

Omgevingsreconstructie met akoestische signalen

Brecht Milis
Jelle Vanolst

Thesis voorgedragen tot het
behalen van de graad van Master
in de ingenieurswetenschappen:
elektrotechniek, optie Multimedia
en signaalverwerking

Promotor:
Prof. dr. ir. Marc Moonen

Academiejaar 2010 – 2011

Omgevingsreconstructie met akoestische signalen

Brecht Milis
Jelle Vanolst

Thesis voorgedragen tot het behalen van de graad van Master in de ingenieurswetenschappen: elektrotechniek, optie Multimedia en signaalverwerking

Promotor:

Prof. dr. ir. Marc Moonen

Assessoren:

Prof. dr. ir. Patrick Wambacq
dr. ir. Alexander Bertrand

Begeleider:

ir. Bram Cornelis

© Copyright K.U.Leuven

Zonder voorafgaande schriftelijke toestemming van zowel de promotor(en) als de auteur(s) is overnemen, kopiëren, gebruiken of realiseren van deze uitgave of gedeelten ervan verboden. Voor aanvragen tot of informatie i.v.m. het overnemen en/of gebruik en/of realisatie van gedeelten uit deze publicatie, wend u tot ESAT, Kasteelpark Arenberg 10 postbus 2440, B-3001 Heverlee, +32-16-321130 of via e-mail info@esat.kuleuven.be.

Voorafgaande schriftelijke toestemming van de promotor(en) is eveneens vereist voor het aanwenden van de in deze masterproef beschreven (originele) methoden, producten, schakelingen en programma's voor industrieel of commercieel nut en voor de inzending van deze publicatie ter deelname aan wetenschappelijke prijzen of wedstrijden.

Voorwoord

Hierbij danken wij onze promotor, begeleider en de voltallige jury. Ook danken wij onze familie en vrienden voor de steun die zij geboden hebben.

Brecht Milis
Jelle Vanolst

Inhoudsopgave

Voorwoord	i
Samenvatting	iii
Lijst van figuren en tabellen	iv
Lijst van afkortingen en symbolen	viii
1 Inleiding	1
2 Schatten van impulsresponsies	3
2.1 Kamerimpulsresponsie als informatiebron over een omgeving	3
2.2 RIR bepalen met gekend bronsignaal: NLMS	5
2.3 RIR blind schatten	6
2.4 Besluit	11
3 Bronlocatie bepalen	13
3.1 Bronlokalisatie in twee dimensies	13
3.2 Bronlokalisatie in drie dimensies	20
3.3 Besluit	26
4 Reflectorlocaties bepalen	27
4.1 2D Lokalisatie van de reflectoren	27
4.2 3D Lokalisatie van de reflectoren	57
4.3 Besluit	70
5 Praktijktesten	71
5.1 De labo opstelling	71
5.2 Resultaten 2D	71
5.3 Resultaten 3D	83
5.4 Besluit	86
6 Objectdetectie en -lokalisatie (2D)	87
6.1 Theoretisch	87
6.2 Simulatie en resultaten	88
6.3 Besluit	98
7 Besluit	99
Bibliografie	101

Samenvatting

Het probleem om een omgeving akoestisch te reconstrueren, krijgt recentelijk hernieuwde aandacht. In deze thesis wordt een manier voorgesteld om de geometrie van een kamer te schatten, in twee stappen. Een kamer met reflecterende wanden bevat een geluidsbron en enkele microfoons, verspreid doorheen deze kamer. Enkel de relatieve posities van de microfoons zijn gekend. De geluidsbron kan algemeen worden beschouwd, zoals een persoon die aan het praten is. Het kan ook een gecontroleerd bronsignaal zijn. Eerst worden de propagatietijden van de geluidsgolven bepaald waarbij de verschillende methodes om dit te bereiken worden toegelicht. Dit zowel voor gekend bronsignaal als ongekend bronsignaal. Uit deze propagatietijden kan de bronlocatie worden bepaald en ook de geometrie van de kamer. Dit wordt zowel uitgewerkt en getest in twee dimensies als in drie dimensies. Ook wordt een korte vergelijking gemaakt met andere methodes om kamerreconstructie te verwezenlijken. Uit de resultaten blijkt dat reconstructie in het driedimensionale geval minder goed verloopt als in het tweedimensionale geval, maar dat de resultaten aanvaardbaar zijn. Als extra toepassing, wordt een methode voorgesteld om op akoestische wijze een object te detecteren en te lokaliseren in de kamer. Vooral wordt geconcentreerd op de locatie van het voorwerp, eerder dan een reconstructie van de vorm ervan.

Lijst van figuren en tabellen

Lijst van figuren

2.1	Voorbeeld van RIR met aanduiding van verschillende zones.	4
2.2	Adaptief filter voor systeemidentificatie, hier de kamerimpulsresponsie. .	5
2.3	Zoeken van direct pad en eerste orde TOA's via kruiscorrelatie.	7
2.4	Voorbeeld van reconstructie van een RIR via het closed-form algoritme.	10
3.1	Gesimuleerde impulsresponsie van het kanaal tussen bron en microfoon 1.	16
3.2	Gesimuleerde impulsresponsie van het kanaal tussen bron en microfoon 2.	17
3.3	Gesimuleerde impulsresponsie van het kanaal tussen bron en microfoon 3.	18
3.4	Gesimuleerde impulsresponsie van het kanaal tussen bron en microfoon 4.	19
3.5	Resultaat van simulatie van bronlokalisatie in 2D-ruimte (o: microfoon; Δ : werkelijke locatie; *: berekende locatie).	20
3.6	Resultaat van simulatie van bronlokalisatie in 3D-ruimte (o: microfoon; Δ : werkelijke locatie; *: berekende locatie).	23
3.7	Resultaat van een simulatie van bronlokalisatie in 3D-ruimte met verspreide opstelling (o: microfoon; Δ : werkelijke locatie; *: berekende locatie).	24
3.8	Resultaat van een simulatie van bronlokalisatie in 3D-ruimte met minder gunstige oplossing (o: microfoon; Δ : werkelijke locatie; *: berekende locatie).	25
4.1	RIR microfoon 1 met eerste orde TOA's via methode 4 maxima.	30
4.2	RIR microfoon 2 met eerste orde TOA's via methode 4 maxima.	31
4.3	RIR microfoon 4 met eerste orde TOA's via methode 4 maxima.	32
4.4	RIR microfoon 4 met eerste orde TOA's via methode 4 maxima en beperking.	34
4.5	RIR microfoon 2 met eerste orde TOA's via alternatieve methode.	35
4.6	RIR micro 1 met eerste orde TOA's en $R=0.9$	36
4.7	RIR micro 1 met eerste orde TOA's en $R=0.5$	37
4.8	Resultaat 4 reflectoren, J_e via (4.19), $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $R=0.5$	44
4.9	Resultaat 4 reflectoren, J_e via (4.19), $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m, $R=0.5$	45

4.10	Resultaat 4 reflectoren, Je via (4.22), $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $R=0.5$	46
4.11	Resultaat 4 reflectoren ideaal, Je via (4.22), $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m.	47
4.12	Resultaat 4 reflectoren, Je via (4.22), $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m, $R=0.5$	48
4.13	Resultaat 4 reflectoren, Je via (4.22), $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m, $R=0.9$	49
4.14	Resultaat 4 reflectoren, TOA's via 4 maxima, $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m, $R=0.7$	50
4.15	Resultaat 4 reflectoren, TOA's via alternatieve methode, $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $R=0.7$	51
4.16	Resultaat 4 reflectoren, TOA's via kruiscorrelatie, $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $R=0.7$	52
4.17	Resultaat 4 reflectoren, Je via (4.22), $\Delta=0.01$ m, $L=3.10$ m, $B=4.30$ m, $R=0.5$	53
4.18	Resultaat 4 reflectoren, Je via (4.22), $\Delta=0.1$ m, $L=3.10$ m, $B=4.30$ m, $R=0.7$	54
4.19	Resultaat 4 reflectoren, Je via (4.22), $\Delta=0.1$ m, $L=3.10$ m, $B=4.30$ m, $R=0.9$	55
4.20	Resultaat 4 reflectoren, Je via (4.22), $\Delta=0.1$ m, $L=9.0$ m, $B=10.0$ m, $R=0.7$	56
4.21	Resultaat 3D lokalisatie, ideale TOA's, $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $H=2$ m.	61
4.22	Resultaat 3D lokalisatie, ideale TOA's, $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m, $H=2$ m.	64
4.23	Resultaat 3D lokalisatie, geschatte TOA's, $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $H=2$ m, $R=0.7$	65
4.24	RIR microfoon 1 met eerste orde TOA's via methode 6 maxima	66
4.25	Resultaat 3D lokalisatie, ideale TOA's en correcte bronpositie, $L=5.3$ m, $B=7.4$ m, $H=2$ m.	67
4.26	Resultaat 3D lokalisatie, geschatte TOA's, $\Delta=0.1$ m, $L=5.3$ m, $B=7.4$ m, $H=2$ m.	68
4.27	Resultaat 3D lokalisatie, ideale TOA's, $\Delta=0.1$, $L=5.3$ m, $B=7.4$ m, $H=2$ m.	69
5.1	Foto van het labo, met de opstelling van bron en microfoons.	72
5.2	2 impulsresponsies van hetzelfde pad, waarbij de luidspreker 90° geroteerd is.	73
5.3	Resultaat van de 2D bronlokalisatie in het labo (\circ : microfoon; Δ : werkelijke locatie; *: berekende locatie).	75
5.4	RIR micro 1 bepaald via NLMS, $N=10000$	76
5.5	RIR micro 2 bepaald via NLMS, $N=10000$	77
5.6	Resultaat 2D lokalisatie met opgenomen microfoonsignalen, luidspreker in 1 richting, lage R , $L=3.76$, $B=6$	78

5.7	Resultaat 2D lokalisatie met opgenomen microfoonsignalen, luidspreker in 1 richting, hoge R , $L=3.76$, $B=6$.	79
5.8	Resultaat 2D lokalisatie met opgenomen microfoonsignalen, gemiddelde over verschillende luidspreker richtingen.	80
5.9	Resultaat 2D lokalisatie met opgenomen microfoonsignalen, 1 TOA per luidsprekerrichting, hoge R , $L=3.76$, $B=6$.	81
5.10	Resultaat 2D reflectorlokalisatie via autocorrelatie, hoge R .	82
5.11	Resultaat van de 3D bronlokalisatie in het labo (\circ : microfoon; Δ : werkelijke locatie; *: berekende locatie).	84
5.12	Resultaat 3D lokalisatie met opgemeten microfoonsignalen.	85
6.1	Resultaat van de stemming voor simulatie 1.	90
6.2	Kamersituatie voor simulatie 1 (\circ : microfoon; Δ : bronlocatie; *: objectlocatie; +: snijpunt).	91
6.3	Resultaat van de stemming voor simulatie 2.	92
6.4	Kamersituatie voor simulatie 2 (\circ : microfoon; Δ : bronlocatie; *: objectlocatie; +: snijpunt).	93
6.5	Resultaat van de stemming voor simulatie 3, met objectstraal 0.15 m.	94
6.6	Kamersituatie voor simulatie 3, met objectstraal 0.15 m (\circ : microfoon; Δ : bronlocatie; *: objectlocatie; +: snijpunt).	95
6.7	Resultaat van de stemming voor simulatie 3, met objectstraal 0.2 m.	96
6.8	Kamersituatie voor simulatie 3, met objectstraal 0.2 m (\circ : microfoon; Δ : bronlocatie; *: objectlocatie; +: snijpunt).	97

Lijst van tabellen

3.1	Locaties van bron en microfoons in 2D-ruimte (in meter).	16
3.2	Locaties van bron en microfoons in 3D-ruimte (in meter).	22
3.3	Locaties van bron en microfoons in 3D-ruimte, verspreide opstelling.	23
3.4	Locaties van bron en microfoons in 3D-ruimte (in meter) voor niet gunstig resultaat.	24
4.1	TOA's in aantal samples, via 4 maxima, $R=0.7$.	33
4.2	TOA's in aantal samples, via spiegeling (ideale waardes).	33
4.3	TOA's in aantal samples, via alternatieve methode, $R=0.7$.	34
4.4	TOA's in aantal samples, via 4 maxima, $R=0.9$.	36
4.5	TOA's in aantal samples, via 4 maxima, $R=0.5$.	36
4.6	Overzicht reflectorlokalisatie.	49
4.7	Locaties van bron en microfoons in 2D-ruimte, opstelling 2.	50
4.8	TOA's in aantal samples, methode 4 maxima, opstelling 2 en vergelijking met ideale. Eerste 4 kolommen met $R=0.5$, laatste vier kolommen zijn de ideale.	50
4.9	TOA's in aantal samples, methode 4 maxima, opstelling 2 voor $R=0.7$ (eerste 4 kolommen) en $R=0.9$ (laatste 4 kolommen).	51
4.10	Locaties van bron en microfoons in 2D-ruimte, opstelling 3.	51

4.11	TOA's in aantal samples, methode 6 maxima en vergelijking met ideale.	62
4.12	Overzicht reflectorlokalisatie in 3D (in meter).	62
4.13	Locaties van bron en microfoons in 3D-ruimte, opstelling 2.	63
5.1	TOA's in aantal samples, luidspreker in één richting, hoge R . Vier linkerkolommen: zonder beperking op afstand tussen TOA's. Vier rechterkolommen: met beperking op afstand tussen TOA's.	79
5.2	TOA's in aantal samples, luidspreker in één richting, lage R .	80
5.3	TOA's in aantal samples, luidspreker in verschillende richtingen, hoge R .	81
5.4	TOA's in aantal samples, 1 TOA per luidsprekerrichting, vier linkerkolommen: hoge R en vier rechterkolommen: lage R .	82
5.5	Overzicht reflectorlokalisatie: praktische metingen.	83
6.1	Locaties van bron en microfoons in 2D-ruimte (in meter).	89

Lijst van afkortingen en symbolen

Afkortingen

<i>DSP</i>	digitale signaalverwerking (Digital Signal Processing)
<i>TOA</i>	aankomsttijd (Time Of Arrival)
<i>TDOA</i>	verschil in aankomsttijd (Time Difference Of Arrival)
<i>RIR</i>	kamerimpulsresponsie (Room Impulse Response)
<i>AED</i>	adaptieve eigenwaarde decompositie
<i>NLMS</i>	genormaliseerde kleinste gemiddelde kwadraten (Normalized Least Mean Squares)
<i>LMS</i>	kleinste gemiddelde kwadraten (Least Mean Squares)
<i>LS</i>	kleinste kwadraten (Least Squares)
<i>CC</i>	kruiscorrelatie (Cross Correlation)
<i>DOA</i>	richting van aankomst (Direction Of Arrival)

Symbolen

F_s	bemonsteringsfrequentie (44.1 kHz indien niet gespecificeerd)
η	geluidssnelheid in lucht (343 m/s bij kamertemperatuur)
R_s	afstand van oorsprong tot bron
R_i	afstand van oorsprong tot microfoon
D_i	afstand van microfoon i tot de bron
$\Delta_{i,j}$	TDOA van direct pad tussen microfoon i en microfoon j
$\tau_{i,j}$	TOA horende bij microfoon i en reflector j
μ	stapgrootte
$d_{i,j}$	bereikverschil
τ_{CC}	tijdsvertraging bij kruiscorrelatie
Φ_{CC}	kruiscorrelatie
\mathbf{w}	filtercoëfficiënten
\mathbf{x}	microfoonsignaal
x	sample van microfoonsignaal
\mathbf{g}_i	impulsresponsie pad bron - microfoon i
F	iteratiestap
$\mathbf{R}_{x_i x_j}$	covariantiematrix van microfoonsignaal i en j
Σ	covariantiematrix
\mathbf{H}	paar van impulsresponsies
L	lengte impulsresponsie
\mathbf{u}	witteruis signaal
κ	factor bij stapgrootte
\mathbf{v}	homogene coördinatenvector
N	aantal reflectoren
M	aantal microfoons
ρ	straal object
γ_i	hoek bij microfoon i bij positieve verdraaiing in xy-vlak
α	polaire coördinaat
α_i	hoek bij microfoon i bij verdraaiing rond positieve z-as
β_i	hoek bij microfoon i bij verdraaiing rond positieve y-as
θ	bolcoördinaat
ϕ	bolcoördinaat
J_e	kostenfunctie
\mathbf{l}	coëfficiëntenvector
\mathbf{C}	coëfficiëntenmatrix
ζ	kegelsnede/3D oppervlak in homogene coördinaten
λ	rechte in homogene coördinaten
ψ	vlak in homogene coördinaten
\mathbf{R}_i	rotatiematrix bij microfoon i
$\mathbf{S}_{i,k}$	schaalmatrix bij microfoon i en reflector k
\mathbf{T}_i	translatiematrix bij microfoon i
$\mathbf{Q}_{i,k}^{maj}$	lengte van semi-hoofdass bij microfoon i en reflector k
$\mathbf{Q}_{i,k}^{maj}$	lengte van semi-bijass bij microfoon i en reflector k
\mathbf{e}	foutsignaal
Θ	bronlocatie
R	reflectiecoëfficiënt
Δ	discretisatiestapgrootte

Hoofdstuk 1

Inleiding

Het oorspronkelijke idee dat geleid heeft tot deze thesis komt uit een film van Batman, nl. The Dark Knight. Wie deze film gezien heeft, weet allicht nog dat Batman in deze film gebruik maakt van een akoestisch systeem om de schuilplaats van de Joker 's nachts te kunnen vinden. Om dit te verwezenlijken, heeft hij elke GSM in de stad getransformeerd tot een hoog frequente zendontvanger, waarmee puur op akoestische basis een 3D-beeld gevormd wordt van de omgeving waarin een toestel zich bevindt. Ook de positie van de Joker in de kamer weet hij op deze manier. De vraag die in deze thesis onderzocht wordt is of dergelijk systeem ook in realiteit gebouwd kan worden, met behulp van DSP technieken.

Een eerste gebruik van akoestische signalen voor het reconstrueren van een omgeving kan men terugvinden bij onderwatertoepassingen. Reeds lange tijd wordt sonar gebruikt als onderwatercamera of als objectdetector [1] [2]. Ook 3D-reconstructie van een bodemlandschap onder water kan op dit moment vrij nauwkeurig uitgevoerd worden via sonar [3]. Nu is het niet zo vanzelfsprekend om deze technieken ook boven water en in gesloten ruimtes toe te passen. Ten eerste is er het verschil tussen propagatie van geluid in lucht en in water. Ten tweede treden er reverberatie effecten op wanneer men in een gesloten ruimte werkt. Vandaar dat in lucht naar andere oplossingen wordt gezocht, waarbij reverberatie eerder als extra informatiebron over de omgeving beschouwd wordt.

Technieken voor akoestische omgevingsreconstructie kunnen ingedeeld worden in twee klassen. De eerste manier bestaat eruit gebruik te maken van een microfoonrooster en bundelsturingstechnieken. Er is reeds onderzoek verricht om via een klein circulair rooster 2D- en ook 3D-reconstructie uit te voeren [4]. Een andere mogelijkheid is het gebruik van een sferisch microfoonrooster [5], ook audiocamera genoemd. Zulke camera's zijn reeds op de markt beschikbaar [6], vaak in combinatie met een optische camera. Deze worden toegepast om van gesloten omgevingen een akoestische intensiteitsmap te verkrijgen, bijvoorbeeld van een concerthal [7]. Een tweede aanpak maakt geen gebruik van dit soort roosters, maar van losse microfoons verspreid over de kamer. Deze aanpak sluit dichter aan bij het originele Batman systeem, de verschillende microfoons kunnen bijvoorbeeld GSM's van personen in de kamer zijn. Deze tweede klasse van technieken worden in deze thesis bestudeerd. De methode

is nog vrij nieuw, publicaties over dergelijke methodes zijn schaars en meestal ook beperkt tot 2D-reconstructies. Gelijkaardig aan het gebruik van meerdere microfoons kan ook één bewegende microfoon worden gebruikt [8]. Ook toont recent onderzoek aan dat het reconstrueren van een kamer in theorie mogelijk is via één microfoon en één bron [9]. Via een kort durend geluid in de kamer, zoals een klap, wordt op dezelfde plaats als de bron de kamerimpulsresponsie (RIR) opgemeten. Zo kan men dus de vorm van een kamer ‘horen’.

Voor de bron kunnen verschillende veronderstellingen gemaakt worden: ofwel is zowel het uitgezonden signaal als de positie ten opzichte van de microfoons ongekend (dit kan dan bijvoorbeeld een pratend persoon zijn), ofwel is enkel het uitgezonden signaal gekend, maar niet de positie (bijvoorbeeld een toestel in de kamer dat gecontroleerd wordt zoals de GSM in het Batman voorbeeld), ofwel zijn beide gekend. Alle manieren worden onderzocht en hebben hun aparte toepassingen. Een eerste stap in onze aanpak is het vinden van de afmetingen van de kamer door de posities van reflecterende wanden (muren, vloer en plafond) te zoeken. In 3D leidt dit dan tot een balkvormig model van de kamer. Naast het zoeken van afmetingen van een kamer zijn we ook geïnteresseerd in het vinden van objecten en/of personen in de kamer. Voor het bepalen van de afmetingen van de kamer worden zowel simulaties als praktische testen uitgevoerd en besproken, voor de objectdetectie en localisatie worden enkel simulaties besproken. Indien het bronsignaal ongekend is, en alleen opgenomen microfoonsignalen ter beschikking zijn, moet er een blinde kanaalschatting uitgevoerd worden. De beperkingen van blind schatten ten opzichte van kanaalschatting met gekend bronsignaal worden ook toegelicht.

Akoestische omgevingsreconstructie kan in verscheidene audiotoeepassingen aangewend worden. Informatie over de geometrie van de kamer kan helpen bij het zoeken naar RIR's. Bijvoorbeeld bij 3D-audio in televisies is het nodig om de RIR vanaf het toestel tot aan de luisteraar te kennen [10]. In plaats van deze te schatten uit opgenomen signalen telkens wanneer de luisteraar zich verplaatst of anders oriënteert, kan men deze laten genereren door een model, dat gebruik maakt van de kamerdimensies. Gelijkaardig kan men via zulk een model dereverberatie effecten modelleren, kennis hiervan kan betere resultaten geven in algoritmes omtrent bronlocalisatie en echoonderdrukking [11]. Dit wordt bijvoorbeeld geïmplementeerd in de Kinect van Xbox. Dit toestel registreert bewegingen in 3D en zorgt ervoor dat men computerspellen kan spelen zonder extra besturingsapparatuur. Ook spraakcontrole is in het toestel geïmplementeerd. Er wordt hierbij gebruik gemaakt van een microfoonrooster en akoestische ruisonderdrukking en echo-cancellatie software om een goed spraaksignaal te verkrijgen. Via een kalibratieprocedure zoekt de Kinect op akoestische wijze de impulsresponsie van de kamer [12]. Informatie over hoe de reflecties verlopen in de kamer helpt dan verder om ongewenste geluiden en echo's te onderscheiden van gewenste spraakcommando's.

Hoofdstuk 2

Schatten van impulsresponsies

Het lineaire filter dat het akoestisch pad tussen geluidsbron en microfoon in een kamer modelleert, wordt de kamerimpulsresponsie (RIR) genoemd. Zoals in het inleidend hoofdstuk besproken, is het doel de posities van reflectoren in een omgeving te bepalen aan de hand van meerdere microfoons en een geluidsbron. Er is een belangrijk verschil met deze aanpak ten opzichte van het gebruik van microfoonroosters. Beschouwend het directe pad van een geluidsbron tot een lineair uniform verdeeld microfoonrooster, dan volgt dat de microfoons hetzelfde signaal ontvangen op een tijdsvertraging na. Deze is afhankelijk van de zogenaamde *Direction Of Arrival* (DOA), nl. de richting van waaruit de geluidsgolf toekomt aan het microfoonrooster. Indien abstractie gemaakt wordt van het directionaliteitspatroon van de microfoons, kan via eenvoudige filter- en-som bundelsturingstechnieken een intensiteitsmap in functie van de stuurrichting van het rooster worden opgesteld. Via een circulair rooster [13] kan deze stuurrichting van 0 tot 360 graden variëren, met een sferisch rooster kan deze intensiteitsmap in 3D worden opgesteld. In [14] wordt een manier om via de maxima in de intensiteitsmap de reflectoren in een omgeving te bepalen besproken.

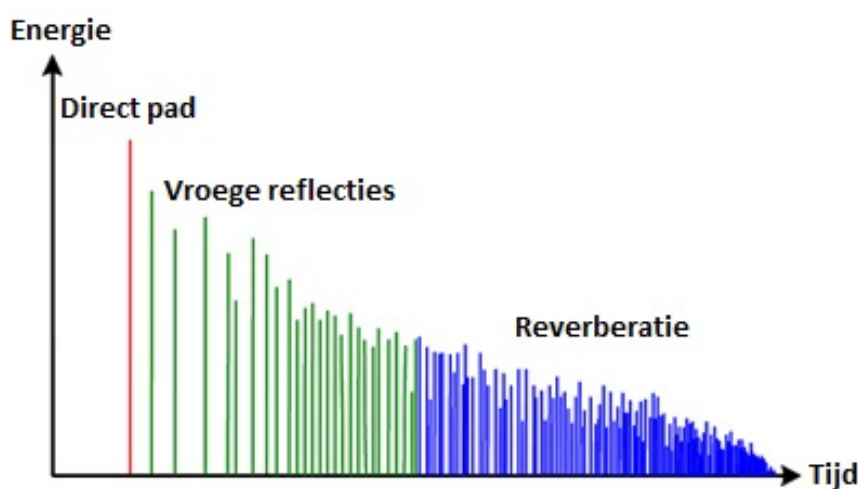
Indien niet met dit soort roosters wordt gewerkt, maar met enkelvoudige microfoons verspreid doorheen de kamer, is het niet zo evident om de DOA te bepalen als met een microfoonrooster. Vandaar dat hier andere technieken gebruikt worden, meestal met behulp van de impulsresponsie van de kamer. Indien het bronsignaal zelf gecontroleerd wordt, kan de responsie berekend worden via adaptieve filtering. Indien enkel de microfoonsignalen beschikbaar zijn, is een blinde aanpak van toepassing. Beide methodes zijn getest in simulaties en worden in dit hoofdstuk vergeleken. Eerst volgt een beknopte bespreking van de voornaamste karakteristieken van dergelijke RIR's.

2.1 Kamerimpulsresponsie als informatiebron over een omgeving

Wanneer een object of persoon een geluidsgolf produceert, gaat deze door verscheidene objecten in de omgeving gereflecteerd en geabsorbeerd worden. Zolang de golf de microfoon niet bereikt spreekt men van de dode tijd. In de veronderstelling dat

er een direct pad bestaat tussen de geluidsbron en de microfoon, zal de golf langs dit pad het eerst de microfoon bereiken. Daarna volgen de reflecties van objecten die in de dichte nabijheid van de microfoons liggen, in de literatuur vroege reflecties genoemd. Na een tiental milliseconden ontstaat een grote dichtheid aan gereflecteerde geluidsgolven in alle richtingen. Dit wordt reverberatie genoemd. De amplitude van deze reverberatiegolven daalt exponentieel. De duur ervan hangt vooral af van de reflectiecoëfficiënten van de reflectoren in de kamer. De vroege reflecties daarentegen worden hoofdzakelijk bepaald door de geometrie van de kamer en in mindere mate door de reflectiecoëfficiënten [10]. Figuur 2.1 toont een voorbeeld van een RIR met de verschillende zones aangeduid.

Indien de reflectiecoëfficiënten van de reflectoren in een grote en een kleine ruimte



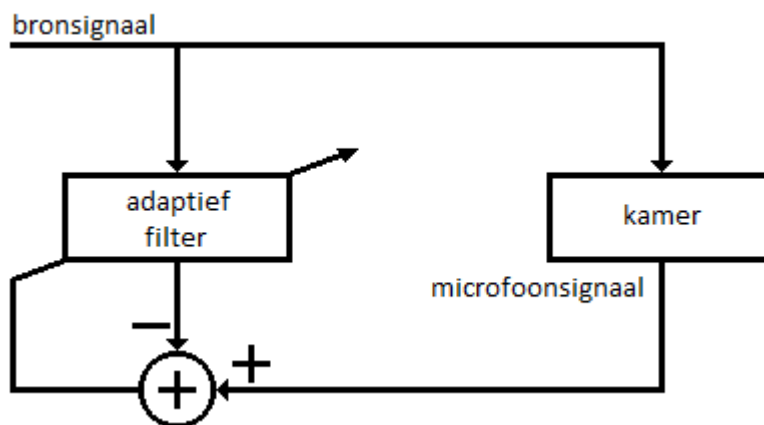
FIGUUR 2.1: Voorbeeld van RIR met aanduiding van verschillende zones.

gelijk zijn, zal de reverberatie sneller in amplitude afnemen in de kleinere dan in de grotere ruimte. Dit komt omdat in kleine ruimtes de golven frequenter gereflecteerd en daardoor ook sneller geabsorbeerd worden. Ook voor het menselijk gehoor speelt reverberatie een rol om informatie te verzamelen over een omgeving. Indien de reverberatietijd lang is (bijvoorbeeld voor een kerkgebouw kan deze tot enkele seconden bedragen) gaat de mens de reverberatie horen als gescheiden van het voorgrond geluid. Voor kleine ruimtes daarentegen bedraagt de reverberatie gemiddeld een honderdtal milliseconden. De reverberatie geeft hier enkel een tonale kwaliteit aan het voorgrondsignaal, maar perceptueel is geen scheiding tussen het voorgrondsignaal en de reverberatie waarneembaar. Ook indiceert een doffe klank dat de reflectiecoëfficiënten van de reflectoren vrij laag zijn, wat wijst op reflectoren van een absorberend materiaal. Dit heeft namelijk tot effect dat hoge frequenties meer geabsorbeerd worden door deze muren, en reverberatie sneller uitdeint. Het omgekeerde kan gezegd worden voor ruimtes met meer reflecterende materialen, zoals betonnen muren. Dit illustreert dat de RIR veel informatie bevat betreffende de geometrie van een omgeving en het materiaal waaruit reflectoren bestaan. De

verhouding van de amplitude van het directe pad ten opzichte van de reverberatie is een indicatie van de afstand waarop een bron zich bevindt. De amplitude van het directe pad zal sterk wijzigen wanneer de bron van positie wijzigt, de reverberatie zal daarentegen weinig wijzigen.

2.2 RIR bepalen met gekend bronsignaal: NLMS

NLMS is een eenvoudig iteratief algoritme om een systeem te identificeren [15]. Dit is een mogelijke methode om de impulsresponsies van elk pad van de bron tot elke microfoon te schatten. Een bemerking dient wel gemaakt te worden: het bronsignaal moet gekend zijn of indien mogelijk door de methode zelf gegenereerd worden. Maar als de bron gekend is, dan is dit een eenvoudige en snelle methode om de impulsresponsies te schatten (complexiteit $O(N)$). Het opzet is als volgt. De ingang van het systeem is het uitgezonden bronsignaal. De ontvangen signalen van de microfoons zijn de *desired* of gewenste signalen. Het adaptief filter, met filtercoëfficiënten \mathbf{w} , tracht dit signaal te reconstrueren vanuit het bronsignaal. Op deze manier kan dan de impulsresponsie gevonden worden als de overdrachtsfunctie van ingangssignaal naar gewenst signaal. Dit is weergegeven in figuur 2.2.



FIGUUR 2.2: Adaptief filter voor systeemidentificatie, hier de kamerimpulsresponsie.

NLMS is zoals reeds gezegd een iteratief, eenvoudig algoritme. De updateformule is de volgende:

$$\mathbf{w}(F+1) = \mathbf{w}(F) + \mu \cdot \mathbf{u}_{F+1} \cdot (x_{F+1} - \mathbf{u}_{F+1}^T \mathbf{w}(F)) \quad (2.1)$$

met

$$\mu = \frac{\bar{\mu}}{\kappa + \mathbf{u}_{F+1}^T \mathbf{u}_{F+1}} \quad (2.2)$$

Hierin is F de iteratiestap, en \mathbf{u}_{F+1} de vector met de ingangssamples op de vertraginglijn van het filter op tijdstip $F+1$, met een zelfde lengte als de geschatte

impulsresponsie. x is het *desired* signaal, dus het microfoonsignaal. \mathbf{w} wordt als een nulvector geïnitieerd, en zal zolang er samples van het microfoonsignaal zijn, worden geüpdatet. De μ is de variabele stapgrootte, en deze wijzigt elke iteratie. De grootte van μ bepaalt hoeveel de huidige benadering wordt gewijzigd in de richting van de juiste oplossing. Dat deze wijziging gebeurt in de richting van de juiste oplossing is een eigenschap van de LMS-algoritme familie, die gebaseerd is op de Wiener filter steilste afdaling iteraties. Deze stapgrootte μ moet zorgvuldig worden gekozen, en dit kan via $\bar{\mu}$, zodat de stabiliteit niet in gedrang komt en er een oplossing gevonden wordt. Deze keuze dient gemaakt te worden in functie van het signaal. Wordt de parameter te groot gekozen, dan gaat er niet geconvergeerd worden naar de oplossing, maar blijft er rond gecirkeld worden. Een te klein gekozen stapgrootte leidt dan weer tot een te grote nodige iteratiehoeveelheid. Dit resulteert in lange wachttijden vooraleer de oplossing convergeert. Een variabele keuze van $\bar{\mu}$ is een goed alternatief. De eerste iteraties gebeuren met grote stapgrootte, daarna wordt deze verkleind, zodat dichter naar de oplossing geconvergeerd wordt. De genormaliseerde LMS, oftewel NLMS heeft als voordeel dat de stapgrootte een indicatie geeft van hoeveel de maximale fout zal zijn (mean square error is maximum 10 procent bij stapgrootte kleiner dan 0.2). Er is stabiliteit zodra deze stapgrootte kleiner is dan twee. De κ is een factor die moet gekozen worden in functie van het probleem. Deze dient om te beletten dat de noemer van (2.2) nul wordt. Als waarde voor de parameter $\bar{\mu}$ is 0.01 gekozen, en 0.0001 is gekozen voor κ .

2.3 RIR blind schatten

In deze sectie worden de blinde schattingsmethodes uitgewerkt en hun voor- en nadelen besproken.

2.3.1 Kruiscorrelatie

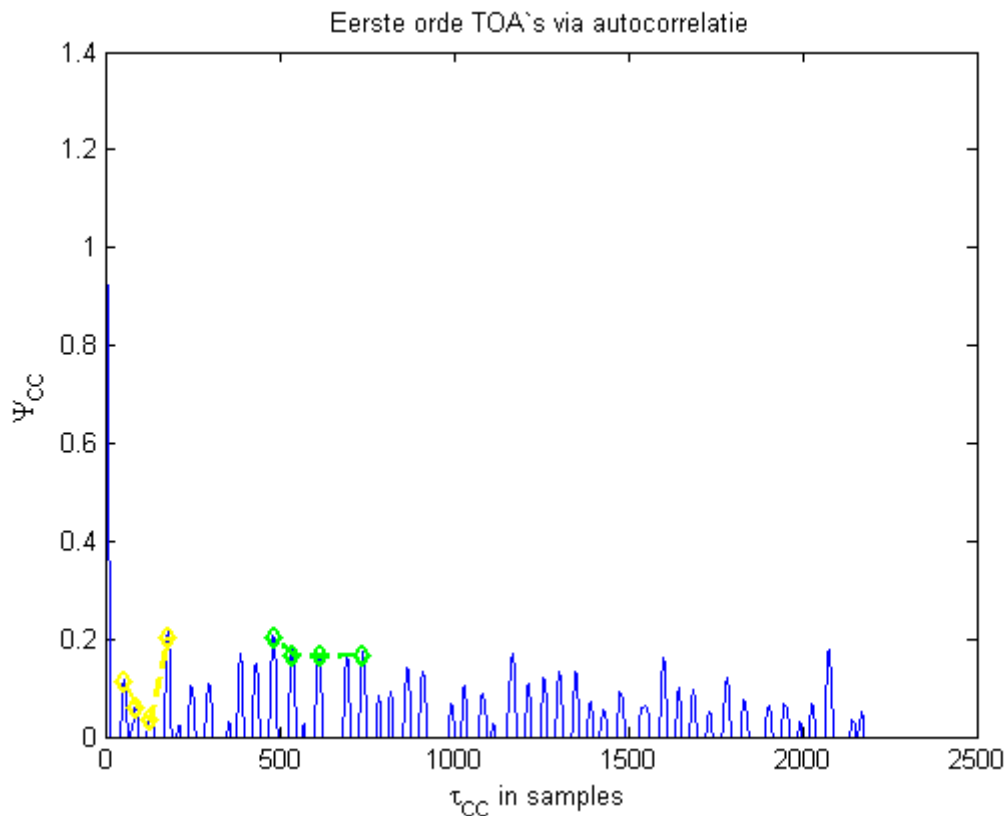
De kruiscorrelatie (CC) methode is een veelgebruikte methode om het verschil in tijdsvertraging te zoeken tussen twee signalen [16]. Voor een bepaald bereik van mogelijke tijdsvertragingen τ_{CC} wordt telkens de kruiscorrelatie berekend tussen beide signalen. De τ_{CC} waarbij deze kruiscorrelatie de hoogste waarde oplevert wordt dan gezien als het resultaat voor de gezochte tijdsvertraging. Via de kruiscorrelatie is er geen nood aan kennis van het bronsignaal. Het algoritme werkt op paren van microfoonsignalen. Op deze manier kan rechtstreeks de TDOA (Time Difference of Arrival) bepaald worden tussen de twee microfoonsignalen. Zoals in hoofdstuk 3 besproken kunnen deze TDOA's gebruikt worden om de bron te lokaliseren. In formulevorm krijgen we:

$$\hat{\tau}_{CC} = \arg \max_m \Phi_{CC}(m), \quad \text{met } \Phi_{CC}(m) = E \{x_0[n]x_1[n+m]\} \quad (2.3)$$

Hierbij is $x_0[n]$ het eerste microfoonsignaal en $x_1[n+m]$ het tweede signaal verschoven over m samples. In de implementatie wordt de verwachte waarde $E \{ \cdot \}$ vervangen door een gemiddelde over een reeks samples uit het signaal. Ook wordt het resultaat

genormaliseerd via de standaarddeviaties van beide signalen.

Zoals uitgewerkt in hoofdstuk 4, is een schatting van de vier eerste orde TOA's benodigd voor de 2D reconstructie van een kamergeometrie. Deze kunnen ook geschat worden via het CC algoritme, waarbij dat een correlatie van elk microfoonsignaal met zichzelf wordt uitgevoerd. Figuur 2.3 toont het resultaat. De grootste waarde in het resultaat ligt uiteraard rond nul. Daarna worden de vier TOA's ofwel gezocht als de vier volgende maxima, ofwel via een methode die kijkt naar de verhouding met de amplitude van de voorgaande pieken. Dit wordt meer in detail besproken in paragraaf 4.1.1. Het resultaat waarbij de eerste orde TOA's gevonden worden als de vier maxima, is aangeduid in het groen. Het resultaat van de andere methode is aangeduid in het geel. In figuur 4.16 wordt het resultaat getoond voor reflectorlokalisatie die gebruik maakt van de TOA's via het CC algoritme.



FIGUUR 2.3: Zoeken van direct pad en eerste orde TOA's via kruiscorrelatie.

2.3.2 Closed-Form schatting

Een tweede methode om de RIR's blind te schatten is via een kleinste kwadraten methode [17]. Dit resulteert in een gesloten uitdrukking, waaruit de impulsresponsies tegelijk berekend kunnen worden. Een voordeel ten opzichte van de kruiscorrelatie

methode, is dat de methode ook reverberatie effecten in rekening brengt. Dit heeft als resultaat dat ook hogere orde TOA's nauwkeurig geschat kunnen worden, waarvoor het kruiscorrelatie algoritme minder geschikt is. Het principe waarop het algoritme gebaseerd is, is beschreven via de volgende relatie:

$$\mathbf{x}_1^T[n]\mathbf{g}_2 = \mathbf{x}_2^T[n]\mathbf{g}_1 \quad (2.4)$$

met

$$\mathbf{x}_i[n] = [x_i[n] \quad x_i[n-1] \quad \cdots \quad x_i[n-L+1]]^T$$

Hierbij zijn \mathbf{x}_1 en \mathbf{x}_2 vectoren die samples van de twee microfoonsignalen bevatten, en \mathbf{g}_1 en \mathbf{g}_2 de bijhorende impulsresponsies met lengte L . De relatie volgt uit het feit dat beide signalen van dezelfde bron afkomstig zijn. De covariantiematrix van deze twee microfoonsignalen kan geschreven worden als:

$$\Sigma = \begin{bmatrix} \Sigma_{x_1x_1} & \Sigma_{x_1x_2} \\ \Sigma_{x_2x_1} & \Sigma_{x_2x_2} \end{bmatrix}, \text{ met } \Sigma_{x_ix_j} = E\{\mathbf{x}_i(n)\mathbf{x}_j^T(n)\} \quad (2.5)$$

Via het definiëren van \mathbf{h} en toepassing van (2.4) levert dit:

$$\mathbf{h} = \begin{bmatrix} \mathbf{g}_2 \\ -\mathbf{g}_1 \end{bmatrix} \Rightarrow \Sigma\mathbf{h} = 0 \quad (2.6)$$

Omdat deze uitdrukking geldt voor elk paar van microfoons i en j en voor elke set van K samples genomen uit deze microfoonsignalen, kan het volgend stelsel van $K - 2L + 3$ vergelijkingen opgesteld worden, waarin de verwachte waarden vervangen worden door de microfoonsignalen voor een aantal samples:

$$\begin{bmatrix} \mathbf{X}_i(L) & | & -\mathbf{X}_j(L) \end{bmatrix} \begin{bmatrix} \mathbf{g}_i \\ \mathbf{g}_j \end{bmatrix} = 0 \quad (2.7)$$

Hierin zijn \mathbf{g}_i en \mathbf{g}_j opnieuw de impulsresponsies, beide met een lengte L . De matrix $\mathbf{X}_m(L')$ wordt gegeven door:

$$\begin{bmatrix} x_m(L') & x_m(L'+1) & \cdots & x_m(2L') \\ x_m(L'+1) & x_m(L'+2) & \cdots & x_m(2L'+1) \\ \vdots & \vdots & \ddots & \vdots \\ x_m(K-L') & x_m(K-L'+1) & \cdots & x_m(K) \end{bmatrix}, \text{ met } L' = L - 1 \quad (2.8)$$

Deze laatste matrix heeft een dimensie van $(K-2L'+1) \times (L'+1) = (K-2L+3) \times (L)$. Dit stel vergelijkingen kan geschreven worden voor elk microfoonpaar, waaruit dan de onbekende \mathbf{g}_i en \mathbf{g}_j kunnen worden bepaald. Zoals in [17] beschreven is het ook mogelijk de RIR's voor alle microfoons tegelijk te schatten via volgende matrix:

$$\mathbf{X}^i(L') = \underbrace{\begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{X}_{i+1}(L') & -\mathbf{X}_i(L') & \mathbf{0} & \mathbf{0} \\ \vdots & & & \vdots & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{X}_M(L') & \mathbf{0} & \cdots & -\mathbf{X}_i(L') \end{bmatrix}}_{M \text{ blokken}} \left. \vphantom{\begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{X}_{i+1}(L') & -\mathbf{X}_i(L') & \mathbf{0} & \mathbf{0} \\ \vdots & & & \vdots & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{X}_M(L') & \mathbf{0} & \cdots & -\mathbf{X}_i(L') \end{bmatrix}} \right\} M - i \text{ blokken} \quad (2.9)$$

Hierbij staat M voor het aantal microfoons en i loopt van 1 tot M . Elk blok in deze matrix heeft een grootte $(K - 2L + 3) \times (L)$. Wanneer de volgende matrix wordt gedefinieerd als:

$$\mathbf{X}(L') = \left[\begin{array}{c} \mathbf{X}^1(L') \\ \vdots \\ \mathbf{X}^{M-1}(L') \end{array} \right] \left. \vphantom{\begin{array}{c} \mathbf{X}^1(L') \\ \vdots \\ \mathbf{X}^{M-1}(L') \end{array}} \right\} \frac{M(M-1)}{2} \text{blokken} \quad (2.10)$$

welke een dimensie heeft van $(K - 2L + 3) \left(\frac{M(M-1)}{2}\right) \times (ML)$, kan een gelijkaardige uitdrukking als (2.6) bekomen worden:

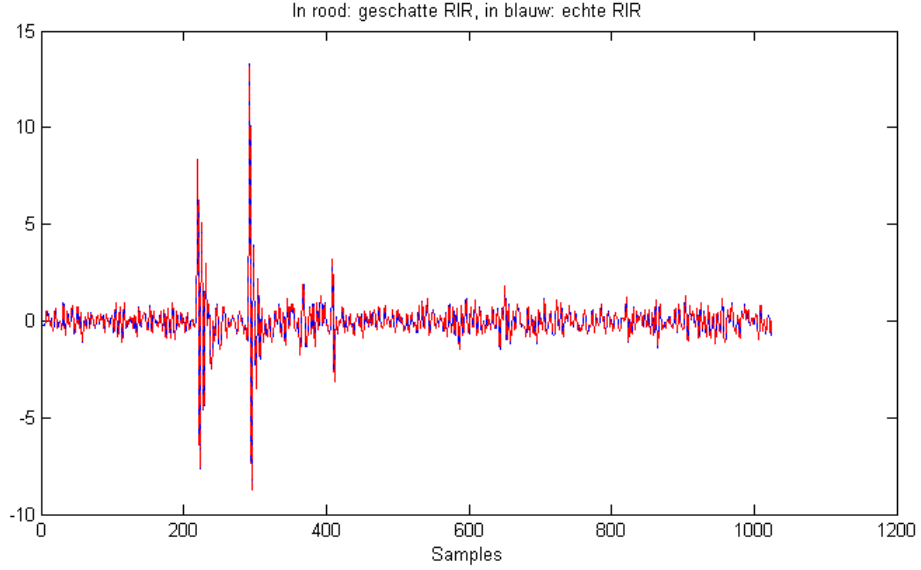
$$\mathbf{X}(L')\mathbf{G} = \mathbf{0}, \quad \text{met } \mathbf{G} = [\mathbf{g}_1^T, \dots, \mathbf{g}_M^T]^T \quad (2.11)$$

Het resultaat is een lineaire vergelijking waaruit de impulsresponsies berekend kunnen worden. De interpretatie van vergelijking (2.11) is dat de gezochte impulsresponsies bepaald kunnen worden als de nulruimte van matrix $\mathbf{X}(L')$. Het probleem hier is het goed kunnen schatten van de parameter L . Indien deze overschat wordt, zal de gevonden nulruimte meerdere vectoren bevatten. In de literatuur [18] zijn methodes beschreven om uit deze vectoren de gewenste vector \mathbf{G} te construeren. Het aantal vectoren in de nulruimte is een indicatie van de overschatting is van L , namelijk het aantal vectoren in de ruimte verminderd met één. Een mogelijkheid is dus ook om opnieuw het algoritme uit te voeren, met de juiste waarde voor L . Het nadeel van deze aanpak is natuurlijk de lange rekentijd. Aangezien met grote matrices gewerkt wordt, waarvan de nulruimte moet worden gevonden, verloopt het algoritme vrij traag.

