BLIND SOURCE SEPARATION USING INDEPENDENT COMPONENT ANALYSIS IN THE SPHERICAL HARMONIC DOMAIN

Nicolas Epain, Craig Jin, André van Schaik

CARLab - Computing and Audio Research Laboratory
The University of Sydney
Sydney, NSW 2006, Australia
craig.jin@sydney.edu.au

ABSTRACT

Spherical microphone arrays provide a new and promising tool for the spatial analysis of complex sound fields. Most current frameworks, including HOA (Higher Order Ambisonics), rely on the spherical harmonic expansion of the sound field to form directional beams. Working in the spherical harmonic domain presents a number of advantages that include scalability (keeping the low-order components leads to a lower spatial resolution) and the ability to rotate the sound scene by a simple matrix operation on the signals. In this paper, we show that the spherical harmonic domain also provides significant advantages for the application of ICA (Independent Component Analysis) to separate and localise multiple sound sources.

1. BLIND SOURCE SEPARATION USING INDEPENDENT COMPONENT ANALYSIS

Independent Component Analysis (ICA) is a statistical method that was developed in the 1980's [1] and in this paper we are primarily interested in its application to the blind source separation (BSS) problem, although its applicability is more general. Assume that a vector of time signals, $\mathbf{X}(t)$, measured by N sensors are observed. In the BSS problem, we assume that these sensor signals consist of a linear mixture of some underlying source signals, *i.e.* there exists an N-by-M matrix $\mathbf{A}$ and a vector of M signals $\mathbf{S}(t)$ so that:

$$\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t)\,, \tag{1}$$

where $\mathbf{A}$ is the mixing matrix, and $\mathbf{X}(t)$ and $\mathbf{S}(t)$ are the vectors of the sensor and source signals, respectively. The signal vectors $\mathbf{X}(t)$ and $\mathbf{S}(t)$ are given by:

$$\begin{aligned} \mathbf{X}(t) &= \left[\mathbf{x}_1(t)\,,\mathbf{x}_2(t)\,,\,\ldots\,,\mathbf{x}_N(t)\right]^T \\ \mathbf{S}(t) &= \left[\mathbf{s}_1(t)\,,\mathbf{s}_2(t)\,,\,\ldots\,,\mathbf{s}_M(t)\right]^T . \end{aligned} \tag{2}$$

In the BSS problem, both the source signals and the mixing matrix are unknown; thus it is clear that the problem requires an additional constraint to obtain a solution, for otherwise the identity matrix is the simplest solution for the mixing matrix. In the ICA framework, this contraint is the additional assumption that the source signals are statistically independent with a non-Gaussian distribution.

Separating sound sources using a microphone array is a difficult task. While it may be assumed that the microphone signals are linear mixtures of the source signals, these mixtures are actually *convolutive* in most cases. The convolutive nature of the mixing may occur because of room reflections and also because the microphones are not co-located in space. Indeed, the only case where linear (instantaneous) mixtures of the sound sources are recorded by a microphone array is when co-located, directional microphones are used to record sounds in the free-field. Nevertheless, there are advantages to having co-located microphones in that the convolution is limited to the room reflections. As an example, consider the application of ICA to the signals recorded by a Soundfield microphone. In this case, there is theoretically the potential to be able to separate up to four sources using the four B-format output signals. In this paper we focus on a spherical microphone array, for which the microphones are not co-located. We show that applying ICA in the spherical harmonic domain removes the difficulty of having microphones that are not co-located in space.

