

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 17 日現在

機関番号：82636

研究種目：基盤研究(C)

研究期間：2011～2014

課題番号：23500162

研究課題名(和文)電子透かしによる拡張音響メディアの実現

研究課題名(英文)Enrichment of audio media by means of information hiding

研究代表者

西村 竜一(Nishimura, Ryouichi)

独立行政法人情報通信研究機構・ユニバーサルコミュニケーション研究所 音声コミュニケーション研究室・研究マネージャー

研究者番号：30323116

交付決定額(研究期間全体)：(直接経費) 3,800,000円

研究成果の概要(和文)：広い範囲で高い精度で音場を再現するには、必要となるマイクロホンやスピーカのチャンネル数が必然的に多くなる。一方、ヘッドホン再生では、聴取者の特性を考慮することにより、わずか2チャンネルの信号でも空間性豊かな音が再生できる。本研究はこの差に着目したもので、スピーカ再生時には5.1チャンネルサラウンド信号として利用でき、かつ、それをヘッドホンやイヤホンで再生した場合には、臨場感豊かなバイノーラル音になるような音響メディアの合成手法を提案した。計算機シミュレーションおよび聴取実験の結果、元々の5.1チャンネル信号との物理的および主観的誤差も小さい、所望の特性を有する信号が合成されていることが確認された。

研究成果の概要(英文)：For sound field reproduction of a wide area with high precision, the number of microphones and loudspeakers required inevitably increases. In contrast, good spatial sound can be reproduced with only two channel signals in headphone listening by considering individual characteristics of listener's ears. This research focuses on this point and proposes a method to compose an audio media that is available as a 5.1 channel surround signal in loudspeaker listening and also becomes a binaural signal for the listener in headphone or earphone listening. It is shown by computer simulations and listening tests that the composed media has little difference from the original 5.1 channel surround signal with respect to both objective measures and subjective impressions. In addition, the method assures that the reproduced binaural signal in headphone listening becomes the desired one.

研究分野：音響信号処理

キーワード：情報ハイディング 音響メディア 機能拡張 5.1サラウンド 立体音響 聴取実験 制約条件・重み付き最小二乗法 イヤホン再生

1. 研究開始当初の背景

(1) 社会的背景 - 音楽聴取環境の変化

今日の音楽聴取環境は、再生ハードウェアの小型・軽量化・大容量化・無線通信機能等の技術革新により、数十年前とは大きく変化している。その結果、図1に示すように、パ

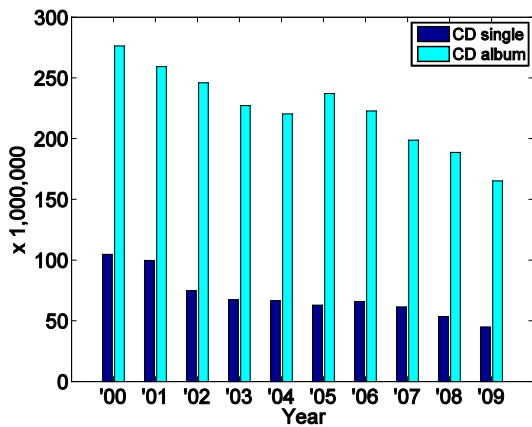


図1 CDの生産量の推移

ッケージメディアとしてのCDの生産量は減衰の一途を辿っているのに対し、音楽データのダウンロード数は着実に数を伸ばしている。しかし、ダウンロードされるデータはCDのデータと同じであり、録音スタジオでモニタースピーカを使ってオーディオエンジニアがミキシングした音を、どのような方法で再生されるのかを考慮せず、一様に配布・配信しているのが現状である。双方向性通信が手軽に利用可能になった現在においては、聴取者の情報を配信元に送付することで、各ユーザーの聴取環境ごとに最適なデジタル音楽情報を個々に配布することも可能である。

(2) 学術的背景 - 電子透かしの応用研究

音楽の不正コピー防止手段として、多くの音響電子透かし手法が国内外を問わず提案されたが、その応用に関する研究は少ない。背景には、多くのアルゴリズムが、公開されると攻撃耐性が低下し、実用的でなくなるという理由もあるが、音楽業界が電子透かしの採用に対して積極的になる材料が乏しい点も挙げられる。レコーディングエンジニアからすると、自分がミキシングした原曲に僅かでも変化を加えることは、歪を与えて音楽を劣化させることに等しい。このため、音楽に対する音響電子透かしは、不正コピーを取り締まる側のみメリットがあり、音楽制作者、聴取者にはデメリットのほうが大きかった。そこで、音響電子透かしをメディアの高機能化に応用し、音楽制作者、聴取者の両方にとっても有益となる利用法が提案できれば、音響電子透かしの応用の可能性拡大に繋がる。

