

(19) 日本国特許庁(JP)

再公表特許(A1)

(11) 国際公開番号

W02017/188141

発行日 平成31年3月7日 (2019.3.7)

(43) 国際公開日 平成29年11月2日 (2017.11.2)

(51) Int.Cl. F I テーマコード (参考)  
**H04S 7/00 (2006.01)** H04S 7/00 300 5D162

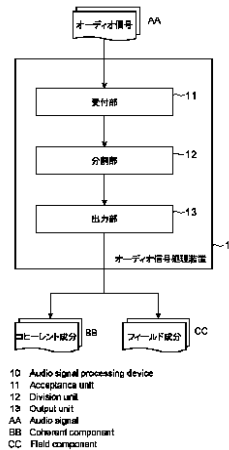
審査請求 未請求 予備審査請求 未請求 (全 26 頁)

<p>出願番号 特願2018-514561 (P2018-514561)</p> <p>(21) 国際出願番号 PCT/JP2017/016019</p> <p>(22) 国際出願日 平成29年4月21日 (2017.4.21)</p> <p>(31) 優先権主張番号 特願2016-89417 (P2016-89417)</p> <p>(32) 優先日 平成28年4月27日 (2016.4.27)</p> <p>(33) 優先権主張国 日本国 (JP)</p>	<p>(71) 出願人 305060567                  国立大学法人富山大学                  富山県富山市五福3190</p> <p>(74) 代理人 100088155                  弁理士 長谷川 芳樹</p> <p>(74) 代理人 100113435                  弁理士 黒木 義樹</p> <p>(74) 代理人 100144440                  弁理士 保坂 一之</p> <p>(72) 発明者 安藤 彰男                  富山県富山市五福3190 国立大学法人                  富山大学内</p> <p>Fターム(参考) 5D162 AA05 BA07 CA27 CB12 DA28                  EG02</p> <p style="text-align: right;">最終頁に続く</p>
--	---

(54) 【発明の名称】 オーディオ信号処理装置、オーディオ信号処理方法、およびオーディオ信号処理プログラム

(57) 【要約】

一実施形態に係るオーディオ信号処理装置は、複数のチャンネルのオーディオ信号を受け付ける受付部と、各チャンネルのオーディオ信号をコヒーレント成分とフィールド成分とに分割する分割部と、各チャンネルのコヒーレント成分およびフィールド成分を出力する出力部とを備える。分割処理では、対象チャンネル以外のチャンネルのオーディオ信号を少なくとも用いて算出される推定信号のうち該対象チャンネルのオーディオ信号との相関が最も高い推定信号が該対象チャンネルのコヒーレント成分として抽出される。そして、対象チャンネルのオーディオ信号およびコヒーレント成分の差分がフィールド成分として抽出される。



**【特許請求の範囲】****【請求項 1】**

複数のチャンネルのオーディオ信号を受け付ける受付部と、  
前記オーディオ信号をコヒーレント成分とフィールド成分とに分割する分割処理を各チャンネルについて実行する分割部であって、前記分割処理が、

前記分割処理の対象となる一つの前記チャンネルを対象チャンネルとした場合に、該対象チャンネル以外のチャンネルの前記オーディオ信号を少なくとも用いて算出される推定信号のうち該対象チャンネルの前記オーディオ信号との相関が最も高い推定信号を該対象チャンネルの前記コヒーレント成分として抽出するステップと、

前記対象チャンネルの前記オーディオ信号と該対象チャンネルの前記コヒーレント成分との差分を該対象チャンネルの前記フィールド成分として抽出するステップと  
を含む、該分割部と、

前記分割部により抽出された各チャンネルの前記コヒーレント成分および前記フィールド成分を出力する出力部と  
を備えるオーディオ信号処理装置。

**【請求項 2】**

前記分割処理が、

窓関数を用いてオーディオ信号を複数のフレームに区切る処理を各チャンネルについて実行するステップと、

連続する少なくとも二つの前記フレームを一つのブロックにまとめる処理を前記複数のフレームの全体に対して実行することで複数の前記ブロックを生成する処理を各チャンネルについて実行するステップと、

前記ブロックのそれぞれにおいて前記対象チャンネルの前記コヒーレント成分を抽出するステップと  
を含む、  
請求項 1 に記載のオーディオ信号処理装置。

**【請求項 3】**

前記分割部が、

各チャンネルのオーディオ信号を複数のサブバンドに分割することで、各チャンネルについて複数のサブバンド信号を生成するステップと、

前記複数のサブバンドのそれぞれにおいて前記対象チャンネルのコヒーレント成分を抽出するステップと、

前記複数のサブバンドにおけるコヒーレント成分を加算することで前記対象チャンネルのコヒーレント成分を抽出するステップと  
を含む、

請求項 1 または 2 に記載のオーディオ信号処理装置。

**【請求項 4】**

オーディオ信号処理装置が、複数のチャンネルのオーディオ信号を受け付ける受付ステップと、

前記オーディオ信号処理装置が、前記オーディオ信号をコヒーレント成分とフィールド成分とに分割する分割処理を各チャンネルについて実行する分割ステップであって、前記分割処理が、

前記分割処理の対象となる一つの前記チャンネルを対象チャンネルとした場合に、該対象チャンネル以外のチャンネルの前記オーディオ信号を少なくとも用いて算出される推定信号のうち該対象チャンネルの前記オーディオ信号との相関が最も高い推定信号を該対象チャンネルの前記コヒーレント成分として抽出するステップと、

前記対象チャンネルの前記オーディオ信号と該対象チャンネルの前記コヒーレント成分との差分を該対象チャンネルの前記フィールド成分として抽出するステップと  
を含む、該分割ステップと、

前記オーディオ信号処理装置が、前記分割ステップにおいて抽出された各チャンネルの前

10

20

30

40

50

記コヒーレント成分および前記フィールド成分を出力する出力ステップとを含むオーディオ信号処理方法。

【請求項 5】

複数のチャンネルのオーディオ信号を受け付ける受付ステップと、

前記オーディオ信号をコヒーレント成分とフィールド成分とに分割する分割処理を各チャンネルについて実行する分割ステップであって、前記分割処理が、

前記分割処理の対象となる一つの前記チャンネルを対象チャンネルとした場合に、該対象チャンネル以外のチャンネルの前記オーディオ信号を少なくとも用いて算出される推定信号のうち該対象チャンネルの前記オーディオ信号との相関が最も高い推定信号を該対象チャンネルの前記コヒーレント成分として抽出するステップと、

前記対象チャンネルの前記オーディオ信号と該対象チャンネルの前記コヒーレント成分との差分を該対象チャンネルの前記フィールド成分として抽出するステップとを含む、該分割ステップと、

前記分割ステップにおいて抽出された各チャンネルの前記コヒーレント成分および前記フィールド成分を出力する出力ステップと

をコンピュータに実行させるオーディオ信号処理プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明の一側面は、オーディオ信号処理装置、オーディオ信号処理方法、およびオーディオ信号処理プログラムに関する。

【背景技術】

【0002】

オーディオ信号のチャンネル数を変更する手法が従来から知られている。具体的には、Mチャンネルのオーディオ信号をNチャンネル（ただし、 $N > M$ ）のオーディオ信号に変換するアップミックスという手法と、Nチャンネルのオーディオ信号をMチャンネルのオーディオ信号に変換するダウンミックスという手法が存在する。例えば、2チャンネル（左チャンネルおよび右チャンネル）のオーディオ信号から5.1チャンネルのオーディオ信号への変換はアップミックスの一例である。また、5.1チャンネルのオーディオ信号から2チャンネルのオーディオ信号への変換はダウンミックスの一例である。

【0003】

例えば下記特許文献1には、テレビ・ラジオのスポーツ実況番組のステレオ放送を、迫力ある臨場感と聴き取りやすいアナウンスとするサラウンド再生装置が記載されている。この装置はフロント左/右チャンネル信号創成手段、フロントセンタチャンネル信号創成手段、およびリア左/右サラウンドチャンネル信号創成手段を有する。フロント左/右チャンネル信号創成手段は、2チャンネル音声信号入力に対して、マトリクス処理を行って得たフロント左/右チャンネル用各音声信号に、残響音を選択的に付加すると共にフロント用音量調整を行い、フロント左/右チャンネル用各音声信号として出力する。フロントセンタチャンネル信号創成手段は、2チャンネル音声信号入力から、同相成分を抽出して得た音声信号に、残響音を付加せずにフロントセンタチャンネル用音声信号としてセンタ用音量調整を行って出力する。リア左/右サラウンドチャンネル信号創成手段は、マトリクス処理を行って得たフロント左/右チャンネル用各音声信号に、残響音を付加すると共にリア用音量調整を行い、リア左/右チャンネル用各音声信号として出力する。

