



(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 I594232 B

(45)公告日：中華民國 106 (2017) 年 08 月 01 日

(21)申請案號：101109568

(22)申請日：中華民國 101 (2012) 年 03 月 20 日

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

(30)優先權：2011/03/21 世界智慧財產權組織 PCT/SE2011/050306

(71)申請人：L M艾瑞克生 (P U B L) 電話公司 (瑞典) TELEFONAKTIEBOLAGET L M
ERICSSON (PUBL) (SE)

瑞典

(72)發明人：桑德蘭 尼拉斯 SANDGREN, NICLAS (SE)

(74)代理人：蔣大中

(56)參考文獻：

US 6459914B1

US 2003/0216909A1

US 2009/0276067A1

US 2010/0182510A1

WO 2010/027509A1

審查人員：黃衍勳

申請專利範圍項數：23 項 圖式數：5 共 26 頁

(54)名稱

處理音訊信號之方法及裝置

METHOD AND APPARATUS FOR PROCESSING OF AUDIO SIGNALS

(57)摘要

本發明揭示一音訊處置實體中用於使一音訊信號之一時間區段中之優勢頻率衰減之方法及配置。獲得一音訊信號之一時間區段，並且導出該時間區段之頻譜密度或「頻譜」之一估計。藉由平滑化該估計導出該估計之一近似，且藉由反轉該近似而導出一頻率遮罩。接著基於該頻率遮罩使該音訊時間區段中包括之頻率衰減。該方法及配置不涉及多頻帶濾波或起音及釋放時間之選擇。

Method and arrangement in an audio handling entity, for damping of dominant frequencies in a time segment of an audio signal. A time segment of an audio signal is obtained, and an estimate of the spectral density or "spectrum" of the time segment is derived. An approximation of the estimate is derived by smoothing the estimate, and a frequency mask is derived by inverting the approximation. Frequencies comprised in the audio time segment are then damped based on the frequency mask. The method and arrangement involves no multi-band filtering or selection of attack and release times.

指定代表圖：

符號簡單說明：

302 . . . 獲得信號區段

304 . . . 導出頻譜密度估計

306 . . . 藉由使頻譜密度估計平滑化而導出近似

308 . . . 藉由反轉近似導出一頻率遮罩

310 . . . 基於頻率遮罩使信號區段之頻率衰減

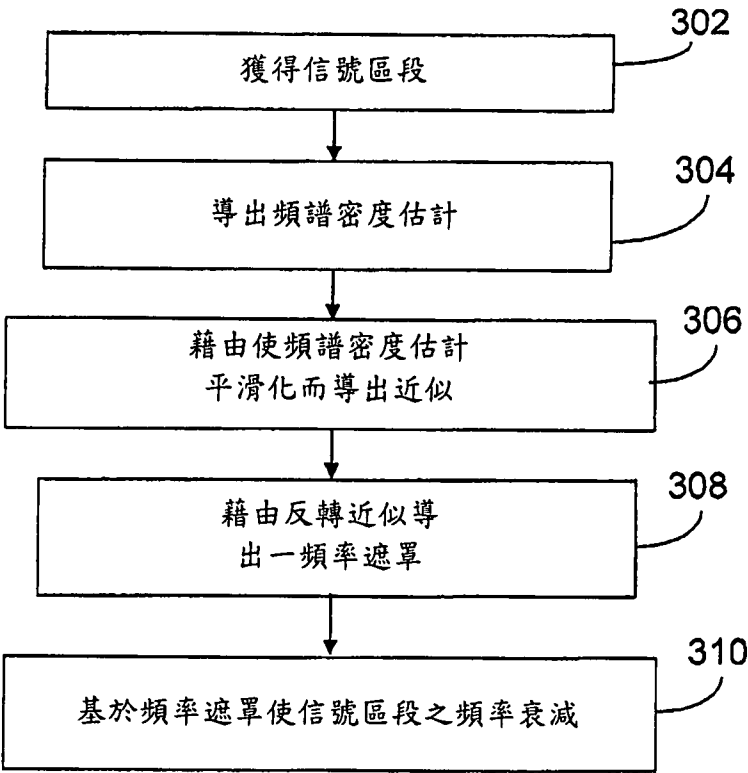


圖 3

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：101109568

※ 申請日：101年3月20日

※IPC 分類：~~G06F~~

G10L 21/02 (2013.01)

一、發明名稱：(中文/英文)

處理音訊信號之方法及裝置

METHOD AND APPARATUS FOR PROCESSING OF AUDIO SIGNALS

二、中文發明摘要：

本發明揭示一音訊處置實體中用於使一音訊信號之一時間區段中之優勢頻率衰減之方法及配置。獲得一音訊信號之一時間區段，並且導出該時間區段之頻譜密度或「頻譜」之一估計。藉由平滑化該估計導出該估計之一近似，且藉由反轉該近似而導出一頻率遮罩。接著基於該頻率遮罩使該音訊時間區段中包括之頻率衰減。該方法及配置不涉及多頻帶濾波或起音及釋放時間之選擇。

三、英文發明摘要：

Method and arrangement in an audio handling entity, for damping of dominant frequencies in a time segment of an audio signal. A time segment of an audio signal is obtained, and an estimate of the spectral density or "spectrum" of the time segment is derived. An approximation of the estimate is derived by smoothing the estimate, and a frequency mask is derived by inverting the approximation. Frequencies comprised in the audio time segment are then damped based on the frequency mask. The method and arrangement involves no multi-band filtering or selection of attack and release times.

