

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7588223号
(P7588223)

(45)発行日 令和6年11月21日(2024.11.21)

(24)登録日 令和6年11月13日(2024.11.13)

(51)国際特許分類	F I		
G 1 0 L 15/30 (2013.01)	G 1 0 L 15/30		
H 0 4 R 3/00 (2006.01)	H 0 4 R 3/00	3 2 0	
G 1 0 L 15/00 (2013.01)	G 1 0 L 15/00	2 0 0 G	
G 1 0 L 15/10 (2006.01)	G 1 0 L 15/00	2 0 0 F	
G 1 0 L 15/28 (2013.01)	G 1 0 L 15/10	2 0 0 W	
請求項の数 23 (全26頁) 最終頁に続く			

(21)出願番号	特願2023-520198(P2023-520198)	(73)特許権者	502208397
(86)(22)出願日	令和3年10月1日(2021.10.1)		グーグル エルエルシー
(65)公表番号	特表2023-544043(P2023-544043 A)		Google LLC
(43)公表日	令和5年10月19日(2023.10.19)		アメリカ合衆国 カリフォルニア州 9 4 0 4 3 マウンテン ビュー アンフィシ
(86)国際出願番号	PCT/US2021/053105		アター パークウェイ 1 6 0 0
(87)国際公開番号	WO2022/072786		1 6 0 0 Amphitheatre P
(87)国際公開日	令和4年4月7日(2022.4.7)		arkway 9 4 0 4 3 Mounta
審査請求日	令和5年6月30日(2023.6.30)		in View, CA U.S.A.
(31)優先権主張番号	63/086,953	(74)代理人	110001195
(32)優先日	令和2年10月2日(2020.10.2)		弁理士法人深見特許事務所
(33)優先権主張国・地域又は機関	米国(US)	(72)発明者	ラミン, マルビン・トビ
			スイス、8 0 0 2 チューリッヒ、ブラ
		(72)発明者	ントシェンケシュトラッセ、1 1 0
			サルニカウ, シャルヘイ
			最終頁に続く

(54)【発明の名称】 リアルタイムおよびファイルベースのオーディオデータ処理

(57)【特許請求の範囲】

【請求項 1】

オーディオデータを処理する方法であって、

第1の電子装置から、第2の電子装置が、オーディオデータサンプルの第1のシーケンスおよびオーディオデータサンプルの第2のシーケンスを受信することを含み、前記オーディオデータサンプルの第2のシーケンスは、前記第1の電子装置のマイクロフォンによってキャプチャされるオーディオ信号において前記オーディオデータサンプルの第1のシーケンスに続き、前記方法はさらに、

前記第2の電子装置が前記オーディオデータサンプルの第1のシーケンスをリアルタイムデータ処理モードに従って処理することと、

前記第2の電子装置が前記リアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートできないと判断することと、

前記第2の電子装置が前記リアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートできないという前記判断に従って、前記第2の電子装置のバッファに、前記オーディオデータサンプルの第2のシーケンスをキャッシュし、バッチデータ処理モードで前記オーディオデータサンプルの第2のシーケンスを含むデータファイルを生成することとを含む、オーディオデータを処理する方法。

【請求項 2】

前記処理されたオーディオデータサンプルの第1のシーケンスと、前記オーディオデータサンプルの第2のシーケンスを含む前記データファイルとをサーバシステムに転送する

ことをさらに含み、

前記オーディオデータサンプルの第1のシーケンスは、前記リアルタイムデータ処理モードに対応する第1のデータ転送速度を有し、前記オーディオデータサンプルの第2のシーケンスは、前記バッチデータ処理モードに対応する第2のデータ転送速度を有する、請求項1に記載の方法。

【請求項3】

前記オーディオ信号は、前記オーディオデータサンプルの第1のシーケンスおよび前記オーディオデータサンプルの第2のシーケンスを取得するようオーディオサンプリングレートでサンプリングされ、請求項2に記載の方法。

【請求項4】

前記方法はさらに、前記第1のシーケンス内の前記オーディオデータサンプルを複数のオーディオデータパケットにグループ化することを含み、各オーディオデータパケットは1つ以上の連続するオーディオデータサンプルを含み、前記方法はさらに、

前記複数のオーディオデータパケットをサーバシステムにストリーミングすることを含む、請求項1～3のいずれか1項に記載の方法。

【請求項5】

前記第2の電子装置が前記リアルタイムデータ処理モードにおいて前記オーディオデータサンプルの処理をサポートできないと判断することは、

前記処理されたオーディオデータサンプルの第1のシーケンスのサブセットのデータサンプルレイテンシがレイテンシ許容値を超えると判断すること、

前記処理されたオーディオデータサンプルの第1のシーケンスの欠落オーディオデータサンプルのレートが欠落レート許容値を超えると判断すること、および

前記処理されたオーディオデータサンプルの第1のシーケンスの順序外オーディオデータサンプルのレートが順序外レート許容値を超えると判断すること、のうちの少なくとも1つをさらに含む、請求項1～4のいずれか1項に記載の方法。

【請求項6】

前記第2の電子装置は、前記第2の電子装置に関連付けられるデータサンプルレイテンシと、欠落オーディオデータサンプルのレートと、順序外オーディオデータサンプルのレートとのうちの少なくとも1つに基づいて、前記リアルタイムデータ処理モードにおけるオーディオデータサンプルの転送をサポートしないと判断される、請求項1～5のいずれか1項に記載の方法。

【請求項7】

前記第1の電子装置が、前記オーディオ信号の記録を要求する第1のユーザアクションを受信することをさらに含み、前記オーディオ信号は、前記第1のユーザアクションにตอบสนองしてキャプチャされる、請求項1～6のいずれか1項に記載の方法。

【請求項8】

前記第1の電子装置が、第1の時間に、前記オーディオ信号の記録を要求する第1のユーザアクションを受信し、前記オーディオ信号をキャプチャすることは、前記第1の時間に続く第2の時間に開始される、請求項1～4のいずれか1項に記載の方法。

【請求項9】

前記第1のユーザアクションにตอบสนองして、オーディオデータ要求を送信するための承認を取得することと、前記第2の電子装置がオーディオデータ要求を生成することとをさらに含み、

前記オーディオデータサンプルの第1のシーケンスを受信することは、前記オーディオデータ要求にตอบสนองしてである、請求項7または8に記載の方法。

【請求項10】

前記データファイルは、第1のデータファイルを含み、

前記第1のデータファイルを生成した後、前記第2の電子装置が、前記バッチデータ処理モードでオーディオデータサンプルの第3のシーケンスを含む第2のデータファイルを生成し続けることをさらに含み、前記オーディオデータサンプルの第3のシーケンスは、

10

20

30

40

50

前記オーディオ信号において前記オーディオデータサンプルの第2のシーケンスの直後に続き、前記オーディオデータサンプルの第2のシーケンスおよび前記オーディオデータサンプルの第3のシーケンスの各々は、事前に規定された数のデータサンプルを有する、請求項1～9のいずれか1項に記載の方法。

【請求項11】

前記第2の電子装置は、オーディオ処理のために、前記処理されたオーディオデータサンプルの第1のシーケンスおよびオーディオデータサンプルの第2のシーケンスをサーバシステムに転送するよう構成され、

前記サーバシステムは、ユーザアカウントを含む仮想ユーザドメインをホストし、

前記第1の電子装置および前記第2の電子装置は、前記ユーザアカウントにリンクされる、請求項1～10のいずれか1項に記載の方法。

10

【請求項12】

前記オーディオ信号は、1つ以上の事前に規定されたキーワードまたはユーザ要求を含み、前記第2の電子装置は、前記オーディオ信号内の前記1つ以上の事前に規定されたキーワードまたはユーザ要求を識別するために、前記オーディオデータサンプルの第1のシーケンスおよび前記オーディオデータサンプルの第2のシーケンスをローカルで処理するよう構成される、請求項1～11のいずれか1項に記載の方法。

【請求項13】

前記ユーザ要求は、前記オーディオ信号をキャプチャすることを中止する要求を含み、前記方法はさらに、

20

前記第2の電子装置が、前記オーディオ信号をキャプチャすることを中止する要求を生成することを含み、前記第1の電子装置は、前記要求に回答して、前記オーディオデータサンプルの第2のシーケンスに続くオーディオデータサンプルのシーケンスを転送することをアポートするよう構成される、請求項12に記載の方法。

【請求項14】

前記方法はさらに、

前記データサンプルの第2のシーケンスを転送する間に、前記オーディオ信号をキャプチャすることを中止するための第2のユーザアクションを受信することと、

前記第2のユーザアクションに回答して、前記オーディオデータサンプルの第2のシーケンスの直後に続くオーディオデータサンプルのシーケンスを受信することとを含む、請求項1～13のいずれか1項に記載の方法。

30

【請求項15】

前記第1の電子装置は、遠隔制御装置を含み、前記第2の電子装置は、前記遠隔制御装置によって制御されるよう構成されるネットワーク接続されたテレビ装置を含む、請求項1～14のいずれか1項に記載の方法。

【請求項16】

前記第2の電子装置は、1つ以上のプロセッサと、前記第2の電子装置上でAndroidオペレーティングシステムおよび1つ以上のユーザアプリケーションを実現するよう構成される1つ以上のプログラムを記憶するメモリとを含む、請求項1～15のいずれか1項に記載の方法。

40

【請求項17】

前記Androidオペレーティングシステムは、前記リアルタイムデータ処理モードに従って前記オーディオデータサンプルの第1のシーケンスを処理するための命令を有するオーディオマネージャモジュールを含む、請求項16に記載の方法。

【請求項18】

前記第2の電子装置の前記1つ以上のユーザアプリケーションは、前記オーディオデータサンプルの第2のシーケンスを前記バッファにキャッシュし、前記バッチデータ処理モードにおいて、前記オーディオデータサンプルの第2のシーケンスを含んで、前記データファイルを生成するための、ファイルベースのオーディオデータ処理モジュールを含む、請求項16に記載の方法。

50

【請求項 19】

前記第2の電子装置は、ユーザアクションによってアクティブにされた記録セッション中に前記リアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートしないと判断され、前記データサンプルの第1のシーケンスおよび前記データサンプルの第2のシーケンスの両方が前記記録セッション中に処理される、請求項1～18のいずれか1項に記載の方法。

【請求項 20】

前記第2の電子装置は、2つの別個のユーザアクションによってアクティブにされる2つの別個の記録セッションを分離する中断において前記リアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートしないと判断され、前記データサンプルの第1のシーケンスおよび前記データサンプルの第2のシーケンスは、前記2つの別個の記録セッション中にキャプチャされる、請求項1～16のいずれか1項に記載の方法。

10

【請求項 21】

オーディオデータを処理する方法であって、

第1の電子装置から、第2の電子装置が、オーディオデータサンプルの第1のシーケンスおよびオーディオデータサンプルの第2のシーケンスを受信することを含み、前記オーディオデータサンプルの第2のシーケンスは、前記第1の電子装置のマイクロフォンによってキャプチャされるオーディオ信号において前記オーディオデータサンプルの第1のシーケンスに続き、前記方法はさらに、

20

前記第2の電子装置が、バッチデータ処理モードに従って前記オーディオデータサンプルの第1のシーケンスを処理することを含み、それは、前記第2の電子装置のバッファに前記オーディオデータサンプルの第1のシーケンスをキャッシュすることと、前記オーディオデータサンプルの第1のシーケンスを含むデータファイルを生成することとを含み、前記方法はさらに、

前記第2の電子装置がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得ると判断することと、

前記第2の電子装置が前記リアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得るという前記判断に従って、前記第2の電子装置が、前記リアルタイムデータ処理モードに従って前記オーディオデータサンプルの第2のシーケンスを処理することとを含む、オーディオデータを処理する方法。

