

Spektralanalyse und Frequenzkonturierung durch Filter mit asymmetrischen Flanken

ARNO UNKRIG, UWE BAUMANN

Fachgebiet Akustische Kommunikation, Technische Universität München

Einleitung

Übliche Spektralanalyseverfahren können spektrale Verdeckungseffekte nicht oder nur ungenau simulieren, da die Selektionscharakteristik des menschlichen Gehörs nicht richtig nachgebildet wird. Die oft verwendete Fast-Fourier-Transformation (FFT) weist für alle Analysefrequenzen dieselbe Bandbreite auf. Bei der Fourier-t-Transformation [4] (FTT) kann die Bandbreite zwar für jede Analysefrequenz separat gewählt werden, die unterschiedlichen Steilheiten der oberen und unteren Mithörschwellenflanke bleiben aber unberücksichtigt. Dies ist ein Nachteil aller Fouriertransformationen mit reellem Zeitfenster. Im vorliegenden Beitrag wird gezeigt, wie die Berücksichtigung der spektralen Verdeckung durch die Verwendung einer komplexen Fensterfunktion näherungsweise erreicht werden kann.

Spektrale Verdeckung

Nicht alle Töne, die in einem Schall objektiv vorhanden sind, sind auch tatsächlich hörbar; schwache Töne werden von stärkeren „verdeckt“ oder „maskiert“.

Zwicker hat diesen Effekt schon vor längerer Zeit quantitativ untersucht [5] und als sogenannte „Mithörschwellen“ dargestellt. In Abbildung 1 erkennt man, daß Töne niedrigerer Frequenz weniger verdeckt werden als Töne höherer Frequenz.

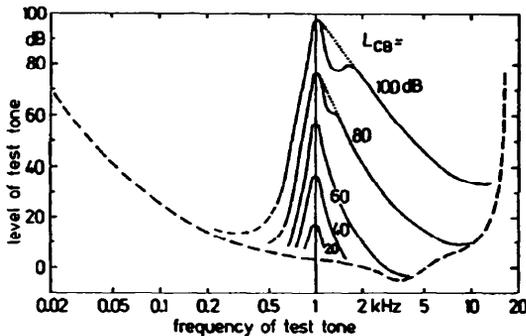


Abbildung 1: Mithörschwellen von Schmalbandrauschen (Mittenfrequenz 1 kHz) unterschiedlichen Pegels L_{CB} nach Zwicker [5].

Die Mithörschwelle läßt sich oberhalb der Ruheshörschwelle durch eine Dreiecksfunktion annähern [3]. Die linke Flanke weist eine Steigung von 27 dB / Bark auf, die Steigung der rechten hängt vom Pegel und von der Frequenz des Maskierers ab. Für einen Maskiererton mit 440 Hz und 55 dB ergibt sich zum Beispiel eine Steigung von $-13,5$ dB / Bark.

Untersucht man einen Schall, der aus dem erwähnten Maskierer und zwei weiteren Tönen besteht, so stellt man fest, daß sich für alle drei Töne Maxima im FTT-Spektrum ausbilden, obwohl die zwei schwächeren Töne subjektiv vom Maskierer verdeckt werden (Abbildung 2).

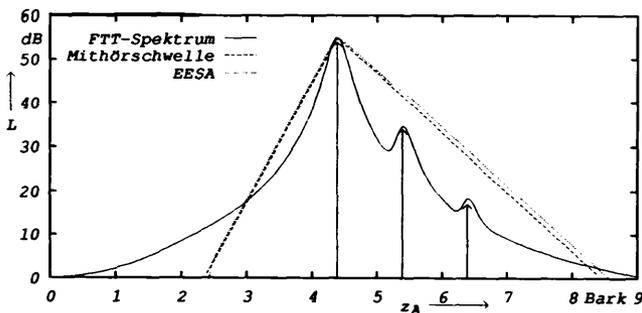


Abbildung 2: Mithörschwelle des Maskierertons und FTT-Spektrum (Parameter nach Heinbach [2]) der drei Töne.

