CHAPTER 9

ANIMATED VOCAL FOLD MODEL TOOLBOX

9.1 INTRODUCTION

There are two types of glottal source models: (1) glottal models with acoustic parameters; and (2) vocal fold vibratory models. Two glottal models are described in Appendix 7: (1) the LF model; and (2) the polynomial model. These models generally describe the dynamic behavior of the glottal volume–velocity waveform, which is often used in speech synthesis. The vocal fold vibratory models depict the vibratory motion of the vocal folds and calculate estimates of data waveforms, such as the projected glottal area, the vocal fold contact area, and an estimate of the electroglottographic waveform. This chapter describes two models of the vocal folds using a graphic user interface and animation that is implemented in a software toolbox in MATLAB. The results are presented in the form of an animated movie of the vibratory motion of the vocal folds. The theory is provided in Appendix 10.

One model is referred to as the two-mass model (Ishizaka and Flanagan, 1972). This model depicts the vocal folds as two masses interconnected with springs and dash pots (dampers) (see Appendix 10). In our implementation of this model, the user can control the pre-phonatory shape (or configuration), the vocal fold tension, and the lung pressure of the vocal folds. Several other parameters can be adjusted by the user, such as the size of the two masses, their thickness, and their maximum excursion.

A second model is called the ribbon model (Titze, 1984). This model presents a more realistic 3-dimensional view of the vocal folds than the two-mass model. The model incorporates a pre-phonatory configuration of the vocal folds, maximum excursion profiles, the length and depth (thickness) of the vocal folds, and time lags in the movement of the vocal folds along their length and depth.

The vocal fold models depict the vibratory motion of the vocal folds in 3 dimensions. The ribbon model also calculates the projected glottal area, an estimate of the vocal fold contact area, and the electroglottographic waveform during vibratory motion. The toolbox creates movies of the vibratory motion that can be viewed by the user.

A digitized sequence of frames of one vibratory cycle of the actual vocal folds along with a movie of this sequence is contained in a folder on the accompanying CDs. See the README file.

9.2 TOOLBOX FOR THE VOCAL FOLD MODELS

The software is to be installed in a subdirectory (e.g., vocal_fold) in MATLAB in a manner similar to that described for the toolboxes introduced in previous chapters. The toolbox contains two software packages: (1) the two-mass model; and (2) the ribbon model. The

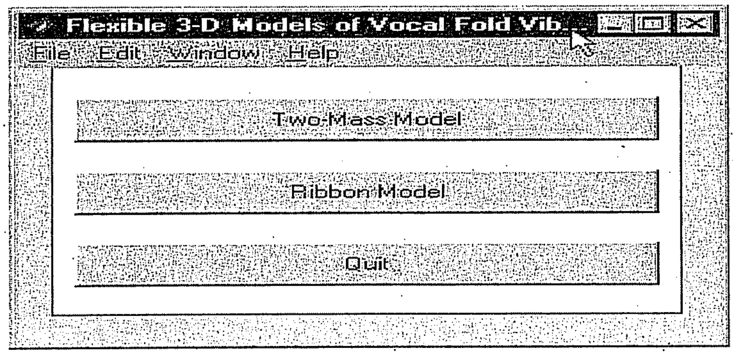

FIGURE 9.1 Main window for the vocal fold models.

two models are independent. To start the software, change directory to the vocal_fold directory and type main in the MATLAB command window. The Main function window shown in Figure 9.1 appears. This window allows the user to select one of the three options: (1) two-mass model; (2) ribbon model; or (3) quit. The Quit button closes all previously opened windows and exits the user from the toolbox.

9.3 TWO-MASS VOCAL FOLD MODEL

The theory for the two-mass vocal fold model is provided in Appendix 10. Here, we present the software toolbox and graphical user interface. To start the two-mass model, press the Two-Mass Model button in Figure 9.1. This opens the two-mass model control window shown in Figure 9.2. From top to bottom, the various parameters that the user can control

