



(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 I882056 B

(45)公告日：中華民國 114 (2025) 年 05 月 01 日

(21)申請案號：109145090

(22)申請日：中華民國 109 (2020) 年 12 月 18 日

(51)Int. Cl. : **H04R3/04 (2006.01)****H03G9/02 (2006.01)****G06F3/16 (2006.01)**

(30)優先權：2019/12/19

歐洲專利局

19306698.2

(71)申請人：法商斯蒂爾塞瑞斯法國公司 (法國) STEELSERIES FRANCE (FR)

法國

(72)發明人：格蘭傑 達米安 GRANGER, DAMIEN (FR)；蘇維拉拉巴斯蒂 內森 SOUVIRAA-

LABASTIE, NATHAN (FR)；德萊普蘭克 安東尼 DELEPLANQUE, ANTOINE

(FR)

(74)代理人：劉法正；尹重君

(56)參考文獻：

TW 201347404A

US 2013/0013096A1

US 2016/0072467A1

US 2017/0011748A1

US 2019/0075395A1

審查人員：林宥榆

申請專利範圍項數：15 項 圖式數：11 共 41 頁

(54)名稱

用以藉由設備進行音訊呈現之方法

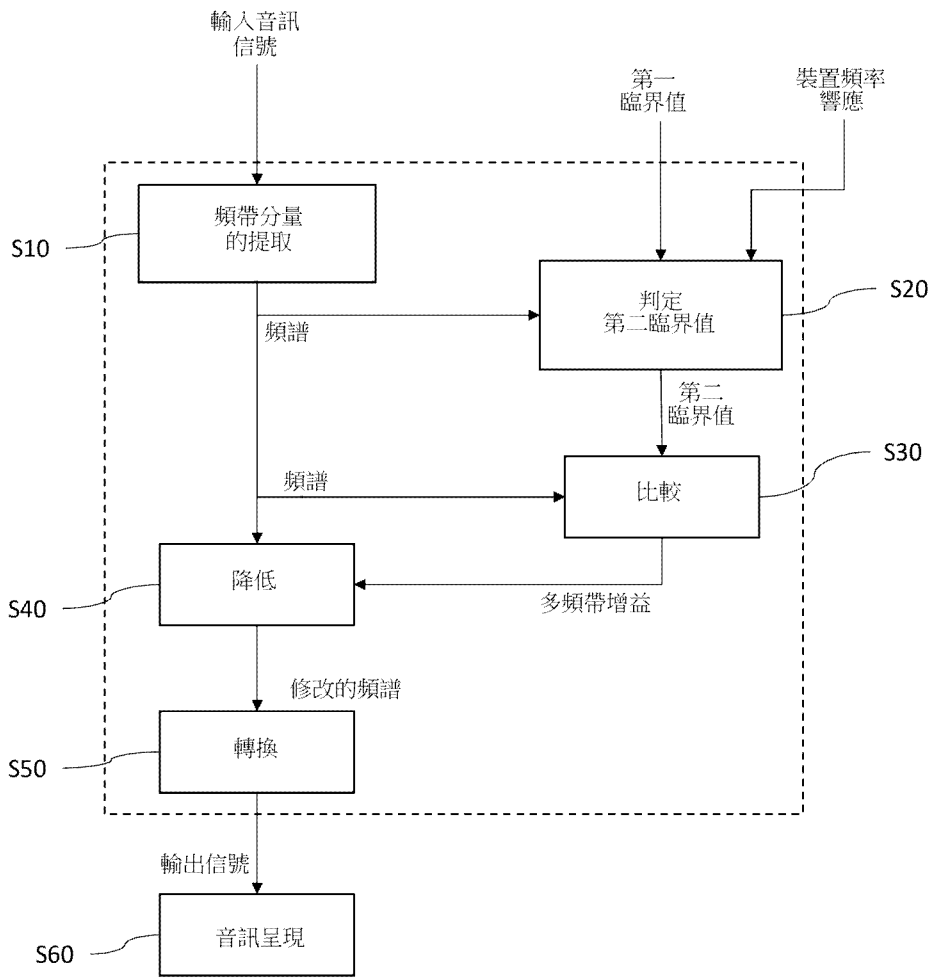
(57)摘要

一種用以藉由一種包含有至少一個音訊呈現裝置的設備進行音訊呈現之方法，該方法包含有：從一輸入音訊信號中提取複數個頻帶分量；從該裝置頻率響應及該等複數個提取頻帶分量中決定至少一個代表掩蔽頻率能量的指示符，掩蔽頻率對應於高於一頻率臨界值之頻帶；從該代表掩蔽頻率能量的指示符決定至少一個校正因子；對於每一個頻帶，藉由使用該校正因子修改與該頻帶相關聯之一預定第一聲級臨界值來決定一第二聲級臨界值，及從在該提取頻帶分量的一聲級和與該提取頻帶相關聯之該第二聲級臨界值之間的一比較決定一降低增益，以及施加該降低增益。

A method for audio rendering by an apparatus comprising at least one audio rendering device, the method comprising: extracting (S10) a plurality of frequency band components from an input audio signal; determining (S15) from the device frequency response and the plurality of extracted frequency band components at least one indicator representative of masking frequencies energy, masking frequencies corresponding to frequency bands that are above a frequency threshold; determining at least one correction factor (S16) from the indicator representative of masking frequencies energy; e) for each frequency band, determining a second acoustic level threshold (S20) by modifying (S17) with the correction factor a predetermined first acoustic level threshold associated with