Akustische Tubenendoskopie mit Methoden der digitalen Sprachverarbeitung

Christiane Antweiler¹, Peter Vary¹, and Ercole Di Martino² ¹ Institut für Nachrichtengeräte und Datenverarbeitung (ind), RWTH Aachen, 52056 Aachen {antweiler, vary}@ind.rwth-aachen.de ² Klinik für Hals-, Nasen- und Ohrenkrankheiten und Plastische Gesichtschirurgie, Ev. Diakonie Krankenhaus, 28239 Bremen, E.DiMartino@diako-bremen.de

Zusammenfassung

In diesem Beitrag wird ein neues Verfahren zur Funktionsprüfung der Eustachischen Röhre vorgestellt. Basierend auf einer akustischen Messung der Übertragungsstrecke Nase-Ohr wird mit Techniken der digitalen Signalverarbeitung ein virtuelles Abbild der Eustachischen Röhre berechnet. Durch die Animation der virtuellen Abbilder über der Zeit lässt sich der dynamische Prozess der Tubenfunktion visualisieren. Dies ermöglicht erstmals eine Art "akustische Tubenendoskopie" der Eustachischen Röhre unter physiologischen Bedingungen, wobei dieses Verfahren sowohl bei intaktem als auch bei defektem Trommelfell eingesetzt werden kann. Im Rahmen einer interdisziplinären Kooperation wurde ein Echtzeit-Prototyp entwickelt, der anhand von neu gewonnenen diagnoseund therapierelevanten Merkmalen völlig neue Möglichkeiten in der otologischen Medizin eröffnet.

1. Einleitung

In der HNO-Medizin ist die Funktionalität der Eustachischen Röhre für die Planung des Eingriffs am Mittelohr von entscheidender Bedeutung.

Im Vergleich zu den in der klinischen Routine bislang überwiegend eingesetzten Methoden (SSTV-Test, Tympanometrie) stellt die Sonotubometrie eine viel versprechende Alternative dar, da sie eine nicht-invasive und rasche Evaluation der Tubenfunktion unter physiologischen Bedingungen ermöglicht. Die Sonotubometrie basiert auf einem Messsystem, bei dem ein 8 kHz Sinuston mittels Lautsprecher in der Nase abgestrahlt und mit einem im Ohr befindlichen Mikrofon synchron aufgenommen wird. Durch den Anstieg der Signalamplitude im gemessenen Mikrofonsignal läßt sich die Funktionalität des Tubenöffnungsvorgangs nachvollziehen und in diesem Rahmen analysieren und klassifizieren (s. Bild 1).

Ein Nachteil dieses Verfahrens ist jedoch seine schlechte Ergebnisqualität, so dass derzeit kein Untersuchungsverfahren alle diagnostischen Erfordernisse erfüllt (z.B. [1]-[3]).

In Abschnitt 2 wird ein Messverfahren vorgestellt, bei dem die Nase-Ohr Strecke mit sogenannten *perfekten Sequenzen* (PSEQ) [4], [5] angeregt und eine Schätzung der Übertragungsstrecke vorgenommen wird. Eine neu eingeführte Messgröße ersetzt und verbessert die bisherigen Ergebnisse der Sonotubometrie entscheidend. Eine weiterführende Idee wird in Abschnitt 3 behandelt. Kern ist



Bild 1: Beispiele zur Sonotubometrie

ITG-Fachtagung Sprachkommunikation 2006

die Entwicklung eines in Anlehnung an die Sprachsignalverarbeitung berechneten Tubenmodells. Das Prinzip der Modellierung des Sprechtraktes [6], [7] mittels unterschiedlicher Röhrenabschnitte und digitalem Ersatzschaltbild wird in umgekehrter Richtung auf die Nase-Ohr Strecke übertragen. Diese neue Technik ermöglicht eine zuverlässige Detektion der Tubenöffnung und die Visualisierung dynamischer Abläufe.

2. Echtzeit-Messsystem mit PSEQ

Bei dem neu entwickelten Messverfahren wird die Eustachische Röhre im technischen Sinne als lineares Übertragungssystem betrachtet und eine Messung der Übertragungsstrecke Nase-Ohr gemäß Bild 2 mit perfekten Sequenzen (PSEQ) [4], [5], d.h. periodisch wiederholten Rauschsignalen, durchgeführt. Die Abtastfrequenz beträgt $f_A = 32$ kHz. Eine anschließende Verarbeitungsstufe führt mit Hilfe des Normalized Least Mean Square (NLMS) Algorithmus eine Systemidentifikation mit hoher zeitlicher Auflösung durch [8]. Auf Grund der speziellen Korrelationseigenschaften der perfekten Sequenz liefert der Adaptionsalgorithmus im Gegensatz zur herkömmlichen Sonotubometrie in jedem Abtasttakt nicht nur einen Amplitudenwert, sondern eine Schätzung der Übertragungsstrecke Nase-Ohr in Form einer Impulsantwort.



Bild 2: Echtzeit-Messsystem mit PSEQ



Bild 3: Norm $\Theta(k)$ für vier Manöver

Mit der quadratischen Norm der Impulsantwort $\mathbf{w}(k)$ gemäß

$$\Theta(k) = ||\mathbf{w}(k)||^2 = \sum_{i=0}^{N-1} w_i^2(k) \qquad (1)$$

steht ein zur Sonotubometrie vergleichbarer Kurvenverlauf zur Verfügung (s. Bild 3). Die Norm $\Theta(k)$ ermöglicht durch Ausnutzung der Information aller Frequenzanteile bis 16 kHz die Darstellung der Tubenöffnung in deutlich verbesserter Qualität. Dies lässt sich unmittelbar für die medizinische Diagnostik nutzen [9]. Bislang wurde bei der Diagnostik vorwiegend berücksichtigt, mit welcher Energie (Lautstärke) ein akustisches Signal von der Nase zum Ohr übertragen wird. Die Norm $\Theta(k)$ beschreibt mit der jeweiligen Energie der Impulsantwort eine im Kern ähnliche Information. Im folgenden Abschnitt wird gezeigt, wie sich aus dem vorgestellten Messverfahren weitere von der Schallenergie unabhängige Merkmale extrahieren lassen.

3. Virtuelles Tubenmodell

Grundlegende Idee ist die Entwicklung eines in Anlehnung an die Sprachsignalverarbeitung berechneten Tubenmodells. Aus akustischen und physikalischen Betrachtungen kann für den Prozess der Spracherzeugung unmittelbar ein digitales Modell abgeleitet werden [6], [7]. Es besteht im Wesentlichen aus den beiden Komponenten Anregung (Luftstrom) und Sprechtraktfilter, welche direkt an der Signalerzeugung und Signalformung beteiligt sind. Die Tatsache, dass den meisten aktuellen Sprachcodieralgorithmen und Standards dieses digitale Modell der Spracherzeugung zugrunde liegt, verdeutlicht die besondere Relevanz dieses Ansatzes. Das digitale Modell lässt sich aus dem sog. Röhrenmodell herleiten. Zu diesem Zweck wird der physikalische Sprechtrakt aus mehreren zylindrischen, verlustlosen Röhrenabschnitten mit unterschiedlichen Querschnitten approximiert. Das zugehörige Flussdiagramm enthält ausschließlich Addierer, Multiplizierer und Verzögerungselemente, und kann geradewegs in ein zeit-diskretes Modell umgerechnet werden.

Das Prinzip der Modellierung des Sprechtraktes mittels unterschiedlicher Röhrenabschnitte und digitalem Ersatzschaltbild wird nun in umgekehrter Richtung auf die Übertragungsstrecke Nase-Ohr übertragen. In einem Schlüsselexperiment wurde in Analogie zum Röhrenmodell des Sprechtraktes ein einfaches Modell für die Übertragungsstrecke Nase-Ohr hergeleitet, in dem ausgehend von der am Probanden gemessenen Impulsantwort die zeitlich veränderlichen Querschnitte eines virtuellen Tubenmodells bestimmt wurden. Das Prinzip ist in Bild 4 wiedergegeben.

Ausgehend von der zu einem Zeitpunkt bestimmten Impulsantwort w werden die zugehörigen Autokorrelationskoeffizienten φ_{ww} bestimmt und das Normalgleichungssystem aufgestellt. Zur Berechnung der Reflektionskoef-



fizienten K kann z.B. der Levinson-Durbin Algorithmus [10] eingesetzt werden. Anschließend werden diese iterativ gemäß

$$A_i = A_{i-1} \frac{1 - K_i}{1 + K_i} \quad i = 2, \dots n$$
 (2)

in Röhrenquerschnittsflächen A umgerechnet. Die Parameter des Modells müssen an die spezielle Aufnahmesituation der Nase-Ohr Messstrecke angepasst werden. Dabei sind insbesondere der Prädiktionsgrad n, der Abschluss-Reflektionskoeffizient am Ohr γ und der Normalisierungsfaktor A_1 für die Form des Modells verantwortlich. Bei entsprechender Wahl dieser Parameter kann bereits mit diesem einfachen virtuellen Modell eine Korrelation zur menschlichen Anatomie erzielt werden.

4. Ergebnisse

Der zeitlichen Verlauf der Norm $\Theta(k)$ aus Bild 5-a visualisiert den durch dreimaliges Gähnen provozierten Öffnungs- und Schließvorgang der Eustachischen Röhre. Die Bilder 5-b,c zeigen exemplarisch für zwei Zeitpunkte die virtuellen Tubenmodelle für einen geschlossenen bzw. offenen Zustand der Tube. Der Vergleich der Bilder verdeutlicht den Unterschied der beiden Tubenmodelle in Bezug auf Größe und Form bei geöffneter und geschlossener Tube.

Um eine Beurteilung vornehmen zu können, sind in Bild 6 das virtuelle Modell einer of-



Bild 5: Ergebnisse einer Probanden-Messung

ITG-Fachtagung Sprachkommunikation 2006



Bild 6: Korrelation von Anatomie und Modell

fenen Tube (Bild 5-c) sowie der approximierte anatomische und in Röhrensegmente umgerechnete Verlauf einer menschlichen Nase-Ohr Strecke gegenübergestellt. Der Vergleich zeigt, dass bereits mit dem ersten einfachen virtuellen Modell eine gute Korrelation zur Anatomie erzielt werden kann. Ziel zukünftiger Forschung ist den Grad dieser Korrelation weiter zu erhöhen.

Durch die Animation der Querschnitte über der Zeit lässt sich der dynamische Öffnungsund Schließvorgang der Eustachischen Röhre visualisieren, was einer Art "akustischer Tubenendoskopie" entspricht.

Von besonderem Interesse ist, dass das virtuelle Tubenmodell von der Schallenergie unabhängig ist und ausschließlich auf der Analyse der spektralen Zusammensetzung der Übertragungsfunktion basiert. Somit wird eine komplementäre Informationsquelle erschlossen, was die Extraktion bisher nicht darstellbarer, diagnoserelevanter Merkmale ermöglicht. Beispielweise ist mit diesem Verfahren bei einem Patienten der Nachweis einer stets geöffneten Eustachischen Röhre gelungen, was allein durch die Auswertung eines Amplitudenoder Norm-Verlaufs nicht gelingt.

5. Zusammenfassung

Im Rahmen einer interdisziplinären Kooperation ist ein neues akustisches Diagnosesystem zur Abbildung der Tubenfunktion entstanden. Der besondere Vorteil dieses Verfahrens ist die Auswertung der zeitlich veränderlichen Impulsantworten, die die unbekannte Übertragungsstrecke Nase-Ohr weitgehend charakterisieren. Während in der Sonotubometrie lediglich betrachtet wird, mit welcher Intensität ein Sinuston (*eine* Frequenz) von der Nase zum Ohr übertragen wird, findet im neuen Verfahren das Übertragungsverhalten bei *allen* Frequenzen bis 16 kHz Berücksichtigung. Die zusätzlich gewonnene Information wird zur Berechnung eines virtuellen Abbilds der Eustachischen Röhre genutzt.

Die neue Methode ermöglicht eine zuverlässige Detektion von Tubenöffnungen und die Visualisierung der dynamische Abläufe unter Berücksichtigung der anatomischen Verhältnisse. Auf diese Weise können neue in die Tubenfunktionsdynamik Einblicke unter physiologischen Bedingungen gewonnen werden.

References

- [1] I. Honjo, "Evaluation of static and dynamic functions of the eustachian tube", in *The Eustachian tube in middle ear diseases*, Springer, Tokyo, pp. 25-38, 1988.
- [2] T. P. McBride, C. Dekray, M. Cunningham M., W. Doyle, "Evaluation of noninvasive Eustachian tube function tests in normal adults", *Laryngoscope*, Vol. 98, pp. 655-658, 1988.
- [3] E. Di Martino, R. Thaden, G. A. Krombach, M. Westhofen, "Eustachian tube function tests. Current knowledge", *HNO*, Vol. 52, pp. 1029-1040, 2004.
- [4] H. D. Lüke, *Korrelationssignale*, Springer, Berlin, 1992.
- [5] H. D. Lüke, H. D. Schotten, "Odd-perfect, almost binary correlation sequences", *IEEE Trans. Aero*space a. Electron. Syst. AES-31, 1995.
- [6] P. Vary, U. Heute, W. Hess, *Digitale Sprachsignalverarbeitung* B. G. Teubner, Stuttgart, 1998
- [7] T. Parsons, *Voice and Speech Processing*. McGraw-Hill, New-York, 1987.
- [8] C. Antweiler, M. Antweiler, "System Identification with Perfect Sequences Based on the NLMS Algorithm", AEÜ, Vol. 49, No. 3, pp. 129-134, 1995.
- [9] C. Antweiler, E. Di Martino, A. Telle, "Akustische Messverfahren zur Funktionsprüfung der Tuba Eustachii mittels perfekter Sequenzen", 38. DGBMT Jahrestagung, BMT, Ilmenau, Vol. 2, pp. 898-899, 2004.
- [10] J. Makhoul, "Linear Prediction: A Tutorial Review", *IEEE Proceedings*, vol. 63, April, pp. 561-580, 1975.