



(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 I466107 B

(45) 公告日：中華民國 103 (2014) 年 12 月 21 日

(21) 申請案號：100115214 (22) 申請日：中華民國 100 (2011) 年 04 月 29 日

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

(30) 優先權：2010/04/29 美國 61/329,322
2010/07/08 美國 12/832,920(71) 申請人：視聽公司 (美國) AUDIENCE, INC. (US)
美國

(72) 發明人：艾佛瑞 馬克 EVERY, MARK (GB)；艾凡達諾 卡洛斯 AVENDANO, CARLOS (US)；索貝齊 路德傑 SOLBACH, LUDGER (DE)；姜擘 JIANG, YE (CN)；默吉亞 卡洛 MURGIA, CARLO (IT)

(74) 代理人：陳長文

(56) 參考文獻：

TW	200933609A	TW	201009817A
US	2008/0019548A1	US	2009/0067642A1

審查人員：李泉河

申請專利範圍項數：15 項 圖式數：5 共 33 頁

(54) 名稱

多麥克風之穩固雜訊抑制

MULTI-MICROPHONE ROBUST NOISE SUPPRESSION

(57) 摘要

一種穩固雜訊減少系統可同時地減少一聲學信號中之雜訊分量及回音分量，同時限制話語失真之位準。該系統可自呈一近講型、手持型或其他組態之兩個或兩個以上麥克風接收聲學信號。將該等經接收聲學信號變換成蝸域副頻帶信號，且可自該等副頻帶信號減去回音分量及雜訊分量。識別該等聲學副頻帶信號中之特徵且使用該等特徵以產生一乘法遮罩。將該乘法遮罩應用於該等減去雜訊之副頻帶信號，且在時域中重新建構該等副頻帶信號。

A robust noise reduction system may concurrently reduce noise and echo components in an acoustic signal while limiting the level of speech distortion. The system may receive acoustic signals from two or more microphones in a close-talk, hand-held or other configuration. The received acoustic signals are transformed to cochlea domain sub-band signals and echo and noise components may be subtracted from the sub-band signals. Features in the acoustic sub-band signals are identified and used to generate a multiplicative mask. The multiplicative mask is applied to the noise subtracted sub-band signals and the sub-band signals are reconstructed in the time domain.

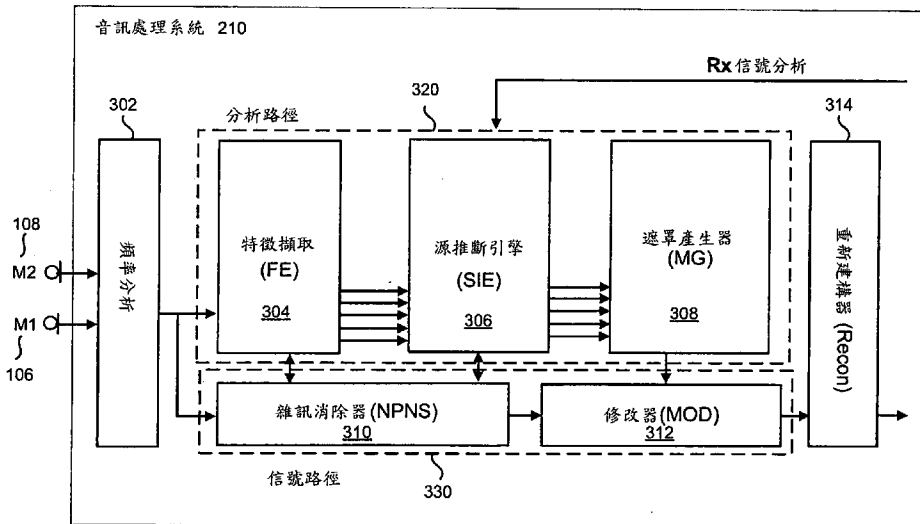


圖3

- 106 . . . 主麥克風
- 108 . . . 副麥克風
- 210 . . . 音訊處理系統
- 302 . . . 頻率分析模組
- 304 . . . 特徵擷取模組
- 306 . . . 源推斷引擎模組
- 308 . . . 遮罩產生器模組
- 310 . . . 雜訊消除器模組/雜訊消除模組/ NPNS 模組
- 312 . . . 修改器模組
- 314 . . . 重新建構器模組
- 320 . . . 分析路徑子系統
- 330 . . . 信號路徑子系統

發明專利說明書

公告本

(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：100115214

※申請日：100.4.29

※IPC 分類：G10L 21/02 (2013.01)

一、發明名稱：(中文/英文)

多麥克風之穩固雜訊抑制

MULTI-MICROPHONE ROBUST NOISE SUPPRESSION

二、中文發明摘要：

一種穩固雜訊減少系統可同時地減少一聲學信號中之雜訊分量及回音分量，同時限制話語失真之位準。該系統可自呈一近講型、手持型或其他組態之兩個或兩個以上麥克風接收聲學信號。將該等經接收聲學信號變換成蝸域副頻帶信號，且可自該等副頻帶信號減去回音分量及雜訊分量。識別該等聲學副頻帶信號中之特徵且使用該等特徵以產生一乘法遮罩。將該乘法遮罩應用於該等減去雜訊之副頻帶信號，且在時域中重新建構該等副頻帶信號。

三、英文發明摘要：

A robust noise reduction system may concurrently reduce noise and echo components in an acoustic signal while limiting the level of speech distortion. The system may receive acoustic signals from two or more microphones in a close-talk, hand-held or other configuration. The received acoustic signals are transformed to cochlea domain sub-band signals and echo and noise components may be subtracted from the sub-band signals. Features in the acoustic sub-band signals are identified and used to generate a multiplicative mask. The multiplicative mask is applied to the noise subtracted sub-band signals and the sub-band signals are reconstructed in the time domain.

四、指定代表圖：

(一)本案指定代表圖為：第(3)圖。

(二)本代表圖之元件符號簡單說明：

106	主麥克風
108	副麥克風
210	音訊處理系統
302	頻率分析模組
304	特徵擷取模組
306	源推斷引擎模組
308	遮罩產生器模組
310	雜訊消除器模組 / 雜訊消除模組 / NPNS 模組
312	修改器模組
314	重新建構器模組
320	分析路徑子系統
330	信號路徑子系統

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】

本發明大體上係關於音訊處理，且更特定言之，係關於音訊信號之雜訊抑制處理。

本申請案主張2010年4月29日申請之名為「Multi-Microphone Noise Suppression」之美國臨時申請案序號61/329,322的優先權利。本申請案係與2010年7月8日申請之名為「Method for Jointly Optimizing Noise Reduction and Voice Quality in a Mono or Multi-Microphone System」之美國專利申請案第xx/xxx,xxx號(代理人案號PA5218US)相關。前述申請案之揭示內容以引用的方式併入本文中。

【先前技術】

當前，存在用於減少不利音訊環境中之背景雜訊之許多方法。靜止雜訊抑制系統將靜止雜訊抑制達固定或變化數目個dB。固定抑制系統將靜止或非靜止雜訊抑制達固定數目個dB。靜止雜訊抑制器之缺點在於將不抑制非靜止雜訊，而固定抑制系統之缺點在於其必須將雜訊抑制達一保守位準，以便避免在低SNR下之話語失真。

另一形式之雜訊抑制為動態雜訊抑制。常見類型之動態雜訊抑制系統係基於信雜比(SNR)。SNR可用以判定抑制值。遺憾地，歸因於在音訊環境中存在不同雜訊類型，因而SNR單獨地不為話語失真之極好預測因子。通常，遍及給定時段之話語能量將包括字、暫停、字、暫停，等等。

