

Kommunikationsstörungen

Berichte aus Phoniatrie und Pädaudiologie

Herausgeber: M. Döllinger

Begründet 1996 von U. Eysholdt

Marion Semmler

**Endoscopic laser-based 3D
imaging for *in vivo* examination
of human phonation**

Endoskopische laserbasierte 3D
Bildgebung zur *in vivo* Untersuchung
der menschlichen Phonation

**SHAKER
VERLAG**

Endoscopic laser-based 3D imaging
for *in vivo* examination of human phonation

Endoskopische laserbasierte 3D Bildgebung
zur *in vivo* Untersuchung der menschlichen Phonation

Der Technischen Fakultät
der Friedrich-Alexander-Universität
Erlangen-Nürnberg

zur
Erlangung des Doktorgrades
DOKTOR-INGENIEUR

vorgelegt von
Marion Semmler
aus Greding

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

Tag der mündlichen Prüfung: 19. März 2018
Vorsitzender des Promotionsorgans: Prof. Dr.-Ing. Reinhard Lerch
Erstgutachter: Prof. Dr.-Ing. Michael Döllinger
Zweitgutachter: Prof. Dr.-Ing. Michael Schmidt

Kommunikationsstörungen - Berichte aus Phoniatrie und
Pädaudiologie

Band 26

Marion Semmler

**Endoscopic laser-based 3D imaging for
in vivo examination of human phonation**

Endoskopische laserbasierte 3D Bildgebung zur
in vivo Untersuchung der menschlichen Phonation

D 29 (Diss. Universität Erlangen-Nürnberg)

Shaker Verlag
Aachen 2018

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., 2018

Copyright Shaker Verlag 2018

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 978-3-8440-5916-8

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

Obwohl ich der Phoniatrie vorerst erhalten bleibe, geht mit dem Abschluss meiner Doktorarbeit ein wichtiger Lebensabschnitt für mich zu Ende. Ich möchte diese Gelegenheit nutzen, um mich für die schöne Zeit in der Abteilung und für all die verschiedenen Arten von Unterstützung und Hilfe von allen Seiten zu bedanken.

Ich bin froh und stolz Teil einer so spannenden Arbeitsgruppe gewesen zu sein, in der die Verbindung von Klinik und Wissenschaft wie selbstverständlich gemeistert wird und in der die verschiedenen Disziplinen so harmonisch und gewinnbringend zusammenarbeiten. Ich habe die vielen verschiedenen Herangehensweisen und Blickwinkel auf Probleme und Lösungen als etwas sehr Wertvolles zu schätzen gelernt.

Das gemütliche Zusammensein bei gemeinsamen Ausflügen genauso wie die lebhaften Unterhaltungen bei unseren Mittags- und Kaffeepausen haben ein sehr angenehmes und für mich fast familiäres Arbeitsklima geschaffen und die Arbeit immer wieder auch zum Vergnügen gemacht. Ich möchte allen aktuellen und ehemaligen Kollegen in der Abteilung für Phoniatrie und Pädaudiologie für dieses besondere Arbeitsumfeld danken.

Für die Gelegenheit in einem so tollen Team zu promovieren, möchte ich mich ganz besonders bei meinem Doktorvater und Betreuer Prof. Michael Döllinger bedanken. Er hat mir über all die Jahre immer das Gefühl gegeben, dass er an mich glaubt und mir oft mehr zutraut als ich mir selbst. Für seine Unterstützung meiner wissenschaftlichen Laufbahn durch stetige Forderung und Förderung bin ich ihm sehr dankbar.

Ein herzliches Dankeschön gilt natürlich meinen Kollegen und Co-Doktoranden aus der Forschung: für Diskussionen und Ratschläge, für Korrekturlesen und Mitdenken, bei Papern, Anträgen und Vorträgen. Vielen Dank für Eure Hilfe und die gute Zusammenarbeit.

Ich bedanke mich ganz herzlich bei Sandra, Veronika und Anke für die produktiven Gespräche und die gegenseitige Unterstützung auf fachlicher und emotionaler Ebene bei unserem gemeinsamen Projekt „Frauen in der Wissenschaft“. Ein besonderer Dank für fachlichen Austausch, persönliches Coaching und Beruhigung meiner Nerven geht an Stefan, die zweite Hälfte des philosophischen Duos in der Phoniatrie. Vielen Dank auch an meine zwei Hiwis, Jonas und Bernhard, für die gute Zusammenarbeit und die zuverlässige Zuarbeit in den verschiedenen Projekten.

