

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5559898号
(P5559898)

(45) 発行日 平成26年7月23日 (2014. 7. 23)

(24) 登録日 平成26年6月13日 (2014. 6. 13)

(51) Int. Cl. F 1
H 0 3 G 3 / 2 0 (2006. 01) H 0 3 G 3 / 2 0 Z

請求項の数 19 外国語出願 (全 22 頁)

(21) 出願番号	特願2013-21272 (P2013-21272)	(73) 特許権者	502208397
(22) 出願日	平成25年2月6日 (2013. 2. 6)		グーグル インコーポレイテッド
(65) 公開番号	特開2013-162525 (P2013-162525A)		アメリカ合衆国 カリフォルニア州 94
(43) 公開日	平成25年8月19日 (2013. 8. 19)		043 マウンテン ビュー アンフィシ
審査請求日	平成25年3月26日 (2013. 3. 26)		アター パークウェイ 1600
(31) 優先権主張番号	13/368, 173	(74) 代理人	110001195
(32) 優先日	平成24年2月7日 (2012. 2. 7)		特許業務法人深見特許事務所
(33) 優先権主張国	米国 (US)	(72) 発明者	ニクラス・エンボム
			スウェーデン、エスー120 68 スト
			ックホルム、シックラ・カナルガタ、87
		(72) 発明者	ヨーン・スコグランド
			アメリカ合衆国、94114 カリフォル
			ニア州、サン・フランシスコ、ノー・スト
			リート、235

最終頁に続く

(54) 【発明の名称】 通信システムにおける音声レベルを変化させるための制御システム、制御方法、および、プログラム

(57) 【特許請求の範囲】

【請求項 1】

通信システムにおける音声レベルを変化させるための制御システムであって、前記制御システムは、

音声信号および映像信号を受信する少なくとも1つの受信ユニットと、

前記映像信号について実施された顔認識において検知された顔の数に基づいて話している個人の数を決定する決定ユニットと、

前記顔認識において検知された顔の数に基づいて単一話し手モードまたは複数話し手モードへと切り替え、前記単一話し手モードまたは前記複数話し手モードかに基づいて前記音声信号の利得を調整する利得制御部とを備え、

検知された前記顔の数が1であることに応じて、前記利得制御部は、前記単一話し手モードに切り替えて、第1の頻度で前記音声信号の利得を調整し、

検知された前記顔の数が1より大きいことに応じて、前記利得制御部は、前記複数話し手モードに切り替えて、前記第1の頻度とは異なる第2の頻度で前記音声信号の利得を調整する、制御システム。

【請求項 2】

前記決定ユニットは、顔認識において検知された前記顔の数と、前記音声信号に対して実行されたスピーチ分析において検知されたアクティブなスピーチとに基づいて、話している前記個人の数を決定する、請求項1に記載の制御システム。

【請求項 3】

10

20

前記音声信号に対してスピーチ分析を実施することによってアクティブなスピーチを検知するスピーチ検知部をさらに備える、請求項2に記載の制御システム。

【請求項4】

前記映像信号に対して顔認識を実施することによって顔の数を検知する顔検知ユニットをさらに備える、請求項1に記載の制御システム。

【請求項5】

前記利得制御部は、前記顔認識において検知された顔の数および前記スピーチ分析において検知されたアクティブなスピーチの数に基づいて、前記単一話し手モードまたは前記複数話し手モードへと切り替える、請求項2に記載の制御システム。

【請求項6】

前記検知された顔の数が1であることに応じて、前記利得制御部は、前記単一話し手モードへ切り替える前に予め定められた時間期間以上前記映像信号において一つの顔が検知されたと判定する、請求項1に記載の制御システム。

【請求項7】

前記検知された顔の数が1より大きいことに応じて、前記利得制御部は、前記複数話し手モードへ切り替える前に予め定められた時間期間以上前記映像信号において1より大きい数の顔が検知されたと判定する、請求項1に記載の制御システム。

【請求項8】

前記第2の頻度は、前記第1の頻度より高い、請求項1に記載の制御システム。

【請求項9】

前記決定ユニットは、検知されたアクティブなスピーチの音量を少なくとも1つの閾値と比較し、

前記利得制御部は、前記検知されたアクティブなスピーチの音量と前記少なくとも1つの閾値との比較の結果に基づいて、前記第1の頻度または前記第2の頻度と判定する、請求項2に記載の制御システム。

【請求項10】

前記利得制御部は、前記検知されたアクティブなスピーチの音量が、予め定められた時間期間に対する閾値の少なくとも1つより高いか低いかに基づいて、前記第1の頻度または前記第2の頻度と判定する、請求項1に記載の制御システム。

【請求項11】

前記少なくとも1つの受信ユニットは、前記音声信号を受信するための第1の受信ユニットを含み、

前記少なくとも1つの受信ユニットは、前記映像信号を受信するための第2の受信ユニットを含む、請求項1に記載の制御システム。

【請求項12】

前記第1の受信ユニットはマイクロフォンであり、

前記第2の受信ユニットはカメラである、請求項11に記載の制御システム。

【請求項13】

通信システムにおける音声レベルを変化させるための制御方法であって、前記制御方法は、

音声信号を受信するステップと、

映像信号を受信するステップと、

前記映像信号について実施された顔認識処理において検知された顔の数に基づいて、話している個人の数を決するステップと、

前記検知された顔の数が1に等しいことに応じて、単一話し手モードに切り替えて、第1の頻度で前記音声信号の利得を調整するステップと、

前記検知された顔の数が1より大きいことに応じて、複数話し手モードに切り替えて、前記第1の頻度とは異なる第2の頻度で前記音声信号の利得を調整するステップとを備える、制御方法。

【請求項14】

前記話している個人の数の決定は、前記顔認識において検知された顔の数と、音声信号について実行されたスピーチ分析において検知されたアクティブなスピーチとに基づく、請求項 1 3 に記載の制御方法。

【請求項 1 5】

話している前記音声信号におけるアクティブスピーチを検知するために、前記音声信号に対してスピーチ分析を実施するステップをさらに備える、請求項 1 3 に記載の制御方法。

【請求項 1 6】

前記映像信号における顔の数を検知するために、前記映像信号に対する顔認識処理を実施するステップをさらに備える、請求項 1 3 に記載の制御方法。

10

【請求項 1 7】

通信システムにおいて音声レベルを変化させるための制御方法であって、前記制御方法は、

映像信号を取得するステップと、

音声信号を取得するステップと、

前記音声信号において少なくとも 1 人のユーザのスピーチを検知するステップと、

1 つの顔または複数の顔のいずれかを検知するために前記映像信号に対して顔認識を実施するステップと、

検知された 1 つの顔または複数の顔の前記数に基づいて、話している個人の顔の前記数を判定するステップと、

20

話している個人について決定された前記数に基づいて、単一話し手モードと複数話し手モードとの間で切り替えるステップと、

複数の顔の前記検知に応答して前記単一話し手モードから前記複数話し手モードに切り替えるステップと、

単一の顔のみの前記検知に応答して前記複数話し手モードから前記単一話し手モードに切り替えるステップと、

前記複数話し手モードにおいて第 1 の頻度で前記音声信号の前記利得を調整するステップと、

前記単一話し手モードにおいて第 2 の頻度で前記音声信号の前記利得を調整するステップとを備え、前記第 1 の頻度は、前記第 2 の頻度よりも大きい、制御方法。

30

【請求項 1 8】

検知されたスピーチの前記音量を少なくとも 1 つの閾値と比較することによって、検知されたスピーチの前記音量が所与の音量範囲外にあるかどうかを判定するステップと、

検知されたスピーチの前記音量が前記所与の音量範囲外にあるという事象の発生に基づいて、前記検知されたスピーチの前記音量が所定長さの時間にわたって前記所与の音量範囲外にあるかどうかを判定するステップと、

