JAIST Repository
https://dspace.jaist.ac.jp/

Title	スピーカの配置に基づいたサラウンド音場補正手法の提案と評価
Author(s)	秋山，大知；金井，秀明
Citation	情報処理学会研究報告，2013-GN-87(5)：1-8
Issue Date	2013-03-11
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/12216
Rights	社団法人　情報処理学会，秋山大知，金井秀明，情報処理学会研究報告，2013-GN-87(5)，2013，1-8.　ここに掲載した著作物の利用に関する注意：本著作物の著作権は（社）情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 Notice for the use of this material: The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.
Description	

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

Vol.2013-GN-87 No.5
2013/3/18

スピーカの配置に基づいたサラウンド音場補正手法の提案と評価

秋山 大知[1]　金井 秀明[†1]

概要：本稿では，サラウンドによる 3 次元音場の提示について，スピーカの配置に依存しない提示法を提案する．具体的には，標準的なサラウンドのスピーカ配置から外れた配置がなされた際に，標準的なスピーカ配置での提示とほぼ同じ音場を体験出来るようなサラウンド音場補正手法を開発する．

Research of a surround-sound field adjustmen system based on loudspeakers arrangement

Akiyama Daichi[1]　Kanai Hideaki[†1]

Abstract: In this paper, we propose a presentation method that does not depend on the placement of the loudspeakers for the presentation of three-dimensional sound field by the surround. Specifically, we propose a method to correct the surround sound field adjustment signal by the signal processing to reproduce surround-loudspeakers arrangement the same as a standard if placement of the loudspeakers is changed.

1. はじめに

近年，映画の音声や音楽作品，イベントやパビリオンの音響効果として，サラウンドによる音の提示が多く行われている．サラウンドとは，リスナーの周囲に複数のスピーカを配置することにより，360° 全ての方向から音の提示を行う提示手法である．サラウンドの音を聴いたリスナーが，ある 1 点に音が位置していると知覚することを「音像が定位する」という．サラウンドによって作られた作品には，大きく分けて 2 種類のものが存在する．コンサートホールなどの音場をリアルに再現するためにサラウンドを活用する「現場音再生型」と，作品の表現の幅を広げるために新しく音場を作り出す「音場創造型」の 2 つである [1]．本研究で対象とするのは，音場創造型のサラウンドである．音場創造型のサラウンドを効果的に使った作品の例として，冨田勲の「交響詩ジャングル大帝[*1]」がある．この作品は，オーケストラの楽器を手塚治虫の漫画「ジャングル大帝」の登場キャラクターに見立て，楽器の音が前後左右に動き回ることで，ジャングルのあちこちから動物が集まってくる様子や，ジャングルの動物がハンターから逃げている様子を表現している．リスナーは，さまざまな方向から聴こえる音や動き回る音を聴いて，その作品の世界観や面白さを「サラウンド音場」として体験することが出来る．

音場創造型の作品は，その作品の制作者の意図が音場に現れることになる．従って，制作者の意図をリスナーが正しく体験するためには，正しいサラウンド環境で音を聴かなくてはならない．正しいサラウンド環境とは，制作者がサラウンドのミキシングをする際に実際にスタジオでなされていたサラウンドセッティングのことである．サラウンドには数種類のフォーマットが定められており，特殊な場合でない限り，そのフォーマット通りのセッティングを行えば正しい（制作者が作品を制作した時に意図した音場と同じ）音場が再現出来る．ITU-R BS 775-01 という規格では，5 チャンネルから 8 チャンネルまでのスピーカ配置に

[1] 北陸先端科学技術大学院大学　知識科学研究科
School of Knowledge Science, Japan Advanced Institute of Science and Technology
[†1] 現在，北陸先端科学技術大学院大学　ライフスタイルデザイン研究センター
Presently with Research Center for Innovative Lifestyle Design, Japan Advanced Institute of Science and Technology

[*1] 日本コロムビア　B002L9QG0Q

ⓒ 2013 Information Processing Society of Japan

ついて検討がなされ，規格として採用されている [2]．図 1 は，家庭の再生機器でも多く用いられる 5.1 ch サラウンドの配置である．5.1 ch サラウンドの場合，この図の通りにスピーカを配置し，リスナーがこの図の位置にいれば，正しい音場が体験出来る*2．

一方，スピーカが規格通りに配置されない場合，リスナーに正しい音場を伝えることは出来ない．実際にサラウンドを導入する場合，スピーカを配置しようとした場所に物や柱があるなどの理由から, 規格通りの配置を諦めざるを得ない場合がある. 我々は，スピーカが規格通りに配置されない場合，リスナーが知覚する音像定位方向がどのように変化するかを，被験者にサラウンド音声を聴かせることにより調べた．22 才〜49 才の男性 18 名, 女性 2 名に対して，サラウンドシステムで 0°，36°，72°，108°，144°，180°，216°，252°，288°，324° の 10 方向からランダムにピンクノイズを提示した．これを被験者 1 人に対して 100 回行い，音を聴いた被験者がどの方向に音像を知覚したかを回答し，その正答率を調べた．その結果，標準的な 5.0 ch サラウンド*3の配置での正答率は 0.5403，ランダムに配置がずれた，非標準的な 5.0 ch サラウンド配置では 0.3948 となり，スピーカ配置がずれた場合は，音像定位の正確さが損なわれることが分かった．

