Numerical Glottal Sound Source Model as Coupled Problem between Vocal Cord Vibration and Glottal Flow

Hideyuki Nomura, Tetsuo Funada
Graduate School of Natural Science and Technology, Kanazawa University
Kakuma-machi, Kanazawa-shi 920-1192, Japan
nomu@t.kanazawa-u.ac.jp

Abstract

Understanding of dynamics of the speech production process may lead to a breakthrough in increasing the quality of synthesized voice and contribute to medical science with respect to treating voice disorder. However, little is known about mechanisms of speech dynamics, in particular about complex interactions of glottal flow and vocal fold vibration. In this paper, a two-dimensional viscoelastic body model of the vocal folds coupled with an unsteady glottal flow is proposed, and we demonstrate the speech production process by using that model. The vocal cord was modeled by distributed mass-spring-damper elements. The speech production process is simulated by numerically solving equations of nonlinear compressible viscous fluid coupled with motion of vocal fold. The results indicated that the upper and lower edges of the vocal fold vibrated with a phase difference. Furthermore, a change in pressure in the larynx synchronized with the vocal fold vibration, and both amplitude and fundamental frequency of the speech wave slightly fluctuated with time. Although we assumed the physical property of left and right part of vocal cords to be symmetry, the each vocal cord vibration indicated some difference. From the obtained results, although there are some problems to be discussed, one may say that the proposed model is valid for the glottal sound source.

1. Introduction

Understanding of mechanisms in the phonation is important, since that may become a breakthrough in improvement of qualities of synthesized voices, and is also expected to contribute to developments of medical engineering against voice disorders.

Some models of glottal sound source have been proposed, and speech production process have been analyzed by using them. Many researchers have used the two-mass model of Ishizaka and Flanagan [1] and improvement models of that [2] for model analysis of speech productions. Their model were simple structures and principal mechanisms of the phonation could be understood using them. However, since a flow in the larynx were assumed to be a linear one-dimensional quasi-steady flow, it is difficult to describe complicated and nonlinear phenomena in a turbulent glottal flow.

Flow measurements [3, 4] and a numerical simulation [5] in a static glottal model have showed unsteady two or three-dimensional complicated flows such as the Coanda effect or turbulence. These results suggest that it is necessary to consider such nonlinear phenomena in the speech production model. However, little is known about a coupled problem of the vocal cord vibration and complicated glottal flow.

The purpose of this study is to propose a glottal sound source model including complicated turbulent flow in the larynx. We assume a vocal cord as distributed elements of mass-spring-damper and glottal flow as two-dimensional compressible viscous fluid. Speech production processes are numerically simulated as a coupled problem between vocal cord model and fluid model. In this report, we show an outline of the proposed model of glottal source and a verification of that model for the phonation.

2. Speech production model

The larynx model in the coronal plane appears in Fig. 1. We will consider not only about the vocal cords as voiced sound source, but also about the false vocal cords. Since the false vocal cords play an important role in the speech production process, other words, glottal jets impinge the false vocal cords, and interaction of the jets and the false vocal cords may create additional sound sources [6]. It is assumed that a lung pressure can be approximated by an air reservoir at the boundary $\Gamma_1$. We assume phenomena in the larynx to be uniformly in the $x$ direction, and we consider a two-dimensional speech production model. In this model, the total pressure $P^*$, the fluid density $\rho^*$, the absolute temperature $T^*$, and the fluid velocity $u^* = (u^*, v^*)$ are calculated in order to analyze the speech production process. Boundaries $\Gamma_3$ and $\Gamma_4$ have been assumed to be non-vibrating rigid bodies in a previous study [5]. In this paper, we model these boundary of vocal cords as viscoelastic bodies including effects of vocal cord vibration.

