

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4800792号
(P4800792)

(45) 発行日 平成23年10月26日(2011.10.26)

(24) 登録日 平成23年8月12日(2011.8.12)

(51) Int.Cl.	F 1
G 06 F 13/00	(2006.01) G 06 F 13/00 351 A
H 04 L 12/18	(2006.01) H 04 L 12/18
H 04 R 3/00	(2006.01) H 04 R 3/00 320
H 04 M 3/00	(2006.01) H 04 M 3/00 B

請求項の数 10 (全 22 頁)

(21) 出願番号	特願2006-47524 (P2006-47524)
(22) 出願日	平成18年2月23日 (2006.2.23)
(65) 公開番号	特開2006-268844 (P2006-268844A)
(43) 公開日	平成18年10月5日 (2006.10.5)
審査請求日	平成21年2月19日 (2009.2.19)
(31) 優先権主張番号	11/066,137
(32) 優先日	平成17年2月23日 (2005.2.23)
(33) 優先権主張国	米国(US)

(73) 特許権者	500046438 マイクロソフト コーポレーション アメリカ合衆国 ワシントン州 9805 2-6399 レッドモンド ワン マイ クロソフト ウェイ
(74) 代理人	100077481 弁理士 谷 義一
(74) 代理人	100088915 弁理士 阿部 和夫
(72) 発明者	チン リー アメリカ合衆国 98052 ワシントン 州 レッドモンド ワン マイクロソフト ウェイ マイクロソフト コーポレーション内

最終頁に続く

(54) 【発明の名称】サーバレスピアツーピアマルチパーティリアルタイム音声通信システムおよび方法

(57) 【特許請求の範囲】

【請求項 1】

ピアツーピアコンピュータネットワーク内の複数のピアノードを使用する、マルチパーティリアルタイム通信セッションの方法であって、

入力ストリームを複数のフレームに区分するステップと、

前記複数のピアノードのうちの1つを選択するステップと、

選択されたピアノードがそのピアノードの利用可能な資源に比例する数のフレームを受け取りミキシング及び再送信の処理を行うステップと、

前記ミキシング及び再配信を、予め定めたスケジュールに従ってフレームバイフレーム方式で前記複数のピアノードの少なくとも一部の間で交代させて、前記マルチパーティリアルタイム通信セッションを提供するステップと

を含むことを特徴とする方法。

【請求項 2】

前記入力ストリームは、(a) 音声コンテンツを含む到着音声ストリーム、(b) 音声コンテンツを含む到着マイクロフォン信号、の少なくとも一方であり、

前記音声コンテンツをミキシングおよび再配信することをさらに含むことを特徴とする請求項1に記載の方法。

【請求項 3】

前記ミキシングは、

(a) 前記到着音声ストリームをエントロピーデコードおよび逆量子化してエンコード

10

20

音声係数パケットを生成すること、(b)前記到着マイクロフォン信号を音声係数ブロックに変換し、前記音声係数ブロックをエンコードしてエンコード音声係数パケットを生成すること、のいずれかを実行して、エンコード音声係数パケットを生成することと、

前記音声係数パケットとブロックを結合して、あるフレームに対する前記係数パケットとブロックの組合せである、合成係数ブロックを生成することと
をさらに含むことを特徴とする請求項2に記載の方法。

【請求項4】

前記合成係数ブロックを量子化することと、
量子化された前記合成係数ブロックをリエンコードしてミックス音声パケットを生成することと
をさらに含むことを特徴とする請求項3に記載の方法。
10

【請求項5】

前記ミックス音声パケットをデコードすることと、
デコードされた前記ミックス音声パケットを逆量子化してミックス変換係数を得ることと、
前記ミックス変換係数に逆変換を適用してミックス音声コンテンツを含む音声ストリームを生成することと
をさらに含むことを特徴とする請求項4に記載の方法。

【請求項6】

前記音声ストリームおよびマイクロフォン信号が送られるあるピアノードを除き、前記ピアノードのそれからの音声コンテンツを、ミキシングおよび再配信することをさらに含むことを特徴とする請求項2に記載の方法。
20

【請求項7】

(a)音声コンテンツを含む前記到着音声ストリーム、(b)音声コンテンツを含む前記到着マイクロフォン信号、の少なくとも一方をミキシングおよび再配信することと、
エコーを低減するために、各ピアノードに、ミックス音声コンテンツからそれ自体の音声コンテンツを差し引かせることと
をさらに含むことを特徴とする請求項2に記載の方法。

【請求項8】

完全接続メッシュを形成するピアノード群としてクリークを定義することと、
前記クリークが前記マルチパーティリアルタイム通信セッションにおいてスーパーゲートウェイノードとしての役割を果たすように、前記クリークを配置することと
をさらに含むことを特徴とする請求項1の記載の方法。
30

【請求項9】

完全接続メッシュを形成するピアノード群としてクリークを定義することと、
前記クリークが、前記マルチパーティリアルタイム通信セッションにおいてクライアントノードとして作用する追加ノードを有するスタートポロジにおいて、スーパーサーバとしての役割を果たすように、前記クリークを配置し、それによって前記クライアントノードがミキシングタスクを免除されるようにすることと
をさらに含むことを特徴とする請求項1の記載の方法。
40

【請求項10】

請求項1に記載のコンピュータにより実施される方法を実行するためのコンピュータ実行可能命令を有することを特徴とするコンピュータ可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般的には、コンピュータネットワーキングに関し、より具体的には、サーバレスピアツーピア(P2P)マルチパーティリアルタイム音声通信システムおよび方法であって、ピアのそれぞれが、交代して音声をミキシングし、圧縮した音声を再配信(redelivery)するとともに、あるピアによってミキシングおよび再配信され
50

る音声フレーム数が、そのピアの利用可能な資源（アップロード帯域幅または計算能力など）に比例する、システムおよび方法に関する。

【背景技術】

【0002】

マルチパーティ音声通信システムは、一群の人人がリアルタイム音声通信セッションに参加することを可能にする。さらに、このシステムは、多数の人が同時に話すことを可能にする。（音声キャプチャ、音響エコーチャンセレーション（AEC）、自動ゲインコントロール（AGC）、および音声／通話圧縮などの）ツーパーティ音声通信システムの音声構成要素に加えて、マルチパーティ音声通信システムには、音声ミキシングおよびネットワーク配信における固有の課題がある。

10

【0003】

一例として、 n 個のピアコンピュータ（またはピア）が、起こりうる複数の同時話者とのマルチパーティ音声通信セッションに参加していることを仮定する。さらに、音声の各ストリームは、 b_w の帯域幅を必要とすると仮定する。マルチパーティ音声通信システムは、様々なトポロジおよびミキシング方策を用いて形成することができる。普及しているトポロジの 1 つは、図 1 A に示すような、スタートポロジである。高性能の中央サーバ S は、すべてのピア (t_1, t_2, t_3, t_4 、および t_5) からの音声ストリームを受信し、この音声ストリームをミキシングし、そしてミキシングおよびリエンコード（re-encode）された音声をすべてのピアに返す。

20

【0004】

スタートポロジの利点は、各パーティが、ツーパーティ通信システムのハードウェアと同じハードウェアを使用し、したがって変更を必要としないことである。マルチパーティ通信セッションをサポートするためには、サーバだけを再設計すればよい。結果的に、スタートポロジは、商業用マルチパーティ通信のソリューションとして一般的に選択されている。そのようなシステムの 1 つが、非特許文献 1 に記載されている。スタートポロジの主たる欠点は、サーバ S 上に大きな演算と帯域の負担がかかることがある。サーバ S は、 n 個のストリームの圧縮音声 ($n \cdot b_w$ のダウンロード帯域幅) を受信し、それらをデコード、ミキシング、リエンコードして、ミキシングした音声を n 個のピア ($n \cdot b_w$ のアップロード帯域幅) に返す必要がある。

30

【0005】

第 2 の一般的なトポロジは、図 1 B に示すような、完全接続ユニキャスト（unicast）ネットワークである。完全接続ネットワークにおいては、すべてのピアがネットワーク内の他のすべてのピアに接続されている。この種のトポロジの一例について、非特許文献 2 で考察されている。このトポロジにおいては、ピア (t_1, t_2, t_3, t_4 、および t_5) は、いかなる音声ミキシングも再配信も実行しない。その代わりに、それぞれの話者は、単に、圧縮音声を他のすべてのピアに送信する。そのようなトポロジでは、それぞれのピアは、音声を残りのピアに送信するのに $(n - 1) \cdot b_w$ のアップロード帯域幅を必要とするとともに、入力音声を受信するのに、最大 $(n - 1) \cdot b_w$ のダウンロード帯域幅を必要とする。このトポロジの 1 つの欠点は、ネットワークトラフィックの大幅な増加であり、このことは、それぞれのピアおよびネットワーク全体に大きな負担を与える。

40

【0006】

第 3 の可能なトポロジは、汎用グラフ（generic graph）であり、これはエンドシステムミキシングを使用する。この種のトポロジの一例が、図 1 C に示されており、また非特許文献 3 に記載されている。図 1 C に示すように、この例において、ピア a、b、f および g は葉ノード（leaf node）であり、いかなるミキシング演算も実行しない。ピア c、d および e は、ゲートウェイノードとしての役割を果たし、このゲートウェイノードは、近傍のピアのために音声をミキシングして再配信する。一般に、 m 個の隣接ノードがあれば、音声を受信して再配信するのに、 $m \cdot b_w$ のアップロードおよびダウンロード帯域幅を必要とする。 m は通常は n よりもずっと小さいので、このトポロ

50

ジの設計は、大規模会議セッションにうまく対応することができる。しかしながら、このトポロジの欠点は、ゲートウェイノードに対する負担が大きくなる可能性があることである。別の欠点は、ゲートウェイの連鎖が長くなると、音声配信の待ち時間が長くなることである。さらに別の欠点は、配信の連鎖に沿って、音声の同期がずれる可能性もあることである。

【0007】

音声通信セッションにおけるトラフィックをさらに低減するネットワークレベルのソリューションの1つは、IPマルチキャストによるものである。IPマルチキャストにおいては、ソースから伝送される単一パケットが、そのソースをルートとする配送木(*distribution tree*)に沿った、ルータ群において複製される。このようにして、コンテンツが任意の数の受信者に配信される。例えば、図1Aに示すスタートポロジにおいて、ピアはユニキャストを経由してサーバに圧縮音声を送信することもできる。しかしながら、サーバSは、ミキシングおよびリエンコードした音声をn個のピアにマルチキャストして返すことができる。そのようなシステムのサンプル実現形態は非特許文献4に記載されている。サーバSのアップロード帯域幅はbwに縮小される。

10

【0008】

しかしながら、IPマルチキャストの1つの欠点は、サーバのダウンロード帯域幅に対する要求条件が、 $n \cdot bw$ で一定のままとなることである。図1Bに示す、完全接続ネットワークにおいて、それぞれの話者は、ネットワークの他のすべてのピアに圧縮音声をマルチキャストすることもできる。ここでも、完全接続ネットワークに対する、IPマルチキャストの欠点は、ピアのアップロード帯域幅がbwに縮小されるのに対して、ピアのダウンロード帯域幅は、 $(n - 1) \cdot bw$ で一定のままであることである。IPマルチキャストの別の欠点は、とりわけ、ドメイン間ルーティングプロトコル、ISPビジネスモデル(課金モデル)、配送木に沿った輻輳制御(*congestion control*)、およびセキュリティなどの問題が理由で、その展開が実世界においては遅いことである。その結果として、ある限られた大学/企業サブネットおよび(インターネット2などの)ネットワークテストベッドを除いて、固有IPマルチキャストのサポートは広く普及していない。ネットワークレベルのマルチキャストサービスの展開におけるこれらの問題のために、今日のインターネットにおける大部分のトラフィックは、ユニキャストベースであり、それによって2つのコンピュータが互いに直接話し合う。

20

【0009】

P2Pネットワークによるファイル転送のための1対多数コンテンツ配布のシステムおよび方法の1種が、2004年7月7日付出願、J. Li、P. ChouおよびC. Zhangの「Efficient One-to-Many Content Distribution in a Peer-to-Peer Computer Network」という名称の米国特許出願第10/887,406に記載されている。しかしながら、その研究では、1対多数ファイル転送および配布を含むが、音声通信セッションは、多数対多数配布を要する。さらに、その研究では、TCP/IPキューを広範囲に使用している。しかしながら、キューの使用は音声会議には実際的ではなく、それはパケットがタイムリーに到着しなくてはならないからである。さらに、異なるソースからの音声がミキシングされる可能性があり、このために、音声配信は、音声通信応用において独特のものとなっている。

30

【0010】

【非特許文献1】K. Singh, G. Nair, and H. Shulzrinne, "Centralized Conferencing Using SIP," Proceedings of the 2nd IP Technology Workshop, April 2001

【非特許文献2】J. Lennox and H. Shulzrinne, "A protocol for reliable decentralized conferencing," Proceedings of the 13th international workshop on network and operating systems support for digital audio and video, (NOSS-DAV'2003), pp. 72-81, 2003, Monterey, California

【非特許文献3】M. Radenkovic, C. Greenhalgh, and S. Benford, "Deployment issues for multi-user audio support in CVEs," Proceedings ACM Symposium on virtual rea

40

50

lity software and technology, pp. 179-185, 2002, Hong Kong, China

【非特許文献 4】"ConferenceXP:wireless classrooms, collaboration and distance learning" at <http://www.conferencexp.net/>

【非特許文献 5】Siren/G.722.1, "Coding at 24 and 32 kbits/s for hands-free operation in systems with low frame loss"

【非特許文献 6】"MP3: The definitive Guide," Scot Hacker press

【非特許文献 7】"Inside Windows (登録商標) Media," Microsoft Press

【発明の開示】

【発明が解決しようとする課題】

【0011】

10

既存のマルチパーティ音声通信システムの欠点の1つは、ピアまたはサーバが果たすミキシングおよび再配信の役割が、ネットワークトポロジによって固定されていることである。既存音声通信システムの別の欠点は、音声ストリーム全体のミキシングを実行することである。したがって、必要とされるのは、ネットワーク資源を最も効率よく使用する、音声通信システムおよび方法である。さらに、必要とされるのは、上記のネットワークトポロジの欠点を回避するとともに、ピアが果たすミキシングおよび再配信の役割においてフレキシブルであるシステムおよび方法である。さらに、必要とされるのは、音声ストリーム全体ではなく、音声ストリームのフレームに対してミキシングを実施する、音声通信システムおよび方法である。さらに、必要とされるのは、キューリングの使用を回避して、ファイル転送技法の遅延問題を克服する、音声通信システムおよび方法である。

20

【課題を解決するための手段】

【0012】

30

本明細書において開示する発明は、リアルタイムマルチパーティ音声通信セッションを最大効率で提供するピアツーピア(P2P)音声通信システムおよび方法を含む。P2Pネットワークは、各コンピュータが概して等価な能力と責任を有する種類のネットワークである。P2P音声通信システムおよび方法では、圧縮音声がパケットに区分され、そのパケットのそれぞれが、ミキシングと再配信のために単一ピアに送信される。あるピアによってミキシングおよび再配信されるパケット数は、そのピアの利用可能な資源に比例する。これらの資源には、アップロード帯域幅を含めてよい。あるいは、資源には、計算能力を含めることもできる。P2P音声通信システムおよび方法は、マルチパーティ音声通信セッションにおいて要求される帯域幅を低減する。さらに、P2P音声通信システムおよび方法では、音声サービス負担とピアアップロード帯域幅を均衡させて、P2Pネットワーク内のすべての参加ピア間でマルチパーティ通信セッションのコストを再分配する。これによって、P2P音声通信システムおよび方法では、高性能のサーバまたはピアを必要とすることなく、マルチパーティ音声通信セッションを実施することが可能となる。

【0013】

ミキシングおよび再配信がネットワークトポロジに基づいている従来技法とは異なり、本明細書に開示するP2P音声通信システムおよび方法では、圧縮音声をパケットまたはフレームに分割(split)または区分(divide)して、それぞれのピアに交代で音声パケットのミキシングおよび再配信を行わせる。P2P音声通信システムおよび方法では、ピアのネットワーク帯域幅負荷を柔軟に均衡させることによって、より多くの資源を有するピアが少ない資源を有するピアを支援できるようになる。さらにP2P音声通信システムおよび方法では、音声ミキシングにより、マルチパーティリアルタイム音声通信システムセッションを行うのに必要な帯域幅が減少する。

40

【0014】

本発明は、以下の説明と本発明の態様を示す添付の図面とを参照することによってより詳細に理解されるであろう。一例として本発明の原理を図示する添付の図面を合わせると、以下に示す本発明の詳細の説明から、その他の特徴および利点は明白になるであろう。

【0015】

図面を参照するが、これらの図面では、同じ参照番号は全体を通して対応する部分を表

50

わしている。

【発明を実施するための最良の形態】

【0016】

本発明の以下の説明において、添付の図面を参照するが、それらは本発明の一部を形成するものであり、それには、一例として、それによって本発明を実施することができる特定の例を示してある。ここで理解すべきことは、本発明の範囲から逸脱することなく、その他の実施形態も利用することが可能であり、また構造的変更も行うことができることである。

【0017】

I . 序文

10

ピアツーピア (P2P) コンピュータネットワークにおける現行の1対多数配布技法は、コンテンツをソースノードからピアノードに直接送るよりも、より効率的であるが、これらの技法は、ネットワークにおける最も効率的なコンテンツ配布を達成することができない。これにはいくつかの要因がある。1つの要因は、これらの技法には、ピアノード間の帯域幅の差を適切に説明してそれに適応するものが1つもないことである。別の要因は、これらの技法では、コンテンツを配布するときに、ネットワーク上のすべてのピアノードの帯域幅能力を完全に利用することができないことである。

