

位置情報による移動音源再現手法の主観的影響*

木村敏幸 (名大・情報科学), 箕一彦 (中京大・情報科学),
武田一哉 (名大・情報科学), 板倉文忠 (名城大・理工)

1 はじめに

波面合成法 [1, 2] による音場再現に必要なチャンネル信号の数は一般的に非常に多いため, 遠隔システムを実現するには伝送量を削減する必要がある. 我々は今までに伝送量をチャンネルの数から音源の数に削減する空間符号化手法を提案し, その有効性を示してきた [3]. 空間符号化手法を移動音源に適用する場合, 位置情報による移動音源再現手法を用いるので, 本報告ではその影響を主観評価実験によって検討する.

2 アルゴリズム

図 1 に移動音源再現手法のブロック図を示す. 原音場内

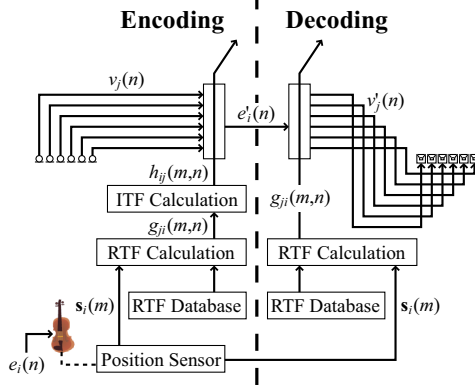


図 1: 移動音源再現手法のブロック図

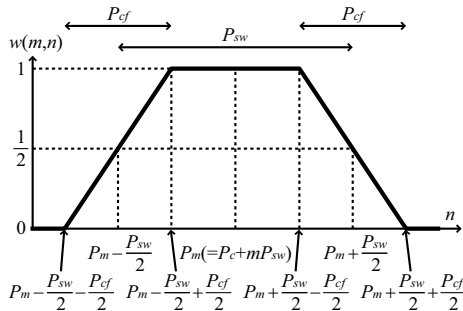


図 2: 窓関数の形状

に配置したマイクロホンアレーでチャンネル信号 $v_j(n)$ を収録し, 同時に移動音源 $e_i(n)$ の位置情報 $s_i(m)$ も位置センサで収録する. 但し, m, n はサンプリング周波数 F_p, F_s [Hz] による離散時間である. 送信側ではまず $s_i(m)$ から時間 m における室内伝達関数 $g_{ji}(m, n)$ を求める. 次に, $g_{ji}(m, n)$ から時間 m における逆伝達関数 $h_{ij}(m, n)$ を算出する [3].

$$\mathbf{H}(m, \omega) = \mathbf{G}^+(m, \omega)\mathbf{D}(\omega) \quad (1)$$

但し, $\mathbf{G}^+(m, \omega)$ は $\mathbf{G}(m, \omega)$ の Moore-Penrose 擬似逆行列で, $\mathbf{H}(m, \omega), \mathbf{G}(m, \omega), \mathbf{D}(\omega)$ は以下のように定義される.

$$\begin{aligned} \mathbf{H}(m, \omega) &= [H_{ij}(m, \omega)] & (M \times N) \\ \mathbf{G}(m, \omega) &= [G_{ji}(m, \omega)] & (N \times M) \\ \mathbf{D}(\omega) &= \text{diag}(e^{-j\omega P_c}, \dots, e^{-j\omega P_c}) & (N \times N) \end{aligned} \quad (2)$$

但し, $H_{ij}(m, \omega), G_{ji}(m, \omega)$ は $h_{ij}(m, n), g_{ji}(m, n)$ の Fourier 変換係数で, $P_c (= F_s T_c)$ は因果律を満たす逆伝達関数を算出するために必要な遅延サンプル数である. 最

*Subjective Affection of the Method to Reproduce Moving Sound Sources by Position Information by T. Kimura, K. Kakehi, K. Takeda and F. Itakura (Nagoya Univ.)

