

1 はじめに

近年、音場再現技術に関する研究が盛んに行われている。この技術を用いることで自室に居ながらにして別の空間にある物体や音を体験することが可能となる。本報告ではこの技術を応用し、自由視点テレビシステム [1] に聴覚情報を組み込むことを目的としている。

自由視点テレビシステムは、対象物の周りにカメラアレーを配置することで、カメラアレーの内側の任意の点における画像を生成することができる技術である。これに対応した音場を対象物の周りに再現するにはマイクロホンアレーを対象物の周りに設置し、収録した仮想音源信号を対応するスピーカアレーから再生させる必要がある [2, 3]。しかし、カメラアレーの内側にマイクロホンアレーを設置するとカメラにマイクロホンが映り込んでしまうので実用上問題が生じる。

そこで、この問題を解決するために、本報告ではカメラアレーと同じ場所に設置したマイクロホンアレーからカメラアレー内部に存在する仮想音源信号を逆フィルタ処理 [4] によって算出する手法を提案し、その性能を評価する。

2 音場再現システム

2.1 提案手法

図 1(左) のようにマイクロホンアレーを実音源の周りに配置し、実音源による音場を収録する。但し、カメラアレーはマイクロホンアレーと同じ位置とする。一方、実音源による音場は図 1(右) のように配置されたスピーカ (仮想音源信号) によって得られると仮定する。ゆえに、仮想音源信号をスピーカアレーから再生すれば、Huygens の原理によって得られた音場は実音源による音場と等しくなると考えられる。

2.2 仮想音源信号の推定

仮想音源信号 $S_k(\omega)$ の推定は以下の手順によって行う。まず原音場におけるマイクロホンアレーから信号 $X_l(\omega)$ を収録する。次に仮想音源からマイクロホンまでの伝達関数 $G_{kl}(\omega)$ から、逆伝達関数 $H_{lm}(\omega)$ を求める [4]。最後に $X_l(\omega)$ に $H_{lm}(\omega)$ を畳み込むことで仮想音源信号 $S_k(\omega)$ を推定する [3]。

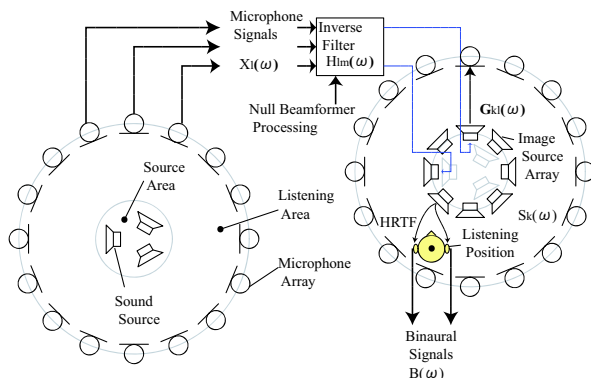


図 1: 提案する音場収録、再現システム

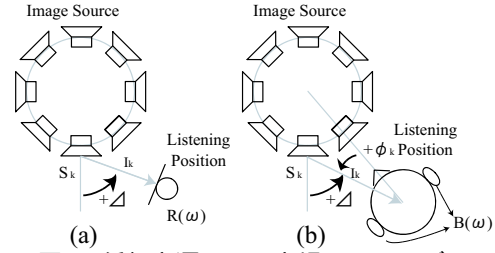


図 2: 仮想音源による音場再現のモデル

2.3 聴取位置における信号の生成

仮想音源と聴取位置の関係を図 2(a) に示す。任意の聴取位置における信号 $R(\omega)$ は仮想音源信号から式 (1) によって算出される。ここで $I_k(\omega)$ は仮想音源から聴取位置までの伝達関数、 $c_k(\Delta)$ は仮想音源における聴取位置の見込み角 Δ に応じた仮想音源の指向係数である。また、仮想音源と聴取者の関係を図 2(b) に示す。任意の聴取位置におけるバイノーラル信号 $B(\omega)$ は式 (2) によって算出される。ここで $D_k(\phi_k, \omega)$ は聴取位置から仮想音源に対する見込み角 ϕ_k に応じた HRTF である。

$$R(\omega) = \sum_{k=1}^N c_k(\Delta) I_k(\omega) S_k(\omega) \quad (1)$$

$$B(\omega) = \sum_{k=1}^N c_k(\Delta) D_k(\phi_k, \omega) I_k(\omega) S_k(\omega) \quad (2)$$