Indien L echter onderschat wordt, resulteert dit in een lege nulruimte. Dit probleem kan van crucialere aard zijn, omdat het niet altijd mogelijk is om L nog te vergroten in praktijk. Dit levert problemen op met het benodigd geheugen voor het algoritme. In hoofdstuk 5 wordt hierop teruggekomen. Figuur 2.4 toont een voorbeeld van een kanaalschatting via dit algoritme. Twee willekeurige inputsignalen, in dit geval gekozen als witte ruis, worden beide eerst gefilterd met een verschillende RIR. In de figuur wordt deze voor het eerste signaal weergegeven in blauw. De lengte van beide RIR's bedraagt $L = 1024$. Op deze twee signalen wordt het algoritme toegepast, met als doel de beide RIR's te kunnen reconstrueren. Hierbij wordt in het algoritme $L = 1024$ gebruikt zodat, zoals eerder vermeld, er slechts één oplossing wordt gevonden. Het resultaat van de geschatte RIR voor het eerste signaal is in de figuur aangegeven in het rood.

2.3.3 AED algoritme

Het AED of Adaptive Eigenvalue Decomposition algoritme gaat de akoestische impulsresponsies voor elk pad van bron naar microfoon trachten te schatten via een adaptief filter. Net zoals het closed-form algoritme modelleert AED niet enkel de tijdsvertraging, maar een impulsresponsie met reverberatie effecten. Dit heeft als gevolg dat de TDOA nauwkeuriger bepaald kan worden [19]. Daarnaast zou



FIGUUR 2.4: Voorbeeld van reconstructie van een RIR via het closed-form algoritme.

volgens de literatuur deze methode ook een schatting van de eerste en hogere orde reflecties moeten opleveren, zoals het NLMS algoritme en het closed-form algoritme. In vergelijking met het closed-form algoritme werkt deze methode op paren van microfoons, en adaptief. Het algoritme wordt herhaald naargelang het aantal te schatten responsies. Het principe waarop het algoritme gebaseerd is, is hetzelfde als (2.4). Het AED-algoritme gaat de uitdrukking (2.6) iteratief evalueren door \mathbf{w} (filtercoëfficiënten) te zoeken met behulp van een kleinste kwadraten oplossing, welke de uitdrukking $\mathbf{w}^T \Sigma \mathbf{w}$ minimaliseert. Uitwerking van het kleinste kwadraten probleem en toepassen van de beperking $\mathbf{w}^T \mathbf{w} = 1$ levert de volgende adaptatieformule, met F de iteratiestap [19]:

$$\mathbf{w}(F+1) = \frac{\mathbf{w}(F) - \mu * e(F) * \mathbf{x}(F)}{\|\mathbf{w}(F) - \mu * e(F) * \mathbf{x}(F)\|} \quad (2.12)$$

waarbij

$$e(F) = \mathbf{w}^T(F) \mathbf{x}(F) \quad (2.13)$$

De startwaarde voor vector \mathbf{w} wordt gekozen als een eenheidsvector met het $\frac{L}{2}$ -de element gelijk aan 1 en de andere elementen gelijk aan 0. L is de gekozen lengte van een impulsresponsie, zodat de lengte van \mathbf{w} $2L$ is. L moet groot genoeg zijn zodat de TOA van het directe pad gevonden kan worden, hier is $L = 2048$ gekozen. Als stapgrootte is $\mu = 2$ gekozen. Voor de bronlokalisatie is alleen de TOA van het directe pad nodig. Deze wordt gevonden als de index waar \mathbf{w} maximaal is. De TOA wordt gevonden door de index om te zetten naar seconden via Fs , de samplefrequentie. In principe zou dit algoritme ook kunnen gebruikt worden om de reflectorlokalisatie uit te voeren. Hiervoor zijn zoals eerder vermeld de eerste

orde TOA's nodig, of zelfs in combinatie met de tweede orde reflecties afhankelijk van de gebruikte methode. Uit testen van het AED algoritme blijkt echter dat deze eerste en hogere orde TOA's moeilijk of niet te zien zijn in de bekomen impulsresponsies. Voor de TOA van het directe pad is deze methode wel bruikbaar, hoewel er bij de praktische metingen ook problemen optreden (zie hoofdstuk 5). Voor het blind schatten van de eerste orde TOA's blijkt de closed-form methode het beste resultaat te geven.

2.4 Besluit

Na een studie en implementatie van de besproken algoritmes om blinde kanaalschatting uit te voeren, blijkt dat hier toch enkele problemen optreden. In het geval van een adaptieve benadering, meer specifiek het AED-algoritme, geeft de methode een goed resultaat voor de schatting van het directe pad maar niet voor de hogere orde reflecties. Hoewel het algoritme rekening houdt met reverberatie effecten, zijn deze weinig of niet terug te vinden in de bekomen schattingen. Dit is wel het geval bij het closed-form algoritme, welke de impulsresponsies nauwkeurig kan terugvinden. Het probleem hier is eerder van numerieke aard, om een goed resultaat te verkrijgen moet de geschatte lengte van de impulsresponsie groot genoeg zijn. Indien deze onderschat wordt, levert de methode geen resultaat. Ook vraagt de methode veel rekenkracht. Kruiscorrelatie is zoals AED geschikt voor de schatting van het directe pad, maar minder geschikt voor eerste orde reflecties of hogere.

Indien het bronsignaal gekend is, kunnen eenvoudigere algoritmes gebruikt worden. Het meest voor de hand liggende is een adaptief algoritme, zoals NLMS. Na simulaties blijkt het NLMS algoritme reeds impulsresponsies te leveren die voldoende nauwkeurig zijn voor verder gebruik bij het lokaliseren van de reflectoren en indien nodig ook de bron.

Hoofdstuk 3

Bronlocatie bepalen

Zoals reeds aangegeven in de inleiding is het eerste opzet van deze thesis een omgeving te kunnen reconstrueren aan de hand van geluidssignalen. Meer specifiek is het eerste onderzoekstopic het bepalen van afmetingen van een kamer of gebouw waarin de geluidsbron en microfoons aanwezig zijn. Het bepalen van de dimensies van de omgeving wordt gedaan door de positie en oriëntatie te zoeken van wanden, vloer, plafond en dergelijke van de kamer of het gebouw. Om dit te bewerkstelligen wordt een strategie in twee stappen voorgesteld. De eerste stap bestaat erin de locatie van de bron ten opzichte van de microfoons te bepalen. De tweede stap is dan het eigenlijke zoeken naar de locatie van de reflectoren. De bronlokalisatie wordt in dit hoofdstuk besproken. De tweede stap wordt besproken in hoofdstuk 4. Telkens wordt verondersteld dat de posities van de microfoons relatief ten opzichte van mekaar gekend zijn, en dat de bron volledig ongekend is (zowel zijn uitgezonden signaal als zijn positie). Een voordeel van deze ‘passieve’ manier is dat de bron nu heel algemeen is, en bijvoorbeeld ook een spreker in de kamer kan zijn. Er bestaat ook vrij recent onderzoek naar ‘actieve’ methodes, waarbij één microfoon en één gecontroleerde bewegende zender wordt gebruikt [8]. Hier wordt verder niet op ingegaan, maar de gebruikte formules om op deze manier de locatie van de reflectoren te bepalen zijn vrij gelijkaardig aan deze voor de passieve methode.

3.1 Bronlokalisatie in twee dimensies

Het tweedimensionale geval wordt eerst bestudeerd. De strategie om de bronlocatie te zoeken is gebaseerd op het werk van J.Filos en S.Abel, [20] en [21]. Aangezien enkel de posities van de microfoons relatief ten opzichte van mekaar gekend zijn, en we slechts één aanwezige bron beschouwen, is het mogelijk om de bronlocatie te berekenen door gebruik te maken van verschil in aankomsttijden van het bronsignaal tussen de verschillende microfoons. Enkel het directe pad van bron naar microfoon dient beschouwd te worden. Deze verschiltijden in aankomst (TDOA’s) worden gevonden met het eerder besproken AED-algoritme. Om de bron te lokaliseren zijn minimaal drie niet collineair gepositioneerde microfoons nodig. Dit kan eenvoudig ingezien worden omdat men rond elke microfoon een cirkel kan construeren als

verzameling van mogelijke locaties van de bron, met de TOA gekend. Cirkels van twee verschillende microfoons geven een maximum van twee snijpunten. Er is dus nog een derde cirkel nodig met een middelpunt dat niet op de rechte door de middelpunten van de eerste twee ligt om één positie van de twee te kiezen. Het bepalen van de bronlocatie aan de hand van TDOA's wordt in de volgende sectie besproken.

3.1.1 Theoretisch

Als de TDOA's voor elk microfoonpaar via het AED algoritme berekend zijn, kan de locatie van de bron gevonden worden. Beschouw M microfoons in het 2D-vlak, op locaties:

$$r_i = [x_i \ y_i]^T, \quad i = 1, \dots, M \quad (3.1)$$

met de referentiemicrofoon geplaatst in de oorsprong

$$r_1 = [0 \ 0]^T \quad (3.2)$$

en de bron op ongekende locatie

$$r_s = [x_s \ y_s]^T \quad (3.3)$$

De afstanden van de i -de microfoon en de bron tot de oorsprong (respectievelijk R_i en R_s) zijn gedefinieerd als volgt:

$$R_i = \| r_i \| = \sqrt{x_i^2 + y_i^2}, \quad i = 2, \dots, M \quad (3.4)$$

$$R_s = \| r_s \| = \sqrt{x_s^2 + y_s^2} \quad (3.5)$$

Het verschil in afstand van microfoon i en microfoon j tot de bron is gedefinieerd als $d_{i,j}$ (bereikverschil) en kan via de geluidssnelheid η gerelateerd worden met de TDOA van het directe pad tussen deze microfoons $\Delta_{i,j}$ als volgt:

$$d_{i,j} = \eta \Delta_{i,j} \quad (3.6)$$

De bronlocatie kan bepaald worden met behulp van kleinste kwadraten. De locatie van de bron zal op de intersectie van een groep van cirkels liggen, waarbij het centrum van elke cirkel ligt op de locatie van een microfoon, en de straal van elke cirkel gerelateerd is tot de afstand tussen bron en microfoon. De beste schatting van de bronlocatie zal dus het punt zijn dat het dichtst bij deze cirkels ligt. De exacte straal van elke cirkel is niet gekend, maar wordt geschat via de bereikverschillen en de veronderstelde afstand van de bron tot de oorsprong. Om dit probleem aan te pakken wordt een foutenfunctie gedefinieerd als het verschil tussen de opgemeten en echte waarden. Deze foutenfunctie zal geminimaliseerd worden via kleinste kwadraten (LS). In wat volgt zal een opgemeten waarde aangegeven worden door $\hat{(\cdot)}$. De geschatte afstand tussen de i -de microfoon en de bron kan in functie van de metingen geschreven worden als:

$$\hat{D}_i = R_s + \hat{d}_{i,1} \quad (3.7)$$

Hierbij wordt R_s als reeds exact gekend verondersteld, maar dit stelt geen probleem omdat R_s later in de berekeningen weg valt, en dus niet meer relevant is. Het kwadraat van de werkelijke afstand tussen de i -de microfoon en de bron kan als volgt worden geschreven (m.b.v. de cosinusregel):

$$D_i^2 = R_i^2 + R_s^2 - 2R_i R_s \frac{r_i^T r_s}{R_i R_s} \quad (3.8)$$

$$D_i^2 = R_i^2 + R_s^2 - 2r_i^T r_s \quad (3.9)$$

De foutenfuntie wordt nu gedefinieerd als $\epsilon_i = \frac{1}{2}(\hat{D}_i^2 - D_i^2)$ (de factor $\frac{1}{2}$ is louter om de notatie later te vereenvoudigen):

$$\epsilon_i = r_i^T r_s + \hat{d}_{i,1} R_s - \frac{1}{2} R_i^2 + \frac{1}{2} \hat{d}_{i,1}^2, \quad i = 2, \dots, M \quad (3.10)$$

Deze $M-1$ fouten kunnen tezamen in matrixvorm worden geschreven:

$$\mathbf{e}(r_s) = \mathbf{A}\Theta - \mathbf{b} \quad (3.11)$$

$$\text{waarbij } \mathbf{A} = [\mathbf{S} \mid \hat{\mathbf{d}}], \quad \mathbf{S} = \begin{bmatrix} x_2 & y_2 \\ x_3 & y_3 \\ \vdots & \vdots \\ x_M & y_M \end{bmatrix}$$

$$\Theta = \begin{bmatrix} x_s \\ y_s \\ R_s \end{bmatrix}, \quad \mathbf{b} = \frac{1}{2} \begin{bmatrix} R_2^2 - \hat{d}_{2,1}^2 \\ R_3^2 - \hat{d}_{3,1}^2 \\ \vdots \\ R_M^2 - \hat{d}_{M,1}^2 \end{bmatrix}$$

$$\hat{\mathbf{d}}^T = [\hat{d}_{2,1} \quad \hat{d}_{3,1} \quad \dots \quad \hat{d}_{M,1}]$$

Het LS-criterium wordt dan gegeven door: $J = \mathbf{e}^T \mathbf{e}$. En de oplossing van dit LS-probleem voorde bronlocatie Θ is dan:

$$\hat{\Theta} = \mathbf{A}^+ \mathbf{b} \quad (3.12)$$

waarbij de $(\cdot)^+$ duidt op de pseudo-inverse.

3.1.2 Simulatie

Om de theorie te toetsen wordt het algoritme gesimuleerd. Via een model van een kamer met vier vlakke wanden kunnen de impulsresponsies voor het kanaal tussen bron en microfoon worden gevonden [22]. Hierbij is de reflectiecoëfficiënt van de wanden ingesteld op $R = 0.7$. De responsies geconvolveerd met een bronsignaal leveren de gesimuleerde microfoonsignalen. Deze signalen fungeren als ingang van het bronlokalisatiealgoritme. De TDOA's worden met het AED-algoritme geschat. Om meer redundantie te hebben, worden voor het tweedimensionale geval vier microfoons

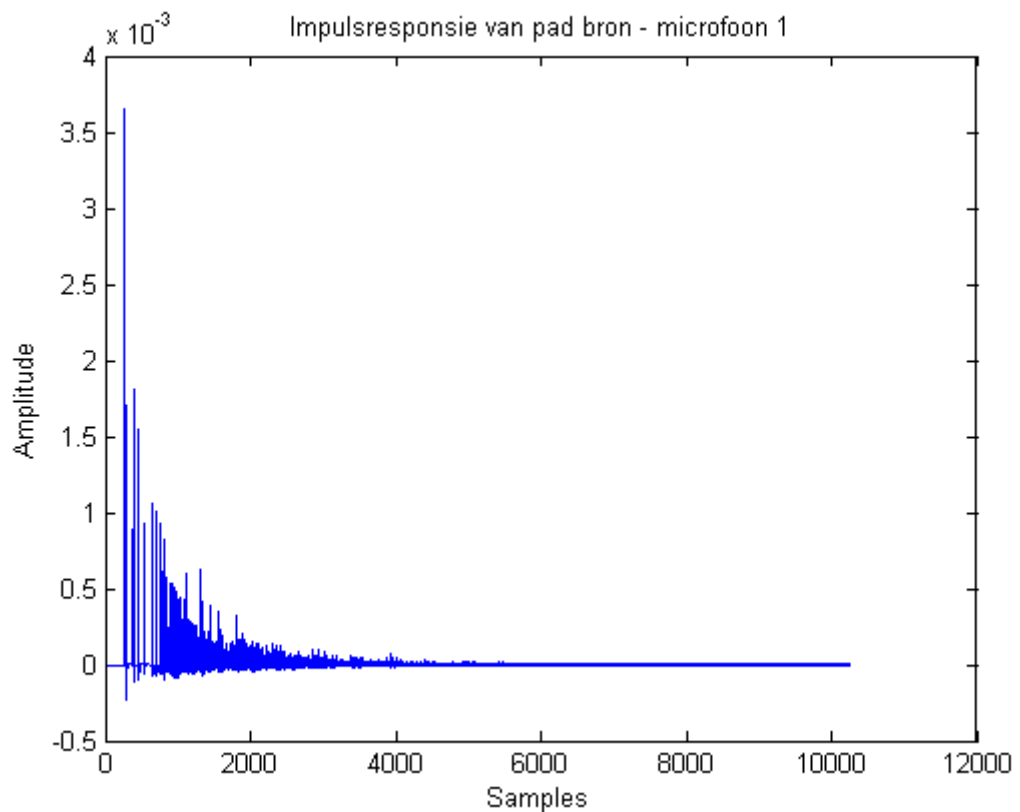
3. BRONLOCATIE BEPALEN

gebruikt, in plaats van het minimum van drie. Het gebruikte bronsignaal is witte ruis. De afmetingen van de simulatieruimte komen overeen met de afmetingen van de kamer waarin later de praktijktesten zullen gebeuren in hoofdstuk 5. De lengte van de kamer (x-as) bedraagt 3.76 m, en de breedte (y-as) bedraagt 6 m. Ook de microfoonposities zijn gekozen zoals in de praktijktest. Deze locaties zijn samengevat in tabel 3.1. De impulsresponsies die gegenereerd worden door het model zijn voor

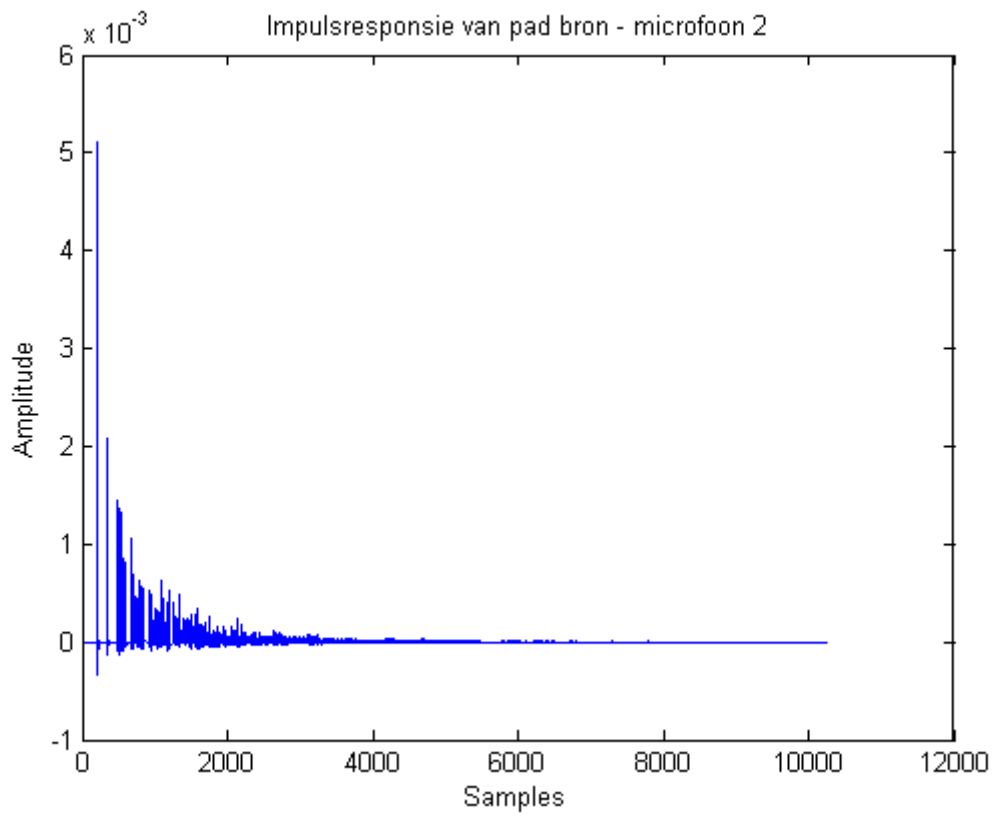
TABEL 3.1: Locaties van bron en microfoons in 2D-ruimte (in meter).

	x	y
bron	1.1	1.9
microfoon 1	1.75	3.9
microfoon 2	2.6	1.93
microfoon 3	0.7	3
microfoon 4	0.3	1.1

elke microfoon weergegeven in afbeeldingen 3.1, 3.2, 3.3 en 3.4.

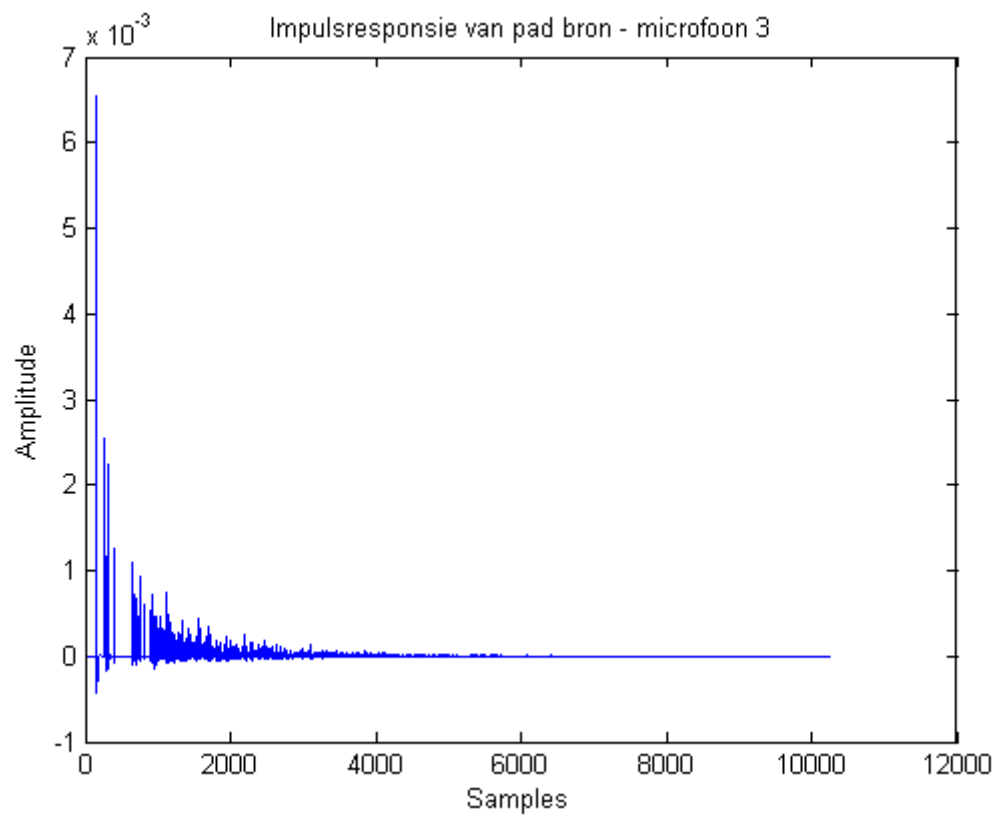


FIGUUR 3.1: Gesimuleerde impulsresponsie van het kanaal tussen bron en microfoon 1.

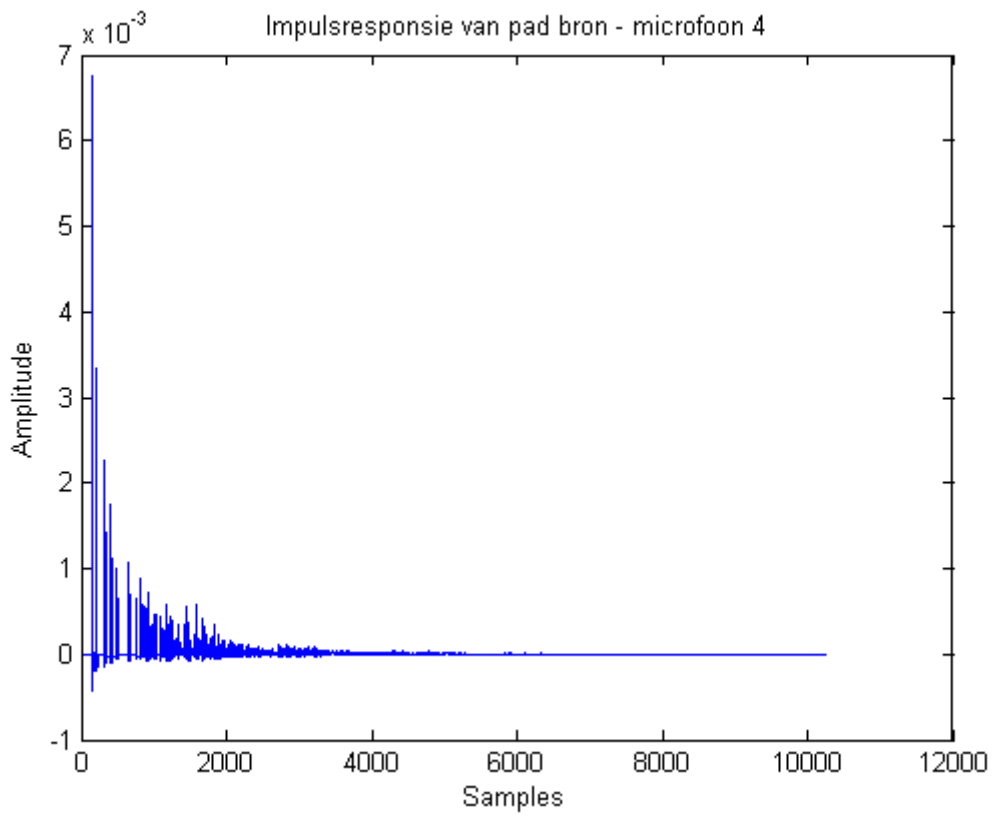


FIGUUR 3.2: Gesimuleerde impulsresponsie van het kanaal tussen bron en microfoon 2.

3. BRONLOCATIE BEPALEN

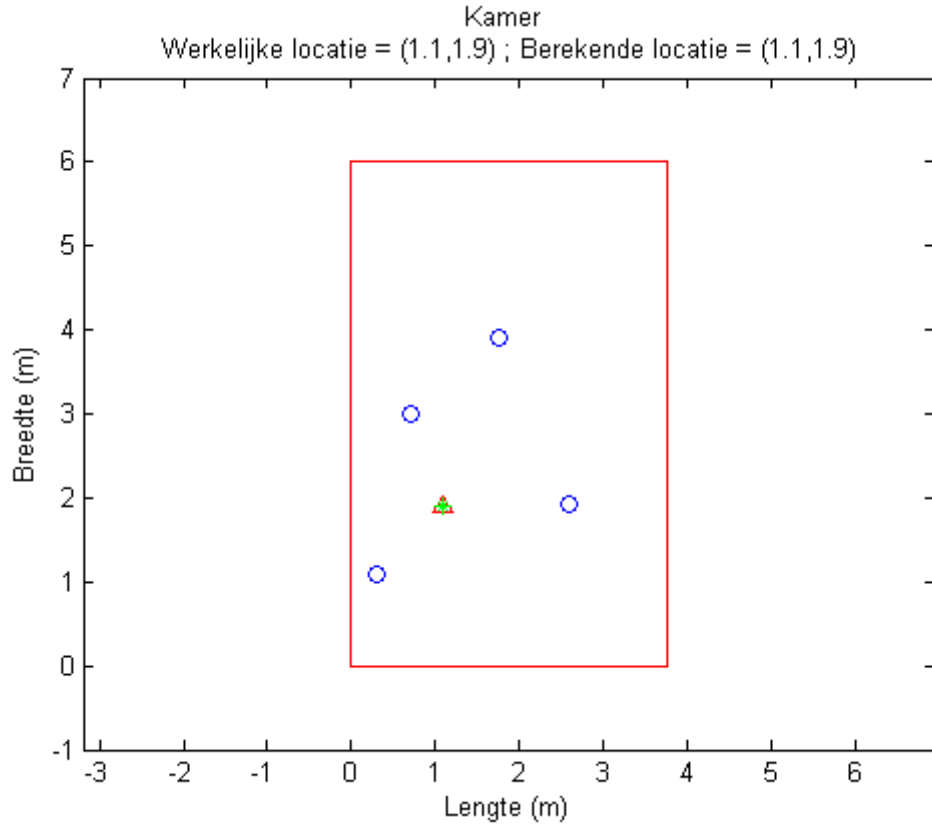


FIGUUR 3.3: Gesimuleerde impulsresponsie van het kanaal tussen bron en microfoon 3.



FIGUUR 3.4: Gesimuleerde impulsresponsie van het kanaal tussen bron en microfoon 4.

De ingangssignalen voor het AED-algoritme zijn zoals reeds besproken de filtering van de witte ruis met de impulsresponsies. Vervolgens wordt de locatie van de bron gezocht, met de voorgestelde methode. Het resultaat is weergegeven in figuur 3.5.



FIGUUR 3.5: Resultaat van simulatie van bronlokalisatie in 2D-ruimte (\circ : microfoon; Δ : werkelijke locatie; $*$: berekende locatie).

3.2 Bronlokalisatie in drie dimensies

De 3D uitbreiding van het bronlokalisatiealgoritme wordt in deze sectie besproken. Als model wordt van een kamer met vier vlakke wanden, een vlak plafond, en een vlakke vloer vertrokken. De strategie is hier net dezelfde als bij het tweedimensionale geval, met dien verstande dat hier met bollen wordt gewerkt, in plaats van cirkels. De strategie is gebaseerd op [21]. Als de TOA van een pad tussen bron en microfoon gegeven is, is geweten dat alle mogelijke locaties van de bron voor deze microfoon op een boloppervlak liggen. Dit impliceert dat er een minimum van vier microfoons aanwezig moet zijn om een oplossing te bekomen. De TDOA's worden via het AED-algoritme bepaald, zoals in het tweedimensionale geval.

3.2.1 Theoretisch

Met de TDOA's gekend kan dus zoals eerder gesteld dezelfde methodiek gevolgd worden als in het tweedimensionale geval. De formules blijven bijna hetzelfde, enkel een z-componente voor de hoogte in de kamer is toegevoegd. Beschouw M microfoons in de 3D-ruimte, op locaties:

$$r_i = [x_i \ y_i \ z_i]^T, \quad i = 1, \dots, M \quad (3.13)$$

met de referentiemicrofoon geplaatst in de oorsprong

$$r_1 = [0 \ 0 \ 0]^T \quad (3.14)$$

en de bron op ongekende locatie

$$r_s = [x_s \ y_s \ z_s]^T \quad (3.15)$$

De afstanden van de i -de microfoon en de bron tot de oorsprong zijn hier gedefinieerd als volgt:

$$R_i = \| r_i \| = \sqrt{x_i^2 + y_i^2 + z_i^2}, \quad i = 2, \dots, M \quad (3.16)$$

$$R_s = \| r_s \| = \sqrt{x_s^2 + y_s^2 + z_s^2} \quad (3.17)$$

Ook hier kan het bereikverschil weer via de geluidssnelheid η met de TDOA van het directe pad tussen microfoon i en microfoon j , $\Delta_{i,j}$, gerelateerd worden via (3.6). De bronlocatie wordt bepaald met behulp van kleinste kwadraten. De locatie van de bron zal op de intersectie van een groep van bollen liggen, waarbij het centrum van elke sfeer op de locatie van een microfoon ligt, en de straal opnieuw gerelateerd is tot de afstand tussen bron en microfoon. Om het probleem aan te pakken wordt eenzelfde foutenfunctie gedefinieerd als in het tweedimensionale geval, nl. het verschil tussen de opgemeten en echte waarden. De formules van toepassing zijn exact dezelfde als in het 2D geval. De $M-1$ fouten worden tezamen in matrixvorm geschreven:

$$\mathbf{e}(r_s) = \mathbf{A}\Theta - \mathbf{b} \quad (3.18)$$

$$\text{waarbij } \mathbf{A} = [\mathbf{S} \mid \hat{\mathbf{d}}], \quad \mathbf{S} = \begin{bmatrix} x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \\ \vdots & \vdots & \vdots \\ x_M & y_M & z_M \end{bmatrix}$$

$$\Theta = \begin{bmatrix} x_s \\ y_s \\ z_s \\ R_s \end{bmatrix}, \quad \mathbf{b} = \frac{1}{2} \begin{bmatrix} R_2^2 - \hat{d}_{2,1}^2 \\ R_3^2 - \hat{d}_{3,1}^2 \\ \vdots \\ R_M^2 - \hat{d}_{M,1}^2 \end{bmatrix}$$

$$\hat{\mathbf{d}}^T = [\hat{d}_{2,1} \ \hat{d}_{3,1} \ \dots \ \hat{d}_{M,1}]$$

Het LS-criterium is gegeven door: $J = \mathbf{e}^T \mathbf{e}$. En de oplossing van dit LS-probleem voor Θ is dan dezelfde als voor het 2D-geval:

$$\hat{\Theta} = \mathbf{A}^+ \mathbf{b} \quad (3.19)$$

waarbij de $(\cdot)^+$ duidt op de pseudo-inverse.

3.2.2 Simulatie

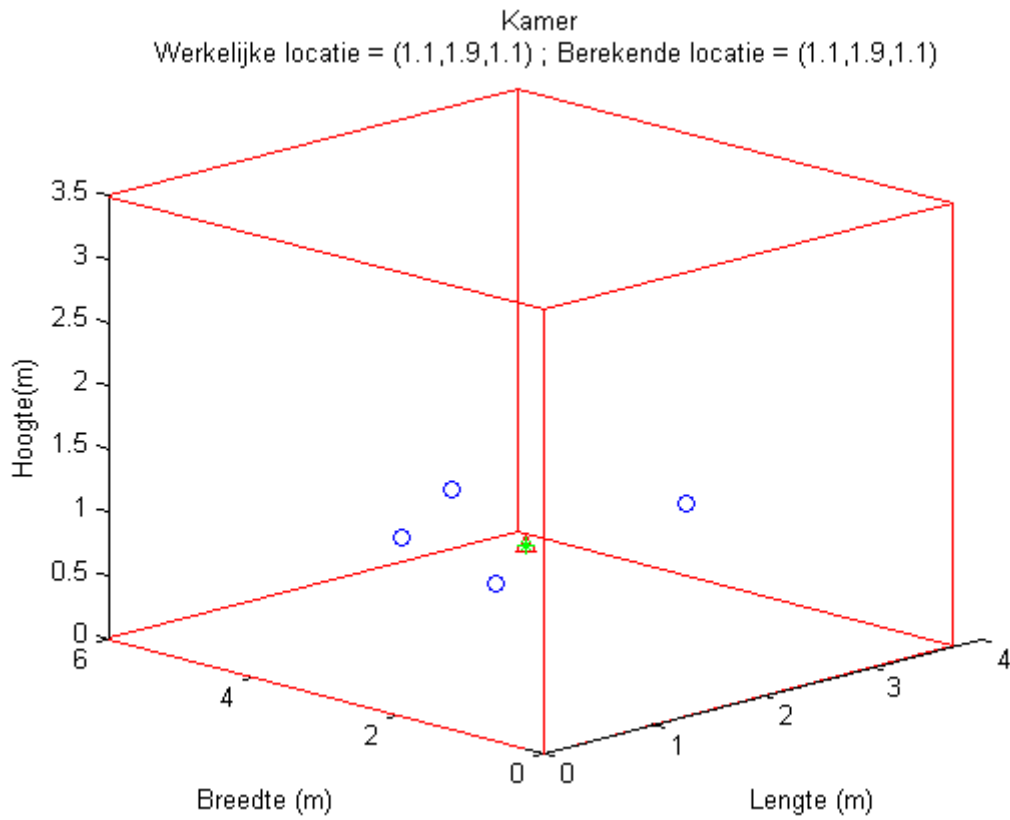
De uitbreiding naar de 3D-theorie wordt ook gesimuleerd. Het model dat nu voor de kamer gebruikt wordt, bevat naast de vier muren ook plafond en vloer. Zo kunnen dus meer reflecties worden waargenomen in de impulsresponsies. Uit het model volgen de impulsresponsies van het pad tussen bron en microfoon. Om de simulatie van de driedimensionale bronlokalisatie uit te voeren worden vier microfoons verondersteld en één bron, op locaties samengevat in tabel 3.2. Het bronsignaal is het witte ruis signaal dat ook eerder in de 2D-simulatie gebruikt is. Lengte en breedte van de kamer blijven ongewijzigd (3.76 m en 6 m respectievelijk). Enkel de hoogte wordt in rekening gebracht. Deze bedraagt 3.5 m.

TABEL 3.2: Locaties van bron en microfoons in 3D-ruimte (in meter).

	x	y	z
bron	1.1	1.9	1.1
microfoon 1	1.75	3.9	1.1
microfoon 2	2.6	1.93	1.1
microfoon 3	0.7	3	1.1
microfoon 4	0.3	1.1	1.1

Het algoritme simuleren voor een 3D-ruimte levert de oplossing die weergegeven is in figuur 3.6. De \circ -symbolen zijn de microfoons, de $*$ stelt de berekende bronlocatie voor, en de Δ is de werkelijke bronlocatie. Het resultaat van figuur 3.6 is gunstig. Maar er dient opgemerkt te worden dat dit een speciaal geval is, aangezien de z-coördinaat van de microfoons dezelfde is. Om meer naar een echte 3D-situatie te gaan, wordt de methode ook getest waarin de microfoons verschillende z-coördinaten hebben. Een voorbeeld hiervan is de opstelling weergegeven in tabel 3.3. De lengte van de kamer bedraagt 9 m en de breedte 10 m. De hoogte wordt gekozen op 4 m. Figuur 3.7 (\circ : microfoon; Δ : werkelijke locatie; $*$: berekende locatie) toont het resultaat, welk opnieuw gunstig is.

Het resultaat is echter niet altijd gunstig. Een voorbeeld van een dergelijke situatie is gegeven in figuur 3.8 (\circ : microfoon; Δ : werkelijke locatie; $*$: berekende locatie), waarbij de locaties van bron en microfoons zijn samengevat in tabel 3.4. De resultaten blijken beter te zijn indien de microfoons voldoende verspreid staan ten opzichte van elkaar, en de bron ongeveer centraal gelegen tussen de microfoons.

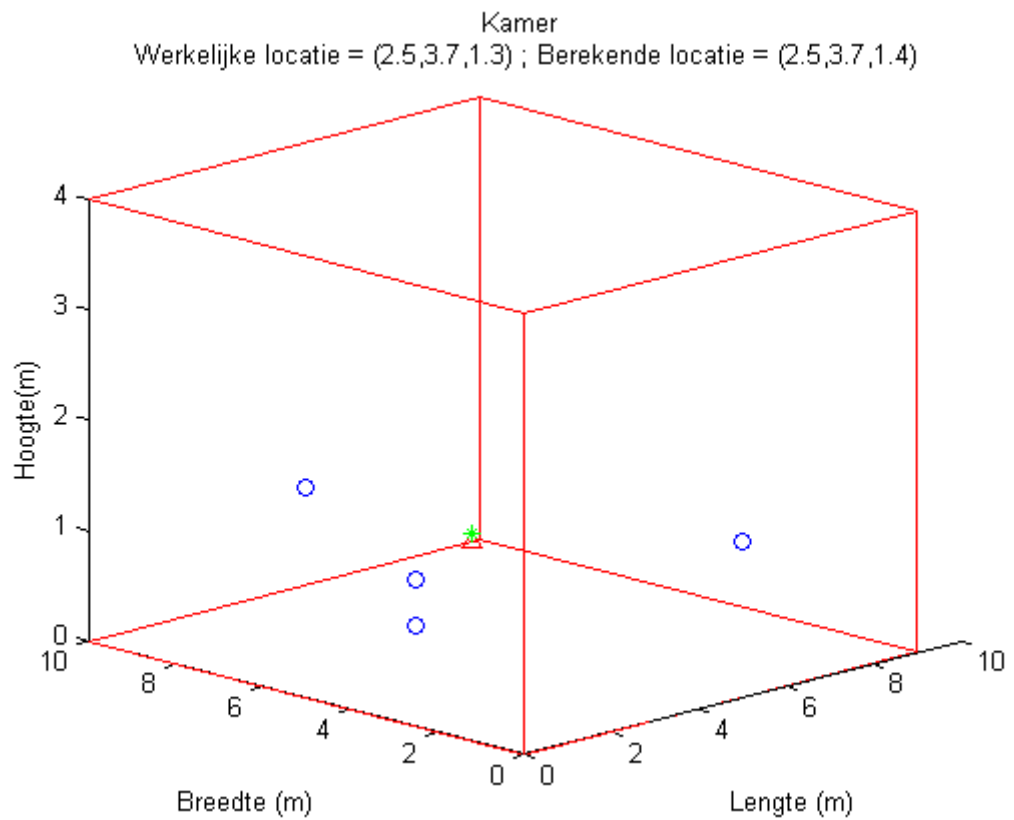


FIGUUR 3.6: Resultaat van simulatie van bronlokalisatie in 3D-ruimte (o: microfoon; Δ : werkelijke locatie; *: berekende locatie).

TABEL 3.3: Locaties van bron en microfoons in 3D-ruimte, verspreide opstelling.

	x	y	z
bron	2.5	3.7	1.3
microfoon 1	2.0	4.5	0.5
microfoon 2	8.0	3.0	0.8
microfoon 3	1.0	6.0	1.7
microfoon 4	5.0	7.5	0.3

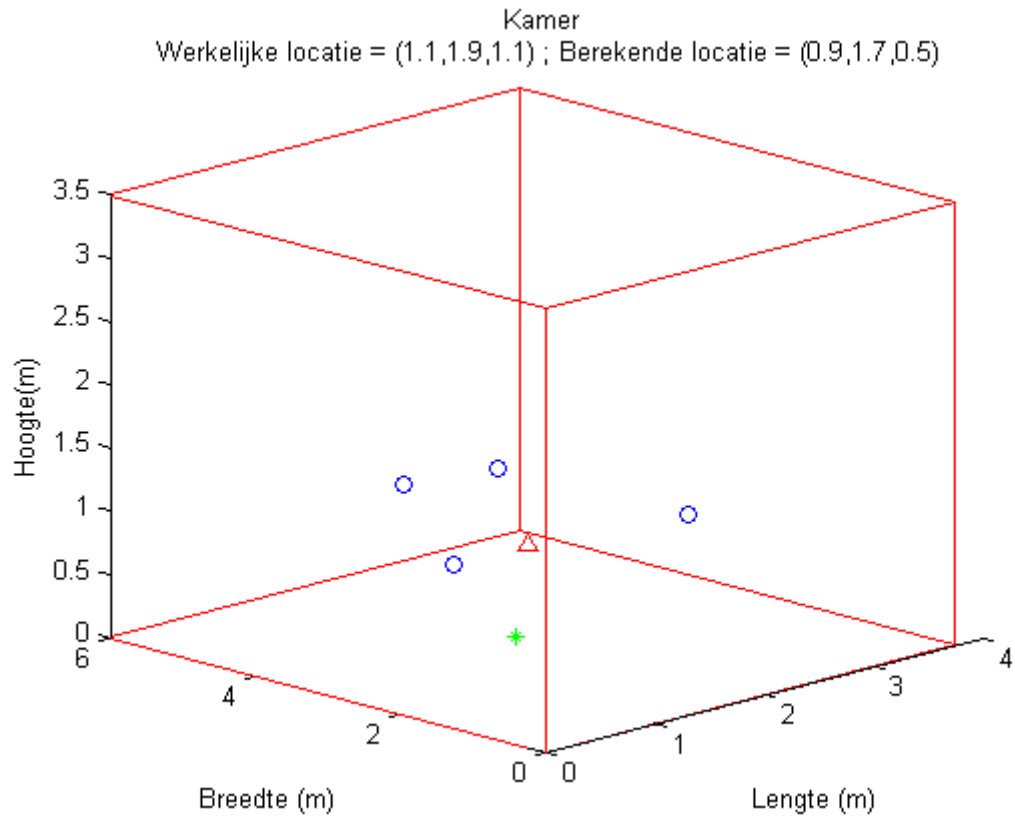
3. BRONLOCATIE BEPALEN



FIGUUR 3.7: Resultaat van een simulatie van bronlokalisatie in 3D-ruimte met verspreide opstelling (○: microfoon; Δ: werkelijke locatie; *: berekende locatie).

