

## 音場再現におけるマルチチャネル信号の空間符号化

木村 敏幸<sup>\*1\*2</sup> 笥 一彦<sup>\*2</sup> 武田 一哉<sup>\*2</sup> 板倉 文忠<sup>\*3</sup>

## Spatial Coding of Multi-channel Audio Signals in Sound Field Reproduction

Toshiyuki Kimura<sup>\*1\*2</sup> Kazuhiko Kakehi<sup>\*2</sup> Kazuya Takeda<sup>\*2</sup> Fumitada Itakura<sup>\*3</sup>

**Abstract** – The sound field reproduction technique based on the Huygens's principle is important to realize auditory virtual reality. However, it is necessary to transmit many multi-channel signals. In this paper, a new spatial coding technique for the multi-channel signals is proposed to reduce transmission bit rate. Source signals are extracted by convolving the inverse transfer function calculated from a room transfer function to the channel signals at an encoded side. At a decoded side, channel signals are reconstructed by convolving the room transfer function to the extracted source signals. As a result, the size of data to be transmitted is reduced from the number of channels to that of sound sources. A coding experiment was performed for a real sound field. It was confirmed by subjective assessment that the perceptual distortion of the coding had little effect on the perceptual quality of sound field even if the source signals are not completely extracted.

**Keywords** : sound field reproduction, spatial coding, room transfer function, sound source extraction, sound field perception

## 1 はじめに

バーチャルリアリティ技術を応用することによって、テレビ電話のような従来のコミュニケーションに比べてより臨場感のあるコミュニケーションを実現することが可能になる。例えば、離れた場所にいる人々があたかも場を共有しているかのように会議を行ったり、まるでコンサートホールにいるかのように音楽を自分の居る場所で聴くことができるようになったりする。このようなことを聴覚的に実現する技術として、音場再現技術が近年盛んに研究されている。

音場再現は原音場の音をマイクロホンアレーによって収録し、複数のスピーカによって原音場の波面を合成する領域制御手法 [1, 2, 3, 4] によって実現することができる。この手法は無響室や非常に多くのスピーカを必要とするが、従来提案されているバイノーラル方式 [5, 6] やトランスオーラル方式 [7, 8] とは異なり、複数の受聴者が制御領域内ならどこに居ても構わないし、頭部も自由に動かすことができるというところに特徴がある。

領域制御手法をリアルタイムで遠隔に実現するためには、マイクロホンアレーで収録したマルチチャネル信号を伝送する必要がある。しかしながら、音場を再

現するために必要なマイクロホンの数は一般に非常に多いため、符号化によって伝送量を削減する必要がある。従来提案されているマルチチャネル符号化手法には DVD-video で用いられている AC-3[9] や BS デジタル放送で用いられている MPEG2 AAC[10] などが挙げられる。しかし、これらの手法は主にチャンネル内信号を対象とした情報圧縮を行っているので、音場再現に適用しても伝送量はチャンネル数に比例したままである。一方でチャンネル間の相互相関行列の主成分分析に基づく空間符号化手法 [11, 12] も提案されているが、音場再現におけるチャンネル間の相関は一般にあまり高くないので、削減できる伝送量はほんのわずかである。そこで、より効率的な符号化を行うためには新たな観点に基づいた空間符号化手法を確立することが必要である。

音場再現における空間符号化手法を確立するに当たって、次の2つの観点に着目する。一つは一般に音源の数はチャンネル数よりも少ないということである。このことは会議室における話者やコンサートホールにおけるピアノ独奏を始めとする一般の室内楽など、かなりの場合に該当していると考えられる。もう一つは全てのチャンネル信号の中には音源に関する情報が必ず含まれているということである。なぜなら、各チャンネル信号は音源信号に音場の室内伝達関数を畳み込んだ信号として表現できるからである。ゆえに、チャンネル信号から音源信号を抽出して伝送すれば、伝送量がチャンネル数から音源の数の削減されることが期待でき

\*1 日本学術振興会 特別研究員 (特別推進研究 COE)

\*2 名古屋大学大学院 情報科学研究科

\*3 名古屋大学 情報メディア教育センター

\*1 JSPS Research Fellow

\*2 Graduate School of Information Science, Nagoya Univ.

\*3 Center for Information Media Studies, Nagoya Univ.

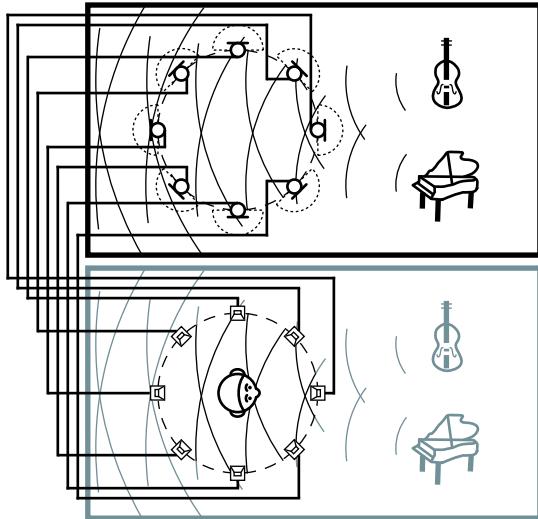


図1 Huygensの原理による音場再現の模式図  
Fig. 1 System configuration of sound field reproduction based on Huygens's principle

る。この場合、各音源信号はその音源(楽器)のすぐそばに置かれたマイクロホンによって収録することも可能である。しかし、実際にはマイクロホンを各音源のすぐそばに設置できなかつたり、他の音源からの音の混入が起こったりする。このような状況においてはマイクロホンアレーで収録したチャンネル信号から音源信号を抽出する方法が必要になる。この場合、ホールの音響特性が既知でありさえすれば、舞台上の音源配置に応じて音源収録用のマイクロホンを設置する必要がなくなるという運用上のメリットも生じる。