【0004】

下記非特許文献1, 2はいずれも、アップミックスの手法を記載する文献である。非特許文献1には、ステレオ信号を帯域分割し、帯域ごとにステレオ信号を主信号とアンビエンス信号とに分割し、アンビエンス信号を5.1チャンネルの後方チャンネルから再生する手法が記載されている。非特許文献2には、ステレオ信号を帯域分割した後に、そのステレオ信号を直接音成分と残響音成分とに分割し、残響音成分を側方から再生する方法が記載されている。

10

20

30

40

50

## 【 0 0 0 5 】

下記非特許文献 3 , 4 はいずれも、多チャンネルのオーディオ信号を 2 チャンネルのオーディオ信号のペアに分割することで、3 チャンネル以上のオーディオ信号を生成する手法を開示する。

## 【 先行技術文献 】

## 【 特許文献 】

## 【 0 0 0 6 】

【 特許文献 1 】 特開 2 0 0 7 - 2 8 0 6 5 号公報

## 【 非特許文献 】

## 【 0 0 0 7 】

【 非特許文献 1 】 C. Avendano and J-M Jot, "A Frequency-Domain Approach to Multichannel Upmix," J. Audio Eng. Soc., Vol. 52, No. 7/8, pp. 740-749, 2004

【 非特許文献 2 】 C. Faller, "Multiple-Loudspeaker Playback of Stereo Signals," J. Audio Eng. Soc., Vol. 54, No. 11, pp. 1051-1064, 2006

【 非特許文献 3 】 J. Thompson, B. Smith, A. Warmer and J-M Jot, "Direct-Diffuse Decomposition of Multichannel Signals Using a System of Pairwise Correlations," Proc. Audio Eng. Soc. 133rd Convention, Paper no. 8807, 2012

【 非特許文献 4 】 C. Faller, L. Altmann, J. Levinson and M. Schmidt, "Multichannel Ring Upmix," Proc. Audio Eng. Soc. 134th Convention, Paper no. 8908, 2013

## 【 発明の概要 】

## 【 発明が解決しようとする課題 】

## 【 0 0 0 8 】

特許文献 1 に記載のサラウンド再生装置は原音に残響音を付加するため、再生音の雰囲気（例えば音色）が原音から変わったり損なわれたりしてしまう。これに対して非特許文献 1 , 2 に記載の手法は残響音を付加するものではないが、原理上、2 チャンネルのオーディオ信号（すなわち、ステレオ信号）にしか適用できない。

## 【 0 0 0 9 】

非特許文献 3 , 4 に記載の手法では、2 チャンネルのオーディオ信号の間で相関が高い成分をコヒーレント成分として抽出するので、二つのスピーカの間付近に位置する音の情報を取得することになる。したがって、3 チャンネル以上のオーディオ・システムでは、任意の二つのスピーカの間付近の音の情報だけしかコヒーレント成分として抽出することができず、全スピーカで囲まれた領域の中央部分に位置する音の情報を抽出することができない。

## 【 0 0 1 0 】

そこで、原音のチャンネル数にかかわらず、オーディオ信号のチャンネル数を変更する際に原音の雰囲気を可能な限り維持する手法が望まれている。

## 【 課題を解決するための手段 】

## 【 0 0 1 1 】

本発明の一側面に係るオーディオ信号処理装置は、複数のチャンネルのオーディオ信号を受け付ける受付部と、オーディオ信号をコヒーレント成分とフィールド成分とに分割する分割処理を各チャンネルについて実行する分割部であって、分割処理が、分割処理の対象となる一つのチャンネルを対象チャンネルとした場合に、該対象チャンネル以外のチャンネルのオーディオ信号を少なくとも用いて算出される推定信号のうち該対象チャンネルのオーディオ信号との相関が最も高い推定信号を該対象チャンネルのコヒーレント成分として抽出するステップと、対象チャンネルのオーディオ信号と該対象チャンネルのコヒーレント成分との差分を該対象チャンネルのフィールド成分として抽出するステップとを含む、該分割部と、分割部により抽出された各チャンネルのコヒーレント成分およびフィールド成分を出力する出力部とを備える。

## 【 0 0 1 2 】

本発明の一側面に係るオーディオ信号処理方法は、オーディオ信号処理装置が、複数の

10

20

30

40

50

チャンネルのオーディオ信号を受け付ける受付ステップと、オーディオ信号処理装置が、オーディオ信号をコヒーレント成分とフィールド成分とに分割する分割処理を各チャンネルについて実行する分割ステップであって、分割処理が、分割処理の対象となる一つのチャンネルを対象チャンネルとした場合に、該対象チャンネル以外のチャンネルのオーディオ信号を少なくとも用いて算出される推定信号のうち該対象チャンネルのオーディオ信号との相関が最も高い推定信号を該対象チャンネルのコヒーレント成分として抽出するステップと、対象チャンネルのオーディオ信号と該対象チャンネルのコヒーレント成分との差分を該対象チャンネルのフィールド成分として抽出するステップとを含む、該分割ステップと、オーディオ信号処理装置が、分割ステップにおいて抽出された各チャンネルのコヒーレント成分およびフィールド成分を出力する出力ステップとを含む。

10

**【0013】**

本発明の一側面に係るオーディオ信号処理プログラムは、複数のチャンネルのオーディオ信号を受け付ける受付ステップと、オーディオ信号をコヒーレント成分とフィールド成分とに分割する分割処理を各チャンネルについて実行する分割ステップであって、分割処理が、分割処理の対象となる一つのチャンネルを対象チャンネルとした場合に、該対象チャンネル以外のチャンネルのオーディオ信号を少なくとも用いて算出される推定信号のうち該対象チャンネルのオーディオ信号との相関が最も高い推定信号を該対象チャンネルのコヒーレント成分として抽出するステップと、対象チャンネルのオーディオ信号と該対象チャンネルのコヒーレント成分との差分を該対象チャンネルのフィールド成分として抽出するステップとを含む、該分割ステップと、分割ステップにおいて抽出された各チャンネルのコヒーレント成分およびフィールド成分を出力する出力ステップとをコンピュータに実行させる。

20

**【0014】**

このような側面においては、対象チャンネル以外のチャンネルのオーディオ信号を用いて推定され、且つ該対象チャンネルの実際のオーディオ信号ごとの相関が最も高い信号が該対象チャンネルのコヒーレント成分として抽出される。また、対象チャンネルの実際のオーディオ信号とそのコヒーレント成分との差分が該対象チャンネルのフィールド成分として抽出される。このコヒーレント成分およびフィールド成分は各チャンネルについて得られる。このように、音を追加することなく元のオーディオ信号のみを用いて各チャンネルのコヒーレント成分およびフィールド成分を求めることで、原音の雰囲気可能な限り維持することができる。加えて、コヒーレント成分およびフィールド成分は元のチャンネル数の分だけ求めることができるので、この手法は原音のチャンネル数にかかわらず適用できる。

30

**【発明の効果】****【0015】**

本発明の一側面によれば、原音のチャンネル数にかかわらず、オーディオ信号のチャンネル数を変更する際に原音の雰囲気を可能な限り維持することができる。

**【図面の簡単な説明】****【0016】**

**【図1】** 実施形態に係るオーディオ信号処理の例を示す図である。

**【図2】** 実施形態に係るオーディオ信号処理装置として機能するコンピュータのハードウェア構成を示す図である。

40

**【図3】** 実施形態に係るオーディオ信号処理装置の機能構成を示す図である。

**【図4】** オーディオ信号を処理する単位であるブロックを示す図である。

**【図5】** ある一つのチャンネルにおける処理を示す図である。

**【図6】** 実施形態に係るオーディオ信号処理装置の動作を示すフローチャートである。

**【図7】** 図6に示すコヒーレント成分の抽出の詳細を示すフローチャートである。

**【図8】** 実施形態に係るオーディオ信号処理プログラムの構成を示す図である。

**【図9】** 従来手法におけるコヒーレント成分の抽出の例を示す図である。

**【図10】** 実施形態におけるコヒーレント成分の抽出の例を示す図である。

**【発明を実施するための形態】****【0017】**

50

以下、添付図面を参照しながら本発明の実施形態を詳細に説明する。なお、図面の説明において同一または同等の要素には同一の符号を付し、重複する説明を省略する。

【0018】

図1～図5を参照しながら、実施形態に係るオーディオ信号処理装置10の機能および構成を説明する。オーディオ信号処理装置10は、複数のチャンネルのオーディオ信号のそれぞれをコヒーレント成分とフィールド成分とに分割するコンピュータである。オーディオ信号は、ヒトが聴くことができる周波数帯域（一般に約20Hz～20000Hz）の音を含むデジタル信号であり、必要に応じてアナログ信号に変換される。オーディオ信号で示される音の例として声、音楽、映像の音、自然音、あるいはこれらの任意の組合せが挙げられるが、これらに限定されるものではない。

