

Multiple Vertical Panning を用いた立体音響システムのための視聴覚実験

木村 敏幸[†] 安藤 広志[†]

[†] 独立行政法人情報通信研究機構 ユニバーサルコミュニケーション研究所
〒 619-0288 京都府相楽郡精華町光台 2-2-2
E-mail: †{t-kimura,h-ando}@nict.go.jp

あらまし 本報告では、大画面裸眼立体映像提示システムに適合するために従来提案していた Multiple Vertical Panning (MVP) 方式に基づいた立体音響システムの視聴覚性能を評価するために、200 インチスクリーンの上下に計 82 個のスピーカを配置したスピーカアレイを用いて視聴覚実験を行った。12 名の視聴者に立体映像と音を視聴させ合成音像の位置を回答させたところ、提案システムの音を立体映像と一緒に提示した際に視聴者はどの場所で視聴しても立体像の位置に音像を感じたため、提案システムがステレオホニックに比べて有効であることが分かった。

キーワード 超臨場感コミュニケーション、立体音響システム、垂直パニング、裸眼立体映像システム、視聴覚実験

Audio-visual experiment for 3D audio system using multiple vertical panning

Toshiyuki KIMURA[†] and Hiroshi ANDO[†]

[†] Universal Communication Research Institute,
National Institute of Information and Communications Technology
2-2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0288 Japan
E-mail: †{t-kimura,h-ando}@nict.go.jp

Abstract In this paper, in order to evaluate the audio-visual performance of the proposed three-dimensional (3D) audio system, which is based on Multiple Vertical Panning (MVP) method and matches to the glasses-free 3D video display system in which the size of screen is very large, the audio-visual experiment was designed by using the loudspeaker array in which eighty-two loudspeakers were placed on the top and bottom of the 200-inch screen. Twelve viewers viewed a 3D video and sound and reported the position of synthesized sound images. As a result, since viewers could always feel the synthesized sound images at the position of the 3D object when the 3D sound of the proposed system was presented with 3D video, it was indicated that the proposed 3D audio system was effect compared with the conventional system such as stereophonic.

Key words Ultra-realistic communication, 3D audio system, vertical panning, glasses-free 3D video display system, audio-visual experiment

1. はじめに

独立行政法人情報通信研究機構 (NICT) では超臨場感コミュニケーション技術に関する研究を進めている [1]. 立体映像技術や立体音響技術によって三次元空間上に映像や音響をよりリアルに表現することができるようになれば、今までの映像・音響メディアでは実現できなかったより臨場感のあるコミュニケーション (立体テレビ電話や立体遠隔通信会議など) が可能になると期待される。

NICT では次世代の立体映像技術として裸眼立体映像技術に

着目し、これまでにプロジェクタアレイを用いた方式を提案し、200 インチの画面を持った大画面裸眼立体映像提示システムを開発してきている [2]. しかし、今までに開発してきたシステムは視覚のみを提示するもので、それ以外の感覚は提示できなかった。よりリアルな臨場感を表現するためには複数の感覚を同時に提示できるシステムを開発する必要がある。特に、聴覚を提示するためには大画面裸眼立体映像提示システムに適合した立体音響システムを開発する必要がある。

我々はこれまでに今までの立体音響システムとは異なる観点に基づいた新たなる立体音響システム (Multiple Vertical

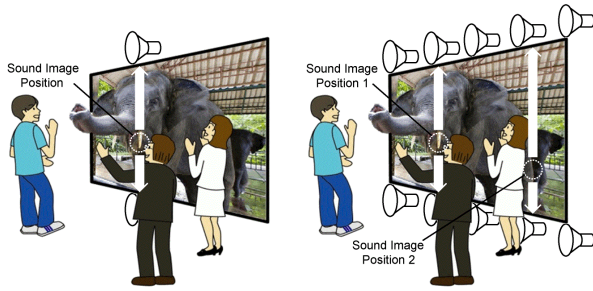


図1 提案する立体音響システムの基本構成

Fig.1 Basic configuration of the proposed 3D audio system.