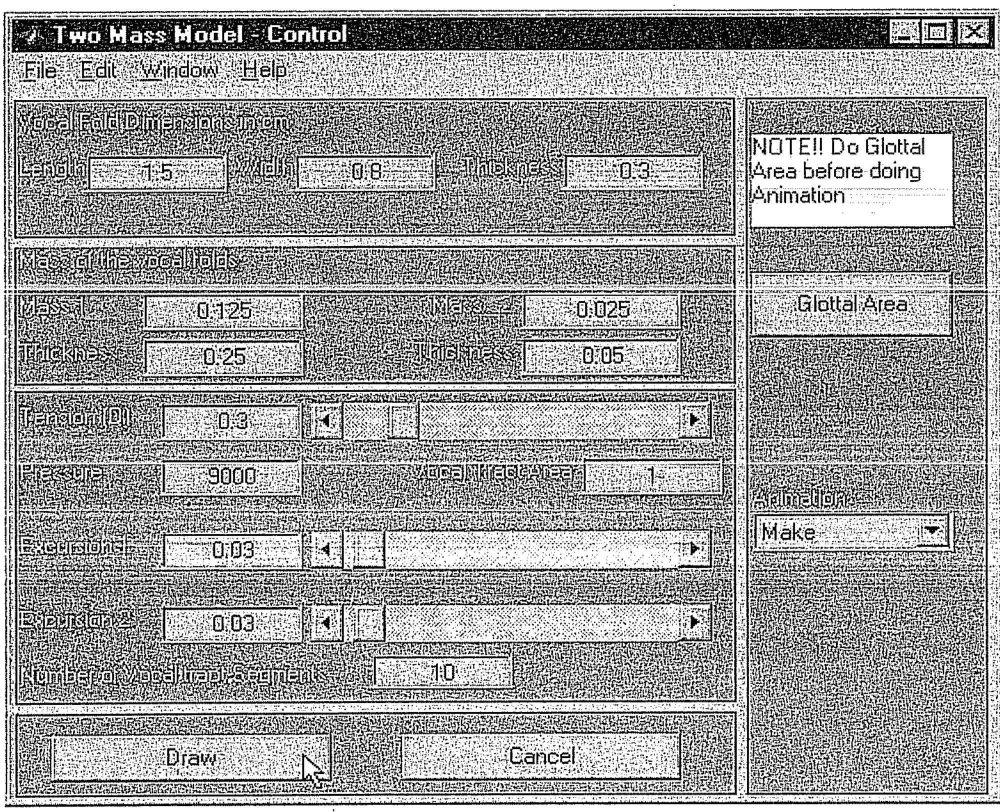

FIGURE 9.2 Two-mass vocal fold model control window.

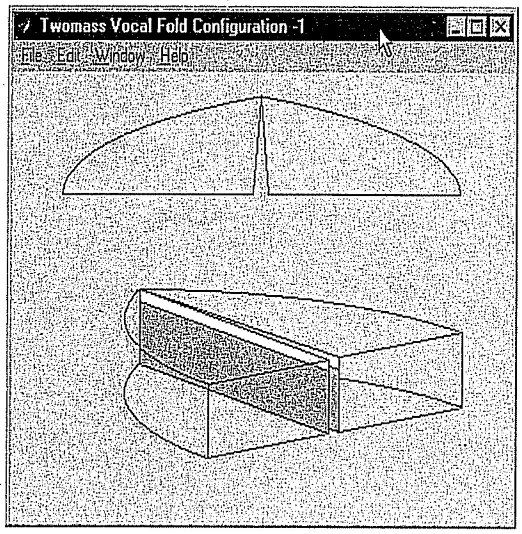

FIGURE 9.3 Static two-mass vocal fold model.

include the vocal fold length, width, and thickness. The length and thickness do not affect the model because of normalization of parameters within the model. However, the width does influence the vibratory motion of the vocal fold model. The size and thickness of the two masses of the model can be adjusted. The total thickness (the sum of the thickness 1 and thickness 2) can exceed 0.3 cm and can in fact, exceed 1 cm. There are no basic restrictions on the length, width, and thickness. However, unrealistic values can give unrealistic vibratory motion results. The tension parameter, Q, is the one suggested by Ishizaka and Flanagan (1972). This parameter is linearly related to the fundamental frequency of vibration of the vocal fold model. The pressure parameter is the lung pressure. The vocal tract area parameter controls the size of the vocal tract area, while the number of vocal tract segments sets the number of vocal tract area segments. The vocal tract area and the number of vocal tract segments present a load on the vocal fold model (see Appendix 10). However, these parameters do not greatly affect the vibratory motion of the vocal folds. The excursion parameters set the maximum excursion of the vocal folds in the pre-phonatory stage. Neither of these values should be zero. The parameter values shown in Figure 9.2 are the default values.

The Cancel button closes the two-mass model control window as well as any windows opened by the model. To start the model, first set the desired parameter values, then press the Draw button. This action opens the windows shown in Figures 9.3 and 9.4, which show the pre-phonatory vocal fold configuration for the two-mass model. Here we have used the default parameter values shown in Figure 9.2. Note that the vocal folds are plotted as though they are rigid ribbons or bands. The upper ribbon represents mass 2 with its corresponding thickness, while the lower ribbon represents mass 1 with its associated thickness. The plot of the vocal folds differs from the type of plot that is often used for the two-mass model. However, the equations for the two-mass model, given in Appendix 10, are used to make the model calculations. We chose this configuration because it is similar to the ribbon model, which is shown next.

