

(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2004-145025

(P2004-145025A)

(43) 公開日 平成16年5月20日(2004.5.20)

(51) Int.Cl.⁷

G 1 O L 17/00

G 1 O L 15/20

G 1 O L 21/02

F I

G 1 O L 3/00

5 4 5 C

G 1 O L 3/00

5 4 5 D

G 1 O L 3/02

3 O 1 D

テーマコード (参考)

5 D O 1 5

審査請求 未請求 請求項の数 5 O L (全 25 頁)

(21) 出願番号 特願2002-310156 (P2002-310156)

(22) 出願日 平成14年10月24日 (2002.10.24)

(71) 出願人 598098526

アルゼ株式会社

東京都江東区有明3丁目1番地25

(74) 代理人 100106002

弁理士 正林 真之

(72) 発明者 富士本 淳

東京都江東区有明3-1-25 有明フロ

ンティアビルA棟

Fターム(参考) 5D015 AA03

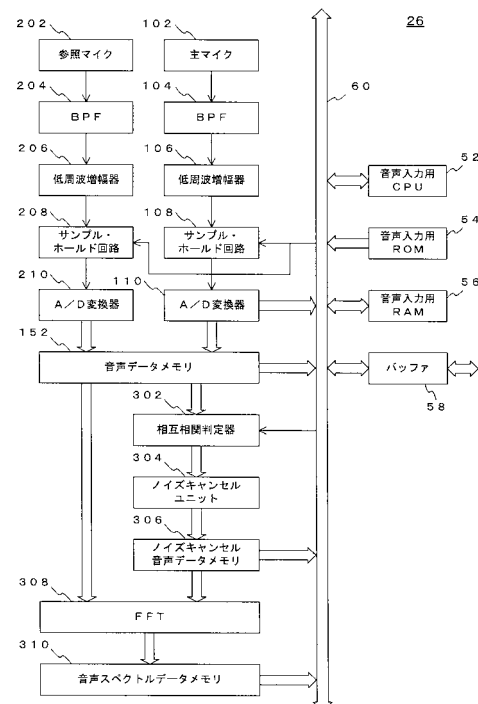
(54) 【発明の名称】 音声照合装置

(57) 【要約】

【課題】 照合される可能性を高めた音声認証装置を提供する。

【解決手段】 音声照合装置10は、照合話者の音声信号を入力するための複数の音声入力手段と、それら複数の音声入力手段の各々において入力された複数の音声信号をそれぞれ相互比較する音声比較手段と、その音声比較手段により比較された結果に基づいて複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、その音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいてその近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、その音声補正手段により補正された音声信号に基づいて声紋、音韻分析を行う音声認識手段と、を備え、音声認識手段により行われた声紋、音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有する。

【選択図】 図4



【特許請求の範囲】

【請求項 1】

登録話者の音声信号と照合話者の音声信号とを照合する音声照合装置であって、照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、
前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有することを特徴とする音声照合装置。

10

【請求項 2】

登録話者の音声信号と照合話者の音声信号とを照合する音声照合装置であって、照合話者の音声信号を入力するための複数の音声入力手段と、入力された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、
前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有するものであり、かつ、
前記音声認識手段が前記複数の入力手段の一より入力された音声信号に基づいて声紋分析及び／又は音韻分析を行った結果が正常に照合されなかった場合には、前記当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、を備え、
前記音声認識手段は、前記音声補正手段により補正された近似の音声信号の少なくとも二つに含まれる一に基づいて声紋分析及び／又は音韻分析を行う音声認識を行うものであり、かつ、
前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを再照合する機能を有することを特徴とする音声照合装置。

20

30

【請求項 3】

登録話者の音声信号と照合話者の音声信号とを照合する音声照合装置であって、照合話者に対して供給される質問情報を記憶している質問情報記憶手段と、当該質問情報記憶手段により記憶されている質問情報を照合話者に供給する質問情報供給手段と、当該質問情報供給手段により供給された質問情報に対する回答情報として照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、
前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有することを特徴とする音声照合装置。

40

【請求項 4】

登録話者の音声信号と照合話者の音声信号とを照合する音声照合装置であって、照合話者に対して供給される質問情報を記憶している質問情報記憶手段と、当該質問情報

50

記憶手段により記憶されている質問情報に対する認証情報を記憶している認証情報記憶手段と、前記質問情報記憶手段により記憶されている質問情報を照合話者に供給する質問情報供給手段と、当該質問情報供給手段により供給された質問情報に対する回答情報として照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、
前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果と、前記認証情報記憶手段により記憶されている認証情報と、に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有することを特徴とする音声照合装置。

10

【請求項 5】

サーバと通信可能に接続されており、登録話者の音声信号と照合話者の音声信号とを照合する音声照合装置であって、
照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、音声認識手段により行われた声紋分析及び／又は音韻分析の結果を前記サーバに対して供給する音声認識結果供給手段と、前記サーバが前記音声認識結果供給手段から供給された声紋分析及び／又は音韻分析の結果に基づいて登録話者の音声信号と照合話者の音声信号とを照合した結果を受信する照合結果受信手段と、を備えたことを特徴とする音声照合装置。

20

【発明の詳細な説明】**【0001】****【発明の属する技術分野】**

30

本発明は、音声が入力されることにより照合を行う音声照合装置に関するものである。

【0002】**【従来の技術】**

一般に音声照合装置では、登録話者の音声信号と照合話者の音声信号とに基づいて照合を行い、その照合の結果に基づいて、何らかの認証処理を行うものである。

【0003】

このような装置は、入力された音声信号をデジタル化し、デジタル化された音声データに対して所定の処理を行うことにより、音韻、声紋等、様々な音声要素を取り出すことができ、それに基づいて、音声認識を行うことが可能となるのである。

【0004】

40

また、本人照合を音声により行う音声照合装置が提案されており（例えば、特許文献1）、短時間の音声信号データに対しても、照合率の高い話者照合を行うことが可能となる。

【0005】**【特許文献1】**

特開2002-23792号公報

【0006】**【発明が解決しようとする課題】**

しかしながら、このような装置では、必ずしも、音声信号が正常に照合されとは限らない。これには、入力された音声信号が良質なものであればよいが、そのような音声ばかりではなく、如何に良質な音声信号を取り込み、如何に必要な音声信号を取り除かない、歪

50

ませないといった各種の要素が必要となってくる。上述した装置においては、この要素を含むには未だ充分ではなく、必ずしも、音声信号が正常に照合されるとは限らないのである。

【0007】

また、このような装置においては、声紋を分析し、照合話者を特定することが可能であるが、照合話者を偽る者により、その照合話者の音声録音した音声信号が入力されるおそれがあり、このような課題が解決されることが望まれている。

【0008】

本発明は、上述した如き課題に鑑みてなされたものであり、照合される可能性を高めた音声照合装置を提供することを目的とする。

10

【0009】

【課題を解決するための手段】

以上のような目的を達成するために、本発明の音声照合装置は、照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有することを特徴とする。

20

【0010】

より具体的には、本発明は、以下のようなものを提供する。

【0011】

(1) 登録話者の音声信号と照合話者の音声信号とを照合する音声照合装置であって、照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有することを特徴とする音声照合装置。

30

【0012】

(1)の発明によれば、「照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有する」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声が集めることができる。これにより、照合が成功する可能性を高めた音声照合装置を提供することができる。

40

【0013】

50

(2) 登録話者の音声信号と照合話者の音声信号とを照合する音声照合装置であって、照合話者の音声信号を入力するための複数の音声入力手段と、入力された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有するものであり、かつ、前記音声認識手段が前記複数の入力手段の一より入力された音声信号に基づいて声紋分析及び／又は音韻分析を行った結果が正常に照合されなかった場合には、前記当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、を備え、前記音声認識手段は、前記音声補正手段により補正された近似の音声信号の少なくとも二つに含まれる一に基づいて声紋分析及び／又は音韻分析を行う音声認識を行うものであり、かつ、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを再照合する機能を有することを特徴とする音声照合装置。 10

【0014】

(2)の発明によれば、「照合話者の音声信号を入力するための複数の音声入力手段と、入力された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有するものであり、かつ、前記音声認識手段が前記複数の入力手段の一より入力された音声信号に基づいて声紋分析及び／又は音韻分析を行った結果が正常に照合されなかった場合には、前記当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、を備え、前記音声認識手段は、前記音声補正手段により補正された近似の音声信号の少なくとも二つに含まれる一に基づいて声紋分析及び／又は音韻分析を行う音声認識を行うものであり、かつ、前記音声認識手段により行われた 20
声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを再照合する機能を有する」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声が集めることができる。これにより、照合が成功する可能性を高めた音声照合装置を提供することができる。 30

