

Microphone arrays and distributed microphone arrays in a vehicle

Oliver Kirstein

Ford Werke Köln GmbH, 50725 Köln, E-Mail: okirstei@ford.com

Einleitung

Der Einzug neuer Technologien in das Fahrzeug wird unweigerlich zu Veränderungen der Kommunikation zwischen Insassen untereinander führen. Dies gilt auch für die Kommunikation zwischen Fahrzeug und Insassen sowie über die Grenzen des Fahrzeugs hinaus.

Beim autonomen Fahren beispielsweise wird die Kommunikation in alle Richtungen stattfinden können, da es niemanden mehr gibt, dessen Hauptaufgabe im Führen des Fahrzeugs besteht. Zudem interagiert nicht mehr der Fahrer unter Verwendung von Sprachbefehlen mit dem Fahrzeug, sondern alle Insassen könnten gleichberechtigt mit dem Fahrzeug kommunizieren.

Für das Fahrzeug ist hierbei entscheidend zu wissen, wer gerade spricht. Dabei kann ein Triangulationsalgorithmus unter Verwendung eines lokalen oder auch verteilten Mikrofonarrays hilfreich sein, welcher unempfindlich gegenüber Fahr- und anderen Nebengeräuschen sein muss.

Stand Heute

Die Angebote für verschiedenste Audiosysteme der Fahrzeughersteller belegen, dass der qualitativ hochwertigen Musikwiedergabe im Fahrzeug eine hohe Bedeutung beigemessen wird. Um dies sicherzustellen werden schon einmal 23 oder sogar noch mehr Lautsprecher verbaut.

Den vielen Lautsprechern steht häufig aber nur ein Mikrofon zur fahroptimierten Sprachaufnahme entgegen.

Befinden sich drei weitere Mikrofone in der zweiten / dritten Sitzreihe des Fahrzeugs, kann man mit hoher Sicherheit davon ausgehen, dass diese derzeit nur zur Minimierung von Innengeräuschen – Active Noise Cancellation (ANC) – verwendet werden.

Dieses Mikrofon-Setup ist in keiner Weise für die obige Aufgabenstellung geeignet, da der Beifahrer nicht berücksichtigt werden kann und die ANC-Mikrofone hinten über keine Richtcharakteristik verfügen.

Sprachaufnahme für alle Sitzplätze

Der schwierige Teil, die Integration der Mikrofone in die Fahrzeugumgebung, wurde im untersuchten Fahrzeug aber bereits gelöst. Für diese Untersuchung wurden somit die Positionen der ANC Mikrofone für die mittleren bzw. hinteren Sitzpositionen verwendet, aber mit vorhandenen Mikrofon-Arrays ausgestattet. Nur die Position für den Fahrer und Beifahrer leicht verändert (siehe Abbildung 1).

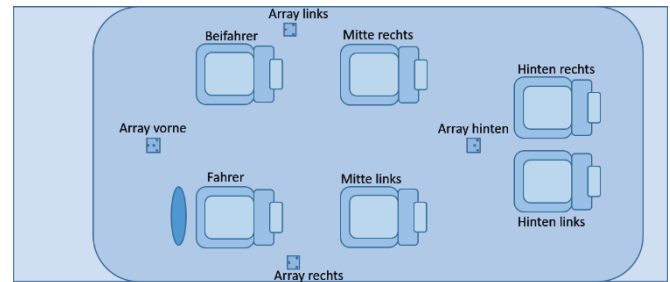


Abbildung 1: Diese zeigt schematisch die Position der einzelnen Mikrofon Arrays in Bezug auf die Sitzpositionen.

Die Mikrofon-Arrays im Detail:

jedes einzelne Mikrofon ist omni-direktional und ihr Abstand variiert zwischen 12mm und 21mm.

Array vorne: in zentraler Position zwischen Fahrer und Beifahrer, so dass eine symmetrische Abdeckung erreicht werden kann.

Array links/rechts: je eins pro Seite, damit keine Kompromisse zwischen ANC und Sprachaufnahme eingegangen werden müssen. Zudem ist die Wahl der Position durch das Panorama-Glasdach sowieso auf die Dachränder beschränkt.

Array hinten: in zentraler Position zwischen dem rechten und linken Passagier, sodass auch hier eine symmetrische Abdeckung erreicht werden kann.

Zeitliche Verzögerung und Einfallswinkel (Triangulation)

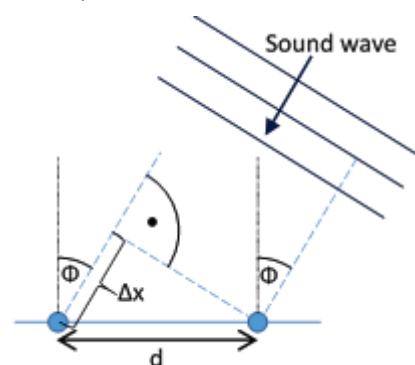


Abbildung 2: Zusammenhang zwischen zusätzlich zurückzulegenden Weg Δx und Einfallswinkel Φ unter der Annahme von Fernfeld.

$$\sin \Phi = c_{\text{air}} \frac{mT}{d} \quad (1)$$

Für Gleichung (1) gilt: $c_{\text{air}}=343,2$ [m/s], $T = 1/f_s$ [s], d [m] ist der Abstand der beiden Mikrofone und m die Anzahl der Samples der zeitlichen Verzögerung.

Verwendeter Ansatz zur Sprecherlokalisierung

Die Sprecherlokalisierung basiert auf der zeitlichen Kreuzkorrelation [1] der Signale an zwei Mikrofonen. Die verwendete Abtastrate (f_s) liegt bei 48 kHz.

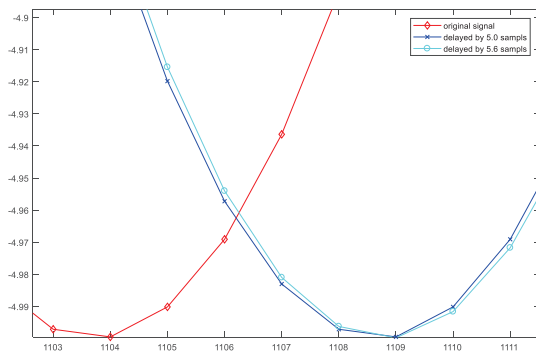


Abbildung 3: Zeitliche Darstellung von den Signalen am Mikrofonpärchen, Mikrofon 2 (cyan) läuft mit 5.6 Samples Mikrofon 1 (rot) nach. Der durch die Kreuzkorrelation gefundene Versatz von 5 Samples ist in dunkel blau gezeigt. X-Achse in Samples, Y-Achse ist die Amplitude des Audiosignales.

Wie in Abbildung 3 zu sehen ist, ist die Genauigkeit nicht hinreichend gut, da die Kreuzkorrelation nur volle Samples finden kann. Um diese zu verbessern wurde eine Spline Interpolation verwendet und somit die Abtastrate auf 480 kHz erhöht. Abbildung 4 zeigt ebenfalls eine Messung mit einem Mikrofonpärchen mit Interpolation, nun ist es möglich auch die erste Nachkommastelle der zeitlichen Verzögerung (in Samples) in der Kreuzkorrelation zu bestimmen.

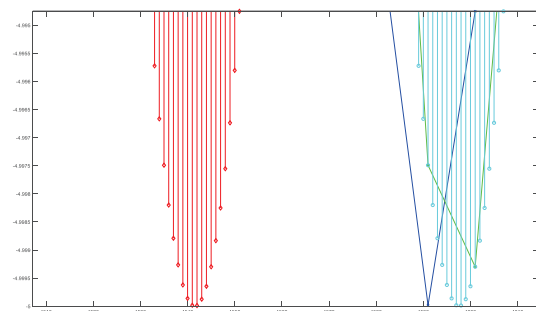


Abbildung 4: Diese zeigt das Signal an Mikrofon 1 (rot) mit Interpolation und das an Mikrofon 2 (cyan) mit Interpolation. Der zuvor gefundene Versatz von 5.0 Samples (dunkel blau) und der nun gefundene Versatz von 5.6 Samples (grün) werden in externer Samplerate von 48kHz gezeigt.

Dieses Interpolationskonzept mit anschließender Neuabtastung wird auch in der Simulation für die Berechnung der Eingangssignale an jedem Mikrofon des Arrays verwendet.

Abbildungen 5 und 6 zeigen die Auswirkung der Interpolation auf die Bestimmung des Einfallswinkels.

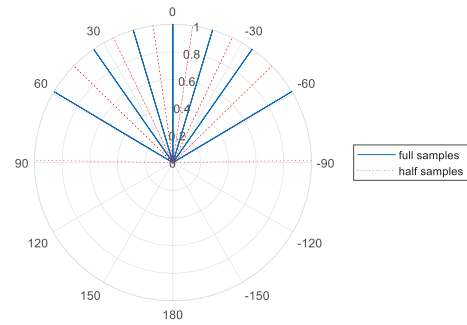


Abbildung 5: Umwandlung der zeitlichen Verzögerung in Einfallswinkelgruppen (Abstand der Mikrofone d ist 25mm). Z.B. bei einer Zeitdifferenz von 0 Samples ergibt sich ein Winkel von -7.5° bis $+7.5^\circ$, da -0.5 bis $+0.5$ Sample nicht unterschieden werden können.

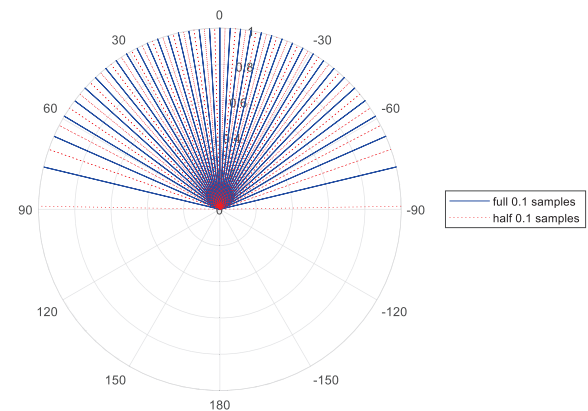


Abbildung 6: Umwandlung der zeitlichen Verzögerung in Einfallswinkelgruppen mit Interpolation auf 480 kHz.

Zur Unterscheidung von Sprachsignal und Hintergrundgeräusch wurden zwei Ansätze evaluiert. Beide zeigen positive Ergebnisse.

- 1) Begrenzung des Signals schmalbandig um die Vokal-Formanten (320 Hz, 500 Hz, 1 kHz).
- 2) Reduktion des Hintergrundgeräusches mit Hilfe eines Wiener Filters.

Ansatz 1 zeigt jedoch eine deutliche Sprachabhängigkeit und eine zeitliche Verzögerung bis der erste Vokal gefunden wird. Der Ansatz 2 benötigt auf den Fahrzeuginnenraum kalibrierte Koeffizienten. Deren Bestimmung für mehrere Fahrzeuge und Varianten (Limousine, Kombi, etc.) bringt einen erheblichen Aufwand mit sich. Weitere Untersuchungen sind geplant um eine praktikable Lösung zu finden.

Simulationsumgebung

Mit Hilfe einer vorhandenen Freifeld-Simulation und unter Verwendung des zuvor beschriebenen Lokalisierungsansatzes wurden verschiedene Sprecherpositionen und verschiedene Hintergrundgeräusche (von Position Noise) untersucht. Zudem wurden auch verschiedene lokale und verteilte Mikrofon-Array Konstellationen (blau) untersucht. Der Nullpunkt des Koordinaten Systems des Fahrzeuges in

den Abbildungen 7, 8 und 9 liegt in der Mitte der Stoßstange vorne, die Einheit ist [mm]. Fahrer (**Drv 50%**) und Beifahrer (**Pass**) sind in der zu erwartenden Position Sprachquelle. **Rear left** und **Rear right** beziehen sich auf die Passagiere in der mittleren Sitzreihe. Eine Hilfslinie (**gelb**) markiert die direkte Richtung von der Quelle zu der Mitte des Arrays.

Ergebnis

Es werden einige ausgewählte Szenarien gezeigt:

Lokales Array

In der Konsole im Dachhimmel (vorne)

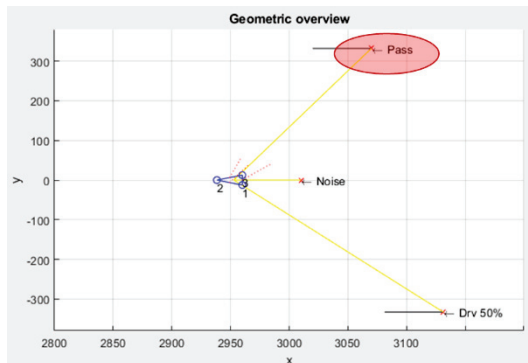


Abbildung 7: Dies zeigt den Fall, dass der Beifahrer (**Pass**) als Quelle identifiziert wurde. Die gefundene Richtung jedes Mikrofonpärchens ist in rot-gestrichelt gezeigt.

Im Dachhimmel über der Kopfstütze der Frontsitze

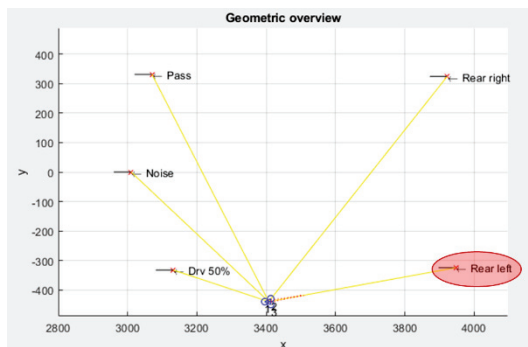


Abbildung 8: Dies zeigt den Fall, dass der Passagier Mitte links (**Rear left**) als Quelle identifiziert wurde. Die gefundene Richtung jedes Mikrofonpärchens ist in rot-gestrichelt gezeigt.

Verteiltes Array

Kombination eines Teils des Arrays vorne mit einem Teil des Arrays hinten



Abbildung 9: Dies zeigt den Fall, dass der Passagier Mitte links (**Rear left**) als Quelle identifiziert wurde. Die gefundene Richtung jedes Mikrofonpärchens ist in rot-gestrichelt gezeigt.

Für ein Mikrofonpärchen ist aufgrund der Symmetrie eine Aussage für die Position der Quelle vor oder hinter den Array unterbestimmt. Unter der Annahme, dass sich die Sprachquelle innerhalb des Fahrzeuges befindet und Hinzuziehen der Ergebnisse anderer Pärchen des Arrays oder anderer Arrays kann jedoch auf die richtige Seite geschlossen werden.

Ausblick

Untersuchung von Beamforming auf die Sprachaufnahme sowie Integration der Simulation in die akustische Gesamtfahrzeuginnenraum-Simulation.

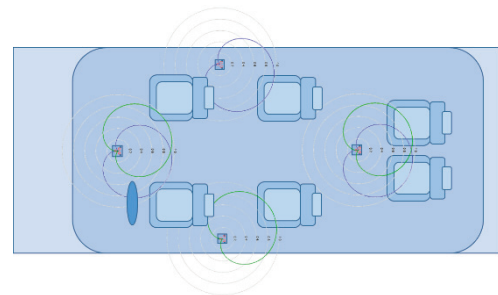


Abbildung 10: Hier werden die Richtcharakteristik eines möglichen Beamformers gezeigt (1kHz).

Literatur

- [1] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," in IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 24, pp. 320–327, 1976.