

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-219506

(P2004-219506A)

(43) 公開日 平成16年8月5日(2004.8.5)

(51) Int.Cl.⁷

G 1 O L 19/00

F I

G 1 O L 9/18

E

テーマコード (参考)

5 D O 4 5

審査請求 未請求 請求項の数 7 O L (全 9 頁)

(21) 出願番号 特願2003-3947 (P2003-3947)
 (22) 出願日 平成15年1月10日 (2003.1.10)

(71) 出願人 000003078
 株式会社東芝
 東京都港区芝浦一丁目1番1号
 (74) 代理人 100083161
 弁理士 外川 英明
 (72) 発明者 井阪 岳彦
 東京都青梅市末広町2丁目9番地 株式会
 社東芝青梅事業所内
 (72) 発明者 三関 公生
 東京都青梅市末広町2丁目9番地 株式会
 社東芝青梅事業所内
 Fターム(参考) 5D045 DA11

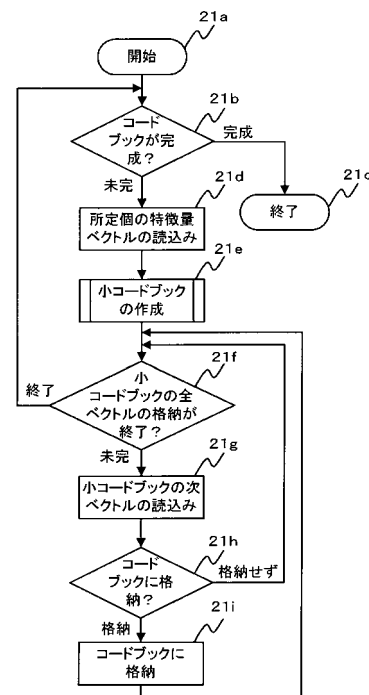
(54) 【発明の名称】 コードブック作成方法、コードブック作成装置及び通信端末装置

(57) 【要約】

【課題】 大きな記憶領域を必要とせず、かつ、多くの計算を必要とせずにコードブックを作成するコードブック作成方法、コードブック作成装置及び通信端末装置を提供する。

【解決手段】 登録用音声を分割し、分割された登録用音声それぞれからその特徴量ベクトルを作成し、代表ベクトル作成装置は、所定個の特徴量ベクトルを読み込み（ステップ21d）、それに対して小コードブックを作成し（ステップ21e）、その中のベクトルをコードブックに格納する（ステップ21i）処理を繰り返すことにより、コードブックを作成する。

【選択図】 図2



【特許請求の範囲】

【請求項 1】

音声信号の特徴量を表すベクトルが格納されるコードブックを作成するコードブック作成方法であって、
前記音声信号を所定時間幅で分割し、
その分割された音声信号毎にその特徴量を表すベクトルを作成し、
その作成された特徴量を表すベクトルからそれらを代表するベクトルを作成し、
その作成された代表するベクトルを前記コードブックに格納することを特徴とするコードブック作成方法。

【請求項 2】

前記代表するベクトルの作成日時を前記コードブックに格納し、
入力された第 2 の音声信号を所定時間幅で分割し、
その分割された第 2 の音声信号毎にその特徴量を表すベクトルを作成し、
その作成された第 2 の音声信号の特徴量を表すベクトルからそれらを代表する第 2 のベクトルを作成し、
前記作成されたベクトルを前記コードブックに格納する際、当該コードブックに格納されてから予め定められた時間が経過しているベクトルの有無を確認し、予め定められた時間を経過しているベクトルがある場合にはこのベクトルを前記コードブックから削除することを特徴とする請求項 1 記載のコードブック作成方法。

【請求項 3】

前記代表するベクトルを前記コードブックに格納する際、
前記代表するベクトルと、前記コードブックに格納されているベクトルとの距離を評価し、
前記距離が予め定められた距離より短い場合は前記代表するベクトルを前記コードブックに格納しないことを特徴とする請求項 1 記載のコードブック作成方法。