Next press the Glottal Area button in Figure 9.2. This action opens the message window shown in Figure 9.5, which informs the user that the glottal area functions are being calculated. This window disappears once the calculations are completed. At the same time, the window shown in Figure 9.6 appears.

The lower graph shows the glottal area for mass 1 (the lower mass) of the two-mass model. The upper graph is the glottal area for mass 2 (the upper mass) of the two-mass

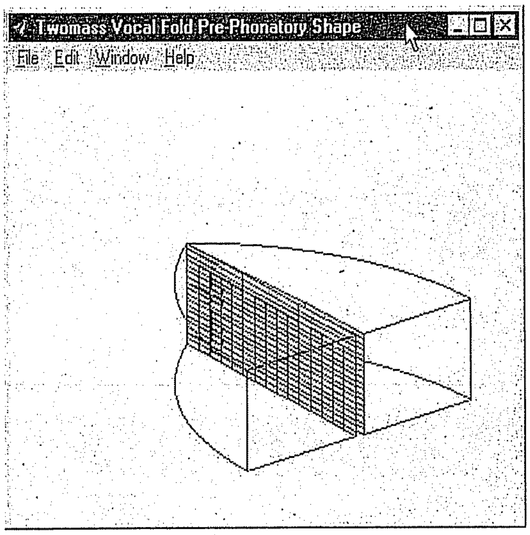

FIGURE 9.4 Pre-phonatory view of two-mass model.

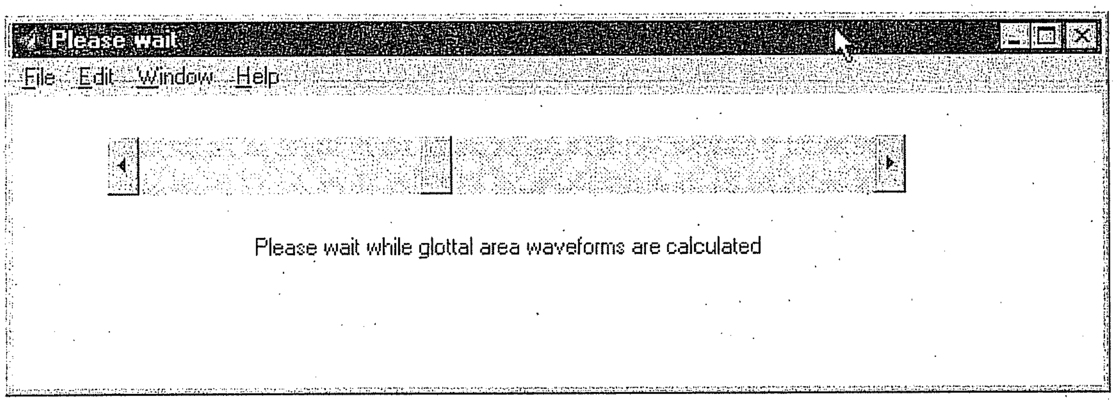

FIGURE 9.5 Message window.

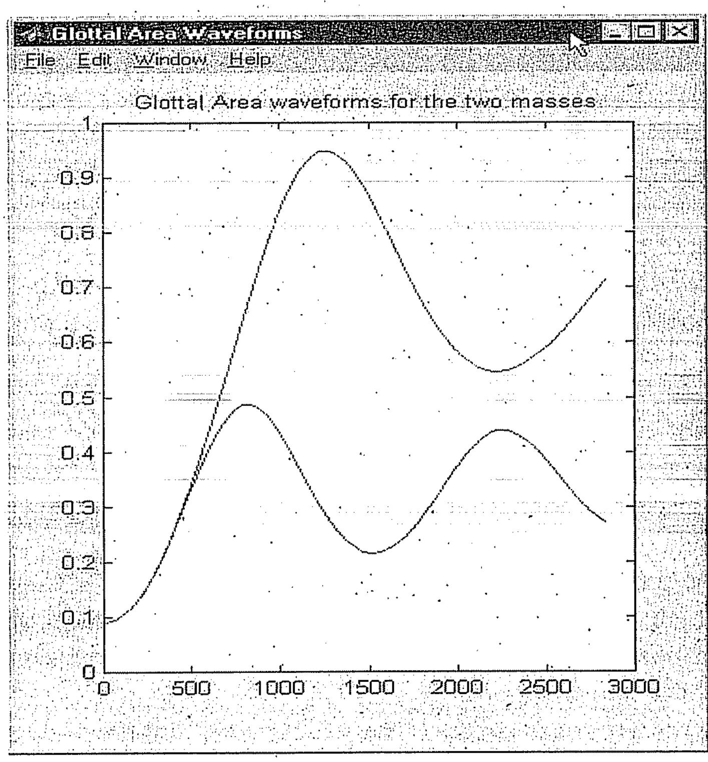

FIGURE 9.6 Glottal area waveforms calculated by the model.

