논문 2010-47SP-5-2

UHD TV를 위한 가상 음상의 인지 위치

(Perceptual Localization of a phantom sound image for Ultrahigh-Definition TV)

이 영 우*, 김 선 민**

(Young Woo Lee and Sunmin Kim)

요 약

이 연구는 다양한 라우드 스피커 구성인 2-horizontal, 2-vertical, triplet 라우드 스피커 구성에서 UHD TV를 위한 가상 음상의 인지 위치 결과를 제시한다. 비등간격 라우드 스피커 구성에서 가상 음상을 생성하기 위해 잘 알려진 vector-based amplitude panning 알고리즘을 수정하여 적용하였다. 실제 상황에서의 위치 인지 성능을 평가하기 위해, 청취 평가는 일반적인 청취환경에서 TV의 on-axis와 off-axis에서 수행되었다. 인지 각도의 모호함에 의한 오차를 줄이기 위해 각도 조정 테스트 방법이 옥타브 밴드 신호의 각도 평가에 이용되었다. 피실험자는 실제 음원과 가상 조정된 음원의 각도가 일치할 때까지 조정 각도를 변화하였다. 공간적 블러링은 각 밴드에서 인지된 조정 각도의 차이를 비교하여 측정되었다. 청취 평가 결과는 triplet panning 방법이 vertical panning 방법보다 on-axis와 off-axis 둘 다, 인지 위치와 공간적 블러링 관점에서 더 나은 성능을 보여주었다.

Abstract

This paper presents a localization perception of a phantom sound image for ultrahigh-definition TV with respect to various loudspeaker configurations; two-horizontal, two-vertical and triplet loudspeakers. Vector base amplitude panning algorithm with modification for non-equidistant loudspeaker setup is applied to create the phantom sound image. In order to practically study the localization performance in real situation, the listening tests were conducted at the on-axis and off-axis positions of TV in normal listening room. A method of adjustment which can reduce the ambiguity of a perceived angle is exploited to evaluate the angles of octave-band signals. The subjects changed the panning angle until the real sound source and virtually panned source were coincident. A spatial blurring can be measured by examining the differences of the panning angles perceived with respect to each band. The listening tests show that the triplet panning method has better performance than vertical panning in view of perceptual localization and spatial blurring at both on-axis and off-axis positions.

Keywords : Triplet panning, Perceived angle, Spatial blurring, GVBAP, UHD TV

I. 서 론

In video image of TV, the sensation of reality grows as the horizontal viewing angle increases from 30° to 100° and saturates at an angle of around 100° [1]. The NHK Super Hi-Vision system has been developed to enhance the sensation of reality and to achieve a 100° horizontal viewing angle. The ultrahigh-definition system with 8k x 4k image pixels has the 16 times-resolution of HDTV[2]. The 22.2 channel sound system provides an exceptional spatial sound quality in wide listening area which, in horizontal plane, consists of five frontal speakers;

* 정회원, ** 정회원, 삼성전자 DMC 연구소
(DMC R&D Center, Samsung Electronics, Co.Ltd)
접수일자: 2010년7월2일, 수정완료일: 2010년8월9일

front left, front left center, front center, front right center, and front right[3]. Since the front left/right center speakers are placed at 22.5 to 30°, the speakers can be hided by TV with large viewing angle. Therefore, it is necessary to create a phantom sound image using non-hided loudspeakers.

Localization perception of virtual source has received a great deal of attention and various panning method has been developed[4~10]. In standard stereophonic setup (equally, symmetrically spaced loudspeaker layout in horizontal plane), it is known that a localization of a virtual source can be easily controlled[5] while the virtual source direction is biased towards the median plane of a listener in arbitrary setups and the direction of the virtual source within the median plane is differently perceived with respect to subjects[6]. Especially, elevation localization has been known to be highly listener-specific due to the pinna, head and torso[7].

Most experiments for virtual source panning assume that all loudspeakers be equidistant from a listener. The vector-base amplitude panning (VBAP) algorithm is well-known method for creating the virtual source[8]. For practical setup such as home theatre system, symmetrical and equidistant layout may not be feasible because the loudspeakers are usually set to be non-equidistant from a listener. Distance-based amplitude panning (DBVP)[9] and Generalized VBAP (GVBAP)[10] have been studied as alternative panning methods for non-equidistant loudspeaker layout.

