

Verfahren zur Trennung und Auralisation harmonischer und stochastischer Signalanteile

Philipp Heck¹, Uwe Letens², Malte Kob¹

¹ Lehr- und Forschungsgebiet für Phoniatrie und Pädaudiologie

RWTH – Universitätsklinikum Aachen, Pauwelsstraße 30, D-52074 Aachen

² DaimlerChrysler AG, Entwicklung PKW (Mercedes-Benz), D-71059 Sindelfingen

Einleitung

Die Problematik der Signaltrennung in einen tonalen und einen stochastischen Anteil ist sowohl in der Fahrzeugakustik als auch im Bereich der Sprach- und Stimmakustik zu finden, dadurch ergab sich die Kooperation zwischen der DaimlerChrysler AG und dem Universitätsklinikum Aachen. Das im Rahmen dieses Projektes entwickelte Verfahren soll für Fahrzeuggeräusche optimiert werden und dabei mit möglichst wenigen Zusatzinformationen bezüglich des Fahrzeuges auskommen. Insbesondere wird keine Kenntnis der Motordrehzahl vorausgesetzt. Da diese jedoch grundlegend für die weiterführende Analyse ist, ist die Bestimmung der Grundfrequenz ein weiteres Ziel dieses Projektes. Nachdem verschiedene bekannte Analyse- und Detektionsverfahren aus dem Bereich der Stimm- und Sprachsignalverarbeitung nur unzureichende Ergebnisse lieferten, wurde ein neues Konzept verfolgt und umgesetzt.

Ein neues Grundfrequenzbestimmungsverfahren stellt die Basis für das nachfolgende Detektionsverfahren der tonalen Komponenten dar. Diese werden zu kontinuierlichen, harmonischen Zeitverläufen zusammengefasst. Die Signaltrennung wird einerseits durch zwei FIR-Filterbänke und andererseits mit einer Sinussynthese umgesetzt, wobei die Parameter dieser Signaltrennung aus den harmonischen Verläufen bestimmt werden.

Methode

Grundfrequenzbestimmung

Standardisierte Grundfrequenzbestimmungsverfahren wie beispielsweise das Autokorrelationsverfahren oder das Cepstrumverfahren [1] haben bei unseren Tests mit Fahrzeuggeräuschen keine zuverlässigen Ergebnisse erbracht. Dies ist darauf zurück zu führen, dass bei Fahrzeuggeräuschen teilweise nur sehr wenige oder sehr schwache harmonische Anteile hörbar und dadurch auch detektierbar sind. Aus diesem Grund wurde ein eigenes Bestimmungsverfahren entwickelt, dass auch bei stark verrauschten Signalen eine sichere Grundfrequenzbestimmung gewährleistet. Des Weiteren ist dieses Verfahren in der Lage aus Signalen mit verschiedenen Obertonreihen die zugehörigen verschiedenen Grundfrequenzen zu detektieren. Dies ermöglicht auch die Detektion der nicht von der Motordrehzahl abhängigen tonalen Komponenten. Um die Grundfrequenzen f_0 eines Signals $y(t)$ mit dem Spektrum $Y(f)$ zu bestimmen wird in diesem Verfahren in einem vorgegebenen Bereich für

alle möglichen Grundfrequenzen ein Modellspektrum $M(f, f_0)$ errechnet (Gleichung 1).

$$M(f, f_0) = \begin{cases} Y(n \cdot f_0) & \text{für } Y(n \cdot f_0) > N(f) \\ 0 & \text{sonst} \end{cases} \quad (1)$$

mit $n \in N$ und $f_{max} > f > f_{min}$

Dabei ist $N(f)$ eine Schätzung für den Rauschpegelverlauf des Signals. Die Grundfrequenzen werden an den Stellen gefunden, an denen die Funktion $R(f_0)$ lokale Maxima besitzt, die durch Gleichung 2 gegeben ist. Darin ist T die Anzahl der Spektrallinien von $M(f, f_0)$, deren Amplituden über dem Rauschpegelverlauf liegen.

$$R(f_0) = \frac{1}{T} \cdot \sum_{f=f_{min}}^{f_{max}} M(f, f_0) \quad (2)$$

In dem hier verfolgten Ansatz wird ein kontinuierlicher und stetiger Grundfrequenzverlauf vorausgesetzt, der aus den blockweise berechneten Grundfrequenzen interpoliert wird.