【0018】

本明細書において説明する、P2P音声通信システムおよび方法は、マルチパーティ音声通信セッションを実行するための新規な解決策である。P2P音声通信システムおよび方法の重要な特徴は、音声ミキシングおよび音声の再配信の作業が、ミューチュアルキャストクリーク (Mutual Cast clique) 内のピア間で交代して行われることである。ミューチュアルキャストクリークは、完全接続メッシュを形成する、少数のピアノードを含む。音声ミキシングを変換ドメイン上でフレームバイフレーム方式で実行することができるという、波形コード化音声 (waveform coded audio) の固有の特性を使用して、このP2P音声通信システムおよび方法では、参加ピア間でミキシングおよび再配信のタスクを交代させる。これによって、ネットワーク帯域幅と計算負荷を共有することができる。このようにして、P2P音声通信システムおよび方法では、高性能のサーバなしに、マルチパーティ音声通信セッションを行うことができる。

20

【0019】

I I . 全体概要

30

図2は、本明細書において開示するP2P音声通信システムおよび方法の例証的実施形態を示すブロック図である。ここで留意すべきことは、図2は、P2P音声通信システムを実施して使用することのできる、いくつかの方法の内の1つにすぎないことである。

【0020】

図2を参照すると、完全接続ピアツーピア (P2P) ネットワーク200が、この例証的実施形態に示してある。P2P音声通信システムおよび方法を実行するP2Pネットワーク200は、ミューチュアルキャストクリークとも呼ばれる。図2に示すこの例証的実施形態において、ミューチュアルキャストクリーク200は、3つのピアノード、すなわちピアノード(1)、ピアノード(2)、およびピアノード(3)を含む。ピアノード(1)、(2)、(3)は、矢印で示すように完全接続されている。ピアノード(1)、(2)、(3)のそれぞれは、P2P音声通信システムおよび方法を含む。

40

【0021】

図3は、図2に示すようなピアノードのそれぞれに含まれるP2P音声通信システム300の一般的な例証的実現形態を示す詳細ブロック図である。図3を参照すると、一般に、P2P音声通信システム300は、各ピアからのネットワーク音声およびローカルピアからのマイクロフォン入力を受信し、その音声をミキシングして、ミックス音声をピアノードならびにローカルピアに出力する。そのローカルピアノードは、そのミックス音声をミキシングして再配信する。ピアノードの少なくとも一部は、ラウンドロビン (round robin) 方式で、ミキシングおよび再配信を交代で実行する。

50

【0022】

具体的には、図3はローカルピアノード(図示せず)上のP2P音声通信システム300を示す。このP2PネットワークにN個のピアノードがあると仮定する。システム300への入力には、ピア(1)310からピア(N)315までの、ネットワーク音声が含まれる。図3における省略符号は、ピアノードのすべては図示していないことを示している。さらに、マイクロフォンストリーム320も、システム300への入力である。マイクロフォンストリーム320は、ローカルピア上の1つまたは複数のマイクロフォンからのものである。P2P音声通信システム300は、入力音声をフレームまたはパケットに区分して、パケットのそれぞれをあるフレームにおいてミキシングする音声ミキサ330を含む。

10

【0023】

ピア(1)310からのエンコードネットワーク音声は、エントロピーデコード(*entropy decoding*)および逆量子化(1)340を実行することによって処理される。以下に詳細を説明するように、これによってピア(1)310からのエンコードネットワーク音声が部分的にデコードされて、MDCT変換係数のブロックが生成される。同様に、それぞれのピアからピア(N)315へのエンコードネットワーク音声は、エントロピーデコードおよび逆量子化(N)345を実行することによって処理される。ここでも、図3における省略記号は、エントロピーデコードおよび逆量子化のすべては図示されていないことを示す。マイクロフォンストリームは、変形離散コサイン変換(MDCT: modified discrete cosine transform)モジュール350を使用して処理されて、MDCT変換係数が生成される。

20

【0024】

音声ミキサは、ピア(1)から(N)までのネットワーク音声からの音声コンテンツおよびマイクロフォンストリーム320をミキシングするのに使用される。これによってミックス音声パケットが生成される。このミックス音声コンテンツは、ピア(1)360へのミックス音声として、ピア(N)へのミックス音声として、および、省略記号で示すように、その他のピアに対するミックス音声として再配信される。

【0025】

ピア(1)からピア(N)の配信されたネットワーク音声は、ローカル再生のためにもミキシングされる。次いで、(ローカル入力が他のピアによってミキシングされている場合には)ローカルマイクロフォン入力コンテンツを引いたミックス音声が、逆MDCTモジュール370に送り込まれる。これによって、ネットワークピアの音声コンテンツが回復され、スピーカ380を介して再生することが可能となる。

30

【0026】

ここで留意すべきことは、その他のピアからの音声は、P2Pネットワーク上で伝送される前にエンコードされることである。さらに、ミックス音声コンテンツがピアノードによって受信されると、そのピアは、ミックス音声をデコードして、音声通信セッションに関連する音声を回復する。したがって、ピアノードのそれぞれが、音声ミキサ330を包含する。さらに、音声ミキサ330は、複数の音声エンコーダおよび音声デコーダ構成要素をさらに包含する。いつどのように、これらの構成要素のそれぞれを使用するかは、任意の時間にピアが行っている処理によって決まる。例えば、ローカルピアがミキシングおよび再配信を実行しているときには、音声ミキサ330は、ミューチュアルキャストクリーク内の各ピアから1フレームのエンコード音声を受け取り、部分デコードを実行し、音声をミキシングして、部分エンコードを実行し、1フレームのミックス音声を各ネットワークピアに送る。そのローカルピアがクライアントピアとなり、ミキシングおよび再配信が別のピアに割り当てられるときには、音声ミキサ330は、単に1フレームのエンコードマイクロフォンストリーム入力を、ミキシングを担当するピアに送り、ミキシングピアから1フレームを受け取り、デコード演算を実行し、ローカルにミックス音声を再生する。P2P音声通信システムおよび方法が使用するラウンドロビン方式は、ローカルピアノードの機能を順繰りで迅速に変更するので、構成要素間の必要なスイッチングは迅速に行

40

50

われる。音声ミキサ 330、音声エンコーダ、音声デコーダの詳細を以下に説明する。

【0027】

I I I . 動作の概要

図3に示すP2P音声通信システム300の動作について、次に説明する。図4は、図3に示すP2P音声通信システム300の一般的動作を示す一般的フロー図である。一般に、本P2P音声通信方法は、入力音声ストリームをフレームまたはパケットに区分して、フレームの処理を、P2Pネットワーク内のピアノードの間で交代させて行う。より具体的には、本方法は、入力音声ストリームを複数のフレーム（ボックス400）に区分することから始まる。次に、P2Pネットワーク内のピアノードの1つが選択される（ボックス410）。 10

【0028】

選択されたピアノードは、その選択されたノードの利用可能な資源に比例するフレームを処理するように割り当てられる（ボックス420）。言い換えると、大量の資源を有するピアノードには、より多数のフレームが割り当てられるのに対して、少量の資源を有するピアノードには、より少数のフレームが割り当てられる。各ピアノードは、交代してフレームを処理する。処理には、入力音声ストリーム内の音声コンテンツのミキシングおよびピアノードへの音声の再配信が含まれる。すなわち、選択されたピアノードは、音声フレームのミキシングを実行して、ミックス音声フレームをP2Pネットワーク内の他のピアノードに再配信する。フレーム処理は、ピアノード間で交代して行われる（ボックス430）。好ましい実施形態では、この交代は、各ノードに順番が回るように、ラウンドロビン方式または方法で実行される。 20

【0029】

I V . 動作の詳細および実用例

図3および図4に示すP2P音声通信システムおよび方法の動作の詳細についてここで考察する。本明細書に開示するP2P音声通信システムおよび方法をより詳細に理解するために、例証的な実用例の動作の詳細を提示する。ここで留意すべきことは、この実用例は、P2P音声通信システムおよび方法を実現することのできる1つの方法にすぎないことである。 30

【0030】

音声コンテンツ

（波形コード化音声のミキシング）

この実用例においては、P2P音声通信システムおよび方法の音声は、波形コーデックを用いてエンコードされている。そのような波形コーデックに非特許文献5がある。代替コーデックとしては非特許文献6がある。さらなる代替コーデックとしては、ウィンドウズ（登録商標）メディアオーディオ（Windows（登録商標）Media Audio）（非特許文献7）がある。図5は、本実用例において使用される音声エンコーダ500の動作の詳細ブロック／フロー図である。 40

【0031】

図5を参照すると、音声波形を含む音声入力510は、最初に、フレームに分割または区分される。この実用例において、各フレームの長さは、約20ミリ秒（ms）であった。ここで留意すべきことは、フレーム間に50%の重複があるために、全アルゴリズム遅延は40msであり、フレームのそれの2倍であることである。しかしながら、それぞれの量子化およびエントロピーコード化されたフレーム係数の長さは、依然として20msである。次に、各フレームは、変形離散コサイン変換（MDCT）モジュール520によって係数ブロック（ $C_{i,j}$ ）に変換される。添え字*i*はピアを示し、添え字*j*はフレーム番号を示す。すなわち、図5において、係数ブロック $C_{i,1}$ は*i*番目のピアの第1フレームを、係数ブロック $C_{i,2}$ は*i*番目のピアの第2フレームを示す。次に、係数ブロックは、量子化器530に送られて、量子化される。次いで、この量子化係数ブロックは、エントロピーエンコーダ540によって、パケット $p_{i,j}$ にエントロピーエンコード（entropy encoded）される。図5において、エンコードパケット $P_{i,j}$ 50