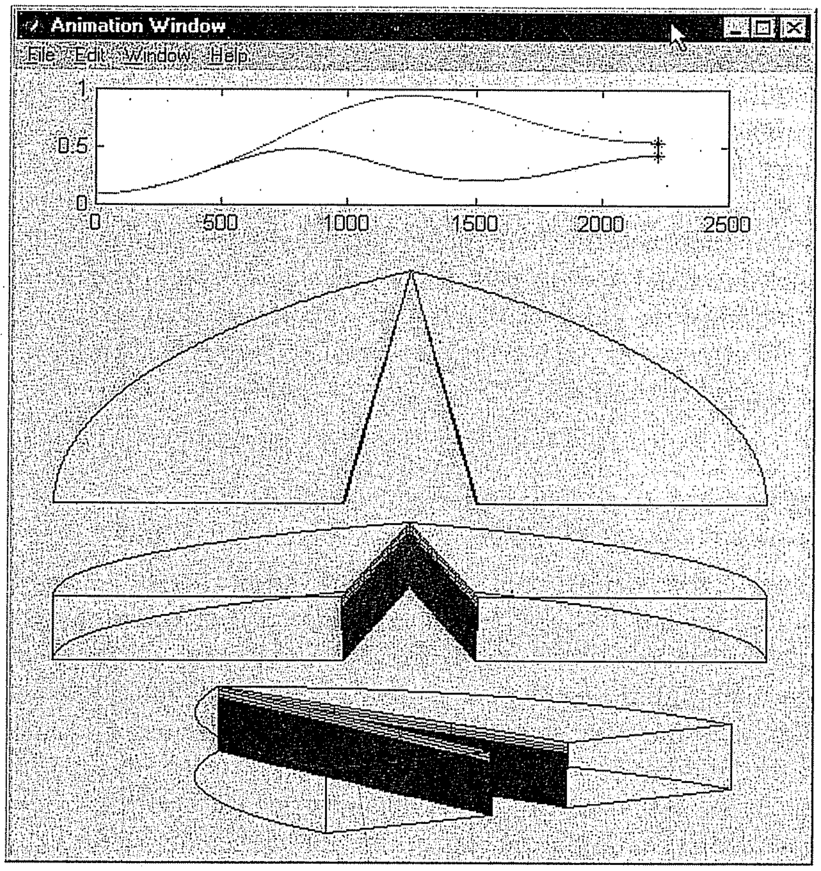

FIGURE 9.7 Animated movie of the vibratory motion of the two-mass vocal fold model.

model. The horizontal scale is arbitrary, with the value 3000 representing the end of a glottal vibratory cycle. Thus, the user can think of 3000 as the pitch period. The vertical scale is normalized to unity and depicts the amplitude of the area functions.

Once the glottal area functions are calculated, press the Make button. This action creates a 3-dimensional vibratory model of the motion of the vocal folds, as shown in Figure 9.7. This window is the animation window. As the vocal fold positions are being generated, a marker moves across the upper graphs of the two glottal area functions, while simultaneously the vibratory motion of the vocal folds is calculated. Three views of the vocal folds are provided. The top view is the same as that seen via a laryngeal mirror as shown in Chapter 3. The top of this view is anterior, while the bottom of this view is the posterior. The middle view of the vocal folds is looking at the vocal folds along a horizontal plane from posterior to anterior. The lower view is an isometric view at a 30-degree rotation from the other two views. The make action repeats the motion of the vocal folds twice upon completion of the necessary calculations. To view the animated movie, press the Play button, which is included in the pull down menu with the Make button. This action plays the animated movie of the vibratory motion of the vocal folds for 20 repetitions, that is, 20 pitch periods. The user may play the animated movie repeatedly. The colors shown for the two masses of the vocal fold model are generated by MATLAB, and are apparently not under user control. If the animated model does not appear in proper motion, that is, it appears as though the horizontal synchronization is not working properly, then reset the color resolution in the Windows 95 display setting to 256 colors or less. Note that the horizontal scale of the top graph (the glottal areas) in Figure 9.7 differs from that shown in Figure 9.6. This is because the values for mass 1, mass 2, the tension, excursion 1, and excursion 2 can affect the pitch period of the vibrating vocal folds. Thus, the final pitch period of the vibrating vocal folds is determined by the values of these parameters and is shown in the animation window. In this example the preliminary pitch period is 3000, while the final pitch period after the calculations by the model are completed is about 2500.

