

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6161962号  
(P6161962)

(45) 発行日 平成29年7月12日(2017.7.12)

(24) 登録日 平成29年6月23日(2017.6.23)

(51) Int.Cl.	F I
<b>H04S 7/00 (2006.01)</b>	H04S 7/00 400
	H04S 7/00 300

請求項の数 5 (全 36 頁)

(21) 出願番号	特願2013-119366 (P2013-119366)	(73) 特許権者	000005049
(22) 出願日	平成25年6月6日(2013.6.6)		シャープ株式会社
(65) 公開番号	特開2014-239269 (P2014-239269A)		大阪府堺市堺区匠町1番地
(43) 公開日	平成26年12月18日(2014.12.18)	(74) 代理人	110000338
審査請求日	平成28年6月6日(2016.6.6)		特許業務法人HARAKENZO WORLD PATENT & TRADEMARK
		(72) 発明者	佐藤 純生
			大阪府大阪市阿倍野区長池町22番22号
			シャープ株式会社内
		(72) 発明者	服部 永雄
			大阪府大阪市阿倍野区長池町22番22号
			シャープ株式会社内
		審査官	北原 昂

最終頁に続く

(54) 【発明の名称】 音声信号再生装置及び方法

(57) 【特許請求の範囲】

【請求項1】

2つの音声信号を、周波数毎または周波数帯域毎の相関信号に分離する分離部と、  
該分離部で分離された各相関信号について、音像方向及び信号電力値を算出する算出部と、

各相関信号について前記音像方向に基づき予め定められた複数の音像方向グループに割り当てられた前記信号電力値に応じて、前記複数の音像方向グループのそれぞれに対応する表示を行う音像表示部と、

を備えたことを特徴とする音声信号再生装置。

【請求項2】

前記音像表示部は、棒グラフ形式で表示することを特徴とする請求項1に記載の音声信号再生装置。

【請求項3】

前記音像表示部は、複数の発光部を有し、前記割り当てられた信号電力値に応じて前記発光部の発光色を変化させることを特徴とする請求項1または2に記載の音声信号再生装置。

【請求項4】

前記音像表示部は、複数の発光部を有し、前記割り当てられた信号電力値に応じて前記発光部の発光強度を変化させることを特徴とする請求項1または2に記載の音声信号再生装置。

10

20

## 【請求項 5】

分離部が、2つの音声信号を、周波数毎または周波数帯域毎の相関信号に分離する分離ステップと、

算出部が、前記分離ステップで分離された各相関信号について、音像方向及び信号電力値を算出する算出ステップと、

音像表示部が、各相関信号について前記音像方向に基づき予め定められた複数の音像方向グループに割り当てられた前記信号電力値に応じて、前記複数の音像方向グループのそれぞれに対応する表示を行う音像表示ステップと、

を含む事の特徴とする音声信号再生方法。

## 【発明の詳細な説明】

10

## 【技術分野】

## 【0001】

本発明は、音声信号を再生する音声信号再生装置及び方法に関する。

## 【背景技術】

## 【0002】

従来から提案されている音響再生方式には、ステレオ(2ch)方式、5.1chサラウンド方式(ITU-R BS.775-1)、7.1ch、9.1ch、22.2chなどのマルチチャンネル再生方式や、音源オブジェクト指向再生方式がある。後者の方式は、全ての音が、いずれかの音源オブジェクトが発する音であるとする方式であり、各音源オブジェクト(以下、「仮想音源」と呼ぶ。)が自身の位置情報と音声信号とを含んでいる。音楽コンテンツを例にとると、各仮想音源は、それぞれの楽器の音と楽器が配置されている位置情報とを含む。

20

## 【0003】

音源オブジェクト指向再生方式は、通常、直線状あるいは面状に並べたスピーカ群によって音の波面を合成する再生方式(すなわち波面合成再生方式)により再生される。このような波面合成再生方式のうち、非特許文献1に記載のWave Field Synthesis(WFS)方式は、直線状あるいは曲線上に並べたスピーカ群(以下、スピーカアレイという)を用いる現実的な実装方法の1つとして近年盛んに研究されている。

## 【0004】

このような波面合成再生方式は、スイートスポットが狭い上述のマルチチャンネル再生方式とは異なり、並べられたスピーカ群の前のどの位置で聴いている受聴者に対しても、良好な音像と音質を両方同時に提示することができるという特長を持つ。つまり、波面合成再生方式でのスイートスポットは幅広くなっている。

30

## 【0005】

特許文献1には、通常、映画コンテンツなどに使用されている5.1chの音声信号を複数のチャンネルの音声信号に変換して波面合成再生方式で再生する再生方法が提案されている。具体的には、左フロントチャンネル信号、右フロントチャンネル信号、センターチャンネル信号、左リアチャンネル信号、右リアチャンネル信号、サブウーファーチャンネル信号のうち、左フロントチャンネル信号と右フロントチャンネル信号とに着目して相関信号成分と無相関信号成分とに分離し、相関信号成分を複数の仮想音源に割り当てた上で、中心の仮想音源にセンターチャンネル信号を重畳することにより、5.1chの音声信号を波面合成再生方式で再生している。

40

## 【0006】

ところで、従来から、音響システムやオーディオソフトウェアにおいて、音楽に連動させ視覚的な表示を変化させ、没入感を増加させるような技術が様々提案されている。例えば、ある一定の周波数帯域毎に、音圧のパワーの絶対値を棒グラフとして表示させる方式や、音の時間的な変化によって色や図形をランダムに変化させて表示させる方式などが挙げられる。

## 【0007】

その他、音圧レベルの表示を音像位置と連動させ、空間的な音像位置を視覚化する方式

50

も提案されている（例えば、特許文献 2 を参照）。特許文献 2 に記載の方式では、左右の音声信号のレベル差の絶対値のみで音像位置を判断している。よって、左右の音声信号が振幅のみ異なる場合、もしくは振幅も同一である場合には、知覚上の音像位置と表示位置が一致することになる。

【先行技術文献】

【特許文献】

【0008】

【特許文献 1】特許第 4 8 1 0 6 2 1 号公報

【特許文献 2】実開昭 6 0 - 1 7 7 6 0 0 号公報

【非特許文献】

10

【0009】

【非特許文献 1】A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis", J. Acoust. Soc. Am. Volume 93(5), アメリカ合衆国, Acoustical Society of America, May 1993, pp. 2764-2778

【発明の概要】

【発明が解決しようとする課題】

【0010】

しかしながら、特許文献 2 に記載の方式では、左右の音声信号のレベル差の絶対値のみで音像位置を判断する方式であるため、音像が左右のどちらに寄っているかというような情報は表示されない。

20

【0011】

また、通常の音楽コンテンツなどにおいては、それぞれの音像位置が異なるようにミキシングされたボーカルや各楽器の音が混ざり合っており、それぞれの構成音毎に音像が異なるが、特許文献 2 に記載の方式では、音像位置は代表値で表示されることになるため、聴覚と知覚で音像位置が一致しないことになる。

【0012】

さらに、左右で異なる音声信号が入力されている場合、知覚上、音像は生じないはずであるが、特許文献 2 に記載の方式では、左右の音声信号のレベル差の絶対値のみで音像位置を判断するため、いずれかの場所に音像が表示されてしまい、聴覚と知覚とで音像位置が一致しないことになる。

30

【0013】

本発明は、上述のような実情に鑑みてなされたものであり、その目的は、聴覚と知覚とで音像位置が一致するように音像位置を表示させることが可能な音声信号再生装置及び音声信号再生方法を提供することにある。

【課題を解決するための手段】

【0014】

上記の課題を解決するために、本発明の第 1 の技術手段は、音声信号再生装置であって、2 つの音声信号を、周波数毎または周波数帯域毎に周波数についての相関信号と無相関信号とに分離する分離部と、該分離部で分離された各相関信号について、音像方向及び信号電力値を算出する算出部と、各相関信号について、前記音像方向に基づき、予め定められた複数の音像方向グループのうちの 1 つに前記信号電力値を割り当て、該音像方向グループ毎に前記信号電力値の加算値を求める加算部と、前記複数の音像方向グループのそれぞれに対応する位置に各音像方向グループについての前記加算値を示す情報を表示する音像表示部と、を備えたことを特徴としたものである。

40

【0015】

本発明の第 2 の技術手段は、第 1 の技術手段において、前記音像表示部は、棒グラフ形式で、前記複数の音像方向グループのそれぞれに対応する位置に各音像方向グループについての前記加算値を示す情報を表示することを特徴としたものである。

【0016】

本発明の第 3 の技術手段は、第 1 の技術手段において、前記音像表示部は、複数の発光

50

部を有し、前記複数の音像方向グループのそれぞれに対応する位置の前記発光部の発光色を、各音像方向グループについての前記加算値を示す情報に応じて変化させることを特徴としたものである。

【 0 0 1 7 】

本発明の第 4 の技術手段は、第 1 の技術手段において、前記音像表示部は、複数の発光部を有し、前記複数の音像方向グループのそれぞれに対応する位置の前記発光部の発光強度を、各音像方向グループについての前記加算値を示す情報に応じて変化させることを特徴としたものである。

【 0 0 1 8 】

本発明の第 5 の技術手段は、音声信号再生方法であって、分離部が、2つの音声信号を、周波数毎または周波数帯域毎に周波数についての相関信号と無相関信号とに分離する分離ステップと、算出部が、前記分離ステップで分離された各相関信号について、音像方向及び信号電力値を算出する算出ステップと、加算部が、各相関信号について、前記音像方向に基づき、予め定められた複数の音像方向グループのうちの 1 つに前記信号電力値を割り当て、該音像方向グループ毎に前記信号電力値の加算値を求める加算ステップと、音像表示部が、前記複数の音像方向グループのそれぞれに対応する位置に各音像方向グループについての前記加算値を示す情報を表示する音像表示ステップと、を含むことを特徴としたものである。

【発明の効果】

【 0 0 1 9 】

本発明によれば、音声信号再生装置において、聴覚と知覚とで音像位置が一致するように音像位置を表示させることが可能になる。

【図面の簡単な説明】

【 0 0 2 0 】

【図 1】本発明に係る音声信号再生装置の一構成例を示すブロック図である。

【図 2】図 1 の音声信号再生装置における音声信号処理部の一構成例を示すブロック図である。

【図 3】図 2 の音声信号処理部における分離抽出部での分離抽出処理の一例を説明するためのフロー図である。

【図 4】受聴者と左右のスピーカと合成音像との位置関係の一例を説明するための模式図である。

【図 5】再生用スピーカと受聴者及び合成音像との位置関係の一例を説明するための模式図である。

【図 6】図 1 の音声信号再生装置において実行される音像表示処理の一例を説明するためのフロー図である。

【図 7】図 1 の音声信号再生装置における表示部の一例を示す図である。

【図 8】図 1 の音声信号再生装置における表示部の他の例を示す図である。

【図 9】5.1ch サラウンドシステムのスピーカ群のうち、LFEを除いた 5 つのスピーカの配置例である。

【図 10】図 9 の配置例におけるダウンミックス後の出力対象スピーカの配置例を示す図である。

【図 11】図 2 の音声信号処理部における分離抽出部での着目ペア判定処理の一例を説明するためのフロー図である。

【図 12】受聴者と左右のスピーカと仮想音源との位置関係の他の例を説明するための模式図である。

【図 13】受聴者と左右のスピーカ及び左右のサラウンドスピーカと仮想音源との位置関係の一例を説明するための模式図である。

【図 14】受聴者と左右のスピーカ及び左右のサラウンドスピーカと仮想音源との位置関係の他の例を説明するための模式図である。

【図 15】受聴者と左右のスピーカ及び左右のサラウンドスピーカと仮想音源との位置関

10

20

30

40

50

係の他の例を説明するための模式図である。

【図 1 6】受聴者と左右のスピーカ及び左右のサラウンドスピーカと仮想音源との位置関係の他の例を説明するための模式図である。

【図 1 7】左右のスピーカ及び左右のサラウンドスピーカと全ての仮想音源との位置関係の一例を説明するための模式図である。

【図 1 8】左右のスピーカ及び左右のサラウンドスピーカと全ての仮想音源との位置関係の他の例を説明するための模式図である。

【図 1 9】6.1 c h サラウンドシステムのスピーカ群において、L F E を除いた 6 つのスピーカのうち、ダウンミックス後の出力対象スピーカの配置例を示す図である。

【図 2 0】7.1 c h サラウンドシステムのスピーカ群において、L F E を除いた 7 つのスピーカのうち、ダウンミックス後の出力対象スピーカの配置例を示す図である。

【図 2 1】非特許文献 1 に記載の技術において、1 つの直線上に並べたスピーカ群の背後に仮想音源を設けた場合に、各仮想音源に対応する音を出力するスピーカについて説明するための模式図である。

【図 2 2】図 1 の音声信号再生装置におけるスピーカ群の配置例を説明するための模式図である。

【図 2 3】図 1 の音声信号再生装置におけるスピーカ群の他の配置例を説明するための模式図である。

【図 2 4】図 1 の音声信号再生装置におけるスピーカ群の他の配置例を説明するための模式図である。

【図 2 5】図 1 の音声信号再生装置を備えた映像表示システムの構成例を示す図である。

【発明を実施するための形態】

【0021】

本発明に係る音声信号再生装置は、2 チャンネル以上のマルチチャンネル再生方式用の音声信号を、そのまま対応するスピーカ群から出力するか、もしくは、波面合成再生方式などの他の再生方式でスピーカ群からより適切な音像を提供できるような音声信号に変換してから出力する装置であり、後述するような音像表示処理を行う点に特徴を有する。本発明に係る音声信号再生装置は、音声再生処理と音像表示処理が実行可能なように構成すればよく、例えば、テレビ装置等の表示装置やオーディオシステムなど、様々な A V (Audio Visual) 機器として構成することができる。

【0022】

音声再生機能に関しては、以下では基本的に、波面合成再生方式により 2 つ以上のチャンネルの入力音声信号を仮想音源に対する音像としてスピーカ群（複数のスピーカ）により再生させるための音声信号に変換して再生する場合を例に挙げて説明する。以下では特に、仮想音源の数と同数・同位置の再生用スピーカを用意して、各仮想音源から出力すべき音声信号を、一対一で対応する再生用スピーカから再生するような場合について説明する。

【0023】

ただし、スピーカの数や位置が仮想音源のそれらと異なるように仮想音源が設定される場合にも、各仮想音源からスピーカ群への割り当てを行うことで、同様に適用できる。また、上述したように、入力された音声信号をそのまま音声出力するような構成を採用することもできる。なお、この場合にも、本発明に係る音像表示処理のための処理、例えば音声信号の分離抽出までの処理は必要となる。

【0024】

以下、図面を参照しながら、本発明に係る音声信号再生装置の構成例及び処理例について説明する。図 1 は、本発明に係る音声信号再生装置の一構成例を示すブロック図で、図 2 は、図 1 の音声信号再生装置における音声信号処理部の一構成例を示すブロック図である。

【0025】

図 1 で例示する音声信号再生装置（音声データ再生装置）10 は、デコーダ 11、音声

10

20

30

40

50

信号抽出部 12、音声信号処理部 13 を備えるとともに、D/A コンバータ 14、複数の増幅器 15、及び複数のスピーカ 16 となるスピーカアレイ 17 を備える。なお、増幅器 15 とスピーカ 16 の数は基本的に同じとする。また、音声信号再生装置 10 は、音像表示処理のために表示制御部 18 及び表示部 19 を備える。