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：

302	獲得信號區段
304	導出頻譜密度估計
306	藉由使頻譜密度估計平滑化而導出近似
308	藉由反轉近似導出一頻率遮罩
310	基於頻率遮罩使信號區段之頻率衰減

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】

本發明係關於音訊信號之處理，特定言之係關於用於使一音訊信號中之優勢頻率衰減之一方法及一配置。

【先前技術】

在語音源係透過麥克風在一特定地點擷取之音訊通信中，所獲得之信號位準(振幅)中之變動可能是明顯的。變動可能與若干因素有關，包含語音源與麥克風之間之距離、語音之響度及音高中之變動以及周圍環境之影響。當數位化所擷取之音訊信號時，信號位準中之明顯變動或波動可造成信號過載及限幅效應。此類不足可導致所擷取之音訊信號之適當後處理變得難以達到，且另外，雜散之資料過載可導致音訊產生地點之不悅收聽體驗。

減小此等不足或缺點之一通常方法為採用所擷取之信號之壓縮，其減小動態範圍，使得得以獲得所關注之信號之更小型振幅表示。一典型壓縮器使用一預定義臨限值以選擇需要關注之信號振幅。對於下行壓縮之考慮情形，減小高於預定義臨限值之信號位準達預設衰減因數或比率。

可以涉及不同級別之數學之若干方法實行動態範圍壓縮(DRC)。衰減因數通常為一固定值，但是其效應通常由「淡入」(起音)及「淡出」(釋放)時間間隔予以平滑化，其可視為衰減之時間變動。壓縮之位準可為頻率獨立的且因此對於信號中存在之全部頻率而言是固定的，或者可針對不同頻寬而動態計算該壓縮位準。

考慮到時變及頻率相依衰減之下行DRC之最先進方法，計算工作量可為龐大的。在實時應用中，若結合通信中常見之短時窗(通常10毫秒)上之全頻帶信號(24 kHz帶寬)之壓縮實行額外語音處理演算法，諸如例如聲響回音消除(AEC)或雜訊移除，則可能無法獲得多頻帶分析。

此外，由於在振幅超過預定義臨限值之每個情形中調變信號，故而時域中之振幅之習知壓縮會引入偽影。儘管此等效果之可聽度可能受限於起音及釋放時間之仔細選擇，然而聲音之波特徵仍被改變。此外，使用者參數之選擇，諸如壓縮比率、臨限值、起音及釋放時間是不明確的，且因此無瑣碎任務。

【發明內容】

期望達成對具有一波動信號位準之音訊信號之改良處理，因此避免根據上述先前技術之音訊信號處理之問題。本發明之一目的為解決上文概述之議題之至少一些。此外，本發明之一目的為提供用於在一音訊信號中實現優勢頻率之衰減之一方法和一配置。此等目的可由根據隨附獨立技術方案之一方法及一裝置滿足。在附屬技術方案中提出實施例。

音訊壓縮之概念已為人所熟知且在實際應用中常用。所提出技術之主要新穎之處在於其係基於一非參數性頻譜分析架構，且其在無需任何多頻帶濾波(濾波器庫)下以一頻率相依方式覆蓋整個頻帶。此外，此可使用具有產生一穩健結果之低計算複雜度之一穩固理論之聲音方法論來完

成。

所提出技術不需要選擇起音及釋放時間，因為振幅之斜率中不存在遽變，且因此在無任何壓縮之「淡入」或「淡出」下保留音訊信號之特性。然而，在針對各信號時間訊框個別計算壓縮之位準時，容許其為時變且充分資料相依的。

根據一第一態樣，提供一種在一音訊處置實體中用於使一音訊信號之一時間區段中之優勢頻率衰減之方法。該方法涉及獲得一音訊信號之一時間區段，以及導出該時間區段之「頻譜」之頻譜密度之一估計。藉由使該估計平滑化而導出該經估計之頻譜密度之一近似。藉由反轉該導出之近似而導出一頻率遮罩，且接著基於該頻率遮罩使該音訊時間區段中包括之頻率衰減。

根據一第二態樣，提供一種在一音訊處置實體中用於使一音訊信號之一時間區段中之優勢頻率衰減之配置。該配置包括經調適以獲得一音訊信號之一時間區段之一功能單元。該配置進一步包括經調適以導出該時間區段之該頻譜密度之一估計的一功能單元。該配置進一步包括經調適以藉由平滑化該估計而導出該頻譜密度估計之一近似的一功能單元，以及經調適以藉由反轉該近似而導出一頻率遮罩的一功能單元。該配置進一步包括經調適以基於該頻率遮罩而使該音訊時間區段中包括之頻率衰減的一功能單元。

可在不同實施例中實施以上方法及配置。在一些實施例中，導出之頻譜密度估計為一週期圖。在一些實施例中，

平滑化涉及倒頻譜分析，其中導出該頻譜密度估計之倒頻譜係數，且其中倒頻譜係數具有低於一特定臨限值之一絕對振幅值；或者，移除具有高於一預設臨限值之指數之連續倒頻譜係數。

在一些實施例中，頻率遮罩係組態成具有1之一最大增益，其使得在使用該頻率遮罩時不放大頻率。該頻率遮罩之最大衰減可預定義為一特定位準，或者經平滑化之估計頻譜密度可由該頻率遮罩中之未經平滑化之估計頻譜密度予以正規化。衰減可涉及在頻域中使該頻率遮罩乘以該估計頻譜密度，或者基於該頻率遮罩組態一有限脈衝回應(FIR)濾波器，用於在該時域中之該音訊信號時間區段上使用。

已主要在一方法方面描述以上實施例。然而，以上實施例亦旨在包含配置之實施例，該配置經調適以實現上述特徵之效能。可根據需要、需求或偏好以不同方式組合以上示例性實施例之不同特徵。

