公告本

申請日期	87. 7. 13
案號	87111338
類別	G10L 5/04

394925
A4
C4
394925

（以上各欄由本局填註）

發明 ~~新型~~ 專利說明書

一、發明 ~~新型~~ 名稱	中文	語音編碼器爲基礎的語音識別器
	英文	"A VOCODER-BASED VOICE RECOGNIZER"
二、發明 ~~創作~~ 人	姓名	1.葉賀達 荷西寇維茲 2.嘉寶雷 艾嵐
	國籍	1-2.以色列
	住、居所	1.以色列莫比波艾姆塞克市克法哈那基街283號 2.以色列德拉法市艾力亞哈金街12號
三、申請人	姓名（名稱）	美商技術促進認可工業公司
	國籍	美國
	住、居所（事務所）	美國加州凱茲沃市塔潘咖峽谷路9574號
	代表人姓名	嘉寶雷 艾嵐

經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

裝 訂 線

（由本局填寫）

承辦人代碼：	A6 B6
大　　類：	
ＩＰＣ分類：	

本案已向：

國（地區）	申請專利，申請日期：	案號：	，□有　□無主張優先權
美國	1998年1月8日	09/002,616	☑有□無主張優先權

有關微生物已寄存於：　　，寄存日期：　　，寄存號碼：

（請先閱讀背面之注意事項再填寫本頁各欄）

裝　訂　線

經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（1）

發明的範圍

本發明有關於一般的語音識別器，並且是使用LPC語音編碼器資料爲輸入的語音識別器。

發明的背景

語音識別器是眾所周知的技術並且在很多應用上使用。例如：語音識別使用在要求和控制的行動通訊元件，在電腦錄音機，在小孩的玩具和在汽車電話上。在所有這些系統中，語音信號被數位化然後參數化。這個參數化的輸入信號和已知語調的參考參數化信號作比較。這個識別語調就是和輸入信號最吻合之參考信號有關的語調。

語音識別系統最常使用於語音撥號系統，其中當一個使用者說出他想打給的人名後，語音識別系統從先前提供的參考資料表列中識別出人名，並提供和此人名有關的電話號碼。電話機之後就撥出這個號碼。其結果就是使用者連接到目的地站而不用查出他要撥的號碼也不用使用手去撥號。

語音撥號對於汽車行動電話更爲重要，因爲使用者通常就是車的駕駛，因此必須長時間注意道路的情況，如果駕駛想要打給某人，駕駛說出他要打出的人名是比他自己撥號較爲安全。

圖1列出參考符號，表示出一個數位行動電話的主要原件。典型的一個行動電話包含一個麥克風10，一個喇叭12，元件14是轉換類比和數位的信號，一個語音編碼器16實施於一個編號爲DSP-1的數位信號處理晶片，一個作

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（ 2 ）

業系統18實施於一個微控制器或一個中央處理單元，一個射頻界面單元(RF unit)19和一個天線20。在傳送過程中，麥克風10產生類比的語音信號，其信號被單元14數位化。語音編碼器16把聲音取樣點作壓縮使得傳送的資料量減低，經由RF單元19和天線20傳送到另一個行動電話。接收端的行動電話的天線20提供了接收信號，而RF單元19和聲音編碼器16相對的反壓縮收到的信號成爲聲音的取樣點。單元14把聲音取樣點轉換成類比信號而由喇叭12送出。作業系統18控制行動電話的運作。

對於聲音撥號系統，行動電話系統額外包含一個聲音識別器22實施於一個分離的DSP元件編號爲DSP-2，其把收到的數位化聲音取樣點當做輸入，把聲音信號參數化，比對參考聲音信號和參數化輸入信號，聲音識別器22典型的提供作業系統18的吻合信號辨識，或是如果一個電話號碼和吻合信號有關，識別器22提供相關的號碼。

圖2列出參考符號，大致例示語音識別器22的運作，數位化的聲音取樣點根據預定的長度分成區段(frames)，例如是5-20 msec且這些區段在步驟28由識別器22提供。對於每個區段，識別器22在步驟30中先計算這個區段的能量。

圖3亦列出參考符號，解釋了說出的字"RICHARD"的每個區段能量爲一個時間函數。能量信號有兩個脈衝31和33相對於這個字的兩個音節。當沒說話時(如參考數字35所示)，或是在兩個音節之間，能量的大小是相對較低的。

因此，在圖2的步驟32中，識別器22在能量信號中找尋

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（ 3 ）

一個字的開端和尾端。一個字的開端字義如點37所示，能量信號在超過一個先定下的時間長度都是較低的時候之後在能量開始處有一個顯著的上升。一個字的尾端定義如點39所示，能量信號在超過一個先定下的時間長度都是較低的時候之後在能量結束處有一個顯著的下降。在圖3中，開端點37發生在約0.37sec而尾端點發生在約0.85sec。

如在步驟34所作的確認，當一個字找到後，語音識別器22在步驟36執行一個線性預測編碼分析來產生說出話的參數。在步驟38中，語音識別器22計算說出話的識別特性，在步驟40中，語音識別器22。在參考字庫中的參考字的識別特性中找尋一個吻合的字。相對的，語音識別器22把識別特性存入參考的字庫，這個過程稱爲"訓練(training)"。

不幸的，語音識別的過程有艱深的計算量，因此必須做在第二個DSP元件，DSP-2。這使得行動電話的成本大大的增加。

本發明的簡述

本發明的目標是提供一個運作在壓縮語音資料之下而不是所有語音資料取樣點的語音識別器，壓縮是根據LPC的語音編碼器，因而使得識別器所須運作的計算量降低。因此，語音識別可以做在微控制器或CPU之中，且也把運作系統做在其中。因爲語音識別器並不分析語音信號，控制器或CPU用了有限的能源，且/或不用收到語音信號。

此外，本發明還提供了特性產生器，其可以在不同型式

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（ 4 ）

的LPC語音編碼中取出相同型式的特性資料以做爲識別之用。因此本發明在壓縮的語音資料中執行相同的識別運算(吻合和訓練)，其壓縮是由不同基於LPC型式的語音編碼器。

因此本發明具體來說就是提供了使用基於LPC語音編碼器資料來識別一個說出的話，而不須完整的重建的語音資料，在此描述語音編碼器爲基礎的識別器如何建立這個方法。此方法包含在語音識別器資料的每個區段中產生至少一個能量估計的步驟，並且使用相關能量估計在語音識別器資料中找尋字的邊界。如果一個字找到後，從語音編碼器資料中取出相對此字的LPC字參數，並且計算這個LPC字參數的識別特性。最後此識別特性吻合之前儲存其它字的識別特性之後，說出的字就被辨識出來。

此外，本發明還具體包含了從語音編碼器資料的剩餘資料中去作能量估計。此估計可以用很多方法來作。一個方式就是剩餘資料從編碼器資料重建，而此估計由剩餘資料的大小(norm)來做。另一個方式，一個基音增益值(pitch-gain)從語音編碼器資料中取出，這個值就作爲能量估計。另一個方式，基音增益值，落後值(lag values)，和剩餘資料從語音編碼器資料中取出。從剩餘資料中一個剩餘信號被產生，一個剩餘能量估計被產生。借由使用基音增益值和先前由落後值所定義的能量估計，從殘餘的非剩餘部份產生一個非剩餘能量估計。最後這兩個能量估計，(剩餘和非剩餘)，被結合。

再者，本發明還具體包含了語音編碼器資料能由下列任

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（ 5 ）

何一個語音編碼器產生：RPE-LTP全和半速率，QCELP 8和13 Kbps，EVRC，LD CELP，VSELP，CS ACELP，加強全速率語音編碼器和LPC 10。

本發明也具體提供了一個數位的蜂巢式電話，其包含一個行動電話的作業系統，一個基於LPC的語音編碼器，和一個語音編碼器爲基礎的語音識別器。識別器包含一個前端處理器來處理語音編碼器資料去決定何時一個字被說，並產生這個被說字的識別特性，而識別器從參考字庫中識別出此說出的字。

此外，本發明還具體包含，這個前端處理器包含一個能量估計器，一個LPC參數取出器，和一個識別特性產生器。此能量估計器用語音編碼器資料的殘餘資訊去估計一個語音信號的能量。此LPC參數取出器把語音編碼器資料的LPC參數取出。此識別特性產生器從LPC參數中去產生識別特性。