#### 【0026】

デコーダ 11 は、音声のみあるいは音声付き映像のコンテンツを復号化し、信号処理可能な形式に変換し音声信号抽出部 12 に出力する。そのコンテンツは、放送局から送信されたデジタル放送のコンテンツや、ネットワークを介してディジタルコンテンツを配信するサーバからインターネットからダウンロードしたり、あるいは外部記憶装置等の記録媒体から読み込んだりすることによって取得する。このように、図 1 では図示しないが、音声信号再生装置 10 は、マルチチャネルの入力音声信号を含むディジタルコンテンツを入力するディジタルコンテンツ入力部を備える。デコーダ 11 は、ここで入力されたディジタルコンテンツを復号化することになる。

#### 【0027】

音声信号抽出部 12 では、得られた信号から音声信号を分離、抽出する。ここで説明する例では、抽出する音声信号を 2ch ステレオ信号とし、音声信号抽出部 12 はその 2 チャネル分の信号を音声信号処理部 13 に出力する。無論、元々入力された信号が 2ch ステレオ信号である場合には、音声信号抽出部 12 では、その 2ch ステレオ信号を抽出すれば済む。

#### 【0028】

一方で、例えば入力音声信号が 5.1ch など、2ch を越えるチャネル数である場合には、音声信号抽出部 12 は、例えば ARIB STD-B21「デジタル放送用受信装置 標準規格」によって定められているような、次の数式(1)の通常のダウンミックス方法によって 2ch にダウンミックスし、音声信号処理部 13 に出力する。

#### 【0029】

##### 【数 1】

$$L_t = a \times \left( L + \frac{1}{\sqrt{2}} \times C + k_d L_s \right)$$

$$R_t = a \times \left( R + \frac{1}{\sqrt{2}} \times C + k_d R_s \right) \quad (1)$$

#### 【0030】

数式(1)で、 $L_t$ 、 $R_t$  はダウンミックス後の左右チャネル信号、 $L$ 、 $R$ 、 $C$ 、 $L_s$ 、 $R_s$  はそれぞれ 5.1ch の各信号（左フロントチャネル信号、右フロントチャネル信号、センターチャネル信号、左リアチャネル信号、右リアチャネル信号）、 $a$  はオーバーロード低減係数で例えば  $1/\sqrt{2}$ 、 $k_d$  はダウンミックス係数で例えば  $1/\sqrt{2}$ 、または  $1/2$ 、または  $1/2$ 、または 0 となる。

#### 【0031】

このように、マルチチャネルの入力音声信号は、3 以上のチャネルをもつマルチチャネル再生方式の入力音声信号であってもよく、その場合、音声信号処理部 13 では、マルチチャネルの入力音声信号を 2 つのチャネルの音声信号にダウンミックスした後の 2 つのチャネルの音声信号について、後述の音像表示処理の対象としてもよい。もしくは、3 以上のチャネルをもつマルチチャネル再生方式の入力音声信号から単に 2 つのチャネルの音声信号を抽出して、その 2 つのチャネルの音声信号についてのみ、後述の音像表示処理の対象としてもよい。

#### 【0032】

音声信号処理部 13 では、得られた 2 チャネル信号から、音声再生処理の一部として、入力音声信号とは異なるマルチチャネル（好ましくは 5 チャネル以上）の音声信号を生成

10

20

30

40

50

する。

【0033】

音声信号処理部13は、それらの音声信号をD/Aコンバータ14に出力する。D/Aコンバータ14では得られた信号をアナログ信号に変換し、それぞれの信号を増幅器15に出力する。各増幅器15では入力されたアナログ信号を拡声し各スピーカ16に伝送し、この拡声されたアナログ信号が各スピーカ16から空間中に音として出力される。

【0034】

スピーカアレイ17において、スピーカ16は一系列または複数列に配置されており、その1つ1つの形状は円形や楕円形や菱形などどのような形状であってもよい。また、配列の方向も直線状に限らず、曲線状に各スピーカ16の中心を配列しておいてもよい。なお、音声信号処理部13では、スピーカアレイ17における各スピーカ16の配置に応じて各スピーカ16に対する遅延量や出力レベルを決定すればよい。

10

【0035】

さらに、音声信号処理部13では、本発明の主たる特徴である音像表示処理の一部として、上記得られた2チャンネル信号から音像方向情報を生成する。音像方向情報とは、周波数毎(線スペクトル毎)または周波数帯域毎(周波数領域毎)についての、音像の推定方向角(音像方向を示す値)とその信号成分の電力値(信号電力値)を指す。音声信号処理部13は、線スペクトル毎または周波数帯域毎に生成した音像方向情報を表示制御部18に出力する。

【0036】

表示制御部18は、音像表示のための情報を求める加算部18aを有する。この加算部18aは、各相関信号について、算出された音像方向に基づき、予め定められた複数の音像方向グループ(音像方向範囲)のうちの1つに算出された信号電力値を割り当て、音像方向グループ毎に信号電力値の加算値を求める。表示制御部18は、加算部18aで求めた音像方向グループ毎の加算値を表示部19に出力する。

20

【0037】

表示部19は、音像表示部の一例であり、音像表示部は、上記複数の音像方向グループのそれぞれに対応する位置(好ましくは各音像方向グループの中央の方向のと一致する位置)に各音像方向グループについての加算値を示す情報を表示する。具体的な表示例については後述する。音像方向グループ毎の信号電力値は、表示対象となる2チャンネル信号により時系列で変化しており、表示部19ではそのような変化に合わせて音像方向の表示を変化させることができる。このように、本発明では、聴覚と知覚とで音像位置が一致するように音像位置を表示させることができ、受聴者の没入感を高めることができる。

30

【0038】

以下、音声再生処理及び音像表示処理を行う音声信号処理部13の詳細な構成例を、図2を参照しながら説明する。音声信号処理部13は、変換部21、分離抽出部22、逆変換部23、及び音声出力信号生成部24を備える。

【0039】

変換部21は、入力された2つのチャンネルの入力信号それぞれを、1セグメントの1/4の長さの音声データ分読み出す。ここで、音声データとは、例えば48kHzなどの標本化周波数で標本化された離散音声信号波形を指すものとする。そして、セグメントとは、ある一定の長さの標本点群からなる音声データ区間であり、ここでは後ほど離散フーリエ変換の対象となる区間長を指すものとし、処理セグメントとも呼ぶ。その値は例えば1024とする。この例では、1セグメントの1/4の長さである256点の音声データが読み出し対象となる。

40

【0040】

読み出した256点の音声データはバッファに蓄えられる。このバッファは、直前の1セグメント分の音声信号波形を保持しておけるようになっており、それより過去のセグメントは捨てていく。直前の3/4セグメント分のデータ(768点)と最新の1/4セグメント分のデータ(256点)を繋げて1セグメント分の音声データを作成し、窓関数を

50

乗算する。すなわち、全ての標本データは窓関数演算に4回読み込まれることになる。

【0041】

ここで、窓関数の乗算とは、従来提案されている次のHann窓を1セグメント分の音声データに乗算する窓関数演算処理を実行する。

【数2】

$$w(m) = \sin^2\left(\frac{m}{M}\pi\right) \quad (0 \leq m < M) \quad (2)$$

【0042】

mは自然数、Mは1セグメント長で偶数とする。変換部21への入力信号をそれぞれ $x_L(m)$ 、 $x_R(m)$ とすると、窓関数乗算後の音声信号 $x_L(m)$ 、 $x_R(m)$ は、

$$\begin{aligned} x_L(m) &= w(m) x_L(m), \\ x_R(m) &= w(m) x_R(m) \end{aligned} \quad (3)$$

と計算される。

【0043】

変換部21は、そうして得られた音声データを、次の数式(4)のように離散フーリエ変換し、周波数領域の音声データを得る。つまり、変換部21は、2つのチャンネル(の音声データ)に対し、離散フーリエ変換を施す。ここで、DFTは離散フーリエ変換を表し、kは自然数で、 $0 < k \leq M/2$ である。 $X_L(k)$ 、 $X_R(k)$ は複素数となる。

$$\begin{aligned} X_L(k) &= \text{DFT}(x_L(m)), \\ X_R(k) &= \text{DFT}(x_R(m)) \end{aligned} \quad (4)$$

【0044】

分離抽出部22は、線スペクトル毎(つまり周波数毎)に、変換部21で変換された2つのチャンネルについて、相関信号と無相関信号を分離(抽出)する。以下では、分離抽出部22のうちこのような分離(抽出)を行う部位を、分離部22aとして説明する。つまり、分離部22aは2つの音声信号を、周波数毎(または周波数帯域毎)に相関信号と無相関信号とに分離する。ここで相関信号とは、周波数について相関する信号を指し、無相関信号とは周波数について相関しない信号を指す。

【0045】

また、分離抽出部22は、画像表示処理のための算出部22bを有し、この算出部22bは、分離部22aで分離された各相関信号について、画像方向及び信号電力値を算出する。より具体的には、算出部22bは、線スペクトル毎に、その相関信号の画像方向の角度の推定を行うと共に、その相関信号の信号電力値を求める。なお、算出部22bは、以下に説明する例のように音声再生処理において画像方向及び信号電力値が必要な場合には、画像表示処理と音声再生処理の双方の処理を兼ねることになる。音声再生処理に画像方向及び信号電力値が必要な場合としては、上述したように、入力された音声信号をそのまま音声出力するような構成が挙げられる。

【0046】

また、分離部22aにおいては、線スペクトル毎でなくても周波数帯域(小帯域)毎に相関信号と無相関信号とを分離、抽出してもよく、その場合、算出部22bでの画像方向及び信号電力値の算出も小帯域毎の相関信号についての算出となる。つまり、ここでは線スペクトル毎に相関係数を取得するなどの処理を行う例を挙げて説明するが、特許文献1に記載のように、Equivalent Rectangular Band(ERB)を用いて分割した帯域(周波数領域であり、小帯域とも呼ぶ)毎に相関係数を取得するなどの処理を実行してもよい。

【0047】

算出部22bでの算出後は、加算部18aが、その推定角度に基づいて、相関信号を方向毎に分離(グループ分け)し、各画像方向グループに属する相関信号の信号電力値について加算値を求める。そして、表示部19が、画像方向グループと加算値を示す情報とを関連付けて表示する。



## 【 0 0 4 8 】

分離抽出部 2 2 における処理の具体的な内容を、図 3 を用いて説明する。分離抽出部 2 2 は、変換部 2 1 で離散フーリエ変換後の 2 つのチャネルの音声信号について、線スペクトル毎にステップ S 3 2 ~ S 3 4 の処理を実行する（ステップ S 3 1 a , S 3 1 b。具体的に個々の処理について説明する。

## 【 0 0 4 9 】

離散フーリエ変換した後の線スペクトルは、直流成分すなわち例えば  $X_L(0)$  を除いて、 $M/2$ （ただし、 $M$  は偶数）を境に対称となっている。すなわち、 $X_L(k)$  と  $X_L(M-k)$  は  $0 < k < M/2$  の範囲で複素共役の関係になる。したがって、以下では  $k < M/2$  の範囲を分析の対象として考え、 $k > M/2$  の範囲については複素共役の関係にある対称の線スペクトルと同じ扱いとする。

10

## 【 0 0 5 0 】

次に、各線スペクトルに対し、左チャネルと右チャネルの正規化相関係数を次式で求めることで、相関係数を取得する。ここで、対象となる線スペクトルの音声信号  $X$  のゲイン（振幅）を  $G$  とすると、 $P(X)$  は、音声信号  $X$  の電力（音圧に相当）を表し、ゲイン  $G$  の二乗値で表現できる。

## 【 0 0 5 1 】

## 【 数 3 】

$$d^{(k)} = \frac{D^{(k)}}{G^{(k)}} \quad (5)$$

20

$$D^{(k)} = |Re\{X_L(k)\} \cdot Re\{X_R(k)\} + Im\{X_L(k)\} \cdot Im\{X_R(k)\}| \quad (6)$$

$$G^{(k)} = \sqrt{P_L^{(k)} P_R^{(k)}} \quad (7)$$

$$P_L^{(k)} = |X_L(k)|^2, \quad P_R^{(k)} = |X_R(k)|^2 \quad (8)$$

## 【 0 0 5 2 】

この正規化相関係数  $d^{(k)}$  は左右のチャネルの音声信号にどれだけ相関があるかを表すものであり、0 から 1 の間の実数の値をとる。全く同じ信号同士であれば 1、そして全く無相関の信号同士であれば 0 となる。ここで、左右のチャネルの音声信号の電力  $P_L^{(k)}$  と  $P_R^{(k)}$  の両方が 0 である場合、その線スペクトルに関して相関信号と無相関信号の抽出は不可能とし、処理を行わず次の線スペクトルの処理に移ることとする。また、 $P_L^{(k)}$  と  $P_R^{(k)}$  のいずれか片方が 0 である場合、数式(5)では演算不可能であるが、正規化相関係数  $d^{(k)} = 0$  とし、その線スペクトルの処理を続行する。

30

## 【 0 0 5 3 】

次に、この正規化相関係数  $d^{(k)}$  を用いて、左右チャネルの音声信号から相関信号と無相関信号をそれぞれ分離抽出するための変換係数を求め（ステップ S 3 2）、ステップ S 3 2 で取得したそれぞれの変換係数を用いて、左右チャネルの音声信号から相関信号と無相関信号を分離抽出する（ステップ S 3 3）。相関信号及び無相関信号は、いずれも推定した音声信号として抽出すればよい。ステップ S 3 2 , S 3 3 の処理は、主に分離抽出部 2 2 のうち分離部 2 2 a が担えばよいが、算出部 2 2 b における処理も兼ねている。

40

## 【 0 0 5 4 】

ステップ S 3 2 , S 3 3 の処理例を説明する。ここで、特許文献 1 と同様、左右チャネルそれぞれの信号は、無相関信号と相関信号から構成され、相関信号については、左右のチャネルからゲインのみ異なる信号波形（つまり同じ周波数成分からなる信号波形）が出力されるものとするモデルを採用する。ここで、ゲインは、信号波形の振幅に相当し、音圧に関連する値である。そして、このモデルでは、左右のチャネル信号から出力される相関信号によって合成される音像は、その相関信号の左右それぞれの音圧のバランスによって方向が決定されるものとする。

50

## 【 0 0 5 5 】

そのモデルに従うと、入力信号  $x_L(m)$ 、 $x_R(m)$  は、

$$\begin{aligned} x_L(m) &= s(m) + n_L(m), \\ x_R(m) &= s(m) + n_R(m) \end{aligned} \quad (9)$$

と表される。ここで、 $s(m)$  は左右の相関信号、 $n_L(m)$  は左チャネルの音声信号から相関信号  $s(m)$  を減算したものであって（左チャネルの）無相関信号として定義できるもの、 $n_R(m)$  は右チャネルの音声信号から相関信号  $s(m)$  に  $d$  を乗算したものを減算したものであって（右チャネルの）無相関信号として定義できるものである。また、 $d$  は相関信号の左右の音圧バランスの程度を表す正の実数である。

## 【 0 0 5 6 】

10

数式(9)により、数式(3)で前述した窓関数乗算後の音声信号  $x_L(m)$ 、 $x_R(m)$  は、次の数式(10)で表される。ただし、 $s(m)$ 、 $n_L(m)$ 、 $n_R(m)$  はそれぞれ  $s(m)$ 、 $n_L(m)$ 、 $n_R(m)$  に窓関数を乗算したものである。

$$\begin{aligned} x_L(m) &= w(m) \{ s(m) + n_L(m) \} = s(m) + n_L(m), \\ x_R(m) &= w(m) \{ s(m) + n_R(m) \} = s(m) + n_R(m) \end{aligned} \quad (10)$$

