MATLAB Apps for Teaching Digital Speech Processing

Lawrence Rabiner, Rutgers University

Ronald Schafer, Stanford University

GUI LITE 2.5 editor written by Maria d'Souza and Dan Litvin

MATLAB coding support from Ye Cheng, Siva Yedithi and Kirty Vedula

Project Goals

- Design and implement a LITE GUI Design and Editing Program with a small but useful set of GUI Objects
- Build a set of MATLAB Apps in support of teaching/learning Digital Speech Processing concepts
- Make GUI LITE Design/Editor and Apps broadly available via MATLAB Central (using the File Exchange program)

Methods of Learning Digital Speech Processing

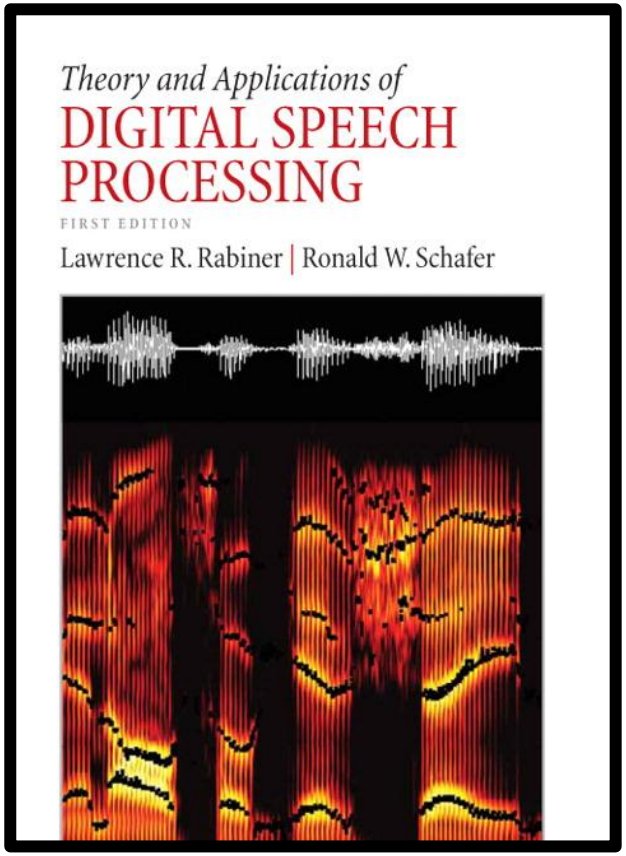
Learn by Teaching

Writing Textbooks

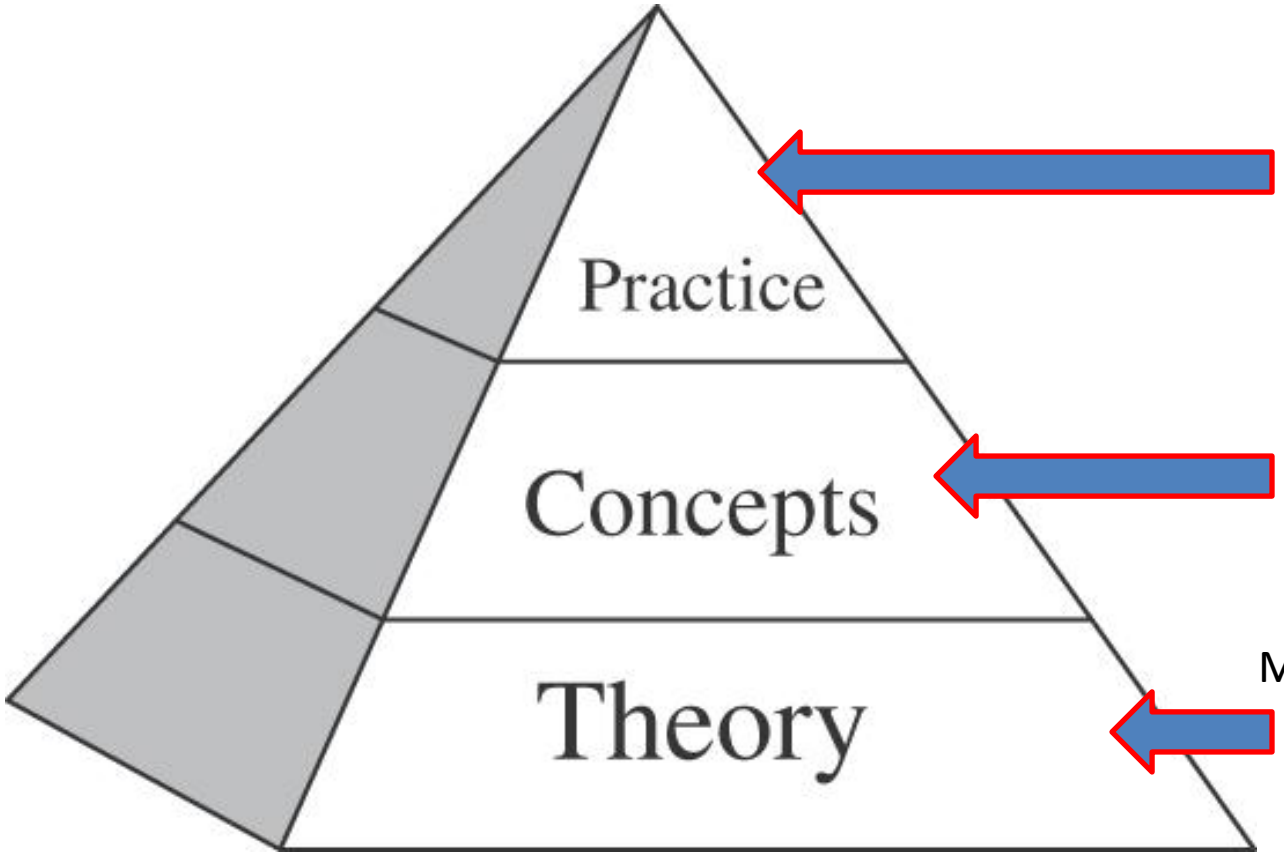
Learn by Doing

Homework Problems

MATLAB Apps



Pyramid of Learning



Ability to implement theory and concepts in working code (MATLAB, C, C++); algorithms, applications

Basic understanding of how theory is applied; autocorrelation, waveform coding, …

Mathematics, derivations, signal processing; e.g., STFT, cepstrum, LPC, …

Need to understand speech processing at all three levels of the pyramid

The Speech Stack

Speech Applications **— coding, synthesis, recognition, understanding, verification, language translation, speed-up/slow-down**
Speech Algorithms **— speech-silence (background), voiced-unvoiced, pitch detection, formant estimation**
Speech Representations **— temporal, spectral, homomorphic, Linear Prediction Coding**
Fundamentals **— acoustics, linguistics, pragmatics, speech production/perception**

Basics – read/write speech/audio files; display speech files; play files

Speech Apps Goals

- give students and instructors **hands-on experience and guidance** with Graphical User Interface-based Apps that cover a large range of speech processing capability

- align the Apps with **deeper explanations of the theory and concepts of each App** from the textbook *Theory and Applications of Digital Speech Processing* (*TADSP*), L. Rabiner and R. Schafer

Elements of Exercises

- The complete set of MATLAB Speech Processing Apps is made available to students and instructors via MATLAB Central, File Exchange, on the MathWorks website, including:
  - all the code that is required to run the complete set of Speech Processing Apps
  - an extensive set of speech and audio files for processing
  - a scripted run with associated graphical output which can be used to validate each individual App
  - one or more results figures generated by the scripted run using the current version of the code
  - the User's Guide for each Speech Processing App
  - the code for the MATLAB GUI Lite 2.5 Design and Editing Program (to be made available later this year)

Graphical User Interface (GUI) Components

- GUI Lite 2.5 (created by students at Rutgers University) simplifies the process of creating viable GUIs for a wide range of speech processing Apps
- GUI Lite Elements
  - basic design tool and editor (GUI Lite 2.5); (runGUI.m)
  - panels; used to block group of buttons/graphical panels/etc., into one or more coherent blocks
  - graphics panels; used to display one or more graphical outputs (figures)
  - text blocks; used to display global information about the specific speech processing App
  - buttons; used to get and set (vary) app parameters; used to display a list of options; used to initiate actions within the code
    - editable buttons – get and/or set parameter value
    - text buttons – display variable values
    - slider buttons – display variable range
    - popupmenu buttons – display list of variable options (e.g., list of speech files)
    - pushbuttons – initiate actions within the code

GUI LITE 2.5 Design Process

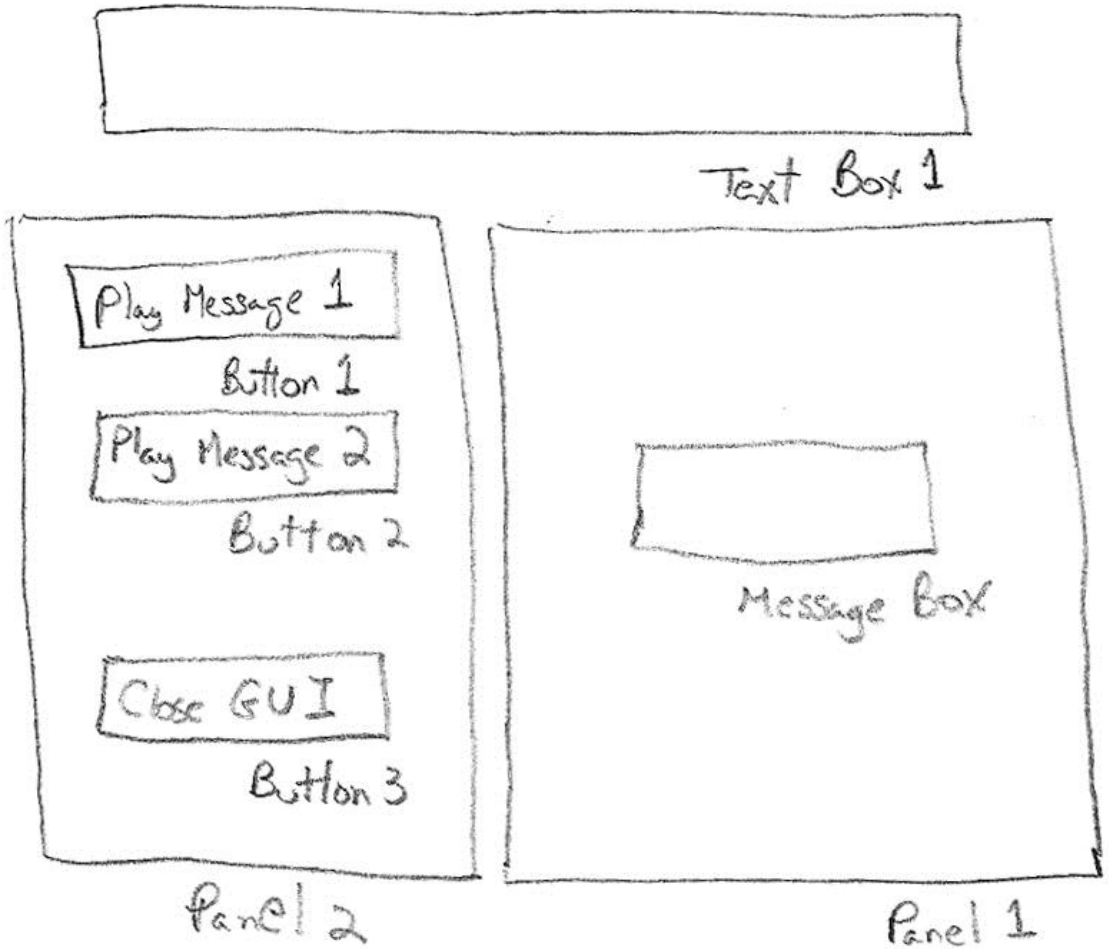
- begin with a rough (hand-drawn) sketch of the desired GUI look, as segmented into button panels, graphics panels, text boxes, and different button types

- run the program 'runGUI.m' to either create the desired set of GUI elements and save the results of the design as a GUI file, or to edit an existing GUI file

- edit the two programs created by GUI LITE 2.5 (in the design mode)
  - rename GUI main program from 'EditrunGUI.m' to 'app_GUI25.m'
  - rename GUI Callbacks program from 'PanelandButtonCallbacks.m' to 'Callbacks_app_GUI25.m'

- run the resulting App and loop on the GUI design and Callbacks implementation until the look and feel of the App meets the design specs

Example: Hello/Goodbye World Plan



Design Specs:

- 2 Panels (for linking inputs and outputs)
- 1 Text Box (for describing the Exercise GUI)
- 3 Buttons (all pushbuttons) (for embedding Callback code to play two messages and to close up the GUI)

GUI LITE 2.5 Initial Screen

hello_goodbye_world

runGUI.m

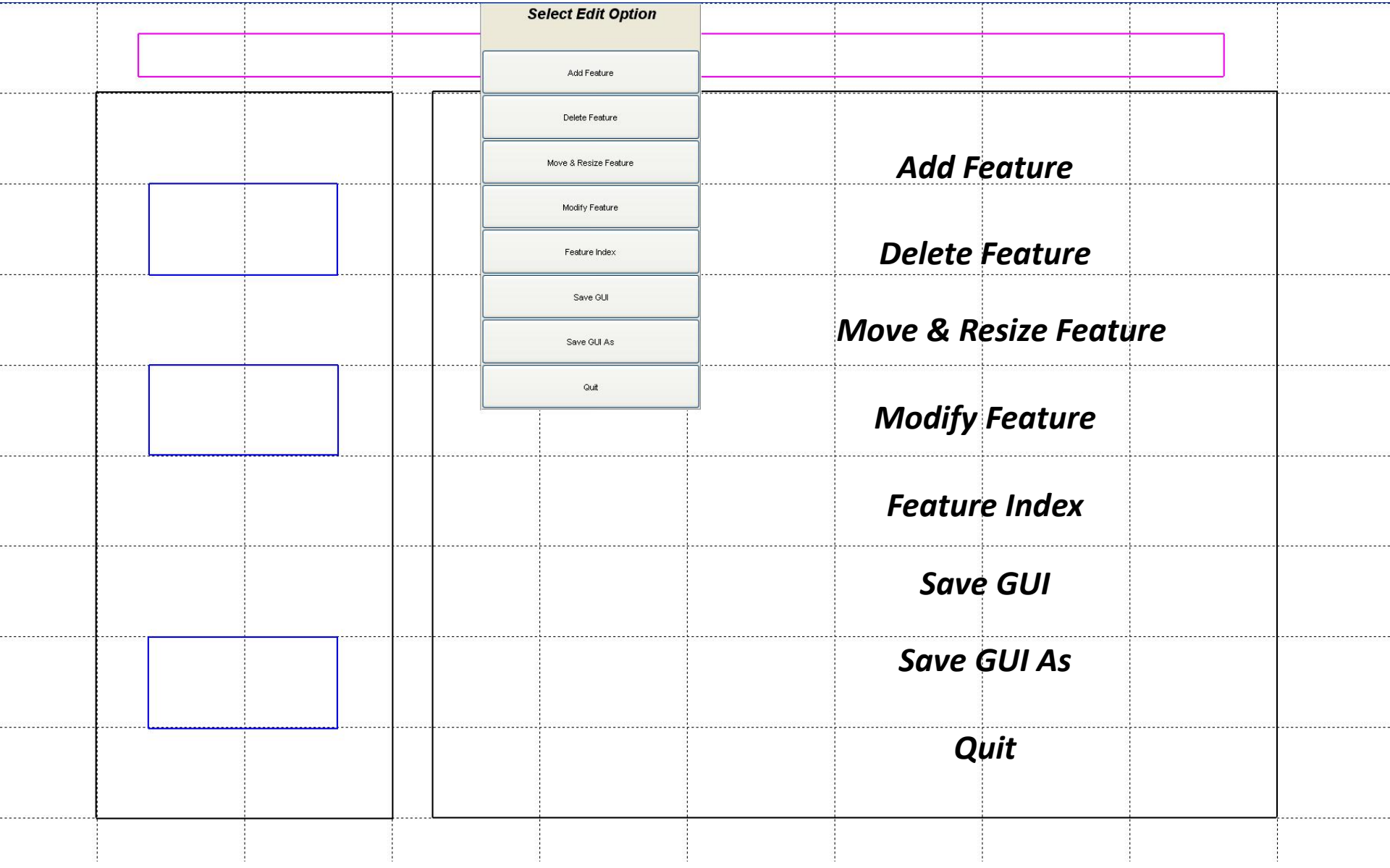
run:

hello_goodbye_world_GUI25.m

directory:

hello_goodbye_world_gui25

GUI LITE 2.5 Edit Screen



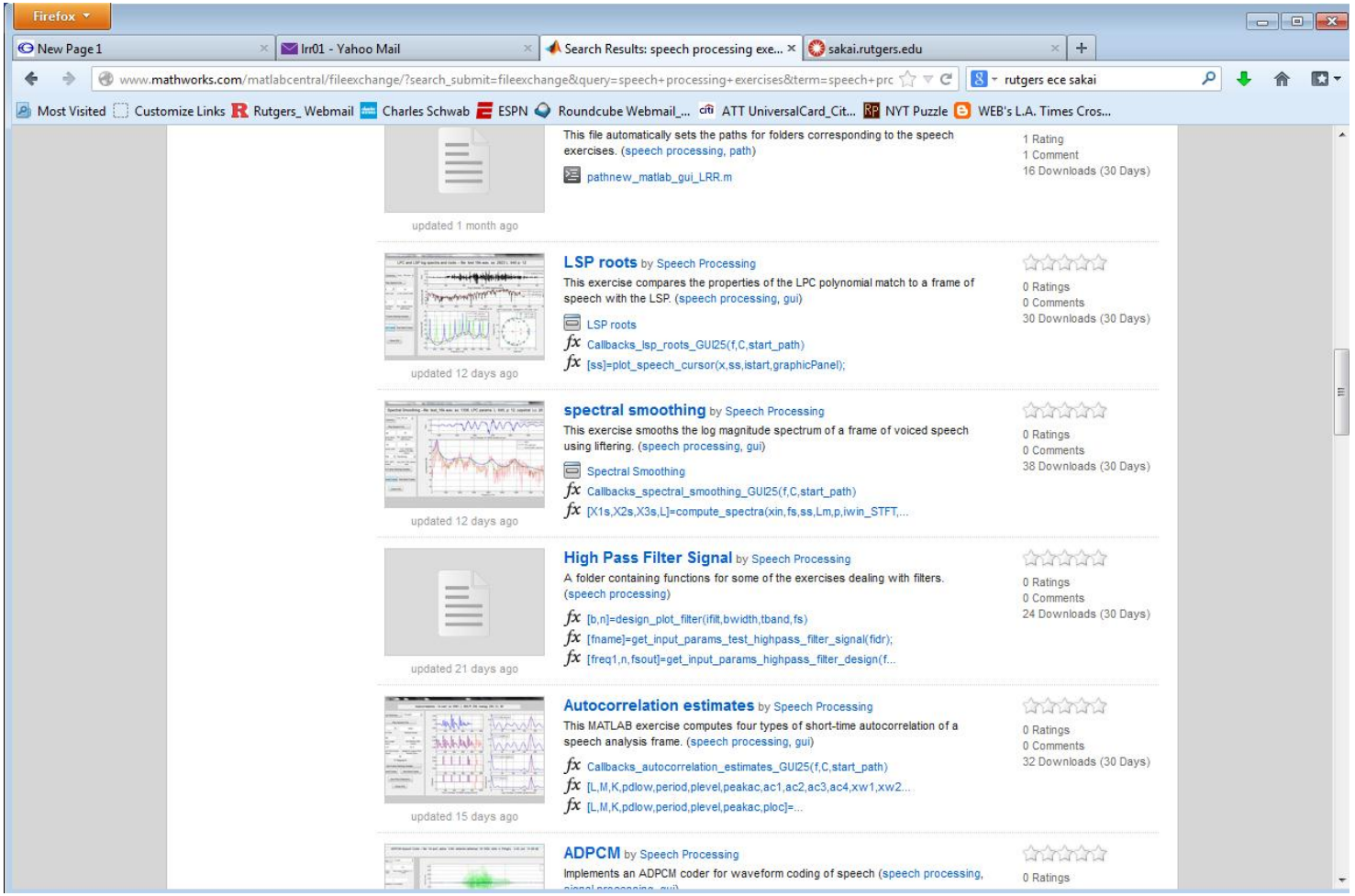
GUI LITE 2.5 Features

- provides a versatile design program for creating LITE GUIs for a range of speech processing Apps
- separates GUI design and editing from Callbacks for each App
- provides a versatile editor for modifying GUI elements without impacting the Callback actions
- provides a GUI element indexing feature that enables the user to identify GUI elements with the appropriate Callback elements
- significantly easier to learn and use than MATLAB GUIDE (Graphical User Interface Design Environment)

Download Each MATLAB Speech Processing App

http://www.mathworks.com/matlabcentral/fileexchange/index?term=authorid%3A373534

Apps downloaded from MATLAB File Exchange;
Search Term 'speech processing exercises'



