



(19) 中華民國智慧財產局

(12) 發明說明書公告本

(11) 證書號數：TW I579834 B

(45) 公告日：中華民國 106 (2017) 年 04 月 21 日

(21) 申請案號：101127284

(22) 申請日：中華民國 101 (2012) 年 07 月 27 日

(51) Int. Cl. : G10L21/02 (2013.01)

(30) 優先權：2011/07/29 美國 61/513,298

(71) 申請人：D T S 股份有限公司 (美國) DTS LLC (US)

美國

(72) 發明人：崔斯 詹姆士 TRACEY, JAMES (US) ; 盧大慶 NOH, DAEKYOUNG (KR) ; 何星 HE, XING (CN)

(74) 代理人：詹銘文；葉璟宗

(56) 參考文獻：

US 4454609

US 2007/0299659A1

審查人員：涂淑惠

申請專利範圍項數：11 項 圖式數：10 共 51 頁

(54) 名稱

調整聲音清晰度強化的方法與系統

METHOD AND SYSTEM FOR ADJUSTING VOICE INTELLIGIBILITY ENHANCEMENT

(57) 摘要

一種適應地處理語音以增加聲音清晰度的系統與方法。此系統與方法可適應性地識別與追蹤共振峰位置，由此隨著共振峰的改變，啟動欲強調的共振峰。基此，即使在吵雜環境，此系統與方法仍可增加近端清晰度。此系統與方法可實作於網路電話(Voice-over IP, VoIP)應用程式、電話與/或視訊應用程式(包括手機、智慧型手機等)、筆記型與平板通訊設備等。系統與方法也可強化包括非以聲道所產生之非語音的聲音，如短音。

Systems and methods for adaptively processing speech to improve voice intelligibility are described. These systems and methods can adaptively identify and track formant locations, thereby enabling formants to be emphasized as they change. As a result, these systems and methods can improve near-end intelligibility, even in noisy environments. The systems and methods can be implemented in Voice-over IP (VoIP) applications, telephone and/or video conference applications (including on cellular phones, smart phones, and the like), laptop and tablet communications, and the like. The systems and methods can also enhance non-voiced speech, which can include speech generated without the vocal track, such as transient speech.

指定代表圖：

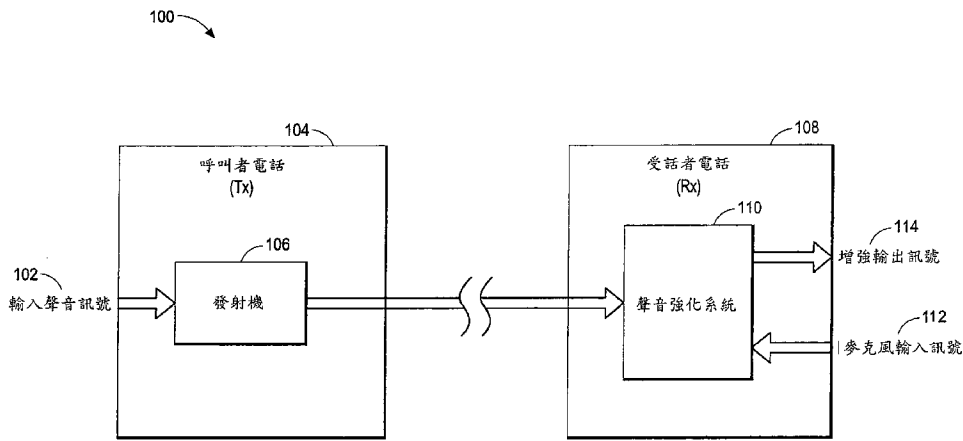


圖 1

符號簡單說明：

100 . . . 行動電話環境

102 . . . 輸入聲音訊號

104 . . . 呼叫者電話

106 . . . 發射機

108 . . . 受話者電話

110 . . . 聲音強化系統

112 . . . 麥克風輸入訊號

114 . . . 強化輸出訊號

105-08-09

公告本

# 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：101127284

※ 申請日期：101.7.27

※IPC 分類：G10L 21/02 (2013.01)

## 一、發明名稱：(中文/英文)

調整聲音清晰度強化的方法與系統

METHOD AND SYSTEM FOR ADJUSTING VOICE  
INTELLIGIBILITY ENHANCEMENT

## 二、中文發明摘要：

一種適應地處理語音以增加聲音清晰度的系統與方法。此系統與方法可適應性地識別與追蹤共振峰位置，由此隨著共振峰的改變，啟動欲強調的共振峰。基此，即使在吵雜環境，此系統與方法仍可增加近端清晰度。此系統與方法可實作於網路電話(Voice-over IP, VoIP)應用程式、電話與/或視訊應用程式(包括手機、智慧型手機等)、筆記型與平板通訊設備等。系統與方法也可強化包括非以聲道所產生之非語音的聲音，如短音。

## 三、英文發明摘要：

Systems and methods for adaptively processing speech to improve voice intelligibility are described. These systems and methods can adaptively identify and track formant locations, thereby enabling formants to be emphasized as

105-08-09

they change. As a result, these systems and methods can improve near-end intelligibility, even in noisy environments. The systems and methods can be implemented in Voice-over IP (VoIP) applications, telephone and/or video conference applications (including on cellular phones, smart phones, and the like), laptop and tablet communications, and the like. The systems and methods can also enhance non-voiced speech, which can include speech generated without the vocal track, such as transient speech.

**四、指定代表圖：**

(一)本案指定代表圖為：圖 1。

(二)本代表圖之元件符號簡單說明：

100：行動電話環境

102：輸入聲音訊號

104：呼叫者電話

106：發射機

108：受話者電話

110：聲音強化系統

112：麥克風輸入訊號

114：強化輸出訊號

**五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：**

無

## 六、發明說明：

### 【發明所屬之技術領域】

本發明是有關於一種聲音的處理，特別是有關於適應性語音清晰度的處理。

### 【先前技術】

手機常被用於高背景雜訊之地區。此雜訊往往已達到會造成來自於手機揚聲器之語音訊息的清晰度大大地被降低的程度。在許多例子中，當受話者接聽時，可能會因為周圍高雜訊遮蔽了呼叫者的聲音或使呼叫者的聲音失真，而使一些訊息遺失或至少其中部分模糊不清。

目前，等化器、截波器或簡單地增加手機音量已被用來在高背景雜訊中最小化清晰度的遺失。等化器與截波器會因本身所產生的背景雜訊，而無法解決此問題。增加手機之整體聲音等級或揚聲器音量無法顯著地改善清晰度，並且會導致其他如回授與受話者不舒服等問題。

### 【發明內容】

為了概述本揭露，在此描述本發明的目的、優點與新特徵。然，必須理解的是，本發明所揭露之任何特定實施例所達到之所有優點並非是必要的。因此，所揭露的發明是以可以達到或最佳化在此所教示之一個優點或一群優點的方式來實施或完成，而無需達到在此所教示或建議的其他優點。

一些實施例提出一種調整聲音清晰度強化的方法，其包括接收輸入聲音訊號並且以線性預測編碼(linear

predictive coding)程序來獲得此輸入聲音訊號的頻譜圖。此頻譜圖可包括一或多個共振峰頻率。此外，此方法也包括以一或多個處理器來調整輸入聲音訊號的頻譜以產生用以強化一或多個共振峰頻率的強化濾波器。此外，此方法還包括施予此強化濾波器至輸入聲音訊號的表述中以產生具有強化共振峰頻率的已修改聲音訊號，根據輸入聲音訊號偵測一包絡，並且分析此已修改聲音訊號之包絡以判斷一或多個暫時強化係數。再者，此方法更包括施予此一或多個暫時強化係數至上述已修改聲音訊號來產生輸出聲音訊號，其中至少施予一或多個暫時強化係數的運作是由一或多個處理器來執行。

在一些實施例中，前述之方法可包括下列特徵之任何組合：上述施予一或多個暫時強化係數至已修改聲音訊號的步驟包括，銳化已修改聲音訊號中一或多個包絡的波峰以強調已修改聲音訊號中被選擇之子音；上述偵測包絡的步驟包括偵測以下一或多種訊號的包絡：輸入聲音訊號與已修改聲音訊號；且更包括施予反向濾波器至輸入聲音訊號以產生激發訊號，如所述施予強化濾波器至輸入聲音訊號的表述，包括：施予強化濾波器至激發訊號。

一些實施例提出一種調整聲音清晰度強化的系統，其包括：分析模組，其用以獲取輸入聲音訊號的至少一部分的頻譜圖。頻譜圖可包括一或多個共振峰頻率。此系統也可包括一共振峰強化模組，其中共振峰強化模組可產生強化濾波器，並且強化濾波器可強調一或多個共振峰頻率。

一或多個處理器可施予強化濾波器至輸入聲音訊號的表述，以產生經修改的聲音訊號。再者，此系統也可包括時間包絡塑形器，其用以至少部分地根據已修改聲音訊號的一或多個包絡，來施予時間強化至已修改聲音訊號。

在一些實施例中，前述之系統可包括以下特徵之任何組合：上述的分析模組更用以使用一個線性預測編碼技術以獲取輸入聲音訊號的頻譜圖，其中線性預測編碼技術用以產生對應於頻譜圖的係數；更包括對映模組，用以映射係數至線頻譜對；更包括修改線頻譜對來增加對應於共振峰頻率之頻譜圖的增益；上述的強化濾波器更用以施予在下述一或多個訊號：輸入聲音訊號和從輸入聲音訊號得到的激發訊號；上述的時間包絡塑形器更用以細分經修改的聲音訊號至數個頻段，而其中在至少一些複數的頻段中，一或多個包絡對應至一包絡；更包括聲音強化控制器，該聲音強化控制器用以至少部分地根據在輸入麥克風聲音訊號中所偵測到的環境雜訊量來調整強化濾波器的增益；更包括聲音動作偵測器，用以偵測輸入麥克風訊號的聲音，以及控制聲音強化控制器來回應所偵測到之聲音；其中聲音動作偵測器更用以根據反應於輸入麥克風訊號之偵測聲音的一先前雜訊輸入來使聲音強化控制器控制強化濾波器的增益；且更包括麥克風校正模組，用以設定用來接收輸入麥克風訊號的麥克風增益，其中麥克風校正模組更用以至少部分地根據參考與錄音雜訊訊號來設定增益。

一些實施例提出一種調整聲音清晰度強化的系統，其

包括線性預測編碼分析模組，該模組使用線性預測編碼技術以獲取對應於輸入聲音訊號頻譜的線性預測編碼係數，其中上述頻譜包括一或多個共振峰頻率。此系統也可包括對映模組，其用以將線性預測編碼係數映射至線頻譜對。此系統也可包括共振峰強化模組，其包括一或多個處理器，其中共振峰強化模組會修改線頻譜對以調整輸入聲音訊號的頻譜，且產生可強調一或多個共振峰頻率強化的濾波器。此強化濾波器可被施予至一輸入聲音訊號之表述，以產生一已修改聲音訊號。

在各種不同之實施例中，前述之系統可包括以下特徵之任何組合：更包括聲音動作偵測器，用以在輸入麥克風訊號中偵測聲音，且反應於輸入麥克風訊號的偵測聲音而引發強化濾波器的增益調整；更包括麥克風校正模組，用以設定可接收輸入麥克風訊號的麥克風的增益，其中麥克風校正模組更用以至少部分地根據參考訊號與錄音雜訊來設定增益並且上述強化濾波器更用以施予至以下一或多個訊號：輸入聲音訊號和從輸入聲音訊號得到的激發訊號；更包括時間包絡塑形器，其用以至少部分地根據已修改聲音訊號的一或多個包絡，來施予時間強化至已修改聲音訊號；並且時間包絡塑形器更用以銳化已修改聲音訊號的一或多個包絡的波峰，以強調已修改聲音訊號的已選擇部分。