後に, (3) 式によって移動音源信号 $e'_i(n)$ を抽出する.

$$e'_i(n) = \sum_{m=0}^{\infty} w(m, n) \sum_{j=1}^M h_{ij}(m, n) * v_j(n) \quad (3)$$

但し, L_h は逆伝達関数の次数で, $w(m, n)$ は図 2 に示される台形窓関数である. 受信側ではまず $s_i(m)$ から $g_{ji}(m, n)$ を求め, (4) 式によってチャンネル信号 $v'_j(n)$ を復元する. 但し, L_g は室内伝達関数の次数である.

$$v'_j(n) = \sum_{m=0}^{\infty} \sum_{i=1}^N g_{ji}(m, n) * [w(m, n)e'_i(n)] \quad (4)$$

3 符号化実験

3.1 チャンネル信号の作成

チャンネル信号 $v_j(n)$ は虚像法 [4] を用いて作成した. 実験に用いた原音場を図 3 に示す. 各マイクロホンは円の外側に指向性を向けている. チャンネル信号の作成条件を表 1

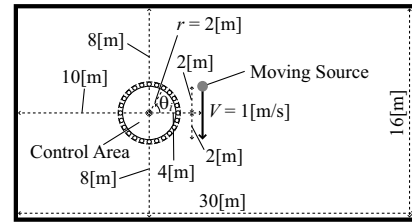


図 3: 符号化実験に用いた原音場

に示す. 音声は無響室で収録し, フルートは MIDI 音源か

表 1: チャンネル信号の作成条件

ドライソース	音声	フルート
F_s (サンプリング周波数)	48[kHz]	
音源の長さ	4[second]	
壁面の反射係数	0.5	0.7
最大の壁面反射回数	6	10
部屋の残響時間 [second]	0.6	1.0
V (音源の移動速度)	1[m/s]=3.6[km/h]	
マイクロホンの指向性	カージオイド型	

ら作成したものを用了.

3.2 室内伝達関数データベースの作成

図 3 中の音源の軌跡を 8.33mm ごとに区切り, 得られた計 481 点の音源位置から 24 点の受音点位置までの室内伝達関数データベースを虚像法によって作成した. その際, 作成条件は表 1 と同じ条件とした.

3.3 音源信号の抽出及びチャンネル信号の復元

まず, 先ほど求めたデータベースから $g_{ji}(m, n)$ を求めた. 次に, $g_{ji}(m, n)$ から $h_{ij}(m, n)$ を算出した. 得られた $h_{ij}(m, n)$ を $v_j(n)$ に畳み込み, $e'_i(n)$ を抽出した. さらに, $e'_i(n)$ に $g_{ji}(m, n)$ を畳み込んで $v'_j(n)$ を復元した. $h_{ij}(m, n)$ の算出や畳み込みの際の $w(m, n)$ に用いた条件を表 2 に示す.

表 2: 逆伝達関数の算出及び窓関数の条件

残響時間 [second]	0.6	1.0
FFT フレーム長 [sample]	65536	131072
符号化遅延時間 T_c ($=P_c = T_c F_s$)	20[ms] ($=960$ [sample])	
フィルタ長 [sample]	28800	48000
サンプリング周波数 F_p ($=P_{sw} = F_s / F_p$)	30, 60, 120[Hz]	($=1600, 800, 400$ [sample])
クロスフェード時間 T_{cf} ($=P_{cf} = T_{cf} F_s$)	1, 4[ms]	($=48, 192$ [sample])

4 主観評価実験

4.1 実験環境

主観評価実験は約 80ms の低残響室にて行った。実験に用いたスピーカレーと被験者の位置を図 4 に示す。暗騒

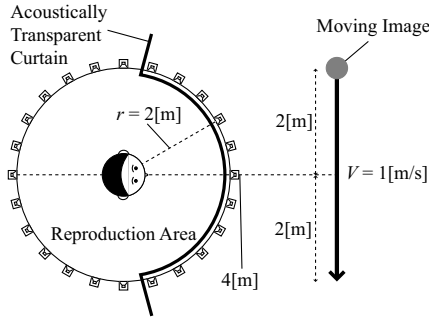


図 4: スピーカレーと被験者の位置

音レベルは 25.0dB(A)，音圧レベルは被験者の位置において約 70dB(A) とした。また，視覚の影響を低減するため，照明を薄暗くし，前方は音響透過カーテンで覆った。

4.2 実験計画

被験者は 8 名の男子学生である。ITU-R 勧告 [5] に従って三刺激二重盲験法を採用した。主観評価実験の流れを図 5 に示す。最初に提示する音 (以下「X」) には常に符号化

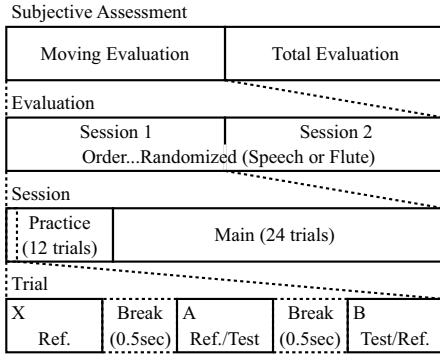


図 5: 主観評価実験の流れ図

しない原音を割り当て，2 番目に提示する音 (以下「A」) と 3 番目に提示する音 (以下「B」) には原音もしくは表 3 に示されるような 6 種類の符号化音を割り当てた。

