FormantMeasurer: Software for efficient human-supervised measurement of format trajectories

© 2008, 2010, 2011 Geoffrey Stewart Morrison & Terrance M. Nearey

Available from http://geoff-morrison.net

Release 2011-05-26	- undo last hand edit command now works again
	- resynthesis amplitude range setting removed, synthesis in now better without this (remove dBdownCriterionStr from
	FormantMeasurer.ini)
Release 2010-11-30	- in Track Edit window if save and advance (right click) command is issued immediately following a manual edit (left click)
	command, a dialogue box checks if this is what the user really wants to do.
	- updated voice source in synthesiser supplied by Terry
Release 2010-05-23	- if requested start and end markers do not exist, or if markers are too close to the beginning or end of the sound file a warning is
	displayed and the marked section is skipped
Release 2010-01-19	- automatically runs a range of minimum and maximum values for f0 calculation, runs a goodness heuristic, and allows user to
	select best track
	- now displays and allows individual track selection from four-tracks-below-cutoff as well as three-tracks-below-cutoff
	- hard selection of formats added: press key {[1], [2], [3]} then click on track to select {F1, F2, F3}
	- old function of keys [1], [2], [3], cycle through next best {F1, F2, F3} track, moved to keys [q], [w], [e] (keys immediately
	below [1], [2], [3])
	- added rfstable to heuristics for deciding best individual tracks, penalty according to amount of change in trajectory when one
	additional coefficient is added
	- process-only-one-file option now works
Release 2008-12-21	 added option to save window positions
Release 2008-12-19	- first general release distributed for trial and feedback, please report bugs

This software is provided as-is without any guarantee that it will work. Geoff is willing to give some end-user support, but please read this document thoroughly and follow all the instructions before contacting him. Please do NOT contact Terry asking for end-user support. This document and the released version of the software were prepared by Geoff.

This software is provided free-of-charge for academic not-for-profit research. Please include appropriate acknowledgments in published papers.

Commercial use of this software in whole or in part is strictly prohibited except with the prior consent of the copyright holders.

Contents

Des	cription	3	
Requirements			
	Matlab code version:	5	
	Compiled version:	5	
	Sound Files and Marker Files:	5	
Installing and Running			
	Matlab code version:	6	
	Compiled version:	6	
Choosing Options			
	File Options	8	
	Marker Options	8	
	Formant Measurement Options	9	
	Amplitude Measurement Options	9	
	Fundamental Frequency Measurement Options	9	
Windows			
Sele	ecting Tracks	12	
	Selecting Formant Tracks	12	
	Synthetic Vowel	14	
	Selecting f0 Track	15	
	Moving On	17	
Editing Tracks			
Mouse Clicks & Keystrokes		21	
	When the mouse cursor is in the TrackSelect window	21	
	When the mouse cursor is in the Fundamental Frequency window	21	
	When the mouse cursor is in the TrackEdit window	22	
Output			
Customisation			
Γroubleshooting			
Ref	References		

Description

The FormantMeasurer software measures formant trajectories using a range of parameters for linear-predictive-coding (LPC), runs some heuristics to attempt to identify the best track for each of the first three formants (F1, F2, F3), and presents the results to a human for checking. The software looks for three (and four) formant tracks below a specified cutoff frequency. The user specifies the lower and upper bounds for the cutoff frequency, and the software tries eight different cutoff frequencies in this range. Eight different tracksets corresponding to the eight different cutoff frequencies are overlayed on the spectrogram and displayed on the screen. The software suggests the best F1, F2, and F3 tracks, and the user can accept these or select others. The software also measures the fundamental frequency (the user can choose from sixteen different f0 tracks) and amplitude of the vowel being measured and combines these with the selected formant tracks to synthesise a vowel. In addition to the visual display, the user can therefore also auditorily compare the original vowel with one synthesised on the basis of the measured formant tracks. Finally, the user has the option of hand editing the formant tracks.