MATLAB App Categories

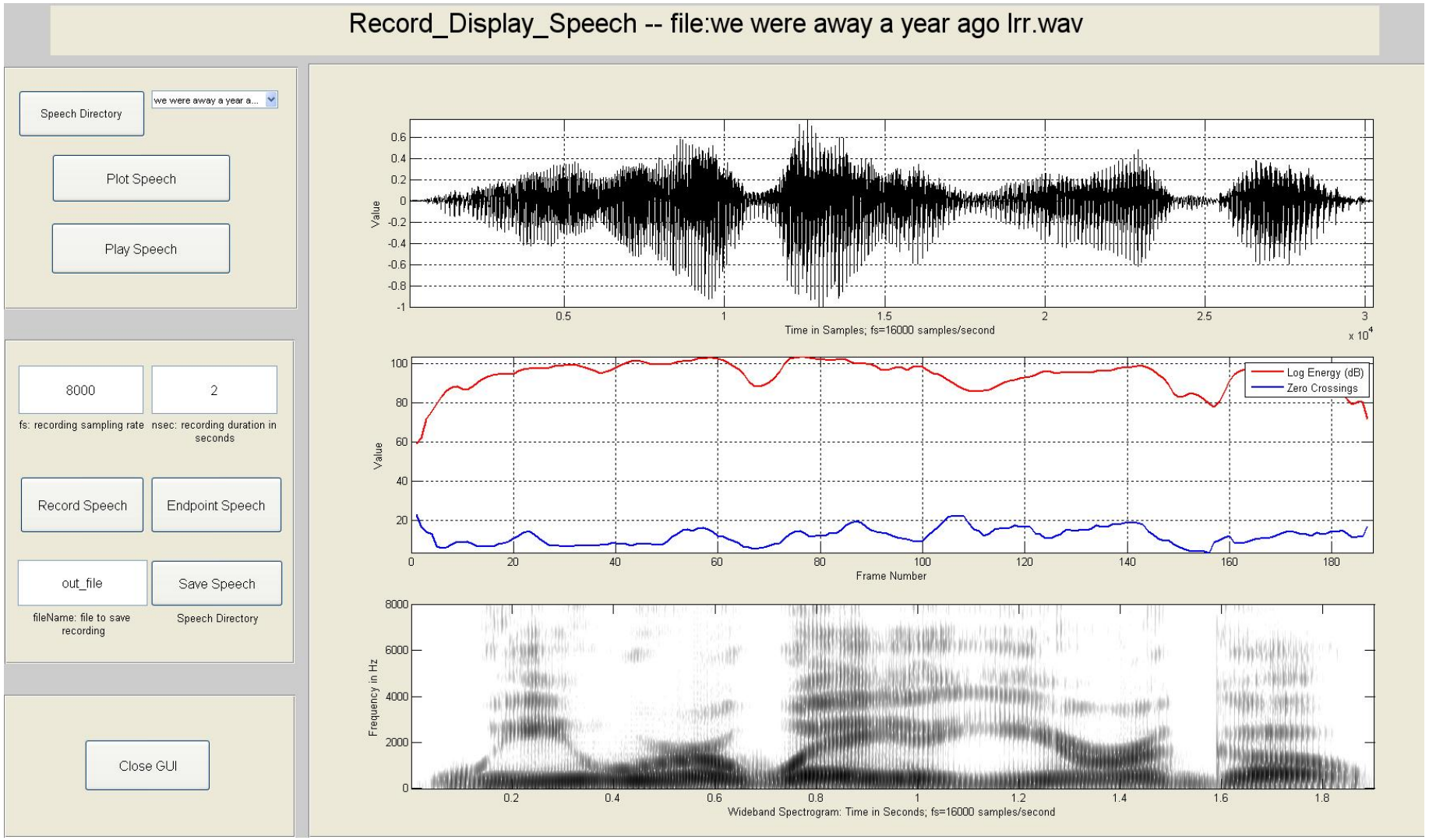
- The speech processing Apps are grouped into the 5 (extended) layers of the Speech Stack, namely:
  - **Basics** of speech processing using MATLAB (5)
  - **Fundamentals** of speech processing (6)
  - **Representations** of speech in time, frequency, cepstrum and linear prediction domains (22)
  - **Algorithms** for speech processing (8)
  - **Applications** of speech processing (17)

Basics Functionality

- read/write a speech file
- play/record a MATLAB array
- plot a speech file
- convert speech file sampling rate
- lowpass/bandpass/highpass filter a speech file
- plot a frame of speech
- plot a spectrogram of a speech file