Another difficulty with applying ICA to solve the BSS problem is that the source signals are ordered and normalised in an arbitrary way, *e.g.*, we have:

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \\ \vdots \\ \mathbf{x}_N(t) \end{bmatrix} &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{s}_1(t) \\ \mathbf{s}_2(t) \\ \vdots \\ \mathbf{s}_N(t) \end{bmatrix} \\ &= \begin{bmatrix} \frac{a_{13}}{2} & a_{11} & \ldots & a_{12} \\ \frac{a_{23}}{2} & a_{21} & \ldots & a_{22} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n3}}{2} & a_{n1} & \ldots & a_{n2} \end{bmatrix} \begin{bmatrix} 2\mathbf{s}_3(t) \\ \mathbf{s}_1(t) \\ \vdots \\ \mathbf{s}_2(t) \end{bmatrix} . \end{aligned} \tag{3}$$

This issue is referred to as the permutation problem and becomes an issue when processing the signals using short and overlapping time frames. It becomes difficult to join the separated signals from one time frame to the next. In the following, we show that using ICA in the spherical harmonic domain solves the permutation problem because it localises the sources in space. The different sources can then be tracked over time, and ordered based on their directions.

2. SPHERICAL HARMONIC DESCRIPTION OF THE SOUND FIELD

The Higher Order Ambisonics (HOA) method of spatial recording and playback is based on the idea that the sound field can be described as a sum of spherical harmonic components. In this section, we briefly review the fundamentals of HOA and show why the spherical harmonic domain provides a powerful environment for blind source separation using ICA.

2.1. Fourier-Bessel representation of the sound field

In the frequency domain, any sound field consisting of incoming sound waves can be expressed as a series of spherical harmonic functions [2]:

$$p\left(r,\theta,\varphi,k\right)=\sum_{l=0}^{\infty}\sum_{m=-l}^{m} i^{l}j_l(kr)\,Y_l^m(\theta,\varphi)\,b_{lm}(k)\,, \qquad (4)$$

where $p\left(r,\theta,\varphi,k\right)$ is the acoustic pressure corresponding to the wave number k and at the point with spherical coordinates (r,θ,φ), and where j_l is the spherical Bessel function of degree l and Y_l^m is the spherical harmonic function of order l and degree m. Equation 4, known as the Fourier-Bessel representation of the sound field, states that the sound field is fully described by the complex coefficients $b_{lm}(k)$. Describing the sound field at any point in space requires an infinite number of b_{lm} coefficients, which is impossible in practice. Nevertheless, a good approximation of the acoustic pressure in the vicinity of the origin can be obtained by truncating the series to an order L that depends upon the wave number, k, and radius, r, from the origin according to the formula [3]:

$$L \geq \frac{ekr-1}{2}, \qquad (5)$$

where e is the mathematical constant known as Euler's number.

In this paper, we consider only the order-L truncated spherical harmonic description of the sound field. As well, in the following we refer to the coefficients b_{lm} up to order L as the order-L HOA components. The corresponding time-domain signals (inverse Fourier transforms of the $b_{lm}(k)$) are referred to simply as the HOA signals.

2.2. HOA signals in the case of plane waves

The HOA components corresponding to a plane wave incoming from direction (θ_i,φ_i) are given by:

$$\mathbf{b}_i(k)=s_i(k)\,\mathbf{y}_i, \qquad (6)$$

where $s_i(k)$ is the Fourier transform of the signal at wavenumber k and

$$\mathbf{y}_i=\left[Y_0^0(\theta_i,\varphi_i)\,,\;Y_1^{-1}(\theta_i,\varphi_i)\,,\dots,\;Y_L^L(\theta_i,\varphi_i)\right]^T. \qquad (7)$$

The Y_l^m are the spherical harmonic functions as described previously for Equation 4. In the case where N plane waves are contributing to the sound field, the resulting HOA components are obtained by a simple matrix-vector product:

$$\mathbf{b}(k)=\sum_i \mathbf{b}_i(k)=\mathbf{Y}\mathbf{s}(k)\,, \qquad (8)$$

where

$$\mathbf{Y}=[\mathbf{y}_1,\mathbf{y}_2,\dots,\mathbf{y}_N] \quad \text{and}$$
$$\mathbf{s}(k)=\left[s_1(k)\,,s_2(k)\,,\dots,s_N(k)\right]^T. \qquad (9)$$

As matrix $\mathbf{Y}$ does not depend on frequency, the time-domain HOA signals resulting from the same N plane waves can be expressed as the following matrix product:

$$\mathbf{B}(t)=\mathbf{Y}\mathbf{S}(t)\,, \qquad (10)$$

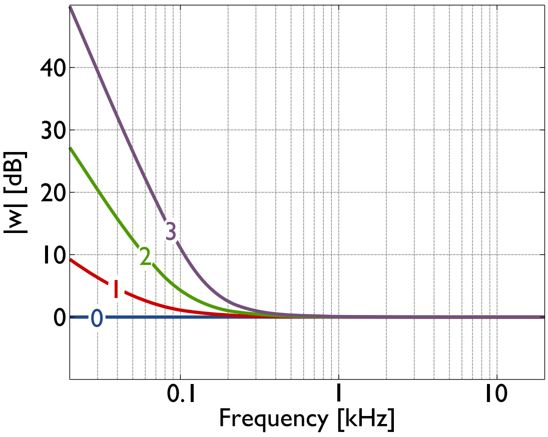

Figure 1: Amplitude of the modal coefficients is shown as a function of frequency for a spherical source positioned 1 m away from the origin.

where $\mathbf{B}(t)$ and $\mathbf{S}(t)$ are the matrices containing the HOA and plane-wave signals, respectively. $\mathbf{B}(t)$ and $\mathbf{S}(t)$ are given by:

$$\mathbf{B}(t)=\left[\mathbf{b}_{00}(t)\,,\mathbf{b}_{1-1}(t)\,,\dots,\mathbf{b}_{LL}(t)\right]^T$$
$$\mathbf{S}(t)=\left[\mathbf{s}_1(t)\,,\mathbf{s}_2(t)\,,\dots,\mathbf{s}_N(t)\right]^T. \qquad (11)$$

Please note that the notations for $\mathbf{b}$ and $\mathbf{s}$ are overloaded in that the Fourier domain and time domain variables use the same symbols, but with the arguments k and t, respectively. As well, $\mathbf{b}$ is used with both a single subscript and a double subscript. With a single subscript, $\mathbf{b}$ denotes the HOA component for a single plane wave direction and the subscript indicates the direction. With a double subscript, $\mathbf{b}$ denotes the HOA component for the complete sound field and the subscripts indicate the order and degree of the spherical harmonic components. Given the above equation, it should be clear that under the assumption that the sound field consists of a sum of N plane waves, the order-L HOA signals form an instantaneous mixture of the source signals. In the case where N is less than the number of harmonics, ICA should perform well at un-mixing the source signals from the HOA signals. It should, of course, be noted if the sound field is reverberant, then some of the N plane waves may be reflections.

2.3. HOA signals in the case of spherical waves

Generally and realistically, a sound field rarely consists of only a sum of plane waves. In a typical sound field recording scenario, such as a music concert or a meeting, sound sources are more appropriately described as spherical sources. Spherical sources differ from plane wave sources in the way they contribute to the spherical harmonic expansion of the sound field. In the case of a unique spherical source at coordinates (r_i,θ_i,φ_i), the corresponding HOA components are given by:

$$\mathbf{b}_i(k)=s_i(k)\,\mathbf{W}(kr_i)\,\mathbf{y}_i\;\;, \qquad (12)$$

where $\mathbf{W}(kr_i)$ is a diagonal matrix whose coefficients along the diagonal are the modal coefficients for source i and are given by [4]:

$$w_l(kr_i)=i^{-l}\frac{h_l(kr_i)}{h_0(kr_i)}\;\;, \qquad (13)$$

where h_l is the degree-l spherical Hankel function of the first kind. Figure 1 illustrates the amplitude of these modal coefficients as a function of the frequency in the case of a spherical source located one meter away from the origin.