$P_{1,i}, P_{2,j}$ は、 i 番目のピアの第 1 および第 2 のフレームに対するエンコードパケットを示す。

【0032】

M D C T は線形演算であるので、波形コード化音声は、変換ドメインにおいてのミキシングができるだけでなく、フレーム毎にミキシングすることもできる。図 6 は、この実用例において使用される音声ミキサ 600 の動作の詳細ブロック / フロー図である。図 6 に、2 つの圧縮音声パケットからなる、あるフレームのミキシングを示してある。一般に、音声ミキサ 600 は、あるフレームの圧縮音声パケットをデコードして係数ブロックを得て、係数ブロックを結合し、ブロックをリエンコードし、1 フレームの音声に対する単一の圧縮音声パケットを得る。P2P 音声通信システムおよび方法の音声ミキシング技法では、現行の音声会議システムにおいて行われるように、入力音声ストリーム全体についてのミキシング演算を実行するのではなく、音声のフレームをミキシングする。

10

【0033】

特に、この実用例においては、エントロピーデコーダ (1) 610 および逆量子化器 (1) 620 を使用して、圧縮音声パケット $P_{1,j}$ がデコードされて、あるピアに対する係数ブロックが得られる。同様に、エントロピーデコーダ (2) 630 および逆量子化器 (2) 640 を使用して、圧縮音声パケット $P_{2,j}$ がデコードされて、別のピアに対する係数ブロックが得られる。結果として得られる係数は、M D C T 変換係数である。これらの M D C T 変換係数は、結合モジュール 650 を使用して、結合モジュール 650 中の「+」記号で表わすように互いに加算されるか、または結合される。ここで留意すべきことは、図 6 には音声に対して「+」記号を示してあるが、結合は、減算（マイナス符号「-」）としても加算（プラス符号「+」）としてもよく、エンコード音声に対する聴覚上の差異はないということである。次に、結果として得られる係数ブロックは、ミキシング量子化器 660 を使用して量子化されて、それからミキシングエントロピーエンコーダ 670 を使用してエントロピーリエンコードされる。これによって、ミックス音声パケット $P_{1,j} + P_{2,j}$ を含む、フレームが生成される。このミキシング工程中に、その他のいずれのフレームの音声パケットにも、アクセスはなかった。

20

【0034】

図 7 は、この実用例において使用される音声デコーダ 700 の動作の詳細ブロック / フロー図である。受信器において、ミキシングされたパケットは、正常にデコードされた。図 7 を参照すると、ミックス音声パケット $P_{1,j} + P_{2,j}$ は、音声デコーダ 700 に入力された。各フレームは、エントロピーデコーダ (3) 710 によって、次いで逆量子化器 (3) 720 によって処理された。これによって、ミキシングされた M D C T 変換係数、 $C_{1,1} + C_{2,1}$ および $C_{1,2} + C_{2,2}$ が生成された。次に、ミキシングされた M D C T 変換係数は、逆 M D C T モジュール 730 を使用して処理された。結果として得られた出力は、ピア 1 およびピア 2 からの、ミックス音声波形（音声 1 + 音声 2）740 であった。

30

【0035】

（音声コンテンツのラウンドロビン式交代ミキシングおよび再配信）

この実用例における別の構成要素は、ラウンドロビン方式による音声コンテンツの交代ミキシングと再配信であった。言い換えると、ピアのそれぞれが、交代で音声コンテンツをミキシングおよび再配信した。この実用例においては、完全接続メッシュを形成する少數のピアノードからなる、ミューチュアルキャストクリークを使用した。波形コード化音声コンテンツは、フレームバイフレーム方式でミキシングすることができるという特性を使用して、P2P 音声会議システムおよび方法では、ミキシングおよび再配信演算がピア間で交代で行われた。これによって、最も効率的な方法で、帯域幅および計算負荷が適当に配分されることが確実になった。好ましい実施形態では、この交代は、ラウンドロビン方式および / または方法で行われ、それによって各ピアノードが交代して音声をミキシングして再配信した。

40

【0036】

50

図8は、この実用例において使用された、3つのピアノード(1、2および3)に対する音声ミキシングセッションを示す。音声ミキシングスケジュールを図8の下部に示してある。この音声ミキシングスケジュールを参照すると、ピアノード1、2、3は、それぞれ、フレーム $3k$ 、 $3k+1$ 、 $3k+2$ において音声ミキシングおよび再配信を担当した。すなわち、1番目のフレームにおいて、ピア2が音声パケットのミキシングおよび再配信を行った。図8に示すように、ピア2がミキシングおよび再配信を担当している間、ピア1およびピア3は、それらのエンコード音声 $P_{1,1}$ および $P_{3,1}$ をピア2に送った。次いで、入力音声パケットは、エントロビーエンコードおよび逆量子化を経て、MDCT係数 $C_{1,1}$ および $C_{3,1}$ に戻された。次いで、ピア2は、それ自体の係数 $C_{2,1}$ を加えた。次いで、ミックス音声は、ピア1およびピア3に送り返された。

10

【0037】

エコーを回避するために、ソース音声はミキシングせずに、送り返された。言い換えると、図8を参照すると、ピア2は、係数 $C_{1,1}$ および $C_{2,1}$ を加算して、この係数の合計を量子化およびエントロビーエンコードした。次いで、ピア2はミックスパケット $P_{1,1} + P_{2,1}$ をピア3に送り返した。同様に、ミックスパケット $P_{3,1} + P_{2,1}$ はピア2によってピア1に送られた。目的地において、異なるピアからのミックス音声パケットは、ソートされ、エントロビーエンコードされ、再生のために、逆量子化および逆MDCT変換された。

【0038】

再び、図8のミキシングスケジュールを参照すると、第2フレームにおいて、ピア3がミキシングの役割についた。ピア1およびピア2は、2番目のフレームにおけるそれらの圧縮音声パケット $P_{1,2}$ および $P_{2,2}$ をピア3に送った。次いで、ピア3は入力音声パケットを、それ自体の係数 $C_{3,2}$ とミキシングした。次いで、ピア3は、ミックスパケット $P_{3,2} + P_{2,2}$ をピア1に送り、ミックスパケット $P_{3,2} + P_{1,2}$ をピア2に送った。

20

【0039】

図8に示すように、3番目のフレームにおいて、ピアノード1がミキシングノードとなり、以下同様であった。ミキシングおよび再配信のタスクをタイムシェアリングすることによって、ミキシングの帯域幅および計算コストは、ピアのそれぞれの間に分配される。その結果として、サーバを必要とすることなく、1群の低性能のピアによって、マルチパーティ音声通信セッションを行うことができる。

30

【0040】

P2P音声通信システムおよび方法では、ミューチュアルキャストクリークはn個のノードからなる。各ピアノードは、 $2(n-1)$ 個のパケットをnフレーム毎に送受信する。その中で、 $(n-1)$ 個のパケットは、それがミキシング演算を実行しない $n-1$ 個のフレームの間に送受信される。さらに、 $(n-1)$ 個のパケットが、それがミキシングおよび再配信の演算を実行する間に、送受信される。したがって、アップロード/ダウンロードに必要な帯域幅は、 $(2-2/n) \cdot bw$ である。また、平均して、 $(2-2/n)$ 個のストリームの音声が各ピアによってデコード、リエンコードされるということも計算することができる。ミキシングの間に、ピアは、 $(n-1)$ 回のエントロビーエンコードおよび逆量子化の演算、および $(n-1)$ 回の前方量子化およびエントロビーエンコードの演算を実行する。

40

【0041】

再配信されたミックス音声パケットに含まれるコンテンツに関して、少なくとも2つの可能性がある。好ましい一実施形態においては、ミックス音声パケットは、ソース(または選択された)ピアノードの音声コンテンツは含まない。言い換えると、交代音声ミキシング技法は、あるピアのソース音声をミキシングせず、かつ送り返さないことを含む。一代替実施形態においては、ミックス音声パケットは、ピアノードのそれからの音声コンテンツを含み、そのミックス音声パケットからそれ自体の音声コンテンツを差し引くのは、各ピアに任せている。例えば、ミックス音声は、フレームjに対してパケット(m

50

$j = p_{1,j} + p_{2,j} + p_{3,j}$ を含む。次いで、同じミックス音声がすべてのピアに返送される。エコーを低減または除去するために、各ピアには、そのミックス音声からそれ自体の音声を差し引く責任がある。例えば、ピア i は m_j から $p_{i,j}$ を差し引き、それは加算の代わりに減算によるミキシング演算である。この代替実施形態の利点は、ピアは、ミキシング中に一回の前方量子化およびエントロピーエンコード演算を実行する必要があるだけであることである。さらに、IPマルチキャストがすべてのピア間でサポートされている場合には、ミキシングピアは、ミックスパケットを残りのピアにマルチキャストすることができます。この代替実施形態の欠点は、ミックス音声は、量子化してエントロピーエンコードする必要があるために、ミックスパケット m_j 内の構成音声 $p_{i,j}$ が、ピア i が保持する音声 $p_{i,j}$ と異なることである。したがって、残留エコーが残る可能性がある。この残留エコーは、ピアの数の増加および / またはミックス音声のコーディングビットレートの減少とともににより明白になる。したがって、この残留エコー問題のため、音声パケットのすべてをミックスしない実施形態が好ましい。