2. 研究の目的

(1) 概要

広い範囲で高い精度で音場を再現するには、

多くのマイクロホンやスピーカが必要となるため、必要となるチャンネル数も必然的に多くなる。一方、ヘッドホン再生では、聴取者の特性を考慮することにより、わずか2チャンネルの信号でも空間性豊かな音を提示できるポテンシャルを有している。本研究はこの差に着目したものであり、スピーカ再生用の空間性豊かな音信号に、それよりもずっと少ない情報量で同等の空間性を再現できるヘッドホン用の信号を生成するのに必要な情報を埋め込み、ひとつのデータでスピーカ再生にもヘッドホン再生にも適した音メディアを実現することである。

(2) 具体的説明

本研究で想定する音信号の配信環境は、次のようなものである。制作側は、業務利用も想定して、非常に多くのチャンネルで空間情報豊かに音を記録する。しかし、パーソナル利用では、通信帯域や再生システム規模の制限により、おのずとチャンネル数を減らしたものとなる。一方、ヘッドホン用信号は、僅か2チャンネルで良いが、空間性を出すには個人適応が必要不可欠である。そこで、聴取者の情報を制作側に送付し、空間情報が豊かな制作側でユーザーのヘッドホン用の音を作成し、それをスピーカ再生用の信号から再合成するのに必要なパラメータを透かしとして埋め込んだスピーカ再生用の信号を生成する方法が考えられる。本研究では、このシステムの実現を目指し、5.1チャンネルサラウンド信号に着目して必要な調査及び研究開発を行う。

3. 研究の方法

(1) 立体音の原理

ある空間で音源から聴取者の耳までに音が受ける変化は頭部伝達関数で表される。この頭部伝達関数と音源信号とを掛けることによりバイノーラル信号を生成することができる。ただし、頭部伝達関数は聴取者の頭部形状と部屋の音響特性の両方の影響を受けるため、聴取本人がその部屋で測定する必要がある。これは頭部形状による特性が主に音の到来方向で決まるのに対し、各方向からの到来音が部屋の特性で変化するためである。5.1チャンネルサラウンドでは、部屋の壁からの反射音も音場を豊かにすると考えられるが、想定する部屋へ各聴取者が赴いて測定するのは、必ずしも現実的ではない。そこで、バイノーラル信号の入手方法として、聴取者の頭部形状に起因する特性と部屋に起因する特性を独立に取得し、それらからバイノーラル信号を生成する手法を使用する。

(2) バイノーラル音の生成

本研究では、図2に示すような理想的なサラウンド信号の聴取環境を考える。各スピーカから境界面上の各マイクロホン位置までのインパルス応答を測定し、聴取者の頭部伝

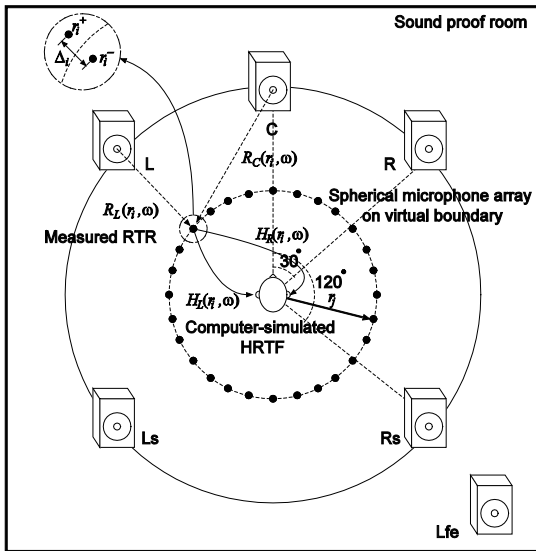


図 2 サラウンドスピーカによる聴取

達関数と組み合わせて、ヘッドホン聴取に適したバイノーラル信号を生成する。問題は、5.1 チャンネルサラウンド信号を 2 チャンネルの再生装置で再生するときに行われる単純なミキシングをしたときにこのバイノーラル信号が生成されるように 5.1 チャンネルサラウンド信号を如何にして合成するかである。この目的を達成するため、拘束条件・重み付き最小二乗法に着目した。

