Vergleich von PCA und MVDR im Falle inkohärenter Rauschumgebung

Wolf Baumann[1], Patrick Vicinus[2], Reinhold Orglmeister[1]

[1] *TU Berlin, Fachgebiet Elektronik und medizinische Signalverarbeitung, email: w.baumann@ee.tu-berlin.de*

[2] *Peiker acustic GmbH, email: patrick.vicinus@peiker.de*

Einleitung

Methoden der mehrkanaligen Signalverarbeitung etablieren sich zunehmend als Standard bei sprachakustischen Anwendungen. Sie sind in der Lage, sowohl zeitliche als auch räumliche Signaleigenschaften zu berücksichtigen und werden im allgemeinen als Beamforming-Verfahren bezeichnet. Die Berücksichtigung der räumlichen Quellenverteilung erfolgt dabei im Fall von Fernfeld-Bedingungen über die Auswertung von Laufzeit- bzw. Phasenunterschieden zwischen den Mikrofonen.

Das einfachste Verfahren dieser Art stellt der Delay-and-Sum-Beamformer dar, dessen Ausgangsignal durch Summation der laufzeitkompensierten Sensorsignale ensteht. Für vollständig inkohärentes Rauschen stellt dieses Verfahren die optimale Lösung im Sinne des besten Signal-Rauschabstandes (SNR) dar [Bitzer and Simmer, 2001]. Es wird deshalb oft im nichtadaptiven Zweig von Beamformern mit GSC-Struktur (Generalized Sidelobe Canceller) verwendet.

Für den Einsatz in realen Umgebungen muss das Modell der Signalerfassung verallgemeinert werden, da sich die Übertragungsfunktion zwischen Signalquelle und Mikrofon auf die Laufzeiten und Amplituden der Sensorsignale auswirkt. Durch Berücksichtigung der Übertragungsfunktion, die sich bei unkorreliertem Rauschen gleicher Leistung mit Hilfe von Statistik zweiter Ordnung schätzen läßt, kann eine verbesserte Rauschunterdrückung erreicht werden.

Die dabei erforderliche Sensorgewichtung läßt sich sowohl anhand der PCA-Lösung (Principal Component Analysis) als auch durch den Minimum-Variance-Distortionless-Response (MVDR) Beamformer-Entwurf herleiten, da sich beide Ansätze lediglich in der Signalskalierung unterscheiden. Dieser Zusammenhang soll nachfolgend anhand eines zweikanaligen Modells erläutert werden.

1 Modell

Das zweikanalige Modell der Signalerfassung läßt sich im Zeitbereich unter Berücksichtigung der Impulsantwort $h_i(t)$ zwischen Quelle und Mikrofon i wie folgt formulieren:

$$x_1(t) = h_1(t) * s(t) + n_1(t) \tag{1}$$
$$x_2(t) = h_2(t) * s(t) + n_2(t), \tag{2}$$

wobei $s(t)$ das Nutzsignal und $n_i(t)$ das jeweilige Rauschsignal darstellen. Bei Annahme eines Referenzsensors bzw. bei Normierung auf Mikrofon 1 vereinfacht sich das Modell zu

$$x_1(t) = s(t) + n_1(t) \tag{3}$$
$$x_2(t) = h(t) * s(t) + n_2(t). \tag{4}$$

Dementsprechend läßt sich das normierte Modell im Frequenzbereich als

$$X_1(\omega) = S(\omega) + N_1(\omega) \tag{5}$$
$$X_2(\omega) = H(\omega) \cdot S(\omega) + N_2(\omega). \tag{6}$$

angeben. Aus Gründen der Übersichtlichkeit wird des Weiteren auf das Mitführen der Parameter verzichtet. D.h., es gilt nachfolgend in jedem Frequenzband

$$X_1 = S + N_1 \tag{7}$$
$$X_2 = H \cdot S + N_2 \tag{8}$$

bzw. in Matrixschreibweise mit $\mathbf{A} = [1\ H]^T$ und $\mathbf{N} = [N_1\ N_2]^T$