The user can experiment with various parameter settings to gain some experience with the manner by which the model parameters influence the vibratory motion of the vocal folds. One variation is to change the tension to $Q = 0.205$, excursion $1 = 0.05$, and excursion $2 = 0.01$, while all of the other parameter values remain as set in Figure 9.2. This will produce a very realistic vocal fold vibratory motion for a normal phonation.

9.4 THE RIBBON VOCAL FOLD MODEL

The ribbon vocal fold model has a control window similar to that of the two-mass vocal fold model, as shown in Figure 9.8. However, some of the model parameters are different. The vocal fold length, width, and thickness can affect the model. Changes in the width can cause strange vocal fold shapes to occur. It is recommended that the default settings of the length, width, and thickness be used. The vertical and horizontal phase difference parameters affect the motion of the vocal folds in the vertical and horizontal planes. The maximum excursion parameter controls the maximum excursion of the vocal folds (see Appendix 10). The fundamental frequency controls the relative pitch period of the vibratory motion of the vocal folds. For example, a setting of 300 causes the vocal fold model to vibrate twice as fast as a setting of 150. The abduction and shape quotients are defined in Appendix 10 and are calculated by the model. These values change with the values of

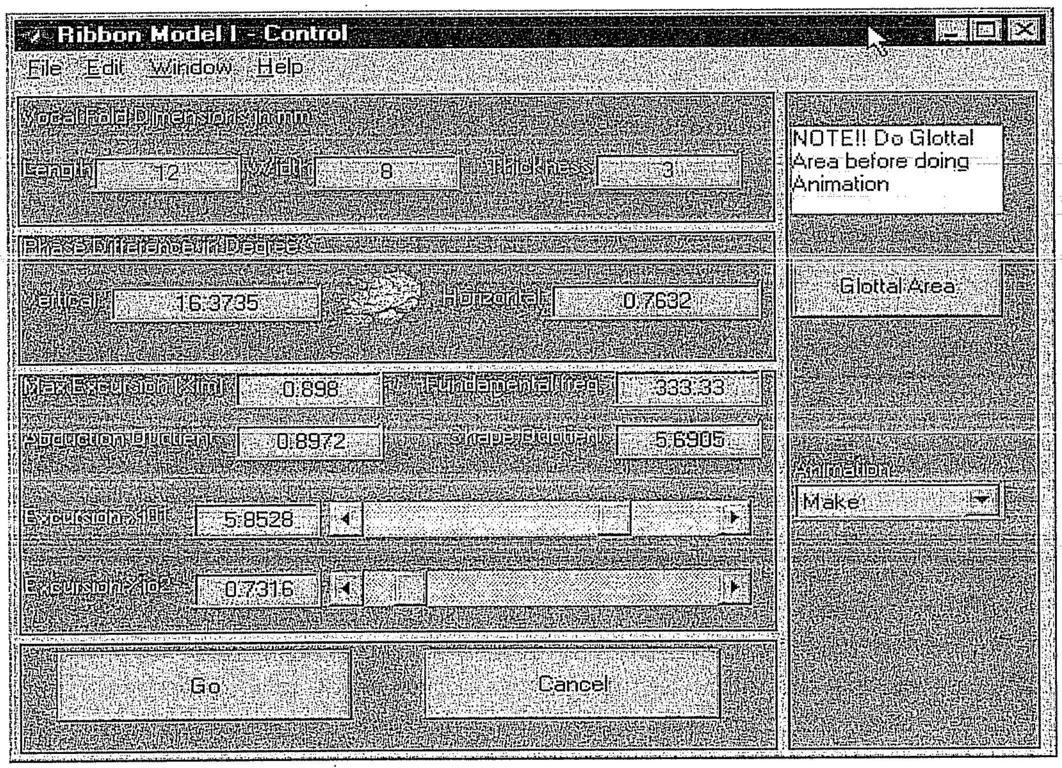

FIGURE 9.8 Ribbon vocal fold model control window.