said frequency band, and determining a reduction gain from a comparison (S30) between an acoustic level of the extracted frequency band and the second acoustic level threshold associated to said extracted frequency band, and applying (S40) the reduction gain.

指定代表圖：

符號簡單說明：  
S10~S60:方塊



【圖1】



I882056

**【發明摘要】****【中文發明名稱】**

用以藉由設備進行音訊呈現之方法

**【英文發明名稱】**

A METHOD FOR AUDIO RENDERING BY AN APPARATUS

**【中文】**

一種用以藉由一種包含有至少一個音訊呈現裝置的設備進行音訊呈現之方法，該方法包含有：從一輸入音訊信號中提取複數個頻帶分量；從該裝置頻率響應及該等複數個提取頻帶分量中決定至少一個代表掩蔽頻率能量的指示符，掩蔽頻率對應於高於一頻率臨界值之頻帶；從該代表掩蔽頻率能量的指示符決定至少一個校正因子；對於每一個頻帶，藉由使用該校正因子修改與該頻帶相關聯之一預定第一聲級臨界值來決定一第二聲級臨界值，及從在該提取頻帶分量的一聲級和與該提取頻帶相關聯之該第二聲級臨界值之間的一比較決定一降低增益，以及施加該降低增益。

**【英文】**

A method for audio rendering by an apparatus comprising at least one audio rendering device, the method comprising: extracting (S10) a plurality of frequency band components from an input audio signal; determining (S15) from the device frequency response and the plurality of extracted frequency band components at least one indicator representative of masking frequencies energy, masking frequencies corresponding to frequency bands that are above a frequency threshold; determining at least one correction factor (S16) from the indicator representative of masking frequencies energy; e) for each frequency band, determining a second acoustic level threshold (S20) by modifying (S17) with the correction factor a predetermined first acoustic level threshold associated with said frequency band, and determining a reduction gain from a comparison (S30) between an acoustic level of the extracted frequency band and the second acoustic level threshold associated to said extracted frequency band, and applying (S40) the reduction gain.

