

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6607895号  
(P6607895)

(45) 発行日 令和1年11月20日 (2019. 11. 20)

(24) 登録日 令和1年11月1日 (2019. 11. 1)

(51) Int. Cl.		F I			
HO 4 S	7/00	(2006. 01)	HO 4 S	7/00	3 6 0
HO 4 S	3/00	(2006. 01)	HO 4 S	3/00	4 0 0

請求項の数 14 外国語出願 (全 42 頁)

(21) 出願番号	特願2017-179893 (P2017-179893)	(73) 特許権者	507236292
(22) 出願日	平成29年9月20日 (2017. 9. 20)		ドルビー ラボラトリーズ ライセンシン
(62) 分割の表示	特願2016-543161 (P2016-543161)		グ コーポレイション
原出願日	平成26年12月18日 (2014. 12. 18)		アメリカ合衆国 9 4 1 0 3 カリフォル
(65) 公開番号	特開2018-14749 (P2018-14749A)		ニア州 サンフランシスコ マーケット
(43) 公開日	平成30年1月25日 (2018. 1. 25)	(74) 代理人	100107766
審査請求日	平成29年12月12日 (2017. 12. 12)		弁理士 伊東 忠重
(31) 優先権主張番号	61/923, 579	(74) 代理人	100070150
(32) 優先日	平成26年1月3日 (2014. 1. 3)		弁理士 伊東 忠彦
(33) 優先権主張国・地域又は機関	米国 (US)	(74) 代理人	100091214
(31) 優先権主張番号	201410178258.0		弁理士 大貫 進介
(32) 優先日	平成26年4月29日 (2014. 4. 29)		
(33) 優先権主張国・地域又は機関	中国 (CN)		

最終頁に続く

(54) 【発明の名称】 少なくとも一つのフィードバック遅延ネットワークを使ったマルチチャネル・オーディオにตอบสนองしたバイノーラル・オーディオの生成

(57) 【特許請求の範囲】

【請求項 1】

マルチチャネル・オーディオ入力信号のチャンネルのある集合にตอบสนองしてバイノーラル信号を生成する方法であって、当該方法は：

前記集合の各チャンネルにバイノーラル室内インパルス応答 (BRIR) を適用し、それによりフィルタリングされた信号を生成する段階と；

フィルタリングされた信号を組み合わせて前記バイノーラル信号を生成する段階とを含み、

前記集合の各チャンネルにBRIRを適用することは、後期残響生成器を使って、該後期残響生成器に呈された制御値にตอบสนองして、共通の後期残響を前記集合のチャンネルのダウンミックスに加えることを含み、前記共通の後期残響は前記集合の少なくともいくつかのチャンネルにわたって共有される単一チャンネルBRIRの後期残響部分の集団的なマクロ属性をエミュレートし、前記制御値は目標両耳間コヒーレンスを達成するように決定されたものであり、

前記ダウンミックスは前記集合のチャンネルのステレオ・ダウンミックスである、方法。

【請求項 2】

前記集合の各チャンネルにBRIRを適用することは、前記集合の各チャンネルに、該チャンネルについての単一チャンネルBRIRの直接応答および早期反射部分を適用することを含む、請求項 1 記載の方法。

【請求項 3】

前記後期残響生成器は、前記ダウンミックスに前記共通の後期残響を加えるためのフィードバック遅延ネットワークのバンクを含み、該バンクの各フィードバック遅延ネットワークは前記ダウンミックスの異なる周波数帯域に後期残響を加える、請求項 1 または 2 記載の方法。

【請求項 4】

前記フィードバック遅延ネットワークのそれぞれは複素直交ミラー・フィルタ領域で実装される、請求項 3 記載の方法。

【請求項 5】

前記後期残響生成器は、前記集合の前記チャンネルの前記ダウンミックスに前記共通の後期残響を加えるための単一のフィードバック遅延ネットワークを含み、前記フィードバック遅延ネットワークは時間領域で実装される、請求項 1 ないし 4 のうちいずれか一項記載の方法。

10

【請求項 6】

前記集団的なマクロ属性は、平均パワー・スペクトル、エネルギー減衰構造、モード密度およびピーク密度の一つまたは複数を含む、請求項 1 ないし 5 のうちいずれか一項記載の方法。

【請求項 7】

前記制御値の一つまたは複数が周波数依存であるおよび / または前記制御値の一つが残響時間である、請求項 1 ないし 6 のうちいずれか一項記載の方法。

【請求項 8】

20

マルチチャンネル・オーディオ入力信号のチャンネルのある集合に応答してバイノーラル信号を生成するシステムであって、当該システムは：

前記集合の各チャンネルにバイノーラル室内インパルス応答 (BRIR) を適用し、それによりフィルタリングされた信号を生成し；

フィルタリングされた信号を組み合わせて前記バイノーラル信号を生成する、一つまたは複数のプロセッサを有しており、

前記集合の各チャンネルにBRIRを適用することは、後期残響生成器を使って、該後期残響生成器に呈された制御値に応答して、共通の後期残響を前記集合のチャンネルのダウンミックスに加えることを含み、前記共通の後期残響は前記集合の少なくともいくつかのチャンネルにわたって共有される単一チャンネルBRIRの後期残響部分の集団的なマクロ属性をエミュレートし、前記制御値は目標両耳間コヒーレンスを達成するように決定されたものであり、

30

前記集合の前記チャンネルの前記ダウンミックスは前記集合のチャンネルのステレオ・ダウンミックスである、システム。

【請求項 9】

前記集合の各チャンネルにBRIRを適用することは、前記集合の各チャンネルに、該チャンネルについての単一チャンネルBRIRの直接応答および早期反射部分を適用することを含む、請求項 8 記載のシステム。

【請求項 10】

前記後期残響生成器は、前記ダウンミックスに前記共通の後期残響を加えるよう構成されたフィードバック遅延ネットワークのバンクを含み、該バンクの各フィードバック遅延ネットワークは前記ダウンミックスの異なる周波数帯域に後期残響を加える、請求項 8 または 9 記載のシステム。

40

【請求項 11】

前記フィードバック遅延ネットワークのそれぞれは複素直交ミラー・フィルタ領域で実装される、請求項 10 記載のシステム。

【請求項 12】

前記後期残響生成器は、時間領域で実装されたフィードバック遅延ネットワークを含み、前記後期残響生成器は、前記共通の後期残響を前記ダウンミックスに加えるために、前記フィードバック遅延ネットワークにおいて時間領域で前記ダウンミックスを処理するよ

50

う構成されている、請求項 8 または 9 記載のシステム。

【請求項 1 3】

前記集団的なマクロ属性は、平均パワー・スペクトル、エネルギー減衰構造、モード密度およびピーク密度の一つまたは複数を含む、請求項 8 ないし 12 のうちいずれか一項記載のシステム。

【請求項 1 4】

前記制御値の一つまたは複数が周波数依存であるおよび / または前記制御値の一つが残響時間である、請求項 8 ないし 13 のうちいずれか一項記載のシステム。

【発明の詳細な説明】

【技術分野】

10

【0001】

関連出願への相互参照

本願は2014年4月29日に出願された中国特許出願第201410178258.0号、2014年1月3日に出願された米国仮特許出願第61/923,579号および2014年5月5日に出願された米国仮特許出願第61/988,617号の優先権を主張するものである。各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。

1. 発明の分野

本発明は、入力信号のチャンネルの集合の各チャンネルに（たとえば全チャンネルに）バイノーラル室内インパルス応答（BRIR: Binaural Room Impulse Response）を適用することによって、マルチチャンネル・オーディオ入力信号に応答してバイノーラル信号を生成するための方法（時にヘッドフォン仮想化方法と称される）およびシステムに関する。いくつかの実施形態では、少なくとも一つのフィードバック遅延ネットワーク（FDN: feedback delay network）がダウンミックスBRIRの後期残響部分を前記チャンネルのダウンミックスに適用する。

20

【背景技術】

【0002】

2. 発明の背景

ヘッドフォン仮想化（またはバイノーラル・レンダリング）は、標準的なステレオ・ヘッドフォンを使ってサラウンド・サウンド経験または没入的な音場を送達することをねらいとする技術である。

30

【0003】

初期のヘッドフォン仮想化器は、バイノーラル・レンダリングにおける空間的情報を伝えるために頭部伝達関数（HRTF: head-related transfer function）を適用した。HRTFは、無響環境において空間内の特定の点（音源位置）から聴取者の両耳に音がどのように伝わるかを特徴付ける方向および距離依存のフィルタ対の集合である。両耳間時間差（ITD: interaural time difference）、両耳間レベル差（ILD: interaural level difference）、頭のシャドーイング効果（head shadowing effect）、肩および耳介反射に起因するスペクトルのピークおよびノッチといった本質的な空間的手がかりが、レンダリングされるHRTFフィルタリングされたバイノーラル・コンテンツにおいて知覚されることができる。人間の頭のサイズの制約条件のため、HRTFは、ほぼ1メートルより先の源距離に関しては十分または堅牢な手がかりを提供しない。結果として、HRTFのみに基づく仮想化器は通例、良好な頭外定位または知覚される距離を達成しない。

40

【0004】

日常生活における音響イベントの多くは残響のある環境で生起する。残響のある環境では、HRTFによってモデル化される（源から耳への）直接経路に加えて、さまざまな反射経路を通じてオーディオ信号が聴取者の耳に達する。反射は、距離、部屋サイズおよび空間の他の属性といった聴覚体験に深遠な影響を導入する。この情報をバイノーラル・レンダリングにおいて伝えるために、仮想化器は、直接経路HRTFにおける手がかりに加えて、部屋残響を適用する必要がある。バイノーラル室内インパルス応答（BRIR）は、特定の音響環境における空間内の特定の点から聴取者の耳までのオーディオ信号の変換を特徴付け

50

る。理論上は、BRIRは空間的知覚に関するすべての音響手がかりを含む。

【 0 0 0 5 】

図 1 は、マルチチャンネル・オーディオ入力信号のそれぞれの全周波数範囲チャンネル ( $X_1, \dots, X_N$ ) にバイノーラル室内インパルス応答 (BRIR) を適用するよう構成された通常のヘッドホン仮想化器の一つの型のブロック図である。チャンネル  $X_1, \dots, X_N$  のそれぞれは、想定される聴取者に対する異なる源方向 (すなわち、対応するスピーカーの想定される位置から想定される聴取者位置への直接経路の方向) に対応するスピーカー・チャンネルであり、そのような各チャンネルは対応する源方向についてのBRIRによって畳み込みされる。各チャンネルからの音響経路は、各耳についてシミュレートする必要がある。したがって、本稿の残りでは、用語BRIRは、一つのインパルス応答または左右の耳に関連付けられたインパルス応答の対のいずれをも指す。よって、サブシステム 2 はチャンネル  $X_1$  を  $BRIR_1$  (対応する源方向についてのBRIR) と畳み込みするよう構成され、サブシステム 4 はチャンネル  $X_N$  を  $BRIR_N$  (対応する源方向についてのBRIR) と畳み込みするよう構成される、などとなる。各BRIRサブシステム (サブシステム 2、...、4 のそれぞれ) の出力は、左チャンネルおよび右チャンネルを含む時間領域信号である。BRIRサブシステムの左チャンネル出力どうしは加算要素 6 において混合され、BRIRサブシステムの右チャンネルどうしは加算要素 8 において混合される。要素 6 の出力は、仮想化器から出力されるバイノーラル・オーディオ信号の左チャンネルLであり、要素 8 の出力は、仮想化器から出力されるバイノーラル・オーディオ信号の右チャンネルRである。

【 0 0 0 6 】

マルチチャンネル・オーディオ入力信号は、低域効果 (LFE: low frequency effects) またはサブウーファー・チャンネルをも含んでいてもよい。これは図 1 では「LFE」チャンネルとして同定されている。通常の仕方では、LFEチャンネルはBRIRと畳み込みされないが、その代わり、図 1 の利得段 5 において (たとえば -3dB 以上) 減衰させられ、利得段 5 の出力が仮想化器のバイノーラル出力信号の各チャンネルに等しく (加算要素 6 および 8 によって) 混合される。段 5 の出力をBRIRサブシステム (2、...、4) の出力と時間整列させるために、LFE経路において追加的な遅延段が必要とされることがある。あるいはまた、LFEチャンネルは単に無視されてもよい (すなわち、仮想化器に呈されないまたは仮想化器によって処理されない)。たとえば、本発明の図 2 の実施形態 (後述) は、それが処理するマルチチャンネル・オーディオ入力信号のいかなるLFEチャンネルをも単に無視する。多くの消費者ヘッドフォンは、LFEチャンネルを正確に再生することができない。

【 0 0 0 7 】

いくつかの通常の仮想化器では、入力信号は、時間領域から周波数領域への変換を受けてQMF (quadrature mirror filter [直交ミラー・フィルタ]) 領域にされ、QMF領域周波数成分の諸チャンネルを生成する。これらの周波数成分は (たとえば図 1 のサブシステム 2、...、4 のQMF領域実装において) QMF領域でフィルタリングを受けて、結果として得られる周波数成分が次いで (たとえば図 1 のサブシステム 2、...、4 のそれぞれの最終段において) 時間領域に変換し戻される。それにより、仮想化器のオーディオ出力は時間領域信号 (たとえば、時間領域バイノーラル信号) である。

【 0 0 0 8 】

一般に、ヘッドホン仮想化器に入力されるマルチチャンネル・オーディオ信号のそれぞれの全周波数範囲チャンネルは、聴取者の耳に対して既知の位置にある音源から放出されるオーディオ・コンテンツを示すと想定される。ヘッドホン仮想化器は、入力信号のそのような各チャンネルにバイノーラル室内インパルス応答 (BRIR) 適用するよう構成される。各BRIRは、直接応答および反射という二つの部分に分解できる。直接応答は、音源の到来方向 (DOA: direction of arrival) に対応するHRTFを、(音源と聴取者の間の) 距離に起因する適正な利得および遅延をもって調整し、任意的には小さな距離についてのパララックス (parallax) 効果をもって増強したものである。

【 0 0 0 9 】

BRIRの残りの部分は反射をモデル化する。早期の反射は通例一次または二次反射であり

10

20

30

40

50

、比較的疎な時間的分布をもつ。各一次または二次反射のミクロ構造（たとえばITDおよびILD）は重要である。後期反射（聴取者に達する前に三つ以上の表面から反射された音）については、反射回数の増大とともにエコー密度が増大し、個々の反射のミクロ属性は観察しにくくなる。ますますより後期の反射については、マクロ構造（たとえば、残響減衰レート、両耳間コヒーレンスおよび全体的な残響のスペクトル分布）がより重要になる。このため、反射は、早期反射および後期残響という二つの部分にさらにセグメント分割できる。

#### 【 0 0 1 0 】

直接応答の遅延は聴取者からの源距離を音速で割ったものであり、そのレベルは（源位置近くの壁または大きな表面がない場合）源距離に反比例する。他方、後期残響の遅延およびレベルは一般に源位置には敏感でない。実際のな事情のため、仮想化器は、異なる距離をもつ源からの直接応答を時間整列させるおよび／またはそのダイナミックレンジを圧縮することを選ぶ。しかしながら、BRIR内での直接応答、早期反射および後期残響の間の時間的およびレベル関係は維持されるべきである。

#### 【 0 0 1 1 】

典型的なBRIRの有効長さは、多くの音響環境において数百ミリ秒以上に達する。BRIRの直接的な適用は、数千のタップのフィルタとの畳み込みを必要とするが、これは計算的に高価である。加えて、パラメータ化なしでは、十分な空間分解能を達成するためには、異なる源位置についての諸BRIRを記憶する大きなメモリ・スペースを必要とする。最後だが軽んじてはならないこととして、音源位置は時間とともに変化するおよび／または聴取者の位置および配向は時間とともに変化する。そのような動きの正確なシミュレーションは時間変化するBRIRインパルス応答を要求する。そのような時間変化するフィルタの適正な補間および適用は、これらのフィルタのインパルス応答が多くのタップをもつ場合には、困難であることがある。

#### 【 0 0 1 2 】

シミュレートされた残響をマルチチャネル・オーディオ入力信号の一つまたは複数のチャネルに適用するよう構成された空間的残響器を実装するために、フィードバック遅延ネットワーク（FDN）として知られる周知のフィルタ構造をもつフィルタが使用されることができる。FDNの構造は単純である。いくつかの残響タンク（たとえば、図4のFDNでは利得要素 $g_1$ および遅延線 $z^{-n_1}$ を有する残響タンク）を有し、各残響タンクは遅延および利得をもつ。FDNの典型的な実装では、すべての残響タンクからの出力は、ユニタリー・フィードバック・マトリクスによって混合され、該マトリクスの出力がフィードバックされて残響タンクの入力と合計される。残響タンク出力に利得調整がなされてもよい。残響タンク出力（またはその利得調整されたバージョン）はマルチチャネルまたはバイノーラル再生のために好適に再混合されることができる。コンパクトな計算およびメモリ・フットプリントをもつFDNによって、自然に聞こえる残響が生成され、適用されることができる。したがって、FDNは、HRTFによって生成された直接応答を補足するよう仮想化器において使用されてきた。

#### 【 0 0 1 3 】

たとえば、商業的に入手可能な「ドルビー・モバイル」ヘッドフォン仮想化器は、（左前方、右前方、中央、左サラウンドおよび右サラウンド・チャネルをもつ）五チャネル・オーディオ信号の各チャネルに残響を加え、五つの頭部伝達関数（「HRTF」）フィルタ対の集合の異なるフィルタ対を使って、それぞれの残響付加されたチャネルをフィルタリングするよう動作可能であるFDNベースの構造をもつ残響器を含む。「ドルビー・モバイル」ヘッドフォン仮想化器は、二チャネル・オーディオ入力信号にตอบสนองして二チャネルの「残響付加された」バイノーラル・オーディオ出力（残響が加えられた二チャネルの仮想サラウンド・サウンド出力）を生成するようにも動作可能である。残響付加されたバイノーラル出力がレンダリングされ、ヘッドフォン対によって再生されるとき、それは聴取者の鼓膜において、左前方、右前方、中央、左後方（サラウンド）および右後方（サラウンド）位置にある五つのラウドスピーカーからのHRTFフィルタリングされた残響付加された音

10

20

30

40

50

として知覚される。仮想化器は、ダウンミックスされた二チャンネル・オーディオ入力を（該オーディオ入力とともに受領されるいかなる空間的手がかりパラメータを使うことなく）アップミックスし、五つのアップミックスされたオーディオ・チャンネルを生成し、アップミックスされたチャンネルに残響を加え、五つの残響付加されたチャンネル信号をダウンミックスして仮想化器の二チャンネルの残響付加された出力を生成する。それぞれのアップミックスされたチャンネルについての残響はHRTFフィルタの異なる対においてフィルタリングされる。

【発明の概要】

【発明が解決しようとする課題】

【0014】

仮想化器では、FDNはある残響減衰時間およびエコー密度を達成するように構成される。しかしながら、FDNは早期反射のミクロ構造をシミュレートする柔軟性を欠く。さらに、通常の仮想化器では、FDNのチューニングおよび構成設定は大半が試行錯誤的なものである。

【0015】

すべての反射経路（早期および後期）をシミュレートするのではないヘッドフォン仮想化器は有効な頭外定位を達成できない。発明者は、すべての反射経路（早期および後期）をシミュレートしようとするFDNを用いる仮想化器は、通例、早期反射および後期残響の両方をシミュレートし、両方をオーディオ信号に加えることにおいて、高々限られた成功しか収めていないことを認識するに至った。発明者はまた、FDNを用いるが残響減衰時間、両耳間コヒーレンスおよび直接対後期比といった空間的な音響属性を適正に制御する能力をもたない仮想化器は、ある程度の頭外定位を達成するかもしれないが、過度の音色の歪みおよび残響を導入するという代償を伴うことをも認識するに至った。