In contrast with the plane-wave case, the mixing matrix for the source signals changes with frequency. Thus, the time-domain HOA signals are no longer instantaneous mixtures of the source signal. Instead, they are obtained by convolving the source signal with filters corresponding to the modal coefficients:

$$\mathbf{B}(t) = \sum_i \mathbf{w}_i(t) * (\mathbf{y}_i \mathbf{s}_i(t)) \tag{14}$$

However, as suggested by Fig. 1, the value of the low-order modal coefficients is close to 1 for most frequency values. Depending on the frequency content of the source signal and on the source distance, it is thus possible to approximate the contribution of a spherical source as a plane wave source in the same direction. For instance, the contribution of a talker located one meter away from the origin to the order-1 HOA signals will be indistinguishable from that of a plane-wave source in the same direction, since the order-1 modal coefficient is close to 1 above 100 Hz and speech has low energy below this frequency. In summary, when the sound field consist of spherical sound sources, the ICA analysis of the HOA signals may still provide reasonable results, when the sources are 'far enough' relative to their frequency content.

2.4. HOA encoding in practice

In the previous sections we have shown that the spherical harmonic domain provides some interesting advantages when using ICA to separate sound sources. However, our arguments hold only when the spherical harmonic expansion of the sound field is known perfectly. This is rarely the case, instead, the operation of obtaining HOA signals, which is known as HOA encoding, is achieved by filtering the signals recorded by a spherical microphone array with encoding filters and is subject to errors [5]. For a given microphone array, the encoding error depends strongly on both the frequency and the order of the decomposition. Fig. 3 illustrates the Signal-to-Noise Ratio (SNR) of the HOA signals for a typical spherical microphone array in the case of a -40 dB RMS measurement noise (note that this array is more precisely described in Section 4). Clearly, there exists only a narrow frequency range where the HOA signals are accurately encoded. At low frequencies, the measurement noise is dramatically amplified by the encoding filters; at high frequencies, spatial aliasing pollutes the different HOA signals with information belonging to higher order harmonics.

As explained above, the HOA domain is appropriate for using independent component analysis in that the HOA signals are instantaneous mixtures of the source signals under certain conditions. However, this is true only in the frequency range where the encoding is almost perfect. A way of avoiding this problem is to design the microphone array so that the encoding quality is accurate enough for the entire spectral content of the source signals. Another possibility is to perform the ICA on band-pass filtered HOA signals that contain only the frequencies where the encoding is accurate.

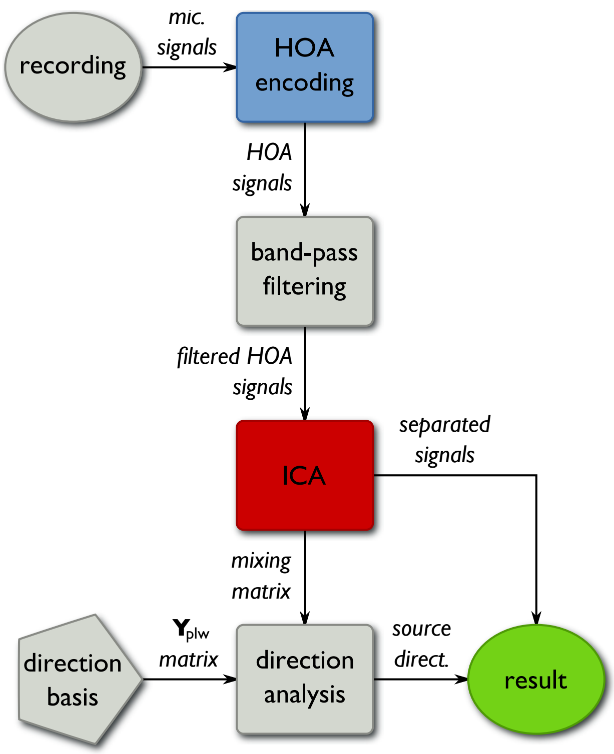

Figure 2: A flow diagram is shown for the proposed ICA method that is performed in the spherical harmonic domain.

3. A NEW METHOD FOR USING ICA WITH A MICROPHONE ARRAY

3.1. General method

Figure 2 shows the proposed general method for using Independent Component Analysis with a microphone array. First, the signals recorded by the microphone array are encoded to the HOA domain using HOA-encoding filters. Second, the obtained noisy HOA signals are band-pass filtered to keep only the frequencies where the HOA signals can be considered as instantaneous mixtures of the source signals, *i.e.*, the frequencies where the encoding is almost perfect and the source modal coefficients are close to 1. In a third step, the ICA is performed on the band-pass filtered HOA signals. The output of this analysis consists of a set of separated signals and a mixing matrix. In the standard ICA framework, there would then be an indeterminacy on the ordering and gain of the sources, known as the permutation problem. In this case, however, the resulting mixing matrix can be analysed to identify the source positions. Finally, we obtain a set of separated source signals and their corresponding angular positions.

