

# TUBENFUNKTIONSPRÜFUNG MIT METHODEN DER DIGITALEN SPRACHSIGNALVERARBEITUNG

*Aulis Telle und Peter Vary*

*Institut für Nachrichtengeräte und Datenverarbeitung, RWTH Aachen University  
{telle,vary}@ind.rwth-aachen.de*

**Kurzfassung:** In der HNO-Medizin kommt der Eustachischen Röhre, die eine Verbindung zwischen Mittelohr und Nasenrachen darstellt, eine besondere Bedeutung für den Verlauf und die Genese akuter und chronischer Mittelohrentzündungen zu. Eine Beurteilung der Funktionsfähigkeit dieser Röhre ist sowohl für die Diagnose von Erkrankungen im Bereich des Nasenrachens und des Ohres als auch für die Planung operativer Maßnahmen von besonderer klinischer Relevanz. Wie in früheren Beiträgen bereits gezeigt, kann aus der akustischen Übertragungsfunktion der Strecke von der Nase zum Ohr ein virtuelles dynamisches Tubenmodell berechnet werden, welches für die Tubenfunktionsprüfung eingesetzt werden kann. Die Abbildungseigenschaften des virtuellen Modells bezüglich der tatsächlichen Geometrie der Übertragungstrecke sind jedoch noch unzureichend. In diesem Beitrag wird betrachtet, welchen Einfluss die Bandpasscharakteristik des verwendeten Lautsprechers auf die Rekonstruktion der Röhrenquerschnitte hat und wie die hervorgerufenen Rekonstruktionsfehler kompensiert werden können.

## 1 Einleitung

Ein vielversprechendes Verfahren zur Funktionsprüfung der Eustachischen Röhre ist die sog. Sonotubometrie [5]. Die Grundidee der Sonotubometrie besteht darin, die Tubenfunktion anhand der Änderung des Übertragungsverhaltens der akustischen Strecke von der Nase zum Ohr über der Zeit auszuwerten. Dafür wird ein Lautsprecher in der Nase und ein Mikrofon im Ohr platziert. Während ein Anregungssignal über den Lautsprecher abgegeben wird, führt der Patient Aktionen wie Schlucken oder Gähnen aus, die in der Regel mit einer Öffnung der Eustachischen Röhre einhergehen. Zur Beurteilung der Tubenfunktion wird die Variation der Leistung des im Ohr aufgenommenen Mikrofonsignals ausgewertet. Bei einem signifikanten Anstieg der Leistung wird von einer Öffnung der Eustachischen Röhre ausgegangen. In dieser Form konnte sich die Sonotubometrie jedoch noch nicht auf breiter Basis durchsetzen, da die Ursachen des Schalleistungsanstiegs im Ohr nicht eindeutig geklärt sind. Eine zuverlässige Interpretation der Kurven ist deshalb nicht immer möglich.

Ein grundlegend neuer Ansatz in der Sonotubometrie ist die Berechnung eines virtuellen, dynamischen Tubenmodells mit Methoden der digitalen Sprachsignalverarbeitung [1, 2, 3]. Hierbei wird nicht die Variation der Schalleistung im Ohr sondern die Variation der spektralen Veränderungen des Übertragungssystems von der Nase zum Ohr analysiert. In ersten Studien konnte bereits gezeigt werden, dass ein solches virtuelles Tubenmodell das Potential hat, die Sonotubometrie um ein zusätzliches, diagnose-relevantes Maß zu erweitern.

In diesem Beitrag wird zunächst das virtuelle, dynamische Tubenmodell kurz vorgestellt. Anschließend wird anhand von Simulationsergebnissen gezeigt, welchen Einfluss eine Bandpassfilterung des Anregungssignals bzw. der ermittelten Impulsantwort auf die Rekonstruktionsergebnisse hat. Dabei werden der Hochpass- und der Tiefpassanteil einzeln betrachtet. Schließlich werden Maßnahmen vorgestellt, mit denen die Genauigkeit der Rekonstruktion erheblich verbessert werden kann.

