

Automatische Diagnose von Stimmstörungen mittels phonovibrographischer Merkmalsbeschreibungen

Daniel Voigt¹, Anxiong Yang¹, Michael Döllinger¹, Ulrich Eysholdt¹, Jörg Lohscheller²

¹ Abteilung für Phoniatrie und Pädaudiologie, Universitätsklinikum Erlangen, Bohlenplatz 21, 91054 Erlangen, Deutschland

² Fachhochschule Trier, Fachbereich Informatik – Medizininformatik, Schneidershof, 54293 Trier, Deutschland

Einleitung

Bei der klinischen Diagnose von Stimmstörungen ist es besonders wichtig, die Bewegungsmuster der beiden Stimmlippen im Kehlkopf zu untersuchen. Als besonders geeignet hat sich hierfür die digitale Hochgeschwindigkeits- (HG) Videoendoskopie erwiesen, da sie als einziges Verfahren eine umfassende Untersuchung der Dynamik in Echtzeit ermöglicht. Eine gesunde Stimme ist durch links-rechts-symmetrische und periodische Schwingungsmuster gekennzeichnet. Die Bewertung dieser dynamischen Stimmlippeneigenschaften hat sich im klinischen Alltag als problematisch erwiesen, da sie zum einen ein hohes Maß an Erfahrung seitens des auswertenden Arztes voraussetzt und zum anderen sehr zeitintensiv ist. Aus dieser Problematik resultiert, dass die mithilfe von HG-Aufnahmen gestellten subjektiven Diagnosen eine verringerte Inter- und Intra-Rater-Reliabilität aufweisen [1].

In dieser Arbeit wird ein objektiver Beschreibungs- und Klassifikationsansatz für Stimmlippenschwingungen vorgestellt, der es erlaubt, zuverlässig und reproduzierbar zwischen gesunden und pathologischen laryngealen Bewegungsmustern zu unterscheiden. Die Grundlage bildet hierbei das sogenannte Phonovibrographie-Verfahren [2], das die Stimmlippenpositionen in endoskopischen HG-Aufnahmen automatisch detektiert und zur weiteren Verarbeitung grafisch aufbereitet. Mithilfe einer Reihe von quantitativen Merkmalen, die die räumlich-zeitlichen Veränderungen der Stimmlippen und ihre Symmetrie erfassen, wird eine automatische Klassenzuordnung mittels maschineller Lernverfahren ermöglicht. Besonderes Augenmerk liegt hierbei auf Krankheitsbildern, die ausschließlich anhand ihrer charakteristischen Dynamik identifiziert werden können.

Methode

Im Folgenden werden die einzelnen Schritte der phonovibrographischen Merkmalsgewinnung und der anschließenden Klassifikation erläutert.

Daten

Von einem Probandenkollektiv, bestehend aus insgesamt 120 Frauen, wurden endoskopische HG-Aufnahmen der schwingenden Stimmlippen während einer gehaltenen Phonation des Vokals /a/ aufgezeichnet (siehe Abbildung 1a+b). Bei jeweils 25 untersuchten Patientinnen lag eine diagnostizierte hyper- bzw. hypofunktionelle Stimmstörung vor. Für jeweils 15 Patientinnen wurde eine links- bzw. rechtsseitige Stimmlippenparese festgestellt. Die verbleibenden 40 Frauen wiesen keinerlei Stimmbeschwerden auf. Sie dienten daher als gesunde Referenzgruppe.