Record/Display Speech

Basics



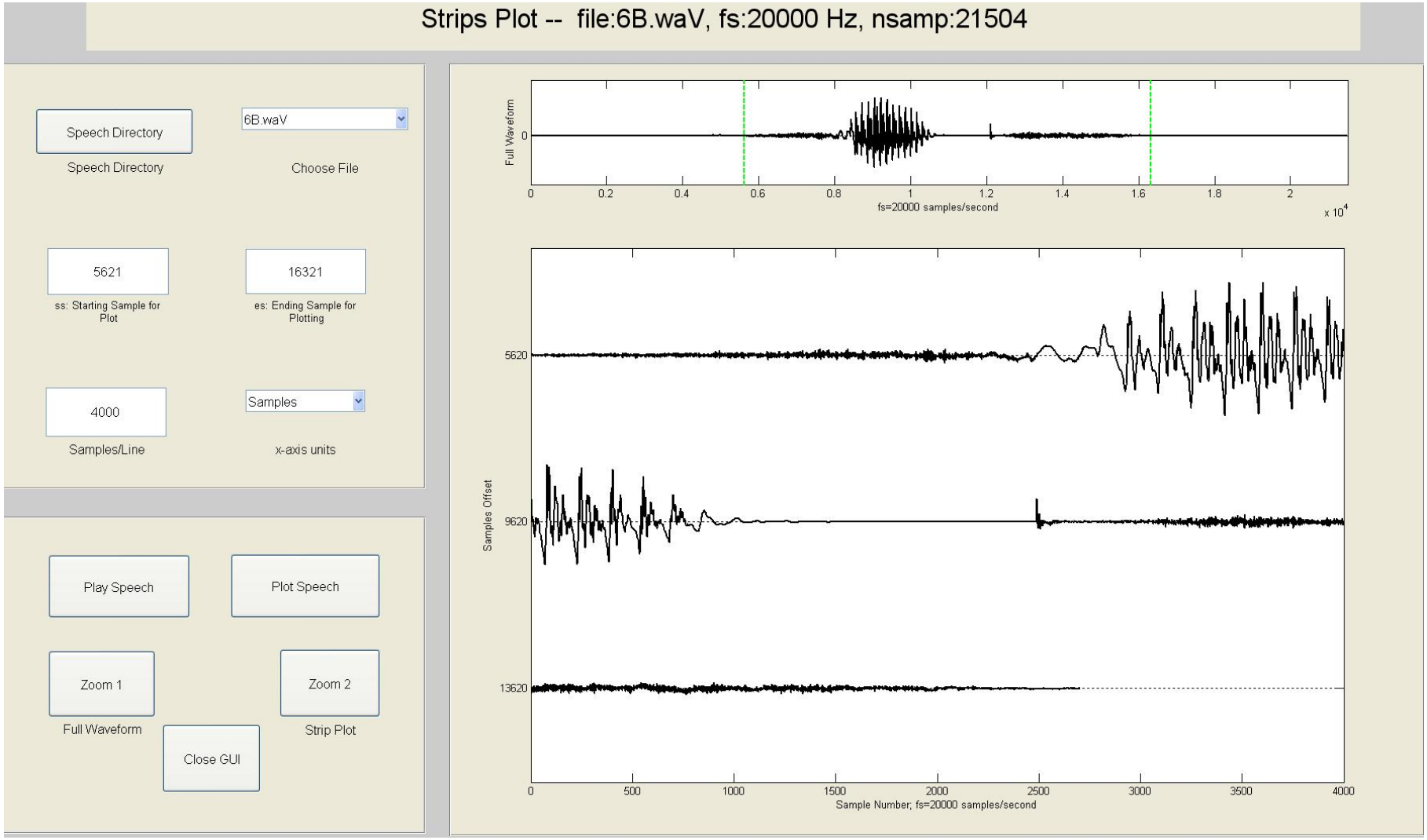
record_display_speech

Waveform Zoom Strips Plot

- Plotting and examining speech/audio waveforms is one of the most useful ways of understanding the properties of speech and audio signals.
- This MATLAB App displays a speech/audio waveform as a single line array and as a running plot of samples (called a Strips Plot).
- Zoom feature allows the user to select a region of signal for display and playback.

Waveform Zoom Strips Plot

Basics



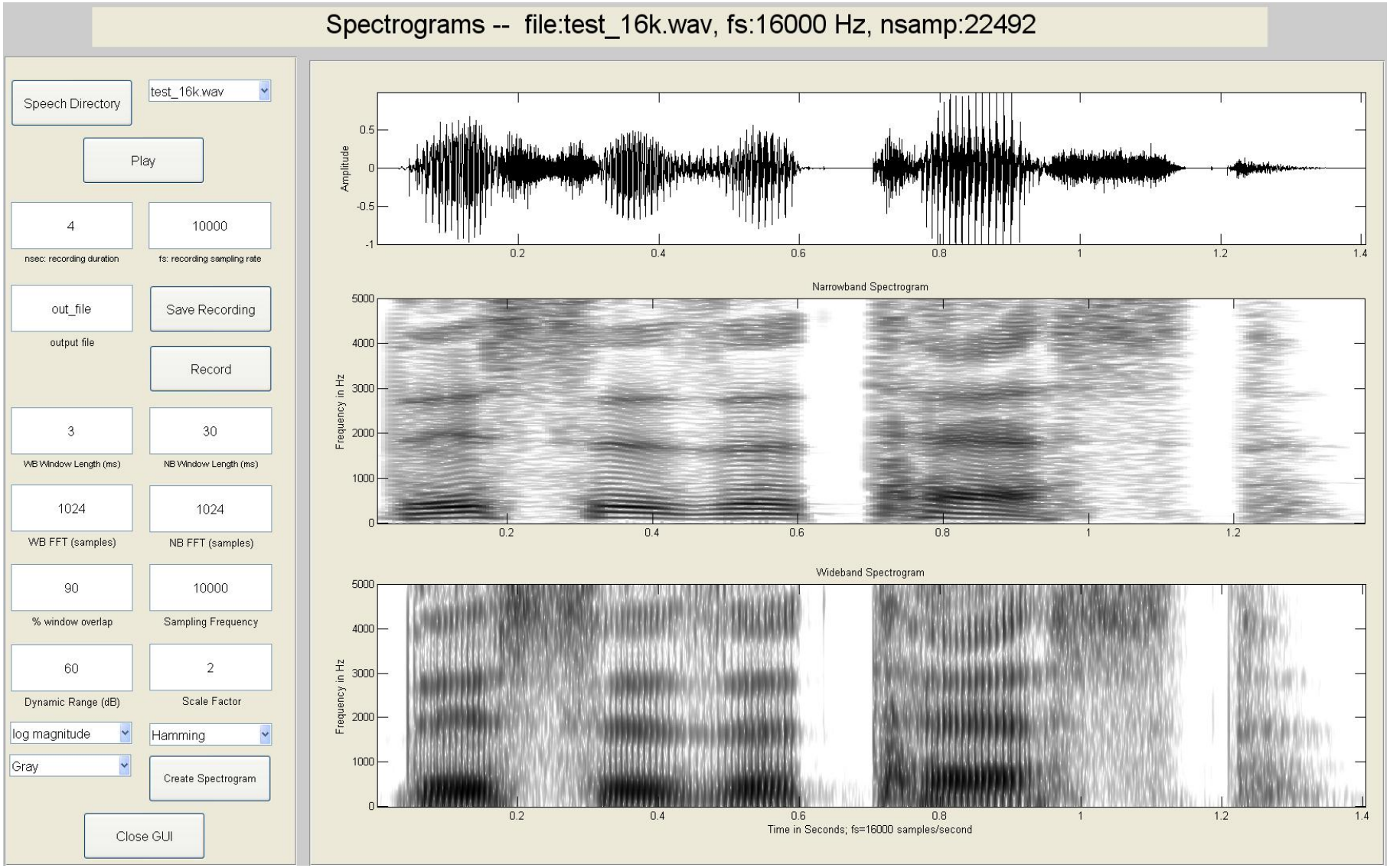
strips_plot

Wideband/Narrowband Spectrogram

- The sound spectrogram is an image map of the sequence of short-time log spectrums of the frames of a speech or audio signal.