再者，本發明還具體提供：此前端處理器是選擇爲適當的多工語音編碼器的型式。

圖形的簡單描述

本發明可以配合下列的圖形所做的詳細描述來了解。

圖1是解釋有語音識別能力的蜂巢式電話技術的方塊圖；

圖2是解釋以LPC爲基礎的語音識別方法的流程圖；

圖3是一個說出字的能量的圖形解釋；

圖4是解釋壓縮語音資料架構的圖表；

圖5是方塊圖解釋了有語音編碼器爲基礎的語音識別器

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（6）

的蜂巢式電話，就是本發明的具體架構和運作；

圖6是解釋語音識別方法的流程圖，也是本發明的具體範圍；

圖7是從殘餘信號所估計的說出字的能量圖形解釋；

圖8是殘餘信號的圖形解釋，有用於了解本發明的運作；

圖9是解釋一個GSM解碼器的方塊圖；

圖10是從估計的殘餘信號所估計的說出字的能量的圖形解釋。

本發明的詳細描述

有很多不同型式的語音壓縮法則，最常見的就是以LPC爲基礎的方法。申請人已經了解到因爲大多數的語音識別法則運用到線性預測編碼分析以使得把語音信號參數化，壓縮語音信號的元素可以提供給語音識別器來有效降低語音識別器所必須執行的分析量。因此本發明是以語音編碼器爲基礎的語音識別器並把它做在蜂巢式行動電話的微控制器或CPU之中，其詳細的架構如下列圖5，6，及7所示。

線性預測分析：

下面是簡單解釋以LPC爲基礎的語音編碼器16的運作。通常在語音編碼的討論中包含了比這邊更詳細的線性預測編碼的描述，這些可以在下面這篇文章發現"Speech Coding: A Tutorial Review" by Andreas S. Spanias, Proceedings of the IEEE, Vol. 82, No. 10, October 1994, pp. 1541-1582。

語音編碼器16把語音信號分成一連串的區段，每個的長度是N，典型的針對語音信號是大約長20 msec。在每個

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）
經濟部中央標準局員工消費合作社印製
（請先閱讀背面之注意事項再填寫本頁）
裝
訂

五、發明說明（ 7 ）

區段，語音編碼器16執行了線性預測編碼(LPC)的分析。

線性預測編碼描述一個語音信號y(n)如下：

方程式1

$$\dot{y}(n)=a_1y(n-1)+a_2y(n-2)+\ldots+a_py(n-p)+\varepsilon(n)$$

其中a_1是LPC係數，而ε(n)是殘餘信號。典型的每個區段有p個LPC係數a_1而殘餘信號ε(n)的長度是N。LPC係數和殘餘物信號形成了每個區段(frame)的參數。語音編碼器典型的更進一步把殘餘信號ε(n)用至少基音(pitch)和增益(gain)值參數化。語音編碼器也能產生在已知的LPC語音編碼器技術上任何基於不同LPC型式的參數，(例如cepstrum係數，MEL cepstrum係數，線頻譜對(LSPs)，反射係數，對數區域比(LAR)係數等)，所有這些都能簡單的由LPC係數計算而來。

這些合成的值接著作編碼，因此產生了典型的語音壓縮區段，例如在參考圖4所示的區段52。語音壓縮區段52包括了編碼過且/或參數化的LPC係數a_1和編碼過的殘餘信號ε(n)。

語音編碼器爲基礎的編碼器識別器

參考圖5其解釋了在一個蜂巢式電話之中的一個語音編碼器爲基礎的語音識別器50。因爲這個蜂巢式電話和圖1所示的電話技術相似，相似的參考數字對應到相似的元件。參考圖6和7也有助於了解語音編碼器爲基礎的語音識別器50的運作。

圖5的蜂巢式電話包含了麥克風10，喇叭12，轉換單元

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（ 8 ）

14，語音編碼器16，作業系統18，RF界面單元19，和天線20。此外圖5所示的蜂巢式電話包含有語音編碼器爲基礎的語音識別器50其接收由語音編碼器16所產生的以LPC爲基礎的壓縮語音信號當作輸入。

根據本發明的具體範圍：此以語音編碼器爲基礎的語音識別器50做在叫作CPU 51之中，而作業系統18也作在其中。元件可以如圖是一個CPU或是一個微控制器。因爲語音識別器50並不分析語音信號，語音識別器50可以做在任何型式的微控制器或CPU包括那些只要有限的處理電源或那些不用收到語音信號的。

圖6以一般的型式解釋了語音編碼器爲基礎的語言識別器50如何對一個壓縮的區段作運作，例如是區段52。

如較早的技術，收到的信號在步驟58中，每個區段做能量的計算。然而在本發明之中，能量的計算是在步驟60中由語音編碼器資料來做而不是從取樣資料來做，且能量的估算不包含重建取樣資料。

申請者了解到殘餘的信號ε(n)可以用來做能量估計，因爲如在技術中所知，殘餘信號描述了發聲區域的空氣壓力，而LPC參數描述了發聲區的架構，因此和語音量無關。因此，殘餘信號和一個人如何大聲或安靜的說話有很高的相關性。

根據本發明的具體範圍，計算能量的一個方法就是去決定在殘餘信號上每個區段的能量，或每個子區段能量如果此區段再分成子區段的話。數學上可以寫成下式：

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（ 9 ）

方式2

$$\widetilde{E} = \sum_{n=1}^{M} \varepsilon(n)^2$$

其中$\widetilde{E}_i$是在第i個區段的能量，殘於信號ε(n)由語音編碼器資料重建，而數字M是在區段或子區段中取樣點的數目。

圖7解釋了由語音字"RICHARD"的重建殘餘信號而產生的估計能量信號。我們可以看到，圖7中估計的能量信號並不是圖3能量信號的複數。然而估計的能量信號跟先前技術的能量信號有很大的相關性。圖7信號的開端和尾端點分別寫成62和63也是大約分別在0.37 sec和0.85 sec。

其它從語音編碼器資料來估計能量的方法跟本發明有關聯，其中一些方法在下文來描述。

回到圖6，語音編碼器爲基礎的語音識別器50在步驟64中從估計的能量信號找尋字的邊界。如果找到了，語音識別器50可以用任何LPC參數的特性(例如平均且/或變異性)在字邊界會快速變得很尖銳的特性來使得字的邊界更精準。

如同步驟66的確認，如果一個字找到之後，識別器50從語音編碼器資料中取出LPC字參數(步驟68)。步驟68包含了把語音壓縮區段52所提供的編碼LPC參數做解壓縮和之後轉換成LPC係數。

識別器50然後從取出的LPC係數去計算它的識別特性(步驟70)。這些識別特性可以是任何以LPC爲基礎的參數，例如：cepstrum參數，MEL cepstrum參數，線頻譜對

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（ 10 ）

(LSPs)，反射係數，對數區域比參數(LAR)，等；所有這些參數都很容易從LPC係數計算。因此如果語音編碼器使用一種型式的LPC參數，而識別器50使用另一種型式的LPC參數，識別器50可以直接或經由LPC係數把它的型式轉成另一種。

最後識別器50使用在步驟70產生的識別特性去識別看看輸入信號是否是在它的參考字庫中的其中一個參考字或者是訓練一個新的字放入它的字庫中。因爲由識別器50所產生的識別特性可以跟先前技術所用的一樣，所以這個步驟是相等於先前技術的辨識/訓練步驟40。由Lawrence和Biing Hwang Juang所寫的書語音識別的基礎，Prentice-Hall, 1993，描述了適當的識別器50並在此合併以做爲參考。

在步驟60-70中把語音識別資料轉換成爲了識別/訓練的步驟所須的識別特性是相當重要的。有很多以LPC爲基礎的語音識別器，其中每個對語音信號都做不同的運算。步驟60-70可以配合任何型式的語音編碼器去產生相同的識別特性而不管語音識別器的特性。因此步驟60-70形成了識別/訓練步驟40的"頭尾端(front end)"。

本發明包括有一個語音編碼器爲基礎的語音識別器，其有一個多數的頭尾端，和一個單獨的識別/訓練單元，這對那些也是在不同型式的數位蜂巢式電話系統下運做的行動電話特別有用，其每個使用不同型式的語音編碼器。因爲有很多的頭尾端，本發明的語音識別器可以和很多型式

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（ 11 ）

的語音識別器運作。

使用在決定邊界的能量估計方法：

一些簡單的語音編碼器，例如在美國國防部標準1015 V.53的LPC10語音編碼器，把殘餘信號ε(n)只用基音和增益值來描述。圖8爲了參考所用，解釋了一個語音信號的標準殘餘信號，其有一串重覆大小幾乎都是一樣的峰值70。在峰值70之間的距離定義爲基音P而峰值的大小70定義爲增益G。非語音信號有一個增益值但沒有基音值。

因此殘餘信號中的區段或子區段的能量可以由增益值G來估算。在這個例子中，區段或子區段的能量不是由重建的殘餘信號ε(n)而是由壓縮的語音信號的殘餘信號ε(n)中區取出的一個增益參數值G來估計。

其它的語音編碼器，例如用於全球行動通訊系統，或是時域分割多工存取(TDMA)和碼域分割多工存取(CDMA)的數位蜂巢通訊系統中的語音編碼器，把現存的殘餘信號的區段或子區段和先前殘餘信號的區段或是子區段做相關(correlate)。現存殘餘信號的區段幾乎非常吻合於先的殘餘信號，當乘上一個基音增益PG後就得到所知的LAG值。然後語音編碼器會決定一個"剩餘信號(remnant signal)"，它和先前殘餘信號乘上基音增益PG及現存的殘餘信號都是不同的。然後現存的殘餘信號就由基音增益值PG和LAG值及剩餘信號來描述。

最近型式的語音編碼器，現存區段或子區段i的能量可以由剩餘信號和殘餘信號的非剩餘部份來估計：

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（ 12 ）

方程式3

$$\tilde{E}_i = \sqrt[m]{E_{LAG}{}^m + E_{rem}{}^m} \quad (m = 1 \quad or \quad 2)$$

$$E_{LAG} = PG \cdot \frac{1}{FL}\left\{(LAG \bmod 40)E_{i-\left\lceil \frac{LAG}{FL} \right\rceil} + (FL - LAG \bmod 40)E_{i-\left\lfloor \frac{LAG}{FL} \right\rfloor}\right\}$$

其中E_{rem}是剩餘信號的能量估計而E_{LAG}是殘餘的非剩餘能量，其由落後現區段或子區段LAG/FL個的區段和子區段的能量和基音增益值決定。前者可以由相對簡單的運算把剩餘信號重建或其它方法來產生。符號「⌈」和「⌊」分別代表（上限）"ceiling"和下限"floor"運算，因此找第m個根的運算可以不用做。

GSM語音編碼器的能量估計

圖9的參考圖是用來解釋GSM標準中在一個語音編碼器中的解碼部份。圖9類以於在第34頁裡的圖3.4；它是歐洲電信標準機關在1992年3月所出的I-ETS 300 036標準。解碼器的詳細說明可以在上述標準提供在此僅列下參考。爲了清楚起見，只有有助於了解本發明的能量和特性的計算的解碼器部份在下文中會提出。

圖9中粗線代表輸入資料和細線代表內部信號。輸入的資料包括了M_{cr}，x_{maxcr}，x_{mcr}，b_{cr}，N_{cr}和LAR_{cr}，這些都由I-ETS標準來定義。

圖9顯示出解碼器包含一個RPE解碼器80，一個長項預測器84，一個短項合成濾波器86，和一個加強放大器88。REP解碼器80收到M_{cr}，x_{maxcr}和x_{mcr}信號並產生剩餘

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）
經濟部中央標準局員工消費合作社印製
（請先閱讀背面之注意事項再填寫本頁）
裝
訂

五、發明說明（ 13 ）

信號e_r'。長項預測器84由剩餘信號e_r'使用b_{cr}和N_{cr}信號去產生殘餘信號d_r'。短項合成濾波器86由殘餘信號d_r'和以LAR_{cr}資料型式傳送的短項LPC參數來產生語音信號。

類似前面所描述的方法，一個計算能量的方式是做殘餘信號的第一和第二個絕大小(norm)如下：

方程式4

$$\tilde{E}=\sum_{n=0}^{39}\left|d_r'[n]\right|^m \quad (m=1或2)$$

其它能量計算是使用剩餘信號e_r'和長項預測器84中的內部資料值b_r'和N_r'。此外，預測器84包含一個參數解碼器90，一個延遲元件92，一個乘法器94和一個加法器96。解碼器90把輸入值b_{cr}和N_{cr}轉換成內部資料值b_r'和N_r'其中b_r'是一個乘法器，類似之前討論的基音增益PG，而N_r'是延遲量類似先前討論的LAG值。長項預測器82把信號d_r''和剩餘信號e_r'相加，其中信號d_r''是先前殘餘信號d_r'(I-Nr')經由延遲器92做延遲。此外在乘法器94中乘上一個b_r'量。

能量可以由方程式3來估計，其中N_r'和b_r'代替了LAG和PG值而FL設成40。此外剩餘信號E_{rem}的能量估計可以由下式計算：

方程式5

$$E_{rem}=\sum_{n=0}^{39}\left|e_r'[k]\right|^m$$

圖10是參考用來簡單顯示如何由前述的計算來估計能量。字的開端和尾端邊界記爲98和99各別發生在和先前

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）
經濟部中央標準局員工消費合作社印製
（請先閱讀背面之注意事項再填寫本頁）
裝
訂

五、發明說明（ 14 ）

技術所發生的位置相同。

其它從取出參數來估計能量的方法也用了前面的N_r'和b_r'，而FL設成40，而剩餘信號的能量估計由下式來估計：

方程式6

$$E_{rem}=|X_{maxcr}|^m$$

回到圖9，LPC字參數由短項合成濾波器86中所傳送的資料取出來，此濾波器含有一個LAR解碼器100，一個插入器102，一個反射係數決定單元104，和一個濾波器106。同樣地，單元100，102，和104把收到的LAR_{cr}資料轉換成反射係數r_r'，其中後者很容易轉換成LPC參數。

參考圖6如前所述，一旦LPC係數取出之後，他們就被轉換成識別器/訓練過程所須的識別特性(步驟70)。

本技術的所用技巧是很重要的，因爲當它們用於所有的由GSM數位蜂巢式通訓系統中，本發明可以運用於所有種類的數位蜂巢通訊系統和所有以LPC爲基礎的語音編碼器。對於每種語音編碼器，儲存於壓縮語音資料中的資訊型式必須被分析以決定如何使用它來做能量和特性的計算。壓縮語音資料由各個語音編碼器的標準定義來描述。

下面的表格列出了現存的蜂巢通訊細系統，它們用的語音編碼器，還有語音編碼器和/或系統的標準定義。

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

五、發明說明（ 15 ）

Digital Cellula Communication System	LPC-based Vocoder	Standard
GSM	RPE-LTP full rate	I-ETS 300 036 6.1
	RPE-LTP half rate	I-ETS 300 581-2 ver.4
CDMA	QCELP 8 Kbps, 13 Kbps	IS 96 A
	EVRC	IS 127
	LD CELP	ITU G.728
TDMA	VSELP	IS 54 B
PHS, PCS	CS ACELP	ITU G.729
PCS-TDMA	Enhanced Full Rate Vocoder	IS 641
PDC(in Japan)	VSELP	RCR STD 27

如熟習此項技術之人士所明瞭，本發明所用的技術不受上面列出和描述的規格所限制。本發明的範圍由下面的申請項目來定：

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

四、中文發明摘要（發明之名稱：語音編碼器爲基礎的語音識別器）

一種語音編碼器爲基礎的語音識別器，其識別一個說出字是以線性預測編碼(LPC)爲基礎的語音編碼器資料而不用完全重建語音資料。此識別器在語音編碼器資料的每個架構(frame)產生致少一個能量估計，並且利用這個相關的估計來找尋語音邊碼器資料中的字的邊界。如果找到一個字後，線性預測編碼(LPC)字的參數將會從相對於這個字的語音編碼器資料中抓出來，並計算從LPC字的參數的識別特性。最後這個識別特性跟先前其它字儲存的識別特性吻合，因此識別出這個說出的字。

英文發明摘要（發明之名稱："A VOCODER-BASED VOICE RECOGNIZER"）

A vocoder based voice recognizer recognizes a spoken word using linear prediction coding (LPC) based vocoder data without completely reconstructing the voice data. The recognizer generates at least one energy estimate per frame of the vocoder data and searches for word boundaries in the vocoder data using the associated energy estimates. If a word is found, the LPC word parameters are extracted from the vocoder data associated with the word and recognition features are calculated from the extracted LPC word parameters. Finally, the recognition features are matched with previously stored recognition features of other words, thereby recognizing the spoken word.