另外，在音訊環境中可存在靜止雜訊及動態雜訊。SNR將

所有此等靜止及非靜止話語以及雜訊分量進行平均化。不考慮雜訊信號之特性之 SNR 之判定(僅考慮雜訊之總位準)。

為了克服先前技術之缺點，需要一種用於處理音訊信號之改良型雜訊抑制系統。

【發明內容】

本技術提供一種穩固雜訊抑制系統，該穩固雜訊抑制系統可同時地減少一聲學信號中之雜訊分量及回音分量，同時限制話語失真之位準。該系統可自呈一近講型、手持型或其他組態之兩個或兩個以上麥克風接收聲學信號。將該等經接收聲學信號變換成蝸域副頻帶信號，且可自該等副頻帶信號減去回音分量及雜訊分量。識別該等聲學副頻帶信號中之特徵且使用該等特徵以產生一乘法遮罩。將該乘法遮罩應用於該等減去雜訊之副頻帶信號，且在時域中重新建構該等副頻帶信號。

一實施例包括一種用於執行一音訊信號中之雜訊減少之系統，該系統可包括一記憶體。儲存於該記憶體中且藉由一處理器執行之一頻率分析模組可自時域聲學信號而在一蝸域中產生副頻帶信號。儲存於該記憶體中且藉由一處理器執行之一雜訊消除模組可消除該等副頻帶信號之至少一部分。儲存於該記憶體中且藉由一處理器執行之一修改器模組可抑制該等經修改副頻帶信號中之一雜訊分量或一回音分量。儲存於該記憶體中且藉由一處理器執行之一重新建構器模組可自藉由該修改器模組提供之該等抑制分量之

副頻帶信號重新建構一經修改時域信號。

亦可將雜訊減少執行為藉由具有一處理器及記憶體之一機器執行之一程序。另外，可實施一種電腦可讀儲存媒體，在該電腦可讀儲存媒體中體現有一程式，該程式可藉由一處理器執行以執行用於減少一音訊信號中之雜訊之一方法。

【實施方式】

本技術提供一種穩固雜訊抑制系統，該穩固雜訊抑制系統可同時地減少一聲學信號中之雜訊分量及回音分量，同時限制話語失真之位準。該系統可自呈一近講型、手持型或其他組態之兩個或兩個以上麥克風接收聲學信號。將該等經接收聲學信號變換成蝸域副頻帶信號，且可自該等副頻帶信號減去回音分量及雜訊分量。識別該等聲學副頻帶信號中之特徵且使用該等特徵以產生一乘法遮罩。將該乘法遮罩應用於該等減去雜訊之副頻帶信號，且在時域中重新建構該等副頻帶信號。本技術為一動態雜訊抑制系統及非靜止雜訊抑制系統，且基於雜訊之特性及使用狀況而提供「感知上最佳」量之雜訊抑制。

經由雜訊消除與雜訊抑制之組合而執行雜訊(及回音)減少會允許音訊裝置設計之靈活性。詳言之，減法階段與乘法階段之組合係有利的，此係因為其允許在音訊裝置上麥克風置放之靈活性及使用狀況(例如，近講型/遠講型)之靈活性兩者，同時最佳化語音品質相對於雜訊抑制之總取捨。麥克風針對「近麥克風」組態可經定位成在彼此之四

公分內，或針對「擴展麥克風」組態或具有兩個以上麥克風之組態的組合可經定位成相隔四公分以上。

圖1為可供使用本技術之實施例之環境的說明。使用者可充當至音訊裝置104之音訊(話語)源102。例示性音訊裝置104包括兩個麥克風：與音訊源102有關之主麥克風106，及經定位成與主麥克風106相隔一距離之副麥克風108。或者，音訊裝置104可包括單一麥克風。在又其他實施例中，音訊裝置104可包括兩個以上麥克風，諸如，三個、四個、五個、六個、七個、八個、九個、十個或甚至更多麥克風。

主麥克風106及副麥克風108可為全向麥克風。或者，實施例可利用其他形式之麥克風或聲學感測器，諸如，定向麥克風。

在麥克風106及108自音訊源102接收聲音(亦即，聲學信號)的同時，麥克風106及108亦拾取雜訊112。雖然在圖1中將雜訊110展示為來自單一位置，但雜訊110可包括來自不同於音訊源102之位置之一或多個位置的任何聲音，且可包括混響及回音。雜訊110可為靜止雜訊、非靜止雜訊，及/或靜止雜訊與非靜止雜訊之組合。

一些實施例可利用藉由兩個麥克風106及108接收之聲學信號之間的位準差(例如，能量差)。因為在近講型使用狀況下主麥克風106比副麥克風108更接近音訊源102，所以主麥克風106之強度位準較高，從而導致在(例如)話語/語音片段期間藉由主麥克風106接收之較大能量位準。

可接著使用位準差以鑑別時頻域中之話語及雜訊。另外實施例可使用能量位準差與時間延遲之組合以鑑別話語。基於雙耳提示編碼，可執行話語信號擷取或話語增強。

圖2為例示性音訊裝置104的方塊圖。在所說明實施例中，音訊裝置104包括接收器200、處理器202、主麥克風106、可選副麥克風108、音訊處理系統210，及輸出裝置206。音訊裝置104可包括音訊裝置104操作所必要之另外或其他組件。類似地，音訊裝置104可包括執行類似於或等效於圖2所描繪之功能之功能的較少組件。

處理器202可執行儲存於音訊裝置104中之記憶體(圖2中未說明)中的指令及模組，以執行本文中所描述之功能性，包括聲學信號之雜訊減少。處理器202可包括實施為處理單元之硬體及軟體，處理單元可處理針對處理器202之浮點操作及其他操作。

例示性接收器200為經組態以自通信網路接收信號之聲學感測器。在一些實施例中，接收器200可包括天線裝置。可接著將信號轉遞至音訊處理系統210，以使用本文中所描述之技術來減少雜訊，且將音訊信號提供至輸出裝置206。可在音訊裝置104之傳輸路徑及接收路徑中之一者或兩者中使用本技術。

音訊處理系統210經組態以經由主麥克風106及副麥克風108而自聲學源接收聲學信號，且處理聲學信號。處理可包括執行聲學信號內之雜訊減少。下文更詳細地論述音訊處理系統210。主麥克風106與副麥克風108可間隔開一距

離，以便允許偵測主麥克風106與副麥克風108之間的能量位準差、時間差或相位差。可將藉由主麥克風106及副麥克風108接收之聲學信號轉換成電信號(亦即，主電信號及副電信號)。根據一些實施例，可藉由類比至數位轉換器(未圖示)將電信號自身轉換成數位信號以供處理。為了出於清晰目的而區別聲學信號，在本文中將藉由主麥克風106接收之聲學信號稱作主聲學信號，而在本文中將藉由副麥克風108接收之聲學信號稱作副聲學信號。可藉由音訊處理系統210處理主聲學信號及副聲學信號以產生具有改良型信雜比之信號。應注意，可僅利用主麥克風106來實踐本文中所描述之技術之實施例。

輸出裝置206為將音訊輸出提供至使用者之任何裝置。舉例而言，輸出裝置206可包括揚聲器、頭戴式耳機或手機之聽筒，或會議裝置上之揚聲器。

在各種實施例中，在主麥克風與副麥克風為緊密間隔(例如，相隔1 cm至2 cm)之全向麥克風時，可使用波束成形技術以模擬面向前及面向後之定向麥克風。可使用位準差以鑑別時頻域中之話語及雜訊，可在雜訊減少中使用位準差。

