(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3561076号
(P3561076)
(45)発行日　平成16年9月2日(2004.9.2)　(24)登録日　平成16年6月4日(2004.6.4)

(51)Int.Cl.[7]	F I		
G10L 15/00	G10L	3/00	551A
G10L 15/18	H04M	3/42	P
H04M 3/42	G10L	3/00	537G

請求項の数 9　(全 11 頁)

(21)出願番号　特願平8-78564
(22)出願日　平成8年4月1日(1996.4.1)
(65)公開番号　特開平8-320696
(43)公開日　平成8年12月3日(1996.12.3)
審査請求日　平成12年1月11日(2000.1.11)
審判番号　不服2003-1713(P2003-1713/J1)
審判請求日　平成15年2月4日(2003.2.4)
(31)優先権主張番号　08/413330
(32)優先日　平成7年3月30日(1995.3.30)
(33)優先権主張国　米国(US)

(73)特許権者　390035493
エイ・ティ・アンド・ティ・コーポレーション
ＡＴ＆Ｔ　ＣＯＲＰ.
アメリカ合衆国　１００１３－２４１２
ニューヨーク　ニューヨーク　アヴェニュー　オブ　ジ　アメリカズ　３２
(74)代理人　100064447
弁理士　岡部　正夫
(74)代理人　100085176
弁理士　加藤　伸晃
(74)代理人　100106703
弁理士　産形　和央
(74)代理人　100096943
弁理士　臼井　伸一

最終頁に続く

(54)【発明の名称】任意に話された単語の自動通話認識方法

(57)【特許請求の範囲】

【請求項１】

複数の音素モデルからなる複数の単語モデルを記憶している第１のデータベースを備える自動音声認識　（ＡＳＲ）　システムにおいて、単語認識は、少なくとも部分的に該単語モデルに基づいて行われ、補助的な第２のデータベースに記憶された情報を利用して該ＡＳＲシステムの能力を高める方法であって、

ユーザから第１の音声入力を受信するステップと、

該ＡＳＲシステムを使用して該ユーザからの該第１の音声入力を認識するステップと、

該認識された第１の音声入力を使用して該補助的な第２のデータベースに記憶されている文書を識別し、そして検索するステップとを含み、該検索された文書は、該第１の音声入力の表現にはなっておらず、該方法は、さらに、

該補助的な第２のデータベースから検索された該文書から、該ユーザからのいかなる音声入力をも使用することなく該音素モデルを使用して、話者に依存しないテンプレートを生成するステップと、

該テンプレートを使用して第２の音声入力を認識するステップとを含む方法。

【請求項２】

請求項１に記載の方法において、ユーザから第１の音声入力を受信するステップは、電話番号を表す音声を受信するステップを含む方法。

【請求項３】

請求項２に記載の方法において、該テンプレートを使用して第２の音声入力を認識するス

テップは、音声認識アルゴリズムを使用するステップを含む方法。
【請求項４】
請求項１に記載の方法において、文書のテンプレートを生成するステップは、文書／音声システムを使用して音素表現を生成するステップを含む方法。
【請求項５】
電話ネットワークにおいて、自動音声認識　（ＡＳＲ）　システムと関連する補助的なデータベースを使用して該ＡＳＲシステムの能力を高める方法であって、
発呼者からの入力呼を受信するステップと、
該発呼者に発声を催促するステップと、
該発呼者が発した発声を認識するステップと、
該認識された発声をインデックスとして使用して、補助的なデータベースに記憶されている文書を検索するステップとを含み、該検索された文書は、該発声の表現にはなっておらず、該方法は、さらに、
該補助的なデータベースから検索された該文書から、該発呼者からのいかなる音声入力をも使用することなく複数の音素モデルを使用して、話者に依存しないテンプレートを生成するステップと、
該話者に依存しないテンプレートを使用して、該補助的なデータベースから検索された該文書に対応する他の発声を認識するステップとを含む方法。
【請求項６】
請求項５に記載の方法において、テンプレートを生成するステップは、文書／音声システムを使用して、該文書の音素表現を生成するステップを含む方法。
【請求項７】
請求項５に記載の方法において、他の発声を認識するステップは、音声認識アルゴリズムと複数の音素モデルとを使用するステップを含む方法。
【請求項８】
電話ネットワークと接続して使用される自動音声認識　（ＡＳＲ）　システムであって、
１つの交換サービスポイントによって取り扱われる発呼局と、
発呼者から音声入力を受信する手段と、
該電話ネットワークを取り扱うホストコンピュータのランダムアクセスメモリに該発呼局から受信した音声入力を送信する電話線インターフェイス装置と、
該ホストコンピュータ内に配置されると共に、該音声入力を使用して補助的なデータベースから情報を検索する中央処理装置とを含み、該補助的なデータベースから検索された情報は、該音声入力の表現にはなっておらず、該システムは、さらに、
該補助的なデータベースから検索された該情報の音素表現を生成する文書／音声手段と、
該補助的なデータベースから検索された該情報の音素表現から、該発呼者からのいかなる音声入力をも使用することなく複数の音素モデルを使用して話者に依存しないテンプレートを生成し、該生成された話者に依存しないテンプレートを使用して該検索された情報と関連する発声を認識する音声認識手段とを含むシステム。
【請求項９】
請求項８に記載のＡＳＲシステムにおいて、該音素表現を話者に依存しないテンプレートとして使用するシステム。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】
本発明は自動通話認識に関する。特に、本発明は、電話網に関してデータベースから読み出された補助情報を用いて、ユーザーが話した単語を認識する際に自動通話認識装置（自動音声認識装置）（ＡＳＲ）を支援する方法に関する。
【０００２】
【従来の技術及び発明が解決しようとする課題】
多数のアプリケーションについて、繰り返される仕事を自動化してより早く効率的に実行

できるようにするためにコンピュータを用いるのが効果的である。音声技術の一形式である通話認識（音声認識）により、人々は話し単語を用いてコンピュータと対話することができる。しかしながら、さまざま人の通話には固有の違いがあるため、通話認識はなかなか難しいものである。

【０００３】

音声認識の１つの応用は電話網である。自動通話認識（ＡＳＲ）装置を用いて、人々は電話を介してやり取りし、オペレータを介在することなく簡単な仕事を実行することができる。例えば、通話認識は、電話のユーザーが電話番号を覚えたり見たり尋ねたりする必要がないようにダイヤリングに用いることができる。ユーザーインターフェースの物理的操作の代わりに通話を使用できるということは、遠隔通信が進歩しても引き続きＡＳＲ技術の需要を高く保っている。一般に、遠隔通信に用いられるＡＳＲ装置には、話者依存式及び話者非依存式という２つのタイプがある。

【０００４】

話者依存式自動通話認識装置の一般的な実行は、特定の話者のデータを使用して訓練（学習）されてこの話者の通話パターンに応答するコンピュータを使用する。この訓練工程は、アナログ通話入力を発生するための音声（すなわち単語）の発声と、この通話入力の信号データへの変換と、この音声を表わす１つのテンプレートの発生と、作動させるためのコンピュータ命令等の適切な特定応答データの記憶からなる。

【０００５】

システムをリアルタイムで動作させた場合には、学習用の話者が発した単語データをデジタル化し、個々の話者に依存するプレートのセットとその単語データをＡＳＲシステム内で比較し、それによって、話者の単語データとテンプレートの間でマッチングをとり、コンピュータから特定の応答を得ることが可能である。話者依存式ＡＳＲ装置は、主に、学習適切に行なえる場合、例えば同一個人が多くの機会に装置をアクセスする場合（多くの学習データを得られる場合）に用いられる

【０００６】

個々の話者による学習を適切に行えないアプリケーションについては、話者非依存式ＡＳＲを用いなければならない。話者非依存式ＡＳＲ装置の一般的な実行は、多数の異なる人達が話した単語を表わす複合テンプレートまたはテンプレート群を記憶するためにコンピュータを用いる。テンプレートは、広範囲の発音と通話特性の変化とを表わす多数のデータサンプル（すなわち、複数の話者が話した単語）から引き出される。話者非依存式通話認識装置は、話者を特定する訓練をすることなく非常にさまざまな人達と対話することができる。

【０００７】

話された番号を認識するために話者非依存式ＡＳＲを用いる電話アプリケーションは技術上知られている。これらのアプリケーションは、話者の語彙がわずかなメニュー命令及び／または番号（例えば０～９）に限られる場合に特に有効である。しかしながら、電話網を介して話された字句を認識するのは非常に難しい。実際には、個々の話者の非常にさまざまな通話パターンがあることや、いろいろなタイプの雑音と帯域幅制限に起因して、電話環境はあらゆるＡＳＲアプリケーション（話者依存式及び話者非依存式）をエラーしやすくする。

【０００８】

それにもかかわらず、ＡＳＲの探究後の商業的アプリケーションは、電話網を介して行なわれる商業的取り引き、例えばクレジットカード取り引きと関連する仕事を自動化する。例えば、顧客が電話を介して商品やサービスを買おうとする場合、関連情報を集めて、最小限度のオペレータ介在を伴う早くかつ効率的な処理を確立するために、ＡＳＲを用いることができる。

【０００９】

電話による商品やサービスの購入は、ＡＳＲ装置に、任意に話された単語（例えば固有名詞）の認識ばかりでなくいたるところでアクセスできることも要求する。例えば、従来の

ＡＳＲ装置を用いて全国を横断する潜在的な顧客との大規模なビジネスの要求に役立つように、上記に説明した手法を用いて全国のあらゆる人の名前に対応する話者非依存テンプレートを作り出して記憶する必要がある。しかしながら、現在の技術によるＡＳＲ装置は、記憶したテンプレートに対応する数千のあり得る名前のうちの１つと話された名前を整合させることができない。
したがって、ＡＳＲ装置において、増加した能力で話された単語を自動的に認識させることができる改善が技術上要求されている。

【００１０】

【課題を解決するための手段】

この要求に取り組み、当該技術分野において、固有名詞等の任意に話された単語を早く正確に認識することができるようにＡＳＲ装置の能力を増大させるために補助データを用いる方法及び装置によって技術的進歩が達成される。
ＡＳＲ装置の電話アプリケーションでは、発呼者から電話番号を求めるのが普通である。発呼者の電話番号は、発呼者ＩＤやタッチトーン入力等の通話または他の既知の方法で得ることができる。したがって、本発明の方法及び装置の模範的な一実施例によれば、発呼者の電話番号は、データベースから文書（テキスト）（例えば、発呼者名）を読み出すための索引（インデックス）として役立つ。次いで、この文書は、ＡＳＲ装置に利用可能な選択を制限または規定するために用いられる。

【００１１】

より詳細には、発呼者の電話番号（第１の音声入力）は、この電話番号と関連する文書を読み出すために補助データベースにアクセスするのに用いられる。上述の模範的な実施例では、発呼者名を含む文書は補助データベースから読み出される。発呼者名の文書（テキスト）は、発呼者の固有名詞のデジタル化英数字表現からなる。文書／通話装置（文章／音声システム）は、技術上知られているように、発呼者名のテキストを音素表現に変換するのに用いられる。名前の音素表現は話者非依存テンプレートとしてＡＳＲ装置に記憶され、この表現に従う通話はＡＳＲ装置で認識することができる。補助的な文書データの音素表現への読み出し及び変換により、ＡＳＲ装置は、特定の話者による訓練なしで読み出されたデータの音素表現に対応する話された単語にただちに応答することができる。

【００１２】

上記の模範的な実施例の実時間アプリケーションの間、電話は、本発明によるＡＳＲ装置を保有するサービスプロバイダーにより受信される。商品やサービスをクレジットカードで買おうとする発呼者は、電話番号及び名前を促される。発呼者より与えられる電話番号に基づき、補助データベース（例えば電子的電話番号簿）がアクセスされ、この電話番号と関連する名前の文書が読み出される。データベースから読み出された名前の文書の音素表現は、文書／通話装置で作り出される。次いで、この音素表現はＡＳＲ装置で話者非依存テンプレートとして記憶されて用いられ、発呼者が話した通りに名前が認識される。

【００１３】

発呼者が話した名前が、名前の普通でない発音や、劣悪な電話伝送品質や、認識するのが難しい声を持つ発呼者などのためにＡＳＲ装置で認識されない場合は、発呼者は名前のつづりを１文字ずつ与えるよう促される。データベースから読み出された通りに、書かれた名前のつづりの音素表現を作り出すために、文書／通話装置を用いることができる。また、データベースから読み出された名前のつづりは話者非依存テンプレートとしてＡＳＲ装置に記憶され、それにより、ＡＳＲ装置は、発呼者の名前のつづりを（発呼者が話した通りに）認識することを試みることができる。発呼者の名前の話された発音と（確立された通話認識アルゴリズムに基づく）音素表現が整合しない場合は、発呼は人間の案内人に送られる。

【００１４】

【発明の実施の形態】

図１は、本発明の模範的な実施例による２つの装置１００Ａ及び１００Ｂを示す。音声電話装置１００Ａには、電話端末１０２と、通信線１０３Ａ及び１０３Ｂと、公衆交換電話

網１０７に配置される交換サービスポイント（ＳＳＰ）１０５が含まれる。明快にするため、１つの交換サービスポイントが示されているが、運転公衆交換電話網は相互接続されたＳＳＰ網からなる。電話線インターフェース装置１０８は、電源変動等の網機能不良からＡＳＲ装置１１０の機器を保護すると共に、原通話が装置に届く前にまだデジタル化されていなければ電話端末１０２から到来する通話をデジタル化する。

【００１５】

ＳＳＰ１０５は分散制御型構内デジタル交換機であり、例えば、ＡＴ＆Ｔテクニカルジャーナル，第６４巻，第６号，１９８５年７月～８月，第１３０３～１５６４ページや、１９８１年１１月，ベルラボラトリーレコード，第２５８ページや、１９８１年１２月，ベルラボラトリーレコード，第２９０ページに開示され、ＡＴ＆Ｔで製造された５ＥＳＳ（登録商標）交換機である。かけがえとして、ＳＳＰ１０５は、１９８６年5月２７日にエム・ダブリュ・ベックナー（Ｍ．Ｗ．Ｂｅｃｋｎｅｒ） に発行された米国特許第４，５９２，０４８ 号に開示されているＩＳＤＮ交換装置のような分散制御型アナログまたはデジタル交換機でも良い。図１では、ＳＳＰ１０５は、顧客識別線１０３を介して電話端末１０２に接続されると共に、以下に説明されるように線１２１Ｂを介してホストコンピュータ１２４と交信する。

【００１６】

また、図１には、マイクロフォン１０４及びマイクロフォンインターフェース装置１０９を含むユーザー対話装置１００Ｂが示されている。マイクロフォン１０４は、技術上知られているように、ＡＳＲ装置と顧客間のリンクとしてサービスプロバイダーにより保有されるキオスク又は自動化金銭出納機に配置することができる。

【００１７】

到来する通話は、マイクロフォン１０４で電気信号に変換され、通信リンク１０６Ａを介してマイクロフォンインターフェース装置１０９に送られる。マイクロフォンインターフェース装置１０９は、通信リンク１０６Ｂを介するＡＳＲ装置１１０への送信前に到来通話信号をデジタルデータに変換する。

【００１８】

ＡＳＲ装置１１０（図２において以下に詳細に説明される）はデータバス１２５を介してホストコンピュータ１２４と交信する。ホストコンピュータ１２４には、コンピュータの動作全体を制御する中央処理装置（ＣＰＵ）１２６と、データを一時的に記憶するランダムアクセスメモリ（ＲＡＭ）１２８と、データを永久記憶するリードオンリーメモリ（ＲＯＭ）１３０と、ホストコンピュータ１２４と関連する制御プログラムを記憶する不揮発性データベース１３４が含まれる。ＣＰＵ１２６はデータバス１３２を介してＲＡＭ１２８及びＲＯＭ１３０と交信する。同様に、ＣＰＵ１２６はデータバス１３３を介して不揮発性データベース１３４と交信する。入力／出力（Ｉ／Ｏ）インターフェース１３６はデータバス１３５を介してホストコンピュータ１２４に接続され、データリンク１３９を介してＩ／Ｏインターフェース１３６と交信する構内情報通信網（ＬＡＮ）１３８と、データリンク１４１を介してＩ／Ｏインターフェース１３６と交信する補助データベース１４０と、通信線１２１Ａ、ＳＳＰ１０５及びデータリンク１２１Ｂを介してホストコンピュータ１２４にデジタルデータを送信するデータサービス網１４２とからのデータフローを促進する。

【００１９】

図２は、図１に示されるようなＡＳＲ装置１１０の模範的な実施例の簡略ブロック図を示す。ＡＳＲ装置１１０は話者非依存または話者依存のどちらかの通話認識を行なうことができ、装置の動作全体を制御するＣＰＵ２０２を含む。ＣＰＵ２０２は一般的に参照数字２０３で表わされた複数のデータバスを有する。また、図には、技術上知られているように、ランダムアクセスメモリ（ＲＡＭ）２０４と、リードオンリーメモリ（ＲＯＭ）２０６と、発呼者にあいさつと助言を発する通話発生装置２１８と、書き込み文書を音素表現に変換する文書／通話（ＴＴＳ）装置２１９（ＣＰＵ２０２及びＲＡＭ２０４と交信する）も示されている。