検知されたスピーチの前記音量に基づいて前記第 1 の頻度と判定するステップと、

検知されたスピーチの前記音量に基づいて前記第 2 の頻度と判定するステップとを備える、請求項 1 7 に記載の制御方法。

【請求項 1 9】

40

通信システムにおける音声レベルを変化させるためのコンピュータによって実行されるプログラムであって、前記プログラムは、前記コンピュータに、

映像信号を取得するステップと、

音声信号を取得するステップと、

前記音声信号において少なくとも 1 人のユーザのスピーチを検知するステップと、

1 つの顔または複数の顔のいずれかを検知するために前記映像信号に対して顔認識を実施するステップと、

検知された 1 つの顔または複数の顔の前記数に基づいて、話している個人の顔の前記数を判定するステップと、

話している個人について決定された前記数に基づいて、単一話し手モードと複数話し

50

手モードとの間で切り替えるステップと、

複数の顔の前記検知に応答して前記単一話し手モードから前記複数話し手モードに切り替えるステップと、

単一の顔のみの前記検知に応答して前記複数話し手モードから前記単一話し手モードに切り替えるステップと、

前記複数話し手モードにおいて第1の頻度で前記音声信号の前記利得を調整するステップと、

前記単一話し手モードにおいて第2の頻度で前記音声信号の前記利得を調整するステップとを実行させ、

前記第1の頻度は、前記第2の頻度よりも大きい、プログラム。

10

【発明の詳細な説明】

【技術分野】

【0001】

発明の詳細な説明

この出願は、2012年2月7日に出願された米国特許出願第13／368,173の利益を主張するものであり、その全ての内容は引用によりここに援用される。

【0002】

発明の分野

本開示は、一般に、単一話し手モードと複数話し手モードとを利用した（二重モードの）会議システムのための自動利得制御（AGC）機構に関する。

20

【背景技術】

【0003】

背景

自動利得制御（AGC）機構は、個々の話し声が適切なレベルで録音されるようにマイクロフォンの利得（デジタルまたはアナログ）を設定することを目的としている。しかしながら、話している個人（単一の話し手）または話している複数の個人（複数の話し手）によってマイクロフォンが使用される場合、AGC機構は、話している個人の数を適切に判断しなければ、話している各個人の利得を適切に調整し得ない。

【0004】

ミーティングの最中に（たとえば、ヘッドセットを用いて）個人が話し、他の人（ミーティングの参加者でない人）が背後で話している場合、利得の変化によって悪影響が起これる。より具体的には、システム（たとえば、マイクロフォンシステム）は、現実／意図的に話しているのは実際には1人のみしかない場合に、話している個人が複数いると判定し、話している個人が複数いるという事に基づいて利得を変化させ得る。このため、話しているのが1人なのか複数人かを判断するだけではなく、現実／意図的に話しているのが1人なのか複数人なのかを適切に判断することができるAGC機構が必要とされている。

30

【先行技術文献】

【特許文献】

【0005】

40

【特許文献1】米国特許出願公開第2009／0002480号

【特許文献2】米国特許出願公開第2009／0015658号

【特許文献3】米国特許第6,125,343号

【特許文献4】米国特許第6,297,846号

【特許文献5】米国特許第6,457,043号

【特許文献6】米国特許第6,606,111号

【特許文献7】米国特許第6,611,281号

【特許文献8】米国特許第6,795,106号

【特許文献9】米国特許第6,804,340号

【特許文献10】米国特許第6,940,545号

50

【特許文献 1 1】米国特許第 7, 6 3 6, 4 5 3 号

【特許文献 1 2】米国特許第 7, 6 6 4, 2 4 6 号

【特許文献 1 3】米国特許第 7, 8 0 6, 6 0 4 号

【発明の概要】

【課題を解決するための手段】

【0006】

概要

この概要では、本開示におけるいくつかの局面についての基本的な理解をもたらすために、選択された概念を単純化して説明する。この概要は、開示についての概観を克明に記載するものではなく、開示される主要な要素もしくは重要な要素を特定したり、開示の範囲を規定することを意図したものでもない。この概要は、以下に記載される詳細な説明の前段階として、開示されるいくつかの概念を単に示すものである。

10

【0007】

本発明の局面は、通信システムにおける音声レベルを変化させるための制御システムを提供する。制御システムは、音声信号および映像信号を受信するための少なくとも 1 つの受信ユニットと、音声信号または映像信号のいずれかについての認識を行うことによって話している個人の数を決定するための決定ユニットと、話している個人について決定された数に基づいて音声信号の利得を調整するための利得調整ユニットとを含む。

【0008】

本発明の実施形態によれば、認識は、話している個人の数を決定するために、顔認識またはスピーチ分析のいずれかを実施することによって行われる。

20

【0009】

本発明の他の実施形態によれば、認識は、話している個人の数を決定するために、音声信号に対してスピーチ分析を実施することによって行われる。

【0010】

本発明のさらに他の実施形態によれば、認識は、映像信号に対して顔認識を実施することによって行われる。

【0011】

本発明の付加的な実施形態によれば、制御システムは、話している個人の数についての決定に基づいて単一話し手モードと複数話し手モードとの間で切り替えるための切替ユニットをさらに含む。

30

【0012】

本発明の他の実施形態によれば、顔認識は、1 つの顔または複数の顔を検知するために実施される。

【0013】

本発明のさらに他の実施形態によれば、制御システムは、検知された顔の数に基づいて単一話し手モードと複数話し手モードとの間で切り替えるための切替ユニットをさらに含む。

【0014】

本発明の付加的な実施形態によれば、切替ユニットは、複数の顔の検知に応答して単一話し手モードから複数話し手モードに切り替え、利得調整ユニットは、複数話し手モードにおいて第 1 の頻度で音声信号の利得を調整し、切替ユニットは、単一の顔のみの検知に応答して複数話し手モードから単一話し手モードに切り替え、利得調整ユニットは、単一話し手モードにおいて第 2 の頻度で音声信号の利得を調整し、第 1 の頻度は、第 2 の頻度とは異なる。

40

【0015】

本発明の他の実施形態によれば、第 1 の頻度は、第 2 の頻度よりも大きい頻度である。

本発明のさらに他の実施形態によれば、検知ユニットは、検知されたスピーチの音量を少なくとも 1 つの閾値と比較することによって、検知されたスピーチの音量が所与の音量範囲外にあるかどうかを判定し、検知ユニットは、検知されたスピーチの音量が所与の音

50

量範囲外となる事象の発生に基づいて、検知されたスピーチの音量が所定長さの時間にわたって所与の音量範囲外にあるかどうかを判定し、検知ユニットは、検知されたスピーチの音量に基づいて第1の頻度と判定し、検知ユニットは、検知されたスピーチの音量に基づいて第2の頻度と判定する。

【0016】

本発明の付加的な実施形態によれば、少なくとも1つの受信ユニットは、音声信号および映像信号の両方を有するデータのストリームを受信する。

【0017】

本発明の他の実施形態によれば、少なくとも1つの受信ユニットは、音声信号を受信するための第1の受信ユニットを含み、少なくとも1つの受信ユニットは、映像信号を受信するための第2の受信ユニットを含む。

10

【0018】

本発明のさらに他の実施形態によれば、第1の受信ユニットはマイクロフォンであり、第2の受信ユニットはカメラである。

【0019】

本発明の局面は、通信システムにおける音声レベルを変化させるための制御方法をさらに提供する。制御方法は、音声信号を受信するステップと、映像信号を受信するステップと、話している個人の数を決定するために映像信号または音声信号のいずれかについての認識を行うステップと、話している個人について決定された数に基づいて音声信号の利得を調整するステップとを含む。

20

【0020】

本発明の実施形態によれば、認識は、話している個人の数を決定するために、顔認識またはスピーチ分析のいずれかを実施することによって行われる。

【0021】

本発明の他の実施形態によれば、認識は、話している個人の数を決定するために、音声信号に対してスピーチ分析を実施することによって行われる。

【0022】

本発明の付加的な実施形態によれば、認識は、映像信号に対して顔認識を実施することによって行われる。

【0023】

30

加えて、本発明の局面は、通信システムにおいて音声レベルを変化させるための制御方法を提供する。制御方法は、映像信号を取得するステップと、音声信号を取得するステップと、音声信号において少なくとも1人のユーザのスピーチを検知するステップと、1つの顔または複数の顔のいずれかを検知するために映像信号に対して顔認識を実施するステップと、検知された1つの顔または複数の顔の数に基づいて、話している個人の数を決定するステップと、話している個人について決定された数に基づいて、単一話し手モードと複数話し手モードとの間で切り替えるステップと、複数の顔の検知に応答して単一話し手モードから複数話し手モードに切り替えるステップと、単一の顔のみの検知に応答して複数話し手モードから単一話し手モードに切り替えるステップと、複数話し手モードにおいて第1の頻度で音声信号の利得を調整するステップと、単一話し手モードにおいて第2の頻度に音声信号の利得を調整するステップとを含み、第1の頻度は、第2の頻度より大きい。

