

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6442079号  
(P6442079)

(45) 発行日 平成30年12月19日(2018.12.19)

(24) 登録日 平成30年11月30日(2018.11.30)

(51) Int.Cl.	F I
G 1 0 L 19/00 (2013.01)	G 1 0 L 19/00 3 3 0 C
G 1 0 L 19/005 (2013.01)	G 1 0 L 19/005

請求項の数 30 (全 54 頁)

(21) 出願番号	特願2017-553222 (P2017-553222)	(73) 特許権者	507364838
(86) (22) 出願日	平成28年3月10日(2016.3.10)		クアルコム、インコーポレイテッド
(65) 公表番号	特表2018-514804 (P2018-514804A)		アメリカ合衆国 カリフォルニア 921
(43) 公表日	平成30年6月7日(2018.6.7)		21 サン ディエゴ モアハウス ドラ
(86) 国際出願番号	PCT/US2016/021869		イブ 5775
(87) 国際公開番号	W02016/167904	(74) 代理人	100108453
(87) 国際公開日	平成28年10月20日(2016.10.20)		弁理士 村山 靖彦
審査請求日	平成30年8月8日(2018.8.8)	(74) 代理人	100163522
(31) 優先権主張番号	62/147,386		弁理士 黒田 晋平
(32) 優先日	平成27年4月14日(2015.4.14)	(72) 発明者	スバシンハ・シャミンダ・スバシンハ
(33) 優先権主張国	米国 (US)		アメリカ合衆国・カリフォルニア・921
(31) 優先権主張番号	62/147,431		21-1714・サン・ディエゴ・モアハ
(32) 優先日	平成27年4月14日(2015.4.14)		ウス・ドライブ・5775
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 デジッタバッファの更新

(57) 【特許請求の範囲】

【請求項1】

第1のデータと第2のデータとを含むパケットを受信するように構成されるデジッタバッファであって、前記第1のデータが、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデータの部分的なコピーを含み、前記第2のデータが、フレームの前記シーケンスのうちの第2のフレームに対応する、デジッタバッファと、

分析器であって、前記パケットを受信したことに応答して、

前記第2のデータと関連付けられる第2のフレーム受信タイムスタンプを生成し、

前記第2のフレーム受信タイムスタンプに基づいて、および前記第1のフレームと前記第2のフレームとの間のフレームオフセットに基づいて、前記第1のデータと関連付けられる第1のフレーム受信タイムスタンプを生成する

ように構成され、前記第1のフレーム受信タイムスタンプが、前記第2のフレーム受信タイムスタンプによって示される第2の時間より早い第1の時間を示す、前記分析器と、を備える、デバイス。

【請求項2】

前記第1のデータおよび前記第1のフレーム受信タイムスタンプに基づいて第1の出力を生成し、

前記第2のデータおよび前記第2のフレーム受信タイムスタンプに基づいて第2の出力を生成し、前記第2のフレーム受信タイムスタンプがクロック出力に基づき生成される

ように構成される発話デコーダと

10

20

をさらに備える、請求項1に記載のデバイス。

【請求項3】

前記発話デコーダがさらに、

前記第2のデータと関連付けられる第2のバッファリング期間の満了の前に満了する第1のバッファリング期間の満了の後で前記第1の出力を生成し、前記第1のバッファリング期間の前記満了が前記第1のフレーム受信タイムスタンプに基づき、前記第2のバッファリング期間の前記満了が前記第2のフレーム受信タイムスタンプに基づき、

前記第2のバッファリング期間の前記満了の後で前記第2のデータに基づいて前記第2の出力を生成するように構成される、請求項2に記載のデバイス。

【請求項4】

前記デジッタバッファにて1または複数のフレームを処理するように構成される発話デコーダと、

前記発話デコーダによって直近に処理された特定のフレームの生成タイムスタンプを記憶するように構成されるメモリをさらに備え、

前記分析器がさらに、前記第1のフレームの第1の生成タイムスタンプを決定するように構成される、請求項1に記載のデバイス。

【請求項5】

前記パケットが前記第2のフレームの第2の生成タイムスタンプを示し、前記第1の生成タイムスタンプがフレームオフセットおよび前記第2の生成タイムスタンプに基づいて決定され、前記分析器がさらに、前記第1の生成タイムスタンプが前記特定のフレームの前記生成タイムスタンプを超えているとの決定にตอบสนองして、第1のバッファエントリに前記第1のフレーム受信タイムスタンプを記憶するように構成される、請求項4に記載のデバイス。

【請求項6】

第1のデータと第2のデータとを含むパケットをデバイスのデジッタバッファにおいて受信するステップであって、前記第1のデータが、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデータの部分的なコピーを含み、前記第2のデータが、フレームの前記シーケンスのうちの第2のフレームに対応する、ステップと、

前記パケットを受信したことに応答して、

前記第2のデータと関連付けられる第2のフレーム受信タイムスタンプを生成するステップと、

前記第2のフレーム受信タイムスタンプに基づいて、および前記第1のフレームと前記第2のフレームとの間のフレームオフセットに基づいて、前記第1のデータと関連付けられる第1のフレーム受信タイムスタンプを生成するステップであって、前記第1のフレーム受信タイムスタンプが、前記第2のフレーム受信タイムスタンプによって示される第2の時間より早い第1の時間を示す、ステップと、

を備える、方法。

【請求項7】

前記第1のデータおよび前記第1のフレーム受信タイムスタンプに基づいて第1の出力を生成するステップと、

前記第2のデータおよび前記第2のフレーム受信タイムスタンプに基づいて第2の出力を生成するステップであって、前記第2のフレーム受信タイムスタンプがクロック出力に基づき生成される、ステップと、

をさらに備える、請求項6に記載の方法。

【請求項8】

前記第1の出力が、前記第2のデータと関連付けられる第2のバッファリング期間の満了の前に満了する第1のバッファリング期間の満了の後で生成され、前記第1のバッファリング期間の前記満了が前記第1のフレーム受信タイムスタンプに基づき、前記第2のバッファリング期間の前記満了が前記第2のフレーム受信タイムスタンプに基づき、前記第2の出力が前記第2のバッファリング期間の前記満了の後で生成される、請求項7に記載の方法。

**【請求項 9】**

前記デジッタバッファに前記パケットをデュアルエンキューするステップをさらに備え、前記パケットをデュアルエンキューするステップが、

前記デジッタバッファの第1のバッファエントリに前記第1のデータと前記第1のフレーム受信タイムスタンプとを記憶するステップであって、前記第1のバッファエントリが前記第1のフレームに対応する、ステップと、

前記デジッタバッファの第2のバッファエントリに前記第2のデータと前記第2のフレーム受信タイムスタンプとを記憶するステップであって、前記第2のバッファエントリが前記第2のフレームに対応する、ステップとを備える、請求項6に記載の方法。

**【請求項 10】**

10

前記第1のバッファエントリに前記第1のフレーム受信タイムスタンプを記憶するステップが、前記第2のフレームに対応する第2のバッファリング期間の満了の前の前記第1のフレームの前記部分的なコピーの処理を可能にする、請求項9に記載の方法。

**【請求項 11】**

前記第1のフレームが前記デジッタバッファに存在しないとの決定に応答して無音の期間を延ばすステップであって、前記第1のフレームが無音の前記期間に後続するトークスパートの第1のフレームであり、無音の前記期間がオーディオ入力のない期間に対応する、ステップと、

無音の前記期間を延ばした後で前記第1のフレームを処理するステップとをさらに備える、請求項6に記載の方法。

20

**【請求項 12】**

前記第1のフレームデータが前記デジッタバッファに存在しない、および前記パケットが前記第1のフレームデータの前記部分的なコピーを含むとの決定に応答して、

第1のメモリ位置に前記パケットの第1のコピーを記憶するステップであって、前記デジッタバッファの第1のバッファエントリが前記第1のメモリ位置を示す、ステップと、

第2のメモリ位置に前記パケットの第2のコピーを記憶するステップであって、前記デジッタバッファの第2のバッファエントリが前記第2のメモリ位置を示す、ステップとをさらに備える、請求項6に記載の方法。

**【請求項 13】**

前記第1のバッファエントリがフレームのデータの部分的なコピーに対応することを示すために、前記第1のバッファエントリと関連付けられる第1のフラグを第1の値に設定するステップと、

30

前記第2のバッファエントリがフレームのデータの完全なコピーに対応することを示すために、前記第2のバッファエントリと関連付けられる第2のフラグを第2の値に設定するステップとをさらに備え、前記第2のデータが前記第2のフレームの第2のフレームデータの完全なコピーを含む、請求項12に記載の方法。

**【請求項 14】**

前記第1のフレーム受信タイムスタンプおよび特定の再生遅延に基づいて、前記第1のフレームに対応する第1のバッファリング期間を決定するステップと、

前記第2のフレーム受信タイムスタンプおよび前記特定の再生遅延に基づいて、前記第2のフレームに対応する第2のバッファリング期間を決定するステップとをさらに備え、

40

前記第1のフレームの第1の再生時間が前記第2のフレームの第2の再生時間より早い、請求項6に記載の方法。

**【請求項 15】**

前記パケットを受信する前に、および前記第1のフレームが前記デジッタバッファに存在しないとの決定に応答して、

消去を再生するステップと、

バッファタイムラインの値を更新するステップとをさらに備える、請求項6に記載の方法。

**【請求項 16】**

50

前記デジッタバッファの第1のバッファエントリに前記第1のデータを記憶するステップであって、前記第1のバッファエントリが前記第1のフレームに対応する、ステップと、  
前記第1のフレームデータを含む第2の packets を受信するステップと、  
前記第1のデータを前記第1のフレームデータで置き換えるために前記デジッタバッファの前記第1のバッファエントリを更新するステップとをさらに備える、請求項6に記載の方法。

【請求項 17】

前記デジッタバッファの前記第1のバッファエントリに前記第1のフレーム受信タイムスタンプを記憶するステップと、

前記第2の packets の更新された受信タイムスタンプを決定するステップと、

前記第1のフレーム受信タイムスタンプを前記第2の packets の前記更新された受信タイムスタンプで置き換えるために前記第1のバッファエントリを更新するステップとをさらに備える、請求項16に記載の方法。

【請求項 18】

複数の packets を受信するステップと、

前記複数の packets に基づいて、フレーム消去レート、オフセット値、またはフレーム結束値のうちの少なくとも1つを決定するステップと、

前記フレーム消去レートが第1の閾値以上であること、前記オフセット値が第2の閾値以上であること、または前記フレーム結束値が第3の閾値以上であることのうちの少なくとも1つにตอบสนองして、バッファと関連付けられるパラメータを更新するステップとをさらに備える、請求項6に記載の方法。

【請求項 19】

複数の packets を受信するステップと、

フレーム消去レートが第1の閾値を満たす、オフセット値が第2の閾値を満たす、およびフレーム結束値が第3の閾値を満たすとの決定にตอบสนองして、バッファと関連付けられるパラメータを更新するステップとをさらに備え、前記フレーム結束値が前記複数の packets に基づく、請求項6に記載の方法。

【請求項 20】

前記パラメータを更新するステップが、第1の値から第2の値に前記パラメータの値を変更するステップを備え、前記パラメータが更新された後で、前記フレーム消去レートが前記第1の閾値より小さいこと、前記オフセット値が前記第2の閾値より小さいこと、または前記フレーム結束値が前記第3の閾値より小さいことにตอบสนองして、前記パラメータの前記値を前記第2の値から前記第1の値に変更するステップとをさらに備える、請求項19に記載の方法。

【請求項 21】

プロセッサによって実行されると、前記プロセッサに動作を実行させる命令を記憶したコンピュータ可読記憶デバイスであって、前記動作が、

第1のデータと第2のデータとを含む packets をデジッタバッファにおいて処理するステップであって、前記第1のデータが、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデータの部分的なコピーを含み、前記第2のデータが、フレームの前記シーケンスのうちの第2のフレームに対応する、ステップと、

前記 packets を処理したことに応答して、

前記第2のデータと関連付けられる第2のフレーム受信タイムスタンプを生成するステップと、

前記第2のフレーム受信タイムスタンプに基づいて、および前記第1のフレームと前記第2のフレームとの間のフレームオフセットに基づいて、前記第1のデータと関連付けられる第1のフレーム受信タイムスタンプを生成するステップであって、前記第1のフレーム受信タイムスタンプが、前記第2のフレーム受信タイムスタンプによって示される第2の時間より早い第1の時間を示す、ステップと、

を備える、コンピュータ可読記憶デバイス。

10

20

30

40

50

## 【請求項 2 2】

前記第1のフレーム受信タイムスタンプを生成するステップが、  
フレームオフセットをフレーム再生時間長と乗算して時間長を得るステップと、  
前記第2のフレーム受信タイムスタンプから少なくとも前記時間長を差し引くステップ  
とを備える、請求項21に記載のコンピュータ可読記憶デバイス。

## 【請求項 2 3】

前記第1のフレームの第1の生成タイムスタンプを決定するステップと、  
前記第1の生成タイムスタンプがフレームの前記シーケンスの以前のフレームの生成タ  
イムスタンプを超えているとの決定にตอบสนองして、第1のバッファエントリに前記第1のフレ  
ーム受信タイムスタンプを記憶するステップとをさらに備え、前記以前のフレームがフレ  
ームの前記シーケンスの直近に処理されたフレームである、請求項21に記載のコンピュ  
ータ可読記憶デバイス。

10

## 【請求項 2 4】

前記第1のフレームの前記第1の生成タイムスタンプが前記以前のフレームの前記生成タ  
イムスタンプを超えていることに基づいて、前記以前のフレームがフレームの前記シーケ  
ンスにおいて前記第1のフレームより早いと決定するステップをさらに備える、請求項23  
に記載のコンピュータ可読記憶デバイス。

## 【請求項 2 5】

前記第1の生成タイムスタンプを生成するステップが、  
前記第1のフレームと前記第2のフレームとの間のフレームオフセットをフレームごとの  
サンプルの数と乗算して第1の値を生成するステップと、  
前記第2のフレームの第2の生成タイムスタンプから前記第1の値を差し引くステップと  
を備え、前記パケットが前記第2の生成タイムスタンプの指示を含む、請求項23に記載の  
コンピュータ可読記憶デバイス。

20

## 【請求項 2 6】

前記パケットが、フレームオフセットの指示、フレームごとのサンプルの数、または両  
方を含む、請求項21に記載のコンピュータ可読記憶デバイス。

## 【請求項 2 7】

前記動作がさらに、前記第1のフレームが前記デジッタバッファに存在しないまたはバ  
ッファタイムラインの値と無関係であるとの決定にตอบสนองして、第1のバッファエントリに  
前記第1のフレーム受信タイムスタンプを記憶するステップを備える、請求項21に記載の  
コンピュータ可読記憶デバイス。

30

## 【請求項 2 8】

前記第1のデータおよび前記第1のフレーム受信タイムスタンプに基づいて第1の出力を  
生成するステップと、

前記第2のデータおよび前記第2のフレーム受信タイムスタンプに基づいて第2の出力を  
生成するステップであって、前記第2のフレーム受信タイムスタンプが、クロック出力に  
基づき生成される、ステップと

をさらに備える、請求項21に記載のコンピュータ可読記憶デバイス。

## 【請求項 2 9】

第1のデータと第2のデータとを含むパケットを受信するための手段であって、前記第1  
のデータが、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデー  
タの部分的なコピーを含み、前記第2のデータが、フレームの前記シーケンスのうちの第2  
のフレームに対応する、手段と、

バッファ管理を実行するための手段であって、

前記パケットを受信したことにตอบสนองして前記第2のデータと関連付けられる第2のフレ  
ーム受信タイムスタンプを生成するための手段と、

前記第2のフレーム受信タイムスタンプに基づいて、および前記第1のフレームと前記  
第2のフレームとの間のフレームオフセットに基づいて、前記第1のデータと関連付けられ  
る第1のフレーム受信タイムスタンプを生成するための手段であって、前記第1のフレーム

40

50

受信タイムスタンプが、前記第2のフレーム受信タイムスタンプによって示される第2の時間より早い第1の時間を示す、手段と、

を備える、前記バッファ管理を実行するための手段と  
を備える、装置。

【請求項 30】

前記受信するための前記手段および前記バッファ管理を実行するための前記手段が、通信デバイス、コンピュータ、音楽プレーヤ、ビデオプレーヤ、エンターテインメントユニット、ナビゲーションデバイス、携帯情報端末(PDA)、デコーダ、またはセットトップボックスのうちの少なくとも1つに組み込まれる、請求項29に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、その内容の全体が参照により組み込まれている、2015年8月18日に出願された「DE-JITTER BUFFER UPDATE」と題する米国特許出願第14/829,560号、2015年4月14日に出願された「DE-JITTER BUFFER UPDATE」と題する米国仮特許出願第62/147,386号、および2015年4月14日に出願された「BUFFER PARAMETER ADJUSTMENT」と題する米国仮特許出願第62/147,431号の利益を主張するものである。

【0002】

本開示は全般に、デジッタバッファに対する更新に関する。

【背景技術】

【0003】

技術の進歩は、より小さくより強力なコンピューティングデバイスをもたらしている。たとえば、現在、小型で、軽量で、ユーザが容易に持ち運べる、ポータブルワイヤレス電話、携帯情報端末(PDA)、およびページングデバイスなどの、ワイヤレスコンピューティングデバイスを含む、様々なポータブルパーソナルコンピューティングデバイスが存在する。より具体的には、携帯電話およびインターネットプロトコル(IP)電話などのポータブルワイヤレス電話は、ワイヤレスネットワークを通じて音声およびデータパケットを通信することができる。さらに、多くのそのようなワイヤレス電話には、内部に他のタイプのデバイスが組み込まれている。たとえば、ワイヤレス電話は、デジタルスチルカメラ、デジタルビデオカメラ、デジタルレコーダ、およびオーディオファイルプレーヤも含むことがある。また、そのようなワイヤレス電話は、ウェブブラウザアプリケーションなどの、インターネットにアクセスするために使用され得るソフトウェアアプリケーションを含む、実行可能な命令を処理することができる。したがって、これらのワイヤレス電話は、かなりのコンピューティング能力を含むことがある。

【0004】

ワイヤレス電話などの電子デバイスは、ネットワークを介してデータを送信して受信し得る。たとえば、オーディオデータは、回線交換ネットワーク(たとえば、公衆交換電話網(PSTN)、モバイル通信用グローバルシステム(GSM(登録商標))ネットワークなど)またはパケット交換ネットワーク(たとえば、ボイスオーバーインターネットプロトコル(VoIP)ネットワーク、ボイスオーバーロングタームエボリューション(VoLTE)ネットワークなど)を介して、送信および受信され得る。パケット交換ネットワークでは、オーディオフィームのシーケンスに対応するオーディオパケットは、ソースデバイスから宛先デバイスへ個々にルーティングされ得る。ネットワーク条件が原因で、オーディオパケットは順序が誤って到来することがある。宛先デバイスは、受信されたパケットをデジッタバッファに記憶することができ、受信されたパケットの順序が誤っている場合には受信されたパケットを並べ替えることができる。

【0005】

宛先デバイスは、受信されたパケットに基づいて、処理されたオーディオ信号を生成し得る。ソースデバイスによって送信されたある特定のパケットは、宛先デバイスによって

10

20

30

40

50

受信されないことがあり、またはエラーを伴って受信されることがある。デジッタバッファのバッファ深度が増大し、処理されるまでのパケットのバッファリング時間が増大するにつれて、宛先デバイスへ到達するための時間をより多くのパケットが有するようになるので、処理されるオーディオ信号の品質は改善することがある。しかしながら、バッファ深度を大きくすることは、エンドツーエンドの遅延を増やし、ユーザ体験に悪影響を及ぼすことがある。バッファ深度を低減することはエンドツーエンドの遅延を減らす、フレームのシーケンスのうちの一部のフレームに対応するパケットが時間内に宛先デバイスへ到達していないので、処理されたオーディオ信号がそれらのフレームに対するデータなしで生成される可能性を高める。

【発明の概要】

10

【課題を解決するための手段】

【0006】

ある特定の態様では、デバイスはデジッタバッファおよび分析器を含む。デジッタバッファは、第1のデータおよび第2のデータを含むパケットを受信するように構成される。第1のデータは、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデータの部分的なコピーを含む。第2のデータは、フレームのシーケンスのうちの第2のフレームに対応する。分析器は、パケットを受信したことに応答して、第1のデータと関連付けられる第1のフレーム受信タイムスタンプと、第2のデータと関連付けられる第2のフレーム受信タイムスタンプとを生成するように構成される。第1のフレーム受信タイムスタンプは、第2のフレーム受信タイムスタンプによって示される第2の時間より早い第1の時間を示す。

20

【0007】

別の態様では、方法は、デジッタバッファにおいてパケットを受信するステップを含む。パケットは、第1のデータおよび第2のデータを含む。第1のデータは、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデータの部分的なコピーを含む。第2のデータは、フレームのシーケンスのうちの第2のフレームに対応する。方法はまた、パケットを受信したことに応答して、第1のデータと関連付けられる第1のフレーム受信タイムスタンプを生成するステップと、第2のデータと関連付けられる第2のフレーム受信タイムスタンプを生成するステップとを含む。第1のフレーム受信タイムスタンプは、第2のフレーム受信タイムスタンプによって示される第2の時間より早い第1の時間を示す。

30

【0008】

別の態様では、コンピュータ可読記憶デバイスは、プロセッサによって実行されると、デジッタバッファにおいてパケットを受信するステップを含む動作をプロセッサに実行させる命令を記憶する。パケットは、第1のデータおよび第2のデータを含む。第1のデータは、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデータの部分的なコピーを含む。第2のデータは、フレームのシーケンスのうちの第2のフレームに対応する。動作はまた、パケットを受信したことに応答して、第1のデータと関連付けられる第1のフレーム受信タイムスタンプを生成するステップと、第2のデータと関連付けられる第2のフレーム受信タイムスタンプを生成するステップとを含む。第1のフレーム受信タイムスタンプは、第2のフレーム受信タイムスタンプによって示される第2の時間より早い第1の時間を示す。

40

【0009】

別の態様では、デバイスは、第1のデータおよび第2のデータを含むパケットを受信するための手段を含む。第1のデータは、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデータの部分的なコピーを含む。第2のデータは、フレームのシーケンスのうちの第2のフレームに対応する。デバイスはまた、パケットを受信したことに応答して第1のデータと関連付けられる第1のフレーム受信タイムスタンプを生成するための手段と、パケットを受信したことに応答して第2のデータと関連付けられる第2のフレーム受信タイムスタンプを生成するための手段とを含む。バッファ管理を実行するための手段を含む。第1のフレーム受信タイムスタンプは、第2のフレーム受信タイムスタンプによって

50

示される第2の時間より早い第1の時間を示す。

【0010】

本開示の他の態様、利点、および特徴は、以下のセクション、すなわち、図面の簡単な説明、発明を実施するための形態、および特許請求の範囲を含む本出願の検討後に明らかになるであろう。

【図面の簡単な説明】

【0011】

【図1】デジッタバッファを更新するように動作可能なシステムの特定の説明のための例のブロック図である。

【図2A】図1のシステムのある特定の例の動作を示す図である。

10

【図2B】図1のシステムのある特定の例の動作を示す図である。

【図3】図1のシステムのある特定の例の動作を示す図である。

【図4A】図1のシステムのある特定の例の動作を示す図である。

【図4B】図1のシステムのある特定の例の動作を示す図である。

【図5】図1のシステムのある特定の例の動作を示す図である。

【図6】図1のシステムのある特定の例の実装形態の動作を示す図である。

【図7】図1のシステムのある特定の例の実装形態の動作を示す図である。

【図8】図1のシステムのある特定の例の実装形態の動作を示す図である。

【図9】バッファパラメータを更新するように動作可能なシステムの特定の説明のための例のブロック図である。

20

【図10】デジッタバッファを更新するある特定の例の方法を示すフローチャートである。

【図11】デジッタバッファを更新するある特定の例の方法を示すフローチャートである。

【図12】バッファパラメータを調整するある特定の例の方法を示すフローチャートである。

【図13】デジッタバッファを更新するように動作可能なデバイスのある特定の説明のための例のブロック図である。

【発明を実施するための形態】

【0012】

本開示の特定の態様が、図面を参照して以下で説明される。説明において、共通の特徴が、図面を通じて共通の参照番号により指定される。

【0013】

30

デジッタバッファを更新するシステムおよび方法が開示される。ソースデバイスはパケットを宛先デバイスに送信し得る。たとえば、宛先デバイスにおける第1のユーザは、ソースデバイスにおける第2のユーザとの音声呼に参与していることがある。音声呼の間、第2のユーザは、ソースデバイスに結合された、またはソースデバイスに含まれるマイクロフォンに向かって話し得る。ソースデバイスは、第2のユーザの発話に対応する入力オーディオ信号を受信し得る。入力オーディオ信号は、オーディオフレームのシーケンスに対応し得る。たとえば、オーディオフレームのシーケンスは、第1のフレームおよび第2のフレームを含み得る。ソースデバイスは、オーディオフレームのシーケンスに対応するパケットを生成し得る。たとえば、ソースデバイスは、第1のフレームに対応する第1のフレームデータを含む第1のパケットを生成し得る。例示すると、第1のフレームデータは第1のフレームの第1の発話パラメータを含み得る。別の例として、ソースデバイスは、第2のフレームに対応する第2のフレームデータを含む第2のパケットを生成し得る。例示すると、第2のフレームデータは第2のフレームの第2の発話パラメータを含み得る。

40

【0014】

パケットは部分的な冗長性を有することがある。たとえば、第2のパケットは第1のパケットの部分的なコピー(たとえば、部分的なデータ)を含むことがある。例示すると、第2のパケットは第1の発話パラメータのサブセットを含むことがある。ある特定の例では、部分的なデータは、第1の発話パラメータと第2の発話パラメータとの間の違いを示すことがある。

【0015】

50



第1の packets は、第1のフレームに対応する第1の生成タイムスタンプ(たとえば、リアルタイムトランスポートプロトコル(RTP)タイムスタンプ)を含み得る。第2の packets は、第2のフレームに対応する第2の生成タイムスタンプ(たとえば、RTPタイムスタンプ)を含み得る。第1の生成タイムスタンプおよび第2の生成タイムスタンプは、オーディオフレームのシーケンスにおいて第1のフレームが第2のフレームより早く出現することを示し得る。ソースデバイスは、フレームのシーケンスに対応する packets (たとえば、第1の packets および第2の packets)を宛先デバイスに送信し得る。ソースデバイスは、ある特定の生成スタンプを符号化し、またはそれをある特定の packets に添付し得る。ソースデバイスは、宛先デバイスのクロックと同期していることもしていないこともあるソースデバイスのクロックに基づいて、その特定の生成タイムスタンプを決定し得る。加えて、または代わりに、ソースデバイスは、ソースデバイスによって符号化されるサンプルの数に基づいて、その特定の生成タイムスタンプを決定し得る。いくつかの例では、生成タイムスタンプはRTPの単位で表される。生成タイムスタンプは、フレームのシーケンスの順序を示し得る。

#### 【0016】

宛先デバイスは、ソースデバイスから1つまたは複数の packets を受信し得る。宛先デバイスは、受信された packets をバッファ(たとえば、デジタルバッファ)に記憶することができ、順序が誤った状態で packets が受信される場合には packets を並べ替えることができる。宛先デバイスは、特定のフレームに対応するデータが最小限の時間長(たとえば、再生遅延)の間バッファリングされた後で、デジタルバッファの中の各々の特定のフレームを処理し得る。特定のフレームを処理する際の遅延は、後続のフレームが処理されるべきであるときにその後続のフレームに対応する packets がデジタルバッファの中にある可能性を高め得る。

#### 【0017】

いくつかの事例では、宛先デバイスは、第1の packets を受信する前に第2の packets を受信することがある。第1の packets は、第2の packets の後で受信されることがあり、または受信されないことがある(たとえば、第1の packets は失われることがある)。宛先デバイスは、第2の packets が第1のフレームに対応する部分的なデータを含むとの決定に回答して、第2の packets をデュアルエンキューし得る。たとえば、宛先デバイスは、第2の packets を受信したことに回答して、第1のフレームがデジタルバッファに存在しないかどうかを決定し得る。宛先デバイスは、第1のフレームが存在しないとの決定に回答して、直近に処理されたフレームがフレームのシーケンスの中の第1のフレームより前にあるかどうかを決定し得る。宛先デバイスは、直近に処理されたフレームがフレームのシーケンスの中の第1のフレームより前にあるとの決定に回答して、第1のフレームに対応する第2の packets の第1のコピーをバッファリングし得る。宛先デバイスは、第2の packets を受信したことに回答して、第2のフレームに対応する第2の packets の第2のコピーをバッファリングし得る。

#### 【0018】

宛先デバイスは、デジタルバッファからのフレームを処理し得る。たとえば、宛先デバイスは、第1のフレームの第1の生成タイムスタンプに基づいて、処理されるべき次のフレームが第1のフレームであると決定し得る。第1の packets が第1のフレームの再生のために時間内に受信されなかった場合、宛先デバイスは第2の packets からの部分的なデータを処理し得る。宛先デバイスは、部分的なデータに基づいて、処理された発話信号を生成し得る。部分的なデータに基づいて生成される処理された発話信号は、第1のフレームに対応するいかなるデータも伴わずに生成される処理された発話信号よりも良好なオーディオ品質を有し得る。

#### 【0019】

宛先デバイスは、デジタルバッファの中の第1のバッファエントリを追加する(または更新する)ことによって、第2の packets の第1のコピーをバッファリングし得る。第1のバッファエントリは、第2の packets の第1のコピーを含み得る(または第1のコピーの位置を示

10

20

30

40

50

し得る)。第1のバッファエントリは、第1のフレームと関連付けられ得る。たとえば、第1のバッファエントリは、第1の生成タイムスタンプを示し得る。部分的なデータが第1のフレームに対応するバッファエントリ(たとえば、第1のバッファエントリ)へ記憶されるように第2の packets をデュアルエンキューすることで、宛先デバイスは、第1のフレームデータを含む第1の packets が失われた(または受信されていない)と決定することなく、より効率的に部分的なデータを処理することが可能になり得る。たとえば、デュアルエンキューを用いずに、宛先デバイスは、デジッタバッファの中の第1のフレームが処理されるべき次のフレームであるとき、その第1のフレームに対応するバッファエントリがないとの決定に回答して、第1の packets が失われたと決定し得る。宛先デバイスは、第1の packets が失われたとの決定に回答して、デジッタバッファが第1のフレームに対応する部分的なデータを含む別のバッファエントリを含むかどうかを決定し得る。宛先デバイスは、他のバッファエントリが部分的なデータを含むとの決定に回答して、他のバッファエントリを処理し得る。デュアルエンキューを用いて、宛先デバイスは、第1のフレームが処理されるべき次のフレームであるとの決定に回答して(たとえば、第1のフレームの部分的なコピーのためのデジッタバッファを探すことなく)、第1のバッファエントリを処理し得る。

10

#### 【0020】

宛先デバイスは、再生遅延(たとえば、最小再生遅延)があることの利点と、処理の遅延の欠点とのバランスをとるように、部分的なデータを処理し得る。再生遅延があることの利点には、後続のフレームと関連付けられる packets が、その後続のフレームが処理されるべきときにすでに受信されている確率を高めることがあり得る。処理の遅延の欠点には、エンドツーエンドの遅延の増大があり得る。

20

#### 【0021】

宛先デバイスは、フレームがデジッタバッファに存在しないとの決定に回答して、消去を再生し得る。たとえば、宛先デバイスは、第2の packets を受信する前に第1のフレームに対応する消去を再生し得る。宛先デバイスは、フレームのシーケンスの中の第1のフレームに後続するフレームが処理されていないとの決定に回答して、第1のフレームに対応する消去を再生した後で第1のフレーム(たとえば、部分的なデータ)を処理し得る。たとえば、宛先デバイスは、直近に処理されたフレームがフレームのシーケンスの中の第1のフレームより前にあるとの決定に回答して、第1のフレーム(たとえば、部分的なデータ)を処理し得る。

30

#### 【0022】

第1のフレームに対応する消去を再生した後で第1のフレーム(たとえば、部分的なデータ)を処理することで、暗黙的バッファ適応(IBA:implicit buffer adaptation)が可能になり得る。たとえば、デジッタバッファのバッファ深度は、第1のフレームに対応する消去を再生した後で第1のフレームの部分的なデータをバッファリングすることによって、(たとえば、1フレームだけ)暗黙的に増やされ得る。第1のフレームに対応する消去は時間 $t_0$ において再生されることがあり、第1のフレームに対応する部分的なデータは時間 $t_1$ において処理されることがある。時間 $t_1$ において処理されることになっていた後続のフレーム(たとえば、第2のフレーム)は、時間 $t_2$ において処理され得る。後続のフレームの処理遅延(たとえば、時間 $t_1$ から時間 $t_2$ への)は、第1のフレームの再生時間長(たとえば、20ミリ秒)に対応し得る。後続のフレーム(たとえば、第2のフレーム)に続くフレーム(たとえば、第3のフレーム)の処理も、処理遅延の分だけ遅延し得る。デジッタバッファは、処理遅延に対応する追加の時間の間、後続のフレーム(たとえば、第2のフレームおよび第3のフレーム)のフレームデータを記憶し得る。デジッタバッファは、暗黙的バッファ適応の結果として、より多くのフレームに対応するフレームデータを記憶し得る。たとえば、時間 $t_1$ におけるデジッタバッファの中の後続のフレーム(たとえば、第2のフレーム)を削除する代わりに、後続のフレームを時間 $t_1$ から時間 $t_2$ に移して維持することで、時間 $t_1$ の後でデジッタバッファに記憶されるフレームの数を(たとえば、1フレームだけ)増やすことができる。

40

50

## 【 0 0 2 3 】

ある特定の例では、デジッタバッファの深度は、入力オーディオ信号における後続の無音の期間の後で元に戻り得る(たとえば、1フレームだけ減り得る)。たとえば、宛先デバイスは、無音の期間の後の後続のフレームを、その後続のフレームと関連付けられるバッファリング期間が満了したとの決定にตอบสนองして(たとえば、追加の遅延を伴わずに)処理し得る。処理されたオーディオ信号は、入力オーディオ信号より短い無音を含み得る。

## 【 0 0 2 4 】

暗黙的バッファ適応の結果として、デジッタバッファの深度は、変化するネットワーク条件に動的に適応し得る。たとえば、デジッタバッファの深度は、ネットワーク条件(たとえば、混雑またはパケット喪失)によりフレーム(たとえば、フレームの部分的なデータ)がフレームに対応する消去の処理の後で受信されるようになるとき、(たとえば、1フレームだけ)増え得る。別の例として、デジッタバッファの深度は、後続の無音の後で元に戻り得る(たとえば、1フレームだけ減り得る)。

## 【 0 0 2 5 】

図1を参照すると、デジッタバッファを更新するように動作可能なシステムの特定の説明のための例が開示され、全体的に100と指定される。システム100は、ネットワーク190を介して1つまたは複数の他のデバイス(たとえば、ソースデバイス104)と通信している宛先デバイス102を含み得る。ソースデバイス104は、マイクロフォン146に結合されることがあり、またはそれと通信していることがある。たとえば、ソースデバイス104はマイクロフォン146を含み得る。宛先デバイス102は、スピーカ142に結合され得る。宛先デバイス102は、メモリ176に結合された、またはそれと通信している、分析器122を含み得る。宛先デバイス102は、受信機124、デジッタバッファ126、発話デコーダ160、またはこれらの組合せを含み得る。ある特定の実装形態では、分析器122は、デジッタバッファ126と発話デコーダ160との間のインターフェースモジュールとして動作するように構成され得る。メモリ176は、分析データ120を記憶するように構成され得る。分析データ120は、(たとえば、1フレーム以上の)フレームオフセット136、最後の生成タイムスタンプ138、再生遅延128、バッファタイムライン106、またはこれらの組合せを含み得る。

## 【 0 0 2 6 】

宛先デバイス102は、図1に示されるものよりも少数または多数の構成要素を含み得る。たとえば、宛先デバイス102は、1つまたは複数のプロセッサ、1つまたは複数のメモリユニット、または両方を含み得る。宛先デバイス102は、ネットワーク化されたコンピューティングシステムまたは分散型のコンピューティングシステムを含み得る。たとえば、メモリ176は、ネットワーク化されたメモリまたは分散型のメモリであり得る。ある特定の説明のための例では、宛先デバイス102は、通信デバイス、ヘッドセット、デコーダ、スマートフォン、セルラー電話、モバイル通信デバイス、ラップトップコンピュータ、コンピュータ、タブレット、携帯情報端末(PDA)、セットトップボックス、ビデオプレーヤ、エンターテインメントユニット、ディスプレイデバイス、テレビジョン、ゲームコンソール、音楽プレーヤ、ラジオ、デジタルビデオプレーヤ、デジタルビデオディスク(DVD)プレーヤ、チューナ、カメラ、ナビゲーションデバイス、またはこれらの組合せを含み得る。

## 【 0 0 2 7 】

動作の間、第1のユーザ152は第2のユーザ154との音声呼に関与していることがある。その音声呼のために、第1のユーザ152は宛先デバイス102を使用することがあり、第2のユーザ154はソースデバイス104を使用することがある。音声呼の間、第2のユーザ154は、ソースデバイス104に結合されたマイクロフォン146に向かって話し得る。入力オーディオ信号130は、第2のユーザ154によって話される語の一部分、語、または複数の語に対応し得る。入力オーディオ信号130は、オーディオフレームのシーケンスに対応し得る。たとえば、オーディオフレームのシーケンスは、第1のフレーム184および第2のフレーム186を含み得る。第1のフレーム184(たとえば、オーディオフレームN)は、オーディオフレームのシーケンスにおいて第2のフレーム186(たとえば、オーディオフレームN+3)より前にあり得

る。ソースデバイス104は、マイクロフォン146を介して、第2のユーザ154から入力オーディオ信号130を受信し得る。ある特定の実装形態では、マイクロフォン146は、オーディオ信号を捕捉することができ、アナログデジタルコンバータ(ADC)は、補足されたオーディオ信号を、アナログ波形から、デジタルオーディオサンプルからなるデジタル信号へと変換することができる。デジタルオーディオサンプルは、ボコーダ(音声エンコーダデコーダ)によって「圧縮」され得る。ボコーダのエンコーダは、デジタルオーディオサンプルを圧縮し、パケット(たとえば、第1のパケット132および第2のパケット134)のシーケンスを形成し得る。第1のパケット132は、第1のフレーム184に対応する第1のフレームデータ164を含み得る。第2のパケット134は、第2のフレーム186に対応する第2のフレームデータ166を含み得る。

10

**【0028】**

第1のパケット132は、第1のフレーム184の第1の生成タイムスタンプ(たとえば、RTPタイムスタンプ)を示し得る。第2のパケット134は、第2のフレーム186の第2の生成タイムスタンプ(たとえば、RTPタイムスタンプ)を示し得る。第1の生成タイムスタンプおよび第2の生成タイムスタンプは、オーディオフレームのシーケンスにおいて第1のフレーム184が第2のフレーム186より前にあることを示し得る。第1の生成タイムスタンプおよび第2の生成タイムスタンプは、ソースデバイス104によって生成され得る。ソースデバイス104は、(宛先デバイス102のクロックと同期していることもしていないこともある)ソースデバイス104のクロックに基づいて、その特定の生成タイムスタンプを決定し得る。加えて、または代わりに、ソースデバイス104は、ソースデバイス104によって符号化されるサンプルの数に基づいて、その特定の生成タイムスタンプを決定し得る。特定の例では、生成タイムスタンプはRTPの単位で表される。ソースデバイス104は、第1の生成タイムスタンプを符号化し、またはそれを第1のパケット132に添付し、第2の生成タイムスタンプを符号化し、またはそれを第2のパケット134に添付し得る。生成タイムスタンプは、フレームのシーケンスの順序を示し得る。

20

**【0029】**

いくつかの実装形態では、1つまたは複数のパケットは部分的な冗長性を有し得る。たとえば、後続のパケット(たとえば、第2のパケット134)は、前のオーディオフレーム(たとえば、オーディオフレームN)を再構築するために使用され得る冗長なデータ(たとえば、第1のフレームデータ164の部分的なコピー)を含み得る。たとえば、第2のパケット134は、第1のフレーム184の再現を、ただし場合によっては第1のフレームデータ164を使用する場合と比較してより低い分解能で可能にするために、第1のフレームデータ164の少なくとも一部分に対応する第1の部分的なデータ174を含み得る。ある特定の実装形態では、冗長なデータ(たとえば、第1の部分的なデータ174)は「重要な」発話フレームに対応し得る。たとえば、重要な発話フレームの喪失は、宛先デバイス102において生成される処理された発話信号において、重大な品質上の影響(たとえば、ユーザが気付くような品質上の影響)を引き起こし得る。

30

**【0030】**

ある特定の実装形態では、ソースデバイス104および宛先デバイス102は、定ビットレート(たとえば、毎秒13.2キロビットパーセカンド(kbps))チャンネル上で動作し得る。この実装形態では、主要データ(たとえば、第2のフレームデータ166)に対応する主要フレームビットレートは、冗長なデータ(たとえば、第1の部分的なデータ174)を収容するために(たとえば、9.6kbpsに)低減され得る。たとえば、定ビットレートの残りのビットレート(たとえば、3.6kbps)が冗長なデータに対応し得る。ある特定の例では、主要フレームビットレートの低減は、全体の発話品質への影響を少なくするように、入力オーディオ信号130の特性に応じてソースデバイス104において実行され得る。ソースデバイス104は、ネットワーク190を介してパケットのシーケンス(たとえば、第1のパケット132、第2のパケット134、または両方)を宛先デバイス102に送信し得る。

40

**【0031】**

宛先デバイス102は、パケットのシーケンスのうちの1つまたは複数のパケット(たとえ

50

ば、第1の packets 132、第2の packets 134、または両方)を受信し得る。分析器122は、デジッタバッファ126に、宛先デバイス102によって受信される packets (たとえば、第1の packets 132、第2の packets 134、または両方)を記憶し得る。たとえば、分析器122は、宛先デバイス102から packets を受信することがあり、デジッタバッファ126に packets を記憶することがある。ある特定の例では、 packets は、宛先デバイス102において順序が誤って受信されることがある。分析器122は、 packets の順序が誤っている場合、デジッタバッファ126の中の1つまたは複数の packets を並べ替えることができる。ソースデバイス104によって送信された packets のシーケンスの1つまたは複数の packets は、宛先デバイス102によって受信されないことがあり、またはエラーを伴って受信されることがある。たとえば、 packets (たとえば、第1の packets 132)は、 packets 喪失が原因で受信されないことがあり、または、ネットワーク条件が原因で受信機124によって部分的に受信されることがある。

10

#### 【 0 0 3 2 】

いくつかの状況では、宛先デバイス102は、第2の packets 134を受信して第1の packets 132を受信しないことがあり、または、第1の packets 132を受信する前に第2の packets 134を受信することがある。そのような状況では、宛先デバイス102は、受信機124から第2の packets 134を受信することができ、第2の packets 134を分析器122に提供することができる。分析器122は、第2の packets 134を受信すると、デジッタバッファ126に第2の packets 134をデュアルエンキューする(たとえば、第2の packets 134の複数のコピーを挿入する)ことができる。たとえば、分析器122は、第1のフレーム184がデジッタバッファ126に存在しない、および第2の packets 134が第1の部分的なデータを含むとの決定にตอบสนองして、第1のフレーム184のための第1のバッファエントリ114に第2の packets 134の第1のコピーを記憶し得る。分析器122はまた、第2のフレーム186のための第2のバッファエントリ118に、第2の packets 134の第2のコピーを記憶し得る。分析器122は、第1のバッファエントリ114がフレームのフレームデータの部分的なコピーに対応することを示すために、第1のバッファエントリ114の第1のフラグ144を特定の値(たとえば、1)に設定し得る。分析器122は、第2のバッファエントリ118がフレームのフレームデータの完全なコピーに対応することを示すために、第2のバッファエントリ118の第2のフラグ156を特定の値(たとえば、0)に設定し得る。

20

#### 【 0 0 3 3 】

ある特定の実装形態では、分析器122は、 packets を解析することによってフレームオフセット136(たとえば、部分的なフレームと対応する完全なフレームとの間の packets の数)を決定し得る。分析器122は、 packets がフレームのフレームデータの部分的なコピーを含むとの決定にตอบสนองして、 packets を解析してフレームオフセット136を決定し得る。たとえば、分析器122は、第2の packets 134を解析して、第1の部分的なデータ174の第1のフレームオフセット(たとえば、1)を決定することがあり、第1のフレームオフセットは第1の部分的なデータ174が第1のフレーム184に対応することを示す。加えて、または代わりに、第2のフレームデータ166は、フレームオフセット136を含むことがあり、またはフレームオフセット136を示すことがある。他の特定の实装形態では、フレームオフセット136は、2、3、または5などの、1より大きな任意の数であり得る。

30

40

#### 【 0 0 3 4 】

分析器122は、第2の packets 134によって示される第2のフレーム186の第2の生成タイムスタンプに基づいて、第2の生成タイムスタンプ158を決定し得る。分析器122は、第2のバッファエントリ118に第2の生成タイムスタンプ158を記憶し得る。分析器122は、第2の生成タイムスタンプ158および第1のフレームオフセット136に基づいて第1の生成タイムスタンプ148を決定し得る(たとえば、第1の生成タイムスタンプ148=(第2の生成タイムスタンプ158-(フレームオフセット136\*フレームごとのサンプル数))).分析器122は、第1のバッファエントリ114に第1の生成タイムスタンプ148を記憶し得る。ある特定の例では、第2の packets 134は、フレームオフセット136、フレームごとのサンプル数(たとえば、フレームごとに640個のサンプル)、または両方を示し得る。

50

## 【 0 0 3 5 】

分析器122は、第2のパケット134の受信タイムスタンプに基づいて、第2のフレーム186の第2の受信タイムスタンプ150を決定し得る。分析器122は、第2のバッファエントリ118に第2の受信タイムスタンプ150を記憶し得る。受信タイムスタンプ(フレーム受信タイムスタンプとも呼ばれ得る)は、フレームが利用可能である時間、またはフレームが宛先デバイス102によって受信される時間を示し得る。受信タイムスタンプはまた、過去の時間(たとえば、特定のフレームが利用可能であるべきであった時間、または受信されるべきであった時間)を示すように調整され得る。受信タイムスタンプは、(ソースデバイス104のクロックと同期していることもしていないこともある)宛先デバイス102のクロックに基づいて、宛先デバイス102によって生成され得る。加えて、または代わりに、受信タイムスタンプは、ソースデバイス104によって示されるサンプルの数(たとえば、それによって符号化されるサンプルの数)に基づき得る。

10

## 【 0 0 3 6 】

分析器122は、過去の時間(たとえば、特定のフレームが利用可能であるべきであった時間、または受信されるべきであった時間)を示すようにパケットの受信タイムスタンプを調整することによって、パケット(たとえば、部分的に冗長なパケット)の特定のフレーム(たとえば、部分的なフレーム)の受信タイムスタンプを生成し得る。たとえば、分析器122は、第2のパケット134の受信タイムスタンプに基づいて、かつフレームオフセット136に基づいて、第1のフレーム184の第1の受信タイムスタンプ140を生成し得る。分析器122は、フレームオフセット136をフレーム再生時間長と乗算することによって、特定の時間長を決定し得る。分析器122は、第2のパケット134の受信タイムスタンプから少なくとも特定の時間長を差し引くことによって、第1の受信タイムスタンプ140を生成し得る。例示すると、分析器122は、 $FRTS = PRTS - (F0 * FPD) - 1$ ミリ秒という式に基づいて第1の受信タイムスタンプ140を決定することがあり、ここで $FRTS$ =第1の受信タイムスタンプ140であり、 $PRTS$ =第2のパケット134の受信タイムスタンプであり、 $F0$ =フレームオフセット136であり、 $FPD$ =フレーム再生時間長である。ある特定の実装形態では、第2のパケット134は、フレームオフセット136(たとえば、3)、フレーム再生時間長(たとえば、20ミリ秒(ms))、または両方を示し得る。分析器122は、第1のバッファエントリ114に第1の受信タイムスタンプ140を記憶し得る。第1の受信タイムスタンプ140は、第2の受信タイムスタンプ150によって示される第2の時間より早い第1の時間を示し得る。