A modelling of vocal cord as viscoelastic body is referred to the distributed model proposed by Ikeda et al. [7]. The vocal cords can be divided into two tissue layers with different mechanical properties, i.e., the body and cover layers [8].
cover layer is formed from a elastic cover with effective mass of the vocal cord. In order to take into account of mechanical properties of vocal cord, the elastic cover is supported by distributed elements with a nonlinear spring and damper. A schematic of these model is shown in Fig. 2 (a). The elastic cover as the cover layer of vocal cord consists of masses and nonlinear springs in Fig. 2 (b).

These motion can be predicted by solving the motion equation of mass–spring–damper. Details of this motion equation are found in reference [9].

An equation of motion of the vocal cord vibration is applied to the boundaries $\Gamma_1$ and $\Gamma_2$ in Fig. 1. In this study, we do not consider a distribution of vocal cord vibration in the depth direction of the cord, $y$.

In some studies [1, 2, 7], a flow motion in the larynx has been treated as quasi-steady or unsteady one-dimensional fluid. On the other hand, complicated and unsteady phenomena of glottal flow in spatial domains has been reported in experimental [3, 4] and numerical [5] studies. Since, these results suggest that a one-dimensional analysis of flow is not appropriated for the prediction of phonation, we assume the glottal flow to be a two-dimensional compressible viscous fluid, which is described as nonlinear partial differential equations. The fluid analysis is based on boundary fitted coordinates along the glottal shape [10].

The size parameters of initial shape of the larynx [11] is shown in Table 1. Suppose that the air in the larynx is uniform and at rest for $t^* < 0$, and the vocal cord is relaxed with no strain.

A pressure function $P_{\text{in}}^L(t^*) = p^L_r(t^*) + P_{\text{atm}}$ is applied to the boundary $\Gamma_1$, where $t^*$ is time, $P_{\text{atm}}$ is the atmospheric pressure, and $p^L_r(t^*)$ is the lung pressure,

$$p^L_r(t^*) = \begin{cases} \frac{P_{\text{Lo}}^L}{2} \left(1 - \cos \left( \frac{\pi t^*}{t^*_r} \right) \right) & (0 < t^* \leq t^*_r), \\ \frac{P_{\text{Lo}}^L}{2} \left(1 - \cos \left( \frac{\pi t^*}{t^*_r} \right) \right) & (t^* > t^*_r), \end{cases}$$

(1)

where $P_{\text{Lo}}^L$ and $t^*_r$ are the steady-state value of $p^L_r(t)$ and the time required for $p^L_r$ to increase to $P_{\text{Lo}}^L$ from zero, referred to as the rise time, respectively. Furthermore, a non-reflecting characteristic boundary condition [12] is imposed at the outflow boundary $\Gamma_2$ in order to minimize acoustic reflection. In reality, although speech waves within the vocal tract involve forward ($+x$) and backward ($-x$, reflecting) waves, which interfere with each other and cause formants, we neglect such reflecting waves, since our focus in the present study is on the sound source production process in the larynx.

Since the governing equations are nonlinear partial differential equations, solving these equations analytically subject to initial and boundary condition is extremely involved. Instead, we employ a numerical computation method for speech production process with alternately operation of vocal cord vibration analysis and fluid analysis. A finite-difference and Runge-Kutta scheme are applied for the fluid and vocal cord vibration analysis, respectively. The fluid analysis is described as a moving boundary problem caused by vocal cord vibration.

Small elements with spring, damper, and mass which characterize physical properties of vocal cord are located at each differential grid. The Young’s modulus $E$ and the surface density $\rho_s$ mainly determine the physical properties of vocal cord. These physical parameters are weighted as spatial function of $x$, in order to control the distribution of vibration in the larynx. Their typical values are $E = 3.0$ kPa and $\rho_s = 2.4$ kg/m$^2$ at glottis $x = 0$. In this study, left and right parameters of vocal cords are same values, in other words, vocal cords are physically symmetric. Details of analysis are found in reference [9].