（請先閱讀背面之注意事項再填寫本頁各欄）
裝
訂
線
經濟部中央標準局員工消費合作社印製
本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

六、申請專利範圍

1. 一種使用以LPC爲基礎的語音編碼器資料而不用完整重建該語音資料來識別口語字的方法，該語音編碼器資料形成由一連串的訊框，該方法包含下列步驟：

每個語音編碼器資料訊框產生至少一個能量估計值；

使用相關的能量估計值尋找語音編碼器資料中的字的邊界；

如果找到一個字，從與此字相關的語音編碼資料中取出LPC字參數；

從該等取出的LPC字參數去計算識別特性；

將該識別特性和先前儲存的其他字的識別特性做比對，以識別出該口語字。

2. 一種準備使用線性預測編碼(LPC)爲基礎的語音編碼器資料而不用完整重建的語音資料去識別一個口語字的方法，該語音編碼器資料形成由一連串的訊框，該方法包含下列步驟：

每個語音編碼器資料訊框產生至少一個能量估計值；

使用相關的能量估計值尋找語音編碼器資料中的字的邊界；

如果找到一個字，從此字的相對語音編碼資料中取出LPC字參數從相對取出的LPC字參數去計算識別特性。

3. 如申請專利範圍第2項的方法，其中該產生步驟包括從該語音編碼器資料中所發現的殘餘信號估計能量的步驟。

4. 如申請專利範圍第3項的方法，其中該估計步驟包括由

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

六、申請專利範圍

該語音編碼器資料重建殘餘信號和產生該殘餘信號範數(norm)的步驟。

5. 如申請專利範圍第3項的方法，其中該估計步驟包括該語音編碼器資料取出基音增益值，並使用所取出的說話的基音增益值去當成說出話的能量估計。

6. 如申請專利範圍第3項的方法，其中該產生步驟包含下面步驟：

從所說話的語音編碼器資料中取出基音增益值，遲滯值(lag values)，和剩餘資料；

從該剩餘資料重建一個剩餘信號；

產生該剩餘信號的能量估計；

利用該基音增益值和該遲滯值所定義的先前能量估計來產生該殘餘信號的非剩餘部份的能量估計；以及

結合該剩餘和非剩餘的能量估計。

7. 如申請專利範圍第1項的方法，其中該語音編碼器資料型式是由下列任一個語音編碼器：RPE-LTP全速率和半速率，QCELP 8和13 Kbps，EVRC，LD CELP，VSELP，CS ACELP，加強全速率語音編碼器和LPC 10所產生的一種型式。

8. 如申請專利範圍第2項的方法，其中該語音編碼器資料型式是由下列任一個語音編碼器：RPE-LTP全速率和半速率，QCELP 8和13 Kbps，EVRC，LD CELP，VSELP，CS ACELP，加強全速率語音編碼器和LPC 10所產生的一種型式。

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

六、申請專利範圍

9. 一種LPC爲基礎的語音編碼器資料，當成一個語音識別系統的輸入。

10. 一種數位蜂巢式電話，包括：

一個行動電話的作業系統；

一個語音編碼器，其使用至少線性預性編碼(LPC)去壓縮語音信號，藉此產生語音編碼器資料；以及

一個語音編碼器爲基礎的語音識別器，包括：

一個前端處理器，其處理該語音編碼器資料，以決定何時說出一個字並產生所說話的識別特性；以及

一個識別器，其辨識所說的字至少是一組參考字之中的一個。

11. 如申請專利範圍第10項的數位蜂巢式電話，其中該前端處理器包括：

一個能量估計器，其使用該語音編碼器資料的殘餘資訊去估計一個語音信號的能量；

一個LPC參數取出器，其從該語音編碼器資料中去取出LPC參數；

一個識別特性產生器，其從該LPC參數中去產生說出話的識別特性。

12. 如申請專利範圍第10項的蜂巢式電話，其中該前端處理器是選擇爲多工運作的語音編碼器型式。

13. 如申請專利範圍第10項的蜂巢式電話，其中該語音編碼器是下列任一個語音編碼器：RPE-LTP全速率和半速率，QCELP 8和13 Kbps，EVRC，LD CELP，VSELP，CS

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

六、申請專利範圍

ACELP，加強全速率語音編碼器和LPC 10。

14. 一種語音編碼器爲基礎的語音識別器，其與LPC爲基礎的語音編碼器所產生的資料運作，該語音識別器包含：

一個前端處理器，其處理該語音編碼器資料以決定何時一個字被說並產生所說話的識別特性；

一個識別器，其辨識所說的字至少是一組參考字之中的一個。

15. 如申請專利範圍第14項的語音識別器，其中該前端處理器包括：

一個能量估計器，其使用說出字的語音編碼器資料的殘餘資訊去估計一個語音信號的能量；

一個LPC參數取出器，其從說出字的語音編碼器資料中去取出LPC參數；以及

一個識別特性產生器，其從說出話的LPC參數中去產生說出話的識別特性。

16. 如申請專利範圍第15項的語音識別器，其中該能量估計器包括一個由說出字的語音編碼器資料之殘餘資料來估計能量的殘餘能量估計器。

17. 如申請專利範圍第16項的語音識別器，其中該能量估計器包括一個殘餘信號重建器，其從該語音編碼器資料去重建殘餘資料；一範數(norm)產生器，其產生該殘餘資料的範數(norm)，藉此產生該能量估計。

18. 如申請專利範圍第16項的語音識別器，其中該殘餘能量估計器包括一個取出器，其從該語音編碼器資料取出基

（請先閱讀背面之注意事項再填寫本頁）
裝
訂
經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

六、申請專利範圍

音增益值，藉此產生該能量估計。

19. 根據申請專利範圍第16項的語音識別器，其中該殘餘能量估計器包括：

一個取出器，其從該語音編碼資料中　取出基音增益值，遲滯值(lag values)，和剩餘信號；

一個重建器，其從所說話的剩餘信號重建一個剩餘信號；

一個剩餘能量估計器，從所說話的剩餘信號產生一個能量估計；

一個非剩餘能量估計器，其借由使用所說的基音增益值和先前由所說話的遲滯值所定義的能量估計來對所說話的殘餘信號的非剩餘部份產生一個能量估計；

一個結合器，其把所說話的剩餘和非剩餘的能量估計結合起來因此去產生對說出的能量估計。

20. 如申請專利範圍第14項的語音識別器，所提的語音編碼器是下列語音編碼器：(RPE-LTP全速率和半速率，QCELP 8和13 Kbps，EVRC，LD-CELP，VSELP，CS ACELP，加強全速率語音編碼器和LPC 10)其中一種。

（請先閱讀背面之注意事項再填寫本頁）
裝
訂

經濟部中央標準局員工消費合作社印製

本紙張尺度適用中國國家標準（CNS）A4規格（210×297公釐）

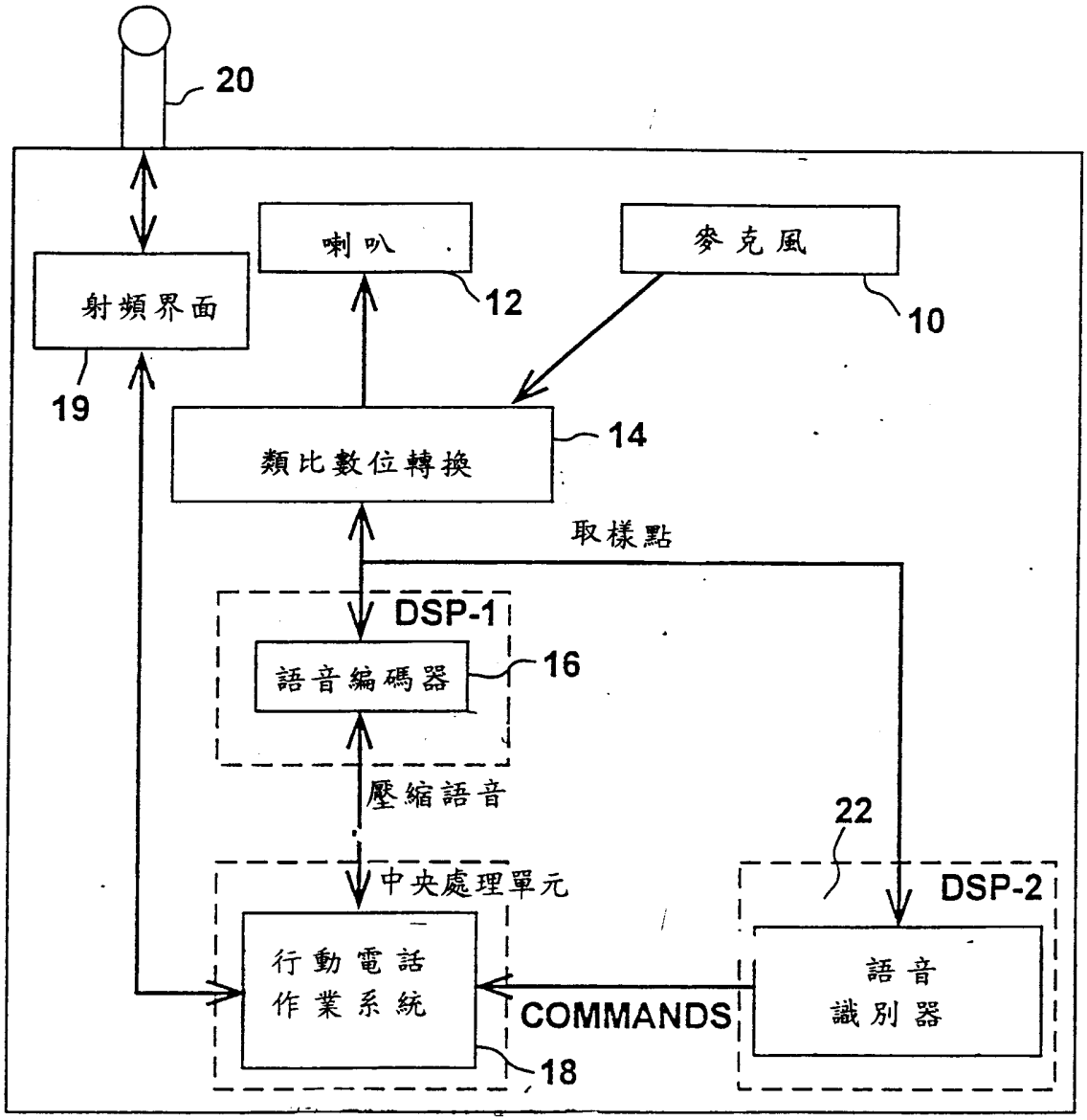

20
喇叭
麥克風
射頻界面
12
10
19
類比數位轉換
14
取樣點
DSP-1
語音編碼器
16
壓縮語音
22
中央處理單元
DSP-2
行動電話
作業系統
COMMANDS
語音
識別器
18

圖1

先前技術

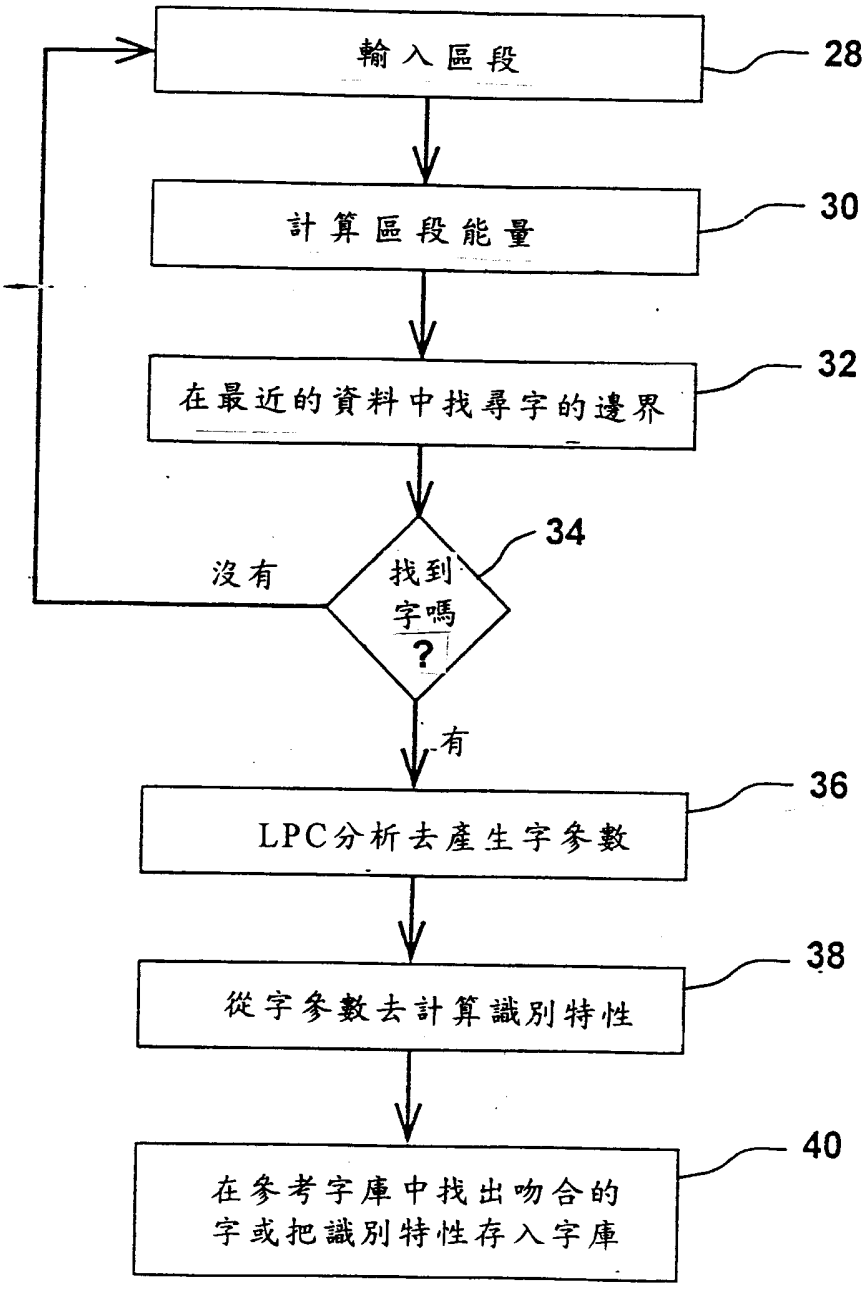

輸入區段
28
計算區段能量
30
在最近的資料中找尋字的邊界
32
找到字嗎?
34
沒有
有
LPC分析去產生字參數
36
從字參數去計算識別特性
38
在參考字庫中找出吻合的字或把識別特性存入字庫
40

