

VIRTUAL SOURCE PANNING USING MULTIPLE-WISE VECTOR BASE IN THE MULTISPEAKER STEREO FORMAT

Se-Woon Jeon, Young-cheol Park*, Seok-Pil Lee**, and Dae Hee Youn

School of Electrical & Electronic Engineering, Yonsei University, Seoul, Korea

*Computer & Telecommunications Engineering Division, Yonsei University, Wonju, Gangwon, Korea

**Korea Electronics Technology Institute (KETI), Seoul, Korea

phone: + (82) 2-2123-4534, fax: + (82) 2-364-4870, email: jdotsw@dsp.yonsei.ac.kr

ABSTRACT

In the last few decades, various panning algorithms have been proposed to generate virtual sound localization in loudspeaker systems. Vector base amplitude panning (VBAP) is the most widely adopted pair-wise amplitude panning method. However, pair-wise amplitude panning has a directional discontinuity problem in the multichannel surround panning. In particular, sound localization for the virtual source does not smoothly vary in the direction of the near speaker. While coincident panning using multiple loudspeakers, such as Ambisonics, performs better, it has less stability in sound localization due to the *precedence effect*. In this paper, a multiple-wise vector base virtual source panning algorithm is proposed. To generate more stable localization of panning sound, the proposed panning algorithm calculates the amplitude panning gains using multiple-wise vector base formulation. Additionally, the angle-dependent and nonnegative gain control function is applied to prevent artifacts caused by negative amplitude gains. The subjective listening tests using the method of adjustment (MOA) are performed to evaluate the sound localization of the proposed algorithm, which is compared with the conventional amplitude panning methods.

1. INTRODUCTION

High-definition (HD) and ultrahigh-definition (UHD) video formats, which provide high-quality optical resolution and a wider angle view, are rapidly being applied to various multimedia systems. The audio format for HD and UHD video must generate a wider sound field and more immersive sound effects in order to optimally accompany these video formats. However, the conventional stereo system cannot live up to the desired sound quality in the latest multimedia systems. The effort to generate a spatial localization of the sound source in the multichannel loudspeaker system has tried and various sound reproduction techniques had been suggested. Using one loudspeaker for one sound source can create the most natural and point-like sound source. But it is difficult to construct such a system because of spatial restrictions or the insufficiency of the loudspeakers. Thus, the panning algorithms based on the psychoacoustic theorem, e.g. *summing localization*, were proposed to acoustically create a directional perception of the sound source in the spatial sound stage [1]-[9].

The most popular panning technique is pair-wise amplitude panning. When two signals with level differences due to gain factors emanate from two loudspeakers, the sound localization of the virtual source is perceived by the listener

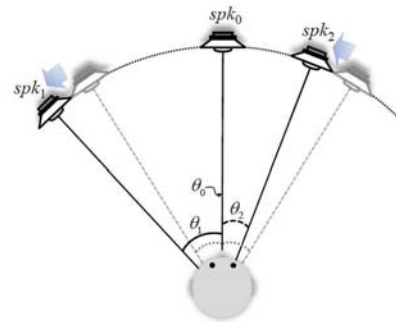


Figure 1: Arbitrary angle-positioned multispeaker stereo format.

as being on the opposite side of the loudspeakers. The directional position of the virtual sound source is controlled by the amplitude level difference in the horizontal stage between two loudspeakers. The sine law and the tangent law are basic pair-wise amplitude panning methods in the stereophonic loudspeaker format [2, 3]. In 1997, Pulkki proposed vector base amplitude panning (VBAP), which was not only available with a symmetric stereo layout, but also for any angle-positioned pair-wise loudspeaker setup [4]. However, pair-wise amplitude panning, including VBAP, has the disadvantage that the virtual source is pulled in the direction of the near speaker; this is called the *detent effect* [5]. It also has a discontinuity problem with respect to the directional spread of virtual source [6].

