

Model-Based Quantification of Pathological Voice Production

Der Technischen Fakultät der
Universität Erlangen-Nürnberg
zur Erlangung des Grades

DOKTOR-INGENIEUR

vorgelegt von
Dipl.-Ing. Raphael Schwarz

Erlangen - 2007

Als Dissertation genehmigt von
der Technischen Fakultät der
Universität Erlangen-Nürnberg

Tag der Einreichung 10.04.2007
Tag der Promotion: 15.05.2007
Dekan : Prof. Dr. A. Leipertz
Berichterstatter : Prof. Dr. Dr. U. Eysholdt
Prof. Dr. R. Lerch

Modellbasierte Quantifizierung der pathologischen Stimmgebung

Der Technischen Fakultät der
Universität Erlangen-Nürnberg
zur Erlangung des Grades

DOKTOR-INGENIEUR

vorgelegt von
Dipl.-Ing. Raphael Schwarz

Erlangen - 2007

Kommunikationsstörungen - Berichte aus Phoniatrie und
Pädaudiologie

Band 16

Raphael Schwarz

**Model-Based Quantification of
Pathological Voice Production**

D 29 (Diss. Universität Erlangen-Nürnberg)

Shaker Verlag
Aachen 2007

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

Zugl.: Erlangen-Nürnberg, Univ., Diss., 2007

Copyright Shaker Verlag 2007

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

Printed in Germany.

ISBN 978-3-8322-6322-5

ISSN 1436-1175

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen

Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9

Internet: www.shaker.de • e-mail: info@shaker.de

Danksagung

Meinem Doktorvater Prof. Dr. Dr. Ulrich Eysholdt danke ich für die Aufnahme in seiner Abteilung und für die Überlassung des Themas. Sein Engagement ermöglichte interdisziplinäre Strukturen, welche eine substantielle Basis meiner Arbeit waren.

Bei Herrn Prof. Dr. Reinhard Lerch möchte ich mich für die gewährte Unterstützung und für die Übernahme des Korreferats bedanken.

Ein großer Dank geht an meine beiden Betreuer Dr. Jörg Lohscheller und PD Dr. Michael Döllinger, die sich stets Zeit für Diskussionen nahmen und viele wertvolle Anregungen zur Arbeit gaben. Insbesondere bin ich ihnen dankbar für ihre Mühen beim Durcharbeiten meiner Manuskripte und die resultierenden konstruktiven Kritiken.

Martin Burger und Tobias Wurzbacher haben den Entstehungsprozess der vorliegenden Arbeit unmittelbar miterlebt und konstruktiv unterstützt. Für die zahlreichen Diskussionen und Anregungen möchte ich mich herzlich bedanken.

Ebenso danke ich Prof. Dr. Dr. Ulrich Hoppe für die intensive Zusammenarbeit.

Weiterhin bin ich den Mitarbeitern der Abteilung für Phoniatrie und Pädaudiologie zu Dank verpflichtet. Namentlich seien hier Dr. Hikmet Toy und PD Dr. Maria Schuster für die Erstellung der Hochgeschwindigkeitsaufnahmen und für die geduldige und aufschlussreiche Beantwortung medizinischer Fragestellungen erwähnt.

Ein ganz besonderer Dank gilt meiner Familie, die meinen bisherigen Lebensweg so liebevoll unterstützt hat. Ebenso bedanke ich mich bei meiner Freundin Dorothee, die mir durch ihr Interesse an meiner Arbeit sowie ihr Verständnis und aufgebrachte Geduld während der Erstellung des Manuskripts eine wertvolle Stütze war.

Die vorliegende Arbeit wurde im Rahmen des Projekts DFG Ey15/10 der Deutschen Forschungsgemeinschaft sowie des Teilprojekts B5 des Sonderforschungsbereichs 603 "Modellierung und Visualisierung komplexer Szenen und Sensordaten" gefördert.

Erlangen, Mai 2007.

Zusammenfassung

Heiserkeit, das primäre Symptom einer Stimmstörung, beruht auf irregulären Stimmlippenschwingungen. Der vollständige Verlust der Stimme ist wohl die extremste Form einer Stimmstörung. Die onkologische Therapie von Kehlkopfkrebs kann solch einen Stimmverlust zur Folge haben, wenn Lage und Zustand des Tumors eine vollständige Entfernung des Kehlkopfes und damit der Stimmlippen notwendig macht. Infolgedessen verliert der Patient zunächst die Fähigkeit zu sprechen. Um in solch einem Fall eine Stimmrehabilitation zu ermöglichen, wird Narbengewebe am oberen Ende der Speiseröhre (PE-Segment) zur Ersatzstimmgebung genutzt. Die Qualität sowohl von Normal- als auch Ersatzstimme hängt maßgeblich von der Anatomie und den Schwingungsmustern der stimmgebenden Elemente ab. Mittels endoskopischer Hochgeschwindigkeitsaufnahmen können die Tongeneratoren während der Stimmgebung in Echtzeit beobachtet werden. Die digitale Aufnahmetechnik ermöglicht eine rechnergestützte Auswertung der erzeugten Bilddaten.

In der vorliegenden Arbeit werden Methoden zur modellbasierten Analyse und objektiven Quantifizierung der aus den digitalen Bilddaten extrahierten Schwingungsmuster entwickelt. Prinzipiell wird die Dynamik eines biomechanischen Modells des Tongenerators an das tatsächlich beobachtete Schwingungsmuster unter Verwendung einer automatischen Parameteroptimierung angepasst. Dabei wird neben der Frequenz und Amplitude auch die spezifische Wellenform der Schwingung berücksichtigt. Die Anpassung der Modelldynamik stellt ein Optimierungsproblem mit nicht-konvexer Zielfunktion dar. Die berechneten Werte der Modellparameter liefern eine objektive Quantifizierung der Dynamik des Tongenerators. Stabilität und Zuverlässigkeit der entwickelten Optimierungsalgorithmen werden mit synthetisch generierten Schwingungsmustern nachgewiesen. Erste Studien mit klinischen Datensätzen belegen die Anwendbarkeit der entwickelten Methoden zur objektiven Quantifizierung und Klassifizierung von Schwingungsmustern der Stimmlippen und des PE-Segments.

Contents

1	Introduction	1
2	Fundamentals	5
2.1	Functional Anatomy of the Larynx	6
2.2	Voice Generating Process	12
2.3	Voice Disorders	13
2.3.1	Organic Voice Disorders	13
2.3.2	Functional Voice Disorders	14
2.4	Total Laryngectomy	14
2.4.1	Laryngeal Cancer	15
2.4.2	Substitute Voice Generation	16
3	High-Speed Recordings	19
3.1	High-Speed Imaging of the Vocal Folds	20
3.2	High-Speed Imaging of the PE Segment	22
4	Image Processing	25
4.1	Extraction of the Vocal Fold Dynamics	25
4.1.1	Single-Line Trajectories	27
4.1.2	Multi-Line Trajectories	29
4.2	Extraction of the PE Dynamics	30
5	Model-Based Classification	33
5.1	Two-Mass Model	35
5.1.1	Anchor Force	37
5.1.2	Vertical Coupling Force	38
5.1.3	Collision Force	38
5.1.4	Driving Force	39

5.2	Simulation of Vocal Fold Vibrations	40
5.2.1	Defining Non-Standard Model Configurations	41
5.2.2	Dynamics of Asymmetric Model Configurations	41
5.3	Objective Function	43
5.4	Automatic Parameter Optimization	45
5.4.1	Initial Value Search	46
5.4.2	Optimization Procedure	49
5.5	Validation of the Parameter Optimization	50
5.6	Application to Clinical Data	51
5.6.1	Clinical Data	51
5.6.2	Adaptation Results	53
5.6.3	Interpretation of the Model-Based Classification	57
6	Spatio-Temporal Quantification	63
6.1	Multi-Mass Model	64
6.1.1	Anchor Force	68
6.1.2	Vertical Coupling Force	68
6.1.3	Longitudinal Coupling Force	68
6.1.4	Collision Force	69
6.1.5	Driving Force	70
6.2	Optimization Parameters	72
6.3	Objective Function	72
6.4	Automatic Parameter Optimization	73
6.4.1	Model Initialization	74
6.4.2	Optimization Procedure	75
6.5	Validation with Synthetic Data Sets	78
6.5.1	Validation Results	79
6.6	Application to Real Vocal Fold Oscillations	81
6.7	Interpretation of the Clinical Results	89
7	Quantification of PE Segment Vibrations	93
7.1	Multi-Mass Model of the PE Segment	93
7.2	Automatic Parameter Optimization	96
7.2.1	Optimization Parameters	96
7.2.2	Objective Function	96
7.2.3	Model Initialization and Optimization Procedure	98
7.2.4	Validation with Synthetic Data Sets	99

7.3 Application to Real PE Segment Oscillations	101
8 Summary and Outlook	105
A PE-MMM Fitting Results	109
List of Abbreviations	113
Bibliography	115
List of Figures	129
List of Tables	131
Curriculum Vitae	133

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen	5
2.1	Funktionelle Anatomie des Kehlkopfes	6
2.2	Stimmgebung	12
2.3	Stimmstörungen	13
2.3.1	Organische Stimmstörungen	13
2.3.2	Funktionelle Stimmstörungen	14
2.4	Vollständige Laryngektomie	14
2.4.1	Kehlkopfkrebs	15
2.4.2	Ersatzstimmgebung	16
3	Hochgeschwindigkeitsaufnahmen	19
3.1	Hochgeschwindigkeitsaufnahmen der Stimmlippen	20
3.2	Hochgeschwindigkeitsaufnahmen des PE-Segments	22
4	Bildverarbeitung	25
4.1	Extraktion der Stimmlippendynamik	25
4.1.1	Einlinien-Trajektorien	27
4.1.2	Mehrlinien-Trajektorien	29
4.2	Extraktion der PE-Dynamik	30
5	Modellbasierte Klassifizierung	33
5.1	Zwei-Massen-Modell	35
5.1.1	Ankerkraft	37
5.1.2	Vertikale Kopplungskraft	38

5.1.3	Kollisionskraft	38
5.1.4	Antreibende Kraft	39
5.2	Simulation von Stimmlippenschwingungen	40
5.2.1	Definition verschiedener Modellkonfigurationen	41
5.2.2	Dynamik asymmetrischer Modellkonfigurationen	41
5.3	Zielfunktion	43
5.4	Automatische Parameteroptimierung	45
5.4.1	Startwertsuche	46
5.4.2	Optimierungsprozedur	49
5.5	Validierung der Parameteroptimierung	50
5.6	Anwendung auf klinische Datensätze	51
5.6.1	Klinische Daten	51
5.6.2	Anpassungsergebnisse	53
5.6.3	Interpretation der modellbasierten Klassifizierung	57
6	Räumlich-Zeitliche Quantifizierung	63
6.1	Mehr-Massen-Modell	64
6.1.1	Ankerkraft	68
6.1.2	Vertikale Kopplungskraft	68
6.1.3	Longitudinale Kopplungskraft	68
6.1.4	Kollisionskraft	69
6.1.5	Antreibende Kraft	70
6.2	Optimierungsparameter	72
6.3	Zielfunktion	72
6.4	Automatische Parameteroptimierung	73
6.4.1	Modellinitialisierung	74
6.4.2	Optimierungsprozedur	75
6.5	Validierung mit synthetischen Datensätzen	78
6.5.1	Validierungsergebnisse	79
6.6	Anwendung auf klinische Datensätze	81
6.7	Interpretation der klinischen Ergebnisse	89
7	Quantifizierung von Schwingungen des PE-Segments	93
7.1	Mehr-Massen-Modell des PE-Segments	93
7.2	Automatische Parameteroptimierung	96
7.2.1	Optimierungsparameter	96
7.2.2	Zielfunktion	96

7.2.3	Modellinitialisierung und Optimierung	98
7.2.4	Validierung mit synthetischen Daten	99
7.3	Anwendung auf reelle Schwingungen des PE-Segments	101
8	Zusammenfassung und Ausblick	105
A	Anpassungsergebnisse des PE-MMM	109
	Abkürzungsverzeichnis	113
	Literaturverzeichnis	115
	Abbildungsverzeichnis	129
	Tabellenverzeichnis	131
	Lebenslauf	133