圖2

先前技術

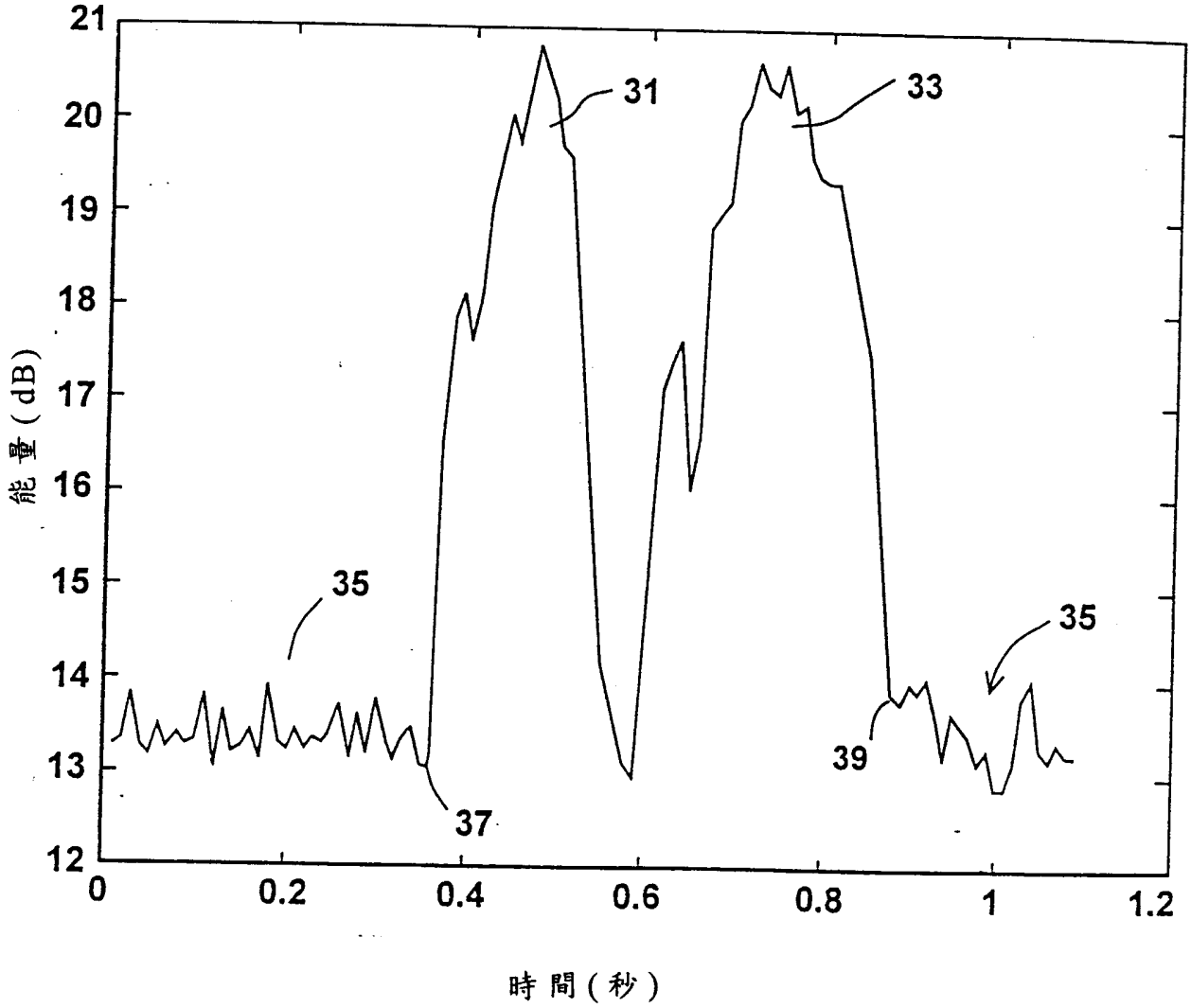

21
20
19
18
17
16
15
14
13
12
能量(dB)
0
0.2
0.4
0.6
0.8
1
1.2
時間(秒)
31
33
35
35
37
39