$$\mathbf{X} = \mathbf{AS} + \mathbf{N}. \tag{9}$$

2 Schätzung der Übertragungsfunktion

Die Schätzung der unbekannten komplexwertigen Übertragungsfunktion kann in jedem Frequenzband durch die Auswertung der Statistik zweiter Ordnung erfolgen. Die Eigenwertzerlegung der Kovarianzmatrix (siehe Abschnitt 3) liefert eine zuverlässige Schätzung, wenn an den Mikrofonen die gleiche Rauschleistung vorherrscht. Abbildung 1 verdeutlicht diesen Zusammenhang anhand eines Streudiagramms für reellwertige Daten bei einem SNR von 0 dB am Referenzsensor.

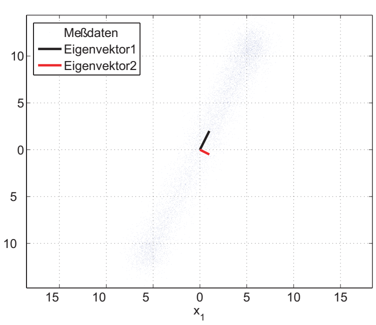

Abbildung 1: Eigenvektoren der Kovarianzmatrix im reellwertigen zweikanaligen Fall bei einer Übertragungsfunktion zwischen den Sensoren von H=2.

3 PCA

Ziel der PCA (*Principal Component Analysis*) oder Hauptkomponentenanalyse ist es, durch eine geeignete Transformation eine neue Darstellung der Meßsignale $\mathbf{x}$ in einem orthogonalen Koordinatensystem zu finden. Das heißt, die Kovarianzmatrix der transformierten Daten ist diagonal.

Mit der Transformationsmatrix $\mathbf{V}$ ergeben sich die transformierten Daten $\mathbf{Y}$ wie folgt (siehe z.B. [Köhler, 2005]):

$$\mathbf{Y} = \mathbf{V}^T\mathbf{X}. \tag{10}$$

Die Transformationsmatrix wird dabei so berechnet, daß der Signalvektor als gewichtete Überlagerung orthogonaler Basisvektoren aufgestellt werden kann. Dieses Approximationsproblem läßt sich mit Hilfe der Lagrange-Multiplikator-Methode lösen. Die Basisvektoren ergeben sich dann als Eigenvektoren der Kovarianzmatrix.

Für das Signalmodell aus Gleichung (9) sind die Eigenvektoren als Spalten der Transformationsmatrix von der Form

$$\mathbf{V} = \begin{bmatrix} 1 & 1 \\ H^* & -1/H \end{bmatrix}, \tag{11}$$

wobei sich zeigen läßt, daß der zum größten Eigenwert gehörende Eigenvektor durch $\mathbf{ev}_1 = [1 \quad H^*]^T$ gegeben ist. Für die daraus entstehende Signalkomponente läßt sich folglich die Vorschrift

$$Y_1 = X_1 + H^* X_2 \tag{12}$$

angeben. Damit ergibt sich für die einzelnen Signalanteile

$$Y_1 = S + N_1 + H^* \cdot HS + H^* N_2 \tag{13}$$

$$= S(1 + |H|^2) + N_1 + H^* N_2. \tag{14}$$

Soll das Nutzsignal mit einer im Vergleich zum Referenzsensor gleichen Amplitude berechnet werden, muss der zusätzliche Skalierungsfaktor

$$K = \frac{1}{1 + |H|^2} \tag{15}$$

in das Ausgangssignal multipliziert werden. Damit läßt sich der PCA-Koeffizientenvektor für das zweikanalige Modell folgendermaßen darstellen

$$\mathbf{w}_{PCA} = K \cdot [1 \quad H^*]^T. \tag{16}$$

4 MVDR

Der Minimum-Variance-Distortionless-Response-Beamformer (MVDR) stellt eine weit verbreitete Methode der mehrkanaligen Signalverarbeitung dar. Er kann in nichtadaptiver Form, z.B. als superdirektiver Beamformer [Bitzer and Simmer, 2001] oder als adaptives Verfahren ausgelegt sein.

Für die Berechnung der Beamformerkoeffizienten wird die Minimierung der Rauschleistung unter Einhaltung der Nebenbedingung einer ungestörten Übertragungsfunktion in Blickrichtung formuliert. Mit $\mathbf{\Phi}_{nn}$ als Matrix der Kreuzleistungsdichtespektren des Rauschens an den Sensoren ergibt sich für die Beamformerkoeffizienten [Simmer et al., 2001]

$$\mathbf{w}_{MVDR} = \frac{\mathbf{\Phi}_{nn}^{-1}\mathbf{d}}{\mathbf{d}^H \mathbf{\Phi}_{nn}^{-1}\mathbf{d}}. \tag{17}$$

Dabei charakterisiert $\mathbf{d}$ den Nutzsignalpfad, der meist in Form des Steering-Vektors durch frequenz- und richtungsabhängige Phasenfaktoren vorgegeben wird.

Für reale Umgebungen müssen statt einfacher Phasenverschiebungen die komplexwertigen Koeffizienten der Übertragungsfunktion berücksichtigt werden. Enstprechend Gleichung (9) nimmt der Vektor $\mathbf{d}$ dann die Form

$$\mathbf{d} = \mathbf{A} = [1 \quad H]^T \tag{18}$$

an. Bei inkohärenter Rauschumgebung mit gleicher Rauschleistung σ_n^2 an den Sensoren gilt

$$\mathbf{\Phi}_{nn} = \sigma_n^2 \cdot \mathbf{I}, \tag{19}$$

wobei $\mathbf{I}$ die Einheitsmatrix darstellt. Damit nimmt Gleichung (17) eine zur PCA-Lösung identische Form an

$$\mathbf{w}_{MVDR} = \frac{\mathbf{I}^{-1}\mathbf{d}}{\mathbf{d}^H\mathbf{I}^{-1}\mathbf{d}} = \frac{\mathbf{d}}{\mathbf{d}^H\mathbf{d}}, \tag{20}$$

da sich für beide Ansätze die gleiche Sensorgewichtung

$$\mathbf{w}_{MVDR} = \mathbf{w}_{PCA} = \frac{1}{1 + |H|^2} \cdot [1 \quad H^*]^T \tag{21}$$

ergibt.

5 Zusammenfassung

In der vorliegenden Arbeit wurde durch Berücksichtigung der Übertragungsfunktion das auf Laufzeitunterschieden basierende Modell des Delay-and-Sum-Beamformers für reale Umgebungen verallgemeinert. Die Schätzung der Übertragungsfunktion ist für inkohärentes Rauschen gleicher Leistung durch die Eigenwertzerlegung der Kovarianzmatrix möglich. Wie gezeigt wurde, lässt sich bei Kenntnis der Übertragungsfunktion die im Sinne des maximalen SNR optimale Sensorgewichtung sowohl durch die PCA als auch durch einen MVDR-Ansatz herleiten.

Literatur

[Bitzer and Simmer, 2001] Bitzer, J. and Simmer, K. U. (2001). *Superdirective Microphone Arrays*, chapter 2, pages 19–38. In [Brandstein and Ward, 2001].

[Brandstein and Ward, 2001] Brandstein, M. and Ward, D., editors (2001). *Microphone Arrays*. Springer, Berlin Heidelberg New York.

[Köhler, 2005] Köhler, B.-U. (2005). *Konzepte der statistischen Signalverarbeitung*. Springer, Berlin.

[Simmer et al., 2001] Simmer, K. U., Bitzer, J., and Marro, C. (2001). *Post-Filtering Techniques*, chapter 3, pages 39–60. In [Brandstein and Ward, 2001].