

發明專利說明書

(本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※申請案號： 96131267

※申請日期： 96.8.23

※IPC 分類：

一、發明名稱：(中文/英文)

聲音驗證裝置

VOICE AUTHENTICATION APPARATUS

G10L 15/02 (2006.01)

G10L 19/06 (2006.01)

二、申請人：(共 1 人)

姓名或名稱：(中文/英文)

日商山葉股份有限公司

YAMAHA CORPORATION

代表人：(中文/英文)

梅村 充

UMEMURA, MITSURU

住居所或營業所地址：(中文/英文)

日本國靜岡縣濱松市中區中澤町10番1號

10-1, NAKAZAWA-CHO, NAKA-KU, HAMAMATSU-SHI,

SHIZUOKA-KEN, JAPAN

國籍：(中文/英文)

日本 JAPAN

三、發明人：(共 2 人)

姓 名：(中文/英文)

1. 吉岡 靖雄
YOSHIOKA, YASUO
2. 川原 毅彦
KAWAHARA, TAKEHIKO

國 籍：(中文/英文)

1. 日本 JAPAN
2. 日本 JAPAN

四、聲明事項：

主張專利法第二十二條第二項 第一款或 第二款規定之事實，其事實發生日期為： 年 月 日。

申請前已向下列國家（地區）申請專利：

【格式請依：受理國家（地區）、申請日、申請案號 順序註記】

有主張專利法第二十七條第一項國際優先權：

1. 日本；2006年09月14日；特願2006-249160

2.

無主張專利法第二十七條第一項國際優先權：

1.

2.

主張專利法第二十九條第一項國內優先權：

【格式請依：申請日、申請案號 順序註記】

主張專利法第三十條生物材料：

須寄存生物材料者：

國內生物材料 【格式請依：寄存機構、日期、號碼 順序註記】

國外生物材料 【格式請依：寄存國家、機構、日期、號碼 順序註記】

不須寄存生物材料者：

所屬技術領域中具有通常知識者易於獲得時，不須寄存。

九、發明說明：

【發明所屬之技術領域】

本發明係關於一種個別聲音之驗證技術。

【先前技術】

已提出一種習知聲音驗證技術，該技術藉由將先前自己授權使用者獲取之聲音(下文中稱作"參考聲音"或"登錄聲音")的特徵量與自待驗證之人(下文中稱作"受驗者")獲取之聲音(下文中稱作"樣本聲音"或"驗證聲音")之特徵量之間的距離與一臨限值進行比較來判定受驗者之真確性。專利文獻日本專利申請案公告第2003-248661號描述一種組態，其中根據驗證之目的或根據驗證之所需精確度水準來改變臨限值。

圖8為用以評估聲音驗證之圖。圖8中之錯誤拒絕率(False Rejection Rate, FRR)意謂即使一受驗者為一授權使用者，該受驗者之真確性亦將在驗證期間被拒絕的可能性(使用者拒絕率)，且錯誤接受率(False Acceptance Rate, FAR)意謂即使受驗者並非授權使用者，該受驗者之真確性亦將在驗證期間被接受之可能性(陌生人接受率)。如自圖8可見，若將驗證中所使用之臨限值設定為圖8中之值"a"，則拒絕未授權受驗者的可能性維持在一足夠高之水準，而拒絕授權使用者之可能性得以充分降低。

然而，驗證聲音與登錄聲音之間的距離根據驗證期間產生之環境聲響(下文中稱作"樣本雜音"或"驗證雜音")而改變，因為驗證雜音影響驗證聲音之特性。因此，圖8中之

FRR及FAR中之每一者的曲線根據驗證雜音之特性而平行於水平軸(距離)移動。舉例而言，若將由圖8中之實線所示之FAR改為虛線L1，則錯誤接受陌生人的真確性之可能性增加(亦即，驗證之精確度水準降低)，且若將圖8中之FRR改為虛線L2，則拒絕授權使用者的真確性之可能性增加(亦即，驗證之便利性降低)。亦即，習知聲音驗證之問題在於驗證之精確度與驗證之便利性之間的平衡根據驗證雜音之特性而被打破。

此問題甚至不可如在專利文獻日本專利申請案公告第2003-248661號中藉由根據驗證之目的或所需精確度水準而改變臨限值來解決。當以由行動電話代表之攜帶型電子設備執行驗證時此問題更為嚴重，因為驗證雜音之特性根據電子設備的使用環境而以各種方式改變。

【發明內容】

已鑒於此等情形作出本發明，且本發明之一目標為解決無關於樣本雜音而維持驗證之精確度與驗證的便利性之間的平衡之問題。

為解決以上問題，根據本發明之一態樣之聲音驗證裝置包含特性分析構件，其分析在受驗者產生用於受驗者的驗證之樣本聲音或驗證聲音時在受驗者周圍產生之樣本雜音或驗證雜音的特性；第一設定構件(例如，圖1中之設定器51)，其根據由特性分析構件分析之驗證雜音之特性來設定第一校正值(例如，圖1中之校正值Aa)；校正構件，其基於由第一設定構件設定之第一校正值來校正指數值(例

如，圖 1 中之距離 D_0)，該指數值指示先前已登錄之參考聲音或登錄聲音的特徵量與自受驗者獲取之驗證聲音之特徵量之間的相似程度；及判定器，其藉由將由校正構件校正後之指數值與預定臨限值進行比較來判定受驗者之真確性。

根據此態樣，在根據驗證雜音之特性校正指數值之後將指示登錄聲音是否與驗證聲音類似之指數值與預定臨限值進行比較，藉此無關於驗證雜音而維持驗證的便利性與驗證之精確度之間的平衡。預定臨限值可為固定值或可變值。

在本發明之一較佳實施例中，第一設定構件根據驗證雜音及登錄聲音與在登錄聲音之登錄期間產生之登錄雜音之間的位準比率(例如，圖 4 中之位準比率 EN_SN)來判定第一校正值，使得第一校正值設定為預定臨限值與一變數(例如，圖 4 中之臨限值 TH)之間的差，該變數係與驗證雜音之特性及登錄聲音與登錄雜音之間的位準比率相關聯而判定的。

舉例而言，第一設定構件將第一校正值設定為預定臨限值與隨登錄聲音與登錄雜音之間的位準比率而沿對應於驗證雜音之特性之線或曲線變化的變數之間的差。

更特定言之，第一設定構件根據登錄聲音與登錄雜音之間的位準比率及作為登錄聲音與登錄雜音之間的位準比率與驗證聲音與驗證雜音之間的位準比率(例如，位準比率 V_SN)之間的位準比率差(例如，圖 4 中之位準比率差

DIF_SN1~DIF_SN3)之驗證雜音之特性來設定第一校正
值。

此實施例可無關於登錄雜音或驗證雜音之特性而藉由簡單處理來維持驗證之便利性與驗證的精確度之間的平衡。第一設定構件可基於界定第一校正值與登錄聲音與登錄雜音之間的位準比率及驗證雜音之特性(更特定言之，登錄聲音與登錄雜音之間的位準比率與驗證聲音與驗證雜音之間的位準比率之間的位準比率差)的關係之表來設定第一校正值，且亦可藉由使用表示此等值之間的關係之等式進行計算而設定第一校正值。該表或該等式之內容可(例如)根據來自使用者之指令而適當改變。

根據本發明之一實施例之聲音驗證裝置進一步包含第二設定構件(例如，圖1中之設定器52)，其根據驗證雜音與登錄雜音之間的差(例如，圖6中之相關值NOISE_DIF)來設定一第二校正值(例如，圖1中之校正值Ab)，其中該校正構件根據第一及第二校正值來校正指數值。

根據此實施例，即使當在實際驗證期間登錄雜音(亦即，參考雜音)與驗證雜音(亦即，樣本雜音)之間的實際相關性不同於在判定第一校正值與登錄雜音及登錄聲音之關係時所採用之登錄雜音與驗證雜音之間的原始相關性，亦可基於根據實際驗證雜音與登錄雜音之間的差之第二校正值來校正指數值，藉此有效地維持驗證之便利性與驗證之精確度之間的平衡。第二設定構件可基於界定第二校正值與驗證雜音與登錄雜音之間的差之關係之表來設定第二校