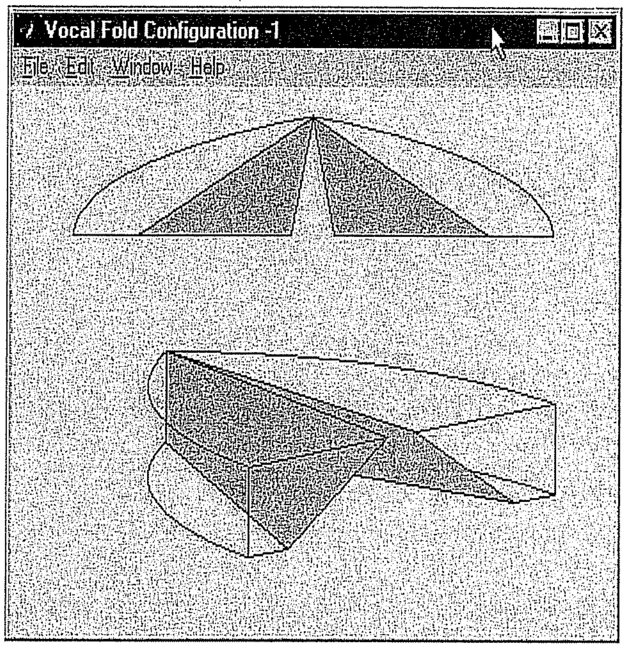

FIGURE 9.9 Static ribbon vocal fold model.

excursion xi01, excursion xi02, and maximum excursion xim. So the user does not set the values of the abduction and shape quotients. The excursion xi01 setting controls the pre-phonatory setting of the lower edge of the vocal fold model, while excursion xi02 controls the pre-phonatory setting of the upper edge of the vocal fold model.

After setting the desired parameter values, start the model by pressing the Go button. This action opens the windows shown in Figures 9.9 and 9.10, which are similar to those for the two-mass vocal fold model.

The parameter control settings used for Figures 9.9 and 9.10 are the default values shown in Figure 9.8. Next press the Glottal Area button in Figure 9.8. This opens a message window (not shown) similar to Figure 9.5. Upon completion of the model calculations, the message window closes and simultaneously opens the glottal area window shown in Figure 9.11. This figure shows three graphs: the projected glottal area, the vocal fold contact area, and the estimated electroglottographic waveform.

The next step is to press the make button in Figure 9.8, which creates a 3-dimensional vibratory ribbon model of the motion of the vocal folds that is similar to the two-mass vocal fold model. This model is shown in Figure 9.12, which is the animation window. As the vocal fold positions are being generated, a marker moves across the upper graph of the projected glottal area function, while simultaneously the vibratory motion of the vocal folds is calculated. Three views of the vocal folds are provided. The top view is the same as that seen via a laryngeal mirror as shown in Chapter 3. The top of this view is anterior, while the bottom of this view is the posterior. The middle view of the vocal folds is looking at the vocal folds along a horizontal plane from posterior to anterior. The lower view is an isometric view at a 30-degree rotation from the other two views. The make action repeats the motion of the vocal folds 20 times upon completion of the necessary calculations. To view the animated movie, press the Play button, which is included in the pull down menu with the Make button. This action plays the animated movie of the vibratory motion of the vocal folds for 20 repetitions; that is, 20 pitch periods. The user may play the animated movie repeatedly. While the figures in the text are black and white, the MATLAB software generates colors that will appear on a color computer monitor. Thus, the two masses of the

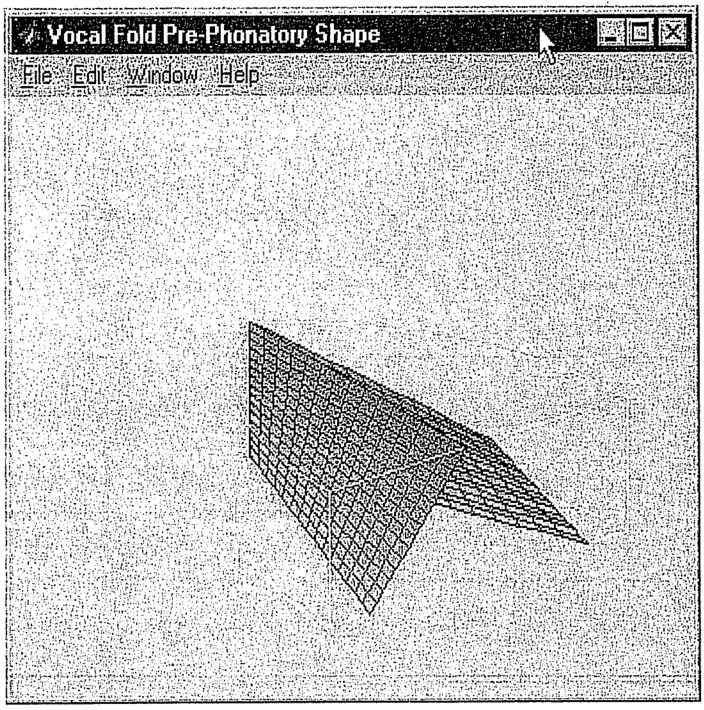

FIGURE 9.10 Pre-phonatory view of ribbon model.

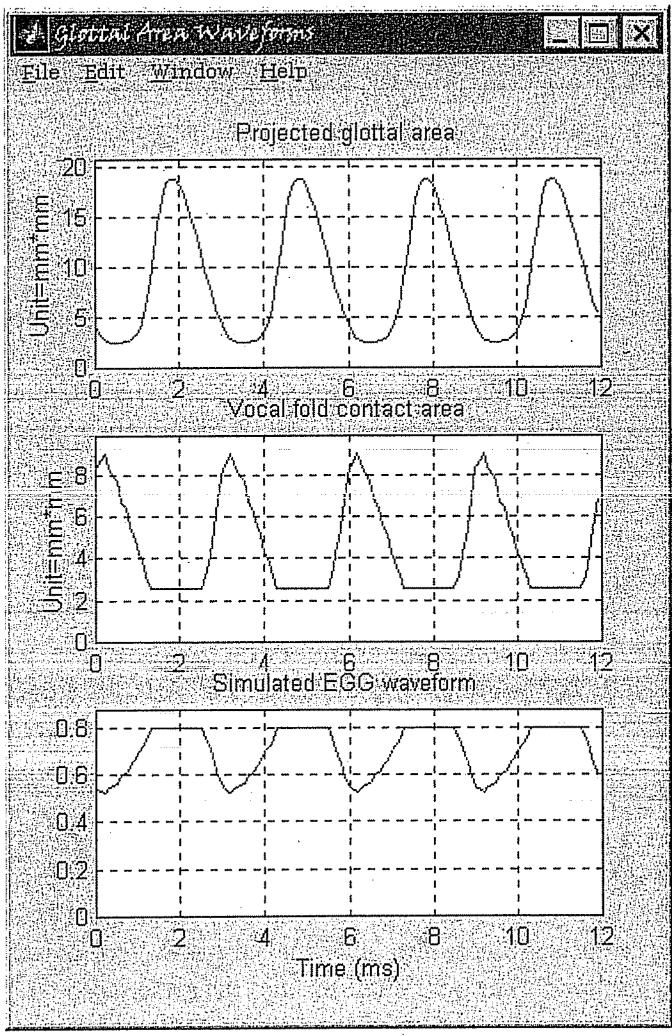

FIGURE 9.11 Glottal area waveforms calculated by the ribbon model.