この結果を踏まえ，我々はサラウンドによる再生システムをより便利で導入しやすくするために，スピーカ配置に応じて再生されるサラウンド音声が自動的に補正されるシステムの提案を行う．スピーカ配置が規格通りの配置からずれていても，規格通りの配置で再生した場合の音像定位方向の知覚とほぼ同程度の正確さで音像定位の知覚ができるシステムを目指す．

以下，第 2 章で関連研究，第 3 章で本研究で提案する補正手法の概要，第 4 章で提案手法の評価実験の結果と考察について述べる．

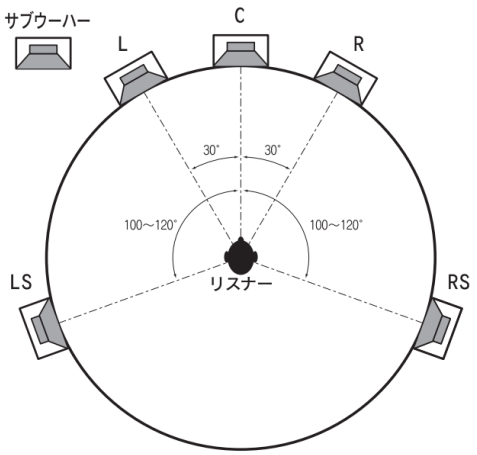

図 1 ITU-R BS 775-01 による 5.1 ch サラウンドのスピーカ配置

*2 サブウーハーは，人間は低音域の音の方向知覚は鈍感であるという理由から，どの位置に置いても良いとされている．
*3 サブウーハーなしのサラウンド

2. 関連研究

ヒトに空間音響を提示するシステムの研究は以前から行われている．ヒトの音像方向知覚は個人に特有の「頭部伝達関数」を学習することで行われることが知られている [3]．頭部伝達関数の情報を利用してヘッドホンで音像を提示するシステムが研究されている．Wightman と Kistler は，被験者本人の頭部伝達関数を計測し，頭部伝達関数の特性をフィルタリングした音をヘッドホンで提示することによる音像定位実験を行った，その結果，実音源と同程度の精度で音源方向を知覚できることを示した [4][5]．また，Wenzel らは同じシステムで，他人の頭部伝達関数を用いた場合，前後・上下誤判定が顕著に増加することを示した [6]．これらの結果から，音像提示に頭部伝達関数を利用する場合は，本人の頭部伝達関数を利用しなければならないことがわかる．つまり，同時に一人に対してしか，音場提示が行えない．

サラウンドのような多チャンネルスピーカを利用した提示は頭部伝達関数を利用しないため，一度に複数のリスナーに対して提示を行うことが可能である．こうしたサラウンドの利点を活かした研究に，音によるアウェアネスの提示手法の提案・検討がある．これは，方向情報を伴った音を提示することでヒトに気付きを与える研究である．山崎らは，公共空間における避難誘導のアナウンスをサラウンドに近い方式で音像定位させることにより，誘導方向をより明確にする避難誘導システムとして提案した [7]．その結果，従来のサラウンドパンニングに加えて，ヒトの音の方向知覚に深く関与している帯域を強調した拡声を行えば，アナウンスで避難誘導を行うのに最適な音場を公共空間で提示できる可能性を示した．また，梅津らは，博物館や舞台，旅行ツアーなどの分野で利用される音声案内システムにおける，音の方向提示の効果を検討した [8]．方向情報付きの音声案内を提示することにより，鑑賞者の興味の対象を特定の方向へ向けることが可能になる．その結果，鑑賞者自身が鑑賞対象に応じて番号を入力するなど，従来の音声案内で必要とされていた機器の操作の代わりに音声案内を活用することができる．このような，公共空間におけるアウェアネスを聴覚情報で与える場合には，従来のサラウンドの提示手法が適していると言える．サラウンドで再生することを想定した芸術作品の鑑賞だけでなく，こうしたアウェアネス提示法としてもサラウンドによる音場提示手法が利用されつつある．

サラウンドによる音場提示の精度は，以下の 4 点で決まる [2]．

(1) リスナーの位置によって，リスナーが知覚する音場が変わる．
(2) スピーカの数によって，再生される音場の精度が変わ

る．（スピーカ数が多い程高精度で音場で維持が可能）

(3) スピーカの配置によって，再生される音場の精度が変わる．（配置が規格通りの配置でないと正しく音像が定位しない）

(4) スピーカ自体の特性[*4]によって再生される音場の精度が変わる．

音場提示システムとしてサラウンドを導入する際は，これらのサラウンド環境を正しく設定しなければ制作者の意図した音場が再生されない．1. でも述べたように，サラウンドには，ITU-R BS 775-01 の規格が決められており，この規格通りにスピーカが設置されることが前提となっている．これらのサラウンド環境を正しく設定できない場合，実際に設定されている環境でもなるべく正しい音像が得られるように，信号処理技術を使って補正を行う必要がある．James は 5.0 ch サラウンドのスピーカ配置を数パターン想定し，配置変化に応じてサラウンドパンニングのアルゴリズムをそれに応じて変えなければ正しい音像が知覚できないことを実験により示した [9]．一方，厳密に高精度の再生は不可能だが，これらのサラウンド環境が正しくなくても極力正しい音場提示ができるように補正する研究がある．James は同研究で，想定したスピーカ配置のパターンに応じて，最適なパンニングの方法を検討した．これにより，サラウンド作品を制作する際に，実際に再生される配置を想定してパンニングを行うことができる．しかし，制作時に想定していない配置がなされる場合は，やはり正しい音場提示をすることはできない．従って，サラウンド環境を補正するためには，作品を制作する段階ではなく，再生する段階で補正を行わなければならない．

