

【11】證書號數：I585756

【45】公告日：中華民國 106 (2017) 年 06 月 01 日

【51】Int. Cl. : G10L25/51 (2013.01) G10L25/66 (2013.01)

發明

全 5 頁

【54】名稱：口吃偵測方法與裝置、電腦程式產品

METHOD AND DEVICE FOR RECOGNIZING STUTTERED SPEECH AND
COMPUTER PROGRAM PRODUCT

【21】申請案號：104124768

【22】申請日：中華民國 104 (2015) 年 07 月 30 日

【11】公開編號：201705127

【43】公開日期：中華民國 106 (2017) 年 02 月 01 日

【72】發明人：楊淑蘭(TW) YANG, SHU LAN；葉品忻(TW) YEH, PIN HSIN；楊智傑(TW)
YANG, CHIH CHIEH

【71】申請人：國立屏東大學

NATIONAL PINGTUNG UNIVERSITY

屏東縣屏東市民生路 4 之 18 號

【74】代理人：李世章；秦建譜

【56】參考文獻：

CN 1714390B

CN 201076475Y

方芳, “基於 ANN 和 HMM 模型的口吃語音識別研究”, 天津大學 2010 年碩士論文,
2010.12.31, <http://m.book118.com/html/2015/1018/27509265.shtm>

審查人員：黃衍勳

【57】申請專利範圍

1. 一種口吃偵測方法，包括：取得一聲音訊號，並將該聲音訊號分為多個音節；根據一第一變數與一第二變數決定多個位置向量，其中每一該些位置向量具有一第一位置與一第二位置，且該第一位置與該第二位置是根據以下方程式(1)所決定：

$$\begin{cases} i \leq a \times (b + 1) \\ i - j = a \times (b - m), m = 0, 1, 2, \dots (b - 1) \dots (1) \end{cases}$$

其中 i 為該第一位置，j 為該第二位置，a 為該第一變數，b 為該第二變數；對於每一該些位置向量，判斷該些音節中第 i 個音節與第 j 個音節是否相似，以判斷該位置向量是否符合一相似條件；以及若該些位置向量中符合該相似條件的位置向量的數目大於一相似數目臨界值，則判斷該聲音訊號存在一口吃現象。

2. 如申請專利範圍第 1 項所述之口吃偵測方法，更包括：判斷是否已嘗試所有的該第一變數與該第二變數，其中 $a \times (b + 1)$ 小於等於該些音節的總數目；若尚未嘗試所有的該第一變數與該第二變數，調整該第一變數與該第二變數，並且根據調整後的該第一變數與該第二變數重新決定該些位置向量。
3. 如申請專利範圍第 1 項所述之口吃偵測方法，其中判斷該第 i 個音節與該第 j 個音節是否相似的步驟包括：根據以下方程式(2)~(7)分別計算該第 i 個音節與該第 j 個音節的多個特徵值：

$$\text{Volume} = \sum_i^n |s_i| \dots (2)$$

$$\text{ZCR} = \sum_i^n \text{sgn}[\text{sgn}[s_i \times s_{i+1}] - 1] \dots (3)$$

(2)

$$\text{Entropy} = - \sum_{k=1}^N p_k \log p_k, p_k = s(f_k) / \sum_{i=1}^N s(f_i) \dots (4)$$

$$\text{HOD} = \sum_{i=1}^{n-1} |(s_{i+1} - s_i)| \dots (5) \quad \text{VH} = \alpha \times \text{Volume} + (1-\alpha) \times \text{HOD} \dots (6) \quad \text{VE} = \text{Volume} \times (1 - \text{entropy}) \dots (7)$$

其中 Volume 為音量特徵值，ZCR 為過零率特徵值，Entropy 為熵值特徵值，HOD 為微分特徵值，VH 為音量微分特徵值，VE 為音量熵值特徵值， s_i 為該聲音訊號在時間點 i 的振幅， n 為音框的長度， $\text{sgn}[]$ 代表一正負號函數， $s(f_k)$ 代表該聲音訊號在頻率域中第 k 個頻率的振幅， N 代表該聲音訊號在該頻率域的長度， α 為常數；以及根據該第 i 個音節的該些特徵值與該第 j 個音節的該些特徵值，來判斷該第 i 個音節與該第 j 個音節的是否相似。

4. 如申請專利範圍第 3 項所述之口吃偵測方法，其中將該聲音訊號分為該些音節的步驟包括：計算該聲音訊號的該音量微分特徵值以取得一音量微分特徵訊號；判斷該音量微分特徵訊號是否大於一音節臨界值；以及從該音量微分特徵訊號中取得大於該音節臨界值的部分以決定該些音節。
5. 如申請專利範圍第 3 項所述之口吃偵測方法，其中根據該第 i 個音節的該些特徵值與該第 j 個音節的該些特徵值，來判斷該第 i 個音節與該第 j 個音節是否相似的步驟包括：對於每一該些特徵值，對該第 i 個音節與該第 j 個音節執行一動態時軸校正演算法以判斷該第 i 個音節與該第 j 個音節在對應的該特徵值是否相似；以及若有 f 個該些特徵值判斷該第 i 個音節與該第 j 個音節為相似，則判斷該第 i 個音節與該第 j 個音節相似，其中 f 為正整數，其小於等於該些特徵值的數目。
6. 一種電腦程式產品，當電腦載入該電腦程式產品並執行後，可完成上述申請專利範圍第 1~5 項中任一項所述之口吃偵測方法。
7. 一種口吃偵測裝置，包括：一記憶體，儲存多個指令；以及一處理器，用以執行該些指令以執行多個步驟：取得一聲音訊號，並將該聲音訊號分為多個音節；根據一第一變數與一第二變數決定多個位置向量，其中每一該些位置向量具有一第一位置與一第二位置，且該第一位置與該第二位置是根據以下方程式(1)所決定：

$$\begin{cases} i \leq a \times (b + 1) \\ i - j = a \times (b - m), m = 0, 1, 2, \dots (b - 1) \end{cases} \dots (1)$$

