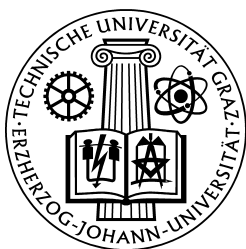


Entwicklung eines drahtlosen, zeitvarianten Binauralsystems auf Ultraschallbasis

Diplomarbeit

durchgeführt und eingereicht von
Nils Peters

Für das:	Institut für elektronische Musik und Akustik (IEM) der Universität für Musik und darstellende Kunst Graz, Österreich
Vorstand:	o.Univ.-Prof. Mag. DI Dr. Robert Höldrich
in Zusammenarbeit mit:	AKG Acoustics GmbH Wien
Begutachter:	o.Univ.-Prof. Mag. DI Dr. Robert Höldrich
Betreuer:	DI David Bajzek (AKG) Univ. Ass. DI Dr. Alois Sontacchi (IEM)



Graz, im Herbst 2005

Danksagung etc.

Hier könnte Ihre Werbung stehen.

Kurzfassung

Die vorliegende Arbeit befasst sich mit der Entwicklung einer unidirektionalen Headtracker-Einheit eines schnurlosen Surround-Kopfhörers. Die Headtracker-Einheit basiert auf einer Ultraschallsendestrecke, welche die azimutale Kopfbewegung relativ zur Bezugsrichtung in Echtzeit erfasst. Die ermittelte Bewegung wird anschließend einer bestehenden Binauraleinheit für die räumliche Aufbereitung des dargebotenen Audiomaterials zugeführt.

Der Kopfhörer ist als Teil eines kombinierten Audio/Video Unterhaltungssystems für Passagiere in Automobilen gedacht. Aufgrund des Einsatzortes ergibt sich die Forderung nach einer robusten und exakten Funktionsweise in akustisch sehr komplizierten Innenräumen. Hierbei gilt es negative Einflüsse von Reflexionen und Rauschen zu minimieren. Ferner sollte durch einen geeigneten Systemaufbau und die Wahl des Detektionsalgorithmus am schnurlosen Kopfhörer ein niedriger Energieverbrauch und damit eine lange Betriebsdauer angestrebt werden.

Darüberhinaus soll das System die Möglichkeit bieten, für mehrere gleichzeitige Benutzer erweitert werden zu können. Die Kopfbewegungen der Anwender müssen in diesem Fall eindeutig separiert und sauber erfasst werden.

Es werden verschiedene Möglichkeiten aufgezeigt und diskutiert, wie diese Ziele konstruktiv umgesetzt werden können. Abschließend werden verschiedenen Tracking-Algorithmen unter der Echtzeitsoftware *Pure Data (PD)* implementiert um die Funktionalität des neuen Headtracker-Systems nach verschiedenen Kriterien zu bewerten und mit einem bestehenden, drahtgebundenen System zu vergleichen.

Abstract

In this thesis a wireless head-tracking system based on ultrasonic is described which was designed for surround headphones. The head tracker detects the azimuthal component of head movements in real-time and provides these data for the auralization unit. The auralization unit will render the audio information to maintain a static audio environment around the listener.

Linked with a video console, these headphones will be part of rear-seat car entertainment system. As a consumer product it has to be robust but accurate enough to work properly in different situations in passenger compartments. For this purpose, the negative effects of reflections and noise should be minimized. Furthermore, the wireless headphone unit has to be optimized for low energy consumption for durability reasons. Above all, the head-tracking system should allow for multi-usability in order to serve a number of passengers in the car simultaneously.

This thesis aims to present a variety of features from different signal processing domains and examines whether they can be employed for this specific task. Finally, the implementation of different tracking algorithms in the real time software *Pure Data (PD)* is described with the aim to evaluate and compare them to a wired reference system under various aspects.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Sinn und Nutzen eines Headtrackers	1
1.2	Bisherige Ansätze und Entwicklungen	2
1.2.1	Beschreibung eines kabelgebundenen Ultraschallsystems	3
1.2.2	Andere Trackingansätze	4
1.3	Entwicklungsziele	6
2	Ultraschall	7
2.1	Allgemein	7
2.2	Physikalische Grundlagen	7
2.2.1	Schallgeschwindigkeit	8
2.2.1.1	Einfluss der Temperatur	8
2.2.1.2	Einfluss des Luftdruckes	9
2.2.1.3	Einfluss der Luftfeuchtigkeit	9
2.2.1.4	Einfluss der Gaszusammensetzung	10
2.2.1.5	Verhalten bei Luftturbulenzen	11
2.2.2	Dämpfung und Frequenzgang	11
2.2.2.1	Bestimmung des Absorptionskoeffizienten	12
2.2.2.2	Frequenzabhängigkeit der Absorption	13
2.2.2.3	Einfluss von Feuchtigkeit und Temperatur	13
2.3	Beschreibung des Schallfeldes	14
2.3.1	Schallausbreitung	14
2.3.2	Reflexion, Brechung, Beugung	15

2.3.3	Bewegte Schallquellen - Dopplereffekte	18
2.4	Gesundheitliche Aspekte	18
3	Geometrische Grundlagen	19
3.1	Situation	19
3.2	Orientierung	21
3.3	Versch. Geometrien	22
3.3.1	Zwei Sender - Vier Empfänger	22
3.3.2	Drei Sender - Drei Empfänger	23
3.3.3	Zwei Sender - Zwei Empfänger	23
3.3.4	Ein Sender - Drei Empfänger	24
3.3.5	Ein Sender - Zwei Empfänger	26
3.3.6	Ein Sender - Ein Empfänger	29
4	Signalformen	30
4.1	Allgemein	30
4.2	Äußere Form	31
4.3	Innere Form, Kodierungsverfahren	32
4.3.1	Modulationsverfahren	34
5	Detektionsmethoden	37
5.1	Schwellwertdetektion	37
5.1.1	Hüllkurvendetektion	39
5.1.2	Verwandte Methoden	40
5.1.2.1	Kombination aus Thresholding und Zero-Crossing De- tection	40
5.1.2.2	Zero-Crossing Detection der Signalhüllkurve	41
5.2	Phasenmessung	41
5.3	Korrelationsbestimmung	42
5.3.1	Korrelationsauswertung im Zeitbereich	43
5.3.1.1	Gewichtsfunktionen	45
5.3.1.2	Digitale Auflösungsgrenzen	49

5.3.2	Korrelationsauswertung im Frequenzbereich	50
5.3.2.1	Verbesserungen	51
5.3.3	1-Bit-Korrelation	53
5.3.3.1	Einfluss der Referenzsignallänge	54
5.3.3.2	Separation verschiedener Signale	56
5.4	Korrelationsalternativen	57
5.5	Dopplerdetektion	60
5.5.1	Messung der Periodendauer	61
5.5.2	Frequenzmessung mittels Kurzzeit Fourier-Transformation	62
5.5.3	Störunterdrückung mittels Wavelet-Filterbank	62
6	Praktische Aspekte	66
6.1	Störgeräusche und Grundrauschen	66
6.2	Psychoakustische Parameter	67
6.2.1	Wahrnehmbare Drehwinkeländerungen	67
6.2.2	Systemlatenzen	68
6.2.2.1	Ergebnisse verschiedener Untersuchungen	68
6.3	Wahl der Sensoren	69
6.3.1	Einfluss der Sensorengröße	70
6.4	Sensorpositionierung	72
6.4.1	Abschattung und Beugung von Schallwellen am Kopf	74
6.4.2	Rotationspunkt des Kopfes in Vergleich zum Modell	74
6.4.3	Abstand der Sensoren am Kopf	76
6.5	Maximale Pulsrate	77
7	Umsetzung	78
7.1	Allgemein	78
7.1.1	Messaufbau	79
7.2	Implementierte Algorithmen	80
7.2.1	Reduzierung der Taktrate durch Einseitenbanddemodulation	81
7.2.2	Schwellwertdetektion der Hüllkurve - Filter	83

7.2.3	1-Bit-Korrelation - BitCor	86
7.2.4	Korrelation im Zeitbereich - ZeitCor	88
7.2.5	AMDF / ASDF	92
7.2.6	Korrelation im Frequenzbereich - PhaseCor	96
7.3	Übersicht	99
7.4	Winkelnachbearbeitung	102
7.4.1	Ergebnisse der Nachbearbeitung	105
7.5	Zwei-System Lösung	108
8	Ausblicke	110
	Bibliographie	111
A	Wichtige Symbole und Bezeichnungen	117
B	Technische Daten der Sensoren	118
C	Technische Daten des Referenzsystems	120
D	Hörbereiche von Haustieren	121

Abbildungsverzeichnis

1.1	Stereo im Vergleich zu einem winkeladaptierten Binauralisierung	2
1.2	Head Mounted Display „ <i>Sword of Damocles</i> “	3
1.3	Logitech Headtracker	4
2.1	Schallgeschwindigkeit als Funktion der Temperatur	9
2.2	Abhängigkeit der Schallgeschwindigkeit von der Luftfeuchtigkeit	10
2.3	Absorption als Funktion der Frequenz und Luftfeuchtigkeit	13
2.4	Nah- und Fernfeld eines Kolbenmembranstrahlers	14
2.5	Reflexion und Brechung von Ultraschall	16
2.6	Geometrischer Strukturen für diffuse Reflexionen	17
3.1	Laufzeitenmessung bei Trackinganwendungen	19
3.2	Horizontalschnitt des Kopfes in der Ebene der oberen Zahnreihe	20
3.3	Relative Orientierung von Hörern im Raum	21
3.4	Absolute Orientierung im Raum	22
3.5	4 Sensoren zur Bestimmung der Einfallsrichtung im Raum	23
3.6	Skizze für die Anordnung von zwei Empfänger und Sender	24
3.7	Skizze für die Anordnung von drei Sensoren am Kopfhörer	25
3.8	Lösungsmenge bei drei ko-linearen Sensoren	26
3.9	Skizze für die Anordnung mit einem Sender und zwei Empfängern	27
3.10	exakte Lösung und Näherung mit zwei Sensoren	28
3.11	Fehler der Näherungsformel bei versch. Drehwinkeln und Distanzen	28
4.1	Strukturbild des realen Übertragungspfad	30

4.2	Chirp- und Burstsinal im Vergleich	32
4.3	binäres lineares Schieberegister	33
5.1	Impulsantwort eines Piezowandlers	38
5.2	Detektion mit angepasstem Schwellwert	38
5.3	Strukturdiagramm der adaptiven Hüllkurvendetektion	40
5.4	kombiniertes Thresholding und Zero-Crossing	40
5.5	Zero-Crossing Detection der Signalhüllkurve	41
5.6	Direkte Korrelationsberechnung im Zeitbereich	44
5.7	Berechnung der Korrelation über die Kreuzleistungsdichte	44
5.8	Kreuzkorrelationsfunktion ohne und mit Vorfilterung	46
5.9	Parabolische Interpolation	49
5.10	Kreuzkorrelationsfunktion und entsprechende Phase des Kreuzleistungs- spektrums	50
5.11	1-Bit-Korrelation eines Signals an zwei Empfängern	53
5.12	Blockschaltbild eines 128 stufigen 1Bit-Korrelators	54
5.13	1-Bit Korrelation mittels FIR-Filter	54
5.14	1-Bit Korrelation mit verschiedenen Eingangssignalen	57
5.15	Varianz als Funktion vom SNR	58
5.16	Varianz bei geringem SNR	60
5.17	Messung der Periodendauer eines Doppler-Tracking Systems	61
5.18	Vergleich Periodendauermessung mit STFT	63
5.19	Erkennen von Störungen durch invertierten der Hüllkurve	64
5.20	Vergleich Echoanfälligkeit der Dopplerverfahren	64
5.21	Blockschaltbild einer adaptiven Wavelet-Filterbank	65
6.1	MAMA als Funktion von Geschwindigkeit und Stimulus	67
6.2	Einfluss der Systemlatenz auf die Lokalisation	69
6.3	Laufzeitenfehler aufgrund der Sensorgröße	70
6.4	Gegenüberstellung der Detektionsfehler bei versch. Sensordurchmessern . .	71
6.5	System mit zwei Sendern am Kopfhörer	72
6.6	System mit zwei Empfängern am Kopfhörer	73

6.7	Winkelfehler bei falscher Sensorposition; $b = 3\text{cm}$	75
6.8	Winkelfehler bei falscher Positionierung des Kopfhörers	75
6.9	Auflösungsgrenzen als Funktion der Aperturgröße	76
6.10	Problematik bei Raumreflexionen und Richtcharakteristik der Sensoren	77
7.1	Systemaufbau zur Umsetzung	78
7.2	Aufbau zur Evaluierung der Algorithmen	79
7.3	Winkelfehler durch 10cm Höhenunterschied zwischen Sender und Empfänger	80
7.4	Schema der Einseitenbanddemodulation	82
7.5	Signalzustand während der ESB-Demodulation	82
7.6	Strukturbild zur Detektion der Hüllkurven mittels Schwellwert	83
7.7	Messergebnis Hüllkurvendetektion	85
7.8	Messergebnis analoge Hüllkurvendetektion	85
7.9	Strukturbild <i>BitCor</i>	86
7.10	Messergebnis <i>BitCor</i>	87
7.11	Strukturbild <i>ZeitCor</i>	88
7.12	<i>ZeitCor</i> mit Medianfilter	89
7.13	<i>ZeitCor</i> mit Schwellwertabfrage	89
7.14	Korrelation an sieben diskreten Verzögerungspunkten	90
7.15	periodische Fehler bei <i>ZeitCor</i>	91
7.16	Messergebnis <i>ZeitCor</i>	91
7.17	Strukturbild, <i>AMDF</i> bzw. <i>ASDF</i>	92
7.18	<i>AMDF</i> -Algorithmus	93
7.19	periodische Fehler während der <i>AMDF</i>	94
7.20	periodische Fehler während der <i>ASDF</i>	94
7.21	Messergebnis <i>AMDF</i>	95
7.22	Messergebnis <i>ASDF</i>	95
7.23	Strukturbild <i>PhaseCor</i>	96
7.24	Ideales Spektrum vom Eingangssignal nach Demodulation	96
7.25	Ableitung der verrauschten Phase des Kreuzleistungsspektrums	97

7.26	Detektionsergebnis mit verschiedenen Glättungsfiltern	98
7.27	Messergebnis <i>PhaseCor</i> mit spektralen Glättungsfilter mf=0.9	98
7.28	Messergebnis <i>PhaseCor</i> mit spektralen Glättungsfilter mf=0.95	99
7.29	gemittelter Winkel und Standardabweichung der Algorithmen	100
7.30	Wege zur Fehlerbereinigung des Detektionsergebnisses	102
7.31	Medianfilterung	103
7.32	Strukturdiagramm der dynamischen Filterung	104
7.33	Winkelergebnisse der Nachbearbeitung	107
7.34	Messung von <i>PhaseCor</i>	107
7.35	Situation im Parallelbetrieb	108
7.36	Signalseparation Dual System Solution	109
B.1	Abmessungen der <i>Murata</i> -Sensoren	118
B.2	Frequenzgang der Sensorstrecke bei zueinander ausgerichteten Sensoren für verschiedene Distanzen	119
B.3	Richtcharakteristik der Sensorstrecke	119

Kapitel 1

Einleitung

1.1 Sinn und Nutzen eines Headtrackers

Der Durchbruch der DVD (*Digital Versatile Disc*) als Nachfolgemedium zur VHS-Kassette, bringt das Surroundkino in die heimischen Wohnzimmer. In Kombination mit den entsprechenden HiFi-Komponenten ist es nun möglich die akustisch perfekte Illusion nicht nur im Kinosaal zu erleben. Entfernung, Richtung und die eigenen Position sollen laut den Herstellern „akustisch fühlbar“ sein.

Allerdings ist die Qualität dieser Illusion oftmals gemindert. So beeinträchtigen die unsachgemäße Positionierung der (häufig klanglich unausgewogenen) Lautsprecher sowie eine Hörposition außerhalb des *sweetspots*¹ das Hörerlebnis stark. Weitere beeinflussende Faktoren sind die Raumakustik des Wohnzimmers sowie eine (nachbarschaftsstreitvermeidende) Lautstärken- und Dynamikreduktion der Tonspur.

Alle angesprochenen Probleme kann ein Kopfhörer lösen, der das dargebotene Audiomaterial für eine Kopfhörerwiedergabe richtungsabhängig aufbereitet (*binauralisiert*) und damit einen räumlichen, natürlichen Höreindruck schafft². Mit jeder Kopfbewegung dreht sich allerdings das erzeugte Klangbild mit - die Position virtueller Schallquellen passt nicht mehr zum Bild und der Höreindruck wirkt plötzlich unreal. Um ein korrektes bewegungsunabhängiges Schallfeld zu schaffen, muss daher bei der Binauralisierung die Kopfbewegung berücksichtigt werden.

Das Verfolgen (engl. „*tracking*“) dieser Kopfbewegung und das Bereitstellen des aktuellen Kopfwinkels für die Binauralisierung ist die Aufgabe des Headtrackers. Da er für die Binauralisierung eine wichtige Information liefert, besitzt der Tracker großen Einfluss auf die naturgetreue, realistische Klangbildumsetzung. Im Rahmen der Wahrnehmungsfähigkeit³ sollte ein Headtracker exakt und schnell die Winkelinformation liefern.

¹der ideale Abhörpunkt eines Mehrkanalsystems

²siehe [IEM1], [BS00], [STU1], [STU2]

³Details zu den menschlichen Wahrnehmungseigenschaften siehe Kap. 6.2 auf Seite 67

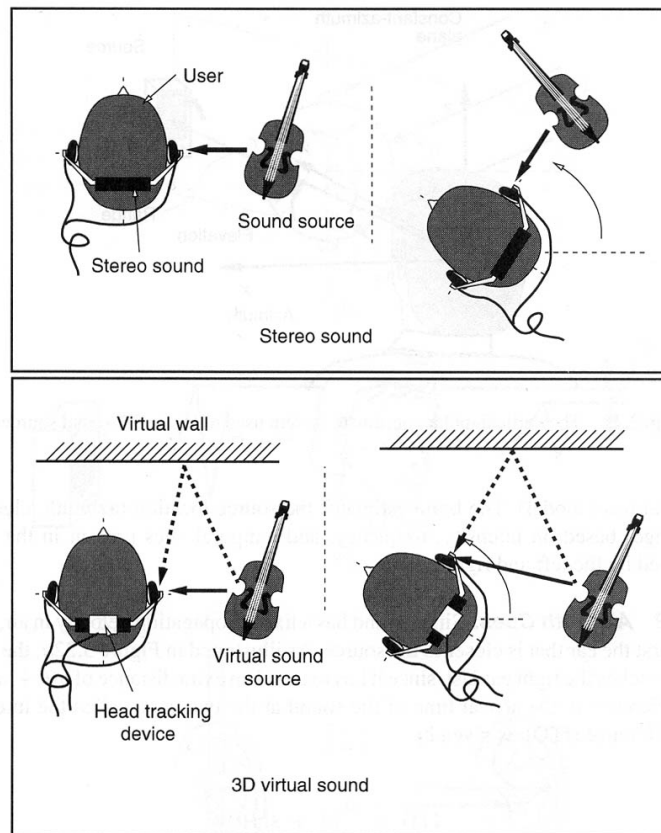


Abb. 1.1: Stereo im Vergleich zu einem winkeladaptierten Binauralisierung ([BU03])

1.2 Bisherige Ansätze und Entwicklungen

Die Entwicklung von Motion- oder Headtracking-Systemen steht eng im Zusammenhang mit der Erzeugung von Virtuellen Realitäten (VR). Die ersten VR-ähnlichen Experimente begannen mit Ivan Sutherland's „*Sword of Damocles*¹“, dem ersten Head Mounted Display (HMD) im Jahr 1968 (siehe Abb.1.2). Ermöglicht durch die Fortschritte in der Computertechnologie und der Entwicklung von leistungsfähigen Signalprozessoren entstanden seit den späten 80er Jahren Headtracking Systeme von *Polhemus*, *Logitech* oder *Ascension*.

¹Das HMD wurde so benannt, da mehrere Kilogramm Hardware von der Decke auf den Benutzer herunterhingen.

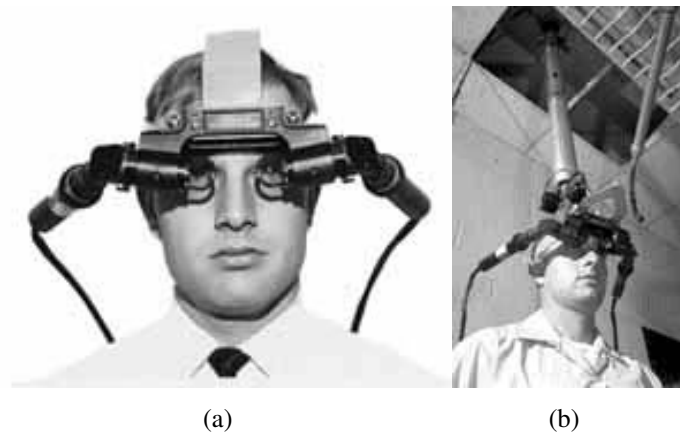


Abb. 1.2: HMD „Sword of Damocles“ (1968)

1.2.1 Beschreibung eines kabelgebundenen Ultraschallsystems

Bisher wurden nur akustische Headtracker gefertigt¹, deren Sende- und Empfangssensoren kabelverbunden sind. Am Beispiel des *Logitech* Headtrackers soll das Funktionsprinzip eines kabelgebundenen Headtrackers erläutert werden.

Das System besteht aus einer dreieckigen, ca 20cm großen Sende- und einer kleineren, am Kopfhörer montierbaren Empfangseinheit mit jeweils 3 Ultraschallsensoren (Abb.1.3 (a)). Sender und Empfänger sind mit einer zentralen Steuer- und Recheneinheit verbunden. Über die Steuereinheit werden nun reihum die Transmitter mit Burstsignalen angeregt und hochtaktige Zählwerke gestartet. Trifft ein Signale an einem Empfängern ein, meldet der Empfänger dies der Steuereinheit über die Kabelverbindung. Die Steuereinheit stoppt den jeweiligen Zähler und speichert den aktuellen Zählerstand. Nachdem von allen 3 Transmitter die Signale übertragen wurden, besitzt die Recheneinheit $3 \cdot 3 = 9$ Zählerstände, welche die 9 Laufzeiten der drei Burstsignale von den Sendern zu den 3 Empfängern repräsentieren. Die Laufzeiten l_n sind über die Schallgeschwindigkeit proportional zu den Wegstrecken w_n den die Burstsignale zurückgelegt haben (Gl.1.1). Mittels einfacher Triangulation erhält man Information über alle 6 Freiheitsgrade² im Raum im Bezug zum Sender.

$$w_n = c \cdot l_n \quad (1 \leq n \leq 9) \quad (1.1)$$

Es besteht die Möglichkeit mit einer Sendereinheit mehrere Empfängereinheiten zu versorgen und somit gleichzeitig mehrere Benutzer zu bedienen³ (siehe Abb.1.3 (c)).

Die relative Genauigkeit ist bis zu einer Distanz bis 1,4 Meter mit 2% angegeben.

Das Funktionsprinzip beruht also auf die Kenntnis der Aussende- und Empfangszeitpunkte der Burstsignale. Ohne die Kabelverbindung der Steuereinheit mit den Sendern und Empfängern ist der Headtracker daher nicht betriebsfähig. Akustische Tracker haben prinzi-

¹ Siehe auch das Studer-System:([STU1], [STU2],[EB02]).

² Position (x,y,z) und Rotation um alle 3 Koordinatenachsen

³ engl.:Single Server- Multi Client

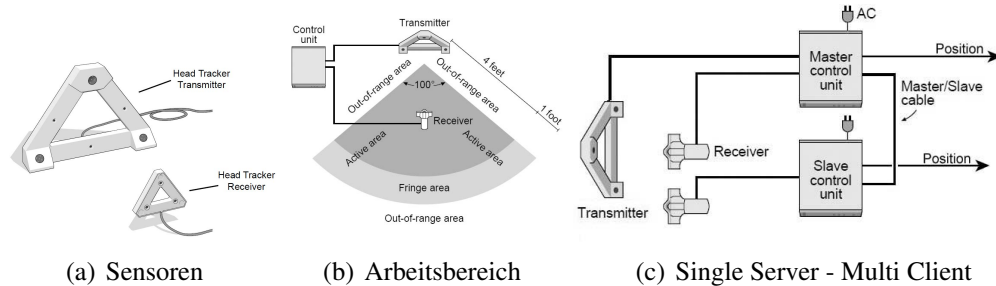


Abb. 1.3: Logitech Headtracker [LO92]

piell einige Schwächen. So kann es bei Schallreflexionen zu Problemen mit Interferenzen kommen. Weiterhin muss eine ständige Sichtverbindung zwischen Sender und Empfänger bestehen und die Schallgeschwindigkeit möglichst exakt bekannt sein (siehe Gl.3.1, Kap.2, Kap.3). Im Vergleich zu anderen Verfahren sind akustische Tracker jedoch kostengünstig, robust und leichtgewichtig.

1.2.2 Andere Trackingansätze

Neben dem Ultraschallverfahren wurden folgende Methoden zur Erfassung von Bewegungsabläufen entwickelt ([BS00], [BH93], [BU03]):

Mechanisch Das System wird direkt am Hals getragen, wobei die Kopfstellung über mechanisch veränderlichen Widerständen (Potentiometer) bestimmt wird. Der gravierendste Nachteil ist die starke Einschränkung der Bewegungsfreiheit. Für Headtracking-Anwendungen sind sie aufgrund der umständlichen Befestigung der Sensoren am Körper daher nur bedingt geeignet (Laborversuche).

Elektromagnetisch Elektromagnetische Tracker können je nach der Art der Erzeugung des Magnetfeldes in zwei Klassen geteilt werden: A.C. magnetic tracker (Wechselstrom) oder D.C. magnetic tracker (Gleichstrom). Beide Klassen funktionieren nach dem gleichen Prinzip: In einem Sender, der aus drei orthogonal angeordneten Spulen besteht, wird ein elektromagnetisches Feld erzeugt und von einem baugleichen Sensor empfangen. Anhand der Stärke und des Verhältnisses der in den Spulen des Empfängers induzierten Ströme kann eine Aussage über Position und Lage getroffen werden. Die Messgenauigkeit ist sehr hoch, sie liegt im Bereich der Lagewinkel bei 0.5 Grad. Da das Magnetfeld nicht durch den menschlichen Körper beeinträchtigt wird, gibt es keine Abschattungsprobleme. Andererseits verursacht die Felderzeugung im A.C.-Betrieb störende Wirbelfelder in allen metallische Objekte in der Umgebung, worunter die Genauigkeit des Trackers leidet. Im D.C.-Betrieb verursachen ferromagnetische Gegenstände einen Genauigkeitsverlust. Sowohl die A.C. als auch die D.C.-Methode wird von fremden elektromagnetischen Feldern

(z.B. CRT-Displays) beeinflusst. Vertreter dieser Art sind der *Isotrak* und *Fastrack* von *Polhemus*¹ oder der *Flock of Birds* der Firma *Ascension*².

Optisch Optische Systeme stehen den akustischen Systemen sehr nahe, wobei anstatt Ultraschallsensoren hier helligkeitsempfindliche Fotodioden verwendet werden. Anhand der Beleuchtungsstärke, die eine Lichtquelle an verschiedenen positionierten Fotodioden erzeugt, kann über Triangulation auf den Ort der Lichtquelle geschlossen werden (*PSD - Position Sensing Detector*). Mit kohärenten Laserstrahlen arbeitende Systeme verwenden strukturiertes Licht, um einen Raum abzutasten. Das an Objekten reflektierte Licht wird mit einer Kamera aufgezeichnet. Aus den Bildern kann auf die Position und Lage geschlossen werden.

Der Vorteil des einfachen Systemaufbaus steht die hohe Störempfindlichkeit bei starker Lichteinstrahlungen als Nachteil gegenüber, sodass eine Anwendung bei Tageslicht bzw. Sonneneinstrahlung schwer möglich ist. Im Vergleich zu den akustischen Trackern kann bei Optischen Systemen durch die extrem hohe Ausbreitungsgeschwindigkeit eine bessere zeitliche Dynamik erzielt werden. Alle optische Systeme haben das Verdeckungsproblem und sind für 3D- Audio Anwendungen nur eingeschränkt geeignet.

Bildbasierend Hierbei werden zwei oder mehr Kameras verwendet, die aus Bewegungsverläufen von Targets auf deren Positionen schließen. Diese Targets müssen aktiv (z.B. durch befestigte LEDs) oder passiv (durch Markierungen oder Rasterungen) gekennzeichnet sein. Voraussetzung ist eine Sichtverbindung zwischen Kamera und Ziel, sowie für eine Realtime-Anwendung eine leistungsfähige Bildverarbeitung (z.B. [GO97]). Besonders beim Einsatz von passiven Targets führen reflektierende Materialien im Beobachtungsraum zu Irritationen am Kamerasystem.

Inertialsysteme Das Prinzip beruht auf der Messung der Kopfbeschleunigung mit anschließender Integration. Andere Systeme bestimmen direkt über Coriolis-Kraft³ die Winkelgeschwindigkeit (Gyroskope). Typischerweise kommt es dabei zu einem Abdriften, was durch regelmäßiges Kalibrieren kontrolliert werden kann. Langzeitdrift und Anfälligkeit sind weitere Schwächen. Inertialsysteme haben jedoch den Vorteil, dass sie keine Quelle, wie etwa elektromagnetische Tracker benötigen. Dadurch ergibt sich eine theoretisch beliebige Reichweite.

¹<http://www.polhemus.com>; copyright ©2005 Polhemus. 40 Hercules Drive, P.O. Box 560, Colchester, VT 05446, USA

²<http://www.ascension-tech.com>; copyright ©2005 Ascension Technology Corporation. P.O. Box 527 Burlington, VT 05402, USA.

³Bezeichnung für eine Trägheitskraft, die scheinbar auf Körper wirkt. Diese Kraft lässt sich feststellen, wenn man den Körper von einer mit dem Bezugssystem rotierenden Stelle aus betrachtet. Ändert ein Objekt in einem rotierenden System seinen Abstand zur Drehachse, ändert sich seine Geschwindigkeit. Sie wird umso kleiner, je näher sich der Gegenstand am Mittelpunkt befindet. Der Betrag dieser Geschwindigkeitsänderung pro Zeit entspricht der Coriolis-Kraft.

1.3 Entwicklungsziele

Das hier zu entwickelnde Echtzeitsystem sollte folgende Punkte erfüllen:

- Unidirektionale, drahtlose Kommunikation zwischen Kopfhörer und Basiseinheit mittels Ultraschallsignalen
- Einsatzbereich: Halbraum mit Radius von 0.3 bis ca. 1.5 Meter
- Sehr hohe azimutale Winkelauflösung (Winkelauflösung 1°)
- Sichere Winkeldetektierung bis $\pm 45^\circ$
- Bestmögliche dynamische Auflösung innerhalb standardisierter menschlicher Wahrnehmungsgrenzen
- Geringe Störanfälligkeit gegenüber Echos und systemfremden Ultraschallsignalen
- Geringer Leistungsbedarf der Headtracking-Komponenten am batteriebetriebenen Kopfhörer
- Kostengünstige Realisierung und Robustheit für Serienproduktion
- Keine hörbaren Frequenzen infolge von subharmonischen Komponenten der Ultraschallsignale
- Möglichkeiten zum Ausbau auf Multiuser-System mit mehreren Kopfhörern
- Ungestörte Funktionalität von Einzelsystemen in unmittelbarer Nachbarschaft

Kapitel 2

Ultraschall

2.1 Allgemein

Als Ultraschall bezeichnet man Frequenzen zwischen 20 kHz bis 1 GHz und damit Schallwellen oberhalb des menschlichen Hörbereiches. Ultraschall breitet sich nicht nur in Gasen oder Flüssigkeiten aus, sondern tritt auch in verschiedenen Formen elastischer Wellen in Festkörpern auf (siehe [KU88] und [SU80]).

Obwohl weiterhin die aus dem Hörschallbereich bekannten physikalischen Gesetze der Schallentstehung und Schallausbreitung gelten, gibt es praktische und technische Gründe den Ultraschall gesondert zu behandeln. So unterscheiden sich die Methoden zur Schallerzeugung und Schallmessung im Allgemeinen von denen im Hörschallbereich. Weiterhin sind die erreichbaren Beschleunigungen, Wechseldrucke, Schallintensitäten oder Energiedichten um ein vielfaches größer als im Hörschallbereich. Dieses Kapitel soll Aufschluss über die Ausbreitungseigenschaften und ihren bestimmenden Faktoren in der Luft liefern. Abschließend wird auf gesundheitliche Aspekte bei Ultraschalleinstrahlung eingegangen.

2.2 Physikalische Grundlagen

Die Ausbreitung von Schall ist im Gegensatz zu elektromagnetischen Wellen stets an ein stoffliches Medium gebunden und wird von diesem entscheidend beeinflusst. Hier werden die Ausbreitungseigenschaften von Schall (im besonderen Fall der Ultraschall) in Luft zusammengefasst. Für alle in Kapitel 5 vorgestellten Detektionsmethoden ist eine genaue Bestimmung der tatsächlichen Schallgeschwindigkeit Grundlage für eine exakte Kopfwinkelbestimmung.

2.2.1 Schallgeschwindigkeit

2.2.1.1 Einfluss der Temperatur

In Gasen ist nur eine Ausbreitung von Longitudinalwellen möglich. Für die Schallgeschwindigkeit folgt aus der eindimensionalen Wellengleichung für den Druck p die Laplacesche Gleichung der Schallgeschwindigkeit:

$$c = \sqrt{\frac{(C_p \cdot p)}{(C_v \cdot \rho)}} = \sqrt{\kappa \frac{p}{\rho}} \quad (2.1)$$

Die spezifische Wärme (C_p und C_v) ist jene Wärme, die man benötigt, um die Temperatur einer Masse von $m = 1\text{kg}$ um $\Delta T = 1^\circ\text{C}$ zu erhöhen. Aus Gleichung 2.1 folgt, dass die Schallgeschwindigkeit vom Luftdruck unabhängig ist, solange Druck und Dichte proportional bleiben. Da die Dichte bei konstantem Druck von der Temperatur abhängt (Gl.2.2), ändert sich auch die Schallgeschwindigkeit mit der Temperatur (siehe Gl.2.3 bis Gl.2.5).

$$\rho(T) = \frac{\rho_0}{1 + \alpha \cdot T} \quad (2.2)$$

ρ_0 : Dichte bei Bezugstemperatur und $p = 1\text{ bar}$

α : Ausdehnungskoeffizient $\left(= \frac{1}{273,16} \text{ Kelvin}^{-1} \right)$

$$c(T) = \sqrt{\kappa \frac{p}{\rho(T)}} \quad (2.3)$$

$$= \sqrt{\frac{\kappa \cdot p \cdot (1 + \alpha \cdot T)}{\rho_0}} \quad (2.4)$$

$$= c_0 \cdot \sqrt{1 + \alpha \cdot T} \quad (2.5)$$

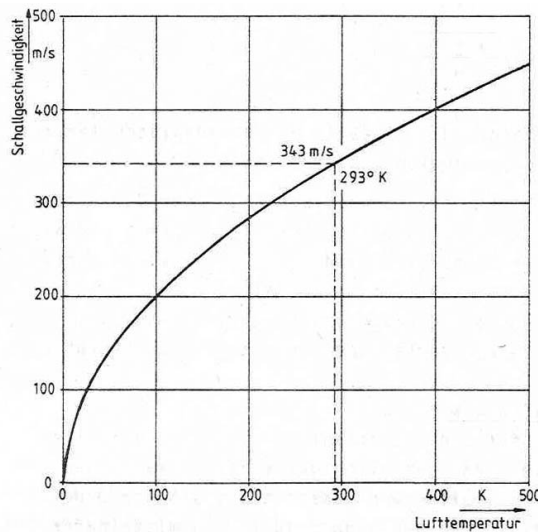
Gleichung 2.5 führt zu Gl.2.6, der Gleichung der temperaturabhängigen Schallgeschwindigkeit in Luft:

$$c(T) = 331,45 \cdot \sqrt{1 + \frac{T}{273,16^\circ\text{K}}} \quad \left[\frac{\text{m}}{\text{s}} \right] \quad (2.6)$$

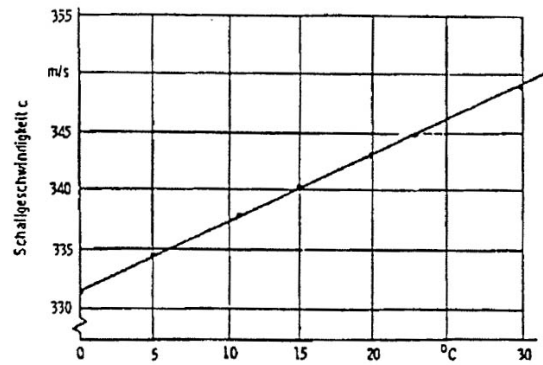
c : Schallgeschwindigkeit $(= 331,45 \frac{\text{m}}{\text{s}}$ bei $\theta_0 = 0^\circ\text{C}$ und $p = 1,013 \cdot 10^5 \text{ Pa}$)

Das bedeutet eine Zunahme der Geschwindigkeit von 0,18% pro $^\circ\text{C}$ (bzw. $0,6 \frac{\text{m}}{\text{s}}$ pro $^\circ\text{C}$) und damit eine wesentliche Genauigkeitsbegrenzung des Systems bei unbekannter Umgebungstemperatur.

Eine Temperaturänderung von beispielsweise $\pm 1^\circ\text{C}$ bewirkt einen Messfehler von ca. $2,5\text{mm}$ auf 1,5 Meter bei 20°C Referenztemperatur. Wie man in Abb.2.1 sieht, ist die Änderung der Schallgeschwindigkeit nicht linear, kann aber in einem schmalen Temperaturbereich als linear angesehen werden (siehe Abb 2.1).



(a) Verlauf im großen Bereich ([EH85])



(b) Änderung im erwarteten Arbeitsbereich des Headtrackers ([ZI92])

Abb. 2.1: Schallgeschwindigkeit als Funktion der Temperatur

2.2.1.2 Einfluss des Luftdruckes

Eine langsame Luftdruckänderung hat auf die Schallgeschwindigkeit keinen Einfluss, da Luftdruck und Dichte der Luft bei gleicher Temperatur zueinander proportional sind. In Gleichung 2.1 ist der Bruch $\frac{p}{\rho}$ daher immer konstant. Eine plötzliche extreme Luftdruckänderung (z.B. eine Explosion) würde theoretisch die Homogenität des Ausbreitungsmediums und damit die Schallgeschwindigkeit beeinflussen.

Eine mittelbare Auswirkung hat der Luftdruck auf die Dispersion (siehe Abb.2.2.2).

2.2.1.3 Einfluss der Luftfeuchtigkeit

Die Abhängigkeit der Schallgeschwindigkeit von der Luftfeuchtigkeit kann folgendermaßen ausgedrückt werden:

$$c_f = \frac{c_t}{1 - \frac{e}{p} \cdot \left(\frac{w}{a} - 0,625 \right)} \quad (2.7)$$

c_t : Schallgeschwindigkeit der trockenen Luft

c_f : Schallgeschwindigkeit der feuchten Luft

e : Dampfdruck

w/a : Verhältnis von spezifischer Wärme bei konstantem Druck zu spezifischer Wärme bei konstantem Volumen für Luft (a) und Wasserdampf (w)

Der Einfluss der Luftfeuchtigkeit ist bei einer Temperatur von 20°C relativ gering. Bei vielen akustischen Laufzeitmessungen wird die Luftfeuchtigkeit daher nicht kompensiert. Mit steigender Temperatur bewirkt die Luftfeuchtigkeit jedoch Schallgeschwindigkeitsänderungen im Prozentbereich.

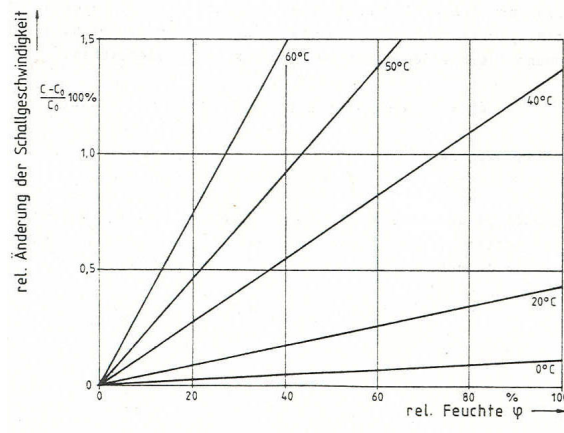


Abb. 2.2: Abhängigkeit der Schallgeschwindigkeit von der Luftfeuchtigkeit [EH85]

2.2.1.4 Einfluss der Gaszusammensetzung

Die Anteile der zahlreichen Gase in der Luft liegen in der freien Atmosphäre in stabiler Zusammensetzung vor. In geschlossenen Räumen kann sich jedoch das Gasgemisch verändern und dabei die Schallgeschwindigkeit beeinflussen.

Folgende Tabelle zeigt die wichtigsten Bestandteile der Luft mit den Werten ihrer Schallgeschwindigkeit ([EH85]):

Gas	Volumenanteil (%)	c bei 0°C
Stickstoff	78,08	334
Sauerstoff	20,95	316
Argon	0,93	319
Kohlendioxyd	0,03	338
Sonstige	0,01	415 - 1284

Tab. 2.1: Zusammensetzung der Luft

Ein der Luft beigemischtes Gas verursacht eine (dem Volumensanteil entsprechende) Än-

derung der Schallgeschwindigkeit.

$$\begin{aligned} c &= (V_L \cdot c_0) + (V_G \cdot c_G) \\ 1 &= V_L + V_G \end{aligned} \quad (2.8)$$

V_L :	Volumensanteil der Luft
V_G :	Volumensanteil des zugeführten Gases
c_0 :	Schallgeschwindigkeit der Luft
c_G :	Schallgeschwindigkeit des zugeführten Gases

Steigt der Anteil von CO_2 infolge der Atmung von Personen in einem Raum auf 1% an, so ergibt sich eine relative Änderung der Schallgeschwindigkeit von:

$$\frac{\Delta c}{c_0} = \frac{c - c_0}{c_0} = 2 \cdot 10^{-4} \quad (2.9)$$

Eine Veränderung der Luftzusammensetzung hat also relativ wenig Auswirkungen auf die Schallgeschwindigkeit und wird daher in den meisten Ultraschallanwendungen nicht berücksichtigt.

2.2.1.5 Verhalten bei Luftturbulenzen

Die Beeinflussung der Schallgeschwindigkeit von der Windgeschwindigkeit kann durch eine Erweiterung von Gleichung 2.6 gut approximiert werden ([WE97]).

$$c_i = 331,45 \cdot \sqrt{1 + \frac{T}{273,16^\circ K}} + \vec{e}_i^T \vec{v}_w \quad \left[\frac{m}{s} \right] \quad (2.10)$$

c_i :	die Schallgeschwindigkeit in Richtung $\vec{e}_i$
$\vec{v}_w$:	Windgeschwindigkeit in $\left[\frac{m}{s} \right]$

Räumliche und zeitliche Fluktuationen der Windgeschwindigkeit sind neben Temperaturschwankungen der Luft eine der größten Fehlerquellen bei Ultraschallanwendungen ([WE97]). Untersuchungen finden sich auch in [MC04] und [FI92]. Inwieweit Luftturbulenzen im Automobil (z.B. durch Gebläse oder eines offenen Seitenfensters) zu Irritationen des Headtrackers führen, ist im jeweiligen Fahrzeug zu untersuchen.

2.2.2 Dämpfung und Frequenzgang

Die Schallabsorption führt zu einer exponentiellen Abnahme der Amplitude mit zurückgelegtem Weg und ist in allen Ausbreitungsmedien gegeben durch ([EH85]):

$$\hat{p}(x) = p_0 \cdot e^{-\alpha x} \quad (2.11)$$

wobei α die Absorptionskonstante ist. Diese hängt von der Art des Ausbreitungsmediums und von der Frequenz ab. Das oft praktischere Dämpfungsmaß gibt die Abnahme des Schalldruckpegel in Dezibel pro Weeinheit an ([KU88]).

$$A = 20\alpha \cdot \lg(e) \approx 8,69 \cdot \alpha \left[\frac{dB}{m} \right] \quad (2.12)$$

2.2.2.1 Bestimmung des Absorptionskoeffizienten

Der Absorptionskoeffizient α setzt sich aus den Teilkoeffizienten der unterschiedlichen Absorptionsursachen zusammen. Bei Schallausbreitung in Gasen sind dies:

α_{WS} :	Wärmestrahlung
α_R :	Reibung und Schubviskosität
α_{WL} :	Wärmeleitung
α_{RX} :	Innermolekulare Relaxation

Der Einfluss der Wärmestrahlung tritt nur bei Frequenzen unter 30 Hz auf und kann daher bei Ultraschall vernachlässigt werden ([EH85]). Somit bestimmt sich der Absorptionskoeffizient für Ultraschall durch:

$$\alpha = \alpha_{WL} + \alpha_R + \alpha_{RX} \quad (2.13)$$

der Anteil der Reibung (α_R) und Wärmeleitung (α_{WL}) werden als klassische Ursachen für die Absorption beschrieben und können folgendermaßen ausgedrückt werden: ([EH85])

$$\alpha_R + \alpha_{WL} = \frac{2\pi^2 f^2}{c^3 \rho_0} \cdot \left(\frac{4}{3} \cdot \eta + \frac{\kappa - 1}{C_p} \cdot K \right) \quad (2.14)$$

η :	Schubviskosität
K :	Wärmeleitzahl

Der von Reibung und Wärmeleitung herrührende Dämpfungsanteil ergibt sich nach Gl. 2.14 für 40 KHz zu $\alpha \approx 2 \cdot 10^{-2} m^{-1}$. Der tatsächliche Wert ist jedoch meist erheblich höher. Einen zusätzlichen, und nicht zu vernachlässigbaren Absorptionsanteil leisten bei mehratomigen Gasen so genannte innermolekulare Relaxationsvorgänge ([KU88] und [BR80]). Ursache der Dämpfung ist die Trägheit der Moleküle und damit ein Zeitverzug in der Einstellung der Schwingungsenergie der Gasmoleküle bei schnellen Druckänderungen. Das bedeutet also, dass Schallgeschwindigkeit und Absorption zum Teil frequenzabhängig sind. Dabei wird vom Gas Arbeit verrichtet, die der Schallwelle Energie (auch Relaxationsenergie genannt) entzieht. Diese Energie wird schließlich als Wärme umgesetzt. Diese Vorgänge sind nicht einfach analytisch beschreibbar, da zwischen den Relaxationserscheinungen und den Absorptionsursachen komplizierte Wechselwirkungen bestehen. In ([EH85]) finden sich eine empirisch bestimmte Gleichung für eine Annäherung des Absorptionsgrades.

