

3D Reconstruction Of Vocal Fold Surface Dynamics In Functional Dysphonia

**Der Technischen Fakultät
der Friedrich-Alexander-Universität
Erlangen-Nürnberg
zur
Erlangung des Doktorgrades**

DOKTOR-INGENIEUR

vorgelegt von
Dipl.-Ing. Georg Luegmair
aus Bombay, Indien

**Als Dissertation genehmigt
von der Technischen Fakultät der
Friedrich-Alexander-Universität Erlangen-Nürnberg**

Tag der mündlichen Prüfung: 26.11.2013

Vorsitzende des

Promotionsorgans:

Gutachter:

Prof. Dr.-Ing. habil. Marion Merklein

Prof. Dr.-Ing. Michael Döllinger

Prof. Dr.-Ing. Günther Greiner

3D Rekonstruktion Von Stimmlippenschwingungen Bei Funktioneller Dysphonie

**Der Technischen Fakultät
der Friedrich-Alexander-Universität
Erlangen-Nürnberg
zur
Erlangung des Doktorgrades**

DOKTOR-INGENIEUR

vorgelegt von
Dipl.-Ing. Georg Luegmair
aus Bombay, Indien

Kommunikationsstörungen - Berichte aus Phoniatrie und
Pädaudiologie

Band 23

Georg Luegmair

**3D Reconstruction of Vocal Fold
Surface Dynamics in Functional Dysphonia**

3D Rekonstruktion von Stimmlippenschwingungen
bei Funktioneller Dysphonie

D 29 (Diss. Universität Erlangen-Nürnberg)

Shaker Verlag
Aachen 2014

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Zugl.: Erlangen-Nürnberg, Univ., Diss., 2013

Copyright Shaker Verlag 2014

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8440-2470-8

ISSN 1436-1175

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen

Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9

Internet: www.shaker.de • E-Mail: info@shaker.de

Danksagung

Nach einiger Zeit der intensiven Arbeit habe ich meine Dissertation abgeschlossen. Auf dem Weg zu diesem Punkt hatte ich viele Begleiter und Unterstützer, bei denen Ich mich bedanken möchte.

Ich danke meinem Doktorvater und Betreuer Prof. Dr. Michael Döllinger herzlichst für die Aufnahme in seine Arbeitsgruppe, zuerst mit einem Thema für die Diplomarbeit und weiterführend dann für die Promotion. Ich danke für die Möglichkeiten und Chancen zu außergewöhnlichen Erfahrungen, die er mir eröffnet hat. Außerdem danke ich für die Freiheit, während der Bearbeitung meines Promotionsthemas neue Ideen zu verfolgen und umsetzen zu können.

Bei Prof. Dr. Günter Greiner bedanke ich mich herzlich für die Übernahme des Koreferats und seine Bereitschaft, das Gutachten unter erheblichem Zeitdruck zu erstellen. Dies hat mir meinen weiteren Weg in der Zukunft erheblich erleichtert.

Bei Prof. Dr. Dr. Ulrich Eysholdt möchte ich mich für die Aufnahme die Abteilung für Phoniatrie und Pädaudiologie, sowie seine Unterstützung bei meinen Unternehmungen bedanken.

Ich möchte ich mich auf diesem Wege auch bei meinen Kolleginnen und Kollegen der letzten Jahre bedanken: Dr. Daniel Voigt, Dr. Anxiong Yang, Dr. Anke Ziethe, Veronika Birk, Denis Dubrovskiy, Simon Petermann und Marion Semmler. All diese Personen haben das Arbeitsumfeld der letzten Jahre sehr angenehm gestaltet und Arbeit nicht nur Arbeit sein lassen.

Ein besonderer Dank gilt meinem ehemaligem Kollegen Dr. Björn Hüttner. Er hat wie kein anderer den Werdegang dieser Arbeit erlebt und mir stets mit Rat und Tat zur Seite gestanden, wenn ich ihn gebraucht habe. Ich habe die langjährige Zusammenarbeit als mehr als angenehm empfunden.

Herausheben möchte ich auch meine Frau Rosa. Sie war, gerade in der Zeit der Fertigstellung dieser Arbeit, der Rückhalt den ich gebraucht habe und ich konnte mich ihrer Unterstützung immer sicher sein.

Meinen Eltern danke ich dafür, dass sie mir beigebracht haben immer volles Herzblut

in eine Sache zu stecken und diese dann auch Konsequenz zu Ende zu führen und dafür, dass sie mich die letzten Jahre unterstützt haben.

Dr. Daryush Mehta und Dr. James Kobler danke ich für die Einladung in ihr Labor und die fruchtbare Zusammenarbeit.

Zu guter Letzt möchte ich mich bei allen ehemaligen und aktuellen nicht namentlich genannten Kollegen der Abteilung für Phoniatrie und Pädaudiologie und der HNO des Universitätsklinikums Erlangen danken, die wissentlich und unwissentlich einen Beitrag zum erfolgreichen Abschluss dieser Arbeit beigetragen haben.

- Erlangen, den 24.11.2013

Summary

Voice and speech are human's primary means of communication. A voice disorder (dysphonia) impairs the ability to speak and consequently, the ability to communicate. Therefore, dysphonia has a grave influence not only on the social life, but also on the work life and finances of the person concerned.

Paramount to an effective treatment of pathologies creating voice disorders is a proper understanding of the underlying morphology and physiology of the voice producing organs. Physiology of phonation includes the oscillations of the vocal folds during phonation. These oscillations generate pressure fluctuations in the air stream from the lungs, which are perceived as sound or more specific, as voice. The focus of this thesis lies on these oscillations. The vibratory behavior in pathological cases was found to differ from healthy vocal folds, i.e. the vocal folds oscillate less periodically, regularly and/or symmetrically.

To gain deeper insight into vocal fold dynamics, a new method is presented in this thesis. Based on stereotriangulation, it allows to measure irregular vocal fold oscillations by reconstruction of the entire vocal fold surface in 3D.

The setup for this purpose consists of a high-speed camera and a LASER projection system. Both elements allow for a easy miniaturization with regard to future in vivo applications through the use of an endoscope and fiber optics. The high-speed camera allows for a high temporal resolution, which is necessary to capture irregular vocal fold motion. The LASER projection system emits a bundle of rays onto the vocal fold surface, creating a regular grid of LASER points. The points are captured with the high-speed camera. For the reconstruction of the surface, the LASER points have to be segmented from the high-speed footage.

In this thesis, the author presents the algorithms necessary for the image processing of the calibration images, calibration of the setup, segmentation of the LASER points from the high-speed footage, reconstruction of the 3D coordinates, and the anatomically correct interpolation of the entire vocal fold surface.

The developed method is applied to two excised human larynges in an experimental

study. The larynges are examined in an in vitro experimental setup, which allows to replicate phonation. Additionally, asymmetric vocal fold tension can be induced, as e.g. is the case of muscle tension dysphonia. An investigation of phonation under these circumstances is unique and the first of this kind.

The relations of vocal fold dynamics, 2D as well as 3D, with the produced acoustical signal is investigated with regard to muscle tension asymmetry. Several significant correlations are found between 2D dynamics and acoustics. An even higher number of significant correlations was found between 3D dynamics and acoustics. This indicates that the presented method adds a dimension to the measured data, which can be significant in future investigations. In addition to the correlations, a list of significant parameters for both larynges was ascertained. The list contains 4-7 different parameters per domain, which is an immense reduction in relation to the number of parameters that were derived initially from the raw data. This helps to reduce the number of parameters that needs to be computed and evaluated and still obtain significant findings in future studies.

Zusammenfassung

Stimme und Sprache sind das wichtigste menschliche Kommunikationsmittel. Eine Stimmstörung (Dysphonie) schränkt die Sprachfähigkeit ein und limitiert dadurch die Kommunikation. Dies hat nicht nur negative Auswirkungen auf das Sozialleben, sondern auch auf das Arbeitsleben der betroffenen Personen und der damit verbundenen finanziellen Konsequenzen.

Vorrangig bei der Behandlung einer Erkrankung die zur Störung der Stimmbildung führen, ist ein umfassendes Verständnis der zugrundeliegenden Morphologie und Physiologie der stimmbildenden Organe. Teil der Stimmgebung sind die Oszillationen der Stimmlippen während der Phonation, welche Druckschwankungen im ausgeatmeten Luftstrom erzeugen und als Ton bzw. Stimme wahrgenommen werden. Der Fokus dieser Arbeit liegt auf diesen Oszillationen. In Pathologien weichen die Vibrationsmuster von denen gesunder Stimmlippen ab, d.h. sie sind weniger periodisch, regulär und/oder symmetrisch.

Um tieferes Verständnis der Bewegungen der Stimmlippen bei der Phonation zu gewinnen, wird eine neue Methode in dieser Arbeit vorgestellt. Basierend auf der Stereotriangulation erlaubt diese es, die dreidimensionalen, irregulären Stimmlippenoszillation durch die Rekonstruktion der gesamten Stimmlippenoberfläche zu erfassen. Der Messaufbau besteht aus einer Hochgeschwindigkeitskamera und einem LASER-Projektor. Beide Elemente erlauben eine einfache Miniaturisierung für einen zukünftig vorgesehenen in vivo Einsatz durch die Verwendung eines Endoskops und Fiberoptiken. Die Hochgeschwindigkeitskamera ermöglicht eine hohe zeitliche Auflösung, welche für die Aufnahme von irregulären Stimmlippenschwingungen notwendig ist. Der LASER-Projektor projiziert ein Strahlenbündel auf die Stimmlippenoberfläche, was ein regelmäßiges Gitter aus LASER-Punkten auf der Oberfläche erzeugt. Dieses Gitter wird dann von der Kamera erfasst. Für die Rekonstruktion der Stimmlippenoberfläche müssen die Punkte aus der Kameraaufnahme extrahiert werden. In dieser Arbeit werden die dafür notwendigen Algorithmen vorgestellt, d.h. die Bildverarbeitung für die Kalibrierung, die Segmentierung der LASER-Punkte aus den Aufnahmen, die Rekonstruktion der 3D

Koordinaten und die anatomisch korrekte Interpolation der gesamten Stimmlippenoberfläche.

Die Rekonstruktion wird in einer experimentellen Studie auf zwei exzidierte Kehlköpfe angewendet. Dazu werden die Kehlköpfe in einem Aufbau für in vitro Versuche untersucht. Der Aufbau erlaubt die Reproduktion der Stimmgebung. Zusätzlich können asymmetrische Stimmlippenspannungen erzeugt werden, wie sie z.B. bei funktionellen Dysphonien vorkommen. Die 3D Rekonstruktion und Einbringung von Asymmetrien zur Untersuchung der Korrelationen zwischen 2D, 3D und akustischen Parametern sind Alleinstellungsmerkmale dieser Studie machen sie damit zur Ersten dieser Art. Signifikante Korrelation zwischen der Stimmlippenbewegung und dem produzierten Stimmsignal sind erkennbar, wobei eine Mehrheit der Korrelationen zwischen 3D Parametern und Akustik existiert. Zusätzlich wird eine Liste an aussagekräftigen Parametern erstellt. Diese enthält 4-7 Parameter pro Datensatz (2D, 3D, Akustik), was eine immense Reduzierung im Gegensatz zur Anzahl der ausgewerteten Ausgangsparameter darstellt. Dadurch kann sich der Aufwand der zu berechnenden und auszuwertenden Größen bei gleicher Aussagekraft in zukünftigen Studien verringern.

Contents

1	Introduction	1
2	Anatomy and physiology of phonation	5
2.1	Anatomy	5
2.1.1	Cartilages	6
2.1.2	Muscles and ligaments	8
2.2	Physiology of phonation	8
3	3D reconstruction from high-speed imaging	11
3.1	Introduction	11
3.2	Basics of stereo-triangulation	13
3.3	General reconstruction setup	14
3.4	Camera model	16
3.5	Laser projection system	17
3.6	Calibration of the setup	18
3.6.1	Overview of the calibration method	18
3.6.2	Homographic transformation	20
3.6.3	Image processing for camera and laser calibration images	21
3.6.4	Calibration of camera	31
3.6.5	Calibration of the LPS	32
3.6.6	Consideration of optical distortion	35
3.6.7	Error estimation method	36
3.7	3D reconstruction of the vocal fold surface	39
3.7.1	Overview of the reconstruction process	39
3.7.2	Tracking	40
3.7.3	Filtering by Principle Component Analysis	45
3.7.4	Reconstruction of 3D coordinates	45
3.7.5	Mean period computation	45

3.7.6	Vocal fold surface interpolation	51
3.8	Computation of parameters	59
3.8.1	Vocal fold edge tracking	59
3.8.2	Mucosal wave tracking	61
3.9	Analyzed parameters	62
3.10	Discussion	63
4	Controlled arytenoids manipulation	67
4.1	Objective	67
4.2	Implementation	68
4.3	Exemplary application	69
5	Application: Hoarseness induced by asymmetry	71
5.1	Introduction	71
5.2	Methods	73
5.3	With glottal gap (Human 2)	75
5.3.1	Mean-period error	75
5.3.2	Correlation: Asymmetry vs. 2D	76
5.3.3	Correlation: Asymmetry vs. acoustics	79
5.3.4	Correlation: Asymmetry vs. 3D	80
5.3.5	Correlation: 2D/3D vs. acoustics	82
5.4	Without glottal gap (Human 3)	86
5.4.1	Correlation: Asymmetry vs. 2D	87
5.4.2	Correlation: Asymmetry vs. acoustics	91
5.4.3	Correlation: Asymmetry vs. 3D	93
5.4.4	Correlation: 2D/3D vs. acoustics	95
5.5	Discussion	99
5.5.1	Phonation parameters	99
5.5.2	2D vs. 3D	106
5.6	Conclusion	108
6	Summary & Outlook	109
A	Parameter definitions	123
A.1	GAW parameters	123
A.1.1	Perturbation factors	123
A.1.2	Jitter related parameters	125

A.1.3	Shimmer related parameters	126
A.1.4	Energy related parameters	127
A.1.5	Noise related parameters	128
A.1.6	Glottal parameters	130
A.1.7	Amplitude to length ratio	131
A.2	Acoustic parameters	132
B	Tables	135