### 【實施方式】

#### I.前言

現存的聲音清晰度系統是嘗試強調語言中的共振

峰，其中共振峰包括由講話者聲帶所產生之對應於某些母音與子音的共振頻率。這些現存的系統一般是使用濾波器組，其具有用以在預期有共振峰發生的不同固定頻率的頻帶強調共振峰的帶通濾波器。此方法的問題是，共振峰的位置會隨著不同個體而不同。再者，一給定之個體之共振峰位置可隨時間改變。固定頻帶帶通濾波器可能會因此強調與給定個體的共振峰不同的共振峰頻率，導致聲音清晰度減弱。

在其他特徵之中，本揭露描述用以強化聲音清晰度的適應性處理語音的系統與方法。在某些實施例中，這些系統與方法可適應性地識別與追蹤共振峰位置，由此當共振峰位置改變時，啟動強調共振峰。基此，即使是在吵雜的環境下，這些系統與方法可增加近端清晰度。此系統與方法也可強化非語音的聲音，其中其可包括非聲帶產生之聲音，如短音。一些可強化非語音的聲音的例子包括阻塞音，如破裂音，摩擦音，以及破擦音。

許多技術可用於適應性地追蹤共振峰位置。適應性濾波器是其中之一種技術。在一些實施例中，用於以線性預測編碼(LPC)為背景的適應性濾波器可用以追蹤共振峰。為了方便說明，本說明書剩下的部分將會以線性預測編碼為背景，描述適應性共振峰追蹤。然而，必須理解的是，在某些實施例中，除了線性預測編碼之外，也可使用許多其他適應性處理技術來追蹤共振峰位置。在此可用於此取代線性預測編碼或與之並用的技術的例子包括多頻帶能量解

碼(multiband energy demodulation)、極點相互作用(pole interaction)、係數自由非線性預測(parameter-free non-linear prediction)、以及上下文相關音位資訊(content-dependent phonemic information)。

## II. 系統架構

圖 1 繪示實作聲音強化系統 110 的行動電話環境 100 的實施例。聲音強化系統 110 包括增加聲音輸入訊號 102 之清晰度的硬體與/或軟體。例如，聲音強化系統 110 以聲音強化來處理聲音輸入訊號 102，其中此聲音強化可凸顯語音聲音的區別特徵，如共振峰和非語音聲音(例如子音，包括，比如破裂音與摩擦音)。

範例行動電話環境 100 顯示呼叫者電話 104 與受話者電話 108。雖然在此範例中，聲音強化系統 110 是安裝在受話者電話中，然而，在其他實施例中，兩支電話皆可具有聲音強化系統。呼叫者電話 104 與受話者電話 108 可為行動電話、網路電話(voice over Internet protocol, VOIP)、智慧型電話(smart phone)、線路電話、公共電話與/或視訊會議電話、其他計算機裝置(如筆記型電腦或平板電腦)等。呼叫者電話 104 可視為在行動電話環境 100 之遠端，且受話者電話可視為在行動電話環境 100 之近端。當受話者電話 108 的使用者講話時，近端與遠端可互相交換。

在所描述之實施例中，呼叫者會提供聲音輸入 102 給呼叫者電話 104。呼叫者電話 104 中的發射機 106 會傳輸聲音輸入訊號 102 至受話者電話 108。發射機 106 會藉由

無線或有線線路或是兩者的組合來傳輸聲音輸入訊號 102。受話者電話 108 中的聲音強化系統 110 會強化聲音輸入訊號 102，以增加聲音清晰度。

聲音強化系統 110 可動態識別表現於聲音輸入訊號 102 中的聲音的共振峰或是其他特徵部分。基此，即使共振峰隨時間變化或是因不同講話者而不同，聲音強化系統 110 也可動態地強化聲音共振峰或是聲音的其他特徵部分。聲音強化系統 110 也可至少部分地依據從受話者電話 108 中的麥克風中所偵測到的麥克風輸入訊號 112 的環境雜訊來適應調整施予至聲音輸入訊號 102 的聲音強化的等級。環境雜訊或內容可能包括背景或週遭雜訊。若環境雜訊增加，聲音強化系統 110 可增加聲音強化量，反之亦然。聲音強化可能因此至少部分追蹤環境雜訊偵測量。同樣地，聲音強化系統 110 可至少部分地根據環境雜訊量來增加施予至聲音輸入訊號 102 的全體增益。

然而，當環境雜訊較少時，聲音強化系統 110 可減少施予之聲音強化量與/或增益增加量。當環境是低等級雜訊時，聲音強化與/或音量增大會令受話者感覺刺耳或不愉悅，所以如此地縮減對於受話者而言是有益的。例如，一旦環境雜訊超過門檻量，聲音強化系統 110 才開始施予聲音強化至聲音輸入訊號 102，以避免導致聲音在無環境雜訊中聽起來刺耳。

因此，在某些實施例中，在環境雜訊等級變化時，聲音強化系統 110 會將聲音輸入訊號轉換為對於收聽者而言

更具清晰度的強化輸出訊號 114。在一些實施例中，呼叫者電話 104 也包括聲音強化系統 110。聲音強化系統 110 可至少部分地根據由呼叫者電話 104 所偵測到之環境雜訊量，來施予此強化至聲音輸入訊號 102。聲音強化系統 110 可因此被使用在呼叫者電話 104、受話者電話 108 或兩者中。

雖然聲音強化系統 110 被顯示為電話 108 的一部分，然而，聲音強化系統 110 亦可實作在任何通訊裝置中。例如，聲音強化系統 110 可實作在電腦、路由器、類比電話轉接器、錄音機等裝置中。聲音強化系統 110 也可被用於公共廣播(Public Address, PA)裝備(包括在網路協定中的公共廣播)、無線電收發器、助聽器(例如，聽力幫助)、擴音器電話、與其他聲音系統中。再者，聲音強化系統 110 可被實作在用以提供聲音輸出給一或多個喇叭的任何以處理器為基礎的系統中。

圖 2 是繪示聲音強化系統 210 更詳細的實施例。聲音強化系統 210 可實現聲音強化系統 110 的一些或全部特徵，並且可以硬體與/或軟體的方式來實作。聲音強化系統 210 可實作於行動電話、手機、智慧型電話、或其他包括任何上述裝置的計算機裝置。聲音強化系統 210 可至少部分地依據偵測到之環境雜訊與/或輸入聲音訊號等級，來適應性地追蹤共振峰與/或聲音訊號的其他部分且調整強化處理。

聲音強化系統 210 包括適應性聲音強化模組 220。適

應性聲音強化模組 220 可包括用以適應性地於施予聲音強化至聲音輸入訊號 202(例如，從呼叫者電話接收、在助聽器或其他裝置中的聲音輸入訊號)的硬體與/或軟體。聲音強化可凸顯聲音輸入訊號 202 中包括語音與/或非語音聲音之口語聲音的區別特徵。

有利的是，在某些實施例中，適應性聲音強化模組 220 適應性地追蹤共振峰，以對於不同的說話者(例如，個體)或是隨時間改變共振峰的相同說話者，強化適當的共振峰頻率。適應性聲音強化模組 220 也可強化聲音中非語音的部分，例如某些子音或由聲帶以外的部分聲道所產生的其他聲音。在一實施例中，適應性聲音強化模組 220 會時域地塑形聲音輸入訊號來強化非語音的聲音。稍後將配合圖 3 更詳細描述這些特徵。

聲音強化控制器 222 是用以控制由聲音強化模組 220 所提供的聲音強化等級。聲音強化控制器 222 可提供一強化等級控制訊號或數值給適應性聲音強化模組 220，其會增加或減少施予之聲音強化的等級。控制訊號可隨著包括環境雜訊的麥克風輸入訊號 204 的增加與減少來批次或逐個來調適。

在某些實施例中，在偵測到麥克風輸入訊號 204 中的環境雜訊的功率門檻值之後，聲音強化控制器 222 會調整聲音強化的等級。在門檻值之上，聲音強化控制器 222 會引發聲音強化等級以追蹤或本質上追蹤在麥克風輸入訊號 204 中的環境雜訊量。例如，在一實施例中，在雜訊門檻

之上所提供的聲音強化等級是正比於雜訊能量(或功率)對門檻之比值。在另一實施例中，聲音強化等級亦可在不使用門檻下被調整。由強化控制器 222 施予的聲音強化的調適等級可隨環境雜訊的增加，以指數或線性方式增加(反之亦然)。

麥克風校正模組 234 會被提供，以確保或試圖確保對於組成聲音強化系統 210 的各個裝置來說，聲音強化控制器 222 將聲音強化等級調適在大約相同的等級。麥克風校正模組 234 會計算且儲存一或多組校正參數，其可調整施予至麥克風輸入訊號 204 的增益，用以使一些或全部裝置之麥克風的整體增益相同或差不多相同。稍後將配合圖 10 詳細描述麥克風校正模組 234 的功能。

當受話者電話 108 的麥克風收到該電話 108 的揚聲器輸出 114 的聲音訊號時，使人不愉悅的效果可能會產生。此揚聲器回授會被聲音強化控制器 222 視為環境雜訊，其會導致聲音強化的自我啟動與因揚聲器回授之聲音強化所造成的調變。調變輸出訊號的結果會造成收聽者的不愉悅。相似之問題可能發生在當受話者電話 108 正在輸出從呼叫者電話 104 接收的聲音訊號的時候，同時受話者講話、咳嗽或是用其他方法發出聲音到受話者電話 108 的情況。在這種說話者與收聽者兩者同時講話(或發出聲音)的雙重講話情況下，適應性聲音強化模組 220 會根據雙重講話調變聲音輸入訊號 202。此調變後的輸出訊號會令收聽者不愉悅。

為了克服此問題，所描述的實施例提供聲音動作偵測器 212。聲音動作偵測器 212 可偵測語音或從麥克風輸入訊號 204 之喇叭發出的其他聲音，且區別聲音與環境雜訊。當麥克風輸入訊號 204 包括環境雜訊，聲音動作偵測器 212 會根據當下量測之環境雜訊，允許聲音加強 222 調整由適應性聲音強化模組 220 所提供的聲音加強量。然而，當聲音動作偵測器 212 偵測到在麥克風輸入訊號 204 中的聲音時，聲音動作偵測器 212 可使用先前環境雜訊的量測來調整聲音強化。

在所描述的實施例中，聲音強化系統 210 包括額外強化控制 226，以更調整聲音強化控制器 222 所提供之控制量。額外強化控制 226 可提供額外強化控制訊號給聲音強化控制器 222，其中此額外強化控制訊號可被用作為一個數值並且強化等級無法少於此數值。額外強化控制 226 可經由使用者介面呈現給使用者。此控制 226 允許使用者增加聲音強化控制器 222 所決定的強化等級。在一實施例中，聲音強化控制器 222 可以從額外強化控制 226 加入額外強化給聲音強化控制器 222 所決定的強化等級。額外強化控制 226 會對想要更多聲音強化處理或是想要常常施予聲音強化處理的聽力損傷者特別有用。

適應性聲音強化模組 220 可提供輸出聲音訊號至輸出增益控制器 230。輸出增益控制器 230 可控制施予至聲音強化模組 220 的輸入訊號的全體增益。輸出增益控制器 230 可以硬體與/或軟體來實作。輸出增益控制器 230 可至少部

分地根據雜訊輸入 204 與聲音輸入訊號 202 的等級，來調整施予至輸出訊號之增益。除了任何使用者設定的增益之外，如電話的音量控制，也可應用此增益。有利的是，依據在麥克風輸入訊號 204 與/或聲音輸入訊號 202 等級的環境雜訊來調適聲音訊號之增益，可幫助收聽者更感知聲音輸入訊號 202。

