

Testsystem zum Vergleich verschiedener Parameter zur Verbesserung der Erkennungsleistung bei der Flugzeuggeräuschidentifikation

D.Hemmer¹, C. Pörschmann²

¹ Fachhochschule Köln, 50679 Köln, e-mail: dominic.hemmer@fh-koeln.de

² Fachhochschule Köln, 50679 Köln, e-mail: christoph.poerschmann@fh-koeln.de

Einleitung

Bei Langzeitlärmmessungen von Flugzeuggeräuschen kommt es je nach Positionierung der Messstelle zu Beeinflussungen des Fluglärmpegels durch verschiedene Fremdgeräusche (z.B.: Verkehrslärm, Regengeräusche, Windgeräusche). Durch deren Überlagerung am Immissionsort wird eine Bereinigung des zu untersuchenden Immissionsanteils von den Fremdeinflüssen nötig. Mit geeigneten Methoden der digitalen Audiosignalverarbeitung ist es möglich, einzelne Signalquellen zu detektieren und den quellspezifischen Anteil weitgehend getrennt von anderen Umgebungsgeräuschen zu betrachten.

Grundlagen

Aus einem aufgezeichneten Audiosignal können verschiedene Merkmale [2] extrahiert werden. Diese Merkmale können dann bei unterschiedlichen Audiosignalen auf Ähnlichkeit abgeglichen und zur Geräuscherkennung genutzt werden. Verwendung finden vergleichbare Verfahren auch in anderen Bereichen, so z.B. in der Sprach- oder Sprechererkennung. Es werden diverse Merkmale der Audiosignale trainiert, um sie dann mit den Merkmalen der zu identifizierenden Signale zu vergleichen. Dazu werden die spektralen und temporalen Eigenschaften eines Audiosignals untersucht und gesammelt, um sie dann in einer Merkmalsmatrix zusammenzufassen. Verschiedenste Parameter können miteinander kombiniert werden und bilden somit eine Vielzahl an Variationsmöglichkeiten. Es lassen sich z.B. folgende veränderbare Parameter verwenden:

- Fensterlänge des zu untersuchenden Zeitabschnitts
- Zeitliche Verschiebung der Fenster
- Vektorquantisierungsstufen der Merkmalsmatrix
- Verschiedene Kombinationen von Merkmalen

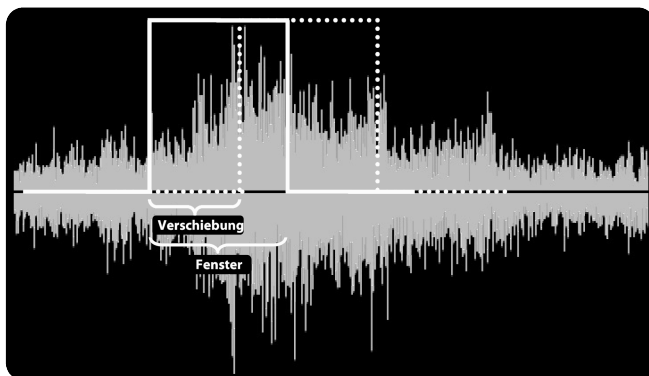


Abbildung 1: Parameter - Fenster und Verschiebung

Um für die Detektion unterschiedlicher Geräuscharten möglichst optimale Merkmale auszuwählen und für diese die passenden Parameter zu bestimmen, wurde ein Testsystem unter Matlab entwickelt. Dieses System erlaubt es, unterschiedliche Kombinationen von Merkmalen und deren Parametern zu testen und die Erkennungsraten zu vergleichen.

Voruntersuchung

Möchte man verschiedene spektrale und temporale Eigenschaften eines Audiosignals als Merkmale zur Klassifikation und Identifikation heranziehen, kommt es zu einer hohen Anzahl von veränderbaren Parametern. Dabei führen jedoch verschiedene Kombinationsmöglichkeiten der Parameter zu unterschiedlichen Untersuchungsergebnissen. Eine Möglichkeit diese zahlreichen Konstellationen zu untersuchen ist die Anwendung der „Brute-Force“ Methode. Durch die Einbeziehung aller hier vorhandenen Parameter ergibt sich eine extrem hohe Anzahl von Kombinationen. So z.B.:

$$N = F \cdot V \cdot VQ \cdot \sum_{k=1}^n \frac{n!}{k!(n-k)!}$$

Mit 29 Merkmalen (n), 5 Fensterbreiten (F), 5 Verschiebungszeiten (V) und 5 Vektorquantisierungsstufen (VQ) ergeben sich demnach $6,71 \cdot 10^{10}$ Kombinationsmöglichkeiten (N).

Dieses Vorgehen beansprucht eine lange Rechenzeit. Um sie zu reduzieren, ist es angemessen, durch eine geeignete Voruntersuchung und eine gezielte Analyse einzelner Merkmale bestimmte Kombinationsmöglichkeiten auszuschließen oder sogar bestimmte Parametervariationen komplett aus der Berechnung herauszunehmen.

Im Rahmen der beschriebenen Untersuchungen wurden unter Zuhilfenahme der folgenden Kriterien die Anzahl der Merkmale und deren Kombinationsmöglichkeiten verringert:

- Kleine Streuung der Merkmale innerhalb einer Klasse
- Großer Abstand der Klassen zueinander

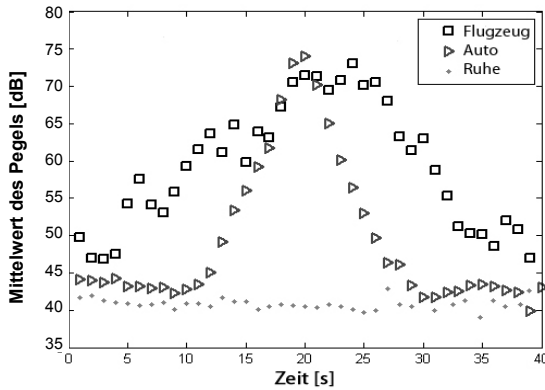


Abbildung 2: zeitlicher Verlauf des Merkmals ‚Pegel‘ der verschiedenen Klassen: Flugzeug, Auto und Ruhe

Aufbau des Testsystems

Um eine optimale Auswahl der verwendeten Parameter, sowie eine Anpassung an die jeweilige Umgebungssituation in der die Geräusche aufgenommen wurden testen zu können, wurden das System in verschiedene Blöcke unterteilt. Diese können je nach Bedarf und Untersuchungsziel angewendet und mit in den Testdurchlauf integriert werden. Dadurch wird es möglich, verschiedene Verfahren miteinander zu vergleichen und den erforderlichen Rechenaufwand an die gewünschte Erkennungsleistung anzupassen.

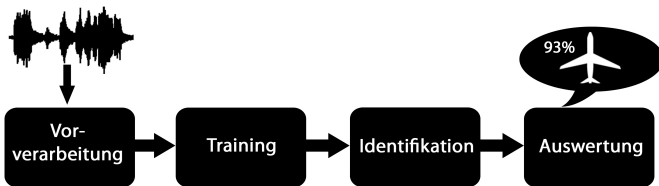


Abbildung 3: Aufbau des Testsystems

Vorverarbeitung: In diesem Block besteht die Möglichkeit, die Audiosignale vor der weiteren Verwendung zur Identifikation oder zum Training zu bearbeiten. Folgende Bearbeitungsmöglichkeiten sind optional auswählbar:

- Filtern der Audiodatei (Hochpass, Tiefpass, Bandpass, Bandsperre)
- Rauschreduktion [1]
- Bewertungsfilter (A, B, C, D)
- Hinzufügen verschiedener Pegel von Hintergrundrauschen

Training: Im Trainingsblock werden die Merkmale der für das Training ausgewählten Audiosignale normiert in einer Matrix zusammengefasst. Diese dient dann als Referenzmatrix für die Identifikation. Die verschiedenen Parameter wie Fensterlänge, Verschiebung, Vektorquantisierungsstufen und Merkmalskombinationen müssen an dieser Stelle schon mit in die Referenzmatrix einfließen. Zusätzlich kann hier zwischen zwei verschiedenen Trainingsmethoden ausgewählt werden. Zum einen die Methode, die eine Referenzmatrix

zum weiteren Vektorvergleich bereitstellt und zum anderen, eine Methode, die ein Hidden-Markov-Modell trainiert, um es für die spätere Identifikation einsetzen zu können.

Identifikation: Bei der Identifikation kann zwischen den beiden o.g. Methoden (Vektorvergleich oder Hidden-Markov-Modell) gewählt werden. Auch hier kommen die verschiedenen Kombinationsmöglichkeiten der oben erwähnten Parameter zum Einsatz und bestimmen maßgeblich die Erkennungswahrscheinlichkeit des zu identifizierenden Audiosignals.

Auswertung: Im letzten Block des Testsystems kann eine Auswertung und Interpretation der gesammelten Untersuchungsergebnisse durchgeführt werden.

Das Testsystem bietet somit die Möglichkeit verschiedene Parameter miteinander zu kombinieren und zwei unterschiedlichen Vorgehensweisen (Vektorvergleich oder Hidden-Markov-Modell) für die Identifikation anzuwenden. Durch die modulare Gestaltung der einzelnen Blöcke kann das System auch leicht mit zusätzlichen Geräuschklassen oder veränderten Geräuschen innerhalb einer Klasse nachtrainiert werden. Zudem wurde die Möglichkeit geschaffen, lange Rechenzeiten durch die Verwendung eines eigenen Rechenclusters (Grid-Computing) drastisch zu reduzieren.

Untersuchungsbeispiel bei der Identifikation von Flugzeuggeräuschen

Mit Hilfe des Testsystems konnte die hohe Anzahl an möglichen Parametern auf eine sinnvolle (hohe Erkennungsleistung) Anzahl eingegrenzt werden. Somit konnte beispielsweise eine Erkennungsleistung von über 80% bei Fensterlängen zwischen 2.0s und 3.5s erzielt werden. Bei der Kombination aus der Fensterlänge von 2.0s und der Verschiebung von 1.0sec konnte eine gewisse Konstanz festgestellt werden. So tritt keine maßgebliche Änderung bezüglich der Erkennungsrate auf, wenn man die Vektorquantisierungsstufen {128, 256, 512, 1024} variiert und verschiedene Merkmalskombinationen verwendet. Bei hohen Erkennungsleistungen fanden sich zudem die Merkmale:

- Schwankung des Spektrums
- Schwerpunkt des Kepstrums
- Schwankung der LPC
- MFCCs

in unterschiedlichen Kombinationen wieder. Betrachtet man hingegen die Merkmale einzeln, wirkt sich die Kombination von der Fensterlänge und Verschiebung meist stark auf die Erkennung aus. Nahezu unbeeinflusst ist dagegen das Merkmal ‚MFCC‘. Die Erkennung von Flugzeuggeräuschen ist von der Variation der Fensterlänge, Verschiebung und der Vektorquantisierung weitestgehend unabhängig.

Literatur

- [1] Zölzer, U. (2002). DAFX Digital Audio Effects. John Wiley & Son, Chichester/England.
- [2] Niemann, H. (2003). Klassifikation von Mustern. 2. überarbeitete Auflage im Internet.