FIXED-POINT FREQUENCY DOMAIN ICA with GENERALIZED GAUSSIAN FUNCTION BASED NEGENTROPY APPROXIMATION for SPEECH SIGNAL SEPARATION

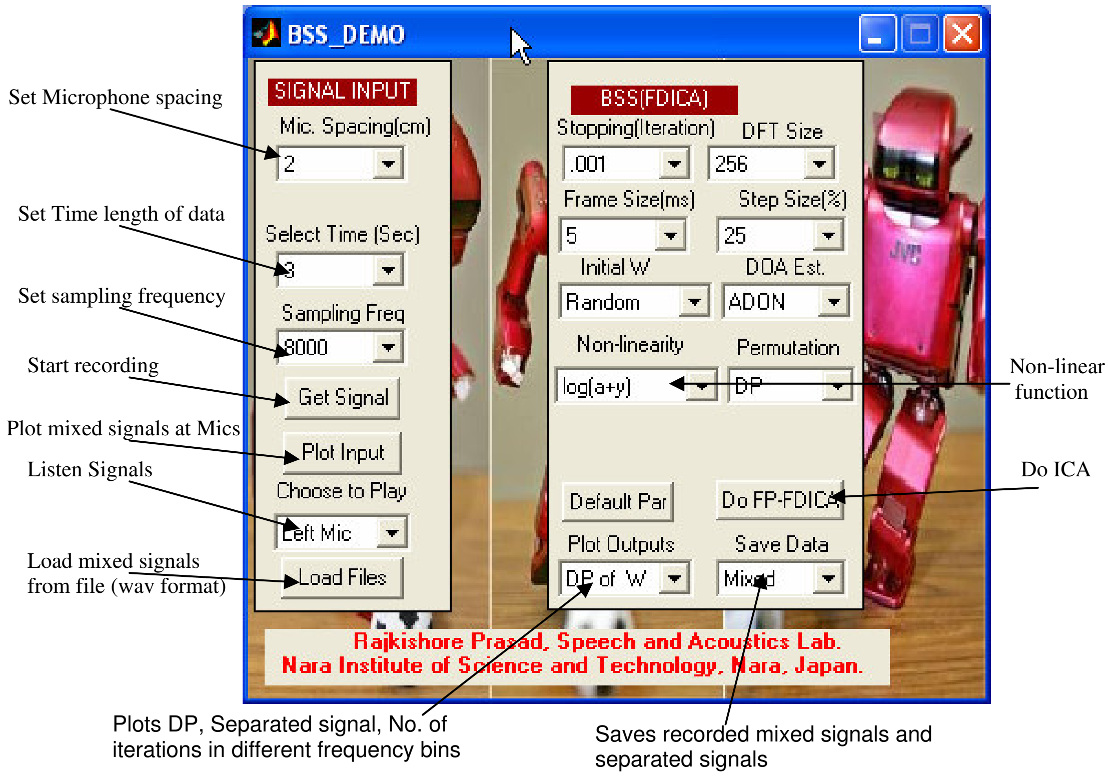

Fig.1 GUI for fixed-point ICA based blind separation of speech signal.

This GUI is for blind separation of speech signal by frequency domain fixed-point Independent Component Analysis (ICA) [1][2]. This demo uses GGD based statistical model for time-frequency series of speech [3] and uses similar nonlinear function for the negentropy approximation [4] which gives better convergence speed than the conventional nonlinear functions proposed in [1]. The general purpose logarithmic nonlinear function is also available for details please see [1].Using this demo you can capture speech signals (only under windows OS) from two speakers with two element linear microphone array, with inter-element spacing of 2cm, 4cm, and 6 cm. (only these are available spacings in menu !.), which can be set by making selection in pop-up menu under title '**Mic. Spacing**'. Some of the used functions have been taken from others (e.g. by John Burkardt and from VOICEBOX, toolbox) which are freely available.

Installation:
First unzip and extract the files to any location of your choice and bring that location in MATLAB path. GUI can be activated by typing BSS_DEMO on the command line. It requires MATLAB's Signal processing Toolbox and Statistics Toolbox to be installed on your PC. If you are going to capture speech data, you need *data acquision toolbox* as well as modifications in accordance with the two channel sound card to be used.