30

【請求項 22】

電子装置であって、

1つ以上のプロセッサと、

前記1つ以上のプロセッサによって実行されると、前記1つ以上のプロセッサに請求項1～6、21のいずれか1項に記載の方法を実行させる命令を記憶したメモリとを備える、電子装置。

【請求項 23】

1つ以上のプロセッサに請求項1～21のいずれか1項に記載の方法を実行させるプログラム。

【発明の詳細な説明】

40

【技術分野】

【0001】

関連出願の参照

本出願は、2020年10月2日提出の米国仮出願63/086,953の利益を主張するものであり、それは、参照によりその全体が本明細書に組み込まれる。

【0002】

技術分野

本出願は、概して、限定はしないが、電子装置におけるデータ転送および音声認識を容易にするためにリアルタイムならびにファイルベースのオーディオデータ処理を提供するための方法ならびにシステムを含む、オーディオデータ転送ならびに処理に関する。

50

【背景技術】

【0003】

背景

マイクロフォンを有する電子装置は、ユーザからの音声入力を収集し、音声入力に従って異なる音声起動型機能を開始するためのアシスタント装置として広く使用されている。例えば、デジタルテレビ装置に結合される多くの遠隔制御装置は、マイクロフォンを統合するよう構成される。これらの遠隔制御装置の音声入力は、デジタルテレビ装置にストリーミングされ、デジタルテレビ装置によって少なくとも部分的に処理される。デジタルテレビ装置は、追加のオーディオ処理のために、（前処理された、またはされていない）音声入力を、遠隔サーバシステムに提出してもよい。ユーザ要求は、テレビ装置および/または遠隔サーバシステムにおけるオーディオ処理の結果として音声起動型機能を開始するための音声入力から抽出される。テレビ装置の通信、キャッシュ、および処理能力におけるいかなる欠陥も、音声起動型機能の開始に遅延を引き起こす可能性がある。これは、制限された能力を有する低コストのテレビ装置にしばしば起こる。これらの装置の通信、キャッシュ、または処理能力の欠陥を補償するために、現行よりも効率的なデータ処理および転送機構を有することが有益であろう。

10

【発明の概要】

【0004】

概要

本出願は、マイクロフォンを有する電子装置（例えば、遠隔制御装置、アシスタント装置）から受信したオーディオデータを処理し転送することに関する。電子装置は、オーディオ処理能力を有する別の電子装置（例えば、テレビ装置）に結合されるか、またはオーディオ処理能力を有するサーバに結合される。これら2つの電子装置は、通信チャンネルを介して接続される。オーディオデータは、通信チャンネルを介してリアルタイムで転送され、受信電子装置の通信、計算、および記憶がオーディオデータサンプルの処理をリアルタイムでサポートし得るかどうかに応じて、受信電子装置によって、リアルタイムで、またはバッチで、処理される。それにより、リアルタイムオーディオデータ処理は、特に、リアルタイムでオーディオデータを通信、キャッシュ、または処理するのに十分なリソースを常に有するわけではないいくつかの電子装置において、バッチオーディオデータ処理によって補足される。

20

30

【0005】

具体的には、一局面では、オーディオデータを処理する、たとえば、リアルタイムデータ処理モードからバッチデータ処理モードに切替えるための方法が実現される。本方法は、第1の電子装置（たとえば、遠隔制御装置）から、第2の電子装置（たとえば、テレビ装置）が、オーディオデータサンプルの第1のシーケンスおよびオーディオデータサンプルの第2のシーケンスを受信することを含む。オーディオデータサンプルの第2のシーケンスは、第1の電子装置のマイクロフォンによってキャプチャされるオーディオ信号においてオーディオデータサンプルの第1のシーケンスに続く。本方法は、第2の電子装置が、リアルタイムデータ処理モードに従ってオーディオデータサンプルの第1のシーケンスを処理することと、第2の電子装置がリアルタイムデータ処理モードにおいてオーディオデータサンプルの処理をサポートできないと判断することとをさらに含む。本方法は、第2の電子装置がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートできないという判断に従って、第2の電子装置のバッファに、オーディオデータサンプルの第2のシーケンスをキャッシュすることと、バッチデータ処理モードでオーディオデータサンプルの第2のシーケンスを含むデータファイルを生成することとをさらに含む。

40

【0006】

代替的に、別の局面では、オーディオデータを処理する、たとえばバッチデータ処理モードからリアルタイムデータ処理モードに切替えるための方法が実現される。本方法は、第1の電子装置（たとえば、遠隔制御装置）から、第2の電子装置（たとえば、テレビ装置）が、オーディオデータサンプルの第1のシーケンスおよびオーディオデータサンプル

50

の第2のシーケンスを受信することを含む。オーディオデータサンプルの第2のシーケンスは、第1の電子装置のマイクロフォンによってキャプチャされるオーディオ信号においてオーディオデータサンプルの第1のシーケンスに続く。本方法は、第2の電子装置が、オーディオデータサンプルの第1のシーケンスをバッチデータ処理モードに従って処理することをさらに含み、それは、第2の電子装置のバッファにおいて、オーディオデータサンプルの第1のシーケンスをキャッシュすることと、オーディオデータサンプルの第1のシーケンスを含むデータファイルを生成することとをさらに含む。本方法は、第2の電子装置がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得ると判断することをさらに含む。本方法は、第2の電子装置がリアルタイムデータ処理モードにおいてオーディオデータサンプルの処理をサポートし得るという判断に従って、リアルタイムデータ処理モードに従って第2の電子装置がオーディオデータサンプルの第2のシーケンスを処理することをさらに含む。

10

【0007】

非一時的コンピュータ可読媒体は、命令を記憶しており、命令は、1つ以上のプロセッサによって実行されると、プロセッサに上記の方法のいずれかの方法を実行させる。電子装置は、1つ以上のプロセッサと、命令が記憶されたメモリとを含み、命令は、1つ以上のプロセッサによって実行されると、プロセッサに上記の方法のいずれかの方法を実行させる。

【0008】

図面の簡単な説明

20

説明される様々な実現例をより良く理解するために、以下の実現例の説明を、以下の、同様の参照番号は対応する部分を指す図面と併せて、参照されたい。

【図面の簡単な説明】

【0009】

【図1】いくつかの実現例による、ネットワーク接続されたTV装置、遠隔制御装置、およびサーバシステムが1つ以上の通信ネットワークを介して互いに対話する例示的なメディア環境を示す図である。

【図2】いくつかの実現例による、第1の電子装置と第2の電子装置との間の例示的なオーディオデータ転送経路を示す図である。

【図3】いくつかの実現例による、リアルタイムデータ処理モードからバッチデータ処理モードに切替わる例示的なオーディオデータ処理プロセスを示す概略図である。

30

【図4】いくつかの実現例による、バッチデータ処理モードからリアルタイムデータ処理モードに切替わる例示的なオーディオデータ処理プロセスを示す概略図である。

【図5】いくつかの実現例による、ユーザアクションまたは音声入力によって開始される例示的な音声アシスタントプロセスを示す概略図である。

【図6】いくつかの実現例による、オーディオデータをテレビ装置に転送するよう構成される例示的な遠隔制御装置を示す図である。

【図7】いくつかの実現例による、2つのオーディオデータ処理モードでオーディオデータを動的に処理するための方法のフローチャートである。

【図8】いくつかの実現例による、2つのオーディオデータ処理モードでオーディオデータを動的に処理するための別の方法のフローチャートである。

40

【0010】

同様の参照番号は、図面のいくつかの図を通して対応する部分を指す。

【発明を実施するための形態】

【0011】

実現例の説明

多くの電子装置（例えば、遠隔制御装置、音声起動型ディスプレイ、またはスピーカ装置）は、電子装置が配置されている環境から音声入力を収集するために使用されるマイクロフォンを含む。そのような電子装置は、オーディオ信号内のホットワードを検出すると、または電子装置の専用アシスタントボタンをユーザが押したことに応答して、音声入力

50

を自動的に収集するよう構成されてもよい。音声入力を受信すると、電子装置は、1つ以上の通信ネットワークを介して遠隔サーバシステム（例えば、アシスタントサーバ）に音声入力を通信し、遠隔サーバシステムは、音声入力におけるユーザ要求を認識し、ユーザ要求に応答する。ある例では、電子装置は遠隔制御装置を含み、音声入力は、遠隔制御装置に結合されたネットワーク接続されたテレビ（TV）装置を制御するために開始される。遠隔制御装置は、TV装置を介して遠隔サーバシステムに音声入力を送信し、音声入力は、TV装置が遠隔サーバシステムに音声入力を送信する前にTV装置において処理される。音声入力を処理するこのプロセスの間、TV装置は、その通信、計算、および記憶能力を使用して、遠隔制御装置と遠隔サーバシステムとをブリッジする。

【0012】

TV装置のオーディオマネージャに配信されるオーディオデータは、遠隔制御装置のマイクロフォンによって収集されるものとは異なる場合がある。これは、様々な要因、例えば、遠隔制御装置とTV装置とを結合する通信チャネルを介するデータパケットの損失および遅延、TV装置のプロセッサ負荷に起因して起こる。本出願の様々な実現例では、リアルタイムデータ処理とファイルベースのバッチデータ処理との組合せが、第1の電子装置（たとえば、遠隔制御装置）のマイクロフォンによって収集されたオーディオデータを処理するために、第2の電子装置（たとえば、TV装置）において実現される。いくつかの実現例では、リアルタイムデータ処理は、ファイルベースのバッチデータ処理よりも優先される。第2の電子装置の通信能力、計算能力、および記憶能力のうちの少なくとも1つがオーディオデータサンプルの処理をリアルタイムでサポートできないと判断された場合、後続のデータサンプルはキャッシュされ、第2の電子装置においてデータファイルに編成される（例えば、オーディオマネージャとは異なるオーディオデータ処理モジュールによって処理される）。代替的に、第2の電子装置の通信、計算、および記憶能力がオーディオデータサンプルの処理をリアルタイムでサポートし得ると判断された場合、後続のデータサンプルは、第2の電子装置によってリアルタイムでデータパケットに処理される（例えば、第2の電子装置のオペレーティングシステムの一部であるオーディオマネージャによって処理される）。そのような制御されたオーディオデータ転送プロセスは、第2の電子装置におけるデータ処理および通信を合理化し、電子装置とのユーザ対話を開始する音声入力に関連付けられるパフォーマンス、オーディオ品質、およびユーザ体験を向上させる。

【0013】

ここで、添付の図面にその例が示される実現例を詳細に参照する。以下の詳細な説明では、説明される様々な実現例を十分理解するために、多数の具体的な詳細が述べられる。しかしながら、説明される様々な実現例は、これらの具体的な詳細なしに実施されてもよいことが当業者には明らかであろう。他の例では、周知の方法、手順、構成要素、回路、およびネットワークは、実現例の局面を不必要に不明瞭にしないように、詳細には説明されていない。

【0014】

図1は、いくつかの実現例による、ネットワーク接続されたTV装置102、遠隔制御装置104、およびサーバシステム106が1つ以上の通信ネットワーク180を介して互いに対話する例示的なメディア環境100である。メディア環境100は、サーバシステム106によって作成およびホストされる仮想ユーザドメインに対応し、仮想ユーザドメインは、複数のユーザアカウントを含む。各ユーザアカウントについて、サーバシステム106は、コンテンツソース110および1つ以上のメディア装置102および116~126に結合され、それぞれのユーザアカウントを介してユーザによるレビューのためにコンテンツソース110によって提供されるメディアコンテンツをストリーミングするよう構成される。任意選択で、コンテンツソース110は、広告ソース、電子番組ガイド（EPG）ソース、およびメディアコンテンツソースのうちの1つ以上を含む。