適應性等級控制 232 也被顯示於所描述之實施例中，其可更進一步調整輸出增益控制器 230 所提供之增益。使用者介面也可呈現適應性等級控制 232 予使用者。隨著進入的聲音輸入訊號 202 等級降低或雜訊輸入 204 增加，增加此適應性等級控制 232 可使輸出增益控制器 230 的增益增加更多。隨著進入的聲音輸入訊號 202 等級降低或雜訊輸入 204 降低，降低此適應性等級控制 232 可使輸出增益控制器 230 的增益增加較少。

在一些實例中，聲音強化模組 220、聲音強化控制器 222 與/或輸出增益控制器 230 所施予的增益，會使聲音訊號截切或飽和。飽和會導致諧波失真，其會令收聽者不愉悅。因此，在某些實施例中，失真控制模組 140 會被提供。失真控制模組 140 可接收輸出增益控制器 230 之增益調整聲音訊號。失真控制模組 140 可包括硬體與/或軟體，以控制失真，同時至少部分保存或甚至更增加聲音強化模組 220、聲音強化控制器 222 與/或輸出增益控制器 230 所提供之訊號能量。雖然在提供給失真控制模組 140 訊號中並無出現削減，在一些實施例中，失真控制模組 140 會引發

至少部分飽和或削減，以更進一步增加音量與訊號清晰度。

在某些實施例中，失真控制模組 140 會藉由對應一或多個聲音訊號樣本至輸出訊號，來控制聲音訊號失真，其中此輸出訊號具有比完全飽和訊號較少之諧波。此對應可線性追蹤聲音訊號或近似線性追蹤未飽和之樣本。對於飽和之樣本，此對應可為採用已控制失真的非線性轉換。基此，在某些實施例中，失真控制模組 140 可允許聲音訊號在較少失真情況下比起全飽和訊號更大聲。因此，在某些實施例中，失真控制模組 140 會將表示物理聲音訊號的資料轉換成為表示有已控制失真之另外物理聲音訊號的資料。

聲音強化系統 110 與 210 之各種特徵可包括描述於美國專利第 8,204,742 號之相同或相似元件對應功能，其中此專利申請日為 2009 年 9 月 14 日，發明名稱為”適用性聲音清晰度處理”，並且其揭露的完整內容可作為本案的參考。此外，聲音強化系統 110 或 210 可包括任何描述於美國專利第 5,459,813 號(‘813 專利)之特徵，其中此專利申請日為 1993 年 6 月 23 日，發明名稱為”公有位置可讀系統”，其揭露的完整內容結合可作為本案的參考。例如，一些聲音強化系統 110 或 210 之實施例，可實現描述於‘813 專利的固定共振峰追蹤特徵，同時實作在此所描述的一些或所有其他特徵(例如，非聲音語音的時間強化、聲音動作偵測、麥克風校正、或此些的結合等)。同樣地，在未實作在此描述的一些或全部其他特徵的情況下，聲音強

化系統 110 或 210 之其他實施例亦可實作此述之適應性共振峰追蹤特徵。

### III. 適應性共振峰追蹤實施例

請參照圖 3，適應性聲音強化模組 320 的一實施例被繪示。適應性聲音強化模組 320 是圖 2 之適應性聲音強化模組 220 的更詳細實施例。因此，適應性聲音強化模組 320 可由聲音強化系統 110 或 210 來實作。基此，適應性聲音強化模組 320 可以軟體與/或硬體來實作。適應性聲音強化模組 320 可具優勢地與適應地追蹤聲音語言(例如，共振峰)且也可時域地強化非語音的聲音。

在適應性聲音強化模組 320 中，輸入語言會被提供給前置濾波器 310。此輸入語言對應到上述之聲音輸入訊號 202。前置濾波器 310 可為一高通濾波器或類似可衰減某些低頻的濾波器。舉例來說，在一實施例中，雖然可能選擇其他截止頻率，前置濾波器 310 衰減低於大約 750Hz 的頻率。藉由衰減約 750Hz 以下低頻頻譜的能量，前置濾波器 310 可創造更多淨空給隨後程序，以啟動較佳的線性預測編碼分析與強化。同樣地，在其他實施例中，前置濾波器 310 可包括取代高通濾波器或與其並用的低通濾波器，其衰減高頻且因此提供額外淨空至增益程序。在一些實作中，前置濾波器 310 可省略。

在所述實施例中，前置濾波器 310 的輸出會被提供給線性預測編碼分析模組 312。線性預測編碼分析模組 312 會使用線性預測技術至頻譜分析，且在頻譜中識別共振峰

位置。雖然在此是描述辨別共振峰位置，但更廣義來說，線性預測編碼分析模組 312 可產生用以表示輸入語音的頻率或輸入語音的功率頻譜的係數。頻譜圖包括在輸入語音中對應共振峰之波峰。所識別之共振峰会對應至頻帶，而非只是波峰本身。例如，所述位於 800Hz 之共振峰可能實際上包括 800Hz 附近之頻帶。藉由產生有此頻譜圖的這些係數，隨著共振峰位置在輸入語音中隨間時改變，線性預測編碼分析模組 312 可適應性地辨別共振峰的位置。適應性聲音強化模組 320 的後續元件因此能夠可適應性地強化這些共振峰。

在一實施例中，當全極點濾波器模型可精確地在語音中模擬出共振峰位置時，線性預測編碼分析模組 312 使用預測演算法來產生全極點濾波器係數。在一實施例中，一自相關方法被用來獲取用於全極點濾波器的係數。可被用來執行此分析的特定演算法是 Levinson-Durbin 演算法。儘管直接型式係數會被產生，然而，Levinson-Durbin 演算法亦會產生網狀格濾波器之係數。產生係數給一群樣本而非單一樣本，可增進處理效率。

線性預測編碼分析所產生之係數會有對於量化雜訊敏感的趨勢。一個非常小的係數誤差可令整個頻譜失真或使濾波器不穩定。為了減少量化雜訊在全極點濾波器上的效應，從線性預測編碼係數到線頻譜對(亦稱為線頻譜頻率)的對應或轉換可由對映模組 314 來執行。對映模組 314 可產生用於每個線性預測編碼係數的一係數。具優勢的是，

在某些實施例中，此對應可產生在單位元(在  $Z$  轉換領域)上的線頻譜對，以增進全極點濾波器的穩定度。另外，除了以線頻譜作為呈現對雜訊之敏感係數的方法外，此係數可以使用對數面積比例(Log Area Ratios, LAR)或其他技術來呈現。

在某些實施例中，共振峰強化模組 316 收到線頻譜對且執行額外處理來製造加強全極點濾波器 326。此強化全極點濾波器 326 是強化濾波器之一範例，其可應用於輸入聲音訊號的呈現，用以產生更清晰的聲音訊號。在一實施例中，共振峰強化模組 316 會以在共振峰頻率中強調頻譜波峰的方式來調整線頻譜對。請參考圖 4，範例繪圖 400 包括一頻率強度頻譜 412(實線)，其具有以波峰 414 與波峰 416 識別的共振峰位置。共振峰強化模組 316 會調整這些波峰 414 與波峰 416，以產生一新頻譜 422(以虛線近似)，其中此新頻譜在相同或大體上相同共振峰位置，具有波峰 424 與波峰 426，但其有更高的增益。在一實施例中，如垂直條 418 所示，共振峰強化模組 316 會藉由降低線頻譜對之間的距離來增加波峰的增益。

在某些實施例中，對應於共振峰頻率的線頻譜對會被調整，以表示緊靠在一起的頻率，由此增加每個波峰的增益。當線性預測多項式在單位圓的任何地方有複數根時，在一些實施例中，線性頻譜多項式的根只在單位圓上。因此，線頻譜對有一些相較於線性預測編碼直接量化更好的特質。由於根在一些實現中交錯，若根單調地增加，即可

達成濾波器的穩定度。不像線性預測編碼係數，線頻譜對可能不會對量化雜訊太過敏感，因此可達到濾波器的穩定度。兩個根越靠近，在相應頻率下，濾波器越會共振。因此，減低對應於線性預測編碼頻譜波峰之兩根(一線頻譜對)之間距離，可有利地在該共振峰位置中增加濾波器增益。

在一實施例中，共振峰強化模組 316 會藉由使用相位轉換運算方式將調變因子  $\delta$  施予在每個根(例如，乘上  $e^{j\Omega\delta}$ )，以降低波峰之間的距離。改變  $\delta$  值可導致此些根，沿著單位圓移動，以使彼此更接近或是更分開。因此，對於一對線性頻譜根而言，藉由施予一個正的調變因子  $\delta$ ，第一個根可移動至更靠近第二個根，而藉由應用一個負的調變因子  $\delta$ ，第二根可移動至更靠近第一個根。在一些實施例中，藉由某些量，根與根之間的距離可被縮小，用以達到要求的強化，例如距離縮減約 10%、或約 25%、或約 30%、或約 50%或其他數值。

根的調整也可藉由聲音強化控制器 222 來控制。如圖 2 所示，聲音強化模組 222 會根據麥克風輸入訊號 204 的雜訊等級，調整施予至聲音清晰度強化的量。在一實施例中，聲音強化控制器 222 輸出控制訊號至適應性聲音強化控制器 220，並且共振峰強化模組 316 可使用適應性聲音強化控制器 220 來調整施予至線性頻譜根的共振峰強化的量。在一實施例中，共振峰強化模組 316 會根據控制訊號來調整調變因子  $\delta$ 。因此，指示應施予更多強化的控制訊號(例如，因為更多雜訊)會使共振峰強化模組 316 改變調

變因子  $\delta$ ，以讓根與根更互相靠近在一起，反之亦然。

請再參考圖 3，共振峰強化模組 316 可將調整過之線頻譜對映射回線性預測編碼係數(點陣或直接型式)，以產生強化全極點濾波器 326。然而，在一些實施例中，並不需要執行此種映射，但更準確地說，強化全極點濾波器 326 可以線頻譜對當作係數來實作。

為了強化輸入語音，在一些實施例中，強化全極點濾波器 326 會在從輸入語音訊號合成的激發訊號 324 上操作。在一些實施例中，此合成是藉由施予全零點濾波器 322 至輸入語音來執行，以產生激發訊號 324。全零點濾波器 322 是由線性預測編碼分析模組 312 所建立且可以當成反向濾波器，其中此反向濾波器是由線性預測編碼分析模組 312 建立之全極點濾波器的反轉換。在一實施例中，全零點濾波器 322 以線頻譜對來實現，並且此線頻譜對是由線性預測編碼分析模組 312 計算而得。藉由施予全極點濾波器之反轉換至輸入語音，且之後施予強化全極點濾波器 326 至反轉換語音訊號(激發訊號 324)，可使原本輸入語音訊號可被回復(至少大部分的原本輸入語音訊號)且強化。當全零點濾波器 322 的係數與強化全極點濾波器 326 可依批次轉換(或甚至逐個樣本)，即使在吵雜的環境，輸入語音的共振峰可被適應性地追蹤且強調，由此增加語音清晰度。因此，在某些實施例中，強化語音可使用分析合成技術來產生。

圖 5 繪示適應性聲音強化模組 520 之另外實施例，其

中適應性聲音強化模組 520 包括圖 3 之適應性聲音強化模組 320 的所有特徵，再加上額外特徵。特別是，在所描述之實施例中，圖 3 的強化全極點濾波器 326 會被施予兩次：一次是施予至激發訊號 324(526a)，且另一次是施予至輸入語音(526b)。施予強化全極點濾波器 526b 至輸入語音可產生一訊號，其具有接近輸入語音頻譜的平方的頻譜。藉由組合器 528，將此接近頻譜平方訊號加入至強化激發訊號輸出相加，以產生強化語音輸出。選擇性增益方塊 510 可被提供來調整頻譜平方訊號。(雖然增益使被施予至頻譜平方訊號，然而，增益也可被施予至強化全極點濾波器 526a 的輸出，或被施予至濾波器 526a 與 526b 兩者的輸出)。使用者介面控制可被提供給使用者以調整增益 510，此使用者諸如組成適應性聲音強化模組 320 裝置的廠商或裝置終端使用者。更多增益被施予至頻譜平方訊號會增加訊號刺耳程度，其在特別吵雜的環境或許會增加清晰度，但在較少雜訊的環境會太刺耳。因此，提供使用者控制可以讓使用者能夠調整強化語音訊號的刺耳感覺。在一些實施例中，此增益 510 可根據環境雜訊輸入，來被聲音強化控制器 222 自動地控制。