The core of the formant measuring procedure was described in the conference presentation Nearey, Assmann, & Hillenbrand (2002). A copy of the poster from this presentation is available on Terry's website http://www.ualberta.ca/~tnearey/. Formants are extracted from autocorrelated LPC coefficients, a procedure which requires one to set an upper frequency bound and specify the number of LPC coefficients to use. Although there are some rules-of-thumb with respect to selecting a cutoff frequency and the number of coefficients, in practice obtaining good results often requires a certain amount of trial-and-error. The FormantMeasurer software fixes the number of coefficients at 9 (this will find four peaks out of which the three best formant candidates will be selected), the sampling frequency is fixed at 10 kHz, and the cutoff frequency is roved. The idea is to put the cutoff frequency between F3 and F4 so that there are exactly three formants below the cutoff frequency, and (hopefully) produce good formant tracking results for F1, F2, and F3.

The software calculates a goodness score for each trackset based on a combination of eight heuristics:

- 1. Presence: To what extent are there good candidates available to fill the time slots?
- 2. BwReason: Are the bandwidths of the peaks reasonable?
- 3. AmpReason: Is the amplitude reasonable?
- 4. ContReason: Is there reasonable continuity within each formant track?

¹ The formant-extraction procedure is from Markel & Gray (1976). It picks the locally best values at each frame, and then interpolates over frames which do not have a full set of values. In the future, we might implement a dynamic programming / Viterbi algorithm approach looking for the globally best track, which would be less susceptible to jumps.

- 5. DistReason: Are the F2-F1 and F3-F2 distances reasonable?
- 6. RangeReason: Are the formant ranges reasonable for given the frequency cutoff?
- 7. RfStable: Are formant tracks relatively stable when the number of LPC coefficients is increased from 9 to 10?
- 8. Rabs: Correlation of resynthesised spectrogram with original.

In addition to calculating a best trackset on the basis of these eight heuristics, the current version of the software also uses heuristics 1, 2, 3, 4, and 7 to independently calculate a best F1, a best F2, and a best F3 across all eight tracksets: the best F1 could come from one trackset, the best F2 from another, and the best F3 from a third.

In addition to finding three formants below the cutoff, the software also finds, and makes available for selection, four formants below the cutoff. The heuristics are not applied to the latter.

For fundamental frequency, the software tries four different minimum f0 values and four different maximum f0 values, and uses the mean of the instantaneous second derivatives as the basis for a heuristic goodness score. Tracks with large mean second derivatives probably have jumps and are poor (unless they include a switch from modal to creaky voice), and tracks with very small mean second derivatives are probably unrealistically smooth.

The software is set up to first calculate formant tracks, fundamental frequency tracks, and amplitude in batch mode, then present the results to a human for checking. Thus the relatively time-consuming calculations do not slow down the user while they are checking the results. The human interface has an ergonomic design: Generally the right hand stays on the mouse, and the most frequent keystrokes are single keystrokes on the left side of the keyboard requiring minimum movement of the left hand. The user can customise the keystrokes and mouse buttons (e.g., to create an ergonomically efficient left-handed version).

The software is available as Matlab code and also in a compiled version; however, in both cases the output consists of Matlab data files (mat files). It would be possible to convert the output to text files, but as present it is envisioned that the software will be used by researchers who themselves use Matlab or by collaborators of researchers who use Matlab. The advantage of the compiled version is that it does not require a Matlab licence, and therefore can be run on more machines.

To a first approximation, the core code was written by Terry and the human interface by Geoff. The code has been in use for several years and its functionality has gradually been altered and improved. The current version reflects Geoff's needs and design philosophy. The combination of two programmers'

different coding styles and the gradual evolution of the software has not resulted in a particularly elegant set of code, but it works and cleaning up the code has a low priority in the research programme.

Requirements

Matlab code version:

- installed and licenced version of Matlab
- installed and licenced version of Matlab Signal Processing Toolbox and Statistics Toolbox
- Windows XP/Vista/7

The software has been tested on 32 and 64 bit versions of Matlab R2008b and R2009b running under 32 and 64 bit versions of Windows XP respectively, and Matlab R2010a running under Windows 7 64 bit. You may have to make changes if you want to run it on a different version of Matlab or using a different operating system. Geoff will not provide end-user support for other versions of Matlab or other operating systems.