TABEL 3.4: Locaties van bron en microfoons in 3D-ruimte (in meter) voor niet gunstig resultaat.

	x	y	z
bron	1.1	1.9	1.1
microfoon 1	1.75	3.9	.5
microfoon 2	2.6	1.93	1
microfoon 3	0.7	3	1.5
microfoon 4	0.3	1.1	2



FIGUUR 3.8: Resultaat van een simulatie van bronlokalisatie in 3D-ruimte met minder gunstige oplossing (o: microfoon; Δ : werkelijke locatie; *: berekende locatie).

3.3 Besluit

Uit de simulaties blijkt dat de bronlokalisatie in 2D een nauwkeuriger resultaat geeft dan deze in 3D. In drie dimensies kan er een probleem optreden. De correctheid van het resultaat is afhankelijk van de gekozen bron- en microfoonposities.

Hoofdstuk 4

Reflectorlocaties bepalen

Zoals ook in de inleiding besproken, zijn er verschillende manieren om de locaties van reflectoren in een omgeving te bepalen. Globaal gezien zijn er twee categorieën: methodes die gebruik maken van bundelsturingstechnieken en microfoonroosters, en methodes waar de microfoons verspreid staan of waar men er slechts één gebruikt. Om dichter naar het concept van de Batman sonar te gaan, wordt de methode waarbij microfoons willekeurig verspreid staan doorheen de omgeving verkozen. Dit maakt de methode meer algemeen toepasbaar. Informatie die gekend is, is net zoals bij het bepalen van de bronlocatie de relatieve positie van de microfoons ten opzichte van elkaar. In eerste instantie wordt het bronsignaal ongekend verondersteld. In het voorgaande hoofdstuk werd reeds besproken hoe via de TOA's van het directe pad de positie van de bron gevonden kan worden. In de aanpak om de reflectoren te bepalen wordt de methode voorgesteld in [20] gevolgd, namelijk enkel met gebruik van de eerste orde TOA's en de reeds berekende bronpositie. In de literatuur is recent ook een manier beschreven die gebruik maakt van 1^{ste} en 2^{de} orde reflecties [9]. Een vergelijking met deze en andere methodes zal kort worden toegelicht.

4.1 2D Lokalisatie van de reflectoren

Zoals ook bij de simulaties van de bronlokalisatie, worden de berekeningen eerst uitgewerkt in 2D en daarna uitgebreid naar 3D. De eerste taak is het vinden van de TOA's horende bij de eerste orde reflecties. Zoals in hoofdstuk 2 uitgewerkt kunnen de RIR's horende bij elke microfoon worden gevonden, ofwel via blinde schatting ofwel via het gekende bronsignaal. Gegeven deze RIR's moeten hierin eerst de eerste orde reflecties geïdentificeerd worden. Deze geven een indicatie van de lengte van het akoestisch pad van bron tot micro via een reflector. Omdat de totale padlengte gekend is, en de positie van de bron en microfoon, kan een ellips bepaald worden die elke mogelijke positie van de reflector weergeeft.

In deze thesis wordt gefocust op toepassingen in een kamer. Een kamer op zich kan ook een complexe omgeving zijn, aanwezigheid van meubels, deuren, personen, niet-standaard hoeken en dergelijke kunnen invloed hebben op de impulsresponsies van de kamer, maar er wordt verondersteld dat deze klein is. De meeste kamers

bestaan uit vier loodrecht op elkaar staande vlakke muren, een vloer en plafond welke parallel zijn ten opzichte van elkaar. In 2D geeft dit als model een rechthoekige geometrie van de vier muren, met als voorwaarde dat deze loodrecht op elkaar staan, zodat schuine muren dus buiten beschouwing worden gelaten.

In de RIR's van elke microfoon kunnen telkens de eerste orde reflecties van deze vier reflectoren gezocht worden. Hierbij wordt verondersteld dat de bron voldoende omnidirectioneel is, en de microfoons een directionaliteitskarakteristiek bezitten die breed genoeg is om reflecties van alle reflectoren te kunnen opvangen. Op het praktische aspect hiervan wordt in hoofdstuk 5 nog teruggekomen. Dit resulteert dan in vier ellipsen voor de vier mogelijke reflectoren per microfoon. Zoals beschreven in [20], liggen de reflectiepunten op een gemeenschappelijke rechte per reflector, in de veronderstelling van vlakke muren. Deze lijn kan worden bepaald als de raaklijn aan de gevonden ellipsen. Wat het probleem complexer maakt, is dat enkel de TOA's gekend zijn. Uit de RIR's kan niet bepaald worden welke TOA bij welke reflector hoort. Hoe dit wordt aangepakt en de resultaten ervan worden verder in de tekst besproken.

4.1.1 Eerste orde TOA's verkrijgen uit de RIR's

De eerste taak is dus de TOA's nauwkeurig te verkrijgen vertrekkende van de RIR's. Qua opstelling van de microfoons en bron wordt dezelfde gebruikt als in het voorgaande hoofdstuk, beschreven in tabel 3.1. De figuren 3.1, 3.2, 3.3 en 3.4 geven de RIR's voor elke microfoon weer gegenereerd door het kamermodel. Hierin is duidelijk het exponentiële verloop van de reverberatie en de maximale piek van het directe pad te zien zoals beschreven in paragraaf 2.1.

In het kamermodel kan voor elke reflector de reflectiviteitscoëfficiënt als parameter meegegeven worden. In deze 2D simulatie is de vloer en het plafond van de ruimte als volledig absorberend verondersteld nl. $R = 0$, en de reflectiviteitscoëfficiënt van de vier overige wanden op $R = 0.7$ ingesteld. In de RIR figuren ziet men buiten de piek corresponderend met het directe pad, nog ongeveer vier andere piekwaardes in de vroege reflecties. Doorgaans heeft een golf die slechts één reflectie ondergaat een hogere amplitude dan een golf die twee reflecties ondergaat, omdat ook absorptie dan tweemaal optreedt. Echter is dit niet altijd zo. Er kan zich bijvoorbeeld een situatie voordoen waarin een microfoon zich dicht bij een hoek in de kamer bevindt en dicht bij de bron, maar relatief ver van de andere reflectoren. Dan kan het gebeuren dat de golf die één reflectie ondergaat, en een lang pad aflegt, meer verzwakt wordt dan een golf die tweemaal gereflecteerd wordt maar slechts een heel kort pad aflegt. Dus de posities van de microfoons en de reflectiviteitscoëfficiënt spelen hierin een grote rol. De verzwakkingsconstante σ is afhankelijk van de frequentie en van de vochtigheid van de kamer. Hoe hoger de frequentie hoe groter de verzwakking. Ter illustratie: de verzwakkingsconstante bedraagt bij een golf met een frequentie van 4 kHz ongeveer 0.1 dB/m. De uitdrukking voor σ in Neper per meter is [23]:

$$\sigma = \frac{1}{2 \cdot d} \cdot \ln \frac{I_0}{I_d}, \quad \text{met } d = \text{afstand, } I_0 \text{ en } I_d \text{ resp. de initiële en verzwakte intensiteit.}$$

In decibel per meter omgezet:

$$\sigma' = \frac{20}{d} \cdot \log \frac{V_0}{V_d} = \frac{10}{d} \cdot \log \frac{I_0}{I_d} = \frac{10}{\frac{\ln 10}{2}} \approx 8.686 \cdot \sigma.$$

De verzwakte amplitude V_d in functie van de originele amplitude V_0 wordt dan gevonden dan via:

$$\frac{V_d}{V_0} = \cdot 10^{\frac{\sigma' \cdot d}{20}}$$

Neem aan dat de padlengte van bron tot de verre reflector 6m bedraagt, van reflector tot microfoon 5m. Stel de reflectiecoëfficiënt van een wand gelijk aan $R = 0.7$. Dit geeft dan een verzwakking van: $10^{-6 \cdot 0.1/20} \cdot 0.7 \cdot 10^{-5 \cdot 0.1/20} \approx 0.6167$. Stel dat de verzwakking van het korte pad verwaarloosbaar is, en de dubbele reflectie de totale verzwakking bepaalt, dan geeft dit $0.7^2 = 0.49$. In dit geval zal dus de dubbele reflectie de meest verzwakte golf zijn. Indien de reflectiviteitscoëfficiënt hoger is, voor betonnen muren is $R = 0.9$ geen uitzonderlijke waarde bij een golf van 4 kHz [24], dan zou de golf die de dubbele reflectie ondergaat net iets minder verzwakt zijn. Voor hogere frequenties zal dit verschil groter zijn.

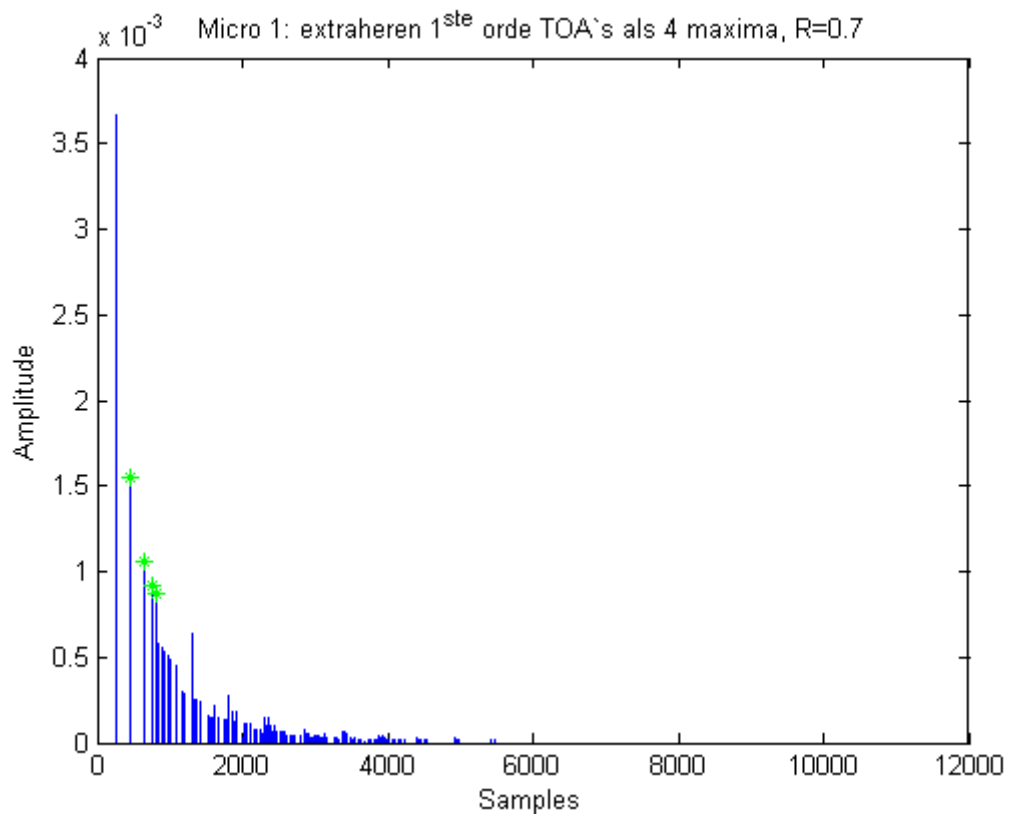
In vele gevallen zullen echter de eerste orde reflecties wel componenten met grotere amplitudes zijn dan tweede en hogere orde reflecties in de RIR's. Een mogelijke aanpak om vertrekkende van deze veronderstelling de eerste orde TOA's te bepalen, is om eerste alle maxima in een rij op te slaan, vervolgens de vier grootste eruit te selecteren op het direct pad na. Zoals in de figuren met de RIR's te zien is, bevat deze ook componenten met een negatieve amplitude. Een eerste stap is het toepassen van een positieve drempelwaarde. Als waarde voor deze drempel is 10^{-5} gekozen, zodat ruiscomponenten met zeer lage amplitudes buiten beschouwing gelaten worden. Elke piek heeft nu twee nuldoorgangen met de x-as. Het algoritme in pseudocode om nu elk maximum op te slaan ziet er als volgt uit:

```

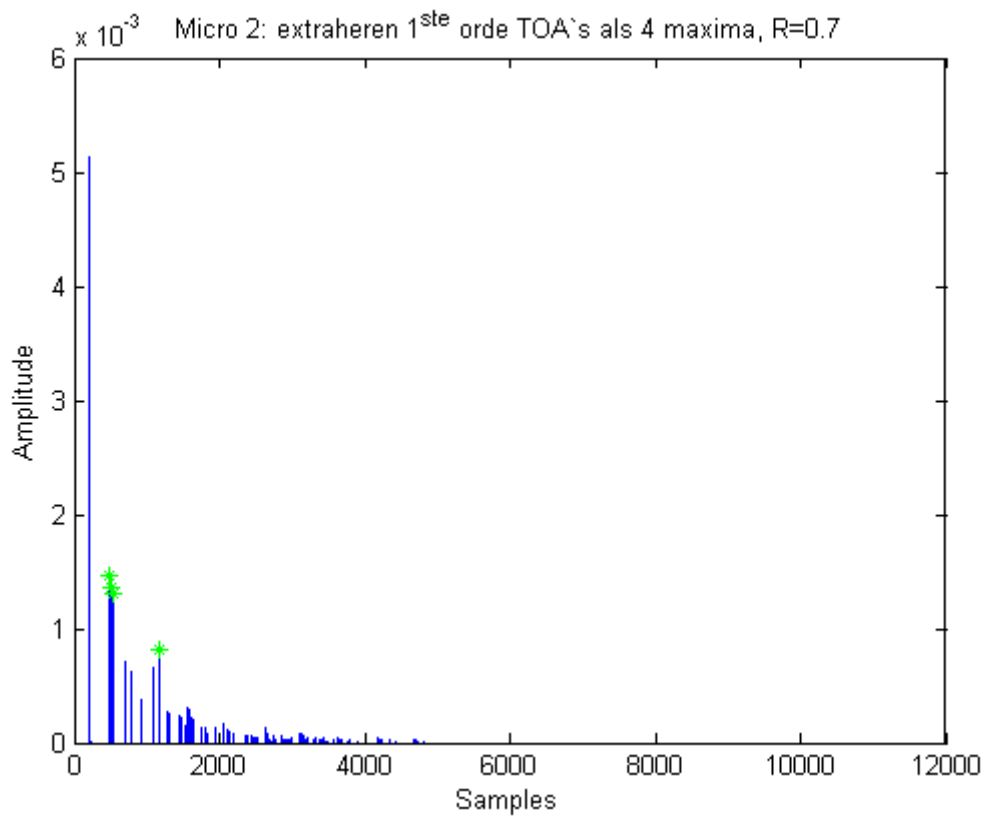
for  $i = 1 \rightarrow L$  do
  if  $RIR(i) == 0$  &&  $findStart == false$  then
     $findStart \leftarrow true$ 
  end if
  if  $RIR(i) \neq 0$  &&  $findStart == true$  then
     $startOfInterval \leftarrow i$ 
     $findStart \leftarrow false$ 
     $findEnd \leftarrow true$ 
  end if
  if  $RIR(i) == 0$  &&  $findEnd == true$  then
     $endOfInterval \leftarrow i$ 
     $findEnd \leftarrow false$ 
     $[value, index] \leftarrow \max AIR(startOfInterval \rightarrow endOfInterval)$ 
    add{ $maxArray, value$ } , add{ $positionArray, index$ }
  end if
end for

```

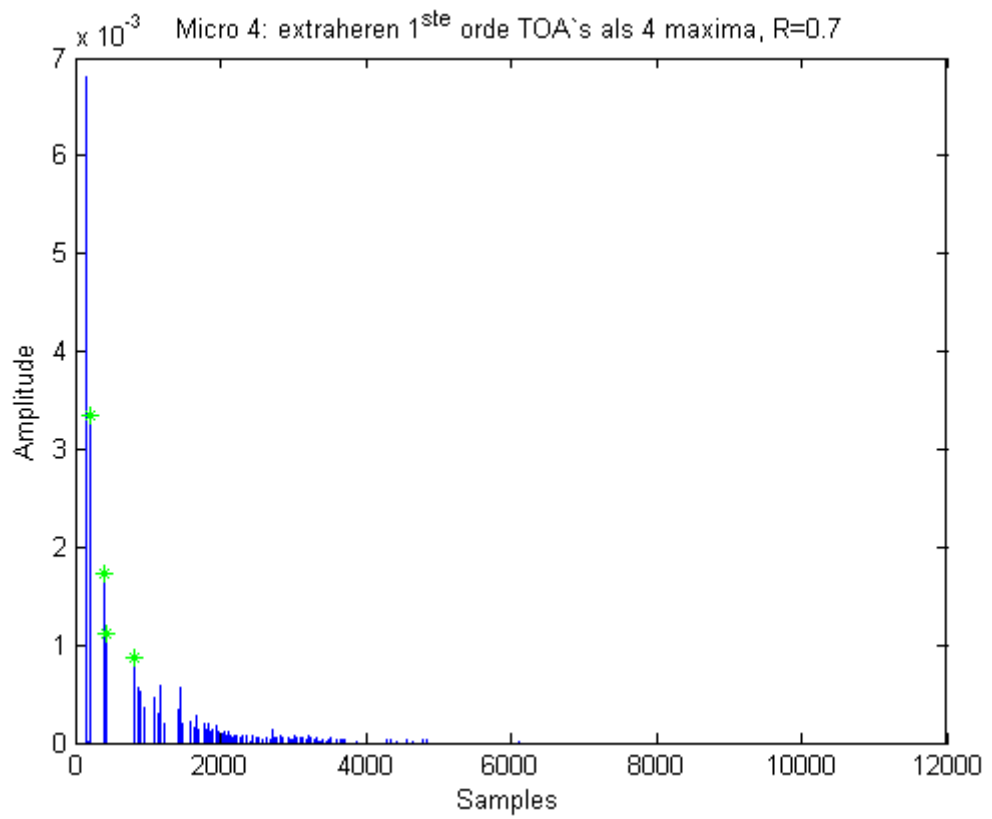
Hierbij is L de geschatte lengte waarin de eerste vier reflecties zullen optreden, het *add* commando voegt een waarde toe aan een reeks. Het resultaat van dit algoritme is een reeks met alle maximum waarden van de RIR tot een opgegeven lengte L , en een reeks met de posities ervan uitgedrukt in aantal samples, gerangschikt in volgorde van voorkomen. Aangezien het directe pad normaal gezien overeenstemt met het globale maximum van de RIR, kan de index hiervan zo worden bepaald en daarna zowel uit de positiereeks als de waardereeks worden verwijderd. Dan wordt telkens het maximum van de waardereeks gezocht, de waarde met zelfde index uit de positiereeks gehaald en bijgehouden. Daarna wordt dit maximum verwijderd uit beide reeksen. Deze procedure wordt vier keer herhaald om de vier eerste orde reflecties te vinden. Het resultaat voor drie microfoons is te zien in figuren 4.1, 4.2 en 4.3. De waarden van de TOA's staan in tabel 4.1. Ook is ter controle een algoritme opgesteld dat de TOA's zoekt gegeven de kamerafmetingen. In dit geval kan de totale padlengte eenvoudig gevonden worden door de coördinaten van de betreffende microfoon te spiegelen ten opzichte van de reflector. De TOA's gevonden op deze manier staan weergegeven in tabel 4.2.



FIGUUR 4.1: RIR microfoon 1 met eerste orde TOA's via methode 4 maxima.



FIGUUR 4.2: RIR microfoon 2 met eerste orde TOA's via methode 4 maxima.



FIGUUR 4.3: RIR microfoon 4 met eerste orde TOA's via methode 4 maxima.

TABEL 4.1: TOA's in aantal samples, via 4 maxima, $R=0.7$.

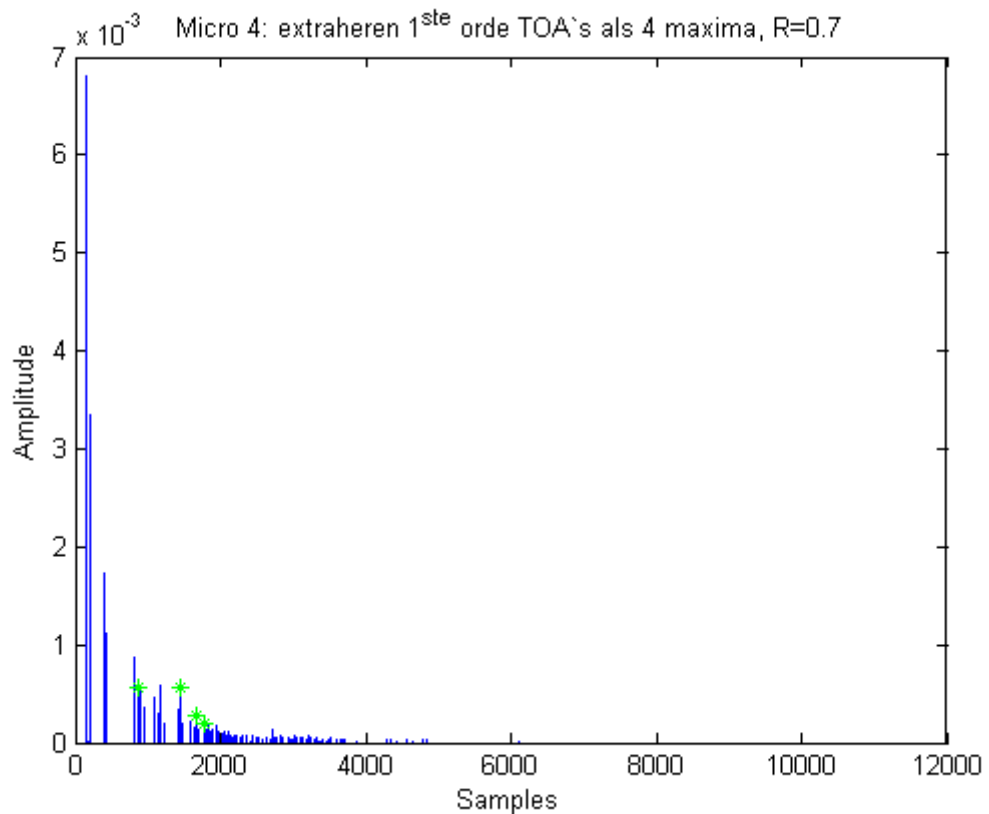
	TOA1	TOA2	TOA3	TOA4
microfoon 1	453	660	759	810
microfoon 2	481	497	535	1171
microfoon 3	275	639	757	924
microfoon 4	211	404	431	802

TABEL 4.2: TOA's in aantal samples, via spiegeling (ideale waardes).

	TOA1	TOA2	TOA3	TOA4
microfoon 1	452	659	757	809
microfoon 2	480	496	534	1077
microfoon 3	274	638	756	922
microfoon 4	209	403	801	1172

Als variant hierop is ook getracht een beperking op te leggen die stelt dat twee opeenvolgende maxima groter dan een bepaalde drempel van elkaar verwijderd moeten liggen. Dit kan gedeeltelijk het eerder beschreven probleem van een microfoon dicht in een hoek oplossen. In figuur 3.5 ligt microfoon 4 met als coördinaten (0.3;1.1) in de linkerhoek en dicht tegen linker verticale wand. Op zicht is te zien dat het pad van bron tot microfoon 4 via de onderste horizontale wand of via dubbele reflectie door onderste horizontale en linkse verticale wand weinig verschilt in lengte. In de onderste rij van tabel 4.2 zijn de verwachte waardes voor de TOA's horende bij microfoon 4 voor elke reflector te zien. Bij inspectie hoort de waarde 403 samples bij de onderste wand, 209 samples bij de rechtse, 1172 samples bij de bovenste en 801 samples tenslotte bij de linkse wand. In figuur 4.3 kan inderdaad een vrij grote piek vlak na de piek die bij de eerste orde reflectie met de onderste wand komt gezien worden. De piek op 1172 samples die de reflectie de bovenste wand aangeeft, is echter in amplitude kleiner dan de tweede orde reflectie met linker hoek, waardoor deze niet te zien is in tabel 4.1. Zoals eerder in deze paragraaf gezegd hoeft dit in de praktijk niet altijd waar te zijn. De afstand tussen de piek op 404 samples en de tweede orde er vlak na op 431 samples (tabel 4.1) is iets minder dan 30 samples. Indien hieruit een voorwaarde gekozen wordt, met name dat gezochte TOA's minstens 30 samples uit elkaar moeten liggen, dan wordt het resultaat van figuur 4.4 verkregen. Nu wordt wel de juiste TOA voor de bovenste reflector gevonden. Het probleem hier is nogal vanzelfsprekend, namelijk dat het best mogelijk is dat een padlengte minder dan 30 samples verschilt. Bijvoorbeeld in tabel 4.1 is te zien dat voor microfoon 2 tussen de eerste twee TOA's slechts een verschil is van 16 samples. Het is dus moeilijk hiervoor een goede waarde te kiezen.

Een andere methode die vertrekt van dezelfde waarde- en positiereeks, gaat de waardereeks doorlopen en telkens vergelijken met de vorige waarde. Een waarde groter als de vorige wordt aangezien als een betere kandidaat voor een eerste orde TOA. Indien deze waarde kleiner is dan de vorige wordt deze gewoon toegevoegd



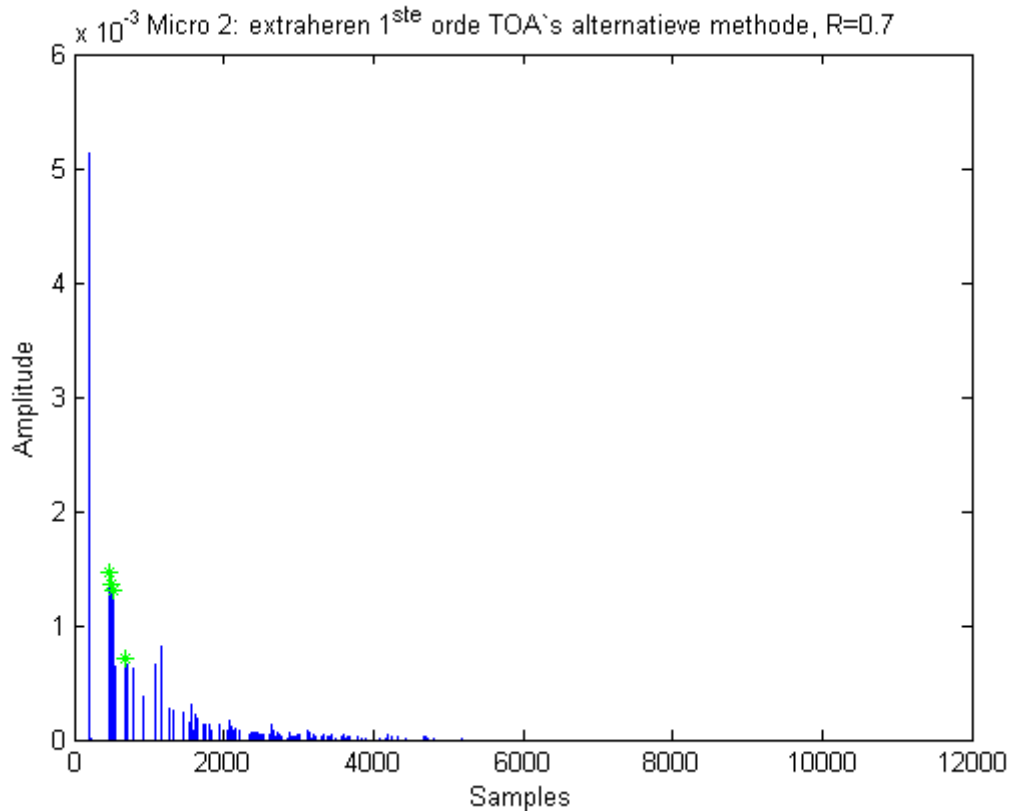
FIGUUR 4.4: RIR microfoon 4 met eerste orde TOA's via methode 4 maxima en beperking.

aan de uiteindelijke reeks. Een duidelijk nadeel van deze methode is dat, indien bijvoorbeeld vier dalende pieken na elkaar vallen, maar de gezochte grotere pieken eigenlijk pas daarna verschijnen, deze niet in de uiteindelijke reeks terecht zullen komen. Het resultaat is getekend in figuur 4.5 en weergegeven in tabel 4.3. Wanneer deze vergeleken wordt met tabel 4.1 is enkel een verschil in de tweede rij waar te nemen, de piek op 692 samples wordt hier als vierde TOA beschouwd, terwijl er nog een piek met hogere amplitude volgt.

De posities van de microfoons in de kamer hebben dus een grote invloed op de

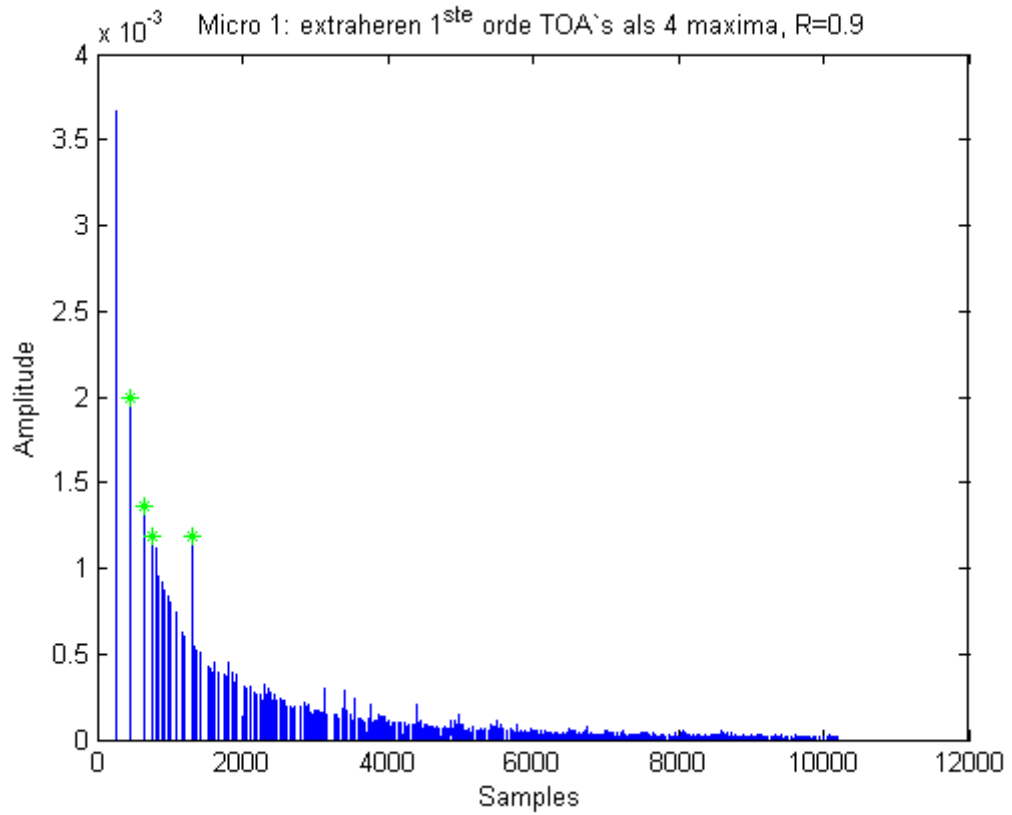
TABEL 4.3: TOA's in aantal samples, via alternatieve methode, $R=0.7$.

	TOA1	TOA2	TOA3	TOA4
microfoon 1	453	660	759	810
microfoon 2	481	497	535	692
microfoon 3	275	639	757	924
microfoon 4	211	404	431	802



FIGUUR 4.5: RIR microfoon 2 met eerste orde TOA's via alternatieve methode.

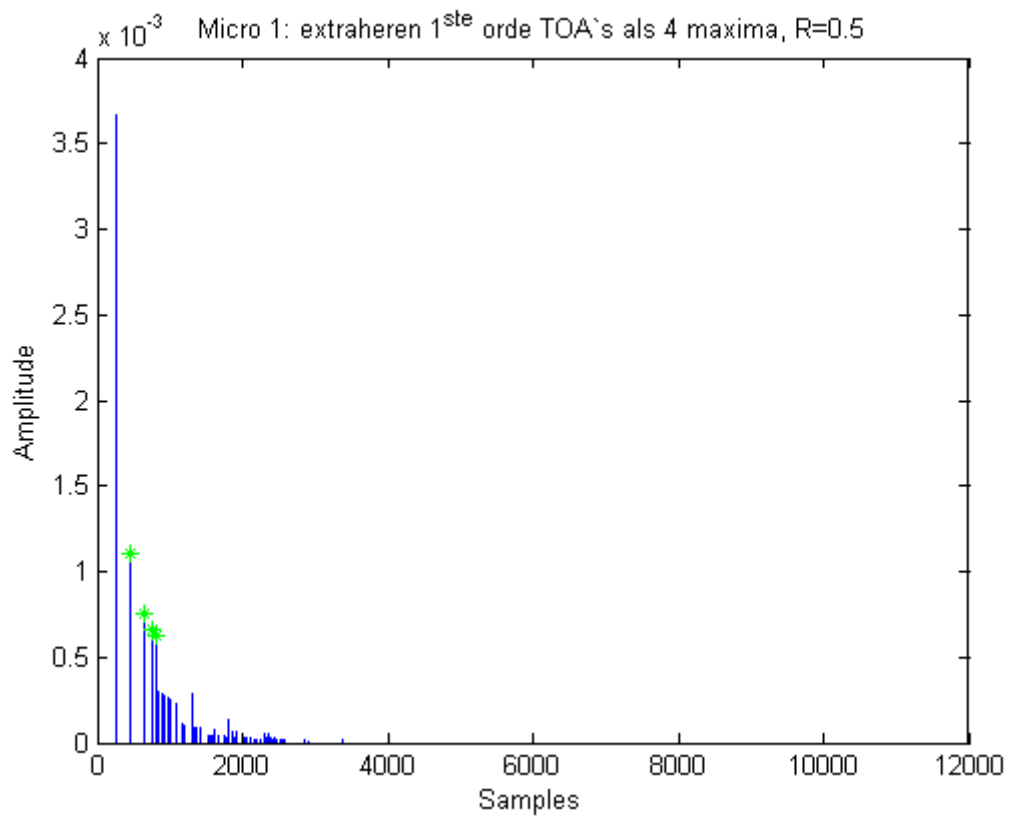
precisie waarmee de eerste orde TOA's kunnen worden geschat. Ideaal zou zijn als de microfoons ver genoeg van hoeken verwijderd staan en verspreid ten opzichte van elkaar. Ook zal bij een hogere reflectiviteitscoëfficiënt, bijvoorbeeld $R = 0.9$ of hoger, het verschil tussen tweede en eerste orde reflecties kleiner zijn dan bij lagere waarden. Vergelijkende figuren voor de RIR van microfoon 1 gevonden met een coëfficiënt $R = 0.9$ en $R = 0.5$ worden getoond in respectievelijk figuur 4.6 en 4.7, met aanduiding van de eerste orde TOA's. Bijhorende tabellen zijn tabel 4.4 en 4.5. Het valt op dat er tussen de gevonden TOA's weinig verschil is, hoewel deze de eerste orde TOA's in figuur 4.7 duidelijker zichtbaar zijn dan in figuur 4.6 en ook figuur 4.1. Uiteindelijke resultaten voor posities van reflectoren worden getoond in de volgende paragraaf, voor de verschillende gevallen.

FIGUUR 4.6: RIR micro 1 met eerste orde TOA's en $R=0.9$.TABEL 4.4: TOA's in aantal samples, via 4 maxima, $R=0.9$.

	TOA1	TOA2	TOA3	TOA4
microfoon 1	453	660	759	810
microfoon 2	481	497	535	1171
microfoon 3	275	639	679	757
microfoon 4	211	404	431	802

TABEL 4.5: TOA's in aantal samples, via 4 maxima, $R=0.5$.

	TOA1	TOA2	TOA3	TOA4
microfoon 1	453	660	759	810
microfoon 2	481	497	535	1079
microfoon 3	275	639	757	924
microfoon 4	211	404	431	802



FIGUUR 4.7: RIR micro 1 met eerste orde TOA's en $R=0.5$.

Tenslotte is te bemerken dat indien naast eerste orde reflecties ook tweede orde reflecties moet kunnen onderscheiden worden van de hogere ordes, dit het probleem om deze precies te lokaliseren nog moeilijker zal maken. In volgende paragraaf en ook in de praktische resultaten zal blijken dat de nauwkeurigheid waarmee deze TOA's bepaald kunnen worden het cruciale element is van de aanpak besproken in deze thesis, ook van gelijkaardige methodes die vertrekken van de RIR's.

4.1.2 Berekenen van de ellips parameters

In deze paragraaf wordt vertrokken van de gevonden eerste orde TOA's. Ook in de vele gerelateerde papers wordt vaak hiervan vertrokken. Meestal wordt gesteld dat men de RIR's ter beschikking heeft en de eerste orde TOA's, of zelfs de eerste en tweede orde TOA's nauwkeurig gekend zijn. Nochtans is reeds aangehaald dat het nauwkeurig schatten van de RIR's en de nodige TOA's geen eenvoudig probleem is. Het gevolg van afwijkingen op de TOA's zal daarom ook behandeld worden en in rekening worden gebracht bij de implementatie.

Een tweedimensionale kegelsnede kan in homogene coördinaten uitgedrukt worden gebruikmakende van de volgende parametrisatie:

$$\zeta = \left\{ (x, y) \in \mathbb{R}^2 \mid ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0 \right\} \quad (4.1)$$

Met

$$\mathbf{v} = \begin{bmatrix} x & y & 1 \end{bmatrix}^T \quad (4.2)$$

en

$$\mathbf{C} = \begin{bmatrix} a & b & d \\ b & c & e \\ d & e & f \end{bmatrix} \quad (4.3)$$

kan deze vergelijking eenvoudig herschreven worden als:

$$\mathbf{v}^T \mathbf{C} \mathbf{v} = 0 \quad (4.4)$$

Opdat de kegelsnede een ellips definieert moet voldaan zijn aan de volgende voorwaarden:

$$\det(\mathbf{C}) \neq 0, \begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0, \frac{\det(\mathbf{C})}{a+c} < 0 \quad (4.5)$$

De coëfficiëntenmatrix \mathbf{C}_I van de eenheidscirkel is gegeven door:

$$\mathbf{C}_I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (4.6)$$

en de vergelijking van de eenheidscirkel luidt dan

$$\mathbf{v}^T \mathbf{C}_I \mathbf{v} = 0 \quad (4.7)$$

Zoals beschreven in [20], kan via Sylvester's wet van inertie en eigenschappen van symmetrische matrices de coëfficiëntenmatrix van een ellips afgeleid worden uit deze eenheidscirkel, door achtereenvolgens deze te scaleren, roteren en translteren. In wat volgt staat i voor de beschouwde microfoon, en per microfoon staat k voor de beschouwde reflector. De gebruikte matrices voor deze transformatie worden verder in de tekst gedefinieerd. Het vertrekpunt is de eenheidscirkel met coëfficiëntenmatrix \mathbf{C}_I . De eerste stap is een schaalverandering, waarbij gesteld wordt dat de hoofdas van de ellips volgens de x-as ligt en de korte as volgens de y-as. De halve lengte van de hoofdas wordt aangeduid als $Q_{i,k}^{\text{maj}}$ en deze van de korte als als $Q_{i,k}^{\text{min}}$. Dit geeft als schalingsmatrix $\mathbf{S}_{i,k}$:

$$\mathbf{S}_{i,k} = \begin{bmatrix} Q_{i,k}^{\text{maj}} & 0 & 0 \\ 0 & Q_{i,k}^{\text{min}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.8)$$

De matrix \mathbf{R}_i is een rotatiematrix. De ellips wordt nu verdraaid over een hoek γ_i rond de positieve z-as, of dus in tegenuurwijzerzin.

$$\mathbf{R}_i = \begin{bmatrix} \cos \gamma_i & -\sin \gamma_i & 0 \\ \sin \gamma_i & \cos \gamma_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.9)$$

De matrix \mathbf{T}_i tenslotte is een translatiematrix, welke ervoor zorgt dat het centrum op $(0,0)$ verschoven wordt naar het punt $(\Delta x_i, \Delta y_i)$. Deze wordt gegeven door:

$$\mathbf{T}_i = \begin{bmatrix} 1 & 0 & \Delta x_i \\ 0 & 1 & \Delta y_i \\ 0 & 0 & 1 \end{bmatrix} \quad (4.10)$$

Wat gewenst is, is een ellips die de meetkundige plaats van mogelijke reflectiepunten beschrijft. Dit is zo indien de twee brandpunten van de ellips liggen op respectievelijk de positie van de bron en de positie van de beschouwde microfoon. Dit houdt in dat het centrum van de ellips bepaald kan worden als het midden van het lijnstuk tussen de bronpositie (x_s, y_s) en de microfoonpositie (x_i, y_i) . De hoek γ_i wordt gegeven door:

$$\gamma_i = \arctan \frac{y_s - y_i}{x_s - x_i} \quad (4.11)$$

waardoor de parameters Δx_i en Δy_i gevonden worden als:

$$\Delta x_i = x_i + \frac{D_i \cos \gamma_i}{2}, \Delta y_i = y_i + \frac{D_i \sin \gamma_i}{2} \quad (4.12)$$

De som van de afstanden tot beide brandpunten kan worden gevonden via de geschatte TOA's, waarbij $\tau_{i,k}$ staat voor de TOA horende bij microfoon i en reflector k uitgedrukt in seconden nl. $\tau_{i,k} = TOA_{i,k} \cdot F_s$. Hieruit volgen volgende uitdrukkingen voor de lengtes van de semi-hoofdas en semi-bijwas met η de geluidssnelheid in m/s en D_i de totale afstand tussen bron en microfoon in meter:

$$Q_{i,k}^{\text{maj}} = \frac{\eta \cdot \tau_{i,k}}{2}, Q_{i,k}^{\text{min}} = \frac{\sqrt{(\eta \cdot \tau_{i,k})^2 - D_i^2}}{2} \quad (4.13)$$

De volledige coëfficiëntenmatrix van de ellips kan dan gevonden worden als:

$$\mathbf{C}_{i,k} = \mathbf{T}_i^{-T} \mathbf{R}_i^{-T} \mathbf{S}_{i,k}^{-T} \mathbf{C}_1 \mathbf{S}_{i,k}^{-1} \mathbf{R}_i^{-1} \mathbf{T}_i^{-1} \quad (4.14)$$

Dit leidt tot een totaal van $M \times N$ ellipsen, met M het aantal microfoons en N het aantal reflectoren. De reflectoren zelf worden nu gezocht als een raaklijn aan telkens M ellipsen.

