A triple microphone array for surround sound recording

Rilin CHEN[1]; Pengxiao Teng[2]; Yichun YANG[3]

[1, 2, 3] Institute of Acoustics, Chinese Academy of Sciences, China

ABSTRACT

B-format signal can be directly by the gradient microphones. In this paper, a small size triple microphone for surround sound recording is proposed, where every two components of triple microphone array compose a differential microphone array. Some direct and indirect methods for B-format signal are discussed. With the virtual microphones method, the loudspeakers feeding signal will be obtained.

Keywords: Surround Sound, Differential Array, Recording,

1. INTRODUCTION

Audio capture is an important issue in the many different application field, such as teleconference, mobile phones, music recording and etc.. It is noted that stereo signal capture is now the key feature of different brand cell phones, which are one of the most common electric devices. With the popularization of 4G communication or even development of the 5G communication in future, the communication bandwidth will be much wider and not only stereo but also surround communication will be realized.

Sound field microphone (1) is firstly applied in the 70's, which will export the B-format signals W, X, Y, and Z that is the first order harmonics. With these signal, the sound field will be reconstructed by the decoder (2, 3). Later, higher order ambisonics has been further researched for 2D and 3D accurate reproduction and extension of listening area by Daniel (4), Elko (5) and etc.. Also, gradient microphone (6, 7) with the character of spatial selectivity is often used in stereo and surround recording. For the small size of the smart phones and other mobile devices, microphones above are unpractical to be used in these devices. Therefore, omnidirectional sound pressure microphones will be a good choice to be built into these devices. Throughout this paper, we will focus on 2D sound field recording with first-order directive microphones techniques which is known as differential microphone array (8, 9], where two or more omnidirectional microphones in end-fire with a spacing combined to obtain direction selection. The technique is widely used in communication (10), sound capture (11), hearing aid (12) and sound field recording (13, 14).

In this paper, we will first review the first order differential array. Next, the different array type is designed to obtain the B-format signals for ambisonics. Also, our triple array type is proposed and the signal processing is described.

2. Analysis of Differential Microphone Array

2.1 Construction of Differential Array

For the construction of a first-order differential array, figure 1 present the subtraction of 2 omnidirectional microphones with inter element spacing d. Supposing the signal at point O is $m(\boldsymbol{r}_O, t)$, the signal at position $-d/2$ and $d/2$, where $d \ll \lambda$, can be expressed as

$$m_1(\boldsymbol{r}_{1,}, t) = m(\boldsymbol{r}_{O,}, t) exp\left(j\frac{kd}{2}cos\theta\right), \tag{1}$$

[1] chenrl@mail.ioa.ac.cn
[2] px.teng@ mail.ioa.ac.cn
[3] yychun@ mail.ioa.ac.cn

$$m_2(\boldsymbol{r}_{2,}, t) = m(\boldsymbol{r}_{O,}, t) exp\left(-j\frac{kd}{2}cos\theta\right), \tag{2}$$

with $\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_O$ being the positions of 2 microphones and the middle position, k the wave number, θ the incident angle. The subtraction of two microphones is

$$\begin{aligned} m_1(\boldsymbol{r}_1, t) - m_2(\boldsymbol{r}_2, t) &= m(\boldsymbol{r}_{O,}, t)\left(exp\left(j\frac{kd}{2}cos\theta\right) - exp\left(-j\frac{kd}{2}cos\theta\right)\right), \\ &= 2jsin\left(\frac{kd}{2}cos\theta\right) m(\boldsymbol{r}_{O,}, t) \end{aligned} \tag{3}$$

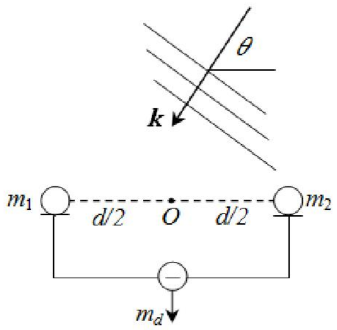

Figure 1 – First order differential array composed of two omnidirectional microphones.