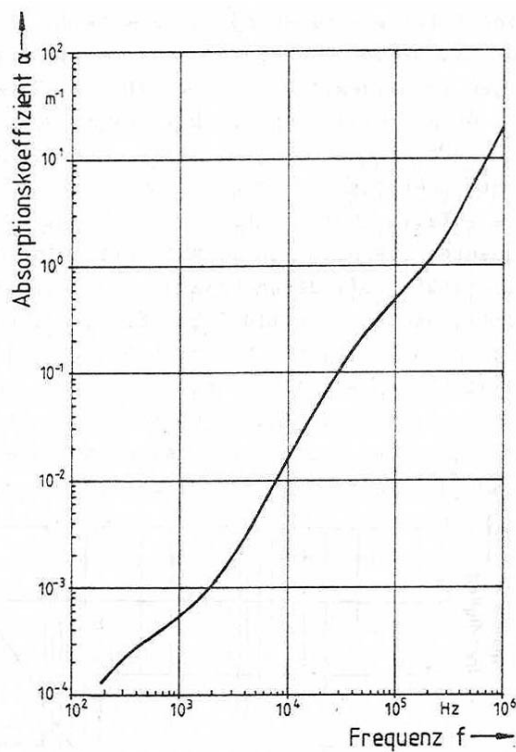
2.2.2.2 Frequenzabhängigkeit der Absorption

Die Absorption steigt mit dem Quadrat der Frequenz und kann im Ausbreitungsmedium Luft näherungsweise angegeben werden ([ZI92]):

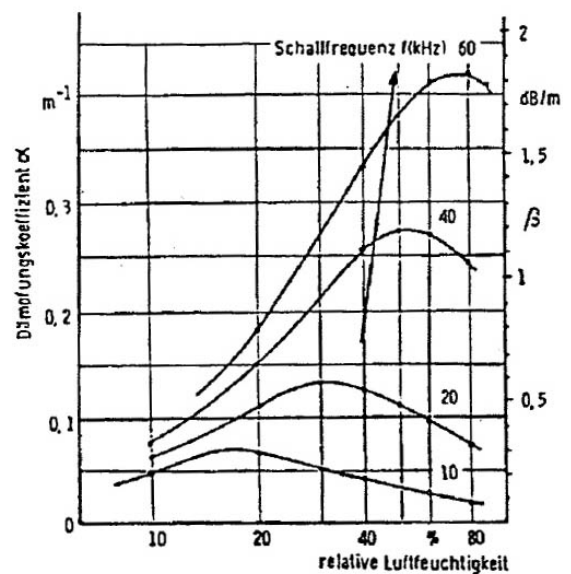
$$d_p = 1,6 \cdot 10^{-10} \cdot f^2 \quad (2.15)$$

2.2.2.3 Einfluss von Feuchtigkeit und Temperatur

Mit zunehmender Luftfeuchtigkeit steigt die Absorption gerade bei hohen Frequenzen stark an.



(a) α als Funktion der Frequenz ([EH85])



(b) α als Funktion der rel. Luftfeuchtigkeit ([ZI92])

Abb. 2.3: Absorption als Funktion der Frequenz und Luftfeuchtigkeit

2.3 Beschreibung des Schallfeldes

2.3.1 Schallausbreitung

Schallwellen breiten sich in Gasen als Longitudinalwellen aus, wobei im homogenen Medium der Ausbreitungsbereich von der Charakteristik der Quelle abhängig ist. Zum Beispiel erzeugt eine Punktquelle sphärische Wellenfronten - auch Kugelschallfeld genannt. Mit wachsenden Abstand von der Schallquelle geht das Kugelschallfeld in ein ebenes Feld über, da die Krümmung der Wellenfläche in einem betrachteten Raumwinkel stetig abnimmt. Aufgrund der kleinen Wellenlängen bei Ultraschallsignalen lassen sich diese mit einem Kolbenmembranschwinger sehr gut bündeln und somit gerichtet abstrahlen. Das erzeugte Schallfeld ähnelt einem kegelförmigen Ausschnitt aus einer Kugelwelle. Zwischen dem divergenten Fernfeld und der Kolbenfläche bildet sich das durch Interferenz gekennzeichnete Nahfeld aus, indem sich abgestrahlte Wellenzüge aus der Kolbenmitte mit Wellen vom Kolbenrand überlagern. Die Ausdehnung vom Nahfeld in Abb. 2.4 lässt sich wie folgt

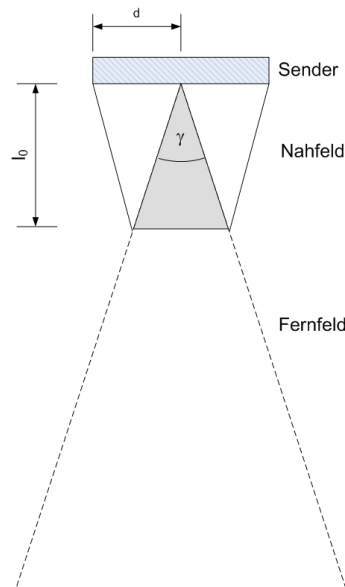


Abb. 2.4: Nah- und Fernfeld eines Kolbenmembranstrahlers

bestimmen:

$$l_0 = \frac{d^2}{2 \cdot \lambda} \quad (2.16)$$

mit:

d : Radius der Kolbenmembran
 λ : Wellenlänge der abgestrahlten Frequenz

Laut Kolbenmembrantheorie gibt es auch auf der Bezugsachse Nullstellen des Druckes (siehe Gl.2.17). Diese können auch über den herkömmlichen Bereich vom Nahfeld hinausreichen ([ZO93]).

$$r_o = \frac{n\lambda}{2} \left[\left(\frac{a}{n\lambda} \right)^2 - 1 \right] \quad (2.17)$$

$$n = 0 \dots \frac{a \cdot f}{c} \quad (2.18)$$

Für tiefe Frequenzen ($\lambda > a$) gibt es auf der Bezugsachse keine Nullstellen des Schalldruckes. Für $\lambda \leq a$ existieren n Nullstellen, wobei $n \leq a/\lambda$ erfüllt sein muss.

In der Praxis nimmt man das Nahfeld bis zu einer Entfernung an, die der doppelten größten Abmessung der Schallquelle entspricht.

Der Ausbreitungswinkel γ des Schallfeldes ¹ ergibt sich aus:

$$\gamma = \arcsin \frac{\lambda}{d} \quad (2.19)$$

Für einen handelsüblichen Ultraschallsender mit 40 kHz Resonanzfrequenz und einem Membranradius von 5 mm ergibt sich ein Nahfeld von:

$$l_0 \approx \frac{0,005^2 m^2}{2 \cdot 8,6 \cdot 10^{-3} m} \approx 2,9 \cdot 10^{-3} m \quad (2.20)$$

Für diese Schallbündel gelten im Fernfeld dann weitgehend die aus der Optik bekannten Gesetzmäßigkeiten.

2.3.2 Reflexion, Brechung, Beugung

Trifft eine Schallwelle im homogenen Medium auf ein Hindernis oder eine Grenzfläche, lassen sich drei Vorgänge beobachten: Reflexion, Brechung und Beugung. Trifft eine fortschreitende Kompressionswelle auf eine Grenzfläche zwischen zwei Medien mit unterschiedlichen Impedanzen z_1 und z_2 , so wird an dieser Stelle die Schallausbreitung gestört. Es erfolgt eine Aufspaltung der Welle in einen reflektierten und in einen weiterlaufenden Anteil. Bei senkrechtem Einfall gilt für den reflektierten Anteil r und den durchgelassenen Anteil d

$$r = \frac{p_r}{p_e} = \frac{z_2 - z_1}{z_2 + z_1} \quad (2.21)$$

$$d = \frac{p_d}{p_e} = \frac{2 \cdot z_2}{z_2 + z_1} \quad (2.22)$$

¹ auch Divergenzwinkel genannt

wobei z die Schallkennimpedanz (oder Wellenwiderstand) kennzeichnet.

$$z = \rho c \quad (2.23)$$

Das Phänomen der Reflexion wird in den verschiedensten Bereichen der Technik und Medizin angewendet, z.B. beim Echolot zur Tiefenbestimmung oder zum Auffinden von Fischschwärmen, in der Materialprüfung zur Suche von Fehlern oder bei der medizinischen Diagnose. Beim Headtracking wirken Reflexionen allerdings störend und sind daher (wenn möglich) zu vermeiden.

Trifft eine Ultraschallwelle schräg auf die Grenzfläche zwischen zwei Medien, so tritt eine Brechung des durchgelassenen Schallbündels auf (siehe Abb.2.5). Ist das 2.

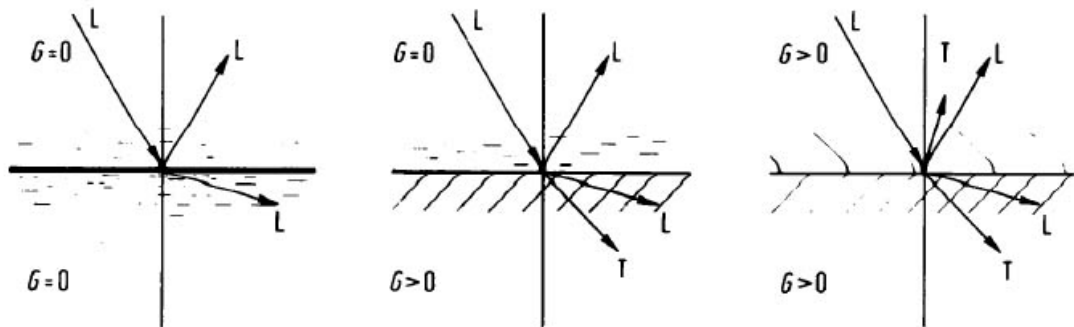


Abb. 2.5: Reflexion und Brechung einer Ultraschallwelle an der ebenen Grenzfläche zweier Medien [BR80]

Medium ein Festkörper, so tritt an der Grenzfläche eine Aufspaltung der gebrochenen Longitudinalwelle in einen Longitudinal- und einen Schubwellenanteil auf. Dabei zeigt die Schubwelle aufgrund ihrer kleineren Geschwindigkeit auch einen kleineren Brechungswinkel. Ist das 1. Medium ebenfalls ein Festkörper, so tritt auch bei der reflektierten Komponente neben der Longitudinalwelle eine zusätzliche Schubwelle auf.

Wie in der Optik existiert auch für Ultraschallwellen das Phänomen der Totalreflexion bei Erreichung der jeweiligen Grenzwinkel. Bei stetiger Vergrößerung des Einfallswinkels wird zuerst der Grenzwinkel für die Longitudinal- und danach der für die Schubwelle erreicht. Der Winkelbereich, bei dem sich nur eine Schubwelle im 2. Medium ausbreitet, hat erhebliche messtechnische Bedeutung. Beim Grenzwinkel für die Schubwelle bildet sich nur noch eine Oberflächenwelle (Rayleigh-Welle) in der Grenzschicht aus.

Eine Streuung der Schallwelle erfolgt, wenn die Dichte oder die elastischen Konstanten des Ausbreitungsmediums ortsabhängig sind. Die Amplitude des Streusignals ist eine Funktion des Streukoeffizienten sowie der Größe und Art des einfallenden Schallsignals. Der Streukoeffizient hängt von dem Verhältnis Wellenlänge zur Abmessung des Streuzentrums ab. Abb.2.6 zeigt die in der Raumakustik am häufigsten zum Erzielen

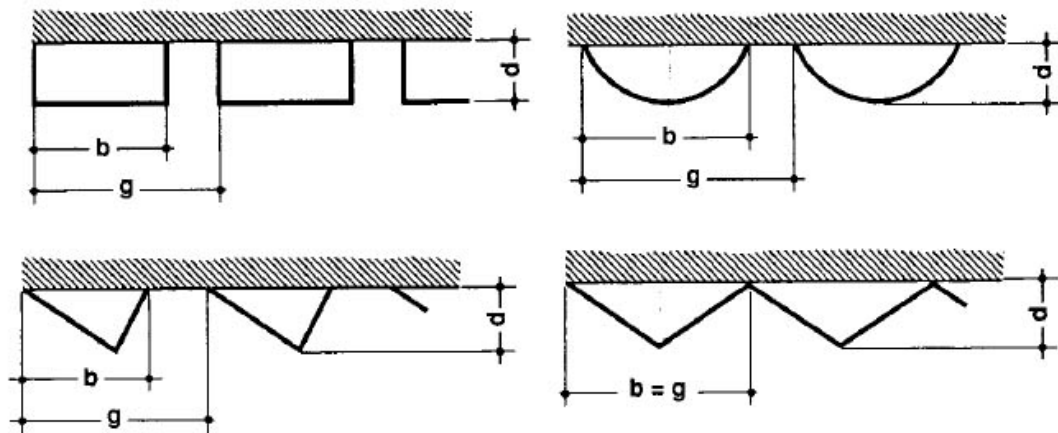


Abb. 2.6: Beispiele geometrischer Strukturen für diffuse Reflexionen ([FA98])
 b Strukturbreite d Strukturtiefe g Strukturperiode

einer hohen Diffusität eingesetzten geometrischen Grundformen ([FA98]). Der höchste Diffusitätsgrad (ungefähr 0,8) wird in dem Frequenzbereich erzielt, in dem die Strukturperiode g etwa ein- bis zweimal so groß ist wie die Wellenlänge. Die Frequenz mit der stärksten Diffusität f_{opt} ergibt sich mit der Näherungsformel 2.24. Die Strukturbreite b sollte hierbei etwas mehr als die Hälfte der Strukturperiode groß sein.

Damit ein Ultraschallsignal von 40 kHz aufgrund der Oberflächenstruktur gestreut wird, sind nach Gl.2.24 nur 1.25 cm Strukturperiode und etwa 0.8 cm Strukturtiefe notwendig. Für Ultraschallfrequenzen wirkt also praktisch fast jede Oberflächenstruktur als Diffusor. Alle strukturlosen Flächen (wie z.B. Glasscheiben) verursachen gerichtete (Einzel-)Reflexionen die aufgrund ihrer hohen Energie für diese Applikation sehr störend sind.

$$f_{opt} = \frac{500}{g} \quad (2.24)$$

$$0.3 \cdot b \leq d \leq 0.5 \cdot b \quad (2.25)$$

g : Strukturperiode in Meter
 b : Strukturbreite
 d : Strukturtiefe

Mit abnehmender Ausdehnung des Objektes im Schallfeld bis zur Größenordnung einer Wellenlänge geht die Reflexion in Beugung der Welle über. Hierbei wird das Objekt von den Schallwellen umgangen - die Schallausbreitung ist nicht mehr gradlinig. Der Schalldruck der reflektierten Welle sinkt ab und die Randzonen verlieren durch Beugung die bislang gradlinige Kontur. Diese Erscheinungen lassen sich in allgemeiner Form durch das Kirchhoffsche Beugungsintegral beschreiben (siehe [EH85]). Wird der Ultraschall auf dem Weg vom Sender zum Empfänger gebeugt, beeinflusst dies die Detektionsgenauigkeit, da immer von einer gradlinigen Schallausbreitung ausgegangen wird.

2.3.3 Bewegte Schallquellen - Dopplereffekte

Tritt während der Abstandsmessung zwischen Sender und Empfänger eine Relativbewegung auf, so verschiebt sich aufgrund des Dopplereffektes die empfangene Frequenz. Der Effekt ist umso stärker, je schneller die Relativgeschwindigkeit v ist ([SR02]).

$$\Delta f = \pm \frac{v}{c - v} f \quad (2.26)$$

Schnelle Kopfbewegungen¹ können so Frequenzverschiebungen Δf von bis zu 35 Hz bewirken. Einige der in Kapitel 5 vorgestellten Detektionsmethoden werden von diesem Effekt negativ beeinflusst, wo hingegen jenes Verfahren in Abschnitt 5.5 diesen Effekt nutzt.

2.4 Gesundheitliche Aspekte

In den 80er Jahren ist über mögliche gesundheitliche Schädigungen durch Ultraschall viel diskutiert und veröffentlicht worden² und es steht außer Frage, dass es bei hohen Ultraschallbelastungen am Menschen zu gesundheitlichen Schäden kommen kann.

Als mögliche Mechanismen für biologische Effekte durch Ultraschalleinwirkung sind die chemische Veränderungen bewirkende Kavitation, das Zerreißen sehr langer Moleküle bei hohen Frequenzen und die Temperaturerhöhung aufgrund der Absorption im beschallten Gewebe anzusehen. Das American Institute for Ultrasound in Medicine (AIUM) untersuchte 1976 die Einwirkung von Ultraschallstrahlung auf den Menschen. Danach sind bei allen Untersuchungen keine signifikanten biologischen Effekte beobachtet worden, wenn die Intensität in dem beschallten Gewebe kleiner als 100 mW/cm^2 war. Weiterhin zeigten sich für Beschallungszeiten zwischen 500 sec und 1 sec auch bei höheren Intensitäten keine Effekte, wenn das Produkt aus Intensität und Beschallungszeit kleiner als 50 W s/cm^2 war. Bei der Intensität handelt es sich um den zeitlich gemittelten räumlichen Maximalwert.

Zu den gelegentlich auftretenden Klagen über Kopfschmerzen, Unwohlsein und Schwindelanfälle durch hohe Ultraschallintensitäten in Luft wird die Ansicht vertreten, dass hierfür die relativ hohen Pegel der subharmonischen Komponenten im Hörbereich als Ursache anzusehen sind. Daraus ergeben sich für den betroffenen Personenkreis auch die notwendigen Schutzmaßnahmen. Bis zu Ultraschallpegeln von 110 dB - 120 dB scheinen noch keine Verluste im Hörvermögen oder sonstige physiologische Veränderungen bei den betroffenen Personen aufzutreten ([BR80]).

¹Bei einer Winkelgeschwindigkeit von $360^\circ/\text{sec}$

²z.B. am American Institute for Ultrasound in Medicine (AIUM) oder der Weltgesundheitsorganisation WHO

Kapitel 3

Geometrische Grundlagen

3.1 Situation

Die Grundlage zur Positionsbestimmung ist die exakte Messung der Laufzeitdifferenz¹. Diese Laufzeit entsteht, wenn von einem Sender das ausgesendete Signal an zwei, zum Sender unterschiedlich entfernten Empfängern eintrifft (siehe Abb.3.1). Eine Laufzeit t_L ist über die Schallgeschwindigkeit c proportional zur der entsprechenden Wegstrecke w_L .

$$t_L = \frac{w_L}{c} \quad (3.1)$$

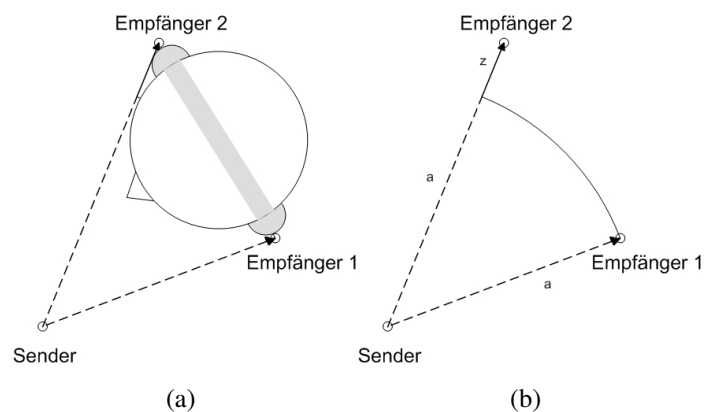


Abb. 3.1: Laufzeitenmessung bei Trackinganwendungen

¹Methoden dazu im Kap. 5

Die geometrischen Modelle im folgenden Kapitel gehen davon aus, dass Rotationspunkt des Kopfes, Ohren und Sensoren auf einer Achse liegen. In Abb.3.2 sind die Positionen des Rotationspunktes des Kopfes und der Gehörgänge markiert. Verbindet man beide Ohren mit einer Linie, so zeigt sich, dass die Ohren sich etwas hinter dem Rotationspunkt befinden. Diese Abweichung ist jedoch sehr gering und wird im Folgenden vernachlässigt¹.

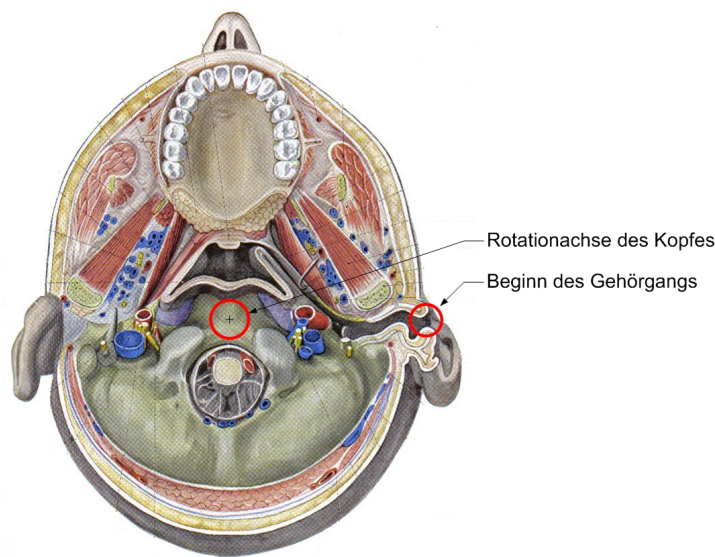


Abb. 3.2: Horizontalschnitt des Kopfes in der Ebene der oberen Zahnreihe (von kaudal gesehen, [HA69])

Kapitel 1.2.1 beschreibt ein bereits bestehendes Headtrackingsystem auf Ultraschallbasis. Dieses System besitzt neben der akustischen (Ultraschallstrecke) auch eine elektrische (kabelgebundene) Verbindung zwischen den Systemkomponenten. Es handelt sich hierbei um eine bidirektionales System, womit es möglich ist, die absoluten Laufzeiten², und damit auch die absoluten Wegstrecken bezeichnet zwischen Sender und Empfänger zu messen (siehe Gl.1.1). Nach Abb.3.1 sind diese Wegstrecken a bzw. $a + z$.

Wenn man diese elektrische Rückleitung vermeiden möchte (der Headtracker soll ja schnurlos sein), sind den Empfängern der exakte Aussendezeitpunkt des Sendesignals unbekannt und es kann nur die Wegstrecke b aus Abb.3.1 ermittelt werden. Es stellt sich die Frage, ob die Kenntnis der Laufzeitdifferenz für die eindeutige, genaue Kopfwinkelbestimmung genügt. Voraussetzung dafür ist ein geschicktes Positionieren aller Komponenten der Sendestrecke (Sender und Empfänger).

Ist das eine Ende der Ultraschallsendestrecke am Bildschirm und das andere Ende der Ultraschallsendestrecke am Kopfhörer installiert, verringert sich der mögliche Positionsbereich auf einen Halbraum, da sich der Kopfhörer im Betrieb immer nur *vor* dem Videobildschirm befinden wird (niemals dahinter).

¹signifikantere Winkelfehler durch Fehlpositionierung der Sensoren siehe Kap. 6.4.2 auf Seite 74

²Aus der Literatur auch als *time of flight* (T.O.F.) bekannt.

In diesem Kapitel wird gezeigt, welche Detektionsmöglichkeiten und konstruktionsbedingte Ungenauigkeiten verschiedene Sensoranordnungen aufweisen.

3.2 Absolute und relative Orientierung

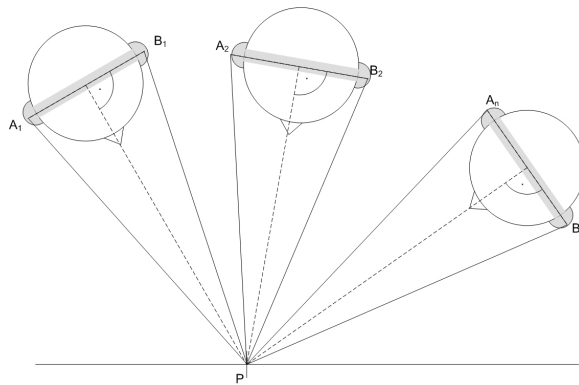


Abb. 3.3: Relative Orientierung - Skizze von drei Hörern, die Sitzposition ist unklar

Die Abb.3.3 zeigt schematisch 3 Hörer, die von unterschiedlichen Punkten hin zum Bezugspunkt P orientiert sind. Die Punkte A_i und B_i sollen dabei die Ohren darstellen, weiterhin ist $\overline{A_i B_i}$ immer gleich groß. Wie nach Abb.3.1 können die Laufzeitdifferenzen D_i aller 3 Hörpositionen ermittelt werden:

$$\overline{A_1 P} - \overline{B_1 P} = \overline{A_2 P} - \overline{B_2 P} = \overline{A_n P} - \overline{B_n P} = 0$$

Es ergeben sich für alle 3 Hörer gleiche Laufzeitdifferenzen - die Hörer haben somit die gleiche relative Orientierung zum Bezugspunkt P , unabhängig zur Entfernung von P und der Position im Raum. Detektiert man ausschließlich diesen relativen Winkel und übergibt ihn an das Binauralmodul, so gleichen sich an allen 3 Hörpositionen die aufbereiteten Audiosignale, da zum Binauralisieren die gleiche Außenohrübertragungsfunktion (head related transfer function, HRTF) angewendet wird. Die Orientierung des Oberkörpers des Hörers wird somit nicht berücksichtigt. Diese Art von Detektierung scheint aus heutiger Sicht ausreichend zu sein, da bis dato eine Abhängigkeit des Torsos auf die Spatialisierung zwar bestätigt ist, aber in Form von Aussenohrübertragungsfunktionen noch unzureichend erfasst wurde.

Ist es möglich die absolute Lage im Raum zu detektieren, kann man diesem Fall Rückschlüsse auf die Sitzposition und die Torsolage ziehen, wenn davon ausgegangen wird, dass der Hörer sich in einer aufrechten, angelehnten Haltung befindet¹. Eine ungefähre Berücksichtigung des Oberkörpers, und damit eine realistischerer Signalaufbereitung mittels exaktere HRTFs wäre also möglich (siehe Abb.3.4).

¹Diese Annahme ist gerade bei angeschnallten Passagieren auf einer Fahrzeuggückbank zulässig

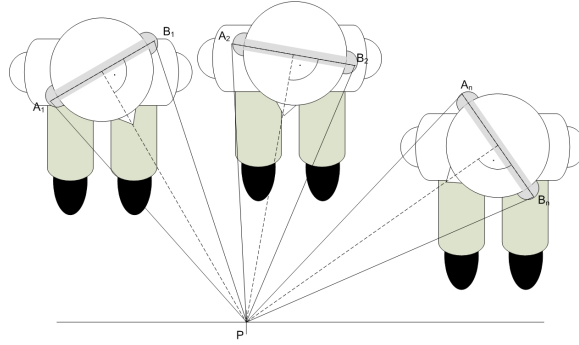


Abb. 3.4: Absolute Orientierung im Raum - Skizze mit drei Hörern, die Sitzposition und die Lage des Torsos kann erahnt werden

3.3 Verschiedene Geometrien

Die im Folgenden vorgestellten geometrischen Ansätze gehen davon aus, dass sich die Empfänger am Kopfhörer befinden. Hierbei u.a. auch der Material- und Berechnungsaufwand der erzielbaren Positionsbestimmbarkeit gegenübergestellt werden.

Die Geometrien sind natürlich ebenso gültig, wenn am Kopfhörer statt der Empfänger mehrere Ultraschallsender positioniert werden. Dies hat vor allem praktische Auswirkungen, deren Vor- und Nachteile in Kap.6.4 gegenübergestellt werden.

3.3.1 Zwei Sender -Vier Empfänger

Positioniert man 4 Empfänger (E_1 bis E_4) an den Mittelpunkten der Seiten eines Rechteckes nach Abb.3.5 kann vom Mittelpunkt des Empfängerarrays eine Gerade g im Raum aufgestellt werden, auf der sich ein Sender S annähernd befinden muss. Dieser Ansatz ist praktisch die vorgestellte Geometrie nach Kap.3.3.5 für den dreidimensionalen Raum.

$$E_1 = \begin{pmatrix} -x \\ 0 \\ 0 \end{pmatrix}, E_2 = \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 \\ -y \\ 0 \end{pmatrix}, E_4 = \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix}$$

$$\cos(\alpha) = \frac{\overline{SE_1} - \overline{SE_2}}{2x}, \cos(\beta) = \frac{\overline{SE_3} - \overline{SE_4}}{2y}, \cos(\gamma) = \sqrt{1 - \cos^2(\alpha) - \cos^2(\beta)} \quad (3.2)$$

$$g = \mu h, h = \begin{pmatrix} \cos(\alpha) \\ \cos(\beta) \\ \pm \cos(\gamma) \end{pmatrix} \quad (3.3)$$

Werden von zwei horizontal, fest installierten Sendern die Einfallsrichtung bestimmt, kann mit den vier Empfängern am Kopf neben dem Azimutwinkel (x-Richtung) auch ein Elevationswinkel (Y-Richtung) bestimmt werden. Dadurch lassen sich Nick-, Dreh- und

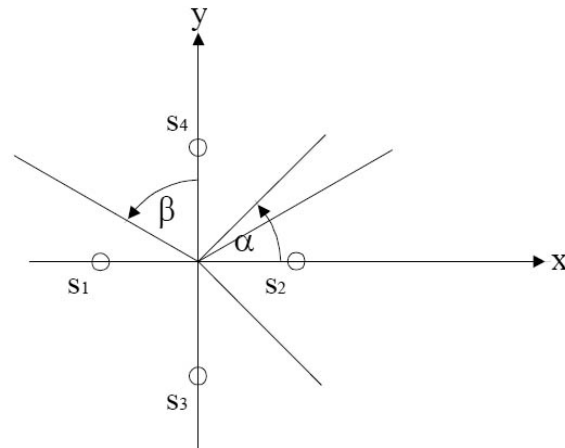


Abb. 3.5: Vier Empfänger zur Bestimmung der Einfallsrichtung im Raum, ([WU97])

Rollbewegungen des Kopfes auseinander halten. Die Positionierung von 4 Empfängern am Kopfhörer nach Abb.3.5 ergibt neben konstruktionsellen Schwierigkeiten wahrscheinlich auch ästhetische Probleme.

3.3.2 Drei Sender - Drei Empfänger

Befinden sich die Sender bzw. Empfänger jeweils an den Eckpunkten eines aufgestellten Dreiecks, kann bei Kenntnis der Laufzeiten jeder Sendestrecke, eine eindeutige Lösung im Halbraum ermittelt werden ([LO92], [BH93] oder [EB02]). Sind jedoch nur Laufzeitdifferenzen bekannt, ergibt sich ein uneindeutiger, dreidimensionaler Lösungsraum.

Mit sinkender Zahl der Sensoren werden auch die Möglichkeit zur Positionsbestimmung geringer. Mit den folgenden Lösungsvorschlägen ist keine Positionsbestimmung im Raum mehr möglich. Ziel ist daher, die Position in einer Ebene zu bestimmen.

3.3.3 Zwei Sender - Zwei Empfänger

In dieser Anordnung befinden sich 2 Sender auf der Bezugsachse und 2 Sender am Kopfhörer. Damit können insgesamt vier Laufzeitdifferenzen ermittelt werden.

$$\overline{S1E1} - \overline{S2E1} = D_1 \quad (3.4)$$

$$\overline{S1E2} - \overline{S2E2} = D_2 \quad (3.5)$$

$$\overline{S1E1} - \overline{S1E2} = D_3 \quad (3.6)$$

$$\overline{S1E2} - \overline{S2E2} = D_4 \quad (3.7)$$

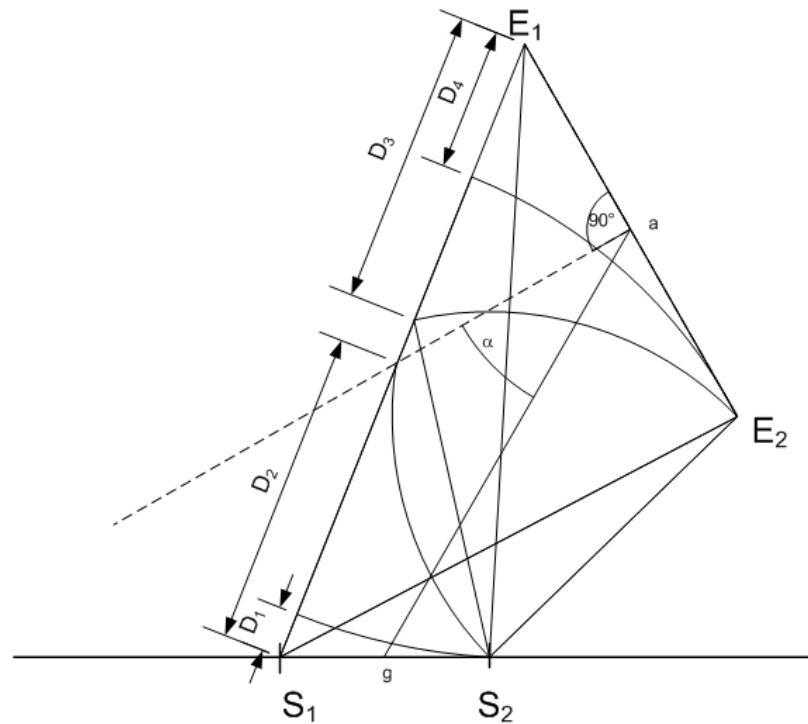


Abb. 3.6: Skizze für die Anordnung von zwei Empfängern am Kopfhörer und zwei Sendern auf der Bezugsachse

Aus diesen Laufzeitdifferenzen lässt sich keine eindeutige Lösung für eine absolute Positionsbestimmung finden. Die Lösungen jedes einzelnen Senders liegen auf einer Hyperbel. Das Problem ist, dass mittels 2 Empfänger nicht auf die Distanz zwischen einem Sender und dem Empfängerpaar geschlossen werden kann. Es kann jedoch wie in Abschnitt 3.3.5 die relative Orientierung des Kopfhörers zu jedem einzelnen Sender bestimmt werden.

3.3.4 Ein Sender - Drei Empfänger

Verwendet man kopfhörerseitig 3 Sensoren, so ist der Ort der Quelle der Schnitt von 2 Hyperbeln. Es kann der Abstand r und der Drehwinkel α bestimmt werden. Über dem Kopfhörermittelpunkt befindet sich der zentraler Ultraschallempfänger B . Die seitlichen Empfänger A und C haben zueinander den bekannten Abstand a um den Kopfmittelpunkt B . Der Sender liegt im Bezugspunkt E . Somit spannen sich 3 Dreiecke zwischen Sender und Empfänger auf:

$$\triangle ABE$$

$$\triangle ACE$$

$$\triangle BCE$$

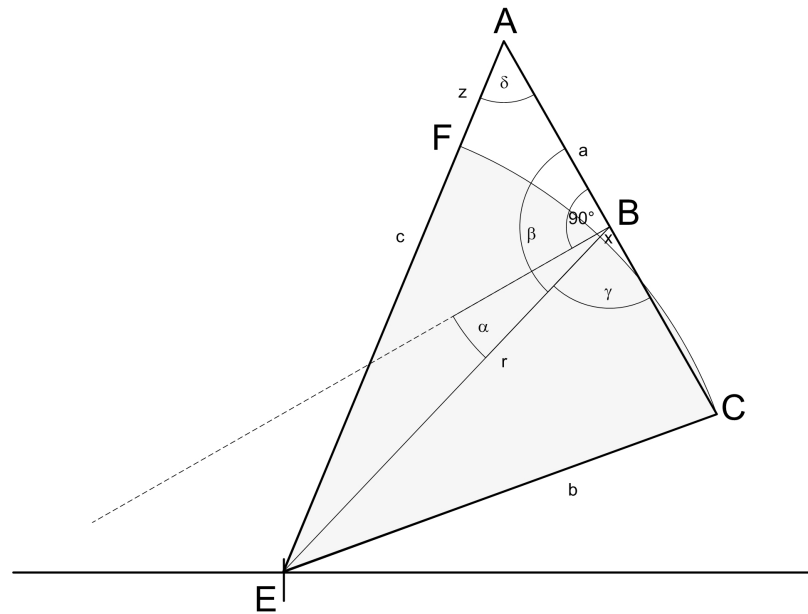


Abb. 3.7: Skizze für die Anordnung von drei Sensoren (A, B, C) am Kopfhörer

Unter der Annahme, dass die Laufzeitdifferenzen

$$x = \overline{CE} - \overline{BE} \quad (3.8)$$

$$z = \overline{AE} - \overline{CE} \quad (3.9)$$

(3.10)

detektiert werden und

$$\overline{AB} \stackrel{!}{=} \overline{BC}$$

gilt, ergeben sich folgende, elementare geometrische Beziehungen:

$$\gamma + \alpha = 90^\circ \quad (3.11)$$

$$\beta - \alpha = 90^\circ \quad (3.12)$$

$$\beta + \gamma = 180^\circ \quad (3.13)$$

$$b+z = c \tag{3.14}$$

sowie unter Anwendung des Cosinussatzes:

$$c^2 = \left(\frac{a}{2}\right)^2 + r^2 - ar \cos \beta \quad (3.15)$$

$$b^2 = \left(\frac{a}{2}\right)^2 + r^2 - a r \cos \gamma \quad (3.16)$$

$$r^2 = \left(\frac{a}{2}\right)^2 + c^2 - ac \cos \delta \quad (3.17)$$

$$b^2 = a^2 + c^2 - 2ac \cos \delta \quad (3.18)$$

Aus Gleichungen 3.11, 3.12, 3.14, 3.15, 3.16, ergibt sich:

$$\sin \alpha = \frac{z^2 + 2bz}{2ar} \quad (3.19)$$

Es wird vorausgesetzt:

$$\overline{AB} \stackrel{!}{=} \overline{BC} \stackrel{!}{=} \text{bekannt}$$

Die Distanz r ergibt sich nach Gl.3.20 (übernommen aus [CA81])

$$r = \frac{a^2 \left[1 - \frac{2z^2 - 4zx}{a^2} \right]}{4(2x - z)} \quad (3.20)$$

Gleichung 3.20 eingesetzt in Gl.3.19 ergibt den Drehwinkel α .

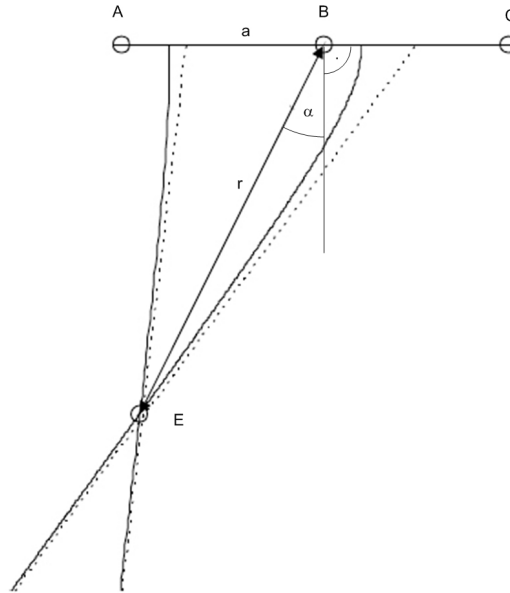


Abb. 3.8: Lösungsmenge bei drei ko-linearen Sensoren

3.3.5 Ein Sender - Zwei Empfänger

Ähnlich dem vorherigen Kapitel verhält es sich bei folgender Anordnung. Beschränkt man sich auf 2 Empfänger am Kopfhörer, ergibt sich die Geometrie in Abb.3.9. Mittels dieser Anordnung ist eine Bestimmung der relativen Kopfhörerorientierung in der Ebene möglich. Die Herleitung der Geometrie ist ganz ähnlich zum vorherigen Abschnitt. Die Elimination von δ in Gl.3.17 und Gl.3.18 führt unter Berücksichtigung der Beziehung 3.14 zu einer quadratischen Lösung für b .

$$b_{1/2} = \frac{-z}{2} \pm \sqrt{\frac{-z^2}{4} + \frac{a^2}{4} + r^2} \quad (3.21)$$

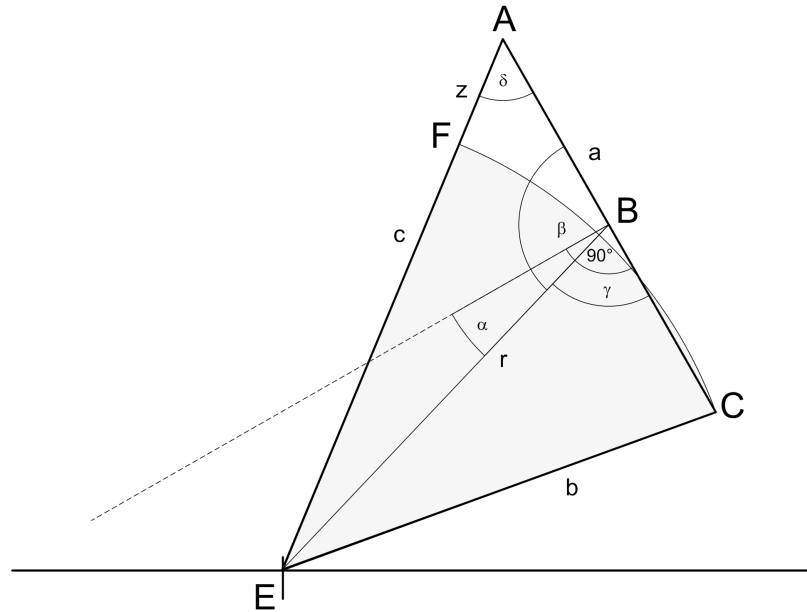


Abb. 3.9: Skizze für die Anordnung einem Sender (E) und zwei Empfängern (A , C)

Die positive Lösung von Gleichung 3.21 in 3.19 ergibt schließlich

$$\sin \alpha = \frac{z \cdot \sqrt{-z^2 + a^2 + 4r^2}}{2 \cdot ar} \quad (3.22)$$

Gleichung 3.22 enthält neben den bekannten Größen a (Abstand der Empfänger) und z (Laufzeitdifferenz) die entfernungsabhängige Unbekannte r . Sie entspricht der Distanz zwischen Referenzpunkt und dem Rotationsmittelpunkt des Kopfes. Dadurch wird die mögliche Position des Kopfhörers auf eine Hyperbel begrenzt (siehe Abb.3.10). Der Verlauf der Hyperbel kann mit Gl.3.23 asymptotisch angenähert werden: Mit der Näherung, $r = b + \frac{1}{2}z$ ergibt sich die exakte Lösung für eine unendlich große Entfernung zwischen Kopfhörer und Referenzpunkt. Hierbei fällt die Variable r heraus.

$$\alpha' = \arcsin\left(\frac{z}{a}\right) \quad (3.23)$$

Die Größe des Fehlers durch diese Näherung ist vom Winkel α abhängig und wird mit zunehmender Entfernung des Kopfhörers zum Bezugspunkt immer geringer (asymptotischen Approximation, siehe dazu Abb.3.11).

Um die relative Kopfhörerorientierung genau detektieren zu können, ist nach Gl.3.22 also die Bestimmung der Strecke zwischen dem Bezugspunkt E und dem Rotationspunkt des Kopfhörers (Punkt B) notwendig.

Diese Strecke könnte ungefähr anhand der Signalstärke des empfangenen US-Signals über die Dämpfungsverluste¹ geschätzt werden. Da die empfangene Signalstärke auch von der

¹ siehe dazu Kap. 2.2.2 auf Seite 11

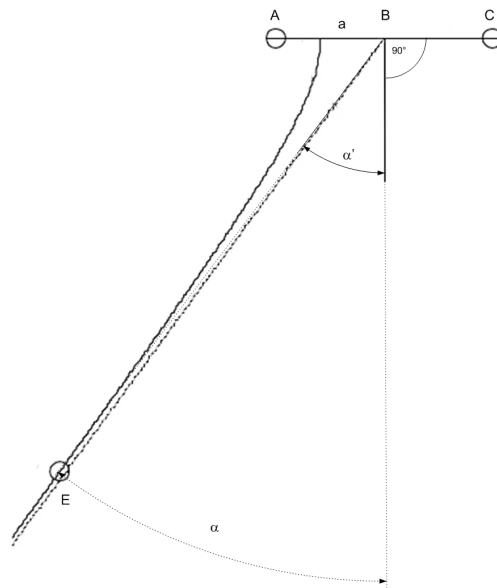


Abb. 3.10: exakte Lösung und Näherung mit zwei Sensoren am Kopfhörer

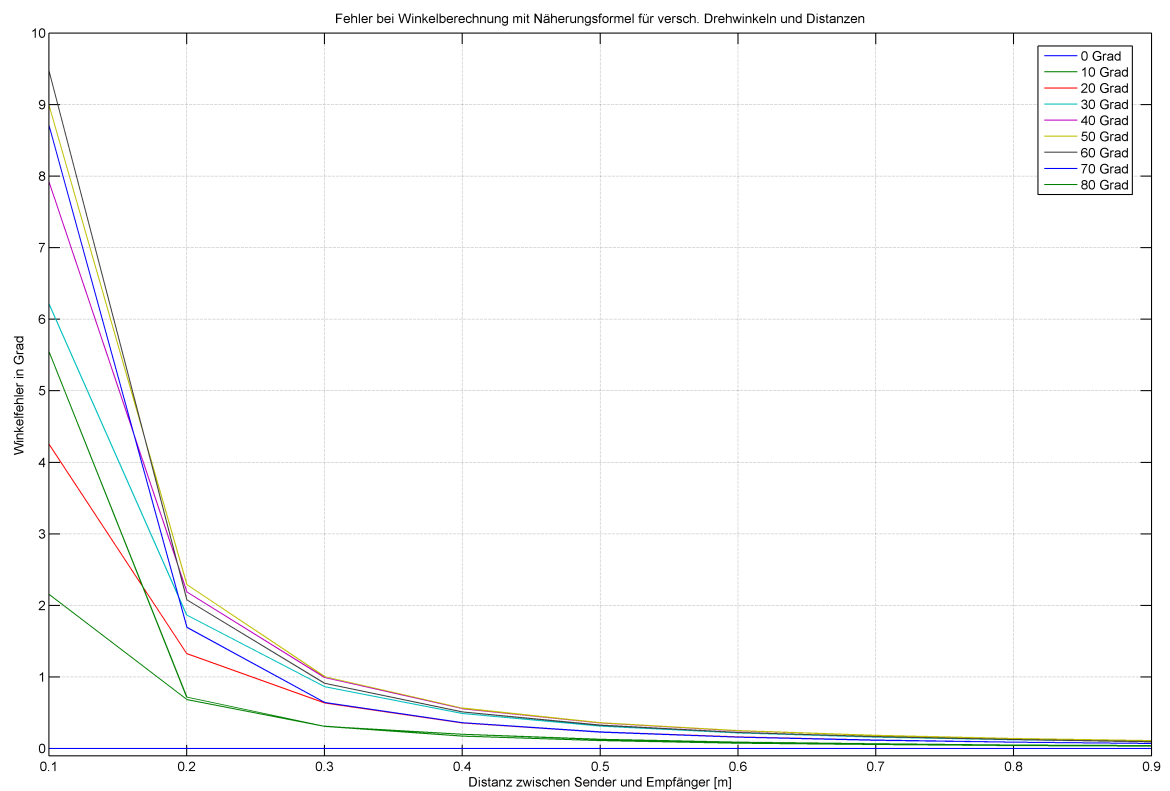


Abb. 3.11: Fehler der Näherungsformel bei versch. Drehwinkeln und Distanzen

Richtcharakteristik der verwendeten Sensoren abhängt, ist dieser Ansatz praktisch nicht umsetzbar.

3.3.6 Ein Sender - Ein Empfänger

Der Versuch den Materialaufwand auf nur eine Sensorstrecke zu beschränken, führt zu einer Anordnung welche keine Laufzeitdifferenz und damit auch keine Kopfwinkelbestimmung zulässt. Mit einer Intensitätsmessung könnte lediglich der Abstand der beiden Sensoren geschätzt werden. Dies ist allerdings sehr ungenau, da Interferenzen und die Richtcharakteristik der Sensoren die Messung beeinflussen. Von weiteren Erklärungen wird hier abgesehen, Informationen zur Intensitätsmessung finden sich in [\[WU97\]](#).

Kapitel 4

Signalformen

4.1 Allgemein

Zum Betreiben der Ultraschallstrecke des Headtrackers können verschiedene Sendesignale verwendet werden. Auf den ersten Blick lässt die Wahl des Signaltyps einem also sehr viele Freiheiten. Bei genauerer Betrachtung stellt man fest, dass die verschiedenen Signaltypen auf ihrem Weg vom Sender zum Empfänger durch Störungen unterschiedlich stark verzerrt werden (siehe Abb.4.1). Diese Störungen ergeben sich aus dem Übertragungsverhalten der verwendeten Sensoren sowie aus den Einflüssen während der Schallausbreitung¹. Werden die Sender² am Kopfhörer betrieben, sollten die begrenzten Energiekapazitäten beachtet werden. Hierbei wäre eine Signalform sinnvoll, die nur eine geringe Verstärkungsleistung benötigt. Die Aufgabe ist es, für den Anwendungsfall ein geeignetes Sendesignal zu wählen. Dieses Kapitel stellt einige Signaltypen vor und zeigt deren Vor- und Nachteile. Weiterhin wird im Hinblick auf den Einsatz mehrerer getrennt-parallelbetriebener Headtracker Möglichkeiten zur Signalkodierung erläutert.

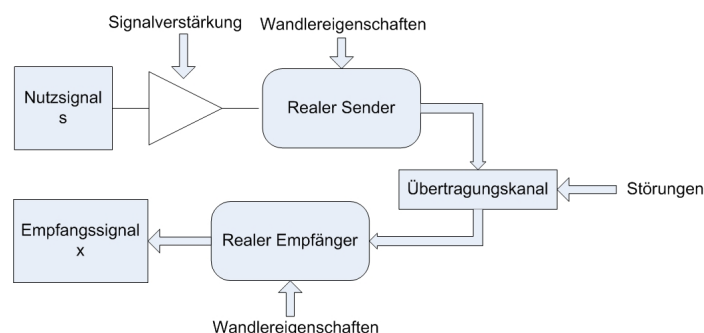


Abb. 4.1: Strukturbild des realen Übertragungspfades

¹ siehe dazu Kap. 2 auf Seite 7

² siehe dazu Kap. 6.4 auf Seite 72

4.2 Äußere Form

Continuous Wave (CW): Hierbei wird vom Ultraschallsender permanent ein Ton ausgesendet. Diese „Dauerbeschallung“ benötigt zum einen senderseitig viel Energie, zum anderen kommt es während der Schallausbreitung zwangsläufig zu Überlagerung mit Echo-signalen (Freifeld ausgeschlossen). Eine empfangsseitige Echofilterung ist nur teilweise oder mit großem Aufwand möglich¹. Neben Reflexionen ist dieser Signaltyp auch extrem beeinflussbar durch Dopplereffekte². Die Detektionsmethode aus Kap.5.5 nutzt genau dies aus.

Dauertöne besitzen eine gute Störunterdrückung gegenüber Rauschen. Ein weiterer Vorteil dieser Signalform ist der geringe Bandbreitenbedarf und damit eine geringe mechanische Beanspruchung der Sensoren (keine hörbaren Nebenfrequenzen). Die geringe Bandbreite erleichtert zudem das Betreiben mehrerer Systeme in unmittelbarer Nachbarschaft: verwenden die Systeme unterschiedlichen Arbeitsfrequenzen, könnte ein empfangsseitiger Bandpass (oder ein adaptiver Notchfilter) das gewünschte Signal herausfiltern.

Da außerdem das Sendesignal ununterbrochen am Empfänger vorliegt, kann eine Winkeldetektion zu jedem Zeitpunkt und ohne Synchronisierung von Sender und Empfänger erfolgen.

Nadelimpuls: Abgeleitet vom *Dirac-Impuls* benötigt der Nadelimpuls aufgrund des raschen Ein- und Ausschaltvorgangs die volle zur Verfügung stehende Bandbreite des Sensors³. Dadurch kommt es zu einer hohen mechanischen Beanspruchung der Sender, was sich in hörbare Nebenfrequenzen widerspiegelt. Aufgrund des Ein- und Ausschwingverhaltens realer Sender kommt es zu einer Verschmierung des Impulses. Generell kann man durch einen einzelnen Nadelimpuls nur wenig Energie transportieren ohne das System zu überansprechen (Verzerrungen). Auch aufgrund ihrer geringen Länge ist die Störunterdrückung schlecht. Dies führt zu einer schwierigeren Detektierbarkeit bei größeren Distanzen. Ein Vorteil der geringen Zeitdauer ist die gute Partialechoseparation. Das macht diesen Signaltypen für kleine Innenräume, in denen schnelle erste Reflexionen zu erwarten sind, interessant. Der Leistungsbedarf ist geringer als bei *Continuous Waves*.

Gated wave (GW): Gated Waves sind Schwingungspakete von kurzer Dauer, bestehend aus ein oder mehreren Sinustönen. Das breitbandige Spektrum, bedingt durch die transienten Ein- und Ausschaltzeitpunkte, kann über weiche fade-in und fade-outs (z.B. *Hanning* oder *Blackman*) kontrolliert und somit gering gehalten werden. Diese Amplitudengewichtungen verringern jedoch die effektive Signallänge und dabei auch die übertragbare Signalenergie. Eine Verlängerung der Gesamtdauer kann dem entgegenwirken. Zusammenfas-

¹ siehe [WE97], [ZI92], [KU02]

² Kap. 2.3.3 auf Seite 18

³ Multiplikation mit rect-Funktion im Zeitbereich führt zu spektralen Faltung mit sinc-Funktion

send kann man sagen, dass das *Gated Wave* die Vorteile von *CW* und *Nadelimpuls* kombiniert.

4.3 Innere Form, Kodierungsverfahren

Chirp-Signale: Ein Chirp ist frequenzmoduliertes Schwingungspaket, wobei die Frequenzmodulation einer definierten Funktion folgt (meist linear oder exponentiell). Die Bandbreite kann über den Frequenzverlauf des Chirps geregelt werden. Chirpsignale zeichnen sich durch besonders günstige Eigenschaften bei der Verwendung in bandbegrenzten Übertragungssystemen aus, denn die Energie die im Chirp enthalten verteilt sich zeitversetzt über das gesamte bandbegrenzte Frequenzband. Damit kann das Signal effektiv mehr Energie enthalten, was zu einem höheren SNR während der Signalübertragung führt. Chirps sind immun gegenüber nichtlinearen Verzerrungen und besitzen gegenüber einem herkömmlichen Sinussignal in Abhängigkeit von der benutzten Bandbreite eine schärfer ausgeprägte Autokorrelationsfunktion ([ZI92]).

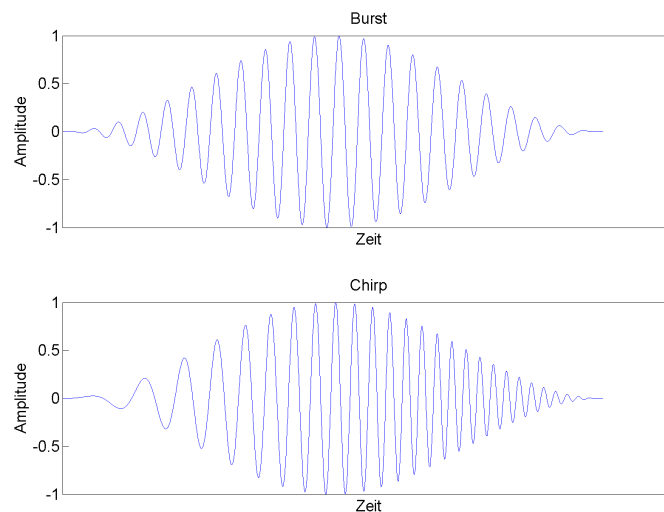


Abb. 4.2: Chirp- und Burstsignal im Vergleich, mit Hanningfenster gewichtet

Burst-Signale: Im Vergleich zu einem *Chirp*-Signal ist die Frequenzinformation des Burstes konstant (siehe Abb.4.2). Der Bandbreitenbedarf ist abhängig von der verwendeten Fensterfunktion.

pseudozufälliges Rauschen: Pseudozufälliges Rauschen ist ein algorithmisch erzeugtes Rauschen, das (wie das weißem Rauschen auch) völlig dekorreliert ist:

$$G_{x,x}(n) = \delta(n) \quad (4.1)$$

Zufallsfolgen werden mit einem Zufallsgenerator erzeugt. Aufgrund von Initialisierungsparametern werden anhand einer Iterationsvorschrift eine Zufallszahlenfolge erzeugt, die sich nicht mehr von einer gleichverteilten Zufallsfolge unterscheiden lässt. Es ist zu sehen, dass die Autokorrelationsfunktion $G_{x,x}$ eines pseudozufälligen Signal immer dann den Wert 1 besitzt, wenn die Zufallsfolge um eine Periode verschoben mit sich selbst korreliert wird. Ein Beispiel für einen Generator ist ein *lineares rückgekoppeltes Schieberegister*: Das Schieberegister besitzt die Länge n und enthält die Funktion f :

$$f(x_{n-1}, \dots, x_0) = c_{n-1}x_{n-1} + \dots + c_0x_0 \quad (4.2)$$

mit

$$c_{n-1}, \dots, c_0 \in F$$

c_j werden als Rückkopplungskoeffizienten bezeichnet. Es gilt also die Bildungsvorschrift:

$$a_j = \sum_{i=0}^{n-1} c_j a_{j-n+i} \quad (4.3)$$

Im binären Fall ($F = 0, 1$) lässt sich das Schieberegister wie in Abb.4.3 darstellen. Ein binäres lineares Schieberegister der Länge n erzeugt für jeden Anfangszustand $\neq 0$ eine periodische Folge mit der Periodendauer $L = 2^n - 1$. Bei dieser Art von Folgen ist es möglich, aus einer Teilfolge der Länge $2n$ (deutlich kleiner als die Periodendauer L) die gesamte Folge zu ermitteln ([HA05]). Sind die Initialisierungsparameter und die Iterationsvorschrift bekannt, ist das Wiedererkennen der Folge natürlich noch einfacher.

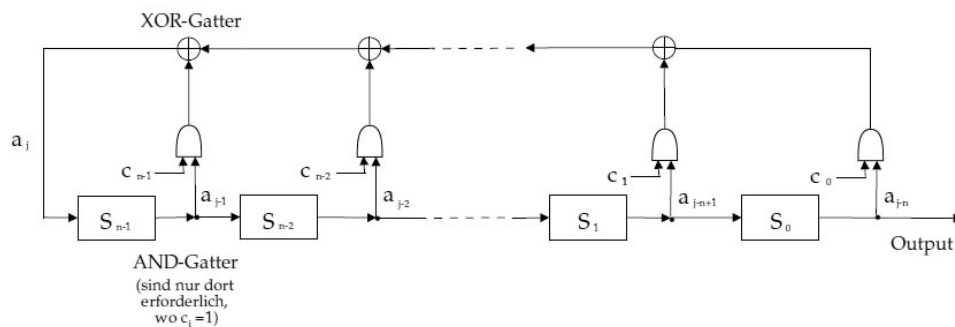


Abb. 4.3: binäres lineares Schieberegister der Länge n ([HA05])

Beispiel: Ein Schieberegister mit 4 Stellen hat den Anfangszustand:

$$(a_0, a_1, a_2, a_3) = (1011) \quad (4.4)$$

Die Bildungsvorschrift lautet:

$$f(x_3, x_2, x_1, x_0) = x_2 + x_0 \quad (4.5)$$

t	S_3	S_2	S_1	S_0	Output
0	1	1	0	1	
1	0	1	1	0	1
2	1	0	1	1	0
3	1	1	0	1	1

Die einzelnen Zustände sind demnach:

Die Outputfolge ist somit 101101101... und mit 3 periodisch.

Die binären Zustände 1 und 0 werden auf die analogen Zustände +1 und -1 bezogen¹. Das binäre Pseudoräuschen benötigt wie ein Dirac-Impuls die gesamte spektrale Bandbreite. Allerdings ist theoretisch die Signalleistung des Pseudoräuschens um den Faktor L größer als bei einem periodischen Impuls mit der gleichen Periode. Dieser Leistungsgewinn führt zu einem besseren SNR . Der theoretische Dynamikgewinn gegenüber einer Messung mit einem Einzelimpuls beträgt $10 \log L \approx n \cdot 3[\text{dB}]$ ([GR02]). Indes verringert die D/A Umsetzung und das Übertragungsverhalten der Sender diesen Gewinn um mehrere dB². Pseudoräuschen wird beispielsweise in der akustischen Messtechnik verwendet³. Das Zeitzeichensignal *DCF77*⁴ von Funkuhren ist ebenfalls ein pseudozufälliges Signal.

Pseudozufälliges Rauschen ist nur eine Form um Signalinformationen zu kodieren. Weitere Methoden können der Fachliteratur entnommen werden (z.B. [HA05]).

4.3.1 Modulationsverfahren

Die Signalmodulation ist ein Vorgang, bei dem abhängig vom Verlauf des Nutzsignals ein oder mehrere Merkmale einer Trägerschwingung verändert werden. Empfangsseitig wird mittels des entsprechenden Demodulationsverfahrens aus der Trägerschwingung dieses Nutzsignal zurückgewonnen.

Diese aus der Nachrichtentechnik bekannten Methoden stellen eine einfach zu realisierende Möglichkeit dar, Signale verschiedener Sender an den Empfängern zu trennen. Dies ist vor allem mit Blick auf den Einsatz mehrerer parallelbetriebener Headtracker interessant. Da im Rahmen dieser Arbeit nicht ausführlich auf die verschiedenen Verfahren eingegangen werden kann, sei an dieser Stelle auf die Vielzahl von Fachliteratur (z.B. [MA02] oder [OH04]) verwiesen.

¹also auf die Vollaussteuerung des D/A Wandlers

²Der genaue Wert hängt stark von der Bandbreite der Sender ab.

³MLS-Verfahren zur Bestimmung raumakustischer Gütemaße.

⁴Die Abkürzung des westeuropäischen Zeitzeichensenders DCF77 steht für: Deutschland (*D*), Langwellensender (*C*), der Nähe zu Frankfurt (*F*) und der Trägerfrequenz (*77.5 kHz*).

Zeitmultiplex-Übertragung Die modulierten Signale nutzen ein vorhandenes begrenztes Frequenzband gemeinsam, indem sie sich in der Signalübertragung zeitlich abwechseln, bzw. nach definierten zeitlichen Mustern übertragen. Recht aufwendig sind in allen Zeitmultiplex-Systemen die Steuereinrichtungen, die im Empfänger die oftmals zeitlich sehr eng beieinanderliegenden Signale ohne Übersprechen separieren soll. zu diesen Verfahren gehören:

- Pulsamplitudenmodulation - PAM
- Pulsphasenmodulation - PPM
- Pulsdauermodulation - PDM

Frequenzmultiplex-Übertragung Bei dieser Modulationsart nutzen alle Signalquellen einen Übertragungskanal permanent aus, indem sie in (vom Empfänger unterscheidbaren) unterschiedlichen Frequenzbändern senden. Die einzelnen Kanäle können somit auch ohne Einhalten bestimmter Synchronisationsbedingungen am Empfänger getrennt werden (*asynchrone Multiplex-Übertragung*). Zu dieser Klasse gehören:

- Amplitudenmodulation - AM
- Frequenzmodulation - FM
- Phasenmodulation - PM

Codemultiplex-Übertragung Mit diesem Verfahren können Signale unterschieden werden, die mehrere Sender im gleichen Frequenzband zeitgleich ausstrahlen. Hierbei werden die Quellsignale mit individuellen Codesequenzen codiert. Zu diesem Zweck werden allen Sendern zueinander unabhängige Codemuster zugewiesen. Mit diesem Muster können die unterschiedlich codierten Signale auseinander gehalten und voneinander getrennt werden. Dieses Übertragungsverfahren wird z.B. für das *GPS*¹ verwendet.

¹Global Positioning System, satellitengestützten Positionsbestimmungssystem

Orthogonal Frequency Division Multiplex (OFDM) Hierbei handelt es sich um ein Verfahren, das mehrere Trägerfrequenzen für die Übertragung eines Digitalsignals benutzt, diese Trägerfrequenzen allerdings nur mit einer verringerten Übertragungsrate moduliert. Zu diesem Zweck wird bei *OFDM* das zur Verfügung stehende Frequenzband in mehrere Trägerbänder unterteilt. *OFDM* ist unempfindlicher gegen Störungen und Echos als andere Modulationsverfahren. Es kommt u.a. bei *WLAN*-Verbindungen¹ zum Einsatz.

¹Wireless Local Area Network

Kapitel 5

Detektionsmethoden

Vorrangigen Einfluss auf die Funktionsfähigkeit und die Systemeigenschaften des Headtrackers werden durch die Wahl der Laufzeitdetektierungsmethode bestimmt. Ziel ist es, die Laufzeitdifferenz der Sendestrecken aus Kap.3 möglichst schnell, genau und mit geringem Berechnungsaufwand zu ermitteln¹. Dabei sind die gemessenen Zeitdifferenzen über Gl.3.1 proportional zu den entsprechenden Wegdifferenzen. Teilweise bedingen die gezeigten Ansätze spezielle, in Kap.4 vorgestellte Signalformen.

5.1 Schwellwertdetektion

Bei dieser recht simplen und daher weit verbreiteten Methode² wird das Eintreffen eines (Nadel)Impulses über eine Schwellwertdetektion erkannt. Wird ein Piezowandler mit einem Impuls angeregt, ergibt dies aufgrund der Massenträgheit des Wandlers keinen idealen Impuls, sondern ein Signal mit Ein- und Ausschwingphase - also die Impulsantwort des Wandlers (siehe Abb.5.1).

Die Breite der Kurvenform bedingt im Schwellwertdetektor eine Entscheidungsunsicherheit. Je nach Entfernung des Senders zum Empfänger variiert die Amplitude des empfangenen Signals und damit auch die Zahl der über dem Schwellwert liegenden Zeitpunkte. Daher ist diese Verfahren sehr ungenau und arbeitet nur in einem sehr begrenzten Entfernungsbereich zufrieden stellend, da die Schallintensität des empfangenen Signals mit zunehmendem Senderabstand proportional sinkt. Außerhalb eines Bereiches übersteigt die Ungenauigkeit mehrere Millimeter.

Abgesehen von der Unschärfe der Detektion schon im ungestörten Fall ergeben sich durch die leichte Beeinflussbarkeit durch Fremdschall und der hohen Beanspruchung des Sendewandlers durch die Impulsübertragung³ noch weitere Probleme. Vorteil dieses Verfahrens

¹ siehe dazu Kap. 1.3 auf Seite 6

² siehe [CI84], [MO87], [BI80]

³ hohe Bandbreite des Impulses, siehe 4.2 auf Seite 31

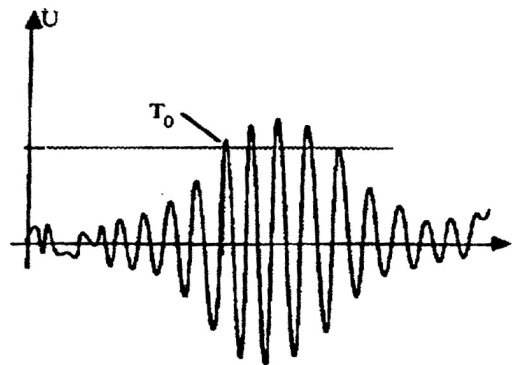


Abb. 5.1: Impulsantwort eines Piezowandlers [ZI92]

ist der geringe Konstruktionsaufwand und damit die niedrigen Produktionskosten.

Eine Verbesserung der Genauigkeit kann entweder durch eine selbstregulierende Verstärkerschaltung (AGC „automatic gain control“) oder durch adaptive Anpassung des Schwellwertes in Abhängigkeit von der Laufzeit erreicht werden. Dies bedarf allerdings eine exakte Kenntnis der *TOF*-Zeit¹ am Empfänger².

Eine Anpassung der Diskriminatorschwelle bewirkt einen relativ konstanten Auslösepunkt am Detektor ([KL81]).

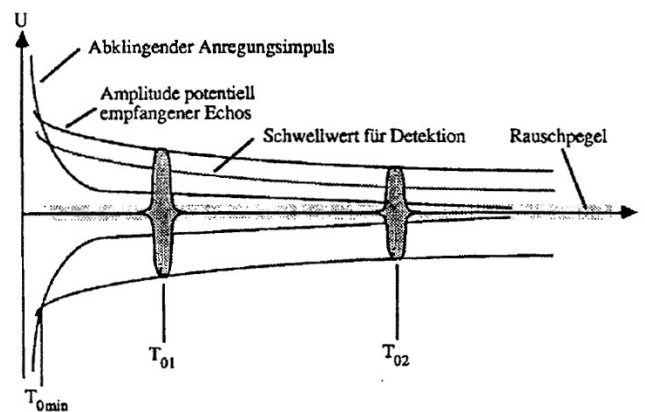


Abb. 5.2: Detektion mit angepasstem Schwellwert [ZI92]

¹Time of flight - Flugzeit

²wie z.B. es bei dem in Abschnitt 1.2.1 auf Seite 3 vorgestellten Modell durch Synchronisierung von Sender- und Empfängerseite möglich ist.

5.1.1 Hüllkurvendetektion

Um die Schwächen der einfachen Schwellwertdetektion über die Signalamplitude zu umgehen, wird ein eintreffendes GW-Paket über die Signaleinhüllende detektiert. Diese Einhüllende ist im Vergleich zum Trägersignal phasen- und frequenzstabiler ([PA91]).

Folgende Verfahren zur digitalen Hüllkurvengewinnung sind aus der Literatur ([PA91] und [ZO02]) bekannt:

quasi analoge Verarbeitung Üblicherweise wird die Hüllkurve mittels eines Gleichrichters und einer anschließenden Tiefpassfilterung extrahiert. Diese Technik erzeugt allerdings ein Signalverzögerung.

Peak Detection Anstatt der Tiefpassfilterung wird im gleichgerichteten Signal von einer Signalspitze zur nächsten interpoliert¹.

Hilbert Transformation Die Hüllkurve eines Signals ergibt sich aus dem Betrag der Hilberttransformierten. Dabei werden die negativen Frequenzen des Spektrums auf Null gesetzt und die Beträge der positiven Frequenzen verdoppelt. Nach einer Transformation in den Zeitbereich erhält man aus dem Absolutbetrag die Hüllkurve (siehe auch [GM81]).

¹zum Interpolieren siehe auch Kap.5.3.1.2.

5.1.2 Verwandte Methoden

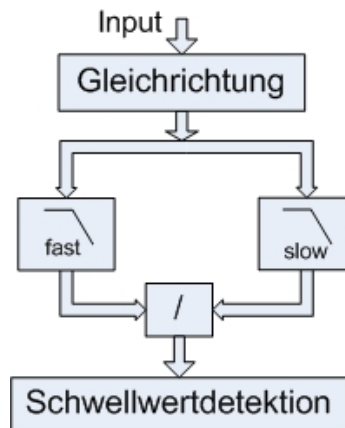


Abb. 5.3: Strukturdiagramm der adaptiven Hüllkurvendetektion

Eine Kombination aus *AGC* und Hüllkurvenerkennung ist in dieser Diplomarbeit entwickelt worden. Hierbei wird mittels eines Tiefpasses mit sehr langer Zeitkonstante die mittlere Schall- und Rauschenergie bestimmt. Ein weiterer Tiefpass, dessen Zeitkonstante auf die Feuerrate der Bursts abgestimmt ist, liefert die Einhüllende (siehe Kap.5.1.1). Bildet man nun das Verhältnis zwischen diesen beiden Filtern, so erhält man ein amplitudenstabileres Hüllkurvensignal, welches eine solidere Schwellwertdetektion ermöglicht.

5.1.2.1 Kombination aus Thresholding und Zero-Crossing Detection

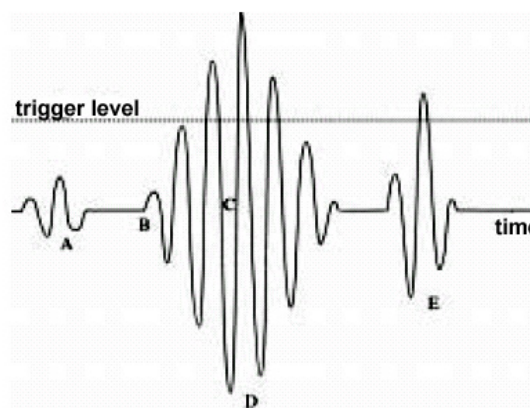


Abb. 5.4: Kombiniertes Thresholding und Zero-Crossing ([WZ97])

Diese nach ([WE97]) sehr robuste Methode detektiert den Nulldurchgang eines empfangenen *GW*-Signals, nachdem dieses Signal einen gewissen Schwellwert überschritten hat.

Da die Methode unempfindlich gegenüber Reflexionen und Amplitudenverzerrungen sein soll, wird dem Autor nach eine Genauigkeit von 0,06% relativ zur Laufzeit erreicht.

Abbildung 5.4 zeigt ein empfangenes Ultraschallsignal, wobei es sich bei D um den zu detektierenden Burst handelt. Das Schwingungspaket A und E stellen reflektierte Bursts dar, die durch stärkere Dämpfung aufgrund ihres Ausbreitungsweges eine kleinere Amplitude aufweisen. Überschreitet das Eingangssignal den Schwellwert, wird der darauf folgende Nulldurchgang detektiert (siehe Position C). Für große Amplitudenschwankungen ergibt sich so ein Fehler vom Vielfachen der Wellenlänge. Eine quasi konstante Amplitude kann wieder durch eine AGC erreicht werden. Nachdem die Schwellwertüberschreitung die Nulldurchgangdetektion ausgelöst hat, ist der Schwellwert für eine der Burstrate angepasste Zeit unsensibel auf Überschreitungen. Dies bewirkt, dass nachfolgende Echos, die wie E in Abb.5.4 über dem Schwellwert liegen, nicht zu Fehldetektionen führen.

5.1.2.2 Zero-Crossing Detection der Signalhüllkurve

Die Autoren von [FI92] und [FI01] stellen ein Detektionssystem vor, welches ein eintreffendes GW-Signal im Nulldurchgang der ersten Ableitung der Signalhüllkurve detektiert (siehe Abb.5.5). Dieser Ansatz ist sowohl digital als auch in einer analogen Schaltung realisierbar.

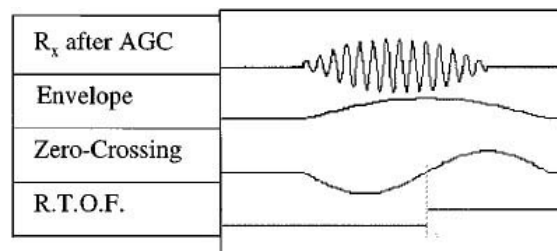


Abb. 5.5: Zero-Crossing Detection der Signalhüllkurve ([FI01])

5.2 Phasenmessung

Die Phasenmessmethode arbeitet mit dem konstanten Schallfeld eines CW-Signals¹. Grundlage ist bei konstanter Frequenz die gleichbleibende örtliche Lage der Schalldruckverteilung einer im ungestörten Medium fortschreitenden Schallwelle. Hierbei wird das ausgesendete Signal an zwei oder mehreren Sensoren erfasst und über die Phasendifferenz des Sendesignals am Ort der Sensoren auf eine Laufzeitdifferenz geschlossen. Überschreitet die Laufzeitdifferenz den Wert der Wellenlänge im Medium, dann wiederholen sich

¹CW - continuous wave, siehe Kap. 4.2 auf Seite 31

Phasenwerte periodisch mit jeder weiteren Wellenlänge (räumliches Aliasing). Daher beschränkt sich die maximal messbare Laufzeitdifferenz auf die halbe Wellenlänge des Sendesignals.

Befinden sich nach Kap. 3.3.5 zwei Empfängersensoren in einem Abstand von $a = 16\text{cm}$ am Kopfhörer, so ergibt sich eine maximale Frequenz des Sendesignals, bei welcher die maximale Laufzeitdifferenz¹ a detektierbar ist:

$$\frac{c}{2 \cdot a} = f \quad (5.1)$$

$$\frac{345\text{m/s}}{2 \cdot 0,16\text{m}} \approx 1000\text{Hz} \quad (5.2)$$

Wie man leicht sieht, kann das Verfahren aufgrund der benötigten Signalfrequenz und den Nachteilen eines CW-Signals nicht überzeugen. Mit handelsüblichen 40kHz Ultraschallsensoren ist eine maximale Laufzeitdifferenz von:

$$\frac{c}{2 \cdot f} = a \quad (5.3)$$

$$\frac{345\frac{\text{m}}{\text{s}}}{2 \cdot 40000\text{Hz}} \approx 3\text{mm} \quad (5.4)$$

möglich.

Prinzipiell bietet die Phasenvergleichsmethode bei geringen Interferenzen im Schallfeld jedoch eine sehr gute Auflösung kleiner Abstände, da die Phasendifferenz zweier elektrischer Signale heutzutage sehr genau bestimmbar ist ([EH85]). In der Literatur lassen sich Methoden zur Erweiterung des Messbereiches finden. Dies erfolgt meistens durch Nachführung der Signalfrequenz entsprechend dem detektierten Abstand ([EH85] sowie [KI84] und [GM81]). Dies bedarf allerdings eine hohe Übertragungsbandbreite und eine direkte Verbindung von Sender und Empfängereinheit, damit die Sendefrequenz auf den aktuellen Messwert justiert werden kann.

5.3 Korrelationsbestimmung

Eine genaue Methode zur Bestimmung der Zeitverzögerung zwischen 2 Signalen ist die Kreuzkorrelation. Das Korrelationsergebnis liefert eine Aussage zur Ähnlichkeit der verglichenen Signale. Das Maß der Übereinstimmung zeigt sich in der Größe des Korrelationsergebnisses, welches von der Ähnlichkeit der Signalform abhängt. Dieses Verfahren kann sowohl im Zeitbereich oder (durch eine Fouriertransformation) auch im Frequenzbereich durchgeführt werden. ([FO83], [ME86])

¹bei einem Drehwinkel von 90°

Erreicht das gesendete Signal $s(t)$ um den Zeitbereich D zeitversetzt 2 Empfänger, so können die empfangenen Signale wie folgt mathematisch beschrieben werden:

$$x_1(t) = s(t) + n_1(t) \quad (5.5)$$

$$x_2(t) = \alpha s(t + D) + n_2(t) \quad (5.6)$$

s :	Sende- bzw. Nutzsignal
a :	Dämpfungsfaktor ($0 \leq a \leq 1$)
n_1, n_2 :	Rauschen
x_1, x_2 :	empfangene Signale am Empfänger 1 und Empfänger 2
D :	Laufzeitdifferenz

Das Signal s sowie die Rauschanteile n_1 und n_2 werden als paarweise unkorrelierte Gauß'sche Zufallsprozesse angenommen.

5.3.1 Korrelationsauswertung im Zeitbereich

Um den Zeitversatz D zu bestimmen, vergleicht man x_1 und x_2 auf Gemeinsamkeiten (Gl.5.7). Dabei resultiert die geschätzte Verzögerung D aus dem Wert τ , an dem $R_{x_1, x_2}(\tau)$ ein Maximum erreicht (Gl.5.8).

$$R_{x_1, x_2}(\tau) = \int_{-\infty}^{\infty} x_1(t) x_2(t + \tau) dt \quad (5.7)$$

$$\hat{D} = \arg \max R_{x_1, x_2}(\tau) \quad (5.8)$$

In Gl.5.7 wird die Kreuzkorrelation über einen unendlich langen Zeitraum betrachtet. Unter dieser idealisierten Voraussetzung besitzt die Kreuzkorrelationsfunktion ihr Maximum bei $\tau = D$. Da die Signale nur über einen begrenzten Zeitraum betrachtet werden können, kann $R_{x_1, x_2}(\tau)$ nur mit einem Fehler geschätzt werden (Gl.5.9). Die Länge des zeitlichen Fensters bewirkt, dass Frequenzanteile des zu korrelierenden Signals, deren Periodendauer größer oder im Bereich der Länge des zeitlichen Fensters liegen schlecht bis gar nicht verglichen werden können. Weiterhin kann D auch nicht größer als dieser Fensterlänge sein. In der Praxis werden daher Fenster benutzt, welche die 10fache Länge der maximal zu erwarteten Zeitdifferenz besitzen.

$$R_{x_1, x_2}(\tau) = E[x_1(t) x_2(t - \tau)] \quad (5.9)$$

Gleichung 5.10 beschreibt die zeitdiskrete Kreuzkorrelation, wie sie in digitalen Systemen Verwendung findet. Für ergodische Prozesse¹ kann die Kreuzkorrelation im Zeitraum $N \cdot T_s$ geschätzt werden ([CA76]).

$$\hat{R}_{x_1, x_2}[m] = \frac{1}{N} \sum_{k=1}^N x_1[k] \cdot x_2[k + m] \quad (5.10)$$

¹bei ergodischen Signalen sind Zeit- und Ensemblemittelwerte identisch

Abb.5.6 zeigt schematisch die direkte Methode zur Gewinnung der Kreuzkorrelationsfunktion. Durch eine auf das Nutzsignal abgestimmte Vorfilterung von x_1 und x_2 kann der Fehler $|\hat{D} - D|$ minimiert werden. Von großem Vorteil ist daher ein *a priori* Wissen über die zu korrelierenden Signale und den zu erwartenden Störgeräuschen, um die Filter H_1 und H_2 optimal zu designen. Die Kreuzkorrelationsfunktion kann indirekt auch über die Berech-

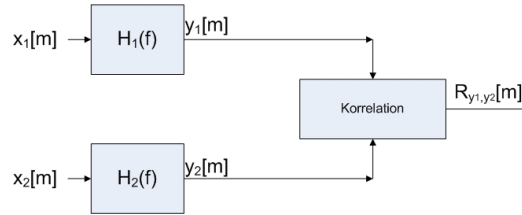


Abb. 5.6: Direkte Korrelationsberechnung im Zeitbereich

nung der spektralen Kreuzleistungsdichte im Frequenzbereich ermittelt werden. Hierbei ergibt eine Multiplikation der fouriertransformierten Eingangssignale¹ das Kreuzleistungsspektrum². Dabei ist zu beachten, dass bei dieser komplexen Multiplikation ein Faktor konjugiert komplex sein muss³. Abb.5.7 zeigt diesen Ansatz.

$$G_{x_1, x_2}(\omega) = X_1^*(\omega) \cdot X_2(\omega) \quad (5.11)$$

Eine Vorfilterung von x_1 und x_2 nach Abb.5.6 ergibt das Kreuzleistungsspektrum der ge-

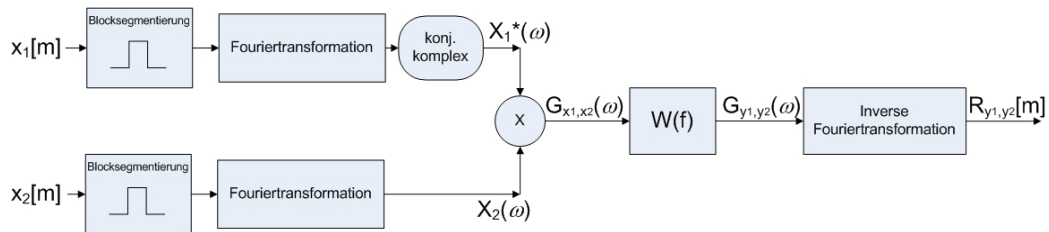


Abb. 5.7: Berechnung der Korrelation über die Kreuzleistungsdichte

filterten Signale y_1 und y_2 :

$$G_{y_1, y_2}(\omega) = H_1(\omega) H_2^*(\omega) G_{x_1, x_2}(\omega) \quad (5.12)$$

¹Die Fouriertransformierte des Eingangssignals ist wie folgt definiert: $X(\omega) = \sum_{-\infty}^{\infty} x(m) \cdot e^{-j\omega m}$ ([OP75])

²Eigentlich muss es „Kreuzleistungsdichtespektrum“ heißen.

³Die Korrelation entspricht einer Faltungsoperation, wobei jedoch die zeitliche Umkehr eines Signals nicht erfolgt. Diese zeitliche Nicht-Umkehr wird durch konjugieren im Frequenzbereich berücksichtigt.

Das Kreuzleistungsspektrum steht mit der Kreuzkorrelationsfunktion über die inverse Fouriertransformation in Beziehung:

$$R_{y_1, y_2}[m] = \frac{1}{2\pi} \int_{-\pi}^{\pi} G_{y_1, y_2}(\omega) e^{j\omega m} d\omega \quad (5.13)$$

$$\omega = \frac{2\pi \cdot f}{f_s} \quad (5.14)$$

Wobei $f \leq \frac{f_s}{2}$ gelten muss. Die Vorfilterung durch H_1 und H_2 kann ebensogut im Frequenzbereich realisiert werden. Dafür wird die Gewichtungsfunktion $W(\omega)$ eingeführt.

$$R_{y_1, y_2}[m] = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) G_{x_1, x_2}(\omega) e^{j\omega m} d\omega \quad (5.15)$$

mit:

$$W(\omega) = H_1(\omega) H_2^*(\omega) \quad (5.16)$$

Ob die Korrelation nach Abb.5.6 oder nach Abb.5.7 berechnet werden sollte, ist hauptsächlich von der Länge des Signalausschnittes N (siehe Gl.5.10) und den verfügbaren Rechenkapazitäten abhängig: Unter Berücksichtigung aller durchzuführenden Rechenschritte kann man sagen, dass ab einer Länge von $N = 32$ eine Berechnung nach Abb.5.7 weniger Rechenoperationen als die direkte Korrelationsberechnung nach Abb.5.6 benötigt.

5.3.1.1 Gewichtungsfunktionen

Die Tabelle 5.1 fasst einige aus der Literatur bekannte Gewichtungsfunktionen zusammen¹. Sie werden in der Literatur auch als „*prewhitening filters*“ bezeichnet. Aufgabe dieser Gewichte ist es, die Korrelationsfunktion zu entrauschen und somit eine bessere Detektion von D zu ermöglichen (siehe Abb.5.8). Es gibt keine Gewichtung, die prinzipiell am besten filtert. Die Gewichte sind auf unterschiedliche Anwendungssituationen, Signal-Störabstände und Bandbreiten der zu korrelierenden Signale ausgelegt. Vielmehr ist daher die Effizienz der Gewichte von der richtigen Filterwahl im Bezug zum Nutzsignal und den Störgrößen abhängig². Durch einige Filter, sowie aufgrund der Hin- und Rücktransformationen kommt es zu einer Verschmierung bzw. Aufweitung der Kreuzkorrelationsfunktion. Je geringer in Gl.5.10 die Länge N der Signalausschnitte ist (und damit auch die Spektrale Information) desto größer wird dieser Effekt. Im Extremfall kann in Gl.5.8 die Position des Maximums nicht mehr exakt bestimmt werden. Die Wahl der Gewichtungsfunktion ist also ein Kompromiss zwischen Genauigkeit und Stabilität.

¹zu finden auch in [CA76]

²siehe dazu Gleichung 5.38 auf Seite 52

Klasse	Gewichtungsfunktion $W(\omega)$	Bemerkung
KORR	1	direkte Korrelation ohne Gewichtung
ROTH (Roth Impulse Response)	$\frac{1}{G_{x_1, x_1}(\omega)}$	beschreibt Wiener-Hopf Filter
SCOT (Smooth COherence Transform)	$\frac{1}{\sqrt{G_{x_1, x_1}(\omega) \cdot G_{x_2, x_2}(\omega)}}$	unterdrückt tonale Anteile des Kreuzspektrums
PHAT (PHase Transform)	$\frac{1}{ G_{x_1, x_2}(\omega) }$	behält nur Phaseninformation des Kreuzspektrums
ML (Maximum-Likelihood)	$\frac{ \gamma_{x_1, x_2}(\omega) ^2}{ G_{x_1, x_2}(\omega) (1 - \gamma_{x_1, x_2}(\omega) ^2)}$	Statistische Schätzfunktion für optimale Gewichtung je nach Energiegehalt

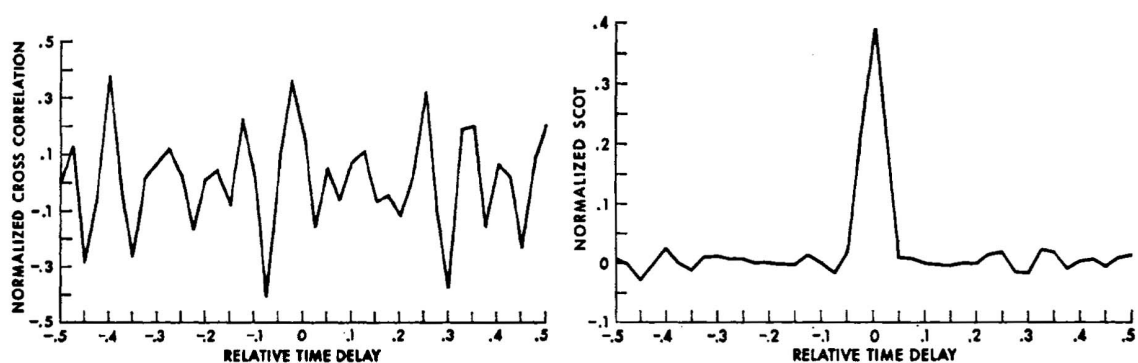
Tab. 5.1: verschiedene Gewichtungsfunktionen $W(\omega)$ 

Abb. 5.8: Kreuzkorrelationsfunktion ohne und mit Vorfilterung (hier SCOT) [CA76]

ROTH: Die Gewichtung der ROTH-Fensters ergibt als Korrelationsfunktion die Schätzung der Impulsantwort des *optimum linear (Wiener-Hopf) filters* (siehe 5.17). Die ROTH-Gewichtung schätzt Spektralbereiche mit der größten Übereinstimmungen zwischen x_1 und

x_2 und hebt diese heraus, indem verrauschte Frequenzbereiche gedämpft werden. Diese Bandbreitenbegrenzung hat allerdings zur Folge, dass die Kreuzkorrelationsfunktion aufgeweitet wird ([CA76]). Ist das Autoleistungsspektrum des Rauschens $G_{n_1, n_1}(\omega)$ gleich einem Vielfachen (oder 0) von $G_{s_1, s_1}(\omega)$, bewirkt diese Gewichtung eine einheitliche Dämpfung über allen Frequenzbins und somit keine Aufweitung der Korrelationsfunktion. Wie in Gl. 5.17 zu sehen geht in die ROTH-Gewichtung nur das Autoleistungsspektrum von x_1 ein. Damit wird hier vorausgesetzt, dass das Autoleistungsspektrum beider empfangenen Signale ident ist. Ist dies nicht der Fall, wird diese Gewichtung nicht optimal wirken können.

$$G_{y_1, y_2}(\omega) = \frac{G_{x_1, x_2}(\omega)}{G_{x_1, x_1}(\omega)} \quad (5.17)$$

SCOT: Die SCOT-Gewichtung baut auf die ROTH-Gewichtung auf. Die Idee von SCOT ist es ebenso, nichtkorrelierende Spektralanteile schwächer zu gewichten als korrelierende Frequenzen. Hierbei wird jedoch angenommen, dass sowohl das Autoleistungsspektrum von x_1 also auch von x_2 die Schätzung $\hat{G}_{x_1, x_2}$ beeinflussen.

Die SCOT-Gewichtung geht von der Kohärenzfunktion zweier stationärer Zufallsprozesse $x_1(t)$ und $x_2(t)$ aus (Gl. 5.18). Die Kohärenzfunktion beschreibt die normierte Kreuzkorrelation als Funktion der Frequenz im Wertebereich von:

$$\gamma_{x_1, x_2}(\omega) = \frac{G_{x_1, x_2}(\omega)}{\sqrt{G_{x_1, x_1}(\omega) G_{x_2, x_2}(\omega)}} \quad (5.18)$$

$$0 \leq |\gamma_{x_1, x_2}(\omega)| \leq 1 \quad (5.19)$$

Eine hohe Kohärenz wird als hoher SNR interpretiert. Das Fenster passt sich somit den zugeführten Signalen an. Ein Vorteil von SCOT ist es, gezielt schmalbandige, tonale Störungen der Eingangssignale zu dämpfen ([CA73]). Jedoch führt auch diese Bandbreitenreduktion zur Aufweitung der Kreuzkorrelationsfunktion.

