

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6013918号  
(P6013918)

(45) 発行日 平成28年10月25日 (2016. 10. 25)

(24) 登録日 平成28年9月30日 (2016. 9. 30)

(51) Int. Cl. F I  
H04S 5/02 (2006.01) H04S 5/02 H

請求項の数 14 (全 20 頁)

(21) 出願番号	特願2012-550544 (P2012-550544)	(73) 特許権者	590000248
(86) (22) 出願日	平成23年1月26日 (2011. 1. 26)		コーニンクレッカ フィリップス エヌ ヴェ
(65) 公表番号	特表2013-519253 (P2013-519253A)		KONINKLIJKE PHILIPS N. V.
(43) 公表日	平成25年5月23日 (2013. 5. 23)		オランダ国 5656 アーエー アイン ドーフエン ハイテック キャンパス 5
(86) 国際出願番号	PCT/IB2011/050334		High Tech Campus 5, NL-5656 AE Eindhoven
(87) 国際公開番号	W02011/095913		
(87) 国際公開日	平成23年8月11日 (2011. 8. 11)	(74) 代理人	110001690
審査請求日	平成26年1月22日 (2014. 1. 22)		特許業務法人M&Sパートナーズ
(31) 優先権主張番号	10152388.4	(74) 代理人	100114753
(32) 優先日	平成22年2月2日 (2010. 2. 2)		弁理士 宮崎 昭彦
(33) 優先権主張国	欧州特許庁 (EP)		

最終頁に続く

(54) 【発明の名称】 空間音声再生

(57) 【特許請求の範囲】

【請求項 1】

マルチチャネル音声信号を受信するための受信器と、マルチチャネル音声信号の空間特性を決定するための回路と、異なる空間処理アルゴリズムを適用する複数のマルチチャネル音声再生モードであって、異なる空間レンダリング技術を使用する当該複数のマルチチャネル音声再生モードから選択される再生モードを、前記空間特性に応じて選択するための回路と、選択された再生モードを使用して、マルチチャネル音声信号を再生するために、ラウドスピーカーのセットにより供給される空間チャンネルのセットを駆動するための再生回路とを有し、前記複数のマルチチャネル音声再生モードが、マルチチャネル信号の空間特徴を維持している再生モードと、少なくとも一つの主要なソース信号と周囲信号とへ分離して前記少なくとも一つの主要なソース信号及び前記周囲信号に互いに異なる空間再生を適用する再生モードとを有する、空間音声再生のための装置。

【請求項 2】

音声再生モードの少なくとも一つが、マルチチャネル音声信号のチャンネルの数より多い数の空間チャンネルにアップミキシングすることと、マルチチャネル音声信号のチャンネルの数より少ない数の空間チャンネルにダウンミキシングすることとの少なくとも一つを有する、請求項 1 に記載の装置。

【請求項 3】

空間チャンネルのセットが、マルチチャネル音声信号とは異なる数のチャンネルを有する、請求項 1 に記載の装置。

## 【請求項 4】

音声再生モード間の切換えのための最大スイッチ周波数が 1 Hz を超える、請求項 1 に記載の装置。

## 【請求項 5】

空間特性を決定するための前記回路がせいぜい 10 秒の時定数を持つ空間特性を決定する、請求項 1 に記載の装置。

## 【請求項 6】

マルチチャネル音声信号に対するコンテンツ特徴を決定するための回路を更に有し、選択するための前記回路が、コンテンツ特徴に応じて選択される再生モードを選択する、請求項 1 に記載の装置。

10

## 【請求項 7】

コンテンツ特徴を決定するための前記回路が、マルチチャネル音声信号と関連したメタデータに応じて、コンテンツ特徴を決定する、請求項 6 に記載の装置。

## 【請求項 8】

マルチチャネル音声信号を再生するための前記再生回路がコンテンツ特徴に応じて選択される再生モードの空間レンダリング技術の特徴を適応させる、請求項 6 に記載の装置。

## 【請求項 9】

マルチチャネル音声信号を再生するための前記再生回路が空間特性に応じて選択される再生モードの空間レンダリング技術の特徴を適応させる、請求項 1 に記載の装置。

## 【請求項 10】

前記特徴がマルチチャネル音声信号の少なくとも 2 つのチャンネルに適用される空間拡散の程度である、請求項 9 に記載の装置。

20

## 【請求項 11】

マルチチャネル音声信号を再生するための前記再生回路が第 1 の選択される再生モードから第 2 の選択される再生モードへ徐々に移行する、請求項 1 に記載の装置。

## 【請求項 12】

空間特性を決定するための前記回路が、マルチチャネル音声信号の少なくとも 2 つのチャンネルの差信号のエネルギー指標に対するマルチチャネル音声信号の少なくとも 2 つのチャンネルの和信号のエネルギー指標に応じて、空間特性を決定する、請求項 1 に記載の装置。

30

## 【請求項 13】

空間特性を決定するための前記回路が、マルチチャネル音声信号を少なくとも一つの主要な音源信号と残差信号とに分解し、残差信号のエネルギー指標に対する主要な音源信号のエネルギー指標に応じて、空間特性を決定する、請求項 1 に記載の装置。

## 【請求項 14】

マルチチャネル音声信号を受信するステップと、マルチチャネル音声信号の空間特性を決定するステップと、異なる空間処理アルゴリズムを適用する複数のマルチチャネル音声再生モードであって、異なる空間レンダリング技術を使用する当該複数のマルチチャネル音声再生モードから選択される再生モードを、前記空間特性に応じて選択するステップと、選択された再生モードを使用して、マルチチャネル音声信号を再生するためにラウドスピーカーのセットを駆動するステップとを有し、前記複数のマルチチャネル音声再生モードが、マルチチャネル信号の空間特徴を維持している再生モードと、少なくとも一つの主要なソース信号と周囲信号とへ分離して前記少なくとも一つの主要なソース信号及び前記周囲信号に互いに異なる空間再生を適用する再生モードとを有する、空間音声再生の方法。

40

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、空間音声再生、排他的ではないが、特に、マルチチャネル音声信号のアップミキシングを含む空間音声再生に関する。

50

**【背景技術】****【0002】**

ステレオ記録及び再生の形式の空間音声再生は、数十年の歴史がある。最近の年代において、改良された空間リスニング経験を提供するために、より高度なアレンジメント及び信号処理が用いられた。特に、例えば5台又は7台の空間スピーカを使用するサラウンドサウンドの使用は、例えば映画又はテレビを視聴することに関連して強化された経験を提供するために主流になった。加えて、「サウンドバー」のようなコンパクトなマルチドライバラウドスピーカーシステムは、従来のステレオ及び5.1システムに対する普及しているオプションになっている。これらの装置は、小さな装置からでさえ広い空間オーディオイメージの経験をリスナーに提供する。これは、信号のデジタル処理及び装置の特別な物理的配置に基づく。

10

**【0003】**

空間サウンド処理は、改良された空間経験を提供するために、音声再生の一部として先進の信号処理を益々利用する。例えば、複雑なアルゴリズムが、音声信号をより高い数のチャンネルにアップミキシングするために用いられる。例えば、5チャンネルサラウンド信号は、送信側でステレオ又はモノラルの信号にダウンミックスされる。この信号は、その後配信され、音声再生は、元の5チャンネル信号へ受信信号をアップミキシングすることを含む。

**【0004】**

別の例として、信号処理は、サウンド拡散効果をステレオ信号へ供給するために用いられ、結果的によりワイドなサウンドをリスナーが経験することになる。通常、これらの方法は、チャンネル間の相関を低減する信号処理オペレーションに基づく。これらの技術は、前述のコンパクトなラウドスピーカーシステムで特に普及している。

20

**【0005】**

別の例として、空間信号の再生は、例えばステレオ信号内の主要な音源の抽出を含む。残りの残差信号は、拡散する周囲ステレオイメージに通常対応する。再生特徴が各信号に対して最適化されるように、その後、主要な信号及び周囲信号が異なって再生される。