## 【 0 0 5 7 】

数式(10)を離散フーリエ変換することによって、次の数式(11)を得る。ただし、 $S(k)$ 、 $N_L(k)$ 、 $N_R(k)$  はそれぞれ  $s(m)$ 、 $n_L(m)$ 、 $n_R(m)$  を離散フーリエ変換したものである。

20

$$\begin{aligned} X_L(k) &= S(k) + N_L(k), \\ X_R(k) &= S(k) + N_R(k) \end{aligned} \quad (11)$$

## 【 0 0 5 8 】

したがって、 $k$  番目の線スペクトルにおける音声信号  $X_L(k)$ 、 $X_R(k)$  は、

$$\begin{aligned} X_L(k) &= S(k) + N_L(k), \\ X_R(k) &= d^{(k)} S(k) + N_R(k) \end{aligned} \quad (12)$$

と表現される。ここで、 $d^{(k)}$  は  $k$  番目の線スペクトルにおける  $d$  を表す。

## 【 0 0 5 9 】

数式(12)から、数式(8)の音圧  $P_L^{(k)}$  と  $P_R^{(k)}$  は、

$$\begin{aligned} P_L^{(k)} &= P_S^{(k)} + P_N^{(k)}, \\ P_R^{(k)} &= [d^{(k)}]^2 P_S^{(k)} + P_N^{(k)} \end{aligned} \quad (13)$$

30

と表される。ここで、 $P_S^{(k)}$ 、 $P_N^{(k)}$  は  $k$  番目の線スペクトルにおけるそれぞれ相関信号、無相関信号の電力であり、

## 【 数 4 】

$$P_S^{(k)} = |S(k)|^2, \quad P_N^{(k)} = |N_L(k)|^2 = |N_R(k)|^2 \quad (14)$$

と表される。ここで、左右のチャネル間の無相関信号の音圧は等しいと仮定している。

## 【 0 0 6 0 】

また、数式(6)～(8)より、 $d^{(k)}$  は、次の数式(15)で表すことができる。ただし、この算出においては  $S(k)$ 、 $N_L(k)$ 、 $N_R(k)$  が互いに直交し、かけ合わされたときの電力は 0 と仮定している。

40

## 【 数 5 】

$$d^{(k)} = \frac{\alpha^{(k)} P_S^{(k)}}{\sqrt{P_L^{(k)} P_R^{(k)}}} \quad (15)$$

## 【 0 0 6 1 】

数式(13)と数式(15)を解くことにより、次の式が得られる。

## 【数 6】

$$\alpha^{(k)} = \frac{\beta}{2\gamma} \quad , \quad P_S^{(k)} = \frac{2\gamma^2}{\beta} \quad , \quad P_N^{(k)} = P_L^{(k)} - P_S^{(k)} = P_L^{(k)} - \frac{2\gamma^2}{\beta} \quad (16)$$

ただし、 $\beta$ 、 $\gamma$ は中間変数であり、

$$\beta = P_R^{(k)} - P_L^{(k)} + \sqrt{(P_L^{(k)} - P_R^{(k)})^2 + 4P_L^{(k)}P_R^{(k)}[d^{(k)}]^2} \quad , \quad \gamma = d^{(k)}\sqrt{P_L^{(k)}P_R^{(k)}} \quad (17)$$

である。

10

## 【0062】

これらの値を用いて、各線スペクトルにおける相関信号と無相関信号を推定する。k 番目の線スペクトルにおける相関信号  $S(k)$  の推定値  $est(S(k))$  を、媒介変数  $\mu_1$ 、 $\mu_2$  を用いて、

$$est(S(k)) = \mu_1 X_L(k) + \mu_2 X_R(k) \quad (18)$$

とおくと、推定誤差 は、

$$= est(S(k)) - S(k) \quad (19)$$

と表される。ここで、 $est(A)$  は  $A$  の推定値を表すものとする。そして二乗誤差<sup>2</sup>が最少になるとき、 $X_L(k)$ 、 $X_R(k)$  はそれぞれ直交するという性質を利用すると、

20

$$E[\quad \cdot X_L(k)] = 0, \quad E[\quad \cdot X_R(k)] = 0 \quad (20)$$

という関係が成り立つ。

## 【0063】

数式(12)、(14)、(16)～(19)を利用すると、数式(20)から次の連立方程式が導出できる

。

$$\begin{aligned} (1 - \mu_1 - \mu_2) P_S^{(k)} - \mu_1 P_N^{(k)} &= 0 \\ (1 - \mu_1 - \mu_2) P_S^{(k)} - \mu_2 P_N^{(k)} &= 0 \end{aligned} \quad (21)$$

## 【0064】

この数式(21)を解くことによって、各媒介変数が次のように求まる。

30

## 【数 7】

$$\mu_1 = \frac{P_S^{(k)}}{(\alpha^{(k)})^2 + 1} \frac{P_S^{(k)}}{P_S^{(k)} + P_N^{(k)}} \quad , \quad \mu_2 = \frac{\alpha^{(k)} P_S^{(k)}}{(\alpha^{(k)})^2 + 1} \frac{P_S^{(k)}}{P_S^{(k)} + P_N^{(k)}} \quad (22)$$

ここで、このようにして求まる推定値  $est(S(k))$  の電力  $P_{est(S(k))}$  が、数式(18)の両辺を二乗して求まる次の式

$$P_{est(S(k))} = (\mu_1 + \mu_2)^2 P_S^{(k)} + (\mu_1^2 + \mu_2^2) P_N^{(k)} \quad (23)$$

を満たす必要があるため、この式から推定値を次式のようにスケールする。なお、 $est(A)$  は  $A$  の推定値をスケールしたものを表す。

40

## 【0065】

## 【数 8】

$$est'(S(k)) = \frac{\sqrt{P_S^{(k)}}}{\sqrt{(\mu_1 + \alpha^{(k)} \mu_2)^2 P_S^{(k)} + (\mu_1^2 + \mu_2^2) P_N^{(k)}}} est(S(k)) \quad (24)$$

## 【0066】

そして、k 番目の線スペクトルにおける左右のチャネルの無相関信号  $N_L(k)$ 、 $N_R$

50

(k) に対する推定値  $est(N_L(k))$ 、 $est(N_R(k))$  はそれぞれ、

$$est(N_L(k)) = \mu_3 X_L(k) + \mu_4 X_R(k) \quad (25)$$

$$est(N_R(k)) = \mu_5 X_L(k) + \mu_6 X_R(k) \quad (26)$$

とおくことにより、上述の求め方と同様にして、媒介変数  $\mu_3 \sim \mu_6$  は、

【数 9】

$$\mu_3 = \frac{[\alpha^{(k)}]^2 P_S^{(k)} + P_N^{(k)}}{([\alpha^{(k)}]^2 + 1) P_S^{(k)} + P_N^{(k)}} \quad , \quad \mu_4 = \frac{-\alpha^{(k)} P_S^{(k)}}{([\alpha^{(k)}]^2 + 1) P_S^{(k)} + P_N^{(k)}} \quad (27)$$

$$\mu_5 = \frac{-\alpha^{(k)} P_S^{(k)}}{([\alpha^{(k)}]^2 + 1) P_S^{(k)} + P_N^{(k)}} \quad , \quad \mu_6 = \frac{P_S^{(k)} + P_N^{(k)}}{([\alpha^{(k)}]^2 + 1) P_S^{(k)} + P_N^{(k)}} \quad (28)$$

10

と求めることができる。このようにして求めた推定値  $est(N_L(k))$ 、 $est(N_R(k))$  も上述と同様に、次の式によってそれぞれスケールリングする。

【0067】

【数 10】

$$est'(N_L(k)) = \frac{\sqrt{P_N^{(k)}}}{\sqrt{(\mu_3 + \alpha^{(k)} \mu_4)^2 P_S^{(k)} + (\mu_3^2 + \mu_4^2) P_N^{(k)}}} est(N_L(k)) \quad (29)$$

20

$$est'(N_R(k)) = \frac{\sqrt{P_N^{(k)}}}{\sqrt{(\mu_5 + \alpha^{(k)} \mu_6)^2 P_S^{(k)} + (\mu_5^2 + \mu_6^2) P_N^{(k)}}} est(N_R(k)) \quad (30)$$

【0068】

数式(22)、(27)、(28)で示した各媒介変数  $\mu_1 \sim \mu_6$  及び数式(24)、(29)、(30)で示したスケールリングの係数が、ステップ S 3 2 で求める変換係数に該当する。そして、ステップ S 3 3 では、これらの変換係数を用いた演算（数式(18)、(25)、(26)）により推定することで、相関信号と無相関信号（右チャンネルの無相関信号、左チャンネルの無相関信号）とを分離抽出する。

【0069】

30

次に、再生用スピーカ 1 6 への割り当て処理を行う（ステップ S 3 4）。この処理は、再生用スピーカ 1 6 のそれぞれに仮想音源を設定した処理に該当する。なお、スピーカ 1 6 とは異なる位置に仮想音源を設定した場合には、音声信号処理部 1 3 において、ここで説明する処理と同様に仮想音源への割り当てを行った後、各スピーカ 1 6 の配置に応じて各スピーカ 1 6 に対する遅延量や出力レベルを決定すればよい。

【0070】

ステップ S 3 4 の割り当て処理では、前処理として、線スペクトル毎に推定した相関信号によって生成される合成音像の方向を推定する。この推定処理は、図 2 の分離抽出部 2 2 のうち算出部 2 2 b が担うことになる。この推定処理について、図 4、図 5 に基づき説明する。図 4 は、受聴者と左右のスピーカと合成音像との位置関係の一例を説明するための模式図、図 5 は、再生用スピーカと受聴者及び合成音像との位置関係の一例を説明するための模式図である。

40

【0071】

いま、図 4 に示す位置関係 4 0 のように、受聴者から左右のスピーカ 4 1 L、4 1 R の中点にひいた線と、同じく受聴者 4 3 からいずれかのスピーカ 4 1 L / 4 1 R の中心までひいた線がなす見開き角を  $\theta_0$ 、受聴者 4 3 から推定合成音像 4 2 の位置までひいた線がなす見開き角を  $\theta$  とする。ここで、左右のスピーカ 4 1 L、4 1 R から同じ音声信号を、音圧バランスを変えて出力した場合、その出力音声によって生じる合成音像 4 2 の方向は、音圧バランスを表す前述のパラメータ  $\theta$  を用いて次の式で近似できることが一般的に知られている（以下、立体音響におけるサインの法則と呼ぶ）。

50

【 0 0 7 2 】

【 数 1 1 】

$$\theta = \sin^{-1} \left( \frac{\alpha - 1}{\alpha + 1} \sin \theta_0 \right) \quad (31)$$

【 0 0 7 3 】

ここで、2 c h ステレオの音声信号を波面合成再生方式で再生できるようにするために、図2に示す分離部22aが2 c hの信号を複数チャネルの信号に変換する。例えば変換後のチャネル数を5つとした場合、それを図5で示す位置関係50のように、スピーカアレイ17として、均等に配置された再生用スピーカ52a～52eを用いて再生する。再生用スピーカ52a～52eは、図1の複数のスピーカ16に該当するものである。既に説明したように、分離部22aは、まず2 c hの音声信号を、線スペクトル毎に1つの相関信号と2つの無相関信号に分離する。分離部22aでは、さらにそれらの信号をどのように再生用スピーカ（ここでは5つの再生用スピーカ）に割り当てるかを事前に決めておかなければならない。なお、割り当ての方法については複数の方法の中からユーザ設定可能にしておいてもよいし、再生用スピーカ数に応じて選択可能な方法を変えてユーザに提示するようにしてもよい。

【 0 0 7 4 】

割り当て方法の1つの例として、次のような方法を採用。それは、まず、左右の無相関信号については、5つの再生用スピーカの両端（スピーカ52a, 52e）にそれぞれ割り当てる。次に、相関信号によって生じる合成音像については、5つのうちの隣接する2つの再生用スピーカに割り当てる。隣接する2つの再生用スピーカに割り当てるかについては、まず、前提として、相関信号によって生じる合成音像が5つの再生用スピーカの両端（再生用スピーカ52a, 52e）より内側になるものとし、すなわち、2 c h ステレオ再生時の2つのスピーカによってなす見開き角内におさまるように5つの再生用スピーカ52a～52eを配置するものとする。そして、合成音像の推定方向から、その合成音像を挟むような隣接する2つの再生用スピーカを決定し、その2つの再生用スピーカへの音圧バランスの割り当てを調整して、その2つの再生用スピーカによって合成音像を生じさせるように再生する、という割り当て方法を採用。

【 0 0 7 5 】

そこで、図5で示す位置関係50のように、受聴者53から両端の再生用スピーカ52a, 52eの midpoint にひいた線と端の再生用スピーカ52eにひいた線とがなす見開き角を $\theta_0$ 、上記 midpoint にひいた線と受聴者53から合成音像51にひいた線とがなす見開き角を $\theta$ とする。さらに、受聴者53から合成音像51を挟む2つの再生用スピーカ52c, 52dの midpoint にひいた線と、受聴者53から再生用スピーカ52dとがなす見開き角を $\theta_0$ 、受聴者53から合成音像51にひいた線とがなす見開き角を $\theta$ とする。ここで、 $\theta_0$ は正の実数である。数式(31)で説明したようにして方向を推定した図4の合成音像42（図5における合成音像51に対応）を、これらの変数を用いて再生用スピーカに割り当てる方法について説明する。

【 0 0 7 6 】

まず、k番目の合成音像の方向 $\theta(k)$ が数式(31)によって推定され、例えば $\theta(k) = 0.15$  [rad]であったとする。そして、再生用スピーカが5つの場合、図5に示すように合成音像51は左から数えて3番目の再生用スピーカ52cと4番目の再生用スピーカ52dの間に位置することになる。また、再生用スピーカが5つである場合、3番目の再生用スピーカ52cと4番目の再生用スピーカ52dの間について、三角関数を用いた単純な幾何的計算により、 $\theta_0 = 0.121$  [rad]となり、k番目の線スペクトルにおける $\theta(k)$ とすると、 $\theta(k) = \theta_0 = 0.088$  [rad]となる。このようにして、各線スペクトルにおける相関信号によって生じる合成音像の方向を、それを挟む2つの再生用スピーカの方向からの相対的な角度で表す。そして上述したように、その2つの再生用スピーカ52c, 52dでその合成音像を生じさせることを考

える。そのためには、2つの再生用スピーカ52c, 52dからの出力音声信号の音圧バランスを調整すればよく、その調整方法については、再び数式(31)として利用した立体音響におけるサインの法則を用いる。

【0077】

ここで、k番目の線スペクトルにおける相関信号によって生じる合成音像を挟む2つの再生用スピーカ52c, 52dのうち、3番目の再生用スピーカ52cに対するスケール係数を $g_1$ 、4番目の再生用スピーカ52dに対するスケール係数を $g_2$ とすると、3番目の再生用スピーカ52cからは $g_1 \cdot \text{est}(S(k))$ 、4番目の再生用スピーカ52dからは $g_2 \cdot \text{est}(S(k))$ の音声信号を出力することになる。そして、 $g_1$ 、 $g_2$ は立体音響におけるサインの法則により、

【数12】

$$\frac{\sin \varphi^{(k)}}{\sin \varphi_0} = \frac{g_2 - g_1}{g_2 + g_1} \quad (32)$$

