

# Multiple Vertical Panning を用いた立体音響システムにおける スピーカ数の臨場感への影響\*

○木村敏幸, △安藤広志 (NICT)

## 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の映像や音響をよりリアルに表現することができるになれば, 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]. 我々はさらに大画面裸眼立体映像提示システムに適合する立体音響システムとして, 今までとは異なる観点に基づいた新たな立体音響システム (Multiple Vertical Panning, 以降「MVP 方式」と呼ぶ) を提案し, MVP 方式が十分な性能を有することを実証してきた [3].

しかし, MVP 方式は横方向に非常に多くのスピーカを要するシステムなので, MVP 方式を立体遠隔通信会議などに応用しようとする場合, 伝送するチャンネル数が非常に多くなり, 実用化の妨げになる可能性がある. そこで, 本発表では MVP 方式においてスピーカの数減らしていった場合にどのくらい臨場感に影響があるのかを視聴覚実験によって検討する.

## 2 MVP 方式の原理 [3]

Fig. 1 に MVP 方式の基本構成を示す. まず, Fig. 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.

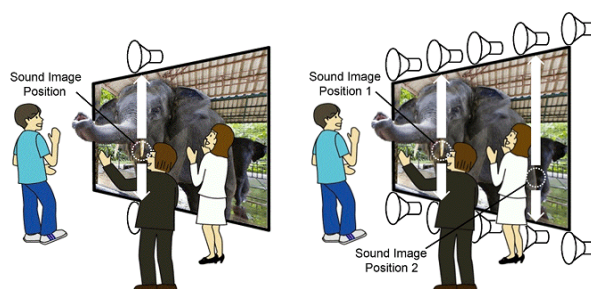


Fig. 1 Basic configuration of the multiple vertical panning (MVP) method.

さらに, Fig. 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.

提案したシステムでは視聴者はヘッドホンを装着する必要がない. また, スピーカはスクリーンの上下に配置しているので, プロジェクタアレイとスクリーンの間には何も配置されていない. また, 通信会議システムとしての応用を考えた場合, 話者の声を直接收音するだけでなく, 收音用マイクロホンの配置場所は限定されないため, スクリーンと聴取位置の間に配置する必要がない. 以上のことが提案したシステムの大きな特徴である.

## 3 視聴覚実験

### 3.1 実験環境・条件

実験は 200 インチ背面投影型映像スクリーンが設置されている会議室にて行った. 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 402 ms, 部屋の暗騒音レベルは 38 dBA だった.

Fig. 2 に示すようにスクリーンの上下に計 42 個の

\*Effect of the number of loudspeakers on sense of presence in 3D audio system using multiple vertical panning. by KIMURA, Toshiyuki, ANDO, Hiroshi (National Institute of Information and Communications Technology)

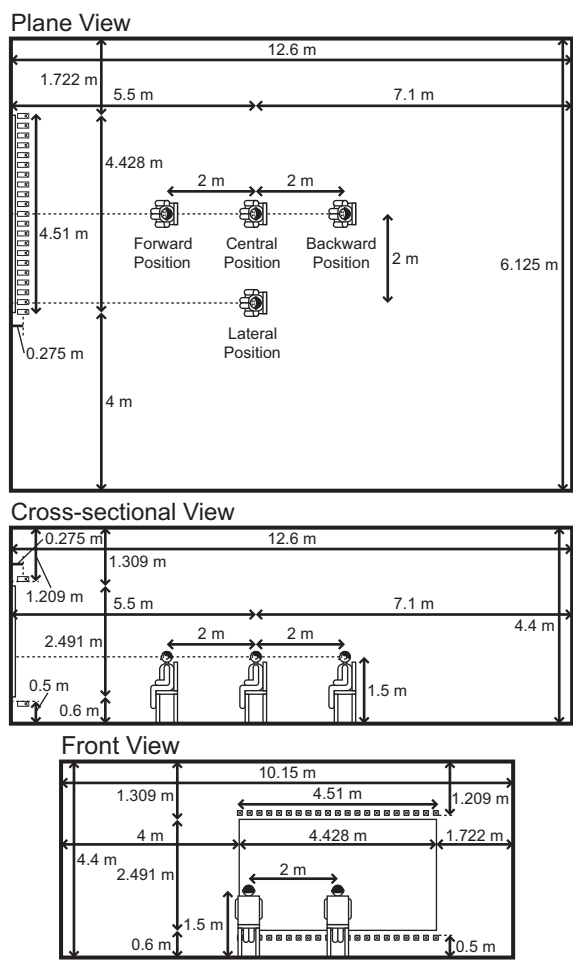


Fig. 2 Position of viewers, screen and the loud-speaker array in the audio-visual experiment.