4.1.3 Lokaliseren van de reflectoren als raaklijnen aan de ellipsen.

In 2D wordt een rechte als volgt gedefinieerd.

$$\lambda = \left\{ (x, y) \in \mathbb{R}^2 \mid l_1 x + l_2 y + l_3 = 0 \right\} \quad (4.15)$$

Met

$$\mathbf{l} = \begin{bmatrix} l_1 & l_2 & l_3 \end{bmatrix}^T \quad (4.16)$$

geeft dit in vectornotatie

$$\mathbf{l}^T \mathbf{v} = 0 \quad (4.17)$$

Zoals beschreven in [25] is een rechte rakend aan een ellips met parametrisatiematrix \mathbf{C} indien:

$$\mathbf{l}^T \cdot \text{adj}(\mathbf{C}) \cdot \mathbf{l} = 0 \quad \text{met} \quad \text{adj}(\mathbf{C}) = \det(\mathbf{C}) \cdot \mathbf{C}^{-1} \quad (4.18)$$

Een mogelijke kostenfunctie die gebruikt kan worden om de drie onbekende parameters van \mathbf{l} te schatten wordt voorgesteld in [20]:

$$J_e(\mathbf{l}, \{\mathbf{C}_{i,k}\}_{i=1}^M) = \sum_{i=1}^M \left\| \mathbf{l}^T \cdot \text{adj}(\mathbf{C}_{i,k}) \cdot \mathbf{l} \right\|^2 \quad (4.19)$$

Merk ook op dat de waardes voor parameters l_1 en l_2 op een cirkel met straal $r = 1$ liggen, zodat deze de richtingscoëfficiënt van λ bepalen, terwijl l_3 dan de loodrechte afstand van λ tot de oorsprong weergeeft. Met α de hoek die λ maakt met de y -as, kunnen l_2 en l_3 ook geschreven worden als:

$$l_1 = \cos \alpha, l_2 = \sin \alpha \quad (4.20)$$

Indien we op zoek zijn naar wanden die parallel liggen met het assenstelsel waarin de microfoon posities zijn bepaald, kan α enkel 0 , $\frac{\pi}{2}$, π of $\frac{3\pi}{2}$ zijn. Dit betekent dus dat de oriëntatie van de reflectoren gekend gesteld wordt. Vier vergelijkingen zijn dan mogelijk voor λ :

$$\lambda = \begin{cases} x + l_3 = 0 \\ y + l_3 = 0 \\ -x + l_3 = 0 \\ -y + l_3 = 0 \end{cases} \quad (4.21)$$

Zoals eerder vermeld gaat in onze aanpak telkens de raaklijn gezocht worden die het best past aan M ellipsen. Indien opnieuw de opstelling zoals in 3.1 gebruikt

wordt met $M = 4$ microfoons, en $N = 4$ reflectoren, wordt een totaal van 16 ellipsen verkregen. In [20] wordt gepostuleerd dat om de ellipsen te groeperen, er $\binom{NM}{M} = \frac{(N \cdot M)!}{M!(N \cdot M - M)!}$ mogelijkheden zijn. In dit geval geeft dit 1820 verschillende combinaties. In de implementatie in deze thesis besproken hebben we dit aantal mogelijke combinaties echter gereduceerd. Indien verondersteld wordt dat elke gevonden TOA bij één microfoon daadwerkelijk bij een verschillende reflector hoort, en dus abstractie gemaakt wordt van de gevallen waarbij de geometrie zodanig is dat paden langs twee verschillende reflectoren exact even lang zijn, is het niet nodig om ook ellipsen per microfoon onderling te combineren. Het is immers de bedoeling van elke microfoon de TOA en bijhorende ellips te zoeken die hoort bij eenzelfde reflector, welke dan gevonden wordt als raaklijn aan deze ellipsen. Per microfoon is er dus de keuze uit N ellipsen en slechts één hiervan raakt aan de beschouwde reflector. Dit geeft een totaal van N^M mogelijke combinaties, in het voorbeeld dus slechts 256 mogelijkheden in tegenstelling tot 1820.

De manier waarop de parameters van de reflector gevonden worden, is door de α en l_3 te zoeken die uitdrukking (4.19) minimaliseert. Een mogelijke oplossing voor dit probleem is gebruik te maken van ingebouwde functies in Matlab om het minimum van een functie te zoeken, zoals *fmincon*. Meestal moet hiervoor een startwaarde meegegeven worden. Een probleem dat hierbij optreedt is dat vaak een lokaal minimum wordt gevonden rond deze startwaarde in plaats van het globale minimum, omdat de kostenfunctie niet convex is. Een tweede oplossing bestaat eruit om dit minimum zoeken zelf te implementeren door exhaustief te zoeken. Dit kan gebeuren op een manier, zoals in volgende pseudocode:

```

for  $k = 1 \rightarrow N$  do
  for  $c = 1 \rightarrow (N - (k - 1))^M$  do
    for  $\alpha = [0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}]$  do
      for  $l_3 = 1 \rightarrow 10$  do
         $\mathbf{l} \leftarrow [\cos \alpha, \sin \alpha, l_3]$ 
        add{ $J_e, J_e(\mathbf{l}, \{\mathbf{C}_{i,k}\}_{i=1}^M$  voor ellips combinatie  $c$ )}
      end for
    end for
     $[value, index] \leftarrow \min J_e$ 
    add{ $J_e$  per combinatie,  $J_e$ }
    add{ $\alpha$  per combinatie,  $\alpha$ } met  $\alpha$  degene die  $J_e$  minimaliseert
    add{ $l_3$  per combinatie,  $l_3$ } met  $l_3$  degene die  $J_e$  minimaliseert
  end for
   $\alpha(k) \leftarrow \alpha$  die  $J_e$  per combinatie minimaliseert
   $l_3(k) \leftarrow l_3$  die  $J_e$  per combinatie minimaliseert
  Vewijder ellipsen gebruikt in combinatie  $c$  uit mogelijke ellipscombinaties
end for

```

Hierbij proberen we eerst alle mogelijke combinaties van ellipsen. Daarna wordt een bepaald bereik van mogelijke α en l_3 waardes overlopen. Als nauwkeurigheid voor de l_3 waardes wordt bijvoorbeeld een discretisatiestap van 0.1 meter gekozen. Indien het bereik van l_3 van 0 tot 10 meter gekozen wordt, en α om het eenvoudig te houden enkel

$0, \frac{\pi}{2}, \pi$ of $\frac{3\pi}{2}$ kan zijn, geeft dit nog steeds 404 combinaties van $[\alpha, l_3]$. Voor elk hiervan moet (4.19) worden uitgerekend, achteraf worden de parameters waarvoor de J_e waarde minimaal is bijgehouden. Dit wordt gedaan voor elke combinatie van ellipsen, uiteindelijk wordt dan hieruit ook de ellipscombinatie gezocht waarvoor J_e minimaal is. Aangezien de J_e waarde per combinatie reeds de minimale is voor elke mogelijke $[\alpha, l_3]$ combinatie, wordt dus de combinatie van ellipsen waarvoor het best een raaklijn te vinden is gevonden, met de parameters van deze raaklijn. Dit toont ook aan dat het beperken van het aantal ellipscombinaties van groot belang is voor de snelheid van het algoritme. In dit geval moeten namelijk $404 \cdot 256 = 103424$ berekeningen van J_e worden gedaan, in vergelijking met $404 \cdot 1820 = 735280$ berekeningen per reflector in vergelijking met [20].

Ook dient opgemerkt te worden dat indien op deze manier de parameters van de eerste reflector gevonden worden, de ellipsen die beschouwd werden om de best passende raaklijn te zoeken, nu niet meer beschouwd te moeten worden bij de zoektocht naar de volgende reflectoren. Immers indien de juiste ellipsen gebruikt zijn, horen deze bij de reflector die reeds bepaald is en niet bij degene die nog gezocht moeten worden. Op deze manier verkleint de zoekruimte voor de tweede reflector naar $(N - 1)^M$ mogelijkheden, in het voorbeeld nog 81 mogelijkheden. Dezelfde redenering gaat op voor de volgende reflectoren, zodat voor de laatste reflector slechts M ellipsen meer overblijven.

Het algoritme is getest voor verschillende microfoonopstellingen en kamerafmetingen. Onze eerste vaststelling hieruit is dat kleine afwijkingen op de TOA's reeds een grote invloed kunnen hebben op de correctheid waarmee een reflector als raaklijn wordt gevonden. Om dit probleem te verlichten, hebben we een wijziging aangebracht in de manier waarop J_e berekend wordt. In formule (4.19) zou $\mathbf{1}^T \cdot \text{adj}(\mathbf{C}) \cdot \mathbf{1}$ exact gelijk zijn aan nul indien elke ellips perfect raakt aan de beschouwde reflector. Het minimaliseren van de sommatie van deze uitdrukking voor de M ellipsen is dan een logische keuze als kostenfunctie. Een probleem stelt zich echter bijvoorbeeld als de beschouwde reflector wel rakend is aan $M - 1$ ellipsen, maar door een fout op de geschatte TOA, er voor één microfoon geen ellips gevonden wordt die goed raakt aan de reflector. Eén van de N mogelijke ellipsen voor deze microfoon zal een minimale J_e waarde geven. De J_e -waarde wordt hierdoor vergroot, met een bijdrage afhankelijk van de fout op de TOA. Dit kan twee gevolgen hebben, namelijk dat de ellips die dan toch gekozen wordt hoort bij een andere reflector, en deze dus te vroeg verwijderd wordt uit de mogelijke ellipsen waaruit de combinaties gekozen worden. Anderzijds is het waarschijnlijk dat er nu andere reflector parameters gevonden worden die een betere raaklijn als resultaat geven dan de parameters die de gezochte reflector beschrijven. Een voorgestelde oplossing hiervoor is om niet te kijken naar de som van de J_e waardes voor de M microfoons, maar waardes apart uit te rekenen, en enkel de twee minimale J_e waardes op te tellen:

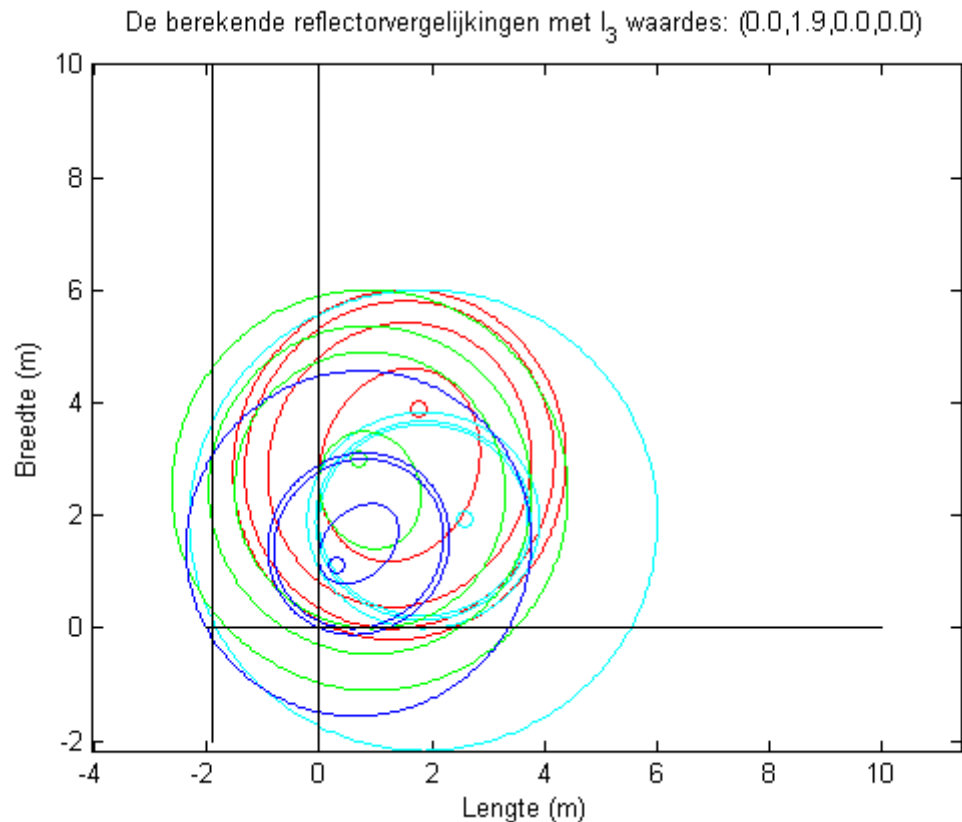
$$J_{e,i}(\mathbf{1}, \mathbf{C}_{i,k}) = \left\| \mathbf{1}^T \cdot \text{adj}(\mathbf{C}_{i,k}) \cdot \mathbf{1} \right\|^2 \quad \text{met } i = 1 \dots M \quad (4.22)$$

$$J_{e,\text{all}} = \text{sort}([J_{e,1} J_{e,2} \dots J_{e,M}]) \quad J_e(\mathbf{1}, \{\mathbf{C}_{i,k}\}_{i=1}^M) = \sum_{i=1}^2 J_{e,\text{all}}(i)$$

Indien dan twee van de M TOA's per reflector juist geschat zijn, zullen nog steeds de juiste parameters als resultaat naar voren komen. Op die manier wordt het algoritme robuuster tegen afwijkingen op de TOA's.

Een tweede aanpassing is ervoor te zorgen dat niet tweemaal dezelfde reflector als resultaat kan worden gevonden. Dit kan gebeuren indien enkele ellipsen weinig verschillen van elkaar, en daardoor voor twee groepen van gecombineerde ellipsen de best passende raaklijn dezelfde is. Dit is ook afhankelijk van de discretisatiestap voor de parameter l_3 , of dus de nauwkeurigheid waarmee de positie bepaald wordt. Een kleinere discretisatiestap, bijvoorbeeld 0.01 m zal een beter resultaat geven maar verlengt ook sterk de rekentijd van het algoritme. Ter vergelijking, in figuur 4.8 staat het resultaat met discretisatiestap voor l_3 gelijk aan 0.1 m en in figuur 4.9 met discretisatiestap 0.01 m. De kleur in deze figuur duidt aan bij welke microfoon de ellipsen horen. De J_e berekening is in beide uitgevoerd zoals in (4.19), dus zonder bovenstaande aanpassingen. Men ziet in figuur 4.9 dat er geen samenvallende reflectoren meer gevonden worden, en het resultaat beter is. De reflectoren coïnciderend met x-as en y-as worden quasi perfect gelokaliseerd, de tweede verticale wand met een afwijking van 1.32 meter en de horizontale met een afwijking van 0.75 meter. Indien het geval bekeken wordt waarbij de reflectoren parallel met het gekozen assenstelsel zijn, kan dit vrij eenvoudig opgelost worden. Er moet dan enkel gezorgd worden dat indien de gevonden α leidt tot een reflector die parallel is met één van de reeds gevonden reflectoren, en de parameter l_3 zodanig is dat beide coïncideren, de J_e waarde op oneindig groot wordt geforceerd. Bij het zoeken van het minimum over alle parametercombinaties naar degene met minimum J_e heeft deze dan geen invloed meer. Concreet, indien er naar (4.21) gekeken wordt, geven de eerste en de derde vergelijking de verticale wanden en de overige twee de horizontale. Hieruit is af te leiden dat indien de gevonden α een som van π geeft met een van reeds eerder gevonden α 's, het over een rechte parallel met de vorige verticale reflector gaat. Gelijkaardig, een som van 2π duidt op de tweede horizontale reflector. De gevonden rechten zullen coïncideren indien de parameter l_3 voor beide gelijk is in absolute waarde en verschillend van teken. Gelijke waarde en teken zou overeenkomen met twee wanden gespiegeld ten opzichte van de x-as of y-as, wat mogelijk is aangezien er geen restrictie is op de positie van de oorsprong van het assenstelsel in de kamer. In de simulatie is het bereik van mogelijke waarden voor l_3 gekozen van 0 m tot 10 m. Zoals volgt uit (4.21), kunnen wanden links van de $y - as$ of onder de $x - as$ zo nog steeds gevonden worden. Het gevolg is dat het probleem enkel nog kan optreden indien $l_3 = 0$. In de simulaties is de oorsprong van het assenstelsel gekozen als de linker onderhoek van de kamer, dus één wand coïnciderend met de x-as en één met de y-as, beide met $l_3 = 0$.

Indien het bereik van mogelijke α 's van 0 tot 2π gekozen wordt, is de kans dat tweemaal exact dezelfde raaklijn voor verschillende groepen van ellipsen gevonden wordt kleiner. Dezelfde oplossing kan gebruikt worden, maar de kans dat de som van α 's precies π of 2π is, is lager. Om te zorgen dat alleen een rechthoekige geometrie gevonden wordt, kan indien er één reflector bepaald is, het bereik van α vernauwd worden. Er moet namelijk telkens een verschil van $\frac{\pi}{2}$ zijn, eventueel met een bepaalde marge. Het probleem dat hier dan nog kan optreden, is dat als de

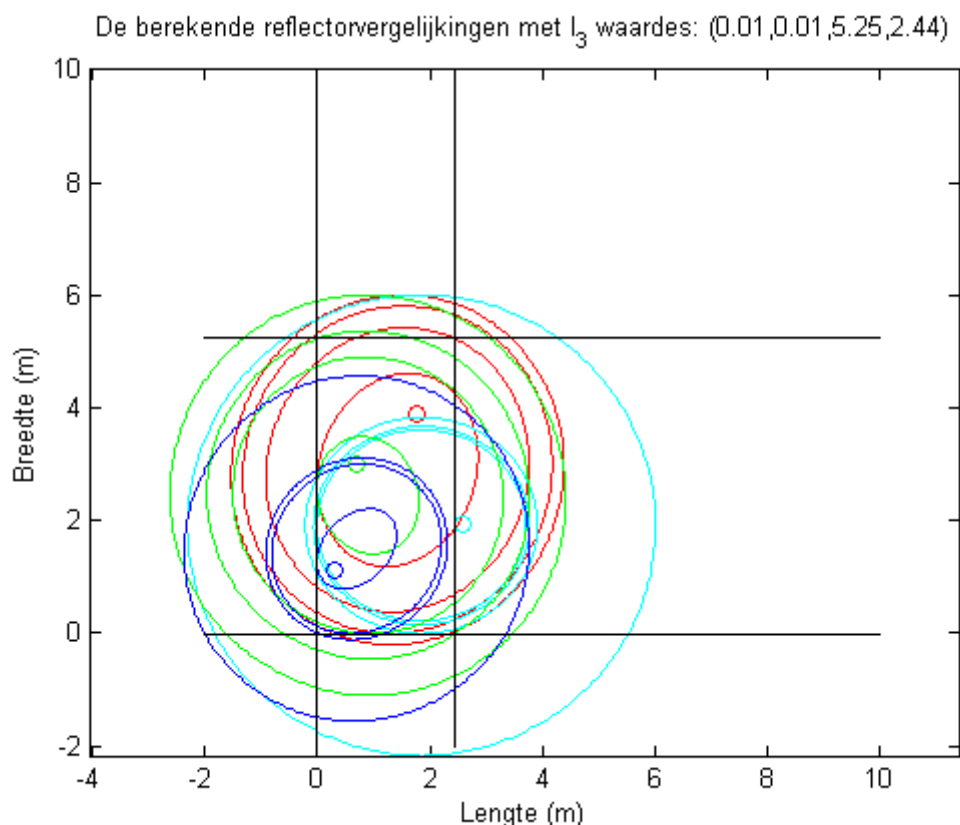


FIGUUR 4.8: Resultaat 4 reflectoren, J_e via (4.19), $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $R=0.5$.

eerste parametrisatie die gevonden wordt een fout vertoond groter dan de voorziene marge, de restrictie ertoe leidt dat de volgende reflectoren ook fout gelokaliseerd zullen worden.

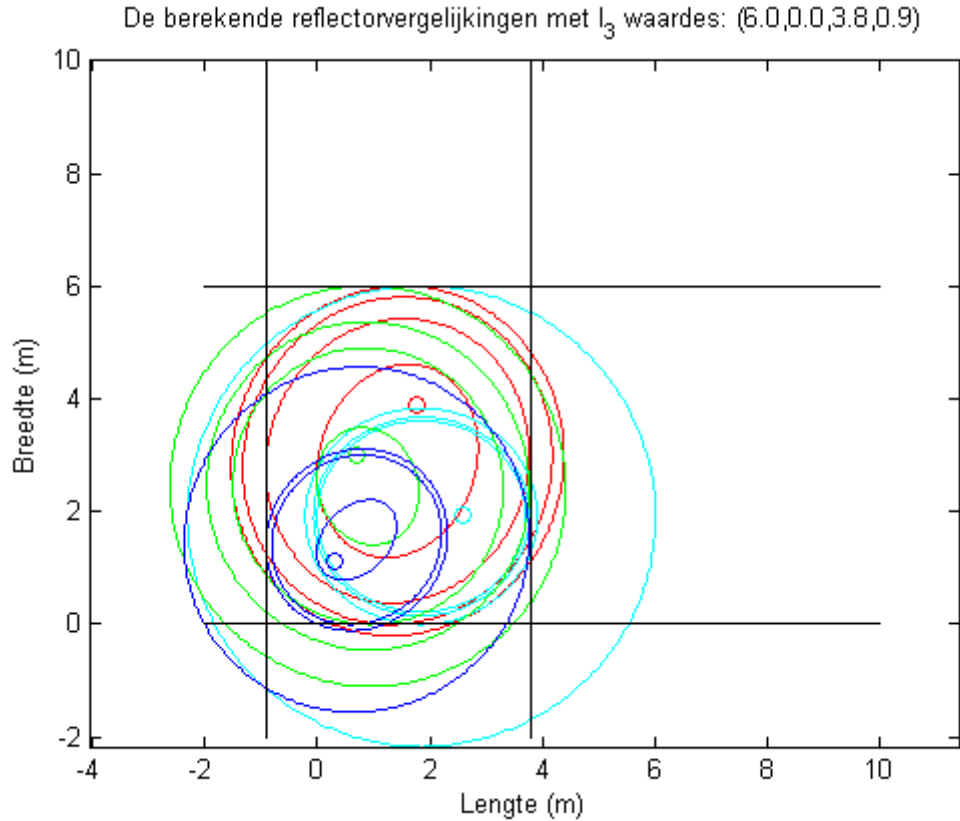
4.1.4 Resultaten 2D simulatie

In vorige sectie werd reeds in figuren 4.8 en 4.9 het resultaat getoond voor de originele J_e berekening met kamerafmetingen gelijk aan $L = 3.76$ en $B = 6$. Als reflectiecoëfficiënt is $R = 0.5$ gekozen. De TOA's worden gevonden uit de RIR's via de methode vier maxima te zoeken, zonder beperking van minimale afstand tussen de TOA's. Deze werden reeds gegeven in tabel 4.5. In deze tabel is te zien dat de enige grote afwijkingen ten opzichte van de ideale TOA's voorkomen bij microfoon 4, via vergelijking met tabel 4.2. De ellipsen horende bij microfoon 1, 2, 3 en 4 zijn respectievelijk weergegeven in het rood, cyaan, groen en blauw. Wanneer dezelfde TOA's gebruikt worden, maar formule (4.22) gebruikt wordt om de reflectoren te bepalen, geeft figuur 4.10 het resultaat. De discretisatiestap voor de parameter l_3 is hier 0.1 m gekozen, en het resultaat is zichtbaar verbeterd, en zelfs beter als het



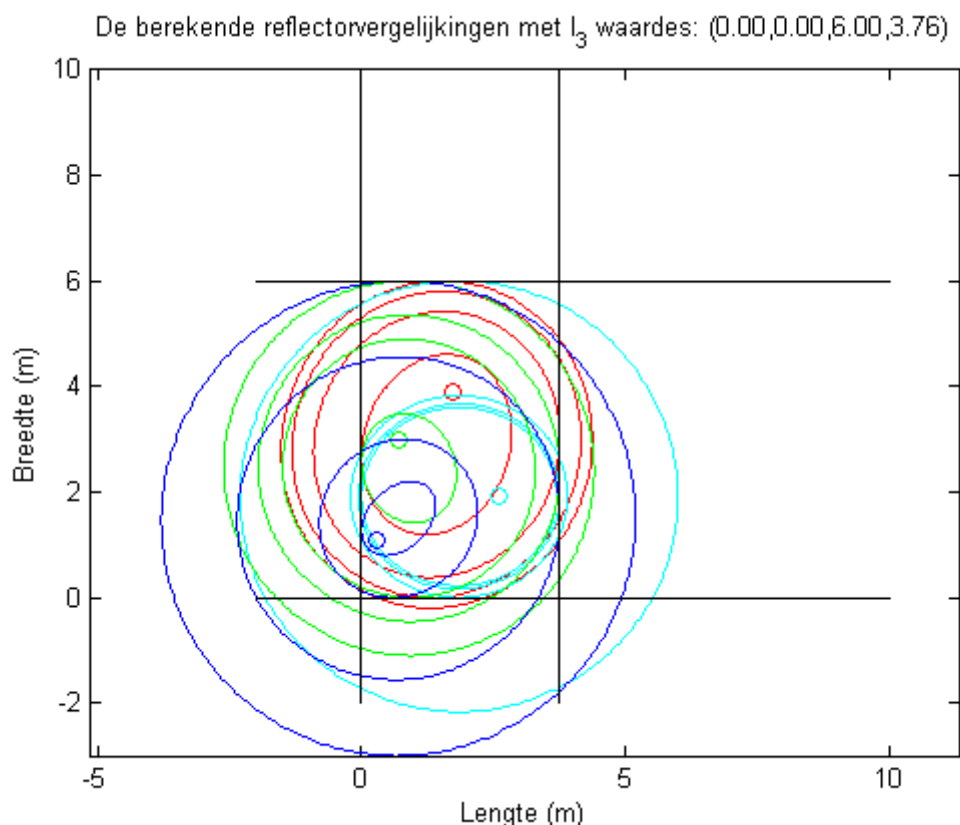
FIGUUR 4.9: Resultaat 4 reflectoren, Je via (4.19), $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m, $R=0.5$.

resultaat via (4.19) met discretisatiestap 0.01 m. De twee horizontale reflectoren worden quasi perfect gelokaliseerd, de rechter verticale wand wijkt slechts 0.04 meter af van de reële locatie wegens de afronding. De linkerwand vertoont een afwijking van 0.9 meter. De afwijking op de linkerwand kan verklaard worden via de fout op de TOA's. Indien namelijk tabel 4.5 en 4.2 vergeleken worden, is het de TOA van 431 samples op rij 4 en kolom 3 in tabel 4.5 die zorgt voor de fout, welke verklaard werd in vorige sectie. Deze genereert de derde grootste blauwe ellips (microfoon 4), waaraan een raaklijn kan gevonden worden die ook goed raakt aan de tweede grootste rode ellips (microfoon 1). Deze leidt tot de verticale reflector op $x = -0.9$. In figuur 4.11 is het resultaat te zien indien de exacte TOA's van tabel 4.2 worden gebruikt. De ellips die de raaklijn met $x = -0.9$ vormt in vorige figuur, is hier afwezig, en vervangen door een ellips die raakt aan de bovenste wand op $y = 6$. De discretisatiestap voor de parameter l_3 is ingesteld op 0.01 m, een perfecte reconstructie is het resultaat. Merk op dat aangezien de TOA's nu correct zijn, ook formule (4.19) gebruikt kan worden voor de berekening van Je . De aanpassing naar robuustheid tegen fouten op TOA's is nu niet nodig. Het resultaat is precies hetzelfde als figuur 4.11. Indien de TOA's uit tabel 4.5 gebruikt worden en formule (4.22) voor Je , met discretisatiestap voor



FIGUUR 4.10: Resultaat 4 reflectoren, J_e via (4.22), $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $R=0.5$.

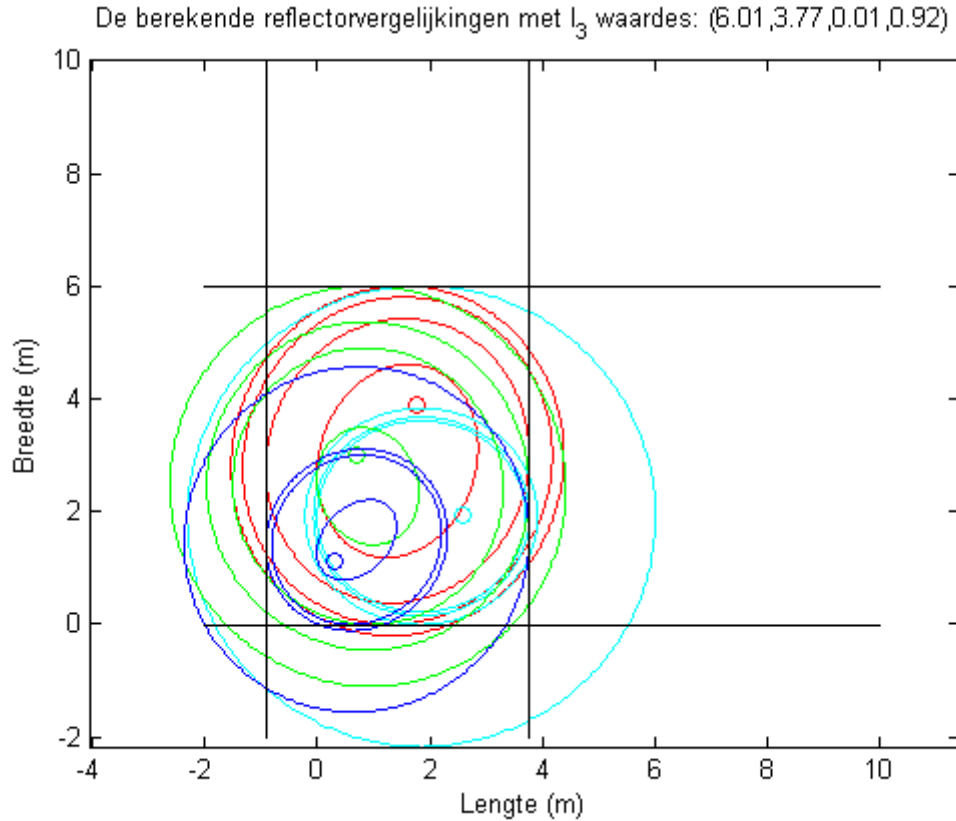
l_3 gelijk aan 0.01 m, wordt het resultaat van figuur 4.12 verkregen. Het resultaat verschilt weinig van figuur 4.10. Als besluit kan gezegd worden dat het resultaat met grotere discretisatiestap, waardoor het algoritme veel sneller uitgevoerd kan worden, praktisch evengoed is als met lagere discretisatiestap dankzij de aanpassingen in de formule voor J_e . Indien de reflectiecoëfficiënt $R = 0.9$ gekozen wordt, wordt het resultaat van figuur 4.13 verkregen. Ook hier is er weinig verschil met figuur 4.10, ondanks de significante verschillen tussen de twee laatste TOA's van rij drie in tabel 4.4 en 4.5. De TOA's in tabel 4.1 voor $R = 0.7$ zijn zo goed als gelijk aan deze in tabel 4.5 en hetzelfde kan gezegd worden over het resultaat voor de reflectorlokalisatie. Ook indien $R = 0.7$ en de alternatieve methode om de TOA's uit de RIR's te extraheren gebruikt wordt, zoals in tabel 4.3, worden dezelfde vergelijkingen gevonden voor de reflectoren. Figuur 4.14 toont het resultaat met $R = 0.7$ via TOA's door 4 maxima en figuur 4.15 voor de alternatieve methode. Ook werd in paragraaf 2.3.1 uitgewerkt dat het mogelijk is de TOA's te schatten via autocorrelatie van de microfoonsignalen. Het resultaat (figuur 4.16) is iets minder nauwkeurig als de methode van 4 maxima zoeken uit de RIR's, maar niet onaanvaardbaar. Het nadeel hier is echter dat de autocorrelatie afhankelijk blijkt te zijn van het gekozen bronsignaal. De gevonden



FIGUUR 4.11: Resultaat 4 reflectoren ideaal, J_e via (4.22), $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m.

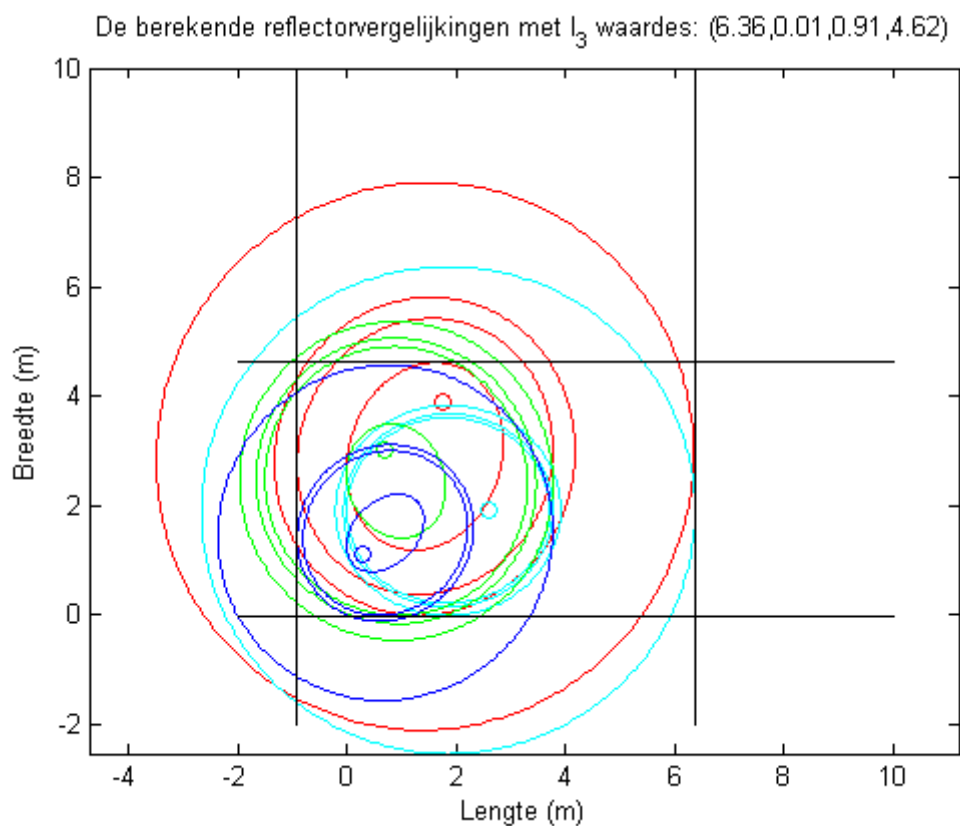
TOA's verschillen indien men een ander bronsignaal gebruikt, waardoor grotere fouten bij de reconstructie optreden.

Vervolgens wordt ook het resultaat getoond indien de kamerafmetingen en de microfoonopstelling veranderd worden. De afmetingen van de kamer zijn nu $L = 3.10$ m (volgens de x-as) en $B = 4.30$ m (volgens de y-as). Tabel 4.7 geeft de microfoonposities aan. De TOA's worden via methode vier maxima gevonden, een vergelijking met de ideale is weergegeven in tabel 4.8. Hierin is te zien dat de TOA's zonder grote afwijkingen teruggevonden worden. Figuur 4.17 geeft dan het uiteindelijke resultaat van de lokalisatie weer. De afmetingen van de kamer worden exact teruggevonden. Dezelfde kamerafmetingen en microfoonopstelling zijn ook getest in figuur 4.18, waarbij nu een reflectiecoëfficiënt $R = 0.7$ gebruikt wordt in tegenstelling tot $R = 0.5$ zoals in 4.17. De TOA's gevonden zijn gelijkaardig aan deze in tabel 4.8, behalve voor microfoon 4. De TOA in de vierde kolom van 790 samples wordt namelijk niet gevonden, maar wel een foute TOA van 373 samples. Dit leidt tot een ellips die aan geen enkele van de echte reflectoren goed raakt, maar zoals in figuur 4.18 te zien wordt de kamer nog steeds goed gereconstrueerd. Bij een reflectiecoëfficiënt van $R = 0.9$ wijken de TOA's iets harder af, en hetzelfde geldt



FIGUUR 4.12: Resultaat 4 reflectoren, J_e via (4.22), $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m, $R=0.5$.

voor de reflectorlokalisatie. De TOA's voor beide staan weergegeven in tabel 4.9. Tenslotte is het algoritme ook getest met een grotere kamer, namelijk $L = 9$ m en $B = 10$ m. De locatie van de microfoons is opnieuw gewijzigd zodat deze verspreid staan doorheen de kamer. Het resultaat blijkt correct te zijn zoals te zien in figuur 4.20 en bijhorende microfoonposities in tabel 4.13. Als reflectiecoëfficiënt is $R = 0.7$ genomen, de TAO's wijken licht af van de ideale, maar het resultaat blijft correct. De afwijking kan ook gezien worden in figuur 4.20, namelijk de derde grootste ellips behorende bij microfoon 1 (rood) raakt aan geen van de gevonden ellipsen, wat in het ideale geval natuurlijk wel zo zou zijn. Hetzelfde kan gezegd worden voor de op één na grootste ellips horende bij microfoon 2 (cyaan). Hieruit kan besloten worden dat de grootte van de kamer niet de bepalende factor is voor de performantie van het algoritme, wel de posities van de microfoons ten opzichte van hoeken in de kamers en de reflectiecoëfficiënt R . In sectie 4.1.1 werd uitgelegd hoe alle maxima in de RIR's verzameld werden over een gekozen maximale impuls lengte L . Nu de kamergeometrie groter ingesteld is, moet deze L ook langer zijn om de TOA's te kunnen vinden, een lengte $L = 2500$ is gekozen in de simulatie. Om alle resultaten op te sommen geeft tabel 4.6 een overzicht weer.

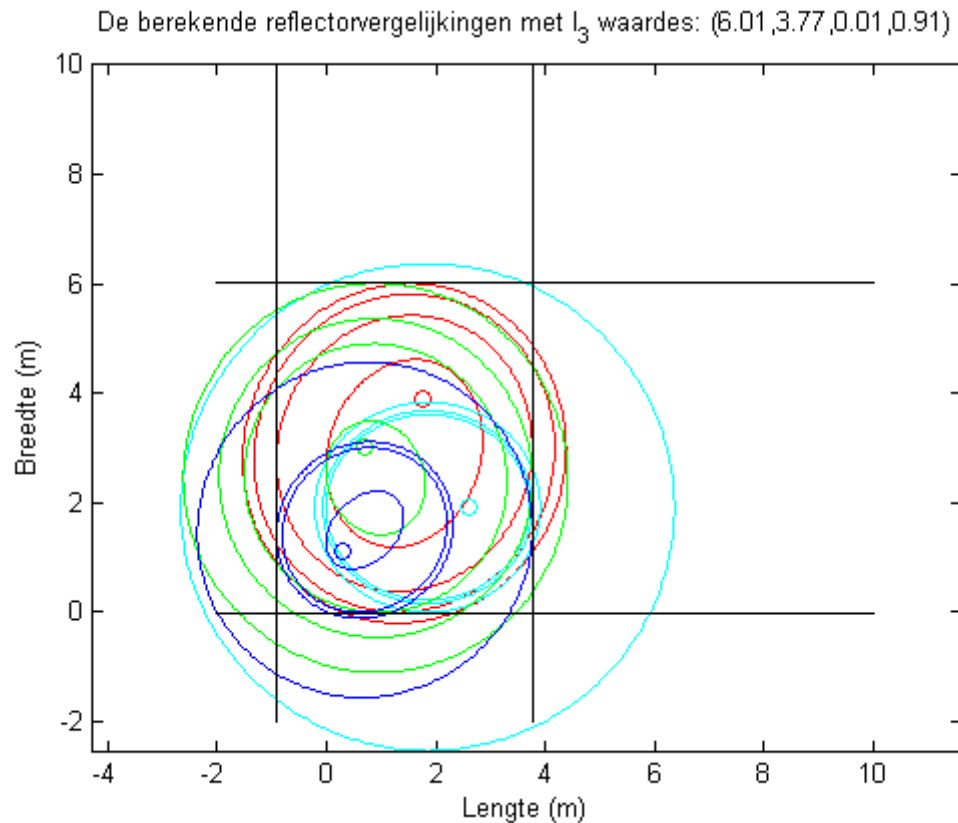


FIGUUR 4.13: Resultaat 4 reflectoren, J_e via (4.22), $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m, $R=0.9$.

TABEL 4.6: Overzicht reflectorlokalisatie

Bepaling l_3	$x = l_3$	$x = -l_3$	$y = l_3$	$y = -l_3$	R	Δ (m)
Echte waarde l_3	3.76	0.00	6.00	0.00		
min. J_e als som M microfoons	0.0	1.9	0.0	0.0	0.5	0.1
min. J_e als som M microfoons	2.44	0.01	5.25	0.01	0.5	0.01
min. J_e als som 2 minima	3.8	0.9	6.0	0.0	0.5	0.1
min. J_e als som 2 minima	3.77	0.92	6.01	0.01	0.5	0.01
min. J_e als som 2 minima	3.77	0.91	6.01	0.01	0.7	0.01
min. J_e als som 2 minima	3.77	0.91	6.03	0.01	0.9	0.01
Echte waarde l_3	3.10	0.00	4.30	0.00		
min. J_e als som 2 minima	3.10	0.00	4.30	0.00	0.5	0.01
min. J_e als som 2 minima	3.1	0.0	4.3	0.0	0.7	0.1
min. J_e als som 2 minima	3.0	0.6	4.7	0.0	0.9	0.1
Echte waarde l_3	9.0	0.0	10.0	0.0		
min. J_e als som 2 minima	9.0	0.0	10.0	0.0	0.7	0.1

4. REFLECTORLOCATIES BEPALEN



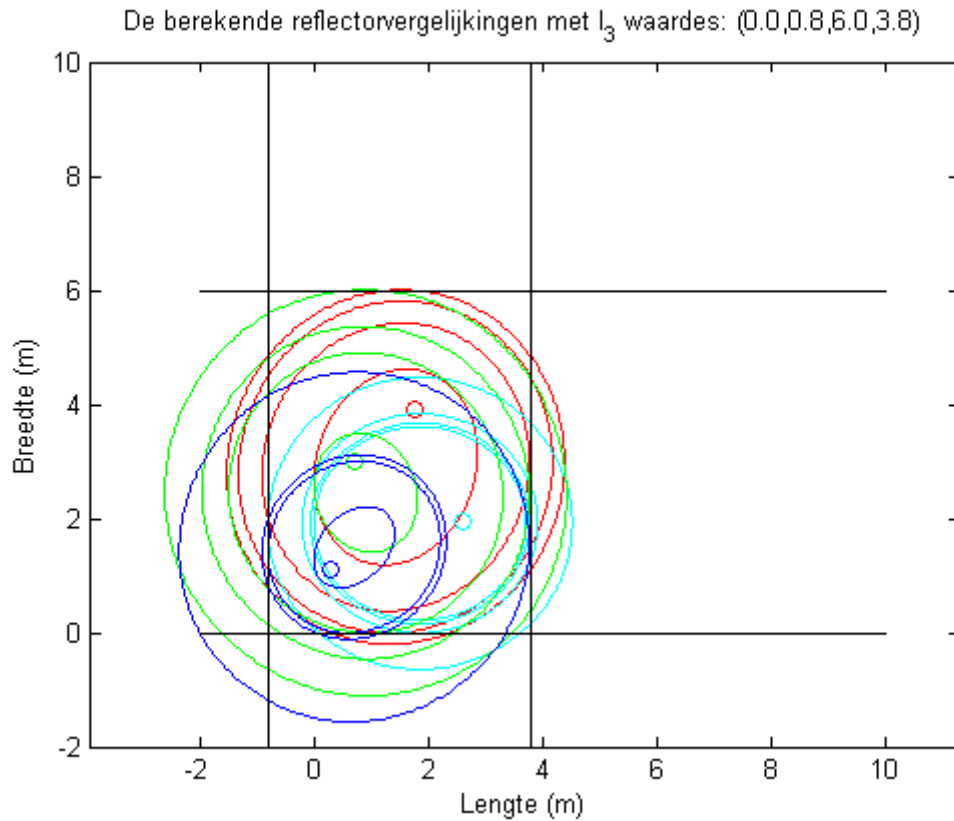
FIGUUR 4.14: Resultaat 4 reflectoren, TOA's via 4 maxima, $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m, $R=0.7$.

	x	y
bron	1.1	1.9
microfoon 1	1.45	3.6
microfoon 2	2.3	1.65
microfoon 3	0.7	3
microfoon 4	0.1	0.7

TABEL 4.7: Locaties van bron en microfoons in 2D-ruimte, opstelling 2.

TABEL 4.8: TOA's in aantal samples, methode 4 maxima, opstelling 2 en vergelijking met ideale. Eerste 4 kolommen met $R=0.5$, laatste vier kolommen zijn de ideale.

	TOA1	TOA2	TOA3	TOA4	TOA1	TOA2	TOA3	TOA4
microfoon 1	399	406	524	716	397.512	404.643	522.278	714.825
microfoon 2	366	444	488	675	364.621	442.191	486.051	673.254
microfoon 3	275	484	590	639	273.615	482.701	588.270	637.673
microfoon 4	222	363	668	790	220.118	361.319	666.946	788.970



FIGUUR 4.15: Resultaat 4 reflectoren, TOA's via alternatieve methode, $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $R=0.7$.

