

# Mikrofonarrays - Methoden und Anwendung

Dirk Püschel<sup>1</sup>

<sup>1</sup> Soundtec GmbH, 37073 Göttingen, E-Mail: d.pueschel@soundtec.eu

## Einleitung

Mikrofone sind die Antennen des Akustikers, und die Kombination mehrerer Mikrofone in einem Mikrofonarray eröffnet vielfältige Möglichkeiten, die Richtcharakteristik dieser Antenne in weiten Grenzen zu gestalten. Durch die Verfügbarkeit immer schnellerer Prozessoren ist es inzwischen möglich, die Mikrofonsignale auf komplexe Weise zu verrechnen und damit so unterschiedliche Dinge zu realisieren wie Richtmikrofone, Akustische Kameras, akustische Holographiesysteme oder auch einen virtuellen Kunstkopf. Im Vortrag werden die wesentlichen Einsatzfelder von verschiedenen Mikrofonarrays dargestellt. Es wird die jeweilige Funktionsweise erläutert, und so ein Einblick in die Einsatzgebiete und die jeweiligen Verfahren gegeben. Dabei werden insbesondere die Algorithmen und Unterschiede zwischen Schallfeld-Holographie, Richtmikrofonen und verschiedene Akustischen Kameras beleuchtet. Darüber hinaus werden aktuelle Entwicklungen von Mikrofonarrays sowie zukünftige Möglichkeiten aufgezeigt.

## Überblick

Im folgenden werden die folgenden Themenbereiche behandelt:

- Fourierakustik mit Holographie und Nahfeldholographie
- Akustische Kameras mit Beamforming, Fokalisation und Identifikation
- Richtmikrofone mit Verfahren und beispielhaften Forschungsfeldern

## Holographie/ Fourierakustik

Das kleinste Array besteht aus zwei Mikrofonen (Abbildung 1), während das derzeit vielleicht grösste akustische Array 1020 Mikrofone umfasst (MIT Forschungsarray – Abbildung 2) [1].

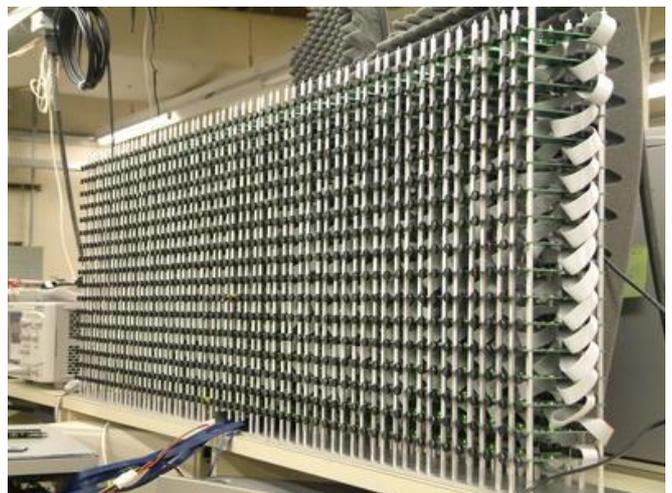
Wenn Arrays zur Lokalisation von Schallquellen eingesetzt werden, dann stellt sich die Frage wie das Schallfeld um eine Anordnung von Quellen aussieht. Denn dieses Schallfeld wird von dem Array aufgenommen und aus ihm soll letztlich die Anordnung und der Pegel der Quellen berechnet werden. Einen guten Einblick zu dem Thema gibt das Buch von E.G.Williams [2].

Um diese Berechnungsverfahren in der Fourierakustik zu betrachten, arbeiten wir mit nur einer Spektrallinie, deren Schallfeld an jedem Punkt durch eine komplexe Zahl mit

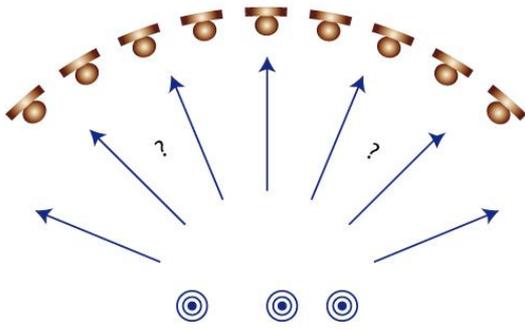
Betrag und Phase gekennzeichnet ist. D.h. Wir stellen uns vor, dass an jedem Punkt des Schallfeldes eine Zerlegung der Zeitfunktion nach Frequenzen stattfindet und wir nur eine frei wählbare Frequenz betrachten. Von dieser Frequenz haben wir somit eine räumliche Verteilung im Schallfeld in 3 Dimensionen, die im weiteren betrachtet wird. Da sich alle Frequenzen ohne gegenseitige Beeinflussung überlagern, kann damit letztlich das gesamte Schallspektrum behandelt werden.



**Abbildung 1:** Das kleinste Array besteht aus zwei Mikrofonen.



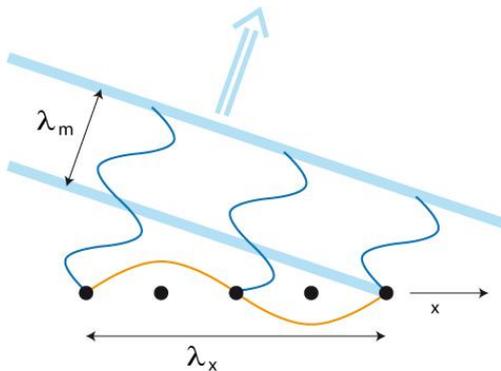
**Abbildung 2:** Eines der derzeit grössten Arrays ist das MIT Forschungsarray mit 1020 Mikrofonen und Rechenknoten.



**Abbildung 3:** Die Abbildung der Quellen von der Objektebene ins Fernfeld ist gesucht. Sie wird durch die Fouriertransformierte der räumlichen Verteilung der Quellen gebildet.

Betrachten wir auf diese Weise die Abbildung einer Anordnung von Quellen auf einer Linie (Abbildung 3) --- also eine eindimensionale Schallverteilung --- dann wird die Verteilung im Fernfeld durch die Fouriertransformation gebildet. D.h. Wenn wir um diese Quellen Mikrofone unter allen Winkeln im Fernfeld aufstellen und dort die komplexen Werte bei einer Frequenz anschauen, dann sind diese die Fouriertransformierte der (komplexwertigen) räumlichen Anordnung in der Quellebene. Somit kann durch einfache Rücktransformation der gemessenen Verteilung über die Winkel auf die Verteilung in der Quellebene zurückgerechnet werden.

Um dies zu sehen, schauen wir uns die Abstrahlung in das Fernfeld in der graphischen Darstellung an (Abbildung 4).

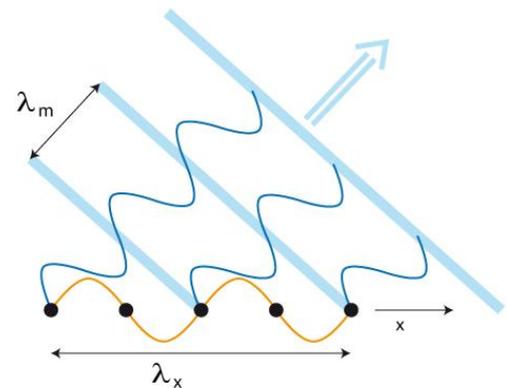


**Abbildung 4:** Abstrahlung der Schallwellen in das Fernfeld aus der Objektebene von den Quellpunkten (schwarz).  $\lambda_m$  ist die Wellenlänge der Ausbreitung im Medium mit  $\lambda_m = c/f$  (Schallgeschwindigkeit  $c$  und Frequenz  $f$ ).  $\lambda_x$  ist hier die die Grundwelle der räumlichen Schallverteilung auf der X-Achse. In Richtung der konstruktiven Überlagerung ergeben sich Wellenfronten, die als blaue Balken eingezeichnet sind.

Die Punkte auf der X-Achse repräsentieren die räumliche Verteilung der komplexen Werte bei einer Frequenz. Diese

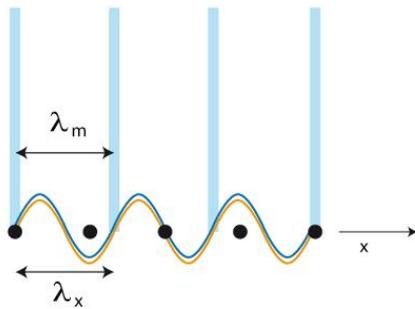
Räumliche Verteilung kann in Fourierkomponenten zerlegt werden, also als Summe von Sinusfunktionen betrachtet werden (tatsächlich als Summe von komplexen e-Funktionen, die als Kombination von Sinus- und Cosinus-Funktionen beschrieben werden können). In dem Bild die die Grundwelle der räumlichen Verteilung orange eingezeichnet. Jeder Punkt dieser Verteilung strahlt eine Welle mit der betrachteten Schallfrequenz  $f$  ab. Diese Abstrahlung überlagert sich nur für eine Richtung konstruktiv, während sie sich für alle anderen Richtungen auslöscht. D.h. Der Anteil der räumlichen Schallverteilung, der sich durch die orange Welle in der Objektebene beschreiben lässt, strahlt unter einem definierten Winkel in das Fernfeld ab, der sich aus dem Verhältnis der Schallwellenlänge im Medium und der Wellenlänge der betrachteten räumlichen Verteilung ergibt.