【實施方式】

現將藉由示例性實施例且參考附圖更詳細描述本發明。

簡要描述，振幅壓縮係在一音訊信號之最優勢之頻率下實行。可藉由使用頻域中之頻譜分析偵測該等最優勢頻率。藉由降低優勢頻率之增益(即，衰減)而非在整個信號之振幅增加高於一特定臨限值時實行壓縮，可保留聲音之正弦波特性。以一自動資料相依方式判定增加之增益(即，當對於全部頻率，增加之增益為0與1之間之一值時

之衰減)。

假定以一特定取樣速率(f_s)在時間上數位取樣一音訊信號。出於後處理及傳輸之原因，經取樣之信號被分成若干時間區段或長度 N 之「訊框」。此後，一個此訊框中之資料將標示為 $y_k(k=0, 2, \dots, N-1)$ 。

使用例如傅利葉分析且尤其快速傅利葉變換(FFT)後，可獲得一頻譜密度估計 Φ_p ，諸如資料 y_k 之週期圖

$$\Phi_p = \frac{1}{N} \left| \sum_{k=0}^{N-1} y_k e^{-i\omega_p k} \right|^2 \quad p=0, \dots, N-1 \quad (1)$$

其中

$$\omega_p = \frac{2\pi}{N} p$$

為傅利葉網格點。

通常，一音訊信號之週期圖具有一不穩定之行為。此可見圖1，其中一週期圖係以一細實線繪示。由於使用頻譜資訊將近似削弱信號中幾乎全部有用資訊，故而使用頻譜資訊，諸如週期圖作為何處實行信號壓縮之先前知識係極不直觀且不明智。

然而，其現已知藉由使用調用一相當量之平滑化，且因此排除細節及尖峰時估計頻譜之「基線」作為關於估優勢頻率之位置之先前資訊之一技術，壓縮可在未引入干擾偽影下於此等相關頻率下實行。對於週期圖之一經平滑化估計之計算，已使用涉及倒頻譜定限之一技術，然而可使用適於達成一經平滑化之頻譜密度估計之替代其他技術。

序列

$$c_k = \frac{1}{N} \sum_{p=0}^{N-1} \ln(\Phi_p) e^{i\omega_k p} + \gamma \delta_{k,0} \quad k = 0, \dots, N-1 \quad (2)$$

其中

$$\delta_{k,0} = \begin{cases} 1 & \text{若 } k=0 \\ 0 & \text{則} \end{cases} \quad \gamma = 0.577216...$$

悉知為與信號 y_k 有關之倒頻譜或倒頻譜係數。另外，已知 N 個倒頻譜係數之多者通常呈現少量值。因此，藉由以一理論上穩固方式(參見[1][2])將此等係數定限或截斷為零，可獲得(1)之一經平滑化估計為

$$\tilde{\Phi}_p = \alpha \hat{\Phi}_p \quad p = 0, \dots, N-1 \quad (3)$$

其中

$$\hat{\Phi}_p = \exp \left[\sum_{k=0}^{N-1} \hat{c}_k e^{-i\omega_p k} \right] \quad p = 0, \dots, N-1 \quad (4)$$

且其中

$$\alpha = \frac{\sum_{p=0}^{N-1} \Phi_p \hat{\Phi}_p}{\sum_{p=0}^{N-1} \hat{\Phi}_p^2}$$

為一正規化常數。在(4)中，序列 $\hat{c}_k$ 對應於(2)中之定限或截斷之序列 c_k 。

在圖1中，其表示48 kHz下取樣之語音信號之一典型10毫秒時間訊框(之頻率成分(frequency content))，使用[1]之倒頻譜定限演算法獲得之經平滑化之頻譜密度估計係展示為一粗虛線。顯然，虛線並非實線之細節之一精確估

計，此為其適切於該目的如此良好之原因所在。粗略估計具有最高頻率功率之頻率，導致一「滾動基線」。

圖1中經平滑化之頻譜密度估計(虛線)之反轉可用作含有在其下需要頻率壓縮之資訊的一頻率遮罩。若經平滑化之頻譜密度估計(虛線)已為頻譜密度估計(實線)之一精確估計，即，若平滑已不存在或極為有限，則將其使用作為信號訊框之一頻率遮罩將得到一極差且實際上毫無用處之結果。

藉由使頻率遮罩具有1之一最大增益值，可確保不會在任何頻率下實行信號之放大。對應於最大衰減之頻率遮罩之最小增益值可設定成一預設位準(5)，以確保佔優勢頻率「總是」衰減達一已知值。或者，可藉由使用例如未經平滑化之頻譜密度估計(例如週期圖)之最大值正規化經平滑化之頻譜密度估計而以一自動方式(6)設定最大壓縮或衰減之位準。

$$F_p = 1 - \lambda \frac{\tilde{\Phi}_p}{\max(\tilde{\Phi}_p)} \quad \text{其中 } 0 < \lambda < 1 \quad (5)$$

$$F_p = 1 - \frac{\tilde{\Phi}_p}{\max(\Phi_p)} \quad (6)$$

其中 $p=0, 2, \dots, N-1$ 。

圖2展示用於圖1中考慮使用由於無需選擇參數而為完全自動之(6)而獲得之信號訊框之所得頻率遮罩。即使(3)之計算可能涉及與一倒頻譜振幅臨限值[1][2]之值有關之一參數之一瑣碎選擇，其亦可視為自動的，使得當頻譜密度

估計具有一不穩定行為時選擇一較低參數值，且當頻譜密度估計具有一較小不穩定行為時選擇一較高參數值。對於音訊信號之情形，可將參數預定義為一恒定值。

若在一特定案例中使用(6)獲得之壓縮之位準不充分，則可使用(5)且使 λ 呈現為0與1之間之一期望值。

接著藉由在頻域中直接乘以估計頻譜密度而使用濾波器遮罩，以計算一經壓縮之資料組 $\hat{y}_k$ ($k=0, 2, \dots, N-1$)，或者例如作為一有限脈衝回應(FIR)濾波器之設計之輸入，其可應用於時域中之 y_k 。

實例程序圖3

現將參考圖3描述一音訊信號之一時間區段中之衰減優勢頻率之程序之一示例性實施例。該程序可在一音訊處置實體中實行，諸如例如一電話會議系統中之一節點及/或一無線或有線通信系統中之一節點或終端機、音訊廣播中涉及之一節點，或音樂製作中使用之一實體或器件。

在一動作302中獲得一音訊信號之一時間區段。假定該音訊信號是由一麥克風或類似者擷取，且假定其將以一取樣頻率取樣。該音訊信號可包括例如由參與一電話會議或一些其他類型之通信會期之一或多個揚聲器產生之語音。時間區段可為例如約10毫秒或適於信號處理之任何其他長度。

所導出之時間區段之頻譜密度之一估計(頻域中)係在一動作304中獲得。此估計可為例如一週期圖，且可藉由使用諸如FFT之一傅利葉變換方法而導出。藉由使頻譜密度

估計平滑化而在一動作306中導出該估計頻譜密度之一近似。該近似實際上應為「粗略的」，即並非極接近頻譜密度估計，其通常對於諸如例如語音或音樂之音訊信號而言是不穩定的(參照圖1)。可例如藉由使用一倒頻譜定限演算法、移除(在倒頻譜域中)具有低於一特定臨限值之一絕對振幅值之倒頻譜係數、或移除具有高於一預設臨限值之一指數之連續倒頻譜係數而導出該近似。

藉由反轉導出之近似(即經平滑化之頻譜密度估計)而在一動作308中從頻譜密度估計之經導出之近似中導出一頻率遮罩。接著在一動作310中使用或應用該頻率遮罩用於使信號時間區段中包括之頻率衰減。衰減可涉及在頻率域中使該頻率遮罩乘以估計之頻譜密度，或者可基於該頻率遮罩組態一FIR濾波器，此FIR濾波器可在時域中用於音訊信號時間區段上。

可以不同方式組態該頻率遮罩。舉例而言，可將該頻率遮罩之最大增益設為1，因此確保無信號頻率將在基於該頻率遮罩而處理時被放大。此外，可將頻率遮罩之最大衰減(最小增益)預定義為一特定位準，或可由頻率遮罩中之未經平滑化之估計頻譜密度而正規化經平滑化之估計頻譜密度。

實例配置，圖4

下文將參考圖4描述一實例配置400，其經調適以實現與一音訊信號之一時間區段中之優勢頻率之衰減有關之上述程序之效能。該配置係繪示成位於一通信系統中之一音訊

處置實體401中。該音訊處置實體可為例如一電話會議系統中之一節點或終端機及/或一無線或有線通信系統中之一節點或終端機、音訊廣播中涉及之一節點、或音樂製作中使用之一實體或器件。該配置400進一步係繪示成經由一通信單元402而與其他實體通信，此可視為包括用於無線及/或有線通信之習知構件。該配置及/或音訊處置實體可進一步包括其他常規功能單元416，以及一或多個儲存單元414。

該配置400包括一獲得單元404，其經調適以獲得一音訊信號之一時間區段。該音訊信號可包括例如由參與一電話會議或一些其他類型之通信會期之一或多個揚聲器產生之語音。舉例而言，可獲得表示例如10毫秒之一時間間隔之一組連續取樣。假定該音訊信號已由一麥克風或類似者擷取且以一取樣頻率取樣。該音訊信號可已經由獲得單元404、由音訊處置實體401中或在另一節點或實體中之其他功能單元擷取且/或取樣。

該配置進一步包括一估計單元406，其經調適以導出時間區段之頻譜密度之一估計。該單元406可經調適以例如藉由使用諸如FFT之一傅利葉變換方法而導出例如一週期圖。此外，該配置包括一平滑化單元408，其經調適以藉由平滑化該估計而導出頻譜密度估計之一近似。該近似應為相當「粗略的」，即並非極接近頻譜密度估計，其通常對於諸如例如語音或音樂之音訊信號而言是不穩定的(參照圖1)。該平滑化單元408可經調適以藉由使用一倒頻譜

定限演算法、移除(在倒頻譜域中)具有低於一特定臨限值之一絕對振幅值之倒頻譜係數、或移除具有高於一預設臨限值之一指數之連續倒頻譜係數而達成經平滑化頻譜密度估計。

該配置400進一步包括一遮罩單元410，其經調適以藉由反轉經估計之頻率密度之近似而導出一頻率遮罩。該遮罩單元410可經調適以將頻率遮罩之最大增益組態成1，因此確保不會放大頻率。該遮罩單元410可進一步經調適以將該頻率遮罩之最大衰減組態成一特定預定義位準，或者在導出該頻率遮罩時由未經平滑化之估計頻譜密度正規化經平滑化之估計頻譜密度。

此外，該配置包括一衰減單元412，其經調適以基於頻率遮罩而使音訊時間區段中包括之頻率衰減。該衰減單元412可經調適以例如使頻率遮罩乘以頻域中之經估計之頻譜密度，或者經調適以基於頻率遮罩組態一FIR濾波器，且經組態以使用該FIR濾波器用於對時域中之音訊信號時間區段進行濾波。

示例性替代配置，圖5

圖5繪示一音訊處置實體中之一替代配置500，其中一電腦程式510係由連接至一處理器506之一電腦程式產品508攜載。該電腦程式產品508包括在其上儲存電腦程式510之一電腦可讀媒體。該電腦程式510可組態成在電腦程式模組中結構化之一電腦程式碼。因此，在所述實例實施例中，電腦程式510中之碼構件包括用於獲得一音訊信號之