本論文では上記のような観点に従って、音源抽出に基づいた音場再現における空間符号化手法について検討する[13, 14, 15]。第2章では領域制御による音場再現手法と提案する空間符号化手法について述べる。第3章では実際の室内での静止音源を対象とした符号化実験について述べる。本手法は室内伝達関数を補足情報として伝送する必要があるが、本実験においては送信側と受信側であらかじめ測定しておいた室内伝達関数の情報が共有されているものと仮定した。第4章では主観評価実験によって符号化手法が人間の聴覚に及ぼす影響について検討し、本手法による知覚的な歪みは十分容認されうるものであることを示す。

## 2 アルゴリズム

### 2.1 領域制御による音場再現

図1に領域制御手法の模式図を示す。まず、原音場にマイクロホンアレーを設置する。この時、各マイクロホンは制御領域の境界上に配置され、その指向性は制御領域の外側に向けておく。そして、原音場の音を

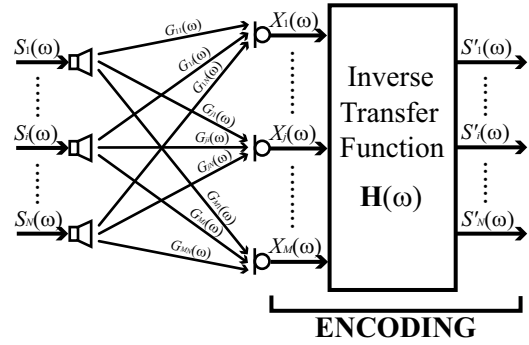


図2 音源抽出のブロック図  
Fig. 2 Block diagram of extraction of sound sources

各マイクロホンで収録する。

一方、自由空間上に制御領域と同じ形状を持った再現領域を設定し、再現領域の境界上にスピーカを配置する。各スピーカの位置はそれぞれのマイクロホンに対応させる。制御領域の境界上が十分な数のマイクロホンで覆われており、マイクロホンやスピーカの特性が十分理想的であるとするならば、収録した音をスピーカから再生すると、Huygensの原理に従って原音場の波面が再現領域内に実現され、再現領域の中にいる人は図1の灰色枠内の原音場にいるような印象を受けることになる。

### 2.2 音源信号の抽出

提案する音源抽出アルゴリズムは既知の室内伝達関数から算出した逆伝達関数をマイクロホンアレーで収録したチャンネル信号に畳み込むことによって音源信号を抽出するものである。詳細を以下に説明していく。

$N$  個の音源と  $M$  個のマイクロホンから成るシステムにおける音源信号抽出のブロック図を図2に示す。但し、 $N < M$  である。 $S_i(\omega)$  は  $i$  番目の音源信号 ( $i = 1, \dots, N$ ) で、これが求めるべき音源信号である。 $X_j(\omega)$  は  $j$  番目のマイクロホンで収録したチャンネル信号 ( $j = 1, \dots, M$ ) である。 $G_{ji}(\omega)$  は  $i$  番目の音源信号から  $j$  番目のチャンネル信号までの室内伝達関数で、これはあらかじめ知られているものとする。 $H_{ij}(\omega)$  は  $j$  番目のチャンネル信号から  $i$  番目の音源信号までの逆伝達関数であり、これは室内伝達関数から算出される。

室内伝達関数から逆伝達関数をどのようにして算出するかを説明する。全てのチャンネル信号は(1)式のように音源信号に室内伝達関数を畳み込んだ信号として表される。

$$X_j(\omega) = \sum_{i=1}^N G_{ji}(\omega) S_i(\omega) \quad (j = 1, \dots, M) \quad (1)$$

チャンネル信号に逆伝達関数を畳み込むことによって、

音源信号  $S'_k(\omega)$  を抽出する.

$$S'_k(\omega) = \sum_{j=1}^M H_{kj}(\omega) X_j(\omega) \quad (k = 1, \dots, N) \quad (2)$$

(1) 式を (2) 式に代入すると, 以下の式が得られる.

$$\sum_{j=1}^M H_{kj}(\omega) G_{ji}(\omega) = \begin{cases} S'_i(\omega)/S_i(\omega) & (i=k) \\ 0 & (i \neq k) \end{cases} \quad (3)$$

(3) 式を行列形式で書き改めると,

$$\begin{pmatrix} G_{11}(\omega) & \dots & G_{M1}(\omega) \\ \vdots & \ddots & \vdots \\ G_{1N}(\omega) & \dots & G_{MN}(\omega) \end{pmatrix} \begin{pmatrix} H_{11}(\omega) & \dots & H_{N1}(\omega) \\ \vdots & \ddots & \vdots \\ H_{1M}(\omega) & \dots & H_{NM}(\omega) \end{pmatrix} = \begin{pmatrix} S'_1(\omega)/S_1(\omega) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & S'_N(\omega)/S_N(\omega) \end{pmatrix} \quad (4)$$

元の音源信号が抽出される場合,  $S'_i(\omega)/S_i(\omega)$  は 1 になるはずである. しかし, 一般的に  $G_{ji}(\omega)$  は初期遅延を持っているので, 1 にしたときには  $H_{ij}(\omega)$  は因果律を満たさないフィルタになってしまう. そこで, 因果律を満たす逆伝達関数を算出するために  $S'_i(\omega)/S_i(\omega)$  に  $\tau_0$  点分の遅延項を設ける.

$$S'_k(\omega)/S_k(\omega) = e^{-j\omega\tau_0} \quad (k = 1, \dots, N). \quad (5)$$

室内伝達関数行列  $\mathbf{G}(\omega)$ , 逆伝達関数行列  $\mathbf{H}(\omega)$  及び単位行列  $\mathbf{D}(\omega)$  を以下の式のように定義すると,

$$\mathbf{G}(\omega) \equiv \begin{pmatrix} G_{11}(\omega) & \dots & G_{M1}(\omega) \\ \vdots & \ddots & \vdots \\ G_{1N}(\omega) & \dots & G_{MN}(\omega) \end{pmatrix} \quad (6)$$

$$\mathbf{H}(\omega) \equiv \begin{pmatrix} H_{11}(\omega) & \dots & H_{N1}(\omega) \\ \vdots & \ddots & \vdots \\ H_{1M}(\omega) & \dots & H_{NM}(\omega) \end{pmatrix} \quad (7)$$

$$\mathbf{D}(\omega) \equiv \begin{pmatrix} e^{-j\omega\tau_0} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-j\omega\tau_0} \end{pmatrix} \quad (8)$$