在某些實施例中，適應性聲音強化模組 320 或 520 以少於所示之所有區塊還少的區塊來實現。在其他實施例中，額外的區塊或濾波器也可加在適應性聲音強化模組 320 或 520。

## VI. 時間包絡塑形器

在一些實施例中，圖 3 中經強化全極點濾波器 326 修改之聲音訊號或如圖 5 中組合器 528 的輸出可提供至時間包絡塑形器 332。經由在時域中的時間包絡塑形，時間包絡塑形器 332 可強化非語音的聲音(包括短音)。在一實施例中，時間包絡塑形器 332 會強化中間頻率，包括頻率低於約 3kHz(且選擇性地高於低音頻率)。時間包絡塑形器 332 同樣可加強中頻以外的頻率。

在某些實施例中，在時域中，時間包絡塑形器 332 可藉由首先偵測從強化全極點濾波器 326 輸出訊號的包絡來強化時間頻率。時間包絡塑形器 332 可使用各種方法來偵測包絡。一個範例為最大值追蹤，其中時間包絡塑形器 332 可將訊號分割成視窗區段，並且然後從每個視窗區段中選擇最大或峰值。時間包絡塑形器 332 可用一直線或曲線將每個值間的最大值連接起來，用以形成包絡。在一些實施例中，為了增加語音清晰度，時間包絡塑形器 332 會將訊號切割成適當數目的頻帶且對每個頻帶執行不同塑形。

範例視窗尺寸可包括 64、128、256 或 512 個樣本，雖然其他視窗尺寸也可選擇(包括非 2 的次方的視窗尺寸)。一般來說，大視窗尺寸可延伸欲強化的時間頻率至較低頻率。再者，其他技術亦可被用來偵測訊號的包絡，例如 Hilbert Transform-related 技術與自解調技術(例如，將訊號平方與通過低通濾波器)。

一旦偵測到包絡，時間包絡塑形器 332 會調整包絡的形狀，以選擇性地銳化或平滑包絡的外觀。在第一個步驟

中，時間包絡塑形器 332 會根據包絡的特性計算增益。在第二個步驟中，時間包絡塑形器 332 會施予增益至實際訊號的取樣，以達到所預期的效果。在一實施例中，所預期的效果是銳化語音的暫態部分，以強調非聲帶語音(如某些子音像是“s”與“t”)，從而增加語音清晰度。在其他應用中，平滑語音對於柔和語音也是有幫助的。

圖 6 繪示時間包絡塑形器 632 的更詳細實施例，其可實現圖 3 時間中包絡塑形器 332 的特徵。時間包絡塑形器 632 也用於不同應用中，其獨立於上述之適應性聲音強化模組。

時間包絡塑形器 632 收到輸入訊號 602(例如，從濾波器 326 或結合器 528 中所接收的訊號)。然後，時間包絡塑形器 632 使用帶通濾波器 610 或諸如此類的濾波器來細分輸入訊號 602 至複數個頻帶。任何數目的頻帶皆可被選擇。如同一範例，時間包絡塑形器 632 可以分割輸入訊號 602 至四個頻帶，包括從約 50Hz 到約 200Hz 的第一頻帶，從約 200Hz 到約 4kHz 的第二頻帶，從約 4kHz 到約 10kHz 的第三頻帶，以及從約 10kHz 到 20kHz 的第四頻帶。在其他實施例中，時間包絡塑形器 332 不會分割訊號至多頻帶，但反而在整體訊號上操作。

最低頻可為低音頻或使用次頻帶通濾波器 610a 而得的次頻。次頻帶對應於典型地由重低音所產生的頻率。在以上的範例，最低頻帶約 50Hz 到約 200Hz。次頻帶通濾波器 610a 的輸出會被提供至一次補償增益方塊 612，其施

予增益至在次頻帶中的訊號。如以下所詳述，增益可被施予至其他頻帶，以銳化或強調輸入訊號 602 的外觀。然而，除了次頻帶 610a 之外，施予此增益會增加頻帶 610b 的能量，導致低音輸出可能減少。用以補償此減少低音效應，次補償增益方塊 612 會根據施予至其他頻帶 610b 的增益量來施予增益至此頻帶 610a。次補償增益的值等於或是大概等於原來輸入訊號 602(或其包絡)與銳化輸入訊號的能量差值。次補償增益可由增益區塊 612 藉加法、平均或是其他組合相加能量或是施予至其他頻帶 610b 的增益來計算。次補償增益也可藉由選擇施予至其中一個頻帶 610b 的波峰增益與使用用於次補償增益的值或諸如此類的值的增益方塊 612 來計算。在另一實施例中，然而，次補償增益亦可以是固定值。次補償增益方塊 612 的輸出會被提供給結合器 630。

每個其他帶通濾波器 610b 的輸出可被提供至實現上述任何包絡偵測演算法的包絡偵測器 622。例如，包絡偵測器 622 可執行最大值追蹤或諸如此類的動作。包絡偵測器 622 的輸出會被提供至包絡塑形器 624，其可調整包絡的形狀，用以選擇性地銳化或平滑包絡的外觀。每個包絡塑形器 624 提供輸出訊號至結合器 630，其中結合器 630 會結合每個包絡塑形器 624 與次補償增益方塊 612 的輸出，以提供輸出訊號 634。

如圖 7 與圖 8 所示，包絡塑形器 624 提供之銳化效應可藉由操縱每個頻帶(或是沒有細分的全部頻帶)包絡的斜

率來達到。請參考圖 7，範例繪圖 700 顯示部分時域包絡 701。在圖 700 中，時域包絡 701 包括兩部分，第一部分 702 與第二部分 704。第一部分 702 有正斜率，而第二部分 704 有負斜率。因此，兩部分 702，704 會形成波峰 708。在包絡上的點 706、708 以及 710 表示波峰數值，其中此波峰數值由上述視窗或框藉由最大值包絡偵測器偵測而得。第一部分 702 與第二部分 704 表示連結波峰點 706、708 與 710 的線，從而形成包絡 701。當波峰 708 被顯示於包絡 701 的同時，包絡 701 的其他部分(無顯示)可能反而有反折點或零斜率。針對包絡 701 的範例部分所描述的分析也可被實作於包絡 701 的此其他部分。

包絡 701 的第一部分 702 與水平形成角度  $\theta$ 。此角度的邊緣陡度可反映包絡 701 的第一部分 702 與第二部分 704 是否代表語音訊號的暫態部分，其中越陡角度越表示暫態。同樣地，包絡 701 的第二部分 702 與水平形成角度  $\phi$ 。此角度也反映暫態表示的可能性，越陡角度越表示暫態。因此，增加角度  $\theta$  與角度  $\phi$  之中的一個或兩個，可有效地銳化或強調暫態，並且尤其是增加  $\phi$ ，可以導致聲音乾燥(比如，較少回音的聲音)，因為回音可能降低。

藉著調整由第一部分 702 與第二部分 704 所形成的每一條線的斜率可增加角度，以產生具有陡峭或銳化部分 712 與 714 的一新包絡。如圖所示，第一部分 702 的斜率可表示為  $dy/dx_1$ ，同時第二部分 704 的斜率可表示為  $dy/dx_2$ 。一增益會被施予以增加每一個斜率的絕對值(例

如，用於  $dy/dx1$  的正增強，以及用於  $dy/dx2$  的負增強)。此增益可相依於依賴每個角度  $\theta$  與  $\phi$  的值。在某些實施例中，為了銳化暫態，增益值會隨著正斜率增加且隨著負斜率減少。提供至包絡的第一部分 702 的增益調整量可能與提供至包絡的第二部分 704 的增益調整量相同，但此非必須。在一實施例中，用於第二部分 704 的增益的絕對值是大於施予至第一部分 702 的增益，藉此可更銳化聲音。在樣本的波峰，增益可能平滑化，以減少因為突然從正增益轉變到負增益造成的副作用。在某些實施例中，每當上述角度低於門檻時，增益會被施予至包絡。在其他實施例中，每當上述角度高於門檻時，增益會被施予。經計算增益(或用於多個樣本與/或多個頻帶的增益)可構成時間強化係數，其中此時間強化係數會銳化在訊號中的波峰，且藉此強化所選擇的子音或聲音訊號的其他部分。

以下為可實作此些特徵具平滑的範例增益方程式：

gain =  $\exp(gFactor * \delta * (i - mBand \rightarrow prev\_maxXL / dx) * (mBand \rightarrow mGainoffset + Offsetdelta * (i - mBand \rightarrow prev\_maxXL)))$ 。在這個範例方程式中，因為包絡與角度都是以對數近位法來計算，因此增益是角度變化的指數函數。gFactor 控制起奏(attack)或衰退的比例。 $(i - mBand \rightarrow prev\_maxXL / dx)$  表示包絡的斜率，而增益方程式的以下部分表示平滑函數，其從之前增益開始且以當前增益結束：  
 $(mBand \rightarrow mGainoffset + Offsetdelta * (i - mBand \rightarrow prev\_maxXL$

))。雖然人類聽覺系統是根據對數近位法，但指數函數可幫助收聽者更佳地辨別暫態語音。

gFactor 的起奏/衰退函數更被描述於圖 8，其中增加起奏斜率 812 的不同等級被繪示於第一圖 810，且減少衰退斜率 822 的不同等級被繪示表示於第二圖 820。起奏斜率 812 可依以上所述增加斜率，以強調對應於圖 7 的陡峭第一部份 712 的暫態聲音。同樣地，衰退斜率 822 可依以上所述減少斜率，用以進一步強調對應於圖 7 的陡峭第二部份的暫態聲音。

#### V. 範例聲音偵測程序

圖 9 繪示據聲音偵測程序 900 的實施例。雜訊偵測程序 900 可藉由上述之聲音強化系統 110 與 210 之一來實現。在一實施例中，雜訊偵測程序 900 是由聲音動作偵測器 212 實現。

聲音偵測程序 900 會偵測在輸入訊號之中的聲音，如麥克風輸入訊號 204。若輸入訊號包括雜訊而非聲音時，聲音偵測程序 900 會允許聲音強化量根據目前量測到之環境雜訊來做調整。然而，當輸入訊號包括聲音時，聲音偵測程序 900 會使先前量測的環境雜訊來被用於調整聲音強化。使用先前雜訊的量測可有利地避免依據聲音輸入調整聲音強化，同時仍然啟動聲音強化以適應環境雜訊情況。

在程序 900 的方塊 902 中，聲音動作偵測器 212 會收到輸入麥克風訊號。在方塊 904 中，聲音動作偵測器 212 會執行麥克風訊號的聲音作動分析。聲音動作偵測器 212

可使用各種技術來偵測聲音動作。在一實施例中，聲音動作偵測器 212 會偵測雜訊動作，而非聲音，並且指出非雜訊活動的週期是對應於聲音。聲音動作偵測器 212 可使用以下技術的任何組合或諸如此類的技術，以偵測聲音與/或雜訊：訊號的統計分析(例如使用標準差、變異數等)，低頻帶能量與高頻帶能量的比例、越零率、頻譜流量或其他頻域近似或自相關。再者，在一些實施例中，聲音動作偵測器 212 使用在美國專利第 7,912,231 中所描述的一些或全部雜訊偵測技術偵測雜訊，此美國專利申請日為 2006 年 4 月 21 日，發明名稱是”減少聲音訊雜訊的系統與方法”，並且其揭露的完整內容可作為本案的參考。

