Faculty of Engineering and Physical Sciences

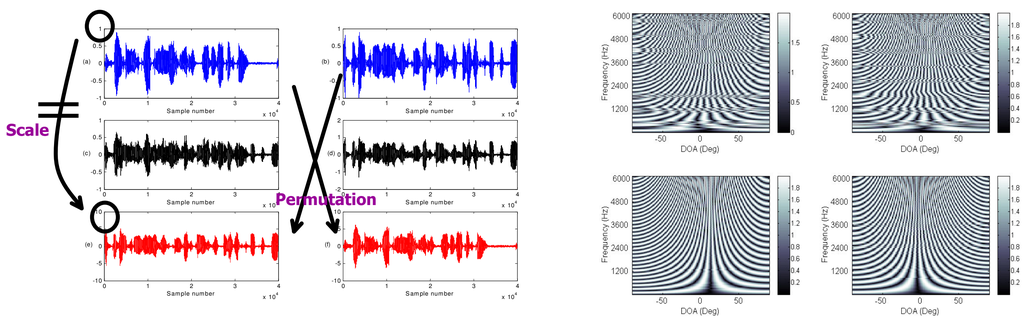
CVSSP Centre for Vision, Speech & Signal Processing

UNIVERSITY OF SURREY

Blind Signal Processing & Machine Audio Perception

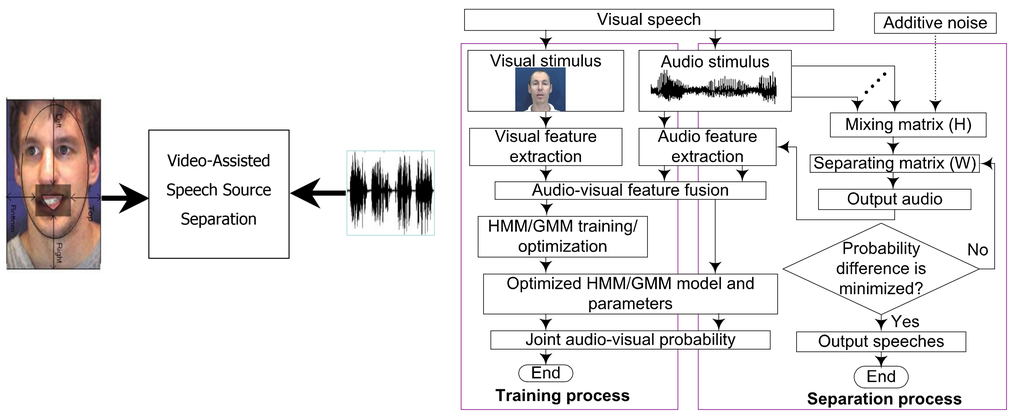
Blind Source Separation

The fundamental problem that BSS attempts to answer is: Given a number of observed (sensor) signals, without any (or very little) knowledge about the mixing channel and source signals, how to recover the source signals? With the help of statistical learning technique independent component analysis (ICA), BSS has been successfully applied to several areas including cocktail party problem, speech enhancement, biomedical signal processing, telecommunications, etc. The following two figures are examples of audio source separation.



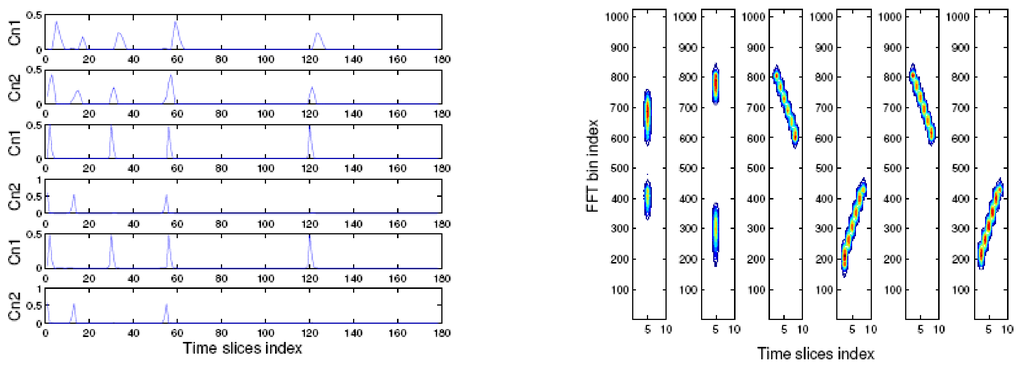
Audio-Visual Speech Separation

Human has the ability to identify the desired speakers even in a very noisy cocktail party environment, due to the natural bi-modality of human speech, where the speech is produced by the coupling of the visual movement of lips and facial muscles. Audio-visual speech separation actually take visual clues as additional information for the extraction of sound signals from a mixture, by the exploitation of the coherence between audio and visual modalities using a statistical model. The following two figures illustrate the basic idea and a practical system implementing this idea.



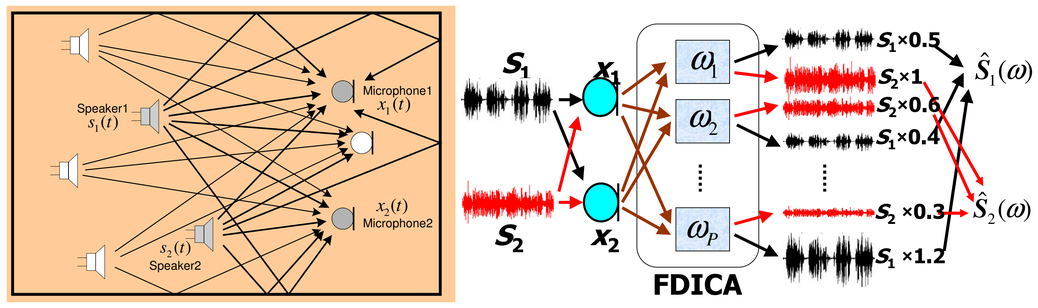
Convolutive Non-negative Sparse Coding

Non-negative sparse coding (NSC) is a machine learning technique for finding a low-rank representation of original data. Due to the non-negativity and sparseness constraints, NSC gives a parts-based sparse representation. We have extended this concept and considered a convolutive model, which has shown to be useful for the extraction of audio objects with time-varying frequency features in the magnitude spectrum domain. This extension can be very useful for the analysis of more complex auditory scenes. It would also be useful for the development of a hearing machine, a major effort for the machine audio perception, machine listening, or machine audition research.



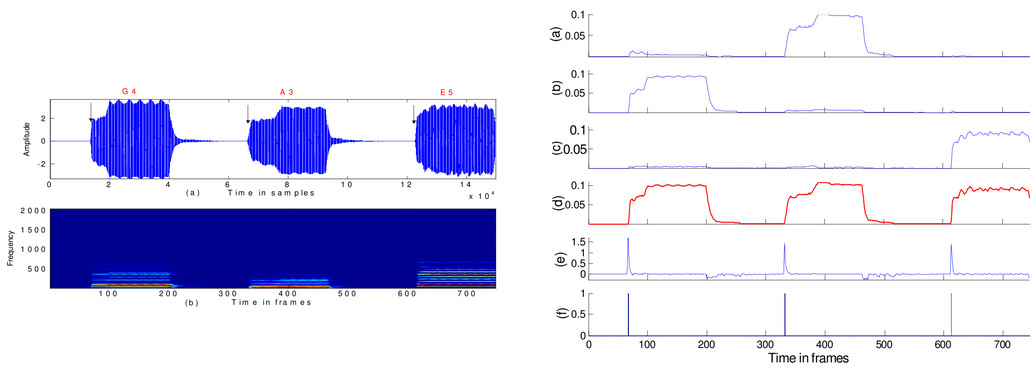
Cocktail Party Problem

Cocktail party problem can be addressed using convolutive BSS, where the observed signals at each microphone pick up not only the direct response from the speakers but also the reflections from walls, ceilings and floors. The desired source signals from the speakers are usually estimated by transforming the mixtures from the time-domain into the frequency domain. Cocktail party problem is closely related to computational auditory scene analysis, which investigates how to mimic the human's perceptual analysis of complex auditory scenes. The following two figures are graphical illustration of cocktail party problem and a frequency domain approach.



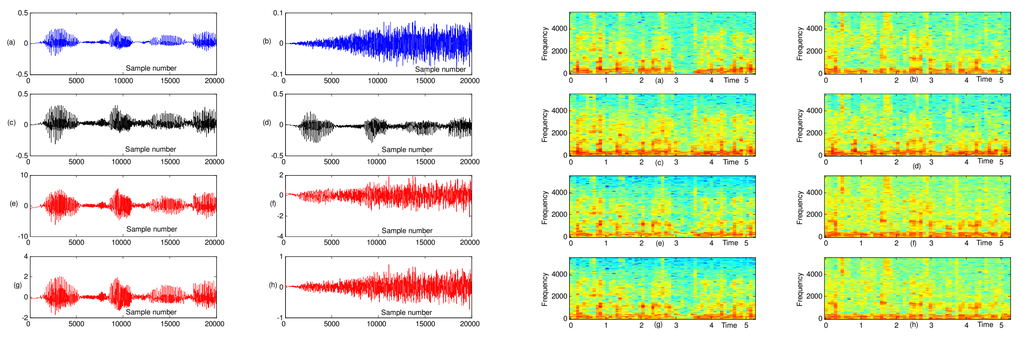
Audio Information Retrieval & Audio Transcription

Audio information retrieval and feature extraction is an important issue for audio transcription and perception, which has many applications in audio industry, such as low bit rate audio coding, audio transmission and machine listening. We have applied non-negative matrix factorisation (NMF) technique for the separation of the audio spectrogram, and the onsets and pitch frequencies can be detected from the separated spectrum of the individual notes. The proposed algorithms can be useful tools for high quality music consoles and audio signal processing software. The following figures give an example of onset detection of musical notes from an audio signal using the NMF.



Speech Enhancement

Speech enhancement is very useful in speech telecommunication, speech recognition, speaker verification, and human-machine interactions. We are particularly interested in applying BSS/ICA techniques for the removal of additive background noise, competitive speech, interference music, reverberations, in order to extract or restore the desired clean speech that we are after. As shown in the following figures, BSS can be useful for this purpose.



For more information, please contact Dr Wenwu Wang via email: w.wang@surrey.ac.uk