

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5820820号
(P5820820)

(45) 発行日 平成27年11月24日 (2015. 11. 24)

(24) 登録日 平成27年10月9日 (2015. 10. 9)

(51) Int. Cl.	F I				
G 1 0 L 19/008 (2013. 01)	G 1 0 L	19/008	1 0 0		
H 0 4 S 1/00 (2006. 01)	H 0 4 S	1/00		Z	
G 1 0 L 19/00 (2013. 01)	H 0 4 S	1/00		K	
	G 1 0 L	19/00	4 0 0 Z		

請求項の数 15 (全 38 頁)

(21) 出願番号	特願2012-548400 (P2012-548400)	(73) 特許権者	591037214
(86) (22) 出願日	平成23年1月11日 (2011. 1. 11)		フラウンホッフアーゲゼルシャフト ツ
(65) 公表番号	特表2013-517518 (P2013-517518A)		ァ フェルダールング デァ アンゲヴァ
(43) 公表日	平成25年5月16日 (2013. 5. 16)		ンテン フォアシュンク エー. ファオ
(86) 国際出願番号	PCT/EP2011/050265		ドイツ連邦共和国 80686 ミュンヘ
(87) 国際公開番号	W02011/086060	(74) 代理人	100079577
(87) 国際公開日	平成23年7月21日 (2011. 7. 21)		弁理士 岡田 全啓
審査請求日	平成24年9月13日 (2012. 9. 13)	(72) 発明者	ヴィルカモ ユハ
(31) 優先権主張番号	10174230. 2		ドイツ連邦共和国 90409 ニュルン
(32) 優先日	平成22年8月26日 (2010. 8. 26)		ベルグ ピルクハイマーシュトラーセ 9
(33) 優先権主張国	欧州特許庁 (EP)		1
(31) 優先権主張番号	61/295, 278	(72) 発明者	プログスティーズ ヤン
(32) 優先日	平成22年1月15日 (2010. 1. 15)		ドイツ連邦共和国 91058 エアラン
(33) 優先権主張国	米国 (US)		ゲン イン デァ ツァイル 9アー
前置審査			最終頁に続く

(54) 【発明の名称】 ダウンミックス信号と空間パラメータ情報からダイレクト/アンビエンス信号を抽出する装置および方法

(57) 【特許請求の範囲】

【請求項 1】

ダウンミックス信号 (1 1 5) と空間パラメータ情報 (1 0 5) から、ダイレクトおよび/またはアンビエンス信号 (1 2 5 1、1 2 5 2) を抽出する装置 (1 0 0) であって、前記ダウンミックス信号 (1 1 5) と前記空間パラメータ情報 (1 0 5) は、前記ダウンミックス信号 (1 1 5) よりも多くのチャンネル ($C h_1 \dots C h_N$) を有する多重チャンネルオーディオ信号 (1 0 1) を表し、前記空間パラメータ情報 (1 0 5) は、前記多重チャンネルオーディオ信号 (1 0 1) のチャンネル間関係を含み、

前記空間パラメータ情報 (1 0 5) に基づいて、前記多重チャンネルオーディオ信号 (1 0 1) のダイレクト部分のダイレクトレベル情報 (1 1 3) を推定し、および/または前記多重チャンネルオーディオ信号 (1 0 1) のアンビエント部分のアンビエンスレベル情報 (1 1 3) を推定する、ダイレクト/アンビエンス推定器 (1 1 0) と、

前記ダイレクト部分の前記推定されたダイレクトレベル情報 (1 1 3) に基づいて、または前記アンビエント部分の前記推定されたアンビエンスレベル情報 (1 1 3) に基づいて、前記ダウンミックス信号 (1 1 5) から、ダイレクト信号部分 (1 2 5 1) および/またはアンビエント信号部分 (1 2 5 2) を抽出する、ダイレクト/アンビエンス抽出器 (1 2 0) と、

を備え、

前記ダイレクト/アンビエンス抽出器 (4 2 0) は、前記ダイレクト部分の前記推定されたダイレクトレベル情報 (1 1 3) または前記アンビエント部分の前記推定されたアン

ピエンスレベル情報(113)をダウンミックスし、前記ダイレクト部分または前記アンビエント部分のダウンミックスされたレベル情報を取得し、前記ダウンミックスされたレベル情報に基づいて、前記ダウンミックス信号(115)から、前記ダイレクト信号部分(125 1)または前記アンビエント信号部分(125 2)を抽出するように構成された、装置。

【請求項2】

前記ダイレクト/アンビエンス抽出器(420)は、前記ダイレクト部分の前記推定されたダイレクトレベル情報(113)をコヒーレント和と結合し、前記アンビエント部分の前記推定されたアンビエンスレベル情報(113)を非コヒーレント和と結合することによって、前記ダイレクト部分の前記推定されたダイレクトレベル情報(113)または前記アンビエント部分の前記推定されたアンビエンスレベル情報(113)のダウンミックスを実行するように更に構成された、請求項1に記載の装置。

10

【請求項3】

前記ダイレクト/アンビエンス抽出器(520)は、前記ダイレクト部分または前記アンビエント部分の前記ダウンミックスされたレベル情報(555 1、555 2)からゲインパラメータ(565 1、565 2)を導き出し、前記導き出されたゲインパラメータ(565 1、565 2)を前記ダウンミックス信号(115)に適用し、前記ダイレクト信号部分(125 1)または前記アンビエント信号部分(125 2)を取得するように更に構成された、請求項1または2に記載の装置。

【請求項4】

20

前記ダイレクト/アンビエンス抽出器(520)は、前記ダイレクト部分または前記アンビエント部分の前記ダウンミックスされたレベル情報(555 1、555 2)から、ダイレクトからトータルへの(DTT)エネルギーレシオまたはアンビエントからトータルへの(ATT)エネルギーレシオを決定し、前記決定されたDTTエネルギーレシオまたはATTエネルギーレシオに基づいて、抽出パラメータを、前記ゲインパラメータ(565 1、565 2)として用いるように更に構成された、請求項3に記載の装置。

【請求項5】

前記ダイレクト/アンビエンス抽出器(520)は、二次の $M \times M$ 抽出マトリックスを前記ダウンミックス信号(115)に適用することによって、前記ダイレクト信号部分(125 1)または前記アンビエント信号部分(125 2)を抽出するように構成され、前記二次の $M \times M$ の抽出マトリックスのサイズ(M)は、ダウンミックスチャンネル($Ch_1 \dots Ch_M$)の数(M)に相当する、請求項1~4のいずれかに記載の装置。

30

【請求項6】

前記ダイレクト/アンビエンス抽出器(520)は、前記ダイレクト信号部分(125 1)を取得するために、第1の複数の抽出パラメータを前記ダウンミックス信号(115)に適用し、前記アンビエント信号部分(125 2)を取得するために、第2の複数の抽出パラメータを前記ダウンミックス信号(115)に適用するように更に構成され、前記第1の複数の抽出パラメータと前記第2の複数の抽出パラメータは、対角行列を構成している、請求項5に記載の装置。

【請求項7】

40

前記ダイレクト/アンビエンス推定器(110)は、前記ダイレクト/アンビエンス推定器(110)によって受信された前記空間パラメータ情報(105)と、前記ダウンミックス信号(115)の少なくとも2つのダウンミックスチャンネル(825)に基づいて、前記多重チャンネルオーディオ信号(101)の前記ダイレクト部分の前記ダイレクトレベル情報(113)または前記多重チャンネルオーディオ信号(101)の前記アンビエント部分の前記アンビエンスレベル情報(113)を推定するように構成された、請求項1~6のいずれかに記載の装置。

【請求項8】

前記ダイレクト/アンビエンス推定器(710)は、前記空間パラメータ情報(105)を用いて、前記多重チャンネルオーディオ信号(101)の各チャンネル(Ch_i)に

50

対して、ステレオアンビエンス推定式を適用するように構成され、

前記ステレオアンビエンス推定式は、 i のデシベル値であるチャンネルレベル差 (CLD_i) と、前記チャンネル Ch_i のチャンネル間コヒーレンス (ICC_i) パラメータに依存する次式によって与えられ、

$$DTT_i = f_{DTT} [\sigma_i(Ch_i, R), ICC_i(Ch_i, R)],$$

$$ATT_i = 1 - DTT_i$$

R は残りのチャンネルの線形結合である、請求項 1 ~ 7 のいずれかに記載の装置。

【請求項 9】

前記ダイレクト/アンビエンス抽出器 (620) は、チャンネルクロスミックスによる最小平均二乗 (LMS) 解法によって、前記ダイレクト信号部分 (125 1) または前記アンビエント信号部分 (125 2) を抽出するように構成され、前記 LMS 解法は等しいアンビエンスレベルを必要としない、請求項 1 ~ 8 のいずれかに記載の装置。

【請求項 10】

前記ダイレクト/アンビエンス抽出器 (620) は、前記 LMS 解法がステレオチャンネルダウンミックス信号に限定されないように、信号モデルを仮定することによって前記 LMS 解法を導き出すように構成された、請求項 9 に記載の装置。

【請求項 11】

前記ダイレクト信号部分 (125 1) を処理し、第 1 のバイノーラル出力信号 (915) を取得する、バイノーラル・ダイレクトサウンド・レンダリングデバイス (910) と、

前記アンビエント信号部分 (125 2) を処理し、第 2 のバイノーラル出力信号 (1015) を取得する、バイノーラル・アンビエントサウンド・レンダリングデバイス (1010) と、

前記第 1 のバイノーラル出力信号 (915) と前記第 2 のバイノーラル出力信号 (1015) を結合し、結合されたバイノーラル出力信号 (1135) を取得する、結合器 (1130) と、

を更に備えた、請求項 1 ~ 10 のいずれかに記載の装置。

【請求項 12】

前記バイノーラル・アンビエントサウンド・レンダリングデバイス (1010) は、前記第 2 のバイノーラル出力信号 (1015) を提供するために、前記アンビエント信号部分 (125 2) に、空間効果および/またはフィルタを適用するように構成され、前記第 2 のバイノーラル出力信号 (1015) は、現実の拡散音場の両耳間のコヒーレンスに適合する、請求項 11 に記載の装置。

【請求項 13】

前記バイノーラル・ダイレクトサウンド・レンダリングデバイス (910) は、前記第 1 のバイノーラル出力信号 (915) を取得するために、頭部関連伝達関数 (HRTF) に基づいてフィルタを通して前記ダイレクト信号部分 (125 1) を供給するように構成された、請求項 11 または 12 に記載の装置。

【請求項 14】

ダウンミックス信号 (115) と空間パラメータ情報 (105) から、ダイレクトおよび/またはアンビエンス信号 (125 1、125 2) を抽出する方法 (100) であって、前記ダウンミックス信号 (115) と前記空間パラメータ情報 (105) は、前記ダウンミックス信号 (115) よりも多くのチャンネル ($Ch_1 \dots Ch_N$) を有する多重チャンネルオーディオ信号 (101) を表し、前記空間パラメータ情報 (105) は、前記多重チャンネルオーディオ信号 (101) のチャンネル間関係を備え、

前記空間パラメータ情報 (105) に基づいて、前記多重チャンネルオーディオ信号 (101) のダイレクト部分のダイレクトレベル情報 (113) および/または前記多重チャンネルオーディオ信号 (101) のアンビエント部分のアンビエンスレベル情報 (113) を推定するステップ (110) と、

10

20

30

40

50

前記ダイレクト部分の前記推定されたダイレクトレベル情報(113)に基づいて、または前記アンビエント部分の前記推定されたアンビエンスレベル情報(113)に基づいて、前記ダウンミックス信号(115)からダイレクト信号部分(125 1)および/またはアンビエント信号部分(125 2)を抽出するステップ(120)と、
を備え、

前記方法は、前記ダイレクト部分の前記推定されたダイレクトレベル情報(113)または前記アンビエント部分の前記推定されたアンビエンスレベル情報(113)をダウンミックスし、前記ダイレクト部分または前記アンビエント部分のダウンミックスされたレベル情報を取得するステップを更に備え、

前記ダウンミックス信号(115)から前記ダイレクト信号部分(125 1)および/またはアンビエント信号部分(125 2)を抽出するステップは、前記ダウンミックスされたレベル情報に基づいて実施される、方法。

【請求項15】

コンピュータプログラムがコンピュータ上で実行されるとき、請求項14に記載の方法(100)を実行するプログラムコードを有する、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、オーディオ信号処理に関し、特に、ダウンミックス信号と空間パラメータ情報からダイレクト/アンビエンス信号を抽出する装置および方法に関する。本発明の更なる実施形態は、オーディオ信号のバイノーラル再生を強化するためのダイレクト/アンビエンス分離の利用に関する。また更なる実施形態は、二つ以上のチャンネルを有する多重チャンネルサウンドのバイノーラル再生に関する。多重チャンネルサウンドを有する代表的なオーディオコンテンツは、映画のサウンドトラックと多重チャンネルの音楽レコーディングである。

【背景技術】

【0002】

人間の空間聴覚システムは、おおよそ2つの部分においてサウンドを処理する傾向がある。一方において、ローカライズ可能な部分、言い換えればダイレクト部分であり、他方において、ローカライズ不能な部分、言い換えればアンビエント部分である。これらの2つのオーディオ成分にアクセスすることが望ましいバイノーラルサウンド再生や多重チャンネルアップミックスのような多くのオーディオ処理アプリケーションが存在する。

【0003】

従来技術において、非特許文献1、非特許文献2、非特許文献3、特許文献1、特許文献2、および特許文献3において解説されたようなダイレクト/アンビエンス分解の方法が知られており、さまざまなアプリケーションに用いることができる。最新技術のダイレクト アンビエンス分離アルゴリズムは、周波数バンドにおけるステレオサウンドのチャンネル間信号比較に基づいている。

【0004】

さらに、非特許文献4において、アンビエンス抽出を備えたバイノーラル再生が言及されている。バイノーラル再生に関連するアンビエンス抽出は、非特許文献5においても言及されている。後者の論文は、各々のチャンネルにおけるダイレクト成分の適応な最小平均二乗クロスチャンネルフィルタリングを用いて、ステレオマイクロホン録音におけるアンビエンス抽出にフォーカスしている。空間オーディオコーデック、例えばMP3サラウンドは、非特許文献6および非特許文献7に記載されているように、通常は、オーディオを多重チャンネルに拡張する空間サイド情報と組合せた1つまたは2つのチャンネルオーディオストリームから成る。

10

20

30

40

50

【先行技術文献】

【非特許文献】

【0005】

【非特許文献1】「空間オーディオ符号化および強化のための主要アンビエンス信号分解とベクトルベースのローカライズ」、Goodwin、Jot、音響・音声および信号処理に関するIEEE国際会議、2007年4月

【非特許文献2】「ステレオ録音からの相関ベースのアンビエンス抽出」、Merimaa、Goodwin、Jot、AES第123回大会、ニューヨーク、2007年

【非特許文献3】「ステレオ信号の多重スピーカ再生」、C. Faller、AESジャーナル、2007年10月

【非特許文献4】「空間オーディオシーン符号化に基づくバイノーラル3次元オーディオレンダリング」、Goodwin、Jot、AES第123回大会、ニューヨーク2007年

【非特許文献5】J. UsherおよびJ. Benesty、「空間音響品質の強化：新規な反響抽出オーディオアップミックス装置」、IEEEオーディオ・音声・言語処理議事録、第15巻、2141-2150頁、2007年9月

【非特許文献6】ISO/IEC23003-1 MPEGサラウンド

【非特許文献7】Breebaart, J., Herre, J., Villemoes, L., Jin, C., Kjoerling, K., Plogsties, J., Koppens, J., 「多重チャンネルはモバイルに進む：MPEGサラウンド・バイノーラル・レンダリング」、第29回AES会議議事録、ソウル、韓国、2006年

【特許文献】

【0006】

【特許文献1】「複合類似度インデックスを用いたステレオオーディオ信号の主要アンビエンス分解」、Goodwin 他、米国特許公報第2009/0198356号公報、2009年8月

【特許文献2】「特許出願名称：ステレオ信号から多重チャンネルオーディオ信号を生成する方法」、発明者：Christof Faller、代理人：FISH & RICHARDSON P.C.、承継人：LG ELECTRONICS, INC.、出所：MINNEAPOLIS, MN US、IPC8クラス：AH04R500FI、USPCクラス：3811

【特許文献3】「ステレオ信号のためのアンビエンス生成」、Avendano 他、発行日：2009年7月28日、出願番号：10/163,158、出願日：2002年6月4日