3.2. Source direction analysis

As shown in Section 2, the relation between the band-pass HOA signals and the source signals is a simple matrix product:

$$\hat{\mathbf{B}}(t) = \mathbf{Y}\mathbf{S}(t)\,, \tag{15}$$

where $\hat{\mathbf{B}}(t)$ is the matrix containing the band-pass filtered HOA signals. Assuming the ICA separates the source signals perfectly, the output signals are proportional to the actual source signals and the columns of the corresponding mixing matrix are proportional to those of matrix $\mathbf{Y}$:

$$\mathbf{A} = \mathbf{YG}, \tag{16}$$

where $\mathbf{G}$ is a diagonal matrix whose non-zero coefficients are the proportionality constants between the actual source signals and the extracted ones. It is then possible to retrieve the source directions by calculating the correlation, Γ_{ij}, between the i-th column of the mixing matrix and the j-th column of a $\mathbf{Y}$ matrix corresponding to a large number of directions in space:

$$\Gamma_{ij} = \frac{\mathbf{a}_i^T \mathbf{y}_j}{\|\mathbf{a}_i\| \, \|\mathbf{y}_j\|} \tag{17}$$

where $\mathbf{a}_i$ and $\mathbf{y}_j$ denote the ith column of $\mathbf{A}$ and a vector of the spherical harmonic function values for direction (θ_j, φ_j), respectively. The source direction is then chosen as the one that shows the maximum correlation value with $\mathbf{a}_i$.

Significantly, the value of the maximum correlation can be used to determine whether or not the extracted signals correspond to an actual source or not. Real sources should show a large correlation with at least one direction in space, providing the direction basis is precise enough. Therefore, signals whose corresponding spatial correlation values are below a certain threshold (typically 0.95) can be considered as residuals of the ICA and discarded.

3.3. Wide-band BSS method

A major drawback of the proposed method is that the output signals from the ICA are band-pass filtered. A simple way of obtaining wide-band separated signals is to apply the un-mixing matrix resulting from the ICA to the full-band HOA signals:

$$\hat{\mathbf{S}}(t) = \mathbf{A}^{-1}\mathbf{B}(t)\,, \tag{18}$$

where $\hat{\mathbf{S}}(t)$ and $\mathbf{A}^{-1}$ denote the matrix of the full-band separated signals and the un-mixing matrix, respectively. As the assumption that the HOA signals are an instantaneous mixture of the source signals holds only for the band-passed HOA signals, this method is unable to separate the source signals efficiently at low and high frequencies. Nevertheless, it is our hypothesis that the perceived quality of the output signals is generally improved when using the wide-band method.

4. FREE-FIELD MULTIPLE TALKER SEPARATION AND LOCALISATION

In this section, we simulate the use of the proposed method to blindly separate simultaneous talkers recorded by a spherical microphone array in anechoic conditions.

4.1. Simulation setup

In the simulation, eight talkers are located at a distance of two meters around the microphone array. The different angular positions of the talkers are shown in Fig. 4. Each talker is modelled as a spherical source. The talker signals are male or female speech signal recorded in free field conditions.