Pulkki proposed multiple-direction amplitude panning (MDAP) to resolve the problem of pair-wise amplitude panning [6]. MDAP generates multiple panning vectors to make the listener perceive a single virtual source. With MDAP, the virtual vector of one virtual source is made by the superposition of two or more virtual vectors panned by VBAP. As a result, the local minima of directional spread at the near-speaker position can be prevented. The localization performance of MDAP has not been fully certified yet. Actually, the localization performance of VBAP has been known to be worse in the lateral side and the rear than in the frontal stage in an angle-dependent manner [7]. Therefore, the localization of a single-perceived virtual source through use of the multiple vectors in MDAP can be degraded by the lateral panning vector.

In Ambisonics, spatial localization of the virtual source is generated by using all of the multiple loudspeakers around the listener [8, 9]. The gain factors of Ambisonics are allowed to be both positive and negative. It can achieve more

correct localization of virtual source than pair-wise amplitude panning. But the directional quality is less robust to the listener outside of the best position due to the *precedence effect* [10]. The minus gain can also cause the out-of-phase problem [8]. Currently, Poletti proposed the panning functions for surround sound systems with nonuniform loudspeaker layout; these panning functions produced a robust sound-field directionality in the ITU 3/2 configuration [11]. It resulted better performance for sound localization, but required additional computation in the asymmetric and arbitrary angle-positioned loudspeakers layout.

In this paper, we present a multiple-wise vector base panning algorithm that maintains the advantages of VBAP, but provides more accurate localization of the virtual source by concurrently using multiple loudspeakers. The proposed panning algorithm is inspired from the vector base formulation of Pulkki's VBAP, with a motive of Gerzon's panpot law in the multispeaker stereo setup as shown Fig. 1 [4, 5]. The main advantages of the proposed algorithm are its stability of the bias problem due to the *precedence effect* in the sound localization, and the simplicity of the panning gain calculation for any arbitrary multichannel format. To prevent the out-of-phase problem due to negative panning gain, a non-negative gain control function is applied to the vector base formulation of the gain factors.

The organization of paper is as follows. First, the multispeaker stereo setup and the vector base panning algorithms are introduced in Section 2. The formulation of the proposed panning algorithm with application of a nonnegative center channel gain function is described in Section 3. In Section 4, we evaluate and discuss the sound localization performance of the proposed panning algorithm by comparing conventional panning techniques using subjective listening tests. Finally, it is summarized and concluded in Section 5.

2. MULTIPLE LOUDSPEAKERS FORMAT AND VECTOR BASE AMPLITUDE PANNING METHODS

2.1 Multispeaker stereo setup

To obtain better performance of the spatial localization in the frontal stage, the center loudspeaker is occasionally added in the middle of the stereo format, e.g. the ITU 3/2 multichannel format. In this format as shown in Fig. 1, the panning algorithm of the frontal multiple loudspeakers should be changed. Gerzon defined such a format as a multispeaker stereo format, and also calculated the optimal panning gain, that is, the panpot law for the multispeaker stereo, to satisfy the coincident localization of the velocity vector and the energy vector [5]. Gerzon's panpot law was certified as the optimal computation method, but it does assume that the loudspeaker angle for the left and the right speakers should be same relative to the center direction, $\theta_1 = \theta_2$ of Fig. 1. Similarly, the loudspeaker format in Ambisonics must be symmetric to achieve optimal results [10]. The proposed amplitude panning algorithm is supposed that the loudspeakers' configuration is arbitrary with respect to angle positioning, that is, the layout could be either symmetric or asymmetric.

2.2 Vector base amplitude panning methods

The vector base formulation to calculate the panning gain is introduced by Pulkki [4]. In the two-dimensional plane, the vector matrix \mathbf{L}_{12} , which consists of a pair of loudspeaker

