

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

○木村敏幸 (東北学院大学)

1 はじめに

著者はこれまでに大画面ディスプレイ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 [1]. さらに、スピーカをディスプレイ上部及び下部に5個ずつ配置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も示してきた [2].

しかしながら、先行研究 [2] におけるスピーカ配置は、ステレオホニックのような従来のスピーカ配置との互換を取るために必ずディスプレイの4隅にスピーカを配置していた. 一方で、MVP方式を用いた遠隔通信会議システムを構築する場合、従来のスピーカ配置と互換をとる必要性がないため、ディスプレイの4隅にスピーカを配置する必要がなくなる.

そこで、本発表ではディスプレイの4隅にスピーカを配置しないスピーカ配置を設定し、臨場感に及ぼす影響を視聴覚実験によって評価する.

2 視聴覚実験

2.1 実験環境・条件

実験は研究室の実験室の一角にて行った. 実験室の残響時間は約 500 ms, 部屋の暗騒音レベルは 38 dBA だった. Fig. 1 に示すように映像ディスプレイの上下に 18 個, 左右に 2 個の計 20 個のスピーカを配置した. 映像ディスプレイには 55 インチの 4K テレビ (TOSHIBA: 55J20X), スピーカには市販のスピーカユニット (ダイトーボイス: AR-10N) を市販の密閉型エンクロージャ (ダイトーボイス: EX-10 BK) に取り付けたものを使用した. 4K テレビの最適視聴距離 (テレビの高さの 1.5 倍) を考慮して、映像ディスプレイから 1.05 m 離れた地点に 2 カ所の視聴位置 (以降「中心位置」「側方位置」と呼ぶ) を設定した. 2 カ所の視聴位置の間隔は 0.54 m とした. 設定した視聴位置に椅子を設置し、視聴者を椅子に座らせた. 視聴位置の高さは視聴者の頭部中心において 1.44 m とした. 音圧レベルは中心位置での視聴者の頭部中心において約 70 dBA に設定した.

実験に用いた映像を Fig. 2 に示す. この映像は先

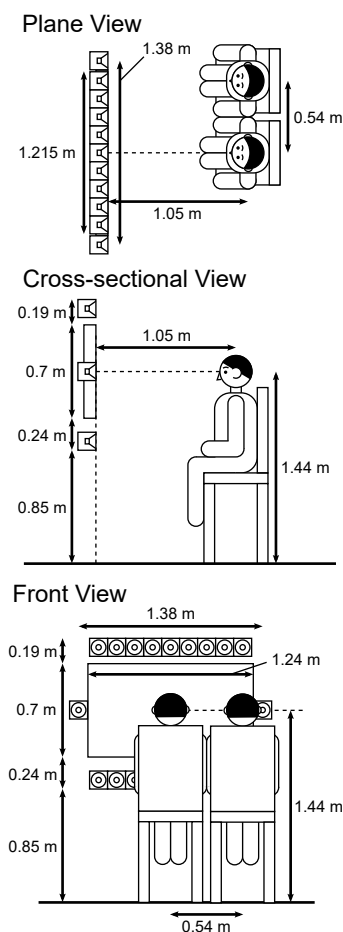


Fig. 1 Position of viewers, visual display and the loudspeaker array in the audio-visual experiment.

行研究 [2] で用いられたものと同じであり、5 秒間の間に UFO (Fig. 2 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. 2 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.

本実験で設定した音の再生条件を Fig. 3 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し、黒色の矢印は再生するスピーカ対による水平及び垂直パニングを表す. 条件 (a) は従来手法 (ステレオホニック方式) の条件であり、手法を比較するために設定した. 一方、条件 (b) は先行研究 [2] において明らかに臨場感の違いがあると判断された条件であり、今回の実験における視聴者の臨場感の識別能力を検査するために設定した.

MVP 方式を用いる条件 (b)-(f) において、時間

* Effect of the loudspeaker arrangement on sense of presence in 3D audio system using multiple vertical panning. by KIMURA, Toshiyuki (Tohoku Gakuin University)

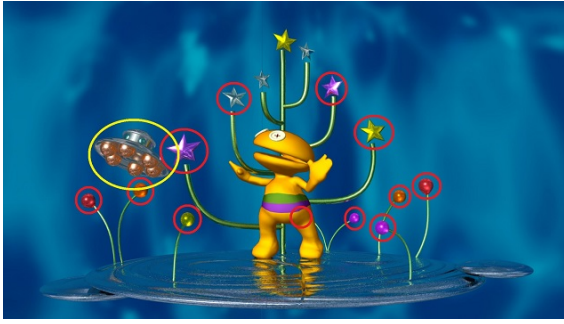


Fig. 2 Video used in the audio-visual experiment.