20

30

## 【 0 0 3 7 】

ある特定の例では、分析器122は、メモリ176の第1の位置に第2のパケット134の第1のコピーを記憶することがあり、メモリ176の第2の位置に第2のパケット134の第2のコピーを記憶することがある。第1のバッファエントリ114は、第1の位置を示し得る(たとえば、第1の位置への第1のポインタを含み得る)。第2のバッファエントリ118は、第2の位置を示し得る(たとえば、第2の位置への第2のポインタを含み得る)。

## 【 0 0 3 8 】

本明細書において説明されるように、分析器122は、第1の再生時間における第1のフレーム184に対応する第1のバッファエントリ114を処理し得る。分析器122は、第1の部分的なデータ174を発話デコーダ160に提供し得る。ある特定の実装形態では、分析器122は、発話デコーダ160から第1のフレーム184に対する要求を受信したことに応答して、第1の部分的なデータ174を発話デコーダ160に提供し得る。本明細書において説明されるように、分析器122は、第2の再生時間における第2のフレーム186に対応する第2のバッファエントリ118を処理し得る。分析器122は、第2のフレームデータ166を発話デコーダ160に提供し得る。ある特定の実装形態では、分析器122は、発話デコーダ160から第2のフレーム186に対する要求を受信したことに応答して、第2のフレームデータ166を発話デコーダ160に提供し得る。発話デコーダ160は、第1の部分的なデータ174および第2のフレームデータ166に基づいて、処理されたオーディオ信号116を生成し得る。たとえば、発話デコーダ160は、第1の部分的なデータ174に基づいて第1の生成されたフレーム194を生成することがあり、第2のフレーム166に基づいて第2のフレーム186を生成することがある。処理されたオー

40

50

ディオ信号116は、第1の生成されたフレーム194および第2のフレーム186に対応し得る。発話デコーダ160は、スピーカ142を介して、第1のユーザ152に処理されたオーディオ信号116を出力し得る。

【 0 0 3 9 】

ある特定の例では、第1のフレームデータ164は、第1のフレーム184の第1の発話パラメータに対応し得る。第1の部分的なデータ174は、第1の発話パラメータのサブセットを含み得る。別の特定の例では、第2のフレームデータ166は第2のフレーム186の第2の発話パラメータに対応することがあり、第1の部分的なデータ174は第1の発話パラメータと第2の発話パラメータとの間の差に対応することがある。この例では、分析器122は、第2の発話パラメータと第1の部分的なデータ174との合計に基づいて、第1の発話パラメータを生成し得る。分析器122は、第1の発話パラメータに基づいて、処理されたオーディオ信号116を生成し得る。処理されたオーディオ信号116は、第1の packets 132 がデジタルバッファ126において欠けているときであっても、第1の発話パラメータに基づいて生成され得ることを理解されたい。ある特定の例では、分析器122は、第2の packets 134、第1の部分的なデータ174、第2のフレームデータ166、第1の発話パラメータ、第2の発話パラメータ、またはこれらの組合せを、発話デコーダ160に提供することがあり、発話デコーダ160は、処理されたオーディオ信号116を生成することがある。失われた packets (たとえば、第1の packets 132) を補償するために第1の部分的なデータ174を使用して生成される処理されたオーディオ信号116は、第2のフレームデータ166だけにに基づいて生成される処理された発話信号よりも良好なオーディオ品質を有し得る。たとえば、処理されたオーディオ信号116は、第2のフレームデータ166に基づいて、かつ第1のフレームデータ164(または第1の部分的なデータ174)に基づかずに生成される処理された発話信号よりも、ユーザが気付くアーティファクトが少ないことがある。

【 0 0 4 0 】

ある特定の装置形態では、分析器122は、デジタルバッファ126の中の各々の特定のフレームを処理するのを、その特定のフレームに対応するバッファリング期間が満了するまで控え得る。分析器122は、特定のフレームの受信タイムスタンプおよび再生遅延128(たとえば、最小再生遅延)に基づいて、特定のフレームに対応するバッファリング期間を決定し得る。

【 0 0 4 1 】

第1のバッファエントリ114に第1の受信タイムスタンプ140を記憶することで、第1のフレーム184に対応する第1のバッファリング期間は、第2のフレーム186に対応する第2のバッファリング期間の満了の前に満了するようになり得る。たとえば、分析器122は、第1の受信タイムスタンプ140と再生遅延128の合計に基づいて、第1のバッファリング期間を決定し得る。分析器122は、第2の受信タイムスタンプ150と再生遅延128の合計に基づいて、第2のバッファリング期間を決定し得る。第1の受信タイムスタンプ140は、第2の受信タイムスタンプ150によって示される第2の時間より早い第1の時間を示し得る。分析器122は、第1のフレーム184に対応するバッファリング期間が満了したとの決定に応答して、第1の再生時間において第1のバッファエントリ114(たとえば、第1の部分的なデータ174)を処理することができるが、第2のフレーム186に対応するバッファリング期間はまだ継続中であり得る。たとえば、第1のフレーム184の第1の再生時間は、第2のフレーム186の第2の再生時間より前にあり得る。

【 0 0 4 2 】

分析器122は、再生されたフレームを追跡するために、かつ処理されるべき次のフレームを示すために、バッファタイムライン106を維持し得る。バッファタイムライン106は、バッファタイムライン106によって示される生成タイムスタンプの後で次にある特定の生成タイムスタンプを、処理されるべき次のフレームが有することを示し得る。たとえば、フレームが処理されデジタルバッファ126から除去されるにつれて、分析器122は、特定のフレームを処理したことに応答して特定のフレームの生成タイムスタンプ(たとえば、RTPタイムスタンプ)を示すように、バッファタイムライン106(たとえば、RTPタイムライン)

の値を更新または設定し得る。分析器122は、加えて、または代わりに、特定のフレームを処理したことに応答して特定のフレームの生成タイムスタンプを示すように、最後の生成タイムスタンプ138を更新し得る。分析器122は、直近に処理されたフレームの生成タイムスタンプを追跡するために、最後の生成タイムスタンプ138を維持し得る。

【0043】

入力オーディオ信号130におけるパケットの喪失または無音により、バッファタイムライン106は、最後の生成タイムスタンプ138と異なるようになり得る。たとえば、分析器122は、特定のフレームがデジッタバッファ126に存在しないとの決定に応答して、消去を再生し得る。分析器122は、消去を再生したことに応答して、フレーム再生時間長(たとえば、20ms)に基づいてバッファタイムライン106の値を更新(たとえば、インクリメント)し得る。この例では、バッファタイムライン106の値は最後の生成タイムスタンプ138を超えることがある。ある特定の実装形態で、分析器122は、特定のフレームがデジッタバッファ126に存在しないことを示す特定の入力を発話デコーダ160に提供することによって、消去を再生し得る。ある特定の実装形態では、分析器122は、対応するバッファリング期間(たとえば、第1のバッファリング期間)が満了した、および対応するフレーム(たとえば、第1のフレーム184)が処理されるべき次のフレームであることをバッファタイムライン106の値が示すとの決定に応答して、特定の再生時間において特定のバッファエントリ(たとえば、第1のバッファエントリ114)を処理し得る。

【0044】

ある特定の実装形態では、分析器122は、特定のフレームの生成タイムスタンプ(たとえば、RTPタイムスタンプ)が直近に処理されたフレームの生成タイムスタンプを超えているとの決定に応答して、その特定のフレームのための特定のパケットのコピーをデジッタバッファ126に記憶し得る。たとえば、分析器122は、第2の生成タイムスタンプ158が最後の生成タイムスタンプ138を超えているとの決定に応答して、第2のバッファエントリ118に第2のパケット134を含めることができ、これは、直近に処理されたフレームがオーディオフレームのシーケンスにおいて第2のフレーム186よりも前にあることを示す。別の例として、分析器122は、第1の生成タイムスタンプ148が最後の生成タイムスタンプ138を超えているとの決定に応答して、第1のバッファエントリ114に第2のパケット134を含めることができ、これは、直近に処理されたフレームがオーディオフレームのシーケンスにおいて第1のフレーム184よりも前にあることを示す。

【0045】

分析器122は、特定のフレームの生成タイムスタンプが最後の生成タイムスタンプ138を超えると、バッファタイムライン106の値とは無関係に、その特定のフレームのための特定のパケットのコピーを記憶するように、デジッタバッファ126を更新し得る。たとえば、第1の生成タイムスタンプ148は、バッファタイムライン106の値以下であることがあり、これは、分析器122が第1のフレーム184に対応する消去を再生した(たとえば、第1のフレーム184の再生時間の後で第1のフレームデータ164または第1の部分的なデータ174がデジッタバッファ126において受信されたので、第1のフレーム184の代わりに消去が再生された)ことを示す。この例では、分析器122は、第2のパケット134の第1のコピーを第1のバッファエントリ114に記憶するように、デジッタバッファ126を更新し得る。分析器122は、第2のパケット134の第1のコピーが第1のバッファエントリ114に記憶されるべきである、および第1のフレーム184に対応する消去が再生されたとの決定に応答して、第1のフレーム184が処理されるべき次のフレームであることを示すように、バッファタイムライン106を更新し得る。この例では、消去の後で第1のバッファエントリ114を処理することは、後続のフレームの処理にさらなる遅延を加えることがあり、それは、図2A~図5および図7を参照してさらに説明されるように、第1のフレーム184に対応する消去がすでに再生されているからである。さらなる遅延は、より多数のフレームをバッファリングするためにデジッタバッファ126のバッファ深度を(たとえば、1フレームだけ)暗黙的に大きくさせることによって、暗黙的バッファ適応をもたらし得る。ある特定の実装形態では、バッファ深度は、入力オーディオ信号130における後続の無音の期間の後で元に戻り得る(たと

10

20

30

40

50



えば、1フレームだけ減り得る)。たとえば、分析器122は、無音の期間の後の後続のフレームを、その後続のフレームと関連付けられるバッファリング期間が満了したとの決定に  
10 応答して(たとえば、追加の遅延を伴わずに)処理し得る。処理されたオーディオ信号116は、入力オーディオ信号130より短い無音を含み得る。暗黙的バッファ適応の結果として、デジッタバッファ126の深度は、変化するネットワーク条件に動的に適応し得る。たとえば、デジッタバッファの深度は、ネットワーク条件(たとえば、混雑またはパケット喪失)によりフレーム(たとえば、フレームの部分的なデータ)がフレームに対応する消去の  
処理の後で受信されるようになるとき、(たとえば、1フレームだけ)増え得る。別の例として、デジッタバッファの深度は、後続の無音の後で元に戻り得る(たとえば、1フレーム  
だけ減少し得る)。

#### 【0046】

ある特定の実装形態では、分析器122は、第2のパケット134を受信した後で、かつ第1の  
バッファエントリ114を処理する前に、第1のパケット132を受信し得る。分析器122は、第  
1のパケット132を受信したことと、第1のパケット132によって示される第1の生成タイム  
スタンプが最後の生成タイムスタンプ138を超えているとの決定したこととに  
20 応答して、第1のバッファエントリ114を更新し得る。たとえば、分析器122は、第2のパケット134の  
第1のコピーを第1のバッファエントリ114から除去し得る。ある特定の実装形態では、分析器122は、メモリ176からの第1のバッファエントリ114に対応する第2のパケット134の  
第1のコピーを削除する(または削除のためにマークする)ことがある。分析器122は、第1の  
バッファエントリ114に第1のパケット132を記憶し得る。ある特定の実装形態では、分析  
器122は、メモリ176の中のある特定の位置に第1のパケット132を記憶することができ、第  
1のバッファエントリ114の中の特定の位置(たとえば、そのアドレス)を示すように第1の  
バッファエントリ114を更新することができる。分析器122は、第1のバッファエントリ114  
がフレームのフレームデータの完全なコピーに対応することを示すために、第1のフラグ1  
44を特定の値(たとえば、0)に更新し得る。分析器122は、第1のフレーム184の第1の生成  
タイムスタンプを示すように第1の生成タイムスタンプ148を更新し得る。分析器122は、  
第1のパケット132の受信タイムスタンプを示すように第1の受信タイムスタンプ140を更新  
し得る。分析器122(または受信機124)は、第1のパケット132を受信したことに応答して、  
第1のパケット132の受信タイムスタンプを決定し得る。分析器122は、第1のバッファエン  
トリ114および第2のバッファエントリ118を処理し得る。発話デコーダ160は、第1のフレ  
ームデータ164および第2のフレームデータ166に基づいて、処理された発話信号を生成し  
30 得る。

#### 【0047】

システム100は、フレームのフレームデータ(たとえば、部分的なコピー)がそのフレ  
ームに対応する消去を再生した後でデジッタバッファにおいてバッファリングされることを  
可能にすることによって、暗黙的バッファ適応を可能にし得る。フレームに対応する消去  
を再生した後でフレームデータを処理することは、フレームシーケンスの後続のフレーム  
の処理を遅らせることがあり、デジッタバッファによるより多数のフレームのバッファリ  
ングをもたらす。デジッタバッファの深度は、消去を再生した後でフレームデータを処理  
40 することによって、暗黙的に増やされ得る。再生遅延の強制によって引き起こされ得る処理  
されたオーディオ信号の中断は、あるバッファエントリが失われたパケットのための部  
分的なデータを記憶しているときにそのバッファエントリに対応する第1の受信タイムス  
タンプを生成することによって回避され得る。第1の受信タイムスタンプは、部分的なデ  
ータを含んでいたパケットの第2の受信タイムスタンプよりも前の第1の時間を示し得る。

#### 【0048】

システム100によって提供される1つの具体的な利点は、ある特定のフレームのフレーム  
データがその特定のフレームに対応する消去を再生した後で受信されるときに、暗黙的  
バッファ適応を可能にすることである。暗黙的バッファ適応はバッファ深度を増やし、これ  
により、ソースデバイス104と宛先デバイス102との間のエンドツーエンドの遅延の永続的  
な増大を伴うことなく、特定のフレームがその特定のフレームに対応する消去を再生した  
50

後で処理されることを可能にする。したがって、暗黙的バッファ適応は、エンドツーエンドの遅延の一時的な増大しか伴わずに、処理されたオーディオ信号116の品質を改善する。

【 0 0 4 9 】

図2A~図5は、フレームのフレームデータの部分的なコピーがフレームに対応する消去を再生した後で受信されるとき、図1のシステム100の動作の非限定的な例を示す。図2A~図3は、無音に続く発話データを含むアクティブフレームにフレームが対応するときの、システム100の動作を示す。たとえば、アクティブフレームは、無音の期間に後続するトークスパートに対応し得る。図4A~図5は、発話データのアクティブフレームのシーケンスの中の発話データを含むアクティブフレームにフレームが対応するときの、システム100の動作を示す。

10

【 0 0 5 0 】

図2Aを参照すると、図1のシステムの特定の例の動作を示す図表が開示され、全体的に200と指定される。図表200は、処理されるフレームのフレーム数を時間の関数として示す。図表200はまた、時間 $t_0$ におけるデジッタバッファ126の状態を示す。図2Aに示されるように、時間 $t_0$ において、フレームオーディオシーケンスのフレームのサブセットは、図1の分析器122によってすでに処理されている可能性がある。たとえば、フレーム1~3はアクティブな発話(たとえば、トークスパート)に対応し得る。分析器122は、フレーム1~3をすでに処理している可能性がある。最後の生成タイムスタンプ138は、直近に処理されたフレーム(たとえば、フレーム3)の生成タイムスタンプ(たとえば、 $3 \times 640 = 1920$ )を示し得る。フレーム4~9は、トークスパートとトークスパートの間の無音の期間に対応し得る。分析器122は、無音フレームを再生することがあり、データフレームを再生しないことがあり、またはフレーム4~9に対応する消去を再生することがある。時間 $t_0$ におけるバッファタイムライン106の値は、フレーム9に対応する生成タイムスタンプ(たとえば、5760)を示し得る。デジッタバッファ126は、フレーム11に対応するバッファエントリ202およびフレーム12に対応するバッファエントリ204を含み得る。バッファエントリ202は、フレーム11の生成タイムスタンプ212(たとえば、7040)を含み得る。バッファエントリ204は、フレーム12の生成タイムスタンプ214(たとえば、7680)を含み得る。バッファエントリ202は、バッファエントリ202がフレームのフレームデータの完全なコピーに対応することを示すための第1の値を有するフラグ222を含み得る。バッファエントリ204は、バッファエントリ204がフレームのフレームデータの完全なコピーに対応することを示すための第1の値を有するフラグ224を含み得る。

20

30

【 0 0 5 1 】

時間 $t_0$ において、分析器122は、バッファタイムライン106(たとえば、5760)に基づいて、処理されるべき次のフレームがフレーム10であると決定し得る。たとえば、分析器122は、バッファタイムライン106の値に基づいて、次のフレーム(たとえば、図1の第1のフレーム184)の次の生成タイムスタンプ(たとえば、6400)を決定し得る(たとえば、次の生成タイムスタンプ=バッファタイムライン106+フレームごとのサンプルの数(たとえば、640))。分析器122は、次のフレーム(たとえば、フレーム10)がデジッタバッファ126に存在しないと決定し得る。たとえば、分析器122は、次の生成タイムスタンプに対応するバッファエントリがデジッタバッファ126において欠けていると決定し得る。分析器122は、フレーム10(たとえば、第1のフレーム184)がデジッタバッファ126に存在しないと決定に応答して、(たとえば、無音フレームを再生すること、またはデータフレームを再生しないことによって)1フレーム長だけ無音の期間を延ばすことができる。分析器122は、フレーム10に対応する生成タイムスタンプを示すように、バッファタイムライン106の値を更新し得る。たとえば、分析器122は、デジッタバッファ126に存在しないフレーム10のために無音の期間を延ばすと、フレームごとのサンプルの数(たとえば、640)だけ、バッファタイムライン106の値をインクリメントすることができる。

40

【 0 0 5 2 】

図2Bを参照すると、図2Aの例を続ける図表が開示され、全体的に250と指定される。時

50

間t1において、フレーム10に対応する無音の期間を延ばした後で、分析器122は、フレーム10のフレームデータ(たとえば、フレームデータの部分的なコピー)を含むパケット(たとえば、図1の第2のパケット134)を受信し得る。パケットはまた、別のフレーム(たとえば、フレーム13)のフレームデータの完全なコピーを含み得る。第2のパケット134は、フレーム13の生成タイムスタンプ(たとえば、8320)を示し得る。分析器122は、フレーム13(たとえば、第2のフレーム186)に対応する第2のバッファエントリ118にパケット(たとえば、第2のパケット134)のコピーを記憶し得る。分析器122は、第2のバッファエントリ118に第2の生成タイムスタンプ158を記憶し得る。第2の生成タイムスタンプ158は、フレーム13の生成タイムスタンプ(たとえば、8320)を示し得る。分析器122は、第2のバッファエントリ118がフレームのフレームデータの完全なコピーに対応することを示す特定の値(たとえば、1)に第2のフラグ156を設定し得る。

10

【 0 0 5 3 】

図1を参照して説明されたように、分析器122は、図1の第2の生成タイムスタンプ158およびフレームオフセット136に基づいて、フレーム10(たとえば、6400)のための第1の生成タイムスタンプ148を決定し得る。分析器122は、第1の生成タイムスタンプ148が最後の生成タイムスタンプ138(たとえば、1920)を超えているとの決定にตอบสนองして、フレーム10(たとえば、第1のフレーム184)に対応するパケット(たとえば、第2のパケット134)の第1のコピーを第1のバッファエントリ114に記憶し得る。分析器122は、バッファタイムライン106の値とは無関係に、第1のバッファエントリ114にパケット(たとえば、第2のパケット134)の第1のコピーを記憶し得る。たとえば、分析器122は、無音の期間が1フレーム長だけ延ばされたことを示すバッファタイムライン106の値(たとえば、6400)とは無関係に、第1のバッファエントリ114にパケット(たとえば、第2のパケット134)の第1のコピーを記憶し得る。ある特定の実装形態では、分析器122は、バッファタイムライン106の値が第1の生成タイムスタンプ148以上であるとの決定にตอบสนองして、第1のフレーム184(たとえば、フレーム10)が処理されるべき次のフレームであることを示すように、バッファタイムライン106を更新し得る。たとえば、分析器122は、バッファタイムライン106の値が第1の生成タイムスタンプ148以上であるとの決定にตอบสนองして、生成タイムスタンプ(たとえば、5760)を示すようにバッファタイムライン106を更新し得る。生成タイムスタンプ(たとえば、5760)は、第1の生成タイムスタンプ148(たとえば、6400)の前にあり得る。第1のフレーム184(たとえば、フレーム10)が処理されるべき次のフレームであることを示すようにバッファタイムライン106を更新することで、分析器122は、無音の期間を1フレーム長だけ延ばした後で第1のバッファエントリ114を処理することが可能になり得る。

20

30

【 0 0 5 4 】

図3を参照すると、図2Aの図2Bの例を続ける図表が開示され、全体的に300と指定される。分析器122は、処理されるべき次のフレームが第1のフレーム184(たとえば、フレーム10)であると決定し得る。たとえば、分析器122は、バッファタイムライン106の値(たとえば、5760)およびフレームごとのサンプルの数(たとえば、640)に基づいて、次の生成タイムスタンプ(たとえば、6400)を決定し得る。分析器122は、第1の生成タイムスタンプ148に基づいて、デジッタバッファ126が次の生成タイムスタンプ(たとえば、6400)に対応する第1のバッファエントリ114を含むと決定し得る。分析器122は、時間t2において、第1のバッファエントリ114を処理し得る。たとえば、図1を参照して説明されるように、分析器122は、フレームデータ(たとえば、第1の部分的なデータ174)、パケット(たとえば、第2のパケット134)、第1のフレーム184(たとえば、フレーム10)に対応する第1の発話パラメータ、またはこれらの組合せを、発話デコーダ160に提供し得る。分析器122は、第1のフレーム184(たとえば、フレーム10)を処理したことにตอบสนองして、第1の生成タイムスタンプ148を示すようにバッファタイムライン106を更新し得る。たとえば、分析器122は、フレームごとのサンプルの数(たとえば、640)だけバッファタイムライン106をインクリメントし得る。

40

【 0 0 5 5 】

図1を参照して説明されたように、発話デコーダ160は処理されたオーディオ信号116を

50

生成し得る。1フレーム長だけ無音の期間を延ばした後で第1のバッファエントリ114を処理することで、バッファエントリ202、バッファエントリ204、および第2のバッファエントリ118は、第1のフレーム184の再生時間長に対応する追加の時間の間デジッタバッファ126に記憶されるようになり得る。たとえば、第1のバッファエントリ114は、第1のバッファエントリ114がデジッタバッファ126に挿入されていなければバッファエントリ202が処理されていたであろう時間に、処理され得る。バッファエントリ202の処理は、第1のフレーム184(たとえば、フレーム10)の再生時間長(たとえば、20ミリ秒)だけ遅れることがあり、後続のフレームの処理遅延をもたらす。処理遅延の間、宛先デバイス102は追加の packets を受信することがあり、デジッタバッファ126は追加の packets に対応するフレームをバッファリングすることがある。デジッタバッファ126によってバッファリングされるフレームの数は増えることがあり、デジッタバッファ126の深度を暗黙的に大きくすることによる暗黙的バッファ適応をもたらす。処理遅延は、後続のフレームが処理されるべき時間にそれらの後続のフレームに対応するデータをデジッタバッファ126が含む確率を高め得る。

10

#### 【0056】

図4Aを参照すると、図1のシステムの別の特定の実装形態の動作を示す図表が開示され、全体的に400と指定される。図表400は、フレーム1~9がアクティブな発話(たとえば、トークスパート)に対応するという点で、図3の図表300と異なる。最後の生成タイムスタンプ138は、直近に処理されたフレーム(たとえば、フレーム9)の生成タイムスタンプ(たとえば、5760)を示し得る。最後の生成タイムスタンプ138は、バッファタイムライン106と同じ値を示し得る。図1および図2Aを参照して説明されたように、分析器122は、時間t0において、フレーム10がデジッタバッファ126に存在しないとの決定に回答して、消去を再生し得る。