Panning (MVP) 方式) を提案し、聴取実験による聴覚性能を検討してきた [3]。本報告では提案システムの視聴覚性能を評価するために実施した視聴覚実験について述べ、提案システムが従来の立体音響システムに比べて有効であることを示す。

2. 提案システムの原理

図1に提案システムの基本構成を示す。まず、図1の左側に示すように、立体像の位置の上下に2個のスピーカを配置する。そして、音源に音量差をつけて2個のスピーカから音を再生(すなわち「垂直パニング」)すると、視聴者は2個のスピーカの間で音が鳴っているように感じるようになる。その際、適切な音量差を設定することによって、視聴者は立体像の位置で音が鳴っているように感じるようになる。音が鳴っているのはスクリーンの上下に配置した2個のスピーカだけなので、複数の視聴者はどこにいても常に立体像の位置で音が鳴っているように感じることができる。

さらに、図1の右側に示すように、スクリーンの上下に複数のスピーカ対を配置することによって、音像を表現できる位置がスクリーンの上下方向のみならず左右方向にも拡大される。これにより、複数の視聴者はどこにいても常に大画面裸眼立体映像提示システムがスクリーン上に描写する立体像の位置で音が鳴っているように感じることができる。

3. 視聴覚実験

本報告では200インチ画面を持ったスクリーンの上下にスピーカアレイを配置して実施した視聴覚実験について述べる。

3.1 実験環境・条件

実験はATR地下1階の200インチ背面投影型映像スクリーンが設置されている映像実験室にて行った。スクリーンの後方には2台のプロジェクタが設置されており、それぞれ左目用と右目用の映像が投影される。プロジェクタとスクリーンの間には偏光フィルタが設置されているので、視聴者は偏光方式の立体メガネをかけることによって立体映像を観ることができる。部屋の残響時間は258ms、部屋の暗騒音レベルは41dBAだった。

図2に示すようにスクリーンの上下に計82個のスピーカを配置した。スクリーンは実験室の天井と床にワイヤーで固定されており、スピーカをスクリーンの真上と真下に配置することができないため、スピーカはスクリーンより0.5m前方に配置

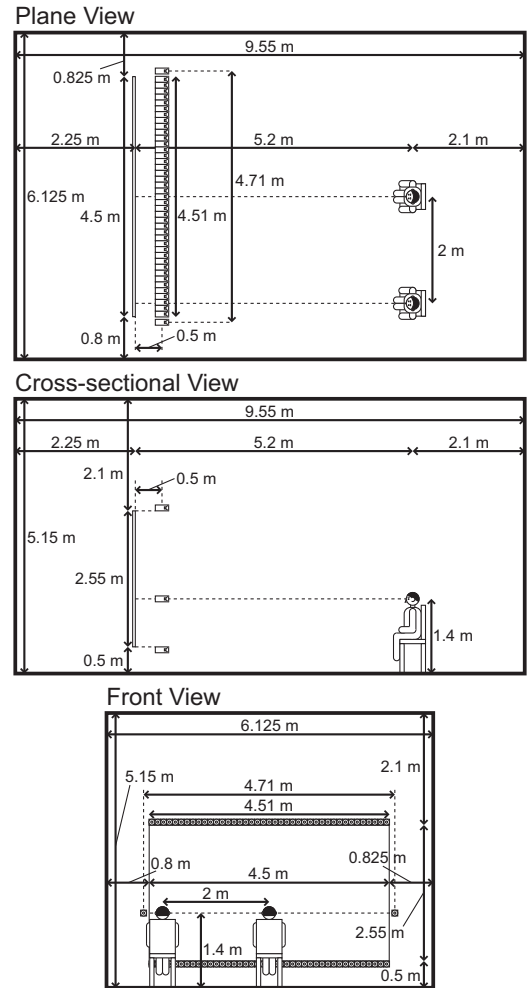


図2 視聴覚実験における視聴者、スクリーン及びスピーカアレイの配置

Fig.2 Position of viewers, screen and the loudspeaker array in the audio-visual experiment.

した。スピーカは市販のスピーカユニット (Fostex: FE103En) を自作の密閉型エンクロージャ(11cm(幅)×25cm(奥行)×11cm(高さ))に取り付けたものを用いた。従って、上下のスピーカアレイ全体の長さは4.51m(=11cm×41)となる。大画面裸眼立体映像提示システムの最適視聴距離を考慮した場合、視聴位置はスクリーンから5.5m離れた地点に設定すべきであるが、実験室の後方には映像処理作業用の机が固定されていたので、視聴位置はスクリーンからなるべく離れた地点(5.2mの距離)に設定した。大画面裸眼立体映像提示システムの最適な視聴幅はスクリーンから5.5m離れた場合、スクリーン正面を中心として左右に2mである。そこで、スクリーン正面の視聴位置(以降「正面位置」と呼ぶ)に加えて、スクリーン正面から2m左側に移動した視聴位置(以降「側方位置」と呼ぶ)をもう1か所設定した。2か所の視聴位置の高さは視聴者の耳の位置において1.4mとした。さらに、提案システムを従来システムと比較するために、スクリーンの左右に2個のスピーカを視聴者の耳の位置と同じ高さに配置した。音圧レベルはスクリーン正面の視聴位置において約70dBAに設定した。