【0015】

具体的には、ユーザおよびユーザアカウントに関連付けられる1つ以上のメディア装置

10

20

30

40

50

は、メディア環境 100 内に配置されて、コンテンツソース 110 に記憶され、そこからストリーミングされるメディアコンテンツをユーザに提供する。コンテンツソース 110 は、任意選択で、第三者メディアコンテンツソースまたはサーバシステム 106 によってホストされる内部メディアソースである。いくつかの実現例では、1つ以上のメディア装置は、遠隔コンテンツソースからメディアコンテンツを直接ストリーミングするか、またはその視聴者に表示するためにメディアコンテンツをストリーミングするよう構成される埋め込みキャストユニットを統合する、ネットワーク接続されたTV装置 102 を含む。ネットワーク接続されたTV装置 102 は、専用の遠隔制御装置 104 および/または遠隔制御アプリケーションを有する電子装置（例えば、携帯電話 122、タブレットコンピュータ 124、ラップトップコンピュータ 126、アシスタント装置 138）に通信可能に結合される。専用の遠隔制御装置 104 は、TV装置 102 に近接して配置されてもよく、デジタル符号化された赤外線信号のパルスを用いてTV装置 102 と通信するよう構成されてもよい。代替として、ある状況では、専用の遠隔制御装置 104 または遠隔制御アプリケーションを有する電子装置は、通信ネットワーク 180 を介して（すなわち、短距離通信リンク、ローカルエリアネットワークおよび/または広域ネットワークを介して）TV装置 102 と通信するよう構成され、TV装置 102 に物理的に近接している必要はない。

【0016】

ネットワーク接続されたTV装置 102 は、1つ以上のプロセッサと、1つ以上のプロセッサによる実行のための命令を記憶するメモリとを含む。ネットワーク接続されたTV装置 102 に記憶された命令は、統合されたTVアプリケーション、ローカルコンテンツキャストアプリケーション、遠隔制御アプリケーション、アシスタントアプリケーション、およびコンテンツソース 110 に関連付けられる1つ以上のメディア再生アプリケーションのうち1つ以上を含む。これらのアプリケーションは、TV装置 102 のオペレーティングシステムとは異なるユーザアプリケーションであり、任意選択で、メディア環境 100 の仮想ユーザドメイン内のユーザアカウントにリンクされる。加えて、ネットワーク接続されたTV装置 102 は、データパケットレベルでオーディオデータを処理するために、そのオペレーティングシステムに統合されたオーディオマネージャ（例えば、図2の234）を含む。

【0017】

代替として、いくつかの実現例では、メディア環境 100 内に配置されるメディア装置は、メディアコンテンツを視聴者に直接出力する表示装置 116 と、表示装置 116 に結合され、メディアコンテンツを表示装置 116 にストリーミングするよう構成されるキャスト装置 118 とを含む。表示装置 116 の例は、限定はしないが、テレビ（TV）表示装置および音楽プレーヤを含む。キャスト装置 118 の例は、セットトップボックス（STB）、DVDプレーヤおよびTVボックスを含むが、これらに限定されない。図1に示されるこの例では、表示装置 116 は、DVDプレーヤまたはセットトップボックス 118 に配線接続されたTVディスプレイを含む。対照的に、いくつかの実現例では、メディア環境 100 に配置されたメディア装置は、メディアコンテンツを視聴者に出力するコンピュータ画面 120A と、メディアコンテンツをコンピュータ画面 120A にストリーミングするデスクトップコンピュータ 120B とを含む。いくつかの実現例では、メディア環境 100 に配置されたメディア装置は、モバイルデバイス、たとえば、携帯電話 122、タブレットコンピュータ 124、およびラップトップコンピュータ 126 を含む。メディア装置 118 ~ 126 の各々は、コンテンツソース 110 またはサーバシステム 106 に関連付けられる内部メディアソースによって提供されるメディアコンテンツ項目を受信および再生するよう構成される1つ以上のメディア再生アプリケーションを含む。

【0018】

サーバシステム 106 は、メディア環境 100 内の1つ以上のメディア装置に対するメディアコンテンツ推奨およびストリーミングを管理するよう構成される統合されたメディ

10

20

30

40

50

プラットフォーム（I - JWP）128を含む。I - AAP128によって生成されたメディアコンテンツ推奨は、サーバ側TVアプリケーション134を介して、ネットワーク接続されたTV装置102上に提示され、サーバ側TVアプリケーション134は、メディアコンテンツ推奨からのユーザ選択にตอบสนองして、TV装置102上の統合されたTVアプリケーション上でメディアコンテンツの表示を可能にする。さらに、UNIP128はまた、メディアコンテンツ推奨を、テレビ装置102に加えて、他のメディア装置118～126に提供するように構成される集中型メディアコンテンツ管理モジュールとして機能してもよい。いくつかの実現例では、各ユーザアカウントに関連付けられる活動データが、TVアプリケーション134およびキャッシングサービスモジュール136から収集され、ユーザアカウントのユーザに提供されるメディアコンテンツ推奨を個人専用にするために使用される。

10

【0019】

いくつかの実現例では、サーバシステム106によってホストされる仮想ユーザドメインのユーザアカウントはまた、メディア装置102, 104, および116～126のうちの1つ以上に加えて、他のタイプの1つ以上の装置、たとえば、メディア環境100にインストールされるネットワーク接続されたアシスタント装置138にも関連付けられる。アシスタント装置138の例は、スピーカアシスタント装置142およびディスプレイアシスタント装置144を含む。スピーカアシスタント装置142は、音声入力を収集し、音声入力からユーザコマンドを認識し、ユーザコマンドにตอบสนองして動作（例えば、音楽を再生すること、質問に答えること）を実施することができる。ディスプレイアシスタント装置144は、オーディオおよび/またはビデオ入力を収集し、オーディオおよび/またはビデオ入力からユーザコマンドを認識し、ユーザコマンドにตอบสนองして動作（例えば、音楽を再生すること、画像または映像クリップを提示すること、質問に答えること）を実施することができる。アシスタント装置138の各々は、専用デバイスアプリケーションまたは汎用ユーザアプリケーション（例えば、ウェブブラウザ）によって任意選択的に管理され、ネットワーク接続されたTV装置102の統合されたTVアプリケーションと連携して仮想ドメイン内のユーザアカウントにリンクされる。

20

【0020】

さらに、いくつかの実現例では、サーバシステム106は、任意選択で、人工知能で動作するアシスタントモジュール140を含む。アシスタントモジュール140は、マイクロフォンによって収集された音声入力からユーザ要求を認識し、インターネットを検索し、イベントおよびアラームをスケジュールし、ハードウェア設定を調整し、パブリックまたはプライベート情報を提示し、メディアコンテンツ項目を再生し、ユーザと双方向会話し、製品を購入し、送金するなどの動作を開始するように構成される。マイクロフォンは、メディア環境100に配置されたメディア装置102, 104, および116～126ならびにアシスタント装置138のいずれかにおいて統合される。いくつかの実現例では、アシスタントモジュール140は、マイクロフォンによって収集された音声入力を処理し、たとえば自然言語処理（NLP）アルゴリズムを使用して音声入力からユーザ要求を識別するように構成される音声認識モジュール160に結合される。

30

【0021】

いくつかの実現例では、サーバシステム106は、サーバシステム106によって管理される1つ以上のユーザアカウントの情報、ならびに1つ以上のユーザアカウントの各々に関連付けられるユーザ装置およびアプリケーションの情報を記憶するように構成される、装置およびアプリケーションレジストリ150を含む。例えば、装置およびアプリケーションレジストリ150は、ネットワーク接続されたTV装置102、遠隔制御装置104、メディア装置116～126、アシスタント装置138の情報と、対応する統合されたTVアプリケーション、遠隔制御アプリケーション、メディア再生アプリケーション、およびアシスタント装置138に関連付けられる専用のデバイスアプリケーションの情報を記憶する。

40

【0022】

50

任意選択的に、同じユーザアカウントに関連付けられるこれらのメディア装置およびアシスタント装置は、異なる地域にわたって分散される。任意選択的に、これらの装置は、同じ物理的位置に配置される。各メディアまたはアシスタント装置は、1つ以上の通信ネットワーク180を使用して、別の装置またはサーバシステム106と通信する。使用される通信ネットワーク180は、限定はしないが、インターネット、イントラネット、ローカルエリアネットワーク(LAN)、セルラーネットワーク、イーサネット(登録商標)、ストレージエリアネットワーク(SAN)、電話ネットワーク、Bluetoothパーソナルエリアネットワーク(PAN)などを含む、1つ以上のタイプのトポロジを有する1つ以上のネットワークであり得る。いくつかの実現例では、サブネットワーク内の2つ以上の装置は有線接続を介して結合され、一方、同じサブネットワーク内の装置のうちの少なくともいくつかはローカル無線通信ネットワーク(例えば、ZigBee、Z-Wave、Insteon、Bluetooth(登録商標)、Wi-Fi、および他の無線通信ネットワーク)を介して結合される。

10

【0023】

様々な実現例では、マイクロフォンを有する第1の電子装置(例えば、遠隔制御装置104、メディア装置116~126のいずれか、アシスタント装置138)は、1つ以上のプロセッサと、通信チャンネルを介して1つ以上のプロセッサによって実行されるべき命令を記憶するメモリとを有する第2の電子装置(例えば、TV装置102またはメディア装置116~126のいずれか)に結合される。第1の電子装置は、そのマイクロフォンを使用してオーディオ信号をキャプチャする。オーディオ信号は、オーディオデータサンプルの第1のシーケンスと、オーディオデータサンプルの第1のシーケンスに続くオーディオデータサンプルの第2のシーケンスとにサンプリングされる。任意選択で、オーディオデータサンプルの第1および第2のシーケンスは、同じ記録セッションにおいて、または2つの別個の記録セッション中に記録される。各記録セッションは、第1の電子装置とのそれぞれのユーザアクション(例えば、ユーザが図6のアシスタントボタン606を押す)によってアクティブにされる。第1の電子装置は、オーディオデータサンプルの第1のシーケンスおよび第2のシーケンスの両方を第2の電子装置に転送し、次いで、第2の電子装置は、リアルタイムデータ処理モードおよびバッチデータ処理モードのうちの1つを介してオーディオデータサンプルの各シーケンスを処理する。

20

【0024】

オーディオデータサンプルの第1のシーケンスを転送する間または転送した後に、第2の電子装置は、第2の電子装置の通信能力、計算能力および記憶能力がリアルタイムデータ処理モードにおいてオーディオデータサンプルの処理をサポートし得るかどうかを判断する。第2の電子装置の通信能力、計算能力、および記憶能力のうちの少なくとも1つがリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートできない場合、第2の電子装置は、(例えば、図2において、そのオペレーティングシステムに、またはユーザアプリケーションとして、インストールされるオーディオデータ処理モジュール238によって、)バッチデータ処理モードでオーディオデータサンプルの第2のシーケンスを処理する。対照的に、第2の電子装置がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得る場合、第2の電子装置は、(例えば、図2で、そのオペレーティングシステムに統合されるオーディオマネージャ234によって、)リアルタイムデータ処理モードでオーディオデータサンプルの第2のシーケンスを処理する。したがって、バッチデータ処理モードは、たとえば、オーディオデータのエラーレートまたはレイテンシが対応するエラーまたはレイテンシ許容値を超えると、第2の電子装置はリアルタイムデータ処理モードにおいてオーディオデータサンプルの処理をサポートできないとの判断に従って、アクティブにされる。