20

#### 【0057】

図4Bを参照すると、図4Aの例を続ける図表が開示され、全体的に402と指定される。図1および図2Bを参照して説明されたように、分析器122は、時間t1において、第1の生成タイムスタンプ148が最後の生成タイムスタンプ138を超えているとの決定に回答して、第1のバッファエントリ114に packets (たとえば、第2の packets 134)の第1のコピーを記憶し得る。分析器122は、第1のフレーム184(たとえば、フレーム10)に対応する消去が再生されたことをバッファタイムライン106の値が示すかどうかとは無関係に、第1のバッファエントリ114に packets (たとえば、第2の packets 134)の第1のコピーを記憶し得る。ある特定の实装形態では、分析器122は、オーディオフレームのシーケンスの中の第1のフレーム184に後続するフレームが処理されるべき次のフレームであることをバッファタイムライン106の値が示すとの決定に回答して、第1のフレーム184が処理されるべき次のフレームであることを示すように、バッファタイムライン106を更新し得る。たとえば、分析器122は、バッファタイムライン106の値(たとえば、6400)が第1の生成タイムスタンプ148以上であるとの決定に回答して、第1の生成タイムスタンプ148(たとえば、6400)の前にある生成タイムスタンプ(たとえば、5760)を示すように、バッファタイムライン106を更新し得る。

30

#### 【0058】

図5を参照すると、図4Aの図4Bの例を続ける図表が開示され、全体的に500と指定される。分析器122は、時間t2において、第1のバッファエントリ114を処理し得る。図1および図3を参照して説明されたように、第1のフレーム184(たとえば、フレーム10)に対応する消去を再生した後で第1のバッファエントリ114を処理することは、暗黙的バッファ適応を可能にし得る。

40

#### 【0059】

図6を参照すると、図1のシステムの特定の实装形態の動作を示す図表が開示され、全体的に600と指定される。図表600は、様々な実装形態においていつバッファエントリが処理されるかの比較を示す。図表600は、4つの場合の比較を示す表652を含む。図表600は、4つの場合の各々において、どの packets がデジッタバッファ126に存在するか(影付きのパターンを使用して図6に示されている)、およびバッファエントリがいつ処理されるかを示

50

す、表650を含む。図表600はまた、デジッタバッファ126のバッファエントリを示す。

【 0 0 6 0 】

表652に示されるように、場合1は、特定のフレームに対応する部分的なデータと別のフレームに対応する完全なフレームデータとを含むパケットがデジッタバッファ126において一度エンキューされる(たとえば、デュアルエンキューではない)実装形態に対応する。場合1において、部分的なデータは再生遅延(たとえば、最小再生遅延)とは無関係に処理される。完全なフレームデータは、場合1~4における再生遅延(たとえば、40ミリ秒(ms))に基づいて処理され得る。表652に示されるように、受信タイムスタンプ調整(Rx\_ts調整)は場合1において適用可能ではない。

【 0 0 6 1 】

表652に示されるように、場合2~4は、特定のフレームに対応する部分的なデータと別のフレームに対応する完全なフレームデータとを含むパケットが二度エンキューされる(たとえば、デュアルエンキュー)実装形態に対応する。たとえば、パケットの第1のコピーは特定のフレームに対応する第1のバッファエントリにエンキューされることがあり、パケットの第2のコピーは他のフレームに対応する第2のバッファエントリにエンキューされることがある。

【 0 0 6 2 】

表652に示されるように、場合2~3では受信タイムスタンプ調整(Rx\_ts調整)がない。たとえば、第1のバッファエントリはパケットの受信タイムスタンプを記憶し得る。第2のバッファエントリもパケットの受信タイムスタンプを記憶し得る。場合2では、部分的なデータが再生遅延(たとえば、最小再生遅延)に基づいて処理され、場合3では、部分的なデータが再生遅延とは無関係に処理される。

【 0 0 6 3 】

表652に示されるように、場合4では受信タイムスタンプ調整(Rx\_ts調整)がある。たとえば、パケットの受信タイムスタンプによって示される第2の時間より早い第1の時間を示す、第1の受信タイムスタンプが生成され得る。第1のバッファエントリは、第1の受信タイムスタンプを記憶し得る。場合4において、部分的なデータは再生遅延(たとえば、最小再生遅延)に基づいて処理される。

【 0 0 6 4 】

表650において示されるように、時間 $t=80$ ミリ秒(ms)および時間 $t=100$ msにおいて、デジッタバッファ126は空である(たとえば、フレームのシーケンスに対応するデータを含まない)。分析器122は、データなしの指示を発話デコーダ160に提供し得る。

【 0 0 6 5 】

時間 $t=120$ msにおいて、デジッタバッファ126は、フレームのシーケンスの第2のフレーム(F2)に対応するバッファエントリ604を含み得る。バッファエントリ604は、第2のフレーム(F2)の完全なフレームデータを含むパケットの受信タイムスタンプに対応するタイムスタンプ(TS)624(たとえば、115ms)を記憶し得る。分析器122は、TS624(たとえば、115ms)および再生遅延(たとえば、40ms)に基づいて、バッファエントリ604に対応するバッファリング期間を決定し得る。たとえば、分析器122は、TS624によって示される第1の時間(たとえば、115ms)と再生遅延(たとえば、40ms)との合計に対応する時間(たとえば、155ms)において、バッファエントリ604に対応するバッファリング期間が満了すると決定し得る。場合1~4では、時間 $t=120$ msにおいて、分析器122は、バッファエントリ604に対応するバッファリング期間が満了していないとの決定にตอบสนองして、バッファエントリ604を処理すること(およびデキューすること)を控え得る。分析器122は、データなしの指示を発話デコーダ160に提供し得る。

【 0 0 6 6 】

時間 $t=140$ msにおいて、デジッタバッファ126は、第2のフレーム(F2)に対応するバッファエントリ604と、フレームのシーケンスの第4のフレーム(F4)に対応するバッファエントリ608とを含み得る。バッファエントリ608は、第4のフレーム(F4)の完全なフレームデータを含むパケットの受信タイムスタンプに対応するTS628(たとえば、138ms)を記憶し得る

10

20

30

40

50

。パケットはまた、フレームのシーケンスの第1のフレーム(F1)の部分的なデータを含み得る。

【0067】

場合1~4では、時間 $t=140\text{ms}$ において、分析器122が、バッファエントリ604および608に対応するバッファリング期間が満了していないと決定し得る。たとえば、バッファエントリ604に対応する第1のバッファリング期間は時間 $t=155\text{ms}$ に満了することがあり、バッファエントリ608に対応する第2のバッファリング期間は時間 $t=178\text{ms}(138\text{ms}+40\text{ms})$ に満了することがある。分析器122は、バッファエントリ604および608に対応するバッファリング期間が満了していないとの決定にตอบสนองして、バッファエントリ604および608を処理するのを控え得る。

10

【0068】

場合1において、第4のフレーム(F4)の完全なフレームデータおよび第1のフレーム(F1)の部分的なデータを含むパケットは、デジッタバッファ126に一度エンキューされ得る。たとえば、デジッタバッファ126は、第4のフレーム(F4)に対応するバッファエントリ608を含むことがあり、第1のフレーム(F1)に対応するバッファエントリを含まないことがある。場合1において、部分的なデータは再生遅延とは無関係に処理される。分析器122は、分析器122がフレームのシーケンスの中のある特定のフレーム(たとえば、F1)より前のフレームを処理しておらず、デジッタバッファ126がその特定のフレーム(たとえば、F1)に対応するバッファエントリを含まないとき、その特定のフレーム(たとえば、F1)が処理されるべき次のフレームであると決定することが可能ではないことがある。その特定のフレーム(たとえば、F1)が処理されるべき次のフレームとして示されないので、分析器122は、その特定のフレーム(たとえば、F1)に対応する完全なフレームデータが失われていると決定することが可能ではないことがあり、他のフレーム(たとえば、F4)に対応するバッファエントリの中のその特定のフレーム(たとえば、F1)に対応する部分的なデータを探さないことがある。したがって、分析器122は、データなしの指示を発話デコーダ160に提供し得る。場合1では、デジッタバッファ126が第1のフレーム(F1)に対応する部分的なデータを含むとしても、第1のフレーム(F1)に対応するデュアルエンキューおよびバッファエントリがないので、分析器122は消去を再生する(たとえば、データなしの指示を提供する)ことがある。

20

【0069】

30

場合2~4では、第4のフレーム(F4)の完全なフレームデータおよび第1のフレーム(F1)の部分的なデータを含むパケットは、デジッタバッファ126にデュアルエンキューされ得る。たとえば、デジッタバッファ126は、第1のフレーム(F1)に対応するバッファエントリ602と、第4のフレーム(F4)に対応するバッファエントリ608とを含み得る。バッファエントリ602はTS622を記憶し得る。受信タイムスタンプ調整がない場合2~3では、TS622は、バッファエントリ608のTS628(たとえば、138ms)(たとえば、パケットの受信タイムスタンプ)に対応し得る。

【0070】

場合2において、部分的なデータは再生遅延(たとえば、40ms)に基づいて処理される。分析器122は、TS622(たとえば、138ms)および再生遅延(たとえば、40ms)に基づいて、バッファエントリ602と関連付けられるバッファリング期間が時間 $t=178\text{ms}$ に満了すると決定し得る。したがって、図1の分析器122は、バッファエントリ602と関連付けられるバッファリング期間が満了していないとの決定にตอบสนองして、時間 $t=140\text{ms}$ においてバッファエントリ602を処理するのを控え得る。分析器122は、データなしの指示を発話デコーダ160に提供し得る。場合2では、デジッタバッファ126が第1のフレーム(F1)に対応する部分的なデータを含むとしても、部分的なデータが再生遅延に基づいて処理され、受信タイムスタンプ調整がないので、分析器122は消去を再生する(たとえば、データなしの指示を提供する)ことがある。

40

【0071】

場合3において、部分的なデータは再生遅延(たとえば、40ms)とは無関係に処理される

50

。バッファエントリ602と関連付けられるバッファリング期間は、TS622(たとえば、138ms)および再生遅延(たとえば、40ms)に基づいて、時間 $t=178\text{ms}$ に満了しているはずである。図1の分析器122がバッファリング期間に基づいてバッファエントリ602を処理していたとすると、バッファリング期間が満了していないので、分析器122は、時間 $t=140\text{ms}$ においてバッファエントリ602を処理するのを控えるであろう。しかしながら、場合3では、分析器122は、バッファエントリ602が部分的なデータを記憶しているので、バッファリング期間とは無関係に、時間 $t=140\text{ms}$ においてバッファエントリ602を処理し得る。

【0072】

受信タイムスタンプ調整がある場合4では、TS622は、TS628によって示される第2の時間(たとえば、138ms)より早い第1の時間(たとえば、77ms)を示し得る。たとえば、図1を参照して説明されるように、分析器122は、TS628、図1のフレームオフセット136(たとえば、3フレーム)、および再生時間長(たとえば、20ms)に基づいて、TS622を生成し得る。場合4において、部分的なデータは再生遅延(たとえば、40ms)に基づいて処理される。分析器122は、TS622(たとえば、77ms)および再生遅延(たとえば、40ms)に基づいて、バッファエントリ602と関連付けられるバッファリング期間が時間 $t=117\text{ms}$ に満了したと決定し得る。図1の分析器122は、バッファエントリ602と関連付けられるバッファリング期間が満了したとの決定にตอบสนองして、時間 $t=140\text{ms}$ においてバッファエントリ602を処理し得る。

【0073】

場合3~4において、分析器122は、時間 $t=140\text{ms}$ においてバッファエントリ602を処理した後で、バッファエントリ602に対応するデータを発話デコーダ160に提供し得る。図1を参照して説明されるように、発話デコーダ160は、バッファエントリ602に対応するデータに基づいて、図1の処理されたオーディオ信号116を生成し得る。

【0074】

時間 $t=160\text{ms}$ において、デジッタバッファ126は、第5のフレーム(F5)に対応するバッファエントリ610を含み得る。バッファエントリ610は、第5のフレーム(F5)の完全なフレームデータを含むパケットの受信タイムスタンプを示すTS630(たとえば、160ms)を含み得る。パケットは、第2のフレーム(F2)の部分的なデータを含み得る。分析器122は、デジッタバッファ126にパケットをデュアルエンキューするのを控え得る。たとえば、場合1では、デュアルエンキューはない。場合2~4では、分析器122は、デジッタバッファ126が第2のフレーム(F2)に対応するバッファエントリ(たとえば、バッファエントリ604)をすでに含んでいる、およびパケットが第2のフレーム(F2)に対応する部分的なデータを含んでいるとの決定にตอบสนองして、パケットをデュアルエンキューするのを控え得る。たとえば、分析器122は、第5のフレーム(F5)に対応するバッファエントリ610にパケットを記憶し得るが、第2のフレーム(F2)に対応するバッファエントリにパケットのコピーを記憶するのを控え得る。

【0075】

時間 $t=160\text{ms}$ において、場合1および3~4において、デジッタバッファ126はバッファエントリ602を含まない。たとえば、場合1において、分析器122は、本明細書において説明されるように、シングルエンキューが原因で、バッファエントリ602をデジッタバッファ126に追加していない可能性がある。場合3~4において、分析器122は、バッファエントリ602を処理したことに応答して、デジッタバッファ126からバッファエントリ602をデキュー(たとえば、削除)した可能性がある。場合1および3~4において、分析器122は、バッファエントリ604に対応するバッファリング期間が満了したとの決定にตอบสนองして、第2のフレーム(F2)に対応するバッファエントリ604を処理し得る。

【0076】

本明細書において説明されるように、場合2において、デジッタバッファ126は、デュアルエンキューが原因で、かつ、分析器122がバッファエントリ602をより早く(たとえば、時間 $t=140\text{ms}$ において)処理するのを控えたので、バッファエントリ602を含むことがある。加えて、分析器122は、バッファエントリ602に対応するバッファリング期間が時間 $t=160\text{ms}$ において満了していないと決定し得る。たとえば、場合2において、バッファエントリ

10

20

30

40

50

602に対応するバッファリング期間は、時間 $t=178\text{ms}$ (たとえば、 $138\text{ms}+40\text{ms}$ )に満了し得る。しかしながら、分析器122は、バッファエントリ604と関連付けられるバッファリング期間が時間 $t=160\text{ms}$ において満了したと決定し得る。たとえば、バッファエントリ604に対応するバッファリング期間は、時間 $t=155\text{ms}$ (たとえば、 $115\text{ms}+40\text{ms}$ )に満了し得る。分析器122は、バッファエントリ602に対応するバッファリング期間が満了していないとの決定に応答して、第1のフレーム(F1)に対応するバッファエントリ602を処理するのを控え得る。分析器122は、フレームのシーケンスにおいてバッファエントリ604と関連付けられる第2のフレーム(F2)よりも早く現れるフレーム(たとえば、第1のフレーム(F1))に対応するバッファエントリ602をデジッタバッファ126が含むとの決定に応答して、バッファエントリ604を処理するのを控え得る。図1を参照して説明されたように、分析器122は、それぞれの生成タイムスタンプに基づいて、フレームのシーケンスにおいて第1のフレーム(F1)が第2のフレーム(F2)より早く現れると決定し得る。したがって、分析器122は消去を再生し得る。たとえば、分析器122は、データなしの通知を発話デコーダ160に送信し得る。

#### 【0077】

時間 $t=180\text{ms}$ において、デジッタバッファ126は、第6のフレーム(F6)に対応するバッファエントリ612を含み得る。バッファエントリ612は、第6のフレーム(F6)の完全なフレームデータを含むパケットの受信タイムスタンプを示すTS632(たとえば、 $178\text{ms}$ )を含み得る。パケットはまた、第3のフレーム(F3)の部分的なデータを含み得る。場合1では、シングルエンキューがある(すなわちデュアルエンキューがない)ので、デジッタバッファ126は、第3のフレーム(F3)に対応するバッファエントリ606を含まないことがある。場合2~4では、分析器122は、第3のフレーム(F3)がデジッタバッファ126に存在しないとの決定に応答して、デジッタバッファ126にパケットをデュアルエンキューし得る。たとえば、分析器122は、第3のフレーム(F3)に対応するバッファエントリ606にパケットの第1のコピーを記憶することができ、第6のフレーム(F6)に対応するバッファエントリ612にパケットの第2のコピーを記憶することができる。バッファエントリ606はTS626を含み得る。受信タイムスタンプ調整がない場合2~3では、TS626はTS632と同じであり得る。たとえば、TS626はパケットの受信タイムスタンプ(たとえば、 $178\text{ms}$ )を示し得る。受信タイムスタンプ調整がある場合4では、TS626は、TS632によって示される第2の時間(たとえば、 $178\text{ms}$ )より早い第1の時間(たとえば、 $117\text{ms}$ )を示し得る。たとえば、第1の時間は、パケットの受信タイムスタンプによって示される時間(たとえば、 $178\text{ms}$ )より早いことがある。

#### 【0078】

時間 $t=180\text{ms}$ において、場合1では、図1を参照して説明されたように、分析器122は、図1のバッファタイムライン106に基づいて、処理されるべき次のフレームが第3のフレーム(F3)であると決定し得る。たとえば、分析器122は、第2のフレーム(F2)の生成タイムスタンプを示すようにバッファタイムライン106を更新し得る。分析器122は、処理されるべき次のフレームが、バッファタイムライン106によって示される生成タイムスタンプの後で次にある特定の生成タイムスタンプを有する第3のフレーム(F3)であると決定し得る。分析器122は、第3のフレーム(F3)が処理されるべき次のフレームである、およびデジッタバッファ126が第3のフレーム(F3)に対応するバッファエントリを含まないとの決定に応答して、第3のフレーム(F3)に対応する完全なフレームデータが失われていると決定し得る。分析器122は、第3のフレーム(F3)に対応する完全なフレームデータが失われているとの決定に応答して、別のバッファエントリが第3のフレーム(F3)に対応する部分的なデータを含むかどうかを決定し得る。たとえば、分析器122は、第6のフレーム(F6)に対応するバッファエントリ612が第3のフレーム(F3)に対応する部分的なデータを含むと決定し得る。分析器122は、時間 $t=180\text{ms}$ において、再生遅延とは無関係に第3のフレーム(F3)に対応する部分的なデータを処理し得る。

#### 【0079】

場合2では、時間 $t=180\text{ms}$ において、分析器122は、バッファエントリ602に対応する再生遅延期間が満了したとの決定に応答して、第1のフレーム(F1)に対応するバッファエントリ602を処理し得る。



## 【 0 0 8 0 】

場合3~4では、時間 $t=180\text{ms}$ において、分析器122は、第3のフレーム(F3)に対応するバッファエントリ606を処理し得る。たとえば、場合3では、分析器122は、再生遅延とは無関係に第3のフレーム(F3)が処理されるべき次のフレームであるとの決定に応答して、バッファエントリ606を処理し得る。場合4では、分析器122は、第3のフレーム(F3)が処理されるべき次のフレームである、およびバッファエントリ606に対応する再生遅延期間が満了したとの決定に応答して、バッファエントリ606を処理し得る。たとえば、バッファエントリ606に対応する再生遅延期間は、時間 $t=157\text{ms}$ (たとえば、 $117\text{ms}+40\text{ms}$ )に満了し得る。

## 【 0 0 8 1 】

時間 $t=200\text{ms}$ において、場合1および3~4では、分析器122は、第4のフレーム(F4)が処理されるべき次のフレームである、および第4のフレーム(F4)に対応する再生遅延期間が満了したとの決定に応答して、第4のフレーム(F4)に対応するバッファエントリ608を処理し得る。場合2では、分析器122は、図1を参照して説明されたように、第2のフレーム(F2)が処理されるべき次のフレームである、およびバッファエントリ604に対応する再生遅延期間が満了したとの決定に応答して、第2のフレーム(F2)に対応するバッファエントリ604を処理し得る。

## 【 0 0 8 2 】

時間 $t=220\text{ms}$ において、場合1および3~4では、分析器122は、第5のフレーム(F5)が処理されるべき次のフレームである、および第5のフレーム(F5)に対応する再生遅延期間が満了したとの決定に応答して、第5のフレーム(F5)に対応するバッファエントリ610を処理し得る。場合2では、分析器122は、図1を参照して説明されたように、第3のフレーム(F3)が処理されるべき次のフレームである、およびバッファエントリ606に対応する再生遅延期間が満了したとの決定に応答して、第3のフレーム(F3)に対応するバッファエントリ606を処理し得る。

## 【 0 0 8 3 】

図表600は、時間 $t=140\text{ms}$ において、場合1を参照して説明されたように、フレームのシーケンスの初期フレームの部分的なデータがデジッタバッファ126にバッファリングされているが、初期フレームの完全なフレームデータが失われていると分析器122が決定しなかったときに、(場合1のような)シングルエンキューが、それらの初期フレームに対応するフレームデータを伴わない処理されたオーディオ信号の生成をもたらし得ることを示す。

## 【 0 0 8 4 】

図表600は、デュアルエンキューがあり、受信時間調整がない場合に、時間 $t=140\text{ms}$ および時間 $t=160\text{ms}$ において、場合2を参照して説明されたように、特定のフレームの部分的なデータがデジッタバッファ126にバッファリングされているが、その部分的なデータに対応するバッファエントリと関連付けられるバッファリング期間が満了していないとき、再生遅延に基づいて部分的なデータを処理することが、フレームを処理するときに必要な消去および遅延をもたらし得ることを示す。

## 【 0 0 8 5 】

図表600は、図6の例では、場合3~4を参照して説明されたように、デュアルエンキューがあり受信時間調整がない状態で再生遅延とは無関係に部分的なデータを処理することが、いくつかの状況では、デュアルエンキューがあり受信時間調整がある状態で再生遅延に基づいて部分的なデータを処理する場合に匹敵する性能を有し得ることを示す。たとえば図8を参照して説明されるように、デュアルエンキューがあり受信時間調整がない状態で再生遅延とは無関係に部分的なデータを処理することは、他の状況においてデュアルエンキューがあり受信時間調整がある状態で(たとえば、高ジッタの場合)再生遅延に基づいて部分的なデータを処理することと比較して、異なる結果を有し得る。

## 【 0 0 8 6 】

図7を参照すると、図1のシステムの特定の実装形態の動作を示す図表が示され、全体的

10

20

30

40

50

に700と指定される。図表700は、様々な実装形態においていつバッファエントリが処理されるかの比較を示す。図表700は、2つの場合の比較を示す表752を含む。図表700は、2つの場合(図6の場合1および4)の各々において、どのパッケージがデジッタバッファ126に存在するか(影付きのパターンを使用して図7に示されている)、およびバッファエントリがいつ処理されるかを示す、表750を含む。図表700は、デジッタバッファ126のバッファエントリを示す。

【 0 0 8 7 】

図7のデジッタバッファ126のバッファエントリは、第6のフレーム(F6)に対応するバッファエントリ612のTS632が、図6におけるTS632によって示される時間(たとえば、178ms)より後の図7における第1の時間(たとえば、183ms)を示すという点で、図6のデジッタバッファ126のバッファエントリと異なる。たとえば、第6のフレーム(F6)の完全なフレームデータを含むパッケージは、図6を参照して説明された例におけるパッケージの受信タイムスタンプより後の、図7を参照して説明された例における第1の受信タイムスタンプと関連付けられ得る。パッケージは、第3のフレーム(F3)に対応する部分的なデータを含み得る。

【 0 0 8 8 】

場合4では、図1を参照して説明されるように、分析器122は、TS632(たとえば、183ms)、図1のフレームオフセット136、および再生時間長(たとえば、20ms)に基づいて、第3のフレーム(F3)に対応するTS626(たとえば、122ms)を生成し得る。図7のTS626は、図6のTS626(たとえば、117ms)によって示される第2の時間より後の第1の時間(たとえば、122ms)を示し得る。