音源には4秒間の音を2種類(白色雑音、音声)用いた。本

実験で設定した音の再生条件を図3に示す。条件(a), (b)における灰色のスピーカは音が再生されないスピーカを表す。両方の条件において、スクリーンには格子点及び線が平面画像として常に提示されている。格子線の縦と横の幅は0.33 mである。(a)の提案システム条件では、まずは提示する立体像の横位置からスクリーン上下に配置した2個の再生用スピーカを選択し、下記の式によって音源信号 $s(n)$ に上下の音量差をつけた音信号を選択した2個の再生用スピーカから再生した。

$$x_U(n) = a_U s(n) \quad (1)$$

$$x_D(n) = a_D s(n) \quad (2)$$

但し、 $x_U(n)$, $x_D(n)$ は上側及び下側のスピーカから再生する音信号、 a_U , a_D はそれぞれのゲイン係数を表し、音量差 ΔA [dB] から下記のように定義される。

$$a_U = \frac{10^{\frac{\Delta A}{20}}}{\sqrt{10^{\frac{\Delta A}{10}} + 1}} \quad (3)$$

$$a_D = \frac{1}{\sqrt{10^{\frac{\Delta A}{10}} + 1}} \quad (4)$$

本実験では音量差 ΔA は前回の報告 [3] に従い、下記のように設定した。ただし、 P_V は立体像の縦位置を表し、-0.33 が視聴者の耳の高さに相当する。

$$\Delta A = \frac{P_V + 0.1437}{0.1065} \quad (5)$$

(b)に示すステレオ条件では、従来のステレオホニック方式に従い、下記の式によって音源信号 $s(n)$ に左右の音量差をつけた音信号をスクリーンの左右に配置した2個のスピーカから再生した。

$$x_L(n) = a_L s(n) \quad (6)$$

$$x_R(n) = a_R s(n) \quad (7)$$

但し、 $x_L(n)$, $x_R(n)$ は左側及び右側のスピーカから再生する音信号、 a_L , a_R はそれぞれのゲイン係数を表す。本実験ではタンゼント則に従い下記の式のように設定した [4]。

$$a_L = \frac{1 - x}{\sqrt{2(1 + x^2)}} \quad (8)$$

$$a_R = \frac{1 + x}{\sqrt{2(1 + x^2)}} \quad (9)$$

但し、 $x = \frac{\tan \theta}{\tan \theta_0}$ で、 θ 及び θ_0 はスクリーン正面の視聴位置における立体像及び左右の2個のスピーカの開き角を表す。本実験では、下記のように設定した。

$$\frac{\tan \theta}{\tan \theta_0} = \frac{P_H}{2.355} \quad (10)$$

P_H は立体像の横位置を表し、0 がスクリーン正面に相当する。

さらに、一部の実験条件では音に加えて図4に示すような立体映像も同時に提示した。音源が白色雑音の時は図4の左側のスピーカ、音源が音声の時は図4の右側のキャラクタの立体映像を提示した。立体映像は Maya [5] で作成し、最適視聴距離は

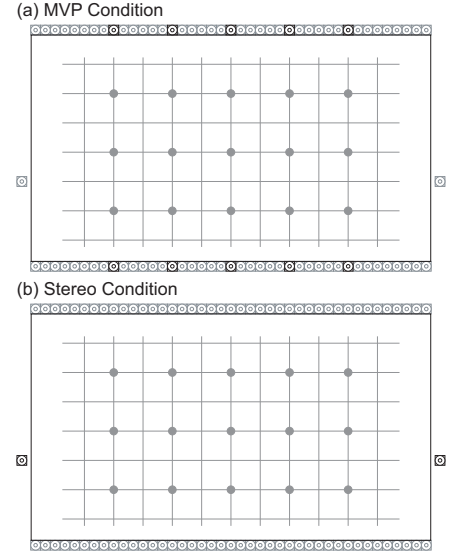


図3 視聴覚実験で用いた音の再生条件

Fig. 3 Sound conditions used in the audio-visual experiment.