リスナーの位置による違いを補正する研究は，Ki-Seung らの研究がある．Ki-Seung らは，リスナーの位置をトラッキングし，リスナーの位置に応じて左右のスピーカのクロストーク・キャンセラ[*5]のレベルを変化させることで，リスナーがどの位置にいても最適なステレオ音場が提示できるトランスオーラル系[*6]の再生システムを作った [10]．

スピーカ数を補正する研究としては，ステレオからサラウンドフォーマットに変換する研究や，サラウンドフォーマットのチャンネル数が変化したフォーマットに変換する研究がある．Eugeniusz はステレオの音声を周波数毎に解析し，その音がどの位置で鳴っているのか，その音が直接音[*7]なのか反射音[*8]なのかを周波数毎に推定し，ステレオから 5.1 ch サラウンドに変換するアルゴリズムを提案している [11]．また，SONY はサラウンドスピーカに出力するレベルを均等にするよう調整することにより，5.1 ch サラウンドから 7.1 ch サラウンドへ変換可能なサラウンド AV アンプ「STR-VZ555ES」を発表している [12]．

スピーカの配置による違いを補正する技術はまだ少ないものの，SONY がスピーカの配置を補正する機能を持ったサラウンド AV アンプ「TA-DA5600ES」を発表している [13]．この機能は，スピーカから出力されるテスト信号をマイクロホンで計測し，各スピーカの出力が最適なレベルになるように調整することでスピーカ配置の補正を実現している．しかし，実際に計測が必要なため，雑音や室内にいる人物等の影響を受け易く，厳密な補正を行うためには測定環境を整えなければならない．本研究では，サラウンド再生における，スピーカ配置に基づいた音場補正をスピーカの位置情報だけを用いて行う．スピーカの位置情報だけを用いることで，スピーカの位置情報が既知である場合には環境の変化に左右されない確実な補正を行うことが可能である．

3. 提案手法

3.1 スピーカ配置のずれ

スピーカ配置に基づいた音場補正手法を提案する前に，本研究で扱う「スピーカ配置のずれ」を定義する．本研究では，各スピーカの位置をリスナーの位置を原点とした極座標系で考える．スピーカの位置はリスナーからスピーカまでの距離 $\boldsymbol{R}$ と，リスナーから見たいずれかの方向[*9]を 0 とした際のスピーカの角度 $\boldsymbol{\Theta}$ で表せることになる．$\boldsymbol{R}$ と $\boldsymbol{\Theta}$ は，それぞれのスピーカの位置情報を要素とする以下のベクトルで表す．

$$\boldsymbol{R} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{pmatrix} \qquad \boldsymbol{\Theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_N \end{pmatrix} \tag{1}$$

ここで，N は，サラウンド再生に用いているスピーカ数である．標準的な 5.0 ch の場合，r_1 と θ_1 は，図 2 のように定義される．この場合，全てのスピーカは被験者の位置を中心とした円上にあるので，r_1〜r_N の値は全て等しい．θ_1〜θ_N の値は，採用するサラウンドの方式によって決まった値が入る．

スピーカ位置の変化は，$\boldsymbol{R}$ の変化と $\boldsymbol{\Theta}$ の変化に分けて考える．$\boldsymbol{R}$ だけが変化した場合，リスナーから見た各スピーカの方向は変わらないが，リスナーから各スピーカへの距離は変化する．一方，$\boldsymbol{\Theta}$ だけが変化した場合は，リスナーから各スピーカへの距離は全て等しいが，リスナーから見た各スピーカの方向はそれぞれ変化する．$\boldsymbol{R}$ と $\boldsymbol{\Theta}$ 両方が変化すると，スピーカの配置は完全にランダムに変化する．

本研究では，問題を単純化するために，スピーカの位置の変化は $\boldsymbol{\Theta}$ の変化のみとする．したがって，スピーカが

*4 周波数特性や指向性，雑多なスピーカを利用した場合の個体差等
*5 左（右）耳に右（左）チャンネルの音が入ることを防ぐ機能
*6 クロストーク・キャンセラを持ったステレオ再生装置
*7 前から聴こえると考えられる音
*8 後ろから聴こえると考えられる音
*9 相対的な位置関係が重要なので，どの方向を 0 とするかはあまり重要ではない