This paper investigates the localization perception of a virtual sound source which is placed at 30° azimuth (position of front left/right speakers) in practical loudspeaker setup. GVBAP method was exploited as a panning method for non-equidistant loudspeaker layouts; two-horizontal, two-vertical and triplet loudspeakers. Results of listening tests are analyzed in view of perceived angle, spatial blurring, listening position, and so on.

Ⅱ. Virtual source panning method

This section presents GVBAP method[10]. Fig. 1 shows the loudspeaker setup in horizontal plane for explanation of two-dimensional model of GVBAP method. This can be easily extended to three dimensions. GVBAP method for non-equidistant loudspeakers compensates the coefficients of VBAP method in view of gain and delay using a free field acoustic model. The reverberant acoustic model can be also used for real environment.

The normalized gain factor gsi obtained by VBAP[8] is compensated using the ratio of the distances of loudspeaker and virtual source where the ri and rs are defined as the distances from i-th loudspeaker and the virtual source to a listener, respectively. Output signals of three loudspeakers can be expressed as

$$p_i(n) = g_{si}\frac{r_i}{r_s}p_s(n-\Delta_i),\ i=1,2,3 \quad (1)$$

where P_s is a sound pressure of the virtual source and P_i is defined as the time delay between *i-th* loudspeaker to the virtual source as follows:

$$\Delta_i = (r_s - r_i)F_s/c \quad (2)$$

where F_s is the sampling rate and the c is the sound speed.

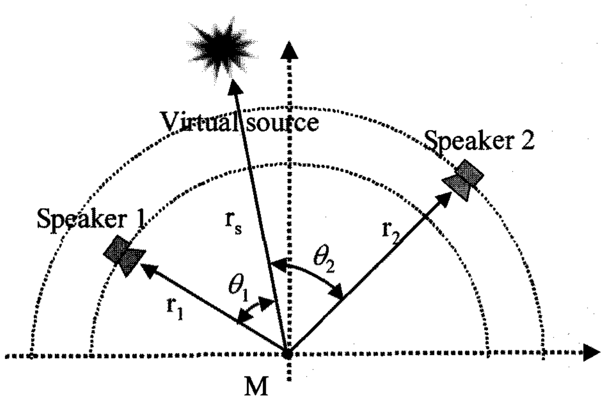


그림 1. 비대칭, 비등간격인 라우드스피커의 배치

Fig. 1. Arrangement of loudspeakers with non-symmetric and non-equidistant.

Ⅲ. Subjective evaluation

1. Listening test setup

The method of adjustment (MOA)[11] was used for the listening tests. MOA is known as a speedy research method compared to pair-wise comparison of a large number of pre-processed signal pairs. The MOA method changes the panning angle until the real sound source and virtually panned source were coincident. The direction of the virtual source was controlled using graphical user interface (GUI) software designed for the experiment (Fig. 3). The virtually panned source was produced with two or three loudspeakers of which the gain factors were calculated from 'GVBAP method. The initial panning direction was selected randomly to minimize the bias due to initial conditions. If the subjects perceived that the angles of the virtual source and real sound source were different, the subjects were asked to change the angle of the virtual source toward the appropriate direction of real sound source using the GUI software by themselves. The panning angle for each octave-band pink noise was controlled with an interval of 3°.

In the test the reference real source was presented for 1.5 s at first. The virtually panned source was followed for 1.5 s consecutively. The pause period between the real source and the virtual source was 0.5 s. This sequence was repeated until the subjects decided a match between the two sources. The GUI software was displayed in the wide TV screen and the subject was requested to gaze the GUI software to fix their head direction during the experiment. The mouse to control GUI software was placed on subject's knee so as not to cause early reflections.

In order to study the localization performance in real situation, the listening tests were conducted in a normal listening room with a reverberation time of 0.39s at 500Hz and a reference sound level was set to 80dB(C) using the pink noise. The test participants were eight young (27-35 years) males with normal hearing, mainly researchers from the laboratory where the tests were conducted. The positions of loudspeakers for the experiments were practically determined to achieve a 100° horizontal viewing angle in the 132 inch wide screen, as in Fig. 2, 5, and 8.

2. Experiment I : Horizontal panning

The horizontal localization of a virtual source was investigated with front left (FL) and front center (FC) loudspeakers in middle layer. The test setup for horizontal panning is shown in Fig. 2. The loudspeaker for reproducing a real sound source was placed at 30° azimuth in the horizontal plane, and the two loudspeakers used for the virtual source production were placed at the azimuth angles of 50° and 0°.

