Computerbasierte Bewegungsanalyse von Stimmlippenschwingungen

Thomas Wittenberg♠, Ingo Popp♠, Monika Tigges♣ und Robert Schmidt♠

♠ Fraunhofer Institut für Intergrierte Schaltungen — Angewandte Elektronik
Am Weichselgarten 3, D-91058 Erlangen
♣ Städtisches Krankenhaus Karlsruhe
Email: {wbg,poi,sch}@iis.fhg.de

Zusammenfassung. Ein System zur automatischen Bewegungsanalyse von Stimmlippenschwingungen wird vor dem Hintergrund der für die Bewegungsganalyse bekannten Korrespondenz–, Verdeckungs–, und Aperturprobleme vorgestellt und diskutiert. Um die genannten Probleme auszugleichen, ist der beschriebene Ansatz wissensbasiert und flächenorientiert.

1 Motivation

Funktionelle Stimmstörungen, also Fehlfunktionen der Stimmlippen (ugs. Stimmbänder) äussern sich in der Regel durch Heiserkeit und können damit im Extremfall durch die Beeinträchtigung der wichtigsten zwischenmenschlichen Kommunikationsform, des Sprechens, eine sog. Kommunikationsstörung auslösen. Solche funktionellen Stimmstörungen können einerseits durch Fehlanwendung der Stimme bei Viel- und Berufssprechern (Veräufern, Politikern, Lehrern, Pfarrern) entstehen und andereseits als direkte Operationsfolge bei der Beatmungs-Intubationen auftreten, bei denen der Intubationsschlauch auf die Nerven der Stimmbänder drückt.

Die Stimmlippen des Menschen schwingen im Normalfall des Sprechens mit einer Grundfrequenz von etwa 120 Hz bei Männern und 250 Hz bei Frauen. Da das visuelle Wahrnehmungssystem des Menschen jedoch nur Bildfolgen mit einer Bildrate von bis zu 25 Einzelbildern pro Sekunde verarbeiten kann, bedarf es zur klinischen Untersuchung von Stimmlippenschwingungen einer Methode, mit der die schnellen Bewegungsabläufe im Kehlkopf entsprechend hochfrequent aufgenommen und in Zeitlupe wiedergegeben werden können.

Die derzeit meistverbreitete konventionelle klinische Untersuchungsmethode für Stimmstörungen, die Video-Stroboskopie [1], liefert für den Arzt zwar räumlich hochaufgelöste (Farb-)Bilder, die jedoch aufgrund das Abtasttheorems nur zeitlich unterabgetastetet Bewegungen aufnehmen und darstellen kann. Abhilfe können an dieser Stelle Hochgeschwindigkeitskamerasysteme liefern, die Stimmlippenschwingungen mit einer Abtastrate von bis zu 10000 Hz aufnehmen können. Das in den vergangenen zehn Jahren entwicklte Diagnosesytem

besteht zum einen aus einer miniaturisierten digitalen Hochgeschwindigkeitskamera mit Endoskop, und zum anderen aus einer Diagnosesoftware zur wissensbasierten semiautomatischen Analyse von Stimmlippenschwingungen. Dieses Diagnose System befindet sich zur Zeit in der Evaluierungsphase im Rahmen einer europäsischen Multicenterstudie.

2 Material und Methode

Aufgrund des Korrespondenz-, des Verdeckungs und des Aperturproblems lassen sich die Standardansätze der Bewegungsanalyse nicht für diese Anwendung anwenden.

