

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7362807号
(P7362807)

(45)発行日 令和5年10月17日(2023.10.17)

(24)登録日 令和5年10月6日(2023.10.6)

(51)国際特許分類 F I
H 0 4 S 7/00 (2006.01) H 0 4 S 7/00 3 0 0

請求項の数 8 外国語出願 (全29頁)

(21)出願番号	特願2022-27836(P2022-27836)	(73)特許権者	507236292
(22)出願日	令和4年2月25日(2022.2.25)		ドルビー ラボラトリーズ ライセンシング
(62)分割の表示	特願2020-117715(P2020-117715))の分割		グ コーポレイション
原出願日	平成28年2月4日(2016.2.4)		アメリカ合衆国 9 4 1 0 3 カリフォル
(65)公開番号	特開2022-65179(P2022-65179A)		ニア州 サンフランシスコ マーケット
(43)公開日	令和4年4月26日(2022.4.26)	(74)代理人	ストリート 1 2 7 5
審査請求日	令和4年2月25日(2022.2.25)		100107766
(31)優先権主張番号	62/113,268		弁理士 伊東 忠重
(32)優先日	平成27年2月6日(2015.2.6)	(74)代理人	100070150
(33)優先権主張国・地域又は機関	米国(US)		弁理士 伊東 忠彦
		(74)代理人	100135079
			弁理士 宮崎 修
		(72)発明者	ランドー, ジョシュア ブランドン
			アメリカ合衆国, カリフォルニア州 9
			4 1 0 3, サンフランシスコ, マーケッ
			最終頁に続く

(54)【発明の名称】 適応オーディオ・コンテンツのためのハイブリッドの優先度に基づくレンダリング・システムおよび方法

(57)【特許請求の範囲】

【請求項1】

適応オーディオをレンダリングする方法であって：

静的なチャンネル・ベースのオーディオおよび少なくとも一つの動的オブジェクトを含む入力オーディオ・ビットストリームを受領する段階であって、前記動的オブジェクトは優先度値をもち、前記入力オーディオ・ビットストリームは、オーディオ・コンテンツおよびレンダリング・メタデータを含むオブジェクト・オーディオ・ベースのデジタル・ビットストリーム・フォーマットに従ってフォーマットされている、段階と；

前記動的オブジェクトが低優先度動的オブジェクトであるか、または前記動的オブジェクトが高優先度動的オブジェクトであるかを判定する段階であって、該判定は、前記優先度値の、優先度閾値との比較に基づいて、前記動的オブジェクトを低優先度動的オブジェクトまたは高優先度動的オブジェクトのいずれかとして分類することを含み、前記優先度閾値は、事前設定された値または自動化されたプロセス選択に基づく、段階と；

前記動的オブジェクトが前記低優先度動的オブジェクトである場合、前記動的オブジェクトを第一のレンダリング・プロセスに基づいてレンダリングする、または前記動的オブジェクトが前記高優先度動的オブジェクトである場合、前記動的オブジェクトを第二のレンダリング・プロセスに基づいてレンダリングする段階とを含み、

前記第一のレンダリング・プロセスは前記第二のレンダリング・プロセスとは異なるメモリ処理を使用し、

前記第一のレンダリング・プロセスまたは前記第二のレンダリング・プロセスは、前記

動的オブジェクトの前記分類に基づいて選択され、前記分類とは独立に前記静的なチャンネル・ベースのオーディオをレンダリングする、方法。

【請求項 2】

スピーカー・システムへの伝送のために後処理する段階をさらに含む、請求項 1 記載の方法。

【請求項 3】

前記後処理する段階は、アップミックス、ボリューム制御、等化および低音管理のうちの少なくとも一つを含む、請求項 2 記載の方法。

【請求項 4】

前記後処理する段階は、前記スピーカー・システムを通じた再生のための前記入力オーディオ・ビットストリームに存在している高さ手がかりのレンダリングを容易にするための仮想化段階をさらに含む、請求項 3 記載の方法。

【請求項 5】

前記第一のレンダリング・プロセスは、前記静的なチャンネル・ベースのオーディオをレンダリングするよう最適化されている第一のレンダリング・プロセッサにおいて実行され、

前記第二のレンダリング・プロセスは、前記第一のレンダリング・プロセッサに比べて、向上したパフォーマンス機能、向上したメモリ帯域幅および向上した伝送帯域幅のうちの少なくとも一つによって前記高優先度動的オブジェクトをレンダリングするよう最適化されている第二のレンダリング・プロセッサにおいて実行される、

請求項 1 記載の方法。

【請求項 6】

前記第一のレンダリング・プロセッサおよび前記第二のレンダリング・プロセッサは、伝送リンクを通じて互いに結合された別個のレンダリング・デジタル信号プロセッサ (DSP) として具現される、請求項 5 記載の方法。

【請求項 7】

プロセッサによって実行されたときに請求項 1 記載の方法を実行する命令を含んでいる非一時的なコンピュータ可読記憶媒体。

【請求項 8】

入力オーディオ・ビットストリームの適応オーディオをレンダリングするためのシステムであって：

静的なチャンネル・ベースのオーディオおよび少なくとも一つの動的オブジェクトを含む入力オーディオ・ビットストリームを受領するためのインターフェースであって、前記動的オブジェクトは優先度値をもち、前記入力オーディオ・ビットストリームは、オーディオ・コンテンツおよびレンダリング・メタデータを含むオブジェクト・オーディオ・ベースのデジタル・ビットストリーム・フォーマットに従ってフォーマットされている、インターフェースと；

前記動的オブジェクトが低優先度動的オブジェクトであるか、または前記動的オブジェクトが高優先度動的オブジェクトであるかを判定するためのデコード段であって、該判定は、前記優先度値の、優先度閾値との比較に基づいて、前記動的オブジェクトを低優先度動的オブジェクトまたは高優先度動的オブジェクトのいずれかとして分類することを含み、前記優先度閾値は、事前設定された値または自動化されたプロセス選択に基づく、デコード段と；

前記動的オブジェクトが前記低優先度動的オブジェクトである場合、前記動的オブジェクトを第一のレンダリング・プロセスに基づいてレンダリングする、または前記動的オブジェクトが前記高優先度動的オブジェクトである場合、前記動的オブジェクトを第二のレンダリング・プロセスに基づいてレンダリングするためのレンダリング段とを有しており、

前記第一のレンダリング・プロセスは前記第二のレンダリング・プロセスとは異なるメモリ処理を使用し、

前記第一のレンダリング・プロセスまたは前記第二のレンダリング・プロセスは、前記

10

20

30

40

50

動的オブジェクトの前記分類に基づいて選択され、前記分類とは独立に前記静的なチャンネル・ベースのオーディオをレンダリングする、システム。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願への相互参照

本願は2015年2月6日に出願された米国仮特許出願第62/113,268号の優先権を主張するものである。同出願の内容はここに参照によってその全体において組み込まれる。

【0002】

技術分野

一つまたは複数の実装は概括的にはオーディオ信号処理に関し、より詳細には適応オーディオ・コンテンツのための、ハイブリッドの優先度に基づくレンダリング戦略に関する。

【背景技術】

【0003】

デジタル映画館の導入および三次元（「3D」）コンテンツまたは仮想3Dコンテンツの発達は、サウンドについての新たなスタンダードを作り出した。たとえば、コンテンツ・クリエイターにとってのより大きな創造性を許容する複数チャンネルのオーディオの組み込みや、聴衆にとってのより包み込むような、リアルな聴覚経験などである。空間的オーディオを配送する手段として伝統的なスピーカー・フィールドおよびチャンネル・ベースのオーディオを超えて拡張することは枢要であり、聴取者が選んだ構成のために特にレンダリングされたオーディオを用いることで聴取者が所望される再生構成を選択することを許容するモデル・ベースのオーディオ記述には多大な関心が寄せられてきた。音の空間的呈示はオーディオ・オブジェクトを利用する。オーディオ・オブジェクトは、見かけの源位置（たとえば3D座標）、見かけの源幅および他のパラメータの、関連付けられたパラメトリックな源記述をもつオーディオ信号である。さらなる進歩として、オーディオ・オブジェクトと伝統的なチャンネル・ベースのスピーカー・フィールドとの混合をオーディオ・オブジェクトのための位置メタデータとともに含む次世代空間的オーディオ（「適応オーディオ」とも称される）フォーマットが開発されている。空間的オーディオ・デコーダでは、チャンネルは関連付けられたスピーカーに直接送られるか、あるいは既存のスピーカー集合にダウンミックス〔下方混合〕され、オーディオ・オブジェクトはデコーダによって、柔軟な（適応的な）仕方でレンダリングされる。各オブジェクトに関連付けられたパラメトリックな源記述、たとえば3D空間における位置軌跡は、デコーダに接続されたスピーカーの数および位置とともに入力として取られる。次いで、レンダラーはパン則のようなある種のアルゴリズムを使って、取り付けられたスピーカーの集合にまたがって各オブジェクトに関連付けられたオーディオを分配する。このようにして、各オブジェクトのオーディオがレンダリングされた空間的意図が、聴取室に存在する特定のスピーカー構成を通じて、最適に呈示される。

【0004】

高度なオブジェクト・ベースのオーディオの到来は、さまざまな異なるスピーカー・アレイに伝送されるオーディオ・コンテンツの性質およびレンダリング・プロセスの複雑さを有意に増した。たとえば、映画サウンドトラックは、スクリーン上の像に対応する多くの異なる音要素、ダイアログ、ノイズおよびサウンド効果を含むことがある。これらの音要素は、スクリーン上の異なる位置から発し、背景音楽および周囲効果（ambient effects）と組み合わせさせて全体的な聴覚体験を作り出す。正確な再生は、音が、音源の位置、強度、動きおよび奥行きに関してスクリーン上に示されるものにできるだけ近く対応する仕方で再現されることを要求する。

【0005】

高度な3Dオーディオ・システム（ドルビー（登録商標）アトモス（商標）システムなど）は主に映画館用途のために設計され、配備されてきたが、映画館の適応オーディオ経験

10

20

30

40

50

を家庭やオフィス環境にもたらず消費者レベルのシステムが開発されつつある。映画館に比べ、これらの環境は会場サイズ、音響特性、システム・パワーおよびスピーカー構成の点で明らかな制約がある。このように、現在の業務用レベルの空間的オーディオ・システムは、高度なオブジェクト・オーディオ・コンテンツを、種々のスピーカー構成および再生機能を備える聴取環境にレンダリングするよう適応される必要がある。この目的に向け、コンテンツ依存レンダリング・アルゴリズム、反射音送出などといった洗練されたレンダリング・アルゴリズムおよび技法の使用を通じて空間的な音の手がかりを再現するよう、伝統的なステレオまたはサラウンドサウンド・スピーカー・アレイの機能を拡張するために、ある種の仮想化技法が開発されている。そのようなレンダリング技法は、オブジェクト・オーディオ・メタデータ・コンテンツ (OAMD: object audio metadata content) ベッドおよびISF (Intermediate Spatial Format [中間空間的フォーマット]) オブジェクトのような種々の型の適応的なオーディオ・コンテンツをレンダリングするよう最適化されたDSPベースのレンダラーおよび回路の開発につながった。個別的なOAMDコンテンツをレンダリングすることに関して適応オーディオの種々の特性を活用する種々のDSP回路が開発されている。しかしながら、そのようなマルチプロセッサ・システムはそれぞれのプロセッサのメモリ帯域幅および処理機能に関する最適化を必要とする。

10

【0006】

したがって、必要とされているのは、適応オーディオのためのマルチプロセッサ・レンダリング・システムにおける二つ以上のプロセッサのためのスケーラブルなプロセッサ負荷を提供するシステムである。

20

【0007】

サラウンドサウンドおよび映画館ベースのオーディオの家庭における採用が増えたことで、標準的なツウエーまたはスリーウエーの床置き型またはブックシェルフ型スピーカーを超えたスピーカーの種々の型および構成が開発されている。5.1または7.1システムの一部としてのサウンドバー・スピーカーのような種々のスピーカーが特定のコンテンツを再生するために開発されている。サウンドバーは二つ以上のドライバーが単一のエンクロージャー (スピーカー・ボックス) 内に集められており、典型的には単一の軸に沿って配置されているスピーカーのクラスを表わす。たとえば、一般的なサウンドバーは典型的には、スクリーンから直接音を送出するために、テレビジョンまたはコンピュータ・モニタの上、下または真正面に収まるよう設計された長方形のボックスにおいて整列されている4~6個のスピーカーを含む。サウンドバーの構成のため、物理的な配置を通じた高さ手がかりを提供するスピーカー (たとえば高さドライバー (height driver)) または他の技法に比べて、ある種の仮想化技法は実現するのが難しいことがある。

30

【0008】

したがって、さらに必要とされているのは、サウンドバー・スピーカー・システムを通じた再生のための適応オーディオ仮想化技法を最適化するシステムである。

【0009】

背景セクションで論じられている主題は、単に背景セクションでの開示のために従来技術であると想定されるべきではない。同様に、背景セクションにおいて言及されているまたは背景セクションの主題に関連する問題は、従来技術において以前から認識されていたと想定されるべきではない。背景セクションにおける主題は単に、種々のアプローチを表わすものであり、それらのアプローチ自身も発明であることがありうる。ドルビー、ドルビー・トゥールHDおよびアトモスはドルビー・ラボラトリーズ・ライセンシング・コーポレーションの商標である。

40

【発明の概要】

【課題を解決するための手段】

【0010】

適応オーディオをレンダリングする方法の実施形態が記述される。該レンダリングは、チャンネル・ベースのオーディオ、オーディオ・オブジェクトおよび動的オブジェクトを含む入力オーディオを受領する段階であって、前記動的オブジェクトは低優先度動的オブジ

50

エクトの集合および高優先度動的オブジェクトの集合として分類される、段階と；前記チャンネル・ベースのオーディオ、前記オーディオ・オブジェクトおよび前記低優先度動的オブジェクトをオーディオ処理システムの第一のレンダリング・プロセッサにおいてレンダリングする段階と；前記高優先度動的オブジェクトを前記オーディオ処理システムの第二のレンダリング・プロセッサにおいてレンダリングする段階とを執行することによる。入力オーディオは、オーディオ・コンテンツおよびレンダリング・メタデータを含むオブジェクト・オーディオ・ベースのデジタル・ビットストリーム・フォーマットに従ってフォーマットされていてもよい。前記チャンネル・ベースのオーディオはサラウンドサウンド・オーディオ・ベッドを含み、前記オーディオ・オブジェクトは中間空間的フォーマットに準拠するオブジェクトを含む。前記低優先度動的オブジェクトおよび高優先度動的オブジェクトは、前記入力オーディオを含むオーディオ・コンテンツの作者、ユーザー選択された値および前記オーディオ処理システムによって実行される自動化されたプロセスのうちの一つによって定義されうる優先度閾値によって区別される。ある実施形態では、優先度閾値は、オブジェクト・オーディオ・メタデータ・ビットストリームにおいてエンコードされる。前記低優先度および高優先度のオーディオ・オブジェクトのオーディオ・オブジェクトの相対的な優先度はオブジェクト・オーディオ・メタデータ・ビットストリームにおけるそれぞれの位置によって決定されてもよい。

10

【0011】

ある実施形態では、本方法はさらに、前記第一のレンダリング・プロセッサにおいて前記チャンネル・ベースのオーディオ、前記オーディオ・オブジェクトおよび前記低優先度動的オブジェクトをレンダリングしてレンダリングされたオーディオを生成する間またはその後、前記高優先度オーディオ・オブジェクトを前記第一のレンダリング・プロセッサを通して前記第二のレンダリング・プロセッサに渡し；前記レンダリングされたオーディオをスピーカー・システムへの伝送のために後処理することを含む。後処理段階は、アップミックス、ボリューム制御、等化、低音管理および前記スピーカー・システムを通じた再生のための前記入力オーディオに存在している高さ手がかりのレンダリングを容易にするための仮想化段階のうち少なくとも一つを含む。

20

【0012】

ある実施形態では、前記スピーカー・システムは、単一の軸に沿って音を送出する複数の共位置のドライバーを有するサウンドバー・スピーカーを有しており、前記第一および第二のレンダリング・プロセッサは、伝送リンクを通じて一緒に結合された別個のデジタル信号処理回路において具現される。優先度閾値は、前記第一および第二のレンダリング・プロセッサの相対的な処理機能、前記第一および第二のレンダリング・プロセッサのそれぞれに関連付けられたメモリ帯域幅および前記伝送リンクの伝送帯域幅のうち少なくとも一つによって決定される。

30

【0013】

実施形態はさらに、適応オーディオをレンダリングする方法であって、該レンダリングは、オーディオ・コンポーネントおよび関連付けられたメタデータを含む入力オーディオ・ビットストリームを受領する段階であって、前記オーディオ・コンポーネントはそれぞれチャンネル・ベースのオーディオ、オーディオ・オブジェクトおよび動的オブジェクトから選択されるオーディオ型をもつ、段階と；各オーディオ・コンポーネントについてのデコード・フォーマットをそれぞれのオーディオ型に基づいて決定する段階と；各オーディオ・コンポーネントの優先度を、該各オーディオ・コンポーネントに関連付けられたメタデータにおける優先度フィールドから決定する段階と；第一のレンダリング・プロセッサにおいて第一の優先度型のオーディオ・コンポーネントをレンダリングする段階と；第二のレンダリング・プロセッサにおいて第二の優先度型のオーディオ・コンポーネントをレンダリングする段階とを執行することによる、方法に向けられる。前記第一のレンダリング・プロセッサおよび第二のレンダリング・プロセッサは、伝送リンクを通じて互いに結合された別個のレンダリング・デジタル信号プロセッサ(DSP)として実装される。第一の優先度型のオーディオ・コンポーネントは低優先度の動的オブジェクトを含み、第二の

40

50

優先度型のオーディオ・コンポーネントは高優先度の動的オブジェクトを含み、本方法はさらに、前記チャンネル・ベースのオーディオ、前記オーディオ・オブジェクトを前記第一のレンダリング・プロセッサにおいてレンダリングすることを含む。ある実施形態では、前記チャンネル・ベースのオーディオはサラウンドサウンド・オーディオ・ベッドを含み、前記オーディオ・オブジェクトは中間空間的フォーマット（ISF）に準拠するオブジェクトを含み、前記低優先度および高優先度動的オブジェクトはオブジェクト・オーディオ・メタデータ（OAMD）フォーマットに準拠するものを含む。各オーディオ・コンポーネントについてのデコーダ・フォーマットは：OAMDフォーマットされた動的オブジェクト、サラウンドサウンド・オーディオ・ベッドおよびISFオブジェクトのうちの少なくとも一つを生成する。本方法はさらに、前記スピーカー・システムを通じた再生のための前記入力オーディオに存在している高さ手がかりのレンダリングを容易にするよう、少なくとも前記高優先度動的オブジェクトに仮想化プロセスを適用してもよく、スピーカー・システムは、単一の軸に沿って音を送出する複数の共位置のドライバーを有するサウンドバー・スピーカーを有していてもよい。

10

【0014】

実施形態はさらに、上述した方法を実装するデジタル信号処理システムおよび/または上述した方法の少なくとも一部を実装する回路を組み込むスピーカー・システムに向けられる。

【0015】

参照による組み込み

本明細書において言及される各刊行物、特許および/または特許出願はここに参照によって、個々の各刊行物および/または特許出願が具体的かつ個別的に参照によって組み込まれることが示されている場合と同じ程度にその全体において組み込まれる。

20

【図面の簡単な説明】

【0016】

以下の図面では、同様の参照符号が同様の要素を指すために使われる。以下の図はさまざまな例を描いているが、前記一つまたは複数の実装は図面に描かれる例に限定されるものではない。

【図1】高さチャンネルの再生のための高さスピーカーを提供するサラウンド・システム（たとえば9.1サラウンド）における例示的なスピーカー配置を示す図である。

30

【図2】ある実施形態のもとでの、適応的なオーディオ混合を生成するためのチャンネルおよびオブジェクト・ベースのデータの組み合わせを示す図である。

【図3】ある実施形態のもとでの、ハイブリッドの優先度に基づくレンダリング・システムにおいて処理されるオーディオ・コンテンツの型を示す表である。

【図4】ある実施形態のもとでの、ハイブリッドの優先度に基づくレンダリング戦略を実装するマルチプロセッサ・レンダリング・システムのブロック図である。

【図5】ある実施形態のもとでの、図4のマルチプロセッサ・レンダリング・システムの、より詳細なブロック図である。

【図6】ある実施形態のもとでの、サウンドバーを通じて適応オーディオ・コンテンツの再生のために優先度に基づくレンダリングを実装する方法を示すフローチャートである。

40

【図7】ハイブリッドの優先度に基づくレンダリング・システムの実施形態とともに使用されうるサウンドバー・スピーカーを示す図である。

【図8】例示的なテレビジョンおよびサウンドバー消費者使用事例における優先度に基づく適応オーディオ・レンダリング・システムの使用を示す図である。

【図9】例示的なフル・サラウンドサウンド家庭環境における優先度に基づく適応オーディオ・レンダリング・システムの使用を示す図である。

【図10】ある実施形態のもとでの、サウンドバーについて優先度に基づくレンダリングを利用する適応オーディオ・システムにおける使用のためのいくつかの例示的なメタデータ定義を示す表である。

【図11】いくつかの実施形態のもとでの、レンダリング・システムと一緒に使う中間空

50

間的フォーマットを示す図である。

【図12】ある実施形態のもとでの、中間空間的フォーマットと一緒に使うための積層環フォーマット・パン空間における環の配置を示す図である。

【図13】ある実施形態のもとでの、ISF処理システムにおいて使うための、諸スピーカーの弧を、ある角度にパンされたオーディオ・オブジェクトとともに示す図である。

【図14】A～Cは、異なる実施形態のもとでの、積層環中間空間的フォーマットのデコードを示す図である。

【発明を実施するための形態】

【0017】

オブジェクト・オーディオ・メタデータ(OAMD)ベッドまたは中間空間的フォーマット(ISF)オブジェクトが第一のDSPコンポーネント上の時間領域オブジェクト・オーディオ・レンダラー(OAR)コンポーネントを使ってレンダリングされ、一方、OAMD動的オブジェクトは第二のDSPコンポーネント上の後処理チェーンにおける仮想レンダラーによってレンダリングされるハイブリッドの優先度に基づくレンダリング戦略のためのシステムおよび方法が記述される。出力オーディオは、一つまたは複数の後処理および仮想化技法によってサウンドバー・スピーカーを通じた再生のために最適化されてもよい。本稿に記載される一つまたは複数の実施形態の諸側面は、ソフトウェア命令を実行する一つまたは複数のコンピュータまたは処理装置を含む混合、レンダリングおよび再生システムにおいて源オーディオ情報を処理するオーディオまたはオーディオビジュアル・システムにおいて実装されうる。記載される実施形態はいずれも、単独でまたは任意の組み合わせにおいて互いと一緒に使用されうる。さまざまな実施形態が、本明細書の一つまたは複数の場所で論じられるまたは暗示されることがありうる従来技術でのさまざまな欠点によって動機付けられていることがありうるが、それらの実施形態は必ずしもこれらの欠点のいずれかに取り組むものではない。つまり、種々の実施形態は本明細書において論じられることがある種々の欠点に取り組むことがある。いくつかの実施形態は、本明細書において論じられることがあるいくつかの欠点または一つだけの欠点に部分的に取り組むだけであることがあり、いくつかの実施形態はこれらの欠点のどれにも取り組まないこともある。

【0018】

本記述の目的のためには、以下の用語は関連付けられた意味をもつ：用語「チャンネル」は、オーディオ信号にメタデータを加えたものを意味する。メタデータにおいて、位置はチャンネル識別子、たとえば左前方または右上方サラウンドとして符号化される。「チャンネル・ベースのオーディオ」は、関連付けられた公称位置をもつスピーカー・ゾーンのあらかじめ定義されたセット、たとえば5.1、7.1などを通じた再生のためにフォーマットされたオーディオである。用語「オブジェクト」または「オブジェクト・ベースのオーディオ」は、見かけの源位置(たとえば3D座標)、見かけの源幅などといったパラメトリックな源記述をもつ一つまたは複数のオーディオ・チャンネルを意味する。「適応オーディオ」は、チャンネル・ベースのおよび/またはオブジェクト・ベースのオーディオ信号に、オーディオ・ストリームに位置が空間内の3D位置として符号化されているメタデータを加えたものを使って、再生環境に基づいてオーディオ信号をレンダリングするメタデータを加えたものを意味する。「聴取環境」は、任意の開けた、部分的に囲まれたまたは完全に囲まれた領域、たとえば部屋であって、オーディオ・コンテンツを単独でまたはビデオまたは他のコンテンツと一緒に再生するために使用できる領域を意味し、自宅、映画館、シアター、講堂、スタジオ、ゲーム・コンソールなどにおいて具現されることができ、そのような領域は、壁またはバツフルのような、そこに配置された一つまたは複数の表面を有していてもよく、それが音波を直接または拡散的に反射する。

【0019】

適応的なオーディオ・フォーマットおよびシステム

ある実施形態では、相互接続システムは、「空間的オーディオ・システム」または「適応オーディオ・システム」と称されうる音フォーマットおよび処理システムとともに機能するよう構成されているオーディオ・システムの一部として実装される。そのようなシス

10

20

30

40

50

テムは、向上した聴衆没入感、より大きな芸術的制御ならびにシステム柔軟性およびスクレーパリティを許容するためのオーディオ・フォーマットおよびレンダリング技術に基づく。全体的な適応オーディオ・システムは一般に、通常のチャンネル・ベースのオーディオ要素およびオーディオ・オブジェクト符号化要素の両方を含む一つまたは複数のビットストリームを生成するよう構成されたオーディオ・エンコード、配送およびデコード・システムを含む。そのような組み合わせられたアプローチは、別個に実施されるチャンネル・ベースまたはオブジェクト・ベースのアプローチのいずれと比べても、より大きな符号化効率およびレンダリング柔軟性を提供する。

