

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2012-501480

(P2012-501480A)

(43) 公表日 平成24年1月19日(2012.1.19)

(51) Int.Cl. F 1 テーマコード (参考)
G 1 0 L 15/28 (2006.01) G 1 0 L 15/28 2 1 0 A 5 D 0 1 5

審査請求 未請求 予備審査請求 未請求 (全 21 頁)

(21) 出願番号 特願2011-525265 (P2011-525265)
 (86) (22) 出願日 平成21年8月31日 (2009. 8. 31)
 (85) 翻訳文提出日 平成23年2月22日 (2011. 2. 22)
 (86) 国際出願番号 PCT/US2009/055479
 (87) 国際公開番号 W02010/025440
 (87) 国際公開日 平成22年3月4日 (2010. 3. 4)
 (31) 優先権主張番号 61/093, 220
 (32) 優先日 平成20年8月29日 (2008. 8. 29)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 12/550, 380
 (32) 優先日 平成21年8月30日 (2009. 8. 30)
 (33) 優先権主張国 米国 (US)

(71) 出願人 507055073
 マルチモーダル・テクノロジーズ・インク
 アメリカ合衆国、ペンシルバニア州 15
 217、ピッツバーグ、マーレイ・アベニ
 ュー 1710
 (74) 代理人 100079108
 弁理士 稲葉 良幸
 (74) 代理人 100109346
 弁理士 大貫 敏史
 (72) 発明者 コール, デトレフ
 アメリカ合衆国、ペンシルバニア州 15
 208、ピッツバーグ、レイノルズ スト
 リート 6954
 Fターム(参考) 5D015 LL00

最終頁に続く

(54) 【発明の名称】 ハイブリッド型音声認識

(57) 【要約】

ハイブリッド型音声認識システムは、クライアント側での音声認識エンジンおよびサーバ側での音声認識エンジンを使用して、それらの音声に対する音声認識結果を生成する。アービトレーション・エンジンが、クライアント側での音声認識結果およびサーバ側での音声認識結果のうちの、1つまたはその両方に基づいて、音声認識出力を生成する。

【選択図】 図1

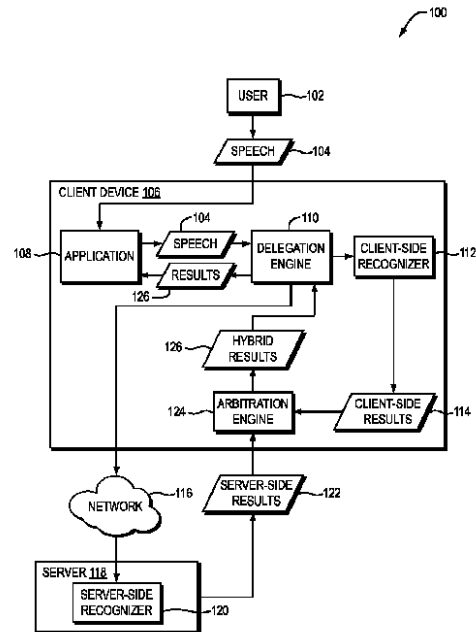


FIG. 1

【特許請求の範囲】**【請求項 1】**

クライアント装置が実行する、コンピュータで実施される方法であって、

(A) オーディオ信号に自動音声認識を適用するよう求める要求を、要求者から受信するステップと、

(B) 前記オーディオ信号を、前記クライアント装置内の第 1 の自動音声認識エンジンに供給するステップと、

(C) 前記第 1 の自動音声認識エンジンから第 1 の音声認識結果を受信するステップと、

(D) サーバ装置内の第 2 の自動音声認識エンジンが、前記クライアント装置にアクセス可能であるかどうか判定するステップと、

(E) 前記第 2 の自動音声認識エンジンが、前記クライアント装置にアクセス可能ではないと判定された場合、前記要求に応答して、前記第 1 の音声認識結果を前記要求者に提供するステップとを含む方法。

10

【請求項 2】

前記要求者は、前記クライアント装置に結合されたマシンを含む、請求項 1 に記載の方法。

【請求項 3】

前記要求者は、前記クライアント装置上で実行されるソフトウェアを含む、請求項 1 に記載の方法。

【請求項 4】

オーディオ信号に自動音声認識を適用するよう求める要求を、要求者から受信するための手段と、

前記オーディオ信号を、クライアント装置内の第 1 の自動音声認識エンジンに供給するための手段と、

前記第 1 の自動音声認識エンジンから第 1 の音声認識結果を受信するための手段と、

サーバ装置内の第 2 の自動音声認識エンジンが、前記クライアント装置にアクセス可能であるかどうか判定するための手段と、

前記第 2 の自動音声認識エンジンが、前記クライアント装置にアクセス可能ではないと判定された場合、前記要求に応答して、前記第 1 の音声認識結果を前記要求者に提供するための手段とを備える装置。

20

30

【請求項 5】

クライアント装置が実行する、コンピュータで実施される方法であって、

(A) オーディオ信号に自動音声認識を適用するよう求める要求を、要求者から受信するステップと、

(B) 前記オーディオ信号を、サーバ装置内の第 1 の自動音声認識エンジンに供給するステップと、

(C) 前記第 1 の自動音声認識エンジンから第 1 の音声認識結果を受信するステップと、

(D) 前記クライアント装置内の第 2 の自動音声認識エンジンが、前記クライアント装置にアクセス可能であるかどうか判定するステップと、

(E) 前記第 2 の自動音声認識エンジンが、前記クライアント装置にアクセス可能ではないと判定された場合、前記要求に応答して、前記第 1 の音声認識結果を前記要求者に提供するステップとを含む方法。

40

【請求項 6】

オーディオ信号に自動音声認識を適用するよう求める要求を、要求者から受信するための手段と、

前記オーディオ信号を、サーバ装置内の第 1 の自動音声認識エンジンに供給するための手段と、

前記第 1 の自動音声認識エンジンから第 1 の音声認識結果を受信するための手段と、

クライアント装置内の第 2 の自動音声認識エンジンが、前記クライアント装置にアクセス可能であるかどうか判定するための手段と、

50

前記第 2 の自動音声認識エンジンが、前記クライアント装置にアクセス可能ではないと判定された場合、前記要求に応答して、前記第 1 の音声認識結果を前記要求者に提供するための手段とを備える装置。

【請求項 7】

クライアント装置が実行する、コンピュータで実施される方法であって、

(A) オーディオ信号に自動音声認識を適用するよう求める要求を、要求者から受信するステップと、

(B) 前記オーディオ信号を、前記クライアント装置内の第 1 の自動音声認識エンジンに供給するステップと、

(C) 前記第 1 の自動音声認識エンジンから第 1 の音声認識結果を、第 1 の時点で受信するステップと、

(D) 前記オーディオ信号を、サーバ装置内の第 2 の自動音声認識エンジンに供給するステップと、

(E) 第 2 の音声認識結果が、前記第 1 の時点の後、所定の期間内に前記クライアント装置により前記第 2 の自動音声認識エンジンから受信されたかどうか判定するステップと、

(F) 前記第 2 の音声認識結果が、前記所定の期間内に前記クライアント装置によって受信された場合に、前記要求に応答して、前記第 2 の音声認識結果を前記要求者に提供するステップと、

(G) 前記第 2 の音声認識結果が、前記所定の期間内に前記クライアント装置によって受信されなかった場合に、前記要求に応答して、前記第 1 の音声認識結果を前記要求者に提供するステップとを含む方法。

10

20

【請求項 8】

(E) が、前記第 2 の音声認識結果のタイプに基づいて、前記所定の期間を選択するステップを含む、請求項 7 に記載の方法。

【請求項 9】

オーディオ信号に自動音声認識を適用するよう求める要求を、要求者から受信するための手段と、

前記オーディオ信号を、クライアント装置内の第 1 の自動音声認識エンジンに供給するための手段と、

前記第 1 の自動音声認識エンジンから第 1 の音声認識結果を、第 1 の時点で受信するための手段と、

前記オーディオ信号を、サーバ装置内の第 2 の自動音声認識エンジンに供給するための手段と、

第 2 の音声認識結果が、前記第 1 の時点の後、所定の期間内に前記クライアント装置により前記第 2 の自動音声認識エンジンから受信されたかどうか判定するための手段と、

前記第 2 の音声認識結果が、前記所定の期間内に前記クライアント装置によって受信された場合に、前記要求に応答して、前記第 2 の音声認識結果を前記要求者に提供するための手段と、

前記第 2 の音声認識結果が、前記所定の期間内に前記クライアント装置によって受信されなかった場合に、前記要求に応答して、前記第 1 の音声認識結果を前記要求者に提供するための手段とを備える装置。

30

40

【請求項 10】

クライアント装置が実行する、コンピュータで実施される方法であって、

(A) オーディオ信号に自動音声認識を適用するよう求める要求を、要求者から受信するステップと、

(B) 前記オーディオ信号を、前記クライアント装置内の第 1 の自動音声認識エンジンに供給するステップと、

(C) 前記オーディオ信号を、サーバ装置内の第 2 の自動音声認識エンジンに供給するステップと、

50

(D) 前記第1の自動音声認識エンジンから第1の音声認識結果を受信するステップと、
 (E) 前記第1の音声認識結果に関連する信頼度測定値が、所定の閾値を超えるかどうか判定するステップと、

(F) 前記信頼度測定値が前記所定の閾値を超える場合に、前記要求に応答して、前記第1の音声認識結果を前記要求者に提供するステップとを含む方法。

【請求項11】

(G) (F)の前に、前記第2の自動音声認識エンジンから第2の音声認識結果を受信するステップをさらに含み、

(F)は、前記要求者に、前記第1の音声認識結果を提供するが、第2の音声認識結果を提供しないステップを含む、請求項10に記載の方法。

【請求項12】

オーディオ信号に自動音声認識を適用するよう求める要求を、要求者から受信するための手段と、

前記オーディオ信号を、クライアント装置内の第1の自動音声認識エンジンに供給するための手段と、

前記オーディオ信号を、サーバ装置内の第2の自動音声認識エンジンに供給するための手段と、

前記第1の自動音声認識エンジンから第1の音声認識結果を受信するための手段と、

前記第1の音声認識結果に関連する信頼度測定値が、所定の閾値を超えるかどうか判定するための手段と、

前記信頼度測定値が前記所定の閾値を超える場合に、前記要求に応答して、前記第1の音声認識結果を前記要求者に提供するための手段とを備える装置。

【請求項13】

クライアント装置が実行する、コンピュータで実施される方法であって、

(A) オーディオ信号に自動音声認識を適用するよう求める要求を、要求者から受信するステップと、

(B) 前記オーディオ信号を、前記クライアント装置内の第1の自動音声認識エンジンに供給するステップと、

(C) 前記第1の自動音声認識エンジンから第1の音声認識結果を受信するステップと、

(D) 前記オーディオ信号を、サーバ装置内の第2の自動音声認識エンジンに供給するステップと、

(E) 前記第2の自動音声認識エンジンから第2の音声認識結果を受信するステップと、

(F) 前記第1の音声認識結果および前記第2の音声認識結果に基づいて、ハイブリッド型音声認識結果を生成するステップと、

(G) 前記要求に応答して、前記ハイブリッド型音声認識結果を前記要求者に提供するステップとを含む方法。

【請求項14】

(F)が、ローバー(ROVER: Recognizer Output Voting Error Reduction)を使用して、前記第1の音声認識結果と前記第2の音声認識結果を組み合わせるステップを含む、請求項13に記載の方法。

【請求項15】

前記クライアント装置が、前記第1および第2の自動音声認識エンジンのうちの一方を好ましい音声認識エンジンとして扱うように構成され、

(C)が、前記第1の音声認識結果を、第1の時点で、前記クライアント装置内のアービトレーション・エンジンにおいて受信するステップを含み、

(E)が、前記第2の音声認識結果を、前記第1の時点よりも後の第2の時点で、前記クライアント装置内の前記アービトレーション・エンジンにおいて受信するステップを含み、

(F)が、

(F)(1) 前記ハイブリッド型音声認識結果内に前記第1の認識結果を含むステップ

10

20

30

40

50

と、

(F)(2) 前記第2の自動音声認識エンジンが前記好ましい音声認識エンジンである場合にのみ、前記ハイブリッド型音声認識結果内に前記第2の音声認識結果を含むステップとを含む、請求項13に記載の方法。

【請求項16】

前記クライアント装置が、前記第1および第2の自動音声認識エンジンのうちの一方を好ましい音声認識エンジンとして扱うように構成され、

(E)が、前記第2の音声認識結果を、第1の時点で、前記クライアント装置内のアービトレーション・エンジンにおいて受信するステップを含み、

(C)が、前記第1の音声認識結果を、前記第1の時点よりも後の第2の時点で、前記クライアント装置内の前記アービトレーション・エンジンにおいて受信するステップを含み、

10

(F)が、

(F)(1) 前記ハイブリッド型音声認識結果内に前記第2の認識結果を含むステップと、

(F)(2) 前記第1の自動音声認識エンジンが前記好ましい音声認識エンジンである場合にのみ、前記ハイブリッド型音声認識結果内に前記第1の音声認識結果を含むステップとを含む、請求項13に記載の方法。

【請求項17】

(C)が、前記第1の音声認識結果を、第1の時点で、前記クライアント装置内のアービトレーション・エンジンにおいて受信するステップを含み、

20

(E)が、前記第2の音声認識結果を、前記第1の時点よりも後の第2の時点で、前記クライアント装置内の前記アービトレーション・エンジンにおいて受信するステップを含み、

(F)が、

(F)(1) 前記ハイブリッド型音声認識結果内に前記第1の認識結果を含むステップと、

(F)(2) 前記第1の音声認識結果内の任意の単語と時間的にオーバーラップしない、前記第2の音声認識結果内の単語を識別するステップと、

(F)(3) 前記ハイブリッド型音声認識結果内に、前記第2の音声認識結果からのオーバーラップしていない単語のみを含むステップとを含む、請求項13に記載の方法。

30

【請求項18】

(E)が、前記第2の音声認識結果を、第1の時点で、前記クライアント装置内のアービトレーション・エンジンにおいて受信するステップを含み、

(C)が、前記第1の音声認識結果を、前記第1の時点よりも後の第2の時点で、前記クライアント装置内の前記アービトレーション・エンジンにおいて受信するステップを含み、

(F)が、

(F)(1) 前記ハイブリッド型音声認識結果内に前記第2の認識結果を含むステップと、

40

(F)(2) 前記第2の音声認識結果内の任意の単語と時間的にオーバーラップしない、前記第1の音声認識結果内の単語を識別するステップと、

(F)(3) 前記ハイブリッド型音声認識結果内に、前記第1の音声認識結果からのオーバーラップしていない単語のみを含むステップとを含む、請求項13に記載の方法。

【請求項19】

(F)が、

(F)(1) 前記ハイブリッド型音声認識結果内に前記第1の認識結果を含むステップと、

(F)(2) 前記第1の認識結果を、前記ハイブリッド型音声認識結果内の前記第2の認識結果で置き換えるステップとを含む、請求項13に記載の方法。

50

【請求項 20】

(F)が、

(F)(1)前記ハイブリッド型音声認識結果内に前記第2の認識結果を含むステップと、

(F)(2)前記第2の認識結果を、前記ハイブリッド型音声認識結果内の前記第1の認識結果で置き換えるステップとを含む、請求項13に記載の方法。

【請求項 21】

オーディオ信号に自動音声認識を適用するよう求める要求を、要求者から受信するための手段と、

前記オーディオ信号を、クライアント装置内の第1の自動音声認識エンジンに供給するための手段と、

前記第1の自動音声認識エンジンから第1の音声認識結果を受信するための手段と、

前記オーディオ信号を、サーバ装置内の第2の自動音声認識エンジンに供給するための手段と、

前記第2の自動音声認識エンジンから第2の音声認識結果を受信するための手段と、

前記第1の音声認識結果および前記第2の音声認識結果に基づいて、ハイブリッド型音声認識結果を生成するための手段と、

前記要求に应答して、前記ハイブリッド型音声認識結果を前記要求者に提供するための手段とを備える装置。

【発明の詳細な説明】**【背景技術】****【0001】**

音声を変換することや、音声に应答してコンピュータの操作を制御することなどの機能を実行するための、様々な自動音声認識装置(ASR)が存在する。自動音声認識装置の用途によっては、エンド・ユーザに应答して出力されるように、他の用途よりも短いターンアラウンド・タイム(音声を発話してから音声認識装置が出力を生成するまでの時間)が求められる。例えば、オンスクリーン・カーソルの動きを制御することなど「ライブ」の音声認識用途に使用される音声認識装置は、医学報告の書き起こしを生成するのに使用される音声認識装置よりも短いターンアラウンド・タイム(「应答時間」とも呼ばれている)を必要とすることがある。