10

【0019】

図1は、オーディオ信号処理装置10によるオーディオ信号の処理の一例を示し、より具体的には、2チャンネル（LチャンネルおよびRチャンネル）、すなわちステレオのオーディオ信号の処理を示す。オーディオ信号処理装置10は各チャンネルの信号をコヒーレント成分とフィールド成分とに分割する。

【0020】

ある一つのチャンネルのコヒーレント成分とは、他のチャンネルのオーディオ信号との相関が高い成分である。ある一つのチャンネルのフィールド成分とは、該チャンネルのオーディオ信号（すなわち、元の信号）と該チャンネルのコヒーレント成分との差分である。より具体的には、フィールド成分はオーディオ信号からコヒーレント成分を差し引くことで得られる成分である。コヒーレント成分は明瞭な方向性を有する音であるのに対して、フィールド成分は、拡散性を持つ、周囲を取り巻くような音（ambient sound）である。以下では、フィールド成分に対応する音を「フィールド音」ともいう。

20

【0021】

図1は、オーディオ信号処理装置10がLチャンネルのオーディオ信号をLチャンネルのコヒーレント成分 $L_c$ およびフィールド成分 $L_f$ に分割し、Rチャンネルのオーディオ信号をRチャンネルのコヒーレント成分 $R_c$ およびフィールド成分 $R_f$ に分割することを示す。コヒーレント成分 $L_c$ はRチャンネルのオーディオ信号との相関が高い成分であり、コヒーレント成分 $R_c$ はLチャンネルのオーディオ信号との相関が高い成分である。

【0022】

図1は2チャンネルのオーディオ信号の処理を示すが、オーディオ信号処理装置10は任意の個数のオーディオ信号を処理してよい。オーディオ信号処理装置10は3以上のチャンネルのオーディオ信号を処理してもよく、例えば、8Kスーパーハイビジョン用の22.2チャンネルのオーディオ信号を処理してもよい。

30

【0023】

三次元空間での音の方向、距離、広がりを再現可能な立体音響効果を実現するために、複数チャンネルのオーディオ信号は、三次元空間内に分散して配置された複数のマイクにより記録される。複数チャンネルのオーディオ信号は、複数の目的音（object sound）が互いに混ざったり目的音がフィールド音と混ざったりしたかたちで記録される。一般に音源からの距離は個々のマイクで異なるため、ある特定の音が到着する時間はマイク毎に異なり、その結果、記録されたオーディオ信号のコヒーレントが低くなる。コヒーレント成分を各チャンネルのオーディオ信号から取り出すことができれば、音の明瞭性および見かけの音源の幅（ASW：Apparent Source Width）を改善することができる。また、フィールド成分を抽出してこれをアップミックスに用いることで、良好なアンビエンス効果（聴取者の周囲を音を取り巻くような感じ）を生み出すことが可能になる。一般に、コヒーレント成分は主たる音源から発せられる目的音（例えば、歌声、楽器の音、スピーカから発せられる音など）に相当し、フィールド成分は、音の方向性が明瞭でない音（例えば、エコー、うなりなど）に相当する。

40

【0024】

N個のチャンネルのうち1番目のチャンネルのオーディオ信号を $x_1(n)$ とすると、この

50

オーディオ信号  $x_l(n)$  は  $M$  個の目的音  $q_{lm}(n)$  ( $m = 1, \dots, M$ ) とフィールド音  $v_l(n)$  とから成る。すなわち、オーディオ信号  $x_l(n)$  は式 (1) で示される。

【数 1】

$$x_l(n) = \sum_{m=1}^M q_{lm}(n) + v_l(n), \quad l=1, \dots, N \quad \dots(1)$$

【0025】

この式 (1) で示されるように、目的音とフィールド音とは互いに統計的に独立と見なすことができる。オーディオ信号  $x_l(n)$  のコヒーレント成分  $\gamma_l(n)$  は式 (2) で示される。

【数 2】

$$\gamma_l(n) = \sum_{m=1}^M q_{lm}(n) \quad \dots(2)$$

【0026】

オーディオ信号  $x_l(n)$  のフィールド成分  $\phi_l(n)$  は式 (3) で示される。

【数 3】

$$\phi_l(n) = x_l(n) - \gamma_l(n) \quad \dots(3)$$

【0027】

オーディオ信号処理装置 10 の具体的な実現方法は限定されない。例えば、オーディオ信号処理装置 10 はパーソナル・コンピュータ、サーバ、携帯端末などのコンピュータに所定のプログラム (例えば、後述するオーディオ信号処理プログラム P1) をインストールすることで実現されてもよい。あるいは、アンプなどの音響機器がオーディオ信号処理装置 10 として機能してもよい。

【0028】

図 2 は、オーディオ信号処理装置 10 として機能するコンピュータ 100 の一般的なハードウェア構成を示す。コンピュータ 100 は、オペレーティングシステムやアプリケーション・プログラムなどを実行するプロセッサ (例えば CPU) 101 と、ROM および RAM で構成される主記憶部 102 と、ハードディスクやフラッシュメモリなどで構成される補助記憶部 103 と、ネットワークカードまたは無線通信モジュールで構成される通信制御部 104 と、キーボードやマウスなどの入力装置 105 と、モニタなどの出力装置 106 とを備える。

【0029】

オーディオ信号処理装置 10 の各機能要素は、プロセッサ 101 または主記憶部 102 の上に所定のソフトウェア (例えば、後述するオーディオ信号処理プログラム P1) を読み込ませてそのソフトウェアを実行させることで実現される。プロセッサ 101 はそのソフトウェアに従って、通信制御部 104、入力装置 105、または出力装置 106 を動作させ、主記憶部 102 または補助記憶部 103 におけるデータの読み出し及び書き込みを行う。処理に必要なデータまたはデータベースは主記憶部 102 または補助記憶部 103 内に格納される。

【0030】

なお、オーディオ信号処理装置 10 は 1 台のコンピュータで構成されてもよいし、複数台のコンピュータで構成されてもよい。複数台のコンピュータを用いる場合には、これらのコンピュータがインターネットやイントラネットなどの通信ネットワークを介して接続されることで、論理的に一つのオーディオ信号処理装置 10 が構築される。

【0031】

図 3 は、オーディオ信号処理装置 10 の機能構成を示す。図 3 に示すように、オーディオ信号処理装置 10 は機能的構成要素として受付部 11、分割部 12、および出力部 13 を備える。

10

20

30

40

50

## 【 0 0 3 2 】

受付部 1 1 は、複数のチャネルのオーディオ信号を受け付ける機能要素である。「オーディオ信号を受け付ける」とは、オーディオ信号処理装置 1 0 がオーディオ信号を任意の手法で取得することである。言い換えると、「オーディオ信号を受け付ける」とは、オーディオ信号がオーディオ信号処理装置 1 0 に入力されることを意味する。各チャネルのオーディオ信号を受け付ける具体的な手法は限定されない。例えば、受付部 1 1 はデータベースまたは他の装置にアクセスしてオーディオ信号のデータファイルを読み出すことでそのオーディオ信号を受け付けてもよい。あるいは、受付部 1 1 は他の装置から通信ネットワーク経由で送られてきたオーディオ信号を受信してもよい。あるいは、受付部 1 1 はオーディオ信号処理装置 1 0 で入力されたオーディオ信号を取得してもよい。いずれにしても、受付部 1 1 は受け付けた各チャネルのオーディオ信号を分割部 1 2 に出力する。

10

## 【 0 0 3 3 】

分割部 1 2 は、各チャネルのオーディオ信号をコヒーレント成分とフィールド成分とに分割する機能要素である。以下の説明は、分割部 1 2 が式 ( 4 ) で示される N チャネルのオーディオ信号  $\{ x_1(n) \mid 1 = 1, \dots, N \}$  を処理することを前提とする。

## 【 数 4 】

$$\mathbf{x}(n) = \begin{pmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_N(n) \end{pmatrix} \dots(4)$$