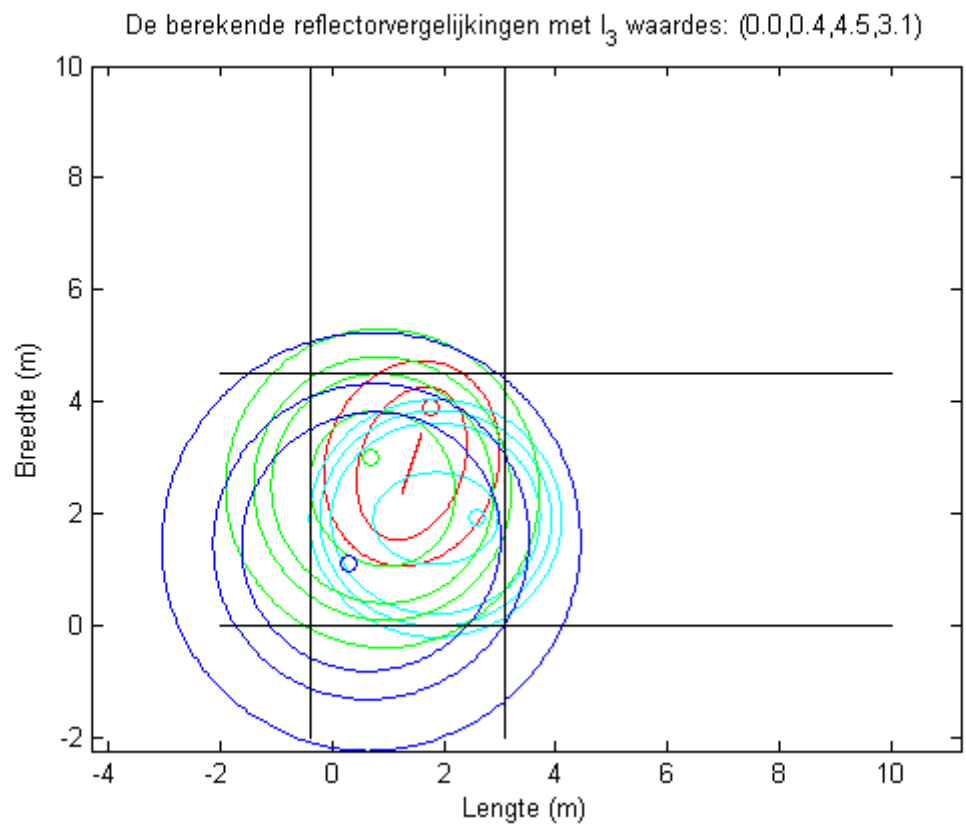
TABEL 4.9: TOA's in aantal samples, methode 4 maxima, opstelling 2 voor $R=0.7$ (eerste 4 kolommen) en $R=0.9$ (laatste 4 kolommen).

	TOA1	TOA2	TOA3	TOA4	TOA1	TOA2	TOA3	TOA4
microfoon 1	399	406	524	716	399	406	522	524
microfoon 2	366	444	488	675	366	444	488	588
microfoon 3	275	484	590	639	275	484	535	590
microfoon 4	222	363	373	668	222	363	373	974

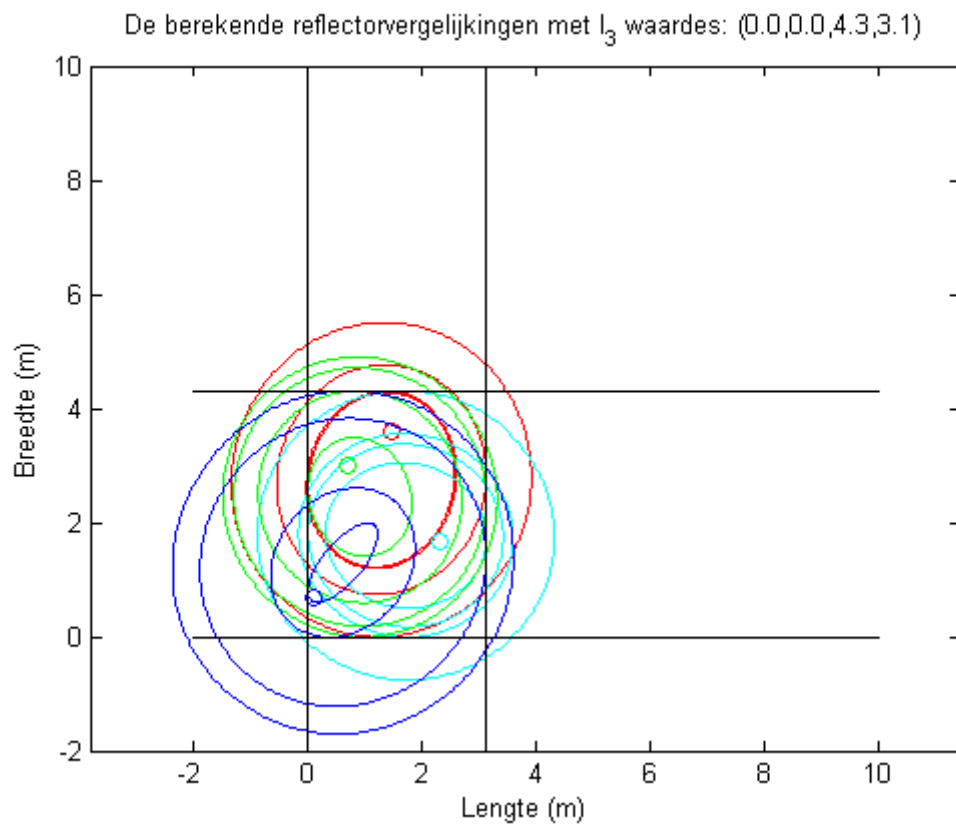
TABEL 4.10: Locaties van bron en microfoons in 2D-ruimte, opstelling 3.

	x	y
bron	2.5	3.7
microfoon 1	2.0	4.5
microfoon 2	8.0	3.0
microfoon 3	1.0	6.0
microfoon 4	5.0	7.5

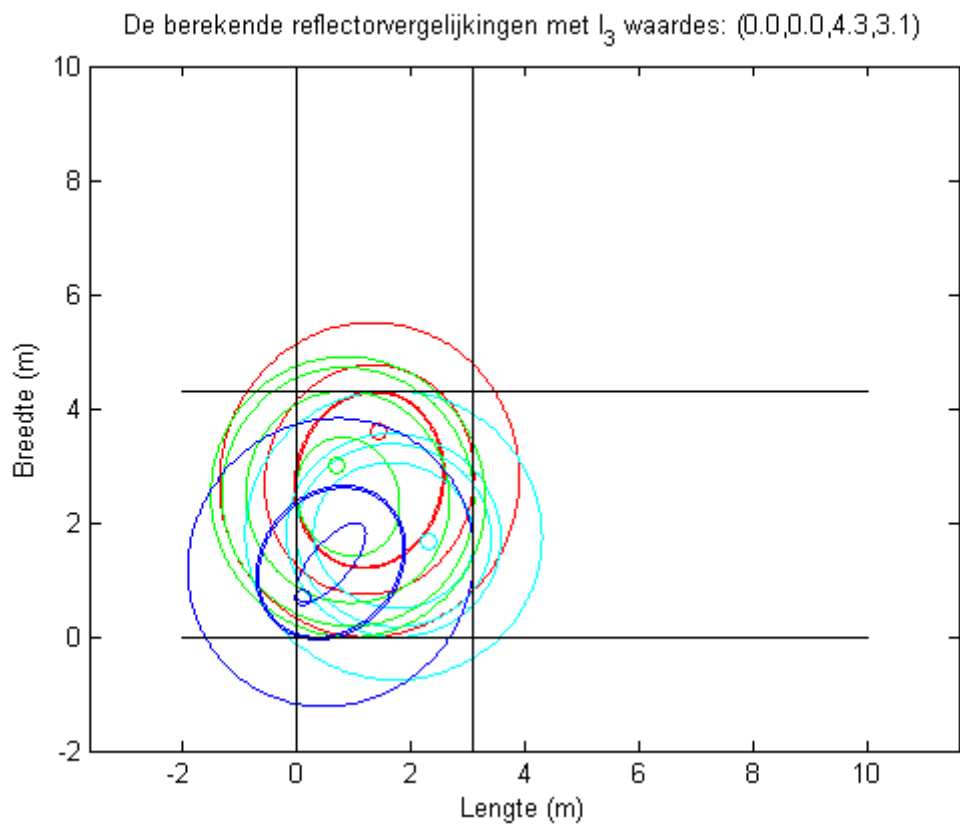
4. REFLECTORLOCATIES BEPALEN



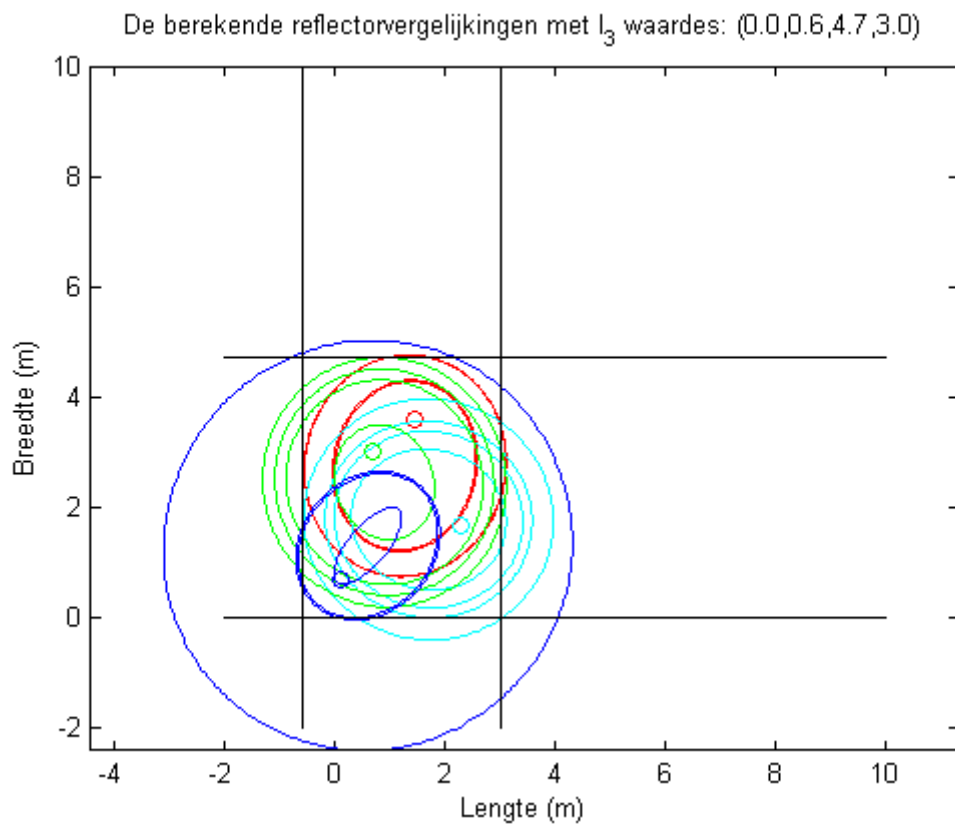
FIGUUR 4.16: Resultaat 4 reflectoren, TOA's via kruiscorrelatie, $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $R=0.7$.



FIGUUR 4.17: Resultaat 4 reflectoren, J_e via (4.22), $\Delta=0.01$ m, $L=3.10$ m, $B=4.30$ m, $R=0.5$.

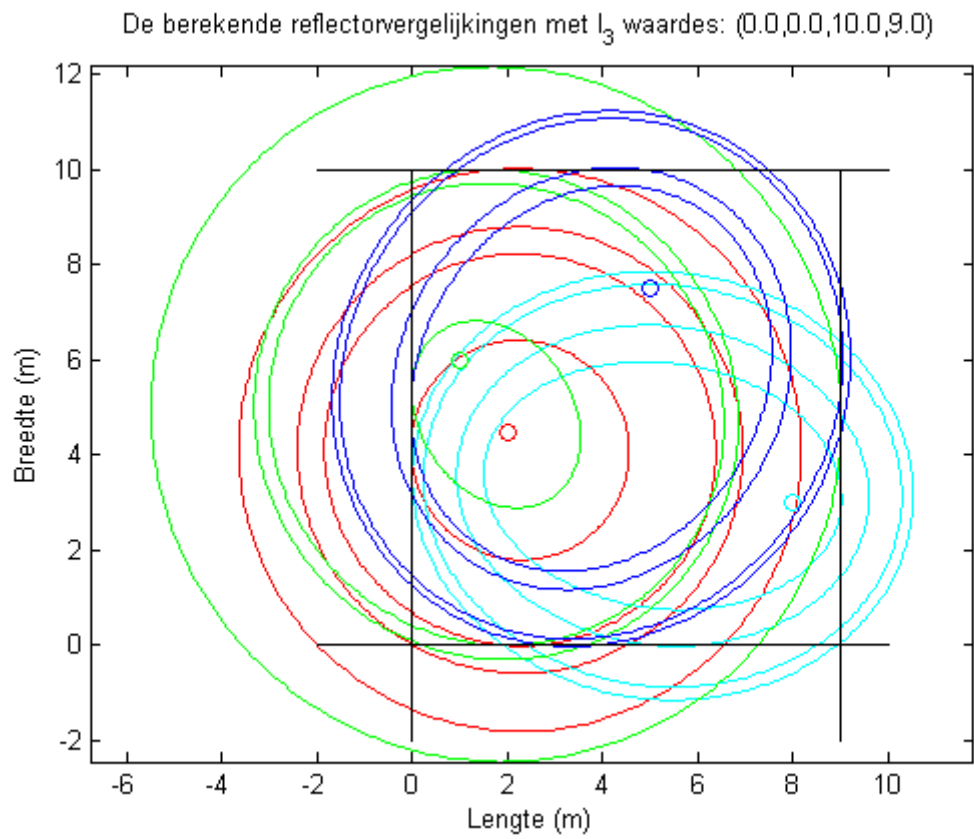


FIGUUR 4.18: Resultaat 4 reflectoren, J_e via (4.22), $\Delta=0.1$ m, $L=3.10$ m, $B=4.30$ m, $R=0.7$.



FIGUUR 4.19: Resultaat 4 reflectoren, J_e via (4.22), $\Delta=0.1$ m, $L=3.10$ m, $B=4.30$ m, $R=0.9$.

4. REFLECTORLOCATIES BEPALEN



FIGUUR 4.20: Resultaat 4 reflectoren, J_e via (4.22), $\Delta=0.1$ m, $L=9.0$ m, $B=10.0$ m, $R=0.7$.

4.2 3D Lokalisatie van de reflectoren

De strategie is gelijkaardig aan het 2D geval. Het verschil zit hem in het feit dat nu zowel microfoon en bron coördinaten in drie dimensies bezitten. Opnieuw komt het aantal reflectoren dat men wil lokaliseren overeen met het aantal TOA's dat men moet schatten uit de RIR's van elke microfoon. Indien zes vlakken van de kamer gezocht worden, zoals bij een balkvormig model voor de kamer, zijn dit zes TOA's. Het extraheren van deze TOA's uit de RIR's werd reeds uitgewerkt in paragraaf 4.1.1. De TOA's geven op hun buurt de totale padlengte tussen een microfoon en de bron weer in drie dimensies, via een reflectie met de muur. De meetkundige plaats van alle mogelijke reflectiepunten is nu geen ellips meer, maar een driedimensionale ellipsoïde. Een eerste paragraaf zal beschrijven hoe de parameters van deze ellipsoïden gevonden kunnen worden. Het zoeken van de reflectoren gebeurt nu niet meer door het zoeken naar raaklijnen, maar wel door het zoeken naar raakvlakken aan de ellipsoïden.

4.2.1 Uitbreiding naar 3D: ellipsen naar ellipsoïden

Een driedimensionaal oppervlak kan beschreven worden via de volgende vergelijking:

$$\zeta = \left\{ (x, y, z) \in \mathbb{R}^3 \mid ax^2 + 2bxy + 2cxz + 2dyz + ey^2 + fz^2 + 2gx + 2hy + 2iz + j = 0 \right\} \quad (4.23)$$

wat leidt tot een symmetrische coëfficiëntenmatrix van de vorm:

$$\mathbf{C} = \begin{bmatrix} a & b & c & g \\ b & e & d & h \\ c & d & f & i \\ g & h & i & j \end{bmatrix} \quad (4.24)$$

Dit kan omgezet worden naar een vectoriële notatie met behulp van homogene coördinaten. Met

$$\mathbf{v} = [x \quad y \quad z \quad 1]^T \quad (4.25)$$

transformeert de vergelijking zich net zoals in het 2D-geval tot

$$\mathbf{v}^T \mathbf{C} \mathbf{v} = 0 \quad (4.26)$$

Hier wordt vertrokken van een eenheidsbol, gedefinieerd door:

$$\mathbf{C}_{\mathbf{I}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (4.27)$$

Een ellipsoïde kan dan op gelijkaardige wijze bekomen worden door achtereenvolgende een schalering, rotatie en een translatie. De schaleringsmatrix $\mathbf{S}_{i,k}$ is van de vorm:

$$\mathbf{S}_{i,k} = \begin{bmatrix} Q_{i,k}^{\text{maj}} & 0 & 0 & 0 \\ 0 & Q_{i,k}^{\text{min}} & 0 & 0 \\ 0 & 0 & Q_{i,k}^{\text{min}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.28)$$

De hoofdas wordt gealigneerd met de x-as, de twee bijassen met de y-as en de z-as. De twee brandpunten van de ellipsoïde liggen opnieuw op de positie van de microfoon en de bron. De totale afstand tot de twee brandpunten wordt gegeven door $\tau_{i,k} \cdot \eta$ met i de beschouwde microfoon, k de reflector en η de geluidssnelheid. Op voorhand is niet geweten of het om een reflector parallel met het xy-vlak gaat, of daarentegen één van de vier wanden welke loodrecht staan op het xy-vlak. Daarom is de enige manier om te zorgen dat de ellipsoïde toch aan de juiste reflector raakt, beide bijassen even lang te maken. Zo wordt een ellipsoïde in de vorm van een ‘rugby bal’ verkregen, ook prolate sferoïde genoemd. De formules voor de semi-lengtes van de assen zijn dezelfde als formules (4.13).

De tweede stap is de rotatie, welke er moet voor zorgen dat de x-as gealigneerd wordt met de rechte die door de coördinaat van de microfoon en van de bron gaat. We kunnen deze rotatie zien als twee achtereenvolgende rotaties. Een eerste rotatie verloopt in het xy-vlak, en is een rotatie rond de positieve z-as. De hoek α waarover verdraaid moet worden, is de hoek tussen de x-as van het referentiestelsel en de projectie van de rechte door microfoon- en broncoördinaat op het xy-vlak. De tweede rotatie is een rotatie rond de negatieve y-as. Met y-as wordt hier de y-as bedoeld van het reeds geroteerde assenstelsel, niet van het referentie assenstelsel. Deze rotatie gebeurt over een hoek β , bepaald door de rechte door microfoon- en broncoördinaat en zijn projectie op het xy-vlak. De uitdrukkingen voor α en β zijn:

$$\tan \alpha_i = \frac{y_s - y_i}{x_s - x_i}, \tan \beta_i = \frac{z_s - z_i}{\sqrt{(x_s - x_i)^2 + (y_s - y_i)^2}} \quad (4.29)$$

De rotatiematrix ziet er dan uit als volgt:

$$\mathbf{R}_i = \mathbf{R}_i(z, \alpha_i) \mathbf{R}_i(y, -\beta_i) = \begin{bmatrix} \cos \alpha_i \cos \beta_i & -\sin \alpha_i & -\cos \alpha_i \sin \beta_i & 0 \\ \sin \alpha_i \cos \beta_i & \cos \alpha_i & -\sin \alpha_i \sin \beta_i & 0 \\ \sin \beta_i & 0 & \cos \beta_i & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.30)$$

Merk op dat de volgorde omwisselbaar is. De eerste rotatie is dan rond de negatieve y-as, over een hoek bepaald tussen de projectie van de rechte door microfoon- en broncoördinaat op het xz-vlak en de x-as. In (4.29) zou de teller van de formule voor α dan wijzigen in $z_s - z_i$. De tweede rotatie rond de z-as moet dan gebeuren over de hoek tussen de rechte door microfoon- en broncoördinaat en zijn projectie op het xz-vlak. Voor β zou dit dan $\frac{y_s - y_i}{\sqrt{(x_s - x_i)^2 + (z_s - z_i)^2}}$ geven.

Tenslotte moet er nog een translatie gebeuren zodanig dat het centrum van de prolate sferoïde op het midden van het lijnstuk tussen de bron en beschouwde microfoon ligt. De translatiematrix is gelijkaardig als (4.10):

$$\mathbf{T}_i = \begin{bmatrix} 1 & 0 & 0 & \Delta x_i \\ 0 & 1 & 0 & \Delta y_i \\ 0 & 0 & 1 & \Delta z_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.31)$$

waarbij Δx_i , Δy_i en Δz_i bepaald worden via:

$$\Delta x_i = x_i + \frac{x_s - x_i}{2}, \Delta y_i = y_i + \frac{y_s - y_i}{2}, \Delta z_i = z_i + \frac{z_s - z_i}{2} \quad (4.32)$$

De coëfficiëntenmatrix voor de prolata sferoïde is te vinden via vergelijking (4.14), met het verschil dat dit nu een 4×4 matrix wordt i.p.v. een 3×3 . Het bepalen van de reflectoren wordt besproken in de volgende paragraaf.

4.2.2 Lokaliseren van de reflectoren als raakvlakken aan de ellipsoïden.

De strategie is gelijkaardig aan deze voor de 2D simulatie. Een vlak in 3D kan geparametriseerd worden door:

$$\psi = \left\{ (x, y, z) \in \mathbb{R}^3 \mid p_1 x + p_2 y + p_3 z + p_4 = 0 \right\} \quad (4.33)$$

Omzetten in een vectoriële vergelijking via

$$\mathbf{p} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \end{bmatrix}^T \quad (4.34)$$

en \mathbf{v} via (4.25) geeft:

$$\mathbf{p}^T \mathbf{v} = 0 \quad (4.35)$$

De uitdrukking die bepaalt of ψ een vlak is, rakend aan de prolata sferoïde met coëfficiëntenmatrix $\mathbf{C}_{i,k}$, is opnieuw vergelijking (4.18). De parameters p_1 , p_2 en p_3 bepalen de oriëntatie van het vlak ψ en liggen binnen eenheidsfeer, parameter p_4 bepaalt de loodrechte afstand van het vlak tot de oorsprong. Een gevolg is dat de parameters p_1 t.e.m. p_3 ook kunnen beschreven worden gebruikmakende van sferische coördinaten:

$$\begin{cases} p_1 = \sin(\phi) \cos(\theta) \\ p_2 = \sin(\phi) \sin(\theta) \\ p_3 = \cos(\phi) \end{cases} \quad (4.36)$$

Hierbij geeft θ de hoek in het xy-vlak t.o.v. de x-as weer, met een bereik van 0 tot 2π , en ϕ de hoek met de z-as met een bereik van 0 tot π . Opnieuw wordt meer in detail het geval bestudeerd waarbij de gezochte wanden parallel zijn met de vlakken van het referentie assenstelsel dat de microfoon- en bronposities definieert. De vier zijwanden van het balkvormige kamermodel komen overeen met parametrisaties waarbij $\phi = \frac{\pi}{2}$ en $\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$. De onderste horizontale reflector of de vloer van de kamer, komt overeen met parametrisatie $\phi = 0$ en de bovenste met $\phi = \pi$. Als vergelijkingen geeft dit:

$$\psi = \begin{cases} x + p_4 = 0 \\ y + p_4 = 0 \\ -x + p_4 = 0 \\ -y + p_4 = 0 \\ z + p_4 = 0 \\ -z + p_4 = 0 \end{cases} \quad (4.37)$$

Merk op dat er in dit geval geen nood is aan twee aparte lussen die gaan over alle mogelijke waarden voor α en β . In de plaats daarvan kan er één reeks opgesteld worden, met de mogelijke combinaties van α en β , en loopt een lus over de lengte

van deze reeks. Net zoals in paragraaf 4.1.3 gaat een bereik gekozen worden voor de parameter p_4 van 0 m tot 10 m. Als discretisatiestap gaat $\Delta = 0.1$ genomen worden, om de rekentijd te beperken. Het verschil met lagere discretisatiestappen wordt verderop getoond. Een eerste verschil met paragraaf 4.1.3 is het aantal mogelijke combinaties van ellipsoïden dat gemaakt kan worden. Aangezien er nu zes reflecterende wanden gezocht worden, zijn er nu ook zes TOA's te schatten per RIR van een microfoon en levert dit zes verschillende ellipsoïden per microfoon. Voor het zoeken van de eerste reflector geeft dit dus $N^M = 6^4 = 1296$ mogelijke combinaties indien $M = 4$. Dit is nog steeds een pak minder dan indien $\binom{NM}{M} = \binom{24}{6} = 10626$ zou gebruikt worden. Een tweede verschil is dat nu drie parameters voor \mathbf{p} gevonden moeten worden, wat een extra lus oplevert in de pseudocode voor het 2D geval. Voor elke parametercombinatie zal opnieuw een kostenfunctie J_e uitgerekend worden, deze is gelijk aan (4.22), met nu vector \mathbf{p} op de plaats van \mathbf{l} . Zoals uitgewerkt in 4.1.3 is deze formule robuuster tegen fouten op de TOA dan formule (4.19). Na de eerste reflector, gaan de gekozen ellipsoïden voor het gevonden raakvlak buiten beschouwing gelaten worden voor de zoektocht naar de volgende reflectoren. Dit verkleint telkens sterk het aantal mogelijke combinaties. De veronderstelling is hier dan ook dat de juiste reflector gevonden is en dus ellipsoïden die hieraan raken niet raken aan de reflectoren die nog gevonden moeten worden.

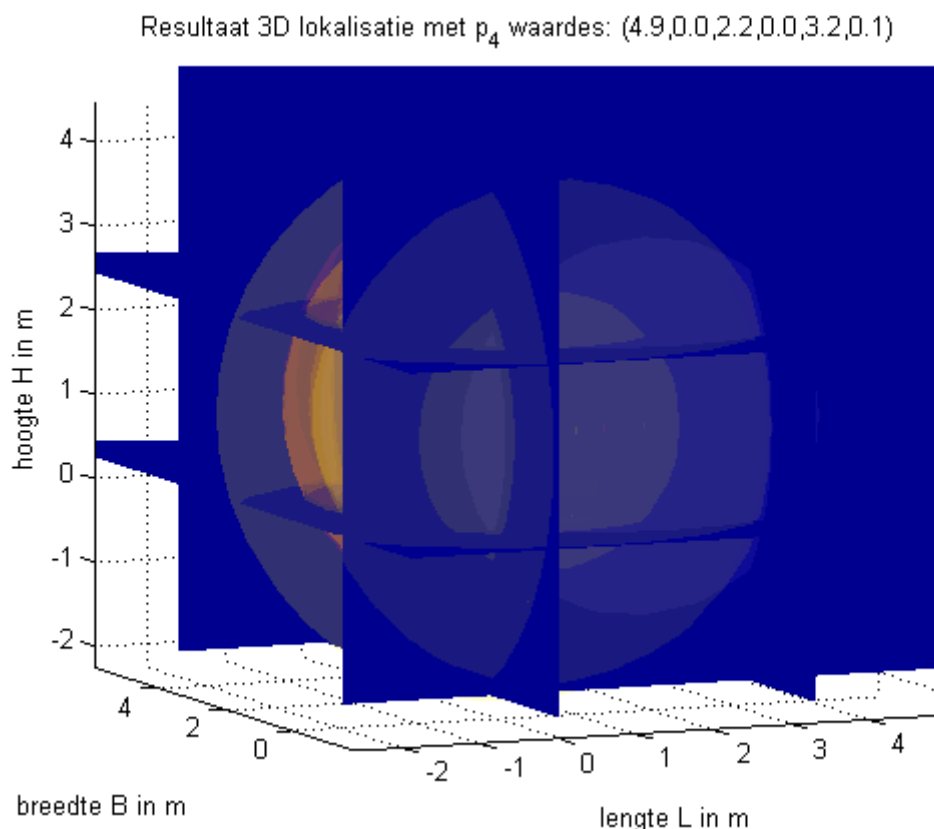
Ook het voorkomen van tweemaal dezelfde reflector als resultaat wordt geïmplementeerd. Voor de vier rechtopstaande wanden gebeurt dit hetzelfde als in paragraaf 4.1.3. Indien een gevonden minimale ϕ -waarde een som van π of 2π vormt met één van de ϕ -waardes voor eerder gevonden reflectoren, moet er gekeken worden of de bijhorende waardes voor de parameter p_4 niet gelijk zijn in absolute waarde en verschillend van teken. Gelijkaardig voor de horizontale vlakken boven en onder, waar moet gekeken worden of de som van de θ -waardes gelijk is aan π . Dan gaat het immers over de twee onderste vergelijkingen uit (4.37). In de volgende paragraaf worden de resultaten van de simulatie getoond voor verschillende kamergeometrieën.

4.2.3 Resultaten van 3D simulatie

Eerst wordt de simulatie uitgevoerd met exacte TOA's. Deze kunnen gevonden worden via een algoritme dat de totale padlengte berekent gegeven de posities van de reflectoren, door het spiegelen van de coördinaten van de microfoon ten opzichte van een bepaalde reflector. Als kamerafmetingen wordt opnieuw $L = 3.76$ m en $B = 6$ m gebruikt voor lengtes langs x-as en y-as, als hoogte wordt $H = 2$ m gekozen. De microfoonposities zijn qua x- en y-coördinaten identiek aan deze in tabel 3.1, de z-coördinaat is voor elke microfoon gelijk aan 1.1 m gekozen. Dit omdat deze opstelling dezelfde is als degene gebruikt in de praktische metingen, zie hoofdstuk 5. In het kamermodel moet de aanpassing gebeuren zodat boven- en onderwand ook reflecterend zijn, in de 2D simulatie werden deze als volledig absorberend ingesteld. Het resultaat is te zien in figuur 4.21. Hierbij is als discretisatiestap $\Delta = 0.1$ m gekozen voor de parameter p_4 . De vergelijkingen die gevonden worden zijn: $y = 4.9$, $y = 0$, $x = 3.1$, $x = -0.1$, $z = 2.2$ en $z = 0$. De positie van de twee horizontale wanden klopt dus vrij goed, deze van de rechter zijwand wijkt 0.66 m af van $x = 3.76$

en de achterste wand toont een fout van 1.1 m op $y = 6$. De resultaten zijn dus minder goed dan in het 2D geval. Omdat de TOA's exact zijn, kan men hier ook formule (4.19) gebruiken. Immers formule (4.22) werd aangepast om robuuster te zijn tegen fouten op TOA's, wat hier niet nodig is. In dit geval wordt dus een raakvlak gezocht dat perfect raakt aan M ellipsoïden. Het resultaat blijkt weinig te verschillen: de gevonden vergelijkingen zijn: $y = 4.9$, $y = 0$, $x = 3.2$, $x = 0$, $z = 2.2$ en $z = 0$.

Een vergelijking met de gevonden reflectorlocaties bij gebruik van $\Delta = 0.01$ m



FIGUUR 4.21: Resultaat 3D lokalisatie, ideale TOA's, $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $H=2$ m.

is te zien in de figuur 4.22. Het resultaat is sterk gelijkende op het resultaat met $\Delta = 0.1$ m. Indien wel wordt vertrokken van de RIR's, moeten er nu zes TOA's uit deze RIR's geschat worden. Opnieuw kan dit gebeuren door de zes grootste pieken uit de RIR's te extraheren, op het directe pad na. In figuur 4.23 is het resultaat te zien indien we een reflectiecoëfficiënt $R = 0.7$ kiezen voor alle wanden. De TOA's staan in tabel 4.11 samen met de ideale TOA's. Hierin zijn toch enkele afwijkingen tussen de geschatte TOA's en de ideale TOA's waarneembaar. Figuur 4.24 toont de positionering van de geschatte eerste orde TOA's in de RIR van microfoon 1. Een vergelijking met een waarde $R = 0.5$ wordt gegeven in tabel 4.12. Tenslotte wordt het

4. REFLECTORLOCATIES BEPALEN

TABEL 4.11: TOA's in aantal samples, methode 6 maxima en vergelijking met ideale.

	Geschat TOA1	TOA2	TOA3	TOA4	TOA5	TOA6
microfoon 1	361	396	453	588	810	840
microfoon 2	305	347	481	497	535	556
microfoon 3	275	280	325	361	397	542
microfoon 4	211	277	315	322	404	541
	Ideaal TOA1	TOA2	TOA3	TOA4	TOA5	TOA6
microfoon 1	376.427	394.752	451.602	479.598	525.629	757.004
microfoon 2	293.161	324.288	345.391	479.928	533.514	798.421
microfoon 3	273.615	300.571	323.225	558.121	637.673	637.673
microfoon 4	209.145	298.041	320.875	402.715	600.492	888.083

resultaat ook getest voor een kamer waarin de microfoonposities meer verspreid zijn in de 3D-ruimte, zie hiervoor tabel 4.13. De afmetingen van deze kamer zijn: $L = 5.3$ m volgens de x-as, $B = 7.4$ m volgens de y-as en $H = 2$ m volgens de z-as. Het resultaat is zichtbaar in onderstaande figuren, met figuur 4.26 die gebruik maakt van geschatte TOA's met $R = 0.7$ en figuur 4.25 die gebruikt maakt van de ideale TOA's. Hierbij dient wel een kleine nuance gemaakt te worden, zoals eerder vermeld vertoont het bronlokalisatiealgoritme in het 3D geval ook meer fouten. Bij de voorgaande microfoonopstelling vormde dit geen probleem, bij de nieuwe opstelling bedraagt de gevonden locatie van de bron (1.3;1.7;1.5) terwijl de echte locatie (2.5;3.7;1.3) is. In figuren 4.25 en 4.26 is hiervan abstractie gemaakt, en handmatig de juiste positie meegegeven aan het algoritme. Indien dit niet gebeurt wordt het resultaat zoals in figuur 4.27 verkregen, gebruik makende van de ideale TOA's.

TABEL 4.12: Overzicht reflectorlokalisatie in 3D (in meter).

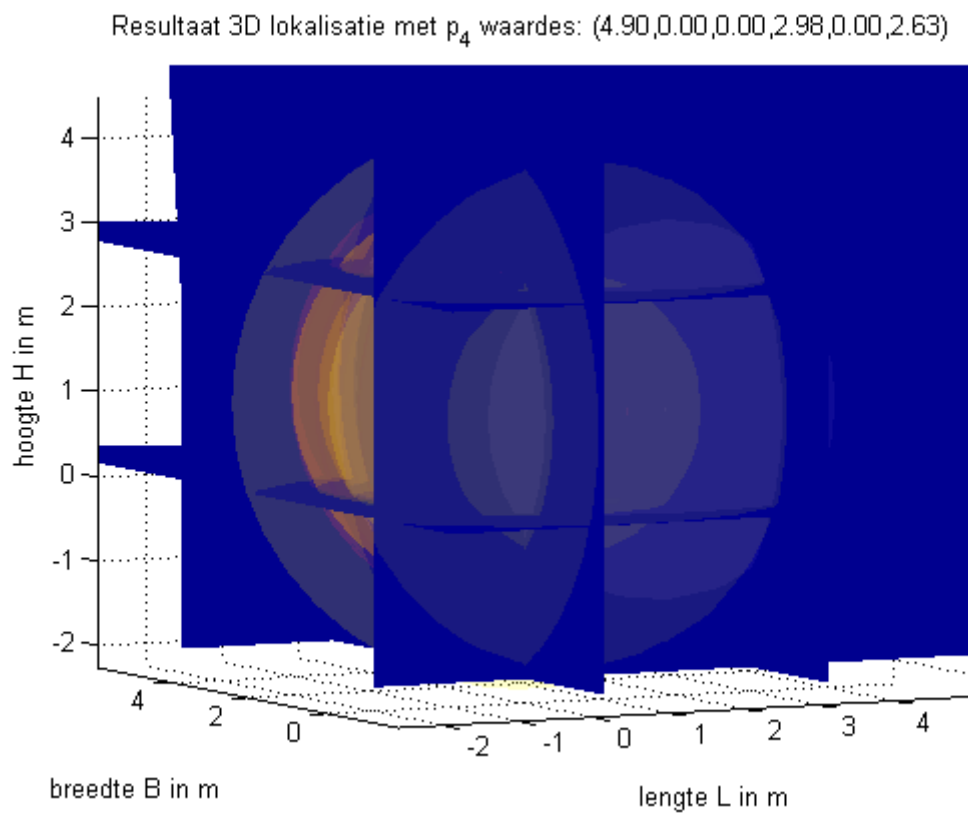
Beschrijving	$x=p_4$	$x=-p_4$	$y=p_4$	$y=-p_4$	$z=p_4$	$z=-p_4$	R	Δ
Echte waarde p_4	3.76	0.0	6.0	0.0	2.0	0.0		
Ideale TOA's, J_e via (4.22)	3.2	0.1	4.9	0.0	2.2	0.0		0.1
Ideale TOA's, J_e via (4.19)	3.1	0.0	4.9	0.0	2.2	0.0		0.1
Ideale TOA's, J_e via (4.19)	2.98	0.0	4.90	0.0	2.63	0.0		0.01
Geschatte TOA's	3.8	0.0	0.4	0.0	3.1	0.9	0.7	0.1
Geschatte TOA's	3.7	0.0	1.9	0.0	3.1	0.9	0.5	0.1
Echte waarde p_4	5.3	0.0	7.4	0.0	2.0	0.0		
Ideale TOA's, $M=4$	5.1	0.0	7.4	0.0	2.0	0.0		0.1
Ideale TOA's, bronloc. fout	3.7	0.1	6.4	1.3	3.2	0.0		0.1
Geschatte TOA's	3.5	0.0	4.7	0.0	1.3	0.3	0.7	0.1
Ideale TOA's, $M=5$	5.1	0.0	7.4	0.0	3.5	0.0		
Geschatte TOA's, $M=5$	5.3	0.0	4.7	0.0	3.3	0.7	0.7	0.1

Vervolgens wordt de simulatie uitgevoerd met $M = 5$ microfoons. De vijfde microfoon krijgt als (3.7;5.3;1.5) coördinaten. De overige vier microfoons behouden

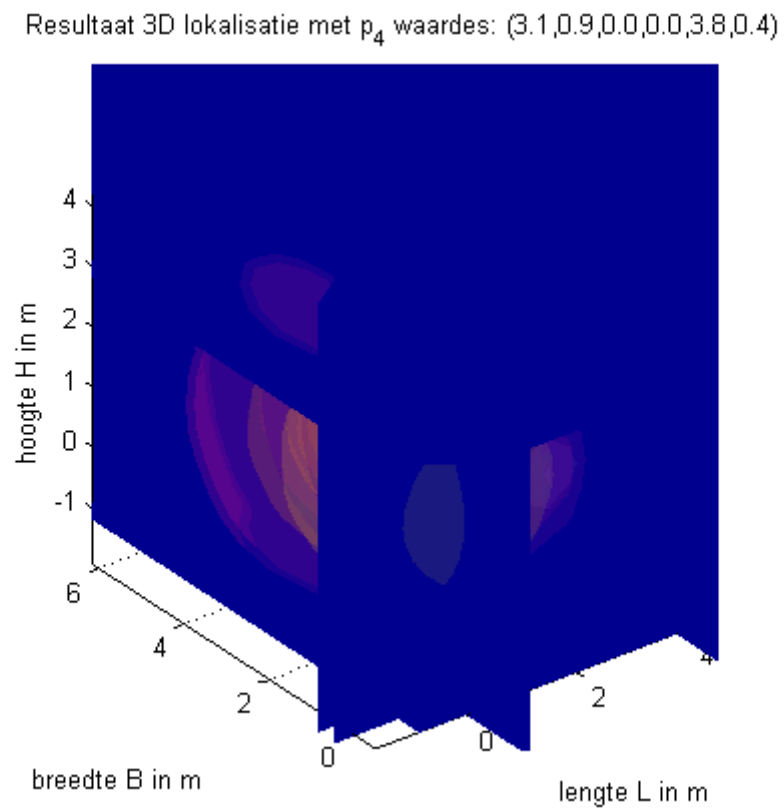
TABEL 4.13: Locaties van bron en microfoons in 3D-ruimte, opstelling 2.

	x	y	z
bron	2.5	3.7	1.3
microfoon 1	1.0	1.0	0.5
microfoon 2	2.5	1.0	0.8
microfoon 3	1.0	0.5	1.7
microfoon 4	0.5	2.0	0.3

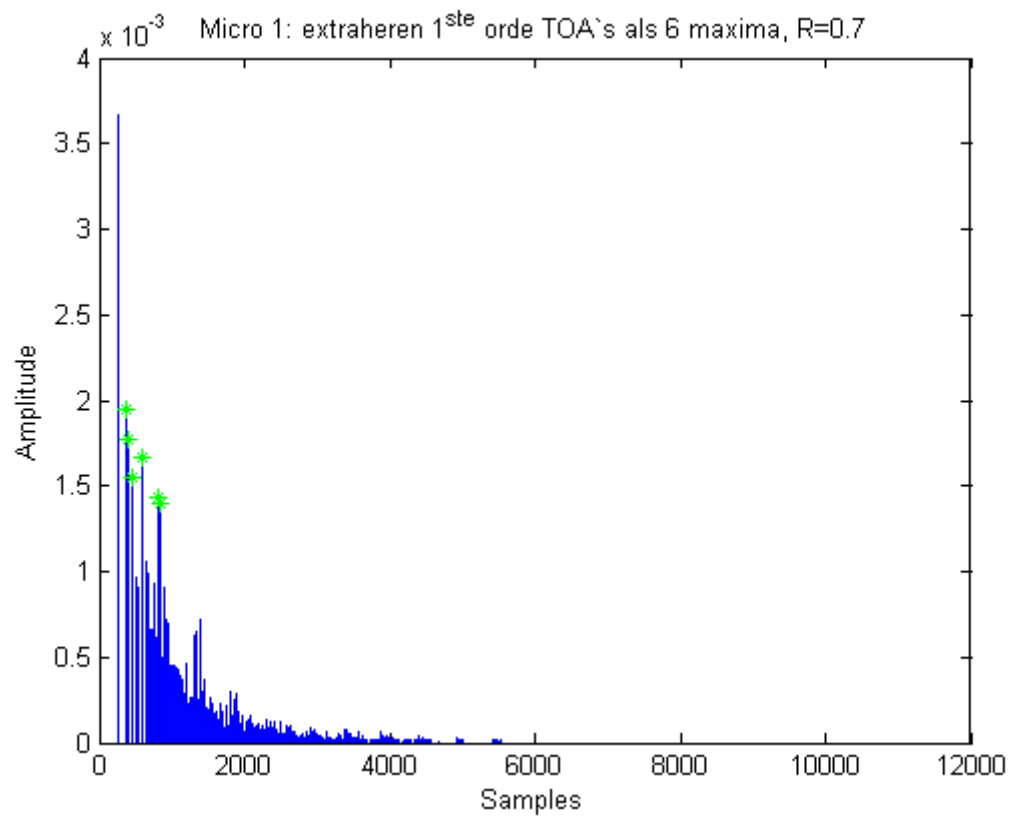
dezelfde locaties als in tabel 4.13. Het resultaat van de bronlokalisering wordt in dit geval beter. In tabel 4.12 staat het verschil aangegeven qua resultaat, zowel via de ideale TOA's als de geschatte TOA's. De gevonden vergelijking voor de rechterwand is beter, in het geval dat de geschatte TOA's worden gebruikt. Deze voor de horizontale vlakken is echter minder goed. Een overzicht van de verschillende simulaties staat in tabel 4.12.



FIGUUR 4.22: Resultaat 3D lokalisatie, ideale TOA's, $\Delta=0.01$ m, $L=3.76$ m, $B=6$ m, $H=2$ m.

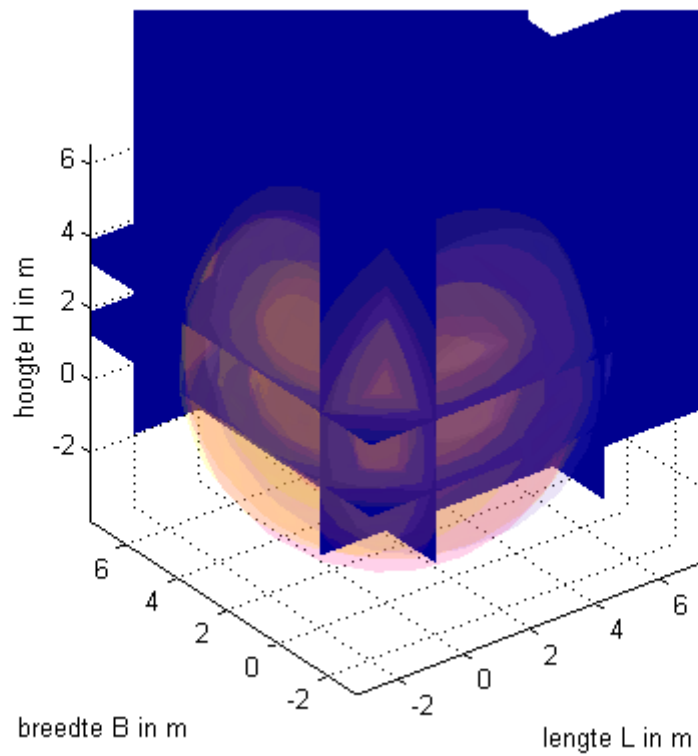


FIGUUR 4.23: Resultaat 3D lokalisatie, geschatte TOA's, $\Delta=0.1$ m, $L=3.76$ m, $B=6$ m, $H=2$ m, $R=0.7$.