圖3為用於執行如本文中所描述之雜訊減少之例示性音訊處理系統210的方塊圖。在例示性實施例中，音訊處理系統210體現於音訊裝置104內之記憶體裝置內。音訊處理系統210可包括頻率分析模組302、特徵擷取模組304、源推斷引擎模組306、遮罩產生器模組308、雜訊消除器模組

310、修改器模組312，及重新建構器模組314。音訊處理系統210可包括比圖3所說明之組件更多或更少的組件，且可將模組之功能性組合或擴充至較少或額外模組中。在圖3之各種模組之間及在本文中之其他圖中說明例示性通信線路。通信線路既不意欲限制哪些模組與其他模組以通信方式耦接，通信線路亦不意欲限制在模組之間所傳達之信號之數目及類型。

在操作中，將自主麥克風106及副麥克風108所接收之聲學信號轉換成電信號，且經由頻率分析模組302處理電信號。在藉由頻率分析模組302處理聲學信號之前，可在時域中預處理聲學信號。時域預處理可包括應用輸入限制器增益、話語時間延伸，及使用FIR或IIR濾波器進行濾波。

頻率分析模組302獲取聲學信號，且模仿藉由濾波器組模擬的蝸(例如，蝸域)之頻率分析。頻率分析模組302將主聲學信號及副聲學信號中之每一者分離成兩個或兩個以上頻率副頻帶信號。副頻帶信號為對輸入信號之濾波操作之結果，其中濾波器之頻寬窄於藉由頻率分析模組302接收之信號之頻寬。可藉由一系列級聯式複值一階IIR濾波器來實施濾波器組。或者，可將諸如短時傅立葉變換(STFT)、副頻帶濾波器組、調變式複數重疊變換、蝸模組、小波等等之其他濾波器用於頻率分析及合成。可將頻率副頻帶信號之樣本依序地分組成若干時間訊框(例如，遍及預定時段)。舉例而言，一訊框之長度可為4毫秒、8毫秒或某其他時間長度。在一些實施例中，可能根本不存

在訊框。結果可包括在快速蝸變換(FCT)域中之副頻帶信號。

將副頻帶訊框信號自頻率分析模組302提供至分析路徑子系統320及信號路徑子系統330。分析路徑子系統320可處理信號以識別信號特徵、區分副頻帶信號之話語分量與雜訊分量，且產生信號修改器。信號路徑子系統330負責藉由減少主聲學信號之副頻帶信號中之雜訊而修改副頻帶信號。雜訊減少可包括應用修改器(諸如，在分析路徑子系統320中所產生之乘法增益遮罩)，或藉由自副頻帶信號減去分量。雜訊減少可減少雜訊且保留副頻帶信號中之所要話語分量。

信號路徑子系統330包括雜訊消除器模組310及修改器模組312。雜訊消除器模組310自頻率分析模組302接收副頻帶訊框信號。雜訊消除器模組310可自主聲學信號之一或多個副頻帶信號減去(例如，消除)雜訊分量。因而，雜訊消除器模組310可輸出主信號中之雜訊分量之副頻帶估計及呈減去雜訊之副頻帶信號之形式的話語分量之副頻帶估計。

雜訊消除器模組310可基於源位置而藉由減法演算法來提供(例如)在具有雙麥克風組態之系統中之雜訊消除。雜訊消除器模組310亦可提供回音消除，且對揚聲器及Rx路徑非線性固有地穩固。藉由在具有很少語音品質降級或無語音品質降級的情況下執行雜訊及回音消除(例如，自主信號副頻帶減去分量)，雜訊消除器模組310可增加自頻率

分析模組302所接收且提供至修改器模組312及後濾波模組之副頻帶信號中之話語對雜訊比(SNR)。所執行之雜訊消除之量可取決於雜訊源之擴散性及麥克風之間的距離，雜訊源之擴散性及麥克風之間的距離兩者皆有助於麥克風之間的雜訊之相干性，其中較大相干性導致較好消除。

可以多種方式實施雜訊消除器模組310。在一些實施例中，可用單一NPNS模組實施雜訊消除器模組310。或者，雜訊消除器模組310可包括兩個或兩個以上NPNS模組，該等NPNS模組可(例如)以級聯方式予以配置。

在一些實施例中藉由雜訊消除器模組310執行之雜訊消除之一實例被揭示於2008年6月30日申請之名為「System and Method for Providing Noise Suppression Utilizing Null Processing Noise Subtraction」之美國專利申請案第12/215,980號、2009年4月13日申請之名為「Adaptive Noise Cancellation」之美國申請案第12/422,917號及2010年1月26日申請之名為「Adaptive Noise Reduction Using Level Cues」之美國申請案第12/693,998號中，該等申請案之揭示內容各自以引用的方式併入本文中。

分析路徑子系統320之特徵擷取模組304接收自藉由頻率分析模組302提供之主聲學信號及副聲學信號所導出的副頻帶訊框信號，以及NPNS模組310之輸出。特徵擷取模組304計算如下各項：副頻帶信號之訊框能量估計；主聲學信號與副聲學信號之間的麥克風間位準差(ILD)、麥克風間時間差(ITD)及麥克風間相位差(IPD)；主麥克風及副麥

克風之自雜訊估計；以及可藉由其他模組利用之其他單耳或雙耳特徵，諸如，麥克風信號之間的時間估計及交叉相關。特徵擷取模組304可提供至NPNS模組310之輸入且處理來自NPNS模組310之輸出。

特徵擷取模組304可產生空處理麥克風間位準差 (null-processing inter-microphone level difference；NP-ILD)。可在本系統中與原始ILD互換地使用NP-ILD。可藉由特徵擷取模組304內之ILD模組來判定主麥克風與副麥克風之間的原始ILD。可藉由如下方程式以算術方式表示在實施例中藉由ILD模組計算之ILD：

$$ILD = \left[\left[c \cdot \log_2 \left(\frac{E_1}{E_2} \right) \right]_{-1} \right]_{+1}$$

其中E1及E2分別為主麥克風106及副麥克風108之能量輸出，該等能量輸出係遍及非重疊時間間隔(「訊框」)而在每一副頻帶信號中予以計算。此方程式描述被正規化達c倍且限於範圍[-1, +1]之dB ILD。因此，當音訊源102對於E1而言接近主麥克風106且不存在雜訊時，ILD=1，但隨著添加更多雜訊，ILD將減少。

在一些狀況下，在麥克風之間的距離相對於在主麥克風與嘴之間的距離較小時，原始ILD可能不會有用於將源與擾亂項(distracter)加以鑑別，此係因為源及擾亂項兩者皆可能具有約略相等的原始ILD。為了避免關於用以將源與擾亂項加以鑑別之原始ILD的限制，可使用雜訊消除模組

330之輸出以導出針對話語信號具有正值且針對雜訊分量具有小值或負值之ILD，此係因為此等者將在雜訊消除模組310之輸出處顯著地衰減。自雜訊消除模組330之輸出所導出的ILD可為空處理麥克風間位準差(NP-ILD)，且藉由如下方程式以算術方式表示該ILD：

$$NP-ILD = \left[\left[c \cdot \log_2 \left(\frac{E_{NP}}{E_2} \right) \right] \right]_{-1}^{+1}$$