Ein besonderer Dank geht an meine Familie und Freunde, von denen viele so mutig und nett waren, ihre Zeit und Stimmlippen für unsere Testmessungen zur Verfügung zu stellen. Ihr habt mir in schwierigen Zeiten Rückhalt, Kraft und Motivation zum Weitermachen gegeben. Ein ganz liebes Dankeschön geht natürlich auch an Beke für extrem gründliches und gleichzeitig schnelles Korrekturlesen meiner Arbeit.

Des Weiteren bedanke ich mich beim Bayerischen Laserzentrum für die erfolgreiche Zusammenarbeit in der Entwicklung des Prototypen für die Laserprojektionseinheit und bei der Deutschen Forschungsgemeinschaft (DO1247/6-1) und der Deutschen Krebshilfe (111332), die diese Arbeit finanziell gefördert haben.

Abstract

Verbal communication enables a direct and efficient exchange of information, in which the voice and articulated speech become key elements of human interaction. A lasting impairment of the voice is always accompanied by severe social and even economic disadvantages. While the primary and secondary economic sectors decrease constantly, the communication-based service sector holds the larger economic volume nowadays. Accordingly, a thorough understanding and adequate examination techniques of the laryngeal dynamics as a primary source of the acoustic signal are not only of medical and social interest but also of economic value.

Conventionally, laryngologic examinations of three-dimensional (3D) structures and movements in the larynx are based on the subjective assessment of two-dimensional (2D) imaging. Therefore, objective quantitative 3D parameters are highly desirable to reflect the complex 3D anatomy and physiology and to allow a comprehensive evaluation in the clinical routine.

Within the scope of this thesis, knowledge and methods from the fields of engineering, computer science, optics and medicine were combined. By this, for the first time, a 3D reconstruction of the complete superior vocal fold surface was successfully obtained from *in vivo* high-speed recordings during phonation. The innovations originating from this thesis cover to the following areas: the hardware of the measuring device, the software of the analytical tool and the acquisition of unknown medical data.

The custom-developed miniaturized measuring device is an endoscopic camera-laser set-up, which is based on the concept of stereo triangulation and suitable for *in vivo* application. The laser projection unit generates a regular laser grid that is projected on the vocal folds and recorded by a high-speed camera. The experimental set-up is developed with regard to future clinical application striving to ensure safety and tolerance for the patients. The design of the laser projection unit represents a reconciliation of the opposing requirements: It is a compromise between laser safety regulations for tissue exposure and sufficient visibility of the projected pattern in the *in vivo* recording situation. It is also a trade-off between the miniaturization of the measuring device for best possible patient's tolerance and the resulting vertical resolution.

With regard to the clinical application, a semi-automatic analytical software tool with a graphic user interface was developed in order to calibrate the measuring device and facilitate the detection of the projected laser points. Within the evaluation process, the vocal fold edge is modelled by means of cubic Bézier curves and a G2-continuous transition to the reconstructed coordinates, in order to recreate the natural behavior of mucosal tissue with a continuous curvature. The vocal fold characterization for clinical

purposes demands a condensation of the recorded 3D surface data to diagnostically conclusive parameters. Two deviating approaches are presented and the parameters are tested in a preliminary study on healthy subjects.

The devised measuring instrument with an endoscopic width of 13.5 mm turned out to be tolerable for the majority of all test subjects. Furthermore, the laser projection unit with exposure values (max. 9.7 kW/m^2 for 0.5 sec) below the internationally accepted limit (11 kW/m^2 for 1 sec) was approved by the local ethics commission. The feasibility of the experimental set-up and the reconstruction routines were successfully demonstrated in the recording and analysis of a healthy test group. In this process, an evaluation is carried out of the accuracy of the presented methods, the decrease in the number of indispensable user interactions and the reduction of the total processing duration. Even if not statistically confirmed, the results of the preliminary study demonstrate and emphasize the available knowledge provided by 3D investigation of the larynx dynamics. It can be concluded that healthy phonation is not necessarily accompanied by symmetric oscillation behavior in all dimensions. Enabled by comparable absolute measures, the obvious discrepancy between the amplitudes on the medial and the superior surface demonstrates the variability of 3D parameters depending on the investigated area.