正值，且亦可藉由使用表示該關係的等式來進行計算而設定第二校正值。該表或該等式之內容可(例如)根據來自使用者之指令而適當改變。

根據本發明之一實施例之聲音驗證裝置進一步包含第三設定構件(例如，圖1中之設定器53)，該第三設定構件根據驗證聲音(亦即，樣本聲音)或登錄聲音(亦即，參考聲音)的持續期間(例如，圖7中之語音長度EN_SPEECH_LEN或V_SPEECH_LEN)來設定第三校正值(例如，圖1中之校正值 A_c)，其中該校正構件根據第一及第三校正值來校正指數值。

根據此實施例，即使當登錄聲音或驗證聲音之語音長度(或持續期間)不同於在判定第一校正值與登錄雜音及登錄聲音之關係時所採用的語音長度時，亦可基於根據驗證聲音或登錄聲音之第三校正值來校正指數值，藉此有效地維持驗證之便利性與驗證之精確度之間的平衡。第三設定構件可基於界定第三校正值與驗證聲音或登錄聲音之持續期間之間的關係之表來設定第三校正值，且亦可藉由使用表示該關係之等式進行計算來設定第三校正值。該表或該等式之內容可(例如)根據來自使用者之指令而適當改變。

用作用於判定用以校正指數值之校正值之基礎的參數不限於驗證聲音或登錄聲音之持續期間及驗證雜音與登錄雜音之間的差。舉例而言，聲音驗證裝置可進一步包含根據包括於登錄聲音或驗證聲音中之有聲與無聲聲響之持續期間之間的比率來設定校正值之設定構件，或根據登錄聲音

或驗證聲音之音節之數目來設定校正值的設定構件。根據如上所述之表或等式來判定校正值之組態或表或等式的內容為可變之組態用於此等實施例中之任一者中。

本發明亦指定為一種控制根據以上實施例中之每一者之聲音驗證裝置的聲音驗證方法。根據本發明之一實施例之聲音驗證方法包含：分析在一受驗者產生一用於該受驗者的驗證之驗證聲音時在該受驗者周圍產生之驗證雜音之特性；根據驗證雜音之所分析特性設定校正值；基於所設定校正值校正指數值，該指數值指示先前已登錄之登錄聲音之特徵量與自受驗者獲取的驗證聲音之特徵量之間的相似程度；及藉由將經校正之指數值與預定臨限值進行比較來判定受驗者之真確性。此方法達成與根據本發明之聲音驗證裝置之操作及優勢相同的操作及優勢。

根據以上實施例中之每一者之聲音驗證裝置藉由專用於每一處理的諸如數位信號處理器(DSP)之硬體(電子電路)來實施，且亦經由程式與諸如中央處理單元(CPU)之通用處理單元之間的協作來實施。根據本發明之程式使得電腦執行以下處理：特性分析處理，以分析在受驗者產生用於受驗者之驗證的樣本聲音時在受驗者周圍產生之驗證雜音之特性；設定處理，以根據在特性分析處理中所分析之驗證雜音之特性來設定校正值；校正處理，以基於在設定處理中設定之校正值來校正指數值，該指數值指示先前已登錄之登錄聲音的特徵量與自受驗者獲取之驗證聲音之特徵量之間的相似程度；及判定處理，以藉由將經校正之指數值

與預定臨限值進行比較來判定受驗者之真確性。此程式亦達成與根據以上實施例中之每一者之聲音驗證裝置的操作及優勢相同之操作及優勢。本發明之程式係以儲存於諸如CD-ROM之機器可讀取媒體或攜帶型記錄媒體中之形式而提供至使用者且接著安裝於電腦上，且亦以在網路上之分布式方式自伺服器裝置提供且接著安裝於電腦上。

【實施方式】

<A：聲音驗證裝置之組態>

圖1為說明根據本發明之實施例之聲音驗證裝置的組態之方塊圖。聲音驗證裝置100為基於當受驗者說出特定詞語時產生之聲音來判定其真確性(亦即，判定受驗者是否為先前已登錄之正式使用者)之設備。聲音驗證裝置100安裝於多種電子設備中，諸如行動電話或資訊處理設備。圖1中所示之特性分析器20、驗證器40或校正值控制器50可各實施為(例如)由諸如CPU之處理單元執行之程式，且亦可由諸如DSP之硬體電路實施。儲存單元35可含有含程式之機器可讀媒體。

聲音驗證裝置100之操作分為初始登錄及驗證。初始登錄為在驗證之前登錄由授權使用者說出之聲音(登錄聲音)的操作。驗證為藉由將登錄聲音與由受驗者說出之聲音(驗證聲音)進行對比來判定受驗者之真確性的操作。操作單元10包括由使用者操縱之複數個操作器。使用者可藉由適當操縱操作單元10而命令聲音驗證裝置100開始初始登錄或驗證。

圖 1 中之輸入單元 15 及特性分析器 20 用以當執行驗證時偵測驗證聲音或聲音驗證裝置 100 之環境雜音(驗證雜音)的特性，且用以當如圖 1 中之虛線 R 所示執行初始登錄時偵測登錄聲音或聲音驗證裝置 100 之環境雜音(登錄雜音)的特性。

輸入單元 15 為根據環境聲響(聲音及雜音)產生聲響信號 S 之聲響收集器。如圖 2 中所說明，聲響信號 S 分為未出聲區段 P1 及出聲區段 P2。出聲區段 P2 為授權使用者在初始登錄期間說出登錄聲音之區段，或受驗者在驗證期間說出驗證聲音之區段。另一方面，未出聲區段 P1 為未說出登錄或驗證聲音之區段。因為在安裝有聲音驗證裝置 100 之環境中可產生多種雜音，所以未出聲區段 P1 並非完全安靜(亦即，聲響信號 S 之振幅在未出聲區段 P1 中並非零)，且輸入單元 15 在未出聲區段 P1 中收集登錄雜音或驗證雜音。

將由輸入單元 15 產生之聲響信號 S 提供至圖 1 之特性分析器 20。特性分析器 20 為用於分析由輸入單元 15 收集之聲響之構件。特性分析器 20 包括區段偵測器 22、開關 23、雜音分析器 25、聲音分析器 26 及特徵分析器 28。區段偵測器 22 將未出聲區段 P1 與出聲區段 P2 加以區分。舉例而言，區段偵測器 22 將聲響信號 S 之振幅不連續增加或減少時之時間偵測為未出聲區段 P1 與出聲區段 P2 之間的邊界。可使用多種已知技術來將未出聲區段 P1 與出聲區段 P2 加以區分。

開關 23 為用於切換由輸入單元 15 產生之聲響信號 S 之供應目的地的構件。此允許將聲響信號 S 由區段偵測器 22 識

別為未出聲區段P1之部分提供至雜音分析器25，且允許將聲響信號S由區段偵測器22識別為出聲區段P2之部分提供至聲音分析器26及特徵分析器28。

雜音分析器25為用於基於未出聲區段P1中之聲響信號S來分析登錄雜音或驗證雜音之特性的構件。此實施例中之雜音分析器25週期性地分析未出聲區段P1中之聲響信號S的特性。當操縱操作單元10以命令聲音驗證裝置100開始初始登錄或驗證時，雜音分析器25判定區段P(稱作偵測區段)(該區段在所命令時間之前一預定時間開始且在所命令時間結束)之經分析特性為如圖2中所示之登錄雜音或驗證雜音的特性。在以下描述中，與登錄雜音或登錄聲音相關聯之項目由包括"登錄(enroll)"之縮寫"EN"的符號表示，且與驗證雜音或驗證聲音相關聯之項目由包括"驗證(verify)"之縮寫"V"的符號表示。