(4) 式から (9) 式が導出される.

$$\mathbf{G}(\omega)\mathbf{H}(\omega) = \mathbf{D}(\omega) \quad (9)$$

$N < M$  の場合には,  $\mathbf{H}(\omega)$  にはいくつかの解が存在する. ここでは Moore-Penrose 擬似逆行列を用いて  $\mathbf{H}(\omega)$  を算出する [8].

$$\mathbf{H}(\omega) = \mathbf{G}^+(\omega)\mathbf{D}(\omega) \quad (10)$$

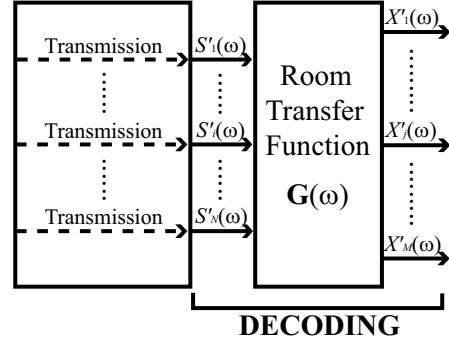


図3 チャネル復元のブロック図  
Fig.3 Block diagram of reconstruction of channel signals

ここで,  $\mathbf{G}^+(\omega)$  は  $\mathbf{G}(\omega)$  の Moore-Penrose 擬似逆行列である.

Moore-Penrose 擬似逆行列は特異値分解を用いることによって算出される.  $\mathbf{G}(\omega)$  を特異値分解すると (11) 式のようになる.

$$\mathbf{G}(\omega) = \mathbf{U}(\omega) \left( \mathbf{A}(\omega) \mid \mathbf{O} \right) \mathbf{V}^H(\omega) \quad (11)$$

ここで,  $\mathbf{U}(\omega)$ ,  $\mathbf{V}(\omega)$  はそれぞれ  $N \times N$ ,  $M \times M$  次元のユニタリ行列,  $^H$  は Hermite 共役を表す.  $\mathbf{A}(\omega)$  は  $N \times N$  次元の対角行列で,  $\mathbf{O}$  は  $N \times (M - N)$  次元の零行列である.  $\mathbf{A}(\omega)$  の対角成分は  $N$  個の特異値  $\lambda_i(\omega) (i = 1 \dots N)$  で,  $\mathbf{G}(\omega)$  のランクが  $d$  の場合,  $N$  個のうち  $d$  個が 0 でない値を持つ.  $\lambda_i(\omega)$ ,  $\mathbf{U}(\omega)$ ,  $\mathbf{V}(\omega)$  を用いると, Moore-Penrose 擬似逆行列  $\mathbf{G}^+(\omega)$  は以下のように表される.

$$\mathbf{G}^+(\omega) = \mathbf{V}(\omega) \begin{pmatrix} \mathbf{A}^+(\omega) \\ \mathbf{O}^T \end{pmatrix} \mathbf{U}^H(\omega) \quad (12)$$

但し,  $^T$  は転置を表し,  $\mathbf{A}^+(\omega)$  は以下のように定義される.

$$\mathbf{A}^+(\omega) \equiv \begin{pmatrix} \sigma_1(\omega) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_N(\omega) \end{pmatrix} \quad (13)$$

$$\sigma_i(\omega) = \begin{cases} 1/\lambda_i & (\lambda_i \neq 0) \\ 0 & (\lambda_i = 0) \end{cases} \quad (14)$$

### 2.3 チャネル信号の復元

チャネル信号復元のブロック図を図3に示す. 伝送されてきた音源信号に室内伝達関数を畳み込むことによってチャネル信号  $X'_j(\omega)$  が復元される.

$$X'_j(\omega) = \sum_{i=1}^N G_{ji}(\omega) S'_i(\omega) \quad (j = 1 \dots M) \quad (15)$$

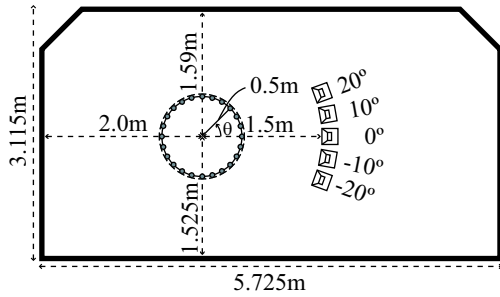


図4 音源用スピーカとマイクロホンアレーの配置図

Fig.4 Position of loudspeakers and microphone array

(5) 式より，音源信号を抽出する際には  $\tau_0$  点の遅延が生じる．一方，チャンネル信号の復元の際には遅延は生じない．従って，提案する符号化手法における符号化全体での遅延は  $\tau_0$  点となる．

また，提案手法によって取り出される信号は音源の波形信号である．ゆえに，抽出した音源信号に従来のチャンネル内符号化手法を適用することによって，伝送量を更に削減することも可能となる．

### 3 符号化実験

この章では提案した空間符号化手法の有効性を検証するために実際に行った実験について述べる．本研究では音源が静止している場合を検討した．

#### 3.1 室内伝達関数の測定

名古屋大学工学部7号館410号室の可変残響室において室内伝達関数を測定した．この部屋は壁面の吸音パネルを取り外すことによって，残響時間を150msから300msに調節することができる．部屋の天井と床を吸音状態にすることによって，2次元の音場を想定している．音源用のスピーカとマイクロホンアレーの配置図を図4に示す．スピーカとマイクロホンアレーの高さは0.835mとした．制御領域は半径0.5mの円とし，その境界上に15度おきに  $\theta = -165^\circ$  から  $180^\circ$  まで計24個のマイクロホンを設置した．円の中心から1.5mの距離に方位角  $-20^\circ \sim 20^\circ$  で10度おきにスピーカを設置した．スピーカは無指向性スピーカ(小野測器 SJ-1810)，マイクロホンは無指向性マイクロホン(SONY ECM-77S)を用いた．アレー用の輪状アルミ製フレームの側壁にマイクロホンを取り付けた．測定条件を表1に示す．音圧レベルはスピーカから1m地点での値である．サンプリング周波数48kHzの65536点TSP信号[16]を各スピーカから再生し，同期加算を16回行ってFIRフィルタ形式の室内伝達関数を測定した．

実環境における会議室やコンサートホールの残響

表1 室内伝達関数の測定条件

Table 1 Measurement conditions of room transfer function

残響時間	150ms	300ms
室温	19.5°C	19.2°C
暗騒音レベル	20.0dB(A)	19.4dB(A)
音圧レベル	90.0dB(A)	91.6dB(A)
フィルタ次数	7200	14400

表2 マルチチャンネル信号の作成条件

Table 2 Synthetic conditions of multi-channel signals

ドライソース	音声	木琴	木管五重奏
音源の数	1	1	5
音源の方位角	$0^\circ$	$0^\circ$	$0^\circ, \pm 10^\circ, \pm 20^\circ$
残響時間	0.6sec	1.2sec	1.2sec

時間に近い値での試聴実験を行うため，以降では測定した室内伝達関数はサンプリング周波数が12kHzのFIRフィルタとして扱う．この操作によって部屋の残響時間は4倍になり，残響時間の値はそれぞれ  $150[\text{ms}] \times 4 = 0.6[\text{sec}]$ ，  $300[\text{ms}] \times 4 = 1.2[\text{sec}]$  となる．これらの値はそれぞれ会議室やコンサートホールにおける典型的な残響時間に相当する．この場合，音の伝達時間や伝達距離も4倍になるため，図4の部屋の大きさも等価的に4倍になり，マイクロホンアレーの半径は  $0.5[\text{m}] \times 4 = 2[\text{m}]$ ，中心から各スピーカまでの距離は  $1.5[\text{m}] \times 4 = 6[\text{m}]$  相当となる．

#### 3.2 マルチチャンネル信号の作成

無響室で収録したドライソースに3.1節で測定した室内伝達関数を畳み込むことによってチャンネル信号を作成した．ドライソースには女性話者の音声，木琴，木管五重奏の3種類を用いた．音声は無響室で収録し，木琴と木管五重奏の各楽器音はMIDI音源から作成した．ドライソースの長さは5秒で，サンプリング周波数は12kHzである．また，帯域を50-5000Hzに制限している．作成条件を表2に示す．音源から音が出ていない場合も考慮して，5個の音源以外の条件も設定した．その際，音源の方位角以外のスピーカには無音信号を割り当てた．

#### 3.3 逆伝達関数の算出

逆伝達関数の算出は残響時間ごとに行った．算出手順を以下に示す．まず，3.1節で測定した室内伝達関数をDFTで周波数領域に変換し，(6)式に従って  $5 \times 24$  次元の  $\mathbf{G}(\omega)$  を導出した．次に，(10)式によって  $24 \times 5$  次元の  $\mathbf{H}(\omega)$  を算出した．このとき， $\mathbf{D}(\omega)$  の対角成分にFIRバンドパスフィルタのDFT変換係数を用いることによって，室内伝達関数の振幅特性が小さい周波数領域においては逆伝達関数の算出を省略している．

表 3 逆伝達関数の算出条件  
Table 3 Calculation conditions of inverse transfer function

残響時間	0.6sec	1.2sec
DFT フレーム長	32768	65536
BPF の遅延	10ms, 20ms, 40ms, 80ms, 160ms, 320ms	
フィルタ長	14400	28800

これによって、より安定な逆伝達関数を算出することができる。本研究ではサンプリング周波数 12kHz で、帯域幅が 50–5000Hz の FIR バンドパスフィルタを適用した。最後に、 $H(\omega)$  を IDFT によって時間領域に戻し、打ち切ることによって逆伝達関数フィルタを算出した。算出条件を表 3 に示す。バンドパスフィルタの遅延が提案する空間符号化手法における遅延に相当する。

### 3.4 符号化及び復号化

3.2 節で作成した 24 チャネルのマルチチャネル信号に 3.3 節で算出した逆伝達関数を畳み込んで 5 チャネルの音源信号を抽出することによって、符号化を行った。これによって、伝送量は 24 チャネルから 5 チャネルに削減された。また、抽出した 5 個の音源信号に室内伝達関数を畳み込んで 24 チャネルのマルチチャネル信号を復元することによって、復号化を行った。

### 3.5 逆伝達関数による符号化への影響

この節では算出した逆伝達関数によって符号化システムにどのような影響が及ぼされるかを考察する。まず、符号化遅延を 320ms (=3840points) とした時の算出結果を図 5 に示す。左列が残響時間が 0.6sec、右列が残響時間が 1.2sec の時の結果である。(a) は測定した方位角  $0^\circ$  のスピーカから方位角  $\theta = 0^\circ$  のマイクロホンまでの室内伝達関数で入力チャネル信号に相当する。(b) は室内伝達関数から算出した逆伝達関数である。(c) は室内伝達関数に逆伝達関数を畳み込んだ波形である。この波形が 3.3 節での  $D(\omega)$  の対角成分に用いた FIR バンドパスフィルタの波形に近ければ、音源信号の抽出精度が良いということを示している。(d) は (c) の波形に更に室内伝達関数を畳み込んだもので、復元したチャネル信号に相当する。

室内伝達関数の初期遅延ポイント数と逆伝達関数の初期遅延ポイント数を足し合わせると、設定した符号化遅延ポイント数になる。このことから、逆伝達関数によって音源信号を抽出することは室内伝達関数におけるスピーカからマイクロホンまでの全ての経路(経路に対応する遅延ポイント数によって直接音, 初期反射音, 残響音に分類される [2]) に対応した遅延和アレー処理 [17] に従って音源信号を抽出することに相当