一時間區段之一獲得模組510a。該電腦程式進一步包括用於導出該時間區段之頻譜密度之一估計之一估計模組510b。該電腦程式510進一步包括用於藉由平滑化該估計而導出該頻譜密度估計之一近似之一平滑化模組510c；以及用於藉由反轉經估計之頻譜密度之近似而導出一頻率遮罩之一遮罩模組510d。該電腦程式進一步包括用於基於頻率遮罩而使音訊時間區段中包括之頻率衰減之一衰減模組510e。

該等模組510a至510e可基本上實行圖3所繪示之流程之動作，以模擬圖4所繪示之一音訊處置實體中之配置。換言之，當在處理單元506中執行不同模組510a至510e時，其等對應於圖4之單元404至412之各自功能。舉例而言，電腦程式產品可為一快閃記憶體、一RAM(隨機存取記憶體)ROM(唯讀記憶體)或一EEPROM(電可擦除可程式化ROM)，且在替代實施例中該電腦程式模組510a至510e可在配置500及/或收發機節點內以多個記憶體形式分佈於不同電腦程式產品中。連接至處理器之單元502及504表示例如輸入及輸出之通信單元。單元502及單元504可配置成整合型實體。

儘管上文結合圖5所揭示之實施例中之碼構件係實施為電腦程式模組，其在處理單元中執行時造成配置及/或收發機節點實行上文結合上述圖式所述之動作，然而該等碼構件之至少一者可在替代實施例中至少部分實施為硬體電路。

應注意，互動單元或模組之選擇，以及該等單元之命名僅為示例性目的，且可以複數個替代方式組態適於執行上述方法之任何者之網路節點以可執行所提出之處理動作。

亦應注意本揭示內容中所述之單元或模組應視為邏輯實體且非必需視為單獨物理實體。

縮寫

AEC	聲響回音控制
DRC	動態範圍壓縮
FIR	有限長度脈衝回應
FFT	快速傅利葉變換

參考

- [1] Stoica, P., Sandgren, N. Smoothed Nonparametric Spectral Estimation via Cepstrum Thresholding. IEEE Sign. Proc. Mag. 2006.
- [2] Stoica, P., Sandgren, N. Total Variance Reduction via Thresholding: Application to Cepstral Analysis. IEEE Trans. Sign. Proc. 2007.

【圖式簡單說明】

圖1展示根據一示例性實施例之一音訊信號區段之一頻譜密度估計(實線)及一經平滑化之頻譜密度估計(虛線)。

圖2展示根據一示例性實施例基於一經平滑化之頻譜密度估計之一頻率遮罩。

圖3係繪示根據一示例性實施例之一音訊處置實體中之一程序之一流程圖。

圖 4 至 圖 5 係繪示根據一示例性實施例之一音訊處置實體中之一各自配置之方塊圖。

【主要元件符號說明】

400	配置
401	音訊處置實體
402	通信單元
404	獲得單元
406	估計單元
408	平滑化單元
410	遮罩單元
412	衰減單元
414	儲存單元
416	常規功能單元
500	替代配置
502	單元
504	單元
506	處理器
508	電腦程式產品
510	電腦程式
510a	獲得模組
510b	估計模組
510c	平滑化模組
510d	遮罩模組
510e	衰減模組

七、申請專利範圍：

1. 一種在一音訊處置實體中用於使一音訊信號之一時間區段中之優勢頻率衰減(damping)之方法，該方法包括：
 - 獲得一音訊信號之一時間區段；
 - 導出該時間區段之頻譜密度之一估計；
 - 藉由平滑化該估計而導出該經估計之頻譜密度之一近似；
 - 藉由反轉該經估計之頻譜密度之該近似而導出一頻率遮罩，該反轉之輸出產生一頻域信號作為該頻率遮罩；
 - 及
 - 基於該頻率遮罩而使該音訊時間區段中包括之頻率衰減。
2. 如請求項1之方法，其中該平滑化涉及導出該頻譜密度估計之倒頻譜(cepstral)係數，以及至少以下一者：
 - 移除具有低於一特定臨限值之一絕對振幅值的倒頻譜係數；及
 - 移除具有高於一預設臨限值之指數之連續倒頻譜係數。
3. 如請求項1之方法，其中該頻率遮罩係組態成具有1之一最大增益。
4. 如請求項1之方法，其中該頻率遮罩之最大衰減係預定義為一特定位準。
5. 如請求項1之方法，其中該頻率遮罩 F_p 定義為：

$$F_p = 1 - \lambda \frac{\bar{\phi}_p}{\max(\bar{\phi}_p)};$$

其中 λ 為 $0 < \lambda < 1$ ，且 $p=0、\dots、N-1$ ；其中 N 為該音訊信號時間區段之取樣之數目；且 $\bar{\phi}_p$ 為該經平滑化之估計頻譜密度。

6. 如請求項1之方法，其中在該頻率遮罩中，該經平滑化之估計頻譜密度係由未經平滑化之估計頻譜密度正規化。

7. 如請求項1之方法，其中該頻率遮罩 F_p 定義為：

$$F_p = 1 - \frac{\bar{\phi}_p}{\max(\phi_p)};$$

其中 $p=0、\dots、N-1$ ；且其中 N 為該音訊信號時間區段之取樣之數目， ϕ_p 為該經估計頻譜密度，且 $\bar{\phi}_p$ 為該經平滑化之估計頻譜密度。

8. 如請求項1之方法，其中該信號區段之該頻譜密度之該估計為一週期圖。

9. 如請求項1之方法，其中該衰減涉及至少以下一者：

在頻域中使該頻率遮罩乘以該估計頻譜密度；及

基於該頻率遮罩組態一有限脈衝回應(FIR)濾波器，用於在時域中之該音訊信號時間區段上使用。

10. 如請求項1之方法，其中該平滑化係非參數式的(non-parametric)。

11. 如請求項5之方法，其中該經平滑化之估計頻譜密度定義為：

$$\tilde{\Phi}_p = \alpha \hat{\Phi}_p, \text{ 其中 } \hat{\Phi}_p = \exp \left[\sum_{k=0}^{N-1} \hat{c}_k e^{-i\omega_p k} \right]$$

其中 ω_p 為一序列傅利葉網格點，其中 $p=0, \dots, N-1$ ，其中 N 為該音訊信號時間區段之取樣之數目，其中 α 為一正規化常數，且其中序列 $\hat{c}_k$ 係經修改倒頻譜係數序列。

12. 如請求項11之方法，其中該正規化常數 α 定義為：

$$\alpha = \frac{\sum_{p=0}^{N-1} \Phi_p \hat{\Phi}_p}{\sum_{p=0}^{N-1} \hat{\Phi}_p^2}, \text{ 其中 } \hat{\Phi}_p = \exp \left[\sum_{k=0}^{N-1} \hat{c}_k e^{-i\omega_p k} \right]$$

其中 ω_p 為一序列傅利葉網格點，其中 $p=0, \dots, N-1$ ，其中 N 為該音訊信號時間區段之取樣之數目，其中 α 為一正規化常數，且其中序列 $\hat{c}_k$ 係該第二倒頻譜係數序列。

13. 如請求項1之方法，其中該反轉該經估計之頻譜密度之該近似包含1除以該經估計之頻譜密度之該近似之函數。

14. 一種音訊信號處理裝置，其包括一處理器及記憶體，該記憶體含有可由該處理器執行之指令，藉此該音訊信號處理裝置可操作以：

獲得一音訊信號之一時間區段；

導出該時間區段之頻譜密度之一估計；

藉由平滑化該估計而導出該頻譜密度估計之一近似；

藉由反轉該估計頻譜密度之該近似而導出一頻率遮罩，該反轉之輸出產生一頻域信號作為該頻率遮罩；及

基於該頻率遮罩而使該音訊時間區段中包括之頻率衰

減。

15. 如請求項14之音訊信號處理裝置，其中該平滑化包含導出該頻譜密度估計之倒頻譜係數及根據一預定義規則移除倒頻譜係數。

16. 如請求項15之音訊信號處理裝置，其中該預定義規則涉及以下一者：

移除具有低於一特定臨限值之一絕對振幅值的倒頻譜係數；及

移除具有高於一預設臨限值之指數之連續倒頻譜係數。

17. 如請求項14之音訊信號處理裝置，其中該頻率遮罩經組態以具有為1之最大增益。

18. 如請求項14之音訊信號處理裝置，其中該該頻率遮罩之最大衰減係經預定義成一特定位準。

19. 如請求項14之音訊信號處理裝置，其中在該頻譜遮罩中，該經平滑化之估計頻譜密度係藉由未經平滑化之估計頻譜密度而正規化。

20. 如請求項14之音訊信號處理裝置，其中該衰減涉及以下至少一者：

在頻域中使該頻率遮罩乘以該估計頻譜密度；及

基於該頻率遮罩組態一FIR濾波器，用於在時域中之該音訊信號時間區段上使用。

21. 如請求項14之音訊信號處理裝置，其中該平滑化係非參數式的。

22. 如請求項19之音訊信號處理裝置，其中該經平滑化之估計頻譜密度定義為：

$$\tilde{\Phi}_p = \alpha \hat{\Phi}_p, \text{ 其中 } \hat{\Phi}_p = \exp \left[\sum_{k=0}^{N-1} \hat{c}_k e^{-i\omega_p k} \right]$$

其中 ω_p 為一序列傅利葉網格點，其中 $p=0, \dots, N-1$ ，其中 N 為該音訊信號時間區段之取樣之數目，其中 α 為一正規化常數，且其中序列 $\hat{c}_k$ 係經修改倒頻譜係數序列。

23. 如請求項22之音訊信號處理裝置，其中該正規化常數 α 定義為：

$$\alpha = \frac{\sum_{p=0}^{N-1} \Phi_p \hat{\Phi}_p}{\sum_{p=0}^{N-1} \hat{\Phi}_p^2}, \text{ 其中 } \hat{\Phi}_p = \exp \left[\sum_{k=0}^{N-1} \hat{c}_k e^{-i\omega_p k} \right]$$

其中 ω_p 為一序列傅利葉網格點，其中 $p=0, \dots, N-1$ ，其中 N 為該音訊信號時間區段之取樣之數目，其中 α 為一正規化常數，且其中序列 $\hat{c}_k$ 係該第二倒頻譜係數序列。

八、圖式：

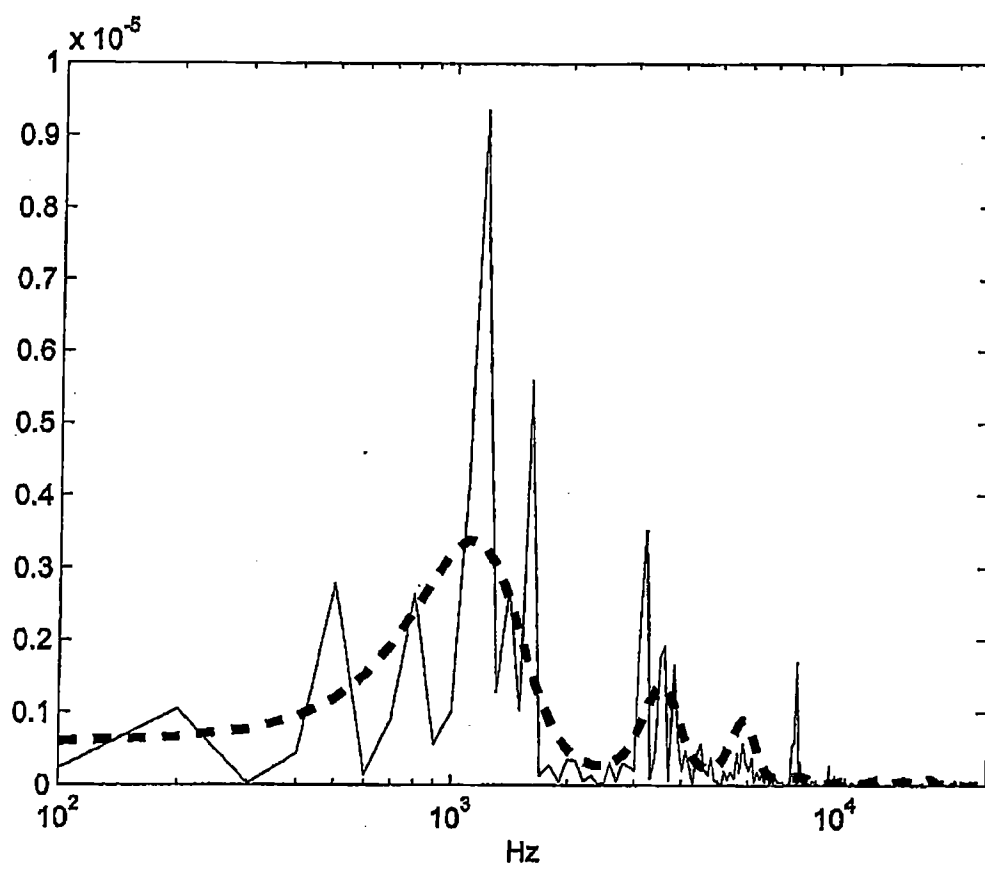


圖 1

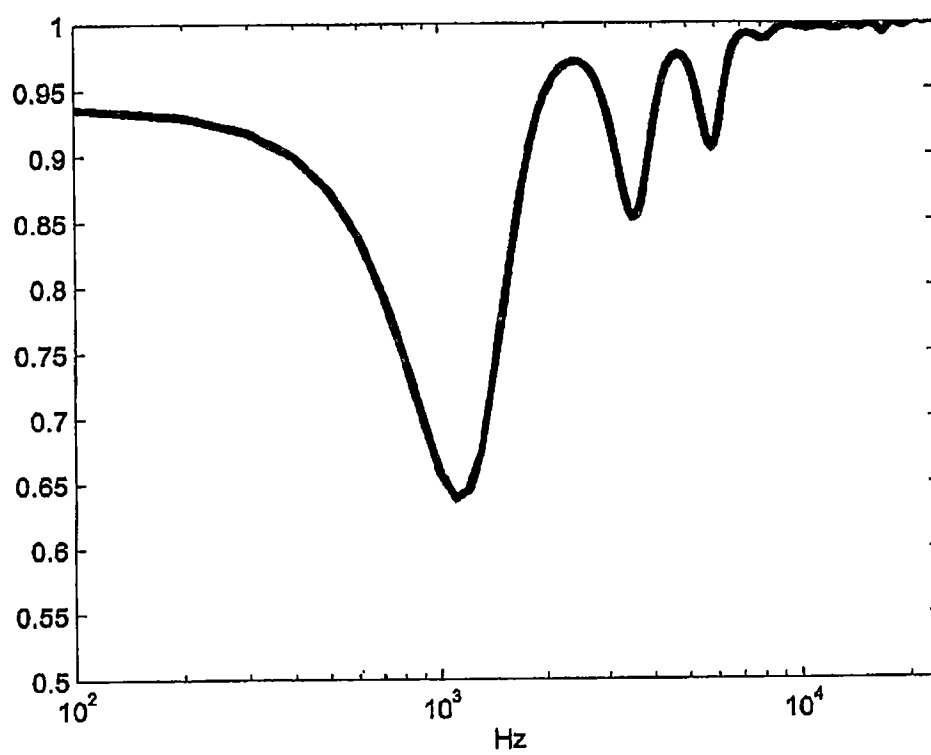


圖 2