【 0 0 8 9 】

時間 $t=100\text{ms}$ から時間 $t=160\text{ms}$ において、分析器122は図6を参照して説明されたように動作し得る。時間 $t=180\text{ms}$ において、分析器122は、第3のフレーム(F3)に対応するバッファエントリがデジッタバッファ126に存在しないと決定し得る。分析器122は、消去を再生し得る(たとえば、データなしを示す通知を図1の発話デコーダ160に提供し得る)。ある特定の例では、発話デコーダ160は、データなしを示す通知を受信したことに応答して、デジッタバッファ126が第3のフレーム(F3)に対応する部分的なデータを記憶するバッファエントリを有するかどうかを決定するように、分析器122に要求し得る。分析器122は、デジッタバッファ126のバッファエントリのいずれもが第3のフレーム(F3)に対応する部分的なデータを含まないとの決定に応答して、データなしを示す別の通知を発話デコーダ160に提供し得る。

【 0 0 9 0 】

時間 $t=200\text{ms}$ において、デジッタバッファ126は、第6のフレーム(F6)に対応するバッファエントリ612を含み得る。バッファエントリ612は、第3のフレーム(F3)に対応する部分的なデータを含み得る。場合4では、デジッタバッファ126は、第3のフレーム(F3)に対応するバッファエントリ606を含み得る。たとえば、図1を参照して説明されたように、分析器122は、直近に処理されたフレームがフレームのシーケンスの中の第3のフレーム(F3)の前にあるとの決定に応答して、第3のフレーム(F3)に対応する消去を再生した後でデジッタバッファ126にバッファエントリ606を含み得る。図1を参照して説明されたように、分析器122は、第3のフレーム(F3)が処理されるべき次のフレームであることを示すように、バッファタイムライン106を更新し得る。

【 0 0 9 1 】

時間 $t=200\text{ms}$ において、場合1では、分析器122は、デジッタバッファ126が第3のフレーム(F3)に対応するバッファエントリを含まないと決定し得る。分析器122は、第3のフレーム(F3)に対応する消去が再生されたとの決定に応答して、別のフレームに対応するバッファエントリが第3のフレーム(F3)に対応する部分的なデータを含むかどうかを決定するのを控え得る。分析器122は、第4のフレーム(F4)が処理されるべき次のフレームである、デジッタバッファ126が第3のフレーム(F3)に対応するバッファエントリを含まない、およびバッファエントリ608に対応するバッファリング期間が満了したとの決定に応答して、第4のフレーム(F4)に対応するバッファエントリ608を処理し得る。

## 【 0 0 9 2 】

時間 $t=200\text{ms}$ において、場合4では、分析器122は、第3のフレーム(F3)が処理されるべき次のフレームである、およびバッファエントリ606に対応するバッファリング期間が満了したとの決定に応答して、第3のフレーム(F3)に対応するバッファエントリ606を処理し得る。図1を参照して説明されたように、第3のフレームに対応する消去を再生した後で第3のフレーム(F3)に対応するバッファエントリ606を処理することは、暗黙的バッファ適応を引き起こし得る。

## 【 0 0 9 3 】

後続のフレーム(たとえば、F4、F5、F6など)の処理は、場合1と比較すると場合4では遅れることがあり、後続のフレームが処理されるべきときにそれらの後続のフレームに対応するデータがデジッタバッファ126の中にある確率が上がる。

10

## 【 0 0 9 4 】

図8を参照すると、図1のシステムの特定の実装形態の動作を示す図表が示され、全体的に800と指定される。図表800は、様々な実装形態においていつバッファエントリが処理されるかの比較を示す。図表800は、図6の場合3および4に対応する2つの場合の比較を示す表852を含む。図表800は、2つの場合の各々において、どのパッケージがデジッタバッファ126に存在するか(影付きのパターンを使用して図8に示されている)、およびバッファエントリがいつ処理されるかを示す、表850を含む。図表800は、デジッタバッファ126のバッファエントリを示す。

20

## 【 0 0 9 5 】

図8に示される例は、図8の再生遅延(たとえば、80ms)が図6の再生遅延(たとえば、40ms)より長いという点で、図6に示される例と異なる。より長い再生遅延(たとえば、80ms)は、宛先デバイス102が高ジッタを受けている状況に対応し得る。図8のデジッタバッファ126のバッファエントリは、図6のデジッタバッファ126のバッファエントリと異なる。

## 【 0 0 9 6 】

時間 $t=60\text{ms}$ において、デジッタバッファ126は空であり得る。図1の分析器122は、デジッタバッファ126が空であるとの決定に応答して、データなしの指示を図1の発話デコーダ160に提供し得る。

## 【 0 0 9 7 】

時間 $t=80\text{ms}$ において、デジッタバッファ126は、第2のフレーム(F2)に対応するバッファエントリ804を含み得る。バッファエントリ804のタイムスタンプ(TS)824は、第2のフレーム(F2)の完全なフレームデータを含むパッケージの受信タイムスタンプ(たとえば、75ms)を示し得る。図1の分析器122は、TS824(たとえば、75ms)および再生遅延(たとえば、80ms)に基づいて、バッファエントリ804のバッファリング期間が時間 $t=155\text{ms}$ に満了すると決定し得る。時間 $t=80\text{ms}$ において、分析器122は、バッファエントリ804に対応するバッファリング期間が満了していないとの決定に応答して、バッファエントリ804を処理することを控え得る。分析器122は、消去を再生し得る(たとえば、データなしの通知を図1の発話デコーダ160に提供し得る)。

30

## 【 0 0 9 8 】

時間 $t=100\text{ms}$ において、デジッタバッファ126は、第3のフレーム(F3)に対応するバッファエントリ806を含み得る。バッファエントリ806のTS826は、第3のフレーム(F3)の完全なフレームデータを含むパッケージの受信タイムスタンプ(たとえば、97ms)を示し得る。分析器122は、バッファエントリ806に対応するバッファリング期間が時間 $t=177\text{ms}$ (たとえば、97ms+80ms)に満了すると決定し得る。時間 $t=100\text{ms}$ において、分析器122は、それぞれのバッファリング期間が満了していないとの決定に応答して、バッファエントリ804および806を処理することを控え得る。分析器122は、消去を再生し得る(たとえば、データなしの通知を図1の発話デコーダ160に提供し得る)。

40

## 【 0 0 9 9 】

時間 $t=120\text{ms}$ において、デジッタバッファ126は、第4のフレーム(F4)に対応するバッファエントリ808を含み得る。バッファエントリ808のTS828は、第4のフレーム(F4)の完全な

50

フレームデータを含むパケットの受信タイムスタンプ(たとえば、103ms)を示し得る。分析器122は、バッファエントリ808に対応するバッファリング期間が時間 $t=183\text{ms}$ (たとえば、 $103\text{ms}+80\text{ms}$ )に満了すると決定し得る。パケットはまた、第1のフレーム(F1)に対応する部分的なデータを含み得る。したがって、分析器122はパケットをデュアルエンキューし得る。たとえば、デジッタバッファ126は、第1のフレーム(F1)に対応するバッファエントリ802を含み得る。バッファエントリ802はパケットの第1のコピーを記憶することがあり、バッファエントリ808はパケットの第2のコピーを記憶することがある。受信タイムスタンプ調整がない場合3では、バッファエントリ802のTS822は、パケットの受信タイムスタンプ(たとえば、103ms)を示し得る。受信タイムスタンプ調整がある場合4では、バッファエントリ802のTS822は、パケットの受信タイムスタンプによって示される第2の時間(たとえば、103ms)より早い第1の時間(たとえば、42ms)を示し得る。

10

【0100】

図1を参照して説明されたように、時間 $t=120\text{ms}$ において、場合3では、分析器122は、バッファエントリ802が第1のフレーム(F1)の部分的なデータを記憶している、および第1のフレーム(F1)が処理されるべき次のフレームであるとの決定に応答して、再生遅延とは無関係にバッファエントリ802を処理し得る。場合4では、分析器122は、バッファエントリ802のバッファリング期間が時間 $t=122\text{ms}$ (たとえば、 $42\text{ms}+80\text{ms}$ )に満了すると決定し得る。したがって、時間 $t=120\text{ms}$ において、場合4では、分析器122は、それぞれのバッファリング期間が満了していないとの決定に応答して、バッファエントリ802、804、806、および808を処理することを控え得る。したがって、分析器122は、消去を再生し得る(たとえば、データなしの通知を図1の発話デコーダ160に提供し得る)。

20

【0101】

時間 $t=140\text{ms}$ において、場合3では、分析器122は、それぞれのバッファリング期間が満了していないとの決定に応答して、バッファエントリ804、806、および808を処理することを控え得る。分析器122は、消去を再生し得る(たとえば、データなしの通知を図1の発話デコーダ160に提供し得る)。場合4では、分析器122は、バッファエントリ802に対応するバッファリング期間が満了したとの決定に応答して、バッファエントリ802を処理し得る。

【0102】

時間 $t=160\text{ms}$ において、デジッタバッファ126は、第5のフレーム(F5)に対応するバッファエントリ810を含み得る。バッファエントリ810のTS830は、第5のフレーム(F5)の完全なフレームデータを含むパケットの受信タイムスタンプ(たとえば、160ms)を示し得る。分析器122は、バッファエントリ810に対応するバッファリング期間が時間 $t=240\text{ms}$ (たとえば、 $160\text{ms}+80\text{ms}$ )に満了すると決定し得る。パケットはまた、第2のフレーム(F2)に対応する部分的なデータを含み得る。分析器122は、デジッタバッファ126が第2のフレーム(F2)に対応するバッファエントリ804をすでに有しているとの決定に応答して、パケットをデュアルエンキューするのを控え得る。たとえば、第2のフレーム(F2)の完全なフレームデータを含むパケットは、第2のフレーム(F2)の部分的なデータを含むパケットより前に受信されている可能性がある。分析器122は、完全なフレームデータを部分的なデータで置き換える(たとえば、上書きする)のを控え得る。

30

40

【0103】

図1を参照して説明されたように、時間 $t=160\text{ms}$ において、分析器122は、バッファエントリ804に対応するバッファリング期間が満了している、および第2のフレーム(F2)が処理されるべき次のフレームであるとの決定に応答して、第2のフレーム(F2)に対応するバッファエントリ804を処理し得る。

【0104】

図1を参照して説明されたように、時間 $t=180\text{ms}$ において、分析器122は、バッファエントリ806に対応するバッファリング期間が満了している、および第3のフレーム(F3)が処理されるべき次のフレームであるとの決定に応答して、第3のフレーム(F3)に対応するバッファエントリ806を処理し得る。

50

## 【 0 1 0 5 】

図1を参照して説明されたように、時間 $t=200\text{ms}$ において、分析器122は、バッファエントリ808に対応するバッファリング期間が満了している、および第4のフレーム(F4)が処理されるべき次のフレームであるとの決定にตอบสนองして、第4のフレーム(F4)に対応するバッファエントリ808を処理し得る。

## 【 0 1 0 6 】

図表800に示されるように、場合3に対応する実装形態は、高ジッタの場合の場合4に対応する実装形態とは異なる性能を有し得る。再生遅延とは無関係にフレームの部分的なデータを処理することは、後続のフレームが処理されるべき時間までに、その後続のフレームに対応するフレームデータが受信されている確率を下げる可能性がある。場合3では、処理された信号は、後続のフレームに対応するフレームデータなしで生成され得る。再生遅延に基づいて部分的なデータを処理することは、後続のフレームが処理されるべき時間までに、その後続のフレームに対応するフレームデータが受信されている確率を上げることがある。処理された信号は、後続のフレームに対応するフレームデータに基づいて生成され得る。場合4において生成される処理された信号は、場合3において生成される処理された信号より良好なオーディオ品質(たとえば、より少数のアーティファクト)を有し得る。

## 【 0 1 0 7 】

図9を参照すると、バッファパラメータを更新するように動作可能なシステムの特定の説明のための例が開示され、全体的に900と指定される。システム900は、ネットワーク944を介してソースデバイス104などの1つまたは複数の他のデバイスと通信している宛先デバイス902を含み得る。たとえば、ソースデバイス104は、オーディオデータを宛先デバイス902に通信するように構成され得る。例示すると、ソースデバイス104および宛先デバイス902は、ボイスオーバーインターネットプロトコル(VoIP)呼などの音声個に参与していることがある。加えて、または代わりに、宛先デバイス902はソースデバイス104にオーディオデータを通信するように構成されることがあり、ソースデバイス104はオーディオデータを受信して処理するように構成されることがある。宛先デバイス902は、図1の宛先デバイス102を含むことがあり、またはそれに対応することがある。

## 【 0 1 0 8 】

ソースデバイス104は、宛先デバイス902に通信されるべきオーディオデータに基づいて、パケットのシーケンスを生成し得る。たとえば、ソースデバイス104は、発話コード、デコーダ(図示されていない)、またはこれらの組合せを含み得る。発話コードは、発話信号をセグメントへと圧縮し、分割し、または圧縮および分割して、フレームを生成するように構成され得る。各セグメント(または「フレーム」)の時間長は、信号のスペクトルエンベロープが比較的静的なままになることが予想され得るのに十分短く選択され得る。説明のための非限定的な例として、1フレーム長は20ミリ秒であることがあり、これは8キロヘルツ(kHz)のサンプリングレートにおける160個のサンプルに対応する。他のフレーム長またはサンプリングレートが使用され得ることに留意されたい。

## 【 0 1 0 9 】

ボコードなどの発話エンコードは、複数のフレーム(たとえば、複数の符号化されたオーディオフレーム)を生成することができ、フレーム符号化シーケンス内の複数のフレームの順序に対応するシーケンス番号により各フレームをタグ付けすることができる。各フレームのシーケンス番号は、フレームのヘッダ(たとえば、ボコードパケットのヘッダ)に含まれ得る。いくつかの実装形態では、各フレームは同じサイズまたは同じ最大フレームサイズを有し得る。複数のフレームの各フレームはボコードパケットに含まれることがあり、またはそれに対応することがあり、各ボコードパケットはあるフレームシーケンス番号に対応することがある。たとえば、第1のフレーム(たとえば、第1のボコードパケット)は第1のフレームシーケンス番号に対応することがあり、第2のフレーム(たとえば、第2のボコードパケット)は第2のフレームシーケンス番号に対応することがある。第1のフレームが第2のフレームより前に生成されるとき、第1のフレームシーケンス番号は第2のフレームシーケンス番号より低いことがあり、これは、第1のフレームがフレームのシーケ

スにおいて第2のフレームより早いフレーム(たとえば、古いフレーム)であることを示す。ソースデバイス104は、フレームのシーケンスを第1のデータパケット940として宛先デバイス902に送信し得る。第1のデータパケット940は、ソースデバイス104によって送信される複数の別個のパケットを含み得る。

【0110】

ソースデバイス104は、ネットワーク944を介して第1のデータパケット940を宛先デバイス902に送信し得る。ネットワーク944は、ソースデバイス104と宛先デバイス902との間の通信を可能にし得る、サーバ(図示されていない)などの1つまたは複数のデバイスを含み得る。いくつかの実装形態では、RTPサーバなどのサーバが、ソースデバイス104によって送信されるフレーム(またはパケット)を、宛先デバイス902に与えられる束ねられたパケットへと束ねるように構成され得る。たとえば、サーバは、複数のボコーダパケットなどの複数のフレームを、束ねられたパケット(たとえば、マルチフレームパケット)へと束ねるように構成され得る。第1のデータパケット940は、図1の第1のパケット132、第2のパケット134、もしくは両方を含むことがあり、またはこれらに対応することがある。ネットワーク944は、図1のネットワーク190を含むことがあり、またはこれに対応することがある。

10

【0111】

いくつかの実装形態では、ボコーダは、以前に生成されたフレームに対応する冗長な情報を含むように、フレームのシーケンスのパケット(たとえば、ボコーダパケット)を生成し得る。たとえば、ボコーダは、第1のボコーダパケットを生成することができ、第1のボコーダパケットを生成した後で第2のボコーダパケットを生成することができる。第1のボコーダパケットは第1のフレームを含み得る。第2のボコーダパケットは、第2のフレームと第1のフレームの部分的なフレーム(たとえば、冗長な情報)とを含み得る。いくつかの実装形態では、部分的なフレームは発話フレームに対応することがあり、発話フレームの喪失は、宛先デバイス902において生成される処理された発話信号に対して重大な品質上の影響を引き起こす。

20

【0112】

第1のデータパケット940により表され得るパケットのシーケンスの第1の例が960に図示されている。第1の例960は、ある期間にわたってソースデバイス104により送信されるフレームのシーケンスを示す。たとえば、フレームのシーケンスは、第1の時間( $t_1$ )に送信される第1のパケット961と、第2の時間( $t_2$ )に送信される第2のパケット962と、第3の時間( $t_3$ )に送信される第3のパケット963と、第4の時間( $t_4$ )に送信される第4のパケット964とを含み得る。いくつかの実装形態では、第1の例960における時間の各々の間の(たとえば、 $t_1$ と $t_2$ の間などの)差は同じであることがあり、たとえば20msである。第1のパケット961は、発話信号の第1のセグメントに対応することがあり、(n)というシーケンス番号を有する第1のフレーム971を含むことがある。第2のパケット962は、第1のセグメントに後続する発話信号の第2のセグメントに対応し得る。第2のパケット962は、シーケンス番号(n+1)を有する第2のフレーム972を含み得る。

30

【0113】

第3のパケット963は、第2のセグメントに後続する第3のセグメントに対応し得る。第3のパケット963は、シーケンス番号(n+2)を有する第3のフレーム973を含み得る。第3のパケット963(たとえば、第3のフレーム973)はまた、第1のフレーム971に対応する部分的なフレーム975を含み得る。部分的なフレーム975は、第1のフレーム971と同じシーケンス番号であるシーケンス番号(n)を有し得る。ある特定のフレームに基づく特定のフレームは主要フレームと呼ばれることがあり、たとえば第1のフレーム971は主要フレームである。部分的なフレーム975が第1のフレーム971から2フレーム離れているので(第1のフレーム971のシーケンス番号(n)と、部分的なフレーム975を含む第3のフレーム973のシーケンス番号(n+2)との差によって示されるように)、部分的なフレームは2というオフセットを有する。部分的なフレーム975は第3のフレーム973の後にあるものとして示されているが、他の実装形態では、部分的なフレームは、第3のフレーム973より前にあることがあり、また

40

50

は第3のフレーム973に含まれることがある。第4のパケット964は、第3のセグメントに後続する第4のセグメントに対応し得る。第4のパケット964は、シーケンス番号(n+3)を有する第4のフレーム974を含み得る。

【 0 1 1 4 】

ネットワーク944によって(たとえば、ネットワーク944に含まれるサーバ(図示されていない)によって)生成される束ねられたパケットと関連付けられ得る、マルチフレームパケットのシーケンスの第2の例が、980において図示されている。第2の例980は、ある期間にわたってソースデバイス104によって送信されるフレームに基づいて、ネットワーク944によって作成されるマルチフレームパケットのシーケンス(たとえば、フレームのシーケンス)を示す。ソースデバイス104によって送信されるフレームは、ネットワーク944によって受信されるパケット940の第1のシーケンスを含むことがあり、またはそれに対応することがある。ネットワーク944は、パケット940の第1のシーケンスのフレームに基づいて、マルチフレームパケットのシーケンスを生成(または送信)し得る。たとえば、マルチフレームパケットのシーケンスは、第1の時間(t1)においてネットワーク944によって送信(または生成)される第1のマルチフレームパケット981と、第2の時間(t2)においてネットワーク944によって送信(または生成)される第2のマルチフレームパケット982と、第3の時間(t3)においてネットワーク944によって送信(または生成)される第3のマルチフレームパケット983と、第4の時間(t4)においてネットワーク944によって送信(または生成)される第4のマルチフレームパケット984とを含み得る。いくつかの実装形態では、第2の例980における時間の各々の間の(たとえば、t1とt2の間、t2とt3の間などの)差は同じであることがあり、たとえば40msである。第1のマルチフレームパケット981は、シーケンス番号(n)を有する第1のフレーム991とシーケンス番号(n+1)を有する第2のフレーム992とを含み得る。第2のマルチフレームパケット982は、シーケンス番号(n+2)を有する第3のフレーム993とシーケンス番号(n+3)を有する第4のフレーム994とを含み得る。

【 0 1 1 5 】

第3のマルチフレームパケット983は、シーケンス番号(n+4)を有する第5のフレーム995とシーケンス番号(n+5)を有する第6のフレーム996とを含み得る。第5のパケット995は、第2のフレーム992に対応する部分的なフレーム999を含み得る。部分的なフレーム999は、第2のフレーム992と同じシーケンス番号であるシーケンス番号(n+1)を有し得る。部分的なフレーム999が第2のフレーム992から3フレーム離れているので(第2のフレーム992のシーケンス番号(n+1)と、部分的なフレーム999を含む第5のフレーム995のシーケンス番号(n+4)との差によって示されるように)、部分的なフレーム999は3というオフセットを有する。第4のマルチフレームパケット984は、シーケンス番号(n+6)を有する第7のフレーム997とシーケンス番号(n+7)を有する第8のフレーム998とを含み得る。フレーム991~999の各々は、ソースデバイス104によって生成され得る。

【 0 1 1 6 】

ネットワーク944は、第2の例980において図示されるマルチフレームパケットのシーケンスを生成して、宛先デバイス902に送信し得る。たとえば、マルチフレームパケットのシーケンスは、通信チャネル(すなわち、有線またはワイヤレスのネットワーク接続)を通じて送信され得る。宛先デバイス902は、ネットワーク944によって生成(または送信)されるマルチフレームパケットの第1のデータパケット940またはシーケンスに基づき得る(またはそれに対応し得る)パケット942の第2のシーケンスを受信し得る。パケット942の第2のシーケンスは、宛先デバイス902によって受信される複数の別個のパケットを含み得る。ネットワーク944の条件が原因で、第1のデータパケット940(または第2の例980のマルチフレームパケットのシーケンス)は、第1のシーケンスと異なるパケットの第2のシーケンスを有するパケット942の第2のシーケンスとして、宛先デバイス902によって受信され得る。たとえば、パケット942の第2のシーケンスが宛先デバイス902において受信されるとき、第2のデータパケットのパケットは、順序が誤っていることがあり、破損している(たとえば、ビットエラーを有する)ことがあり、欠けていることがあり、またはこれらの組合せであることがある。別の例として、第2の例980のマルチフレームパケットのシーケ

スは、宛先デバイス902によって順序が誤った状態で受信されることがある。

【0117】

宛先デバイス902は、受信機924と、バッファ926と、バッファマネージャ910と、メモリ930と、発話デコーダ956とを含み得る。たとえば、受信機924は、図1の受信機124を含むことがあり、またはそれに対応することがある。受信機924は、1つまたは複数のボコードパケットまたは1つまたは複数のマルチフレームパケットなどの、1つまたは複数のパケットを受信するように構成され得る。パケットは、可変のネットワーク遅延が原因で、ランダムに、または予測不可能な到達時間に、受信機924に到達し得る。加えて、または代わりに、パケットは、順序が誤って(たとえば、パケット940の第1のシーケンスに関して順序が誤って、第2の例980のマルチフレームパケットのシーケンスに関して順序が誤って、またはこれらの組合せで)到達し得る。受信されたパケットは、バッファマネージャ910に提供され得る。

10

【0118】

バッファマネージャ910は、受信されたパケットを構文解析し、各パケットに含まれる1つまたは複数のフレームを識別する(たとえば、抽出する)ように構成され得る。たとえば、バッファマネージャ910は、図1の分析器122を含むことがあり、またはそれに対応することがある。バッファマネージャ910は、バッファ926に1つまたは複数の受信されたパケットのフレームを記憶する(たとえば、エンキューする)ように構成され得る。たとえば、バッファ926は、図1のデジッタバッファ126を含むことがあり、またはそれに対応することがある。デジッタバッファなどのバッファ926は、フレームがフレームのシーケンスに従って順序付けられることを可能にするために、かつパケット到達時間と関連付けられるジッタを減らすために、受信されたフレームを記憶するように構成され得る。バッファ926のサイズ(たとえば、最小再生遅延)を大きくすることは、遅れて失われるパケットの数を減らすことができるので、バッファ926の中のフレームの数を増やすことができる。バッファ926の中のフレームの数の増大は、オーディオ品質の向上につながり得る。