## 2 Virtuelles Tubenmodell

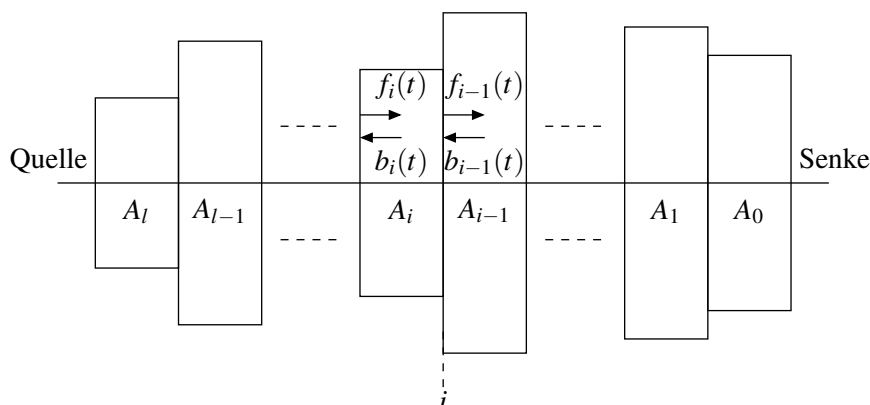
Der Querschnittsverlauf einer gesunden Eustachischen Röhre verändert sich während der Durchführung von Aktionen wie Schlucken oder Gähnen in der Art, dass ein Druckausgleich zwischen Mittelohr und Rachenraum stattfinden kann. In dieser Phase ändern sich auch die akustischen Übertragungseigenschaften dieser Strecke signifikant. Für die Tubenfunktionsprüfung wird bisher nur die Änderung der Schallleistung am Ohr genutzt, die sich durch Öffnen der Eustachischen Röhre in der Regel deutlich vergrößert. Die spektralen Veränderungen des Übertragungssystems bleiben dabei jedoch unberücksichtigt.

Wie in [2], [3] und [1] vorgeschlagen, können die Änderungen in den Spektralkomponenten durch die Berechnung eines virtuellen Tubenmodells ausgenutzt werden. Es ist abgeleitet vom sog. *Röhrenmodell* aus der Sprachverarbeitung, auf dem viele der aktuellen und standardisierten Algorithmen der Sprachcodierung basieren.

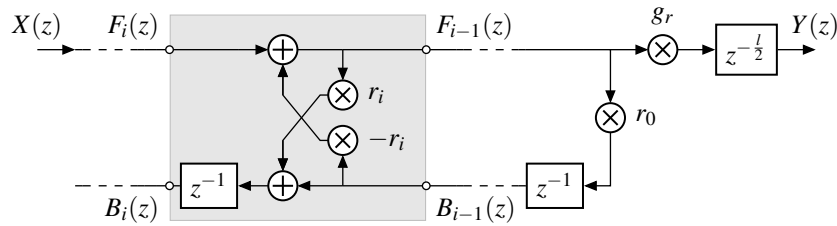
Im Bereich der Sprachcodierung wird durch das Röhrenmodell der Prozess der Spracherzeugung modelliert. Aus akustischen und physikalischen Betrachtungen des Vokaltraktes heraus kann unmittelbar ein digitales Modell abgeleitet werden [4, 8, 9]. Die beiden wesentlichen Komponenten dieses Modells sind eine Anregung, die den Luftstrom durch die Glottis nachbildet, und das Vokaltraktfilter, welches die Signalformung durch den Rachenraum berücksichtigt.

Das digitale Vokaltraktfilter lässt sich aus einem idealisierten Röhrenmodell herleiten und als autoregressiver (AR) Prozess darstellen. Zu diesem Zweck wird die Anatomie des Vokaltraktes durch eine Röhre mit kreisförmigem, abschnittsweise konstantem Querschnitt, schallharten Wänden und homogenem Ausbreitungsmedium approximiert [7] (siehe Abbildung 1). Daraus kann über mehrere Zwischenschritte ein Signalflussdiagramm in Form eines Lattice-Filters abgeleitet werden [9], wie es in Abbildung 2 gezeigt ist. Dieses Modell kann zur Analyse eines vorliegenden (Sprach-)Signals genutzt werden, indem aus dem Zeitsignal abschnittsweise die zeitveränderlichen Modellparameter geschätzt werden.

Übertragen auf die Tubenfunktionsprüfung bedeutet dies, dass die Glottis als Schallquelle durch einen Lautsprecher in der Nase ersetzt wird. Die Senke ist statt der Mundöffnung das Mikrophon im äußeren Gehörgang. Im Gegensatz zur Sprachverarbeitung ist die Anregung exakt bekannt, so dass eine Impulsantwort des Übertragungssystems bestimmt werden kann, die als Ausgangssignal für die Bestimmung der Modellparameter dient. Aus diesen kann anschließend der dynamische Querschnittsverlauf der Übertragungstrecke rekonstruiert werden. Auf Basis der Veränderung des Querschnittsverlaufes über der Zeit können dann ggf. Aussagen über die Aktivität der Eustachischen Röhre während der Messung getroffen werden.



**Abbildung 1** - Idealisiertes Röhrenmodell nach Kelly und Lochbaum [7].  $f_i(t)$  und  $b_i(t)$  bezeichnen jeweils die vor- bzw. zurücklaufende Schallwelle am Anfang des Röhrenabschnittes  $i$ , der einen Querschnitt  $A_i$  hat. Die Übergangsstelle zwischen den Segmenten  $i$  und  $i - 1$  wird mit dem Index  $i$  gekennzeichnet. Die Gesamtzahl der Röhrensegmente ist  $l + 1$ .



**Abbildung 2** - Digitales Ersatzschaltbild für das Röhrenmodell in Lattice-Struktur-Darstellung [9]. Die hinterlegte Struktur wird entsprechend der Anzahl der Röhrensegmente wiederholt.

Eine wichtige Eigenschaft der Abbildung der Impulsantwort bzw. deren Autokorrelationsfunktion auf einen Querschnittsverlauf ist, dass diese Abbildung nach [6] eineindeutig ist, wenn die Ordnung des der Impulsantwort zugrunde liegenden AR-Prozesses kleiner oder gleich der Ordnung des rekonstruierten Filters ist. Im Folgenden wird angenommen, dass diese Bedingung erfüllt ist.

### 3 Einfluss einer Bandpassfilterung auf die Rekonstruktionsergebnisse

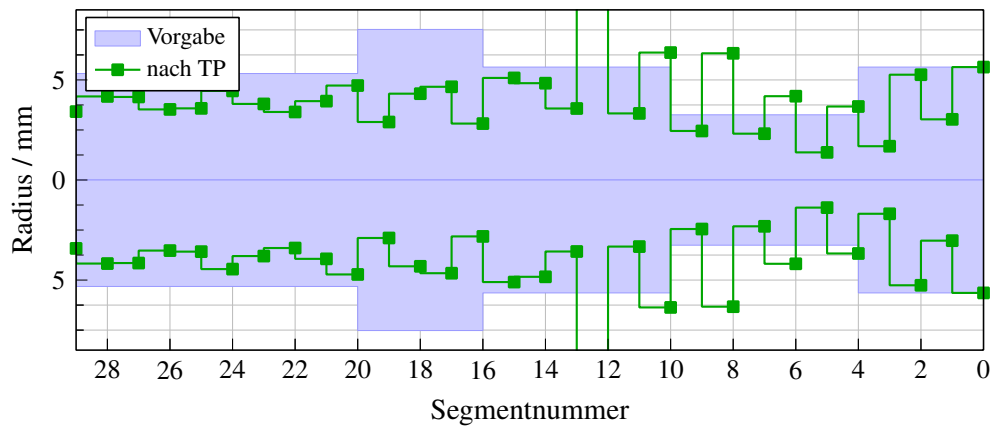
Bei der Messung der akustischen Eigenschaften der Übertragungsstrecke Nase-Ohr wird ein dynamischer Miniaturlautsprecher (Ohrhörer) eingesetzt. Die Übertragungsfunktion eines solchen Lautsprechers ist spektral sowohl zu tiefen wie auch zu hohen Frequenzen hin deutlich begrenzt. Wegen massiver tieffrequenter Störungen z.B. während des Schluckaktes wird zu deren Unterdrückung zusätzlich ein Hochpass mit einer Grenzfrequenz von etwa 1 kHz eingesetzt.

Die Folge ist, dass die ermittelten Impulsantwort unabhängig von den tatsächlichen Übertragungseigenschaften einen Bandpasscharakter aufweisen. Das Fehlen der tiefen und hohen Komponenten führt zwangsläufig zu Fehlern in der Rekonstruktion des Querschnittsverlaufes, wenn nicht die tatsächliche Impulsantwort der betrachteten Röhre zufällig ebenfalls einen entsprechenden Bandpasscharakter aufweist. Der Einfluss der Tiefpass- und Hochpassanteile des Bandpasses wird im Folgenden separat betrachtet.

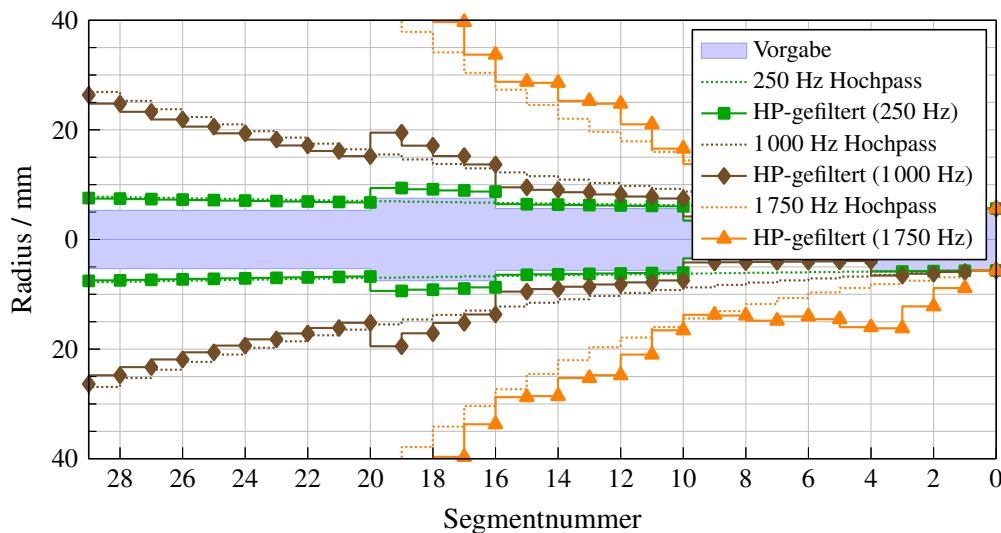
Durch das Fehlen der hochfrequenten Anteile der Impulsantwort der betrachteten Röhre wird bei der Rekonstruktion des Querschnittsverlaufes implizit ein Tiefpassverhalten der Röhre angenommen. Dies führt dazu, dass der berechnete Querschnittsverlauf das typische Merkmal eines akustischen Tiefpasses aufweist, nämlich zwischen großen und kleinen Querschnitten alternierende Röhrenabschnitte.

In Abbildung 3 ist dieser Effekt am Beispiel einer willkürlich vorgegebenen und durch ein Lattice-Filter mit entsprechenden Filterkoeffizienten simulierten Röhre veranschaulicht. Die Kurven zeigen einen Längsschnitt durch die vorgegebene bzw. rekonstruierte Röhre. Die Schallquelle ist auf der linken Seite, die Senke auf der rechten Seite. Die Nummerierung der Röhrensegmente erfolgt von rechts nach links. Die Kurve mit ausgefüllter Fläche zeigt den Querschnittsverlauf der vorgegebenen Röhre an. Für die Unterdrückung der hochfrequenten Anteile in der Impulsantwort wurde ein linearphasiges FIR-Tiefpassfilter (Least-Squares Design) der Länge  $N = 1000$  verwendet, dessen Grenzfrequenz bei  $f_g = 12000$  Hz lag. Die Abtastrate beträgt in allen Beispielen  $f_a = 48000$  Hz. Der Querschnittsverlauf der aus der tiefpassgefilterten Impulsantwort gewonnenen Röhre zeigt insbesondere am Anfang (rechts) den erwarteten alternierenden Verlauf.

Zur Betrachtung des Hochpasseinflusses wurde die Impulsantwort vor der Rekonstruktion des Querschnittsverlaufes mit einem linearphasigen FIR-Hochpassfilter (Least-Squares Design) der Länge  $N = 1500$  gefiltert. Für die in Abbildung 4 gezeigten Beispiele wurden drei Hochpässe mit den Grenzfrequenzen  $f_{g1} = 250$  Hz,  $f_{g2} = 1000$  Hz und  $f_{g3} = 1750$  Hz verwendet. Zusätzlich zu den rekonstruierten Querschnittsverläufen sind die zu den Impulsantworten der Hochpassfilter korrespondierenden Querschnittsverläufe aufgetragen.



**Abbildung 3** - Vorgegebene und nach Tiefpassfilterung ( $f_g = 12 \text{ kHz}$ ) aus der Impulsantwort rekonstruierte Röhrenquerschnitte.



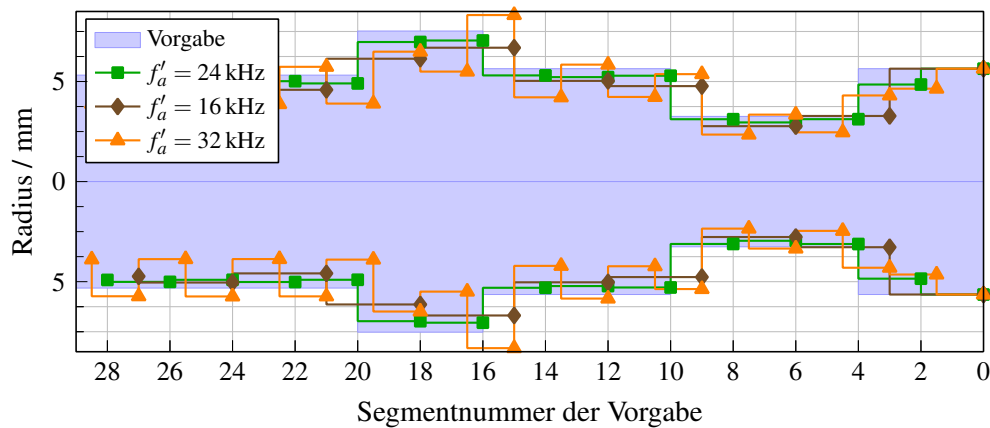
**Abbildung 4** - Vorgegebene und nach Hochpassfilterung aus der Impulsantwort rekonstruierte Röhrenquerschnitte. Zum Vergleich sind die zu den Impulsantworten der Hochpässe korrespondierenden Querschnittsverläufe aufgetragen (gepunktete Linien).

Wie die Beispiele zeigen, weisen die rekonstruierten Röhren eine ausgeprägte Exponentialtrichterform auf, die für einen akustischen Hochpass typisch ist. Bei der Rekonstruktion wird also wiederum implizit ein Hochpassverhalten der betrachteten Röhre angenommen. Es fällt auf, dass die Querschnittsverläufe der rekonstruierten Röhren grob dem Querschnittsverlauf des jeweils verwendeten Hochpasses folgen.

## 4 Kompensation des Tiefpasseinflusses

Wie bereits angedeutet, liegt das Problem der Rekonstruktion eines Querschnittsverlaufes aus einer tiefpassgefilterten Impulsantwort darin, dass die vorhandene Tiefpasseigenschaft implizit der zu rekonstruierenden Röhre zugeschrieben wird. Für eine Kompensation des Tiefpasseinflusses gilt es, dem entgegenzuwirken.

Dadurch, dass aufgrund des Tiefpasscharakters bei hohen Frequenzen keine signifikanten Komponenten in der Impulsantwort enthalten sind, kann ohne Verlust an Informationen eine kritische Unterabtastung der Impulsantwort durchgeführt werden.



**Abbildung 5** - Vorgabe und nach Tiefpassfilterung ( $f_g = 12 \text{ kHz}$ ) und Unterabtastung der Impulsantwort rekonstruierte Röhrenquerschnitte. Die Unterabtastung erfolgte zu den Abtastraten  $f'_a = 24 \text{ kHz}$ ,  $f'_a = 16 \text{ kHz}$  und  $f'_a = 32 \text{ kHz}$

Eine solche Unterabtastung hat zwei Auswirkungen. Zum einen verringert sich durch die Reduktion der Abtastrate die höchste darstellbare Frequenz der Impulsantwort. Entsprechend sinkt allerdings auch die größte im digitalen Röhrenmodell anregbare Frequenz. Das Fehlen der hohen Frequenzkomponenten in der ursprünglichen, hoch abgetasteten Impulsantwort kann sich somit nach Reduktion der Abtastrate nicht mehr auf das Röhrenmodell auswirken, da dieser Frequenzbereich nicht darstellbar ist.

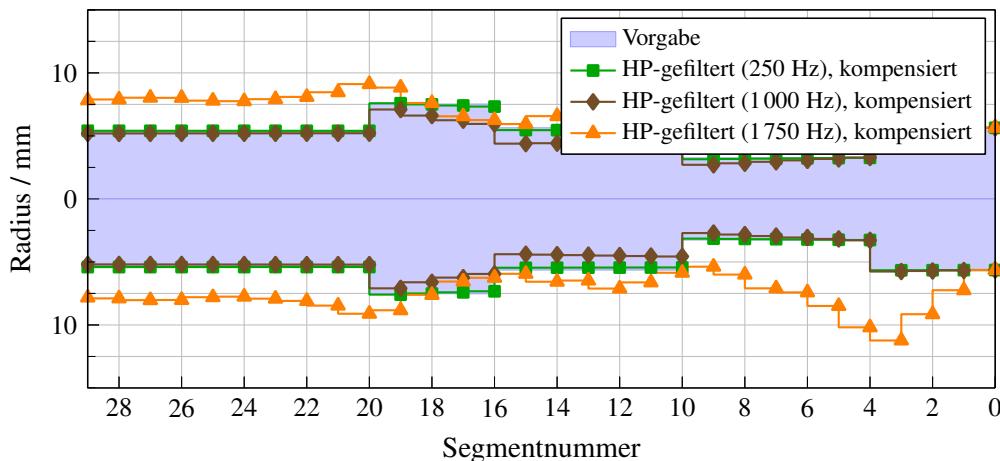
Zum anderen ist die Länge eines Röhrenabschnitts im digitalen Modell umgekehrt proportional zur Abtastrate, d.h. die Verringerung der Abtastrate führt zu einer größeren Länge der einzelnen Segmente. Folglich verringert sich die örtliche Auflösung des digitalen Modells, was bei der Darstellung und dem Vergleich der Ergebnisse zu berücksichtigen ist.

Abbildung 5 zeigt das Ergebnis der Kompensation durch Unterabtastung für obiges Beispiel. Dafür wurde die Impulsantwort nach entsprechender Antialiasing-Filterung von  $f_a = 48 \text{ kHz}$  auf  $f'_a = 24 \text{ kHz}$  (kritische Unterabtastung),  $f'_a = 16 \text{ kHz}$  (zu starke Unterabtastung) und  $f'_a = 32 \text{ kHz}$  (zu schwache Unterabtastung) unterabgetastet und anschließend das korrespondierende Röhrenmodell berechnet. Man erkennt deutlich, dass nach der Unterabtastung der Impulsantwort die Abweichung zwischen dem rekonstruierten Querschnittsverlauf von der Vorgabe erheblich geringer ist als ohne Unterabtastung. Im Fall der zu schwachen Unterabtastung ist außerdem gut zu sehen, dass weiterhin eine alternierende Struktur erhalten bleibt, da nicht stark genug unterabgetastet wurde und somit ein Tiefpassanteil erhalten geblieben ist. Eine zu starke Unterabtastung ist weniger kritisch.

## 5 Kompensation des Hochpasseinflusses

Wie in Abbildung 4 gezeigt, weisen sowohl die Querschnittsverläufe der Hochpässe als auch die der hochpassgefilterten Impulsantworten wie erwartet die Form eines Exponentialtrichters auf. Die Querschnittsverläufe zu den hochpassgefilterten Impulsantworten folgen dabei weitgehend dem Querschnittsverlauf des entsprechenden Hochpassfilters. Ein direkter mathematischer Zusammenhang konnte zwar nicht nachgewiesen werden, es hat sich aber gezeigt, dass der ursprüngliche Querschnittsverlauf gut approximiert werden kann, wenn die zur gefilterten Impulsantwort korrespondierenden Querschnitte durch die des entsprechenden Hochpasses dividiert werden.