### Table 1: Initial size parameters of the larynx (unit in mm).

<table>
<thead>
<tr>
<th>$L_{\text{Nec}}$</th>
<th>$D_x$</th>
<th>$D_y$</th>
<th>$D_z$</th>
<th>$D_1$</th>
<th>$D_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.8</td>
<td>16.4</td>
<td>0.4</td>
<td>10.3</td>
<td>5.1</td>
<td>19.0</td>
</tr>
</tbody>
</table>

3. Results and discussion

Examples of obtained waveforms based on the proposed model are shown in Fig. 3. The lung pressure $P_{\text{Lo}}^L$ and the rise time $t^*_r$ are set to 1000 Pa and 10 ms, respectively. A waveform in (a) is a glottal area $A_{\text{gmin}}$. A shape of vocal cord is continuously changed with coordinate $x$. The plotted vale is a minimum area around the glottis at each time. A volume velocity at glottis $x = 0$ and a pressure difference, $p = P - P_{\text{atm}}$, between the total pressure and the atmospheric pressure at $x^* \approx 40$ mm are shown in (b) and (c), respectively. Each plotted value $A_{\text{gmin}}, U$, and $p$ is nondimensionalised, and the
In Fig. 4. Top and bottom parts indicate the vibrations without vocal cord collision, such as a falsetto. Vocal cord collision, it is effective that the model apply to the intermediate vocal cords. Although this model is not appropriate when vocal cords collide each other and the space vanishes between vocal cords, since the results did not perfectly closed and always have a closing duration in a period. However, obtained results did not perfectly closed and always $A_{gmin} > 0$ in Fig. 3 (a). This imperfect closing is caused by that we use the numerical model which avoid colliding both vocal cords, since it is not able to divide a fluid space in a finite-difference method, when vocal cords collide each other and the space vanishes between vocal cords. Although this model is not appropriate with vocal cords collision, it is effective that the model apply to a phonation without vocal cord collision, such as a falsetto.

The locations of vibrations each vocal cord surface appear in Fig. 4. Top and bottom parts in Fig 4 indicate the vibrations on boundaries $\Gamma_3$ and $\Gamma_4$, respectively. Solid and dotted lines show the location of upper (near vocal tract) and lower (near trachea) portion of the vocal cords. The upper portion vibrates with different phase to the lower portion. However, the maximum vibrational amplitude is about 0.5 mm. This value is slightly less than a vocal cord vibration on ordinary phonation at this lung pressure (1000 Pa). This less vibration is caused by inappropriate elastic parameters for vocal cords.

In this model, we assumed the physical condition of both vocal cords to be symmetric. However, little difference between both vocal cord vibrations are observed in Fig. 4. In symmetric static glottis model without vibration, asymmetric glottal flows have been observed [3, 4, 5]. The asymmetric flow caused the asymmetric vocal cord vibration. Flows in the vibrating vocal cord model also show asymmetric patterns. The physical property of vocal cord is asymmetry in voice disorder, and it is possible to observe asymmetric vocal cord vibration. The result in Fig. 4 suggests that different vibrations between both vocal cords occur for even normal organs.

Fig. 5 shows sequence of the shapes at every 1 ms in a period of vocal cord vibration $t^* = 81 \sim 90$ ms. As we mentioned above, the upper portion vibrates with different phase to the lower portion, and waves at the surface of vocal cord propagate from the trachea to the vocal tract side, that is to say, the vocal cord shape changes from converging glottis to diverging glottis with time similarly observed real vocal cord vibration.

4. Conclusions

We proposed a two-dimensional glottal sound source model. The vocal cord was modeled by distributed mass-spring-damper elements. The speech production process was numerically simulated as a coupled problem between the vibrational elastic body and the two-dimensional glottal flow in the larynx. Obtained result indicated a slightly fluctuation in the envelope of speech wave. Although we assumed the physical property of left and right part of vocal cords to be symmetry, the each vocal cord vibration indicated some difference. From the obtained results, one may say that the proposed model is valid for the glottal sound source, however, the collision of both vocal cords and less vibrational amplitude remain as problems to be discussed further.

5. Acknowledgments

This work was, in part, supported by the Nakatani Electronic Measuring Technology Association of Japan.
6. References