如圖1中所示，此實施例中之雜音分析器25在執行初始登錄時計算登錄雜音之頻率特性EN_NOISE_FC及雜音位準EN_NOISE_LEVEL，且在執行驗證時計算驗證雜音之頻率特性V_NOISE_FC及雜音位準V_NOISE_LEVEL。雜音位準EN_NOISE_LEVEL及V_NOISE_LEVEL中之每一者為偵測區段P中之聲響信號S在未出聲區段P1中的分量中包括於特定頻帶中之分量之量值(聲壓)的平均值。頻率特性EN_NOISE_FC或V_NOISE_FC為表示將偵測區段P中之聲響信號S所分為的複數個頻帶之分量之各別量值的資訊。因此，雜音分析器25包括(例如)具有不同通頻帶之複數個

帶通濾波器(濾波器組)。頻率特性 EN_NOISE_FC 及 V_NOISE_FC 之資訊僅需要反映登錄雜音或驗證雜音之特性。舉例而言，雜音分析器 25 可為用於藉由諸如快速傅立葉變換(FFT)之頻率分析來計算頻譜以作為頻率特性 EN_NOISE_FC 及 V_NOISE_FC 之構件。

聲音分析器 26 基於出聲區段 P2 中之聲響信號 S 來分析登錄聲音或驗證聲音之特性。此實施例中之聲音分析器 26 在執行初始登錄時計算登錄聲音之語音位準 EN_SPEECH_LEVEL 及語音長度(或持續期間)EN_SPEECH_LEN，且在執行驗證時計算驗證聲音的語音位準 V_SPEECH_LEVEL 及語音長度 V_SPEECH_LEN。語音位準 EN_SPEECH_LEVEL 及 V_SPEECH_LEVEL 中之每一者為包括於出聲區段 P2 中的聲響信號 S 中之特定頻帶中的分量之量值的平均值。語音長度 EN_SPEECH_LEN 及 V_SPEECH_LEN 中之每一者表示出聲區段 P2 之持續期間(亦即，語音持續之時段)。自聲響信號 S 之振幅快速增加時之時間(其對應於出聲區段 P2 之開始)至聲響信號 S 的振幅快速減少時之時間(其對應於出聲區段 P2 之結束)之時段被偵測為語音長度 EN_SPEECH_LEN 及 V_SPEECH_LEN。

特徵分析器 28 為用於分析登錄聲音或驗證聲音之特徵之構件。此實施例中之特徵分析器 28 在執行初始登錄時計算登錄聲音之特徵量 EN_DATA，且在執行驗證時計算驗證聲音的特徵量 V_DATA。特徵量 EN_DATA 及 V_DATA 中之每一者為自出聲區段 P2 中之聲響信號 S 計算的倒譜之向量之

時序(time-series)。因此，用於執行包括頻率分析(例如，FFT方法)之多種計算之構件較佳用作特徵分析器28。

儲存設備32為用於儲存驗證中所使用之多種資訊之構件。舉例而言，如圖1中所示，儲存設備32將特性分析器20指定用於登錄聲音及登錄雜音之多種資訊儲存為驗證辭典。更特定言之，由雜音分析器25指定之頻率特性EN_NOISE_FC及雜音位準EN_NOISE_LEVEL、由聲音分析器26指定之語音位準EN_SPEECH_LEVEL及語音長度EN_SPEECH_LEN，及由特徵分析器28指定的特徵量EN_DATA在執行驗證之前儲存於儲存設備32中。儲存設備32可為固定至聲音驗證裝置100之設備，且亦可為可拆卸地附接至聲音驗證裝置100之攜帶型設備(記憶體)。

驗證器40為用於藉由將驗證聲音與登錄聲音進行對比來判定受驗者之真確性之構件。驗證器40包括距離計算器42、校正器44及判定器46。距離計算器42計算儲存於儲存設備32中之特徵量EN_DATA與特徵分析器28針對驗證聲音產生之特徵量V_DATA之間的距離D0。距離計算器42使用諸如計算特徵量EN_DATA及V_DATA之向量序列之間的正規化距離之DP匹配之多種圖案匹配技術來計算距離D0。隨著距離D0減小，驗證聲音與登錄聲音更類似(亦即，受驗者更可能為一授權使用者)。

校正器44為用於基於校正值A(Aa、Ab及Ac)來校正由距離計算器42計算出之距離D0之構件。在此實施例中，校正器44藉由自距離D0減去由校正值控制器50設定之每一校正

值 A_a 、 A_b 及 A_c 來計算經校正的距離 D_1 。

判定器 46 藉由將由校正器 44 校正之距離 D_1 與臨限值 TH_0 進行比較來判定受驗者之真確性。更特定言之，若距離 D_1 低於臨限值 TH_0 (亦即，若登錄聲音與驗證聲音類似)，則判定器 46 接受受驗者之真確性，且若距離 D_1 高於臨限值 TH_0 (亦即，若登錄聲音不同於驗證聲音)，則判定器 46 拒絕受驗者之真確性。臨限值 TH_0 為預定的固定值。判定器 46 之判定結果經由輸出單元 60 輸出。舉例而言，輸出指示驗證結果之影像之顯示器或輸出指示驗證結果的聲響之音訊輸出設備較佳用作輸出單元 60。

校正值控制器 50 為用於設定在校正器 44 之校正中使用之校正值 A (A_a 、 A_b 及 A_c) 的構件。校正值控制器 50 包括設定器 51、52 及 53。設定器 51 基於特性分析器 20 在初始登錄及驗證期間所產生之資訊來設定校正值 A_a 。設定器 52 根據驗證雜音之特性與登錄雜音之特性之間的差來設定校正值 A_b 。設定器 53 根據初始登錄期間的語音長度 EN_SPEECH_LEN 來設定校正值 A_c 。稍後將描述校正值控制器 50 之操作之細節。圖 1 中之儲存單元 35 儲存校正值控制器 50 用以設定校正值 A 之表。儲存設備 32 中之特定儲存區域可用作儲存單元 35。

<B：聲音驗證裝置之操作>

現將描述聲音驗證裝置 100 之詳細操作，集中於校正值控制器 50 在驗證期間設定校正值 A_a 、 A_b 及 A_c 之過程。每次有必要進行驗證時，校正值控制器 50 執行圖 3 之過程。

驗證變得必要時之實例包括當包括聲音驗證裝置100之電子設備隨著施加電力而開始操作時，及當電子設備開始特定操作(例如，存取特定資訊之操作)時。當起始驗證時，受驗者在藉由操縱操作單元10而發出語音啟動指令之後向輸入單元15說出特定詞語。雜音分析器25自結束於操縱操作單元10時之偵測區段P中的聲響信號S指定頻率特性V_NOISE_FC及雜音位準V_NOISE_LEVEL，且自緊接偵測區段P之出聲區段P2中之聲響信號S指定語音位準V_SPEECH_LEVEL及語音長度V_SPEECH_LEN。

如圖3中所示，校正值控制器50計算登錄聲音與登錄雜音之間的位準比率EN_SN(步驟S10)。位準比率EN_SN為儲存設備32中儲存之語音位準EN_SPEECH_LEVEL與雜音位準EN_NOISE_LEVEL之間的比率。舉例而言，使用等式(1)來計算位準比率EN_SN。

$$EN_SN = \log(EN_SPEECH_LEVEL/EN_NOISE_LEVEL) \dots \dots (1)$$

接著，校正值控制器50計算驗證聲音與驗證雜音之間的位準比率V_SN(步驟S11)。位準比率V_SN為自雜音分析器25提供之雜音位準V_NOISE_LEVEL與自聲音分析器26提供之語音位準V_SPEECH_LEVEL之間的比率。與位準比率EN_SN類似，使用等式(2)來計算位準比率V_SN。

$$V_SN = \log(V_SPEECH_LEVEL/V_NOISE_LEVEL) \dots \dots (2)$$

接下來，在步驟S12，校正值控制器50計算在步驟S10計算出的初始登錄中之位準比率EN_SN與在步驟S11計算出之驗證中之位準比率V_SN之間的差DIF_SN(DIF_SN=