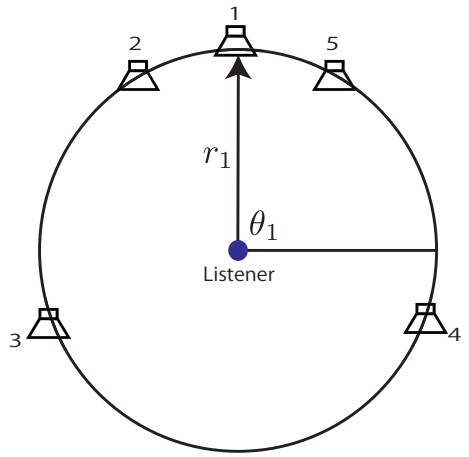

図 2 標準的な 5.0 ch 配置での r_1 と θ_1

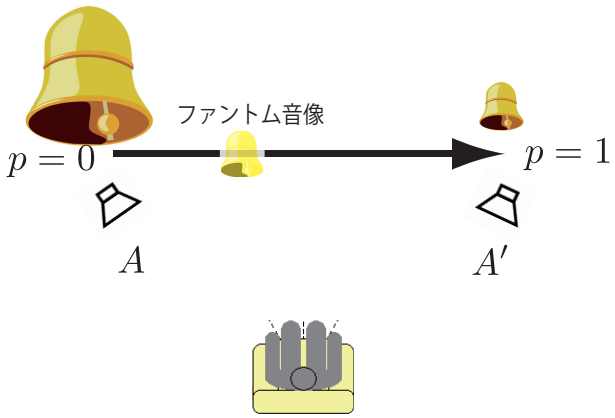

図 3 ステレオ再生の場合のファントム音像の出現

リスナーを中心とした円周上で位置が変化することを，スピーカの位置の変化として定義する．

3.2 要素技術：2 つのスピーカによる音像定位

サラウンドやステレオ再生のような音場提示手法は，2 つのスピーカから同じ音が音量差を伴って再生されることにより，リスナーの聴感上ではそのスピーカの間に音像が生じるという現象が理論的基盤になっている（図 3）．この時，リスナーの聴感上で生じている音像を「ファントム音像」と呼ぶ [2]．2 つのスピーカからの出力音に音量差がない場合，ファントム音像は，2 つのスピーカの中間地点に出現する．右（左）チャンネルのスピーカの出力音が左（右）チャンネルの出力音に比べて大きくなるにつれて，ファントム音像は右（左）側に移動する．ステレオの場合，音像は左右の移動のみだが，これを前後，後ろ側の左右で行うのがサラウンドによる音提示である．本研究では，この考え方を音場補正に利用する．ランダムに位置が変化して配置されたスピーカを用いて，実際はそこには配置されていない標準的な配置のスピーカからの出力音をファントム音像で提示すれば，標準的な配置で再生した音場と同程度の音場が提示できると考えられる．

2 つのスピーカによってファントム音像を出現させる場合，2 つのスピーカの音量比をどのように決定するかは様々な方法が考えられる．3.1 で述べたように，本研究では，リスナーからスピーカへの距離は常に一定である．このような，常に一定の距離にファントム音像が出現する音量比の決定法には，Reveillon の「等距離パンニング」がある [14]．2 つあるスピーカのうちの 1 つのスピーカからの音量を A，反対のスピーカからの音量を A'，どれだけ A' 側に音像を定位させるかを表す変数を p（図 3 参照）とすると，等距離パンニングでは A と A' は次のよう決定される．

$$A = \cos\left(2\pi\frac{p}{4}\right) \tag{2}$$

$$A' = \sin\left(2\pi\frac{p}{4}\right) \tag{3}$$

等距離パンニングのようなパンニング手法を使うことにより，ファントム音像の位置によって，音の大きさの感じ方が変化することを防ぐことができる．

3.3 音場補正手法

3.2 で述べた，ファントム音像を利用したサラウンド音場補正手法の詳細を述べる．先に述べたように，リスナーから各スピーカへの距離は全て等しいとしたため，等距離パンニングによってファントム音像を提示すれば，標準的な位置に配置されたスピーカから音が聴こえている状態を再現できることになる．すなわち，スピーカの位置（本研究では，リスナーを中心とした円周上の角度のこと）がずれた場合，ずれの角度に応じて，そのずれの方向と反対方向のスピーカから音を出力すれば，元々スピーカがあった位置に，ファントム音像が出現する（図 4）．

ファントム音像による補正手法の詳細を述べる．3.1 で定義した標準的な配置でのスピーカの角度 $\boldsymbol{\Theta}$ に対して，実際に配置されているスピーカの角度を $\boldsymbol{\Theta}'$ と定義すると，実際に配置されているスピーカの角度がどれだけずれているかを以下のベクトルで表す．

$$\boldsymbol{D} = \boldsymbol{\Theta} - \boldsymbol{\Theta}' = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{pmatrix} \tag{4}$$

補正前の各スピーカへの入力信号 $\boldsymbol{X}$，補正後の各スピーカからの出力信号 $\boldsymbol{Y}$ を以下のベクトルで表す．

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \qquad \boldsymbol{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} \tag{5}$$

補正手法は，2 つのスピーカによってファントム音像をつくることにより行われる．ファントム音像は，2 つのスピーカから音量差のある同じ信号を出力することによって行わ