【指定代表圖】 圖1

【代表圖之符號簡單說明】

S10～S60:方塊

【特徵化學式】

無

## 【發明說明書】

### 【中文發明名稱】

用以藉由設備進行音訊呈現之方法

### 【英文發明名稱】

A METHOD FOR AUDIO RENDERING BY AN APPARATUS

### 【技術領域】

發明領域

【0001】 本發明涉及一種用以藉由包含有至少一個音訊呈現裝置之設備進行音訊呈現的方法。

### 【先前技術】

發明背景

【0002】 一音訊呈現裝置係一種接收音訊資料並播放對應於該接收音訊資料之聲音的裝置。通常，一音訊呈現裝置係一揚聲器，即，把一電氣音訊信號轉換成一相應聲音的電聲換能器。

【0003】 許多現代電子設備設有嵌入式揚聲器，以允許使用者收聽多種類型的內容。例如，該設備可以是一膝上型電腦、一平板電腦、或一智慧型手機。為了最小化此類設備的尺寸，製造商通常會選擇小型揚聲器，但這會導致低的響度。製造商已試圖設計該等包含有揚聲器的盒子，並進行其他的聲學最佳化以利用共振現象來改善該低頻的聲級，但是不足夠的。除此之外，較高頻率的聲級常常係不足夠的。這導致一低的響度以及該聲音有一不良的平衡。

【0004】 嵌入式揚聲器的該低響度通常使用放大器來補償。但這經常會導致該等揚聲器被過載，導致出有偽音諸如在該呈現聲音中有失真或有嗡嗡聲。透過使用諸如等化、壓縮器、或限制器之類的幾種效果，可以改善該所呈現聲音的響度及平衡，同時限制了在該等揚聲器上的該等約束。這些效果可以用硬體實

現，例如藉由所謂的智慧型放大器；或者用軟體實現，藉由一處理方法來實現。此外，該等智慧型放大器還監控該揚聲器偏移及該揚聲器溫度，以自適應該放大增益從而避免過載。

**【0005】** 此外，該等揚聲器通常與其他組件諸如鍵盤按鍵絕緣不良，當在該等揚聲器上播放音訊內容時，鍵盤組件會振動。這些振動係可聞的並會使聽者對聲音品質的整體感覺惡化，即使有了良好的響度及平衡也是如此。該等智慧型放大器並不考慮該等揚聲器的環境；它們僅用於解決揚聲器的過載、響度以及平衡問題，因此無助於減少此類聲音的惡化。

**【0006】** 該等振動通常係由於該音訊內容激發頻率(例如349 Hz、740 Hz)之一有限列表所引起的。一種簡單的處理方法包括對包含有該等問題頻率之一頻帶的該聲級施加一固定的降低增益，例如藉由施加一陷波濾波器或頻帶切割器(具有一非常窄的頻帶)。但是，這種方法可能會使得聽眾所聽到的聲音品質下降。

**【0007】** 美國專利第US9762198號描述了一種用於動態調整一壓縮器之臨界值的方法。具有多個頻帶分量之一輸入音訊信號被處理。隨時間變化的臨界值可被決定。一壓縮器在每一個頻帶分量上執行具有一相應時變臨界值之一壓縮操作以產生增益。每一個增益被施加到一經延遲之對應頻帶分量以產生處理後的頻帶分量，該等處理後的頻帶分量被相加以產生一輸出信號。在一些實施方式中，該輸出信號之一感知頻譜的一時變估計及由該感知頻譜估計所引起之一失真頻譜的一時變估計被決定，例如，使用一種基於諧波產生之失真可聞度模型。該模型可以透過預先在該裝置上進行測量而獲得。但是，這種方法不能準確地反映每一個揚聲器的行為。在同一模型的樣本之間有太多的變化，並且在同一裝置上所做的兩個測量之間也有太多的變化。也可以透過一偏移量更簡單地來估計該模型，但這會導致過多的近似。而且，該失真可聞度係透過一聽覺心理遮蔽效應的一估計來計算的。此外，這一種計算非常耗費資源，因此需要近似來減輕該