を満たせばよい。

【0078】

一方、3番目の再生用スピーカ52cと4番目の再生用スピーカ52dからの電力の合計が、元の2chステレオの相関信号の電力と等しくなるように $g_1$ 、 $g_2$ を正規化すると、

$$g_1^2 + g_2^2 = 1 + [\alpha^{(k)}]^2 \quad (33)$$

となる。

【0079】

これらを連立させることで、数式(34)が求められる。

【数13】

$$g_1 = \frac{1}{\sqrt{1+q^2}} \cdot \sqrt{1+[\alpha^{(k)}]^2}, \quad g_2 = \frac{q}{\sqrt{1+q^2}} \cdot \sqrt{1+[\alpha^{(k)}]^2}$$

$$\text{ただし、} q = \frac{\sin \varphi_0 + \sin \varphi^{(k)}}{\sin \varphi_0 - \sin \varphi^{(k)}} \quad (34)$$

【0080】

この数式(34)に上述の $\alpha^{(k)}$ 、 $\varphi_0$ を代入することによって、 $g_1$ 、 $g_2$ を算出する。このようにして算出したスケール係数に基づき、上述したように3番目の再生用スピーカ52cには $g_1 \cdot \text{est}(S(k))$ の音声信号を、4番目の再生用スピーカ52dからは $g_2 \cdot \text{est}(S(k))$ の音声信号を割り当てる。そして、これも上述したように、無相関信号は両端の再生用スピーカ52a, 52eに割り当てられる。すなわち、1番目の再生用スピーカ52aには $\text{est}(N_L(k))$ を、5番目の再生用スピーカ52eには $\text{est}(N_R(k))$ を割り当てる。

【0081】

この例とは異なり、もし合成音像の推定方向が1番目と2番目の再生用スピーカの間であった場合には、1番目の再生用スピーカには $g_1 \cdot \text{est}(S(k))$ と $\text{est}(N_L(k))$ の両方が割り当てられることになる。また、もし合成音像の推定方向が4番目と5番目の再生用スピーカの間であった場合には、5番目の再生用スピーカには $g_2 \cdot \text{est}(S(k))$ と $\text{est}(N_R(k))$ の両方が割り当てられることになる。

【0082】

上述のような処理を、ステップS31a, S31bのループにより全ての線スペクトルについて行う。例えば、256点の離散フーリエ変換を行った場合は1~127番目の線スペクトルまで、512点の離散フーリエ変換を行った場合は1~255番目の線スペクトルまで、セグメントの全点(1024点)について離散フーリエ変換を行った場合は1

10

20

30

40

50

～ 5 1 1 番目の線スペクトルまで、となる。

【 0 0 8 3 】

その結果、再生用スピーカの数  $J$  (この例では  $J = 5$ ) とすると、各再生用スピーカ (出力チャネル) に対する周波数領域の出力音声信号  $Y_1(k)$ 、 $\dots$ 、 $Y_J(k)$  が求まる。これら出力が、図 2 における分離抽出部 2 2 の逆変換部 2 3 への出力結果となる。以上のようにして図 2 における分離抽出部 2 2 の処理がなされる。

【 0 0 8 4 】

次に、図 2 における逆変換部 2 3 での処理がなされる。逆変換部 2 3 は、分離部 2 2 a で抽出された相関信号に対して、離散フーリエ逆変換を施す。逆変換部 2 3 は、分離部 2 2 a で抽出された相関信号の代わりに、( a 1 ) その相関信号及び無相関信号 (その相関信号を除く信号) に対して、もしくは ( a 2 ) その相関信号から生成された音声信号、または ( a 3 ) その相関信号及びその無相関信号から生成された音声信号に対して、離散フーリエ逆変換を施すようにしてもよい。

【 0 0 8 5 】

具体的に逆変換部 2 3 の処理を例示すると、逆変換部 2 3 では、分離抽出部 2 2 から出力された各出力チャネルを離散フーリエ逆変換することによって、時間領域の出力音声信号  $y_j(m)$  を求める。ここで、 $DFT^{-1}$  は離散フーリエ逆変換を表す。

【 0 0 8 6 】

$$y_j(m) = DFT^{-1}(Y_j(k)) \quad (1 \leq j \leq J) \quad (35)$$

ここで、数式 (3)、(4) で説明したように、離散フーリエ変換した信号は、窓関数乗算後の信号であったため、逆変換して得られた信号  $y_j(m)$  も窓関数が乗算された状態になっている。したがって、そうして得られた信号に、数式 (2) に示す窓関数を再度乗算し、1 つ前に処理したセグメントの先頭から 1 / 4 セグメント長ずつずらしながら出力バッファに加算していくことにより変換後のデータを得る。

【 0 0 8 7 】

図 2 における音声出力信号生成部 2 4 では、逆変換部 2 3 で生成した各仮想音源の音声信号から、上述の非特許文献 1 に記載の技術に則って、各スピーカに対する音声信号を生成する。ここでは、各仮想音源の音声信号から、各スピーカの配置に応じて各スピーカに対する遅延量や出力レベルを決定し、それに応じた音声信号が生成される。このようにして、音声信号処理部 1 3 は、マルチチャネル再生方式のチャネルの入力音声信号をスピーカ群により再生させるために変換することができる。ただし、ここで挙げた例のように再生用スピーカのそれぞれに仮想音源を設定した場合には、音声出力信号生成部 2 4 におけるこのような処理は不要となる。仮想音源を再生用スピーカと一致させる場合でもそうでない場合でも、音声出力信号生成部 2 4 は、必要に応じて、例えば入力音声信号のうち使用しなかったチャネルの音声信号を加算する処理など、他の処理を行うように構成しておけばよい。

【 0 0 8 8 】

以上のようにして音声再生処理がなされる。上述したように、本発明では音像表示処理もなされる。音像表示処理に関し、図 1 における表示制御部 1 8 に音声信号処理部 1 3 より出力する音像方向情報について、説明する。上述したように、音像方向情報は、音像方向と信号電力値を含み、分離抽出部 2 2 により生成される。

【 0 0 8 9 】

具体的には、音像方向は図 4 に として示した推定方向の角度 (推定角度) で例示したものであり、信号電力値は数式 (14) に示した  $P_s(k)$  で例示したものであり、それらはいずれも、線スペクトル毎に (つまり各線スペクトルの相関信号のそれぞれについて) 生成された値である。なお、特に言及しなかったが、上述したような信号電力値を算出する処理も、図 2 の分離抽出部 2 2 のうち算出部 2 2 b が担うことになる。算出部 2 2 b は、各線スペクトルの相関信号について算出した推定角度と信号電力値とを、表示制御部 1 8 に出力する。

【 0 0 9 0 】

10

20

30

40

50

次に、表示制御部 18 及び表示部 19 での処理について、図 6 及び図 7 を参照しながら説明する。図 6 は、図 1 の音声信号再生装置 10 において主に加算部 18 a で実行される音像表示処理の一例を説明するためのフロー図である。図 7 は、図 1 の音声信号再生装置 10 における表示部 19 の一例を示す図である。

【0091】

表示制御部 18 は、線スペクトル毎にステップ S 6 2 , S 6 3 の処理を実行する（ステップ S 6 1 a , S 6 1 b ）。ステップ S 6 2 では、加算部 18 a が、対象となる線スペクトルの相関信号について、その推定角度 を基に音像方向グループを特定する。音像方向グループというのは、とりうる音像方向の角度 をグループに分けたもので、以下、単にグループと呼ぶ。 のとりうる範囲は、想定している聴取位置での最右端スピーカから最左端スピーカまでの見開き角の範囲であり、予め定められた範囲である。

10

【0092】

例えばこの範囲を  $-30^{\circ} \sim 30^{\circ}$  とし、表示部 19 で音像方向を図 7 のように 15 本の棒（バー）グラフで表示する場合には、この範囲を 15 のグループに分ければよい。そのとき、第一のグループが  $-30^{\circ}$  以上  $-26^{\circ}$  未満、第二のグループが  $-26^{\circ}$  以上  $-22^{\circ}$  未満、というように  $4^{\circ}$  ずつのグループに分ける。ここでは均等に分割する例を示すが、均等でなくてもよく、表示部 19 での表示の仕方に合わせてグループに分割すればよい。

【0093】

ステップ S 6 3 では、加算部 18 a が、ステップ S 6 2 で特定したグループにその相関信号についての信号電力値を加算する。より具体的には、加算部 18 a は、予め用意しておいたグループ毎のパッファに相関信号の電力値  $P_s(k)$  を加算する。ここでは、単純に加算する例を示したが、等ラウドネス曲線（ISO 226:2003）に基づいて電力値を重み付けするなどして加算すれば、聴感上受ける音圧に、この数値を近づけることが可能となる。つまり、加算値として単なる総和を例に挙げたが、これに限らず、例えば線スペクトル毎（または周波数帯域毎）に重み付けを行った重み付け加算値を採用することもできる。

20

【0094】

ステップ S 6 2 , S 6 3 を全ての線スペクトルについて繰り返すことにより、各グループ毎に信号電力値が加算され、グループ毎の信号電力値の総和が算出される。

30

算出された加算値は、表示部 19 で表示されることになる。この表示のための制御も表示制御部 18 が行う。

【0095】

表示制御部 18 は、全ての線スペクトルについて電力値加算が終了した後、各グループについて、ステップ S 6 5 , S 6 6 の処理を実行する（ステップ S 6 4 a , 6 4 b ）。ステップ S 6 5 では、表示制御部 18 が、グループ毎に信号電力値の加算値を表示するための値に変換する。つまり、グループ毎の信号電力値の総和を、表示するための値に変換する。ステップ S 6 6 では、表示制御部 18 が、変換後の値を表示させるための制御信号を生成し、表示部 19 に出力する。表示部 19 は、その制御信号に従い、音像を表示させる。

40

【0096】

図 7 を参照して、表示部 19 に表示される音像の具体例について説明する。

図 7 で例示するテレビ装置 70 は、表示装置の一例であり、この表示装置は、音声信号を再生しスピーカアレイ 17 に渡す再生部（図示せず）と、映像信号が示す映像を表示する映像表示部（液晶パネルや有機エレクトロルミネッセンスパネルなど）の一例としての表示パネル 71 とを備えている。表示パネル 71 は、表示部（音像表示部）19 の一例でもある。実際、映像表示部の画面には映像が表示可能であり、その映像に各グループについての加算値を示す情報、つまり音像を示す情報を含めればよい。もしくは、この情報を OSD（On Screen Display）画像として映像表示部の画面に重畳して表示させることもできる。よって、ステップ S 6 6 における制御信号の一例としては、このような映像また

50



はOSD画像を表示させるための信号が挙げられる。

【0097】

テレビ装置70には、音声出力のために、図1で示した複数のスピーカ16（スピーカ16a, 16b, 16c, . . .）が配列されたスピーカアレイ17が内蔵されている。ただし、スピーカアレイ17は外部に接続されていてもよい。外部に接続される場合、スピーカアレイ17は、テレビ台（テレビボード）に埋め込むこともでき、またサウンドバーと呼ばれるテレビ装置の下に置く一体型のスピーカシステムとして埋め込むこともできる。なお、この例では、スピーカアレイ17を表示パネル71の下側に設けているが、上側、あるいは上側及び下側に設けるなどしてもよい。

【0098】

表示パネル71には、加算値を示す情報が棒グラフ72として表示されている。棒グラフ72では、複数のグループのそれぞれに対応する位置（この例では各スピーカ16a, 16b, 16c, . . .の設置位置）に各グループについての加算値を示す情報73a, 73b, 73c, . . .を表示したグラフである。つまり、棒グラフ72は、複数のグループのそれぞれに対応する位置を一方の軸に配し、各グループについての加算値（各グループに含まれる関連信号の信号電力値）に対応する長さの棒を他方の軸とするグラフとなっている。

【0099】

1つのグループに含まれる関連信号を出力するための信号電力値の加算値は音圧の強さに対応している。したがって、このような棒グラフ72は、棒の高さの変化によりその棒の存在する方向からの音圧の強さを表現することができる。

【0100】

このように、表示パネル71にグループ毎の加算値を示す情報を棒グラフ形式で表示させたい場合には、ステップS65において、グループ毎の加算値（信号電力値の加算値）をそれぞれの棒の高さに変換し、その加算値が大きいほど、棒が高く描画されるようにすればよい。

【0101】

グループ毎の加算値は、上述のように処理セグメント毎に算出されるため、処理セグメント毎に描画を実行することによって、複数のスピーカ16を有するスピーカアレイ17から出力される音と同期して、棒グラフ72が動くことになる。これにより、音像表示を音の再生に合わせて変化させることができ、聴取できる音像の方向とその視覚的効果の場所が一致し（聴覚と知覚とで音像位置が一致し）、受聴者に没入感を与えることができる。特に、棒グラフ72のような棒グラフ形式での表示によれば、受聴者が一目で音像を認識できるため、他の音像表示方法に比べて受聴者の没入感をより増すことができる。

【0102】

また、棒グラフにおける棒の数（つまりグループの数）は、図7で例示したように基本的にスピーカ群の解像度（すなわち設置されるスピーカの個数）だけ用意しておく、聴取できる音像と視認できる棒とを一致させることができるため好ましい。ただし、スピーカの数とグループ数とは異ならせておいてもよい。

【0103】

また、以上では、本発明に係る音声信号再生装置が、スピーカアレイから音声信号を再生する装置であることを前提として説明しているが、例えば、単に左右の2つのスピーカから出力する場合であっても、音像表示についてはここで説明するように実行し、音声再生については対象となる2つの音声信号を左右のそれぞれのスピーカから出力するだけでよい、同様に適用できる。その他、例えば5.1ch用のスピーカシステムからマルチチャンネル再生方式で出力する場合など、他の再生方式で出力する場合にも同様に適用できる。

【0104】

次に、図8を参照しながら、表示部（音像表示部）19の他の構成例について説明する。図8は、図1の音声信号再生装置10における表示部19の他の例を示す図である。

10

20

30

40

50

ここで例示する表示部 19 は、音声信号再生装置 10 が映像表示部を備えている場合、備えていない場合に拘わらず、複数の発光部を有し、これらの発光部で音像表示を行う。なお、映像表示部を備えないような音声信号再生装置 10 としては、例えば様々なオーディオシステムが挙げられる。

【0105】

図 8 で例示するスピーカアレイ 17 は、各スピーカ 16 a, 16 b, 16 c, . . . の傍にそれぞれ 8 つの LED (Light Emitting Diode) 81 a, 81 b, 81 c, . . . を配設している。つまり、上記発光部の 1 つを 8 つの LED 81 a で構成している。ただし、上記発光部に対応する LED 81 a の個数は 8 つに限らず、例えば単に 1 つだけでもよい。また、LED に限らず他種の発光部を設けることもできる。

10

【0106】

また、上記発光部は基本的にスピーカ群の解像度（すなわち設置されるスピーカの個数）だけ用意しておく、聴取できる音像と発光とを一致させることができるため好ましいが、これに限ったものではない。スピーカアレイ 17 の場合、その長手方向に沿って上記発光部を複数並べておけばよい。ただし、スピーカアレイ 17 の長手方向に沿って並べられたものに限ったものでもなく、グループに対応する位置に配置されていればよい。また、1 つのスピーカに対する LED の配置も図 8 で例示するようにスピーカを取り囲むような配置に限ったものではない。なお、各発光部をグループ毎に複数並べて棒グラフ形式で表示させることも可能である。

20

【0107】

このような構成の複数の発光部を用い、表示部 19 は、複数のグループのそれぞれに対応する位置の発光部の発光色を、各グループについての加算値を示す情報に応じて変化させる。このように、表示部 19 は、LED の色変化によりその LED の存在する方向からの音圧の強さを表現してもよい。