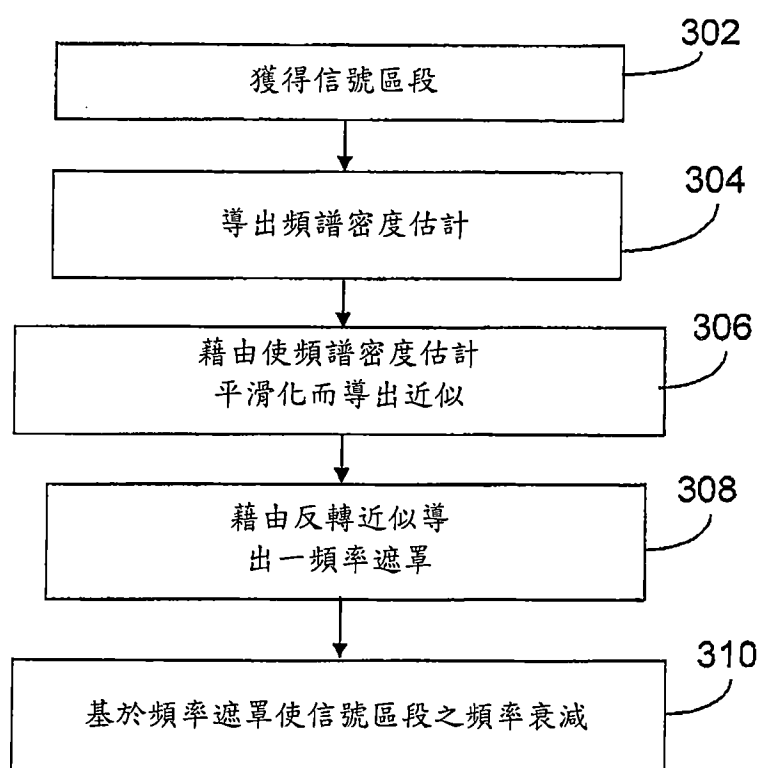


圖 3

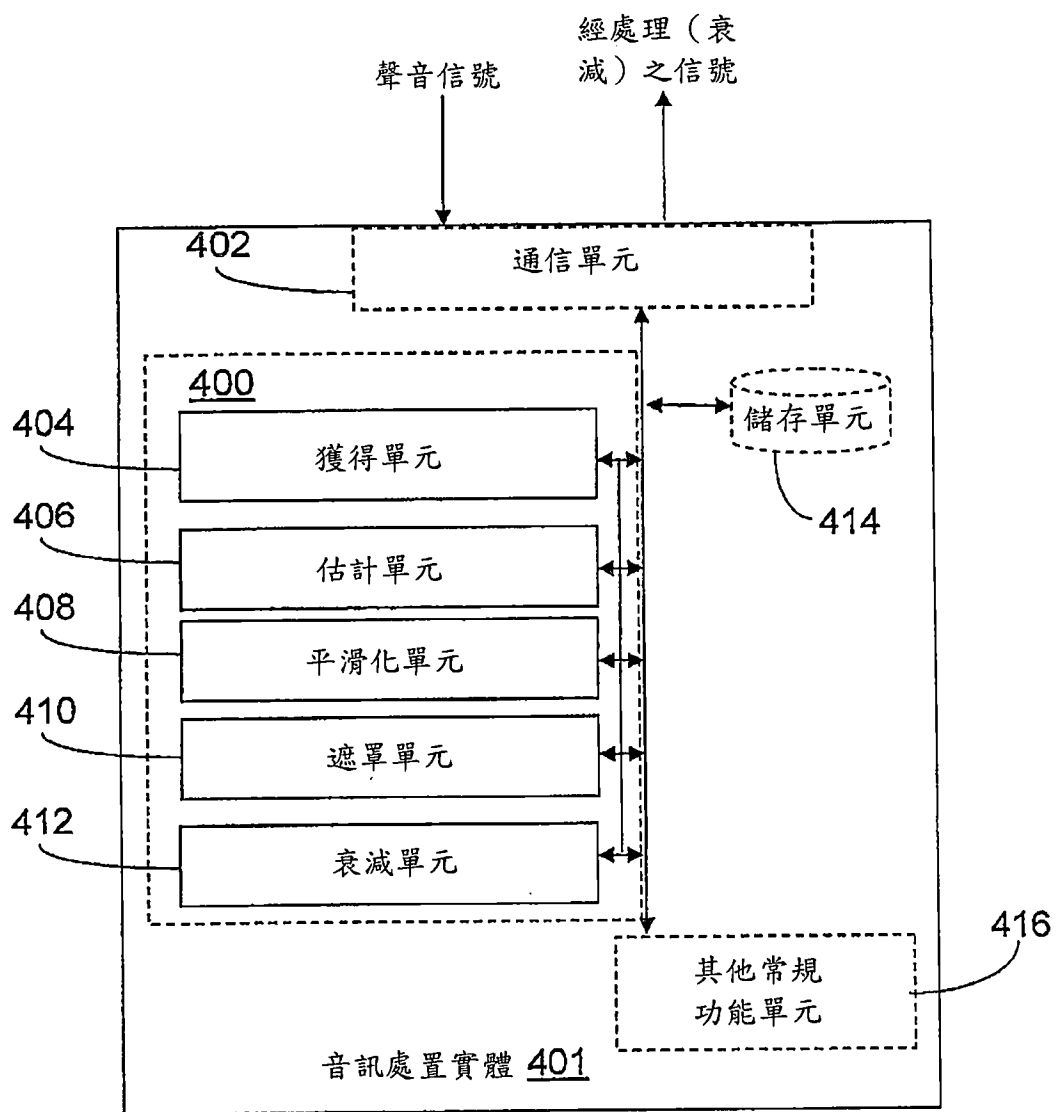


圖 4

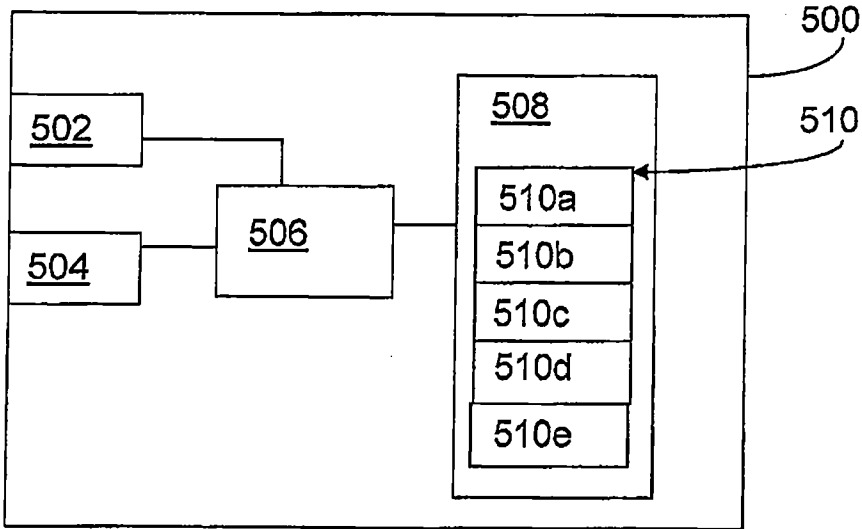


圖 5