

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6400739号
(P6400739)

(45) 発行日 平成30年10月3日 (2018. 10. 3)

(24) 登録日 平成30年9月14日 (2018. 9. 14)

(51) Int. Cl.

F I

G 1 O L 19/008 (2013. 01)

G 1 O L 19/008 2 O O

G 1 O L 19/00 (2013. 01)

G 1 O L 19/00 3 3 O B

請求項の数 13 (全 15 頁)

(21) 出願番号	特願2016-570945 (P2016-570945)	(73) 特許権者	503260918
(86) (22) 出願日	平成27年2月26日 (2015. 2. 26)		アップル インコーポレイテッド
(65) 公表番号	特表2017-509932 (P2017-509932A)		Apple Inc.
(43) 公表日	平成29年4月6日 (2017. 4. 6)		アメリカ合衆国 95014 カリフォル
(86) 国際出願番号	PCT/US2015/017830		ニア州 クパチーノ アップル パーク
(87) 国際公開番号	W02015/148046		ウェイ ワン
(87) 国際公開日	平成27年10月1日 (2015. 10. 1)		One Apple Park Way,
審査請求日	平成28年8月23日 (2016. 8. 23)		Cupertino, Californ
(31) 優先権主張番号	61/970, 284		ia 95014, U. S. A.
(32) 優先日	平成26年3月25日 (2014. 3. 25)	(74) 代理人	100094569
(33) 優先権主張国	米国 (US)		弁理士 田中 伸一郎
(31) 優先権主張番号	14/613, 203	(74) 代理人	100088694
(32) 優先日	平成27年2月3日 (2015. 2. 3)		弁理士 弟子丸 健
(33) 優先権主張国	米国 (US)	(74) 代理人	100067013
前置審査			弁理士 大塚 文昭

最終頁に続く

(54) 【発明の名称】 ダッキング制御のためのメタデータ

(57) 【特許請求の範囲】

【請求項 1】

ビデオコンテンツに関連するオーディオコンテンツをエンコードするための方法であって、

(i) 第1のチャンネルグループ、オブジェクト、又は前記ビデオコンテンツにおいて起きているアクションを記述するための視認性の記述的オーディオコンテンツを含むステム、(ii) 第2のチャンネルグループ、オブジェクト、又はダイアログ、音楽及び効果の少なくとも1つを含むステム、(iii) 第3のチャンネルグループ、オブジェクト、又はダイアログ、音楽及び効果の別の1つを少なくとも含むステム、を有するサウンドプログラムコンテンツ片を受信すること、

前記サウンドプログラムコンテンツ片内の前記第1のチャンネルグループ、オブジェクト、又はステムが前記サウンドプログラムコンテンツ片の再生中に強調されるべきかを決定することと、

(i) 前記第2のチャンネルグループ、オブジェクト又はステムに適用すべきダッキング値の第1のセット、及び(ii) 前記第3のチャンネルグループ、オブジェクト又はステムに適用すべきダッキング値の第2のセットは、前記第1のチャンネルグループ、オブジェクト又はステム内の活性が検出されたとき各ダッキング値が決定され、前記ダッキング値の第1のセットは、前記ダッキング値の第2のセットとは異なり、前記ダッキング値の第1のセット及び前記ダッキング値の第2のセットが前記第2及び第3のチャンネルグループ、オブジェクト、又はステムにそれぞれ適用されるとき、前記第2及び第3のチャンネルグルー

ブ、オブジェクト、又はステムは、前記サウンドプログラムコンテンツの再生中に強調されず、前記ダッキング値の第1のセット及び第2のセットを音声アセット内の前記第1のチャンネルグループ、オブジェクト、又はステムに関連付けることであって、前記音声アセットは、(i)前記第1、第2及び第3のチャンネルグループ、オブジェクト又はステム、(ii)前記第1のチャンネルグループ、オブジェクト又はステムに関連付けられた前記ダッキング値の第1のセット及び第2のセットを含む、方法。

【請求項2】

前記ダッキング値の第1及び第2のセットが、前記第1のチャンネルグループ、オブジェクト、又はステム内における高活性の期間中には、前記第1のチャンネルグループ、オブジェクト、又はステム内における低活性の期間中よりも、前記第2及び第3のチャンネルグループ、オブジェクト、又はステムのダッキングを多く提供するように、前記ダッキング値の第1及び第2のセットは、前記第1のチャンネルグループ、オブジェクト、又はステム内における前記活性に基づいて時間とともに変化する、請求項1に記載の方法。

【請求項3】

前記第2のチャンネルグループ、オブジェクト又はステムがダイアログを含み、且つ前記第3のチャンネルグループ、オブジェクト又はステムが音楽及び効果を含むとき、前記ダッキング値の第1及び第2のセットが生成され、前記サウンドプログラムコンテンツの再生中に前記第2のチャンネルグループ、オブジェクト、又はステムは、前記第3のチャンネルグループ、オブジェクト、又はステムよりも強調されないようにする、請求項1に記載の方法。

【請求項4】

前記ダッキング値の第1のセット及び第2のセットが、再生中に前記第2及び第3のチャンネルグループ、オブジェクト又はステムのダイナミックレンジを縮小させる、請求項1に記載の方法。

【請求項5】

前記ダッキング値の第1のセットが、再生中に前記第2のチャンネルグループ、オブジェクト、又はステムを音場内で移動させる、請求項1に記載の方法。

【請求項6】

前記ダッキング値の第1のセットが、前記第2のチャンネルグループ、オブジェクト、又はステムに適用されるときに、前記第2のチャンネルグループ、オブジェクト、又はステムに関連付けられた音響を減衰させる、0デシベル以下のゲイン値である、請求項1に記載の方法。

【請求項7】

前記活性は発語（スピーチ）活性を含み、

前記第2のチャンネルグループ、オブジェクト、又はステムの移動は、前記第1のチャンネルグループ、オブジェクト、又はステム内に発語活性があるときの限られる、請求項5に記載の方法。

【請求項8】

ビデオコンテンツに関連するオーディオコンテンツを再生するための方法であって、

(i)第1のチャンネルグループ、オブジェクト、又は前記ビデオコンテンツにおいて起きているアクションを記述するための視認性の記述的オーディオコンテンツを含むステム、(ii)第2のチャンネルグループ、オブジェクト、又はダイアログ、音楽及び効果の少なくとも1つを含むステム、