【発明の概要】

【0007】

しかしながら、MPEGサラウンド(MPS)およびパラメータステレオ(PS)のような最新のパラメータオーディオ符号化技術は、付加的な空間サイド情報に加えて、低減された数のいくつかのケースではただ1つのオーディオダウンミックスチャンネルを提供するだけである。「オリジナル」の入力チャンネル間の比較は、サウンドの意図された出力フォーマットへの最初の復号化の後で可能となるだけである。

【0008】

それ故、ダウンミックス信号および空間パラメータ情報からダイレクト信号部分またはアンビエント信号部分を抽出するコンセプトが必要である。しかしながら、パラメータサイド情報を用いたダイレクト/アンビエンス抽出に対する既存の解法はない。

【0009】

それ故に、本発明の目的は、空間パラメータ情報の使用によって、ダウンミックス信号からダイレクト信号部分またはアンビエント信号部分を抽出するコンセプトを提供することである。

【0010】

この目的は、請求項1に記載の装置、請求項15に記載の方法、または請求項16に記載のコンピュータプログラムによって達成される。

【0011】

10

20

30

40

50

本発明の根底にある基本概念は、空間パラメータ情報に基づいて多重チャンネルオーディオ信号のダイレクト部分またはアンビエント部分のレベル情報が推定され、推定されたレベル情報に基づいてダウンミックス信号からダイレクト信号部分またはアンビエント信号部分が抽出される時に、上述のダイレクト/アンビエンス抽出を達成することができるということである。ここで、ダウンミックス信号と空間パラメータ情報は、ダウンミックス信号より多くのチャンネルを有する多重チャンネルオーディオ信号を表す。この尺度は、空間パラメータサイド情報を用いることによって、1つ以上の入力チャンネルを有するダウンミックス信号からダイレクトおよび/またはアンビエンス抽出を可能とする。

【0012】

本発明の一実施形態によれば、ダウンミックス信号と空間パラメータ情報からダイレクト/アンビエンス信号を抽出する装置は、ダイレクト/アンビエンス推定器およびダイレクト/アンビエンス抽出器を備える。ダウンミックス信号と空間パラメータ情報は、ダウンミックス信号より多くのチャンネルを有する多重チャンネルオーディオ信号を表す。さらに、空間パラメータ情報は、多重チャンネルオーディオ信号のチャンネル間関係を備える。ダイレクト/アンビエンス推定器は、空間パラメータ情報に基づいて、多重チャンネルオーディオ信号のダイレクト部分またはアンビエント部分のレベル情報を推定するように構成される。ダイレクト/アンビエンス抽出器は、ダイレクト部分またはアンビエント部分の推定レベル情報に基づいて、ダウンミックス信号から、ダイレクト信号部分またはアンビエント信号部分を抽出するように構成される。

【0013】

本発明の他の実施形態によれば、ダウンミックス信号と空間パラメータ情報からダイレクト/アンビエンス信号を抽出する装置は、バイノーラル・ダイレクトサウンド・レンダリングデバイス、バイノーラル・アンビエントサウンド・レンダリングデバイスおよび結合器を更に備える。バイノーラル・ダイレクトサウンド・レンダリングデバイスは、ダイレクト信号部分を処理し、第1のバイノーラル出力信号を取得するように構成される。バイノーラル・アンビエントサウンド・レンダリングデバイスは、アンビエント信号部分を処理し、第2のバイノーラル出力信号を取得するように構成される。結合器は、第1のバイノーラル出力信号と第2のバイノーラル出力信号を組合せ、複合バイノーラル出力信号を取得するように構成される。従って、オーディオ信号のダイレクト信号部分およびアンビエント信号部分が別々に処理されるオーディオ信号のバイノーラル再生を提供することができる。

【図面の簡単な説明】

【0014】

以下において、本発明の実施形態が、以下の添付図面を参照して説明される。

【図1】多重チャンネルオーディオ信号を表すダウンミックス信号と空間パラメータ情報からダイレクト/アンビエンス信号を抽出する装置の一実施形態のブロック図を示す。

【図2】パラメータステレオオーディオ信号を表すモノラルダウンミックス信号と空間パラメータ情報からダイレクト/アンビエンス信号を抽出する装置の一実施形態のブロック図を示す。

【図3 a】本発明の一実施形態に係る多重チャンネルオーディオ信号のスペクトル分解の図解図を示す。

【図3 b】図3 aのスペクトル分解に基づく多重チャンネルオーディオ信号のチャンネル間関係を計算する図解図を示す。

【図4】推定レベル情報のダウンミックスを備えるダイレクト/アンビエンス抽出器の実施形態のブロック図を示す。

【図5】ゲインパラメータをダウンミックス信号に適用することによるダイレクト/アンビエンス抽出器の更なる実施形態のブロック図を示す。

【図6】チャンネルクロスミックスを備えたLMS解法に基づくダイレクト/アンビエンス抽出器の更なる実施形態のブロック図を示す。

【図7 a】ステレオアンビエンス推定式を用いたダイレクト/アンビエンス推定器の実施

10

20

30

40

50

形態のブロック図を示す。

【図 7 b】例示的なダイレクトからトータルへのエネルギーレシオ対チャンネル間コヒーレンスのグラフを示す。

【図 8】本発明の実施形態に係るエンコーダ/デコーダシステムのブロック図を示す。

【図 9 a】本発明の実施形態に係るバイノーラル・ダイレクトサウンド・レンダリングの概観のブロック図を示す。

【図 9 b】図 9 a のバイノーラル・ダイレクトサウンド・レンダリングの詳細のブロック図を示す。

【図 10 a】本発明の実施形態に係るバイノーラル・アンビエントサウンド・レンダリングの概観のブロック図を示す。

10

【図 10 b】図 10 a のバイノーラル・アンビエントサウンド・レンダリングの詳細のブロック図を示す。

【図 11】多重チャンネルオーディオ信号のバイノーラル再生の実施形態の概念的ブロック図を示す。

【図 12】バイノーラル再生を含むダイレクト/アンビエンス抽出の実施形態の全体ブロック図を示す。

【図 13 a】フィルタバンクドメインにおいてモノラルダウンミックス信号からダイレクト/アンビエント信号を抽出する装置の実施形態のブロック図を示す。

【図 13 b】図 13 a のダイレクト/アンビエンス抽出の実施形態のブロック図を示す。

【図 14】本発明の更なる実施形態に係る例示的な M P E G サラウンド復号化スキームの図解図を示す。

20

【発明を実施するための形態】

【0015】

図 1 は、ダウンミックス信号 115 と空間パラメータ情報 105 から、ダイレクト/アンビエンス信号 125₁、125₂ を抽出する装置 100 の実施形態のブロック図を示す。図 1 に示されるように、ダウンミックス信号 115 と空間パラメータ情報 105 は、ダウンミックス信号 115 より多いチャンネル $C h_1 \dots C h_N$ を有する多重チャンネルオーディオ信号 101 を表す。空間パラメータ情報 105 は、多重チャンネルオーディオ信号 101 のチャンネル間関係を備えることができる。特に、装置 100 は、ダイレクト/アンビエンス推定器 110 とダイレクト/アンビエンス抽出器 120 を備える。ダイレクト/アンビエンス推定器 110 は、空間パラメータ情報 105 に基づいて、多重チャンネルオーディオ信号 101 のダイレクト部分またはアンビエント部分のレベル情報 113 を推定するように構成することができる。ダイレクト/アンビエンス抽出器 120 は、ダイレクト部分またはアンビエント部分の推定レベル情報 113 に基づいて、ダウンミックス信号 115 からダイレクト信号部分 125₁ またはアンビエント信号部分 125₂ を抽出するように構成することができる。

30

【0016】

図 2 は、パラメータステレオオーディオ信号 201 を表すモノラルダウンミックス信号 215 と空間パラメータ情報 105 からダイレクト/アンビエンス信号 125₁、125₂ を抽出する装置 200 の実施形態のブロック図を示す。図 2 の装置 200 は、図 1 の装置 100 と同じブロックを基本的に備える。それ故、同様の実施態様および/または機能を有する同一のブロックは、同じ符号で示される。さらに、図 2 のパラメータステレオオーディオ信号 201 は、図 1 の多重チャンネルオーディオ信号 101 に相当することができる。図 2 のモノラルダウンミックス信号 215 は、図 1 のダウンミックス信号 115 に相当することができる。図 2 の実施形態において、モノラルダウンミックス信号 215 と空間パラメータ情報 105 は、パラメータステレオオーディオ信号 201 を表す。パラメータステレオオーディオ信号は、『L』で示される左チャンネルと、『R』で示される右チャンネルを備えることができる。ここで、ダイレクト/アンビエンス抽出器 120 は、ダイレクト/アンビエンス推定器 110 を用いることによって空間パラメータ情報 105 から導き出すことができる推定レベル情報 113 に基づいて、モノラルダウンミックス

40

50

信号 2 1 5 からダイレクト信号部分 1 2 5 1 またはアンビエント信号部分 1 2 5 2 を抽出するように構成される。

【 0 0 1 7 】

実際には、図 1 または図 2 の実施形態における空間パラメータ（空間パラメータ情報 1 0 5）は、それぞれ、MPEG サラウンド（MPS）またはパラメータステレオ（PS）サイド情報に特に関連する。これらの 2 つの技術は、最新技術の低ビットレートステレオまたはサラウンドオーディオ符号化方法である。図 2 を参照すると、PS は、空間パラメータを備えた 1 つのダウンミックスオーディオチャンネルを提供し、図 1 を参照すると、MPS は、空間パラメータを備えた 1 つ、2 つ、またはそれ以上のダウンミックスオーディオチャンネルを提供する。

10

【 0 0 1 8 】

具体的には、図 1 と図 2 の実施形態は、空間パラメータサイド情報 1 0 5 を、1 つ以上の入力チャンネルを有する信号（すなわち、ダウンミックス信号 1 1 5 ; 2 1 5）からのダイレクトおよび/またはアンビエンス抽出の分野において直ちに用いることができることを明らかに示している。

【 0 0 1 9 】

ダイレクトおよび/またはアンビエンスレベル（レベル情報 1 1 3）の推定は、レベル差および/または相関のようなチャンネル間関係またはチャンネル間差分に関する情報に基づいている。これらの値は、ステレオ信号または多重チャンネル信号から計算することができる。図 3 a は、 $Ch_1 \dots Ch_N$ のそれぞれのチャンネル間関係の計算に用いられる多重チャンネルオーディオ信号（ $Ch_1 \dots Ch_N$ ）のスペクトル分解 3 0 0 の図解図を示す。図 3 a から分かるように、多重チャンネルオーディオ信号（ $Ch_1 \dots Ch_N$ ）の検査チャンネル Ch_i または残りのチャンネルの線形結合 R のスペクトル分解は、それぞれ複数のサブバンド 3 0 1 を備え、時間/周波数グリッドの小さなボックスで示されるように、複数のサブバンド 3 0 1 の各サブバンド 3 0 3 は、サブバンド値 3 0 5 を有する水平軸（時間軸 3 1 0）に沿って伸展する。さらに、サブバンド 3 0 3 は、フィルタバンクの異なる周波数領域に対応する垂直軸（周波数軸 3 2 0）に沿って、連続的に配置される。図 3 a において、それぞれの時間/周波数タイル $X_i^{n,k}$ または $X_R^{n,k}$ は、破線によって示される。ここで、インデックス i は、チャンネル Ch_i を示し、 R は、残りのチャンネルの線形結合を示し、一方インデックス n と k は、特定のフィルタバンクタイムスロット 3 0 7 とフィルタバンクサブバンド 3 0 3 に相当する。検査チャンネル Ch_i のチャンネル間コヒーレンス（ ICC_i ）またはチャンネルレベル差（ CLD_i ）のようなチャンネル間関係 3 3 5 は、図 3 b に示されるように、例えば、時間/周波数軸 3 1 0、3 2 0 に関して同じ時間/周波数ポイント（ t_0 、 f_0 ）に位置する時間/周波数タイル $X_i^{n,k}$ と $X_R^{n,k}$ に基づいて、ステップ 3 3 0 において計算される。ここで、チャンネル間関係 ICC_i と CLD_i の計算は、以下の関係を用いることによって実行することができる。

20

30

【 0 0 2 0 】

$$ICC_i = \frac{\langle Ch_i R^* \rangle}{\sqrt{\langle Ch_i Ch_i^* \rangle \langle RR^* \rangle}}$$

40

$$\sigma_i = \frac{\langle Ch_i Ch_i^* \rangle}{\langle RR^* \rangle}$$

【 0 0 2 1 】

ここで、 Ch_i は検査チャンネルであり、 R は残りのチャンネルの線形結合であり、 $\langle \dots \rangle$ は時間平均を示す。残りのチャンネルの線形結合 R の例は、それらのエネルギー正規化された合計である。なお、チャンネルレベル差（ CLD_i ）は、通常はパラメータのデシベル値である。

50

【 0 0 2 2 】

上記の式に関して、チャンネルレベル差 (CLD_i) またはパラメータ α_i は、残りのチャンネルの線形結合 R のレベル P_R に正規化されたチャンネル Ch_i のレベル P_i に相当することができる。ここで、レベル P_i または P_R は、チャンネル Ch_i のチャンネル間レベル差パラメータ $ICLD_i$ と、残りのチャンネルのチャンネル間レベル差パラメータ $ICLD_j$ ($j \neq i$) の線形結合 $ICLD_R$ から導き出すことができる。

【 0 0 2 3 】

ここで、 $ICLD_i$ と $ICLD_j$ は、それぞれ基準チャンネル Ch_{ref} に関係づけることができる。更なる実施形態において、チャンネル間レベル差パラメータ $ICLD_i$ と $ICLD_j$ は、基準チャンネル Ch_{ref} である多重チャンネルオーディオ信号 ($Ch_1 \dots Ch_N$) の他のいかなるチャンネルにも関係づけることができる。これは、結局、チャンネルレベル差 (CLD_i) またはパラメータ α_i に対して、同じ結果に導く。

【 0 0 2 4 】

更なる実施形態によれば、図 3 b のチャンネル間関係 3 3 5 は、また、多重チャンネルオーディオ信号 ($Ch_1 \dots Ch_N$) の入力チャンネルの異なるまたは全てのペア Ch_i 、 Ch_j に関して演算することによって導き出すことができる。この場合、ペアに関して計算されたチャンネル間コヒーレンスパラメータ $ICC_{i,j}$ またはチャンネルレベル差 ($CLD_{i,j}$) またはパラメータ $\alpha_{i,j}$ (または $ICLD_{i,j}$) を取得することができ、インデックス (i, j) は、それぞれチャンネル Ch_i と Ch_j の特定のペアを示す。

【 0 0 2 5 】

図 4 は、推定レベル情報 1 1 3 のダウンミックスを含むダイレクト/アンビエンス抽出器 4 2 0 の実施形態 4 0 0 のブロック図を示す。図 4 の実施形態は、図 1 の実施形態と同じブロックを基本的に備える。それ故、同様の実施形態および/または機能を有する同一のブロックは、同じ符号で示される。しかしながら、図 1 のダイレクト/アンビエンス抽出器 1 2 0 に相当することができる図 4 のダイレクト/アンビエンス抽出器 4 2 0 は、多重チャンネルオーディオ信号のダイレクト部分またはアンビエント部分の推定レベル情報 1 1 3 をダウンミックスし、ダイレクト部分またはアンビエント部分のダウンミックスされたレベル情報を取得し、ダウンミックスされたレベル情報に基づいてダウンミックス信号 1 1 5 からダイレクト信号部分 1 2 5 1 またはアンビエント信号部分 1 2 5 2 を抽出するように構成される。図 4 に示されるように、空間パラメータ情報 1 0 5 は、例えば、図 1 の多重チャンネルオーディオ信号 1 0 1 ($Ch_1 \dots Ch_N$) から導き出すことができ、図 3 b において導入された $Ch_1 \dots Ch_N$ のチャンネル間関係 3 3 5 を備えることができる。図 4 の空間パラメータ情報 1 0 5 は、また、ダイレクト/アンビエンス抽出器 4 2 0 に供給されるダウンミックス情報 4 1 0 を備えることができる。実施形態において、ダウンミックス情報 4 1 0 は、オリジナルの多重チャンネルオーディオ信号 (例えば図 1 の多重チャンネルオーディオ信号 1 0 1) のダウンミックス信号 1 1 5 へのダウンミックスを特徴づけることができる。ダウンミックスは、例えば、時間ドメインまたはスペクトルドメインのようないかなる符号化ドメインにおいても動作するダウンミキサ (図示せず) を用いて実行することができる。