40

【0024】

本発明のさらに他の実施形態によれば、制御方法は、検知されたスピーチの音量を少なくとも1つの閾値と比較することによって、検知されたスピーチの音量が所与の音量範囲外にあるかどうかを判定するステップと、検知されたスピーチの音量が所与の音量範囲外となる事象の発生に基づいて、検知されたスピーチが所定長さの時間にわたって所与の音量範囲外にあるかどうかを判定するステップと、検知されたスピーチの音量に基づいて第1の頻度と判定するステップと、検知されたスピーチの音量に基づいて第2の頻度と判定するステップとをさらに含む。

50

【0025】

本発明の適用可能性についてのさらに他の範囲は、以下に記載される詳細な説明によって明らかになるであろう。本発明の精神および範囲内における様々な変更および修正は、詳細な説明から当業者にとって明らかなものであることから、詳細な説明および具体例は、本発明の好ましい実施形態を示す一方で、例示のみを目的としていることを理解されたい。

【0026】

本開示におけるこれらおよび他の目的、特徴、および特性は、明細書の一部を構成する添付の請求項および図面と併せて以下の詳細な説明を検討することによって当業者にとってより明らかなものとなるであろう。

【図面の簡単な説明】

【0027】

【図1】ここに記載される1つ以上の実施形態に係る会議システムの一局面を示す回路図である。

【図2】ここに記載される1つ以上の実施形態に係る映像分析方法の一局面を示すフローチャートである。

【図3】ここに記載される1つ以上の実施形態に係る音声分析方法の一局面を示すフローチャートである。

【図4】ここに記載される1つ以上の実施形態に係る会議システムの制御部（たとえば、利得制御部150）の一局面を示す回路図である。

【発明を実施するための形態】

【0028】

ここに設けられる見出しは、便宜のみを目的としたものであり、請求項に記載の発明の範囲または意味に影響を与えない。

【0029】

図面において、同じ参照符号および頭字語は、理解の容易化および利便性のために、同じまたは類似の構造または機能を伴う要素または行為を特定するものである。図面については、以下の詳細な説明の中で詳細に説明される。

【0030】

詳細な説明

本発明の様々な例が記載される。以下の記載は、これらの例の完全な理解のため、および実施可能な記載のために、具体的な詳細を提供するものである。当業者は、これらの多くの詳細が無くとも本発明を実施することができることを理解するであろう。同様に、当業者は、本発明がここに詳細に記載されていない多くの他の自明な特徴も含み得ることを理解するであろう。関連する記載を不必要に不明確化すること避けるために、いくつかの周知の構造または機能については、以下では詳細に示されない、または記載されない。

【0031】

図1は、本発明の1つ以上の実施形態に係る会議システム100の一局面を示す回路図である。図1において、会議システムは、画像取得ユニット110（または画像取得回路／回路素子110）と、スピーチ取得ユニット120（またはスピーチ取得回路／回路素子120）と、顔検知ユニット130（または顔検知回路／回路素子130）と、スピーチ検知ユニット140（またはスピーチ検知回路／回路素子140）と、利得制御部150（モードを切り替えるための切替ユニットを内部または外部に含み得る）と、映像エンコーダ160と、音声エンコーダ170と、ネットワーク180とを含む。

【0032】

画像取得ユニット110は、画像の取得、検知、および／または感知を行う装置（たとえば、カメラまたは他の類似の装置）である。画像取得ユニット110は、画像センサを含んでもよい。たとえば、画像取得ユニット110は、CCD（電荷結合素子）画像センサ、CMOS（相補型金属酸化物半導体）画像センサ、または他の類似の画像センサのような任意のタイプの画像センサであってもよい。

10

20

30

40

50

【0033】

画像取得ユニット110は、カメラを介して画像を取得、検知、および／または感知してもよく、入力信号または受信信号から画像データを受信、取得、検知、感知、および／または抽出してもよい。取得、検知、感知、および／または抽出された画像は、顔検知ユニット130に提供される。画像は、有線または無線による伝送によって顔検知ユニット130に提供されてもよい。

【0034】

スピーチ取得ユニットまたは装置120は、音声もしくはスピーチを取得および／または感知するための、音声もしくはスピーチの取得および／または感知を行う装置（たとえば、マイクロフォンまたは他の類似の装置）である。

10

【0035】

スピーチ取得ユニット120は、音声もしくはスピーチ（データまたは信号）をマイクロフォンを介して取得および／または感知してもよい、または音声のデータ／信号もしくはスピーチのデータ／信号を入力信号または受信信号から受信、取得、感知、および／または抽出してもよい。取得、感知、および／または抽出された音声もしくはスピーチ（以下では、音声データまたは音声信号という）は、有線または無線の伝送を介してスピーチ検知ユニット140に提供される。

【0036】

画像取得ユニット110およびスピーチ取得ユニット120は、2つの別個のユニットまたは装置として開示されているが、画像取得ユニット110（たとえば、カメラ）およびスピーチ取得ユニット120（たとえば、マイクロフォン）は、（いくつかまたは全ての開示される実施形態において）単一の装置として一体化、または互いに結合されてもよい。

20

【0037】

加えて、画像および音声／スピーチは、単一の装置において同時に取得、検知、感知、および／または抽出されてもよく、複数の装置から同時に取得、検知、感知、および／または抽出されてもよい。

【0038】

画像および音声／スピーチは、会議システム100へと伝送されてもよい（すなわち、併せて単一の信号として）。画像および音声／スピーチが併せて単一の信号として伝送される場合（または別個の信号として伝送される場合）、画像取得ユニット110およびスピーチ取得ユニット120は、受信信号から画像データを抽出する単一の画像抽出ユニットまたは装置110（または別個の信号として伝送される場合は、2つの画像抽出ユニット110、120）および音声もしくはスピーチを受信信号から抽出する音声もしくはスピーチ抽出ユニットまたは装置120とそれぞれ置き換えられてもよい。このため、画像抽出ユニット110は、受信信号から画像データを抽出し、抽出した画像を顔検知ユニット130に提供し、音声もしくはスピーチ抽出ユニット120は、音声もしくはスピーチを受信信号から抽出し、抽出した音声もしくはスピーチをスピーチ検知ユニット140に提供する。

30

【0039】

画像取得／抽出ユニット110およびスピーチ取得／抽出ユニット120は、2つの別個のユニットもしくは装置として開示されているが、画像取得／抽出ユニット110および音声もしくはスピーチ取得／抽出ユニット120は、（いくつかまたは全ての開示される実施形態において）単一の装置として一体化、または互いに結合されてもよい。

40

【0040】

加えて、以下のステップ210についての詳細な記載は、その全体または一部が画像取得ユニット110に対応する。このため、ステップ210の詳細がここに援用される（ステップ210に関する詳細な記載は、その全体または一部が画像取得ユニット110に援用される）。

【0041】

50

以下のステップ 3 1 0 についての詳細な記載は、その全体または一部が音声もしくはスピーチ取得／抽出ユニット 1 2 0 に対応する。このため、ステップ 3 1 0 の詳細がここに援用される（ステップ 3 1 0 に関する詳細な記載は、その全体または一部が音声もしくはスピーチ取得／抽出ユニット 1 2 0 に援用される）。

【0 0 4 2】

顔検知ユニット 1 3 0（または顔検知回路／回路素子 1 3 0）は、画像取得ユニット 1 1 0 によって取得された話し手の数を判定するために、画像内の人の数を検知する。たとえば、顔検知ユニット 1 3 0 は、画像取得ユニット 1 1 0 によって取得された人の顔を検知する。顔検知ユニット 1 3 0 は、画像取得ユニット 1 1 0 によって取得された人の頭（または人体、人）を代わりに検知することができる。顔検知ユニット 1 3 0 は、検知された顔、頭、および人などの数を利得制御部 1 5 0 に提供する。

10

【0 0 4 3】

以下のステップ 2 2 0 および／またはステップ 2 3 0 についての詳細な記載は、その全体または一部が顔検知ユニット 1 3 0 に対応する。このため、ステップ 2 2 0 および／またはステップ 2 3 0 についての詳細がここに援用される（ステップ 2 2 0 および／ステップ 2 3 0 に関する詳細な記載の全体または一部が、顔検知ユニット 1 3 0 に援用される）。

【0 0 4 4】

加えて、画像取得ユニット 1 1 0 によって顔検知ユニット 1 3 0 に提供される映像（もしくは画像）データまたは映像（もしくは画像）信号は、顔検知ユニット 1 3 0 によって映像エンコーダ 1 6 0 へと伝送される。

20

【0 0 4 5】

スピーチ検知ユニット 1 4 0（もしくはスピーチ検知回路／回路素子 1 4 0）は、取得された音声もしくはスピーチの信号またはデータ内のスピーチを検知する。スピーチ検知ユニット 1 4 0 は、検知されたスピーチもしくは音声を利得制御部 1 5 0 に提供する。スピーチ検知ユニット 1 4 0 は、能動的なスピーチと考えられるものを保持（および利得制御部に転送）し、能動的なスピーチと考えられないものを無視してもよい。たとえば、全てのスピーチが利得制御部 1 5 0 に送られ、全ての雑音が消去される。最後に、スピーチ検知ユニット 1 4 0 は、信号内の異なる声の数を検知するために使用されてもよい。

【0 0 4 6】

30

以下のステップ 3 2 0 および／またはステップ 3 3 0 についての詳細な記載は、その全体または一部が音声もしくはスピーチ検知ユニット 1 4 0 に対応する。このため、ステップ 3 2 0 および／またはステップ 3 3 0 についての詳細がここに援用される（ステップ 3 2 0 および／または 3 3 0 に関する詳細な記載は、その全体または一部が音声もしくはスピーチ検知ユニット 1 4 0 に援用される）。

【0 0 4 7】

利得制御部 1 5 0 は、検知された顔または頭の数や顔検知ユニット 1 3 0 から受け取り、検知されたスピーチ／音声の信号またはデータをスピーチ検知ユニット 1 4 0 から受け取る。受け取った情報（たとえば、検知された顔または頭の数、ならびに検知されたスピーチ／音声のデータ／信号）に基づき、利得制御部 1 5 0 は、受け取った（スピーチ取得ユニット 1 2 0 またはスピーチ検知ユニット 1 4 0 から受け取った）音声の利得を調整し、利得の調整された音声信号を音声エンコーダ 1 7 0 に出力する。

40

【0 0 4 8】

以下のステップ 2 2 0、ステップ 2 3 0、ステップ 2 4 0、ステップ 2 5 0、ステップ 3 3 0、ステップ 3 4 0、および／またはステップ 3 5 0 に記載の詳細は、その全体または一部が利得制御部 1 5 0 に対応する。このため、ステップ 2 2 0、ステップ 2 3 0、ステップ 2 4 0、ステップ 2 5 0、ステップ 3 3 0、ステップ 3 4 0、および／またはステップ 3 5 0 の詳細がここに援用される（ステップ 2 2 0、ステップ 2 3 0、ステップ 2 4 0、ステップ 2 5 0、ステップ 3 3 0、ステップ 3 4 0、および／またはステップ 3 5 0 に関する詳細な記載の全体または一部が利得制御部 1 5 0 に援用される）。

50

【0049】

映像エンコーダ160は、映像信号を顔検知ユニット130から受信し、映像信号を符号化し、符号化された映像信号を出力する。映像エンコーダ160は、デジタル映像用に映像の圧縮および／または伸長を可能にする装置である。映像エンコーダ160は、受信した映像信号に対して映像符号化を行い、映像符号化信号を生成してネットワーク180に出力する。

【0050】

音声エンコーダ170は、利得の調整された音声信号を利得制御部150から受信し、利得の調整された音声信号を符号化し、符号化された音声信号を提供する。音声エンコーダ170は、データ（音声）圧縮を可能にする装置である。音声エンコーダ170は、利得の調整された音声信号に対して音声符号化を行い、音声符号化信号を生成してネットワーク180に提供する。

10

【0051】

図2は、上記の会議システムのうちの少なくとも1つによって行われ得る例示的な映像分析方法を示すフローチャートである。図2において、映像分析方法は、映像信号を受信するステップ（ステップ210）、映像分析ステップ（ステップ220）、比較ステップ（反復型のステップとなり得るステップ230）、および／またはAGC-T値を設定するステップ（ステップ240および／または250）を含んでもよい。

【0052】

ステップ210において、会議システム100は、少なくとも画像取得ユニット110に関して詳細に記載されたように、映像信号を受信する。このため、画像取得ユニット110に関する詳細な記載がここに援用される。

20

【0053】

ステップ220において、会議システム100は、少なくとも顔検知ユニット130に関して詳細に記載されたように、受信した映像信号について映像分析を行う。このため、顔検知ユニット130に関する詳細な記載がここに援用される（顔検知ユニット130に関する詳細な記載の全体または一部がステップ220に援用される）。より具体的には、ステップ210において（たとえば、画像取得ユニット110によって）取得された、話している個人の数を判定するために、ステップ220において、画像中の人の数が（たとえば、顔検知ユニット130によって）検知される。

30

【0054】

ステップ220における顔（もしくは頭や体など）検知は、（デジタル）画像中の人間の顔（もしくは頭や体など）の場所と大きさを判定することによって行われる。たとえば、顔検知においては、顔の特徴が検知され、顔の特徴と考えられないもの（体、椅子、机、木など）は無視される。加えて、ステップ220において、検知は従来の方法によって行われてもよい。

【0055】

ステップ230において、所定の時間（より長い時間）にわたって映像中に複数の顔が存在するかどうか、および／または所定の時間（以上の時間）にわたって映像中に単一の顔が存在するかどうかについての判定が行われる（所定の時間は、1秒、2秒、3秒などであってもよい）。ステップ230は、ステップ240および／または250においてAGC閾値（AGC-T）を出力することができるように行ってもよく、これにより、単一の顔が検知されたかどうか（たとえば、話している単一の個人のみを検知する）、または複数の顔が検知されたかどうか（たとえば、話している複数の個人を検知する）についての判定をレベル分析ユニット、スピーチ検知ユニット140、および／または利得制御部150に対して知らせるための手段が提供される。

40

【0056】

AGC-T値は、2つの値（たとえば、2進値／論理値）を含むことができる。第1のAGC-T値は、複数の個人が話している旨の判定（もしくは検知）を表わす（または複数話し手モードに切り替えるための判定／命令を表わす）「真」値（たとえば、0または

50

1の値)であり、第2のAGC-T値は、単一の個人が話しているという判定(もしくは検知)を表わす(または単一話し手モードに切り替えるための判定/命令を表わす)「偽」値(たとえば、1または0の値)である。AGC-T値は、顔検知ユニット130(たとえば、ステップ230)からの単一の出力または2つの異なる出力として、レベル分析ユニット(またはスピーチ検知ユニット140および/もしくは利得制御部150)の単一の入力または2つの異なる入力に提供されてもよい。

【0057】

言い換えると、ステップ230においては、所定の時間(以上)にわたって映像中に単一の顔が検知されたかどうかの判定、または複数の顔が検知されたかどうかの判定に基づき、出力されてレベル分析ユニット、スピーチ検知ユニット140、および/または利得制御部150に提供される(たとえば、レベル分析ステップ330への入力)AGC-T値に基づいて単一話し手モードまたは複数話し手モード(複数の話し手モードともいわれる)に切り替えるかどうかについての判定を行ってもよい。

【0058】

会議システム100は、単一話し手モードまたは複数話し手モードで自動的に開始してもよい。代替的に、会議システム100は、初期化モードで開始してもよい(すなわち、特定のモードで開始するように自動的に設定されていない場合)。たとえば、ステップ230において、初期化の間(現状で単一話し手モードまたは複数の話し手モードのいずれにもない)、所定の時間(以上)にわたって映像中に単一の顔が検知されたか(否か)、または複数の顔が検知されたか(否か)についての判定が行われる(たとえば、初期化期間は、たとえば1秒、2秒、3秒などである)。初期化期間において複数の顔が映像中に検知されたと判定された場合(または単一の顔が検知されなかったと判定された場合)、利得制御部は、(たとえば、複数の話し手モードの値に対応するAGC-T値の受け取りに基づいて)システムを複数の話し手モードに設定する。初期化期間において単一の顔のみが映像中に検知されたと判定された場合(または複数の顔が検知されなかったと判定された場合、もしくは複数より小さい数の顔が検知された場合)、利得制御部は、(たとえば、単一話し手モードの値に対応するAGC-T値の受け取りに基づいて)システムを単一話し手モードに設定する。

【0059】

ステップ230において、初期化期間の後(単一話し手モードまたは複数話し手モードのいずれかの最中にある)、所定の時間(たとえば、1秒、2秒、3秒など)(以上)にわたって映像中に単一の顔が検知されたか(否か)、または複数の顔(または複数の顔より小さい数)が検知されたか(否か)についての判定が行われ、現状のモードは切り替えることができる(単一話し手モードから複数話し手モード、およびその逆もまた同様)。

【0060】

単一の顔のみが映像中に検知されたと判定された場合(または複数の顔が検知されなかったと判定された場合、もしくは複数の顔より小さい数の顔が検知されたと判定された場合)、利得制御部は、(たとえば、単一話し手モードの値に対応するAGC-T値の受け取りに基づいて)システムを単一話し手モードに切り替える。

【0061】

映像中に複数の顔が検知されたと判定された場合(または単一の顔が検知されなかったと判定された場合)、利得制御部は、(たとえば、複数の話し手モードの値に対応するAGC-T値の受け取りに基づいて)システムを複数の話し手モードに切り替える。

【0062】

利得制御部は、いずれのモードにおいてもスピーチ信号の利得を調整(変更)することが可能であってもよい。利得制御部がスピーチの利得を調整し得る頻度は、いずれのモードにおいても同じ頻度で行われてもよい。代替的な実施形態において、検知されたスピーチ信号に対して単一話し手モードにおいて与えられる利得の変化は、検知されたスピーチ信号に対して複数話し手モードにおいて与えられる利得変化と比して遅い頻度で与えられてもよい。なぜなら、複数の顔が検知された場合と比して、単一の顔が検知された場合に

10

20

30

40

50

において実際の入力信号の大きさは急速に変化しそうにないためである。たとえば、利得制御部が単一話し手モードにおいてスピーチ信号の利得を変化させる頻度は0.5秒ごととなり得る一方で、利得制御部が複数話し手モードにおいてスピーチ信号の利得を変化させる頻度は0.1秒ごととなり得る。このため、複数話し手モードにおいて利得制御部をより早く応答させることにより、利得制御は、話している複数の個人の音量をより急速に（ほぼ）同じレベルにすることができる。複数の話し手に対して急速な利得の変化を与え、結果として複数の個人の声が（ほぼ）同じレベルで聞こえるようにすることによって、一人の個人がマイクロフォンの近くに居て、他の話し手がマイクロフォンから大きく離れている場合において、システム全体が少なくとも恩恵を受け得る。

【0063】

10

代替的な実施形態において、AGC-T値を単一話し手モードに設定すると判定された場合、（利得制御部150を介した）自動利得制御は、話している個人のみ「固定」（lock onto）し得て（選択／検知された話している個人のみに対して利得を増大させる制御を行う）、話している個人の信号に所定量の（増大した）利得が与えられ得る（話している個人の利得のみを変化／増大させる、または話している個人の利得を増大させる一方で、検知／固定された話している個人を除いた全ての利得、他の検知された話をしてる複数の個人の利得、および／または検知された雑音の利得を減少させる）。

【0064】

同様に（上記の段落に記載された代替的な実施形態において）、AGC-T値を複数話し手モードに設定すると判定された場合、（利得制御部150を介した）自動利得制御は、検知された話している複数の個人に固定され（検知された話している複数の個人に対する増大した利得制御を維持する）、声（または音声）と考えられる一部および全ての信号に対して所定量の利得が与えられ得る。全ての開示された時間（全ての与えられた例において、たとえば「所定の時間」）は、たとえばユーザによる設定など、全ての実用的手段によっていつでも設定され得る、または装置によって予め定められるもしくは予め設定され得る、または以前の判定時間を用いた適合アルゴリズムに基づいて定められてもよい。

20

【0065】

ステップ230において、所定の時間にわたって複数の顔（または単一の顔など）が映像中に存在するか（否か）についての判定は、顔検知ユニット130および／または利得制御部150によって行われ得る。このため、顔検知ユニット130および／または利得制御部に関する詳細な記載がここに援用される（顔検知ユニット130および／または利得制御部に関する詳細な記載は、その全体または一部がステップ230に援用される）。図3は、上記の会議システムの少なくとも1つによって行われ得る音声分析方法の例を表わすフローチャートである。

30

【0066】

ステップ310において、会議システム100は、少なくともスピーチ取得ユニット120に関して詳細に記載したように、音声信号を受信する。このため、スピーチ取得ユニット120に関する詳細な記載がここに援用される。

【0067】

ステップ320において、会議システム100は、少なくともスピーチ検知ユニット140に関して詳細に記載したように、受信した映像信号についてのスピーチ分析を行う。このため、スピーチ検知ユニット140に関する詳細な記載がここに援用される（スピーチ検知ユニット140に関する詳細な記載の全体または一部がステップ320に援用される）。より具体的には、（たとえばスピーチ取得ユニット120によって）ステップ310において取得されたスピーチまたは音声の全てを判定するために、ステップ320において、一部および全てのスピーチ／音声（たとえば、スピーチ検知ユニット140によって）取得される。簡単にいえば、（ステップ320において）スピーチ検知ユニット140は、単に能動的なスピーチを検知してもよい。これに加え、ステップ320において、従来の方法によって検知が行われてもよい。

40

【0068】

50

代替的な実施形態において、ステップ 320 では、スピーチ検知ユニット 140 は、検知されたスピーチ／音声を使用して、話している個人の数の判定を支援してもよい（または、図 2 に示されるように映像分析に置き換えてもよい）。たとえば、複数のスピーチ取得ユニット（複数のマイクロフォン、または空間的に分離された複数のマイクロフォン）を使用することにより、話している異なる個人について受信した音声信号の時間的な遅れの差異を用いて、話している個人の数を複数の話し手信号から判定してもよい。より具体的には、ステップ 320 において、スピーチ検知ユニット 140 が話している個人の数（1 人の個人、2 人の個人など）を正確に判定できる場合、（話している単一の個人または話している複数の個人を示す）AGC-T 値をスピーチ検知ユニット 140 が（ステップ 320 において）提供すると考えられることから、図 2 に示される映像分析の全てが不要となる。

10

【0069】

能動的なスピーチの検知（のみ）に基づき、処理はステップ 320 からステップ 330 に進み得る。そうでなければ、能動的なスピーチが検知されるまで、システムはステップ 320 を継続する。

【0070】

ステップ 330 において、会議システム 100 は、少なくともスピーチ検知ユニット 140 および／または利得制御部 150 に関して詳細に記載したように、受信した音声／スピーチ信号についてのレベル分析を行う。このため、スピーチ検知ユニット 140 および／または利得制御部 150 に関する詳細な記載がここに援用される（スピーチ検知ユニット 140 および／または利得制御部 150 に関する詳細な記載の全体または一部がステップ 330 に援用される）。これに加え、ステップ 330 におけるレベル分析は、スピーチ検知ユニット 140 および／または利得制御部 150 と別個にまたは連動して作動するレベル分析ユニットによって行われてもよい。

20

【0071】

たとえば、ステップ 330（ステップ 330 a ともよばれる）において、各音声／スピーチ信号のレベル（または音量）が判定される。より具体的には、ステップ 330（またはステップ 330 a）において、検知された（能動的な）音声は、（検知されたスピーチが特定のレベルを上回るかどうか、音量が大きすぎるかどうかを示すために）上限と比較され、（検知されたスピーチの音量が特定のレベルを下回るかどうか、音量が小さすぎるかどうかを示すために）下限と比較される。

30

【0072】

ステップ 330 において（ステップ 330 b ともよばれ、ステップ 330 b はステップ 330 a の後に行われる）、特定の閾値を上回る音量または下回る音量が検知された場合、スピーチ検知ユニット 140 および／または利得制御部 150 は、検知された音量が特定の時間にわたって特定の閾値を上回るかどうか、または検知された音量が特定の閾値を下回るかどうかを判定する（たとえば、特定の時間は、1 秒、2 秒、3 秒などであってもよい）。

【0073】

最後に、ステップ 330（ステップ 330 a および 330 b）において（たとえば）利得制御部 150 によって行われる分析は、利得制御部 150 による利得変化値の判定（ステップ 340）および／または利得変化の付与（ステップ 350）の前に与えられる AGC-T 値も考慮する。

40

【0074】

検知された（能動的な）スピーチの音量が特定の時間にわたって特定の閾値よりも高いおよび／または低いという判定（のみ）に基づいて、処理はステップ 330 からステップ 340 に進んでもよい。そうでなければ、システムは、検知された（能動的な）音声特定の時間にわたって特定の範囲外となるまで（特定の時間にわたって特定の閾値を上回る、または下回る）ステップ 330 を継続する。

【0075】

50

ステップ340において、会議システム100は、少なくともスピーチ検知ユニット140および／または利得制御部150に関して詳細に記載されたように、検知された音声／スピーチ信号の各々の利得調整値についての判定を行う。このため、スピーチ検知ユニット140および／または利得制御部150に関する詳細な記載がここに援用される（スピーチ検知ユニット140および／または利得制御部150に関する詳細な記載の全体または一部がステップ330に援用される）。より具体的には、ステップ340において、複数話し手モードとすることに基づいてより急速／急激に利得を変化させるかどうか、これに対して単一話し手モードとすることに基づいて利得を急激に変化させないかどうかを判定する。このため、ステップ340において、単一話し手モードおよび複数話し手モードにおける利得変化の頻度が判定される。

10

【0076】

代替的に、単一話し手モードにある場合、ステップ340では、利得制御部が単一の個人（話し手）のスピーチ信号の利得を調整し得るように、利得制御部に対する利得調整値を判定および付与することもできる。この代替的な方法において、複数話し手モードにある場合、ステップ340では、利得制御部が個人（話し手）のスピーチ信号の各々の利得を調整し得るように、利得制御部に対する利得調整値を判定および付与することもできる。

【0077】

ステップ350において、会議システム100は、スピーチ取得ユニット120によって取得された受信音声／スピーチまたはスピーチ検知ユニット140によって検知された音声／スピーチにおけるスピーチ信号に対する利得調整を行う。

20

【0078】

ステップ350において、少なくとも利得制御部150に関して詳細に記載されたように利得調整を行うことが、ここに援用される（利得調整部150に関して記載された詳細の全体または一部がステップ350に援用される）。

【0079】

図4は、本発明の実施形態に係る利得制御部150（コンピュータ装置1000ともよばれる）の一局面を示す回路図である。

【0080】

コンピュータ装置1000（たとえば、利得制御部150）のごく基本的な構成において、コンピュータ装置1000は、一般的に1つ以上のプロセッサ1010とシステムメモリ1020とを含む。メモリバス1030は、プロセッサ1010とシステムメモリ1020との間の通信に使用することができる。

30

【0081】

所望の構成に応じて、コンピュータ装置1000の1つ以上のプロセッサ1010は、マイクロプロセッサ、マイクロコントローラ、デジタル信号プロセッサ、またはこれらの任意の組み合わせを含む任意のタイプであってもよいが、これらに限定されるものではない。プロセッサ1010は、レベル1キャッシュ1011およびレベル2キャッシュ1012などの1つ以上のレベルのキャッシングと、プロセッサコア1013と、レジスタ1014とを含むことができる。プロセッサコア1013は、算術論理演算ユニット（ALU）、浮動小数点演算ユニット（FPU）、デジタル信号処理コア（DSPコア）、またはこれらの任意の組み合わせを含むことができる。メモリ制御部1015は、プロセッサ1010と併せて使用することもでき、一部の適用例においては、メモリ制御部1015は、プロセッサ1010の内部部品とすることもできる。

40

【0082】

所望の構成に応じて、システムメモリ1020は、揮発性メモリ（RAM）、不揮発性メモリ（ROM、フラッシュメモリなど）、またはこれらの任意の組み合わせを含むタイプとすることができるが、これらに限定されるものではない。システムメモリ1020は、一般的に、オペレーティングシステム1021と、1つ以上のアプリケーション1022と、プログラムデータ1024とを含む。アプリケーション1022は、認証アルゴリ

50

ズム 1023 を含む。プログラムデータ 1024 は、サービスデータ 1025 を含む。

【0083】

コンピュータ装置 1000 は、付加的な特徴または機能、および付加的なインターフェイスを有し、基本的構成 1001 と任意の必要な装置およびインターフェイスとの間の通信を容易にすることができる。たとえば、バス／インターフェイス制御部 1040 を使用して、記憶インターフェイスバス 1041 を介した基本的構成 1001 と 1 つ以上のデータ記憶装置 1050 との間の通信を容易にすることができる。データ記憶装置 1050 は、取り外し可能な記憶装置 1051、取り外し不可能な記憶装置 1052、またはこれらの組み合わせとすることができる。取り外し可能な記憶装置および取り外し不可能な記憶装置の例をいくつか挙げると、フレキシブルディスクドライブおよびハードディスクドライブ (HDD) などの磁気ディスク装置、コンパクトディスク (CD) ドライブまたはデジタル多用途ディスク (DVD) ドライブなどの光ディスクドライブ、半導体ドライブ (SSD)、およびテープドライブなどがある。例示的なコンピュータ記憶媒体は、コンピュータ読み取り可能な指示、データ構造、プログラムモジュール、または他のデータなどの情報を記憶するための任意の方法または技術に適用される、揮発性および不揮発性媒体ならびに取り外し可能な媒体および取り外し不可能な媒体を含むことができる。

10

【0084】

システムメモリ 1020、取り外し可能な記憶装置 1051、および取り外し不可能な記憶装置 1052 は、すべてがコンピュータ記憶媒体の例である。コンピュータ記憶媒体は、RAM、ROM、EEPROM、フラッシュメモリ、もしくは他のメモリ技術、CD-ROM、デジタル多用途ディスク (DVD) もしくは他の光記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置、もしくは他の磁気記憶装置、または必要な情報を記憶することができ、他のコンピュータ装置 1000 によってアクセスすることができる任意の他の媒体を含むが、これらに限定されるものではない。任意のこのようなコンピュータ記憶装置は、コンピュータ装置 1000 の一部とすることができる。

20

【0085】

コンピュータ装置 1000 は、バス／インターフェイス制御部 840 を介して様々なインターフェイス装置 (たとえば、出力インターフェイス、周辺インターフェイス、通信インターフェイスなど) から基本的構成 1001 への通信を容易にするためのインターフェイスバス 1042 も含むことができる。例示的な出力装置 1060 は、ディスプレイまたはスピーカなどの様々な外部装置に対して 1 つ以上の A/V ポート 1063 を介して通信するように構成することができるグラフィクス処理ユニット 1061 と音声処理ユニット 1062 とを含む。例示的な周辺インターフェイス 1070 は、入力装置 (たとえば、キーボード、マウス、ペン、音声入力装置、タッチ入力装置など) または他の周辺装置 (たとえば、プリンタ、スキャナなど) などの外部装置に対して 1 つ以上の I/O ポート 1073 を介して通信するように構成することができるシリアルインターフェイス制御部 1071 またはパラレルインターフェイス制御部 1072 を含む。例示的な通信装置 1080 は、1 つ以上の通信ポート 1082 を介してネットワーク通信によって 1 つ以上の他のコンピュータ装置 1090 との通信を容易にするように構成することができるネットワーク制御部 1081 を含む。通信接続は、通信媒体の一例である。通信媒体は、コンピュータ読み取り可能な指示、データ構造、プログラムモジュール、または搬送波もしくは他の搬送機構などの変調データ信号における他のデータとして一般的に具現化され得て、任意の情報配信媒体を含む。「変調データ信号」は、信号において情報を符号化するように 1 つ以上の特徴が設定または変更される信号とすることができる。限定ではなく一例として、通信媒体は、有線ネットワークまたは直接有線接続などの有線媒体、ならびに音響、ラジオ周波数 (RF)、赤外線 (IR) および他の無線媒体などの無線媒体を含むことができる。ここで使用されるコンピュータ読み取り可能媒体の用語は、記憶媒体および通信媒体の両方を含むことができる。

30

40

【0086】

コンピュータ装置 1000 は、携帯電話、パーソナルデータアシスタント (PDA)、

50

パーソナルメディアプレーヤ装置、無線ウェブ視聴装置、パーソナルヘッドセット装置、特定用途向け装置、または上記の機能のいくつかを含むハイブリッド装置など、省スペースの持ち運び可能な（または携帯）電子機器の一部として適用することができる。コンピュータ装置1000は、ラップトップ型コンピュータおよびラップトップ型でないコンピュータの構成の両方を含むパーソナルコンピュータとして適用することもできる。

【0087】

システムの局面に関し、ハードウェアへの適用とソフトウェアへの適用との間の差異は小さなものである。ハードウェアまたはソフトウェアの使用は、概して（常にというわけではなく、ハードウェアとソフトウェアとの間の選択は、特定の状況においては重要なものとなり得る）費用対効果の妥協点を示す設計上の選択的事項である。ここに記載の処理および／またはシステムおよび／または他の技術を利用することができる様々な手段があり（たとえば、ハードウェア、ソフトウェア、および／またはファームウェア）、好ましい手段は、処理および／またはシステムおよび／または他の技術が展開される背景によって変わる。たとえば、速度および精度が最重要であると開発者が判断した場合、開発者は、主にハードウェアおよび／またはファームウェアの手段を選択してもよい。柔軟性が最重要である場合、開発者は、主にソフトウェアによる適用を選択してもよい。1つ以上の他の場合において、開発者は、ハードウェア、ソフトウェア、および／またはファームウェアのいくつかの組み合わせを選択してもよい。

【0088】

上記の詳細な記載は、装置および／または処理の様々な実施形態について、ブロック図、フローチャート、および／または例を使用して規定した。このようなブロック図、フローチャート、および／または例が1つ以上の機能および／または動作を含む限りにおいて、当業者は、このようなブロック図、フローチャート、または例の範囲内における機能および／または動作の各々が、幅広いハードウェア、ソフトウェア、ファームウェア、または実質的にこれらの任意の組み合わせによって個々および／または集合的に実施することができることを理解するであろう。

【0089】

ある実施形態において、ここに記載される主題のいくつかの部分は、特定用途向け集積回路（ASIC）、フィールドプログラマブルゲートアレイ（FPGA）、デジタル信号プロセッサ（DSP）または他の集積形態によって実施してもよい。当業者は、ここに記載される実施形態のいくつかの局面について、その全体または一部が、集積回路において、1つ以上のコンピュータ上で実行される1つ以上のコンピュータプログラムとして（たとえば、1つ以上のコンピュータシステム上で実行される1つ以上のプログラムとして）、1つ以上のプロセッサ上で実行される1つ以上のプログラムとして（たとえば、1つ以上のマイクロプロセッサ上で実行される1つ以上のプログラムとして）、ファームウェアとして、または実質的にこれらの任意の組み合わせとして均等に実施することができると認識し、回路の設計および／またはソフトウェアおよび／またはファームウェアのコードの記載がこの開示を鑑みて当業者の技術範囲に入ると認識するであろう。

【0090】

これに加え、当業者は、ここに記載される主題の機構が、様々な形態でプログラム製品として配布することができ、ここに記載される主題の例示的な実施形態は、実際に配布を行うために使用される信号担持媒体の特定のタイプに関わらず適用することができることを理解するであろう。信号担持媒体の例は、フロッピー（登録商標）ディスク、ハードディスクドライブ、コンパクトディスク（CD）、デジタル映像ディスク（DVD）、デジタルテープ、コンピュータメモリなどの記録可能型媒体、ならびにデジタルおよび／またはアナログ通信媒体などの伝送型媒体（たとえば、光ファイバケーブル、導波管、有線通信リンク、無線通信リンクなど）を含むが、これらに限定されるものではない。

【0091】

当業者は、ここに規定される方法で装置および／または処理を記載し、工学的にこのような記載の装置および／または処理をデータ処理システムに統合することは当該技術分野

10

20

30

40

50

において通常なされるものであると認識するであろう。すなわち、ここに記載の装置および／または処理の少なくとも一部は、妥当な量の実験を介してデータ処理システムに統合することができる。当業者は、一般的なデータ処理システムが、システムユニット筐体、映像表示装置、揮発性および不揮発性メモリなどのメモリ、マイクロプロセッサおよびデジタル信号プロセッサなどのプロセッサ、オペレーティングシステム、ドライバ、グラフィカルユーザインターフェイス、およびアプリケーションプログラムなどのコンピュータ上の実態のうちの1つ以上、タッチパッドまたはタッチスクリーンなどの1つ以上の相互作用装置、および／またはフィードバックループと制御モータとを含む制御システム（たとえば、位置および／速度を感知するためのフィードバック、部品および／または量を移動および／または調整するための制御モータ）を概して含むことを理解するであろう。一般的なデータ処理システムは、データコンピューティング／通信システムおよび／またはネットワークコンピューティング／通信システムに一般的に見られる任意の適切な市販の部品を利用することにより実施され得る。

10

【0092】

ここで使用される実質的に複数形および／または単数形の用語に関し、当業者は、文脈および／または適用において適切と認める場合に、複数形から単数形、および／または単数形から複数形に変換することができる。様々な単数形／複数形の置き換えが、明瞭化のためにここに明示的に規定されてもよい。

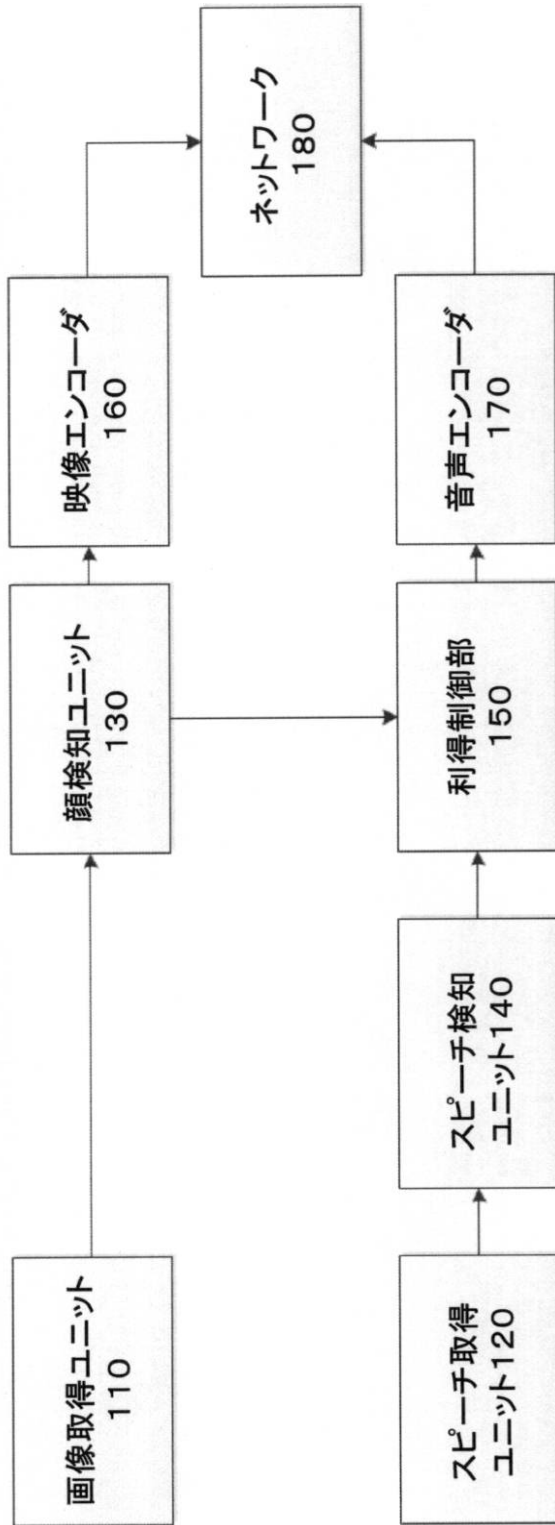
【0093】

例示的な実施形態が、本開示において示され、記載される。実施形態は、様々な他の組み合わせおよび環境において使用することができ、ここに示される発明の概念の範囲内において変更または修正することができる。いくつかのこのような変化形態は、一時的でないコンピュータ読み取り可能媒体に記憶されたプログラムを使用し、コンピュータおよび／またはコンピュータシステムが上記の方法の変化形の一部またはすべてを実行することができるようにすることを含んでもよい。このような変形例は、本発明の精神および範囲から逸脱しているとは捉えられず、当該技術において当業者によって自明となり得るこのような変形例の全ては、以下の特許請求項の範囲に含まれることが意図される。