Compiled version:

- Matlab Compiler Runtime version 7.9 (this can be downloaded from Geoff's ftp site; however, you will need to ask Geoff for the username and password the licencing agreement doesn't allow it to be posted to a generally accessible location on the internet).
- 32 bit Windows XP/Vista

Sound Files and Marker Files:

FormantMeasurer will measure wav sound files only (these should be raw PCM), and each wav file to be measured must have an associated mat file which contains markers indicating the portions of the wav file which are to be measured. These mat files are created using Geoff's SoundLabeller software which is available from his website http://geoff-morrison.net. There must be no other mat files in the folder in which the sound and marker files are located. If you want to test FormantMeasurer without using SoundLabeller you can download a set of sample sound and marker files from Geoff's website.

Installing and Running

Matlab code version:

Unzip the files in the archive FormantMeasurer.zip into a folder named FormantMeasurer.

- To start the software, run FormantMeasurer script.m or FormantMeasurer function.m in Matlab.

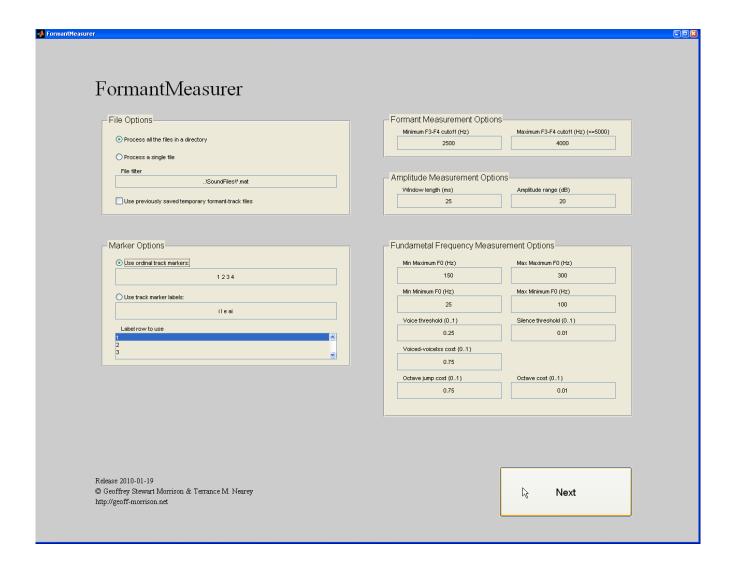
Compiled version:

- 1. Install Matlab Compiler Runtime version 7.9
- 2. Place FormantMeasurer_pkg.exe into a folder named FormantMeasurer, and double click on FormantMeasurer pkg.exe.
- To start the software, double click on FormantMeasurer.exe. (The first time you use the software, it might take a while to start.)

Choosing Options

When you start the software the screen shown below should appear.

- Select your options then press Next.
- Select the sound file to process or the folder containing the sound files to process.
- Wait for the files to be processed in batch mode.



File Options

- Choose to process all or a number of sound files which are all located in the same folder, or choose to process a single file.
- The File Filter is particularly useful if you want to process a subset of files within a folder. You can enter a string providing the folder in which your sound files are located and a filename filter indicating which files to process within that folder. The default filter is set to *.mat, i.e., all the mat files in the folder. If you entered ..\SoundFiles\h*t.mat you would be processing all the files with names beginning with h and ending with t and having file extension .mat in the folder SoundFiles which is a sister folder to the FormantMeasurer folder (.\ specifies the same level in the folder hierarchy, ..\ up one level, ..\..\ up two levels, etc.). Make sure you specify .mat as the file extension. You can override the filter by later selecting a folder or file which does not match the folder and filter options specified here.
- If you have previously processed files and have not erased the temporary formant-track files, you can select the option to use the previously saved files rather than recalculating them.