其中 i 為該第一位置， j 為該第二位置， a 為該第一變數， b 為該第二變數；對於每一該些位置向量，判斷該些音節中第 i 個音節與第 j 個音節是否相似以判斷該位置向量是否符合一相似條件；以及若該些位置向量中符合該相似條件的位置向量的數目大於一相似數目臨界值，判斷該聲音訊號存在一口吃現象。

圖式簡單說明

[圖 1]是根據一實施例繪示口吃偵測裝置的示意圖。

[圖 2]是根據一實施例繪示將聲音訊號分為多個音節的示意圖。

[圖 3]是根據一實施例繪示音節矩陣的示意圖。

[圖 4]是根據一實施例繪示口吃偵測方法的流程圖。

(3)

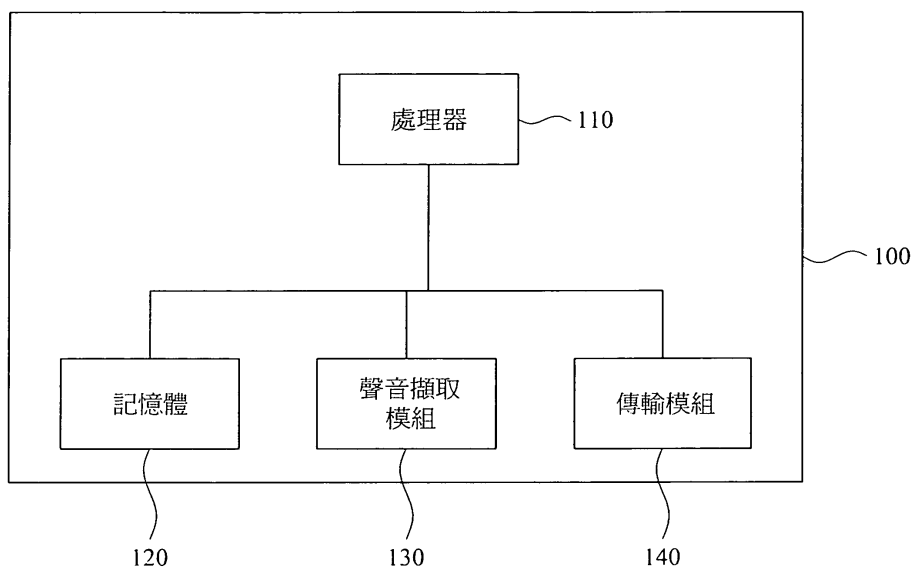


圖 1

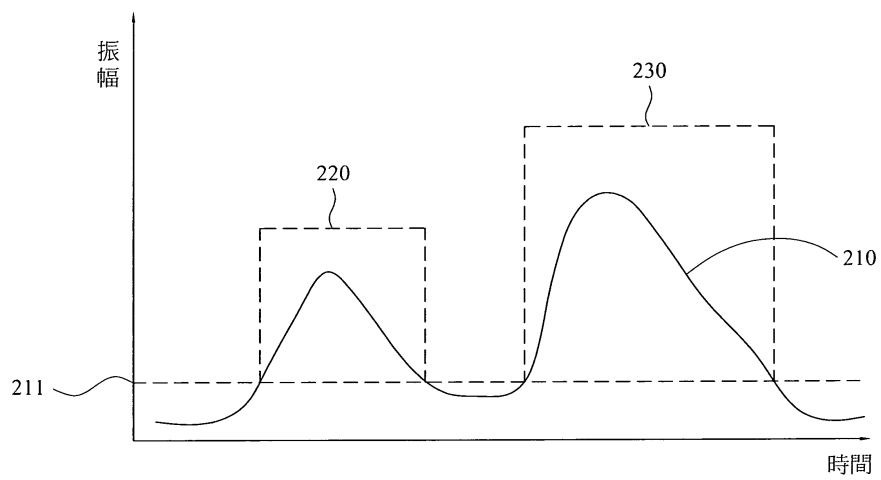


圖 2

(4)

311

i	1	2	3	4	5	6	7	8	9	10	11	12	13
j	有	八	個	有	八	個	有	八	個	有	八	個	人
1	有	1	2	3	4	5	6	7	8	9	10	11	12
2	八		1	2	3	4	5	6	7	8	9	10	11
3	個			1	2	3	4	5	6	7	8	9	10
4	有				1	2	3	4	5	6	7	8	9
5	八					1	2	3	4	5	6	7	8
6	個						1	2	3	4	5	6	7
7	有							1	2	3	4	5	6
8	八								1	2	3	4	5
9	個									1	2	3	4
10	有										1	2	3
11	八											1	2
12	個												1
13	人												

312
313
300

圖 3