The GUI software for the test is shown in Fig. 3. The direction of the virtual source was controlled as

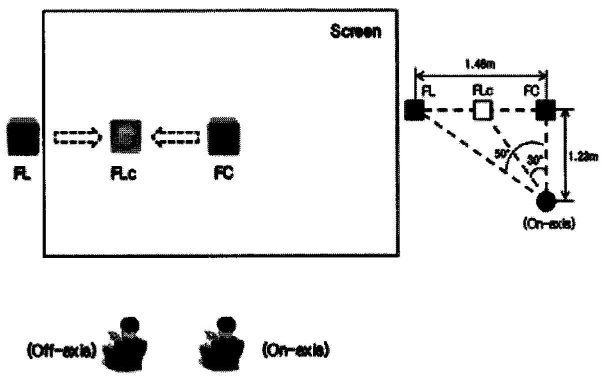


그림 2. Horizontal panning을 위한 라우드스피커 구성
Fig. 2. Loudspeaker setup for horizontal panning.

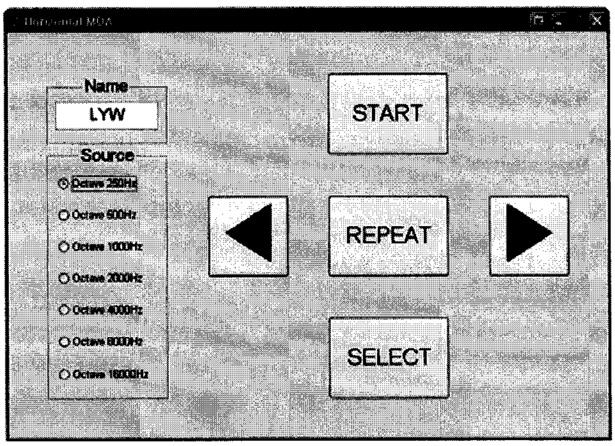
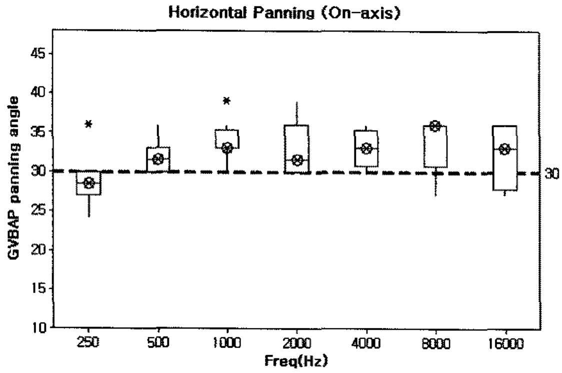
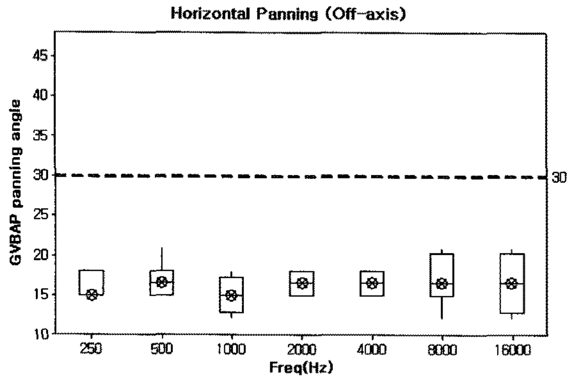


그림 3. Horizontal 테스트를 위한 GUI 소프트웨어
Fig. 3. GUI software for horizontal test.



(a) Azimuth perceived at on-axis



(b) Azimuth perceived at off-axis

그림 4. Octave-band Pink 잡음에서의 수평 패닝의 테스트 결과

Fig. 4. Test results of horizontal panning with octave-band pink noise.

adjusting the panning angle with left or right buttons in the GUI software. The test was run using octave-band pink noise with seven center frequencies at both on-axis and off-axis (71cm away from the listening sweet spot of TV).

The adjusted panning angles determined by the subjects are shown with respect to frequency in Fig. 4. In on-axis test, the adjusted panning angles were matched well with the real source directions. In off-axis test, the adjusted angles were biased towards the front center loudspeaker, which is coincided with the previous studies about stereophonic panning[5]. Note that the adjusted angles have an opposite meaning of the perceived angles in view of the panning direction.

