Mixed-order Ambisonics encoding of cylindrical microphone array signals

Jorge Trevino1,2,* Shoichi Koyama3,† Shuichi Sakamoto1,2,‡ and Yōiti Suzuki1,2,§

1Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, 980–8577 Japan
2Graduate School of Information Sciences, Tohoku University, 6–3–09 Aramaki Aza Aoba, Aoba-ku, Sendai, 980–8577 Japan
3NTT Media Intelligence Laboratories, 3–9–11 Midori-cho, Musashino, 180–8585 Japan

(Received 29 November 2013, Accepted for publication 18 December 2013)

Keywords: Mixed-order Ambisonics, Cylindrical array, Spherical harmonics, Sound field reproduction

PACS number: 43.60.Fg [doi:10.1250/ast.35.174]
pressure field, cylinder can be treated as the sum of an incident sound recording system \[8\]. The sound pressure field sensed by pressure sensed by them. This leads to a more robust magnify the effects of the direction of arrival on the sound arrays an efficient choice for MOA.

The possibility of independently choosing the horizontal and polar resolutions makes cylindrical arrays an efficient choice for MOA. The microphones are placed over a rigid cylinder. This is done by continuing the spherical harmonic expansion up to a higher Ambisonic order, the horizontal Ambisonic order. However, only the horizontally-oriented spherical harmonics are considered. That is, only the expansion coefficients for which \( m = \pm n \) are used, the rest of the coefficients are not calculated.

3D Ambisonic order is used to encode the entire sound field by applying Eq. (1). This encoding is complemented with a high-resolution description of the sound field in the horizontal plane. This is done by introducing a relationship between each of the plane waves and its corresponding observed total field. Arbitrary incident and total fields will be related by the linear combination of these results.

The first step is to consider that the cylinder defines a rigid boundary; that is

\[
\frac{\partial}{\partial r} \psi_i(k, r) \bigg|_{r = R_{\text{cyl}}} = 0. \tag{5}
\]

Here, \( R_{\text{cyl}} \) stands for the radius of the cylinder. After introducing this constraint, the total sound field on the surface of the rigid cylinder, when the incident field consists of a plane wave arriving from an elevation angle \( \theta_{\text{src}} \) and azimuth angle \( \phi_{\text{src}} \), is given by the following formula:

\[
\tilde{\psi}_i^{\text{src}}(k, \theta, z) = \frac{i \exp[i k \sin \theta_{\text{src}} z]}{\pi^2 k \sin \phi_{\text{src}} R_{\text{cyl}}} \sum_{n=-\infty}^{\infty} i^n \exp[i n(\theta - \theta_{\text{src}})] H_n^{(1)}(k \sin \phi_{\text{src}} R_{\text{cyl}}). \tag{6}
\]

We use the notation \( \tilde{\psi}_i^{\text{src}} \) to emphasize this is the total field due to a plane wave incoming from a particular direction. \( H_n^{(1)} \) stands for the derivative of the Hankel function of the first kind and order \( n \). The total sound field, in general, is given by the linear combination:

\[
\psi_i = A \cdot \tilde{\psi}_i. \tag{7}
\]

The weights \( A \) are the same for both the incident and the total field. Calculating them will allow us to remove the scattered field.

4. Proposed system

Our proposal uses a cylindrical microphone array to record sound field information and encodes it using MOA. A block diagram for the proposed encoder is shown in Fig. 3. We introduce a measuring grid which does not necessarily match the microphones’ layout. A set of beamformers [9] are
used to isolate the sound arriving from each direction in the measuring grid. These measurements are encoded according to their angles of arrival using the spherical harmonic functions. Finally, the encoding is normalized using the Furse-Malham weighting coefficients [10] to ensure optimal use of the system’s dynamic range before the signals are broadcasted or stored. The resulting encoding is in line with any system capable of reproducing MOA [2,3,11].