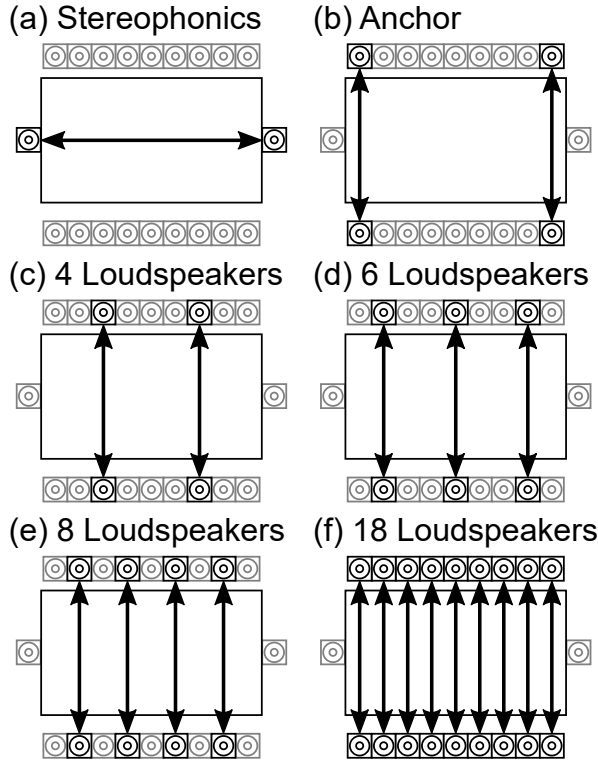


Fig. 3 Sound conditions used in the audio-visual experiment.

$T = \frac{m-1}{F_v}$ に提示オブジェクト (UFO, 星, ボール) 位置 (P_H, P_V) より再生する音は以下の手順により合成した. 但し, $F_v (=30 \text{ fps})$, $m (=1, \dots)$ は映像信号のフレームレートとフレーム番号を表す. 一方, $P_H (= -0.62 \sim 0.62)$, $P_V (= -0.35 \sim 0.35)$ はオブジェクトの水平及び垂直位置を表し, P_H が 0 なら, 水平位置はディスプレイの横中心位置に相当し, P_V が 0 なら, 垂直位置は視聴位置の高さに相当する.

まずは実験条件におけるスピーカ配置を基に, オブジェクトの水平位置 P_H を下記の式のように補正する.

$$P'_H = \begin{cases} P_L & (\text{if } P_H < P_L) \\ P_H & (\text{if } P_L \leq P_H \leq P_R) \\ P_R & (\text{if } P_H > P_R) \end{cases} \quad (1)$$

但し, P'_H は補正された水平位置を表し, P_L, P_R は

スピーカ配置における最も左側及び右側のスピーカの水平位置を表す.

次に, P'_H を基に, 下記の式によってディスプレイ上下に配置した 2 個の再生用スピーカを選択する.

$$P''_H = \Delta d \text{round} \left(\frac{P'_H - P_L}{\Delta d} \right) + P_L \quad (2)$$

但し, $P''_H (= -0.54, \dots, 0.54)$ は選択された 2 個の再生用スピーカの水平位置を表す. また, Δd は左右のスピーカ間隔を表し, Fig. 3 の条件 (b)-(f) の順に 1.08, 0.54, 0.405, 0.27, 0.135 m となる.

最後に, 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.

$$x_U(n) = a_U w(n) s(n) \quad (3)$$

$$x_D(n) = a_D w(n) s(n) \quad (4)$$

$$\left(n = \frac{F_s}{F_v} (m-1), \dots, \frac{F_s}{F_v} m + LF_s \right)$$

但し, $F_s (=48 \text{ kHz})$, $n (=0, \dots)$ は音源信号の標本化周波数及び標本化時間を表し, $x_U(n)$, $x_D(n)$ は上側及び下側のスピーカから再生する音信号を表す, 音源信号には Web サイト [3] にて公開されている効果音素材を使用した. $w(n)$ は音源信号の窓関数で, 以下のように定義される.

$$w(n) = \begin{cases} \frac{1}{LF_s} \{n - \frac{F_s}{F_v} (m-1)\} & (n = \frac{F_s}{F_v} (m-1), \dots, \frac{F_s}{F_v} (m-1) + LF_s) \\ 1 & (n = \frac{F_s}{F_v} (m-1) + LF_s, \dots, \frac{F_s}{F_v} m) \\ -\frac{1}{LF_s} (n - \frac{F_s}{F_v} m) + 1 & (n = \frac{F_s}{F_v} m, \dots, \frac{F_s}{F_v} m + LF_s) \end{cases} \quad (5)$$

但し, $L (=1 \text{ ms})$ は窓関数のクロスフェード時間を表す. 一方, a_U, a_D はそれぞれのゲイン係数を表し, 音量差 $\Delta A [\text{dB}]$ から下記のように定義される.

$$a_U = \frac{10^{\frac{\Delta A}{20}}}{\sqrt{10^{\frac{\Delta A}{10}} + 1}} \quad (6)$$

$$a_D = \frac{1}{\sqrt{10^{\frac{\Delta A}{10}} + 1}} \quad (7)$$

本実験では音量差 ΔA は先行研究 [1] に従い, 下記のように設定した.

$$\Delta A = \frac{P'_V + 0.1437}{0.1065} \quad (8)$$

但し, $P'_V (= \frac{2.5}{0.86} P_V + 0.0654)$ は先行研究 [1] とは上下のスピーカ間隔が異なることにより設けた補正垂直位置である.

一方, 条件 (a) 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ディスプレイの左右に配置した2個のスピーカから再生した。

$$x_L(n) = a_L w(n)s(n), \quad (9)$$

$$x_R(n) = a_R w(n)s(n) \quad (10)$$

$$\left(n = \frac{F_s}{F_v}(m-1), \dots, \frac{F_s}{F_v}m + LF_s \right)$$

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を表す。 a_L 、 a_R はそれぞれのゲイン係数を表し、本実験ではタンゼント則 [4] に従い下記の式のように設定した。

$$a_L = \frac{1 - \frac{\tan\theta}{\tan\theta_0}}{\sqrt{2 \left\{ 1 + \left(\frac{\tan\theta}{\tan\theta_0} \right)^2 \right\}}} \quad (11)$$

$$a_R = \frac{1 + \frac{\tan\theta}{\tan\theta_0}}{\sqrt{2 \left\{ 1 + \left(\frac{\tan\theta}{\tan\theta_0} \right)^2 \right\}}} \quad (12)$$

但し、 θ 及び θ_0 はディスプレイ正面の視聴位置における音像及び左右の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。

$$\frac{\tan\theta}{\tan\theta_0} = \frac{P_H}{1.38} \quad (13)$$

但し、 P_H はオブジェクトの水平位置を表す。

2.2 実験計画・手順

実験の流れ図を Fig. 4 に示す。まず、先行研究 [2] と同じように評価指標を2種類(「音の位置の一致度」「音の動きの一致度」)設定し、評価指標ごとに視聴者を7名ずつ割り当てた。「音の位置の一致度」とは「星やボール (Fig. 2 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. 2 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。評価手法には Scheffé の一対比較法 (浦の変法) [5] を用いた。実験では、視聴位置が異なる計2つのセッションを設定した。各セッションにおいて1回の練習試行の後、30回の本試行を行った。本試行は Fig. 3 における6つの実験条件 [条件 (a)-(f)] の順列の総数 (6×5=30) である。視聴位置及び本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 Table 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。

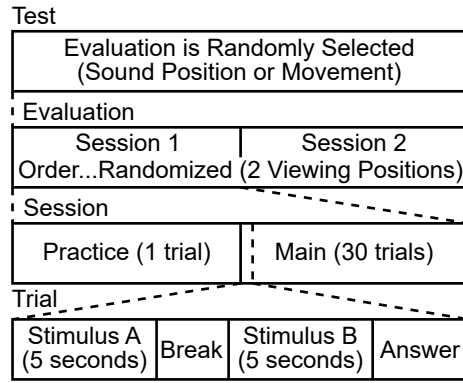


Fig. 4 Flowchart of the audio-visual experiment.

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

2.3 実験結果及び考察

得られた実験結果を計4種類 [=評価指標 (2)×視聴位置 (2)] 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 (音の動き、側方位置) 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、スピーカ配置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配置による影響を検討することにした。視聴位置及び評価指標ごとの平均評定値を Figs. 5-6 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。

MVP方式はスピーカの数が多いほどより臨場感を表現することができるため、スピーカ数が最も多い条件 (f) が最も臨場感が高い条件と言える。一方、臨場感を保持するためには、すべてのセッションにおいて視聴者が条件 (f) と同等の臨場感を感じていることが必要となる。そこで、条件 (f) を基に臨場感への影響を評価することにした。条件 (f) を基準とした境界線を Figs. 5-6 中の黒色の太点線で示す。さらに、5%水準での条件 (f) との有意差を Figs. 5-6 中の矢印で示す。

従来手法である条件 (a) においては、中心位置では条件 (f) との有意差は見られないものの、側方位置では音の位置の一致度を評価する場合において条件 (f)

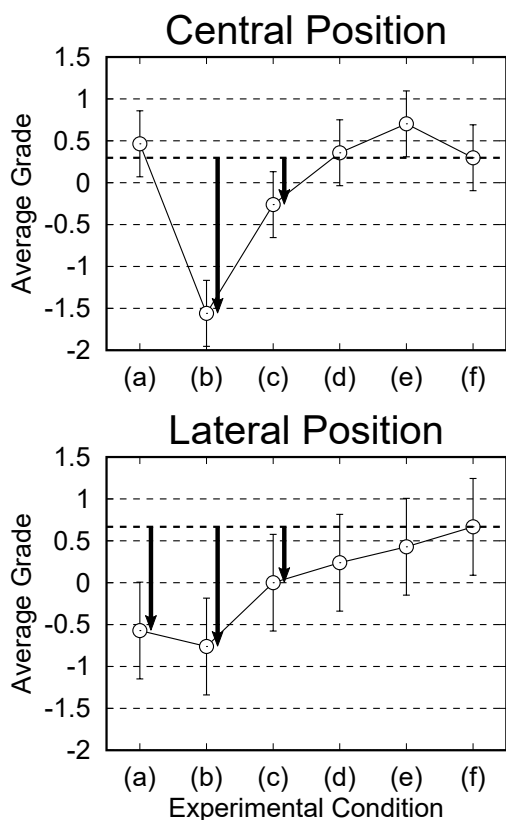


Fig. 5 Results of the audio-visual experiment (Sound location).

より有意に平均評定値が低下した。従って、先行研究 [1] においても示したように、従来手法は中心位置でしか臨場感を表現できない手法であるのに対して、MVP 方式はどの視聴位置においても臨場感を表現できる手法であると言える。

一方、MVP 方式による条件 (b)-(e) においては、条件 (b) や (c) では一部のセッションにおいて条件 (f) よりも有意に平均評定値が低下しているが、条件 (d) と (e) では条件 (f) よりも有意に平均評定値が低下しなかった。従って、MVP 方式においてディスプレイの 4 隅にスピーカを配置しないスピーカ配置を適用すれば、必要なスピーカ数は 6 個 (ディスプレイ上部及び下部にそれぞれ 3 個) にまで削減することができると言える。

3 おわりに

本発表では、著者がこれまでに提案している立体音響システム (MVP 方式) において、ディスプレイの 4 隅にスピーカを配置しないスピーカ配置が臨場感に及ぼす影響を視聴覚実験によって検討した。その結果、ディスプレイの 4 隅にスピーカを配置しなければ、必要なスピーカの数 6 個 (スクリーン上部及び下部にそれぞれ 3 個) にまで減らすことができると言うことが分かった。

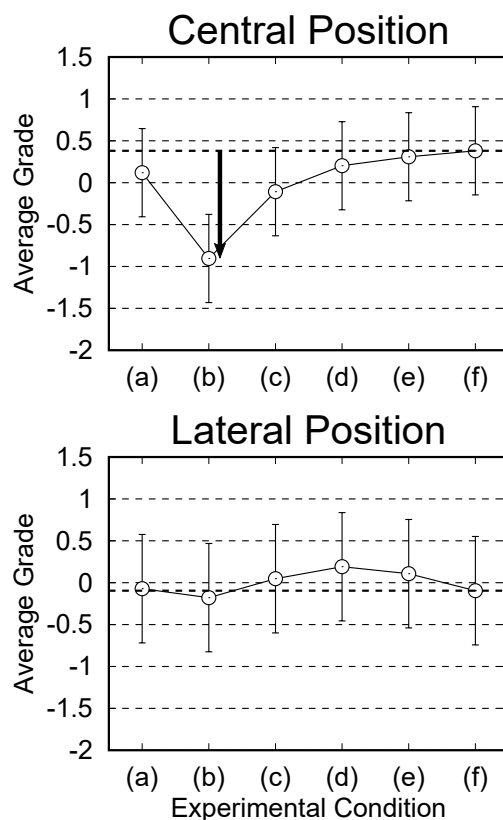


Fig. 6 Results of the audio-visual experiment (Sound movement).

今後はより実用的なシステムを確立するために、水平パニングを取り入れるなどの工夫によっ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 研究の遂行にあたり、視聴覚実験を実施して下さった大葉翼氏に感謝の意を申し上げる。

参考文献

- [1] T. Kimura and H. Ando, "3D Audio System Using Multiple Vertical Panning for Large-screen Multi-view 3D Video Display," *ITE Trans. on Media Tech. and App.*, 2 (1), pp. 33–45, 2014.
- [2] T. Kimura and H. Ando, "Effect of Discretization of Sound Source Position on Sense of Presence in 3D Audio Systems Using Multiple Vertical Panning," *Trans. VR Soc. Jpn.*, 20 (3), pp. 179–188, 2015.
- [3] Pocket Sound: <https://pocket-se.info/>.
- [4] B. Bernfeld, "Attempts for Better Understanding of the Directional Stereophonic Listening Mechanism", *Proc. 44th Audio Eng. Soc. Convention, C-4*, pp. 1–24, Feb. 1973.
- [5] 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.