Marker Options

- You can choose to measure between predefined sets of ordinal markers, e.g., if you have four markers in every file and the first vowel occurs between the 1st and 2nd markers, and the second vowel occurs between the 3rd and 4th markers, then select this option and type 1 2 3 4 in the box. Numbers must be separated by spaces and for each pair of numbers the second number must be higher than the first. Sequences such as 1 2, 1 2 3 4 5 6, 1 2 5 6, 10 11 23 24, 3 4 1 2, 1 2 2 3 and 1 3 2 4 are permitted, but sequences such as 1 1 2 2, 2 1 4 3, 1, and 1 2 3 are not.
- Alternatively, you can choose to measure between pairs of adjacent markers corresponding to markers labels, e.g., if you have an unpredictable number of markers in each file but have labelled some sections as i, I, e, and ai, then you can measure those sections by selecting this option and typing i I e ai in the box. Labels must be separated by spaces. You can include as many or as few labels as you like.

SoundLabeller allows multiple rows of markers and labels. Select the number of the row which contains the markers and labels you want to use.

Formant Measurement Options

– Enter the lower and upper bounds for the F3-F4 cutoff frequency. A lower range, e.g., 2500–4000 Hz, should be more effective for adult male voices, and a higher range, e.g., 3000–4500 Hz, more effective for adult female voices. Note that there may be vowel tokens in which, for example, the best F1 track is obtained using an F3-F4 cutoff which is actually below F3.

Amplitude Measurement Options

– Enter the window length for calculating the RMS amplitude of the signal. 25 Hz should work well for adult male voices and 20 Hz for adult females.

Fundamental Frequency Measurement Options

- The algorithm for measuring fundamental frequency is that described in Boersma (1993). The paper can be downloaded from http://www.fon.hum.uva.nl/Proceedings/IFA-Proceedings.html. A detailed description of the formant options can be found in the paper. Usually most of the default options will work well, and it may be sufficient to adjust the minimum and maximum f0 values to obtain good results.
- The minimum f0 value is roved over four values between Min Minimum F0 and Max Minimum F0.
 Try 25 Hz to 100 Hz.
- The maximum f0 value is roved over four values between Min Maximum F0 and Max Maximum F0.
 Try 150 Hz to 300 Hz.
- All the options listed can also be changed and f0 recalculated when tracking results are being checked

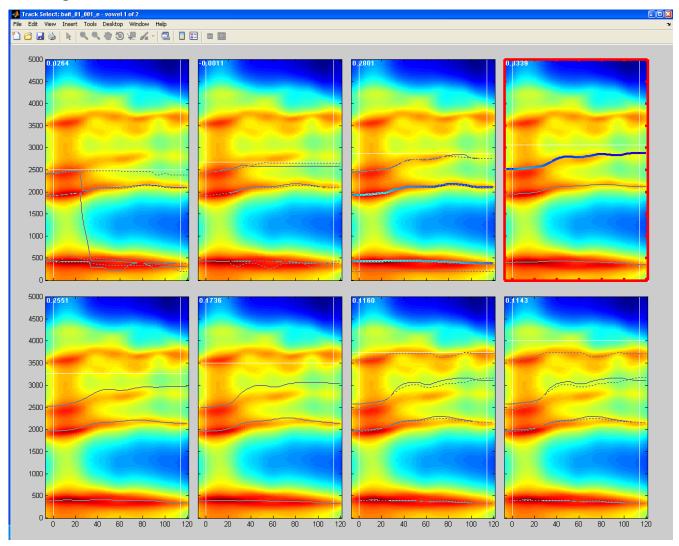
– A number of parameters described in Boersma (1993) are not included in the options list, but can be changed by altering their values in the FormantMeasurer.ini file.

Windows

- Once the files have been processed, the original and resynthesised vowels will be played and four windows will appear:
 - 1. Track Select displays eight tracksets and allows the user to select formant tracks
 - 2. Fundamental Frequency displays the f0 tracks
 - 3. Original and Synthesised Waveforms displays the original and synthesised waveforms
 - 4. Track Edit will display the selected formant tracks and allow hand editing
- The blue bar at the top of the window gives the window title and the name an number of the vowel being measured.
- The software is easiest to use if you have a lot of screen space. Geoff's desktop computers have dual-head video cards, he moves the Track Select window to full-screen on the monitor on the right, the Fundamental Frequency and Original and Synthesised Waveforms non-overlapping less-than-full-screen on the monitor on the left, and, after the first track has been selected, moves the Track Edit window to full-screen on the monitor on the left.
- If you are happy with the position of the windows, you can save the window positions. The save-window-positions option appears when you quit the software using the Escape key. A file WindowPositions.ini is saved. The next time you use the software, the windows will automatically move to the positions you saved. Selecting the save-window-positions option again will overwrite the exiting WindowPositions.ini file. The positions are dependent on the number of monitors you have, the number of pixels the monitor is set to display, and the width-to-heigh ratio of the monitor. If you change your monitor set up, delete the existing WindowPositions.ini file before running the software, otherwise it is possible that some of the windows may not be visible.

