Study of Higher Order Ambisonic Microphone

Sébastien Moreau[1] and Jérôme Daniel[1]

[1] *France Telecom R&D, 2 avenue Pierre Marzin, 22307 Lannion Cedex, France*
sebastien.moreau@francetelecom.com, jerome.daniel@francetelcom.com

Introduction

Ambisonics is a sound spatialization technology which provides a rational and flexible way to encode and reproduce 3D sound fields on the basis of their spherical harmonics decomposition. Formerly restricted to order 1 (B-Format), it was more recently extended to higher spatial resolutions (High Order Ambisonics or HOA), thus allowing to enlarge the area of acoustic reconstruction [1]. The present paper proposes a way to design Higher Order Ambisonic microphone with a rigid sphere microphone array. Due to small array size, low frequency estimation may imply an excessive bass-boost. To limit it, we introduce a criterion based on the targeted size of the reproduction area.

Natural sound field encoding with HOA

Harmonic decomposition of a 3D sound field

In the spherical coordinate system (azimuth θ, elevation δ, radius r), solving the Helmholtz equation $(\Delta+k^2)p=0$ (with wave number $k=2\pi f/c$, frequency f and sound speed c) in a source-free region of space leads to the so-called Fourier-Bessel decomposition:

$$p(\theta,\delta,kr)=\sum_{m=0}^{\infty} j^m j_m(kr) \sum_{O\le n\le m,\sigma=\pm1} B_{mn}^{\sigma} Y_{mn}^{\sigma}(\theta,\delta) \quad (1)$$

where $Y_{mn}^{\sigma}(\theta,\delta)$ are angular functions called *spherical harmonics*, $j_m(kr)$ are radial functions called *spherical Bessel functions*, and B_{mn}^{σ} are HOA components that describe entirely the sound field (for more details, see [1]). Pressure sound $p(\theta,\delta,kr)$ is expressed from a reference point: the listener point.

To record HOA components, we introduce in the sound field a rigid sphere centered on the reference point, and measure the sound pressure at its surface. For a given radius a, the pressure at the surface can also be expressed in terms of spherical harmonics:

$$p_a(\theta,\phi)=\sum_{m=0}^{\infty} \sum_{0\le n\le m,\sigma=\pm1} W_m(ka) B_{mn}^{\sigma} Y_{mn}^{\sigma}(\theta,\phi) \quad (2)$$

with

$$W_m(ka)=\frac{i^{m-1}}{(ka)^2 h_m^{-}{}'(ka)}.$$

Sound field approximation by truncation

An approximation of a 3D sound field can be achieved by truncating the Fourier-Bessel series (1):

$$\hat{p}_M(\theta,\delta,kr)=\sum_{m=0}^{M} j^m j_m(kr) \sum_{O\le n\le M,\sigma=\pm1} B_{mn}^{\sigma} Y_{mn}^{\sigma}(\theta,\delta) \quad (3)$$

where M is the truncation order. Therefore the number of retained HOA components B_{mn}^{σ} is $K=(M+1)^2$. The spatial area of correct approximation depends on the order M and the frequency (wave number).

To characterize the approximation accuracy, we define the normalized quadratic truncation error ε_M associated to a finite order of truncation M:

$$\varepsilon_M(kr)=\frac{\int_S \left|p(\theta,\delta,kr)-\hat{p}_M(\theta,\delta,kr)\right|^2 dS}{\int_S \left|p(\theta,\delta,kr)\right|^2 dS}, \quad (4)$$

where integrals apply over a sphere S of radius kr.

Encoding natural sound field HOA components

Practically, to encode a 3D natural sound field, we need to *sample spatially the surface of the sphere* and make pressure measures only at a finite number of points Q. Each recorded signal contains a "portion" of spatial components B_{mn}^{σ}. Equation (2) can be rewritten for order of truncation M and for each measure q $(1 \le q \le Q)$:

$$p_a(\theta_q,\phi_q)=\sum_{m=0}^{M} W_m(ka) \sum_{0\le n\le m,\sigma=\pm1} B_{mn}^{\sigma} Y_{mn}^{\sigma}(\theta_q,\phi_q) \quad (5)$$

For correct HOA components estimation, we set Q ≥ K and arrange the Q microphones as uniformly as possible over the surface of the sphere.

In a global matrix formulation, (5) becomes:

$$\begin{pmatrix} p_a(\theta_1,\phi_1) \\ p_a(\theta_2,\phi_2) \\ \vdots \\ p_a(\theta_Q,\phi_Q) \end{pmatrix} = \begin{pmatrix} Y_{00}^1(\theta_1,\phi_1) & Y_{11}^1(\theta_1,\phi_1) & \dots & Y_{mn}^{\sigma}(\theta_1,\phi_1) \\ Y_{00}^1(\theta_2,\phi_2) & Y_{11}^1(\theta_2,\phi_2) & \dots & Y_{mn}^{\sigma}(\theta_2,\phi_2) \\ \vdots & \vdots & \vdots & \vdots \\ Y_{00}^1(\theta_Q,\phi_Q) & Y_{11}^1(\theta_Q,\phi_Q) & \dots & Y_{mn}^{\sigma}(\theta_Q,\phi_Q) \end{pmatrix}$$

$$\times \begin{pmatrix} W_0(ka) & 0 & \cdots & 0 \\ 0 & W_1(ka) & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & W_m(ka) \end{pmatrix} \begin{pmatrix} B_{00}^1 \\ B_{11}^1 \\ \vdots \\ B_{mn}^{\sigma} \end{pmatrix}$$

Or, by shortly namming matrixes and vectors:

$$\mathrm{p}_a = \mathrm{Y.W.B} \quad (6)$$

By inverting this spatial sampling equation, it is possible to estimate HOA components B_{mn}^{σ} from Q microphone signals p_a up to finite order M:

$$B = W^{-1}.(Y^t Y)^{-1} Y^t .p_a \qquad (7)$$

In other words, we can record HOA components of a natural sound field by matrixing microphone signals then applying equalizers $EQ_m = W_m^{-1}$.