4. 研究成果

(1) 拘束条件・重み付き最小二乗法

拘束条件・重み付き最小二乗法を用いることにより、単純なミキシングをした結果が、所望のバイノーラル信号になるという拘束条件を課した上で、5.1 チャンネルの各チャンネル信号と元々の 5.1 チャンネル信号との差が最小になるように 5.1 チャンネルサラウンド信号を合成することが可能である。さらに、各チャンネル信号のエネルギーに応じて、それぞれのチャンネルの誤差に重みづけを行うことにより、聴感上での歪を低減するように工夫した。数学的に記述とすると次のようになる。

$$\hat{X} = \arg \min_X (X - S)^* P^{-1} (X - S) \\ \text{subject to } QX - Y = 0$$

ここで、 X が求めたい 5.1 チャンネルサラウンド信号、 S が元々の 5.1 チャンネルサラウンド信号、 Q が再生装置におけるサラウンド信号のミキシング処理を表す行列、 Y が所望のバイノーラル信号である。また、 P は、各チャンネルの誤差に対する重み付けを表す。この式をラグランジュの未定乗数法を用いて解くことで、5.1 チャンネルサラウンド信号を合成することができる。

(2) 誤差最小化係数

さらに本研究では、元々の 5.1 チャンネル信

号との音の大きさの違いは、コンテンツの再生装置のボリューム操作で補償できるため、その誤差を許容する誤差最小化係数の導入も検討した。この場合の式は、

$$\hat{X} = \arg \min_X (X - \alpha S)^* P^{-1} (X - \alpha S) \\ \text{subject to } QX - Y = 0$$

となる。ただし、この誤差最小化係数 α は、着目している時間区間ごとに変化するため、それを全時間区間で平均した平均誤差最小化係数を用いた場合についても検討した。それぞれの手法を適用した際に、元々の 5.1 チャンネルサラウンド信号との差がどの程度生じているのかを図示したのが図 3 である。図中の各結果は、(A) 拘束条件付き最小二乗法、(B) 拘束条件・重み付き最小二乗法、(C) 誤差最小化係数付き拘束条件・重み付き最小二乗法、(D) 平均誤差最小化係数付き拘束条件・重み付き最小二乗法のそれぞれの誤差である。また、図中の**のマークは、これらの条

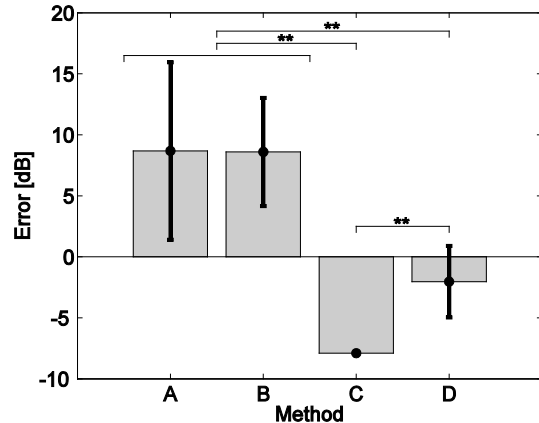


図 3 各手法における平均残差

件の間には統計的に意味のある差があることを示している。この図から、拘束条件付き最小二乗法と拘束条件・重み付き最小二乗法に統計学的に有意な差はないが、誤差最小化係数を導入することにより、これらの手法よりも有意に誤差が低下していることが分かる。

(3) 聴取実験による評価

以上の方法では、求めるパラメータの数よりもそのパラメータを規定している方程式の数のほうが少ない劣決定系の問題として解を導出しているため、サラウンド信号をヘッドホンで聴取した場合には、理想的な立体音と完全に一致する音が聴取される。一方、サラウンドスピーカで聴取した場合には、サラウンド信号の合成手法の違いに依存した音の劣化が生じるため、その劣化の程度を聴取実験により調査した。実験参加者は、20 歳代の男性 5 名、女性 6 名の計 11 名であり、聴取は防音室で行った。実験参加者には、原信号基準音として聞かせたうえで、表 1 の各

手法をランダムな順番で聞かせ、それぞれについて、基準音を 100 点だとして各音の「音

表 1 実験条件

要因	条件
手法	(1)原信号、(2)拘束条件付き最小二乗法、(3)拘束条件・重み付き最小二乗法、(4)平均誤差最小化係数付き拘束条件・重み付き最小二乗法、(5)アンカー
刺激	5.1ch サラウンドのシネマコンテンツから、各 10 秒で切り出した 10 区間

質」がいくらかなのかを判断させた。なお、アンカーとは、ITU-R BS.1534 (Multi stimulus test with hidden reference and anchor: MUSHRA) で求められる、一定の方法で音質を劣化させた信号である。回答は、図