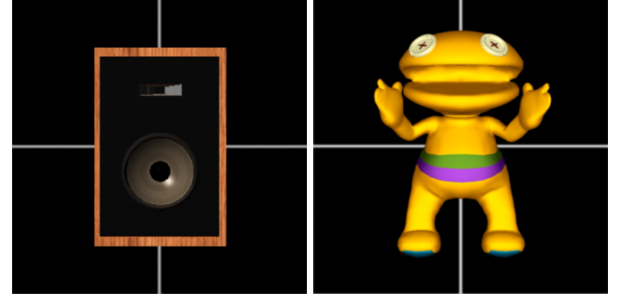


図4 視聴覚実験で用いた立体映像 (左: 白色雑音, 右: 音声)

Fig. 4 3D videos used in the audio-visual experiment (Left: white noise, Right: speech).

表1 視聴覚実験における実験条件

Table 1 Experimental conditions in the audio-visual experiment.

Index	Sound	3D video
(i)	Stereo	Sound only
(ii)	Proposed system	Sound only
(iii)	Stereo	Sound & video
(iv)	Proposed system	Sound & video

5.5 m, 視差は0.0625 mとした。従って、本実験では表1に示すような4つの実験条件が設定された。また、大画面裸眼立体映像提示システムは視聴位置が変わるとそれに合わせて立体像の見え方が変わるので、本実験においても視聴位置が変わるたびにその位置に合わせた立体映像を提示した。

3.2 実験計画・手順

視聴者は日常生活において聴力に異常の見られない21–40歳の計12名(男性6名, 女性6名)である。聴取実験の流れを図5に示す。まず、実験全体を視聴位置ごとに2つに分割した。さらに、音源ごとに分割を行い、計4つのセッションを設定した。各セッションにおいて10回の練習試行の後、100回の本試行を行った。その際、50試行ごとに休憩時間を設けた。視聴位置、音源及び試行の提示順序は視聴者ごとにランダム化し

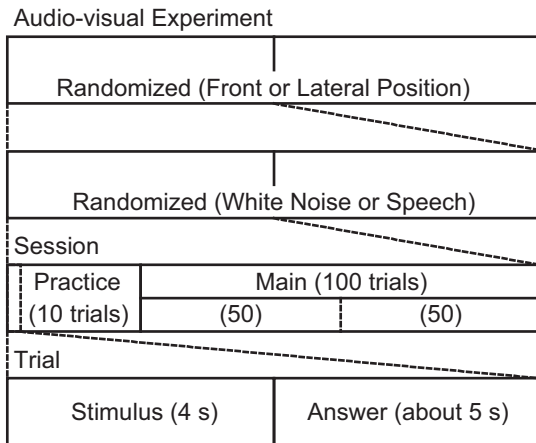


図 5 視聴覚実験の流れ図

Fig. 5 Flowchart of the audio-visual experiment.

表 2 視聴覚実験における立体像の位置

Table 2 3D image positions in the audio-visual experiment.

Index	P_H	P_V	Index	P_H	P_V	Index	P_H	P_V
1	-1.32	0.66	6	-1.32	0	11	-1.32	-0.66
2	-0.66	0.66	7	-0.66	0	12	-0.66	-0.66
3	0	0.66	8	0	0	13	0	-0.66
4	0.66	0.66	9	0.66	0	14	0.66	-0.66
5	1.32	0.66	10	1.32	0	15	1.32	-0.66

表 3 視聴覚実験における練習試行及び本試行

Table 3 Practice and main trials in the audio-visual experiment.

	Element	Note
Practice (10)	= 2 conditions × 5 positions	(ii) & (iv) of Table 1 1, 5, 8, 11 & 15 in Table 2
Main (100)	= [1 conditions × 5 positions + 3 conditions × 15 positions] × 2 repetitions	(i) in Table 1 $P_H = -1.32, -0.66, 0, 0.66, 1.32$ (ii)-(iv) in Table 1 1-15 in Table 2

た。立体像の位置を表 2 に、練習試行及び本試行の詳細を表 3 に示す。

視聴者には音を聞いた後に知覚した音像の位置を回答用紙に記入するように教示した。但し、音と同時に立体像が提示された時には立体像を注視して音を聞くように教示した。音像の位置は図 6 に示すような格子点及び線に従って回答させた。これは図 3 における格子点及び線と対応しており、横番号は-5 から 5 までの 11 通り、縦番号は-3 から 3 までの 7 通りで回答できる。2 つ以上の音が鳴っていると思った場合は、2 か所以上の位置で回答させた。また、回答の際には視聴者は頭部及び上半身を自由に動かすことができた。