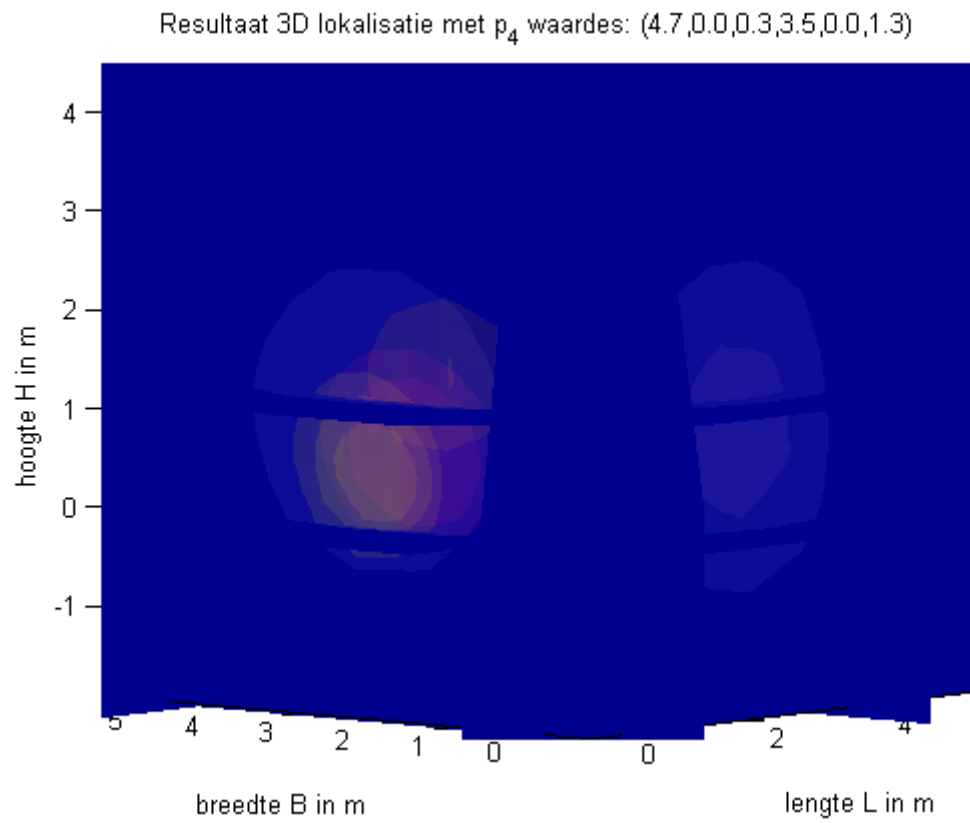


FIGUUR 4.24: RIR microfoon 1 met eerste orde TOA's via methode 6 maxima .

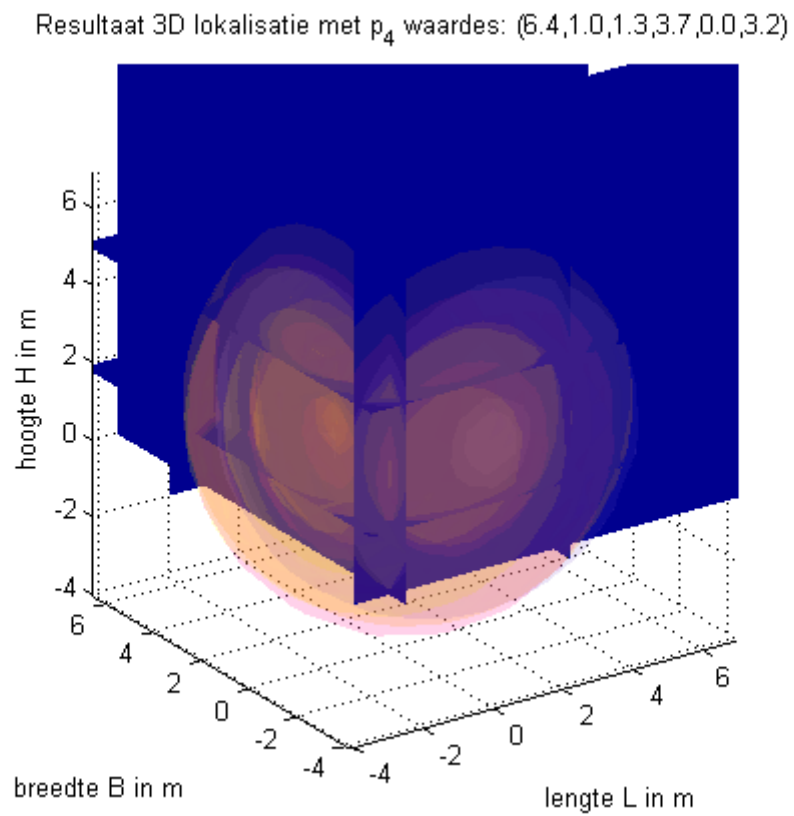
Resultaat 3D lokalisatie met p_4 waarden: (7.4,0.0,0.0,5.1,0.0,2.0)



FIGUUR 4.25: Resultaat 3D lokalisatie, ideale TOA's en correcte bronpositie, $L=5.3$ m, $B=7.4$ m, $H=2$ m.



FIGUUR 4.26: Resultaat 3D lokalisatie, geschatte TOA's, $\Delta=0.1$ m, $L=5.3$ m, $B=7.4$ m, $H=2$ m.



FIGUUR 4.27: Resultaat 3D lokalisatie, ideale TOA's, $\Delta=0.1$, $L=5.3$ m, $B=7.4$ m, $H=2$ m.

4.3 Besluit

Een eerste besluit dat kan gemaakt worden is dat zowel voor het 2D- als het 3D-geval de resultaten van de simulaties sterk afhankelijk zijn van de correctheid waarmee de eerste orde TOA's geschat kunnen worden uit de RIR's. Deze kunnen in het algemeen beter geschat worden indien de wanden van de omgeving een lagere reflectiecoëfficiënt hebben, bijvoorbeeld $R = 0.5$ of $R = 0.7$. De grootte van de kamers blijkt voor de simulaties van minder belang te zijn op het verkrijgen van een al dan niet correcte reconstructie. Indien de microfoons op een korte afstand van hoeken in de kamer gepositioneerd staan, kan dit het schatten van de TOA's ook negatief beïnvloeden. Een voordeel van deze methode in vergelijking met methodes waarbij ook tweede orde TOA's correct geschat moeten worden, is dat deze laatste doorgaans moeilijker te bepalen zijn dan de eerste orde TOA's. Mogelijk kunnen er betere methodes gevonden worden om de TOA's uit de RIR's te extraheren dan het zoeken van de maxima in de RIR's, eventueel met extra voorwaarden zoals in paragraaf 4.1.1 uitgewerkt of via filtertechnieken.

Een tweede opmerking is dat de resultaten voor het 3D-geval iets minder nauwkeurig zijn, zelfs indien de correcte TOA's gebruikt worden in de simulatie. Ook is de benodigde rekentijd beduidend langer dan voor het 2D-geval, omdat nu meer reflectoren gezocht worden, en er dus meer combinaties van ellipsoïden gemaakt kunnen worden waarvoor telkens een passend raakvlak dient gevonden te worden. Het aantal combinaties wordt gegeven door N^M zoals uitgewerkt in 4.1.3. Ook het aantal microfoons M dat men gebruikt, heeft dus ook een sterke invloed op de rekentijd van het algoritme. Verder is ook te besluiten dat het gebruik van (4.22) als kostenfunctie in het algemeen een beter resultaat geeft dan (4.19) bij het zoeken van de raaklijnen of raakvlakken.

Hoofdstuk 5

Praktijktesten

De theorie en simulaties uit voorgaande hoofdstukken zijn ook getest in de praktijk, om te kijken of het mogelijk is de geometrie van een kamer te vinden aan de hand van opgenomen akoestische signalen. Ook nu worden de testen eerst voor het 2D scenario uitgevoerd. Natuurlijk worden deze in een driedimensionale ruimte uitgevoerd, maar gebruikmakende van de algoritmes die een 2D reconstructie opleveren. De microfoons worden op eenzelfde hoogte geplaatst, en het plafond en de vloer worden zoveel mogelijk gedempt. Het resultaat in praktijk is verschillend van wat verwacht wordt volgens de simulaties. Een aantal aanpassingen zijn nodig om het resultaat te optimaliseren. Zo zijn bijvoorbeeld de bron en microfoons niet omnidirectioneel, waardoor een aanpassing nodig is voor het verkrijgen van de impulsresponsies en de TOA's. Deze problemen en andere worden in dit hoofdstuk besproken, eveneens als de resultaten. Na het tweedimensionale geval volgt de bespreking van akoestische omgevingsreconstructie in drie dimensies.

5.1 De labo opstelling

De testen worden uitgevoerd in het spraaklabo van ESAT. Eén luidspreker en vier microfoons worden gebruikt. Deze worden opgesteld volgens de afmetingen die terug te vinden zijn tabel 3.1. De signalen die de microfoons ontvangen worden eerst versterkt met een versterker die een vlakke respons heeft, en daarna in een mengpaneel gevoerd. Met Adobe Audition worden de signalen verwerkt op de computer, en worden ze opgeslagen. De kamer zelf beschikt over gordijnen die geopend en gesloten kunnen worden, teneinde de wanden reflecterend of eerder dempend te kunnen maken. De reflectiecoëfficiënt van een geplooid gordijn begraagt ongeveer $R = 0.4$ voor een golf van 4 kHz [24]. Ook het plafond is voorzien van een gordijn, en op de vloer ligt een tapijt. Een foto van het labo, met de bron en microfoons is te zien op figuur 5.1.

5.2 Resultaten 2D

Eerst wordt het bronlokalisatie algoritme, zoals uitgewerkt in hoofdstuk 3, uitgevoerd op de opgenomen microfoonsignalen. Gegeven deze bronlocatie, trachten we een

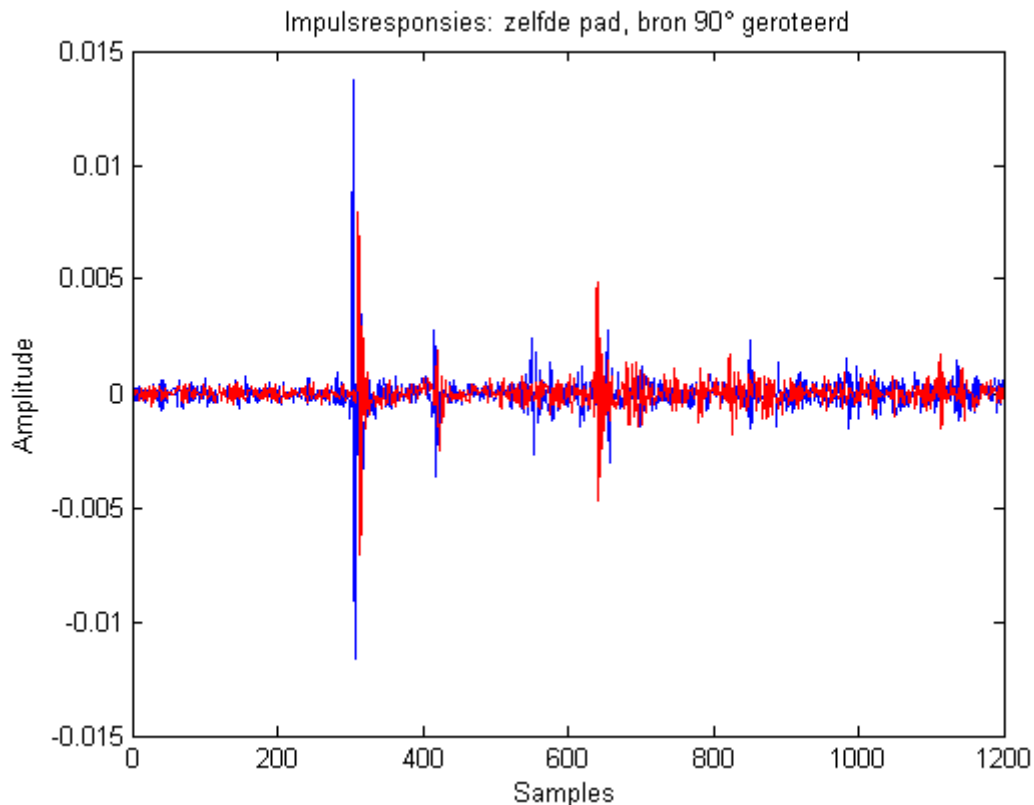


FIGUUR 5.1: Foto van het labo, met de opstelling van bron en microfoons.

akoestische reconstructie uit te voeren van de geometrie van de kamer. In de metingen is als bronsignaal witte ruis gebruikt, zodat alle frequenties even sterk aanwezig zijn.

5.2.1 Bronlokalisatie: praktische problemen

Een reeds vermeld probleem is dat de microfoons en de luidspreker niet ideaal zijn, ze zijn niet omnidirectioneel. Microfoons die achter de luidspreker geïnstalleerd zijn, pikken minder signaal op. De microfoons hadden daarenboven ook een cardioïde responsie karakteristiek. In de testen in de kamer wordt er daarom voor gezorgd dat de microfoons gericht worden in de richting van waar het geluid wordt verwacht. Om met de niet-omnidirectionele luidspreker een gunstig resultaat te verkrijgen worden vier opnames gemaakt, waarbij de luidspreker telkens naar een andere muur wordt gericht. De vier opgenomen signalen per microfoon worden daarna samen behandeld. Voor het directe pad wordt dan van elk signaal de juiste piek gezocht, en daarna per microfoon het gemiddelde genomen van deze vier indices. Als de apparatuur ideaal zou zijn, zouden de vier pieken op dezelfde plaats vallen, maar hier is dit niet het geval. Dit ook omdat wanneer de luidspreker wordt gedraaid, hij niet exact op dezelfde positie blijft staan. Een vergelijking van twee impulsresponsies op de besproken manier gevonden zijn weergegeven in figuur 5.2. Het betreft hier de impulsresponsies van luidspreker naar microfoon 1, waarbij een 90° rotatie van de luidspreker is gebeurd. De pieken vallen ongeveer op dezelfde plek, sommigen zijn minder zichtbaar in de andere responsie door een te felle demping (directionaliteit). Het gemiddelde van de 4 TOA's gevonden per microfoon levert een aanvaardbare oplossing, zoals gezien kan worden in de resultaten.



FIGUUR 5.2: 2 impulsresponsies van hetzelfde pad, waarbij de luidspreker 90° geroteerd is.

5.2.2 Bronlokalisatie: resultaten

Zoals in de simulatie besproken is de eerste stap om het lokaliseren van de bron te verwezenlijken, de TDOA's van elk microfoonsignaal bepalen ten opzichte van een referentie microfoonsignaal. In eerste instantie veronderstellen we dat het bronsignaal niet gekend is, zodat gebruik moet gemaakt worden van blinde schattingsmethodes.

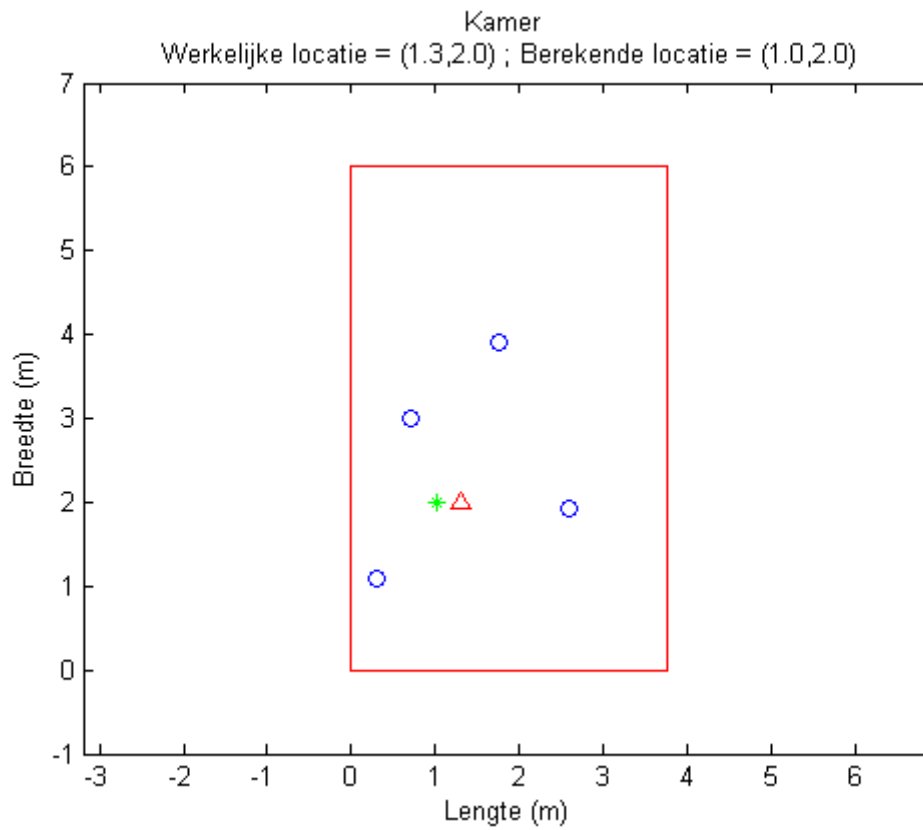
Bronlokalisatie: luidsprekersignaal niet gekend

Bij het uitvoeren van het AED-algoritme, uitgewerkt in hoofdstuk 2, komen we tot de conclusie dat het resultaat minder goed is dan verwacht. De afwijkingen van de bronlocatie gevonden via gebruik van het AED-algoritme zijn behoorlijk groot. Ook met het closed-form algoritme treden er problemen op. Zoals reeds vermeld in paragraaf 2.3.2, wordt er via deze methode geen oplossing gevonden indien de impulsresponsielengte van de RIR's onderschat wordt. Om de impulsresponsie met voldoende nauwkeurigheid te benaderen, is een lange filter vereist. Deze filterlengte is echter niet realiseerbaar in het closed-form algoritme omwille van de te hoge compu-

tationele complexiteit. Ook heeft de conditie van de signalen invloed op het resultaat dat deze methode levert. Indien de gezochte impulsresponsies gemeenschappelijke nulpunten hebben, kunnen deze niet uniek bepaald worden [17]. Dan rest er nog de kruiscorrelatie methode besproken in paragraaf 2.3.1, welke ook een blinde methode is. Deze methode levert goede resultaten voor het vinden van de TDOA's. Het nadeel van deze methode is echter dat deze in eerste instantie geen informatie geeft over de eerste en hogere orde TOA's. Dit is wel het geval bij het closed-form algoritme, en volgens de literatuur ook bij het AED-algoritme hoewel dit in praktijk niet het geval blijkt te zijn. Dit betekent dat men nog een aparte methode moet gebruiken om de eerste orde TOA's te zoeken teneinde de reflectoren te bepalen.

Bronlokalisatie: luidsprekersignaal gekend

Een goede methode die zowel de TOA's geeft nodig voor bronlokalisatie, als deze voor reflectorlokalisatie is de NLMS procedure. Hierbij wordt dan wel het luidsprekersignaal als gekend verondersteld. Uit deze procedure verkrijgt men meteen de impulsresponsies. Hierdoor is er maar één berekening op de signalen nodig voor beide delen (lokalisatie luidspreker en reflectoren). De locatie van de piek, horende bij het directe pad, wordt gevonden via een methode met een drempelwaarde. De drempelwaarde is bepaald door het gemiddelde van de impulsrespons over de eerste 600 samples. Vervolgens wordt elke waarde hiermee vergeleken en wanneer de waarde van de impulsresponsie 5 keer de drempelwaarde overschrijdt, wordt de plaats van maximale waarde in de 10 volgende samples gerelateerd met de TOA van het directe pad gezocht. Het resultaat is weergegeven in figuur 5.3. De locaties van de microfoons zijn samengevat in tabel 3.1. De berekende locatie (1.0;2.0), weergegeven door het sterretje, komt vrij goed overeen met de werkelijke (1.3;2.0). De afwijking van 30 cm is te wijten aan het feit dat de luidspreker, die zelf een zekere omvang heeft, niet exact op het aangegeven punt geplaatst kan worden. Het spreekt voor zich dat deze afwijking ook zijn impact zal hebben op de lokalisering van de muren.

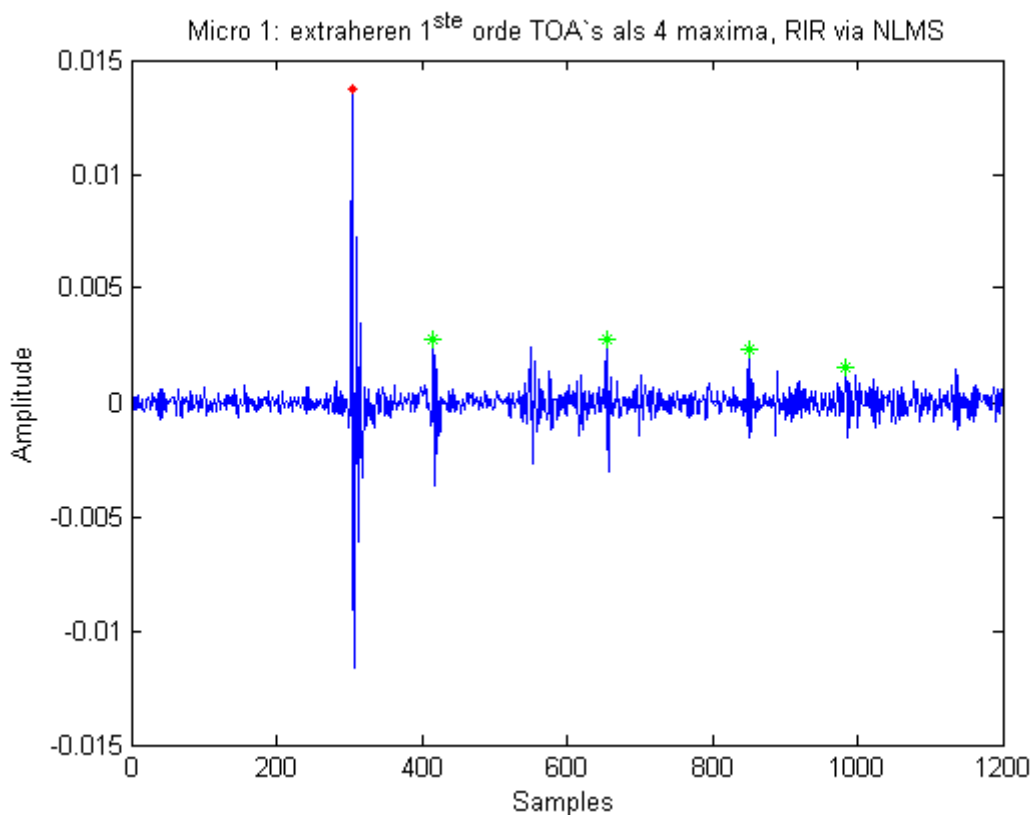


FIGUUR 5.3: Resultaat van de 2D bronlokalisatie in het labo (\circ : microfoon; Δ : werkelijke locatie; *: berekende locatie).

5.2.3 Reflectorlokalisatie: resultaten

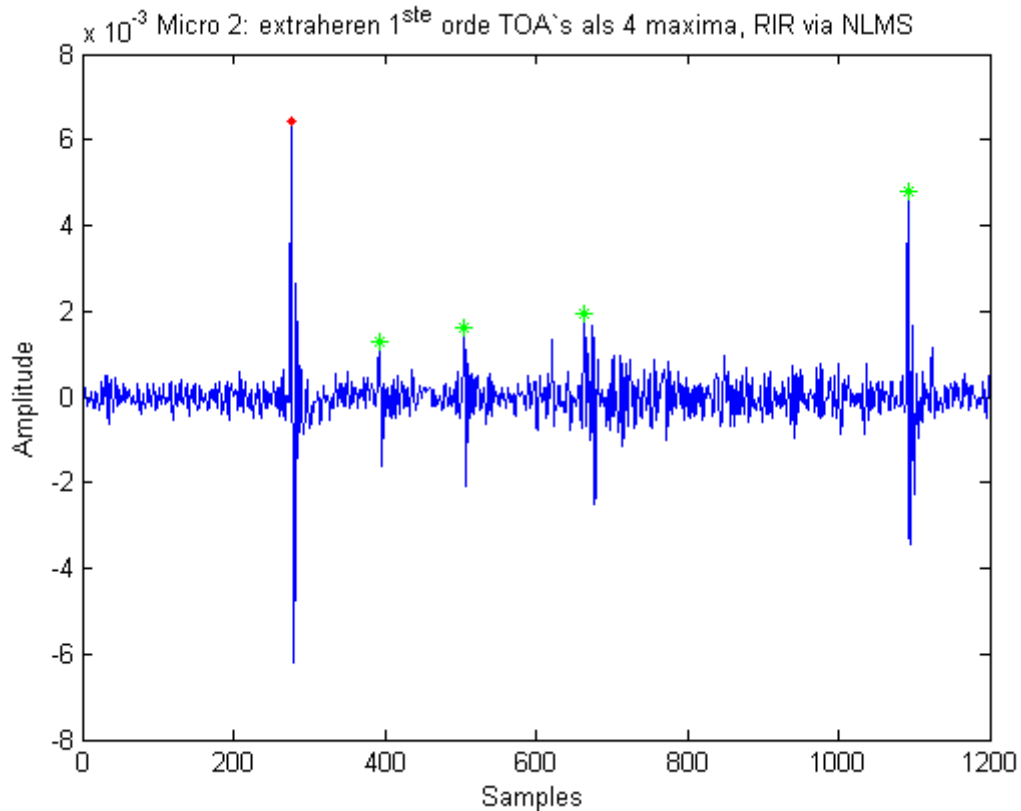
In vorige sectie haalden we reeds de problemen met het blind schatten van de nodige impulsresponsies aan. Het AED-algoritme blijkt hiervoor niet geschikt, het closed-form algoritme levert alleen een berekenbare oplossing indien we korte impulsresponsies hanteren. Indien het luidsprekersignaal wel gekend is, kan rechtstreeks NLMS gebruikt worden zowel voor het lokaliseren van de luidspreker als van de reflectoren. Figuren 5.4 en 5.5 geven twee RIR's weer gevonden via de NLMS methode, met aanduiding van de geschatte eerste orde TOA's. De signalen zijn 10000 samples lang, wat neerkomt op 227 ms. Merk de verschillen op met de twee RIR's die gegeven worden door het room model nl. 4.1 en 4.2.

Het verwachte exponentiële verloop is niet altijd even mooi zichtbaar, ook gebeurt



FIGUUR 5.4: RIR micro 1 bepaald via NLMS, N=10000.

het weleens dat er nog een piek groter dan deze van het directe pad voorkomt in de vroege reflecties. Zoals eerder toegelicht is er nu ook afhankelijkheid van het directionaliteitspatroon van de microfoons en de luidspreker. De reflector die het meest in de richting van de luidspreker ligt, zal een grotere piek in de RIR's vertonen dan de andere. Een reflector achter de luidspreker is zelfs nauwelijks te herkennen. Het resultaat zal beter zijn voor de lagere frequenties, omdat de microfoons en de

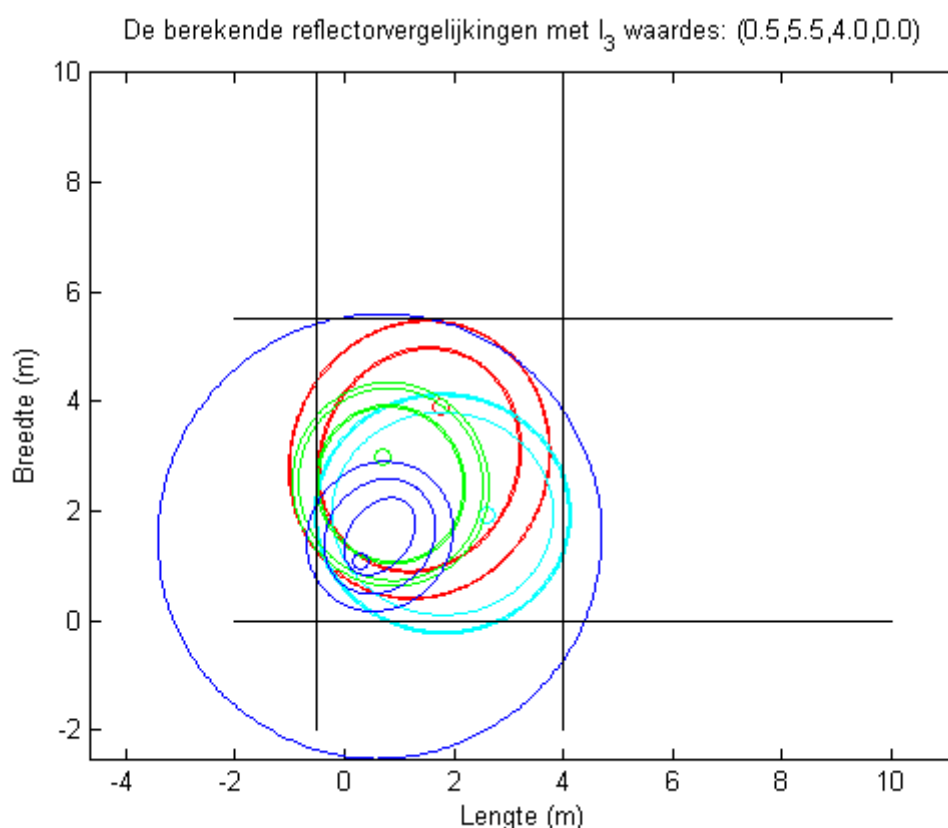


FIGUUR 5.5: RIR micro 2 bepaald via NLMS, $N=10000$.

luidspreker een meer omnidirectionele karakteristiek vertonen voor lagere frequenties. De kamerafmetingen zijn dezelfde als deze in figuur 4.11 uit vorig hoofdstuk, namelijk $L = 3.76$ m en $B = 6$ m. De opstelling van de microfoons is dezelfde als deze in tabel 3.1. Indien we vanuit één meting, met de luidspreker in een bepaalde richting, het algoritme van paragraaf 4.1.1 gebruiken om de TOA's van de vier gezochte reflectoren te schatten, krijgen we het resultaat van tabel 5.1. Hierbij zijn de gordijnen van het labo open gelaten, wat neerkomt op een hogere reflectiecoëfficiënt R . De TOA's worden gezocht door via de methode vier maxima te zoeken, aangezien deze in de simulaties ook het beste resultaat gaf. Indien we de gordijnen sluiten, dus een lage R , krijgen we de TOA's van tabel 5.2. De ideale TOA's werden reeds gegeven in vorig hoofdstuk in tabel 4.2. Men kan deze ook vergelijken met het resultaat gevonden uit de simulaties, zoals in tabellen 4.1, 4.4 en 4.5 voor verschillende waarden van R . In paragraaf 4.1.1 werd reeds gesproken over de mogelijkheid een beperking op te leggen op de minimale afstand tussen opeenvolgende TOA's. In de hoger vermelde figuren van de RIR's, bekomen via het toepassen van het NLMS-algoritme op de opgenomen signalen, zien we dat opeenvolgende pieken kort achter elkaar volgen. Vandaar dat in dit geval een lage minimale afstand, zoals 15 samples, als extra beperking instellen toch een verschil maakt. Het verschil met beperking is ook weergegeven in tabel 5.1

in de vier laatste kolommen. Het algoritme om de reflectoren te zoeken is gelijk aan hetgene beschreven in 4.1.3. Het resultaat voor lage R wordt geïllustreerd in figuur 5.6, met de TOA's uit tabel 5.2. Als kostenfunctie is formule 4.22 gebruikt. Het resultaat vertoont afwijkingen maar is vrij goed. Het resultaat voor hoge R wordt getoond in figuur 5.7. Hieruit blijkt dat het resultaat met lagere R beter is. Het toepassen van een beperking op de afstand tussen de TOA's blijkt weinig voordeel op te leveren. Ter vergelijking staan de gevonden reflectorvergelijkingen opgesomd in tabel 5.5.

Een tweede manier van aanpakken is, zoals ook in voorgaande paragraaf over

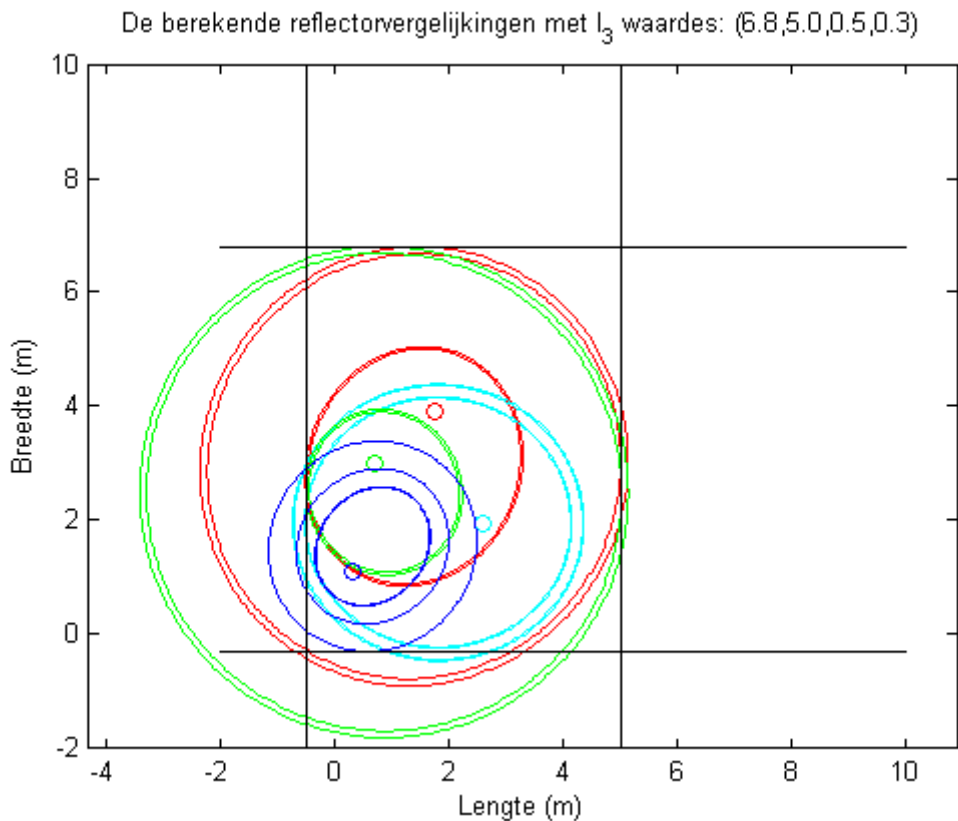


FIGUUR 5.6: Resultaat 2D lokalisatie met opgenomen microfoonsignalen, luidspreker in 1 richting, lage R , $L=3.76$, $B=6$.

de luidsprekerlocatie, de directionaliteit van de luidspreker en microfoons te compenseren door de luidspreker telkens te verdraaien. Men krijgt dan vier signalen per microfoon, uit elk van deze worden de impulsresponsies geschat via NLMS en de eerste orde TOA's bepaald. Deze worden dan uitgemiddeld over de vier signalen. De TOA's die men hiermee verkrijgt staan weergegeven in tabel 5.3. Figuur 5.8 toont het resultaat van de 2D reflectorlokalisatie. Hierbij zijn de gordijnen open, dus een hoge R . Opnieuw staat een vergelijking met lagere R in de overzichtstabel 5.5. In dit geval blijken de signalen met hogere R een beter resultaat op te leveren. Indien we stellen

TABEL 5.1: TOA's in aantal samples, luidspreker in één richting, hoge R . Vier linkerkolommen: zonder beperking op afstand tussen TOA's. Vier rechterkolommen: met beperking op afstand tussen TOA's.

	TOA1	TOA2	TOA3	TOA4	TOA1	TOA2	TOA3	TOA4
microfoon 1	543	553	971	1003	543	971	1003	1159
microfoon 2	602	607	656	662	602	656	838	1005
microfoon 3	366	379	1087	1117	268	379	1087	1117
microfoon 4	283	287	365	487	287	365	487	1390

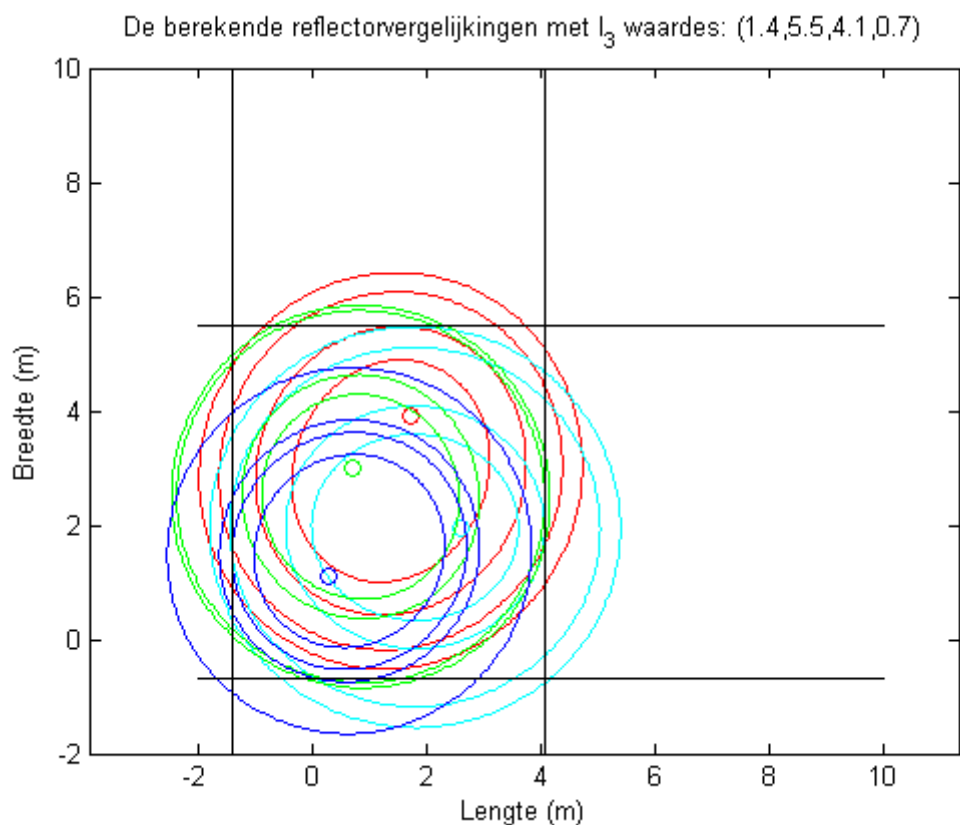


FIGUUR 5.7: Resultaat 2D lokalisatie met opgenomen microfoonsignalen, luidspreker in 1 richting, hoge R , $L=3.76$, $B=6$.

dat per verdraaiing van 90 graden van de luidspreker, deze telkens hoofdzakelijk naar een bepaalde reflector gericht is, kan het probleem nog eenvoudiger aangepakt worden. Deze veronderstelling klopt als we een rechthoekige kamergeometrie veronderstellen, en de luidspreker niet gedraaid staat richting een hoek van de kamer. In dit geval moeten niet vier maxima gezocht worden in de RIR's, maar is slechts één maximum voldoende. Men weet nu immers dat bij elke meting waarbij de luidspreker verdraaid wordt, de TOA, gevonden als het maximum in de RIR en verschillend van

TABEL 5.2: TOA's in aantal samples, luidspreker in één richting, lage R .

	TOA1	TOA2	TOA3	TOA4
microfoon 1	534	543	663	668
microfoon 2	524	597	602	608
microfoon 3	368	378	457	484
microfoon 4	203	287	365	1057



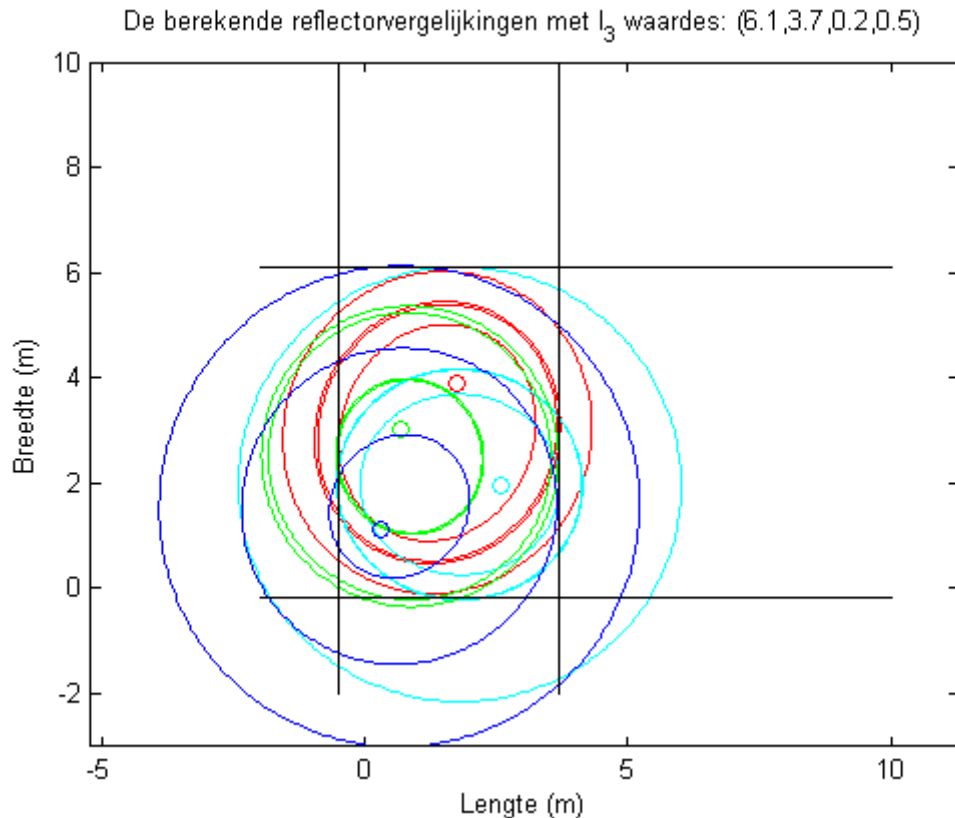
FIGUUR 5.8: Resultaat 2D lokalisatie met opgenomen microfoonsignalen, gemiddelde over verschillende luidspreker richtingen.

het directe pad, hoort bij een reflector in het ‘gezichtsveld’ van de luidspreker. Indien we dezelfde opgenomen signalen gebruiken, met gordijnen open (hoge R), krijgen we de TOA's in de eerste vier kolommen van tabel 5.4 als oplossing. Het resultaat van de reflectorlokalisatie kan men zien in figuur 5.9. In vergelijking met de voorgaande manieren vertoont dit het beste resultaat. De gevonden TOA's liggen ook dichter bij de ideale in tabel 4.2 dan deze in voorgaande tabellen. De methode is ook uitgetest met gesloten gordijnen, hier blijkt het resultaat minder goed te zijn. Dit is in contrast met de simulatie in hoofdstuk 4. De reden hiervoor is dat de reflectiecoëfficiënt van de gordijnen te laag is om alle TOA's te vinden door de te grote verzwakking. Dit levert

TABEL 5.3: TOA's in aantal samples, luidspreker in verschillende richtingen, hoge R .

	TOA1	TOA2	TOA3	TOA4
microfoon 1	513	662	818	905
microfoon 2	470	588	841	933
microfoon 3	467	555	846	869
microfoon 4	448	546	602	836

de TOA's in de laatste vier kolommen van tabel 5.4 welke opnieuw harder afwijken van de ideale. Merk op dat indien een zo goed als omnidirectionele speaker gebruikt zou worden, deze manier van aanpak niet nodig is, en men slechts één signaal per microfoon moet opnemen om de reflectorlokalisatie te verwezenlijken. Anderzijds in een praktische toepassing met bijvoorbeeld een automatisch draaiende microfoon een mogelijke oplossing. Tenslotte kunnen we ook de eerste orde TOA's trachten te

FIGUUR 5.9: Resultaat 2D lokalisatie met opgenomen microfoonsignalen, 1 TOA per luidsprekerrichting, hoge R , $L=3.76$, $B=6$.