【００２０】
ＲＡＭ２０４はバス２０３でＣＰＵ２０２に接続され、発呼者が電話端末１０２またはマイクロフォン端末１０４で話した単語、話者依存テンプレート２１４、話者非依存テンプレート２１６などの通話データを一時的に記憶する。ＲＯＭ２０６は同様にデータバス２０３でＣＰＵ２０２に接続され、通話認識アルゴリズム２０８及び音素モデル２１０を含む通話認識及び検定データを永久記憶する。多くの他の有効な通話認識方法が技術上知られているが、この模範的実施例では、音素に基づいた通話認識アルゴリズム２０８が用いられる。
【００２１】
音素は通話の最小単位群の１つを指す技術用語であり、より大きな通話セグメント、例えば形態素を構成するために他の前記単位と結合することができる。例えば、話された単語“オペレータ（ｏｐｅｒａｔｏｒ）”の音声を表わすセグメントは“ａａ”，“ｐ”，“ａｘｒ”，“ｅｙ”，“ｄｘ”及び“ａｘｒ”等の音素の組み合わせで表わすことができる。音素モデル２１０は、従来のオフライン処理で話者のサンプルの発音から引き出された通話認識分類データを用いてコンパイルされる。この処理の間、言語の全ての音素を含むように選択された単語が多数の訓練話者（例えば１０００人）により話される。発音は、発音の内容の書き込み文書を発生させる訓練された個人により処理される。
【００２２】
次いで、この単語の書き込み文書は、ＴＴＳ装置２１９等の文書／通話装置で受信され、技術上知られているように、文書／通話変換のルールを用いて書き込み文書の音素表現を作り出すことができる。次いで、書き込み文書の音素表現は、発音を音素モデル２１０と比較する通話認識アルゴリズム２０８の動作から引き出された音素と比較される。音素モデル２１０は、発音の文書／通話変換から引き出された音素と通話認識アルゴリズム２０８で認識された音素間で十分な整合が得られるまで、技術上知られているような調整方法を用いてこの“モデル訓練”処理の間調整される。
【００２３】
音素モデル２１０は、認識処理の間通話認識アルゴリズム２０８と共に用いられる。より詳細には、通話認識アルゴリズム２０８は、話された単語を確定した音素モデルと整合させる。通話認識アルゴリズムが整合ありと決定すれば（すなわち、話された発音が予め定められたパラメータにしたがって音素モデルを統計的に整合すれば）、音素のリストが生成される。
【００２４】
音素モデル２１０は、多数の話者にわたる話された単語の特性の分布を表わすので、訓練話者により代表される同一話者集団（すなわち、生粋のアメリカ人、スペイン語を話す人々等）に役立つＡＳＲ装置へのいたるところでのアクセスに用いることができる。
【００２５】
話者非依存テンプレート２１６は、予想される発音または成句を表わす音素のリストである。話者非依存テンプレート２１６はＴＴＳ装置２１９を介して書き込み文書を処理することにより作り出され、書き込み単語または成句の予想される発音を例示する音素のリストが生成される。一般に、多数のテンプレートがＲＡＭメモリ２０４に記憶され、通話認識アルゴリズム２０８に利用可能になる。アルゴリズム２０８の仕事は、話された発音における音素とどのテンプレートが最も良く整合するかを選ぶことである。
【００２６】
話者依存テンプレート２１４は、単語または成句の発音を与える話者を持ち、通話認識アルゴリズム２０８及び音素モデル２１０を用いて発音を処理してアルゴリズムで認識された音素からなる音素のリストを作ることにより、生成される。この音素のリストは、その特定の発音のための話者依存テンプレート２１４になる。
【００２７】
実時間通話認識動作の間、発音は、音素のリストが生成されるように音素モデル２１０を用いて通話認識アルゴリズム２０８で処理される。この音素のリストは、技術上知られて

いる方法を用いて、話者非依存テンプレート２１６及び話者依存テンプレート２１４で処理されたリストと整合される。通話認識アルゴリズム２０８は整合結果を報告する。

【００２８】

図３は、ＡＳＲ装置が本発明の方法の模範的な実施例の話者非依存モードで動作している場合に、ＡＳＲ装置で行なわれる作用を説明するフロー図である。

本発明の商業的アプリケーションの一例として、顧客はホームテレホン（電話端末１０２）から電話をかけ、ＡＳＲ装置１１０及びホストコンピュータ１２４を用いるサービスプロバイダーより提供されるサービスのクレジットカードによる購入を希望していると仮定する。この例では、顧客はこのサービスを以前に購入していなかったので、ＡＳＲ装置１１０はこの顧客の特定の通話パターンを認識するように訓練されていない（すなわち、この顧客用に確定された話者依存テンプレート２１４が存在しない。）。しかしながら、クレジットカード取り引きが認められるようにするために、ＡＳＲ装置１１０はこの顧客の名前を受信して認識しなければならない。

【００２９】

この例は、ステップ３００において、ＡＳＲ装置１１０が公衆交換電話網１０７の通信線１０３Ａ，１０３Ｂ及びＳＳＰ１０５を介して送られた顧客が発した到来電話を受信する。

代りに、顧客は、マイクロフォン端末１０４を含むユーザー対話装置を内蔵するキオスクからサービスの注文を出すことができる。もしそうならば、到来する“電話”は、顧客入力（例えば会話）がマイクロフォン端末１０４で検出されて通信リンク１０６Ｂを介して装置に送られる時に、ＡＳＲ装置１１０で受信される。