**【発明の概要】****【発明が解決しようとする課題】****【0006】**

しかしながら、斯様な空間音声再生技術はリスニング経験を改善するにもかかわらず、幾つかの関連する不利な点がある傾向がある。特に、再生は、全ての状況において最適空間経験を提供するというわけではなく、信号処理は、場合によっては実際のところ劣化する空間経験に結果としてなる。

30

**【0007】**

従って、空間音声再生のための改良されたシステムは有利であり、特に、増大された柔軟性、容易にされたオペレーション、容易にされた実行、改良された空間リスニング経験及び/又は改良されたパフォーマンスを可能にするシステムは有利である。

**【0008】**

従って、本発明は、好ましくは単独で又は任意の組合せで、上述の不利な点の一つ以上を緩和し、軽減し又は除去しようとする。

40

**【課題を解決するための手段】****【0009】**

本発明の一態様によると、マルチチャンネル音声信号を受信するための受信器と、マルチチャンネル音声信号の空間特性を決定するための回路と、複数のマルチチャンネル音声再生モードが異なる空間レンダリング技術を使用する当該複数の音声再生モードから選択される再生モードを選択するための回路と、選択された再生モードを使用して、マルチチャンネル音声信号を再生するために、ラウドスピーカーのセットにより供給される空間チャンネルのセットを駆動するための再生回路とを有する、空間音声再生のための装置が提供される。

**【0010】**

50

本発明は、多くの実施例で改良された音声再生を提供する。特に、改良された空間経験は、多くのシナリオにおいて提供される。通常、空間再生は、特定の音声信号に対して改善される。このアプローチは、更に、多くの実施例において複雑さが低い実行及び容易にされたオペレーションを可能にする。

【 0 0 1 1 】

適切な再生方法の選択は、低い複雑さを維持すると共に、経験される特定の条件に対して最適化される。

【 0 0 1 2 】

空間特性は、信号の空間的構成及び/又は空間複雑さを表す。例えば、空間特性は、主要な音源を抽出するための適切な基準又はプロセスに従う一つ以上の主要な音源の存在を表す。ある実施例では、空間特性は、マルチチャネル信号により表されるサウンドイメージの音源の空間分布を表す。

10

【 0 0 1 3 】

ラウドスピーカーのセットは、具体的には、（おそらく空間的でない低周波効果のスピーカ又はサブウーファを加えて）例えば3台、5台又は7台の空間スピーカを有するサラウンドサウンドセットアップのラウドスピーカーである。ラウドスピーカーのセットは、1台の物理的装置の通常3つ以上の個別に駆動されるラウドスピーカー（又は、ラウドスピーカーアレイ）を持つマルチドライバラウドスピーカーシステムである。ラウドスピーカーのセットは、複数の斯様な装置を有してもよい。

【 0 0 1 4 】

20

本発明の任意の特徴によると、音声再生モードの少なくとも1つが、マルチチャネル音声信号のチャンネルの数より多い数の空間チャンネルにアップミキシングすることと、マルチチャネル音声信号のチャンネルの数より少ない数の空間チャンネルにダウンミキシングすることとの少なくとも一つを有する。

【 0 0 1 5 】

本発明は、改良された空間経験を提供する。例えば、ステレオ信号の幾つかのサウンドイメージは、モノラル信号として再生されるとき改良された空間経験を提供する。ステレオ信号の他のサウンドイメージは、センター信号と結合された拡張ステレオ信号として再生されるとき、すなわち3つの空間チャンネルを使用して再生されるとき、改良された空間経験を提供する。

30

【 0 0 1 6 】

本発明の任意の特徴によると、空間チャンネルのセットは、マルチチャネル音声信号とは異なる数のチャンネルを有する。

【 0 0 1 7 】

本発明は、音声再生システムに対する改良された空間経験を提供し、特に、音声再生を特定のサウンドイメージ及び空間特徴に適応させて、追加の自由度を可能にする。

【 0 0 1 8 】

本発明の任意によると、音声再生モード間の切換えの最大スイッチ周波数は、1 Hz を超える。

【 0 0 1 9 】

40

これは、音声の様々な特徴に密接にマッチする動的な適合及び最適化を提供し、これによって改良されたりスニング経験を提供する。

【 0 0 2 0 】

このフィーチャは、音声信号に対する再生モードの改良されたパフォーマンス及び改良された適合を可能にし、これによって強化されたりスニング経験を提供する。このアプローチは、信号特徴に対する再生の短期間の適合を可能にする。

【 0 0 2 1 】

幾つかの実施例では、再生モード間の切換えのための最大スイッチ周波数は、0.01 Hz、0.1 Hz 又は10 Hz さえも超える。

【 0 0 2 2 】

50

最大スイッチ周波数は、装置が再生モード間を切換え可能な最大周波数である。最大周波数は、空間特性推定及び切換え機能の特徴を含むシステムの設計パラメータにより制限される。

【0023】

本発明の任意の特徴によると、空間特性を決定するための前記回路がせいぜい10秒の時定数を持つ空間特性を決定する。

【0024】

これは、音声の様々な特徴に密接にマッチする動的な適合及び最適化を提供し、これにより改良されたリスニング経験を提供する。

【0025】

この特徴は、音声信号に対する再生モードの改良されたパフォーマンス及び改良された適合を可能にし、これによって強化されたリスニング経験を提供する。このアプローチは、信号特性に対する再生の短期間の適合を可能にする。

【0026】

幾つかの実施例では、空間特性を決定するための回路は、500秒未満、100秒、1秒、500ms、100ms又は更に50msの時定数を持つ空間特性を決定するように有利に設けられる。

【0027】

時定数は、空間特性がステップ変化に従う最終的な（漸近的な）値の  $1 - 1/e = 63\%$  に達する時間を表わす。

【0028】

幾つかの実施例では、空間特性を決定するための回路は、空間特性のローパスフィルタを含み、ローパスフィルタは0.001Hz、0.01Hz、0.1Hz、1Hz、10Hz又は50Hzを超える3dBカットオフ周波数を持つ。

【0029】

本発明の任意の特徴によると、複数の音声再生モードは、モノラル再生モードと、マルチチャンネル信号の空間特徴を維持している再生モードと、空間拡散処理を有する再生モードと、少なくとも一つの主要なソース信号と周囲信号とへ分離して少なくとも一つの主要なソース信号及び周囲信号に互いに異なる空間再生を適用する再生モードとの少なくとも二つを有する。

【0030】

これらの再生技術は、異なる音声特徴に対する改良されたリスニング特徴を提供するために特に有利であり適している。多くの実施例では、複数の音声再生モードは、2つ、3つ又は4つ全ての再生モードを好適に含み、これらが異なる特徴に特に適しているので、よって、音声特徴のかなりの範囲に対する改良された再生を供給する一組のモードも提供する。この技術は、広範囲の音声信号に対する適切な再生特徴も提供する。

【0031】

本発明の任意の特徴によると、装置は、マルチチャンネル音声信号に対するコンテンツ特徴を決定するための回路を更に有し、選択するための前記回路が、コンテンツ特徴に応じて選択される再生アルゴリズムを選択する。

【0032】

これは、再生の適合を更に改良し、多くの実施例で改良された空間経験を提供する。コンテンツ特性は、例えば、マルチチャンネル音声信号及び/又は関連するビデオ信号のコンテンツ分析により決定される。

【0033】

本発明の任意の特徴によると、コンテンツ特徴を決定するための前記回路は、マルチチャンネル音声信号と関連したメタデータに応じて、コンテンツ特徴を決定する。

【0034】

これは、多くの実施例において有利である特に正確で複雑さが低いアプローチを提供する。

10

20

30

40

50

## 【 0 0 3 5 】

本発明の任意の特徴によると、マルチチャネル音声信号を再生するための前記回路は、コンテンツ特徴に応じて選択される再生モードの空間レンダリング技術の特徴を適応させる。