Despite pending improvements like the increase in accuracy and acceleration of the analytical process, I am convinced of the tremendous potential of endoscopic laser-based 3D imaging for evidence-based medicine in laryngology. A comprehensive examination of the anatomy and dynamics in phonation should incorporate all available degrees of freedom, i.e. a full 3D model of the vocal folds. This fresh perspective of the phonatory process will enable new insights on the functional cause and effect chain in healthy and pathological phonation. Additionally, an objective quantification of the oscillation dynamics has a clear advantage over a subjective evaluation. This new possibility to identify different severity levels supports the diagnostic process as well as a quantification of the therapeutics success.

Zusammenfassung

Die verbale Kommunikation erlaubt einen direkten und effizienten Austausch von Informationen, wodurch unsere Stimme und Sprache eine Schlüsselrolle in der zwischenmenschlichen Interaktion einnehmen. Eine dauerhafte Schädigung der Stimme stellt immer auch eine soziale und sogar wirtschaftliche Beeinträchtigung dar. Während der erste und zweite Wirtschaftssektor stetig zurückgehen, erwirtschaftet der dritte, auf Kommunikation basierende Sektors, heutzutage immer größere Umsätze. Dementsprechend sind ein gründliches Verständnis und passende Untersuchungsmethoden für den Kehlkopf als primäre Akustikquelle nicht nur von medizinischem und sozialem Interesse, sondern auch von wirtschaftlicher Bedeutung.

Herkömmlicherweise basieren laryngologische Untersuchungen der dreidimensionalen (3D) Strukturen und Bewegungen im Kehlkopf auf einer subjektiven Bewertung von zweidimensionaler (2D) Bildgebung. Um aber die komplexe 3D Anatomie und Physiologie angemessen abzubilden und eine umfassende, klinische Beurteilung zu ermöglichen, benötigt man objektive und quantitative 3D Parameter.

Im Rahmen dieser Doktorarbeit wurden Kenntnisse und Methoden aus verschiedenen Fachgebieten wie Technik, Informatik, Optik und Medizin angewendet. Dadurch konnte die endoskopisch einsehbare Stimmlippenoberfläche zum ersten Mal aus *in vivo* Hochgeschwindigkeitsaufnahmen von der menschlichen Phonation vollständig in drei Dimensionen rekonstruiert werden. Die entwickelten Neuerungen dieser Arbeit erstrecken sich auf die Hardware des Messaufbaus, die Software der Auswertung und die Erhebung medizinischer Daten.

Die Spezialanfertigung für das miniaturisierte Messgerät ist ein endoskopischer Kamera-Laser-Aufbau auf dem Prinzip der Stereotriangulation, der jetzt für die *in vivo* Anwendung geeignet ist. Die Laser-Projektionseinheit erzeugt ein reguläres Lasergitter, das auf die Stimmlippen projiziert wird und dabei von einer Hochgeschwindigkeitskamera aufgenommen wird. Der experimentelle Aufbau wurde im Hinblick auf den zukünftigen klinischen Einsatz entwickelt, wobei die Sicherheit und Toleranz des Patienten gewährleistet werden müssen. Das Design der Projektionseinheit vereint die gegensätzlichen Anforderungen aus Medizin und Technik: Es stellt einen Kompromiss zwischen den Laserschutzrichtlinien für die Bestrahlung der Haut und einer ausreichenden Sichtbarkeit des projizierten Musters in der *in vivo* Aufnahme dar. Des Weiteren musste zwischen der Miniaturisierung des Messgeräts für die bestmögliche Toleranz des Patienten und der resultierenden vertikalen Auflösung abgewogen werden.

Im Hinblick auf die klinische Anwendung wurde eine teilautomatisierte Analyse-Software mit einer graphischen Benutzeroberfläche entwickelt um die Berechnung

der Kalibrierung und die Detektion der Laserpunkte zu vereinfachen. Innerhalb der Auswertung wurde die Stimmlippenkante mit kubischen Bézierkurven und einem G2-stetigen Übergang zu den rekonstruierten Koordinaten modelliert um das natürliche Verhalten der Schleimhaut mit stetiger Krümmung nachzubilden.

Eine Charakterisierung der Stimmlippen für den klinischen Einsatz erfordert eine Kondensierung der 3D-Oberflächeninformation zu diagnostisch aussagekräftigen Parametern. Es werden zwei unterschiedliche Ansätze vorgestellt und die Parameter in einer vorläufigen Studie an gesunden Probanden getestet.