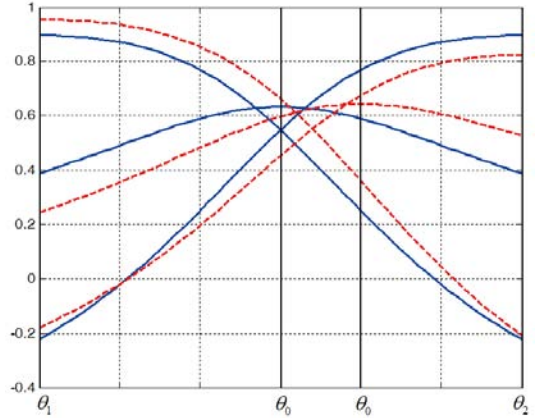


Figure 2: Optimal multiple-wise panning gain by *pseudo-inverse* matrixing; *solid line* - symmetric layout ($|\theta_1 - \theta_0| = |\theta_2 - \theta_0|$) and *dashed line* - asymmetric layout ($|\theta_1 - \theta_0| \neq |\theta_2 - \theta_0|$).

vectors, is multiplied by the gain vector \mathbf{g} and forms the virtual source vector $\mathbf{p} = [p_1 \ p_2]^T$.

$$\mathbf{p}^T = \mathbf{g} \mathbf{L}_{12}, \quad \text{where } \mathbf{g} = [g_1 \ g_2], \mathbf{L}_{12} = [l_1 \ l_2]^T. \quad (1)$$

The panning gain of two loudspeakers in the two-dimensional space can be calculated by the inverse matrix of \mathbf{L}_{12} as

$$\mathbf{g} = \mathbf{p}^T \mathbf{L}_{12}^{-1}. \quad (2)$$

In the multichannel format, the surround panning gains are computed by VBAP's formulation with paired loudspeakers. In the multispeaker stereo format as Fig. 1, the loudspeaker pairs are different at left stage ($\theta_1 \sim \theta_0$) and right stage ($\theta_0 \sim \theta_2$). The directional perception of the virtual source is point-like in the direction of the center loudspeaker. Point-like perception of the virtual sound source is the best result of sound localization, but can be affected by the directional spread, such as spatial blurring. The directional spread increases in further directions to the loudspeaker [6, 10]. It means that the directional spread varies with panning direction. For example, when the virtual source moves across the center loudspeaker in the horizontal plane, the variation of directional spread causes a degradation of spatial sound quality.

In MDAP, the variation problem of directional spread is improved by generating the virtual source using multiple loudspeakers [6]. It is implemented by the superposition of multiple-panned virtual vectors. However, the pair-wise panning of VBAP was verified as the sound localization is worse in the lateral or rear sides than in the frontal stage [7]. As a result, the multiple directional-vectors created by MDAP result in the different accuracy of localization. Thus, the perception of the virtual source can be biased in the direction of the center because of the localization error for the more laterally positioned virtual vector. This effect is verified by the subjective listening test in Section 4.

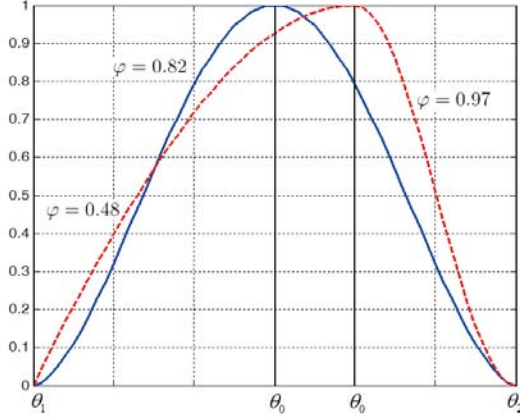


Figure 3: Angle-dependent gain control function for nonnegative panning gain; *solid line* - symmetric layout ($|\theta_1 - \theta_0| = |\theta_2 - \theta_0|$) and *dashed line* - asymmetric layout ($|\theta_1 - \theta_0| \neq |\theta_2 - \theta_0|$).

3. PROPOSED AMPLITUDE PANNING ALGORITHM USING MULTIPLE-WISE VECTOR BASE FORMULATION

In this section, we introduce the vector formulation to compute the optimal multiple-wise panning gain in the arbitrary multispeaker stereo format and the nonnegative and controllable angle-dependent function to prevent the out-of-phase effect by negative panning gain.

