

公告本

申請日期	88. 7. 23
案 號	38104608
類 別	4106 9/2

436759

A4
C4

(以上各欄由本局填註)

發明專利說明書

一、發明名稱	中 文	用於多雜音狀況之語音檢測系統
	英 文	SPEECH DETECTION SYSTEM FOR NOISY CONDITIONS
二、發明人	姓 名	(1)趙 翊 (2)珍-克勞德·姜奎爾
	國 籍	(1)中國 (2)法國
	住、居所	(1)美國加州高雷塔市蓋里大道715號#203 (2)美國加州聖大芭芭拉市努埃塞斯道4543號
三、申請人	姓 名 (名稱)	日商·松下電器產業股份有限公司
	國 籍	日 本
	住、居所 (事務所)	日本國大阪府門真市大字門真1006番地
	代 表 人 姓 名	森下洋一

裝

訂

線

436759

(由本局填寫)

承辦人代碼：
大類：
I P C分類：

A6
B6

本案已向：

美 國 (地 區) 申請專利，申請日期： 案號： ， ☒有 ☐無主張優先權
1998,3,24 09/047,276

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

五、發明說明(1)

本發明之背景及概述：

本發明大致上有關於語音處理及語音識別系統。更特別地有關於一探測系統，用以探測一輸入信號內語音之起始及終結。

用於語音識別及用於其他目的之自動語音處理係時下電腦能實施之最具挑釁之任務之一。例如，語音識別引用一高度複雜之模式匹配技術，它可以是對變化性非常敏感。在消費者之用途中，識別系統需要能應付不同語言之多種範圍，並要在廣闊變化之環境狀況下操作。體外信號和雜音之呈現可大大地降低識別品質及語音處理性能。

最自動之語音識別系統之聲音之第一模型圖樣來工作，以及隨後使用那些圖形來識別音位，字母和最後文字。為了精確地識別，要排除任何在實際語音之前或跟隨在後之體外聲音(雜音)係至為重要。有甚多習知技術試圖探測語音之開始及終結，雖然如此，但仍被視為有待改良之空間。

本發明劃分進來之信號成為頻率帶，各帶代表一不同範圍之頻率。各帶內之短期用能量係隨後與多個臨限比較，以及比較之結果係用來驅動一狀態機，當帶之至少一個之限制之帶信號能量係高於其相關臨限之至少一個時，它自一“無語音”狀態轉換至一“語音出現”狀態。當帶之至少一個之限制之帶信號能量係低於其相關臨限之至少一個時，此狀態機同樣地自一“語音出現”狀態轉換至“無語音”狀態。此系統亦包括一局部語音探測機構以語音之實際開始

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

頁

五、發明說明(2)

之前一假定之“無聲分段”為基礎。

一梯級頻率數據結構累積有關此頻率帶內能量之平均值及變化之長期數據，以及此一資料係用來調整適應性臨限。此頻率帶係根據噪音特性而分配。此梯級頻率表示法分別地在語音信號，無聲及噪音之間供給強烈之區分。在語音信號本身內，此無聲部分(僅具有背景噪音)典型地支配，同時它係在梯級頻率上強烈地被反映。背景雜音，係比較地正常者，當梯級頻率上一明顯尖光時即顯現。

此系統係極適應於在雜音狀況中探測語音，同時它將探測語音之開始及終結兩者，以及應付語音之開始中可能經過捨位而已丟失之情勢。

為了對本發明，其目的和優點之更完整之瞭解，可能必須以下列說明及附圖為基準。

附圖之簡要說明：

第1圖係在一目前較佳之2-帶實施例中語音探測系統之一方塊圖；

第2圖係使用以調整此適應性臨限之系統之詳細方塊圖；

第3圖係局部語音探測系統之詳細方塊圖；

第4圖說明本發明之語音信號狀態機；

第5圖係一線圖，說明一範例性梯級頻率，有用於對本發明之瞭解；

第6圖係一波形圖，說明多個臨限使用於為語音探測之比較信號能量上；

(請先閱讀背面之注意事項再填寫本頁)

訂

五、發明說明(3)

第7圖係一波形圖，說明使用以避免強烈雜音脈衝之探測失誤之開始語音延遲之探測機構；

第8圖係一波形圖，說明使用以提供脈動於連續語音裡面之語音之終結延遲之探測機構；

第9A圖係一波形圖，說明局部語音探測機構之一觀點；

第9B圖係一波形圖，說明局部語音探測機構之另一觀點；

第10圖係波形圖之收集，說明此多帶臨限分析係如何組合以選擇相當於一語音呈現狀態之最終範圍；

第11圖係一波形圖，說明在強烈雜音之出現中此S臨限之使用；以及

第12圖說明此適當性臨限當其適用於背景雜音位準時之性能。

較佳實施例之說明：

本發明分離輸入信號成為多個信號線路，各代表一不同頻率帶。第1圖說明本發明引用兩個帶之一實施例，一帶相當於輸入信號之整個頻譜，以及另一帶相當於整個頻譜之一高頻率支組。此說明之實施例係特別地適用以檢測有一低信/噪比之輸入信號，諸如對在一移動中之汽車內或一喧嘩辦公室環境內所發現之狀況。在這些公用環境中，甚多雜音能量係經低於2000Hz而分佈。

當一兩帶系統係在此說明時，本發明可隨時延伸至其他多帶配置。一般而言，涵蓋不同頻率範圍之個別帶，經

(請先閱讀背面之注意事項再填寫本頁)

訂

五、發明說明(4)

設計以自此雜音隔離此信號。時下之實現係數位。當然，類比實現亦可以使用本文所含之說明來達成。

參看第1圖，此輸入信號含一可能之語音信號以及雜音經已在20處表示。此輸入信號係已數位化並通過一漢明窗口作處理以次分此輸入信號數據成為幀。本較佳實施例引用一預先界定之採樣率之10ms幀(在此一情況為8000 Hz)，產生每幀80數位樣本。此說明之系統係經設計以操作在有一頻率伸展於300 Hz至3,400 Hz之範圍內之輸入信號上。因此一兩倍於此上頻限之採樣率($2 \times 4000 = 8000$)業已選擇。如果一不同頻率內容係在輸入信號之資料輸送部分內被發現時，那麼此採樣率和頻帶可以適當地調整。

漢明窗口22之輸出係代表此輸入信號之數位採樣之順序(語音加雜音)，並經配置成為預定大小之幀。這時幀隨後係進給至快速傅里葉變換轉換器24，它變換輸入信號數據自時間領域成為頻率領域。在此一點此信號係分裂成為多個線路，一第一線路在26處以及第二線路在28處。此第一線路相當於含所有輸入信號之頻率之一頻率帶，而此第二線路28相當於輸入信號之全頻譜之高頻率支組。因為頻率領域內含係由數位數據所表示，故頻率帶分裂係分別由求和模塊30和32來完成。

應予說明者即此求和模塊30概括範圍10至108上面之譜分量，然而求和模塊32則概括範圍64至108，以此一方式，此求和模塊30選擇輸入信號中所有頻率帶，而模塊32僅選擇高頻率帶。在此一情況中，模塊32抽取由模塊30所

(請先閱讀背面之注意事項再填寫本頁)

訂

線

五、發明說明(5)

選擇之帶之一支組。此係目前用以探測在一移動車輛或喧嘩辦公室內所共同發現之一類嘈雜輸入信號內之語音內涵之較佳配置。其他嘈雜情況可指定其他分裂之頻率帶配置。例如，多個信號線路可以構形以隨要求而涵蓋個別，永不重疊之頻率帶以及局部重疊之頻率帶。

此求和模塊30和32一次一幀地概括此頻率分量，因此所產生之模塊30及32之輸出代表此信號內之有限之頻帶，短期能量。如果須要時，此一原始數據可以傳送通過一修勻濾波器，諸如濾波器34和36。在本較佳實施例中，一三抽頭平均數係在兩者位置中用作此修勻濾波器。

