

逆フィルタを用いた 音場空間符号化方式の主観評価*

木村敏幸, 箕一彦(名大・人間情報/CIAIR), 武田一哉, 板倉文忠(名大・工/CIAIR)

1.はじめに

遠隔通信会議や遠隔コンサートホールなどのよりリアルなコミュニケーションを実現するために、スピーカアレーを用いた音場再現技術が近年盛んに研究されている。しかし、この方式は非常に多くのチャンネルを要するので、伝送量がチャンネル数に比例する結果となる。このためチャンネル当たりの伝送量を削減するには MPEG2 AAC[1] のような方式では不十分である。

我々は今までに逆フィルタを用いた音場空間符号化方式を提案してきた[2, 3]。本報告では主観評価実験によって、提案した手法の有効性を検討した。

2.符号化実験

2.1.室内インパルス応答の測定

室内インパルス応答の測定は名古屋大学工学部 7号館 410 室の可変残響室内で行った。スピーカとマイクロホンアレーの配置図を Fig.1 に示す。

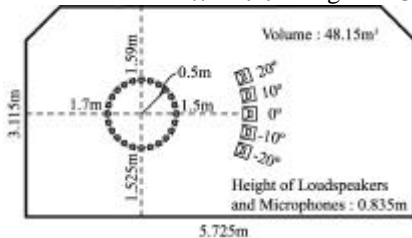


Fig.1: Position of loudspeakers and microphone array

測定条件を Table 1 に示す。音圧レベルはスピーカから正面軸上に 1m 離れた地点での値である。

Table 1: Measurement conditions of room impulse response

Reverberation time	150ms	300ms
Room temperature	19.5°C	19.2°C
Noise level	20.0dB(A)	19.4dB(A)
Sound pressure level	90.0dB(A)	91.6dB(A)
Sampling frequency	48kHz	
Reference signal	TSP (65536 points)	
Repetitions	16	
FIR filter order	7200	14400

2.2.マルチチャンネル信号の作成

測定した室内インパルス応答を音源信号に畳み込むことによってマルチチャンネル信号を作成した。それぞれのマルチチャンネル信号の作成条件を Table 2 に示す。音源信号のサンプリング周波数が室内インパルス応答測定時の 4 分の 1 であるので、残響時間は測定時の 4 倍になる。ゆえに、マルチチャンネル信号の残響時間は測定時の 4 倍になり、値は 0.6sec と 1.2sec になる。また、音源が 1 個の場合には他の方位角からの再生信号には振幅

が 0 の信号を割り当てた。

Table 2: Synthesis conditions of multi-channel signals

Source signals	Speech	Xylophone	Woodwind Quintet
Durations	5sec		
Sampling frequency	12kHz		
Bandwidth	50Hz-5kHz		
Number of sources	1	5	
Azimuth angle	0°	0°, ±10°, ±20°	
Reverberation time	0.6sec	1.2sec	

2.3.逆フィルタの算出

符号化に用いる逆フィルタは以下の手順によって算出した[3, 4]。まず、測定した室内インパルス応答を DFT によって周波数領域に変換し、室内伝達関数行列を算出した。本研究では 5×24 次元の行列とした。次に、Moore-Penrose 擬似逆行列を用いて 24×5 次元の逆伝達関数行列を算出した。このときの単位行列の対角成分には FIR バンドパスフィルタを DFT 変換したものをを用いた。バンドパスフィルタの特性はサンプリング周波数 12kHz、帯域 50Hz-5kHz である。最後に、逆伝達関数行列を IDFT によって時間領域に戻し、先頭部分で打ち切って逆フィルタを算出した。逆フィルタの算出条件を Table.3 に示す。バンドパスフィルタの遅延は因果律を満たすために必要なものである。このバンドパスフィルタの遅延が符号化遅延に相当する。

Table 3: Calculation conditions of inverse filter

Reverberation time	0.6sec	1.2sec
FFT points	32768	65536
Delay of BPF	10ms, 20ms, 40ms	
Truncated FIR points	14400	28800

3.主観評価実験

主観評価実験の方法は ITU-R BS.1116-1 の勧告 [5]を参考にした。

3.1.実験条件

低残響室(残響時間約 80ms)において Fig.2 のようにスピーカアレーを配置し、円の中心に被験者を配置した。灰色のスピーカは被験者が感じる音像である。被験者には 5 名の学生を用いた。

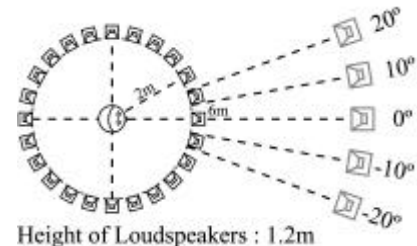


Fig.2: Position of loudspeaker array and sound images

*Subjective assessment of spatial audio coding using inverse filters, By T. Kimura, K. Kakehi, K. Takeda and F. Itakura (Nagoya University)

暗騒音レベルは 25.0dB(A)で、音圧レベルは受聴位置において約 70dB(A)に設定した。実験の構成図を Fig.3 に示す。セッションと試行の提示順序は被験者ごとにランダムイズした。各セッションにおける試行の組み合わせの一覧を Table 4 に示す。

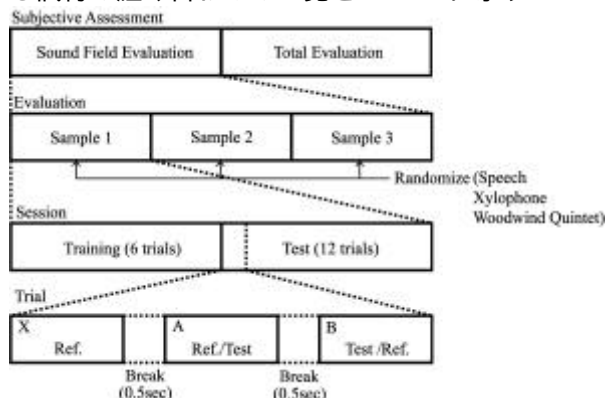


Fig.3: Organization of subjective assessment

Table 4: Combinations of trials
(O: Original sound, C: Coding sound)

X	A	B	Trials
O	O	C (Delay = 10ms)	2
O	C (Delay = 10ms)	O	2
O	O	C (Delay = 20ms)	2
O	C (Delay = 20ms)	O	2
O	O	C (Delay = 40ms)	2
O	C (Delay = 40ms)	O	2

3.2. 評定手順

評定に用いた指標を Table 5 に示す。

Table 5: Grade of impairment

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

3.2.1. 音場評価

最初に音場の評価を行った。「X は基準となる部屋で鳴っています。A と B のどちらかは基準となる部屋で鳴っています。A と B のうち、X と同じ部屋で鳴っていると思った方を「5.0」と評定して下さい。もう一方は部屋の差の程度を Table 5 に従って「1.0」から「4.9」までで評定して下さい。」と教示した。

3.2.2. 全体評価

次に全体の評価を行った。「X は原音です。A と B のどちらかは原音です。A と B のうち、原音だと思った方を「5.0」と評定して下さい。もう一方は原音との差の程度を Table 5 に従って「1.0」から「4.9」までで評定して下さい。」と教示した。

3.3. 実験結果

符号化音の評定点から原音の評定点を引いた値を音源、符号化遅延の条件ごとに平均したグラフを Fig.4, 5 に示す。バーは 95%信頼区間を示す。木琴のような立ち上がりの早い衝撃音に対してはすべての場合において音場や全体的な差を感じてい

る。しかし、音声や木管五重奏のような持続音においては、符号化遅延が 20ms と 40ms のときは信頼区間の下限が -1 以上である。つまり、この条件においては符号化による音場や全体的な音質の差を感じてはいるが、その差を気にしてはいないということが分かる。このことから実用上想定される音源においては符号化遅延をある程度許容すれば提案した符号化方式による知覚への影響は見られないと言える。

Sound Field Evaluation

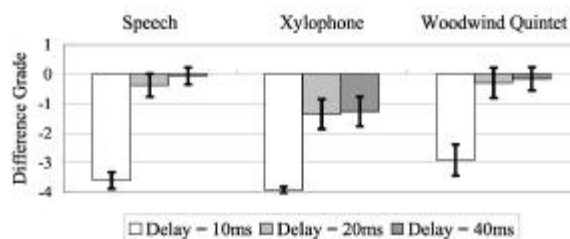


Fig.4: Difference grade in sound field evaluation

Total Evaluation

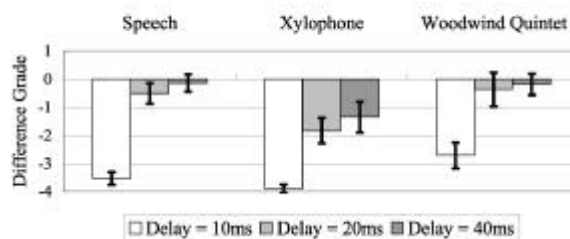


Fig.5: Difference grade in total evaluation

4. まとめ

提案した音場空間符号化方式の知覚への影響を主観評価実験によって検討した。その結果、実用上想定される音源に対してはある程度の符号化遅延を許容すれば影響が見られないことが分かった。

今後は、音楽家や専門知識を持った人たちを被験者とした主観評価実験や、移動音源の場合や残響時間が更に長い場合のようなより実環境に近い条件での影響を検討する必要がある。

5. 参考文献

- [1] ISO/IEC 13818-7, "Information technology - Generic coding of moving pictures and associated audio information - Part 7 Advanced Audio Coding"
- [2] 木村, 算, 武田, 板倉, "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, 711-712 (2001.10)
- [3] 木村, 算, 武田, 板倉, "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, 595-596 (2002.3)
- [4] J. Bauck and D. H. Cooper, "Generalized transaural stereo and applications," J. Audio Eng. Soc., **44**, no.9, pp.683-705, 1996
- [5] ITU-R Recommendation BS.1116-1 "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems"