ISCA Archive
http://www.isca-speech.org/archive

Models and Analysis of Vocal Emissions for
Biomedical Applications (MAVEBA)
2nd International Workshop
Florence, Italy
September 13-15, 2001

Detecting Asymmetries in High Speed Observations of Vocal Folds during Phonation

Michael Döllinger Ulrich Hoppe Stefan Schuberth
Frank Hettlich* Ulrich Eysholdt

Dept. of Phoniatrics and Pediatric Audiology, University of Erlangen-Nuremberg,Bohlenplatz 21, D-91054 Erlangen, Germany, Tel: ++49-9131-8532603, Fax: ++49-9131-8539272,
email: boert-rdm@gmx.de
*Institute of Mathematics, Englerstr. 2, D-76128 Karlsruhe, Germany

1 Abstract

In this paper we describe an approach for extracting asymmetry proportions of vocal folds which affect vibrations. Combining high speed observations with the two-mass-model by Ishizaka & Flanagan (1972) as modified by Steinecke & Herzel (1995) an inversion algorithm is presented which allows the identification of physiological parameters of the vocal folds. The problem can be regarded as an optimization problem with a non-convex objective function. Therefore, the choice of appropriate initial values is important. It can be automized by using spectral features of the mass trajectories. The applicability of the inversion procedure to high speed observations of vocal fold vibrations is demonstrated in a case of functional dysphonia and in a healthy voice.

Keywords: *high speed glottography, two mass model, inversion*

2 Introduction

The two mass model as introduced by Ishizaka & Flanagan [1] and simplified by Steinecke & Herzel [2] describes the most important features of the vocal fold motion [2, 3, 4]. It has been successfully applied to explain high speed observations of vocal fold vibrations. In a recent work [5] several parameters of the model as mass, vocal fold stiffness and subglottal pressure have been estimated by visual comparison of experimental observations with model predictions. Until now there is a lack of an automatic algorithm which allows the identification of the model parameters from the high speed recordings without user interaction.

The aim of the present work is to develop an automatic algorithm which allows identification of the degree and amount of vocal fold asymmetry from high speed observations.

3 Two-Mass-Model

The Two-Mass-Model (2MM) consists of a pair of two coupled oscillators representing each one vocal fold. Masses are set into vibrations by the Bernoulli-Force. Hence, the model consists of two parts describing the myoelastic and the aerodynamic properties. The model is scetched as a frontal section in Fig. 1. According to the simplifications by Herzel & Steinecke [4], the dynamics of the system can be described by a system of eight differential equations (1).

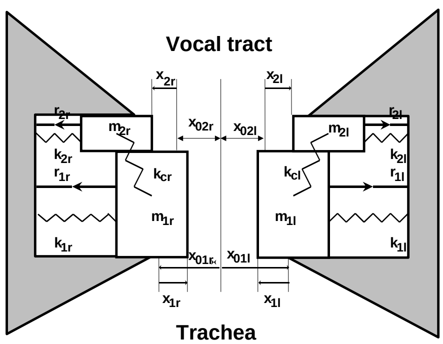

Figure 1: *Schematic representation of the two-mass model of the vocal folds.*

$x_{i\alpha}$ denote oscillation amplitudes measured as the distance from the glottal midline (see Fig. 1) and $v_{i\alpha}$ are the corresponding velocities. The indices (i, α) represent lower $(i = 1)$ and upper $(i = 2)$ masses and left $(\alpha = l)$ and right $(\alpha = r)$ masses. The matrix comprises tissue properties of the vocal folds, i.e. masses $m_{i\alpha}$, stiffness coefficients $k_{i\alpha}$ and damping coefficients $r_{i\alpha}$. Nonlinearities of the elastic forces and interactions with the vocal tract have been neglected. Hence, the only nonlinearities result from the driving Bernoulli force F_1 (2) which is assumed to act only on the lower masses, and the impact forces $I_{1\alpha}$ and $I_{2\alpha}$. These impact forces (3) act as additional restoring forces when the vocal folds get contact. $\Theta(x)$ describes the Heaviside function and P_s the subglottal pressure. The parameters $c_{i\alpha}$ represent additional stiffness coefficients describing the influence of the impact of the left and right masses. l denotes the length of the glottis and a_i denotes the instantaneous glottal area between the lower $(i = 1)$ and upper $(i = 2)$ masses, respectively. The presented model describes the most important vibratory patterns of vocal fold oscillations including mucosal waves when using the standard parameters presented by Ishizaka & Flanagan [1]. Asymmetry can be imposed by expressing the parameters $m_{i\alpha}, k_{i\alpha}, \ldots$ in terms of the standard parameters $m_{i\alpha 0}, k_{i\alpha 0} \ldots$ using the asymmetry parameters Q_r and Q_l which are defined as follows [6]:

$$
\begin{aligned}
k_{i\alpha} &= Q_\alpha k_{i\alpha 0}\,, \\
k_{c\alpha} &= Q_\alpha k_{c\alpha 0}\,, \\
c_{i\alpha} &= Q_\alpha c_{i\alpha 0}\,, \\
m_{i\alpha} &= m_{i\alpha}/Q_{\alpha 0}.
\end{aligned}
$$

Using this representation, the vibratory pattern of the vocal fold oscillations can be described by a set of three parameters: Q_l, Q_r, and the subglottal pressure P_s.

$$\frac{d}{dt}\begin{pmatrix} x_{1\alpha} \\ v_{1\alpha} \\ x_{2\alpha} \\ v_{2\alpha} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\left(\frac{k_{1\alpha}+k_{c\alpha}}{m_{1\alpha}}\right) & -\frac{r_{1\alpha}}{m_{1\alpha}} & \frac{k_{c\alpha}}{m_{1\alpha}} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_{c\alpha}}{m_{2\alpha}} & 0 & -\left(\frac{k_{2\alpha}+k_{c\alpha}}{m_{2\alpha}}\right) & -\frac{r_{2\alpha}}{m_{2\alpha}} \end{pmatrix}\begin{pmatrix} x_{1\alpha} \\ v_{1\alpha} \\ x_{2\alpha} \\ v_{2\alpha} \end{pmatrix} + \begin{pmatrix} 0 \\ I_{1\alpha}(x_{1r}, x_{2r}) + \frac{1}{m_{1\alpha}}F_1(x_{1r}, x_{2r}, x_{1l}, x_{2l}) \\ 0 \\ I_{2\alpha}(x_{2r}, x_{2l}) \end{pmatrix} . \tag{1}$$

$$F_i = l d_i P_1 \quad \text{with} \quad P_1 = P_s\left[1 - \Theta(a_{min})\left(\frac{a_{min}}{a_1}\right)^2\right]\Theta(a_1). \tag{2}$$

$$I_{i\alpha}(x_{i\alpha}) = -\Theta(-a_i)\frac{c_{i\alpha}}{m_{i\alpha}}\frac{a_i}{2l} \quad \text{with} \quad \Theta(x) = \begin{cases} 1 & , \quad x > 0 \\ 0 & , \quad x \leq 0 \end{cases} . \tag{3}$$

4 The Inverse Problem

One calls two problems inverse to each other if the formulation of one problem involves the other one. For mostly historic reasons, one might call one of these problems (usually the simpler one) the *direct problem*, the other one the *inverse problem* [7].
For our application, the direct problem is defined by solving the system of differential equations (1) and the inverse problem means the finding of corresponding parameters representing the pathophysiology from experimental time series (HGG sequences). While the 2MM provides trajectories for both upper and lower masses the HGG sequences usually not allow this separation. For comparison of experimental and theoretic curves only the trajectories of the lower masses resulting from the 2MM are used. In the following, the most important steps of the inversion procedure are outlined.

4.1 Objective Function

Our aim is to minimize the error between the trajectories obtained from HGG sequences and theoretical curves generated with the 2MM. Assuming periodic curves it is reasonable to calculate fourier coefficients via discrete fourier transform (DFT) of both, the experimental time series and the theoretical time series. The fourier representation is unique and indicates amplitudes and phases from involved partial vibrations [8]. From the experimental time series, we pick up the local maxima and their direct neighbours which are stressed out the backround noise (Fig. 2 and Fig. 4). The coefficients for the left and right masses constitute the sets I_l, I_r with $3* \mid I_l \mid = K_l$, $3* \mid I_r \mid = K_r \in \mathbb{N}$, respectively. $\mid I_l \mid$ means the cardinality of the

$$
\begin{aligned}
& \| (| c_{1l} |, \ldots, | c_{Kl} |) - (| d_{1l} |, \ldots, | d_{Kl} |) \|_2 \\
+ & \| (\arg c_{1l}, \ldots, \arg c_{Kl}) - (\arg d_{1l}, \ldots, \arg d_{Kl}) \|_2 \\
+ & \| (| c_{1r} |, \ldots, | c_{Kr} |) - (| d_{1r} |, \ldots, | d_{Kr} |) \|_2 \\
+ & \| (\arg c_{1r}, \ldots, \arg c_{Kr}) - (\arg d_{1r}, \ldots, \arg d_{Kr}) \|_2,
\end{aligned} \tag{4}
$$

set I_l. The multiplier three is necessary because of taking the next neighbours of the characteristic coefficients, too. The other fourier coefficients are neglected.

Since representing the curves with fourier coefficients, it is essential to scale the amplitudes so that the absolute and phase values of the fourier coefficients are balanced. Now we can define our objective function (4). The objective function has to be minimized on $(Q_l, Q_r, P_s) \in \mathbb{R}^3_+$. c_{il}, c_{ir} denote the coefficients for the experimental trajectories. d_{il}, d_{ir} represent the fourier coefficients of the theoretical curves generated with th 2MM. For the objective function, fourier coefficients of the oscillations of the 2MM are neglected if they are less then 25 percent of the absolute values of the experimental curves. The reason for doing this, is that the objective function gets a much more better structure: The objective function is constant in uninteresting regions which can be immediately neglected for the optimization procedure.

4.2 Nelder-Mead-Algorithm

Conventional algorithms as the Levenberg-Marquardt-Algorithm [9] have proven inadequate, since the objective function of the problem is non-convex. Due to the flexibility concerning the problem requirements the Nelder-Mead-Algorithm [10] has proven quite more stable. As an advantage, the objective function needs not to be differentiable. Additionally, the NM-Algorithm is rather fast when only a few parameters have to be optimized. On the other hand, the NM-Algorithm does not provide any information of convergence. Briefly, the algorithm works as follows: Four starting vectors are necessary for the optimization. These vectors define a simplex. A new point is defined by moving away from the centroid of the simplex in the direction which is opposite to the point with the largest error. When this new point gives a better solution, the worst of the other solutions is omitted. The algorithm stops, when the extension of the simplex is below a predefined threshold ϵ. A detailed description of the NM-Algorithm can be found in [10].

4.3 Initial Values

Since a non-convex objective function has to be assumed, it is very important to find suitable initial values for Q_l, Q_r, P_s. For a simple mass-spring oscillator the fundamental frequency can be determined by:

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m}}. \tag{5}$$

Applying (5) to our problem, we get

$$f_{0l} = \frac{1}{2\pi}\sqrt{\frac{k_{1l0} * Q_l}{\frac{m_{1l0}}{Q_l}}} = \frac{1}{2\pi} * Q_l \qquad (6)$$

for the left vocal fold and also for the right vocal fold. Although the 2MM is more complicated, equation (5) holds for a large parameter range. By combining all elements within I_l with all elements within I_r one obtains $\mid I_l \mid * \mid I_r \mid$ different sets of initial values for Q_l, Q_r.

For each Q_l, Q_r an area with a rectangular shape is defined. Within this area a search for the initial values is performed along the diagonals and edges, with increment 0.01. During the search, the subglottal pressure P_s is permanently modified in predefined intervals with length 7 $cm\ H_2O$. Within these intervals for the subglottal pressure the algorithm searches parallel to the Q_l, Q_r rectangle for initial values. Four parameter sets which provide the best approximations to the experimental curves constitute the initial values for the NM-Algorithm.

5 Application to High Speed Observations

Now we want to demonstrate the applicability of the inversion procedure on high speed recordings. The patients were examined by endoscopic investigation, stroboscopy, and High Speed Glottography. High speed recordings were performed during 'regular' phonation: The patients were instructed to articulate /i/ in a 'comfortable' way. From the HGG sequence trajectories of the left and right vocal folds were extracted. Fourier coefficients were calculated from these trajectories. The theoretical trajectories as retrieved by the 2MM describe many features of the experimental curves as we will see now:

First we want to describe the application of the inversion algorithm to vocal fold oscillations measured in a female patient (38 years old) suffering from a *functional dysphonia*.

The best approximation (Fig. 3) to the experimental trajectories can be achieved when taking coefficient no. 140 and no. 71 for the left and right side, respectively. These coefficients correspond to fundamental frequencies of 560 Hz for the left vocal fold and 284 Hz for the right vocal fold. The subglottal pressure P_s was varied between 14 $cm\ H_2O$ and 21 $cm\ H_2O$ for searching initial values. The algorithm stops after 40 iterations with $Q_l = 4.5498$ and $Q_r = 2.0953$. This means the masses are the fifth part respectively the half of the standard masses. The spring constants are 4.5 times respectively 2 times larger than the standard spring constants. Since the standard parameters of the 2MM describe vibratory patterns in male subjects both Q_l and Q_r are increased, this means values greater than 1. The subglottal pressure was determined as $P_s = 21.1\ cm\ H_2O$, which corresponds to an increased pressure by a factor of 2.6 compared to the standard pressure. The value of the objective function for the obtained parameter set is 25.9, this means an update of 40 percent compared to the initial values.

The second example is a healthy female logopedia student with normal voice. The corresponding absolute values of the fourier

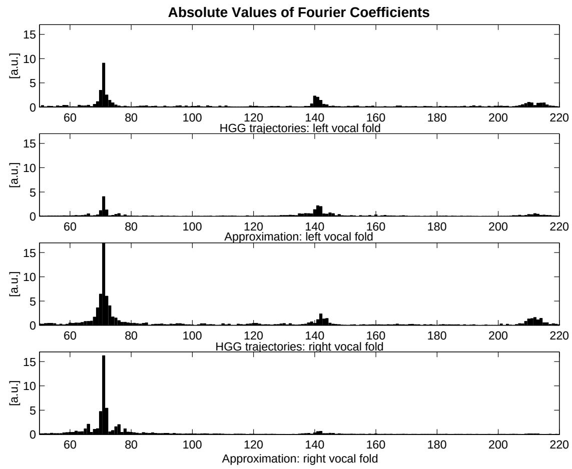

Figure 2: *Absolute values of fourier coefficients for experimental HGG trajectories and their approximation. The charcteristic coefficients are* $I_l = \{71, 140, 214\}$ *and* $I_r = \{71, 142\}$.

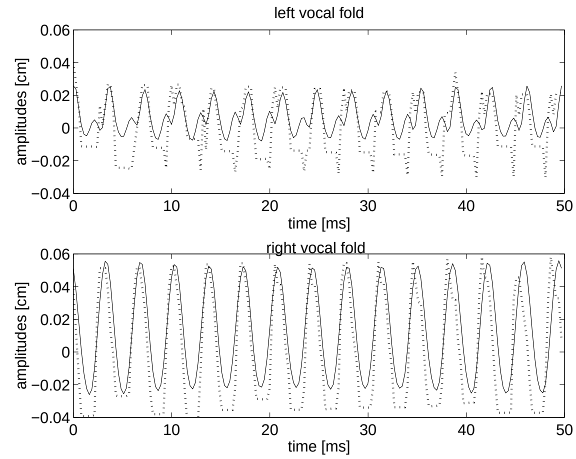

Figure 3: *Result of the approximation during a time slot of* $50\,ms$*: Upper picture shows the left vocal fold trajectory together with its approximation. Within the lower figure the corresponding curves for the right vocal folds are shown. The solid lines represent the result of the algorithm and the dotted lines the experimental HGG trajectories.*

coefficients are shown in Fig. 4. The characteristic coefficients correspond to a fundamental frequency of 352 Hz for the left and right vocal folds. The subglottal pressure P_s was varied between 8 $cm\ H_2O$ and 15 $cm\ H_2O$ for searching initial values.

In this example the algorithm stops after 15 iterations with $Q_l = 2.68$ and $Q_r = 2.64$. The subglottal pressure was detected at $P_s = 12.8\ cm\ H_2O$. The value of the objective function for the obtained parameter set of the NM-Algorithm is 20.6, this means an update of 15 percent compared to the initial values. Because of the female patient, the values for Q_l and Q_r are increased again. The similar values for the asymmetry parameters verify the healthy voice of the female patient. The subglottal pressure is slightly increased by comparison with the standard pressure.

6 Discussion

For perturbed high speed trajectories, it is difficult to find the characteristic fourier coefficients which are the most important properties for the working of the algorithm. If the high speed recordings are of high quality, the optimization procedure provides good results:
For the *functional dysphonia* both the reduced amplitude of the right vocal fold and the slight phase delay of the left vocal fold is reproduced in the theoretical curves (Fig. 3). In this case high speed sequences in combination with the inverse solution of the two-mass-model allow to determine the presence of laryngeal asymmetry, the

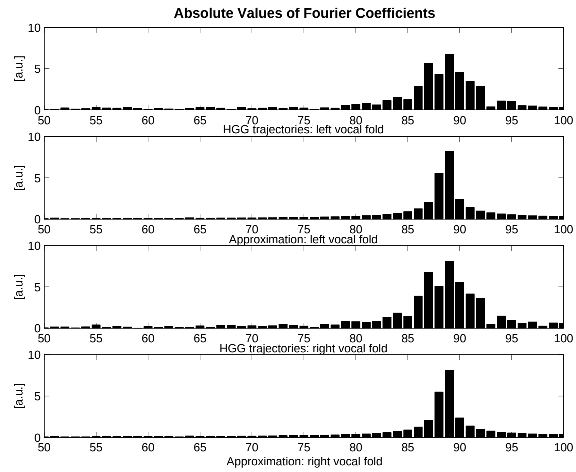

Figure 4: *Absolute values of fourier coefficients for extracted fold trajectories and their approximation. The characteristic coefficient for the left and right vocal fold is no. 89 .*

kind of laryngeal asymmetry i.e. mass or tension, and the degree of asymmetry. Also, the optimization algorithm verified the healthy voice of the logopedia student (Fig. 5).

However, some differences between experimental and theoretical curves have to be mentioned. One important difference is due to the permeation of the masses of the 2MM experimental curves have to be interpolated during glottal closure. Glottal closure can be identified by the horizontal segments of the dotted curves (Fig. 3 and Fig. 5). Other differences occur because of the short-term irregularities in the experimental curves. These time variations may either be caused by real physiological changes or by the limited accuracy of the high speed measurement.

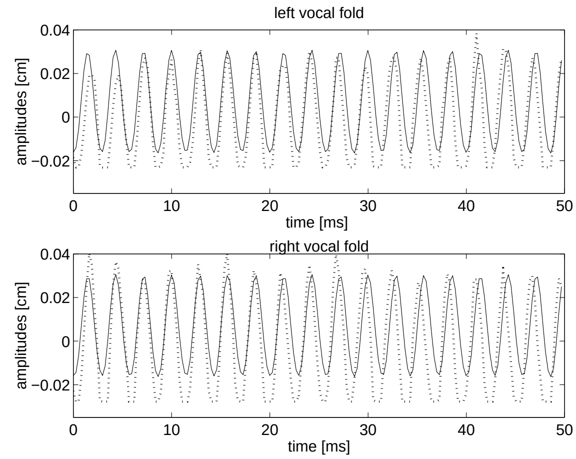

Figure 5: *Result of the approximation during a time slot of* 50 *ms: Upper picture shows the left vocal fold trajectory together with its approximation. Within the lower figure the corresponding curves for the right vocal folds are shown. The solid lines represent the result of the algorithm and the dotted lines the experimental HGG trajectories.*

In view of the algorithm, it is to mention that the finding of good initial values is essential for an acceptable optimization result. If the preprocess does not find good initial values, the algorithm is unable to find a local or even the global minimum of the objective function.

References

[1] K. Ishizaka and J. L. Flanagan, “Synthesis of voiced sounds from a two-mass model of the vocal cords,” *Bell Syst. Techn. J.*, vol. 51, pp. 1233–1268, 1972.

[2] H. Herzel, “Bifurcations and chaos in voice signals,” *Applied Mechanical Revues*, vol. 46, pp. 399–413, 1993.

[3] H. Herzel, D. A. Berry, I. R. Titze, and I. Steinecke, “Nonlinear dynamics of the voice: Signal analysis and biomechanical modeling,” *Chaos*, vol. 5, pp. 30–34, 1995.

[4] I. Steinecke and H. Herzel, “Bifurcations in an asymmetric vocal fold model,” *J. Acoust. Soc. Am.*, vol. 97, pp. 1571–1578, 1995.

[5] P. Mergell, *Nonlinear Dynamics of Phonation - High-Speed Glottography and Biomechanical Modeling of Vocal Fold Oscillations.* Kommunikationsstörungen - Berichte aus Phoniatrie und Pädaudiologie, Aachen: Shaker Verlag, 1998.

[6] I. Steinecke, “Untersuchungen an einem vereinfachten Stimmlippenmodell,” diplomarbeit, Institut für Thoretische Physik, Humboldt-Universität Berlin, 1994.

[7] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems.* P.O. Box 17, 3300 AA Dordrecht, The Netherlands: Kluwer Academic Publisher, 1996.

[8] R. Brigola, *Fourieranalysis, Distributionen und Anwendungen.* Braunschweig/Wiesbaden: Vieweg Verlag, 1997.

[9] J. Moré, “The leven-marquardt alghorithm: Implementation and theory,” in *Lecture Notes in Mathematics 630* (Watson, ed.), Numerical Anal., 1977.

[10] W. Murray, *Numerical Methods for Unconstrained Optimization.* London: Academic Press, 1972.