Betrachten wir die harmonischen der räumlichen Verteilung in Abbildung 5, so sehen wir, daß die Abstrahlung mit zunehmender Neigung aus der Orthogonalen zur Quellebene erfolgt.



**Abbildung 5:** Bei höheren Wellenzahlen  $k=c/\lambda$  erfolgt die Abstrahlung mit zunehmender Neigung aus der Orthogonalen. D.h. Die Anteile der höheren Raumfrequenzen aus der räumlichen Verteilung in der Quellebene finden sich unter zunehmendem Winkel im Fernfeld.

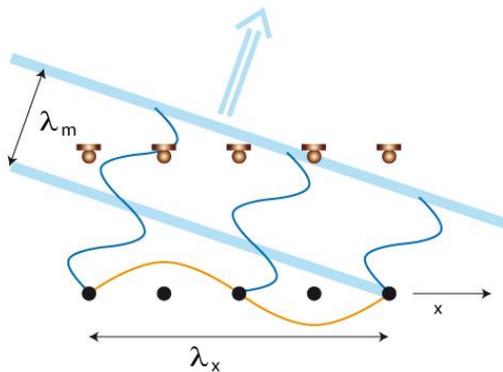
Bei noch höheren räumlichen Frequenzen wird irgendwann die räumliche Wellenlänge gleich und dann kleiner der Wellenlänge im Medium. Abbildung 6 zeigt den Fall, dass die Wellenlängen identisch sind. Die Abstrahlung ist dem Fall parallel zur Objektebene. Bei noch kürzeren Wellenlängen der räumlichen Verteilung kann dann keine Abstrahlung mehr in das Fernfeld erfolgen.



**Abbildung 6:** Wenn die Wellenlänge der räumlichen Verteilung gleich der Wellenlänge im Medium ist, erfolgt die Abstrahlung parallel zur Objektebene.

### Holographie bei Messung in definiertem Abstand

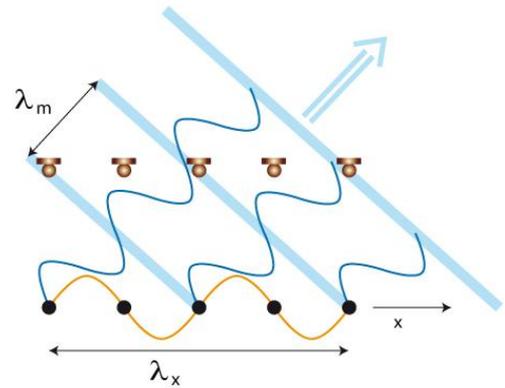
Da die Messung im Fernfeld unpraktisch bzw. in vielen Fällen unmöglich ist, wird das Schallfeld mit Mikrofonen in einer Ebene mit definiertem Abstand zur Objektebene gemessen, wie es in Abbildung 7 gezeigt ist.



**Abbildung 7:** Messung des Schallfeldes mit Mikrofonen in einer Mikrofonebene mit definiertem Abstand von der Objektebene. Hier dargestellt für die Grundwelle der räumlichen Verteilung.

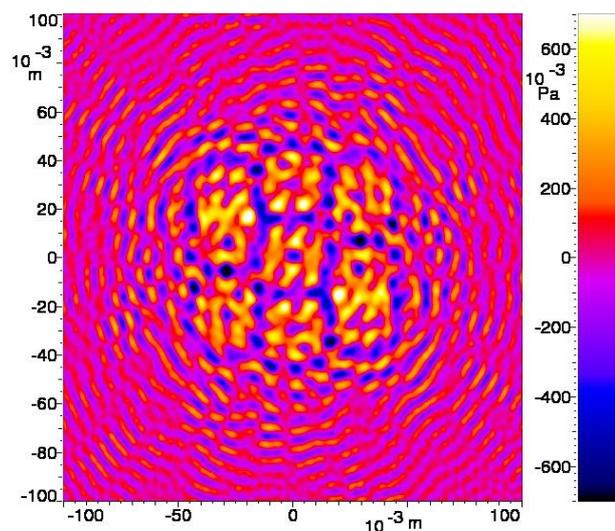
In diesem Fall ist die Verteilung in der Mikrofonebene nicht einfach die Fouriertransformierte der Verteilung in der Objektebene. Stattdessen sehen wir hier dass die Raumfrequenzen in der Mikrofonebene genauso vorhanden sind, wie in der Objektebene. Sie sind lediglich räumlich verschoben. Die Verschiebung ergibt sich aus dem Abstrahlungswinkel und dem Abstand zwischen Mikrofon- und Objektebene. Das bedeutet, wenn wir die Wellenlänge des Schalls im Medium kennen und die Wellenlänge der betrachteten Raumfrequenz, kennen wir den Abstrahlungswinkel und mit dem Abstand kennen wir dann auch die (Phasen-)Verschiebung der Raumfrequenz zwischen den Ebenen.

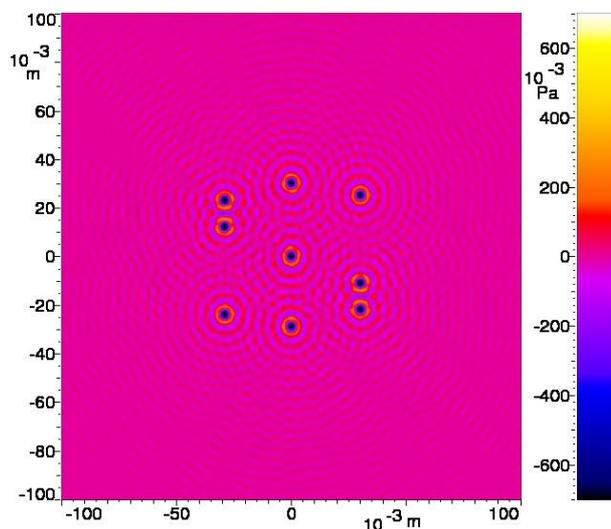
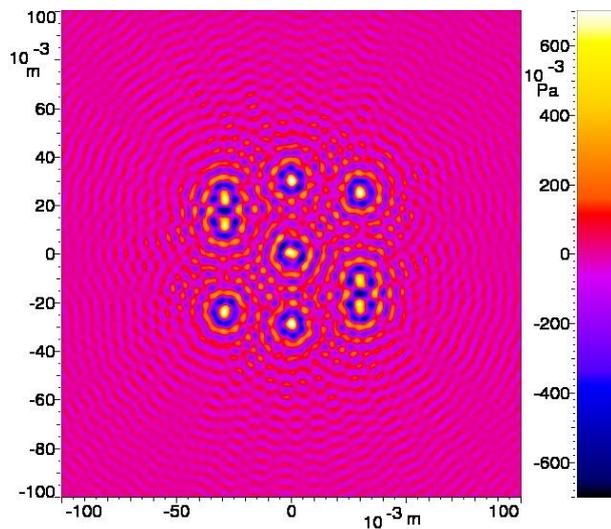
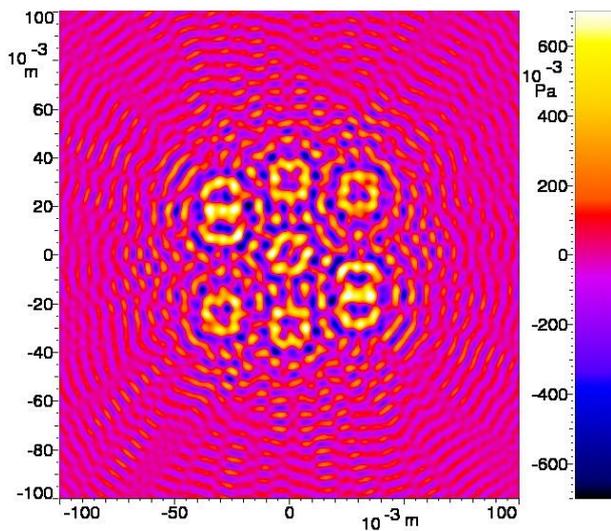
Diese Verschiebung ist für jede Raumfrequenz eine andere wie wir in dem Beispiel der nächsthöheren Raumfrequenz in Abbildung 8 sehen können.



**Abbildung 8:** Stärkere Verschiebung der räumlichen Verteilung bei Betrachtung einer höheren Raumfrequenz aufgrund der stärker geneigten Abstrahlung für diese Raumfrequenz.