運算量。因此，該結果非常劣質。另外，該等增益被施加在其上之該等頻帶與被使用於該臨界值的那些頻帶相同。然而，該等頻帶可以包含各種不同的聲音，並且該降低導致了在該音訊內容之該音調上產生一明顯的影響。

**【0008】** 專利申請第US2019/075395號描述了一種方法，其中一音訊信號的頻譜傾斜被使用來決定一揚聲器在該音訊信號的回放期間是否將引入可感知到的失真。該頻譜傾斜係指能量係如何地被分配在該音訊信號的整個頻譜中。可藉由決定在一失真產生頻帶中的能量與在一失真掩蔽頻帶中的能量間的一比值來指出該頻譜傾斜。基於該決定的頻譜傾斜，該失真產生頻帶可被衰減以降低由該揚聲器所引入的該失真。更具體地說，該信號準位在該輸入音訊信號的一衰減頻帶中被衰減以基於在該輸入音訊信號之一失真掩蔽頻帶的該信號準位與該失真掩蔽頻帶的該信號準位之間的一比值來產生一經修改的音訊信號。

## **【發明內容】**

### 發明概要

**【0009】** 本發明提出了在保持最佳可能之響度及平衡的同時，使得諸如振動、失真或嗡嗡聲之類的偽音不可聞，即藉由最小化在該呈現聲音上的影響。為此，該方法僅在必要時以取決於該聲音內容的一種動態方式來調整該聲音。

**【0010】** 為達此目的，本發明係一種用以藉由包含有至少一個音訊呈現裝置的設備進行音訊呈現之方法，該方法包含有：

- a) 接收一輸入音訊信號，
- b) 從該輸入音訊信號中提取複數個頻帶分量，每一個頻帶分量具有與一頻帶相關聯的一聲級，
- c) 從該等複數個提取的出頻帶分量中決定至少一個代表掩蔽頻率能量的指示符，掩蔽頻率對應於高於一頻率臨界值之頻帶，代表掩蔽頻率能量的該指示符僅從高於該頻率臨界值的頻帶來決定，

- d) 從該至少一個代表掩蔽頻率能量的指示符中決定至少一個校正因子，
- e) 對於每一個頻帶，藉由使用該校正因子修改與該頻帶相關聯之一預定第一聲級臨界值來決定一第二聲級臨界值，其中該第一聲級臨界值獨立於該輸入音訊信號的該內容，該第一聲級臨界值指出可以在沒有明顯偽音的情況下由該設備播放之該頻帶的一最高聲級，以及其中該等第二聲級臨界值取決於該輸入音訊信號，
- f) 對於每一個頻帶，從在該頻帶之該提取頻帶分量的一聲級和與該提取頻帶相關聯之該第二聲級臨界值之間的一比較，決定一降低增益，以及
- g) 把該降低增益施加到該等頻帶分量的該聲級，
- h) 把該等不同的頻帶分量轉換為一輸出音訊信號，
- i) 藉由該音訊呈現裝置對該輸出音訊信號進行音訊呈現。

**【0011】** 當該環境允許時，該方法可以限制該振動降低的使用。例如，當感知上不需要時，即當該環境的該聲級與該等振動之該估計聲級相比係足夠高時，該方法將施加很少的降低或不降低，因此可以從一感知的角度將其掩蔽。以這種方式，該音訊內容將被較少地改變，而有一較高的音訊呈現保真度。