【００３０】

電話装置及びユーザー対話装置の両方において、到来する電話はインターフェース装置（すなわち、それぞれ電話線インターフェース装置１０８及びマイクロフォンインターフェース装置１０９）で処理され、ＡＳＲ装置１１０において受信された全入力が通常のデジタルフォーマットになることが保証される。

ステップ３０２に示されるように、ＡＳＲ装置１１０の通話発生装置２１８は、所定のインデックス（例えば、ホームテレホンナンバー）、このインデックスと関連する名前又はこの名前のつづり等の入力を顧客に促す。代りに、本装置は、以下に説明されるように、処理に必要になるまで名前のつづりを顧客に促すのを引き延ばすことができる。

【００３１】

次いで、処理は決定ステップ３０４に進み、要求された入力が受信されたか否かを決定する。ステップ３０４の結果が“ＮＯ”決定ならば、処理はステップ３０６に進み、電話は人間の案内人に送られ、次いで、処理は終了ステップ３０８で終わる。

ステップ３０４の結果が“ＹＥＳ”決定ならば、処理はステップ３１０に進み、インデックス、名前及び名前のつづりの顧客の発音はＡＳＲ装置１１０のＲＡＭ２０４に記憶される。上記の例では、顧客は話すことによりインデックスを与えた。インデックスが顧客のホームテレホンナンバーの場合は、発呼者ＩＤまたはタッチトーン入力等の他の既知の方法で取り出すことができる。

【００３２】

次いで、処理はステップ３１２に進み、通話認識アルゴリズム２０８及び音素モデル（数字に関するもの）２１０を用いて発呼者のインデックスの認識を試みる。このインデックスは、以下に説明されるように、補助データベースから情報を読み出すために用いられる。そのインデックスが付いているいくつかの数字について不確かさがある場合は、本装置は多数の可能性を認識するようにプログラムすることができる。決定ステップ３１６では、顧客のインデックスが前のステップで認識されていたか否かが決定される。この決定ステップの結果が“ＮＯ”決定ならば、処理はステップ３０６に進み、電話は人間の案内人に送られ、次いで、処理は終了ステップ３０８で終わる。

【００３３】

決定ステップ３１６の結果が“ＹＥＳ”決定ならば、処理はステップ３１８に進み、ＡＳ

Ｒ装置１１０のＣＰＵ２０２はホストコンピュータ１２４に補助データを要求する。上記の例では、望ましい補助データは、ＡＳＣＩＩ文書フォーマットにおけるような顧客名のデジタル的に記憶された表現（例えば、顧客より受信されたホームテレホンナンバーと関連する名前）である。

【００３４】

補助データの要求がホストコンピュータ１２４のＣＰＵ１２６で受信された時、ＣＰＵ１２６は、インデックス（例えば、顧客より与えられたテレホンナンバー）を用いて、顧客の名前のデジタル表現（例えば、ＡＳＣＩＩ文書フォーマット）を読み出すために、何れの補助データベースにアクセスしなければならないかを決定する。ＣＰＵ１２６は、不揮発性データベース１３４から受信した何れのデータベースにアクセスするべきかの命令に基づいて決定を行なう。

【００３５】

例えば、サービスプロバイダーは大規模エンティティならばＣＤ－ＲＯＭデータベース等の補助データベース１４０を保有することができ、これはデータリンク１４１及びＩ／Ｏインターフェース１３６を介してホストコンピュータ１２４と通信する。データベース１４０は、顧客の住所及び名前、クレジットカード口座番号、及びテレホンナンバーで検索される購入歴等の広範囲にわたる顧客情報を含むことができる。しかしながら、サービスプロバイダーが早く変化する顧客ベースを有する小規模エンティティならばホストコンピュータ１２４内に（例えば、ＲＡＭ１２８内に）制限された補助データベースを記憶することができる。

【００３６】

代りに、サービスプロバイダーのなかには、補助データをデータリンク１３９及びＩ／Ｏインターフェース１３６を介してホストコンピュータ１２４にダウンロードすることができるコンピュータ網（例えばＬＡＮ１３８）を保有することができるものもある。

【００３７】

上記の例において、サービスプロバイダーは、公衆交換電話網１０７を保有する通信網より提供されるデータサービスに加入していると仮定する。データサービス１４２は、合衆国における全住宅電話のテレホンナンバーと対応する顧客名を含む電子的電話帳を記憶しているデータベースを含む。上記の例では、データサービス１４２は、情報（例えば顧客名の文書）のデジタルデータパケットを通信線１２１Ａを介してＳＳＰ１０５に送信する。ＳＳＰ１０５は、このデジタル情報を通信線１２１Ｂ乃至Ｉ／Ｏインターフェース１３６を介してホストコンピュータ１２４に送り、その結果、顧客名の文書は、ホストコンピュータ１２４で受信されてＲＡＭ１２８に記憶される。多数の名前が（ホームテレホンナンバーまたはテレホンナンバーと関連する多数の名前の多数の可能性により）読み出される場合、全ての可能性のある名前がホストコンピュータに与えられる。

【００３８】

次いで、処理はステップ３２０に進み、ＣＰＵ２０２から受信された要求に応じて、補助データベースから読み出された名前の文書がＲＡＭ１２８から読み出されてＴＴＳ装置（ｔｅｘｔ－ｔｏ－ｓｐｅｅｃｈｓｙｓｔｅｍ：文章／通話装置）２１９で処理され、その結果文書の音素表現が生成されて、話者非依存テンプレートとしてＲＡＭ２０４に記憶される。ステップ３２２に示されるように、発呼者が話した（及びＲＡＭ２０４に記憶された）通りの発呼者の名前の認識が、通話アルゴリズム２０８と音素モデル２１０とステップ３２０で作り出された話者非依存テンプレートとを用いて試みられる。

【００３９】

決定ステップ３２４では、ステップ３２２で認識がされたか否かに関する決定が行なわれる。ステップ３２４の結果が“ＹＥＳ”決定ならば、処理はステップ３２６に進み、取り引きが承認され、次いで、処理はステップ３２８で終る。ステップ３２４の結果が“ＮＯ”ならば、処理はステップ３３０に進み、（データベースから読み出された通り）読み出された顧客名のつづりの音素表現がＴＴＳ装置２１９で作り出され、第２の話者非依存テンプレートとして記憶される。ステップ３３２では、顧客が話した通りの顧客名のつづり

の認識が、通話認識アルゴリズム２０８と音素モデル２１０とステップ３３０で作り出された話者非依存テンプレートを用いて試みられる。次いで、処理はステップ３３４に進み、発呼者名のつづりが認識されたか否かが決定される。ステップ３３４で“ＮＯ”決定ならば、処理はステップ３０６に進み、発呼は人間の案内人に送られ、次いで処理はステップ３０８で終る。ステップ３３４で“ＹＥＳ”決定ならば、処理はステップ３２６に進み、取り引きが承認され、次いで処理はステップ３２８で終わる。

【００４０】

上記の例は、電話機またはユーザー対話装置を用いた顧客と、ＡＳＲ装置１１０と、ホストコンピュータ１２４と、補助データベースとの間の実時間対話を示す。しかしながら、ＡＳＲ装置１１０がＬＡＮ１３８でアクセスされる実施例や、顧客名がある期間にわたってデータベースに登録、記憶され、かつデータベース１４２で与えられるデータサービスが定期的にＡＳＲ装置１１０及びホストコンピュータ１２４でアクセスされる実施例等の他の実施例でも良い。

【００４１】

本発明の方法及び装置は、ＡＳＲ装置が話者特定訓練をすることなく任意に話された単語を認識することができるという点で、先行技術に勝る効果が得られる。上記に説明した実施例は例示のためのものにすぎず、本発明の多くの他の改作が、特許請求の範囲で定義されるように本発明の範囲から逸脱することなく当業者により引き出され得ることが理解されるべきである。

【図面の簡単な説明】

【図１】本発明の模範的な実施例によるＡＳＲ装置と関連する電話とユーザーの対話装置の簡略ブロック図である。

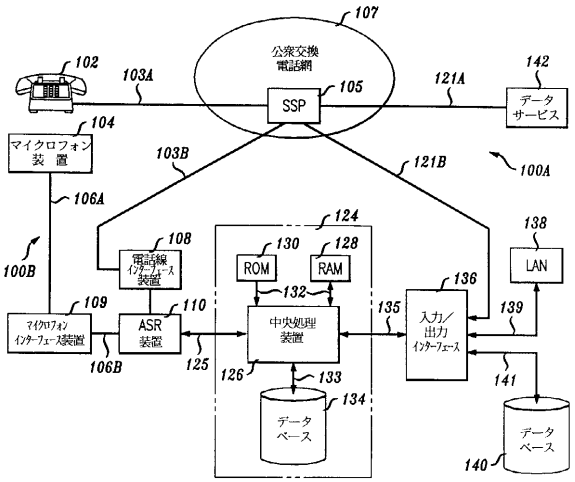
【図２】本発明の模範的な実施例による通話認識及び検定を提供するために用いられる図１のＡＳＲ装置の簡略ブロック図である。