【請求項 4】

前記代表するベクトルと最も近い距離にある前記ベクトルの前記コードブックからの削除と、そのベクトルを前記代表するベクトルの方向へ所定距離の移動を行ったベクトルの前記コードブックへの格納を行うことを特徴とする請求項 3 記載のコードブック作成方法。

【請求項 5】

前記代表するベクトルを前記コードブックに格納する際、
前記代表するベクトルと、前記コードブックに格納されているベクトルとの距離を評価し、
前記距離が予め定められた距離より長い場合は前記代表するベクトルを前記コードブックに格納しないことを特徴とする請求項 1 記載のコードブック作成方法。

【請求項 6】

音声信号の特徴量を表すベクトルが格納されるコードブックを作成するコードブック作成装置であって、
前記音声信号を所定時間幅で分割し、当該分割された音声信号毎にその特徴量を表すベクトルを作成する手段と、
前記作成手段によって作成された特徴量を表すベクトルからそれらを代表するベクトルを作成し、当該代表するベクトルを前記コードブックに格納する手段とを具備することを特徴とするコードブック作成装置。

【請求項 7】

送話音声からアナログ送話音声信号を出力するマイクロフォンと、
アナログ受話音声信号から受話音声を発生するスピーカと、
前記マイクロフォンから出力される音声信号を所定時間幅で分割し、当該分割された音声信号毎にその特徴量を表すベクトルを作成する手段と、
前記作成手段によって作成された第 1 の音声信号の特徴量を表すベクトルからそれらを代表するベクトルを作成し、当該代表するベクトルをコードブックに格納する手段と、

前記作成手段によって作成された第2の音声信号の特徴量を表すベクトルと前記コードブックに記憶された前記代表するベクトルとの間の距離を評価する手段とを具備し、前記評価手段によって得られた評価結果が所定の条件を満たす場合に動作を終了することを特徴とする通信端末装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、小さな記憶領域と少ない計算量でコードブックを作成するコードブック作成方法、コードブック作成装置及び通信端末装置に関する。

【0002】

【従来の技術】

音声を入力して、それを発話した利用者を特定する音声認証の手法として、ベクトル量子化の手法が用いられている。この手法を用いる音声認証方法では、利用者が登録用に発話した音声信号の特徴量を抽出して特徴量ベクトルを作成し、それらを代表する代表ベクトルにより構成されるコードブックが利用者毎に設けられている。

【0003】

そして、音声認証の際は、認証用に発話された音声信号の特徴量ベクトルを、あらかじめコードブックに記憶された代表ベクトルと比較し、特徴量ベクトルとそのコードブック中のある代表ベクトルとの間の距離が小さければ、発話者は、そのコードブックに対応する利用者であると判断する。

【0004】

代表ベクトルは、各利用者が登録用に発話した音声信号から作成された特徴量ベクトルを所定の数にクラスタリングし、それぞれのクラスを代表するベクトルとして作成される。ここで、登録用の音声は、あらゆる認証用の発話に含まれる音素を含む必要があり、登録用には長時間の発話が必要である。

【0005】

そのため、コードブックの作成には、多量の音声信号を分析して特徴量ベクトルを作成する処理、及び、特徴量ベクトルのクラスタリングのためにそれらの間の距離を計算する処理などが必要となり、その結果、大きな記憶領域を用いた多くの計算が必要となる。

【0006】

そこで、コードブックの作成には学習方式が用いられている。即ち、仮の代表ベクトルを設定し、登録用音声信号から逐次特徴量ベクトルを作成し、それに基づいて仮の代表ベクトルを改良していくことによって、大きな記憶領域や多くの計算を必要としない処理が行われている（例えば、特許文献1参照。）。

【0007】

【特許文献1】

特開平5 - 207297号公報（第2 - 3頁、図1）

【0008】

【発明が解決しようとする課題】

しかしながら、上述した従来の方法では、仮の代表ベクトルの初期値が適切でない場合には、それが最適な代表ベクトルに収束するまでに長時間を要する問題点があった。

【0009】

本発明は、上記問題点を解決するためになされたもので、大きな記憶領域を必要とせず、かつ、多くの計算を必要とせずにコードブックを作成するコードブック作成方法、コードブック作成装置及び通信端末装置を提供することを目的とする。

【0010】

【課題を解決するための手段】

上記目的を達成するために、本発明のコードブック作成方法は、音声信号の特徴量を表すベクトルが格納されるコードブックを作成するコードブック作成方法であって、前記音声信号を所定時間幅で分割し、その分割された音声信号毎にその特徴量を表すベクトルを作

10

20

30

40

50

成し、その作成された特徴量を表すベクトルからそれらを代表するベクトルを作成し、その作成された代表するベクトルを前記コードブックに格納することを特徴とする。

【0011】

本発明によれば、登録用音声を分割し、分割された登録用音声それぞれからそれを代表するベクトルを作成し、それらを統合してコードブックを作成することにより、大きな記憶領域を用いず、かつ、短い処理時間でコードブックを作成することができる。

【0012】

【発明の実施の形態】

以下に、本発明によるコードブック作成方法の実施の形態を、図面を参照して説明する。

図1は、本発明の実施形態に係るコードブック作成方法が適用された移動通信端末装置の構成を示すブロック図である。 10

【0013】

この移動通信端末装置は、装置全体の制御を行う制御装置1と、アンテナ2aと、通信装置2bと、送受信装置3と、スピーカ4aと、マイクロフォン4bと、通話装置4cと、特徴量抽出装置4dと、特徴量ベクトル記憶部4eと、利用者からの操作指示を入力する入力装置5と、利用者へ情報の提示等を行う出力装置6と、代表ベクトル作成装置7と、小コードブック記憶部7aと、コードブック記憶装置8と、音声認証装置9とからなる。

【0014】

上記のように構成された、本発明の実施形態に係る移動通信端末装置の動作を説明する。制御装置1は、装置全体の制御を行う。通信装置2bは、アンテナ2aが受信した高周波信号を送受信装置3へ出力し、また、送受信装置3が出力する高周波信号をアンテナ2aより送信する。 20

【0015】

次に、送受信装置3の動作を説明する。送受信装置3は、通信装置2bからの高周波信号を増幅、周波数変換及び復調し、それによって得たデジタル音声信号を通話装置4cへ送り、また、制御信号を制御装置1へ送る。送受信装置3は、更に、通話装置4cが出力するデジタル音声信号、及び制御装置1が出力する制御信号を変調、周波数変換及び増幅し、高周波信号を得てそれを通信装置2bへ送る。

【0016】

次に、通話装置4cの動作を説明する。通話装置4cは、送受信装置3が出力するデジタル音声信号をアナログ音声信号に変換し、それを増幅してスピーカ4aに送る。また、マイクロフォン4bが出力するアナログ音声信号を増幅し、それをデジタル音声信号に変換して送受信装置3に送る。 30

【0017】

次に、特徴量抽出装置4dの動作を説明する。特徴量抽出装置4dは、マイクロフォン4bが出力するアナログ音声信号を分析して、その特徴量ベクトルを特徴量ベクトル記憶部4eに格納する装置である。

【0018】

特徴量抽出装置4dは、利用者が入力装置5のキー（図示せず）の操作によってコードブック作成を要求した際、及び、この移動通信端末装置の電源が投入された際に、それによって制御装置1が起動して、動作を開始する。 40

【0019】

その際、制御装置1は、利用者に対して、マイクロフォン4bに向かって発話を促す旨の誘導メッセージを出力装置6に表示し、或いは、スピーカ4aから発生する。また、更に、発話すべきメッセージそのものを出力装置6に表示しても良い。