2. Experiment II : Vertical panning

The vertical localization of a virtual source was investigated with top front left center (TpFLc) and bottom front left center (BtFLc) loudspeakers. The test setup for vertical panning is shown in Fig. 5. The loudspeaker for reproducing a real sound source was placed at 30° azimuth in the horizontal plane, and the two loudspeakers used for the virtual source production were placed at elevation directions of 37° and -28° with the azimuth of 30°. The GUI software for the test is shown in Fig. 6.

The direction of the virtual source was controlled as adjusting the panning angle with up or down buttons in the GUI software. The test materials and procedures were the same as those of the experiment I. As known in Fig. 7, the adjusted panning angles showed a big deviation with respect to frequency at both on-axis and off-axis which is similar to the

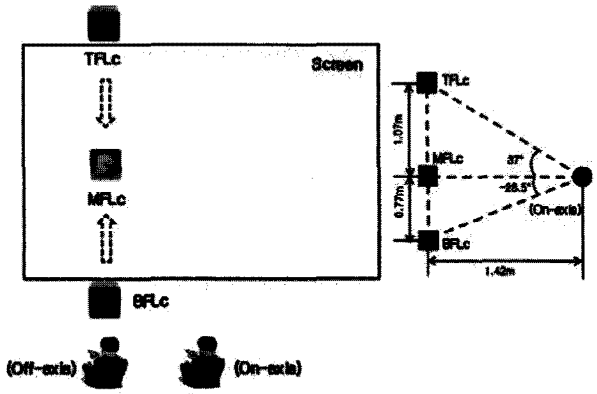


그림 5. Vertical panning을 위한 라우드스피커 구성

Fig. 5. Loudspeaker setup for vertical panning.

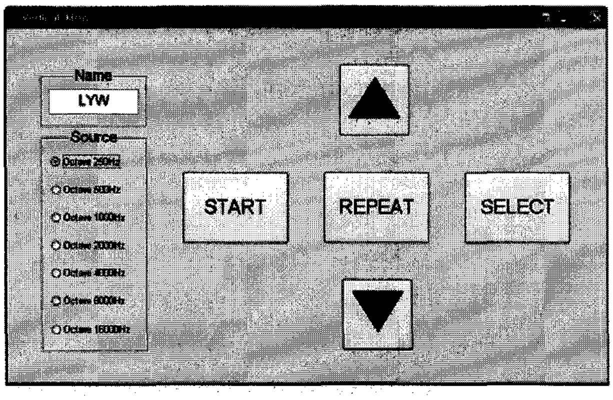
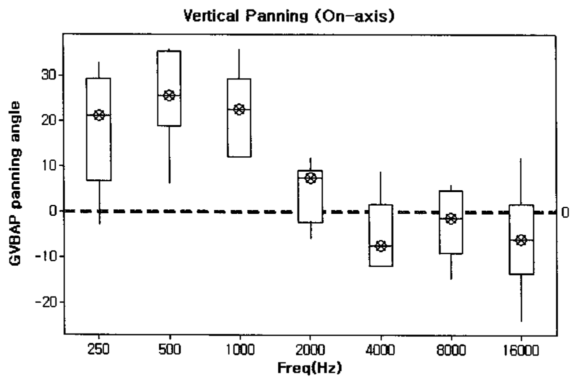
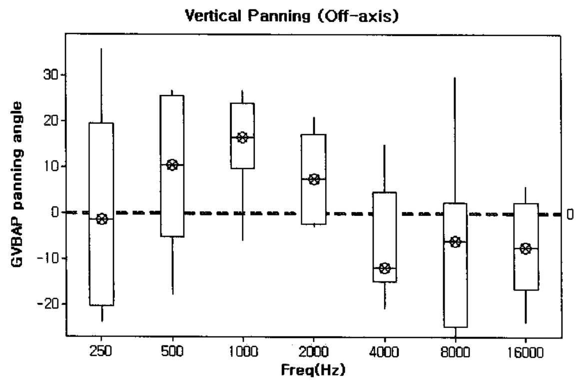


그림 6. Vertical 테스트를 위한 GUI 소프트웨어

Fig. 6. GUI software for vertical test.



