

# Vergleich verschiedener Methoden zur Geräuschklassifizierung

Silvia Allegro\*, Stefan Launer, Phonak AG, Laubisrütistr. 28, CH-8712 Stäfa

Michael Büchler, Norbert Dillier, Universitätsspital Zürich, ORL-Klinik, Frauenklinikstr. 24, CH-8091 Zürich

\*Email: silviaa@phonak.ch

## Zusammenfassung

Heute sind drei Methoden zur Geräuschklassifizierung im Hörgerät bekannt: die Analyse der Amplitudenstatistik (Ludvigsen, 1997), die Analyse des Modulationsspektrums (Ostendorf et al., 1997) und die Analyse der zeitlichen Fluktuationen und der spektralen Form (Kates, 1995, Büchler, 2000).

Diese drei Ansätze basieren auf verschiedenen mathematischen Grundlagen. Sie werden hier kurz vorgestellt und dann miteinander verglichen. Ziel ist, herauszufinden, inwieweit sich die drei Methoden unterscheiden und ob eine Kombination der einzelnen Ansätze die Geräuschklassifizierungsleistung insgesamt verbessern würde.

Wie gezeigt werden kann, erlauben zwar alle drei Methoden die eindeutige und robuste Erkennung von reinen Sprachsignalen, nicht aber die Identifikation weiterer Geräuschklassen. Ein Vergleich einzelner Merkmale ergibt, dass alle drei Methoden auf unterschiedliche Art und Weise Amplitudenmodulationen und damit dieselbe physikalische Eigenschaft des Signals beschreiben.

## Amplitudenstatistik

Zur Unterscheidung von kontinuierlichen und impulsförmigen Signalen (z.B. Sprache) wird nach Ludvigsen die Amplitudenstatistik der Einhüllenden gebildet und daraus die Merkmale *Breite* und *Symmetrie* (links- / rechtssteil) der Verteilung ermittelt (Ludvigsen, 1997, Figur 1).

Reine Sprachsignale weisen eine breite und asymmetrische Amplitudenverteilung auf und können somit eindeutig identifiziert werden. Figur 2 zeigt die Amplitudenverteilung nach Ludvigsen für ein reines Sprachsignal und ein Störgeräusch.

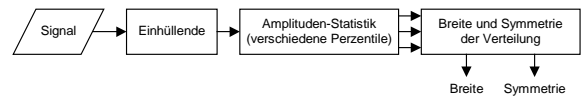
## Modulationsfrequenzanalyse

Für die Unterscheidung von Sprache, Sprache im Störgeräusch und Geräusch ohne Sprache berechnet Ostendorf die *Modulationstiefe* der Einhüllenden für *niedrige, mittlere* und *hohe Modulationsfrequenzen* ( $m_1$ ,  $m_2$ ,  $m_3$ ) (Ostendorf et al., 1997, Figur 3).

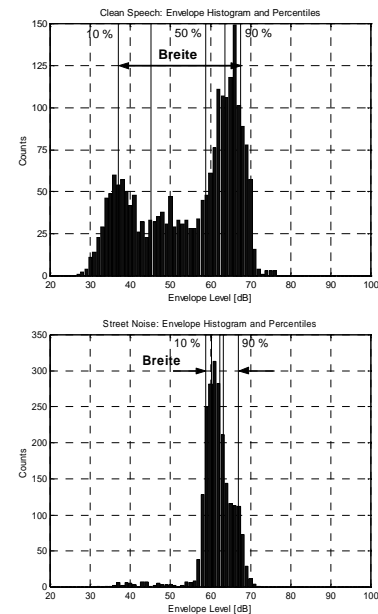
Reine Sprachsignale sind durch starke Modulationen bei tiefen Modulationsfrequenzen gekennzeichnet und können mit der Modulationsanalyse eindeutig von Sprache im Störgeräusch und Geräusch ohne Sprache unterschieden werden. Figur 4 zeigt die Ostendorf-Modulationstiefe  $m_1$  bei niedrigen Modulationsfrequenzen für ein reines Sprachsignal und ein Störgeräusch.