## 【 0 0 3 6 】

これは、再生の適合を更に改善し、多くの実施例で改良された空間経験を提供する。

## 【 0 0 3 7 】

本発明の任意の特徴によると、マルチチャネル音声信号を再生するための前記回路は、空間特性に応じて選択される再生モードの空間レンダリング技術の特徴を適応させる。

## 【 0 0 3 8 】

これは、再生の適合を更に改善し、多くの実施例で改良された空間経験を提供する。

## 【 0 0 3 9 】

本発明の任意の特徴によると、空間処理特徴は、マルチチャネル音声信号の少なくとも2つのチャンネルに適用される空間拡散の程度である。

## 【 0 0 4 0 】

これは、空間的拡散が、ある音声特徴に対して著しく強化された空間経験を提供するが、他の音声特徴に対して空間経験を劣化させるので、特に有利な最適化を提供する。従って、音声特徴に対する空間的拡散の最適化は、特に有利なパフォーマンスを提供する。

## 【 0 0 4 1 】

本発明の任意の特徴によると、マルチチャネル音声信号を再生するための前記回路は、第1の選択される再生アルゴリズムから第2の選択される再生アルゴリズムへ徐々に移行する。

## 【 0 0 4 2 】

これは、改良されたパフォーマンスを提供し、特に、異なる再生モード間の変化の認知度を低減する。装置は、具体的には、遷移間隔の間に、第1の選択された再生アルゴリズム及び第2の選択された再生アルゴリズム両方を使用してラウドスピーカーのセットのための駆動信号を生成し、加重が遷移間隔の間に動的に変えられる駆動信号の加重組合せとして生成される信号によりラウドスピーカーのセットを駆動するように設けられる。

## 【 0 0 4 3 】

本発明の任意の特徴によると、空間特性を決定するための前記回路は、マルチチャネル音声信号の少なくとも2つのチャンネルの差信号のエネルギー指標に対するマルチチャネル音声信号の少なくとも2つのチャンネルの和信号のエネルギー指標に応じて、空間特性を決定する。

## 【 0 0 4 4 】

これは、空間再生を適合させるため特に有利な空間特性である。特に、これは、多くのシナリオに対して正確さと複雑さとの間の有利なトレードオフを提供する。

## 【 0 0 4 5 】

本発明の任意の特徴によると、空間特性を決定するための前記回路は、マルチチャネル音声信号を少なくとも一つの主要な音源信号と残差信号とに分解し、残差信号のエネルギー指標に対する主要な音源信号のエネルギー指標に応じて、空間特性を決定する。

## 【 0 0 4 6 】

これは、空間再生を適合させるための特に有利な空間特性である。特に、これは、多くのシナリオに対して正確さと複雑さとの間の有利なトレードオフを提供する。

## 【 0 0 4 7 】

本発明の一態様によると、マルチチャネル音声信号を受信するステップと、マルチチャネル音声信号の空間特性を決定するステップと、マルチチャネル音声再生モードが異なる空間レンダリング技術を使用して、複数の音声再生モードから選択される再生モードを選択するステップと、選択された再生モードを使用して、マルチチャネル音声信号を再生するためにラウドスピーカーのセットを駆動するステップとを有する、空間音声再生の方法が提供される。

10

20

30

40

50

## 【 0 0 4 8 】

本発明のこれら及び他の態様、特徴及び利点は、これ以降説明される実施例を参照して明らかに説明されるだろう。

## 【 0 0 4 9 】

本発明の実施例は、図面を参照して単なる例示として説明されるだろう。

## 【図面の簡単な説明】

## 【 0 0 5 0 】

【図 1】図 1 は、本発明の幾つかの実施例による空間音声再生のためのシステムの例の具体例である。

【図 2】図 2 は、本発明の幾つかの実施例による空間音声再生のためのシステムの要素の例の具体例である。

【図 3】図 3 は、本発明の幾つかの実施例による空間音声再生のためのシステムの例の具体例である。

## 【発明を実施するための形態】

## 【 0 0 5 1 】

以下の説明は、3つのチャンネルへのアップミキシングを使用して、ステレオ信号の空間音声再生に適用できる本発明の実施例に焦点を当てる。しかしながら、本発明は、このアプリケーションに限定されず、多くの他の音声信号及び再生方法に適用されてもよいことは、理解されるだろう。

## 【 0 0 5 2 】

図 1 は、本発明の幾つかの実施例による音を再生するためのシステムの例を示す。システムは、複数の音声チャンネルを有する空間音声信号を受信する受信器 101 を有する。この例では、入力信号はステレオ信号であるが、他の実施例では、他のチャンネル数が使用されてもよいことは、理解されるだろう。例えば、入力信号は、5チャンネルサラウンドサウンド入力信号でもよい。幾つかのシナリオでは、入力信号はコード化信号であり、受信器 101 はシステムによる他の処理のため入力信号を部分的に又は完全にデコードする。例えば、各コード化セグメントに対して、入力信号の周波数表現は、コード化スキームにより使用される中間周波数表現として生成されてもよい。入力信号の複数のチャンネルが、単一のコード化された音声信号及び関連するパラメータデータにより表されてもよいことは、理解されるだろう。例えば、多チャンネル入力信号は、コード化されたモノラル信号及び空間パラメータデータである。具体例として、入力信号は、パラメトリックステレオ信号である。

## 【 0 0 5 3 】

入力マルチチャンネル音声信号は、内部又は外部ソースから受信される。

## 【 0 0 5 4 】

受信器 101 は、受信器 101 からマルチチャンネル（具体例では、ステレオ信号）を受信するドライバ回路 103 に結合されている。ドライバ回路 103 は、一組のラウドスピーカー 105 のための駆動信号を生成する。ラウドスピーカーのセットは、多くの空間チャンネルを供給する。この例では、ラウドスピーカーは、左チャンネル、右チャンネル及びセンターチャンネルを提供するが、他の実施例では、より多くの（又は少ない）空間チャンネルが供給されてもよいことは、理解されるだろう。例えば、幾つかの実施例では、ラウドスピーカーは、左チャンネル及び右チャンネルを提供するだけである。他の実施例では、フルサラウンドシステムは、例えば5つ又は7つの空間チャンネルを具備する。

## 【 0 0 5 5 】

幾つかの例では、ラウドスピーカー 105 のセットのスピーカにより供給される空間チャンネルの数は、マルチチャンネル信号のチャンネルの数に等しい。しかしながら、この例では、ラウドスピーカー 105 のセットにより供給される空間チャンネルの数は、マルチチャンネル信号のチャンネルの数より多い。この例では、ドライバ回路 103 は、マルチチャンネル信号のチャンネルを空間チャンネルの数にアップミキシングすることを含む幾つかの再生モードで動作する。代わりに又は追加的に、ドライバ回路 103 は、少なくとも幾つかの再生モ

ードで利用可能なチャンネルのサブセットを選択するための機能を含み、当該サブセットは異なる再生モードで異なる。これらのモードの一つ以上は、更に、入力チャンネルのダウンミキシングを含む。例えば、ステレオ入力信号に対して、ある再生モードでは、空間チャンネルの2つ（例えば左と右）を使用して出力を供給し、他の再生モードでは、一つの空間チャンネル（例えばセンターチャンネル）だけを使用し、更に他の再生モードでは、3つの空間的なチャンネル（例えば左、右及びセンターチャンネル）を使用する。

【0056】

具体例では、ラウドスピーカー105のセットは、空間配置に3つのラウドスピーカーを有し、これにより3つの空間チャンネルを供給する。よって、ラウドスピーカー105のセットのスピーカは、左、右及びセンタースピーカーに対応する。