【 0 0 2 6 】

更なる実施形態によれば、ダイレクト/アンビエンス抽出器 4 2 0 は、また、ダイレクト部分の推定レベル情報をコヒーレント和と結合し、アンビエント部分の推定レベル情報を非コヒーレント和と結合することによって、多重チャンネルオーディオ信号 1 0 1 のダイレクト部分またはアンビエント部分の推定レベル情報 1 1 3 のダウンミックスを実行するように構成することができる。

【 0 0 2 7 】

推定レベル情報は、それぞれダイレクト部分またはアンビエント部分のエネルギーレベルまたはパワーレベルを表すことができることが指摘される。

【 0 0 2 8 】

特に、推定されたダイレクト/アンビエント部分のエネルギー (すなわちレベル情報 1

10

20

30

40

50

13) のダウンミックスは、チャンネル間の完全な非コヒーレンスまたは完全なコヒーレンスを仮定することによって実行することができる。非コヒーレント和またはコヒーレント和に基づくダウンミックスの場合に適用することができる2つの数式は、それぞれ以下の通りである。

【0029】

非コヒーレント信号に対して、ダウンミックスエネルギーあるいはダウンミックスレベル情報は、 $E_{DMX} = \sum_{i=1}^N g_i^2 E_{Ch_i}$ によって計算することができる。

【0030】

コヒーレント信号に対して、ダウンミックスエネルギーあるいはダウンミックスレベル情報は、 $E_{DMX} = \left(\sum_{i=1}^N g_i \sqrt{E_{Ch_i}} \right)^2$ によって計算することができる。

【0031】

ここで、 g は、ダウンミックス情報から取得することができるダウンミックスゲインであり、 $E(C h_i)$ は多重チャンネルオーディオ信号のチャンネル $C h_i$ のダイレクト/アンビエント部分のエネルギーを示す。非コヒーレントダウンミックスの代表的な例として、5.1チャンネルを2チャンネルにダウンミックスする場合に、左側のダウンミックスのエネルギーは、次式とすることができる。

$$E_{L_DMX} = E_{Left} + E_{Left_surround} + 0.5 * E_{Center}$$

【0032】

図5は、ゲインパラメータ g_D 、 g_A をダウンミックス信号115に適用することによるダイレクト/アンビエンス抽出器520の更なる実施形態を示す。図5のダイレクト/アンビエンス抽出器520は、図4のダイレクト/アンビエンス抽出器420に相当する。最初に、前に記載されているように、ダイレクト部分545₁またはアンビエント部分545₂の推定レベル情報は、ダイレクト/アンビエンス推定器から受信することができる。受信されたレベル情報545₁、545₂は、ステップ550において、それぞれ、結合され/ダウンミックスされ、ダイレクト部分555₁またはアンビエント部分555₂のダウンミックスされたレベル情報を取得することができる。次に、ステップ560において、ゲインパラメータ g_D 、565₁または g_A 、565₂は、それぞれ、ダイレクト部分またはアンビエント部分に対するダウンミックスされたレベル情報555₁、555₂から導き出すことができる。最後に、ダイレクト/アンビエンス抽出器520は、ダイレクト信号部分125₁またはアンビエント信号部分125₂が取得されるように、導き出されたゲインパラメータ565₁、565₂をダウンミックス信号115に適用するために用いることができる(ステップ570)。

【0033】

ここで、図1、4、5の実施形態において、ダウンミックス信号115は、それぞれ、ダイレクト/アンビエンス抽出器120；420；520の入力に存在する複数のダウンミックスチャンネル($C h_1 \dots C h_M$)から構成できることに注意すべきである。

【0034】

更なる実施形態において、ダイレクト/アンビエンス抽出器520は、ダイレクト部分またはアンビエント部分のダウンミックスされたレベル情報555₁、555₂から、ダイレクトからトータルへの(DTT)またはアンビエントからトータルへの(ATT)エネルギーレシオを決定し、決定されたDTTまたはATTエネルギーレシオに基づいて、抽出パラメータを、ゲインパラメータ565₁、565₂として用いるように構成される。

【0035】

さらに他の実施形態において、ダイレクト/アンビエンス抽出器520は、ダウンミッ

10

20

30

40

50

クス信号 1 1 5 に、第 1 の抽出パラメータ $sqr t (D T T)$ を乗算してダイレクト信号部分 1 2 5 1 を取得し、第 2 の抽出パラメータ $sqr t (A T T)$ を乗算してアンビエント信号部分 1 2 5 2 を取得するように構成される。ここで、ダウンミックス信号 1 1 5 は、図 2 の実施形態 (モノラルダウンミックスケース) に示されるように、モノラルダウンミックス信号 2 1 5 に相当する。

【 0 0 3 6 】

モノラルダウンミックスケースにおいて、アンビエンス抽出は、 $sqr t (A T T)$ と $sqr t (D T T)$ を適用することによってなすことができる。しかしながら、同じアプローチは、特に、各チャンネル $C h_i$ に対して $sqr t (A T T_i)$ と $sqr t (D T T_i)$ を適用することによって、多重チャンネルダウンミックス信号に対しても有効である。

10

【 0 0 3 7 】

更なる実施形態によれば、ダウンミックス信号 1 1 5 が複数のチャンネルを備える場合 (多重チャンネルダウンミックスケース)、ダイレクト/アンビエンス抽出器 5 2 0 は、第 1 の複数の抽出パラメータ、例えば $sqr t (D T T_i)$ をダウンミックス信号 1 1 5 に適用してダイレクト信号部分 1 2 5 1 を取得し、第 2 の複数の抽出パラメータ、例えば $sqr t (A T T_i)$ をダウンミックス信号 1 1 5 に適用してアンビエント信号部分 1 2 5 2 を取得するように構成することができる。ここで、第 1 と第 2 の複数の抽出パラメータは、対角行列を構成することができる。

【 0 0 3 8 】

一般に、ダイレクト/アンビエンス抽出器 1 2 0 ; 4 2 0 ; 5 2 0 は、また、二次の $M \times M$ 抽出マトリックスをダウンミックス信号 1 1 5 に適用することによってダイレクト信号部分 1 2 5 1 またはアンビエント信号部分 1 2 5 2 を抽出するように構成ことができ、二次の $M \times M$ 抽出マトリックスのサイズ (M) はダウンミックスチャンネル ($C h_1 \dots C h_M$) の数 (M) に相当する。

20

【 0 0 3 9 】

アンビエンス抽出のアプリケーションは、それ故、二次の $M \times M$ 抽出マトリックスを適用することによって記述することができ、ここで、 M はダウンミックスチャンネル ($C h_1 \dots C h_M$) の数である。これは、入力信号を操作してダイレクト/アンビエンス出力を得る全ての可能な方法を含むことができ、対角行列として構成される二次の $M \times M$ 抽出マトリックスの主要要素を表す $sqr t (A T T_i)$ と $sqr t (D T T_i)$ パラメータに基づく比較的簡単な方法、または完全なマトリックスとして構成される LMS クロスミックスアプローチを含む。後者は、以下に記載される。ここで、 $M \times M$ 抽出マトリックスを適用する上記のアプローチは、1 つを含むいかなるチャンネル数もカバーすることに注意すべきである。

30

【 0 0 4 0 】

更なる実施形態によれば、抽出マトリックスは、より少ない出力チャンネル数を有することができるので、必ずしもマトリックスサイズ $M \times M$ の二次マトリックスである必要はない。それ故、抽出マトリックスは、低減された行数を有することができる。この例は、 M の代わりに単一のダイレクト信号を抽出する。

【 0 0 4 1 】

また、抽出マトリックスの M 列を有することに対応する入力として、全ての M 個のダウンミックスチャンネルを常にとる必要はない。これは、特に、入力として全てのチャンネルを有することを必要としないアプリケーションに関連することができる。

40

【 0 0 4 2 】

図 6 は、チャンネルクロスミックスを備える LMS (最小平均二乗) 解法に基づくダイレクト/アンビエンス抽出器 6 2 0 の更なる実施形態 6 0 0 のブロック図を示す。図 6 のダイレクト/アンビエンス抽出器 6 2 0 は、図 1 のダイレクト/アンビエンス抽出器 1 2 0 に相当することができる。図 6 の実施形態において、図 1 の実施形態と同様の実施態様および/または機能を有する同一のブロックは、それ故、同じ符号で示される。しかしながら、図 1 のダウンミックス信号 1 1 5 に相当することができる図 6 のダウンミックス信

50

号 6 1 5 は、複数のダウンミックスチャンネル $Ch_1 \dots Ch_M$ 6 1 7 を備えることができ、ダウンミックスチャンネル数 (M) は、多重チャンネルオーディオ信号 1 0 1 のチャンネル $Ch_1 \dots Ch_N$ (N) のそれより小さい (すなわち $M < N$)。具体的には、ダイレクト/アンビエンス抽出器 6 2 0 は、チャンネルクロスミックスを備える最小平均二乗 (LMS) 解法によってダイレクト信号部分 1 2 5 1 またはアンビエント信号部分 1 2 5 2 を抽出するように構成され、LMS 解法は等しいアンビエンスレベルを必要としない。等しいアンビエンスレベルを必要とせず、いかなるチャンネル数にも拡張可能である LMS 解法は、以下において提供される。ちょうど言及された LMS 解法は、必須ではないが、上記に対してより正確な変形例を表す。

【 0 0 4 3 】

10

ダイレクト/アンビエンス抽出のためのクロスミックス重みに対して、LMS 解法において使用されるシンボルは、以下の通りである。

【 0 0 4 4 】

Ch_i : チャンネル i

a_i : チャンネル i におけるダイレクトサウンドのゲイン

D および D^{\wedge} : サウンドのダイレクト部分およびその推定

A_i および A_i^{\wedge} : チャンネル i のアンビエント部分およびその推定

$P_x = E [X X^*]$: X の推定エネルギー

$E [\]$: 期待値

E_x : X の推定誤差

20

w_{Di} : チャンネル i に対するダイレクト部分への LMS クロスミックス重み

$w_{Ai,n}$: チャンネル n に対するチャンネル i のアンビエンスへの LMS クロスミックス重み

【 0 0 4 5 】

この文脈において、LMS 解法の導出は、多重チャンネルオーディオ信号のそれぞれのチャンネルのスペクトル表現に基づくことができ、それはすべてが周波数バンドにおいて機能することを意味することに注意すべきである。

【 0 0 4 6 】

信号モデルは、次式によって与えられる。

$$Ch_i = a_i D + A_i$$

30

【 0 0 4 7 】

導出は、最初に、a) ダイレクト部分を、次に、b) アンビエント部分を取扱う。最後に、重みに対する解法が導き出され、重みの正規化の方法が記述される。

【 0 0 4 8 】

a) ダイレクト部分

【 0 0 4 9 】

重みのダイレクト部分の推定は、次式の通りである。

$$\hat{D} = \sum_{i=1}^N w_{Di} Ch_i = \sum_{i=1}^N w_{Di} (a_i D + A_i)$$

40

【 0 0 5 0 】

推定誤差は次式のように読み取れる。

$$E_{\hat{D}} = D - \hat{D} = D - \sum_{i=1}^N w_{Di} (a_i D + A_i)$$

【 0 0 5 1 】

LMS 解法を持つため、入力信号に直交な $E_{\hat{p}}$ を必要とする。

$$E[E_{\hat{p}}Ch_k]=0, \text{ for all } k$$

$$\begin{aligned} & E\left[\left(D - \sum_{i=1}^N w_{\hat{D}i}(a_i D + A_i)\right)(a_k D + A_k)^*\right] \\ &= \left(a_k - \sum_{i=1}^N w_{\hat{D}i} a_i a_k\right) P_D - w_{\hat{D}k} P_{Ak} = 0 \\ &\Leftrightarrow \sum_{i=1}^N w_{\hat{D}i} a_i a_k P_D + w_{\hat{D}k} P_{Ak} = a_k P_D \end{aligned} \quad 10$$

【 0 0 5 2 】

マトリックス形式において、上記の関係は次式のように読み取れる。

$$A\bar{w} \equiv \bar{P}$$

$$\begin{bmatrix} (a_1 a_1 P_D + P_{A1}) & a_1 a_2 P_D & \cdots & a_1 a_N P_D \\ a_1 a_2 P_D & (a_2 a_2 P_D + P_{A2}) & & \vdots \\ \vdots & & \ddots & \\ a_1 a_N P_D & \cdots & & (a_N a_N P_D + P_{AN}) \end{bmatrix} \begin{bmatrix} w_{\hat{D}1} \\ w_{\hat{D}2} \\ \vdots \\ w_{\hat{D}N} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} P_D \quad 20$$

【 0 0 5 3 】

b) アンビエント部分

【 0 0 5 4 】

同じ信号モデルから始めて、次式から重みを推定する。

$$\hat{A}_i = \sum_{n=1}^N w_{\hat{A}i,n} Ch_n = \sum_{n=1}^N w_{\hat{A}i,n} (a_i D + A_i) \quad 30$$

【 0 0 5 5 】

推定誤差は、次の通りである。

$$E_{\hat{A}i} = A_i - \hat{A}_i = A_i - \sum_{n=1}^N w_{\hat{A}i,n} (a_i D + A_i)$$

【 0 0 5 6 】

そして、直交性は、次の通りである。

$$E[E_{\hat{A}_i} C h_k] = 0, \text{ for all } k$$

$$E \left[\left(A_i - \sum_{n=1}^N w_{\hat{A}_i, n} (a_n D + A_n) \right) (a_k D + A_k)^* \right]$$

$$= \begin{cases} -\sum_{n=1}^N w_{\hat{A}_i, n} a_n a_k P_D - w_{\hat{A}_i, k} P_{A_k} = 0 & , \text{ if } i \neq k \\ -\sum_{n=1}^N w_{\hat{A}_i, n} a_n a_k P_D - w_{\hat{A}_i, k} P_{A_k} + P_{A_k} = 0 & , \text{ if } i = k \end{cases} \quad 10$$

$$\Leftrightarrow \begin{cases} \sum_{n=1}^N w_{\hat{A}_i, n} a_n a_k P_D + w_{\hat{A}_i, k} P_{A_k} = 0 & , \text{ if } i \neq k \\ \sum_{n=1}^N w_{\hat{A}_i, n} a_n a_k P_D + w_{\hat{A}_i, k} P_{A_k} = P_{A_k} & , \text{ if } i = k \end{cases}$$

【 0 0 5 7 】

マトリックス形式において、上記の関係は次式のように読み取れる。

$$AW = P$$

$$\begin{bmatrix} (a_1 a_1 P_D + P_{A1}) & a_1 a_2 P_D & \cdots & a_1 a_N P_D \\ a_1 a_2 P_D & (a_2 a_2 P_D + P_{A2}) & & \vdots \\ \vdots & & \ddots & \\ a_1 a_N P_D & \cdots & & (a_N a_N P_D + P_{AN}) \end{bmatrix} \begin{bmatrix} w_{\hat{A}1,1} & w_{\hat{A}2,1} & \cdots & w_{\hat{A}N,1} \\ w_{\hat{A}1,2} & w_{\hat{A}2,2} & & \vdots \\ \vdots & & \ddots & \\ w_{\hat{A}1,N} & \cdots & & w_{\hat{A}N,N} \end{bmatrix} = \begin{bmatrix} P_{A1} & 0 & \cdots & 0 \\ 0 & P_{A2} & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & P_{AN} \end{bmatrix} \quad 20$$

【 0 0 5 8 】

重みに対する解法

【 0 0 5 9 】

重みは、ダイレクト部分とアンビエント部分の両方の計算において同一のマトリックス A を反転することによって解くことができる。ステレオ信号の場合、解法は次式の通りである。

$$w_{D1} = \frac{a_1 P_D P_{A2}}{a_2 a_2 P_D P_{A1} + a_1 a_1 P_D P_{A2} + P_{A1} P_{A2}} = \frac{a_1 P_D P_{A2}}{\text{div}} \quad 30$$

$$w_{D2} = \frac{a_2 P_D P_{A1}}{\text{div}}$$

$$w_{\hat{A}1,1} = \frac{a_2 a_2 P_D P_{A1} + P_{A1} P_{A2}}{\text{div}}$$

$$w_{\hat{A}1,2} = \frac{a_1 a_2 P_D P_{A1}}{\text{div}}$$

$$w_{\hat{A}2,1} = \frac{a_1 a_2 P_D P_{A2}}{\text{div}} \quad 40$$

$$w_{\hat{A}2,2} = \frac{a_1 a_1 P_D P_{A2} + P_{A1} P_{A2}}{\text{div}}$$

【 0 0 6 0 】

ここで、d i v は、除数 $a_2 a_2 P_D P_{A1} + a_1 a_1 P_D P_{A2} + P_{A1} P_{A2}$ である。

【 0 0 6 1 】

重みの正規化

【 0 0 6 2 】

重みは、LMS 解法に対するものであるが、エネルギーレベルは保存されなければならないので、重みは正規化される。これは、また、上記の式において不必要な項 d i v による除算をなす。正規化は、出力ダイレクトおよびアンビエントチャンネルのエネルギーが

P_D および P_{Ai} （ここで、 i はチャンネルインデックス）であることを確保することによって生ずる。

【0063】

これは、チャンネル間コヒーレンス、混合係数およびチャンネルエネルギーを知っていると直接的に仮定する。簡単のため、2チャンネルの場合であって、特に1つの重みのペア $w_{\hat{A}1,1}$ と $w_{\hat{A}1,2}$ であり、それが第1と第2の入力チャンネルから第1のアンビエンスチャンネルを生成するゲインであった場合にフォーカスする。ステップは、以下の通りである。

【0064】

ステップ1：出力信号エネルギーを計算する（コヒーレント部分は振幅に関して合計し、非コヒーレント部分はエネルギーに関して合計する） 10

$$P_{\hat{A}1} = \left(w_{\hat{A}1,1} \sqrt{|ICC| \cdot P_1} + \text{sign}(ICC) w_{\hat{A}1,2} \sqrt{|ICC| \cdot P_2} \right)^2 + (1 - |ICC|) P_1 w_{\hat{A}1,1}^2 + (1 - |ICC|) P_2 w_{\hat{A}1,2}^2$$

【0065】

ステップ2：正規化ゲイン係数を計算し、

$$g = \sqrt{\frac{P_{\hat{A}1}}{P_{\hat{A}1}}}$$

【0066】

その結果をクロスミックス重み係数 $w_{\hat{A}1,1}$ と $w_{\hat{A}1,2}$ に適用する。ステップ1において、入力チャンネルが否定的にコヒーレントである場合も考慮するために、 ICC に対する絶対値と符号オペレータが含まれる。残りの重み係数も同じやり方で正規化される。 20

【0067】

特に、上記を参照して、ダイレクト/アンビエンス抽出器620は、LMS解法がステレオチャンネルのダウンミックス信号に限定されないように、安定な多重チャンネル信号モデルとみなすことによってLMS解法を導出するように構成することができる。

【0068】

図7aは、ステレオアンビエンス推定式に基づくダイレクト/アンビエンス推定器710の実施形態700のブロック図を示す。図7のダイレクト/アンビエンス推定器710は、図1のダイレクト/アンビエンス推定器110に相当することができる。特に、図7のダイレクト/アンビエンス推定器710は、多重チャンネルオーディオ信号101の各チャンネル(Ch_i)に対する空間パラメータ情報105を用いて、ステレオアンビエンス推定式を適用するように構成される。ここで、ステレオアンビエンス推定式は、次式の関数依存性として表され、 30

$$DTT_i = f_{DTT}[\sigma_i(Ch_i, R), ICC_i(Ch_i, R)],$$

$$ATT_i = 1 - DTT_i$$

【0069】

チャンネル Ch_i のチャンネルレベル差(CLD_i)またはパラメータ σ_i とチャンネル間コヒーレンスパラメータ(ICC_i)の依存性を明確に示すことができる。図7において図示されるように、空間パラメータ情報105は、ダイレクト/アンビエンス推定器710に供給され、各チャンネル Ch_i に対するチャンネル間関係パラメータ ICC_i と σ_i を備えることができる。このステレオアンビエンス推定式を、ダイレクト/アンビエンス推定器710を用いて適用した後、ダイレクトからトータルへの(DTT_i)またはアンビエントからトータルへの(ATT_i)エネルギーレシオは、それぞれ、その出力715において取得される。それぞれの DTT または ATT エネルギーレシオを推定するために用いられる上記のステレオアンビエンス推定式は、等しいアンビエンスの状態に基づいていないことに注意する必要がある。 40

【0070】

特に、ダイレクト／アンビエンスレシオ推定は、チャンネル内のダイレクトエネルギーのそのチャンネルのトータルエネルギーと比較したレシオ（D T T）を次式によって数式化することができることで実行することができる。

$$Ratio = \frac{1}{2} \left[\left(1 - \frac{1}{\sigma} \right) + \sqrt{\left(\frac{1}{\sigma} - 1 \right)^2 + 4 \frac{ICC^2}{\sigma}} \right]$$

【 0 0 7 1 】

ここで、 $\sigma = \frac{\langle ChCh^* \rangle}{\langle RR^* \rangle}$ 、 $ICC = \frac{\langle ChR^* \rangle}{\sqrt{\langle ChCh^* \rangle \langle RR^* \rangle}}$ であり、Chは検査チャンネルであり、

10

Rは残りのチャンネルの線形結合である。<>は時間平均である。アンビエンスレベルが、チャンネル内および残りのチャンネルの線形結合において等しく、そのコヒーレンスがゼロとみなされるときに、この式に従う。

【 0 0 7 2 】

図7bは、例示的なD T T（ダイレクトからトータルへの）エネルギーレシオ760のグラフ750を、チャンネル間コヒーレンスパラメータICC 770の関数として示す。図7bの実施形態において、チャンネルCh_iのレベルP（Ch_i）と残りのチャンネルの線形結合RのレベルP（R）が等しくなるように、チャンネルレベル差（CLD）またはパラメータは、例示的に1セットされている（= 1）。この場合、D T Tエネルギーレシオ760は、D T T ~ ICCによってマークされた直線775によって示されるように、ICCパラメータに線形に比例する。図7bにおいて、完全に非コヒーレントなチャンネル間関係に相当することができるICC = 0の場合には、完全にアンビエントな状況（ケース『R₁』）に相当することができるD T Tエネルギーレシオ760は0になることがわかる。しかしながら、完全にコヒーレントなチャンネル間関係に相当することができるICC = 1の場合には、完全にダイレクトな状況（ケース『R₂』）に相当することができるD T Tエネルギーレシオ760は1とすることができる。それ故、R₁のケースにおいて、基本的にチャンネル内にそのチャンネルのトータルエネルギーに関してダイレクトエネルギーがなく、一方R₂のケースにおいて、アンビエントエネルギーがない。

20

【 0 0 7 3 】

図8は、本発明の更なる実施形態に係るエンコーダ／デコーダシステム800のブロック図を示す。エンコーダ／デコーダシステム800のデコーダ側において、図1の装置100に相当することができるデコーダ820の実施形態が示される。図1と図8の実施形態の類似性のため、これらの実施形態において同様の実施態様および／または機能を有する同一のブロックは、同じ符号で示される。図8の実施形態に示されるように、ダイレクト／アンビエンス抽出器120は、複数のダウンミックスチャンネルCh₁... Ch_Mを有するダウンミックス信号115について動作することができる。図8のダイレクト／アンビエンス推定器110は、多重チャンネルオーディオ信号101のダイレクト部分またはアンビエント部分のレベル情報113が、空間パラメータ情報105に加えて、受信される少なくとも2つのダウンミックスチャンネル825に基づいて推定されるように、ダウンミックス信号815の少なくとも2つのダウンミックスチャンネル825を受信する（オプション）ように更に構成することができる。最後に、ダイレクト信号部分125 1またはアンビエント信号部分125 2は、ダイレクト／アンビエンス抽出器120による抽出の後に取得される。

30

40

【 0 0 7 4 】

エンコーダ／デコーダシステム800のエンコーダ側において、多重チャンネルオーディオ信号（Ch₁... Ch_N）を多数のダウンミックスチャンネルCh₁... Ch_Mを有するダウンミックス信号115にダウンミックスし、チャンネル数がNからMに低減されるダウンミキサー815を備えることができるエンコーダ810の実施形態が示される。ダウンミキサー815は、また、多重チャンネルオーディオ信号101からチャンネル間関係を計算することによって、空間パラメータ情報105を出力するように構成することができる

50

。図8のエンコーダ/デコーダシステム800において、ダウンミックス信号115と空間パラメータ情報105は、エンコーダ810からデコーダ820に送信することができる。ここで、エンコーダ810は、エンコーダ側からデコーダ側への送信に対して、ダウンミックス信号115と空間パラメータ情報105に基づいて符号化信号を導き出すことができる。さらに、空間パラメータ情報105は、多重チャンネルオーディオ信号101のチャンネル情報に基づいている。

【0075】

一方では、チャンネル間関係パラメータ i (Ch_i , R)と ICC_i (Ch_i , R)は、エンコーダ810において、チャンネル Ch_i と残りのチャンネルの線形結合 R の間で計算し、符号化信号の中で送信することができる。デコーダ820は、次に、符号化信号を受信し、送信されたチャンネル間関係パラメータ i (Ch_i , R)と ICC_i (Ch_i , R)について動作することができる。

10

【0076】

他方では、エンコーダ810は、また、送信される異なるチャンネルのペア (Ch_i , Ch_j) の間のチャンネル間コヒーレンスパラメータ $ICC_{i,j}$ を計算するように構成することができる。この場合、デコーダ810は、以前に記載されている対応する実施形態を実現することができるように、送信されたペアに関して計算された $ICC_{i,j}$ (Ch_i , Ch_j) パラメータから、チャンネル Ch_i と残りのチャンネルの線形結合 R の間のパラメータ ICC_i (Ch_i , R)を導き出すことが可能でなければならない。この文脈において、デコーダ820は、ダウンミックス信号115の知識だけからパラメータ ICC_i (Ch_i , R)を復元することができない点に注意すべきである。

20

【0077】

実施形態において、送信された空間パラメータは、ペアに関するチャンネル比較についてだけではない。

【0078】

たとえば、最も代表的なMPSのケースは、2つのダウンミックスチャンネルがあるということである。MPS復号化における空間パラメータの第1のセットは、2つのチャンネルを3つ：センター、レフトおよびライトにする。このマッピングを導くパラメータのセットは、センター予測係数(CPC)と、この2から3への構成に特有のICCパラメータと呼ばれる。

30

【0079】

空間パラメータの第2のセットは、各々を次の2つに分ける。すなわち、サイドチャンネルを対応する前後のチャンネルに、センターチャンネルをセンターとLfeチャンネルに分ける。このマッピングは、前に導入されたICCとCLDパラメータに関する。

【0080】

すべての種類のダウンミックス構成とすべての種類の空間パラメータに対する計算ルールを作るとは、実際的でない。しかしながら、仮想的に、ダウンミックスステップに従うことは、実際的である。我々は、どのようにして2つのチャンネルが3つになり、3つが6つになるかを知っているのだから、結局は、どのように2つの入力チャンネルが6つの出力に割り振られるかの入出力関係を見つける。出力は、ダウンミックスチャンネルの線形結合に加えてそれらの非相関化バージョンの線形結合のみである。出力信号を実際に復号化し、それを測る必要はないが、我々は、この「復号化マトリックス」を知っているのだから、パラメータドメインにおいて、いかなるチャンネルまたはチャンネルの組合せの間のICCとCLDパラメータも演算処理上能率的に計算することができる。

40

【0081】

ダウンミックス信号および多重チャンネル信号の構成に関係なく、復号化信号の各出力は、ダウンミックス信号の線形結合に加えてそれらの各々の非相関化バージョンの線形結合である。

$$Ch_out_i = \sum_{k=1}^{dmx_channels} (a_{k,i} Ch_dmx_k + b_{k,i} D[Ch_dmx_k])$$

【 0 0 8 2 】

ここで、オペレータ D [] は、非相関化器、すなわち入力信号の非コヒーレントな複製を作るプロセスに相当する。係数 a と b は、それらがパラメータサイド情報から直接的に導き出せるので既知である。これは、定義上、パラメータ情報は、どのようにダウンミックス信号から多重チャンネル出力を生成するかとのデコーダに対するガイドであるからである。上記の数式は、全ての非相関化部分をエネルギー性／コヒーレンスの比較に対して組合せることができることから、次のように単純化することができる。

10

$$Ch_out_i = \sum_{k=1}^{dmx_channels} (a_{k,i} Ch_dmx_k) + D_i$$

係数 b も最初の数式において既知であるので、D のエネルギーは既知である。

【 0 0 8 3 】

これにより、出力チャンネルの間のまたは出力チャンネルの異なる線形結合間のいかなる種類のコヒーレンスとエネルギーの比較をもすることができる点に注意すべきである。2つのダウンミックスチャンネルと、例えば、チャンネル番号3と5が互いに比較される出力チャンネルのセットの簡単な例の場合に、シグマは以下のように計算される。

20

$$\sigma_{3,5} = \frac{E[Ch_out_3^2]}{E[Ch_out_5^2]}$$

【 0 0 8 4 】

ここで、E [] は、期待値（実際には：平均）オペレータである。両項は、以下のように数式化することができる。

$$\begin{aligned} E[Ch_out_i^2] &= E\left[\left(\sum_{k=1}^2 (a_{k,i} Ch_dmx_k) + D_i\right)^2\right] \\ &= E[D_i^2] + \sum_{k=1}^2 (a_{k,i}^2 E[Ch_dmx_k^2]) + 2a_{1,i}a_{2,i} (E[Ch_dmx_1 Ch_dmx_2]) \end{aligned}$$

30

【 0 0 8 5 】

上記の全てのパラメータは、既知またはダウンミックス信号から測定可能である。交差項の E [C h _ d m x * D] は定義上ゼロであり、それ故それらは数式の下列にない。同様に、コヒーレンス式は、次の通りである。

$$ICC_{3,5} = \frac{E[Ch_out_3 Ch_out_5]}{\sqrt{E[Ch_out_3^2] E[Ch_out_5^2]}}$$

40

【 0 0 8 6 】

また、上記の数式の全ての部分は入力の線形結合に加えて非相関化された信号であるので、解法は直接的に利用可能である。

【 0 0 8 7 】

上記の例は2つの出力チャンネル比較を備えていたが、同様に、後述される例示的プロセスのような出力チャンネルの線形結合間の比較をすることができる。

【 0 0 8 8 】

前の実施形態の概要において、提案された技術／コンセプトは、以下のステップを備えることができる。

50

1. ダウンミックスチャンネル数より多いことができる「オリジナル」のチャンネルセットのチャンネル間関係（コヒーレンス、レベル）を取り出す。

2. この「オリジナル」のチャンネルセットのアンビエンスエネルギーとダイレクトエネルギーを推定する。

3. この「オリジナル」のチャンネルセットのダイレクトエネルギーとアンビエンスエネルギーを、少ないチャンネル数にダウンミックスする。

4. ダウンミックスされたエネルギーを用い、ゲイン係数またはゲインマトリックスを適用することによって、提供されたダウンミックスチャンネルにおけるダイレクト信号とアンビエンス信号を抽出する。

【0089】

空間パラメータサイド情報の使用は、図2の実施形態によって最もよく説明され、要約されている。図2の実施形態において、単一のオーディオチャンネルとそれが表すステレオサウンドのチャンネル間差分（コヒーレンス、レベル）に関する空間サイド情報を含む、パラメータステレオストリームを持つ。ここで、我々は、チャンネル間差分を知っているので、上記のステレオアンビエンス推定式をそれらに適用し、オリジナルのステレオチャンネルのダイレクトエネルギーとアンビエンスエネルギーを得ることができる。次に、ダイレクトエネルギーを（コヒーレント和と）、アンビエンスエネルギーを（非コヒーレント和と）一緒に加算することによって、チャンネルエネルギーを「ダウンミックス」することができ、単一のダウンミックスチャンネルのダイレクトからトータルへのエネルギーレシオとアンビエントからトータルへのエネルギーレシオを抽出することができる。

【0090】

図2の実施形態を参照して、空間パラメータ情報は、それぞれパラメータステレオオーディオ信号の左チャンネル（L）と右チャンネル（R）に対応するチャンネル間コヒーレンスパラメータ（ ICC_L 、 ICC_R ）とチャンネルレベル差パラメータ（ CLD_L 、 CLD_R ）を基本的に備える。ここで、チャンネル間コヒーレンスパラメータ ICC_L と ICC_R は、等しい（ $ICC_L = ICC_R$ ）が、チャンネルレベル差パラメータ CLD_L と CLD_R は、 $CLD_L = -CLD_R$ の関係にある点に注意すべきである。同様に、チャンネルレベル差パラメータ CLD_L と CLD_R は、通常はそれぞれパラメータ L と R のデシベル値であるので、左チャンネル（L）と右チャンネル（R）に対するパラメータ L と R は、 $L = 1/R$ の関係にある。これらのチャンネル間差分パラメータは、ステレオアンビエンス推定式に基づいて、両方のチャンネル（L、R）に対して、それぞれのダイレクトからトータルへのエネルギーレシオ（ DTT_L 、 DTT_R ）とアンビエントからトータルへのエネルギーレシオ（ ATT_L 、 ATT_R ）を計算するために直ちに用いることができる。ステレオアンビエンス推定式において、左チャンネル（L）のダイレクトからトータルへのエネルギーレシオとアンビエントからトータルへのエネルギーレシオ（ DTT_L 、 ATT_L ）は、左チャンネルLに対するチャンネル間差分パラメータ（ CLD_L 、 ICC_L ）に依存し、一方、右チャンネル（R）のダイレクトからトータルへのエネルギーレシオとアンビエントからトータルへのエネルギーレシオ（ DTT_R 、 ATT_R ）は、右チャンネルRに対するチャンネル間差分パラメータ（ CLD_R 、 ICC_R ）に依存する。さらに、パラメータステレオオーディオ信号の両方のチャンネルL、Rに対するエネルギー（ E_L 、 E_R ）は、それぞれ、左チャンネル（L）と右チャンネル（R）に対するチャンネルレベル差パラメータ（ CLD_L 、 CLD_R ）に基づいて導き出すことができる。ここで、左チャンネルLに対するエネルギー（ E_L ）は、左チャンネルLに対するチャンネルレベル差パラメータ（ CLD_L ）をモノラルダウンミックス信号に適用することによって取得することができ、一方、右チャンネルRに対するエネルギー（ E_R ）は、右チャンネルRに対するチャンネルレベル差パラメータ（ CLD_R ）をモノラルダウンミックス信号に適用することによって取得することができる。次に、両チャンネル（L、R）に対するエネルギー（ E_L 、 E_R ）を対応する DTT_L 、 DTT_R 、および ATT_L 、 ATT_R ベースのパラメータと乗算することによって、両チャンネル（L、R）に対するダイレクトエネルギー（ E_{DL} 、 E_{DR} ）とアンビエンスエネルギー（ E_{AL} 、 E_{AR} ）が取得される。次に、両チャンネル（L、R）に対するダ

10

20

30

40

50

イレクトエネルギー (E_{DL} 、 E_{DR}) は、コヒーレントダウンミックスルールを用いることによって結合/加算し、モノラルダウンミックス信号のダイレクト部分に対するダウンミックスされたエネルギー ($E_{D, mono}$) を取得することができ、一方、両チャンネル (L、R) に対するアンビエンスエネルギー (E_{AL} 、 E_{AR}) は、非コヒーレントダウンミックスルールを用いることによって結合/加算し、モノラルダウンミックス信号のダイレクト部分に対するダウンミックスされたエネルギー ($E_{A, mono}$) を取得することができる。次に、ダイレクト信号部分およびアンビエント信号部分に対するダウンミックスされたエネルギー ($E_{D, mono}$ 、 $E_{A, mono}$) を、モノラルダウンミックス信号のトータルエネルギー (E_{mono}) に関連づけることによって、モノラルダウンミックス信号のダイレクトからトータルへのエネルギーレシオ (DTT_{mono}) とアンビエントからトータルへのエネルギーレシオ (ATT_{mono}) が取得される。最後に、これらの DTT_{mono} エネルギーレシオと ATT_{mono} エネルギーレシオに基づいて、モノラルダウンミックス信号からダイレクト信号部分またはアンビエント信号部分を基本的に抽出することができる。

10

【0091】

オーディオの再生において、サウンドをヘッドホン上に再生するニーズがしばしば生ずる。ヘッドホンリスニングは、スピーカリスニングおよびあらゆる自然音響環境に対して大幅に異なる特別な特徴を有する。オーディオは、左右の耳に対して直接にセットされる。生成されるオーディオコンテンツは、通常はスピーカ再生に対して生成される。それ故、オーディオ信号は、我々の聴覚システムが空間音響知覚において用いる属性と頭出しを含まない。それは、バイノーラル処理がシステムに導入されない限り、当てはまる。

20

【0092】

バイノーラル処理は、基本的に、入力サウンドに取り入れ、それを修正し、それが (我々の聴覚システムが空間サウンドを処理する方法に関して) 知覚的に正しい両耳間およびモノラルの属性のようなものだけを含むようにするプロセスであるということができる。バイノーラル処理は、直接的なタスクではなく、技術水準による存在する解法は、多くの副次的な最適性を有する。

【0093】

多重チャンネルオーディオ信号をヘッドホンに対するバイノーラル対応物に変換するように設計されたメディアプレーヤおよび処理デバイスのような、音楽と映画再生に対するバイノーラル処理が既に含まれた多数のアプリケーションがある。代表的なアプローチは、頭部関連伝達関数 (HRTF) を用い、仮想スピーカを作成し、信号に対して空間効果を加える。これは、理論的に、特定の空間においてスピーカでリスニングすることと等価とすることができた。

30

【0094】

しかしながら、実際は、このアプローチが一貫してリスナーを満足させていないことを繰り返し示している。この直接的な方法を備えた良好な空間化が、音色または音質における好ましくない変化、空間効果の悩ましい知覚およびダイナミクスの損失を有するようなオーディオ品質を損失する代償がついてくるという妥協があるように思われる。更なる問題は、不正確なローカライゼーション (例えば頭部内のローカライゼーション、前後の混乱)、音源の空間距離の欠如、および両耳間の mismatch、すなわち間違った両耳の頭出しによる耳の近くの聴覚を含む。

40

【0095】

異なるリスナーは、非常に異なって問題を判断するかもしれない。感度は、また、音楽 (音色に関する厳格な品質判定基準)、映画 (より厳格でない) およびゲーム (さらに厳格でないがローカライゼーションが重要である) のような入力素材に依存して変化する。また、コンテンツに依存して通常は異なる設計目標が存在する。

【0096】

それ故、次の記述は、平均的な全体の認知品質を最大化するために上記の問題をできるだけうまく克服するアプローチを取扱う。

【0097】

50

図9 aは、本発明の更なる実施形態に係るバイノーラル・ダイレクトサウンド・レンダリングデバイス910の概観900のブロック図を示す。図9 aに示されるように、バイノーラル・ダイレクトサウンド・レンダリングデバイス910は、図1の実施形態のダイレクト/アンビエンス抽出器120の出力に存在することができるダイレクト信号部分125 1を処理し、第1のバイノーラル出力信号915を取得するように構成される。第1のバイノーラル出力信号915は、Lで示される左チャンネルとRで示される右チャンネルを備えることができる。

【0098】

ここで、バイノーラル・ダイレクトサウンド・レンダリングデバイス910は、ダイレクト信号部分125 1を、頭部関連伝達関数(HRTF)を通して供給し、変換されたダイレクト信号部分を取得するように構成することができる。バイノーラル・ダイレクトサウンド・レンダリングデバイス910は、更に、変換されたダイレクト信号部分に空間効果を適用し、最終的に第1のバイノーラル出力信号915を取得するように構成することができる。

10

【0099】

図9 bは、図9 aのバイノーラル・ダイレクトサウンド・レンダリングデバイス910の詳細905のブロック図を示す。バイノーラル・ダイレクトサウンド・レンダリングデバイス910は、ブロック912で示される「HRTF変換器」と、ブロック914で示される空間効果処理デバイス(並列リバーブあるいは初期反射のシミュレーション)を備えることができる。図9 bに示されるように、HRTF変換器912と空間効果処理デバイス914は、頭部関連伝達関数(HRTF)と空間効果を並列に適用することによってダイレクト信号部分125 1について動作することができ、第1のバイノーラル出力信号915が取得される。

20

【0100】

具体的には、図9 bを参照して、この空間効果処理は、また、非コヒーレントな反響したダイレクト信号919を提供することができ、その信号919は、次のクロスミックスフィルタ920によって処理され、信号を拡散音場の両耳間コヒーレンスに適合させることができる。ここで、フィルタ920とHRTF変換器912の結合出力は、第1のバイノーラル出力信号915を構成する。更なる実施形態によれば、ダイレクトサウンドに関する空間効果処理は、初期反射のパラメータ表現とすることもできる。

30

【0101】

それ故、実施形態において、空間効果は、好ましくはHRTFに並列に適用することができ、直列(すなわち、信号をHRTFを通して供給した後に空間効果を適用することによって)には適用されない。具体的には、音源からダイレクトに伝搬するサウンドのみが通過するかまたは対応するHRTFによって変換される。間接的な/反響したサウンドは、至る所で耳に入るように、すなわち統計的方法で(HRTFの代わりにコヒーレンス制御を使用することによって)、近似することができる。直列の実施態様があってもよいが、並列の方法が好ましい。

【0102】

図10 aは、本発明の更なる実施形態に係るバイノーラル・アンビエントサウンド・レンダリングデバイス1010の概観1000のブロック図を示す。図10 aに示されるように、バイノーラル・アンビエントサウンド・レンダリングデバイス1010は、例えば、図1のダイレクト/アンビエンス抽出器120から出力されるアンビエント信号部分125 2を処理し、第2のバイノーラル出力信号1015を取得するように構成することができる。第2のバイノーラル出力信号1015は、また、左チャンネル(L)と右チャンネル(R)を備えることができる。

40

【0103】

図10 bは、図10 aのバイノーラル・アンビエントサウンド・レンダリングデバイス1010の詳細1005のブロック図を示す。図10 bにおいて、バイノーラル・アンビエントサウンド・レンダリングデバイス1010は、非コヒーレントな反響したアンビエン

50

ト信号1013が取得されるように、アンビエント信号部分125 2に、「空間効果処理」で表示されたブロック1012で示される空間効果を適用するように構成することができる。バイノーラル・アンビエントサウンド・レンダリングデバイス1010は、現実の拡散音場の両耳間コヒーレンスに適合する第2のバイノーラル出力信号1015が提供されるように、ブロック1014で示されるクロスミックスフィルタのようなフィルタを適用することによって、非コヒーレントな反響したアンビエント信号1013を処理するように更に構成することができる。「空間効果処理」で示されるブロック1012は、また、現実の拡散音場の両耳間コヒーレンスを直接生成するように構成することができる。この場合、ブロック1014は使用されない。

【0104】

更なる実施形態によれば、バイノーラル・アンビエントサウンド・レンダリングデバイス1010は、第2のバイノーラル出力信号1015の提供のために、アンビエント信号部分125 2に対して空間効果および/またはフィルタを適用し、第2のバイノーラル出力信号1015が現実の拡散音場の両耳間コヒーレンスに適合するように構成される。

【0105】

上記実施形態において、非相関性とコヒーレンス制御は、2つの連続的なステップにおいて実行することができるが、これは必要条件でない。非コヒーレント信号の中間の数式化なしに、単一ステップのプロセスで同じ結果を得ることも可能である。両方法は等しく有効である。

【0106】

図11は、多重チャンネル入力オーディオ信号101のバイノーラル再生の実施形態1100の概念的なブロック図を示す。具体的には、図11の実施形態は、多重チャンネル入力オーディオ信号101のバイノーラル再生の装置を表し、第1の変換器1110（「周波数変換」）と、分離器1120（「ダイレクト アンビエンス分離」）と、バイノーラル・ダイレクトサウンド・レンダリングデバイス910（「ダイレクト音源レンダリング」）と、バイノーラル・アンビエントサウンド・レンダリングデバイス1010（「アンビエントサウンドレンダリング」）と、『プラス』で示される結合器1130と、第2の変換器1140（「逆周波数変換」）を備える。特に、第1の変換器1110は、多重チャンネル入力オーディオ信号101をスペクトル表現1115に変換するように構成することができる。分離器1120は、スペクトル表現1115からダイレクト信号部分125 1またはアンビエント信号部分125 2を抽出するように構成することができる。ここで、分離器1120は、特にダイレクト/アンビエンス推定器110とダイレクト/アンビエンス抽出器120を含む図1の装置100に相当することができる。前に説明されたように、バイノーラル・ダイレクトサウンド・レンダリングデバイス910は、ダイレクト信号部分125 1について動作し、第1のバイノーラル出力信号915を取得することができる。同様に、バイノーラル・アンビエントサウンド・レンダリングデバイス1010は、アンビエント信号部分125 2について動作し、第2のバイノーラル出力信号1015を取得することができる。結合器1130は、第1のバイノーラル出力信号915と第2のバイノーラル出力信号1015を結合し、結合信号1135を取得するように構成することができる。最後に、第2の変換器1140は、結合信号1135を時間ドメインに変換し、ステレオ出力オーディオ信号1150（「ヘッドホンに対するステレオ出力」）を取得するように構成することができる。

【0107】

図11の実施形態の周波数変換演算は、システムが、空間オーディオの知覚処理における固有のドメインである周波数変換ドメインにおいて機能することを示す。既に周波数変換ドメインにおいて機能しているシステムにおいてアドオンとして使用されている場合、システム自身は必ずしも周波数変換を有しない。

【0108】

上記のダイレクト/アンビエンス分離プロセスは、2つの異なる部分に細分化することができる。ダイレクト/アンビエンス推定部において、ダイレクトアンビエント部分のレ

10

20

30

40

50

ベルおよび/またはレシオは、オーディオ信号の信号モデルと属性の組合せに基づいて推定される。ダイレクト/アンビエンス抽出部において、既知のレシオおよび入力信号は、出力ダイレクト/アンビエンス信号の生成に用いることができる。

【0109】

最後に、図12は、バイノーラル再生の使用のケースを含むダイレクト/アンビエンス推定/抽出の実施形態1200の全体ブロック図を示す。特に、図12の実施形態1200は、図11の実施形態1100に相当することができる。しかしながら、実施形態1200において、空間パラメータ情報105に基づく推定/抽出プロセスを含む、図1の実施形態のブロック110、120に対応する図11の分離器1120の詳細が示される。加えて、図11の実施形態1100とは対照的に、図12の実施形態1200においては、異なるドメイン間の変換プロセスが示されていない。実施形態1200のブロックは、また、多重チャンネルオーディオ信号101から導き出すことができるダウンミックス信号115について明らかに動作する。

10

【0110】

図13aは、フィルタバンクドメインにおいてモノラルダウンミックス信号からダイレクト/アンビエンス信号を抽出する装置1300の実施形態のブロック図を示す。図13aに示されるように、装置1300は、解析フィルタバンク1310と、ダイレクト部分に対する合成フィルタバンク1320と、アンビエント部分に対する合成フィルタバンク1322を備える。

【0111】

20

特に、装置1300の解析フィルタバンク1310は、短時間フーリエ変換(STFT)を実行するように実施することができるか、または、例えば、解析QMFフィルタバンクとして構成することができるか、一方、装置1300の合成フィルタバンク1320、1322は、逆短時間フーリエ変換(ISTFT)を実行するように実施することができるか、または、例えば、合成QMFフィルタバンクとして構成することができる。

【0112】

解析フィルタバンク1310は、図2の実施形態に示されたようなモノラルダウンミックス信号215に相当することができるモノラルダウンミックス信号1315を受信し、モノラルダウンミックス信号1315を複数のフィルタバンクサブバンド1311に変換するように構成される。図13aに見られるように、複数のフィルタバンクサブバンド1311は、それぞれ複数のダイレクト/アンビエンス抽出ブロック1350、1352に接続され、複数のダイレクト/アンビエンス抽出ブロック1350、1352は、それぞれ DTT_{mono} または ATT_{mono} ベースのパラメータ1333、1335をフィルタバンクサブバンドに適用するように構成される。

30

【0113】

DTT_{mono} 、 ATT_{mono} ベースのパラメータ1333、1335は、図13bに示されるように、 DTT_{mono} 、 ATT_{mono} 演算器1330から供給することができる。特に、図13bの DTT_{mono} 、 ATT_{mono} 演算器1330は、 DTT_{mono} 、 ATT_{mono} エネルギーレシオを計算するか、または、前に同様に記載されたパラメータステレオオーディオ信号(例えば、図2のパラメータステレオオーディオ信号201)の左チャンネルと右チャンネル(L、R)に対応する提供されたチャンネル間コヒーレンスパラメータとチャンネルレベル差パラメータ(ICC_L 、 CLD_L 、 ICC_R 、 CLD_R)105から DTT_{mono} 、 ATT_{mono} ベースのパラメータを導き出すように構成することができる。ここで、単一のフィルタバンクサブバンドに対して、対応するパラメータ105と、 DTT_{mono} 、 ATT_{mono} ベースのパラメータ1333、1335を用いることができる。この文脈において、それらのパラメータは周波数を通じて一定でないことが指摘される。