(a) Elevation perceived at on-axis



(b) Elevation perceived at off-axis

그림 7. Octave-band Pink 잡음에서의 수직 패닝의 테스트 결과

Fig. 7. Test results of vertical panning with octave-band pink noise.

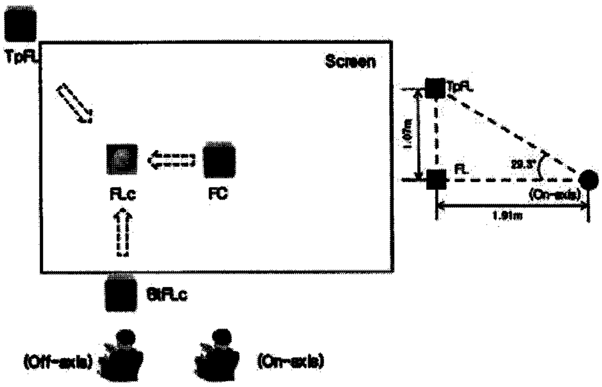


그림 8. Triplet panning을 위한 라우드스피커 구성

Fig. 8. Loudspeaker setup for triplet panning.

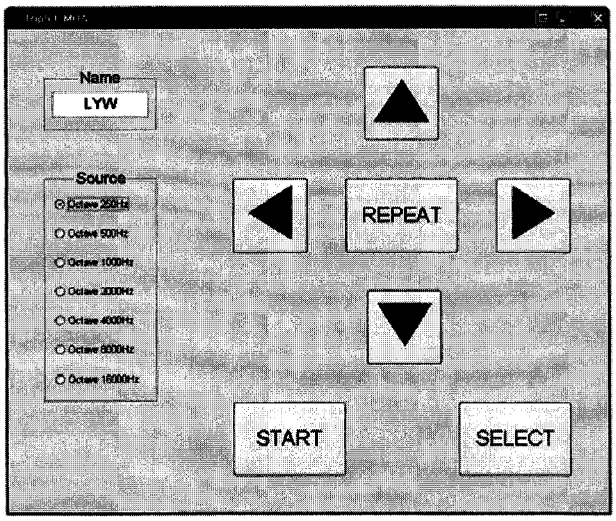


그림 9. Triplet 테스트를 위한 GUI 소프트웨어

Fig. 9. GUI software for triplet test.

result of the previous studies on the median plane listening test[6].

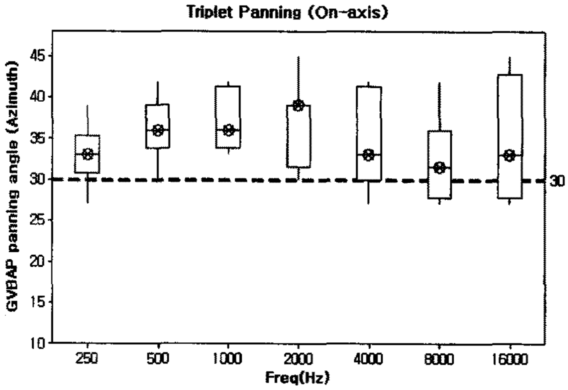
3. Experiment III : Triplet panning

The triplet localizations (azimuth, elevation) of a virtual source were investigated with top front left (TpFL), middle front center (FC) and bottom front left center (BtFLc) loudspeakers. Fig. 8 illustrates the loudspeaker setup for the triplet panning. The three loudspeakers were placed at (50°, 29°), (30°, -27°), and (0°, 0°), respectively.

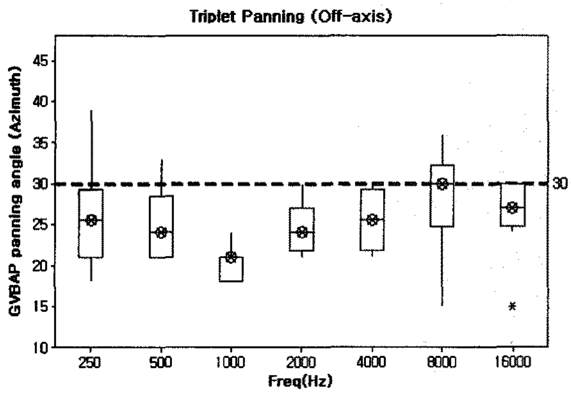
The GUI software for the triplet test is shown in Fig. 9. The direction of the virtual source was controlled as adjusting the panning angle with left, right, up, or down buttons in the GUI software. The test materials and procedures were the same as those of the experiments I and II.