3.1 Multiple-wise vector base amplitude panning

First, to calculate the optimal multiple-wise panning gain, we defined the multiple-vector matrix \mathbf{L}_{012} to compute the coincident panning gain for the multispeaker stereo format. The vector of virtual source \mathbf{p} is then decided by the angle-dependent configuration of loudspeakers \mathbf{L}_{012} with the vector of panning gain \mathbf{g} as

$$\mathbf{p}^T = \mathbf{g}^T \mathbf{L}_{012}, \quad \text{where } \mathbf{g} = [g_0 \ g_1 \ g_2], \mathbf{L}_{012} = [l_0 \ l_1 \ l_2]^T. \quad (3)$$

It is similar to Eq. (1), but we employ *pseudoinverse* method, also known as the Moore-Penrose inverse matrix, to compute the inverse of non-regular square matrix \mathbf{L}_{012} [12]. By using the multiple-wise vector base formulation, the optimal panning gain vector of multiple loudspeakers is

$$\mathbf{g} = \mathbf{p}^T \mathbf{L}_{012}^+, \quad (4)$$

where the *pseudoinverse* matrixing is $\mathbf{X}^+ = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$. As a result, the optimal panning gains in the loudspeaker format of Fig. 1 are described in Fig. 2 with energy normalization as

$$\mathbf{g}^{\text{scaled}} = \frac{\mathbf{g}}{\sqrt{g_0^2 + g_1^2 + g_2^2}}. \quad (5)$$

3.2 Nonnegative gain control function

In Section 3.1, the optimal panning gains using a multiple-vector base formulation were calculated for any loudspeakers format, i.e. symmetric layout or asymmetric layout, as

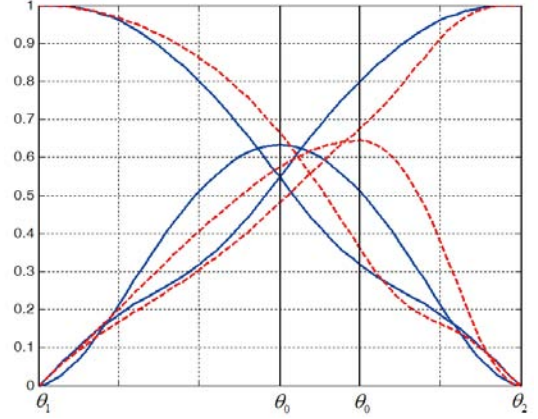


Figure 4: Proposed multiple-wise panning gain with nonnegative gain control function; *solid line* - symmetric layout ($|\theta_1 - \theta_0| = |\theta_2 - \theta_0|$) and *dashed line* - asymmetric layout ($|\theta_1 - \theta_0| \neq |\theta_2 - \theta_0|$).

VBAP [4]. But in Eq. (4) the panning gain at far directions from the loudspeaker is negative as Fig. 2. Panning algorithms using multiple loudspeakers, e.g. Ambisonics, were proposed to achieve the best perception of sound localization [5, 9]. Especially, Gerzon introduced the optimal panning gain to satisfy the perceptual angle of Gerzon vectors in the multispeaker stereo setup [5]. To obtain the optimal gain factors for the listener in the middle of loudspeakers, these panning techniques allow for negative panning gains for the loudspeakers far from the direction of virtual source. However, negative panning gain occasionally causes an out-of-phase effect, degrading the sound localization. To resolve this problem, an in-phase panning algorithm that only used positive panning gains was proposed in Ambisonics [8].

Therefore, we additionally applied the angle-dependent gain control function in the procedure of vector formulation of Eq. (4) to obtain the nonnegative panning gain.

$$\mathbf{L}'_{012} = \mathbf{W}(\theta) \mathbf{L}_{012}, \quad \text{where } \mathbf{W}(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \delta(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (6)$$

In Eq. (4), the vector matrix of loudspeakers, \mathbf{L}_{012} , is substituted with \mathbf{L}'_{012} of Eq. (6). It means that the vector of the middle loudspeaker is angle-dependently regulated by a nonnegative gain control function, $\delta(\theta)$. This control function is defined as

$$\delta(\theta) = (0.5 - 0.5 \cos(\frac{2\pi\theta}{N-1}))^\varphi, \quad (7)$$

where $N = 2|\theta_0 - \theta_p|$, $p = 1$ or 2 .

The panning angle θ is $[-|\theta_0 - \theta_p|, +|\theta_0 - \theta_p|]$. In Fig. 3, the angle-dependent nonnegative gain control function in the multispeaker stereo format as like Fig. 1 is described. This function gradually reduces the loudness of the middle loudspeaker as the panning angle of the virtual source is further from the center. The function also very effectively prevents the negative panning gain and is also useful for incorporating minute adjustments of the virtual source's panning angle.

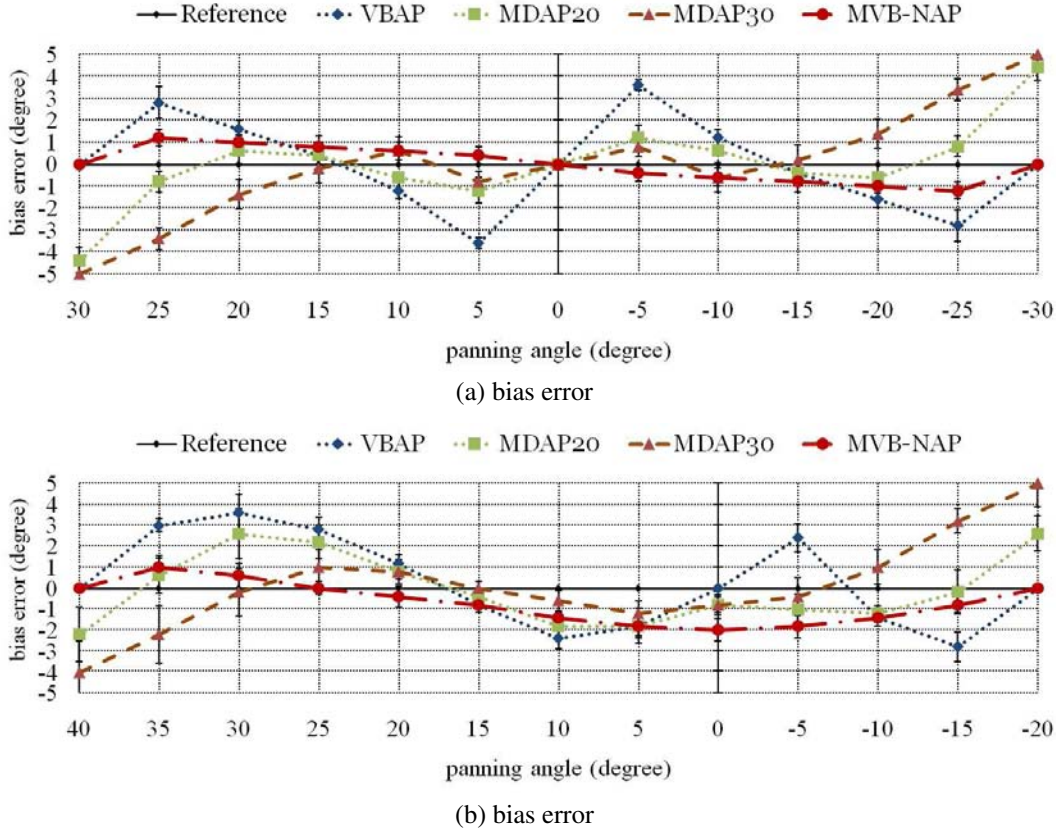


Figure 5: Subjective test results; bias error (degree) of panning angle. (a) symmetric layout (30°, 0°, -30°) (b) asymmetric layout (40°, 0°, -20°).

The optimal value of φ , in Eq. (7) can be decided by the configuration of loudspeakers. If the angle-interval of the loudspeakers, i.e. $|\theta_1 - \theta_0|$ or $|\theta_0 - \theta_2|$, increases, the larger value is perceptually better to improve the quality of sound localization from degradation by the out-of-phase or by the *detent effect*. So, in the asymmetric layout, i.e. $|\theta_1 - \theta_0| \neq |\theta_2 - \theta_0|$ in Fig. 1, the value φ of left side from the centered loudspeaker is experimentally decided to be larger than the right side as like Fig. 3. In Fig. 4, the proposed panning gain results for the symmetric and asymmetric layouts applied with nonnegative center gain control function are described. The formulation of the proposed panning algorithm is derived from the vector form of multiple loudspeakers and it is termed multiple-wise vector base nonnegative amplitude panning (MVB-NAP).

4. PERFORMANCE EVALUATIONS AND DISCUSSIONS

In the subjective listening tests, sound localization performance was evaluated according to the method of adjustment (MOA) [7, 13]. The test signal consisted of 2 s pink noise. The reference signal was emanated in the desired panning direction by monophonic loudspeakers that had an angle interval of 5 degrees (°) between the left loudspeaker at θ_1° and the right loudspeaker at θ_2° . The loudspeakers for the multispeaker stereo were positioned as Fig. 1 and the distance from the listener to each loudspeaker was equally 1.5 m. The five experienced audio engineers participated in the

tests. They were asked to adjust the panning angle with 1 degree resolution to match the direction of reference by self-manipulating the control buttons in PC software as shown in Fig. 6.

The three conventional panning methods, i.e. pair-wise VBAP and MDAP with 20° or 30° spread angle, and the proposed panning method, MVB-NAP, were compared [4, 6]. The results of subjective listening test are described in Fig. 5. The bias error presents the difference between the perceived panning angle and the desired panning angle. Fig. 5-(a) denotes the result of bias error in the symmetric layout (30°, 0°, -30°), and Fig. 5-(b) denotes the result of bias error in the asymmetric layout (40°, 0°, -20°). To generate the panning angle at all of directions from θ_1° to θ_2° , the additional loudspeakers at 90° and -90° were positioned only for MDAP as [6].

In the symmetric layout, the sound localization of VBAP results in increases in the bias error in closer directions to the loudspeaker. This bias is caused by the *detent effect* and is a common problem in the pair-wise amplitude panning [5]. But in the direction to the loudspeakers, the sound localization of VBAP is best because the signal is emanated by only one loudspeaker. MDAP results in better performances in closer directions to the center-positioned loudspeaker. However, at the far directions from the center loudspeaker, the localization performance remarkably decreases. It is understood that the virtual vector for the lateral direction demonstrates worse performance with respect to localization than the virtual vector in the relative frontal direction. In addition,

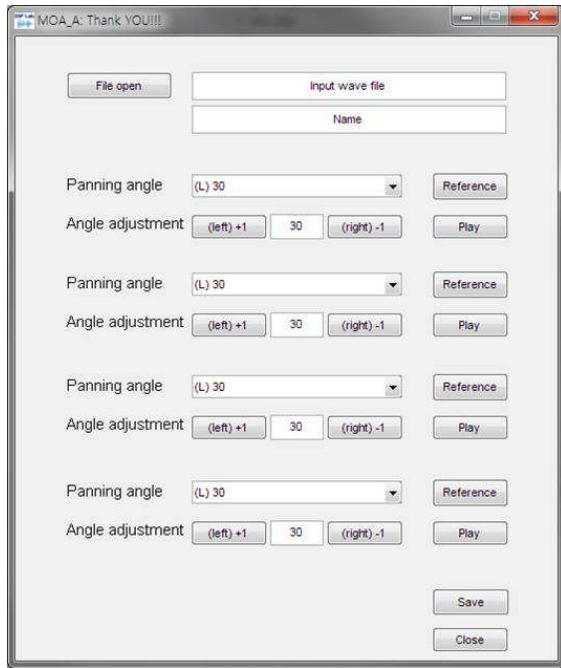


Figure 6: User interface of the test program using the method of adjustment(MOA).

the position of additional loudspeakers at $\pm 90^\circ$ can influence sound localization. The bias error of MVBNA is the smallest in the frontal direction between about 10° and -10° . In the direction of the side loudspeakers, MVBNA also shows the *detent effect* as being similar to that for pair-wise VBAP. In this test of symmetric layout, we experimentally chose the optimal parameter of nonnegative gain control function as $\varphi = 0.82$ as Fig. 3. If the curve of Eq. (7) is designed to obtain better sound localization, its results can be improved.

In the asymmetric layout, pair-wise VBAP results the similar effect with a symmetric layout. MDAP also shows better localization in closer directions to the center-positioned loudspeaker. However, at the side directions, the sound localization inclines toward the center direction. The reason is that the localization performance of the virtual source of VBAP is angle-dependent. This can be verified by the test result that MDAP30 presents worse than the MDAP20 in the lateral directions. Although MVBNA totally results in the more robust localization than the other panning methods, the bias error slightly increases within -2 degree at the near directions of the frontal loudspeaker. Its bias to the right direction is caused by the proximity of the center and right loudspeakers, when the multiple loudspeakers concurrently emanate the signal as like MDAP. In the asymmetric layout, the parameters (φ) of nonnegative gain control function were set as 0.48 and 0.97 as Fig. 3. To reduce the bias error with close loudspeakers, the curve of Eq. (7) should be sharper. If more higher-order sinusoidal function is applied into the formulation of Eq. (7), its sound localization can be improved.

5. CONCLUSIONS

In the loudspeaker system, the point-like perception of the virtual sound source is the best result. However, the per-

formance of panning algorithms is easily influenced by the configuration of loudspeakers and the listener's position. Additionally, the negative gain factor can affect the stability of sound localization. In this paper, we introduced a nonnegative amplitude panning algorithm using a multiple-wise vector base formulation to compute panning gains for the multiple loudspeakers. It has the advantages of both vector base formulation and multiple-loudspeaker panning techniques. The proposed panning method can be improved by the optimization of nonnegative gain control function in any configuration of loudspeakers. The results of subjective listening tests indicate that the localization performance of MVBNA is more robust in almost every directions. However, the proposed panning algorithm can be further improved through the optimization of a nonnegative gain control function. A clear and theoretical decision of control function should be verified with a psychoacoustic theorem and subjective test results in the future.

REFERENCES

- [1] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA: MIT Press, 1997.
- [2] A. D. Blumlein, U.K. patent 394,325, 1931.
- [3] B. Bernfeld, "Attempts for better understanding of the directional stereophonic listening mechanism," *AES 44th Conv.*, Rotterdam, Mar. 1973.
- [4] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456-466, Jun. 1997.
- [5] M. A. Gerzon, "Panpot laws for multispeaker stereo," *AES 92nd Conv.*, Vienna, Mar. 1992.
- [6] V. Pulkki, "Uniform spreading of amplitude panned virtual sources," in *Proc. IEEE Workshop on Application of Signal Processing to Audio and Acoustics (WASPAA)*, New York, USA, Oct. 1999.
- [7] V. Pulkki and T. Hirvonen, "Localization of virtual sources in multichannel audio reproduction," *IEEE Trans. on Speech and Audio Processing*, vol. 13, no. 1, pp. 105-119, Jan. 2005.
- [8] G. Monro, "In-phase corrections for ambisonics," in *Proc. Int. Computer Music Conf. (ICMC)*, Berlin, Germany, 2000.
- [9] P. G. Craven, "Continuous surround panning for 5-speaker reproduction," *AES 24th Int. Conf. on Multichannel Audio*, Jun. 2003.
- [10] V. Pulkki, "Spatial sound generation and perception by amplitude panning technique," PhD thesis, Helsinki University of Technology, Espoo, Finland, 2001.
- [11] M. Poletti, "Robust two-dimensional surround sound reproduction for nonuniform loudspeaker layouts," *J. Audio Eng. Soc.*, vol. 55, no. 7/8, pp. 598-610, July/August 2007.
- [12] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2002.
- [13] B. L. Cardozo, "Adjusting the method of adjustment: SD vs DL," *J. Acoust. Soc. Am.*, vol. 37, no. 5, pp. 768-792, May 1965.