

# Die Akustische Kamera

## Teil 1: Bildgebende Analysemethoden in der Akustik– Stärken und Grenzen des akustischen Beamformings

Dr. Ralf Schröder

In den letzten Jahren haben Mikrofonarraysysteme zur Lokalisierung von Schallquellen vermehrt in den industriellen Anwendungen Einzug gehalten. Eine der ersten praxistauglichen mobilen Entwicklungen dieser Art ist die Akustische Kamera der GFaI<sup>1</sup> aus Berlin. Anwendungsgebiete sind in der Technik vorrangig die akustische Fehlerdetektion und die Nutzung während der Produktentwicklung, aber auch das Gewinnen umweltbezogener Aussagen anhand von Messungen lärmemittierender Objekte aller Art. Dabei ist das bisher abgedeckte Einsatzspektrum außerordentlich vielfältig und reicht von der kleinen elektrischen Zahnbürste über Messungen an Werkzeugen und Haushaltsgeräten, an Kraft- und Schienenfahrzeugen, Motoren, Maschinen und Aggregaten bis hin zu großen Windkraft- und Industrieanlagen.

Das hier vorgestellte System ist portabel ausgelegt und besteht aus dem Mikrofonarray, einem hochkanaligen Datenrekorder zur zeitsynchronen Aufzeichnung und Zwischenspeicherung aller Messdaten sowie einem Rechner mit Auswertesoftware („NoiseImage“). Wahlweise sind für verschiedene Objektgrößen, Messabstände und Frequenzbereiche unterschiedlich große Mikrofonarrays verwendbar, welche über einen einheitlichen Bus mit nur wenigen Steckern (24 Kanäle je Stecker) kontaktiert werden können.

Mit dem im Zeitbereich arbeitenden GFaI-System können akustische Fotos und Filme der interessierenden Objekte relativ einfach und schnell erstellt werden. Diese Bilder geben Auskunft über die örtliche Lage der Hauptschallquellen am Objekt. Die graphische Darstellung erfolgt, ähnlich einem Wärmebild, als Überlagerung des kantenextrahierten optischen Bildes mit der berechneten Farbdarstellung der Schalldruckverteilung. In der aktuellen Version der Auswertesoftware („NoiseImage 3“) sind nun auch dreidimensionale Kartierungen möglich. Hier erfolgt die Überlagerung des farbigen Schalldruckbildes nicht mehr mit einem zweidimensionalen Photo, sondern mit einem vom Anwender gelieferten 3D-Modell des Objektes (Bild 1).

Die zeitliche Abtastrate des GFaI-Systems ist mit 192 kHz je Kanal sehr hoch und erlaubt genaue Kartierungen über den gesamten interessierenden Audiofrequenzbereich, wobei die untere Grenzfrequenz durch den maximalen Mikrofonabstand im Array begrenzt ist. Mit dem System sind neben akustischen Kartierungen stationärer Quellen erstmalig auch transiente Ereignisse und instationäre Signalverläufe in zeitlich sehr hochaufgelösten (theoretisch bis zur Abtastfrequenz, praktisch bis etwa 50000 Bildern je Sekunde) akustischen Filmen (Movies) visuell anschaulich darstellbar. Die Kanaldaten aller Mikrofone stehen nach der Messung weiterhin in voller zeitlicher Auflösung zur Verfügung und können in interaktiven Auswertungen für übliche akustische Berechnungen verwendet werden, dazu gehören z.B. verschiedene Norm-Filterungen und spektrale Analysen, wobei letztere auch ortsselektiv möglich sind. Auf diese Weise kann der Anwender sich sehr schnell einen Überblick über die zeitlichen, örtlichen und spektralen Verteilungen der Hauptschallquellen seines Messobjekts verschaffen.

---

<sup>1</sup> GFaI – Gesellschaft zur Förderung angewandter Informatik e.V.

Bild 1: Akustisches Foto am Beispiel einer 3D-Innenraumkartierung (Bilder/bsp\_3D.jpg)

## Funktionsprinzipien der Akustischen Kamera

Die Funktion der Akustischen Kamera beruht auf der Auswertung der relativen Laufzeiten von den Punkten einer Bildebene bzw. der 3D-Objektoberfläche zu den einzelnen Mikrofonen des Arrays. Die Rekonstruktion der Zeitfunktion für einen Ort  $\mathbf{x} = (x', y', z')^T$  in der Bildebene wird dabei nach folgender Beziehung vorgenommen:

$$\hat{f}(\mathbf{x}, t) = \frac{1}{M} \sum_{i=1}^M f_i(t - \Delta_i) \quad (1)$$

Dabei bezeichnet  $t$  die Zeit,  $M$  ist die Anzahl der Mikrofone im Array, die  $f_i$  sind die aufgezeichneten Zeitfunktionen der einzelnen Mikrofone an den Positionen  $\mathbf{x}_i$ , und die  $\Delta_i$  sind die entsprechenden relativen Zeitverzögerungen (also Laufzeitdifferenzen), welche aus den absoluten Laufzeiten  $\tau_i$  gemäß  $\Delta_i = \tau_i - \min(\tau_i)$  berechnet werden. Die absoluten Laufzeiten  $\tau_i$  ergeben sich trivial aus der Beziehung  $\tau_i = |\mathbf{r}_i|/c$ , wobei  $c$  die Schallgeschwindigkeit in Luft ( $c \approx 344$  m/s bei 20°C) ist und  $|\mathbf{r}_i| = |\mathbf{x}_i - \mathbf{x}|$  den geometrischen Abstand von Mikrofon Nummer  $i$  zum gerade berechneten Quellort  $\mathbf{x}$  in der Bildebene bezeichnet. Die einzelnen Mikrofon-Zeitfunktionen  $\pi_i$  werden also erst laufzeitkorrigiert und dann zeitkohärent aufaddiert. Das so verstärkte Signal wird danach wieder auf die Anzahl  $M$  der Sensoren normiert, um ein Ergebnis zu erhalten, welches vom energetischen Gehalt dem der Zeitfunktion eines einzelnen Mikrofons entspricht (Bild 2).

