This user guide accompanies the "HSpot_Detect_GUI" software that accompanies the paper "Filter-Based Methodology for the Location of Hot Spots in Proteins and Exons in DNA" by Ramachandran et al., "IEEE Transactions on Biomedical Engineering", Volume: 59, Issue: 6, June 2012.

General Description

The software implements the filter-based hot-spot location methodology described in the above paper with the aid of user-friendly graphical user interfaces (GUIs). The main GUI, namely, "HSpot Detect GUI", can be used to design and apply an invese-Chebyshev filter and an opti-mized bandpass notch (BPN) filter. The sub-GUI, namely, "HSpot FSlider GUI", can be used for manually tuning the BPN filter, if desired. By default, however, the filter is tuned using the au-tomatic tuning strategy discussed in the paper.

Contents

The package consists of two m-files and two fig files namely

```
HSpot_Detect_GUI.m,
HSpot_FSlider_GUI.m,
HSpot_Detect_GUI.fig, and
HSpot_FSlider_GUI.fig,
```

this user guide, and an example protein FASTA file containing a number of protein sequences.

The fig files contain the GUI layouts, and the m-files contain the callback functions along with all the other necessary functions.

Note: For the GUIs to work, it is mandatory for a GUI m-file and its matching fig file to have identical names. They differ only by filename extension.

Minimum Requirements

The software has been written and tested on the Microsoft Windows platform. Versions of Windows above XP are compatible.

At the minimum, the software requires Matlab as well as the Signal Processing and the Optimization toolboxes to execute.

Compatibility

The software would be compatible with any of the versions of Matlab and the toolboxes released since 2009.

Input

The protein sequences to be analyzed must be grouped according to functionality, and all sequences belonging to a particular functional group must be stored in a single text file in FASTA format. Please see accompanying test data for exact formatting.

GUI Component Descriptions

Main GUI

File Listbox: This listbox appears at the top left corner of the GUI and displays the contents of the current directory. The items in this box are selectable, although multiple selections are not permitted. If the selected item is a directory, then double-clicking on it opens it. The box can be used to navigate to the required directory containing the FASTA file(s) to be processed. The name and the full path of the current directory are displayed below the box.

      The two buttons just below the file listbox can be used to toggle between displaying only text files and displaying all files.

"Load FASTA" Pushbutton: After an input file in the required format has been selected in the file listbox, a press of this button scans the input file for protein IDs and displays all the IDs found in the listbox titled "Protein IDs Found". Multiple selections are permitted in this listbox. These can be made in the standard manner either by using Ctrl+Click operations or Select+Drag operations.

"Read Seq." Pushbutton: After one or more protein IDs have been selected in the "Protein IDs Found" listbox, a press of this button loads the sequences corresponding to the selected IDs into the working memory.

"To EIIP" Pushbutton: A press of this button converts the current protein character sequences in the working memory into EIIP numerical sequences and stores them in the working memory.

The corresponding protein IDs are then displayed in the listbox "EIIP seqs. in memory". Multiple selections are permitted in this listbox.

"Compute Consensus Spectrum" Pushbutton: A press of this button computes the consensus spectrum corresponding to the group of protein sequences whose IDs have been currently selected in the "EIIP seqs. in memory" listbox. The spectrum is displayed in a new figure window, and, in addition, the characteristic frequency is automatically identified and displayed in the figure title. Furthermore, the default values for the filter specifications are automatically computed and the corresponding text fields are filled in.

"Filter Specifications/Parameters" Panel: This panel contains the text fields corresponding to the specifications for the two types of filters. Default values for these fields are automatically assigned after computing the consensus spectrum. However, if desired, these values can be changed by editing the appropriate fields.

"Design Filter" Pushbutton: A press of this button designs the corresponding filter taking the current values for the specifications.

"Filter Info" Pushbutton: A press of this button prints a report in the command window with the details of the corresponding filter that is currently in the working memory.

"Freq. response" Pushbutton: A press of this button plots the frequency response of the corresponding filter that is currently in the working memory.

"Predict" Pushbutton: A press of this button predicts the hot-spot locations for the currently selected proteins in the "EIIP seqs. in memory" listbox. For each of the selected proteins, the power of the filtered signal is displayed in a separate figure window. The peak locations of this signal denote hot spots, and these are displayed in the command window in descending order of peak magnitudes.