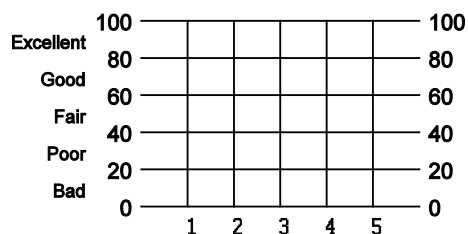


図 4 回答用紙

4 に示す回答用紙に記入させ、実験参加者の要求に応じて、繰り返し刺激音を聞き返すことを許可した。また、参加者にいい加減な回答をさせないように、刺激の中に原信号と同じものが一つ以上入っていることを告げて、それには 100 点を付すように教示した。一方、最も悪い音には、0 点を付けるように指示した。

(4) 聴取実験結果

評価結果を箱ひげ図で示したものが図 5 である。この図で、R: 原信号、A: 拘束条件付

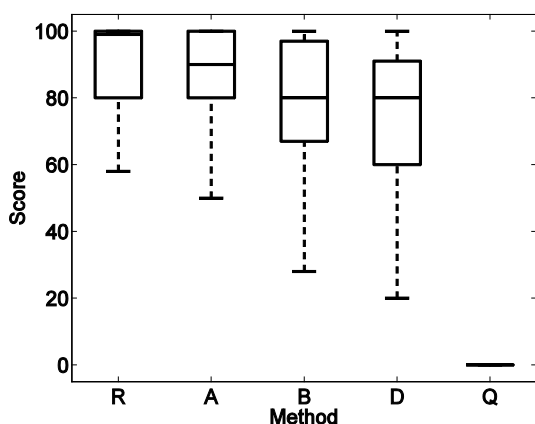


図 5 聴取実験の評価結果

き最小二乗法、B: 拘束条件・重み付き最小二乗法、D: 平均誤差最小化係数付き拘束条件・重み付き最小二乗法、Q: アンカーの結果である。原信号が一番評価が高く、拘束条

件付き最小二乗法、拘束条件・重み付き最小二乗法、平均誤差最小化係数付き拘束条件・重み付き最小二乗法の順に評価が下がっている。ただし、統計的な検定を行った結果、実験に用いた 10 種類の音源のうち 6 種類で差があるとは言えないということが確認された。今回、「音質」に対して評価をさせたが、音質は本来、多次的なものであり、マルチチャンネル信号に対する評価項目として、「前方音像感」や「サラウンド感」も存在する。したがって、この実験においてどのような音の側面に聴取者が着目して評価をしていたかについては、慎重に検討する必要がある。例えば、知覚された音質の劣化は、空間的な音像のゆがみや、音のひずみに起因するものではなく、音響情動的には空間性が豊かになっているにも関わらず、普段ヘッドホンで聞きなれている音と音色が異なっているために、低い評点を下している可能性も考えられる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 1 件)

- ① 西村 竜一、藪田 光太郎、制約条件・重み付き最小二乗法によるヘッドホン受聴を志向したサラウンド信号の再合成、電子情報通信学会論文誌、査読有、印刷中

[学会発表] (計 8 件)

- ① R. Nishimura, K. Sonoda, Surround sound signals generated by constrained weighted least squares for headphone listening, Forum Acusticum, Sep. 9, 2014, Krakow (Poland)
- ② 西村 竜一、藪田 光太郎、高臨場感ヘッドホン再生のための制約条件・重み付き最小二乗法によるサラウンド信号合成、電子情報通信学会マルチメディア情報ハイディング・エンリッチメント研究会、2014 年 3 月 8 日、大江町交流プラザ (石川県・金沢市)
- ③ 西村 竜一、加藤 宏明、竹本 浩典、Mokhtari Parham、バイノーラル Ambisonics の合成係数の導出に関する検討、日本音響学会 2012 年秋季研究発表会、2012 年 9 月 19 日、信州大学 (長野県・松本市)

6. 研究組織

(1) 研究代表者

西村 竜一 (NISHIMURA, Ryouichi)

独立行政法人情報通信研究機構・ユニバーサルコミュニケーション研究所・研究マネージャー

研究者番号: 30323116

(2) 研究分担者

藺田 光太郎 (SONODA, Kotaro)
長崎大学・工学部・助教
研究者番号：9 0 4 1 5 8 5 2

(3) 連携研究者

坂本 修一 (SAKAMOTO, Shuichi)
東北大学・電気通信研究所・准教授
研究者番号：6 0 3 3 2 5 2 4