其中 E_{NP} 為NPNS之輸出能量。NP-ILD之使用允許在音訊裝置內麥克風之置放的較大靈活性。舉例而言，NP-ILD可允許以前後組態置放麥克風，該組態具有在2 cm至15 cm之間的分離距離，且在總抑制位準方面具有幾個dB之效能變化。

NPNS模組可將消除雜訊之副頻帶信號提供至特徵擷取模組304中之ILD區塊。因為可將ILD判定為NPNS輸出信號能量對副麥克風能量之比率，所以ILD常常可與空處理麥克風間位準差(NP-ILD)互換。可使用「原始ILD」以將自「原始」主麥克風信號及副麥克風信號計算ILD之狀況進行歧義消除。

判定能量位準估計及麥克風間位準差被更詳細地論述於名為「System and Method for Utilizing Inter-Microphone Level Differences for Speech Enhancement」之美國專利申請案第11/343,524號中，該申請案以引用的方式併入本文中。

源推斷引擎模組306可處理藉由特徵擷取模組304提供之訊框能量估計以計算雜訊估計且導出副頻帶信號中之雜訊及話語之模型。源推斷引擎模組306調適性地估計聲學源之屬性，諸如，NPNS模組310之輸出信號的聲學源之能譜。可利用能譜屬性以在遮罩產生器模組308中產生乘法遮罩。

源推斷引擎模組306可自特徵擷取模組304接收NP-ILD，且追蹤目標音訊源102、背景雜訊及(視情況)回音之NP-ILD機率分佈或「叢集」。

接著，連同其他聽覺提示使用此資訊，以在源與雜訊類別之間界定分類邊界。歸因於改變環境條件、音訊裝置104之移動、使用者之手及/或臉之位置、與音訊裝置104有關之其他物件，及其他因素，話語、雜訊及回音之NP-ILD分佈可隨著時間推移而變化。叢集追蹤器調適於話語或雜訊源之時變NP-ILD。

當忽略回音時，在無任何一般性損失之情況下，當源及雜訊ILD分佈非重疊時，有可能在該兩個分佈之間指定分類邊界或顯性臨限值，使得在SNR為足夠正時將信號分類為話語，或在SNR為足夠負時將信號分類為雜訊。可按照副頻帶及時間訊框將此分類判定為顯性遮罩(dominance mask)，且藉由叢集追蹤器模組將此分類輸出至源推斷引擎模組306內之雜訊估計器模組。

叢集追蹤器可至少部分地基於自聲學信號所導出之聲學特徵而判定聲學特徵之全域概述，以及基於聲學特徵之全

域執行估計及全域概述而判定瞬時全域分類。可更新該等全域執行估計，且基於至少該一或多個聲學特徵而導出瞬時局域分類。可接著至少部分地基於該瞬時局域分類及該一或多個聲學特徵而判定譜能量分類。

在一些實施例中，叢集追蹤器模組基於此等局域叢集及觀測而將能譜中之點分類為話語或雜訊。因而，將能譜中之每一點之局域二進位遮罩識別為話語或雜訊。

叢集追蹤器模組可按照副頻帶產生雜訊/話語分類信號且將分類提供至NPNS模組310。在一些實施例中，該分類為指示在雜訊與話語之間的區別的控制信號。雜訊消除器模組310可利用分類信號以估計經接收麥克風信號中之雜訊。在一些實施例中，可將叢集追蹤器模組之結果轉遞至源推斷引擎模組306內之雜訊估計模組。換言之，提供當前雜訊估計連同能譜中可經定位有雜訊之位置以用於處理音訊處理系統210內之雜訊信號。

藉由叢集追蹤器模組來追蹤叢集之一實例被揭示於2007年12月21日申請之名為「System and method for Adaptive Classification of Audio Sources」之美國專利申請案第12/004,897號中，該申請案之揭示內容以引用的方式併入本文中。

源推斷引擎模組306可包括一雜訊估計模組，該雜訊估計模組可自叢集追蹤器模組及雜訊消除器模組310之輸出接收雜訊/話語分類控制信號以估計雜訊 $N(t,w)$ ，其中 t 為時間點，且 W 表示頻率或副頻帶。將藉由雜訊估計模組判

定之雜訊估計提供至遮罩產生器模組308。在一些實施例中，遮罩產生器模組308接收雜訊消除器模組310之雜訊估計輸出及叢集追蹤器模組之輸出。

源推斷引擎模組306中之雜訊估計模組可包括NP-ILD雜訊估計器及靜止雜訊估計器。可將雜訊估計(諸如)與 $\max()$ 運算進行組合，使得由組合式雜訊估計導致之雜訊抑制效能至少為個別雜訊估計之雜訊抑制效能。

可自顯性遮罩及雜訊消除器模組310之輸出信號能量導出NP-ILD雜訊估計。當在特定副頻帶中顯性遮罩為1(指示話語)時，使雜訊估計凍結，且當在特定副頻帶中顯性遮罩為0(指示雜訊)時，將雜訊估計設定成等於NPNS輸出信號能量。靜止雜訊估計追蹤變化得比話語通常變化得更緩慢的NPNS輸出信號之分量，且至此模組之主要輸入為NPNS輸出能量。

遮罩產生器模組308接收如藉由源推斷引擎模組306估計的副頻帶話語分量及雜訊分量之模型，且產生乘法遮罩。將乘法遮罩應用於藉由NPNS 310提供至修改器312的所估計之減去雜訊之副頻帶信號。修改器模組312使增益遮罩與藉由NPNS模組310輸出的主聲學信號之減去雜訊之副頻帶信號相乘。應用該遮罩會減少主聲學信號之副頻帶信號中雜訊分量之能量位準，且會導致雜訊減少。

藉由溫納濾波器(Wiener filter)及語音品質最佳化抑制系統來界定乘法遮罩。溫納濾波器估計可基於雜訊之功率譜密度及主聲學信號之功率譜密度。溫納濾波器基於雜訊估

計而導出增益。考慮到有雜訊信號，使用所導出之增益以產生清潔話語信號之理論MMSE的估計。為了限制由於遮罩應用而導致之話語失真之量，可使用感知上導出之增益下限而在下端處限制溫納增益。

自遮罩產生器模組308所輸出之增益遮罩之值係時間及副頻帶信號相依的，且以每副頻帶為基礎而最佳化雜訊減少。雜訊減少可經受話語損失失真遵守容許臨限極限之約束。臨限極限可基於許多因素，諸如，語音品質最佳化抑制(VQOS)位準。VQOS位準為副頻帶信號中藉由雜訊減少引入之話語損失失真的估計最大臨限位準。VQOS係可調諧的且考量副頻帶信號之性質，且向系統及聲學設計者提供充分設計靈活性。將在副頻帶信號中所執行之雜訊減少之量的下限判定為經受VQOS臨限值，藉此限制副頻帶信號之話語損失失真之量。結果，當可能時可在副頻帶信號中執行大量雜訊減少，且當諸如不可接受高之話語損失失真的條件不允許大量雜訊減少時雜訊減少可較小。

在實施例中，可將副頻帶信號中之雜訊分量之能量位準減少至不小於殘餘雜訊目標位準，殘餘雜訊目標位準可為固定的或緩慢時變的。在一些實施例中，殘餘雜訊目標位準針對每一副頻帶信號係相同的；在其他實施例中，殘餘雜訊目標位準可橫越諸副頻帶而變化。此目標位準可為雜訊分量不再成為可聽到或可感知時之位準、低於用以俘獲主聲學信號之麥克風之自雜訊位準之位準，或低於在實施雜訊減少技術之系統內基頻晶片上之分量或內部雜訊門限

(noise gate)之分量之雜訊門限的位準。

修改器模組312自雜訊消除器模組310接收信號路徑蝸樣本，且將自遮罩產生器308所接收之增益遮罩應用於經接收樣本。信號路徑蝸樣本可包括主聲學信號之減去雜訊之副頻帶信號。藉由溫納濾波器估計提供之遮罩可快速地變化(諸如，自訊框至訊框)，且雜訊及話語估計可在諸訊框之間變化。為了幫助處理該變化，可藉由修改器312將遮罩之向上及向下時間跳越速率(slew rate)約束成在合理極限內。可使用簡單線性內插將遮罩自訊框速率內插至樣本速率，且藉由乘法雜訊抑制將遮罩應用於副頻帶信號。修改器模組312可輸出經遮罩之頻率副頻帶信號。

重新建構器模組314可將經遮罩之頻率副頻帶信號自蝸域轉換回成時域。該轉換可包括添加經遮罩之頻率副頻帶信號及相移信號。或者，該轉換可包括使經遮罩之頻率副頻帶信號與蝸頻道之反頻率相乘。一旦完成至時域之轉換，隨即可經由輸出裝置206將經合成聲學信號輸出至使用者及/或將經合成聲學信號提供至編解碼器以供編碼。

在一些實施例中，可執行經合成時域聲學信號之額外後處理。舉例而言，可在將經合成聲學信號提供至使用者之前將藉由舒適雜訊產生器產生之舒適雜訊添加至該信號。舒適雜訊可為通常不能為收聽者所辨別的均一恆定雜訊(例如，粉紅雜訊(pink noise))。可將此舒適雜訊添加至經合成聲學信號以加強可聽度之臨限值且以遮罩低位準非靜止輸出雜訊分量。在一些實施例中，可將舒適雜訊位準選

擇成恰好高於可聽度之臨限值，且可由使用者設定。在一些實施例中，遮罩產生器模組308可以使用舒適雜訊之位準，以便產生將會將雜訊抑制成處於或低於舒適雜訊之位準的增益遮罩。

圖3之系統可處理藉由音訊裝置接收的若干類型之信號。可經由一或多個麥克風將該系統應用於聲學信號。該系統亦可處理經由天線或其他連接所接收的信號，諸如，數位Rx信號。

圖4及圖5包括用於執行本技術之例示性方法的流程圖。可以任何次序執行圖4及圖5之每一步驟，且圖4及圖5之方法可各自包括相較於所說明之步驟為額外的步驟或比所說明之步驟更少的步驟。

圖4為用於執行聲學信號之雜訊減少之例示性方法的流程圖。在步驟405處，可接收麥克風聲學信號。藉由麥克風106及108接收之聲學信號可各自包括話語及雜訊之至少一部分。在步驟410處，可對聲學信號執行預處理。預處理可包括將增益、等化及其他信號處理應用於聲學信號。

在步驟415處，在蝸域中產生副頻帶信號。可使用複數濾波器之級聯而自時域信號產生副頻帶信號。

在步驟420處，執行特徵擷取。特徵擷取可自用以消除雜訊分量、推斷副頻帶是否具有雜訊或回音且產生遮罩之副頻帶信號擷取特徵。執行特徵擷取係關於圖5予以更詳細地論述。

在步驟425處，執行雜訊消除。可藉由NPNS模組330對

自頻率分析模組302所接收之一或多個副頻帶信號執行雜訊消除。雜訊消除可包括自主聲學信號副頻帶減去雜訊分量。在一些實施例中，可自主聲學信號副頻帶消除回音分量。可將消除雜訊(或消除回音)之信號提供至特徵擷取模組304以判定雜訊分量能量估計且將該信號提供至源推斷引擎306。

在步驟430處，可判定副頻帶之雜訊估計、回音估計及話語估計。可判定聲學信號中之每一副頻帶的每一估計及聲學音訊信號中之每一訊框的每一估計。可至少部分地自藉由源推斷引擎306接收之Rx信號判定回音。將關於特定時間訊框內之副頻帶是被判定為雜訊、話語或是回音之推斷提供至遮罩產生器模組308。

在步驟435處，產生遮罩。可藉由遮罩產生器308產生遮罩。可產生遮罩，且在每一訊框期間基於關於特定副頻帶是被判定為雜訊、話語或是回音之判定而將遮罩應用於每一副頻帶。可基於語音品質最佳化抑制(經判定為針對特定語音失真位準而最佳化的抑制位準)而產生遮罩。在步驟440處，可接著將遮罩應用於副頻帶。可藉由修改器312將遮罩應用於藉由NPNS 310輸出之副頻帶信號。可藉由修改器312將遮罩自訊框速率內插至樣本速率。

在步驟445處，自副頻帶信號重新建構時域信號。可藉由重新建構器模組314將一系列延遲及複數乘運算應用於副頻帶信號而重新建構時間頻帶信號。在步驟450處，可接著對經重新建構時域信號執行後處理。可藉由後處理器

執行後處理，且後處理可包括將輸出限制器應用於經重新建構信號、應用自動增益控制，及其他後處理。在步驟455處，可接著輸出經重新建構輸出信號。

圖5為用於自音訊信號擷取特徵之例示性方法的流程圖。圖5之方法可提供針對圖4之方法之步驟420的更多細節。在步驟505處，接收副頻帶信號。特徵擷取模組304可自頻率分析模組302接收副頻帶信號且自雜訊消除器模組310接收輸出信號。在步驟510處，判定二階統計，諸如，副頻帶能量位準。可判定每一訊框之每一副頻帶的能量副頻帶位準。在步驟515處，可計算麥克風之間的交叉相關及麥克風信號之自相關。在步驟520處，判定麥克風間位準差(ILD)。在步驟525處，判定空處理麥克風間位準差(NP-ILD)。至少部分地自副頻帶信號能量及雜訊估計能量判定ILD及NP-ILD兩者。接著藉由音訊處理系統利用經擷取特徵以減少副頻帶信號中之雜訊。

上述模組(包括關於圖3所論述之模組)可包括儲存於諸如機器可讀媒體(例如，電腦可讀媒體)之儲存媒體中的指令。可藉由處理器202擷取及執行此等指令以執行本文中所述之功能性。指令之一些實例包括軟體、程式碼及韌體。儲存媒體之一些實例包括記憶體裝置及積體電路。

雖然參考上文所詳述之較佳實施例及實例來揭示本發明，但應理解，此等實例意欲呈說明性而非限制性意義。應預期，熟習此項技術者將容易地想到修改及組合，該等修改及組合將在本發明之精神內及在以下申請專利範圍之

範疇內。

【圖式簡單說明】

圖1為可供使用本技術之實施例之環境的說明。

圖2為例示性音訊裝置的方塊圖。

圖3為例示性音訊處理系統的方塊圖。

圖4為用於執行聲學信號之雜訊減少之例示性方法的流程圖。

圖5為用於自音訊信號擷取特徵之例示性方法的流程圖。

【主要元件符號說明】

102	目標音訊源
104	音訊裝置
106	主麥克風
108	副麥克風
112	雜訊
200	接收器
202	處理器
206	輸出裝置
210	音訊處理系統
302	頻率分析模組
304	特徵擷取模組
306	源推斷引擎模組
308	遮罩產生器模組
310	雜訊消除器模組/雜訊消除模組/NPNS模組