3 再現信号に対する分析

3.1 信号生成条件

図 3 に示すように実音源、仮想音源、マイクロホンアレーを配置した。実音源は同心円の中心からの距離 $r_1 (= 0.6\text{m})$ 、方位角 $\theta = 45^\circ$ の位置に配置した。仮想音源の数を N とするとき、中心から N 番目の仮想音源の向かって延ばした半直線を方位角の基準 $\theta = 0^\circ$ とした。仮想音源は半径 $r_2 (= 0.8\text{m})$ の円周上に均等に配置し、検討する仮想音源の数は 12、18、24、36、48 の 5 種類とした。マイクロホンは半径 $r_3 (= 2.1\text{m})$ 上に均等に配置し、検討する仮想音源の数の 4 倍とした [4]。なお、サンプリング周波数は 32kHz、音速を 340m/s とした。また、伝達関数 $G_{kl}(\omega)$ は各機器の幾何学的な距離から算出される遅延と距離減衰を与えたデルタ関数として近似した。また、式 (1)、(2) における $c_k(\Delta)$ を次のように与えた。式 (3) より、仮想音源は外向きの指向性を持つ。

$$c_k(\Delta) = \begin{cases} \cos(\Delta) & -90^\circ < \Delta < 90^\circ \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

3.2 推定した仮想音源信号に対する評価

推定した波形の評価は式 (4) の Signal to Deviation Ratio (SDR) を用いて評価した。ここで、 $s(n)$ は基準となる仮想音源信号、 $\hat{s}(n)$ は比較対象となる仮想音源信号である。

$$\text{SDR} = 10 \log_{10} \frac{\sum_n s^2(n)}{\sum_n [s(n) - \hat{s}(n)]^2} \quad [\text{dB}] \quad (4)$$

図 4 に仮想音源数 12 のときのパワーで正規化した音源信号推定精度 (上) と各仮想音源におけるパワー比

* Sound field auralization system in free listening positions using head related transfer functions by W. Mizuno, T. Kimura, T. Nishino and K. Takeda (Nagoya Univ.)

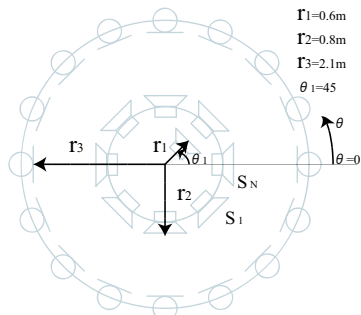


図 3: 各機器の配置図

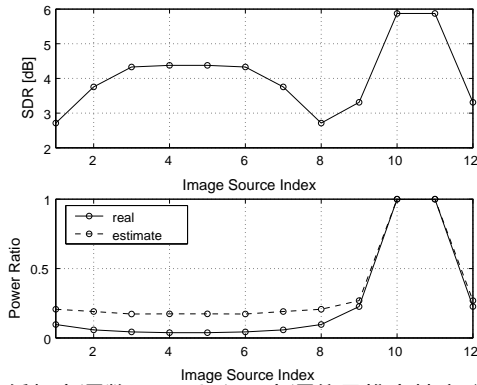


図 4: 仮想音源数 12 のときの音源信号推定精度 (上) と各仮想音源におけるパワー比 (下)

(下) を示す。実音源に近い位置ほど仮想音源信号の推定精度が高くなっており、平均して 4dB 程度であることが分かる。また、仮想音源間のパワー比においては実際に収録した信号ではパワーが低い仮想音源に対し、推定した信号ではパワーが十分に低くなっていないことが分かる。そこで、この結果が方向定位に及ぼす影響について主観評価実験による調査を行った。

3.3 聴取位置で再現された音圧信号に対する評価

図 5 に方位角 $\theta = 0^\circ$ における $r(n)$ の SDR を距離別に示す。図より、再現位置が仮想音源に近いときは仮想音源数による SDR の差は小さいが、距離が離れるにつれて仮想音源数が多いほど高い SDR を持つことが分かる。これは、再現位置が仮想音源に近いときは音場再現に用いる仮想音源数が少なく、正しい再現信号が合成されていないためであると考えられる。また、距離が離れることで仮想音源数が多い場合には音場再現に寄与する仮想音源の数が増え、再現精度が上昇すると考えられる。