The microphone array has been designed to provide high quality 3D, order 2, HOA signals from about 300 to 3500 Hz,

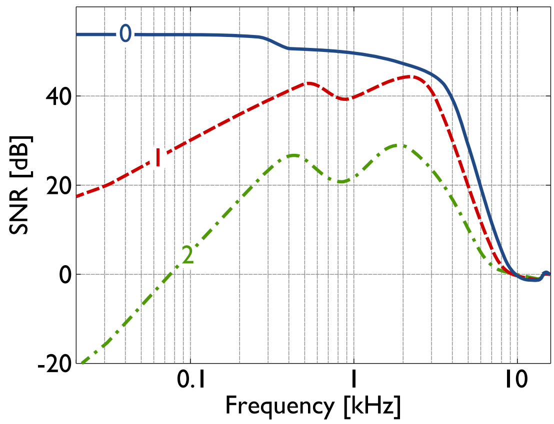

Figure 3: The signal-to-noise ratio of the HOA signals at order 0 (solid line), 1 (dashed line) and 2 (dash-dotted line), is shown when recording with the microphone array described in Section 4 in the presence of a -40 dB RMS measurement noise.

which approximately corresponds to the narrow-band frequency range of a telephone. Its design is similar to the one of the microphone array presented in [6]. It consists of two concentric arrays of 12 omnidirectionnal microphones. There are 12 microphones located on the surface of a rigid sphere with a radius of 3 cm; the other 12 microphones are located on the surface of a imaginary sphere with a radius of 15 cm. For both arrays, the angular positions of the microphones correspond to the corners of an icosahedron. The 512-sample long encoding filters are calculated to maximise the SNR of the different order HOA signals for every frequency value. The resulting SNR is presented in Fig. 3.

The sound wave propagation between the sources and the sensors is modelled using a very high order spherical harmonic expansion, which simulates the diffractive effect of the rigid sphere accurately. In addition, the effect of measurement noise is modelled by adding a -40 dB RMS uncorrelated white noise to the microphone signals. The obtained HOA are band-pass filtered so that only the 300–3500 Hz frequency range is used for the analysis. The ICA is then performed on the narrow-band HOA signals using FastICA [7], an Independent Component Analysis package for the MATLAB environment. Finally, the source direction analysis is performed using a regular angular distribution with a constant step of 1° for both the azimuth and elevation.

4.2. Simulation results

Fig. 4 shows the results of the source direction analysis performed on the mixing matrix, as described in Section 3.2. The resulting error for the source directions is of the order of 1°, which is on the order of the accuracy of the direction basis vectors that were used. Although the output of the ICA contains 9 signals, which corresponds to the number of input HOA signals, one of them has been discarded due to the low correlation between the corresponding mixing matrix column and all of the direction basis vectors.

In order to evaluate the quality of the source separation, the PESQ (Perceptual Evaluation of Speech Quality) scores [8] have been calculated for the output signals of the proposed ICA

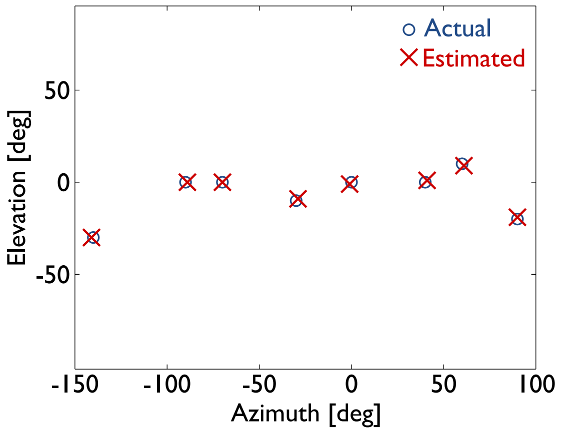

Figure 4: The true and estimated source positions for the eight simultaneous talkers is shown for the free field sound condition.

Talker	1	2	3	4	5	6	7	8	avg.
Sph. beam.	1.1	1.5	1.7	1.3	1.5	1.3	1.7	1.8	1.5
Nar. ICA	2.1	2.5	2.4	2.2	2.4	2	3	3.1	2.5
Wid. ICA	1.9	2.3	2.2	1.9	2.2	1.9	2.9	3	2.3

Table 1: PESQ scores for talkers 1–8 and average PESQ score obtained with: an order-2 spherical beamformer steered in the talker directions; the proposed narrow-band ICA method; the proposed wide-band ICA method.