Somit kann für jede Raumfrequenz die Phasenverschiebung zwischen Objekt und Mikrofonebene sehr einfach berechnet werden. Damit aber wird die Rekonstruktion der Schallverteilung in der Objektebene ermöglicht. Wir müssen lediglich die Fouriertransformierte der räumlichen Verteilung in der Mikrofonebene bilden, die Phasenverschiebung zwischen den Ebenen für jede Raumfrequenz passend zurück drehen und dann durch Fourier-Rücktransformation die Verteilung in der Objektebene ermitteln. Damit kann dann aus einer gemessenen Verteilung der komplexen Spektralwerte bei einer Frequenz in der Mikrofonebene auf die Verteilung dieser Spektralwerte in der Objektebene zurück gerechnet werden. In Abbildung 9 ist dies für verschiedene Abstände von der Objektebene beispielhaft gezeigt.





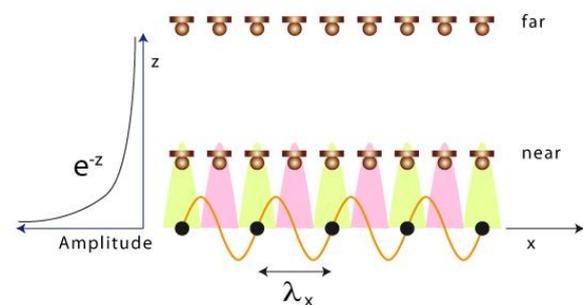
**Abbildung 9:** Schallfeld in verschiedenen Abständen von der Objektebene. Von der Verteilung in der Mikrofonenebene wurde auf verschiedene Ebenen zurück gerechnet, die sich der Objektebene nähern. Am Ende ist die Verteilung der Quellen in Form eines S klar erkennbar.

Wird dies für alle Spektrallinien einzeln durchgeführt und das Ergebnis aller Frequenzen an jedem Raumpunkt in die

Zeit zurück transformiert, liegt das Schallfeld in der Objektebene komplett räumlich und zeitlich vor.

## Nahfeld-Holographie

Bei der oben beschriebenen Holographiemethode können räumliche Details bis zur Schallwellenlänge in der Objektebene aufgelöst werden. Feinere Details bilden sich in Raumfrequenzen ab, die nicht mehr ins Fernfeld abgestrahlt werden und daher nicht rekonstruiert werden können. Was aber passiert im Nahfeld? Abbildung 10 skizziert die Abstrahlung in das Nahfeld für den Fall, dass die Wellenlänge der räumlichen Verteilung kleiner wird als die Wellenlänge des abgestrahlten Schalls.



**Abbildung 10:** Abstrahlung in das Nahfeld mit räumlichen Wellenlängen die kleiner sind, als die Wellenlänge des Schalls im Medium. Das Schallfeld breitet sich orthogonal zur Objektebene aus und bildet dabei direkt die räumliche Struktur in der Objektebene ab. Allerdings wird diese Ausbreitung exponentiell gedämpft, so dass diese Anteile bald im Rauschen versinken. Es sind zwei Mikrofonenebenen gezeigt. Eine ist nah am Objekt und kann die Nahfeldverteilung erfassen, während die andere so fern ist, dass sich die abgestrahlten Anteile für diese Raumfrequenzen schon zu stark ausgelöscht haben, um sie zu rekonstruieren.

Die feinen Anteile der räumlichen Verteilung löschen sich mit zunehmendem Abstand von der Objektebene aus, da es keine Winkel gibt, unter dem sie ausbreitungsfähig sind. Daher werden sie exponentiell gedämpft senkrecht zur Objektebene in den Raum abgestrahlt. Liegt die Mikrofonenebene nah genug an der Objektebene, so können diese Anteile noch erfasst werden. Dies ist i.d.R. Der Fall, wenn der Abstand kleiner ist als das zweifache der Wellenlänge.

Bei der Rückrechnung muss die exponentielle Dämpfung umgekehrt werden, was das Verfahren sehr anfällig für Rauschen und alle Arten von Störung macht. Die Rückrechnung wird natürlich umso schwieriger, je kleiner die Details sind. Die tatsächliche Grenze muss oft experimentell ermittelt werden. Man verzichtet dann bei der Rekonstruktion auf die höheren Raumfrequenzen.

Das Verfahren kann allgemeiner für beliebig gekrümmte Flächen durchgeführt werden, wenn man eine 3-dimensionale Verteilung von Mikrofonen und Objektebene zulässt. Damit ist dann allgemein die Ausbreitungsrechnung und Rückrechnung mit Hilfe eines Oberflächenmodells des Objektes möglich. Aus der Schalldruckverteilung können durch Ableitung die Schnellevektoren an jedem Punkt berechnet werden und damit ist auch die Bestimmung der Schwingamplitude in der Objektebene möglich.

## Zusammenfassung Holographie

- Fourierakustik ermöglicht uns das Verständnis von Wellenfeldern und damit Schallfeldholographie
- Mit Nahfeldholographie ist die Auflösung der Schallverteilung genauer als eine Wellenlänge im Medium möglich
- In Kombination mit Modellen der Objektoberfläche ist die Berechnung der Schwingung von komplexen Oberflächengemetrien möglich

## Akustische Kameras

### Nahfeldkamera

Statt die Schwingung an der Objektoberfläche mittels Nahfeldholographie zu berechnen, kann sie in vielen Fällen auch direkt gemessen werden, wenn eine akustische Kamera mit echten Schnellesonden verwendet wird. Solch eine Kamera mit 20 Sonden ist in Abbildung 11 gezeigt.



**Abbildung 11:** Akustische Nahfeldkamera bestehend aus 20 Schnellesonden. Damit wird direkt die Schallschnelle vor der Objektoberfläche gemessen. Eine Rückrechnung auf die Objektebene ist i.d.R. nicht nötig.

Mit den Schnellesonden wird direkt vor der Objektoberfläche die Schallschnelle bestimmt. Die Messung der Schnelle ist an der Oberfläche sehr vorteilhaft, weil dort der abgestrahlte Schalldruck niedrig ist und von den

reflektierten Anteilen des Schalls aus dem Raum überlagert wird. Alternativ kann aus Druck und Schnelle auch der Schallfluss, also die Intensität bestimmt werden. Das Verfahren ist aufgrund der direkten Messung der physikalischen Größen des Schallfeldes sehr robust.

### Fernfeldkamera

Der häufigere Fall sind akustische Beamforming-Kameras für das Fernfeld, welche anders arbeiten als die Holographiemethode. Die Abbildungsfunktion wird in diesem Fall wie bei einem Kameraobjektiv realisiert. Die verschiedenen Schallwege werden durch unterschiedliche Verzögerung so überlagert, dass sie sich für einen Objektpunkt konstruktiv überlagern. Im optischen Fall leistet dies die Linse. Im akustischen Fall wird diese Verzögerung nach der Wandlung des Mikrofonsignals in einen digitalen Datenstrom im Rechner durchgeführt. Die verschiedenen Bildpunkte ergeben sich lediglich durch Verzögern und Aufsummieren der Mikrofonsignale.

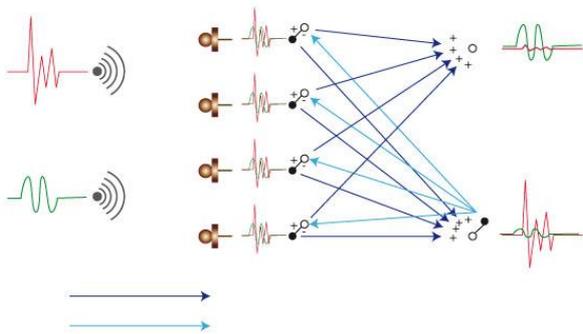
Schwachpunkt des Verfahrens ist die relativ geringe Zahl von Schallwegen/Mikrofonen die praktisch verwendet werden können. Während im optischen Fall eine extrem hohe Zahl von Lichtwegen zur Verfügung steht und die Dynamik nur durch die Abbildungsleistung der Linse beeinflusst wird, beträgt die nutzbare Dynamik bei einer akustischen Beamformingkamera mit 32 Mikrofonen maximal 15 dB. Jede Verdoppelung bringt lediglich 3 dB Gewinn, da das Prinzip der Abbildung allein auf dem Unterschied zwischen kohärenter und inkohärenter Addition der Schallwege basiert. Selbst mit 1024 Mikrofonen ergeben sich damit maximal 30 dB. Dieser theoretische Wert wird aber für viele Frequenzen nicht erreicht, da die Mikrofone vor allem bei tiefen Frequenzen zu nahe stehen und damit redundant werden. Aufgrund der großen Bandbreite des Schalls mit einem typischen Faktor von 1000 zwischen kleinster und größter Frequenz ist eine gleichmäßig effizient funktionierende räumliche Abtastung für alle Frequenzen nicht erreichbar. Im optischen Fall ist das Verhältnis der Wellenlängen von Rot zu Blau kleiner als 2 und damit viel unproblematischer.