3.3 実験結果及び考察

実験結果を分析するために、視聴者の回答から 2 か所以上の番号が記入されているものを全て除去した上で、視聴者が回答した横番号及び縦番号の平均値を算出した。実験条件ごとの結果を図 7~10 に示す。縦及び横方向のエラーバーは水平及び垂

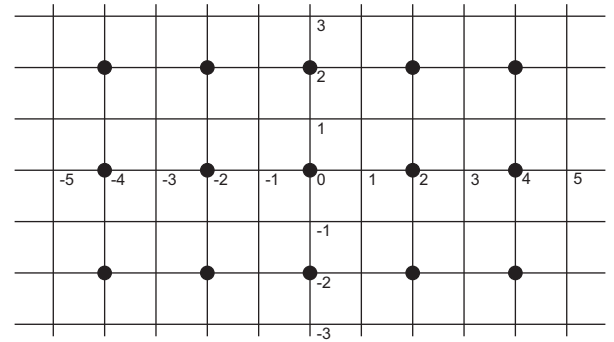


図 6 知覚した音像の位置と回答番号の関係図

Fig. 6 Relation between perceived positions of sound images and answer indexes.

直方向の 95%信頼区間を表す。提示した立体像の位置は灰色の丸で示されているので、知覚した音像の位置が灰色の丸に近ければ、視聴者は立体像の位置に正しく音像を感じているということを表している。

音のみを提示した場合、ステレオホニックでは正面位置での音像の横位置は概ね正しく定位されているが、縦位置は入力 (-0.33) よりも高い位置に定位されている。また、側方位置では音像の横位置も正しく定位されていない。これは、ステレオホニックが正面位置での視聴を前提としたシステムであるためである。

これに対して、提案システムにおいては、正面位置では 15 か所の音像位置の間に有意な差が生じている。また、側方位置では 15 か所の音像の横位置の間に有意な差が生じている。従って、音のみを提示した場合、提案システムはどの場所で視聴してもステレオホニックに比べて音像の横位置を正しく定位できるし、正面位置では音像の縦位置もステレオホニックに比べて正しく定位できるということが言える。

一方、音と立体映像を提示した場合、腹話術効果が発生するため、ステレオホニックにおいても正面位置では視聴者は概ね立体像の位置と同じ位置に音像を感じている。しかし、視聴位置が側方位置になると、ステレオホニックでは音像の位置が左側に偏り、腹話術効果が起こらなくなるため、視聴者は立体像の位置に音像を感じていない。

これに対して、提案システムでは、どの視聴位置においても視聴者は立体像の位置と同じ位置に音像を感じている。従って、音と立体映像を提示した場合、提案システムでは視聴者はどの場所で視聴しても正確な位置に音像を感じるため、提案システムは従来のステレオホニックに比べて有効であるということが言える。

4. まとめ

本報告では、従来提案されている大画面裸眼立体映像提示システムに適合するために、Multiple Vertical Panning (MVP) 方式に基づいた新たな立体音響システムを提案した。提案システムの視聴覚性能を評価するために、200 インチスクリーンの上下に計 82 個のスピーカを配置したスピーカアレイを用い