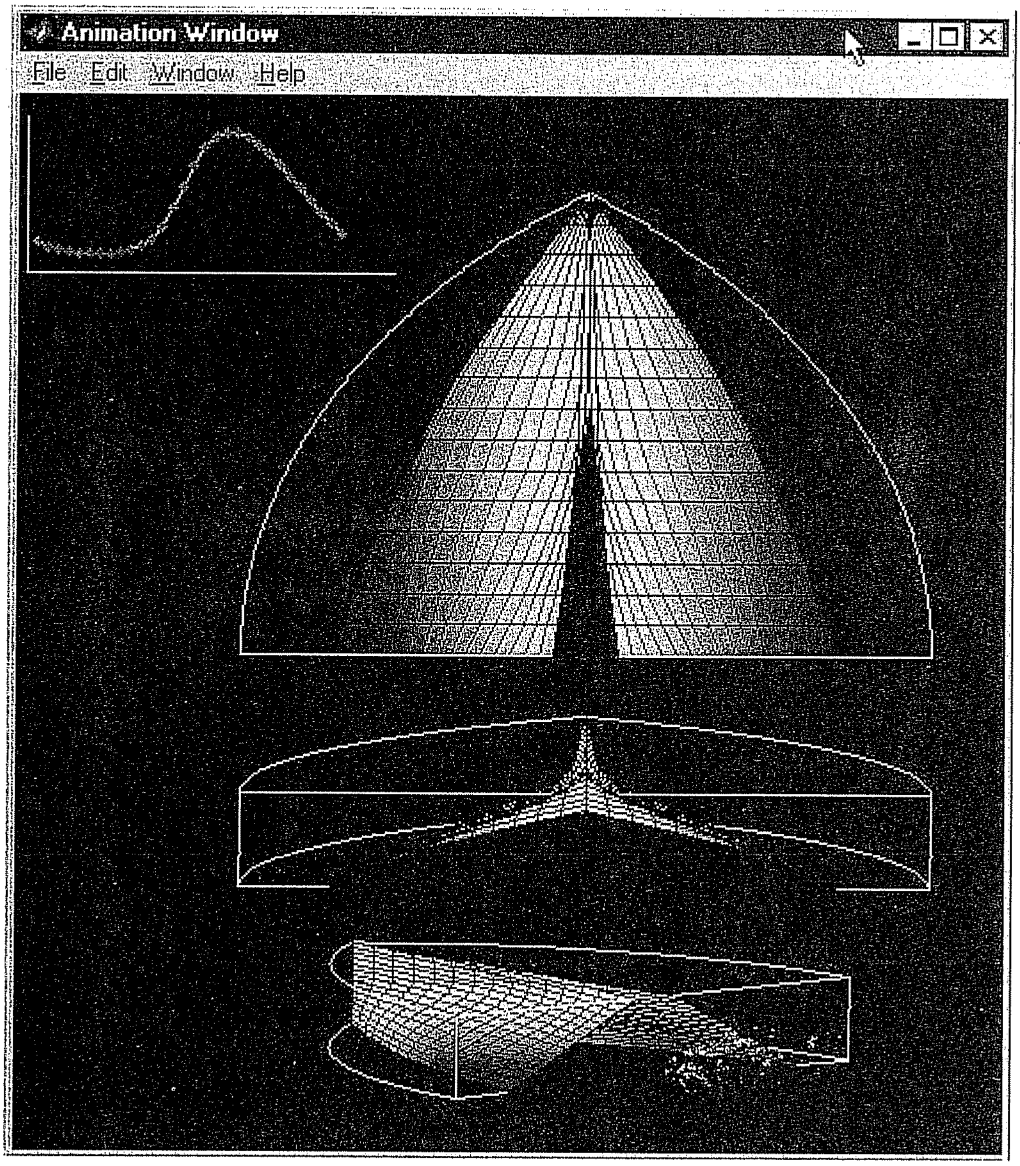

FIGURE 9.12 Animated movie of the vibratory motion of the ribbon vocal fold model.

vocal fold model will appear in color. If the animated model does not appear in proper motion, that is, it appears as though the horizontal synchronization is not working properly, then reset the color resolution in the Windows 95 display setting to 256 colors or less. Note that the horizontal scale of the top graph (the projected glottal area) in Figure 9.11 differs from that shown in Figure 9.12. However, the only difference is that the graph in Figure 9.12 is for one pitch period, otherwise the graphs are the same.

Several parameter values that result in a very realistic vocal fold vibratory motion are the following. Use the default settings except change vertical phase difference to 8. Another setting is maximum excursion xim = 1.3 (a setting of 2 is also good), excursion xi01 = 4.57 (or 3.09), and excursion xi02 = 0.73. The fundamental frequency can be left at 333 or changed to 150. The vertical phase difference can be left at 16 or changed to 8.

As the maximum excursion xim gets large, the user can see gaps occur between the vocal folds during vibration. As the excursion xi01 and excursion xi02 are made large, the user can see the vibratory motion of the vocal folds very clearly. No vocal fold contact takes place. The top view of the vocal folds is not good; however, the isometric view is very good.

A slight modification to the software allows the projected glottal area waveform to be saved and used as an excitation waveform for speech synthesis. Similarly, the movies of the vibratory motion of the vocal folds can be saved. However, this is not recommended since the movie files are very large (many MB) and can easily be regenerated at will.

Several other models of the vibratory motion of the vocal folds are briefly mentioned in Appendix 10. However, these models are not implemented here.

PROBLEMS

9.1 For the two-mass vocal fold model, change the thickness settings for mass 1 and mass 2 to 0.3 for each mass. Observe the pre-phonatory shape and the glottal area functions and the movie. Describe the differences between these results and those using the default settings.

9.2 For the two-mass vocal fold model, change the settings for mass 1 to 0.025 and that for mass 2 to 0.125. Observe the pre-phonatory shape and the glottal area functions and the movie. Describe the differences between these results and those using the default settings.

9.3 For the two-mass vocal fold model, experiment with various values for the pressure setting. Discuss the effect this parameter has on the vibratory motion of the vocal folds.

9.4 For the two-mass vocal fold model, experiment with various values of the vocal tract area and the number of vocal tract segments. Discuss the effect this parameter has on the vibratory motion of the vocal folds.

9.5 For the ribbon vocal fold model, change the vertical phase setting as follows: 2.0, 4.0, 8.0, 16.0, and 32.0. Let the other values remain at their default setting. Describe the effect this parameter has on the vibratory motion of the vocal folds.

9.6 For the ribbon vocal fold model, change the horizontal phase setting as follows: 0.1, 0.5, 1.0, 4.0, 8.0, 16.0, and 32.0. Let the other values remain at their default setting. Describe the effect this parameter has on the vibratory motion of the vocal folds.

9.7 For the ribbon vocal fold model, set the excursion xi02 to 2.5. Let the excursion xi01 be 1.0, 2.5, and 5.0. Let the other values remain at their default setting. Describe the effect this parameter has on the vibratory motion of the vocal folds.