【0015】

また、照合が失敗した場合において、このノイズを除去する補正が行われるため、余計な処理を行うことなく、必要に応じて、自動的に良質な音声を収集することとなり、音声照合装置の処理効率が高まるケースが増加する。

【0016】

(3) 登録話者の音声信号と照合話者の音声信号とを照合する音声照合装置であって、照合話者に対して供給される質問情報を記憶している質問情報記憶手段と、当該質問情報記憶手段により記憶されている質問情報を照合話者に供給する質問情報供給手段と、当該質問情報供給手段により供給された質問情報に対する回答情報として照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分 50

析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有することを特徴とする音声照合装置。

【0017】

(3)の発明によれば、「照合話者に対して供給される質問情報を記憶している質問情報記憶手段と、当該質問情報記憶手段により記憶されている質問情報を照合話者に供給する質問情報供給手段と、当該質問情報供給手段により供給された質問情報に対する回答情報として照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有する」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声が集めることができる。これにより、照合が成功する可能性を高めた音声照合装置を提供することができる。

【0018】

また、この音声認識手段は、登録、記憶された登録話者の声紋情報と、照合話者の声紋情報と、を比較することにより、照合を行うこととなる。このため、より多くの音声を入力できれば、より一層、照合が成功する可能性を高めた音声照合装置を提供することができる。

【0019】

更には、前記質問情報記録手段においては、照合話者に対して供給される質問を複数記憶することが可能であり、当該質問情報記憶手段により記憶されている複数の質問情報から、照合話者に対して供給される一の質問情報を決定する質問決定手段を備えてもよく、この場合には、各種の質問を照合話者に対して供給することが可能であり、それらの各種の質問に基づく、種々の回答情報を得ることができる。このため、より多くの音声が入力され、より一層、照合が成功する可能性を高めた音声照合装置を提供することができる。

【0020】

(4)登録話者の音声信号と照合話者の音声信号とを照合する音声照合装置であって、照合話者に対して供給される質問情報を記憶している質問情報記憶手段と、当該質問情報記憶手段により記憶されている質問情報に対する認証情報を記憶している認証情報記憶手段と、前記質問情報記憶手段により記憶されている質問情報を照合話者に供給する質問情報供給手段と、当該質問情報供給手段により供給された質問情報に対する回答情報として照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果と、前記認証情報記憶手段により記憶されている認証情報と、に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有することを特徴とする音声照合装置。

【0021】

(4)の発明によれば、「照合話者に対して供給される質問情報を記憶している質問情報

記憶手段と、当該質問情報記憶手段により記憶されている質問情報に対する認証情報を記憶している認証情報記憶手段と、前記質問情報記憶手段により記憶されている質問情報を照合話者に供給する質問情報供給手段と、当該質問情報供給手段により供給された質問情報に対する回答情報として照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果と、前記認証情報記憶手段により記憶されている認証情報と、に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有する」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声が集めることができる。これにより、照合が成功する可能性を高めた音声照合装置を提供することができる。

10

【0022】

また、この音声認識手段は、登録、記憶された登録話者の声紋情報と、照合話者の声紋情報と、を比較することにより、照合を行うこととなる。このため、より多くの音声を入力できれば、より一層、照合が成功する可能性を高めた音声照合装置を提供することができる。

20

【0023】

更には、前記質問情報記録手段においては、照合話者に対して供給される質問を複数記憶することが可能であり、当該質問情報記憶手段により記憶されている複数の質問情報から、照合話者に対して供給される一の質問情報を決定する質問決定手段を備えてもよく、この場合には、各種の質問を照合話者に対して供給することが可能であり、それらの各種の質問に基づく、種々の回答情報を得ることができる。このため、より多くの音声が入力され、より一層、照合が成功する可能性を高めた音声照合装置を提供することができる。

【0024】

更にまた、この音声認識手段は、登録、記憶された登録された音韻と、照合話者の音韻と、を比較するだけでなく、登録、記憶された登録話者の声紋情報と、照合話者の声紋情報と、を比較することにより、照合を行うこととなる。

30

【0025】

このため、照合の際に、その登録話者とは異なる照合話者が回答情報を入力した場合であっても、照合することがないため、行われ得る「なりすまし」を防止することが可能であり、セキュリティが高まる。

【0026】

(5) サーバと通信可能に接続されており、登録話者の音声信号と照合話者の音声信号とを照合する音声照合装置であって、照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、音声認識手段により行われた声紋分析及び／又は音韻分析の結果を前記サーバに対して供給する音声認識結果供給手段と、前記サーバが前記音声認識結果供給手段から供給された声紋分析及び／又は音韻分析の結果に基づいて登録話者の音声信号と照合話者の音声信号とを照合した結果を受信する照合結果受信手段と、を備えたことを特徴とする音声照合装置。

40

【0027】

50

(5)の発明によれば、「照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、音声認識手段により行われた声紋分析及び／又は音韻分析の結果を前記サーバに対して供給する音声認識結果供給手段と、前記サーバが前記音声認識結果供給手段から供給された声紋分析及び／又は音韻分析の結果に基づいて登録話者の音声信号と照合話者の音声信号とを照合した結果を受信する照合結果受信手段と、を備えた」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声

10

【0028】

また、このようなサーバに対して、複数の音声照合装置を接続することが可能となるので、サーバに登録話者として登録しておくことにより、そのサーバに接続された音声照合装置に対して、照合を行うことが可能となり、簡便である。

【0029】

【発明の実施の形態】

以下に、本発明に好適な一実施形態について図面に基づいて説明する。

【0030】

〔音声照合装置の構成〕

図1は、本発明の音声照合装置を示すブロック図である。

【0031】

音声照合装置10は、図1に示す如く、データバスBUSに、メインCPU21、メインメモリ22、液晶表示パネルやCRT等で構成された表示部23、キーボードやジョグダイヤル等である入力部24、通信インターフェイス（以降、通信I/Fと称する。）25、音声入力装置26、ハードディスクとそのドライブ等で構成された記憶部27が、データバスBUSを介して接続されて構成されている。

30

【0032】

メインCPU21はメインメモリ22に格納されている動作プログラムに従って種々の動作を実行するようになされており、当該動作に応じて各回路部を制御する。メインCPU21の各種処理内容は必要に応じて表示部23に表示される。

【0033】

また、通信I/F25には、後述する如く、サーバ81等を接続可能であり、そのサーバ81を介して照合処理を行ってもよい。

【0034】

〔音声照合装置における記憶部の構成〕

また、上述した記憶部27には、図2に示す如く、データベースが記憶されている。

【0035】

この記憶部27には、図2に示す如く、登録話者データベースが記憶されている。

【0036】

この登録話者データベースは、この音声照合装置10を用いる登録話者の情報が関連付けて記憶されるものであり、これらの項目には、「ID番号」、「質問ID」、「回答情報」、「声紋情報ID」等、各種の情報が関連付けられている。

【0037】

この「質問ID」とは、質問番号を示すものであり、図3に示す如く、記憶部27に記憶されている、「質問ID」と、「質問情報」と、を関連付けた質問テーブルを参照するこ

50

とにより、その質問の内容が決定される。

【0038】

また、この「回答情報」とは、上述した「質問ID」に対応した情報であり、登録話者により利用登録時に登録されたものであり、後述する音声入力装置26から入力される音声信号に基づいて登録され、又は、入力部24の操作に基づいて登録されるものである。

【0039】

尚、この「質問ID」により関連付けられている「質問情報」と、「回答情報」と、が意味合いを持つように関連付けられていてもよく、もちろん、何ら意味合いを持つことなく関連付けられていてもよい。例えば、「質問情報」が、「出身地は？」というものであった場合には、意味合いを持つ「岐阜県です」という回答情報であっても、意味合いを持たない「水曜日です」という回答情報であってもよい。 10

【0040】

更には、「声紋情報ID」とは、声紋情報の番号を示すものであり、記憶部27に記憶されている、「声紋情報ID」と、「声紋情報」と、を関連付けた声紋情報テーブルを参照することにより、その声紋情報の内容が決定される。この声紋情報テーブルは、記憶部27に記憶されており、登録話者により利用登録時に登録されたものであり、後述する音声入力装置26から入力される音声信号に基づいて登録され、又は、入力部24の操作に基づいて登録されるものである。