【 0 0 2 0 】

適応オーディオ・システムおよび関連するオーディオ・フォーマットの例示的実装は、ドルビー（登録商標）・アトモス（商標）・プラットフォームである。そのようなシステムは、9.1サラウンド・システムまたは同様のサラウンドサウンド構成として実装されてもよい高さ（上下）次元を組み込む。図1は、高さチャンネルの再生のための高さスピーカーを提供する現在のサラウンド・システム（たとえば9.1サラウンド）におけるスピーカー配置を示している。9.1システム100のスピーカー構成は、床面における五つのスピーカー102および高さ面における四つのスピーカー104から構成される。一般に、これらのスピーカーは、室内で多少なりとも正確に任意の位置から発するよう設計された音を生じるために使用されうる。図1に示されるようなあらかじめ定義されたスピーカー構成は、当然ながら、所与の音源の位置を正確に表現する能力を制限することがある。たとえば、音源は左スピーカー自身よりさらに左にパンされることはできない。これはすべてのスピーカーにあてはまり、よってダウンミックスがその中に制約される一次元（たとえば左右）、二次元（たとえば前後）または三次元（たとえば左右、前後、上下）の幾何形状をなす。そのようなスピーカー構成において、さまざまな異なるスピーカー構成および型が使用されうる。たとえば、ある種の向上されたオーディオ・システムは、9.1、11.1、13.1、19.4または他の構成にあるスピーカーを使ってもよい。スピーカー型はフルレンジ直接スピーカー、スピーカー・アレイ、サラウンド・スピーカー、サブウーファー、ツイーターおよび他の型のスピーカーを含みうる。

【 0 0 2 1 】

オーディオ・オブジェクトは、聴取環境における特定の物理的位置（単数または複数）から発するよう知覚されうる音要素の群と考えることができる。そのようなオブジェクトは静的（すなわち定常）または動的（すなわち動いている）であることができる。オーディオ・オブジェクトは、他の機能とともに所与の時点における音の位置を定義するメタデータによって制御される。オブジェクトが再生されるとき、オブジェクトは、必ずしもあらかじめ定義された物理チャンネルに出力されるのではなく、位置メタデータに従って、存在している諸スピーカーを使ってレンダリングされる。セッションにおけるトラックはオーディオ・オブジェクトであることができ、標準的なパン・データは位置メタデータに似ている。このように、スクリーン上に配置されたコンテンツはチャンネル・ベースのコンテンツと同じ仕方で効果的にパンしうるが、サラウンドに配置されたコンテンツは望むなら個別のスピーカーにレンダリングされることができる。オーディオ・オブジェクトの使用が離散的な諸効果についての所望される制御を提供する一方、サウンドトラックの他の側面がチャンネル・ベースの環境において効果的に機能しうる。たとえば、多くの周囲効果または残響は、スピーカーのアレイに供給されることから実際に裨益する。これらはアレイを満たすために十分な幅をもつオブジェクトとして扱われることができるが、いくつかのチャンネル・ベースの機能を保持することが有益である。

【 0 0 2 2 】

適応オーディオ・システムは、オーディオ・オブジェクトに加えてオーディオ・ベッドをサポートするよう構成されている。ここで、ベッドとは、事実上、チャンネル・ベースのサブミックスまたはステムである。これらは、コンテンツ・クリエイターの意図に依存して、個々に、あるいは単一のベッドに組み合わせられて、最終的な再生（レンダリング）のために送達されることができる。これらのベッドは、5.1、7.1および9.1ならびに図1に

10

20

30

40

50

示したような頭上スピーカーを含むアレイのような、異なるチャンネル・ベースの構成で生成されることができる。図2は、ある実施形態のもとでの、適応的なオーディオ混合を生成するための、チャンネルおよびオブジェクト・ベースのデータの組み合わせを示している。プロセス200において示されるように、たとえばパルス符号変調された(PCM)データの形で提供された5.1または7.1サラウンドサウンド・データでありうるチャンネル・ベースのデータ202が、オーディオ・オブジェクト・データ204と組み合わせられて、適応オーディオ混合208を生成する。オーディオ・オブジェクト・データ204は、もとのチャンネル・ベースのデータを、オーディオ・オブジェクトの位置に関するある種のパラメータを指定する関連するメタデータと組み合わせることによって生成される。図2に概念的に示されるように、オーサリング・ツールは、スピーカー・チャンネル・グループおよびオブジェクト・チャンネルの組み合わせを同時に含むオーディオ・プログラムを生成する能力を提供する。たとえば、オーディオ・プログラムは、任意的にグループ(またはトラック、たとえばステレオまたは5.1トラック)に編成されている一つまたは複数のスピーカー・チャンネルと、一つまたは複数のスピーカー・チャンネルについての記述メタデータと、一つまたは複数のオブジェクト・チャンネルと、一つまたは複数のオブジェクト・チャンネルについての記述メタデータとを含むことができる。

10

【0023】

ある実施形態では、図2のベッドおよびオブジェクト・オーディオ・コンポーネントは、特定のフォーマット標準に準拠するコンテンツを含んでいてもよい。図3は、ある実施形態のもとでの、ハイブリッドの優先度に基づくレンダリング・システムにおいて処理されるオーディオ・コンテンツの型を示す表である。図3のテーブル300に示されるように、コンテンツの二つの主要な型がある。軌跡に関して比較的静的であるチャンネル・ベースのコンテンツと、システムにおいてスピーカーまたはドライバーの間で動く動的なコンテンツである。チャンネル・ベースのコンテンツはOAMDベッドにおいて具現されてもよく、動的なコンテンツは、少なくとも二つの優先度レベル、すなわち低優先度および高優先度に優先順位付けされるOAMDオブジェクトである。動的なオブジェクトはある種のフォーマット・パラメータに従ってフォーマットされてもよく、ISFオブジェクトのようなある種の型のオブジェクトとして分類されてもよい。ISFフォーマットは本稿でのちにより詳細に述べる。

20

【0024】

動的オブジェクトの優先度は、コンテンツ型(たとえばダイアログか効果か周囲音(ambient sound)か)、処理要件、メモリ要件(たとえば高帯域幅か低帯域幅か)および他の同様の特性といった、オブジェクトのある種の特性を反映する。ある実施形態では、各オブジェクトの優先度はあるスケールに沿って定義され、オーディオ・オブジェクトをカプセル化するビットストリームの一部として含まれる優先度フィールドにおいてエンコードされる。優先度は1(最低)から10(最高)の整数値のようなスカラー値として、あるいは二値フラグ(0低/1高)として設定されてもよく、あるいは他の同様のエンコード可能な優先度設定機構でもよい。優先度レベルは一般に、オブジェクト毎に一度、コンテンツ作者によって設定される。コンテンツ作者は、上述した特性の一つまたは複数に基づいて各オブジェクトの優先度を決定してもよい。

30

40

【0025】

代替的な実施形態では、前記オブジェクトのうち少なくともいくつかのオブジェクトの優先度レベルはユーザーによって、あるいは自動化された動的プロセスを通じて設定されてもよい。該プロセスは、動的プロセッサ負荷、オブジェクト・ラウドネス、環境変化、システム障害、ユーザー選好、音響的な調整などといったある種のランタイムの基準に基づいてオブジェクトのデフォルト優先度レベルを修正してもよい。

【0026】

ある実施形態では、動的オブジェクトの優先度レベルは、マルチプロセッサ・レンダリング・システムにおけるオブジェクトの処理を決定する。各オブジェクトのエンコードされた優先度レベルは、デュアルまたはマルチDSPシステムのどのプロセッサ(DSP)がそ

50

の特定のオブジェクトをレンダリングするために使われるかを決定するためにデコードされる。これは、優先度に基づくレンダリング戦略が、適応オーディオ・コンテンツをレンダリングすることにおいて使用されることができるようになる。図4は、ある実施形態のもとでの、ハイブリッドの優先度に基づくレンダリング戦略を実装するためのマルチプロセッサ・レンダリング・システムのブロック図である。図4は、二つのDSPコンポーネント406および410を含むマルチプロセッサ・レンダリング・システム400を示している。二つのDSPは二つの別個のレンダリング・サブシステム、すなわちデコード/レンダリング・コンポーネント404およびレンダリング/後処理コンポーネント408内に含まれる。これらのレンダリング・サブシステムは一般に、オーディオがさらなる後処理および/または増幅およびスピーカ一段に送られる前に、レガシーの、オブジェクトおよびチャンネル・オーディオ・デコード、オブジェクト・レンダリング、チャンネル再マッピングおよび信号処理を実行する処理ブロックを含む。

10

【0027】

システム400は、入力オーディオをデジタル・ビットストリーム402としてエンコードする一つまたは複数の捕捉、前処理、オーサリングおよび符号化コンポーネントを通じて生成されるオーディオ・コンテンツをレンダリングおよび再生するよう構成される。適応オーディオ・コンポーネントは、源分離およびコンテンツ型のような因子を調べることによる入力オーディオの解析を通じて適切なメタデータを自動的に生成するために使われてもよい。たとえば、チャンネル対の間の相関付けられた入力の相対的なレベルの解析を通じてマルチチャンネル記録から位置メタデータが導出されてもよい。発話または音楽といったコンテンツ型の検出はたとえば特徴抽出および分類によって達成されてもよい。ある種のオーサリング・ツールは、サウンドエンジニアの創造的な意図の入力およびコード化を最適化し、それによりひとたびそれが事実上任意の再生環境における再生のために最適化されたらサウンドエンジニアが最終的なオーディオ混合を作り出せるようにすることによって、オーディオ・プログラムのオーサリングを許容する。これは、オーディオ・オブジェクトと、もとのオーディオ・コンテンツに関連付けられ、それと一緒にエンコードされている位置データとの使用を通じて達成できる。ひとたび適応オーディオ・コンテンツがオーサリングされて適切なコーデック装置において符号化されたら、それはスピーカ414を通じた再生のためにデコードされ、レンダリングされる。

20

【0028】

図4に示されるように、オブジェクト・メタデータを含むオブジェクト・オーディオおよびチャンネル・メタデータを含むチャンネル・オーディオが入力オーディオ・ビットストリームとしてデコード/レンダリング・サブシステム404内の一つまたは複数のデコーダ回路に入力される。入力オーディオ・ビットストリーム402は、図3に示されるような、OAMDベッド、低優先度動的オブジェクトおよび高優先度動的オブジェクトを含むさまざまなオーディオ・コンポーネントに関するデータを含んでいる。各オーディオ・オブジェクトに割り当てられた優先度が、二つのDSP 406または410のうちのどちらがその特定のオブジェクトに対してレンダリング・プロセスを実行するかを決定する。OAMDベッドおよび低優先度オブジェクトはDSP 406 (DSP1) においてレンダリングされ、一方、高優先度オブジェクトはDSP 410 (DSP2) でのレンダリングのためにレンダリング・サブシステム404を素通しにされる。次いで、レンダリングされたベッド、低優先度オブジェクトおよび高優先度オブジェクトはサブシステム408内の後処理コンポーネント412に入力されて、スピーカ414を通じた再生のために伝送される出力オーディオ信号413を生成する。

30

40

【0029】

ある実施形態では、低優先度オブジェクトを高優先度オブジェクトから区別する優先度レベルは、それぞれの関連付けられたオブジェクトについてのメタデータをエンコードするビットストリームの優先度内に設定されている。低優先度と高優先度の間のカットオフまたは閾値は優先度範囲に沿ったある値、たとえば1から10の優先度スケールに沿った値5または7、あるいは二値の優先度フラグ0または1についての単純なディテクターとして

50

設定されてもよい。各オブジェクトについての優先度レベルは、各オブジェクトをレンダリングするために適切なDSP（DSP1またはDSP2）にルーティングするために、デコード・サブシステム402内の優先度決定コンポーネントにおいてデコードされてもよい。

【0030】

図4のマルチプロセッシング・アーキテクチャは、DSPの特定の構成および機能ならびにネットワークおよびプロセッサ・コンポーネントの帯域幅/処理機能に基づいて、種々の型の適応オーディオ・ベッドおよびオブジェクトの効率的な処理を容易にする。ある実施形態では、DSP1はOAMDベッドおよびISFオブジェクトをレンダリングするために最適化されるが、OAMD動的オブジェクトを最適にレンダリングするには構成されないこともある。一方、DSP2はOAMD動的オブジェクトをレンダリングするために最適化される。この応用については、入力オーディオにおけるOAMD動的オブジェクトは高優先度レベルを割り当てられ、それによりレンダリングのためにDSP2へと素通しにされる。一方、ベッドおよびISFオブジェクトはDSP1においてレンダリングされる。これは、最もよくレンダリングできる適切なDSPがオーディオ・コンポーネント（単数または複数）をレンダリングすることを許容する。