V_SN-EN_SN)。在下文中，差DIF_SN將稱作"位準比率差"。校正值控制器50計算表示登錄雜音之特性與驗證雜音之特性之間的相關性之相關值NOISE_DIF(步驟S13)。此處，舉例而言，相關性為登錄雜音與驗證雜音之各別波譜形狀是否類似。舉例而言，使用等式(3)來計算相關值NOISE_DIF。

$$NOISE_DIF = \frac{\sum_{i=M}^L (EN_MAG(i) - EN_MAG_AVE)(V_MAG(i) - V_MAG_AVE)}{\sqrt{\sum_{i=M}^L (EN_MAG(i) - EN_MAG_AVE)^2} \sqrt{\sum_{i=M}^L (V_MAG(i) - V_MAG_AVE)^2}} \dots \dots (3)$$

等式(3)中之值EN_MAG(i)為複數個頻帶中由變數i指定之頻帶中的登錄雜音之量值，且自值EN_MAG(i)減去之值EN_MAG_AVE為由變數i指定之頻帶中的登錄雜音之量值之平均值。類似地，等式(3)中之值V_MAG(i)為由變數i指定之頻帶中的驗證雜音之量值，且自值V_MAG(i)減去之值V_MAG_AVE為該頻帶中的驗證雜音之量值之平均值。因此，若登錄雜音與驗證雜音完全匹配，則校正值NOISE_DIF為1，且校正值NOISE_DIF隨著兩個雜音之特性之間的差增加而減少(-1≤NOISE_DIF≤1)。

接著，校正值控制器50之設定器51基於在步驟S10計算出之位準比率EN_SN及在步驟S12計算出之位準比率差DIF_SN來設定校正值Aa(步驟S14)。校正值Aa與位準比率EN_SN及位準比率差DIF_SN之關係以以下方式來判定。

首先，吾人考慮根據由距離計算器42計算出之距離D0大於還是小於臨限值TH來判定受験者之真確性之狀況。藉

由進行測試，此申請案之發明者已發現，用於將驗證之精確度維持在一高水準的臨限值TH滿足與位準比率EN_SN及位準比率差DIF_SN之特定關係。更特定言之，如圖4中所示，臨限值TH經設定使得對於位準比率EN_SN(其為登錄雜音與登錄聲音之間的比率)之值改變為一新值且接著對水平軸上之每一相應位準比率EN_SN繪製垂直軸上之臨限值TH的每一狀況，驗證之精確度滿足一特定條件(例如，使得FRR及FAR各低於一特定位準)。當對所繪製之點進行統計處理時，其傾向於根據位準比率差DIF_SN而沿一線分布。舉例而言，當將位準比率差DIF_SN設定為值DIF_SN1時，滿足所要條件之臨限值TH隨位準比率EN_SN而沿一具有諸如對應於值DIF_SN1之梯度及截距的屬性之線改變。另外，如圖4中所示，用於將驗證精確度維持在一所要水準之臨限值TH隨著位準比率EN_SN或位準比率差DIF_SN(DIF_SN1>DIF_SN2>DIF_SN3)增加而增加。

因此，若針對在步驟S10計算出之位準比率EN_SN來設定臨限值TH以使得臨限值TH及位準比率EN_SN滿足對應於一在步驟S12計算出之位準比率差DIF_SN的線之關係，則可將驗證維持在一所要精確度水準。舉例而言，在位準比率EN_SN計算為圖4中之值SNa之狀況下，若位準比率差DIF_SN為DIF_SN1，則將臨限值TH設定為THa，且若位準比率差DIF_SN為DIF_SN2，則將臨限值TH設定為THb。

因為距離大於還是小於臨限值係在驗證期間判定，所以用於將待與距離D0比較之臨限值TH改變一特定值之處理

等效於用於在臨限值TH固定為TH0的情形下將距離D0校正特定值Aa之處理。因此，此實施例之設定器51將預定臨限值TH0與臨限值TH之間的差設定為校正值Aa，臨限值TH與位準比率EN_SN滿足一對應於位準比率差DIF_SN之線之關係。亦即，舉例而言，當位準比率EN_SN在步驟S10計算為值SNa(如圖4中所示)時，若位準比率差DIF_SN在步驟S12計算為DIF_SN1，則臨限值TH0與值THa之間的差a1設定為校正值Aa，且若位準比率差DIF_SN在步驟S12計算為DIF_SN2，則臨限值TH0與值THb之間的差a2設定為校正值Aa。藉由自距離D0減去以此方式設定之校正值Aa而計算出之距離D1在判定器46處與臨限值TH0進行比較，藉此將此實施例的驗證維持在所要精確度水準。

更特定言之，此實施例之設定器51基於已建立以滿足以上條件之表來自位準比率EN_SN及位準比率差DIF_SN判定校正值Aa。圖5為說明在步驟S14使用之表之內容之示意圖。如圖5中所示，對應於不同位準比率差DIF_SN(DIF_SN1、DIF_SN2、DIF_SN3.....)之複數個表儲存於儲存單元35中。在每一表中，每一位準比率EN_SN與滿足圖4之與位準比率EN_SN之關係的相應校正值Aa相關聯。亦即，在對應於位準比率差DIF_SN之表中，每一位準比率EN_SN與一相應校正值Aa相關聯，使得藉由將校正值Aa與臨限值TH0相加獲取之臨限值TH與位準比率EN_SN具有一對應於位準比率差DIF_SN的線之關係。

在步驟S14，設定器51搜尋對應於在步驟S12計算出之位

準比率差 DIF_SN 之表，且將對應於包括於所擷取表中之在步驟 S10 計算出之位準比率 EN_SN 的校正值 Aa 輸出至校正器 44。若不存在對應於在步驟 S12 計算出之位準比率差 DIF_SN 之表，則設定器 51 根據來自對應於先於及緊隨所計算之位準比率差 DIF_SN 之位準比率差 DIF_SN 之表之位準比率 EN_SN 來指定校正值 Aa，且藉由在所指定之校正值 Aa 之間內插來計算一校正值 Aa，該校正值 Aa 實際上待應用於校正器 44 之校正。因此，舉例而言，若位準比率差 DIF_SN 在步驟 S12 計算為圖 2 之值 DIF_SN1 與值 DIF_SN2 之平均數，則校正值 Aa 計算為對應於值 DIF_SN1 的校正值 Aa(值 a1) 與對應於值 DIF_SN2 之校正值 Aa(值 a2) 之平均數。

然而，臨限值 TH 與位準比率 EN_SN 及位準比率差 DIF_SN 之關係係(例如)在登錄雜音之特性與驗證雜音之特性一致的假設下判定。然而，在許多狀況下，登錄雜音之特性不同於驗證雜音之特性。因此，設定器 52 根據登錄雜音與驗證雜音之間的相關值 NOISE_DIF 來計算校正值 Ab 以校正距離 D0(步驟 S15)。校正器 44 自距離 D0 減去校正值 Ab。

圖 6 為說明相關值 NOISE_DIF 與校正值 Ab 之間的關係之圖。使用等式(3)計算出之相關值 NOISE_DIF 根據登錄雜音與驗證雜音之間的相關性而在 "-1" 至 "1" 之範圍內變化。當兩個雜音之特性完全一致時，相關值 NOISE_DIF 為 "1"。若登錄雜音與驗證雜音之間的關係等效於在判定圖 4 之關

係時的關係(若兩個雜音在此實施例中一致)，則無需根據相關值 NOISE_DIF 來校正距離 D0。當登錄雜音與驗證雜音之間的關係與判定圖 4 之關係時的關係差別較大時，必須更大地校正距離 D0。因此，當相關值 NOISE_DIF 為 "1" 時，設定器 52 將校正值 Ab 設定為零(亦即，未校正距離 D0)，且設定器 52 隨著相關值 NOISE_DIF 減小至 "1" 以下而將校正值 Ab 設定為較高。更特定言之，儲存單元 35 儲存一表，其中每一相關值 NOISE_DIF 與一相應校正值 Aa(其滿足與相關值 NOISE_DIF 之此關係)相關聯，且設定器 52 基於此表判定校正值 Ab。