30

40

【0025】

図2は、いくつかの実現例による、第1の電子装置202と第2の電子装置204との間の例示的なオーディオデータ転送経路200である。第1の電子装置202は、オーディオ信号220をキャプチャするように構成されたマイクロフォン206を有するか、ま

50

たはそれに結合され、第2の電子装置204は、通信チャンネル208を介して第1の電子装置202に結合される。オーディオ信号220は、第1の電子装置202においてサンプリングされ、オーディオデータサンプルは、通信チャンネル208を介して第1の電子装置202から第2の電子装置204に転送される。通信チャンネル208は、限定はしないが、ローカル無線通信ネットワーク（例えば、ZigBee, Z-Wave, Insteon, Bluetooth（登録商標）、Wi-Fi、および他の無線通信ネットワーク）を含む1つ以上の通信ネットワーク180によって有効にされる。この例示的なオーディオデータ転送経路200では、通信チャンネル208は、第1の電子装置202の第1のBluetoothスタック208Aおよび第2の電子装置204の第2のBluetoothスタック208Bによって一括して有効にされるBluetooth通信リンクを介して形成される。

10

【0026】

第1の電子装置202は、マイクロフォン206によってキャプチャされたオーディオ信号220のオーディオデータサンプルを取得し、通信チャンネル208を介して第2の電子装置に転送されるようにオーディオデータサンプルを編成するよう構成されるオーディオストリーミングモジュール210を含む。いくつかの実現例では、オーディオストリーミングモジュール210は、オーディオデータサンプルのサブセットを、順序付けられたオーディオデータパケットのシーケンスにグループ化する。各データパケットは、1つ以上の連続するオーディオデータサンプルを含み、任意選択で、事前に規定されたデータフォーマット（例えば、MPEG-4 HE-AACコーデックフォーマット、EVC音声コーデックフォーマット）に従って、プリアンブルと、メッセージヘッダと、符号化パケットデータと、ダミーフィールドと、完全性チェックフィールドとを有する。第1のBluetoothスタック208Aへの入力では、複数の順序付けられたオーディオデータパケットのシーケンスが、通信チャンネル208を介した送信のためにオーディオデータ230のストリームに順次配列される。いくつかの実現例では、同じデータパケットシーケンス中の各データパケットは事前に規定されたデータフォーマットに対応し、2つの異なるデータパケットシーケンスは、任意選択で、同じデータフォーマットまたは異なるデータフォーマットに対応する。

20

【0027】

第2の電子装置204は、3つのレベルのプログラム、すなわち、ハードウェア抽象化層（HAL）上のカーネル216と、デバイスファームウェア218と、アプリケーションおよびサービスプログラム222とを含む。HAL上のカーネル216およびデバイスファームウェア218は、第2の電子装置204のオペレーティングシステムの一部であり、アプリケーションおよびサービスプログラム222は、オペレーティングシステムの外部にあり、専用コンピュータ動作（例えば、ゲーミング、ウェブブラウジング、文書編集、メディア再生）を実現するために製造業者またはユーザによってインストールされる。いくつかの実現例では、第2のBluetoothスタック208Bを介して第1の電子装置202または任意の他の電子装置から任意選択的に生じる動作データを受信した後、第2の電子装置204は、受信された動作データをHAL上のカーネル216の入力イベント識別部226に渡す。入力イベント識別部226は、第2のBluetoothスタック208Bによって受信されたデータから動作データを識別し、識別された動作データをデバイスファームウェア218の入力ディスパッチャ228に与える。入力ディスパッチャ228は、第2の電子装置204上にインストールされたアシスタントアプリケーション232に動作データを割り当てて、動作データ内のユーザ要求を認識し、ユーザ要求に応答して動作を開始する。

30

40

【0028】

いくつかの実現例では、オーディオデータ232のストリームは、第1の電子装置202のマイクロフォン206によって収集され、カーネル/HAL216によって、第1の電子装置202に関連付けられる遠隔制御アプリケーション236に直接ディスパッチされる。遠隔制御アプリケーション236は、デバイスファームウェア218のオーディオマネージャ234および第2の電子装置204上にインストールされたアシスタントアプ

50

リケーション 232 と協働して (260)、オーディオデータ 230 のストリーム内のユーザ要求を認識し、それに応答する。オーディオマネージャ 234 は、リアルタイムデータ処理モードを実現し、オーディオデータ 230 をデータパケットレベルで処理する。すなわち、オーディオマネージャ 234 は、第 1 の電子装置 202 から受信された複数の順序付けられたオーディオデータパケットのシーケンス内のデータパケットからオーディオデータサンプルを識別し、誤ったデータパケット、欠落したデータパケット、および順序外データパケットを補償し、受信されたオーディオデータサンプルを別のオーディオデータパケットのシーケンスに編成し、そのオーディオデータパケットのシーケンスを後のオーディオ処理または転送のためにアシスタントアプリケーション 232 に渡す (262) よう構成される。

10

【0029】

代替的に、いくつかの実現例では、第 2 の電子装置 204 は、バッチデータ処理モードを実現し、第 1 の電子装置 202 から受信したオーディオデータ 230 をデータファイルに処理する、ファイルベースのオーディオデータ処理モジュール 238 を含む。ファイルベースのオーディオデータ処理モジュール 238 は、デバイスファームウェア 218 (すなわち、オペレーティングシステム) に統合される (238A) か、またはアプリケーションおよびサービスプログラム 222 の中でユーザアプリケーションとしてインストールされる (238B) かのいずれかである。遠隔制御アプリケーション 236 は、オーディオデータ 230 を処理モジュール 238 に与える (270 または 280)。処理モジュール 238 は、第 1 の電子装置 202 から受信した複数の順序付けられたオーディオデータパケットのシーケンス内のデータパケットからオーディオデータサンプルを識別し、データパケットをデータファイルにキャッシュし、後のオーディオ処理または転送のために、データファイルのアシスタントアプリケーション 232 に与える (272 または 282) よう構成される。第 2 の電子装置 204 は、データファイルを記憶するためのオーディオデータバッファ 214 をさらに含む。

20

【0030】

オーディオデータサンプルは、オーディオマネージャ 234 および処理モジュール 238 によって、それぞれリアルタイムデータ処理モードおよびバッチデータ処理モードで交互に処理される。第 2 の電子装置 204 は、第 2 の電子装置 204 の通信、計算、および記憶能力がオーディオデータサンプルのリアルタイム転送をサポートし得るかどうかに応じて、これらの 2 つのオーディオデータ処理モードの間で切替わる。いくつかの実現例では、順序付けられたオーディオデータパケットのシーケンスがオーディオマネージャ 234 または処理モジュール 238 によって処理されている間、そのような判断およびモード切替えは動的に実現される。代替として、そのような判断およびモード切替えは、2 つの別個の記録セッション間で、すなわち、リアルタイム処理モードおよびバッチデータ処理モードのうちの一つによって各々が独立して処理される、オーディオデータパケットの 2 つの別個のシーケンス間で実現される。各記録セッションは、任意選択的に、第 1 の電子装置 202 上でのユーザアクションまたは電子装置 202 によって収集されたオーディオデータ 230 から検出された音声起動によってアクティブにされる。

30

【0031】

より具体的には、オーディオデータ 230 のストリームを受信すると、第 2 の電子装置 204 は、オーディオデータ 230 を監視して、第 2 の電子装置 204 の通信、計算、および記憶能力がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得るかどうかを判断する。たとえば、第 2 の電子装置 204 は、たとえば、オーディオマネージャ 234 がオーディオデータ 230 中のデータサンプルを処理する前または処理した後に、オーディオデータ 230 中のデータサンプルのデータサンプルレイテンシがレイテンシ許容値を超えるかどうかをリアルタイムで判断する。別の例では、第 2 の電子装置 204 は、オーディオマネージャ 234 がオーディオデータ 230 内のデータサンプルを処理する前または処理した後に、オーディオデータ 230 中のデータサンプルの欠落オーディオデータサンプルのレートが欠落レート許容値を超えるかどうか、またはオー

40

50

オーディオデータ 230 中のデータサンプルの順序外オーディオデータサンプルのレートが順序外レート許容値を超えるかどうかをリアルタイムで判断する。いくつかの実現例では、第 2 の電子装置 204 は、その中央処理装置 (CPU) 利用率を監視し、CPU 利用率が所定の利用率 (たとえば、85%) を超える場合、それはオーディオデータサンプルの処理をリアルタイムでサポートできないと判断する。判断結果に従って、第 2 の電子装置 204 は、オーディオマネージャ 234 またはファイルベースのオーディオデータ処理モジュール 238 を選択して、オーディオデータ 230 のストリームをそれぞれリアルタイムまたはバッチで処理する。

【0032】

いくつかの実現例では、第 2 の電子装置 204 は、オーディオマネージャ 234 または処理モジュール 238 の出力をローカルで処理して、ユーザのプライバシーを保護する目的でそこからユーザ要求を認識し、任意選択で、処理された出力を、第 1 の電子装置 202 を制御するために遠隔制御アプリケーション 236 に提供するように構成される。代替的に、いくつかの実現例では、第 2 の電子装置 204 は、オーディオマネージャ 234 または処理モジュール 238 の出力を、遠隔サーバシステム 106 に送信する前に前処理して、そこからユーザ要求を認識するように構成される。代替的に、いくつかの実現例では、たとえば、第 2 の電子装置 204 が低コスト装置であるよう意図されるとき、第 2 の電子装置 204 は、限定された音声認識能力を有する。第 2 の電子装置 204 は、オーディオマネージャ 234 または処理モジュール 238 の出力全体を遠隔サーバシステム 106 に送信し、サーバシステム 106 に依拠してそこからユーザ要求を認識するように構成される。

【0033】

図 3 は、いくつかの実現例による、リアルタイムデータ処理モードからバッチデータ処理モードに切替わる例示的なオーディオデータ処理プロセス 300 を示す概略図である。オーディオ信号 220 は、第 1 の電子装置 202 のマイクロフォン 206 を使用してキャプチャされ、オーディオサンプリングレートでサンプリングされて、オーディオデータサンプルの第 1 のシーケンス 302 と、オーディオデータサンプルの第 1 のシーケンス 302 に続くオーディオデータサンプルの第 2 のシーケンス 304 とを取得する。オーディオデータサンプルの第 1 のシーケンス 302 は、通信チャネル 208 を介して第 2 の電子装置 204 に転送される。いくつかの実現例では、第 1 のシーケンス 302 および第 2 のシーケンス 304 中のオーディオデータサンプルは、複数のオーディオデータパケットにグループ化され、各データパケットは、1 つ以上の連続するオーディオデータサンプルを含む。任意選択で、オーディオデータサンプルの第 2 のシーケンス 304 は、第 1 のシーケンス 302 の直後に続く。任意選択で、オーディオデータサンプルの第 2 のシーケンス 304 は、中断によって第 1 のシーケンス 302 から分離される。複数のオーディオデータパケットは、第 2 の電子装置 204 にストリーミングされる。いくつかの状況では、1 つ以上のデータパケットが、通信チャネル 208 を介して転送される過程で、第 1 のシーケンス 302 および第 2 のシーケンス 304 から並べ替えられるか、または破棄される。

【0034】

オーディオデータパケット 308 は、オーディオデータサンプルの第 1 のシーケンス 302 からグループ化され、第 2 の電子装置 204 によってリアルタイム処理モードで処理され、オーディオデータパケット 308 における誤ったデータパケット、欠落したデータパケット、および順序外データパケットが両方とも発生し、訂正されてもよい。すなわち、1 つ以上のデータパケット 308 は、第 2 の電子装置 204 によって (具体的には、オーディオマネージャ 234 によって) 破棄されるかまたは並べ替えられる。オーディオデータサンプルの第 1 のシーケンス 302 を処理する間または処理した後に、第 2 の電子装置 204 は、第 2 の電子装置 204 の通信能力、計算能力、および記憶能力のうちの少なくとも 1 つがリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートできないと判断する。そのような判断に回答して (時間 t_F において)、第 2 の電子装置 204 は、リアルタイムデータ処理モードをアボートし、オーディオデータサンプルの第 2 のシーケンス 304 を処理するためにバッチデータ処理モードを開始する。具体的に

は、第2の電子装置204は、オーディオデータバッファ214にオーディオデータサンプルの第2のシーケンス304をキャッシュし、バッチデータ処理モードでオーディオデータサンプルの第2のシーケンス304を含むデータファイル310を生成する。いくつかの実現例では、第2の電子装置204は、第2の電子装置204に関連付けられるデータサンプルレイテンシ、欠落オーディオデータサンプルのレート、順序外オーディオデータサンプルのレート、およびCPU利用率、の少なくとも1つに基づいて、リアルタイムデータ処理モードにおいてオーディオデータサンプルの処理をサポートする対応する能力を判断する。

【0035】

いくつかの実現例では、第2の電子装置204は、第2の電子装置204の通信、計算、および記憶能力がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得るという判断に従って、オーディオデータサンプルの第2のシーケンス304の処理を中止する。代替的に、いくつかの実現例では、第2の電子装置204は、事前に規定された数のオーディオデータサンプルを含むように第2のシーケンス304を制限するよう構成される。第2の電子装置204は、事前に規定された数に達すると、オーディオデータサンプルの第2のシーケンス304を、第1のデータファイル310にキャッシュされるよう処理する。第2の電子装置204は、次いで、オーディオデータサンプルの第2のシーケンス304の直後に続くオーディオデータサンプルの第3のシーケンス306を第2のデータファイル312に編成する。第2のデータファイル312は、事前に規定された数のオーディオデータサンプルがオーディオデータサンプルの第3のシーケンス306に含まれるとき、または第2の電子装置204がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得ると判断された場合に、転送される。

【0036】

いくつかの状況下では、データファイル310または312に含まれる現在のオーディオデータサンプルの数は、第2の電子装置204がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得ると判断されるとき、事前に規定された数に達していない。判断に従って、データファイル310または312は、現在の数のオーディオデータサンプルとともに直ちに転送されてもよい。任意選択で、データファイル310または312は転送されることからアポートされ、現在の数のオーディオデータサンプルは、第2の電子装置204のオーディオマネージャ234によるリアルタイムオーディオデータ処理のためにデータパケットに再編成される。

【0037】

いくつかの実現例では、処理されたオーディオデータサンプルの第1のシーケンス308、およびオーディオデータサンプルの第2のシーケンスを含むデータファイル310は、サーバシステム106に転送される(320)。処理されたオーディオデータサンプルの第1のシーケンス302は、リアルタイムデータ処理モードに対応する第1のデータ転送速度を有し、オーディオデータサンプルの第2のシーケンス304は、バッチデータ処理モードに対応する第2のデータ転送速度を有する。第2のデータ転送速度は、第1のデータ転送速度よりも大きい。いくつかの実現例では、第1のデータ転送速度はオーディオ信号220のオーディオサンプリングレートよりも遅く、第2のデータ転送速度はそのオーディオサンプリングレートよりも大きい。

【0038】

図4は、いくつかの実現例による、バッチデータ処理モードからリアルタイムデータ処理モードに切替わる例示的なオーディオデータ処理プロセス400を示す概略図である。オーディオ信号220は、第1の電子装置202のマイクロフォンを使用してキャプチャされ、あるオーディオサンプリングレートでサンプリングされて、オーディオデータサンプルの第1のシーケンス402と、オーディオデータサンプルの第1のシーケンス402に続くオーディオデータサンプルの第2のシーケンス404とを取得する。オーディオデータサンプルの第1のシーケンス402およびオーディオデータサンプルの第2のシーケンス404の両方は、通信チャネル208を介して第2の電子装置204に転送される。

オーディオデータサンプルの第1のシーケンス302は、第2の電子装置204によって、バッチデータ処理モードに従って処理される。第1のシーケンス402中のオーディオデータサンプルは、第1の電子装置202のオーディオデータバッファ214にキャッシュされ、データファイル406に編成される。

【0039】

オーディオデータサンプルの第1のシーケンス402を処理する間または処理した後に、第2の電子装置204は、第2の電子装置204がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得ることを検出または判断する。任意選択で、(第1の時間 t_A における)そのような判断に従って、第2の電子装置204は、第1のシーケンス402のオーディオデータサンプルの数が第2の時間 t_B において事前に規定された数に達するまで、第1のシーケンス402に、より多くのオーディオデータサンプルを追加し続ける。第2の電子装置204は、オーディオデータサンプルの第1のシーケンス402をデータファイル406にキャッシュすることを完了した後、オーディオデータサンプルの第1のシーケンス402に続いて収集されるオーディオデータサンプルの第2のシーケンス404をリアルタイムデータ処理モードにおいて処理し始める。代替的に、いくつかの実現例では、(時間 t_C における)判断に従って、第2の電子装置204は、第1のシーケンス402に対するオーディオデータサンプルの数が事前に規定された数に達したかどうかとは無関係に、第1のシーケンス402に対するオーディオデータサンプルの追加を中止する。第2の電子装置204は、時刻 t_C でデータファイル406を作成し、それにより、バッチデータ処理モードを終了する。第2の電子装置204は、リアルタイムデータ処理モードにおいて、オーディオデータサンプルの第1のシーケンス402に続いて収集されたオーディオデータサンプルの第2のシーケンス404の転送を直ちに開始する。さらに、いくつかの実現例(図示せず)では、(たとえば、時間 t_C における)判断で、第2の電子装置204は、バッチデータ処理モードでオーディオデータサンプルの第1のシーケンス402を処理することをアボートし、リアルタイムデータ処理モードで直ちにオーディオデータサンプルの第1のシーケンス402をデータパケット408に処理し始める。オーディオデータサンプルの第1のシーケンス402が転送された後、第2の電子装置204は、リアルタイムデータ処理モードでオーディオデータサンプルの第2のシーケンス404を処理し続ける。

【0040】

オーディオデータサンプルの第1のシーケンス302および402の各々は、任意選択で、第2の電子装置204に送信されるオーディオデータ230のストリームを開始するか、またはオーディオデータ230のストリームの途中にある。同様に、オーディオデータサンプルの第2のシーケンス304および404の各々ならびにオーディオデータサンプルの第3のシーケンス306は、任意選択で、第2の電子装置204に送信されるオーディオデータ230のストリームにおける最後のシーケンスであるか、またはオーディオデータ230のストリームの途中にある。いくつかの実現例では、オーディオデータサンプルの第1のシーケンス302または402は、オーディオデータサンプルの第2のシーケンス304または404の直前ではないことに留意されたい。第1および第2のシーケンスは、中断によって分離された2つの別個の記録セッション中にキャプチャされる。第2の電子装置204は、2つの記録セッションを分離する中断中にリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得るかどうかを判断する。

【0041】

図5は、いくつかの実現例による、ユーザアクションまたは音声入力によって開始される例示的な音声アシスタントプロセス500を示す概略図である。音声アシスタントプロセス500は、第1の電子装置202および第2の電子装置204によって協働して実現される。いくつかの実現例では、第1の電子装置202は物理的なアシスタントボタン(例えば、図6のアシスタントボタン606)を含み、ユーザは、アシスタントボタンに対するユーザアクションを介して音声アシスタント機能の開始または終了を要求することができる。例えば、ユーザは、音声アシスタントプロセス500(録音セッションとも呼ば

10

20

30

40

50

れる)を開始するために物理的なアシスタントボタンを押す。ユーザがアシスタントボタンを押している間、オーディオ信号が、音声アシスタントプロセス500を介する音声データのサンプリング、転送および認識のために第1の電子装置202のマイクロフォン206によって継続的に収集される。マイクロフォン206は、ユーザがアシスタントボタンの押下を解除して対応する記録セッションを完了するまで、オーディオ信号の収集を停止しない。

【0042】

いくつかの実現例では、第1のユーザアクションの検出にตอบสนองして、第1の電子装置202は、第2の電子装置204の遠隔制御アプリケーション236およびアシスタントアプリケーション232にアシスタント呼出し要求502を送信する。アシスタント呼出し要求502にตอบสนองして、アシスタントアプリケーション232は、第1の電子装置202が第2の電子装置204とともに音声アシスタントプロセス500を実施することを許可されていることを検証し、アシスタントを開始する命令504を遠隔制御アプリケーション236に送信する。アシスタントを開始する命令504にตอบสนองして、遠隔制御アプリケーション236は、マイクロフォン開命令506を第1の電子装置202に送信する。第1の電子装置202のマイクロフォン206がオンにされた後、オーディオデータサンプル230が収集され、第2の電子装置204に転送される。オーディオデータサンプル230は、第2の電子装置204に転送された後、リアルタイムデータ処理モードに関連付けられるオーディオマネージャ234によって、またはバッチデータ処理モードに関連付けられるファイルベースのオーディオデータ処理モジュール238によって、処理される(520および530)。いくつかの実現例では、リアルタイムデータ処理モードおよびバッチデータ処理モードは、第2の電子装置204において、たとえば、第2の電子装置204に関連付けられるデータサンプルレイテンシ、欠落オーディオデータサンプルのレート、順序外オーディオデータサンプルのレート、またはCPU利用率に基づいて、動的に交替する。いくつかの実現例では、リアルタイムデータ処理モードおよびバッチデータ処理モードの一方のみが、第1のユーザアクションの検出にตอบสนองしてアクティブにされる各記録セクション中にアクティブにされる。

【0043】

いくつかの実現例では、アシスタントアプリケーション232は、遠隔制御アプリケーション236が、第1の電子装置202によって収集されたオーディオデータをキャプチャするように第1の電子装置202を制御してもよいように、記録を開始する命令508を遠隔制御アプリケーション236に送信する。録音を開始する命令508は、任意選択で、アシスタントを開始する命令504と共に発行され、マイクロフォン開命令506をトリガするよう構成される。記録を開始する命令508にตอบสนองして、オーディオデータサンプル230は、第1の電子装置202から転送された後、第2の電子装置204によって記録される。代替的に、いくつかの実施形態において、記録を開始する命令508'は、オーディオデータサンプル230のサブセットが第2の電子装置204によって受信された後にアシスタントアプリケーションによって発行される。録音を開始する命令508'は、オーディオデータサンプルのサブセットの内容(例えば、内容におけるユーザ要求)に基づいて発行されてもよく、オーディオデータサンプル230のサブセットは、第2の電子装置204によって録音されない。

【0044】

いくつかの実現例では、アシスタントアプリケーション232は、録音を停止する命令510をオーディオマネージャ234および遠隔制御アプリケーション236の両方に送信し、その結果、遠隔制御アプリケーション236は、マイクロフォン206をオフにするように第1の電子装置202を制御するためのマイクロフォン閉命令512を発行してもよい。代替的に、いくつかの実現例では、第1の電子装置202のユーザは、音声アシスタントプロセス500を開始した第1のユーザアクションを終了するか、または第2のユーザアクション(例えば、アシスタントボタンの解放)を適用して音声アシスタントプロセス500を終了する。第2のユーザアクションにตอบสนองして、第1の電子装置202の

10

20

30

40

50

マイクロフォン 206 がオフにされ、アシスタントを終了するための要求 514 が第 2 の電子装置 204 の遠隔制御アプリケーション 236 およびアシスタントアプリケーション 232 に送信される。

【0045】

図 3 および図 4 を参照すると、第 1 の電子装置 202 のマイクロフォン 206 は、オンにされ、第 1 のユーザ押下に関して、異なる時間に、オーディオ信号をキャプチャし始めるように、設定される。いくつかの実現例では、第 1 の電子装置 202 は、第 1 の時間 t_1 においてオーディオ信号の記録を要求する第 1 のユーザアクションを受信し、そのマイクロフォン 206 は、マイクロフォン開命令 506 が第 2 の電子装置 204 から発行されるかどうかとは無関係に、第 2 の時間 t_2 において第 1 のユーザアクションに回答して直ちにオーディオ信号をキャプチャするようにオンにされる。第 2 の時間 t_2 は、第 1 の時間 t_1 と実質的に同時である。代替的に、いくつかの実現例では、第 1 の電子装置 202 は、第 1 の時間 t_1' にオーディオ信号の記録を要求する第 1 のユーザアクションを受信する。オーディオ信号は、第 1 の時間 t_1' に先行する第 2 の時間 t_2 にマイクロフォンによってキャプチャされる。時間 t_2 と t_1' との間の持続時間に対応するオーディオデータはキャッシュされるが、依然として第 1 のユーザアクションに回答して第 2 の電子装置 204 に転送される。加えて、いくつかの実現例では、第 1 の電子装置 202 は、第 1 の時間 t_1'' においてオーディオ信号の記録を要求する第 1 のユーザアクションを受信し、たとえば、第 1 の時間 t_1'' から事前に規定されたバッファ時間（5 秒など）だけ遅延された、第 1 の時間 t_1'' に続く第 2 の時間 t_2 においてオーディオ信号をキャプチャすることを開始する。

10

20

【0046】

いくつかの実現例では、第 1 の電子装置 202 は、第 1 のユーザアクション時間（ t_1 、 t_1' 、または t_1'' ）を受信することと、キャプチャされたオーディオデータを t_2 において第 2 の電子装置 204 に転送することとの間の持続時間において、第 2 の電子装置 204 からマイクロフォン開命令 506 を含むオーディオデータ要求を受信するよう待つことに留意されたい。第 2 の電子装置 204 は、第 1 のユーザアクションに回答してオーディオデータ要求を送信するための承認を取得し、この承認は、第 2 の電子装置 204 のアシスタントアプリケーション 232 によって、または遠隔サーバシステム 106 によって付与される。

30

【0047】

いくつかの状況では、音声アシスタントプロセス 500 は、音声入力によって開始される。第 1 の電子装置 202 のマイクロフォン 206 は、第 1 の電子装置 202 がスリープモードにあるかアクティブモードにあるかにかかわらず、オーディオ信号を連続的に収集し、対応するオーディオデータを第 2 の電子装置 204 に提供するよう構成される。スリープモードでは、オーディオデータは、アクティブモードを有効にするよう 1 つ以上の事前に規定されたホットワード（たとえば、「Hi Google（登録商標）」）が検出されるまで、メディア装置またはユーザアプリケーションを制御するために使用される音声アシスタント機能に対するどのようなユーザ要求も認識するようには処理されない。第 2 の電子装置 204 は、オーディオデータ内の 1 つ以上の事前に規定されたホットワードをローカルで検出するよう構成され、1 つ以上の事前に規定されたホットワードの検出に回答して音声アシスタントプロセス 500 を開始する。

40

【0048】

図 5 を参照すると、いくつかの実施形態において、アシスタント呼出し要求 502 は、1 つ以上の事前に規定されたホットワードを含む。第 2 の電子装置 204 は、第 1 の電子装置 202 から受信したオーディオデータ内においてホットワードを検出すると、アシスタント呼出し要求 502 を確認する。アシスタント呼出し要求 502 に回答して、アシスタントアプリケーション 232 は、第 1 の電子装置 202 が第 2 の電子装置 204 とともに音声アシスタントプロセス 500 を実施することを許可されていることを検証し、アシスタントを開始する命令 504 を遠隔制御アプリケーション 236 に送信する。第 2 の電

50

子装置 204 は、追加のユーザ要求を識別するために、リアルタイムデータ処理モードまたはバッチデータ処理モードでオーディオデータサンプル 230 をローカルで処理するよう、アクティブモードで動作するように制御される。さらに、いくつかの実現例では、アシスタントアプリケーション 232 は、第 1 の電子装置によって提供されるオーディオデータ 230 において、記録を停止するユーザ要求（例えば、図 3 および図 4 の「Google, stop recording (グーグル、記録を止めて)」) を識別する。アシスタントアプリケーション 232 は、オーディオデータサンプル 230 の記録 510 を停止する命令を生成し、命令 510 をオーディオマネージャ 234 および遠隔制御アプリケーション 236 に送信して、遠隔制御アプリケーション 236 がマイクロフォン閉命令 512 を発行して、第 1 の電子装置 202 を制御してそのマイクロフォンをオフにし得るようにする。これらの手段によって、音声アシスタントプロセス 500 は、いかなる物理的なユーザアクションも用いることなく第 1 の電子装置 202 によってキャプチャされたオーディオデータに基づいて開始および終了される。

10

【0049】

図 3 を参照すると、一例では、オーディオ信号をキャプチャすることを中止する要求 314 がオーディオ信号 220 において認識され、第 2 の電子装置 204 から第 1 の電子装置 202 に送信される。要求 314 に応答して、第 2 の電子装置 204 は、第 2 のデータサンプルのシーケンス 304 を含む第 1 のデータファイル 310 の処理を完了するが、第 2 のシーケンス 304 の直後に続く第 3 のデータサンプルのシーケンス 306 を含む第 2 のデータファイル 312 を転送することをアボートする。対照的に、図 4 を参照すると、別の例では、要求 314 を受信すると、第 2 の電子装置 204 は、オーディオデータサンプルの第 2 のシーケンス 404 中において要求 314 に続く後続のデータパケットの処理をアボートする。

20

【0050】

図 6 は、いくつかの実現例による、オーディオデータをテレビ装置に転送するよう構成される例示的な遠隔制御装置 104 を示す。遠隔制御装置 104 の複数のユーザボタンは、電源ボタン 602、ホームボタン 604、アシスタントボタン 606、サイクルボタン 608 (再生/サイクルボタンとも呼ばれる)、前ボタン 610、次ボタン 612、プレビュー/バックドロップボタン 614、および音量制御ボタン 616 のうちの 1 つ以上を含む。電源ボタン 602 上のユーザ作動は、遠隔制御装置 104 に結合されたメディア装置がユーザ作動前にオフである場合は、当該メディア装置の電源をオンにし、当該メディア装置がユーザ作動前にオンである場合は、当該メディア装置の電源をオフにする。ホームボタン 604 に対するユーザ作動は、遠隔制御装置 104 に結合されたメディア装置を制御してホーム画面を表示する。例えば、ホーム画面は、所定のインターネットコンテンツチャンネルによって提供される特定の広告クリップまたはランダムに選択されたメディアプログラムを表示する。いくつかの実現例では、電源ボタン 602 またはホームボタン 604 は、特定のインターネットコンテンツチャンネルによって提供されるメディアコンテンツの即時再生を可能にするよう構成されるクイックキャストボタンとして機能する。

30

【0051】

アシスタントボタン 606 に対するユーザアクションは、メディア環境 100 内のオーディオ信号を収集するように遠隔制御装置 104 に統合されたマイクロフォン 206 を制御し、メディア環境 100 内に位置する 1 つ以上のメディア再生装置 (例えば、TV 装置 102) を制御するように、ユーザ要求がオーディオ信号から抽出される。いくつかの実現例では、遠隔制御装置 104 のマイクロフォン 206 は、アシスタントボタン 606 が第 1 の短時間の間押下げられると、環境 100 からのオーディオ信号の収集を開始し、オーディオ信号の収集を中止するには、第 2 の短時間の押下またはユーザ要求が適用される。代替的に、いくつかの実現例では、遠隔制御装置 104 のマイクロフォン 206 は、アシスタントボタン 606 が押されているときに環境 100 からオーディオ信号を収集するだけであり、アシスタントボタン 606 が解放されているときにオーディオ信号の収集を中止する。加えて、いくつかの実現例では、遠隔制御装置 104 のマイクロフォン 206

40

50

は、環境 100 からのオーディオ信号を継続的にキャプチャし、オーディオ信号は、1つ以上の事前に規定されたキーワードおよび/またはユーザ要求を含む。ユーザ要求は、遠隔制御装置 104 または遠隔制御装置 104 に結合された 1つ以上のメディア装置もしくはアプリケーションを制御するために使用され得る。

【0052】

このアプリケーションでは、オーディオデータは、リアルタイムデータ処理モードおよびバッチデータ処理モードを含む 2つのオーディオデータ処理モードを有する第 2の電子装置 204 (例えば、ネットワーク接続されたテレビ装置 102) において処理される。第 2の電子装置 204 は、第 2の電子装置 204 の通信、キャッシュ、および処理能力がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得るかどうかが判断する。第 2の電子装置 204 がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得るという判断に従って、第 2の電子装置 204 は、後続のオーディオデータサンプルを、リアルタイムデータ処理モードに従って、データパケットレベルで処理する。第 2の電子装置 204 がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートできないという判断に従って、第 2の電子装置 204 は、第 2の電子装置 204 のバッファに後続のオーディオデータサンプルをキャッシュし、バッチデータ処理モードでこれらのオーディオデータサンプルを含むデータファイルを生成する。いくつかの実現例では、判断およびモード有効化は、2つの記録セッション間の中断中に実現される。代替として、いくつかの実現例では、判断およびモード切替えは、同じ記録セッション中に動的に実現される。

【0053】

図 7 は、いくつかの実現例による、オーディオデータを 2つのオーディオデータ処理モードにおいて (リアルタイムデータ処理モードからバッチデータ処理モードに) 動的に処理するための方法 700 のフローチャートである。方法 700 は、第 1の電子装置 202 および第 2の電子装置 204 によって実行され、任意選択で、非一時的コンピュータ可読記憶媒体に記憶され、それぞれの電子装置の 1つ以上のプロセッサによって実行される命令によって管理される。図 7 に示す動作の各々は、コンピュータメモリまたは非一時的コンピュータ可読記憶媒体に記憶された命令に対応してもよい。コンピュータ可読記憶媒体は、磁気ディスク記憶装置もしくは光ディスク記憶装置、フラッシュメモリなどのソリッドステート記憶装置、または他の不揮発性メモリデバイスを含んでもよい。コンピュータ可読記憶媒体に記憶される命令は、1つ以上のプロセッサによって解釈されるソースコード、アセンブリ言語コード、オブジェクトコード、または他の命令フォーマットのうちの 1つ以上を含んでもよい。方法 700 のいくつかの動作は組み合わせられてもよく、および/またはいくつかの動作の順序は変更されてもよい。

【0054】

オーディオ信号は、第 1の電子装置 202 のマイクロフォン 206 を使用してキャプチャされる (702)。第 1の電子装置 202 は、オーディオ信号から、オーディオデータサンプルの第 1のシーケンス 302 と、オーディオデータサンプルの第 1のシーケンス 302 に続くオーディオデータサンプルの第 2のシーケンス 304 とを取得し (704)、リアルタイムデータ処理モードに従って、通信チャネル 208 を介して第 2の電子装置 204 にオーディオデータサンプルの第 1のシーケンス 302 を転送する。第 2の電子装置 204 は、通信チャネル 208 を介して第 1の電子装置 202 から、オーディオデータサンプルの第 1のシーケンス 302 とオーディオデータサンプルの第 2のシーケンス 304 とを受信する (706)。オーディオデータサンプルの第 1のシーケンス 302 の間またはその後に、第 2の電子装置 204 は、第 2の電子装置 204 がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートできないと判断する (708)。第 2の電子装置 204 がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートできないという判断に従って、第 2の電子装置 204 は、第 2の電子装置のバッファにオーディオデータサンプルの第 2のシーケンス 304 をキャッシュし (712)、バッチデータ処理モードでオーディオデータサンプルの第 2のシーケンス 304 を含む

データファイル 310 を生成する (714)。

【0055】

いくつかの実現例では、第2の電子装置204は、処理されたオーディオデータサンプルの第1のシーケンスと、オーディオデータサンプルの第2のシーケンスを含むデータファイルとをサーバシステム106に転送する(たとえば、図5の540)。オーディオデータサンプルの第1のシーケンス302は、リアルタイムデータ処理モードに対応する第1のデータ転送速度を有し、オーディオデータサンプルの第2のシーケンス304は、バッチデータ処理モードに対応する第2のデータ転送速度を有する。第2のデータ転送速度は、第1のデータ転送速度よりも大きい。さらに、ある例では、第1の電子装置202によってキャプチャされたオーディオ信号は、オーディオサンプリングレートでサンプリングされて、オーディオデータサンプルの第1のシーケンス302および第2のシーケンス304を取得する。第1のデータ転送速度はオーディオサンプリングレートよりも遅く、第2のデータ転送速度はオーディオサンプリングレートよりも大きい。

10

【0056】

いくつかの実現例では、第1のシーケンス302におけるオーディオデータサンプルは、リアルタイムデータ処理モードにおいて複数のオーディオデータパケットにグループ化される。各オーディオデータパケットは、一貫したデータフォーマットに従って任意選択で編成される1つ以上の連続するオーディオデータサンプルを含む。複数のオーディオデータパケットは、サーバシステム106にストリーミングされる。

【0057】

いくつかの実現例では、第2の電子装置204は、第2の電子装置204に関連付けられるデータサンプルレイテンシと、欠落オーディオデータサンプルのレートと、順序外オーディオデータサンプルのレートとのうちの少なくとも1つに基づいて、リアルタイムデータ処理モードにおけるオーディオデータサンプルの転送をサポートしないと判断される。具体的には、ある例では、処理されたオーディオデータサンプルの第1のシーケンスのサブセットのデータサンプルレイテンシがレイテンシ許容値を超える。別の例では、処理されたオーディオデータサンプルの第1のシーケンスの欠落オーディオデータサンプルのレートが、欠落レート許容値を超える。さらに別の例では、処理されたオーディオデータサンプルの第1のシーケンスの順序外オーディオデータサンプルのレートが、順序外レート許容値を超える。

20

30

【0058】

いくつかの実現例では、オーディオ信号の記録を要求する第1のユーザアクションが第1の電子装置202において受信される。オーディオ信号は、第1のユーザアクションにตอบสนองしてキャプチャされる。例えば、第1のユーザアクションは、第1の電子装置202のアシスタントボタンの押下である。押下は、第2の電子装置204のアシスタントアプリケーション232またはサーバシステム106による承認を取得するためのプロセスを開始する。承認を受信すると、オーディオ信号は、キャプチャされ、処理され、記録される。具体的には、ある例では、第1の電子装置202は、第1の時間 t_1 'においてオーディオ信号の記録を要求する第1のユーザアクションを受信し、オーディオ信号をキャプチャすることは、第1の時間 t_1 'に続く第2の時間 t_2 において開始される。第2の時間 t_2 は、第1の時間から事前に規定されたバッファ時間だけ遅延される。具体的には、いくつかの状況では、第1のユーザアクションにตอบสนองして、第1の電子装置202は、第2の電子装置204からオーディオデータ要求を受信する。第2の電子装置204は、第1のユーザアクションにตอบสนองしてオーディオデータ要求を送信するための承認を取得するよう構成される。オーディオデータサンプルの第1のシーケンスを転送することは、オーディオデータ要求にตอบสนองして開始される。

40

【0059】

いくつかの実現例では、データファイル310は第1のデータファイル310を含む。第1のデータファイル310を生成した後、第2の電子装置204は、バッチデータ処理モードにおいてオーディオデータサンプルの第3のシーケンス306を含む第2のデータ

50

ファイル 3 1 2 を生成し続ける。オーディオデータサンプルの第 3 のシーケンス 3 0 6 は、オーディオ信号においてオーディオデータサンプルの第 2 のシーケンス 3 0 4 の直後に続き、オーディオデータサンプルの第 2 および第 3 のシーケンス 3 0 4 および 3 0 6 の各々は、事前に規定された数のデータサンプルを有する。

【 0 0 6 0 】

いくつかの実現例では、第 2 の電子装置 2 0 4 は、オーディオ処理（たとえば、音声認識）のために、処理されたオーディオデータサンプルの第 1 のシーケンス 3 0 2 および第 2 のシーケンス 3 0 4 をサーバシステム 1 0 6 に転送するよう構成される。サーバシステム 1 0 6 は、ユーザアカウントを含む仮想ユーザドメインをホストする。第 1 の電子装置 2 0 2 および第 2 の電子装置 2 0 4 は、ユーザアカウントにリンクされる。

10

【 0 0 6 1 】

代替として、いくつかの実現例では、オーディオ信号は、1 つ以上の事前に規定されたホットワードまたはユーザ要求を含む。第 2 の電子装置 2 0 4 は、オーディオデータサンプルの第 1 のシーケンスおよび第 2 のシーケンスをローカルで処理して、オーディオ信号における 1 つ以上の事前に規定されたホットワードまたはユーザ要求を識別するよう構成される。さらに、いくつかの実現例では、ユーザ要求は、オーディオ信号をキャプチャすることを中止する要求を含む。オーディオ信号をキャプチャすることを中止する要求は、第 2 の電子装置 2 0 4 によって認識され、それによって第 1 の電子装置 2 0 2 に提供される。要求に応答して、第 1 の電子装置 2 0 2 は、オーディオデータサンプルの第 2 のシーケンス 3 0 4 に続くオーディオデータサンプルのシーケンスを転送することをアボートする。

20

【 0 0 6 2 】

いくつかの実現例では、データサンプルの第 2 のシーケンス 3 0 4 を転送している間に、第 2 の電子装置 2 0 4 は、オーディオ信号をキャプチャすることを中止するための第 2 のユーザアクションを受信する。第 2 のユーザアクションに回答して、第 2 の電子装置 2 0 4 は、オーディオデータサンプルの第 2 のシーケンス 3 0 4 の直後に続くオーディオデータサンプルのシーケンスを受信することをアボートする。

【 0 0 6 3 】

図 8 は、いくつかの実現例による、オーディオデータを 2 つのオーディオデータ処理モードにおいて（バッチデータ処理モードからリアルタイムデータ処理モードに）動的に処理するための方法 8 0 0 のフローチャートである。方法 8 0 0 は、第 1 の電子装置 2 0 2 および第 2 の電子装置 2 0 4 によって実行され、任意選択で、非一時的コンピュータ可読記憶媒体に記憶され、それぞれの電子装置の 1 つ以上のプロセッサによって実行される命令によって管理される。図 8 に示す動作の各々は、コンピュータメモリまたは非一時的コンピュータ可読記憶媒体に記憶された命令に対応してもよい。コンピュータ可読記憶媒体は、磁気ディスク記憶装置もしくは光ディスク記憶装置、フラッシュメモリなどのソリッドステート記憶装置、または他の不揮発性メモリデバイスを含んでもよい。コンピュータ可読記憶媒体に記憶される命令は、1 つ以上のプロセッサによって解釈されるソースコード、アセンブリ言語コード、オブジェクトコード、または他の命令フォーマットのうちの 1 つ以上を含んでもよい。方法 8 0 0 のいくつかの動作は組み合わされてもよく、および/またはいくつかの動作の順序は変更されてもよい。

30

40

【 0 0 6 4 】

オーディオ信号は、第 1 の電子装置 2 0 2 のマイクロフォンを使用してキャプチャされる（8 0 2）。第 1 の電子装置 2 0 2 は、オーディオ信号中において、オーディオデータサンプルの第 1 のシーケンス 4 0 2、およびオーディオデータサンプルの第 1 のシーケンス 4 0 2 に続くオーディオデータサンプルの第 2 のシーケンス 4 0 4 を取得する（8 0 4）。第 2 の電子装置 2 0 4 は、第 1 の電子装置 2 0 2 から、オーディオデータサンプルの第 1 のシーケンス 4 0 2 とオーディオデータサンプルの第 2 のシーケンス 4 0 4 とを受信する（8 0 6）。第 2 の電子装置 2 0 4 は、第 2 の電子装置のバッファにオーディオデータサンプルの第 1 のシーケンス 4 0 2 をキャッシュすること（8 1 0）と、オーディオデ

50

ータサンプルの第1のシーケンス402を含むデータファイルを生成すること(812)とを含んで、バッチデータ処理モードに従ってオーディオデータサンプルの第1のシーケンス402を処理する。オーディオデータサンプルの第1のシーケンスを処理する間または処理した後に、第2の電子装置204は、第2の電子装置204がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得ると判断する(814)。第2の電子装置204がリアルタイムデータ処理モードでオーディオデータサンプルの処理をサポートし得るという判断に従って、第2の電子装置204は、リアルタイムデータ処理モードに従ってオーディオデータサンプルの第2のシーケンスを処理する(816)。

【0065】

いくつかの実現例では、オーディオデータサンプル402および404は、サーバシステム106に転送される。オーディオデータサンプルの第1のシーケンス402は、リアルタイムデータ処理モードに対応する第1のデータ転送速度を有し、オーディオデータサンプルの第2のシーケンス404は、バッチデータ処理モードに対応する第2のデータ転送速度を有する。第1のデータ転送速度は、第2のデータ転送速度よりも大きい。さらに、ある例では、第2のデータ転送速度はオーディオサンプリングレートよりも遅く、第1のデータ転送速度はオーディオサンプリングレートよりも大きい。いくつかの実現例では、第2のシーケンス404中のオーディオデータサンプルは、リアルタイムデータ処理モードにおいて複数のオーディオデータパケットにグループ化され、複数のオーディオデータパケットは、任意選択で、リアルタイムでサーバシステム106にストリーミングされる。

【0066】

図7および図8を参照すると、いくつかの実現例では、第1の電子装置202は遠隔制御装置104を含み、第2の電子装置204は、遠隔制御装置104によって制御されるよう構成されるネットワーク接続されたTV装置102を含む。いくつかの実現例では、第2の電子装置204は、1つ以上のプロセッサと、第2の電子装置204上でAndroidオペレーティングシステムおよび1つ以上のユーザアプリケーションを実現するよう構成される1つ以上のプログラムを記憶するメモリとを含む。いくつかの実現例では、第1の電子装置202は電池で動作する。

【0067】

図7および図8の各々における動作が説明された特定の順序は、単なる例示であり、説明された順序が、動作が実行され得る唯一の順序であることを示すことを意図していないことを理解されたい。当業者であれば、本明細書で説明されるように、統合されたユーザインターフェース内に情報項目を表示し、コンテンツに焦点を合わせるための種々の方法を認識するであろう。加えて、方法700, 750, 800および850のうちの1つに関して説明される詳細はまた、方法700, 750, 800および850のうちの任意の他のものに類似する方式で適用可能であることに留意されたい。簡潔性のために、類似の詳細は繰り返さない。