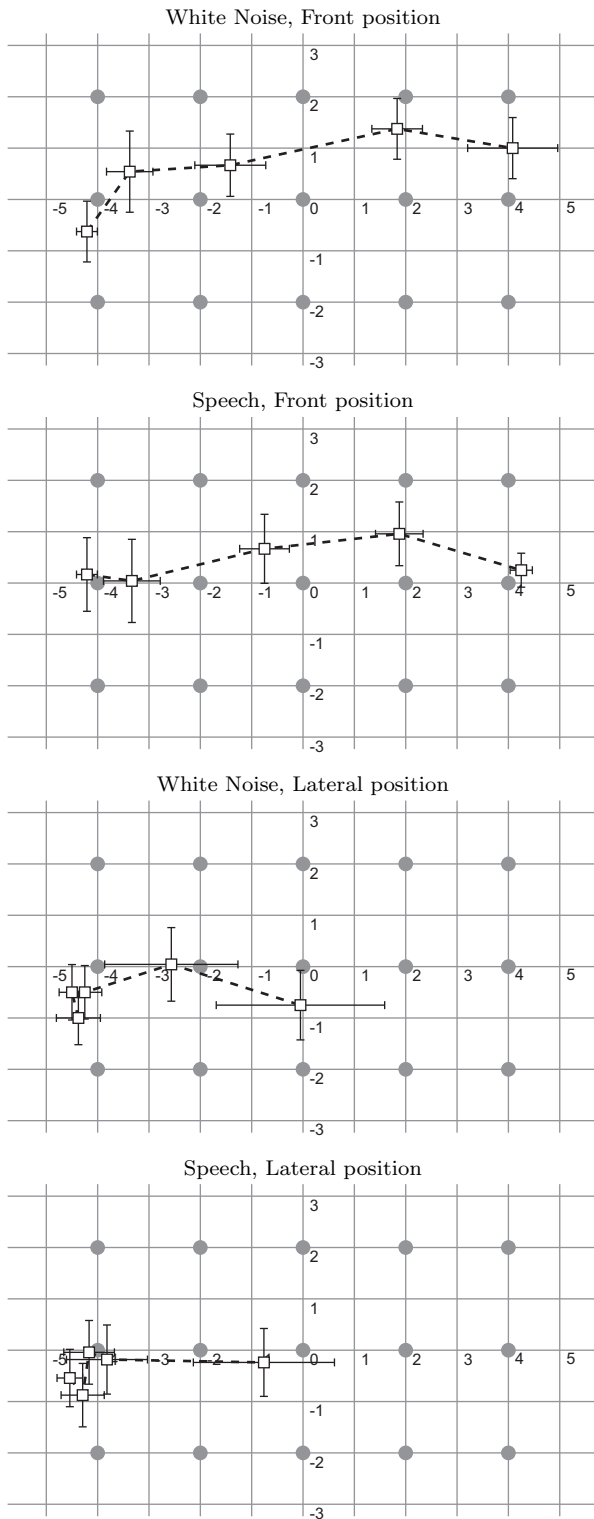


図 7 音のみ提示条件におけるステレオホニクの結果
Fig. 7 Results of the stereo in the sound only condition.

て視聴覚実験を行ったところ、提案システムの音を立体映像と一緒に提示した際に視聴者はどの場所で視聴しても立体像の位置に音像を感じたため、提案した方式が従来のステレオホニクに比べて有効であることが分かった。

今後はスピーカの数減らすことや取音・伝送手法を確立することにより、提案したシステムの実用化の可能性を検討する必要があります。また、音量の変化によって立体音の距離感をより

