Modellbasierte Bestimmung physiologischer Stimmparameter aus digitalen Hochgeschwindigkeitsaufnahmen

Michael Döllinger, Ulrich Hoppe, Jörg Lohscheller, Stefan Schuberth, Frank Hettlich[1] und Ulrich Eysholdt

Abteilung für Phoniatrie und Pädaudiologie
Friedrich-Alexander Universität Erlangen-Nürnberg, 91054 Erlangen
[1]Mathematisches Institut II, Universität Karlsruhe (TH), 76128 Karlsruhe
Email: boert-rdm@gmx.de

Zusammenfassung. Für die quantitative Bewertung pathologischer oder gesunder Stimmen ist die endoskopische Beobachtung der Stimmlippenbewegungen von zentraler Bedeutung. In dieser Arbeit wird ein Verfahren vorgestellt, das ermöglicht Asymmetrien zwischen rechter und linker Stimmlippe durch physiologische Parameter zu beschreiben. Ausgangspunkt ist ein Zwei-Massen-Modell in Verbindung mit digitalen Hochgeschwindigkeitsaufnahmen von Stimmlippen. Das vorhandene Problem wird als Optimierungsproblem mit nicht konvexer Zielfunktion betrachtet. Hierbei ist die Wahl geeigneter Startwerte von großer Bedeutung.

1 Einleitung

Eine heisere Stimme entsteht oft durch irreguläre, im Extremfall durch chaotische Stimmlippenschwingungen. In vielen Fällen entstehen diese irregulären Schwingungen durch Asymmetrien zwischen linker und rechter Stimmlippe. Bei einer Reihe von Stimmstörungen sind diese Asymmetrien nur während der Stimmgebung zu beobachten. Deshalb kann dieses Verhalten nur mit digitalen Hochgeschwindigkeitsaufnahmen beobachtet werden.

Ziel dieser Arbeit ist es, einen Algorithmus zur automatischen und objektiven Beurteilung von Stimmlippenbewegungen auf der Basis der Hochgeschwindigkeitsaufnahmen zu erstellen. Es sollen sowohl Art als auch Ausmaß möglicher Asymmetrien bestimmt werden.

2 Methode

Die Stimmlippenbewegungen werden mit einer digitalen Hochgeschwindigkeitskamera aufgezeichnet ($3704\,Hz$ und $128\,x\,64$ Pixel Auflösung). Die Bewegungskurven der Stimmlippen werden zunächst aus den erzeugten digitalen Hochgeschwindigkeitsaufnahmen extrahiert (*experimentelle Kurven*). Um Stimmlippenbewegungen quantifizieren zu können, wird ein biomechanisches Modell benutzt.

Ausgangsmodell ist das vereinfachte Zwei-Massen-Modell (2MM), welches aus zwei miteinander gekoppelten Oszillatoren besteht [1,2]. Durch aerodynamische Kräfte (subglottaler Druck) wird das Modell in Bewegung gesetzt . Das 2MM ist in der Lage viele Schwingungseigenschaften von Stimmlippen nachzubilden. In einer vorangehenden Arbeit wurde das 2MM bereits an Stimmlippenbewegungen manuell angepasst [3].

Der Fehler zwischen den experimentellen Kureven und den mit dem 2MM errechneten Kurven (*theoretische Kurven*) soll minimiert werden. Für die Minimierung werden zwei Faktoren Q_l und Q_r eingeführt, welche auf der linken/rechten Seite die Federkonstanten und Massen des 2MM beeinflussen. Somit können Asymmetrien der Muskelspannungen und der Stimmlippenmassen nachgebildet werden. Neben diesen Faktoren wird noch der subglottale Druck P_s variiert. Dieser kann als Maß für die Energie, die dem System zugeführt wird, gesehen werden kann. Die übrigen im 2MM enthaltenen Parameter werden aus einem Standardparametersatz bezogen [1].

Bei dem zu lösenden Problem handelt es sich um ein Optimierungsproblem, mit zu minimierender Zielfunktion über die Parameter (Q_l, Q_r, P_s). Die experimentellen Kurven und die theoretischen Kurven werden zunächst einer Fouriertransformation unterzogen. Da das Spektrum der experimentellen Kurven nur von wenigen Oberschwingungen dominiert wird, werden nur diese in der Zielfunktion berücksichtigt. Da pathologische Stimmlippenbewegungen keine vollständige Periodizität aufweisen, werden zusätzlich die Nachbarkoeffizienten der dominanten Oberschwingungen mit in die Zielfunktion einbezogen. Somit werden kleinere Frequenzänderungen der experimentellen Kurven, bedingt durch Pathologie oder kleinere Messfehler (Pixelrauschen, Fehler der Kantenerkennung) berücksichtigt. Die Fourierkoeffizienten werden in Absolut- und Phasenwerte getrennt. In der Zielfunktion werden Absolutwerte und Phasenwerte der Fourierkoeffizienten getrennt voneineander betrachtet. Die Absolutwerte der Fourierkoefizienten können in Abhängigkeit der Amplituden beliebig große oder kleine Werte annehmen. Da sich die Phasen maximal um 2π unterscheiden, werden die Kurven skaliert. Der Skalierungsfaktor ergibt sich aus dem Angleichen des maximal möglichen Fehlers der Absolutwerte der Fourierkoeffizienten und des maximalen Phasenfehlers der experimentellen Kurven.