Fig. 10 shows the perceived azimuth and elevation behaviour of the triplet panning. The localization perception at on-axis of the horizontal panning (Fig. 4-a) seems to be better than that of the triplet panning (Fig. 10-a) while the horizontal panning at off-axis (Fig. 4-b) has a much biased localization perception as compared with the triplet panning at off-axis (Fig. 10-b).

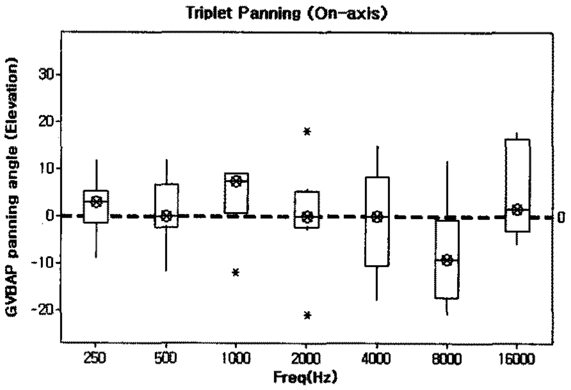
Also, the horizontal panning method had a smaller deviation with respect to the subjects than the triplet panning. As for the elevation perception, the triplet panning method (Fig. 10-c,d) showed more stable behaviour with respect to the subjects and better localization perception than those of the vertical panning (Fig. 7).



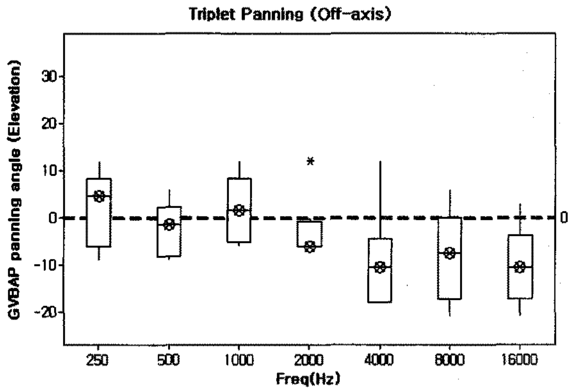
(a) Azimuth perceived at on-axis



(b) Azimuth perceived at off-axis



(c) Elevation perceived at on-axis



(d) Elevation perceived at off-axis

그림 10. Octave-band Pink 잡음에서의 Triplet panning의 테스트 결과
Fig. 10. Test results of triplet panning with octave-band pink noise.

IV. Discussion

1. Perceived angle

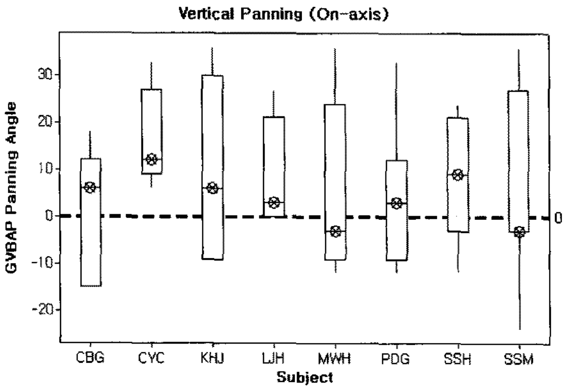
The elevation angles perceived by the vertical and triplet panning at on-axis and off-axis are statistically analyzed in order to investigate that the averages of the median values of the adjusted elevation angle with respect to each frequency band as in Figs. 7 and 10(c,d) are significantly different between the two panning methods. Since the collected data were normally distributed, the 2-sample t-test was conducted with 95% confidence levels. Table 1 shows that the localization angle of the triplet panning at on-axis is nearer to the reference angle (0°) than that of vertical panning at on-axis with a statistical significance while behaviours at off-axis are not statistically significant. This means that the triplet panning method at on-axis has better performance than the vertical panning in view of the perceived elevation angle.

표 1. 위치 인지의 통계 분석
Table 1. Statistical analysis on localization perception.

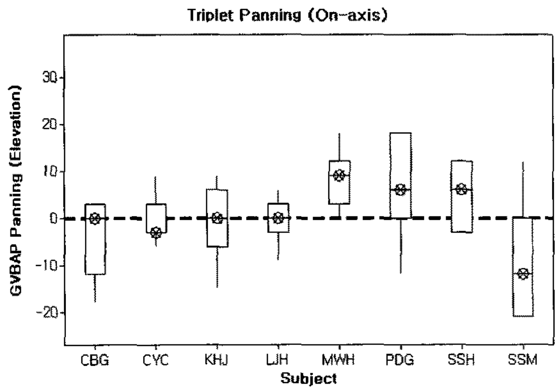
Setup	Mean	StDev	p-value
Vertical on-axis	13.07	9.59	0.036
Triplet on-axis	3.00	3.77	
Vertical off-axis	8.79	4.78	0.249
Triplet off-axis	6.00	3.77	

2. Spatial blurring

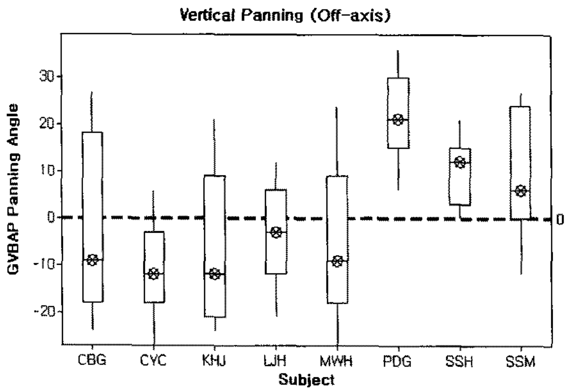
A spatial blurring in the virtual source panning seems to be caused by the localization angle differently perceived with respect to each frequency band. It makes the perceived source width large and the perceived position ambiguous which can be easily



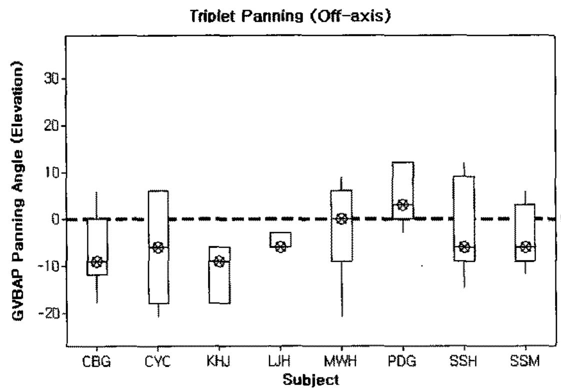
(a) Elevation perceived at on-axis by vertical panning



(b) Elevation perceived at on-axis by triplet panning



(c) Elevation perceived at off-axis by vertical panning



(d) Elevation perceived at off-axis by triplet panning

그림 11. 피실험자에 따른 테스트 결과
Fig. 11. Test results with respect to subjects.

observed from the experiment using full-band pink noise as the test materials. Therefore, the spatial blurring can be measured as examining the differences of the panning angles perceived with respect to each band. In order to find out the spatially-blurred characteristics of virtually panned sources, the adjusted elevation angle for the vertical and triplet panning was plotted as a function of the subject as in Fig. 11. The interquartile range of the panning angles with respect to the frequency band adjusted by each subject represents the spatial blurring of the subject. The median and data range of the adjusted elevation angle seems to be more stable in the triplet panning as compared with the vertical panning.

Statistical tests were run to investigate whether the spatial blurring of the adjusted elevation angle was significantly different between the vertical and triplet setups. The averages of the interquartile ranges obtained from each subject were calculated for the two panning methods at on-axis and off-axis.

Since the collected data were normally distributed, the 2-sample t-test was conducted with 95% confidence level. Table 2 shows that the spatial

표 2. 공간적 블러링의 통계 분석
Table 2. Statistical analysis on spatial blurring.

Setup	Mean	StDev	p-value
Vertical on-axis	26.25	7.30	0.001
Triplet on-axis	12.75	5.50	
Vertical off-axis	22.13	8.48	0.034
Triplet off-axis	13.50	6.00	

blurring of the triplet panning is, with a statistical significance, less than that of the vertical panning at both on-axis and off-axis. This means that the triplet panning method has better performance than the vertical panning in view of the spatial blurring.