【図３】本発明による模範的な方法のフロー図である。

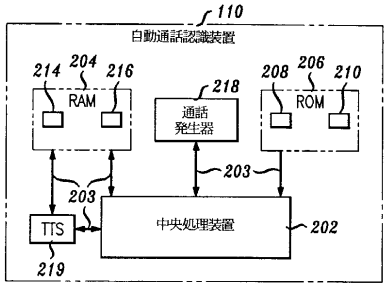
【符号の説明】

１０２　電話端末
１０３Ａ、１０３Ｂ　通信線
１０４　マイクロフォン装置
１０７　公衆交換電話網
１０８　電話線インターフェース装置
１１０　ＡＳＲ装置
１２４　ホストコンピュータ
１３５　データバス
１４１　データリンク

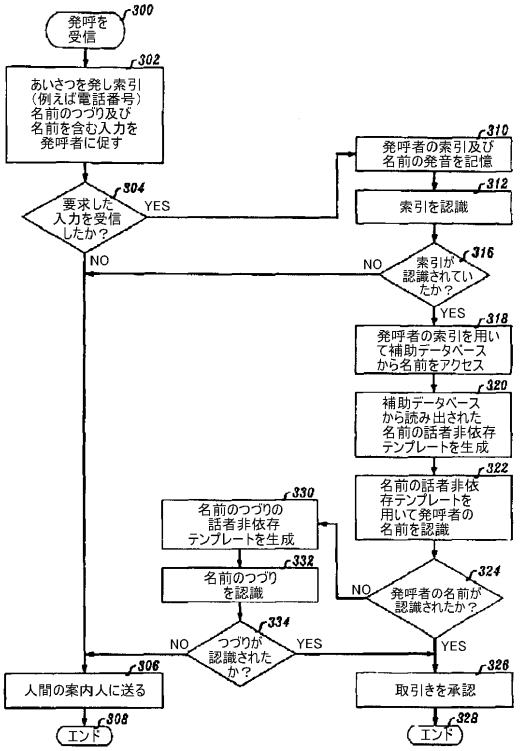
【図１】


107
公衆交換電話網
102
103A
SSP
105
121A
142
データサービス
104
マイクロフォン装置
103B
121B
100A
106A
124
108
電話線インターフェース装置
130
ROM
128
RAM
138
LAN
100B
132
136
109
ASR装置
110
中央処理装置
135
入力／出力インターフェース
139
106B
125
126
133
141
134
データベース
140
データベース

【図２】


110
自動通話認識装置
214
204
216
RAM
218
通話発生器
208
206
210
ROM
203
203
TTS
中央処理装置
202
219

【図３】


300
発呼を受信
302
あいさつを発し索引（例えば電話番号）名前のつづり及び名前を含む入力を発呼者に促す
304
要求した入力を受信したか？
YES
NO
310
発呼者の索引及び名前の発音を記憶
312
索引を認識
316
索引が認識されていたか？
NO
YES
318
発呼者の索引を用いて補助データベースから名前をアクセス
320
補助データベースから読み出された名前の話者非依存テンプレートを生成
322
名前の話者非依存テンプレートを用いて発呼者の名前を認識
330
名前のつづりの話者非依存テンプレートを生成
332
名前のつづりを認識
324
発呼者の名前が認識されたか？
NO
YES
334
つづりが認識されたか？
NO
YES
306
人間の案内人に送る
326
取引きを承認
308
エンド
328
エンド

フロントページの続き

(74)代理人 100091889
弁理士 藤野 育男
(74)代理人 100101498
弁理士 越智 隆夫
(74)代理人 100096688
弁理士 本宮 照久
(74)代理人 100102808
弁理士 高梨 憲通
(74)代理人 100104352
弁理士 朝日 伸光
(74)代理人 100107401
弁理士 高橋 誠一郎
(74)代理人 100106183
弁理士 吉澤 弘司
(72)発明者 ロジャー ボアガン ガーバー
アメリカ合衆国 ６０５４０ イリノイズ，ネイパーヴィル，ノース ライト ストリート ２１９
(72)発明者 マイケル アレン ユドコウスキー
アメリカ合衆国 ６０６４５ イリノイズ，シカゴ，ウエスト ファーゴ ２９５２

合議体
審判長 新宮 佳典
審判官 小松 正
審判官 橋爪 正樹

(56)参考文献 特開昭５８－１８９７００（ＪＰ，Ａ）
特開昭６３－３０４２９８（ＪＰ，Ａ）
特開昭５２－１２７００７（ＪＰ，Ａ）
特開平２－１３５９６２（ＪＰ，Ａ）
特開平２－１４１０５６（ＪＰ，Ａ）
特開平４－３８５５８（ＪＰ，Ａ）
実開昭６３－１７１５７３（ＪＰ，Ｕ）