【0108】

例えば、音像がスピーカ 16 b の方向のみに存在する場合には、対応する 8 つの LED 81 b の色を、例えば加算値が弱い場合には青色を、強い場合には赤色を表示させるなどして、音と同期して色を変えればよい。また、スピーカ 16 b の方向から強い音（加算値の大きい音）が出力され、スピーカ 16 c の方向から弱い音（加算値の小さい音）が出力され、他のスピーカの方向からは音出力されなかった場合には、8 つの LED 81 b の色を赤色にし、8 つの LED 81 c の色を青色にするなどして、音と同期して色を変えればよい。

30

【0109】

また、表示部 19 は、複数のグループのそれぞれに対応する位置の LED の発光強度（発光輝度）を、各グループについての加算値を示す情報に応じて変化させるようにしてもよい。加算値を LED の明るさに変換するようにすれば、音像位置に同期して LED の明るさが変化することになる。つまり、LED の強弱変化によりその LED の存在する方向（音像位置）からの音圧の強さを表現してもよい。

【0110】

例えば、音像がスピーカ 16 b の方向のみに存在する場合には、対応する 8 つの LED 81 b の全てを加算値に応じた輝度で発光させるか、もしくは 8 つの LED 81 b のうち加算値に応じた数の LED 81 b を発光させるなどして、音と同期して発光輝度を変えればよい。また、スピーカ 16 b の方向から強い音出力され、スピーカ 16 c の方向から弱い音出力され、他のスピーカの方向からは音出力されなかった場合には、例えば 8 つの LED 81 b を発光させ、1 つの LED 81 c を発光させるなどして、音と同期して発光輝度を変えればよい。

40

【0111】

このような複数の発光部を用いた音像表示方法を採用した場合にも、音像表示を音の再生に合わせて変化させることができ、聴取できる音像の方向とその視覚的效果の場所が一致し、受聴者に没入感を与えることができる。特に、音の強さに応じて発光色を変化させ

50

たり発光強度を変化させたりすることで、例えば音楽ビデオを再生している場合や音楽を再生している場合に、ライブ会場での感覚を受聴者に与えることができる。

【0112】

以上の例では、一方向から受聴者が音像位置を視認する場合、つまり音像表示が一方向（前方向）にのみなされる場合について説明した。しかし、例えば5.1~7.1chの入力音声信号に対しては、仮想音源を設定して三方や四方に用意したアレイスピーカから音声を出力するとともに、各方向において図8で例示したように発光部を変化させて音像を視認させることもできる。このような例を以下に説明する。

【0113】

まず、このような構成における、音声信号の分離、音像方向及び信号電力値の算出などの処理について、再び図1及び図2を併せて参照しながら説明する。ここでは、基本的に上述した処理例と異なる点を説明し、同様の処理を行う部分についてはその説明を省略する。

【0114】

まず、音声信号抽出部12では、得られた信号から音声信号を分離、抽出するが、この例では得られた信号は5.1chとし、そのうちのL、R、C、L<sub>s</sub>、R<sub>s</sub>の5つのチャンネルの音声信号を音声信号処理部13に出力する。残りのLFEのチャンネルの音声信号は、別途設けた遅延処理部（図示せず）で前述の信号L、R、C、L<sub>s</sub>、R<sub>s</sub>を音声信号処理部13で処理するのに要する時間分遅延させ、D/Aコンバータ14に出力する。

【0115】

本例における音声信号処理部13は、マルチチャンネル再生方式の5つ以上のチャンネルの入力音声信号を、スピーカ群により再生させるために変換する。ここで複数のスピーカ16は、より適切な音像を表現できるようにするため、その入力音声信号のチャンネル数以上の数のスピーカでなることが好ましい。この場合、複数のスピーカ16は、例えば5.1chのうちの5chを入力音声信号とした場合には、そのチャンネル数が後述のダウンミックス処理により1つ少なくなり、その少なくなったチャンネル数より多い数（換言すれば入力音声信号のチャンネルの数と同じかそれより多い数）のスピーカでなる。

【0116】

具体的に説明すると、音声信号処理部13では、得られた5チャンネル信号から、入力音声信号とは異なるマルチチャンネルの音声信号を生成する。つまり、音声信号処理部13では、入力音声信号を別のマルチチャンネルの音声信号に変換する。入力音声信号のチャンネル数以上のスピーカから出力させるためには、変換後のマルチチャンネルのチャンネル数は、入力されたチャンネル数（この例では5つ）以上とすることが好ましい。ただし、仮想音源用のチャンネルとすることで入力されたチャンネル数より小さくても複数のスピーカ（スピーカ群）16への割り当てはできる。以下、仮想音源の数分の信号を生成するものとして説明する。

【0117】

音声信号処理部13は、生成した音声信号をD/Aコンバータ14に出力する。仮想音源の数は、ある一定以上の数があれば予め決めておいても性能上差し支えはないが、仮想音源数が多くなるほど演算量も多くなる。そのため実装する装置の性能を考慮してその数を決定することが望ましい。ここで説明する例では、その数を16として、後述の図17のように円周上に仮想音源を配置する場合について説明するが、この例に限ったものではない。

【0118】

D/Aコンバータ14では得られた信号をアナログ信号に変換し、それぞれの信号を増幅器15及びサブウーファー用増幅器（図示せず）に出力する。各増幅器15では入力されたアナログ信号を拡声し各スピーカ16に伝送し、この拡声されたアナログ信号が各スピーカ16から空間中に音として出力される。サブウーファー用増幅器では入力されたLFE用のアナログ信号を拡声しサブウーファー（図示せず）に伝送し、この拡声されたアナログ信号がサブウーファーから空間中に音として出力される。

10

20

30

40

50

## 【0119】

本例における変換部21は、入力音声信号である5つ以上のチャンネルのうち1つの特定チャンネルを、その1つの特定チャンネルに隣合う2つのチャンネルにダウンミックスするダウンミックス部（図示せず）を有する。ここでは、特定チャンネルとしてCチャンネルの信号を採用した例を挙げて説明する。ダウンミックス部は、Cの信号のゲインにダウンミックス係数をかけたものを、RとLの2チャンネルの信号それぞれに加算する。ダウンミックス係数は時間的に変化しない、0より大きい実数であり、例えば $1/2$ や $1/\sqrt{2}$ などの値をとる。

## 【0120】

変換部21は、ダウンミックス部から出力された2つのチャンネルと、 $L_S$ 、 $R_S$ とを合わせた4つの入力信号それぞれを、1セグメントの $1/4$ の長さの音声データ分読み出す。

読み出した256点の音声データはバッファに蓄えられる。このバッファは、直前の1セグメント分の音声信号波形を保持しておけるようになっており、それより過去のセグメントは捨てていく。直前の $3/4$ セグメント分のデータ（768点）と最新の $1/4$ セグメント分のデータ（256点）を繋げて1セグメント分の音声データを作成し、数式(2)の窓関数を乗算する。すなわち、全ての標本データは窓関数演算に4回読み込まれることになる。

## 【0121】

ダウンミックス部からの出力信号をそれぞれ $x_L(m)$ 、 $x_R(m)$ 、 $x_{L_S}(m)$ 、 $x_{R_S}(m)$ とすると、変換部21において、窓関数乗算後の音声信号 $x_L(m)$ 、 $x_R(m)$ 、 $x_{L_S}(m)$ 、 $x_{R_S}(m)$ は、

$$\begin{aligned} x_L(m) &= w(m) x_L(m), \\ x_R(m) &= w(m) x_R(m), \\ x_{L_S}(m) &= w(m) x_{L_S}(m), \\ x_{R_S}(m) &= w(m) x_{R_S}(m) \end{aligned} \quad (36)$$

と計算される。

## 【0122】

変換部21は、そうして得られた音声データを、次の数式(37)のように離散フーリエ変換し、周波数領域の音声データを得る。つまり、変換部21は、上記特定チャンネルを除いた、ダウンミックス後の2つのチャンネルと残りのチャンネル（の音声データ）に対し、離散フーリエ変換を施す。ここで、DFTは離散フーリエ変換を表し、 $k$ は自然数で、 $0 < k < M/2$ である。 $X_L(k)$ 、 $X_R(k)$ 、 $X_{L_S}(k)$ 、 $X_{R_S}(k)$ は複素数となる。

$$\begin{aligned} X_L(k) &= \text{DFT}(x_L(m)), \\ X_R(k) &= \text{DFT}(x_R(m)), \\ X_{L_S}(k) &= \text{DFT}(x_{L_S}(m)), \\ X_{R_S}(k) &= \text{DFT}(x_{R_S}(m)) \end{aligned} \quad (37)$$

## 【0123】

分離抽出部22は、線スペクトル毎に、変換部21で変換された4つ以上のチャンネルにおける隣合う2つのチャンネルの組み合わせのうち着目した組み合わせについて、相関信号と無相関信号を分離、抽出する。なお、説明の簡略化のため、本例では分離抽出部22の処理を分離部22aでの処理と算出部22bでの処理とに分けずに説明する。

## 【0124】

本例における分離抽出部22は、線スペクトル毎に（つまり周波数毎に）、上記隣合う2つのチャンネルの組み合わせのうち、どの組み合わせに着目するかを、各チャンネルの電力の大きさに基づいて判定する。分離抽出部22においては、上述したように線スペクトル毎でなくても周波数領域（小帯域）毎に分離、抽出してもよく、その場合には判定も小帯域毎に行う。

## 【0125】

分離抽出部22における分離抽出処理の具体的な内容を、再度図3を用いて説明する。

ここで説明する 5.1ch の例では、離散フーリエ変換が数式(37)のように 4 つのチャンネルに対して施されている。よって、分離抽出部 22 は、変換部 21 で離散フーリエ変換後の 4 つのチャンネルの音声信号について、線スペクトル毎にステップ S32 ~ S34 の処理を行うが、それに先立ち、線スペクトル毎に着目ペア判定処理(図 3 では図示せず)を実行する(ステップ S31a, S31b)。具体的に個々の処理について説明する。

#### 【0126】

上記の着目ペア判定処理では、元のマルチチャンネル再生方式で再生を想定しているスピーカ群のうち、スピーカ配置が隣合う 2 つの入力信号のいずれに着目するかを判定する処理を実行する。

#### 【0127】

ここで、隣合うスピーカ配置の定義について、図 9 及び図 10 を参照して説明する。図 9 は、5.1ch サラウンドシステムのスピーカ群のうち LFE を除いた 5 つのスピーカの配置例を示す図で、図 10 は、図 9 の配置例におけるダウンミックス後の出力対象スピーカの配置例を示す図である。

#### 【0128】

図 9 で示すように、5.1ch サラウンドシステムでは、LFE 用のスピーカを除き、左フロントスピーカ 91、右フロントスピーカ 92、センタースピーカ 93、左サラウンドスピーカ 94、右サラウンドスピーカ 95 が存在する。上述の、スピーカ配置が隣合うか否かを判定する際、上述のダウンミックス部で L と R それぞれに加算された C の信号は、対象から外す。すなわち、ダウンミックス部で他の信号に加算される対象となった C 信号に係るセンタースピーカ 93 を除き、図 10 に示す 4 つのスピーカ(左フロントスピーカ 91、右フロントスピーカ 92、左サラウンドスピーカ 94、右サラウンドスピーカ 95)の配置を考える。

#### 【0129】

図 10 に示すように、左フロントスピーカ 91 と隣合うのは右フロントスピーカ 92 と左サラウンドスピーカ 94 となる。同様に、左サラウンドスピーカ 94 と隣合うのは左フロントスピーカ 91 と右サラウンドスピーカ 95、右フロントスピーカ 92 と隣合うのは左フロントスピーカ 91 と右サラウンドスピーカ 95、右サラウンドスピーカ 95 と隣合うのは右フロントスピーカ 92 と左サラウンドスピーカ 94、となる。したがって、この場合、図 10 中に両矢印で示したように組合せ(ペア)が 4 つ存在する。

#### 【0130】

着目ペア判定処理では、これらの各ペアのうち、どのペアに着目するかを判定を図 11 で例示するように行う。図 11 は、分離抽出部 22 での着目ペア判定処理の一例を説明するためのフロー図である。

#### 【0131】

まず、各ペアについて、線スペクトル毎に、次の数式(38)のように電力の和  $P_1 \sim P_4$  を算出する(ステップ S111)。ここで、対象となる線スペクトルの音声信号 X のゲイン(振幅)を G とすると、 $P(X)$  は、音声信号 X の電力(音圧に相当)を表し、ゲイン G の二乗値で表現できる。

$$\begin{aligned} P_1(k) &= P(X_L(k)) + P(X_R(k)), \\ P_2(k) &= P(X_{LS}(k)) + P(X_L(k)), \\ P_3(k) &= P(X_{RS}(k)) + P(X_{LS}(k)), \\ P_4(k) &= P(X_R(k)) + P(X_{RS}(k)) \end{aligned} \quad (38)$$

#### 【0132】

次に、各ペアの線スペクトルに対し、両方のチャンネル間の正規化相関係数を次の数式(39)で求めることで、相関係数を取得する(ステップ S112)。なお、 $\text{Re}\{BB\}$ 、 $\text{Im}\{CC\}$  はそれぞれ「BB」の実部、「CC」の虚部を表す。

10

20

30

40

【数 1 4】

$$\begin{aligned}
 d^{(k)}_1 &= \frac{|\operatorname{Re}\{X_L(k)\} \cdot \operatorname{Re}\{X_R(k)\}| + |\operatorname{Im}\{X_L(k)\} \cdot \operatorname{Im}\{X_R(k)\}|}{\sqrt{|X_L(k)|^2 |X_R(k)|^2}}, \\
 d^{(k)}_2 &= \frac{|\operatorname{Re}\{X_{LS}(k)\} \cdot \operatorname{Re}\{X_L(k)\}| + |\operatorname{Im}\{X_{LS}(k)\} \cdot \operatorname{Im}\{X_L(k)\}|}{\sqrt{|X_{LS}(k)|^2 |X_L(k)|^2}}, \\
 d^{(k)}_3 &= \frac{|\operatorname{Re}\{X_{RS}(k)\} \cdot \operatorname{Re}\{X_{LS}(k)\}| + |\operatorname{Im}\{X_{RS}(k)\} \cdot \operatorname{Im}\{X_{LS}(k)\}|}{\sqrt{|X_{RS}(k)|^2 |X_{LS}(k)|^2}}, \\
 d^{(k)}_4 &= \frac{|\operatorname{Re}\{X_R(k)\} \cdot \operatorname{Re}\{X_{RS}(k)\}| + |\operatorname{Im}\{X_R(k)\} \cdot \operatorname{Im}\{X_{RS}(k)\}|}{\sqrt{|X_R(k)|^2 |X_{RS}(k)|^2}} \quad (39)
 \end{aligned}$$

【0 1 3 3】

この正規化相関係数  $d^{(k)}_1 \sim d^{(k)}_4$  は両方のチャンネルの音声信号にどれだけ相関があるかを表すものであり、0 から 1 の間の実数の値をとる。全く同じ信号同士であれば 1、そして全く無相関の信号同士であれば 0 となる。ここで、両方のチャンネルの音声信号の電力が 0 である場合、その線スペクトルに関して相関信号と無相関信号の抽出は不可能とし、処理を行わず次の線スペクトルの処理に移ることとする。また、いずれか片方の電力が 0 である場合、数式(39)では演算不可能であるが、正規化相関係数  $d^{(k)}_r = 0$  とし、その線スペクトルの処理を続行する。ただし、 $r = 1 \sim 4$  である。