【0031】

レンダリングされるオーディオ・コンポーネントの型（すなわちベッド/ISFオブジェクトかOAMD動的オブジェクトか）に加えてまたはその代わりに、オーディオ・コンポーネントのルーティングおよび分散式のレンダリングは、ある種のパフォーマンスに関係した指標、たとえば前記二つのDSPの相対的な処理機能および/または前記二つのDSPの間の伝送ネットワークの帯域幅に基づいて実行されてもよい。こうして、一方のDSPが他方のDSPより著しく強力であり、ネットワーク帯域幅がレンダリングされていないオーディオ・データを伝送するのに十分であれば、より強力なほうのDSPが前記オーディオ・コンポーネントのうちより多くをレンダリングするために頼られるよう優先度レベルが設定されてもよい。たとえばDSP2がDSP1よりずっと強力であれば、DSP2がOAMD動的オブジェクトのすべてを、あるいは他の型のオブジェクトをレンダリングできるとすればフォーマットに関わりなくすべてのオブジェクトを、レンダリングするよう構成されてもよい。

【0032】

ある実施形態では、オブジェクト優先度レベルの動的な変更を許容するために、ある種の用途固有のパラメータ、たとえば部屋構成情報、ユーザー選択、処理/ネットワーク制約条件などがオブジェクト・レンダリング・システムにフィードバックされてもよい。すると、優先順位付けされたオーディオ・データは、スピーカー414を通じた再生のための出力に先立って、等化器およびリミッターといった一つまたは複数の信号処理段を通じて処理される。

【0033】

システム400は適応オーディオのための再生システムの例を表わしているのであって、他の構成、コンポーネントおよび相互接続も可能であることを注意しておくべきである。たとえば、二つの型の優先度に区分された動的オブジェクトを処理するために図3においては二つのレンダリングDSPが示されている。より大きな処理パワーおよびより多くの優先度レベルのために追加的な数のDSPも含まれてもよい。こうして、N個の異なる優先度の区別のためにN個のDSPが使用されることができる。たとえば、高、中、低の優先度レベルについての三つのDSPなどである。

【0034】

ある実施形態では、図4に示されるDSP 406および410は、物理的な伝送インターフェースまたはネットワークによって一緒に結合された別個の装置として実装されている。DSPはそれぞれ別個のコンポーネントまたはサブシステム、たとえば図のようなサブシステム404および408内に含まれてもよく、あるいは同じサブシステム、たとえば統合されたデコーダ/レンダラー・コンポーネントに含まれる別個のコンポーネントであってもよい。あるいはまた、DSP 406および410は、モノリシックな集積回路デバイス内の別個の処理コンポーネントであってもよい。

10

20

30

40

50

【 0 0 3 5 】

例示的実装

上述したように、適応オーディオ・フォーマットの初期の実装は、新規なオーサリング・ツールを使ってオーサリングされ、適応的なオーディオ・シネマ・エンコーダを使ってパッケージングされ、PCMもしくは既存のデジタル映画館イニシアチブ（DCI: Digital Cinema Initiative）頒布機構を使う独自の無損失コーデックを使って頒布されるコンテンツ・キャプチャー（オブジェクトおよびチャンネル）を含むデジタル映画館コンテキストにおいてであった。この場合、オーディオ・コンテンツはデジタル映画館においてデコードされ、レンダリングされて、没入的な空間的オーディオ映画館体験を作り出すことが意図される。しかしながら、今不可欠なのは、適応オーディオ・フォーマットによって提供される向上したユーザー経験を、自宅にいる消費者に直接届けることである。これは、フォーマットおよびシステムのある種の特徴が、より制限された聴取環境での使用のために適応されることを要求する。説明の目的のため、用語「消費者ベースの環境」は、家、スタジオ、部屋、コンソール・エリア、講堂などといった通常の消費者またはプロフェッショナルによる使用のための聴取環境を含む、任意の映画館ではない環境を含むことが意図されている。

10

【 0 0 3 6 】

消費者オーディオのための現在のオーサリングおよび頒布システムは、オーディオ・エッセンス（すなわち、消費者再生システムによって再生される実際のオーディオ）において伝達されるコンテンツの型の限られた知識でのあらかじめ定義された固定されたスピーカー位置への再生のために意図されたオーディオを生成し、送達する。しかしながら、適応オーディオ・システムは、固定されたスピーカー位置固有のオーディオ（左チャンネル、右チャンネルなど）と位置、サイズおよび測度を含む一般化された3D空間情報を有するオブジェクト・ベースのオーディオ要素との両方についてのオプションを含むオーディオ生成への新たなハイブリッド・アプローチを提供する。このハイブリッド・アプローチは、（固定したスピーカー位置によって提供される）忠実性とレンダリングにおける柔軟性（一般化されたオーディオ・オブジェクト）とのためのバランスの取れたアプローチを提供する。このシステムは、コンテンツ生成/オーサリングの時点でコンテンツ作成者によってオーディオ・エッセンスと対にされた新たなメタデータを介してオーディオ・コンテンツについての追加的な有用な情報をも提供する。この情報は、レンダリングの間に使用できる前記オーディオの属性についての詳細な情報を提供する。そのような属性はコンテンツ型（たとえばダイアログ、音楽、効果、効果音（Foley）、背景音/周囲音等）ならびにオーディオ・オブジェクト情報、たとえば空間的属性（たとえば3D位置、オブジェクト・サイズ、速度など）および有用なレンダリング情報（たとえば、スピーカー位置にスナップ、チャンネル重み、利得、ベース〔低音〕管理情報など）を含みうる。オーディオ・コンテンツおよび再生意図メタデータは、コンテンツ作成者によって手動で作成されるか、あるいはオーサリング・プロセスの間にバックグラウンドで実行できる自動的なメディア・インテリジェンス・アルゴリズムの使用を通じて生成されて望むなら最終的な品質管理フェーズの間にコンテンツ作成者によって確認されることができる。

20

30

【 0 0 3 7 】

図5は、チャンネルおよびオブジェクト・ベースのコンポーネントという異なる型をレンダリングするための優先度に基づくレンダリング・システムのブロック図であり、図4に示したシステムの、より詳細な図である。図5に示されるように、システム500は、ハイブリッドのオブジェクト・ストリーム（単数または複数）およびチャンネル・ベースのオーディオ・ストリーム（単数または複数）両方を担持するエンコードされたビットストリーム506を処理する。ビットストリームは、レンダリング/信号処理ブロック502および504によって処理され、これらはそれぞれ別個のDSP装置を表わすまたはそれによって実装される。これらの処理ブロックにおいて実行されるレンダリング機能は、適応オーディオのためのさまざまなレンダリング・アルゴリズムおよびアップミックスなどといったある種の後処理アルゴリズムを実装する。

40

50

【0038】

優先度に基づくレンダリング・システム500は、デコード/レンダリング段502およびレンダリング/後処理段504という二つの主要なコンポーネントを有する。入力オーディオ506はHDMI (high-definition multimedia interface [高精細度マルチメディア・インターフェース]) を通じてデコード/レンダリング段に与えられる。ただし、他のインターフェースも可能である。ビットストリーム検出コンポーネント508は前記ビットストリームをパースして、異なるオーディオ・コンポーネントを、ドルビー・デジタル・プラス・デコーダ、MAT2.0デコーダ、トゥルーHDデコーダなどといった適切なデコーダに差し向ける。それらのデコーダは、OAMDベッド信号およびISFもしくはOAMD動的オブジェクトといったさまざまなフォーマットされたオーディオ信号を生成する。

10

【0039】

デコード/レンダリング段502はOAR (object audio renderer [オブジェクト・オーディオ・レンダラー]) インターフェース510を含み、これはOAMD処理コンポーネント512、OARコンポーネント514および動的オブジェクト抽出コンポーネント516を含む。動的抽出ユニット516はデコーダ全部からの出力を受け、ベッドおよびISFオブジェクトをもしあれば低優先度動的オブジェクトとともに、高優先度動的オブジェクトから分離する。ベッド、ISFオブジェクトおよび低優先度動的オブジェクトはOARコンポーネント514に送られる。図示した例示の実施形態については、OARコンポーネント514はプロセッサ(たとえばDSP)回路502のコアを表わし、固定の5.1.2チャンネル出力フォーマット(たとえば標準的な5.1+二つの高さチャンネル)にレンダリングする。ただし、7.1.4など、他のサラウンドサウンドに高さを加えた構成も可能である。OARコンポーネント514からのレンダリングされた出力513は次いで、レンダリング/後処理段504のデジタル・オーディオ・プロセッサ(DAP)コンポーネントに伝送される。この段は、アップミックス、レンダリング/仮想化、ボリューム制御、等化、低音管理および他の可能な機能といった機能を実行する。段504からの出力522はある例示の実施形態では5.1.2スピーカー・フィードを有する。段504は、プロセッサ、DSPまたは同様の装置といったいかなる適切な処理回路として実装されてもよい。

20

【0040】

ある実施形態では、出力信号522はサウンドバーまたはサウンドバー・アレイに伝送される。図5に示したような特定の使用事例については、二つの段502と504の間のメモリ帯域幅をおとしめることなく、31.1オブジェクトをもつMAT 2.0入力の使用事例をサポートするために、サウンドバーも優先度に基づくレンダリング戦略を用いる。ある例示の実装では、メモリ帯域幅は、最大32個のオーディオ・チャンネルについて48kHzで外部メモリから読まれるまたは書き込まれることを許容する。OARコンポーネント514の5.1.2チャンネル・レンダリングされた出力513のためには8個のチャンネルが必要とされるので、最大で24個のOAMD動的オブジェクトが後処理チェーン504において仮想レンダラーによってレンダリングされうる。24個より多いOAMD動的オブジェクトが入力ストリーム506に存在する場合には、追加的な低優先度オブジェクトが第一段502でOARコンポーネント514によってレンダリングされる必要がある。動的オブジェクトの優先度は、OAMDストリームにおけるその位置に基づいて決定される(たとえば最高優先度のオブジェクトが最初、最低優先度のオブジェクトが最後)。

30

40

【0041】

図4および図5の実施形態は、OAMDおよびISFフォーマットに準拠するベッドおよびオブジェクトとの関係で記述されているが、マルチプロセッサ・レンダリング・システムを使う優先度に基づくレンダリング方式は、チャンネル・ベースのオーディオおよび二つ以上の型のオーディオ・オブジェクトを含む任意の型の適応オーディオ・コンテンツとともに使用されることができる。ここで、オブジェクト型は相対的な優先度レベルに基づいて区別できる。適切なレンダリング・プロセッサ(たとえばDSP)は、オーディオ・オブジェクト型および/またはチャンネル・ベースのオーディオ・コンポーネントの全部またはただ一つの型を最適にレンダリングするよう構成されうる。

50

【 0 0 4 2 】

図5のシステム500は、チャンネル・ベースのベッド、ISFオブジェクトおよびOAMD動的オブジェクトに関わる個別的なレンダリング・アプリケーションならびにサウンドバーを通じた再生のためのレンダリングとともに機能するようOAMDオーディオ・フォーマットを適応させるレンダリング・システムを示している。システムは、サウンドバーまたは同様の共位置のスピーカー・システムを通じて適応オーディオ・コンテンツを再現することに関するある種の実装上の複雑さ問題に対処する優先度に基づくレンダリング戦略を実装する。図6は、ある実施形態のもとでの、サウンドバーを通じた適応オーディオ・コンテンツの再生のための優先度に基づくレンダリングを実装する方法を示すフローチャートである。図6のプロセス600は概括的には、図5の優先度に基づくレンダリング・システム500において実行される方法段階を表わしている。入力オーディオ・ビットストリームを受信後、チャンネル・ベースのベッドおよび種々のフォーマットのオーディオ・オブジェクトを含むオーディオ・コンポーネントがデコードのために適切なデコーダ回路に入力される(602)。オーディオ・オブジェクトは、異なるフォーマット方式を使ってフォーマットされていてもよく、各オブジェクトと一緒にエンコードされる相対的な優先度に基づいて区別(604)されうる動的オブジェクトを含む。プロセスは、定義された優先度閾値と比較しての各動的オーディオ・オブジェクトの優先度レベルを、そのオブジェクトについてビットストリーム内の適切なメタデータ・フィールドを読むことによって決定する。低優先度オブジェクトを高優先度オブジェクトから区別する優先度閾値は、コンテンツ作成者によって設定された固定構成値としてシステムにプログラムされていてもよく、あるいはユーザー入力、自動化された手段または他の適応機構によって動的に設定されてもよい。チャンネル・ベースのベッドおよび低優先度動的オブジェクトは、もしあればシステムの第一のDSPにおいてレンダリングされるべく最適化されたオブジェクトと一緒に、その第一のDSPにおいてレンダリングされる(606)。高優先度の動的オブジェクトは第二のDSPに渡され、そこでレンダリングされる(608)。レンダリングされたオーディオ・コンポーネントは次いで、サウンドバーまたはサウンドバー・アレイを通じた再生のために、ある種の任意的な後処理段階を通じて伝送される(610)。

10

20

【 0 0 4 3 】

サウンドバー実装

図4に示されるところでは、二つのDSPによって生成される優先順位付けされ、レンダリングされたオーディオ出力は、ユーザーへの再生のためにサウンドバーに伝送される。サウンドバー・スピーカーは、フラットスクリーン・テレビジョンの普及を受けて人気が増した。そのようなテレビジョンは非常に薄く、比較的軽くなってきており、可搬性および取り付けオプションが最適化され、それでいて手の出せる価格で増大し続ける画面サイズを提供している。しかしながら、これらのテレビジョンの音質は、スペース、電力およびコストの制約のため、しばしば非常に貧弱である。サウンドバーは、フラットパネル・テレビジョンの下に置かれてテレビジョン・オーディオの品質を改善するしばしばスタイリッシュな、電源付きスピーカーであり、それ自身で、あるいはサラウンドサウンド・スピーカー・セットアップの一部として使用できる。図7は、ハイブリッドの優先度に基づくレンダリング・システムの実施形態とともに使用されうるサウンドバー・スピーカーを示している。システム700において示されるように、サウンドバー・スピーカーは、いくつかのドライバー703を収容するキャビネット701を有する。これらのドライバーは、キャビネットの前面から直接、音を駆出するよう水平(または垂直)軸に沿って配列されている。サイズおよびシステム制約条件に依存して、いかなる実際的な数のドライバー701が使用されてもよく、典型的な数は2~6個の範囲のドライバーである。ドライバーは同じサイズおよび形であってもよく、あるいは異なるドライバーのアレイであってもよい。たとえばより低周波音のための、より大きな中央ドライバーなど。高精細度オーディオ・システムへの直接的なインターフェースを許容するために、HDMI入力インターフェース702が設けられる。

30

40

【 0 0 4 4 】

50

サウンドバー・システム700は、搭載電源または増幅がなく、最小限の受動回路をもつ受動スピーカー・システムであってもよい。キャビネット内に設置された、あるいは外部コンポーネントを通じて緊密に結合された一つまたは複数のコンポーネントをもつ電源付きのシステムであってもよい。そのような機能およびコンポーネントは電源および増幅704、オーディオ処理（たとえばEQ、低音制御など）706、A/Vサラウンドサウンド・プロセッサ708および適応オーディオ仮想化710を含む。本稿の目的のためには、用語「ドライバー」は電気的なオーディオ入力信号に応答して音を生じる単一の電気音響トランスデューサを意味する。ドライバーは、いかなる適切な型、幾何構成およびサイズで実装されてもよく、ホーン、コーン、リボン・トランスデューサなどを含みうる。用語「スピーカー」はユニット的なエンクロージャー内の一つまたは複数のドライバーを意味する。

10

【0045】

サウンドバー710のためのコンポーネント710において、あるいはレンダリング・プロセッサ504のコンポーネントとして提供される仮想化機能は、テレビジョン、コンピュータ、ゲーム・コンソールまたは同様のデバイスといった局所化されたアプリケーションにおける適応オーディオ・システムの実装を許容するとともに、閲覧画面またはモニター表面に対応する平坦な面内に配置されたスピーカーを通じたこのオーディオの空間的な再生を許容する。図8は、例示的なテレビジョンおよびサウンドバー消費者使用事例における優先度に基づく適応オーディオ・レンダリング・システムの使用を示している。一般に、テレビジョン使用事例は、設備（テレビ・スピーカー、サウンドバー・スピーカーなど）のしばしば低下した品質および空間的分解能の点で限定されていることがある（たとえばサラウンドまたは後方スピーカーがない）スピーカー位置/構成（単数または複数）に基づいて、没入的な消費者体験を作り出すことに対して困難を呈する。図8のシステム800は、標準的なテレビジョンの左および右の位置にあるスピーカー（TV-LおよびTV-R）ならびに可能性としては任意的な左および右の上方発射ドライバー（TV-LHおよびTV-RH）を含んでいる。システムは図7に示したサウンドバー700をも含んでいる。先述したように、テレビジョン・スピーカーのサイズおよび品質は、コスト制約および設計選択に起因して、単独のまたは家庭シアター・スピーカーに比べて低下している。しかしながら、サウンドバー700との関連での動的仮想化の使用がこうした不足を克服する助けとなりうる。図8のサウンドバー700は、みなサウンドバー・キャビネットの水平軸に沿って配列された前方発射ドライバーおよび可能な側方発射ドライバーを有するものとして示されている。図8では、動的仮想化効果は、サウンドバー・スピーカーについて示されている。これにより、特定の聴取位置804にいる人々は、水平面内で個々にレンダリングされる適切なオーディオ・オブジェクトに関連付けられた水平要素を聞くことになる。適切なオーディオ・オブジェクトに関連付けられた高さ要素が、適応オーディオ・コンテンツによって与えられるオブジェクト空間情報に基づいたスピーカー仮想化アルゴリズム・パラメータの動的制御を通じてレンダリングされてもよい。少なくとも部分的に没入的なユーザー体験を提供するためである。サウンドバーの共位置のスピーカーについては、この動的仮想化は、部屋の辺に沿って動くオブジェクトの知覚または他の水平面音軌跡効果を作り出すために使用されてもよい。これは、サラウンド・スピーカーや後方スピーカーがないために普通なら存在しない空間手がかりをサウンドバーが提供することを許容する。

20

30

40

【0046】

ある実施形態では、サウンドバー700は、高さ手がかりを提供する仮想化アルゴリズムを許容するために音の反射を利用する上方発射ドライバーのような、共位置でないドライバーを含んでいてもよい。ドライバーのうちあるものは、他のドライバーとは異なる方向に音を放射するよう構成されてもよい。たとえば、一つまたは複数のドライバーが別個に制御される音ゾーンをもつ操縦可能な音ビームを実装してもよい。

【0047】

ある実施形態では、サウンドバー700は高さスピーカーまたは高さ対応の床置きスピー

50

ーカーをもつフル・サラウンドサウンド・システムの一部として使われてもよい。そのような実装は、サウンドバー仮想化がサラウンド・スピーカー・アレイによって提供される没入的な音を增強することを許容する。図9は、例示的なフル・サラウンドサウンド家庭環境における優先度に基づく適応的なオーディオ・レンダリング・システムの使用を示している。システム900において示されるように、テレビジョンまたはモニター802に付随するサウンドバー700は、図示した5.1.2構成のようなスピーカー904のサラウンドサウンド・アレイとの関連で使われる。この場合、サウンドバー700は、サラウンド・スピーカーを駆動し、レンダリングおよび仮想化プロセスの少なくとも一部を提供するためにA/Vサラウンドサウンド・プロセッサ708を含んでいてもよい。図9のシステムは、適応オーディオ・システムによって提供されうるコンポーネントおよび機能のほんの1つの可能なセットを示すものであり、ある種の側面はユーザーのニーズに基づいて低減または除去されてそれでいて向上された経験を提供することがありうる。

【0048】

図9は、サウンドバーによって提供されるものに加えて聴取環境において没入的なユーザー経験を提供するための動的スピーカー仮想化の使用を示している。それぞれの関連するオブジェクトについて別個の仮想化器が使われてもよく、組み合わせられた信号はLおよびRスピーカーに送られて多重オブジェクト仮想化効果を作り出すことができる。例として、LおよびRスピーカーについて動的仮想化効果が示されている。これらのスピーカーは、オーディオ・オブジェクトのサイズおよび位置情報と一緒に、拡散的なまたは点源のニアフィールド・オーディオ経験を作り出すために使用できる。同様の仮想化効果は、システム内の他のスピーカーの任意のものまたは全部に適用されることもできる。

【0049】

ある実施形態では、適応オーディオ・システムは、もとの空間的オーディオ・フォーマットからメタデータを生成するコンポーネントを含む。システム500の方法およびコンポーネントは、通常のチャンネル・ベースのオーディオ要素およびオーディオ・オブジェクト符号化要素の両方を含む一つまたは複数のビットストリームを処理するよう構成されたオーディオ・レンダリング・システムを有する。オーディオ・オブジェクト符号化要素を含む新たな拡張層が定義され、チャンネル・ベースのオーディオ・コーデック・ビットストリームまたはオーディオ・オブジェクト・ビットストリームのいずれか一方に加えられる。このアプローチは、拡張層を含むビットストリームが、既存のスピーカーおよびドライバ設計または個々にアドレッシング可能なドライバおよびドライバ定義を利用する次世代スピーカーと一緒に使うためのレンダラーによって処理されることができるようになる。空間的オーディオ・プロセッサからの空間的オーディオ・コンテンツは、オーディオ・オブジェクト、チャンネルおよび位置メタデータを有する。オブジェクトがレンダリングされるとき、オブジェクトは位置メタデータおよび再生スピーカーの位置に従って、サウンドバーまたはサウンドバー・アレイの一つまたは複数のドライバに割り当てられる。エンジニアの混合入力にตอบสนองしてオーディオ・ワークステーションにおいてメタデータが生成される。このメタデータは、空間的パラメータ（たとえば位置、測度、強度、音色など）を制御するレンダリング・キューを提供するとともに、展示の際に聴取環境におけるどのドライバ（単数または複数）またはスピーカー（単数または複数）がそれぞれの音を再生するかを指定する。メタデータは、空間的オーディオ・プロセッサによるパッケージングおよび転送のためにワークステーションにおいてそれぞれのオーディオ・データに関連付けられる。図10は、ある実施形態のもとでの、サウンドバーのための優先度に基づくレンダリングを利用する適応オーディオ・システムにおいて使うためのいくつかの例示的なメタデータ定義を示す表である。図10のテーブル1000において示されるように、メタデータの一部は、オーディオ・コンテンツ型（たとえば、ダイアログ、音楽など）およびある種のオーディオ特性（たとえば直接音、拡散音など）を定義する要素を含んでいてもよい。サウンドバーを通じて再生する優先度に基づくレンダリング・システムについては、メタデータに含まれるドライバ定義は、再生サウンドバーおよびサウンドバーと一緒に使用されうる他のスピーカー（たとえば他のサラウンド・スピーカーまたは

10

20

30

40

50

仮想化対応スピーカー)の構成設定情報(たとえば、ドライバー型、サイズ、パワー、組み込みA/V仮想化など)を含んでいてもよい。図5を参照するに、メタデータはデコーダ型(たとえばデジタル・プラス、トゥルーHDなど)を定義するフィールドおよびデータをも含んでいてもよく、それからチャンネル・ベースのオーディオおよび動的オブジェクト(たとえばOAMDベッド、ISFオブジェクト、動的OAMDオブジェクトなど)の具体的なフォーマットが導出できる。あるいはまた、各オブジェクトのフォーマットは、個別的な関連付けられたメタデータ要素を通じて明示的に定義されてもよい。メタデータは動的オブジェクトについて優先度フィールドをも含み、関連付けられたメタデータはスカラー値(たとえば1から10)または二値の優先度フラグ(高/低)として表現されてもよい。図10に示されるメタデータ要素は、適応オーディオ信号を伝送するビットストリームにおいてエンコードされる可能なメタデータ要素のほんの一部を示すことが意図されており、他の多くのメタデータ要素およびフォーマットも可能である。

10

【0050】

中間空間的フォーマット(Intermediate Spatial Format)

一つまたは複数の実施形態について上記したように、システムによって処理されるある種のオブジェクトはISFオブジェクトである。ISFは、パン動作を時間変化する部分と静的な部分の二つの部分に分割することによってオーディオ・オブジェクト・パンナーの動作を最適化するフォーマットである。一般に、オーディオ・オブジェクト・パンナーは、モノフォニック・オブジェクト(たとえばObject_i)をN個のスピーカーにパンすることによって動作する。ここで、パン利得はスピーカー位置 $(x_1, y_1, z_1), \dots, (x_N, y_N, z_N)$ およびオブジェクト位置 $XYZ_i(t)$ の関数として決定される。オブジェクト位置が時間変化するので、これらの利得値は時間的に連続的に変化する。中間空間的フォーマットの目標は、単にこのパン動作を二つの部分に分けることである。(時間変化する)第一の部分はオブジェクト位置を利用する。(固定した行列を使う)第二の部分はスピーカー位置のみに基づいて構成される。図11は、いくつかの実施形態のもとでレンダリング・システムと一緒に使うための中間空間的フォーマットを示している。描画1100に示されるように、空間的パンナー1102は、スピーカー・デコーダ1106によるデコードのためにオブジェクトおよびスピーカー位置情報を受領する。これら二つの処理ブロック1102および1106の間でオーディオ・オブジェクト・シーンはKチャンネルの中間空間的フォーマット(ISF)1104において表現される。複数のオーディオ・オブジェクト(1 ≤ i ≤ N_i)が個々の空間的パンナーによって処理され、これらの空間的パンナーの出力が足し合わされてISF信号1104をなしてもよく、一つのKチャンネルISF信号集合はN_i個のオブジェクトの重畳を含みうる。ある種の実施形態では、エンコーダは高度制約データを通じてスピーカー高さに関する情報をも与えられてもよく、再生スピーカーの高さの詳細な知識が空間的パンナー1102によって使用されうる。

20

30

【0051】

ある実施形態では、空間的パンナー1102は、再生スピーカーの位置についての詳細な情報を与えられない。しかしながら、いくつかのレベルまたは層に制約された一連の「仮想スピーカー」の位置と、各レベルまたは層内での近似的な分布について想定がされる。こうして、空間的パンナーは再生スピーカーの位置についての詳細な情報を与えられないものの、しばしば、可能性の高いスピーカー数およびそれらのスピーカーの可能性が高い分布に関していくつかの合理的な想定がある。

40

【0052】

結果として得られる再生経験の品質(すなわち、図11のオーディオ・オブジェクト・パンナーにどのくらいよく一致するか)は、ISF内のチャンネルの数Kを増すことによって、あるいは最も確からしい再生スピーカー配置についてのより多くの知識を集めることによって、改善できる。特に、ある実施形態では、図12に示されるようにスピーカー高さがいくつかの面に分割される。所望される合成音場は、聴取者のまわりの任意の方向から発する一連の音イベントと考えることができる。それらの音イベントの位置は、聴取者を中心とする球1202の表面上に定義されると考えられることができる。(高次アンビソニ

50

ックス (Higher Order Ambisonics) のような) 音場フォーマットは、音場が (かなり) 任意のスピーカー・アレイを通じてさらにレンダリングされることを許容するような仕方で定義される。しかしながら、考えられている典型的な再生システムは、スピーカーの高さが三つの面 (耳高さ面、天井面および床面) において固定されているという意味で制約される可能性が高い。よって、理想的な球状音場の概念は修正されることができる。ここで、音場は、聴取者のまわりの球の表面上のさまざまな高さのところにある環内に位置される音オブジェクトから構成される。たとえば、天頂環、上層環、中層環および低位環をもつ、一つのそのような環の配置が図 12 に示されている (1200)。必要であれば、完全性 (completeness) のため、球の底部の追加的な環も含められることもできる (天底; これも厳密に言えば環ではなく点である)。さらに、他の実施形態においては、追加的なまたはより少数の環が存在していてもよい。

10

【0053】

ある実施形態では、積層環フォーマット (stacked-ring format) は BH9.5.0.1 と名付けられ、ここで、四つの数字はそれぞれ中部、上部、下部および天頂の環におけるチャンネル数を示す。マルチチャンネル・バンドルにおけるチャンネルの総数はこれら四つの数の和に等しい (よって、BH9.5.0.1 フォーマットは 15 個のチャンネルを含む)。四つの環すべてを利用するもう一つの例示的なフォーマットは BH15.9.5.1 である。このフォーマットについては、チャンネルの命名および順序付けは次のようになる: [M1, M2, ..., M15, U1, U2, ..., U9, L1, L2, ..., L5, Z1] ここで、チャンネルは環 (M、U、L、Z の順) に配置されており、各環内では単に昇順に基数で番号付けられる。各環は、該環のまわりに一様に広がっている公称スピーカー・チャンネルの集合に入れられると考えられることができる。よって、各環におけるチャンネルは特定のデコード角に対応し、0° の方位角 (真正面) に対応するチャンネル 1 で始まり、反時計回りに数える (よってチャンネル 2 は聴取者から見て中央の左になる)。よって、チャンネル n の方位角は $(n - 1) / N \times 360^\circ$ である (ここで、N はその環におけるチャンネル数であり、n は 1 から N までの範囲内である)。

20

【0054】

ISF に関係したオブジェクト優先度 (object_priority) についてのある種の使用事例に関し、OAMD は一般に、ISF における各環が個別のオブジェクト優先度値をもつことを許容する。ある実施形態では、これらの優先度値は追加的な処理を実行するために複数の仕方で使われる。第一に、高さ面および下部面の環は極小 / 非最適レンダラーによってレンダリングされ、一方、重要な聴取者面の環はより複雑な / 高精度の高品質レンダラーによってレンダリングされることができる。同様に、エンコードされたフォーマットにおいて、聴取者面の環についてはより多くのビット (すなわちより高い品質のエンコード)、高さ面および地上面の環についてはより少数のビットが使用されることができる。ISF は環を使うので、これは ISF において可能である。一方、これは伝統的な高次アンビソニクス・フォーマットでは一般には可能ではない。相異なる各チャンネルが、全体的なオーディオ品質を損なう仕方で相互作用する極パターンだからである。一般に、高さ環または床環についてのやや低下したレンダリング品質は過度に有害ではない。それらの環におけるコンテンツは典型的には雰囲気コンテンツを含むだけだからである。

30

【0055】

ある実施形態では、レンダリングおよび音処理システムは、空間的オーディオ・シーンをエンコードするための二つ以上の環を使用する。ここで、異なる環は、音場の異なる空間的に別個の成分を表わす。オーディオ・オブジェクトは、環内では、転用可能なパン曲線に従ってパンされ、オーディオ・オブジェクトは、環どうしの間では、転用可能でないパン曲線を使ってパンされる。異なる空間的に別個の成分は、その垂直軸に基づいて分離される (すなわち、垂直方向に積層された環)。音場要素は「公称スピーカー」の形での各環内で伝送される: 各環内での音場要素は空間周波数成分の形で伝送される。環の諸セグメントを表わす事前計算されたサブマトリクスをはぎ合わせることによって、各環についてデコード行列が生成される。音がある環から別の環へ、第一の環にスピーカーが存在しない場合、リダイレクトされることができる。

40

50

【 0 0 5 6 】

ISF処理システムにおいて、再生アレイにおける各スピーカ-の位置は(x,y,z)座標（これは、アレイの中心に近い候補聴取位置に対する各スピーカ-の位置である）を使って表現できる。さらに、(x,y,z)ベクトルは単位ベクトルに変換されることができ、事実上、各スピーカ-位置を単位球の表面に投影する。

【 0 0 5 7 】

【数 1】

$$\text{スピーカ-位置} : V_n = \begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} \quad \{1 \leq n \leq N\} \tag{1}$$

10

$$\text{スピーカ-単位ベクトル} : U_n = \frac{1}{\sqrt{V_n^T \times V_n}} \times V_n \tag{2}$$

図 1 3 は、ある実施形態のもとでの、ISF処理システムにおいて使うための、スピーカ-の弧を、ある角度にパンされたオーディオ・オブジェクトとともに示している。描画 1 3 0 0 は、オーディオ・オブジェクト (o) がいくつかのスピーカ- 1 3 0 2 を通じて逐次的にパンされるシナリオを示している。これにより、聴取者 1 3 0 4 は各スピーカ-を順次通過する軌跡を通じて動いているオーディオ・オブジェクトの印象を経験する。一般性を失うことなく、これらのスピーカ- 1 3 0 2 の単位ベクトルは水平面内の環に沿って配列されているとする。よって、オーディオ・オブジェクトの位置はその方位角 の関数として定義されうる。図 1 3 では、角度 におけるオーディオ・オブジェクトはスピーカ-A、BおよびCを通過する（これらのスピーカ-はそれぞれ方位角 A、 Bおよび Cに位置している）。オーディオ・オブジェクト・パンナ-（たとえば図 1 1 のパンナ- 1 1 0 2 ）は典型的には、角度 の関数であるスピーカ-利得を使って、オーディオ・オブジェクトを各スピーカ-にパンする。オーディオ・オブジェクト・パンナ-は、次のような性質をもつパン曲線を使用してもよい：（ 1 ）オーディオ・オブジェクトが物理的なスピーカ-位置に一致する位置にパンされるときは、他のすべてのスピーカ-を排除してその一致するスピーカ-が使用される；（ 2 ）オーディオ・オブジェクトが二つのスピーカ-位置の間にある角度 にパンされるときは、それら二つのスピーカ-のみがアクティブであり、こうしてオーディオ信号のスピーカ-・アレイ上での最小量の「広がり」を提供する；（ 3 ）パン曲線は、高レベルの「離散性」を示してもよい。該「離散性 (discreteness)」とは、パン曲線エネルギーの、あるスピーカ-とその最近接スピーカ-との間の領域内に制約されている割合を指す。よって、図 1 3 を参照するに、スピーカ-Bについて、

20

30

【数 2】

$$\text{離散性} : d_B = \frac{\int_{\phi_A}^{\phi_C} \text{gain}_B(\phi)^2 d\phi}{\int_0^{2\pi} \text{gain}_B(\phi)^2 d\phi} \tag{3}$$

40

よって、 $d_B = 1$ である。 $d_B = 1$ のとき、これは、スピーカ-Bについてのパン曲線は、 A と C （それぞれスピーカ-AとCの角位置）の間の領域のみで非0になるよう（空間的に）完全に制約されることを含意する。対照的に、上記の「離散性」属性を示さない（すなわち、 $d_B < 1$ ）パン曲線は一つの他の重要な属性を示しうる：パン曲線が空間的に平滑化されており、空間周波数において制約されておりナイキスト・サンプリング定理を満たすのである。

【 0 0 5 8 】

空間的に帯域制限されているいかなるパン曲線もその空間的なサポートにおいてコンパ

50

クトであることはできない。換言すれば、これらのパン曲線は、より幅広い角度範囲に分散される。用語「阻止帯域リップル」は、パン曲線において生起する（望ましくない）非0の利得をいう。ナイキスト・サンプリング基準を満たすことによって、これらのパン曲線は、より「離散的」でなくなってしまう。適正に「ナイキスト・サンプリングされ」ることで、これらのパン曲線は代替的なスピーカー位置にシフトされることができる。つまり、（円において均等に離間されている）N個のスピーカーのある特定の配置について生成されたスピーカー信号の集合が、異なる角度位置にあるN個のスピーカーの代替的な集合に（ $N \times N$ 行列によって）リミックスされることができる；すなわち、スピーカー・アレイは角度スピーカー位置の新たな集合に回転させられることができ、もとのN個のスピーカー信号はN個のスピーカーの該新たな集合に転用されることができる。一般に、この「転用可能性」属性は、N個のスピーカー信号を、 $S \times N$ 行列を通じて、S個のスピーカーにマッピングし直すことを許容する。ただし、 $S > N$ の場合、新たなスピーカー・フィールドはもとのNチャンネルよりも「離散的」であることはないことは受け入れられるとする。

【0059】

ある実施形態では、積層環中間空間的フォーマット（Stacked Ring Intermediate Spatial Format）は、以下の段階によって（時間変化する） (x, y, z) 位置に従って各オブジェクトを表わす、を提供する。

1. オブジェクト i が (x_i, y_i, z_i) に位置しており、この位置は立方体内（よって $|x| \leq 1$ 、 $|y| \leq 1$ および $|z| \leq 1$ ）または単位球内（ $x^2 + y_i^2 + z_i^2 \leq 1$ ）にあると想定される。
2. 転用可能でないパン曲線に従って、オブジェクト i についてのオーディオ信号を、ある数（ R ）の空間的領域のそれぞれにパンするために、垂直位置（ z_i ）が使われる。
3. 各空間的領域（たとえば領域 $r: 1 \leq r \leq R$ ）（これは図4のように、空間の環状領域内にあるオーディオ成分を表わす）は、オブジェクト i の方位角（ θ_i ）の関数である転用可能なパン曲線を使って生成される N_r 個の公称スピーカー信号の形で表現される。

【0060】

サイズ0の環（図12では天頂環）という特殊な場合については、環が最大で一つのチャンネルを含むので、段階3は不要である。

【0061】

図11に示されるように、K個のチャンネルについてのISF信号1104はスピーカー・デコード1106においてデコードされる。図14のA~Cは、異なる実施形態のもとの、積層環中間空間的フォーマットのデコードを示している。図14のAは別個の環としてデコードされる積層環フォーマットを示す。図14のBは天頂スピーカーなしでデコードされる積層環フォーマットを示す。図14のCは天頂スピーカーや天井スピーカーなしでデコードされる積層環フォーマットを示す。

【0062】

上記ではISFオブジェクトを動的OAMDオブジェクトに対する一つの型のオブジェクトとして実施形態が記述されているが、異なるフォーマットでフォーマットされているが動的OAMDオブジェクトとは区別可能なオーディオ・オブジェクトが使われることもできることは注意しておくべきである。

【0063】

本稿に記述されるオーディオ環境の諸側面は、適切なスピーカーおよび再生装置を通じたオーディオまたはオーディオ/ビジュアル・コンテンツの再生を表わし、聴取者が捕捉されたコンテンツの再生を経験している任意の環境、たとえば映画館、コンサートホール、屋外シアター、家庭または部屋、聴取ブース、自動車、ゲーム・コンソール、ヘッドフォンまたはヘッドセット・システム、公衆アナウンス（PA: public address）システムまたは他の任意の再生環境を表わしうる。実施形態は主として、空間的オーディオ・コンテンツがテレビジョン・コンテンツに関連付けられているホームシアター環境における例および実装に関して記述されてきたが、実施形態は、ゲーム、スクリーニング・システムおよび他の任意のモニター・ベースのA/Vシステムといった他の消費者ベースのシステムにおいて実装されてもよいことを注意しておくべきである。オブジェクト・ベースのオー

10

20

30

40

50

オーディオおよびチャンネル・ベースのオーディオを含む空間的オーディオ・コンテンツは、いかなる関係するコンテンツ（関連付けられたオーディオ、ビデオ、グラフィックなど）との関連で使われてもよく、あるいは単独のオーディオ・コンテンツをなしていてもよい。再生環境は、ヘッドフォンまたはニア・フィールド・モニターから大小の部屋、自動車、屋外アリーナ、コンサートホールなどまでのいかなる適切な聴取環境であってもよい。

【0064】

本稿に記載されるシステムの諸側面は、デジタルまたはデジタイズされたオーディオ・ファイルを処理するための適切なコンピュータ・ベースの音処理ネットワーク環境において実装されてもよい。適応オーディオ・システムの諸部分は、コンピュータ間で伝送されるデータをバッファリングおよびルーティングするはたらきをする一つまたは複数のルーター（図示せず）を含め、任意の所望される数の個々の機械を含む一つまたは複数のネットワークを含んでいてもよい。そのようなネットワークは、さまざまな異なるネットワーク・プロトコル上で構築されてもよく、インターネット、広域ネットワーク（WAN）、ローカル・エリア・ネットワーク（LAN）またはその任意の組み合わせであってもよい。ネットワークがインターネットを含む実施形態では、一つまたは複数の機械がウェブ・ブラウザ・プログラムを通じてインターネットにアクセスするよう構成されてもよい。

10

【0065】

上記のコンポーネント、ブロック、プロセスまたは他の機能構成要素の一つまたは複数、システムのプロセッサ・ベースのコンピューティング装置の実行を制御するコンピュータ・プログラムを通じて実装されてもよい。本稿に開示されるさまざまな機能は、ハードウェア、ファームウェアのいくつもある組み合わせを使っておよび/またはさまざまな機械可読もしくはコンピュータ可読媒体において具現されたデータおよび/または命令として、挙動上の、レジスタ転送、論理コンポーネントおよび/または他の特性を用いて記載されることがあることを注意しておくべきである。そのようなフォーマットされたデータおよび/または命令が具現されうるコンピュータ可読媒体は、光学式、磁気式もしくは半導体記憶媒体のようなさまざまな形の物理的（非一時的）、不揮発性記憶媒体を含むがそれに限定されない。

20

【0066】

文脈がそうでないことを明確に要求するのではないが、本記述および請求項を通じて、単語「有する」「含む」などは、排他的もしくは網羅的な意味ではなく包含的な意味に解釈されるものとする。すなわち、「.....を含むがそれに限定されない」の意味である。単数または複数を使った単語は、それぞれ複数または単数をも含む。さらに、「本稿で」「以下で」「上記で」「下記で」および類似の意味の単語は、全体としての本願を指すのであって、本願のいかなる特定の部分を指すものでもない。単語「または」が二つ以上の項目のリストを参照して使われるとき、その単語は該単語の以下の解釈のすべてをカバーする：リスト中の項目の任意のもの、リスト中の項目のすべておよびリスト中の項目の任意の組み合わせ。

30

【0067】

本明細書を通じて「一つの実施形態」「いくつかの実施形態」または「ある実施形態」への言及は、その実施形態との関連で記述されている特定の特性、構造または特性が開示されるシステムおよび方法の少なくとも一つの実施形態に含まれることを意味する。よって、本稿を通じた随所に「一つの実施形態では」「いくつかの実施形態では」または「ある実施形態では」という句が現われるのは、同じ実施形態を指すこともあれば、必ずしもそうでないこともある。さらに、具体的な特徴、構造または特性は、当業者には明白であろう任意の好適な仕方でも組み合わせられてもよい。

40

【0068】

一つまたは複数の実装が、例として、個別的な実施形態を用いて記載されているが、一つまたは複数の実装は開示される実施形態に限定されないことは理解されるものとする。逆に、当業者に明白であろうさまざまな修正および類似の構成をカバーすることが意図されている。したがって、付属の請求項の範囲は、そのようなすべての修正および類似の構

50

成を包含するような最も広い解釈を与えられるべきである。

【 0 0 6 9 】

いくつかの態様を記載しておく。

〔 態 様 1 〕

適応オーディオをレンダリングする方法であって：

チャンネル・ベースのオーディオ、オーディオ・オブジェクトおよび動的オブジェクトを含む入力オーディオを受領する段階であって、前記動的オブジェクトは低優先度動的オブジェクトの集合および高優先度動的オブジェクトの集合として分類される、段階と；

前記チャンネル・ベースのオーディオ、前記オーディオ・オブジェクトおよび前記低優先度動的オブジェクトをオーディオ処理システムの第一のレンダリング・プロセッサにおいてレンダリングする段階と；

前記高優先度動的オブジェクトを前記オーディオ処理システムの第二のレンダリング・プロセッサにおいてレンダリングする段階とを含む、方法。

10

〔 態 様 2 〕

前記入力オーディオは、オーディオ・コンテンツおよびレンダリング・メタデータを含むオブジェクト・オーディオ・ベースのデジタル・ビットストリーム・フォーマットに従ってフォーマットされている、態様 1 記載の方法。

〔 態 様 3 〕

前記チャンネル・ベースのオーディオはサラウンドサウンド・オーディオ・ベッドを含み、前記オーディオ・オブジェクトは中間空間的フォーマットに準拠するオブジェクトを含む、態様 2 記載の方法。

20

〔 態 様 4 〕

前記低優先度動的オブジェクトおよび高優先度動的オブジェクトは、優先度閾値によって区別される、態様 2 記載の方法。

〔 態 様 5 〕

前記優先度閾値は、前記入力オーディオを含むオーディオ・コンテンツの作者、ユーザー選択された値および前記オーディオ処理システムによって実行される自動化されたプロセスのうちの一つによって定義される、態様 4 記載の方法。

〔 態 様 6 〕

前記優先度閾値は、前記オブジェクト・オーディオ・メタデータ・ビットストリームにおいてエンコードされている、態様 5 記載の方法。

30

〔 態 様 7 〕

前記低優先度および高優先度のオーディオ・オブジェクトのオーディオ・オブジェクトの相対的な優先度は前記オブジェクト・オーディオ・メタデータ・ビットストリームにおけるそれぞれの位置によって決定される、態様 5 記載の方法。

〔 態 様 8 〕

前記第一のレンダリング・プロセッサにおいて前記チャンネル・ベースのオーディオ、前記オーディオ・オブジェクトおよび前記低優先度動的オブジェクトをレンダリングしてレンダリングされたオーディオを生成する間またはその後、前記高優先度オーディオ・オブジェクトを前記第一のレンダリング・プロセッサを通して前記第二のレンダリング・プロセッサに渡し；

前記レンダリングされたオーディオをスピーカー・システムへの伝送のために後処理することをさらに含む、

態様 1 記載の方法。

40

〔 態 様 9 〕

前記後処理する段階は、アップミックス、ボリューム制御、等化および低音管理のうちの少なくとも一つを含む、態様 8 記載の方法。

〔 態 様 1 0 〕

前記後処理する段階は、前記スピーカー・システムを通じた再生のための前記入力オー

50

ディオに存在している高さ手がかりのレンダリングを容易にするための仮想化段階をさらに含む、態様 9 記載の方法。

〔態様 1 1〕

前記スピーカー・システムは、単一の軸に沿って音を送出する複数の共位置のドライバーを有するサウンドバー・スピーカーを有する、態様 1 0 記載の方法。

〔態様 1 2〕

前記第一および第二のレンダリング・プロセッサは、伝送リンクを通じて一緒に結合された別個のデジタル信号処理回路において具現される、態様 4 記載の方法。

〔態様 1 3〕

前記優先度閾値は、前記第一および第二のレンダリング・プロセッサの相対的な処理機能、前記第一および第二のレンダリング・プロセッサのそれぞれに関連付けられたメモリ帯域幅および前記伝送リンクの伝送帯域幅のうちの少なくとも一つによって決定される、態様 1 2 記載の方法。

10

〔態様 1 4〕

適応オーディオをレンダリングする方法であって：

オーディオ・コンポーネントおよび関連付けられたメタデータを含む入力オーディオ・ビットストリームを受領する段階であって、前記オーディオ・コンポーネントはそれぞれチャンネル・ベースのオーディオ、オーディオ・オブジェクトおよび動的オブジェクトから選択されるオーディオ型をもつ、段階と；

各オーディオ・コンポーネントについてのデコーダ・フォーマットをそれぞれのオーディオ型に基づいて決定する段階と；

20

各オーディオ・コンポーネントの優先度を、該各オーディオ・コンポーネントに関連付けられたメタデータにおける優先度フィールドから決定する段階と；

第一のレンダリング・プロセッサにおいて第一の優先度型のオーディオ・コンポーネントをレンダリングする段階と；

第二のレンダリング・プロセッサにおいて第二の優先度型のオーディオ・コンポーネントをレンダリングする段階とを含む、

方法。

〔態様 1 5〕

前記第一のレンダリング・プロセッサおよび第二のレンダリング・プロセッサは、伝送リンクを通じて互いに結合された別個のレンダリング・デジタル信号プロセッサ (DSP) として実装される、態様 1 4 記載の方法。

30

〔態様 1 6〕

前記第一の優先度型のオーディオ・コンポーネントは低優先度の動的オブジェクトを含み、第二の優先度型のオーディオ・コンポーネントは高優先度の動的オブジェクトを含み、本方法はさらに、前記チャンネル・ベースのオーディオ、前記オーディオ・オブジェクトを前記第一のレンダリング・プロセッサにおいてレンダリングすることを含む、態様 1 5 記載の方法。

〔態様 1 7〕

前記チャンネル・ベースのオーディオはサラウンドサウンド・オーディオ・ベッドを含み、前記オーディオ・オブジェクトは中間空間的フォーマット (ISF) に準拠するオブジェクトを含み、前記低優先度および高優先度動的オブジェクトはオブジェクト・オーディオ・メタデータ (OAMD) フォーマットに準拠するものを含む、態様 1 5 記載の方法。

40

〔態様 1 8〕

各オーディオ・コンポーネントについてのデコーダ・フォーマットは：OAMD フォーマットされた動的オブジェクト、サラウンドサウンド・オーディオ・ベッドおよび ISF オブジェクトのうちの少なくとも一つを生成する、態様 1 7 記載の方法。

〔態様 1 9〕

前記低優先度および高優先度動的オブジェクトのオーディオ・オブジェクトの相対的な優先度は前記入力オーディオ・ビットストリームにおけるそれぞれの位置によって決定さ

50

れる、態様 16 記載の方法。

〔態様 20〕

前記スピーカー・システムを通じた再生のための前記入力オーディオに存在している高さ手がかりのレンダリングを容易にするよう、少なくとも前記高優先度動的オブジェクトに仮想化プロセスを適用することをさらに含む、態様 19 記載の方法。

〔態様 21〕

前記スピーカー・システムは、単一の軸に沿って音を送出する複数の共位置のドライバーを有するサウンドバー・スピーカーを有する、態様 20 記載の方法。

〔態様 22〕

適応オーディオをレンダリングするシステムであって：

オーディオ・コンテンツおよび関連付けられたメタデータを有するビットストリームにおいて入力オーディオを受領するインターフェースであって、前記オーディオ・コンテンツは、チャンネル・ベースのオーディオ、オーディオ・オブジェクトおよび動的オブジェクトを含み、前記動的オブジェクトは低優先度動的オブジェクトの集合および高優先度動的オブジェクトの集合として分類される、インターフェースと；

前記チャンネル・ベースのオーディオ、前記オーディオ・オブジェクトおよび前記低優先度動的オブジェクトをレンダリングする、前記インターフェースに結合された第一のレンダリング・プロセッサと；

前記高優先度動的オブジェクトをレンダリングする、伝送リンクを通じて前記第一のレンダリング・プロセッサに結合された第二のレンダリング・プロセッサとを有する、システム。

〔態様 23〕

前記チャンネル・ベースのオーディオはサラウンドサウンド・オーディオ・ベッドを含み、前記オーディオ・オブジェクトは中間空間的フォーマット (ISF) に準拠するオブジェクトを含み、前記低優先度および高優先度動的オブジェクトはオブジェクト・オーディオ・メタデータ (OAMD) フォーマットに準拠するオブジェクトを含む、態様 22 記載のシステム。

〔態様 24〕

前記低優先度動的オブジェクトおよび高優先度動的オブジェクトは、優先度閾値によって区別され、前記優先度閾値は、前記メタデータ・ビットストリームの適切なフィールドにおいてエンコードされており、前記入力オーディオを含むオーディオ・コンテンツの作者、ユーザー選択された値および前記オーディオ処理システムによって実行される自動化されたプロセスのうちの一つによって決定される、態様 23 記載のシステム。

〔態様 25〕

前記第一のレンダリング・プロセッサおよび第二のレンダリング・プロセッサにおいてレンダリングされたオーディオに対して一つまたは複数の後処理段階を実行する後処理器をさらに有し、前記後処理段階は、アップミックス、ボリューム制御、等化および低音管理のうち少なくとも一つを含む、態様 24 記載のシステム。

〔態様 26〕

単一の軸に沿って音を送出する複数の共位置のドライバーを有するサウンドバー・スピーカーを通じた再生のための前記レンダリングされたオーディオに存在している高さ手がかりのレンダリングを容易にするための少なくとも一つの仮想化段階を実行する、前記後処理器に結合された仮想化器コンポーネントをさらに有する、態様 25 記載のシステム。

〔態様 27〕

前記優先度閾値は、前記第一および第二のレンダリング・プロセッサの相対的な処理機能、前記第一および第二のレンダリング・プロセッサのそれぞれに関連付けられたメモリ帯域幅および前記伝送リンクの伝送帯域幅のうちの一つによって決定される、態様 24 記載の方法。

〔態様 28〕

聴取環境における仮想化されたオーディオ・コンテンツの再生のためのスピーカー・シ

10

20

30

40

50

ステムであって：

エンクロージャーと；

前記エンクロージャー内に配置され、前記エンクロージャーの前面を通じて音を投射するよう構成された複数の個別ドライバーと；

オーディオ・コンポーネントおよび関連付けられたメタデータを含むオーディオ・ビットストリームに含まれる第一の優先度型のオーディオ・コンポーネントをレンダリングする第一のレンダリング・プロセッサならびに前記オーディオ・ビットストリームに含まれる第二の優先度型のオーディオ・コンポーネントをレンダリングする第二のレンダリング・プロセッサによって生成されたレンダリングされたオーディオを受領するインターフェースとを有する、

スピーカー・システム。

〔態様 29〕

前記第一のレンダリング・プロセッサおよび第二のレンダリング・プロセッサが、伝送リンクを通じて互いに結合された別個のレンダリング・デジタル信号プロセッサ（DSP）として実装される、態様 28 記載のスピーカー・システム。

〔態様 30〕

前記第一の優先度型のオーディオ・コンポーネントは低優先度動的オブジェクトを含み、前記第二の優先度型のオーディオ・コンポーネントは高優先度動的オブジェクトを含み、前記チャンネル・ベースのオーディオはサラウンドサウンド・オーディオ・ベッドを含み、前記オーディオ・オブジェクトは中間空間的フォーマット（ISF）に準拠するオブジェクトを含み、前記低優先度および高優先度動的オブジェクトはオブジェクト・オーディオ・メタデータ（OAMD）フォーマットに準拠するものを含む、態様 29 記載のスピーカー・システム。

〔態様 31〕

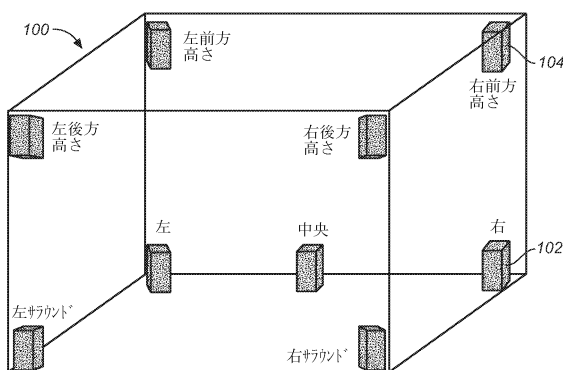
当該スピーカー・システムを通じた再生のための前記入力オーディオに存在している高さ手がかりのレンダリングを容易にするために少なくとも前記高優先度動的オブジェクトに仮想化プロセスを適用する仮想化器をさらに有する、態様 30 記載のスピーカー・システム。

〔態様 32〕

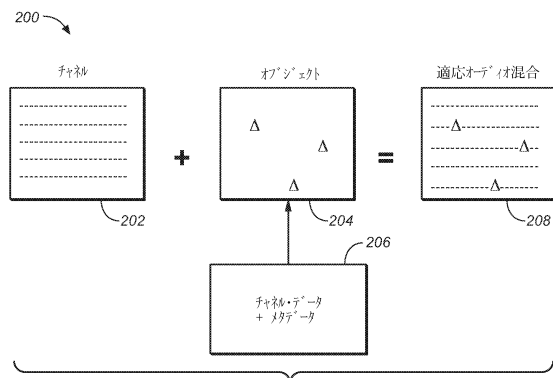
前記仮想化器、前記第一のレンダリング・プロセッサおよび前記第二のレンダリング・プロセッサのうちの少なくとも一つは当該スピーカー・システムの前記エンクロージャーに緊密に結合されているまたは該エンクロージャーに囲まれている、態様 31 記載のスピーカー・システム。

【図面】

【図 1】



【図 2】



10

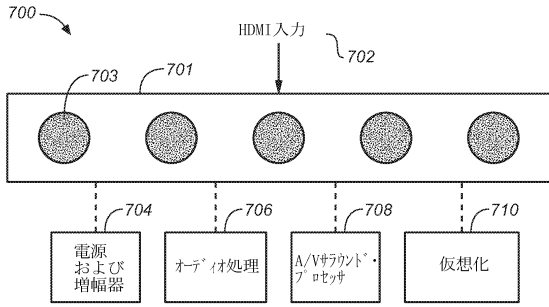
20

30

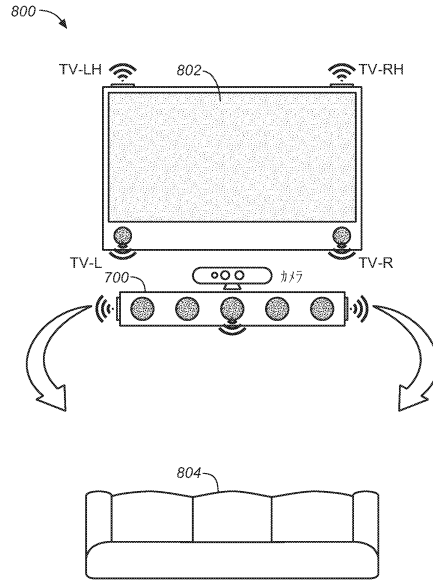
40

50

【図7】



【図8】

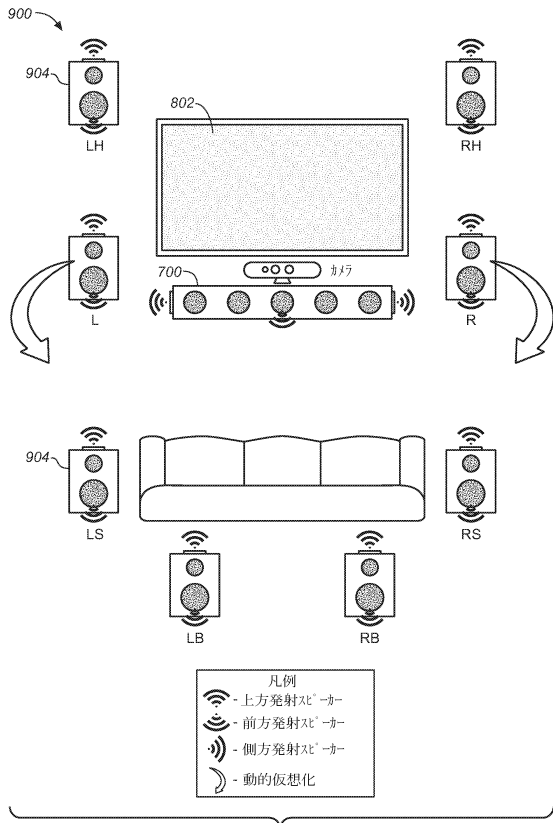


10

- 凡例
- 上方発射スピーカー
 - 前方発射スピーカー
 - 側方発射スピーカー
 - 動的仮想化

20

【図9】



30

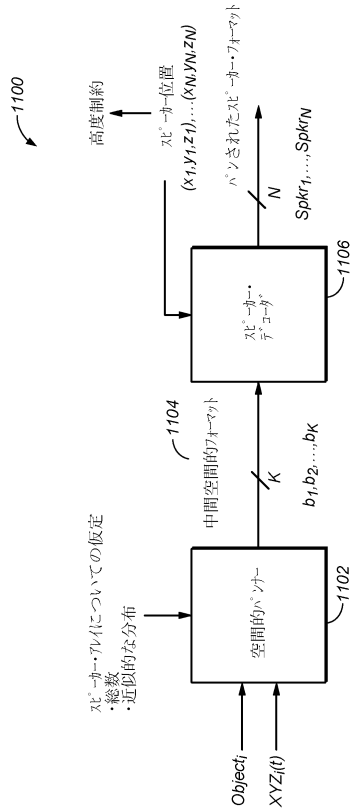
【図10】

モード型	モード要素
オーディオ・コンテンツ型	ダイアログ / 音楽 / 周囲音 / 効果 直接音 / 拡散音 / 反射音
ドライバー定義	サウンズバー構成 サウンズバー構成
デコーダ型	デジタル・プラス トータルHD MAT2.0 任意のSRC
動的オブジェクト優先度	スケール値 二値ブーム

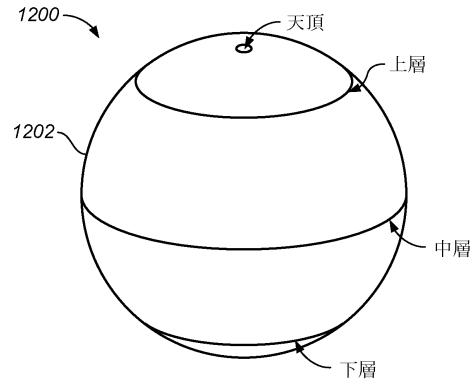
40

50

【図 1 1】



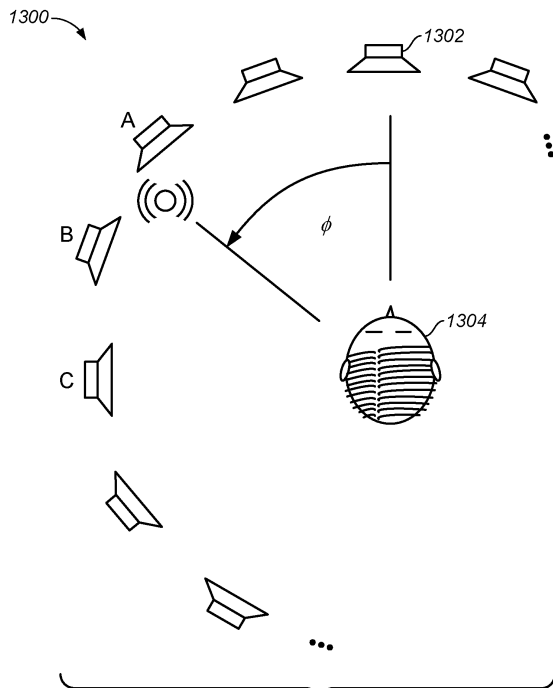
【図 1 2】



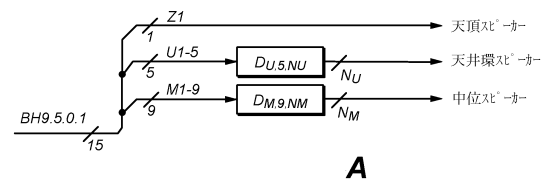
10

20

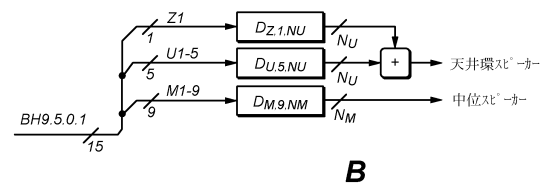
【図 1 3】



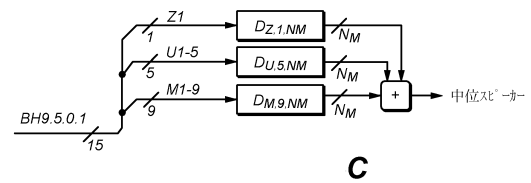
【図 1 4】



30



40



50

フロントページの続き

- ト・ストリート 1275, ドルビー ラボラトリーズ インコーポレイテッド内
- (72)発明者 サンチェス, フレディ
アメリカ合衆国, カリフォルニア州 94103, サンフランシスコ, マーケット・ストリート 1275, ドルビー ラボラトリーズ インコーポレイテッド内
- (72)発明者 シーフェルト, アラン ジェイ.
アメリカ合衆国, カリフォルニア州 94103, サンフランシスコ, マーケット・ストリート 1275, ドルビー ラボラトリーズ インコーポレイテッド内
- 審査官 上田 雄
- (56)参考文献 米国特許出願公開第2012/0230497 (US, A1)
特表2014-525048 (JP, A)
特表2013-502184 (JP, A)
特開平07-092981 (JP, A)
- (58)調査した分野 (Int.Cl., DB名)
H04S 7/00
H04R 3/00