【0042】

ミキシングタスクの割り付け

P2P音声通信システムおよび方法は、音声のミキシングおよび再配信のタスクをフレームバイフレーム方式で割り付ける。このようにして、P2P音声通信システムおよび方法は、より多くの資源を有するピアにより多くのミキシングタスクを、より少ない資源を有するピアにより少ないミキシングタスクを割り当てる。P2P音声通信システムおよび方法において、最も重要な資源と考えられるのは、ピアのアップロード帯域幅である。普及が進むネットワークにおいては、P2Pネットワークの全アップロード帯域幅は、全ダウンロード帯域幅よりもはるかに小さい。このことは、ケーブルモデムおよびADSLネットワークのエンドユーザノードにおいて特に当てはまり、これらのエンドユーザノードに対しては、アップロード帯域幅とダウンロード帯域幅との釣合いがダウンロード帯域幅の方に非対称に偏っている。大学 / 企業ネットワークのユーザノードに対してさえ、ユーザがアップロード帯域幅に上限を設けるために、ダウンロード帯域幅は、利用可能なアップロード帯域幅よりも大きい可能性がある。したがって、より高い利用可能帯域幅を有するピアに、より多くのミキシングおよび再配信タスクを割り付け、より低いアップロード帯域幅を有するピアにより少ないタスクを割り付けるのが有利である。

【0043】

P2P音声通信システムおよび方法が考慮する第2の資源は、ピアのピークアップロード帯域幅（または物理リンク帯域幅）である。ミキシングの間に、P2P音声通信システムおよび方法のピアは、 $(n - 1)$ 個のパケットを受信し、 $(n - 1)$ 個のピアに送信する。P2P音声通信システムにおけるピアのトラフィック特性は、バースト性（bursty）である。より高速の物理リンクを有するピアに、または比較的多くのトーカンパケット（token bucket）を有するルータに接続されているピアに、より多くのミキシングおよび再配信タスクを割り当てる、パケットを複数のピアに送ることにより生じる遅延を低減することができるようになるのが有効である。

【0044】

通常、ピアのダウンロード帯域幅および計算資源はボトルネックではない。それでも、P2P音声通信システムおよび方法は、このことも割り付けにおいて考慮に入れることもできる。ここで留意すべきことは、低速または低性能のノードが、より少ないパケットを配信することを許容すると、それらのノードは、より高速で高性能のノードの寄生虫（leech）となることである。そのような寄生挙動を許容するかどうかは、良好な音声通信性能と寄与の公平性との間の設計選択である。

【0045】

遅延

ここで、P2P音声通信システムおよび方法の最大遅延を計算して、提示する。ピアノード i と j との間のネットワーク伝送遅延を $d_{i,j}$ とする。ミキシング演算で生ずる遅延は無視できると仮定して、ピア k によってミキシングされた音声フレームを受け取るた

10

20

30

40

50

めのピア i での遅延の量は次式となる。

【0046】

【数1】

$$D_{ik} = \left(\max_{j \neq k} d_{jk} \right) + d_{ki}, \quad (1)$$

【0047】

ピア i での最大遅延は次のように計算される。

【0048】

【数2】

$$D_i = \max_{k \neq i} \left(d_{ik} + \left(\max_{j \neq i, k} d_{jk} \right) \right), \quad (2)$$

10

【0049】

それに対して、P2P音声通信システムおよび方法の最大遅延は、

【0050】

【数3】

$$D_{\max} = 2 \max_{i \neq j} d_{ij}, \quad (3)$$

【0051】

であり、これは、ミューチュアルキャストクリークにおける最も遠いピア対のネットワーク遅延の2倍である。

20

【0052】

V. スーパーゲートウェイおよびスーパーサーバ

P2P音声通信システムのミューチュアルキャストクリークは、1組の完全接続されたピアノードによって形成されなければならない。上述のように、遅延は、クリークが大きくなるほど増大する。したがって、ミューチュアルキャストクリーク内のノード数は、大きすぎないことが理想的である。合理的な数としては、3から7の間である。しかしながら、ミューチュアルキャストクリークは、スーパーゲートウェイまたはスーパーサーバとしての役割を果たし、したがって、本明細書に記載したP2P音声通信システムおよび方法のすべての機能を保持しながら、より大規模なネットワークにおいて機能することができる。

30

【0053】

スーパーゲートウェイノード

P2P音声通信システムは、スーパーゲートウェイノードとして機能することができる。これによって、ミューチュアルキャストクリークは、大規模マルチパーティ通信セッションにおいて、「スーパー」ゲートウェイノードとしての役割を果たすことができる。この場合に、残りのノードは、汎用グラフ(generic graph)を形成して、エンドシステムミキシングを使用する。図9は、汎用マルチパーティ通信セッションにおけるゲートウェイノードとしての役割を果たすP2P音声通信システムおよび方法のミューチュアルキャストクリークの例証的例を示す。この構成は、多数のピアノードを有するマルチパーティ通信セッションであって、少数の近接するノードが完全接続されてマルチキャストクリークを形成するものに特に適している。

40

【0054】

図9の例証的例に示すように、ノードa、b、cによって形成されるミューチュアルキャストクリーク900は、ノードd、e、f、g、iに対する「スーパー」ゲートウェイノードとしての役割を果たす。ミューチュアルキャストクリーク900内の各ピアノードは、取り付けられたノードに対してゲートウェイとしての役割を果たす。例えば、ピアノードaは、マルチキャストクリーク900の外側の2つのノードd、eに取り付けられている。すなわち、ミューチュアルキャストクリーク900において、ノードaは、dおよびeの音声をそれ自体の音声と併合して、結合音声(a+d+e)を、P2P音声通信システムおよび方法を使用して、ノードb、cに配信する。ノードaがミューチュアルキャ

50

ストクリーク 900 のためにミキシングを行っているフレームにおいて、結合音声 (a + d + e) が、ノード b からの入力とミックスされて (音声 b + i) 、ノード c に送られる。同様に、結合音声 (a + d + e) は、ノード c からの入力とミックスされて (音声 c + f + g + h) 、ノード b に送られる。ノード a は、また、ノード b 、 c からの入力をミックスして、それをそれ自体と結合し、

$$m = a + b + i + c + f + g + h \quad (4)$$

m + d をノード e に送り、 m + e をノード d に送る。

【0055】

ノード a がミキシングしていないフレームにおいて、結合音声 (a + d + e) は、その時点のミキシングノード (b または c) に送信される。また、ノード a は、ノード b 、 c からのミックス入力も受信し、それらをそれ自体のものと結合してミックスフレーム m を形成し、ノード e に m + d を送信し、ノード d に m + e を送信する。 10

【0056】

スーパーサーバ

P2P 音声通信システムは、スーパーサーバとして機能することもできる。この場合に、残りのノードは、スタートポロジにおけるクライアントノードであり、ミューチュアルキャストクリークは、最低 2 つのノードで構成することができる。図 10 は、ミューチュアルキャストクリークの、スタートポロジを有するマルチパーティ通信セッションにおける、スーパーサーバとしての例証的例を示す。 20

【0057】

図 10 に示すように、ピアノード a 、 b は、 2 ノードミューチュアルキャストクリーク 1000 を形成し、これは、残りのクライアントノード c 、 d 、 e 、 f 、 g に対してスーパーサーバとしての役割を果たす。ミューチュアルキャストクリーク 1000 内のピアノードは、また完全接続メッシュを形成する。さらに、外側の各クライアントノードは、ミューチュアルキャストクリーク 1000 内のすべてのピアノードに接続されている。この構成は、分担してサーバとしての役割をする、いくつかの高性能ブロードバンドノードが存在する、 (4 から 16 の間などの) 小規模～中規模のネットワークにより適している。

【0058】

この構成に伴う別のシナリオは、ネットワークアドレス変換機構 (NAT : n e t w o r k address translator) またはファイアウォールがあるときに発生する。具体的には、クライアントノードは、NAT / ファイアウォールの背後に置くことができる。これらは、インターネットに直接的に接続されたノード、言い換えると、ミューチュアルキャストクリーク 1000 のノードに接続することができる。しかしながら、それらは互いに接続することはできない。そのようなネットワークのミキシングおよび再配信演算は、上述の P2P 音声通信システムおよび方法に非常に類似している。唯一の差は、この実施形態においては、クラントノードは、ミキシングおよび再配信のタスクから免除されていることである。各フレームにおいて、ミューチュアルキャストクリーク 1000 の 1 つのピアは、ミューチュアルキャストクリーク 1000 の内側および外側の両方における、残りのピアのために音声パケットをミキシングして再配信する。 30

【0059】

V I . 帯域幅および計算負荷解析

この章においては、 P2P 音声通信システムおよび方法を使用する、異なる音声通信セッションシナリオの帯域幅要件を計算して、 P2P 音声通信システムおよび方法を使用しないシナリオと比較する。最初に、 n パーティ通信セッションを考える。すべてのピアの帯域幅が等しい場合に、 P2P 音声通信システムおよび方法は、それぞれすべてのピアノードに対して、 (2 - 2 / n) · bw のアップロード / ダウンロード帯域幅を必要とする。特に、 3 ノードミューチュアルキャストクリークに対して、必要な帯域幅は 1.34 bw である。スタートポロジ、または汎用グラフを用いるマルチパーティ通信セッションを行うために、同じ 3 ノード通信セッションは、少なくとも 2 bw の帯域幅のノードを必要とする。すなわち、 P2P 音声通信システムおよび方法は、すべてのピアノードがより少 40

ない資源しか有さない場合でも、マルチパーティ通信セッションを実施することができる。

【0060】

第2に、ミューチュアルキャストクリークが、大規模グラフにおいてスーパーゲートウェイとして働く場合を考える。ゲートウェイノードが、m個のノードに接続されているとする。通常、ゲートウェイノードは、音声トラフィックをミキシングして再配信するのに、 $m \cdot bw$ のアップロードおよびダウンロード帯域幅を必要とする。ゲートウェイノードをnノードミューチュアルキャストクリークで置換し、 m/n ノードが各ノードに取り付けられていると仮定することによって、各ノードのアップロード/ダウンロード帯域幅要件は、このスーパーゲートウェイにおいては、以下のように低減される：

10

【0061】

【数4】

$$\left(2 + \frac{m-2}{n}\right)bw. \quad (5)$$

【0062】

一例として、 $m = 6$ および $n = 3$ とする。この場合に、P2P音声通信システムおよび方法は、帯域幅要件を、 $6bw$ から $3.333bw$ へと低減する。ミューチュアルキャストクリークの、スーパーゲートウェイとしての使用は、また、ゲートウェイノードの帯域幅要件を低下させることができる。

20

【0063】

最後に、ミューチュアルキャストクリークをスーパーサーバとして使用する場合を考える。ここでも、m個のクライアントがあると仮定する。P2P音声通信システムおよび方法を使用しないと、サーバは、m個のクライアントにサービスを提供するのに、 $m \cdot bw$ のアップロードおよびダウンロード帯域幅を必要とする。しかしながら、P2P音声通信システムおよび方法ならびにn個のノードのミューチュアルキャストクリークを使用する場合に、ここでも、平均で、ミューチュアルキャストクリーク内の各ピアノードは、

【0064】

【数5】

$$\left(2 + \frac{m-2}{n}\right)bw$$

30

【0065】

のアップロード/ダウンロード帯域幅を必要とするだけであると計算される。一例として、 $m = 5$ および $n = 2$ とする。P2P音声通信システムおよび方法は、サーバノードの帯域幅要件を、 $5bw$ から $3.5bw$ へと低減する。ここで、ミューチュアルキャストは、ピアの計算負荷、スーパーゲートウェイノードおよびスーパーサーバノードと同じ割合で低減することは、容易に推論できる。

【0066】

VII. 例証的動作環境

P2P音声通信システムおよび方法は、計算機環境内および計算機装置上で動作するように設計されている。P2P音声通信システムおよび方法が動作する計算機環境について考察する。以下の考察は、P2P音声通信システムおよび方法を実装することのできる、適当な計算機環境の簡潔な、一般的記述を行うものである。

40

【0067】

図11は、図3および図4に示したP2P音声通信システムおよび方法を実装することのできる適当な計算機システム環境の例を示している。この計算機システム環境1100は、適当な計算機システム環境の一例にすぎず、本発明の使用または機能に対するいかなる限定も示唆するものではない。また、計算機環境1100は、この例証的動作環境1100に示す構成要素のいずれかまたはそれらの組合せに関して、何らかの依存性または要件を有すると解釈すべきではない。

50

【0068】

P2P音声通信システムおよび方法は、その他多数の汎用または専用の計算機システム環境または構成とともに動作することができる。P2P音声通信システムおよび方法に使用するのに適するよく知られている計算機システム、環境、および/または構成の例としては、それに限定はされないが、パーソナルコンピュータ、サーバコンピュータ、ハンドヘルド、ラップトップもしくはモバイルコンピュータ、またはセルフォンやPDAなどの通信デバイス、マルチプロセッサシステム、マイクロプロセッサベースシステム、セットトップボックス(set top boxes)、プログラム可能家電製品、ネットワークPC、ミニコンピュータ、メインフレームコンピュータ、上記のシステムまたはデバイスの任意のものを含む分散型計算機環境、その他が挙げられる。

10

【0069】

P2P音声通信システムおよび方法は、プログラムモジュールなどの、コンピュータによって実行されているコンピュータ実行可能命令の一般的文脈で説明することができる。一般に、プログラムモジュールとしては、特定のタスクを実行するか、または特定の抽象データ型を実施する、ルーチン、プログラム、オブジェクト、コンポーネント、データ構造、その他が挙げられる。P2P音声通信システムおよび方法は、通信ネットワークによってリンクされたリモート処理装置によってタスクが実行される、分散型計算機環境において実施することもできる。分散型計算機環境においては、プログラムモジュールは、メモリ記憶装置を含む、ローカルおよびリモートのコンピュータ記憶媒体の両方に配置することができる。図11を参照すると、P2P音声通信システムおよび方法を実施するための例証的システムには、コンピュータ1110の形態の汎用計算機装置が含まれる。図11に示すピアノードは、コンピュータ1110の例である。

20

【0070】

コンピュータ1110の構成要素としては、それに限定はされないが、処理ユニット1120、システムメモリ1130、およびシステムメモリを含む様々なシステム構成要素を処理ユニット1120に結合する、システムバス1121を含めることができる。システムバス1121は、メモリバスもしくはメモリコントローラ、周辺バス、および様々なバスアーキテクチャの任意のものを使用するローカルバスを含む、いくつかの種類のバス構造の任意のものとすることができます。一例であり、限定ではないが、そのようなアーキテクチャには、ISA(Industrial Standard Architecture)バス、MCA(Micro channel Architecture)バス、EISA(Enhanced ISA)バス、VESA(Video Electronics Standards Association)ローカルバス、およびMezzanineバスとも呼ばれる、PCI(Peripheral Component Interconnect)バスがある。

30

【0071】

コンピュータ1110は、通常、様々なコンピュータ可読媒体を含む。コンピュータ可読媒体は、コンピュータ1110がアクセス可能であり、揮発性および不揮発性媒体、移動型および固定型媒体の両方を含む、任意の利用可能な媒体とすることができます。一例であって、限定ではないが、コンピュータ可読媒体には、コンピュータ記憶媒体および通信媒体を含めることができます。コンピュータ記憶媒体としては、コンピュータ可読命令、データ構造、プログラムモジュール、その他のデータなどの情報を記憶するために、任意の方法または技術で実現される、揮発性および不揮発性、移動型および固定型の媒体が挙げられる。

40

【0072】

コンピュータ記憶媒体としては、それに限定はされないが、RAM、ROM、EEPROM、フラッシュメモリもしくはその他のメモリ技術、CD-ROM、DVD(デジタルバーサタイルディスク)もしくはその他の光学ディスク記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置もしくはその他の磁気記憶装置、または所望の情報を記憶するのに使用可能であるとともにコンピュータ1110によるアクセスが可能である、その

50

他任意の媒体を挙げることができる。通信媒体は、通常、コンピュータ可読命令、データ構造、プログラムモジュール、または他のデータを、搬送波または他のトランスポート機構などの変調されたデータ信号内に具体化し、通信媒体には、すべての情報配布媒体が含まれる。

【0073】

ここで、「変調データ信号(modulated data signal)」という用語は、信号内に情報をエンコードするように、その1つまたは複数の特性を設定または変更した信号を意味する。一例であり、限定ではないが、通信媒体としては、有線ネットワークまたは直接線接続(direct-wired connection)、および音響、RF、赤外線およびその他の無線媒体などの無線媒体を挙げることができる。上記の任意のものの組合せも、コンピュータ可読媒体の範囲に含めるべきである。10

【0074】

システムメモリ1130には、読み取り専用メモリ(ROM)1131およびランダムアクセスメモリ(RAM)1132などの、揮発性および/または不揮発性メモリの形態のコンピュータ記憶媒体が含まれる。例えば起動時に、コンピュータ1110内の要素間で情報を転送するのを助ける、基本ルーチンを含む、基本入出力システム1133(BIOS)は、通常、ROM1131に記憶される。RAM1132は、通常、データおよび/またはプログラムモジュールを含み、それらは、処理ユニット1120に対して直接アクセス可能であり、かつ/またはその上で動作している。一例であって、限定ではなく、図11には、オペレーティングシステム1134、アプリケーションプログラム1135、その他のプログラムモジュール1136、およびプログラムデータ1137を示してある。20

【0075】

コンピュータ1110には、その他の移動型/固定型、揮発性/不揮発性コンピュータ記憶媒体を含めることができる。一例としてだけであるが、図11には、固定型、不揮発性磁気媒体の読み取り/書き込みを行うハードディスクドライブ1141、移動型、不揮発性磁気ディスク1152の読み取り/書き込みを行う磁気ディスクドライブ1151、CD-ROMまたはその他の光学媒体などの移動型、不揮発性光学ディスク1156の読み取り/書き込みを行う光学ディスクドライブ1155を示してある。

【0076】

例証的動作環境において使用することのできる、その他の移動型/固定型、揮発性/不揮発性コンピュータ記憶媒体としては、それに限定はされないが、磁気テープカセット、フラッシュメモリカード、DVD、デジタルビデオテープ、ソリッドステートRAM、ソリッドステートROM、その他が挙げられる。ハードディスクドライブ1141は、通常、インターフェース1140などの固定型メモリインターフェースを介してシステムバス1121に接続されており、磁気ディスクドライブ1151および光学ディスクドライブ1155は、通常、インターフェース1150などの、移動型メモリインターフェースによってシステムバス1121に接続されている。30

【0077】

上述して図11に示した、ドライブ類およびそれに関連するコンピュータ記憶媒体は、コンピュータ可読命令、データ構造、プログラムモジュールおよびコンピュータ1110用のその他のデータを記憶する。図11には、例えば、ハードディスクドライブ1141は、オペレーティングシステム1144、アプリケーションプログラム1145、他のプログラムモジュール1146、およびプログラムデータ1147を記憶する状態で示してある。ここで留意すべきことは、これらの構成要素は、オペレーティングシステム1134、アプリケーションプログラム1135、他のプログラムモジュール1136、およびプログラムデータ1137と同じでも、異なってもよいことである。オペレーティングシステム1144、アプリケーションプログラム1145、他のプログラムモジュール1146、およびプログラムデータ1147には、それらが少なくとも異なるコピーであることを示すために、ここでは異なる番号を与えてある。ユーザは、キーボード4050

1162 および一般にマウスと呼ばれるポインティングデバイス 1161、トラックボールまたはタッチパッドなどの入力装置を介して、コマンドおよび情報をコンピュータ 1110 に入力することができる。

【0078】

その他の入力装置（図示せず）としては、マイクロフォン、ジョイスティック、ゲームパッド、サテライトディッシュ、スキャナー、無線受信器、またはテレビジョンもしくはビデオ受信機、その他を挙げることができる。これらおよびその他の入力装置は、システムバス 1121 に結合されているユーザ入力インターフェース 1160 を介して、処理ユニット 1120 に接続されることが多いが、例えば、パラレルポート、ゲームポートまたはユニバーサルシリアルバス（USB）などの、その他のインターフェースおよびバス構造によって接続することもできる。モニタ 1191 またはその他の種類のディスプレイ装置も、ビデオインターフェース 1190 などのインターフェースを経由してシステムバス 1121 に接続される。モニタに加えて、コンピュータには、スピーカ 1197 およびプリンタ 1196 などのその他の周辺出力装置を含めてよく、これらは出力周辺インターフェース 1195 を介して接続することができる。10

【0079】

コンピュータ 1110 は、リモートコンピュータ（複数を含む）1180 などの、1つまたは複数のリモートコンピュータへの論理接続を使用して、ネットワーク化された環境で動作可能である。リモートコンピュータ 1180 は、パーソナルコンピュータ、サーバ、ルータ、ネットワーク PC、ピアデバイスまたはその他の共通ネットワークノードとしてもよく、図 11 にはメモリ記憶装置 1181 だけを示してあるが、通常は、コンピュータ 1110 に関して記述した要素の多数または全部を含む。図 11 に示す論理接続は、ローカルエリアネットワーク（LAN）1171 およびワイドエリアネットワーク（WAN）1173 を含むが、その他のネットワークを含めることもできる。そのようなネットワーク環境は、オフィス、企業内コンピュータネットワーク、インターネットおよびインターネットにおいて普及している。20

【0080】

LAN ネットワーク化環境において使用される場合には、コンピュータ 1110 は、ネットワークインターフェースまたはアダプタ 1170 を介して、LAN 1171 に接続される。WAN ネットワーク化環境において使用される場合には、コンピュータ 1110 は、通常、モデム 1172 またはインターネットなどの WAN 1173 上での通信を確立するための、その他の手段を含む。モデム 1172 は、内部式でも外部式でもよく、ユーザ入力インターフェース 1160、またはその他の適当な機構を経由してシステムバス 1121 に接続することができる。ネットワーク化環境において、コンピュータ 1110 に関して図示されたプログラムモジュールまたはその一部を、リモート記憶装置に格納することもできる。限定ではなく一例として、図 11 には、メモリ装置 1181 上に常駐するものとしてリモートアプリケーションプログラム 1185 を示してある。ここで、図示したネットワーク接続は、例証的なものであり、コンピュータ間の通信リンクを確立するためのその他の手段を使用することに気づくであろう。30

【0081】

上記の本発明の説明は、例証と説明のために提示したものである。それは網羅的とするものでも、開示した厳密な形態に本発明を限定するものでもない。上記の教示に照らせば、多くの修正形態および変形形態が可能である。本発明の範囲は、上記の発明の詳細な説明によってではなく、添付の特許請求の範囲によって限定することを意図するものである。40

【図面の簡単な説明】

【0082】

【図 1A】スタートポロジを有するコンピュータネットワークを示すブロック図である。

【図 1B】完全接続ユニキャストトポロジを有するコンピュータネットワークを示すブロック図である。50

【図1C】エンドシステムミキシングを使用する汎用グラフトポロジを有するコンピュータネットワークを示すブロック図である。

【図2】本明細書で開示するP2P音声通信システムおよび方法の例証的実現形態を示すブロック図である。

【図3】図2に示すようなピアノードのそれぞれに収納されている、P2P音声通信システムの例証的実現形態を示す、詳細ブロック図である。

【図4】図3に示すP2P音声通信システムの一般的動作を示す、全体フロー図である。

【図5】この実用例において使用される、音声エンコーダの動作の詳細ブロック／フロー図である。

【図6】この実用例において使用される、図3に示す音声ミキサの動作の詳細ブロック／フロー図である。 10

【図7】この実用例において使用される、音声デコーダの動作の詳細ブロック／フロー図である。

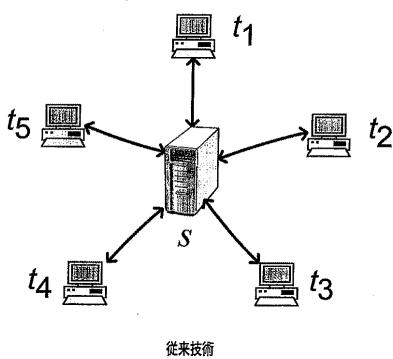
【図8】この実用例において使用される、3つのピアノード（1、2および3）に対する、図6に示す音声ミキサの動作を示す図である。

【図9】汎用マルチパーティ会議グラフにおいてゲートウェイノードとして働く、P2P音声通信システムおよび方法のミューチュアルキャストクリークの例証的例を示す図である。

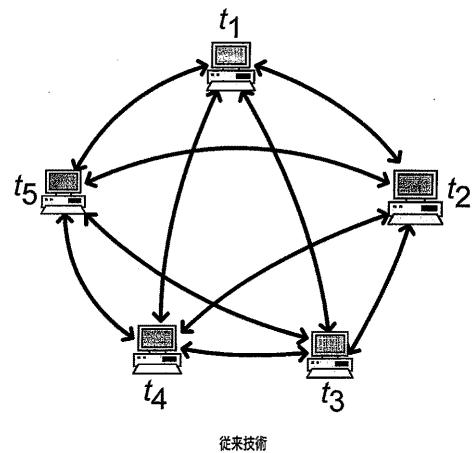
【図10】スタートポロジを備えるマルチパーティ通信セッションにおいてスーパーサーバとして働く、P2P音声通信システムおよび方法のミューチュアルキャストの例証的例を示す図である。 20

【図11】図3および図4に示すP2P音声通信システムおよび方法を実装することのできる、適当な計算機システム環境の例を示す図である。

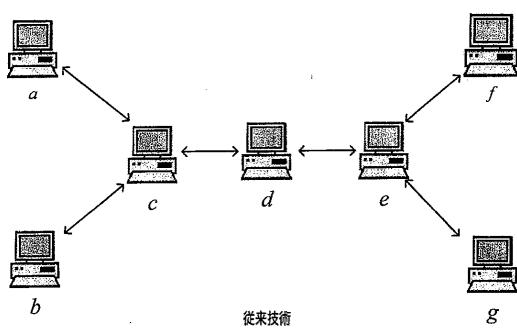
【図1A】



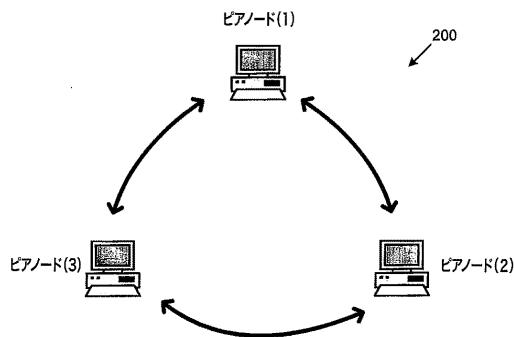
【図1B】



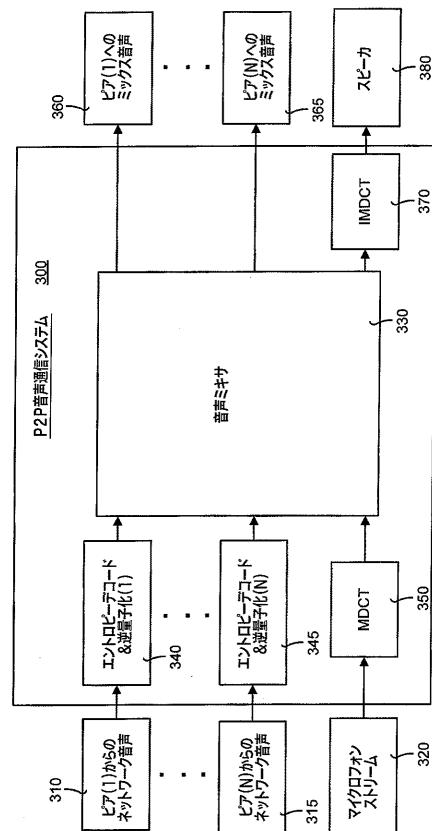
【図1C】



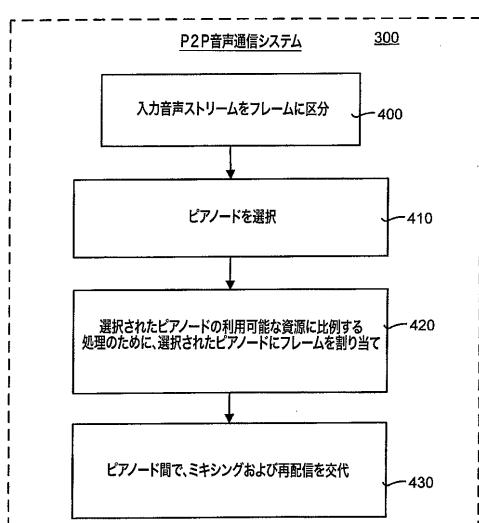
【図2】



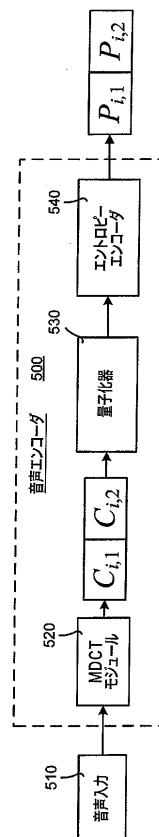
【図3】



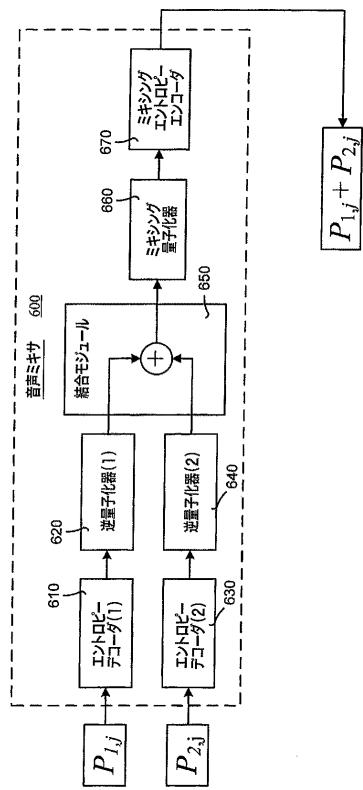
【図4】



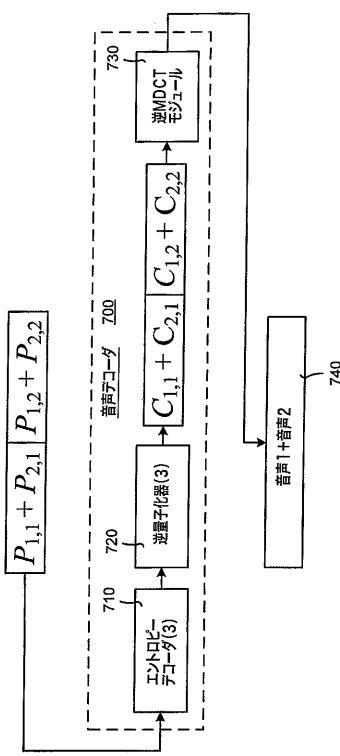
【図5】



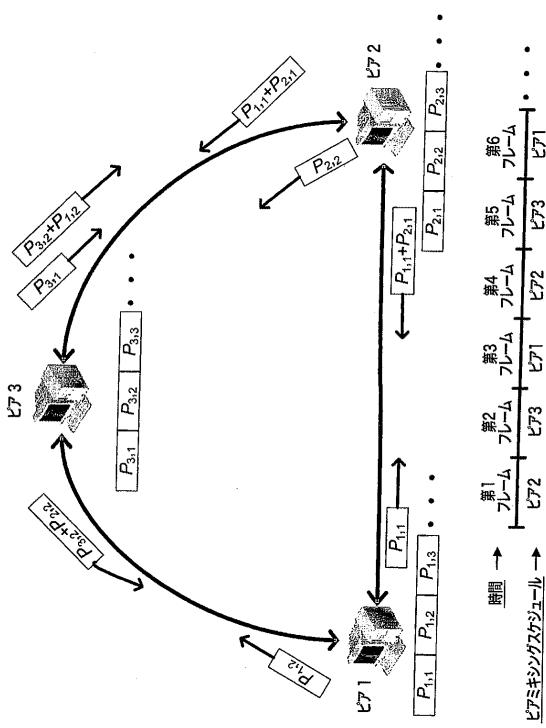
【図6】



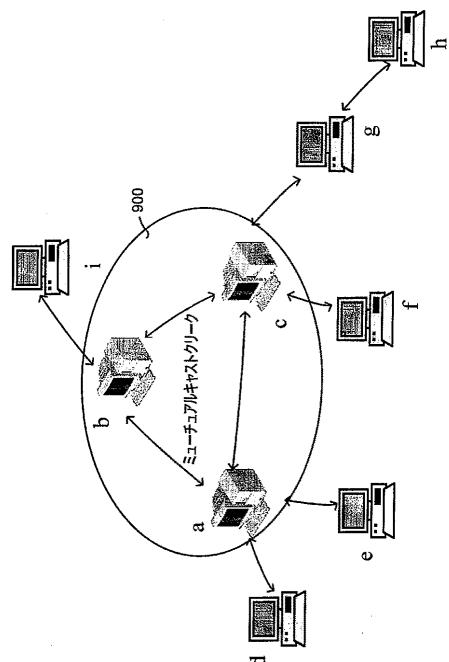
【図7】



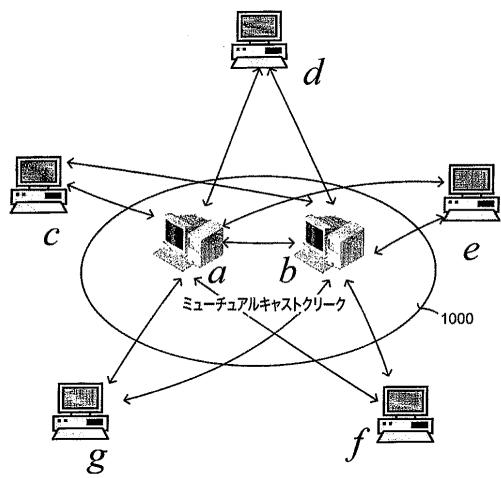
【図8】



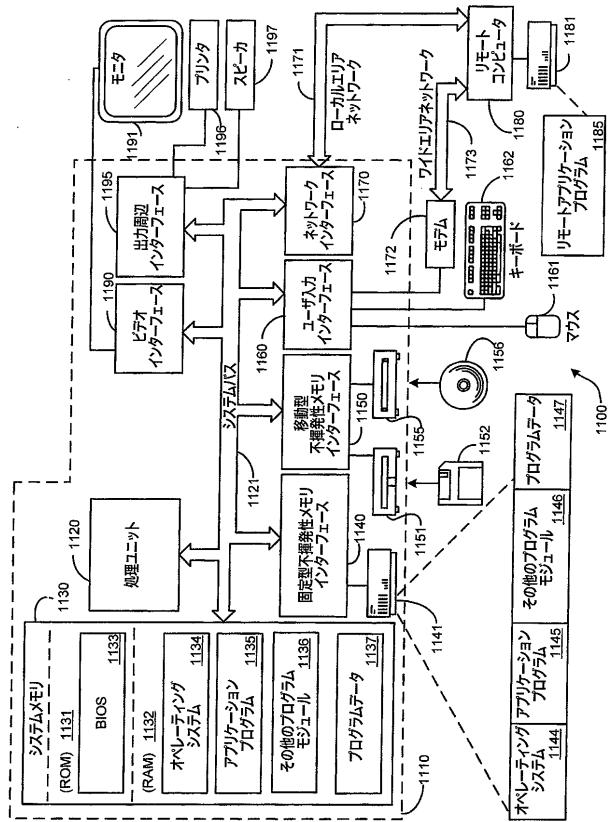
【図9】



【図10】



【図11】



フロントページの続き

審査官 木村 雅也

(56)参考文献 特開2003-323173(JP,A)
特開2003-124953(JP,A)
米国特許出願公開第2003/0182001(US,A1)
米国特許出願公開第2002/0178260(US,A1)

(58)調査した分野(Int.Cl., DB名)

G 06 F	13 / 00
H 04 L	12 / 18
H 04 R	3 / 00
H 04 M	3 / 00