20

## 【 0 0 3 4 】

まず、分割部 1 2 は各チャネルのオーディオ信号を複数の時間区間の信号に分割する。具体的には、分割部 1 2 は窓関数 (例えば、カイザー・ベッセル窓) を用いてオーディオ信号を短い時間間隔 (これを「フレーム」という) の信号に区切る。例えば、後述する変形離散コサイン変換 (MDCT) において 1 0 2 4 個の周波数点を用いるのであれば、分割部 1 2 は 2 0 4 8 点分の長さに対応するカイザー・ベッセル窓を用いてオーディオ信号を複数のフレームに分割する。通常、1 フレーム内のサンプル数は適切な周波数分解能が得られるように決められるが、そのサンプル数はコヒーレント成分を推定するには十分ではない。そこで、分割部 1 2 は連続する複数のフレーム (例えば 2 4 個のフレーム) を一つの時間区間 (これを「ブロック」という) の信号として設定する。図 4 はこのようなブロックの生成の概念を示し、より具体的には、2 チャネル (L チャネルおよび R チャネル) のオーディオ信号のそれぞれを複数のブロックに分割する処理を示す。

30

## 【 0 0 3 5 】

各チャネルのオーディオ信号を複数のブロックに分割すると、分割部 1 2 は各チャネルの各ブロックに対して以下の処理を実行する。本明細書では、オーディオ信号をコヒーレント成分とフィールド成分とに分ける対象 (すなわち、分割処理の対象) となるチャネルを「対象チャネル」という。ここでは、ある一つの対象チャネルにおける処理を説明する。

40

## 【 0 0 3 6 】

分割部 1 2 は、対象チャネルのコヒーレント成分を抽出し、その後該対象チャネルのフィールド成分を抽出する。図 5 は、その一連の処理の前半に相当する、コヒーレント成分の抽出の概念を示す。分割部 1 2 は、フィルタバンクを用いて、対象チャネルである l 番目のチャネルのオーディオ信号  $x_l(n)$  を K 個の周波数帯域 (サブバンド) の信号 (これを「サブバンド信号」という。) に分割する。そして、分割部 1 2 は各サブバンドにおいて、対象チャネル以外の他のチャネルのオーディオ信号を用いてコヒーレント成分  $x_l^{(k)}(n)$  ( $k = 1, \dots, K$ ) を抽出する。分割部 1 2 はこの抽出の際に最小二乗法を用いる。そして、分割部 1 2 は全サブバンドのコヒーレント成分を加算することで、対象チャネルのコヒーレント成分  $x_l(n)$  を抽出する。その後、分割部 1 2 は、元のオー

50



オーディオ信号  $x_1(n)$  からコヒーレント成分  $c_1(n)$  を差し引くことでフィールド成分  $f_1(n)$  を抽出する。

【0037】

分割部12は対象チャネルのオーディオ信号の各ブロックについて以下の処理を実行する。

【0038】

分割部12はフィルタバンクを用いて各チャネルのオーディオ信号  $x_1(n)$  を  $K$  個のサブバンド信号  $x_1^{(k)}(n)$  に分割する。この分割は式(5)で示される。

【数5】

$$x_l^{(k)}(n) = \begin{pmatrix} x_l^{(1)}(n) \\ x_l^{(2)}(n) \\ \vdots \\ x_l^{(K)}(n) \end{pmatrix} \quad \dots(5)$$

10

【0039】

なお、式(5)で示されるサブバンド信号  $x_1^{(k)}(n)$  は時間領域での信号であり、したがって、時間領域サブバンド信号である。周波数領域での信号を用いる上記の非特許文献1~4の手法と異なり、オーディオ信号処理装置10は時間領域サブバンド信号を用いるので、連続する任意のフレーム数の信号を一つのブロック信号として処理することで推定区間長を伸ばすことができる。この結果、得られたコヒーレント成分の音質を損なうことなく各チャネルのオーディオ信号を処理することができる。

20

【0040】

続いて、分割部12はこのサブバンド信号  $x_1^{(k)}(n)$  を、対象チャネル以外の  $N-1$  個のチャネルの同帯域(同じサブバンド)のサブバンド信号  $\{x_m^{(k)}(n) \mid m = 1, \dots, l-1, l+1, \dots, N\}$  の線形結合から推定する。ある1ブロックに対応するこの線形結合は式(6)で示される。

【数6】

$$\tilde{x}_l^{(k)}(n) = \sum_{\substack{m=1 \\ m \neq l}}^N a_m^{(k)} x_m^{(k)}(n) \quad \dots(6)$$

30

【0041】

推定信号

【数7】

$$\tilde{x}_l^{(k)}(n)$$

は、他チャネル(対象チャネル以外の  $N-1$  個のチャネル)の同帯域の信号との相関が高い成分と考えることができる。対象チャネルのサブバンド信号とこの推定信号との推定誤差  $e_1^{(k)}(n)$  は式(7)で示される。

【数8】

$$e_l^{(k)}(n) = x_l^{(k)}(n) - \tilde{x}_l^{(k)}(n) \quad \dots(7)$$

40

【0042】

分割部12は、この推定誤差を最小にする係数  $\{a_m^{(k)} \mid m = 1, \dots, l-1, l+1, \dots, N\}$  を最小二乗法で求める。最小化すべき誤差関数は式(8)で示される。

【数 9】

$$J_l^{(k)} = \sum_{n \in T} (e_l^{(k)}(n))^2 = \sum_{n \in T} \left( x_l^{(k)}(n) - \sum_{\substack{m=1 \\ m \neq l}}^N a_m^{(k)} x_m^{(k)}(n) \right)^2 \quad \dots(8)$$

 $(l=1, \dots, N, \quad k=1, \dots, K)$ 

【0043】

ここで、

【数10】

$$X^{(k)}(l, m) = \sum_{n \in T} x_l^{(k)}(n) x_m^{(k)}(n)$$

とすると、最適な係数群

【数11】

$$\{\hat{a}_l^{(k)} | m=1, \dots, l-1, l+1, \dots, N\}$$

は式(9)を満たす。

【数12】

$$\left( X^{(k)}(1, m) \quad \dots \quad X^{(k)}(l-1, m) \quad X^{(k)}(l+1, m) \quad \dots \quad X^{(k)}(N, m) \right) \times \begin{pmatrix} \hat{a}_1^{(k)} \\ \vdots \\ \hat{a}_{l-1}^{(k)} \\ \hat{a}_{l+1}^{(k)} \\ \vdots \\ \hat{a}_N^{(k)} \end{pmatrix} = X^{(k)}(l, m) \quad \dots(9)$$

【0044】

この式(9)を  $m=1, \dots, l-1, l+1, \dots, N$  で連立させると式(10)が得られる。

【数13】

$$M^{(k)} \hat{\mathbf{a}}_l^{(k)} = \mathbf{v}^{(k)} \quad \dots(10)$$

ここで、

【数14】

$$M^{(k)} = \begin{pmatrix} X^{(k)}(1,1) & \dots & X^{(k)}(l-1,1) & X^{(k)}(l+1,1) & \dots & X^{(k)}(N,1) \\ \vdots & & \vdots & \vdots & & \vdots \\ X^{(k)}(1,l-1) & \dots & X^{(k)}(l-1,l-1) & X^{(k)}(l+1,l-1) & \dots & X^{(k)}(N,l-1) \\ X^{(k)}(1,l+1) & \dots & X^{(k)}(l-1,l+1) & X^{(k)}(l+1,l+1) & \dots & X^{(k)}(N,l+1) \\ \vdots & & \vdots & \vdots & & \vdots \\ X^{(k)}(1,N) & \dots & X^{(k)}(l-1,N) & X^{(k)}(l+1,N) & \dots & X^{(k)}(N,N) \end{pmatrix},$$

10

20

30

40

【数 1 5】

$$\hat{\mathbf{a}}_l^{(k)} = \begin{pmatrix} \hat{a}_1^{(k)} \\ \vdots \\ \hat{a}_{l-1}^{(k)} \\ \hat{a}_{l+1}^{(k)} \\ \vdots \\ \hat{a}_N^{(k)} \end{pmatrix},$$

【数 1 6】

$$\mathbf{v}^{(k)} = \begin{pmatrix} X^{(k)}(l,1) \\ \vdots \\ X^{(k)}(l,l-1) \\ X^{(k)}(l,l+1) \\ \vdots \\ X^{(k)}(l,N) \end{pmatrix}$$

10

である。

【0045】

20

k 番目のサブバンドにおける対象チャネルの係数ベクトル  $\hat{\mathbf{a}}_l^{(k)}$  は式 (11) により得られる。

【数 1 7】

$$\hat{\mathbf{a}}_l^{(k)} = (M^{(k)})^{-1} \mathbf{v}^{(k)} \quad \dots(11)$$

【0046】

k 番目のサブバンドにおける対象チャネルのコヒーレント成分  $\gamma_l^{(k)}(n)$  は式 (12) により得られる。このコヒーレント成分  $\gamma_l^{(k)}(n)$  は、対象チャネル以外のチャネルのオーディオ信号を用いて算出される推定信号のうち該対象チャネルのオーディオ信号との相関が最も高い推定信号に相当する。