10

【0057】

ラウドスピーカーのセットは、このように、空間経験を提供するように設けられる。幾つかの実施例では、ドライバ回路103は、リスニング位置に対するラウドスピーカーの正確な位置を知っているが、通常は、この場合でなく、従来のサラウンド及びステレオシステムから知られているように、空間音声再生はラウドスピーカーの推定位置に基づく。ラウドスピーカーのセットは、複数の空間チャンネルを供給し、例えば、これらは左、右及びセンター空間チャンネルを提供し、空間経験をリスナーに提供するために用いられる。しかしながら、ラウドスピーカーのセットは、各チャンネルに対する単一の別々のラウドスピーカーを持つ必要はない。例えば、ラウドスピーカーのセットは、ラウドスピーカーアレイを有し、音声ビーム形成技術を使用して空間チャンネルを供給するための関連する駆動機能

20

【0058】

ラウドスピーカー105を駆動するとき、ドライバ回路103は異なる音声再生モードを使用するように配される。異なる音声再生モードは、異なる空間レンダリング技術を使用する。このように、異なる音声再生モードは異なる空間処理アルゴリズムを適用し、よって、異なる音声再生モードは異なる空間音声特徴を持つ。例えば、1つの音声再生モードは、単一のラウドスピーカー105だけ（すなわちモノラル再生として）を使用してマルチチャンネル信号を示し、他の音声再生モードは、空間処理なしに対応する空間チャンネルの信号で各ラウドスピーカーを単に駆動し、これにより入力信号の空間特徴を維持する。更に他の再生モードは、全てのラウドスピーカーにわたって入力チャンネルを広げて、空間的拡散を導入する。よって、ドライバ回路103は、非常に異なる空間処理を提供でき、非常に異なる特性でラウドスピーカー105のセットを駆動可能に設計されている。実際、異なる再生モードは、所与の空間処理のために異なるパラメータ設定を使用するだけでなく、異なる基礎をなす原理を適用し、特に異なる空間処理アルゴリズム及び方法を使用する。

30

【0059】

斯様な様々な再生モードは、非常に異なる効果をシステムにより供給可能にし、リスナーの空間経験の高い可変性を可能にする。しかしながら、発明者は、空間信号処理が強化された経験を提供するのに対して、空間信号処理は、ある場合には結果的に減少した空間経験にもなると認識した。例えば、（空間拡散、アップミキシング、モノラル信号への変換等のような）感知したステレオイメージ上の音声フォーマット変換アルゴリズムの効果は、異なるコンテンツ及び信号特徴に対して異なる。

40

【0060】

例えば、方法はアクション映画シーンに適している広い空間イメージを提供するが、単一の楽器を持つ音楽又はニュース番組の場合には、同じ方法は、落ち着きがなく、はっきりしないと知覚される。すなわち、あるタイプのコンテンツに適しているアップミキシング又はステレオ拡大は、異なるタイプのコンテンツのために使用されるとき、不所望な効

50

果を作り出す。

【 0 0 6 1 】

別の例として、ステレオ信号からセンターチャネルを抽出する目的のアップミキシングアルゴリズムは、ステレオミックス内にクリアなセントラル音源がないとき、必ずしも最適に働かない。センターチャネル抽出方法が斯様なコンテンツに対して用いられる場合、結果的にステレオイメージの幅の低減になる。

【 0 0 6 2 】

エンドユーザが再生モードを手動で選択するか又は調整可能にすることは、ユーザが最も気持ちの良い空間経験を提供するモードを選択することができるので、この感度を緩和可能にできる。しかしながら、発明者は、斯様な解決策が、遅くて非常に扱いにくい適合を可能にするだけなので、しばしば実際的でないと認識した。

10

【 0 0 6 3 】

解決策は、可能性のあるタイプの音声ごとに再生モードを定めることである。例えば、ニュース番組のために、ある特定の再生モードが用いられる、映画のために他の特定の再生モードが用いられる等である。しかしながら、発明者は、好適な空間再生が音声の特定タイプと直接リンクしないので、斯様なアプローチが不正確になりそうであることを認識した。

【 0 0 6 4 】

実際、発明者は、大幅に改良された経験が、適切な再生モードの動的なリアルタイム選択を実行することによりしばしば達成できることを認識した。発明者は、有利なパフォーマンスが入力信号の空間特性に基づいて斯様な動的な選択を実行することにより達成できることを更に認識した。よって、図 1 のシステムにおいて、再生モードは、入力信号の空間特性に基づいて動的に選択される。このことにより、入力信号の特定の変化に対する再生モードのリアルタイム且つ速い適合が達成される。

20

【 0 0 6 5 】

斯様なアプローチは、音声再生を信号の現在の特徴に自動的に且つ動的に適合可能にし、このことにより強化されたリスニング経験を可能にする。当該アプローチは、更に、例えば音声の特定タイプ又は音声を表す特定のプログラムタイプに対する平均的又は予想される特徴のためよりもむしろ、現在の特徴及び好みのために再生モードが最適化できる非常に速い適合を可能にする。例えば、当該アプローチは、例えばダイアログ及びアクション音声両方が、その特定の音に対して最も適切な再生アルゴリズムにより再生されるように、映画のサウンドトラックの間、再生モードを動的且つ自動的に変化可能にする。例えば、空間イメージがメディアアイテムの期間にわたって連続的にしばしば変化することが知られている。例えば、映画の音声シーンは、俳優の声のような 1 つの音源だけが聞き取れるときと、ワイドステレオ音声シーンとの間での交番を含む。ステレオイメージがワイドで実体験のように感知できることを所望する第 1 の場合がある一方、音声に対して明らかに局所的な空間位置を持つことが自然である第 2 の場合がある。図 1 のシステムは、斯様な好みを反映するために、再生モードの自動的調整を供給する。

30

【 0 0 6 6 】

特に、図 1 のシステムは、マルチチャネル音声信号の空間特性を決定するように設けられるアナライザ 107 を有する。空間特性は、具体的には、入力信号に存在する空間的構成又は複雑さの程度の指標である。空間特性は、空間拡散の程度を表し、特に、入力信号が、一つ以上の単一の明確な音源により特徴づけられるか又は強い方向キューがない周囲音により特徴づけられるかを表す。

40

【 0 0 6 7 】

アナライザ 107 は、空間特性を供給し、再生モードをドライバ回路 103 により使用できる複数の音声再生モードから選択するように設けられる選択プロセッサ 109 に結合される。選択プロセッサ 109 は、更に、ドライバ回路 103 に結合され、選択された再生モードを使用するためにドライバ回路 103 を制御する。よって、空間特性が変化すると、選択プロセッサ 109 は、現在の特徴に対する最適再生処理を提供するために再生モ

50

ード間を動的且つ自動的に切り換える。よって、改良された空間経験が達成される。

【0068】

システムは、特に、信号特徴に対する再生モードの短期適合を可能にするように設けられる。よって、速いスイッチングが可能であり、これにより、（長期間で）平均的に最適化されるだけでなく、より多くの瞬間的な信号変化に合うような空間再生を可能にする。

【0069】

従って、アナライザ107は、比較的高い周波数以外、ローパスフィルタをかけられるか、又は平均化される空間特性の形式で推定値を生成するように設けられる。同様に、再生モード間の実際のスイッチングは、比較的高い周波数で実施される。このように、再生モードを選択し、例えばプログラムを通じてこの再生モードを使用するよりはむしろ、図1のシステムは、信号特徴の短期間の変化に合うように、動的に再生モードを適応させる。

【0070】

システムの好適な動的特徴は、個々の実施例の特定の特徴及び好みに依存する。

【0071】

しかしながら、多くの実施例では、特に有利なパフォーマンスは、通常は約50msから5分までの範囲にある間隔で再生モードの更新を許容するシステムで達成される。正確な動的性質は、現在の信号特徴に対する適応の正確さとシステムの信頼性との間のトレードオフと、異なるモード間の切換えに関連したアーチファクトの程度に基づいて選択される。

【0072】

多くの実施例において、空間特性を決定するとき含まれるローパスフィルタは、個々の実施例の特定の好みに依存して、0.001Hzを超えて、0.01Hz、0.1Hz、1Hz、10Hz又は50Hzの3dBカットオフ周波数を好適に持つ。同様に、空間特性は、500秒未満、100秒、10秒、1秒、500ms、100ms又は50msの時定数で好適に決定される。時定数は、ステップ変化をフォローする（漸近的な）最終の $1 - 1/e = 63\%$ に達するまで空間特性がかかる時間として規定される。例えば、空間特性は、追跡するか、又はマルチチャネル信号の一つ以上の空間特徴に依存している。他の全てのパラメータを一定に維持しながら空間特徴のステップ変化は、結果的に空間特性の変化となる。このとき、空間特性を決定するための時定数は、その最終的な（漸近的な）値の領域 $1 - 1/e = 63\%$ に到達するために、この変化がかかる時間として測定される。

【0073】

同様に、切換えは、同様の動特性に従って設けられる。特に、再生モードの切換えのための最大スイッチ周波数は、0.01Hzを超え、0.1Hz、1Hz又は10Hzでさえある。最大周波数は、空間特性の決定及び/又は実際の切換え動作により可能である最も速い切換えである。よって、最大切換え周波数は、システムがフォローできる音声信号の根底にある空間特徴の最も高い周波数変化である。

【0074】

特定の実施例において、ドライバ回路103は、4つの異なる再生モード間を切り換えるように設けられる。

【0075】

第1の再生モードでは、ドライバ回路103は、元のステレオ信号を単に維持し、何れの空間変更も導入しない。よって、この動作モードは、マルチチャネル入力信号の空間特徴を維持する。具体例において、ステレオ入力信号は単にステレオ信号として再生され、すなわち左入力チャネルが左のラウドスピーカーに入力し、右の入力チャネルが右のラウドスピーカーに入力し、何れの信号もセンターのラウドスピーカーに入力されない。よって、この再生モードでは、ドライバ回路103は、元の音声チャネルのステレオ再生を供給する。

【0076】

第2の再生モードでは、ドライバ回路103は、モノラル信号として入力信号を再生する。例えば、2つのステレオチャンネルは（例えば単純な加算により）結合され、結果として生じるモノラル信号はセンターのラウドスピーカーに供給され、何れの信号も左又は右のラウドスピーカーに供給されない。よって、ドライバ回路103の第2の再生モードは、入力信号のダウンミキシングを含み、モノラル再生モードである。斯様な再生モードは、例えばニュース番組のためのニュースを読む人が音源のような中心に配置された単一の音源に音声に対応するシナリオにおいて、特に好適等である。

【0077】

第3の再生モードでは、ドライバ回路103は、空間拡散処理を導入するように設けられる。具体例では、第3の再生モードは、ステレオ拡散アルゴリズムを入力ステレオ信号に適用することを有する。斯様なステレオ拡散は、拡大された空間イメージの知覚が達成されるようにステレオチャンネルの非相関性を提供する傾向がある。様々な空間拡散技術が当業者により知られるだろうし、任意の適切なアルゴリズムが本発明を損なうことなく用いられることは、理解されるだろう。

【0078】

斯様な処理は、サウンドイメージが特定の局所的な音源よりもむしろ周囲音により支配されるとき、特に有利である。例えば、斯様な処理は、多くの楽器を持つ大規模なオーケストラにより作られる音楽を再生するとき、強化された経験を提供する。

【0079】

第4の再生モードでは、ドライバ回路103は、入力信号を一つ以上の主要なソース信号に分け、各一次信号は、特定の主要な音源からだけの音を有するようにしている。当業者は、主要な音源を検出し抽出するための種々異なるアルゴリズムを知っているだろうし、任意の適切なアルゴリズムが本発明を損なわずに用いられることは、理解されるだろう。ドライバ回路103は、更に、主要な音源の抽出の後の信号に対応する残差信号を生成する。第4の再生モードでは、入力ステレオ信号は、このように、一つ以上の主要なサウンドソース信号及び周囲ステレオ又はサラウンド信号に分解される。

【0080】

主要なサウンドソース信号及び残差信号は、その後、異なる空間処理が信号に適用されるように、異なって処理される。単純な例として、空間拡散は、残差信号に適用されるが、主要なサウンドソース信号に適用されない。このように、主要な音源の空間的に明確な位置決めは変更されないのに対し、強化されたサウンドイメージが、周囲音環境に通常は対応する残差信号に対して達成される。更にまた、主要なサウンドソース信号は例えばセンターの空間チャンネルにあり、残差信号は右及び左の空間チャンネルにある。よって、この再生モードでは、ラウドスピーカーのセットにより供給される全ての空間チャンネルが使われ、当該再生モードは入力信号のアップミキシングを有する。

【0081】

音声チャンネルから空間ソース分布を推定する方法が提案されてきた。例えば、マルチチャンネル音声データからの主要な音源の方向の決定及び周囲の音声レベルの推定のための方法は、M.Goodwin及びJ-M.Jotによる'Multichannel surround format conversion and generalized upmix', AES30th int. Conference, Finland, March 2007で提案された。ステレオ混合の複数の音源の分布の推定のための2つの他の方法は、例えば、A.Harma及びC.Fallerによる"Spatial decomposition of time-frequency regions: subbands or sinusoids", AES116th Convention, Berlin, Germany, 8-11 May 2004で検討された。

【0082】

第4の再生モードは、例えば特定の音源と周囲の音又はノイズとの混合である信号に特に適している。

【0083】

アナライザ107による入力信号の音源の空間分布の分析は、例えば、各チャンネルの音声エネルギーの周波数選択分析及び/又はチャンネル間の類似を表す幾つかの適切な数値的

10

20

30

40

50

尺度の変化の周波数選択分析に基づいている。例えば、アナライザ 107 は、MPEG サラウンド規格で使用される分析法と同様の分析法を用いる。よって、これらは、入力信号のサブバンド分解と、異なるチャネルの周波数サブバンド間のエネルギー及び共分散値の計算とに基づいている。しかしながら、例えば信号のパラメータの表現に関する相関計量及び/又は異なるチャネル間の類似性を特徴づけている相互情報量のような、多くの他のアプローチが使用されてもよいことは、理解されるだろう。

【0084】

図2は、図1のシステムで使われる特定のアプローチを例示する。

【0085】

この例では、アナライザ 107 は、入力の左及び右の信号が供給される加算器 201 及び減算器 203 を有する。加算器は2つの信号を合計する、減算器 203 は1つをその他から減算する。加算器 201 は、加算器 201 により生成される和信号の信号エネルギーを計算する第1のエネルギー推定器 205 に結合される。減算器 203 は、減算器により生成される差信号の信号エネルギーを測定する第2のエネルギー推定器 207 に結合される。第1及び第2のエネルギー推定器 205、207 は、和エネルギー及び差エネルギーの空間特性指標に基づいて再生モードを選択する選択プロセッサ 109 に結合される。

【0086】

よって、この例では、再生モードの選択は、左右のチャネル信号間の和信号及び差信号の計算とこれらの信号の短時間エネルギーの比較とに基づく。和信号のエネルギーが差信号のエネルギーより著しく大きいとき、入力ステレオ信号が実質的にモノラルであると推定される。和信号及び差信号のエネルギーが同じレベルにある、又は差信号のエネルギーが和信号のエネルギーより大きいとき、入力信号は、通常のステレオ音声信号であると考えられる。

【0087】

よって、各エネルギー分析期間内の検出値は、以下の式により与えられる。

$$\rho = \begin{cases} 1, & \text{if } E_{\text{sum}} > AE_{\text{diff}} \\ 0, & \text{if } E_{\text{sum}} \leq AE_{\text{diff}} \end{cases}$$

ここで、 $E_{\text{sum}}$  及び  $E_{\text{diff}}$  は、それぞれ和信号及び差信号の短時間エネルギーであり、 $A$  は、通常は、1より著しく大きいスカラー係数（例えば、 $A = 100$ ）である。

【0088】

ドライバ回路 103 の動作、特に異なる再生モード間のスイッチは、動的なマトリックス動作として以下の式のように実行される。

$$\begin{pmatrix} y_l(n) \\ y_r(n) \\ y_c(n) \end{pmatrix} = \begin{pmatrix} 1-p(n) & 0 \\ 0 & 1-p(n) \\ \frac{p(n)}{2} & \frac{p(n)}{2} \end{pmatrix} \begin{pmatrix} x_l(n) \\ x_r(n) \end{pmatrix}$$

ここで、 $x_l(n)$  及び  $x_r(n)$  は元の左及び右ステレオ信号であり、 $n$  はサンプルデジタル信号の指標である。出力  $y_l(n)$ 、 $y_r(n)$  及び  $y_c(n)$  は、それぞれ左、右及びセンターのスピーカに対する駆動値である。

【0089】

よって、この例では、和信号及び差信号の信号エネルギーは、センターのスピーカを使用して実質的にモノラルの再生と、左及び右のスピーカを使用してステレオの再生との間を切り換えるために用いられる。

【0090】

別の例として、和及び差動作は、より汎用の動作により置き換えられてもよい。例えば、主要な音源の方向は、主成分分析法（PCA）（又は、適応固有値分解のような他の同様の方法）により推定される。更に、主要な音源が差信号から除去されるように、加重さ

10

20

30

40

50

れた和及び差が使用されてもよい。これは、構造的に非常に類似しているが、図2の例よりもっと汎用的な解決策を導く。

【0091】

説明されたアプローチは、例えばフーリエ変換により生成された個々の周波数分類、又はフィルタバンクの周波数サブバンドのような異なる周波数間隔で例えば独立して適用されてもよい。

【0092】

具体例では、上記アプローチは、入力信号が実質的にモノラルの特徴を持つかを決定するために最初に用いられる。持つ場合は、第2の再生モード（モノラル表現）が用いられる。持たない場合、すなわち  $\rho = 0$  の場合、更なる処理が、他の再生モードのどれが使用されるべきかを選択するために実施される。これらの再生方法は、具体的には、 $x_l(n)$  及び  $x_r(n)$  に適用される処理を適切に切り換えることにより、切り換わる。例えば、第1の再生モード（入力信号の空間特徴を維持する）に対して、入力チャンネルが  $x_l(n)$  及び  $x_r(n)$ （よって、 $y_l(n)$  及び  $y_r(n)$ ）として直接使われるのに対し、第3の再生モード（拡散）に対して、入力信号が  $x_l(n)$  及び  $x_r(n)$ （よって、 $y_l(n)$  及び  $y_r(n)$ ）として使われ、ラウドスピーカーに供給される前に、空間拡散が、入力信号に最初に適用される。

【0093】

幾つかの実施例では、アナライザ107は、一つ以上の主要な音源を有する主要な音源信号を決定する。その後、主要な音源が抽出された後の残りの信号を表わす残差信号が生成される。最後に、空間特性は、残差信号のエネルギー指標に対する主要な音源信号のエネルギー指標に応じて決定される。

【0094】

例えば、方向性フィルタリング技術は、入力信号のステレオ混合から、主要な音源を抽出するために使用される。この抽出は、当業者に知られているような、ビーム形成アルゴリズム、適応可能アルゴリズム、ブラインドソース分離アルゴリズムを含むマルチチャンネル信号分解のための任意の適当な技術と、マルチチャンネルノイズ抑制のための方法とを使用する。

【0095】

混合からの主要な（又は優位な）音源の抽出の後、主要な音源が除去されたか又は抑制されたマルチチャンネル残差信号が決定される。

【0096】

この場合、検出値は、以下のように計算される。

$$\rho = \begin{cases} 1, & \text{if } E_{\text{prim}} > BE_{\text{res}} \\ 0, & \text{if } E_{\text{prim}} < BE_{\text{res}} \end{cases}$$

ここで、 $E_{\text{prim}}$  は優位な又は主要な音源信号のエネルギー尺度であり、 $E_{\text{res}}$  は残差信号のエネルギー尺度である。パラメータBの値は、通常は、主要な信号抽出の特定の特徴に依存する単位ユニット周辺である。抽出された主要な音源のエネルギーが残余のエネルギーと比較して低い場合、システムは、混合が優位な／主要な音源を含まないと決定する。この場合、第3の再生方法が、強化された空間イメージを提供するために選択される。

【0097】

さもないければ、装置は、残差信号が他の優位な音源を含むかどうか評価するための処理をする。これは、例えば、主要な音源分離を繰り返し残差信号に適用することによりなされる。別の例として、決定は、マルチチャンネル信号間の類似性尺度の計算に基づく。典型的類似性尺度は、ピアソン相関のような様々なタイプの加重相関計量であり、相関関数又は正規化相関関数の最大値に対する推定値である。様々なタイプの大きさ差関数又は相互情報量のような情報理論的尺度を使用することも可能である。尺度が2つの残差信号間の

低い類似性を示す場合、これは、（信号が以前に実質的にモノラルではないとわかったので）ある周囲信号を持つ単一の主要な音源の存在を表す。従って、優位な又は主要な音源信号が、（例えば、センターチャンネルに供給されるモノラル信号として）空間拡散を持たないで再生されるのに対し、空間拡散は左及び右のラウドスピーカーに供給される残余のステレオ信号に適用されて、第4の再生モードは用いられる。

【0098】

しかしながら、残差信号のチャンネルが高い類似性を持つことがわかる場合、これは、入力信号が第1の再生方法により良好に再生された2つの主要な音源から成り、よってこれが選択されることを反映しているようである。

【0099】

多くの実施例で好適にも、異なる再生モード間の切換えは、円滑且つ漸進的な遷移である。これは、異なる再生モードの異なる空間特徴から生じるアーチファクトを低減し緩和する。

【0100】

一例として、モノラルモードからステレオ再生モードへの切換えは、以下の式に従う。

$$\begin{pmatrix} y_l(n) \\ y_r(n) \\ y_c(n) \end{pmatrix} = \begin{pmatrix} 1-p(n) & 0 \\ 0 & 1-p(n) \\ \frac{p(n)}{2} & \frac{p(n)}{2} \end{pmatrix} \begin{pmatrix} x_l(n) \\ x_r(n) \end{pmatrix}$$

ここで、 $p(n) = ap(n-1) + (1-a)$  であり、時間積分係数  $a$  は間隔  $[0, 1]$  内の値である。典型的値は、例えば  $a = 0.95$  である。

【0101】

より一般的な例として、装置は、再生モードの2つ（又はそれ以上）同時に動作するように設けられる。システムが切換えを行って2つの再生モードから生成された信号は、その後、以前の再生モードから新規な再生モードに徐々に変わるような2つのモードの加重で混合される。例えば、各ラウドスピーカーに対して、2つの再生モードにより生成される対応する信号は、

$$y(n) = p(n) \cdot x_p(n) + (1-p(n)) \cdot x_n(n)$$

により加算される。

ここで、 $y(n)$  はスピーカに対する駆動信号であり、 $x_p$  は以前の再生モードにより生成されるサンプルであり、 $x_n$  は新規な再生モードにより生成されるサンプルであり、 $n$  はサンプル指標であり、 $p(n)$  は適切な時間的特徴を持って1から0に徐々に変化する値である。

【0102】

多くの実施例では、10msから1秒までの間隔の遷移時間は、有利なパフォーマンスを提供する傾向がある。遷移時間は、新規な再生モードが、結果として生じる結合信号の10%の加重から90%の加重に変化する時間として測定される。

【0103】

幾つかの実施例では、ドライブ回路103は、空間特性に応答して、選択された再生モードの空間レンダリング技術の特徴を適応させるように更に設けられる。例えば、第3の再生モードに対して、適用される空間拡散の程度は、空間の優先順位に依存して調整される。よって、斯様な例では、入力信号の空間混合の分析は、非相関性の量を制御する、すなわち空間拡散アルゴリズムの「ステレオ拡散パラメータ」を制御するためにも用いられる。例えば、入力信号が複数のソースを持つリッチ且つワイドな空間イメージ、又は例えば識別可能な音源を持たない拡散信号を含むことを空間特性が示す場合、両方のチャンネルに基本的に同じコンテンツがあるときより、再生の際、より多くのステレオ拡散が適用される。第1の場合は、2つの音声チャンネル間の相関の量を評価することにより、第2の場合と区別できる。

【0104】

他の例として、2つの別々の源が左及び右それぞれのチャンネルで主要である信号が考えられてもよい。この場合、意図された空間イメージは、ステレオイメージの2つの明らかに局所的な分離された源（例えば、左にデュエットの歌手及び右にギター）から成る。この場合、チャンネル間の相関は低い。ステレオ拡散が信号に対する相関のため当該信号に適用される場合、生成される空間イメージはワイドになるだろう。しかしながら、この場合、ステレオイメージは、2つの意図されたステレオイメージの明らかに局所的な特徴を欠いてぼやける。従って、多分そうであるだろうイメージの明らかに局所的な源を保存するために、このタイプのコンテンツのための直接的な（拡散しない）ステレオ再生を使用することが、おそらくより良好であろう。ステレオイメージが少数の無相関の源の単純な混合を持つ場合か、又は複数の音源の複雑な混合である場合かを検出することは可能である。これを実施する単純な方法は、左と右のチャンネル間の正規化されたクロス相関Cを分析することである。斯様な推論に基づいて、幾つかの実施例における再生モードの選択は、以下の論理に基づく。

$C < T_{low}$  の場合、コンテンツは左右の2つの相関がない源から成ると考えられ、標準の（拡散がない）ステレオ再生が2つの源の位置を維持するために選択されている。 $T_{low} < C < T_{high}$  の場合、コンテンツは通常の複雑なステレオマテリアルであると考えられる。ステレオ拡散アプローチが、このタイプのコンテンツのための再生のためにしかるべく使われる。

$T_{high} < C$  の場合、コンテンツは一つの明確な源を持つと考えられる。従って、ステレオ再生方法又はモノラルコンテンツのための特定の再生が、このタイプの入力のために選択される。

正規化された相関関数は、例えば以下により与えられるピアソン相関であるか、

$$C = E[x_l(n)x_r(n)] / \sqrt{(E[x_l(n)x_l(n)]E[x_r(n)x_r(n)])}$$

又は以下の式により与えられるAvenidaにより提案される正規化された相関尺度である(C.Avenida, Frequency-domain source identification and manipulation in stereo mixes for enhancement, suppression and re-panning applications, IEEE Proc. WASPAA, NY, USA, 2003)。

$$C = 2E[x_l(n)x_r(n)] / (E[x_l(n)x_l(n)] + E[x_r(n)x_r(n)])$$

#### 【0105】

検出は、入力信号の小さな時間周波数セグメントのチャンネル間の相関及びレベル差の統計値に基づき得る。

#### 【0106】

図1のシステムは、多くのシナリオで多くの実生活信号に対して改良されたリスニング経験を提供する。特に、アップミキシングに基づくシステムに対する空間経験は、多くのシナリオで改善される。例えば、ステレオ信号からセンターチャンネルを抽出しようとするアップミキシングアルゴリズムは、センターの音源がサウンドイメージに存在するが、ステレオ混合内に明確なセンターイメージがない場合には理想的に必ずしも働かないとき、非常に良好なパフォーマンスを提供する。実際、センターチャンネル抽出方法が斯様なコンテンツのために用いられる場合、結果的にステレオイメージの幅の低減になる。説明されたアプローチは、適切なアップミックスアプローチを使用するために入力信号の再生を動的に適合可能にする。

#### 【0107】

幾つかの実施例では、再生モードの選択は、更に、入力信号に対するコンテンツ特性を考える。斯様な例は、信号に対するコンテンツ特徴を決定するように設けられるコンテン

10

20

30

40

50

ツプロセッサ 3 0 1 を含むために修正された図 1 のシステムを示す図 3 に図示される。

【 0 1 0 8 】

コンテンツ特徴は、例えば、ジャンル、音声信号と関連したプログラムタイプ（例えば、音声、例えばテレビ又はラジオ番組のようなメディアアイテムと関連している場合）、音声と関連したアーティスト等を示す。コンテンツ特徴は、例えば入力信号と関連したメタデータから決定される。このように、幾つかのシナリオでは、メタデータは、個別に受信されるか、例えば音声信号に埋め込まれている。コンテンツプロセッサ 3 0 1 は、入力信号のコンテンツを記述するデータを抽出するように設けられる。

【 0 1 0 9 】

他の実施例では、コンテンツプロセッサ 3 0 1 は、受信入力信号のコンテンツ分析を実施し、斯様なコンテンツ分析に基づいてコンテンツ特徴を決定するように設けられる。例えば、コンテンツプロセッサ 3 0 1 は、信号が主にスピーチ、音楽又は例えば大きい爆発を含むかどうかを決定するために、信号を分析する。その後、コンテンツプロセッサ 3 0 1 は、対応するタイプのコンテンツを推定し、分析に基づいて、例えば、ニュース番組、音楽プログラム及びアクションフィルムを選択する。種々異なるコンテンツ分析アプローチが当業者に知られていて、適切なアルゴリズムが用いられることは理解されるだろう。視聴信号（すなわち入力音声信号がビデオ信号と結合されている）に対して、コンテンツ分析は、代替的に又は追加的に、入力信号と関連したビデオ信号に基づいてもよい。

【 0 1 1 0 】

コンテンツ特徴は、使用するため再生モードの選択にそれを含める選択プロセッサ 1 0 9 に供給される。特に、異なる再生モード間の短期切換えは、空間特性の短期変化に基づいて依然決定されるが、正確な切換え基準は、コンテンツが何であるかに依存して修正される。例えば、システムは、ニュース番組のためより、アクション映画のための空間拡散アプローチへ切換えそうである。

【 0 1 1 1 】

よって、コンテンツタイプを表すデータが、使用のため最適空間再生方法を選択する際に使われる。特に、コンテンツ特徴は、再生モード選択戦略の信頼性を強化するために用いられる。決定の際、コンテンツ特徴を含むことは、不適当な再生モードが選択されるリスクを低下できる。

【 0 1 1 2 】

例えば、場合によっては、信号の空間分析は、結果的に適切な再生モードを明らかに示さない空間特性になる。この場合、再生モードを選択するとき、コンテンツを考慮することが望ましい。よって、空間信号分析が 4 つの再生クラスの 1 つの信号の空間混合を明らかに分類せず、これらのうちの 2 つ以上の間で不確かな「グレイ」領域にある場合、コンテンツ特徴が考慮される。幾つかの実施例では、再生モードの各々に対応する空間特性の間隔は、例えば特定の特性に依存する。これは、例えばニュース番組及びアクションフィルムに対して異なる、変更なしのステレオ再生モードと拡散ステレオ再生モードとの間の選択に例えば結果としてなる。よって、拡散は、アクションフィルムに対してよりニュース番組に対してあまり使用されない。