40

【0114】

DTT_{mono} または ATT_{mono} ベースのパラメータ1333、1335のアプリケーションの結果として、それぞれ複数の修正されたフィルタバンクサブバンド1353、1355が取得される。引き続き、複数の修正されたフィルタバンクサブバンド1353、

50

1355は、それぞれ合成フィルタバンク1320、1322に供給され、それぞれモノラルダウンミックス信号1315のダイレクト信号部分1325₁またはアンビエント信号部分1325₂を取得するために複数の修正されたフィルタバンクサブバンド1353、1355を合成するように構成される。ここで、図13aのダイレクト信号部分1325₁は、図2のダイレクト信号部分125₁に相当することができ、一方、図13aのアンビエント信号部分1325₂は、図2のアンビエント信号部分125₂に相当することができる。

【0115】

図13bを参照して、図13aの複数のダイレクト/アンビエンス抽出ブロック1350、1352の1つのダイレクト/アンビエンス抽出ブロック1380は、特にDTT_{mono}、ATT_{mono}演算器1330と乗算器1360を備える。乗算器1360は、複数のフィルタバンクサブバンド1311の単一のフィルタバンク(FB)サブバンド1301を、対応するDTT_{mono}/ATT_{mono}ベースのパラメータ1333、1335と乗算するように構成することができ、複数のフィルタバンクサブバンド1353、1355の修正された単一のフィルタバンクサブバンド1365を取得される。特に、ダイレクト/アンビエンス抽出ブロック1380は、ブロック1380が複数のブロック1350に帰属する場合に、DTT_{mono}ベースのパラメータを適用するように構成され、一方、ブロック1380が複数のブロック1352に帰属する場合に、ATT_{mono}ベースのパラメータを適用するように構成される。修正された単一のフィルタバンクサブバンド1365は、ダイレクト部分またはアンビエント部分に対して、それぞれの合成フィルタバンク1320、1322に更に供給することができる。

【0116】

実施形態によれば、空間パラメータおよび導き出されたパラメータは、一般にフィルタバンクの分解能より低い人間の聴覚システム、例えば28バンドのクリチカルバンドによる周波数分解能において与えられる。

【0117】

それ故、図13aの実施形態に係るダイレクト/アンビエンス抽出は、図3bのチャンネル間関係パラメータ335に相当することができるサブバンドに関して計算されたチャンネル間コヒーレンスとチャンネルレベル差パラメータに基づいて、フィルタバンクドメインにおいて異なるサブバンドについて基本的に動作する。

【0118】

図14は、本発明の更なる実施形態に係る例示的なMPEGサラウンド復号化スキーム1400の図解図を示す。特に、図14の実施形態は、ステレオダウンミックス1410から6つの出力チャンネル1420への復号化を記述する。ここで、「res」で示される信号は残差信号であり、それは(「D」で示されるブロックからの)非相関化信号に対するオプションの置き換えである。図14の実施形態によれば、図8のエンコーダ810のようなエンコーダから図8のデコーダ820のようなデコーダに対して、MPSストリームの中で伝送される空間パラメータ情報またはチャンネル間関係パラメータ(ICC、CLD)は、それぞれ「前非相関化マトリックスM1」と「混合マトリックスM2」で示される復号化マトリックス1430、1440を生成するために用いることができる。図14の実施形態に特有のことは、混合マトリックスM2₁₄₄₀を用いたサイドチャンネル(L、R)とセンターチャンネル(C)(L、R、C₁₄₃₅)からの出力チャンネル1420(すなわちアップミックスチャンネルL、LS、R、RS、C、LFE)の生成が、MPSサラウンド標準による特定のチャンネル間関係パラメータ(ICC、CLD)を備える、図1の空間パラメータ情報105に相当することができる空間パラメータ情報1405によって、基本的に決定されるということである。

【0119】

ここで、左チャンネル(L)の対応する出力チャンネルL、LSへの分割と、右チャンネル(R)の対応する出力チャンネルR、RSへの分割と、センターチャンネル(C)の対応する出力チャンネルC、LFEへの分割は、それぞれ、対応するICC、CLDパラ

10

20

30

40

50

メータに対してそれぞれの入力を有する 1 から 2 への (O T T) 構成によって表すことができる。

【 0 1 2 0 】

特に「 5 2 5 構成」に相当する例示的な M P E G サラウンド復号化スキーム 1 4 0 0 は、例えば、次のステップを備えることができる。第 1 ステップにおいて、空間パラメータまたはパラメータサイド情報は、既存の M P S サラウンド標準に従って、図 1 4 に示される復号化マトリックス 1 4 3 0、1 4 4 0 に数式化することができる。第 2 ステップにおいて、復号化マトリックス 1 4 3 0、1 4 4 0 は、パラメータドメインにおいて、アップミックスチャンネル 1 4 2 0 のチャンネル間情報を提供するために用いることができる。第 3 ステップにおいて、このように提供されたチャンネル間情報によって、各アップミックスチャンネルのダイレクト/アンビエンスエネルギーを計算することができる。第 4 ステップにおいて、このように取得されたダイレクト/アンビエンスエネルギーは、ダウンミックスチャンネル 1 4 1 0 の数にダウンミックスすることができる。第 5 ステップにおいて、ダウンミックスチャンネル 1 4 1 0 に適用される重みを計算することができる。

10

【 0 1 2 1 】

更に進める前に、ちょうど言及された例示のプロセスが、ダウンミックスチャンネルから、ダウンミックスチャンネルの平均パワーである

$$E[L_{dmx}^2], E[R_{dmx}^2]$$

20

と、クロススペクトルと称することができる

$$E[L_{dmx}R_{dmx}^*]$$

の測定を必要とすることが指摘される。ここで、用語「平均パワー」は使用される共通の用語ではないので、ダウンミックスチャンネルの平均パワーは意図的にエネルギーと称される。

【 0 1 2 2 】

中括弧で示される予測オペレータは、実用的なアプリケーションにおいて、再帰的または非再帰的な時間平均によって置き換えることができる。エネルギーとクロススペクトルは、ダウンミックス信号から直接的に測定可能である。

30

【 0 1 2 3 】

また、2つのチャンネルの線形結合のエネルギーは、チャンネルのエネルギー、混合係数およびクロススペクトルから(すべて信号演算が必要でないパラメータドメインにおいて)数式化できる点に注意すべきである。

【 0 1 2 4 】

線形結合

$$Ch = aL_{dmx} + bR_{dmx}$$

は、次のエネルギーを有する。

$$\begin{aligned} E[|Ch|^2] &= E[|aL_{dmx} + bR_{dmx}|^2] = a^2 E[|L_{dmx}|^2] + b^2 E[|R_{dmx}|^2] + ab(E[L_{dmx}R_{dmx}^*] + E[R_{dmx}L_{dmx}^*]) \\ &= a^2 E[|L_{dmx}|^2] + b^2 E[|R_{dmx}|^2] + 2ab(\operatorname{Re}\{E[L_{dmx}R_{dmx}^*]\}) \end{aligned}$$

40

【 0 1 2 5 】

以下は、例示のプロセス(すなわち復号化スキーム)の個々のステップを記述する。

【 0 1 2 6 】

第 1 ステップ(混合マトリックスに対する空間パラメータ)

【 0 1 2 7 】

前述のように、M 1 マトリックスと M 2 マトリックスは、M P S サラウンド標準に従って作成される。M 1 の第 a 行、第 b 列の要素は M 1 (a , b) である。

【 0 1 2 8 】

50

第2ステップ(アップミックスチャンネルのチャンネル間情報へのダウンミックスのエネルギーとクロススペクトルを備えた混合マトリックス)

【0129】

ここで、混合マトリックスM1およびM2を有する。左のダウンミックスチャンネル(L_{dmx})と右のダウンミックスチャンネル(R_{dmx})から出力チャンネルが生成される方法を数式化する必要がある。非相関化器が用いられることを仮定する(図14のグレイ領域)。MPS標準における復号化/アップミックスは、最終的には、全部のプロセスにおいて、全体の入出力関係に対して以下の数式を基本的に提供する。

$$L = a_L L_{dmx} + b_L R_{dmx} + c_L D_1[S_1] + d_L D_2[S_2] + e_L D_3[S_3]$$

10

【0130】

上記は、アップミックスされた正面左チャンネルに対する例である。他のチャンネルは、同様に数式化することができる。D要素は、非相関化器であり、a eは、M1とM2のマトリクスエントリから計算可能な重みである。

【0131】

特に、係数a - eは、次のように、マトリクスエントリから直接的に数式化することができる。

$$a_L = \sum_{i=1}^3 M1_{i,1} M2_{1,i}$$

$$b_L = \sum_{i=1}^3 M1_{i,2} M2_{1,i}$$

$$c_L = M2_{1,4}$$

$$d_L = M2_{1,5}$$

$$e_L = M2_{1,6}$$

そして、他チャンネルに対しても同様である。

【0132】

S信号は、次の通りである。

$$S_n = M1_{n+3,1} L_{dmx} + M1_{n+3,2} R_{dmx}$$

30

【0133】

これらのS信号は、図14における左辺マトリクスからの非相関化器への入力である。エネルギー

$$E[D[S_n]^2] = E[|S_n|^2]$$

は、上で説明されたように計算することができる。非相関化器は、エネルギーに影響を及ぼさない。多重チャンネルのアンビエンス抽出をする知覚的に動機づけられた方法は、1つのチャンネルを他のすべてのチャンネルの合計と比較することである。(これは、多くの方法のうちのオプションの1つである点に注意されたい)ここで、例示的にチャンネルLのケースを考慮すると、残りのチャンネルは、次のように読み取れる。

$$X_L = \sum_{Ch=(REST)} a_{Ch} L_{dmx} + \sum_{Ch=(REST)} b_{Ch} R_{dmx} + \sum_{Ch=(REST)} c_{Ch} D_1[S_1] + \sum_{Ch=(REST)} d_{Ch} D_2[S_2] + \sum_{Ch=(REST)} e_{Ch} D_3[S_3]$$

40

【0134】

ここで「残りのチャンネル」に対して「R」を使用することは混乱させるかもしれないので、記号「X」を使用している。

【0135】

次に、チャンネルLのエネルギーは、次の通りである。

$$E[L^2] = a_L^2 E[L_{dmx}^2] + b_L^2 E[R_{dmx}^2] + c_L^2 E[S_1^2] + d_L^2 E[S_2^2] + e_L^2 E[S_3^2] + 2ab \operatorname{Re}\{E[L_{dmx} R_{dmx}^*]\}$$

50

【 0 1 3 6 】

次に、チャンネル X のエネルギーは、次の通りである。

$$E[|X_L|^2] = \left(\sum_{Ch=(REST)} a_{Ch} \right)^2 E[|L_{dmx}|^2] + \left(\sum_{Ch=(REST)} b_{Ch} \right)^2 E[|R_{dmx}|^2] + \left(\sum_{Ch=(REST)} c_{Ch} \right)^2 E[|S_1|^2] + \left(\sum_{Ch=(REST)} d_{Ch} \right)^2 E[|S_2|^2] \\ + \left(\sum_{Ch=(REST)} e_{Ch} \right)^2 E[|S_3|^2] + 2 \left(\sum_{Ch=(REST)} a_{Ch} \sum_{Ch=(REST)} b_{Ch} \right) \text{Re} \{ E[L_{dmx} R_{dmx}^*] \}$$

【 0 1 3 7 】

そして、クロススペクトルは、次の通りである。

$$E[LX_L^*] = \sum_{Ch=(REST)} a_{Ch} a_L E[|L_{dmx}|^2] + \sum_{Ch=(REST)} b_{Ch} b_L E[|R_{dmx}|^2] + \sum_{Ch=(REST)} c_{Ch} c_L E[|S_1|^2] + \sum_{Ch=(REST)} d_{Ch} d_L E[|S_2|^2] \\ + \sum_{Ch=(REST)} e_{Ch} e_L E[|S_3|^2] + \sum_{Ch=(REST)} a_L b_{Ch} E[L_{dmx} R_{dmx}^*] + \sum_{Ch=(REST)} a_{Ch} b_L E[L_{dmx} R_{dmx}^*]^*$$

10

【 0 1 3 8 】

ここで、ICC を次のように数式化することができ、

$$ICC_L = \frac{\text{Re} \{ E[LX_L^*] \}}{\sqrt{E[|L|^2] E[|X_L|^2]}}$$

20

シグマは、次の通りである。

$$\sigma_L = \frac{E[|L|^2]}{E[|X_L|^2]}$$

【 0 1 3 9 】

第3ステップ（アップミックスチャンネルにおけるアップミックスチャンネルの DTT パラメータに対するチャンネル間情報）

【 0 1 4 0 】

ここで、次式に従ってチャンネル L の DTT を計算することができる。

$$DTT_L = \frac{1}{2} \left[\left(1 - \frac{1}{\sigma_L} \right) + \sqrt{\left(\frac{1}{\sigma_L} - 1 \right)^2 + 4 \frac{ICC_L^2}{\sigma_L}} \right]$$

30

【 0 1 4 1 】

L のダイレクトエネルギーは、次の通りである。

$$E[|D_L|^2] = DTT \cdot E[|L|^2]$$

L のアンビエンスエネルギーは、次の通りである。

40

$$E[|A_L|^2] = (1 - DTT) \cdot E[|L|^2]$$

【 0 1 4 2 】

第4ステップ（ダイレクト/アンビエントエネルギーをダウンミックスする）

【 0 1 4 3 】

例示的に非コヒーレントのダウンミックスルールを用いると、左のダウンミックスチャンネルのアンビエンスエネルギーは、次の通りであり、

$$E[A_{L,dmx}|^2] = E[A_{L,l}|^2] + E[A_{L,s}|^2] + \frac{E[A_C|^2] + E[A_{LF}|^2]}{2}$$

ダイレクト部分および右チャンネルのダイレクト部分とアンビエント部分に対しても同様である。上記はまさに1つのダウンミックスルールである点に注意されたい。同様に、他のダウンミックスルールが存在することができる。

【0144】

第5ステップ(ダウンミックスチャンネルにおけるアンビエンス抽出に対する重みを計算する) 10

【0145】

左のダウンミックスDTTレシオは、次の通りである。

$$DTT_{L,dmx} = 1 - \frac{E[A_{L,dmx}|^2]}{E[L_{dmx}|^2]}$$

【0146】

重み係数は、次に、図5の実施形態に記載されたように(すなわち、 \sqrt{DTT} または $\sqrt{1 - DTT}$)アプローチを用いて)、または図6の実施形態のように(すなわち、クロスミックスマトリックス法を用いて)計算することができる。 20

【0147】

基本的に、上述した例示的プロセスは、MPSストリームにおけるCPC、ICCおよびCLDパラメータを、ダウンミックスチャンネルのアンビエンスレシオに関連づける。

【0148】

更なる実施形態によれば、通常は同様の目的を達成する他の手段があり、同様に他の条件がある。例えば、特定のチャンネルが残りのチャンネルと比較される前に記載されたものと比べて、ダウンミックスに対する他のルール、他のスピーカレイアウト、他の復号化方法および多重チャンネルアンビエンス推定をなす他の方法が存在することができる。

【0149】

本発明は、ブロックが実際のまたは論理的なハードウェアコンポーネントを表すブロック図の前後関係において記載されていたが、本発明は、また、コンピュータで実施される方法で実施することができる。後者の場合、ブロックは対応する方法ステップを表し、これらのステップは対応する論理的または物理的ハードウェアブロックによって実行される機能を表す。 30

【0150】

記載された実施形態は、単に本発明の原理に対して例証を示したものである。本願明細書に記載された構成および詳細の修正および変更は、他の当業者にとって明らかであると理解される。本発明は、それ故に、特許クレームの範囲のみによって制限され、本願明細書の実施形態の記述および説明によって提案された特定の詳細によって制限されないことを意図する。 40

【0151】

発明の方法の特定の実施要求に依存して、発明の方法はハードウェアにおいてまたはソフトウェアにおいて実施することができる。実施は、格納された電子的に読込可能な制御信号を有し、発明の方法が実行されるようにプログラム可能なコンピュータシステムと協働する、デジタル記録メディア、特に、ディスク、DVDまたはCD、を使用して実行することができる。一般に、本発明は、それ故に、機械読取可能なキャリアに格納されるプログラムコードを有するコンピュータプログラム製品として実施することができ、プログラムコードは、コンピュータプログラム製品がコンピュータ上で動作するとき、発明の方法を実行するように動作する。言い換えれば、発明の方法は、それ故、コンピュータ 50