Das läßt sich verhindern, wenn der Maskiererton ein Spektrum hervorruft, dessen Form der der Mithörschwelle gleicht (Abbildung 3). Damit gelangt man zu einer „Erregungs-äquivalenten Spektralanalyse“ („Excitation-Equivalent Spectral Analysis“ — EESA).

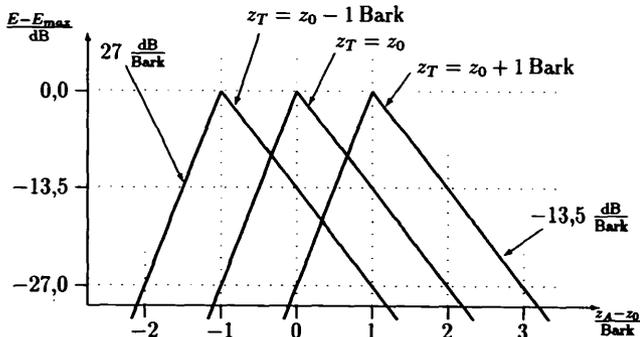


Abbildung 3: Vereinfachte Erregungsmuster einzelner Sinustöne mit unterschiedlichen Tonheiten z_T .

Der Spektralanalysator wurde als eine Bank von Analysefiltern realisiert. Aus dem Erregungsmuster (Abbildung 3) ergibt sich durch Vertauschung der Abszissengröße z_A mit dem Parameter z_T der Frequenzgang der Analysefilter (Abbildung 4).

Die Fensterfunktion für die Fourieranalyse läßt sich durch spektrale Verschiebung und Fouriertransformation des Verstärkungsfrequenzganges gewinnen. Abbildung 5 zeigt verschiedene Fensterformen im Vergleich.

Da die Analysefilter als lineare Systeme realisiert sind, wird die Pegel- und Frequenzabhängigkeit der Steigung der oberen Erregungsflanke nicht nachgebildet; sie wurde auf einen Wert von $-13,5 \text{ dB/Bark}$ festgelegt. Dies stellt einen guten Kompromiß für Maskierer mittlerer Tonhöhe und mittleren Pegels dar.

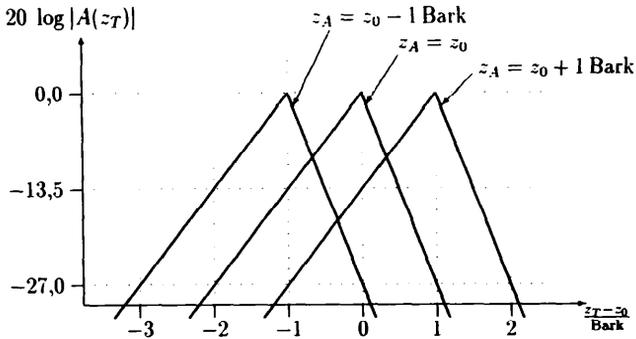


Abbildung 4: Analysefilterfrequenzgänge des vereinfachten Erregungsmodells.

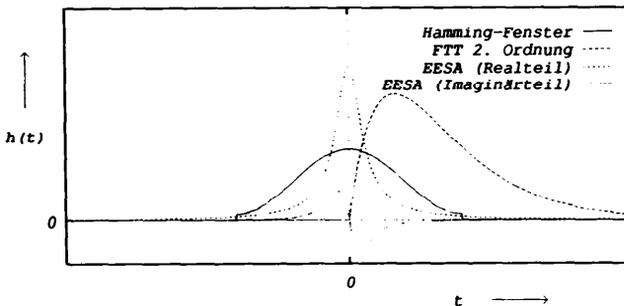


Abbildung 5: Hamming-Fenster, FTT-Fenster zweiter Ordnung und die komplexe Fensterfunktion im Vergleich (schematisch).

Die Filterbank wurde auf einer UNIX-Workstation (Silicon Graphics Iris Indigo) mit komplexen FIR-Filtern implementiert. Abbildung 6 zeigt die Amplitudenfrequenzgänge mehrerer Analysefilter; der gewünschte Verlauf wird in einem großen Dynamikbereich gut erreicht.