する。符号化遅延が 320ms の場合は残響音まで含めて音源信号を抽出しているため、(c) に示される波形は 3.3 節で用いた FIR バンドパスフィルタに近い波形となっている。また、(d) の波形も (a) の波形と全く同じ形になっている。このように、完全に音源を抽出するためには直接音や初期反射音に加えて残響音の遅延ポイントまで含まれる必要があるため、符号化遅延は部屋の残響時間を考慮したポイント数が必要となる。しかし、会議室のような双方向通信を実現しようとする場合には符号化遅延が長すぎると円滑な会話の妨げとなる恐れがある [18]。そのため、双方向通信を実現するには符号化遅延をなるべく短く設定することが必要となる。

そこで、符号化遅延を会話の支障としないと考えられる 40ms (=480points) とした時の算出結果を図 6 に示す。この場合には、直接音や初期反射音の遅延ポイント数は含まれるが、残響音の遅延ポイント数は含まれないため、(c) で示されるように抽出された音源信号は残響音が残留したままの波形になる。しかし、(d) に示される復元したチャネル信号は (a) の波形と比べると概形があまり変化していない。これは抽出された音源信号が残留した残響音に比べて十分にパワーが大きいため、復元の際に付加された残響成分が残留した残響音をマスクしてしまうためであると考えられる。これより、符号化遅延を音源信号の抽出が完全に行われなような比較的短い時間に設定しても、チャネル信号の聴感上の違いはほとんど出て来ないことが予想される。以降では双方向通信の応用も考慮して、符号化遅延を会話に支障のない時間に設定した条件のみで主観評価実験を行うことにする。

## 4 主観評価実験

提案した符号化手法が知覚的にどのような影響を及ぼすのかを主観評価実験によって検討した [15]。

### 4.1 実験環境

主観評価実験は本来なら無響室で行う必要があるが、実験設備の都合上、なるべく残響時間が短い部屋(約 80ms の低残響室)で行った。スピーカアレーと受聴者の位置を図 7 に示す。3.1 節で説明したように、マイクロホンアレーの半径は 2m になるので、スピーカアレーの半径も 2m にした。また、各スピーカからチャネル信号を再生すると、スピーカアレーの中に音場が再現され、図 7 の灰色で示されるように中心から 6m 離れた 5 個のスピーカの位置に音像が生じることになる。スピーカの周波数特性は 50–5000Hz でほぼ一定であったので、この研究では特にスピーカに関する周波数特性の補正は行わなかった。暗騒音レベルは 25.0dB(A)、音圧レベルは受聴者の位置において約 70dB(A) とし

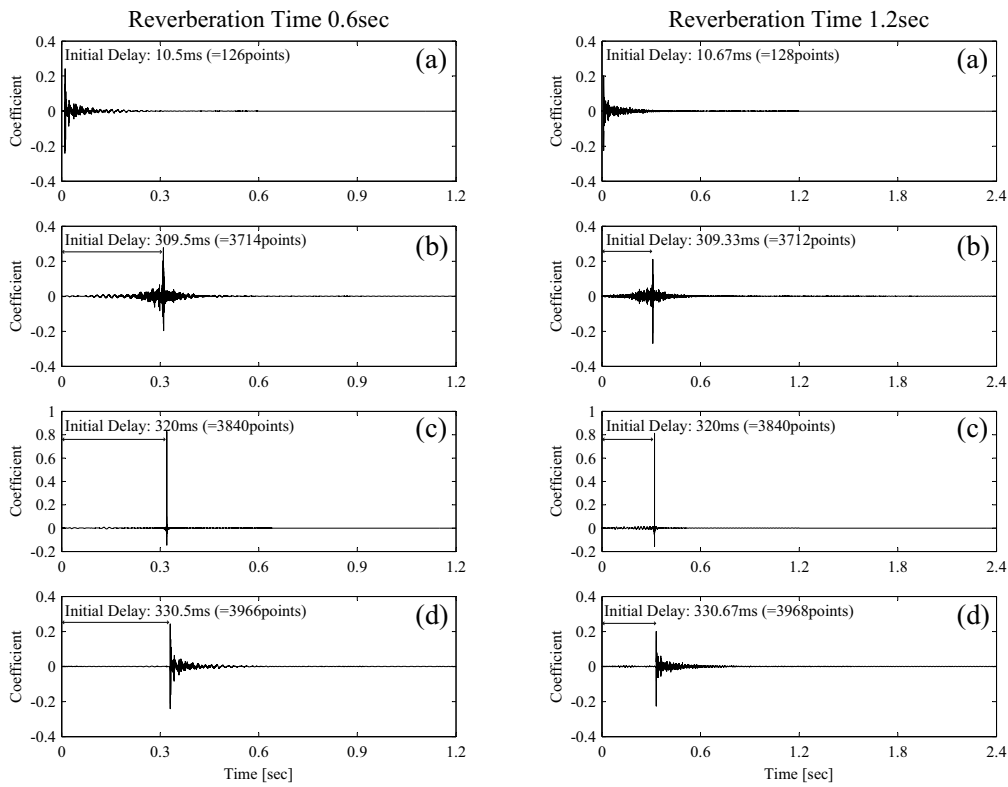


図5 符号化遅延が 320ms (=3840 points) の時の逆伝達関数の算出結果  
 Fig. 5 Computation results when the coding delay is 320ms (=3840 points)