The horizontal accuracy of MOA encodings is set by the horizontal Ambisonic order \( M \). In MOA, \( 2M + 1 \) channels are used to characterize the sound field in the horizontal plane. It is possible to specify a total number of channels for the full encoding with some freedom. This determines the amount of information used to encode elevation by setting the full-3D encoding with some freedom. This determines the amount of information used to encode elevation by setting the full-3D Ambisonic order \( M \) according to the following expression:

\[
\text{No. channels} = N^2 + 2M + 1. \tag{8}
\]

The special case of \( N = M \) corresponds to conventional HOA, where both azimuth and elevation resolutions are equal. Most practical applications of MOA use relatively small values for \( N \).

### 4.1. The measuring grid and beamforming

Our proposal defines a measuring grid as an almost uniform sampling of all directions. The density of the sampling is given by the constraint:

\[
\text{No. directions} \geq (M + 1)^2. \tag{9}
\]

It is not possible to define a uniform measuring grid for an arbitrary number of directions. Nevertheless, a minimum-energy sampling of the sphere [12] can be used as a basis for the measuring grid. An illustration of this is shown in Fig. 4 with some points on the sphere omitted for clarity.

To isolate the sound field arriving from some direction \((\theta_{\text{dir}}, \varphi_{\text{dir}})\) in the measuring grid, our proposal uses a least-squares approach. The task is to calculate the corresponding weight \( A_{\text{dir}} \) from Eqs. (4) and (7) using the microphone signals. To accomplish this, we calculate the pseudo-inverse of a matrix formed by evaluating Eq. (6) for all microphone positions and for all the directions in the measuring grid. This results in the following expression:

\[
W_{LS}(k, \theta_{\text{mic}}, \varphi_{\text{mic}}, \theta_{\text{dir}}, \varphi_{\text{dir}}) = \left[ \frac{i \exp[i k \sin \varphi_{\text{dir}} \sin \theta_{\text{mic}}]}{\pi^2 k \sin \varphi_{\text{dir}} R_{\text{cyl}}} \right].
\]

4.2. Directional encoding and normalization

The spatial encoding stage of our proposal specifies a directional encoding matrix, \( C \), defined by its elements as follows:

\[
p_{\text{dir}}(k) = \sum_{\text{mic}} W_{LS}(k, \theta_{\text{mic}}, \varphi_{\text{mic}}, \theta_{\text{dir}}, \varphi_{\text{dir}}) \psi_t(k, \theta_{\text{mic}}, \varphi_{\text{mic}}). \tag{11}
\]
The spherical harmonics are evaluated at all directions represented in the measuring grid. The encoding matrix \( C \) has a total of \( N^2 + 2M + 1 \) rows and as many columns as there are directions in the measuring grid. An encoding of the sound field can be easily calculated from \( C \):

\[
\hat{B}(k) = C p(k).
\]

The components of vector \( p \) are the results of Eq. (11). Finally, the proposed encoder normalizes the signals \( \hat{B} \). This step prevents artifacts such as clipping or perceptible quantization noise. Normalization is done by multiplying the corresponding Furse-Malham coefficient \( N_{FM}^{nm} \) by each of the signals in \( \hat{B} \). The final encoding produced by our proposal is given by the following expression:

\[
B_{nm}(k) = N_{FM}^{nm} \hat{B}_{nm}(k).
\]

This normalized coefficients \( B_{nm} \) define a MOA encoding of the sound field recorded by the cylindrical microphone array using a full-3D Ambisonic order \( N \) and a horizontal Ambisonic order \( M \).

5. Conclusions

In this paper, a new method to record sound fields and encode them using mixed-order Ambisonics was presented. The proposal uses a cylindrical microphone array to sample the sound field. The cylindrical geometry is more natural to sample the horizontal plane to high detail while retaining a low spatial resolution for elevation. The results of the proposed method are conventional MOA encodings that can be reproduced using any loudspeaker array with a suitable decoder.

Acknowledgements

Parts of this research were supported by a Grant-in-Aid for Scientific Research (A) (no. 24240016) to YS.

References