【0020】

特徴量抽出装置4dは、まず、マイクロフォン4bから出力されるアナログ音声信号を増幅し、それを所定時間幅、例えば20m秒毎に分割し、分割された各々のアナログ音声信号を分析する。

【0021】

即ち、分割されたアナログ音声信号を線形予測分析 (LPC、Linear predictive coding) し、それによって得られたパワースペクトルから計算したケプストラムを分割されたアナログ音声信号の特徴量とし、それらの特徴量をベクトル形式で表した特徴量ベクトルを特徴量ベクトル記憶部 4 e に格納する。

【0022】

続いて、次の分割されたアナログ音声信号の分析を繰り返す。なお、アナログ音声信号の分析方法は、上記の LPC ケプストラム法に限るものではない。

【0023】

次に、入力装置 5 の動作を説明する。入力装置 5 は、複数のキーからなり、制御装置 1 の起動により動作を開始し、利用者が操作したキーを識別するコードを制御装置 1 へ通知して動作を終了する。 10

【0024】

次に、出力装置 6 の動作を説明する。出力装置 6 は、例えば LCD の表示装置であり、制御装置 1 より起動されることにより、表示動作を開始する。入力装置 5 からの入力操作に応答して制御装置 1 より起動されるまで、同じ表示動作を続ける。

【0025】

次に、代表ベクトル作成装置 7 の動作を説明する。代表ベクトル作成装置 7 は、特徴量抽出装置 4 d によって特徴量ベクトル記憶部 4 e に記憶された特徴量ベクトルから、それらを代表する代表ベクトルを作成し、その作成された代表ベクトルをコードブック記憶装置 8 に格納することにより、コードブックを作成する装置である。 20

【0026】

図 2 は、代表ベクトル作成装置 7 の動作のフローチャートを示す。代表ベクトル作成装置 7 は、利用者が入力装置 5 のキーの操作によってコードブックの作成を要求した際に、それに従って制御装置 1 が起動して、動作を開始する (ステップ 21 a)。

【0027】

続いて、コードブックが完成したかを判断し (ステップ 21 b)、もし、それが完成していたら、動作を終了する (ステップ 21 c)。ステップ 21 b のコードブックが完成したかの判断基準は、コードブック記憶装置 8 に所定個のベクトルが記憶されていること、または、コードブックが所定量以上の記憶容量を占めていることである。

【0028】

また、コードブック記憶装置 8 に記憶されたベクトルの中に、作成されてから一定期間経過したものがあれば、それを削除し、それに代わる代表ベクトルを格納することによって完成と判断しても良い。これによって、コードブックの更新が行われ、利用者の発話する音声の経年変化への対応が可能となる。 30

【0029】

一方、ステップ 21 b で、コードブックが未完と判断されるなら、代表ベクトル作成装置 7 は、続いて、所定時間幅に分割された音声信号から作成された特徴量ベクトルを、特徴量ベクトル記憶部 4 e から、所定個読込む (ステップ 21 d)。

【0030】

続いて、読込んだ特徴量ベクトルから、それらを代表する所定個の代表ベクトルを作成し、それを小コードブック記憶部 7 a に格納することによって、小コードブックを作成する (ステップ 21 e)。 40

【0031】

ここで、代表ベクトルは、読込んだ全ての特徴量ベクトルに対して、それぞれから最も近い距離にある代表ベクトルとの間の距離を求め、当該距離の総和が最小になる所定個のベクトルとして計算することによって求めるが、これに限るものではない。

【0032】

なお、2 つのベクトル間の距離は、例えばユークリッド距離、即ち、ベクトルの対応する各要素間の差を二乗し、その総和の平方根として求めるが、これに限るものではない。

【0033】

続いて、代表ベクトル作成装置 7 は、小コードブック記憶部 7 a に記憶された全ての代表ベクトルのコードブック記憶装置 8 への格納が完了したかを判断し（ステップ 2 1 f）、もし、完了していればステップ 2 1 b に戻って処理を繰り返す。一方、完了していなければ、小コードブック記憶部 7 a から次の代表ベクトルを読み込む（ステップ 2 1 g）。