一如將在下文中更詳細解釋者，語音探測係以多個臨限比較此多個受限頻率帶，短期能量為根據。這些臨限係根據與預測語音沈靜部分有關聯(當此系統係有效但在揚聲器發音之前假設要予呈現)之能量之長期平均值及變化而適應性地更新。此實施在產生此適應性臨限上使用一梯級頻率數據結構。第1圖內複合方塊38和40代表分別供信號線路26和28用之適應性臨限更新模塊。這些模塊之進一步細節將與第2圖和若干有關聯波形圖相關聯地提出。

雖然分開之信號線路係經保持在快速傅里葉變換模塊24之下游，但通過此適應性臨限更新模塊38和40，此語音是否係出現或未呈現於輸入信號內之最後決定之兩者信號線路一起作考慮而產生。因此，此語音狀態探測模塊42和其相關聯之局部語音探測模塊44思考自兩者線路26和28之信號能量數據。此語音狀態模塊42實施一狀態機，其細節

(請先閱讀背面之注意事項再填寫本頁)

表

訂

五、發明說明(6)

係在第4圖內進一步地說明。此局部語音探測模塊係詳細地顯示於第3圖內。

現在參看第2圖，此適應性臨限更新模塊38將予以解釋。本較實現為每一能量使用第三種不同臨限。因此，在所說明之實施例中總共有六個臨限。每一臨限之目的將藉思考波形圖和相關討論而清楚地解釋。為每一能量帶此三個臨限係經識別：Threshold，WThreshold以及SThreshold。第一所列臨限Threshold係一基本臨限用來探測語音之開始。此WThreshold係一弱臨限用以探測語音之終結。此SThreshold係一強臨限，用以評估語音探測決定之效力。此等臨限係更正式地界定如下：

Threshold=雜音_位準加上偏移量

WThreshold=雜音_位準加上偏移量*R1；(R1= 0.2..1, 0.5係目前較適當者)

SThreshold=雜音_位準加上偏移量*R2；(R2=2..4, 2係目前較適當者)

此處：

雜音_位準係長期平均值，亦即，在此梯級頻率中所有已過去之輸入能量之最大值。

偏移量=雜音_位準*R3加變化R4；(R3=0.2..1, 0.5係目前較適當者；R4=2..4, 4係目前較適當者)。

變化係短期變化，亦即M個過去輸入幀之變化。

第6圖說明三個臨限置於一示範性信號之上面之關係。應予說明者即SThreshold係較高於Threshold，同時

五、發明說明(7)

WThreshold係大致上較Threshold為低。這些臨限係以雜音位準為根據使用一梯級頻率數據結構以測定被含於輸入信號之預先設定無語音部分內之所有已過去之輸入能量之最大值。第5圖說明一示範性梯級頻率置於說明一示範性雜音位準之波形上面。此梯級頻率作為“計數”而記錄預置無語音部分含一預定雜音位準能量之次數。此梯級頻率因而測繪計數之數目(在Y軸上)作為此能量位準之功能(在X軸上)。應說明者，即第5圖內所說明之範圍中，此最普遍(最高計數)雜音位準能量有Ea之能量值。此值Ea會相當於一預定雜音位準能量。

記錄於梯級頻率(第5圖)內之雜音位準能量數據係自輸入信號之預置無語音部分抽取。就此有關言，吾人假定該供應此輸入信號之聲頻頻道係有效並在實際語音開始之前發送數據至語音探測系統。因此，在此一預置無語音區，此系統係有效地抽樣周圍雜音位準本身之能量特性。

本較佳實現使用一固定大小之梯級頻率以減少電腦記憶體需求。梯級頻率數據結構之適當構形代表為精確預測之理想(意指小梯級頻步進)和廣濶動態範圍(意指大梯級頻步進)之間之一平衡。要從事於精確預估(小梯級頻步進)和廣濶動態範圍(大梯級頻步進)之間之爭執，現用系統根據實際操作情況而適應性地調整梯級頻率進。被應用於調整梯頻步進大小上之算法係說明於下列偽代碼中，該處M係步進大小(代表梯級頻之各步進中能量值之範圍)。

用於適應性梯級頻步進之偽代碼：

五、發明說明(8)

起始階段之後：

計算緩衝器裡面已過幀之平均值

$M = \text{先前平均值之十分之一}$

如果($M < \text{MIN_梯級頻_步進}$)

$M = \text{MIN_梯級頻_步進}$

終結

在上述偽代碼中，應予說明者，即此梯級頻步進M係根據在開始時假定無語音部分之平均值而採用，它係在起始階段中被置於緩衝器內。該平均值係經假定來顯示實際背景雜音狀況。應予說明者即此梯級頻步進係受限於MIN_HISTOGRAM_STEP作為下界限。此一梯級頻步進係在此一刻之後被固定。

此梯級頻係藉為每一幀插入一新值而更新。要適應於緩慢改變之背景雜音，一忽略因數(在現時實現中為0.90)係經引進供每10幀用

用以更新此梯級頻之偽代碼

如果(值 < 梯級頻_大小 * M)

{

//藉忽略因數來更新梯級頻

如果(幀_in_梯級頻 % 10 == 0)

{

for(1=0; 1 < 梯級頻_大小; 1++)

梯級頻[1] *= 梯級頻_忽略_因數

//藉堪入新值更新梯級頻

(請先閱讀背面之注意事項再填寫本頁)

表

訂

後

五、發明說明(9)

梯級頻 $[(值+M/2)/M]+1$;

梯級頻 $[(值-M/2)/M]+1$;

}

現在參看第2圖，適應性臨限更新機構之基本方塊圖係經明。此一方塊圖說明由模塊38和40所實施之操作(第1圖)。此短期(電流數據)能量係經貯存於更新緩衝器50內，並亦係使用於模塊52內以更新梯級頻數據結構一如前述。

此更新緩衝器隨後係由模塊54檢查，它計算貯存於緩衝器50內數據之過去之幀上面之變化。

此際，模塊56識別梯級頻內最大能量值(例如，第5圖內之值Ea)，並供應此值至臨限更新模塊58。此臨限更新模塊使用自模塊54之最大能量值和靜態數據(變化)以修正主臨限Threshold。一如前文所討論者，Threshold係相當於雜音位準加上一預定偏移量。此一偏移量係以由梯級頻內最大值所測定及由模塊54所供應之變化上之雜音位準為根據。其餘之臨限，WThreshole和SThreshold，係依照上文所宣佈之等式自THreshold所計算。

在正常操作中，此臨限適應性地調整，大致上依循預置語音區內雜音位準之軌跡。第12圖說明此一理念。在第12圖中此預置語音區係經顯示於100處，以及語音之開始係大致上顯示於200處。在此一波形上此THreshold位準業已被置於最上面。應予說明者，即此一臨限之位準追蹤預置語音區內雜音位準加上一偏移量。因此，以Threshold(以

(請先閱讀背面之注意事項再填寫本頁)

長

訂

五、發明說明(10)

及此SThreshold和此WThreshold)之可應用於一指定語音分段者將是那些在語音之開始之前立刻地有效之臨限。

回頭參看第1圖，此語音狀態探測和局部語音探測模塊42和44現在將予以說明。取代以一幀之數據為根據而形成語音出現/無語音決定者，此決定係以現時幀加上緊隨現時幀之少許幀為根據而形成。就有關語音開始探測之言，緊隨現時幀之附加幀之考慮(向前看)避免了在一短而強之雜音脈之呈現中，諸如一電脈衝之假探測。就有關語音之終結之探測言，幀之向前看防止在一不同之連續語音信號中之停止或暫短之無聲而提供語音之終結之假探測。此一延遲之決定或向前看之策略係藉在此更新緩衝器50中緩衝此數據而實現(第2圖)，並應用由下列偽代碼所說明之程序：

開始_語音測試：

開始延遲之決定=偽

Loop M跟隨之幀(M=3;30ms)

如果要就是(Energy_All)抑或(Energy_HPF)>Threshold

然後開始延遲之決定=真

語音之終結測試：

終結延遲之決定=偽

Loop M跟隨之幀(M=30;30ms)