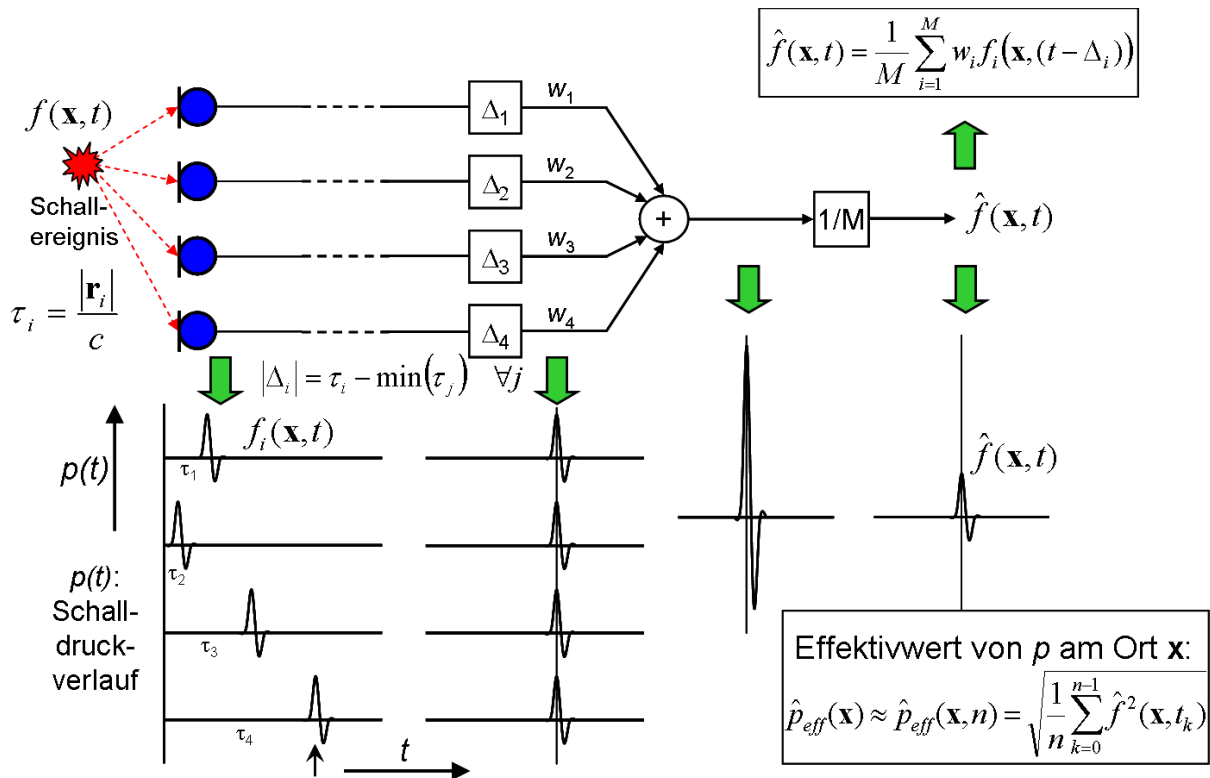


Bild 2: Zum Funktionsprinzip der Akustischen Kamera (Bilder/Beamforming\_Prinzip.emf)

Die Gleichung (1) ist (mit zusätzlichen Wichtungen  $w_i$  der  $f_i$ ) in der Literatur [1] unter der Bezeichnung „Time Domain Delay-and-Sum-Beamforming“ bekannt. Aus der so rekonstruierten Zeitfunktion  $y = \hat{f}(\mathbf{x}, t)$  wird nun die Schätzung des Effektivwertes des Schalldruckes am Ort  $\mathbf{x}$  über eine endliche Anzahl von  $n$  zeitdiskreten Abtastzeitpunkten  $t_k$  berechnet:

$$\hat{p}_{\text{eff}}(\mathbf{x}) = \sqrt{\frac{1}{n} \sum_{k=0}^{n-1} \hat{f}^2(\mathbf{x}, t_k)}. \quad (2)$$

Die einzelnen Effektivwerte aller Orte  $\mathbf{x}$  nach (2) werden abschließend entsprechend einer vorgegebenen Skala farbkodiert und wie oben beschrieben als akustische Karte am Bildschirm dargestellt. Dieses Prinzip lässt sich unverändert auch auf nichtplanare Arrays anwenden.

## Auswertung im Zeit- und Frequenzbereich

Unter Nutzung des Verschiebungssatzes der Fouriertransformation kann Gleichung (1) ebenso im Frequenzbereich berechnet werden. Die meisten in der klassischen Literatur zur Array-Signalverarbeitung und zum Beamforming beschriebenen Verfahren arbeiten traditionsgemäß nur in dieser Domäne. Sie bietet viele Vorteile, will man z.B. weiterentwickelte Signalverarbeitungsalgorithmen (adaptive oder robuste Verfahren, Eigenraumzerlegungen etc.) anwenden. Die Mikrofonsignale sind im Frequenzbereich nur noch über einfache komplexe Multiplikationen und nicht mehr über Faltungen wie im Zeitbereich mit der Ortsinformation verknüpft. Bei vielen spezielleren Algorithmen

werden darüber hinaus noch Schmalbandannahmen gemacht, da es nur für eine einzelne Frequenzlinie möglich ist, eine konstante Zeitverzögerung auch als eine einzige konstante Phasenverschiebung darzustellen. Für breitbandige Signale ergibt konstantes  $\Delta t$  für jede einzelne Spektralkomponente eine spezielle Phasenverzögerung. Deshalb wird der Rechenaufwand im Frequenzbereich erheblich höher, will man Breitbandsignale auswerten. Der Aufwand wächst außerdem im Zeitbereich nur linear mit der Mikrofonzahl, im Frequenzbereich dagegen quadratisch, bedingt durch die Verwendung komplexer Kreuzspektralmatrizen. Beide Effekte zusammen geben der direkten Implementierung im Zeitbereich einen erheblichen Effizienzvorteil, welcher bei der Akustischen Kamera der GFaI gerade bei hochkanaligen Messungen deutlich zum Tragen kommt. Das Zeitbereichsverfahren besitzt außerdem den Vorteil, dass es explizit bereits ein Breitbandverfahren ist, da immer alle Frequenzkomponenten im Signal enthalten sind. Vorheriges Herausfiltern störender, vorrangig tiefer Frequenzen und nachträgliches Beamforming im Zeitbereich sind einfach und ohne Veränderung des Rechenalgorithmus realisierbar.