另外，圖 4 中所示之臨限值 TH 與位準比率 EN_SN 及位準比率差 DIF_SN 之關係係(例如)基於在特定持續期間(時長) L0 期間連續說出登錄聲音時所收集的資料而判定。然而，實際初始登錄中之登錄聲音之語音長度 EN_SPEECH_LEN 係不同的。因此，設定器 53 根據語音長度 EN_SPEECH_LEN 來計算校正值 Ac 以校正距離 D0(步驟 S16)。校正器 44 自距離 D0 減去校正值 Ac。

圖 7 為展示語音長度 EN_SPEECH_LEN 與校正值 Ac 之間的關係之圖。若判定圖 4 之關係時語音長度 EN_SPEECH_LEN 與持續期間 L0 一致，則無需根據語音長度 EN_SPEECH_LEN 來校正距離 D0。因此，若儲存單元 32 中儲存之語音長度 EN_SPEECH_LEN 與持續期間 L0 一致，則設定器 53 將校正值 Ac 設定為 "0"。隨著語音長度 EN_SPEECH_LEN 增加，登錄聲音之語音更穩定，使得特徵量 EN_DATA

更可靠地反映授權使用者之固有的基本特性，藉此增加距離 D_0 的精確度。因此，即使相對於距離 D_0 降低臨限值 TH_0 以便減少接受陌生人之真確性之可能性，亦不會不當地增加拒絕授權使用者之可能性。

因此，若語音長度 EN_SPEECH_LEN 大於持續期間 L_0 ，則設定器 53 根據語音長度 EN_SPEECH_LEN 而選擇負值作為校正值 A_c 。因為校正器 44 自距離 D_0 減去校正值 A_c ，所以若校正值 A_c 為負，則經校正之距離 D_1 大於距離 D_0 。亦即，因為臨限值 TH_0 相對於距離 D_0 減小，所以錯誤地接受陌生人之真確性之可能性得以減小。相反，若語音長度 EN_SPEECH_LEN 小於持續期間 L_0 ，則設定器 53 根據語音長度 EN_SPEECH_LEN 而選擇正值作為校正值 A_c 。舉例而言，設定器 53 基於每一語音長度 EN_SPEECH_LEN 與相應校正值 A_c 相關聯之表來判定校正值 A_c 。

如上所述，在此實施例中，無關於驗證雜音或登錄雜音之特性，可將驗證維持在一所要精確度水準，因為距離 D_0 係根據驗證雜音與驗證聲音 (V_SN) 之間的關係或登錄雜音與登錄聲音 (EN_SN) 之間的關係來校正。亦即，驗證之便利性之增加 (FRR 之減少) 及驗證的精確度之增加 (FAR 之減少) 皆可在不受驗證雜音或登錄雜音影響之情形下達成。

在此實施例中，尤其校正值 A_a 係基於滿足所要條件之臨限值 TH 與位準比率 EN_SN 成一對應於位準比率差 DIF_SN 之線的關係之認知來調整。因此，此實施例之優勢為有可能以高精確度指定最佳校正值 A_a ，同時充分減少設定校正

值 Aa 所需之變數之數目。亦有可能藉由充分考慮初始登錄或驗證時聲音驗證裝置 100 之環境來實現高度精確的驗證，因為距離 D0 係根據相關值 NOISE_DIF 或語音長度 EN_SPEECH_LEN 來校正。

<C：經修改實施例>

以上實施例可以各種方式修改。以下說明特定經修改實施例。以下經修改實施例之適當組合亦為可能的。

(1) 經修改實施例 1

儘管在以上實施例中所說明之組態使用儲存於儲存單元 35 中之表，但藉由使用預定等式的計算方法計算校正值 A(Aa、Ab 及 Ac) 之組態亦為可能的。舉例而言，儲存單元 35 儲存分別表示根據位準比率差 (DIF_SN) 之線 (亦即，圖 4 之線，每一者界定位準比率 EN_SN 與臨限值 TH 之間的關係) 之複數個公式，且設定器 51 藉由將在步驟 S10 計算出之位準比率 EN_SN 代入等式而根據在步驟 S12 計算出的位準比率差 DIF_SN 來計算臨限值 TH，且藉由自所計算之臨限值 TH 減去臨限值 TH0 來計算校正值 Aa。此實例亦使用設定器 52 基於表示校正值 NOISE_DIF 與校正值 Ab 之間的關係之等式來計算校正值 Ab 之組態，或設定器 53 基於表示語音長度 EN_SPEECH_LEN 與校正值 Ac 之間的關係之等式來計算校正值 Ac 之組態。

(2) 經修改實施例 2

判定將距離 D0 校正至何程度之變數不限於校正值 NOISE_DIF 及語音長度 EN_SPEECH_LEN。舉例而言，可

替代或結合根據語音長度EN_SPEECH_LEN設定校正值Ab之組態來使用根據驗證聲音之語音長度V_SPEECH_LEN來計算校正值Ab的組態或根據語音長度EN_SPEECH_LEN與語音長度V_SPEECH_LEN之平均值來計算校正值Ab之組態。舉例而言，校正值Ab經設定使得經校正之距離D1隨著語音長度V_SPEECH_LEN增加而增加，此類似於校正值Ab與語音長度EN_SPEECH_LEN之間的關係。

亦使用校正值控制器50基於除以上所述之變數以外之變數來判定每一校正值的組態。舉例而言，可根據包括於登錄聲音或驗證聲音中之有聲聲響與無聲聲響之持續期間之間的比率來設定每一校正值(在出聲區段P2中)。隨著有聲聲響之比率增加，特徵量EN_DATA及V_DATA之值更可靠地反映說話者的特性，藉此增加距離D0之精確度。因此，即使距離D1自距離D0增加，FRR亦不會不當地增加。因此，使用每一校正值經設定使得距離D1隨著登錄聲音或驗證聲音中之有聲聲響之比率增加而增加的組態。另外，亦可使用每一校正值經設定使得距離D1隨著登錄聲音或驗證聲音之音節之數目增加而增加的組態，因為距離D0之精確度隨著登錄聲音或驗證聲音之音節的數目增加而增加。

(3)經修改實施例3

亦使用變數與校正值之間的關係為可變之組態。舉例而言，可提供藉由根據操作單元10之操縱更新表來改變校正值Aa與位準比率EN_SN或位準比率差DIF_SN之關係的組態。類似地，可根據操作單元10之操縱來修改對應於相關

值 NOISE_DIF 之校正值 A_b 或對應於語音長度 EN_SPEECH_LEN 的校正值 A_c 。此等組態可根據使用者之需要而達成驗證。在藉由如經修改實施例 1 中所說明之等式的計算來計算校正值 $A(A_a、A_b$ 及 $A_c)$ 之組態中，可根據操作單元 10 之操縱來修改該等等式中之每一者的內容(例如，變數之係數)。

(4) 經修改實施例 4

儘管以上實施例中所說明之組態在驗證中使用距離 D_0 (經校正之距離 D_1)，但指示登錄聲音與驗證聲音之間的相似性之值不限於距離 D_0 。舉例而言，使用基於任何變數(亦即，任何指數值)執行驗證之組態，該變數之值隨著登錄聲音的特性與驗證聲音之特性彼此接近而增加。在此組態中，與以上實施例相比，校正值 $A(A_a、A_b$ 及 $A_c)$ 根據每一變數之增加或減小反轉。舉例而言，與圖 4 之組態相反，此組態中之校正值 A_a 隨著位準比率 EN_SN 增加而減小。

(5) 經修改實施例 5

儘管在以上實施例中所說明之組態中將三種類型之校正值 $A_a、A_b$ 及 A_c 個別地輸出至校正器 44，但亦可使用校正值控制器 50 輸出校正值 $A_a、A_b$ 與 A_c 的總和之組態。另外，可適當地修改校正器 44 校正距離 D_0 之計算操作。舉例而言，使用校正器 44 將每一校正值 $A(A_a、A_b$ 及 $A_c)$ 與距離 D_0 相加或將距離 D_0 乘以每一校正值 $A(A_a、A_b$ 及 $A_c)$ 之組態。在藉由將每一校正值 A 與距離 D_0 相加來計算 D_1 之組態

中，與圖1的組態相比，每一校正值A之正負號反轉。

(6)經修改實施例6

儘管以上實施例中所說明之組態將預定臨限值TH0與隨位準比率EN_SN線性變化之變數(臨限值)TH之間的差計算為校正值Aa，但可適當地改變位準比率EN_SN與臨限值TH之間的關係。舉例而言，此經修改實施例使用一組態，其中根據位準比率EN_SN及位準比率差DIF_SN來指定臨限值TH，以使得位準比率EN_SN及臨限值TH滿足對應於位準比率差DIF_SN之曲線的關係(亦即，使得臨限值TH隨著位準比率EN_SN而沿對應於位準比率差DIF_SN之曲線改變)，且臨限值TH與臨限值TH0之間的差判定為校正值Aa。其將足以設定臨限值TH使得臨限值TH與位準比率EN_SN如上所述根據驗證雜音而具有不同關係(通常，根據位準比率差DIF_SN而成線性或曲線關係)。舉例而言，根據導出位準比率EN_SN與對應於一位準比率差DIF_SN之臨限值TH之間的關係之測試之結果或根據對該測試結果的統計處理結果來適當地改變此等變數之間的特定關係。因此，本發明之一較佳實施例中之校正值Aa判定為預定臨限值TH0與臨限值TH之間的差，臨限值TH根據驗證雜音滿足與位準比率EN_SN之關係。

(7)經修改實施例7

儘管在以上實施例中說明聲音驗證裝置100用以區分授權使用者與陌生人之狀況，但根據以上實施例之聲音驗證裝置100亦可用於聲音密碼驗證，其判定由驗證聲音指定

的密碼是否匹配先前基於授權使用者之語音而登錄之密碼。

【圖式簡單說明】

圖1為說明根據本發明之一實施例之聲音驗證裝置的組態之方塊圖。

圖2為由輸入單元產生之聲響信號之波形圖。

圖3為說明由校正值控制器執行之程序之流程圖。

圖4為說明校正值Aa與位準比率EN_SN及位準比率差DIF_SN之關係之圖。

圖5為說明用以判定校正值Aa之表之內容之示意圖。

圖6為說明相關值NOISE_DIF與校正值Ab之間的關係之圖。

圖7為說明語音長度EN_SPEECH_LEN與校正值Ac之間的關係之圖。

圖8為說明臨限值與FAR及FRR之關係之圖。

【主要元件符號說明】

10	操作單元
15	輸入單元
20	特性分析器
22	區段偵測器
23	開關
25	雜音分析器
26	聲音分析器
28	特徵分析器

32	儲存設備
35	儲存單元
40	驗證器
42	距離計算器
44	校正器
46	判定器
50	校正值控制器
51	設定器
52	設定器
53	設定器
60	輸出單元
100	聲音驗證裝置
Aa	校正值
Ab	校正值
Ac	校正值
D ₀	距離
D ₁	距離
L ₀	持續期間
L ₁	虛線
L ₂	虛線
P	偵測區段
P1	未出聲區段
P2	出聲區段
R	虛線

S

聲響信號

TH₀

臨限值

五、中文發明摘要：

在一種聲音驗證裝置中，一特性分析器分析在一受驗者產生一用於該受驗者之驗證之樣本聲音時在該受驗者周圍產生的一樣本雜音之特性。一設定單元根據由該特性分析器分析之該樣本雜音之該等特性來設定一校正值。一校正單元基於該所設定校正值來校正一指數值，該指數值指示先前已登錄之一參考聲音之一特徵量與自該受驗者獲取的該樣本聲音之一特徵量之間的一相似程度。一判定器藉由將該經校正之指數值與一預定臨限值進行比較來判定該受驗者之真確性。

六、英文發明摘要：

In a voice authentication apparatus, a characteristics analyzer analyzes characteristics of a sample noise which is generated around a subject while the subject generates a sample voice for authentication of the subject. A setting unit sets a correction value according to the characteristics of the sample noise analyzed by the characteristics analyzer. A correction unit corrects an index value, which indicates a degree of similarity between a feature quantity of a reference voice which has been previously registered and a feature quantity of the sample voice obtained from the subject, based on the set correction value. A determinator determines authenticity of the subject by comparing the corrected index value with a predetermined threshold value.

十一、圖式：

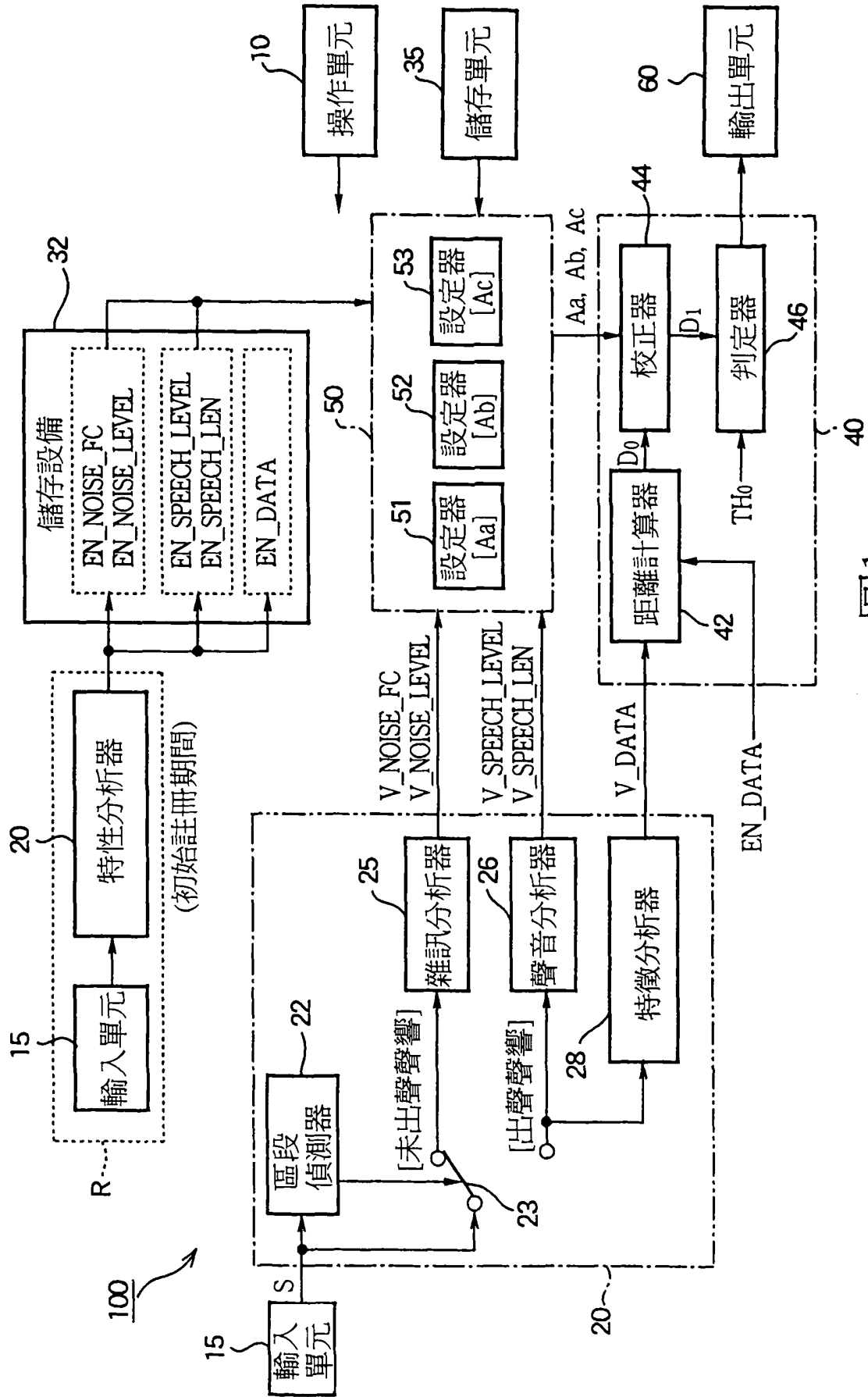


圖1

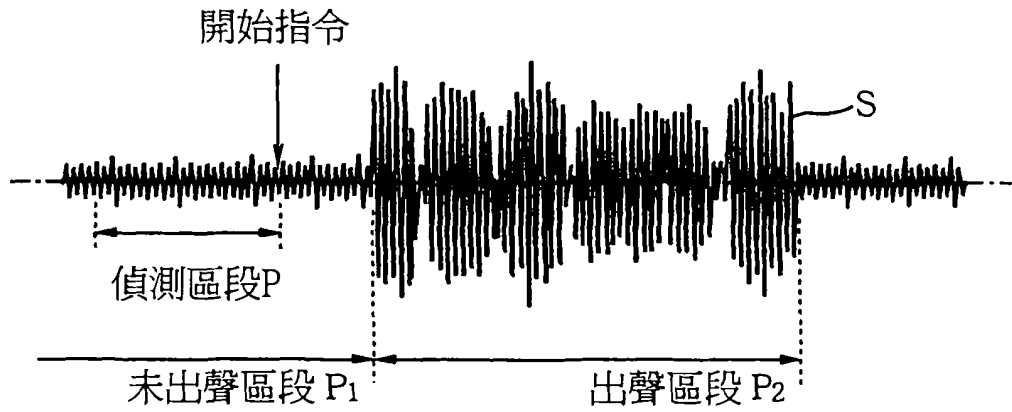


圖2

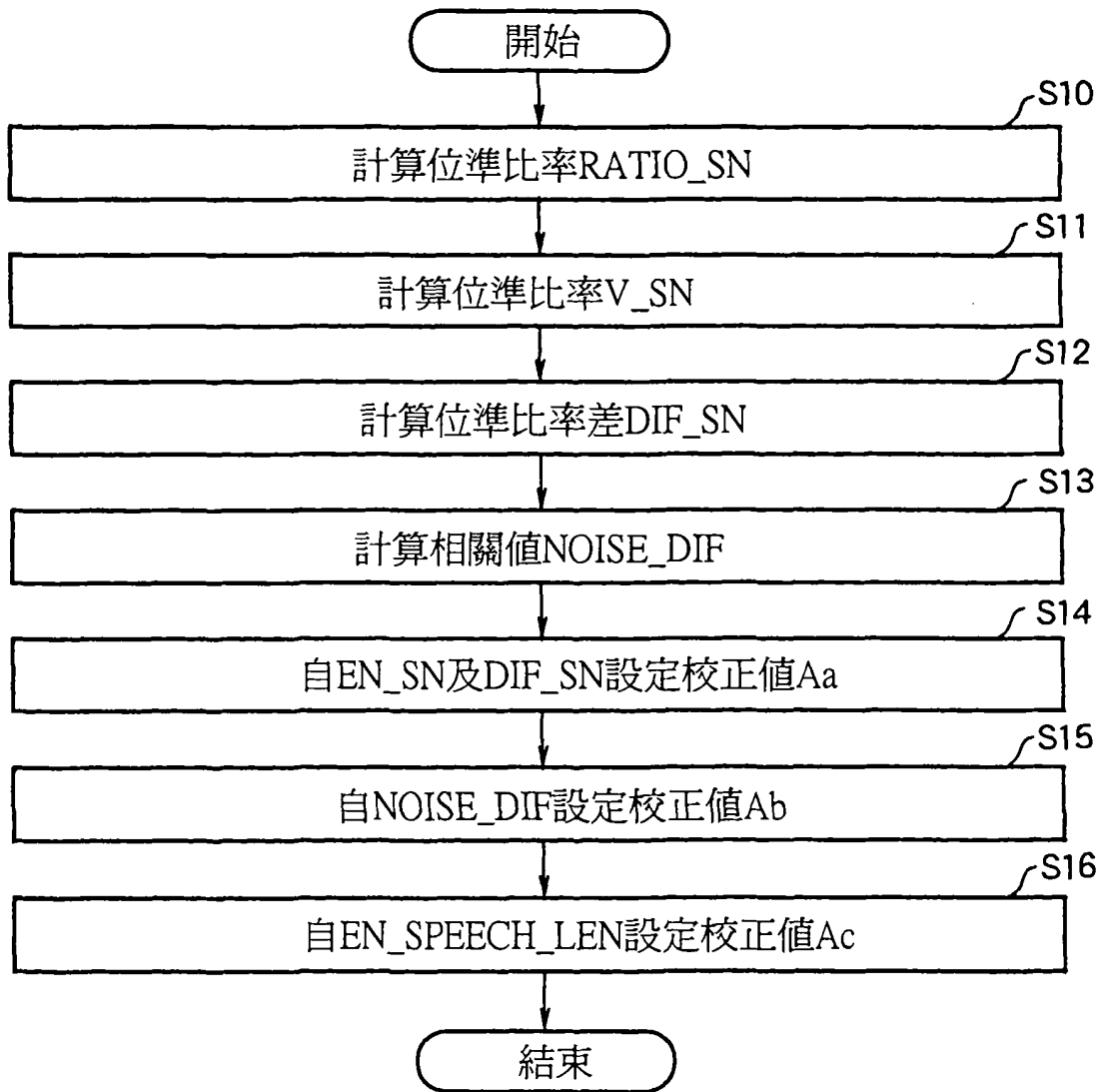


圖3

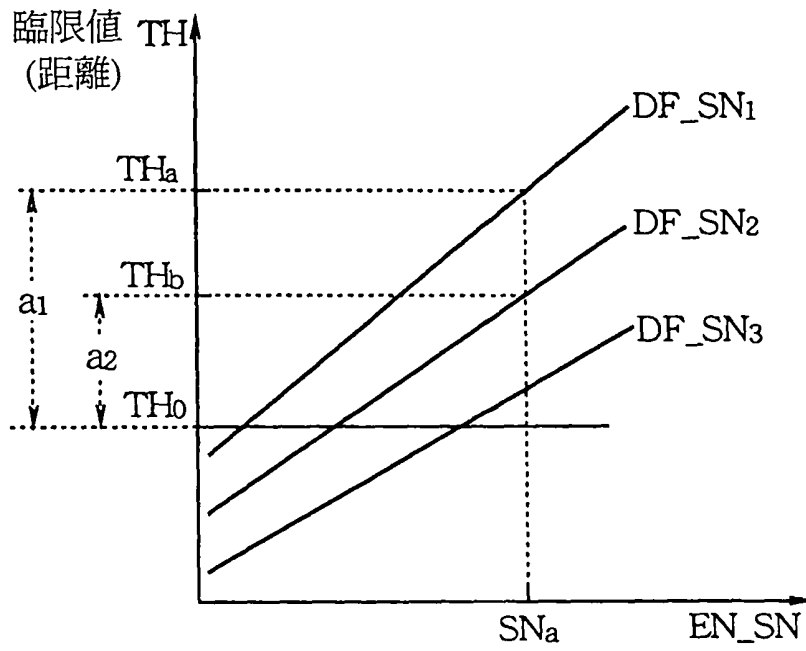


圖4

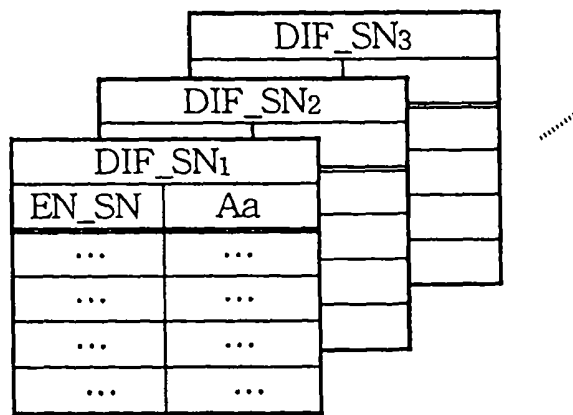


圖5

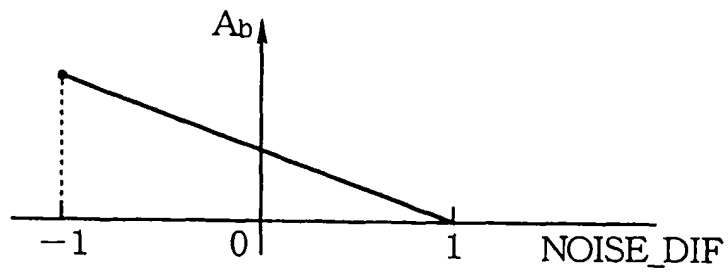


圖6

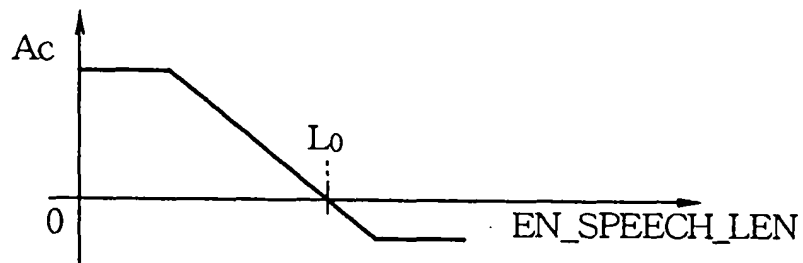


圖7

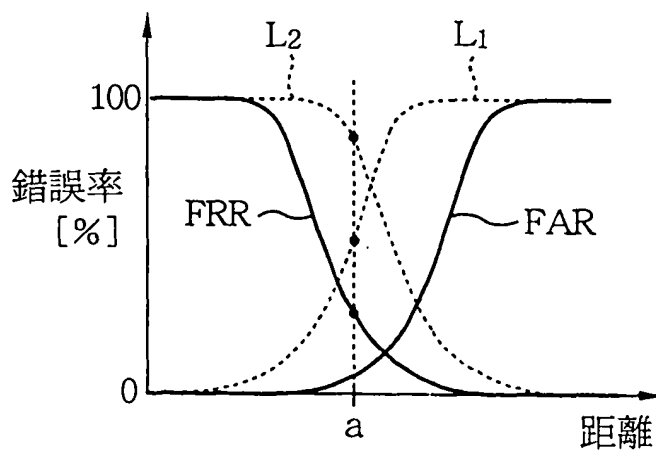


圖8

七、指定代表圖：

(一)本案指定代表圖為：第（ 1 ）圖。

(二)本代表圖之元件符號簡單說明：

10	操作單元
15	輸入單元
20	特性分析器
22	區段偵測器
23	開關
25	雜音分析器
26	聲音分析器
28	特徵分析器
32	儲存設備
35	儲存單元
40	驗證器
42	距離計算器
44	校正器
46	判定器
50	校正值控制器
51	設定器
52	設定器
53	設定器
60	輸出單元
100	聲音驗證裝置
Aa	校正值

Ab	校正值
Ac	校正值
D ₀	距離
D ₁	距離
R	虛線
S	聲響信號
TH ₀	臨限值

八、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

(無)

十、申請專利範圍：

1. 一種聲音驗證裝置，其包含：

特性分析構件，其分析一受驗者產生一用於該受驗者之驗證之樣本聲音時在該受驗者周圍產生的一樣本雜音之特性；

第一設定構件，其根據經該特性分析構件分析之該樣本雜音之該等特性來設定一第一校正值；

校正構件，其基於由該第一設定構件所設定之該第一校正值來校正一指數值，該指數值指示先前已登錄之一參考聲音之一特徵量與自該受驗者獲取之該樣本聲音之一特徵量之間的一相似程度；及

一判定器，其藉由將該由該校正構件校正之指數值與一預定臨限值進行比較來判定該受驗者之真確性；

其中該第一設定構件根據該樣本雜音之該等特性及該參考聲音與一在該參考聲音的登錄期間產生之參考雜音之間的一位準比率(level ratio)來設定該第一校正值，使得該第一校正值設定為該預定臨限值與一變數之間的一差，該變數係與該樣本雜音之該等特性及該參考聲音與該參考雜音之間的該位準比率相關聯而決定。

2. 如請求項1之聲音驗證裝置，其中該第一設定構件將該第一校正值設定為該預定臨限值與該變數之間的該差，該變數隨著該參考聲音與該參考雜音之間的該位準比率而沿一對應於該樣本雜音之該等特性之線或曲線變化。

3. 如請求項1或2之聲音驗證裝置，其中該第一設定構件根

99年11月4日修正替換頁

據該參考聲音與該參考雜音之間的該位準比率及該樣本雜音之該等特性來設定該第一校正值，該樣本雜音之該等特性係作為該參考聲音與該參考雜音之間的該位準比率及該樣本聲音與該樣本雜音之間的一位準比率之間的一位準比率差。

4. 如請求項1之聲音驗證裝置，其進一步包含第二設定構件，該第二設定構件根據該樣本雜音與該參考雜音之間的一差來設定一第二校正值，

其中該校正構件根據該第一校正值及該第二校正值來校正該指數值。

5. 如請求項1之聲音驗證裝置，其進一步包含第三設定構件，該第三設定構件根據該樣本聲音或該參考聲音之一期間來設定一第三校正值，

其中該校正構件根據該第一校正值及該第三校正值來校正該指數值。

6. 一種聲音驗證方法，其包含：

分析在一受驗者產生一用於該受驗者之驗證之樣本聲音時在該受驗者周圍產生的一樣本雜音之特性；

根據該樣本雜音之該等分析後之特性設定一校正值；

基於該所設定校正值校正一指數值，該指數值指示先前已登錄之一參考聲音之一特徵量與自該受驗者獲取的該樣本聲音之一特徵量之間的一相似程度；及

藉由將該經校正指數值與一預定臨限值進行比較來判定該受驗者之真確性；且

其中根據該樣本雜音之該等特性及該參考聲音與一在該參考聲音的登錄期間產生之參考雜音之間的一位準比率(level ratio)來設定該校正值，使得該校正值設定為該預定臨限值與一變數之間的一差，該變數係與該樣本雜音之該等特性及該參考聲音與該參考雜音之間的該位準比率相關聯而決定。

7. 一種用於一電腦中之機器可讀媒體，該媒體含有可由該電腦執行以執行以下處理之程式指令：

- 一特性分析處理，其分析在一受驗者產生一用於該受驗者之驗證之樣本聲音時在該受驗者周圍產生的一樣本雜音之特性；

- 一設定處理，其根據在該特性分析處理中分析後之該樣本雜音之該等特性來設定一校正值；

- 一校正處理，其基於該在該設定處理中所設定之該校正值來校正一指數值，該指數值指示先前已登錄之一參考聲音之一特徵量與自該受驗者獲取之該樣本聲音之一特徵量之間的一相似程度；及

- 一判定處理，其藉由將該經校正指數值與一預定臨限值進行比較來判定該受驗者之真確性；

其中根據該樣本雜音之該等特性及該參考聲音與一在該參考聲音的登錄期間產生之參考雜音之間的一位準比率(level ratio)來設定該校正值，使得該校正值設定為該預定臨限值與一變數之間的一差，該變數係與該樣本雜音之該等特性及該參考聲音與該參考雜音之間的該位準比率相關聯而決定。