

Helmut WITTEK [Institut für Rundfunktechnik, München, Germany]

## **Räumliche Wahrnehmung von virtuellen Quellen bei Wellenfeldsynthese**

### ***Spatial perception of virtual sources in Wavefield Synthesis***

Wellenfeldsynthese (WFS) ist eine Wiedergabetechnik, die es erlaubt, virtuelle Schallfelder mit Hilfe von Lautsprecherarrays entstehen zu lassen. Die Eigenschaften des virtuellen Schallfelds sind dabei dem Original sehr ähnlich, wenn auch nicht kongruent. Das wirft die Frage auf, welche Übereinstimmungen und Unterschiede in Bezug auf die auditive Wahrnehmung bestehen. Im Gegensatz zu konventionellen Wiedergabemethoden wie der Stereophonie bietet WFS allerdings erhebliche Vorteile speziell für die räumliche Qualität des reproduzierten Schallfelds. Bisherige und aktuelle Untersuchungen zur räumlichen Wahrnehmung von WFS werden präsentiert und die gefundenen Eigenschaften diskutiert.

#### **1. Einleitung**

Die Eigenschaften der Wellenfeldsynthese werden aus der Sicht der Wahrnehmung beleuchtet. Das heißt, dass im folgenden nicht die physikalischen Eigenschaften der Wellenfeldsynthese die dominierende Rolle spielen, sondern das Augenmerk denjenigen Eigenschaften zugewandt wird, die tatsächlich für die Wahrnehmung relevant sind. Über die physikalischen Eigenschaften der WFS ist in zahlreichen Publikationen berichtet worden (für eine umfassende Literaturübersicht siehe z.B.[5]), offene Fragen bleiben allerdings bei der Bewertung dieser Eigenschaften. Wie übersetzt sich z.B. die „Alias-Frequenz“ in Wahrnehmungs-bezogene Attribute wie Lokalisationsschärfe oder Klangfarbe? Diese zum Teil allerdings unbeantworteten Fragen werden im folgenden vorgestellt.

#### **2. Auswahl der geeigneten Wiedergabetechnik**

In den letzten Jahrzehnten ist es gelungen, die *klangliche* Qualität einer Tonaufnahme bis zur Perfektion zu bringen. So sind seit langem Geräte und Techniken vorhanden, um das klangli-

che Geschehen perfekt beim Hörer zu reproduzieren und es wird sogar nach weiterer Perfektionierung gesucht (z.B. Erhöhung der Sampling-Frequenz).

Bei der *räumlichen* Qualität der Wiedergabe liegt der Fall anders. Hier ist man womöglich erst am Anfang und es sind noch deutliche Verbesserungen möglich. Die bestehenden Wiedergabearten unterscheiden sich in punkto Klangqualität im Prinzip wenig, sie besitzen aber äußerst unterschiedliche Eigenschaften in Bezug auf die räumliche Qualität der Wiedergabe.

Die Übertragung der räumlichen Eigenschaften eines Klangbilds kann als eine Art Kodierung aufgefasst werden. Es gibt verschiedene Wege, den Ort der Schallquelle im Raum für eine Übertragung zu kodieren. Im Wiedergaberaum findet dann eine Synthese statt, die die Audiosignale möglichst adäquat mit ihren räumlichen Charakteristika zusammenführt. Die räumliche Synthese kann auf verschiedene Arten geschehen. Es gibt drei grundsätzlich unterschiedliche Methoden ([1]):

1. Lautsprecher-Stereofonie
2. Reproduktion der Ohrsignale („Kunstkopftechnik“)
3. Synthese des umgebenden Schallfeldes (z.B. Wellenfeldsynthese, Ambisonics)

Alle bekannten räumlichen Tonübertragungsverfahren lassen sich darauf zurückführen oder stellen Mischformen dar. Die drei genannten Wiedergabetechniken basieren auf völlig unterschiedlichen Wahrnehmungsmechanismen. So kann auch erklärt werden, dass für jede Technik andere Gesetze für die Qualität der Wahrnehmung räumlicher Parameter gelten. Jedes System hat auch spezifische Vor- und Nachteile – eine ganz allgemeine Bewertung der räumlichen Qualität ist ohne Kenntnis der speziellen Aufgabenstellung nicht möglich.

Die Frage, welche Technik für eine bestimmte Anwendung am besten geeignet ist, wird meistens nicht nur durch einen Vergleich der Qualität räumlicher Parameter entschieden. So sind auch grundsätzliche Systemeigenschaften und praktische Aspekte für den Erfolg eines Systems verantwortlich. Diese sind in Tabelle 1 beispielhaft skizziert. Der Vergleich der Qualität räumlicher Parameter (Kapitel 5.1) macht deutlich, dass Nachholbedarf auf wissenschaftlicher Seite in Bezug auf die Erforschung der Wellenfeldsynthese besteht. So sind die wahrnehmungsbezogenen Vorzüge der Wellenfeldsynthese noch nicht grundlegend untersucht und auf praktischer Seite noch zu wenig Erfahrungen vorhanden.

Tabelle 1 gibt eine Auswahl von Parametern wieder, die als Grundlage für die Bewertung der Systeme für eine bestimmte Anwendung dienen.

	<b>Stereofonie</b>	<b>Binaural- technik</b>	<b>Wellenfeld- synthese</b>
1. <i>Technischer Aufwand</i>	+++	+ / - *	---
2. <i>Einfluss des Wiedergabe- raums</i>	0	+++	0
3. <i>Größe der möglichen Hör- zone, Anzahl der Zuhörer</i>	+	-	+++
4. <i>Verfügbarkeit von Tools</i>	+++	+ / - - *	--
5. <i>Akzeptanz beim Hörer, Notwendigkeit von Kopf- hörern</i>	+++	-	++
6. <i>Qualität der Lokalisation</i>	siehe Kapitel 4		
7. <i>Räumliche Qualität</i>	siehe Kapitel 5		
8. <i>Klangliche Qualität</i>	0 - (Klangfarbenunterschiede sollten aber separat diskutiert werden!)		

**Tabelle 1: Praktische Aspekte zur Auswahl der geeigneten Wiedergabetechnik**

+ = gut, 0 = keine Bewertung, - = schlecht

Eine Vielzahl von unterschiedlichen Anwendungen ist denkbar, die jeweils unterschiedliche Anforderungen an die genannten Parameter stellen. Beispielsweise liegen die offensichtlichsten Vorzüge der Wellenfeldsynthese bei Parametern, die mit der Größe der Hörzone und der korrekten Wiedergabe der Quellenperspektive zu tun haben. Das heißt zum Beispiel, WFS wird dann besonders zu bevorzugen sein, wenn eine Bewegung der Hörer im Schallfeld möglich sein soll. Dies ist besonders in Situationen wie z.B. in Simulationsumgebungen oder Installationen in Vergnügungsparks gefragt. Im Gegensatz dazu lässt sich ein perfektes akustisches Abbild in allen Dimensionen am einfachsten durch die Binauraltechnik darstellen – sofern der Kopfhörer nicht stört und keine Bewegung des Hörers im Schallfeld erlaubt ist. Ist die realgetreue Abbildung eines Raums in einer großen Hörzone nicht der wichtigste Punkt, so ist die Stereofonie möglicherweise die beste Wahl – zumal der technische Aufwand minimal ist.

Eine bessere Kenntnis der Eigenschaften der Wellenfeldsynthese insbesondere bezüglich der räumlichen Parameter ist unverzichtbar, wenn ein Systemvergleich, wie er hier skizziert ist,

\* mit/ohne Headtracking

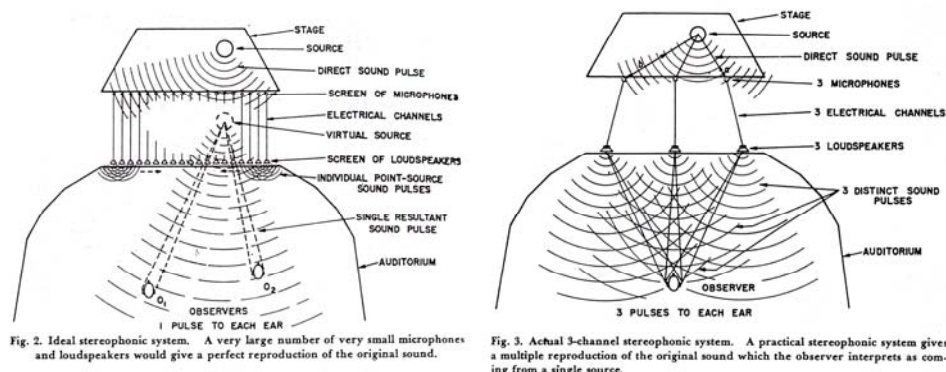
verlässliche Aussagen zu Tage bringen soll. Einige Untersuchungen dazu sind in den folgenden Kapiteln beschrieben.

Besondere Aufmerksamkeit gilt natürlich auch der Fähigkeit eines Systems, die Klangfarbe der Quelle zu reproduzieren. Da die Klangfarbe nur indirekt eine räumliche Eigenschaft des Schallfelds ist, wird diese Frage zunächst bewusst ausgeklammert. Separate Untersuchungen zu diesem Thema sind allerdings notwendig, da die Klangfarbe im Vergleich zu den räumlichen Eigenschaften eine oft entscheidende Rolle bei der Bewertung der allgemeinen Qualität des Schallfelds spielt. Die System-inhärenten Unterschiede bei der Wiedergabe der Klangfarbe stellen deshalb einen erheblichen Faktor dar.

### 3. Grundlagen der Wellenfeldsynthese

WFS basiert auf dem Huygen'schen Prinzip, das durch das sog. Kirchhoff-Helmholtz-Integral mathematisch beschrieben wird (siehe [1] und [3]). Es drückt aus, dass in jedem Punkt innerhalb eines quellenfreien Volumen der Schalldruck berechnet werden kann, sofern Schalldruck und -schnelle auf der das Volumen umgebenden Oberfläche bekannt sind. Daraus leitet sich das Wiedergabeprinzip ab: Sorgt man für eine entsprechende Schallwiedergabe auf der Oberfläche, so erhält man im Volumen das korrekte Schallfeld. Mit Abstrichen gilt dieses Prinzip auch, wenn man statt der kompletten Oberfläche eines Volumens nur den Umfang einer Fläche wiedergibt. Die Lautsprecher eines Lautsprecherarrays übernehmen diese Aufgabe.

Schon vor 70 Jahren hatten die Bell-Lab Ingenieure Steinberg, Snow und Fletcher ([4]) die Idee eines „akustischen Vorhangs“ (Abbildung 1 linkes Bild).



**Abbildung 1: Vorgänger von WFS: Erwünschtes (links) und implementiertes stereophones System (rechts) von Steinberg, Snow und Fletcher [4] aus dem Jahr 1934.**

In den späten 80er Jahren ersetzten die Forscher der TU Delft diese intuitive Konzept mit einer fundierten physikalischen Theorie und führten somit die Wellenfeldsynthese ein. Grund-

legende Arbeiten zum Thema sind z.B. [1],[3]. Eine Einführung in die Wellenfeldsynthese bietet z.B. [5]. Die Tonmeister-bezogene Anwendung der WFS erläutern [21],[23].

Zum Verständnis des Beitrags sind nun einige Grundbegriffe der WFS beschrieben:

#### *Virtuelle Quelle, Fokussierte Quelle*

Die durch WFS reproduzierte Quelle wird virtuelle Quelle genannt. Ihre Position kann hinter und auch vor dem Array sein. In letzterem Fall heißt die virtuelle Quelle auch fokussierte Quelle, da sich die Arraysignale an der Quellenposition fokussieren.

#### *Alias-Effekt, Alias-Frequenz (Spatial Aliasing, Spatial Aliasing Frequency $f_{\text{alias}}$ )*

Oberhalb einer gewissen Frequenz, der sog. Alias-Frequenz  $f_{\text{alias}}$ , kann das Schallfeld nicht mehr korrekt reproduziert werden. Es treten Fehler in der räumlichen Darstellung der Quellen sowie im Frequenzspektrum auf, die sog. Alias-Effekte (Spatial Aliasing). Die Alias-Frequenz ( $f_{\text{alias}}$ ) wird hauptsächlich durch den Abstand der Array-Lautsprecher sowie Quellen- und Hörerposition bestimmt.