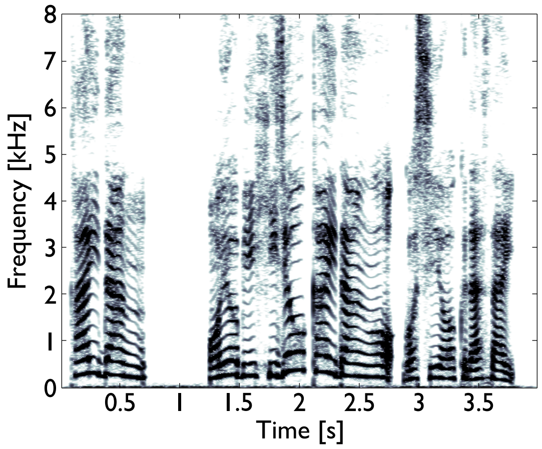

Figure 5: A spectrogram of the signal corresponding to the source located in direction $(-30°,-10°)$ is shown.

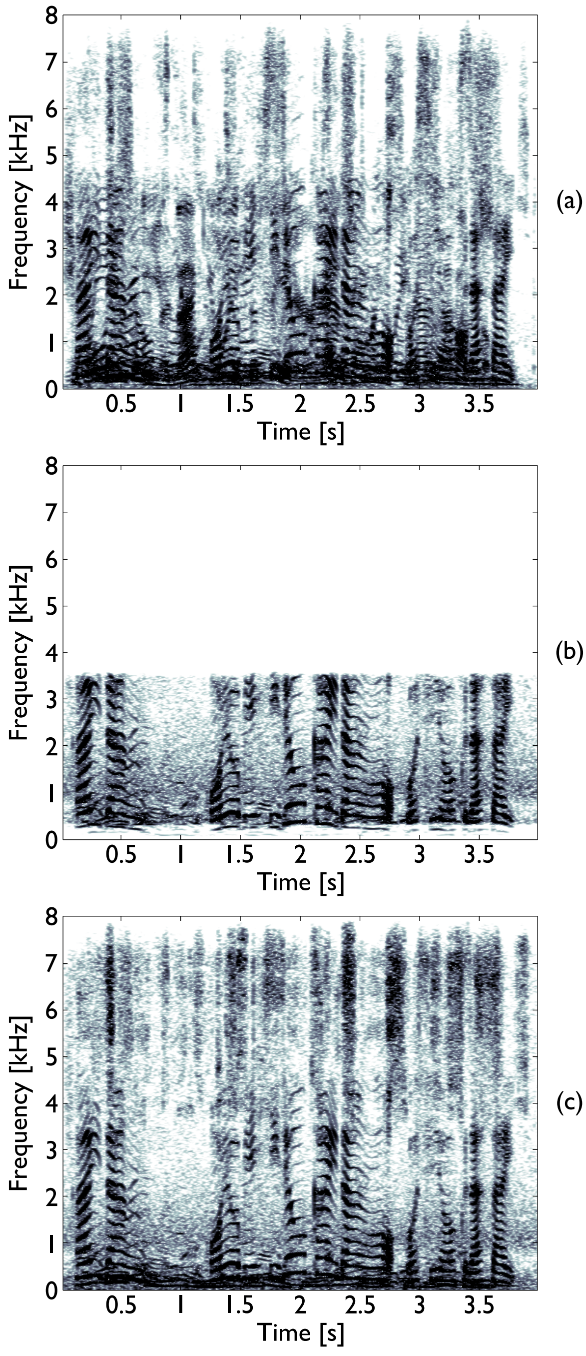

Figure 6: From top to bottom, spectrograms are shown of: (a) the output of a second-order spherical beamformer steered to the direction $(-30°,-10°)$; (b) the output of the proposed narrow-band ICA method corresponding to the source found close to direction $(-30°,-10°)$; (c) the output of the proposed full-band ICA method corresponding to the source found close to direction $(-30°,-10°)$.

method for every talker. The calculation was performed using the MATLAB PESQ package included in [9]. The results are presented in Table 1, as compared with the PESQ scores obtained when using an order-2 spherical beamformer [10] steered in the talker directions. Both ICA methods show a much larger PESQ score than the beamformer, suggesting that the sources are better separated. On the other hand, the scores obtained by the ICA methods are low in absolute terms, which is partly due to a high level of noise in the signals. This noise results from the HOA-encoding of noisy microphone signals.