【0034】

続いて、読み込んだ代表ベクトルをコードブック記憶装置 8 に格納するか否かの判断を行う（ステップ 2 1 h）。図 3 は、代表ベクトル 3 1 をコードブック記憶装置 8 に格納するか否かの判断処理の概念を、代表ベクトル 3 1 及びコードブック記憶装置 8 に記憶されたベクトルが 2 次元であると仮定して、即ち、それらのベクトルを点として示す。

【0035】

代表ベクトル 3 1 をコードブック記憶装置 8 に格納するか否かの第 1 の判断基準として、まず、当該代表ベクトル 3 1（ベクトル X とする）とコードブック記憶装置 8 に記憶されたベクトル 3 2（3 2 a はベクトル Y a、3 2 b は Y b とする）の間の距離 3 3（3 3 a の距離は $|X - Y a|$ 、3 3 b の距離は $|X - Y b|$ ）を計算し、最も近いベクトル間の距離を求める。上記の例では、距離 3 3 a $|X - Y a| >$ 距離 3 3 b $|X - Y b|$ であることから、距離 3 3 b が求められる。

【0036】

ここで、この最も近いベクトル間の距離 3 3 b が、所定数 C 1 より小さい場合に、即ち、 $|X - Y b| < C 1$

が成り立つものであるならば、この代表ベクトル 3 1 をコードブック記憶装置 8 に格納すると、コードブック中にお互いの間の距離が短い複数のベクトル（3 1、3 2 b）が存在することになり、無駄なベクトルを多数記憶することになる。そこで、代表ベクトル作成装置 7 は、代表ベクトル 3 1 をコードブック記憶装置 8 に格納せずにステップ 2 1 f に戻り、処理を繰り返す。一方、 $|X - Y b| \geq C 1$ が成り立つ場合は、その代表ベクトル 3 1 をコードブック記憶装置 8 に格納する。

【0037】

なお、所定数 C 1 の値は、音声認証装置 9 の認証誤り率を考慮して、実験により求められる。

【0038】

ここで、代表ベクトル 3 1 をコードブック記憶装置 8 に格納しないのみならず、この代表ベクトル 3 1 に基づいて学習処理を行っても良い。例えば、代表ベクトル 3 1 に最も近いコードブック記憶装置 8 に記憶されたベクトル 3 2 b を、両ベクトルの間を所定の比率で内分するベクトルによって置き換えても良い。

【0039】

また、ステップ 2 1 h における代表ベクトル 3 1 をコードブック記憶装置 8 に格納するか否かの第 2 の判断基準として、代表ベクトル 3 1 とコードブック記憶装置 8 に記憶されたベクトル 3 2 の間の上記最短距離を計算し、それが所定数 C 2 より大きければ、即ち、 $|X - Y b| > C 2$

が成り立つものであるならば、この代表ベクトル 3 1 は発話者の発音ではない、即ち、雑音であると判断して、代表ベクトル作成装置 7 は、代表ベクトル 3 1 をコードブック記憶装置 8 に格納せずにステップ 2 1 f に戻り、処理を繰り返すようにしても良い。一方、 $|X - Y b| \leq C 2$ が成り立つ場合は、その代表ベクトル 3 1 をコードブック記憶装置 8 に格納する。

【0040】

なお、所定数 C 2 の値は、音声認証装置 9 の認証誤り率を考慮して、実験により求められる。

【0041】

更に、ステップ 2 1 h における第 3 の判断基準として、コードブックが既に完成している場合にも、代表ベクトル作成装置 7 は、代表ベクトル 3 1 をコードブック記憶装置 8 に格納せずにステップ 2 1 f に戻り、処理を繰り返すようにしても良い。コードブックが既に

10

20

30

40

50

完成しているか否かの判断基準は、ステップ 2 1 b で説明した通りである。

【 0 0 4 2 】

一方、ステップ 2 1 h の上記の判断基準 1 ~ 3 (1 つの判断基準でも、2 つ以上の判断基準を併せ持っても良い) に基づいて、代表ベクトル 3 1 をコードブック記憶装置 8 に格納すると判断された場合には、代表ベクトル作成装置 7 は、それをコードブック記憶装置 8 に格納して (ステップ 2 1 i) 、ステップ 2 1 f に戻り、処理を繰り返す。

【 0 0 4 3 】

次に、音声認証装置 9 の動作を説明する。音声認証装置 9 は、認証の際の発話者がコードブック作成時の発話者と同じであるか否かを判断する装置である。図 4 は、音声認証装置 9 の動作のフローチャートを示す。

10

【 0 0 4 4 】

音声認証装置 9 は、移動通信端末装置の電源が投入された際に、制御装置 1 がその他の動作に先立って起動することによって動作を開始し (ステップ 4 1 a) 、まず、認証の際の発話者から発せられた音声から特徴量抽出装置 4 d によって上述した手順で作成された特徴量ベクトルを、特徴量ベクトル記憶部 4 e から、所定個読込む (ステップ 4 1 b) 。

【 0 0 4 5 】

続いて、読込んだ特徴量ベクトルのそれぞれに対してコードブック記憶装置 8 に記憶されたベクトルの中で最も近い距離にあるベクトルとの間の距離を求める (ステップ 4 1 c)

。

【 0 0 4 6 】

20

続いて、当該距離の合計値を評価し (ステップ 4 1 d) 、この合計値が所定値以下であれば、認証の際の発話者はコードブック作成時の発話者と同じと判断し、それを制御装置 1 に通知して (ステップ 4 1 e) 、また、それ以外の場合には、異なると判断し、それを制御装置 1 に通知して (ステップ 4 1 f) 、終了する (ステップ 4 1 g) 。

【 0 0 4 7 】

制御装置 1 は、その判断結果の通知を音声認証装置 9 から受けると、その通知が認証の際の発話者はコードブック作成時の発話者と同じとの情報であれば、認証の際の発話者は正当な利用者と判断して、移動通信端末装置の動作を継続させる。

【 0 0 4 8 】

一方、その通知が認証の際の発話者はコードブック作成時の発話者と異なるとの情報であれば、制御装置 1 は、認証の際の発話者は正当な利用者ではないと判断して、正当な利用者の権利の侵害や秘密の漏洩を防止すべく、移動通信端末装置の動作を終了させる。なお、制御装置 1 は、上記動作の終了に先だって、利用者に再度の発話を促して音声認証装置 9 を再度起動して認証を行わせる、また、別の手段によって利用者を特定する、などをして良い。

30

【 0 0 4 9 】

本実施形態では、代表ベクトル 3 1 をコードブック記憶装置 8 に格納するか否かの判断は、代表ベクトル 3 1 と最も近い距離にあるベクトルをコードブック記憶装置 8 に記憶されているベクトルの中から求め、両ベクトル間の距離のみに基づいて行ったが、代表ベクトル 3 1 とコードブック記憶装置 8 に記憶されている複数のベクトルとの間の距離に基づいて行っても良い。

40

【 0 0 5 0 】

また、コードブック記憶装置 8 には、代表ベクトル作成装置 7 の動作の前に、平均的な音声の特徴量を表すベクトルが初期値として記憶されていても良い。それらのベクトルの作成日時としては、十分に古い日時を併せて記憶させる。十分に古い日時を記憶させることによって、代表ベクトル作成装置 7 が利用者の発話する音声の特徴量を表す代表ベクトルをコードブック記憶装置 8 に格納する際に、初期値ベクトルが必ず削除され、利用者の発話する音声の特徴量に基づいたコードブックを効率良く作成できる。