【課題を解決するための手段】

【0016】

第一のクラスの実施形態では、本発明は、マルチチャンネル・オーディオ入力信号のチャンネルのある集合（たとえば、それらのチャンネルのそれぞれまたは全周波数範囲チャンネルのそれぞれ）にตอบสนองしてパイノーラル信号を生成する方法である。本方法は：（a）前記集合の各チャンネルに（たとえば前記集合の各チャンネルを前記チャンネルに対応するBRIRと畳み込みすることによって）パイノーラル室内インパルス応答（BRIR）を適用し、それによりフィルタリングされた信号を生成する段階であって、前記集合のチャンネルのダウンミックス（たとえばモノフォニック・ダウンミックス）に共通の後期残響を加えるよう少なくとも一つのフィードバック遅延ネットワーク（FDN）を使うことによることを含む、段階と；（b）フィルタリングされた信号を組み合わせるパイノーラル信号を生成する段階とを含む。典型的には、前記ダウンミックスに前記共通の後期残響を加えるために、FDNのバンクが使用される（たとえば、各FDNが異なる周波数帯域に共通の後期残響を加える）。典型的には、段階（a）は前記集合の各チャンネルに、該チャンネルについての単一チャンネルBRIRの「直接応答および早期反射」部分を適用する段階を含み、前記共通の後期残響は、前記単一チャンネルBRIRの少なくとも一部（たとえば全部）の後期残響部分の集団的なマクロ属性をエミュレートするように生成されたものである。

【0017】

マルチチャンネル・オーディオ入力信号にตอบสนองして（またはそのような信号のチャンネルのある集合にตอบสนองして）パイノーラル信号を生成する方法は、本稿では時に、「ヘッドフォン仮想化」方法と称され、そのような方法を実行するよう構成されたシステムは本稿では時に「ヘッドフォン仮想化器」（または「ヘッドフォン仮想化システム」または「パイノーラル仮想化器」）と称される。

【0018】

第一のクラスの典型的な実装では、各FDNはフィルタバンク領域（たとえば、ハイブリッド複素直交ミラー・フィルタ（HCQMF: hybrid complex quadrature mirror filter）領域または直交ミラー・フィルタ（QMF）領域または間引きを含みうる他の変換もしくは

10

20

30

40

50

はサブバンド領域)において実装される。いくつかのそのような実施形態では、バイノーラル信号の周波数依存の空間的な音響属性は、後期残響を加えるために用いられる各FDNの構成を制御することによって制御される。典型的には、マルチチャネル信号のオーディオ・コンテンツの効率的なバイノーラル・レンダリングのために、チャネルのモノフォニック・ダウンミックスがFDNへの入力として使われる。第一のクラスの典型的な実施形態は、たとえば各FDNの入力利得、残響タンク利得、残響タンク遅延または出力マトリクス・パラメータのうちの少なくとも一つを設定するよう制御値をフィードバック遅延ネットワークに呈することによって、周波数依存の属性(たとえば、残響減衰時間、両耳間コヒーレンス、モード密度および直接対後期比)に対応するFDN係数を調整する段階を含む。これは、音響環境のよりよいマッチングおよびより自然に聞こえる出力を可能にする。

10

#### 【0019】

第二のクラスの実施形態では、本発明は、諸チャネルを有するマルチチャネル・オーディオ入力信号に応答してバイノーラル信号を生成する方法である。これは、入力信号のチャネルのある集合の各チャネル(たとえば、入力信号のチャネルのそれぞれまたは入力信号のそれぞれの全周波数範囲チャネル)にバイノーラル室内インパルス応答(BRIR)を適用することによる。これは、前記集合の各チャネルを、該チャネルについての単一チャネルBRIRの直接応答および早期反射をモデル化して該各チャネルに適用するよう構成された第一の処理経路において処理し、前記集合のチャネルのダウンミックス(たとえばモノフォニック(モノ)・ダウンミックス)を、該ダウンミックスへの共通の後期残響をモデル化して適用するよう構成された(前記第一の処理経路と並列な)第二の処理経路において処理することによることを含む。典型的には、前記共通の後期残響は、前記単一チャネルBRIRのうち少なくともいくつか(たとえば全部)の後期残響部分の集団的なマクロ属性をエミュレートするよう生成されたものである。典型的には、第二の処理経路は少なくとも一つのFDN(たとえば複数の周波数帯域のそれぞれについて一つのFDN)を含む。典型的には、第二の処理経路によって実装される各FDNのすべての残響タンクへの入力として、モノ・ダウンミックスが使われる。典型的には、音響環境をよりよくシミュレートし、より自然に聞こえるバイノーラル仮想化を生じるために、各FDNのマクロ属性の系統的な制御のための機構が提供される。たいていのそのようなマクロ属性は周波数依存なので、各FDNは典型的にはハイブリッド複素直交ミラー・フィルタ(HCQMF)領域、周波数領域、領域または別のフィルタバンク領域において実装され、各周波数帯域について異なるまたは独立なFDNが使われる。FDNをフィルタバンク領域において実装することの主要な恩恵は、周波数依存の残響属性をもつ残響の適用を許容するということである。さまざまな実施形態において、FDNは、多様なフィルタバンクの任意のものを使って、幅広い多様なフィルタバンク領域の任意のものにおいて実装される。それは、実または複素数値の直交ミラー・フィルタ(QMF)、有限インパルス応答フィルタ(FIRフィルタ)、無限インパルス応答フィルタ(IIRフィルタ)、離散フーリエ変換(DFT)、(修正)コサインまたはサイン変換、ウェーブレット変換またはクロスオーバー・フィルタを含むがそれに限られない。ある好ましい実装では、用いられるフィルタバンクまたは変換は、FDNプロセスの計算上の複雑さを低減するために間引き(たとえば、周波数領域信号表現のサンプリング・レートの減少)を含む。

20

30

40

#### 【0020】

第一のクラス(および第二のクラス)のいくつかの実施形態は、以下の特徴の一つまたは複数を実装する。

#### 【0021】

1. フィルタバンク領域(たとえばハイブリッド複素直交ミラー・フィルタ領域)のFDN実装またはハイブリッドのフィルタバンク領域FDN実装および時間領域後期残響フィルタ実装。これは典型的には、各周波数帯域についてのFDNのパラメータおよび/または設定の独立な調整を許容する(これは、周波数依存の音響属性の単純で柔軟な制御を可能にする)。これはたとえば、モード密度を周波数の関数として変化させるよう異なる帯域における残響タンク遅延を変化させる能力を提供することによる。

50

## 【 0 0 2 2 】

2. (マルチチャンネル入力オーディオ信号から) 第二の処理経路において処理される、ダウンミックスされた(たとえばモノフォニック・ダウンミックスされた)信号を生成するために用いられる特定のダウンミックス・プロセスは、各チャンネルの源距離ならびに直接応答と後期応答の間の適正なレベルおよびタイミング関係を維持するための直接応答の扱いに依存する。

## 【 0 0 2 3 】

3. 結果として生じる残響のスペクトルおよび/または音色を変えることなく位相多様性(diversity)および増大したエコー密度を導入するために、第二の処理経路において(たとえばFDNのバンクの入力または出力において)全域通過フィルタ(APF: all-pass filter)が適用される。

10

## 【 0 0 2 4 】

4. ダウンサンプル因子格子(downsample-factor grid)に量子化された遅延に関連した問題を克服するために、複素数値のマルチレート構造における各FDNのフィードバック経路において、端数遅延(fractional delay)が実装される。

## 【 0 0 2 5 】

5. FDNにおいて、残響タンク出力は、各周波数帯域における所望される両耳間コヒーレンスに基づいて設定される出力混合係数を使って、バイノーラル・チャンネル中に直接、線形に混合される。任意的に、残響タンクの、バイノーラル出力チャンネルへのマッピングは、バイノーラル・チャンネル間の均衡した遅延を達成するために、諸周波数帯域を横断して交互する。また任意的に、残響タンク出力には、端数遅延および全体的なパワーを保存しつつそのレベルを等化するために、規格化因子が適用される。

20

## 【 0 0 2 6 】

6. 周波数依存の残響減衰時間および/またはモード密度が、実際の部屋をシミュレートするよう各周波数帯域における残響タンク遅延および利得の適正な組み合わせを設定することによって制御される。

## 【 0 0 2 7 】

7. 周波数帯域毎に(たとえば関連する処理経路の入力または出力のいずれかにおいて)一つのスケール因子が適用される。これは:

実際の部屋のDLRにマッチする周波数依存の直接対後期比(DLR: direct-to-late ratio)を制御する(目標DLRおよび残響減衰時間、たとえばT60に基づいて、必要とされるスケール因子を計算するために、単純なモデルが使用されてもよい);

30

過剰なコーミング(combining)アーチファクトおよび/または低周波数のごろごろ音(low-frequency rumble)を緩和するための低周波数減衰を提供する; および/または

FDN応答に拡散場スペクトル整形(diffuse field spectral shaping)を適用するためである。

## 【 0 0 2 8 】

8. 残響減衰時間、両耳間コヒーレンスおよび/または直接対後期比といった後期残響の本質的な周波数依存の属性を制御するために単純なパラメトリック・モデルが実装される。

40

## 【 0 0 2 9 】

本発明の諸側面は、オーディオ信号(たとえば、オーディオ・コンテンツがスピーカー・チャンネルからなるオーディオ信号および/またはオブジェクト・ベースのオーディオ信号)のバイノーラル仮想化を実行する(または実行するよう構成されているまたはその実行をサポートする)方法およびシステムを含む。

## 【 0 0 3 0 】

別のクラスの実施形態では、本発明は、マルチチャンネル・オーディオ入力信号のチャンネルのある集合にตอบสนองしてバイノーラル信号を生成する方法およびシステムである。これは、前記集合の各チャンネルにバイノーラル室内インパルス応答(BRIR)を適用し、それによりフィルタリングされた信号を生成する段階であって、前記集合のチャンネルのダウンミッ

50



クスに共通の後期残響を加えるよう単一のフィードバック遅延ネットワーク (FDN) を使うことによることを含む、段階と；フィルタリングされた信号を組み合わせるバイノーラル信号を生成する段階とを実行することによることを含む。FDNは時間領域で実装される。そのようないくつかの実施形態では、時間領域FDNは：

前記ダウンミックスを受領するよう結合された入力をもつ入力フィルタであって、該入力フィルタは前記ダウンミックスにตอบสนองして第一のフィルタリングされたダウンミックスを生成するよう構成されている、入力フィルタと；

前記第一のフィルタリングされたダウンミックスにตอบสนองして第二のフィルタリングされたダウンミックスをするよう結合され、構成された全域通過フィルタと；

第一の出力および第二の出力をもつ残響適用サブシステムであって、前記残響適用サブシステムは残響タンクの集合を含み、各残響タンクは異なる遅延をもち、該残響適用サブシステムは、前記第二のフィルタリングされたダウンミックスにตอบสนองして第一の未混合バイノーラル・チャンネルおよび第二の未混合バイノーラル・チャンネルを生成し、前記第一の未混合バイノーラル・チャンネルを前記第一の出力において呈し、前記第二の未混合バイノーラル・チャンネルを前記第二の出力において呈するよう結合され、構成されている、残響適用サブシステムと；

前記残響適用サブシステムに結合され、前記第一の未混合バイノーラル・チャンネルおよび第二の未混合バイノーラル・チャンネルにตอบสนองして第一の混合済みバイノーラル・チャンネルおよび第二の混合済みバイノーラル・チャンネルを生成するよう構成されている、両耳間相互相関係数 (IACC: interaural cross-correlation coefficient) フィルタリングおよび混合段とを含む。

#### 【0031】

入力フィルタは、各BRIRが少なくとも実質的に目標DLRにマッチする直接対後期比 (DLR) をもつよう前記第一のフィルタリングされたダウンミックスを生成するよう（好ましくは、それを生成するよう構成された二つのフィルタのカスケードとして）実装されてもよい。

#### 【0032】

各残響タンクは、遅延された信号を生成するよう構成されていてもよく、前記各残響タンクにおいて伝搬する信号に利得を加えて、遅延された信号が少なくとも実質的に目標の遅延された利得にマッチする利得をもつようにするよう結合され、構成された残響フィルタ（たとえば、シェルフ・フィルタまたはシェルフ・フィルタのカスケードとして実装される）を含んでいてもよい。各BRIRの目標残響減衰時間特性（たとえば $T_{60}$ 特性）を達成するためである。

#### 【0033】

いくつかの実施形態では、前記第一の未混合バイノーラル・チャンネルは前記第二の未混合バイノーラル・チャンネルより進んでおり、前記残響タンクは、最も短い遅延をもつ第一の遅延された信号を生成するよう構成された第一の残響タンクと、二番目に短い遅延をもつ第二の遅延された信号を生成するよう構成された第二の残響タンクとを含む。前記第一の残響タンクは前記第一の遅延された信号に第一の利得を適用するよう構成され、前記第二の残響タンクは前記第二の遅延された信号に第二の利得を適用するよう構成され、前記第二の利得は前記第一の利得とは異なり、前記第二の利得は前記第一の利得とは異なり、前記第一の利得および前記第二の利得の適用により、前記第二の未混合バイノーラル・チャンネルに対して前記第一の未混合バイノーラル・チャンネルの減衰が帰結する。典型的には、前記第一の混合済みバイノーラル・チャンネルおよび前記第二の混合済みバイノーラル・チャンネルは、再センタリングされた (re-centered) ステレオ像を示す。いくつかの実施形態では、前記IACCフィルタリングおよび混合段は、前記第一の混合済みバイノーラル・チャンネルおよび前記第二の混合済みバイノーラル・チャンネルが少なくとも実質的に目標IACC特性に一致するIACC特性をもつよう前記第一の混合済みバイノーラル・チャンネルおよび前記第二の混合済みバイノーラル・チャンネルを生成するよう構成されている。

#### 【0034】

本発明の典型的な実施形態は、スピーカー・チャンネルからなる入力オーディオおよびオブジェクト・ベースの入力オーディオの両方をサポートするための単純で統一された枠組みを提供する。オブジェクト・チャンネルである入力信号チャンネルにBRIRが適用される実施形態では、各オブジェクト・チャンネルに対して実行される「直接応答および早期反射」処理は、そのオブジェクト・チャンネルのオーディオ・コンテンツと一緒に提供されたメタデータによって示される源方向を想定する。スピーカー・チャンネルである入力信号チャンネルにBRIRが適用される実施形態では、各スピーカー・チャンネルに対して実行される「直接応答および早期反射」処理は、そのスピーカー・チャンネルに対応する源方向（すなわち、対応するスピーカーの想定される位置から想定される聴取者位置への直接経路の方向）を想定する。入力チャンネルがオブジェクト・チャンネルであるかスピーカー・チャンネルであるかに関わりなく、「後期残響」処理は、入力チャンネルのダウンミックス（たとえばモノフォニック・ダウンミックス）に対して実行され、ダウンミックスのオーディオ・コンテンツについてのいかなる特定の源方向も想定しない。

10

#### 【0035】

本発明の他の側面は、本発明の方法の任意の実施形態を実行するよう構成された（たとえばプログラムされた）ヘッドフォン仮想化器、そのような仮想化器を含むシステム（たとえばステレオ、マルチチャンネルまたは他のデコーダ）および本発明の方法の任意の実施形態を実装するためのコードを記憶するコンピュータ可読媒体（たとえばディスク）である。

20

#### 【図面の簡単な説明】

#### 【0036】

【図1】通常のヘッドフォン仮想化システムのブロック図である。

【図2】本発明のヘッドフォン仮想化システムのある実施形態を含むシステムのブロック図である。

【図3】本発明のヘッドフォン仮想化システムのもう一つの実施形態のブロック図である。

【図4】図3のシステムの典型的な実装に含まれる型のFDNのブロック図である。

【図5】二つの特定の周波数（ $f_A$ および $f_B$ ）のそれぞれにおける $T_{60}$ の値が、 $f_A = 10\text{Hz}$ で $T_{60,A} = 320\text{ms}$ および $f_B = 2.4\text{kHz}$ で $T_{60,B} = 150\text{ms}$ のように設定されている本発明の仮想化器のある実施形態によって達成されうる、Hz単位の周波数の関数としてのミリ秒単位での残響減衰時間（ $T_{60}$ ）のグラフである。

30

【図6】制御パラメータ $\text{Coh}_{\max}$ 、 $\text{Coh}_{\min}$ および $f_C$ が $\text{Coh}_{\max} = 0.95$ 、 $\text{Coh}_{\min} = 0.05$ および $f_C = 700\text{Hz}$ の値をもつよう設定されている本発明の仮想化器のある実施形態によって達成されうる、Hz単位の周波数の関数としての両耳間コヒーレンス（ $\text{Coh}$ ）のグラフである。

【図7】制御パラメータ $\text{DLR}_{1K}$ 、 $\text{DLR}_{\text{slope}}$ 、 $\text{DLR}_{\min}$ 、 $\text{HPF}_{\text{slope}}$ および $f_T$ が $\text{DLR}_{1K} = 18\text{dB}$ 、 $\text{DLR}_{\text{slope}} = \text{周波数10倍毎に6dB}$ 、 $\text{DLR}_{\min} = 18\text{dB}$ 、 $\text{HPF}_{\text{slope}} = \text{周波数10倍毎に6dB}$ および $f_T = 200\text{Hz}$ の値をもつよう設定されている本発明の仮想化器のある実施形態によって達成されうる、Hz単位の周波数の関数としての、1メートルの源距離でのdB単位での直接対後期比（ $\text{DLR}$ ）のグラフである。

【図8】本発明のヘッドフォン仮想化システムの後期残響処理サブシステムのもう一つの実施形態のブロック図である。

40

【図9】本発明のシステムのいくつかの実施形態に含まれる型のFDNの時間領域実装のブロック図である。

【図9A】図9のフィルタ400の実装の例のブロック図である。

【図9B】図9のフィルタ406の実装の例のブロック図である。

【図10】後期残響処理サブシステム221が時間領域で実装される本発明のヘッドフォン仮想化システムのある実施形態のブロック図である。

【図11】図9のFDNの要素422、423および424の実施形態のブロック図である。Aは、フィルタ500の典型的な実装の周波数応答（R1）、フィルタ501の典型的な実装の周波数応答（R2）およびフィルタ500と501を並列に接続したものの周波数応

50

答のグラフである。

【図 1 2】図 9 の FDN のある実装によって達成されうる IACC 特性（曲線「I」）および目標（target）IACC 特性（曲線「I<sub>T</sub>」）の例のグラフである。

【図 1 3】フィルタ 4 0 6、4 0 7、4 0 8 および 4 0 9 のそれぞれをシェルフ・フィルタとして適切に実装することによって図 9 の FDN のある実装によって達成されうる T60 特性のグラフである。

【図 1 4】フィルタ 4 0 6、4 0 7、4 0 8 および 4 0 9 のそれぞれを二つの IIR シェルフ・フィルタのカスケードとして適切に実装することによって図 9 の FDN のある実装によって達成されうる T60 特性のグラフである。