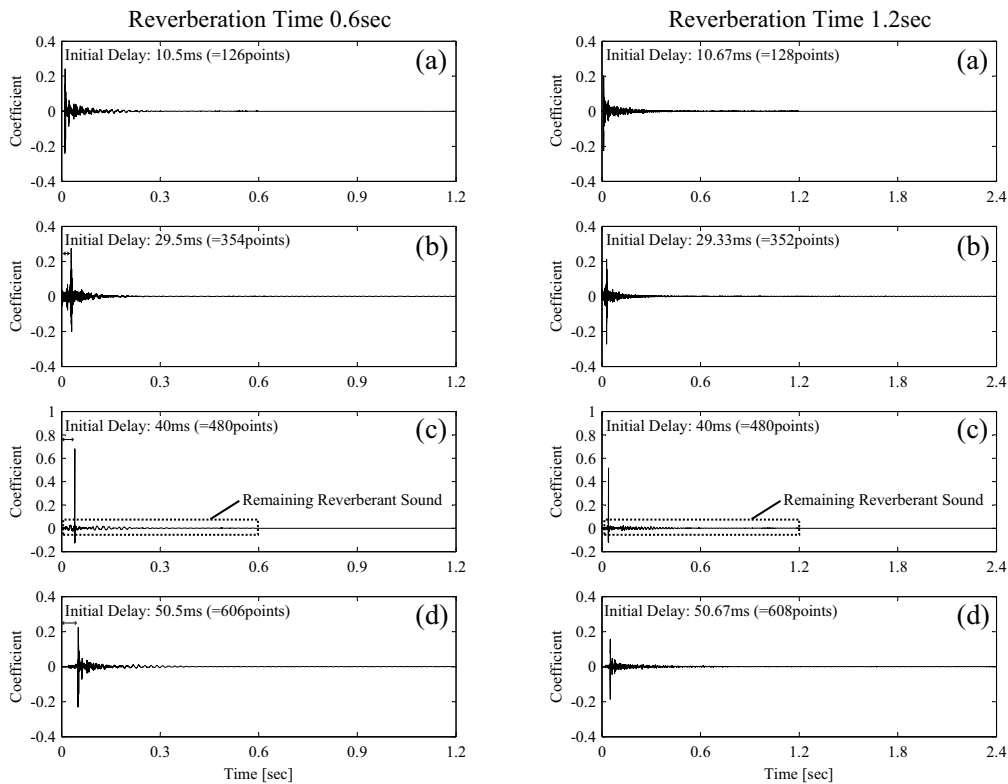


図6 符号化遅延が 40ms (=480 points) の時の逆伝達関数の算出結果  
 Fig. 6 Computation results when the coding delay is 40ms (=480 points)



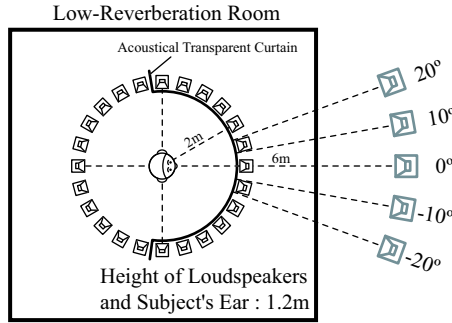


図 7 スピーカアレーと受聴者の位置  
Fig. 7 Position of loudspeaker array and subject

表 4 刺激の組み合わせ  
Table 4 Combinations of stimuli

X	A	B	試行数	
			練習試行	本試行
原音	原音	符号化音 (遅延=10ms)	1	2
原音	符号化音 (遅延=10ms)	原音	1	2
原音	原音	符号化音 (遅延=20ms)	1	2
原音	符号化音 (遅延=20ms)	原音	1	2
原音	原音	符号化音 (遅延=40ms)	1	2
原音	符号化音 (遅延=40ms)	原音	1	2

た. また, 視覚の影響を低減するため, 照明を薄暗くし, スピーカアレーの前方は音響カーテンで覆った.

#### 4.2 実験計画

被験者は日常生活において聴力的に異常の見られない 7 名の男子学生と 1 名の女子学生である. 実験は ITU-R 勧告 [19] に従って三刺激二重盲験法を用いた. 各試行において 3 種類の刺激 (以下 “X”, “A”, “B” とする) が提示される. “X” には常に符号化しない原音が割り当てられ, “A” と “B” には原音もしくは符号化した音がランダムに割り当てられる. 刺激の全ての組み合わせ及び試行数を表 4 に示す.

実験計画は以下のように設定した. まず, 符号化による影響の違いを検討するために, 実験全体を 2 つに分割し, それぞれを “音場評価”, “全体評価” と名付けた. 次に, それぞれの評価を 3 つのセッションに分割した. それぞれのセッションにおいて音源の異なる 3 種類のチャンネル信号 (“音声”, “木琴”, “木管五重奏”) を提示した. その際, 音源の提示順序は被験者ごとにランダム化した. 各セッションにおいて 6 回の練習試行と 12 回の本試行を行い, 1 回の試行において “X”, “A”, “B” の順に提示した. 試行の提示順序も被験者ごとにランダム化した. 実験計画の流れ図を

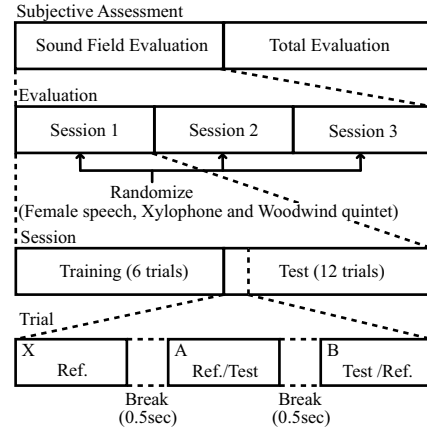


図 8 主観評価実験計画の流れ図  
Fig. 8 Flowchart of Design of subjective assessment

表 5 違いの評定表  
Table 5 Scale table of impairment

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 8 に示す.

#### 4.3 実験手順

前半の音場評価では符号化による音場への影響を評価した. 被験者に以下のように教示した. “X” は基準となる部屋で鳴っています. “A” と “B” のどちらかは必ず基準となる部屋と同じ部屋で鳴っています. “A” と “B” のうち, 基準となる部屋で鳴っていると思った方を “5.0” と評定して下さい. もう一つは表 5 に従って, 部屋の差がどれくらいあるかを小数点第 1 位単位で “1.0” から “4.9” の範囲で評定して下さい.

後半の全体評価では先に評価した音場への影響に加え, 符号化による音質への影響をも含んだ全体的な影響を検討した. 被験者に以下のように教示した. “X” は原音です. “A” と “B” のどちらかは必ず原音と同じです. “A” と “B” のうち, 原音と思った方を “5.0” と評定して下さい. もう一つは表 5 に従って, 原音との差がどれくらいかを小数点第 1 位単位で “1.0” から “4.9” までで評定して下さい.

音を聞いている間, 被験者は頭部を自由に動かすことができた. また, 判断がつくまでは何度も “X”, “A”, “B” を聞くことができた.