ⓒ 2013 Information Processing Society of Japan

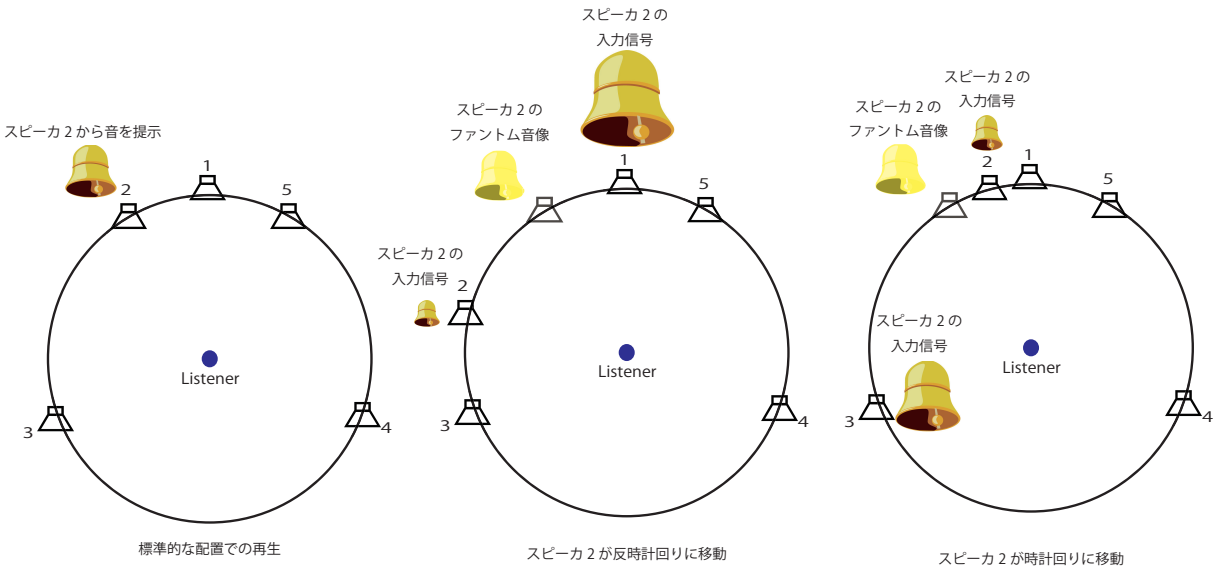

図 4 ファントム音像を提示することによる音場補正

れる．よって，あるスピーカへの入力信号は，そのスピーカがどれだけ標準的な位置からずれているかに応じて入力信号の振幅が変化することになる．スピーカの位置情報 $\boldsymbol{D}$ の要素から決まる各チャンネルの振幅情報を $\boldsymbol{W}$ とし，このように定義する．

$$\boldsymbol{W} = \begin{pmatrix} w(d_1) \\ w(d_2) \\ \vdots \\ w(d_N) \end{pmatrix} \tag{6}$$

さらに，$N \times N$ の正方行列の対角成分以外を 0 にする関数 $\mathrm{diag}(\boldsymbol{M})$ を定義する[*10]．すると，各チャンネルの入力信号の振幅を $w(d_n)$ により変化させた信号を要素とするベクトルは，以下で求まる．

$$\begin{pmatrix} w(d_1)x_1 \\ w(d_2)x_2 \\ \vdots \\ w(d_N)x_N \end{pmatrix} = \left(\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \mathrm{diag}(\boldsymbol{W}\boldsymbol{X}^{\mathrm{T}}) \right)^{\mathrm{T}} \tag{7}$$

また，そのチャンネルの信号は，スピーカの角度のずれの大きさに応じて，角度がずれた方向とは反対のスピーカへ加えられることになる．あるスピーカのインデックスを n として，の反時計回り方向にある隣のスピーカを $n+1$，時計回り方向にある隣のスピーカを $n-1$ とすると，$n+1$ 番目のスピーカへの入力信号 $\boldsymbol{X}'$，$n-1$ 番目のスピーカへの入力信号 $\boldsymbol{X}''$ は次のようになる．

$$\boldsymbol{X}' = \begin{pmatrix} x_N \\ x_{2-1} \\ x_{3-1} \\ \vdots \\ x_{N-1} \end{pmatrix} \qquad \boldsymbol{X}'' = \begin{pmatrix} x_{1+1} \\ x_{2+1} \\ \vdots \\ x_N \\ x_1 \end{pmatrix} \tag{8}$$

同様に，角度のずれに応じて，両隣のスピーカへ入力する振幅のベクトル $\boldsymbol{W}'$，$\boldsymbol{W}''$ を定義する[*11]．すると，最終的な補正後の各スピーカからの出力信号 $\boldsymbol{Y}$ は，次のように定義される．

$$\begin{aligned} \boldsymbol{Y} = &\left(\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \mathrm{diag}(\boldsymbol{W}\boldsymbol{X}^{\mathrm{T}}) \right)^{\mathrm{T}} \\ &+ \left(\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \mathrm{diag}(\boldsymbol{W}'\boldsymbol{X}'^{\mathrm{T}}) \right)^{\mathrm{T}} \\ &+ \left(\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \mathrm{diag}(\boldsymbol{W}''\boldsymbol{X}''^{\mathrm{T}}) \right)^{\mathrm{T}} \end{aligned} \tag{9}$$

ここで，$\boldsymbol{W}$，$\boldsymbol{W}'$，$\boldsymbol{W}''$ は，スピーカの位置情報 $\boldsymbol{D}$ によって決まる各チャンネルの振幅を要素とする列ベクトルである．残された問題は，これらの振幅情報を，スピーカの位置情報のみからどのように決めるかということである．これらの振幅情報は，等距離パンニングの振幅の決め方をそのまま利用できる．

スピーカの位置のずれが大きくればなるほど，そのスピーカ自身の入力信号は，ずれに応じて小さくなる．よって，そのスピーカ自身の入力信号の振幅 $\boldsymbol{W}$ の要素は，式 3 を使って，以下の式で計算できる．

$$w(d_n) = \cos\left(2\pi \frac{a_n |d_n|}{4}\right) \tag{10}$$

a_n は，角度のずれを 0〜1 に正規化する係数である．

一方，位置がずれた場合，そのずれた方向とは反対方向

*10 $\boldsymbol{M}$ は $N \times N$ の正方行列

*11 $\boldsymbol{X}'$，$\boldsymbol{X}''$ と同様のインデックスの付け方

ⓒ 2013 Information Processing Society of Japan

にある隣のスピーカの入力信号に加えられる．スピーカが時計回りにずれた場合は，反時計回り方向の隣にあるスピーカの入力信号に加えられる．反対に，スピーカが反時計回りにずれた場合は，時計回り方向の隣にあるスピーカの入力信号に加えられる．これを等距離パンニングの手法を用いて表すと，$\boldsymbol{W}'$ と $\boldsymbol{W}''$ の要素は次のように決まる．

$$w'(d_n) = g(d_n)\sin\left(2\pi\frac{a_n|d_n|}{4}\right) \tag{11}$$

$$w''(d_n) = (1-g(d_n))\sin\left(2\pi\frac{a_n|d_n|}{4}\right) \tag{12}$$

$$g(d_n) = \begin{cases} 1 & (d_n > 0) \\ 0 & (otherwise) \end{cases} \tag{13}$$

a_n は，角度のずれを 0〜1 に正規化する係数である．式 10，式 12 および式 13 の a_n をどのように決めるかによって，ファントム音像の位置が変わることは，等距離パンニングの定義から明らかである．よって，この値をどのように決めるかという問題は重要な問題である．この値は，実際に想定されるサラウンド環境によって最適な値が決まると考えられ，簡単に定義することはできない．本研究では，スピーカのずれの範囲をそのスピーカの標準的な位置とずれた方向にある隣のスピーカまでとしたため，標準的なスピーカ配置の隣同士の角度を利用して，以下とした．

$$a_n = \begin{cases} \dfrac{1}{|d_n - d_{n+1}|} & (d_n > 0) \\ \dfrac{1}{|d_n - d_{n-1}|} & (otherwise) \end{cases} \tag{14}$$

以上が，スピーカの位置のずれをファントム音像を生成することによって補正するアルゴリズムである．位置がずれているスピーカが 1 つだけの場合，ずれの角度をそのまま利用して補正を行うことができる．

一方，位置がずれているスピーカが 2 つ以上ある場合は，n 番目のスピーカの両隣のスピーカのずれ $\boldsymbol{D}'$，$\boldsymbol{D}''$ を用いて[*12]，スピーカ n のずれの角度 $\boldsymbol{D}$ を次のように修正した角度 $\boldsymbol{A}$ を変数として式 9 の補正を行う．

$$\boldsymbol{A} = \boldsymbol{D} + \boldsymbol{D}' + \boldsymbol{D}'' \tag{15}$$

4. 評価実験

我々は，提案手法をアプリケーションとして実装し，実際に被験者にサラウンド音声を聴かせることによる評価実験を行った．本章では，評価実験の概要と結果について述べる．

4.1 実験 1

4.1.1 実験概要

実験は，James らが [9] で行ったサラウンドの音場提示の評価実験を参考にした．実験内容は，サラウンド上である 1 方向から音を提示し，被験者が聴こえたと感じた方向を回答するという実験である．今回，被験者に提示する音として用いたのは，1 秒程度のドラムスの音である[*13]．サラウンド上で 0°，36°，72°，108°，144°，180°，216°，252°，288°，324° の 10 方向からランダムに音を提示した．これを被験者 1 人に対して 100 回行い，音を聴いた被験者がどの方向に音像を知覚したかを回答し，その正答率を調べた．実験は 5.0 ch サラウンドを対象として行い，標準的な配置，補正なしの非標準的な配置[*14]，補正ありの非標準的な配置（補正なしと同じ配置）の 3 つの条件で行った（図 5）．実験は 23 歳〜24 歳の男性 5 名，女性 1 名に対して行った．

4.1.2 結果と考察

実験の結果を，図 6，図 7 および図 8 に示す．各グラフの横軸はサラウンド上で提示した音像の角度，縦軸は被験者が回答した角度である．被験者全員の角度毎の回答の平