20

【0119】

いくつかの実装形態では、バッファ926にフレームを記憶するのではなく、バッファマネージャ910は、フレームが記憶されるメモリ930などのメモリの中の位置を指し示すポインタ(たとえば、識別子)をバッファ926に記憶し得る。たとえば、メモリ930は、図1のメモリ176を含むことがあり、またはそれに対応することがある。バッファマネージャ910は、受信されたパケットに含まれる部分的なフレームを特定する場合、バッファマネージャ910は、その部分的なフレームと同じシーケンス番号を有する対応するフレーム(たとえば、完全なフレーム)がバッファ926に含まれるかどうかを決定し得る。対応するフレームがバッファ926に記憶されている場合、バッファマネージャ910はその部分的なフレームを廃棄し得る。対応するフレームがバッファ926に記憶されていない場合、バッファマネージャ910は、受信されたパケットの第1のコピーおよび受信されたパケットの第2のコピーをバッファ926に挿入することによって、受信されたパケットをデュアルエンキューし得る。たとえば、バッファマネージャ910は、受信されたパケットの第1のコピーとしてその部分的なフレームをバッファ926に挿入し得る。受信されたパケットの第1のコピーは、その部分的なフレームのシーケンス番号に従ってバッファ926に挿入され得る。バッファマネージャ910は、受信されたパケットの第2のコピーとして、受信されたパケットに含まれる第2のフレームをバッファ926に挿入し得る。受信されたパケットの第2のコピーは、第2のフレームの第2のシーケンス番号に従ってバッファ926に挿入され得る。部分的なフレームがバッファ926に記憶された後で対応するフレームが受信される場合、バッファマネージャ910は部分的なフレームを完全なフレームで置き換え得る。

30

40

【0120】

バッファマネージャ910は、バッファ926から発話デコーダ956にフレームを提供するように構成され得る。たとえば、発話デコーダ956は、図1の発話デコーダ160を含むことがあり、またはそれに対応することがある。発話デコーダ956は、20msのブロックなどの、周期的な時間のブロックにおいて、フレームを受信し受信されたフレームを処理して、オ

50



オーディオ出力958を作成するように構成され得る。たとえば、オーディオ出力958は、図1の処理されたオーディオ信号116を含むことがあり、またはそれに対応することがある。連続的なオーディオ出力を作成するために、発話デコーダ956は、新しいフレームを定期的に(たとえば、一定の間隔で)受信するように構成され得る。たとえば、発話デコーダ956は、20msごとに新しいフレームを受信するように構成され得る。

【0121】

フレームを適切な順序(たとえば、順次的な順序)で発話デコーダ956に提供するために、バッファマネージャ910は、発話デコーダ956に提供されるべき次のフレームのシーケンス番号を示すカウンタ(図示されていない)を保持し得る。カウンタの値に基づいて、バッファマネージャ910は、カウンタの値と一致するシーケンス番号をフレームが有するとき、バッファ926からフレームをデキューし得る。デキューされたフレームは次いで、発話デコーダ956に提供され得る。フレームが発話デコーダ956に提供された後で、バッファマネージャ910は、カウンタの値をインクリメントし得る。

【0122】

フレームは、バッファパラメータ934に従ってバッファ926に保持され得る。たとえば、バッファパラメータ934は、図1の再生遅延128を含むことがあり、またはそれに対応することがある。バッファパラメータ934は、メモリ930に記憶され得る。いくつかの実装形態では、バッファパラメータ934は、再生遅延(たとえば、バッファ深度)または再生遅延パラメータと呼ばれ得る。再生遅延は、フレームがバッファ926にとどまるべき最小の時間の長さまたは最大の時間の長さを示し得る。いくつかの実装形態では、再生遅延は、バッファマネージャ910において受信されているフレームと、発話デコーダ956に提供されているフレームとの間の、最小の時間の長さまたは最大の時間の長さを示し得る。最小の時間の長さは最小再生遅延と呼ばれることがあり、最大の時間の長さは最大再生遅延と呼ばれることがある。宛先デバイス902の動作の間、バッファマネージャ910は、ジッタに基づいて、バッファパラメータ934の1つまたは複数の値を計算して更新し得る。たとえば、ジッタ値がより高いときにより高い再生遅延が使用されることがあり、ジッタ値がより低いときまたは0であるときにより低い再生遅延が使用されることがある。最小再生遅延の増大は、ソースデバイス104と宛先デバイス902との間のエンドツーエンド遅延の増大と引き換えに、遅れて失われるパケット(たとえば、発話デコーダ956により使用されるには到達が遅すぎるパケット)の数を減らし得ることに留意されたい。

【0123】

カウンタの値に基づいて発話デコーダ956に提供されるべきフレームが欠けている(たとえば、フレームがバッファ926の中にない)場合、バッファマネージャ910は暗黙的バッファ適応技法を使用し得る。暗黙的バッファ適応技法は、パケット再生の実効的な遅延を増やすために使用され得る。暗黙的バッファ適応技法は、バッファパラメータ934を明示的に増やさないが、パケットがバッファ926の中にとどまる時間を増やす。暗黙的バッファ適応技法を実施するために、バッファマネージャ910は、カウンタの値に対応する特定のフレームが欠けている場合に消去を再生する。しかしながら、特定のフレーム(または特定のフレームに対応する部分的なフレーム)が閾値の時間の長さ(たとえば、20ms)以内に到達する場合、特定のフレームは、発話デコーダ956に提供され、再生され得る。暗黙的バッファ適応技法は、消去またはその特定のフレームに後続するフレームの全体的な遅延を増やし得る。たとえば、フレームのシーケンスの第1の部分は、発話(たとえば、トークスパート)に対応するオーディオデータを含み得る。この例では、増大した全体的な遅延からの回復は、フレームのシーケンスの第1の部分(たとえば、トークスパート)に続く、たとえばフレームのシーケンスが発話デコーダ956によって受信され処理された後の、無音の期間に対応するフレームのシーケンスの第2の部分の間に行われ得る。

【0124】

いくつかの実装形態では、バッファパラメータ934(たとえば、最小再生遅延)は、説明のための非限定的な例として、約40msという値を有し得る。バッファパラメータが40msあるとき、宛先デバイス902は、部分的なフレームが3というオフセットを有しマルチフレ

10

20

30

40

50

ームパケットに含まれるとき、消去の増大を経験し得る。たとえば、第2の例980を参照すると、部分的なフレーム999(第2のフレーム992に対応する)が、第5のフレーム995が生成され第3のマルチフレームパケット983に含められた後で送信される。したがって、宛先デバイス902は、遅延がないと仮定すると、第2のフレーム992が失われた場合または破損している場合に部分的なフレーム999を受信するために、80ms(たとえば、2つの40msの間隔)待機しなければならない。40msというバッファパラメータ934では、部分的なフレーム999は、欠けている第2のフレーム992の代わりに使用されるための時間内に受信されないことがあり、消去が発生することがある。加えて、マルチフレームパケットが2つ以上のボコーダパケットを含むので、マルチフレームパケットの喪失は、各々が消去をもたらし得る2つのフレームの欠損をもたらす。したがって、マルチフレームパケットの使用および2より大きなオフセット値は、バッファパラメータ934が40msに維持される場合、宛先デバイス902の消去レートの増大という状況を生み出し得る。

#### 【0125】

消去レートの増大をもたらす状況特定のために、バッファマネージャ910は、結束検出器912、オフセット検出器914、および消去検出器916を含み得る。本明細書において説明されるように、バッファマネージャ910は、結束検出器912、オフセット検出器914、または消去検出器916によって生成される1つまたは複数のインジケータにตอบสนองして、バッファパラメータ934の値を調整するように構成され得る。結束検出器912、オフセット検出器914、および消去検出器916の各々は、メモリ930に記憶されている対応するデータを生成するように構成され得る。たとえば、結束検出器912は結束データ931を生成し得る。オフセット検出器914はオフセットデータ932を生成することができ、消去検出器916は消去データ933を生成することができる。オフセットデータ932は、図1のフレームオフセット136を含むことがあり、またはそれに対応することがある。

#### 【0126】

結束データ931は、マルチフレームパケットの中で一緒に束ねられるフレーム(たとえば、発話フレーム)の数と関連付けられる第1のデータを含み得る。たとえば、結束データ931は、ある期間の間に受信される複数のパケットに含まれるフレームの平均の数と関連付けられ得る。たとえば、第1の期間の間に、宛先デバイス902は、第1のパケットを、続いて第2のパケットを、続いて第3のパケットを受信し得る。結束データ931は、第1のパケットに含まれるフレームの第1の数を示す第1のデータエントリと、第2のパケットに含まれるフレームの第2の数を示す第2のデータエントリと、第3のパケットに含まれるフレームの第3の数を示す第3のデータエントリとを記憶し得る。各パケットに含まれるフレームの平均の数が増大すると、連続するパケットとパケットの間の平均の時間間隔も増大する。オフセットデータ932は、第2の期間の間に受信される各々の部分的なフレームのためのオフセット値を追跡し得る。消去データ933は、第3の期間の間に発話デコーダ956に提供される消去(たとえば、消去の再生、無音のフレーム、またはデータなしのフレーム)の数を追跡し得る。

#### 【0127】

結束検出器912、オフセット検出器914、および消去検出器916の各々は、対応するインジケータを生成するように構成され得る。たとえば、結束検出器912は結束インジケータ913を生成することができ、オフセット検出器914はオフセットインジケータ915を生成することができ、消去検出器916は消去インジケータ917を生成することができる。例示すると、結束検出器912は、結束データ931に基づいて結束インジケータ913を生成することができる。結束インジケータ913は、宛先デバイス902によって受信されるパケットごとのフレームの平均の数(たとえば、ボコーダパケットの平均の数)を示し得る。結束インジケータ913は、所定の数の直近に受信されたパケット(たとえば、5つの直近に受信されたパケット)に基づいて、または第1の期間の間に受信されたパケット(たとえば、直近の400msの間に受信されたパケット)に基づいて、パケットごとのフレームの平均の数を計算し得る。加えて、または代わりに、結束インジケータ913は、マルチフレームパケットが宛先デバイス902によって受信されているかどうかを示し得る。たとえば、結束インジケータ913は

、マルチフレームパケットが受信されていることを示すために論理値1を有することがあり、マルチフレームパケットが受信されていないことを示すために論理値0を有することがある。加えて、または代わりに、結束インジケータ913は、宛先デバイス902によって受信される連続するパケットとパケットとの間の平均の時間間隔を示し得る。

【0128】

オフセット検出器914は、オフセットデータ932に基づいてオフセットインジケータ915を生成することができる。オフセットインジケータ915は、第2の期間の間に受信される1つまたは複数の部分的なフレームの平均のオフセットを示し得る。消去検出器916は、消去データ933に基づいて消去インジケータ917を生成することができる。消去インジケータ917は、第3の期間の間の消去の平均の数を示し得る。加えて、または代わりに、消去インジケータ917は、消去の百分率としてフレーム消去レートを示し得る。消去の百分率は、第3の期間の間に発話デコーダ956へ提供される消去の数を第3の期間の間に発話デコーダ956へ提供されるべきフレームの総数で割ったものを示し得る。

【0129】

バッファマネージャ910は、インジケータ913、915、917のうちの1つまたは複数に対応する閾値と比較し得る。メモリ930は、結束閾値937、オフセット閾値938、および消去閾値939などの、1つまたは複数の閾値936を記憶し得る。結束検出器912は、結束インジケータ913を結束閾値937と比較し、オフセットインジケータ915をオフセット閾値938と比較し、消去インジケータ917を消去閾値939と比較し、またはこれらの組合せであり得る。

【0130】

説明のための非限定的な例として、結束閾値937は1.5という値を有し得る。結束インジケータ913が1.5という値以上である場合、結束検出器912は、マルチフレームパケット(たとえば、束ねられたパケット)が宛先デバイス902によって受信されていると決定し得る。別の例として、結束インジケータ913が連続するパケットとパケットの間の平均の時間間隔に対応するとき、結束閾値937は30msという値を有し得る。たとえば、ソースデバイス104がフレームをボコーダパケットとして生成して20msごとに送信するとき、遅延がないと仮定すると、フレームは20msごとに宛先デバイス902において受信され得る。しかしながら、ネットワーク944がマルチフレームパケットを生成するとき、マルチフレームパケットは、(たとえば、2つのボコーダパケットを生成するための時間の長さに対応する)少なくとも40msごとに生成(または送信)され得る。マルチフレームパケットは、遅延がないと仮定すると、40msごとに宛先デバイス902において受信され得る。したがって、結束インジケータ913(宛先デバイス902によって受信される連続するフレームとフレームとの間の平均の間隔を示す)が結束閾値937(たとえば、30ms)以上であるとき、バッファマネージャ910は、マルチフレームパケットが宛先デバイス902によって受信されていると決定し得る。非マルチフレームパケットが受信されるときと比較すると、マルチフレームパケットが受信されるとき、パケットはより稀に受信されるので、バッファマネージャ910は、再生遅延(たとえば、バッファ深度)を、したがってパケットがバッファ926の中にとどまる時間を増やすために、バッファパラメータ934を大きくすることができる。バッファ深度の増大により、バッファマネージャ910は、フレーム復元に利用可能なより多数の部分的なフレームを有することが可能になる。いくつかの実装形態では、説明のための非限定的な例として、宛先デバイス902は、第3世代パートナーシッププロジェクト(3GPP)規格(たとえば、Enhanced Voice Services(EVS) 3GPP規格)などの1つまたは複数の規格に従って、部分的なフレームを使用して、欠けているフレームまたは破損したフレームを復元することができる。

【0131】

説明のための非限定的な例として、オフセット閾値938は2.5という値を有し得る。たとえば、バッファパラメータ934が40msであり、(遅延がないと仮定して)2つのボコーダパケットを含むマルチフレームパケットが40msごとに受信されるとき、2というオフセット値を有する部分的なフレームが、宛先デバイス902によって受信され、対応する完全なフレームが欠けている場合に利用可能であり得る。しかしながら、部分的なフレームが3とい

うオフセット値を有し、(遅延がないと仮定して)2つのボコーダパケットを含むマルチフレームパケットが40msごとに受信される場合、部分的なフレームは80ms後に受信されることがあり、これは、部分的なフレームが欠けているフレームを復元するために必要とされるときには遅すぎることもある。したがって、オフセットインジケータ915(たとえば、宛先デバイス902によって受信される1つまたは複数の部分的なフレームの平均のオフセット)がオフセット閾値938(たとえば、2.5)以上であるとき、バッファマネージャ910は、追加のバッファ深度を与えて(たとえば、再生遅延を増やして)フレームがバッファ926の中にとどまる時間を増やすことによって、バッファパラメータ934を大きくすると決定し得る。フレームの再生の前にフレームがバッファの中にとどまる時間を増やすことで、バッファマネージャ910は、フレームの再生の前により多くの量の部分的なフレームを受信することが可能になる。バッファマネージャ910は、フレームの復元のためにより多くの量の部分的なフレームを使用できる。

10

**【 0 1 3 2 】**

説明のための非限定的な例として、消去閾値939は3%~5%の間の値を有し得る。バッファマネージャ910は、3%~5%に対応する消去レートを、オーディオ出力958の品質の低下をもたらし得る高い消去レートであると考え得る。したがって、消去インジケータ917(消去レート)が消去閾値939(たとえば、4%)以上であるとき、バッファマネージャ910は、フレームの復元のために部分的なフレームを使用するための時間内に部分的なフレームを受信するために追加のバッファ容量を与えることによってバッファパラメータ934を大きくすることができ、これは消去レートを下げることがある。

20

**【 0 1 3 3 】**

インジケータ913、915、917の1つまたは複数に対応する閾値を満たす(たとえば、それ以上である)ことに応答して、バッファマネージャ910はバッファパラメータ934を更新し得る。たとえば、バッファマネージャ910は、バッファパラメータ934を明示的に大きくすることができ、たとえば最小再生値を大きくする。例示すると、バッファパラメータ934は、インジケータ913、915、917の1つまたは複数に対応する閾値を満たすことに応答して、第1の最小再生値(たとえば、40ms)から第2の最小再生値(たとえば、80ms)へ大きくされ得る。バッファマネージャ910は、無音期間の間にバッファパラメータ934を調整し得る。無音期間の間、(符号化されたオーディオデータを含む)パケットは宛先デバイス902により受信されないことがある。たとえば、無音期間は、2つの異なる発話期間と発話期間との間に存在し得る。バッファパラメータ934を増やすことにより、より多くの部分的なフレームがフレームの復元のためにバッファ926において利用可能になることがあり、消去レートが低下することがある。

30

**【 0 1 3 4 】**

バッファパラメータ934が第2の最小再生値に設定された後で、1つまたは複数のインジケータ913、915、917に対応する閾値をもはや満たさないとき、バッファパラメータ934は第1の最小再生値に戻され得る。たとえば、消去インジケータ917が消去閾値939以上であることに応答してバッファパラメータ934が第2の値に設定されたとき、バッファパラメータ934は、消去インジケータ917が消去閾値939より小さいことに応答して第1の値に設定され得る。

40

**【 0 1 3 5 】**

動作の間、宛先デバイス902は、20msという初期値などの第1の値にバッファパラメータ934を設定し得る。宛先デバイス902は、発話のある期間に対応するパケット942の第2のシーケンスを受信し得る。バッファマネージャ910は、バッファ926を使用してパケット942の第2のシーケンスにおいて受信されるフレームを順序付け得る。バッファマネージャ910は、発話デコーダ956がオーディオ出力958を生成することを可能にするために、順序付けられたフレームをバッファ926(またはメモリ930)から発話デコーダ956に提供し得る。

**【 0 1 3 6 】**

フレームを順序付けることと、フレームを発話デコーダ956に提供することに加えて、バッファマネージャ910は、インジケータ913、915、917を生成し、インジケータ913、915

50

、917の各々を対応する閾値と比較し得る。インジケータ913、915、917の1つまたは複数  
が対応する閾値を満たすことに応答して、バッファマネージャ910はバッファパラメータ9  
34を第2の値へ更新し得る。たとえば、バッファマネージャ910は、その発話の期間に続く  
無音の期間の間にバッファパラメータ934を更新し得る。したがって、第2の値を有するバ  
ッファパラメータ934が、無音の期間に続く第2の発話の期間の間に使用され得る。

【0137】

たとえば、ある特定の時間において、バッファパラメータ934は、ボコーダパケット(た  
とえば、束ねられていないRTPパケット)が受信されるとき、および部分的なフレームが2  
以下の値などの特定のオフセット値(たとえば、部分的なフレームと対応する完全なフレ  
ームとの間のパケットの数)を有するとき、受け入れ可能なフレーム消去レートを可能に  
するように選択される、デフォルト値などの第1の値を有し得る。バッファマネージャは  
、フレーム消去レートが受け入れ可能ではないこと、RTPパケット(たとえば、マルチフレ  
ームの束ねられたパケット)が受信されていること、部分的なフレームが特定のオフセッ  
ト値より大きなオフセット値を有すること、またはこれらの組合せをバッファマネージャ  
910が決定する場合、バッファパラメータ934の値を第2の値に更新する(たとえば、大きく  
する)ように構成され得る。

【0138】

いくつかの実装形態では、図12を参照して説明されるように、バッファパラメータ934  
は、インジケータ913、915、917の3つすべてが対応する閾値を満たすことに応答して更新  
され得る。他の実装形態では、バッファパラメータ934は、3つのインジケータ913、915、  
917のうちの2つが対応する閾値を満たすことに応答して更新され得る。他の実装形態では  
、特定のインジケータが複数の閾値と比較され得る。バッファパラメータ934は、特定の  
インジケータが第1の閾値および第2の閾値より小さいことに応答して、第1の値に更新さ  
れ得る。バッファパラメータ934は、特定のインジケータが第1の閾値以上であることに  
応答して、または、特定のインジケータが第2の閾値以上であり別のインジケータがある条  
件を満たすことに応答して、第2の値に更新され得る。複数の閾値を使用することで、バ  
ッファマネージャは、別のインジケータを対応する閾値と比較する必要なく、特定のイン  
ジケータが第1の閾値以上であることに応答して、バッファパラメータを迅速に更新する  
ことが可能になり得る。

【0139】

結束データ931、オフセットデータ932、消去データ933、またはこれらの組合せなどの  
履歴データを追跡することによって、宛先デバイス902は、オーディオ出力958の品質を低  
下させ得る消去レートの増大を示す(たとえば、それにつながり得る)1つまたは複数の条  
件を特定することが可能であり得る。1つまたは複数の条件を特定したことに応答して、  
宛先デバイス902は、フレームの復元のために部分的なフレームを使用するための時間内  
に部分的なフレームを受信するために追加のバッファ容量を与えるように、バッファパラ  
メータ934を動的に更新し得る。たとえば、部分的なフレームがパケットの受信されたス  
トリームの中に存在するとき、宛先デバイス902は、高い消去レートをもたらし得る、マ  
ルチフレームパケットの結束の使用などの、1つまたは複数の条件に応答して、最小再生  
遅延を増やし得る。たとえば、更新されたバッファパラメータにより、より多数の部分的  
なフレームがバッファ926に含まれることが可能になり得る。したがって、バッファ926に  
対して、欠けているフレームを復元するためにより多数の部分的なフレームが使用可能に  
なる。バッファ926の中により多数の利用可能な部分的なフレームを有することによって  
、より多数の欠けているフレームが復元されることがあり、より少数の消去が生成され  
ることがある。より多数の欠けているフレームを復元し、消去の数を減らすことで、フレ  
ーム消去レートを下げることができ、オーディオ出力958の品質を改善することができる。  
加えて、バッファパラメータ934を永続的に大きくすることはソースデバイス104と宛先デ  
バイス902との間のエンドツーエンドの遅延に悪影響を与えることがあるので、バッファ  
パラメータ934は、1つまたは複数の条件がもはや存在しないときには元の値にリセットさ  
れ得る。

## 【 0 1 4 0 】

システム900によってもたらされる1つの具体的な利点は、バッファ926のバッファパラメータ934が、1つまたは複数の検出された条件に対応する1つまたは複数のインジケータに基づいて動的に更新され得るということである。バッファパラメータ934の増大は一時的であり得るので、更新されたパラメータは、ソースデバイス104と宛先デバイス902との間のエンドツーエンドの遅延を永続的に増やさないことがある。バッファパラメータ934は、1つまたは複数の検出された条件が改善したことに応答して、第1の(より低い)値にリセットされ得る。加えて、更新されたパラメータにより、より高いレベルの品質を有するオーディオ出力を生成することができる。

## 【 0 1 4 1 】

10

図10を参照すると、デジッタバッファを更新する方法のある特定の説明のための実装形態のフローチャートが示されており、全体的に1000と指定される。ある特定の实装形態では、方法1000の1つまたは複数の動作は、図1の宛先デバイス102、デジッタバッファ126、分析器122、図9の宛先デバイス902、バッファ926、またはバッファマネージャ910のうちの少なくとも1つによって実行され得る。

## 【 0 1 4 2 】

1002において、方法1000は、デジッタバッファにおいてパケットを受信するステップを含む。たとえば、図1を参照して説明されるように、図1のデジッタバッファ126は、第2のパケット134を受信し得る。図1を参照して説明されるように、第2のパケット134は、第1の部分的なデータ174および第2のフレームデータ166を含み得る。第1の部分的なデータ174は、入力オーディオ信号130に対応するフレームのシーケンスのうちの第1のフレーム184に対応する第1のフレームデータ164の部分的なコピーを含み得る。第2のフレームデータ166は、フレームのシーケンスのうちの第2のフレーム186に対応し得る。

20

## 【 0 1 4 3 】

1004において、方法1000はまた、第1のデータと関連付けられる第1のフレーム受信タイムスタンプを生成するステップを含む。たとえば、図1を参照して説明されたように、図1の分析器122は、第1の受信タイムスタンプ140を生成し得る。図1を参照して説明されたように、図1の分析器122はまた、デジッタバッファ126の第1のバッファエントリ114に第1の受信タイムスタンプ140を記憶し得る。図1を参照して説明されたように、第1のバッファエントリ114は第1のフレーム184に対応し得る。