【0068】

第1、第2などの文言は、いくつかの例では、本明細書では様々な要素を説明するために使用されるが、これらの要素はこれらの文言によって限定されるべきではないことも理解されよう。これらの文言は、ある要素を別の要素と区別するためにのみ使用される。例えば、説明される様々な実現例の範囲から逸脱することなく、第1の電子装置は第2の電子装置と呼ばれることがあり、同様に、第2の電子装置は第1の電子装置と呼ばれることがある。第1の電子装置および第2の電子装置は両方とも電子装置であるが、それらは同じ電子装置ではない。

【0069】

本明細書で説明される様々な実現例の説明で使用される専門用語は、特定の実現例を説明することのみを目的としており、限定することを意図するものではない。様々な説明される実現例の説明および特許請求の範囲において使用されるように、単数形「a」、「an」、および「the」は、文脈が明確に別段の指示がない限り、複数形も含むことが意図さ

10

20

30

40

50

れる。本明細書で使用される「および/または」という文言は、関連付けられる列挙された項目の1つ以上の任意およびすべての可能な組み合わせを指し、それらを包含することも理解されよう。「含む」、「含み」、「備える」、および/または「備え」という文言は、本明細書で使用される場合、述べられた特徴、完全体、ステップ、動作、要素、および/または構成要素の存在を特定するが、1つ以上の他の特徴、完全体、ステップ、動作、要素、構成要素および/またはその群の存在または追加を排除しないことがさらに理解されよう。

【0070】

本明細書で使用される場合、「if」という文言は、任意に、文脈に応じて、「・・・するとき」または「・・・すると」または「判断に応答して」または「検出に応答して」または「・・・という判断に従って」を意味すると解釈される。同様に、「判断される場合」または「[定められた条件または事象]が検出される場合」という表現は、任意に、文脈に応じて、「判断すると」または「判断することに応じて」または「[定められた条件または事象]を検出すると」または「[定められた条件または事象]を検出することに応じて」または「[定められた条件または事象]が検出されたとの判断に従って」を意味すると解釈される。

10

【0071】

様々な図面は、特定の順序でいくつかの論理ステージを示すが、順序に依存しないステージは並べ替えられてよく、他のステージは組み合わせられてよく、または分解されてよい。いくつかの並べ替えまたは他のグループ分けが具体的に言及されているが、他の並べ替えまたは他のグループ分けが当業者には明らかであり、したがって、本明細書で提示される並べ替えおよびグループ分けは、代替物の網羅的なリストではない。さらに、論理ステージは、ハードウェア、ファームウェア、ソフトウェア、またはそれらの任意の組合せで実現され得ることを認識されたい。

20

【0072】

前述の説明は、説明の目的で、特定の實現例を参照して説明されている。しかしながら、上記の例示的な説明は、網羅的であること、または特許請求の範囲を開示された厳密な形態に限定することを意図するものではない。上記の教示を考慮して、多くの修正および変形が可能である。實現例は、特許請求の範囲の基礎となる原理およびそれらの実用的な用途を最良に説明し、それによって、他の当業者が、企図される特定の用途に適するように様々な修正を伴って實現例を最良に使用できるようにするために選択された。

30

40

50

【 図 面 】

【 図 1 】

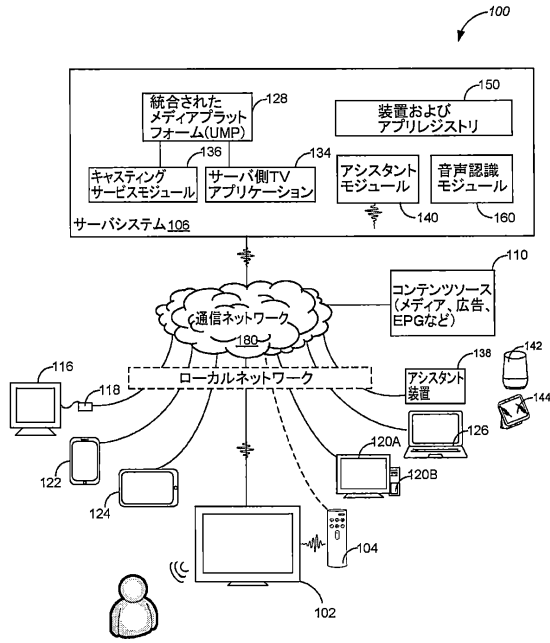


Figure 1

【 図 2 】

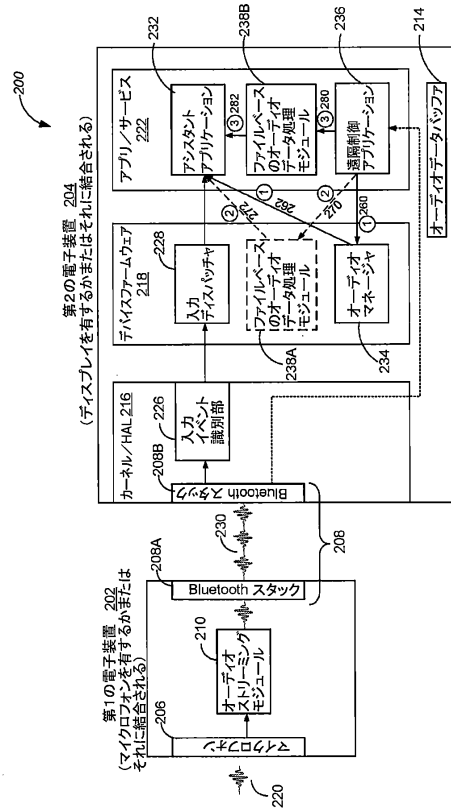


Figure 2

【 図 3 】

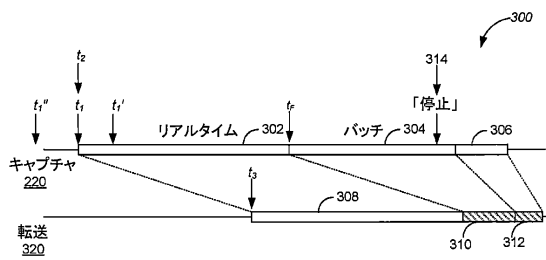


Figure 3

【 図 4 】

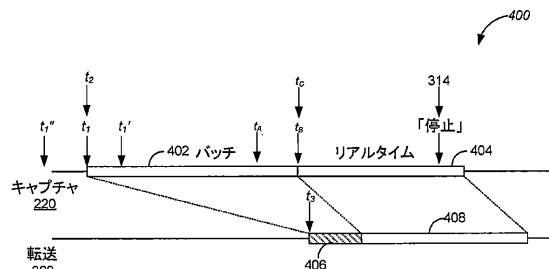


Figure 4

10

20

30

40

50

【 図 5 】

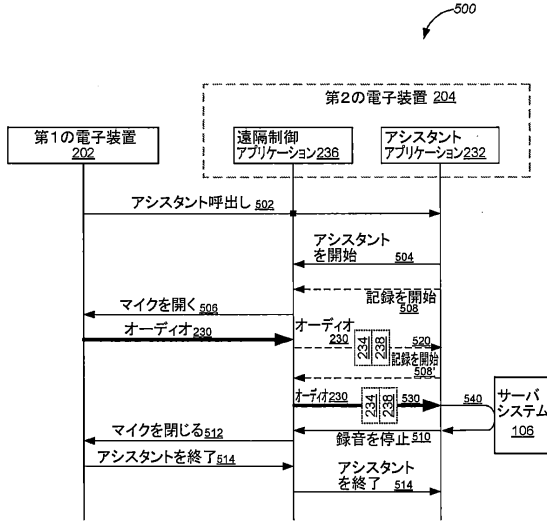


Figure 5

【 図 6 】

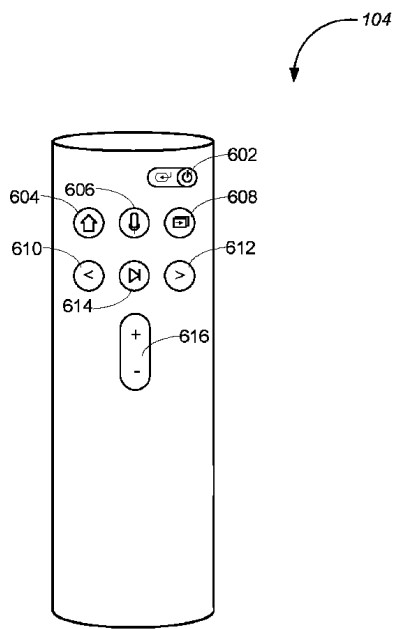


Figure 6

【 図 7 】

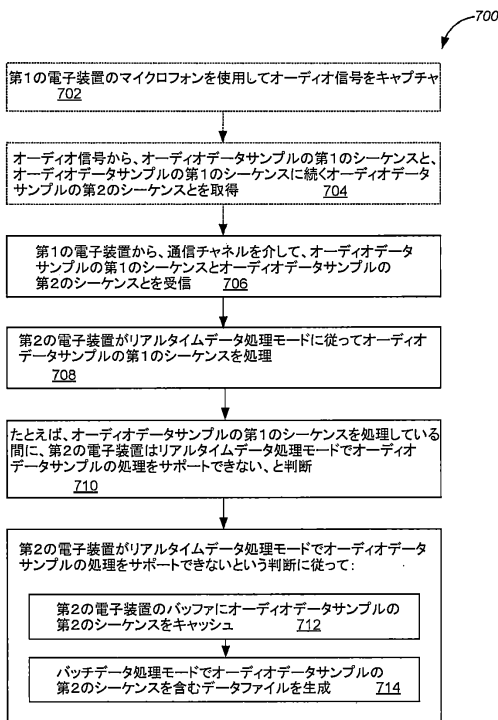


Figure 7

【 図 8 】

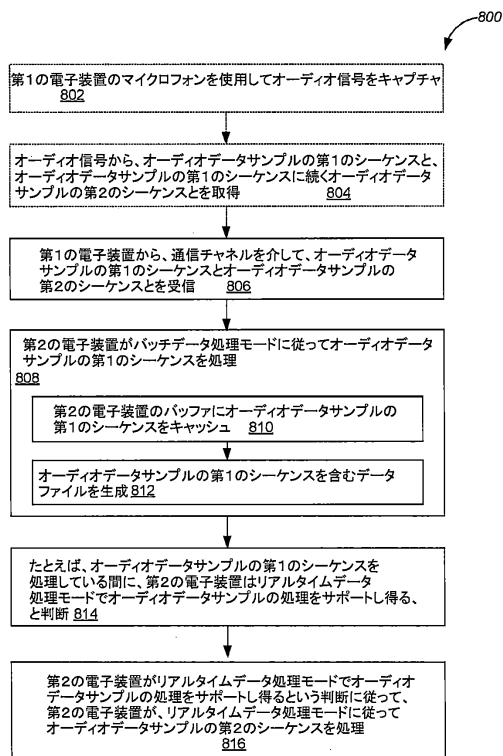


Figure 8

10

20

30

40

50

フロントページの続き

(51)国際特許分類

F I

G 1 0 L 15/28 2 3 0 K

(72)発明者 スイス、8 0 0 2 チューリッヒ、プラントシェンケシュトラーゼ、1 1 0
スリニバサン、マダバ・ガウランガ

スイス、8 0 0 2 チューリッヒ、プラントシェンケシュトラーゼ、1 1 0

審査官 土井 悠生

(56)参考文献 特表2 0 2 0 - 5 2 1 9 9 5 (J P , A)

特開2 0 1 9 - 1 6 9 8 1 2 (J P , A)

特表2 0 1 9 - 5 0 5 0 2 3 (J P , A)

特開2 0 1 7 - 2 2 0 8 0 7 (J P , A)

(58)調査した分野 (Int.Cl. , D B名)

G 1 0 L 1 3 / 0 0 - 9 9 / 0 0

H 0 4 R 3 / 0 0