圖3

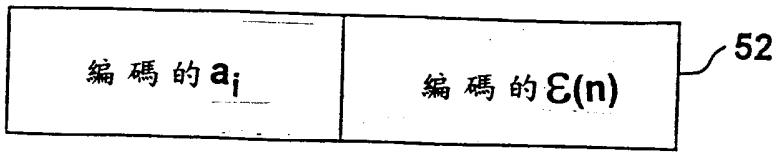

編碼的a_i
編碼的$\varepsilon(n)$
52

圖4

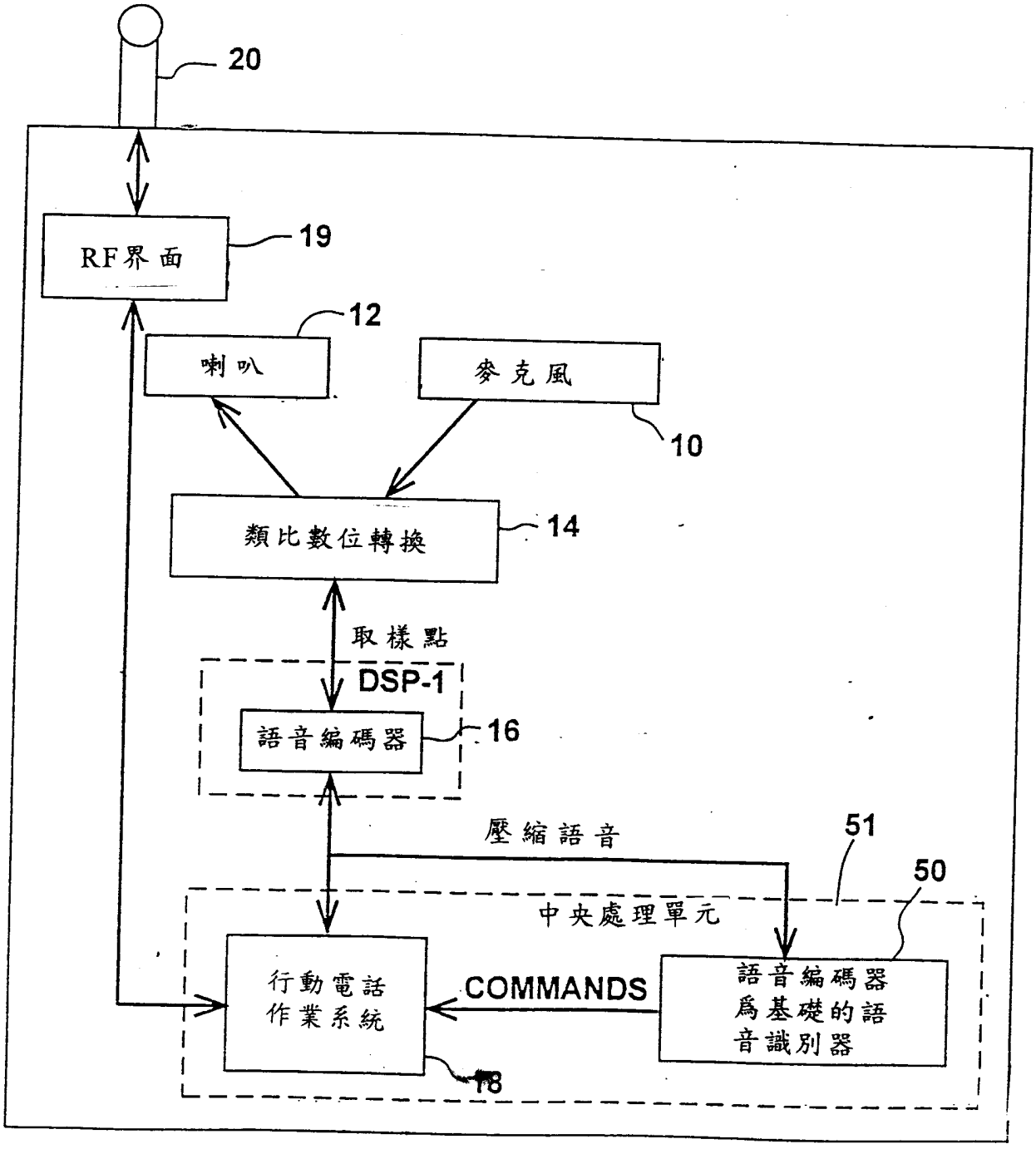

20
RF界面
19
喇叭
12
麥克風
10
類比數位轉換
14
取樣點
DSP-1
語音編碼器
16
壓縮語音
51
50
中央處理單元
行動電話
作業系統
COMMANDS
語音編碼器
爲基礎的語
音識別器
18

圖5

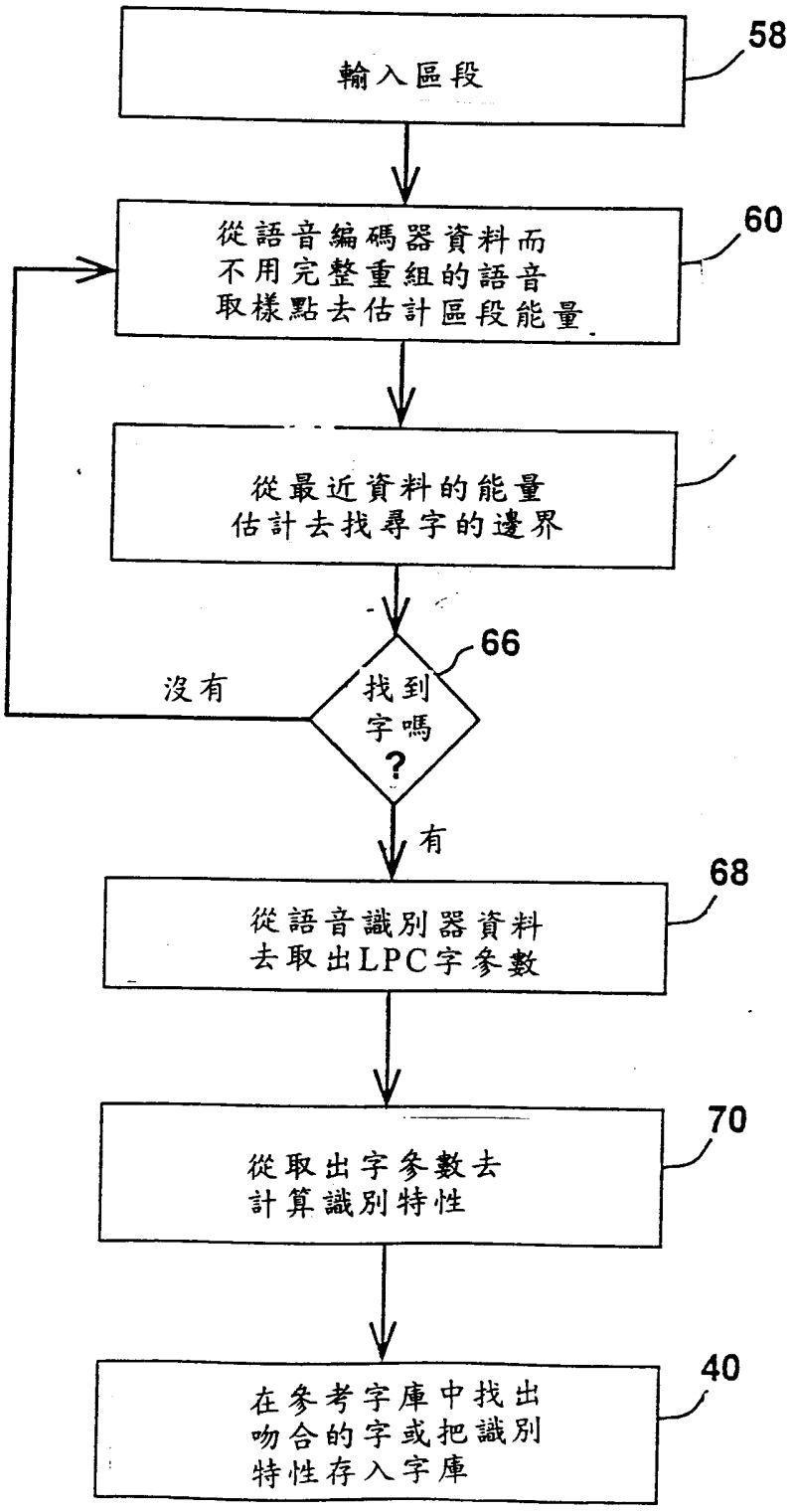

輸入區段
58
從語音編碼器資料而
不用完整重組的語音
取樣點去估計區段能量
60
從最近資料的能量
估計去找尋字的邊界
找到
字嗎
?
66
沒有
有
從語音識別器資料
去取出LPC字參數
68
從取出字參數去
計算識別特性
70
在參考字庫中找出
吻合的字或把識別
特性存入字庫
40