Selecting Tracks

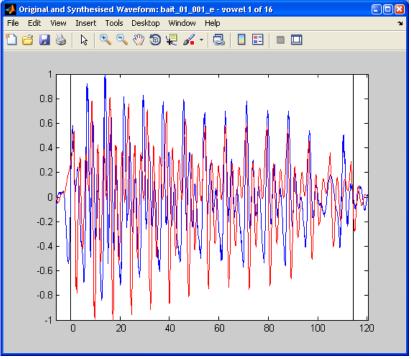
Selecting Formant Tracks



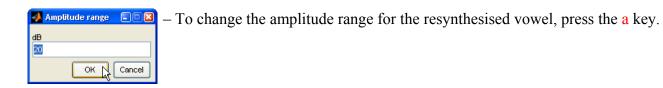
– The Track Select window displays eight tracksets corresponding to the eight different F3-F4 cutoff values. Each trackset is overlayed on the spectrogram of the original vowel. Thin vertical white lines indicate the location of the SoundLabeller markers. A thin horizontal white line indicates the F3-F4 cutoff frequency for each trackset, and a number in the top left corner indicates its goodness score. The tracks from **three-formants-below-the-cutoff** appear as **solid lines**, and the tracks from **four-formants-below-the-cutoff** appear as **dotted lines**. The F1, F2, and F3 tracks with **thick lines** are those which were determined to be the best on the basis of the heuristics, or which were selected by the user. The trackset in which the last selection was made has a red boarder, this is initially set to the trackset determined to be the best on the basis of the heuristics.

- To change the selected formant tracks, click on a solid-line track with the primary mouse button (left click). You don't have to be particularly accurate in you click, you just have to be closer to this solid-line track than any other solid-line track within this trackset. If the track you clicked on was an F1, then the selected F1 track will change, if the track you clicked on was an F2, then the selected F2 track will change, if the track you clicked on was an F3, then the selected F3 track will change.
- The currently selected tracks will be displayed as thick lines in the Track Select window, and will also be displayed in the Track Edit window.
- Alternatively, to select any formant track, either solid or dotted, as an F1: press and release the 1 key, then click with the primary mouse button (left click) on the desired track (you have to be closer to this track than any other solid-line or dotted-line track within this trackset). For F2 press 2, and for F3 press 3. In addition to allowing selection of dotted-line tracks, this allows, for example, the third solid-line track from a low-cutoff trackset to be selected as F2 (this could be a better fit to F2 than any of the second solid-line tracks from any of the tracksets).
- Alternatively, to cycle through from the best F1 track to the second-best F1 track, to the third-best F1 etc., press the q key. For F2 press w, and for F3 press e (these are the keys immediately below the 1, 2, and 3). This only cycles though solid-line tracks.

Synthetic Vowel

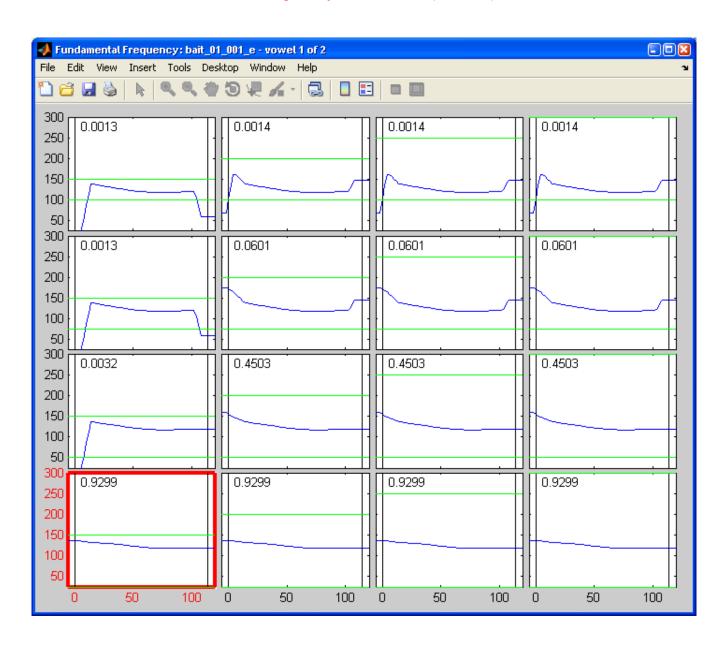


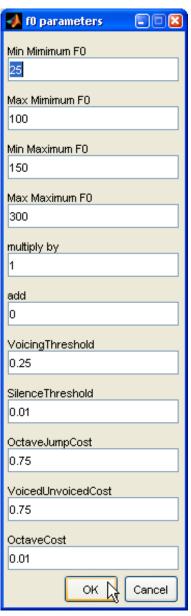
- To listen to the original vowel and the vowel synthesised on the basis of the selected formant tracks, press the spacebar. The original waveform appear in blue, and the synthesised waveform in red.



Selecting f0 Track

- To select a different f0 track, press the s key (the Fundamental Frequency window will be selected) then click on the desired track with the primary mouse button (left click).





– To change the f0 tracking options, while in the Fundamental Frequency window press the f key. To reset the f0 tracking options to the default values, press the d key. A dialogue box will appear. Type in the desired values and click OK.

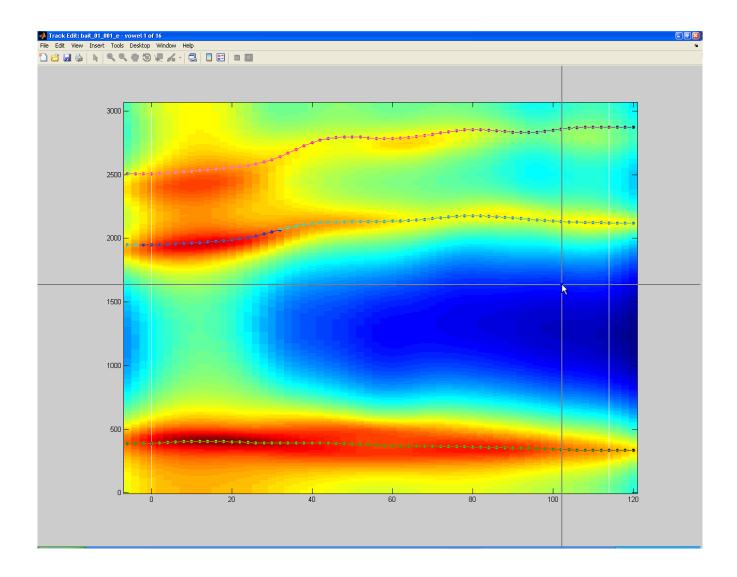
- To return to the Track Select window click the secondary mouse button (right click) anywhere in the Fundamental Frequency window.

Moving On

- When you are satisfied with the selection of formant tracks, click in the Track Select window using the secondary mouse button (right click).
- If you want to move forward to the next vowel or backwards to the previous vowel without saving the results, press the right arrow or left arrow respectively.
- If you want to removing (delete) any exiting results for the present vowel, press the R key.
- To exit FormantMeasurer, press the Escape key. Do NOT attempt to exit by closing all the windows, and avoid exiting via control-c. Exiting via the Escape key allows the software to run clean-up activities and save settings.