若在方塊 906 中偵測到訊號包括聲音時，聲音動作偵測器 212 會使聲音強化控制器 222 使用之前雜訊緩衝以控制適應性聲音強化模組 220 的聲音強化。雜訊緩衝可包括由聲音動作偵測器 212 或聲音強化控制器 222 儲存的一或多個麥克風輸入訊號 204 的雜訊樣本的方塊。假設從先前樣本被儲存在雜訊緩衝中的時間起，環境雜訊沒有被顯著改變的情況下，則從輸入訊號 204 的先前部分中儲存的先前雜訊緩衝可被使用。因為對話中的中斷時常發生，此假設在許多場合中可能正確。

另一方面，若訊號不包括聲音，聲音動作偵測器 212 會使聲音強化控制器 222 使用目前雜訊緩衝來控制適應性聲音強化模組 220 的聲音強化。目前雜訊緩衝可表示一或多個近期收到的雜訊樣本的方塊。在方塊 914 中，聲音動

作偵測器 212 會判斷額外訊號是否已被接收。若是，程序 900 會回至方塊 904。反之，程序 900 會結束。

因此，在某些實施例中，聲音偵測程序 900 可減少聲音輸入調變或其他自我啟動施予至遠端聲音訊號的聲音清晰度強化的等級的非預期影響。

## VI. 範例麥克風校正程序

圖 10 繪示麥克風校正程序 1000 的實施例。麥克風校正程序 1000 是至少部分實作在上述之聲音強化系統 110 與 210 的任何一個中。在一實施例中，麥克風校正程序 1000 是至少部分實作在麥克風校正模組 234 中。如圖所示，程序 1000 的一部分可於實驗室或設計場中實現，同時程序 1000 的剩餘部份，可在場區中實現，例如像是製造聲音強化系統 110 或 210 裝置公司的場區。

如上所述，麥克風校正模組 234 可計算與儲存一或多個校正參數，其可校正參數調整施予至麥克風輸入訊號 204 的增益，以使麥克風全體增益對於一些或是全部裝置來說是相同或是幾乎相同。相對地，現有用來使所有麥克風裝置增益相等的方法會趨於不一致，導致不同雜訊等級在不同裝置中，而啟動聲音強化。在現有麥克風校正方法中，一個現場工程師(例如，在裝置製造公司或其他地方)藉由在測試裝置中啟動一回放喇叭以製造被電話的麥克風或其他裝置所蒐集的雜訊來使用試誤法。現場工程師接下來會嘗試校正麥克風，以使麥克風訊號在聲音強化控制器 222 視為到達雜訊門檻的等級，從而導致聲音強化控制器

222 觸發或啟動聲音強化。每個現場工程師有對於哪種雜訊程度麥克風應該要蒐集，來達到觸發聲音強化，有不同的感覺，因此會產生不一致的問題。再者，很多麥克風有寬廣的增益範圍(例如，-40dB 到+40dB)，且因此當調整麥克風時要找到一個精確的增益數值來使用是困難的。

麥克風校正程序 1000 可計算增益值給每個麥克風，其相對於現場工程師試誤法來說，麥克風會更為一致。從實驗室開始，在方塊 1002，雜訊隨著測試裝置被輸出，其中此裝置可能是擁有或耦接適當喇叭的任何計算裝置。在方塊 1004 中，此雜訊訊號會被錄製成為參考訊號，且在方塊 1006 中平滑能量會從標準參考訊號中被計算。此以 RefPwr 表示的平滑能量可為黃金標準值，其用於在場區中的自動麥克風校正。

在場區中，自動校正可使用黃金參考值 RefPwr 來做動。在方塊 1008 中，例如，現場工程師使用測試裝置以標準音量播放參考訊號。參考訊號是以相同於在方塊 1002 中在實驗室裡所播放雜訊的音量來被播放。在方塊 1010 中，麥克風校正模組 234 會記錄從測試麥克風接收到的聲音。接下來，在方塊 1012 中，麥克風校正模組 234 會紀錄訊號的平滑能量，其以 CaliPwr 表示。在方塊 1014 中，麥克風校正模組 234 會根據參考訊號的能量與所記錄之訊號的能量來計算麥克風偏移量。例如如下式所示： $\text{MicOffset} = \text{RefPwr}/\text{CaliPwr}$ 。

在方塊圖 1016 中，麥克風校正模組 234 會將麥克風

偏移量設定為用於麥克風的增益。當麥克風輸入訊號 204 被接收時，此麥克風偏移量用作為施予至麥克風輸入訊號 204 的校正增益。基此，在所有裝置中，用於相同門檻等級之使聲音強化控制器 222 驅動聲音強化的雜訊等級會相同或近似相同。

## VII. 術語

從揭露，此述之外的很多其他變化是顯而易見的。例如，根據實施例中，某些動作、事件或於此描述之演算法的任何功能可以不同順序執行、可相加、組合或全部一起省略(例如，對演算法的實行，非全部描述的動作或事件皆是必要的)。此外，在某些實施例中，動作或事件可同時地執行，例如，經由多執行緒程序、中斷處理或多處理器或是多核處理器或是其他平行結構，而不是連續地。另外，不同的任務或程序可藉由可一起運行的不同機器與/或計算系統來完成。

各種說明邏輯方塊、模組與演算法步驟連接在一起說明此實施例發明可被以電器硬體、電腦軟體或兩者結合來實現。為了清楚說明此硬體與軟體的可交換性，不同的說明元件、方塊、模組與步驟會大部分依功能描述如下。不管此功能以硬體或是軟體實現，此功能依據整體系統的特定應用與設計來限制。例如，車輛管理系統 110 或 210 可藉由一或多個計算機系統或擁有一或多個處理器的計算機系統，來加以實現。描述的功能可以針對各個特定的應用，以各種不同的方法實現，但如此實現結論不應該解釋為脫

離本揭露的範圍。

在實施例中的各種說明邏輯方塊與模組可藉由機器來實作或執行，例如，一般用途的處理器、數位訊號處理器、特殊應用積體電路(ASIC)、現場可程式化邏輯閘陣列(FPGA)或其他可程式化邏輯裝置、離散閘或電晶體邏輯、分散硬體組件或任何設計用以執行此描述之功能的組合。一般用途處理器可為一微處理器，在不同情況下，此處理器可為控制器、微控制器或狀態機，相同或諸如此類的組合。處理器也可由計算機裝置的組合來實現，例如，數位訊號處理器與微處理器的組合、複數個微處理器、一或多個結合數位訊號處理器核心的微處理器、或任何其他此類的結構。計算環境可包括任何形式的計算機系統，包括：以微處理器為基礎的電腦系統、大型電腦，數位訊號處理器、可攜式計算機裝置、個人記事本、裝置控制器，以及在設備中的計算機的引擎，僅舉數例，但不限於此。

與揭露於此之實施例有關的方法、程序或演算法之步驟描述，可直接以硬體、由處理器執行的軟體模組、或兩者的組合來實施。軟體模組可存在於隨機存取記憶體(RAM)中、快閃記憶體(Flash memory)、唯讀記憶體(ROM memory)、可抹除可程式唯讀記憶體(EPROM memory)、電性可抹除可程式唯讀記憶體(EEPROM memory)、暫存器、硬碟、可移除式碟片、CD-ROM，或任何其他型式非暫時性電腦可讀取媒體，或該領域中已知的物理計算機儲存。一範例儲存媒體可結合到處理器，如此處理器可從此儲存

媒體讀取或寫入此儲存媒體。換句話說，此儲存媒體可被整合到處理器。處理器與儲存媒體可歸屬於特殊應用積體電路。特殊應用積體電路可歸屬於使用者終端機。換句話說，處理器與儲存媒體可歸屬成使用者終端機的離散元件。

於此使用之條件語言，例如，“可以”、“能”、“可能”、“比如”諸如此類，除非用其他方法特別聲明，或在內文中用其他方式理解，否則通常打算表達某些實施例包括某些特徵、元件與/或狀態，而其他實施例不包括。因此，此類條件語言通常不打算表示：對於一或多個實施例而言，此特徵、元件與/或狀態是必要的；或者，一或多個實施例必然地要包括用以決定之邏輯、有無創作者輸入或提示，不論這些特徵、元件與/或狀態被包括或被執行於任何特殊實施例與否。此“包括”、“包括”、“有”與諸如此類用詞皆同意，且使用在一開放端方式，且不排除額外的元件，特徵，動作，操作，等等諸如此類。同樣地，“或”，該詞用於該包括意義(非排除意義)，例如，當使用來連接列表元件，此詞“或”表示一、一些、或全部在表中的元件。再者，此詞“每個”，如此所用，除了原本意義之外，可用此詞“每個”表示一組元件的任何子群組。

隨著以上詳細描述所示、描寫、與指出應用在不同實施例中新穎的特徵，可理解的是，允許在無背離發明精神的情況下，各種刪除、取代、以及改變裝置形式與描述的細節或演算法。如被認清的，本發明描述於此的某些實施例可在不提供所有於此提出的特徵與好處的形式的範圍內

實施，其中一些特徵可與其他分開使用或實施。

### 【圖式簡單說明】

圖 1 是繪示可實現聲音強化系統的行動電話環境的實施例。

圖 2 是繪示聲音強化系統的更詳細實施例。

圖 3 是繪示適應性聲音強化模組的實施例。

圖 4 是繪示語音頻譜的範例。

圖 5 是繪示適應性聲音強化模組的另一實施例。

圖 6 是繪示時間包絡塑形器的實施例。

圖 7 是繪示一例圖之時域語音包絡的範例繪圖。

圖 8 是繪示起奏與衰退包絡的範例繪圖。

圖 9 是繪示聲音檢測程序的一實施例。

圖 10 是繪示麥克風校正程序的一實施例。

### 【主要元件符號說明】

100：行動電話環境

102：輸入聲音訊號

104：呼叫者電話

106：發射機

108：受話者電話

110：聲音強化系統

112：麥克風輸入訊號

114：強化輸出訊號

202：聲音輸入訊號

204：麥克風輸入訊號

- 210：聲音強化系統
- 212：聲音動作偵測器
- 220：適應性聲音強化模組
- 222：聲音強化控制器
- 226：額外強化控制
- 230：輸出增益控制器
- 232：適應性等級控制
- 234：麥克風校正模組
- 240：截波抑制模組
- 250：輸出
- 310：前置濾波器
- 312：線性預測編碼分析模組
- 314：對映模組
- 316：共振峰強化模組
- 320、520：適應性聲音強化模組
- 322：全零點濾波器
- 324：激發訊號
- 326、526a、526b：強化全極點濾波器
- 332：時間包絡塑形器
- 400、700：範例繪圖
- 412：頻率強度頻譜
- 414、416、418、424、426：波峰
- 422：新頻譜
- 602：輸入訊號

- 610a：次頻帶帶通濾波器
- 610b：其他帶帶通濾波器
- 612：次增益補償方塊
- 622：包絡偵測器
- 624：包絡塑形器
- 630、528：結合器
- 632：時間包絡塑形器：
- 634：輸出訊號
- 701：時間包絡
- 702：第一部分
- 704：第二部分
- 706、708、710：鋒值點
- 712、714：陡峭部分或銳化部分
- 810：第一圖
- 812：起奏斜率
- 820：第二圖
- 822：衰退斜率
- 900：聲音偵測程序
- 902、904、906、908、910、912、914、1002、1004、  
1006、1008、1010、1012、1014、1016：方塊
- 1000：麥克風校正程序