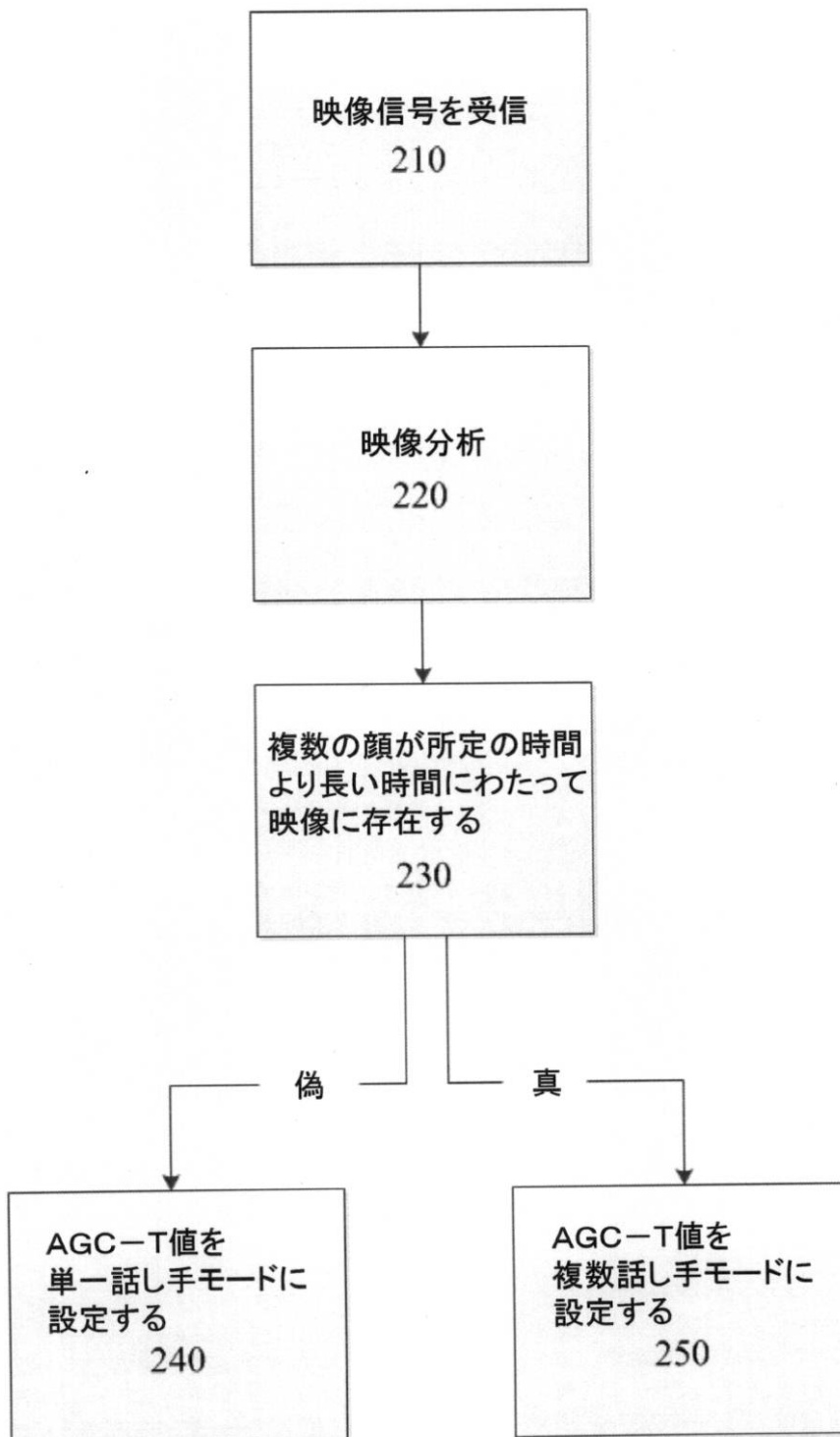
20

【図 1】

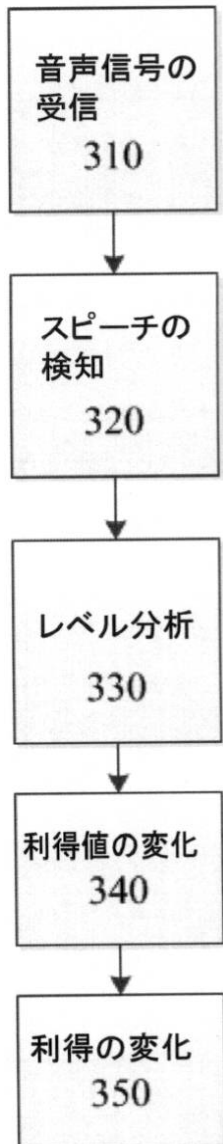
Figure 1

100

【図 2】

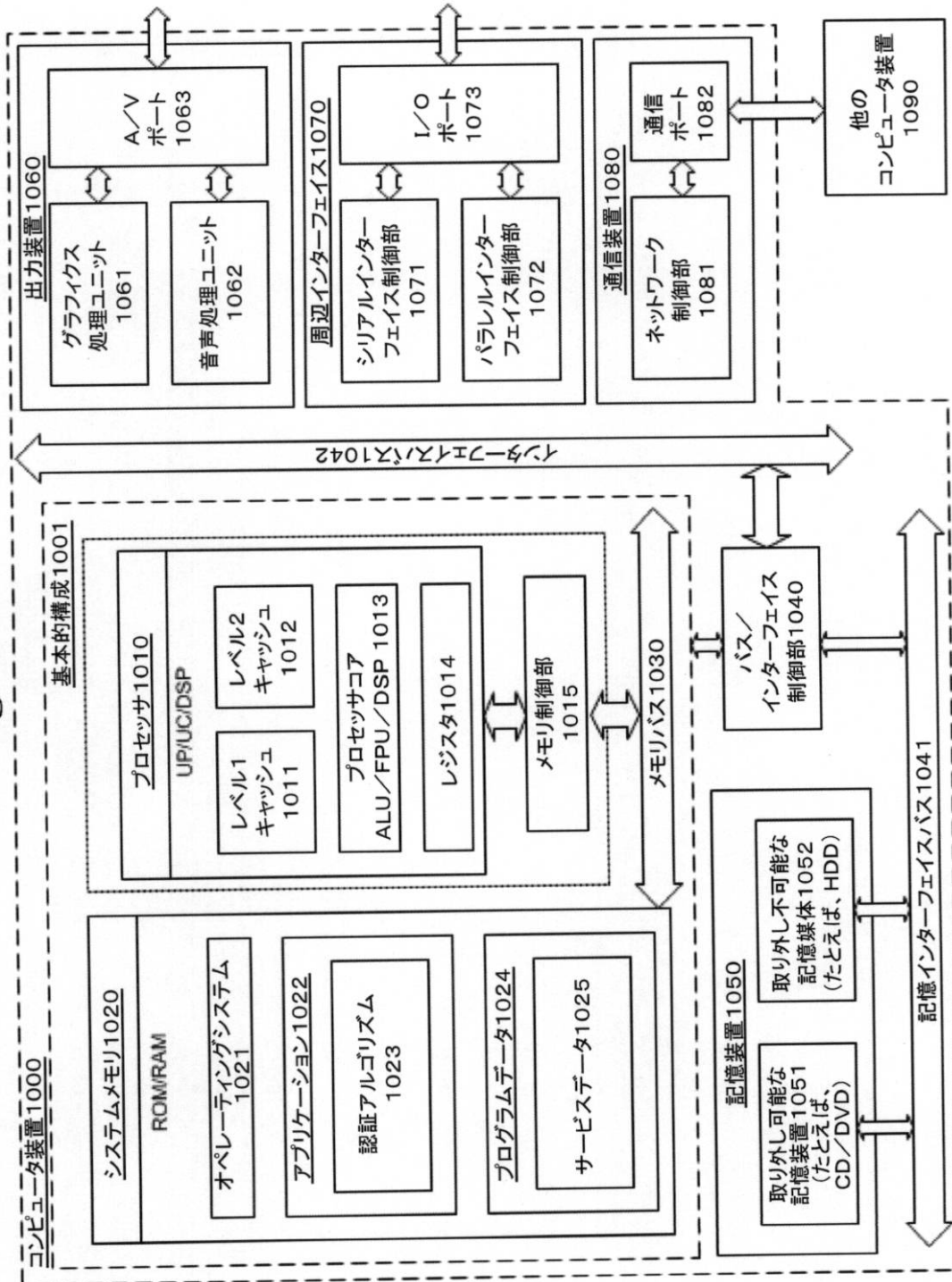
Figure 2

【図 3】

Figure 3

【図 4】

Figure 4



フロントページの続き

- (72)発明者 アンドリュー・ジョン・マクドナルド
アメリカ合衆国、94117 カリフォルニア州、サン・フランシスコ、ハイト・ストリート、6
85、アパートメント・11
- (72)発明者 ビョルン・ボルカー
スウェーデン、エスー167 33 ブロンマ、オーケリエルムスガタン、24

審査官 ▲徳▼田 賢二

- (56)参考文献 特開2000-174909 (JP, A)
特開平07-226930 (JP, A)
特表2011-512772 (JP, A)
特開2007-147762 (JP, A)
特開2006-197505 (JP, A)
特開2003-230049 (JP, A)

- (58)調査した分野(Int.Cl., DB名)
H03G 3/20