PHAT: Die PHAT-Gewichtung ist bei Sprachsignalen laut [Z004] sehr robust in Bezug auf störende Einflüsse durch Raumreflexionen. Hierbei wird jeder einzelne Frequenzbin von G_{x_1, x_2} mit dem jeweiligen inversen Betrag gewichtet (siehe Gl. 5.20). Zur Rücktransformation in den Zeitbereich werden also nur die Phasenanteile des Kreuzspektrums herangezogen. Dies hat den Vorteil, dass die KKF nicht verschmiert wird, wenn das überlagerte Rauschen aus Gl. 5.5 und 5.6 unkorreliert ist ($G_{n_1, n_2} \stackrel{!}{=} 0$). Bei einem schmalbandigen Nutzsignal $s_1(t)$ enthalten viele Frequenzbins keine nutzbaren Phaseninformation (sondern nur Rauschen). Da die Gewichtung ohne Beachtung der Signalenergie vorgenommen wird, gehen somit alle Frequenzen in die Kreuzkorrelation mit ein. Diese verrauschten Frequenzbins beeinträchtigen die Genauigkeit der KKF ([CA76]). Soll dieser Effekt vermieden werden, muss das Frequenzspektrum abhängig der Signalenergie noch zusätzlich gewichtet werden.

$$\frac{\hat{G}_{x_1, x_2}(\omega)}{|G_{x_1, x_2}(\omega)|} = e^{j\hat{\theta}(\omega)} \quad (5.20)$$

ML: Der Maximum-Likelihood-Schätzer (in der Literatur auch *Hannan-Thomson processor* genannt) gewichtet das Frequenzspektrum abhängig der Kohärenzfunktion in Gl.5.18. Unter der Voraussetzung, dass $|\gamma_{x_1, x_2}(\omega)|^2 \neq 1$ sei, gilt Gl.5.21.

$$G_{y_1, y_2}^{ML}(\omega) = \frac{\hat{G}_{x_1, x_2}(\omega)}{|G_{x_1, x_2}(\omega)|} \cdot \frac{|\gamma_{x_1, x_2}(\omega)|^2}{[1 - |\gamma_{x_1, x_2}(\omega)|^2]} \quad (5.21)$$

Wie aus Gl.5.21 zu sehen ist, bestehen Parallelen zwischen der *ML*-Gewichtung und weiter oben gezeigten Schätzern. So berechnet *ML* genau wie *PHAT* die Phase des Kreuzleistungsspektrums wie in Gl.5.20. Zum Anderen erfolgt wie bei *SCOT* eine Gewichtung aufgrund der Kohärenzfunktion γ_{x_1, x_2} . Der *ML*-Schätzer ist somit ein mit der Kohärenz der Phase invers gewichteter *PHAT*-Schätzer ([CA76]). Ausgedrückt durch Signal- und Rauschleistungsdichte ergibt sich ([CA73]):

$$W_{ML}(\omega) = \frac{\frac{G_{s_1, s_1}(\omega)}{G_{n_1, n_1}(\omega) \cdot G_{n_2, n_2}(\omega)}}{1 + \frac{G_{s_1, s_1}(\omega)}{G_{n_2, n_2}(\omega)} + \frac{G_{s_1, s_1}(\omega)}{G_{n_1, n_1}(\omega)}} \quad (5.22)$$

Ist $|G_{x_1, x_2}(f)|$ und $|\gamma_{x_1, x_2}(f)|^2$ bekannt, erhält man die bestmögliche Lösung, da dadurch die geringstmögliche Varianz des Schätzers¹ erreicht wird. In der Praxis sind die jeweiligen Spektren jedoch meist nicht exakt bekannt, sodass sie geschätzt werden müssen. (z.B. mittels Segmentierung und Mittelung ([OP75] oder [CA73])) Die Varianz des angenäherten LS-Schätzers ist dann höher als die des exakten *ML*-Schätzers:

$$\text{Var}(\hat{D}) \geq \left[N \int_{-\infty}^{\infty} (2\pi f)^2 \frac{|\gamma_{x_1, x_2}(\omega)|^2}{[1 - |\gamma_{x_1, x_2}(\omega)|^2]} df \right]^{-1} \quad (5.23)$$

wird vorausgesetzt, dass die Rauschanteile n_1 und n_2 zueinander unkorreliert sind, kann im Allgemeinen die Varianz zur Schätzung des gesuchten Zeitversatzes D in Abhängigkeit der Fensterfunktion $W(\omega)$ angegeben werden ([CA76]):

$$\text{var}[\hat{D}] = \frac{\int_{-\infty}^{\infty} |W(\omega)|^2 (2\pi f)^2 G_{x_1, x_1}(\omega) G_{x_2, x_2}(\omega) [1 - |\gamma_{x_1, x_2}(\omega)|^2] df}{N \int_{-\infty}^{\infty} (2\pi f)^2 [|G_{x_1, x_2}(\omega)| W(\omega)]^2 df} \quad (5.24)$$

Eine genaue Evaluierung verschiedener Gewichtsfunktionen findet man z.B. in [BJ03] oder in [SC81].

¹engl. *Cramer-Rao lower bound*, CRLB siehe [WU97]

5.3.1.2 Digitale Auflösungsgrenzen

Wird die Korrelationsbestimmung auf digitaler Ebene durchgeführt, so werden die zu korrelierenden Signale x_1 und x_2 bei der Analog/Digitalumsetzung mit f_s zeitdiskret abgetastet. Eine Korrelation kann demzufolge die Laufzeitdifferenz D nur mit einer maximalen Genauigkeit von T_s bestimmen.

$$m_{\max} = \arg \max (\hat{R}(m)) \quad (5.25)$$

$$\hat{D} = \frac{m_{\max}}{f_s} \quad (5.26)$$

Benötigt man ein genaueres Ergebnis, so muss entweder die Abtastrate erhöht werden

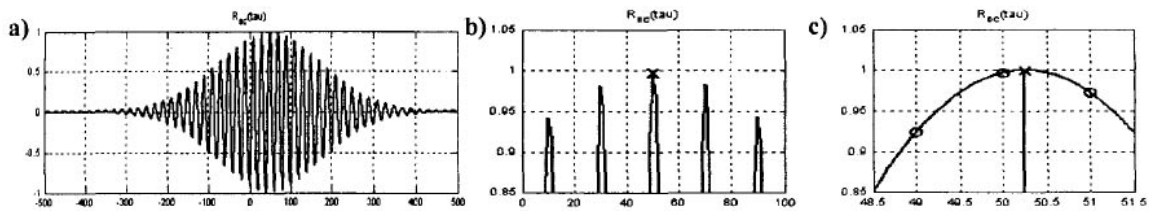


Abb. 5.9: Parabolische Interpolation: a) gesamte Korrelationsfunktion b) Signalausschnitt um das Maximum c) Interpolation mit neuem Maximalwert ([BL02])

(wirklich oder künstlich¹), was natürlich den Rechenlast beträchtlich erhöht oder man interpoliert die berechnete Kreuzkorrelation $\hat{R}$. Dies kann unter hohem Rechenaufwand mit einem geeigneten FIR-Filter² geschehen oder man verwendet Interpolationsfunktionen mit wenigen Stützstellen. Das Einsparen an Rechenoperationen durch eine approximierte Interpolation wiegt in den meisten Fällen den geringen Genauigkeitsverlust auf (z.B. Lagrange, parabolische Interpolation ([BL02])). Die Autoren von [MA92] empfehlen eine parabolische Interpolation mit drei Stützstellen. Dabei wird durch das ganzsamplige, groben Maximum $m_{\max}$ und den angrenzenden Punkten $m_{\max+1}$ und $m_{\max-1}$ eine Parabel gelegt. Das Maximum dieser Parabel liegt zwischen dem groben Maximum und seinem Nachbarsample (siehe Abb.5.9).

$$m_{\max} = \arg \max \hat{R}_{KKF}(m) \quad (5.27)$$

$$\hat{\delta} = -\frac{\hat{R}_{(m_{\max+1})} - \hat{R}_{(m_{\max-1})}}{2[\hat{R}_{(m_{\max+1})} - 2\hat{R}_{(m_{\max})} + \hat{R}_{(m_{\max-1})}]} \quad (5.28)$$

$$\hat{D}_{\text{frac}} = \frac{m_{\max} + \hat{\delta}}{f_s} \quad (5.29)$$

¹Damit ist gemeint, dass entweder ein A/D-Wandler mit einer höheren Abtastrate verwendet wird oder nach der Wandlung durch ein oversampling-processing die Signale interpoliert werden.

²Zwischen den Datenpunkten werden Nullen eingefügt, ein idealer Tiefpass bewirkt dann eine verschmierung der Datenpunkte über diese Nullpunkte.

Diese Interpolationsmethode enthält jedoch einen systematischen Fehler (Bias), der vom Signalspektrum und der Gewichtungsfunktion abhängt ([WU97]). Er hat die größte Unsicherheit bei $\delta = \pm 0.25$. Dieser Fehler ist aber so marginal, dass er als vernachlässigbar angesehen werden kann ([MA92]).

5.3.2 Korrelationsauswertung im Frequenzbereich

Im vorherigen Kapitel wurde die Gewinnung der Kreuzkorrelationsfunktion durch die Berechnung des Kreuzleistungsdichtespektrums und einer anschließenden Rücktransformation in den Zeitbereich vorgestellt. Eine Verzögerung zwischen 2 Signalen ist jedoch ebenso direkt im Frequenzbereich (ohne Rücktransformation) ermittelt werden. Wieder ausgehend von Gleichung 5.5 und 5.6 auf Seite 43 und ihrer Definition, folgt für das Kreuzspektrum:

$$G_{x_1, x_2}(\omega) = G_{s, s}(\omega) e^{j\omega D} \quad (5.30)$$

Die Kreuzleistungsdichte besitzt (im Gegensatz zur Autoleistungsdichte) im Frequenzbereich einen linearen Phasenzug, dessen Steigung proportional zur Laufzeitdifferenz D ist. Ist also die Gruppenlaufzeit von G_{x_1, x_2} bekannt, kann man daraus D schätzen (Gl. 5.32).

$$\omega \hat{D} = \arg[G_{x_1, x_2}(\omega)] = \theta(\omega) \quad (5.31)$$

$$-\hat{D} = \frac{d}{d\omega} \arg[G_{x_1, x_2}(\omega)] \quad (5.32)$$

Da für die Ermittlung von $\hat{D}$ also nur die Phaseninformation des Kreuzleistungsspektrums

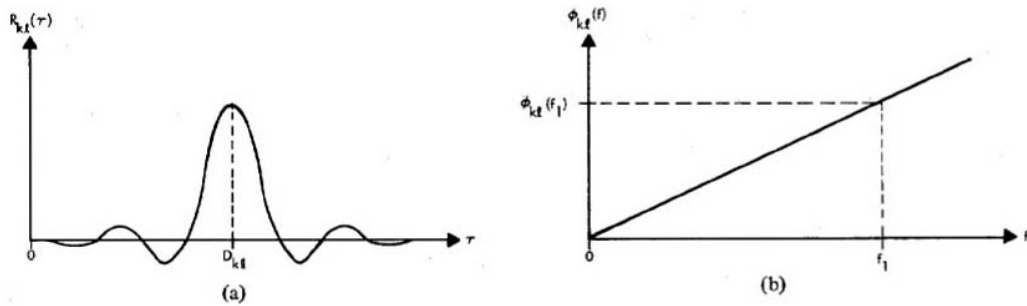


Abb. 5.10: (a) Kreuzkorrelationsfunktion und (b) entsprechende Phase des Kreuzleistungsspektrums ([PI81])

benötigt wird, entspricht diese Methode einer Kreuzkorrelation mit *PHAT*-Gewichtung¹. Allerdings erspart man sich hier die inverse Fouriertransformation. Ein weiterer Vorteil dieser Methode ist, dass subsample-Präzision ohne Interpolation erreicht wird. Für die Berechnung der Gl. 5.32 muss jedoch die zwischen $-\pi$ und π gefaltete Phase in eine kontinuierliche, lineare Funktion gebracht werden. Dies erfolgt durch sog. *phase unwrapping*.

¹ siehe Kap. 5.3.1.1 auf Seite 45

Hierbei werden die Unstetigkeiten (Sprünge von π zu $-\pi$ oder umgekehrt) der Phase durch Addition von $2n\pi$ ausgleichen. Ist das Sendesignal durch einen Rauschbelag gestört, kann das Rauschen zu einem ungewollten Phasensprung führen, da zwischen wirklichen Sprungstellen und zufälligen Schwankungen nicht mehr unterschieden werden kann. Daher ist dieser Methode nur bei geringen oder kleinen Störungen zuverlässig.

5.3.2.1 Verbesserungen

- Falsche Phasensprünge während des *phase unwrappings* aufgrund einer verrauschten Phase können durch Mittelung der Phase über mehrere FFT-Segmente verhindert werden. Die Anzahl der dabei zur Mittelung genutzten Segmente ist proportional zur Latenz des Algorithmus.
- Ist das Nutzsignal s_1 schmalbandig, enthält das Kreuzleistungsspektrum Frequenzstellen ohne relevanter Phaseninformation. In Gleichung 5.30 wird aber die Phaseninformation des gesamten Spektralbereiches herangezogen, was im Allgemeinen zu großen Detektionsfehlern führt ([BL02]). Im Falle von schmalbandigen Nutzsignalen darf zur Berechnung der Gruppenlaufzeit in Gl.5.32 nur die Phaseninformationen des von s_1 genutzten Frequenzbereichs betrachtet werden. Für schmalbandige Signale hat diese Detektionsmethode daher nur mit einem Vorwissen zur spektralen Lage von s_1 Erfolg.

Ist dieses Vorwissen nicht gegeben, können statistische Gewichtungungsverfahren eine Verbesserung schaffen:

LS Schätzer Vergleicht man die *Least Square* Schätzung in Gl.5.34 mit dem Korrelationsgewicht des *ML*-Schätzers¹ ist festzustellen, dass beide die Phase dort stärker gewichten, wo die Kohärenz zwischen x_1 und x_2 groß ist (und damit die Varianz der Phase gering [CH78]). Ausgehend von der Kostenfunktion in Gl.5.33 ergibt eine Minimierung von J die Schätzung der Laufzeitdifferenz $\hat{D}$ in Gl.5.34.

$$J = \sum_{k=1}^{N-1} \psi_k (\omega_k - D\theta_k)^2 \quad (5.33)$$

$$\hat{D} = \frac{\sum_{k=1}^{N-1} \psi_k \theta_k \omega_k}{\sum_{k=1}^{N-1} \psi_k \omega_k^2} \quad (5.34)$$

¹ siehe Kap. 5.3.1.1 auf Seite 45

mit:

$$\theta_k = \tan^{-1} \frac{\text{Imag} \left\{ \sum_{l=1}^M X_{1l}(k) X_{2l}^*(k) \right\}}{\text{Real} \left\{ \sum_{l=1}^M X_{1l}(k) X_{2l}^*(k) \right\}} \quad (5.35)$$

$$\omega_k = \frac{\pi \cdot k}{N-1} \quad (5.36)$$

X_{il} : Fouriertransformierter Signalblock des Signals x_i mit der Signallänge N
 $x = [x_1 \dots x_{N-1}]$

wobei die Varianz ψ_k mit Hilfe der Kohärenzfunktion γ_k gebildet wird. Die Exaktheit der Schätzung steigt mit Erhöhung der Segmentanzahl M der verwendeten Kreuzspektren. Unter der Voraussetzung, dass der Fehler (ε_k), der durch endliche Beobachtungszeit und Rauschen entsteht normalverteilt ist, gilt:

$$|\hat{\gamma}_k|^2 = \frac{\left| \sum_{l=1}^M X_{1l}(k) X_{2l}^*(k) \right|^2}{\sum_{l=1}^M |X_{1l}(k)|^2 \cdot \sum_{l=1}^M |X_{2l}(k)|^2} \quad (5.37)$$

$$\psi_k = \frac{|\gamma_k|^2}{1 - |\gamma_k|^2} \quad (5.38)$$

$$E(\varepsilon_k) = 0 \quad (5.39)$$

Eine andere Schätzung der Varianz zeigt Gl.5.40. Hierbei wird der momentane SNR über ein gemessenes Grundrauschen n_1 und n_2 bestimmt. Diese Schätzung ist also nur dann möglich, wenn das Rauschen ohne Nutzsignal gemessen werden kann ([WU97]).

$$\psi_k = \frac{\lambda_{1k}}{|X_1(k)|^2} + \frac{\lambda_{2k}}{|X_2(k)|^2} \quad (5.40)$$

$$\lambda_{ik} = \frac{1}{J} \sum_{l=1}^J |n_{il}(k)|^2 \quad i = 1, 2$$

Complex Mutal Power Operator Die Probleme, die sich bei Verwendung des *phase unwrappings* ergeben, werden hier umgangen, indem Gruppenlaufzeiten ([BL02]: „*mutual group delay*“) zwischen den Phasensprüngen bestimmt werden (Gl.5.42). Die Gewichtung der Leistungsdichte mit den ermittelten Gruppenlaufzeiten führt in Gl.5.43 zur Laufzeitdifferenz $\hat{D}$ ([BL02]).

$$\tau(k) = \angle G_{x_1, x_2}(k) - \angle G_{x_1, x_2}(k-1) \quad (5.41)$$

$$\tau(k) = \angle [G_{x_1, x_2}(k) \cdot G_{x_1, x_2}^*(k-1)] \quad (5.42)$$

$$\hat{D}_{CMPO} = \frac{\sum_{k=0}^{N/2-1} |G_{x_1, x_2}(k)| \tau(k)}{\sum_{k=0}^{N/2-1} |G_{x_1, x_2}|} \quad (5.43)$$

5.3.3 1-Bit-Korrelation

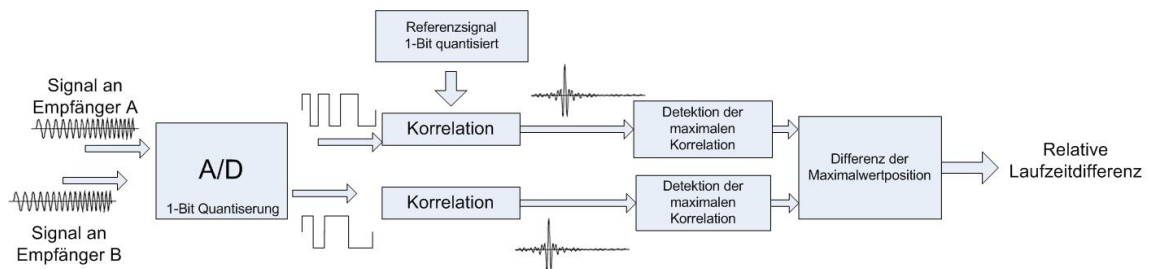


Abb. 5.11: 1-Bit-Korrelation eines Signals an zwei Empfängern

Die 1-Bit-Korrelation ist eine spezielle Art der Korrelation im Zeitbereich. Hierbei werden die empfangenen Signale x_1 und x_2 nicht miteinander, sondern jeweils mit einem Referenzsignal kreuzkorreliert. Eine Zeitverzögerung wird dabei über die Phasenlage der Korrelate zueinander detektiert. Als Nutzsignal verwendet man Schwingungspakete, die sich zeitlich in ihrer Frequenz ändern (*Chirps*¹). Das Empfangssignal wird am Eingang möglichst hochfrequent abgetastet und mit einer Auflösung von einem Bit quantisiert. Es wird also nur Polarität des Empfangssignals erfasst. Die digitalen Abtastwerte werden anschließend mit einem im Empfänger gespeicherten Referenzsignal kreuzkorreliert. Dieses Referenzsignal liegt dem Empfänger als quasi Referenzsignal in 1-Bit quantisierter Form vor. Die größte Korrelation erfolgt dann, wenn das erwartete Signal am Empfänger eingetroffen ist. Eine Maximalwertsuche oder eine Schwellwertdetektion am Korrelationsergebnis kann anschließend diesen Eintreffzeitpunkt ermitteln. Wird das Verfahren auf mehrere Sensoren angewandt, können somit Laufzeitdifferenzen zwischen diesen Sensoren detektiert werden (siehe Abb. 5.11). Bei einer 1-Bit Quantisierung erfährt das bearbeitete Signal eine

x	y	$x \text{ XNOR } y$
1	0	0
1	1	1
0	0	1
0	1	0

Tab. 5.2: Wahrheitstabelle XNOR Element

¹ siehe Kap. 4.3 auf Seite 32

nichtlinere Verzerrung. Die Breite des Hauptimpulses als auch der Abstand der Nullstellen bleiben jedoch unverändert. Chirpsignale besitzen die Eigenschaft, robust gegenüber nicht-linearen Verzerrungen zu sein ([ZI92] siehe auch Kap.4.3).

Die Kreuzkorrelation zwischen Eingangssignal und Referenzsignal kann mittels XNOR-Gattern realisiert werden. Die XNOR-Gatter vergleichen den Eingangsbitstrom mit dem Referenzsignal, anschließendes Summieren aller Ergebnis der Vergleichsoperatoren liefert das Korrelat. Auf einem DSP kann die Kreuzkorrelation mittels eines FIR-Filters reali-

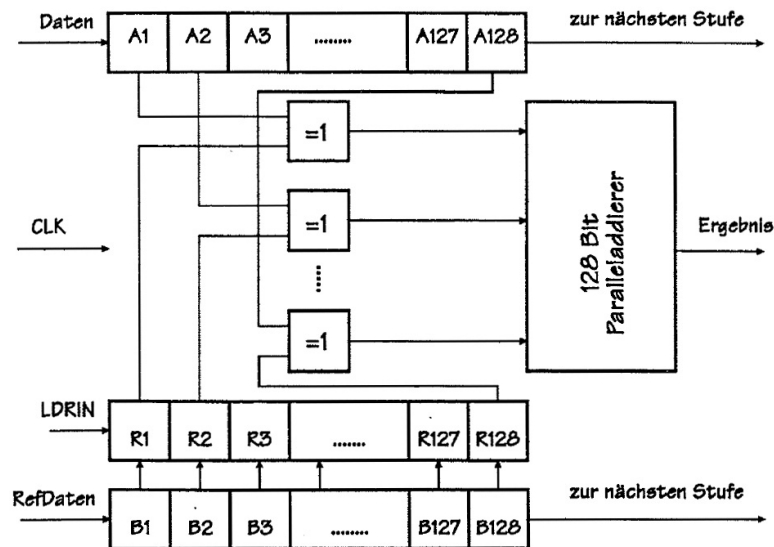


Abb. 5.12: Blockschaltbild eines 128 stufigen 1Bit-Korrelators ([ZI92])

siert werden, wobei das Referenzsignal (um die Zeitachse gedreht) die Filterkoeffizienten $b_0 - b_{N-1}$ liefert (siehe Abb.5.13).

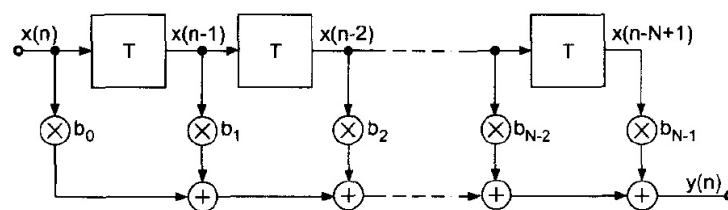


Abb. 5.13: 1-Bit Korrelation mittels FIR-Filter ([ZO02])

5.3.3.1 Einfluss der Referenzsignallänge

Die Wahrscheinlichkeit einer Signalerkennung ist proportional zur Länge des Referenzsignals.

Für den 1Bit-Korrelator mit der Stufenzahl N , der die Operation

$$y(m) = \sum_{n=0}^{N-1} f(n) \equiv n(m) \quad (5.44)$$

$f(m)$: Korrelationsfolge
 $n(m)$: Zufallsfolge vom weißen Rauschen

durchführt, erhält man den Erwartungswert $E(y)$.

$$E(y) = E\left(\sum_{m=0}^{N-1} f(m) \equiv n(m)\right) \quad (5.45)$$

$$= \sum_{m=0}^{N-1} p_f(m) \equiv s(m) \quad (5.46)$$

$$= N \cdot p_f(m) \equiv n(m) \quad (5.47)$$

In dieser Gleichung bezeichnet $p_f(m) \equiv s(m)$ die Wahrscheinlichkeit dafür, dass ein Bit des quantisierten Eingangssignals $n(m)$ mit dem entsprechenden Vergleichsbit im Korrelator übereinstimmt. Bei weißem Rauschen erhält man für $n(m)$ eine gleichverteilte Zufallsfolge, somit beträgt $p_f(m) \equiv n(m) = p_m = 0.5$.

Daraus folgt:

$$E(y) = \frac{N}{2} \quad (5.48)$$

Das zu erwartende Ausgangssignal mit einem reinen Rauschsignal $s(t)$ am Eingang und einer beliebigen Vergleichsfolge des Korrelators beträgt also $\frac{N}{2}$, ein Korrelationsergebnis von 0 würde bedeuten, dass die Eingangsfolge mit dem Einerkomplement von $c(m)$ identisch ist.

Für die folgende Überlegungen wird der Korrelationsbereich von $0..N$ auf $-\frac{N}{2}$ bis $\frac{N}{2}$ entsprechend -100 bis 100% umgerechnet. Eine Korrelation von 0% entspricht somit einem Empfangssignal, welches nur weißes Rauschen enthält.

Bei einer Korrelatorstufenzahl N und einer Wahrscheinlichkeit der Übereinstimmung pro Stufe von p_m erhält man für eine 100%ige Korrelationswahrscheinlichkeit, die allein durch das Rauschsignal bewirkt wird:

$$p_{100\%} = p_m^N \quad (5.49)$$

Dies bedeutet, dass dies bei kleinen Stufenzahlen bis $N = 15$ ($p_{100\%} = \frac{1}{2}^{15} = \frac{1}{32768}$) bei 30000 Taktzyklen pro Messung durchschnittlich einmal pro Messvorgang vorkommt. Für

eine einwandfreie Funktion muss die Stufenzahl daher deutlich höher sein (z.B. $N > 32$). Ein Korrelator mit N Stufen kann insgesamt 2^N verschiedene Zustände annehmen. Bei zufälligem, gleichverteiltem Eingangssignal $n(m)$ ist die Wahrscheinlichkeit, dass ein spezieller Zustand angenommen wird, p_m^N . Von allen möglichen Zuständen gibt es nur einen, bei der alle Stufen mit der Vergleichsfunktion übereinstimmen. Die Anzahl der möglichen Kombinationen für k Übereinstimmungen berechnet sich aus dem Binomialkoeffizienten

$$\binom{N}{k} = \frac{N!}{k! \cdot (N-k)!} \quad (5.50)$$

Die Wahrscheinlichkeit p_k für das Korrelationsergebnis k erhält man durch Multiplikation der Anzahl der Kombinationsmöglichkeiten und der Wahrscheinlichkeit für eine Kombination.

$$p_k = \binom{N}{k} \cdot p_m^N \quad (5.51)$$

Eine Vergrößerung des Referenzsignals erhöht also zum einen den SNR, zum anderen aber auch die benötigte Rechenleistung und die Latenz der Detektion.

5.3.3.2 Separation verschiedener Signale

Bei Verwendung mehrerer Ultraschallsender in enger örtlicher Nachbarschaft kann es zu Überlagerung der verschiedenen Signale und damit zu ungenauer oder gar unmöglicher Detektion kommen. Das hier gezeigte Verfahren besitzt die günstige Eigenschaft, im gleichen Frequenzbereich liegende Störsignale (je nach Grad der Korrelation zum Nutzsignal) zeitlich zu dehnen und in der Amplitude zu verkleinern. Ein gleichzeitiger Betrieb mehrerer Systemen in unmittelbarer Nachbarschaft ist bei geeigneter Wahl der Nutzsignale im gleichen Frequenzband möglich. Je unkorrelierter die Nutzsignale der gleichzeitig betriebenen Systeme zueinander sind, desto weniger wird das Nachbarsystem Einfluss auf die Detektion haben. Für diesen Anwendungsfall müssen also Chirpsignale verwendet werden, deren Vektoren orthogonal zueinander sind. Abb. 5.14 zeigt die Reaktion einer 1-Bit Korrelation auf verschiedene Störsignale: links ein Chirp mit fallender Frequenz, in der Mitte mit steigender und fallender Frequenz sowie rechts das vom Korrelator erwartete Nutzsignal. Wie in Kap. 2.2.2 und Kap. 6 erwähnt wird, erfährt das ideale Sendesignal auf dem Weg zum Empfänger Verzerrungen und Störungen. Dabei zeigt sich, dass diese Verfahren sehr unempfindlich gegen Amplitudenmodulationen ist. Es detektiert auch bei sehr geringen SNR ein empfangene Signal richtig.

Eine Optimierung dieses Verfahrens findet man in [WA96]. Hier handelt es sich um eine 2-Bit Korrelation, wobei für kleinen Signalamplituden das zusätzliche Bit eine bessere Auflösung gegenüber dem Grundrauschen besitzt. Der Schwellwert zum Setzen dieses zusätzlichen Bits ist adaptiv. Damit wird ein höherer Signal/Rausch-Abstand gegenüber der 1-Bit Korrelation erzielt.

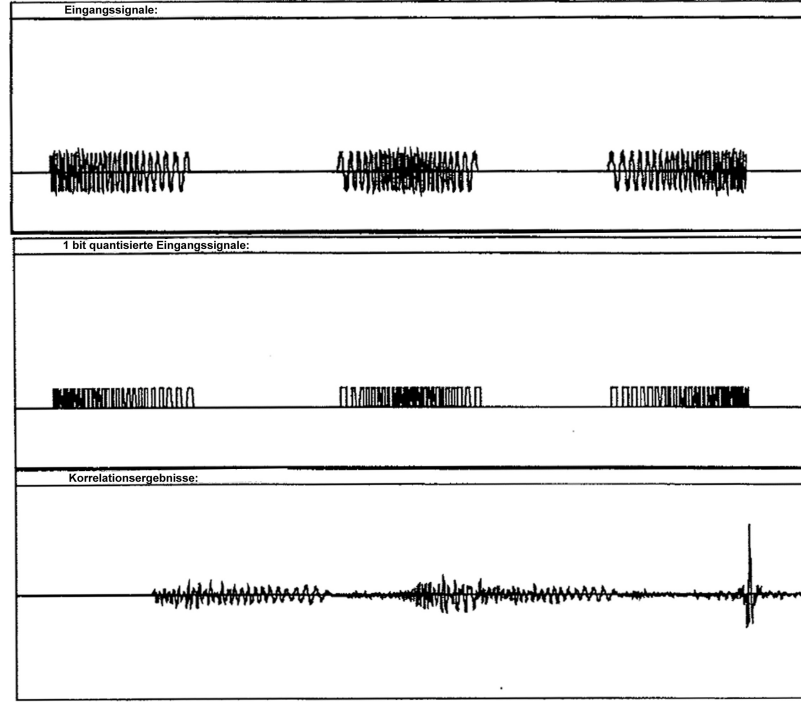


Abb. 5.14: Korrelation mit verschiedenen Eingangssignalen ([ZI92])

5.4 Alternativen zur Kreuzkorrelation

Average Square Difference Function: Der Korrelation vergleichbare Ansätze werden in ([JA93]) vorgestellt und evaluiert. Dabei handelt es sich um Verfahren, die dem Autor nach, weniger Rechenaufwand als die klassische Kreuzkorrelation benötigen. Ob der Rechenaufwand in der Praxis wirklich minimiert werden kann, liegt jedoch in erster Linie an der Architektur des verwendeten Signalprozessors. Bei beiden Ansätzen ergibt die Position des Minimums der berechneten Funktion die Laufzeitdifferenz D . Die Auflösung kann danach durch Interpolation verbessert werden.

$$\hat{R}_{ASDF}(m) = \frac{1}{N} \sum_{k=1}^N [x_1(k) - x_2(k+m)]^2 \quad (5.52)$$

$$m_{min} = \arg \min \hat{R}_{ASDF}(m) \quad (5.53)$$

$$\hat{\delta}_{ASDF} = -\frac{\hat{R}_{ASDF}(m_{min}+1) - \hat{R}_{ASDF}(m_{min}-1)}{2 [\hat{R}_{ASDF}(m_{min}+1) - 2\hat{R}_{ASDF}(m_{min}) + \hat{R}_{ASDF}(m_{min}-1)]} \quad (5.54)$$

$$\hat{D}_{frac} = \frac{m_{min} + \hat{\delta}_{ASDF}}{f_s} \quad (5.55)$$

Average Magnitude Difference Function: Hier sei besonders hervorzuheben, dass die *AMDF* praktisch ohne Multiplikationen auskommt.

$$\hat{R}_{AMDF}(n) = \frac{1}{N} \sum_{k=1}^N |x_1(k) - x_2(k+m)| \quad (5.56)$$

$$m_{min} = \arg \min \hat{R}_{AMDF}(m) \quad (5.57)$$

$$\hat{\delta}_{AMDF} = -\frac{\hat{R}_{AMDF}^2(m_{min} + 1) - \hat{R}_{AMDF}^2(m_{min} - 1)}{2 [\hat{R}_{AMDF}^2(m_{min} + 1) + \hat{R}_{AMDF}^2(m_{min} - 1) - 2\hat{R}_{AMDF}^2(m_{min})]} \quad (5.58)$$

$$\hat{D}_{frac} = \frac{m_{min} + \hat{\delta}_{AMDF}}{f_s} \quad (5.59)$$

Die Varianz dieser beiden Schätzer ist im niedrigen und mittleren SNR-Bereich etwa gleich der Varianz der Kreuzkorrelation ohne Fensterfunktion. Mit steigendem SNR erreichen *ASDF* und *AMDF* jedoch eine geringere Varianz, die sich sogar 0 annähert. Die Varianz der *AMDF* ist geringfügig höher als die der *ASDF*. Siehe dazu Abb.5.15.

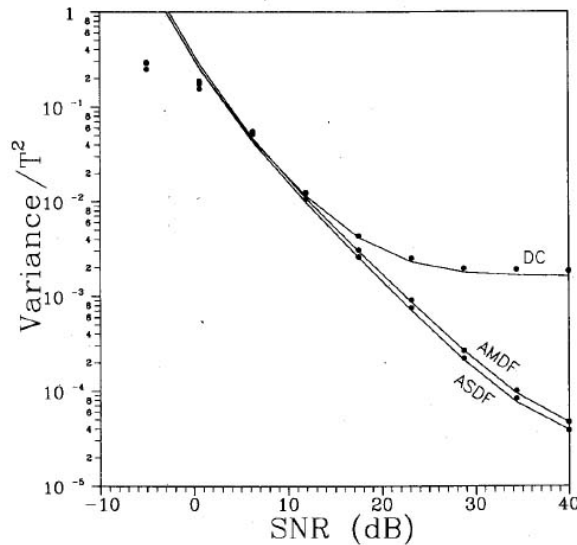


Abb. 5.15: Varianz als Funktion vom SNR ([JA93])

DC: direkte Kreuzkorrelation ohne Vorfilterung; AMDF: Average Magnitude Difference Function; ASDF: Average Magnitude Difference Function

MSX und MXS Um auch bei einem besonders geringem SNR eine verlässliche Aussage über die Laufzeitdifferenz zu erhalten, bieten sich die *MSX*- und *MXS*-Schätzer an (siehe Abb.5.16). Hierbei wird die Kreuzkorrelationsfunktion von x_1 und x_2 mit den Autokorrelationsfunktionen der beiden Signale verglichen ([NA95]).

Mit der mathematischen Beschreibung der Eingangssignale aus Gl.5.5 und 5.6 kann für

das schmalbandige Nutzsignal s_1 eine Beziehung von Kreuz- (Gl.5.60) und Autokorrelation (Gl.5.61) hergestellt werden. Da man hier von unkorreliertem (weißem) Rauschen ausgeht, tritt die Rauschleistung nur in Form eines gewichteten Diracimpulses in der Autokorrelationsfunktion in Gl.5.61 auf.

$$\begin{aligned}\varepsilon \{ \hat{R}_{x_1, x_2}(\tau) \} &= \varepsilon \left\{ \frac{1}{N-\tau} \sum_{t=0}^{N-1-\tau} s(t) s(t-D+\tau) \right\} \\ &= R_{s,s}(\tau-D)\end{aligned}\quad (5.60)$$

$$\begin{aligned}\varepsilon \{ \hat{R}_{x_1, x_1}(\tau) \} &= \varepsilon \left\{ \frac{1}{N-\tau} \sum_{t=0}^{N-1-\tau} s(t) s(t+\tau) + \frac{1}{N-\tau} \sum_{t=0}^{N-1-\tau} n_1(t) n_1(t+\tau) \right\} \\ &= R_{s,s}(\tau) + \sigma_{n_1}^2 \delta(\tau)\end{aligned}\quad (5.61)$$

Somit gilt für alle $\tau \neq 0$:

$$\varepsilon \{ \hat{R}_{x_1, x_1}(\tau) \} = R_{s,s}(\tau) = \varepsilon \{ \hat{R}_{x_1, x_2}(\tau+D) \} \quad (5.62)$$

und genauso für jedes $\tau \neq D$:

$$\varepsilon \{ \hat{R}_{x_1, x_2}(\tau) \} = R_{s,s}(\tau-D) = \varepsilon \{ \hat{R}_{x_1, x_1}(\tau-D) \} \quad (5.63)$$

Anhand obiger Beziehungen ist es nun möglich¹ τ zu finden, wofür $\hat{R}_{x_1, x_2}(\tau_{x_1, x_2}) = \hat{R}_{x_1, x_1}(\tau_{x_1, x_1})$ ist. Nach einer Interpolation des Wertes τ_{x_1, x_1} in Gl.5.66 und τ_{x_1, x_2} erhält man nun die Laufzeitdifferenz:

$$\hat{D}_{MXS} = \tau_{x_1, x_2} - \tau_{x_1, x_1} \quad (5.64)$$

$$\hat{D}_{MSX} = \tau_{x_1, x_1} - \tau_{x_2, x_1} \quad (5.65)$$

$$\tau_{x_1, x_1} = \frac{-(f_1 - f_{-1}) \pm \sqrt{(f_1 - f_{-1})^2 - 8(f_0 - f_D)(f_1 - 2f_0 + f_{-1})}}{2(f_1 - 2f_0 + f_{-1})} \quad (5.66)$$

mit:

$$\begin{array}{ll} f_{-1}: & \hat{R}_{x_1, x_1}(\tau-1) \\ f_0: & \hat{R}_{x_1, x_1}(\tau) \\ f_1: & \hat{R}_{x_1, x_1}(\tau+1) \\ f_D: & \hat{R}_{x_1, x_2}(\tau_{x_1, x_2}) \end{array}$$

Aus der Literatur ([NA95]) sind nur Genauigkeitsvergleiche mit anderen Algorithmen für kleine Zeitdifferenzen bis zu 2 Samples bekannt. In Abb.5.16 lässt sich gut erkennen, welche Genauigkeitsverbesserung MSX und MXS hinsichtlich eines geringen SNR hat. Diese

¹Der Autor von [NA95] führt dafür zwei Ansätze an.

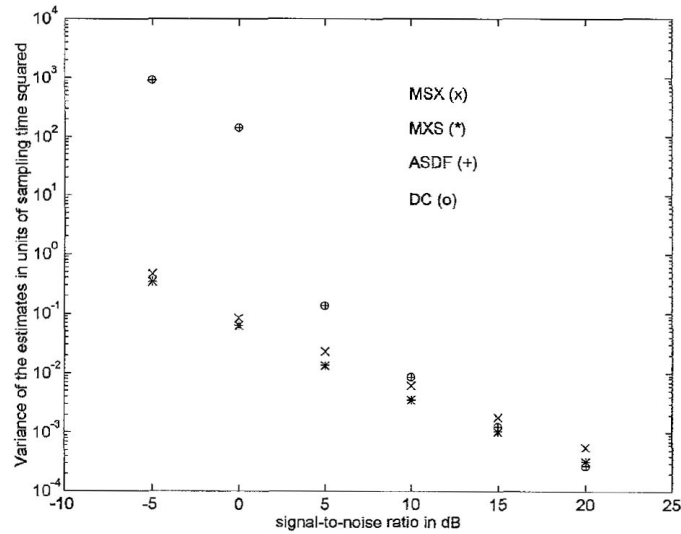


Abb. 5.16: Varianz als Funktion vom SNR für die Detektion einer Laufzeitdifferenz von 0.1 Sample ([NA95])

Genauigkeitssteigerung erfordert allerdings etwa den dreifachen Rechenaufwand¹ im Vergleich zu einer direkten Kreuzkorrelation. Fundamentaler Grundgedanke dieser Methode ist, dass die Rauschbeläge n_1 und n_2 zueinander sowie zu s vollständig unkorreliert sind. Wird in der Praxis diese Voraussetzung nicht eingehalten, steigt auch die Varianz dieser Schätzungsmethode.

5.5 Detektion des Dopplereffektes

Wie in Kap.2.3.3 beschrieben, verändert sich bei einer Bewegung von Sender bzw. Empfänger abhängig von der relativen Bewegungsgeschwindigkeit Δv die empfangene Frequenz mit einer Frequenzverschiebung von Δf ([RU97]). Das Sendesignal wird also frequenzmoduliert. Ziel des hier beschriebenen Verfahrens ist es, von dieser Frequenzmodulation auf die Geschwindigkeit der Kopfbewegung zu schließen. Vorausgesetzt die Kopfposition sei zu einem Referenzzeitpunkt t_0 bekannt, kann von der relativen Positionsänderung auf den momentanen Kopfwinkel des Hörers geschlossen werden. Bei diesem Verfahren wird ein schmalbandiges CW-Signal² mit konstanter Frequenz verwendet, damit die Dopplerverschiebung zu jedem Zeitpunkt ermittelt werden kann. Laut den Autoren von [KU99] und [KU00] kann somit eine bessere zeitliche Auflösung als bei Gated-Wave-Methoden erreicht werden. Raumreflexionen können bei diesem Ansatz zu großen Problemen führen, da CW-Signale sehr reflexionsanfällig sind. In [KU02] werden mit Hilfe einer Wavelet-Filterbank dieses Problem minimiert.

¹Hier muss zusätzlich zu G_{x_1, x_2} noch die Autokorrelationen G_{x_1, x_1} und G_{x_2, x_2} gerechnet werden.

²CW - Continuous Wave siehe Kap. 4.2 auf Seite 31

5.5.1 Messung der Periodendauer

Da Periodendauer P und Frequenz f eines Signals über $P = 1/f$ im Zusammenhang stehen, führt die Dopplergleichung 2.26 zu:

$$v = \frac{c \cdot \Delta P}{P} \quad (5.67)$$

v : momentane Relativgeschwindigkeit
 P : Periodenlänge der versendeten Signals
 ΔP : Abweichung der gemessenen Periodendauer von P

Die Verschiebung des Sensors ist damit das Produkt von Schallgeschwindigkeit und ΔP :

$$d = v \cdot P = c \cdot \Delta P \quad (5.68)$$

In [KU99] und [KU00] wird ein empfangenes 40 kHz CW-Signal mittels einer Anordnung nach Abb.5.17 unterabgetastet und in ein Rechtecksignal umgewandelt. Steigende Flanken des Rechtecksignals werden von einem hochtaktigen Zählwerk gezählt. Die Zahl der gezählten Flanken K pro Zeiteinheit $N \cdot P_{clock}$ ist ein Maß für die Periodendauer P . Die Verschiebung des Sensors d kann daher nach Gl.5.69 berechnet werden. Der Wert K_{Ruhe} ist dabei der Referenzzählerstand, wenn Sender bzw. Empfänger nicht bewegt werden. Wie schon erwähnt bewirken Reflexionen Störungen der Periodenmessung, die bei diesem Verfahren zu Messfehlern führen. Sind keine Reflexionen zu erwarten, hängt die Genauigkeit des Verfahrens hauptsächlich von der Taktrate des Zählers ab.

$$d(N \cdot P) = \frac{c}{P_{clock}} \sum_{n=1}^N K_{Ruhe} - K_{gemessen}(n) \quad (5.69)$$

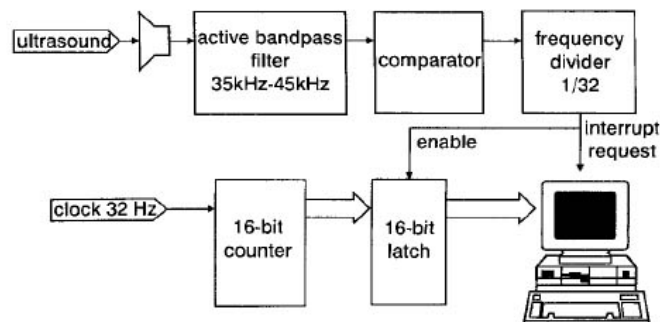


Abb. 5.17: Strukturdiagramm eines Doppler-Trackingsystems bei Messung der Periodendauer ([KU00])

5.5.2 Frequenzmessung mittels Kurzzeit Fourier-Transformation

Mittels der Kurzzeit Fourier-Transformation (*Short Time Fourier Transformation* - *STFT*) wird das Signalspektrum vom Signal x zum Zeitpunkt t mittels eines kurzen Zeitfensters der Länge N um t herum geschätzt:

$$Y_{short}(t, \omega) = \sum_{m=t-0.5N}^{t+0.5N} x(m) e^{\frac{-j\omega m}{f_s}} \quad (5.70)$$

In Gl.5.70 wird als Fensterfunktion ein Rechteckfenster benutzt.

Da die eintreffenden Echos eine geringere Signalamplitude als das Nutzsignal aufweisen, kann im *STFT*-Spektrum die Frequenz des Hauptsignals am Frequenzbin mit der größten Intensität erkannt werden¹. Ist die Frequenz der Reflexionen allerdings sehr nahe an der Frequenz des CW-Signals verursachen die überlappenden Spektren Fehler in der Frequenzbestimmung. Eine Verbesserung in der spektralen Darstellung können eventuell andere Fensterfunktionen wie Hanning- oder Blackmanfenster bewirken ([KU02]).

Abbildung 5.18 zeigt die Ergebnisse einer Simulation mit einer bewegten Schallquelle und 2 Wandreflexionen. Im Vergleich zur Bestimmung der Momentanfrequenz durch eine Periodendauermessung (Abb.5.18(c)) ist die *STFT*-Methode nicht so störanfällig. Wie in der Abb.5.18(d) gut zu sehen, stört aber auch bei der *STFT* eine der beiden Raumreflexion das Messergebnis.

5.5.3 Störunterdrückung mittels Wavelet-Filterbank

Störende Reflexionen während der *STFT* können in [KU02] mit Hilfe einer Wavelet-Filterbank (Abb.5.21) beseitigt werden. Dabei wird von einer direkten Beziehung zwischen Amplitudenhüllkurve und Frequenzverzerrungen ausgegangen. Es wird angenommen, das empfangene CW-Signal sei eine Summe aus Direktschall und Raumreflexionen:

$$x_{noise} = \sum_{k=1}^m x_{reflexion,k} = A_{noise}(t) e^{j\omega_{noise}(t) \cdot t} \quad (5.71)$$

$$x_{direct} = e^{j\omega_{direct}(t) \cdot t} \quad (5.72)$$

$$x = x_{direct} + x_{noise} \quad (5.73)$$

$$x = A e^{j \cdot [\omega_{direct}(t) + \omega_{distortion}(t)] \cdot t} \quad (5.74)$$

¹In der Praxis muss diese Annahme nicht immer stimmen. Durch die Richtcharakteristik des Empfängers können bei bestimmten Einfallswinkeln die Reflexionen größere Amplituden aufweisen als der Direktschall. Siehe dazu Abb. 6.10 auf Seite 77.