The response of the first order differential array is

$$E_d(\theta) = 2jsin\left(\frac{kd}{2}cos\theta\right), \tag{4}$$

While d <<λ, the response will be simplified as follows

$$E_d(\theta) = cos\theta\ , \tag{5}$$

The solid line in figure 2 shows the pattern of figure eight. Then, the other patterns such as cardioid, supercardioid, hypercardioid, and subcardioid can be designed by the combination of pattern of omnidirectional and figure eight (6) with the different value of α.

$$E_d(\theta) = \alpha + (1-\alpha)cos\theta, \tag{6}$$

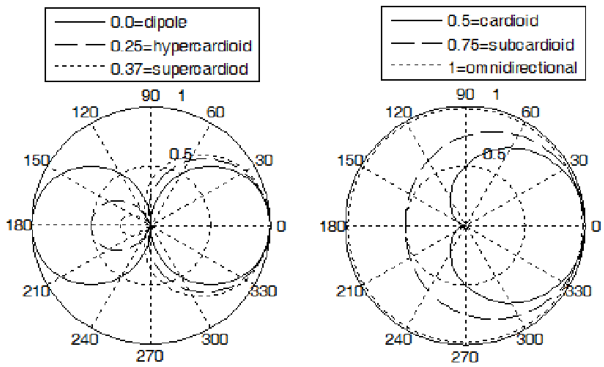

Figure 2 – Patterns of different differential array of first order.

2.2 Compensation for Differential Array

While the spacing d <<λ, the pattern type is independent with frequency just as Eqn. (5). Otherwise, as Eqn. (4), the pattern is relevant with frequency. For the wave number $k = \frac{2\pi}{\lambda}$, the Eqn. (4) can be written as

$$E_d(\theta) = 2jsin\left(\pi\frac{d}{\lambda}cos\theta\right), \tag{6}$$

Figure 3 depicts the pattern change with the different ratio of d and λ. It suggests that the value of $\frac{d}{\lambda}$ should be chosen below 0.5. Without loss of generality, the value of d is assumed to be 0.02 m, the frequency responses of the first order differential array with different incident angles will be described in figure 4. It shows the high pass character up to the frequency $f_m = \frac{c}{2d}$, where c is the speed of sound, so the compensation for low frequency is needed to keep recording distortionless. Here, f_m is 8500 Hz for d=0.02 m and c=340m/s.

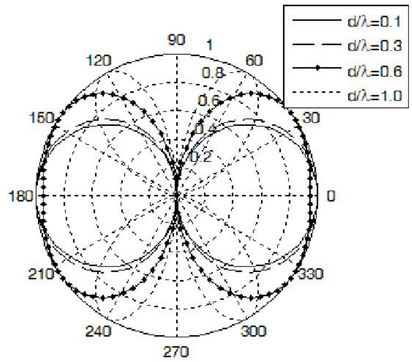

Figure 3 – Patterns change with the ratio d/λ

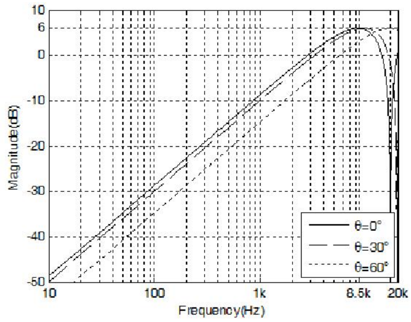

Figure 4 – Frequency responses of first order differential array with different incident angles.

The compensation can be calculated by

$$H = \begin{cases} \min\left(G, \left(2jsin\left(\frac{kd}{2}\right)\right)^{-1}\right), & f < f_m \\ \left(2jsin\left(\frac{\pi f_m d}{c}\right)\right)^{-1}, & f \geq f_m \end{cases}, \tag{7}$$

where G, here 35 dB is chosen, is the maximum gain of the compensation filter to avoid over-amplifying for low frequency. And, the compensation gain for high frequency is set to equal to the gain at frequency f_m. The detailed discussion can be referred to (14).

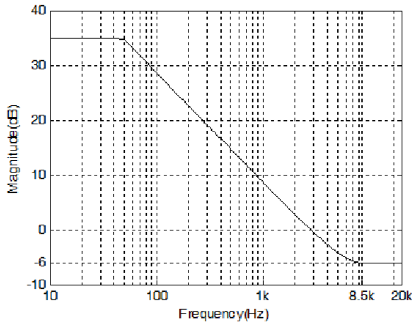

Figure 5 – Compensation filter for first order differential array

3. Microphone Array for Surround Sound Recording

3.1 Four-microphones Array for Recording

With the method of delay and subtraction, the superdirective differential array steering to some direction will be designed. Different value of α result to different pattern types. And, the cardioid type can be obtain by

$$x_1(t) = h(t) * \left(m_1(\boldsymbol{r}_1, t) - m_2(\boldsymbol{r}_2, t-\tau)\right), \tag{8}$$

$$x_2(t) = h(t) * \left(m_2(\boldsymbol{r}_2, t) - m_1(\boldsymbol{r}_1, t-\tau)\right), \tag{9}$$

Here, τ is the travel delay between two omnidirectional microphones, which can be calculated as follows,

$$\tau = \frac{d}{c}\frac{\alpha}{1-\alpha}. \tag{10}$$

And, $h(t)$ is the compensation filter in time domain for Eqn. (7), the symbol $*$ denotes linear convolution, $x_1(t)$ and $x_2(t)$ are the differential signal pointing to the left and right directions respectively (13). Then, two omnidirectional microphones will get the signal from left and right directional and the stereo recording is achieved. Similarly, four microphones in figure 6 can be used

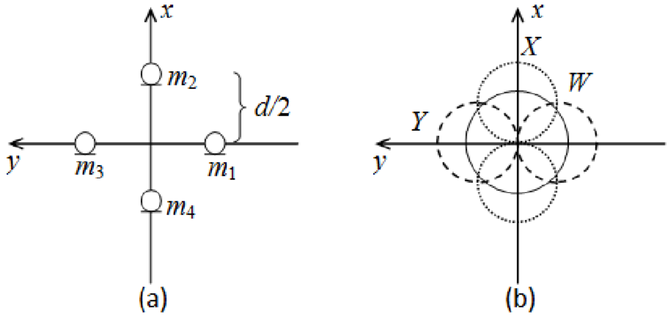

Figure 6 – Four-microphones array for surround recording (a) and its B-format signals (b)

to get the B-format signalss W, X, and Y, where

$$X = h(t) * \left(m_2(t) - m_4(t)\right), \tag{11}$$

$$Y = h(t) * \left(m_3(t) - m_1(t)\right), \tag{12}$$

$$W = \frac{1}{4}\sum_{i=1}^{4} m_i(t), \tag{13}$$

It is desirable to acquire figure eight pattern with the opposite two-microphone pair, where m_2 and m_4 to get the X signal, m_1 and m_3 to get the Y signal. The W signal is not recorded by any microphone, but it is replaced by the mean of four microphones to instead of the microphone at the centre of the array.

3.2 Triangle Microphone Array for Recording

3.2.1 Special Triangle Microphone Array

In the design of the mobile device like cell phones, four microphones may be too many and it is not very flexible to configure the microphones for its explicit placement. Instead, triple microphone array is a flexible configuration that three microphones can be placed at any position to form a triangle. It is quite nature to constitute a right triangle firstly as figure 7(a), which is easily to be positioned on the cell phone. As Eqns. (11), (12), (13), the B-format signals will be calculated by

$$X = h_x(t) * \left(m_1(t) - m_2(t)\right), \tag{11}$$

$$Y = h_y(t) * \left(m_2(t) - m_3(t)\right), \tag{12}$$

$$W = \frac{1}{3}\sum_{i=1}^{3} m_i(t), \tag{13}$$

where $h_x(t)$, $h_y(t)$ are the compensation filter for X and Y. It is obvious that W, X and Y here are not recorded at the same point, as shown in figure 8. Therefore, the sound field reconstruction will not be accurate.

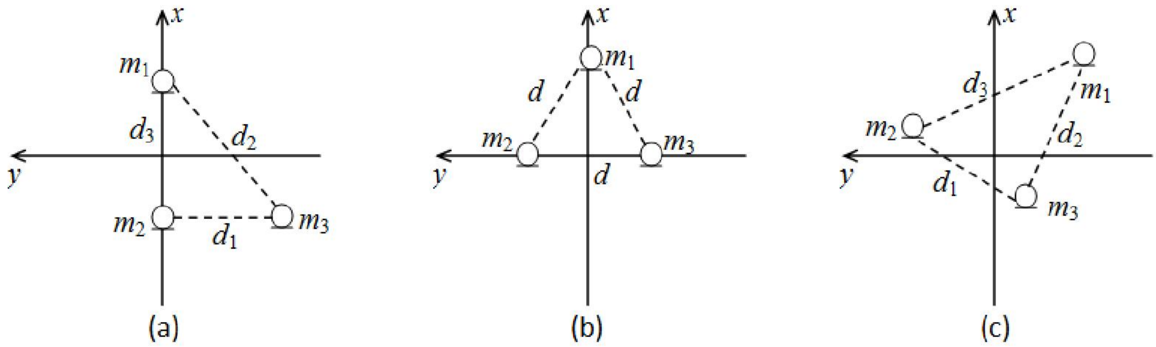

Figure 7 – Different triple microphone array.
(a) Right triangle; (b)Equilateral triangle; (c)General triangle.

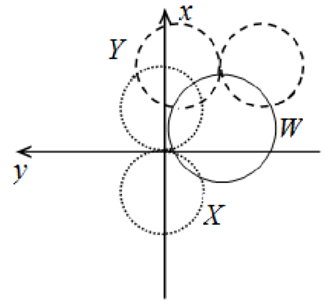

Figure 8 – B-format signals by right triangle which is not on the same position

The better choice of array type is equilateral triangle. The Y and W signals are the same with Eqns. (12) and (13), but the X signal cannot be obtained straightly. The differential signal along m1 and m2 or m1 and m3 is

$$C_1 = h(t) * \left(m_1(t) - m_2(t)\right) \tag{14}$$

$$\text{or } C_2 = h(t) * \left(m_1(t) - m_3(t)\right). \tag{15}$$

The patterns in the direction m_1m_2 and m_1m_3 are $cos(\theta - \psi)$, with ψ $-30°$ and $30°$ respectively, so

$$X = \frac{2C_1 + Y}{\sqrt{3}} \tag{16}$$

$$\text{or } X = \frac{2C_1 - Y}{\sqrt{3}} \ . \tag{17}$$

There will be only one of Eqns. (14) and (15) to be calculated for B-format signals *X*, which means that only two microphone pairs will be utilized while information from the third pair will not be exploited.

3.2.2 **General Triple Microphone Array**

Also, as in the figure 7(c), the third equation as Eqns. (14) and (15) can be written, while any two of these three equation can be used to acquire the B-format signals *X* and *Y*. Nevertheless, it is known that the third pair information will be benefit for reproduction.

3.3 Virtual Microphone to B-format signals

In the decoder of B-format, if B-format signals is known, the virtual microphone or speaker signal will be resulted (15). The three virtual microphone signals in figure 7 can be described as

$$\boldsymbol{C} = \mathbf{D}\boldsymbol{B}, \tag{18}$$

where D is the decoder matrix, B is the B-format signals and C is the virtual signal, and

$$\mathbf{D}=\begin{pmatrix} \frac{\sqrt{2}}{2} & \cos\theta_1 & sin\theta_1 \\ \frac{\sqrt{2}}{2} & \cos\theta_2 & sin\theta_2 \\ \frac{\sqrt{2}}{2} & \cos\theta_3 & sin\theta_3 \end{pmatrix}, \tag{19}$$

$$\boldsymbol{B} = \begin{pmatrix} W \\ X \\ Y \end{pmatrix}, \tag{20}$$

$$\boldsymbol{C} = \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix}, \tag{21}$$

where $\theta_i (i = 1,2,3)$ are the angles between the directions of virtual cardioids microphones $C_i (i = 1,2,3)$ and positive *x* axis. Then, B-format signals are

$$\boldsymbol{B} = \mathbf{D}^{-1}\boldsymbol{C}. \tag{22}$$

As Eqns. (8, 9), two virtual cardioids microphones pointing to opposite direction will be produced with two microphones, so three microphones will generate six virtual microphones, three of which will be chosen for calculation.

$$C_1(t) = h_1(t) * \left(m_1(t) - m_2(t - \tau_1)\right), \tag{23}$$

$$C_2(t) = h_2(t) * \left(m_2(t) - m_3(t - \tau_2)\right), \tag{24}$$

$$C_3(t) = h_3(t) * \big(m_3(t) - m_1(t - \tau_3)\big). \tag{25}$$

where $h_i(t)(i = 1,2,3)$.are the compensation filters for virtual microphones with spacing $d_i(i = 1,2,3)$ and $\tau_i = \frac{d_i}{c}(i = 1,2,3)$.

3.4 B-format to Loudspeaker Array

Without loss of generality, while $\theta_1 = -30°, \theta_2 = 90°, \theta_3 = 210°$, it is the case of figure 7(b), the differential patterns of which is depicted in figure 9. The B-format signals will be gained by Eqn. (22) and it will be exploited to produce the signal to fed loudspeaker array by Eqn. (18). The decoders for 2.0, 4.0 and 5.1 and other loudspeaker systems have been researched for many years (1, 3, 15).