Diese Vorteile werden ergänzt durch weniger örtliche Aliasing-Probleme als bei Schmalbandbetrachtungen. Unter Aliasing versteht man das Auftreten von signalfrequenzabhängigen Nebenmaxima (auch als Nebenkeulen oder „Sidelobes“ bezeichnet) um den Ort der eigentlichen Quelle. Die genaue Ausprägung dieser Aliasing-Muster ist dabei eine komplizierte Funktion, welche vorrangig von der Arraygröße und -geometrie sowie der Kanalzahl abhängig ist. Eine wichtige Kenngröße ist hier der sogenannte Arraykontrast, welcher die Differenz zwischen der Hauptkeule des Arraymusters und dem Maximum der ersten Nebenkeule angibt. Für ein rein hypothetisches Ringarray mit unendlich vielen Mikrofonen liegt diese Kontrastgrenze theoretisch bei -7,9 dB, siehe hierzu auch [2]. Das bisher an der GFaI standardmäßig eingesetzte Ringarray mit 32 Kanälen erreicht diese Grenze nicht, hier tritt praktisch bereits ab 5,5 dB bis 6 dB Aliasing auf. Beim künftig eingesetzten neuen Standardring mit 48 Mikrofonkanälen wird sich dieser Kontrast-Wert auf etwa 6,5 dB bis 7 dB steigern lassen. Die gewählte Ringgeometrie stellt für die praktische Anwendung insofern einen guten Kompromiss dar, als sie bei für viele Fälle ausreichendem Kontrast die schmalste Hauptkeule besitzt, also die beste Ortsauflösung aufweist. Höhere Kontrastwerte sind in der Regel nur mit hochkanaligen, in ihrer Geometrie entsprechend aufwendig optimierten Arrays möglich. Da die Schallquelle selbst in den meisten Fällen ortsfest ist (ihre Lage ist nicht abhängig von der Frequenz), die Aliasing-Muster aber ihre Lage im Bildfeld frequenzabhängig verschieben, mitteln sich die Aliasing-Figuren breitbandiger Signale annähernd aus. Lediglich bei den tiefen Frequenzen ist damit eine Verschlechterung des Bildkontrastes verbunden. Grundsätzlich empfiehlt es sich deshalb, die für die gegebene Arraygröße zu tief liegenden Störfrequenzen vor der Berechnung der akustischen Karten herauszufiltern.

Die Rechnung im Zeitbereich besitzt immer dann besondere Vorteile, wenn die Signale stark impulshaltig und transient sind. Transiente Signale lassen sich zwar grundsätzlich einfach lokalisieren, im Frequenzbereich wird hier jedoch aus einem einfachen Problem ein unnötig kompliziertes gemacht. Zunächst müssen willkürlich Zeitabschnitte und Überlappungsbereiche für die Kurzzeit-Fourier-Transformationen festgelegt werden, welche die erreichbare Zeitauflösung begrenzen. Die Signale sind mit entsprechenden Fensterfunktionen zu wichten, was ihren tatsächlichen Energiegehalt verringert, und anschließend in den Frequenzbereich zu transformieren. Kreuzspektralmatrizen müssen aus den Daten geschätzt werden, da diese Schätzung aber insbesondere bei kurzem Zeitintervall nicht konsistent ist, ist eine Mittelung über mehrere Zeitabschnitte

erforderlich, was wiederum die Detailinformation in den Pegeln sowie die erreichbare Zeitauflösung noch weiter reduziert und gewisse Stationaritätsannahmen voraussetzt, welche aber gerade bei den gesuchten Transienten meist nicht gegeben sind. Die spektrale Breitbandigkeit erfordert das aufwendige Analysieren sehr vieler Spektrallinien bzw. die Zusammenfassung der Daten mehrerer Frequenzbänder. In Verbindung mit niedrigen Abtastraten führen diese Frequenzbereichsverfahren für transiente Signale oft nicht zum gewünschten Ergebnis. Auch sind die Emissionen vieler technischer Geräte und Anlagen überwiegend breitbandig und eher selten nur aus ganz wenigen tonalen Komponenten zusammengesetzt, so dass ein reines Zeitbereichsverfahren zur Schallquellenlokalisierung auch hier sehr vorteilhaft einsetzbar ist.