【0 1 3 4】

次に、各ペアに対し、数式(38)と数式(39)で求めた値を乗算した  $r^{(k)}$  を、次のように計算する。

$$\begin{aligned}
 1^{(k)} &= P_1^{(k)} \times d^{(k)}_1, \\
 2^{(k)} &= P_2^{(k)} \times d^{(k)}_2, \\
 3^{(k)} &= P_3^{(k)} \times d^{(k)}_3, \\
 4^{(k)} &= P_4^{(k)} \times d^{(k)}_4
 \end{aligned} \quad (40)$$

【0 1 3 5】

そして、 $r^{(k)}$  が最大値となる場合の  $r$  である  $r_{\max}$  を求め、着目すべきペアを決定する(ステップ S 1 1 3)。例えば、 $r_{\max} = 1$  の場合、図 1 0 における左フロントスピーカ 9 1 と右フロントスピーカ 9 2 が着目すべきペアとなる。以下では、ステップ S 1 1 3 の判定の結果(すなわち着目ペア判定処理の結果)が、左フロントスピーカ 9 1 と右フロントスピーカ 9 2 が着目すべきペアであった場合を例に挙げて説明するが、他のペアが着目すべきペアであっても同様である。

【0 1 3 6】

このように、分離抽出部 2 2 は、線スペクトル毎に、隣合う 2 つのチャンネルの組み合わせのうち、どの組み合わせに着目するかを、隣合う 2 つのチャンネルの組み合わせについての電力の大きさと相関係数との乗算結果に基づいて判定することが好ましい。無論、分離、抽出を小帯域毎に行う場合にはこの判定も小帯域毎に行うことになる。

【0 1 3 7】

また、数式(40)において、 $r^{(k)}$  の計算の際、 $d^{(k)}_r = 1$  として、相関係数の計算を省略することもできる。 $d^{(k)}_r = 1$  として相関係数の計算を省略することは、隣合う 2 つのチャンネルの電力の大きさの和に基づいて、その和が大きいものを着目すべきペアに決定することを意味する。このように分離抽出部 2 2 では、相関係数を用いなくても、線スペクトル毎に、隣合う 2 つのチャンネルの電力の大きさに基づいて判定を行えばよい。ただし、相関係数も併せて判定に用いることにより、判定の精度を上げることができ

る。

#### 【 0 1 3 8 】

その他の判定の例を挙げる。最終的に  $r(k)$  の最大値を求めることができればよい  
ため、 $d^{(k)}_r = 1$  との省略を行う場合もその省略を行わない場合にも、 $P_r(k)$  と  
して、隣合う 2 つのチャンネルの電力の積を採用することができる。さらに別の例として、  
一番大きな電力をもつチャンネルと、それに隣合う 2 つのチャンネルのうち大きい電力をもつ  
チャンネルとで構成されるペアに着目するように判定することもできる。

#### 【 0 1 3 9 】

着目ペア判定処理の結果、電力が最大となり、なおかつ相関係数が大きいペアについて  
のみ、つまり着目ペアについてのみ、後述するように相関信号・無相関信号に分離した上  
での仮想音源への割り当てを行う（ステップ S 3 2 ~ S 3 4）。この処理は、同じ線スペ  
クトルあるいは小帯域に属する音像は 1 つである、という仮定を置くことを意味しており  
、その仮定において、ステップ S 3 2 ~ S 3 4 は正確な音像位置を再現するための変換処  
理であると言える。なお、上記の判定を電力の大きさのみ実行した場合にも同様に、着目  
ペアについてのみステップ S 3 2 ~ S 3 4 の処理を行えばよい。

#### 【 0 1 4 0 】

まず、上述の正規化相関係数  $d^{(k)}_1$  を用いて、両方のチャンネルの音声信号から相関  
信号と無相関信号をそれぞれ分離抽出するための変換係数を求め（ステップ S 3 2）、取  
得したそれぞれの変換係数を用いて、両方のチャンネルの音声信号から相関信号と無相関信  
号を分離抽出する（ステップ S 3 3）。相関信号及び無相関信号は、いずれも推定した音  
声信号として抽出すればよい。

#### 【 0 1 4 1 】

具体的には、 $d^{(k)}_1$  を  $d^{(k)}$  とおくことで、図 1 ~ 図 8 を参照しながら説明した  
例と同様にステップ S 3 2、S 3 3 での処理を行えばよく、その説明を省略する。ただし  
、本例では、左右のチャンネルの代わりに着目ペアの両方のチャンネルについて処理が実行さ  
れる。つまり、本例で使用するモデルでは、両方のチャンネルそれぞれの信号は、無相関信  
号と相関信号から構成され、相関信号については、両方のチャンネルからゲインのみ異なる  
信号波形（つまり同じ周波数成分からなる信号波形）が出力されるものとする。ここで、  
上述したように、ゲインは、信号波形の振幅に相当し、音圧に関連する値である。そして  
、このモデルでは、両方のチャンネル信号から出力される相関信号によって合成される音像  
は、その相関信号の両方それぞれの音圧のバランスによって方向が決定されるものとする  
。

#### 【 0 1 4 2 】

本例では、ステップ S 3 4 において再生用スピーカに直接割り当てを行う代わりに、ま  
ず仮想音源への割り当てを行う。この説明については、「再生用スピーカ」を「仮想音源  
」に読み換えることで容易に理解できるため、基本的に省略する。また、上述したように  
この割り当ての前処理として、線スペクトル毎に推定した相関信号によって生成される合  
成音像の方向が推定される。

#### 【 0 1 4 3 】

なお、上記の読み換えにより、仮想音源が図 5 の再生用スピーカと同様に直線上に並ん  
でいる場合については説明できる。しかし、図 1 2 で受聴者と左右のスピーカと仮想音源  
との位置関係の他の例を示すように、仮想音源 1 2 1 a ~ 1 2 1 e を、図 9 におけるスピ  
ーカ配置の円（図 1 2 中の破線で示す円）と同心円の円弧上に並べた場合についても、例  
えばその中心点に受聴者 1 2 2 を配置することによって、同様に仮想音源への信号割り当  
てが可能である。

#### 【 0 1 4 4 】

以上、上記着目ペア判定処理で図 1 0 における左フロントスピーカ 9 1 と右フロントス  
ピーカ 9 2 のペアに着目した場合について、説明してきた。この仮想音源割り当てにより  
、左フロントスピーカ 9 1 と右フロントスピーカ 9 2 の音声信号の k 番目の線スペクトル  
については図 5 の再生用スピーカ 5 2 a ~ 5 2 e と同様に直線上に並んでいる仮想音源、

あるいは、図 12 の仮想音源 121a ~ 121e に割り当てられたが、図 10 における左サラウンドスピーカ 94 と右サラウンドスピーカ 95 の音声信号それぞれの k 番目の線スペクトルについては、まだ割り当てが行われていない。

#### 【0145】

それらの割り当てについて図 13 を参照して説明する。図 13 は、受聴者と左右のスピーカ及び左右のサラウンドスピーカと仮想音源との位置関係の例を説明するための模式図である。それらの割り当ては、図 13 に示すように、左サラウンドスピーカ 94 の音声信号の k 番目の線スペクトルは、受聴者 132 から見たときの左サラウンドスピーカ 94 と同じ方向の仮想音源 131a に割り当てる。同様に、右サラウンドスピーカ 95 の音声信号の k 番目の線スペクトルは、受聴者 132 から見たときの右サラウンドスピーカ 95 と同じ方向の仮想音源 131b に割り当てる。

10

#### 【0146】

以上のようにして、ステップ S34 における、k 番目の線スペクトルについて、4 つのチャンネルの、仮想音源への割り当てが行われる。ここでは、着目ペア判定処理の結果、図 10 における左フロントスピーカ 91 と右フロントスピーカ 92 に着目して相関信号と無相関信号の分離処理を行ったが、着目ペア判定処理の結果、着目する対象が異なることもある。そのような場合について、図 14 ~ 図 16 を参照して説明する。図 14 ~ 図 16 は、受聴者と左右のスピーカ及び左右のサラウンドスピーカと仮想音源との位置関係の他の例を説明するための模式図である。

#### 【0147】

20

着目ペア判定処理の結果、例えば左フロントスピーカ 91 と左サラウンドスピーカ 94 に着目するという結果になった場合は、図 14 に示すように、左フロントスピーカ 91 と左サラウンドスピーカ 94 についての k 番目の線スペクトルの相関信号と無相関信号が仮想音源 141a ~ 141e に割り当てられ、右フロントスピーカ 92 の k 番目の線スペクトルが仮想音源 141f に、右サラウンドスピーカ 95 の k 番目の線スペクトルが仮想音源 141g に、それぞれ割り当てられる。

#### 【0148】

着目するペアが他のペアとなった場合にも同様である。左サラウンドスピーカ 94 と右サラウンドスピーカ 95 に着目するという結果になった場合、図 15 に示すように、左サラウンドスピーカ 94 と右サラウンドスピーカ 95 についての k 番目の線スペクトルの相関信号と無相関信号が仮想音源 151a ~ 151e に割り当てられ、左フロントスピーカ 91 の k 番目の線スペクトルが仮想音源 151f に、右サラウンドスピーカ 95 の k 番目の線スペクトルが仮想音源 151g に、それぞれ割り当てられる。また、右サラウンドスピーカ 95 と右フロントスピーカ 92 に着目するという結果になった場合、図 16 に示すように、右サラウンドスピーカ 95 と右フロントスピーカ 92 についての k 番目の線スペクトルの相関信号と無相関信号が仮想音源 161a ~ 161e に割り当てられ、左サラウンドスピーカ 94 の k 番目の線スペクトルが仮想音源 151f に、左フロントスピーカ 91 の k 番目の線スペクトルが仮想音源 151g に、それぞれ割り当てられる。

30

#### 【0149】

上述のような処理を、ステップ S31a, S31b のループにより全ての線スペクトルについて行う。例えば、256 点の離散フーリエ変換を行った場合は 1 ~ 127 番目の線スペクトルまで、512 点の離散フーリエ変換を行った場合は 1 ~ 255 番目の線スペクトルまで、セグメントの全点 (1024 点) について離散フーリエ変換を行った場合は 1 ~ 511 番目の線スペクトルまで、となる。

40

#### 【0150】

その結果、図 17 に示すような仮想音源 171 の数を J (この例では J = 16) とすると、各仮想音源 (出力チャンネル) に対する周波数領域の出力音声信号  $Y_1(k)$ 、 $\dots$ 、 $Y_J(k)$  が求まる。これら出力が、図 2 における分離抽出部 22 の出力結果となる。

#### 【0151】

ここで、図 17 は、左右のスピーカ及び左右のサラウンドスピーカと全ての仮想音源と

50



の位置関係の例を説明するための模式図で、図 18 は図 17 とは異なる例を説明するための模式図である。図 17 の例では、16 個の仮想音源 171 を図 9 における各スピーカ 91 ~ 95 を結ぶ円と同心円周上に配置したが、図 18 の例のように、16 個の仮想音源 181 を四角形（この例では台形）上に並ぶように配置してもよい。図 18 の例は、4 つの辺のそれぞれにおいて仮想音源 181 が直線上に並ぶような配置例であり、図 5 で説明すると再生用スピーカ 52a ~ 52e の位置に配置した仮想音源を各辺毎に組み合わせたものである。

#### 【0152】

また、ここでは例として 5.1ch サラウンドシステムの信号の変換処理について説明したが、6.1ch や 7.1ch サラウンドシステムでも、同様に変換処理が可能である。この点について図 19 及び図 20 を参照して説明する。図 19 は、6.1ch サラウンドシステムのスピーカ群において、LFE を除いた 6 つのスピーカのうち、ダウンミックス後の出力対象スピーカの配置例を示す図である。図 20 は、7.1ch サラウンドシステムのスピーカ群において、LFE を除いた 7 つのスピーカのうち、ダウンミックス後の出力対象スピーカの配置例を示す図である。

#### 【0153】

6.1ch システムにおいても、センター（C）チャンネルを L チャンネルと R チャンネルにダウンミックスし、 $L/R/L_S/R_S/C_B$  の 5ch のうち、図 19 において矢印で示すような各チャンネルのペアを考慮して、前述と同様の処理を行えばよい。7.1ch システムにおいても、センター（C）チャンネルを L チャンネルと R チャンネルにダウンミックスし、 $L/R/L_S/R_S/L_B/R_B$  の 6ch のうち、図 20 において矢印で示すような各チャンネルのペアを考慮して、前述と同様の処理を行えばよい。

#### 【0154】

このように、入力音声信号は、5.1ch あるいは 6.1ch あるいは 7.1ch サラウンド音響信号であり、ダウンミックスする 1 つのチャンネルは前方中央のチャンネルの入力音声信号であることが好ましい。これは、前方中央のチャンネルは上述したように左右のフロントチャンネルとの間で音圧パニングを行っているが、そのような場合でも本例では適切な音像が得られるためである。なお、サブウーファの音声信号は通常、別に処理されるため、入力音声信号はそれらのいずれかのサラウンド音響信号からサブウーファのチャンネルを除いた音響信号と捉えることもできる。そして、このようなサラウンド音声信号を、センターチャンネルをダウンミックスした上で、線スペクトル毎または小帯域毎に、隣合うチャンネルのペアのうち、着目するペアを判定し、着目するペアについてのみ相関信号/無相関信号分離を行うことにより、左右のフロントチャンネルとの間で音圧パニングを行っていたとしても、サラウンド音響信号を、波面合成再生方式で適切な音像として再生することができる。また、これらのサラウンド音響信号に限らず、9.1ch などのサラウンド音響信号を入力音声信号として適用することもできる。

#### 【0155】

以上のようにして本例における分離抽出部 22 の処理がなされる。

次に、逆変換部 23 の処理がなされる。逆変換部 23 は、分離抽出部 22 で抽出された上記着目した組み合わせについての相関信号（またはその相関信号及び無相関信号）に対して、もしくはその相関信号から生成された音声信号に対して、もしくはその相関信号及びその無相関信号から生成された音声信号に対して、離散フーリエ逆変換を施す。また、逆変換部 23 は、上記着目した組み合わせ以外のチャンネルについて変換部 21 で変換された音声信号に対しても、離散フーリエ逆変換を施す。

#### 【0156】

具体的には、逆変換部 23 では、分離抽出部 22 から出力された各出力チャンネルを離散フーリエ逆変換することによって、上述した数式(35)により時間領域の出力音声信号  $y_j(m)$  を求める。なお、ここでは、上記着目した組み合わせについての相関信号及び無相関信号に対して離散フーリエ逆変換を施すとともに、上記着目した組み合わせ以外のチャンネルについて変換部 21 で変換された音声信号に対して離散フーリエ逆変換を施す例を

挙げているが、他の場合も同様である。

【0157】

数式(35)に関し、数式(36)、(37)で説明したように、本例においても離散フーリエ変換した信号は、窓関数乗算後の信号であったため、逆変換して得られた信号 $y_j(m)$ も窓関数が乗算された状態になっている。したがって、そうして得られた信号に、数式(2)に示す窓関数を再度乗算し、1つ前に処理したセグメントの先頭から1/4セグメント長ずつずらしながら出力バッファに加算していくことにより変換後のデータを得る。

【0158】

次に、図21～図24を参照して各スピーカに対する音声信号の生成について説明する。図21は、非特許文献1に記載の技術において、1つの直線上に並べたスピーカ群の背後に仮想音源を設けた場合に、各仮想音源に対応する音を出力するスピーカについて説明するための模式図である。図22～図24は、図1の音声信号再生装置におけるスピーカ群の配置例を説明するための模式図である。