"About the Parameters" Pushbutton: This button is specific to the optimized BPN filter technique. When pressed, an information box opens with details about the parameters specific to this filter.

"No. of passes" text field: This field represents the number of times the input EIIP sequence is filtered through the chosen filter.

"More …" Pushbutton: A press of this button displays an information box describing the "No. of passes" text field.

"Predict & AutoTune" Pushbutton: A press of this button predicts the hot-spot locations for the selected protein sequences after tuning the BPN filter using the automatic procedure described in the paper. The hot-spot locations are displayed in the command window in descending order of peak magnitudes.

"FSliderGUI" Pushbutton: When one or more predictions are made using the "Predict" pushbutton corresponding to the BPN filter, and the resulting figure windows are kept open, then a press of the "FSliderGUI" pushbutton invokes the secondary GUI containing a slider for manually tuning the BPN filter. This also opens a separate figure window showing the frequency response of the filter.

Note: The FSliderGUI can be invoked only by predicting without auto-tuning, i.e., by using the "Predict" pushbutton. In effect, it is associated only with manual tuning.

FSlider GUI

Slider: This can be used to change the notch frequency of the BPN filter, and thereby perform manual tuning of the filter. By default, the slider is positioned at the characteristic frequency. As the slider is moved either by "click and drag" operations or by using the arrows at the edges, the coefficients of the BPN filter are adjusted to reflect the change in the notch frequency. This is carried out using the model described in the paper. Simultaneously, the hot-spot locations are recomputed and the energy curves in the open figures are updated. Users can thus tune the BPN filter and observe the results in real time.

Tip: For best viewing, position the figure windows and the slider such that they do not overlap.

"Print current hot-spot locations" pushbutton: After positioning the slider at a desired frequency, a press of this button will print the current hot-spot locations in the command window in descending order of peak magnitudes.

Typical Operating Sequence

A typical step-by-step operating sequence of the GUIs would be as follows:

1) After downloading and unzipping, it is necessary to ensure that all the unzipped Matlab files are located in a single directory. Set Matlab's working directory to this directory.

2) Run the m-file "HSpot_Detect_GUI". This will invoke the main GUI.

3) Navigate to and select the text file containing the desired FASTA-formatted protein sequences using the "Current Directory" listbox. All sequences belonging to a given functional group must be present in a single file.

4) Press "Load FASTA". This will load all the protein IDs found in the selected text file to the "Protein IDs Found" listbox.

5) Select one or more protein IDs from the listbox, and press "Read Seq.". This will load the character sequences into memory.

6) Select one of more protein IDs again from the "Protein IDs Found" listbox, and press "To EIIP". This will convert the selected protein sequences into EIIP numerical sequences and store in memory.

7) Select one or more protein IDs from the "EIIP seqs. in memory" listbox, and press "Compute Consensus Spectrum". This will compute the consensus spectrum, display it in a figure window, and fill all the filter-specifications fields with corresponding default values. The default values for the filter specifications are carefully chosen to suit the application, and hence the user can choose to keep these values to obtain quick results with good accuracy. If the values are changed and, subsequently, it is desired to reset them to the default values, simply press the "Consensus Spectrum" button again after selecting the desired protein IDs. It is recommended to use as many protein sequences in a functional group as possible to obtain an unambiguous characteristic frequency.

8) Design either one or both the filters by pressing the appropriate "Design Filter" button.

9) Press the appropriate "Predict" button to locate hot spots of the selected protein IDs in the "EIIP seqs. in memory" listbox using the desired filter type.

10) For the BPN filter, the prediction step provides two choices. The user can choose to either to perform manual tuning using the FSliderGUI or let the software perform automatic tuning. In the former case, the "Predict" button must be used, which will open the hot-spot figure windows in conjunction with the slider and the frequency response of the filter. The filter can be tuned by adjusting the slider. As the notch frequency changes, the open figure windows are automatically updated to reflect the changes. In the latter case, i.e., for automatic tuning, the "Predict & AutoTune" button must be used. This will perform automatic tuning and display the results in appropriate figure windows as well as the command window.