The choice of the radius of the spherical array is a compromise. Indeed, distance between microphones introduces high frequency estimation artefacts called *spatial aliasing*. Array radius should be small to confine this artefact above a frequency limit as high as possible. On the other hand, the radius should be large enough to reduce estimation effort that results in excessive bass boost of high order components (dashed curves of Figure 1). Additionally, the next method introduces considerations on reconstruction properties to limit low frequency estimation effort and therefore the bass-boost.

Limiting excessive bass-boost

Excessive bass-boost involved in sound field recording with relatively small microphone array can be interpreted as follows: especially for higher orders and low frequencies, the system tries to catch spatial information that is very poor at the measurement points and is substantial only at a distance from the microphone array. Now it appears from spherical Bessel functions $j_m(kr)$ in equation (1) that, considering the sound field reconstruction up to a given radius, higher order B_{mn}^{σ} become unnecessary at relatively low frequencies. Therefore we introduce a criterion to characterize useless low frequency band of HOA components depending on the order and targeted reconstruction area dimensions. This criterion is based on the quadratic error ε_M. Excessive amplification in EQ_m can be limited in these frequency bands thanks to high-pass filters.

Characterizing HOA components usefulness for a given radius of representation area

We assume that microphone array dimensions are small compared to the distance of sources and thus, that recorded waves are plane. Therefore, the normalized error ε_M becomes [2]:

$$\varepsilon_M(kr) = 1 - \sum_{m=0}^{M} (2m+1)(j_m(kr))^2 \qquad (8)$$

Note that ε_M is independent of the source direction. By considering a targeted radius r_t of reproduction area, ε_M can be expressed as a function of frequency. For a given error ε, we can deduce from (8) a corresponding limit frequency $f_{lim}^{(M)}$ for each order. Then, we consider that low frequency band with upper limit $f_{lim}^{(M)}$ of M order component don't significantly contribute to the sound field in the targeted reproduction area compared to lower orders.

Application to sound field recording by high-passing HOA components

Low frequency bands defined previously don't need excessive amplification regarding to their usefulness. We can therefore associate high-pass filters H_m to EQ_m (Fig. 1). These filters must compensate (or even inverse) EQ_m slopes in useless low frequency bands. Another essential requirement is that the filters have to preserve EQ_m phase in the pass-band, *i.e.* that they must have a linear phase response. If the resulting amplification is judged still excessive, notably regarding to noise amplification, ambitions of estimation must be reduced by choosing a smaller target radius or a bigger error ε. Note that noise amplification is not only determined by EQ_m but also by matrixing operations.

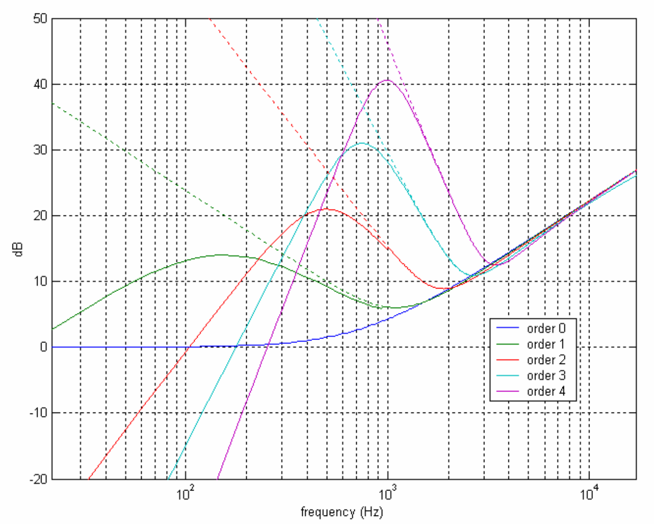

Figure 1: Order 0 to 4 HOA components equalization EQ_m (dashed lines) associated to high-pass filters (cont. lines) for $a = 0{,}25$m, $r_t = 0.25$m, $\varepsilon = 0.25$.

Conclusion

The present paper introduced a practicable way to encode natural 3D sound fields with Higher Order Ambisonics. It consists in recording signals at the surface of a rigid sphere and applying on them operations of matrixing and filtering that theoretically contain excessive bass-boost due to small microphone array dimensions. This excessive amplification is compensated thanks to high-pass filters which discard useless frequency bands of HOA components. These useless frequency bands are determined by a targeted radius of reproduction area and an error criterion. A 4th order prototype is actually under validation in our laboratories.

References

[1] J. Daniel, R. Nicol, and S. Moreau (2003) "Further Investigations of Higher Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging", *AES 114th Convention*, Amsterdam

[2] D.B. Ward, T.D. Abhayapala (2001) "Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers", *IEEE Trans. on Speech and Audio Proc.*, vol. 9, n°6, pp 697-707