#### 4.4 実験結果及び考察

“A” と “B” のうち原音を割り当てた方に “5.0” と評定したかを確認することによって被験者の評定能力を検証する. 被験者の評定能力が十分でない場合, そ

表 6 二項検定の結果  
Table 6 Result of binomial test

被験者	性別	標本数	音場評価		全体評価	
			正答数	危険率	正答数	危険率
A	男	36	33	0.0000***	36	0.0000***
B	女	36	34	0.0000***	35	0.0000***
C	男	36	32	0.0000***	35	0.0000***
D	男	36	31	0.0000***	32	0.0000***
E	男	36	32	0.0000***	31	0.0000***
F	男	36	25	0.0057**	28	0.0002***
G	男	36	25	0.0057**	27	0.0006***
H	男	36	26	0.0020**	25	0.0057**

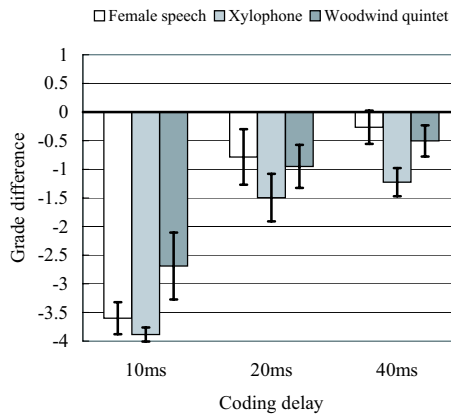


図 9 音場評価におけるランク差  
Fig.9 Grade difference in the sound field evaluation test.

の被験者は当て推量で回答しているため、正答率がチャンスレベルに落ち着くと考えられる。各被験者の全体の正答数を二項検定した結果を表 6 に示す。\*\* は 1%有意を表し、\*\*\* は 0.1%有意を表す。以降の分析では 0.1%有意と判定された被験者のデータのみを用いることとする。

各試行において符号化音を割り当てた音の評定値から原音を割り当てた音の評定値を引いたランク差を算出する。ランク差の平均値を図 9, 10 に示す。音場評価については 20 標本、全体評価については 28 標本のデータである。図中のエラーバーは 95%信頼区間を表す。もし、符号化による影響がない場合には、ランク差平均は 0 に近づくことになる。

音場評価の結果は全体評価の結果とほとんど同じである。このことは、劣化は符号化による音場への影響によるものがほとんどであるということを意味している。

符号化遅延が 10ms のときは被験者は少なくとも符号化による劣化を“わずかに気にしている”。これは設定した符号化遅延が室内伝達関数の初期遅延よりも小さいために、音源抽出の際に 3.5 節で示したような遅延とアレー処理ができなかったためだと考えられる。

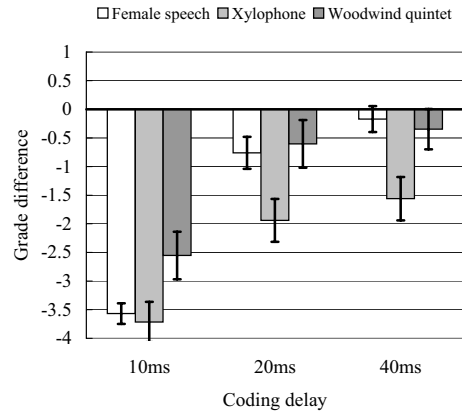


図 10 全体評価におけるランク差  
Fig.10 Grade difference in the total evaluation test.

一方、符号化遅延が 20ms や 40ms の場合には、音声や木管五重奏においてランク差が -1 より大きい。つまり、被験者は音声や木管五重奏のように持続している音に対しては符号化による劣化を“知覚できるが、気にしていない”ということが言える。これは 3.5 節で説明したように、直接音や初期反射音を含めた遅延とアレー処理を行い、抽出した音源信号が残留した残響音に比べて十分に大きくなるため、復元の際に付加された残響成分が抽出の際に残留してしまった残響音をマスクすることに因るものと考えられる。

しかし、木琴のような衝撃音に対しては多少評定値が下がっている。これは逆伝達関数を有限長にしたために起こるプリエコーのような時間的な歪み [20] が衝撃音に対しては鋭敏に感じられることが原因であり、より正確な逆伝達関数を算出する手法 [21, 22, 23] を適用すればこのような問題は解決されると考えられる。

## 5 まとめ

本論文では音場再現における空間符号化手法を検討した。そして、音源抽出に基づく空間符号化手法を提案した。室内伝達関数の逆伝達関数をチャンネル信号に畳み込むことによって、伝送量をチャンネル数から音源の数に削減することができた。実環境にて測定された室内伝達関数を基に符号化実験を行ったところ、24 チャンネルのマルチチャンネル信号を 5 チャンネルの音源信号に圧縮することができた。また、符号化による知覚への影響を検討するため、主観評価実験を行った。その結果、符号化による知覚への影響は多くの場合において問題にならない程度であることが示された。これにより、提案した空間符号化技術がより臨場感のある遠隔コミュニケーションを実現するための技術の一部として役に立つ可能性が示唆された。

今後の課題としては以下のことが挙げられる。まず、



本研究では測定された室内伝達関数が送信側と受信側で共有されていることを前提として検討してきたが, 室内伝達関数の測定が実用上では手間がかかるため, 未知の室内伝達関数を推定する実用的手法を検討することによって, より現実的な空間符号化手法を実装していく必要がある. 次に, 本研究では少数の音源が静止している場合を検討したが, 実際に想定される場面としては演劇のように音源が動いている場合やオーケストラのようにマイクロホンよりも音源の数の方が多く場合が考えられる. ゆえに, 今後はそれらの場合も検討していく必要がある.

### 謝辞

本研究は文部科学省科学研究費補助金 COE 形成基礎研究費 (課題番号 11CE2005), 及び基礎研究 B(課題番号 11480075) の補助を受けて行われた.

