

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4943418号  
(P4943418)

(45) 発行日 平成24年5月30日 (2012.5.30)

(24) 登録日 平成24年3月9日 (2012.3.9)

(51) Int.Cl.

F I

G 1 O L 19/00 (2006.01)

G 1 O L 19/00 2 5 0

G 1 O L 19/00 2 1 3

請求項の数 24 (全 16 頁)

(21) 出願番号 特願2008-503630 (P2008-503630)  
 (86) (22) 出願日 平成18年3月16日 (2006.3.16)  
 (65) 公表番号 特表2008-535014 (P2008-535014A)  
 (43) 公表日 平成20年8月28日 (2008.8.28)  
 (86) 国際出願番号 PCT/IB2006/050819  
 (87) 国際公開番号 W02006/103581  
 (87) 国際公開日 平成18年10月5日 (2006.10.5)  
 審査請求日 平成21年3月16日 (2009.3.16)  
 (31) 優先権主張番号 05102506.2  
 (32) 優先日 平成17年3月30日 (2005.3.30)  
 (33) 優先権主張国 欧州特許庁 (EP)  
 (31) 優先権主張番号 05103077.3  
 (32) 優先日 平成17年4月18日 (2005.4.18)  
 (33) 優先権主張国 欧州特許庁 (EP)

(73) 特許権者 590000248  
 コーニンクレッカ フィリップス エレク  
 トロニクス エヌ ヴィ  
 オランダ国 5 6 2 1 ベーアー アイン  
 ドーフェン フルーネヴァウツウェッハ  
 1  
 (74) 代理人 100087789  
 弁理士 津軽 進  
 (74) 代理人 100114753  
 弁理士 宮崎 昭彦  
 (74) 代理人 100122769  
 弁理士 笛田 秀仙  
 (72) 発明者 マイブルグ フランソワ ピー  
 オランダ国 5 6 5 6 アーアー アイン  
 ドーフェン プロフ ホルストラーン 6  
 最終頁に続く

(54) 【発明の名称】 スケーラブルマルチチャネル音声符号化方法

(57) 【特許請求の範囲】

【請求項 1】

マルチチャネル音声信号を符号化するように構成される音声符号化器であって、  
 - 第 1 及び第 2 音声信号の合成された表現である主要信号部分及び残余信号部分を発生させる符号化器合成モジュールであって、前記主要及び残余信号部分が、前記第 1 及び第 2 音声信号に数学的演算を適用することによって得られ、前記数学的演算が、前記第 1 及び第 2 音声信号の空間特性の記述を含む第 1 空間パラメータを用いた演算である、符号化器合成モジュールと、  
 - パラメータ発生器であって、  
 - 前記第 1 及び第 2 音声信号の空間特性の記述を含む、前記第 1 空間パラメータとは異なる第 2 空間パラメータ、及び  
 - 前記第 1 及び第 2 音声信号の空間特性の記述を含む、前記第 1 空間パラメータ及び前記第 2 空間パラメータとは異なる第 3 空間パラメータ、  
 を発生させるパラメータ発生器と、  
 - 出力発生器であって、  
 - 前記主要信号部分及び前記第 2 空間パラメータを含む第 1 出力部分、及び  
 - 前記残余信号部分及び前記第 3 空間パラメータを含む第 2 出力部分、  
 を含む符号化出力信号を発生させる出力発生器と、  
 を備える音声符号化器。

【請求項 2】

10

20

前記第 3 空間パラメータが、前記第 2 空間パラメータと前記第 1 空間パラメータとの間の差を含む、請求項 1 に記載の音声符号化器。

【請求項 3】

前記第 2 空間パラメータが、コーヒレントに基づくパラメータを含む、請求項 1 に記載の音声符号化器。

【請求項 4】

前記第 3 空間パラメータが、コーヒレントに基づくパラメータと相関に基づくパラメータとの間の差を含む、請求項 1 に記載の音声符号化器。

【請求項 5】

前記残余信号部分が、前記第 1 及び第 2 音声信号の間の差を含む、請求項 1 から 4 いずれか 1 項 に記載の音声符号化器。

10

【請求項 6】

前記符号化器合成モジュールが、前記主要および残余信号部分を、これらの信号部分が、前記第 1 および第 2 音声信号よりも相関していないように発生させる、請求項 1 から 5 いずれか 1 項 に記載の音声符号化器。

【請求項 7】

第 3、第 4、第 5、及び第 6 音声信号を受信し、これらの信号を前記第 1 及び第 2 音声信号とともにダウンミックスし、これらの信号に応答して前記第 1 及び第 2 出力部分を発生させるように更に構成される、請求項 1 から 6 いずれか 1 項 に記載の音声符号化器。

【請求項 8】

20

請求項 1 から 7 いずれか 1 項 に記載の音声符号化器により発生させられた前記符号化出力信号から、マルチチャネル音声信号を復号する音声復号器であって、

- 前記主要信号部分、前記残余信号部分、前記第 2 空間パラメータ及び前記第 3 空間パラメータに基づき、前記第 1 及び第 2 音声信号を発生させる復号器合成ユニットであって、前記残余信号部分及び前記第 2 空間パラメータが、前記第 1 および第 2 音声信号を発生させるのに用いられる混合マトリクスを決定するのに関与される、復号器合成ユニット、を備える音声復号器。

【請求項 9】

前記復号器が、前記主要信号部分を受信し、これに応答して非相関化された主要信号部分を発生させる非相関化器を含む、請求項 8 に記載の音声復号器。

30

【請求項 10】

前記残余信号部分及び前記非相関化された主要信号部分の加算が、前記混合マトリクスを決定するのに関与される、請求項 9 に記載の音声復号器。

【請求項 11】

前記復号器が、前記非相関化された主要信号部分を前記残余信号部分に加える前に、前記非相関化された主要信号部分を減衰する減衰器を含む、請求項 10 に記載の音声復号器。

【請求項 12】

第 2 空間パラメータ及び第 3 空間パラメータの複数の群並びに複数の残余信号部分を受信し、これにより、これらの信号に応答して複数の群の第 1 及び第 2 音声信号を発生させるように構成される、請求項 8 から 11 いずれか 1 項 に記載の音声復号器。

40

【請求項 13】

復号器が、第 2 空間パラメータ及び第 3 空間パラメータの 3 つの群と 3 残余信号部分を受信し、これにより、これらの信号に応答して 3 つの群の第 1 及び第 2 音声信号を生成するように構成される、請求項 12 に記載の音声復号器。

【請求項 14】

マルチチャネル音声信号を符号化する方法であって、  
1) 第 1 及び第 2 音声信号の合成された表現である主要信号部分及び残余信号部分を発生させるステップであって、前記主要及び残余信号部分が、前記第 1 及び第 2 音声信号に数学的演算を適用することによって得られ、前記数学的演算が、前記第 1 及び第 2 音声信号