Die aus den Absolutbeträgen der Fourierkoeffizienten und deren Phasen entstandene Zielfunktion ist nicht konvex. Deshalb muss vor der eigentlichen Optimierung eine wissensbasierte Startwertsuche gestellt werden. Für die erste Approximation von Q_l, Q_r wird die Bewegungsgleichung eines ungekoppelten Oszillators benutzt [3]. Da für P_s keinerlei Anhaltspunkte gegeben sind, wird hierfür ein definiertes, physiologisch sinnvolles Intervall durchlaufen. Hieraus resultieren die Startwerte (Q_l^i, Q_r^i, P_s^i), die dem Optimierungsalgorithmus übergeben werden.

Da die Zielfunktion eine 'schlechte' Struktur aufweist, werden Optimierungsalgorithmen verwendet, die keine Gradienten benutzen. 'Schlechte Struktur' heißt, dass sehr viele lokale Minima in der Zielfunktion enthalten sind. Bei 'schlechten' Startwerten würden Gradientenverfahren sehr schnell in lokalen Minima abbrechen. Besser geeignet für diese Art von Problemen sind gradientenunabhängige Verfahren wie z.B. der Nelder-Mead-Algorithmus oder der Algorithmus von Powell [4].

3 Ergebnis

Das entwickelte Verfahren wurde bereits bei gesunden Stimmlippen und spezifischen Störungsbildern [5] erfolgreich getestet. Exemplarisch wird die Funktionalität des Algorithmus und dessen Vorgehensweise an einer, aus digitalen Hochgeschwindigkeitaufnahmen extrahierten, gesunden Stimmlippenbewegung (Abb. 1) gezeigt. Der verwendete Nelder-Mead-Algorithmus bricht nach 15 Iterationen mit folgendem Ergebnis ab:

$$Q_l^* = 2.6801, \quad Q_r^* = 2.6429, \quad P_s^* = 12.8 \; cm \; H_2O.$$

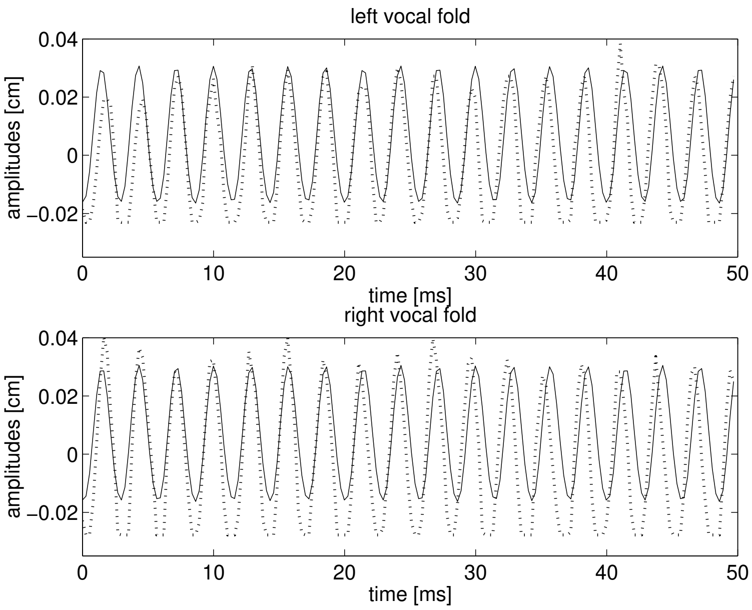

Abb. 1. Ergebnis der Approximation (durchgezogene Linien) an gesunde Stimmlippenbewegung (gestrichelt) in einem Zeitfenster von 50 ms. Das obere Bild entspricht der linken und das untere der rechten Seite der Stimmlippen.

Die Ähnlichkeit der gefundenen Asymmetrieparameter Q_l und Q_r bestätigt die als gesund diagnostizierte gesunde Stimmlippenbewegung. Der Druck nimmt in vielen Parameterbereichen keinen Einfluß auf das Symmetrieverhalten des 2MM und ist auch in diesem Beispiel nur für die Stärke der Auslenkungen der Bewegungskurven des 2MM verantwortlich.

4 Diskussion

Die bisherige Arbeit hat gezeigt, dass aus extrahierten Stimmlippenbewegungen prinzipiell physiologische Parameter [5] gewonnen werden können. Für die Erweiterung des Algorithmus auf komplexere Krankheitsbilder, wie z.B. einer Lähmung des N. laryngeus superior oder bei Stimmlippenpolypen sind allerdings noch Modifikationen einzubinden. Bisher gefundene Approximationen des Optimierungsalgorithmus stimmen mit den pathologischen Stimmlippenbewegungen dieser Krankheitsbilder nicht überein. Ein Grund hierfür ist wahrscheinlich die beschränkte Anzahl der zu optimierenden Parameter.

Der Schwerpunkt der Arbeit liegt in nächster Zeit deshalb vor allem darin, die Zahl der Optimierungsparameter so zu erweitern, dass weitere Krankheitsbilder nachgebildet werden können. Ein weiterer Punkt ist die korrekte physiologische Interpretation der erhaltenen Parametersätze.

Literatur

1. K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," *Bell Syst. Techn. J.*, vol. 51, pp. 1233–1268, 1972.
2. H. Herzel, "Bifurcations and chaos in voice signals," *Applied Mechanical Revues*, vol. 46, pp. 399–413, 1993.
3. P. Mergell, *Nonlinear Dynamics of Phonation - High-Speed Glottography and Biomechanical Modeling of Vocal Fold Oscillations.* Kommunikationsstörungen - Berichte aus Phoniatrie und Pädaudiologie, Aachen: Shaker Verlag, 1998.
4. W. Murray, *Numerical Methods for Unconstrained Optimization.* London: Academic Press, 1972.
5. M. Döllinger, U. Hoppe, J. Lohscheller, S. Schuberth, F. Hettlich, and U. Eysholdt, "Detecting asymmetries in high speed observations of vocal folds," in *2nd International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications*, University of Firenze, 2001.