【0041】

〔音声入力装置の構成〕

20

図4は、本発明の音声入力装置を示すブロック図である。

【0042】

音声入力装置26は、図4に示す如く、主マイク102と、参照マイク202と、を備えている。尚、図4は、一つの参照マイク202を示したが、実際には、参照マイク202をはじめとする複数の参照マイクが備えられている。

【0043】

この主マイク102は、主となる音声信号（以降、主音声信号と称する。）を受け取るものである。

【0044】

また、参照マイク202は、主となる音声信号を受け取る主マイク102とは異なり、主マイク102から供給される主音声信号を補正するために参照する音声信号（以降、参照音声信号と称する。）を受け取るものである。 30

【0045】

主マイク102には、バンドパスフィルタ（以降、BPFと称する。）104が接続されている。このBPF104は、所定の周波数帯に含まれる信号のみを通過させるフィルタであり、男性、女性を問わず、遊技者の発する音声の周波数帯域を通過させ、その周波数帯域とは異なる周波数帯域の信号を予め減衰させることができる。このため、主マイク102から供給された主音声信号を受け取り、その主音声信号をフィルタリングした後、出力することとなる。

【0046】

40

また、参照マイク202にも、バンドパスフィルタ（以降、BPFと称する。）204が接続されている。

【0047】

このBPF104も、所定の周波数帯に含まれる信号のみを通過させるフィルタであり、男性、女性を問わず、遊技者の発する音声の周波数帯域を通過させ、その周波数帯域とは異なる周波数帯域の信号を予め減衰させることができる。このため、参照マイク202から供給された参照音声信号を受け取り、その参照音声信号をフィルタリングした後、出力することとなる。

【0048】

このBPF104には、低周波増幅器106が接続されている。この低周波増幅器106 50

は、B P F 1 0 4 を通過した主音声信号の所定の低周波成分を増幅するものである。これにより、この処理をすることにより、音声認識を的確に行うことができる信号に変換することができるのである。

【 0 0 4 9 】

また、B P F 2 0 4 にも、低周波増幅器 2 0 6 が接続されている。この低周波増幅器 2 0 6 も、B P F 2 0 4 を通過した参照音声信号の所定の低周波成分を増幅するものである。これにより、この処理をすることにより、音声認識を的確に行うことができる信号に変換することができるのである。

【 0 0 5 0 】

この低周波増幅器 1 0 6 には、サンプル・ホールド回路 1 0 8 が接続されている。このサンプル・ホールド回路 1 0 8 は、入出力バス 6 0 に接続されており、入出力バス 6 0 から発せられる制御信号に応じて、変換すべきアナログ信号の値を的確なものするために行う。

【 0 0 5 1 】

入出力バス 6 0 は、音声入力用 C P U 5 2 にデータ信号又はアドレス信号が入出力されるようになされている。

【 0 0 5 2 】

上述した入出力バス 6 0 には、音声入力用 R O M 5 4 及び音声入力用 R A M 5 6 も接続されている。音声入力用 R O M 5 4 は、各種の処理を実行するためのプログラムや、このプログラムを実行する際に用いられる初期データを記憶する。

【 0 0 5 3 】

また、音声入力用 R A M 5 6 は、上述したプログラムの実行中に用いられるフラグや変数の値を記憶する。

【 0 0 5 4 】

上述したサンプル・ホールド回路 1 0 8 は、入出力バス 6 0 に接続されており、音声入力用 C P U 5 2 は、所定のタイミング毎にサンプル・ホールド回路 1 0 8 に制御信号を発するのである。

【 0 0 5 5 】

また、低周波増幅器 2 0 6 にも、サンプル・ホールド回路 2 0 8 が接続されている。このサンプル・ホールド回路 2 0 8 も、入出力バス 6 0 に接続されており、入出力バス 6 0 から発せられる制御信号に応じて、変換すべきアナログ信号の値を的確なものするために行う。

【 0 0 5 6 】

サンプル・ホールド回路 2 0 8 も、入出力バス 6 0 に接続されており、音声入力用 C P U 5 2 は、所定のタイミング毎にサンプル・ホールド回路 1 0 8 に制御信号を発するのである。

【 0 0 5 7 】

このようなサンプル・ホールド回路 1 0 8 及び 2 0 8 を用いることにより、各音声信号の立ち上がり時等においても所定のタイミングで主音声信号及び参照音声信号を出力することとなり、主音声信号及び参照音声信号の同期をとることとなるのである。

【 0 0 5 8 】

このサンプル・ホールド回路 1 0 8 には、A / D 変換器 1 1 0 が接続されている。A / D 変換器 1 1 0 は、アナログ信号をデジタル変換するものであり、サンプル・ホールド回路 1 0 8 から主音声信号を受け取り、デジタル変換し、主音声データを出力する。

【 0 0 5 9 】

また、サンプル・ホールド回路 2 0 8 にも、A / D 変換器 2 1 0 が接続されている。A / D 変換器 2 1 0 は、アナログ信号をデジタル変換するものであり、サンプル・ホールド回路 2 0 8 から参照音声信号を受け取り、デジタル変換し、参照音声データを出力する。

【 0 0 6 0 】

これらの A / D 変換器 1 1 0 及び 2 1 0 には、音声データメモリ 1 5 2 が接続されている

。この音声データメモリ152は、上述した如く、フィルタリング処理、増幅処理等の各種のノイズ除去処理を行った後に、デジタル変換された主音声データ、参照音声データが、記録されるためのものである。また、音声データメモリ152は、入出力バス60にも接続されており、音声入力用CPU52により、音声データが読み取られることもある。尚、この音声データメモリ152は、主音声データ及び参照音声データを共用して記録することとなる。

【0061】

このように、処理を実行することにより、主マイク102及び参照マイク202に入力された音声信号を音声データとして取り扱うことができるのである。

【0062】

また、この音声データメモリ152には、相互相関判定器302が接続されている。相互相関判定器302は、主マイク102から供給された主音声データと、参照マイク202を含む複数の参照マイクから供給された複数の参照データと、を音声データメモリ152から読み出し、主音声データと、複数の参照データのいずれかと、を相互比較するものである。

【0063】

また、相互相関判定器302は、その比較の結果に基づいて、複数の参照データから、主音声データと一番類似している一を抽出し、主音声データと、その抽出された参照音声データ（以降、抽出参照音声データ）と、を出力するものである。

【0064】

この相互相関判定器302には、ノイズキャンセルユニット304が接続されている。このノイズキャンセルユニット304は、相互相関判定器302から供給された主音声データと、抽出参照音声データと、を受け取り、それらを比較することにより、主音声データを補正し、ノイズを除去するものである。尚、このノイズキャンセルユニット304は、主音声データと抽出参照音声データと、の差を少なくするように補正するものである。

【0065】

このノイズキャンセルユニット304には、ノイズキャンセル音声データメモリ306が接続されている。このノイズキャンセル音声データメモリ306は、上述した如く、ノイズキャンセルユニット304により主音声信号に対してノイズキャンセル処理を行った後の主音声データを記録するものである。また、このノイズキャンセル音声データメモリ306は、入出力バス60に接続されており、音声入力用CPU52により、ノイズキャンセル済みの音声データが読み取られることもある。

【0066】

このノイズキャンセル音声データメモリ306には、高速フーリエ変換処理部（以降、FFTと称する。）308が接続されている。このFFT308は、ノイズキャンセル音声データメモリ306に記録された主音声データを読み出し、高速フーリエ変換を行い、変換後のデータを音声スペクトルデータとして出力することとなる。これにより、音声スペクトルデータを抽出することとなるのである。

【0067】

また、FFT308は、音声データメモリ152にも直接接続されており、上述した相互相関処理及びノイズキャンセル処理を行うことなく、高速フーリエ変換処理を実行することも可能である。

【0068】

このFFT308には、音声スペクトルデータメモリ310が接続されている。この音声スペクトルデータメモリ310は、ノイズキャンセルされた主音声データを高速フーリエ変換を行った音声スペクトルデータを記録するものである。このようにすることにより、主マイク102から取り込まれた音声信号を周波数とその強度との関係に変換することができるのである。また、音声スペクトルデータメモリ310は、入出力バス60に接続されており、音声入力用CPU52により、高速フーリエ変換済みの音声データが読み取られることもある。