50

の空間特性の記述を含む第 1 空間パラメータを用いた演算である、ステップと、  
 2) 前記第 1 及び第 2 音声信号の空間特性の記述を含む、前記第 1 空間パラメータとは異なる第 2 空間パラメータを発生させるステップと、  
 3) 前記第 1 及び第 2 音声信号の空間特性の記述を含む、前記第 1 空間パラメータ及び前記第 2 空間パラメータとは異なる第 3 空間パラメータを発生させるステップと、  
 4) 前記主要信号部分及び前記第 2 空間パラメータを含む第 1 出力部分、並びに前記残余信号部分及び前記第 3 空間パラメータを含む第 2 出力部分、を含む符号化出力信号を発生させるステップと、  
 を含む方法。

【請求項 15】

10

請求項 14 記載の方法により発生させられた前記符号化出力信号から、マルチチャネル音声信号を復号する方法であって、

1) 前記主要信号部分、前記残余信号部分、前記第 2 空間パラメータ及び前記第 3 空間パラメータ、を含む前記符号化出力信号を受信するステップと、  
 2) 前記残余信号部分及び前記第 2 空間パラメータ群に基づき、混合マトリクスを決定するステップと、  
 3) 前記決定された混合マトリクスに基づき、前記第 1 および第 2 音声信号を発生させるステップと、  
 を含む方法。

【請求項 16】

20

前記主要信号部分を非相関化させ、これにตอบสนองして非相関化された主要信号部分を発生させるステップを含む、請求項 15 に記載の方法。

【請求項 17】

前記残余信号部分及び前記非相関化主要信号部分を加えるステップを更に含む、請求項 16 に記載の方法。

【請求項 18】

前記混合マトリクスを決定するステップが、前記加えられた残余信号部分及び前記非相関化主要信号部分に基づく、請求項 17 に記載の方法。

【請求項 19】

第 2 空間パラメータ及び第 3 空間パラメータの複数の群並びに複数の残余信号部分を受信し、これにより、これらの信号にตอบสนองして複数の群の第 1 及び第 2 音声信号を発生させるステップを含む、請求項 15 から 18 いずれか 1 項に記載の方法。

30

【請求項 20】

第 2 空間パラメータ及び第 3 空間パラメータの 3 つの群と 3 つの残余信号部分を受信し、これにより、これらの信号にตอบสนองして 3 つの群の第 1 及び第 2 音声信号を生成するステップを含む、請求項 19 に記載の方法。

【請求項 21】

請求項 14 に記載の方法を実行するように構成される計算機実行可能プログラム。

【請求項 22】

請求項 15 から 20 いずれか 1 項に記載の方法を実行するように構成される計算機実行可能プログラム。

40

【請求項 23】

請求項 1 から 7 いずれか 1 項に記載の符号化器を備える装置。

【請求項 24】

請求項 8 から 13 いずれか 1 項に記載の復号器を備える装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、高音質音声符号化の分野に関する。特に、本発明は、マルチチャネル音声データの高音質音声符号化の分野に関する。より特に、本発明は、マルチチャネル音声デー

50

タを符号化及び復号する符号化器及び復号器並びに方法に関する。

【背景技術】

【0002】

多くのマルチチャンネル機器構成/セットアップが可能であるものの、5.1機器構成/セットアップが、最も普及している(図1も参照)。通常のマルチチャンネル5.1セットアップは、5つのスピーカ、すなわち左フロント(Lf)、右フロント(Rf)、センタ(C)、左サラウンド(Ls)、及び右サラウンド(Rs)のスピーカ、並びにこれらを補完する、任意の角度で配置される追加的なLFE(低周波数エンハンス)スピーカからなる。過去において、5.1マルチチャンネル音声データなどのマルチチャンネル音声データを圧縮するいくつかの対策が検討されてきた。簡単な概略は以下に与えられる。

10

【0003】

MPEG-2音声規格である、「ISO/IEC 13818-3:1998 Information technology - 動画信号及び付随する音声信号のはん用符号化 第3部: 音声」においては、マルチチャンネル音声を符号化する規定であって、一方でMPEG-1音声への後方互換性を維持している規定が定められ、「ISO/IEC 11172-3:1993 Information technology--ディジタル記録媒体のための動画信号及び付随する音声信号の1.5Mbit/s 符号化 第3部: 音声」は、モノ及びステレオ音声の符号化に関してのみ提供する。後方互換性は、MPEG-1ビットストリームのデータ部分に配置されるマルチチャンネルコンテンツから導出される基本ステレオ信号を形成することによって、達成される。3つの追加的な信号は、この場合、MPEG-1ビットストリームの補助データ部分に配置される。この技術は、マトリクスিংと称される。MPEG-1音声復号器は、ビットストリームから意味のあるステレオ信号(Lo、Ro)を発生し得る一方で、MPEG-2音声復号器は、追加的なチャンネルを抽出し、5つの入力チャンネルの復号されたものを再構築し得る。後方互換性は、高ビットレートの犠牲により生じられる。通常、MPEG-2レイヤーIIを用いた5チャンネルのマテリアルに関して高音声品質を得るためには、640kbit/sのビットレートが必要とされる。

20

【0004】

MPEG-2 Advanced Audio Coding (AAC)、ISO/IEC TR 13818-5:1997/Amd 1:1999 Advanced Audio Coding(AAC)において、マルチチャンネル音声は、非後方互換性形式で符号化される。このことは、符号器に対してより自由を可能にし、640kbit/sのMPEG-2 Layer IIと比較して、より高い音声品質(透過性)が320kbit/sのビットレートで達成され得るという有利な点を有する。5(.1)チャンネル機器設定において、AACは、Mid-Side(MS)ステレオツール:(Lf、Rf)及び(Ls、Rs)を用いることによって、聴取者に対称であるチャンネル対を符号化し得る。センタ(C)及び(追加的な)LFEチャンネルは、別々に符号化される。代替的に、インテンシティステレオ(IS)符号化は、いくつかの音声チャンネルを1つのチャンネルに合成するために用いられ得、追加的に、各チャンネルに対してスケールン

30

【0005】

パラメトリックマルチチャンネル音声符号化において、知覚的に関連するキュー(cue)(すなわち空間パラメータ)、例えばチャンネル間インテンシティ差異(IID)、チャンネル間時間差異(ITD)、及びチャンネル間コーヒレント(ICC)などは、マルチチャンネル信号におけるチャンネル間で測定される。空間パラメータのより詳細な記載は、Christof Faller: "Coding of Spatial Audio Compatible with Different Playback Formats", AES Convention Paper, AES 117<sup>th</sup> Convention, San Francisco, USA,

40

2004 October 28-31において確認される。更に、マルチチャンネル表現は、標準モノ又はステレオ符号化器を用いて符号化され得るステレオ又はモノ信号へダウンミックスされる。重要な要件は、ステレオ又はステレオダウンミックスが、十分な音声品質である、例えば、少なくともITU-R勧告BS.775-1ダウンミックスと同程度であるべきであることである。したがって、伝送される情報は、モノ又はステレオ信号の符号化されたもの及び空間パラメータを含む。モノ又はステレオダウンミックスは、元のマルチチャンネル音声信号を符号化するのに必要とされるよりも大幅に低いビットレートで符号化され、空間パラメー

50

タは、非常に小さい伝送帯域を必要とする。したがって、ダウンミックス及び空間パラメータは、全てのチャンネルが符号化される場合に必要とされるビットレートの一部のみである全体ビットレートで符号化され得る。パラメトリック復号器は、伝送されるモノ又はステレオダウンミックス及び空間パラメータから元のマルチチャンネル音声信号の高品質に近似するものを生成する。

【発明の開示】

【発明が解決しようとする課題】

【 0 0 0 6 】

本発明の目的は、高効率を提供し、高信号品質を提供し、且つ同時に後方互換性を有する符号化信号を提供するスケーラブルマルチチャンネル音声信号符号化器を提供することである。

10

【課題を解決するための手段】

【 0 0 0 7 】

第 1 態様によると、本発明は、マルチチャンネル音声信号を符号化するように構成される音声符号化器であって、第 1 及び第 2 音声信号の合成された表現である主要信号部分及び残余信号部分を発生させる符号化器合成モジュールであって、前記主要及び残余信号部分が、前記第 1 及び第 2 音声信号に数学的手順を適用することによって得られ、前記数学的手順が、前記第 1 及び第 2 音声信号の空間特性の記述を含む第 1 空間パラメータを含む、符号化器合成モジュールと、パラメータ発生器であって、第 2 空間パラメータを含む第 1 パラメータ群、及び第 3 空間パラメータを含む第 2 パラメータ群、を発生させるパラメータ発生器と、出力発生器であって、前記主要信号部分及び前記第 1 パラメータ群を含む第 1 出力部分、及び、前記残余信号部分及び前記第 2 パラメータ群を含む第 2 出力部分、を含む符号化出力信号を発生させる出力発生器と、を備える音声符号化器を提供する。

20

【 0 0 0 8 】

符号化器合成モジュールにおいて、第 1 及び第 2 音声信号は、主要及び残余部分に合成される。「主要及び残余信号部分」によって、2つの音声信号は、主要信号が、第 1 及び第 2 音声信号の主要部分すなわち主な部分を含む一方で、残余信号が、第 1 及び第 2 音声信号の残余すなわちあまり重要でない部分を含むことを理解される。「空間パラメータ」によって、数学的に表現されるとともに、信号対の 1 つ又は複数の空間特性に基づく又はから導出されるパラメータが理解される。計算され得る斯様な空間特性の非包括的なリストは、チャンネル間インテンシティ差異 ( I I D )、チャンネル間時間差異 ( I T D )、及びチャンネル間コーヒレント ( I C C ) である。前記符号化器合成モジュールは、主要及び残余信号部分を、これらの信号部分が第 1 及び第 2 音声信号よりあまり関連しないように、発生させることが好ましい。好ましくは、主要及び残余信号部分は、これらの信号が、関連しない、すなわち直交するように、又は少なくともこれら信号が、可能な限り関連しないようにあるべきように、発生される。

30

【 0 0 0 9 】

残余信号部分は、ビットストリームで表現され、よって非常に限られた量のビットレートのみを必要とするようにするために、出力ビットストリームに変換される前に、ローパスフィルタ処理され得る。このようなローパスフィルタ処理のカットオフ周波数は、500Hz から 10kHz の区間の間、例えば 2kHz であり得る。

40

【 0 0 1 0 】

符号化器合成モジュールは、上述されるように 2 つの音声信号を 1 つの主要信号に合成する代わりに、第 1、第 2、及び第 3 音声信号を、第 1 及び第 2 主要信号部分に合成するように構成され得る。

【 0 0 1 1 】

第 1 の態様による符号化器は、第 1 及び第 2 音声信号のスケーラブルな符号化表現を提供する。第 1 出力部分又はベースレイヤー部分を用いると、既存の符号化器を用いて、第 1 及び第 2 音声信号を、容認できる生じる音声品質に復号することが可能である。しかし、第 2 出力部分、すなわち洗練されたレイヤー (refined layer) 部分を活用することが可

50

能である復号器を用いることによって、より高い信号品質を得ることが可能である。したがって、第2出力部分は、任意追加的としてみられ得、可能な最良な音声品質が所望である場合にのみ必要である。

【0012】

好ましい実施例において、残余信号部分が、前記第1及び第2音声信号の間の差を含む。残余信号部分は、前記第1及び第2音声信号の間の差として正確に規定され得る。

【0013】

好ましい実施例において、数学的手順は、2次元信号空間における回転を含む。

【0014】

第3空間パラメータは、第2空間パラメータ及び第1空間パラメータ間の差を含み得る。第3空間パラメータは、差分符号化を含み得る。

10

【0015】

第2空間パラメータは、コーヒレントに基づくICCパラメータを含み得る。第3空間パラメータは、コーヒレントに基づくICCパラメータ及び相関に基づくICCパラメータ間の差を含み得る。好ましい実施例において、第2空間パラメータが、コーヒレントに基づくパラメータを含む一方で、第3空間パラメータが、第2空間パラメータと相関に基づくパラメータとの間の差を含む。

【0016】

前記符号化器は、第3、第4、第5、及び第6、又はより更なる音声信号を、これらの音声信号を前記第1及び第2音声信号とともに合成することによって、第1の態様の原理に従い符号化し、これらの信号に応答して前記第1及び第2出力部分を発生させるように更に構成され得る。好ましくは、斯様な符号化器は、複数の符号化合成モジュールを含む構成を用いることによって5.1音声信号を符号化するように構成される。原則的に、第1態様に従う符号化器の原理は、いかなるマルチチャンネルフォーマット音声データを符号化するのにも用いられ得る。

20

【0017】

第2態様において、本発明は、符号化信号に基づきマルチチャンネル音声信号を発生させる音声復号器であって、主要信号部分、残余信号部分、並びに第1及び第2空間パラメータ群に基づき、第1及び第2音声信号を発生させる復号器合成モジュールであって、前記空間パラメータが、前記第1及び第2音声信号の空間特性の記述を含み、前記残余信号及び第2空間パラメータが、前記第1および第2音声信号を発生させるのに用いられる混合マトリクスを決定するのに関与する、復号器合成モジュール、を備える音声復号器を提供する。

30

【0018】

第1態様と関連して説明されるように、既存の復号器は、主要信号部分及び第1空間パラメータのみを活用することによって、本発明に従う符号化器からの符号化出力信号を復号するように用いられ得る。しかし、第2態様に従う復号器は、符号化処理に含まれる符号化合成に対して一様に逆である混合マトリクスを決定するために、第2符号化出力部分、すなわち残余信号部分及び空間パラメータを活用することが可能であり得、したがって、第1及び第2音声信号の完全な再生成が獲得され得る。

40

【0019】

好ましい実施例において、前記復号器は、前記主要信号部分を受信する非相関化器であって、これに応答して非相関化された主要信号部分を発生させる非相関化器を含む。好ましくは、前記残余信号部分及び前記非相関化された主要信号部分の加算が、前記混合マトリクスを決定するのに関与される。前記復号器は、前記非相関化された主要信号部分を前記残余信号部分に加える前に、前記非相関化された主要信号部分を減衰する減衰器を含み得る。

【0020】

好ましい実施例において、前記混合マトリクスは、2次元信号空間における回転を、主要及び残余信号部分に適用する。

50

## 【 0 0 2 1 】

復号器は、パラメータの第 1 及び第 2 群の複数の群並びに複数の残余信号部分を受信し、これにより、これらの信号に応答して複数の群の第 1 及び第 2 音声信号を発生させるように構成され得る。好ましい実施例において、復号器は、パラメータの第 1 及び第 2 群の 3 つの群と 3 つの残余信号部分を受信し、これにより、これらの信号に応答して 3 つの群の第 1 及び第 2 音声信号を生成するように構成される。この実施例において、復号器は、5.1フォーマット又は他のマルチチャネルフォーマットになどに従い、6 つの独立した音声チャンネルを生成し得る。

## 【 0 0 2 2 】

好ましい実施例において、復号器は、復号器が 2 つより多い音声信号を表す符号化信号を復号するのを可能にするように、適切な構成で配置される複数の 1 対 2 チャネル混合マトリクスを含む。例えば、復号器は、6 つ音声信号を生成し、例えば符号化された 5.1 音声信号を復号するように構成される 5 つの混合マトリクスの構成を含み得る。

## 【 0 0 2 3 】

第 3 態様において、本発明は、マルチチャネル音声信号を符号化する方法であって、  
1) 第 1 及び第 2 音声信号の合成された表現である主要信号部分及び残余信号部分を発生させるステップであって、前記主要及び残余信号部分が、前記第 1 及び第 2 音声信号に数学的手順を適用することによって得られ、前記数学的手順が、前記第 1 及び第 2 音声信号の空間特性の記述を含む第 1 空間パラメータを含む、ステップと、  
2) 第 2 空間パラメータを含む第 1 パラメータ群を発生させるステップと、  
3) 第 3 空間パラメータを含む第 2 パラメータ群を発生させるステップと、  
4) 前記主要信号部分及び前記第 1 パラメータ群を含む第 1 出力部分、並びに前記残余信号部分及び前記第 2 パラメータ群を含む第 2 出力部分、を含む符号化出力信号を発生させるステップと、  
を含む方法を提供する。

## 【 0 0 2 4 】

第 1 態様に関連して説明されるのと同じの有利な点及びコメントは、第 3 態様に適用される。

## 【 0 0 2 5 】

第 4 態様において、本発明は、符号化信号に基づきマルチチャネル音声信号を発生させる方法であって、  
1) 主要信号部分、残余信号部分、並びに第 1 及び第 2 音声信号の空間特性の記述を含む第 1 及び第 2 空間パラメータ、を含む前記符号化信号を受信するステップと、  
2) 前記残余信号部分及び第 2 空間パラメータ群に基づき、混合マトリクスを決定するステップと、  
3) 前記決定された混合マトリクスに基づき、前記第 1 および第 2 音声信号を発生させるステップと、  
を含む方法を提供する。

## 【 0 0 2 6 】

前記方法は、前記主要信号部分を非相関化させ、これに応答して非相関化された主要信号部分を発生させるステップを含み得る。前記方法は、前記残余信号部分及び前記非相関化主要信号部分を加えるステップを更に含み得る。前記混合マトリクスを決定するステップが、前記加えられた残余信号部分及び前記非相関化主要信号部分に基づき得る。

## 【 0 0 2 7 】

好ましくは、前記方法は、パラメータの第 1 及び第 2 群の複数の群並びに複数の残余信号部分を受信し、これにより、これらの信号に応答して複数の群の第 1 及び第 2 音声信号を発生させるステップを含む。好ましい実施例において、前記方法は、パラメータの第 1 及び第 2 群の 3 つの群と 3 つの残余信号部分を受信し、これにより、これらの信号に応答して 3 つの群の第 1 及び第 2 音声信号を生成するステップを含む。この実施例において、前記方法は、5.1マルチチャネルフォーマット又は同等なものなどの 6 つの独立した音声

チャンネルを生成することが可能である。

【 0 0 2 8 】

第 2 態様に対して説明されるのと同じの有利な点及びコメントは、第 4 態様に適用される。

【 0 0 2 9 】

第 5 態様において、本発明は、主要信号部分と、第 1 及び第 2 音声信号の空間特性の記述を含む第 1 パラメータ群とを含む第 1 信号部分と、残余信号部分と、第 1 及び第 2 音声信号の空間特性の記述を含む第 2 パラメータ群とを含む第 2 信号部分と、を含む符号化マルチチャンネル音声信号を提供する。

【 0 0 3 0 】

第 5 態様に従う音声信号は、第 1 態様に関連して説明されるのと同じの有利な点を提供する。なぜなら、この信号は、第 1 態様に従う符号化器からの符号化出力信号と同一であるからである。したがって、第 5 態様に従う符号化マルチチャンネル音声信号は、スケーラブル信号である、なぜなら、ベースレイヤーに関して適合された第 1 信号部分は必須である一方で、洗練されたレイヤーに関して適合された第 2 信号部分は選択的であり、選択的な信号品質に関してのみ必要とされるからである。

【 0 0 3 1 】

第 6 態様において、本発明は、第 5 態様に記載の信号を記憶させた記憶媒体を提供する。記憶媒体は、ハードディスク、フロッピー（登録商標）ディスク、CD、DVD、SDカード、メモリースティック、及びメモリチップなどであり得る。

【 0 0 3 2 】

第 7 態様において、本発明は、第 1 態様に記載の方法を実行するように構成される計算機実行可能プログラムコードを提供する。

【 0 0 3 3 】

第 8 態様において、本発明は、第 7 態様に記載の計算機実行可能プログラムコードを含む計算機読み取り可能記憶媒体を提供する。記憶媒体は、ハードディスク、フロッピー（登録商標）ディスク、CD、DVD、SDカード、メモリースティック、及びメモリチップなどであり得る。

【 0 0 3 4 】

第 9 態様において、本発明は、第 4 態様に記載の方法を実行するように構成される計算機実行可能プログラムコードを提供する。

【 0 0 3 5 】

第 10 態様において、本発明は、第 9 態様に記載の計算機実行可能プログラムコードを含む計算機読み取り可能記憶媒体を提供する。記憶媒体は、ハードディスク、フロッピー（登録商標）ディスク、CD、DVD、SDカード、メモリースティック、及びメモリチップなどであり得る。

【 0 0 3 6 】

第 11 態様において、本発明は、第 1 の態様に従う符号化器を含む装置を提供する。前記装置は、例えばサラウンド音声アンプ、サラウンド音声レシーバ、及びDVDプレーヤ/レコーダなどのホームエンターテインメントオーディオ機器などでありえる。原理的に、前記装置は、例えば5.1フォーマットなどのマルチチャンネル音声データを扱うことが可能であるいかなる音声装置でもあり得る。

【 0 0 3 7 】

第 12 態様において、本発明は、第 2 の態様に従う復号器を含む装置を提供する。前記装置は、例えばサラウンド音声アンプ、サラウンド音声レシーバ、A/Vレシーバ、セットトップボックス、及びDVDプレーヤ/レコーダなどのホームエンターテインメントオーディオ機器などでありえる。

【 0 0 3 8 】

第 5 態様に従う信号は、伝送チェーンを介した伝送に適している。このような伝送チェーンは、前記信号を記憶するサーバ、前記信号の配信用のネットワーク、及び前記信号を

10

20

30

40

50



受信するクライアントを含み得る。クライアント側は、例えば、コンピュータ、A/Vレシーバ、セットトップボックス、等のハードウェアを有し得る。したがって、第5態様に従う信号は、デジタルビデオ放送、デジタルオーディオ放送、又はインターネットラジオ等の伝送に適している。

【0039】

上述の態様の全ては、第1及び第2音声信号が、全帯域幅信号であり得ることを理解され得る。任意追加的に、第1及び第2音声信号は、それぞれの全帯域幅音声信号のサブバンド表現を表す。言い換えれば、本発明に従う信号処理は、全帯域幅信号に適用され得る、又はサブバンド毎に適用され得る。

【0040】

以下において、本発明は、添付の図面を参照にしてより詳細に説明される。

【0041】

本発明が様々な変形態様及び代替形式に従い得る一方で、特定の実施例が、例として図面に示されており、以下に詳細に説明される。しかし、本発明が開示される特定の形式に制限されるようには意図されないことを理解されるべきである。むしろ、本発明は、添付の請求項によって規定される本発明の精神及び範囲に含まれる全ての変形態様、同等な物、代替態様を含むこととされる。

【発明を実施するための最良の形態】

【0042】

図1は、独立した音声信号を受信する5つのスピーカC、Lf、Ls、Rf及びRsの中央に位置される聴取人物LPを含む、通常の5.1マルチチャンネル音声設定の概略図を示す。これらは、聴取人物LPに空間的な音声の印象を生じさせるために提供される。加えて、5.1セットアップは、個別のサブウーハーLFE信号を提供する。したがって、このようなマルチチャンネルセットアップに関する完全な信号表現は、合計6つの独立した音声チャンネルを必要とし、つまり、大きなビットレートが、このようなシステムに関する音声信号を完全な音声品質で表現するために必要である。以下に、低ビットレートで、5.1システムにおいて高音声品質を提供することが可能である、本発明の実施例が説明される。

【0043】

図2は、本発明に従う2-1符号化器合成ユニットEUを示す。第1及び第2音声信号x1、x2は、符号化器合成モジュールECMに入力され、符号化器合成モジュールECMにおいて、第1及び第2音声信号x1、x2を合成するとともに主要信号部分m及び残余部分sを含む第1及び第2音声信号x1、x2のパラメトリック表現を生成するために、好ましくは信号回転を含む数学的手順が、第1及び第2音声信号x1、x2に実行される。第1空間パラメータSP1、すなわち第1及び第2音声信号x1、x2の空間的信号特性を記述するパラメータは、数学的符号化合成手順に含まれる。

【0044】

パラメータ発生器PGは、第1及び第2音声信号x1、x2に基づき、第1及び第2パラメータ群PS1、PS2を生成する。第1パラメータ群PS1は、第2空間パラメータSP1を含み、第2パラメータ群PS2は、第3空間パラメータSP3を含む。符号化出力信号は、主要信号部分m及び第1パラメータ群PS1を含む第1出力部分OP1を含む一方で、第2出力部分OP2は、残余信号部分s及び第2パラメータ群PS2を含む。

【0045】

第1空間パラメータSP1に対しての、第2及び第3空間パラメータSP2及びSP3の適切な選択により、復号器側において、符号化器合成又は回転手順の逆順を実行することが可能であり、よって、第1及び第2音声信号x1、x2は、透過的に復号され得る。

【0046】

好ましくは、符号化器は、第1出力部分を、出力ビットストリームのベースレイヤーに加える一方で、第2出力部分は、出力ビットストリームの洗練レイヤーに加えられる。復号ステップにおいて、低減された信号品質が許容される場合、ベースレイヤーのみを使用することが可能である一方で、可能な最良な信号品質は、洗練レイヤーも復号処理に含ま

10

20

30

40

50

れる場合に達成され得る。

【 0 0 4 7 】

説明される符号化原理は、完全な後方互換性を有するスケーラブルハイブリッドマルチチャネル音声符号化器を提供する。復号器は、以下の状況に関して使用され得る： 1 ) 復号モノ又はステレオ信号のみ、 2 ) 残余信号を使用しない復号マルチチャネル出力、及び 3 ) 残余信号を使用する復号マルチチャネル出力。

【 0 0 4 8 】

以下において、符号化器合成モジュール及び空間パラメータの好ましい実施例が説明される。好ましい符号化器合成モジュールは、

【 数 1 】

$$(Eq\ 1) \quad \begin{pmatrix} m[k] \\ s[k] \end{pmatrix} = \begin{pmatrix} sc_{corr} & sc_{corr} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1[k] \\ x_2[k] \end{pmatrix}, \text{ where}$$

$$sc_{corr} = \min \left\{ sc_{corr, \max}, \frac{1}{c_l \cos(\chi + \beta) + c_r \cos(-\chi + \beta)} \right\}$$

に従い、回転された信号の合計の振幅を最大化することによって、第 1 及び第 2 音声信号  $x_1$ 、 $x_2$  を、主要信号部分  $m$  及び残余信号部分  $s$  に合成する。

【 0 0 4 9 】

$sc_{corr}$  に含まれる振幅回転係数は、 $ICC$  及び  $IID$  から導出され、すなわち、これらは、第 1 及び第 2 音声信号  $x_1$ 、 $x_2$  の空間特性に基づく。これらの振幅回転係数は、好ましくは、

$$\chi = \frac{1}{2} \cos^{-1}(ICC), \beta = \tan^{-1} \left( \tan(\chi) \frac{c_r - c_l}{c_r + c_l} \right), c_l = \sqrt{\frac{IID}{1 + IID}}, c_r = \sqrt{\frac{1}{1 + IID}}$$

に従い計算される。

【 0 0 5 0 】

残余信号  $s$  は、 $x_1$  及び  $x_2$  の間の差であるように選択される。このマトリクスは、 $sc_{corr}$  が 0 であり得ないので、常に可逆であり、つまり、このことは、 $sc_{corr}$  が既知である限りは、完全な再構築が達成され得ることを意味することを注意されたい。定数  $sc_{corr, \max}$  を取得するのに適した値は、1.2 である。

【 0 0 5 1 】

復号器において  $sc_{corr}$  を導出するために、第 2 パラメータ群  $PS2$  は、コーヒレント及び相関パラメータの間の差を有することが好ましく、したがって、スケーラブルなビットストリームにおいて洗練レイヤーにおいて対応する残余信号  $s$  と一緒に伝送される。第 1 パラメータ群  $PS1$  は、コーヒレントパラメータ又は相関パラメータの一方を含み、主要信号部分  $m$  と一緒にベースレイヤーにおいて伝送されるように選択される。

【 0 0 5 2 】

残余信号  $s$  が復号器に対して入手可能である場合、 $sc_{corr}$  の計算を促進する相関パラメータが導出され、数式 1 の混合マトリクスの逆の関数：

10

20

30

40

$$\begin{pmatrix} x_1[k] \\ x_2[k] \end{pmatrix} = \begin{pmatrix} \frac{1}{2sc_{corr}} & 1 \\ \frac{1}{2sc_{corr}} & -1 \end{pmatrix} \begin{pmatrix} m[k] \\ s[k] \end{pmatrix}$$

10

が決定される。

【 0 0 5 3 】

別の好ましい実施例において、符号化器合成モジュールは、主成分分析（PCA）に基づくものであり、第1及び第2音声信号 $x_1$ 、 $x_2$ を、

$$\begin{pmatrix} m[k] \\ s[k] \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} x_1[k] \\ x_2[k] \end{pmatrix}$$

20

に従い混合し、ここで、好ましい係数は、

$$\alpha = \frac{1}{2} \tan^{-1} \left( \frac{2ICC \cdot c}{c^2 - 1} \right), c = 10^{\frac{IID}{20}}$$

30

に従いICC及びIIDに基づく。

【 0 0 5 4 】

洗練レイヤーに含まれるべき第2パラメータ群PS2の符号化に関する好ましい選択肢は、

1) ベースレイヤーにおけるコーヒレントパラメータとは独立した相関パラメータの時間又は周波数差分符号化、

40

2) ベースレイヤーにおけるコーヒレントパラメータに関する相関パラメータの差分符号化（例えば、 $ICC = ICC_{correlation} - ICC_{coherence}$ ）、

ここで、1及び2の合成は、どちらが最低ビット量を必要とするかに依存する、

3) 図3及び4は、符号化されたモノ信号への符号化合成に基づく5.1フォーマット符号化器及び対応する5.1復号器それぞれの好ましい機器構成を例示する。図5及び6は、符号化ステレオ信号への符号化合成に基づく、代替的な5.1フォーマット符号化器及び対応する復号器のそれぞれを示す。

【 0 0 5 5 】

図3は、6つの独立した音声信号lf、ls、rf、rs、co、lfeのモノ信号mへの合成に基づく符号化構成を示し、例えば6つの音声信号は、5.1フォーマットにおけるにおける信号l

50

f、ls、rf、rs、co、lfeを表す。符号化器は、前述されるような5つの符号化合成ユニットEUを含み、これらのユニットEUは、6つの信号lf、ls、rf、rs、co、lfeを単一のモノ信号mへ連続的に合成するように構成される。初期セグメント化及び変換ステップSTは、符号化合成の前に信号対に対して実行される。このステップSTは、時間-ドメイン音声信号を重複するセグメントにセグメント化するステップと、その後、これらの重複する時間-ドメインセグメントを（大文字で示される）周波数ドメイン表現に変換するステップと、を含む。

#### 【0056】

セグメント化及び変換ステップSTの後で、2つの左チャンネルLf及びLsは、主要信号部分L、第1及び第2パラメータ群PS1a、PS1b、及び残余信号ResLに合成される。2つの右のチャンネルRf及びRsは、主要信号部分R、第1及び第2パラメータ群PS2a、PS2b、並びに残余信号ResRに合成される。生じる主要信号部分L及びRは、その後、主要信号部分LR、残余信号部分ResLR、並びに第1及び第2パラメータPS4a及びPS4bに合成される。中央信号C0及びサブウーハーチャンネルLFEは、主要信号部分C、第1及び第2パラメータPS3a、PS3b、並びに残余信号ResCに合成される。最終的に、主要信号部分C及びLRは、主要信号部分M、残余信号部分ResM、並びに第1及び第2パラメータPS5a、PS5bに合成される。

#### 【0057】

好ましくは、第1及び第2のパラメータの群PS1a-PS5a、PS1b-PS5bは、量子化、符号化、及び伝送の前に、セグメントにおけるある数の周波数帯域（サブバンド）に関して独立して決定されるが、好ましくは、この処理は、全部の帯域幅信号に実行され得る。信号分析及び処理が適用された後で、選択的な処理が適用され（ITOLA）、セグメントは、時間ドメインに変換されて戻され得IT、またセグメントは、時間-ドメインモノ音声信号mを得るために、重複及び加えられ得OLA。全体として、符号化器は、主要信号部分m及び5つのパラメータ群PS1a-PS5aを含む第1出力部分と、5つの残余信号部分ResL、ResR、ResLR、ResM、ResC及び5つのパラメータ群PS1b-PS5bを含む第2出力部分を生成する。

#### 【0058】

図4は、図3の符号化器に対応する復号器を示し、すなわち、この復号器は、図3の符号化器からの出力信号を受信するように適合される。復号器は、本質的に、図3に記載の処理の逆工程を適用する。復号器は、（選択的な）処理セグメント化を含み、周波数変換STは、主要信号部分mに適用される。復号器は、5つの同様な復号器合成ユニットDUを含み、これらのうちの1つは、点線で示されている。復号器合成ユニットDUは、主要信号部分に基づく第1及び第2信号を生成する混合マトリクスMMを含む。混合マトリクスMM、すなわち符号化器合成モジュールECMにおいて適用された混合マトリクスの逆処理は、受信される主要信号部分、残余部分、並びに第1及び第2パラメータ群に基づき決定される。

#### 【0059】

図4に示される第1復号合成ユニットDUにおいて、主要信号Mは、始めに、非相関化器Decにおいて非相関化され、その後、減衰気Attにおいて減衰される。非相関化及び減衰された主要信号部分は、残余信号部分ResMに加えられる。この加えられた信号は、その後、混合マトリクスMMを決定するのに用いられる。減衰器Attは、残余信号部分ResM及び第1パラメータ群PS5aに応答して設定される。最終的に、混合マトリクスMMは、第1及び第2パラメータ群PS5a、PS5bを用いて決定される。決定された混合マトリクスMMは、その後、主要信号部分Mを、第1出力信号LR及び第2出力信号Cへ合成する。これらの第1及び第2出力信号LR、Cは、その後、それぞれの符号化器合成ユニットへ供給され、連続的にL、R、及びC0、LFEのそれぞれへと合成される。最終的に、Lは、Lf及びLrを生成するように復号器合成される一方で、Rは、Rf及びRrを生成するように復号器合成される。信号分析及び処理が適用された後で、セグメントは、時間ドメインに変換されて戻されIT、またセグメントは、時間ドメイン表現lf、lr、rf、rr、co、lfeを得るために、重複及び加えられるOLA。この逆変換及び重複-加算IT、OLAは、任意選択的である。

10

20

30

40

50

## 【 0 0 6 0 】

図 5 は、符号化器実施例を示し、図 3 の符号化器に関連して説明された原理に従いそれぞれ機能する 3 つの符号化器合成ユニットが、3 つの主要信号部分 L、R、C の対にある 6 つの音声信号部分 Lf、Lr、Rf、Rr、C0、LFE を、関連する第 1 パラメータ群 PS1a-PS3a、第 2 パラメータ群 PS1b-PS3b、及び残余信号部分 ResL、ResR、ResC と合成するように用いられる。3-2 符号化器合成ユニットは、この場合、3 つの主要信号部分 L、R 及び C に適用され、2 つの主要信号部分 L0、R0、残余信号部分 ResEo、及びパラメータ群 PS4 を生じさせる。任意追加的に、初期セグメント化及び周波数ドメイン変換 S T が適用され、図 3 にも関連して説明されるような、最終的な逆変換 IT 及び重複 - 加算 OLA が（選択的に）適用される。

10

## 【 0 0 6 1 】

図 6 は、図 5 の符号化器からの出力を復号するように構成される復号器の構成を示す。入力信号 lo、ro の（選択的な）初期セグメント化及び周波数ドメイン変換 S T の後で、2-3 復号器合成モジュールは、パラメータ群 PS4 と一緒に、主要信号部分 Lo、Ro、残余信号部分 ResEo に応答して主要信号部分 L、R、C を生成する。これらの主要信号部分 L、R、C は、この場合、図 4 の復号器と関連して説明される復号器合成ユニット D U に類似なそれぞれの復号器合成ユニットにおいて処理される。最終的な逆変換 IT 及び重複 - 加算 OLA は、（任意選択的に）上述のようにも適用される。

## 【 0 0 6 2 】

図 7 は、5 人の訓練された聴取者に関して実行された聴取テストの結果を例示する。用いられた音楽項目 A - K は、M P E G 「空間音声符号化」の作品項目において特定されるものである。各項目 A - K に関して、1）残余無しの復号器 - 左に示される、2）残余ありの空間符号化器、すなわち本発明による復号器 - 中央に示される、及び 3）参照（隠し） - 右に示される、3 つの符号化バージョンに関する結果が検査において含まれていた。項目 A - K の全体の平均は、T O T として示される。各符号化されたバージョンに関して、平均等級 G R D がアスタリスク(\*)を用いて示される一方で、聴取者における返答に関する +/- 標準偏差がここで示される。

20

## 【 0 0 6 3 】

2）及び 3）の状況に関して、図 5 及び図 6 に例示される符号化器 / 復号器の原理が用いられる。状況 2）において、残余信号部分は破棄されていた。状況 3）に関して、2kHz に帯域制限されている 3 つの残余信号部分、すなわち左チャンネル ResL に関する残余信号部分、右チャンネル ResR に関する残余信号部分、及び復号器合成モジュール 3-2 に関する残余信号部分 ResEo が使用されていた。残余信号部分のそれぞれ ResL、ResR、ResEo は、8kbit/s で符号化されており、（相関（洗練レイヤー）とコーヒレント（ベースレイヤー）との間の差である）余分な空間パラメータは、700kbit/s の推定ビットレートを必要としている。したがって、全体的な余分な残余に関連するビットレートは、この場合、25kbit/s である。（ベースレイヤーに配置されるべき）標準空間パラメータは、推定 10kbit/s を必要としていた。全体空間データレートは、この場合、約 35kbit/s である。何の核となるコーデックも、ステレオ信号 lo、ro には適用されていなかった。

30

## 【 0 0 6 4 】

この結果から、低ビットレートで符号化される 3 つの残余信号を活用することによって、大きな品質向上が達成され得ることが明らかである。更に、全体平均品質等級は、+/- 92 であり、「透過性」音声品質と考慮されるものに非常に近い。

40

## 【 0 0 6 5 】

本発明による符号化器及び復号器は、デジタルビデオ放送（D V B）、デジタルオーディオ放送（D A B）、インターネットラジオ、電子音楽配信を含むマルチチャンネル音声符号化を含む全てのアプリケーションにおいて適用され得る。

## 【 0 0 6 6 】

請求項における参照符号は、可読性を向上させるために作用するに過ぎない。これらの参照符号は、請求項の範囲を制限するようにどのようにしても解釈されるべきでなく、例

50

のみを例示するのに含まれるのみである。

【図面の簡単な説明】

【0067】

【図1】図1は、5.1マルチチャンネルスピーカセットアップの概略図を示す。

【図2】図2は、本発明による符号化器合成ユニットを示す。

【図3】図3は、符号化器合成に基づき5.1音声信号をモノ信号へ符号化する好ましい符号化器を示す。

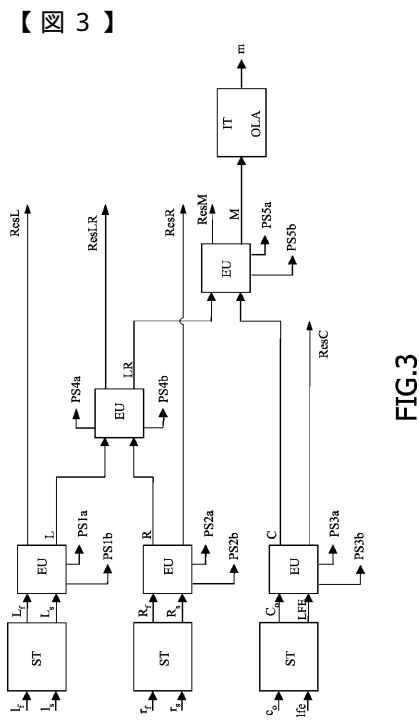
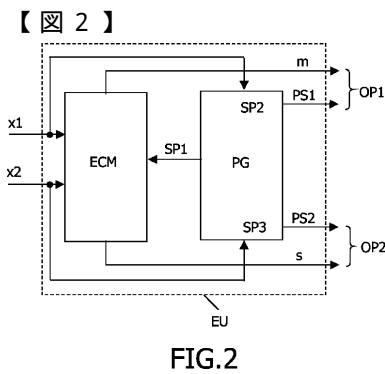
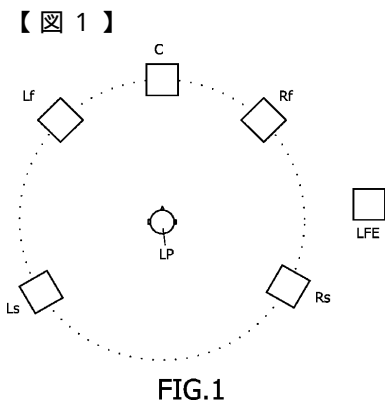
【図4】図4は、図3の符号化器に対応する好ましい復号器を示す。

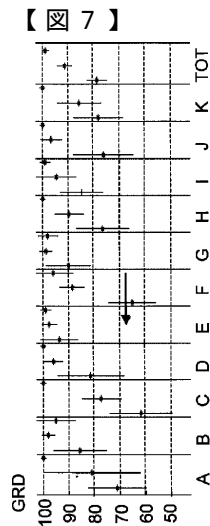
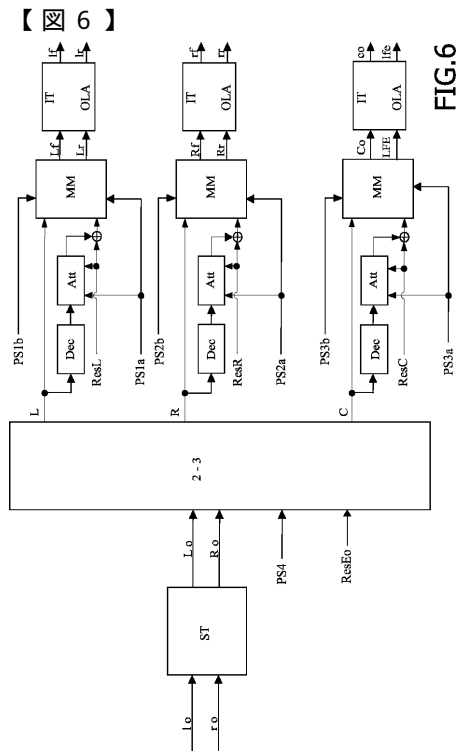
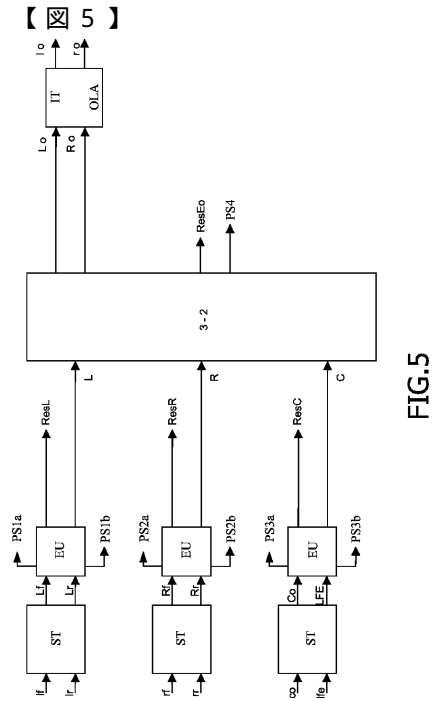
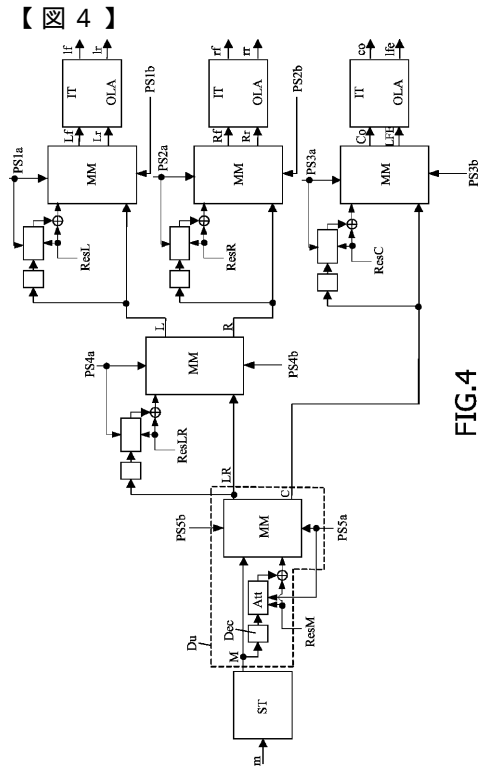
【図5】図5は、符号化器合成に基づき5.1音声信号をステレオ信号へ符号化する好ましい符号化器を示す。

10

【図6】図6は、図5の符号化器に対応する好ましい復号器を示す。

【図7】図7は、本発明による符号化原理を用いて実施される聴取テストの結果を例示するグラフを示す。





---

フロントページの続き

(72)発明者 スフエイエルス エリク ジー ピー  
オランダ国 5 6 5 6 アーアー アインドーフエン プロフ ホルストラーン 6

審査官 井上 健一

(56)参考文献 国際公開第 0 1 / 0 4 3 5 0 3 ( W O , A 1 )  
特開 2 0 0 1 - 2 5 5 8 9 2 ( J P , A )  
特表 2 0 0 0 - 5 0 5 2 6 6 ( J P , A )  
特開平 0 7 - 0 2 3 0 0 9 ( J P , A )  
国際公開第 2 0 0 6 / 0 4 8 8 1 5 ( W O , A 1 )  
国際公開第 2 0 0 5 / 1 0 1 3 7 1 ( W O , A 1 )  
特開 2 0 0 4 - 1 7 3 2 5 0 ( J P , A )

(58)調査した分野(Int.Cl. , D B 名)  
G10L 19/00