【0002】

所望のターンアラウンド・タイムは、例えば、音声認識装置によって処理される音声発話の内容に依存する。例えば、「ウィンドウを閉じよ」など、短い命令および制御の発話においては、約500msのターンアラウンド・タイムは、エンド・ユーザにとって反応が遅く感じられることがある。対照的に、ユーザがテキストに書き起こしたいと望む長い口述文においては、1000msの应答時間は、エンド・ユーザにとって許容できることがある。実際、後者の場合では、ユーザは、その音声に应答してテキストが直ちに表示されることにより、普通ならば、その音声に割り込まれていると感じることがあるので、より長い方の应答時間を好むことがある。段落全体など、口述されるより長い文節では、何秒間にもわたるより長い应答時間でさえも、エンド・ユーザによって許容できることがある。

【0003】

典型的な従来技術の音声認識システムでは、認識確度を維持しながら应答時間を改善するには、音声認識の実行専用であるコンピューティング資源(処理サイクルおよび/またはメモリ)を増大させる必要がある。同様に、典型的な従来技術の音声認識システムでは、通常、音声認識の実行専用であるコンピューティング資源を増大させることによるのみ、应答時間を犠牲にすることなく認識確度を向上させることができる。これらトレードオフの結果の一例として、所与の音声認識装置を、デスクトップ・コンピュータのプラットフォームから、コンピューティング資源の少ない携帯電話などの組み込みシステムにポーティングするときには、同じ应答時間を維持しようとするれば、通常、認識確度が犠牲

10

20

30

40

50

になるはずである。

【0004】

組込み式装置との関連でこれら資源の制約条件を克服するための知られた技法の1つは、音声認識処理の負担のいくらかまたはすべてを、組込み式装置から離れて配置され、この組込み式装置よりもコンピューティング資源がはるかに多い音声認識サーバに委ねることである。この状況でユーザが組込み式装置に発話すると、この組込み式装置は、それ自体のコンピューティング資源を使用して音声を認識しようとはしない。代わりに、この組込み式装置は、音声（または、処理された形式の音声）を、ネットワーク接続を介して音声認識サーバに伝送し、この音声認識サーバは、そのより豊富なコンピューティング資源を使用して音声を認識し、したがって、この組込み式装置が同じ確度で生成することのできる場合よりも迅速に認識結果を生成する。次いで、音声認識サーバは、ネットワーク接続を介して、この結果を組込み式装置に伝送して戻す。理想的には、この技法は、組込み式装置のみを使用して他の方法で実現可能になる場合よりも迅速に、非常に正確な音声認識結果を生成する。

10

【0005】

しかし実際には、このサーバ側での音声認識技法の使用は、様々な欠点がある。具体的には、サーバ側での音声認識は、高速で高信頼のネットワーク接続の可用性に依存するので、必要なときにこうした接続が利用可能でない場合、この技法は破綻する。例えば、十分に広帯域でないネットワーク接続を使用することにより、サーバ側での音声認識によって可能になる潜在的な速度の増加が無効になることがある。一例として、遠隔サーバに対するHTTPコールの典型的なネットワーク待ち時間は、100ms～500msの範囲になることがある。発話データが、発話されて500ms後に音声認識サーバに到達する場合、そのサーバは、命令および制御のアプリケーションが必要とする最小限のターンアラウンド・タイム(500ms)を満足させるのに十分なだけ迅速に結果を生成することが不可能になる。その結果、最速の音声認識サーバでさえ、低速のネットワーク接続とともに使用される場合には、生成される結果が遅く感じられることになる。

20

【発明の概要】

【発明が解決しようとする課題】

【0006】

したがって、必要となるのは、埋込み式装置において、それら装置に必要とされるターンアラウンド・タイム内で、ただし低遅延で可用性の高いネットワーク接続を必要とすることなく、高品質の音声認識結果を生成するための改良された技法である。

30

【課題を解決するための手段】

【0007】

ハイブリッド型音声認識システムは、クライアント側での音声認識エンジンおよびサーバ側での音声認識エンジンを使用して、それらの音声に対する音声認識結果を生成する。アービトレーション・エンジンが、クライアント側での音声認識結果およびサーバ側での音声認識結果のうちの、1つまたはその両方に基づいて、音声認識出力を生成する。

【0008】

本発明の様々な態様および実施形態の他の特徴ならびに利点が、以下の説明および特許請求の範囲から明らかになる。

40

【図面の簡単な説明】

【0009】

【図1】本発明の一実施形態による、音声認識システムのデータ流れ図である。

【図2】本発明の一実施形態による、図1のシステムによって実行される方法の流れ図である。

【図3A】本発明の様々な実施形態による、ハイブリッド型音声認識出力を生成するためにアービトレーション・エンジンが実行する方法の流れ図である。

【図3B】本発明の様々な実施形態による、ハイブリッド型音声認識出力を生成するためにアービトレーション・エンジンが実行する方法の流れ図である。

50

【図 3 C】本発明の様々な実施形態による、ハイブリッド型音声認識出力を生成するためにアービトレーション・エンジンが実行する方法の流れ図である。

【図 3 D】本発明の様々な実施形態による、ハイブリッド型音声認識出力を生成するためにアービトレーション・エンジンが実行する方法の流れ図である。

【図 3 E】本発明の様々な実施形態による、ハイブリッド型音声認識出力を生成するためにアービトレーション・エンジンが実行する方法の流れ図である。

【図 4 A】本発明の様々な実施形態による、複数の音声認識エンジンからのオーバーラップする認識結果を生成するために音声認識システムが実行する方法の流れ図である。

【図 4 B】本発明の様々な実施形態による、複数の音声認識エンジンからのオーバーラップする認識結果を生成するために音声認識システムが実行する方法の流れ図である。

【図 4 C】本発明の様々な実施形態による、複数の音声認識エンジンからのオーバーラップする認識結果を生成するために音声認識システムが実行する方法の流れ図である。

【図 4 D】本発明の様々な実施形態による、複数の音声認識エンジンからのオーバーラップする認識結果を生成するために音声認識システムが実行する方法の流れ図である。

【図 4 E】本発明の様々な実施形態による、複数の音声認識エンジンからのオーバーラップする認識結果を生成するために音声認識システムが実行する方法の流れ図である。

【図 4 F】本発明の様々な実施形態による、複数の音声認識エンジンからのオーバーラップする認識結果を生成するために音声認識システムが実行する方法の流れ図である。

【発明を実施するための形態】

【0010】

図 1 を参照すると、本発明の一実施形態による音声認識システム 100 のデータ流れ図が示してある。図 2 を参照すると、本発明の一実施形態による、図 1 のシステム 100 によって実行される方法 200 の流れ図が示してある。

【0011】

クライアント装置 106 のユーザ 102 が発話し、それにより、クライアント装置に音声 104 を供給する（ステップ 202）。クライアント装置 106 は、デスクトップ・コンピュータもしくはラップトップ・コンピュータ、携帯電話、携帯型情報端末（PDA）、または電話など、どんな装置でもよい。しかし、本発明の各実施形態は、低速なプロセッサもしくは少量のメモリを有するコンピュータまたはモバイル・コンピューティング装置など資源が限られたクライアント、または資源を必要とするソフトウェアを走らせるコンピュータとともに用いると特に有用である。装置 106 は、サウンド・カードに接続されたマイクロフォンなどを介して、任意のやり方でユーザ 102 から音声 104 を受け取ってもよい。音声 104 は、コンピュータ読取り可能な媒体に具体的に格納され、かつ/またはネットワーク接続もしくは他のチャネルを介して伝送されるオーディオ信号内に取り込んでよい。

【0012】

クライアント装置 106 は、トランスクリプション・アプリケーションまたは音声 104 を認識する必要のある他のアプリケーションなど、アプリケーション 108 を含む。アプリケーション 108 は、音声 104 を委譲エンジン 110 に伝送する（ステップ 204）。あるいは、アプリケーション 108 は、何らかのやり方で音声 104 を処理し、音声 104 の処理済みバージョンまたはこの音声 104 から得られる他のデータを、委譲エンジン 110 に供給してもよい。委譲エンジン 110 自体は、認識するために音声を送送するのに備えて、（アプリケーションによって音声に対して実行される任意の処理に加えて、またはその代わりに）音声 104 を処理してもよい。

【0013】

委譲エンジン 110 は、従来の自動音声認識エンジンが提示するインターフェースと同じインターフェースをアプリケーション 108 に提示してもよい。その結果、アプリケーション 108 は、音声 104 を従来の音声認識エンジンに直接供給するのと同じ方式で、音声 104 を委譲エンジン 110 に供給してもよい。したがって、アプリケーション 108 の作成者は、委譲エンジン 110 それ自体が従来の音声認識エンジンではないことを知

10

20

30

40

50

る必要がない。以下で詳細に述べるように、委譲エンジン 110 はまた、従来の音声認識エンジンと同じようにして、音声認識結果をアプリケーション 108 に戻す。したがって、委譲エンジン 110 は、アプリケーション 108 の観点からすれば、従来の音声認識エンジンと同じ機能を実行するよう見える。

【0014】

委譲エンジン 110 は、音声 104 (または、音声 104 の処理済みの形態または音声 104 から得られる他のデータ) を、クライアント装置 106 内のクライアント側での自動音声認識エンジン 112 に供給するとともに (ステップ 206)、ネットワーク 116 を介して遠隔に配置されているサーバ 118 内のサーバ側での自動音声認識エンジン 120 に供給する (ステップ 208)。サーバ 118 は、クライアント装置よりもはるかに豊富なコンピューティング資源を有するコンピューティング装置でもよい。

10

【0015】

クライアント側での音声認識装置 112 およびサーバ側での音声認識装置 120 は、従来の音声認識装置でもよい。しかし、クライアント側での音声認識装置 112 およびサーバ側での音声認識装置 120 は、互いに異なってもよい。例えば、サーバ側での音声認識装置 120 は、クライアント側での音声認識装置 112 が使用するコンピューティング資源よりも豊富なコンピューティング資源を必要とする、より複雑な音声認識モデルを使用してもよい。他の例として、音声認識装置 112 および 120 のうちの一方は発話者から独立していてもよく、もう一方はユーザ 102 の声にも適合してもよい。クライアント 106 およびサーバ 118 のコンピューティング資源の差と、音声認識装置 112 および 120 それ自体の差と、ネットワーク 116 を介してサーバ側での認識装置 120 からの結果をクライアント装置 106 に戻さなければならないという事実との組合せにより、クライアント側での認識装置 112 およびサーバ側での認識装置 120 は、互いに異なる応答時間を有してもよく、それによって待ち時間が生じるが、これはクライアント側での認識装置 112 が原因ではない。

20

【0016】

クライアント側での認識装置 112 とサーバ側での認識装置 120 が互いに異なってもいなくても、責任は、こうした認識装置 112 と 120 の間で様々な方式で分割してもよい。例えば、命令および制御の音声認識のために、もっぱらクライアント側での音声認識装置 112 を使用してもよいが、命令および制御の認識と口述認識の両方のために、サーバ側での音声認識装置 112 を使用してもよい。他の例として、クライアント側での認識装置 112 は、クライアント装置 106 でのプロセッサ時間の所定の最大パーセンテージまで利用できるようにするだけでよい。委譲エンジン 110 は、クライアント側での認識装置 112 およびサーバ側での認識装置 120 に、それぞれの責任に従って適切な音声を送信するように構成してもよい。

30

【0017】

クライアント側での認識装置 112 は、音声 104 に基づいたテキストなどの音声認識結果 114 を生成する (ステップ 210)。同様にして、サーバ側での認識装置 120 は、音声 104 に基づいたテキストなどの音声認識結果 122 を生成する (ステップ 212)。結果 114 は、最適な単語候補のセット、それらの単語に関連する信頼度測定値、音声認識エンジンが通常もたらす他の出力など、他の情報を含んでもよい。

40

【0018】

クライアント側での結果 114 およびサーバ側での結果 122 は、互いに異なってもよい。クライアント側での認識装置 112 およびサーバ側での認識装置 120 は両方とも、それぞれ、その結果 114 および 122 を、クライアント装置 106 内のアービトレーション・エンジン 124 に提供する。アービトレーション・エンジン 124 は、結果 114 および 122 のうちの、一方または両方を解析して、この 2 つの結果 114 および 122 のうちのいずれを (結果 126 として) 委譲エンジン 110 に供給すべきか決定する (ステップ 214)。以下でより詳細に述べるように、アービトレーション・エンジン 124 は、結果 114 と 122 の両方を受信した後、または、結果 114 および 122 のうちの

50

一方を受信して後ただし他方を受信しない場合、のいずれかにおいてステップ 2 1 4 を実行してもよい。したがって、一般に、アービトレーション・エンジン 1 2 4 は、クライアント側での結果 1 1 4 および / またはサーバ側での結果 1 2 2 に基づいて出力 1 2 6 を生成する。

【 0 0 1 9 】

委譲エンジン 1 1 0 は、選択された結果 1 2 6 を、要求アプリケーション 1 0 8 に戻す (ステップ 2 1 6)。結果として、要求アプリケーション 1 0 8 は、委譲エンジン 1 1 0 が単一の統合された音声認識エンジン 1 1 0 であると仮定した場合と同様に、委譲エンジン 1 1 0 から音声認識結果 1 2 6 を受信して戻す。すなわち、委譲エンジン 1 1 0 およびアービトレーション・エンジン 1 2 4 が実行する動作の詳細は、要求アプリケーション 1 0 8 には隠されている。

10

【 0 0 2 0 】

アービトレーション・エンジン 1 2 4 は、様々な技法のうち任意のものを使用して、クライアント側での結果 1 1 4 およびサーバ側での結果 1 2 2 のうちのいずれを委譲エンジン 1 1 0 に供給するのが選択してもよい。例えば、図 3 A の方法 3 0 0 で示すように、サーバ側での認識装置 1 2 0 がネットワークを介してアクセス不能である場合 (例えば、クライアント 1 0 6 とネットワーク 1 1 6 の間の接続がダウンしている場合) (ステップ 3 0 4 ~ 3 0 6)、アービトレーション・エンジン 1 2 4 は、クライアント側での結果 1 1 4 が利用可能になるとすぐ、それらの結果 1 1 4 を選択してもよい (ステップ 3 0 2)。

【 0 0 2 1 】

逆に言えば、図 3 B の方法 3 1 0 で示すように、クライアント側での認識装置 1 1 2 がアクセス不能である場合 (ステップ 3 1 4 ~ 3 1 6)、アービトレーション・エンジン 1 2 4 は、サーバ側での結果 1 2 2 が利用可能になるとすぐ、それらの結果 1 2 2 を選択してもよい (ステップ 3 1 2)。例えば、優先度の高い CPU タスクがクライアント装置 1 0 6 上で実行されている結果として、クライアント側での認識装置 1 1 2 が使用不可になってしまった場合、このことが起きる可能性がある。

20

【 0 0 2 2 】

他の例として、また、サーバ側での認識装置 1 2 0 が、クライアント側での認識装置 1 1 2 よりも平均して高品質の認識結果を提供すると仮定すると、アービトレーション・エンジン 1 2 4 は、クライアント側での認識装置の結果 1 1 4 が利用可能になって後、所定の待ち時間よりも遅れることなくサーバ側での認識装置の結果 1 2 2 が利用可能になる場合には、それらの結果 1 2 2 を選択してもよい。すなわち、図 3 C の方法 3 2 0 で示すように、クライアント側での認識装置の結果 1 1 4 が利用可能になると (ステップ 3 2 2)、アービトレーション・エンジン 1 2 4 は、サーバ側での結果 1 2 2 が、所定の待ち時間が経過する (ステップ 3 2 6) 前に受信される (ステップ 3 2 4) 場合にのみ、それらの結果を戻してもよい (ステップ 3 3 0)。そのときまでにサーバ側での結果 1 2 2 が利用可能でない場合、アービトレーション・エンジン 1 2 4 は、クライアント側での結果 1 1 4 を戻してもよい (ステップ 3 2 8)。

30

【 0 0 2 3 】

所定の待ち時間は、任意の方式で選択してもよい。例えば、所定の待ち時間は、認識結果のタイプに依存してもよい。例えば、方法 3 2 0 によって命令および制御の語法に適用される所定の待ち時間は、口述の語法に適用される所定の待ち時間よりも短くなるように選択してもよい。ほんの一例として、5 0 0 m s の所定の待ち時間を命令および制御の語法に適用してもよく、1 0 0 0 m s の所定の待ち時間を口述の語法に適用してもよい。

40

【 0 0 2 4 】

さらに他の例として、また、図 3 D の方法 3 4 0 で示すように、サーバ側での認識装置 1 2 0 が、クライアント側での認識装置 1 1 2 よりも平均して高品質の認識結果を提供すると仮定しても、クライアント側での認識装置の結果 1 1 4 に関連する信頼度測定値がある所定の閾値を超える場合 (ステップ 3 4 4)、アービトレーション・エンジン 1 2 4 は、クライアント側での認識装置の結果 1 1 4 が利用可能になるとすぐ (ステップ 3 4 2)

50

、それらの結果 1 1 4 を選択してもよい（ステップ 3 4 6）。

【 0 0 2 5 】

アービトレーション・エンジン 1 2 4 は、クライアント側での認識装置 1 1 2 およびサーバ側での認識装置 1 2 0 それぞれによって生成される結果 1 1 4 および 1 2 2 のうちの、一方または他方を「選択すること」に限定されない。むしろ、例えば、図 3 E の方法 3 5 0 で示すように、アービトレーション・エンジン 1 2 4 は、結果 1 1 4 および 1 2 2 を受信してもよく（ステップ 3 5 2 および 3 5 4）、また、それらの結果 1 1 4 および 1 2 2 を組み合わせても、またはそうでなければ様々な方式で処理して（ステップ 3 5 6）、要求アプリケーション 1 0 8 に戻される出力 1 2 6 を生成してもよい（ステップ 3 5 8）。例えば、アービトレーション・エンジン 1 2 4 は、ローバー（ROVER: Recognizer Output Voting Error Reduction）と呼ばれるよく知られた技術を使用して、または他の技法を使用して、結果 1 1 4 および 1 2 2 を組み合わせると出力 1 2 6 を生成してもよい。

10

【 0 0 2 6 】

アービトレーション・エンジン 1 2 4 は、図 3 A ~ 3 E に関してこれまでに開示された各技法を他の技法とどのような組合せで組み合わせてもよい。例えば、図 3 C のステップ 3 2 2 の後に方法 3 4 0 のステップ 3 4 4 および 3 4 6 を実行し、ステップ 3 4 4 での信頼度測定値が閾値を超えない場合に図 3 C のステップ 3 2 4 に進むことにより、図 3 D の方法 3 4 0 を図 3 C の方法 3 2 0 と組み合わせてもよい。

【 0 0 2 7 】

図 4 A の方法 4 0 0 で示すように、認識装置 1 1 2 および 1 2 0 のうちの一方からの結果が、もう一方の認識装置からの結果と時間的にオーバーラップする可能性がある。例えば、音声 1 0 4 は持続時間が 5 秒であり、音声 1 0 4 の最初の 2 秒に対してクライアント側での認識装置 1 1 2 が高信頼度の結果 1 1 4 を生成する（ステップ 4 0 2）と仮定する。結果 1 1 4 の高信頼度測定の結果として、アービトレーション・エンジン 1 2 4 は、それらの結果 1 1 4 を委譲エンジン 1 1 0 に渡してもよく、委譲エンジン 1 1 0 は、サーバ側での結果 1 2 2 が利用可能になる前に、それらの結果 1 1 4 をコミットする（すなわち、アプリケーション 1 0 8 に戻される結果 1 2 6 の中に結果 1 1 4 を含む）（ステップ 4 0 4）。次いで、音声 1 0 4 のそれら 5 秒のうちの何秒か、または 5 秒すべてに対するサーバ側での認識結果 1 2 2 が利用可能になるとき、それらの結果 1 2 2 のうちのいくつかまたはすべてが、クライアント側での結果 1 1 4 のいくつかまたはすべてと競合する（時間的にオーバーラップする）ことがある（ステップ 4 0 6）。アービトレーション・エンジン 1 2 4 は、こうしたオーバーラップに回答して、処置をとってもよい（ステップ 4 0 8）。

20

30

【 0 0 2 8 】

例えば、図 4 B の方法 4 1 0 で示すように、クライアント側での結果 1 1 4 とサーバ側での結果 1 2 2 が、何らかの所定の閾値期間（例えば 1 0 0 m s）未満だけオーバーラップする場合（ステップ 4 1 2）、アービトレーション・エンジン 1 2 4 は、結果 1 1 4 と 1 2 2 がオーバーラップしていないものと考え、図 3 A ~ 3 E に関連して前述した方式のうちの任意の方式でそれらの結果を処理してもよい。そうでない場合、アービトレーション・エンジン 1 2 4 は、結果 1 1 4 と 1 2 2 がオーバーラップしているものと考え、したがって、以下の例で説明する方式などで、それらの結果を処理してもよい（ステップ 4 1 6）。

40

【 0 0 2 9 】

例えば、図 4 B の方法 4 2 0 で示すように、アービトレーション・エンジン 1 2 4 は、各認識装置のうち的一方（例えば、サーバ側での認識装置 1 2 0）が、もう一方の認識装置よりも好ましいと考えてもよい。この場合、好ましくない認識装置からの結果（例えば、クライアント側での結果 1 1 4）が最初に到達し（ステップ 4 2 2）、最初にコミットされ（ステップ 4 2 4）、次いで、前にコミットされた好ましくない結果とオーバーラップする、好ましい認識装置からの結果（例えば、サーバ側での結果 1 2 2）が到達する場合（ステップ 4 2 8）、アービトレーション・エンジン 1 2 4 は、好ましい結果（例えば、サーバ側での結果 1 2 2）もまたコミットしてよい（すなわち、ハイブリッドの結果 1 2 6 の中に含んでもよい）（ステップ 4 3 0）。これにより、音声 1 0 4 のある部分が 2 度

50

コミットされるが、これにより、好ましい認識装置の結果を破棄する場合よりも望ましい結果が得られることがある。後に受信された結果が好ましい認識装置からのものではない場合、それらの結果は、コミットするのではなく破棄してもよい（ステップ 4 3 2）。

【 0 0 3 0 】

さらに他の例として、図 4 D の方法 4 4 0 で示すように、好ましい認識装置からの結果（例えば、サーバ側での結果 1 2 2）が最初に到達し（ステップ 4 4 2）、最初にコミットされ（ステップ 4 4 4）、次いで、前にコミットされた好ましい結果とオーバラップする、好ましくない認識装置からの結果（例えば、クライアント側での結果 1 1 4）が到達する場合（ステップ 4 4 6 および 4 4 8）、アービトレーション・エンジン 1 2 4 は、好ましくない結果を破棄してもよい（ステップ 4 5 0）。そうでない場合、アービトレーション・エンジン 1 2 4 は、後に受信した結果をコミットしてもよく、また別のやり方でそれら进行处理してもよい（ステップ 4 5 2）。

10

【 0 0 3 1 】

より一般には、図 4 E（図 4 A のステップ 4 0 8 の一実施形態を表す）に示すように、アービトレーション・エンジン 1 2 4 が、（そのまたはそれとは異なる）音声認識装置から受信された、前にコミットされた任意の結果とオーバラップする認識結果を受信する場合、アービトレーション・エンジン 1 2 4 は、（両方の認識結果内のそれぞれの単語に関連するタイム・スタンプを使用して）古い認識結果からの単語と時間的にオーバラップする新規の認識結果からの単語を無視し（ステップ 4 6 2）、次いで、新規の認識結果からの残りの（オーバラップしていない）単語をコミットしてもよい（ステップ 4 6 4）。

20

【 0 0 3 2 】

さらに他の例として、図 4 F（図 4 A のステップ 4 0 8 の一実施形態を表す）で示すように、アービトレーション・エンジン 1 2 4 が、（そのまたはそれとは異なる）音声認識装置から受信された、前にコミットされた任意の結果とオーバラップする認識結果を受信する場合、アービトレーション・エンジン 1 2 4 は、新規に受信した結果を使用して、前にコミットされた結果を更新してもよい（ステップ 4 7 2）。例えば、アービトレーション・エンジン 1 2 4 は、新規に受信した結果に関連する信頼度測定値が、前にコミットされた結果に関連する信頼度測定値を超えるかどうか判定してもよく（ステップ 4 7 4）、超える場合には、前にコミットされた結果を新規に受信した結果で置き換えてもよい（ステップ 4 7 6）。

30

【 0 0 3 3 】

本発明の各実施形態には、様々な利点がある。一般に、本発明の各実施形態により、資源が限られている携帯電話などのクライアント側での装置が、可用性の高い、広帯域のネットワーク接続を必要とすることなく、所定のターンアラウンド・タイム要求内で高品質の音声認識結果を得ることができるようになる。本明細書において開示する各技法は、クライアント側での認識装置 1 1 2 とサーバ側での認識装置 1 1 8 の両方を使用して、それらの認識装置のいずれかが個別に生成できる場合よりも良好な結果を生成する、ハイブリッド型音声認識エンジンを効果的に生成する。より具体的には、結果として生じるハイブリッド型の結果は、システムの可用性、認識品質、および応答時間に関して、構成している認識装置 1 1 2 および 1 2 0 のいずれかから個別に得ることができる場合よりも良好な動作特性を有することができる。

40

【 0 0 3 4 】

例えば、ネットワーク 1 1 6 の可用性が時間の経過とともに変動しても、またクライアント装置 1 0 6 の CPU に対する処理負荷が時間の経過とともに変動しても、本明細書において開示する各技法を使用して、ユーザのターンアラウンド・タイム要求を満足させることができる。こうした柔軟性は、アービトレーション・エンジン 1 2 4 が、クライアント側での認識装置 1 1 2 およびサーバ側での認識装置 1 2 0 のターンアラウンド・タイムにおける変化、および時間で変化する他の要因にตอบสนองする変化に対処することができる結果として生じる。それにより、本発明の各実施形態は、ネットワークの速度が落ち、またはネットワークが利用不能になると破綻する、従来のサーバ側での音声認識技法を上回る

50

明確な利益をもたらす。

【 0 0 3 5 】

本発明の各実施形態に従って実施されるハイブリッド型音声認識システムは、構成している2つの認識装置のうち高速な方の装置（例えば、図1のサーバ側での認識装置120）が実現する場合よりも高い音声認識確度を実現することができる。これは、従来のサーバ側での音声認識技法を上回る明確な利点であり、この従来の技法は、サーバ側での認識装置がシステムの使用する唯一の認識装置なので、この認識装置の確度を有する結果を実現するのみである。

【 0 0 3 6 】

同様に、本発明の各実施形態に従って実施されるハイブリッド型音声認識システムは、構成している2つの認識装置のうち低速な方の装置（例えば、図1のクライアント側での認識装置112）が実現する場合よりも高速な平均応答時間を実現することができる。これは、従来のサーバ側での音声認識技法を上回る明確な利点であり、この従来の技法は、サーバ側での認識装置がシステムの使用する唯一の認識装置なので、この認識装置の応答時間を有する結果を実現するのみである。

10

【 0 0 3 7 】

さらに、本発明の各実施形態は、ハイブリッド・システムを形成するのに使用してもよい認識装置のタイプまたは組合せに制約条件を課すものではない。クライアント側での認識装置112およびサーバ側での認識装置120のそれぞれは、どんな種類の認識装置でもよい。それら装置のそれぞれは、もう一方の装置の特性を知ることなく選択してもよい。場合によっては様々なタイプの複数のクライアント側での認識装置を、単一のサーバ側での認識装置とともに使用して、複数のハイブリッド型認識システムを効果的に形成してもよい。クライアント側での認識装置112またはサーバ側での認識装置120のいずれも、ハイブリッド・システムを破綻させることなく、修正してもよく、また置き換えてもよい。結果として、本明細書において開示する各技法により、それらの技法を、多種多様なクライアント側での認識装置およびサーバ側での認識装置とともに使用するのに適したものに柔軟性が広がる。

20

【 0 0 3 8 】

さらに、音声認識エンジンを利用する既存のアプリケーションにどんな修正も要求することなく、本発明で開示する各技法を実施してよい。前述の通り、例えば、委譲エンジン110は、従来の音声認識エンジンと同じインターフェースをアプリケーション108に提供してもよい。結果として、アプリケーション108は、委譲エンジン110が従来の音声認識エンジンである場合と同様に、委譲エンジン110に入力を供給し、そこから出力を受け取ってもよい。したがって、委譲エンジン110は、アプリケーション108にどんな修正も要求することなく、従来の音声認識エンジンの代わりにクライアント装置106に入れてもよい。

30

【 0 0 3 9 】

これまで、具体的な実施形態に関して本発明を説明してきたが、前述の各実施形態は例示的なものとしてのみ示しており、本発明の範囲を限定または定義するものではないことを理解されたい。他の様々な実施形態もまた、以下のものを含むがそれには限定されず、特許請求の範囲に記載の範囲内にある。例えば、本明細書に記載の各要素および各構成部品は、さらなる構成要素にさらに分割してもよく、また一緒に結合して、同じ機能を実行するための構成部品をより少なくなるよう形成してもよい。

40

【 0 0 4 0 】

前述の各技法は、例えば、ハードウェア、コンピュータ読取り可能な媒体に具体的に格納されたソフトウェア、ファームウェア、またはそれらのどんな組合せで実施してもよい。前述の各技法は、プロセッサ、プロセッサが読取り可能な記憶媒体（例えば、揮発性メモリと不揮発性メモリ、および/または記憶素子）、少なくとも1つの入力装置、ならびに少なくとも1つの出力装置を備える、プログラム可能なコンピュータ上で実行される1つまたは複数のコンピュータ・プログラムで実施してもよい。入力装置を使用して入って

50

くる入力にプログラム・コードを適用して、説明した各機能を実行し、出力を生成してもよい。出力は、1つまたは複数の出力装置に供給してもよい。

【0041】

以下の特許請求の範囲に記載の範囲内にある各コンピュータ・プログラムは、アセンブリ言語、機械語、高水準手続き型プログラミング言語、またはオブジェクト指向プログラミング言語など、どんなプログラミング言語で実施してもよい。プログラミング言語は、例えば、コンパイラ型またはインタープリタ型のプログラミング言語でもよい。

【0042】

このようなコンピュータ・プログラムはそれぞれ、コンピュータ・プロセッサで実行するために、機械読取り可能な記憶装置内に具体的に組み入れられたコンピュータ・プログラム製品で実施してもよい。本発明の方法ステップは、コンピュータ読取り可能な媒体上に具体的に組み入れられたプログラムを実行して、入力に演算を施し出力を生成することにより本発明の各機能を実行するコンピュータ・プロセッサによって実行してもよい。適切なプロセッサには、一例として、汎用マイクロプロセッサと特殊目的のマイクロプロセッサが両方含まれる。一般に、プロセッサは、読取り専用メモリおよび/またはランダム・アクセス・メモリから、命令およびデータを受信する。コンピュータ・プログラム命令を具体的に組み入れるのに適した記憶装置には、例えば、EPROM、EEPROM、およびフラッシュ・メモリ・デバイスを含む半導体記憶装置、内部ハード・ディスクや取外し可能ディスクなどの磁気ディスク、光磁気ディスク、CD-ROMなど、あらゆる形態の不揮発性メモリが含まれる。前述のいかなるものも、特別に設計されたASIC（特定用途向け集積回路）またはFPGA（フィールド・プログラマブル・ゲート・アレイ）によって補ってもよく、またそれらに組み込んでもよい。一般に、コンピュータはまた、内部ディスク（図示せず）または取外し可能ディスクなどの記憶媒体から、プログラムおよびデータを受信することができる。これらの要素はまた、従来のデスクトップ・コンピュータまたはワークステーション・コンピュータ、ならびに本明細書において説明した各方法を実施するコンピュータ・プログラムを実行するのに適した他のコンピュータ内に見られ、これらのコンピュータは、任意のデジタル印刷エンジンもしくはマーキング・エンジン、表示モニタ、または、カラーもしくはグレー・スケールの画素を紙、フィルム、表示画面、もしくは他の出力媒体上に生成することのできる他のラスタ出力装置とともに使用してもよい。

10

20

30

【 図 1 】

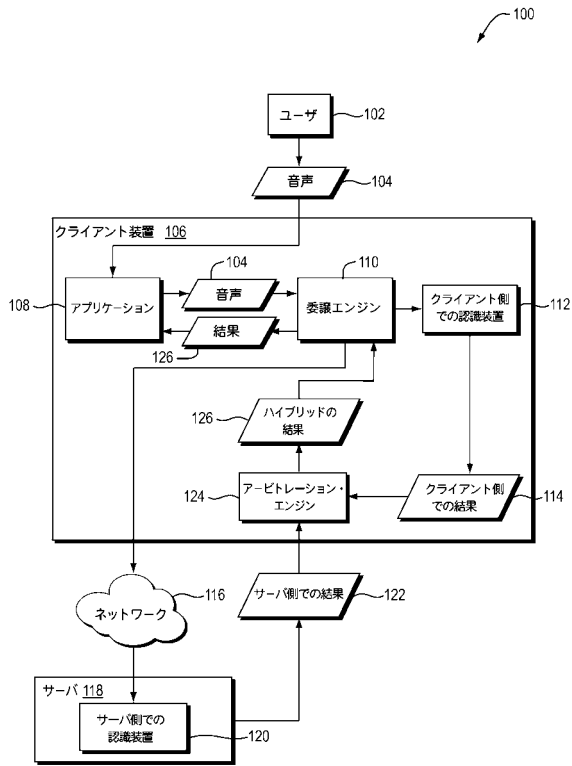


図 1

【 図 2 】

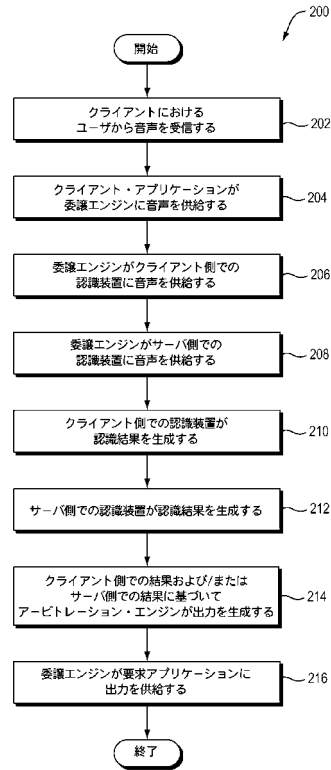


図 2

【 図 3 A 】

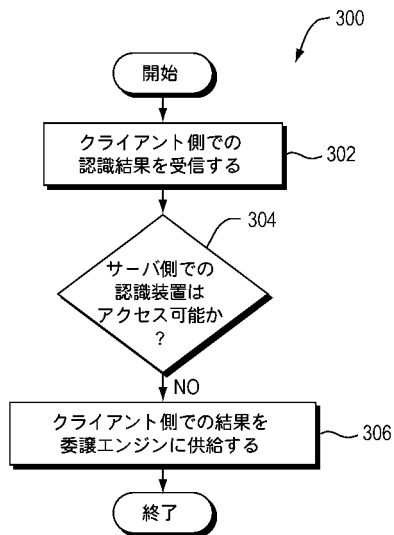


図 3A

【 図 3 B 】

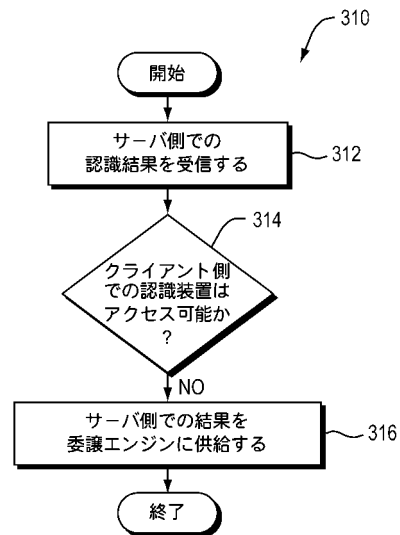
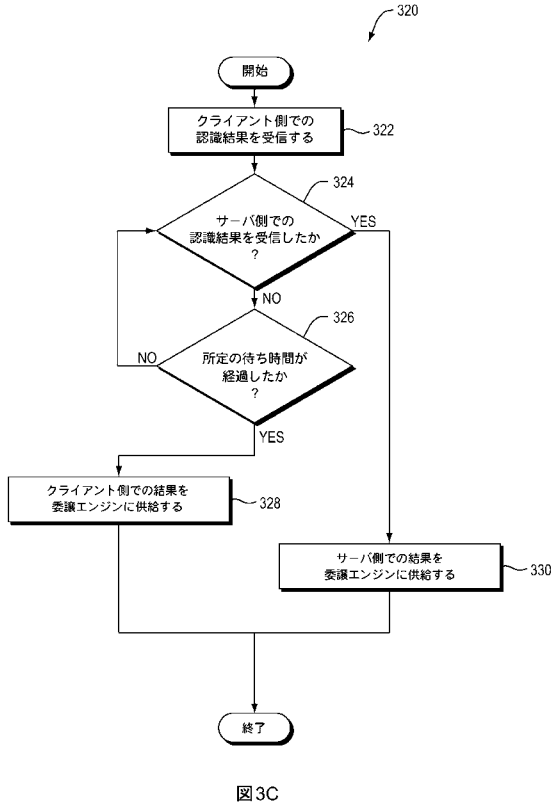
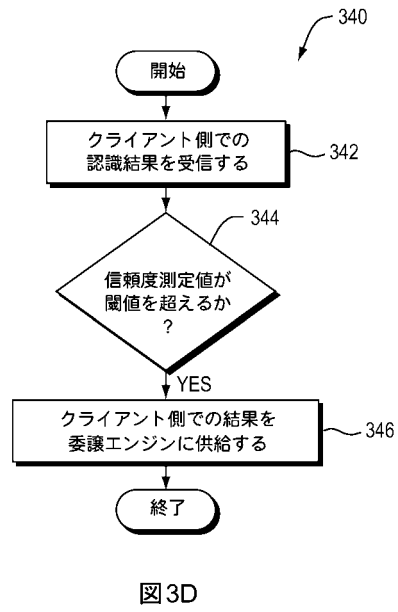


図 3B

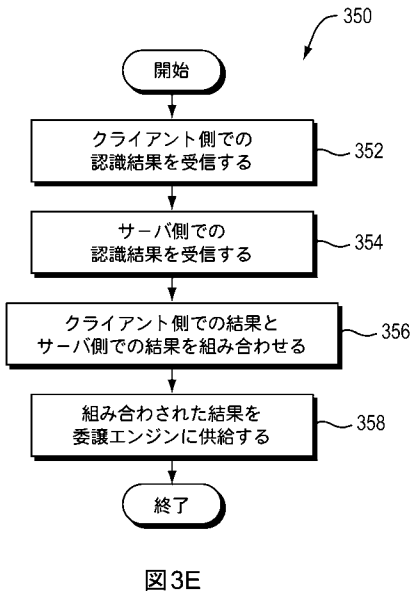
【 図 3 C 】



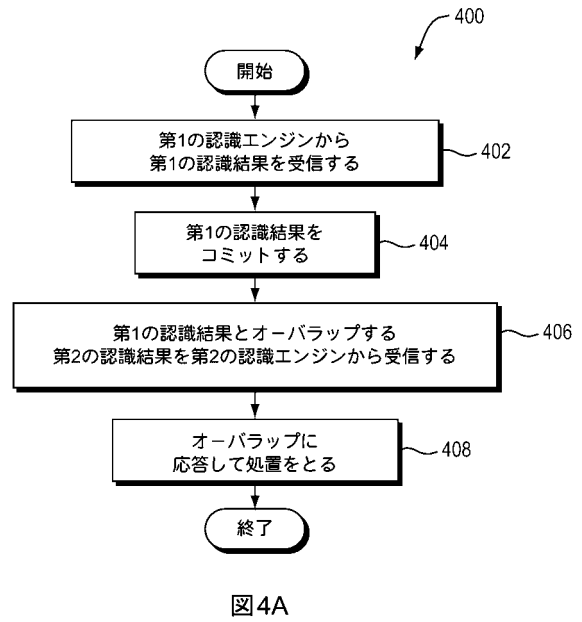
【 図 3 D 】



【 図 3 E 】



【 図 4 A 】



【 図 4 B 】

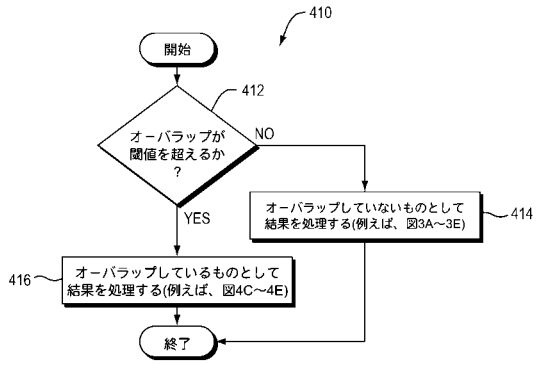


図4B

【 図 4 C 】

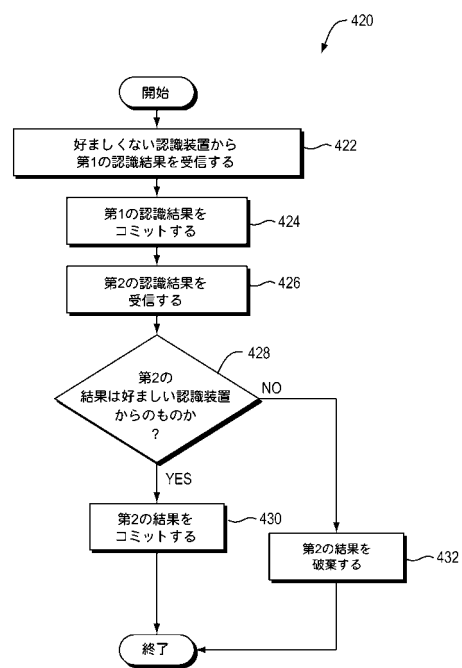


図4C

【 図 4 D 】

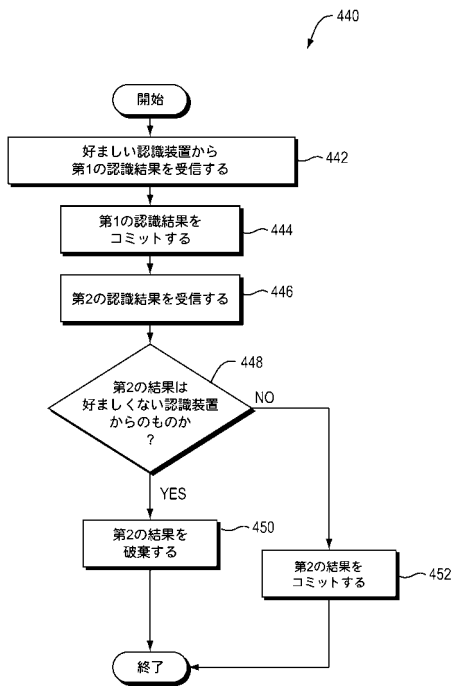


図4D

【 図 4 E 】

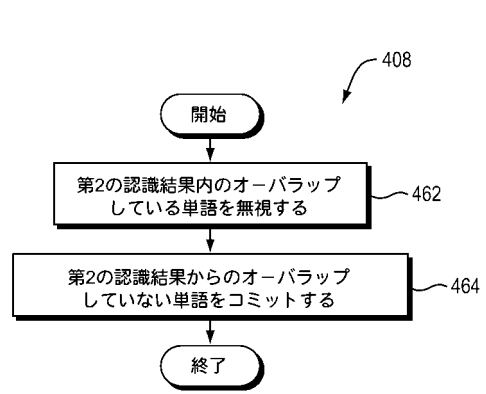


図4E

【 図 4 F 】

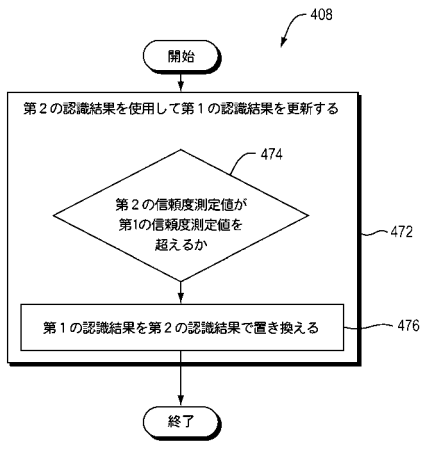




図4F

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2009/055479
A. CLASSIFICATION OF SUBJECT MATTER		
<i>G10L 19/00(2006.01)i</i>		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) IPC : G10L 19/00; G06F 17/00; G06F 3/16; G10L 15/00; G10L 15/06; G10L 15/22; G10L 15/26; G10L 15/28		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models Japanese utility models and applications for utility models		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS(KIPO internal) & Keywords: audio, speech, server, network, recongni*, client, distribute, hybrid		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y A	US 6898567 B2 (SENAKA BALASURIYA) 24 May 2005 See column 4, line 15-column 6, line 3; column 8, lines 28-38; figures 3, 5.	13,21 1-6,10-12 7-9,14-20
Y	KR 10-2006-0089288 A (HYUNDAI AUTONET CO., LTD.) 09 August 2006 See claim 3, figures 1, 3, and 6.	1-6,10-12
A	JP 2005-249829 A (ADVANCED MEDIA INC) 15 September 2005 See abstract, claim 1, figures 1-3.	1-21
A	JP 2007-083901 A (NEC CORP) 08 February 2007 See abstract, claim 1-2, 7-8.	1-21
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 08 APRIL 2010 (08.04.2010)		Date of mailing of the international search report 09 APRIL 2010 (09.04.2010)
Name and mailing address of the ISA/KR  Korean Intellectual Property Office Government Complex-Daejeon, 139 Sconsa-ro, Seo-gu, Daejeon 302-701, Republic of Korea Facsimile No. 82-42-472-7140		Authorized officer JANG Ho Keun Telephone No. 82-42-481-8187 

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/US2009/055479

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6898567 B2	24.05.2005	AU 2002-367354 A1 CN 1633679 A CN 1633679 C0 FI 20040872 A FI 20040872 D0 JP 2005-524859 A KR 10-0632912 B1 US 2003-0139924 A1 WO 03-058604 A1	24.07.2003 29.06.2005 06.06.2007 25.08.2004 23.06.2004 18.08.2005 13.10.2006 24.07.2003 17.07.2003
KR 10-2006-0089288 A	09.08.2006	None	
JP 2005-249829 A	15.09.2005	None	
JP 2007-033901 A	08.02.2007	CN 1905007 A CN 1905007 C0 US 2007-0027693 A1	31.01.2007 31.01.2007 01.02.2007

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW