

SoundLabeller: Ergonomically designed software for marking and labelling sections of sound files

© 2008, 2009, 2010, 2011, 2012 Geoffrey Stewart Morrison

<http://geoff-morrison.net>

Release 2012-07-30	<ul style="list-style-type: none">– Calculation of time_axis now correct when spectrogram not preprocessed.– no longer crashes attempting to draw no spectrogram after a partial deletion– no longer saving empty mat files– no longer crashes if backspace key is pressed when the label field is empty– fixed a bug if there is another already saved label file whose name starts the same as the label file being searched for
Release 2011-04-03	<ul style="list-style-type: none">– Added option to zero out a portion of the recording.
Release 2010-11-01	<ul style="list-style-type: none">– Added option to review already-saved markers only.
Release 2010-10-28	<ul style="list-style-type: none">– Fixed bug which may cause crash when a file is deleted.
Release 2010-10-16	<ul style="list-style-type: none">– If spectrogram smoothing fails, reverts to non-smoothed spectrogram.
Release 2010-01-15	<ul style="list-style-type: none">– Display of total number of sound files now updates when a file is deleted.
Release 2009-12-04	<ul style="list-style-type: none">– In batch mode in previous versions all the sound files and their spectrograms were held in memory, which could lead to insufficient-memory errors. This information is now written to temporary files in a temporary folder and accessed file by file.– Remove command now works correctly– Remove command when the whole sound is selected deletes the sound file and associated marker file
Release 2009-01-31	<ul style="list-style-type: none">– Beep if not requested number of markers now works correctly
Release 2008-12-19	<ul style="list-style-type: none">– Option of multiple rows of labels, number of rows is user selectable, multiple rows is backwards compatible and disagreement between exiting and specified numbers of rows is handled– Default options and key customisation ini files have been simplified– Added option to smooth spectrogram display– Does not crash if there is no zero crossing between the marker and the edge of the sound file
Release 2008-05-18	<ul style="list-style-type: none">– First release distributed for trial and feedback, please report bugs

This software is provided as-is without any guarantee that it will work. I'm willing to give some end-user support, but please read this document thoroughly and follow all the instructions before contacting me.

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 holder.

Contents

Purpose 3

Requirements 3

 Matlab code version: 3

 Compiled version: 3

Installing and Running 4

 Matlab code version: 4

 Compiled version: 4

Choosing Options 5

 File Options 5

 Calculation Options 6

 Other Options 6

 Spectrogram Options 7

Marking Sound Files 8

Mouse Clicks & Keystrokes 12

Output 13

Customisation 13

Purpose

This software displays waveforms and spectrograms of “.wav” sound files (PCM), and the user can place markers indicating the beginning and end of sections of the recording. Labels can also be associated with sections. The purpose of the software is therefore the same as the TextGrid function in Boersma & Weenink’s Praat Software (<http://www.praat.org>). My software differs from the Praat TextGrid function in that it has an ergonomically efficient design which makes it easier and faster to mark-up sound files: The right hand stays on the mouse, mouse clicks do not have to be particularly precise. The most frequent keystrokes are single keystrokes in the bottom left hand corner of the keyboard requiring minimum movement of the left hand. The user can customise the keystrokes and mouse buttons. The software also runs in Matlab which makes it easier for me to integrate the input and output with other Matlab software. This software works well for relatively short sound files but poorly for long sound files because it holds the whole file in memory – my data is typically either a series of sentence-length sound files or a long sound file which I have already cut up into utterance-sized chunks ([Sound File Cutter Upper](#)).

Requirements

Matlab code version:

- installed and licenced version of Matlab
- installed and licenced version of Matlab Signal Processing Toolbox
- Windows XP

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. You may have to make changes if you want to run it on a different version of Matlab or using a different operating system. I 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 my ftp site; however, you will need to ask me 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

Installing and Running

Matlab code version:

Unzip the files in the archive **SoundLabeller.zip** into a folder named **SoundLabeller**.

– To start the software, run **SoundLabeller_script.m** or **SoundLabeller_function.m** in Matlab.

Compiled version:

1. Install Matlab Compiler Runtime version 7.9

2. Place **SoundLabeller_pkg.exe** into a folder named **SoundLabeller**, and double click on **SoundLabeller_pkg.exe**.

– To start the software, double click on **SoundLabeller.exe**.