Data Input

It can load ***.wav*** files only. So if you have recorded data from two speakers (of course two channels, with microphone spacing of 4 cm) just press '**Load files'** button. If the microphone spacing is different for the loaded files, you need to set it accordingly from corresponding menu. Some speech data recorded with microphone spacing of 4cm and sampling frequency of 8000 Hz are placed in the directory named ***sample_mixed_data.*** These data were recorded in real environment with distance of speakers ranging from 1m to 2m from the linear microphone array. If you are running GUI on Linux/Unix platforms only prerecorded data can be used. Under windows OS, if you like to record data , set

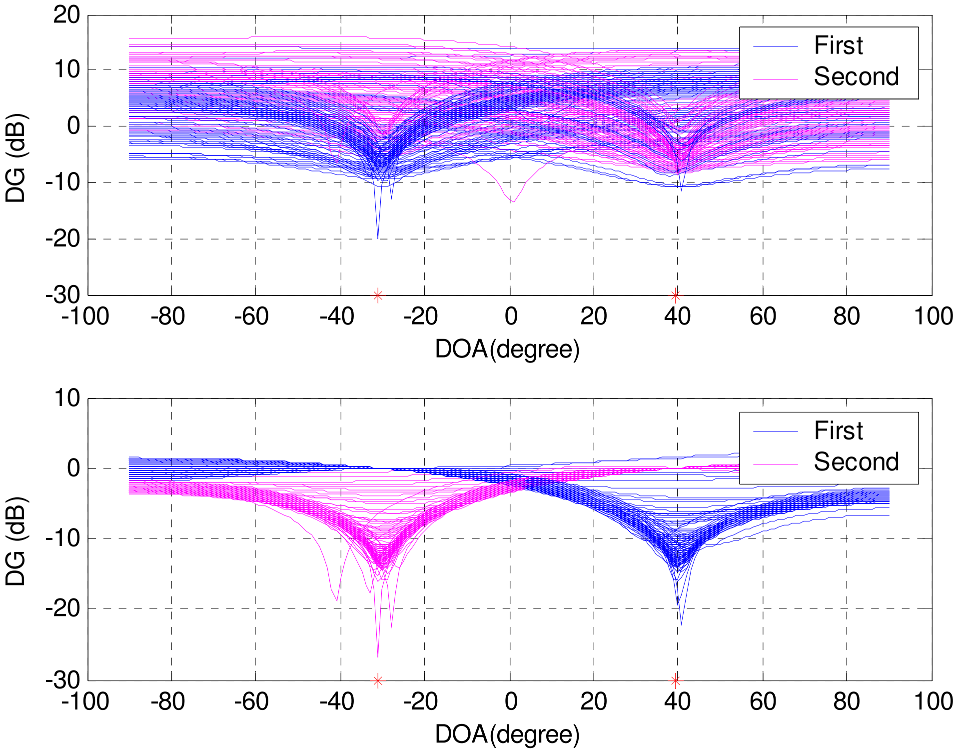

Fig.2 Directivity pattern of the separation matrix. Estimated directions of sources are shown by star on the plot. Upper plot shows DP for the permuted and unscaled matrix while the lower plot is for de-permuted and scaled separation matrix. Such plots can be plotted by choosing option ***DP of W*** from menu under **Plot outputs**.

the microphone spacing, sampling frequency and time length of the data and press push button with the label **'Get Signal'**. However, this will need data acquisition toolbox of MATLAB to be installed on your PC.
You can plot recorded/loaded speech signal by pressing **'Plot Input'** button and to listen loaded signal select from pop-up menu under **'Choose to Play'.**

FDICA
In order to do separation you need to set stopping criterion, DFT size, frame size and step size or shift size of analysis window (it uses Hanning window) as the % of analysis window size. First try with default parameters. However, after pressing **'Default par'** button, microphone spacing will be changes to 4cm so you need to change it again if it is not so with your pre-recorded data. Also you can change DFT size and non-linear functions from default values. Higher values of DFT size give better separation in the reverberant conditions. Finally, press **DO FP-FDICA** button. Before going to ICA loop, it will present before you a list of parameters it is going to use. If these are correct press OK, otherwise you can change it. It will communicate you again after completing separation and then you can play separated signals. You can also plot separated signals, Directivity Patterns (DP) and number of iterations consumed. In the plot for *DP of W* estimated DOA are shown by '*' as is shown in Fig.2. Plots for number of iterations show histograms for the number of iterations taken in different frequency bins.
If you need further help or to report bugs please send e-mail to Rajkishore Prasad at kishor-p@is.naist.jp or at rkishore2k@hotmail.com

References:

1. Appo Hyvarinen ,J. Karahunen, E. Oja,, *'Independent Component Analysis',* John Wiley and Sons Inc., 2001.
2. Rajkishore Prasad, Hiroshi Saruwatari, Akinov Lee and Kiyohiro Shikano, "A fixed point ICA algorithm for convoluted speech separation," Proc. International Symposium on ICA & BSS, pp. 579-584, Nara, Japan, 2003.
3. Rajkishore Prasad, Hiroshi Saruwatari, and Kiyohiro Shikano, ``Probability Distribution of Time-Series of Speech Spectral Components," IEICE Trans. Fundamentals, Vol.E87-A, No.3, pp.584--597. 2004.
4. Rajkishore Prasad, Hiroshi Saruwatari, and Kiyohiro Shikano," Blind Separation of Speech by Fixed-point ICA with Source adaptive Negentropy approximation," IEIEC Journal (to appear in, July 2005)