g = 0.001$ en $\sigma_p = 15$, de segmentatie wordt weergegeven in Fig. 10.6. Het probleem is dat deze parameters voor andere beelden geen goede segmentatie geven, ook de andere parametercombinaties die volgen uit Fig. 10.5 werken enkel goed op het testbeeld en geven voor de andere beelden geen goede resultaten. Wegens tijdsgebrek hebben we geen methode kunnen opstellen die voor alle beelden goede segmentatie toelaat.



Figuur 10.4: Links zien we de modulus van het discrete spectrum van een sample van het weefsel rond de ventrikels, rechts zien we de modulus van het discrete spectrum van een sample van niet relevant weefsel. De spectra werden herschaald, en dan beperkt tot het interval $[0\ 255]$ om als grijswaardebeeld voorgesteld te kunnen worden.



Figuur 10.5: Links zien we de absolute waarde van het verschil tussen de spectra van het relevante en niet-relevante weefsel, rechts zien we de frequenties waarvoor dit verschil een bepaalde drempel (hier 5000) overschrijdt.

Classificatie Door tijdsgebrek hebben we ook de classificatie niet grondig kunnen bestuderen. Goede resultaten worden bekomen met $U = V = 0$, $\sigma_g = 0.1$. Dit was te verwachten, aangezien op die manier de DC-component (de gemiddelde grijswaarde) uitgefilterd wordt. De bandbreedte van het postfilter is niet belangrijk, omdat we toch een uniform beeld hebben na het prefilteren. Verder onderzoek moet nog uitgevoerd worden naar andere filterparameters die de classificatie nog kunnen verbeteren.



Figuur 10.6: Links zien we het gebruikte testbeeld, rechts de segmentatie ervan met $(U, V) = (20, 24)$, $\sigma_g = 0.001$ en $\sigma_p = 15$.

Hoofdstuk 11

Besluit

In deze thesis werden enkele textuurmethodes onderzocht, waarbij het vooral belangrijk was om op basis van de textuurparameters een zo optimaal mogelijke classificatie tussen gezond en ziek weefsel te bekomen. Hieronder wordt samengevat wanneer de laagste classificatiefouten worden bekomen:

- Met een venster van 60x60 pixels. Deze venstergrootte geeft meestal lagere classificatiefouten dan vensters van 30x30, 40x40 of 50x50 pixels. De beste resultaten worden dus bekomen wanneer we voor het grootste venster kiezen, omdat de ruis dan maximaal wordt uitgemiddeld. Grotere vensters worden niet gebruikt, omdat het dan moeilijk wordt om de vensters zo te kiezen dat ze enkel relevant weefsel bevatten.
- Met een hoek van 45 graden bij de coronale beelden. Dit is te verklaren door het feit dat we alle vensters uit de rechterkant van de beelden hebben genomen, en de textuur daar een voorkeursrichting van 45 graden heeft.

Bij de sagittale beelden werden de vensters zowel links, in het midden, als rechts uit het beeld genomen, zodat er geen voorkeurshoek zal zijn. Met een grotere testset zou het wel mogelijk zijn alle vensters aan dezelfde kant van het beeld te kiezen, zodat ook dan een voorkeurshoek kan optreden.

- Het scheiden van de coronale en sagittale beelden levert ook lagere classificatiefouten op. De reden hiervoor is dat de textuur van coronale en sagittale doorsneden te verschillend is om de beelden samen te behandelen.
- Bij de methodes die gebruik maken van eerste orde statistieken leidt het compensatie-algoritme tot lagere classificatiefouten. De reden hiervoor is dat de eerste orde statistieken, vooral de gemiddelde grijswaarde, heel belangrijk zijn voor een goede classificatie en deze parameters opmerkelijk beter werken bij de gecompenseerde

beelden. Bij de methodes die geen eerste orde parameters gebruiken is compensatie weinig voordelig.

De meeste methodes zijn aan elkaar gewaagd: de drie beschouwde varianten van de coöcurrentiematrix moeten niet onderdoen voor de coöcurrentiematrix, hoewel ze meestal veel eenvoudiger zijn. Wanneer we de parameters berekend uit de looplengtematrix aanvullen met de gemiddelde grijswaarde is ook met deze methode even goede classificatie mogelijk. Laws' methode, die geen eerste orde statistieken gebruikt, geeft geen goede resultaten.

De beste classificatie wordt dus bekomen door de eerste orde statistieken gemiddelde grijswaarde en eventueel variantie op de grijswaarde aan te vullen met één of meerdere textuurparameters. Op basis van enkel textuurparameters is veel minder goede classificatie mogelijk (zie bijvoorbeeld de looplengtematrix zonder gemiddelde grijswaarde). Dit kan erop wijzen dat de textuur van het zieke en gezonde weefsel te weinig verschillend is om er enkel op basis van textuur onderscheid tussen te kunnen maken.

Het zou ook kunnen dat de methodes die we gebruikt hebben niet optimaal geschikt zijn om de textuur van de echografiebeelden te beschrijven, en er daarom op basis van enkel textuurparameters niet zo een goede classificatie mogelijk is. Daarom is verder onderzoek van de gaborfilters en onderzoek met totaal andere methodes (zoals waveletanalyse) zeker nuttig.

Bibliografie

- [1]
- [2]
- [3] A. Baraldi and F. Parmiggiani. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing*, 33(2):293–304, Maart 1995.
- [4] O. Basset, Z. Zun, J.L. Mestas, and G. Giminez. Texture analysis of ultrasonic images of the prostate by means of co-occurrence matrices. *Ultrasonic Imaging*, 15:218–237, 1993.
- [5] Z. Dapeng and L. Zhongrong. *Applications of artificial intelligence*, volume 657, chapter Digital image texture analysis using gray level and energy cooccurrence, pages 152–156. SPIE, Bellingham, WA, 1986.
- [6] Marjan De Dene. Textuurclassificatie in echografiebeelden. Master’s thesis, Universiteit Gent, Faculteit Toegepaste Wetenschappen, 2002.
- [7] M.M. Galloway. Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, 4:172–179, 1975.
- [8] R. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6):610–621, November 1973.
- [9] K. Laws. Rapid texture identification. In *Proceedings of SPIE Image Processing for Missile Guidance*, volume 238, pages 376–380, 1980.
- [10] W. Philips. *Nota’s bij de cursus beeldverwerking*. Universiteit Gent, Faculteit Toegepaste Wetenschappen, 2000.
- [11] R. Putz and R. Pabst, editors. *Sobotta atlas van de menselijke anatomie, deel1: Hoofd, hals, bovenste extremiteit*. Bohn Stafleu Van Loghum, Houten/Diegem, 2000.

- [12] M. Unser. Sum and difference histograms for texture analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(1):118–125, Januari 1986.
- [13] C. Vertan, C. Fernandez-Maloigne, N. Richard, and M. Chaker Larabi. Compact color-texture run-length description for ornamental stones recognition and indexing. In *Proceedings of ACIVS (Advanced Concepts for Intelligent Vision Systems) 2002. Gent, Belgie*, pages 1–7, 2002.
- [14] T.P. Weldon, W.E. Higgins, and D.F. Dunn. Gabor filter design for multiple texture segmentation. *Optical Engineering*, 35(10):1–17, Oktober 1996.