10

【0159】

本例における音声出力信号生成部24では、逆変換部23で生成した各仮想音源の音声信号を、上述の非特許文献1に記載の技術に則って、各スピーカに対する音声信号を生成する。ただし、非特許文献1に記載の技術では、図21に示すように、1つの直線上に並べたスピーカ群(スピーカアレイ)211の背後に仮想音源212a～212eがある場合に、「どのスピーカが、どの仮想音源に対応する音を出力するか」を判定するために、仮想音源(仮想音源212bについて図示)からスピーカアレイ211の配列方向を示す直線に下ろした垂線と、仮想音源212bとスピーカを結んだ線とがなす角度 $\theta$ が、ある一定値より小さい場合に、そのスピーカはその仮想音源212bの音を出力する、という方法が採用されている。

20

【0160】

しかし、図12～図16で適用した仮想音源の配置例や図17や図18で説明した仮想音源の配置例では、少なくとも一部で一直線上に並ばないように配列させたスピーカ群が必要であり、例えば図22で例示するような各仮想音源222と各スピーカ221の配置が必要となる。そのため、非特許文献1に記載の方法では、ある仮想音源(例えば仮想音源222a)の音が、それと対面するスピーカ群221bからも出力されてしまい、音像定位に悪影響を及ぼす。

30

【0161】

したがって、例えば図22の仮想音源222aの音を、スピーカ群221のうちどのスピーカが出力するかどうかの判定には、各仮想音源222を配置した円の中心点224と仮想音源222aとを結んだ直線と、判定するスピーカ(例えば図22のスピーカ221a)と中心点224とを結んだ直線とがなす角度 $\alpha$ が、ある一定値より小さい場合に、スピーカ221aは、仮想音源222aの音を出力すると判定する。これを全てのスピーカと全ての仮想音源の組み合わせに対して行うことによって、図22のような配置をする場合でも各スピーカの出力する音声信号を決定することが可能となる。

【0162】

このようにして、音声信号処理部13は、マルチチャネル再生方式の5つ以上のチャネルの入力音声信号を、仮想的に存在する音源である仮想音源に対する音像としてスピーカ群により再生させるために変換することができる。ここで、図22で例示し後述の図23、図24で例示するように、本例におけるスピーカ群は、少なくとも一部で一直線上に並ばないように配列されており、仮想音源とスピーカそれぞれと中心点とを結んだ直線がなす角度によって、出力対象とするスピーカを決定する。

40

【0163】

また、図23で例示するように、スピーカ群231を円周上に並べた場合についても、同じ方法を用いることによって、各仮想音源232に対する音声信号を決定することが可能となる。

【0164】

50

また、例えば図10における左サラウンドスピーカ94と右サラウンドスピーカ95のペアには着目しないという制約をつけた上で、前述の処理を行うと、音声信号が割り当てられる仮想音源は、図24に示すような仮想音源242（この例では13個の仮想音源242）に限定され、それらの仮想音源242を、全周囲ではなく図24に示すように三方のみを囲う配置のスピーカ群241で出力することも可能である。

【0165】

以上、説明したように、本例における音声信号処理部13でなした処理は、着目ペアについてのみ相関信号・無相関信号に分離した上で仮想音源（または実在のスピーカ）への割り当てを行う処理であり、この処理は、同じ線スペクトルあるいは小帯域に属する音像は1つであるといった仮定を置くことを意味している。そして、その仮定では正確な音像位置を再現するための変換処理が可能となる。よって、このような変換処理により、マルチチャンネル再生方式の5つ以上のチャンネルの入力音声信号を、スピーカ群を用いて再生する際に適切な音像を提供できる音声信号に変換することができる。

【0166】

特に、本発明では着目ペアの判定を、電力の大きさを用いて、もしくは電力の大きさ及び相関係数（つまり相関係数の大きさ）を用いて行っている。そのため、ダウンミックス元のチャンネルの信号（この例ではC信号）に含まれる周波数成分については、ダウンミックス先のペアの電力が他のペアの電力に比べて大きくなり、ダウンミックス先のペアが着目ペアに決定される。したがって、センターチャンネル信号と左フロントチャンネル信号との間、あるいは、センターチャンネル信号と右フロントチャンネル信号との間で、音圧パニングを行っている場合にも、適切な音像に変換することができる。無論、他の周波数成分（この例では、C信号に含まれない周波数成分であって、元のL、R、 $L_s$ 、 $R_s$ の信号に含まれる周波数成分）については、他のペアが選択された結果、同様に適切な音像に変換することができる。

【0167】

次に、本発明の実装について簡単に説明する。本発明は、例えばホームシアターシステムやミニシアターシステムなど、映像の伴う装置に利用できる。図25は、図1の音声信号再生装置を備えた映像表示システムの構成例を示す図である。図25で示す部屋250のように、本発明に係る音声信号再生装置は、部屋の壁の三方（あるいは四方）にスピーカ群251～253を並べて取り付け付けたスピーカシステムに適用することができる。そして、図25で例示するように、テレビ装置等の映像表示装置254にこのスピーカシステムを接続し、映像表示装置254で表示させた映像に対応する音声信号をスピーカ群251～253から出力することもできる。

【0168】

さらに、本例においても、線スペクトル毎（または小帯域毎）に図4で例示した式(14)の $P_s^{(k)}$ が得られる。よって、本例においても、スピーカ群251～253の各スピーカの傍にLED255～257を配置したように三方や四方（あるいは円形状）に配置したアレイスピーカに複数の発光部を設けておけば、図8で説明したように音の強さに応じて発光色を変化させたり発光強度を変化させたりすることが、三方や四方（あるいは円形状）で可能となる。また、受聴者の前方のみ、LED256の代わりにあるいはLED256に加えて、図7で例示した棒グラフ72を映像表示装置254に表示させることもできる。

【0169】

なお、上述した各例において適用可能な波面合成再生方式としては、上述したようにスピーカアレイ（複数のスピーカ）を備えて仮想音源に対する音像としてそれらのスピーカから出力するようにする方式であればよく、非特許文献1に記載のWFS方式の他、人間の音像知覚に関する現象としての先行音効果（ハース効果）を利用した方式など様々な方式が挙げられる。ここで、先行音効果とは、同一の音声を複数の音源から再生し、音源それぞれから聴取者に到達する各音声に小さな時間差がある場合、先行して到達した音声の音源方向に音像が定位する効果を指し示したものである。この効果を利用すれば、仮想音

源位置に音像を知覚させることが可能となる。ただし、その効果だけで音像を明確に知覚させることは難しい。ここで、人間は音圧を最も高く感じる方向に音像を知覚するという性質も持ち合わせている。したがって、音声信号再生装置において、上述の先行音効果と、この最大音圧方向知覚の効果とを組み合わせ、これにより、少ない数のスピーカでも仮想音源の方向に音像を知覚させることが可能になる。

#### 【0170】

以上、本発明に係る音声信号再生装置の様々な例について、入力音声信号がマルチチャネル再生方式であることを前提に説明した。入力音声信号が波面合成再生方式の音声信号である場合について補足的に説明する。この場合には、本発明の主たる特徴である音像表示処理のように相関信号の分離等の処理を行わずとも、入力音声信号が示す音源位置の情報に合う位置にその信号電力値（または音圧レベル）を示す情報を表示し、かつ例えば波面合成再生方式でスピーカアレイから音声出力すればよい。これにより、入力音声信号が波面合成再生方式の音声信号である場合においても、聴覚と知覚とで音像位置が一致するように音像位置を表示させることは可能である。

#### 【0171】

また、例えば図1で例示した音声信号処理部13や表示制御部18における各構成要素など、本発明に係る音声信号再生装置の各構成要素は、例えばマイクロプロセッサ（またはDSP: Digital Signal Processor）、メモリ、バス、インターフェイス、周辺装置などのハードウェアと、これらのハードウェア上にて実行可能なソフトウェアとにより実現できる。表示部19を除き、上記ハードウェアの一部または全部は集積回路/IC(Integrated Circuit)チップセットとして搭載することができ、その場合、上記ソフトウェアは上記メモリに記憶しておければよい。また、本発明の各構成要素の全てをハードウェアで構成してもよく、その場合についても同様に、そのハードウェアの一部または全部を集積回路/ICチップセットとして搭載することも可能である。

#### 【0172】

また、上述した様々な構成例における機能を実現するためのソフトウェアのプログラムコードを記録した記録媒体を、音声信号再生装置となる汎用コンピュータ等の装置に供給し、その装置内のマイクロプロセッサまたはDSPによりプログラムコードが実行されることによって、本発明の目的が達成される。この場合、ソフトウェアのプログラムコード自体が上述した様々な構成例の機能を実現することになり、このプログラムコード自体や、プログラムコードを記録した記録媒体（外部記録媒体や内部記憶装置）であっても、そのコードを制御側が読み出して実行することで、本発明を構成することができる。外部記録媒体としては、例えばCD-ROMまたはDVD-ROMなどの光ディスクやメモリカード等の不揮発性の半導体メモリなど、様々なものが挙げられる。内部記憶装置としては、ハードディスクや半導体メモリなど様々なものが挙げられる。また、プログラムコードはインターネットからダウンロードして実行することや、放送波から受信して実行することもできる。

#### 【0173】

以上、本発明に係る音声信号再生装置について説明したが、処理の流れをフロー図で例示したように、本発明は、音声信号を再生する音声信号再生方法としての形態も採り得る。この音声信号再生方法は、次の分離ステップ、算出ステップ、加算ステップ、及び音像表示ステップを含む。

#### 【0174】

分離ステップは、分離部が、2つの音声信号を、周波数毎または周波数帯域毎に周波数についての相関信号と無相関信号とに分離するステップである。算出ステップは、算出部が、分離ステップで分離された各相関信号について、音像方向及び信号電力値を算出するステップである。加算ステップは、加算部が、各相関信号について、音像方向に基づき、予め定められた複数の音像方向グループのうちの1つに信号電力値を割り当て、音像方向グループ毎に信号電力値の加算値を求めるステップである。音像表示ステップは、音像表示部が、複数の音像方向グループのそれぞれに対応する位置に各音像方向グループについ

ての加算値を示す情報を表示するステップである。その他の応用例については、音声信号再生装置について説明した通りであり、その説明を省略する。

【0175】

なお、上記プログラムコード自体は、換言すると、この音声信号再生方法をコンピュータに実行させるためのプログラムである。すなわち、このプログラムは、コンピュータに、2つの音声信号を、周波数毎または周波数帯域毎に周波数についての相関信号と無相関信号とに分離する分離ステップと、分離ステップで分離された各相関信号について、音像方向及び信号電力値を算出する算出ステップと、各相関信号について、音像方向に基づき、予め定められた複数の音像方向グループのうちの1つに信号電力値を割り当て、音像方向グループ毎に信号電力値の加算値を求める加算ステップと、複数の音像方向グループのそれぞれに対応する位置に各音像方向グループについての加算値を示す情報を表示する音像表示ステップと、を実行させるためのプログラムである。その他の応用例については、音声信号再生装置について説明した通りであり、その説明を省略する。

10

【0176】

以上説明したように、本発明に係る音声信号再生装置は、2つの音声信号を、周波数毎または周波数帯域毎に周波数についての相関信号と無相関信号とに分離する分離部と、該分離部で分離された各相関信号について、音像方向及び信号電力値を算出する算出部と、各相関信号について、前記音像方向に基づき、予め定められた複数の音像方向グループのうちの1つに前記信号電力値を割り当て、該音像方向グループ毎に前記信号電力値の加算値を求める加算部と、前記複数の音像方向グループのそれぞれに対応する位置に各音像方向グループについての前記加算値を示す情報を表示する音像表示部と、を備えたことを特徴としたものである。これにより、聴覚と知覚とで音像位置が一致するように音像位置を表示させることができ、受聴者の没入感を高めることができる。

20

【0177】

前記音像表示部は、棒グラフ形式で、前記複数の音像方向グループのそれぞれに対応する位置に各音像方向グループについての前記加算値を示す情報を表示するようにしてもよい。これにより、受聴者の没入感をより増すことができる。

【0178】

もしくは、前記音像表示部は、複数の発光部を有し、前記複数の音像方向グループのそれぞれに対応する位置の前記発光部の発光色を、各音像方向グループについての前記加算値を示す情報に応じて変化させるようにしてもよい。音の強さに応じて発光色を変化させることで、例えば音楽ビデオを再生している場合や音楽を再生している場合に、ライブ会場での感覚を受聴者に与えることができる。

30

【0179】

もしくは、前記音像表示部は、複数の発光部を有し、前記複数の音像方向グループのそれぞれに対応する位置の前記発光部の発光強度を、各音像方向グループについての前記加算値を示す情報に応じて変化させるようにしてもよい。音の強さに応じて発光強度を変化させることで、例えば音楽ビデオを再生している場合や音楽を再生している場合に、ライブ会場での感覚を受聴者に与えることができる。

【0180】

40

本発明に係る音声信号再生方法は、分離部が、2つの音声信号を、周波数毎または周波数帯域毎に周波数についての相関信号と無相関信号とに分離する分離ステップと、算出部が、前記分離ステップで分離された各相関信号について、音像方向及び信号電力値を算出する算出ステップと、加算部が、各相関信号について、前記音像方向に基づき、予め定められた複数の音像方向グループのうちの1つに前記信号電力値を割り当て、該音像方向グループ毎に前記信号電力値の加算値を求める加算ステップと、音像表示部が、前記複数の音像方向グループのそれぞれに対応する位置に各音像方向グループについての前記加算値を示す情報を表示する音像表示ステップと、を含むことを特徴としたものである。これにより、聴覚と知覚とで音像位置が一致するように音像位置を表示させることができ、受聴者の没入感を高めることができる。

50

## 【 0 1 8 1 】

本発明に係るプログラムは、コンピュータに、2つの音声信号を、周波数毎または周波数帯域毎に周波数についての相関信号と無相関信号とに分離する分離ステップと、前記分離ステップで分離された各相関信号について、音像方向及び信号電力値を算出する算出ステップと、各相関信号について、前記音像方向に基づき、予め定められた複数の音像方向グループのうちの1つに前記信号電力値を割り当て、該音像方向グループ毎に前記信号電力値の加算値を求める加算ステップと、前記複数の音像方向グループのそれぞれに対応する位置に各音像方向グループについての前記加算値を示す情報を表示する音像表示ステップと、を実行させるためのプログラムである。これにより、本発明の機能をプログラムとして提供することができる。