如果兩者(Energy_All)和(Energy_HPF)>Threshold

然後終結延遲之決定=真

End of Loop

(請先閱讀背面之注意事項再填寫本頁)

衣

訂

五、發明說明 (11)

參看第7圖，它說明在開始語音測試中之30ms之延遲如何避免臨限以上一雜音尖峰脈衝110之偽探測。同時亦參看第8圖，它說明300 ms延遲此終結語音測試如何防止在語音信號中之一暫短停止120不會觸發語音之終結狀態。

上述偽代碼設定兩個標誌，開始延遲之決定標誌和終結延遲之決定標誌。這些標誌係由第4圖內所示之語音信號狀態機所使用。應予說明者，即此語音之開始使用30ms延遲，相當於3個幀($M=3$)，由於短雜音尖峰脈衝，故此係正常地適用以篩除假探測。此終結則使用一長延遲，在300ms之範圍，此範圍業經被發現足夠以應付連接之語音裡面正常暫停之發生。此300ms延遲相當於30個幀($N=30ms$)。要避免由於語音信號之裁剪或截短之錯誤，此數據可以為開始及終結兩者之探測之語音部分為根據而填充以額外之幀。

語音之開始探測算法假定至少一指定長度之一預置無語音部分之存在。實際上，亦曾有當此一假定可能不正確之時刻，諸如由於信號漏失或電路轉換假信號而使輸入信號係被裁剪之情況中，由是而縮短或消除此假定之“無語音分段”。當此情況發生時，此臨限可能是不正確地被採用，由於此臨限係以雜音位準能量為依據，假定地以聲音信號未出現。此外，當此輸入信號係被裁剪至一點，即沒有語音分段時，此語音探測系統可能不足以識別此輸入信號為含有語音，可能導致輸入階段內語音之丟失，那使得

五、發明說明(12)

後續之語音處理無用。

要避免此局部語音狀態，一捨選策略係經引用如第3圖內所示。第3圖說明由局部語音探測模塊44(第1圖)所引用之機構。此局部語音探測機構藉監控此臨限(Threshold)以測定是否有一突移在適應性臨限位準內。此轉移模塊60藉首先累積可指示在一串聯之幀上面臨限上之改變之值而實施此一分析。此一步驟係由模塊62來實施，它產生累積之臨限改變 Δ 。此一累積之臨限改變 Δ 係與模塊64內之預定絕對值Athrd比較，以及此處理通過任一支組66或支組68而進行，耽視此 Δ 是否係大於或不大於Athrd。如果不大於Athrd，模塊70係發動(如果如此，模塊72係即發動)。模塊70和72保持並更新臨限值T1，相當於探測之轉移之前之臨限值，以及模塊72保持並更新Threshold 2，相當於轉移之後之臨限。這兩個臨限(T1/T2)之比例係隨後與模塊74內之第三臨限Rthrd作比較。如果此比例係較此第三臨限為大時，那麼一正常語音標誌(ValidSpeech flag)係設定。此正確語音標誌係使用於第4圖之語音信號狀態機內。

第9A和9B圖說明操作中之局部語音探測機構。第9A圖相當於會採取Yes支線68(第3圖)之狀況，然而第9B圖相當於會採取No支線66之狀況。參看第9A圖，應予說明者，即有一轉移在自150至160之臨限中。在此所說明之範例中此一轉移係較絕對值Athrd為大。在第9B圖中此轉移在臨限中，自位置152至位置162代表一轉移係不大於Athrd

(請先閱讀背面之注意事項再填寫本頁)

訂

五、發明說明 (13)

。在第9A和9B兩者圖中此轉移位置業已由虛線170所說明。轉移位置之前之平均臨限值係經指示為T1以及轉移位置之後之平均臨限係經指示為T2。比例T1/T2隨後係與比例臨限Rthrd(第3圖內方塊74)比較。ValidSpeech係自預置語音區內簡單寄生雜音辨別如下。如果臨限內之此轉移係較Athrd為小時，或者如果T1/T2比例係較Rthrd為小時，那麼，可回應臨限轉移之信號係被辨別為雜音。另一方面，如果T1/T2比例係較Rthrd為大時，那麼，可回應臨限轉移之信號係被視為局部語音處理，以及它係不用來更新此臨限。

現在參看第4圖，此語音信號狀態機開始，如在起始狀態310中300處所指示者。它隨後進行至無聲狀態320，在此處它保持此狀態直到無聲狀態中所實施之步驟中指定一轉換至語音狀態330。一旦在語音狀態330中，當一定狀況係符合如由語音狀態330方塊內所說明者之步驟所指示時，此狀態機將轉換回至無聲狀態320。

在初始狀態310內數據幀係貯存於緩衝器50內(第2圖)，以及此梯級頻步進大小係經更新。吾人將可記憶該較佳實施例開始以一極小之步進大小 $M=20$ 來操作。此一步進大小可以在初始狀態中被採用一如由上文提供之偽代碼所說明者。同時在初始狀態中此梯級頻數據結構係經啟始以有早期操作移除任何先前貯存之數據。經這些步驟係已實施後，此狀態機轉換至無聲狀態320。

在無聲狀態中每一頻帶受限之短期能量值係與基本臨

五、發明說明(14)

其本之一組臨限。在第4圖內，可應用於信號線路26之臨限(第1圖)係經指定為Threshold-All，以及可應用於信號線路28之臨限係經指定為Threshold-HPF。類似之學名係用作應用於語音狀態330內之其他臨限值。

如果任一短期能量值超過其臨限時，那麼此“開始之延遲決定標誌”係經測試。如果該標誌係經設定為真(TRUE)，一如前文所討論者，一“語音之開始”之訊息係回行，以及此狀態機轉換至語音狀態330。否則，此狀態機保持於無聲狀態中，以及梯級頻數據結構係被更新。

本較佳實施例使用一0.99之忽略因數更新此梯級頻以造成無現時數據在這段時間消失之效果。此係以在添加與現時幀能量相關聯之“計數數據”之前以0.99倍增在梯級頻內現有值來完成。以此一方式，歷史數據之效果係逐漸地隨時間而減小。

雖然不同組之臨限值係使用，但語音狀態330內程序沿著同一直線進行。此語音狀態以WThreshold比較信號線路26和28內各自之能量。如果任一信號線路係高於WThreshold時，那麼一類似之比較以面對而此SThreshold來完成。如果在任一信號線路中之能量係高於SThreshold時，那麼正確語音標誌係設定為“真”。此一標誌在後續之比較步驟中使用。

如果此終結之延遲決定標誌係先前地設定至“真”時，一如上文所述，以及如果此“正確語音”標誌亦已設定為“真”時，那麼，語音之終結之一訊息係回行，以及此狀態

訂

五、發明說明(15)

真”時，那麼，語音之終結之一訊息係回行，以及此狀態機轉回至無聲狀態320。另一方面，如果此“正確語音”標誌係未曾設定為“真”時，一訊息係經發送以取消先前之語音探測，以及此狀態機轉換回至無聲狀態320。

第10和第11圖顯示各種位準如何地影響狀態機操作。第10圖兩者信號線路之同一時間之操作，此全_頻帶(all-frequency band)，帶_全(Band-All)，以及高頻率帶，帶_HPF。應予說明者，即此信號波形因為它們含不同之頻率內涵而不同。在所說明之範例中此最終範圍係經識別為相當於由橫越在b1處之臨限之所有頻帶所產生之語音之開始和相當於橫越在e2處之高頻帶之語音之終結之測得之語音。不同之輸入波形當然會產生依照第4圖內所說明之算法而產生不同之結果。

第11圖顯示此強臨限SThreshold係如何回來證實正確語音之存在於強雜音位準之出現中。如所說明者，一強雜音它落入SThreshold之下者係為區域R負責，它會相當於一正確語音標誌係設定至“偽”。

自前文所述，吾人將瞭解，即本發明提供一系統，它將探測在一輸入信號內之語音之開始及終結，應付甚多在雜音環境中消費者使用上所遭遇之問題。同時本發明業經說明於其目前之較佳形態中，應予瞭解者，即本發明係具有某些變更之能力而不背離本發明之精神，一如在增列之申請專利範圍中所宣佈者。