Choosing Options

- When you start **SoundLabeller** 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.

The screenshot shows the SoundLabeller application window with a blue title bar. The main area is gray and contains several option panels. At the top left is the title 'SoundLabeller'. Below it are three panels: 'File Options', 'Calculation Options', and 'Other Options'. To the right is a 'Spectrogram Options' panel. At the bottom right is a 'Next' button. In the bottom left corner, there is release information.

File Options

- ☒ Process all the files in a directory
- ☐ Process a single file
- File filter:

Calculation Options

- ☒ Precalculate the spectrogram before presenting the sound files for markup.
- ☐ Calculate spectrograms on demand.

Other Options

- Requested number of markers per sound file:
- Number of rows of markers:
- ☒ Allow input of text labels
- ☐ Review already-saved markers only

Spectrogram Options

- Nyquist frequency (Hz):
- Pre-emphasis:
- Window length (ms):
- Window shift (ms):
- Amplitude display floor (proportion):
- ☒ Smooth Spectrograms
- Colours:
 - Jet (blue-red)
 - Jet (red-blue)
 - Gray (light-dark)
 - Gray (dark-light)
 - Bone (light-dark)
 - Bone (dark-light)

Release 2010-11-01
© Geoffrey Stewart Morrison
<http://geoff-morrison.net>

Next

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 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 `*.wav`, i.e., all the wave files in the folder. If you entered `..\SoundFiles\h*t.wav` you would be processing all the files with names beginning with `h` and ending with `t` and having file extension `.wav` in the folder `SoundFiles` which is a sister folder to the `SoundLabeller` folder (`.\` specifies the same level in the folder hierarchy, `..\` up one level, `..\..\` up to levels, etc.). Make sure you specify `.wav` as the file extension. You can manually override the filter by later selecting a folder or file which does not match the folder and filter options specified here.

Calculation Options

- Choose `Precalculate the spectrograms` if you want the spectrograms to be calculated before any of the sound files are presented to you for marking. This is the preferred option if you have a number of short sound files which you wish to process in batch mode. Precalculating the spectrograms delays you being able to start processing the sound files but allows you to work faster once you have started.
- Choose `Calculate spectrograms on demand` if you want the sound files to be initially presented to you as a waveform only, and have the spectrogram presented only when you request it. This is the preferred option if you have a long sound file and you do not need to see the spectrogram of the whole sound file, but when you zoom in on a portion of the sound file you can request a spectrogram. This saves time overall by only calculating the spectrogram for the portion of interest.

Other Options

- `Requested number of markers per sound file:`

This is useful in if you need to put the same number of markers in each sound file. When you save the markers associated with a sound file (or move on to the next sound file or move back to the previous sound file), you will hear a beep if the number of permanent markers on the first row of markers does not match the number entered here. Enter 0 to disable the beeps.

- `Number of rows or markers:`

Specify the number of rows of markers you require, e.g., you could specify 2 and use the first row for segments and the second row for words. The first row will appear the highest in the display.

- `Allow input of text labels:`

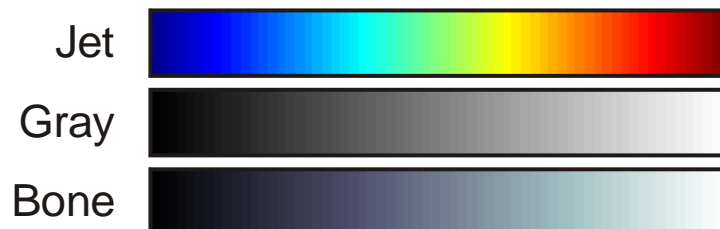
Not allowing input of text labels is useful if adding labels would be redundant, e.g., if the file name indicate the label for the segments which are to be marked in the sound file. Not allowing input of text labels prevents you from accidentally adding unnecessary information, and from accidentally typing a command in the text labelling space.

– **Review already-saved markers only:**

Check this option if you only want to work with sound files for which you have previously saved marker files. Other sound files will not be presented so you can review your previously saved work faster.

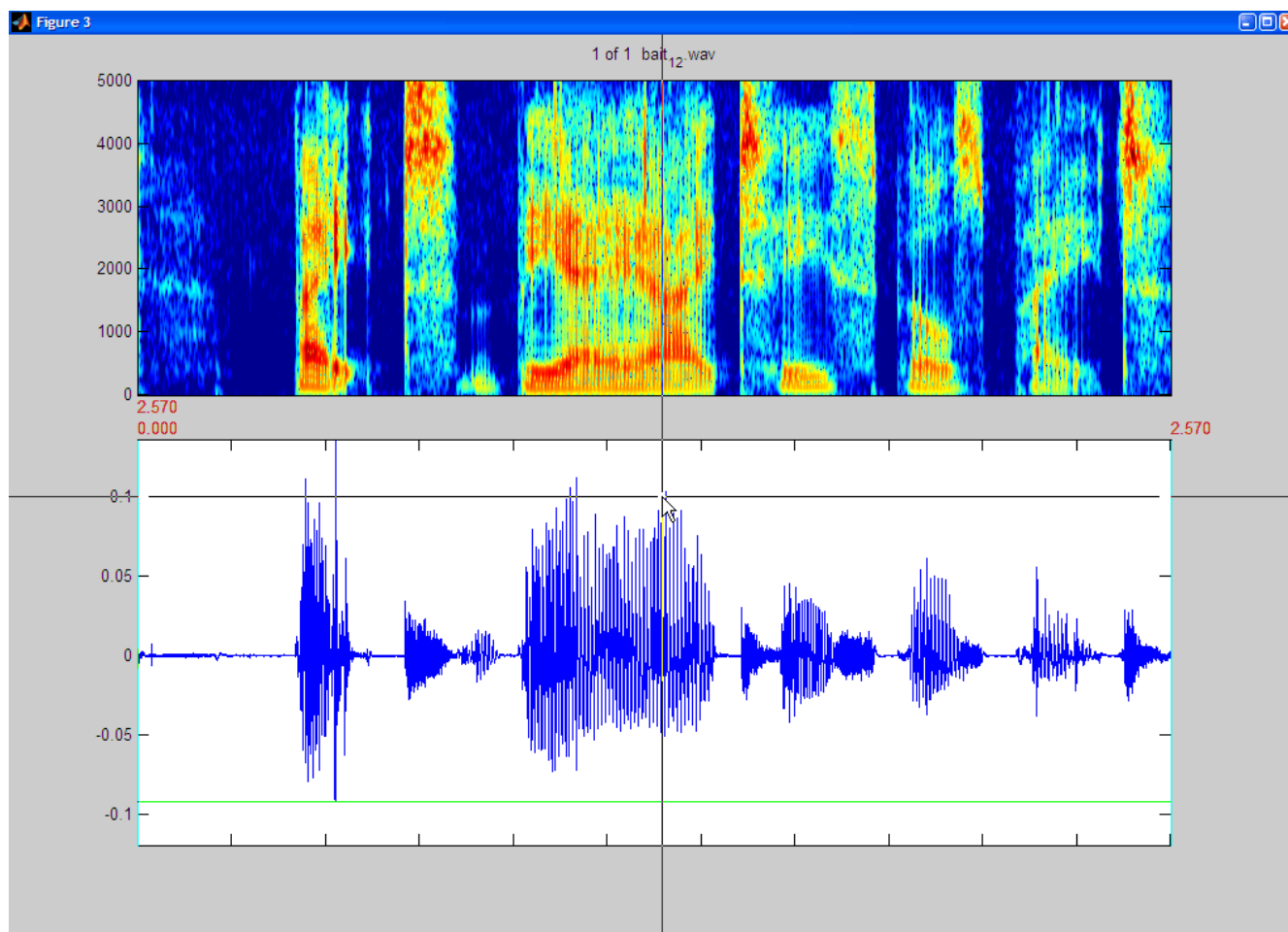
Spectrogram Options

- **Nyquist frequency (Hz):** The upper frequency limit of the spectrogram.
- **Pre-emphasis:** Adjustment to spectral tilt prior to spectrogram calculation. Larger numbers increase higher frequencies by a greater amount. Try values from -1 to $+1$.
- **Window length (ms):** The length of the window over which each time slice of the spectrogram is calculated. Use larger values for wide-band spectrograms and smaller values for narrow-band spectrograms. The window used is a power-four-cosine window.
- **Window shift (ms):** The spacing between each time slice.
- **Amplitude display floor (proportion):** We are usually only interested in the higher amplitude features in the spectrogram, so it is more efficient to use our range of colours to represent amplitude values in upper part of the amplitude range. This option specifies a proportion of the amplitude range - all amplitudes below this proportion are displayed as the same colour.
- **Smooth Spectrograms:** If this option is checked, then a cubic spline interpolation is applied to the displayed spectrogram. It may look prettier when you zoom in but will take more on-line processing time.
- **Colours:** This option allows you to select the colour scheme which will be used to represent amplitudes in the spectrogram.



Marking Sound Files

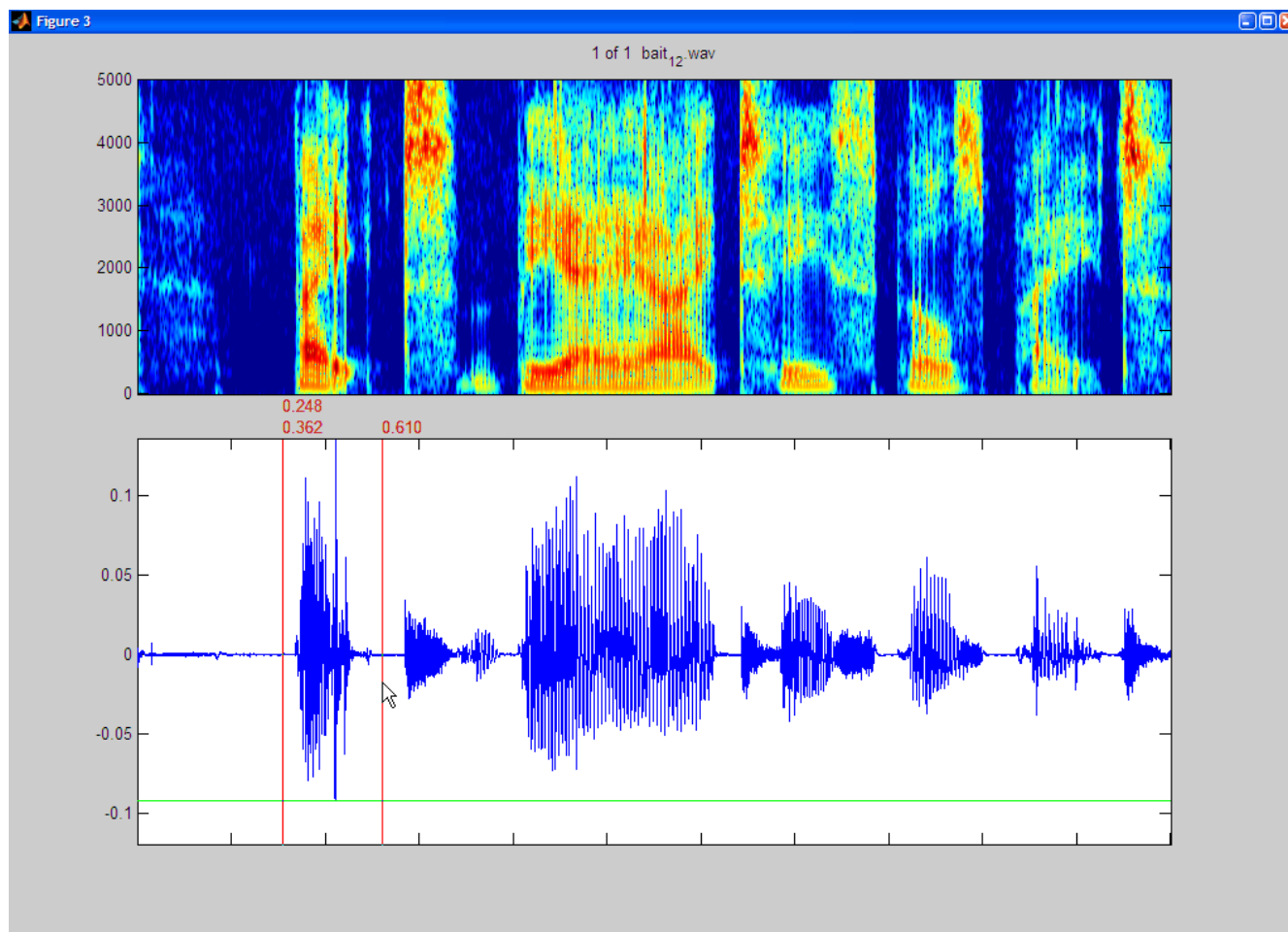
- A waveform, and if requested a spectrogram, will be displayed.