Unter dem *Korrespondenzproblem* versteht man die Zuordnung von Objektpositionen zwischen zwei zeitlich verschiedenen Abbildungen eines Objektes [3], d.h. das Ziel einer Bewegungsanalyse ist die Abschätzung der Bewegung (Translation und Rotation) eines Objektpunktes zwischen diesen Einzelbildern. Alle bekannten Verfahren der Bewegunsganalyse basieren auf eine Wiedererkennung, also einer Korrespondenz eines Objektes im Nachfolgebild. Daraus ergeben sich die beiden anderen Probleme[4]:

- Die Existenz einer Lösung wird durch das *Verdeckungsproblem* beschrieben,
- Die Eindeutigkeit einer Lösung durch das *Aperturproblem*

Unter dem Begriff der *Verdeckung* versteht man die komplette oder partielle Verdeckung eines zu verfolgenden Objektes durch ein anderes Objekt innerhalb der Bildfolge. Durch eine solche Verdeckung lässt sich kein Bewegungsvektor mehr für das verdeckte Objekt berechnen. Im vorliegenden Fall verschwinden die beiden zu untersuchenden Stimmlippenkanten beim glottalen Verschluss durch den gegenseitigen Kontakt.

Als *Aperturproblem* wird die Mehrdeutigkeit des Verschiebungsvektors eines Punktes an Kanten und homogenen Flächen bezeichnet, d.h. innerhalb einer kleinen Umgebung (der *Apertur*) um den Punkt des Interesses, lässt sich ein Verschiebungsvektor lediglich senkrecht zu der Kante in dieser Umgebung berechen, auch wenn an dieser Stelle zuätzlich eine Bewegung parallel zu dieser Kante stattfindet. Bei der Bewegungsanalyse von Stimmlippen tritt dieses Problem speziell im geöffneten Zustand auf an den freien Kanten der Stimmlippen auf, da deren Gradienten nicht immer den Bewegungsrichtungen der Stimmlippen entsprechen [3].

Entsprechend der beschriebenen Problematik wurde ein flächenbasiertes Verfahren implementiert [2]. Die computergestützte Bewegungsanalyse von Stimmlippenschwingungen basiert danach auf der Hypothese, dass die Glottis (die Stimmritze) in den geöffneten Zuständen der einzelnen Schwingungsperioden als mehr oder weniger zusammenhängende dunkle Fläche sichtbar ist, die von den zu verfolgenden Stimmlippen eingeschlossen ist. Auf der Hypothese lassen sich die nachfolgenden Schritte durchgeführen:

- Identifikation der derjenigen Bilder einer Periode mit geöffneter Glottis. Dies geschieht mit der Berechnung der Bildenergie in den ROI's aus den Einzelbildern aus einem Zeitfenster, das zeitlich symmetrisch um das aktuell betrachtete Bild liegt.
- Aus Bildern mit geöffneter Glottis werden die benötigten Parameter (Startkoordinaten, Zusammenhangskriterium, etc...) für die Segmentierung der Glottis geschätzt.
- Für jedes Bild mit geöffneter Glottis erfolgt eine Flächensegmentierung der Glottis durch ein adaptives Regionen-Wachstumsverfahren, anschliessend wird eine Verifikation und Korreketur der glottalen Fläche mit einem elliptischen Modell durchgeführt.
- Aus der glottalen Fläche wird die glottalen Hauptachse (die unabhängig von der Hauptträgheitschse des Fflächesegmentes verläuft) bestimmt, die von der vorderen zur hinteren Kommissur des Kehlkopfes reicht. Anschliessend Approximation der orthogonal dazu liegenden Bewegungsachsen der Stimmlippen an drei für den untersuchenden Phoniater wichtigen Stellen.
- Bei geschlossenen Stimmlippen Schätzung der Trajektorienpunkte mittels der glottalen Hauptachse.
- Extraktion der Trajektorienpunkte der Stimmlippen auf den Bewegungsachsen und Konkatenation der Einzelpunkte zu Bewegungskurven. Anschliessend Normierung der XY-Komponenten der Bewegung im Karthesischen auf den Abstand der Stimmlippen von der glottalen Hauptachse

Alle Schritte lassen sich in der Regel vollautomatisch durchführen, jedoch kann der Untersucher im Bedarfsfall die Berechnung jederzeit unterbrechen und mit manuell geänderten Parametern (ROI, Segmentierungsschwelle, etc..) reinitialisieren. Aus den berechneten, auf die glottale Hauptachse normierten Trajektorien lassen sich interaktiv weitere Parameter wie Einschwingzeit, Amplituden und Phasen, sowie Grundfrequenzverläufe extrahieren.

3 Ergebnisse und Ausblick

In den vergangenen zehn Jahren wurde das Aufnahme- und Auswerteverfahren zusammen mit der phonatrischen Abteilung der Erlanger HNO-Klinik routinemässig eingesetzt und in dieser Zeit auf Grund der aus über 6000 Aufnahmen und Auswertungen gesammelten Erfahrungen iterativ verfeinert und verbessert. Auf der Basis der extrahierten Bewegungskurven entstanden u.a. Studien zur Verträglichkeit von Anästhetikum, zur Klassifikation von funktionellen Stimmstörungen, und zur Evalierung von Stimmbandstörungen auf Grund von Stimmbandlähmungen.

Als repräsentatives Beispiel sei hier eine retrospektive Studie über 123 Patienten mit einseitiger Stimmlippenparese genannt. Alle Bildsequenzen wurden zunächst in Zeitlupenwiedergabe von zwei erfahrenen Phoniatern visuell beurteilt, wobei die aus der Video Stroboskopie bekannten und etablierten Kriterien zur Beurteilung verwendet wurden, namentlich die Berurteilung der Randkantenverschiebung sowie die Art bzw. das Vorhandensein einer supraglottische

Kompensation. Zusätzlich zur subjektiven Beurteilung wurden die Bildsequenzen quantitativ mittels des o.g. automatischen Bildverarbeitungssystems ausgewertet. Quantitativ wurden der Beginn der Stimmlippenschwingungen im Seitenvergleich, der Glottisschluß, die Amplitudenrelation von gesunder und gelähmter Seite, die Einschwingzeit, die Grundfrequenzen und die Öffnungsquotienten analysiert [5].

In den vergangenen zwei Jahren entstand auf der Basis der langjährigen Vorarbeiten ein CI-und MPG konformer, klinischer Prototyp einer miniaturisierten Hochgeschwindigkeitskamera inklusive der oben beschriebenen Auswerte-Software zur Datenreduktion und Parameterquantifizierung. Derzeit wird dieser Prototyp im Rahmen einer derzeit beginnenden europäschien Multicenterstudie mit 14 Teilnehmern (aus Deuschland, Holland, Schweden, Österreich und der Schweiz) evaluiert. Als Referenz für die Evaluierung der neuen Untersuchungsmethode wird der aktuelle weltweit etablierte 'Goldstandard' für die phoniatrische Untersuchung des Kehlkopfes — die Videostroboskopie — verwendet, um quantifizierbare Ergenisse zu erhalten.

Literatur

1. Böhme G, Gross M: *Stroboskopie und andere Verfahren zur Analyse der Stimmlippenschwingungen*, Median Verlag, Heidelberg, 2001.
2. Wittenberg Th: *Wissensbasierte Bewegungsanalyse von Stimmlippenschwingungen anhand digitaler Hochgeschwindigkeitsaufnahmen*, Shaker Verlag Aachen 1998.
3. Kirchner H: *Bewegungsanalyse in Bildfolgen: Ein mehrstufiger Ansatz*, Deutscher Universitäts-Verlag, Wiesbaden, 1993.
4. Tekalp AM: *Digital Video Processing*, Prentice Hall, 1995.
5. Tigges M, Rydell R et al : *Einseitige Stimmlippenparese - wie bewegt sich die gelähmte Stimmlippe?* In: Aktuelle Phoniatrisch Pädaudiologische Aspekte, 1999/2000, Band 7, E. Kruse (Hrsg.), Median Verlag Heidelberg, 2000.