#### *Truncation Effect, Diffraction Effect (Endeffekte, Beugungseffekte)*

In der Theorie wird ein unendlich langes Lautsprecher-Array vorausgesetzt. Da dies in der Praxis unmöglich ist, fehlen dem reproduzierten Schallfeld gewisse Anteile. Dies macht sich dadurch bemerkbar, dass von den Enden des Arrays räumlich fehlerhafte Signale ausgehen, die allerdings meist von geringem Pegel sind.

#### *Hörzone (Listening Area)*

WFS ist im Normalfall mit eindimensionalen Lautsprecherarrays realisiert, das heißt, in der Ebene innerhalb des geschlossenen Kreises wird das Schallfeld optimal synthetisiert. Diese Fläche, in der die Hörer das optimale räumliche Abbild wahrnehmen können, nennt sich Hörzone. Es ist auch möglich, in nur einer Richtung ein Array zu installieren. Die Hörzone reduziert sich entsprechend der Länge und Position des Arrays.

#### *WFS driving function*

Das Signal für die Arraylautsprecher ergibt sich aus der sog. WFS driving function. Sie besteht im Prinzip aus den Signalen, die Mikrofone an den Orten der Arraylautsprecher von der virtuellen Quelle aufnehmen würden. Dazu kommen WFS-Theorie-bedingte Pegel- und Frequenzkorrekturen.

## 4. Lokalisationseigenschaften

### 4.1 Grundprinzip der Lokalisation virtueller Quellen

Das Ziel der WFS ist es, eine korrekt lokalisierbare Wellenfront zu erzeugen. Dabei greift man auf das Grundprinzip zurück, nach dem die interauralen Zeitdifferenzen (ITD) für die Lokalisation entscheidend sind. Diese sind im unteren Frequenzbereich (bis ca. 2,5 kHz) wirksam. Im oberen Frequenzbereich sind interaurale Pegeldifferenzen (ILD) für die Lokalisation verantwortlich. In [6] zeigten Wightman und Kistler, dass selbst bei fehlerhaften ILD noch korrekte Lokalisation stattfindet, und wiesen somit die Dominanz der interauralen Zeitdifferenzen nach. Das bedeutet für WFS, dass die Frequenzbereiche, in denen ITD ausgewertet werden, korrekt erzeugt werden müssen. Die Grenze für die Auswertung von ITD ist individuell unterschiedlich sowie signalabhängig und liegt bei ca. 2,5 bis 5 kHz. Wo bei WFS die Obergrenze des korrekt zu reproduzierenden Frequenzbereichs liegt, ist daraus nicht genau abzuleiten. Ebenso wird aus dieser Untersuchung nicht klar, welche Qualitätseinbußen bei fehlerhafter Wiedergabe im oberen Frequenzbereich zu erwarten sind.

### 4.2 Lokalisation und Alias-Frequenz

Eines der ersten „WFS-ähnlichen“ Systeme wurde von Vogel ([7]) im Jahr 1993 zu Experimenten genutzt. Sein lineares Array bestand aus 12 Lautsprechern, die in einem Abstand von je 45cm zueinander aufgestellt waren. Die theoretische Alias-Frequenz dieses Systems wäre  $f_{\text{alias}}=380$  Hz. Vogel stellte fest, dass das erzeugte Wellenfeld durchaus die gewünschte Richtungsinformation enthielt. Er folgerte, dass die Alias-Effekte diese Information nicht zerstörten. Anscheinend ist also schon bei sehr niedrigen Alias-Frequenzen korrekte Lokalisation möglich. Dies beruht allerdings auch wesentlich auf zwei weiteren Faktoren:

- Der Präzedenz-Effekt („Gesetz der ersten Wellenfront“) hilft bei der Lokalisation. Dies ist in Abbildung 1 (linkes Bild) zu sehen: Die erste Wellenfront, die beim Hörer ankommt, trifft aus der Richtung der virtuellen Quelle ein, da der Schall hier die kürzeste Weglänge hat.
- Auch über der Alias-Frequenz ist die korrekte Richtungsinformation im virtuellen Schallfeld noch enthalten. Man kann den Alias-Effekt auch als Überlagerung zwischen Signalanteilen mit korrekter und mit inkorrekt er Richtungsinformation verstehen. Es ist wahrscheinlich, dass auch in Bereichen oberhalb der Alias-Frequenz noch korrekte Lokalisation erreicht wird.

Allerdings bewirkt der Alias-Effekt wohl Qualitätseinbußen bei der Lokalisation (als auch bei der Klangfarbe) der virtuellen Quelle. Diese Qualitätseinbußen machen sich dadurch bemerkbar, dass die virtuelle Quelle im Vergleich zu einer realen Quelle schlechter und diffuser lokalisierbar ist. Dies wurde in der Vergangenheit mit verschiedenen Methoden gemessen, wie im folgenden beschrieben.

#### **4.3 Lokalisationsqualität, Lokalisationsschärfe, Lokalisiertheit**

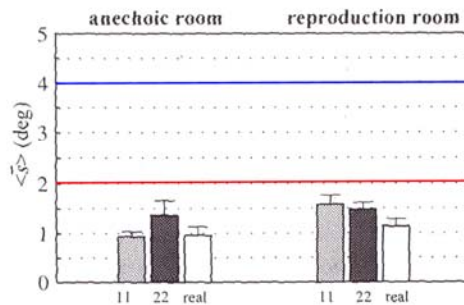
In Experimenten von Vogel ([7]), Start ([8]) und Verheijen ([9]) wird die Streuung der gesammelten subjektiven Richtungsdaten als Maß für die Lokalisationsqualität angesehen. Die Versuchspersonen sollten in mehreren Durchgängen einer virtuellen Quelle eine bestimmte Hörereignisrichtung zuordnen. Die Streuung zwischen den Messwerten aus den verschiedenen Durchgängen und der Versuchspersonen wird mit der Standardabweichung des Mittelwerts gemessen. Vogel ließ die virtuellen Quellen (Lautsprecherabstand  $\Delta x=12\text{cm}$ ) sowohl Breitband-Rauschen als auch Tiefpass-Signale ( $<1,5\text{kHz}$ ) wiedergeben. Start analysierte diese Experimente und fand keinen signifikanten Unterschied zwischen den Standardabweichungen der Mittelwerte der beiden Testsignale (jeweils  $s=1,0^\circ$ ). Da die Alias-Frequenz etwa  $1,5\text{ kHz}$  betrug, folgerte er, dass die Alias-Effekte, die nur im Breitbandsignal vorhanden waren, die Lokalisationsqualität nicht beeinflussten.

Zur Lokalisationsqualität bei Sprache als Testsignal merkte schon Vogel an, dass bestimmte Artefakte, die bei Breitbandrauschen auftreten, verschwinden. So seien z.B. bei seinem WFS-System mit  $f_{\text{alias}}=380\text{ Hz}$  die hoch- und tieffrequenten Anteile nur bei Breitbandrauschen getrennt lokalisierbar. Er führt dies auf die gemeinsame Hüllkurve aller Frequenzen im Sprachsignal zurück.

Mithilfe von Versuchen mit Kunstkopfaufnahmen verglich Start reale Quellen mit virtuellen Quellen. In seinen Experimenten gab es zwischen einem WFS-System mit  $f_{\text{alias}}=1,5\text{ kHz}$  und einer realen Quelle keinen messbaren Unterschied in Bezug auf den kleinsten wahrnehmbaren Winkel bei Richtungsänderungen. Da in einem analog durchgeführten Versuch mit verdoppeltem Lautsprecherabstand und halbiertem Alias-Frequenz durchaus Unterschiede messbar waren, folgerte er, dass  $1,5\text{ kHz}$  als Minimum für die Alias-Frequenz ausreichen würden.

Sowohl Start als auch Verheijen fanden bei Versuchen im reflektionsarmen Raum ebenso keinen Unterschied zwischen realen und virtuellen Quellen der Alias-Frequenz  $f_{\text{alias}}=1,5\text{ kHz}$ . Allerdings fanden auch beide, dass in echten Räumen das WFS-System erheblich schlechter als die natürlichen Quellen abschnitt.

Abbildung 2 zeigt die Ergebnisse Verheijens:



**Abbildung 2: Ergebnisse der Experimente Verheijens: Die Blöcke zeigen die Standardabweichung des Mittelwerts bei wiederholten Messungen der Richtung verschiedener Quellen. Sie Standardabweichung wird als Maß der Streuung und demnach als Maß für die Lokalisationsqualität interpretiert.**

**Graue Blöcke: virtuelle Quellen  $\Delta x=11$  cm,  $f_{\text{alias}}=1500$  Hz.**

**Schwarze Blöcke: virtuelle Quellen  $\Delta x=22$  cm,  $f_{\text{alias}}=750$  Hz.**

**Weißer Blöcke: reale Quellen.**

**Virtuelle Quellen hinter dem Array,**

**Messungen aus zwei verschiedenen Räumen (siehe Beschriftung)**

Diese Ergebnisse deuten darauf hin, dass WFS in reflektiven Umgebungen schlechtere Lokalisationseigenschaften besitzt. Allerdings gilt das Gegenteil für die Wahrnehmung der Klangfarbe: Hier sind im reflektionsarmen Raum wesentlich mehr Artefakte hörbar, die im normalen Raum unhörbar bleiben. Der Grund dafür dürfte darin zu suchen sein, dass die Arraylautsprecher im Raum Reflektionen erzeugen, die in Pegel und Zeit nicht zur virtuellen Quelle passen. Dies verursacht sowohl negative Einflüsse auf die Lokalisation wie auch eine Nivellierung von störenden Klangfarbeneffekten.

#### **4.4 Ein Hybrid-System zur Vermeidung von Alias-Effekten: OPSI**

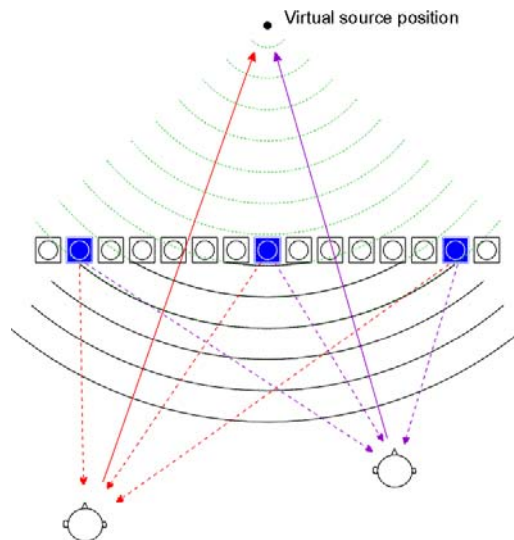
##### *4.4.1 Grundprinzip*

Eine fehlerlose Reproduktion des virtuellen Schallfelds ist nur bis zur sog. Alias-Frequenz möglich. In diesem Kapitel wird eine Lösung beschrieben, bei der störende Alias-Effekte trotzdem komplett vermieden werden.

Die grundlegende Idee besteht darin, oberhalb der Alias-Frequenz ganz auf Wellenfeldsynthese zu verzichten. Stattdessen wird eine Technik verwendet, die bekanntermaßen wenig oder keine störenden Klangfarbeneffekte erzeugt, nämlich die Stereophonie. Oberhalb der Alias-Frequenz hat WFS keine entscheidenden Vorteile mehr gegenüber der Stereophonie, da die Alias-Effekte sowohl Klangfarbe als auch Lokalisation beeinträchtigen. Insofern liegt es nahe, in diesem Bereich die Stereophonie einzusetzen.

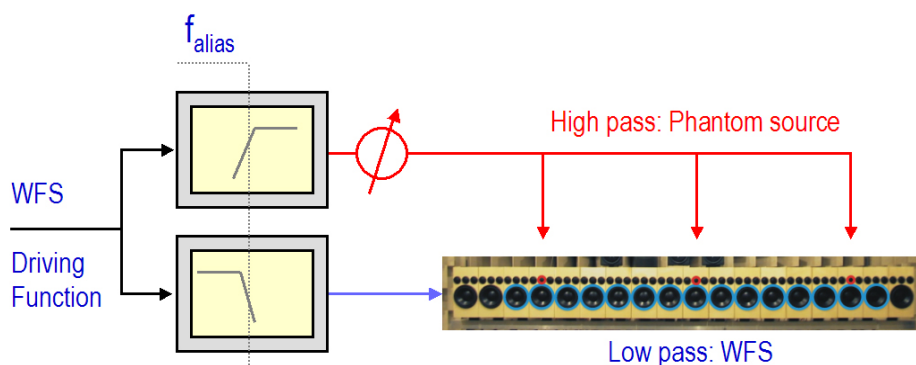


Dieses hybride Prinzip ist in Abbildung 3 illustriert. Es wird OPSI (=“*Optimised Phantom Source Imaging of the high frequency content of virtual sources in Wave Field Synthesis*”) genannt. Es besteht aus einer virtuellen WFS-Quelle, die den Frequenzbereich unterhalb  $f_{\text{alias}}$  wiedergibt, und einer Phantomschallquelle für den Frequenzbereich oberhalb  $f_{\text{alias}}$ :



**Abbildung 3: Grundprinzip von OPSI:**  
Hybrid-Lösung aus einer virtuellen WFS-Quelle (schwarze Wellenlinien) unterhalb  $f_{\text{alias}}$  und einer stereofonen Phantomschallquelle (rote und lila Ausbreitungspfeile) oberhalb  $f_{\text{alias}}$ .

Das Resultat ist ein gemeinsames Hörereignis, zu dem virtuelle Quelle und Phantomschallquelle verschmelzen. Die Position der gemeinsam wahrgenommenen Quelle wird durch die tieffrequente WFS-Quelle bestimmt. Es ist also im Vergleich zur konventionellen WFS-Quelle kein Nachteil in Bezug auf die Lokalisation zu erwarten. Ein Hörversuch (siehe Kapitel 4.5) prüft diese Annahme.

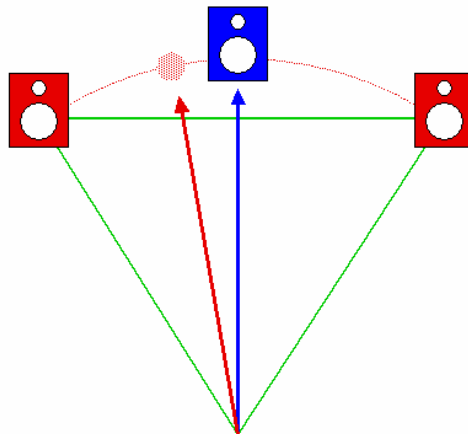


**Abbildung 4: Erzeugung von OPSI-Signalen:**  
Das WFS-Array wird durch die Tiefpass-gefilterten Anteile gespeist. Die Stereo-Lautsprecher werden durch die Hochpass-gefilterten Anteile nach einer Pegelanpassung gespeist.

Die Erzeugung von OPSI-Signalen erläutert Abbildung 4. Das normale WFS-Signal, das aus der WFS driving function gewonnen wird, wird nach einem Tiefpass (mit  $f_{\text{grenz}} \cong f_{\text{alias}}$ ) dem Array zugeführt (blauer Signalweg). Das stereofone Signal (roter Signalweg) wird aus einer Hochpass-Filterung (ebenso mit  $f_{\text{grenz}} \cong f_{\text{alias}}$ ) des WFS-Signals gewonnen. Bestimmte einzelne Lautsprecher (Stereo-Lautsprecher), die auch aus dem Array selbst gewählt werden können, werden nach einer Pegelanpassung mit diesen Signalen gespeist.

#### 4.4.2 Der „OPSI-Lokalisationsfehler“

Die virtuelle WFS-Quelle und die stereofone Phantomschallquelle verschmelzen bei der Wahrnehmung zu einem Hörereignis. Um dies zu erreichen, darf allerdings der Unterschied zwischen den Hörereignisrichtungen der einzelnen Quellen nicht zu groß werden. Ist dies der Fall, würden die Einzelquellen separat wahrgenommen werden oder das gemeinsame Hörereignis hätte eine wesentlich größere und diffusere Ausdehnung. Der Unterschied der Einzelrichtungen wird „OPSI-Lokalisationsfehler“ genannt. Um negative Effekte zu vermeiden, sollte der OPSI-Lokalisationsfehler geringer sein als der „maximale OPSI-Lokalisationsfehler“. Dieser „maximale OPSI-Lokalisationsfehler“ wurde in einem Experiment ermittelt (Abbildung 5).



**Abbildung 5:**

**Experiment zu Ermittlung des maximal erlaubten OPSI-Lokalisationsfehlers:**

Eine Tiefpass-gefilterte Einzelquelle wurde im mittleren (blauen) Lautsprecher wiedergegeben. Sie simulierte die virtuelle WFS-Quelle. Gleichzeitig wurde eine Hochpass-gefilterte Phantomschallquelle zwischen den roten Lautsprechern erzeugt. Im Hörversuch schalteten die Teilnehmer zwischen einer Referenz und dem Testsignal um.

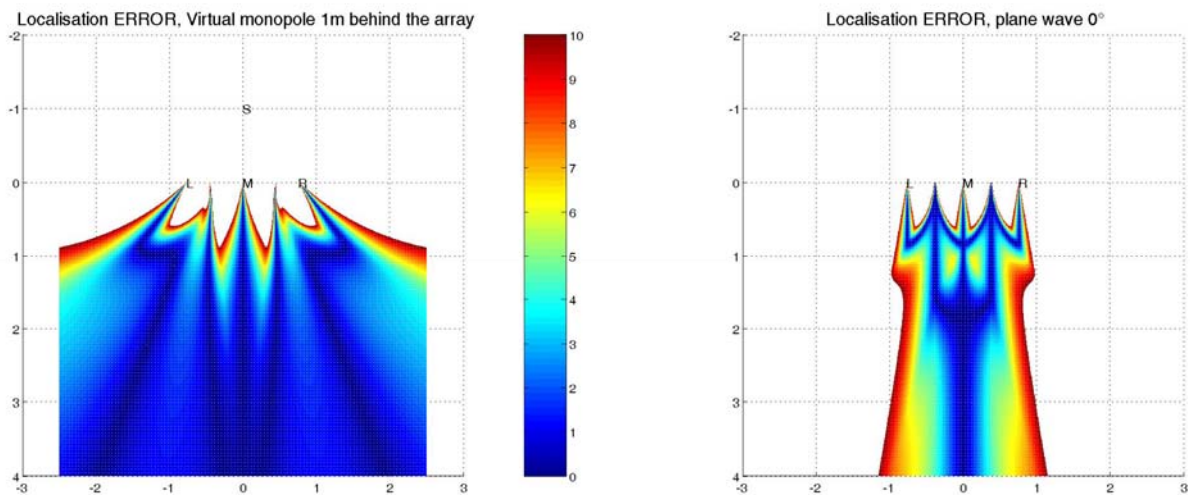
Referenz und Testsignal unterschieden sich nur in Bezug auf den OPSI-Lokalisationsfehler (Unterschied der Einzelrichtungen).

Dabei stellte sich heraus, dass eine Abweichung von mehr als  $5^\circ$  eine hörbare Veränderung des Klangbilds (insbesondere Hörereignisrichtung) zur Folge hat. Ein OPSI-Lokalisationsfeh-

ler von  $5^\circ$  sollte deshalb als kritische Grenze angesehen werden, der in der Hörzone nicht überschritten werden sollte.

In Simulationen kann der OPSI-Lokalisationsfehler in der Hörzone dargestellt werden. Wie in früheren Untersuchungen vorgestellt wurde (siehe „Image Assistant“, [10], [11]) kann die stereofone Abbildung mit psycho-akustischen Messdaten errechnet werden. Aus bekannten Daten sowie neu gewonnenen Messdaten, die speziell für Hochpass-gefilterte Signale gelten, wurde die Richtung der Phantomschallquelle an jedem Punkt errechnet.

Der OPSI-Lokalisationsfehler ist von Quellenposition, Arraygeometrie und Anzahl und Position der Stereo-Lautsprecher abhängig. Für zwei verschiedene Situationen ist der simulierte OPSI-Lokalisationsfehler in Abbildung 6 und Abbildung 7 dargestellt. Die Abbildungen zeigen einen Hörbereich von etwa  $4 \times 5$  m. Das WFS-Array befindet sich auf der Geraden  $y=0$ . Die Stereo-Lautsprecher L, M und R (Stereo-Basis je 0.75m) erzeugen die Phantomschallquelle, die virtuelle Quelle ist an der Position S. Die farbliche Skala gibt den Lokalisationsfehler an, wobei blau für einen optimal geringen Lokalisationsfehler von  $0^\circ$  steht, grün genau die kritische Grenze von  $5^\circ$  markiert und der rote bzw. weiße Bereich einen zu hohen Lokalisationsfehler beschreibt.



**Simulation des OPSI Lokalisationsfehlers:**  
**Die Farben beziehen sich auf die Grad-Skala in der Mitte. Der kritische OPSI-Lokalisationsfehler von  $5^\circ$  ist in den grünen Flächen erreicht.**  
**Angezeigt sind Quelle (S) und Stereo-Lautsprecher (L,M,R). Das WFS-Array befindet sich auf der Geraden  $y=0$ .**

**Abbildung 6:**  
**virtuelle Quelle ist 1m hinter dem Array ( $x=0, y=-1$ )**

**Abbildung 7:**  
**virtuelle ebene Welle, Quelle ist unendlich weit hinter dem Array ( $x=0, y=-\infty$ )**

Abbildung 6 zeigt, dass für diese Quelle in großen Bereichen der OPSI-Lokalisationsfehler unter dem kritischen Wert bleibt. Auch für die ebene Welle, die in Abbildung 7 simuliert ist,

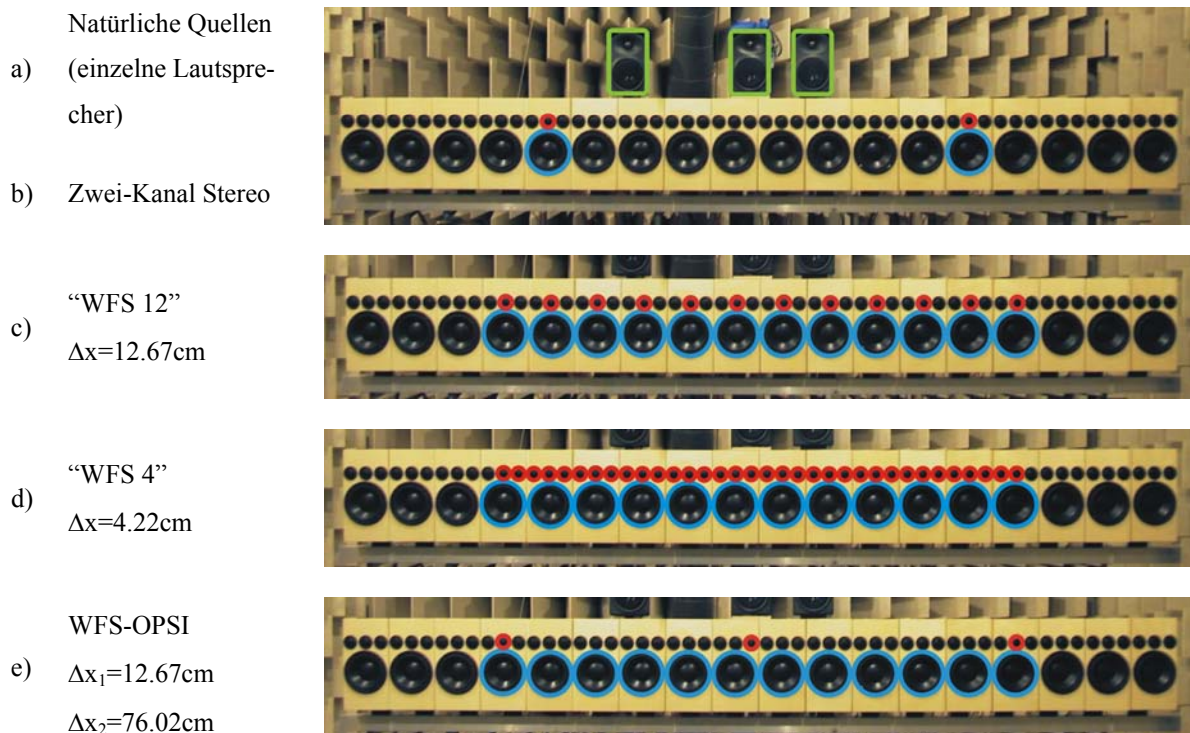
ergibt sich ein Bereich, der in der Breite in etwa den Abmessungen der Stereo-Basis entspricht. Somit lässt sich auch für diesen Fall der OPSI-Lokalisationsfehler in der gesamten Hörzone minimieren, wenn statt nur drei Lautsprechern ein Netz aus Stereo-Lautsprechern angesteuert wird.

## 4.5 Vergleich verschiedener WFS-Arrays

### 4.5.1 Versuchsaufbau

Das OPSI-System und zwei verschiedene WFS-Arrays, sowie normale Phantomschallquellen wurden in einem Hörversuch mit natürlichen Quellen (einzelne Lautsprechern) verglichen, um Rückschlüsse auf die erzielte Qualität der Lokalisation bei den verschiedenen Systeme zu erhalten (für detaillierte Beschreibungen des Experiments siehe [10]).

Abbildung 8 zeigt die Systeme, die am Vergleich beteiligt waren:

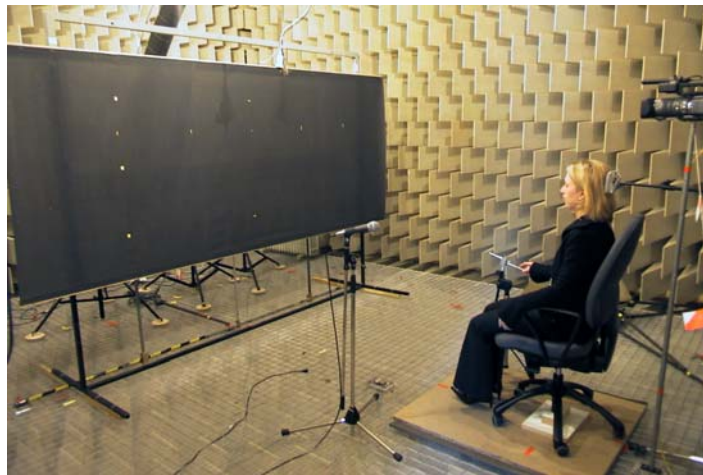


**Abbildung 8: Alle Systeme des Experiments:**

- a) Natürliche Quellen
- b) Zwei-Kanal Stereophonie
- c) „WFS 12“ mit  $\Delta x=12.67\text{cm}$
- d) „WFS 4“ mit  $\Delta x=4.22\text{cm}$
- e) WFS-OPSI siehe vorangegangenes Kapitel 4.4

Die „Natürlichen Quellen“ bestanden aus aktiven Lautsprechern Klein&Hummel O100. Sie waren an den drei Positionen aufgestellt, die auch mit den WFS-Systeme synthetisiert wur-

den, wie Abbildung 10 zeigt. In Abbildung 8 sind zudem die bei den anderen Systemen jeweilig aktiven Lautsprecher aus dem Array farblich markiert. Die verwendeten Arraylautsprecher waren selbstgebaute 2-Weg-Lautsprecher, in deren Gehäuse je 3 Hochtöner Platz fanden. Dadurch waren WFS-Systeme mit den Lautsprecherabständen  $\Delta x=12.67\text{cm}$  („WFS 12“) und sogar  $\Delta x=4.22\text{cm}$  („WFS 4“) realisierbar. Die 18 Versuchspersonen saßen auf einem Stuhl im reflektionsarmen Raum vor einer undurchsichtigen, akustisch transparenten Leinwand (siehe Abbildung 9). Hinter der Leinwand waren die verschiedenen Systeme angeordnet (siehe auch Abbildung 8 und Abbildung 10).



**Abbildung 9: Versuchsdurchführung: Der Hörereignisort wurde mit einem Laserpointer angepeilt.**

Es wurden jeweils 3 Rauschpulse (Rosa Rauschen, Breitband) von je 1s Länge auf einem der Systeme reproduziert. Die Aufgabe der Versuchsteilnehmer bestand darin, mit einem Laserpointer die wahrgenommene Hörereignisrichtung anzupeilen. Die Ergebnisse wurden vom Versuchsleiter abgelesen und gesammelt.

Außerdem waren die Versuchsteilnehmer angehalten, im Anschluss daran ein Urteil über die Sicherheit abzugeben, mit der sie die Quelle lokalisiert hatten. Dieser Parameter, der dem Attribut Lokalisiertheit entspricht, wurde den Teilnehmern zuvor mündlich erklärt. Folgende Skala stand für die Beurteilung der Lokalisiertheit zur Verfügung:

*Wie einfach war es für sie, die Quelle zu lokalisieren? Wie gut lässt sich eine bestimmte Richtung dem Hörereignis zuordnen?*

*1 – sehr gut*

*2 – gut*

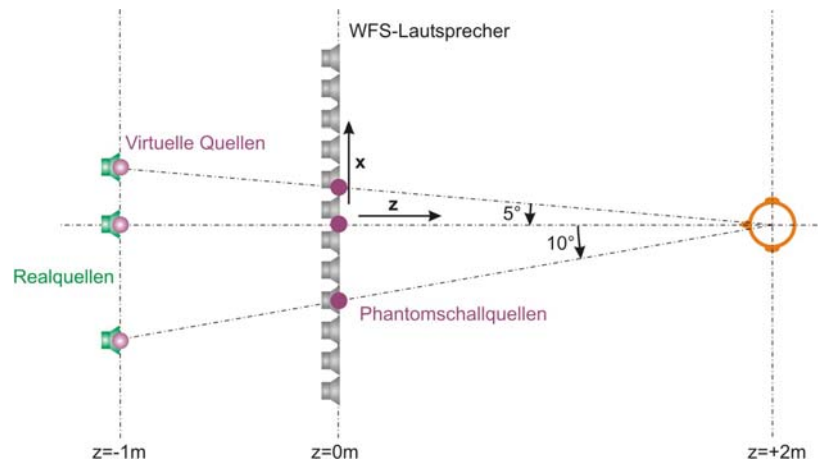
*3 – mittelmäßig*

*4 – schlecht*

*5 – sehr schlecht*



Die reproduzierten Quellen waren in drei verschiedenen Richtungen angeordnet, wie Abbildung 10 zeigt. Jedes Testsignal wurde im Verlauf des Hörversuchs viermal wiederholt.



**Abbildung 10: Versuchsaufbau im reflektionsarmen Raum:**  
Der Kopf des Versuchsteilnehmers ist in oranger Farbe am rechten Bildrand skizziert.

- grün: Natürliche Quellen (einzelne Lautsprecher)
- grau: WFS Array – die virtuellen Quellen befinden sich genau an den Positionen der natürlichen Quellen
- lila: Positionen der stereofonen Phantomschallquellen

#### 4.5.2 Ergebnisse

In Abbildung 11 sind Scatterplots der Versuchsergebnisse zu sehen, die maßstabsgetreu vor einem Foto des Versuchsaufbaus abgebildet sind. Es ist deutlich zu sehen, dass die Lokalisation trockener Quellen im reflektionsarmen Raum sogar für natürliche Quellen (einzelne Lautsprecher) nicht einfach ist.

Die Standardabweichungen der Mittelwerte, in Kapitel 4.3 als Maß für die Lokalisationsqualität benützt, sind in Abbildung 12 und Abbildung 13 angegeben. Auf der x-Achse sind die verschiedenen Systeme angegeben („Real“ steht für die natürlichen Quellen, „PSQ“ für die stereofonen Phantomschallquellen). Die drei Quellenrichtungen  $-10^\circ$ ,  $0^\circ$  und  $5^\circ$  sind separat in verschiedenen Farben angegeben. Es zeigt sich, dass die Streuungen erheblich von der jeweiligen Quellenrichtung abhängen. Die Unterschiede zwischen den Systemen sind größtenteils nicht signifikant, lediglich bei den Phantomschallquellen können deutlich höhere Streuungen festgestellt werden, was auf eine erhöhte Ausdehnung oder eine geringere Lokalisiertheit deuten könnte.

Auffallend ist, dass es erhebliche Unterschiede zwischen den Versuchsteilnehmern in Bezug auf die wahrgenommenen Hörereignisrichtungen gab. Dies erhöht den festgestellten mittleren Fehler. Allerdings sind die Versuchsteilnehmer bei ihren Angaben konsistent, zeigen also wenig Abweichungen zwischen den Durchgängen. Dies ist in Abbildung 14 und

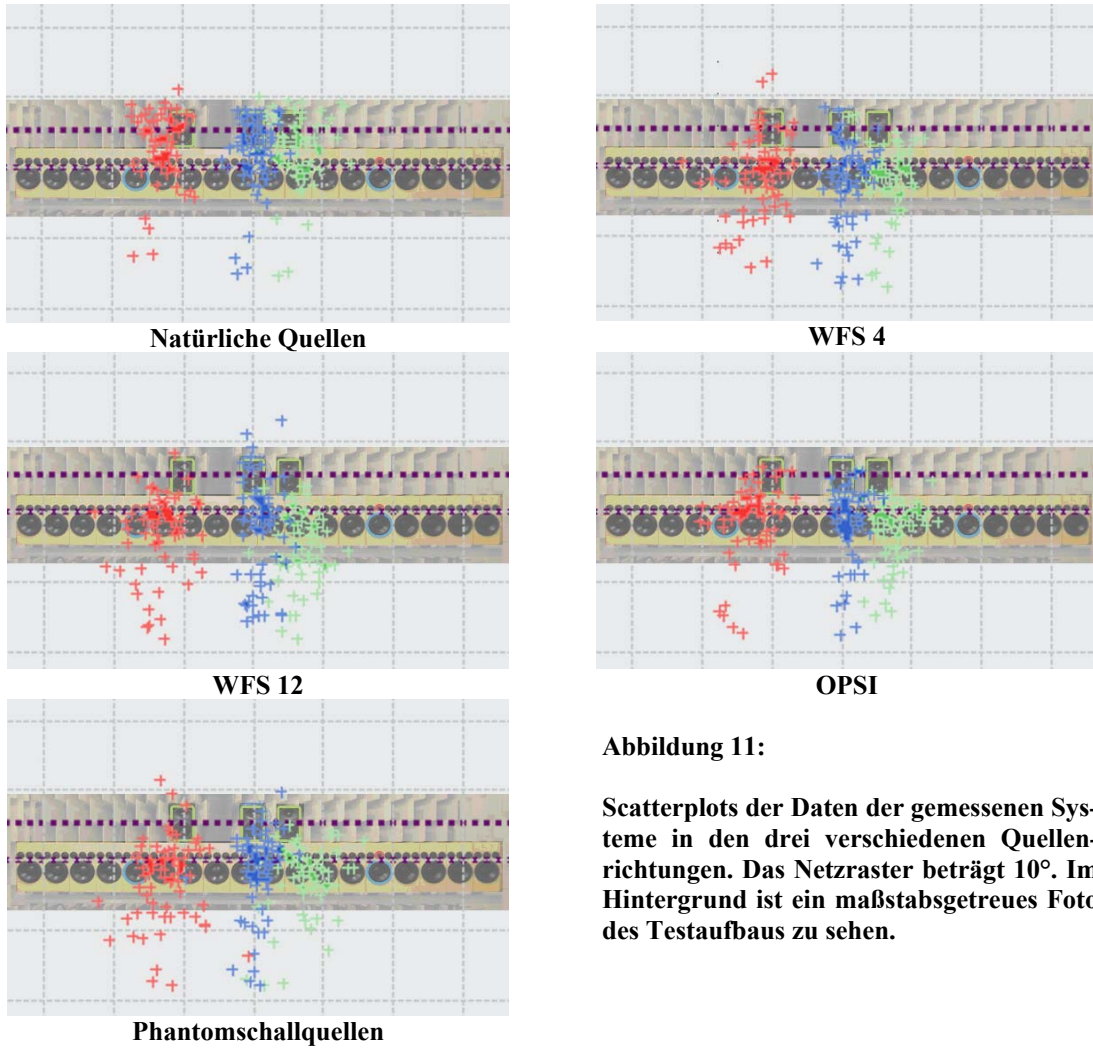


Abbildung 11:

Scatterplots der Daten der gemessenen Systeme in den drei verschiedenen Quellenrichtungen. Das Netzraster beträgt  $10^\circ$ . Im Hintergrund ist ein maßstabsgereutes Foto des Testaufbaus zu sehen.

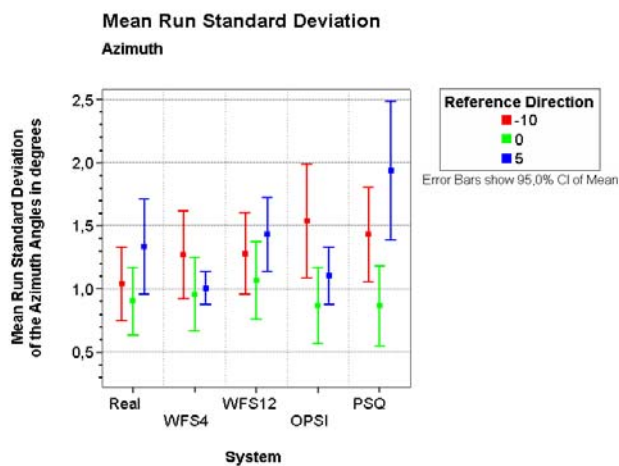


Abbildung 12: Azimuth  
Mittlere Standardabweichungen in Grad (und 95% Konfidenzintervall) der wahrgenommenen horizontalen Richtungen

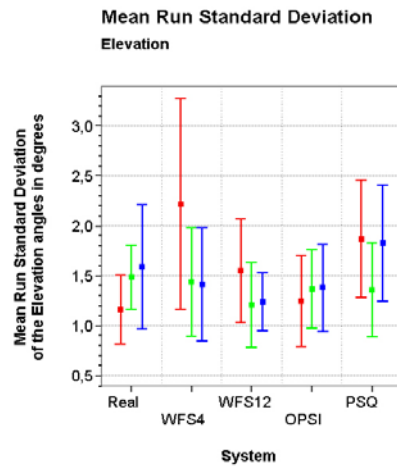
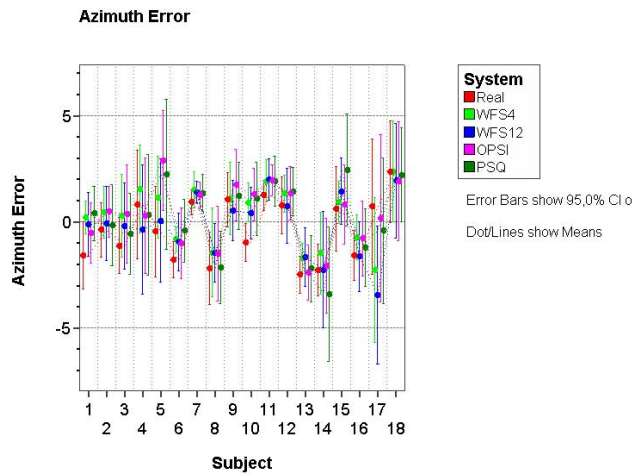
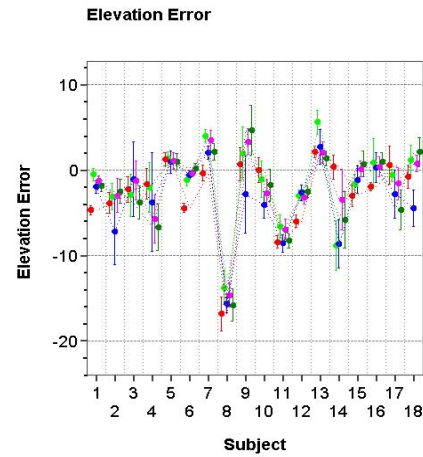


Abbildung 13: Elevation  
Mittlere Standardabweichungen in Grad (und 95% Konfidenzintervall) der wahrgenommenen vertikalen Richtungen

Abbildung 15 zu sehen. Hier sind die mittleren individuellen Azimuth- und Elevationsfehler für jede Versuchsperson angegeben, das heißt die Abweichungen von der tatsächlich wiedergegeben Richtung in horizontaler (Azimuth) und vertikaler (Elevation) Richtung.



**Abbildung 14:**  
**Mittlerer Individueller Azimuth-Fehler**



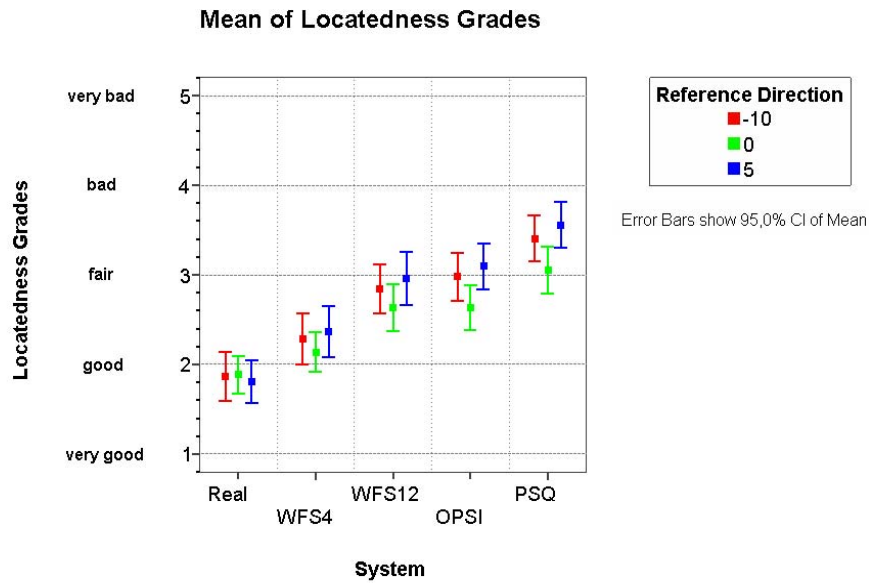
**Abbildung 15:**  
**Mittlerer Individueller Elevations-Fehler**

Es liegt nahe, dass diese Messungen wenig Rückschluss auf Attribute wie Lokalisierungsschärfe oder Lokalisiertheit zulassen. Dazu diene die zusätzliche Frage nach der Lokalisiertheit. Die Ergebnisse dieser Messung zeigt Abbildung 16.

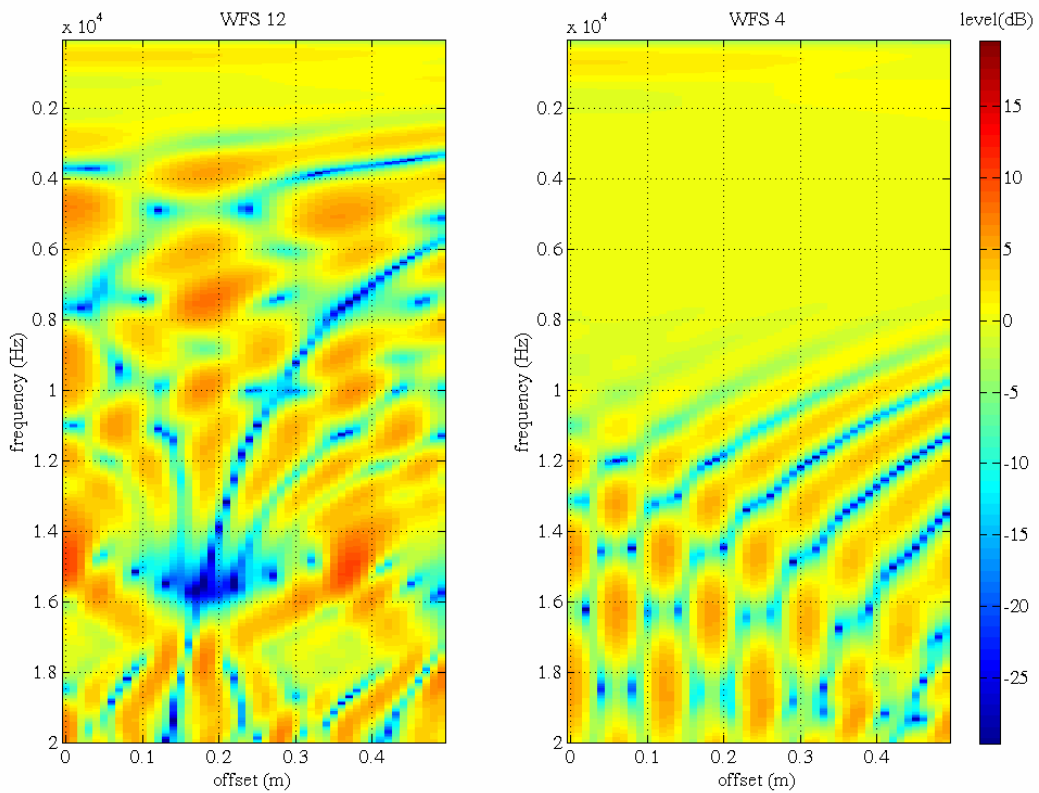
Die Ergebnisse zeigen signifikante Unterschiede zwischen den Systemen und ermöglichen folgende Schlussfolgerungen:

1. Die Lokalisiertheit der natürlichen Quellen ist allen anderen Systemen überlegen. Auch das WFS-System mit hoher Alias-Frequenz kann diese gute Beurteilung nicht erreichen. Die Beurteilung der WFS-Systeme ist möglicherweise im realen Raum noch etwas schlechter, da hier die Array-Reflexionen die Lokalisierung behindern. Im Gegensatz dazu wird die natürliche Quelle im realen Raum wohl eher besser lokalisiert.
2. Die Lokalisiertheit des „WFS 4“-Systems ist signifikant besser als die des „WFS 12“-Systems. Das bedeutet, dass auch eine Erhöhung der Alias-Frequenz jenseits von 2,5 kHz noch positive Einflüsse auf die Lokalisiertheit hat. Eine Analyse dazu ist in Abbildung 17 zu sehen.
3. Die OPSI-Quellen werden nicht schlechter lokalisiert als die vergleichbaren Quellen des „WFS 12“-Systems. Das belegt die Wirksamkeit des OPSI-Prinzips, das in Bezug auf die Klangfarbenwahrnehmung sogar Vorteile zeigen könnte.
4. Die reine Phantomschallquelle wird von allen Quellen am schlechtesten lokalisiert.





**Abbildung 16: Subjektive Beurteilung der Lokalisiertheit.**  
 Die Versuchspersonen beurteilten die Lokalisiertheit auf einer 5-stufigen Skala zwischen 1 (sehr gut) und 5 (sehr schlecht).  
 Parameter: Quellenrichtung (-10°,0°,5°).  
 Die Grafik zeigt Mittelwerte mit 95%-Konfidenzintervall.



**Abbildung 17: Frequenzspektren einer virtuellen WFS-Quelle (1m hinter dem Array), gemessen auf einer Linie von Mikrofonpositionen im Abstand von 2m zum Array.**

- links: „WFS 12“  $\Delta x=12.67\text{cm}$ ;  $\rightarrow f_{\text{alias}} \approx 2.5 \text{ kHz}$ .
- rechts: „WFS 4“  $\Delta x=4.22\text{cm}$ ;  $\rightarrow f_{\text{alias}} \approx 7.5 \text{ kHz}$ .

Zur Erklärung der wahrgenommenen Unterschiede zwischen den Systemen „WFS 4“ und „WFS 12“ sind in der folgenden Abbildung 17 Simulationen der Alias-Effekte der zwei Systeme dargestellt: Die Grafiken zeigen die Frequenzspektren auf einer Linie von Mikrofonpositionen parallel zum Array im Abstand von 2m zum Array. Diese erheblichen spektralen Veränderungen treten bei einer seitlichen Bewegung des Kopfes beim Hörer auf.

Die Alias-Frequenzen der Systeme sind ca. 2,5 kHz (links, „WFS 12“) und 7,5 kHz (rechts, „WFS 4“). Deutlich zu sehen ist auch, dass sich die Alias-Frequenz und die Gestalt der Alias-Effekte schon bei kleinen Bewegungen verschiebt. Diese Verschiebungen sind zumindestens im reflektionsarmen Raum hörbar, bei kritischen Signalen auch im normalen Raum.

## 5. Räumliche Eigenschaften

### 5.1 Unterschiede zwischen den Wiedergabeprinzipien

In Kapitel 2 wurde bereits ein allgemeiner Vergleich zwischen den Wiedergabearten angestellt. Nun soll versucht werden, näher auf die räumlichen Eigenschaften der verschiedenen Wiedergabeprinzipien einzugehen. Tabelle 2 gibt einen groben Überblick über einige räumliche Parameter, die bei der Bewertung der Systeme eine Rolle spielen.

Wie in Kapitel 4 beschrieben, sind die Vorteile der WFS gegenüber der Stereophonie in Bezug auf die Lokalisation von Schallquellen (Punkt 4 in Tabelle 2) signifikant. Zur weiteren Analyse ist es sinnvoll, die Merkmale zunächst grob in zwei Gruppen zu unterteilen:

- a) räumliche Eigenschaften des Schallfelds an einem festen Hörort
- b) räumliche Eigenschaften des Schallfelds für einen sich bewegenden Hörer (und/oder für mehrere sich an unterschiedlichen Orten befindliche Hörer)

Die zu b) gehörenden Eigenschaften wurden in Tabelle 2 mit gelber Farbe hinterlegt. Schon an dieser groben und nicht auf Vollständigkeit beharrenden Tabelle wird klar, dass gerade diese Eigenschaften für WFS einzigartig sind. Gerade diese Eigenschaften machen WFS für bestimmte Anwendungen überlegen. Ist also eine Hörsituation gegeben, bei der es wesentlich ist, dass sich der/die Hörer im Schallfeld bewegen und eine realgetreue Perspektive erfahren kann, ist WFS die richtige Wahl. Die Perspektive, die WFS erzeugen kann, wird auch „*intra*-aktiv“ genannt. Diese Eigenschaft bezeichnet eine Szene, in der die Wahrnehmung der Bewegung der agierenden Person folgt. Im Gegensatz dazu folgt die *inter*-aktive Szene bestimmten Aktionen der agierenden Person. Eine *intra*-aktive Verschiebung der Perspektive durch eine Bewegung des Hörers ist in Abbildung 18 dargestellt.

	Stereofonie	Binaural-technik	Wellenfeld-synthese
1) Wiedergabe in allen 3 Dimensionen	-	+++	0
2) Präsenz, Umhüllung	+ ...	++	+ ... ?
3) Raumeindruck	+	+++	+ ... ?
4) Lokalisation (Schärfe, Fokus), siehe Kap. 4	+ ... -	- / +++ *	++
5) Klangfarbe	++	0 / ++ *	+ ... ?
6) Präsenz durch „Self-Motion“-Cues	-	-	+++
7) (Intra-aktive) Perspektive	-	-	+++
8) Distanz	+	+++	++
9) Darstellung von Quellen vor den Lautsprechern	-	0	+ ... ?

Tabelle 2: Praktische Aspekte zur Auswahl der geeigneten Wiedergabetechnik

+ = gut, 0 = keine Bewertung, - = schlecht

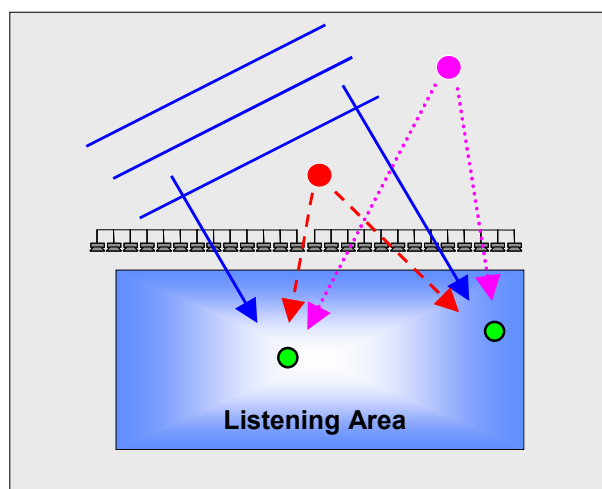


Abbildung 18: WFS an unterschiedlichen Hörpositionen (grüne Punkte): WFS produziert stabile Punktquellen (rot, pink) und ebene Wellen (blau). Die Perspektive ist abhängig von der Position des Zuhörers. (aus [21])

\* mit/ohne Headtracking

Ähnliches gilt, wenn eine Wiedergabe für mehrere Hörer gedacht ist. Wird hier realgetreue Perspektive verlangt (z.B. bei Beschallung im Theater), kommen die Eigenschaften der WFS zum Tragen. Einen Spezialfall nimmt eine Situation ein, in der jedem Hörer in einer größeren Hörzone das *selbe* Klangbild angeboten werden soll (z.B. im Kino). Hier wird bewusst die Darstellung der Perspektive vermieden, das Hauptaugenmerk liegt auf möglichst konstantem Pegel der Schallquellen innerhalb der gesamten Hörzone. Dies kann mit Erfolg durch Lautsprecharrays erreicht werden – WFS hat freilich mehr Potential.

Ein weiterer, potentiell entscheidender Vorteil der Wellenfeldsynthese ist in obiger Tabelle 2 in Punkt 8) und 9) aufgeführt und blau hinterlegt. Eine wesentliche Verbesserung der Wahrnehmung der *Distanz* von virtuellen Quellen gegenüber der Stereophonie führt zu einem deutlichen Gewinn an räumlicher Qualität. Dies gilt vor allem, wenn es gelingt, Quellen *vor* den Lautsprechern darzustellen. Dies wird in den nächsten Kapiteln diskutiert.

## 5.2 Räumliche Tiefe

Rumsey definiert räumliche Tiefe mit dem „Gefühl für die Perspektive in der wiedergegebenen akustischen Szene“ ([13]). Die Vermutung liegt nahe, dass die optimale Wiedergabe räumlicher Perspektive auch zu optimaler Wahrnehmung räumlicher Tiefe führt. Allerdings sollte man auch hier zwischen dem Hörer am festen Hörort und dem sich bewegenden Hörer unterscheiden.

Eine Analogie zur visuellen Tiefenwahrnehmung ist sinnvoll: Man unterscheidet hier zwischen 2D-Wiedergabe (flaches Bild ohne Tiefe), 2½D-Wiedergabe (Perspektive wird erlangt durch Bewegung des Beobachters) und 3D-Wiedergabe (echte Tiefe des Bildes durch Stereoskopie). Das heißt, eine visuelle Szene kann Tiefe erlangen entweder durch die realgetreue Perspektive, die sich durch Beobachterbewegungen erschließen lässt oder durch intuitive Merkmale wie einem stereoskopischem Bild. Man trennt bewusst zwischen 2½D- und 3D-Wiedergabe, da echte, intuitiv erlebte Tiefe nur im zweiten Fall zu erlangen ist.

Im akustischen Zusammenhang lässt sich ein ähnlicher Zusammenhang definieren: Echte, intuitiv erlebte Tiefe entsteht durch Merkmale, die sich wohl auch am festen Hörort erleben lassen. Dies sind die Merkmale der reflektiven Umgebung wie Reflektionsmuster und Nachhall. Beweis dafür ist die Tatsache, dass die Binauraltechnik und auch die Mehrkanalstereophonie es vermögen, Tiefe herzustellen. Im Gegensatz dazu lässt ein Klangbild, das solche Merkmale nicht besitzt, eine Wahrnehmung echter Tiefe vermissen, auch wenn die Perspektive erfolgreich erschlossen wird.

Nur Wellenfeldsynthese ist in der Lage, beides zu liefern und somit in der Lage, echte Tiefe mit intra-aktiver Perspektive zu vereinigen.

### 5.3 Merkmale zur Distanzwahrnehmung

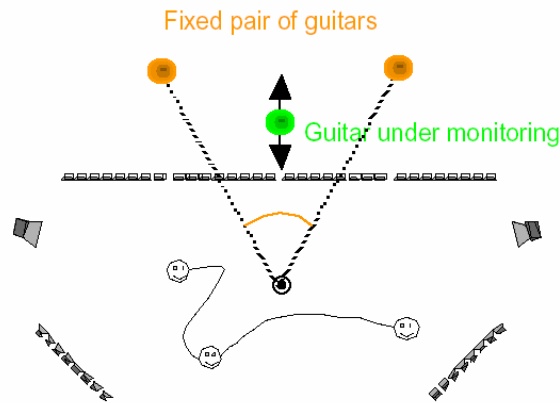
Die Wahrnehmung der Distanz virtueller Quellen beruht auf einer Reihe auswertbaren Merkmalen (siehe z.B. [17]). Diese sind in folgender Tabelle 3 dargestellt. Außerdem ist in dieser Tabelle analysiert, welche Merkmale von den drei Wiedergabesystemen Monofonie, Stereophonie und WFS dargestellt werden können.

Merkmale zur Distanzwahrnehmung	<i>Mono</i>	<i>Stereo</i>	<i>WFS</i>
Lautstärke, Direkt-/ Diffusschallverhältnis, spektrale Merkmale, etc.	✓	✓	✓
Reflektionsmuster	✗	✓	✓
(Intra-aktive) Perspektive	✗	✗	✓
Binaurale Merkmale – Krümmung der Wellenfront	✗	✗	✓

**Tabelle 3: Wiedergabe von räumlichen Attributen zur Distanzwahrnehmung auf bestimmten Wiedergabesystemen**

Einige Merkmale für die Distanzwahrnehmung sind, wie der Tabelle zu entnehmen ist, auch in monofoner Wiedergabe darstellbar. Untersuchungen zeigen z.B., dass allein durch die Veränderung der Lautstärke eine Quelle eine Veränderung bei der Distanzwahrnehmung zu erreichen ist ([18]).

WFS ist den konventionellen Wiedergabesystemen vor allem überlegen, wenn es um die Darstellung der Perspektive geht. Die Untersuchungen zur intra-aktiven Perspektive in der Literatur beweisen, dass implizite Distanzwahrnehmung durch eine Auswertung der relativen Perspektive möglich ist. Echte Distanzwahrnehmung, analog zur echten Wahrnehmung von Tiefe, ist in Abwesenheit der jeweiligen Reflektionsmuster nicht möglich. So untersuchen Noguès et al. ([14]), ob in Anwesenheit zweier „Ankerquellen“ die Distanz einer justierbaren Quelle korrekt einstellbar war. Abbildung 19 zeigt die festen Ankerquellen (orange) und die ihrer Distanz variierbare Quelle (grün). Die variable Quelle veränderte nur ihre Position, nicht aber alle anderen positions-relevanten Parameter wie relative Lautstärke am Referenzort, Verhältnis Direkt-/ Diffusschall etc. . Die Aufgabe für die Versuchspersonen bestand darin, die Distanz aller drei Quellen in Übereinstimmung zu bringen. Es stellte sich heraus, dass es möglich war, sofern sich die Versuchspersonen im Raum bewegen konnten, nicht aber bei festen Hörposition in der Mitte der Hörzone.



**Abbildung 19: Experiment von Noguès et al. [14]:**  
**Nur durch die Bewegung der Hörer konnte die Distanz der grünen Quelle**  
**mit der der orangen Quellen in Übereinstimmung gebracht werden.**

Ähnliches lässt sich aus der Untersuchung von Usher et al. ([15]) folgern. Die Versuchspersonen konnten sich in einem begrenzten Bereich bewegen konnten (siehe Abbildung 20, Versuchspersonen befinden sich innerhalb des mit „inner curtain“ bezeichneten Kreises).

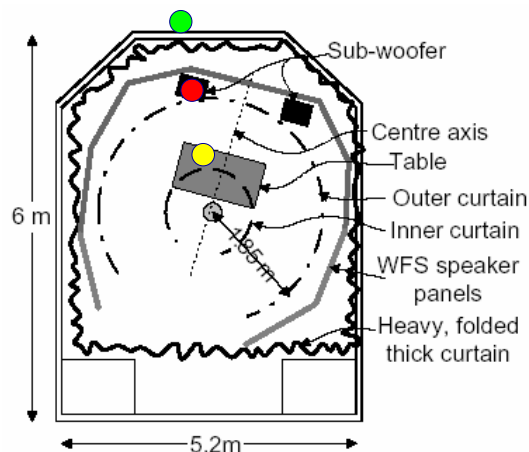


Figure 1: Plan view of the experimental set-up.

**Abbildung 20: Experiment von Usher et al. [15]:**  
**Die Versuchsperson, in der Mitte im „Inner Curtain“ befindlich, soll die relative**  
**Distanz der drei (vom Verfasser farblich markierten) virtuellen Quellen ermitteln.**  
**Dies gelingt nur für die nahe (gelbe) Quelle durch die Wahrnehmung der perspektivischen**  
**Veränderungen und der Lautstärke-Veränderungen bei Bewegung innerhalb des erlaubten Bereichs.**

Drei verschiedene virtuelle Quellen kamen zur Wiedergabe, die in Abbildung 20 mit einem gelben, roten und grünen Kreis markiert sind – die Quellen enthielten keinerlei räumliche Merkmale bis auf ihre Position im Schallfeld. Aufgabe war es, in paarweisen Vergleichen die jeweils näher positionierte Quelle herauszufinden. Dies gelang nur für die sehr nahe, gelb markierte virtuelle Quelle. Auch hier war es für die Versuchspersonen möglich, implizit die Perspektive zur Ermittlung der Quellendistanz zu benutzen. Außerdem kommt ein weiteres Merkmal des virtuellen WFS-Schallfelds zum Tragen: Nähert man sich einer virtuellen Quel-

le, so wird die wahrgenommene Lautstärke dementsprechend größer. Auch dies lässt sich zur indirekten Distanzermittlung nutzen.

#### 5.4 Distanzwahrnehmung bei nahen Quellen

In Tabelle 3 sind als letzter Punkt zur Distanzwahrnehmung die binauralen Merkmale aufgeführt. Diese sind unter gewissen Voraussetzungen in der Lage, Distanzwahrnehmung zu erzeugen. Die binauralen Merkmale entstehen durch die Krümmung der Wellenfront, die, wie Abbildung 21 zeigt, vom Abstand (Radius) des Hörers zur Quelle abhängt.

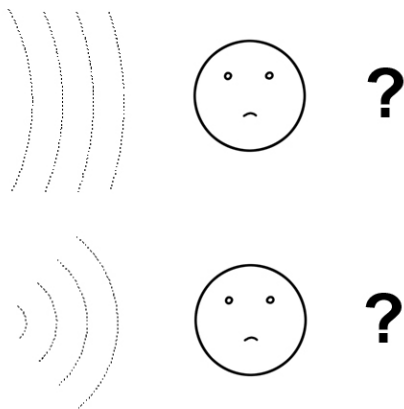


Abbildung 21:

Ist eine Distanzwahrnehmung nur aufgrund der Krümmung der Wellenfront erreichbar?

Dies gilt allerdings nur für sehr nahe Quellen. Schon die vorher zitierten Untersuchungen zur WFS zeigen, dass eine Distanz von Quellen, die weiter als ca. 1m entfernt sind, nicht allein aufgrund der Krümmung der Wellenfront wahrnehmbar ist. Die Distanz von Quellen in weniger als 1m Entfernung kann hingegen durchaus erkannt werden, auch wenn keine weiteren Merkmale vorhanden sind (siehe z.B. [19] und [20]). Vor allem tieffrequente interaurale Pegeldifferenzen bis 3500 Hz, die bei weiter entfernten Quellen weniger stark auftreten, sind für diese Phänomen verantwortlich.

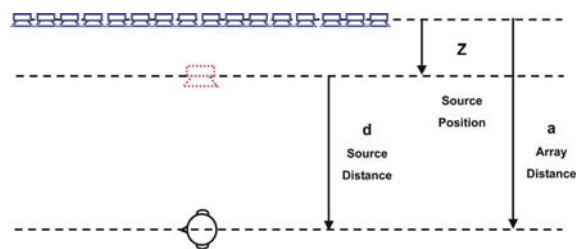
Im Prinzip ist es durch WFS möglich, Quellen im Raum zu platzieren. Diese sog. *fokussierten Quellen* ermöglichen eine kleine Distanz zwischen Quelle und Hörer. Dementsprechend könnten binaurale Merkmale der Distanzwahrnehmung zumindest helfen. Es wäre es durchaus von Vorteil, wenn binaurale Merkmale so stark wirken, dass eine nahe Distanzwahrnehmung zustande kommt, obwohl andere Merkmale fehlerhaft sind. Bei fokussierten Quellen ist es z.B. generell ein Problem, dass vom Lautsprecherarray fehlerhafte Reflektionsmuster ausgehen, die bei der Wahrnehmung der korrekten (geringeren) Distanz stören.

#### 5.5 Distanzwahrnehmung fokussierter WFS-Quellen

Eine Untersuchung zu diesem Thema wurde jüngst am IRT durchgeführt ([16]), welche hier in kurzer Form präsentiert sei. In theoretischer wie praktischer Hinsicht wurde die Natur der

Distanzwahrnehmung naher, fokussierter Quellen bei WFS beleuchtet. Die Fragestellung der Untersuchung war, ob fokussierte Quellen bei WFS aufgrund ihrer Position, das heißt durch die im Vergleich zu entfernteren Quellen gekrümmte Wellenfront eine nahe Distanzwahrnehmung bewirken können. Weitere Merkmale zur Distanzwahrnehmung wurden deshalb ausgeblendet, weil sie im speziellen Fall der fokussierten Quelle entweder fehlerhaft reproduziert werden (Reflektionsmuster) oder keinen Unterschied zu konventionellen Wiedergabemethoden darstellen (Lautstärke).

Abbildung 22 zeigt die Geometrie von Hörversuch und Simulationen im reflektionsarmen Raum:



**Abbildung 22: Array-Quelle-Hörer Geometrie von Wittek ([16]):**  
Die (virtuelle) Quelle (rot gepunktet) ist auf der Ohrachse platziert.

Der Hörer wurde quer zum Array platziert, damit die wichtigen binauralen Merkmale maximal auftraten. Die verschiedenen Quellen - sowohl echte Lautsprecher als Referenz als auch virtuelle, fokussierte WFS-Quellen - wurden somit auf der Ohrachse rechts vom Hörer in verschiedenen Distanzen von 0,25m bis 1,90m platziert. Mithilfe einer speziellen Anordnung (siehe Abbildung 23) wurde die subjektiv wahrgenommene Distanz der jeweiligen Quellen ermittelt:



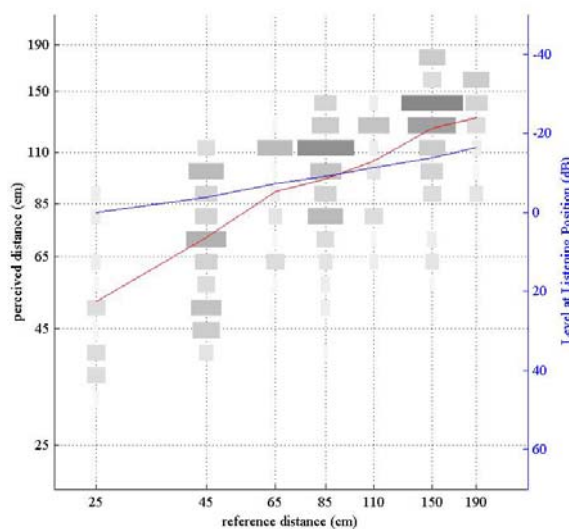
**Abbildung 23:**

**Die Versuchsperson bringt die Entfernung eines beweglichen, stummen Hilfslautsprechers in Deckung mit der subjektiv wahrgenommenen Entfernung der Testquelle (hinter dem akustisch transparenten Vorhang). Der Versuchsleiter kann daraufhin die jeweilige Entfernung ablesen.**

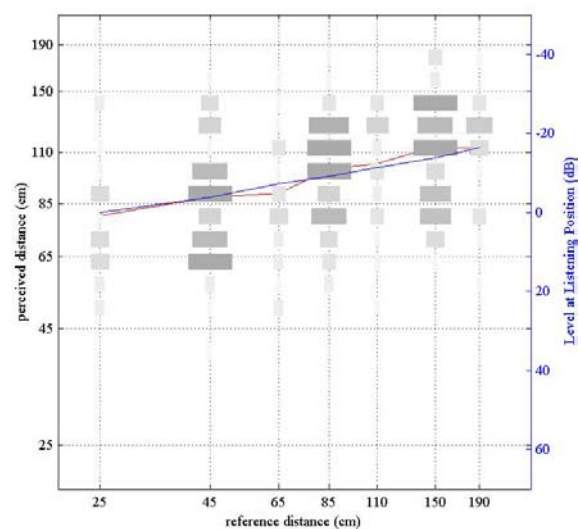


Das Experiment kam zu folgenden Ergebnissen:

Die Abbildungen zeigen die subjektiv wahrgenommene Distanz der jeweiligen Quellen (y-Achse linke) in Abhängigkeit von der tatsächlich eingestellten Distanz (x-Achse). Der Pegel des Lautsprechers wurde konstant gehalten, was bedeutet, dass der resultierende Pegel am Hörort (aufgrund des  $1/r$ -Gesetzes) mit steigender Entfernung abnimmt. Der Pegel am Hörort ist in den Abbildungen auf der blauen y-Achse auf der rechten Seite angetragen. Die rote Kurve zeigt den Mittelwert aus allen Messdaten, außerdem ist die Verteilung der Einzelergebnisse in einem Histogramm in die Abbildungen eingefügt (graue Boxen).



**Abbildung 24: Natürliche Quellen,  
natürliche Lautstärke**



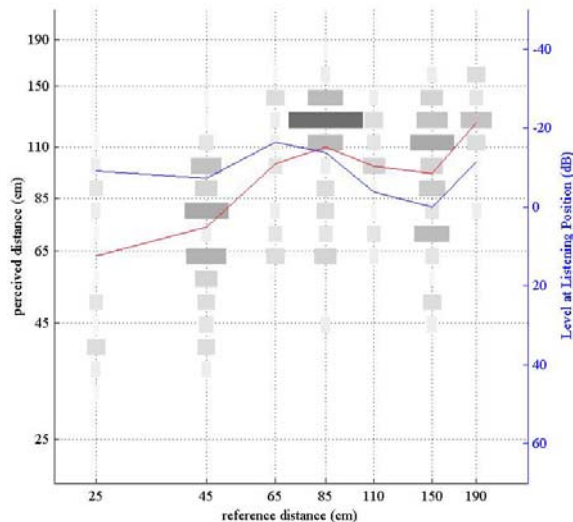
**Abbildung 25: Virtuelle Quellen,  
natürliche Lautstärke**

Abbildung 24 zeigt die Ergebnisse für die natürlichen Quellen (Lautsprecher). Es ist zu erkennen, dass die Distanz der Quellen einen deutlichen Einfluss auf die wahrgenommene Distanz hat. Es kommt zur typischen Überschätzung der Distanz für Abstände  $<1\text{m}$  und zu leichter Unterschätzung für Abstände  $>1\text{m}$ .

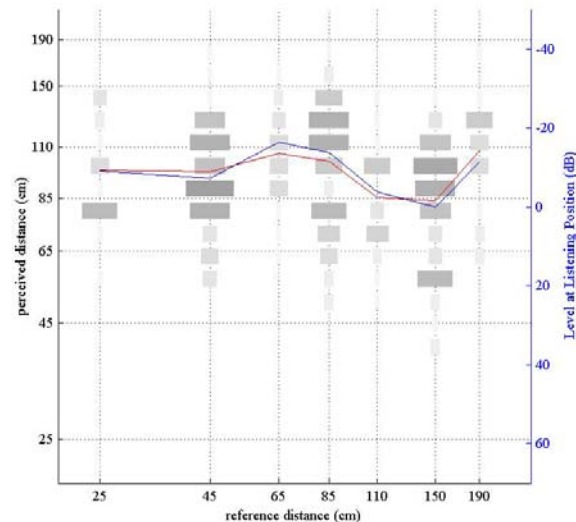
Die Ergebnisse für die virtuellen Quellen sind ähnlich. Abbildung 25 zeigt, dass die Kurve lediglich flacher verläuft, die tatsächliche Position der Quelle also weniger Einfluss auf das Resultat hat.

Dieses Ergebnis allein ist allerdings wenig aussagekräftig. Der Verdacht liegt nahe, dass alleine die jeweilige Lautstärke der Quellen die gemessenen Distanzunterschiede verursachte. Deshalb bediente man sich einer speziellen Methodik, um die Lautstärke aus dem Ergebnis zu eliminieren: Die Lautstärke der Quellen wurde zufällig variiert und hatte somit keinen direkten Bezug mehr zur tatsächlichen Entfernung der Quellen. Nun lagen also zwei Merkmale

vor, die in Konflikt miteinander standen („conflicting cues“): Die Lautstärke und die Krümmung der Wellenfront würden zu jeweils anderen Wahrnehmungen der Distanz führen. Das gemessene Ergebnis würde nun Aufschluss darüber geben können, ob nur die Lautstärke allein zur Distanzwahrnehmung verantwortlich ist oder nicht. Die Versuchspersonen wurden in einem Vorversuch instruiert und trainiert, das Merkmal Lautstärke soweit wie möglich außer Acht zu lassen.



**Abbildung 26:** Natürliche Quellen,  
zufällige Lautstärke

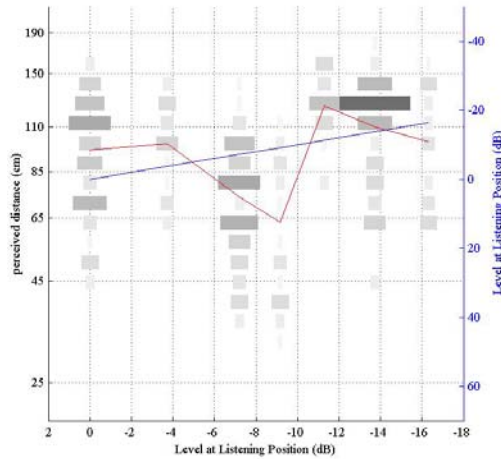


**Abbildung 27:** Virtuelle Quellen,  
zufällige Lautstärke

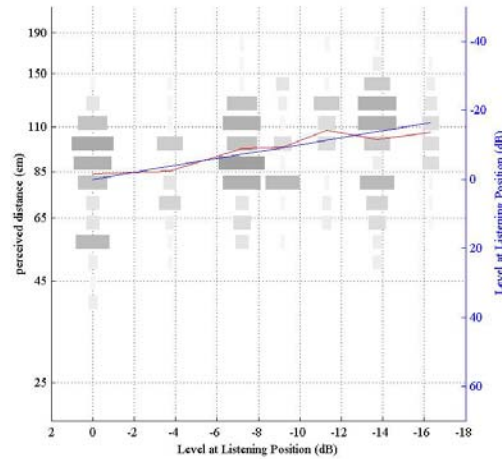
Abbildung 26 zeigt die so entstandenen Ergebnisse für die natürlichen Quellen. Hier ist zunächst zu erkennen: Die Lautstärke der Quelle (blaue Kurve) bestimmt das Resultat für Quellen im Abstand  $>ca.0,85m$ . Die näheren Quellen ( $<ca.0,85m$ ) allerdings weichen von dieser Regel ab und werden tatsächlich näher wahrgenommen. Dies bestätigt die in Kapitel 5.4 zitierten Gesetzmäßigkeiten, nach denen binaurale Merkmale nur für sehr nahe Quellen im Abstand  $<1m$  wirken können. Dies gilt allerdings nicht für alle Versuchspersonen: Besonders am Histogramm für das Signal in 25cm Entfernung zeigt sich, dass die Versuchsteilnehmer teilweise gut (ca.45cm), teilweise gar nicht (ca.1m) die tatsächlich nahe Distanz erkennen. Hier macht sich bemerkbar, dass die Aufgabe für die Versuchspersonen sehr schwer war: Nur etwas mehr als die Hälfte der 18 Versuchspersonen war in der Lage, die Distanzunterschiede aufgrund binauraler Merkmale zu erkennen.

Abbildung 27 schließlich zeigt das Ergebnis für die virtuellen WFS-Quellen: Die Distanz folgt offensichtlich allein dem Pegel statt der Position der Quelle, es kann kein Zusammenhang zwischen Position der Quelle und wahrgenommener Distanz festgestellt werden. Beson-

ders deutlich lässt sich das noch einmal an den Abbildungen 28 und 29 erkennen, die die Daten aus Abbildung 26 und 27 enthalten, aber nun nach dem Pegel am Hörort sortiert sind.



**Abbildung 28: Natürliche Quellen,  
zufällige Lautstärke.  
Daten aus Abbildung 26, sortiert nach Lautstärke**



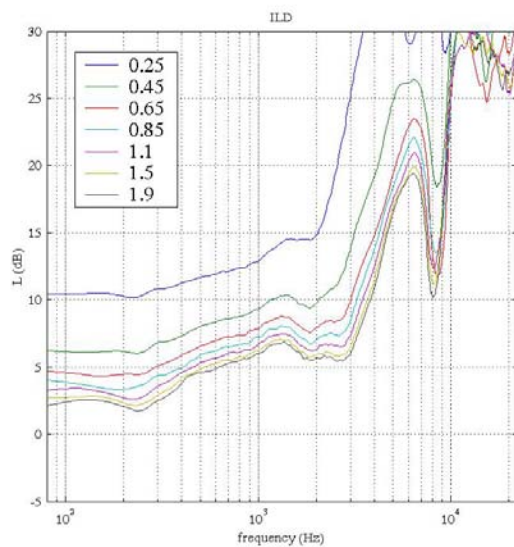
**Abbildung 29: Natürliche Quellen,  
zufällige Lautstärke.  
Daten aus Abbildung 27, sortiert nach Lautstärke**

Es ist also nicht möglich, die Entfernung einer WFS-Quelle aufgrund ihrer Wellenform an einem festen Hörort zu ermitteln. Der Grund dafür ist in der Natur der fokussierten Quellen zu suchen. Sie entspricht in einigen, teils entscheidenden Punkten nicht der Natur einer natürlichen Quelle:

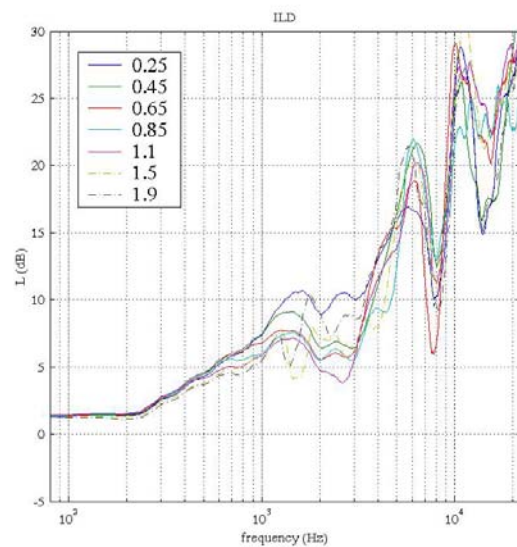
- „1/r-Gesetz“ gilt nicht: Die Lautstärke einer Quelle ist nicht bei WFS nur vom Abstand zwischen Quelle und Hörer verantwortlich. Auch der Abstand des Hörers zum Array spielt eine Rolle. Grund dafür ist die Reduzierung der Wiedergabe auf ein Lautsprecherarray statt einer zweidimensionalen Wiedergabefläche (sog. Amplitudenfehler).
- Fehlende Fokussierung tiefer Frequenzen: Nach Fink ([22]) kann keine Fokussierung mehr stattfinden, wenn der Durchmesser des Fokuspunkts kleiner wird als die halbe Wellenlänge  $\lambda/2$ .

Eine Messung bzw. realgetreue Simulation veranschaulicht, dass aufgrund dieser Fehler die gerade für Distanzwahrnehmung naher Quellen verantwortlichen Merkmale kaum vorhanden sind. Die in Kapitel 5.4 erwähnten nötigen tieffrequenten interauralen Pegeldifferenzen bestimmen maßgeblich die Distanz naher Quellen. Abbildung 30 (natürliche Quellen) und Abbildung 31 (virtuelle fokussierte WFS-Quellen) zeigen die deutlichen Unterschiede. Es sind die interauralen Pegeldifferenzen für die jeweilige Quelle in einem Abstand von 0,25m

bis 1,90m dargestellt. Nur im Bereich zwischen etwa 1000 und 3000 Hz bleiben geringe Unterschiede in den ILD für die virtuelle Quelle übrig.



**Abbildung 30: Interaurale Pegeldifferenzen ILD im Schallfeld einer natürlichen Quelle in einer Distanz d (siehe Legende)**



**Abbildung 31: Interaurale Pegeldifferenzen ILD im Schallfeld einer virtuellen Quelle in einer Distanz d (siehe Legende)**

## 6. Zusammenfassung

In den Untersuchungen, die in den Kapitel 4 und 5 beschrieben sind, wurden Vorzüge der Wellenfeldsynthese (WFS) in Bezug auf die räumliche Wahrnehmung benannt und Probleme diskutiert. Eine bessere Kenntnis der wahrnehmungsbezogenen Eigenschaften der WFS hilft wesentlich bei einer zuverlässigen Einschätzung des Potenzials dieser Wiedergabetechnik für eine bestimmte Anwendung. Nicht immer wird dabei WFS die beste Wahl sein, ihre speziellen Vorzüge räumen ihr aber eine Ausnahmestellung ein.

Die Lokalisationseigenschaften der WFS sind denen einer natürlichen Quelle im Vergleich zu konventionellen Wiedergabemethoden am nächsten und es liegt nahe, dass eine realistische räumliche Abbildung (z.B. „augmented reality“) mit dieser Technik erreicht werden kann. Allerdings sind der WFS auch Grenzen gesetzt, z.B. in Bezug auf die Wahrnehmung der Distanz einer virtuellen Quelle.

## 7. Referenzen

- [1] Theile, G., Wittek, H., Reisinger, M. „Wellenfeldsynthese-Verfahren: Ein Weg für neue Möglichkeiten der räumlichen Tongestaltung“, 22. Tonmeistertagung Hannover, 2002, Tagungsband.

- [2] Berkhout, A.J., de Vries, D. and Vogel, P. "Acoustic control by wave field synthesis", Journal of the Acoustical Society of America, Vol.93, pp 2764-2778, 1993 .
- [3] Boone, M.M., Verheijen, E.N.G. and van Tol, P.F. "Spatial sound field reproduction by wave field synthesis", Journal of the AES, Vol. 43(12), p.1003-1012, 1995 .
- [4] Steinberg and Snow. „Physical Factors.“, Bell Systems Technical Journal, Vol.132, 1934, pp.245-258.
- [5] Wittek, H. “Perception of Spatially synthesized sound fields – Literature review about WFS”, <http://www.hauptmikrofon.de/wfs.htm>, 2003.
- [6] Wightman and Kistler. „The dominant role of low-frequency interaural time differences in sound localization.“, Journal of the Acoustical Society of America, Vol. 91(3), 1992.
- [7] Vogel, P. „Application of wave field synthesis in room acoustics.“, Dissertation, Technische Universität Delft, 1993.
- [8] Verheijen, E. „Sound reproduction by wave field synthesis.“, Dissertation, Technische Universität Delft, 1997.
- [9] Start, E. „Direct Sound Enhancement by Wave Field Synthesis.“, Technische Universität Delft, ISBN 90-9010708-8, 1997.
- [10] Wittek, H. and Theile, G. „Untersuchungen zur Richtungsabbildung mit L-C-R-Hauptmikrofonen.“, 20. Tonmeistertagung, Hannover, 2000, Tagungsband.
- [11] Wittek, H. „Image Assistant“ auf <http://www.hauptmikrofon.de> , 2000.
- [12] Huber, T. „Untersuchungen zur Lokalisation virtueller Quellen bei Wellenfeldsynthese“. Entwurf zur Diplomarbeit IRT/TU München, 2001, auf <http://www.hauptmikrofon.de/wfs.htm>
- [13] Rumsey, F. „Spatial Quality Evaluation for reproduced Sound: Terminology, Meaning, and a Scene-based paradigm.“, Journal of the Audio Engineering Society, Vol.50, No.9, 2002.
- [14] Noguès, M., Corteel, E. and Warusfel, O. „Monitoring Distance Effects with Wave Field Synthesis.“ Proceedings of the 6<sup>th</sup> Int. Conf. on Digital Audio Effects (DAFX-03), London, UK, 2003.
- [15] Usher, J., Martens, W.L., Woszczyk, W. „The influence of the presence of multiple sources on auditory spatial imagery by a graphical response technique“, Proceedings of the ICA 2004.

- [16] Wittek, H., Kerber, S., Rumsey, F., Theile, G. „Spatial perception in Wave Field Synthesis rendered sound fields: Distance of real and virtual nearby sources.“ Proceedings 116<sup>th</sup> AES Convention, Berlin, Germany, Preprint No.6000, 2004.
- [17] Zahorik, P., 1996: “Auditory distance perception: A literature review”, Paper for fulfillment of PhD, Wisconsin University.
- [18] Blauert, J., 1997: “Spatial Hearing”, MIT Press, ISBN 0-262-02413-6.
- [19] Brungart and Rabinowitz. „Auditory localization of nearby sources, I, II and III”, J. Acoust. Soc. Am., Vol. 106, No. 3,4 and 6, 1999.
- [20] Martens, W.L. „Perceptual evaluation of filters controlling source direction: Customized and generalized HRTFs for binaural synthesis.“, Acoustical Science and Technology, 24 (5), 2003, pp.220-232.
- [21] Theile, G., Wittek, H. and Reisinger, M. „Potential Wavefield Synthesis Applications in the Multichannel Stereophonic World.“, Proceedings AES 24th Int. Conference on Multichannel Sound, Banff, Canada, 2003.
- [22] Fink, M. „Acoustic Time-Reversal Mirrors.“ in Fink, M. et al. (ed) „Imaging of Complex Media with Acoustic and Seismic Waves“, Topics Appl. Phys 84, 17-43, Springer-Verlag, Berlin, 2002.
- [23] Theile, G., Wittek, H. and Reisinger, M. “Wellenfeldsynthese-Verfahren: Ein Weg für neue Möglichkeiten der räumlichen Tongestaltung.“, 21.Tonmeistertagung, Hannover, 2002, Tagungsband.