プログラムがコンピュータ上で動作するとき、発明の方法の少なくとも1つを実行するためのプログラムコードを有するコンピュータプログラムである。発明の符号化されたオーディオ信号は、デジタル記憶媒体のようないかなる機械読取可能な記憶媒体上にも記憶することができる。

【0152】

新規なコンセプトおよび技術の効果は、上述の実施形態、すなわちこの出願に記載された装置、方法またはコンピュータプログラムが、パラメータ空間情報の助けによりオーディオ信号からダイレクトおよび/またはアンビエンス成分を推定および抽出することを可能とすることである。特に、本発明の新規な処理は、通常はアンビエンス抽出の分野において、周波数バンドにおいて機能する。提案されたコンセプトは、オーディオ信号からのダイレクトおよびアンビエンス成分の分離を必要とする多くのアプリケーションがあるので、オーディオ信号処理に関連する。

10

【0153】

先行技術のアンビエンス抽出方法に対立して、本発明のコンセプトは、ステレオ入力信号にだけ基づいているものではなく、モノラルダウンミックスにも適用することができる。単一のチャンネルダウンミックスに対しては、一般に、チャンネル間差分を演算することができない。しかしながら、アンビエンス抽出は、空間サイド情報を考慮に入れることによって、この場合においても可能になる。

【0154】

本発明は、「オリジナル」の信号のアンビエンスレベルを推定するために、空間パラメータを利用するという点で有利である。それは、空間パラメータが既に「オリジナル」のステレオまたは多重チャンネル信号のチャンネル間差分に関する情報を含んでいるというコンセプトに基づいている。

20

【0155】

一旦オリジナルのステレオまたは多重チャンネルのアンビエンスレベルが推定されると、提供されたダウンミックスチャンネルにおけるダイレクトおよびアンビエンスレベルも抽出することができる。これは、アンビエント部分に対するアンビエンスエネルギーとダイレクト部分に対するダイレクトエネルギーまたは振幅の線形結合(すなわち重み付き和)によってなすことができる。それ故、本発明の実施形態は、空間サイド情報の助けによりアンビエンス推定および抽出を提供する。

30

【0156】

このコンセプトのサイド情報ベースの処理から拡張して、次の有益な特性または効果が存在する。

【0157】

本発明の実施形態は、空間サイド情報と提供されたダウンミックスチャンネルの助けによりアンビエンス推定を提供する。このようなアンビエンス推定は、サイド情報とともに提供される1つ以上のダウンミックスチャンネルがある場合に重要である。サイド情報と、ダウンミックスチャンネルから測定される情報は、アンビエンス推定において一緒に用いることができる。ステレオダウンミックスを備えるMP EGサラウンドにおいて、これらの2つの情報源は、オリジナルの多重チャンネルサウンドのチャンネル間関係の全部の情報を一緒に提供し、アンビエンス推定はこれらの関係に基づいている。

40

【0158】

本発明の実施形態は、また、ダイレクトエネルギーとアンビエンスエネルギーのダウンミックスを提供する。サイド情報ベースのアンビエンス抽出の記述された状況において、提供されたダウンミックスチャンネルより多いチャンネル数でアンビエンスを推定する中間ステップがある。それ故、このアンビエンス情報は、ダウンミックスオーディオチャンネル数に有効な方法でマップされなければならない。このプロセスは、オーディオチャンネルダウンミックスへの一致によりダウンミックスと称することができる。これは、提供されたダウンミックスチャンネルがダウンミックスされたのと同じ方法で、ダイレクトエネルギーとアンビエンスエネルギーを結合することによって、最も直接的になすことがで

50

きる。

【0159】

ダウンミックスルールは、1つの理想的な解を持たないが、アプリケーションに依存し
そうである。例えば、MPEGサラウンドにおいて、チャンネル（センター、正面のスピー
カー、後部スピーカ）を、それらの通常は異なる信号コンテンツにより異なって処理する
ことは有益になり得る。

【0160】

さらに、実施形態は、他のチャンネルに関して、多重チャンネルアンビエンス推定を各
チャンネルにおいて独立に提供する。この特性/アプローチは、提案されたステレオアン
ビエンス推定式を、全ての他のチャンネルに関連する各チャンネルに簡単に使用すること
を可能にする。この測定によって、全てのチャンネルにおいて等しいアンビエンスレベル
を仮定する必要がある。提案された方法は、各チャンネルにおけるアンビエント成分が全
ての他のチャンネルのいくつかにおいてコヒーレント相当物を持つ成分であるという空間
知覚についての仮定に基づいている。この仮定の有効性を示唆する例は、ノイズ（アンビ
エンス）を放射する2つのチャンネルうちの1つが、認知されたサウンドシーンに有意に
影響を及ぼさずに、各々半分のエネルギーで更に2つのチャンネルに分割することができ
ることである。

【0161】

信号処理に関して、実際のダイレクト/アンビエンスレシオ推定が、提案されたアンビ
エンス推定式を各チャンネル対全ての他のチャンネルの線形結合に適用することによって
生ずることは、有益である。

【0162】

最後に、実施形態は、実際の信号を抽出するために、推定されたダイレクトアンビエン
スエネルギーのアプリケーションを提供する。一旦ダウンミックスチャンネルにおけるア
ンビエンスレベルが知られると、アンビエンス信号を取得するために2つの発明の方法を
適用することができる。第1の方法は、簡単な乗算に基づいており、各ダウンミックスチ
ャンネルに対するダイレクトおよびアンビエント部分は、信号を $s_{q r t}$ （ダイレクトから
トータルへのエネルギーレシオ）および $s_{q r t}$ （アンビエントからトータルへのエネ
ルギーレシオ）で乗算することによって生成することができる。これは、各ダウンミッ
クスチャンネルに対して、お互いにコヒーレントであるが、ダイレクトおよびアンビエント
部分が持つと推定されたエネルギーを持つ2つの信号を提供する。

【0163】

第2の方法は、チャンネルのクロスミックスによる最小平均二乗解法に基づいており、
チャンネルクロスミックス（負符号によっても可能な）は、上記の解法より良好なダイレ
クト/アンビエンス信号の推定を可能にする。非特許文献3および特許文献2において提
供されるステレオ入力とチャンネルにおける等しいアンビエンスレベルに対する最小平均
解法と対比して、本発明は、等しいアンビエンスレベルを必要としない最小平均二乗解法
を提供し、またいかなるチャンネル数にも拡張可能である。

【0164】

新規な処理の付加的な特性は、以下の通りである。バイノーラルレンダリングに対する
アンビエンス処理において、アンビエンスは、現実の拡散音場における両耳間のコヒーレ
ンスと同様の周波数バンドにおける両耳間のコヒーレンスを提供する特性を有するフィル
タによって処理することができ、当該フィルタは、空間効果をも含むことができる。バイ
ノーラルレンダリングに対するダイレクト部分の処理において、ダイレクト部分は、初期
の反射および/または反響のような追加可能な空間効果を備えた頭部関連伝達関数（HRTFs）
を通して供給することができる。

【0165】

この他に、ドライ/ウェット制御に対応する「分離レベル」の制御は、更なる実施形態
において実現することができる。特に、完全な分離は、急激な変化、変調効果、その他の
ような聞き取れる人工産物に導く可能性があるため、多くのアプリケーションにおいて好

10

20

30

40

50

ましくない可能性がある。それ故、記載されたプロセスの全ての関連する部分は、所望の、有用な分離の量を制御する「分離レベル」の制御によって実施することができる。図11に関して、このような分離レベルの制御は、それぞれ、ダイレクト/アンビエンス分離1120および/またはバイノーラルレンダリングデバイス910、1010を制御する破線ボックスの制御入力1105によって示される。この制御は、オーディオエフェクト処理におけるドライ/ウェット制御と同様に働くことができる。

【0166】

提案された解法の主な利点は、以下の通りである。システムは、ダウンミックス情報のみに依存する前の解法と異なって、すべての状況において、パラメータステレオおよびモノラルダウンミックスを備えるMPEGサラウンドとも連動する。このシステムは、更に、ダウンミックスチャンネルの簡単なチャンネル間解析によるよりも正確にダイレクトおよびアンビエンスエネルギーを推定するために、空間オーディオビットストリームにおいてオーディオ信号とともに伝達される空間サイド情報を利用することが可能である。それ故、バイノーラル処理のような多くのアプリケーションは、サウンドのダイレクトおよびアンビエント部分に対して異なる処理を適用することによって利益を得ることができる。

【0167】

実施形態は、次のサウンド心理学的仮説に基づいている。人間の聴覚システムは、時間周波数タイル（特定の周波数と時間範囲に限定された領域）における両耳間のキューに基づいて音源をローカライズする。時間と周波数においてオーバーラップする2つ以上の非コヒーレントな共存する音源が異なる位置において同時に存在する場合、聴覚システムは音源の位置を認知することができない。これは、これらの音源の合計が、リスナー上に、信頼できる両耳間のキューを生成しないからである。聴覚システムは、このように、信頼できるローカライズ情報を提供するクローズした時間周波数タイルをピックアップし、残りをローカライズできないものとして取り扱うように記述することができる。これらの手段によって、聴覚システムは、複雑なサウンド環境において音源をローカライズすることができる。同時に起こるコヒーレント音源は、異なる効果を有し、コヒーレント音源の間に単一の音源を形成される同じ両耳間のキューを近似的に形成する。

【0168】

これは、また、実施形態が利用する特性である。ローカライズ可能な（ダイレクト）サウンドとローカライズ不能な（アンビエンス）サウンドのレベルは、推定することができ、次にこれらの成分が抽出される。空間化信号処理は、ローカライズ可能な/ダイレクト部分にのみ適用され、一方、拡散性/開放性/エンベロープ処理は、ローカライズ不能な/アンビエント部分に適用される。これは、多くのプロセスが必要とされるところでのみ適用され、残りの信号が影響されないままにすることができるので、バイノーラル処理システムの設計において有意の利益を与える。全ての処理は、人間の聴覚の周波数分解能を近似する周波数バンドにおいて起こる。

【0169】

実施形態は、知覚品質を最大にするが、認知される問題点を最小にする信号の分解に基づいている。このような分解によって、オーディオ信号のダイレクト成分とアンビエント成分を分離して取得することが可能である。2つの成分は、次に、所望の効果または表現を得るために、更に処理することができる。

【0170】

具体的には、本発明の実施形態は、符号化ドメインにおいて空間サイド情報の助けによってアンビエンス推定を可能にする。

【0171】

本発明は、また、オーディオ信号のヘッドホン再生の代表的な問題を、信号をダイレクト信号とアンビエント信号に分離することによって低減することができるという点で有益である。実施形態は、既存のダイレクト/アンビエンス抽出方法を改善し、ヘッドホン再生のためのバイノーラルサウンドレンダリングに適用されるようにすることを可能にする。

【 0 1 7 2 】

空間サイド情報ベースの処理の主要な使用のケースは、当然MPEGサラウンドとパラメータステレオ（および同様のパラメータコーディング技術）である。アンビエンス抽出から利益を受ける代表的なアプリケーションは、異なる程度の空間効果を異なる部分のサウンドに適用する能力によるバイノーラル再生と、異なる成分のサウンドを異なって位置決めおよび処理する能力による高いチャンネル数へのアップミックスである。また、例えば、音声了解度を強化する目的のために、ユーザがダイレクト/アンビエンスレベルの修正を必要とするアプリケーションが存在することができる。

【 図 1 】

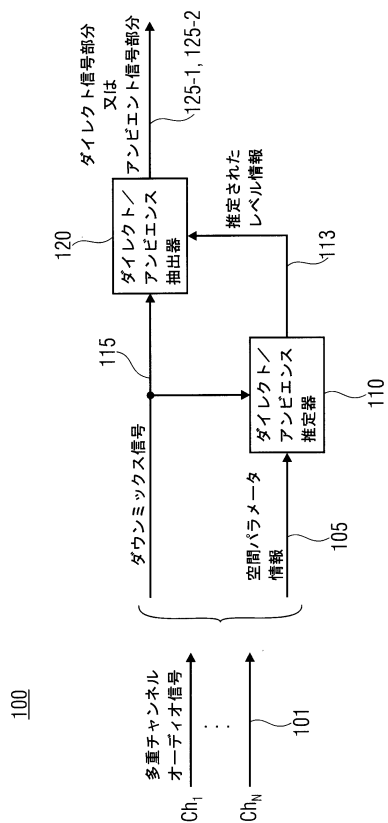


FIG 1

(ダイレクト/アンビエンス信号を抽出する装置)

【 図 2 】

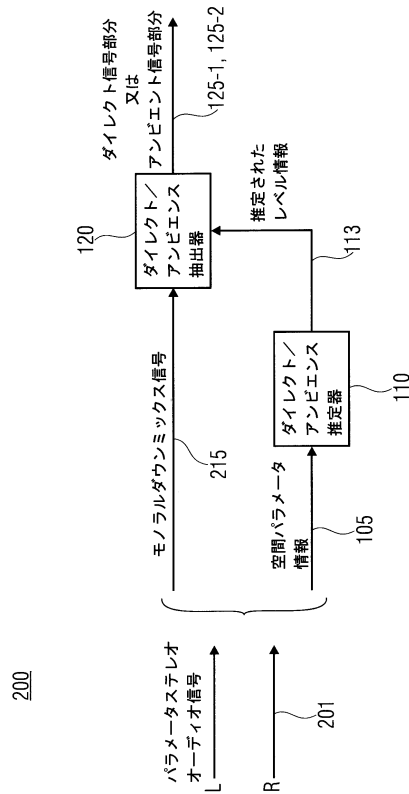


FIG 2

(パラメータステレオストリーム)

【図3a】

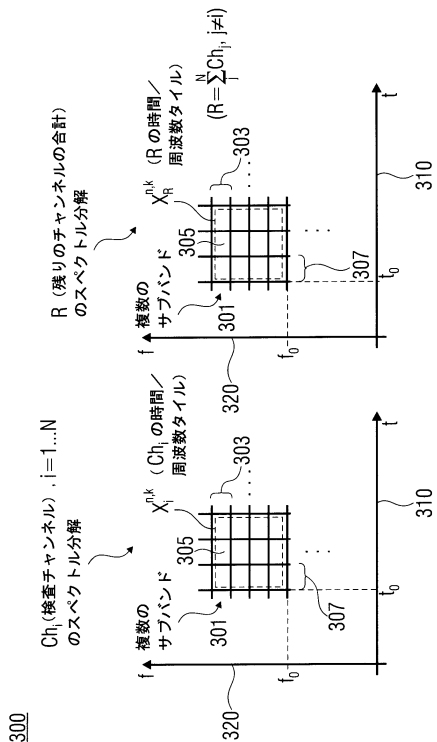


FIG 3A
(多重チャンネルオーディオ信号 Ch₁...Ch_N のスペクトル分解)

【図3b】

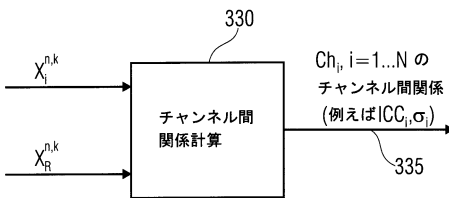


FIG 3B
(チャンネル間関係)

【図4】

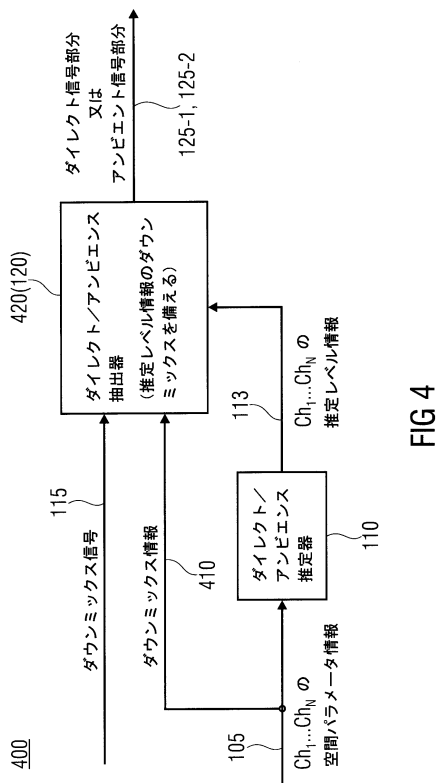


FIG 4
(推定レベル情報のダウンミックスを備えるダイレクト/アンビエンス抽出)