## 七、申請專利範圍：

1. 一種調整聲音清晰度強化的方法，該方法包括：

藉由使用一前置濾波器來接收一輸入聲音訊號；

藉由根據該前置濾波器的輸出而使用一線性預測編碼分析模組以一線性預測編碼程序來獲取該輸入聲音訊號的一頻譜圖，該頻譜圖包括一或多個共振峰頻率；藉由使用一或多個處理器來調整該輸入聲音訊號的該頻譜圖以產生一強化濾波器的係數，以強調該一或多個共振峰頻率；

施予一反向濾波器至該輸入聲音訊號，以獲得一激發訊號；

施予該強化濾波器至該激發訊號來產生具有強化共振峰頻率的一第一已修改聲音訊號；

施予該強化濾波器至該輸入聲音訊號，以產生一第二已修改聲音訊號；

將該第一已修改聲音訊號的至少一部份與該第二已修改聲音訊號的至少一部份進行組合，以產生一經組合已修改聲音訊號；

根據該經組合已修改聲音訊號偵測一包絡；

分析該經組合已修改聲音訊號的該包絡，以決定一或多個時間強化係數；以及

施予該一或多個時間強化係數至該經組合已修改聲音訊號來產生一輸出聲音訊號。

2. 如申請專利範圍第 1 項所述之調整聲音清晰度強化的方法，其中所述施予該一或多個時間強化係數至該經

組合已修改聲音訊號的步驟包括在該經組合已修改聲音訊號的該一或多個包絡中銳化波峰，以強調在該經組合已修改聲音訊號中的所選擇的子音。

3. 一種調整聲音清晰度強化的系統，該系統包括：

一分析模組，用以獲取一輸入聲音訊號的至少一部份的一頻譜圖，該頻譜圖包括一或多個共振峰頻率；

一反向濾波器，用以被施予至該輸入聲音訊號，以獲得一激發訊號；

一共振峰強化模組，用以產生一強化濾波器的係數，其中該強化濾波器用以強調該一或多個共振峰頻率；

該強化濾波器，用以藉由一或多個處理器來被施予至該激發訊號，以產生一第一已修改聲音訊號，該強化濾波器更藉由該一或多個處理器來被施予至該輸入聲音訊號，以產生一第二已修改聲音訊號；

一組合器，用以組合該第一已修改聲音訊號的至少一部份與該第二已修改聲音訊號的至少一部份，以產生一經組合已修改聲音訊號；以及

一時間包絡塑形器，用以至少部分地依據在該經組合已修改聲音訊號的一或多個包絡來施予一時間強化至該經組合已修改聲音訊號。

4. 如申請專利範圍第 3 項所述之調整聲音清晰度強化的系統，其中該分析模組更用以使用一線性預測編碼技術來獲取該輸入聲音訊號的該頻譜，其中該線性預測編碼技術用以產生對應該頻譜圖的係數。

5. 如申請專利範圍第 4 項所述之調整聲音清晰度強化的系統，更包括一對映模組，用以映射該係數至線頻譜對。

6. 如申請專利範圍第 5 項所述之調整聲音清晰度強化的系統，更包括修改該線頻譜對，以增加在對應該共振峰頻率的該頻譜圖的增益。

7. 如申請專利範圍第 3 項所述之調整聲音清晰度強化的系統，其中該時間包絡塑形器更用以細分該經組合已修改聲音訊號成複數個頻帶，且其中該一或多個包絡對應至用於該些頻帶中的其中至少一些的一包絡。

8. 如申請專利範圍第 3 項所述之調整聲音清晰度強化的系統，更包括一聲音強化控制器，用以至少部分地根據在一輸入麥克風訊號中所偵測到的環境雜訊的一量來調整該強化濾波器的一增益。

9. 如申請專利範圍第 8 項所述之調整聲音清晰度強化的系統，更包括一聲音動作偵測器，用以偵測在該輸入麥克風訊號中的聲音且控制該聲音強化控制器以回應所偵測到的聲音。

10. 如申請專利範圍第 9 項所述之調整聲音清晰度強化的系統，其中該聲音動作偵測器更用以根據回應在該輸入麥克風訊號中對偵測到的聲音敏感的一先前雜訊輸入來使該聲音強化控制器調整該強化濾波器的該增益。

