Applicability of Independent Subspace Analysis on Sweep Measurements

Xun Wang, Michael Vorländer

Institut für Technische Akustik, 52056 RWTH Aachen, E-Mail: xwa @akustik.rwth-aachen.de

Introduction

Impulse response and transfer functions are the main properties of linear acoustic systems, and sweeps are frequently used as excitation signals in measurements. Usually the measurement is contaminated with noise. Blind source separation methods such as independent subspace analysis (ISA) has already been successfully applied in the audio signal separation[1,2]. The principle of ISA is based on independent component analysis (ICA). However, ICA requires that the number of the sensors should be no less than the number of sources and it is difficult for ICA to solve the delay between sensors. This drawback makes ICA not applicable in the sweep measurement. ISA extends ICA by projecting one channel signal to multiple independent subspaces, and separates the sources in these independent subspaces. This allows ISA to separate multiple sources with only one-channel signal. In this paper, ISA is implemented to separate sweep signals from sweeps in narrow-band noise.

Methods

The basic ISA model is shown in Eq.1 and Eq.2. The recorded signal is a mixture of several sources.

$$x_{total} = \sum_{p=1}^{l} x_p \tag{1}$$

where the different sources x_p can be projected (transformed) into l independent subspaces.

$$x_p \xrightarrow{project} \sum_{i=1}^{l} a_{p,i} s_i \tag{2}$$

where the subspace function s_i is independent from each other. The goal of ISA is to find out the subspace functions s_i and the factors $a_{p,i.}$ Usually the projection in Eq.(2) is considered as Short Time Fourier Transformation, and the source x_p is transformed into a spectrogram. Assuming that the spectrogram of each source is composed of a few time invariant spectrums multiplied by time-variant weighting factors means that the shape of spectrum of each source remain unchanged with time and only the magnitudes of the spectrum changes with time. Assuming that S is the spectrogram of the source, it is described by an N × M matrix. The spectrogram of each source can also be written as $\mathbf{S}_j = \mathbf{f}_j \cdot \mathbf{t}_j^{\,T}$, where $\mathbf{f}_j = (f_1, f_2, \cdots f_N)^T$ is the short time spectrum and $\mathbf{t}_j = (t_1, t_2, \cdots t_M)^T$ is time variant weighting factors. The overall spectrogram of the mixture of the l sources is shown Eq.(3). In this case, only the absolute value of $\mathbf{f}_j$ is taken into consideration and the phase information of the spectrogram is kept for the subsequent time-domain signal reconstruction after separation.

$$\mathbf{S} = \sum_{j=1}^{l} S_j = \sum_{j=1}^{l} \mathbf{f}_j \mathbf{t}_j^{\,T} \tag{3}$$

ISA also assumes that this time-invariant frequency basis $\mathbf{f}_j$ are mutually independent, and this independence assumption could be described as Eq.(4)

$$\left(\mathbf{f}_i^T \mathbf{f}_j\right) = \begin{cases} E_i & i = j \\ 0 & i \neq j \end{cases} \tag{4}$$

E_i is a constant number of the power of the frequency basis $\mathbf{f}_j$. In fact, Eq.(4) shows the orthogonality of the frequencies basis $\mathbf{f}_j$, however, Eq.(4) is not the strictly sufficient condition for the independence, but it is a necessary condition for the independence assumption. Only if the condition Eq.(4) is satisfied, the different sources could possibly be separated. In order to extract the independent frequency basis $\mathbf{f}_j$, Singular Value decomposition (SVD) and ICA method is implemented.

At first, SVD is used to find out the orthogonal frequency basis and reduce the dimension of the sources. The SVD is obtained from the correlation matrix of spectrogram, which is $\mathbf{C} = \mathbf{S}^T\mathbf{S}$. The SVD of this correlation matrix is described $\mathbf{U\Sigma} V^T = \mathbf{C}$. And then reconstruct a new matrix $\mathbf{X} = \mathbf{\Sigma V}^T \mathbf{S}^T$, the rows of $\mathbf{X}$ can be considered as frequency basis which are orthogonal with each other.

In audio signal separation such as music or speech, the number of sources is much smaller than the number of frequency basis, the number of the frequency basis can be reduced though taking only the biggest singular values. [1] In contrast, in sweep measurements, because the frequency of sweep increases over time and every time-frequency block of the sweep on the spectrogram can be a independent frequency basis, all the singular values have to be taken into calculation. Meanwhile this orthogonal frequency basis in reconstructed matrix X are not sufficient for independent source separation, because any orthogonal transformation on $\mathbf{X}$ does not change the orthogonality of the rows of $\mathbf{X}$, but in fact the frequency basis has already been changed. Therefore the ICA algorithm should be implemented. The reconstructed matrix $\mathbf{X}$ can be written as the linear mixture $\mathbf{A}$ of a few independent vectors $\mathbf{F}^T$, as shown in Eq.(5)

$$\mathbf{X} = \mathbf{A}\mathbf{F}^T \tag{5}$$

where $\mathbf{A} = \mathbf{\Sigma V}^T \mathbf{T}_{\text{weight}}$, $\mathbf{T}_{\text{weight}}$ is the time-variant weighting factor matrix, and $\mathbf{F}^T$ is assumed to be the matrix of independent frequency bases. Eq.(5) satisfies the standard ICA model, and proceeded by ICA, the inverse matrix of A, $\mathbf{W} = \mathbf{A}^{-1}$ can be calculated [3], and the independent frequency bases and time-variant factor are obtained by

$$\mathbf{F}^T = \mathbf{W}\mathbf{X},\ \ \mathbf{F} = \left(\mathbf{f}_1, \mathbf{f}_2, \ldots \mathbf{f}_l\right) \tag{6}$$

$$\mathbf{T}_{\text{weight}} = \left(\mathbf{W\Sigma} V^T\right)^{-1},\ \ \mathbf{T}_{\text{weight}} = (\mathbf{t}_1, \mathbf{t}_2, \ldots \mathbf{t}_l)^T \tag{7}$$

Then $\mathbf{S}_j = \mathbf{f}_j \cdot \mathbf{t}_j^T$ is the spectrogram of the j-th separated sources. The correct separation of Eq.(6) and Eq.(7) must be on the condition that Eq.(4) is satisfied, otherwise ISA will lead to a incorrect separation.