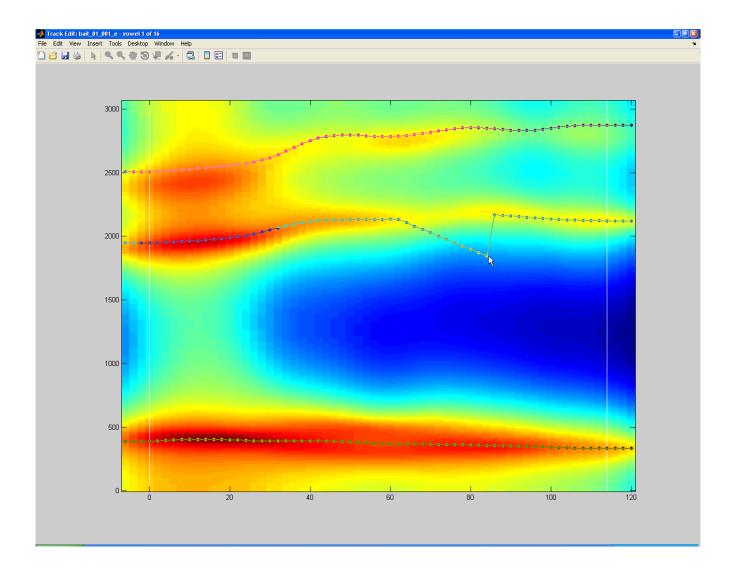
Editing Tracks

– After you have clicked in the Track Select window using the secondary mouse button (right click), the spectrogram and the selected formant tracks will appear in the Track Edit window.



– If you are happy with the results click in the Track Edit window using the secondary mouse button (right click), and the selected formant tracks, the f0 track, and the amplitude will be saved. The software will then advance to the next vowel.

– If you want to hand edit a formant track, click with the primary mouse button (left click) on the track you want to edit. This will select a time value on the closest track. Click again and the software will draw a linear interpolation between the first time point and the time and frequency point of the new click. Keep clicking to add additional linear interpolations between the new point and closest previously clicked points.



- To undo the last interpolation, press the z key (only the last interpolation can be undone).
- To edit a different track, or a different part of the same track without interpolating from the previously clicked points, press the x key or the spacebar.

- To listen to the original vowel and the vowel synthesised on the basis of the edited formant tracks, press the spacebar.
- To change the amplitude range for the resynthesised vowel, press the a key.
- To select a different f0 track, press the s key.
- To return to the Track Select window and change or reset the selected tracks, press the c key.
- If you want to remove (delete) any exiting results for the present vowel, press the R key.
- To exit FormantMeasurer, press the Escape key.
- When you are happy with the results click in the Track Edit window using the secondary mouse button (right click), and the selected formant tracks, the f0 track, and the amplitude will be saved. The software will then advance to the next vowel.

Mouse Clicks & Keystrokes

When the mouse cursor is in the TrackSelect window:

left click - select track right click - go to TrackEditor window - play original and resynthesied vowel spacebar 1 - make the closest track to the <u>next</u> <u>left click</u> the F1 track 2 - make the closest track to the next left click the F2 track 3 - make the closest track to the <u>next</u> <u>left click</u> the F3 track - cycle through best, second best, third best, etc. F1 track q - cycle through best, second best, third best, etc. F2 track W - cycle through best, second best, third best, etc. F3 track e - go to Fundamental Frequency window to select f0 track S - change amplitude range option a - delete saved results for current vowel, if any R - move forwards to next vowel without saving - move backwards to previous vowel without saving Esc - quit

When the mouse cursor is in the Fundamental Frequency window:

left click - select track right click - go to TrackSelect window spacebar - play original and resynthesied vowel f - change f0 measurement options d - change f0 measurement options to default values a - change amplitude range option - delete saved results for current vowel, if any R - move forwards to next vowel without saving - move backwards to previous vowel without saving Esc - quit