圖6

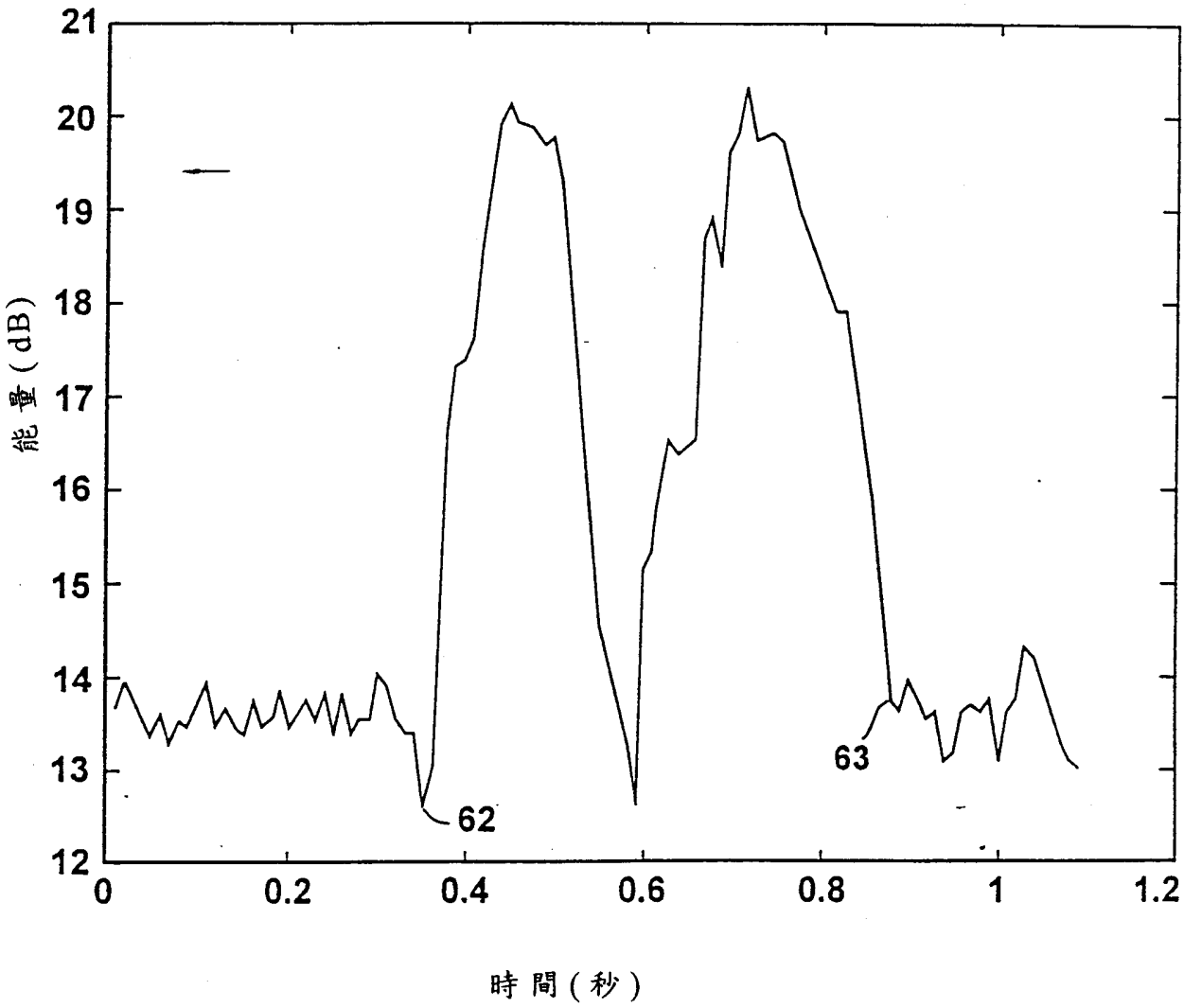


圖7

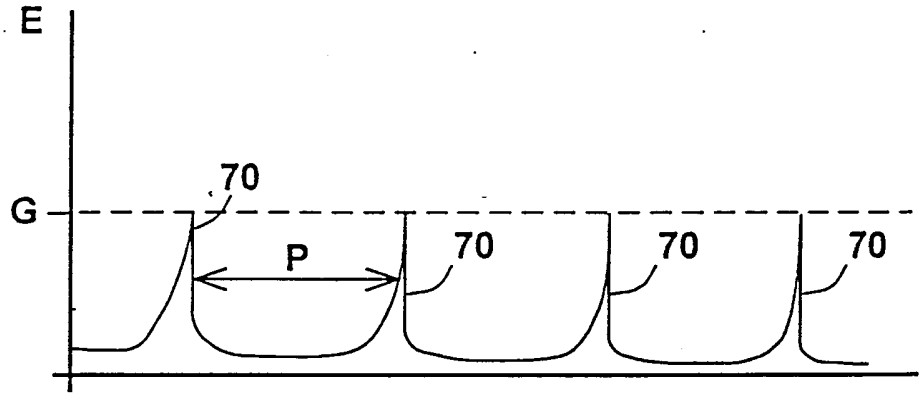


圖8

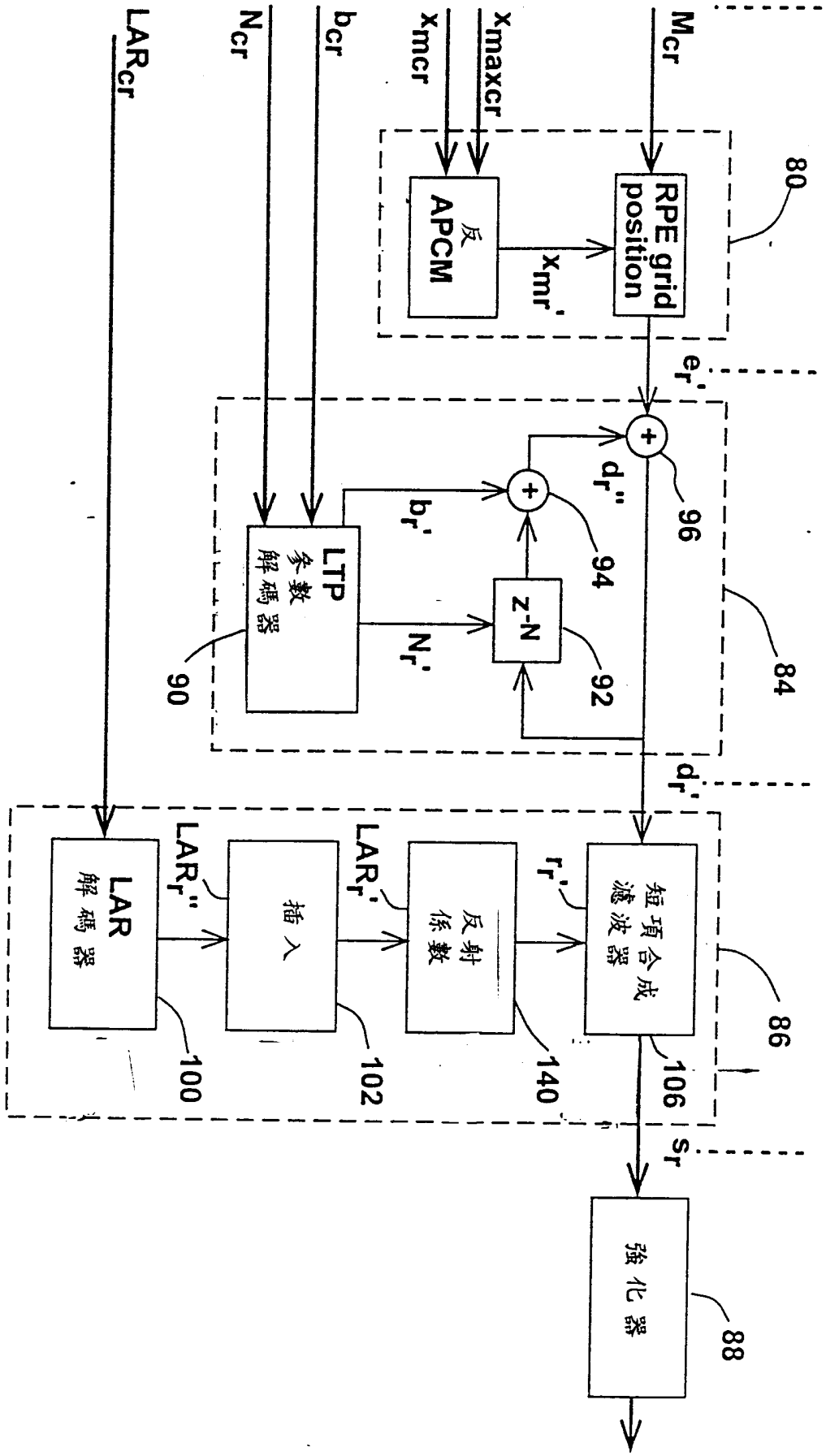

M_{cr}
x_{maxcr}
x_{mcr}
b_{cr}
N_{cr}
LAR_{cr}
RPE grid position
反 APCM
x_{mr}'
80
e_r'
d_r''
96
94
b_r'
LTP 參數 解碼器
N_r'
z-N
92
90
84
d_r'
短項合成濾波器
r_r'
反射係數
LAR_r'
插入
LAR_r''
LAR 解碼器
100
102
140
106
86
s_r
強化器
88

圖 9

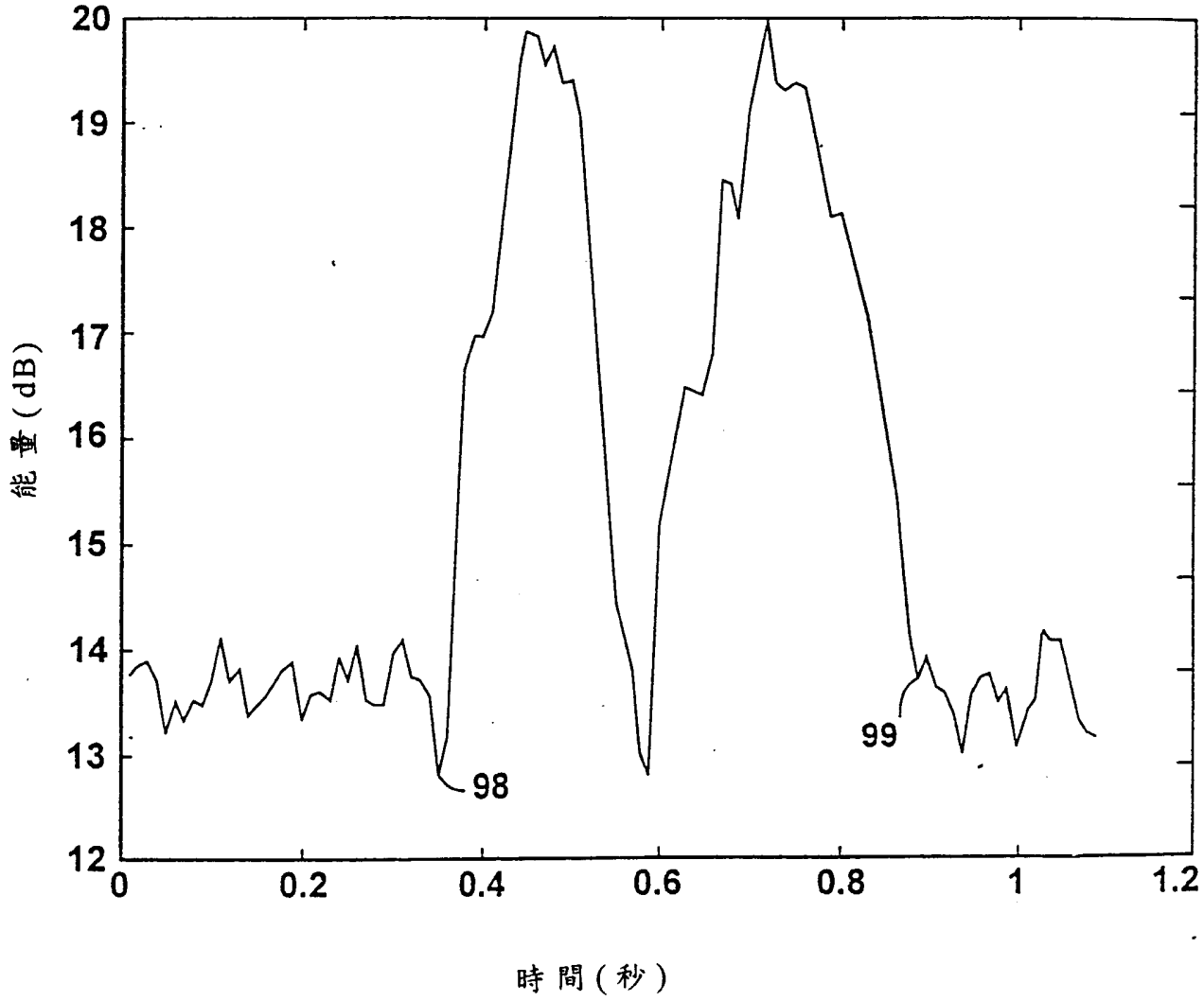


圖10