Detektion der tonalen Anteile

Um eine Entscheidung darüber treffen zu können, ob ein Signalanteil tonal ist, wird das Signal im Frequenzbereich mit einem geschätzten Rauschpegelverlauf verglichen. In den Bereichen, wo das Spektrum Maxima oberhalb des Rauschpegelverlaufes hat, werden tonale Anteile erkannt und deren Parameter wie Amplitude, Mittenfrequenz, SNR und Bandbreite extrahiert. Da als Grundlage für diese Erkennung die Grundfrequenz genutzt wird, kann der Detektionsbereich auf die Frequenzlinien im Bereich der ganzzahligen Vielfachen der Grundfrequenz eingeschränkt werden. In einem weiteren Schritt werden die erkannten Bereiche zeitlich zu harmonischen Verläufen zusammengefasst und bilden die Grundlage für die anschließende Synthese.

Signaltrennung und Synthese

Aus den Eigenschaften der harmonischen Verläufe werden die Filterkoeffizienten für eine FIR-Filterbank errechnet. Da einerseits eine phasentreue Überlagerung der blockweise gefilterten Signalabschnitte notwendig ist und andererseits der Rechenaufwand in akzeptablen Grenzen bleiben sollte, wurden die FIR-Filterkoeffizienten mit der „frequency response masking technique“ berechnet [2],

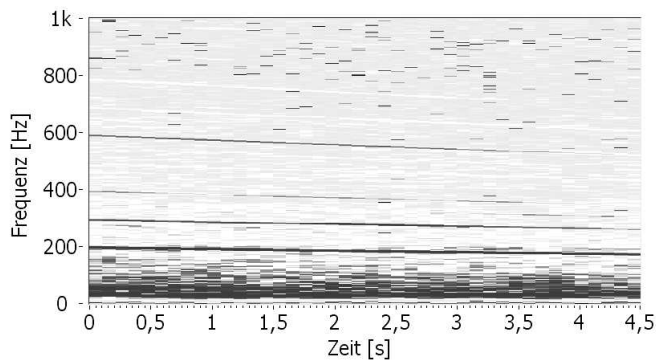


Abbildung 1: Spektrogramm eines Fahrzeuggeräusches, deutlich sind die Hauptmotorordnungen bei 200, 300 und 600 Hz zu sehen

dadurch wurde die Rechenzeit deutlich gesenkt. Die Filterbank extrahiert mit schmalbandigen Filtern den harmonischen und den stochastischen Anteil aus dem Signal. Um das harmonische Signal noch deutlicher zu extrahieren wurde neben der Filterung eine Sinussynthese implementiert. Aus den Informationen jedes harmonischen Verlaufes werden die Ausgangssignale von Sinusoszillatoren berechnet, die überlagert den rein tonalen Anteil des Originalsignals synthetisieren.

Ergebnisse

Die oben erläuterten Verfahren wurden mit der Programmierumgebung Labview implementiert und getestet. Dies ermöglicht eine gute Erweiterbarkeit der Implementierung und eine leichte Integration in andere Programmierumgebungen.

Die Grundfrequenz konnte bei allen getesteten Fahrzeuggeräuschen sicher bestimmt werden. Auch Fahrzeuginnengeräusche, die nur sehr gedämpfte tonale Anteile enthalten, konnten analysiert werden. In Abbildung 1 ist das Spektrogramm für ein exemplarisches Fahrzeuggeräusch zu sehen. Im Vergleich mit Abbildung 2, die das Spektrogramm des mit der Multisinussynthese generierten Signalanteils zeigt, sind die identischen harmonischen Frequenz- und Amplitudenverläufe zu erkennen. Die Filterbank synthetisiert den harmonischen Anteil natürlicher als die Sinussynthese, verursacht allerdings bei schnell veränderlichem Grundfrequenzverlauf hörbare Artefakte durch das Überblenden der gefilterten Signalblöcke. Für die Restsignalsynthese wird durch die exakte Anpassung der Filtereinstellungen an die detektierten Harmonischen der tonale Anteil komplett aus dem Originalsignal entfernt (Bild 3).