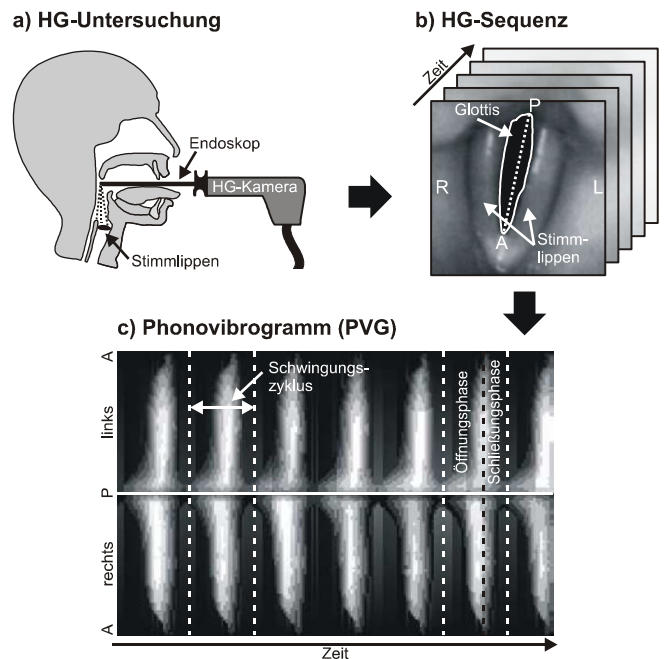


Abbildung 1: Endoskopische HG-Untersuchung der Stimmlippenschwingungen während der Phonation (a), die resultierende HG-Sequenz (b) und das zugehörige Phonovibrogramm (c).

Phonovibrographie

Für die aufgezeichneten HG-Videos wurden die Positionen der beiden Stimmlippen mittels eines speziellen Bildanalyseverfahrens segmentiert [2] und die zugehörigen Phonovibrogramme (PVG) berechnet (siehe Abbildung 1c). Ein PVG visualisiert hierbei kompakt die räumlich-zeitlichen Veränderungen entlang der gesamten Stimmlippenkante und erleichtert somit die klinische Bewertung der Symmetrie und Periodizität des Stimmlippenschwingungsmusters [2].

Merkmalsextraktion

Um die im PVG enthaltenen Bewegungsinformationen quantitativ zu erfassen, wurden zyklensweise normierte PVG-Konturmerkmale berechnet [1]. Indem sie Stimmlippenbereiche mit gleichartigen Auslenkungszuständen verbinden, erfassen diese Merkmalsbeschreibungen das charakteristische Öffnungs- und Schließungsverhalten der Stimmlippen sowie die Links-Rechts-Symmetrie (Abbildung 2). In der vorliegenden Studie wurden Konturlinien für 50%-Stimmlippenauslenkungen extrahiert, da diese die dynamischen Schwingungsvorgänge am besten abbilden [1]. Durch Berechnung der Mittelwerte und zugehörigen Standardabweichungen wurde zusätzlich die zeitliche Veränderlichkeit der berechneten Merkmale erfasst. Auf diese Weise wurde für jedes PVG ein beschreibender Merkmalsvektor bestimmt.

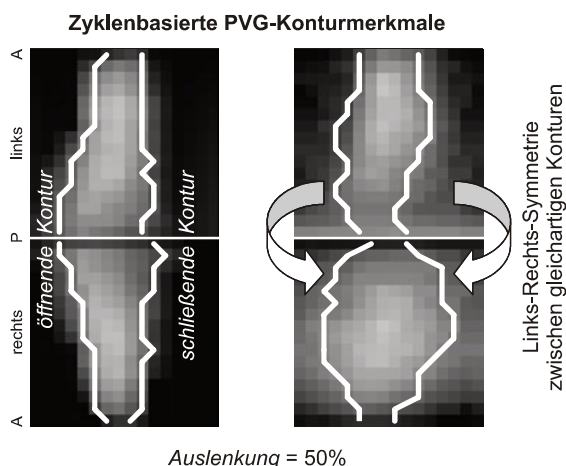


Abbildung 2: Zwei exemplarische normierte PVG-Zyklen mit eingezeichneten öffnenden und schließenden Konturen, die den Zustand der halben Stimmlippenauslenkung markieren (links). Zudem werden Symmetriemerkmal zwischen den Konturen berechnet (rechts).

Klassifikation und Evaluierung

Die resultierenden quantitativen Merkmalsbeschreibungen wurden mit verschiedenen Datenanalyseverfahren aus dem Bereich des maschinellen Lernens hinsichtlich ihrer automatischen Unterscheidbarkeit untersucht. Zum Einsatz kamen hierbei das k-Nächste-Nachbarn-Verfahren (k-NN), der C4.5-Entscheidungsbaum-Klassifikator und eine Support-Vektor-Maschine (SVM) mit linearer, quadratischer, kubischer und RBF-Kernelfunktion [3]. Für die beiden untersuchten Pathologien (Funktionell, Parese) wurden die in Tabelle 1 aufgeführten 2- und 3-Klassenprobleme betrachtet. Die freien Parameter der Lernalgorithmen wurden durch eine ($\mu+\lambda$)-Evolutionstrategie optimiert [4]. Die dabei erzeugten Klassifikationsmodelle ermöglichten die automatische Zuordnung der klinischen Daten zu den Krankheitsklassen. Zur Bewertung der erreichten Klassifikationsgüte wurden mittels 10-fach stratifizierter Kreuzvalidierung die Korrektklassifikationsrate, die Sensitivität und die Spezifität der Klassifikationsmodelle ermittelt.

Ergebnisse und Diskussion

In Abbildung 3 sind die für die einzelnen maschinellen Lernverfahren erreichten Korrektklassifikationsraten dargestellt. Hierbei wurde für beide Krankheitsbilder jeweils über alle zugehörigen 2-Klassenprobleme gemittelt.

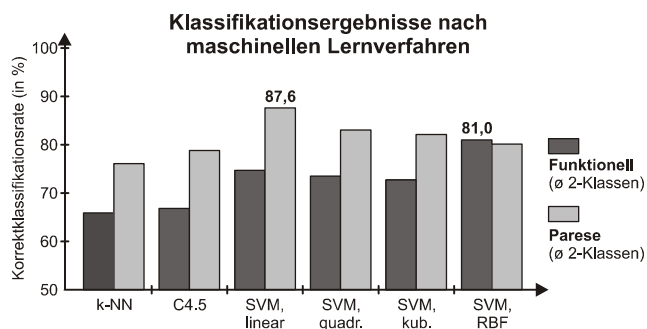


Abbildung 3: Mittlere Korrektklassifikationsraten, aufgeschlüsselt nach Klassifikationsalgorithmen.

Für die jeweils mit höchster Genauigkeit klassifizierenden Lernalgorithmen sind die zugrunde liegenden Einzelergebnisse in Tabelle 1 dargestellt.

Tabelle 1: Kreuzvalidierte Klassifikationsergebnisse

| Klassifikationsproblem | Klassifikationsgüte (in %) | | |
|------------------------------------|----------------------------|----------------------|----------------------|
| | Korrektklassif.rate | Sensitivität | Spezifität |
| Gesund vs. Hyperfunktion | 79,3 | 72,0 | 82,7 |
| Gesund vs. Hypofunktion | 83,3 | 85,3 | 77,3 |
| Hyper- vs. Hypofunktion | 80,3 | 80,0 | 70,7 |
| Gesund vs. Hyper- vs. Hypofunktion | 69,0 | 76,0 56,0 72,7 | 74,0 89,3 82,0 |
| Gesund vs. Parese links | 94,7 | 93,3 | 84,4 |
| Gesund vs. Parese rechts | 92,8 | 91,1 | 91,1 |
| Parese links vs. rechts | 75,3 | 62,2 | 80,0 |
| Gesund vs. Parese links vs. rechts | 70,3 | 75,5 51,1 55,6 | 86,7 78,9 75,6 |

Am zuverlässigsten konnten die Klassifikationsprobleme mit dem SVM-Lernverfahren gelöst werden. Die beste Performance zeigte hierbei für das Krankheitsbild der funktionellen Stimmstörung der RBF-Kernel, für die Paresen die lineare Kernelfunktion. Die Unterscheidung zwischen gesunden und pathologischen Stimmlippenschwingungen konnte für die Paresen mit höher Genauigkeit getroffen werden (~92%) als für die funktionellen Störungen (~81%). Als problematisch erwies sich die Identifikation der von der Parese betroffenen Stimmlippenseite (~75%). Mithilfe des vorgestellten PVG-Beschreibungs- und Klassifikationsansatzes kann der Arzt in Zukunft schneller, zuverlässiger und objektiv klinische Diagnosen stellen. In kommenden Studien wird die Analyse um Merkmale des akustischen Stimmsignals erweitert.

Literatur

- [1] Voigt D., Döllinger M., Yang A., Eysholdt U., Lohscheller J.: Automatic diagnosis of vocal fold palsy by employing phonovibrogram features and machine learning methods. *Comput Methods Programs Biomed* (2010), im Druck
- [2] Lohscheller J., Eysholdt U., Toy H., Döllinger M.: Phonovibrography: mapping high-speed movies of vocal-fold vibrations into 2-D diagrams for visualizing and analyzing the underlying laryngeal dynamics. *IEEE Trans Med Imaging* 27(3) (2008), 300-309
- [3] Duda R.O., Hart P.E., Stork D.G.: *Pattern classification*, 2. Aufl., John Wiley & Sons, New York, 2001
- [4] Beyer H.G., Schwefel H.P.: Evolution strategies – a comprehensive introduction. *Natural Computing* 1(1) (2002), 3-52