30

## 【 0 1 4 4 】

1006において、方法1000はさらに、第2のデータと関連付けられる第2のフレーム受信タイムスタンプを生成するステップを含む。たとえば、図1を参照して説明されたように、図1の分析器122は、第2の受信タイムスタンプ150を生成し得る。図1を参照して説明されたように、図1の分析器122はまた、デジッタバッファ126の第2のバッファエントリ118に第2の受信タイムスタンプ150を記憶し得る。第2のバッファエントリ118は第2のフレーム186に対応し得る。図1を参照して説明されたように、第1のフレーム受信タイムスタンプ140は、第2の受信タイムスタンプ150によって示される第2の時間より早い第1の時間を示し得る。

## 【 0 1 4 5 】

40

いくつかの実装形態では、方法1000は複数のパケットを受信するステップを含み得る。方法1000は、複数のパケットに基づいて、フレーム消去レート、オフセット値、またはフレーム結束値のうちの少なくとも1つを決定するステップを含み得る。方法1000はさらに、フレーム消去レートが第1の閾値以上であること、オフセット値が第2の閾値以上であること、またはフレーム結束値が第3の閾値以上であることのうちの少なくとも1つに応答して、バッファと関連付けられるパラメータを更新するステップを含み得る。

## 【 0 1 4 6 】

いくつかの実装形態では、方法1000は、フレーム消去レートが第1の閾値を満たす、オフセット値が第2の閾値を満たす、およびフレーム結束値が第3の閾値を満たすとの決定に応答して、バッファと関連付けられるパラメータを更新するステップを含み得る。フレー

50

ム結束値は、複数のパケットに基づき得る。パラメータを更新することは、パラメータの値を第1の値から第2の値に変更することを含み得る。方法1000はまた、パラメータが更新された後で、フレーム消去レートが第1の閾値より小さいこと、オフセット値が第2の閾値より小さいこと、またはフレーム結束値が第3の閾値より小さいことに応答して、パラメータの値を第2の値から第1の値に変更するステップを含み得る。

【0147】

方法1000は、第2の受信タイムスタンプ150に基づくバッファリング期間の満了の前の、第1のバッファエントリ114の処理を可能にし得る。第1の受信タイムスタンプ140を生成することで、第1のフレーム184は、第1のフレーム184が処理されるべき次のフレームであるときに、第2のパケット134と関連付けられるバッファリング期間の満了の前に処理されることが可能になり得る。

10

【0148】

図11を参照すると、デジッタバッファを更新する方法のある特定の説明のための実装形態のフローチャートが示されており、全体的に1100と指定される。ある特定の实装形態では、方法1100の1つまたは複数の動作は、図1の宛先デバイス102、デジッタバッファ126、分析器122、図9の宛先デバイス902、バッファ926、またはバッファマネージャ910のうちの少なくとも1つによって実行され得る。

【0149】

702において、方法1100は、デジッタバッファにおいてパケットを受信するステップを含む。たとえば、図1を参照して説明されるように、図1のデジッタバッファ126は、第2のパケット134を受信し得る。図1を参照して説明されるように、第2のパケット134は、入力オーディオ信号130のフレームのシーケンスのうちの第1のフレーム184に対応する第1の部分的なデータ174を含み得る。1104において、方法1100はまた、第1のフレームに対応する消去が再生されたことを検出するステップを含む。たとえば、図1を参照して説明されたように、図1の分析器122は、第1のフレーム184に対応する消去が再生されたことを検出し得る。

20

【0150】

1106において、方法1100はさらに、デジッタバッファの第1のバッファエントリに第1のデータを記憶するステップを含む。たとえば、図1を参照して説明されたように、図1の分析器122は、デジッタバッファの第1のバッファエントリ114に第1の部分的なデータ174を記憶し得る。

30

【0151】

方法1100は、フレームのフレームデータ(たとえば、完全なフレームデータまたは部分的なデータ)がそのフレームに対応する消去を再生した後でデジッタバッファに含められることを可能にし得る。フレームに対応する消去を再生した後でフレームのフレームデータ(たとえば、完全なフレームデータまたは部分的なデータ)を処理することは、デジッタバッファのバッファ深度を暗黙的に大きくすることによる暗黙的バッファ適応を可能にし得る。

【0152】

図12を参照すると、バッファパラメータを調整する方法のある特定の説明のための実装形態のフローチャートが開示されており、全体的に1200と指定される。バッファパラメータは、図1の再生遅延128もしくは図9のバッファパラメータ934を含むことがあり、またはそれらに対応することがある。方法1200は、図1のソースデバイス104、宛先デバイス102(たとえば、分析器122)、または図9の宛先デバイス902(たとえば、バッファマネージャ910)によって実行され得る。図12は、結束インジケータ、オフセットインジケータ、および消去インジケータの組合せに応答した、バッファパラメータの調整を示す。

40

【0153】

1202において、方法1200は、結束インジケータが結束閾値以上であるかどうかを決定するステップを含む。結束インジケータは、図9の結束データ931、結束インジケータ913、もしくはこれらの組合せを含むことがあり、またはこれらに対応することがある。結束閾

50

値は、図9の結束閾値937を含むことがあり、またはそれに対応することがある。結束インジケータが結束閾値より小さいという決定に回答して、方法1200は1208に進み得る。1208において、バッファパラメータは第1の値に設定され得る。いくつかの実装形態では、説明のための非限定的な例として、第1の値は、40msという値などのデフォルト値を含むことがあり、またはそれに対応することがある。バッファパラメータがすでに第1の値に設定されている場合、バッファパラメータの値は変化しないことがある。バッファパラメータを第1の値に設定した後で、方法1200は1202に進み得る。結束インジケータが結束閾値以上であるという決定に回答して、方法1200は1204に進み得る。

【0154】

1204において、方法1200は、オフセットインジケータがオフセット閾値以上であるかどうかを決定するステップを含む。オフセットインジケータは、図1のフレームオフセット136、図9のオフセットデータ932、オフセットインジケータ915、もしくはこれらの組合せを含むことがあり、またはこれらに対応することがある。オフセット閾値は、図9のオフセット閾値938を含むことがあり、またはそれに対応することがある。オフセットインジケータがオフセット閾値より小さいという決定に回答して、方法1200は1208に進み得る。オフセットインジケータがオフセット閾値以上であるという決定に回答して、方法1200は1206に進む。

【0155】

1206において、方法1200は、消去インジケータが消去閾値以上であるかどうかを決定するステップを含む。消去インジケータは、図9の消去データ933、消去インジケータ917、もしくはこれらの組合せを含むことがあり、またはこれらに対応することがある。消去閾値は、図9の消去閾値939を含むことがあり、またはそれに対応することがある。消去インジケータが消去閾値より小さいという決定に回答して、方法1200は1208に進み得る。消去インジケータが消去閾値以上であるという決定に回答して、方法1200は1210に進み得る。

【0156】

1210において、バッファパラメータは第2の値に設定され得る。いくつかの実装形態では、第2の値は第1の値より大きいことがある。バッファパラメータがすでに第2の値に設定されている場合、バッファパラメータの値は変更されないことがある。バッファパラメータを第2の値に設定した後で、方法1200は1202に進み得る。したがって、方法1200は、3つの条件が満たされるときにバッファパラメータを第2の値に設定するステップと、3つよりも少ない条件が満たされるときにバッファパラメータを第1の値に設定するステップとを示す。図12は比較が実行される特定の順序を示すが、比較は別の順序で実行されてもよく、または比較の1つまたは複数は並行して(たとえば、同時に)実行されてもよい。

【0157】

ある代替的な実装形態では、バッファパラメータは、3つの条件のうちの2つが満たされたことに回答して更新され得る。条件は、結束インジケータ、オフセットインジケータ、および消去インジケータが対応する閾値以上であることを含み得る。たとえば、バッファパラメータは、3つの条件のうちの2つが満たされるときに第2の値に設定され、2つ以上の条件が満たされないときに第1の値に設定される。

【0158】

別の代替的な実装形態では、複数の閾値がインジケータのために使用され得る。たとえば、複数の閾値がオフセットインジケータのために使用され得る。代替的な実装形態は、オフセットインジケータが第1のオフセット閾値以上であるかどうかを決定するステップと、オフセットインジケータが第2のオフセット閾値以上であるかどうかを決定するステップとを含み得る。オフセットインジケータが第1のオフセット閾値以上であるという決定に回答して、バッファパラメータは第2の値に設定され得る。バッファパラメータがすでに第2の値に設定されている場合、バッファパラメータの値は変更されないことがある。オフセットインジケータが第1のオフセット閾値未満であるという決定に回答して、オフセットインジケータは第2のオフセット閾値と比較され得る。

【0159】

10

20

30

40

50



オフセットインジケータが第2のオフセット閾値未満であるという決定に応答して、バッファパラメータは第1の値に設定され得る。バッファパラメータがすでに第1の値に設定されている場合、バッファパラメータの値は変更されないことがある。オフセットインジケータが第2のオフセット閾値以上であるという決定に応答して、消去インジケータが消去閾値と比較され得る。

【0160】

消去インジケータが消去閾値未満であるという決定に応答して、バッファパラメータは第1の値に設定され得る。消去インジケータが消去閾値以上であるという決定に応答して、バッファパラメータは第2の値に設定され得る。したがって、バッファパラメータは、3つの条件のうちの2つ(たとえば、第1のオフセット閾値以上のオフセットインジケータ、第2のオフセット閾値以上のオフセットインジケータ、および消去閾値以上の消去インジケータ)が満たされるときは第2の値に設定され、2つ以上の条件が満たされないときは第1の値に設定される。

【0161】

方法1200は、パラメータの動的な調整を可能にする。たとえば、パラメータは、より多数の部分的なフレームがバッファに記憶されフレーム復元動作に利用可能となるように調整され得る(たとえば、大きくされ得る)。より多数の部分的なフレームをバッファに記憶できるようにすることは、フレーム消去レートを下げることがある。

【0162】

図10～図12の方法は、フィールドプログラマブルゲートアレイ(FPGA)デバイス、特定用途向け集積回路(ASIC)、中央処理装置(CPU)などの処理ユニット、デジタル信号プロセッサ(DSP)、コントローラ、別のハードウェアデバイス、ファームウェアデバイス、またはこれらの任意の組合せによって実施され得る。例として、図1および図9に関連して説明されるように、図10～図12の方法の1つまたは複数は、個々に、または組合せで、命令を実行するプロセッサによって実行され得る。

【0163】

図13を参照すると、デバイス(たとえば、ワイヤレス通信デバイス)の特定の説明のための実装形態のブロック図が示されており、全体的に1300と指定される。様々な実装形態では、デバイス1300は、図13に示されるものよりも多数または少数の構成要素を有し得る。説明のための実装形態では、デバイス1300は、図1の宛先デバイス102、ソースデバイス104、および図9の宛先デバイス902のうちの1つまたは複数に対応し得る。説明のための実装形態において、デバイス1300は、図1～図12を参照して説明されたシステムまたは方法の1つまたは複数に従って動作し得る。

【0164】

ある特定の实装形態では、デバイス1300はプロセッサ1306(たとえば、中央処理装置(CPU))を含む。デバイス1300は、1つまたは複数の追加のプロセッサ1310(たとえば、1つまたは複数のデジタル信号プロセッサ(DSP))を含み得る。プロセッサ1310は、発話および音楽コーデック1308とエコーキャンセラ1312とを含み得る。発話および音楽コーデック1308は、ボコーデック1336、ボコーデック1338、または両方を含み得る。

【0165】

デバイス1300は、図1のメモリ176およびコーデック1334を含み得る。メモリ176は、分析データ120を含み得る。図示されていないが、メモリ176はまた、図9の履歴データ(たとえば、931～933)、閾値936、およびバッファパラメータ934を含み得る。デバイス1300は、トランシーバ1350を介してアンテナ1342に結合されるワイヤレスコントローラ1340を含み得る。ある特定の实装形態では、トランシーバ1350は、図1の受信機124を含み得る。

【0166】

デバイス1300は、ディスプレイコントローラ1326に結合されたディスプレイ1328を含み得る。図1のスピーカ142およびマイクロフォン1346または両方が、コーデック1334に結合され得る。コーデック1334は、デジタルアナログコンバータ1302およびアナログデジタル

コンバータ1304を含み得る。ある説明のための実装形態では、マイクロフォン1346は図1のマイクロフォン146に対応し得る。ある特定の実装形態では、コーデック1334は、マイクロフォン146からアナログ信号を受信し、アナログデジタルコンバータ1304を使用してアナログ信号からデジタル信号に変換し、デジタル信号を発話および音楽コーデック1308に提供し得る。発話および音楽コーデック1308はデジタル信号を処理し得る。ある特定の実装形態では、発話および音楽コーデック1308はデジタル信号をコーデック1334に提供し得る。コーデック1334は、デジタルアナログコンバータ1302を使用してデジタル信号をアナログ信号に変換することができ、アナログ信号をスピーカ142に提供することができる。

【0167】

10

デバイス1300は、分析器122、デジッタバッファ126、発話デコーダ160、またはこれらの組合せを含み得る。分析器122、デジッタバッファ126、発話デコーダ160はそれぞれ、図9のバッファマネージャ910、バッファ926、発話デコーダ956を含むことがあり、またはそれらに対応することがある。ある特定の実装形態では、分析器122、発話デコーダ160、または両方が、プロセッサ1306、プロセッサ1310、コーデック1334、発話および音楽コーデック1308、またはこれらの組合せに含まれ得る。ある特定の実装形態では、分析器122、発話デコーダ160、または両方が、ボコーダエンコーダ1336、ボコーダデコーダ1338、または両方に含まれ得る。

【0168】

20

分析器122、デジッタバッファ126、発話デコーダ160、またはこれらの組合せは、本明細書において説明されるデジッタバッファ技法を更新することのハードウェア実装を実装するために使用され得る。代わりに、または加えて、ソフトウェア実装(または組み合わされたソフトウェア/ハードウェア実装)が実装され得る。たとえば、メモリ176は、デバイス1300のプロセッサ1310または他の処理ユニット(たとえば、プロセッサ1306、コーデック1334、または両方)によって実行可能な命令1356を含み得る。命令1356は、図10～図12の方法の1つまたは複数を実行するようにプロセッサ1306および/またはプロセッサ1310によって実行可能な1つまたは複数の命令を含み得る。命令1356は、分析器122、発話デコーダ160、または両方に起因する動作を実施するように実行可能であり得る。

【0169】

30

ある特定の実装形態では、デバイス1300は、システムインパッケージデバイスまたはシステムオンチップデバイス1322に含まれ得る。ある特定の実装形態では、分析器122、デジッタバッファ126、発話デコーダ160、メモリ176、プロセッサ1306、プロセッサ1310、ディスプレイコントローラ1326、コーデック1334、およびワイヤレスコントローラ1340は、システムインパッケージまたはシステムオンチップデバイス1322に含まれる。ある特定の実装形態では、入力デバイス1330および電源1344がシステムオンチップデバイス1322に結合される。その上、ある特定の実装形態では、図13に示されるように、ディスプレイ1328、入力デバイス1330、スピーカ142、マイクロフォン1346、アンテナ1342、および電源1344は、システムオンチップデバイス1322の外部にある。ある特定の実装形態では、ディスプレイ1328、入力デバイス1330、スピーカ142、マイクロフォン1346、アンテナ1342、および電源1344の各々が、インターフェースまたはコントローラなどのシステムオンチップデバイス1322の構成要素に結合され得る。

40

【0170】

デバイス1300は、ヘッドセット、モバイル通信デバイス、スマートフォン、セルラー電話、ラップトップコンピュータ、コンピュータ、タブレット、携帯情報端末、ディスプレイデバイス、テレビジョン、ゲームコンソール、音楽プレーヤ、ラジオ、デジタルビデオプレーヤ、デジタルビデオディスク(DVD)プレーヤ、チューナ、カメラ、ナビゲーションデバイス、またはこれらの任意の組合せを含み得る。

【0171】

説明のための実装形態において、プロセッサ1310は、図1～図12を参照して説明された方法または動作のすべてまたは一部分を実行するように動作可能であり得る。説明のため

50

の例として、プロセッサ1310は、メモリ176に記憶されたソフトウェア(たとえば、1つまたは複数の命令1356からなるプログラム)を実行するように構成され得る。たとえば、プロセッサ1310は、図10の方法1000、図11の方法1100、図12の方法1200、またはそれらの組合せに従って、動作するように構成され得る。例示すると、プロセッサ1310は、デジッタバッファにおいて第1の packets を処理する動作をプロセッサ1310に実行させる命令1356を実行するように構成されることがあり、第1の packets は第1のデータおよび第2のデータを含む。第1のデータは、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデータの部分的なコピーを含み、第2のデータは、フレームのシーケンスのうちの第2のフレームに対応する。プロセッサはまた、第1のデータと関連付けられる第1のフレーム受信タイムスタンプを生成することと、第2のデータと関連付けられる第2のフレーム受信タイムスタンプを生成することとの動作を実行し得る。第1のフレーム受信タイムスタンプは、第2のフレーム受信タイムスタンプによって示される第2の時間より早い第1の時間を示す。プロセッサはまた、第1のデータおよび第2のデータを処理する動作を実行して、処理された発話信号を生成し得る。

#### 【0172】

別の説明のための例では、マイクロフォン1346は、ユーザ発話信号に対応するオーディオ信号を捕捉し得る。ADC1304は、アナログ波形からデジタルオーディオサンプルからなるデジタル信号へと、捕捉されたオーディオ信号を変換し得る。プロセッサ1310はデジタルオーディオサンプルを処理し得る。利得調整器はデジタルオーディオサンプルを調整し得る。エコーキャンセラ1312は、マイクロフォン1346に入るスピーカ142の出力により作り出された可能性のあるエコーを低減し得る。

#### 【0173】

ボコーダエンコーダ1336は、処理された発話信号に対応するデジタルオーディオサンプルを圧縮することができ、packets のシーケンス(たとえば、デジタルオーディオサンプルの圧縮されたビットの表現)を形成することができる。packets のシーケンスはメモリ176に記憶され得る。シーケンスの1つまたは複数のpackets は、シーケンスの他のpackets の部分的なコピーを含み得る。トランシーバ1350は、シーケンスの各packets の何らかの形態を変調することができ(たとえば、他の情報がpackets に付加されることがあり)、アンテナ1342を介して変調されたデータを送信することができる。

#### 【0174】

さらなる例として、アンテナ1342は、ネットワークを介して別のデバイスによって送信されるpackets のシーケンスに対応する入来 packets を受信し得る。受信された packets は、ユーザ発話信号のフレームのシーケンスに対応し得る。

#### 【0175】

分析器122は、デジッタバッファ126(たとえば、デジッタバッファ)に入来 packets を記憶し得る。分析器122は、入来 packets の1つまたは複数の複数のコピーをデジッタバッファ126に記憶し得る。たとえば、分析器122は、デジッタバッファ126に、第1のフレームに対応する第1のバッファエントリに packets の第1のコピーを記憶し、第2のフレームに対応する第2のバッファエントリに packets の第2のコピーを記憶し得る。packets は、第1のフレームのフレームデータの部分的なコピーを含むことがあり、第2のフレームのフレームデータの完全なコピーを含むことがある。別の例では、デジッタバッファは、第1のデータおよび第2のデータを含む packets を受信するように構成され得る。第1のデータは、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデータの部分的なコピーを含み得る。第2のデータは、フレームのシーケンスの中の第2のフレームに対応し得る。分析器122は、packets を受信したことに応答して、第1のデータと関連付けられる第1のフレーム受信タイムスタンプを生成し、第2のデータと関連付けられる第2のフレーム受信タイムスタンプとを生成するように構成され得る。分析器122は、第2のフレーム受信タイムスタンプに基づいて、および第1のフレームと第2のフレームとの間のフレームオフセットに基づいて、第1のフレーム受信タイムスタンプを生成するように構成され得る。たとえば、分析器122は、 $FRTS = PRTS - (FO * FPD) - 1$ ミリ秒という式に基づいて第1の受信

タイムスタンプを決定することがあり、ここでFRTS=第1の受信タイムスタンプであり、PR TS=第2の packets の受信タイムスタンプであり、FO=フレームオフセットであり、FPD=フレーム再生時間長である。分析器122は、第1のバッファエントリに第1のフレーム受信タイムスタンプを記憶するように構成されることがあり、第2のバッファエントリに第2のフレーム受信タイムスタンプを記憶するように構成されることがある。第1のフレーム受信タイムスタンプは、第2のフレーム受信タイムスタンプによって示される第2の時間より早い第1の時間を示し得る。

【0176】

パケットはまた、第2のフレームの第2の生成タイムスタンプを示し得る。分析器122はさらに、フレームオフセットおよび第2の生成タイムスタンプに基づいて第1のフレームの第1の生成タイムスタンプを決定し、第1の生成タイムスタンプが特定のフレームの生成タイムスタンプを超えているとの決定に応答して、第1のバッファエントリに第1のフレーム受信タイムスタンプを記憶するように構成されることがあり、その特定のフレームは直前に処理されたフレームである。たとえば、分析器122は、第1の生成タイムスタンプ=(第2の生成タイムスタンプ158-(フレームオフセット136\*フレームごとのサンプル数))という式に基づいて、第1の生成タイムスタンプを決定し得る。

【0177】

分析器122は、再生遅延および第1のフレーム受信タイムスタンプまたは第1の生成タイムスタンプに基づいて、第1のバッファエントリを処理し得る。分析器122は、第1のフレームのフレームデータの部分的なコピーまたはパケットの第1のコピーをボコーダデコーダ1338に提供し得る。ボコーダデコーダ1338は部分的なコピーを解凍し得る。

【0178】

解凍された波形は、再構築されたオーディオサンプルと呼ばれ得る。エコーキャンセラ1312は、再構築されたオーディオサンプルからエコーを除去し得る。発話デコーダ160は、再構築されたオーディオサンプルに基づいて、処理された発話信号を生成し得る。利得調整器は、処理された発話信号を増幅または抑制し得る。DAC1302は、処理された発話信号をデジタル信号からアナログ波形に変換することができ、変換された信号をスピーカ142に提供することができる。

【0179】

分析器122は、フレームに対応する消去が再生されたことを検出することとは無関係に、デジッタバッファ126にフレームのフレームデータを記憶し得る。消去を再生した後にフレームを処理することは、デジッタバッファ126の深度を暗黙的に大きくすることによる暗黙的バッファ適応を可能にし得る。加えて、分析器122は、結束インジケータ、オフセットインジケータ、消去インジケータ、またはこれらの組合せを1つまたは複数の閾値と比較したことに応答して、バッファパラメータ934を調整し得る。

【0180】

説明された実装形態とともに、装置は、第1のデータおよび第2のデータを含むパケットを受信するための手段を含み得る。たとえば、パケットを受信するための手段は、図1の受信機124、図9の受信機924、トランシーバ1350、パケットを受信するように構成される1つまたは複数の他のデバイス、回路、モジュール、もしくは命令、またはこれらの組合せを含み得る。第1のデータは、フレームのシーケンスのうちの第1のフレームに対応する第1のフレームデータの部分的なコピーを含み得る。第2のデータは、フレームのシーケンスのうちの第2のフレームに対応し得る。

【0181】

装置はまた、パケットを受信したことに応答して第1のデータと関連付けられる第1のフレーム受信タイムスタンプを生成するための手段と、パケットを受信したことに応答して第2のデータと関連付けられる第2のフレーム受信タイムスタンプを生成するための手段とを含む、バッファ管理を実行するための手段を含み得る。たとえば、バッファ管理を実行するための手段、第1のフレーム受信タイムスタンプを生成するための手段、および第2のフレーム受信タイムスタンプを生成するための手段は、分析器122、バッファマネージャ9

10、プロセッサ1306、プロセッサ1310、コーデック1334、またはこれらの組合せを含み得る。第1のフレーム受信タイムスタンプは、第2のフレーム受信タイムスタンプによって示される第2の時間より早い第1の時間を示し得る。いくつかの実装形態では、バッファ管理を実行するための手段はさらに、デジッタバッファの第1のバッファエントリに第1のフレーム受信タイムスタンプを記憶し、デジッタバッファの第2のバッファエントリに第2のフレーム受信タイムスタンプを記憶する機能を実行し得る。第1のバッファエントリは、第1のフレームに対応し得る。第2のバッファエントリは、第2のフレームに対応し得る。バッファ管理を実行するための手段はさらに、第1のバッファエントリを処理して第2のバッファエントリを処理する機能を実行し得る。説明のための実装形態では、バッファ管理を実行するための手段はさらに、1つまたは複数のインジケータが対応する閾値を満たすことに応答して、バッファパラメータを更新するように構成され得る。装置はまた、第1のフレームデータおよび第2のフレームデータに基づいて処理された発話信号を生成するように構成される、発話を復号するための手段を含み得る。たとえば、発話を復号するための手段は、図1の発話デコーダ160、図9の発話デコーダ956、図13のボコーダデコーダ1338、もしくはこれらの組合せを含むことがあり、またはそれらに対応することがある。ある説明のための実施形態では、発話を復号するための手段はさらに、第1のフレームデータを解凍して、解凍された第1のフレームデータに基づいて処理された発話信号を生成する、機能を実行し得る。