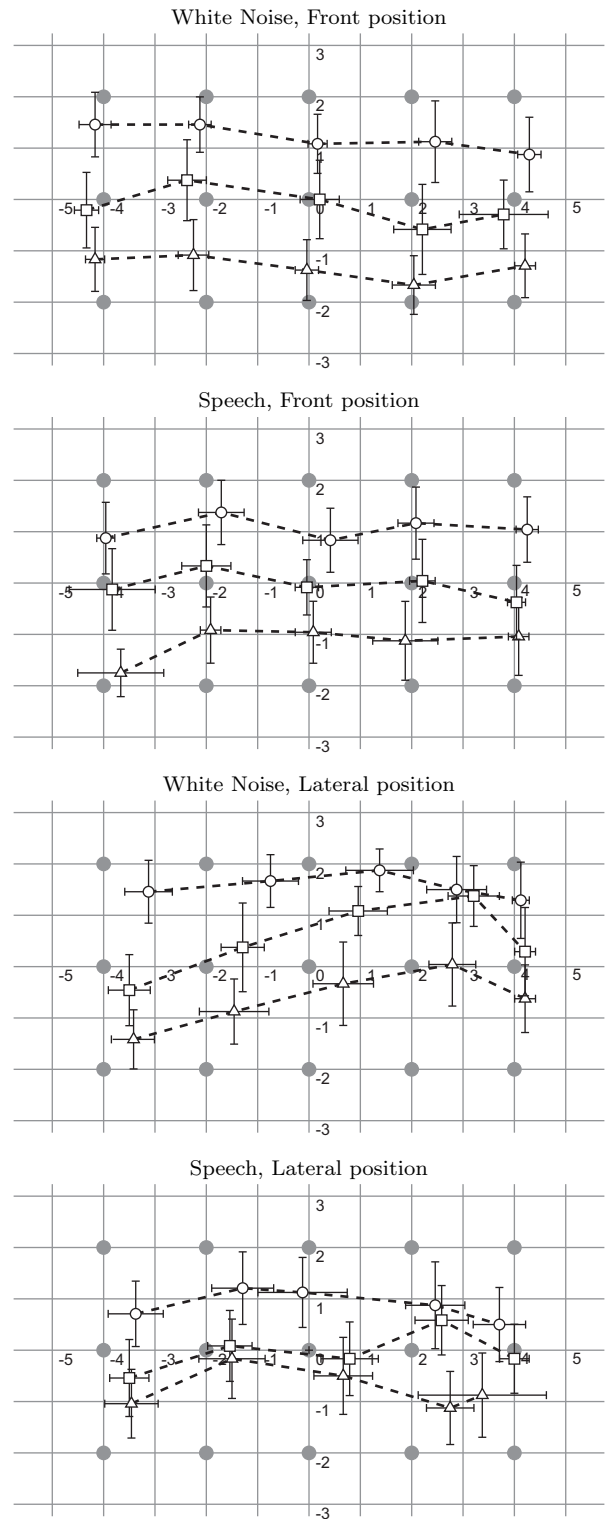


図 8 音のみ提示条件における提案システムの結果
Fig. 8 Results of the proposed system in the sound only condition.

良く表現する手法についても検討する必要があります。

5. 謝 辞

視聴覚実験の実施にあたり、実験環境の構築にご協力下さった岩澤昭一郎主任研究員、立体映像を作成して下さいました Lopez-Gulliver Roberto 主任研究員、牧野真緒研究員に感謝の意を申

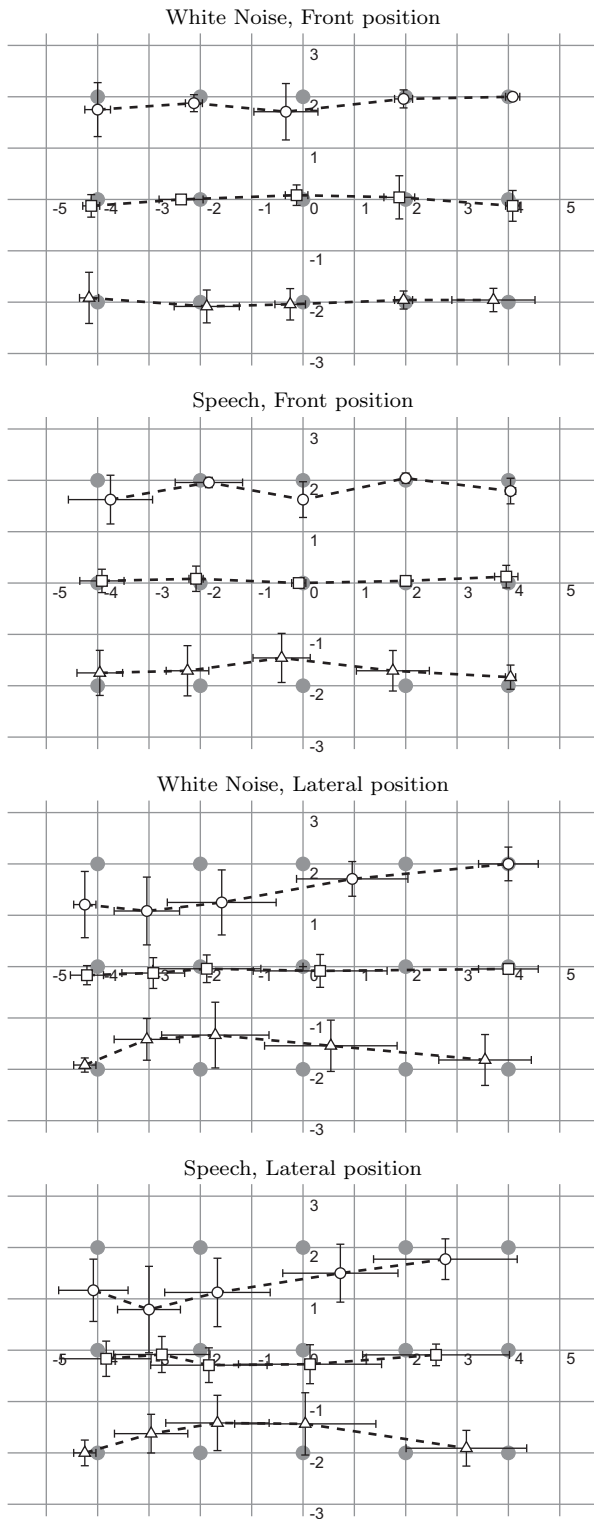


図 9 音映像提示条件におけるステレオホニックの結果
Fig. 9 Results of the stereo in the sound and video condition.