- You have two temporary markers, a left temporary marker and a right temporary marker.
 - To set the left temporary marker, click above the green line in the waveform window using the left mouse button.
 - To set the right temporary marker, click above the green line in the waveform window using the right mouse button.

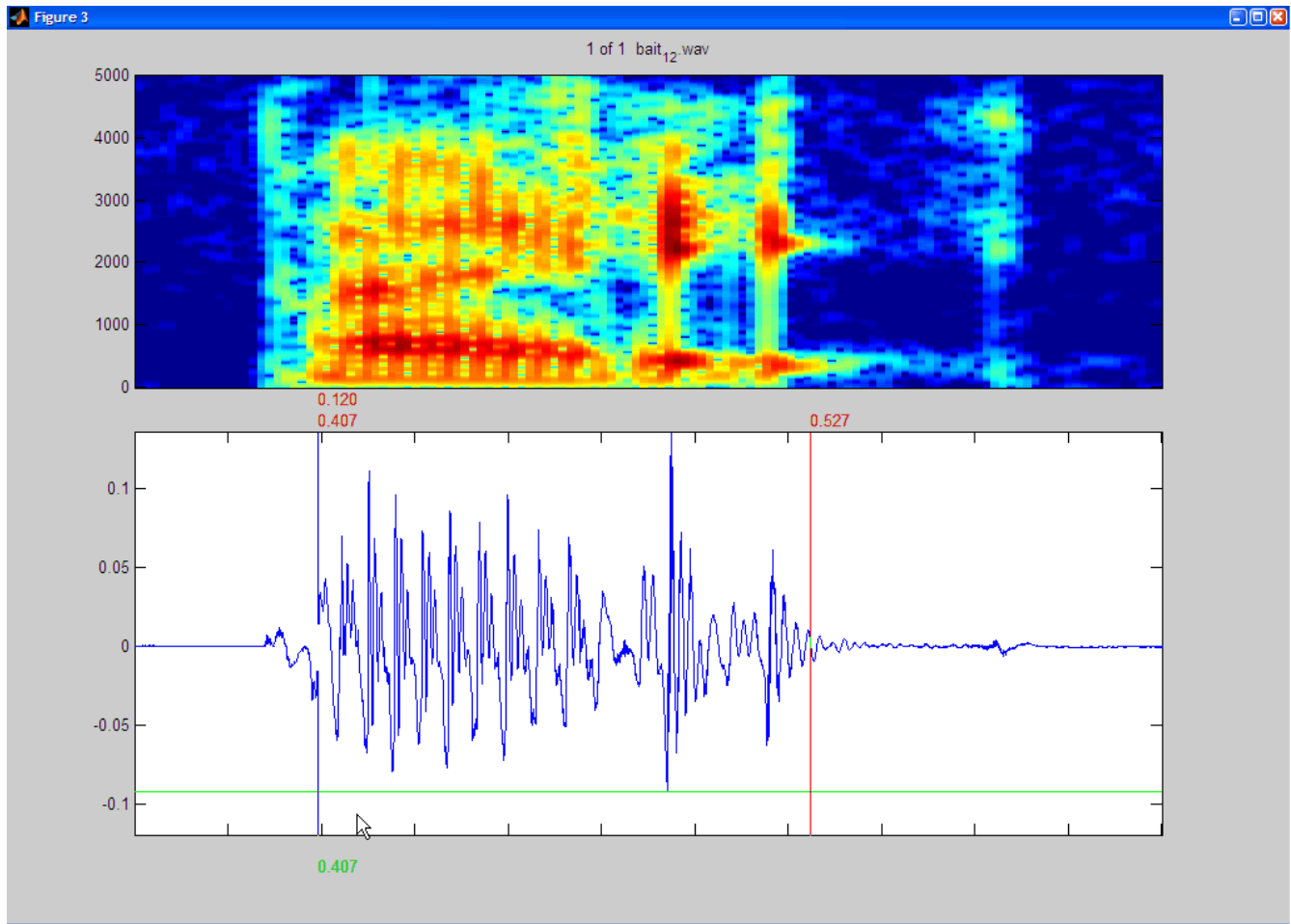
Note that you cannot place the left marker to the right of the right marker and vice versa. The use of the left and right mouse buttons ensures that there is never any doubt about which temporary marker you want to move and which one you want to leave where it is.