【図5】

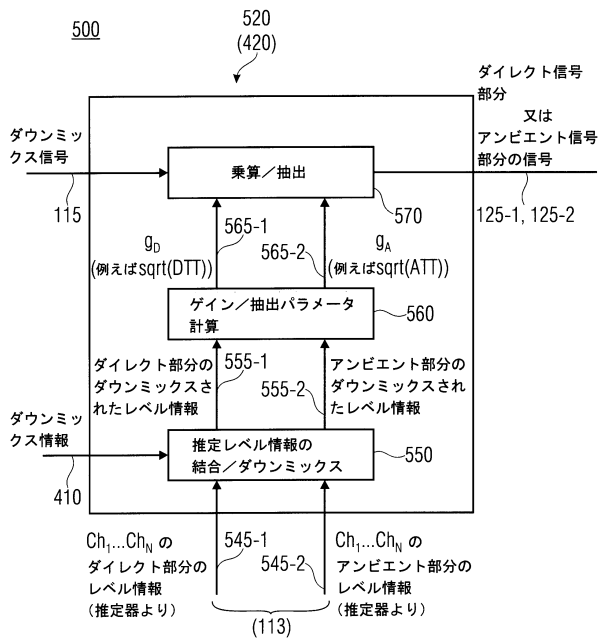


FIG 5
(ゲインパラメータの適用によるダイレクト/アンビエンス抽出)

【図6】

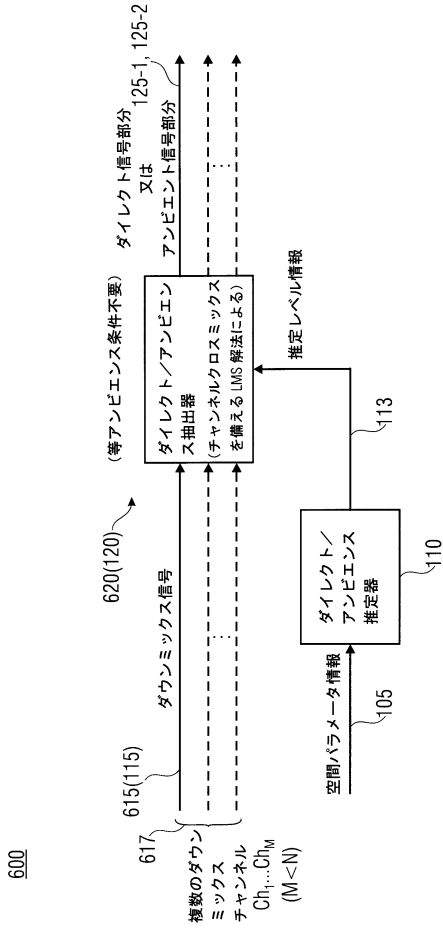


FIG 6
(LMS 解法によるダイレクト/アンビエンス抽出)

【図7a】

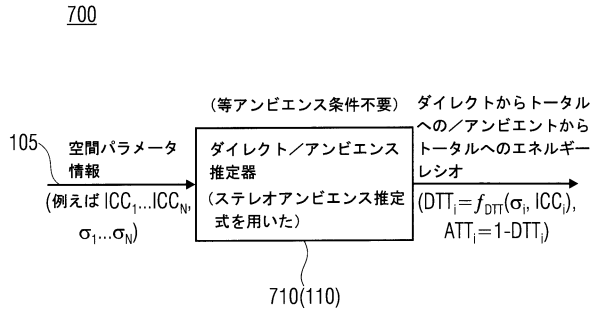
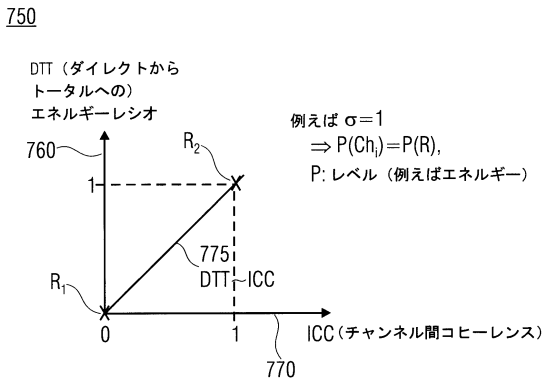


FIG 7A
(ステレオアンビエンス推定式を用いたダイレクト/アンビエンス推定)

【図7b】



R_1 : ICC=0 (完全に非コヒーレント)
 \Rightarrow DTT=0 (完全にアンビエント)

R_2 : ICC=1 (完全にコヒーレント)
 \Rightarrow DTT=1 (完全にダイレクト)

FIG 7B

(ダイレクトからトータルへのエネルギーレシオ
 対チャンネル間コヒーレンス)

【図8】

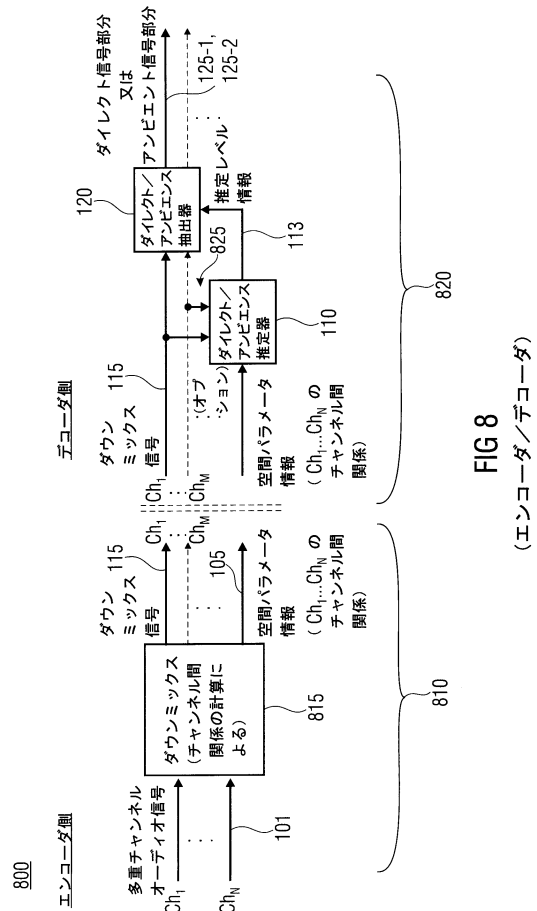


FIG 8
(エンコーダ/デコーダ)

【図 9 a】
900

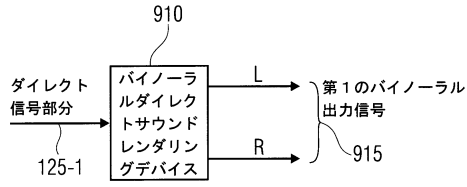
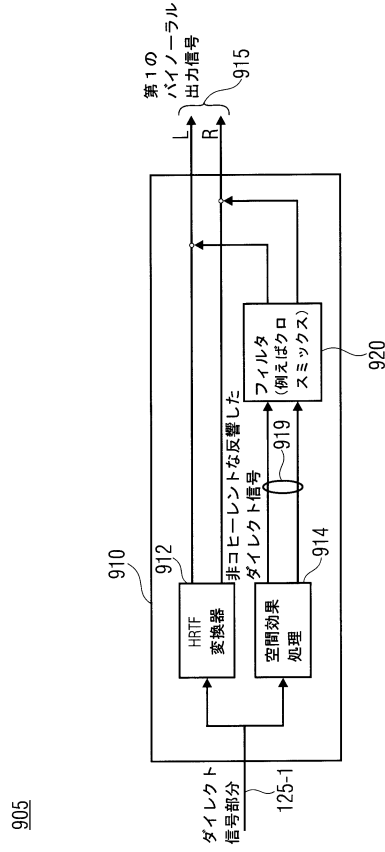


FIG 9A

(バイノーラルダイレクトサウンドレンダリングの概観)

【図 9 b】



905

FIG 9B

(バイノーラルダイレクトサウンドレンダリングの詳細)

【図 10 a】

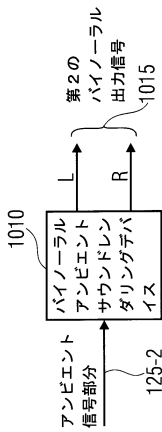
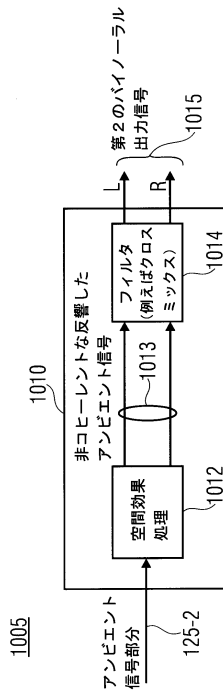


FIG 10A

(バイノーラルアンビエントサウンドレンダリングの概観)

【図 10 b】



1005

FIG 10B

(バイノーラルアンビエントサウンドレンダリングの詳細)

【図 1 1】

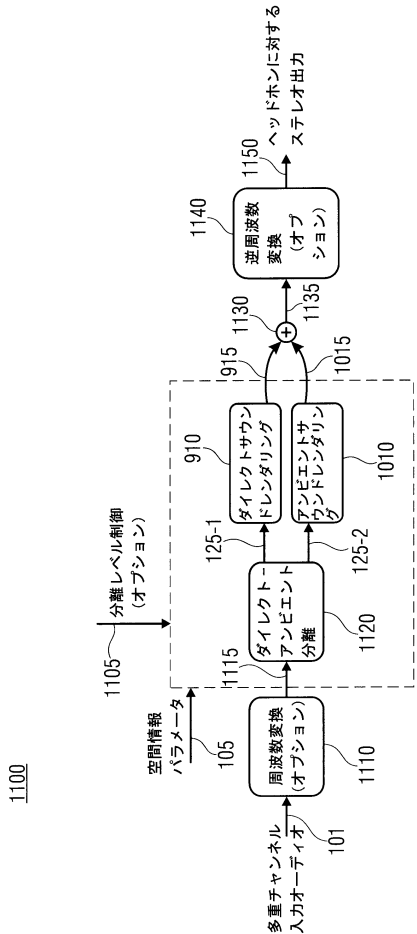


FIG 11

(提案されたバイノーラル再生の概念ブロック図)

【図 1 2】

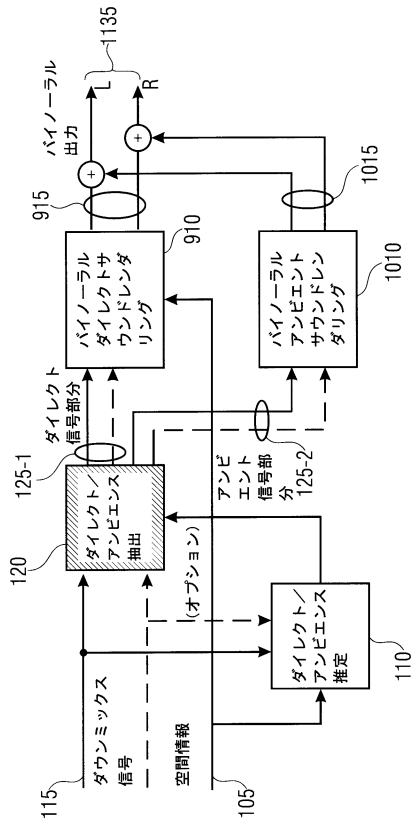
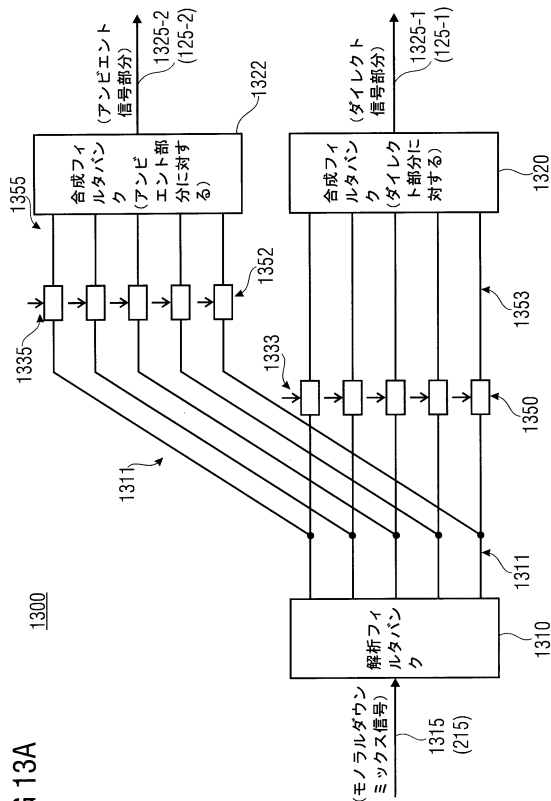


FIG 12

(バイノーラル再生の使用ケースを含む全体ブロック図)

【図 1 3 a】



(フィルタバンクドメインにおけるモノラルダウンミックスからのダイレクト/アンビエンス抽出)

FIG 13A

【図 1 3 b】

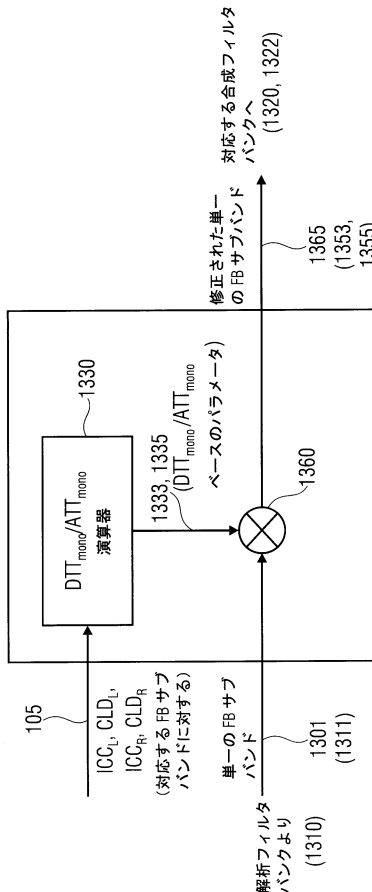


FIG 13B

(ダイレクトアンビエンス抽出ブロック)

【 図 1 4 】

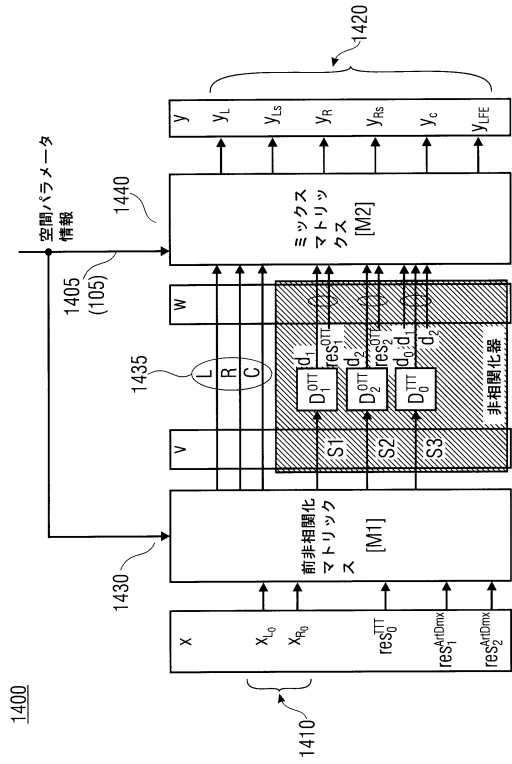


FIG 14

(MPEG サラウンド復号化スキーム)

フロントページの続き

- (72)発明者 ノイゲバウア ベルンハルト
ドイツ連邦共和国 91058 エアランゲン フランツェンスパーダー シュトラーセ 3
- (72)発明者 ヘレ ユールゲン
ドイツ連邦共和国 91054 ブッケンホフ ハーラーシュトラーセ 24

審査官 毛利 太郎

- (56)参考文献 米国特許出願公開第2007/0236858 (US, A1)
特表2009-531724 (JP, A)
則松武志, 低ビットレート高音質マルチチャンネルオーディオ符号化技術: MPEGサラウンド, パ
ナソニック技報, 日本, パナソニック株式会社, 2009年 1月15日, Vol.54, No.4, p.55
-59

- (58)調査した分野(Int.Cl., DB名)
- | | |
|------|---------------|
| G10L | 19/00 - 19/26 |
| H04S | 1/00 - 7/00 |
| H03M | 7/00 - 7/50 |