(iii)第3のチャンネルグループ、オブジェクト、又は(ii)のダイアログ、音楽及び効果とは異なる別のダイアログ、音楽及び効果の少なくとも1つを含むステム、

(iv)ダッキング値の第1のセット、及び前記第1のチャンネルグループ、オブジェクト、又はステムに関連するが当該ダッキング値の第1のセットとは異なるダッキング値の第2のセットであって、前記第1のチャンネルグループ、オブジェクト、又はステム内の活性が検出されたとき各ダッキング値が決定された前記ダッキング値の第1及び第2のセット、を含むサウンドプログラムコンテンツを表す音声アセットを受信することと、

前記ダッキング値の第 1 及び第 2 のセットを、(i)前記第 1 のチャンネルグループ、オブジェクト、又はステム、(ii)前記第 2 のチャンネルグループ、オブジェクト、又はステム、及び(iii)前記第 3 のチャンネルグループ、オブジェクト、又はステムとともに前記音声アセットから抽出することと、

複数のラウドスピーカを介した前記サウンドプログラムコンテンツの再生中に、

前記ダッキング値の第 1 のセットを、前記第 2 のチャンネルグループ、オブジェクト、又はステムに適用することと、

前記ダッキング値の第 2 のセットを、前記第 3 のチャンネルグループ、オブジェクト、又はステムに適用することと、を含み、

前記ダッキング値の第 1 及び第 2 のセットを適用することは、前記第 3 のチャンネルグループ、オブジェクト、又はステムとは異なる前記第 2 のチャンネルグループ、オブジェクト、又はステムを強調しないことである、方法。

10

【請求項 9】

再生中に、(i)前記第 2 のチャンネルグループ、オブジェクト、又はステム、及び(ii)前記第 3 のチャンネルグループ、オブジェクト、又はステムのダイナミックレンジを縮小することにより、前記ダッキング値の第 1 及び第 2 のセットの適用が強調されない、請求項 7 に記載の方法。

【請求項 10】

前記ダッキング値の第 1 及び第 2 のセットの適用前に、スケール係数を前記ダッキング値の第 1 及び第 2 のセットに適用することを更に含む、請求項 8 に記載の方法。

20

【請求項 11】

前記複数のラウドスピーカを駆動して音場内に音を反響させるために、前記第 1、第 2 及び第 3 のチャンネルグループ、オブジェクト、又はステムに基づきドライブ信号のセットを生成することを更に含む、

前記ダッキング値の第 1 のセットの適用は、前記音場内における第 2 のチャンネルグループ、オブジェクト、又はステムの反響場所を、前記音場内の異なる場所に再生中に移動させることを生じさせる、請求項 8 に記載の方法。

【請求項 12】

音声装置内のプロセッサにより、請求項 1 ~ 11 の何れか 1 項に記載の方法を実行させる命令を記憶した非一時的コンピュータ読取り可能な記録媒体。

30

【請求項 13】

サウンドプログラムコンテンツを再生するための音声装置であって、

ハードウェアプロセッサと、

前記ハードウェアプロセッサによって、請求項 1 ~ 11 の何れか 1 項に記載の方法を実行させる命令を記憶したメモリユニット、

を備えた音声装置。

【発明の詳細な説明】

【技術分野】

【0001】

【関連事項】

40

本出願は、2014年3月25日に出願された、米国特許仮出願第 61 / 970 , 284 号の、先の出願日の利益を主張する。

【0002】

サウンドプログラムコンテンツを表す音声アセット内にダッキング値を追加するためのシステム及び方法が記載される。ダッキング値は特定のチャンネル/チャンネルグループに関連付けられてもよく、関連付けられたチャンネル/チャンネルグループを除くサウンドプログラムコンテンツ内の他の全てのチャンネル/チャンネルグループをダッキングするために用いられてもよい。他の実施形態もまた記載されている。

【背景技術】

【0003】

50

音声チャンネルは聞き手によって「オン」及び「オフ」に選択的に切り替えられてもよい。例えば、映画又はテレビ番組は、ユーザによって選択されてもよいナレーションチャンネルを含んでもよい。ナレーションチャンネルは、対応する映像内で起こっている視覚的動作及び活動の音声描写を提供してもよい。したがって、このナレーションチャンネルは、視覚障害を持ったユーザにとって有益になり得る。

【0004】

ナレーションは、視覚障害を持ったユーザが、映像内で起こっている視覚的動作及び活動をより良く理解することを可能にし得るが、このナレーションチャンネルを他の音声チャンネルにかぶせて直接再生すると、これらのチャンネルの全体的な明瞭性を損なう場合がある。例えば、メインのダイアログチャンネルは、これらのチャンネルが両方とも同時に再生されると、ナレーションチャンネルに干渉し得る。

10

【0005】

本節で説明された取り組み方は、遂行することができたであろう取り組み方ではあるが、必ずしも以前に考案され又は遂行されたことのある取り組み方とは限らない。したがって、特に断りのない限り、本節で説明された取り組み方のいずれも、ただ単に本節に含まれているからというだけで、先行技術であるとみなすべきではない。

【発明の概要】

【0006】

本明細書においては、音声エンコード装置及び音声デコード装置が記載される。音声エンコード装置は、サウンドプログラムコンテンツ片に対応する音声チャンネルのセットを調べ、チャンネル又はチャンネルグループのうちの1つに関連付けるべきダッキング値のセットを作成してもよい。例えば、音声チャンネルは、視覚障害を持ったユーザが、映画又はテレビ放送内で起こっている動作を理解することを支援するナレーションチャンネルを含んでもよい。例えば、ナレーションチャンネルは、映像内の人物の移動、又は視覚障害を持ったユーザには明らかでないであろう他の動作の音声描写を含んでもよい。ダッキング値は、エンコードされたサウンドプログラムコンテンツ片を表す音声アセット内のこのナレーションチャンネルに関連付けられてもよい。

20

【0007】

音声アセットは、ラウドスピーカのセットを通した再生のために1つ以上の音声デコード装置へ転送されてもよい。一実施形態では、ダッキング値は、再生中に、サウンドプログラムコンテンツ片内のナレーションチャンネルとは別の他の全てのチャンネルのダイナミックレンジを縮小してもよい。その結果、ナレーションチャンネルは、再生されている他のチャンネルのレンジの縮小を通じて強調されてもよい。一実施形態では、ダッキングが、ナレーションチャンネル内で活性が検出される期間中に実行され、不活性の期間中に実行されないように（例えば、発声/発語活性）、ダッキング値は時間とともに変化してもよい。

30

【0008】

いくつかの実施形態では、ダッキングは音声チャンネル/チャンネルグループに別個に適用されてもよい。例えば、ダイアログチャンネルは音楽/効果チャンネルよりも大きくダッキングされてもよい。この可変性のおかげで、ナレーションチャンネルの明瞭性にさほどひどく影響を与えないと思われる他のチャンネルをダッキングすることを必要とせずに、ナレーションチャンネルの明瞭性を改善することが可能になる。

40

【0009】

いくつかの実施形態では、ダッキング値はまた、ダッキング値に関連付けられていない1つ以上のチャンネルを、音場内の異なる位置において発音させてもよい。例えば、ダッキング値は、再生中にチャンネルをサラウンドラウドスピーカによって発音させてもよい。チャンネルのダイナミックレンジをダッキング及び縮小することに加えて、発音位置のこの変更は別のチャンネル（例えば、ナレーションチャンネル）の明瞭性を改善し得る。

【0010】

上述されたように、ダッキング値は、エンコードされたサウンドプログラムコンテンツ片内に含まれ、特定のチャンネル/チャンネルグループに関連付けられてもよい。エンコード

50

されたサウンドプログラムコンテンツの再生中に、ダッキング値は、ダッキング値に関連付けられていない他の全てのチャンネル／チャンネルグループに適用されてもよい。このダッキングは、ダッキングされないチャンネル／チャンネルグループ内の音声の明瞭性を改善し得る。例えば、上述されたように、ナレーションチャンネル／チャンネルグループは、再生中における他のチャンネル／チャンネルグループの選択的ダッキングの利用を通じて、聞き手によりはっきりと聞こえるようになり得る。

【 0 0 1 1 】

上記概要には、本発明の全ての態様の網羅的なリストを挙げてはいない。本発明には、前述でまとめた種々の態様の全ての好適な組み合わせからの実施可能な全てのシステム及び方法が含まれ、並びに以下の詳細な説明で開示されるもの、特に本出願とともに提出された請求項において指摘されるものが含まれると考えられる。このような組み合わせには、上記概要では具体的には説明されていない特定の優位性がある。

【 0 0 1 2 】

本発明の実施形態を、限定としてではなく例として、添付図面の図に示し、図面中、同様の参照符号は同様の要素を示す。本開示での、本発明の「 a n 」又は「 1 つの 」実施形態への言及は、必ずしも同じ実施形態に対するものではなく、それらは、少なくとも 1 つを意味していることに留意されたい。

【 図面の簡単な説明 】

【 0 0 1 3 】

【 図 1 】一実施形態に係る、音声エンコード装置、及び音声デコード装置のセットを含む、音声システムを示す。

【 図 2 】一実施形態に係る音声エンコード装置の構成要素図を示す。

【 図 3 】一実施形態に係る音声デコード装置の構成要素図を示す。

【 図 4 】 (1) 多重チャンネルサウンドプログラムコンテンツ片を、第 1 のチャンネル／チャンネルグループに関連付けられたダッキング値を含むようにエンコードし、 (2) ダッキング値に関連付けられた第 1 のチャンネル／チャンネルグループのためになるように、第 2 のチャンネル／チャンネルグループがダッキングされるよう、エンコードされた多重チャンネルサウンドプログラムコンテンツ片を再生するための、一実施形態に係る方法を示す。

【 発明を実施するための形態 】

【 0 0 1 4 】

いくつかの実施形態について、添付の図面を参照しながら説明する。詳細について多く説明されるが、当然のことながら、本発明のいくつかの実施形態は、これらの詳細なしに実施してよい。他の例では、本説明の理解を不明瞭にすることがないように、周知の回路、構造、及び技術について、詳細には示されていない。

【 0 0 1 5 】

図 1 は、一実施形態に係る音声システム 1 0 0 を示す。音声システム 1 0 0 は、音声エンコード装置 1 0 1、及び音声デコード装置のセット 1 0 3₁ ~ 1 0 3_Nを含んでもよい。音声エンコード装置 1 0 1 及び音声デコード装置 1 0 3₁ ~ 1 0 3_Nは、分散ネットワーク 1 0 5 を通じて通信可能に結合されてもよい。具体的には、音声エンコード装置 1 0 1 は、1 つ以上のサウンドプログラムコンテンツ片をエンコードし、エンコードされたデータを、分散ネットワーク 1 0 5 を通じて音声デコード装置 1 0 3₁ ~ 1 0 3_Nのうちの 1 つ以上へ転送してもよい。

【 0 0 1 6 】

一実施形態では、音声エンコード装置 1 0 1 は多重チャンネルサウンドプログラムコンテンツ片をエンコードしてもよい。以下において更に詳細に説明されるように、音声エンコード装置 1 0 1 によって作成された、エンコードされたデータは、サウンドプログラムコンテンツ片を構成する別個のチャンネル又はチャンネルグループのためのメタデータを含んでもよい。具体的には、特定のチャンネル／チャンネルグループのためのメタデータは、他の全てのチャンネル／チャンネルグループが、指定された継続時間／期間の間、「ダッキングされる」(即ち、強度／音量を低下させられる) べきであることを指示してもよい。このダッ

10

20

30

40

50

キングは、ダッキングされていないチャンネル／チャンネルグループ（即ち、ダッキング値に関連付けられたチャンネル／チャンネルグループ）が、ユーザ／聞き手によりはっきりと聞こえること／理解されることを可能にし得る。

【 0 0 1 7 】

次に、音声システム 1 0 0 の各要素が例として説明される。他の実施形態では、音声システム 1 0 0 は、図 1 に示され、本明細書において説明されているものよりも多くの要素を含んでもよい。

【 0 0 1 8 】

図 2 は、一実施形態に係る音声エンコード装置 1 0 1 の構成要素図を示す。音声エンコード装置 1 0 1 は、多重チャンネルサウンドプログラムコンテンツ片をエンコードする能力を有する任意のコンピューティングデバイスであってよい。例えば、音声エンコード装置 1 0 1 は、ラップトップコンピュータ、デスクトップコンピュータ、コンピュータサーバ、タブレットコンピュータ、ゲームシステム、並びに／又はモバイル機器（例えば、携帯電話若しくはモバイルメディアプレーヤ）であってもよい。次に、図 2 に示される音声エンコード装置 1 0 1 の各要素が説明される。

【 0 0 1 9 】

音声エンコード装置 1 0 1 は、メインシステムプロセッサ 2 0 1 及びメモリユニット 2 0 3 を含んでもよい。プロセッサ 2 0 1 及びメモリユニット 2 0 3 は、ここでは、音声エンコード装置 1 0 1 の様々な機能及び動作を実施するのに必要な操作を行うプログラムブルデータ処理構成要素及びデータ記憶装置の任意の好適な組み合わせを指すために、総称的に使用される。プロセッサ 2 0 1 は、特定用途向け集積回路（ASIC）、汎用マイクロプロセッサ、フィールドプログラマブルゲートアレイ（FPGA）、デジタル信号コントローラ、又は 1 組のハードウェア論理構造（例えば、フィルタ、論理演算ユニット、専用ステートマシン）などの専用プロセッサであり得るのに対し、メモリユニット 2 0 3 は、マイクロエレクトロニクス不揮発性ランダムアクセスメモリを表し得る。

【 0 0 2 0 】

オペレーティングシステムは、音声エンコード装置 1 0 1 の種々の機能に固有のアプリケーションプログラムとともにメモリユニット 2 0 3 内に記憶されてもよい。これらのプログラムは、音声エンコード装置 1 0 1 の種々の機能を実行するために、プロセッサ 2 0 1 によって走らされるか、又は実行される。例えば、メモリユニット 2 0 3 は、音声エンコード装置 1 0 1 の他のハードウェア及びソフトウェア要素と併せて、多重チャンネルサウンドプログラムコンテンツ片をエンコードする、エンコーダ 2 0 5 を含んでもよい。上述されたように、及び以下において更に詳細に説明されるように、エンコーダ 2 0 5 によって作成された、エンコードされたデータは、サウンドプログラムコンテンツ片の別個のチャンネル又はチャンネルグループのためのメタデータを含んでもよい。メタデータは、音声デコード装置 1 0 3₁ ~ 1 0 3_N のうちの 1 つ以上によって、メタデータに関連付けられていない他のチャンネル／チャンネルグループに適用されるべきダッキング値を指示してもよい。具体的には、以下において更に詳細に説明されるように、ダッキング値は、ダッキングされないチャンネル／チャンネルグループのユーザに対する明瞭性を改善するべく他のチャンネル／チャンネルグループをダッキングするために、音声デコード装置 1 0 3₁ ~ 1 0 3_N のうちの 1 つ以上によって用いられてもよい。

【 0 0 2 1 】

一実施形態では、音声エンコード装置 1 0 1 は、1 つ以上の接続を通じて他の構成要素と通信するための通信インタフェース 2 0 7 を含んでもよい。例えば、通信インタフェース 2 0 7 は、Bluetooth（登録商標）、IEEE（登録商標） 8 0 2 . 1 1 x 規格群、IEEE（登録商標） 8 0 2 . 3、セルラー方式による移動通信のグローバルシステム（GSM（登録商標））規格、セルラー方式による符号分割多元接続（CDMA）規格、及び／又はロングタームエボリューション（LTE（登録商標））規格を使用して通信する能力を有してもよい。一実施形態では、通信インタフェース 2 0 7 は、分散ネットワーク 1 0 5 を通じた映像、音声、及び／又は他のデータ片の送信／受信を促進す

10

20

30

40

50

る。例えば、音声エンコード装置 101 は、通信インタフェース 207 を介して、エンコードされるべき 1 つ以上のサウンドプログラムコンテンツ片を受信してもよい。以下において更に詳細に説明されるように、サウンドプログラムコンテンツ片は、エンコード / 処理され、同じく通信インタフェース 207 を介して再生のために音声デコード装置 103₁ ~ 103_N のうちの 1 つ以上へ送信されてもよい。

【0022】

次に、図 3 を参照して、音声デコード装置 103₁ が説明される。音声デコード装置 103₁ に関して説明されているが、音声デコード装置 103₂ ~ 103_N の各々は、同様又は同一の要素を含んでもよい。音声デコード装置 103₁ は、エンコードされたサウンドプログラムコンテンツ片を受信、デコード、及び再生する能力を有する任意のコンピューティングデバイスであってよい。例えば、音声デコード装置 103₁ は、ラップトップコンピュータ、デスクトップコンピュータ、タブレットコンピュータ、セットトップボックス、マルチメディアプレーヤ、ゲームシステム、並びに / 又はモバイル機器（例えば、携帯電話若しくはモバイルメディアプレーヤ）であってもよい。以下において更に詳細に説明されるように、音声デコード装置 103₁ は、音声エンコード装置 101 から、エンコードされたサウンドプログラムコンテンツ片を表す音声アセットを受信してもよい。音声デコード装置 103₁ は、特定のチャンネル / チャンネルグループのためのエンコードされたサウンドプログラムコンテンツ片内のメタデータを読み出ししてもよい / 抽出してもよい。メタデータは、他のチャンネル / チャンネルグループに適用されるべきダッキング値を含んでもよい。その結果、上述されたように、ダッキング値に関連付けられていないこれらのチャンネル / チャンネルグループは、ダッキングされないチャンネル / チャンネルグループの明瞭性を改善するべく、ダッキング値を用いてダッキングされてもよい。

【0023】

音声デコード装置 103₁ は、メインシステムプロセッサ 301 及びメモリユニット 303 を含んでもよい。プロセッサ 201 及びメモリユニット 203 と同様に、プロセッサ 301 及びメモリユニット 303 は、ここでは、音声デコード装置 103₁ の様々な機能及び動作を実施するのに必要な操作を行うプログラマブルデータ処理構成要素及びデータ記憶装置の任意の好適な組み合わせを指すために、総称的に使用される。プロセッサ 301 は、ASIC などの専用プロセッサ、汎用マイクロプロセッサ、FPGA、デジタル信号コントローラ、又は 1 組のハードウェア論理構造（例えば、フィルタ、算術論理演算装置、専用ステートマシン）であってもよいのに対し、メモリユニット 303 は、マイクロエレクトロニクス不揮発性ランダムアクセスメモリを指してもよい。

【0024】

オペレーティングシステムは、音声デコード装置 103₁ の種々の機能に固有のアプリケーションプログラムとともにメモリユニット 303 内に記憶されてもよい。これらのプログラムは、音声デコード装置 103₁ の種々の機能を実行するために、プロセッサ 301 によって走らされるか、又は実行される。例えば、メモリユニット 303 はデコード 305 を含んでもよい。デコード 305 は、音声デコード装置 103₁ の 1 つ以上の他の構成要素とともに、エンコードされたサウンドプログラムコンテンツ片を表す音声アセットからダッキング値を抽出 / 導出してもよい。例えば、ダッキング値は、一次チャンネルグループの分析を通じて導出され、音声アセットに挿入されていてもよい。デコード 305 は、ダッキング値を抽出するか、又は混合信号内のダッキング値のエンコードされたバージョンにアルゴリズムを適用してダッキング値を作成してもよい。その後、デコード 305 は、これらのダッキング値を、エンコードされたサウンドプログラムコンテンツ片の 1 つ以上のチャンネル / チャンネルグループに適用してもよい。具体的には、ダッキング値は、エンコードされたサウンドプログラムコンテンツ片内において第 1 のチャンネル / チャンネルグループに関連付けられていてもよく、デコード 305 は、ダッキング値を、第 1 のチャンネル / チャンネルグループと異なる第 2 のチャンネル / チャンネルグループに適用してもよい。したがって、デコード 305 によって実行されるダッキングは、ダッキング値に関連付けられたチャンネル / チャンネルグループではなく、音声アセット内のダッキング値に関連付けら

れていない他のチャンネル／チャンネルグループに対するものとなる。

【 0 0 2 5 】

一実施形態では、音声デコード装置 1 0 3₁は、エンコードされたサウンドプログラムコンテンツ片に基づいて音響を出力するための1つ以上のラウドスピーカ 3 0 9を含んでもよい。ラウドスピーカ 3 0 9は、フルレンジドライバ、ミッドレンジドライバ、サブウーファ、ウーファ、ツイータの任意の組み合わせであり得る。ラウドスピーカ 3 0 9の各々は、円筒状磁気ギャップ内を軸方向に運動するよう、電線コイル（例えば、ボイスコイル）を拘束する可撓性サスペンションを介して、硬質バスケット、又はフレームに接続された、軽量ダイアフラム、又はコーンを用いてよい。電気音声信号がボイスコイルに印加されると、電流によってボイスコイル内に磁界が生成され、可変電磁石となる。コイルとラウドスピーカ 3 0 9の磁気システムとが双方向作用し、コイル（ひいては取り付けられたコーン）を前後に運動させる機械力を発生させる。これによって、発信源から到来する印加電気音声信号の制御の下で音響を再現する。

10

【 0 0 2 6 】

一実施形態では、音声デコード装置 1 0 3₁は、1つ以上の接続を通じて他の構成要素と通信するための通信インタフェース 3 0 7を含んでもよい。例えば、通信インタフェース 3 0 7は、Bluetooth（登録商標）、IEEE（登録商標） 8 0 2 . 1 1 x規格群、IEEE（登録商標） 8 0 2 . 3、セルラー方式による移動通信のグローバルシステム（GSM（登録商標））規格、セルラー方式による符号分割多元接続（CDMA）規格、及び／又はロングタームエボリューション（LTE（登録商標））規格を使用して通信する能力を有してもよい。一実施形態では、通信インタフェース 3 0 7は、映像、音声、及び／又は他のデータ片の送信／受信を促進する。例えば、音声デコード装置 1 0 3₁は、音声エンコード装置 1 0 1から通信インタフェース 3 0 7を介して、エンコードされたサウンドプログラムコンテンツ片を表す音声アセットを受信してもよい。この受信された音声アセットは、以下において更に詳細に説明されるように、音声デコード装置 1 0 3₁によってデコードされ、再生されてもよい。

20

【 0 0 2 7 】

上述されたように、他の音声デコード装置 1 0 3₂ ~ 1 0 3_Nは、音声デコード装置 1 0 3₁と同一であってもよい。具体的には、音声デコード装置 1 0 3₂ ~ 1 0 3_Nは各々、プロセッサ 3 0 1、メモリユニット 3 0 3、デコーダ 3 0 5、通信インタフェース 3 0 7、及び1つ以上のラウドスピーカ 3 0 9を含んでもよい。

30

【 0 0 2 8 】

上述されたように、音声エンコード装置 1 0 1及び音声デコード装置 1 0 3₁ ~ 1 0 3_Nは、分散ネットワーク 1 0 5を通じて通信してもよい。分散ネットワーク 1 0 5は、スイッチ、ルータ、コントローラ、アクセスポイントなどを含む、ネットワーク機器の任意の組み合わせで構成されてもよい。分散ネットワークは、Bluetooth（登録商標）、IEEE（登録商標） 8 0 2 . 1 1 x規格群、IEEE（登録商標） 8 0 2 . 3、セルラー方式による移動通信のグローバルシステム（GSM（登録商標））規格、セルラー方式による符号分割多元接続（CDMA）規格、及び／又はロングタームエボリューション（LTE（登録商標））規格を含む、1つ以上の規格／プロトコルを使用して動作してもよい。

40

【 0 0 2 9 】

次に、図 4 を参照して、(1) 多重チャンネルサウンドプログラムコンテンツ片を、第 1 のチャンネル／チャンネルグループに関連付けられたダッキング値を含むようにエンコードし、(2) ダッキング値に関連付けられた第 1 のチャンネル／チャンネルグループのためになるように、第 2 のチャンネル／チャンネルグループがダッキングされるよう、エンコードされた多重チャンネルサウンドプログラムコンテンツ片を再生するための方法 4 0 0 が説明される。方法 4 0 0 の動作はチャンネル／チャンネルグループに関して説明されているが、他の実施形態では、ダッキングは、サウンドプログラムコンテンツ片によって表されたサウンドオブジェクト又はステムに対して同様の仕方で行われてもよい。

50

【 0 0 3 0 】

方法 4 0 0 の各動作は、音声エンコード装置 1 0 1、1 つ以上の音声デコード装置 1 0 3₁ ~ 1 0 3_N、及び / 又は別の装置の 1 つ以上の構成要素によって実行されてもよい。例えば、以下の説明において用いられるように、音声エンコード装置 1 0 1 のエンコーダ 2 0 5 及び音声デコード装置 1 0 3₁ のデコーダ 3 0 5 のうちの 1 つ以上が、方法 4 0 0 の動作を実行するために用いられてもよい。しかし、他の実施形態では、異なる要素及び装置が、方法 4 0 0 を実行するために用いられてもよい。エンコーダ 2 0 5 及びデコーダ 3 0 5 は、メモリユニット 2 0 3 及び 3 0 3 内にそれぞれ存在するように説明されているが、他の実施形態では、エンコーダ 2 0 5 及びデコーダ 3 0 5 は、フィルタ、算術論理演算装置、及び専用ステートマシンを含む、1 つ以上のハードウェア構造によって実装されてもよい。

10

【 0 0 3 1 】

方法 4 0 0 の動作は特定の順序で示され、説明されているが、他の実施形態では、方法 4 0 0 の動作は異なる順序で実行されてもよい。例えば、動作のうちの 1 つ以上は、並行して、又は一部重なった期間中に実行されてもよい。次に、以下において、方法 4 0 0 の各動作が例として説明される。

【 0 0 3 2 】

一実施形態では、方法 4 0 0 は、動作 4 0 1 において、サウンドプログラムコンテンツ片のための 1 つ以上のチャンネルを表す音声信号のセットの受信によって開始してもよい。例えば、音声エンコード装置 1 0 1 は、サウンドプログラムコンテンツ片（例えば、楽曲又は、映画のサウンドトラック / 音声トラック）に対応する N 個の音声チャンネルを受信してもよい。ここで、N は 2 以上である。例えば、サウンドプログラムコンテンツ片を表す 4 つの音声チャンネルが動作 4 0 1 において受信されてもよい。いくつかの実施形態では、受信されるチャンネルのうちの 1 つは、視覚障害を持ったユーザが、映画又はテレビ放送内で起こっている動作を理解することを支援するナレーションチャンネルに対応してもよい。これらの実施形態では、他のチャンネルはいずれかの空間チャンネル（例えば、左、右、及び中央音声チャンネル）に対応し、及び / 又はオブジェクト / ステム（例えば、ダイアログ、音楽、効果音声オブジェクト / ステム）を表してもよい。音声信号 / チャンネルは、動作 4 0 1 において、外部システム又は装置（例えば、外部コンピュータ又はストリーミング音声サービス）から通信インタフェース 2 0 7 を介して受信されてもよい。他の実施形態では、音声信号 / チャンネルは、音声エンコード装置 1 0 1 上にローカルに記憶され（例えば、メモリユニット 2 0 3 内に記憶され）、動作 4 0 1 において取得されてもよい。

20

30

【 0 0 3 3 】

いくつかの実施形態では、動作 4 0 1 において受信される音声信号 / チャンネルは、映像と同時の、それと同期した再現を意図されてもよい。例えば、上述されたように、動作 4 0 1 において受信される音声信号 / チャンネルは、映画又はテレビ番組のための音声トラックであってもよい。本実施形態では、音声信号 / チャンネルは、動作 4 0 1 において、対応する映像コンテンツとともに、又は映像コンテンツとは別に、送信及び受信されてもよい。

【 0 0 3 4 】

動作 4 0 3 において、動作 4 0 1 において受信されたチャンネルを処理 / エンコードし、サウンドプログラムコンテンツ片を表す音声アセットを生成してもよい。一実施形態では、チャンネルはグループ化され、これらのグループ内で処理されてもよい。各チャンネルグループは、互いに関連付けられた 1 つ以上の個々のチャンネルを含んでもよい。チャンネルグループは、エンコードされたサウンドプログラムコンテンツ片のデコード / 再生中にチャンネルの各々のそれぞれのグループに適用されるべきダイナミックレンジ制御 / 圧縮（DRC : Dynamic Range Control / Compression）ゲイン値を決定するために、まとめて分析されてもよい。DRC は、音声信号のダイナミックレンジを狭めるか、又は圧縮することによって、大きな音の音量を低下させるか、又は静かな音を増幅する。DRC ゲイン値は、エンコードされたサウンドプログラムコンテンツ片を表す音声アセットにメタデータとして

40

50

追加されてもよい。

【0035】

一実施形態では、サウンドプログラムコンテンツ片内の単一のチャンネル／チャンネルグループのためのダッキング値が決定されてもよい。ダッキング値は、他のチャンネル／チャンネルグループがダッキングされることになる（例えば、音量を低下させられることになる）量を表す。例えば、動作401において受信されるサウンドプログラムコンテンツ片は、4つのオブジェクトを含んでもよい：（1）ナレーションオブジェクト、（2）ダイアログオブジェクト、（3）音楽オブジェクト、及び（4）効果オブジェクト。いくつかの実施形態では、これらの4つのオブジェクトは処理前にチャンネルグループに割り当てられてもよい（例えば、ナレーションオブジェクトは第1のチャンネルグループに割り当てられてもよく、ダイアログオブジェクトは第2のチャンネルグループに割り当てられてもよく、音楽オブジェクト及び効果オブジェクトは両方とも第3のチャンネルグループに割り当てられてもよい）。本例では、ダッキング値のセットが生成され、第1のチャンネルグループに関連付けられてもよい。第1のチャンネルグループに関連付けられるダッキング値のセットは、音声デコード装置103₁～103_Nのうちの1つ以上の内部のデコーダ305によって他のチャンネルグループのうちの1つ以上に適用されるべき減衰量を指示する。

10

【0036】

いくつかの実施形態では、チャンネル／チャンネルグループの元のレイアウト、及びチャンネル／チャンネルグループの各々の可能なダウンミックスに、個々のダッキングシーケンスが関連付けられてもよい。例えば、チャンネル／チャンネルグループの元のレイアウト、及びチャンネル／チャンネルグループの各々の可能なダウンミックスに、識別子が関連付けられてもよい。本例では、各識別子にダッキング値の別個のシーケンスが関連付けられてもよい。その結果、チャンネル／チャンネルグループに、それらのレイアウト／ダウンミックス構成に基づいて、ダッキング値の適切なシーケンスが適用されてもよい。

20

【0037】

一実施形態では、ダッキング値は、これらのチャンネルグループのダイナミックレンジを狭めること／圧縮することによって、チャンネルグループに適用されてもよい。例えば、ダッキング値は、他のチャンネルグループが、再生中に、音声デコード装置103₁～103_Nのうちの1つ以上によって10dBだけダッキングされてもよいことを指示してもよく、DRCゲイン値と同様に適用されてもよい。このダッキングは、指定された期間にわたって変動してもよい。例えば、第1の期間中に第1のチャンネルグループ内における活性（例えば、発声又は発語活性）を検出している間に、動作403は、この第1の期間中には、第1のチャンネルグループに関連付けられたダッキング値を第1のレベルに設定してもよい。逆に、第2の期間中に第1のチャンネルグループ内における活性を検出しなければ、動作403は、この第2の期間中には、第1のチャンネルグループに関連付けられたダッキング値を第2のレベルに設定してもよい。本例では、第1のレベルは第2のレベルより大きくてもよく、それにより、第2及び第3のチャンネルグループは、第1のチャンネルグループ内における活性期間中には、不活性期間中よりも大きくダッキングされる。したがって、第1のチャンネルグループは、活性期間中に強調されることができ、その一方で、第1のチャンネルグループが比較的不活性である期間中には、第2及び第3のチャンネルグループが聞こえ、及び変更されないことを許容する。

30

40

【0038】

上述されたように、ダッキング値はDRCゲイン値と同様に適用されてもよい。これらの実施形態では、ダッキング値の適用によって、チャンネルグループの減衰を生じさせないか（即ち、第1のチャンネルグループ内における不活性期間中）、又はチャンネルグループの減衰を生じさせるように、ダッキング値は、0dB以下であってよい。いくつかの実施形態では、ダッキング値の適用は、チャンネル／チャンネルグループの「クリッピング」又は「フェーディング」を含む、音声アセット内に指示されている他の効果を無効にしてもよい。

【0039】

50

いくつかの実施形態では、ダッキング値は、各チャンネルグループに対して異なるレベルのダッキングの適用を指示してもよい。上述された3つの例示的チャンネルグループを用いると、動作403は、ダイアログに関連付けられた第2のチャンネルグループは、再生中に、音楽及び効果に関連付けられた第3のチャンネルグループよりも大きくダッキングされる必要があり得ると判定してもよい。ダイアログのこのより大きなダッキングによって、第1のチャンネルグループ内のナレーションは、第3のチャンネルグループ内の音楽及び効果とともに前景音声内に存在することが可能になり、その一方で、第2のチャンネルグループ内のダイアログはちょうどそれに隠れて（即ち、背景内で）再生されることができ。ダイアログは、音楽及び効果と比べてナレーションの明瞭性によりひどく干渉し得るため、第2及び第3のチャンネルグループの別個のダッキングを可能にすることで、ユーザに対するナレーションの明瞭性を依然として高めながらも、ダッキング量を最小限に抑える。

10

【0040】

上述されたように、ダッキング値は、チャンネル/チャンネルグループの再生レベルを低下させるために用いられてもよい。いくつかの実施形態では、ダッキング値は、再生のレベルを変えるだけでなく、チャンネル/チャンネルグループの音場内における発音位置を追加的に、又は別個に変更するように、拡張されてもよい。例えば、上述された3つの例示的チャンネルグループを用いると、動作403において音声アセットに追加されるダッキング値は、第2のチャンネルグループの発音を、音声デコード装置103₁~103_Nによって再生中にフロントラウドスピーカ309からサラウンドラウドスピーカ309へ移動させるべきであることを指示してもよい。このチャンネル/チャンネルグループの移動は、ある期間中に限定されてもよい。例えば、チャンネル/チャンネルグループの移動は、第1のチャンネルグループ（例えば、ナレーションチャンネルグループ）内で活性が検出されたときのみ、生じさせてもよい。音場内におけるチャンネル/チャンネルグループの発音位置を移動させることによって、目的の特定のチャンネル/チャンネルグループを他のチャンネル/チャンネルグループに対して強調させることができる。

20

【0041】

一実施形態では、動作403において生成され、音声アセット内に挿入されるダッキング値は、単一のチャンネル/チャンネルグループに関連付けられてもよい。したがって、各音声アセットは、再生中にダッキングされない単一のチャンネル/チャンネルグループを含み、その一方で、音声アセット内の他の全てのチャンネル/チャンネルグループは、ダッキング値を用いてダッキングされる。いくつかの実施形態では、動作403において生成されるダッキング値は複数のチャンネル/チャンネルグループに関連付けられてもよいが、これらの複数のチャンネル/チャンネルグループのうちの1つのみが一度に再生のためにアクティブになることができる。例えば、音声アセットは、異なる言語に対応する複数の異なるナレーションチャンネル/チャンネルグループを含んでもよい。これらの実施形態では、1つの言語のみが一度に再生のために選択されてもよい（即ち、異なるナレーションチャンネル/チャンネルグループの中で、アクティブなチャンネル/チャンネルグループは1つのみ）。したがって、ダッキング値は、ダッキング値に関連付けられていない他の全てのアクティブなチャンネル/チャンネルグループに適用される（即ち、ダッキング値は、アクティブ及び非アクティブなナレーションチャンネル/チャンネルグループに適用されない）。

30

40

【0042】

いくつかの実施形態では、音声アセットは、異なるチャンネル/チャンネルグループに各々関連付けられるダッキング値の複数のセットを含んでもよい。例えば、上述の例を用いると、音声アセットは、1つのチャンネル/チャンネルグループのみが一度にアクティブになることを許可される、異なる言語に対応する複数の異なるナレーションチャンネル/チャンネルグループを含んでもよい。本実施形態では、ナレーションチャンネル/チャンネルグループの各々は、再生中に他の全てのアクティブなチャンネル/チャンネルグループに適用されることになるダッキング値の異なるセットに関連付けられてもよい。

【0043】

ダッキング値を含む、エンコードされたサウンドプログラムコンテンツの生成に続き

50

、動作405は、サウンドプログラムコンテンツ片を表す音声アセットを音声デコード装置103₁~103_Nのうちの1つ以上へ転送してもよい。一実施形態では、音声アセットの転送は、音声エンコード装置101の通信インタフェース207及び音声デコード装置103₁~103_Nのうちの1つ以上の通信インタフェース307を用いて分散ネットワーク105を通じて実行されてもよい。他の実施形態では、音声アセットは、動作405において、他の技法を通じて（例えば、ポータブルメモリデバイスを介して）転送されてもよい。

【0044】

動作407において、音声デコード装置103₁~103_Nのうちの1つ以上は、音声アセットをデコードし、2つ以上のチャンネル/チャンネルグループ、チャンネル/チャンネルグループのうちの1つに関連付けられたダッキング値、及び/又は動作403において追加されたその他の音声情報を明らかにしてもよい。上述されたように、ダッキング値は、（1）特定のチャンネル/チャンネルグループをダッキングするため、及び/又は（2）特定のチャンネルを音場内で移動させるために用いられてもよい。

10

【0045】

動作409において、方法400は、ダッキング値が音声アセット内に存在していたかどうか、及びダッキング値に関連付けられたチャンネル/チャンネルグループが再生のために選択されていたかどうか（例えば、ユーザによって再生のために選択されていたかどうか）を判定してもよい。動作413において、ダッキング値が音声アセット内に含まれていないか、又はダッキング値に関連付けられたチャンネルが再生のために選択されていない場合には、再生のために選択されたチャンネル/チャンネルグループのために、ダッキングを用いずに駆動信号を生成してもよい。動作413において生成されたこれらの駆動信号は、動作415において、ラウドスピーカ309を通した再生のために用いられてもよい。

20

【0046】

逆に、動作411において、ダッキング値が音声アセット内に含まれ、関連付けられたチャンネル/チャンネルグループが再生のために選択されていた場合には、デコードされたダッキング値を、関連付けられていないチャンネル/チャンネルグループに適用し、動作415におけるラウドスピーカ309を通した再生のための駆動信号のセットを作成してもよい。具体的には、ダッキング値は、ダッキング値に関連付けられていないチャンネル/チャンネルグループのダイナミックレンジを縮小してもよい。上述された3つの例示的チャンネルグループでは、ダッキング値は第2及び/又は第3のチャンネルグループに適用されてもよく、それにより、第1のチャンネルグループからの音声はダッキングされないままとなり、その結果、ユーザに対してより明瞭になる。いくつかの実施形態では、動作411における適用の前に、スケール係数がダッキング値に適用されてもよい。

30

【0047】

動作403に関して上述されたように、ダッキング値はまた、ダッキング値に関連付けられていない1つ以上のチャンネル/チャンネルグループを音場内で移動させてもよい。例えば、動作411において、ダッキング値は、上述の例では、第2のチャンネルグループをサラウンドラウドスピーカ309内へ移動させてもよい。第2及び/又は第3のチャンネルグループのダイナミックレンジをダッキングすることに加えて、この移動は第1のチャンネルグループの明瞭性を改善し得る。

40

【0048】

上述されたように、ダッキング値は、サウンドプログラムコンテンツ片を表す音声アセット内に含まれ、特定のチャンネル/チャンネルグループに関連付けられてもよい。エンコードされたサウンドプログラムコンテンツ片の再生中に、ダッキング値は、他の全てのチャンネル/チャンネルグループに適用されてもよい。このダッキングは、ダッキングされないチャンネル/チャンネルグループ内の音声の明瞭性を改善し得る。例えば、ナレーションチャンネル/チャンネルグループは、再生中における他のチャンネル/チャンネルグループの選択的ダッキングの利用を通じて、ユーザによりはっきりと聞こえるようになり得る。

【0049】

50

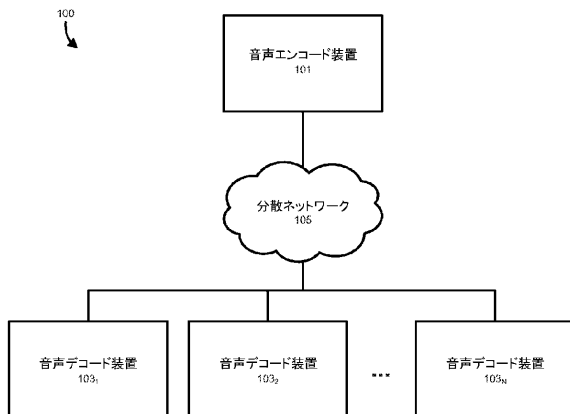
先に説明されたように、本発明の一実施形態は、製造物品であって、その内部において、機械可読媒体（マイクロエレクトロニックメモリなど）が、その上に、上述された動作を実行するように１つ以上のデータ処理構成要素（ここでは総称的に「プロセッサ」と呼ばれる）をプログラムする命令を記憶した、製造物品であってもよい。他の実施形態では、これらの動作の一部は、ハードワイヤード論理（例えば、専用デジタルフィルタブロック及びステートマシン）を含む特定のハードウェア構成要素によって実行されてもよいであろう。それらの動作は、或いは、プログラムされたデータ処理構成要素及び固定されたハードワイヤード回路構成要素の任意の組み合わせにより実行されることがあり得る。

【 0 0 5 0 】

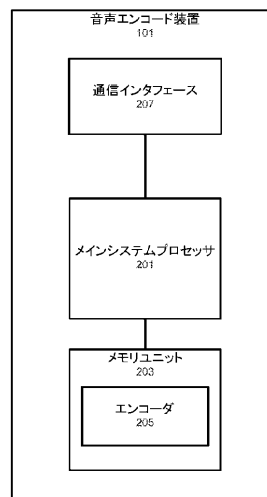
ある実施形態について説明し添付の図面に示してきたが、当然のことながら、このような実施形態は大まかな発明を単に例示するものであってそれを限定するものではなく、また、本発明は図示及び説明した特定の構成及び配置には限定されない。なぜならば、他の種々の変更が当業者に想起され得るからである。したがって、説明は、限定的ではなく例示的であるとみなされるべきである。

10

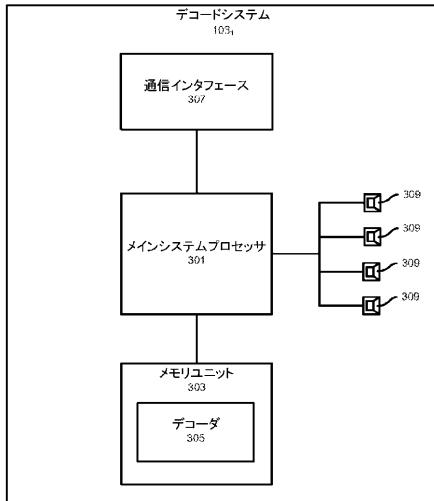
【 図 1 】



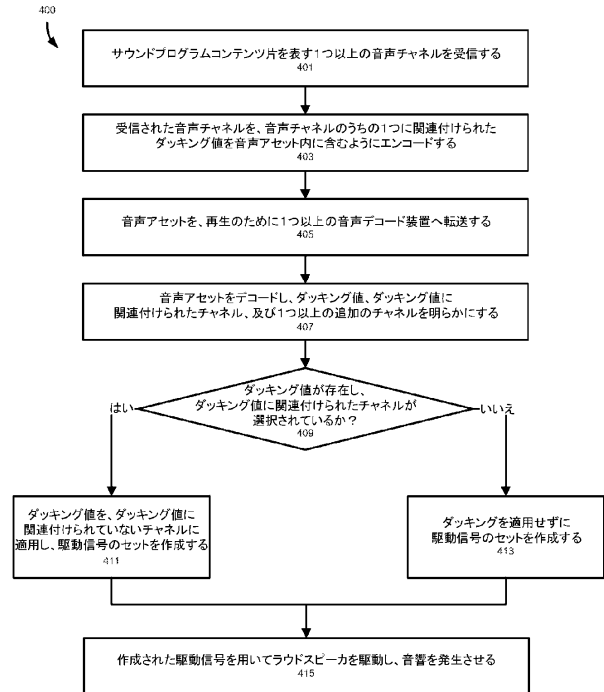
【 図 2 】



【図 3】



【図 4】



フロントページの続き

(74)代理人 100086771

弁理士 西島 孝喜

(74)代理人 100139712

弁理士 那須 威夫

(74)代理人 100122563

弁理士 越柴 絵里

(72)発明者 ホルマン トムリンソン エム

アメリカ合衆国 95014 カリフォルニア州 クパチーノ インフィニット ループ 1 メ
ール ストップ 81-2ピーティー

(72)発明者 バウムガルテ フランク エム

アメリカ合衆国 95014 カリフォルニア州 クパチーノ インフィニット ループ 1 メ
ール ストップ 38-2アイエムジー

(72)発明者 アラマンヒェ エリック エイ

アメリカ合衆国 95014 カリフォルニア州 クパチーノ インフィニット ループ 1 メ
ール ストップ 38-2アイエムジー

審査官 五貫 昭一

(56)参考文献 特表2016-519788(JP,A)

特表2016-520858(JP,A)

特許第4686684(JP,B1)

(58)調査した分野(Int.Cl., DB名)

G10L 19/008

G10L 19/00