10

## 【 0 1 8 2 】

本発明に係る記録媒体は、上記のプログラムを記録したコンピュータ読み取り可能な記録媒体である。これにより、上記プログラムを記録媒体で流通させることができる。

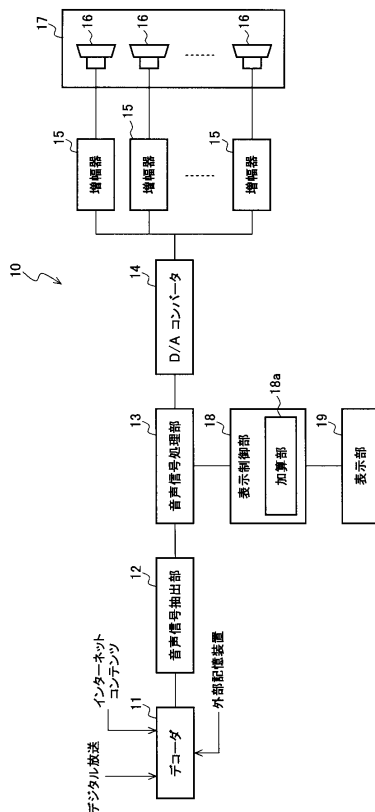
## 【符号の説明】

## 【 0 1 8 3 】

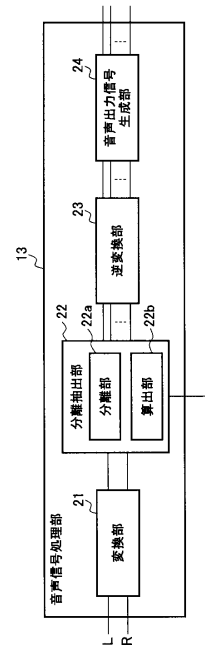
1 0 ... 音声信号再生装置、1 1 ... デコーダ、1 2 ... 音声信号抽出部、1 3 ... 音声信号処理部、1 4 ... D/Aコンバータ、1 5 ... 増幅器、1 6 , 1 6 a , 1 6 b , 1 6 c ... スピーカ、1 7 ... スピーカアレイ、1 8 ... 表示制御部、1 8 a ... 加算部、1 9 ... 表示部、2 1 ... 変換部、2 2 ... 分離抽出部、2 2 a ... 分離部、2 2 b ... 算出部、2 3 ... 逆変換部、2 4 ... 音声出力信号生成部、7 0 ... テレビ装置、7 1 ... 表示パネル、7 2 ... 棒グラフ、7 3 a , 7 3 b , 7 3 c , ... 情報、8 1 a , 8 1 b , 8 1 c ... L E D。

20

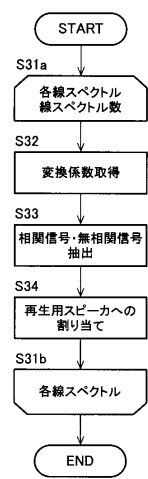
【 図 1 】



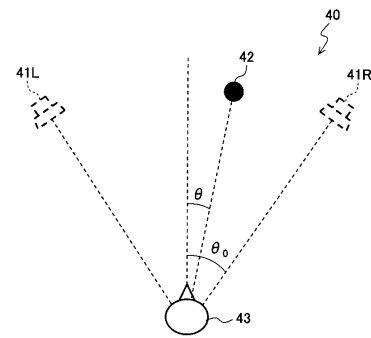
【 図 2 】



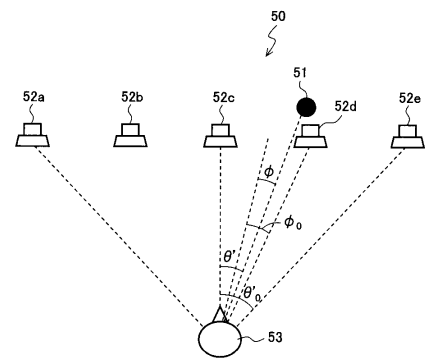
【図 3】



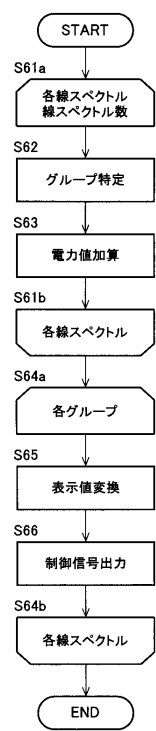
【図 4】



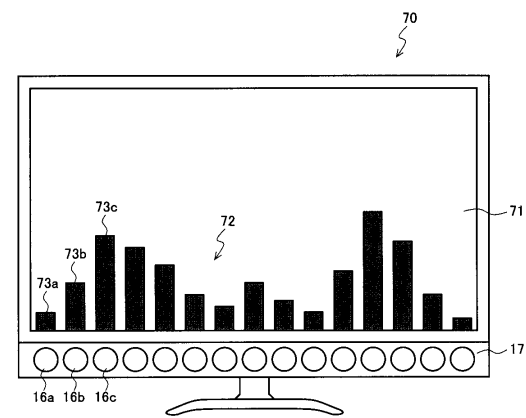
【図 5】



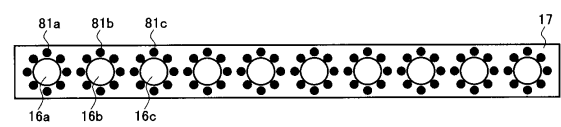
【図 6】



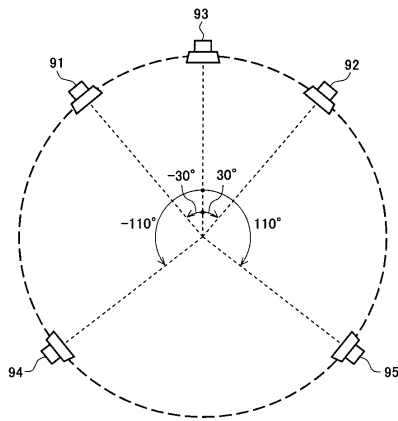
【図 7】



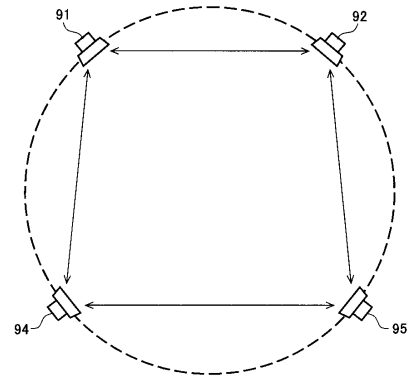
【図 8】



【図 9】



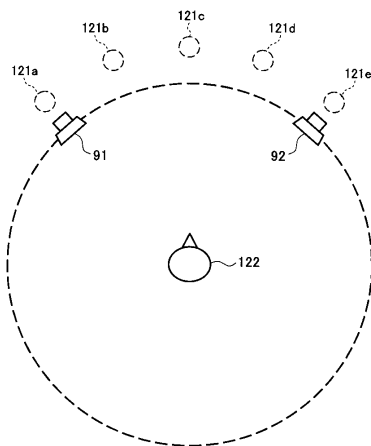
【図 10】



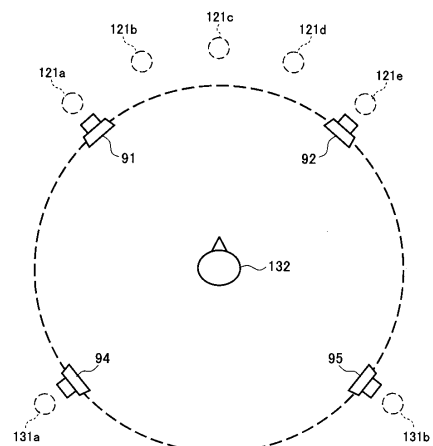
【図 11】



【図 12】

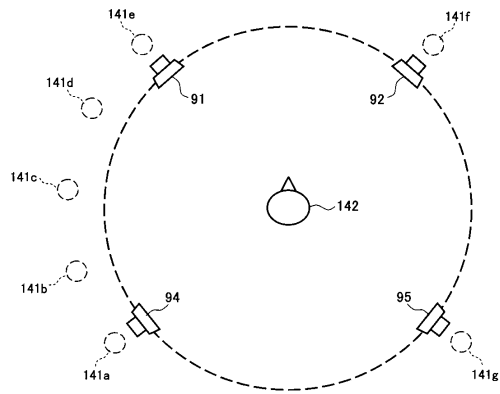


【図 13】

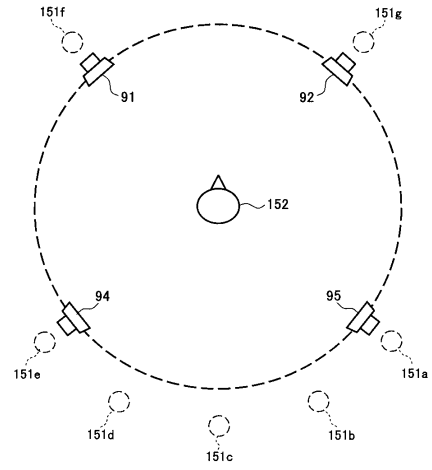




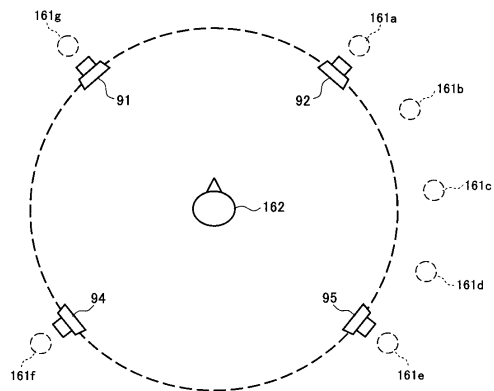
【図 14】



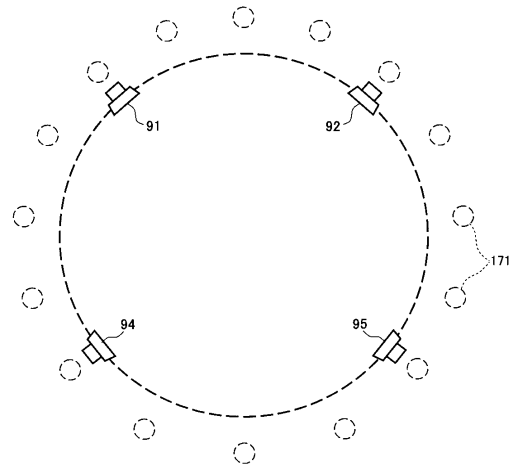
【図 15】



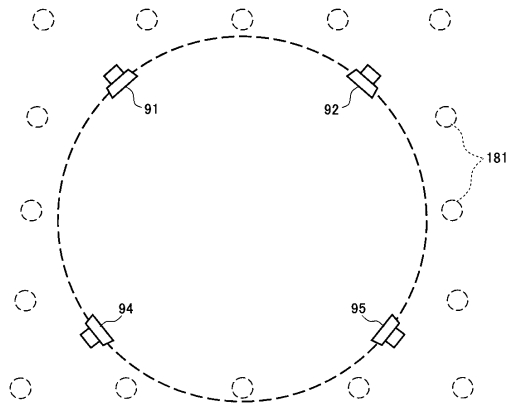
【図 16】



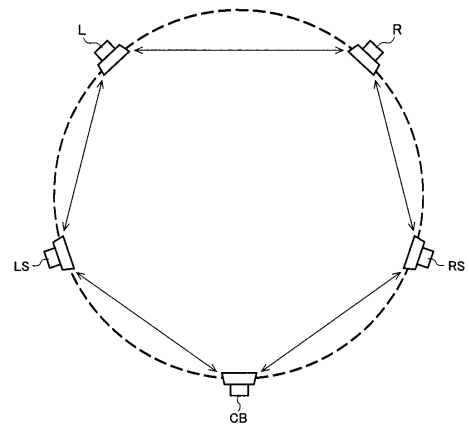
【図 17】



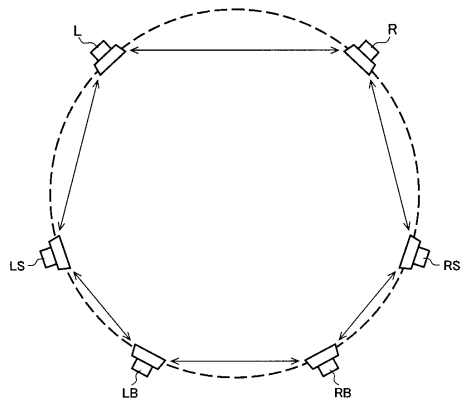
【図 18】



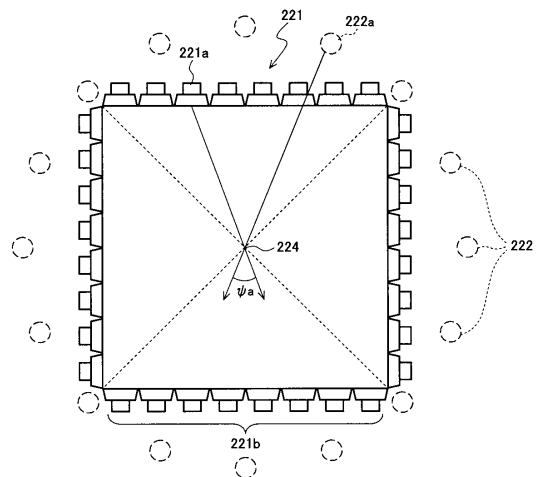
【図 19】



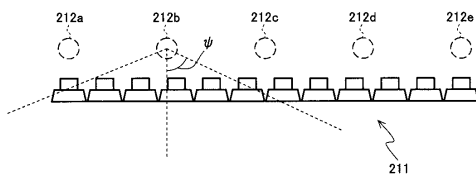
【図 20】



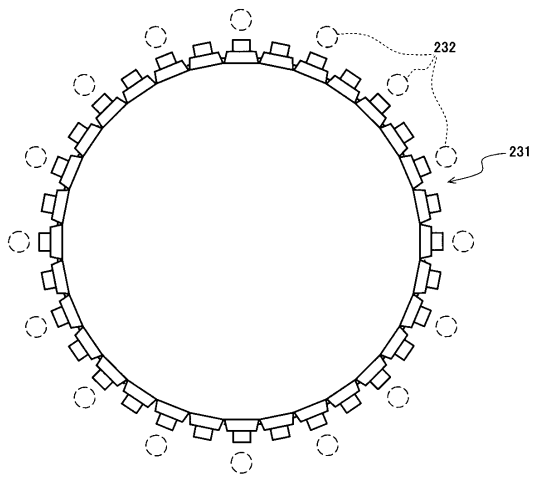
【図 22】



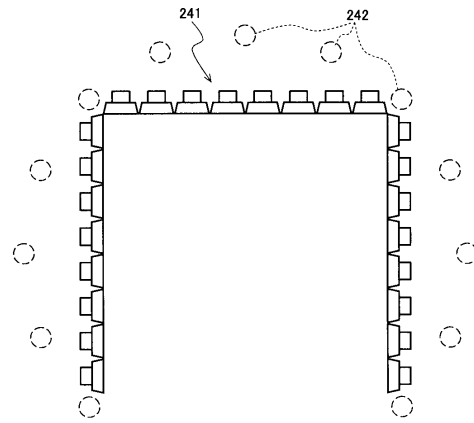
【図 21】



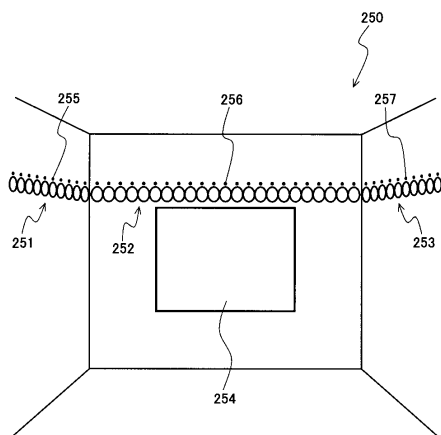
【図 2 3】



【図 2 4】



【図 2 5】



---

フロントページの続き

(56)参考文献 特開2001-268698(JP,A)  
実開昭60-177600(JP,U)  
特開平10-164700(JP,A)  
特開2011-244196(JP,A)  
特開2012-147337(JP,A)  
特開2012-060301(JP,A)  
特開2011-239036(JP,A)  
米国特許出願公開第2008/0253592(US,A1)

(58)調査した分野(Int.Cl., DB名)  
H04S 7/00  
H04R 29/00