11. 如申請專利範圍第 8 項所述之調整聲音清晰度強化的系統，更包括一麥克風校正模組，用以設定一麥克風

的一增益，其中該麥克風用以接收該輸入麥克風訊號，其中該麥克風校正模組更用以根據至少部分的一參考訊號與一錄音雜訊訊號來設定該增益。

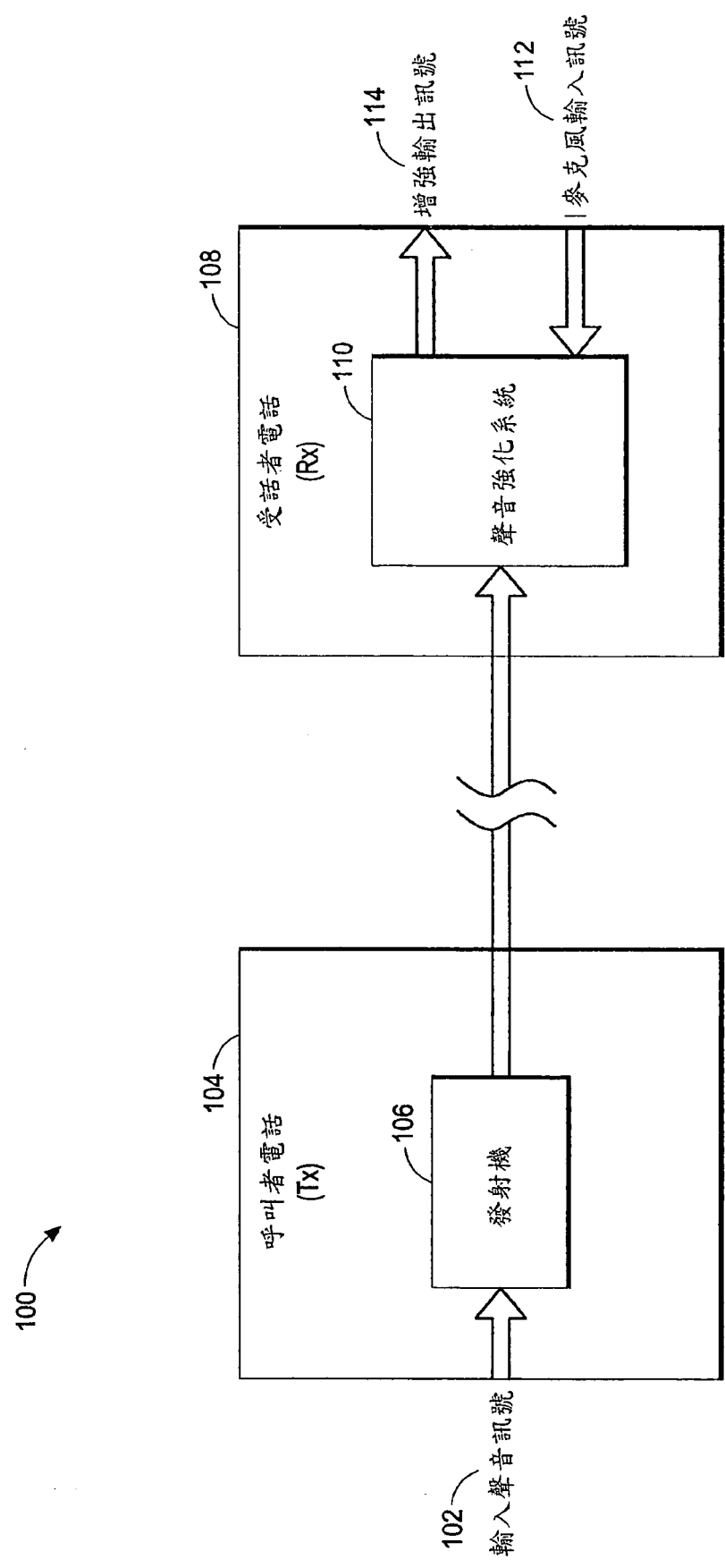


圖 1

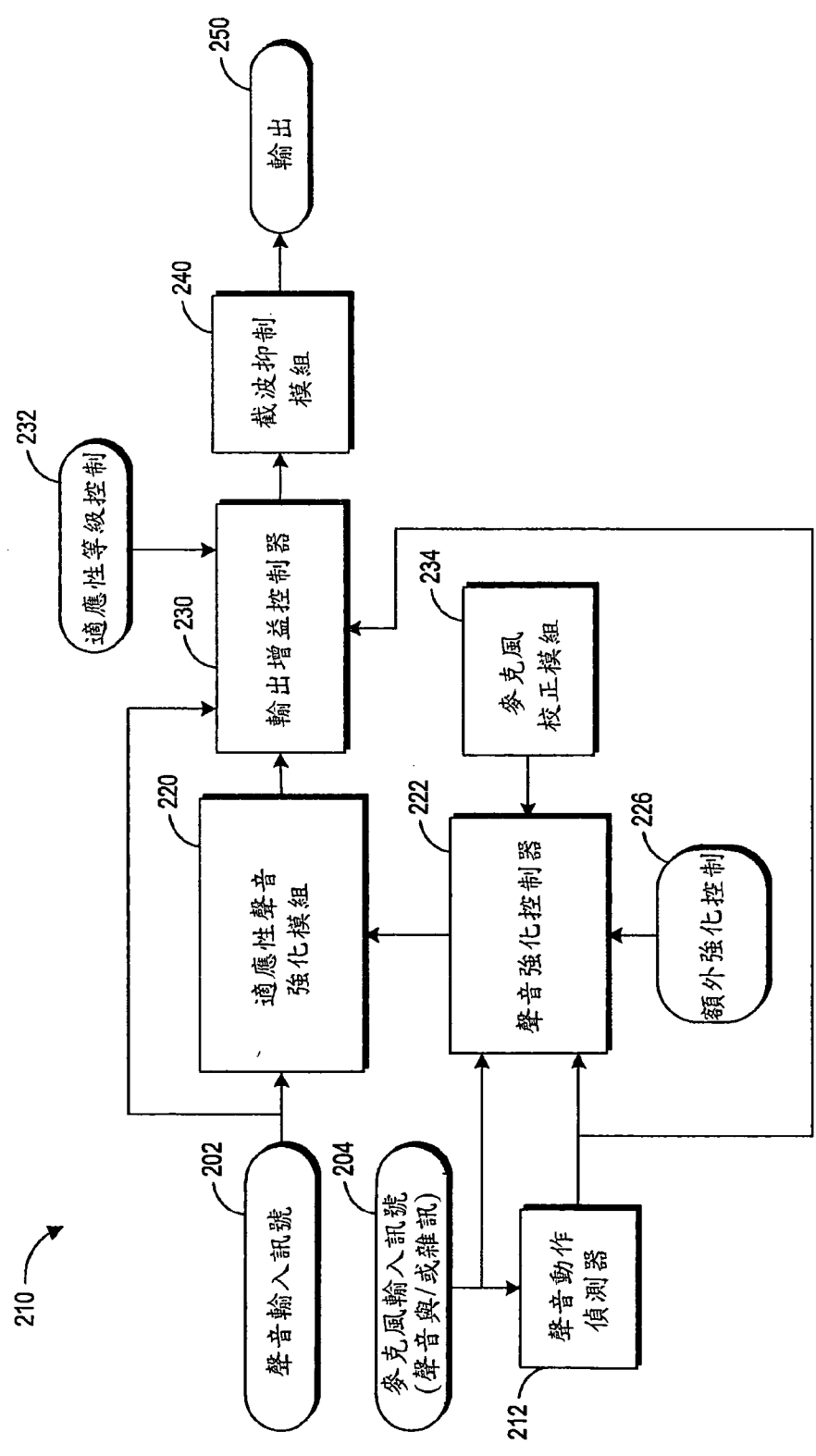


圖 2

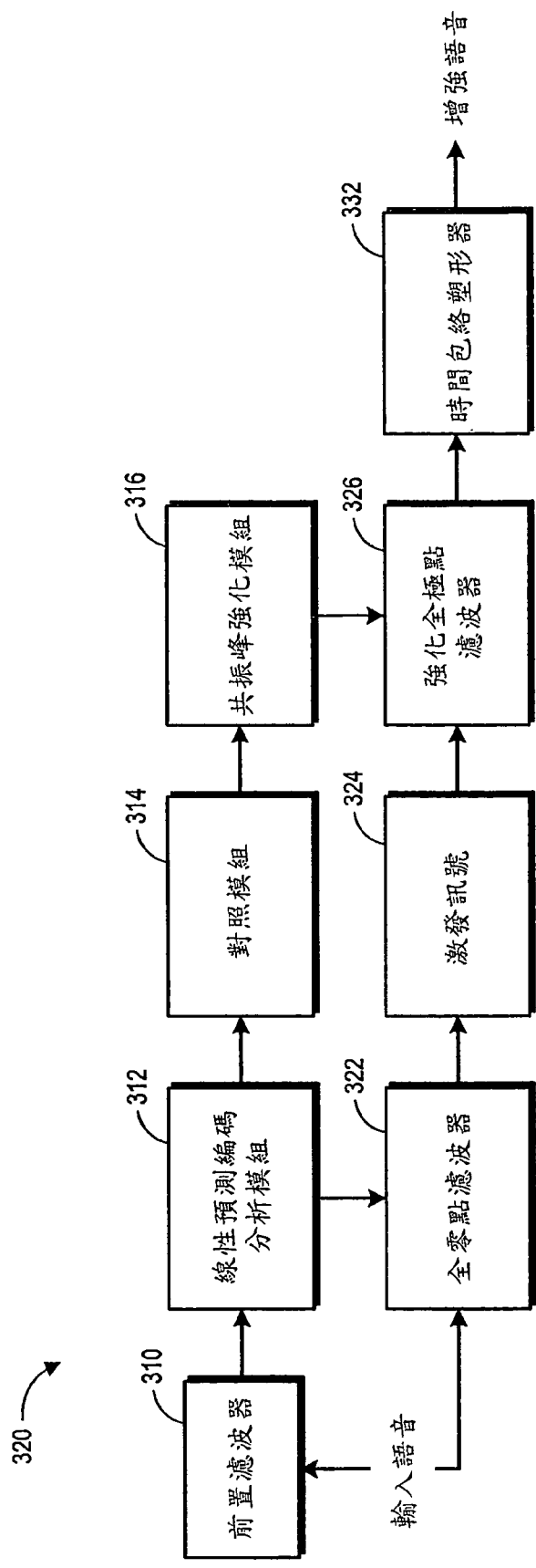


圖 3

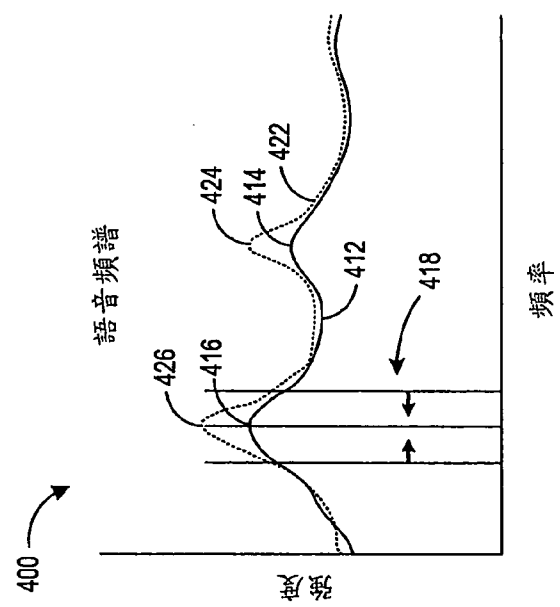