- The temporary markers appear as red lines.

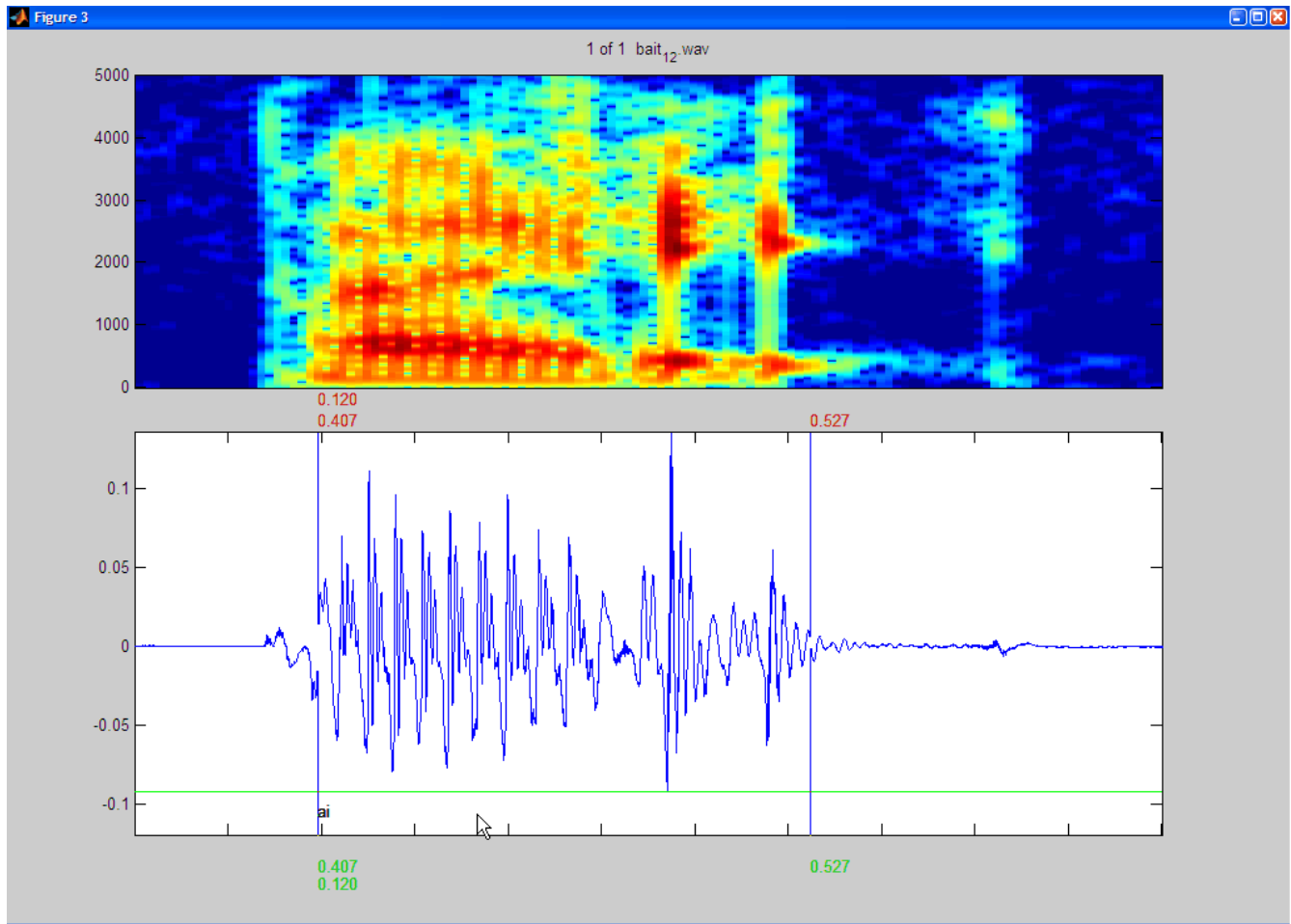


- Pressing the **spacebar** will cause the portion of the sound between the markers to be played. (Pressing **p** will play the whole of the visible waveform.)
- To zoom to the area bounded by the temporary markers, press the **z** key while the mouse cursor is above the green line.
- To move the temporary marker closest to the mouse cursor to the nearest zero crossing, press the **x** key while the mouse cursor is above the green line.
- To make a marker permanent, click below the green line using the primary (left) mouse button. The temporary marker which is nearest to the mouse cursor will become permanent and change colour. You

don't have to be precise - you just have to be below the green line and closer to the temporary marker you want to make permanent than to the other temporary marker.



- To delete a permanent marker, click below the green line using the secondary (right) mouse button.
- To add a text label, place the cursor between two permanent markers and below the green line, then type the text for the label. Phonetic fonts are not supported – the design philosophy is that a label should be distinct from any other label (it helps if it has a mnemonic value) and it should be quick and easy to type, i.e., by pressing one (or two) keys rather than some complex sequence of keystrokes.



- To zoom back out to the whole sound file, press **a** while the mouse cursor is above the green line.
- To save the permanent markers, press **s** while the mouse cursor is above the green line.
- To save the permanent markers and continue to the next sound file, press **c** while the mouse cursor is above the green line.
- To exit, press the **escape** key while the mouse cursor is above the green line.

Mouse Clicks & Keystrokes

The most frequent keystrokes are located in the bottom left hand corner of the keyboard so as to achieve maximum ergonomic efficiency for a right handed user.

When the mouse cursor is in the waveform area above the green line:

- left click - set left temporary marker
- right click - set right temporary marker
- x - move nearest temporary marker to zero crossing
- z - **z**oom to selection
- a - zoom to **a**ll
- q - zoom out 150%
- w - zoom in 50%
- f - move **f**orwards in time 50% of window
- d - move backwards in time 50% of window
- spacebar - play selection
- p - **p**lay visible waveform
- R - **r**emove (delete) selected portion of sound file (also deletes markers) [capital letter makes it difficult to press this key accidentally]. Will delete the entire sound file if the entire file is selected. This action does not become permanent until a save command (**s** or **c**) is issued.
- T - **t**zero out selected portion of sound [capital letter makes it difficult to press this key accidentally]. This action does not become permanent until a save command (**s** or **c**) is issued.
- u - **u**ndo last remove / zero out [must be executed prior to save command]
- s - **s**ave label file (& sound file if changes have been made)
- c - save and **c**ontinue to next sound file
- right arrow - continue to next sound file without saving
- left arrow - back to previous sound file without saving
- g - draw spectro**g**ram
- o - spectrogram **o**ptions
- O - spectrogram **o**ptions with all values reset to default
- escape key - quit

When the mouse cursor is in the waveform area below the green line:

- left click - set permanent marker at nearest temporary marker

- right click** - remove nearest permanent marker
- space** - play marked section
- other** - enters text for marked section (don't type too quickly)

Output

The output of the **SoundLabeller** is a series of **.mat** files with the same filenames as the sound files but with the file extension changed from **.wav** to **.mat**. These are saved in the same folder as the sound files.

The files contain three variables:

- **mark_fs** is a cell array with one cell per row, each cell contains a vector containing the locations of the permanent markers expressed as the nearest time frame in the original sound file
- **mark_time** is a cell array with one cell per row, each cell contains a vector containing the locations of the permanent markers expressed as the time in seconds of the nearest time frame in the original sound file
- **mark_text** is a cell array with one cell per row, each cell contains a cell array containing the labels associated with the time markers. The first cell corresponds to the section beginning with the first marker.

A temporary folder is created while processing the files and is automatically deleted on normal termination, e.g., by pressing the **escape key**. Should this not happen, the folder and its contents can be manually deleted.

Customisation

- If you want to change the default options you can change the file **SoundLabeller.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.
- 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 my 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. 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`.