【 0 1 1 3 】

幾つかの実施例では、ドライバ回路 1 0 3 は、コンテンツ特徴に応答して、選択された再生モードの空間レンダリング技術の特徴を適応させる。よって、入力信号のコンテンツタイプに関する情報を反映するコンテンツ特徴が、選択された空間再生モードのパラメータを制御するために使われる。例えば、ステレオ拡散が最適再生方法であるとシステムが決定するとき適用される拡散の量は、コンテンツタイプに依存して調整される。この目的のために、コンテンツタイプの分類は、例えば、「ニュース」、「映画」、「音楽」、「ドキュメンタリー」などのような分類を識別する高いレベルでなされてもよい。しかしながら、分類は、サブタイプ、例えば音楽の異なるジャンル又は映画の異なるタイプの分類をするためにも有益である。例えば、音楽の特定のジャンルは、通常、どちらかといえば密接な防音スタジオ及び音響雰囲気（例えばシンガーソングライター又は室内楽）と関連す

10

20

30

40

50

る一方、他のジャンルは広い防音スタジオ及び非常にゆったりしたルーム音響効果（例えばクワイア音楽）と関連する。音楽のジャンルを知っていることは、音声信号の空間混合の分析に加えて、適切な再生モードを選択し、及び／又は空間再生モードのパラメータを設定するのを助ける。

【 0 1 1 4 】

上記説明は、ラウドスピーカーのセットが、入力信号（特に２つのチャンネル）より多くの空間チャンネル（特に３つの空間チャンネル）を供給する実施例に焦点を当てた。しかしながら、他の実施例では、ラウドスピーカーのセットは、入力信号より多くの空間チャンネルを提供しないことも理解されるだろう。

【 0 1 1 5 】

実際、多くの実施例で、ラウドスピーカーのセットが入力信号より少ない空間チャンネルを提供することが有利である。例えば、７つのチャンネルサラウンドサウンド入力信号は、３つの空間チャンネルで再生される。このような実施例では、潜在的に複雑な空間処理は、有利なパフォーマンスを提供するために用いられ、説明された原理は、どの再生モードを入力信号の特定の空間特性に適用すべきかを選択するために用いられる。よって、異なるダウンミキシングアルゴリズムが、入力信号の空間特徴に依存して用いられる。

【 0 1 1 6 】

明快さのための上記説明は、種々の機能的回路、ユニット及びプロセッサを参照して本発明の実施例を説明されてきたことは理解されるだろう。しかしながら、種々の機能的回路、ユニット又はプロセッサ間の機能性の任意の適切な配信が本発明を損なうことなく使用されてもよいことは理解されるだろう。例えば、プロセッサ又はコントローラにより実施されるために例示された機能は、同じプロセッサ又はコントローラにより実施されてもよい。よって、特別な機能性ユニット又は回路の参照は、厳格に論理的若しくは物理的構造又は組織を示すよりむしろ説明された機能を提供するための適切な手段の参照として見られるべきである。

【 0 1 1 7 】

本発明は、ハードウェア、ソフトウェア、ファームウェア又はこれらの任意の組み合わせを含む任意の適切な形式で実行できる。本発明は、オプション的には、一つ以上のデータプロセッサ及び／又はデジタル信号プロセッサで実行するコンピュータソフトウェアとして少なくとも部分的に実行されてもよい。本発明の実施例の要素及び部品が、任意の適当な態様で物理的に、機能的に及び論理的に実行されてもよい。実際、機能は、単一のユニット、複数のユニット、又は他の機能ユニットの一部として実行されてもよい。例えば、本発明は、単一のユニットで実行されてもよいし、種々のユニット、回路及びプロセッサ間で物理的及び／又は機能的に配信されてもよい。

【 0 1 1 8 】

本発明が幾つかの実施例と関連して説明されてきたが、ここで説明される特定の形式に限定されることを意図していない。むしろ、本発明の範囲は、添付の請求項によってのみ限定される。加えて、特徴が特定の実施例と関連して説明されるように見えるが、当業者は、説明される実施例の様々な特徴が本発明に従って組み合わせられてもよいことを認識するだろう。請求項において、「有する」という用語は、他の要素又はステップの存在を除外しない。

【 0 1 1 9 】

更にまた、個々にリストされたが、複数の手段、要素、回路又は方法のステップが単一の回路、ユニット又はプロセッサにより実行されてもよい。加えて、個々の特徴が異なる請求項に含まれているが、これらは、好適に組み合わせられてもよく、異なる請求項に含まれていることは、特徴の組み合わせが可能ではない及び／又は好適ではないことを暗示していない。一つのカテゴリの請求項内の特徴に含まれることは、このカテゴリへの限定を暗示しているわけではなく、むしろ特徴が適当に他の請求項のカテゴリに等しく適用可能であることを示す。更にまた、請求項内の特徴の順番は、特徴が働かなければならない任意の特定の順番を暗示するわけではなく、特に、方法の請求項の個別のステップの順番

10

20

30

40

50

は、これらのステップがこの順番に実施されなければならないことを暗示するものではない。むしろ、これらのステップは、任意の好適な順番で実施されてもよい。加えて、単一の参照符号は複数を除外しない。このように、「a」「an」「第1」「第2」等の用語は、複数であることを妨げない。請求項内の参照符号は、単に明確な例として提供されていて、いずれの態様においても請求項の範囲を限定するものとして考慮されるべきではない。

【図1】

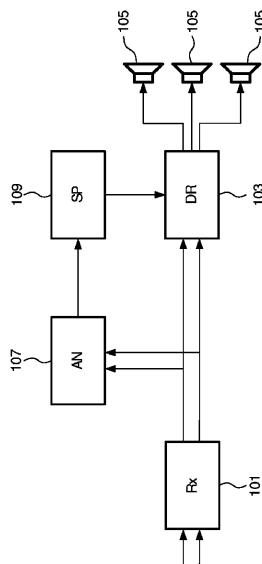


FIG. 1

【図2】

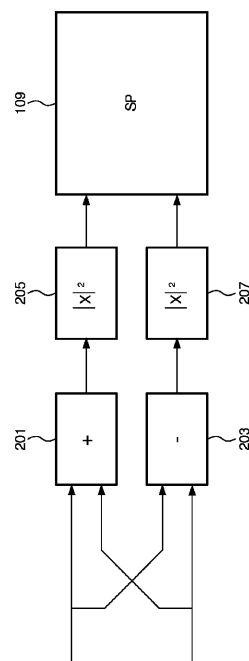
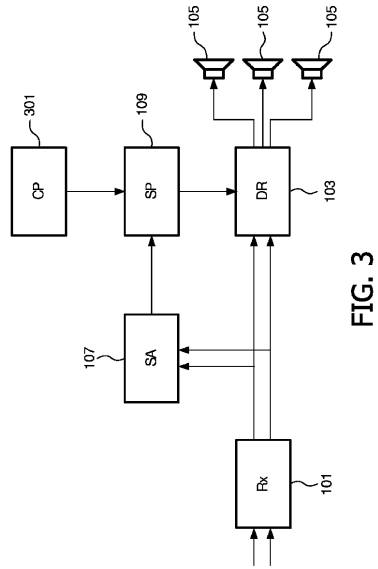


FIG. 2

【図 3】



---

フロントページの続き

(72)発明者 ハルマ アキ サカリ

オランダ国 5 6 5 6 アーエー アインドーフエン ハイ テック キャンパス ビルディング  
4 4

(72)発明者 デ ブルアイン ワーナー パウルス ジョセフス

オランダ国 5 6 5 6 アーエー アインドーフエン ハイ テック キャンパス ビルディング  
4 4

審査官 大野 弘

(56)参考文献 特開 2 0 0 6 - 2 5 4 1 8 7 ( J P , A )

特表 2 0 0 4 - 5 0 4 7 8 7 ( J P , A )

特表 2 0 0 8 - 5 2 1 0 4 6 ( J P , A )

米国特許出願公開第 2 0 0 9 / 0 1 3 6 0 4 8 ( U S , A 1 )

特表 2 0 0 8 - 5 1 2 8 9 0 ( J P , A )

(58)調査した分野(Int.Cl. , D B 名)

H 0 4 S 5 / 0 2