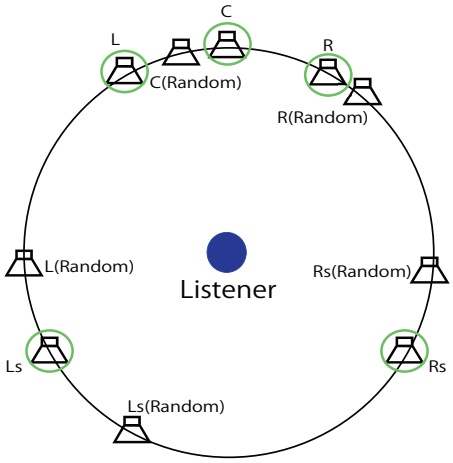

図 5 実験で用いたスピーカ配置．丸で囲んだスピーカが標準的な配置．

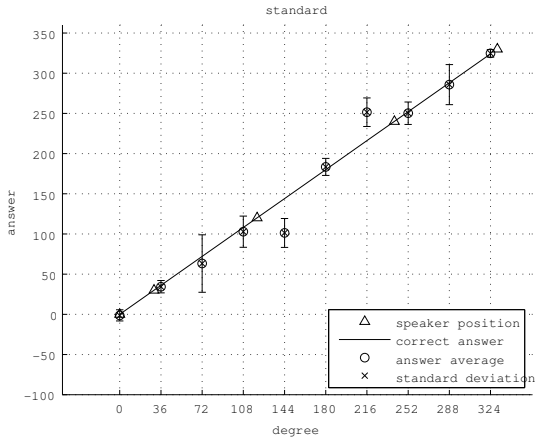

図 6 標準的な 5.0 ch 配置での回答

*12 $\boldsymbol{X}'$，$\boldsymbol{X}''$ と同様のインデックスの付け方

*13 この実験を行う前に数種類の音で音像定位実験を行った結果，ドラムスの音が最も音像を知覚しやすいことが分かった．

*14 標準的なスピーカ配置から乱数を用いて角度をずらした配置

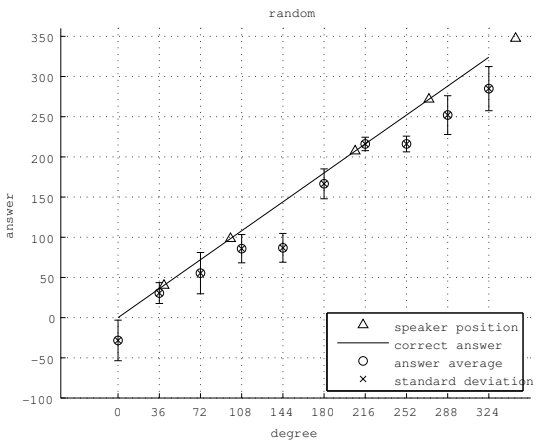

図 7 非標準的な 5.0 ch 配置（補正なし）での回答

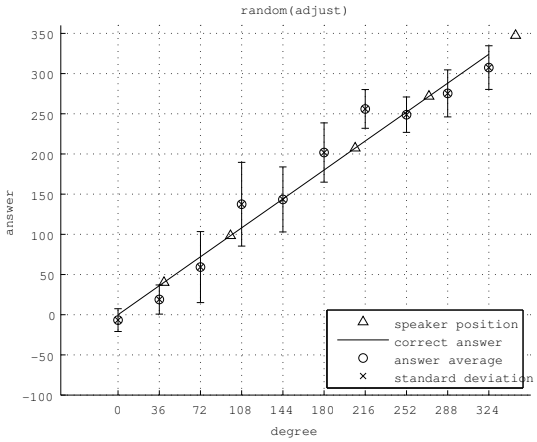

図 8 非標準的な 5.0 ch 配置（補正あり）での回答

均値と標準偏差をプロットしている．また，正解の値を分かりやすくするために，縦軸の値と横軸の値が等しい所に直線を引いた．この直線上に近い所に結果がプロットされているほど多くの被験者が正しい角度を回答したということが分かる．この結果から，0°，252°，288°，324° の音像は，標準的な配置から非標準的な配置（補正なし）になった場合，回答の正確さが損なわれていることが分かった．しかし，非標準的な配置（補正あり）では，そうした角度の正確さが回復していることが分かった．

これらの結果を全体の正答率として比較すると，表 1 のようになった．この結果から，有意水準を 5% とすると，標準的な配置から非標準的な配置（補正なし）に配置が変化した場合は，有意に音像定位方向知覚の正確さが損なわれることが分かった．一方，非標準的な配置（補正なし）と非標準的な配置（補正あり）を比較すると，正答率は有意に上昇していないことが分かった．補正が成功している角度は被験者の前側の角度である．そこで，前側だけで正答率を比較した結果，非標準的な配置（補正なし）と非標準的な配置（補正あり）の正答率は，有意に上昇していた．

表 1 全体の正答率

	正答率	t 値
標準	0.635	-14.0935
非標準（補正なし）	0.4167	
非標準（補正あり）	0.4183	-0.0585

表 2 前側の正答率

	正答率	t 値
標準	0.7545	-9.3442
非標準（補正なし）	0.3901	
非標準（補正あり）	0.5017	-2.7082

4.2 実験 2

我々は，4.1 の結果を詳しく考察するために，更に実験を行った．以下にその詳細を述べる．

4.2.1 実験概要

被験者に対して 1 秒程度のドラムスの音を 2 回連続で提示した．この 2 回の音は，以下の条件からランダムに提示した．

条件 1:
標準的な配置→標準的な配置

条件 2:
標準的な配置→非標準的な配置（補正なし）

条件 3:
標準的な配置→非標準的な配置（補正あり）

条件 4:
非標準的な配置（補正なし）→非標準的な配置（補正あり）

これらの条件から各条件が 2 回ずつ現れるような確率テーブルを作成し，提示する順番を決定した．

被験者は 1 回目と 2 回目の音を聴き比べ，音がどちら方向に移動して聴こえたかを回答し，その回答内容を調べた．被験者は実験 1 と同じ被験者である．

4.2.2 結果と考察

紙面の都合上，条件 3 と条件 4 の結果を図 9，図 10 に示

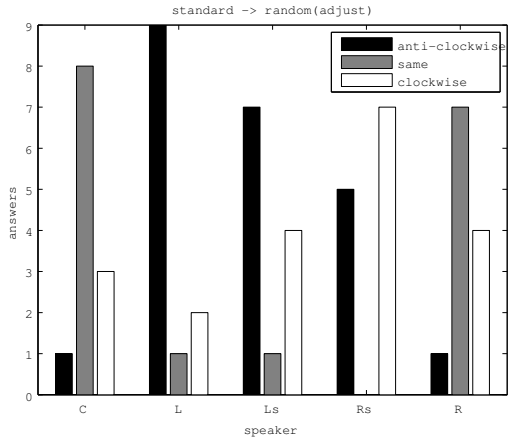

図 9 条件 3 の回答

す．条件 3 では，標準的な配置の音の後，補正ありの非標準的な配置の音を提示している．そのため，被験者が「同じ方向から音が聴こえた」と回答すれば補正が成功していることになる．図 9 から，C と R チャンネルからの提示以外では，「同じ方向から音が聴こえた」という回答はほとんどないことが分かった．この結果は，提案手法ではスピーカ配置がずれた場合に，標準的な配置での提示と同じ位置から音を提示することは成功していないことを示している．

一方，条件 4 では補正なしの非標準的な配置の音の後，補正ありの非標準的な配置の音を提示している．そのため，「スピーカのずれとは反対方向に音が動いた」と回答すれば，補正が成功していることになる．図 5 より，各スピーカは C，L，Ls，Rs チャンネルが反時計回り，R チャンネルが時計回り方向のずれている．図 10 より，「スピーカのずれている方向とは逆方向に音が移動した」という回答が多かったことが分かった．

