

Entwicklung eines Algorithmus für die Detektion von Registerübergängen

Felix Langenbruch¹, Judith Thoma², Tobias Weißgerber³,
Jörn Loviscach⁴, Katrin Neumann⁵, Malte Kob⁶

¹Fachhochschule Bielefeld, 33602 Bielefeld, Deutschland, Email: felix.langenbruch@arcor.de

²Universität Frankfurt, 60325 Frankfurt am Main, Deutschland, Email: registerwechsel@googlemail.com

³Universität Frankfurt, 60325 Frankfurt am Main, Deutschland, Email: Tobias.Weissgerber@kgu.de

⁴Fachhochschule Bielefeld, 33602 Bielefeld, Deutschland, Email: joern.loviscach@fh-bielefeld.de

⁵Universität Frankfurt, 60325 Frankfurt am Main, Deutschland, Email: Katrin.Neumann@em.uni-frankfurt.de

⁶Hochschule für Musik Detmold, 32756 Detmold, Deutschland, Email: kob@hfm-detmold.de

Einleitung

Registerwechsel sind ein vielfach und vielfältig behandeltes Thema, insbesondere im Bereich der Stimmphysiologie und der Gesangspädagogik. Diese Arbeit betrachtet das Phänomen für weibliche Gesangsstimmen aus signalanalytischer Sicht. Hierfür wurden Algorithmen untersucht, die Features aus Stimmsignalen extrahieren und auf einen Zusammenhang mit Registerwechseln prüfen. Ziel ist die Operationalisierung des Begriffs „Registerwechsel“, sowie die Evaluierung der Möglichkeiten einer rechnergestützten Merkmalsextraktion für die Detektion der Wechsel. Zudem sollen die Ergebnisse im Idealfall den Grundstein für neue Erkenntnisse im Bereich der Gesangspädagogik legen. Es ergibt sich die Frage, welche Eigenschaften der Signale charakteristisch für Registerwechsel sein könnten und wie solche Eigenschaften sich aus den Signalen extrahieren lassen. Zudem soll ein Urteil über die Robustheit der Algorithmen möglich sein, das einen Ausblick auf mögliche Herangehensweisen an das Thema in zukünftigen Arbeiten zulässt.

Vorgehensweise

Das Programm liest strukturierte Folgen von Einzeltonen ein, aus denen es Merkmale mit Hilfe von Funktionen aus MATLAB und der MIRtoolbox⁷, sowie eigenen Funktionen extrahiert. Grafische Ausgaben zeigen die Merkmalsverläufe für jede Einzeltonfolge. Als Vergleichswerte für die Erkennung von Registerwechseln dienen subjektive Ermittlungen der Wechsel, die von Judith Thoma und Katrin Neumann im Rahmen eines Online-Verfahrens durchgeführt wurden.

Das Programm extrahiert die folgenden Parameter aus den Signalen:

- **Energien der Partialtöne**

Die Energien der Partialtöne sollen Änderungen der Klangfarbe zeigen. Die Partialtöne werden dabei mit Hilfe eines Bandpass-Filters extrahiert. Abbildung 1 zeigt die grafische Ausgabe beispielhaft für eine Folge von Tönen, in der der Registerwechsel vor dem 15. Ton ermittelt wurde.

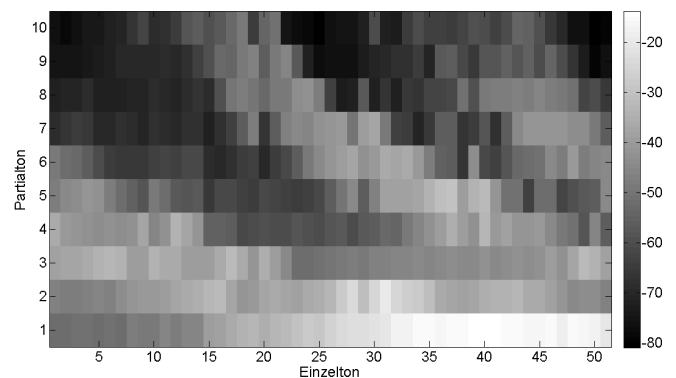


Abbildung 1: Energien der Partialtöne in dB

- **Phasenlage aller Partialtöne relativ zu F0**

Aus den gefilterten Partialtönen werden die Phasen relativ zum Grundton ermittelt. Phasenverschiebungen sind für das menschliche Gehör meist nicht wahrnehmbar, könnten jedoch im Bereich von Registerwechseln eine Rolle spielen.

- **Spektraler Schwerpunkt**

Der spektrale Schwerpunkt beschreibt das Spektrum eines Signals grob durch einen einzelnen Wert. Im Verlauf der Werte für eine Folge von Einzeltonen könnte sich eine besonders starke Verschiebung nach Registerwechseln ergeben.

- **MFCCs**

MFCCs (Mel Frequency Cepstral Coefficients) bieten eine Darstellung des Spektrums auf verschiedenen groben Detailstufen. Sie sind die dominanten Merkmale im Bereich der Spracherkennung. [1]

- **Amplituden- und Periodenmodulation**

Amplituden- und Periodenmodulation beziehen sich hier direkt auf die Wellenform eines Signals.

- **Zeitliche Dichte der Nulldurchgänge**

Die zeitliche Dichte der Nulldurchgänge beschreibt die Anzahl an Nulldurchgängen in einem Signal bezogen auf eine feste Anzahl an Samples. Sie kann beispielsweise ein Indikator für eine geänderte Obertonstruktur sein.

⁷<https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>

• Formanten

Die Ermittlung der Formanten des Signals erfolgt über eine LPC-Analyse. Neben den Frequenzen wurden auch die Magnituden der Formanten betrachtet. Neumann et al. [2] haben bereits gefunden, dass sich die Lage der ersten beiden Formanten von Männerstimmen am Registerübergang auf charakteristische Weise ändert.

Aus diesen Parametern bildet das Programm Verläufe über die jeweilige Tonfolge, aus denen es wiederum insgesamt 21 Merkmale extrahiert. Bei 16 dieser Merkmale handelt es sich um globale Minima oder Maxima der Verläufe, bei den übrigen fünf um Merkmale mit teilweise relativem Bezug. Letzterer ist durch einen dynamischen Schwellwert gegeben, der sich durch die Streuweite der Werte in einem Parameterverlauf ergibt. Die Software hält fest, für welchen Ton in der Reihe der Schwellwert erstmals überschritten wird. Der Index dieses Tons wird später als Vergleichswert für den jeweiligen subjektiv ermittelten Index herangezogen.

Testmaterial

Als Testmaterial dienen neun Ordner mit strukturierten Einzeltonfolgen, die Aufnahmen von weiblichen Gesangsstimmen enthalten. Die Töne sind dabei in mehreren Abschnitten chromatisch aufwärts gesungen und ab eins indiziert. Auf die Stimmgattungen Alt, Mezzosopran und Sopran entfallen jeweils drei Ordner. In allen Aufnahmen singen die Sängerinnen jeweils den Vokal „a“. Das Material wurde von Judith Thoma aufgenommen und von Katrin Neumann, sowie Tobias Weißgerber in der Klinik für Hals-, Nasen-, Ohrenheilkunde (Schwerpunkt Phoniatrie und Pädaudiologie) der Universität Frankfurt bearbeitet.

Resultate

Die Resultate der Berechnungen zeigt das Programm unter anderem in grafischen Ausgaben, die die Inhalte der Matrizen für Erkennungen und Abweichungen durch eine Grauwertdarstellung veranschaulichen sollen. Abbildung 2 zeigt die Abweichungen der durch das Programm ermittelten von den subjektiv ermittelten Indizes.

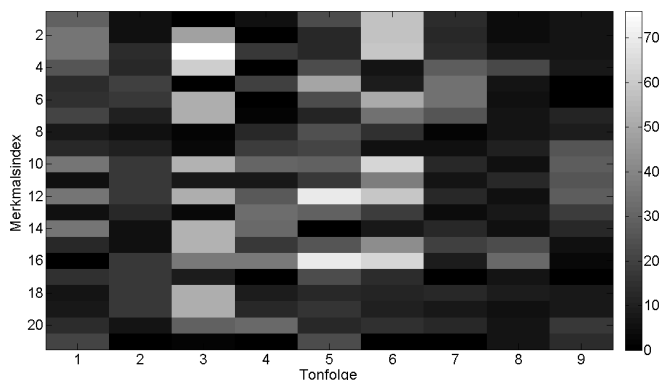


Abbildung 2: Abweichungen zwischen den per Software und den subjektiv ermittelten Registerwechseln (Abstand in Einzeltönen)

Jede Zeile der Matrix steht für ein Merkmal, jede Spalte für einen Ordner mit Einzeltönen. Die Grauwertskala am rechten Rand des Bildes zeigt die Zuordnung der Werte. Abbildung 2 zeigt insbesondere für die Merkmale 17 und 21 vergleichsweise geringe Abweichungen. Merkmal 17 ist die erste negative Differenz des spektralen Schwerpunktes über dem gegebenen Schwellwert (Differenz zwischen Minimum und Maximum des Verlaufs multipliziert mit 0,25). Die mittlere Abweichung beträgt hier zehn Töne, während sie für Merkmal 21 (erstmaliges Auftreten eines Phasenunterschieds über dem Schwellwert = größter Parameterunterschied zweier aufeinanderfolgende Töne multipliziert mit 0,5) acht Töne beträgt. Gemessen an der durchschnittlichen Anzahl an Tönen in einem Ordner, die 63 beträgt, sind dies geringe Werte. Jedoch stimmten die subjektiv ermittelten Registerwechsel für beide Merkmale nur in jeweils zwei von neun Fällen exakt überein, für alle weiteren Merkmale nur in einem oder in keinem Fall.

Hieraus lässt sich schließen, dass eine exakte Erkennung von Registerwechseln durch ein einzelnes Merkmal insgesamt unwahrscheinlich ist, sich jedoch anhand der Abweichungen klare Tendenzen für die Eignung bestimmter Parameter erkennen lassen. Diese sind insbesondere der spektrale Schwerpunkt und die Phasen einzelner Partialtöne relativ zur Phase des Grundtons. Durch die Änderung der Klangfarbe im Bereich der Registerwechsel liegt es nahe, zur Erkennung spektrale Parameter zu nutzen. Andererseits minimieren professionelle Sängerinnen und Sänger gerade diese hörbaren Änderungen. Umso mehr bietet sich die Phasenlage zur Erkennung an.

Ausblick

Das Programm ist auf die Ermittlung eines Registerwechsels pro Einzeltonfolge ausgelegt, so dass eine Erweiterung für die Ermittlung mehrerer Wechsel nahe liegt. Dies könnte zudem belastbarere Aussagen über die Eignung bestimmter Parameter ermöglichen. Zudem sollten bei der Extraktion von Merkmalen aus Parameterverläufen neben globalen auch lokale Minima und Maxima eine Rolle spielen, so dass ein Bezug zu den Verläufen selbst gegeben ist. Eine Kombination einzelner Merkmale könnte zu höheren Erkennungsquoten führen, da sich gezeigt hat, dass je nach Testsequenz verschiedene Merkmale anschlagen. Der Einsatz von Maschinlernen, das mit den subjektiven Analysen antrainiert werden könnte, wäre ebenso eine Option, die die Erkennungsquoten erhöhen und die Abweichungen verringern könnte.

Literatur

- [1] Beth Logan. *Mel Frequency Cepstral Coefficients for Music Modelling*. Tagungsband ISMIR 2000.
- [2] Katrin Neumann, Patrick Schunda, Sebastian Hoth, Harald A. Euler. *The Interplay between Glottis and Vocal Tract during the Male Passagio*. *Folia Phoniatrica et Logopaedica*, Vol. 57, Nr. 5-6, Seiten 308-327, 2005.