### 参考文献

[1] M. Camras : Approach to recreating a sound field , J. Acoust. Soc. Am., Vol.43, No.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., Vol.93, No.5, pp.2764-2778 (1993)

[3] M. A. Poletti : A Unified Theory of Horizontal Holographic Sound Systems , J. Audio Eng. Soc., Vol.48, No.12, pp.1155-1182 (2000)

[4] 平井徹、本地由和、池田雅弘、清水寧、川上福司、東山三樹夫 : 遠隔地音楽レッスン用 AVW システムとその原理, 信学技報, EA2000-17, pp.25-32 (2000)

[5] 浜田晴夫 : バイノーラル音場再生系について, 音響学会誌, Vol.48, No.4, pp.250-257 (1992)

[6] 高根昭一、鈴木陽一、高橋秀介、宮島徹 : 仮想球モデル聴覚ディスプレイ (ADVISE) の基礎的実時間システムの開発, 信学技報, EA2002-50, pp.25-30 (2002)

[7] 安藤四一、設楽貞樹、前川純一、城戸健一 : 電子計算機による室内音響設計法に関する二、三の基礎的考察, 音響学会誌, Vol.29, No.3, pp.151-159 (1973)

[8] J. Bauck and D. H. Cooper : Generalized transaural stereo and applications , J. Audio Eng. Soc., Vol.44, No.9, pp.683-705 (1996)

[9] S. Vernon : Design and implementation of AC-3 coders , IEEE Trans. CE, Vol.41, No.3, pp.754-759 (1995)

[10] ISO/IEC 13818-7 : Information technology - Genetic coding of moving pictures and associated audio information - Part 7 Advanced Audio Coding

[11] D. Yang, H. Ai, C. Kyriakakis and C. C. J. Kuo : An inter-channel redundancy removal approach for high-quality multichannel audio compression , AES 109th Convention, 5238 (2000)

[12] Y. Wang, M. Vilermo, M. Väänänen and L. Yaroslavsky : A multichannel audio coding algorithm for inter-channel redundancy removal , AES 110th Convention, 5295 (2001)

[13] 木村敏幸、寛一彦、武田一哉、板倉文忠 : 空間情報を利用したマルチチャンネル音響信号の符号化と音源定位, 音講論集, pp.711-712 (2001.10)

[14] 木村敏幸、寛一彦、武田一哉、板倉文忠 : 逆フィルタを用いた空間音響圧縮技術の拡がり感, 音講論集,

pp.595-596 (2002.3)

[15] 木村敏幸、寛一彦、武田一哉、板倉文忠 : 逆フィルタを用いた音場空間符号化方式の主観評価, 音講論集, pp.563-564 (2002.9)

[16] Y. Suzuki, F. Asano, H. Y. Kim and T. Sone : An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses , J. Acoust. Soc. Am., Vol.97, No.2, pp.1119-1123 (1995)

[17] 北脇信彦編著 : デジタル音声・オーディオ技術, 電気通信協会, pp.172-174 (1999)

[18] 伊藤憲三、北脇信彦 : 双方向通信系における伝搬遅延の許容限に関する検討, 聴覚研究会資料, H-84-57 (1984)

[19] ITU-R Recommendation BS.1116-1 : Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems

[20] 中島弘史、東山三樹夫 : 周波数領域処理を繰り返して利用する逆フィルタの時間領域設計法, 音講論集, pp.431-432 (1999.9)

[21] M. Miyoshi and Y. Kaneda : Inverse Filtering of Room Acoustics , IEEE Trans. ASSP, Vol.36, No.2, pp.145-152 (1988)

[22] S. J. Elliott, I. M. Stothers and P. A. Nelson : A Multiple Error LMS Algorithm and Its Application to the Active Control of Sound and Vibration , IEEE Trans. ASSP, Vol.35, No.10, pp.1423-1434 (1987)

[23] Y. Tatekuta, H. Saruwatari and K. Shikano : An Iterative Inverse Filter Design Method for the Multichannel Sound Field Reproduction System , IEICE Trans. Fundamentals, Vol.E84-A, No.4, pp.991-998 (2001)

(2003年6月23日受付)

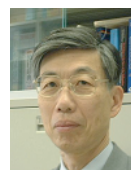
### [著者紹介]

#### 木村 敏幸 (正会員)



1998年名大・工・物化卒。2003年同大学院・人間情報・博士課程了。同年より、日本学術振興会特別研究員(特別推進研究COE)。音場再現に関する研究に従事。日本音響学会, 電子情報通信学会, 日本バーチャルリアリティ学会各会員。

#### 寛 一彦



1965年早大・理工・電気卒。1967年同大学院修士課程了。同年日本電信電話公社(現NTT)入社。同社研究所において電話機回路・音響機器の研究実用化, 通話品質, 聴知覚の基礎研究に従事。1994年名大・人間情報・教授。2003年名大・情報科学・教授。工博。1981年度日本ITU賞, 1984年度電子情報通信学会論文賞, 1985年度科学技術庁長官賞受賞。日本音響学会, 電子情報通信学会, 日本認知科学会, 日本音聲学会, 日本高次脳機能障害学会, 日本語学学会等会員。

武田 一哉



1983年名大・工・電気卒。1985年同大学院博士(前期)課程了。同年国際電信電話株式会社入社。ATR自動翻訳電話研究所, KDD研究所において音声合成・認識システムの研究を行う。1994年名大・工・助教授。2003年名大・情報科学・教授。工博。日本音響学会, IEEE, 電子情報通信学会, 情報処理学会各会員。

板倉 文忠



1963年名大・工・電気卒。1968年同大学院博士課程満了。同年電電公社(現NTT)武蔵野通研入所。音声処理の研究に従事。1973-1975年ベル研究所にて音声認識・音声分析の研究を行う。1984年名大・工・教授。1998年名大・情報メディア教育センター教授。工博。1970年, 1978年, 1981年度電子情報通信学会論文賞, 1972年, 1981年度同業績賞受賞。1996年IEEE Signal Processing Society Award受賞。2000年IEEE Third Millennium Medals受賞。電子情報通信学会フェロー。2003年紫綬褒章受章。日本音響学会, 電子情報通信学会, IEEE各会員。