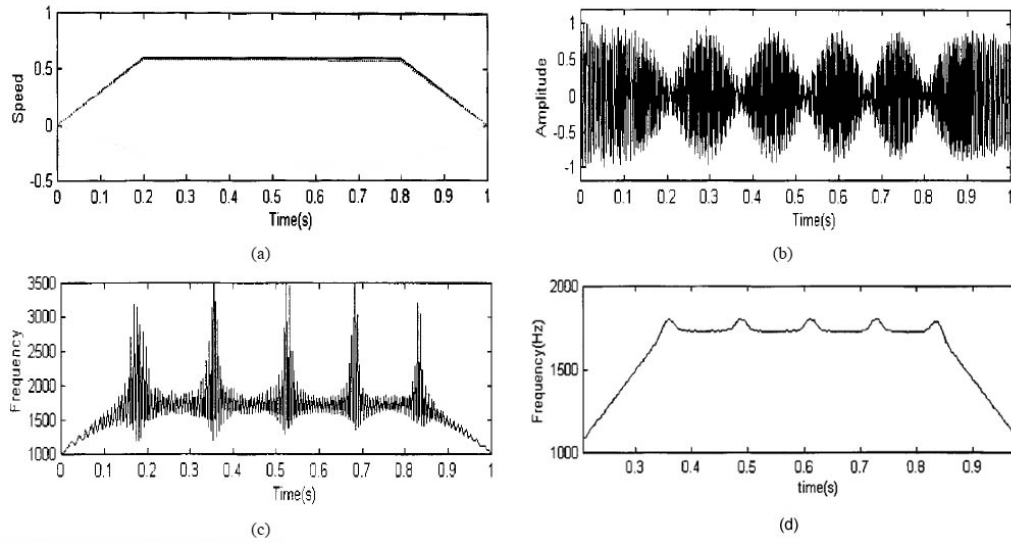


Abb. 5.18: (a) Relativgeschwindigkeit v (b) empfangenes Ultraschallsignal (c) Ergebnis nach Periodendauermessung (d) Zeit-Frequenzspektrum aus STFT-Analyse ([KU02])

mit:

$$\omega_{distortion} = \frac{d \left[\arccos \left(\frac{1 + A_{noise}(t) \cdot \cos(\Delta\omega(t) \cdot t)}{A} \right) \right]}{dt}$$

$$A = \sqrt{1 + A_{noise}^2(t) + 2A_{noise}(t) \cdot \cos(\Delta\omega(t) \cdot t)}$$

$$\Delta\omega(t) = \omega_{direct}(t) - \omega_{noise}(t)$$

A_{noise} : Dämpfung aller Störungen

Aus Gl.5.74 erkennt man, dass sowohl die Amplitude A vom empfangenen Signal x als auch die Frequenz der Störung $\omega_{distortion}$ von der Differenz der Frequenzen von Direktsignal ω_{direct} zu Reflexionen ω_{noise} abhängen und mit $2\pi/\Delta\omega(t)$ periodisch sind.

Zur Bestimmung der Gewichte in der Wavelet-filterbank wird die invertierte Hüllkurve des empfangenen Signals herangezogen. Durch die Invertierung bleibt das Signalspektrum erhalten, es können aber hierdurch besonders gut höhere Störfrequenzen erkannt werden (siehe 5.19). Das Spektrum der invertierten Hüllkurve ist somit eine spektrale Schätzung aller Störungen. Dieses Störspektrum gewichtet nun adaptive Filter, die aus dem empfangenen Signal aus Abb.5.18(b) den Direktschall aufdeckt (siehe 5.21). Abb.5.20 zeigt die Optimierung der Wavelet-Filterbank im Vergleich.

Eine Umsetzung des Doppleransatzes innerhalb eines Headtrackers ist nicht bekannt. Um auch kurze, langsame Kopfbewegungen mit der Dopplermethode exakt erfassen zu können,

ist ein sehr schnelles und gleichzeitig feines Frequenztracking notwendig. Wie bei bestehenden Inertialsystemen (siehe Kap. 1.2.2) wird es wahrscheinlich auch zu einem Driftproblem kommen. Auch im Hinblick auf die Störunsicherheiten ist dieses Verfahren wahrscheinlich nur mit sehr viel Aufwand zu realisieren.

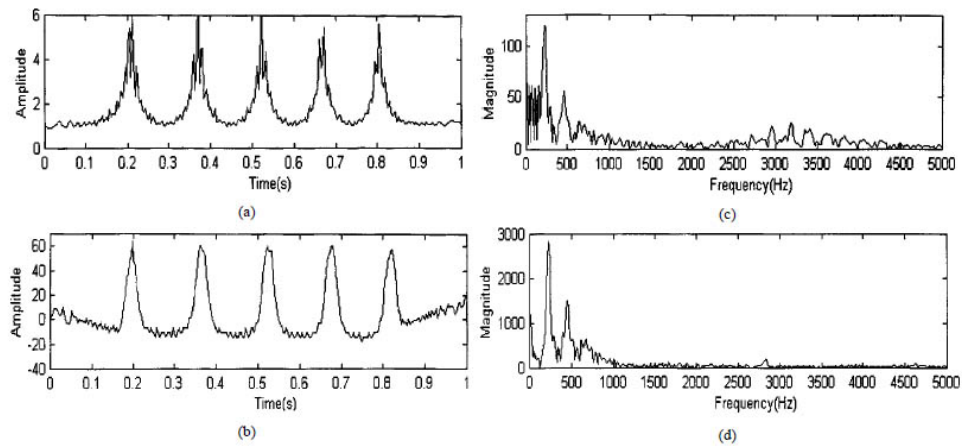


Abb. 5.19: (a) invertierte Hüllkurve des Signals aus Abb. 5.18(b); (b) Verzerrungen *STFT* Frequenzmessung; (c) Spektrum von (a); (d) Spektrum von (b) ([KU02])

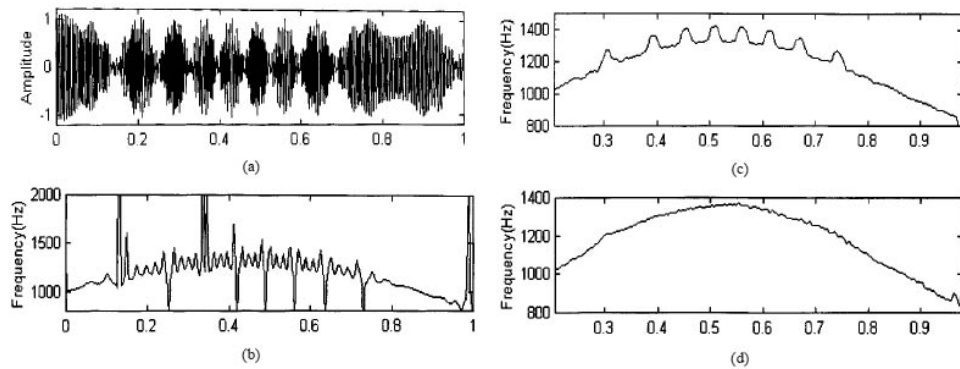


Abb. 5.20: Vergleich Echoanfälligkeit der Dopplerverfahren; (a) Empfangenes Signal; (b) Messung der Periodendauer; (c) *STFT*; (d) adaptiver Wavelet-Filter ([KU02])

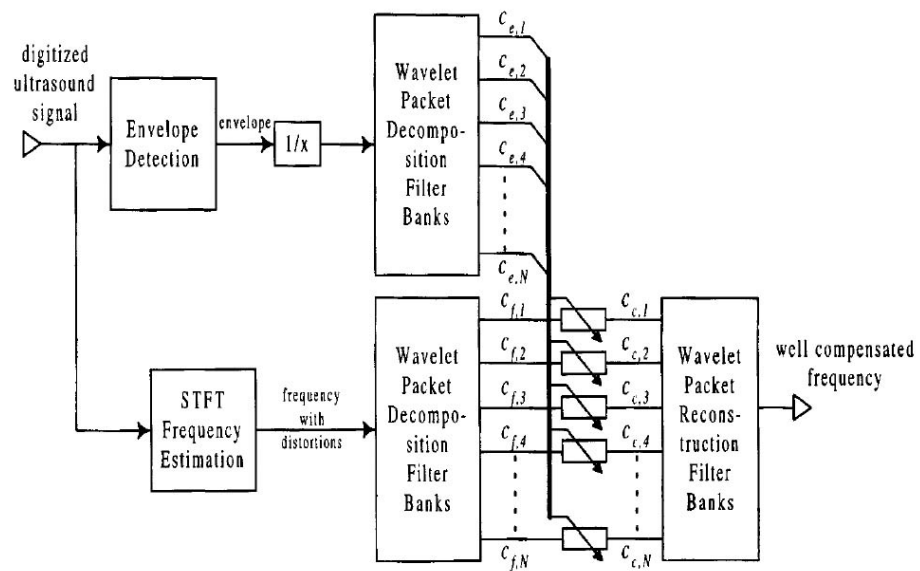


Abb. 5.21: Blockschaltbild einer adaptiven Wavelet-Filterbank ([KU02])

Kapitel 6

Praktische Aspekte für die Realisierung

6.1 Störgeräusche und Grundrauschen

Der Frequenzbereich von Ultraschall ist im Vergleich zum Hörschallbereich nur sehr wenig genutzt. Es gibt jedoch durchaus einige elektrische Geräte, die Ultraschallfrequenzen produzieren. Dies können neben Fernbedienungen vor allem Computerbildschirme¹, Staubsauger und andere elektrische Reinigungsgeräte sein. Ultraschall kann auch an Luftdruckschläuchen, Rollräder sowie durch schnell rotierende Ventilatoren erzeugt werden. Alarmanlagen und Abstandssensoren, wie sie häufig in Automobilen Verwendung finden, arbeiten meistens ebenfalls auf der Basis von Ultraschall. Auch nutzen einige Tiergattungen (z.B. Fledermäuse und Ratten) Ultraschall zur Orientierung und Kommunikation². Aufgrund der Dämpfungsverluste bei der Schallausbreitung können glücklicherweise nur Ultraschallquellen in direkter Nachbarschaft zum Headtracker als Störer fungieren. Das Störungsrisiko lässt sich durch die Verwendung von schmalbandige Ultraschallsignale und entsprechender empfängerseitiger Bandpassfilterung minimieren. Bei Versuchen mit der Sensorstrecke im Auto konnten so ohne aktivierte Bordelektronik mehr als 70 dB Signal/Rauschabstand im interessierenden Frequenzbereich von 39 bis 41kHz erreicht werden. Es ist zu klären inwieweit eine aktivierte Bordelektronik den Signal/Rauschabstand minimiert.

¹Eletronenstrahlröhren in älteren Bildschirmen können Störfrequenzen erzeugen

²siehe dazu auch Anhang D auf Seite 121

6.2 Psychoakustische Parameter

6.2.1 Wahrnehmbare Drehwinkeländerungen

Für die notwendige Winkelgenauigkeit des Headtrackers ist die menschliche Wahrnehmungsfähigkeit maßgebend. Um die Wahrnehmungsfähigkeit beschreiben zu können sind in der Literatur u.a. folgende Definitionen geläufig:

MAA Der *Minimum Audible Angle* beschreibt den kleinsten Winkel, bei dem zwei Schallquellen getrennt voneinander wahrgenommen werden können ([DA91]). Diese Größe ist abhängig von Frequenz, Dauer und Bandbreite der Stimuli.

In [WZ97] wird ein azimuthaler MAA von 1° bis 5° (frontal) bzw. 5° bis 10° für die 90° -Richtung angegeben.

MAMA Der *Minimum Audible Movement Angle* beschreibt die minimale Winkelauslenkung einer pendelnden Schallquelle, bei der die Richtung der Bewegung wahrgenommen werden kann ([DA91]). Dieser Winkel ist neben Frequenz und Bandbreite auch von der Bewegungsrichtung¹ und -geschwindigkeit abhängig.

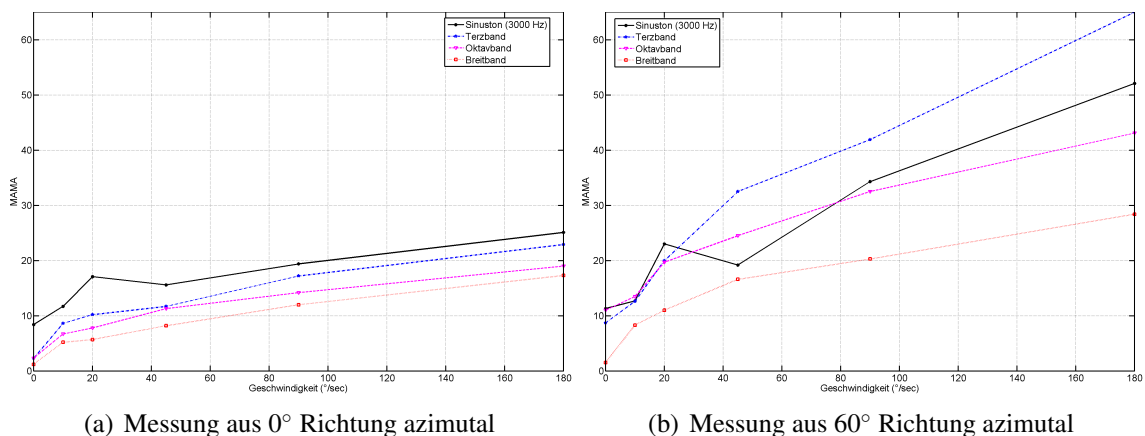


Abb. 6.1: MAMA als Funktion von Geschwindigkeit und Stimulus (nach [DA91])

Abb.6.1 zeigt den MAMA als Funktion von Geschwindigkeit bei verschiedenen Bandbreiten eines Stimulus. Je größer die Bandbreite und langsamer eine Schallquelle bewegt wird, desto besser ist die Wahrnehmungsfähigkeit des Ohres. Der MAMA für einen Breitbandstimulus beträgt bei geringen Geschwindigkeiten ($20^\circ/\text{sec}$) schon über 5° . Ein Vergleich von Abb.6.1(a) mit Abb.6.1(b) zeigt, dass die Wahrnehmungsfähigkeit bei größeren Azimutalwinkeln sinkt - die MAMAs sind durchschnittlich 1.5 bis 3 mal größer. Das Auflösungsvermögen des Ohres ist also seitlich geringer.

¹Damit sind Bewegungen in der horizontalen, vertikalen und diagonalen Ebene gemeint (siehe [GR03]).

6.2.2 Systemlatenzen

Die Dynamik des Systems wird hauptsächlich durch die Verzögerungen des Systems begrenzt. Damit der Anwender kein verzögertes, unreales Klangbild wahrnimmt, darf das Headtrackersystem (zusammen mit dem Binauralmodul) eine maximale Systemverzögerung nicht überschreiten. Die hauptsächlichsten Faktoren, die diese Verzögerung bestimmen, sind die Gesamtlatenz des Systems (*end-to-end latency*) sowie das Aktualisierungsintervall (*update-rate*).

update-rate Die *update-rate* gibt an, wie oft eine Kopfposition erfasst und berücksichtigt wird (zeitliche Diskretisierung). Sie ergibt sich aus dem Aktualisierungsintervall der Kopfbewegung, der Taktung der Schnittstelle zwischen Headtracker und Binauralmodul sowie dem Intervall in dem die Binauraleinheit neue Kopfpositionen aufnehmen kann. Da die Systemkomponenten meist nicht synchron getaktet sind, bestimmt sich die effektive *update-rate* aus der Taktrate des langsamsten Moduls innerhalb des Systems.

end-to-end latency Die Gesamtlatenz bezieht sich auf die Zeit die das System benötigt, um auf ein Ereignis eine passende Reaktion zu erzeugen. Bei einem Headtracker ist dies die Zeit ab der Kopfbewegung bis zum Vorliegen des entsprechend aufbereiteten Audiosignals beim Hörer.

6.2.2.1 Ergebnisse verschiedener Untersuchungen

Ergebnisse zur Untersuchung dieser Parameter finden sich u.a. in [MA98], [FE98], [WZ97], [SA96] und [ME92]. Die Aussagen weichen teilweise stark voneinander ab, da die Parameterwerte nur anhand von psychoakustischen Tests empirisch ermittelt werden können. Die Empfehlungen für eine maximale Gesamtlatenz variieren je nach Autor von 29ms bis 500ms¹. In [WZ97] wird ein Headtracker mit einer Update-Rate von 60Hz, einer Winkelauflösung von 1° und einer Systemlatenz von 29ms empfohlen. In [SA96] behauptet der Autor, dass Systeme bis 96ms Latenz und einer Update Rate von 10Hz nicht signifikant schlechter wahrgenommen werden als schnellere Geräte. Hierbei beeinflusst hauptsächlich die Winkelauflösung die Qualität des Höreindrucks. Eine aktuellere Untersuchung aus dem Jahr 2001 zeigt die Auswirkungen der Systemlatenz auf die Lokalisation bei unterschiedlich langen Stimuli [WZ01]. Diese Ergebnisse sind in Abb. 6.2 zu sehen. Ein Vergleich der vorgestellten Versuche ist nur bedingt möglich, da die Testergebnisse nicht ausschließlich vom Headtracker selbst, sondern von vielen weiteren Faktoren abhängen. (genutzter Stimulus, Binauralsystem, Probanden,...) Wird die Auralisation mit Hilfe eines Displays zusätzlich visuell unterstützt, kann eine geringere *update-rate* akzeptiert

¹Eine Latenz bis 500ms ist bei langsamen Bewegungen bis 8°/sec vertretbar ([PE88] zitiert in [WZ97])

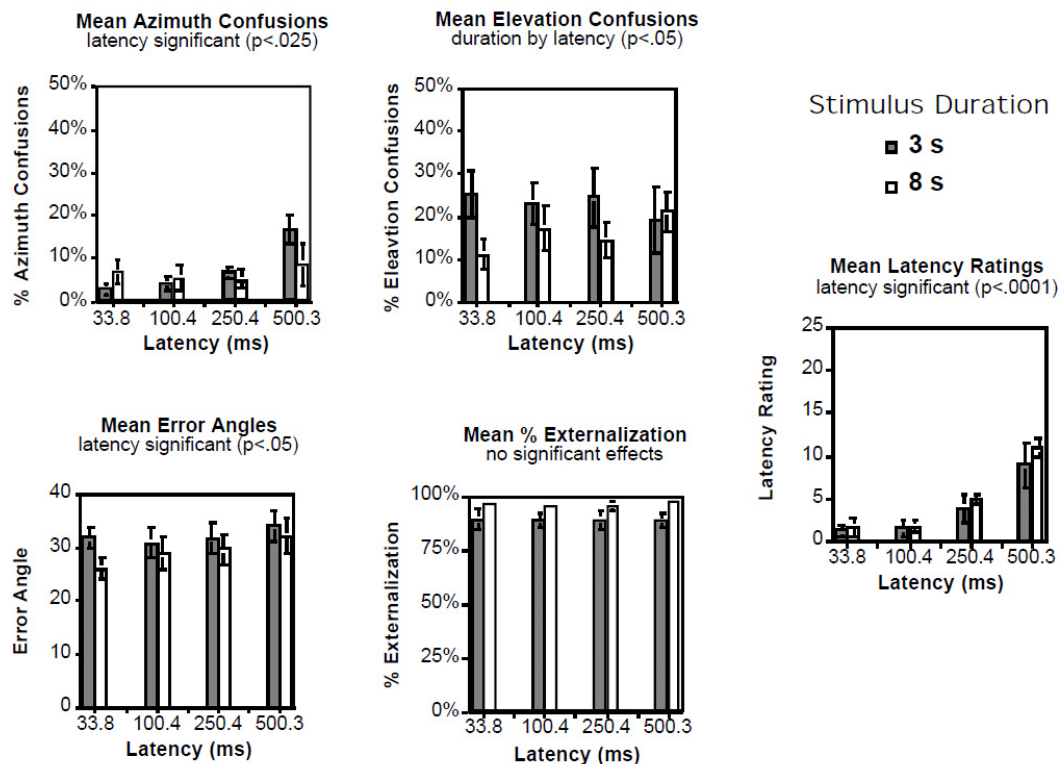


Abb. 6.2: Einfluss der Systemlatenz auf die Lokalisation - Statistische Auswertung der psychoakustischen Tests aus [WZ01]

werden ([WZ97]). Hierbei ist es besonders wichtig, die Lippensynchronität¹ zu bewahren.

6.3 Wahl der Sensoren

Zur Entfernungsmessung werden hauptsächlich piezoelektrische Ultraschallwandler benutzt, wenig verbreitet sind elektrostatische und magnetische Wandler ([SR02]).

Reale Ultraschallsensoren besitzen Eigenschaften, die sich im Vergleich zu idealen Sensoren bereits bei einfachen Ultraschallmessungen negativ bemerkbar machen. Besonders stark wird der Meßvorgang von den elektrischen Parametern der Schallwandler beeinflusst. (hauptsächlich Bandbreitenbegrenzung, Resonanzerscheinungen, Ein- und Ausschwingvorgänge)

Die Verringerung des Wandlereinflusses kann durch geeignete Signalverarbeitungsmethoden erfolgen.

¹Ein Zeitversatz zwischen Bild und Ton ist gerade bei der Darstellung sprechender Personen auffällig und störend.

6.3.1 Einfluss der Sensorengröße

In den in Kapitel 3 vorgestellten geometrischen Ansätzen wird von idealen Punktquellen- und Senken mit unendlich kleiner Oberfläche ausgegangen. Die Membrangröße realer Sensoren beeinflussen die Genauigkeit der Distanzmessung immer dann, wenn Sender und Empfänger nicht zueinander ausgerichtet sind. Abbildung 6.3 verdeutlicht das Problem:

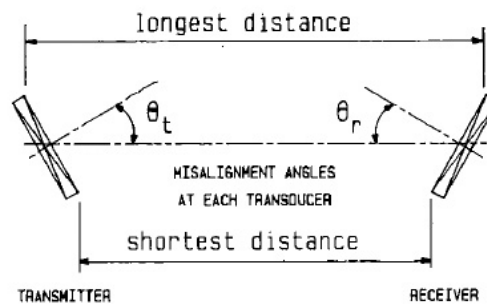


Abb. 6.3: Laufzeitenfehler aufgrund der Sensorgröße ([FI90])

Es entstehen folgende Situationen:

- Ein realer Wandler besitzt eine reale Membran um den geometrischen Mittelpunkt des Sensors. Trifft eine Schallwelle auf einen Teil dieser Membran, noch bevor sie das geometrische Zentrum des Sensors erreicht, wird dieser verfrühte Zeitpunkt detektiert, was in weiterer Folge eine zu kurze Wegstrecke ergibt.
- Eine auf der Bauteiloberfläche eintreffende Schallwelle verursacht eine transiente, unsymmetrische Auslenkung der Bauteiloberfläche. Diese Schwingung pflanzt sich bis zum eigentlichen Sensorelement fort und erreicht dieses noch vor dem eigentlichen akustischen Signal. Wiederum ist die ermittelte Distanz geringer als der räumliche Abstand zwischen den geometrischen Zentren der Wandler.

Sind also die Sensoren nicht zueinander ausgerichtet, wird praktisch immer ein zu kleiner Weg gemessen. Der Fehler zwischen tatsächlichem und gemessenem Weg steigt mit zunehmendem Winkel zwischen den Normalen der Sender und Empfänger an.

Bei einer Sensoranordnung nach Abb.3.9 wird der verursachte Laufzeitenfehler wahrscheinlich nicht sehr groß sein, da sich die Empfänger am Kopfhörer immer ungefähr im gleichen Drehwinkel zum Sender befinden. Jedoch soll dies als mögliche Fehlerursache nicht unerwähnt bleiben.

Ein guter Kompromiss für eine Näherung an einen idealen Kugelstahler erhält man, wenn der Durchmesser der transmittierenden Oberfläche in der Größe der zu sendenden Wellenlänge ist. Für gewöhnlich werden in Ultraschallsystemen Wandler benutzt, deren Oberfläche groß im Bezug auf die verwendeten Wellenlängen sind. Das bedeutet auch,

dass bei US-Systemen keine Beugungserscheinungen auftreten.¹

Dies führt sowohl zu einer Richtwirkung der Bauteile, als auch zu einer Phasenverschiebung zwischen der Welle und dem entsprechenden elektrischen Signal. Aufgrund der Phasenverschiebung ist der geometrische Ursprung des Signals auf der Sensoroberfläche nicht eindeutig bestimmbar ([FI90] und [FI92]).

Abbildung 6.4 zeigt den gemessenen Fehler abhängig von dem relativen Winkel zwischen Sender und Empfänger anhand von zwei unterschiedlich großen Bauteilpaaren ([FI90]²). Man sieht, dass kleinere Sensoren einen geringeren Fehler verursachen als

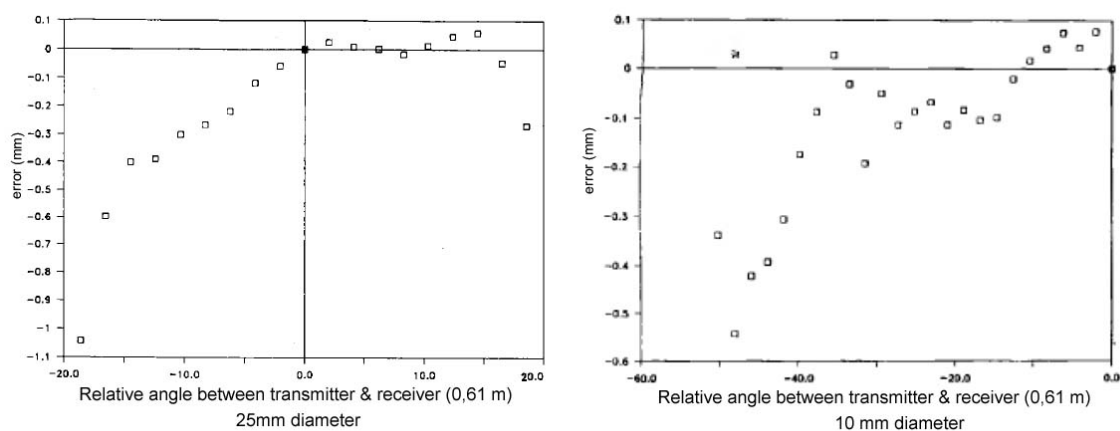


Abb. 6.4: Gegenüberstellung der Detektionsfehler bei versch. Sensordurchmessern ([FI90])

große Sensoren. Lässt man einen Fehler bis zu einem gewissen Grad zu, so haben kleinere Sensoren (bei gleichem Fehler) einen größeren „Arbeitswinkel“. Für größere Sensoren spricht hingegen die Tatsache, dass sie empfindlicher sind, da sie bei einer größeren Empfängermembran mehr Schallenergie aufnehmen.

¹Dies unterscheidet Ultraschallapplikationen von gewöhnlichen niederfrequenten Akustikanwendungen, wo die relevanten Wellenlängen meist viel größer als die Wandleroberfläche sind und sich daher die Schallwellen um Lautsprecher und Mikrofon herumbeugen.

²Die in der Untersuchung verwendeten Sensoren: SQ-40TE2/RE2-18 (10 mm Durchmesser); SQ-40TE1/RE1-25 (25 mm Durchmesser) beide von Projects Unlimited Inc.

6.4 Sensorpositionierung

Sender am Kopf: Befinden sich die Ultraschallsender am Kopfhörer und die Empfänger am Videobildschirm (siehe Skizze 6.5) ergeben sich folgende Vor- und Nachteile:

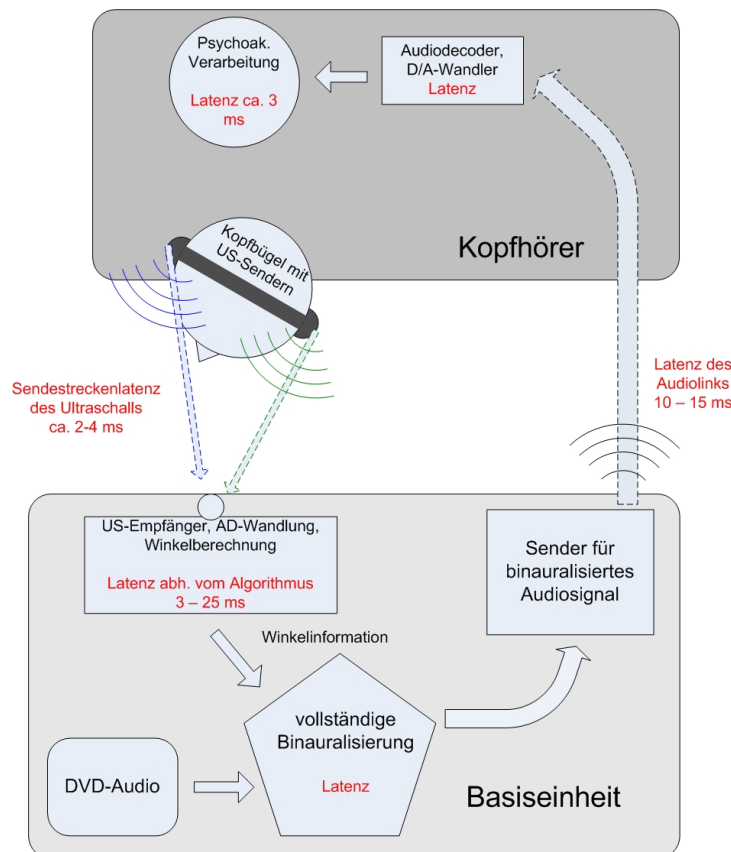


Abb. 6.5: System mit zwei Sendern am Kopfhörer

- Die Sender werden über den Kopfhörer-akku gespeist. Die Empfänger und der DSP sind in der DVD-Konsole eingebaut und werden von dort betrieben. Der Leistungsverbrauch am schnurlosen Kopfhörer ist somit gering.
- Eine erhöhte Anzahl von Empfängern steigert die Eindeutigkeit und die Auflösung des Systems. Diese Optimierung ist aber unabhängig vom Leistungsverbrauch des Kopfhörers.
- Die *end-to-end Latenz* des Systems steigt je nach Länge des Abstandes zwischen Sender und Empfänger unnötig an, da zusätzlich nach detektierter Kopfbewegung das aufbereitete Audiosignal zum Kopfhörer gefunkt werden muss. Die Funkstreckenlatenz fällt also zweimal an.

- Bei Verwendung von zwei oder mehr Kopfhörer steigt die Zahl der ausgesendeten und an der Empfangseinheit wieder zu trennenden Ultraschallsignale mit dem Faktor 2 an. Das bedeutet eine erhöhte Schwierigkeit bei der Signalseparation, was zu einer Steigerung von Rechenoperationen und damit möglicherweise zu Verzögerungen führt. Mitunter leidet dabei auch die Exaktheit und das Auflösungsvermögen des Headtrackers darunter.

Empfänger am Kopf: Befinden sich die Ultraschallempfänger am Kopfhörer und die Sender am TFT-Display (siehe Skizze 6.6) bedeutet dies:

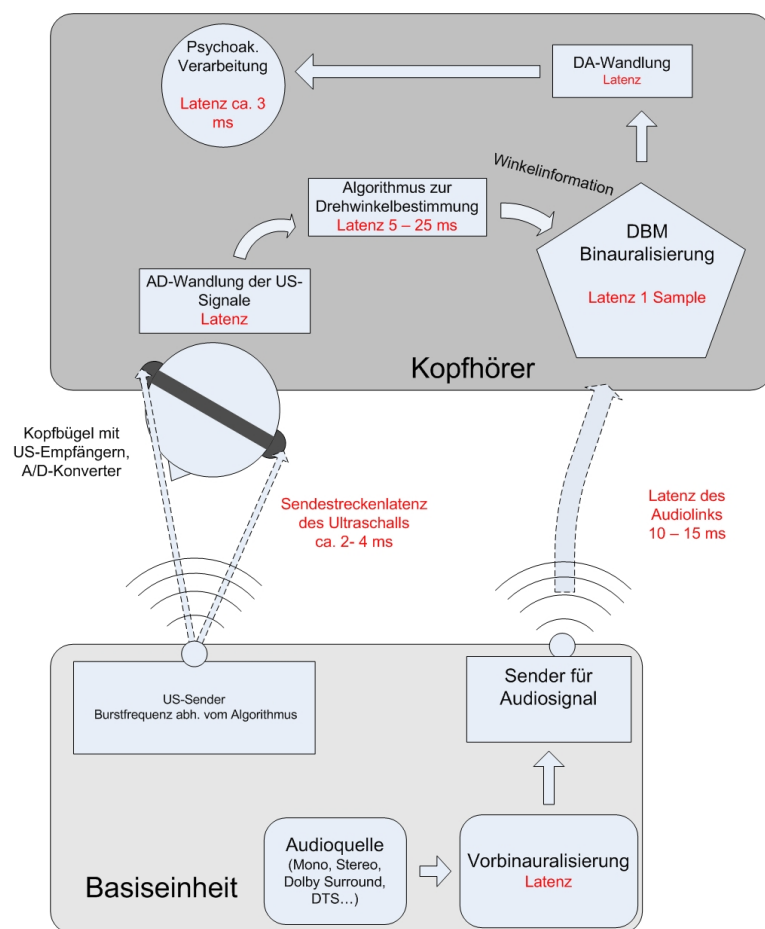


Abb. 6.6: System mit zwei Empfängern am Kopfhörer in Kombination mit dem DBM-Binauralmodul

- Die gesamte Signalverarbeitung (Verstärkung, Filterung, Winkeldetektion, Binauralisierung) findet in einem DSP am Kopfhörer statt. Dadurch wird der Akku des Kopfhörers stärker beansprucht. Für den Empfang des Audio-Funksignals ist am Kopfhörer bereits ein Signalprozessor vorhanden, der eventuell bei entsprechender Modi-

fikation für die Berechnungen des Headtrackingalgorithmus mitverwendet werden könnte. Der zusätzliche Materialaufwand wäre damit gering.

- Beim Betrieb mehrerer Headtracker bleibt die Zahl der Ultraschallsignale im Raum konstant gering¹. Eine eindeutige Signalerkennung am Empfänger wird so erleichtert. Außerdem ist das System im Nachhinein problemlos um mehr Kopfhörer erweiterbar.
- Die Gesamtlatenz ist im Vergleich zum erstgenannten Systemaufbau geringer, da das US-Signal zur Winkeldetektion zeitgleich mit dem Audiosignal zum Kopfhörer gefunkt wird. Es fällt also nur einmal die Latenz der Funkstrecke ins Gewicht.

6.4.1 Abschattung und Beugung von Schallwellen am Kopf

Wie in Kap.2.3.2 erklärt wird, können Schallwellen sich um Hindernisse herumbeugen, deren Ausmaße sich im Bereich der Wellenlänge der Schallwelle befinden. Für eine Ultraschallfrequenz von 40kHz beträgt die Wellenlänge ca 8 mm (Gl.6.2). Dies bedeutet, dass diese Frequenz praktisch an allen Hindernissen reflektiert oder gestreut wird. Objekte mit sehr kleinen Durchmesser (z.B. Haare am Kopf) stellen demnach zwar keine Hindernis dar, jedoch ist dabei der Ausbreitungsweg der Schallwelle nicht mehr geradlinig. Diese Nichtlinearität führt somit theoretisch zu einer längeren Laufzeit des Schalls und zu Messfehlern, da die Winkelbestimmung nach Kap.3 von einer ungehinderten geradlinigen Schallausbreitung ausgeht.

$$\lambda = \frac{c}{f} \quad (6.1)$$

$$8,625\text{mm} = \frac{345\text{m/s}}{40000\text{Hz}} \quad (6.2)$$

6.4.2 Rotationspunkt des Kopfes in Vergleich zum Modell

Die geometrischen Modelle aus Kap.3 gehen davon aus, dass Rotationspunkt des Kopfes, Ohren und Sensoren auf einer Achse liegen. (Vgl. Position E_1 und E_2 Abb.6.7) liegen die Sensoren fälschlicherweise auf einer anderen Linie, ergibt sich ein Berechnungsfehler. Der errechnete Winkel α' ist größer als der wahre Winkel α . Dieser Winkelfehler ist umso höher je größer der wahre Winkel α ist (siehe Gl.6.3). Einen großen Fehler kann eine Fehlposition der Sensoren bewirken. Abb. 6.8 zeigt den Fehler, bei $b = 3\text{ cm}$ und $r = 70\text{ cm}$.

$$\alpha' = \pi - \arccos \left(\frac{b - r \cdot \cos(\alpha)}{\sqrt{r^2 + b^2 - 2(b \cdot r \cdot \cos(\alpha))}} \right) \quad (6.3)$$

¹Single Server - Multi Client Betrieb ähnlich Abb.3.3

- b : Offset zwischen Sender zum Drehmittelpunkt
 r : Distanz von Drehmittelpunkt zur Kopfhörermitte
 α : wahrer Drehwinkel des Kopfes
 α' : berechneter Drehwinkel

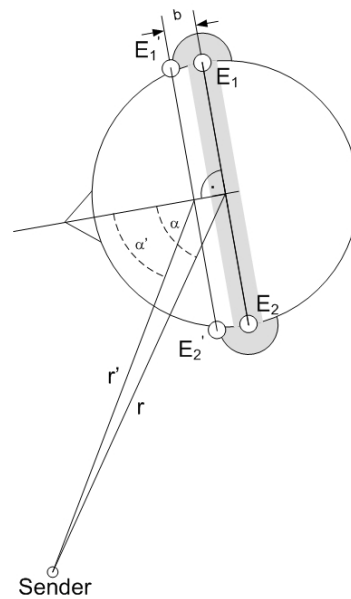


Abb. 6.7: Winkelfehler bei falscher Sensorposition; $b = 3\text{cm}$

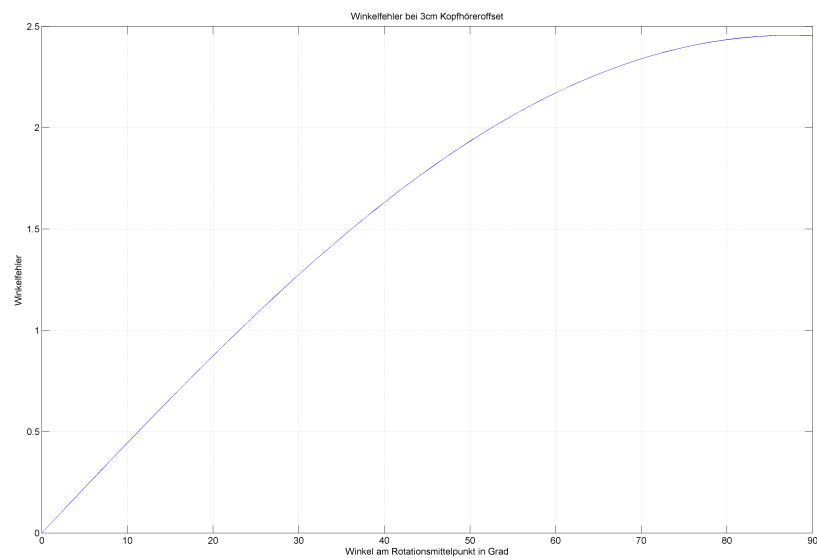


Abb. 6.8: Winkelfehler bei falscher Positionierung des Kopfhörers

6.4.3 Abstand der Sensoren am Kopf

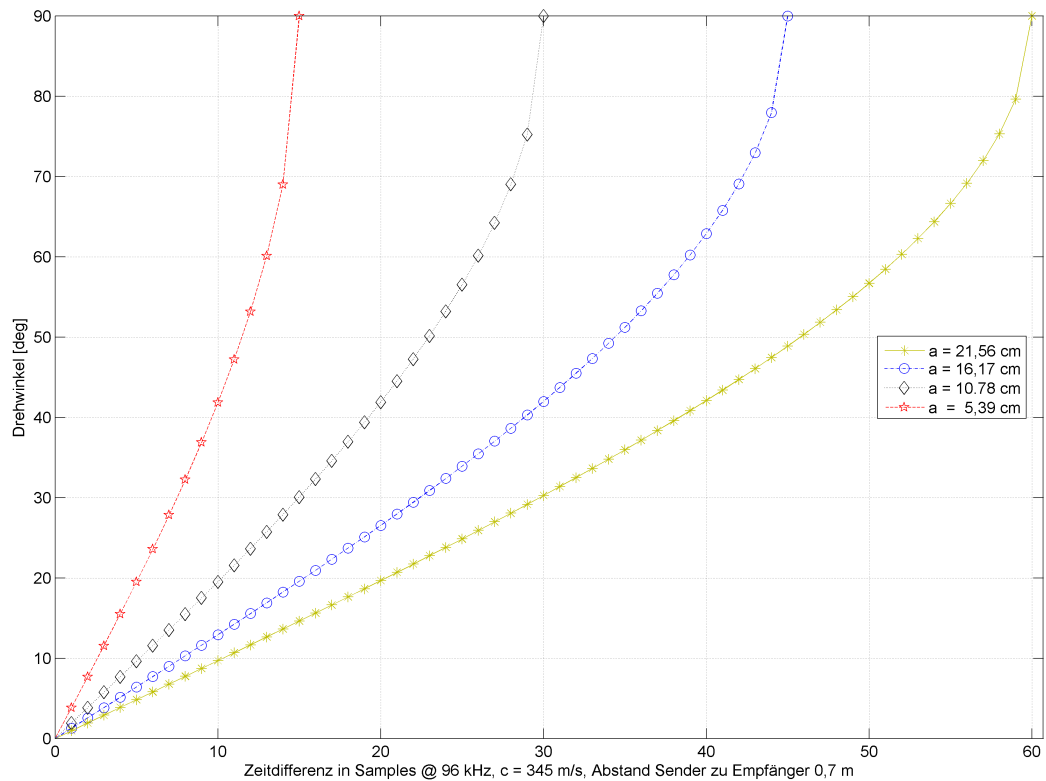


Abb. 6.9: Auflösungsgrenzen als Funktion der Aperturgröße

Die Aperturgröße, also der Abstand der Sensoren am Kopfhörer, besitzt in Relation zur Abtastfrequenz großen Einfluss auf das Auflösungsvermögen des Headtrackers. Eine kleine Aperturgröße hat zur Folge, dass auch eine entsprechend geringe maximale Laufzeitdifferenz detektiert werden kann. Der gleiche Messfehler bei der Bestimmung der relativen Laufzeitdifferenz hat bei kleiner Aperturgröße eine stärkere Auswirkung als bei einer größeren Apertur. Abb. 6.9 zeigt Laufzeitdifferenzen als Funktion der Aperturgröße bei einer Abtastrate von $f_s = 96\text{ kHz}$. Wird (ausgehend der Geometrie von Kap. 3.3.5) eine Laufzeit nur in ganzen Samples gemessen, kann nur eine Winkelauflösung entsprechend der Markierungen in den Graphen Abb. 6.9 erreicht werden. Diese Auflösung ist nichtlinear und verschlechtert sich mit zunehmendem Drehwinkel. Eine Möglichkeit zum Verbessern der Auflösung wäre die Erhöhung der Abtastfrequenz f_s , was jedoch auch mehr Rechenaufwand bedeuten würde. Einen besseren Ansatz stellt für viele Algorithmen eine Interpolation nach Kap. 5.3.1.2 dar, hierbei kann durch geringeren Mehraufwand die Auflösung gesteigert werden.

6.5 Maximale Pulsrate

Werden für die Detektierung gepulste Schwingungspakete¹ benutzt, darf eine maximale Pulsrate nicht überschritten werden ([WE97]). Ansonsten führt dies zu Uneindeutigkeiten bei der Laufzeitenbestimmung. Die maximale Pulsrate ist abhängig von der maximal zu erwartenden Wegdifferenz w_{max} :

$$f_{max} = \frac{c}{w_{max}} \quad (6.4)$$

f_{max} : maximale Pulsrate

w_{max} : maximale Entfernung

Bei einer maximalen Laufzeitdifferenz von ca 10 bis 20 cm² und einer Schallgeschwindigkeit von 345m/s ergibt sich so eine Pulsrate von ca. 1700 Hz. Dieser Wert scheint mit Blick auf die psychoakustischen Anforderungen³ mehr als ausreichend zu sein und ist praktisch in keinster Weise beeinträchtigend.

Ein anderer Faktor der die Pulsrate beschränkt ist die Raumakustik. Bei zu kurz aufeinander folgenden Pulsen können sich Raumreflexionen mit dem Direktschall später abgesendeter Pulse am Empfänger überlagern. Direktschall und Reflexion können so nicht mehr sauber unterschieden werden - die Signalinformation des interessierende Direktschalls wird unter Umständen nicht detektiert. Daher senden viele Ultraschall-Tracker ihre GW-Signale mit einer Senderate von 50 Hz aus ([BU03]). Der Effekt kann sogar dadurch verstärkt werden, dass bei einer extremen Kopfdrehung der Direktschall, aufgrund der Richtwirkung⁴ des Empfängers, bereits gedämpft ist als eine durch eine nahe Wand oder Glasscheibe hervorgerufene Schallreflexion (siehe Abb.6.10).

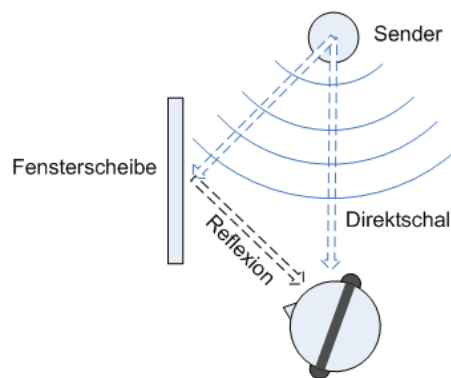


Abb. 6.10: Problematik bei Raumreflexionen und Richtcharakteristik der Sensoren

¹gated waves, siehe Kapitel 4 auf Seite 30

²im Fall der Headtrackerkonstruktion nach 3.3.5 ist $D_{max} \cdot c = a$ bei $\alpha = \pm 90^\circ$

³laut [WZ97] sind Updateraten von 60 Hz ausreichend; siehe Kap. 6.2.2 auf Seite 68

⁴Richtwirkung der verwendeten Sensoren siehe Anhang B

Kapitel 7

Umsetzung

7.1 Allgemein

Im folgenden Kapitel werden die Detektionsmethoden aus Kap.5 in die Praxis umgesetzt und miteinander verglichen. Der Systemaufbau ist in Abb.7.1 zu sehen. Die Empfänger sind am Kopfhörer positioniert, der Sender an einem Referenzpunkt. Nach den Vorüberlegungen aus Kap.3.3.5 und Kap.6.4 erscheint diese Anordnung als beste Lösung, denn Bauteil- und Rechenaufwand steht in einem guten Verhältnis zur möglichen Detektionsgenauigkeit. Genauer zu den verwendeten Sensoren findet sich im Anhang. Bei der Umrechnung von Laufzeit- in Wegdifferenzen ist die aktuelle Lufttemperatur nach Gl.2.6 berücksichtigt. Die Audio-Echtzeitsoftware *Pure Data* (PD) stellt eine gute Entwicklungsumgebung zur Programmierung und Evaluierung der einzelnen Algorithmen dar. Somit konnte ein lauffähiges Binauralsystem geschaffen werden, welches auch in einem Funkkopfhörer eingebunden werden kann¹. Die Binauralisierung erfolgt durch das am IEM entwickelte *DBM*-Binauralmodul ebenfalls unter PD. Das *DBM*-System bietet sich gerade für diese Anwendung an, da es mit 1,1 MIPS (*Million instructions per second*) auskommt und dadurch leicht auf einem DSP am Kopfhörer implementierbar ist (siehe Abb.6.6).

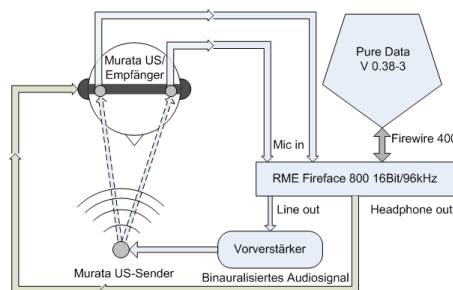


Abb. 7.1: Systemaufbau zur Umsetzung

¹In diesem Entwicklungsstadium ist ein drahtgebundener Kopfhörer (AKG K501) in Verwendung.

7.1.1 Messaufbau

Um die Algorithmen objektiv und realistisch vergleichen zu können, wurde bei der Evaluierung des Systems ein Styroporkopf auf einem Drehstativ mit Schrittmotor fixiert. Zu Vergleichszwecken wurde zusätzlich als Referenzsystem der Positionstracker „*Flock of Birds*¹“ auf den Mittelpunkt dieses Drehstatives montiert. Aufgrund der unterschiedlichen Funktionsprinzipien kann eine Beeinflussung von Referenzsystem und Ultraschallheadtracker ausgeschlossen werden. Der Styroporkopf besitzt außerdem eine Perücke, damit der reflektierende bzw. beugende Einfluss von Haaren in diesem Frequenzbereich auch berücksichtigt wird (siehe Abb. 7.2(a)). Gemessen wurde in einer Distanz von 70cm zwischen Sender und Drehstativ. Das entspricht in etwa dem Kopfabstand einer sitzenden Person auf einer Fahrzeugrückbank zur Kopfstütze des Vordermannes. Der Abstand der Sender am Kopfhörer beträgt 15,85cm. Weiterhin wurde auch mit unterschiedlichen Höhenunterschieden zwischen Sender und Empfänger experimentiert. Dabei ließ sich feststellen, dass aufgrund von Abschattungen und Reflexionen der Sendesignale der Headtracker besser funktioniert, wenn sich der Sender nicht auf gleicher Höhe mit den Empfängern befindet, sondern etwas höher positioniert wird. Auch wenn dadurch Sender und Empfänger nicht mehr in einer Ebene liegen und Gl. 3.22 zu Berechnungsfehler führt (siehe Abb. 7.3), wird eine vertrauenswürdige Detektion auch bei größeren Winkeln erst ermöglicht. Für die dargestellten Messreihen befindet sich der Sender 10cm über den Empfängern. Ferner ist zu bemerken, dass sich die Ultraschallempfänger ca 37mm vor dem Rotationsmittelpunkt des Styroporkopfes befanden². Die Messungen wurden bei einer Temperatur von 28°C und



(a) Kopfhörer mit Empfängern und Referenzsensor (mitte)



(b) Messanordnung: links Kopfhörer, rechts Ultraschallsender und Sendereinheit des *Flock of Birds*

Abb. 7.2: Aufbau zur Evaluierung der Algorithmen

¹ Arbeitsprinzip eines elektromagnetischen Trackingsystems siehe Kap. 1.2.2 auf Seite 4; Technische Daten des *Flock of Birds* siehe Anhang C auf Seite 120

einer Luftfeuchtigkeit von 70% durchgeführt. Am Schrittmotor muss bemängelt werden, dass die Motormechanik eine Winkelunsicherheit von ca 3° zuließ. Darüber hinaus bestand der Boden des Messraumes aus Metallplatten, die das Referenzsystem stören¹.

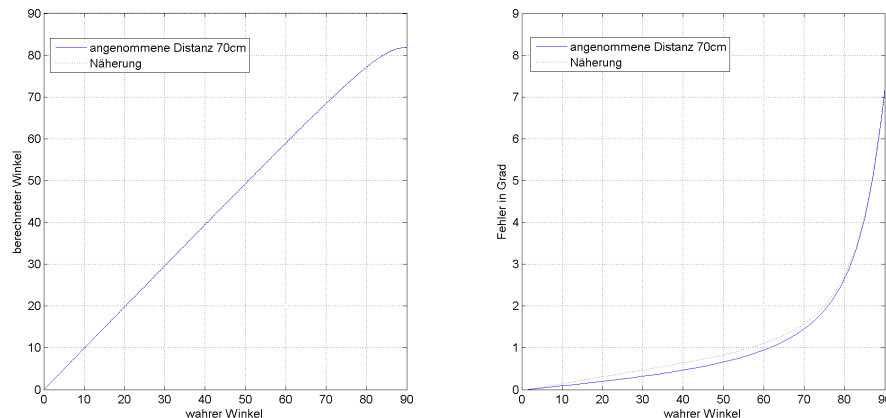


Abb. 7.3: Winkelfehler durch 10cm Höhenunterschied zwischen Sender und Empfänger

7.2 Implementierte Algorithmen

Nicht alle in Kap.5 vorgestellten Ansätze sind auch implementiert worden, da einige Methode der Literatur nach nicht exakt genug arbeiten (z.B. einfache Schwellwertdetektionen) oder der hohe Rechenaufwand die Kapazitäten des DSPs am Kopfhörer übersteigen würde (z.B. MXS oder Detektion der Dopplerfrequenz). Die unter *PD* implementierten und hier zu vergleichenden Algorithmen sind:

- *Filter*: Schwellwertdetektion über die Signalhüllkurve
- *Diode*: Analoge Schwellwertdetektion über die Signalhüllkurve (Simulation in *PD*)
- *ZeitCor*: Korrelation im Zeitbereich
- *PhaseCor*: Korrelation im Frequenzbereich
- *ASDF*: Average Square Difference Function
- *AMDF*: Average Magnitude Difference Function
- *BitCor*: 1-Bit-Korrelation

²Auch dieser Offset führt zu Umrechnungsfehlern von Laufzeiten in Winkelwerte. siehe dazu Kap.6.4.2

¹Die elektromagnetischen Felder des Sensors vom *Flock of Birds* verursachen in Metallen Wirbelfelder, worunter die Genauigkeit des Referenzsystems leidet.

Als Sendesignale s werden 40kHz Bursts signale mit einer Dauer von 2ms verwendet, die mit einem Hanningfenster belegt worden sind. Diese Fades verringern die effektive Signallänge zwar auf ca. 1ms , jedoch wird die mechanische Beanspruchung des Senders durch das weichen Ein- und Ausschaltvorgänge minimiert. Dadurch werden hörbare Nebenfrequenzen vermieden (Vgl. Entwicklungsziele Kap.1.3 sowie gesundheitliche Aspekte Kap.2.4). Eine Ausnahme stellt die 1-Bit-Korrelation dar. Hierbei ist ein logarithmischer Chirp in Verwendung.

7.2.1 Reduzierung der Taktrate durch Einseitenbanddemodulation

Das Nyquist-Shannon-Abtasttheorems¹ führt dazu, dass 40kHz -Ultraschallsignale mit einer Samplingfrequenz $f_s > 2 \cdot 40\text{kHz}$ abgetastet werden müssen. In den Versuchen wurde mit $f_s = 96\text{kHz}$ gearbeitet. Die hohe Abtastfrequenz hat den Nachteil, dass dementsprechend viele Samples vorliegen, welche vom gewählten Detektionsalgorithmus auch alle verarbeiten werden müssen. Dies erfordert letztendlich einen teuren, hochtaktigen und leistungsstarken DSP. Da die Signalverarbeitung am Kopf stattfinden wird, lässt jedoch der Wunsch nach einer langen Betriebsdauer des Kopfhörers keinen DSP mit hohem Leistungsverbrauch zu.

Da aufgrund der Sensorcharakteristik² nur eine sehr geringe Bandbreite genutzt werden kann, besteht die Möglichkeit das empfangene Signal ohne Informationsverlust durch eine Einseitenbanddemodulation in das Basisband zu bewegen und mit einer geringeren Frequenz abzutasten. Dafür wird das empfangene 40kHz Signal $x(t)$ mit dem Spektrum $X(\omega)$ mit einem Signal der Form $\cos(\omega_0 t)$ mit $\omega_0 = 40\text{kHz}$ multipliziert (siehe Gl.7.2). Das Spektrum ergibt sich nach Gl.7.3. Es enthält 2 Versionen des Ursprungssignals: eine herauf- und eine heruntermodulierte Version von X . Mittels eines Tiefpassfilters kann das heruntermodulierte Signal $X(\omega - \omega_0)$ isoliert werden. Es befindet sich in einem Bereich um 0 Herz herum und repräsentiert die Einhüllende von x . (Die Frequenz ω_0 bestimmt somit die Schrittweite der Modulation.)

$$x_{\text{mix}}(t) = x(t) \cdot (\omega_0 t) \quad (7.1)$$

$$= \frac{1}{2} [x(t) e^{j\omega_0 t} + x(t) e^{-j\omega_0 t}] \quad (7.2)$$

$$X_{\text{mix}}(\omega) = \frac{1}{2} [X(\omega - \omega_0) + X(\omega + \omega_0)] \quad (7.3)$$

Das neue Signal könnte nun mit einer sehr geringen Frequenz abgetastet werden (z.B. 500Hz) Durch die Verwendung von Bursts signalen bewirkt die Form und Länge der Einhüllenden jedoch eine Verbreiterung des Spektrums. Das heruntermodulierte Signal kann

¹Das Abtasttheorem besagt, dass ein kontinuierliches, bandbeschränktes Signal mit einer Minimalfrequenz von 0Hz und einer Maximalfrequenz f_{max} mit einer Frequenz größer als $2 \cdot f_{\text{max}}$ abgetastet werden muss, damit man aus dem so erhaltenen zeitdiskreten Signal das Ursprungssignal ohne Informationsverlust approximieren kann.

²Die Sensorstrecke besitzt eine 3dB -Bandbreite von ca. 65Cent , siehe Frequenzgang im Anhang B auf Seite 118

somit mit 3 oder 6 kHz abgetastet werden (Downsampling), womit mit einer 16 bzw. 32 fach geringere Datenmenge Laufzeitdifferenzen ermittelt werden können. Die maximal erreichbare Detektionsauflösung verringert sich jedoch auch um den gleichen Faktor. Daher ist eine Interpolation des Ergebnisses nach Kap. 5.3.1.2 notwendig. Das fraktionale Delay $\hat{D}_{frac}$ wird noch zusätzlich mit dem Downsamplingfaktor multipliziert, um wieder im ursprünglichen 96kHz-Samplingbereich zu sein. Aufgrund der Tiefpassfilterung entsteht eine Verzögerung, welche die *end-to-end Latenz*¹ erhöht (Vgl. Abb. 7.5).

Anstatt der oszillierenden Sinusfunktion $\cos(\omega_0 t)$ kann das Signal auch mit sich selbst multipliziert werden. Dabei wird das Modulationsergebnis immer auf 0 Hz herum fallen. Bei einigen Algorithmen wird diese Methode angewendet, wobei die Einseitenbanddemodulation auf digitaler Ebene ausgeführt wurde (siehe Abb. 7.4): Nach dem Digitalisieren der Eingangssignale mit 96 kHz wird die Einseitenbanddemodulation ausgeführt. Dann erst beginnt die eigentliche Signalverarbeitung. Möchte man die Modulation noch vor der A/D-Wandlung vollziehen, treten die klassischen analogen Probleme einer Multiplizierschaltung auf (Bauteilabhängigkeit und Alterung, Temperaturinstabilität...). Eine Schaltung eines analoger Vierquadrantenmultiplizierers findet sich in [TS99].



Abb. 7.4: Schema der Einseitenbanddemodulation

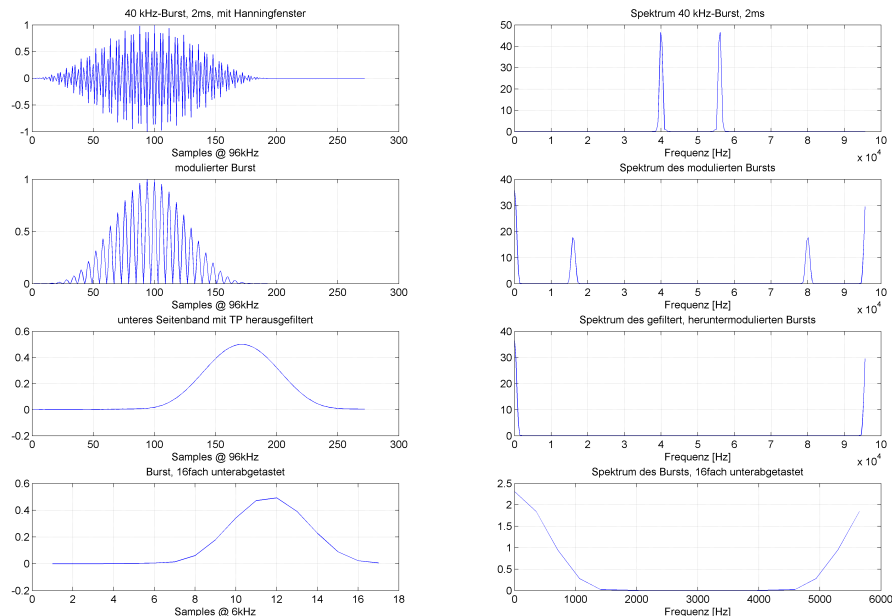


Abb. 7.5: Signal während der ESB-Demodulation im Zeit- und Frequenzbereich

¹Zu den Auswirkungen von Latenzen auf den Höreindruck siehe Kap. 6.2.2

7.2.2 Schwellwertdetektion der Hüllkurve - Filter

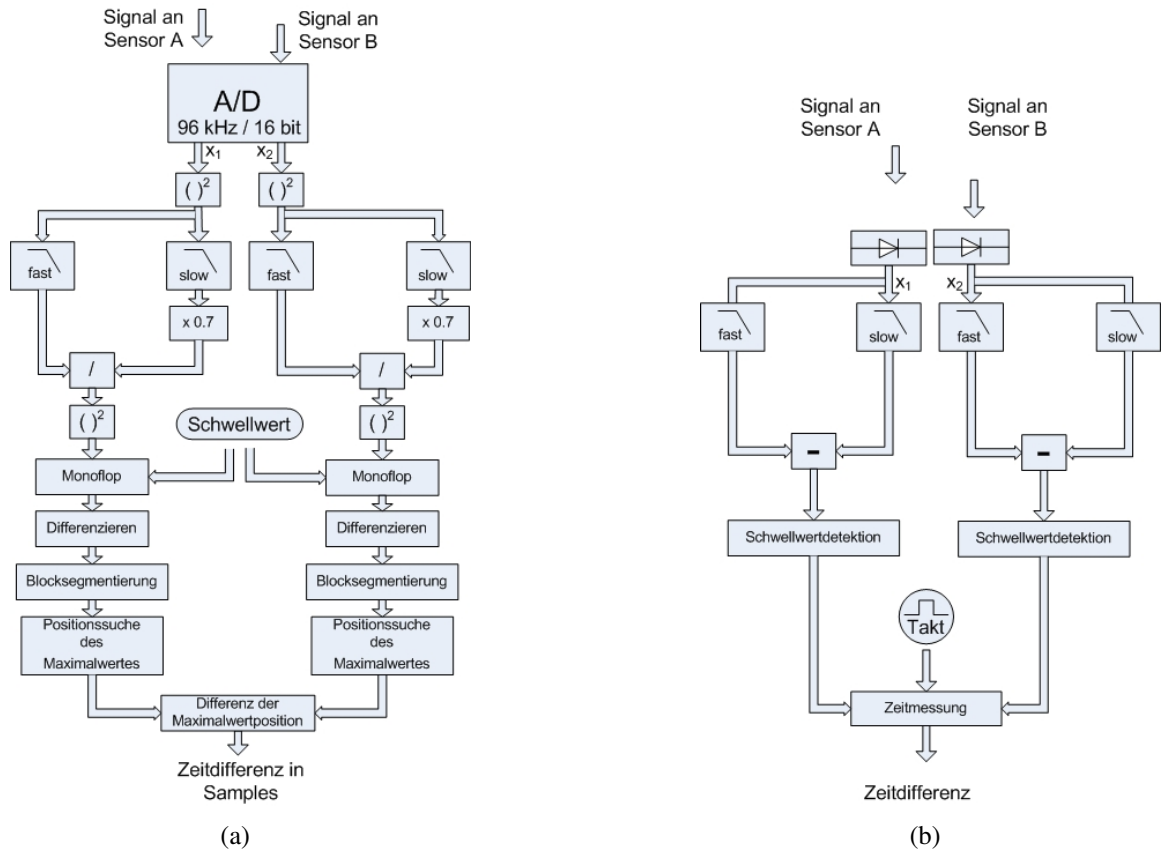


Abb. 7.6: Strukturbild zur Detektion der Hüllkurven mittels Schwellwert; digitale (a) bzw. analoge (b) Umsetzung

Der Algorithmus nach Abb.7.6(a) ist sehr einfach in einem DSP zu implementieren, da einige der verwendeten Operationen in vielen DSPs hardwareseitig integriert sind. Zuerst wird das empfangene Signal durch Quadrieren gleichgerichtet und von 2 verschiedenen Tiefpässen gefiltert. Das Quadrat des Quotient dieser Filtersignale ergibt einen geschärften¹ Hüllkurvenverlauf. In Abb.7.6 sind diese Tiefpassfilter mit *slow* und *fast* bezeichnet und werden durch folgende Differenzengleichungen charakterisiert:

$$y_{[n]}^{\text{fast}} = 0.95 \cdot y_{[n-1]} + 0.05 \cdot x_{[n]} \quad (7.4)$$

$$y_{[n]}^{\text{slow}} = 0.9999 \cdot y_{[n-1]} + 0.0001 \cdot x_{[n]} \quad (7.5)$$

Die große Zeitkonstante des langsamen Tiefpasses ergibt eine Schätzung über die Energie des Grundrauschens. Der schnelle Tiefpass reagiert auf schnelle Pegelschwankungen im Empfangssignal. Die Zeitkonstante von *fast* ist entsprechend der Länge des verwendeten

¹ Alternativ könnten hier auch andere nichtlineare Funktionen (z.B. $e^{(x)}$) verwendet werden.

Signals ($2ms$) und der Burstfrequenz ($f_B = 187Hz$) dimensioniert. Schnellere Pegeländerungen innerhalb des Empfangenen Signals (z.B. von kurzzeitigen Störungen anderer Ultraschallquellen) sollen so unterdrückt werden. Das Verhältnis der beiden Filtersignale ergibt somit ein stabiles Hüllkurvensignal, welches sich dem Pegel des Grundrauschens anpasst.

Dieses Hüllkurvensignal wird anschließend von einem Monoflop in ein binäres Signal konvertiert. Eine ansteigende Flanke des *High*-Zustandes repräsentiert dabei das Eintreffen des Burstes, wobei der Auslösezeitpunkt von einem festgelegtem Schwellwert bestimmt wird. In *PD* wird aus praktischen Gründen innerhalb von Blocksegmenten auf die ansteigende Flanke des *High*-Zustandes getriggert. Die Blockgröße N ist umgekehrt proportional zur Burstfrequenz f_B nach Gl.7.6. Die Zeit, die zwischen dem Auslösen der Trigger von den beiden Empfangssignalen x_1 und x_2 verstreicht, ergibt die Laufzeitdifferenz $\hat{D}$.

Abb.7.7 zeigt die Winkeldetektion des Algorithmus im Vergleich zum Referenzsystem. Die Genauigkeit des Algorithmus ist vor allem von der Taktrate des Systems abhängig, denn durch die zeitliche Diskretisierung kann der Komparator nur mit einer zeitlichen Genauigkeit von T_s den Trigger auslösen. Eine Erhöhung der Taktrate verbessert also einerseits die Genauigkeit, andererseits steigt auch die Anzahl an Berechnungen pro Zeiteinheit an.¹ Um diese groben Winkelsprünge zu vermeiden, sollte ein moderater Tiefpass der Detektion nachgeschaltet werden. Siehe dazu Kap.7.4.

Dieses Algorithmus ist prinzipiell auch rein analog realisierbar: Da eine Multiplikation im Analogen nur recht aufwendig erzeugt werden kann, ersetzt eine Diodenschaltung zum Gleichrichten der Signale die Quadrierer. Zum Anderen wird statt dem Verhältnis der Filterausgänge die Differenz dieser Signale gebildet (siehe Abb.7.6(b)).

¹ Alternative: Eine Vergrößerung des Abstandes der Sensoren am Kopf kann bei gleichbleibender Abtastrate die Genauigkeit steigern (siehe Abb.6.9).

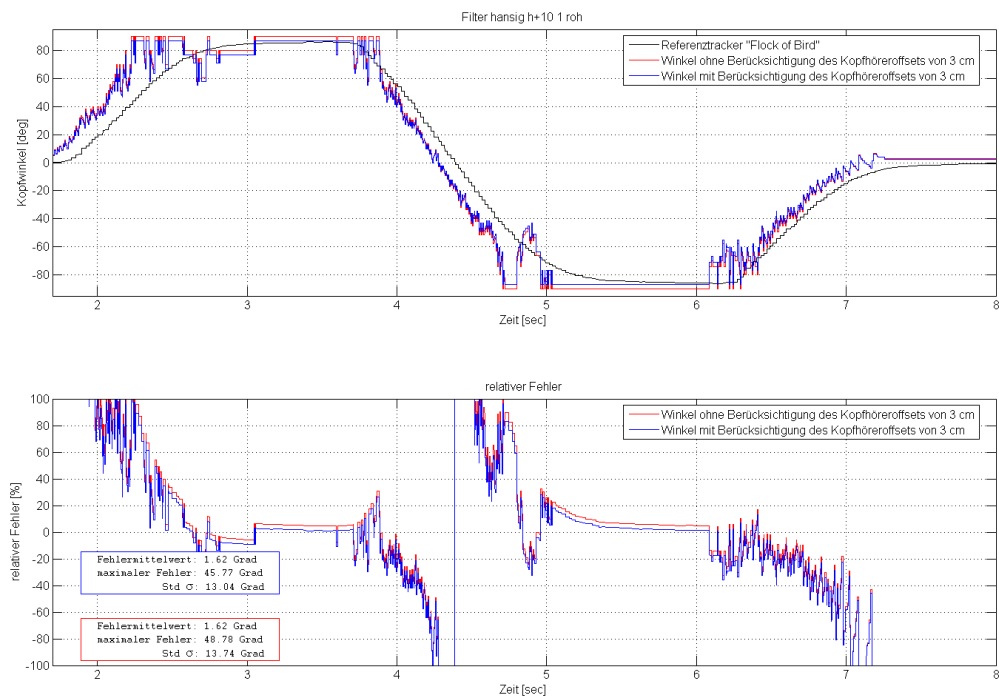


Abb. 7.7: Auswertung der Messung mit Kunstkopf auf Schrittmotor für Hüllkurvendetektion nach Abb.7.6(a)

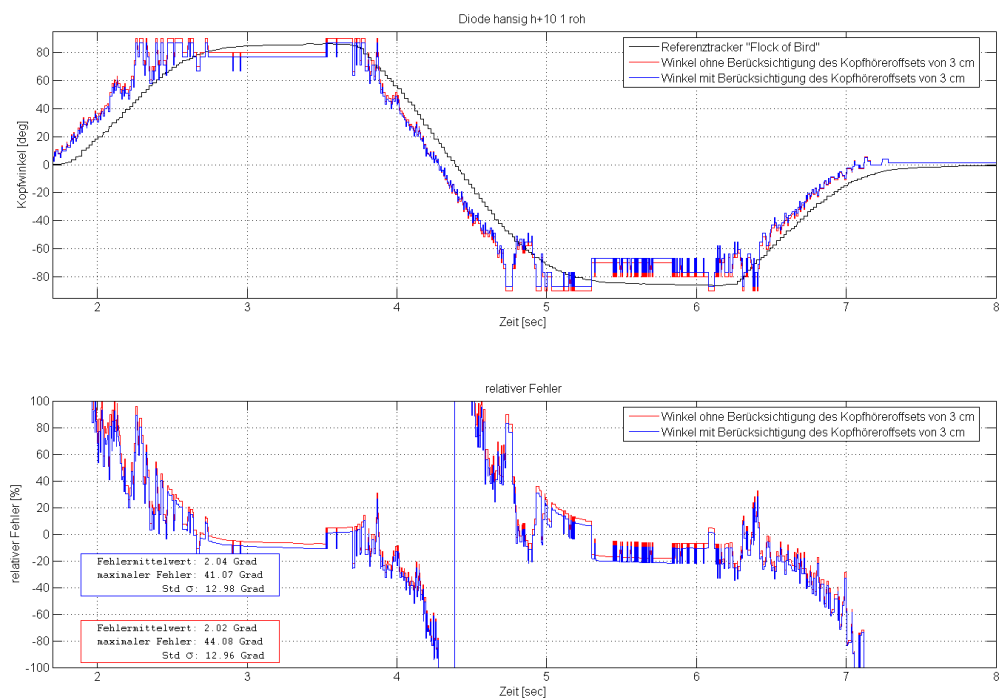


Abb. 7.8: Auswertung der Messung mit Kunstkopf auf Schrittmotor für Hüllkurvendetektion nach Abb.7.6(b)

7.2.3 1-Bit-Korrelation - BitCor

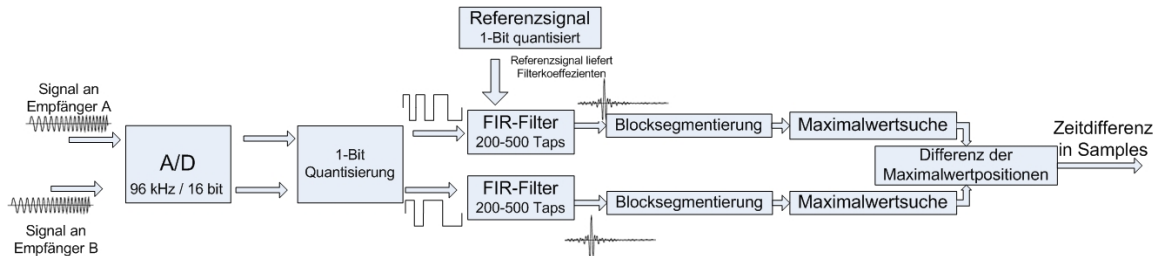


Abb. 7.9: Strukturbild *BitCor*

Als Sendesignal wird bei der 1-Bit-Korrelation ein Chirp genutzt, der über einer Dauer von 5ms von 39kHz auf 41kHz logarithmisch ansteigt. Um das Korrelationsergebnis in den FIR-Filtern zu verbessern, ist das Referenzsignal für die FIR-Filter eine 1-Bit quantisierte Version des Nutzsignals s , welches zuvor mit der Übertragungsfunktion der Sendestrecke bei einem Azimutwinkel von 0° gefaltet wurde. Die Übertragungsfunktion wurde im Messraum nach der *swept-sine technique* gemessen ([FA00]). Das gesamte Referenzsignal hat eine Länge von 510 Samples. Die Performance einer geringeren FIR-Filterlänge wurde mit Signalausschnitten mit einer Länge von 200 und 300 Samples untersucht.

Aus praktischen Gründen wird in *PD* durch eine Blocksegmentierung innerhalb eines Signalausschnittes die zeitliche Position der Maximalwerte beider Filterausgänge ermittelt. Dabei ist die Signalblockgröße N abhängig von der Frequenz, wie oft ein Chirp ausgesendet wird (f_B) und beträgt nach Gl.7.6 bei einer Sendefrequenz $f_B = 93.75\text{Hz}$ und einer Abtastrate von $f_s = 96\text{kHz}$ 1024 Samples.

$$N = \frac{f_s}{f_B} \quad (7.6)$$

Die Laufzeitdifferenz berechnet sich dann aus der Differenz der Maximalwertpositionen der Signalblöcke (siehe Abb.7.9).

Wie man in Abb. 7.10 sehen kann, detektiert dieser Algorithmus nicht gut genug. Das könnte neben der in Kap. 7.2.2 schon angesprochenen zeitlichen Diskretisierung folgende Ursachen haben:

Aufgrund der Richtwirkung der Sensoren verändert sich auch das Übertragungsverhalten bei einer Verdrehung der Sensoren zueinander¹. Aus diesem Grund stimmt das empfangene Signal nur noch bedingt mit dem vom Empfänger in Form des Referenzsignals erwarteten Chirp überein. Die 1-Bit-Korrelation wird dadurch gestört. Die Koeffizienten der FIR-Filter müsste abhängig vom aktuellen Kopfwinkel angepasst werden. Vom Kopfwinkel abhängige Referenzsignale müssen dafür zusätzlich am Empfänger in einem Festspeicher vorliegen und ständig in die FIR-Filter geladen werden.

Einen Einfluss hat wahrscheinlich auch der Dopplereffekt². Bei einer Kopfdrehung bewegt sich ein Empfänger zu und der andere hinweg vom Sender. Der Dopplereffekt bewirkt, dass der versendete Chirp gestaucht bzw. gestreckt an den Empfängern eintrifft. Dadurch verschieben sich auch die Frequenzverläufe und die Momentanfrequenz innerhalb des Chirps entgegengesetzt zueinander. Da die FIR-Filter die Polarität der Momentanfrequenz mit dem Referenzsignal vergleichen, werden die Chirps somit auf einer Empfängerseite verfrüht, auf der anderen Seite verspätet detektiert, was zu einer längeren Laufzeitdifferenz $\hat{D}$ und somit zu größeren Kopfwinkeln führt.

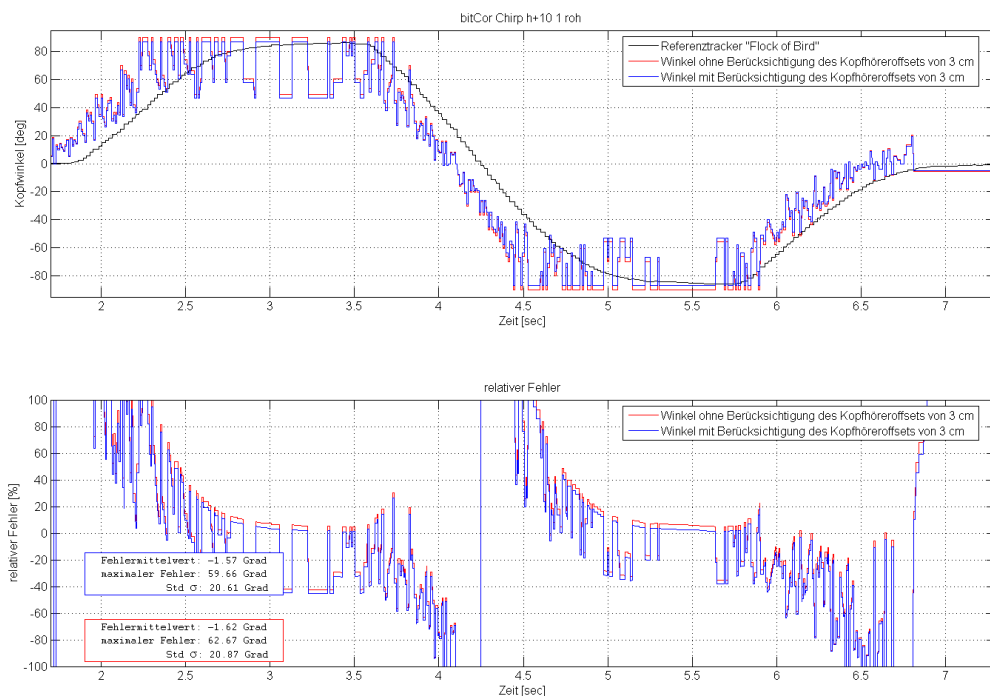


Abb. 7.10: Auswertung der Messung mit Kunstkopf auf Schrittmotor, FIR-Filterlänge 510 Taps

¹Siehe Polardiagramm Anhang B.3 auf Seite 119.

²Siehe Kap. 2.3.3 auf Seite 18.

7.2.4 Korrelation im Zeitbereich - ZeitCor

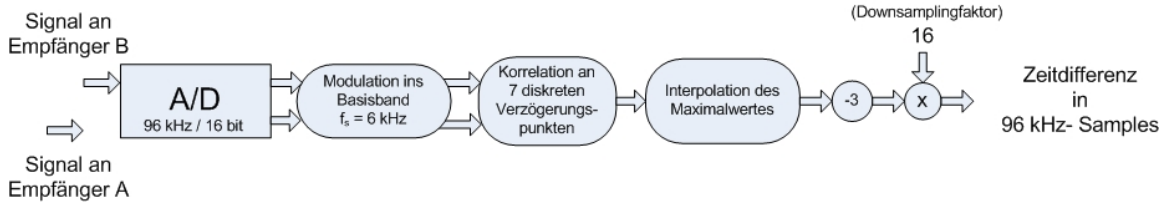


Abb. 7.11: Strukturbild *ZeitCor*

Wie in Abb.7.11 zu sehen, ist der Algorithmus ausgehend von Gl.5.10 implementiert worden. Um den Rechenaufwand gering zu halten werden die empfangenen Signale (wie in Kap.7.2.1 beschrieben) ins Basisband verschoben und mit 6 kHz abgetastet. Durch die Unterabtastung wird neben der zeitlichen Auflösung auch die geometrischen Dimensionen verändert: Angenommen, die Empfänger am Kopfhörer haben einen Abstand von $a = 16\text{cm}$, beträgt die maximale Laufzeitdifferenz¹ in Samples ausgedrückt:

$$D_{\text{Samples}} = \frac{a \cdot f_s}{c} \quad (7.7)$$

$$44,5 \text{ Samples} \approx \frac{0.16 [\text{m}] \cdot 96000 [1/\text{sec}]}{345 [\text{m/sec}]} \quad (7.8)$$

Eine Reduzierung der Abtastfrequenz f_s um den Faktor 16 wird somit auch die maximale Samedifferenz D_{Samples} um den gleichen Faktor reduzieren:

$$2,8 \text{ Samples} \approx \frac{0.16 [\text{m}] \cdot 6000 [1/\text{sec}]}{345 [\text{m/sec}]} \quad (7.9)$$

Die maximale Samedifferenz ist also abhängig von der Aperturgröße a und liegt bei $f_s = 6\text{kHz}$ zwischen $-3 \leq D_{\text{Samples}} \leq 3$ für einen Drehwinkel $-90 \leq \alpha \leq 90$. Das Nutzsignal s das in den Empfangssignalen x_1 und x_2 (nach Gl.5.5 und Gl.5.10) enthalten ist kann also auch nur innerhalb dieses Zeitbereiches verzögert sein. In Gl.7.10 drückt sich diese Vereinfachung durch die Begrenzung der Laufvariablen m aus. Um für das System die Kausalität herzustellen, muss in der Korrelation Signal x_1 verzögert und die Laufvariable m entsprechend angepasst werden (siehe Gl.7.11).

$$R_{x_1, x_2}[m] = \sum_{k=-\infty}^{\infty} x_1[k] \cdot x_2[k+m] \quad (7.10)$$

$$\begin{aligned} -3 \leq m \leq 3 \\ R_{x_1, x_2}[m-3] &= \sum_{k=-\infty}^{\infty} x_1[k-3] \cdot x_2[k+m] \\ 0 \leq m \leq 6 \end{aligned} \quad (7.11)$$

¹Die maximale Laufzeitdifferenz tritt bei einem Kopfwinkel von 90° auf; siehe Kap. 6.5 auf Seite 77

Der komplette Korrelationsalgorithmus ist in Abb.7.14 zu sehen. Anstatt der Summation über k Samples nach Gl.7.11 wird an den interessierenden sieben zeitdiskreten Punkten das Produkt der Signale x_1 und x_2 tiefpassgefiltert. Diese Tiefpässe haben eine lange Zeitkonstante (ca. 50 ms), da sie als „leckende“ Integratoren wirken und über einen längeren Zeitraum die Korrelationsergebnisse aufsummieren sollen¹. Eine Maximalwertsuche über die Ausgangswerte dieser sieben Tiefpässe führt zur zeitlichen Position der maximalen Korrelation. Eine anschließende Interpolation nach Kap.5.3.1.2 ergibt die geschätzten Laufzeitdifferenz $\hat{D}_{frac}$. Obwohl nach Gl.7.9 nur maximal 6 Zeitpunkte für eine Laufzeitdifferenz möglich sind, wird die Korrelation an 8 diskreten Verzögerungspunkten bestimmt. Dieser zusätzliche Rechenaufwand ist notwendig, damit auch für die Zeitpunkte x_2 und x_6 (Vgl. Abb.7.14) beide Nachbarsamples für eine parabolische Interpolation vorliegen. Die für die Kausalität benötigte Zeitverzögerung wird mit der Subtraktion von 3 Samples wieder ausgeglichen. Mit einer abschließenden Multiplikation mit dem Downsamplingfaktor erreicht man die fraktionale Samedifferenz im 96kHz Zeitbereich. Simulationen unter

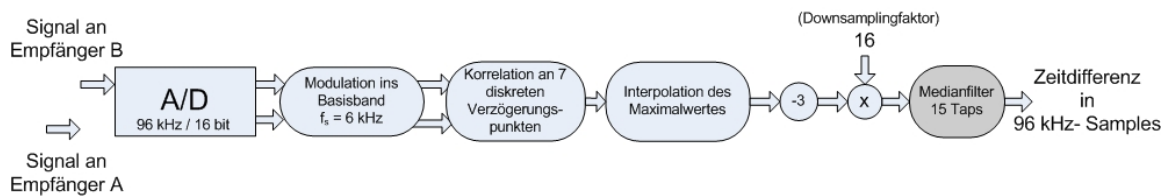


Abb. 7.12: ZeitCor mit Medianfilter

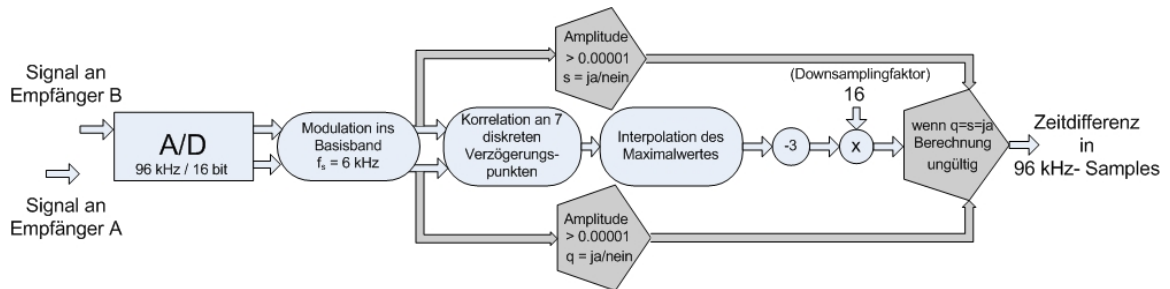


Abb. 7.13: ZeitCor mit Schwellwertabfrage

MATLAB haben gezeigt, dass das Korrelationsergebnis periodische Fehler aufweist (siehe Abb.7.15). Diese Fehler treten immer zu den Zeitpunkten auf, wenn an beiden Empfängern ein Burst eintrifft. Dies hat folgenden Grund: Durch die Hanningfensterung erhält der Burst eine nichtlineare Hüllkurve, welche während der Korrelationsberechnung den Zustand der Tiefpässe kurzzeitig verändert und somit während der Maximalwertinterpolation eine Fehldetektion verursacht. Ein Medianfilter (z.B. mit 15 Taps) kann dieses peaks glätten und eventuell gleichzeitig andere kurzzeitige Fehler unterdrücken (siehe Abb.7.12).

¹ Maximale Laufzeitdifferenz bei 3 Samples und $f_s = 6\text{kHz}$: $D \pm 0.5\text{ms}$. Abhängig von der Signalfrequenz müssen jedoch die Signale über ein entsprechend längeres Zeitfenster korreliert werden. Siehe Kap.5.3.1.

Eine andere Lösung des Problems besteht, wenn nur die Korrelationsergebnisse zwischen den Zeitpunkten, bei denen an beiden Empfängern der Burst eintrifft ausgegeben werden. Dies kann eine Schwellwertabfrage an den Empfängern lösen und ist im Vergleich zur Medianfilterung sehr einfach auf einem DSP zu implementieren (siehe dazu Abb.7.13).

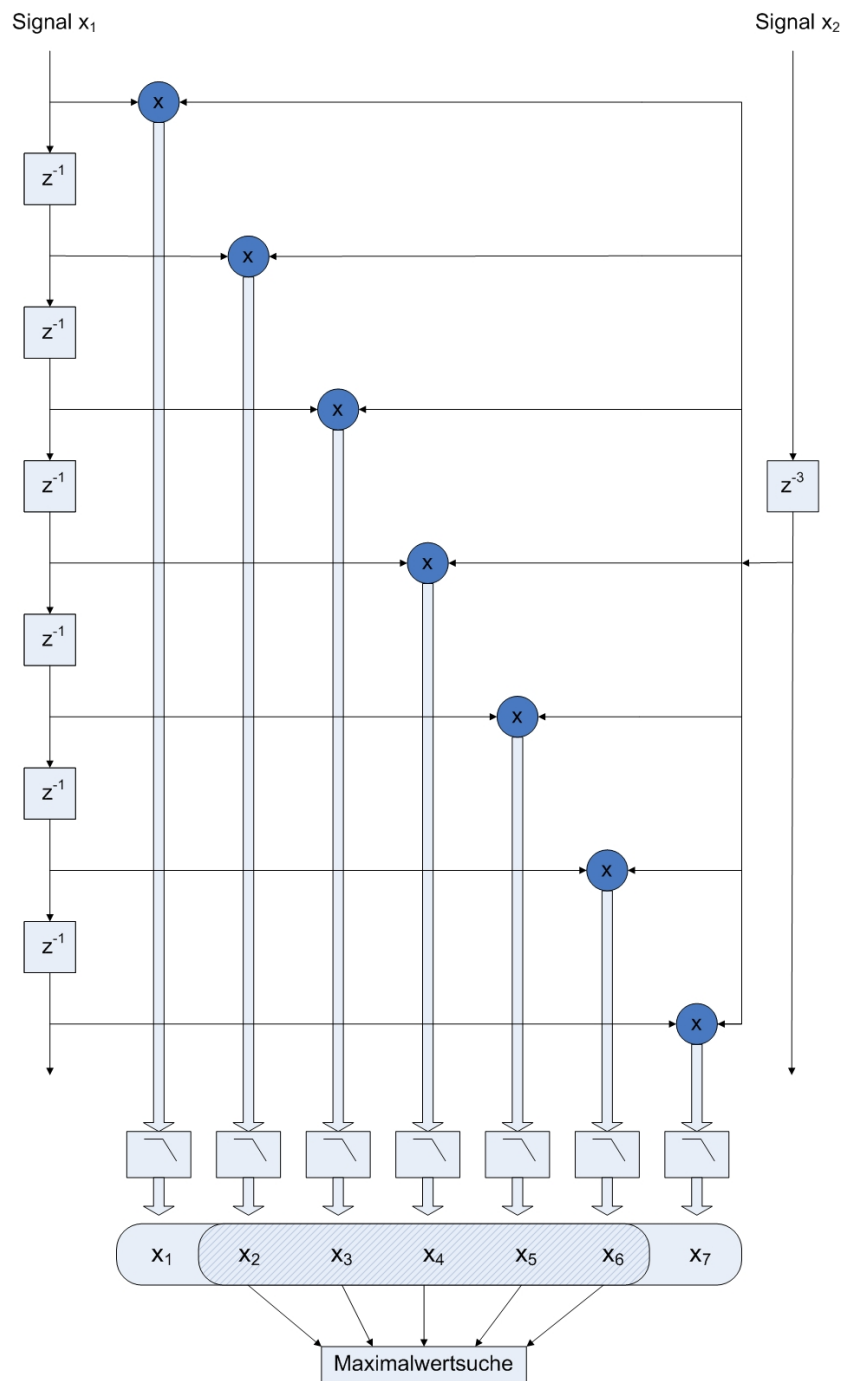


Abb. 7.14: Korrelation an sieben diskreten Verzögerungspunkten

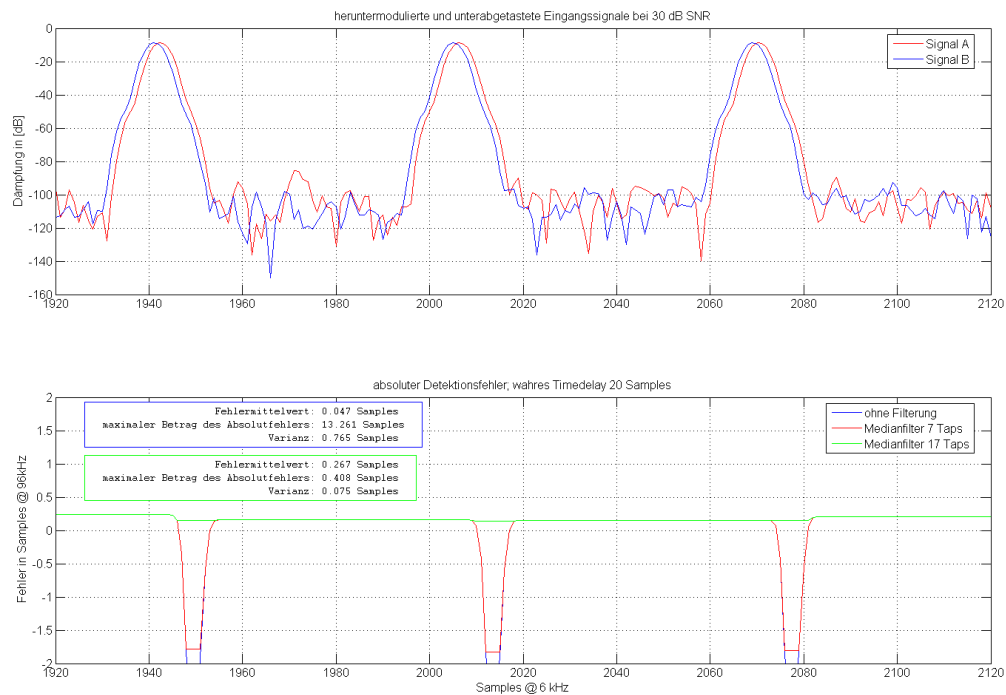


Abb. 7.15: periodische Fehler aufgrund der nichtlinearen Sendesignalform

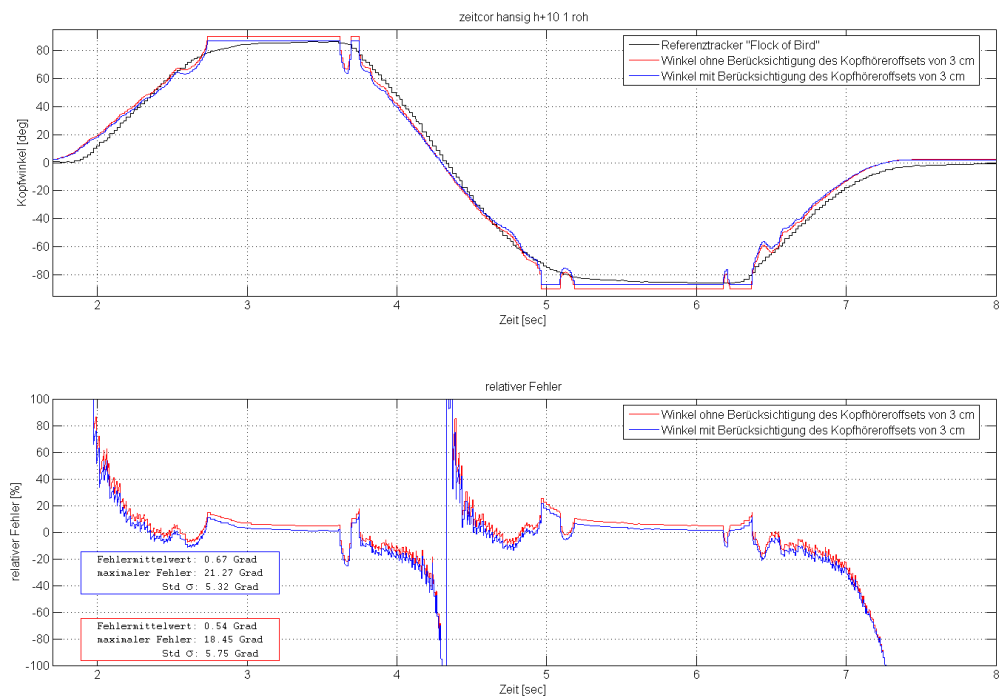


Abb. 7.16: Auswertung der Messung mit Kunstkopf auf Schrittmotor

7.2.5 AMDF / ASDF

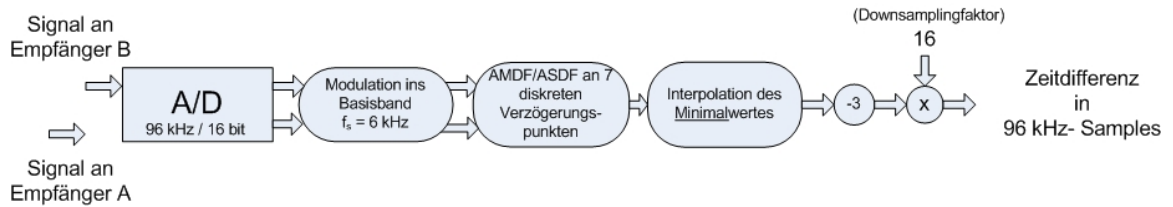


Abb. 7.17: Strukturbild, *AMDF* bzw. *ASDF*

Ähnlich wie in Kap. 7.2.4 sind diese Algorithmen umgesetzt. Anstatt der Multiplikation der beiden Signale (Vgl. Abb. 7.14) wird wie im Kap. 5.4 beschrieben die Differenz der Signale gebildet und diese quadriert (*ASDF*); bzw. der Absolutbetrag dieser Differenz gebildet (*AMDF*) (siehe Abb. 7.18). Eine anschließende Minimalwertsuche und Interpolation dieses Minimalwertes führt zur Schätzung der Laufzeitdifferenz. Wie auch bei der Korrelation im Zeitbereich (siehe Abb. 7.15) bewirkt die nichtlineare Hüllkurve des Nutzsignals s einen periodischen Fehler im Schätzergebnis. Dieser Fehler ist jedoch nicht so ausgeprägt wie in Abb. 7.15 (Vgl. mit Abb. 7.19 und 7.20). Er kann auf gleicher Art wie oben beschrieben bereinigt werden. Im Vergleich zur Korrelation im Zeitbereich ist die Varianz des Schätzers bei großem SNR geringer. Dies bewirkt eine bessere absolute Detektionsgenauigkeit im Vergleich zu Abb. 7.15.

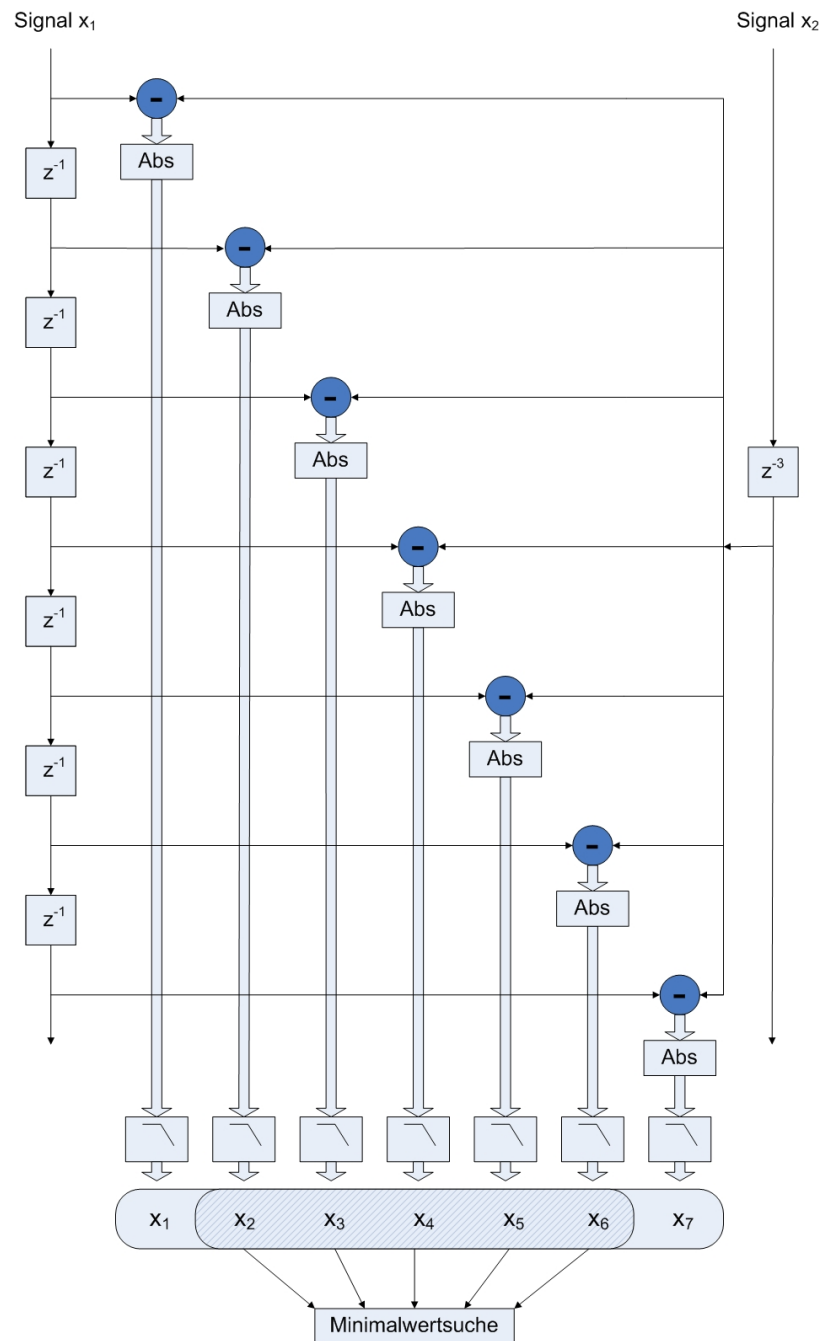


Abb. 7.18: AMDF-Algorithmus

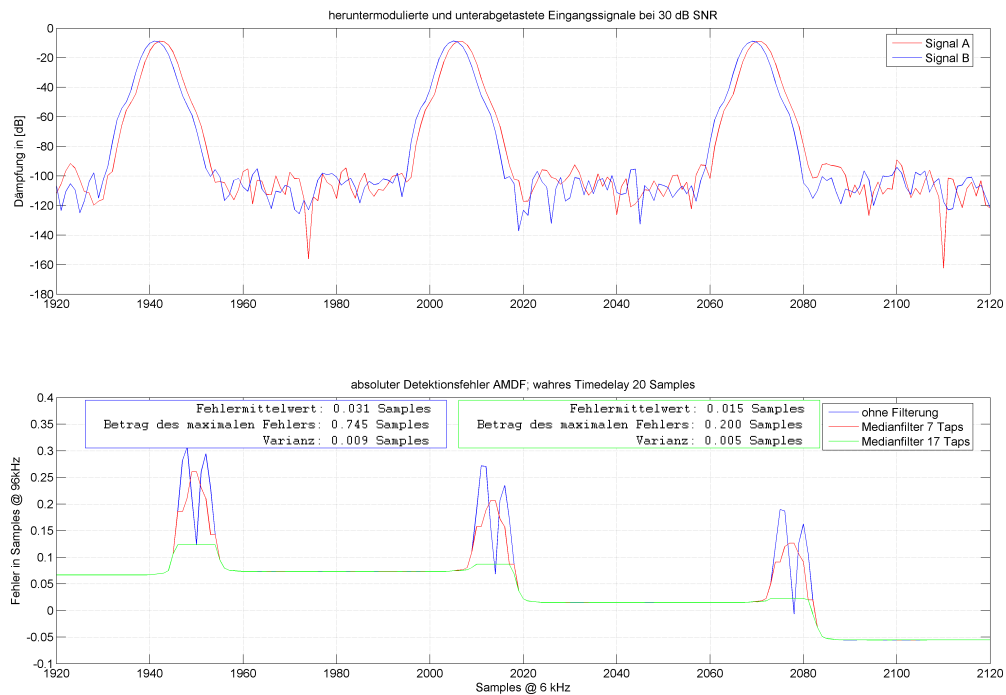


Abb. 7.19: periodische Fehler während AMDF

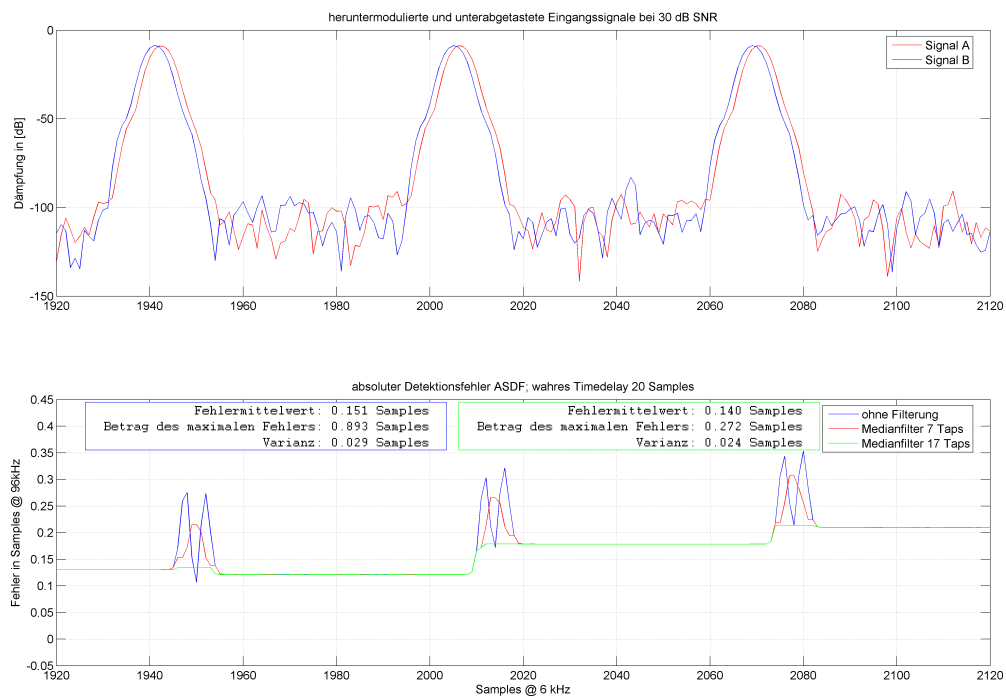


Abb. 7.20: periodische Fehler während ASDF

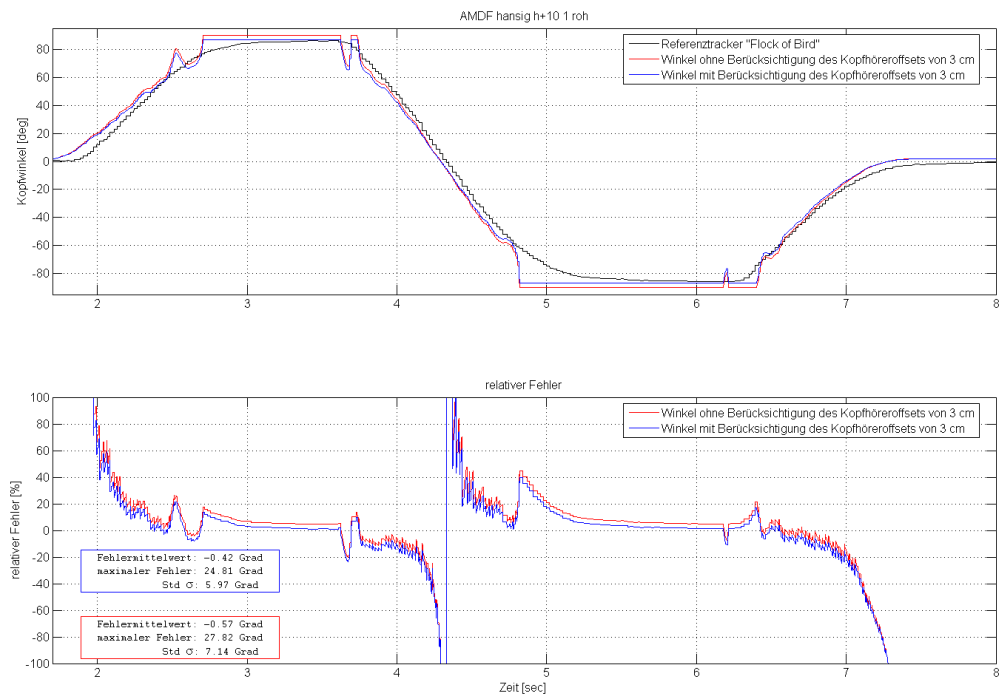


Abb. 7.21: Messung mit Kunstkopf auf Schrittmotor AMDF

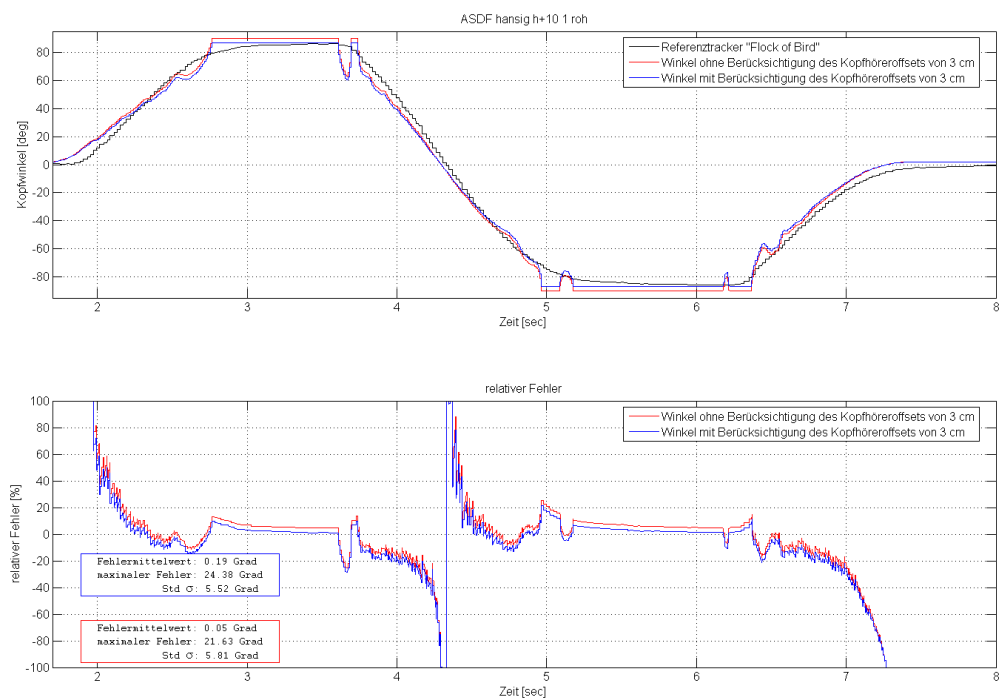


Abb. 7.22: Messung mit Kunstkopf auf Schrittmotor ASDF

7.2.6 Korrelation im Frequenzbereich - PhaseCor

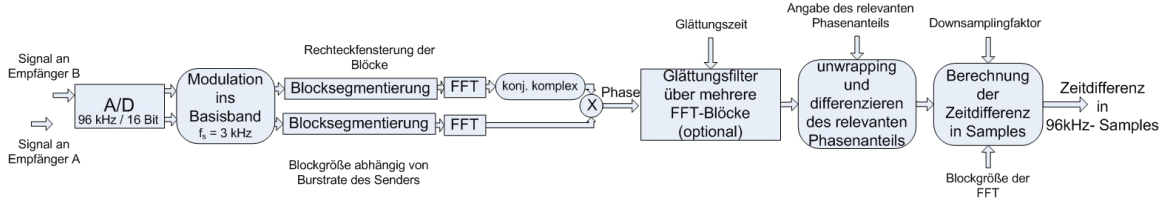


Abb. 7.23: Strukturbild *PhaseCor*

Ausgehend von der Theorie aus Kap.5.3.2 wird hier die Laufzeitdifferenz über die Phase des Kreuzleistungsspektrums der beiden Eingangssignale geschätzt. Der Rechenaufwand ist wie in Kap.7.2.1 beschreiben reduziert, wobei die Abtastfrequenz 3kHz beträgt. Der Rechenaufwand kann damit um das 32fache reduziert werden. Nach der notwendigen Blocksegmentierung für die FFT ist darauf zu achten, dass sich in jedem Signalblock genau ein vollständig empfangener Burst befindet. Nur so ist auch eine Korrelationsberechnung möglich. Die Blockgröße N muss also mit der Burstrate f_B abgestimmt sein (Gl.7.6). Bei einer Taktrate von $f_s = 3\text{kHz}$ und einer Burstrate $f_B = 93.75\text{Hz}$ ergibt sich eine Blockgröße von 32 Samples. Durch die Demodulation und anschließende 3kHz Abtastung wird das verfüg-

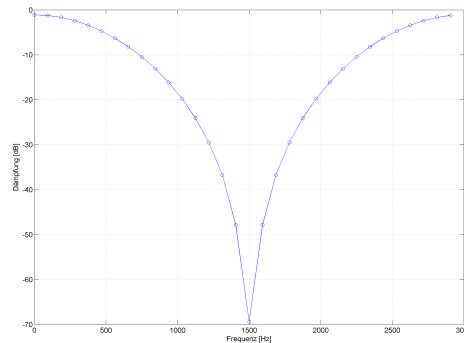


Abb. 7.24: Idealisertes Spektrum vom Eingangssignal nach Demodulation

bare Spektrum viel besser ausgenutzt als bei einer 96kHz Abtastung (siehe Abb.7.24). Da das Spektrum ω_s des Nutzsymbols s bekannt und konstant ist, ist es möglich, ohne statistische Gewichtungungsverfahren wie einem *LS-Schätzer* den Frequenzbereich des Kreuzleistungsspektrums mit der größten Signalenergie und damit dem besten SNR zu bestimmen. Dadurch können Frequenzen, die neben Rauschen keine relevanten Phaseninformation enthalten die Detektion nicht stören (Gl.7.12). In Abb.7.25 ist zu sehen, wie die differenzierte Phase an den Frequenzstellen verrauscht ist, bei welchen keine Energie vom Nutzsymbols s vorhanden ist. Im Detektionsalgorithmus wird daher der Mittelwert der Steigung des Phasengangs an den ersten 4 Frequenzbins als Schätzwert für die Steigung des gesamten Phase genutzt.

$$-\hat{D} = \frac{d}{d\omega} \arg [G_{x_1, x_2}(\omega_s)] \quad (7.12)$$

Ein exponentieller Glättungsfilter der Form nach Gl.7.13, der über mehrere Blöcke die Signalleistung in jedem einzelnen Frequenzbins diskret mittelt, kann zusätzlich noch Rauschanteile unterdrücken (siehe Abb.7.26). Dabei ist jedoch zu beachten, dass ein Glättungsfilter mit zu langer Zeitkonstante die Bewegungsdynamik des Headtrackers einschränkt und der Headtracker bei schnellen Kopfbewegungen merklich nachziehen würde (Vgl. dazu Abb.7.28 mit Abb.7.27)

$$Y_{[n]} = \text{mf} \cdot Y_{[n-1]} + (1 - \text{mf}) \cdot X_{[n]} \quad (7.13)$$

$$0 \leq \text{mf} \leq 1 \quad (7.14)$$

Ist die Steigung der Phase geschätzt, kann nach Gl.7.15 die Laufzeit in Samples ermittelt werden. Durch Multiplikation mit dem Downsamplingfaktor S (hierbei $S = 32$) erreicht man wieder die ursprüngliche 96kHz Zeitauflösung.

$$\hat{D}_{\text{Samples}} = -\frac{d}{d\omega} \arg[G_{x_1, x_2}(\omega_s)] \cdot \frac{S \cdot N}{2 \cdot \pi} \quad (7.15)$$

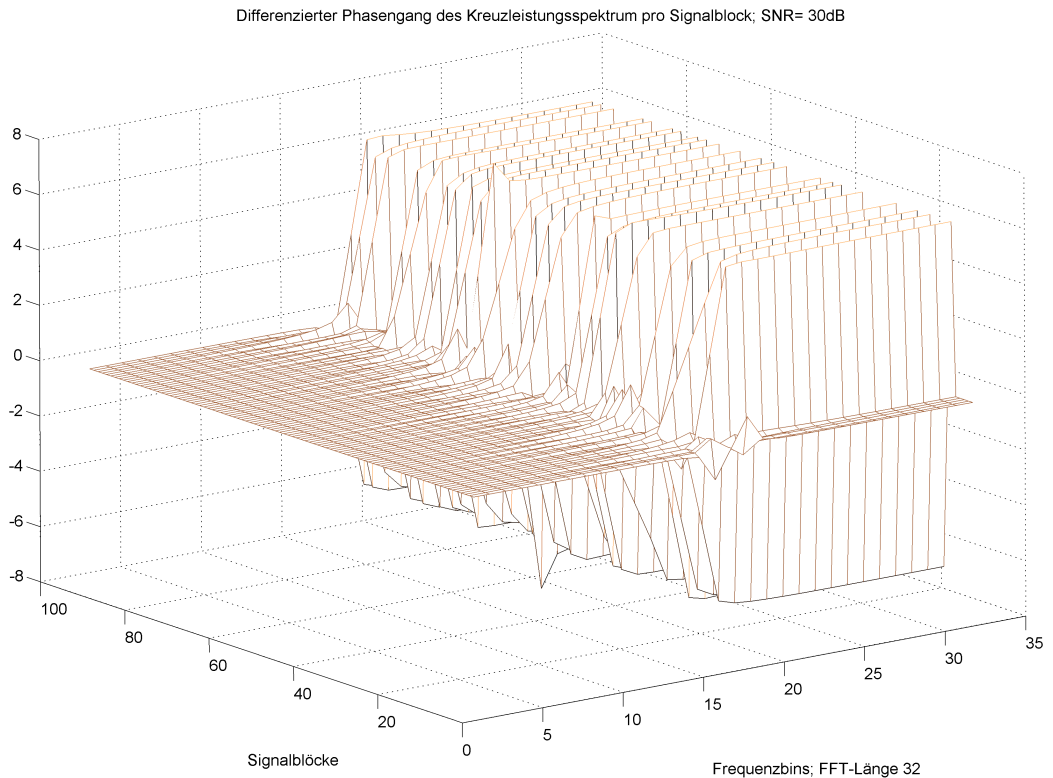


Abb. 7.25: Ableitung der verrauschten Phase des Kreuzleistungsspektrums

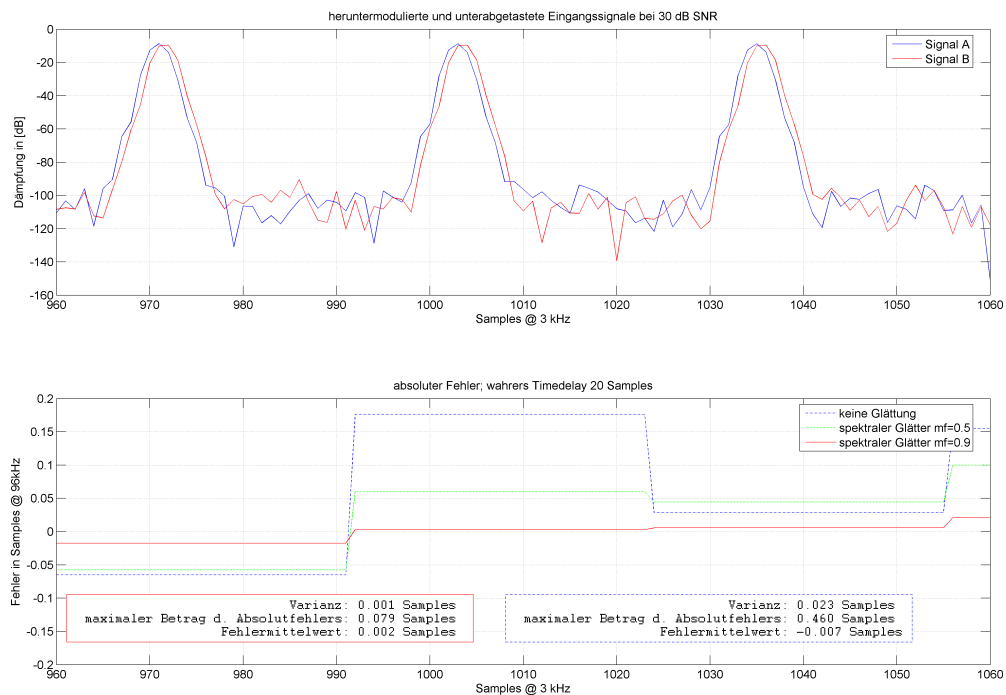


Abb. 7.26: Detektionsergebnis mit verschiedenen Glättungsfiltern

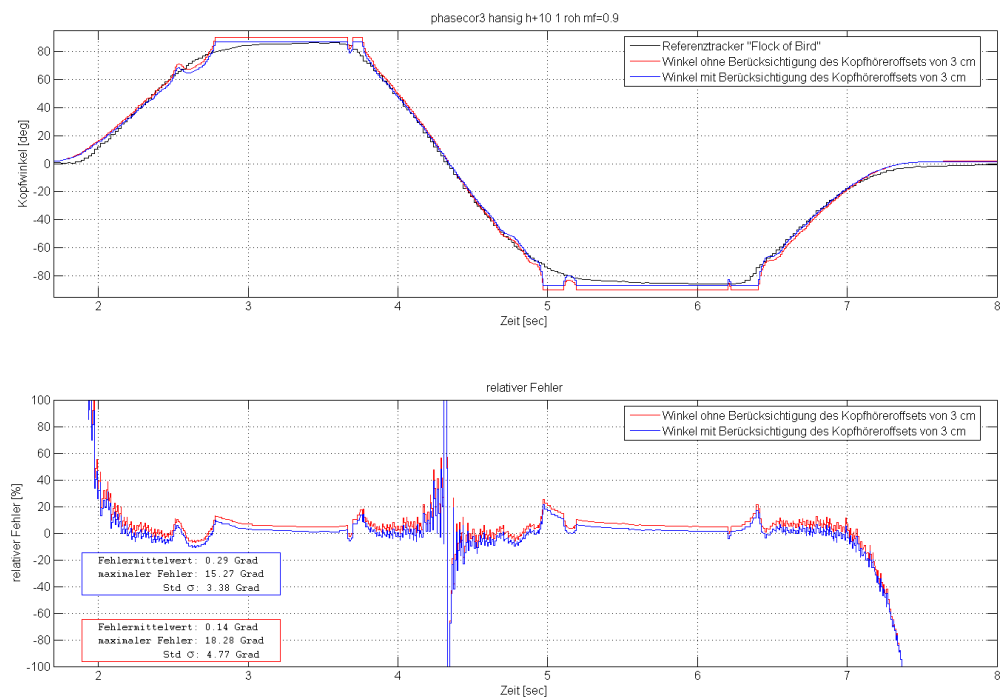


Abb. 7.27: Messung mit Kunstkopf auf Schrittmotor PhaseCor mit spektralen Glättungsfilter mf=0.9

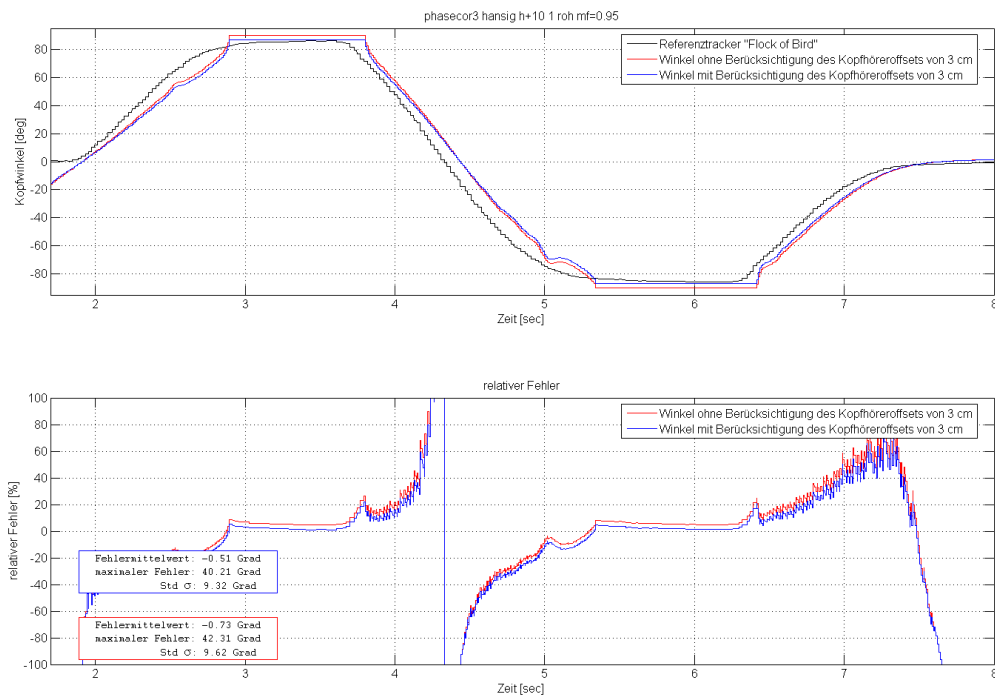


Abb. 7.28: Messung mit Kunstkopf auf Schrittmotor PhaseCor mit spektralen Glättungsfilter mf=0.95

7.3 Übersicht

Abb.7.29 stellt die Genauigkeit der Detektionsalgorithmen gegenüber. Als wahrer Winkel wird dabei immer der Winkel des Referenzsystems *Flock of Birds* genutzt. Die Algorithmen, die eine Laufzeitdifferenz in fraktionalen Samples bestimmen können (*ZeitCor*, *AMDF*, *ASDF*, *PhaseCor*) besitzen im Gegensatz zu den restlichen Ansätzen¹ eine geringere sowie relativ konstante Standardabweichung, die bei großen Winkeln leicht zunimmt. Besonders auffällig ist bei allen Algorithmen die Abweichung vom Referenztracker bei ca. 60°. Bei diesen Winkeln ist der größere Fehler wahrscheinlich auf den Durchmesser, Richtwirkung und gegenseitige Abschattung der Empfänger zurückzuführen. Die vorangegangenen Abbildungen (z.B. Abb.7.8) lassen vermuten, dass einige der getesteten Algorithmen (*Filter*, *Diode*) dem Referenzsystem zeitlich vorausseilen. Das Referenzsystem wäre damit also langsamer als diese Detektionsalgorithmen, womit auch zum Teil die insgesamt höhere Standardabweichung zu begründen wären. Für eine genauere Untersuchung diese Vermutung müssten jedoch die Winkeldaten des Schrittmotors zur Verfügung stehen². Der Algorithmus *PhaseCor* mit einem spektralen Glätter mf= 0.9 besitzt dabei die ge-

¹Die Algorithmen *Filter*, *Diode*, und *BitCor* sind aufgrund der zeitlichen Messauflösung von 96kHz im Detektionsvermögen nach Abb.6.9 beschränkt.

²Bei den Messungen war dies nicht möglich, da das Schrittmotorsystem keinen Export der Winkelinformation vorsieht.

ringste Abweichung gegenüber der Referenz. Hierbei liegt die Standardabweichung unter dem MAMA für geringe Bewegungsgeschwindigkeiten¹. Der geringe Winkelfehler wird also während einer Kopfbewegung nicht wahrgenommen. Nur im statischen Zustand (dieser tritt praktisch nie ein, da Menschen sich immer leicht bewegen) würde die Abweichung registriert. Durch eine Nachbearbeitung der Detektionsergebnisse kann die Qualität des Headtrackerdaten gesteigert werden.

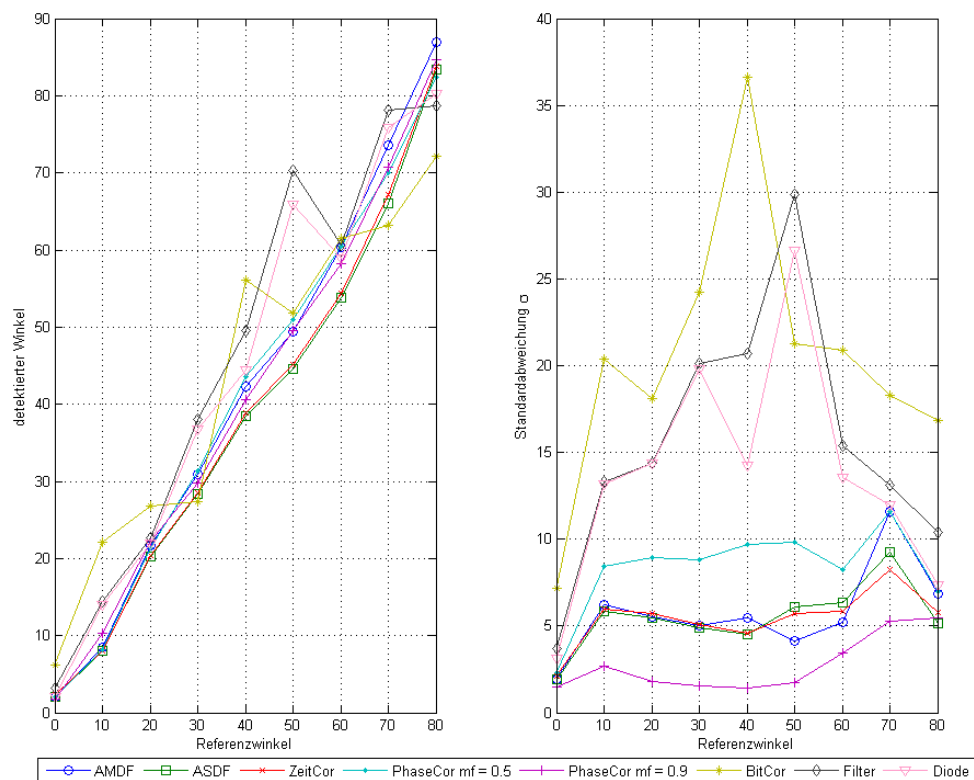


Abb. 7.29: gemittelter Winkel und Standardabweichung der Algorithmen

¹ Siehe Abb. 6.1 auf Seite 67

	Filter	Diode	ZeitCor	AMDF	ASDF	PhaseCor ¹	BitCor ²
Taktrate:	96kHz	analog (hier 96kHz)	6kHz	6kHz	6kHz	3kHz	96kHz
Sendesignal:	40kHz Burstsinal, Länge:2ms hanningfenstert						
Pulsrate / update rate:	187.5 Hz	187.5 Hz	93.75 Hz	93.75 Hz	93.75 Hz	93.75 Hz	39-41kHz Chirp 5ms mit log. Anstieg
Latenz⁴ durch...:	Tiefpassfilter		Demodulation ins Basisband				
Speicher:	Blockverarbeitung 2·1024; 2·z ⁻¹	Integrierer f. Zähler	16·z ⁻¹ ; 4 Zwi- schenspeicher	16·z ⁻¹ ; 4 Zwi- schenspeicher	16·z ⁻¹ ; 4 Zwi- schenspeicher	32-Punkt FFT	FIR- Filterlänge
Add:	0.5	—	0.86 ⁷ + 0.08 ⁶	0.86 ⁷ + 0.12 ⁶	0.86 ⁷ + 0.12 ⁶	0.86 ⁷ + 0.5	98 ²
MOPS⁵ Mult:	1.4	—	1 ⁷ + 0.14	1 ⁷ + 0.14	1 ⁷ + 0.12	1 ⁷ + 0.8	98 ²
sonstiges:			Maxwertsuche/Takt	0.02 Abs(x); Min- wertsuche/Takt	Minwertsuche/Takt	0.01 tan (x)	Maxwertsuche/Takt
-86° - 86°⁴; σ -45° - 45°³; -45° - 45°⁴;	13° 14.7° 13.9°	13° 13.2° 12.5°	5.3° 5.1° 5.1°	6° 5.4° 5.0°	5.5° 4.9° 4.9°	3.3° 2.8° 2.5°	20.6° 20.7° 19.7°
einstellbare Parameter:	Koeff. der Tiefpässe, Schärfungsfunktion, Schwellwert für Trigger	Zählertakt	Medianfilterlänge bzw. Schwellwert bei Fehlerkorrektur; Interpolationsalgorithmus				Schwelle f. 1-Bit- Quantisierung; Referenzsignal

¹Messergebnisse für Algorithmus mit Glättungskoeffizient mf= 0.9

²Angaben für FIR mit 510 Filtertaps

³Standardabweichung ohne Ausgleich des Kopfhöreroffsets

⁴Standardabweichung mit Ausgleich des Kopfhöreroffsets (siehe Kap. 6.4.2 auf Seite 74)

⁵Additionen und Multiplikationen getrennt; ohne Winkelnachbearbeitung

⁶Wird der periodische Fehler innerhalb des Algorithmus nach Abb.7.12 oder Abb. 7.13 kompensiert, erhöht sich Speicher- und Rechenaufwand.

⁷Berechnungsaufwand für eine ESB-Demodulation mit Tschebycheff TP-Filter 10. Ordnung nach [TS99], siehe auch Abb.7.5

7.4 Winkelnachbearbeitung

Um den detektierten Kopfwinkel vor grobe Fehldetektionen und Unsicherheiten zu bewahren, sollte eine Nachfilterung des ermittelten Winkels erfolgen, wie sie z.B. in Abb.7.30 zu sehen ist. Fehler können unabhängig vom Detektionsalgorithmus z.B. durch eintreffende Signalreflexionen oder durch Luftturbulenzen hervorgerufen werden. Die Nachfilterung sollte möglichst latenzfrei sein, da ansonsten die *end-to-end latenz* des Headtrackers zusätzlich ansteigt. Der zu betreibende Aufwand (und damit auch die Anzahl der zusätzlichen Rechenoperationen) hängt natürlich in erster Linie von der Qualität der Daten ab, die der verwendeten Detektionsalgorithmus liefert.

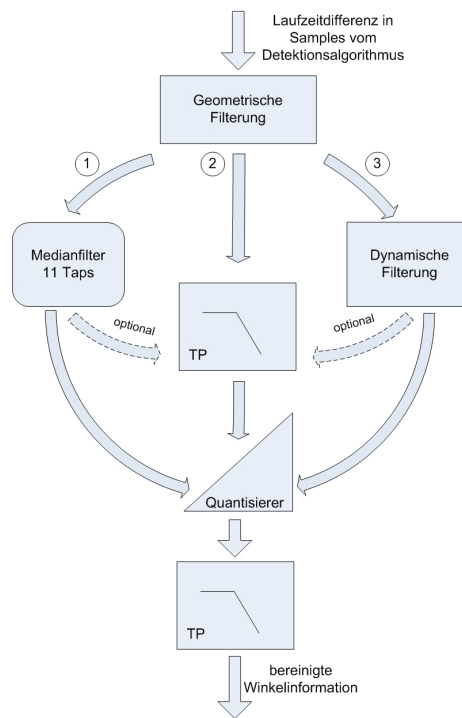


Abb. 7.30: Wege zur Fehlerbereinigung des Detektionsergebnisses

Geometrischen Filterung Um aus einer Laufzeitdifferenz einen Drehwinkel schätzen zu können, muss nach Gl.3.22 die Aperturgröße a bekannt sein. Die Strecke a ist damit der maximale Wegunterschied, den das Sendesignal bei der Schallausbreitung zurücklegen muss. Anhand dieser Vereinfachung kann die detektierte Laufzeitdifferenz nach Gl.7.16 von zu großen Werten befreit werden. Im Falle einer Fehldetektion wird die letzte akzeptierte Laufzeitdifferenz angenommen. Die *update-rate* steigt also kurzzeitig an. Der Algorithmus könnte gleichzeitig dazu verwendet werden, den Arbeitsbereich von vornherein

auf z.B. $\pm 60^\circ$ einzuschränken.

$$D_{\text{Samples}} \leq \frac{a \cdot f_s}{c} \quad (7.16)$$

Medianfilter Durch den Medianfilter werden die Informationen von kurzzeitigen Ausreißern (Impulsrauschen) befreit, ohne dass das Signal wie durch ein Tiefpass verschmiert wird (siehe Abb. 7.31(b)). Im Gegensatz zum Impulsrauschen wird Gauß'sches Rauschen schlecht unterdrückt.

Ein Medianfilter ist ein nichtlinearer Filter, der zu der Klasse der Rangordnungsfilter zählt. Bei einem Medianfilter der Länge $2k + 1$ werden die Werte der Laufzeitdifferenzen in einen $2k + 1$ großen Ringbuffer geschrieben und der Größe nach sortiert. Der mittlere Wert (also der Wert an der Stelle $k + 1$) wird dann ausgegeben. Dadurch können Ausreißerpulse bis zu einer Pulslänge von k entfernt werden. Der Medianfilter erzeugt somit eine nichtkonstante Verzögerung. Längere Störungen oder Änderungen der Signale werden nicht beeinflusst sondern nur um k verzögert (siehe Abb. 7.31(a)). Der hier verwendete Medianfilter hat eine Länge von 11 Werten.

Das Konzept des Medianfilters lässt sich noch durch Einführung von Gewichtungen oder Veränderung der Ausgabe (Minimum- und Maximumfilter) modifizieren ([YI96]). Nachteil des Medianfilters ist die aufwendige Hardwarerealisierung des Sortieralgorithmus in einem Echtzeitsystem.

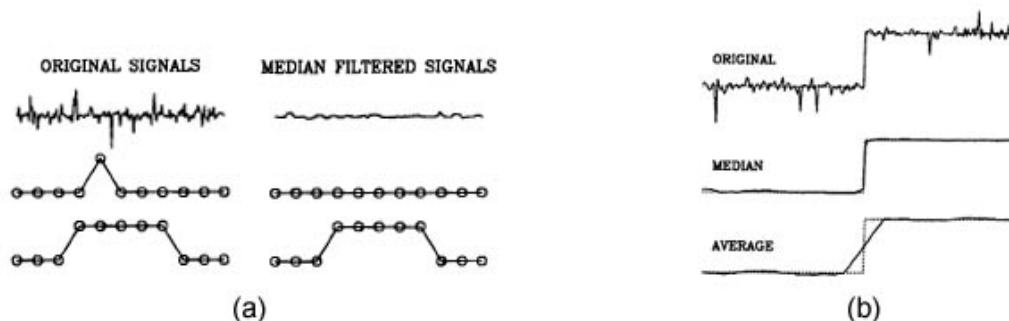


Abb. 7.31: Medianfilterung

Tiefpassfilter In Kombination oder anstelle des Medianfilters kann ein gewöhnlicher Tiefpass 1.Ordnung der Form nach Gl. 7.17 die Winkelinformationen glätten. Dabei bewirkt der Grad des eingestellten Filterkoeffizienten q eine Stabilisierung im statischen Zustand (keine bzw. leichte Kopfbewegungen) aber auch Verschmierung und Trägheit im

dynamischen Fall. Gl.7.18 berechnet die Zeitkonstante τ eines solchen Tiefpasses.

$$y[n] = q \cdot y[n-1] + (1 - q) \cdot x[n] \quad (7.17)$$

$$\tau = \frac{1}{\ln(q)} \cdot \frac{1}{f_s} \quad (7.18)$$

$$0 \leq q \leq 1 \quad (7.19)$$

Dynamische Filterung Nach der geometrischen Bereinigung sollen mittels dieser Heuristik Kopfwinkelveränderungen, die schneller eintreten als ein Benutzer den Kopf bewegen kann, eliminiert werden. Der Algorithmus geht von einer maximalen Kopfdrehung pro Zeiteinheit aus (z.B. $15^\circ/40ms^1$). Ein leicht zu realisierender *moving-average-filter* mittelt über $40ms$ die aktuellen detektierten Kopfwinkel. Gleichzeitig wird der aktuelle Kopfwinkel mit diesem Mittelwert verglichen. Ist die Abweichung der beiden Werte außerhalb eines vorher definierten Toleranzbereiches, wird die aktuelle Winkelinformation verworfen und der letztgültige Winkel beibehalten. Der Mittelwert wird also nur als Vergleichswert genutzt und verändert nicht den Ausgabewinkel (siehe Abb.7.32).

Wird dieser Toleranzbereich zu gering angenommen, kann es passieren, dass der Headtracker bei einer schnellen Kopfdrehung „einfriert“ und erst wieder an dieser Position vom Benutzer „abgeholt“ werden muss. Für diesen Fall sorgt eine Steuerung für eine automatischen Reset der Heuristik: Ein Zähler zählt die verworfenen Winkelinformationen in einer Zeiteinheit und ersetzt beim Erreichen eines maximalen Zählerstandes den Wert des *Moving-Average Filters* durch den aktuellen Winkelwert. Die Heuristik ist als Alternative

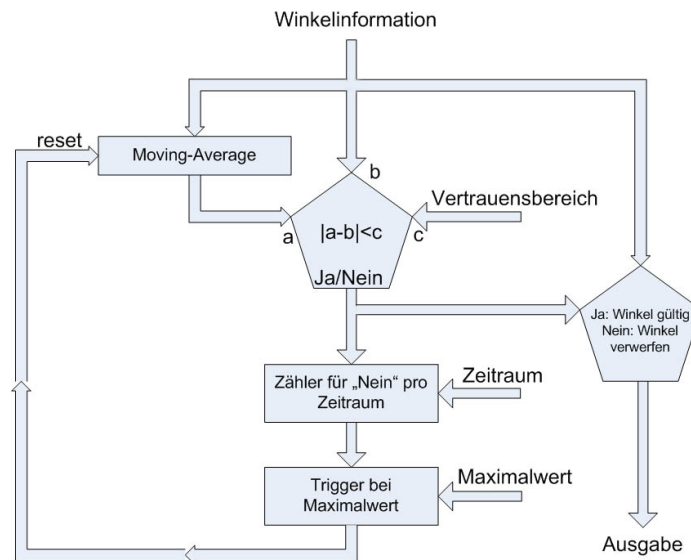


Abb. 7.32: Strukturdiagramm der dynamischen Filterung

zu einer rechenintensiveren Medianfilterung gedacht, wobei die Funktionsfähigkeit von der

¹ Dies entspricht einer kompletten Kopfdrehung von -90 auf 90 Grad in einer halben Sekunde.

geeigneten Wahl der Parameter abhängt. Anders als der Medianfilter kann diese Heuristik nur sehr grobe Ausreißer unterdrücken und keine kleinen Schwankungen herausfiltern.

Winkelquantisierer Durch den Winkelquantisierer wird die azimutale Winkelauflösung reduziert. Zeitliche Unsicherheiten der Detektion können innerhalb der Quantisierungsschwelle unterdrückt werden. Argumentierend mit dem MAMA¹ für langsame Kopfbewegungen hat eine Winkelquantisierung von 3-5 Grad auch keine signifikanten hörbaren Auswirkungen. Vorteil dieser Maßnahme ist jedoch, dass die Winkelinformation für das Binauralmodul sich wirklich nur noch dann ändert, wenn dies auch wahrnehmbar ist. Ein nachgeschalteter leichter Tiefpass soll für langsame Bewegungen von einem quantisierten Winkelwert zum Nächsten überblenden und dadurch zu harte (hörbare) Sprünge vermeiden. Wird der Winkel zwischen Headtracker und Binauralmodul nur dann übermittelt wenn dieser sich auch ändert², kann der Winkelquantisierer den Datenfluss verringern. Abhängig von der Architektur des Binauralmoduls könnten dadurch auch unnötige Koeffizientenberechnungen für das Audiorendering bei unmerklichen Winkeländerungen vermieden werden.

7.4.1 Ergebnisse der Nachbearbeitung

Aufgrund der guten Messergebnisse nach Abb.7.29 wurden die Algorithmen *ASDF*, *AMDF*, *PhaseCor* und *ZeitCor* ausgewählt und mit verschiedenen Kombinationen der vorgestellten Nachbearbeitungsmethoden getestet. Es wurde nach Abb.7.30 verschiedene Verbindungen aus Medianfilter (11 Taps), Tiefpass mit verschiedenen Filterkoeffizienten ($q = 0, 0.9, 0.99$) und Winkelquantisierer (1, 3 oder 5 Grad) getestet. Eine geometrische Filterung wurde immer durchgeführt, da „akzeptierte“ Laufzeitdifferenzen keine Latenzen und Veränderungen verursachen und dieser Filter daher generell zu empfehlen ist.

Die erreichten Standardabweichungen sind in Tabelle 7.1 zu sehen. Es zeigte sich, dass der Medianfilter bei den Algorithmen *AMDF*, *ASDF* und *ZeitCor* keine signifikante Verbesserung bewirkt, da ein Medianfilter mit 15 Taps bereits im eigentlichen Detektionsalgorithmus integriert ist³. Im Algorithmus *PhaseCor* bewirkt der Medianfilter eine Verzögerung und damit eine höhere Standardabweichung⁴. Ein Medianfilter mit einem geringeren Zeitfenster sollte hierbei besser funktionieren. Da im Algorithmus von *PhaseCor* ein spektraler Glättungsfilter eingesetzt wird, bewirkt eine zusätzliche Tiefpassfilterung der Winkelergebnisse sofort eine Verzögerung. Auch bei den anderen Algorithmen war zu beobachten, dass starke Tiefpässe über $q = 0.999$ prinzipiell zu einem Nachziehen führen.

¹Zu den Untersuchungsergebnisse von [DA91] und [WZ97] siehe Kap. 6.2 auf Seite 67.

²Der Logitech Ultraschallheadtracker kann auch in diesem Modus senden. Er wird nach [LO92] als *Incremental Reporting Mode* bezeichnet.

³Um den periodischen Fehler nach Abb.7.15 bzw. 7.20 und 7.19 zu unterdrücken wird innerhalb der getesteten Algorithmen ein Medianfilter nach Abb.7.12 verwendet.

⁴*PhaseCor* liefert nur alle 10,6ms eine Winkelinformation (im Gegensatz zu den drei anderen Algorithmen (0,6ms)). Ein Medianfilter von 11 Taps besitzt somit ein Zeitfenster von 117ms.

Die Resultate der Kombinationen mit der geringsten Standardabweichung von jedem der vier Algorithmen ist in Abb.7.33 zu sehen (auch markiert in Tabelle 7.1). Die Messung des Algorithmus mit der geringsten Standardabweichung (*PhaseCor*) ist abschließend in Abb.7.34 zu sehen. Als Alternative zum *PhaseCor*-Algorithmus kann *AMDF* betrachtet werden. Hierbei ist der Rechenaufwand geringer und es entfällt das Problem¹ der FFT-Blockverarbeitung.

Median- filter	TP- Koeff.	Quanti- sierung [Grad]	ASDF		PhaseCor		ZeitCor		AMDF	
			σ		σ		σ		σ	
			-86°-86°	-45°-45°	-86°-86°	-45°-45°	-86°-86°	-45°-45°	-86°-86°	-45°-45°
—	0	1	5,36	4,77	2,88	2,30	5,12	4,87	5,60	4,83
—	0	3	5,63	5,23	3,37	2,60	5,46	4,97	,96	4,89
—	0	5	5,43	4,99	3,06	2,87	5,16	5,05	5,59	4,98
—	0,9	1	5,15	4,41	2,93	2,37	4,88	4,53	5,31	4,39
—	0,9	3	5,41	4,54	3,41	2,66	5,17	4,65	5,68	5,37
—	0,9	5	5,23	4,63	3,08	2,87	4,91	4,74	5,29	4,56
—	0,99	1	4,68	3,36	5,48	5,89	4,32	3,40	3,86	2,72
—	0,99	3	4,93	3,63	5,73	6,06	4,59	3,64	4,30	2,94
—	0,99	5	4,78	3,94	5,59	6,28	4,41	3,94	3,98	3,34
x	0	1	5,25	4,57	4,79	5,06	5,00	4,67	5,46	4,59
x	0	3	5,52	4,68	5,08	5,21	5,34	4,78	5,83	4,65
x	0	5	5,32	4,80	4,94	5,40	5,04	4,86	5,45	4,76
x	0,9	1	5,05	4,48	5,15	5,49	4,77	4,34	5,17	4,16
x	0,9	3	5,32	4,37	5,42	5,65	5,07	4,46	5,56	5,14
x	0,9	5	5,13	4,46	5,32	5,84	4,80	4,56	5,16	5,05
x	0,99	1	4,79	3,47	9,10	10,02	4,43	3,50	3,89	2,86
x	0,99	3	5,03	3,73	9,25	10,13	4,69	3,74	4,31	3,07
x	0,99	5	4,87	4,03	9,16	10,31	4,50	4,01	4,01	3,34

Tab. 7.1: Standardabweichung bei Nachbearbeitung des Winkelergebnisses

¹Bei der FFT-Blocksegmentierung muss das Empfangssignal vollständig in beiden zu korrelierenden Signalblöcken enthalten sein.

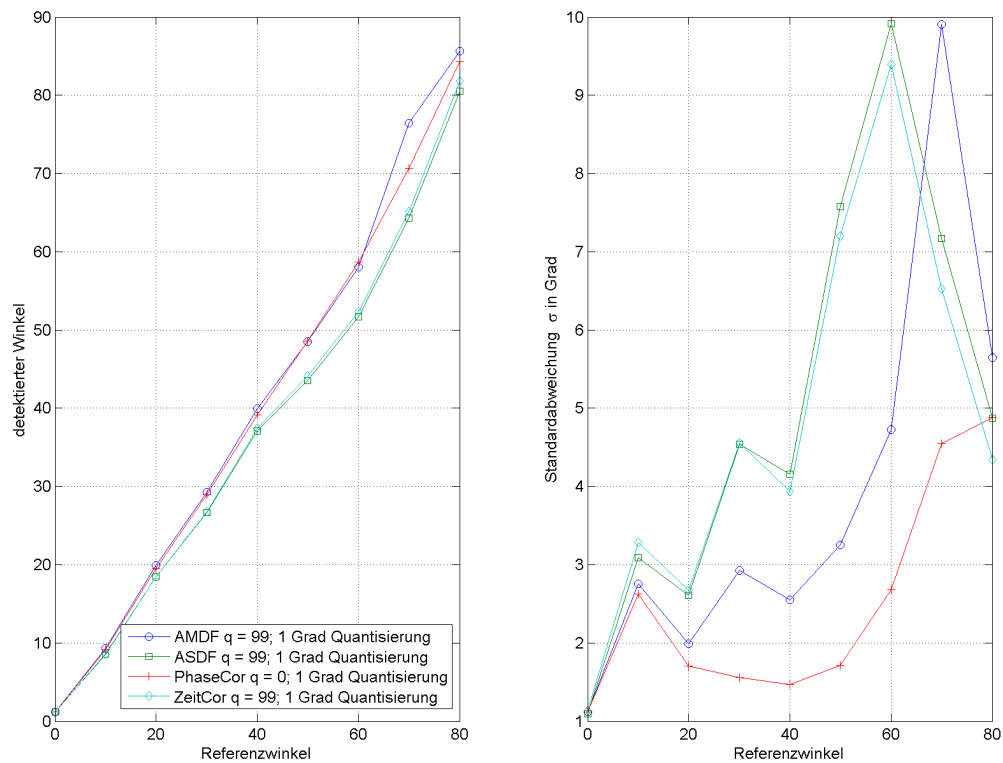
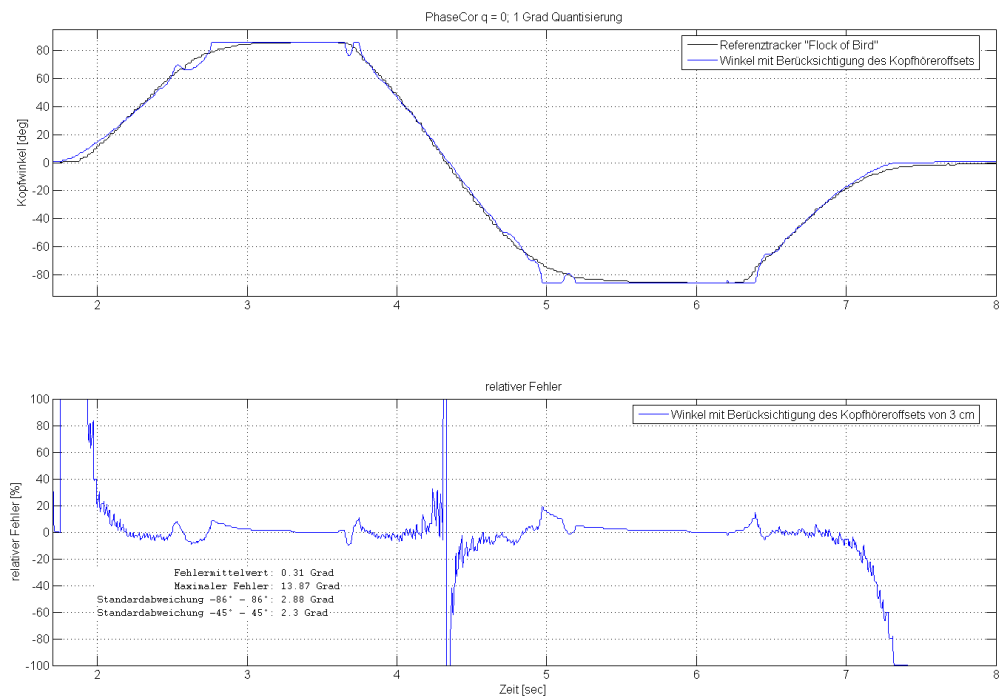


Abb. 7.33: Winkelergebnisse der Nachbearbeitung

Abb. 7.34: Messung von *PhaseCor*

7.5 Zwei-System Lösung

Werden 2 Headtrackersysteme in unmittelbarer örtlicher Nachbarschaft gleichzeitig betrieben, vermischen sich an den Empfängern die Signale beider Sender. Abhängig von der Signallänge, der Laufzeit, der Aussendefrequenz f_B und dem Aussendezeitpunkt treffen die Signale beider Sender zeitversetzt oder (teilweise) zeitlich überlagert am Empfänger ein (siehe Abb.7.35). Die Aufgabe des Headtrackers ist es, aus dem empfangenen Mischsignal die Sendesignale von Sender A und Sender B zu trennen und richtig zuzuordnen. Erst danach kann die eigentliche Kopfwinkeldetektion mit Hilfe der vorgestellten Algorithmen unter Verwendung eines extrahierten Sendesignals erfolgen.

Eine einfache Möglichkeit zum Unterscheiden der Signale ergibt sich, wenn die Transmitter mit unterschiedlichen Trägerfrequenzen arbeiten: Befinden sich die Sendesignale in unterschiedlichen Frequenzbändern, können durch geeignete Bandpassfilter die Burstsignale Empfängerseitig separiert werden. Der Vorteil dieses Verfahrens ist, dass eine zeitliche Überlagerung der verschiedenen Sendesignale am Empfänger niemals zu einer falschen Separation führt (Siehe Frequenzmultiplex-Übertragung Kap.4.3.1). Da die zur Verfügung stehenden Sensoren¹ nicht die erforderliche Bandbreiten zulassen, konnte dieser viel versprechende Ansatz nicht weiter verfolgt werden. Ein mit den verfügbaren Sensoren umsetz-

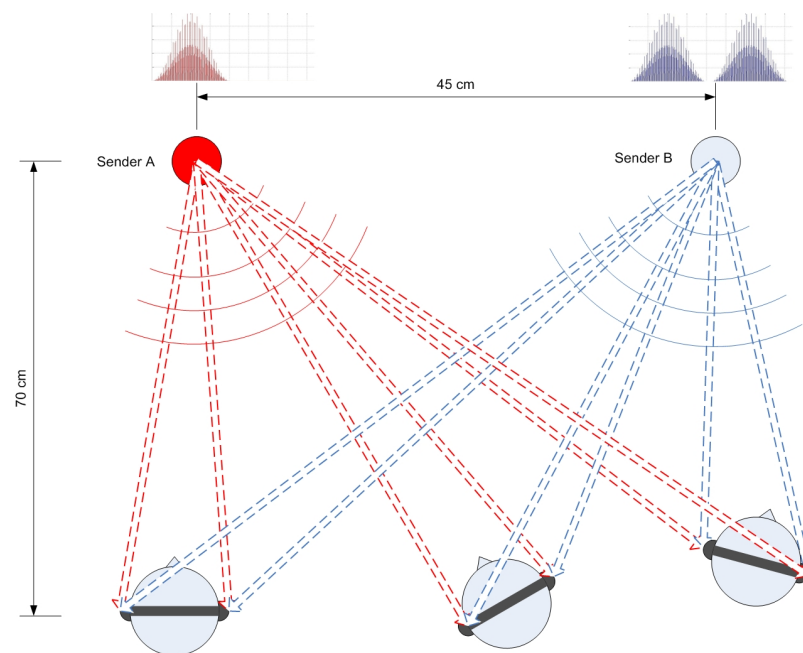


Abb. 7.35: Situation im Parallelbetrieb

barer Lösungsansatz stellt das Zeitmultiplexverfahren dar (siehe Kap.4.3.1). Dabei senden die Transmitter kontrolliert zeitversetzt ihre Burstsignale aus - die Sender müssen also synchronisiert sein. Aufgrund dieser Gegentaktfrequenz und den Laufzeiten der Signale ergibt

¹Übertragungsbandbreite der Sensorstrecke siehe Anhang B auf Seite 118

sich um die Transmitter herum ein Raum, indem die Signale noch zeitversetzt die Sender erreichen. Senden die Transmitter gegentaktig mit einer Frequenz von 50Hz , ergibt sich unter Berücksichtigung der Signallänge der verwendeten Signale und einem Senderabstand von 45cm ein Detektionsbereich von ca. 70cm (siehe Abb.7.35). Sollen die Signale in einem größeren Bereich eindeutig bleiben oder möchte man den Abstand der Sender vergrößern, muss die Taktrate weiter reduziert werden. Dadurch wird auch die *update-rate* des Systems weiter verringert (siehe Kap.6.2.2). Wie in Abb.7.35 angedeutet ist, wird am Sender A ein 2ms langer hanngingefensterter Burst verwendet, am Sender B wird dieses Signal hingegen 2 mal pro Takt hintereinander abgeschickt. Mittels einer Schwellwertdetektion des Hüllkurvenverlaufes nach Abb.7.36 lassen sich diese Signale recht einfach unterscheiden (Vgl. auch mit Kap.5.1.1 und 7.2.2). Die Hüllkurvenanalyse erzeugt eine Latenz von ca 3ms , die natürlich auch die *end-to-end Latenz* des Systems steigert (siehe ebenfalls Kap.6.2.2).

Aufgrund der Verminderung der *update-rate*, der Erhöhung der *end-to-end Latenz* sowie der Notwendigkeit zur Synchronisation der Sender stellt das Zeitmultiplex-Verfahren nicht gerade die optimale Lösung dar. Das prinzipiell bessere Verfahren wäre z.B. eine Frequenzmultiplex-Übertragung. Allerdings stellt das Zeitmultiplex-Verfahren bei der Bandbreite der verwendeten Sensoren wohl die einzige Möglichkeit zur Realisierung einer Zwei-System Lösung dar.

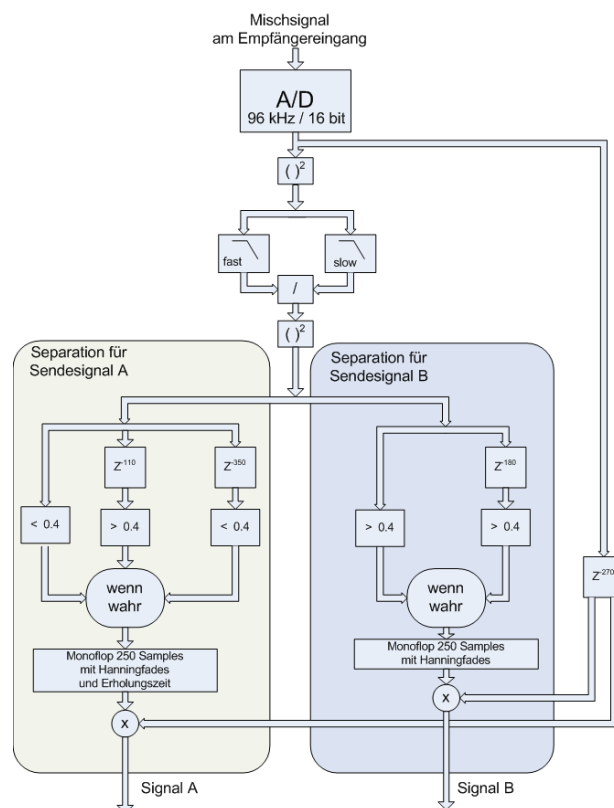


Abb. 7.36: Strukturbild Signalseparation einer Zwei-System Lösung

Kapitel 8

Ausblicke

Diese Arbeit hat gezeigt, wie ein schnurloses, zeitvariantes Kopfhörersystem kostengünstig realisiert werden kann. Dabei konnten praktisch alle Anforderungen nach Kap.1.3 erfüllt werden. Es wurde eine Lösung für ein Mehrbenutzersystem gezeigt, wobei für diesen Anwendungsfall Sensoren mit größerer Bandbreite von Vorteil wären (z.B. von Measurement Specialties¹ oder vom Fraunhofer Institut²).

Unter Laborbedingungen braucht der Ultraschall-Headtracker einem Vergleich mit einem teuren kabelgebundenen magnetischen Trackingsystem nicht scheuen. In der Praxis muss diese Funktionalität jedoch nicht ohne weiteres gegeben sein. Im Automobil werden in erster Linie Reflexionen von den Seitenscheiben die Winkeldetektion stören. Durch eine geschickte Positionierung des Senders sollten unnötige Reflexionen vermieden werden. Vormessungen im Auto haben gezeigt, dass eine Sender am Fahrzeughimmel weniger kritisch ist als eine Senderposition im Bereich der Kopfstütze. Eine am Sender angebrachte „Schallblende“ könnte den Beschallungsbereich zur Seite der Fensterscheibe hin mechanisch einschränken und damit auch Reflexionen reduzieren.

Steht am DSP des Kopfhörers noch genügend Rechenleistung zur Verfügung ist auch eine Optimierung der Winkelergebnisse mit einem stochastischen Zustandsschätzer denkbar. Für Positionstracking bietet sich ein Kalman Filter an: [WE97], [FO96] oder [EM94].

¹http://www.meas-spec.com/piezo/ultrasonic_transducers.htm, copyright © Measurement Specialties, Inc., 1000 Lucas Way Hampton VA 23666 (2005)

²http://www.ibmt.fraunhofer.de/Produktblaetter/US_ps_uswandler_de.pdf, copyright © Fraunhofer Institut für Biomedizinische Technik, Abteilung Ultraschall, Arbeitsgruppe Piezosysteme & Entwicklung

Literaturverzeichnis

- [BH93] BHATNAGAR D. K.: „*Position trackers for Head Mounted Display system: A survey*“ (1993)
- [BI80] BIBER C., ELLIN, S., SHENK E., STEMPECK J.: „*The Polaroid ultrasonic ranging system*“ AES preprint 1696 (A-8) (1980)
- [BJ03] BJÖRKLUND Svante: „*Experimental Evaluation of some Cross Correlation Methods for Time-Delay Estimation in Linear Systems*“ Universität Linköping Schweden (2003)
- [BL02] BLOK Ewa: „*Classification and evaluation of discrete subsample time delay estimation algorithms*“ 14th International Conference on Microwaves, Radar and Wireless Communications (2002)
- [BR80] BRENDDEL K.: „*Ultraschall - Phänomene und ihre Anwendung*“ PTB Braunschweig zur DAGA '80, VDE Verlag (1980)
- [BS00] BREITSCHÄDEL Hannes: „*Aufbau eines binauralen Raumsimulators für fünf Schallquellen unter Berücksichtigung der Kopfdrehung*“ Diplomarbeit, TU-Graz (2000)
- [BU03] BURDEA G.C., COIFFET P.: „*Virtual Reality Technology*“ second edition, John Wiley & Sons, New Jersey (2003)
- [CA73] CARTER G.C., NUTTALL A.H., CABLE P.G.: „*The Smoothed Coherence Transform*“ IEEE Processing, Vol. 61 (1973)
- [CA76] CARTER G.C., KNAPP C.H.: „*The Generalized Correlation Method for Estimation of Time Delay*“ IEEE Transactions on Speech and Audio Processing, Vol. ASSP-24, Nr. 4 (1976)
- [CA87] CARTER G.C.: „*Coherence and time delay estimation*“ IEEE (1987)
- [CA81] CARTER G.C.: „*Time Delay Estimation for Passive Sonar Signal Processing*“ IEEE Transactions on acoustic, speech, and signal processing, VOL. ASSP-29, Nr. 3 (1981)

- [CI84] CIARCIA S.: „*An ultrasonic ranging system*“ Byte Mag. 9, 113-123 (1984)
- [CH78] CHAN Y. T., v. HATTIN R., PLANT J.B.: „*The least squares estimation of time delay and its use in Signal detection*“ IEEE Transactions on acoustics, speech and signal processing, Vol. 26 Nr.3 (1978)
- [CR62] CRON B.F., SHERMAN C.H.: „*Spatial-correlation functions of various noise models*“ J. Acoust. Soc. Am. Vol.34 Nr.11, 1732-1736 (1968)
- [DA91] DAVID W., GRANTHAM C., GRANTHAM D.W.: „*Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity*“ J. Acoust. Soc. Am. 91/3, 1624-1636 (1991)
- [EB02] EBISAWA Yoshinobu: „*A Pilot Study on Ultrasonic Sensor-Based Measurement of Head Movement*“ IEEE Transaction on instrumentation and measurement Vol. 51, Nr. 5 (2002)
- [EH85] EHLERS Dirk: „*Ein Beitrag zur Bestimmung von Abständen in Luft mit Ultraschall*“ Dissertation der TU Berlin (1985)
- [EM94] EMURA S., TACHI, S.: „*Sensor fusion based measurement of human head motion*“ Workshop on Robot and Human Communication IEEE (1994)
- [FA89] FAY Richard R.: „*Hearing in Vertebrates: a Psychophysics Databook.*“ J. Acoust. Soc. Am. Volume 86, Issue 5, p. 2044 (1989)
- [FA98] FASOLD Wolfgang, VERES Eva: „*Schallschutz und Raumakustik in der Praxis*“ Verlag für Bauwesen, Berlin (1998)
- [FA00] FARINA A.: „*Simultaneous measurement of impulse response and distortion with a swept-sine technique*“ preprint of 110th AES Convention, Paris (2000)
- [FE98] FELDERHOFF Uwe, REICHENAUER Klaus, THEILE Günther: „*Stabilität der Lokalisation bei verfälschter Reproduktion verschiedener Merkmale der binauralen Signale*“ Vortrag zur VDT-Tagung (1998)
- [FI90] FIGUEROA J.Fernando, LAMANCUSA John S.: „*Ranging errors caused by angular misalignment between ultrasonic transducer pairs*“ J. Acoust. Soc. Am. 87, 1327-1335 (1990)
- [FI92] FIGUEROA J.Fernando, LAMANCUSA John S.: „*A method for accurate detection of time of arrival: Analysis and design of an ultrasonic ranging system*“ J. Acoust. Soc. Am. 91/1, 486-494 (1992)
- [FI01] FIGUEROA J.F., TONG C. C., BARBIERI E.: „*A Method for Short or Long Range Time-of-Flight Measurements Using Phase-Detection With an Analog Circuit*“ IEEE Transaction on instrumentation and measurement, VOL. 50, Nr. 5 (2001)

- [FO83] FOX J.D., KHURI-YAKUB B.T., KINO G.S.: „*High frequency acoustic wave measurement in air*“ IEEE 1983 Ultrasonic Symposium Atlanta 581-584 (1983)
- [FO96] FOXLIN E.: „*Inertial Head-Tracker Sensor Fusion by a Complementary Separate-Bias Kalman Filter*“ IEEE Proceedings of VRAIS'96 (1996)
- [GM81] GAMMELL P.M.: „*Improved ultrasonic detection using the analytic signal magnitude*“ Ultrasonics 19(2) 73-76 (1981)
- [GO97] GOLDSTEIN S.R., DAUBE-WITHERSPOON M.E., GREEN M.V., EIDSATH A.: „*A Head Motion Measurement System Suitable for Emission Computed Tomography*“ IEEE Transactions on medical imaging Vol. 16 Nr. 1 (1997)
- [GR02] GRABER Gerhard, WESELAK Werner: „*Raumakustik*“ Vorlesungsskript TU-Graz (2002)
- [GR03] GRANTHAM D.W., HORNSBY B.W.Y., ERPENBECK E.A.: „*Auditory spatial resolution in horizontal, vertical, and diagonal planes* J. Acoust. Soc. Am. 114/2 (2003)
- [HA69] HAFFERL Anton: „*Lehrbuch der topographischen Anatomie*“ Springer Verlag 3. Auflage (1969)
- [HA05] HAUCK Peter: „*Kryptologie und Datensicherheit*“ Vorlesungsskript Universität Tübingen (2005)
- [IEM1] LEITNER S., SONTACCHI A., HÖLDRICH R.: „*Kopfbezogene Wiedergabe für Binauralsignale mit Ambisonic*“ Institut für Elektronische Musik und Akustik Graz
- [JA93] JACOVITTI G., SCARANO G.: „*Discrete Time Techniques for Time Delay Estimation*“ IEEE Transactions on acoustics, speech and signal processing, Vol. 41 Nr.2 (1993)
- [KL81] KLEINSCHMITT P., MÁGORI V.: „*Ultraschalldistanzsensoren zur berührungslosen Objekterfassung*“ Siemens Forschungs- und Entwicklungsbericht Band 10, Nr. 2, S.110-118 (1981)
- [KI84] KIM W.K., PARK S.B., JOHNSON S.A.: „*Signal-to-noise ratio and bandwidth for pseudo-random codes in an ultrasonic image system*“ Ultrason.Imag. 6, 113-123 (1984)
- [KO85] KOCIS S.: *ultrasonic numerical spatial measurement of the robot arm movement* Imeko 10th world congress (1985)
- [KU88] KUTTRUFF Heinrich: „*Physik und Technik des Ultraschalls*“ S. Hirzel Verlag Stuttgart (1988)

- [KU99] KUANG W.T., MORRIS A.S.: „*Ultrasonic Doppler distance measurement technique for robot tracking system*“ IEEE Electronic letters Vol. 35 Nr. 11 (1999)
- [KU00] KUANG W.T., MORRIS A.S.: „*Combined time-of-flight and Doppler ultrasonic tracking system for better performance in robot tracking*“ IEEE Proc.-Sci. Meas. Technol. Vol. 147 Nr. 5 (2000)
- [KU02] KUANG W.T., MORRIS A.S.: „*Using Short-Time Fourier Transform and Wavelet Packet Filter Banks for Improved Frequency Measurement in a Doppler Robot Tracking System*“ IEEE Transactions on Instrumentation and Measurement, Vol. 51 Nr. 3 (2002)
- [LO92] „*3D Mouse & Head Tracker Technical Reference Manual*“ Logitech Inc., Fremont, Kalifornien USA (1992)
- [MA92] MARIOLI D., NARDUZZI C., OFFELLI C., PETRI D., SARDINI E., TARONI A.: „*Digital Time-of-Flight Measurement for Ultrasonic Sensor*“ IEEE Transaction on Instrumentation and Measurement. Vol. 41 Nr. 1 (1992)
- [MA98] MACKENSEN Philip, REICHENAUER Klaus, THEILE Günther: „*Einfluß der spontanen Kopfdrehungen auf die Lokalisation beim binauralen Hören*“ Vortrag zur VDT-Tagung (1998)
- [MA02] MÄUSL Rudolf, GÖBEL Jürgen: „*Analoge und digitale Modulationsverfahren*“ Verlag Jürgen Göbel Hüthig Telekommunikation (2002)
- [MC04] MCLEOD I. D., DON C. G., SWENSON G. G.: „*Acoustic pulse propagation in an outdoor turbulent atmosphere*“ J. Acoust. Soc. Am. 116/5 (2004)
- [ME86] MENNE D., HACKBARTH H.: „*Accuracy of distance measurement in the bat Eptesicus Fuscus: Theoretical aspects and computer simulations*“ J. Acoust. Soc. Am. 79, 386-397 (1986)
- [ME92] MEYER K., APPLEWHITE H.L., BIOCCA F.A.: „*A survey of position trackers.*“ Presence 1 p.173 - 200 (1992)
- [MO87] MOREL J.B.: „*Acoustic ranging on an air table*“ Mech. Eng. 109(8) 72-73 (1987)
- [NA95] NANDI Asoke K.: „*On the Subsample Time Delay Estimation of Narrowband Ultrasonic Echoes*“ IEEE Transactions on ultrasonic, Ferroelectrics and Frequency Control, Vol. 42 Nr. 6 (1995)
- [OH04] OHM Jens-Reiner, Lüke Hans D.: „*Signalübertragung*“ Springer Verlag (2004)
- [ON84] ONO T., KOHATA M., MIYAMOTO T.: „*Ultrasonic phase-sensitive rangfinder with double modulation Doppler-free methode for shallow seafloor survey*“ in IEEE 1984 Ultrasonic Symposium Dallas 480-483 (1984)

- [OP75] OPPENHEIM A.V., SCHAFER: „*Digital Signal Processing*“ Prentice-Hall (1975)
- [PA91] PARILLA M., ANAYA J.J., FRISTSCH C.: *Digital Signal processing techniques for high accuracy ultrasonic measurements* IEEE (1992)
- [PE88] PERROTT D. R.: „*MAMA as a function of signal frequency and the velocity of the source*“ J. Acoust. Soc. Am. 83, 1522-1527 (1988)
- [PI81] PIERSOL A. G.: „*Time Delay Estimation Using Phase Data*“ IEEE Transactions on Speech and Audio Processing, Vol. ASSP-29 Nr.3 (1981)
- [RU97] RUSER H., HEIDE P., VOSSIEK M., JENA A.v., Magori V., Tränkler H.-R.: „*A high-sensitive microwave-ultrasonic sensor for reliable measurement of motion and range*“ Preprint Eurosensors XI Warschau, vol.2 (1997)
- [SA96] SANDVAD J.: „*Dynamic aspects of auditory virtual environments*“ 100th AES Convention, Copenhagen (1996)
- [SC81] SCARBROUGH K., AHMED N., CARTER G. C.: „*On the Simulation of a Class of Time Delay Estimation Algorithms*“ IEEE Transactions on Speech and Audio Processing, Vol. ASSP-29, Nr. 3 (1981)
- [SR85] SORGE George, HAUPTMANN Peter: „*Ultraschall in Wissenschaft und Technik*“ Deutsch-Taschenbücher Band 46, Frankfurt am Main (1985)
- [SR02] SORGE George: „*Faszination Ultraschall*“ B.G. Teubner Verlag, (2002)
- [STU1] HORBACH U., PELLEGRINI R., FELDERHOFF U., THEILE G.: „*Ein virtueller Surround Sound Abhörraum im Ü-Wagen*“
- [STU2] „*Binaural Room Scanning (BRS) Virtuelle Reproduktion eines Regieraumes über Kopfhörer*“ Artikel zur Präsentation auf der AES-Convention Paris (2000)
- [SU80] SUTILOV Vladimir A.: „*Physik des Ultraschalls: Grundlagen*“ aus d. Russ. übers. u. hrsg. von Peter Hauptmann. Springer Verlag Wien, New York (1984)
- [TS99] TIETZE U., SCHENK Ch.: „*Halbleiter-Schaltungstechnik*“ 11.Auflage Springer-Verlag (1999)
- [WA73] WARFIELD D.: „*The study of hearing in animals. In: W Gay, ed., Methods of Animal Experimentation*“ IV. Academic Press, London, pp 43-143. (1973)
- [WA96] WANG L.-M., SHUNG K.K., CAMPS O.: „*Two Bit Correlation - An Adaptive Time Delay Estimation*“ IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol. 43 No. 3 (1996)

- [WE94] WEBSTER David: „*A pulsed ultrasonic distance measurment system based upon phase digitizing*“ IEEE Transactions on instrumentation and measurement, Vol. 43 Nr. 4 (1994)
- [WE97] WEHN Hans W., BÉLANGER Pierre R.: „*Ultrasound-Based Robot Position Estimation*“ IEEE Transactions on robotics and automation Vol.13 Nr. 5 (1997)
- [WU97] WÜRFEL Wolfgang: „*Passive akustische Lokalisation*“ Diplomarbeit, TU-Graz (1997)
- [WZ97] WENZEL Elizabeth M.: „*Analysis of the role of update rate and system latency in interactive virtual acoustic environments*“ 103. AES Convention, New York, Preprint 4633 (1997)
- [WZ01] WENZEL Elizabeth M.: „*Effect of increasing system latency on localisation of virtual sounds with short and long duration*“ Processing of the 2001 International Conference on Auditory Display, Espoo, Finland (2001)
- [YI96] YIN L., YANG R., GABBOUJ M., NEUVO Y.: „*Weighted Median Filters: A Tutorial*“, IEEE Trans. on Circuits and Systems, Vol. 43 Nr. 3 (1996)
- [ZI92] ZIMMERMANN Gerhard: „*Hochauflösende Ultraschallentfernungsmessung mit echtzeitfähiger, digitaler Signalverarbeitung für den industriellen Einsatz*“ Dissertation, TU-Wien (1992)
- [ZO93] ZOLLNER Manfred, ZWICKER Eberhard: „*Elektroakustik*“ Springer Verlag, 2.Auflage (1993)
- [ZO02] ZÖLZER Udo: „*DAFX: Digital Audio Effects*“ John Wiley & Sons (2002)
- [ZO04] ZOTKIN Dmitry, DURAI SWAMI Ramani: „*Accelerated Speech Source Localization via a Hierarchical Search of Steered Response Power*“ IEEE Transactions on Speech and Audio Processing, Vol. 12 Nr. 5 (2004)

Anhang A

Wichtige Symbole und Bezeichnungen

c :	Schallgeschwindigkeit
C_p :	spezifische Wärme bei konstantem Druck; in Luft $C_p = 0,24$
C_v :	spezifische Wärme bei konstantem Volumen; in Luft $C_v = 0,171$
D :	Laufzeitdifferenz
$\hat{D}$:	geschätzte Laufzeitdifferenz
D_{Samples} :	Laufzeitdifferenz in Samples
f :	Frequenz $[Hz]$
f_s :	Samplingfrequenz $[Hz]$
G_{x_1, x_2} :	Kreuzleistungsdichtespektrum zwischen den Signalen x_1 und x_2
G_{x_1, x_1} :	Autoleistungsdichtespektrum des Signals x_1
k :	Wellenzahl $k = \frac{\omega}{c} = \frac{2\pi f}{c} = \frac{2\pi}{\lambda} \left[\frac{1}{m} \right]$
p :	Druck
R_{x_1, x_2} :	Kreuzkorrelationsfunktion zwischen den Signalen x_1 und x_2
R_{x_1, x_1} :	Autokorrelationsfunktion des Signals x_1
T :	Temperatur $[^\circ C]$
T_s :	Samplingperiode $[sec]$
v :	Geschwindigkeit $\left[\frac{m}{s} \right]$
κ :	Adiabatexponent $\frac{C_p}{C_v}$; in Luft $\kappa = 1,4$
λ :	Wellenlänge $[m]$
ρ :	Dichte
ω :	Kreisfrequenz; $\omega = 2\pi f \ [Hz]$

Anhang B

Technische Daten der Sensoren

Name	Freq. [kHz]	Sensitivity [dB]	S.P.L. [dB]	Directivity [°]	Detectable Range [m]	Resolution [mm]	Max. Input [V _{pp}]
MA40S4R Receiver	40	-63 typ. (0dB=10V/Pa)	—	80	0.2 - 4	9	—
MA40S4S Sender	40	—	120 typ. (0dB=0.02mPa)	80	0.2 - 4	9	20 Cont. signal

Distance: 30cm, Sensivity: 0dB=1V_{RMS}/μbar, S.P.L.: 0dB=2 · 10⁻⁴ μbar, 1 μbar = 0.1Pa

Tab. B.1: Technische Daten der *Murata* Sensoren (dem Datenblatt entnommen)

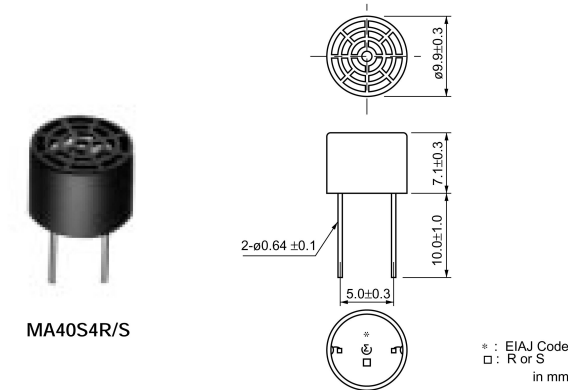


Abb. B.1: Abmessungen der *Murata*-Sensoren

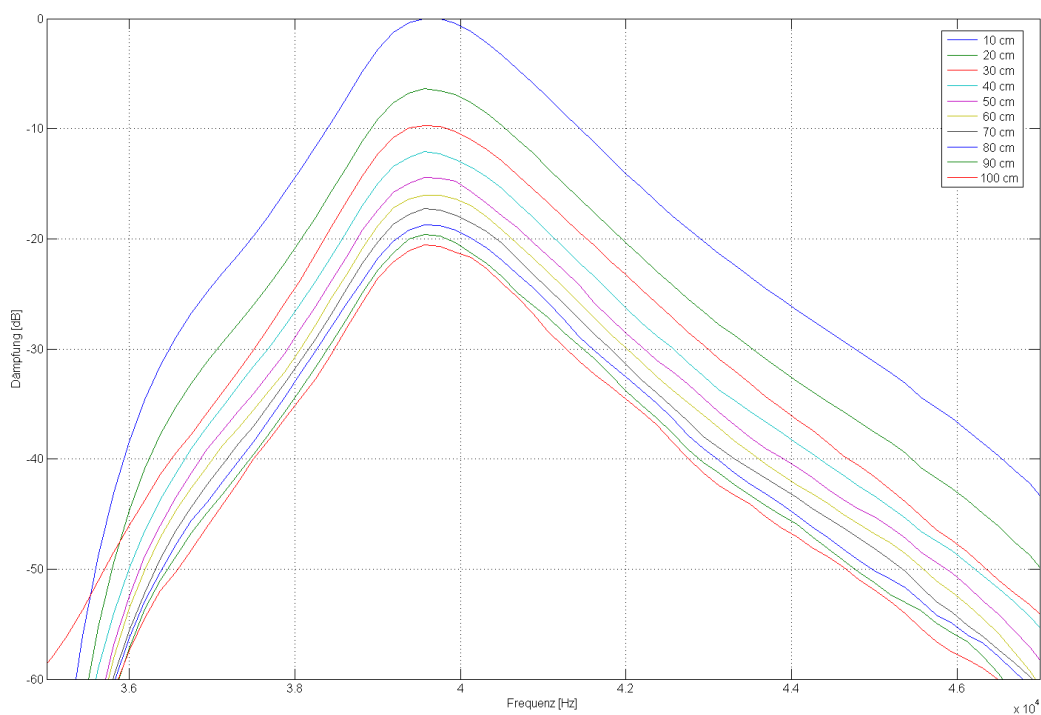


Abb. B.2: Frequenzgang der Sensorstrecke bei zueinander ausgerichteten Sensoren für verschiedene Distanzen (normiert)

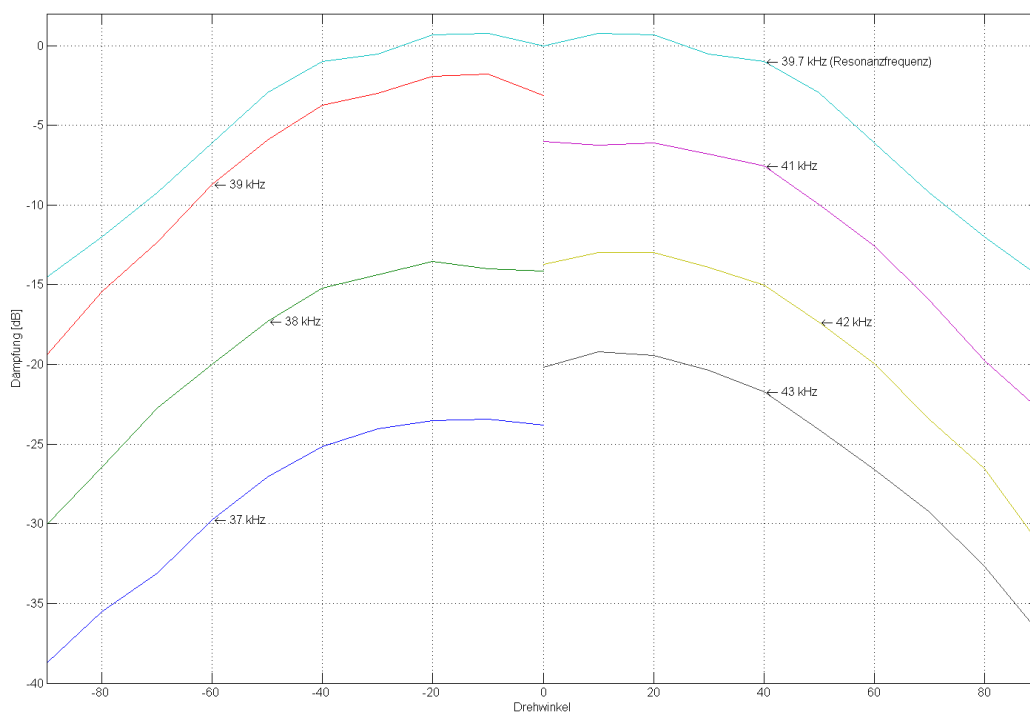


Abb. B.3: Richtcharakteristik der Sensorstrecke, Dämpfung bei Verdrehung der Sensoren zueinander für verschiedene Frequenzen (normiert)

Anhang C

Technische Daten des Referenzsystems *Ascension Flock of Birds*

Prinzip:	Aufbau eines elektromagnetischen Feldes. Positionsbestimmung durch Messung der Feldstärke am Empfänger. gepulste DC-Methode
Arbeitsbereich:	In einem Umkreis von 1,2 m vom Sender
Positionsauflösung:	8mm bei 31cm Distanz zum Sender
Statistische Positionsabweichung:	durchschnittlich 2.5mm RMS innerhalb Arbeitsbereich
Winkelbereich:	$\pm 180^\circ$ Azimuth und Rollen $\pm 90^\circ$ Elevation
Winkelauflösung:	0.1° RMS bei 31cm Distanz zum Sender
Statistische Winkelabweichung:	durchschnittlich 0.5° RMS innerhalb Arbeitsbereich
Update rate:	100 Hz
Gültigkeitsbereich:	Alle Angaben gültig für $30^\circ\text{C} \pm 10^\circ$ in einer metallfreien und elektromagnetisch ungestörten Umgebung

Anhang D

Hörbereiche von Haustieren

Anders als für den Menschen ist Ultraschall für viele Tierarten durchaus hörbar. Die vom Headtrackingsystem ausgesendeten Ultraschallsignale könnten also für bestimmte Tierarten irritierend und lästig empfunden werden. Die folgende Tabelle zeigt den ungefähren Hörbereich bei unterschiedlichen Spezies. Verwendet wurden dabei die Daten von Fay ([FA89]) und Warfield ([WA73]). Da leider keine genaueren Angaben zur Empfindlichkeit innerhalb des Hörbereiches vorliegen, sind die Werte nur als grobe Anhaltspunkte zu sehen. Da in den beiden Studien unterschiedliche experimentelle Methoden angewendet wurden, ist ein genauer Vergleich der einzelnen Tierarten untereinander nicht sinnvoll.

Spezies	ungefährer Hörbereich [Hz]
Frettchen	16 - 44.000
Goldfisch	20 - 3.000
Hase	360 - 2.000
Hund	67 - 45.000
Igel	250 - 45.000
Kanarienvogel	250 - 8.500
Katze	45 - 64.000
Maus	1.000 - 91.000
Ratte	200 - 76.000
Weißwal	1.000 - 123.000

Tab. D.1: Hörbereich von Haustieren ([FA89], [WA73])