圖 4

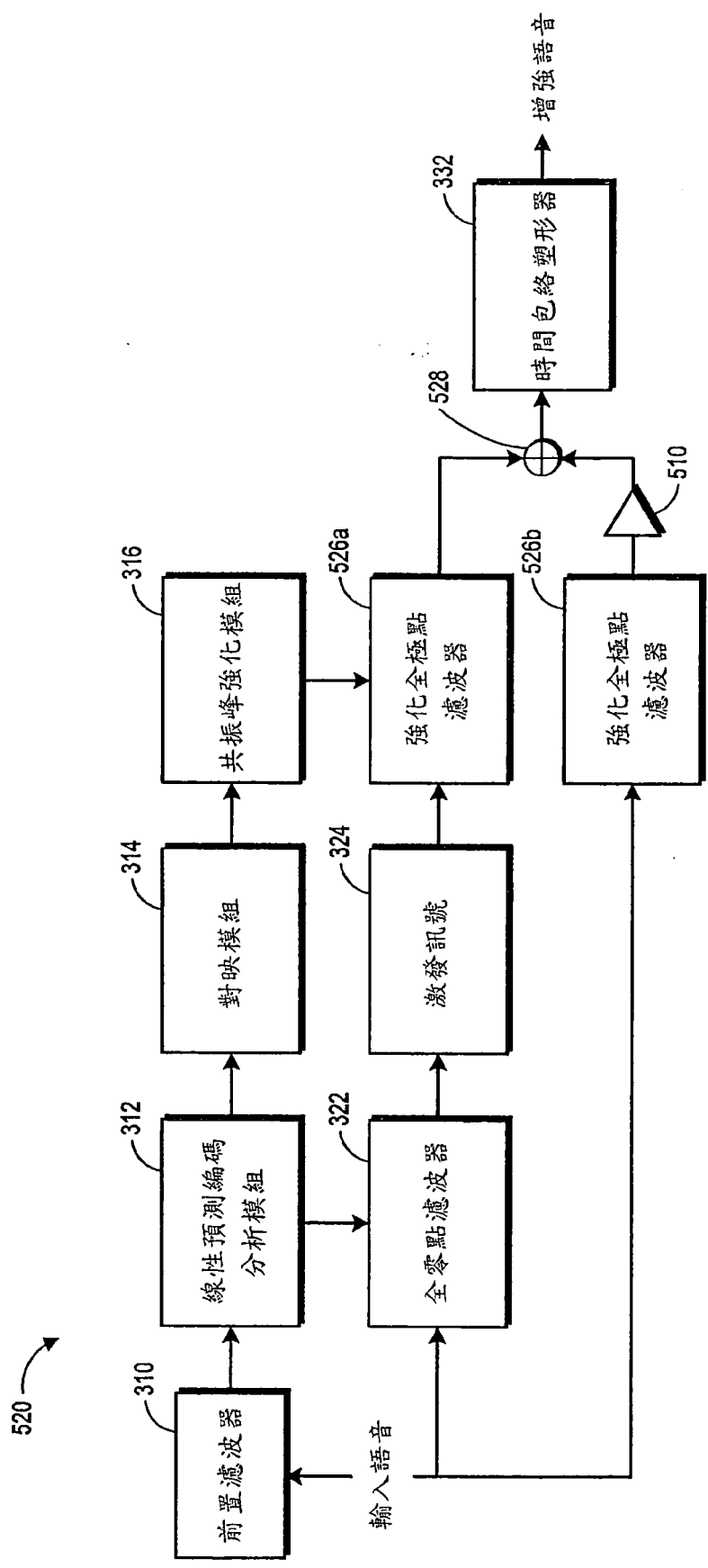


圖 5

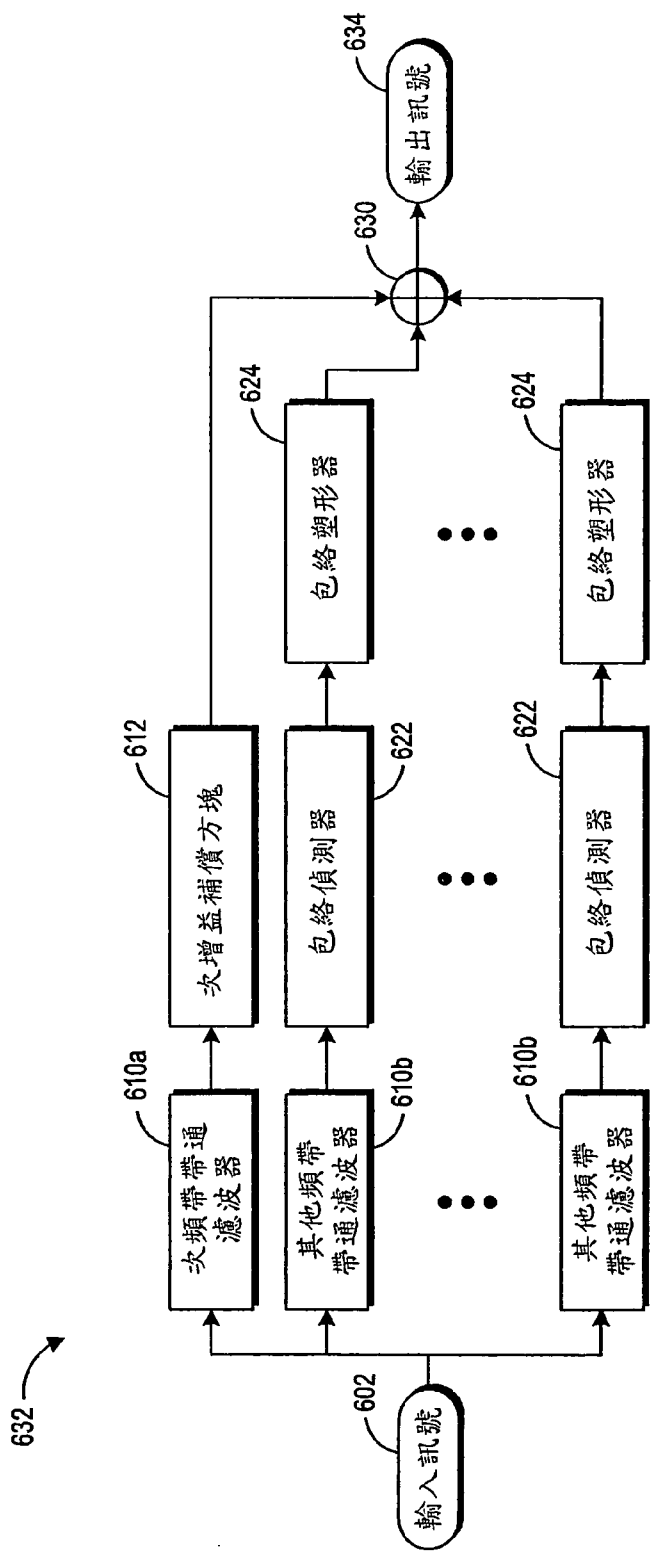


圖 6

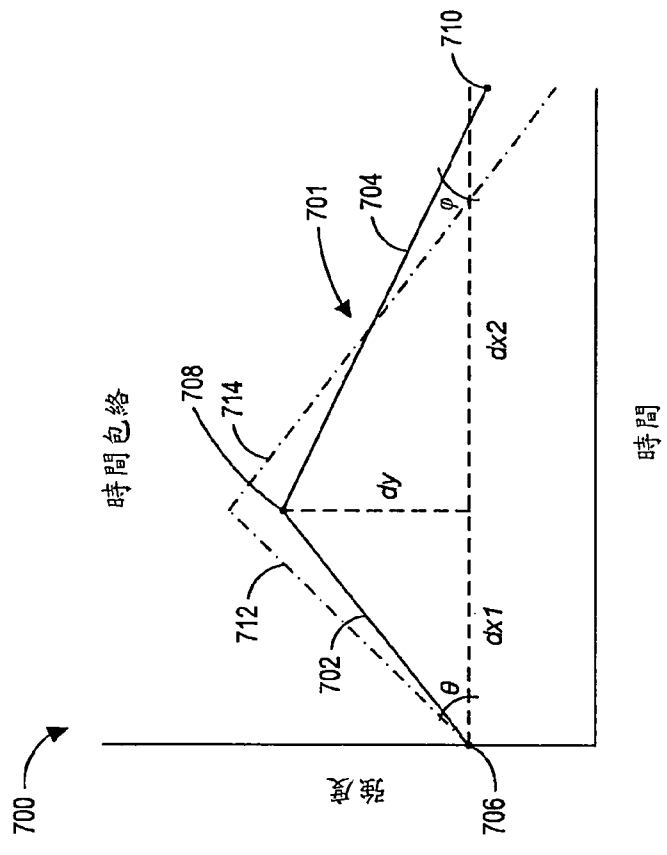


圖 7

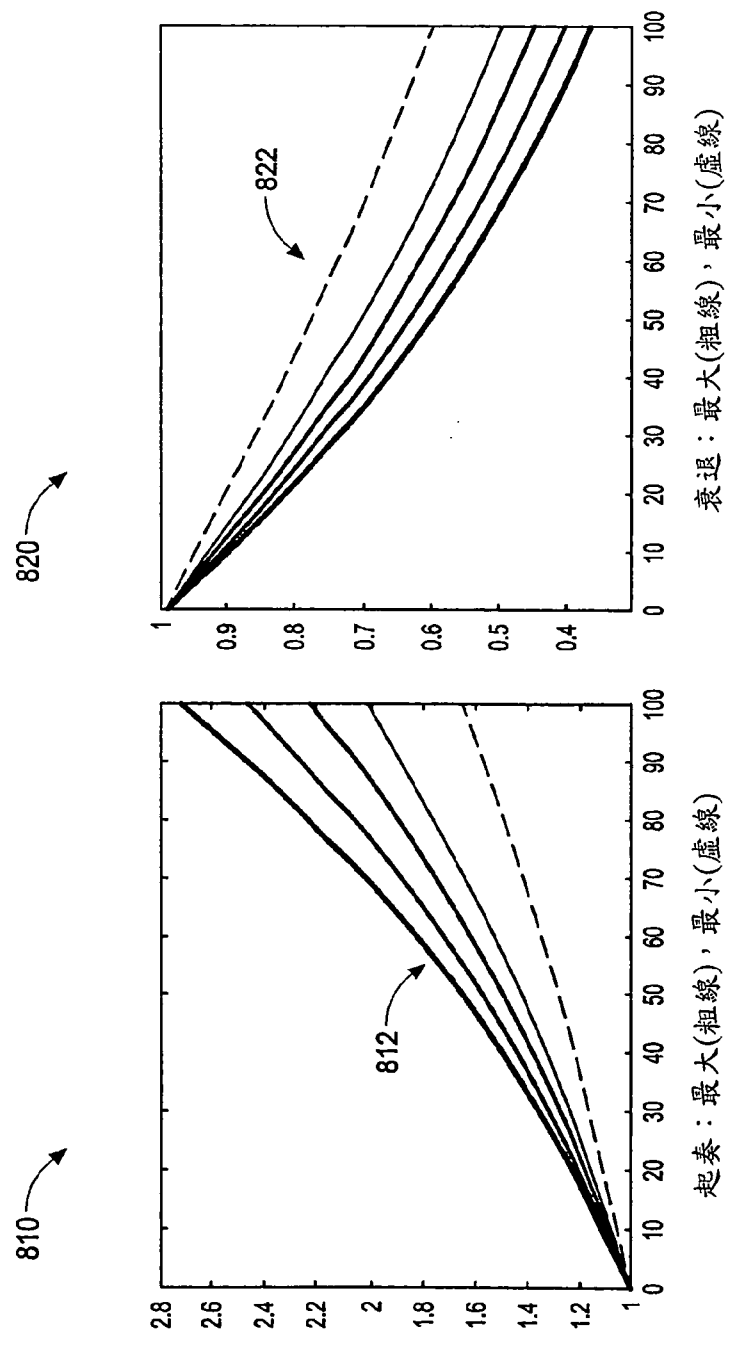


圖 8

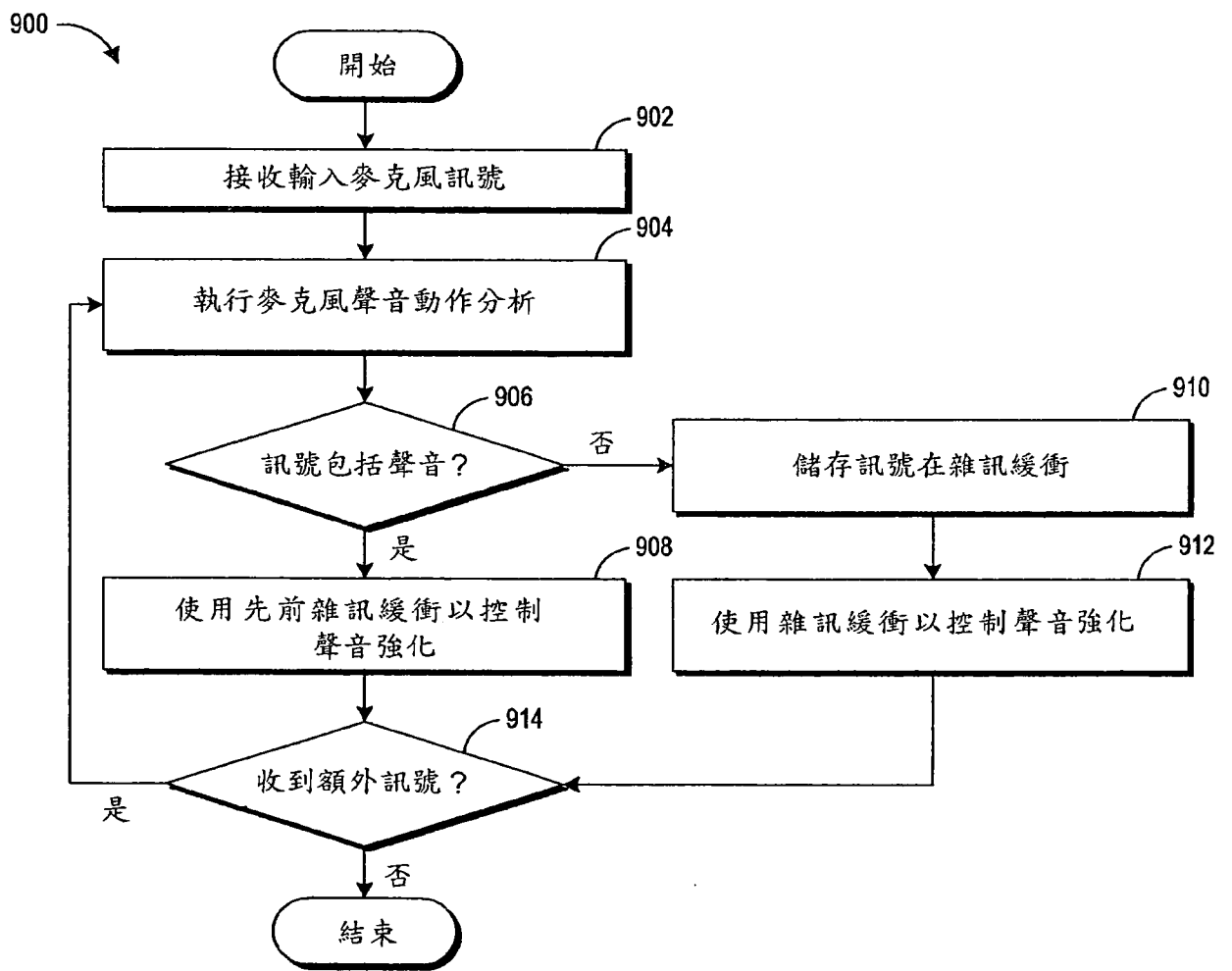


圖 9

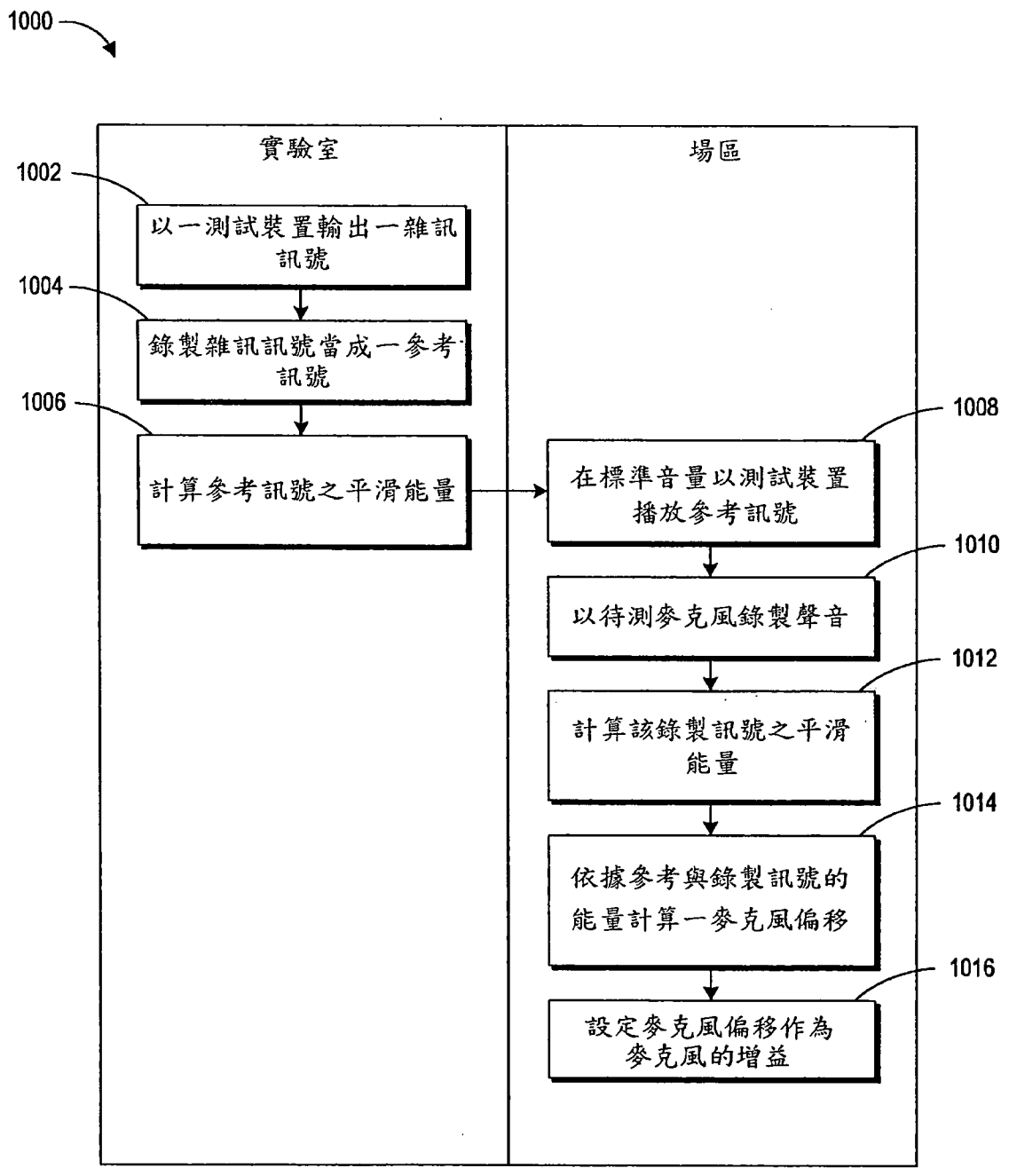


圖 10