30

【数 1 8】

$$\gamma_l^{(k)}(n) = \sum_{\substack{m=1 \\ m \neq l}}^N \hat{a}_m^{(k)} x_m^{(k)}(n) \quad \dots(12)$$

【0047】

分割部 12 はすべてのサブバンドについてコヒーレント成分を求める。そして、分割部 12 は全サブバンドのコヒーレント成分を加算することで対象チャネルのコヒーレント成分を求める。この処理は式 (13) で示される。

【数 1 9】

$$\gamma_l(n) = \sum_{k=1}^K \gamma_l^{(k)}(n) \quad \dots(13)$$

40

【0048】

さらに、分割部 12 は対象チャネルの元のオーディオ信号からそのコヒーレント成分を差し引くことで、対象チャネルのフィールド成分を求める。この処理は上記式 (3) で示される。

【0049】

なお、分割部 12 は、各サブバンドにおいてオーディオ信号からコヒーレント成分を差し引くことでフィールド成分を求め、全サブバンドのフィールド成分を加算することで対象チャネルのフィールド成分を求めてもよい。具体的には、k 番目のサブバンドにおける

50

対象チャンネルのフィールド成分  $\phi_l^{(k)}(n)$  は式 (14) により得られる。

【数 2 0】

$$\phi_l^{(k)}(n) = x_l^{(k)}(n) - \gamma_l^{(k)}(n) = x_l^{(k)}(n) - \sum_{\substack{m=1 \\ m \neq l}}^N \hat{a}_m^{(k)} x_m^{(k)}(n) \quad \dots(14)$$

対象チャンネルのフィールド成分  $\phi_l(n)$  は式 (15) により得られる。

【数 2 1】

$$\phi_l(n) = \sum_{k=1}^K \phi_l^{(k)}(n) \quad \dots(15)$$

【0050】

分割部 12 は上記の処理を対象チャンネルのオーディオ信号の各ブロックに対して実行する。そして、分割部 12 は全ブロックのコヒーレント成分を連結することで対象チャンネルのコヒーレント成分を抽出する。また、分割部 12 は全ブロックのフィールド成分を連結することで対象チャンネルのフィールド成分を生成する。

【0051】

分割部 12 は複数のチャンネルのそれぞれを対象チャンネルとして設定して上記の処理を実行することで、全チャンネルについてコヒーレント成分およびフィールド成分を生成する。そして、分割部 12 は全チャンネルのコヒーレント成分およびフィールド成分を出力部 13 に出力する。

【0052】

このように、分割部 12 は各チャンネルのオーディオ信号に別の信号を追加することなく（すなわち、原音に別の音を追加することなく）、各チャンネルのオーディオ信号をコヒーレント成分とフィールド成分とに分割する。

【0053】

出力部 13 は、分割部 12 により生成された各チャンネルのコヒーレント成分およびフィールド成分を処理結果として出力する機能要素である。この処理結果は、Nチャンネルから2Nチャンネルへのアップミックスを実現したものであるということが出来る。処理結果の出力方法は何ら限定されない。例えば、出力部 13 は処理結果をメモリやデータベースなどの記憶装置に格納してもよいし、通信ネットワークを介して他の装置に送信してもよい。あるいは、出力部 13 は各チャンネルのコヒーレント成分およびフィールド成分を対応するスピーカに出力してもよい。いずれにしても、オーディオ信号処理装置 10 による処理結果を用いて、既存の音声素材を、より多くのチャンネル数を持つコンテンツの制作に利用したり、より多くのチャンネルを有するオーディオ・システムで再生したりすることが可能になる。

【0054】

オーディオ信号処理装置 10 は、Nチャンネルのオーディオ信号を2Nより大きい数のチャンネルにアップミックスしてもよい。具体的には、オーディオ信号処理装置 10 は、抽出した複数のフィールド成分を下記参考文献に記載の手法で無相関化することで、チャンネル間の相関が互いに異なる信号を生成する。これにより、Nより多い個数のフィールド成分が得られる。例えば、ステレオの音声素材を5.1チャンネルの音声素材に変換したり、5.1チャンネルのオーディオ・システムを用いてより高い臨場感で再生したりすることができる。あるいは、5.1チャンネルの音声素材を2.2チャンネルの音声素材に変換したり、2.2チャンネルのオーディオ・システムを用いてより高い臨場感で再生したりすることができる。

（参考文献）J. Breebaart and C. Fallar, "Spatial Audio Processing - MPEG Surround and Other Applications," Wiley, 2007.

【0055】

オーディオ信号処理装置 10 は、Nチャンネルのオーディオ信号を、2Nより小さいJ個のオーディオ信号（ただし、J > N）のオーディオ信号にアップミックスしてもよい。具

10

20

30

40

50

体的には、オーディオ信号処理装置 10 は N 個のフィールド成分をミキシングすることで、N チャンネルから J チャンネルへのアップミックスを実現する。

【0056】

オーディオ信号処理装置 10 による処理結果はアップミックスだけでなくダウンミックスにも利用可能である。

【0057】

次に、図 6 および図 7 を参照しながら、オーディオ信号処理装置 10 の動作を説明するとともに本実施形態に係るオーディオ信号処理方法について説明する。オーディオ信号処理装置 10 では、まず、受付部 11 が複数のチャンネルのオーディオ信号を受け付ける（受付ステップ）。続いて、分割部 12 がオーディオ信号をコヒーレント成分とフィールド成分とに分割する分割処理を各チャンネルについて実行する（分割ステップ）。そして、出力部 13 が各チャンネルのコヒーレント成分およびフィールド成分を出力する（出力ステップ）。以下では、特に重要な分割部 12 の処理（分割ステップ）について詳しく説明する。

10

【0058】

図 6 は、一つの対象チャンネルのコヒーレント成分およびフィールド成分を生成する処理を示す。

【0059】

まず、分割部 12 は各チャンネルのオーディオ信号を複数のブロックに分割する（ステップ S 11）。なお、ステップ S 11 において分割した各チャンネルおよび各ブロックのオーディオ信号を保存することで、2 番目以降の対象チャンネルを処理するにはステップ S 11 を省略することができる。

20

【0060】

続いて、分割部 12 は対象チャンネルの複数のブロックのうちの一つを処理対象として設定する（ステップ S 12）。続いて、分割部 12 は、対象チャンネル以外のチャンネルのオーディオ信号を用いて算出される推定信号のうち、対象チャンネルのオーディオ信号との相関が最も高い推定信号を、対象チャンネルのコヒーレント成分として抽出する（ステップ S 13）。続いて、分割部 12 は、対象チャンネルのオーディオ信号とそのコヒーレント成分との差分を、対象チャンネルのフィールド成分として抽出する（ステップ S 14）。このような処理により、分割部 12 は対象チャンネルの 1 ブロックのコヒーレント成分およびフィールド成分を得る。

30

【0061】

分割部 12 は一つのブロックを処理すると次のブロックの処理に移る（ステップ S 15 参照）。すなわち、分割部 12 は次のブロックを処理対象として設定し（ステップ S 12）、そのブロックのコヒーレント成分およびフィールド成分を生成する（ステップ S 13 および S 14）。分割部 12 はすべてのブロックについてステップ S 12 ~ S 14 の処理を実行し、全ブロックのコヒーレント成分およびフィールド成分を生成する（ステップ S 15 において YES）。そして、分割部 12 は全ブロックのコヒーレント成分を連結することで対象チャンネルの最終的なコヒーレント成分を得ると共に、全ブロックのフィールド成分を連結することで対象チャンネルの最終的なフィールド成分を得る。

【0062】

図 7 は、図 6 におけるステップ S 13 の処理の詳細、すなわち、対象チャンネルのコヒーレント成分を生成する処理の詳細を示す。図 7 に示す処理は対象チャンネルのオーディオ信号の各ブロックについて実行される。

40

【0063】

まず、分割部 12 は各チャンネル（対象チャンネルおよびすべての他チャンネル）について、ブロック信号を複数のサブバンドに分割することで複数のサブバンド信号を生成する（ステップ S 131）。続いて、分割部 12 は複数のサブバンドのうちの一つを処理対象として設定する（ステップ S 132）。続いて、分割部 12 は、対象チャンネル以外のチャンネルのサブバンド信号を用いて算出される推定信号のうち、対象チャンネルのサブバンド信号との相関が最も高い推定信号を、処理対象であるサブバンドにおける対象チャンネルのコヒー