Bedingung für ein örtlich hochauflösendes Beamforming ist jedoch eine sehr hohe zeitliche Abtastrate. Diese Voraussetzung ist zunächst unabhängig von der Domäne, in welcher die Berechnung vorgenommen werden soll. Die auf das Array eintreffende Schallwelle hat bedingt durch die Projektion auf die Arrayebene fast immer eine höhere scheinbare Laufgeschwindigkeit zwischen den einzelnen Sensoren, als es ihrer tatsächlichen Ausbreitungsgeschwindigkeit entspricht. Somit müssen extrem kurze absolute Laufzeitdifferenzen detektierbar sein, um die durch die Diskretisierung bedingten relativen Laufzeit- bzw. Phasenfehler klein zu halten. Nimmt man die Abtastung der einzelnen Mikrofonkanäle hier nur mit der für das Quellsignal nach Nyquist unbedingt nötigen Rate vor, würde das zu sehr schlechten Winkel- bzw. Ortsauflösungen und sehr stark frequenz- und winkelabhängigen, instabil schwankenden und analytisch nicht eindeutig bestimmbar Aliasingmustern und damit zu erheblich verschlechtertem Signal-Störabstand in den akustischen Bildern führen. Die diskrete Zeitauflösung der Einzelkanäle muss bei Rechnung im Zeitbereich für eine möglichst gute Ortszuordnung wesentlich (ca. Faktor 10) höher sein als die oberste im Signal zu kartierende Grenzfrequenz.

## Anwendungsbeispiele

Die Möglichkeit der Analyse sehr kurzer Impulse im Zeitbereich eröffnet neue, z.T. exotisch erscheinende Anwendungsfelder für die Technik der Akustischen Kamera. So zeigt Bild 3 das Spektrogramm eines Mikrofonkanals einer Aufnahme einer ägyptischen Grabfledermaus.

Bild 3: Spektrogramm der Chirpsignale einer ägyptischen Grabfledermaus. Auffällig sind die Emissionen in mehreren getrennten Frequenzbändern.  
(Bilder/Bat\_Spektrogramm.jpg)

Erkennbar sind extrem breitbandige und hochfrequente Emissionen in verschiedenen Teilbändern, welche dem Tier offenbar zu unterschiedlichen Erkennungs- und Klassifikationszwecken dienen. Die Chirps besitzen eine Zeitdauer von nur etwa 5 ms bis 10 ms. Die akustische Kartierung im Zeitbereich (über alle enthaltenen Frequenzen) zeigt eine sehr präzise Ortung der Schallquelle (Bild 4 Mitte). Die Aufnahme des Tieres (Bild 4 links) erfolgte in einem Innenraum aus einem Meter Entfernung mit dem klassischen Standard-Ringarray (s. Bild 1, 32 Mikrofone und 72,5 cm Durchmesser) bei einer Abtastrate von 192 kHz. Wertet man nur das untere Frequenzband um 15 kHz aus, tritt verstärktes Aliasing auf (Bild 4 rechts). Ein ähnliches Ergebnis würde man auch bei geringerer Systemabtastrate (48 kHz) erhalten. Eine Auswertung dieser

niederfrequenten Chirp-Pulse mit Beamforming im Frequenzbereich wäre aber mit einer FFT-Länge von 512 ebenfalls noch möglich.

--	--	--

Bild 4: Akustische Bilder einer ägyptischen Grabfledermaus (links), Auswertung über alle Frequenzen (Mitte) und gefiltert von 14 kHz bis 18 kHz (rechts). Die Schmalbandauswertung zeigt deutlich stärkeres Aliasing bei gleichem Bildkontrast von jeweils 7dB.