【0069】

また、このような音声データ、ノイズキャンセルされた音声データ、ノイズキャンセルされた音声スペクトルデータを受け取った音声入力用CPU52は、バッファ58を介して、メインCPU21に供給することとなり、これらのデータを受け取ったメインCPU21は、図2及び図3のデータベース及びテーブルを用いて、声紋分析及び音韻分析を行った結果、照合処理を行うこととなるのである。

【0070】

また、上述した如く、バッファ58を構成としたことにより、音声データや抽出音声データ等のデータのデータ量が大きいものであっても、的確にメインCPU21に供給することができるのである。

10

【0071】

尚、上述した音声入力装置26は、デジタル信号プロセッサ(DSP)により構成することとしてもよい。

【0072】

[遊技機の動作説明]

以下においては、音声照合装置10は起動しており、上述したメインCPU21、音声入力用CPU52において用いられる変数は所定の値に初期化され、定常動作しているものとする。

【0073】

上述したメインCPU21において実行されるサブルーチンを図5に示す。このサブルーチンは、所定のタイミング毎に呼び出されて実行されるものである。

20

【0074】

最初に、図5に示す如く、照合開始信号を受信したか否かを判断する(ステップS11)。この処理において、メインCPU21は、入力部24からの入力信号に基づいて、ID番号が入力された場合には、照合開始信号を受信したと判断することとなる。この処理において、メインCPU21は、照合開始信号を受信したと判別した場合には、ステップS12に処理を移し、照合開始信号を受信したとは判別しなかった場合には、本サブルーチンを終了する。

【0075】

次いで、ステップS11の判断により照合開始信号を受信したと判別した場合には、質問情報選択処理を実行する(ステップS12)。この処理において、メインCPU21は、ランダムで質問IDをメインメモリ22に記録する。そして、メインCPU21は、メインメモリ22に記録した質問IDと、質問テーブルと、に基づいて、質問情報を決定することとなる。この処理が終了した場合には、ステップS13に処理を移す。

30

【0076】

次いで、質問情報供給処理を実行する(ステップS13)。この処理において、メインCPU21は、ステップS12の処理によりメインメモリ22に記録された質問IDと、質問テーブルと、に基づいて、質問情報を照合話者に供給することとなる。つまり、メインCPU21は、その質問情報を示す画像を表示部23に表示させたり、質問情報を示す音声を音声発生部(図示せず)から発しさせたりする。この処理が終了した場合には、ステップS21に処理を移す。

40

【0077】

次いで、複数のマイクにより音声入力処理を実行する(ステップS21)。この処理において、主マイク102及び参照マイク202は、各々、音声を収音し、その収集した音声を示す、主音声信号、参照音声信号を、BPF104、204に供給する。この処理が終了した場合には、ステップS22に処理を移す。

【0078】

次いで、音声信号フィルタ処理を実行する(ステップS22)。この処理において、BPF104、204は、各々、主マイク102、参照マイク202から受け取った、主音声信号、参照音声信号に対して、所定の周波数帯に含まれる信号のみを通過させ、その周波

50

数帯域とは異なる周波数帯域の信号を予め減衰させ、フィルタリングされた主音声信号、参照音声信号を低周波増幅器 106、206 に供給する。この処理が終了した場合には、ステップ S 23 に処理を移す。

【0079】

次いで、低周波増幅処理を実行する（ステップ S 23）。この処理において、低周波増幅器 106、206 は、BPF 104、204 から供給された主音声信号、参照音声信号に対して、所定の低周波成分を増幅させる。これにより、この処理をすることにより、音声認識を的確に行うことができる信号に変換することができるのである。そして、低周波増幅器 106、206 は、増幅処理が行われた主音声信号、参照音声信号をサンプル・ホールド回路 108、208 に供給する。この処理が終了した場合には、ステップ S 24 に処理を移す。 10

【0080】

次いで、サンプル・ホールド処理を実行する（ステップ S 24）。この処理において、サンプル・ホールド回路 108、208 は、低周波増幅器 106、206 から供給された主音声信号、参照音声信号を、音声入力用 CPU 52 から供給される制御信号に基づいたタイミングで A/D 変換器 110、210 に供給することとなり、主音声信号、参照音声信号の同期をとることとなるのである。この処理が終了した場合には、ステップ S 25 に処理を移す。

【0081】

次いで、A/D 変換処理を実行する（ステップ S 25）。この処理において、A/D 変換器 110、210 は、サンプル・ホールド回路 108、208 から供給された主音声信号、参照音声信号を受け取り、デジタル変換する。この処理が終了した場合には、ステップ S 26 に処理を移す。 20

【0082】

次いで、音声データ記録処理を実行する（ステップ S 26）。この処理において、A/D 変換器 110、210 は、デジタル変換された主音声データ、参照音声データを音声データメモリ 152 に記録する。この処理が終了した場合には、ステップ S 27 に処理を移す。

【0083】

次いで、図 6 に示す如く、相互相関判定処理を実行する（ステップ S 27）。この処理において、相互相関判定器 302 は、主マイク 102 から供給された主音声データと、参照マイク 202 を含む複数の参照マイクから供給された複数の参照データと、を音声データメモリ 152 から読み出す。そして、相互相関判定器 302 は、主音声データと、複数の参照データのいずれかと、を比較し、その比較の結果、複数の参照データから、主音声データと一番類似している一を抽出する。そして、相互相関判定器 302 は、主音声データと、抽出された参照音声データ（以降、抽出参照音声データ）と、をノイズキャンセルユニット 304 に供給する。この処理が終了した場合には、ステップ S 28 に処理を移す。 30

【0084】

次いで、ノイズキャンセル処理を実行する（ステップ S 28）。この処理において、ノイズキャンセルユニット 304 は、相互相関判定器 302 から供給された主音声データと、抽出参照音声データと、を受け取り、それらを比較する。そして、ノイズキャンセルユニット 304 は、主音声データに対して、主音声データと、抽出参照音声データと、の差を少なくするように補正する。この処理が終了した場合には、ステップ S 29 に処理を移す。 40

【0085】

次いで、音声データ記録処理を実行する（ステップ S 29）。この処理において、ノイズキャンセルユニット 304 は、ステップ S 28 の処理によりノイズキャンセルされた主音声データを、ノイズキャンセル音声データメモリ 306 に記録する。この処理が終了した場合には、ステップ S 30 に処理を移す。

【0086】

50

次いで、音声データ高速フーリエ変換処理を実行する（ステップS30）。この処理において、FFT308は、ノイズキャンセル音声データメモリ306に記録された主音声データを読み出し、その主音声データに対して、高速フーリエ変換を行う。この処理が終了した場合には、ステップS31に処理を移す。

【0087】

次いで、音声スペクトルデータ記録処理を実行する（ステップS31）。この処理において、FFT308は、ステップS30の処理により変換された音声スペクトルデータを、音声スペクトルデータメモリ310に記録する。この処理が終了した場合には、ステップS32に処理を移す。

【0088】

次いで、所定時間経過したか否かを判断する（ステップS32）。この処理において、メインCPU21は、メインCPU21に内蔵されているタイマにより、質問情報を提供してから所定時間経過したか否かを判断することとなる。この処理において、メインCPU21は、所定時間経過したと判別した場合には、ステップS41に処理を移し、所定時間経過したとは判別しなかった場合には、ステップS21に処理を移す。

【0089】

次いで、ステップS32の判断により所定時間経過したと判別した場合には、音声スペクトルデータ抽出処理を実行する（ステップS41）。この処理において、メインCPU21は、音声入力用CPU52に対して、音声スペクトルデータを読み出すように命令する。この命令を受けた音声入力用CPU52は、音声スペクトルデータメモリ310に記録されている音声スペクトルデータを読み出し、メインCPU21に供給する。そして、メインCPU21は、その音声スペクトルデータから声紋情報を抽出する。この処理が終了した場合には、ステップS42に処理を移す。

【0090】

次いで、回答情報生成処理を実行する（ステップS42）。この処理において、メインCPU21は、ステップS41の処理により受け取った音声スペクトルデータから音韻分析を行い、回答情報を生成することとなる。この処理が終了した場合には、ステップS43に処理を移す。

【0091】

次いで、データベース参照処理を実行する（ステップS43）。この処理において、メインCPU21は、記憶部27に記録されている登録話者データベースを参照する。尚、この処理においては、メインCPU21は、ステップS11の処理により、ID番号が入力された場合には、そのID番号と、ステップS12の処理により決定された質問番号と、に基づいて、登録話者データベースの所定のデータを読み出すこととなる。この処理が終了した場合には、ステップS44に処理を移す。

【0092】

次いで、音声照合処理を実行する（ステップS44）。この処理において、メインCPU21は、ステップS41の処理により抽出された声紋情報と、ステップS43の処理により行われた参照結果と、を比較し、類似しているか否かを決定する。また、メインCPU21は、ステップS42の処理により生成された音韻情報と、ステップS43の処理により行われた参照結果と、を比較し、類似しているか否かを決定する。このように、メインCPU21は、声紋情報、音韻情報、両者の照合が正常に行われることにより、照合結果が成立したことになる。

【0093】

尚、本実施形態においては、声紋情報、音韻情報、両者の照合が正常に行われることにより、照合結果が成立したことになる好適な一例であったが、本発明はこれに限らず、声紋情報、音韻情報、いずれかの照合が正常に行われることにより、照合結果が成立したとしても何ら問題なく、声紋情報の照合が正常に行われることにより、照合結果が成立することが好ましい。この処理が終了した場合には、ステップS51に処理を移す。

【0094】

次いで、図 7 に示す如く、照合が成立したか否かを判断する（ステップ S 5 1）。この処理において、メイン C P U 2 1 は、ステップ S 4 4 の処理により照合された結果に基づいて判断することとなる。この処理において、メイン C P U 2 1 は、照合が成立したと判別した場合には、ステップ S 5 2 に処理を移し、照合が成立したとは判別しなかった場合には、ステップ S 5 3 に処理を移す。

【0095】

次いで、ステップ S 5 1 の判断により照合が成立したと判別した場合には、照合成立表示処理を実行する（ステップ S 5 2）。この処理において、メイン C P U 2 1 は、照合が成立した旨の画像を表示部 2 3 に表示させる。この処理が終了した場合には、本サブルーチンを終了する。

10

【0096】

一方、ステップ S 5 1 の判断により照合が成立したとは判別しなかった場合には、照合不成立表示処理を実行する（ステップ S 5 3）。この処理において、メイン C P U 2 1 は、照合が成立しなかった旨の画像を表示部 2 3 に表示させる。この処理が終了した場合には、本サブルーチンを終了する。

【0097】

これにより、「照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有する」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声

20

【0098】

また、「照合話者に対して供給される質問情報を記憶している質問情報記憶手段と、当該質問情報記憶手段により記憶されている質問情報を照合話者に供給する質問情報供給手段と、当該質問情報供給手段により供給された質問情報に対する回答情報として照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有する」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声

30

40

【0099】

また、この音声認識手段は、登録、記憶された登録話者の声紋情報と、照合話者の声紋情報と、を比較することにより、照合を行うこととなる。このため、より多くの音声を入力できれば、より一層、照合が成功する可能性を高めた音声照合装置を提供することができる。

50

【0100】

更には、前記質問情報記録手段においては、照合話者に対して供給される質問を複数記憶することが可能であり、当該質問情報記憶手段により記憶されている複数の質問情報から、照合話者に対して供給される一の質問情報を決定する質問決定手段を備えてもよく、この場合には、各種の質問を照合話者に対して供給することが可能であり、それらの各種の質問に基づく、種々の回答情報を得ることができる。このため、より多くの音声が入力され、より一層、照合が成功する可能性を高めた音声照合装置を提供することができる。

【0101】

更には、「照合話者に対して供給される質問情報を記憶している質問情報記憶手段と、当該質問情報記憶手段により記憶されている質問情報に対する認証情報を記憶している認証情報記憶手段と、前記質問情報記憶手段により記憶されている質問情報を照合話者に供給する質問情報供給手段と、当該質問情報供給手段により供給された質問情報に対する回答情報として照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果と、前記認証情報記憶手段により記憶されている認証情報と、に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有する」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声が集めることができる。これにより、照合が成功する可能性を高めた音声照合装置を提供することができる。

【0102】

また、この音声認識手段は、登録、記憶された登録話者の声紋情報と、照合話者の声紋情報と、を比較することにより、照合を行うこととなる。このため、より多くの音声を入力できれば、より一層、照合が成功する可能性を高めた音声照合装置を提供することができる。

【0103】

更には、前記質問情報記録手段においては、照合話者に対して供給される質問を複数記憶することが可能であり、当該質問情報記憶手段により記憶されている複数の質問情報から、照合話者に対して供給される一の質問情報を決定する質問決定手段を備えてもよく、この場合には、各種の質問を照合話者に対して供給することが可能であり、それらの各種の質問に基づく、種々の回答情報を得ることができる。このため、より多くの音声が入力され、より一層、照合が成功する可能性を高めた音声照合装置を提供することができる。

【0104】

更にまた、この音声認識手段は、登録、記憶された登録された音韻と、照合話者の音韻と、を比較するだけでなく、登録、記憶された登録話者の声紋情報と、照合話者の声紋情報と、を比較することにより、照合を行うこととなる。

【0105】

このため、照合の際に、その登録話者とは異なる照合話者が回答情報を入力した場合であっても、照合することがないため、行われ得る「なりすまし」を防止することが可能であり、セキュリティが高まる。

【0106】

また、本実施形態においては、必ずノイズキャンセル処理を行うようにしたが、本発明はこれに限らず、照合結果が照合不成立となった場合においてノイズキャンセル処理を行うようにしてもよい。

【0107】

この場合の動作処理を図 8 から図 10 を用いて説明する。尚、図 8 から図 10 においては、図 5 から図 7 に示すステップ番号と、同じような処理を行うものには同じ符号を付した。

【0108】

最初に、図 8 に示す如く、照合開始信号を受信したか否かを判断する（ステップ S 1 1）。この処理において、メイン CPU 2 1 は、入力部 2 4 からの入力信号に基づいて、ID 番号が入力された場合には、照合開始信号を受信したと判断することとなる。この処理において、メイン CPU 2 1 は、照合開始信号を受信したと判別した場合には、ステップ S 1 6 に処理を移し、照合開始信号を受信したとは判別しなかった場合には、本サブルーチンを終了する。

10

【0109】

次いで、ステップ S 1 1 の判断により照合開始信号を受信と判別した場合には、質問情報選択処理を実行する（ステップ S 1 6）。この処理において、メイン CPU 2 1 は、メインメモリ 2 2 に位置付けられた照合不成立回数データをリセットする、即ち「00」と記録する。この照合不成立回数データとは、照合結果が不成立であった回数を記憶するためのものである。この処理が終了した場合には、ステップ S 1 2 に処理を移す。

【0110】

次いで、質問情報選択処理を実行する（ステップ S 1 2）。この処理において、メイン CPU 2 1 は、ランダムで質問 ID をメインメモリ 2 2 に記録する。そして、メイン CPU 2 1 は、メインメモリ 2 2 に記録した質問 ID と、質問テーブルと、に基づいて、質問情報

20

【0111】

次いで、質問情報供給処理を実行する（ステップ S 1 3）。この処理において、メイン CPU 2 1 は、ステップ S 1 2 の処理によりメインメモリ 2 2 に記録された質問 ID と、質問テーブルと、に基づいて、質問情報を照合話者に供給することとなる。つまり、メイン CPU 2 1 は、その質問情報を示す画像を表示部 2 3 に表示させたり、質問情報を示す音声を音声発生部（図示せず）から発しさせたりする。この処理が終了した場合には、ステップ S 1 9 に処理を移す。

【0112】

次いで、照合不成立であったか否かを判断する（ステップ S 1 9）。この処理において、メイン CPU 2 1 は、メインメモリ 2 2 に位置付けられた照合不成立回数データを読み出し、そのデータが「00」であるか否かを判断することとなる。この処理により、メイン CPU 2 1 は、照合不成立であったと判別した場合には、ステップ S 2 1 に処理を移し、照合不成立であったとは判別しなかった場合には、ステップ S 2 0 に処理を移す。

30

【0113】

次いで、ステップ S 1 9 の判断により照合不成立であったと判別した場合には、複数のマイクにより音声入力処理を実行する（ステップ S 2 1）。この処理において、主マイク 1 0 2 及び参照マイク 2 0 2 は、各々、音声を收音し、その収集した音声を示す、主音声信号、参照音声信号を、B P F 1 0 4、2 0 4 に供給する。この処理が終了した場合には、ステップ S 2 2 に処理を移す。

40

【0114】

一方、ステップ S 1 9 の判断により照合不成立であったとは判別しなかった場合には、主マイクにより音声入力処理を実行する（ステップ S 2 0）。この処理において、主マイク 1 0 2 は、各々、音声を收音し、その収集した音声を示す主音声信号を B P F 1 0 4 に供給する。この処理が終了した場合には、ステップ S 2 2 に処理を移す。

【0115】

次いで、音声信号フィルタ処理を実行する（ステップ S 2 2）。この処理において、B P F 1 0 4、2 0 4 は、各々、主マイク 1 0 2、参照マイク 2 0 2 から受け取った、主音声信号、参照音声信号に対して、所定の周波数帯に含まれる信号のみを通過させ、その周波数帯域とは異なる周波数帯域の信号を予め減衰させ、フィルタリングされた主音声信号、

50

参照音声信号を低周波増幅器 106、206 に供給する。この処理が終了した場合には、ステップ S 23 に処理を移す。

【0116】

次いで、低周波増幅処理を実行する（ステップ S 23）。この処理において、低周波増幅器 106、206 は、BPF 104、204 から供給された主音声信号、参照音声信号に対して、所定の低周波成分を増幅させる。これにより、この処理をすることにより、音声認識を的確に行うことができる信号に変換することができるのである。そして、低周波増幅器 106、206 は、増幅処理が行われた主音声信号、参照音声信号をサンプル・ホールド回路 108、208 に供給する。この処理が終了した場合には、ステップ S 24 に処理を移す。

10

【0117】

次いで、サンプル・ホールド処理を実行する（ステップ S 24）。この処理において、サンプル・ホールド回路 108、208 は、低周波増幅器 106、206 から供給された主音声信号、参照音声信号を、音声入力用 CPU 52 から供給される制御信号に基づいたタイミングで A/D 変換器 110、210 に供給することとなり、主音声信号、参照音声信号の同期をとることとなるのである。この処理が終了した場合には、ステップ S 25 に処理を移す。

【0118】

次いで、A/D 変換処理を実行する（ステップ S 25）。この処理において、A/D 変換器 110、210 は、サンプル・ホールド回路 108、208 から供給された主音声信号、参照音声信号を受け取り、デジタル変換する。この処理が終了した場合には、ステップ S 26 に処理を移す。

20

【0119】

次いで、音声データ記録処理を実行する（ステップ S 26）。この処理において、A/D 変換器 110、210 は、デジタル変換された主音声データ、参照音声データを音声データメモリに記録する。この処理が終了した場合には、ステップ S 36 に処理を移す。

【0120】

次いで、図 9 に示す如く、照合不成立であったか否かを判断する（ステップ S 36）。この処理において、メイン CPU 21 は、メインメモリ 22 に位置付けられた照合不成立回数データを読み出し、そのデータが「00」であるか否かを判断することとなる。この処理により、メイン CPU 21 は、照合不成立であったと判別した場合には、ステップ S 27 に処理を移し、照合不成立であったとは判別しなかった場合には、ステップ S 30 に処理を移す。

30

【0121】

次いで、ステップ S 36 の判断により照合が不成立であったと判別した場合には、相互相関判定処理を実行する（ステップ S 27）。この処理において、相互相関判定器 302 は、主マイク 102 から供給された主音声データと、参照マイク 202 を含む複数の参照マイクから供給された複数の参照データと、を音声データメモリ 152 から読み出す。そして、相互相関判定器 302 は、主音声データと、複数の参照データのいずれかと、を比較し、その比較の結果、複数の参照データから、主音声データと一番類似している一を抽出する。そして、相互相関判定器 302 は、主音声データと、抽出された参照音声データ（以降、抽出参照音声データ）と、をノイズキャンセルユニット 304 に供給する。この処理が終了した場合には、ステップ S 28 に処理を移す。

40

【0122】

次いで、ノイズキャンセル処理を実行する（ステップ S 28）。この処理において、ノイズキャンセルユニット 304 は、相互相関判定器 302 から供給された主音声データと、抽出参照音声データと、を受け取り、それらを比較する。そして、ノイズキャンセルユニット 304 は、主音声データに対して、主音声データと、抽出参照音声データと、の差を少なくするように補正する。この処理が終了した場合には、ステップ S 29 に処理を移す。

50

【0123】

次いで、音声データ記録処理を実行する（ステップS29）。この処理において、ノイズキャンセルユニット304は、ステップS28の処理によりノイズキャンセルされた主音声データを、ノイズキャンセル音声データメモリ306に記録する。この処理が終了した場合には、ステップS30に処理を移す。

【0124】

次いで、音声データ高速フーリエ変換処理を実行する（ステップS30）。この処理において、FFT308は、照合不成立でない初期状態の場合には、音声データメモリ152に記録された主音声データを読み出し、その主音声データに対して、高速フーリエ変換を行う。一方、FFT308は、照合不成立でない初期状態の場合には、ノイズキャンセル音声データメモリ306に記録された主音声データを読み出し、その主音声データに対して、高速フーリエ変換を行う。この処理が終了した場合には、ステップS31に処理を移す。

【0125】

次いで、音声スペクトルデータ記録処理を実行する（ステップS31）。この処理において、FFT308は、ステップS30の処理により変換された音声スペクトルデータを、音声スペクトルデータメモリ310に記録する。この処理が終了した場合には、ステップS32に処理を移す。

【0126】

次いで、所定時間経過したか否かを判断する（ステップS32）。この処理において、メインCPU21は、メインCPU21に内蔵されているタイマにより、質問情報を提供してから所定時間経過したか否かを判断することとなる。この処理において、メインCPU21は、所定時間経過したと判別した場合には、ステップS41に処理を移し、所定時間経過したとは判別しなかった場合には、ステップS21に処理を移す。

【0127】

次いで、ステップS32の判断により所定時間経過したと判別した場合には、音声スペクトルデータ抽出処理を実行する（ステップS41）。この処理において、メインCPU21は、音声入力用CPU52に対して、音声スペクトルデータを読み出すように命令する。この命令を受けた音声入力用CPU52は、音声スペクトルデータメモリ310に記録されている音声スペクトルデータを読み出し、メインCPU21に供給する。そして、メインCPU21は、その音声スペクトルデータから声紋情報を抽出する。この処理が終了した場合には、ステップS42に処理を移す。

【0128】

次いで、回答情報生成処理を実行する（ステップS42）。この処理において、メインCPU21は、ステップS41の処理により受け取った音声スペクトルデータから音韻分析を行い、回答情報を生成することとなる。この処理が終了した場合には、ステップS43に処理を移す。

【0129】

次いで、データベース参照処理を実行する（ステップS43）。この処理において、メインCPU21は、記憶部27に記録されている登録話者データベースを参照する。尚、この処理においては、メインCPU21は、ステップS11の処理により、ID番号が入力された場合には、そのID番号と、ステップS12の処理により決定された質問番号と、に基づいて、登録話者データベースの所定のデータを読み出すこととなる。この処理が終了した場合には、ステップS44に処理を移す。

【0130】

次いで、音声照合処理を実行する（ステップS44）。この処理において、メインCPU21は、ステップS41の処理により抽出された声紋情報と、ステップS43の処理により行われた参照結果と、を比較し、類似しているか否かを決定する。また、メインCPU21は、ステップS42の処理により生成された音韻情報と、ステップS43の処理により行われた参照結果と、を比較し、類似しているか否かを決定する。このように、メイン

C P U 2 1 は、声紋情報、音韻情報、両者の照合が正常に行われることにより、照合結果が成立したことになる。

【0131】

尚、本実施形態においては、声紋情報、音韻情報、両者の照合が正常に行われることにより、照合結果が成立したことになる好適な一例であったが、本発明はこれに限らず、声紋情報、音韻情報、いずれかの照合が正常に行われることにより、照合結果が成立したとしても何ら問題なく、声紋情報の照合が正常に行われることにより、照合結果が成立することが好ましい。この処理が終了した場合には、ステップ S 5 1 に処理を移す。

【0132】

次いで、図 10 に示す如く、照合が成立したか否かを判断する（ステップ S 5 1）。この処理において、メイン C P U 2 1 は、ステップ S 4 4 の処理により照合された結果に基づいて判断することとなる。この処理において、メイン C P U 2 1 は、照合が成立したと判別した場合には、ステップ S 6 1 に処理を移し、照合が成立したとは判別しなかった場合には、ステップ S 6 2 に処理を移す。 10

【0133】

次いで、ステップ S 5 1 の判断により照合が成立したと判別した場合には、照合成立記録処理を実行する（ステップ S 6 2）。メイン C P U 2 1 は、メインメモリ 2 2 に位置付けられた照合不成立回数データを読み出し、そのデータを「00」と記録する。この処理が終了した場合には、ステップ S 5 2 に処理を移す。

【0134】

次いで、照合成立表示処理を実行する（ステップ S 5 2）。この処理において、メイン C P U 2 1 は、照合が成立した旨の画像を表示部 2 3 に表示させる。この処理が終了した場合には、本サブルーチンを終了する。 20

【0135】

一方、ステップ S 5 1 の判断により照合が成立したとは判別しなかった場合には、照合不成立記録処理を実行する（ステップ S 6 2）。この処理において、メイン C P U 2 1 は、メインメモリ 2 2 に位置付けられた照合不成立回数データを読み出し、そのデータをインクリメントし、照合不成立回数データとして記録する。この処理が終了した場合には、ステップ S 6 3 に処理を移す。

【0136】

次いで、照合不成立の回数が所定数以下であるか否かを判断する（ステップ S 6 3）。この処理において、メイン C P U 2 1 は、メインメモリ 2 2 に位置付けられた照合不成立回数データを読み出し、そのデータが所定数以下であるか否かを判断することとなる。この処理により、メイン C P U 2 1 は、照合不成立の回数が所定数以下であると判別した場合には、ステップ S 1 3 に処理を移し、照合不成立の回数が所定数以下でないと判別しなかった場合には、ステップ S 5 3 に処理を移す。 30

【0137】

次いで、ステップ S 6 3 の判断により照合不成立の回数が所定数以下でないと判別しなかった場合には、照合不成立表示処理を実行する（ステップ S 5 3）。この処理において、メイン C P U 2 1 は、照合が成立しなかった旨の画像を表示部 2 3 に表示させる。この処理が終了した場合には、本サブルーチンを終了する。 40

【0138】

これにより、「照合話者の音声信号を入力するための複数の音声入力手段と、入力された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有するものであり、かつ、前記音声認識手段が前記複数の入力手段の一より入力された音声信号に基づいて声紋分析及び／又は音韻分析を行った結果が正常に照合されなかった場合には、前記当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少 50

なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、を備え、前記音声認識手段は、前記音声補正手段により補正された近似の音声信号の少なくとも二つに含まれる一に基づいて声紋分析及び／又は音韻分析を行う音声認識を行うものであり、かつ、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを再照合する機能を有する」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声が集められる。これにより、照合が成功する可能性を高めた音声照合装置を提供することができる。

10

【0139】

また、照合が失敗した場合において、このノイズを除去する補正が行われるため、余計な処理を行うことなく、必要に応じて、自動的に良質な音声を収集することとなり、音声照合装置の処理効率が高まるケースが増加する。

【0140】

また、本実施形態においては、音声照合装置10のみの構成としたが、本発明はこれに限らず、サーバを介して複数の音声照合装置を接続可能としてもよい。

【0141】

例えば、図11に示す如く、サーバ81には複数の音声照合装置10が接続されており、サーバ81は、それらの音声照合装置10から供給された音声信号に関する情報に基づいて、照合を行い、その照合の結果を音声照合装置10に供給するようにしてもよい。この場合には、通信I/F25がサーバ81との送受信を行うこととなり、更には、登録話者に関する登録話者データベースは、サーバ81における記憶部（図示せず）に記録されておればよく、照合処理をサーバ81で行うことが可能となる。

20

【0142】

このように構成された場合には、「照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、音声認識手段により行われた声紋分析及び／又は音韻分析の結果を前記サーバに対して供給する音声認識結果供給手段と、前記サーバが前記音声認識結果供給手段から供給された声紋分析及び／又は音韻分析の結果に基づいて登録話者の音声信号と照合話者の音声信号とを照合した結果を受信する照合結果受信手段と、を備えた」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声が集められる。これにより、照合が成功する可能性を高めた音声照合装置を提供することができる。

30

40

【0143】

また、このようなサーバに対して、複数の音声照合装置を接続することが可能となるので、サーバに登録話者として登録しておくことにより、そのサーバに接続された音声照合装置に対して、照合を行うことが可能となり、簡便である。

【0144】**【発明の効果】**

本発明によれば、「照合話者の音声信号を入力するための複数の音声入力手段と、当該複数の音声入力手段の各々において入力された複数の音声信号を、それぞれ相互比較する音声比較手段と、当該音声比較手段により比較された結果に基づいて、複数の音声信号から近似の音声信号の少なくとも二つを抽出する音声抽出手段と、当該音声抽出手段により抽

50

出された近似の音声信号の少なくとも二つに基づいて、当該近似の音声信号の少なくとも二つに含まれる一を補正する音声補正手段と、当該音声補正手段により補正された音声信号に基づいて声紋分析及び／又は音韻分析を行う音声認識手段と、を備え、前記音声認識手段により行われた声紋分析及び／又は音韻分析の結果に基づいて、登録話者の音声信号と照合話者の音声信号とを照合する機能を有する」ので、複数の音声信号により、それらの音声信号に含まれるものとの比較により、その差異を埋めるようにノイズを除去する補正が行われるため、従来入力した音声信号よりも、周辺環境による影響も少なく、良質な音声が集めることができる。これにより、照合が成功する可能性を高めた音声照合装置を提供することができる。

【図面の簡単な説明】

10

【図 1】本発明の一実施形態における音声照合装置の電気的な構成を示すブロック図である。

【図 2】本発明の一実施形態における音声照合装置のデータベースを示す説明図である。

【図 3】本発明の一実施形態における音声照合装置のテーブルを示す説明図である。

【図 4】本発明の一実施形態における音声照合装置の音声入力装置の電気的な構成を示すブロック図である。

【図 5】本発明の一実施形態における音声照合装置のフローチャートを示す図である。

【図 6】本発明の一実施形態における音声照合装置のフローチャートを示す図である。

【図 7】本発明の一実施形態における音声照合装置のフローチャートを示す図である。

【図 8】本発明の一実施形態における音声照合装置のフローチャートを示す図である。

20

【図 9】本発明の一実施形態における音声照合装置のフローチャートを示す図である。

【図 10】本発明の一実施形態における音声照合装置のフローチャートを示す図である。

【図 11】本発明の一実施形態におけるサーバを含んだ音声照合装置の電気的な構成を示すブロック図である。

【符号の説明】

10 音声照合装置

21 メインCPU

22 メインメモリ

23 表示部

24 入力部

25 通信 I/F

26 音声入力装置

27 記憶部

52 音声入力用CPU

54 音声入力用ROM

56 音声入力用RAM

58 バッファ

60 入出力バス

81 サーバ

102 主マイク

104 バンドパスフィルタ (BPF)

106 低周波増幅器

108 サンプル・ホールド回路

110 変換器

152 音声データメモリ

202 参照マイク

206 低周波増幅器

208 サンプル・ホールド回路

210 変換器

302 相互相関判定器

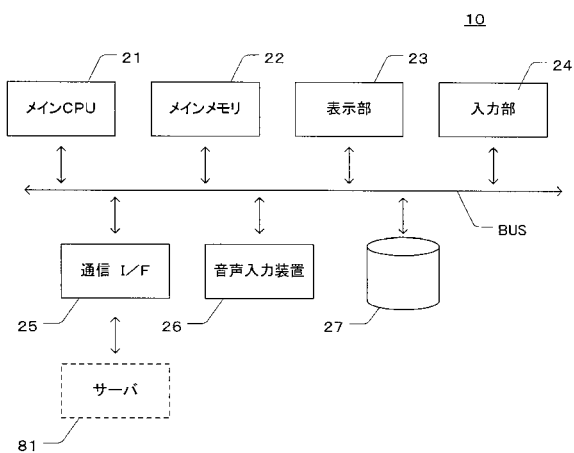
30

40

50

- 304 ノイズキャンセルユニット
 306 ノイズキャンセル音声データメモリ
 308 高速フーリエ変換処理部 (FFT)
 310 音声スペクトルデータメモリ

【図 1】



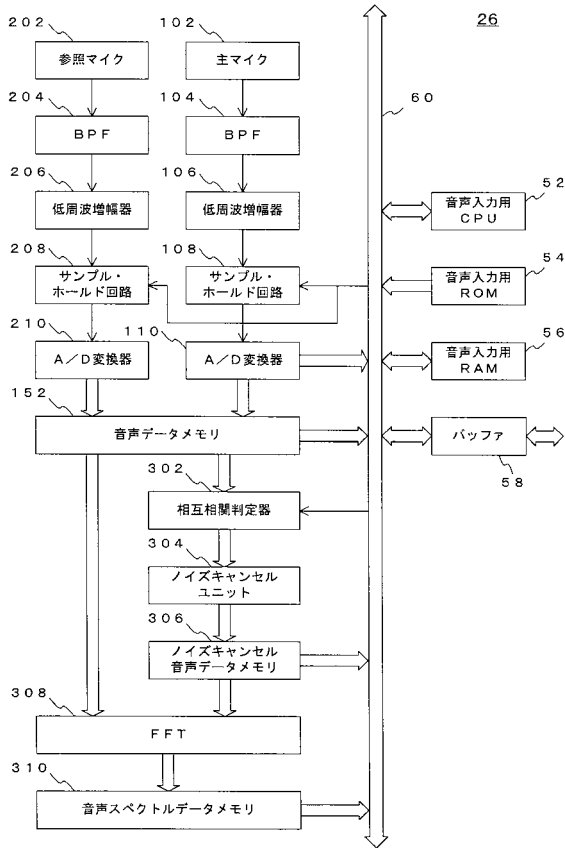
【図 3】

質問ID	質問情報
0001	出身地はどこですか？
0002	お父さんの名前は？
0003	お母さんの名前は？
0004	旅行に行きたいところは？
⋮	⋮

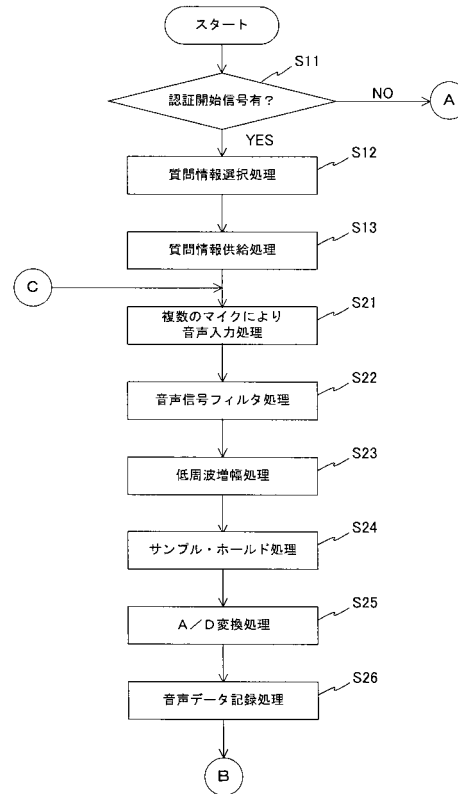
【図 2】

ID	質問ID	回答情報	声紋情報ID	⋮
0001	0001	東京都です	00010001	⋮
	0002	幸則です	00010002	⋮
	0003	澄子です	00010003	⋮
	0004	デカン高原です	00010004	⋮
	⋮	⋮	⋮	⋮
0002	0001	岐阜県	00020001	⋮
⋮	⋮	⋮	⋮	⋮

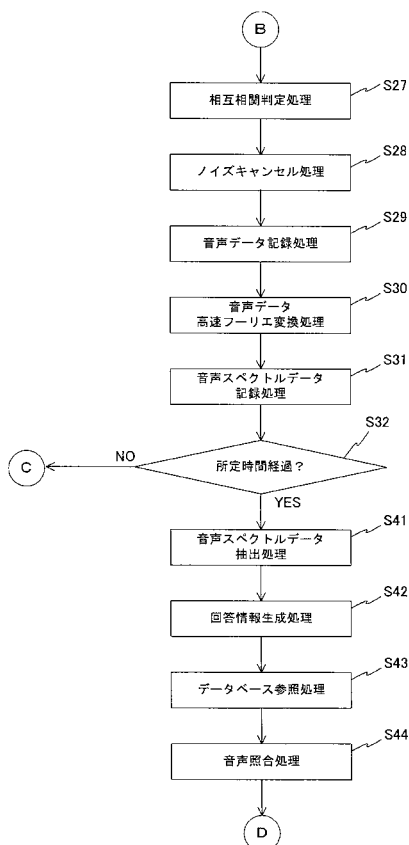
【図 4】



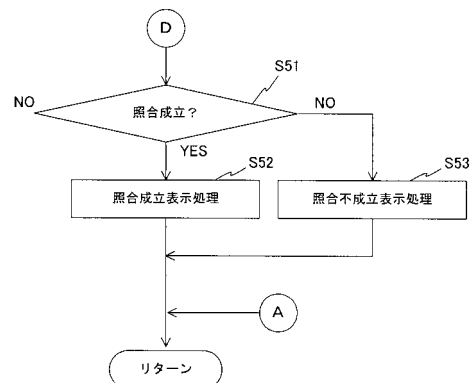
【図 5】



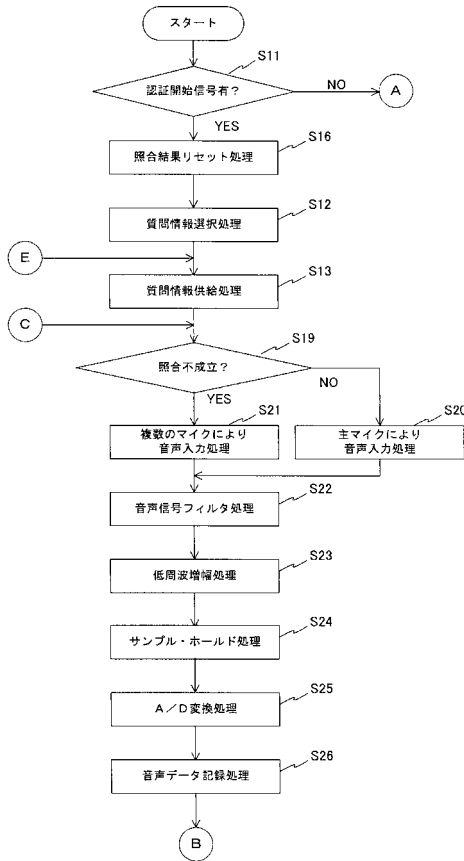
【図 6】



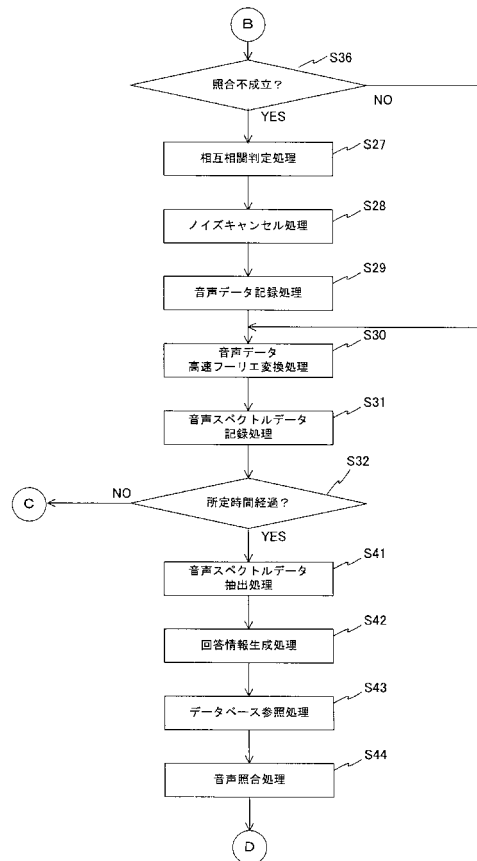
【図 7】



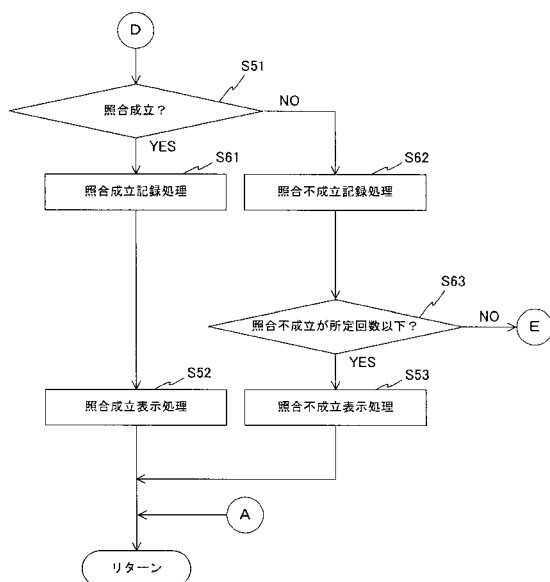
【図 8】



【図 9】



【図 10】



【図 11】