50

レント成分として抽出する（ステップS 1 3 3）。分割部 1 2 はすべてのサブバンドについてステップS 1 3 2 およびS 1 3 3 の処理を実行する（ステップS 1 3 4 参照）。対象チャンネルについて全サブバンドのコヒーレント成分を生成すると（ステップS 1 3 4 においてYES）、分割部 1 2 はそれらのコヒーレント成分を加算することで対象チャンネルのコヒーレント成分（より具体的には、1 ブロック分のコヒーレント成分）を生成する（ステップS 1 3 5）。

【0064】

次に、図8を参照しながら、コンピュータをオーディオ信号処理装置10として機能させるためのオーディオ信号処理プログラムP1を説明する。

【0065】

オーディオ信号処理プログラムP1はメインモジュールP10、受付モジュールP11、分割モジュールP12、および出力モジュールP13を含む。メインモジュールP10は、オーディオ信号の処理を統括的に実行する部分である。受付モジュールP11、分割モジュールP12、および出力モジュールP13を実行することにより実現される機能はそれぞれ、上記の受付部11、分割部12、および出力部13の機能と同様である。

【0066】

オーディオ信号処理プログラムP1は、例えば、CD-ROMやDVD-ROM、半導体メモリなどの有形の記録媒体に固定的に記録された上で提供されてもよい。あるいは、オーディオ信号処理プログラムP1は、搬送波に重畳されたデータ信号として通信ネットワークを介して提供されてもよい。

【0067】

以上説明したように、本発明の一側面に係るオーディオ信号処理装置は、複数のチャンネルのオーディオ信号を受け付ける受付部と、オーディオ信号をコヒーレント成分とフィールド成分とに分割する分割処理を各チャンネルについて実行する分割部であって、分割処理が、分割処理の対象となる一つのチャンネルを対象チャンネルとした場合に、該対象チャンネル以外のチャンネルのオーディオ信号を少なくとも用いて算出される推定信号のうち該対象チャンネルのオーディオ信号との相関が最も高い推定信号を該対象チャンネルのコヒーレント成分として抽出するステップと、対象チャンネルのオーディオ信号と該対象チャンネルのコヒーレント成分との差分を該対象チャンネルのフィールド成分として抽出するステップとを含む、該分割部と、分割部により抽出された各チャンネルのコヒーレント成分およびフィールド成分を出力する出力部とを備える。

【0068】

本発明の一側面に係るオーディオ信号処理方法は、オーディオ信号処理装置が、複数のチャンネルのオーディオ信号を受け付ける受付ステップと、オーディオ信号処理装置が、オーディオ信号をコヒーレント成分とフィールド成分とに分割する分割処理を各チャンネルについて実行する分割ステップであって、分割処理が、分割処理の対象となる一つのチャンネルを対象チャンネルとした場合に、該対象チャンネル以外のチャンネルのオーディオ信号を少なくとも用いて算出される推定信号のうち該対象チャンネルのオーディオ信号との相関が最も高い推定信号を該対象チャンネルのコヒーレント成分として抽出するステップと、対象チャンネルのオーディオ信号と該対象チャンネルのコヒーレント成分との差分を該対象チャンネルのフィールド成分として抽出するステップとを含む、該分割ステップと、オーディオ信号処理装置が、分割ステップにおいて抽出された各チャンネルのコヒーレント成分およびフィールド成分を出力する出力ステップとを含む。

【0069】

本発明の一側面に係るオーディオ信号処理プログラムは、複数のチャンネルのオーディオ信号を受け付ける受付ステップと、オーディオ信号をコヒーレント成分とフィールド成分とに分割する分割処理を各チャンネルについて実行する分割ステップであって、分割処理が、分割処理の対象となる一つのチャンネルを対象チャンネルとした場合に、該対象チャンネル以外のチャンネルのオーディオ信号を少なくとも用いて算出される推定信号のうち該対象チャンネルのオーディオ信号との相関が最も高い推定信号を該対象チャンネルのコヒーレント成分

10

20

30

40

50

として抽出するステップと、対象チャネルのオーディオ信号と該対象チャネルのコヒーレント成分との差分を該対象チャネルのフィールド成分として抽出するステップとを含む、該分割ステップと、分割ステップにおいて抽出された各チャネルのコヒーレント成分およびフィールド成分を出力する出力ステップとをコンピュータに実行させる。

【0070】

このような側面においては、対象チャネル以外のチャネルのオーディオ信号を用いて推定され、且つ該対象チャネルの実際のオーディオ信号ごとの相関が最も高い信号が該対象チャネルのコヒーレント成分として抽出される。また、対象チャネルの実際のオーディオ信号とそのコヒーレント成分との差分が該対象チャネルのフィールド成分として抽出される。このコヒーレント成分およびフィールド成分は各チャネルについて得られる。このように、音を追加することなく元のオーディオ信号のみを用いて各チャネルのコヒーレント成分およびフィールド成分を求めることで、原音の雰囲気（例えば本来の音色）を可能な限りまたは完全に維持することができる。加えて、コヒーレント成分およびフィールド成分は元のチャネル数の分だけ求めることができるので、この手法は原音のチャネル数にかかわらず適用できる。例えば、本発明の一側面は2チャネル、3チャネル、5.1チャネル、2.2.2チャネルなどの任意のチャネル数のオーディオ信号に対して適用できる。

【0071】

図9および図10を用いて上記側面の優位性を説明する。図9は従来の手法におけるコヒーレント成分の抽出の例を示す図であり、図10は上記側面におけるコヒーレント成分の抽出の例を示す図である。図9、10共に、三角形状に配置された三つのスピーカ90からオーディオ信号が出力される例を示し、したがって、この例は3チャネルのオーディオ・システムを示す。

【0072】

図9に示すように、上記の非特許文献3、4に記載の手法では、2チャネルのオーディオ信号の間で相関が高い成分をコヒーレント成分91として抽出する（なお、破線92はフィールド成分を示す）。したがって、このような従来の手法では、二つのスピーカ（チャネル）90の中間部分93に位置する音の情報しか取得することができず、三つのスピーカ（チャネル）90で囲まれた領域の中央部分94に位置する音の情報を抽出することができない。

【0073】

これに対して上記側面では、あるスピーカ（チャネル）90のコヒーレント成分が他のスピーカ（チャネル）90の信号から推定される。そのため、図10に示すように、三つのスピーカ（チャネル）90で囲まれた領域の中央部分95に位置する音の情報を抽出することができる。この中央部分95は、図9における部分93、94の和に相当し得る。

【0074】

他の側面に係るオーディオ信号処理装置では、分割処理が、窓関数を用いてオーディオ信号を複数のフレームに区切る処理を各チャネルについて実行するステップと、連続する少なくとも二つのフレームを一つのブロックにまとめる処理を複数のフレームの全体に対して実行することで複数のブロックを生成する処理を各チャネルについて実行するステップと、ブロックのそれぞれにおいて対象チャネルのコヒーレント成分を抽出するステップとを含んでもよい。

【0075】

複数のフレームで構成されるブロックを採用することで、コヒーレント成分の推定のためのサンプル数が多くなるので、コヒーレント成分をより精度良く抽出することが可能になる。

【0076】

他の側面に係るオーディオ信号処理装置では、分割部が、各チャネルのオーディオ信号を複数のサブバンドに分割することで、各チャネルについて複数のサブバンド信号を生成するステップと、複数のサブバンドのそれぞれにおいて対象チャネルのコヒーレント成分を抽出するステップと、複数のサブバンドにおけるコヒーレント成分を加算することで対

10

20

30

40

50

象チャンネルのコヒーレント成分を抽出するステップとを含んでもよい。

【 0 0 7 7 】

一般に、音声処理では一部の周波数が他の周波数よりも重要であることが多い。サブバンド毎に処理することで、それぞれの周波数帯で要求される精度に応じてコヒーレント成分を抽出することができ、ひいてはコヒーレント成分およびフィールド成分を精度良く抽出することができる。

【実施例】

【 0 0 7 8 】

以下、実施例に基づいて本発明を具体的に説明するが、本発明はそれらに何ら限定されるものではない。

【 0 0 7 9 】

表 1 に示される 7 個のステレオ音声素材（すなわち、2 チャンネルのオーディオ信号）を用意した。いずれの音声素材も市販の CD から入手したものであり、サンプリング周波数は 44.1 kHz であった。表 1 の名前欄は曲名または楽曲の種類を示し、説明欄は演奏の形態を示す。ミキシング欄における「Artificial」はミキシング処理が施された素材であることを示し、「Natural」はミキシング処理が施されていない素材であることを示す。長さ欄は再生時間を示す。