【発明を実施するための形態】

【0 0 3 7】

#### 記法および命名法

請求項を含む本開示を通じて、信号またはデータ「に対して」動作を実行する（たとえば信号またはデータをフィルタリングする、スケーリングする、変換するまたは利得を適用する）という表現は、信号またはデータに対して直接的に、または信号またはデータの処理されたバージョンに対して（たとえば、予備的なフィルタリングまたは前処理を該動作の実行に先立って受けている前記信号のバージョンに対して）該動作を実行することを表わすために広義で使用される。

【0 0 3 8】

請求項を含む本開示を通じて、「システム」という表現は、装置、システムまたはサブシステムを表わす広義で使用される。たとえば、仮想化器を実装するサブシステムは、仮想化器システムと称されてもよく、そのようなサブシステムを含むシステム（たとえば、複数の入力にตอบสนองして X 個の出力信号を生成するシステムであって、前記サブシステムが入力のうちの M 個を生成し、他の X - M 個の入力は外部源から受領されるもの）も仮想化器システム（または仮想化器）と称されることがある。

【0 0 3 9】

請求項を含む本開示を通じて、用語「プロセッサ」は、データ（たとえばオーディオまたはビデオまたは他の画像データ）に対して動作を実行するよう（たとえばソフトウェアまたはファームウェアを用いて）プログラム可能または他の仕方で構成可能であるシステムまたは装置を表わす広義で使用される。プロセッサの例は、フィールド・プログラム可能なゲート・アレイ（または他の構成可能な集積回路またはチップセット）、オーディオまたは他のサウンド・データに対してパイプライン化された処理を実行するようプログラムされたおよび/または他の仕方で構成されたデジタル信号プロセッサ、プログラム可能な汎用プロセッサもしくはコンピュータおよびプログラム可能なマイクロプロセッサ・チップまたはチップセットを含む。

【0 0 4 0】

請求項を含む本開示を通じて、表現「分解フィルタバンク」は、時間領域信号に対して変換（たとえば時間領域から周波数領域への変換）を適用して、一組の周波数帯域のそれぞれにおいて該時間領域信号の内容を示す値（たとえば周波数成分）を生成するよう構成されたシステム（たとえばサブシステム）を表わす広義で使用される。請求項を含む本開示を通じて、表現「フィルタバンク領域」は、変換または分解フィルタバンクによって生成される周波数成分の領域（たとえばそのような周波数成分が処理される領域）を表わす広義で使用される。フィルタバンク領域の例は（これに限られないが）周波数領域、直交ミラー・フィルタ（QMF）領域およびハイブリッド複素直交ミラー・フィルタ（HCQMF）領域を含む。分解フィルタバンクによって適用されうる変換の例は（これに限られないが）離散コサイン変換（DCT）、修正離散コサイン変換（MDCT）、離散フーリエ変換（DFT）およびウェーブレット変換を含む。分解フィルタバンクの例は（これに限られないが）直交ミラー・フィルタ（QMF）、有限インパルス応答フィルタ（FIR フィルタ）、無限インパルス応答フィルタ（IIR フィルタ）、クロスオーバー・フィルタおよび他の好適なマルチレート構造をもつフィルタを含む。

10

20

30

40

50

## 【 0 0 4 1 】

請求項を含む本開示を通じて、「メタデータ」という用語は、対応するオーディオ・データ（メタデータをも含むビットストリームの、オーディオ・コンテンツ）とは別個の異なるデータを指す。メタデータは、オーディオ・データに関連付けられ、該オーディオ・データの少なくとも一つの特徴または特性（たとえばそのオーディオ・データに対してどの型（単数または複数）の処理がすでに実行されているか、あるいは実行されるべきかまたはそのオーディオ・データによって示されるオブジェクトの軌跡）を示す。メタデータのオーディオ・データとの関連付けは、時間同期的である。このように、現在の（最も最近受領または更新された）メタデータは、対応するオーディオ・データが同時的に、示される特徴をもつおよび／または示される型のオーディオ・データ処理の結果を含むことを示しうる。

10

## 【 0 0 4 2 】

請求項を含む本開示を通じて、「結合する」または「結合される」という用語は、直接的または間接的な接続を意味するために使われる。よって、第一の装置が第二の装置に結合する場合、その接続は、直接接続を通じてであってもよいし、他の装置および接続を介した間接的な接続を通じてであってもよい。

## 【 0 0 4 3 】

請求項を含む本開示を通じて、以下の表現は以下の定義をもつ。

## 【 0 0 4 4 】

スピーカーおよびラウドスピーカーは、任意の音を発するトランスデューサを表わすものとして同義に使われる。この定義は、複数のトランスデューサ（たとえばウーファーおよびツイーター）として実装されるラウドスピーカーを含む。

20

## 【 0 0 4 5 】

スピーカー・フィード：ラウドスピーカーに直接加えられるオーディオ信号または直列の増幅器およびラウドスピーカーに加えられるオーディオ信号。

## 【 0 0 4 6 】

チャンネル（または「オーディオ・チャンネル」）：モノフォニック・オーディオ信号。そのような信号は典型的には、該信号を所望されるまたは公称上の位置にあるラウドスピーカーに直接加えるのと等価であるようにレンダリングされることができる。所望される位置は、物理的なラウドスピーカーでは典型的にそうであるように静的であってもよく、あるいは動的であってもよい。

30

## 【 0 0 4 7 】

オーディオ・プログラム：一つまたは複数のオーディオ・チャンネル（少なくとも一つのスピーカー・チャンネルおよび／または少なくとも一つのオブジェクト・チャンネル）および任意的には関連するメタデータ（たとえば、所望される空間的オーディオ呈示を記述するメタデータ）の集合。

## 【 0 0 4 8 】

スピーカー・チャンネル（または「スピーカー・フィード・チャンネル」）：（所望されるまたは公称上の位置にある）指定されたラウドスピーカーに関連付けられているまたは定義されたスピーカー配位内での指定されたスピーカー・ゾーンに関連付けられているオーディオ・チャンネル。スピーカー・チャンネルは、該オーディオ信号を（所望されるまたは公称上の位置にある）指定されたラウドスピーカーにまたは指定されたスピーカー・ゾーン内のスピーカーに直接加えるのと等価であるようにレンダリングされる。

40

## 【 0 0 4 9 】

オブジェクト・チャンネル：オーディオ源（時にオーディオ「オブジェクト」と称される）によって発される音を示すオーディオ・チャンネル。典型的には、オブジェクト・チャンネルは、パラメトリックなオーディオ源記述を決定する（たとえば、パラメトリックなオーディオ源記述を示すメタデータがオブジェクト・チャンネル内に含まれるまたはオブジェクト・チャンネルと一緒に提供される）。源記述は、（時間の関数としての）源によって発された音、時間の関数としての源の見かけの位置（たとえば、3D空間座標）および任意的

50

には源を特徴付ける少なくとも一つの追加的パラメータ（たとえば見かけの源サイズまたは幅）を決定してもよい。

【0050】

オブジェクト・ベースのオーディオ・プログラム：一つまたは複数のオブジェクト・チャンネルの集合を（および任意的には少なくとも一つのスピーカー・チャンネルも）および任意的には関連するメタデータ（たとえば、オブジェクト・チャンネルによって示される音を発するオーディオ・オブジェクトの軌跡を示すメタデータ、あるいは他の仕方でオブジェクト・チャンネルによって示される音の所望される空間的オーディオ呈示を示すメタデータまたはオブジェクト・チャンネルによって示される音の源である少なくとも一つのオーディオ・オブジェクトの識別情報を示すメタデータ）も含むオーディオ・プログラム。

10

【0051】

レンダリング：オーディオ・プログラムを一つまたは複数のスピーカー・フィールドに変換するプロセスまたはオーディオ・プログラムを一つまたは複数のスピーカー・フィールドに変換し、該スピーカー・フィールドを一つまたは複数のラウドスピーカーを使って音に変換するプロセス。（後者の場合、レンダリングは本稿では時にラウドスピーカー「による」レンダリングと称される。）オーディオ・チャンネルは、信号を所望される位置にある物理的なラウドスピーカーに直接加えることによって（所望される位置「において」）トリビアルにレンダリングされることができる。あるいは、一つまたは複数のオーディオ・チャンネルは、（聴取者にとって）そのようなトリビアルなレンダリングと実質的に等価であるよう設計された多様な仮想化技法の一つを使ってレンダリングされることができる。この後者の場合、各オーディオ・チャンネルは、一般には所望される位置とは異なる既知の位置にあるラウドスピーカー（単数または複数）に加えられるべき一つまたは複数のスピーカー・フィールドに変換されてもよく、それによりフィールドに応答してラウドスピーカーによって発される音は、所望される位置から発しているように知覚されることになる。そのような仮想化技法の例は、ヘッドフォンを介したバイノーラル・レンダリング（たとえばヘッドフォン装着者のために7.1チャンネルまでのサラウンド・サウンドをシミュレートする「ドルビー・ヘッドフォン」処理を使う）および波面合成（wave field synthesis）を含む。

20

【0052】

マルチチャンネル・オーディオ信号が「x.y」または「x.y.z」チャンネル信号であるという本稿での記法は信号が「x」個の全周波数スピーカー・チャンネル（想定される聴取者の耳の水平面に公称上位置されているスピーカーに対応）と、「y」個のLFE（またはサブウーファ）チャンネルと、任意的にはまた「z」個の全周波数頭上スピーカー・チャンネル（想定される聴取者の頭の上方に、たとえば部屋の天井またはその近くに位置されるスピーカーに対応）とを有することを表わす。

30

【0053】

表現「IACC」は、本稿では、その通常の意味での両耳間相互相関係数を表わす。これは、聴取者の耳でのオーディオ信号到達時刻の間の差の指標であり、典型的には、到達する信号が大きさにおいて等しく正確に逆相であることを示す第一の値から到達する信号が類似性をもたないことを示す中間的な値を経て、同じ振幅および位相をもつ同一の到達する信号を示す最大値までの範囲内の数によって示される。

40

【0054】

好ましい実施形態の詳細な説明

本発明の多くの実施形態が技術的に可能である。本開示からそれらをどのように実装するかは当業者には明確であろう。本発明のシステムおよび方法の実施形態を図2～図14を参照して記述する。

【0055】

図2は、本発明のヘッドフォン仮想化システムのある実施形態を含むシステム（20）のブロック図である。本ヘッドフォン仮想化システム（時に仮想化器と称される）は、マルチチャンネル・オーディオ入力信号のN個の全周波数範囲チャンネル（ $X_1, \dots, X_N$ ）にバイノ

50

ーラル室内インパルス応答 (BRIR) を適用するよう構成されている。チャンネル  $X_1, \dots, X_N$  (これらはスピーカー・チャンネルまたはオブジェクト・チャンネルでありうる) のそれぞれは、想定される聴取者に対する特定の源方向および距離に対応し、図 2 のシステムは、そのような各チャンネルを、対応する源方向および距離についての BRIR によって畳み込みするよう構成されている。

【 0 0 5 6 】

システム 20 は、エンコードされたオーディオ・プログラムを受領するよう結合されており、それから  $N$  個の全周波数範囲チャンネル ( $X_1, \dots, X_N$ ) を復元することによることを含め該プログラムをデコードし、それらを (図のように結合された要素 12、... 14、15、16、18 を有する) 仮想化システムの要素 12、...、14、15 に提供するよう結合され、構成されているサブシステム (図 2 には示さず) を含むデコーダであってもよい。デコーダは、追加的なサブシステムを含んでいてもよく、そのいくつかは、仮想化システムによって実行される仮想化機能に関係しない機能を実行し、そのいくつかは仮想化機能に関係する機能を実行してもよい。たとえば、後者の機能は、エンコードされたプログラムからのメタデータの抽出と、該メタデータを、該メタデータを用いて仮想化器システムの要素を制御する仮想化制御サブシステムに提供することとを含んでいてもよい。

【 0 0 5 7 】

サブシステム 12 は (サブシステム 15 とともに) チャンネル  $X_1$  を  $BRIR_1$  (対応する源方向および距離についての BRIR) と畳み込みするよう構成されており、サブシステム 14 は (サブシステム 15 とともに) チャンネル  $X_N$  を  $BRIR_N$  (対応する源方向についての BRIR) と畳み込みするよう構成されており、 $N-2$  個の他の BRIR サブシステムのそれぞれについても同様である。サブシステム 12、...、14、15 のそれぞれの出力は、左チャンネルおよび右チャンネルを含む時間領域信号である。加算要素 16 および 18 は要素 12、...、14、15 の出力に結合される。加算要素 16 は、諸 BRIR サブシステムの左チャンネル出力どうしを組み合わせる (混合する) よう構成されており、加算要素 18 は、諸 BRIR サブシステムの右チャンネル出力どうしを組み合わせる (混合する) よう構成されている。要素 16 の出力は、図 2 の仮想化器から出力されるバイノーラル・オーディオ信号の左チャンネル  $L$  であり、要素 18 の出力は、図 2 の仮想化器から出力されるバイノーラル・オーディオ信号の右チャンネル  $R$  である。

【 0 0 5 8 】

本発明の典型的な実施形態の重要な特徴は、本発明のヘッドフォン仮想化器の図 2 の実施形態を図 1 の通常のヘッドフォン仮想化器と比べることから明白になる。比較のために、図 1 および図 2 のシステムは、そのそれぞれに同じマルチチャンネル・オーディオ入力信号が呈されるとき、それらのシステムが同じ直接応答および早期反射部分 (すなわち、図 2 の関連する  $EBRIR_i$ ) をもつ  $BRIR_i$  を入力信号のそれぞれの全周波数範囲チャンネル  $X_i$  に適用するよう (必ずしも同じ度合いの成功ではないが)、構成されているとする。図 1 または図 2 のシステムによって適用される各  $BRIR_i$  は、直接応答および早期反射部分 (たとえば図 2 のサブシステム 12 ~ 14 によって適用される  $EBRIR_1$ 、...、 $EBRIR_N$  の一つ) と後期残響部分という二つの部分に分解できる。図 2 の実施形態 (および本発明の他の典型的な実施形態) は、複数の単一チャンネル BRIR、すなわち  $BRIR_i$  の後期残響部分が源方向を横断して、よってすべてのチャンネルを横断して共有されることができ、入力信号のすべての全周波数範囲チャンネルのダウンミックスに同じ後期残響 (すなわち共通の後期残響) を適用できることを想定する。このダウンミックスは、すべての入力チャンネルのモノフォニック (モノ) ダウンミックスであることができるが、代替的には、入力チャンネルから (たとえば入力チャンネルの部分集合から) 得られるステレオまたはマルチチャンネルのダウンミックスであってもよい。

【 0 0 5 9 】

より具体的には、図 2 のサブシステム 12 は、入力信号チャンネル  $X_1$  を  $EBRIR_1$  (対応する源方向についての直接応答および早期反射 BRIR 部分) と畳み込みするよう構成され、サブシステム 14 は、入力信号チャンネル  $X_N$  を  $EBRIR_N$  (対応する源方向についての直接応答およ

び早期反射BRIR部分)と畳み込みするよう構成される、などとなる。図2の後期残響サブシステム15は、入力信号のすべての全周波数範囲チャンネルのモノ・ダウンミックスを生成し、該ダウンミックスをLBRIR(ダウンミックスされるチャンネルのすべてについての共通の後期残響)と畳み込みするよう構成されている。図2の仮想化器の各BRIRサブシステム(サブシステム12、...、14、15のそれぞれ)の出力は、(対応するスピーカー・チャンネルまたはダウンミックスから生成されたバイノーラル信号の)左チャンネルおよび右チャンネルを含む。それらのBRIRサブシステムの左チャンネル出力は加算要素16において組み合わせられ(混合され)、それらのBRIRサブシステムの右チャンネル出力は加算要素18において組み合わせられる(混合される)。

【0060】

10

適切なレベル調整および時間整列がサブシステム12、...、14、15において実装されていると想定して、加算要素16は、対応する左バイノーラル・チャンネル・サンプル(サブシステム12、...、14、15の左チャンネル出力)を単に合計してバイノーラル出力信号の左チャンネルを生成するよう実装されることがができる。同様に、やはり適切なレベル調整および時間整列がサブシステム12、...、14、15において実装されていると想定して、加算要素18も、対応する右バイノーラル・チャンネル・サンプル(サブシステム12、...、14、15の右チャンネル出力)を単に合計してバイノーラル出力信号の右チャンネルを生成するよう実装されることがができる。

【0061】

20

図2のサブシステム15は、多様な仕方の任意のもので実装できるが、典型的には、それに呈される入力信号チャンネルのモノフォニック・ダウンミックスに共通の後期残響を加えるよう構成された少なくとも一つのフィードバック遅延ネットワークを含む。典型的には、サブシステム12、...、14のそれぞれが、処理対象のチャンネル( $X_i$ )についての単一チャンネルBRIRの直接応答および早期反射部分( $EBRIR_i$ )を適用する場合、共通の後期残響は、(その「直接応答および早期反射部分」がサブシステム12、...、14によって適用される)それらの単一チャンネルBRIRの少なくともいくつか(たとえば全部)の後期残響部分の集団的なマクロ属性をエミュレートするよう生成されている。たとえば、サブシステム15のある実装は、それに呈される入力信号チャンネルのモノフォニック・ダウンミックスに共通の後期残響を適用するよう構成されているフィードバック遅延ネットワーク(203、204、...、205)のバンクを含む、図3のサブシステム200と同じ構造をもつ。

30

【0062】

同様に、図2のサブシステム12、...、14は、(時間領域またはフィルタバンク領域の)多様な仕方の任意のもので実装でき、何らかの特定の用途のための好ましい実装は、(たとえば)パフォーマンス、計算およびメモリのようなさまざまな事情に依存する。ある例示的実装では、サブシステム12、...、14のそれぞれは、それに呈されるチャンネルを、そのチャンネルに関連付けられた直接および早期応答に対応するFIRフィルタと畳み込みするよう構成される。利得および遅延は、サブシステム12、...、14の出力がサブシステム15の出力と単純にかつ効率的に組み合わせられてもよいように適正に設定される。

【0063】

40

図3は、本発明のヘッドフォン仮想化システムのもう一つの実施形態のブロック図である。図3の実施形態は図2の実施形態と同様であり、二つの(左および右チャンネルの)時間領域信号が直接応答および早期反射処理サブシステム100から出力され、二つの(左および右チャンネルの)時間領域信号が後期残響処理サブシステム200から出力される。加算要素210がサブシステム100および200の出力に結合される。要素210は、サブシステム100および200の左チャンネル出力を組み合わせ(混合して)図3の仮想化器から出力されるバイノーラル・オーディオ信号の左チャンネルLを生成し、サブシステム100および200の右チャンネル出力を組み合わせ(混合して)図3の仮想化器から出力されるバイノーラル・オーディオ信号の右チャンネルRを生成するよう構成される。適切なレベル調整および時間整列がサブシステム100および200において実装されて

50

いると想定して、要素 2 1 0 は、サブシステム 1 0 0 および 2 0 0 から出力される対応する左チャンネル・サンプルを単純に合計してバイノーラル出力信号の左チャンネルを生成し、サブシステム 1 0 0 および 2 0 0 から出力される対応する右チャンネル・サンプルを単純に合計してバイノーラル出力信号の右チャンネルを生成するよう実装されることができ。

#### 【 0 0 6 4 】

図 3 のシステムでは、マルチチャンネル・オーディオ入力信号のチャンネル  $X_i$  は、二つの並列な処理経路に向けられ、そこで処理を受ける。一方は直接応答および早期反射処理サブシステム 1 0 0 を通り、他方は後期残響処理サブシステム 2 0 0 を通る。図 3 のシステムは、各チャンネル  $X_i$  に  $BRIR_i$  を適用するよう構成されている。各  $BRIR_i$  は、直接応答および早期反射部分（サブシステム 1 0 0 によって適用される）と後期残響部分（サブシステム 2 0 0 によって適用される）という二つの部分に分解できる。動作では、直接応答および早期反射処理サブシステム 1 0 0 はこうして仮想化器から出力されるバイノーラル・オーディオ信号の直接応答および早期反射部分を生成し、後期残響処理サブシステム（「後期残響生成器」）2 0 0 はこうして仮想化器から出力されるバイノーラル・オーディオ信号の後期残響部分を生成する。サブシステム 1 0 0 および 2 0 0 の出力は（加算サブシステム 2 1 0 によって）混合され、バイノーラル・オーディオ信号を生成し、該バイノーラル・オーディオ信号は典型的にはサブシステム 2 1 0 からレンダリング・システム（図示せず）に呈され、レンダリング・システムにおいてヘッドフォンによる再生のためのバイノーラル・レンダリングを受ける。

#### 【 0 0 6 5 】

典型的には、一対のヘッドフォンによってレンダリングされ、再生されるとき、要素 2 1 0 から出力される典型的なバイノーラル・オーディオ信号は聴取者の鼓膜において、聴取者の前方、背後および上方の位置を含む幅広い多様な位置の任意のところにある「N」個のラウドスピーカーからの音として知覚される（ここで N 2 であり、N は典型的には 2、5 または 7 である）。図 3 のシステムの動作において生成された出力信号の再生は、聴取者に、二つより多くの（たとえば五個または七個の）「サラウンド」源からくる音の経験を与えることができる。これらの源の少なくともいくつかは仮想的である。

#### 【 0 0 6 6 】

直接応答および早期反射処理サブシステム 1 0 0 は、（時間領域またはフィルタバンク領域の）多様な仕方の任意のもので実装でき、何らかの特定の用途のための好ましい実装は、（たとえば）パフォーマンス、計算およびメモリのようなささまざまな事情に依存する。ある例示的な実装では、サブシステム 1 0 0 は、それに呈される各チャンネルを、そのチャンネルに関連付けられた直接および早期応答に対応する FIR フィルタと畳み込みするよう構成される。利得および遅延は、サブシステム 1 0 0 の出力がサブシステム 2 0 0 の出力と（要素 2 1 0 において）単純にかつ効率的に組み合わせられてもよいように適正に設定される。

#### 【 0 0 6 7 】

図 3 に示されるように、後期残響生成器 2 0 0 は、ダウンミックス・サブシステム 2 0 1、分解フィルタバンク 2 0 2、FDN（FDN 2 0 3、2 0 4、...、2 0 5）のバンクおよび合成フィルタバンク 2 0 7 を図のように結合したものを含む。サブシステム 2 0 1 は、マルチチャンネル入力信号のチャンネルをモノ・ダウンミックスにダウンミックスするよう構成されており、分解フィルタバンク 2 0 2 はモノ・ダウンミックスに変換を適用して、モノ・ダウンミックスを「K」個の周波数帯域に分割するよう構成されている。ここで、K は整数である。それぞれの異なる周波数帯域における（フィルタバンク 2 0 2 から出力される）フィルタバンク領域値は、FDN 2 0 3、2 0 4、...、2 0 5 のうちの異なるものに呈される（これらの FDN は「K」個あり、それぞれそれに呈されたフィルタバンク領域値に  $BRIR$  の後期残響部分を適用するよう結合され、構成されている）。フィルタバンク領域値は好ましくは、FDN の計算上の複雑さを軽減するよう、時間において間引きされる。

#### 【 0 0 6 8 】

原理的には、（図 3 のサブシステム 1 0 0 およびサブシステム 2 0 1 への）各入力チャ

10

20

30

40

50



ネルは、そのBRIRの後期残響部分をシミュレートするよう独自のFDN（またはFDNのバンク）によって処理されることができる。異なる音源位置に関連付けられたBRIRの後期残響部分が典型的にはインパルス応答における二乗平均平方根の点では非常に異なっているという事実にもかかわらず、その平均パワー・スペクトル、そのエネルギー減衰構造、モード密度、ピーク密度などといった統計的な属性はしばしば非常に似通っている。したがって、一組のBRIRの後期残響部分は典型的には、チャンネルを横断して知覚的にきわめて似通っている。二つ以上のBRIRの後期残響部分をシミュレートするために一つの共通のFDNまたはFDN（たとえば、FDN 203、204、...、205）のバンクを使うことが可能である。典型的な実施形態では、そのような一つの共通のFDN（またはFDNのバンク）が用いられ、それへの入力、入力チャンネルから構築された一つまたは複数のダウンミックスから構成される。図2の例示的な実装では、ダウンミックスはすべての入力チャンネルのモノフォニック・ダウンミックス（サブシステム201の出力において呈される）である。

10

#### 【0069】

図2の実施形態を参照するに、FDN 203、204、...、205のそれぞれは、フィルタバンク領域において実装され、分解フィルタバンク202から出力される値のうちの異なる周波数帯域を処理して、各帯域についての左および右の残響付加された信号を生成するよう結合され、構成される。各帯域について、左の残響付加された信号はフィルタバンク領域値のシーケンスであり、右の残響付加された信号はフィルタバンク領域値の別のシーケンスである。合成フィルタバンク207は、周波数領域から時間領域への変換を、フィルタバンク領域値（たとえばQMF領域の周波数成分）の2K個のシーケンスに適用し、変換された値を集めて（後期残響が適用されたモノ・ダウンミックスのオーディオ・コンテンツを示す）左チャンネル時間領域信号および（やはり後期残響が適用されたモノ・ダウンミックスのオーディオ・コンテンツを示す）右チャンネル時間領域信号にする。これらの左チャンネルおよび右チャンネルの信号は要素210に出力される。

20

#### 【0070】

典型的な実装では、FDN 203、204、...、205のそれぞれはQMF領域で実装され、フィルタバンク202はサブシステム201からのモノ・ダウンミックスをQMF領域（たとえば、ハイブリッド複素直交ミラー・フィルタ（HCQMF）領域）に変換し、それにより、フィルタバンク202からFDN 203、204、...、205のそれぞれの入力に呈される信号はQMF領域周波数成分のシーケンスとなる。そのような実装では、フィルタバンク202からFDN 203に呈される信号は第一の周波数帯域におけるQMF領域周波数成分のシーケンスであり、フィルタバンク202からFDN 204に呈される信号は第二の周波数帯域におけるQMF領域周波数成分のシーケンスであり、フィルタバンク202からFDN 205に呈される信号は第「K」の周波数帯域におけるQMF領域周波数成分のシーケンスである。分解フィルタバンク202がそのように実装されるとき、合成フィルタバンク207はQMF領域から時間領域への変換をFDNからの出力QMF領域周波数成分の2K個のシーケンスに適用し、要素210に出力される左チャンネルおよび右チャンネルの後期残響付加された時間領域信号を生成する。

30

#### 【0071】

たとえば、図3のシステムにおいてK=3であれば、合成フィルタバンク207に対する六つの入力（FDN 203、204および205のそれぞれから出力される周波数領域またはQMF領域サンプルを含む、左および右のチャンネル）および207からの二つの出力（それぞれ時間領域サンプルからなる左および右のチャンネル）がある。この例では、フィルタバンク207は典型的には二つの合成フィルタバンクとして実装される。一つ（FDN 203、204および205からの三つの左チャンネルが呈されるもの）はフィルタバンク207から出力される時間領域左チャンネル信号を生成するよう構成され、第二のもの（FDN 203、204および205からの三つの右チャンネルが呈されるもの）はフィルタバンク207から出力される時間領域右チャンネル信号を生成するよう構成される。

40

#### 【0072】

任意的に、制御サブシステム209は、FDN 203、204、...、205のそれぞれ

50

に結合され、サブシステム 200 によって適用される後期残響部分 (LBRIR) を決定するためにそれらFDNのそれぞれに対して制御パラメータを呈するよう構成される。そのような制御パラメータの例を以下で述べる。いくつかの実装では、制御サブシステム 209 は、サブシステム 200 によって入力チャネルのモノフォニック・ダウンミックスに適用される後期残響部分 (LBRIR) のリアルタイム変動を実装するよう、リアルタイムで（すなわち、入力装置によってそれに呈されるユーザー・コマンドに応答して）動作可能であることが考えられる。

【0073】

たとえば、図 2 のシステムへの入力信号が 5.1 チャンネル信号（その全周波数範囲チャンネルは次のチャンネル順：L,R,C,Ls,Rsである）であれば、すべての全周波数範囲チャンネルは同じ源距離をもち、ダウンミックス・サブシステム 201 は次のダウンミックス行列として実装されることができる。これは単に全周波数範囲チャンネルを合計してモノ・ダウンミックスを形成する。

【0074】

【数 1】

$$D = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

（FDN 203、204、...、205 のそれぞれにおける要素 301 内の）全域通過フィルタリング後、モノ・ダウンミックスはパワーを保存する仕方で四つの残響タンクにアップミックスされる。

【0075】

【数 2】

$$U = \begin{bmatrix} 1/\sqrt{4} \\ 1/\sqrt{4} \\ 1/\sqrt{4} \\ 1/\sqrt{4} \end{bmatrix}$$

あるいはまた、（一例として）左側の諸チャンネルを最初の二つの残響タンクにパンし、右側の諸チャンネルを最後の二つの残響タンクにパンし、中央チャンネルをすべての残響タンクにパンすることを選ぶことができる。この場合、ダウンミックス・サブシステム 201 は二つのダウンミックス信号を形成するよう実装されることになる。

【0076】

【数 3】

$$D = \begin{bmatrix} 1 & 0 & 1/\sqrt{2} & 1 & 0 \\ 0 & 1 & 1/\sqrt{2} & 0 & 1 \end{bmatrix}$$

この例では、（FDN 203、204、...、205 のそれぞれにおける）残響タンクへのアップミックスは次のようになる。

【0077】

10

20

30

40

## 【数 4】

$$U = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

二つのダウンミックス信号があるので、(FDN 203、204、...、205のそれぞれにおける要素301内の)全域通過フィルタリングは二度適用される必要がある。(L,Ls)、(R,Rs)およびCの後期応答について、そのすべてが同じマクロ属性をもつにもかかわらず、多様性が導入される。入力信号チャンネルが異なる源距離をもつときは、いまだダウンミックス・プロセスにおいて適正な遅延および利得が適用される必要がある。

10

## 【0078】

次に、図3の仮想化器のダウンミックス・サブシステム201ならびにサブシステム100および200の個別的な実装についての考察を述べる。

## 【0079】

サブシステム201によって実装されるダウンミックス・プロセスは、ダウンミックスされるべき各チャンネルについての(音源と想定される聴取者位置との間の)源距離と、直接応答の扱いとに依存する。直接応答の遅延 $t_d$ は：

$$t_d = d/v_s$$

20

である。ここで、 $d$ は音源と聴取者との間の距離であり、 $v_s$ は音速である。さらに、直接応答の利得は $1/d$ に比例する。これらのルールが異なる源距離をもつチャンネルの直接応答の扱いにおいて保存されるならば、サブシステム201は、すべてのチャンネルのストレートなダウンミックスを実装できる。後期残響の遅延およびレベルは一般に、源位置に敏感ではないからである。

## 【0080】

実際の事情のため、仮想化器(たとえば図3の仮想化器のサブシステム100)は、異なる源距離をもつ入力チャンネルについての直接応答を時間整列させるよう実装されてもよい。各チャンネルについての直接応答と後期残響との間の相対的な遅延を保存するために、源距離 $d$ をもつチャンネルは他のチャンネルとダウンミックスされる前に $(d_{\max} - d)/v_s$ だけ遅延させられるべきである。ここで、 $d_{\max}$ は最大可能な源距離を表わす。

30

## 【0081】

仮想化器(たとえば図3の仮想化器のサブシステム100)は、直接応答のダイナミックレンジを圧縮するようにも実装されてもよい。たとえば、源距離 $d$ をもつチャンネルについての直接応答は、 $d^{-1}$ の代わりに因子 $d^{-1}$ によってスケールされてもよい。ここで、 $0 < d^{-1} < 1$ である。直接応答と後期残響との間のレベル差を保存するために、ダウンミックス・サブシステム201は、源距離 $d$ をもつチャンネルを、他のスケールされたチャンネルとダウンミックスする前に、因子 $d^{-1}$ によってスケールするよう実装される必要があることがある。

## 【0082】

40

図4のフィードバック遅延ネットワークは図3のFDN 203(または204または205)の例示的な実装である。図4のシステムは四つの残響タンク(それぞれ利得段 $g_i$ および遅延線 $z^{-n_i}$ を含む)をもつが、このシステムの変形(および本発明の仮想化器の実施形態において用いられる他のFDN)は四つより多いまたは四つより少ない残響タンクを実装する。

## 【0083】

図4のFDNは、入力利得要素300と、要素300の出力に結合された全域通過フィルタ(APF: all-pass filter)301と、APF 301の出力に結合された加算要素302、303、304および305と、それぞれ要素302、303、304および305の異なるものの出力に結合された四つの残響タンクとを含む(各残響タンクは、利得要素

50

$g_k$  (要素 3 0 6 の一つ) と、それに結合された遅延線  $z^{-Mk}$  (要素 3 0 7 の一つ) と、それに結合された利得要素  $1/g_k$  (要素 3 0 9 の一つ) とを有し、 $0 \leq k \leq 3$  )。ユニタリー・マトリクス 3 0 8 が遅延線 3 0 7 の出力に結合され、要素 3 0 2、3 0 3、3 0 4 および 3 0 5 のそれぞれの第二の入力に対してフィードバック出力を呈するよう構成されている。利得要素 3 0 9 のうちの二つのもの (第一および第二の残響タンク) の出力は、加算要素 3 1 0 の入力に呈され、要素 3 1 0 の出力は出力混合マトリクス 3 1 2 の一方の入力に呈される。利得要素 3 0 9 のうちの他の二つのもの (第三および第四の残響タンク) の出力は、加算要素 3 1 1 の入力に呈され、要素 3 1 1 の出力は出力混合マトリクス 3 1 2 の他方の入力に呈される。

#### 【0084】

要素 3 0 2 は、遅延線  $z^{-n1}$  に対応するマトリクス 3 0 8 の出力を、第一の残響タンクの入力に加える (すなわち、マトリクス 3 0 8 を介した遅延線  $z^{-n1}$  の出力からのフィードバックを適用する) よう構成されている。要素 3 0 3 は、遅延線  $z^{-n2}$  に対応するマトリクス 3 0 8 の出力を、第二の残響タンクの入力に加える (すなわち、マトリクス 3 0 8 を介した遅延線  $z^{-n2}$  の出力からのフィードバックを適用する) よう構成されている。要素 3 0 4 は、遅延線  $z^{-n3}$  に対応するマトリクス 3 0 8 の出力を、第三の残響タンクの入力に加える (すなわち、マトリクス 3 0 8 を介した遅延線  $z^{-n3}$  の出力からのフィードバックを適用する) よう構成されている。要素 3 0 5 は、遅延線  $z^{-n4}$  に対応するマトリクス 3 0 8 の出力を、第四の残響タンクの入力に加える (すなわち、マトリクス 3 0 8 を介した遅延線  $z^{-n4}$  の出力からのフィードバックを適用する) よう構成されている。

#### 【0085】

図 4 の FDN の入力利得要素 3 0 0 は、図 3 の分解フィルタバンク 2 0 2 から出力される変換されたモノフォニック・ダウンミックス信号 (フィルタバンク領域信号) の一つの周波数帯域を受領するよう結合されている。入力利得要素 3 0 0 は、それに呈されるフィルタバンク領域信号に、利得 (スケーリング) 因子  $G_{in}$  を適用する。集団的に、すべての周波数帯域についての (図 3 の FDN 2 0 3、2 0 4、...、2 0 5 すべてによって実装される) スケーリング因子  $G_{in}$  は、後期残響のスペクトル整形およびレベルを制御する。図 3 の仮想化器のすべての FDN における入力利得  $G_{in}$  を設定することは、しばしば以下の目標を考慮に入れる:

実際の部屋にマッチする、各チャネルに適用される BRIR の直接対後期比 (DLR);

過剰なコーミング・アーチファクトおよび / または低周波数のごろごろ音を緩和するための必要な低周波数減衰;

拡散場スペクトル包絡のマッチング。

#### 【0086】

(図 3 のサブシステム 1 0 0 によって適用される) 直接応答がすべての周波数帯域において単位利得 (unitary gain) を提供するとすると、特定の DLR (パワー比) は:

$$G_{in} = \sqrt{\ln(10^6) / (T60 * DLR)}$$

となるよう  $G_{in}$  を設定することによって、達成できる。ここで、T60 は、残響が 60dB 減衰するのにかかる時間として定義される残響減衰時間 (これは以下で論じる残響遅延および残響利得によって決定される) であり、「ln」は自然対数関数を表わす。

#### 【0087】

入力利得因子  $G_{in}$  は処理されているコンテンツに依存してもよい。そのようなコンテンツ依存性の一つの応用は、入力チャネル信号間に存在するいかなる相関にもかかわらず、各時間 / 周波数セグメントにおけるダウンミックスのエネルギーが、ダウンミックスされる個々のチャネル信号のエネルギーの和に等しいことを保証することである。その場合、入力利得因子は

【数 5】

$$\sqrt{\frac{\sum_i \sum_j x_i^2(j)}{\sum_j y^2(j)}}$$

と似たまたはこれに等しい項であることができる（あるいはそのような項を乗算されることができる）。ここで、 $i$ は所与の時間／周波数タイルまたはサブバンドのすべてのダウンミックス・サンプルにわたるインデックスであり、 $y(i)$ はそのタイルについてのダウンミックス・サンプルであり、 $x_i(j)$ はダウンミックス・サブシステム 201 の入力に呈される（チャンネル $X_i$ についての）入力信号である。

10

【0088】

図4のFDNの典型的なQMF領域実装では、全域通過フィルタ（APF）301の出力から残響タンクの入力に呈される信号はQMF領域周波数成分のシーケンスである。より自然に聞こえるFDN出力を生成するために、利得要素300の出力にAPF 301が適用されて、位相多様性および増大したエコー密度を導入する。代替的または追加的に、一つまたは複数の全域通過フィルタが、（図3の）ダウンミックス・サブシステム201への個々の入力に、該入力サブシステム201においてダウンミックスされてFDNによって処理される前に適用されてもよく、あるいは図4に描かれる残響タンク・フィードフォワードまたはフィードバック経路において（たとえば、各残響タンクにおける遅延線 $z^{-M_k}$ に加えてまたはその代わりに）適用されてもよく、あるいはFDNの出力に（すなわち、出力マトリクス312の出力に）適用されてもよい。

20

【0089】

残響タンク遅延 $z^{-n_i}$ を実装する際、残響モードが同じ周波数で整列するのを避けるために、残響遅延 $n_i$ は互いに素であるべきである。遅延の合計は、人工的に聞こえる出力を避けるために、十分なモード密度を提供するよう十分大きいべきである。だが、最短の遅延は、後期残響とBRIRの他の成分との間の過剰な時間ギャップを避けるために、十分短いべきである。

【0090】

30

典型的には、残響タンク出力は、初期には、左または右のバイノーラル・チャンネルのいずれかにパンされる。通常、二つのバイノーラル・チャンネルにパンされている残響タンク出力のセットは同数であり、相互排他的である。二つのバイノーラル・チャンネルのタイミングを均衡させることも望まれる。よって、最短の遅延をもつ残響タンク出力が一方のバイノーラル・チャンネルに行くならば、二番目に短い遅延をもつ残響タンク出力は他方のチャンネルに行くことになる。

【0091】

周波数の関数としてモード密度を変えるよう、残響タンク遅延は周波数帯域を横断して異なることができる。一般に、より低い周波数帯域はより高いモード密度を必要とし、よってより長い残響タンク遅延を必要とする。

40

【0092】

残響タンク利得 $g_i$ の振幅および残響タンク遅延は、合同して図4のFDNの残響遅延時間を決定する：

$$T_{60} = -3n_i / \log_{10}(|g_i|) / F_{FRM}$$

ここで、 $F_{FRM}$ は（図3の）フィルタバンク202のフレーム・レートである。残響タンク利得の位相は、残響タンク遅延がフィルタバンクのダウンサンプル因子格子に量子化されていることに関係する問題を克服するよう、端数遅延を導入する。

【0093】

ユニタリー・フィードバック・マトリクス308は、フィードバック経路における諸残響タンクの間の均等な混合を提供する。

50

## 【 0 0 9 4 】

残響タンク出力のレベルを等化するために、利得要素 3 0 9 は規格化利得  $1/|g_i|$  を各残響タンクの出力に適用し、残響タンク利得のレベル効果を除去する一方でその位相によって導入される端数遅延を保存する。

## 【 0 0 9 5 】

出力混合マトリクス 3 1 2 ( 行列  $M_{out}$  としても特定される ) は、初期パニングからの未混合バイノーラル・チャンネル ( それぞれ要素 3 1 0 および 3 1 1 の出力 ) を混合して、所望される両耳間コヒーレンスをもつ出力の左および右のバイノーラル・チャンネル ( マトリクス 3 1 2 の出力において呈される L および R 信号 ) を達成するよう構成された  $2 \times 2$  のマトリクスである。未混合バイノーラル・チャンネルは、初期パニング後には、共通の残響タンク出力を全く含まないので、ほとんど無相関である。所望される両耳間コヒーレンスが  $Coh$  であり、 $|Coh| = 1$  とすると、出力混合マトリクス 3 1 2 は

## 【 数 6 】

$$M_{out} = \begin{bmatrix} \cos \beta & \sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}, \text{ ここで } \beta = \arcsin(Coh)/2$$

と定義されてもよい。残響タンク遅延が異なるので、未混合バイノーラル・チャンネルの一方が常時他方より進んでいる。残響タンク遅延およびパニング・パターンの組み合わせが周波数帯域を横断して同一であれば、音像バイアスが帰結するであろう。このバイアスは、混合済みバイノーラル・チャンネルが交互の周波数帯域において互いに進んだり遅れたりするよう、パニング・パターンが周波数帯域を横断して交互にされるならば、緩和できる。これは、出力混合マトリクス 3 1 2 を、奇数番目の周波数帯域においては ( たとえば、第一の周波数帯域 ( 図 3 の FDN 2 0 3 によって処理される )、第三の周波数帯域などにおいては ) 前の段落で述べた形をもつよう、偶数番目の周波数帯域においては ( たとえば、第二の周波数帯域 ( 図 3 の FDN 2 0 4 によって処理される )、第四の周波数帯域などにおいては )

## 【 数 7 】

$$M_{out,alt} = \begin{bmatrix} \sin \beta & \cos \beta \\ \cos \beta & \sin \beta \end{bmatrix}$$

の形をもつよう、実装することによって、達成されることができる。ここで、 の定義は同じままである。マトリクス 3 1 2 はすべての周波数帯域において同一であるよう実装されることができるが、交互の周波数帯域についてその入力チャンネル順が切り換えられてもよいことを注意しておくべきである。( たとえば、奇数周波数帯域では要素 3 1 0 の出力がマトリクス 3 1 2 の第一の入力に呈されてもよく、要素 3 1 1 の出力がマトリクス 3 1 2 の第二の入力に呈されてもよく、偶数周波数帯域では要素 3 1 1 の出力がマトリクス 3 1 2 の第一の入力に呈されてもよく、要素 3 1 0 の出力がマトリクス 3 1 2 の第二の入力に呈されてもよい。 )

周波数帯域が ( 部分的に ) 重なり合う場合には、それについてマトリクス 3 1 2 の形が交互に変えられるような周波数範囲の幅を増すことができる ( たとえば、二つまたは三つの連続する帯域ごとに一度変えることができる )。あるいは、連続する周波数帯域のスペクトル重なりについて補償するよう平均コヒーレンスが所望される値に等しいことを保証するために、( マトリクス 3 1 2 の形についての ) 上記の式における の値が調整されることができる。

## 【 0 0 9 6 】

本発明の仮想化器におけるそれぞれの個別の周波数帯域についてのFDNについて、上記で定義した目標音響属性 $T_{60}$ 、 $Coh$ および $DLR$ が既知であれば、各FDN（各FDNは図4に示した構造を有していてもよい）は目標属性を達成するように構成されることができる。特に、いくつかの実施形態では、本稿に記載される関係に従って目標属性を達成するように、各FDNについての入力利得（ $G_{in}$ ）および残響タンクの利得および遅延（ $g_i$ および $n_i$ ）ならびに出力マトリクス $M_{out}$ のパラメータが（たとえば図3の制御サブシステム209によってそれに呈される制御値により）設定されることができる。實際上、特定の音響環境にマッチする自然に聞こえる後期残響を生成するために、単純な制御パラメータをもつモデルによって周波数依存の属性を設定することが十分であることがしばしばである。

【0097】

10

次に、本発明の仮想化器のある実施形態の各特定の周波数帯域についてのFDNについての目標残響減衰時間（ $T_{60}$ ）が少数の周波数帯域のそれぞれについて目標残響減衰時間（ $T_{60}$ ）を決定することによってどのように決定できるかの例を述べる。FDN応答のレベルは時間とともに指数関数的に減衰する。 $T_{60}$ は減衰因子（decay factor） $df$ （単位時間に対するdB減衰として定義される）に反比例する、すなわち：

$T_{60} = 60/df$   
である。

【0098】

減衰因子 $df$ は周波数に依存し、一般に、対数周波数スケールに対して線形に増大する。よって、残響減衰時間も、周波数の関数であり、周波数が増加するにつれて一般に減少する。したがって、二つの周波数点についての $T_{60}$ の値を決定（たとえば設定）すれば、すべての周波数についての $T_{60}$ 曲線が決定される。たとえば、周波数点 $f_A$ および $f_B$ についての残響減衰時間がそれぞれ $T_{60,A}$ および $T_{60,B}$ であれば、 $T_{60}$ 曲線は次のように定義される。

20

【0099】

【数8】

$$T_{60}(f) = \frac{T_{60,A} T_{60,B} \log(f_B / f_A)}{T_{60,A} \log(f / f_A) - T_{60,B} \log(f / f_B)}$$

30

図5は、二つの特定の周波数（ $f_A$ および $f_B$ ）のそれぞれにおいて $T_{60}$ 値が $f_A = 10\text{Hz}$ において $T_{60,A} = 320\text{ms}$ および $f_B = 2.4\text{kHz}$ において $T_{60,B} = 150\text{ms}$ に設定される本発明の仮想化器のある実施形態によって達成されうる $T_{60}$ 曲線の例を示している。

【0100】

次に、本発明の仮想化器のある実施形態の各特定の周波数帯域についてのFDNについての目標両耳間コヒーレンス（ $Coh$ ）が少数の制御パラメータを設定することによってどのように達成できるかの例を述べる。後期残響の両耳間コヒーレンス（ $Coh$ ）はおおむね拡散音場のパターンに従う。それはクロスオーバー周波数 $f_C$ までの $\text{sinc}$ 関数およびクロスオーバー周波数より上での定数によってモデル化できる。 $Coh$ 曲線についての単純なモデルは次のようなものである。

40

【0101】

【数9】

$$Coh(f) = \begin{cases} Coh_{min} + (Coh_{max} - Coh_{min}) \text{sinc}(f / f_C) & , f \leq f_C \\ Coh_{min} & , f \geq f_C \end{cases}$$

ここで、パラメータ $Coh_{min}$ および $Coh_{max}$ は  $-1 \leq Coh_{min} < Coh_{max} \leq 1$  を満たし、 $Coh$ の範囲

50

を制御する。最適なクロスオーバー周波数 $f_c$ は聴取者の頭のサイズに依存する。高すぎる $f_c$ は頭の中に定位される音源像につながり、一方、小さすぎる $f_c$ は拡散したまたは分割された音源像につながる。図6は、制御パラメータ $Coh_{max}$ 、 $Coh_{min}$ および $f_c$ が次の値： $Coh_{max} = 0.95$ 、 $Coh_{min} = 0.05$ および $f_c = 700\text{Hz}$ をもつよう設定された本発明のある実施形態によって達成されうるCoh曲線の例である。

【0102】

次に、本発明の仮想化器のある実施形態の各特定の周波数帯域についてのFDNについての目標直接対後期比（DLR）が少数の制御パラメータを設定することによってどのように達成できるかの例を述べる。dB単位での直接対後期比（DLR）は一般に、対数周波数に対して線形に増大し、 $DLR_{1K}$ （1kHzでのdB単位でのDLR）と $DLR_{slope}$ （周波数10倍当たりのdB単位）を設定することによって制御される。しかしながら、低周波数範囲での低いDLRはしばしば過剰なコーミング・アーチファクトにつながる。該アーチファクトを緩和するために、DLRを制御する二つの修正機構が加えられる：

最小DLRフロア、 $DLR_{min}$ （dB単位）；および

遷移周波数 $f_T$ およびそれより下の減衰曲線の傾き $HPF_{slope}$ （周波数10倍当たりのdB単位）によって定義される高域通過フィルタ（high-pass filter）。

【0103】

dB単位での、結果として得られるDLR曲線は、次のように定義される。

【0104】

【数10】

$$DLR(f) = \max(DLR_{1K} + DLR_{slope} \log_{10}(f/1000), DLR_{min}) \\ + \min(HPF_{slope} \log_{10}(f/f_T), 0)$$

DLRはたとえ同じ音響環境にあっても源距離とともに変化することを注意しておくべきである。したがって、ここでの $DLR_{1K}$ および $DLR_{min}$ は1メートルなどの公称源距離についての値である。図7は、制御パラメータ $DLR_{1K}$ 、 $DLR_{slope}$ 、 $DLR_{min}$ 、 $HPF_{slope}$ および $f_T$ が次の値： $DLR_{1K} = 18\text{dB}$ 、 $DLR_{slope} = 6\text{dB/周波数10x}$ 、 $DLR_{min} = 18\text{dB}$ 、 $HPF_{slope} = 6\text{dB/周波数10x}$ および $f_T = 200\text{Hz}$ をもつよう設定された本発明の仮想化器のある実施形態によって達成される、1メートルの源距離についてのDLR曲線の例である。

【0105】

本稿に開示される実施形態の変形は次の特徴のうちの一つまたは複数をもち：

本発明の仮想化器は、時間領域で実装される、あるいはFDNベースのインパルス応答捕捉およびFIRベースの信号フィルタリングをもつハイブリッド実装をもつ；

本発明の仮想化器は、後期残響処理サブシステムのためのダウンミックスされた入力信号を生成するダウンミックス段階の実行中に、周波数の関数としてエネルギー補償の適用を許容するよう実装される；

本発明の仮想化器は、外部因子にตอบสนองして（すなわち、制御パラメータの設定にตอบสนองして）適用される後期残響属性の手動または自動的な制御を許容するよう実装される。

【0106】

システム・レイテンシーが決定的であり、分解および合成フィルタバンクによって引き起こされる遅延が禁止的である用途については、本発明の仮想化器の典型的な実施形態のフィルタバンク領域FDN構造は時間領域に変換されることができ、各FDN構造は本仮想化器のあるクラスの実施形態では時間領域で実装されることができる。時間領域実装では、入力利得因子（ $G_{in}$ ）、残響タンク利得（ $g_i$ ）および規格化利得（ $1/|g_i|$ ）を適用するサブシステムは、周波数依存の制御を許容するために同様の振幅応答をもつフィルタによって置き換えられる。出力混合マトリクス（ $M_{out}$ ）もフィルタのマトリクスによって置き換えられる。他のフィルタと異なり、フィルタのこのマトリクスの位相応答は枢要である。該位相応答によってパワー保存および両耳間コヒーレンスが影響されうるからである。時間

10

20

30

40

50



領域実装における残響タンク遅延は、共通因子としてフィルタバンク・ストライドを共有することを避けるために（フィルタバンク領域実装における値とは）わずかに変えられる必要があることがある。さまざまな制約条件のため、本発明の仮想化器のFDNの時間領域実装の実行は、そのフィルタバンク実装の場合に正確にマッチしないことがある。

【0107】

図8を参照して、次に、本発明の仮想化器の本発明の後期残響処理サブシステムのハイブリッド（フィルタバンク領域および時間領域）実装を記述する。本発明の後期残響処理サブシステムのこのハイブリッド実装は、図4の後期残響処理サブシステム200に対する変形であり、FDNに基づくインパルス応答捕捉およびFIRに基づく信号フィルタリングを実装する。

【0108】

図8は、図3のサブシステム200の同一の符号を付けられた要素と同一である要素201、202、203、204、205および207を含む。これらの要素の上記の記述は図8の参照では繰り返さない。図8の実施形態では、単位インパルス生成器211が分解フィルタバンク202への入力信号（パルス）を呈するよう結合される。FIRフィルタとして実装されるLBRIRフィルタ208（モノ入力、ステレオ出力）は該BRIR（LBRIR）の適切な後期残響部分を、サブシステム201から出力されたモノフォニック・ダウンミックスに対して適用する。こうして、要素211、202、203、204、205および207は、LBRIRフィルタ208に対する処理サイドチェーンである。

【0109】

後期残響部分LBRIRの設定が修正されるときはいつも、インパルス生成器211は、単位インパルスを要素202に対して呈するよう動作させられ、フィルタバンク207からの結果的な出力が捕捉され、（フィルタバンク207の出力によって決定された新たなLBRIRを適用するようフィルタ208を設定するため）フィルタ208に呈される。LBRIR設定変更から新たなLBRIRが有効になる時間までの時間経過を加速するために、新たなLBRIRのサンプルは、利用可能になるにつれて古いLBRIRを置き換えはじめることができる。FDNの内在的なレイテンシーを短縮するため、LBRIRの最初の諸ゼロは破棄できる。これらのオプションは、柔軟性を提供し、該ハイブリッド実装がFIRフィルタリングから追加される計算を代償として、（フィルタバンク領域実装によって提供されるパフォーマンスに比して）潜在的なパフォーマンス改善を提供することを許容する。

【0110】

システム・レイテンシーが枢要であるが計算パワーがそれほど問題ではない用途については、サイドチェーン・フィルタバンク領域後期残響処理器（たとえば、図8の要素211、202、203、204、...、205によって実装されるもの）が、フィルタ208によって適用される有効FIRインパルス応答を補足するために使われることができる。FIRフィルタ208はこの捕捉されたFIR応答を実装し、（入力チャネルの仮想化の間に）入力チャネルのモノ・ダウンミックスに直接適用することができる。

【0111】

さまざまなFDNパラメータ、よって結果として得られる後期残響属性は、手動でチューニングされ、その後、本発明の後期残響処理サブシステムの実施形態に固定構成として組み込まれることができる。たとえば、システムのユーザーによって（たとえば図3の制御サブシステム209を操作することによって）調整されることのできる一つまたは複数のプリセットによってである。しかしながら、後期残響の高レベルの記述、FDNパラメータとのその関係およびその挙動を修正する能力を与えられれば、FDNベースの後期残響処理器のさまざまな実施形態を制御するための幅広い多様な方法が構想される。それは以下のものを含む（ただしそれに限られない）。

【0112】

1. エンドユーザーは、たとえば（たとえば図3の制御サブシステム209の実施形態によって実装される）ディスプレイ上のユーザー・インターフェースによってFDNパラメータを手動で制御し、あるいは（たとえば図3の制御サブシステム209の実施形態によ

10

20

30

40

50

って実装される)物理的なコントロールを使ってプリセットを切り換えてもよい。このようにして、エンドユーザーは、好み、環境またはコンテンツに応じて部屋シミュレーションを適応させることができる。

【0113】

2. 仮想化されるべきオーディオ・コンテンツの作者が、たとえば入力オーディオ信号と一緒に提供されるメタデータによって、コンテンツ自身と一緒に伝達される設定または所望されるパラメータを提供してもよい。そのようなメタデータは、パースされ、関連するFDNパラメータを制御するために(たとえば図3の制御サブシステム209の実施形態によって)用いられてもよい。したがって、メタデータは、残響時間、残響レベル、直接対残響比などといった属性を示してもよく、これらの属性は時間変化して、時間変化するメタデータによって示されてもよい。

10

【0114】

3. 再生装置が、一つまたは複数のセンサーによってその位置または環境を認識してもよい。たとえば、モバイル装置は、該装置がどこにあるかを判別するために、GSMネットワーク、全地球測位システム(GPS)、既知のWiFiアクセスポイントまたは他の任意の位置特定サービスを使ってもよい。その後、位置および/または環境を示すデータが、関連するFDNパラメータを制御するために(たとえば図3の制御サブシステム209の実施形態によって)用いられてもよい。こうして、FDNパラメータは、装置の位置に応答して、たとえば物理的環境を模倣するよう、修正されうる。

【0115】

20

4. 再生装置の位置に関して、ある種の環境において消費者たちが使っている最も一般的な設定を導出するために、クラウド・サービスまたはソーシャル・メディアが使われてもよい。さらに、ユーザーは自分の現在の設定を、(既知の)位置と関連付けて、クラウドまたはソーシャル・メディア・サービスにアップロードして、他のユーザーまたは自分自身のために利用可能にしてもよい。

【0116】

5. 再生装置が、ユーザーの活動およびユーザーがいる環境を判別するために、カメラ、光センサー、マイクロフォン、加速度計、ジャイロスコープといった他のセンサーを含んでいてもよい。その特定の活動および/または環境についてFDNパラメータを最適化するためである。

30

【0117】

6. FDNパラメータは、オーディオ・コンテンツによって制御されてもよい。オーディオ分類アルゴリズムまたは手動で注釈付けされたコンテンツが、オーディオの諸セグメントが発話、音楽、サウンド効果、無音などを含むかどうかを示してもよい。FDNパラメータはそのようなラベルに従って調整されてもよい。たとえば、直接対残響比は、ダイアログ理解性を改善するために、ダイアログについては低減されてもよい。さらに、現在のビデオ・セグメントの位置を判別するためにビデオ解析が使われてもよく、FDNパラメータはビデオにおいて描かれている環境をよりよくシミュレートするためにしかるべく調整されてもよい。および/または

7. 半導体再生システムは、モバイル装置とは異なるFDN設定を使ってもよい。たとえば、設定は装置依存であってもよい。居間にある半導体システムは、典型的な(かなり残響のある)遠方の源をもつ居間シナリオをシミュレートしてもよく、一方、モバイル装置は聴取者により近くコンテンツをレンダリングしてもよい。

40

【0118】

本発明の仮想化器のいくつかの実装は、整数サンプル遅延のほか端数遅延を適用するよう構成されているFDN(たとえば、図4のFDNの実装)を含む。たとえば、そのようなある実装では、整数個のサンプル期間に等しい整数遅延を加える遅延線と直列に、各残響タンク内で端数遅延要素が接続される(たとえば、各端数遅延要素は遅延線の一つの後に、または他の仕方ですれと直列に位置される)。端数遅延は、各周波数帯域において、 $f$ が遅延割合(fraction)、 $T$ がその帯域についての

50

サンプル期間であるとして、サンプル期間のある割合  $f = 1/T$  に対応する位相シフト（単位複素数乗算）によって近似できる。QMF領域において残響を適用するコンテキストにおいて、どのようにして端数遅延を加えるかはよく知られている。

#### 【0119】

第一のクラスの実施形態では、本発明は、マルチチャネル・オーディオ入力信号のチャネルのある集合（たとえば、それらのチャネルのそれぞれまたは全周波数範囲チャネルのそれぞれ）にตอบสนองしてバイノーラル信号を生成するヘッドフォン仮想化方法である。本方法は：（a）前記集合の各チャネルに（たとえば図3のサブシステム100および200においてまたは図2のサブシステム12、...、14、15において前記集合の各チャネルを前記チャネルに対応するBRIRと畳み込みすることによって）バイノーラル室内インパルス応答（BRIR）を適用し、それによりフィルタリングされた信号（たとえば、図3のサブシステム100および200の出力または図2のサブシステム12、...、14、15の出力）を生成する段階であって、前記集合のチャネルのダウンミックス（たとえばモノフォニック・ダウンミックス）に共通の後期残響を加えるよう少なくとも一つのフィードバック遅延ネットワーク（たとえば図3のFDN 203、204、...、205）を使うことによることを含む、段階と；（b）フィルタリングされた信号を（たとえば図3のサブシステム210または図2の要素16および18を含むサブシステムにおいて）組み合わせてバイノーラル信号を生成する段階とを含む。典型的には、前記ダウンミックスに前記共通の後期残響を加えるために、FDNのバンクが使用される（たとえば、各FDNが異なる周波数帯域に後期残響を加える）。典型的には、段階（a）は（たとえば図3のサブシステム100または図2のサブシステム12、...、14において）前記集合の各チャネルに、該チャネルについての単一チャネルBRIRの「直接応答および早期反射」部分を適用する段階を含み、前記共通の後期残響は、前記単一チャネルBRIRの少なくとも一部（たとえば全部）の後期残響部分の集団的なマクロ属性をエミュレートするよう生成されたものである。

#### 【0120】

第一のクラスの典型的な実装では、各FDNはハイブリッド複素直交ミラー・フィルタ（HCQMF: hybrid complex quadrature mirror filter）領域または直交ミラー・フィルタ（QMF）領域において実装される。いくつかのそのような実施形態では、バイノーラル信号の周波数依存の空間的な音響属性は、後期残響を加えるために用いられる各FDNの構成を制御することによって（たとえば図3の制御サブシステム209を使って）制御される。典型的には、マルチチャネル信号のオーディオ・コンテンツの効率的なバイノーラル・レンダリングのために、チャネルのモノフォニック・ダウンミックス（たとえば、図3のサブシステム201によって生成されたダウンミックス）がFDNへの入力として使われる。典型的には、ダウンミックス・プロセスは、各チャネルについての源距離（すなわち、チャネルのオーディオ・コンテンツの想定される源と想定されるユーザー位置との間の距離）に基づいて制御され、各BRIR（すなわち、あるチャネルについての単一チャネルBRIRの直接応答および早期反射部分ならびにそのチャネルを含むダウンミックスについての共通の後期残響によって決定される各BRIR）の時間的およびレベル構造を保存するために源距離に対応する直接応答の扱いに依存する。ダウンミックスされるべきチャネルはダウンミックスの間に種々の仕方で時間整列され、スケーリングされることができるが、各チャネルについてのBRIRの直接応答、早期反射および共通の後期残響部分の間の適正なレベルおよび時間的關係が維持されるべきである。（ダウンミックスを生成するよう）ダウンミックスされるすべてのチャネルについて共通の後期残響部分を生成するために単一のFDNバンクを使う実施形態では、ダウンミックスの生成の間に（ダウンミックスされる各チャネルに対して）適正な利得および遅延が適用される必要がある。

#### 【0121】

このクラスの典型的な実施形態は、周波数依存の属性（たとえば、残響減衰時間、両耳間コヒーレンス、モード密度および直接対後期比）に対応するFDN係数を調整する段階を含む。これは、音響環境のよりよいマッチングおよびより自然に聞こえる出力を可能にする。

10

20

30

40

50

## 【 0 1 2 2 】

第二のクラスの実施形態では、本発明は、マルチチャンネル・オーディオ入力信号にตอบสนองしてバイノーラル信号を生成する方法である。これは、入力信号のチャンネルのある集合の各チャンネル（たとえば、入力信号のチャンネルのそれぞれまたは入力信号のそれぞれの全周波数範囲チャンネル）にバイノーラル室内インパルス応答（BRIR）を適用する（たとえば各チャンネルを対応するBRIRと畳み込みすることによって）ことによる。これは、前記集合の各チャンネルを、該チャンネルについての単一チャンネルBRIRの直接応答および早期反射（たとえば、図2のサブシステム12、14または15によって適用されるEBRIR）をモデル化して該各チャンネルに適用するよう構成された第一の処理経路（たとえば、図3のサブシステム100または図2のサブシステム12、...、14によって実装される）において処理し、前記集合のチャンネルのダウンミックス（たとえばモノフォニック・ダウンミックス）を、前記第一の処理経路と並列な第二の処理経路（たとえば、図3のサブシステム200または図2のサブシステム15によって実装される）において処理することによることを含む。第二の処理経路は、共通の後期残響（たとえば、図2のサブシステム15によって適用されるLBRIR）をモデル化して該ダウンミックスに適用するよう構成されている。典型的には、前記共通の後期残響は、前記単一チャンネルBRIRのうち少なくともいくつか（たとえば全部）の後期残響部分の集団的なマクロ属性をエミュレートする。典型的には、第二の処理経路は少なくとも一つのFDN（たとえば複数の周波数帯域のそれぞれについて一つのFDN）を含む。典型的には、第二の処理経路によって実装される各FDNのすべての残響タンクへの入力として、モノ・ダウンミックスが使われる。典型的には、音響環境をよりよくシミュレートし、より自然に聞こえるバイノーラル仮想化を生じるために、各FDNのマクロ属性の系統的な制御のための機構が提供される（たとえば図3の制御サブシステム209）。たいていのそのようなマクロ属性は周波数依存なので、各FDNは典型的にはハイブリッド複素直交ミラー・フィルタ（HCQMF）領域、周波数領域、領域または別のフィルタバンク領域において実装され、各周波数帯域について異なるFDNが使われる。FDNをフィルタバンク領域において実装することの主要な恩恵は、周波数依存の残響属性をもつ残響の適用を許容するということである。さまざまな実施形態において、FDNは、多様なフィルタバンクの任意のものを使って、幅広い多様なフィルタバンク領域の任意のものにおいて実装される。それは、直交ミラー・フィルタ（QMF）、有限インパルス応答フィルタ（FIRフィルタ）、無限インパルス応答フィルタ（IIRフィルタ）またはクロスオーバー・フィルタを含むがそれに限られない。

## 【 0 1 2 3 】

第一のクラス（および第二のクラス）のいくつかの実施形態は、以下の特徴の一つまたは複数を実装する。

## 【 0 1 2 4 】

1. フィルタバンク領域（たとえばハイブリッド複素直交ミラー・フィルタ領域）のFDN実装（たとえば図4のFDN実装）またはハイブリッド・フィルタバンク領域のFDN実装および時間領域の後期残響フィルタ実装（たとえば図8を参照して記述した構造）。これは典型的には、各周波数帯域についてのFDNのパラメータおよび/または設定の独立な調整を許容する（これは、周波数依存の音響属性の単純で柔軟な制御を可能にする）。これはたとえば、モード密度を周波数の関数として変化させるよう種々の帯域における残響タンク遅延に変化をつける能力を提供することによる。

## 【 0 1 2 5 】

2. （マルチチャンネル入力オーディオ信号から）第二の処理経路において処理されたダウンミックスされた（たとえばモノフォニック・ダウンミックスされた）信号を生成するために用いられる特定のダウンミックス・プロセスは、各チャンネルの源距離ならびに直接応答と後期応答の間の適正なレベルおよびタイミング関係を維持するための直接応答の扱いに依存する。

## 【 0 1 2 6 】

3. 結果として生じる残響のスペクトルおよび/または音色を変えることなく位相多様

10

20

30

40

50

性 (diversity) および増大したエコー密度を導入するために、第二の処理経路において (たとえばFDNのバンクの入力または出力において) 全域通過フィルタ (たとえば図4のAPF 301) が適用される。

【0127】

4. ダウンサンプル因子格子 (downsample-factor grid) に量子化された遅延に関係した問題を克服するために、複素数値のマルチレート構造における各FDNのフィードバック経路において、端数遅延 (fractional delay) が実装される。

【0128】

5. FDNにおいて、残響タンク出力は、各周波数帯域における所望される両耳間コヒーレンスに基づいて設定される出力混合係数を使って、(たとえば図4のマトリクス312によって) バイノーラル・チャンネル中に直接、線形に混合される。任意的に、残響タンクの、バイノーラル出力チャンネルへのマッピングは、バイノーラル・チャンネル間で均衡した遅延を達成するために、諸周波数帯域を横断して交互する。また任意的に、残響タンク出力には、端数遅延および全体的なパワーを保存しつつそのレベルを等化するために、規格化因子が適用される。

【0129】

6. 周波数依存の残響減衰時間が、実際の部屋をシミュレートするよう各周波数帯域における残響タンク遅延および利得の適正な組み合わせを設定することによって制御される。

【0130】

7. 周波数帯域毎に (たとえば関連する処理経路の入力または出力のいずれかにおいて) 一つのスケーリング因子が (たとえば図4の要素306および309によって) 適用される。これにより:

実際の部屋のDLRにマッチする周波数依存の直接対後期比 (DLR: direct-to-late ratio) を制御する (目標DLRおよび残響減衰時間、たとえばT60に基づいて、必要とされるスケーリング因子を計算するために、単純なモデルが使用されてもよい);

過剰なコーミング (combing) アーチファクトを緩和するための低周波数減衰を提供する; および / または

FDN応答に拡散場スペクトル整形 (diffuse field spectral shaping) を適用する。

【0131】

8. 残響減衰時間、両耳間コヒーレンスおよび / または直接対後期比といった後期残響の本質的な周波数依存の属性を制御するために (たとえば図3の制御サブシステム209によって) 単純なパラメトリック・モデルが実装される。

【0132】

いくつかの実施形態では (たとえば、システム・レイテンシーが決定的であり、分解および合成フィルタバンクによって引き起こされる遅延が禁止的である用途については)、本発明のシステムの典型的な実施形態のフィルタバンク領域FDN構造 (たとえば各周波数帯域における図4のFDN) は時間領域で実装されるFDN構造 (たとえば、図9に示されるように実装されうる図10のFDN 220) によって置き換えられる。本発明のシステムの時間領域実施形態では、入力利得因子 ( $G_{in}$ )、残響タンク利得 ( $g_i$ ) および規格化利得 ( $1/|g_i|$ ) を適用するフィルタバンク領域実施形態のサブシステムは、周波数依存の制御を許容するために時間領域フィルタ (および / または利得要素) によって置き換えられる。典型的なフィルタバンク領域実装の出力混合マトリクス (たとえば、図4の出力混合マトリクス312) は (典型的な時間領域実施形態では) 時間領域フィルタの出力集合 (たとえば、図9の要素424の図11の実装の要素500 ~ 503) によって置き換えられる。典型的な時間領域実施形態の他のフィルタと異なり、フィルタのこの出力集合の位相応答は典型的には枢要である (該位相応答によってパワー保存および両耳間コヒーレンスが影響されうるから)。いくつかの時間領域実施形態では、残響タンク遅延は、(たとえば、共通因子としてフィルタバンク・ストライドを共有することを避けるために) 対応するフィルタバンク領域実装における値から変えられる (たとえばわずかに変えられる)。

## 【 0 1 3 3 】

図 1 0 は、図 3 と同様の本発明のヘッドホン仮想化システムの実施形態のブロック図であるが、図 3 の要素 2 0 2 ~ 2 0 7 が図 1 0 のシステムでは、時間領域で実装される単一の FDN 2 2 0 によって置き換えられている（たとえば、図 1 0 の FDN 2 2 0 は図 9 の FDN と同様に実装されてもよい）。図 1 0 では、二つの（左および右チャンネルの）時間領域信号が、直接応答および早期反射処理サブシステム 1 0 0 から出力され、二つの（左および右チャンネルの）時間領域信号が、後期残響処理サブシステム 2 2 1 から出力される。サブシステム 1 0 0 および 2 0 0 の出力に加算要素 2 1 0 が結合されている。要素 2 1 0 は、サブシステム 1 0 0 および 2 2 1 の左チャンネル出力を組み合わせ（混合して）図 1 0 の仮想化器から出力されるバイノーラル・オーディオ信号の左チャンネル L を生成し、サブシステム 1 0 0 および 2 2 1 の右チャンネル出力を組み合わせ（混合して）図 1 0 の仮想化器から出力されるバイノーラル・オーディオ信号の右チャンネル R を生成するよう構成される。適切なレベル調整および時間整列がサブシステム 1 0 0 および 2 2 1 において実装されていると想定して、要素 2 1 0 は、サブシステム 1 0 0 および 2 2 1 から出力される対応する左チャンネル・サンプルを単純に合計してバイノーラル出力信号の左チャンネルを生成し、サブシステム 1 0 0 および 2 2 1 から出力される対応する右チャンネル・サンプルを単純に合計してバイノーラル出力信号の右チャンネルを生成するよう実装されることができ

10

## 【 0 1 3 4 】

図 1 0 のシステムでは、（チャンネル  $X_i$  をもつ）マルチチャンネル・オーディオ入力信号は、二つの並列な処理経路に向けられ、そこで処理を受ける。一方は直接応答および早期反射処理サブシステム 1 0 0 を通り、他方は後期残響処理サブシステム 2 2 1 を通る。図 1 0 のシステムは、各チャンネル  $X_i$  に  $BRIR_i$  を適用するよう構成されている。各  $BRIR_i$  は、直接応答および早期反射部分（サブシステム 1 0 0 によって適用される）と後期残響部分（サブシステム 2 2 1 によって適用される）という二つの部分に分解できる。動作では、直接応答および早期反射処理サブシステム 1 0 0 はこうして仮想化器から出力されるバイノーラル・オーディオ信号の直接応答および早期反射部分を生成し、後期残響処理サブシステム（「後期残響生成器」）2 2 1 はこうして仮想化器から出力されるバイノーラル・オーディオ信号の後期残響部分を生成する。サブシステム 1 0 0 および 2 2 1 の出力は（サブシステム 2 1 0 によって）混合され、バイノーラル・オーディオ信号を生成し、該バイノーラル・オーディオ信号は典型的にはサブシステム 2 1 0 からレンダリング・システム（図示せず）に呈され、レンダリング・システムにおいてヘッドホンによる再生のためのバイノーラル・レンダリングを受ける。

20

30

## 【 0 1 3 5 】

（後期残響処理サブシステム 2 2 1 の）ダウンミックス・サブシステム 2 0 1 は、マルチチャンネル入力信号のチャンネルをモノ・ダウンミックス（これは時間領域信号）にダウンミックスするよう構成されており、FDN 2 2 0 は後期残響部分をモノ・ダウンミックスに適用するよう構成されている。

## 【 0 1 3 6 】

図 9 を参照して、次に、図 1 0 の仮想化器の FDN 2 2 0 として用いることのできる時間領域 FDN の例を記述する。図 9 の FDN は、マルチチャンネル・オーディオ入力信号のすべてのチャンネルのモノ・ダウンミックス（たとえば図 1 0 のシステムのサブシステム 2 0 1 によって生成される）を受領するよう結合された入力フィルタ 4 0 0 を含む。図 9 の FDN は、フィルタ 4 0 0 の出力に結合された（図 4 の APF 3 0 1 に対応する）全域通過フィルタ（APF）4 0 1 と、フィルタ 4 0 1 の出力に結合された入力利得要素 4 0 1 A と、要素 4 0 1 A の出力に結合された加算要素 4 0 2、4 0 3、4 0 4 および 4 0 5（これらは図 4 の加算要素 3 0 2、3 0 3、3 0 4 および 3 0 5 に対応する）と、四つの残響タンクとを含む。各残響タンクは、要素 4 0 2、4 0 3、4 0 4 および 4 0 5 の異なるものの出力に結合され、残響フィルタ 4 0 6 および 4 0 6 A、4 0 7 および 4 0 7 A、4 0 8 および 4 0 8 A ならびに 4 0 9 および 4 0 9 A のうちのひとつと、それに結合された遅延線 4 1 0

40

50

、411、412および413のうちの一つ（図4の遅延線307に対応）と、これらの遅延線の一つの出力に結合された利得要素417、418、419および420のうちの一つとを有する。

#### 【0137】

ユニタリー・マトリクス415（図4のユニタリー・マトリクス308に対応し、典型的にはマトリクス308と同一であるよう実装される）が遅延線410、411、412および413の出力に結合される。マトリクス415は、要素402、403、404および405のそれぞれの第二の入力に対してフィードバック出力を呈するよう構成されている。

#### 【0138】

線410によって加えられる遅延（ $n1$ ）が線411によって加えられる遅延（ $n2$ ）より短く、線411によって加えられる遅延が線412によって加えられる遅延（ $n3$ ）より短く、線412によって加えられる遅延が線413によって加えられる遅延（ $n4$ ）より短いとき、（第一および第三の残響タンクの）利得要素417および419の出力が、加算要素422の入力に呈され、（第二および第四の残響タンクの）利得要素418および420の出力が、加算要素423の入力に呈される。要素422の出力はIACCおよび混合フィルタ424の一方の入力に呈され、要素423の出力はIACCフィルタリングおよび混合段424の他方の入力に呈される。

#### 【0139】

図9の利得要素417～420および要素422、423および424の実装の例を、図4の要素310および311ならびに出力混合マトリクス312の典型的な実装を参照しつつ述べる。図4の出力混合マトリクス312（行列 $M_{out}$ としても特定される）は、初期パニングからの未混合バイノーラル・チャンネル（それぞれ要素310および311の出力）を混合して、所望される両耳間コヒーレンスをもつ左および右のバイノーラル出力チャンネル（マトリクス312の出力において呈される左耳「L」および右耳「R」信号）を生成するよう構成された $2 \times 2$ のマトリクスである。この初期パニングは要素310および311によって実装される。そのそれぞれは二つの残響タンク出力を組み合わせる未混合バイノーラル・チャンネルの一つを生成し、最も短い遅延をもつ残響タンク出力は要素310の入力に呈され、二番目に短い遅延をもつ残響タンク出力は要素311の入力に呈される。図9の実施形態の要素422および423は、（それらの入力に対して呈された時間領域信号に対して、）図4の実施形態の（各周波数帯域における）要素310および311がそれらの入力に呈された（関連する周波数帯域における）フィルタバンク領域成分のストリームに対して実行するのと同じ型の初期パニングを実行する。

#### 【0140】

共通の残響タンク出力を全く含まないので、ほとんど無相関である前記未混合バイノーラル・チャンネル（図4の要素310および311からまたは図9の要素422および423から出力されるもの）は、左右のバイノーラル出力チャンネルについての所望される両耳間コヒーレンスを達成するパニング・パターンを実装するよう（図4のマトリクス312または図9の段424によって）混合されてもよい。しかしながら、残響タンク遅延が各FDN（すなわち、図9のFDNまたは図4におけるそれぞれの異なる周波数帯域について実装されるFDN）において異なるので、一方の未混合バイノーラル・チャンネル（要素310および311または422および423の一方の出力）が常時他方の未混合バイノーラル・チャンネル（要素310および311または422および423の他方の出力）より進んでいる。

#### 【0141】

このように、図4の実施形態では、残響タンク遅延およびパニング・パターンの組み合わせがすべての周波数帯域を横断して同一であれば、音像バイアスが帰結するであろう。このバイアスは、混合済みバイノーラル出力チャンネルが交互の周波数帯域において互いに進んだり遅れたりするよう、パニング・パターンが周波数帯域を横断して交互にされるならば、緩和できる。たとえば、所望される両耳間コヒーレンスがCohであり、 $|Coh| = 1$ と

10

20

30

40

50

すると、奇数番目の周波数帯域における出力混合マトリクス 3 1 2 はそれに呈される二つの入力を次の形

【数 1 1】

$$M_{out} = \begin{bmatrix} \cos \beta & \sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}, \text{ ここで } \beta = \arcsin(Coh)/2$$

をもつ行列によって乗算するよう実装されてもよく、偶数番目の周波数帯域における出力混合マトリクス 3 1 2 はそれに呈される二つの入力を次の形

【数 1 2】

$$M_{out,alt} = \begin{bmatrix} \sin \beta & \cos \beta \\ \cos \beta & \sin \beta \end{bmatrix}$$

をもつ行列によって乗算するよう実装されてもよい。ここで、 $\beta = \arcsin(Coh)/2$ である。

【0 1 4 2】

あるいはまた、バイノーラル出力チャネルにおける上記の音像バイアスは、交互の周波数帯域についてその入力のチャネル順が切り換えられるならば（たとえば、奇数周波数帯域では要素 3 1 0 の出力がマトリクス 3 1 2 の第一の入力に呈されてもよく、要素 3 1 1 の出力がマトリクス 3 1 2 の第二の入力に呈されてもよく、偶数周波数帯域では要素 3 1 1 の出力がマトリクス 3 1 2 の第一の入力に呈されてもよく、要素 3 1 0 の出力がマトリクス 3 1 2 の第二の入力に呈されてもよい）、すべての周波数帯域についてのFDNにおいて同一であるようマトリクス 3 1 2 を実装することによって緩和できる。

【0 1 4 3】

図 9 の実施形態（および本発明のシステムのFDNの他の時間領域実施形態）では、要素 4 2 2 から出力される未混合バイノーラル・チャネル出力が常に要素 4 2 3 から出力される未混合バイノーラル・チャネル出力より進んでいる（遅れている）ときに普通なら帰結するであろう音像バイアスに対処するために周波数に基づいてパニングを交互させることはトリビアルではない。この音像バイアスは、本発明のシステムのFDNの典型的な時間領域実施形態では、本発明のシステムのFDNのフィルタバンク領域実施形態において典型的に対処されるのとは異なる仕方で対処される。特に、図 9 の実施形態（および本発明のシステムのFDNの他の時間領域実施形態）において、未混合バイノーラル・チャネル（たとえば図 9 の要素 4 2 2 および 4 2 3 からの出力）の相対利得は、利得要素（たとえば図 9 の要素 4 1 7、4 1 8、4 1 9 および 4 2 0）によって、上記の均衡しないタイミングのために普通なら帰結するであろう音像バイアスを補償するよう決定される。ある利得要素（たとえば要素 4 1 7）を最も早期に到達する信号（これはたとえば要素 4 2 2 によって一方の側にパンされている）を減衰させるよう実装し、ある利得要素（たとえば要素 4 1 8）をその次に早期の信号（これはたとえば要素 4 2 3 によって他方の側にパンされている）をブーストするよう実装することにより、ステレオ像がセンタリングし直される。こうして、利得要素 4 1 7 を含む残響タンクは要素 4 1 7 の出力に第一の利得を適用し、利得要素 4 1 8 を含む残響タンクは要素 4 1 8 の出力に（第一の利得とは異なる）第二の利得を適用する。それにより、第一の利得および第二の利得は（要素 4 2 2 から出力される）第一の未混合バイノーラル・チャネルを、（要素 4 2 3 から出力される）第二の未混合バイノーラル・チャネルに対して減衰させる。

【0 1 4 4】

より具体的には、図 9 のFDNの典型的な実装では、四つの遅延線 4 1 0、4 1 1、4 1

10

20

30

40

50



2 および 4 1 3 は順次大きくなる長さをもち、それぞれ順次大きくなる遅延値  $n_1$ 、 $n_2$ 、 $n_3$  および  $n_4$  をもつ。この実装では、フィルタ 4 1 7 は  $g_1$  の利得を適用する。こうして、フィルタ 4 1 7 の出力は、 $g_1$  の利得が適用された、遅延線 4 1 0 への入力に遅延されたバージョンである。同様に、フィルタ 4 1 8 は  $g_2$  の利得を適用し、フィルタ 4 1 9 は  $g_3$  の利得を適用し、フィルタ 4 2 0 は  $g_4$  の利得を適用する。こうして、フィルタ 4 1 8 の出力は、 $g_2$  の利得が適用された、遅延線 4 1 1 への入力に遅延されたバージョンであり、フィルタ 4 1 9 の出力は、 $g_3$  の利得が適用された、遅延線 4 1 2 への入力に遅延されたバージョンであり、フィルタ 4 2 0 の出力は、 $g_4$  の利得が適用された、遅延線 4 1 3 への入力に遅延されたバージョンである。

【0 1 4 5】

10

この実装では、次の利得値の選択： $g_1 = 0.5$ 、 $g_2 = 0.5$ 、 $g_3 = 0.5$ 、 $g_4 = 0.5$  は、（要素 4 2 4 から出力されるバイノーラル・チャンネルによって示される）出力音像の一方の側への（すなわち、左または右チャンネルへの）望ましくないバイアスにつながることもありうる。本発明のある実施形態によれば、（それぞれ要素 4 1 7、4 1 8、4 1 9 および 4 2 0 によって適用される）値  $g_1$ 、 $g_2$ 、 $g_3$ 、 $g_4$  は、音像をセンタリングするために次のように選ばれる： $g_1 = 0.38$ 、 $g_2 = 0.6$ 、 $g_3 = 0.5$ 、 $g_4 = 0.5$ 。こうして、出力ステレオ像は、本発明のある実施形態によれば、最も早期に到達する信号（これは今の例では要素 4 2 2 によって一方の側にパンされている）を二番目に遅く到達する信号に対して減衰させ（すなわち、 $g_1 < g_3$  のように選ぶ）、二番目に早期の信号（これは今の例では要素 4 2 3 によって他方の側にパンされている）を最も遅く到達する信号に対してブーストする（すなわち、 $g_4 < g_2$  のように選ぶ）ことにより、センタリングし直される。

20

【0 1 4 6】

図 9 の時間領域 FDN の典型的な実装は、図 4 のフィルタバンク領域（CQMF 領域）FDN に対して、以下の相違点および類似点をもつ。

【0 1 4 7】

同じユニタリー・フィードバック・マトリクス A（図 4 のマトリクス 3 0 8 および図 9 のマトリクス 4 1 5）。

【0 1 4 8】

類似の残響タンク遅延  $n_i$ （すなわち、図 4 の CQMF 実装における遅延は、 $1/T_s$  がサンプリング・レートであるとして（ $1/T_s$  は典型的には 48KHz に等しい）、 $n_1 = 17 \cdot 64T_s = 1088 \cdot T_s$ 、 $n_2 = 21 \cdot 64T_s = 1344 \cdot T_s$ 、 $n_3 = 26 \cdot 64T_s = 1664 \cdot T_s$ 、 $n_4 = 29 \cdot 64T_s = 1856 \cdot T_s$  であってもよく、一方、時間領域実装における遅延は  $n_1 = 1089 \cdot T_s$ 、 $n_2 = 1345 \cdot T_s$ 、 $n_3 = 1663 \cdot T_s$ 、 $n_4 = 185 \cdot T_s$  であってもよい。典型的な CQMF 実装では、各遅延が 64 サンプルのブロックの継続時間の何らかの整数倍であるという実際上の制約条件があるが、時間領域では、各遅延の選択に関してより柔軟性があり、よって各残響タンクの遅延の選択に対してより柔軟性があることを注意しておく）。

30

【0 1 4 9】

類似の全域通過フィルタ実装（すなわち、図 4 のフィルタ 3 0 1 および図 9 のフィルタ 4 0 1 の同様の実装）。たとえば、全域通過フィルタは、いくつかの（たとえば三つの）全域通過フィルタの縦続〔カスケード〕によって実装されることができる。たとえば、それぞれの縦続された全域通過フィルタは、 $g = 0.6$  であるとして、

40

【数 1 3】

$$\frac{g - Z^{-n_i}}{1 - g \cdot Z^{-n_i}}$$

の形であってもよい。図 4 の全域通過フィルタ 3 0 1 は、サンプル・ブロックの好適な遅延（たとえば、 $n_1 = 64 \cdot T_s$ 、 $n_2 = 128 \cdot T_s$  および  $n_3 = 196 \cdot T_s$ ）をもつ三つの縦続された全域通過フィルタによって実装されてもよく、一方、図 9 の全域通過フィルタ 4 0 1（時間領

50

域の全域通過フィルタ)は、同様な遅延(たとえば、 $n_1 = 61 \cdot T_s$ 、 $n_2 = 127 \cdot T_s$ および $n_3 = 191 \cdot T_s$ )をもつ三つの縦続された全域通過フィルタによって実装されてもよい。

#### 【0150】

図9の時間領域FDNのいくつかの実装では、入力フィルタ400は、図9のシステムによって適用されるBRIRの直接対後期比(DLR)を目標DLRに(少なくとも実質的に)マッチさせるとともに、図9のシステムを含む仮想化器(たとえば図10の仮想化器)によって適用されるBRIRのDLRがフィルタ400を置換する(またはフィルタ400の構成設定を制御する)ことによって変更できるように、実装される。たとえば、いくつかの実施形態では、フィルタ400は、目標DLRを実装し、任意的には所望されるDLR制御を実装するフィルタの縦続(たとえば、図9Aに示されるように結合された、第一のフィルタ400Aおよび第二のフィルタ400B)として実装される。たとえば、該縦続のフィルタはIIRフィルタである(たとえば、フィルタ400Aは、目標低周波数特性にマッチするよう構成された一次バターワース高域通過フィルタ(IIRフィルタ)であり、フィルタ400Bは、目標高周波数特性にマッチするよう構成された二次の低シェルフIIRフィルタ)。もう一つの例として、この縦続のフィルタは、IIRおよびFIRフィルタである(たとえば、フィルタ400Aは、目標低周波数特性にマッチするよう構成された二次バターワース高域通過フィルタ(IIRフィルタ)であり、フィルタ400Bは、目標高周波数特性にマッチするよう構成された14次のFIRフィルタ)。典型的には、直接信号は固定されており、フィルタ400は後期信号を目標DLRを達成するよう修正する。全域通過フィルタ(APF)401は好ましくは、図4のAPF301と同じ機能を実行するよう、つまり位相多様性および増大したエコー密度を導入してより自然に聞こえるFDN出力を生成するよう実装される。入力フィルタ400は振幅応答を制御する一方、APF401は典型的には位相応答を制御する。

#### 【0151】

図9では、フィルタ406および利得要素406Aは一緒になって残響フィルタを実装し、フィルタ407および利得要素407Aは一緒になって別の残響フィルタを実装し、フィルタ408および利得要素408Aは一緒になって別の残響フィルタを実装し、フィルタ409および利得要素409Aは一緒になって別の残響フィルタを実装する。図9のフィルタ406、407、408および409のそれぞれは、好ましくは、1に近い最大利得値(単位利得)をもつフィルタとして実装され、利得要素406A、407A、408Aおよび409Aのそれぞれは、(関連する残響タンク遅延 $n_i$ 後に)所望される減衰にマッチする、フィルタ406、407、408および409の対応するものの出力への減衰利得を適用するよう構成される。具体的には、利得要素406Aは、要素406Aの出力に、(残響タンク遅延 $n_1$ 後の)遅延線410の出力が第一の目標の減衰した利得をもつような利得をもたせるよう、フィルタ406の出力に減衰利得(decaygain<sub>1</sub>)を適用するよう構成され、利得要素407Aは、要素407Aの出力に、(残響タンク遅延 $n_2$ 後の)遅延線411の出力が第二の目標の減衰した利得をもつような利得をもたせるよう、フィルタ407の出力に減衰利得(decaygain<sub>2</sub>)を適用するよう構成され、利得要素408Aは、要素408Aの出力に、(残響タンク遅延 $n_3$ 後の)遅延線412の出力が第三の目標の減衰した利得をもつような利得をもたせるよう、フィルタ408の出力に減衰利得(decaygain<sub>3</sub>)を適用するよう構成され、利得要素409Aは、要素409Aの出力に、(残響タンク遅延 $n_4$ 後の)遅延線413の出力が第四の目標の減衰した利得をもつような利得をもたせるよう、フィルタ409の出力に減衰利得(decaygain<sub>4</sub>)を適用するよう構成される。

#### 【0152】

図9のシステムのフィルタ406、407、408および409のそれぞれおよび要素406A、407A、408Aおよび409Aのそれぞれは、好ましくは、図9のシステムを含む仮想化器(たとえば図10の仮想化器)によって適用されるBRIRの目標T60特性を達成するよう実装される(フィルタ406、407、408および409のそれぞれは好ましくはIIRフィルタ、たとえばシェルフ・フィルタまたはシェルフ・フィルタの縦続

として実装される)。ここで、T60は、残響減衰時間( $T_{60}$ )を表わす。たとえば、いくつかの実施形態では、フィルタ406、407、408および409のそれぞれは、シェルフ・フィルタ(たとえば、図13に示されるT60特性を達成するよう $Q=0.3$ およびシェルフ周波数500Hzをもつシェルフ・フィルタ;図13でT60は秒の単位をもつ)として、あるいは二つのIIRシェルフ・フィルタ(たとえば、図14に示されるT60特性を達成するようシェルフ周波数100Hzおよび1000Hzをもつもの;図14でT60は秒の単位をもつ)の縦続として、実装される。各シェルフ・フィルタの形状は、低周波数から高周波数への所望される変化曲線にマッチするよう決定される。フィルタ406がシェルフ・フィルタ(または複数のシェルフ・フィルタの縦続)として実装されるとき、フィルタ406および利得要素406Aを有する残響フィルタも、シェルフ・フィルタ(またはシェルフ・フィルタの縦続)である。同様に、フィルタ407、408および409のそれぞれがシェルフ・フィルタ(またはシェルフ・フィルタの縦続)として実装されるとき、フィルタ407(または408または409)および対応する利得要素(407A、408Aまたは409A)を有する各残響フィルタも、シェルフ・フィルタ(またはシェルフ・フィルタの縦続)である。

#### 【0153】

図9Bは、図9Bに示されるように結合された第一のシェルフ・フィルタ406Bおよび第二のシェルフ・フィルタ406Cの縦続として実装されたフィルタ406の例である。フィルタ407、408、409のそれぞれは、フィルタ406の図9Bの実装と同様に実装されてもよい。

#### 【0154】

いくつかの実施形態では、要素406A、407A、408A、409Aによって適用される減衰利得( $\text{decaygain}_i$ )は次のように決定される。

#### 【0155】

#### 【数14】

$$\text{decaygain}_i = 10^{((-60 \cdot (n_i / F_s) / T) / 20)}$$

ここで、 $i$ は残響タンク・インデックスであり(すなわち、要素406Aは $\text{decaygain}_1$ を適用し、要素407Aは $\text{decaygain}_2$ を適用し、などとなる)、 $n_i$ は $i$ 番目の残響タンクの遅延である(たとえば、 $n_1$ は遅延線410によって適用される遅延)。 $F_s$ はサンプリング・レートであり、 $T$ は、あるあらかじめ決められた低い周波数における所望される残響遅延時間( $T_{60}$ )である。

#### 【0156】

図11は、図9の以下の要素:要素422および423ならびにIACC(両耳間相互相関係数)フィルタリングおよび混合段424、の実施形態である。要素422は、(図9の)フィルタ417および419の出力を合計し、合計された信号を低シェルフ・フィルタ500の入力に呈するよう結合され、構成されており、要素422は、(図9の)フィルタ418および420の出力を合計し、合計された信号を高域通過フィルタ501の入力に呈するよう結合され、構成されている。フィルタ500および501の出力は要素502において加算(混合)され、バイノーラル左耳出力信号を生成し、フィルタ500および501の出力は要素502において混合され(フィルタ500の出力がフィルタ501の出力から要素502において減算される)、バイノーラル右耳出力信号を生成する。要素502および503は、フィルタ500および501のフィルタリングされた出力を混合(加算および減算)して、(受け入れ可能な精度の範囲内で)目標IACC特性を達成するバイノーラル出力信号を生成する。図11の実施形態では、低シェルフ・フィルタ500および高域通過フィルタ501のそれぞれは、典型的には一次IIRフィルタとして実装される。フィルタ500および501がそのような実装をもつ一例では、図11の実施形態は、図12において曲線「I」としてプロットされている例示的なIACC特性を達成しうる

。これは、図 1 2 において「 $I_T$ 」としてプロットされている目標 IACC 特性に対する良好なマッチである。

【 0 1 5 7 】

図 1 1 の A は、図 1 1 のフィルタ 5 0 0 の典型的な実装の周波数応答 (R1)、図 1 1 のフィルタ 5 0 1 の典型的な実装の周波数応答 (R2) および並列に接続したフィルタ 5 0 0 および 5 0 1 の応答のグラフである。図 1 1 の A から、組み合わせられた応答が 100Hz ~ 10,000Hz の範囲を横断して望ましいように平坦であることが明白である。

【 0 1 5 8 】

このように、あるクラスの実施形態では、本発明は、マルチチャネル・オーディオ入力信号のチャネルのある集合にตอบสนองしてバイノーラル信号 (たとえば、図 1 0 の要素 2 1 0 の出力) を生成するシステム (たとえば図 1 0 のシステム) および方法である。これは、前記集合の各チャネルにバイノーラル室内インパルス応答 (BRIR) を適用し、それによりフィルタリングされた信号を生成する段階であって、前記集合のチャネルのダウンミックスに共通の後期残響を加えるよう単一のフィードバック遅延ネットワーク (FDN) を使うことによることを含む、段階と; フィルタリングされた信号を組み合わせる前記バイノーラル信号を生成する段階とを実行することによることを含む。FDN は時間領域で実装される。そのようないくつかの実施形態では、時間領域 FDN (たとえば、図 9 のように構成された、図 1 0 の FDN 2 2 0) は:

前記ダウンミックスを受領するよう結合された入力をもつ入力フィルタ (たとえば図 9 のフィルタ 4 0 0) であって、該入力フィルタは前記ダウンミックスにตอบสนองして第一のフィルタリングされたダウンミックスを生成するよう構成されている、入力フィルタと;

前記第一のフィルタリングされたダウンミックスにตอบสนองして第二のフィルタリングされたダウンミックスをするよう結合され、構成された全域通過フィルタ (たとえば図 9 の全域通過フィルタ 4 0 1) と;

第一の出力 (たとえば要素 4 2 2 の出力) および第二の出力 (たとえば要素 4 2 3 の出力) をもつ残響適用サブシステム (たとえば図 9 の、要素 4 0 0、4 0 1 および 4 2 4 以外のすべての要素) であって、前記残響適用サブシステムは残響タンクの集合を含み、各残響タンクは異なる遅延をもち、該残響適用サブシステムは、前記第二のフィルタリングされたダウンミックスにตอบสนองして第一の未混合バイノーラル・チャネルおよび第二の未混合バイノーラル・チャネルを生成し、前記第一の未混合バイノーラル・チャネルを前記第一の出力において呈し、前記第二の未混合バイノーラル・チャネルを前記第二の出力において呈するよう結合され、構成されている、残響適用サブシステムと;

前記残響適用サブシステムに結合され、前記第一の未混合バイノーラル・チャネルおよび第二の未混合バイノーラル・チャネルにตอบสนองして第一の混合済みバイノーラル・チャネルおよび第二の混合済みバイノーラル・チャネルを生成するよう構成されている、両耳間相互相関係数 (IACC: interaural cross-correlation coefficient) フィルタリングおよび混合段 (たとえば、図 1 1 の要素 5 0 0、5 0 1、5 0 2、5 0 3 として実装されてもよい図 9 の段 4 2 4) とを含む。

【 0 1 5 9 】

入力フィルタは、各 BRIR が少なくとも実質的に目標 DLR にマッチする直接対後期比 (DLR) をもつよう前記第一のフィルタリングされたダウンミックスを生成するよう (好ましくは、それを生成するよう構成された二つのフィルタのカスケード [縦続] として) 実装されてもよい。

【 0 1 6 0 】

各残響タンクは、遅延された信号を生成するよう構成されていてもよく、前記各残響タンクにおいて伝搬する信号に利得を加えて、遅延された信号が少なくとも実質的に前記遅延された信号についての目標の遅延された利得にマッチする利得をもつようにするよう結合され、構成された残響フィルタ (たとえば、シェルフ・フィルタまたはシェルフ・フィルタのカスケードとして実装される) を含んでいてもよい。各 BRIR の目標残響減衰時間特性 (たとえば  $T_{60}$  特性) を達成するためである。

## 【 0 1 6 1 】

いくつかの実施形態では、前記第一の未混合バイノーラル・チャンネルは前記第二の未混合バイノーラル・チャンネルより進んでおり、前記残響タンクは、最も短い遅延をもつ第一の遅延された信号を生成するよう構成された第一の残響タンク（たとえば、遅延線 4 1 0 を含む図 9 の残響タンク）と、二番目に短い遅延をもつ第二の遅延された信号を生成するよう構成された第二の残響タンク（たとえば、遅延線 4 1 1 を含む図 9 の残響タンク）とを含む。前記第一の残響タンクは前記第一の遅延された信号に第一の利得を適用するよう構成され、前記第二の残響タンクは前記第二の遅延された信号に第二の利得を適用するよう構成され、前記第二の利得は前記第一の利得とは異なり、前記第二の利得は前記第一の利得とは異なり、前記第一の利得および前記第二の利得の適用により、前記第二の未混合バイノーラル・チャンネルに対する前記第一の未混合バイノーラル・チャンネルの減衰が帰結する。典型的には、前記第一の混合済みバイノーラル・チャンネルおよび前記第二の混合済みバイノーラル・チャンネルは、再センタリングされた（re-centered）ステレオ像を示す。いくつかの実施形態では、前記 IACC フィルタリングおよび混合段は、前記第一の混合済みバイノーラル・チャンネルおよび前記第二の混合済みバイノーラル・チャンネルが少なくとも実質的に目標 IACC 特性に一致する IACC 特性をもつよう前記第一の混合済みバイノーラル・チャンネルおよび前記第二の混合済みバイノーラル・チャンネルを生成するよう構成されている。

10

## 【 0 1 6 2 】

本発明の諸側面は、オーディオ信号（たとえば、オーディオ・コンテンツがスピーカー・チャンネルからなるオーディオ信号および/またはオブジェクト・ベースのオーディオ信号）のバイノーラル仮想化を実行する（または実行するよう構成されているまたはその実行をサポートする）方法およびシステム（たとえば、図 2 のシステム 2 0 または図 3 または図 1 0 のシステム）を含む。

20

## 【 0 1 6 3 】

いくつかの実施形態では、本発明の仮想化器は、マルチチャンネル・オーディオ入力信号を示す入力データを受領するまたは生成するよう結合され、該入力データに対して、本発明の方法の実施形態を含む多様な処理の任意のものを実行するようソフトウェア（またはファームウェア）をもってプログラムされたまたは（たとえば制御データにตอบสนองして）他の仕方で構成された汎用プロセッサであるまたはそれを含む。そのような汎用プロセッサは典型的には入力装置（たとえばマウスおよび/またはキーボード）、メモリおよび表示装置に結合される。たとえば、図 3 のシステム（または図 2 のシステム 2 0 またはシステム 2 0 の要素 1 2、...、1 4、1 5 を有する仮想化器システム）は、汎用プロセッサにおいて実装されることができ、入力の前記オーディオ入力信号の N 個のチャンネルを示すオーディオ・データであり、出力はバイノーラル・オーディオ信号の二つのチャンネルを示すオーディオ・データである。通常のデジタル アナログ変換器（DAC: digital-to-analog converter）が前記出力データに対して作用して、スピーカー（たとえばヘッドフォン対）による再生のための、バイノーラル信号チャンネルのアナログ・バージョンを生成することができる。

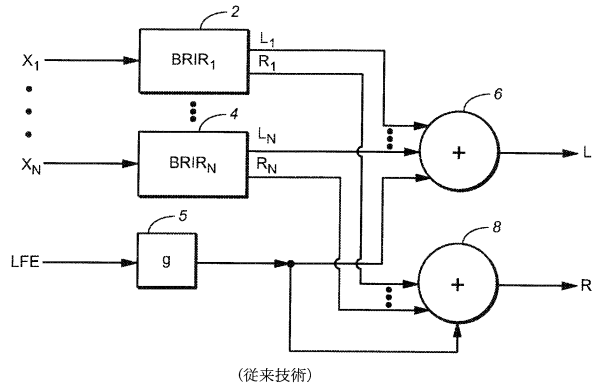
30

## 【 0 1 6 4 】

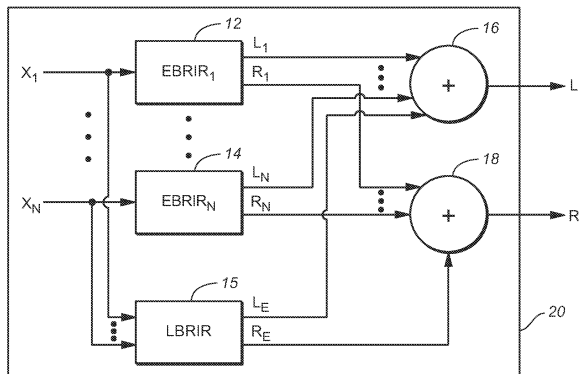
本発明の個別的な実施形態および本発明の応用が本稿に記載されているが、本願において記載され特許請求される発明の範囲から外れることなく、本稿に記載されるこれらの実施形態および応用に対する多くの変形が可能であることは、当業者には明白であろう。本発明のある種の形が示され、記述されているが、本発明は記載され、示されている特定の実施形態や記載される特定の方法に限定されないことは理解されるべきである。

40

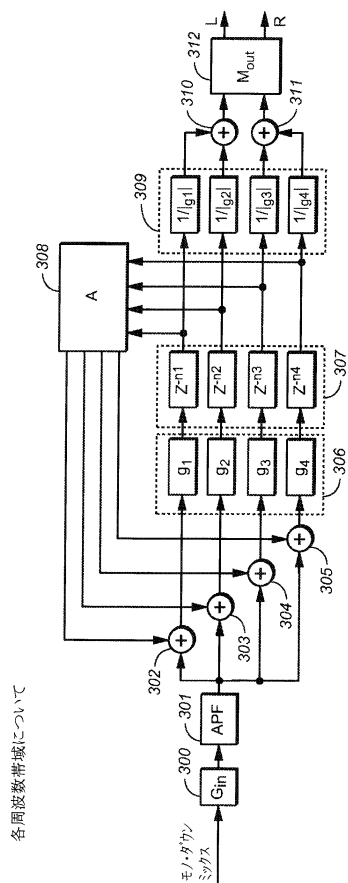
【図 1】



【図 2】

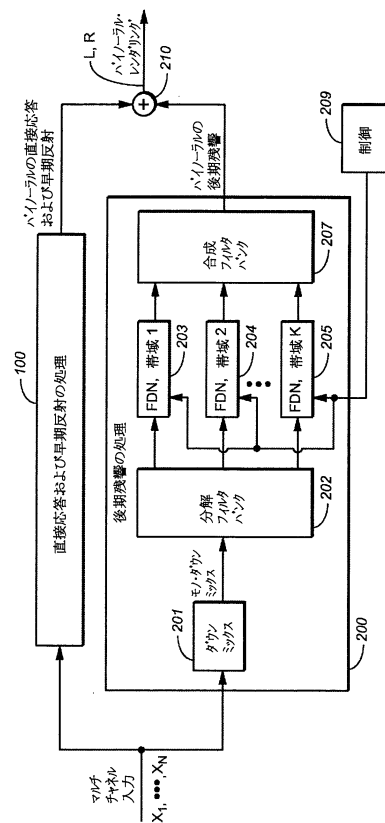


【図 4】

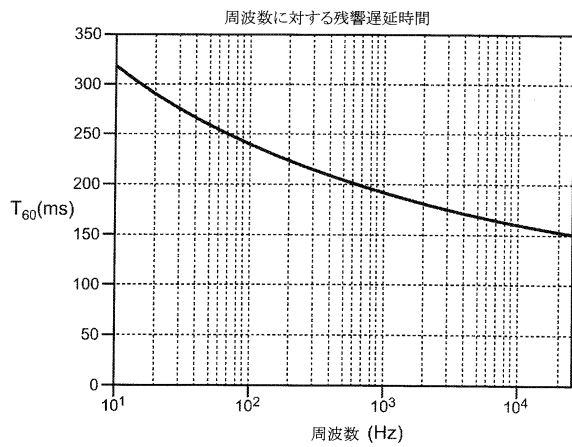


各周波数帯域について

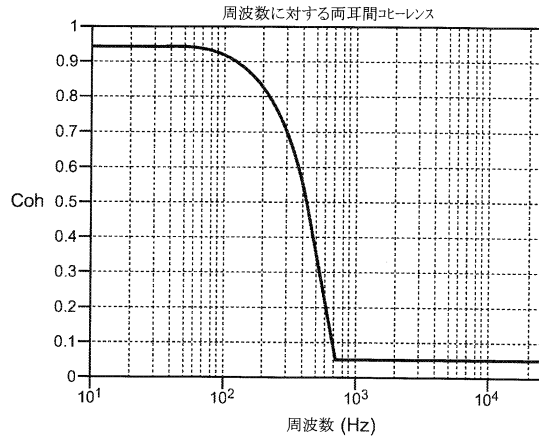
【図 3】



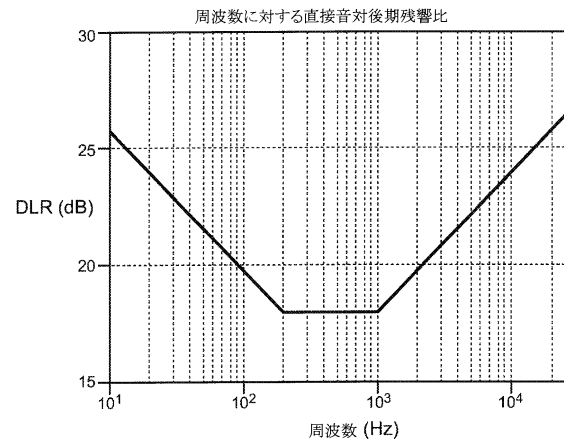
【図 5】



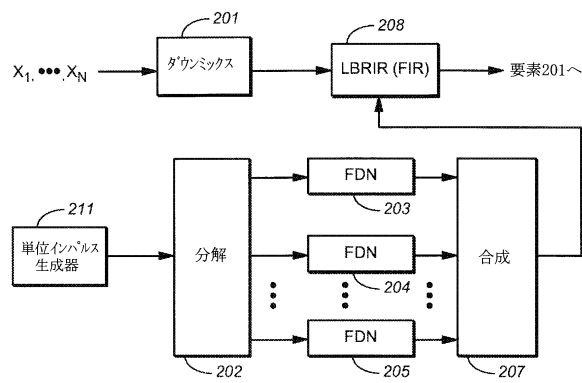
【図 6】



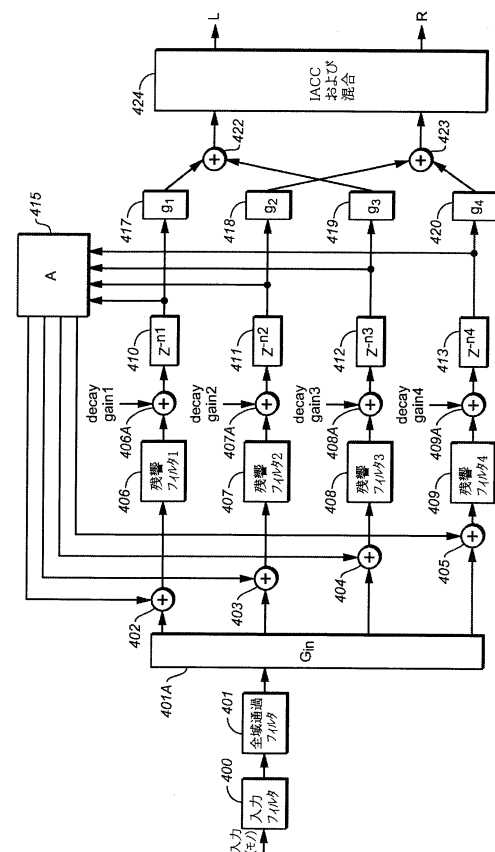
【図 7】



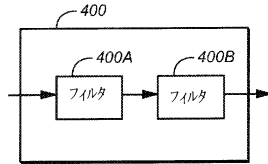
【図 8】



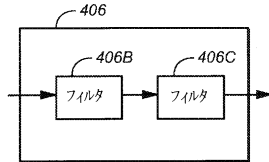
【図 9】



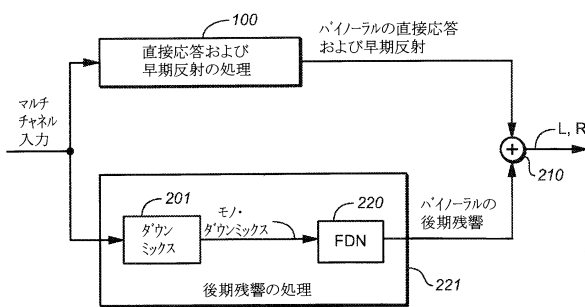
【図 9 A】



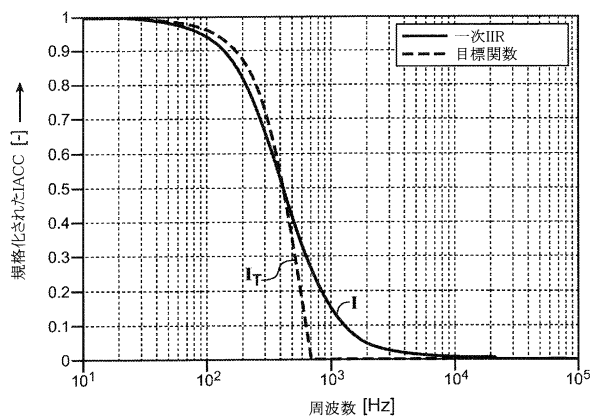
【図 9 B】



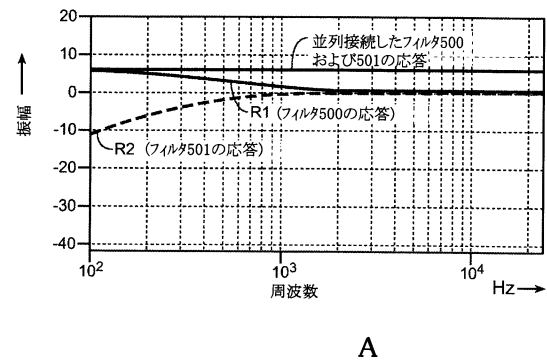
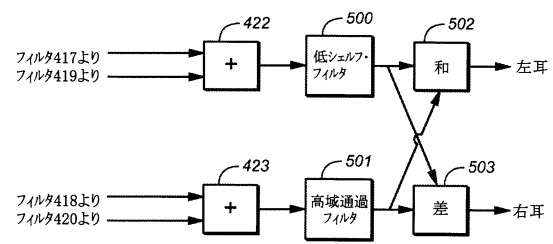
【図 10】



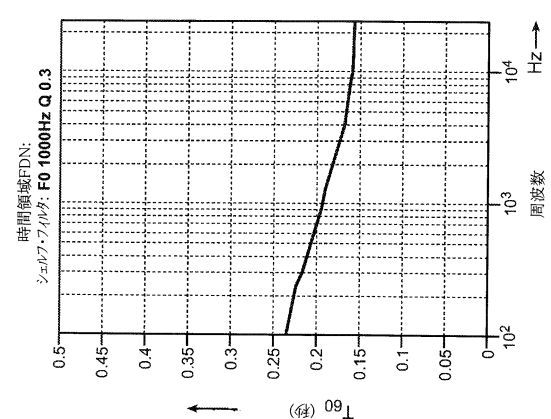
【図 12】



【図 11】

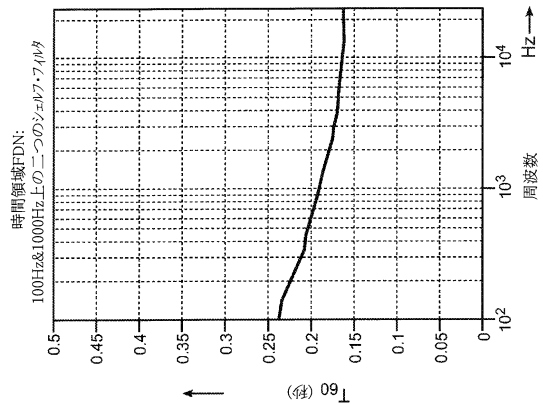


【図 13】





【図 14】



## フロントページの続き

(31)優先権主張番号 61/988,617

(32)優先日 平成26年5月5日(2014.5.5)

(33)優先権主張国・地域又は機関  
米国(US)

(72)発明者 イェン, クアン チック

アメリカ合衆国, カリフォルニア州 94404, フォスター・シティ, キャラベル・レーン 735

(72)発明者 ブリーバルト, ディルク ジェイ.

オーストラリア国, ニュー・サウス・ウェールズ 2060, シドニー, マクマホンズ・ポイント, ミッチェル・ストリート 35 レベル 3, ドルビー・オーストラリア・プロプライエタリー・リミテッド内

(72)発明者 デヴィッドソン, グラント エイ.

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

(72)発明者 ウィルソン, ロンダ

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 クーパー, デイヴィッド エム.

オーストラリア国, ニュー・サウス・ウェールズ 2060, シドニー, マクマホンズ・ポイント, ミッチェル・ストリート 35 レベル 3, ドルビー・オーストラリア・プロプライエタリー・リミテッド内

(72)発明者 シュアン, ズーウェイ

中華人民共和国, 베이징 100020, シャオヤン・ディストリクト, イースト・サード・リング・ミドル・ロード, ワールド・フィナンシャル・センター ナンバー 1, ドルビー ラボラトリーズ インターナショナル サービスズ内

審査官 大野 弘

(56)参考文献 特開2007-336080(JP, A)

特表2013-508760(JP, A)

(58)調査した分野(Int.Cl., DB名)

H04S 7/00

H04S 3/00