Um den Rechenaufwand zu reduzieren, wurden die Analysefilter in Teilbänke mit optimal angepasster Abtastfrequenz gruppiert. Dadurch konnten die sehr großen Ordnungen der FIR-Filter mit niedriger Analysefrequenz stark verringert werden. Als weitere Optimierung wurde die Faltung des Eingangssignals mit den FIR-Filterkoeffizienten durch eine „schnelle Faltung“ [1] beschleunigt.

Von den berechneten Spektren wird der Betrag gebildet und mit einem Tiefpaß geglättet. In Abbildung 2 ist das so berechnete Spektrum punktiert eingezeichnet. Die beiden schwächeren Töne verursachen keine lokalen Maxima im Spektrum, die Verdeckung wurde richtig simuliert.

Frequenzkonturierung

Von dem geglätteten Betragsspektrum werden Frequenzkonturen gebildet, um die Teiltonfrequenzen zu bestimmen. Dazu werden lokale Maxima im frequenzdiskreten Spektrum gesucht und eine Parabelinterpolation durch drei Punkte vorgenommen (Abbildung 7).

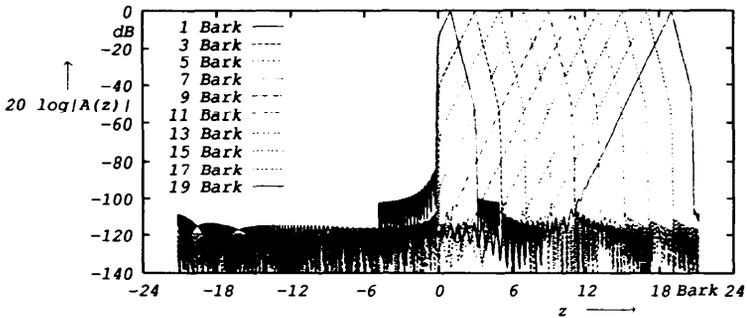


Abbildung 6: Amplitudenfrequenzgänge mehrerer Analysefilter ($z_A = 1, 3, 5, \dots, 19$ Bark).

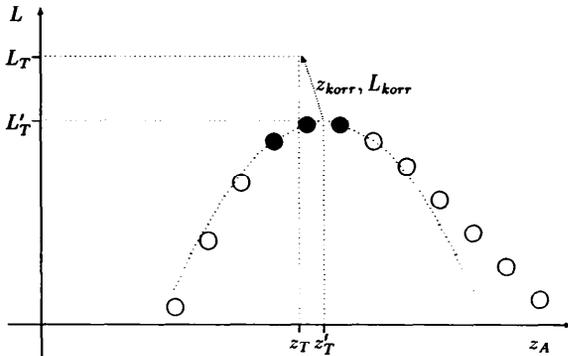


Abbildung 7: Schema zur Parabelinterpolation im frequenzdiskreten Spektrum.

Zu diesem interpolierten Maximum müssen noch z_{korr} und L_{korr} addiert werden, da die Amplitudenfrequenzgänge der Analysefilter in ihrem Maximum einen etwas „runden“ Verlauf aufweisen als in Abbildung 4 vorgegeben. Die erreichte Tonheitsgenauigkeit liegt bei 0,002 Bark, das bedeutet bei Testtonfrequenzen unter 500 Hz eine absolute Abweichung von bis zu 0,2 Hz, darüber eine relative Abweichung von bis zu 0,5 cent.

Literatur

- [1] Azizi, S. A. A.: Entwurf und Realisierung digitaler Filter. Oldenbourg 1981.
- [2] Heinbach, W.: Gehörgerechte Repräsentation von Audiosignalen durch das Teiltonzeitmuster. Dissertation am Lehrstuhl für Elektroakustik der Technischen Universität München 1988.
- [3] Terhardt, E.: Calculating Virtual Pitch. Hearing Research 1 (1979), S. 155-182.
- [4] Terhardt E.: Fourier Transformation of Time Signals: Conceptual Revision. Acustica 57 (1985), S. 242-256.
- [5] Zwicker, E., Fastl, H.: Psychoacoustics, Facts and Models. Springer-Verlag 1990.