When the mouse cursor is in the TrackEdit window:

left click - edit track

right click - save and advance to next vowel

spacebar - play original and resynthesied vowel

x - change track / part of track being edited

z - undo last edit

c - return to TrackSelect window to reset or change track selections

d - delete saved results for current vowel, if any, and advance to next vowel

s - go to Fundamental Frequency window to select f0 track

a - change amplitude range option

R - delete saved results for current vowel, if any

→ - move forwards to next vowel without saving

- move backwards to previous vowel without saving

Esc - quit

Output

- The output of the FormantMeasurer is a series of mat files which are written to a subfolder of the folder in which the wav files and mat files from SoundLabeller are located. The subfolder is named FormantMeasures. During batch processing the software saves temporary files prefixed with tmp, and after checking the final results are prefixed with formants. The next part of the filename is the name of the original sound and marker files, and the end of the filename included the number of the marker at which the vowel began followed by the label for that vowel, if any, from the marker file, e.g., formants_bait_01_001_e.mat is the final result from the analysis of sound and marker files bait_01_.wav and bait_01_.mat, and the vowel beginning at the first marker in the marker file, and the label associated with this marker is 'e'.
- The contents of final output mat files consists of two variables:
 - v_duration is a scalar giving the duration of the vowel as measured between the two markers from SoundLabeller.
 - f_measures is a two-dimensional array.
 - Each row is a time frame, note that the first three and last three rows extend beyond the markers from SoundLabeller.
 - Columns are:
 - 1. time index in milliseconds relative to the marker at the beginning of the vowel
 - 2. f0 in hertz
 - 3. F1 in hertz
 - 4 F2 in hertz
 - 5 F3 in hertz
 - 6. RMS amplitude in decibels relative to the peak amplitude in the sound file.

Customisation

- If you want to change the default options you can change the file FormantMeasurer.ini in a text editor. Save a backup copy of the original file first in case you need to restore it. Save the edited file as a plain ANSI text file. The file consists of two columns. The first column is the variable name and the second column is the default value. The columns must be separated by a tab. Elements in the values of trackmarks_to_use_str and labels_to_measure_str must be separated by spaces.
- If you want to change the mouse-click and keyboard command keys (e.g., if you want to make an ergonomically efficient left-handed version of the software, or if you just don't like Geoff's default options) you can change the file KeyboardMouse.ini in a text editor. Save a backup copy of the original file first in case you need to restore it. Save the edited file as a plain ANSI text file. The file consists of two columns. The first column is the variable name and the second column is the default value. The columns must be separated by a tab. The values can be entered as characters or their corresponding Unicode number expressed in base ten. If you want to use a number key or an arrow key, you will have to enter its Unicode number, e.g., key "1" is Unicode 49. To find the value of a button press in Matlab execute the command [x, y, button] = ginput(1), press the key, and read the value of button.

Warning: If you make an error in editing either of these files, FormantMeasurer may crash or not function correctly.

Troubleshooting

— With the exception of dialogue boxes, all mouse click and keystroke inputs should be either in the Track Select window, Fundamental Frequency window, or the Track Edit window. You should see cross-hairs when you move the mouse pointer over the window which will currently accept input. To move between windows, use the commands listed under *Mouse Clicks & Keystrokes* above. Do not try to change windows by left clicking. If you accidentally left click on the wrong window, then you will not see cross-hairs on that windows. In this case, move the cursor over and left click on the Track Select window, Fundamental Frequency window, or the Track Edit window — once you have clicked on the correct window, the one on which cross-hairs (re)appear, FormantMeasurer should accept input again.

- If FormantMeasurer has frozen and reselecting the Track Select window, Fundamental Frequency window, or the Track Edit window does not work, Ctrl + c should exit the software. The temporary formant-track files should be available the next time you start the software.
- If there is a bug in the software and you are running the complied version, you may be able to get access to error messages by running the software from a Command Prompt window. Please report bugs to Geoff once you have verified that they are not due to problems with the a badly edited ini file or problems with the input way or mat files.

References

- Boersma, P. (1993) 'Accurate short-term analysis of the fundametal frequency and the harmonics-to-noise ratio of a sampled sound', *Proceedings of the Institute of Phonetic Sciences, Amsterdam*, 17, 97–110.
- Markel, J. D. and Gray, A. H. (1976) Linear Prediction of Speech, Berlin: Springer-Verlag.
- Nearey, T. M., Assmann, P. F. and Hillenbrand, J. M. (2002) 'Evaluation of a strategy for automatic formant tracking', *Journal of the Acoustical Society of America*, 112, 2323.