- Each spectrum is obtained from a short-time Fourier transform (STFT) analysis of a frame of speech.

Wideband/Narrowband Spectrogram



Spectrogram

Fundamentals (6 Apps)

- 2-tube vocal tract model
- 3-tube vocal tract model
- p-tube vocal tract model
- glottal pulse model and spectrum
- composite vocal tract model and spectrum
- ideal vocal tract model and spectrum

No App demos from this layer of the Speech Stack

Representations (22 Apps)

- **time domain Apps**
  - windows; features; autocorrelation estimates; amdf
- **frequency domain Apps**
  - phase/magnitude; overlap-add windows; WSOLA
- **cepstral domain Apps**
  - analytical cepstrum; single pole cepstrum; FIR sequence cepstrums; cepstrum aliasing; cepstrum liftering; cepstral waterfall
- **linear prediction Apps**
  - LPC frames; LPC error; LPC varying p; LPC varying L; LSP roots; plot roots

Waveform Similarity and Overlap Add (WSOLA)

- MATLAB App for high quality, time scale modification of speech

- This App can speed-up or slow-down speech or audio by a factor of about 3

WSOLA

Algorithms (8 Apps)

- endpoint detector
- Voiced-Unvoiced_Background estimation method
- autocorrelation pitch detector
- log harmonic spectral waterfall plots
- cepstral pitch detector
- SIFT pitch detector
- formant estimation method

No App demos from this layer of the Speech Stack

Applications – Part 1 (11 Apps)

- speech waveform coding;
  - statistical properties of speech; quantization characteristics of a B-bit uniform or mu-law compressed and quantized speech file; uniform quantization; mu-law compression; mu-law quantization; Signal-to-Noise Ratio (SNR) of uniform and mu-law quantizers
- Automatic Gain Control (AGC)
- Adaptive Differential Pulse Code Modulation (ADPCM) waveform speech coder
- Vector Quantizer (VQ); VQ Cells
- synthetic vowel synthesizer

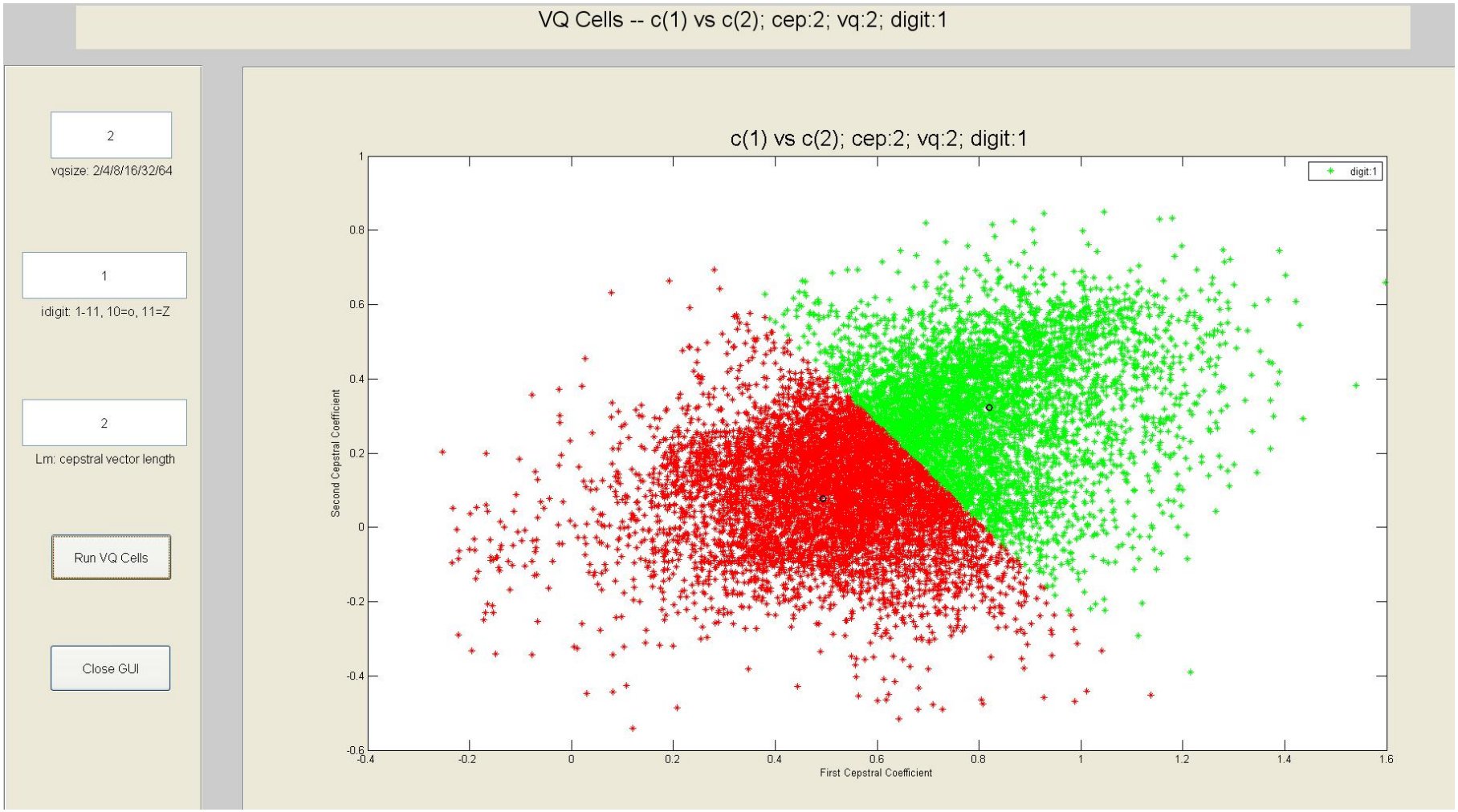
Applications – Part 2 (6 Apps)

- LPC error synthesis
- LPC vocoder
- play pitch period contour
- two-band subband coder
- phase vocoder
- isolated, speaker-trained, digit recognizer

Vector Quantization (VQ) Cells

- App based on VQ codebook designed using 2-element cepstral vectors from a library of isolated digits from a large number of talkers
- VQ codebooks designed for each of 11 digits (0-9 plus oh), and for VQ sizes of 2/4/8/16/32/64 cells
- The App plots the Voronoi cells and centroids of the training set vectors

Vector Quantization (VQ) - 2 Cells



VQ_Cells

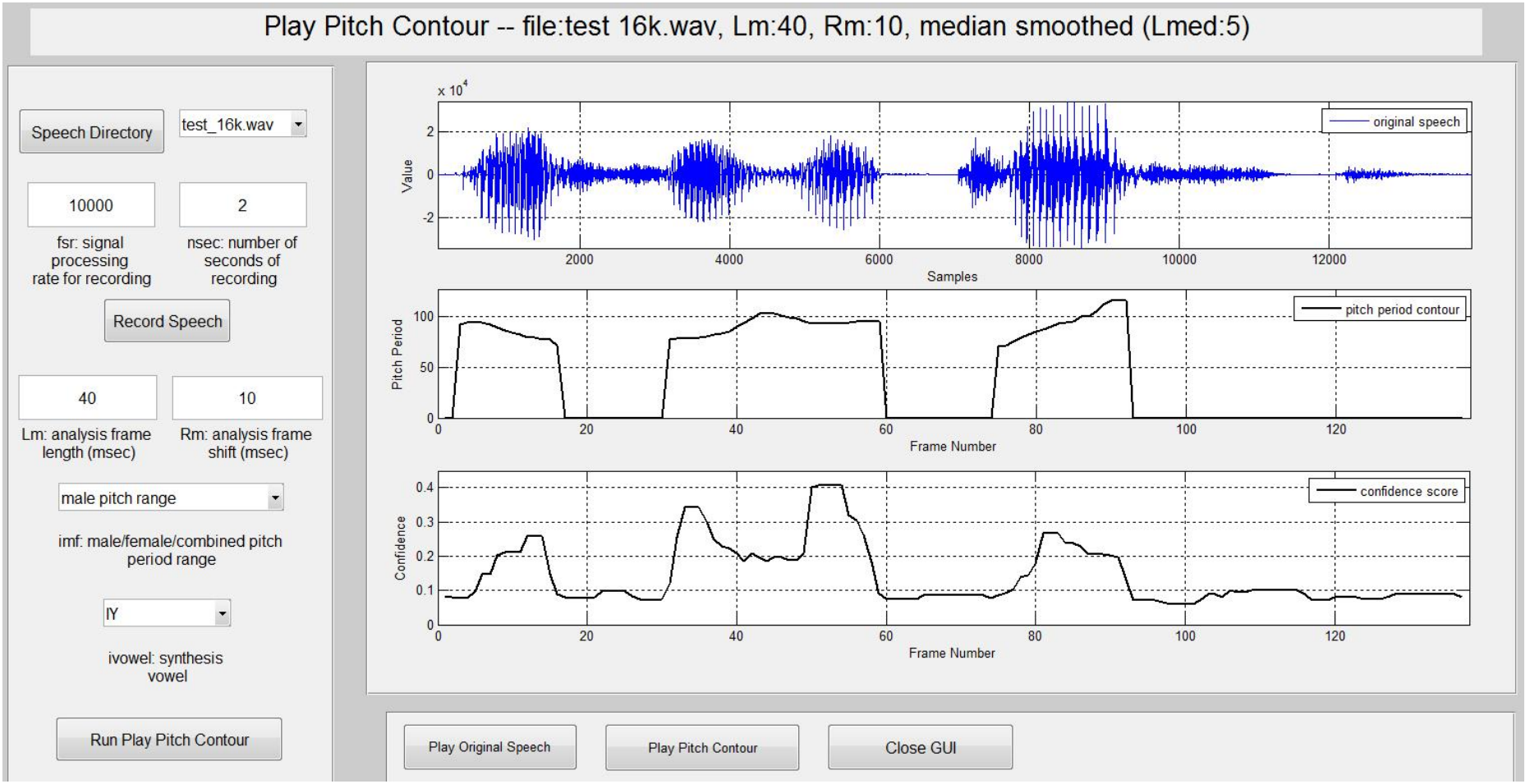
Play Pitch Period Contour

- This App computes the pitch period and gain contours of a speech utterance and synthesizes a designated vowel sound with the same pitch and intensity as the original speech signal, but with resonances appropriate for a designated vowel sound

- The pitch period and gain contours of the spoken utterance are estimated using any conventional pitch period detection method

- The two-state excitation sequence (pitch pulses or random noise) is then convolved with the impulse response of the designated vowel

Play Pitch Period Contour



play_pitch

Summary

- Set of 58 MATLAB speech processing Apps
- Apps aligned with distinct sections in the textbook *Theory and Applications of Digital Speech Processing* (*TADSP*) by Rabiner/Schafer, but also can be used with other speech processing texts
- Each App has an associated LITE Graphical User Interface created using a GUI LITE Design and Editing program and created expressly for these speech processing Apps
- GUI LITE design and implementation Callbacks are totally separated in different m-files
- GUI LITE editor allows the user to modify GUI features without needing to make changes to Callbacks files and vice versa