Finally the spectrogram is transformed into time domain, and original phase information of the overall spectrogram is used. Depending on the priori knowledge about the source different clustering methods can be used. [2]

Simulation results

To test the applicability of ISA on the de-noising of sweep measurement, sweeps mixed with different noises are simulated. Fig.1 (a) shows the spectrogram of linear sweep mixed with 1500Hz and 4000Hz sine signal. For this spectrogram, the sampling rate is 44.1 kHz and the time block size is 1024 samples, and 50% overlap is used. Since the frequency of sweep increases over time, each time-frequency block of sweep is separated into a individual independent frequency basis. Fig. (b) and (c) are two examples of the separated sweep components. The 1500Hz and 4000Hz sine signal is separated into another component, as shown in Fig.1(d). As a result of ISA, the overall spectrogram Fig.1(a) is separated into 62 components, 61 components belong to sweep and one component is the sine signal.

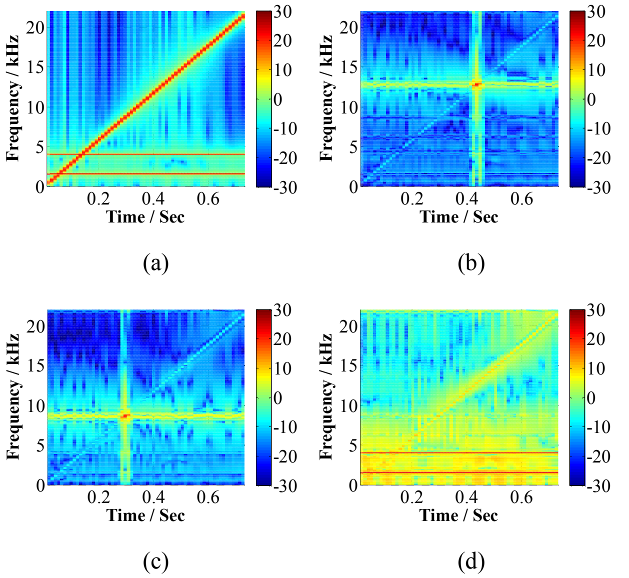

Figure 1 (a) Sweep mixed with 1500Hz and 4000Hz sine signal. (b)(c) Two examples of the separated sweep components. (d) The separated 1500Hz and 4000Hz sine signal component.

After the separation, the components which belong to the sweep are grouped together to synthesize the overall sweep (as shown in Figure 2). However, there is still a little residual 1500 and 4000Hz sine signal. If the noise is a random noise which covers certain frequency range and the frequency range covers more than one frequency bin the spectrogram. ISA can not separate it any more, because ISA cannot find out the "time-invariant" frequency basis for noise. If the noise is narrowband noise, one possible solution is to use short time-block size, because the frequency resolution is reciprocal to block size. When the short time block size is implemented, the frequency resolution is low, the narrow-band noise will transformed into the same frequency bin, the separation becomes possible. Figure 3 shows the separation results of sweep mixed with the noise of 1400~1600Hz and 3900~4100Hz. With the block size of 128 samples, ISA can separate still separate the sweep. At the blocks, where the blocks overlaps with the noise, the sweep component is separated into the noise signal.

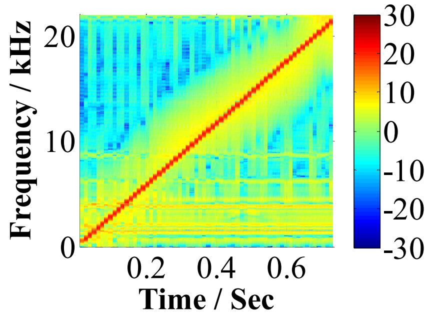

Figure 2 Separated sweep from 1500 and 4000 Hz sine signal

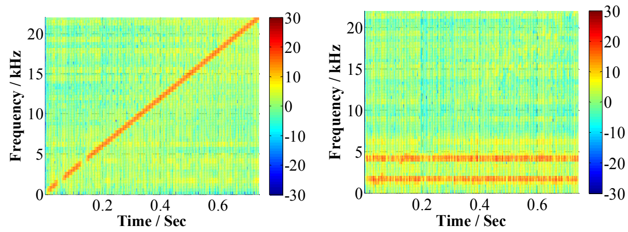

Figure 3 Separation of sweep and noise with frequency range of 14000~1600Hz and 3900~4100 Hz

Discussion

ISA can separate sweep with harmonic signal, however, but still some residual harmonic signal is left in sweep. As for the noise with certain frequency range, the short time-block size of spectrogram should be implemented, however, in this case at time-frequency blocks where the sweep overlaps with the noise, ISA still cannot separate the sweep from noise at these blocks.

Reference

[1] Casey, M. A. Separation of mixed audio sources by independent subspace analysis *Proceedings of the International Computer Music Conference, Berlin, August 2000*

[2] Wellhausen, J. Audio Signal Separation Using Independent Subspace Analysis and Improved Subspace Grouping *Signal Processing Symposium, 2006. NORSIG 2006. Proceedings of the 7th Nordic,* 2006, 310-313

[3] Aapo Hyvarinen, Juha Karhunen, Erkki Oja Independent component analysis *John Wiley & Sons,* 2001