【表 1】

名前	説明	ミキシング	長さ
Over the rainbow	ボーカルと楽器	Artificial	6:40
Quiet night	ジャズ三重奏	Artificial	2:53
Zueignung	ボーカルとピアノ	Natural	1:52
Symphony no.3	オーケストラ	Natural	0:57
Capriccio	バイオリンのソロ	Natural	4:28
Impromptus	ピアノのソロ	Natural	7:58
O mio babbino caro	ボーカルとオーケストラ	Natural	2:38

【 0 0 8 0 】

オーディオ信号を完全に再構築できるフィルタバンクを構築するために、変形離散コサイン変換 (MDCT) を用いた重畳加算法を採用した。オーディオ信号を複数のフレームに分割するための窓関数としてカイザー・ベッセル窓を用いた。フレーム長は 2048 点とし、これは、MDCT において 1024 個の周波数点を得られることを意味する。その周波数点を表 2 に示すように 23 個のサブバンドにまとめた。これらのサブバンドは、MP EG - 2 AAC 標準を参考に、48 kHz long FFT (高速フーリエ変換) における 69 個のサブバンドを三つの連続するサブバンド毎の一つにまとめたものである。24 個のフレームを 1 ブロックとした。サンプリング周波数が 44.1 kHz であれば、ブロック長は 0.58 秒に相当するものであった。



【表 2】

バンド番号	開始周波数番号	終了周波数番号
1	1	6
2	7	12
3	13	18
4	19	24
5	25	30
6	31	36
7	37	44
8	45	53
9	54	62
10	63	74
11	75	87
12	88	103
13	104	123
14	124	148
15	149	179
16	180	218
17	219	266
18	267	325
19	326	398
20	399	491
21	492	613
22	614	782
23	783	1024

10

20

30

## 【 0 0 8 1 】

実験結果をチャンネル間の相互相関係数で評価した。原音、コヒーレント成分、およびフィールド成分の相互相関係数を表 3 に示す。コヒーレント成分は原音よりも高い相互相関を示した。このようなコヒーレント成分は原音よりも狭い音場の雰囲気をもたらす。一方、フィールド成分は、一個の素材（“ Quiet Night ”）を除いて負の相互相関を示した。負の相互相関を示すフィールド成分を側方もしくは後方に設置したスピーカで再生すれば、良好なアンビエンス効果が得られる。その結果として、臨場感の高い音を再生することができる。

40

【表 3】

名前	原音	コヒーレント成分	フィールド成分
Over the rainbow	0.7521	0.9101	-0.3998
Quiet night	0.2386	0.7983	0.0691
Zueignung	0.5474	0.6484	-0.4938
Symphony no.3	0.4172	0.4745	-0.3865
Capriccio	0.7685	0.8620	-0.5840
Impromptus	0.3581	0.6012	-0.2674
O mio babbino caro	0.7002	0.7657	-0.6053

10

## 【0082】

以上、本発明をその実施形態に基づいて詳細に説明した。しかし、本発明は上記実施形態に限定されるものではない。本発明は、その要旨を逸脱しない範囲で様々な変形が可能である。

## 【0083】

上記実施形態では、分割部12が、ある一つの対象チャンネルのコヒーレント成分を、該対象チャンネル以外のチャンネルのオーディオ信号を用いて推定した。この変形例として、分割部は、当該他チャンネルのオーディオ信号と、対象チャンネルの過去のオーディオ信号および当該他チャンネルの過去のオーディオ信号の少なくとも一方とを用いて、該対象チャンネルのコヒーレント成分を推定してもよい。ここで、「過去のオーディオ信号」とは、処理対象のブロックより時間的に前のブロックのオーディオ信号である。対象チャンネルおよび他チャンネルのうち的一方または双方の過去のオーディオ信号も用いて、処理対象のブロックにおける対象チャンネルのオーディオ信号を推定することで、コヒーレント成分をより精度良く抽出することが期待できる。

20

## 【0084】

少なくとも一つのプロセッサにより実行されるオーディオ信号処理方法の手順は上記実施形態での例に限定されない。例えば、オーディオ信号処理装置は上述したステップ（処理）の一部を省略してもよいし、別の順序で各ステップを実行してもよい。また、上述したステップのうち任意の2以上のステップが組み合わされてもよいし、ステップの一部が修正又は削除されてもよい。あるいは、オーディオ信号処理装置は上記の各ステップに加えて他のステップを実行してもよい。

30

## 【0085】

オーディオ信号処理装置は、二つの数値の大小関係を比較する際に、「以上」および「よりも大きい」という二つの基準のどちらを用いてもよく、「以下」および「未満」の二つの基準のうちどちらを用いてもよい。このような基準の選択は、二つの数値の大小関係を比較する処理についての技術的意義を変更するものではない。

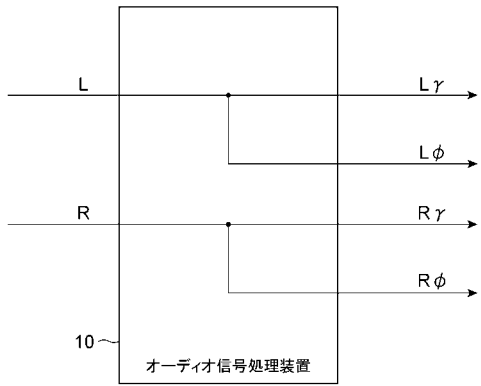
## 【符号の説明】

40

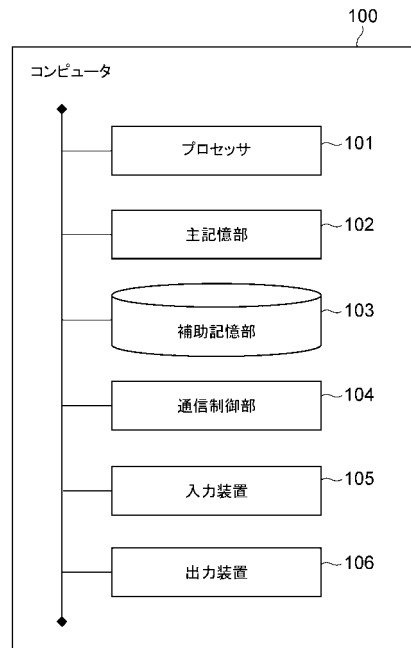
## 【0086】

10 ... オーディオ信号処理装置、11 ... 受付部、12 ... 分割部、13 ... 出力部、e1 ... 推定誤差、P1 ... オーディオ信号処理プログラム、P10 ... メインモジュール、P11 ... 受付モジュール、P12 ... 分割モジュール、P13 ... 出力モジュール。