Figs. 5 and 6 illustrate the differences in the quality of the output signals obtained with the ICA methods, as compared with the output of the spherical beamformer. Fig. 5 shows the spectrogram of the original signal corresponding to the source located in direction $(-30^\circ, -10^\circ)$. This spectrogram can be compared to the spectrograms of the corresponding outputs obtained with the spherical beamformer, the narrow-band ICA and the wide-band ICA, shown in Fig. 6. Although it shows a significant amount of noise, the output corresponding to the narrow-band ICA method is clearly less polluted by the other talkers than the spherical beamformer output. On the other hand, due to band-pass filtering, it has no energy apart from the 300–3500 Hz frequency band. Finally, the wide-band ICA method still provides good source separation in the frequency band 300–4000 Hz, while its high and low frequency content are less polluted by the other talkers than the output of the spherical beamformer.

5. CONCLUSIONS AND FUTURE WORK

Blind source separation using ICA is usually a difficult task when based on the signals recorded by a microphone array. This is due to the convolutive nature of the source-to-microphone impulse responses. In this paper, we propose a general framework that allows for ICA-based BSS with a large number of microphones. The underlying idea is that the HOA signals provide instantaneous mixtures of the source signals. Simulation results demonstrate that the method is able to accurately localise and separate up to 8 simultaneous talkers in an anechoic environment, when using a microphone array with both a reasonable size and number of microphones.

The quality of the HOA encoding is a significant factor for the proposed method and highlights the importance of the microphone array design. For instance, being able to separate up to 8 talkers requires a high quality encoding up to order 2 for a large frequency range. These performance characteristics can be achieved using a dual, concentric array design, such as that described in [11].

The simulation results demonstrate that our ICA algorithms perform better at separating the talkers than the classic spherical beamforming. However, due to the presence of measurement noise, the quality of the output signals can still be improved. For example, this could be improved by using denoising filters, such as the Wiener filter. Significantly, the proposed source localisation method demonstrates high accuracy. Although the use of the proposed novel method has been simulated in an anechoic environment, sound source separation is typically performed on signals recorded in a reverberant environment. The performance of the method in the presence of reverberation will be evaluated in future work. Preliminary simulations show that slightly degraded performance is achieved both in terms of source localisation and the maximum number of talkers when there is a moderate amount of reverberation.

6. REFERENCES

[1] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York, USA: Wiley Interscience, 2001.

[2] E. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography*. London, UK: Academic Press, 1999.

[3] N. Gumerov and R. Duraiswami, *Fast multipole methods for the Helmholtz equation in three dimensions*. The Netherlands: Elsevier, 2005.

[4] J. Daniel, “Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia,” Ph.D. dissertation, Université Paris 6, Paris, France, 2000.

[5] B. Rafaely, “Analysis and design of spherical microphone arrays,” *IEEE Transactions on speech and audio processing*, vol. 13, no. 1, pp. 135–143, January 2005.

[6] A. Parthy, C. Jin, and A. van Schaik, “Acoustic holography with a concentric rigid and open spherical microphone array,” in *Proceedings of the ICASSP 2009*, Taiwan, April 2009.

[7] “The FastICA package for MATLAB.” [Online]. Available: http://www.cis.hut.fi/projects/ica/fastica/

[8] ITU-T, “Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs,” in *ITU-T Recommendation P.862*, 2001.

[9] P. Loizou, *Speech enhancement: theory and practice*. Boca Raton, USA: CRC Press, 2007.

[10] J. Meyer and T. Agnello, “Spherical microphone array for spatial sound recording,” in *Proceedings of the 115th AES convention*, New York, USA, October 2003.

[11] A. Parthy, C. Jin, and A. van Schaik, “Optimisation of co-centred rigid and open spherical microphone arrays,” in *Proceedings of the AES 120th Convention*, Paris, France, May 2006.