Wie Ergebnisse in Abbildung 6 zeigen, sind für Grenzfrequenzen bis  $1 \text{ kHz}$  die Ergebnisse dieser Methode sehr nah am Originalverlauf. Je höher die Grenzfrequenz des Hochpasses jedoch wird, desto schlechter wird die Approximation. Bei einer Grenzfrequenz von  $1750 \text{ kHz}$  sind die Abweichungen z.B. schon so deutlich, dass der vorgegebene Verlauf kaum zu erkennen ist.



**Abbildung 6** - Vorgabe und Querschnittsverlauf nach Kompensation des Hochpasseinflusses durch Division der rekonstruierten Querschnitte durch die Querschnitte des entsprechenden Hochpasses.

## 6 Zusammenfassung

In diesem Beitrag wurden die Auswirkungen einer Bandpasscharakteristik des verwendeten Messsystems auf die Rekonstruktion von Röhrenquerschnitten aus einer ermittelten Impulsantwort betrachtet. Es hat sich gezeigt, dass der Tiefpassanteil dieser Bandpasscharakteristik zu einem für einen akustischen Tiefpass typischen alternierenden Querschnittsverlauf führt. Der Hochpassanteil wiederum hat zur Folge, dass dem eigentlichen Querschnittsverlauf eine Exponentialtrichterform überlagert wird.

Beide Phänomene können teilweise kompensiert werden. Der Ausbildung eines alternierenden Querschnittsverlaufes kann durch kritische Unterabtastung entgegengewirkt werden. Der durch Hochpassfilterung der Impulsantwort entstehende Exponentialtrichter kann bis zu einer Grenzfrequenz von etwa 1000 Hz durch Division der Querschnittswerte durch die zur Impulsantwort des jeweiligen Hochpassfilters korrespondierenden Querschnittswerte erheblich abgemildert werden.

Darüber hinaus kann festgestellt werden, dass allein durch eine Erhöhung der Abtastrate des Messsystems trotz der dann feineren Ortsauflösung des digitalen Röhrenmodells keine Verbesserung der Rekonstruktionsgenauigkeit erzielt werden kann. Dies wird erst durch eine zusätzliche entsprechende Erhöhung der oberen Grenzfrequenz der verwendeten akustischen Komponenten (insb. Lautsprecher und Mikrofon) möglich.

## Literatur

- [1] Christiane Antweiler, Aulis Telle, Peter Vary und Ercole Di Martino. "A New Otological Diagnostic System Providing a Virtual Tube Model". In: *Proceedings of Biomedical Circuits and Systems Conference BIOCAS*. London, Great Britain, Nov. 2006.
- [2] Christiane Antweiler, Peter Vary und Ercole Di Martino. "Akustische Tubenendoskopie mit Methoden der digitalen Sprachverarbeitung". In: *ITG-Fachtagung Sprachkommunikation*. 2006.
- [3] Christiane Antweiler, Peter Vary und Ercole Di Martino. "Virtual Time-Variant Model of the Eustachian Tube". In: *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS)*. Island of Kos, Greece, Mai 2006, S. 5559–5562.
- [4] G. Fant. *Acoustic theory of speech production*. 2. Aufl. The Hague: Mouton, 1970. ISBN: 90-279-1600-4.

- [5] Niels van Heerbeek, Stijn J. C. van der Avoort, Gerhard A. Zielhuis und Cor W. R. J. Cremers. “Sotubometry: A Useful Tool to Measure Intra-individual Changes in Eustachian Tube Ventilatory Function”. In: *Arch Otolaryngol Head Neck Surg* 133.8 (Aug. 2007), S. 763–766.
- [6] Steven M. Kay. *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ 07632: Prentice Hall, 1988. ISBN: 0-13-598582-X.
- [7] J. L. Kelly und C. C. Lochbaum. “Speech synthesis”. In: *Proc. Fourth Int. Congress on Acoustics*. Copenhagen, Dänemark, 1962.
- [8] Gerold Ungeheuer. *Elemente einer akustischen Theorie der Vokalartikulation*. Berlin: Springer-Verlag, 1962.
- [9] Peter Vary und Rainer Martin. *Digital Speech Transmission: Enhancement, Coding and Error Concealment*. Chichester, UK: John Wiley & Sons, Ltd., Jan. 2006. ISBN: 978-0-471-56018-9.