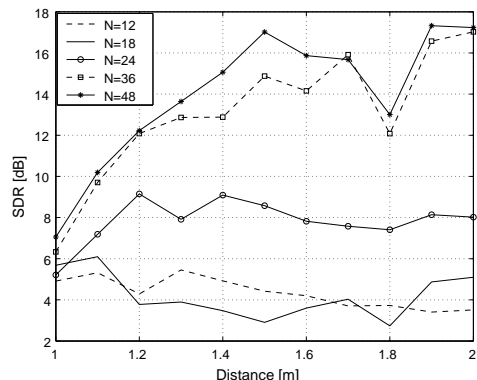


図 5: 方位角 $\theta = 0^\circ$ における距離別の SDR

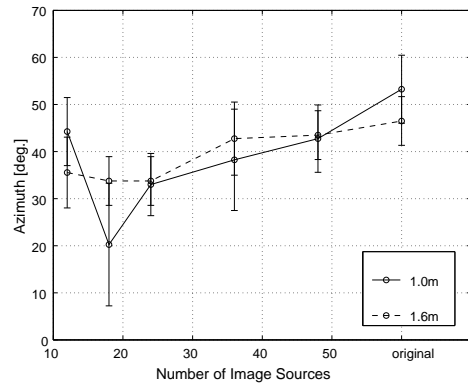


図 6: 方位各 $\theta = 135^\circ$ における実験結果

3.4 聴取実験によるバイノーラル信号の評価

バイノーラル信号の評価を聴取実験により行った。被験者は 22 歳から 24 歳までの男子学生 5 名であり、1 名につき 4 試行の実験を行った。実験に対する回答は音像が定位したと感じられる方位に印を打つ形式とした。図 3 における方位角 $\theta = 135^\circ$ に対して距離を変えて音場を再現した信号を提示音として用いた。得られた回答方向を図 6 に示す。図 6 において、横軸は仮想音源数、縦軸は全 20 回答における平均値であり、エラーバーは 95% 信頼区間を示す。original は実音源に直接 HRTF を畳み込んだ信号にを用いた場合の回答であり、正しく音場が再現された場合にはこの値が回答されると考えられ、本実験における真値とした。評価結果より、仮想音源数 18、距離 1.0m においてのみ真値との誤差が非常に大きい。この現象は仮想音源数が 18 という条件では仮想音源は円周上に 20° ごとに配置され、方位角 $\theta = 135^\circ$ においては、正面に仮想音源が配置されないため、聴取位置に最も近い仮想音源が $\theta = 140^\circ$ に配置されたために方向定位に悪影響を及ぼしたと考えられる。このような結果は仮想音源を密に配置することによって避けられる。また、図 6 において、仮想音源数 18 という条件でも距離が 1.6m になることによって再現信号の合成に寄与する仮想音源の数が増えることで再現性が向上する。

4 まとめ

本報告では自由視点テレビシステムに組み込むための音場再現システムを提案し、シミュレーションと主観実験によって性能を検討した。その結果、本システムにおいて音場再現の可能性が示唆された。今後はさらに再現精度向上について検討する必要がある。

参考文献

- [1] T.Fujii and M.Tanimoto, "Free-viewpoint Television based on the Ray-Space representation", SPIE ITCOM 2002, pp.175-189, August 2002.
- [2] 伊勢史郎, "キルヒホッフ-ヘルムホルツ積分方程式と逆システム理論に基づく音場制御の原理," 音響学会誌, Vol.53, No.9, pp.706-713, 1997
- [3] 木村敏幸, 箕一彦, 武田一哉, 板倉文忠, "音場再現におけるマルチチャンネル信号の空間符号化," 日本バーチャルリアリティ学会論文誌, Vol.8, No.4, pp.433-442, 2003
- [4] 神沼充伸, 伊勢史郎, 鹿野清宏, "周波数領域における最小ノルム解を利用した多チャンネル音場再現システムにおける逆フィルタの設計," 音響学会誌, Vol.57, No.3, pp.175-183, 2001