五、發明說明(16)

元件標號對照

30...輸入信號

22...漢明窗口

24...快速傅里葉變換轉換器

26, 28...信號線路

34, 36...濾波器

50...更新緩衝器

100...預置語音區

110...雜音尖峰脈衝

120...暫短停止

150, 152...轉移位置

160, 162...轉移位置

170...轉移位置

200...語音之開始

訂

檢

四、中文發明摘要(發明之名稱： 用於多雜音狀況之語音檢測系統)

輸入信號係經變換成為頻率領域，並隨後次分成為相當於不同頻率範圍之帶。適應性臨限係個別地應用於自各頻率帶之數據。因此，短期限制帶之能量係為一語音信號之出現與否而被測試。此適應性臨限值係使用梯級頻率數據結構以累積代表各自之頻率帶內能量之平均及變化之長期數據而為每一信號線路獨立地更新。端點探測係由一狀態機器來實施，它自無語音狀態轉變至語言出現狀態以及反之亦然，耽視臨限比較之結果而定。一局部語音探測系統應付輸入信號係經在其中捨位之情況。

英文發明摘要(發明之名稱： SPEECH DETECTION SYSTEM FOR NOISY CONDITIONS)

The input signal is transformed into the frequency domain and then subdivided into bands corresponding to different frequency ranges. Adaptive thresholds are applied to the data from each frequency band separately. Thus the short-term band-limited energies are tested for the presence or absence of a speech signal. The adaptive threshold values are independently updated for each of the signal paths, using a histogram data structure to accumulate long-term data representing the mean and variance of energy within the respective frequency band. Endpoint detection is performed by a state machine that transitions from the speech absent state to the speech present state, and vice versa, depending on the results of the threshold comparisons. A partial speech detection system handles cases in which the input signal is truncated.

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

六、申請專利範圍

1. 一種語音檢測系統，用以檢測一輸入信號以確定一語音信號是否係出現或未出現者，包含：

一頻帶分裂器，用以分裂該輸入信號成為多個頻帶，各帶代表限制之帶信號能量相當於頻率之不同範圍；

一能量比較器系統，以多個臨限用以比較該多個頻帶之受限制之帶信號能量，如此，該每一頻帶係與該帶相關聯之至少一個臨限作比較；以及

一語音信號狀態機聯接至該能量比較器系統，它轉換；

(a)當該帶之至少一個之受限制之帶信號能量係高於其相關聯之臨限之至少一個時，自一無語音狀態至語音呈現狀態，以及

(b)當該帶之至少一個之受限制之帶信號能量係低於其相關聯之臨限之至少一個時，自一語音呈現狀態至無語音狀態。

2. 如申請專利範圍第 1 項之系統，另包含適應性臨限更新系統，它引用梯級頻數據結構以累積可指示該頻帶之至少一個內之能量之歷史數據。
3. 如申請專利範圍第 1 項之系統，另包含一分開之適應性臨限更新系統與每一該頻帶相關聯。
4. 如申請專利範圍第 1 項之系統，另包含適應性臨限更新系統，它根據每一該頻帶內能量之平均值和變化而改正該多個臨限。

六、申請專利範圍

5. 如申請專利範圍第 1 項之系統，包含局部語音探測系統可回應於該多個臨限之至少一個內之改變速率上之預定轉移，如果該一個臨限之平均值之轉移之前至該轉移之後之比例超越一預定值時，該局部語音探測系統抑制該狀態機之轉換至一語音呈現狀態。
6. 如申請專利範圍第 1 項之系統，另包含一多臨限系統，它界定：
- 一 第一臨限作為高於雜音層之一預定偏移量；
 - 一 第二臨限作為該第一臨限之一預定百分比，該第二臨限係較小於該第一臨限；以及
 - 一 第三臨限作為該第一臨限之一預定倍數，該第三臨限係該第一臨限為大；以及
- 其中該第一臨限控制自無語音狀態至語音呈現狀態之轉換；以及
- 其中該第二及第三臨限控制自語音呈現狀態至無語音狀態之轉換。
7. 如申請專利範圍第 6 項之系統，其中如果該帶之至少一個之受限制之帶信號能量係低於該第二臨限，以及如果該帶之至少一個之受限制之帶信號能量低於第三臨限時，該狀態機自該語音呈現狀態轉換至無語音狀態。
8. 如申請專利範圍第 1 項之系統，另包含延遲決定緩衝器，它貯存代表該輸入信號之預定時間增加之數據，以及如果該多個頻帶之至少一個之受限制之帶信號能

(請先閱讀背面之注意事項再填寫本頁)