### Dynamikverbesserung

Diesem Umstand versucht man u.A. durch Entfaltung zu begegnen. Jede Quelle verursacht nicht nur einen Bildpunkt, sondern ist abgeschwächt überall im Bild entsprechend der Abbildungsstreuung vorhanden. Dies kann betrachtet werden wie die Verbreiterung von Spektrallinien bei der Spektralanalyse durch Fensterung im Zeitbereich. Entsprechend kann man versuchen eine verbesserte Abbildung durch Entfaltung des Beamformingbildes zu erreichen. Dabei variiert allerdings die Entfaltungsfunktion von Punkt zu Punkt, was das Verfahren aufwändig macht und eine Realisierung sehr erschwert.

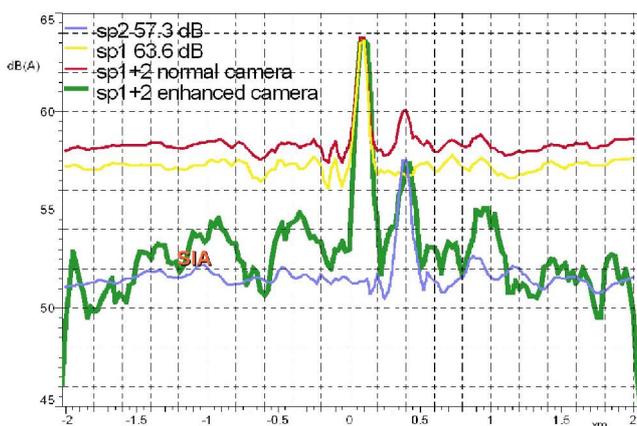
Ein anderes Verfahren für eine stark verbesserte Signaldynamik in der akustischen Kamera wurde von

Soundtec entwickelt und patentiert. Das Schema dazu ist in Abbildung 12 skizziert.



**Abbildung 12:** Akustische Kamera bei der die rekonstruierten Quellen auf das Messsignal zurückgerechnet und von diesem abgezogen werden. Die Länge der dunkel- und hellblauen Pfeile soll die Verzögerungslänge andeuten. Das Verfahren erhöht die Signaldynamik sehr stark und kann im Idealfall eine vollständige Trennung der Quellen erreichen.

Dabei wird zuerst die stärkste Quelle rekonstruiert. Nachdem das Schallsignal für diese Quelle bekannt ist, kann zurück gerechnet werden, wie dieser Anteil in den ursprünglichen Mikrofonensignalen ausgesehen hat und von diesen abgezogen werden. Aus den so bereinigten Mikrofonensignalen kann dann die nächst schwächere Quelle rekonstruiert werden. Die Fehler, die dabei gemacht werden, lassen sich ebenfalls korrigieren, so dass im Idealfall eine komplette Trennung aller Quellen möglich ist. Dazu muss die Zahl der Mikrofone grösser sein, als die Zahl der Quellen. Das Verfahren wird als **Source Identification Algorithm** bezeichnet, also **SIA**. Abbildung 13 zeigt einen Vergleich der Abbildungsleistung einer Beamforming Kamera mit dem SIA Verfahren.

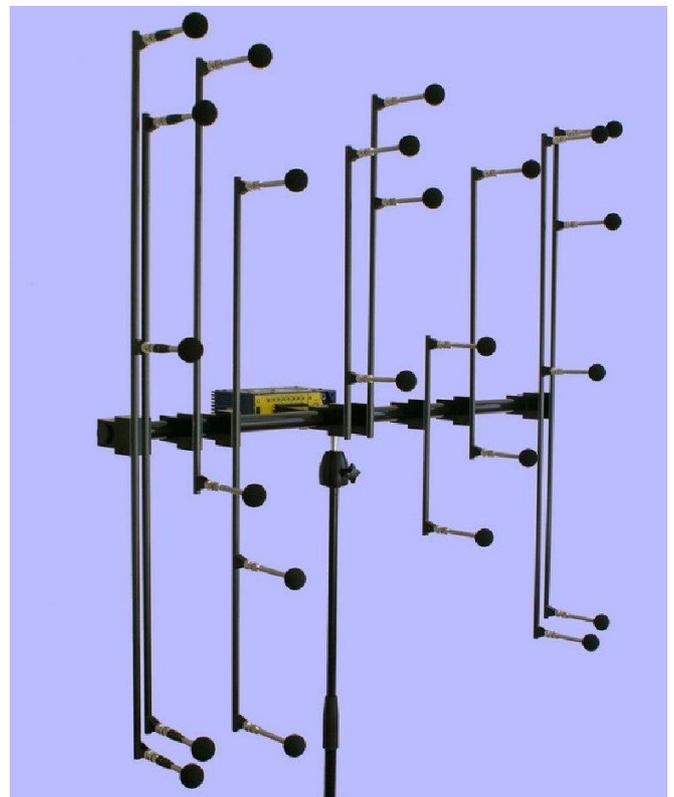


**Abbildung 13:** Vergleich der Abbildungsleistung zwischen Beamforming und SIA Verfahren mit 10 Mikrofonen. Abgebildet werden breitbandig zwei Schallquellen mit Rauschen. Die gelbe Kurve zeigt Quelle 1 allein mittels Beamforming. Die Blaue Kurve entsprechend Quelle 2 und die rote Kurve die Abbildung beider Quellen zusammen.

Die Grüne Kurve zeigt die Abbildung beider Quellen mit dem SIA Verfahren.

Aufgenommen wurde der Schall zweier Quellen mit einer Anordnung aus 10 Mikrofonen. Damit ist mit Beamforming maximal eine Dynamik von 10dB erzielbar. Werden die Quellen einzeln aufgezeichnet und mit Beamforming abgebildet ergibt sich hier eine Signaldynamik von 6-7 dB in der gelben und blauen Kurve. Werden beide Quellen zusammen betrieben, so ist in der Beamforming Abbildung (rote Kurve) die zweite Quelle noch erkennbar, wird aber mit einem falschen Pegel abgebildet. Das SIA Verfahren (grüne Kurve) hingegen bildet auch den Pegel der zweiten Quelle korrekt ab.

Darüber hinaus erreicht die Signaldynamik tatsächlich Werte von mehr als 10 dB. In der grünen Kurve sind Strukturen erkennbar, die durch Reflektionen im Raum entstanden sind und die in der roten Kurve nur andeutungsweise zu sehen sind. Eine akustische Kamera, die nach diesem Verfahren arbeitet ist in Abbildung 14 zu sehen.



**Abbildung 14:** Akustische Kamera von Soundtec mit 24 Mikrofonen und SIA Algorithmus.

Abbildung 15 zeigt beispielhaft die Analyseleistung einer solchen Kamera im Bereich von 500 Hz. Zwei Quellen mit Schmalbandrauschen im Bereich von 500 Hz können hier klar getrennt werden. Sie sind im Abstand von 30cm zueinander angebracht und befinden sich 5m von der SIA Kamera entfernt.



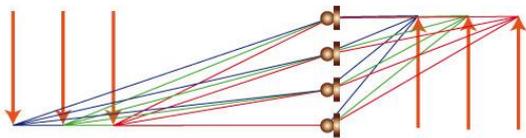
**Abbildung 15:** Abbildungsleistung des SIA Verfahrens an einem Beispiel mit 500Hz Schmalbandrauschen in 5m Abstand zur Kamera. Die Quellen sind 30cm voneinander entfernt.

Die Quellen lassen sich mit dem Verfahren nicht nur trennen, sondern auch Ihr Pegel kann durch Summation von Flächenbereichen korrekt bestimmt werden.

### Autofokus und Multifokus

Ein weiteres Problem bei allen akustischen Kameras ist die korrekte Wahl der Schärfenebene. Wird sie falsch gewählt, sind Schallquellen u.U. gar nicht sichtbar. Aus diesem Grund bieten viele akustische Kameras optische Messsysteme für die Distanz an. Dies hilft nicht weiter, wenn mehrere akustische Objekte in unterschiedlichen Abständen vorhanden sind oder wenn der akustische Abstand ein anderer ist als der optische, wie bei Reflektionen.

Dem kann mit einem akustischen Autofokus begegnet werden, bei dem die Schärfenebene so justiert wird, dass der maximale Kontrast im akustischen Bild erzielt wird. Noch wirkungsvoller wird das Verfahren, wenn für jede einzelne Quelle eine andere Schärfenebene möglich ist. Dies ist in Kombination mit dem SIA Verfahren möglich, wenn nach jeder Quellidentifikation in mehreren Ebenen nach der nächst stärksten Quelle gesucht wird.



**Abbildung 16:** Autofokus und Multifokus erlauben die korrekte Wahl der akustischen Schärfenebene für jede einzelne Schallquelle.

Werden alle genannten Verfahren kombiniert, ist eine sehr hohe Rechenleistung nötig, damit die Kamera in Echtzeit realisierbar ist. Die Implementation auf modernen Grafikprozessoren (GPU) bringt hier einen Faktor 10 gegenüber der Implementation auf der normalen CPU.



**Abbildung 17:** Architektur des Grafikprozessors auf dem die aufwändigen Rechenverfahren für die Echtzeitfunktion implementiert wurden.

Der akustische Multifokus ist bei dem genannten System in einem Abstandsbereich von 1:2 wirksam, so dass nur der ungefähre Abstand zum Objekt angegeben werden muss. Abbildung 18 zeigt zwei Schallquellen an einem Backherd, die aus unterschiedlichen Abständen abgestrahlt werden. Beide Quellen sind scharf fokussiert und damit klar lokalisiert.

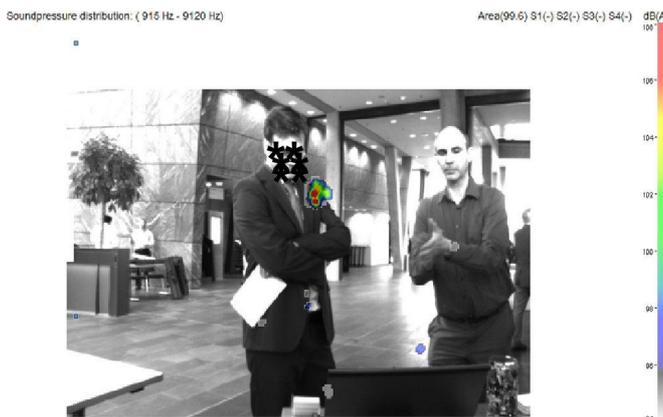


**Abbildung 18:** Gleichzeitige scharfe und klar fokussierte Abbildung mehrere Schallquellen aus unterschiedlichen Abständen an einem Backherd.

Das Verfahren kann dadurch auch eingesetzt werden um zeitlich versetzte Reflexionen zu analysieren. In Abbildung 19 und 20 ist eine Analyse mit sehr hoher Zeitauflösung gezeigt, bei der das Klappgeräusch einer Hand an einer schräg davor stehenden Person reflektiert wird. Dies wäre ohne Autofokus nur sehr mühsam zu analysieren, da zu jedem Zeitpunkt der Fokuspunkt gesucht werden muss, um die Reflexionen zu finden.



**Abbildung 19:** Klappgeräusch einer Hand mit Reflexionen am Körper der Geräusch erzeugenden Person.

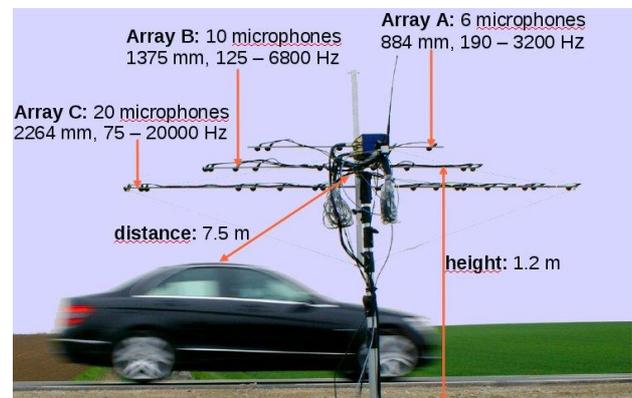


**Abbildung 20:** Reflexion des Klappgeräusches an der schräg davor in 1m Abstand stehenden Person 3ms später.

### Anwendung in der statistischen Vorbeifahrtmessung

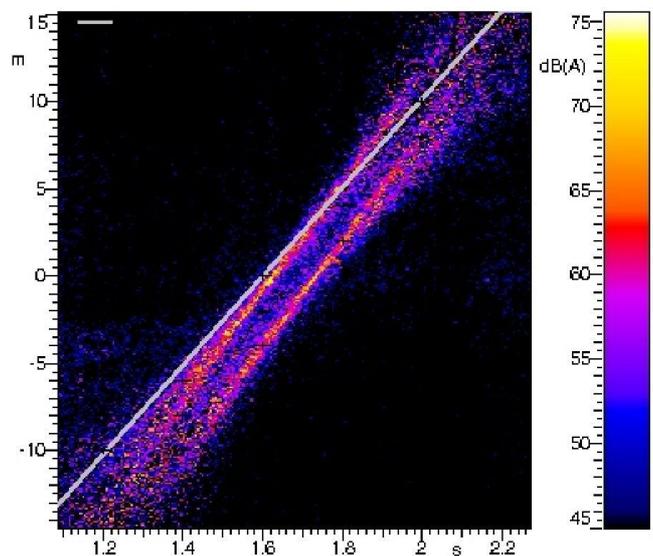
Eine Anwendung für akustische Kameras mit dem oben beschriebenen Eigenschaften ist die statistische Vorbeifahrtmessung, bei der der Pegel von einzelnen Fahrzeugen bestimmt werden soll, während sich direkt davor und dahinter andere Fahrzeuge befinden. Die pegelbestimmenden Schallquellen sind dabei die Rollgeräusche der Reifen. Aus vielen solchen Messungen wird dann eine Statistik über die Eigenschaften des Strassenbelages gebildet. Um nun die hintereinander

fahrenden Fahrzeuge akustisch trennen zu können, wurden Versuche mit 1 dimensionalen Mikrofonarrays mit 6, 10 und 20 Mikrofonen durchgeführt (Abbildung 19), die in [3] und [4] beschreiben sind.



**Abbildung 21:** Messaufbau mit 1 dimensionalen Mikrofonarrays mit 6,10 und 20 Mikrofonen. Gemessen wird in 7.5m von der Strassenmitte und in 1.2m Höhe über der Fahrbahn.

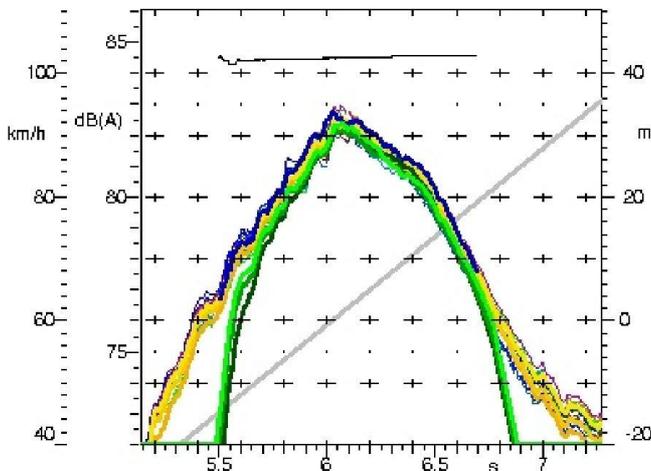
Mit solchen Arrays kann die Position und der Pegel der Schallquellen entlang der Fahrbahn bestimmt werden. Abbildung 22 zeigt die Auftragung über Position und Zeit bei der Messung von einem Fahrzeug, welches mit ca. 100km/h vorbei fährt.



**Abbildung 22:** Auftragung des Pegels der Schallquellen über Position und Zeit bei Messung mit einem eindimensionalen Array. Deutlich sind zwei sich bewegende Quellen erkennbar.

Erkennbar ist, dass sich zwei Schallquellen über den Messbereich von -15m bis +15m bewegen. Die graue Linie zeigt die Position der Fahrzeugfront an, die mittels Lichtschranken und Radar gemessen wurde. Die

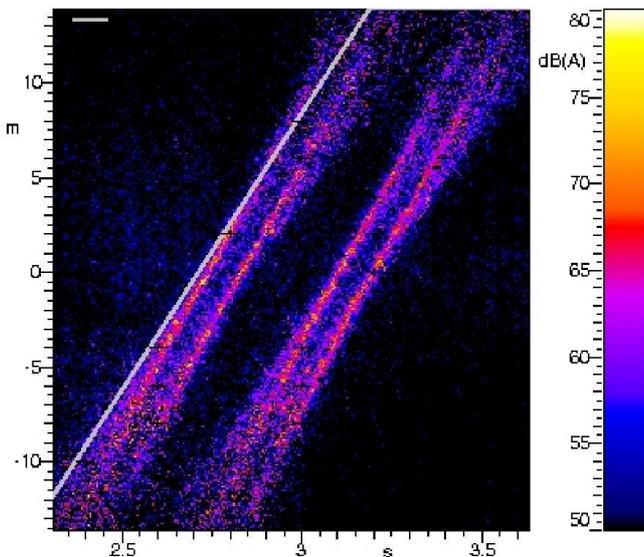
Schallquellen scheinen sich schneller als die Front zu bewegen. Das liegt daran, dass die berechneten Positionen auf die Fahrbahnmitte bezogen sind und die Räder dem Arrays näher sind und dadurch scheinbar schneller sind.



**Abbildung 23:** Auftragung des Summen-Pegels aus dem Pegel Zeit Diagramm (grün) im Vergleich zu einer 1 Mikrofon Standard Pegelmessung (schwarz) und den Pegelverläufen aller beteiligten Mikrofone (gelb).

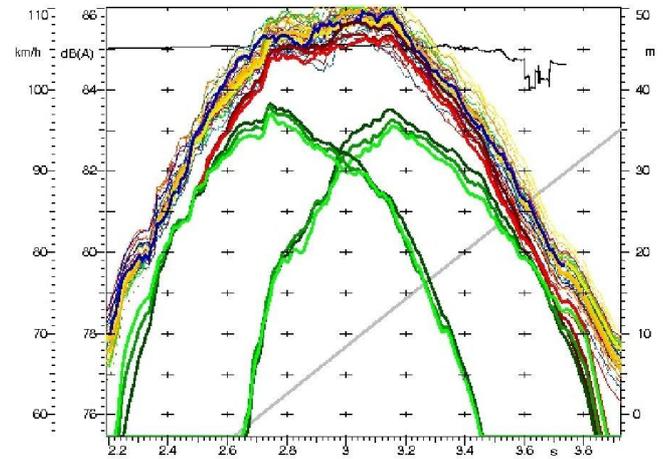
Wird der Pegel aller Schallquellen um die Fahrzeugposition zu jedem Zeitpunkt summiert, so ergibt sich eine Pegelkurve, die in diesem Fall mit dem Verlauf einer Standard Pegelmessung verglichen werden kann (Abbildung 23).

In Abbildung 24 ist das Positions/Zeit Diagramm für den Fall dargestellt, dass zwei Fahrzeuge kurz hintereinander fahren. Deutlich sind nun 4 Schallquellen, also die vier dem Array zugewandten Reifen erkennbar.



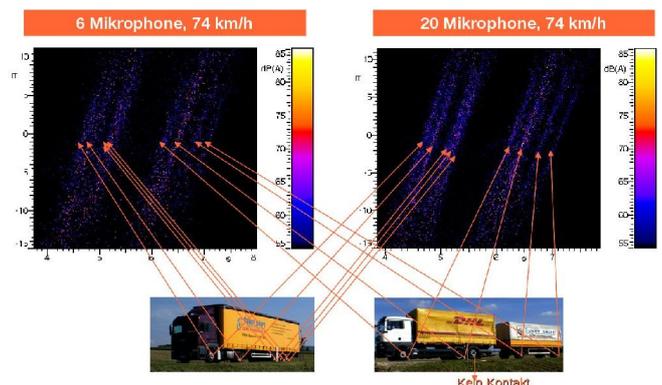
**Abbildung 24:** Auftragung des Pegels der Schallquellen über Position und Zeit bei Messung von zwei Fahrzeugen mit 100km/h kurz hintereinander. Erkennbar sind vier Quellen (vier Reifen), die sich entlang der Fahrbahn bewegen.

Nun können die Pegel der einzelnen Fahrzeuge dadurch gebildet werden, dass um jede Fahrzeugposition ein Fenster in dem Positions Zeit-Diagramm gelegt wird und die Pegel über der Position innerhalb dieses Fensters summiert werden. Somit erhalten wir neben dem Summen-Pegel auch die Einzel-Pegel der Fahrzeuge (Abbildung 25).



**Abbildung 25:** Auftragung der gefensterten Fahrzeug Einzel-Pegel (grün), des Summen-Pegels (rot) und der Pegel aller einzelnen Mikrofone (gelb) aus dem Pegel Zeit Diagramm. Deutlich erkennbar ist die Trennung der Fahrzeugpegel, die im Summenpegel bei dem geringen Fahrzeugabstand längst nicht mehr erkennbar ist.

Die so gewonnenen Einzel-Pegel entsprechen dabei sehr gut den Pegeln, die für die Fahrzeuge einzeln gemessen wurden. Die dargestellten Pegelverläufe enthalten die Messungen mit 6, 10 und 20 Mikrofonen. Schon mit 6 Mikrofonen funktioniert das Verfahren ausreichend präzise. Das Verfahren liefert auch bei der Messung von LKW's mit mehreren Achsen korrekte Ergebnisse. Die Zeit-Positions-Diagramme bei der Messung mit 6 und 20 Mikrofonen für eine Vorbeifahrt von 2 LKW ist in Abbildung 26 dargestellt.



**Abbildung 26:** Positions-Zeit-Diagramm bei der Messung von 2 LKW gemessen mit 6 und mit 20 Mikrofonen im Vergleich. Sowohl das 6er als auch das 20er Array können alle Achsen als erkennbare Quellen abbilden, wenn auch die Auflösung mit 20 Mikrofonen deutlich besser ist.

## Trennung der Quellen bei der Vorbeifahrt bei Messung vor reflektierender Wand

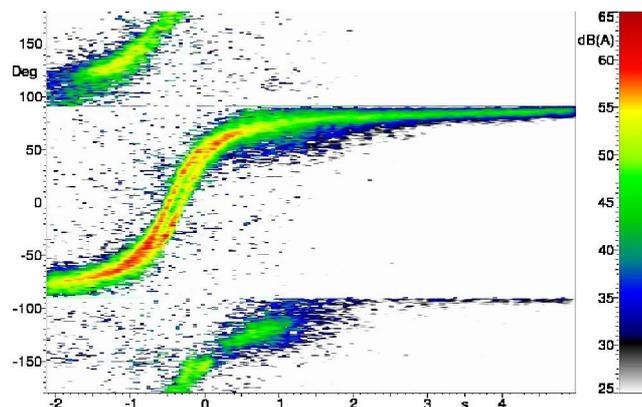
Das oben beschriebene Verfahren wird bereits routinemässig auf Autobahnen und Landstrassen angewendet. Allerdings ist es nicht für Messungen in der Stadt verwendbar, da dort die Reflektionen durch Hauswände dazu kommen. Diese Reflektionen werden bei der Messung mit einem Mikrofon durch ein sogenanntes Backing-Board vermieden. Um aber nun dicht hintereinander fahrende Fahrzeuge trennen zu können, wird ein 2 dimensionales Array benötigt, mit dem alle Richtungen und damit auch die Reflektionen der Hauswand getrennt werden können. Eine Messanordnung zum Vergleich von Backing Board und 2D Mikrofonarrays ist in Abbildung 27 gezeigt.

Mit den 2D Arrays kann nun der Schallpegel rundum nach Richtungen differenziert bestimmt werden. Die Auftragung über der Position entlang der Fahrbahn wird daher nun durch eine Auftragung über Einfallswinkel ersetzt.

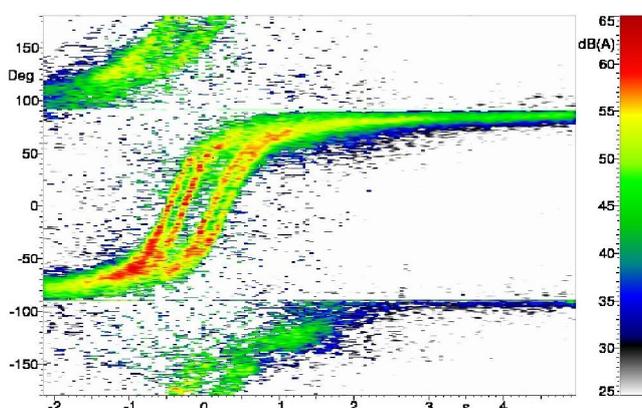


**Abbildung 27:** Messaufbau mit 2-dimensionalen Arrays und Backing Board zur Messung des Vorbeifahrt Pegels vor einer stark reflektierenden Hauswand.

Abbildung 28 zeigt eine solche Messung mit Auftragung des Pegels über Winkel und Zeit durchgeführt mit einem 24 Mikrofon Array. Wie zuvor ist die Bewegung zweier Schallquellen von links nach rechts zu erkennen, hier also von -90 Grad nach +90 Grad. Zusätzlich ist die Reflektion an der Wand hinter dem Array sichtbar in den Winkelbereichen -90 bis -180 und +90 bis +180 Grad. Wird die gleiche Messung mit 2 PKW durchgeführt, so sind wie zuvor beide Fahrzeuge, also alle vier Reifen, klar erkennbar (Abbildung 29). Wird in diesen Diagrammen wiederum der Winkelbereich um jeden PKW summiert und als Pegel über Zeit aufgetragen, so lassen sich auch hier die korrekten Pegelkurven bestimmen. Die Ergebnisse werden hier noch nicht gezeigt, da die zugehörige Forschungsarbeit noch nicht ganz abgeschlossen ist.



**Abbildung 28:** Winkel-Zeit-Diagramm bei der Messung eines PKW. Sichtbar sind die zwei Räder als Schallquellen und zusätzlich die Reflektionen an der Wand in den Winkelbereichen grösser 90 und kleiner -90 Grad.



**Abbildung 29:** Winkel-Zeit-Diagramm bei der Messung zweier PKW. Alle vier Räder sind als Schallquellen erkennbar. Zusätzlich ist die Reflektion an der Hauswand sichtbar.

## Zusammenfassung akustische Kameras

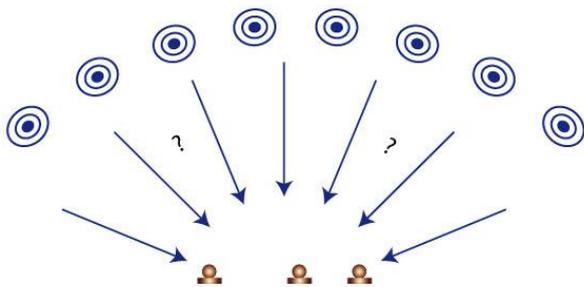
- Die Schwingung der Oberfläche eines Objektes kann mit echten Schnellesensoren direkt vor der Oberfläche gemessen werden
- Akustische Kameras zeigen Schallquellen mit Auflösungen bis ca.  $\frac{1}{2}$  Wellenlänge je nach Verfahren und Apertur
- Akustischer Autofokus, Multifokus und Quellenidentifikation ermöglichen Hohe Abbildungsleistung bei einfacher Anwendung
- Ein Anwendungsbeispiel ist die exakte Pegelbestimmung einzelner Quellen mit rundum Auflösung in der statistischen Vorbeifahrt

## Richtmikrofone

Ein anderes Einsatzfeld von Mikrofonarrays sind Richtmikrofone. Die Anwendungsbereiche von Richtmikrofonen sind z.B.:

- Freisprecheinrichtungen im Auto
- Audiomikrofone am Computer und Handy
- Hörgeräte
- Telekonferenzsysteme
- Arrays für Sprecher in Vortragssälen

Hier ist die Fragestellung gegenüber akustischen Kameras in gewisser Weise umgekehrt. Während im ersten Fall mit einer möglicherweise großen Zahl von Mikrofonen die Schallverteilung in einem ggf. kleinen Bereich möglichst präzise bestimmt werden soll (Holographie), ist die Aufgabe beim Richtmikrofon, mit einer möglichst kleinen Mikrofonanordnung den Schall aus einer Einfallsrichtung zu selektieren oder eine gewünschte Gewichtung des Schalles nach Einfallsrichtung durchzuführen (Abbildung 30).



**Abbildung 30:** Umkehrung der Fragestellung beim Richtmikrofon gegenüber Schallfeldabbildung: Wie werden Schallquellen aus verschiedenen Richtungen bei Aufnahme mittels Array Mikrofon gewichtet?

Diese Unterscheidung wird dann besonders auffällig, wenn die Abmessungen des Arrays für ein Richtmikrofon kleiner als die Wellenlänge des selektierenden Schalles wird. Die Frage ist u.A. ob und wie das überhaupt möglich ist. Methoden für das Design von Richtmikrofonen sind:

- Beamforming --- Richtwirkung durch Verzögerung und Summierung
- Super Directivity --- Schmalere Richtkeulen durch Differenzbildung und damit Nullstellen in der Richtwirkung wie bei Gradientenmikrofonen (Niere, Superniere) und der klassischen Schallintensitätssonde (P-P Sonde).
- Generalisierte Seitenlinienunterdrückung – das ist eine adaptive Methode um die Nullstellen auf die Störkomponenten zu legen

- Sowie viele Variationen der Verfahren in Kombinationen mit automatischer Quelllokalisierung

Einen guten Überblick über die Methoden bei Richtmikrofonen gibt das Buch von M.Brandstein und D.Ward [5]. Die gleichen Verfahren wie für akustische Kameras (z.B. Beamforming) können auch für Richtmikrofone angewendet werden. Ein Beispiel für ein Array-Mikrofon welches Beamforming für die Richtwirkung einsetzt, sind die vier Mikrofone in dem Sensor von Microsofts Kinect Spielekonsole (Abbildung 31).

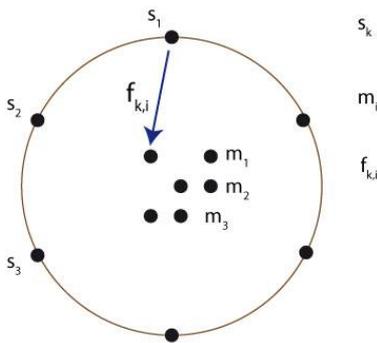


**Abbildung 31:** Microsofts Kinect enthält 4 Mikrofone, die mittels Beamforming zu einem Richtmikrofon werden.

Für den Kinect Sensor wird auf dem Computer ein paralleles Beamforming für 11 Richtungen im Bereich von -50 bis +50 Grad durchgeführt. Das heisst die vier Mikrofonsignale durchlaufen jeweils 11 verschiedene Verzögerungen bevor sie zu 11 Varianten summiert werden. Das entspricht einer eindimensionalen akustischen Kamera mit 11 Bildpunkten. Die Richtung mit der höchsten Energie wird selektiert und deren Signal ausgewertet. Dadurch erzielt das System eine Signal/Rauschverbesserung und kann die akustische Position der Quelle vor dem Sensor grob bestimmen.

Beamforming im Richtmikrofon hat allerdings den Nachteil, dass die Richtmikrofone relativ gross sein müssen oder in Ihrer Selektivität stark begrenzt bleiben. Es gibt aber andere altbekannte Möglichkeiten, die z.B. beim Nierenmikrofon oder bei der klassischen Schallintensitätssonde eingesetzt werden. Dabei wird werden die Mikrofonsignale nicht nur verzögert aufsummiert, sondern in erster Linie subtrahiert. Dies hat zur Folge, dass eine Nullstelle im Richtdiagramm entsteht, nämlich für die Richtung, bei der sich die Signale exakt aufheben. Durch solche Nullstellen kann die Richtkeule sehr effizient geformt werden, selbst wenn das Mikrofon-Array klein gegen die Schallwellenlänge bleibt. Man spricht in dem Fall von Super-Directivity. Beim Nierenmikrofon erfolgt dies durch den mechanischen Aufbau, aber bei der P-P Sonde werden die Signale zweier Mikrofone voneinander abgezogen. Damit dies wie gewünscht funktioniert, ist ein präziser Abgleich zwischen den Mikrofonen notwendig. Und selbst dann hat das Verfahren seine Grenzen, so dass je nach Frequenzbereich ein gewisser Mindestabstand eingehalten werden muss, damit die interessierenden Signalanteile nicht durch Fehler verfälscht werden oder im Rauschen untergehen.

Etwas allgemeiner beschrieben, kann zur Bildung eines Richtmikrofones einfach jede Phasendrehung und Summierung zugelassen werden, womit im Prinzip alle Möglichkeiten gegeben wird. Solange die Zahl der Schallquellen begrenzt bleibt, kann mit einer entsprechenden Anzahl von Mikrofonen theoretisch auch eine vollständige Trennung aller Quellen erfolgen. Das allgemeine Prinzip für solch ein Richtmikrofon ist in Abbildung 32 dargestellt. Für die Erläuterung betrachten wir wieder nur einer komplexwertige Spektrallinie und stellen und die Anwendung für alle spektralen Komponenten getrennt vor.



**Abbildung 32:** Schema für die Abbildung verschiedener Quellen  $s$  auf eine Anzahl von Mikrofonen, die jeweils mit einer Wichtung  $m$  multipliziert werden, bevor Ihre Signale aufaddiert werden. Der Übertragungsfaktor  $f$  wirkt jeweils von Quelle zum Mikrofon (alle Zahlen komplexwertig bei einer Frequenz betrachtet).

Jeder Quellpunkt  $k$  erhält eine gewünschte Wichtung  $S_k$ . Von jedem Quellpunkt zu jedem Mikrofon  $i$  haben wir einen Übertragungsfaktor  $f_{k,i}$  und jedes Mikrofon erhält einen Wichtungsfaktor  $m_i$ , mit dem die Mikrofonensignale aufsummiert werden. Nun sollen die summierten Mikrofonensignale die gewünschten Wichtungen für die Quellpositionen ergeben:

$$s_k = \sum_i f_{k,i} \cdot m_i$$

Wir können dies mit den Vektoren  $S$  und  $M$ , sowie der Matrix  $F$  einfacher schreiben:

$$S = F \cdot M$$

Die Lösung für unser Richtmikrofon ergibt sich dann einfach durch Matrizeninvertierung, so dass wir die richtigen Wichtungen der Mikrofonensignale erhalten:

$$M = F^{-1} \cdot S$$

Die Tücke steckt allerdings im Detail. So ist schon die Matrizeninvertierung nicht immer einfach durchzuführen, aber das größere Problem stellen die Genauigkeiten der Mikrofone und das Rauschen der damit erfassten Signale dar.

Dazu bedarf es dann geschickter Anordnungen (Positionierungen) der Mikrofone selbst, um die Fehler so klein wie möglich zu halten. Darüber hinaus ist oft die Anzahl der gewünschten Wichtungen für die Richtungen größer als die Anzahl der Mikrofone und in dem Fall sind Näherungsverfahren zur Bestimmung einer optimierten Richtwirkung anzuwenden.

### Anwendung eines Richtmikrofon-Arrays als virtueller Kunstkopf

Die Kunstkopftechnik hat derzeit zwei wesentliche Nachteile:

- Die Aufnahme ist nicht an die Übertragungsfunktionen der eigenen Ohren angepasst und reduziert dadurch die Ortbarkeit (insbesondere die Vorne-Ortung) des Schalles
- Beim Abhören ist die Kopfposition festgelegt und der Kopf kann nicht sinnvoll bewegt werden, wie es ein Mensch im realen Schallfeld tun würde

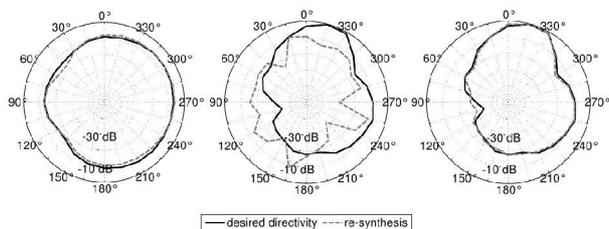
Beide Nachteile können vermieden werden, wenn das Schallfeld im Bereich des Kopfes mit einem Array ausreichend detailliert aufgezeichnet wird und daraus die Ohrsignale für den spezifischen Hörer berechnet werden. Wenn die Übertragungseigenschaften aus verschiedenen Raum-Richtungen zu den Eingängen des Gehörs bei einem spezifischen Hörer bekannt sind (vermessen wurden), dann ist die Aufgabe, aus den Array Signalen ein Mikrofon mit just der zu den Ohren passenden Richtwirkung zu formen.

Ein geeignetes Array für diese Aufgabe ist in Abbildung 33 gezeigt.



**Abbildung 33:** Array aus 24 Mikrofonen zur Aufzeichnung des Schallfeldes im Kopfbereich zur Realisierung eines virtuellen Kunstkopfes

Die zu erzielenden Richtwirkungen sind exemplarisch für die Frequenz bei 1kHz und 11kHz in Abbildung 34 zu sehen. Dort ist auch eine jeweils angepasste Richtwirkung des Arrays mit 8 bzw. 24 Mikrofonen gezeigt. Erkennbar ist, dass bei 1kHz 8 Mikrofone ausreichen, um die Richtwirkung in einer Ebene nachzubilden, während bei 11kHz die Anpassung mit 24 Mikrofonen wesentlich besser funktioniert.



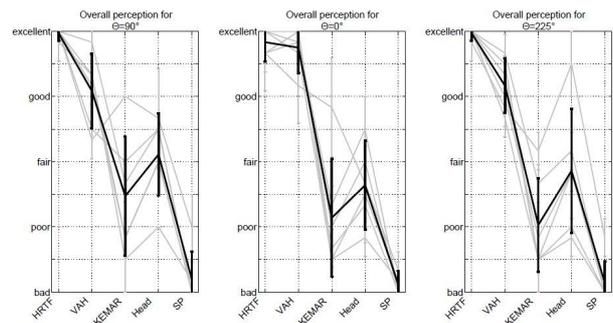
**Abbildung 34:** Richtwirkung des Gehörs in einer Ebene bei 1kHz (Links) und 11kHz (Mitte und Rechts). Dazu jeweils Anpassung der Richtwirkung des Arrays realisiert mit 8 Mikrofonen (links und Mitte) und im Fall 11kHz alternativ mit 24 Mikrofonen (Rechts).

Um zu sehen, ob die Schallfeldwiedergabe nach Aufnahme mit einem solchen Array tatsächlich besser funktioniert als mit einem Kunstkopf, wurden Hörvergleiche durchgeführt [6]. Dabei wurde die Original Schallquelle mit verschiedenen Aufzeichnungsmethoden verglichen. D.h. Die Probanden konnten sich das Original (Rauschen vom Lautsprecher aus einer Richtung) anhören und dies mit verschiedenen Darbietungen über Kopfhörer vergleichen. Die Methoden zur Schallfeldreproduktion sind in Abbildung 35 dargestellt.



**Abbildung 35:** Verschiedene Methoden zur Schallfeldreproduktion über Kopfhörer. 1. Wiedergabe der Ohrsignale wie sie mit den eigenen Gehörmuscheln aufgenommen wurden. 2. Wiedergabe aus Array Aufnahme, die entsprechend der Übertragungsfunktion der eigenen Ohrmuscheln verarbeitet wurde. 3. Wiedergabe einer Aufnahme mittels KEMAR 4. Wiedergabe einer Aufnahme mittels Head Kunstkopf. 5. Wiedergabe einer Aufnahme mittels Petra Kunstkopf

Die Probanden beurteilten dann für drei Richtungen die Gesamt-Qualität der Wiedergabe von schlecht bis exzellent. Die Ergebnisse für die drei Richtungen sind in Abbildung 36 zu sehen.



**Abbildung 36:** Exemplarische Beurteilung der Schallfeldreproduktionsverfahren für drei Einfallsrichtungen 0,0, 90 und 225 Grad.

Die Wiedergabe der Aufzeichnung mit den eigenen Ohren wird dabei durchweg als exzellent empfunden. Dies zeigt, was mit einem idealen Aufzeichnungsverfahren möglich wäre. Die Aufzeichnungen mit den verschiedenen Kunstköpfen werden bestenfalls als ausreichend (fair), aber nicht als gut empfunden. Dagegen wird die Methode mit dem virtuellen Kunstkopf, also Aufzeichnung über das Array und Verrechnung angepasst an die eigenen kopfbezogenen Übertragungsfunktionen mindestens als gut und teilweise als exzellent empfunden. Es ist also möglich mittels eines solchen Arrays eine realistische Wiedergabe des akustischen Feldes zu erzielen, wie sie mit Kunstköpfen nicht erreichbar ist. Dabei ist in diesem Vergleich die Möglichkeit der Kopfbewegung noch nicht realisiert. Es steht zu erwarten, dass damit die Schallfeldwiedergabe sehr realitätsnah gelingt.

### Zusammenfassung Richtmikrofone

- Es gibt verschiedene Verfahren für die Erzeugung von Richtcharakteristiken, die im Zusammenhang mit den verschiedenen akustischen Abbildungsmethoden stehen
- Für die Anwendung sind vor allem kleine kompakte Arrays von Bedeutung, die allein keine Wunder vollbringen können
- Die Entwicklung geht daher in Richtung kompletter Anwendungen in denen die Richtwirkung des Arrays mit nichtlinearen Methoden der Signalverarbeitung kombiniert wird
- Ein Anwendungsbeispiel für ein komplexes Richtmikrofon ist der virtuelle Kunstkopf

## Zusammenfassung

- Die Fourierakustik ist der Schlüssel zum Verständnis von akustischen Schallfeldern und den Abbildungsmethoden mittels Mikrofonarrays
- Die Methoden für akustische Kameras und Schallfeldholographie wurden erläutert und Anwendungsbeispiele dargestellt
- Richtmikrofone können als die Umkehrung akustischer Abbildungssysteme betrachtet werden, bei denen man von kleinen Sensoranordnungen eine komplexe Richtwirkung in das Fernfeld erwartet

## Literatur

- [1] LOUD: Large acOustic Data Array Project, URL: <http://groups.csail.mit.edu/cag/mic-array/>
- [2] Earl G. Williams: "Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography", ACADEMIC PRESS, London, 1999
- [3] D.Püschel, M.Auerbach, W.Bartolomaeus: Pass-by Measurements using array techniques, Euronoise 2009, 1600-1840
- [4] D.Püschel, M.Auerbach, W.Bartolomaeus: Einsatz eines Mikrofon-Arrays für Statistische Vorbeifahrt-Mesungen (SPB), Bundesministerium für Verkehr, Bau und Stadtentwicklung, Wirtschaftsverlag NW, Bonn, 2011
- [5] M.Brandstein, D.Ward: "Microphone Arrays", Springer-Verlag, Berlin Heidelberg New York, 2001
- [6] E.Rasumow, M.Blau, S.Doclo, M.Hansen, S.v.d.Par, D.Püschel, V.Mellert: Individualized binaural reproduction using a virtual artificial head, DAGA 2014