表 3: 符号化音の種類

	1	2	3	4	5	6
F_p	30Hz	30Hz	60Hz	60Hz	120Hz	120Hz
T_{cf}	1ms	4ms	1ms	4ms	1ms	4ms

4.3 実験手順

前半の動き評価では「A」と「B」のうち「X」と同じ動き方をする音だと思った方を「5.0」，もう一方の音に関しては動き方の差を表 4 に従って「1.0」から「4.9」までで評定させた。後半の全体評価では「A」と「B」のうち「X」と同じ音だと思った方を「5.0」，もう一つの音に関しては原音との差を前半と同じように評定させた。被験者は実験中頭部を自由に動かすことができ，評定の判断がつくまでは何度も「X」「A」「B」を繰り返し聞くことができた。

4.4 実験結果及び考察

「A」と「B」のうち原音を割り当てた刺激を「5.0」と評定している被験者を検査し，8 名の内成績が上位の 3 名のデータを分析に用いることにする。符号化音の評定値から原音の評定値の差を算出した結果を図 5, 7 に示す。エラーバーは 95%信頼区間を表す。フルートの場合，評定差が音声よりも全体的に低くなっているが， F_p が 120[Hz] の場合には音の動きや全体的な知覚への影響は「差は分かるが，気にならない」程度であるので，十分実用的なシステムを構築できると考えられる。一方，音声の

表 4: 違いの評定表

差の程度	評定
差が分からない	5.0
差は分かるが，気にならない	4.0
差がわずかに気になる	3.0
差が気になる	2.0
差が非常に気になる	1.0

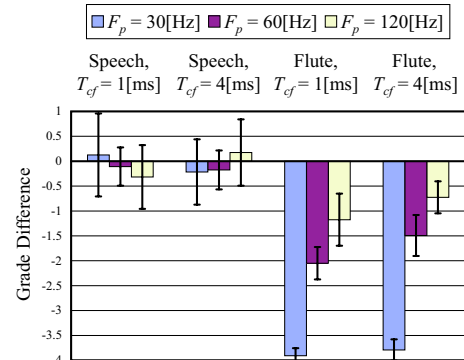


図 6: 動き評価の結果

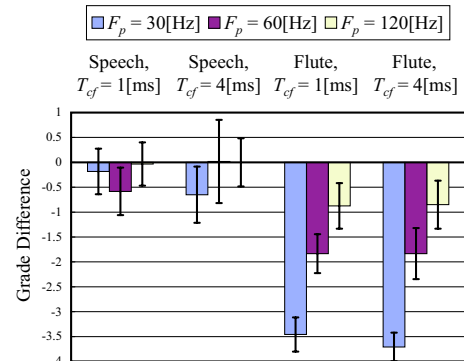


図 7: 全体評価の結果

場合， F_p が 30[Hz] の場合でも被験者は知覚的な影響を識別できていないことが多い。ゆえに，音声を対象とした遠隔通信会議システムを構築する場合にはフルートを対象としたシステムを構築する場合よりも簡易な位置検出システムによって実用的なシステムを構築できると考えられる。

5 まとめ

音源の位置情報を基に移動音源を抽出することで伝送量を削減する空間符号化手法を提案した。虚像法によって作成したチャンネル信号を対象に符号化実験を行ったところ，音の動き知覚や全体的な知覚への影響を十分許容できるシステムを実用化できるということが分かった。今後は実環境で実際に収録したチャンネル信号を用いて検討する必要がある。

参考文献

- [1] M. Camras: “Approach to recreating a sound field”, J. Acoust. Soc. Am., **43**, 6, pp. 1425–1431 (1968).
- [2] A. J. Berkhout, D. de Vries and P. Vogel: “Acoustic control by wave field synthesis”, J. Acoust. Soc. Am., **93**, 5, pp. 2764–2778 (1993).
- [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 VR 学会論文誌, **8**, 4, pp. 433–442 (2003).
- [4] J. B. Allen and D. A. Berkley: “Image method for efficiently simulating small-room acoustics”, J. Acoust. Soc. Am., **65**, 4, pp. 943–950 (1979).
- [5] ITU-R Recommendation BS.1116-1: “Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems”.