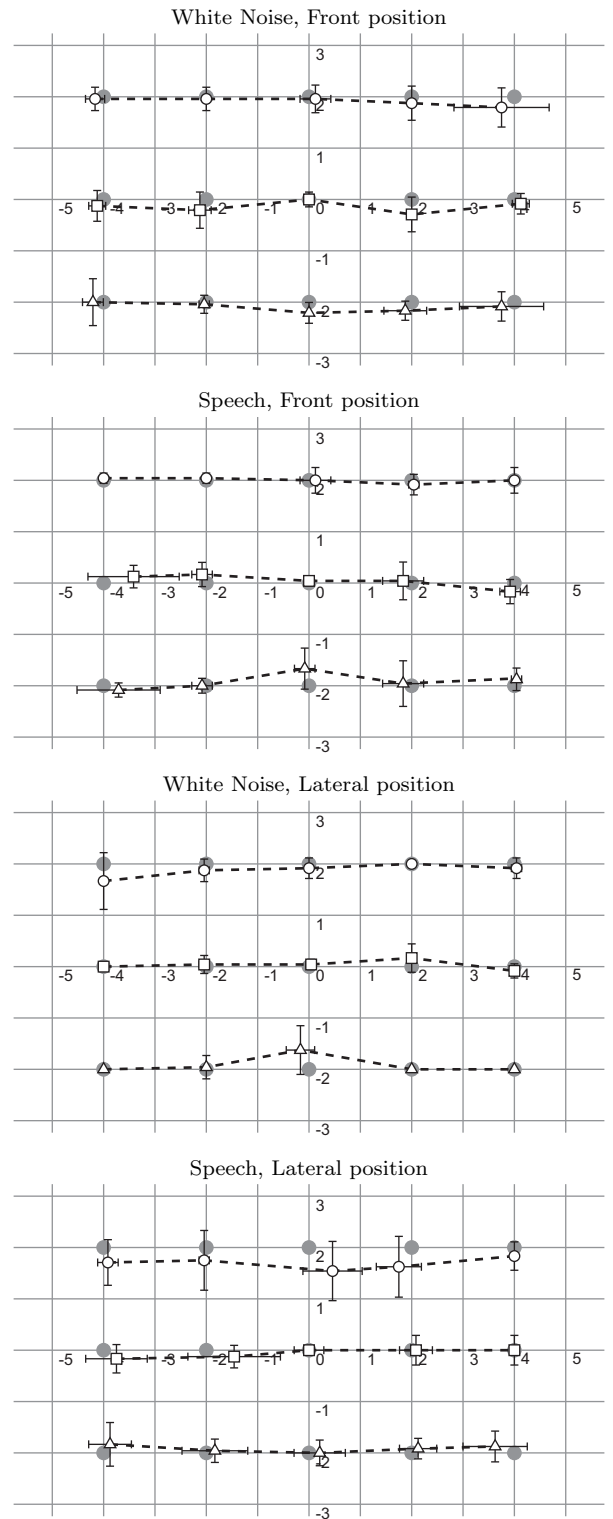


図 10 音映像提示条件における提案システムの結果
Fig. 10 Results of the proposed system in the sound and video condition.

し上げる。

文 献

- [1] 榎並和雅, “超臨場感システムの研究推進に向けて- (独) 情報通信研究機構の研究と、産学官フォーラム URCF の活動,” 画像ラボ, vol.18, no.9, pp.41-45, Sept. 2007.
- [2] 安藤広志, 河北真宏, 岩澤昭一郎, “世界初「200 インチの自然な裸眼立体表示技術」の開発に成功,” 情報通信研究機構報道発表, <http://www.nict.go.jp/press/2011/01/25-1.html>, Jan. 2011.

- [3] 木村敏幸, 安藤広志, “Multiple vertical panning を用いた立体音響システムのための聴取実験,” 信学技報, no.EA2011-122, pp.13-18, March 2012.
- [4] B. Bernfeld, “Attempts for better understanding of the directional stereophonic listening mechanism,” Proc. 44th Audio Eng. Soc. Convention, no.C-4, pp.1-24, Feb. 1973.
- [5] Website of Maya Software, <http://www.autodesk.co.jp/maya>.