- 312 修改器模組
- 314 重新建構器模組
- 320 分析路徑子系統
- 330 信號路徑子系統

七、申請專利範圍：

1. 一種用於執行一音訊(audio)信號中之雜訊減少之系統，該系統包含：
 - 一記憶體；
 - 一頻率分析模組，該頻率分析模組儲存於該記憶體中且藉由一處理器執行以自時域聲學信號(time domain acoustic signals)而在一頻域(frequency domain)中產生若干副頻帶信號(sub-band signals)；
 - 一特徵擷取器(feature extractor)模組，其儲存在記憶體中並由處理器執行，以判定該等副頻帶信號之一或多個特徵，該一或多個特徵係針對該等聲學信號之一系列訊框中之每一訊框而判定；
 - 一雜訊消除(noise cancellation)模組，該雜訊消除模組儲存於該記憶體中且藉由處理器執行以消除該等副頻帶信號之至少一部分，並產生若干雜訊消除的副頻帶信號(noise-cancelled sub-band signals)；
 - 一遮罩產生器模組，其儲存於記憶體中且藉由該處理器執行以產生一遮罩，該遮罩之判定係至少部分地根據由該特徵擷取模組所判定的該一或多個特徵，及該遮罩經組態以藉由一修改器(modifier)模組應用於該等雜訊消除的副頻帶信號；
 - 該修改器模組，該修改器模組儲存於該記憶體中且藉由處理器執行以抑制該等雜訊消除的副頻帶信號中之一雜訊分量(noise component)及一回音(echo)分量中至少之

一者，以產生若干經修改的副頻帶信號；及

一重新建構器(reconstructor)模組，該重新建構器模組儲存於該記憶體中且藉由處理器執行以自該等經修改的副頻帶信號重新建構一經修改時域(time domain)信號。

2. 如請求項1之系統，其中該等時域聲學信號係自一音訊裝置上之一或多個麥克風信號予以接收。
3. 如請求項1之系統，該特徵擷取模組經組態以控制該雜訊消除模組及該修改器模組中至少之一者之調適(adaptation)。
4. 如請求項1之系統，該雜訊消除模組藉由自該等副頻帶信號減去(subtracting)一雜訊分量及一回音分量中至少之一者而消除該等副頻帶信號之至少一部分。
5. 如請求項4之系統，該一或多個特徵係一空處理麥克風間位準差(null-processing inter-microphone level difference)，其係在該特徵擷取模組中自該雜訊消除模組之輸出及自該等經接收副頻帶信號導出(derived)。
6. 如請求項1之系統，其中至少部分地基於話語損失失真之一臨限值準、雜訊或回音抑制之一所要位準或該主聲學信號之每一副頻帶中之一估計信雜比而判定該遮罩。
7. 如請求項3之系統，其中該一或多個特徵包含在一主聲學信號及一第二、第三或其他聲學信號之間的麥克風間位準差、麥克風間時間差及麥克風間相位差中至少之一者。
8. 一種用於執行一音訊信號中之雜訊減少之方法，該方法

包含：

藉由一處理器執行一儲存頻率分析模組以自時域聲學信號而在一頻域中產生若干副頻帶信號；

藉由處理器執行一特徵擷取器模組，以判定該等副頻帶信號之一或多個特徵，該一或多個特徵係針對該等聲學信號之一系列訊框中之每一訊框而判定；

藉由處理器執行一雜訊消除模組以消除該等副頻帶信號之至少一部分，並產生若干雜訊消除的副頻帶信號；

執行一遮罩產生器模組，以產生一遮罩，該遮罩之判定係至少部分地根據由該特徵擷取器模組所判定的該一或多個特徵，及該遮罩經組態以藉由一修改器模組應用於該等雜訊消除的副頻帶信號；

藉由處理器執行該修改器模組以抑制該等雜訊消除的副頻帶信號中之一雜訊分量及一回音分量中至少之一者，以產生若干經修改的副頻帶信號；及

藉由處理器執行一重新建構器模組以自該等經修改的副頻帶信號重新建構一經修改時域信號。

9. 如請求項8之方法，其進一步包含自一音訊裝置上之一或多個麥克風信號接收時域聲學信號。
10. 如請求項8之方法，其進一步包含控制該雜訊消除模組及該修改器模組中至少之一者之調適。
11. 如請求項8之方法，其進一步包含藉由減去一雜訊分量及一回音分量中至少之一者而消除該等副頻帶信號之至少一部分。

12. 如請求項11之方法，該一或多個特徵係在該特徵擷取模組中自該雜訊消除模組之輸出及自該等經接收副頻帶信號導出。
13. 如請求項8之方法，其中至少部分地基於話語損失失真之一臨限值準、雜訊或回音抑制之一所要位準或該主聲學信號之每一副頻帶中之一估計信雜比而判定該遮罩。
14. 如請求項10之方法，其中該一或多個特徵包含在一主聲學信號及一第二、第三或其他聲學信號之間的麥克風間位準差、麥克風間時間差及麥克風間相位差中至少之一者。
15. 一種非暫態電腦可讀儲存媒體，在該電腦可讀儲存媒體上體現有(embodied)一程式，該程式可藉由一處理器執行以執行用於減少一音訊信號中之雜訊之一方法，該方法包含：
 - 自時域聲學信號而在一頻域中產生若干副頻帶信號；
 - 判定該等副頻帶信號之一或多個特徵，該一或多個特徵係針對該等聲學信號之一系列訊框中之每一訊框而判定；
 - 消除該等副頻帶信號之至少一部分，以產生雜訊消除的副頻帶信號；
 - 產生一遮罩，該遮罩之判定係至少部分地根據由該特徵擷取模組所判定的該一或多個特徵，及該遮罩經組態以藉由一修改器模組應用於由該雜訊消除模組輸出的副頻帶信號；