(zusammengetzt aus: Bilder/Bat\_Photo1\_1\_all\_optBild.jpg,  
Bilder/Bat\_Photo1\_1\_all\_akuPhoto7dB.jpg und  
Bilder/Bat\_Photo1\_14bis18kHz\_akuPhoto7dB.jpg)

Die Möglichkeiten der enormen zeitlichen Auflösung, die mit der Auswertung im Zeitbereich verbunden sind, verdeutlicht ein Beispiel aus dem Technikbereich. Dabei wurde ein Feld von Zündkerzen (Bild 5), mit dem die Möglichkeiten der zeitgenauen Zündung demonstriert werden können, mit der Akustischen Kamera (Konfiguration wie im ersten Beispiel) aufgenommen.

Bild 5 Zündkerzenfeld mit Ansteuerung jeder einzelnen Zündkerze  
(Bilder/Foto\_Zuendfeld\_ausschnitt.jpg)

Bild 6 Daten der 32 Mikrofone (Ausschnitt über 12 Zündungen)  
(Bilder/Mik\_Signale\_Zuendk.jpg)

Bild 6 zeigt einen Ausschnitt aus den aufgezeichneten 32 Mikrofonkanälen (genau 12 Zündungen, je Zündkerze eine Zündung). Markiert ist ein Zeitbereich über 68,8 ms, für den ein Akustisches Foto (Bild 7) errechnet wurde.

Bild7 Akustisches Foto über den im Bild 6 markierten Bereich  
(Bilder/akust\_foto\_kd.jpg)

Für einen besseren Bildkontrast wurde das akustische Foto auf ein per Kantenextraktion bearbeitetes Bild der Messszene projiziert. Sehr schön sind alle zwölf Emissionen bei einem Kontrast von 5 dB zu erkennen. Selbst eine Reflexion an einer Montagekante wird deutlich.

Bild 8 Signal eines Mikrofons (Zündung genau einer Kerze)  
(Bilder/Mik\_Signal\_1Zuend\_1Mik.jpg)

Im Bild 8 wird das Signal gezeigt, dass von einem Mikrofon für eine Zündung aufgezeichnet wurde. Die Ursprungsanregung (der markierte Bereich) ist nur 0,18 ms lang. Durch Laufzeitunterschiede wird dieser Bereich auf 0,31 ms (bei der Berücksichtigung aller Mikrofone) verlängert.

Bild 9 Akustisches Foto einer Zündung (Bilder/akust\_foto\_1Zuend\_kd.jpg)

Das Akustische Foto einer einzelnen Zündung zeigt Bild 9 (berechnet über 0,31 ms). Bei einem Kontrast von 5 dB tritt noch kein Aliasing auf. Dieses Beispiel zeigt

eindrucksvoll die Stärken der Akustischen Kamera bei der Ortung transienter akustischer Ereignisse.

Weitere Beispiele für die Funktionalität und den technischen Einsatz der Akustischen Kamera sollen in einem Folgebeitrag im Sensor Report 2/2007 vorgestellt werden.

L0711

[www.gfai.de](http://www.gfai.de)

## **Zu Personen**

Dr. Ralf Schröder ist promovierter Physiker. Er arbeitet seit 2003 in der GFaI und ist seit Anfang 2005 Leiter des Bereiches Signalverarbeitung, in dem die Akustische Kamera seit den 90-er Jahren entwickelt wurde.

(Bilder/schroeder\_color.jpg)

Dr.-Ing. Olaf Jaeckel ist seit 1999 an der GFaI und seit 2003 als wissenschaftlicher Mitarbeiter in der Arbeitsgruppe Signalverarbeitung tätig. Er ist vorrangig für die Betreuung verschiedener Forschungsprojekte zur Akustischen Kamera verantwortlich.

(Bilder/jaeckel\_color.jpg)

## **Literatur**

- [1] Johnson, D.H.; Dudgeon, D.E.:  
Array Signal Processing. Concepts and Techniques.  
PTR Prentice Hall 1993.
- [2] Van Trees, H.L.:  
Optimum Array Processing.  
J. Wiley & Sons 2002.