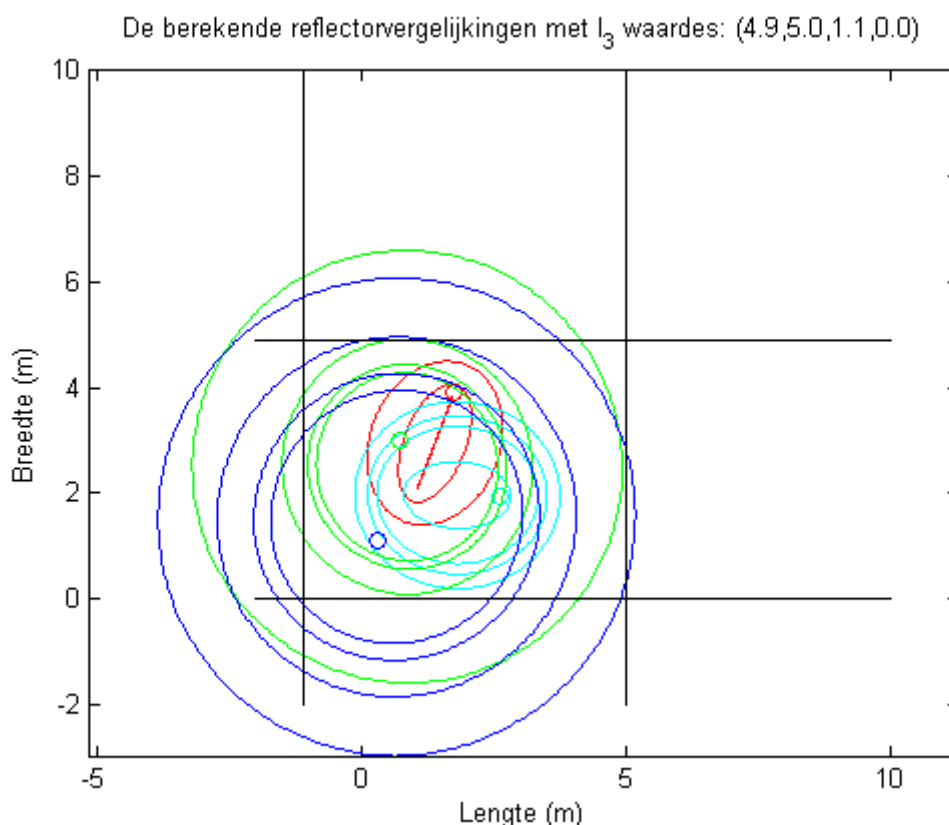
schatten via een berekening van autocorrelaties van de opgenomen signalen. Deze methode kan gebruikt worden indien blind schatten noodzakelijk is, en methodes

5. PRAKTIJKTESTEN

TABEL 5.4: TOA's in aantal samples, 1 TOA per luidsprekerrichting, vier linkerkolommen: hoge R en vier rechterkolommen: lage R .

	TOA1	TOA2	TOA3	TOA4	TOA1	TOA2	TOA3	TOA4
microfoon 1	543	640	654	801	412	418	543	797
microfoon 2	491	602	806	1092	438	505	604	608
microfoon 3	397	387	709	743	363	378	717	749
microfoon 4	365	365	786	1190	365	367	383	727

zoals het closed-form algoritme niet mogelijk zijn wegens te weinig rekenkracht. Deze methode heeft echter enkele nadelen. Ten eerste kan men er de luidsprekerlocatie niet in eenzelfde stap uithalen zoals bij NLMS, hiervoor zijn kruiscorrelaties nodig tussen de microfoonsignalen. Ten tweede is de nauwkeurigheid afhankelijk van het luidsprekersignaal. Ten derde blijkt uit de resultaten dat de TOA-schattingen minder nauwkeurig zijn dan via NLMS. Het resultaat met hoge R is getekend in figuur 5.10 en beter dan verwacht. Het resultaat met lage R daarentegen vertoont grote fouten. Ter overzicht worden de resultaten samengevat in tabel 5.5.



FIGUUR 5.10: Resultaat 2D reflectorlokalisatie via autocorrelatie, hoge R .

TABEL 5.5: Overzicht reflectorlokalisatie: praktische metingen.

Beschrijving	$x = l_3$	$x = -l_3$	$y = l_3$	$y = -l_3$	Refl. coëff.
Echte waarde l_3	3.76	0.00	6.00	0.00	
Eén luidsprekerrichting	5.0	0.3	6.8	0.5	hoog
Eén luidsprekerrichting met constraint	5.0	0.5	6.8	1.7	hoog
Gem. over 4 richtingen	4.1	1.4	5.5	0.7	hoog
Eén TOA per richting	3.7	0.5	6.1	0.2	hoog
Autocorrelatie	4.0	1.4	5.3	0.0	hoog
Eén luidsprekerrichting	4.0	0.5	5.5	0.0	laag
Gem. over 4 richtingen	3.8	1.9	4.2	0.0	laag
Eén TOA per richting	0.1	0.3	5.0	0.5	laag
Autocorrelatie	5.0	1.0	4.9	0.0	laag

5.3 Resultaten 3D

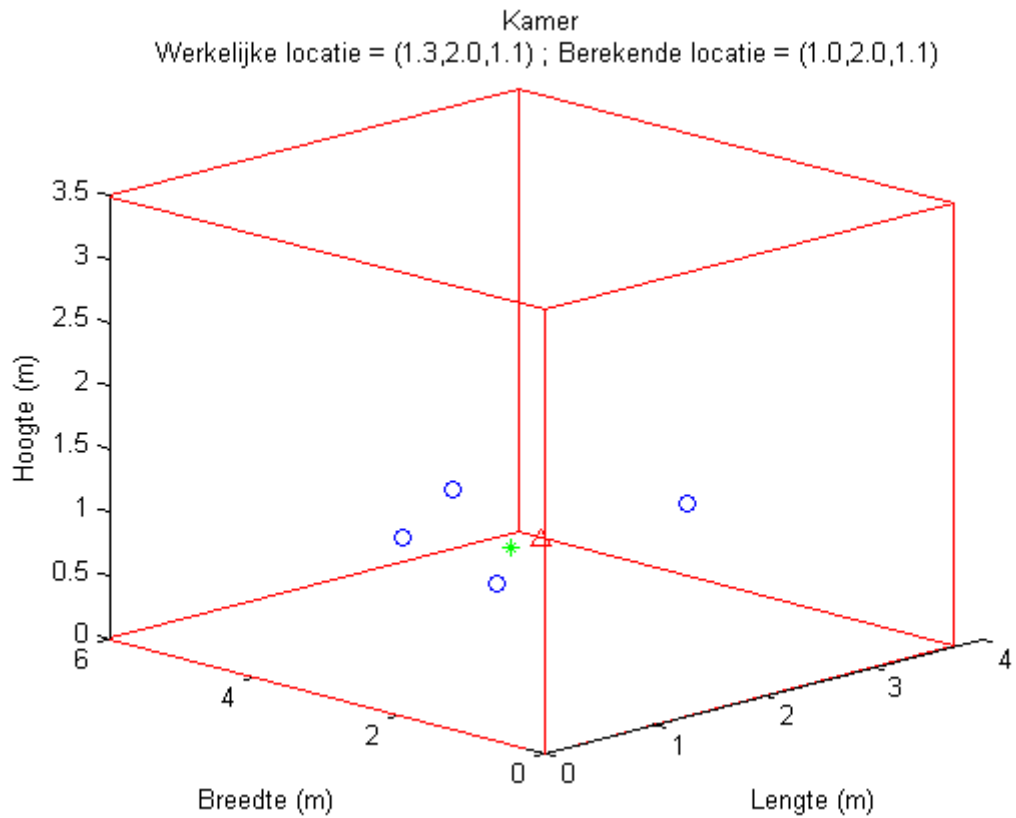
Na de testen van het twee dimensionale geval is het nu de beurt aan de uitbreiding naar drie dimensies. Alle opmerkingen uit het vorige hoofdstuk zijn natuurlijk ook in 3D aanwezig. De labo opstelling is dezelfde, met dien verstande dat plafond en vloer ook in rekening worden gebracht. De locaties van de microfoons, zijn dezelfde als in de simulatie, samengevat in tabel 3.2.

5.3.1 Bronlokalisatie: resultaten

Om dezelfde redenen als in het 2D geval, maken we gebruik van het NLMS algoritme om de TOA's van het directe pad te vinden. Zoals in figuur 5.11 kan gezien worden, wordt een redelijk goed resultaat gevonden met de methodes besproken in dit en vorige hoofdstukken. De gevonden locatie (1.0;2.0;1.1) komt kort bij de echte locatie van de luidspreker (1.3;2.0;1.1), waarmee in het labo getest is. Op de figuur komt de werkelijke locatie overeen met het driehoekje, en de berekende met het sterretje. De cirkeltjes geven de microfoons weer. De reden van de afwijking is hier ook te wijten aan de niet ideale luidspreker. Als hij gedraaid wordt, verplaatst automatisch ook de oorsprong van het geluid.

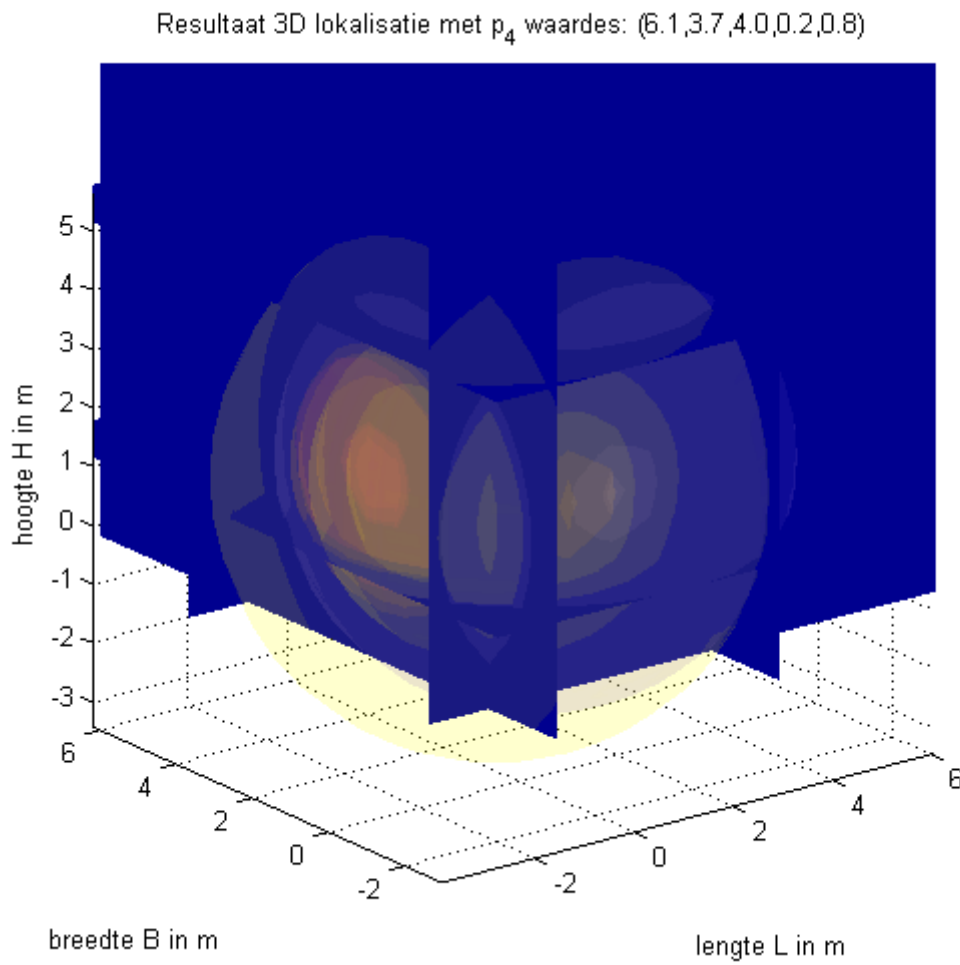
5.3.2 Reflectorlokalisatie: resultaten

In de paragraaf over de lokalisatie in 2D vertrekkende van de praktische metingen, werd reeds getoond dat het beste resultaat verkregen wordt indien een reeks microfoonsignalen worden opgenomen met de luidspreker telkens verdraaid. Daarom dat we in deze paragraaf enkel deze methode gaan bespreken. Om het resultaat te testen hebben we een extra microfoonsignaal toegevoegd, opgenomen met de speaker gericht naar boven. Wegens beperkte middelen in het spraaklabo, was het moeilijker om hetzelfde te doen voor de vloer van de kamer. Daarom dat in de uitvoering van het algoritme de vloer op $z = 0$ wordt ingesteld, en er slechts 5 reflectoren meer gezocht worden. Het resultaat hiervan kan men zien in figuur 5.12. Het gebruikte algoritme



FIGUUR 5.11: Resultaat van de 3D bronlokalisatie in het labo (○: microfoon; Δ: werkelijke locatie; *: berekende locatie).

is hetzelfde als hetgene beschreven in de 3D-simulatie. De verticale wanden worden vrij goed teruggevonden, enkel het plafond dat een hoogte heeft van 3.5 m wordt iets te hoog geschat.



FIGUUR 5.12: Resultaat 3D lokalisatie met opgemeten microfoonsignalen.

5.4 Besluit

Betreffende de bronlokalisatie kan besloten worden dat het AED-algoritme, zoals ook gebruikt in de simulaties, niet tot goede resultaten leidt in de praktijk. Indien het bronsignaal gekend is, en NLMS kan gebruikt worden, is het resultaat echter wel goed. Om de eerste orde TOA's te schatten, nodig voor de schatting van de kamergeometrie, blijkt de NLMS-methode opnieuw de beste methode te zijn. De computationele complexiteit is laag, in tegenstelling tot de blinde closed-form methode. Relatief lange RIR's moeten geschat kunnen worden, maar wegens de complexiteit is de closed-form methode hiervoor minder geschikt. Qua resultaat van de reflectorlokalisatie, valt te besluiten dat het verschil in de praktische resultaten ten opzichte van de simulaties grotendeels te wijten is aan de niet-omnidirectionaliteit van luidspreker en microfoons. Dit kon verholpen worden door aparte metingen uit te voeren, waarbij de bron telkens gericht is naar een andere reflector. Op deze wijze is het bekomen resultaat vrij nauwkeurig. Een tweede vaststelling is dat het resultaat verbetert indien metingen gebeuren op wanden van een hard materiaal, in tegenstelling tot absorberend materiaal zoals een stoffen gordijn. Dit is in tegenspraak met het besluit uit de simulaties, namelijk dat lagere reflectiecoëfficiënten tot betere extractie van de eerste orde TOA's uit de RIR's leiden. De verklaring hiervoor is dat een te lage reflectiecoëfficiënt als gevolg heeft dat de amplitude van de geluidsgolven horende bij paden via een reflector te laag is voor detectie.

Hoofdstuk 6

Objectdetectie en -lokalisatie (2D)

Een mogelijke uitbreiding van de techniek, is het detecteren van objecten in een ruimte. Dit kan bijvoorbeeld als bewakingstoepassing gebruikt worden. Men kan detecteren of een voorwerp of een (onbevoegd) persoon aanwezig is in de ruimte. Als er continu kan gemeten worden is het ook mogelijk om een persoon zijn pad te volgen, of een verandering in de omgeving nauwkeurig op te meten. Eventueel kan dit gecombineerd worden met akoestische bronlokalisatie, om een spreker te lokaliseren en te kunnen volgen, ook na het spreken. Het kan ook toegepast worden in computerspellen waarbij lichaamsbewegingen kunnen gebruikt worden om het spel te controleren.

6.1 Theoretisch

Een manier om te weten welke pieken in de impulsresponsies deze zijn die toebehoren aan de reflectie op het object bestaat erin eerst de impulsresponsies op te meten van de kamer zonder het object, en daarna met het object. Na aftrekking van deze twee zou de piek behorend tot de reflectie gevonden moeten kunnen worden. Er is echter wel een groot nadeel. De vorm van het voorwerp moet op voorhand gekend zijn om deze te lokaliseren met dezelfde methode als bij het lokaliseren van de muren. Omdat het niet praktisch is alle mogelijke vormen te testen, gaat hier gewoon gezocht worden naar de meest waarschijnlijke locatie van het gedetecteerde object. De aanpak die we voorstellen is de volgende. De padlengtes van bron tot microfoon via reflectie op het object, worden gebruikt voor de constructie van ellipsen, zoals ook bij de lokalisering van de kamerwanden. Er valt te verwachten dat nabij het object veel snijpunten van ellipsen met mekaar zullen vallen. De ruimte wordt opgedeeld in kleine deelruimtes, en via een stemming wordt voor elke deelruimte bepaald hoeveel snijpunten erin vallen. De deelruimte met het grootste aantal stemmen zal hoogstwaarschijnlijk de plaats van een object benaderen of raken. De snijpunten van de ellipsen worden bepaald door gebruik te maken van de Bézout-determinant, zoals besproken in [26]. Hierbij dienen enkele bemerkingen te worden gemaakt. Als de deelgebieden te

klein worden gemaakt, zullen er niet veel stemmen per deelruimte zijn, waardoor er geen onderscheid te maken valt tussen alle deelruimten, en niet beslist kan worden welke de meeste stemmen kreeg. Ze zullen dan allemaal ofwel geen ofwel één stem hebben. Worden de deelruimtes echter te groot genomen, dan is er geen nauwkeurige plaatsbepaling meer. Nog een bemerking gaat over het feit dat het object op het rechtstreeks pad tussen bron en microfoon kan staan. Dan is er geen direct pad, en ook geen pad via één reflectie op het object. Bijgevolg gaat de gevonden piek in de impulsresponsie van deze microfoon niet de correcte zijn. Ook zal er dan geen ellips gevonden worden, maar zal deze gereduceerd worden tot een rechte of andere kegelsnede. Het zal blijken uit simulaties dat als er één pad verbroken wordt, nog steeds een aanvaardbare oplossing gevonden wordt. Het simuleren van de TOA's is niet mogelijk via het vorige gebruikte model van de kamer. De TOA's worden bijgevolg rechtstreeks geschat door gebruik te maken van een benaderende formule, waarbij het object als een cilinder zal beschouwd worden. De notaties voor de locatie van microfoon i , bronlocatie en objectlocatie zijn respectievelijk de volgende:

$$r_i = \begin{bmatrix} x_i & y_i \end{bmatrix}^T, \quad i = 1, \dots, M \quad (6.1)$$

$$r_s = \begin{bmatrix} x_s & y_s \end{bmatrix}^T \quad (6.2)$$

$$r_o = \begin{bmatrix} x_o & y_o \end{bmatrix}^T \quad (6.3)$$

De schatting van de TOA tot elke microfoon via reflectie op het object wordt bepaald met volgende formule, waarin η de geluidssnelheid is, en ρ de straal van het object is:

$$TOA_i = \frac{1}{\eta} \left\{ \sqrt{(x_i - x_o)^2 + (y_i - y_o)^2} + \sqrt{(x_s - x_o)^2 + (y_s - y_o)^2} \right\} - 2 \frac{\rho}{\eta} \cos \left(\frac{1}{2} \left(\arctan \left(\frac{y_i - y_o}{x_i - x_o} \right) - \arctan \left(\frac{y_s - y_o}{x_s - x_o} \right) \right) \right) \quad (6.4)$$

Het voorgaande wordt gesimuleerd, en de resultaten ervan zijn besproken in de volgende sectie.

6.2 Simulatie en resultaten

In deze sectie worden enkele situaties getest, en de resultaten ervan besproken. De deelruimtes hebben een grootte van 20 cm op 20 cm in elke simulatie. De locaties van de vier gebruikte microfoons en de bron zijn dezelfde als in de andere experimenten en testen in voorgaande hoofdstukken, en worden nog eens samengevat in tabel 6.1. Er worden hier drie verschillende situaties besproken. In de eerste situatie wordt het pad tussen bron en microfoon niet gehinderd, zodat er voor elke microfoon een pad met reflectie is. Dit wordt besproken in simulatie 1. In een tweede situatie zal een rechtstreeks pad gehinderd worden. Deze situatie is gesimuleerd in simulatie 2. In een derde en laatste situatie worden twee paden geblokkeerd, gesimuleerd in simulatie 3.

TABEL 6.1: Locaties van bron en microfoons in 2D-ruimte (in meter).

	x	y
bron	1.1	1.9
microfoon 1	1.75	3.9
microfoon 2	2.6	1.93
microfoon 3	0.7	3
microfoon 4	0.3	1.1

6.2.1 Simulatie 1

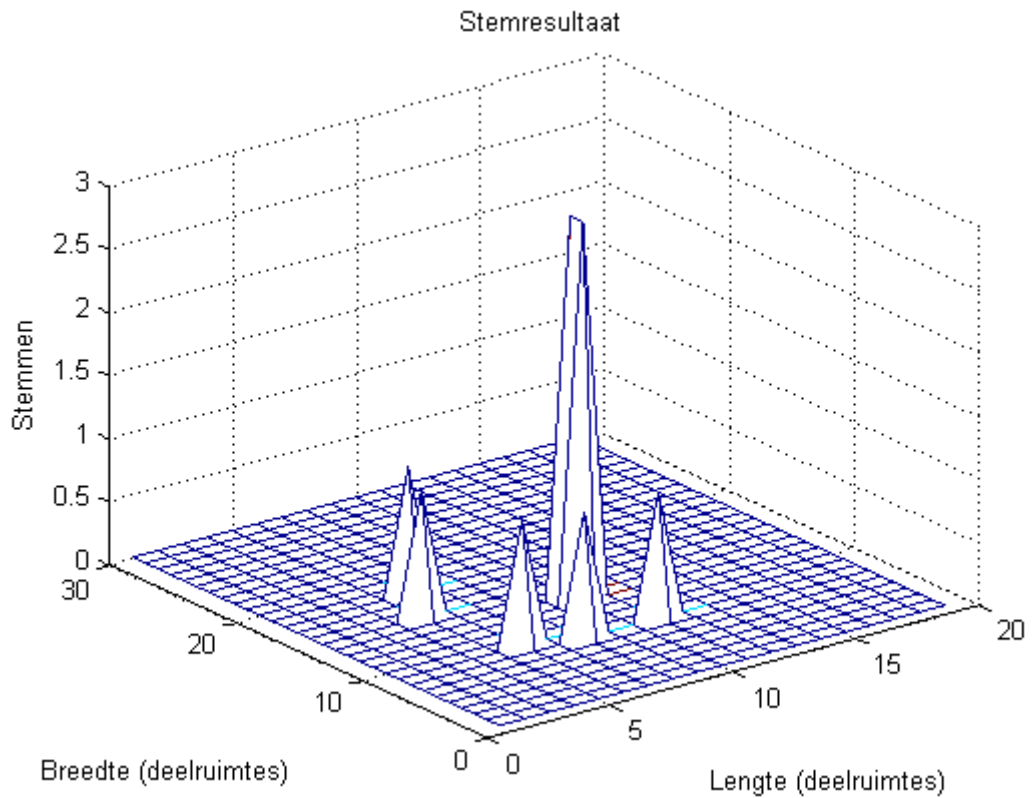
In de eerste simulatie is getracht een situatie te vinden waarbij het voorwerp geen van de directe paden tussen bron en microfoons hindert. Dit is het geval voor een voorwerp met straal van 0.1 m, op een locatie $x_o = 2$, $y_o = 2.5$. Het resultaat van de stemming is weergegeven in figuur 6.1, en de situatie in de kamer, met snijpunten, is weergegeven in figuur 6.2. De driehoek stelt de bron voor, de ster de echte locatie van het object. Het gevonden resultaat (2;2.6) komt overeen met de werkelijke locatie van het object. De afwijking is te wijten aan de grootte van de deelruimtes waarin de kamer is verdeeld. Een afwijking tot 20 cm is dus getolereerd.

6.2.2 Simulatie 2

In een tweede simulatie is ervoor gezorgd dat het object net één direct pad tussen bron en microfoon zal hinderen. Dit gaat ongetwijfeld het aantal gevonden snijpunten tussen de ellipsen verminderen en eventuele foutieve snijpunten toevoegen. Toch zal, zoals verder wordt besproken, een gunstig resultaat gevonden worden. Wanneer het voorwerp op $x_o = 2$, $y_o = 2$ geplaatst wordt, en het een straal krijgt toegewezen van 0.2 m, zal exact één pad worden geblokkeerd. De stemruimte is weergegeven in figuur 6.3, en de situatie in de kamer in figuur 6.4. (De driehoek stelt de bron voor, de ster de echte locatie van het object.) Het resultaat dat gevonden is (1.8;2.2), valt binnen de tolerantiegrenzen van ± 0.2 m, waaruit kan besloten worden dat de methode ook werkt indien er één direct pad wordt gehinderd.

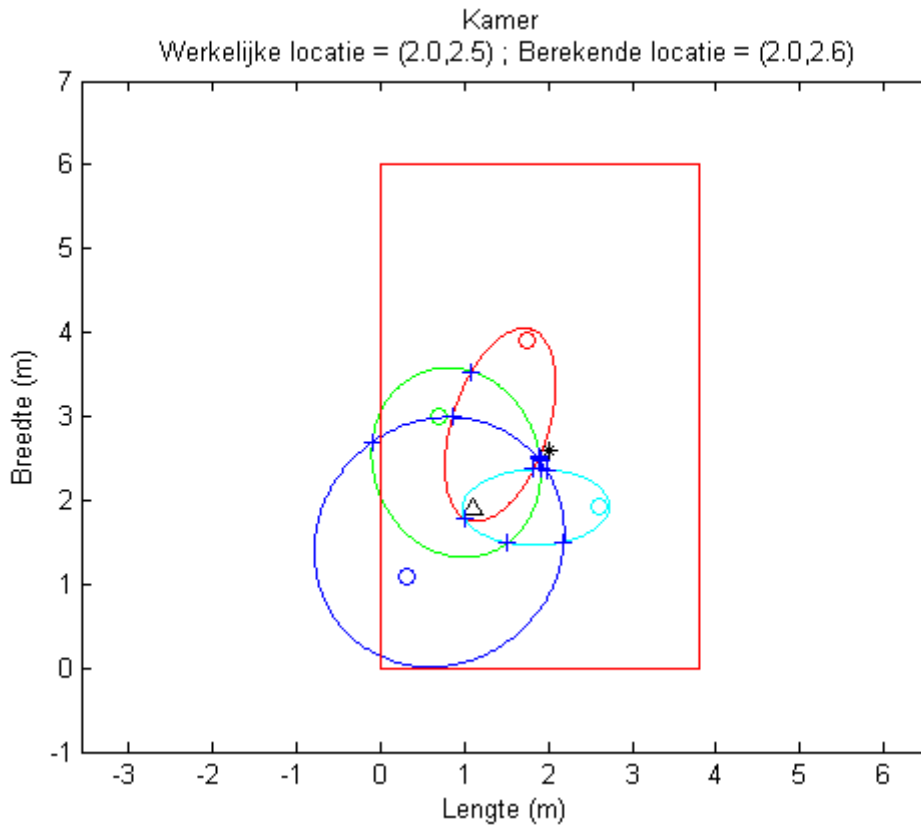
6.2.3 Simulatie 3

In een derde simulatie zijn twee directe paden gehinderd door het object, waardoor er dus geen pad zal zijn van twee microfoons naar de bron via één reflectie op het object. Dit gaat het aantal gevonden snijpunten tussen de ellipsen nog meer verminderen en eventuele foutieve snijpunten toevoegen. Het resultaat dat nu wordt gevonden is minder betrouwbaar. Deze voorwaarden worden bereikt als het voorwerp op $x_o = 1$, $y_o = 2.5$ geplaatst wordt, en het een straal toegewezen krijgt van 0.15 m. De votingruimte is weergegeven in figuur 6.5, en de situatie in de kamer in figuur 6.6. (De driehoek stelt de bron voor, de ster de echte locatie van het object.) Het resultaat dat gevonden is (1.2;2.6), valt binnen de tolerantiegrenzen van ± 0.2 m, maar wanneer de stemruimte bestudeerd wordt, is er niet echt uitsluitel te

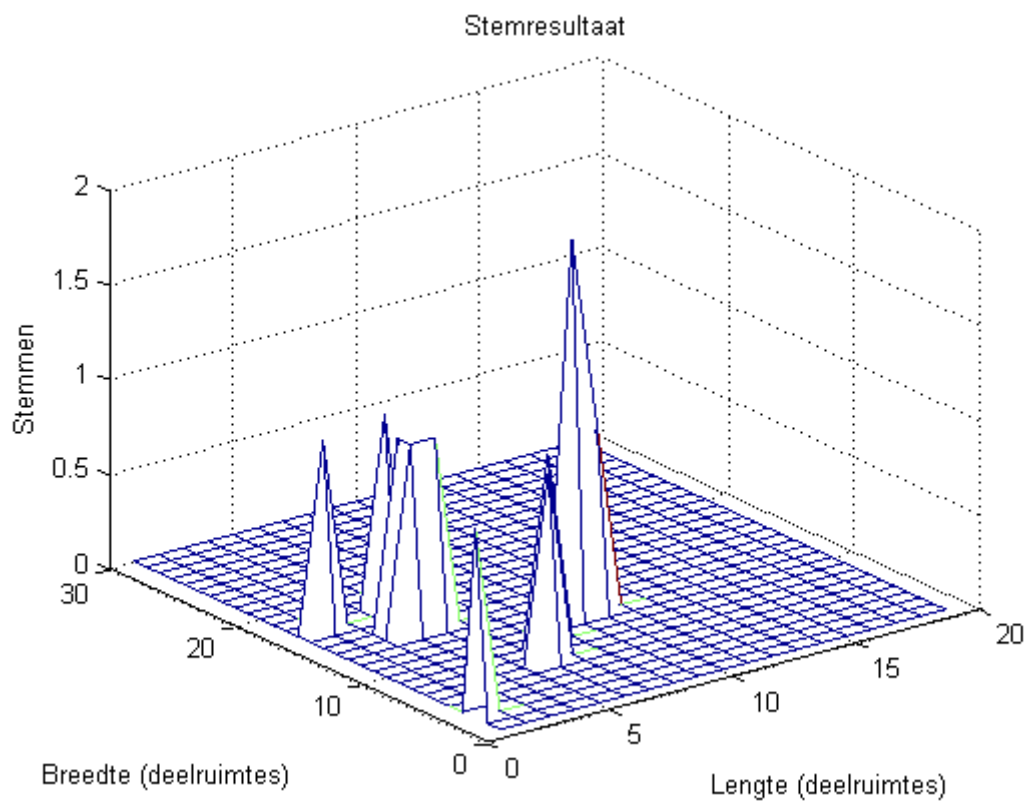


FIGUUR 6.1: Resultaat van de stemming voor simulatie 1.

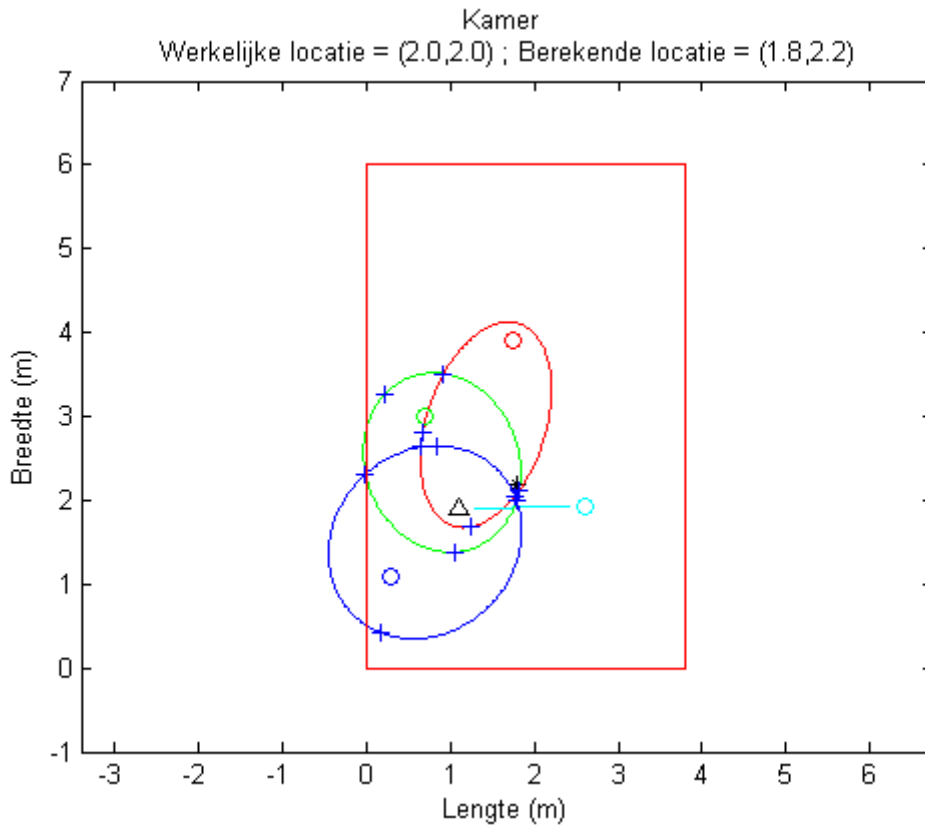
maken over welke deelruimte nu de meeste stemmen heeft gekregen. De resultaten moeten dus zeer zorgvuldig gecontroleerd worden. Inderdaad, wanneer het object 0.2 m in straal wordt gekozen, dan is de gevonden locatie (1.8;4.2) fout. De locatie volgens de y-as ligt ver buiten de tolerantiegrens. De stemming en de kamersituatie is weergegeven in [6.7](#), respectievelijk [6.8](#)



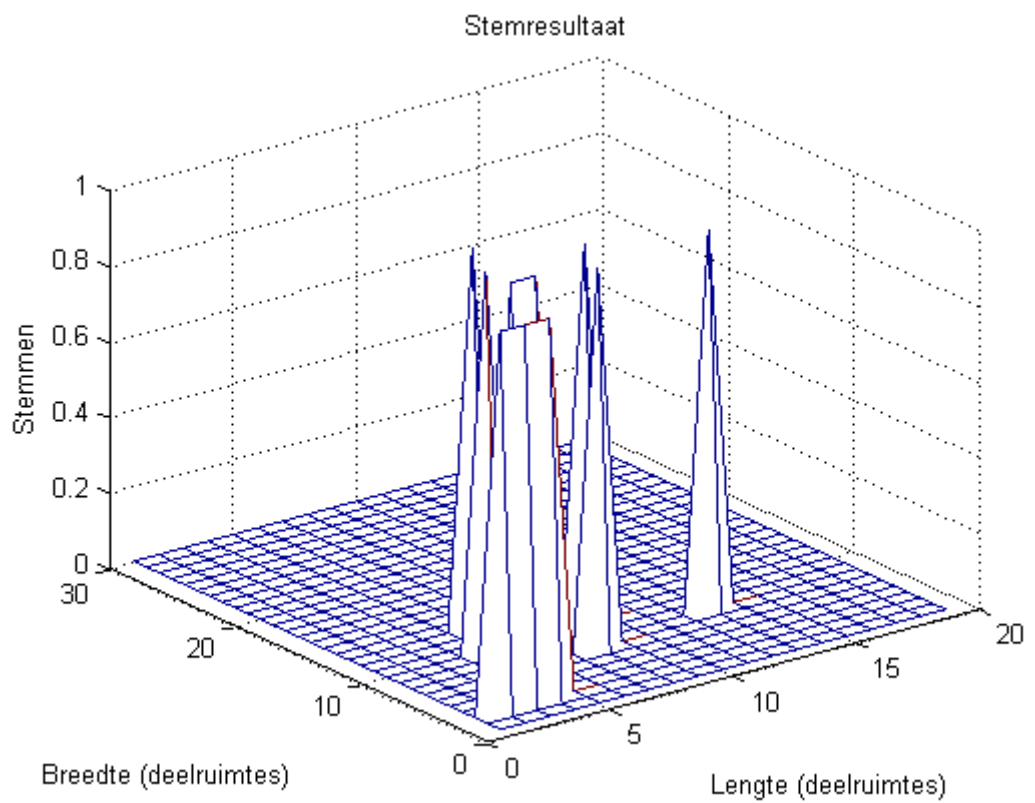
FIGUUR 6.2: Kamersituatie voor simulatie 1 (o: microfoon; Δ : bronlocatie; *: objectlocatie; +: snijpunt).



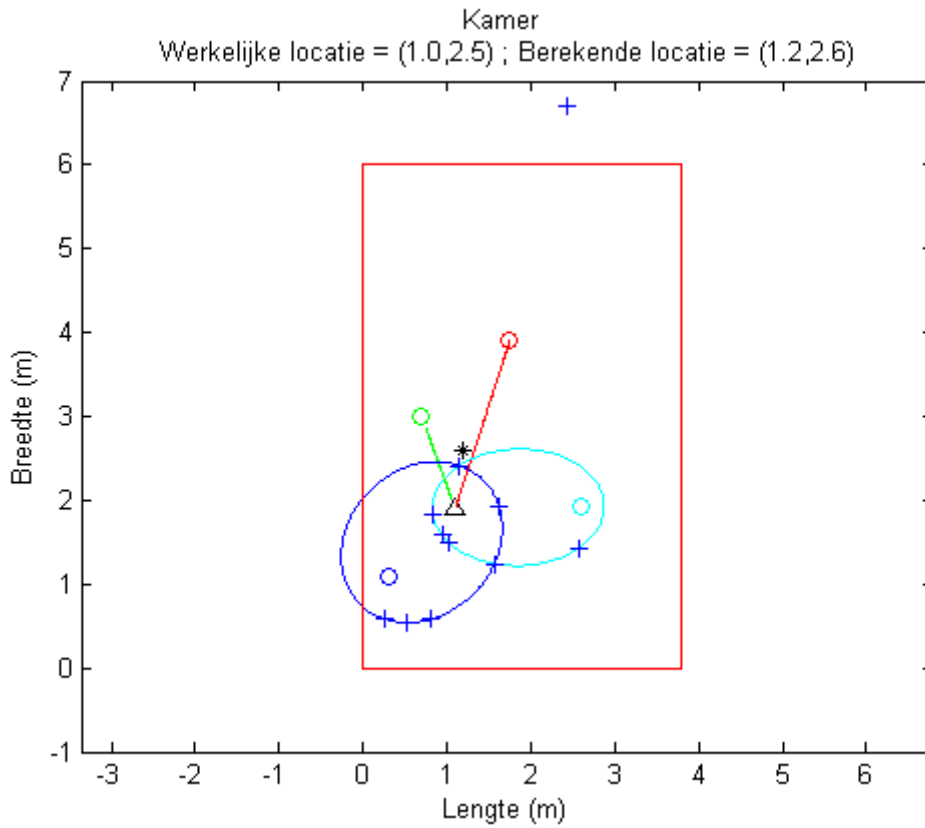
FIGUUR 6.3: Resultaat van de stemming voor simulatie 2.



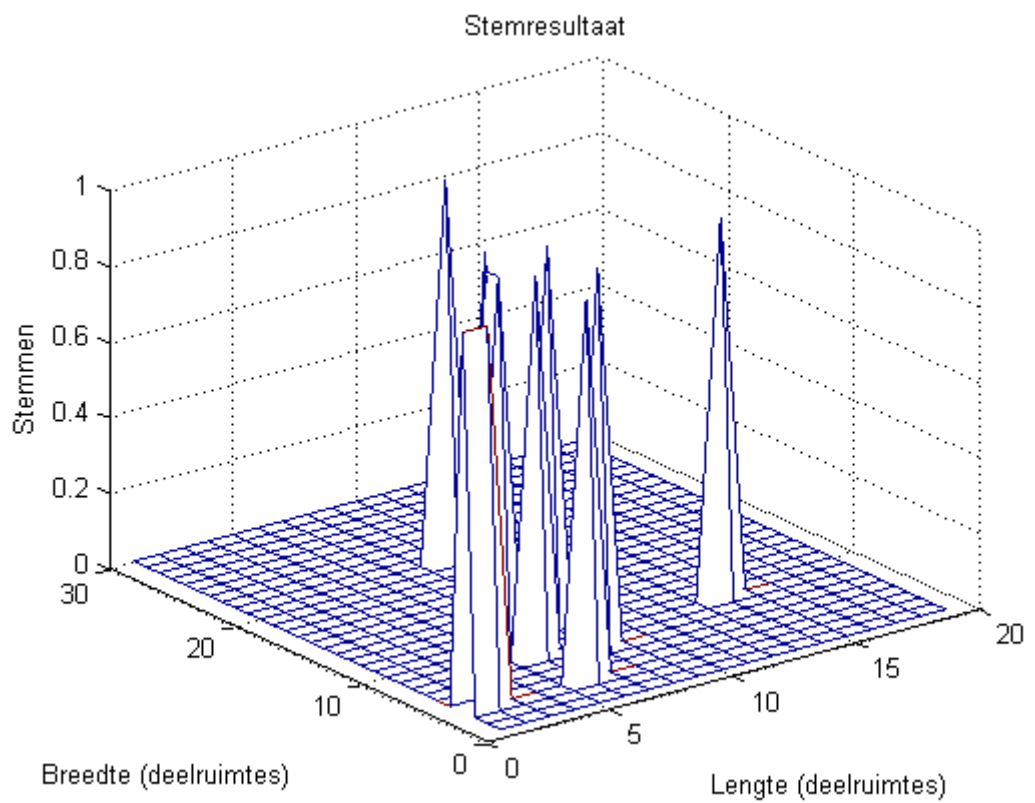
FIGUUR 6.4: Kamersituatie voor simulatie 2 (o: microfoon; Δ : bronlocatie; *: objectlocatie; +: snijpunt).



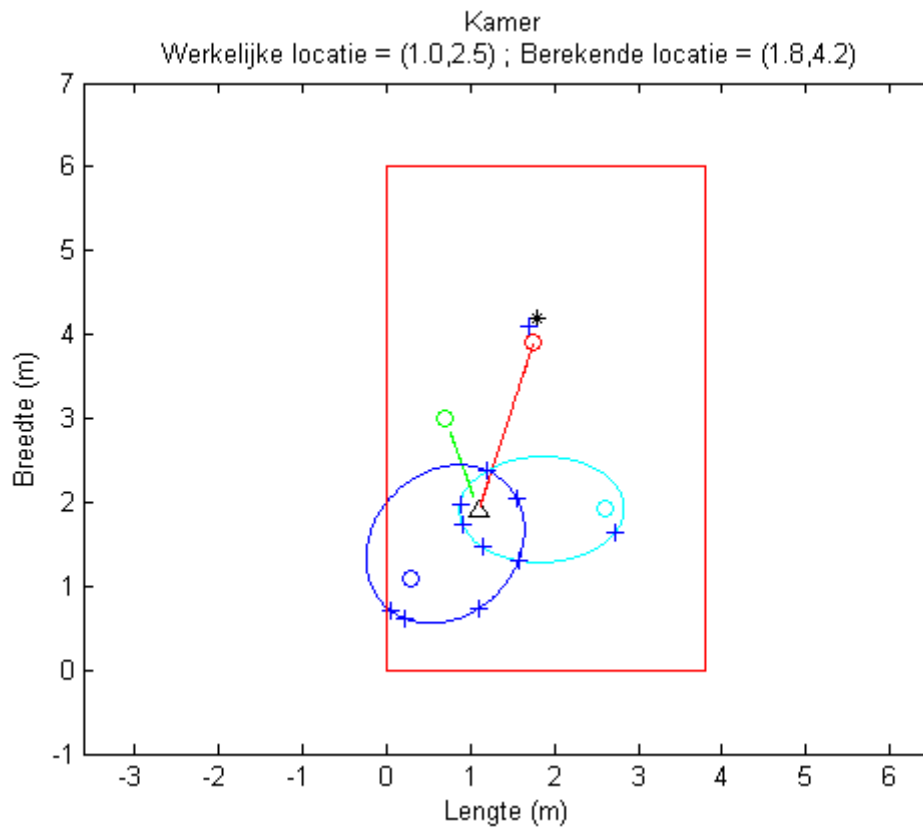
FIGUUR 6.5: Resultaat van de stemming voor simulatie 3, met objectstraal 0.15 m.



FIGUUR 6.6: Kamersituatie voor simulatie 3, met objectstraal 0.15 m (o: microfoon; Δ : bronlocatie; *: objectlocatie; +: snijpunt).



FIGUUR 6.7: Resultaat van de stemming voor simulatie 3, met objectstraal 0.2 m.



FIGUUR 6.8: Kamersituatie voor simulatie 3, met objectstraal 0.2 m (\circ : microfoon; Δ : bronlocatie; $*$: objectlocatie; $+$: snijpunt).

6.3 Besluit

Uit deze testen kunnen een aantal besluiten en bedenkingen getrokken worden. Uit de testen blijkt dat de plaatsing van de microfoons ten opzichte van het object cruciaal is. Er mogen niet teveel directe paden geblokkeerd worden. Ook is het nodig om op voorhand te weten wat de grootteorde van het object is, zodanig dat een juiste deelruimtegrootte gekozen kan worden. Als het object veel groter is dan de deelruimten, zullen ook de stemmen meer verspreid liggen, en niet gegroepeerd in een deelruimte, of deelruimtegroep. Dit zijn nadelen van deze methode. Willen we een exacte positie van het object, dan zal ook de exacte vorm en grootte gekend moeten zijn van het voorwerp, om zo een zelfde methode te gebruiken als bij het zoeken van de locatie van de kamerwanden. Dit gaat dan heel rekenintensief zijn, en te lang duren voor real-time applicaties zoals computerspellen. Voor real-time is de in dit hoofdstuk besproken methode echter beter geschikt, daar deze een lage computationele complexiteit heeft. Om de nauwkeurigheid te verbeteren, kunnen er bijvoorbeeld metingen met verschillende groottes van deelruimten achtereenvolgens gebeuren, en daaruit een gemiddelde genomen worden. Een andere mogelijke uitbreiding, is de spreiding van de snijpunten van de ellipsen in rekening te brengen, en niet enkel het aantal snijpunten in een deelruimte. Op deze wijze kan men steeds met dezelfde grootte van deelruimtes werken, een groter object kan dan herkend worden als een gebied van deelruimtes met hoge dichtheid aan snijpunten. Op deze wijze kan zowel de locatie als de grootte van het object geschat worden, en hoeft de grootte van het te detecteren object niet a priori gekend te zijn.

Hoofdstuk 7

Besluit

De eerste stap in het probleem is het bepalen van de propagatietijden van de geluidsgolven, teneinde de bronlokalisatie en geometrische reconstructie uit te voeren. Als besluit hierover kan gezegd worden dat indien het bronsignaal gekend is, de benodigde RIR's voor elk pad van bron tot microfoon met lage computationele complexiteit geschat kunnen worden via NLMS. Indien het bronsignaal ongekend is, treden er problemen op qua complexiteit bij gebruik van het closed-form algoritme zoals beschreven in hoofdstuk 2. De grootste reden hiervoor is de nood aan het schatten van relatief lange impulsresponsies. Een andere mogelijkheid, het AED-algoritme, levert wel goede resultaten voor schattingen van aankomsttijden via het directe pad in de simulaties maar niet in de praktijk. In theorie zou het AED-algoritme zelfs voor de schatting van de eerste orde reflecties bruikbaar zijn, maar in praktijk blijkt deze dus voor geen van beide bruikbaar. Het besluit voor de kruiscorrelatiemethode is dat deze goede resultaten levert voor de bronlokalisatie, maar niet voor de omgevingsreconstructie. Als conclusie volgt hieruit dat de voorgestelde methode beter toepasbaar is in situaties waarin het bronsignaal gecontroleerd wordt. Voorbeelden hiervan zijn 3D-audio, echo-cancellatie, videospellen, en andere toepassingen waarin kennis over de kamergeometrie voor betere resultaten kan zorgen. Het voordeel van de methode die in deze thesis wordt besproken, is dat er geen restrictie wordt opgelegd op de posities van de microfoons. Enkel de relatieve positie moet gekend zijn. Ook op het bronsignaal worden in eerste instantie geen restricties gelegd, hoewel uit de praktijkresultaten blijkt dat de methode beter werkt indien een signaal met lage frequentie-inhoud gebruikt wordt. Ook blijkt uit deze praktijkresultaten, dat de geluidsbron en microfoons zo goed mogelijk een omnidirectionele karakteristiek moeten bezitten.

De voorgestelde methode voor akoestische omgevingsreconstructie werkt in twee stappen. De eerste taak is het lokaliseren van de bron, daarna kan aan de hand hiervan en van de eerste orde TOA's de geometrie bepaald worden. De bronlokalisatie gebeurt via het NLMS-algoritme, de resultaten hiervan zijn gunstig voor het 2D geval. Deze vertrekt van een schatting van de aankomsttijden van de golven via het directe pad van bron tot microfoon, deze kunnen vrij goed geëxtraheerd worden uit de geschatte RIR's. Bij uitbreiding naar 3D treden er vaken afwijkingen op

de gevonden bronlocaties. Om de geometrie te kunnen schatten, moeten ook de eerste orde TOA's gekend zijn. De resultaten van de schatting hiervan zijn minder nauwkeurig als deze van het directe pad. Dit omdat deze niet altijd even prominent te zien zijn in de RIR's. Dit is ook afhankelijk van de reflectiecoëfficiënten van de kamerwanden. Een te hoge reflectiecoëfficiënt leidt tot een lange reverberatietijd, en moeilijker onderscheid tussen eerste orde reflecties en hogere orde reflecties. Dit probleem blijkt niet zo cruciaal te zijn, aangezien simulaties met een reflectiecoëfficiënt van 0.9 nog steeds goede resultaten opleveren. Indien deze echter te laag is, zijn de amplitudes van de gereflecteerde golven te laag om goede detectie te bekomen. In de praktische resultaten bleek ook dat de resultaten beter zijn met wanden van een hard materiaal, en minder met meer absorberende materialen zoals een gordijn.

Vandaar dat een aanpassing voorgesteld werd aan de methode waarop de reflectoren bepaald worden, zodat deze robuuster is tegen fouten op de TOA's. Deze worden bepaald als raaklijnen aan ellipsen die de meetkundige plaats beschrijven van elk mogelijk reflectiepunt, indien men een bepaald pad beschouwd. De kostenfunctie aan deze raaklijn is cruciaal voor een goede werking van het algoritme, het is hier dat de aanpassing gebeurt zoals beschreven in hoofdstuk 4. In het driedimensionale geval wordt een gelijkaardige methode toegepast, waarbij nu raakvlakken aan ellipsoïden worden gezocht. De resultaten zijn behoorlijk maar minder gunstig als voor het tweedimensionale geval. Ook is de benodigde rekentijd significant groter. Factoren die hierin een invloed hebben zijn het aantal microfoons in de kamer, het aantal te zoeken reflectoren en de nauwkeurigheid waarmee de reflectoren gezocht moeten worden. De afhankelijkheid van aantal microfoons en reflectoren is exponentieel via de relatie N^M , met N het aantal reflectoren en M het aantal microfoons. Tenslotte kan men nog besluiten dat de grootte van de kamer geen grote invloed heeft op de correctheid van de geometrische reconstructie.

Naast het schatten van een kamergeometrie werd ook een nieuwe methode voorgesteld om op akoestische wijze een object of persoon in de kamer te detecteren. Uit testen van de methode blijkt dat de plaatsing van de microfoons ten opzichte van het object belangrijk is. Blokkage van directe paden tussen bron en microfoons beïnvloedt het resultaat negatief. Ook werkt de methode beter indien op voorhand een estimatie gemaakt kan worden van de grootte van het te lokaliseren object. Voordelen van de methode zijn dat er opnieuw enkel gebruik gemaakt wordt van losse microfoons doorheen de kamer, en een willekeurig bronsignaal. Toepassingen van deze methode kunnen o.a. bewakingstoepassingen zijn, controle bij computerspellen of traceren van een persoon.

Bibliografie

- [1] Sonar. [Online]. Available: <http://en.wikipedia.org/wiki/Sonar>
- [2] J. L. Sutton, "Underwater acoustic imaging," *Proc. IEEE*, vol. 67, pp. 554–566, Apr. 1979.
- [3] N. Brahim, D. Guériot, S. Daniel, and B. Solaiman, "3D reconstruction of underwater scenes using image sequences from acoustic camera," in *IEEE OCEANS 2010*, May 2010.
- [4] D. Ba, F. Ribeiro, C. Zhang, and D. Florêncio, "L1 regularized room modeling with compact microphone arrays," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2010.
- [5] A. M. O'Donovan, D. N. Zotkin, and R. Duraiswami, "Spherical microphone array based immersive audio scene rendering," in *Proceedings of the 14th International Conference on Auditory Display*, Jun. 2008.
- [6] acoustic camera. [Online]. Available: <http://www.acoustic-camera.com>
- [7] A. O'Donovan, R. Duraiswami, and D. Zotkin, "Imaging concert hall acoustics using visual and audio cameras," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2008.
- [8] F. Antonacci, A. Sarti, and S. Tubaro, "Geometric reconstruction of the environment from its response to multiple acoustic emissions," in *Acoustics Speech and Signal Processing (ICASSP)*, Mar. 2010.
- [9] I. Dokmanić, Y. M. Lu, and M. Vetterli, "Can one hear the shape of a room: The 2D-polygonal case," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011.
- [10] W. G. Gardner, "3D audio and acoustic environment modeling," Wave Arts inc., 99 Massachusetts Avenue, Suite 7, Arlington, MA 02474, Tech. Rep., 1999.
- [11] F. Ribeiro, D. Ba, C. Zhang, and D. Florêncio, "Turning enemies into friends: Using reflections to improve sound source localization," in *IEEE International Conference on Multimedia and Expo (ICME)*, Jul. 2010.
- [12] Xbox - kinect. [Online]. Available: <http://www.xbox.com/nl-NL/Kinect/Home>

- [13] A. Canclini, P. Annibale, F. Antonacci, A. Sarti, R. Rabenstein, and S. Tubaro, "From direction of arrival estimates to localization of planar reflectors in a two dimensional geometry," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Jan. 2011.
- [14] E. Mabande, H. Sun, K. Kowalczyk, and W. Kellermann, "On 2D localization of reflectors using robust beamforming techniques," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011.
- [15] S. S. Haykin and B. Widrow, *Least-Mean-Square Adaptive Filters*. New York, USA: Wiley, 2003.
- [16] J. Chen, J. Benesty, and Y. Huang, "Time delay estimation in room acoustic environments: An overview," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. 1–19, 2006.
- [17] G. Xu, H. Lui, L. Tong, and T. Kailath, "A least squares approach to blind channel identification," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 43, pp. 2982–2993, Dec. 1995.
- [18] P. A. Naylor and N. D. Gaubitch, *Speech Dereverberation*, 1st ed. London, GB: Springer Verlag, 2010, pp. 129–143.
- [19] J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization," *Acoustical Society of America*, vol. 107, pp. 384–391, Jan. 2000.
- [20] J. Filos, E. A. Habets, and P. A. Naylor, "A two-step approach to blindly infer room geometries," in *Proc. of the International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Aug. 2010.
- [21] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 1661–1669, Dec. 1987.
- [22] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *JASA*, vol. 65, pp. 943–950, Apr. 1979.
- [23] Acoustics: The speed and attenuation of sound. [Online]. Available: http://www.kayelaby.npl.co.uk/general_physics/2_4/2_4_1.html
- [24] Absorption coefficient chart. [Online]. Available: http://www.sae.edu/reference_material/pages/CoefficientChart.htm
- [25] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, ISBN: 0521540518, 2004.
- [26] D. Abermy, "Intersection of ellipses," Geometric Tools, LCC, <http://www.geometrictools.com/>, Tech. Rep., 2010.

Fiche masterproef

Studenten: Brecht Milis
Jelle Vanolst

Titel: Omgevingsreconstructie met akoestische signalen

Engelse titel: Reconstruction of environment with acoustic signals

UDC: 621.3

Korte inhoud:

Het probleem van het reconstrueren van een omgeving via akoestische signalen krijgt recentelijk hernieuwde aandacht. In deze thesis wordt een manier voorgesteld om de geometrie van een kamer te schatten, in twee stappen. Een kamer met reflecterende wanden bevat een geluidsbron en enkele microfoons, verspreid doorheen deze kamer. Enkel de relatieve posities van de microfoons zijn gekend. De geluidsbron kan algemeen worden beschouwd, zoals een persoon die aan het praten is. Het kan ook een gecontroleerd bronsignaal zijn. Eerst worden de propagatietijden van de geluidsgolven bepaald. Hieruit kan de bronlocatie worden bepaald en ook de geometrie van de kamer. Dit wordt zowel uitgewerkt en getest in twee dimensies als in drie dimensies. Als extra toepassing, wordt een methode voorgesteld om op akoestische wijze een object te detecteren en te lokaliseren in de kamer.

Thesis voorgedragen tot het behalen van de graad van Master in de ingenieurswetenschappen: elektrotechniek, optie Multimedia en signaalverwerking

Promotor: Prof. dr. ir. Marc Moonen

Assessoren: Prof. dr. ir. Patrick Wambacq
dr. ir. Alexander Bertrand

Begeleider: ir. Bram Cornelis

Reconstruction of an Environment using Acoustic Signals

Jelle Vanolst, Brecht Milis, Marc Moonen, Bram Cornelis
Electrical Engineering dept. KULeuven
Leuven, Belgium

Abstract—The problem of reconstructing an environment is gaining more interest lately. This paper will present a way to infer the geometry of a room, in two major steps. The room with reflective walls, contains some microphones, of which the relative positions are known. An input to the system can be anything, say a speaker in the room. Also the use of a controlled source signal is considered. First, the propagation times of the sound need to be estimated. Both the direct-path as the paths of first-order reflections are of interest. The ways on how to find these will be presented here as well. Using the propagation times of the direct-path, the source of the sound can be localized. The propagation times of the paths of the first-order reflections can be used in a second step to locate the walls of the rooms. Both in 2-D and in 3-D, these theories will be examined and tested. As an extra application, a possible approach for acoustic object detection is proposed.

I. INTRODUCTION

A first use of acoustic signals to reconstruct an environment can be found in underwater applications. A sonar can be used as underwater camera or object detector [1], [2]. An accurate 3-D reconstruction of a seabed is also possible [3]. Above the water surface, it's not straightforward to apply the same techniques. Firstly, there is a difference in propagation between sound waves in water and in air. Secondly, when working in a closed environment, there will be some reverberation. Hence other solutions are needed in this case.

Techniques for reconstruction of an environment using acoustic signals can be categorized in two classes. The first class uses microphone arrays and beamforming techniques. Research has been done in 2-D and 3-D reconstruction using a small circular array [4]. Also spherical arrays can be used, called audio cameras [5], [6]. These can be used to retrieve a map of acoustic intensity, to determine the most reflecting surfaces in a 3-D environment. A possible application is measuring the acoustic performance of a concert hall [7]. The second class of techniques does not use microphone arrays, but makes use of single microphones, spread out inside a room. This method is quite new, and most publications about this subject are limited to 2-D cases. Similar to multiple microphones, it's also possible to use one moving microphone [8]. Recent research shows that it's even possible to reconstruct the shape of a room, using only one microphone and one source [9]. By emitting and recording a short sound pulse at the same location, the Room Impulse Response (RIR) can be extracted and the shape of the room can be estimated. In this way, one can hear the shape of a room.

For the source, several assumptions can be made. The location and broadcasted signal of the source could be fully known, only the signal could be known or both could be unknown. All possibilities will be examined, and each will have its own applications. If the source signal is unknown, then a blind channel estimation is necessary.

The major goal is to infer the geometry of the room, by searching the positions of the walls, floor and ceiling. This leads to a shoebox model of the room. We are also interested in detecting and locating objects in a room. A new method will be proposed for acoustic object detection, results will be described using simulations.

Information about the geometry of the room can be employed in several audio applications. It can simplify the estimation of Room Impulse Responses (RIRs). For example in 3-D audio applications, it's necessary to know the RIR from a speaker to a spectator [10]. Instead of estimating the RIR using recorded microphone signals, it can be generated by a model, using the room dimensions. Similarly, dereverberation effects can be simulated, which leads to better results in algorithms for source localization and echo cancellation [11]. An application that makes use of this is the Kinect from Xbox. This device registers movements in 3-D, and provides gameplay without the use of extra controllers [12]. Voice control is one of the features of this device. The Kinect uses a microphone array for acoustic noise suppression and echo cancellation for retrieving a good speech signal. Information on how reflected sound waves propagate through the room, can help in separating speech commands from background noise and echos.

II. ESTIMATING THE RIRS

For estimating the Room Impulse Responses, four methods will be proposed. They all have different advantages and disadvantages. First, NLMS will be discussed. In this algorithm, the source signal needs to be known, or generated by the system itself. Then, the blind estimation algorithms are considered. In these, the source signal is unknown, and the impulse responses are estimated using only the received microphone signals. Possible algorithms are the cross correlation algorithm, closed form estimation and Adaptive Eigenvalue Decomposition (AED).

A. NLMS

In this algorithm, the source signal has to be known. It will be given as an input to the algorithm. The recorded signals

are considered as the desired signals, because the goal is to recreate them by filtering the source signal. In that way, the filter coefficients will be the impulse response. This happens in an iterative way. The update formula, with μ the stepsize, is the following:

$$\mathbf{w}(F+1) = \mathbf{w}(F) + \mu \cdot \mathbf{u}_{F+1} \cdot (x_{F+1} - \mathbf{u}_{F+1}^T \mathbf{w}(F)) \quad (1)$$

with

$$\mu = \frac{\bar{\mu}}{\kappa + \mathbf{u}_{F+1}^T \mathbf{u}_{F+1}} \quad (2)$$

where \mathbf{w} are the filter coefficients, and \mathbf{u} contains the input samples on the delay line of the filter. x is a sample of the desired signal, and F is the iteration step. An all-zero vector is used as start vector for \mathbf{w} , $\bar{\mu}$ and κ have to be chosen in function of the signals. The number of iterations is determined by the length of the microphone signal.

B. Cross Correlation

The principle of this algorithm is very simple. It can be used to calculate the Time Difference Of Arrival (TDOA) between two microphone signals, in a certain possible delay range. One signal is shifted over the other, and then the similarity between both is evaluated. The TDOA is the delay for which the correlation is maximized. The algorithm performs well in finding the TDOA between the direct-paths, but not in finding the TOAs of the paths along first-order reflections. Using the direct-path TDOAs, the source can be localized.

C. Closed Form

A second method to estimate the RIRs blindly, is by using a least squares method, as described in [13]. In this case, one tries to estimate all impulse responses at once, using a closed expression. The principal on which the algorithm is based, is:

$$\mathbf{x}_1^T(n) \mathbf{g}_2 = \mathbf{x}_2^T(n) \mathbf{g}_1 \quad (3)$$

in which \mathbf{x}_1 and \mathbf{x}_2 are vectors from length L , the estimated length of the impulse responses. They contain samples of two microphone signals. \mathbf{g}_1 and \mathbf{g}_2 are the accompanying impulse responses. The equation is valid since the source signal originates from the same source. With the correlation matrix

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{x_1 x_1} & \mathbf{R}_{x_1 x_2} \\ \mathbf{R}_{x_2 x_1} & \mathbf{R}_{x_2 x_2} \end{bmatrix}, \text{ met } \mathbf{R}_{x_i x_j} = E\{\mathbf{x}_i(n) \mathbf{x}_j^T(n)\} \quad (4)$$

and by defining \mathbf{h} , (3) can be applied:

$$\mathbf{h} = \begin{bmatrix} \mathbf{g}_2 \\ -\mathbf{g}_1 \end{bmatrix} \Rightarrow \mathbf{R} \mathbf{h} = 0 \quad (5)$$

Since these equations are valid for each pair of microphones i and j , and each set of K samples, the next set of equations can be made, in which the expected values will be replaced by the microphone signals:

$$[\mathbf{X}_i(L) \quad | \quad -\mathbf{X}_j(L)] \begin{bmatrix} \mathbf{g}_i \\ \mathbf{g}_j \end{bmatrix} = 0 \quad (6)$$

The matrix $\mathbf{X}_m(L')$ is given by (with $L' = L - 1$):

$$\begin{bmatrix} x_m(L') & x_m(L'+1) & \cdots & x_m(2L') \\ x_m(L'+1) & x_m(L'+2) & \cdots & x_m(2L'+1) \\ \vdots & \vdots & \ddots & \vdots \\ x_m(K-L') & x_m(K-L'+1) & \cdots & x_m(K) \end{bmatrix} \quad (7)$$

To estimate all the RIRs at once, the next matrices are defined

$$\mathbf{X}^i(L') = \quad (8)$$

$$\begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{X}_{i+1}(L') & -\mathbf{X}_i(L') & \mathbf{0} & \mathbf{0} \\ \vdots & & & \vdots & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{X}_M(L') & \mathbf{0} & \cdots & -\mathbf{X}_i(L') \end{bmatrix}$$

(which is M blocks wide, and $M - 1$ blocks high)

$$\mathbf{X}(L') = \left[\begin{array}{c} \mathbf{X}^1(L') \\ \vdots \\ \mathbf{X}^{M-1}(L') \end{array} \right] \left. \vphantom{\begin{array}{c} \mathbf{X}^1(L') \\ \vdots \\ \mathbf{X}^{M-1}(L') \end{array}} \right\} \frac{M(M-1)}{2} \text{ blocks} \quad (9)$$

to get the final equation, according to (5)

$$\mathbf{X}(L') \mathbf{G} = \mathbf{0}, \quad \text{met } \mathbf{G} = [\mathbf{g}_1^T, \cdots, \mathbf{g}_M^T]^T \quad (10)$$

M is the number of microphones, with i from 1 to M . The advantage over the cross correlation method, is that this algorithm also takes reverberation into account. This makes the algorithm also useful for room geometric estimation, as will be described later in the text. The problem here, is that the algorithm needs a lot of computational power, because large matrices are used. Also the estimation of the length of the impulse response needs to be precise, in order to find exactly one vector in the nullspace of $\mathbf{X}(L')$.

D. Adaptive Eigenvalue Decomposition

This algorithm calculates the impulse responses per couple of microphones. It's based on the same principle as the closed form method, but evaluates the expression in an adaptive way. According to the literature [14], this algorithm also estimates the first and higher order reflections like the closed form method. The update formula is the following:

$$\mathbf{w}(F+1) = \frac{\mathbf{w}(F) - \mu * e(F) * \mathbf{x}(F)}{\|\mathbf{w}(F) - \mu * e(F) * \mathbf{x}(F)\|} \quad (11)$$

in which

$$e(F) = \mathbf{w}^T(F) \mathbf{x}(F) \quad (12)$$

The startvector for \mathbf{w} is an all-zero vector of length L , with only the value in the middle set to 1. The stepsize μ also needs to be chosen carefully.

III. LOCATING THE SOURCE

After having the TOAs or TDOAs belonging to the direct path extracted, the source can be localized. In this section, an approach in 3-D will be given. The 2-D case is easy to extract from the theory here, and can also be found in [15]. The 3-D version is based on [16]. When the distance or TOA between the microphone and source is known, spheres

centered on the microphone positions can be constructed, defining all possible locations of the source. In the case that only the TDOAs are known, a workaround has to be made. Consider M microphones in the 3-D space, on locations:

$$r_i = [x_i \ y_i \ z_i]^T, \quad i = 1, \dots, M \quad (13)$$

and a reference microphone in the origin

$$r_1 = [0 \ 0 \ 0]^T \quad (14)$$

The location of the source is unknown and defined as

$$r_s = [x_s \ y_s \ z_s]^T \quad (15)$$

The distances from the i -th microphone R_i to the source and from the source to the origin R_s are defined as:

$$R_i = \|r_i\| = \sqrt{x_i^2 + y_i^2 + z_i^2}, \quad i = 2, \dots, M \quad (16)$$

$$R_s = \|r_s\| = \sqrt{x_s^2 + y_s^2 + z_s^2} \quad (17)$$

Using η (speed of sound), the range-difference $d_{i,j}$ can be related to the TDOA of the direct path $\Delta_{i,j}$ between microphone i and j

$$d_{i,j} = \eta \Delta_{i,j} \quad (18)$$

The source will be located on the intersection of the spheres as defined above. An error function will be constructed to find the closest point to the spheres and hence minimizing the error in a least squares method. In following equations, measurements will be indicated by $(\hat{\cdot})$. One can write for both the measured and the real distances:

$$\hat{D}_i = R_s + \hat{d}_{i,1} \quad (19)$$

and

$$D_i^2 = R_i^2 + R_s^2 - 2R_i R_s \frac{r_i^T r_s}{R_i R_s} \quad (20)$$

$$D_i^2 = R_i^2 + R_s^2 - 2r_i^T r_s \quad (21)$$

The error function then becomes $\epsilon_i = \frac{1}{2}(\hat{D}_i^2 - D_i^2)$

$$\epsilon_i = r_i^T r_s + \hat{d}_{i,1} R_s - \frac{1}{2} R_i^2 + \frac{1}{2} \hat{d}_{i,1}^2, \quad i = 2, \dots, M \quad (22)$$

In matrix notation, this gives for all $M-1$ errors together:

$$\mathbf{e}(r_s) = \mathbf{A}\Theta - \mathbf{b} \quad (23)$$

$$\text{in which } \mathbf{A} = [\mathbf{S} \mid \hat{\mathbf{d}}], \quad \mathbf{S} = \begin{bmatrix} x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \\ \vdots & \vdots & \vdots \\ x_M & y_M & z_M \end{bmatrix}$$

$$\Theta = \begin{bmatrix} x_s \\ y_s \\ z_s \\ R_s \end{bmatrix}, \quad \mathbf{b} = \frac{1}{2} \begin{bmatrix} R_2^2 - \hat{d}_{2,1}^2 \\ R_3^2 - \hat{d}_{3,1}^2 \\ \vdots \\ R_M^2 - \hat{d}_{M,1}^2 \end{bmatrix}$$

$$\hat{\mathbf{d}}^T = [\hat{d}_{2,1} \ \hat{d}_{3,1} \ \dots \ \hat{d}_{M,1}]$$

Minimizing the least squares criterion ($J = \mathbf{e}^T \mathbf{e}$) yields:

$$\hat{\Theta} = \mathbf{A}^+ \mathbf{b} \quad (24)$$

where $(\cdot)^+$ is a pseudo inverse.

IV. RECONSTRUCTING THE ROOM GEOMETRY

This section describes the methodology to locate the walls and ceiling. Only the 3-D method will be explained, since again, the 2-D version is easy to extract from this method. The method does not make use of any microphone arrays, but only uses single microphones, located randomly in the room. Only relative positions from one microphone to another need to be known. The method used is an expansion of the 2-D version in [15]. Another approach uses the first and second order reflections [9], but this is not handled here.

The method described here uses two steps. First, using simple geometry, it can be seen that if the length of a path from source to microphone via one reflection is known, then all the possible reflection points lie on an ellipsoid with its foci positioned on the source and the microphone location. The first step is to define those ellipsoids. In a second step, the common tangents of these ellipsoids will be calculated. The tangential planes should model the searched room walls. Grouping the correct ellipsoids together is a critical step, which will be discussed too.

A. Ellipsoid definition

Any three dimensional surface can be written as:

$$\zeta = \{(x, y, z) \in \mathbb{R}^3 \mid ax^2 + 2bxy + 2cxz + 2dyz + ey^2 + fz^2 + 2gx + 2hy + 2iz + j = 0\} \quad (25)$$

which gives in matrix notation, using homogenous coordinates:

$$\mathbf{v}^T \mathbf{C} \mathbf{v} = 0 \quad (26)$$

with

$$\mathbf{v} = [x \ y \ z \ 1]^T \quad (27)$$

and

$$\mathbf{C} = \begin{bmatrix} a & b & c & g \\ b & e & d & h \\ c & d & f & i \\ g & h & i & j \end{bmatrix} \quad (28)$$

For the calculation of the parameters of the ellipsoids, we start from a unit sphere

$$\mathbf{C}_I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (29)$$

which is transformed into an ellipsoid. First, the sphere is morphed into a rugby ball shape, then rotated, and finally translated to the correct position, making sure that the foci correspond to the microphone and the source position. The scaling matrix to morph the circle is:

$$\mathbf{S}_{i,k} = \begin{bmatrix} Q_{i,k}^{\text{maj}} & 0 & 0 & 0 \\ 0 & Q_{i,k}^{\text{min}} & 0 & 0 \\ 0 & 0 & Q_{i,k}^{\text{min}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (30)$$

with

$$Q_{i,k}^{\text{maj}} = \frac{\eta \cdot \tau_{i,k}}{2}, Q_{i,k}^{\text{min}} = \frac{\sqrt{(\eta \cdot \tau_{i,k})^2 - D_i^2}}{2} \quad (31)$$

These Q s are the lengths of the semi-axes of the ellipsoid. η is the speed of sound, and $\tau_{i,k}$ is the TOA corresponding to microphone i and reflector k . The rotation matrix per microphone is the following:

$$\mathbf{R}_i = \begin{bmatrix} \cos \alpha_i \cos \beta_i & -\sin \alpha_i & -\cos \alpha_i \sin \beta_i & 0 \\ \sin \alpha_i \cos \beta_i & \cos \alpha_i & -\sin \alpha_i \sin \beta_i & 0 \\ \sin \beta_i & 0 & \cos \beta_i & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (32)$$

which corresponds to a positive rotation around the z-axis over α_i , followed by a negative rotation around the rotated y-axis over β_i . The following relationships hold:

$$\tan \alpha_i = \frac{y_s - y_i}{x_s - x_i}, \tan \beta_i = \frac{z_s - z_i}{\sqrt{(x_s - x_i)^2 + (y_s - y_i)^2}} \quad (33)$$

The translation matrix is defined as

$$\mathbf{T}_i = \begin{bmatrix} 1 & 0 & 0 & \Delta x_i \\ 0 & 1 & 0 & \Delta y_i \\ 0 & 0 & 1 & \Delta z_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (34)$$

in which Δx_i , Δy_i and Δz_i are determined via

$$\Delta x_i = x_i + \frac{x_s - x_i}{2}, \Delta y_i = y_i + \frac{y_s - y_i}{2}, \Delta z_i = z_i + \frac{z_s - z_i}{2} \quad (35)$$

The full transformation from sphere to ellipsoid is then given by

$$\mathbf{C}_{i,k} = \mathbf{T}_i^{-T} \mathbf{R}_i^{-T} \mathbf{S}_{i,k}^{-T} \mathbf{C}_1 \mathbf{S}_{i,k}^{-1} \mathbf{R}_i^{-1} \mathbf{T}_i^{-1} \quad (36)$$

B. Common Tangent Estimation

A plane can be parameterized as

$$\psi = \{(x, y, z) \in \mathbb{R}^3 \mid p_1 x + p_2 y + p_3 z + p_4 = 0\} \quad (37)$$

and after defining $\mathbf{p} = [p_1 \ p_2 \ p_3 \ p_4]^T$, this can be written as

$$\mathbf{p}^T \mathbf{v} = 0 \quad (38)$$

The equation that determines if a plane ψ is tangent to an ellipsoid is the following:

$$\mathbf{p}^T \cdot \text{adj}(\mathbf{C}) \cdot \mathbf{p} = 0 \quad \text{with} \quad \text{adj}(\mathbf{C}) = \det(\mathbf{C}) \cdot \mathbf{C}^{-1} \quad (39)$$

The plane that is tangent to many ellipsoids is most likely a surface of the room. Those particular plane parameters are found minimizing the following cost function, where k stands for the number of a reflector:

$$J_e(\mathbf{p}, \{\mathbf{C}_{i,k}\}_{i=1}^M) = \sum_{i=1}^M \|\mathbf{p}^T \cdot \text{adj}(\mathbf{C}_{i,k}) \cdot \mathbf{p}\|^2 \quad (40)$$

One has to be careful not to get trapped in a local minimum. A simplification can be made, and that is using spherical coordinates

$$\begin{cases} p_1 = \sin(\phi) \cos(\theta) \\ p_2 = \sin(\phi) \sin(\theta) \\ p_3 = \cos(\phi) \end{cases} \quad (41)$$

so only three parameters need to be considered to minimize J_e instead of four.

V. ACOUSTIC OBJECT DETECTION

A possible expansion of the technique, is the detection and localization of objects inside a room. The method we propose, starts by measuring the RIRs for each pad from source to microphone of an empty room. Next, when an object is inside the room, the RIRs will be different. By subtracting the original RIR from the other, the peak of the path via reflection on the object should become visible. Like in the reflector localization, ellipses can be constructed. The philosophy is that the location of the object will be surrounded by intersections of ellipses. The room will be divided into subsections, and for every subsection, the number of intersections that lie into that section will be counted. The subsection with the most counts, will most likely approximate the real object location. It's therefore crucial to choose good dimensions of the subsections, in function of the expected size of the object. It's best to try multiple dimensions and compare the results. In figure 1, a result of a simulation of the method is shown, in which the object is considered as a cylinder with radius 0.1 m, on location (2;2.5). The size of the subsections is chosen to be 0.2 m \times 0.2 m. The location that is found (2;2.6) lies within the width of a subsection, which can be seen as a tolerance. Care should be taken when an object blocks the direct path. In that case, there is no reflection, and possibly wrong intersections will be found. Simulations with four microphones showed that a maximum of one direct path can be blocked to obtain a reliable result.

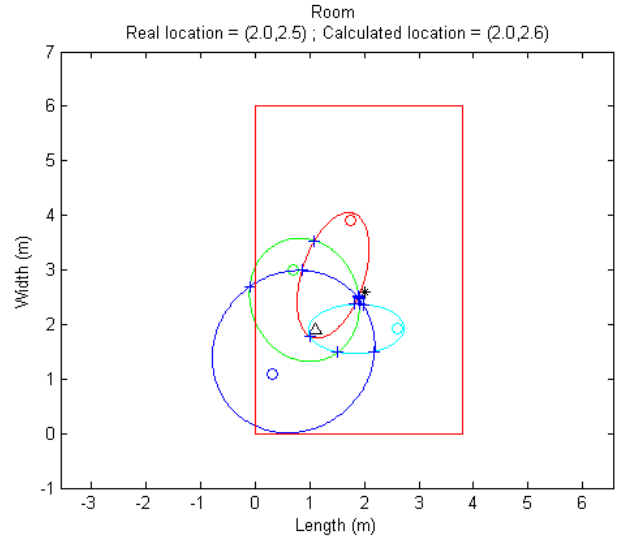


Fig. 1. Result of 2D acoustic object detection, real object location: (2.0;2.5).

VI. RESULTS

Both the source localization algorithm and the room geometry estimation are evaluated using simulations and practical measurements. First, the results of the 3-D source localization

will be described. For achieving this, the AED algorithm is used. As described above, this method also takes reverberation effects into account, which is not the case for the cross correlation method. This should lead to a better performance of the source localization [14]. Another advantage is that knowledge of the source signal is not required for this method. In simulations, using a room model for generating the Room Impulse Responses, the AED algorithm performs well. Although, the performance depends on the source and microphone positions. When microphones are located close to corners of the considered room, the performance of the algorithm degrades. In the practical measurements however, the AED algorithm seems to perform less, when compared to the simulations. Another blind method, that also incorporates reverberation effects, is the closed form method. Again the source signal can be unknown. The disadvantage of this method is the computational complexity, as mentioned before. This complexity increases with the estimated impulse response length L . If the estimated impulse response length L is too short, the nullspace of matrix $\mathbf{X}(L')$ in expression (10) will be empty. This means no result is found for the RIRs. The RIRs which have to be estimated, are rather long. As a consequence, in practice the method is not suitable for our problem.

If the source signal is known, or can be controlled by the application, the NLMS method can be applied. The resulting estimations of the RIRs, contain information about both the direct paths and the first-order reflection paths. This means these results can be used for both the source localization and the room geometry estimation. The results with practical measurements are positive. The direct path is retrieved by comparing the amplitude of an estimated RIR to a specified threshold. This threshold is found as the mean value for a chosen range of the RIR. The first peak that is considerably higher than this threshold, will be a good estimation for the TOA corresponding to the direct path. To improve the results, multiple measurements are performed, with different orientations of the speaker we used as source. This is done to avoid the influence of the non-onmidirectionality of the speaker. Figure 2 shows the result, when four microphones are used. Microphones and source are positioned as in table I. The real source location is (1.3;2.0;1.1), the resulting estimation (1.0;2.0;1.1). When the

TABLE I
LOCATIONS OF SOURCE AND MICROPHONE (IN METERS).

	x	y	z
source	1.3	2.0	1.1
microphone 1	1.75	3.9	1.1
microphone 2	2.6	1.93	1.1
microphone 3	0.7	3	1.1
microphone 4	0.3	1.1	1.1

source location is known, this information together with the TOAs of the first-order reflection paths can be used to infer the room geometrics. When six reflectors are searched (four walls, ceiling and floor) this means that for every microphone, six ellipsoids can be constructed as described earlier. The problem is that one doesn't know which ellipsoid belongs to

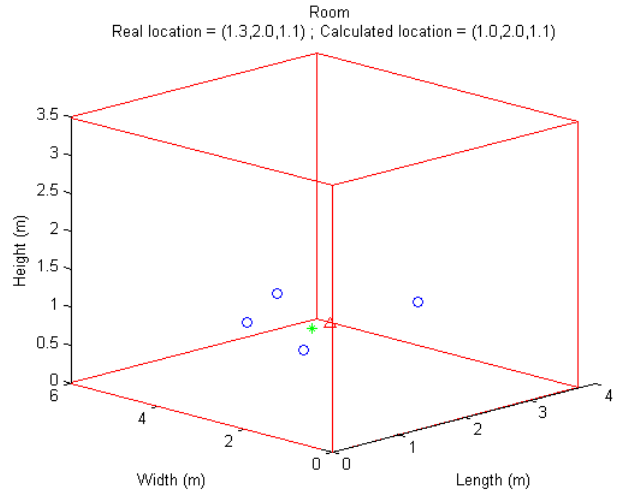


Fig. 2. Result of 3D source localization, real source location: (1.3;2.0;1.1).

which reflector. This means every combination of M ellipsoids needs to be evaluated, where M is the number of microphones. We state that every ellipsoid belongs to a different reflector, which means that there are N^M possible ways to combine the ellipsoids. For each combination then, the parameters of the plane that minimizes expression (40) are searched in an exhaustive manner.

The six first order TOAs, each corresponding to a path along a different reflector, need to be extracted from the estimated RIRs. Under the assumption that a sound wave reaching a microphone being reflected by two walls, has a lower amplitude than sound waves reflected by only one wall, a suitable way to extract the TOAs is by searching the maximum peaks in the RIRs. Of course the peak belonging to the direct path needs to be excluded from those six maxima, which can be done by deleting all contributions in the RIR in a chosen neighbourhood around the direct path. When performing this step on the estimated RIRs, the resulting TOAs still contain some errors. The correctness of the TOAs is crucial in the described method for inferring the geometry. To improve the result, another cost function than (40) is proposed. This cost function only takes into account the calculated cost for two ellipsoids per evaluated combination, as stated below.

$$J_{e,i}(\mathbf{p}, \mathbf{C}_{i,k}) = \|\mathbf{p}^T \cdot \text{adj}(\mathbf{C}_{i,k}) \cdot \mathbf{p}\|^2 \quad \text{met } i = 1 \dots M \quad (42)$$

$$J_{e,all} = \text{sort}([J_{e,1} J_{e,2} \dots J_{e,M}])$$

$$J_e(\mathbf{l}, \{\mathbf{C}_{i,k}\}_{i=1}^M) = \sum_{i=1}^2 J_{e,all}(i)$$

The result is, when one or more TOAs contain errors and lead to false dimensioned ellipsoids, their influence on the resulting parameters of the tangential plane will be minimized. If two or more ellipsoids are constructed with correct TOAs, it's likely that for these ellipsoids a tangential plane will be found which minimizes the cost function (42). If the ellipsoids

are correctly dimensioned, the cost function will be very close to zero for these ellipsoids. Because the false dimensioned ellipsoids have more random dimensions, it's less likely to find a tangential plane for these ellipsoids with a lower cost function, than for the correctly dimensioned ellipsoids. A second improvement to eliminate false results, is to make sure a specific plane can only be found once as result. As final

seem to be better than in 3-D, although the 3-D results are still acceptable.

VII. CONCLUSION

The first subject we described, was the link between the extractions of TOAs out of RIRs with acoustic source localization and geometry estimation. The conclusion that can be drawn here, is that methods which use knowledge of the source signal or a controlled source, give better results than blind estimation methods. Reasons for this are the computational complexity or accuracy of the discussed blind methods. A method that seems to perform well is the NLMS method. The 3-D source localization results are acceptable using this method. The same can be stated for the 3-D room geometry estimation. Results depend on the reflection coefficient of the walls. The proposed acoustic object detection algorithm gives accurate results for the location of the object. The condition for a correct performance of the proposed method, is that some information about the size of the object is known.

REFERENCES

- [1] Sonar. [Online]. Available: <http://en.wikipedia.org/wiki/Sonar>
- [2] J. L. Sutton, "Underwater acoustic imaging," *Proc. IEEE*, vol. 67, pp. 554–566, Apr. 1979.
- [3] N. Brahim, D. Guériot, S. Daniel, and B. Solaiman, "3D reconstruction of underwater scenes using image sequences from acoustic camera," in *IEEE OCEANS 2010*, May 2010.
- [4] D. Ba, F. Ribeiro, C. Zhang, and D. Florêncio, "L1 regularized room modeling with compact microphone arrays," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2010.
- [5] A. M. O'Donovan, D. N. Zotkin, and R. Duraiswami, "Spherical microphone array based immersive audio scene rendering," in *Proceedings of the 14th International Conference on Auditory Display*, Jun. 2008.
- [6] acoustic camera. [Online]. Available: <http://www.acoustic-camera.com>
- [7] A. O'Donovan, R. Duraiswami, and D. Zotkin, "Imaging concert hall acoustics using visual and audio cameras," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2008.
- [8] F. Antonacci, A. Sarti, and S. Tubaro, "Geometric reconstruction of the environment from its response to multiple acoustic emissions," in *Acoustics Speech and Signal Processing (ICASSP)*, Mar. 2010.
- [9] I. Dokmanić, Y. M. Lu, and M. Vetterli, "Can one hear the shape of a room: The 2D-polygonal case," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011.
- [10] W. G. Gardner, "3D audio and acoustic environment modeling," Wave Arts inc., 99 Massachusetts Avenue, Suite 7, Arlington, MA 02474, Tech. Rep., 1999.
- [11] F. Ribeiro, D. Ba, C. Zhang, and D. Florêncio, "Turning enemies into friends: Using reflections to improve sound source localization," in *IEEE International Conference on Multimedia and Expo (ICME)*, Jul. 2010.
- [12] Xbox - Kinect. [Online]. Available: <http://www.xbox.com/nl-NL/Kinect/Home>
- [13] G. Xu, H. Lui, L. Tong, and T. Kailath, "A least squares approach to blind channel identification," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 43, pp. 2982–2993, Dec. 1995.
- [14] J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization," *Acoustical Society of America*, vol. 107, pp. 384–391, Jan. 2000.
- [15] J. Filos, E. A. Habets, and P. A. Naylor, "A two-step approach to blindly infer room geometries," in *Proc. of the International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Aug. 2010.
- [16] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 1661–1669, Dec. 1987.

Result of the 3D localisation with values for p_4 : (6.1,3.7,4.0,0.2,0.8)

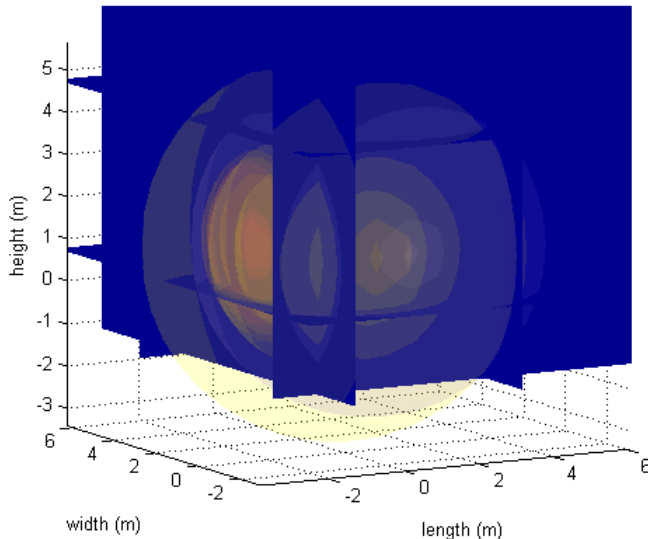


Fig. 3. Result of 3D room reconstruction, real room dimensions: $L = 3.8$ m, $B = 6$ m, $H = 3.5$ m.

remark, the correctness of the extracted TOAs depends on the reflectivity of the walls. A higher reflectivity coefficient, for example $R = 0.9$ or higher, will lead to longer reverberation times. As a consequence, it will in general be harder to recognize the first order reflection TOAs out of the RIR, because waves undergoing more reflections will still possess high amplitudes. The opposite can be said for lower reflection coefficients. Although, as follows from the practical tests, if the reflection coefficient is too low the amplitude of first-order reflection waves will be too low to detect them properly. The dimension of the room also has influence on the detectability of the TOAs, because the paths to be traversed by the waves are longer. This influence seems to be less important than the influence of the reflection coefficients. From simulations with different room sizes, we can conclude that larger rooms still can be reconstructed well. Figure 3 shows the result, using four microphones and one source, in a test room with dimensions $L = 3.8$ m, $B = 6$ m and $H = 3.5$ m. The test room has concrete walls. As source signal white noise is used. The measured signals have a length of 10000 samples, or 227 ms. Microphone and source are positioned as stated in table I. From the result follows that the four straight walls are reconstructed almost correctly, only the ceiling is found at $z = 4$ instead of $z = 3.5$ ($x = 3.7$, $y = 6.1$). Also 2-D simulations were performed, where only four straight walls are considered without floor and ceiling. The results in 2-D