3. Subject

As shown in Fig. 11, the results of the triplet panning show less deviation with respect to each subject as compared with the vertical panning (see the median values of the subjects). Most of subjects reported that it was easier to adjust the perceived angle of the horizontal panning method than other panning methods. As for the vertical and triplet panning, some subjects reported that they perceived the virtual source to be spread at low frequency, but other subjects reported that the virtual source was spread at high frequency. In the vertical panning at on-axis, some subjects reported that they perceived the virtual source was biased towards the center direction in azimuth.

4. Listening position

In off-axis test, the perceived azimuth angle in horizontal panning was more biased towards the front left loudspeaker as compared with the triplet panning due to precedence effect[4]. Also, the spatial blurring at off-axis of the triplet panning was less than that of the vertical panning. This means that a listening sweet spot of the triplet loudspeaker setup is wider than that of other loudspeaker setups.

V. Conclusion

This paper investigated a localization perception of a phantom sound image for ultrahigh-definition TV with respect to various loudspeaker configurations and the corresponding panning methods. For practical setup, the loudspeakers with non-equidistance from a listener were placed around the TV in a normal listening room. Virtual source of octave-band pink noise for replacing a front left center loudspeaker was reproduced using the generalized vector base amplitude panning method in two-horizontal, two-vertical, and triplet loudspeaker layouts. Subjective listening test using the method of adjustment were carried out and the results were statistically analyzed using the 2-sample t-test. Consequently, the triplet loudspeaker layout showed better performance in view of the perceived angle, the spatial blurring and the listening sweet spot as compared with the two vertically-spaced loudspeaker layout. On the other hand, the two horizontally-spaced loudspeaker layout with lack of azimuth perception at off-sweet spot showed better performance rather than the triplet setup in view of the perceived azimuth at on-axis, the spatial blurring, and the stability with respect to the subject.

VI. Future work

Since the front center speaker is positioned at 0°, it can be also hided by TV. The effect of a phantom sound image panned from non-hided loudspeakers excluding the center loudspeaker will be investigated. Also, new panning method as using the reverberant acoustic model instead of using the free field acoustic model will be studied in order to decrease the panning error caused in various real environments.

Reference

[1] T. Hatada, et al., "Psychophysical analysis of the sensation of reality induced by a visual wide-field display", SMPTE Journal, Vol. 89, pp. 560-569 (1980).

[2] F. Okano, et al., "Ultrahigh-definition Television System with 4000 Scanning Lines", presented at NAB 2004, Las Vegas, USA (2004).

[3] K. Hamasaki, et al., "Wide Listening Area with Exceptional Spatial Sound Quality of a 22.2 Multichannel Sound System", presented at 122nd AES Convention, Vienna, Austria (2007).

[4] R. Y. Litovsky, et al., "The precedence effect", J. Acoust. Soc. Am. 106(4), pp. 1633-1654 (1999).

[5] V. Pulkki and M. Karjalainen, "Localization of

Amplitude-Panned Virtual Sources, I: Stereophonic Panning", J. Audio Eng. Soc., (2001).

[6] V. Pulkki, "Localization of Amplitude-Panned Virtual Sources, II: Two- and Three-Dimensional Panning", J. Audio Eng. Soc. (2001).

[7] J. Blauert, "Spatial Hearing, The Psychophysics of Human Sound Localization", MIT Press (1997).

[8] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning", J. Audio Eng. Soc. (1997).

[9] L. Trond and B. Pascal, "DBAP - Distance-based amplitude panning", International Computer Music Conference (2009).

[10] K. Sunmin and P. Youngjin, "Driving method of multiple loudspeakers for virtual sound localization system", Domestic conference, Acoustic Society of Korea (2004).

[11] B. L. Cardozo, "Adjusting the Method of Adjustment: SD vs. DL", J. Acoust. Soc. Am., vol. 37, pp. 768-792 (1965).

저 자 소 개

이 영 우(정회원)

2004년 인제대학교 의용공학과 학사 졸업.

2006년 한양대학교 의용생체공학과 석사졸업.

2007년~삼성전자 DMC 연구소 선임연구원.

<주관심분야 : 음향신호처리, Spatial Hearing>

김 선 민(정회원)-교신저자

1996년 KAIST 기계공학과 학사 졸업.

1998년 KAIST 기계공학과 석사 졸업.

2004년 KAIST 기계공학과 박사 졸업.

2004년~삼성전자 DMC 연구소 책임연구원.

<주관심분야 : 음향신호처리, Spatial Hearing>