スピーカを配置した。スクリーンは会議室の壁面に設置されており、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 0.275 m 前方に配置した。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(11 cm(幅)×25 cm(奥行)×11 cm(高さ))に取り付けたものを用いた。大画面裸眼立体映像提示システムの視聴範囲(最適視聴距離は 5.5 m)を考慮して、スクリーンから 3.5, 5.5, 7.5 m 離れた地点に 3 か所の視聴位置(以降「前方位置」「中心位置」「後方位置」と呼ぶ)を設定した。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 5.5 m 離れた場合、スクリーン正面を中心として左右に 2 m である。そこで、スクリーン正面から 5.5 m 離れ、さらに 2 m 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。計 4 か所の視聴位置の高さは視聴者の耳の位置において 1.5 m とした。音圧レベルは中心位置において約 70 dBA に設定した。

実験に用いた立体映像を Fig. 3 に示す。この立体

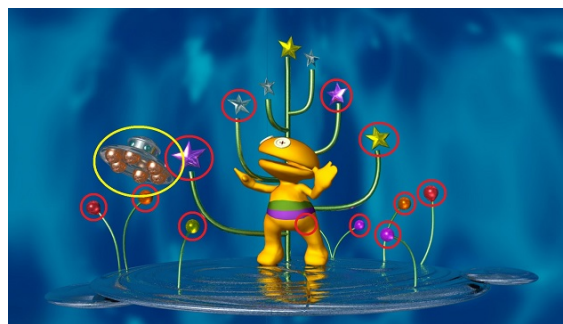


Fig. 3 3D video used in the audio-visual experiment.

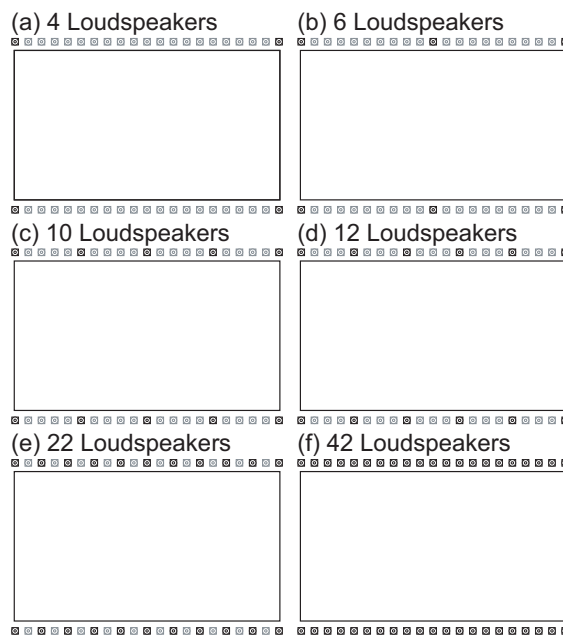


Fig. 4 Sound conditions used in the audio-visual experiment.

映像(5秒間)は UFO(Fig. 3 内の黄色の枠線)が音を出しながら画面内を動き回り、UFO が星やボール(Fig. 3 内の赤色の枠線)に当たると星やボールの場所から音が再生されるという内容である。立体映像の最適視聴距離は 5.5 m, 視差は 0.0625 m とした。また、大画面裸眼立体映像提示システムは視聴位置が左右方向に移動す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左右方向に移動するたびにその位置に合わせた立体映像を提示した。

本実験で設定した音の再生条件を Fig. 4 に示す。条件 (a)-(f) 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。各条件における再生音の合成は以下の手順により実施した。まずは提示する立体像の横位置を元に下記の式によってスクリーン上下に

配置した2個の再生用スピーカを選択する.

$$P'_H = \Delta d \text{round} \left( \frac{P_H + 2.2}{\Delta d} \right) - 2.2 \quad (1)$$

但し,  $P_H$  ( $=-2.214 \sim 2.214$ ) は立体像の横位置,  $P'_H$  ( $=2.2, \dots, 2.2$ ) は選択した2個の再生用スピーカの横位置を表し, 0 がスクリーンの横中心位置に相当する. また,  $\Delta d$  は左右のスピーカ間隔を表し, Fig. 4 の条件順に 4.4, 2.2, 1.1, 0.88, 0.44, 0.22 m となる. 次に, 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する.

$$x_U(n) = a_U s(n) \quad (2)$$

$$x_D(n) = a_D s(n) \quad (3)$$

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  $\Delta A$  [dB] から下記のように定義される.

$$a_U = \frac{10^{\frac{\Delta A}{20}}}{\sqrt{10^{\frac{\Delta A}{10}} + 1}} \quad (4)$$

$$a_D = \frac{1}{\sqrt{10^{\frac{\Delta A}{10}} + 1}} \quad (5)$$

本実験では音量差  $\Delta A$  は先行研究 [3] に従い, 下記のように設定した.

$$\Delta A = \frac{\alpha P_V + 0.1437}{0.1065} \quad (6)$$

但し,  $\alpha$  ( $=\frac{2.7}{2.5}$ ) は先行研究 [3] では上下のスピーカ間隔が 2.5 m であったのに対し, 今回の実験における上下のスピーカ間隔が 2.7 m であることにより設けた補正係数である. また,  $P_V$  は立体像の縦位置を表し, -0.3455 が視聴者の耳の高さに相当する.

### 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 29-39 歳の計 9 名 (男性 5 名, 女性 4 名) である. 評価手法には Scheffé の一対比較法 (浦の変法) [4] を用いた. 聴取実験の流れ図を図 5 に示す. まず, 評価指標を 2 種類 (「音の位置の一致度」「音の動きの一致度」) 設定し, 実験全体を評価指標ごとに 2 つに分割した. 「音の位置の一致度」とは「星やボール (Fig. 3 内の赤色の枠線) の音が常に映像の位置で鳴っているように感じるか」ということを表し, 「音の動きの一致度」は「UFO (Fig. 3 内の黄色の枠線) の音が常に映像の位置に合わせて動いているように感じるか」ということを表す. さらに, 視聴位置ごとに分割を行い, 計 8 つのセッションを設定した. 各セッションにおいて 6 回の練習試行の後, 30 回の本試行を行った. 練習試行は図 4 における 3 つの実験条件 [条件 (a),

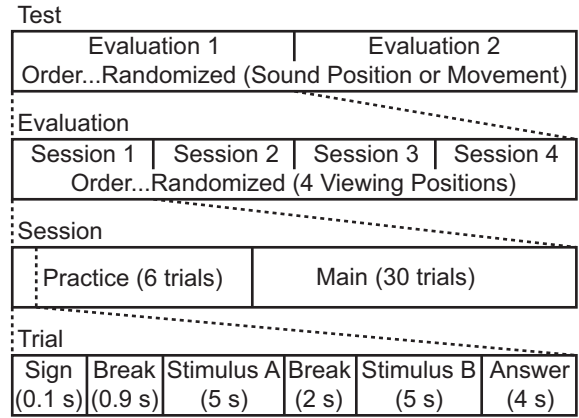


Fig. 5 Flowchart of the audio-visual experiment.

Table 1 Scale of Scheffé's paired comparison.

Grade	Judgment
3	Very good
2	Fairly good
1	Little good
0	The same
-1	Little bad
-2	Fairly bad
-3	Very bad

(b), (f) の順列の総数 ( $3 \times 2 = 6$ ) であり, 本試行は図 4 における 6 つの実験条件 [条件 (a)-(f)] の順列の総数 ( $6 \times 5 = 30$ ) である. 評価指標, 視聴位置及び試行の提示順序は視聴者ごとにランダム化した.

視聴者には先に提示した刺激音 (A) を基準にして後に提示した刺激音 (B) の音の位置や動きの一致度を聴取者に評定させた. その際, 評定は表 1 に示すような 7 段階とした. また, 評定の際には視聴者は頭部や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

得られた実験結果を 8 種類 [=評価指標 (2)  $\times$  視聴位置 (4)] のセッションごとに Scheffé の一対比較法 (浦の変法) に基づいて分散分析を行ったところ, 1 つのセッション (音の位置, 後方位置) を除くすべてのセッションにおいて実験条件に基づく主効果が 0.1% 有意であった. その結果, スピーカ数の変化による臨場感への影響に有意差が見られると判断されたので, セッションごとに平均評定値を算出し, 評定値の差によってスピーカ数による影響を検討することにした.

実験条件ごとの平均評定値を評価指標ごとに Figs. 6-7 に示す. エラーバーはヤードスティックによる 95% 信頼区間を表す. 条件 (a) では音像が左端や右端に偏っており, 映像位置と音像位置のずれを明らかに知覚で

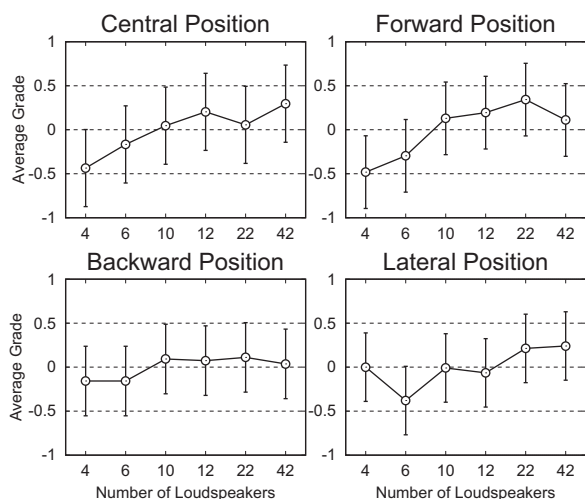


Fig. 6 Results of the audio-visual experiment (Sound location).

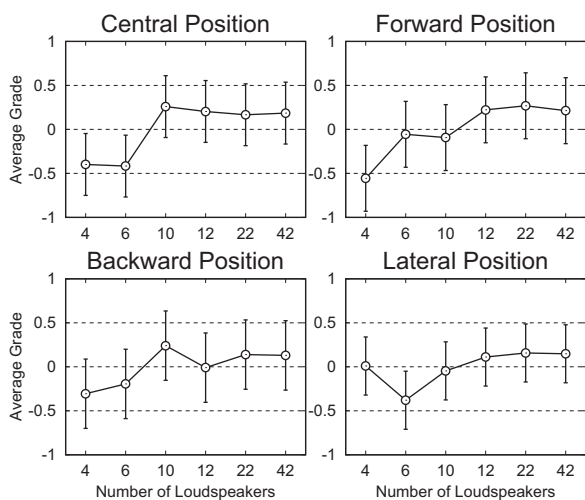


Fig. 7 Results of the audio-visual experiment (Sound movement).

きるの、側方位置を除いて他の条件よりも平均評定値が低くなっている。ゆえに、この結果から見ると視聴者は臨場感の違いをおおむね正しく識別していると考えられる。

検討した実験条件の中で最も平均評定値が高い条件を基準としてスピーカ数による臨場感への影響を検討する。スピーカ数が4と6の場合は最も平均評定値が高い条件よりも5%有意で平均評定値が低い場合が存在する。一方、スピーカ数が10, 12, 22, 42の場合は全てのセッションにおいて最も平均評定値が高い条件との有意差は見られない。つまり、スピーカ数が10以上になると、視聴者はそれ以上にスピーカ数を増やしても臨場感の違いを識別できないということが言える。従って、MVP方式を用いて視聴覚システムを構築する場合、スピーカ数は10個(スクリーン上部及び下部にそれぞれ5個)にまで減らす

ことが可能であると考えられる。

#### 4 おわりに

本発表では、我々が従来提案している立体音響システム(MVP方式)において伝送するチャンネル数を減らすために、スピーカ数を減らしていった場合どのくらい臨場感に影響があるのかを視聴覚実験によって検討した。その結果、スピーカを10個(スクリーン上部及び下部にそれぞれ5個)にまで減らしても視聴者は臨場感の違いを識別することができないということが分かった。

今後は実用的な伝送システムの確立のために、スピーカの配置を工夫するなどして伝送するチャンネル数をさらに減らすことができないか検証を行う。

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、立体映像を提供して下さいた牧野真緒研究員に感謝の意を申し上げます。本発表における視聴覚実験は独立行政法人情報通信研究機構倫理規定に従い実施した。

#### 参考文献

- [1] 榎並, 画像ラボ, 18 (9), 41-45, 2007.
- [2] 安藤ら, 情報通信研究機構報道発表, 2011. <http://www.nict.go.jp/press/2011/01/25-1.html>
- [3] Kimura *et al.*, ITE Trans. on Media Tech. and App., 2 (1), pp. 33-45, 2014.
- [4] 日科技連官能検査委員会編, “官能検査ハンドブック,” 日科技連出版社, 1973, pp. 366-374.