以上の結果から，本研究の提案手法は，スピーカが非標準的に配置された時，それらのスピーカが提示する音像は，提案手法がない場合と比較して，標準的な配置にあるスピーカと全く同じ方向に音が定位して知覚させることに成功しているとは言えない．しかし，そのスピーカのずれた方向とは反対方向に音を定位させることはできるということが分かった．

5. おわりに

本研究では，サラウンドによる音場提示について，スピーカ配置がずれても正確な音像定位が可能な音場補正手法を提案し，その評価を行った．提案手法では，配置がずれているスピーカの周辺のスピーカから補助的に音を出すことにより，標準的な配置のスピーカから聴こえている筈の音をファントム音像として提示する．その手法をアプリケーションとして実装し，ヒトに対してサラウンド音声を聴かせることによる評価実験を行った．その結果，スピーカ配置がずれたことによって正確に伝わらなくなった音像定位は，提案手法によってある程度正確さが回復することが分かった．一方で，提案手法では，標準的な配置での音像定位方向と同じ方向に知覚させることは難しいことも分かった．

今後は，補正手法の中でも決定法が不明瞭なパラメータ（式 10〜式 13 の a_n）の決定法の確立や幅広い評価実験（より多くの被験者に対する実験，無響室での客観評価など）を行っていく予定である．

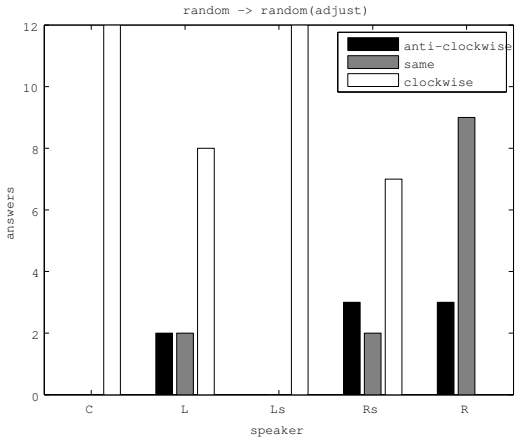

図 10 条件 4 の回答

参考文献

[1] 古山俊一 ”5.1 サラウンド・サウンド セミナー報告”, AMEI NEWS 3, 2004
[2] 沢口真生，中原雅考，亀川徹 ”サラウンド入門”, 東京藝術大学出版, 2010
[3] Morimoto Masayuki,Ando Yoichi, ”On the simulation of sound localization”, Journal of the Acoustical Society of Japan (E), Acoustical Society of Japan, (1980)
[4] Frederic L. Wightman, Doris J. Kistler, ”Headphone simulation of free-field listening. I: Stimulus synthesis”, J. Acoust. Soc. Am. 85 (2), February (1989)
[5] Frederic L. Wightman, Doris J. Kistler, ”Headphone simulation of free - field listening. II: Psychophysical validation”, J. Acoust. Soc. Am. Volume 85, Issue 2, pp. 868-878 (1989)
[6] Wenzel EM, Arruda M, Kistler DJ, Wightman FL, ”Localization using nonindividualized head-related transfer functions”, J Acoust Soc Am. (1993)
[7] 山崎 大輔, 橋本 修, 井上 勝夫, 天野 圭介 ,「明瞭度と音像定位方向知覚に対して貢献する拡声音周波数特性の検討：残響，ノイズ下における避難誘導に対する拡声音周波数制御法の適用 その 1.(音声伝達と明瞭性, 環境工学 I)」, 学術講演梗概集. D-1, 環境工学 I, 室内音響・音環境, 騒音・固体音, 環境振動, 光・色, 給排水・水環境, 都市設備・環境管理, 環境心理生理, 環境設計, 電磁環境, 社団法人日本建築学会, 2004
[8] 梅津 直貴, 井ノ上 寛人, 堀内 恒, 佐藤 美恵, 小黒 久史, 春日 正男,「空間把握性に注目した音響案内システムの開発に関する研究 (感性とメディア及び一般)」, 映像情報メディア学会技術報告 13426893 一般社団法人映像情報メディア学会 2011
[9] James R. West, ”Five-Channel Panning Laws:An Analytical and Experimental Comparison”, Submitted to the Faculty of the University of Miami in partial fulfillment of the requirements for the degree of Master of Science in Music Engineering Technology, (1998)
[10] Ki-Seung Lee, Seok-Pil Lee, ”A real-time audio system for adjusting the sweet spot to the listener's position”, IEEE, (2010)
[11] Eugeniusz Kornatwski, ”Stereo to the “Real Surround Sound” Conversion Algorithm”, IEEE, (2011)
[12] 「5.1ch，6.1ch 信号を変換し，7.1ch サラウンドの豊かな音場を実現 高音質な音声フォーマット”DTS 96/24”にも対応した AV レシーバー 発売」http://www.sony.jp/CorporateCruise/Press/200209/02-0905/
[13] 「TA-DA5600ES 商品の特長」http://www.sony.jp/audio/products/TA-DA5600ES/feature_1.html
[14] Curtis Roads; 青柳龍也，小坂直敏，平田圭二， 堀内靖雄 訳・監修; 後藤真孝，引地孝文，平 野砂峰旅，松島俊明 訳:『コンピュータ音楽 歴史・テクノロジー・アート』東京電機大学 出版局 (2001)