Diskussion

Damit eine zufriedenstellende Signaltrennung vorgenommen werden kann, ist es momentan notwendig, die Parameter des Algorithmus individuell an die zu analysierenden Signale anzupassen. Zukünftig werden Parametervoreinstellungen gesucht, die für verschiedene Signalklassen, wie beispielsweise Motordrehzahl-Hochlauf, optimal angepasst sind. Im Besonderen muß ein Kompromiss

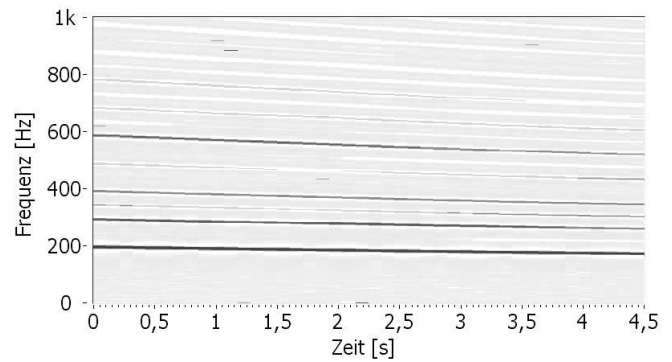


Abbildung 2: Spektrogramm der Sinussynthese, alle Motorordnungen wurden erkannt und synthetisiert

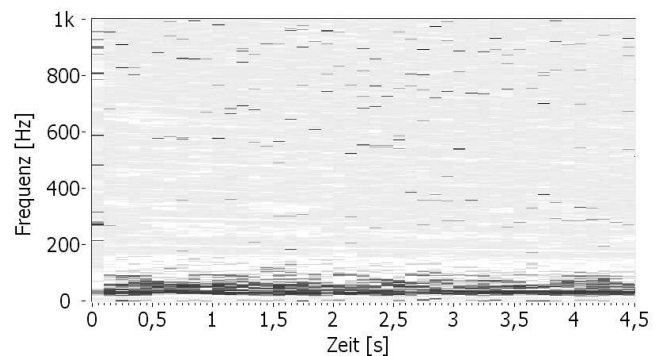


Abbildung 3: Spektrogramm des gefilterten stochastischen Anteils, es sind keine tonalen Anteile mehr zu erkennen

zwischen einer frequenzgenauen und empfindlichen Detektion, die große FFT-Blocklängen erfordert, und einer zeitlich korrekten Abbildung der harmonischen Verläufe getroffen werden.

Der Rauschpegelverlauf $N(f)$ aus Gleichung 1 wird sowohl für die Grundfrequenzbestimmung, als auch für die Detektion der tonalen Anteile als Gewichtungskriterium genutzt. Die Schätzung von $N(f)$ wird momentan durch eine starke Glättung des Originalspektrums realisiert. Dies ist allerdings nicht für alle Signalabschnitte eine gute Schätzung und kann zu Entscheidungsfehlern zwischen tonalen und nicht tonalen Spektralwerten führen. Bei der Weiterentwicklung des Verfahrens könnte eine bessere Schätzmethode des Rauschpegelverlaufes gefunden werden.

Für eine Optimierung des Verfahrens für die Trennung von Stimm- oder Musiksignalen müsste die Grundfrequenzanalyse auf un stetige Verläufe erweitert werden. In der Verbindung mit der Fähigkeit, mehrere Grundfrequenzzuordnungen zu bestimmen, wäre es dann möglich mehrstimmige Signale zu trennen und ihre Grundfrequenzverläufe zu bestimmen.

Literatur

- [1] Vary, P., Heute, U., Hess, W. (1998): Digitale Sprachsignalverarbeitung, Teubner Verlag Stuttgart
- [2] Lim, Y.C. (1986): Frequency-response masking approach for the synthesis of sharp linear phase digital filters, IEEE Trans. Circuits and Syst. vol. CAS-33, Seite 357-364