## Zeitliche Fluktuationen und spektrale Form

Basierend auf den Merkmalen von Kates für die Unterscheidung verschiedener Störgeräusche (Kates, 1995) verwendet Büchler (Büchler, 2000) zur allgemeinen Geräuschklassifizierung den *mittleren Pegel* (*TotPow*), die *Pegelschwankungen* (*MLFS*), den *spektralen Schwerpunkt* (*CGAV*) sowie die *Schwankungen des spektralen Schwerpunktes* (*CGFS*) (Figur 5).



Figur 1: Analyse der Amplitudenstatistik nach Ludvigsen. Ermittelte Merkmale: Breite und Symmetrie der Verteilung.



Figur 2: Amplitudenverteilung nach Ludvigsen für reine Sprache (oben) und Strassenlärm (unten). Das Sprachsignal weist eine breite und asymmetrische Amplitudenverteilung auf.

Reine Sprachsignale weisen starke Pegelschwankungen auf und können damit eindeutig identifiziert werden. Figur 6 zeigt die Pegelschwankungen MLFS nach Büchler für ein reines Sprachsignal und ein Störgeräusch.

## Interpretation und Vergleich der Merkmale

Um die ähnliche Klassifizierungsleistung der drei verschiedenen Methoden zu erklären, wurden die einzelnen Merkmale genauer untersucht.

Ludvigsen beschreibt die Amplitudenverteilung der Einhüllenden durch Breite und Symmetrie. Die Breite der Verteilung ist ein Mass für die *zeitlichen Schwankungen* des Signals.

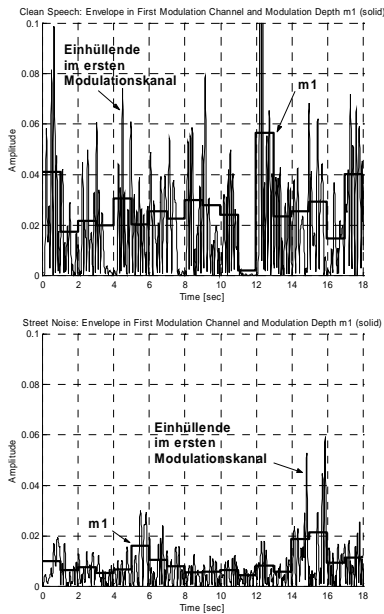
Ostendorf ermittelt die Modulationstiefe der Einhüllenden bei niedrigen, mittleren und hohen Modulationsfrequenzen. Die Stärke der Amplitudenmodulation ist ebenfalls ein Mass für die *zeitlichen Signalschwankungen*.

Büchler beschreibt die Verteilung der Amplituden durch die Pegelschwankungen MLFS. Diese sind umgekehrt proportional zur Varianz der Amplituden. Die Varianz ist ein Mass für die Amplitudenschwankungen und beschreibt somit gleichfalls *zeitliche Schwankungen* des Signals.

Alle drei Geräuschklassifizierungsmethoden charakterisieren demnach akustische Signale durch die zeitlichen Schwankungen der Amplitude. Die jeweils für die Ana-



Figur 3: Analyse des Modulationsspektrums nach Ostendorf. Merkmale: Modulationstiefe der Einhüllenden für niedrige, mittlere und hohe Modulationsfrequenzen.



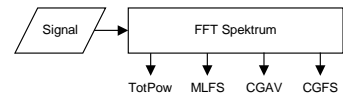
Figur 4: Einhüllende im ersten Modulations-Kanal und Modulationstiefe  $m_1$  (fett) für niedrige Modulationsfrequenzen nach Ostendorf. Oben: reine Sprache, unten: Strassenlärm. Die Modulationen des Sprachsignals bei niedrigen Modulationsfrequenzen sind stärker als diejenigen des Störgeräuschs.

lyse verwendeten Zeitfenster sind vergleichbar. Damit wird also immer dieselbe physikalische Eigenschaft der Signale beschrieben, nämlich die *Amplitudenmodulation*. Die untersuchten Geräuschklassifizierungsmethoden sind jedoch keinesfalls identisch: Im Gegensatz zur Ludvigsen-Breite und dem Büchler-MLFS, welche nur Angaben über alle insgesamt vorhandenen Amplitudenmodulationen erlauben, ermöglicht der Ansatz nach Ostendorf zusätzlich zur Modulationstiefe Aussagen über den *Frequenzgehalt der Modulationen*. Dafür erlaubt die Methode nach Ludvigsen mit der Berechnung beliebiger Perzentilwerte eine detaillierte Beschreibung der *Amplitudenverteilung*. Der Ansatz nach Büchler misst zusätzlich zur Amplitudenmodulation den *Energiegehalt* des Signals sowie *spektrale Eigenschaften*.

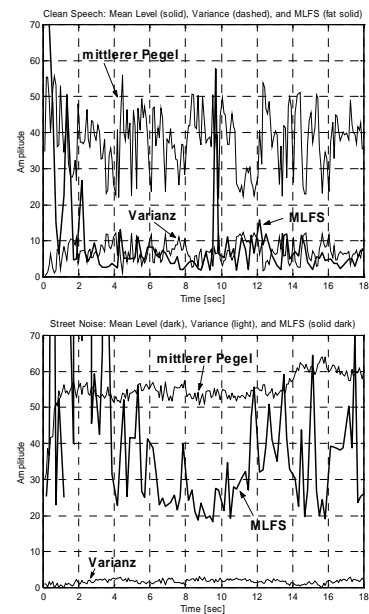
### Schlussfolgerungen

Die systematische Untersuchung der drei heute bekannten Methoden zur Geräuschklassifizierung im Hörgerät hat gezeigt, dass zwar jeder Algorithmus eine eindeutige und robuste Trennung von Sprachsignalen und Nicht-Sprachsignalen erlaubt, dass aber eine differenziertere Unterscheidung verschiedener akustischer Umgebungen mit keinem der drei Ansätze möglich ist. Eine Kombination der diskutierten Methoden verbessert also die Klassifizierungsleistung nicht.

Die genauere Analyse der Merkmale der drei Geräuschklassifizierungsmethoden zeigt, dass die Breite der Amplitudenverteilung (Ludvigsen), die Modulationstiefen der Einhüllenden (Ostendorf) und die Pegelschwankungen (Kates / Büchler) bezüglich Klassifizierungsleistung



Figur 5: Analyse der zeitlichen Struktur und der spektralen Form nach Büchler. Merkmale: mittlerer Pegel, Pegelschwankungen, spektraler Schwerpunkt und dessen Schwankungen.



Figur 6: Mittlerer Pegel, Varianz und Pegelschwankungen MLFS (fett) nach Büchler für reine Sprache (oben) und Strassenlärm (unten). MLFS ist umgekehrt proportional zur Varianz der Pegelschwankungen und somit für das Sprachsignal kleiner als für das Störgeräusch.

aequivalent sind. Alle diese Merkmale werden mit vergleichbaren Analyse-Fenstern berechnet und beschreiben somit in der einen oder anderen Form die zeitlichen Schwankungen der Signalamplitude bzw. die Amplitudenmodulation des Signals. Trotzdem sind die drei Methoden nicht identisch: nur der Ostendorf-Ansatz ermöglicht Aussagen über die Modulationstiefe bei verschiedenen Modulationsfrequenzen. Dafür erlaubt der Ludvigsen-Ansatz die detaillierte Beschreibung der Amplitudenverteilung, während der Ansatz von Büchler Angaben über den Energiegehalt und das Frequenzspektrum ermöglicht.

Die Tatsache, dass immer dieselbe physikalische Eigenschaft der Geräusche extrahiert wird erklärt die ähnliche Klassifizierungsleistung der drei verschiedenen Klassifizierungsmethoden. Insbesondere weisen Sprachsignale wegen den Pausen starke Schwankungen auf; daher ist die Analyse der Amplitudenmodulationen hervorragend geeignet, reine Sprache zu erkennen.

### Referenzen

- Ludvigsen, C. (1997). „Schaltungsanordnung für die automatische Regelung von Hörhilfsgeräten“, Europäische Patentschrift, EP 0 732 036 B1.
- Ostendorf, M. et al. (1997). „Empirische Klassifizierung verschiedener akustischer Signale und Sprache mittels einer Modulationsfrequenzanalyse“, DAGA 97, 608-609.
- Kates, J. M. (1995). „Classification of background noises for hearing-aid applications“, Journal of the Acoustical Society of America, 97 (1), 461-470.
- Büchler, M. (2000). „Klassifizierung der akustischen Umgebung für Hörgeräte-Anwendungen“, DAGA 2000.