Das entwickelte Messinstrument mit einer Breite von 13.5mm an der endoskopischen Spitze war für den Großteil aller Probanden gut verträglich. Die Laser-Projektionseinheit liegt mit ihren Bestrahlungswerten von maximal 9.7 kW/m^2 für weniger als 0.5Sek. unter dem international akzeptierten Limit (11 kW/m^2 für 1Sek.) und wurde von der zuständigen Ethikkommission genehmigt. Die Anwendbarkeit des experimentellen Aufbaus und der Rekonstruktion wurden durch die Aufnahme und Analyse einer gesunden Testgruppe erfolgreich gezeigt. In diesem Zuge wurde die Genauigkeit der Methode, die Verringerung der Anzahl erforderlicher Nutzer-Interaktionen und die Verkürzung der kompletten Auswertungsdauer bewertet. Obwohl die Daten keine statistische Aussagekraft haben, untermauert die Studie trotzdem die Möglichkeit neue Erkenntnisse durch die 3D Untersuchung der Stimmlippendynamik zu erhalten. So kann geschlussfolgert werden, dass eine gesunde Phonation nicht zwingend mit symmetrischem Schwingungsverhalten in allen Dimensionen verknüpft ist. Durch die Berechnung von vergleichbaren Messgrößen in Absolutwerten kann man einen offensichtlichen Unterschied zwischen den Amplituden der medialen und superioren Oberfläche erkennen, der die Variabilität von 3D Parametern in Abhängigkeit der untersuchten Fläche belegt.

Trotz ausstehender Verbesserungen wie einer Erhöhung der Genauigkeit und einer Beschleunigung der Auswertung bin ich vom außerordentlichen Potential der endoskopischen laserbasierten 3D Bildgebung für eine evidenzbasierte Medizin in der Laryngologie überzeugt. Eine umfassende Untersuchung der Anatomie und Dynamik während der Stimmgebung sollte alle zugänglichen Freiheitsgrade, wie z.B. eine vollständige 3D Darstellung der Stimmlippen, miteinbeziehen. Diese neue Perspektive auf den Phonationsprozess wird neue Erkenntnisse über das funktionale Zusammenspiel in gesunder und pathologischer Stimmgebung ermöglichen. Zudem ist eine objektive Quantifizierung der Schwingungsdynamik im Vergleich zu einer subjektiven Einschätzung klar von Vorteil. Die so mögliche Identifikation verschiedener Schweregrade erleichtert die Diagnosestellung ebenso wie die Quantifizierung von Therapieerfolgen.

Contents

1	Introduction	1
2	Experimental set-up for laser-based stereo triangulation	9
2.1	Concept of stereo triangulation in camera-laser set-up	9
2.2	Requirements for <i>in vivo</i> laser projection	11
2.3	Laser safety regulations in laryngoscopy	12
2.4	Realization of miniaturized measuring device	14
2.4.1	First generation	14
2.4.2	Second generation	15
2.5	Challenges and potential of the experimental set-up	20
3	Image processing for 3D reconstruction of the vocal fold surface	25
3.1	Overview of 3D reconstruction process	25
3.2	Calibration process	28
3.2.1	Automation of calibration algorithm	28
3.2.2	Evaluation of automated calibration algorithm	30
3.3	Laser point detection	34
3.3.1	Automation of initial grid detection and tracking	34
3.3.2	Evaluation of automated initial grid detection and tracking	38
3.4	Validation of 3D reconstruction	42
3.5	G2-continuous modelling of the vocal fold edge	44
3.6	Challenges and potential of the reconstruction process	48
4	Dynamic 3D parameters for vocal fold characterization	49
4.1	Deviating definitions of 3D parameters	49
4.2	Preliminary study on dynamic 3D parameters	53
4.2.1	Hypothesis and objective of the study	53
4.2.2	Study participants	54
4.2.3	Results	54
4.3	Challenges and potential of 3D parameters	62

5 Conclusion	65
5.1 Status quo of the project	65
5.2 Future perspectives of the project	67
5.2.1 Technological advances	67
5.2.2 Preclinical trials	68
5.2.3 Versatility of technology	69
5.3 Impact of the project	70
Bibliography	71
List of Figures	83
List of Tables	90