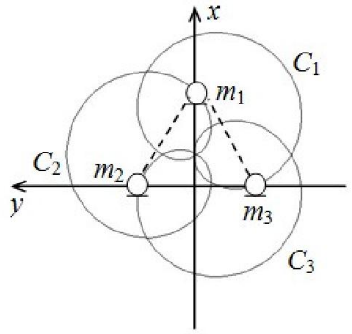

Figure 9 – Virtual microphone pattern of equilateral triangle array

4. Future work

In this paper, small size microphone configuration for surround sound recording in the mobile device is considered. For less microphone and flexible processing, a triple microphone array and its encoder for B-format signal is proposed. With this method, surround sound recording can be realized for the mobile device. In the future, the B-format signal recorded by the method proposed here should be applied in the sound field reproduction.

REFERENCES

1. Beranek LL. Criteria for noise and vibration in buildings and communities. In: Beranek LL, Ver IL, editors. Noise and Vibration Control Engineering - Principles and Applications. 2nd ed. New York, USA: Wiley; 2006. p. 887-909.
2. Widrow B, Stearns SD. Adaptive Signal Processing. Englewood Cliffs, New Jersey, USA: Prentice-Hall; 1985.
3. Embleton TFW. Tutorial on sound propagation outdoors. J Acoust Soc Am. 1996;100(1):31-48.
4. Nishimura M, Kondo M, Iwamoto K, Tsujimoto Y. Active attenuation of impulsive noise from a railway tunnel exit. Proc INTER-NOISE 94; 29-31 August 1994; Yokohama, Japan 1994. p. 1383-8.
5. Stinson M, Daigle G. Meteorological measurements for use in sound propagation calculations. Proc 7th Long Range Sound Propagation; 24-26 July 1996; Ecully, France 1996. p. 137-47.

1. M. A. Gerzon. The design of precisely coincident microphone arrays for stereo and surround sound. In: Preprint 50th Conv. Aud. Eng. Soc., Mar. 1975.
2. Michael A. Gerzon. Design of Ambisonic Decoders for Multispeaker Surround Sound. In: 58th Audio Engineering Society Convention, N.Y., Nov. 1977
3. Michael A. Gerzon. Ambisonic decoders for HDTV. In: 92nd Convention of the Audio Engineering Society, Vienna, Austria, 1992.
4. J. Daniel. Acoustic field representation, application to the transmission and the reproduction of complex sound environments in a multimedia context, Ph.D. thesis, Universit´e Paris 6, 2000, French.
5. Meyer J. and Elko G.W. A spherical microphone array for spatial sound recordings. J. Acoust. Soc. Am., 2002, 111(5): 2346-2346.
6. Harry F. Olson. Gradient microphones. J. Acoust. Soc. Am.,1946, 17(3): 192–198.
7. Harry F. Olson. Directional microphones. Journal of the Audio Engineering Society, 1967, 15(4): 420–430.

8. G.W. Elko and A.T. Nguyen Pong. A simple first-order differential microphone. In: Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), Mohonk Mountain Resort (NY), Oct. 1995, p. 169-172.
9. R.M,M. Derkx and C.P, Janse. Theoretical analysis of a first-order azimuth-steerable superdirective microphone array. IEEE Trans. On Audio, Speech and Language Processing, 17(I), 2009, p. 150-162.
10. G. W. Elko and Anh-Tho Nguyen Pong. A steerable and cariable first-order differential microphone array. In: IEEE International Conference on Acoustics, Speech, and Signal Processing ,1997, p. 223-226.
11. R.M.M. Derkx. First-order superdirectional acoustic zooming in the presence of directional interferences. In: 2011 Joint Workshop on Hands-free Speech Communication and Microphone Arrays, p. 176-179.
12. D. Preves, T. Peterson, and M. Bren. In-the-Ear Hearing Aid with Directional Microphone System. U.S. patent 5, 1998, p. 757-933.
13. Christof Faller. Conversion of two closedly spaced omnidirectional microphone signals to XY stereo signal. In: AES 129th Convention, San Francisco, CA, USA, 2010, November, 4-7.
14. Mihailo Kolundzija, Christof Faller, Martin Vetterli. Sound field recording by measuring gradients. In: AES 128th Convention, London, UK, 2010, May 22-25.
15. B. Wiggins. Has ambisonics come of age?. In: Proceedings of the Institute of Acoustics, 30(6). 2008.