訂
線

六、申請專利範圍

量在遍及該預定時間增加中並未超越至少一個臨限時，它抑制狀態機之自無語音狀態轉換至語音呈現狀態。

9. 一種在一輸入信號中測定一語音信號是否係呈現或未出現之方法，其包含之步驟為：

分裂該輸入信號成為多個頻帶，各帶代表一受限制之帶信號能量相當於不同範圍之頻率；

比較該多個頻帶之受限制之帶信號能量與多個臨限，如此，該每一頻帶係與至少一個與該帶相關聯之臨限作比較；以及

測定該：

(a)當該帶之至少一個之受限制之帶信號能量係高於其相關聯臨限之至少一個時，一語音呈現狀態即存在，以及

(b)當該帶之至少一個之受限制之帶信號能量係低於其相關聯臨限之至少一個時，一無語音狀態係存在。

10. 如申請專利範圍第 9 項之方法，另包含使用一梯級頻以累積可指示該頻帶之至少一個內之能量之歷史數據而界定該多個臨限之至少一個。

11. 如申請專利範圍第 9 項之方法，另包含分別地為每一該頻率通應性地更新該多個臨限之至少一個。

12. 如申請專利範圍第 9 項之方法，另包含根據每一該頻帶內能量之平均值及變化而修正該多個臨限。

13. 如申請專利範圍第 9 項之方法，另包含探測該多個臨

六、申請專利範圍

限之至少一個內改變速率上之一預定轉移，並測定，
如果該一臨限之平均值之轉移之前對轉移之後之比例
超越一預定值時，該語音呈現狀態即並不存在。

14. 如申請專利範圍第 9 項之方法，另包含界定；

一第一臨限作為高於此雜音層之一預定偏移量；

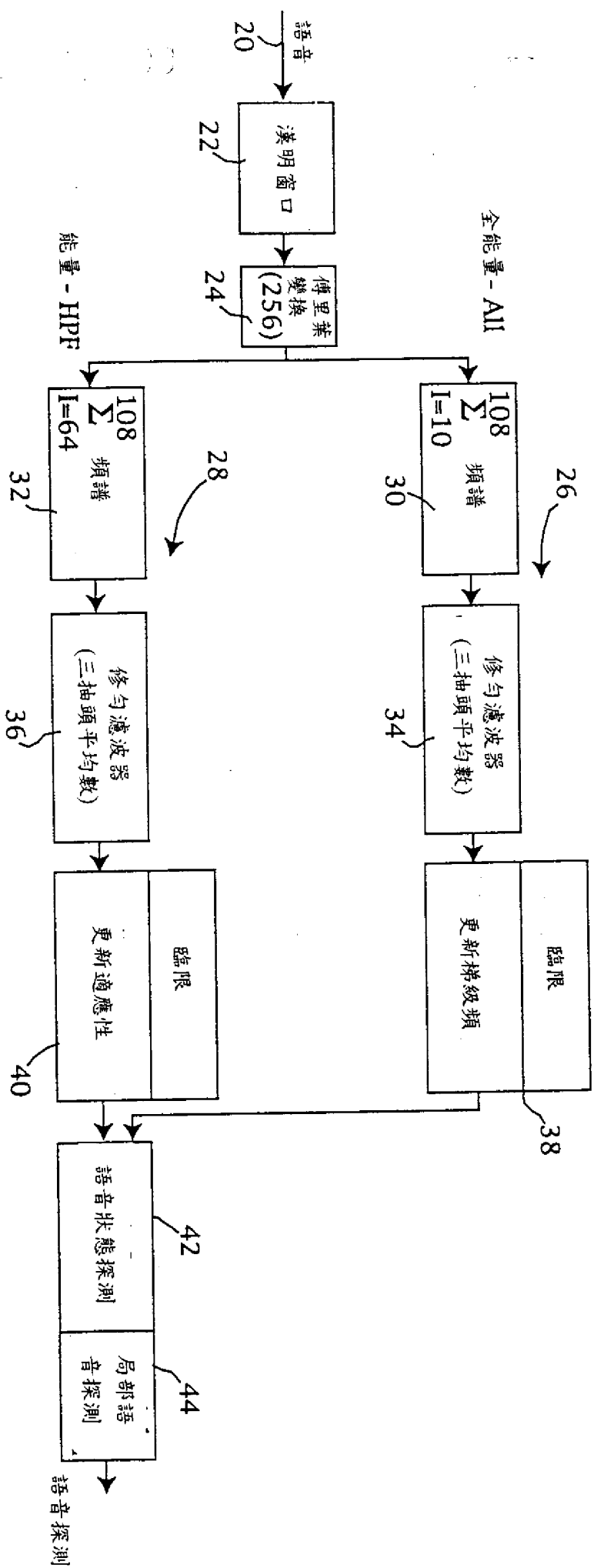
一第二臨限作為該第一臨限之一預定百分比，該
第二臨限係較該第一臨限為小；以及

一第三臨限作為該第一臨限之一預定倍數，該第
三臨限係較該第一臨限為大；以及

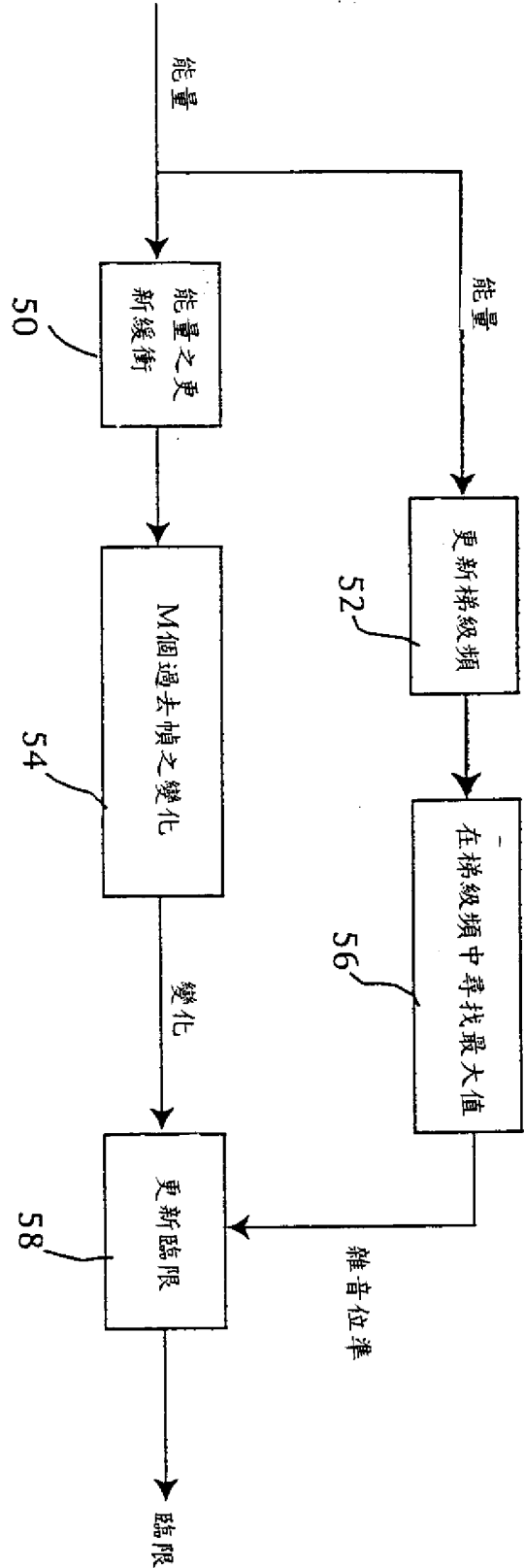
根據該第一臨限測定該語音呈現狀態之存在以及
根據該第二和第三臨限測定該無語音狀態之存
在。

15. 如申請專利範圍第 14 項之方法，其中如果該帶之至少
一個之受限制之帶信號能量係高於該第二臨限，以及
如果該帶之至少一個之受限制之帶信號能量係高於該
第三臨限時，該無語音狀態係經測定為存在。

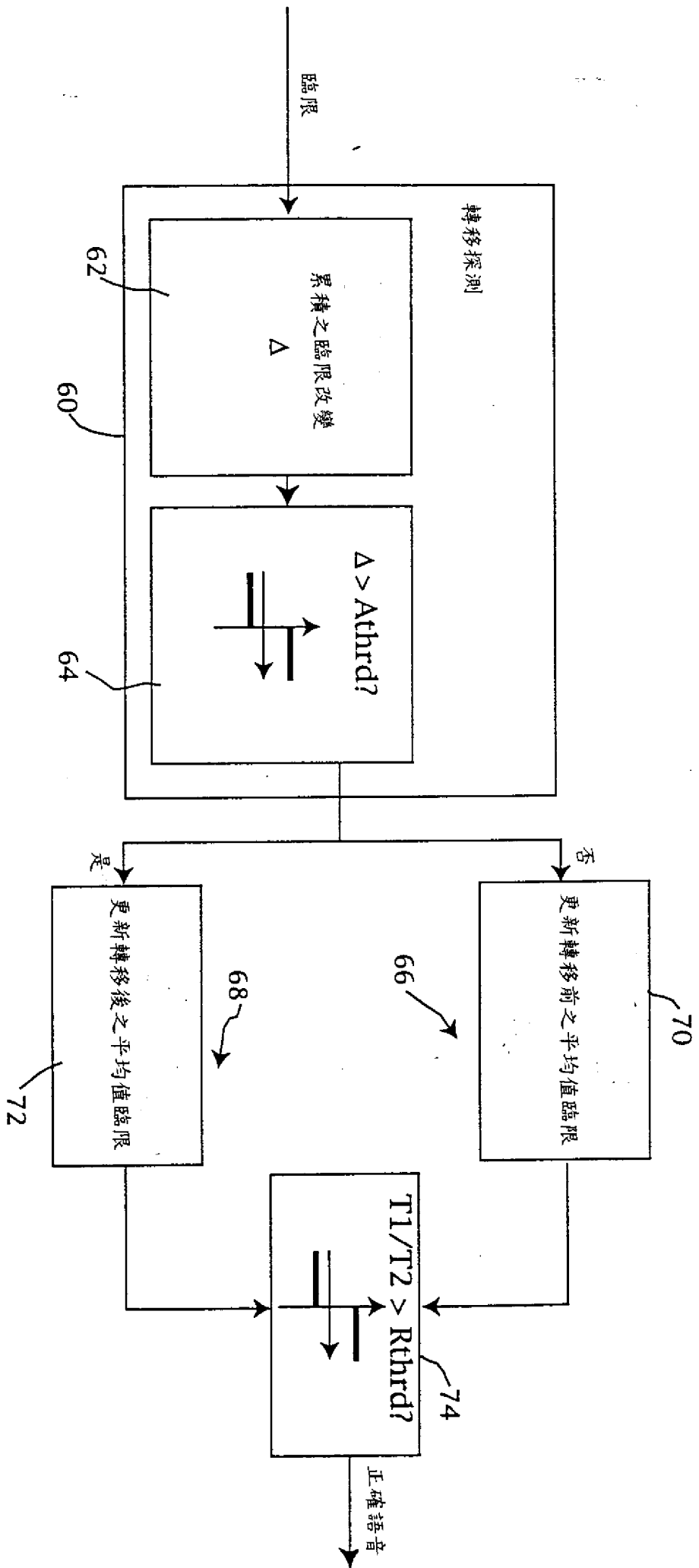
16. 如申請專利範圍第 9 項之方法，另包含如果該多個頻
帶之至少一個之受限制之帶信號能量在遍及一預定之
增加之時間中並未超越至少一個臨限時，即測定該語
音呈現狀態並不存在。



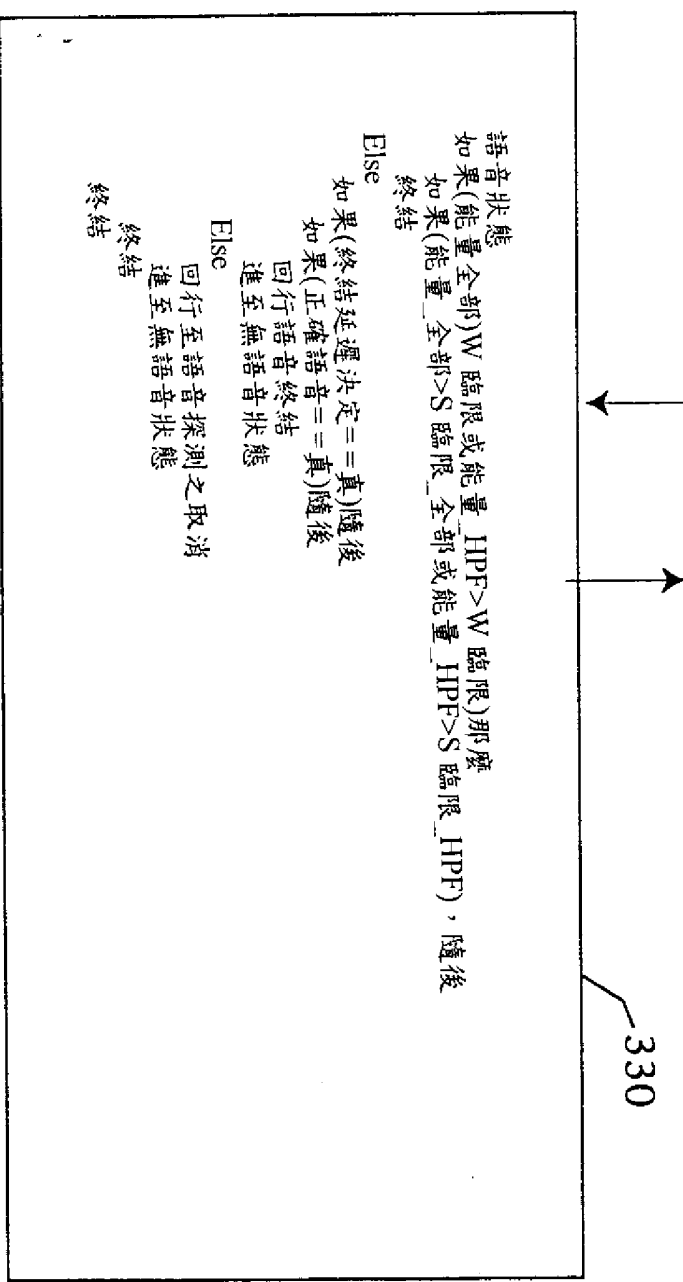
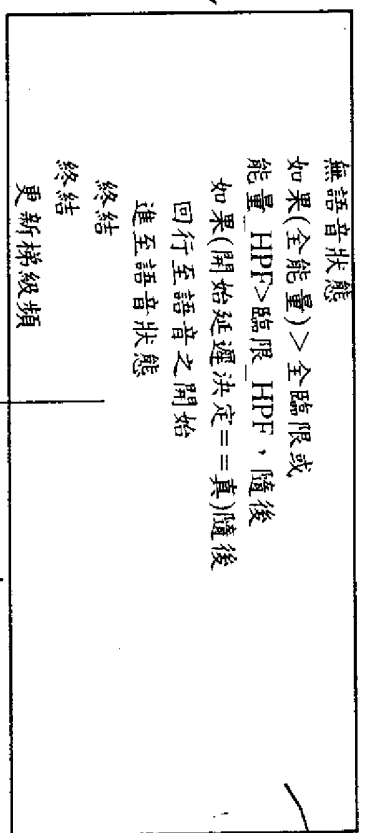
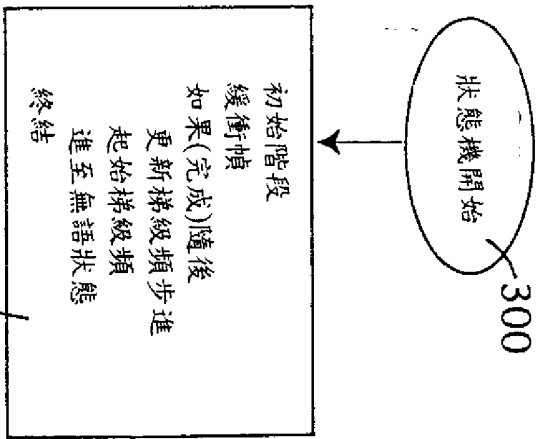
第 1 圖