抑制該等雜訊消除的副頻帶信號中之一雜訊分量及一回音分量中至少之一者，以產生若干經修改的副頻帶信號；及

自該等經修改的副頻帶信號重新建構一經修改時域信號。

八、圖式：

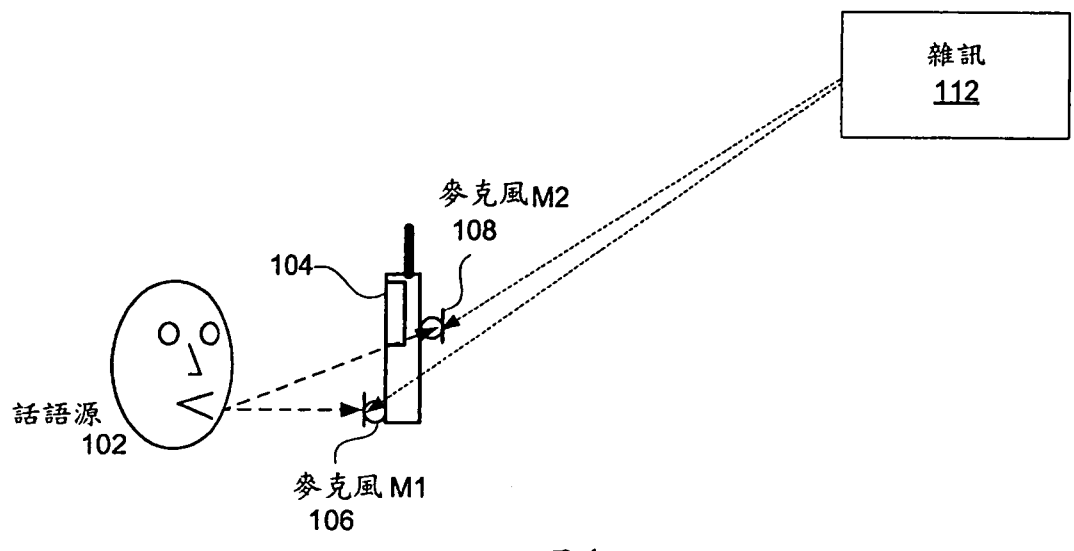


圖1

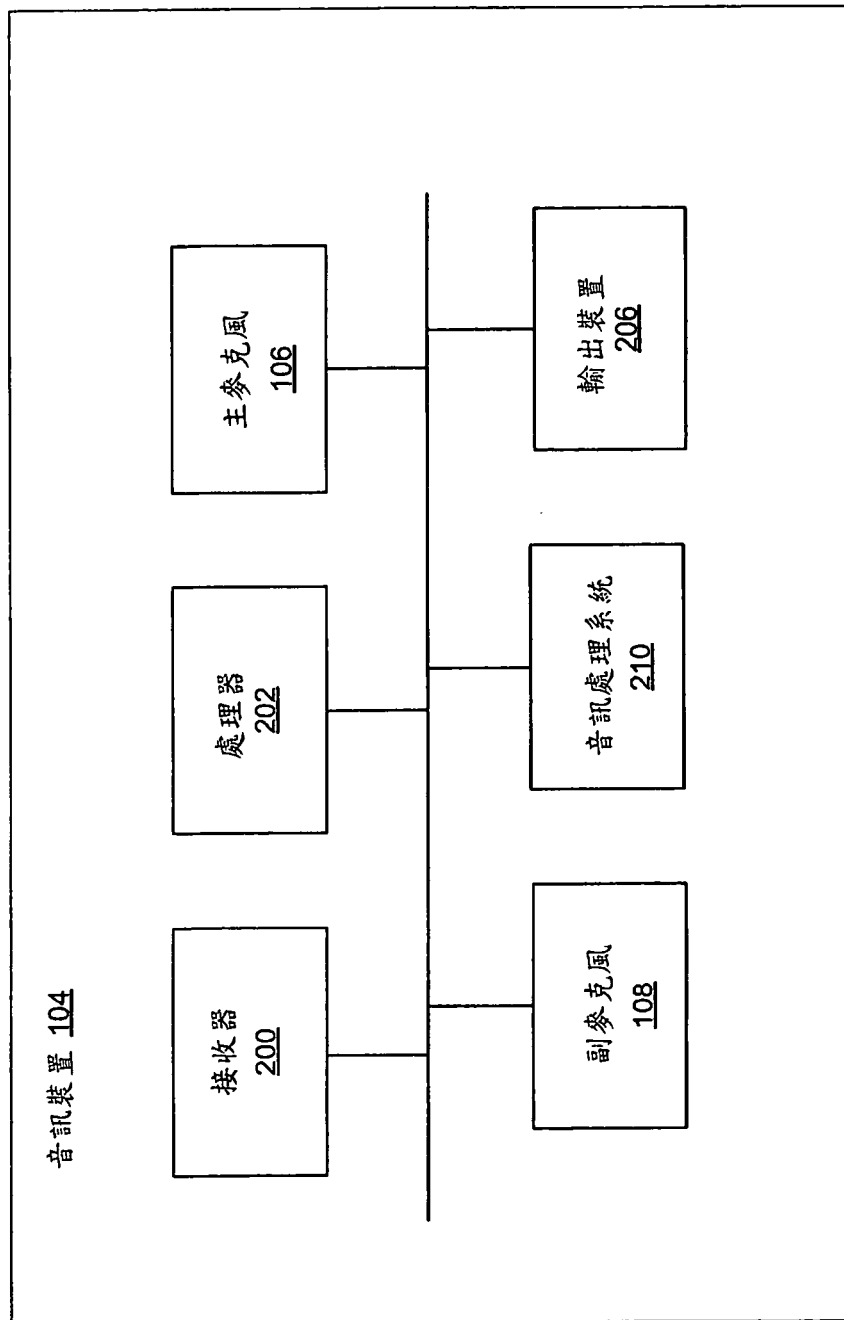


圖2

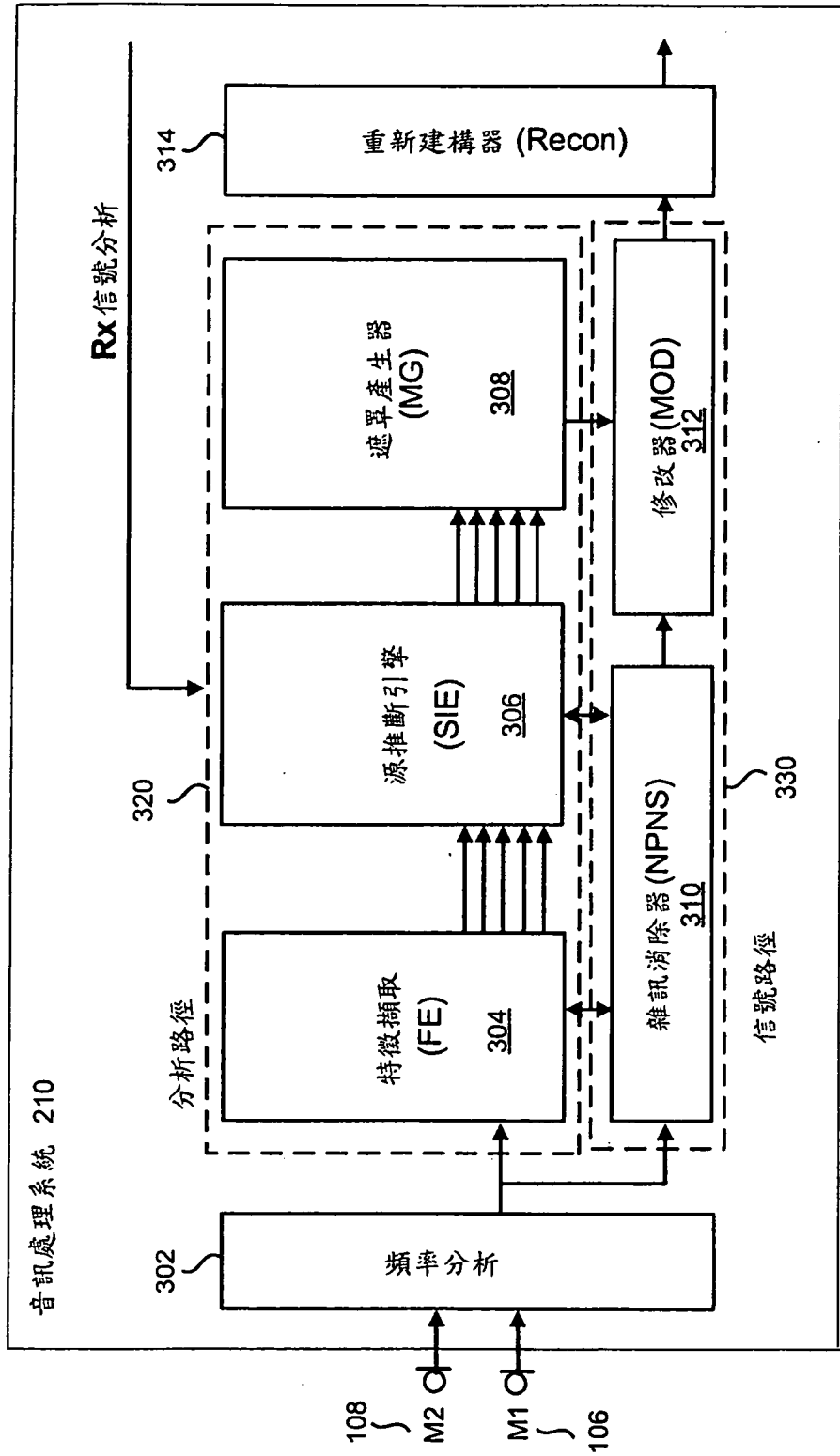


圖3

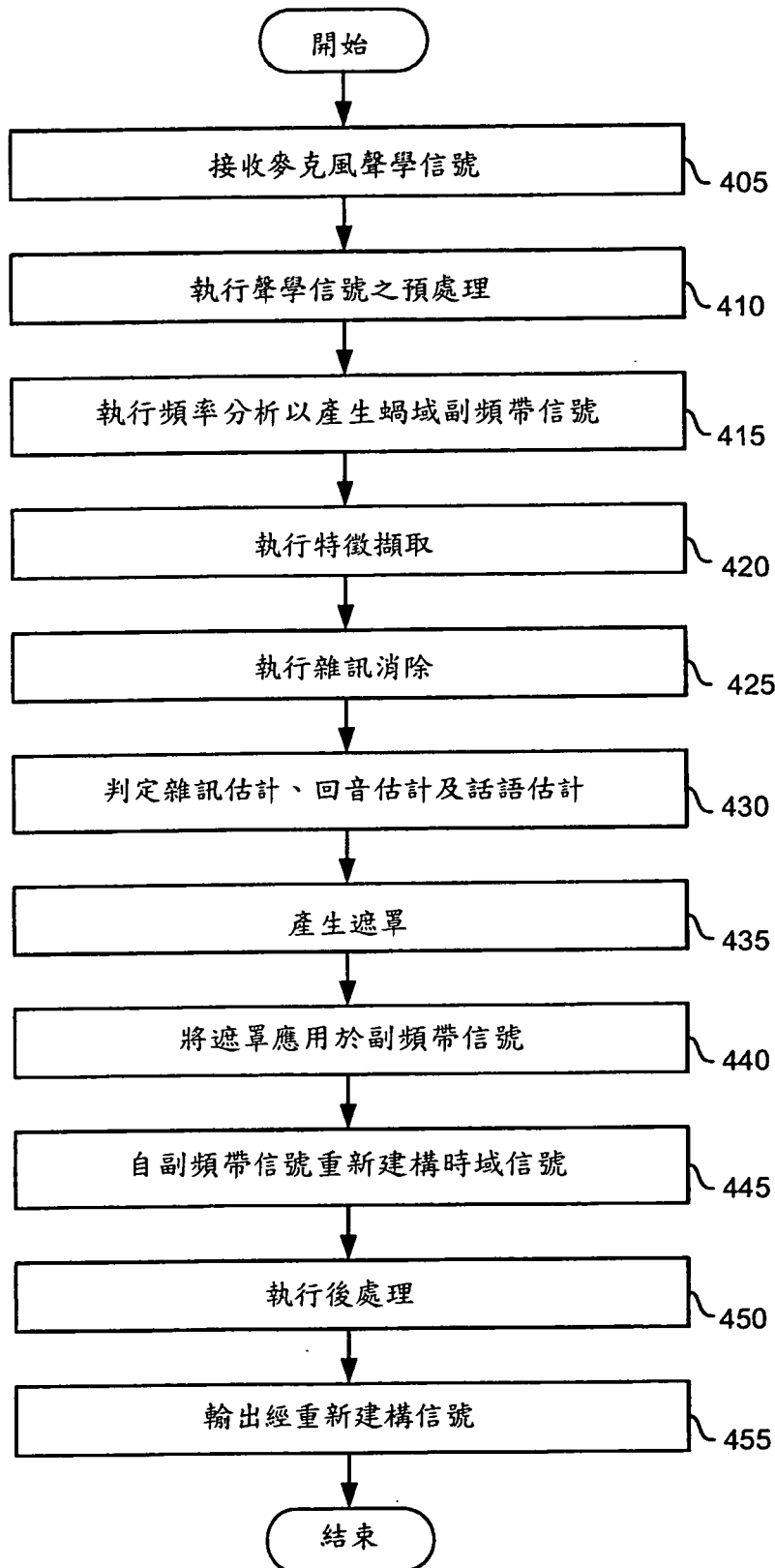


圖4

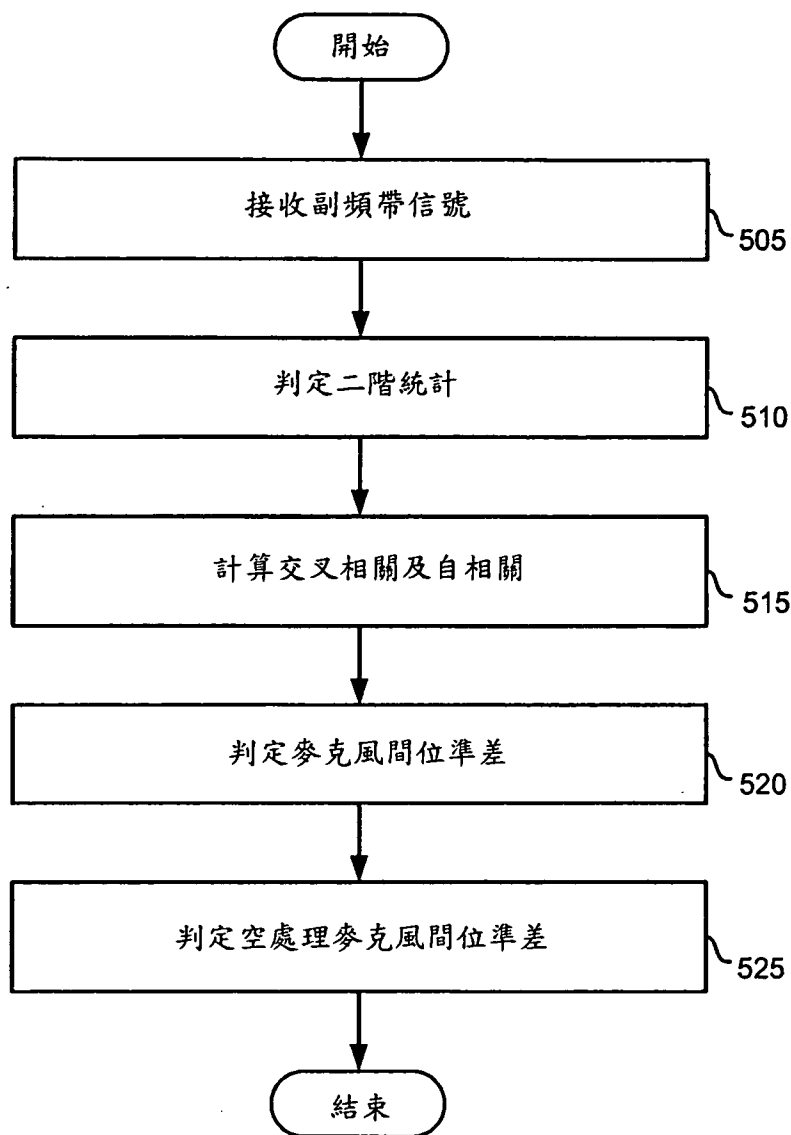


圖5