#### 【0182】

開示される態様の1つまたは複数が、通信デバイス、固定位置のデータユニット、移動位置のデータユニット、携帯電話、セルラー電話、衛星電話、コンピュータ、タブレット、ポータブルコンピュータ、ディスプレイデバイス、メディアプレーヤ、またはデスクトップコンピュータを含み得る、デバイス1300などのシステムまたは装置において実装され得る。代わりに、または加えて、デバイス1300は、セットトップボックス、エンターテインメントユニット、ナビゲーションデバイス、携帯情報端末(PDA)、モニタ、コンピュータモニタ、テレビジョン、チューナ、ラジオ、衛星ラジオ、音楽プレーヤ、デジタル音楽プレーヤ、ポータブル音楽プレーヤ、ビデオプレーヤ、デジタルビデオプレーヤ、デジタルビデオディスク(DVD)プレーヤ、ポータブルデジタルビデオプレーヤ、衛星、車両、プロセッサを含む任意の他のデバイス、もしくは、データもしくはコンピュータ命令を記憶するか取り出す任意の他のデバイス、またはそれらの組合せを含み得る。別の説明のための非限定的な例として、システムまたは装置は、ハンドヘルドパーソナル通信システム(PCS)ユニットなどの遠隔ユニット、全地球測位システム(GPS)対応デバイス、メータ読取り機器などのポータブルデータユニット、または、プロセッサを含む任意の他のデバイス、もしくはデータもしくはコンピュータ命令を記憶するか取り出す任意の他のデバイス、またはそれらの任意の組合せを含み得る。

#### 【0183】

図1～図13の1つまたは複数は、本開示の教示に従ったシステム、装置、および/または方法を示し得るが、本開示は、これらの示されたシステム、装置、および/または方法に限定されない。本明細書において例示または説明されるような、図1～図13のいずれの図の1つまたは複数の機能または構成要素も、図1～図13の別の機能または構成要素の1つまたは複数の他の部分と組み合わせられ得る。したがって、本明細書において説明されるいずれの単一の例も、限定するものとして見なされるべきではなく、本開示の例は、本開示の教示から逸脱することなく適切に組み合わせられ得る。例として、図10～図12の方法の1つまたは複数は、個々に、または組合せで、図13のプロセッサ1310によって実行され得る。例示すると、図10～図12の方法の1つの一部分が、図10～図12の方法の1つの第2の部分と組み合わせられ得る。加えて、図10～図12を参照して説明される1つまたは複数の動作は、任意選択であってよく、少なくとも部分的に同時に実行されてよく、かつ/または、示された順序または説明された順序とは異なる順序で実行されてよい。

#### 【0184】

当業者は、本明細書において開示される実装形態に関して説明された様々な例示的な論

10

20

30

40

50

理ブロック、構成、モジュール、回路、およびアルゴリズムステップが、電子ハードウェア、プロセッサによって実行されるコンピュータソフトウェア、または両方の組合せとして実装され得ることをさらに理解されよう。様々な説明のための構成要素、ブロック、構成、モジュール、回路、およびステップが、全般にそれらの機能に関して上で説明された。そのような機能がハードウェアとして実装されるかプロセッサ実行可能命令として実装されるかは、システム全体に課される特定の用途および設計制約に依存する。当業者は、説明された機能を具体的な適用例ごとに様々な方法で実装し得るが、そのような実装の決定は、本開示の範囲からの逸脱をもたらすものと解釈されるべきではない。

#### 【 0 1 8 5 】

本明細書において開示される実装形態に関して説明される方法またはアルゴリズムのステップは、直接ハードウェアにおいて、プロセッサによって実行されるソフトウェアモジュールにおいて、またはその2つの組合せにおいて具現化され得る。ソフトウェアモジュールは、ランダムアクセスメモリ(RAM)、フラッシュメモリ、読出し専用メモリ(ROM)、プログラマブル読出し専用メモリ(PROM)、消去可能プログラマブル読出し専用メモリ(EPROM)、電氣的消去可能プログラマブル読出し専用メモリ(EEPROM)、レジスタ、ハードディスク、取外し可能ディスク、コンパクトディスク読出し専用メモリ(CD-ROM)、または当技術分野で既知の任意の他の形態の非一時的記憶媒体に存在し得る。例示的な記憶媒体は、プロセッサが記憶媒体から情報を読み取り、記憶媒体に情報を書き込むことができるように、プロセッサに結合される。代替的に、記憶媒体はプロセッサと一体化されてもよい。プロセッサおよび記憶媒体は、特定用途向け集積回路(ASIC)に存在し得る。ASICは、コンピューティングデバイスまたはユーザ端末に存在し得る。代替的に、プロセッサおよび記憶媒体は、コンピューティングデバイスまたはユーザ端末に個別の構成要素として存在し得る。

#### 【 0 1 8 6 】

開示された態様の前述の説明は、開示された態様を当業者が作成または使用できるようにするために提供される。これらの態様への様々な修正が当業者には容易に明らかになり、本明細書で定義された原理は、本開示の範囲から逸脱することなく、他の態様に適用され得る。したがって、本開示は、本明細書に示される態様に限定されることを意図するものではなく、以下の特許請求の範囲によって定義される原理および新規な特徴と可能な限り一致する最も広い範囲が与えられるべきである。

#### 【 符号の説明 】

#### 【 0 1 8 7 】

- 100 システム
- 102 宛先デバイス
- 104 ソースデバイス
- 106 バッファタイムライン
- 114 第1のバッファエントリ
- 116 処理されたオーディオ信号
- 118 第2のバッファエントリ
- 120 分析データ
- 122 分析器
- 124 受信機
- 126 デジッタバッファ
- 128 再生遅延
- 130 入力オーディオ信号
- 132 第1のパケット
- 134 第2のパケット
- 136 フレームオフセット
- 138 最後の生成タイムスタンプ
- 140 第1の受信タイムスタンプ

10

20

30

40

50

142	スピーカ	
144	第1のフラグ	
146	マイクロフォン	
148	第1の生成タイムスタンプ	
150	第2の受信タイムスタンプ	
152	第1のユーザ	
154	第2のユーザ	
156	第2のフラグ	
158	第2の生成タイムスタンプ	
160	発話デコーダ	10
164	第1のフレームデータ	
166	第2のフレームデータ	
174	第1の部分的なデータ	
176	メモリ	
184	第1のフレーム	
186	第2のフレーム	
190	ネットワーク	
194	第1の生成されたフレーム	
200	図表	
202	バッファエントリ	20
204	バッファエントリ	
212	生成タイムスタンプ	
214	生成タイムスタンプ	
222	フラグ	
224	フラグ	
250	図表	
300	図表	
400	図表	
402	図表	
500	図表	30
600	図表	
602	バッファエントリ	
604	バッファエントリ	
606	バッファエントリ	
608	バッファエントリ	
610	バッファエントリ	
612	バッファエントリ	
622	タイムスタンプ	
624	タイムスタンプ	
626	タイムスタンプ	40
628	タイムスタンプ	
630	タイムスタンプ	
632	タイムスタンプ	
650	表	
652	表	
700	図表	
750	表	
752	表	
800	図表	
802	バッファエントリ	50

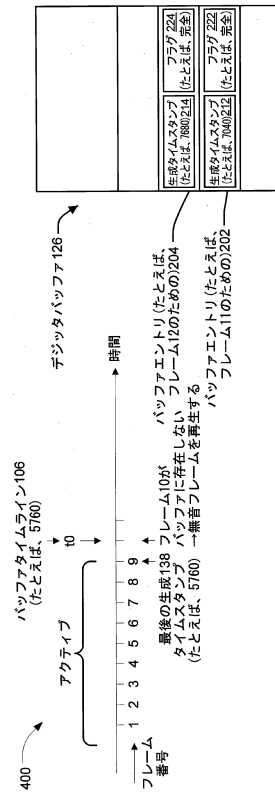
804	バッファエントリ	
806	バッファエントリ	
808	バッファエントリ	
810	バッファエントリ	
822	タイムスタンプ	
824	タイムスタンプ	
826	タイムスタンプ	
828	タイムスタンプ	
830	タイムスタンプ	
850	表	10
852	表	
900	システム	
902	宛先デバイス	
910	バッファマネージャ	
912	結束検出器	
913	結束インジケータ	
914	オフセット検出器	
915	オフセットインジケータ	
916	消去検出器	
917	消去インジケータ	20
924	受信機	
926	バッファ	
930	メモリ	
931	結束データ	
932	オフセットデータ	
933	消去データ	
934	バッファパラメータ	
936	閾値	
937	結束閾値	
938	オフセット閾値	30
939	消去閾値	
940	パケットの第1のシーケンス	
942	パケットの第2のシーケンス	
944	ネットワーク	
956	発話デコーダ	
958	出力	
960	マルチフレームパケットのシーケンス	
961	第1のパケット	
962	第2のパケット	
963	第3のパケット	40
964	第4のパケット	
971	フレーム $n$	
972	フレーム $n+1$	
973	フレーム $n+2$	
974	フレーム $n+3$	
975	部分的なフレーム	
980	マルチフレームパケットのシーケンス	
981	第1のパケット	
982	第2のパケット	
983	第3のパケット	50



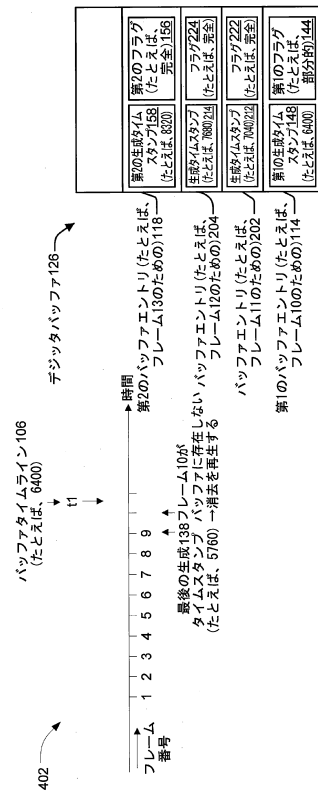
984	第4のパケット	
991	フレーム $n$	
992	フレーム $n+1$	
993	フレーム $n+2$	
994	フレーム $n+3$	
995	フレーム $n+4$	
996	フレーム $n+5$	
997	フレーム $n+6$	
998	フレーム $n+7$	
999	フレーム $n+1$	10
1000	方法	
1100	方法	
1200	方法	
1300	デバイス	
1302	DAC	
1304	ADC	
1306	プロセッサ	
1308	発話および音楽コーデック	
1310	プロセッサ	
1312	エコーキャンセラ	20
1322	システムオンチップデバイス	
1326	ディスプレイコントローラ	
1328	ディスプレイ	
1330	入力デバイス	
1334	コーデック	
1336	ボコーダエンコーダ	
1338	ボコーダデコーダ	
1340	ワイヤレスコントローラ	
1342	アンテナ	
1344	電源	30
1346	マイクロフォン	
1350	トランシーバ	
1356	命令	



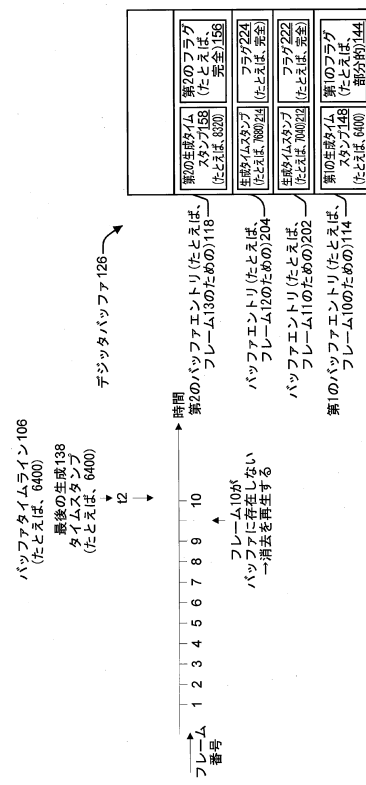
【図 4 A】



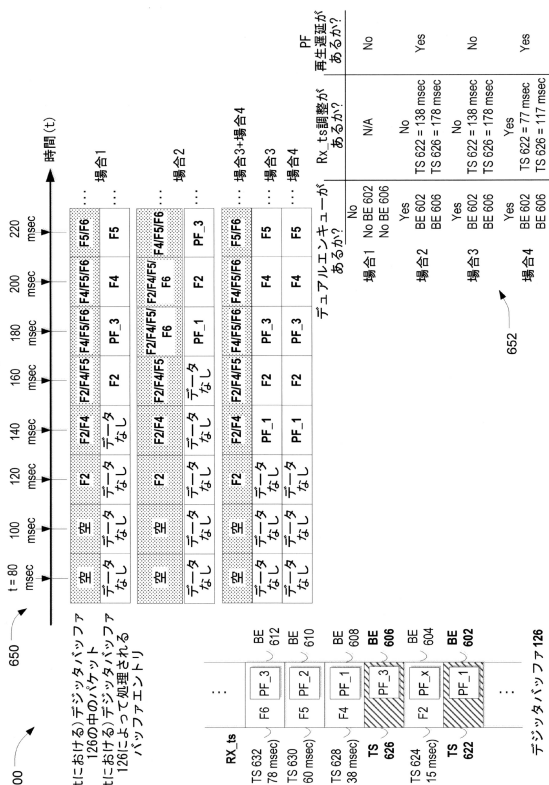
【図 4 B】



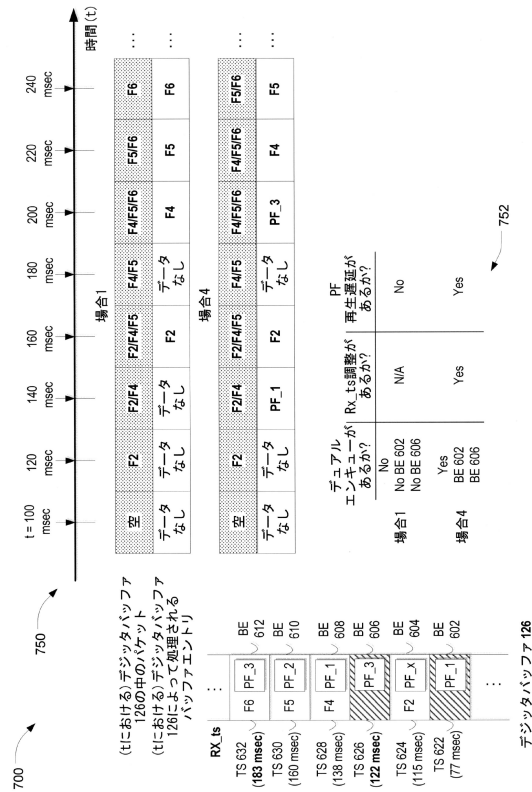
【図 5】



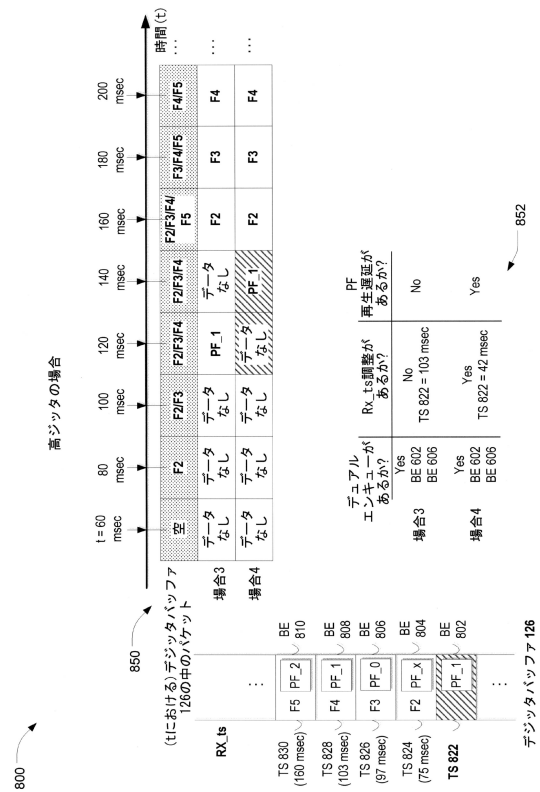
【図 6】



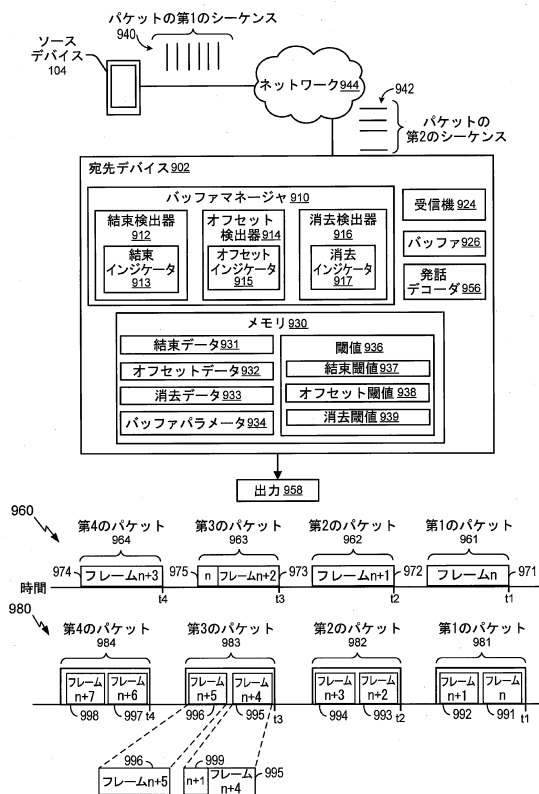
【圖 7】



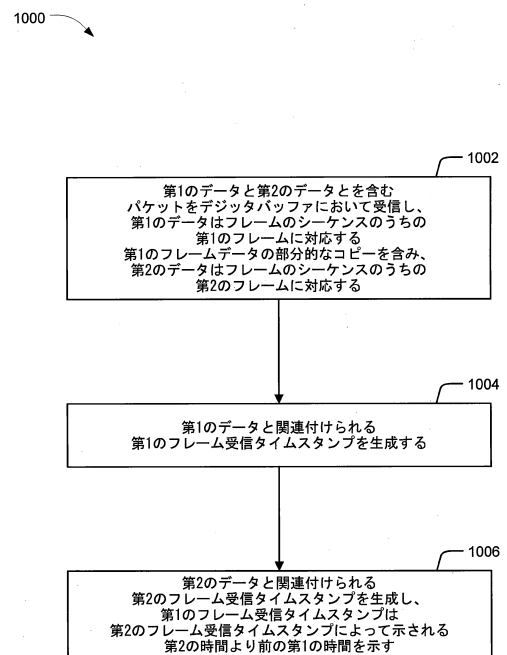
【 図 8 】



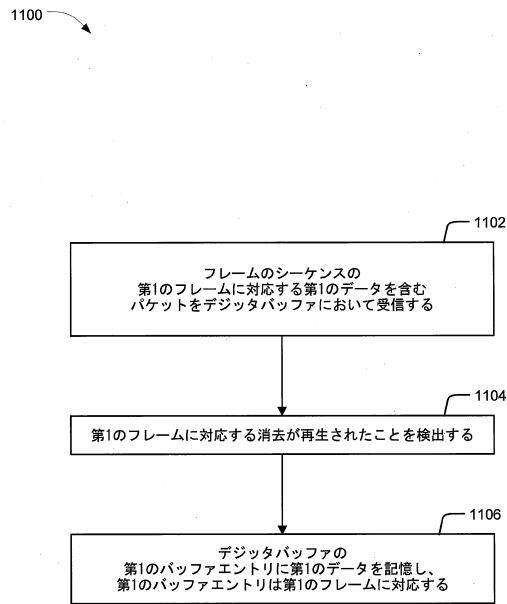
【 図 9 】



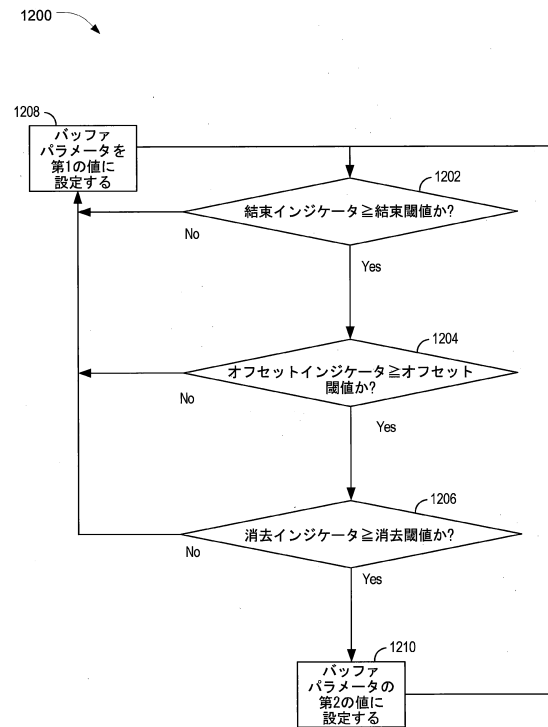
【 図 1 0 】



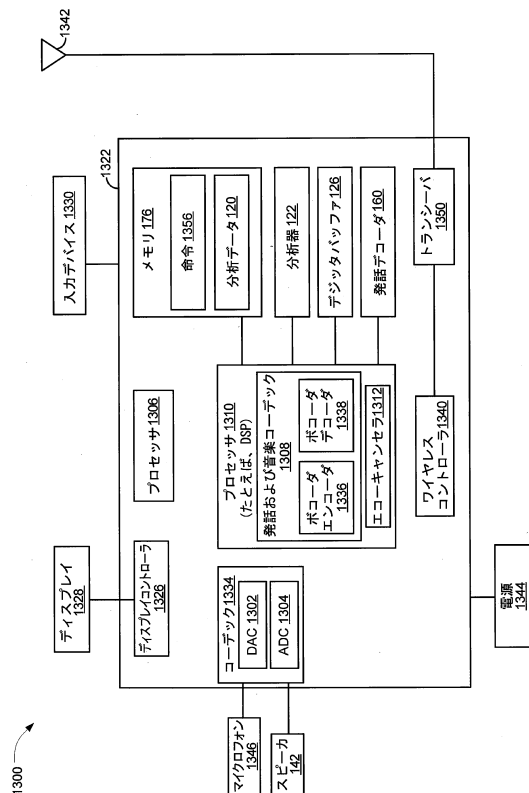
【図 1 1】



【図 1 2】



【図 1 3】



## フロントページの続き

- (31)優先権主張番号 14/829,560  
(32)優先日 平成27年8月18日(2015.8.18)  
(33)優先権主張国 米国(US)

## 早期審査対象出願

- (72)発明者 ヴィヴェク・ラジェンドラン  
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ  
ヴ・5775
- (72)発明者 ドゥミンダ・デウスレンドラ  
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ  
ヴ・5775
- (72)発明者 チャンドラ・マウリ・ポリセッティ  
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ  
ヴ・5775
- (72)発明者 ヴェンカトラマン・エス・アッティ  
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ  
ヴ・5775
- (72)発明者 ヴェンカタ・スブラマニウム・チャンドラ・セカール・チェピーヤム  
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ  
ヴ・5775

審査官 山下 剛史

- (56)参考文献 国際公開第2008/053806(WO,A1)  
国際公開第2014/202647(WO,A1)  
国際公開第2014/202672(WO,A2)  
国際公開第2008/052888(WO,A2)  
国際公開第2011/036123(WO,A1)  
特表2012-529243(JP,A)

## (58)調査した分野(Int.Cl.,DB名)

G10L 19/00-19/26  
H04L 1/00-29/00