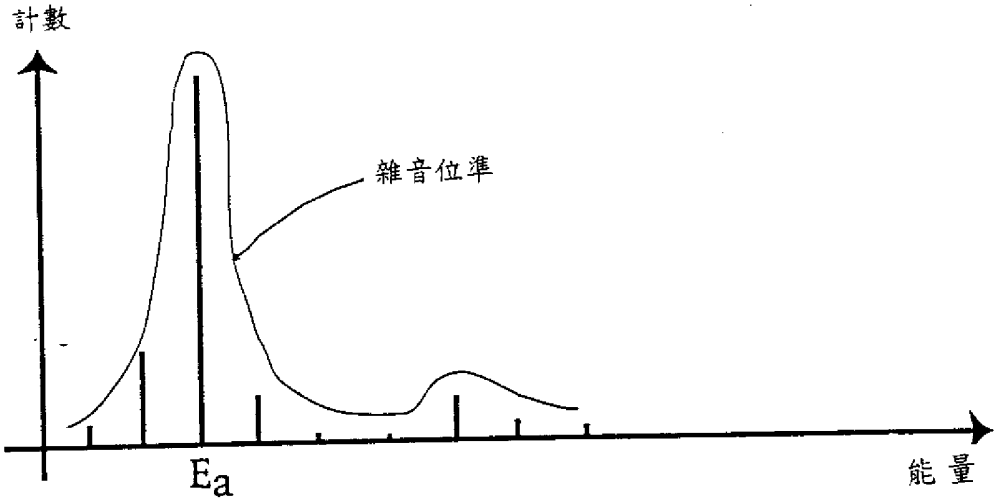


第 2 圖

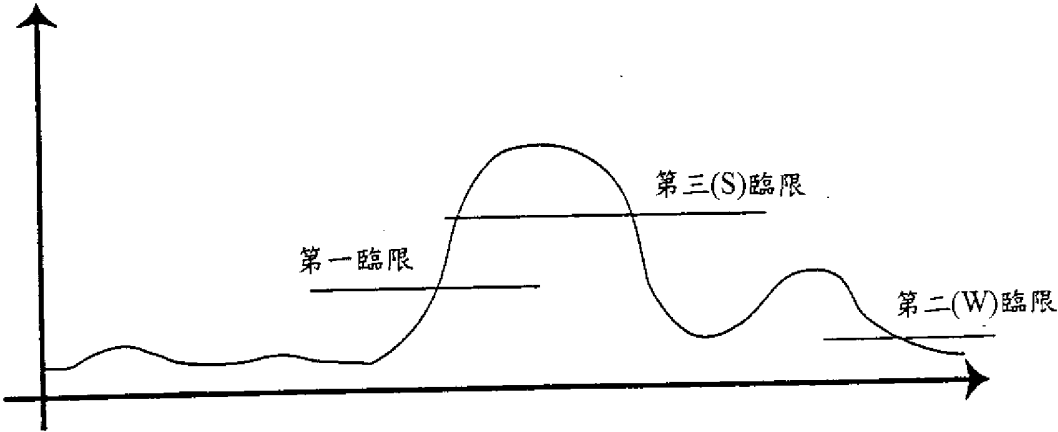


第 3 圖

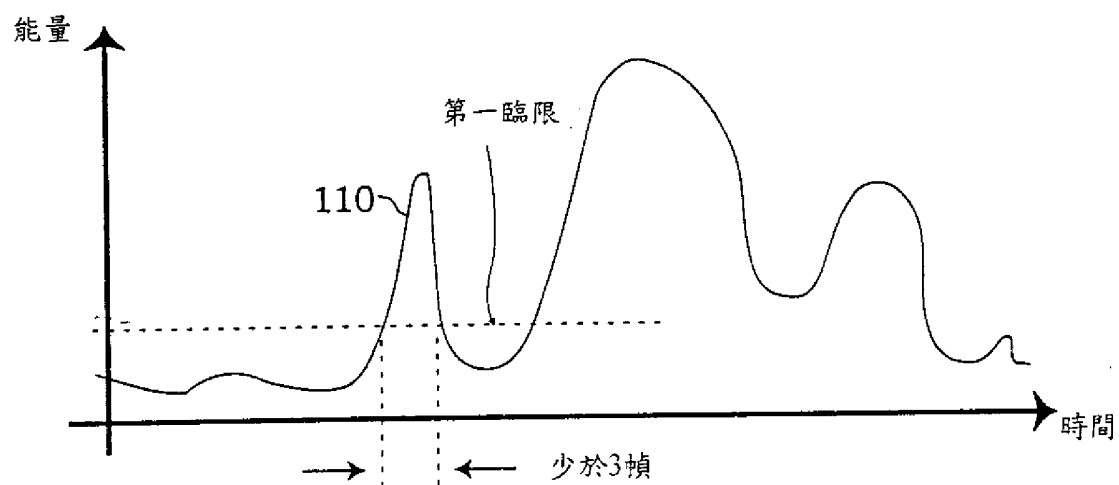




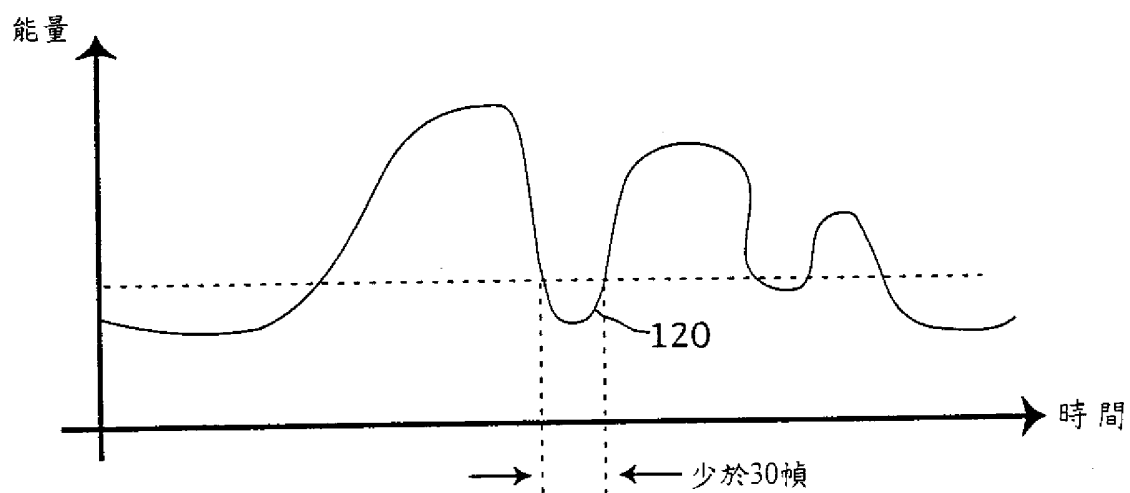
第 5 圖



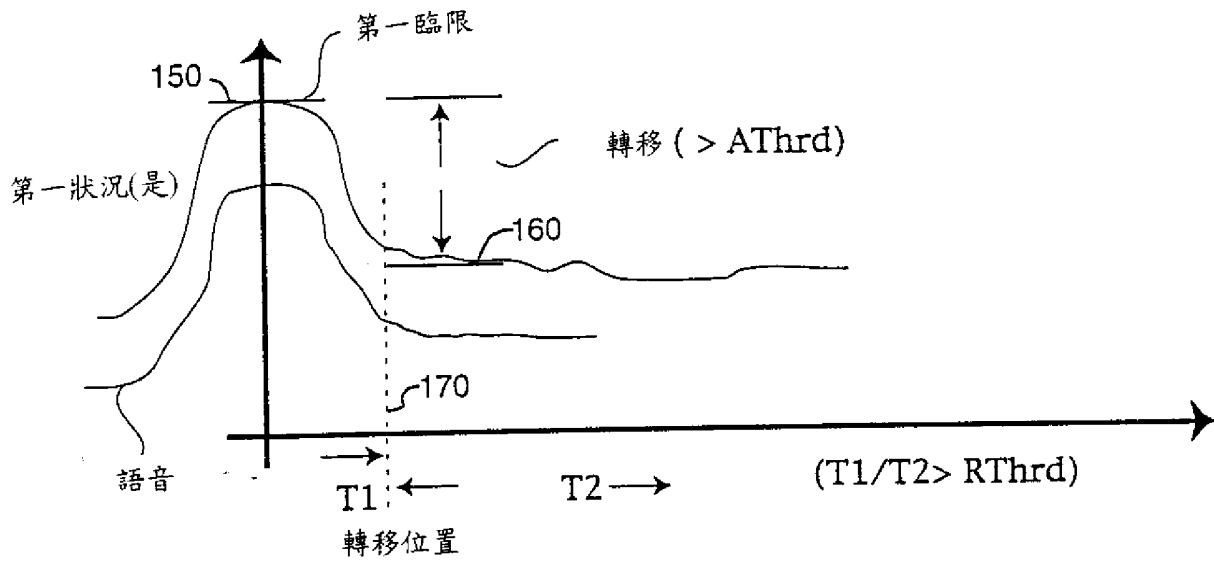
第 6 圖



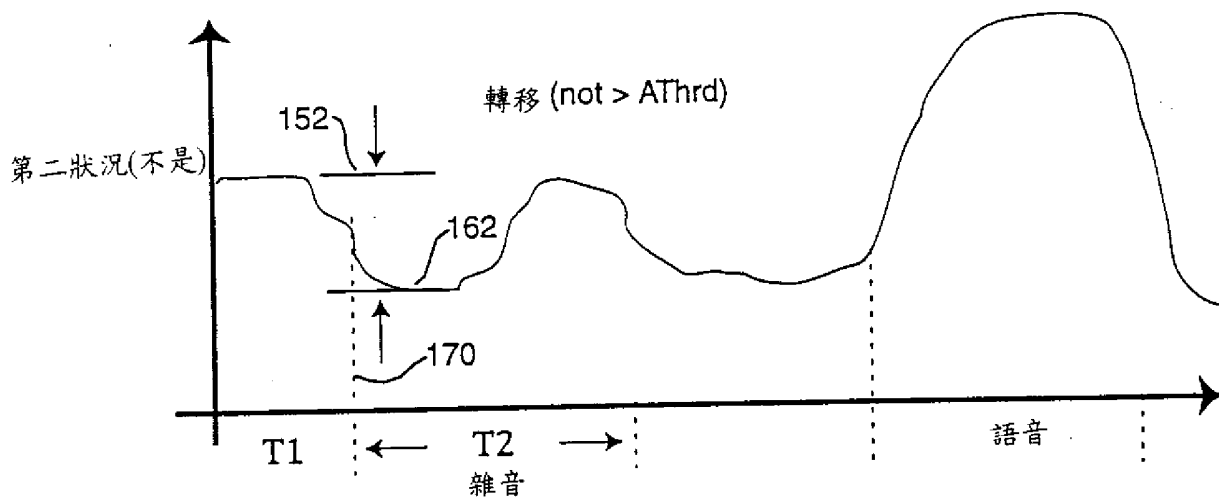
第 7 圖



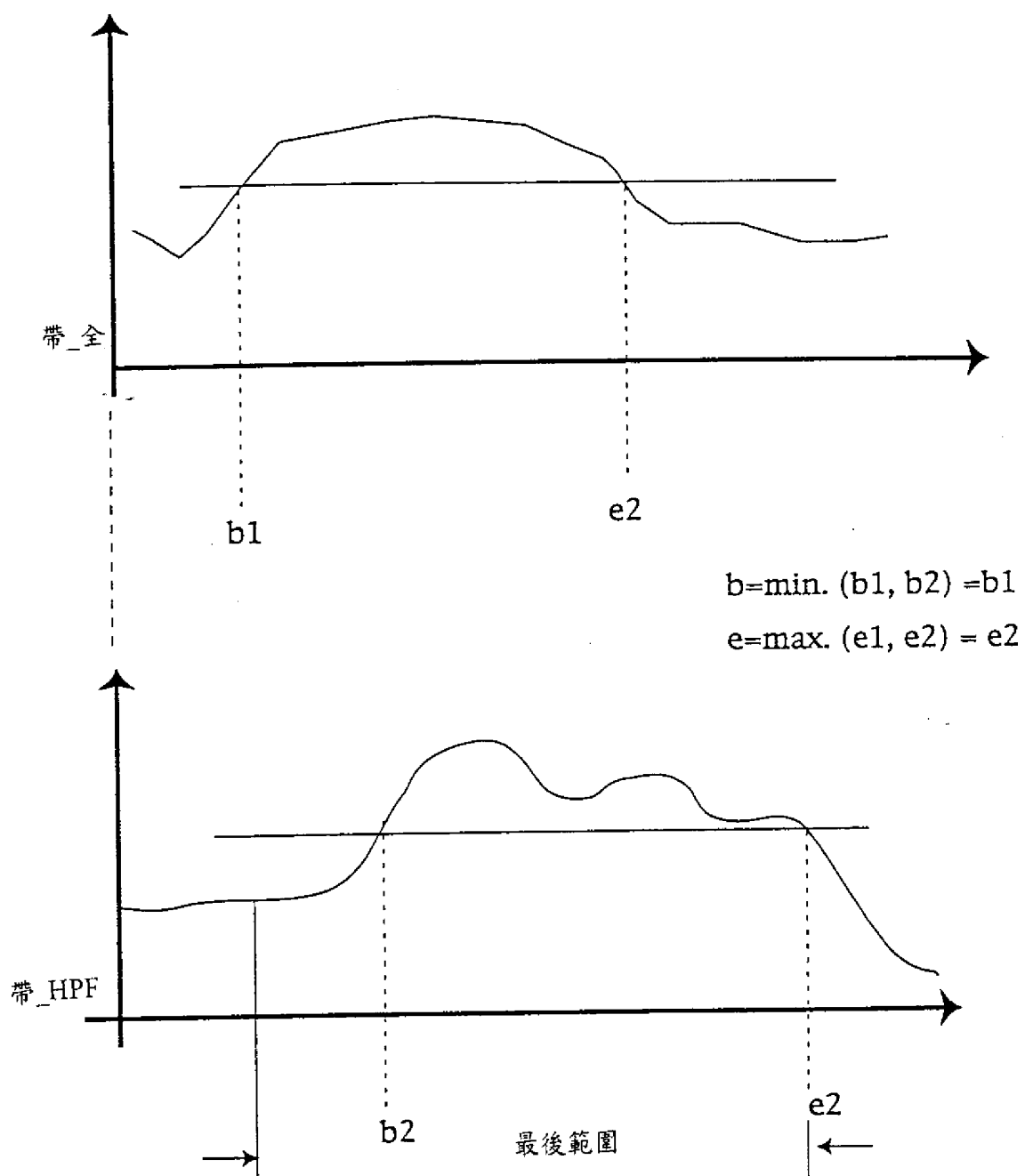
第 8 圖



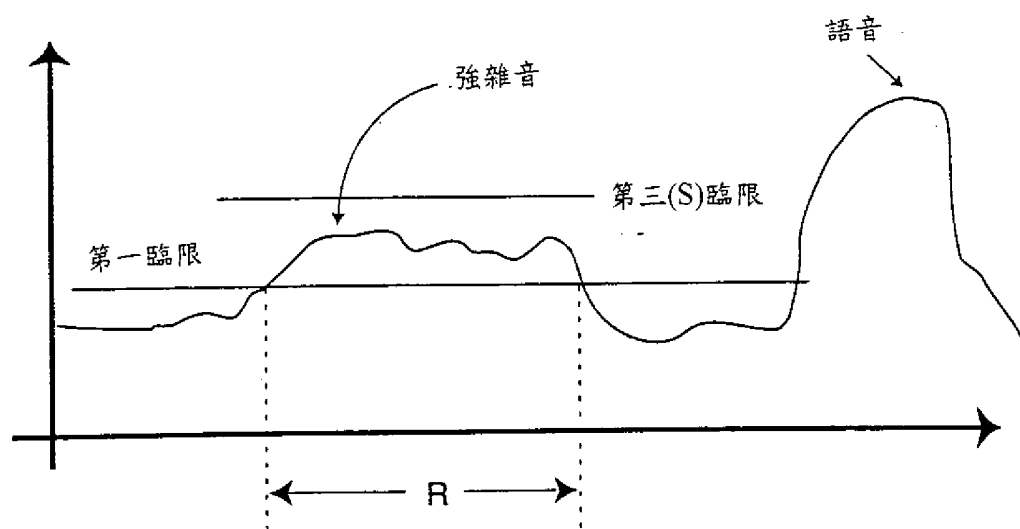
第9A圖



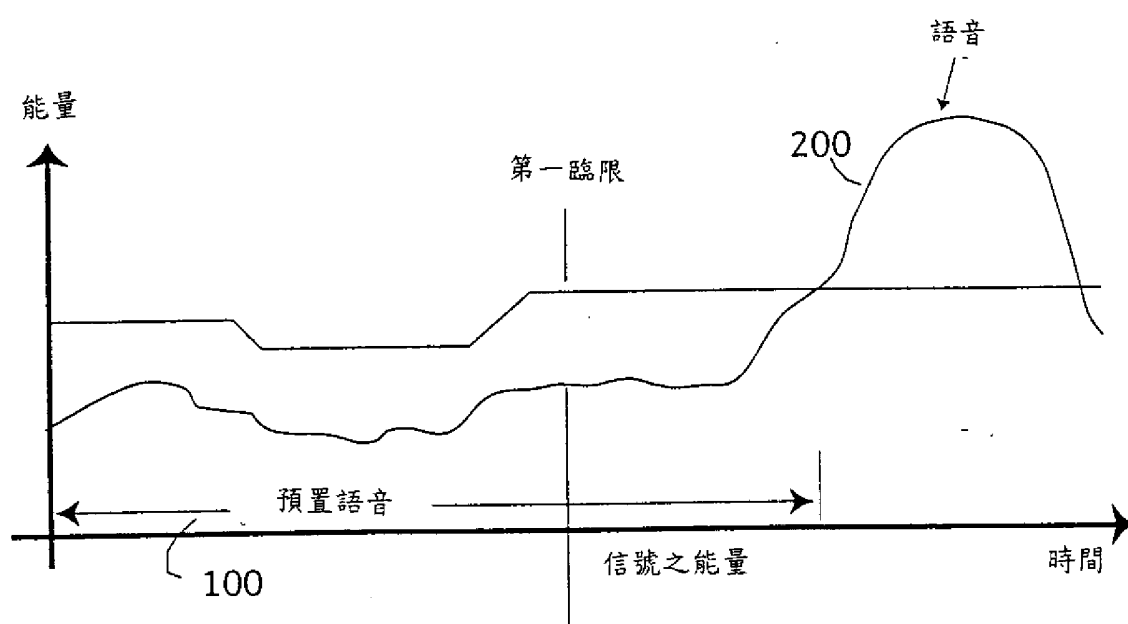
第9B圖



第10圖



第11圖



第12圖