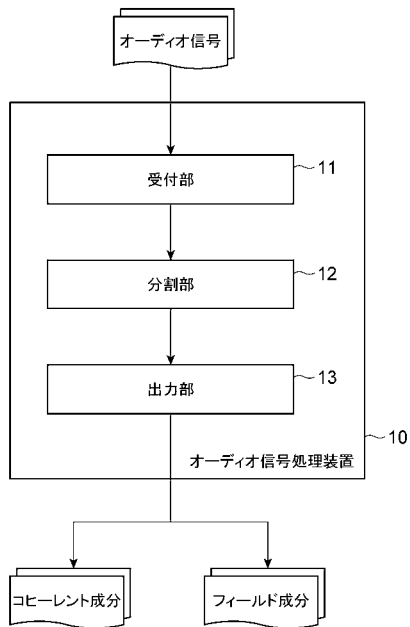
【 図 1 】



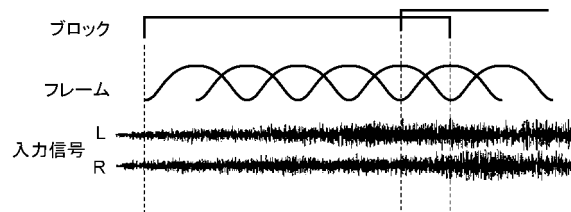
【 図 2 】



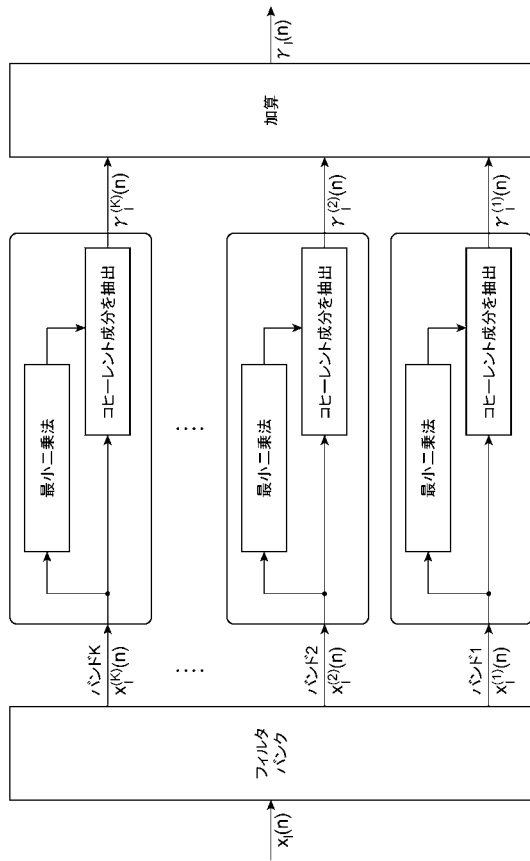
【 図 3 】



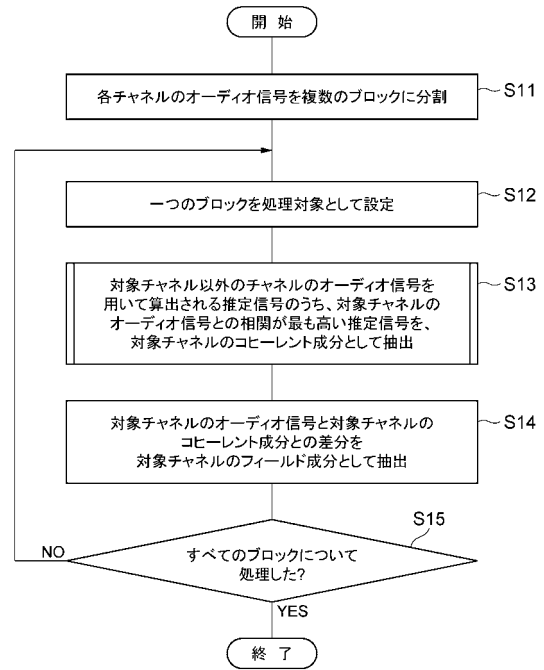
【 図 4 】



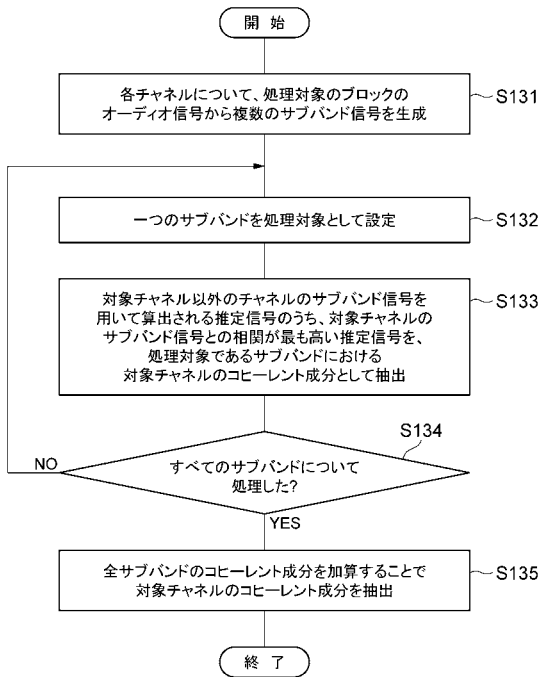
【 図 5 】



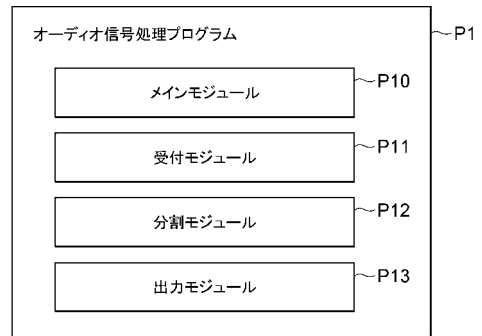
【 図 6 】



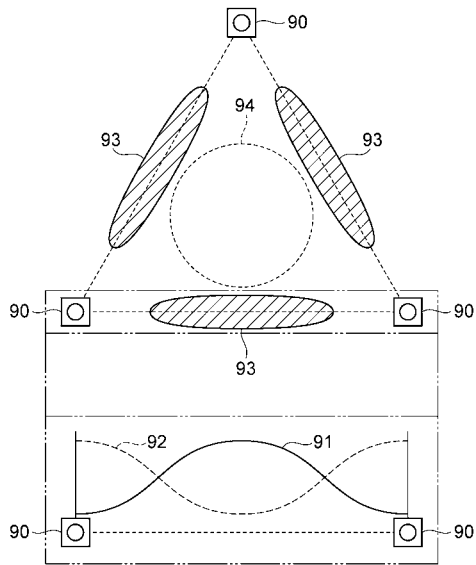
【 図 7 】



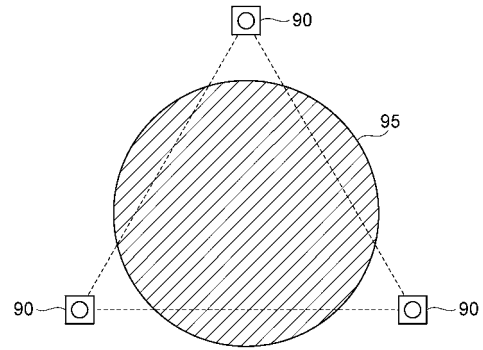
【 図 8 】



【 図 9 】



【 図 10 】



## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/JP2017/016019
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> H04S7/00(2006.01)i, G10L19/00(2013.01)i  According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) H04S7/00, G10L19/00  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2017 Kokai Jitsuyo Shinan Koho 1971-2017 Toroku Jitsuyo Shinan Koho 1994-2017  Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2016-501472 A (Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.), 18 January 2016 (18.01.2016), entire text; all drawings & US 2015/0248891 A1 entire text; all drawings & WO 2014/076030 A1	1-5
A	JP 2013-517518 A (Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.), 16 May 2013 (16.05.2013), entire text; all drawings & US 2012/0314876 A1 entire text; all drawings & WO 2011/086060 A1	1-5
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 31 May 2017 (31.05.17)		Date of mailing of the international search report 13 June 2017 (13.06.17)
Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan		Authorized officer  Telephone No.

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/JP2017/016019

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2008-536183 A (Coding Technologies AB.), 04 September 2008 (04.09.2008), entire text; all drawings & US 2006/0239473 A1 entire text; all drawings & WO 2006/108543 A1	1-5

国際調査報告		国際出願番号 PCT/J P 2 0 1 7 / 0 1 6 0 1 9	
A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int.Cl. H04S7/00(2006.01)i, G10L19/00(2013.01)i			
B. 調査を行った分野 調査を行った最小限資料 (国際特許分類 (IPC)) Int.Cl. H04S7/00, G10L19/00			
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2017年 日本国実用新案登録公報 1996-2017年 日本国登録実用新案公報 1994-2017年			
国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)			
C. 関連すると認められる文献			
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号	
A	JP 2016-501472 A (フラウンホーファーゲゼルシャフト・ツール・フェルデルング・デル・アンゲヴァンテン・フォルシュング・アイングトラーゲネル・フェライン) 2016.01.18, 全文, 全図 & US 2015/0248891 A1, 全文, 全図 & WO 2014/076030 A1	1-5	
A	JP 2013-517518 A (フラウンホッフアーゲーゼルシャフト ツァ フェルダールング デア アンゲヴァンテン フォアシュンク エー. ファオ) 2013.05.16, 全文, 全図	1-5	
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。			
* 引用文献のカテゴリー		の日の後に公表された文献	
「A」 特に関連のある文献ではなく、一般的な技術水準を示すもの		「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの	
「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの		「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの	
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)		「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの	
「O」 口頭による開示、使用、展示等に言及する文献		「&」 同一パテントファミリー文献	
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願			
国際調査を完了した日 31.05.2017		国際調査報告の発送日 13.06.2017	
国際調査機関の名称及びあて先 日本国特許庁 (ISA/J P) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号		特許庁審査官 (権限のある職員) 菊池 充	5Z 4545
		電話番号 03-3581-1101 内線 3591	



国際調査報告		国際出願番号 PCT/J P 2017/016019
C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	& US 2012/0314876 A1, 全文, 全図 & WO 2011/086060 A1  JP 2008-536183 A (コーディング テクノロジーズ アクチボラゲ ット) 2008.09.04, 全文, 全図 & US 2006/0239473 A1, 全文, 全図 & WO 2006/108543 A1	1-5

## フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ

(注) この公表は、国際事務局(WIPO)により国際公開された公報を基に作成したものである。なおこの公表に係る日本語特許出願(日本語実用新案登録出願)の国際公開の効果は、特許法第184条の10第1項(実用新案法第48条の13第2項)により生ずるものであり、本掲載とは関係ありません。