【 0 0 5 1 】

更に、コードブック記憶装置 8 には、1 名の利用者の発話に基づいて作成されたコードブ

50

ック１つが記憶されているとしたが、複数のコードブックが、それぞれを作成する際に発話した利用者の識別情報と共に、記憶されていても良い。その場合には、音声認証装置９は、認証用の発話を行った利用者が上記複数の利用者のいずれであるかを判断する。本発明は以上の構成に限定されるものではなく、種々の変形が可能である。

【００５２】

なお、以上に本発明を移動通信端末装置に適用した例を説明したが、本発明はＰＤＡ（Personal Digital Assistant）などの携帯端末装置にも当然適用可能である。

【００５３】

【発明の効果】

10

以上述べたように、本発明によれば、登録用音声を分割し、分割された登録用音声それぞれからそれを代表するベクトルを作成し、それらを統合してコードブックを作成することにより、大きな記憶領域を用いず、かつ、短い処理時間でコードブックを作成することができる。

【図面の簡単な説明】

【図１】本発明の実施形態に係る移動通信端末装置を示すブロック図。

【図２】本発明の実施形態に係る代表ベクトル作成装置の動作を示すフローチャート。

【図３】代表ブロックとコードブック記憶装置に記憶されたベクトルとの距離を示す概念図。

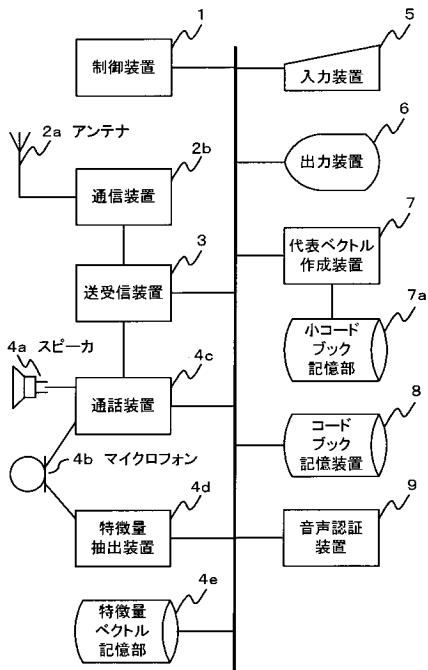
【図４】本発明の実施形態に係る音声認証装置の動作を示すフローチャート。

20

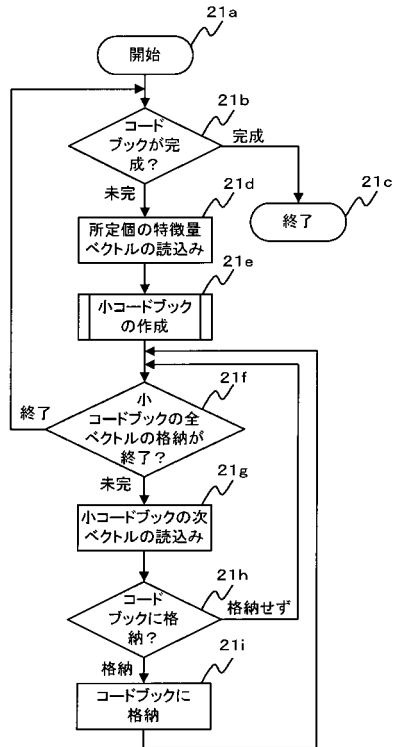
【符号の説明】

- １ 制御装置
- ４ｂ マイクロフォン
- ４ｄ 特徴量抽出装置
- ４ｅ 特徴量ベクトル記憶部
- ７ 代表ベクトル作成装置
- ７ａ 小コードブック記憶部
- ８ コードブック記憶装置
- ９ 音声認証装置

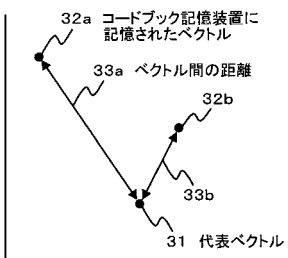
【図 1】



【図 2】



【図 3】



【図 4】