**【0012】** 本發明方法之其他優選的但非限制性的方面如下，它們係獨立地被使用或使用在一技術上可行的組合中：

-該音訊呈現裝置具有一預定的裝置頻率響應，並且還根據該裝置頻率響應來決定該表示掩蔽頻率能量的指示符；

-該頻率臨界值高於或等於6 kHz，優選地高於或等於8 kHz，並且更優選地高於或等於10kHz；

-複數個代表掩蔽頻率能量的指示符被決定，該等代表掩蔽頻率能量的指示符係從對應於由不同頻率臨界值所限定之不同頻率子範圍的頻帶來決定的；

-從該等複數個代表掩蔽頻率能量的指示符決定數個校正因子，並且使用不

同的校正因子來決定不同頻帶的第二聲級臨界值，取決於該頻帶來使用該校正因子；

-至少一個頻帶與若干個代表掩蔽頻率能量的指示符相關聯，並且被使用來決定與該頻帶相關聯之該第二聲級臨界值的該校正因子係從與該頻帶相關聯之該等若干個代表掩蔽頻率能量的指示符中至少一個來決定的；

-從每一個代表掩蔽頻率能量的指示符決定一校正因子，得到與該頻率子範圍之頻帶相關聯的複數個校正因子；

-藉由使用一選擇的校正因子來決定該第二聲級臨界值，該選擇的校正因子係從與該第二聲級相關聯之至少一個頻帶相關聯之複數個校正因子中所選擇出的，該選擇的校正因子被選擇用於產生在第二聲級臨界值中一最低的第二聲級臨界值，該等第二聲級臨界值係使用該等複數個校正因子所決定的；

-頻率的兩個子範圍具有一個共同的頻帶，導致會有兩個代表掩蔽頻率的指示符與該共同頻帶相關聯，不是該等兩個代表掩蔽頻率之指示符中非該所選指示符的一個指示符被設置為零，就是該第一聲級臨界值被未修改；

-該等頻帶分量被提取作為頻帶信號分量，並且一頻帶信號分量的該聲級係該頻帶信號分量的一振幅，或者

-該等頻帶分量被提取作為頻帶分格分量，並且每一個頻帶係一頻率分格，並且其中步驟b)包含把該輸入音訊信號轉換成在一頻域中的頻帶分格分量，其中一頻帶分格分量的該聲級係一頻帶分格分量的一大小，並且其中步驟h)包含把頻帶分格分量轉換為一時域；

-在步驟c)中所決定之該代表掩蔽頻率能量的指示符係高於該頻率臨界值之頻帶分量之聲級總和的一函數；

在步驟c)中所決定之該代表掩蔽頻率能量的指示符係該頻譜之統計特性諸如變異數的一函數；

-在該輸入音訊信號中該掩蔽頻率能量越高，則該校正因子越低；反之，在該輸入音訊信號中該掩蔽頻率能量越低，則該校正因子越高；

-該校正因子受到與該輸入音訊信號內容無關之邊界的限制；

-每一個第一聲級臨界值被包含在0到-60dB之間；

-該等第一聲級臨界值係在接收該輸入音訊信號之前由一自動校準程序或一手動校準程序所產生的；

-該輸入音訊信號係一多聲道輸入音訊信號，每一個聲道與一音訊呈現裝置相關聯，並且至少步該驟b)至e)係針對每一個聲道被獨立執行的；

-該等頻帶分量被提取作為頻帶分格分量，每一個頻帶為一頻帶分格，並且其中步驟g)包含：

-g1)對於每一個頻帶，從在該頻帶之該提取頻帶分量的一聲級和與該頻帶相關聯之該第二聲級臨界值之間的一比較，決定一臨時增益，

-g2)在該等頻帶分量之頻率的聲級上檢測複數個局部最小值，

-g3)把位於兩個局部最小值之間的頻帶分組在一起，以形成複數個頻率分組，

-g4)對於每一個分組，決定與該分組之該等頻帶之該等初始增益中一最低初始增益相對應之該分組的一降低增益，

g5)把該分組的該降低增益影響到該分組之每一個頻帶；

-當由該音訊呈現裝置呈現該輸出音訊信號時，一麥克風在該設備附近錄音一麥克風信號，並且該麥克風信號被使用來根據與該設備附近環境聲音相對應之該麥克風信號的一部分來修改該指示符。

**【0013】** 根據另一方面，本發明提供了一種設備，該設備包含有至少一個音訊呈現裝置及一處理單元，該設備被組配來執行根據本發明之方法的該等步驟。

【0014】 根據另一方面，本發明提供了一種非暫時性的電腦可讀取媒體，其包含有存儲在其上的程式指令，當該等程式指令被一電腦讀取時，會致使該電腦執行根據本發明之方法的該等步驟。

### 【圖式簡單說明】

【0015】 在閱讀了以下針對本發明之較佳實施例的詳細描述之後，本發明的其他方面、目的以及優點將變得更加明顯，這些實施例以作為非限制性的實例被給出，並參考了該等附圖，其中：

-圖1根據一可能的實施例係一示意圖，其展示出該方法的步驟；

-圖2根據一可能的實施例係一示意圖，其展示出該方法之一步驟之一詳細視圖；

-圖3展示出當播放一振動頻率時，具有可聞振動之一第一播放音訊內容之一頻譜；

-圖4展示出當播放一振動頻率時，具有可聞振動之一第二播放音訊內容之一頻譜；

-圖5根據一可能的實施例展示出在施加該增益函數之前及之後該第一播放音訊內容之一頻譜；

-圖6根據一可能的實施例展示出在施加該增益函數之前及之後該第二播放音訊內容之一頻譜；

-圖7展示了一校正因子隨時間推移而演變的一實例，該校正因子係從對振動免疫之一音訊內容之一代表掩蔽頻率能量的指示符所決定的；

-圖8展示了一校正因子隨時間推移而演變的一實例，該校正因子係從被振動汙染之一音訊內容之一代表掩蔽頻率能量的指示符所決定的；

-圖9展示出一代表掩蔽頻率能量的指示符作為一測試音訊信號聲級之一函數的一演變實例，以及從其導出之該第一聲級臨界值；

-圖10展示出該由校準過程所產生之第一聲級臨界值的一實例；

-圖11根據本發明之一可能的實施例展示出在一多頻帶後處理中一音訊內容之一頻譜的一分組以及對應自適應臨界值的一實例。

## 【實施方式】

### 【0016】 較佳實施例之詳細說明

該方法係由一種包含有至少一個音訊呈現裝置及一處理單元的設備來執行的。該設備可以例如是一電腦、一膝上型電腦或平板電腦、一智慧型手機、由數個實體元件連接在一起所構成的一聲音系統、或任何其他類型的類似裝置。一音訊呈現裝置例如係一揚聲器，或可產生可聞聲音的一類似裝置。

【0017】 參考圖1，一輸入音訊信號首先被接收。該輸入音訊信號可以具有任何類型的音訊內容(音樂、歌曲、聲音、等等)，並且可具有各種來源(數位音訊諸如MP3、音訊CD、視訊遊戲、電影DVD、網際網路、等等)。該輸入音訊信號的該內容意圖由諸如一揚聲器之類的一音訊呈現裝置被呈現(即播放)給一使用者。該輸入音訊信號可以是一多聲道輸入音訊信號，該多聲道輸入音訊信號的每一個聲道被分配給諸如一揚聲器的一音訊呈現裝置。

【0018】 一輸入音訊信號係由隨時間變化的資料(通常是用於數位信號的二進制數位序列)所組成。該等資料變化定義了該輸入音訊信號的該音訊內容。該等資料變化根據可被分組為頻帶之各種頻率來發生。因此，可以從該輸入音訊信號中提取出複數個頻帶分量(步驟10)。每一個頻帶分量具有與一頻帶相關聯的一聲級。

【0019】 該方法可以在時域中進行。在這種情況下，該等頻帶分量被提取成為頻帶信號分量，並且一頻帶信號分量的該聲級係該頻帶信號分量的一振幅，其在此指定在每一個頻帶上的該短時間能量。可以藉由使用一濾波器組，即把該輸入音訊信號分割成多個頻帶信號分量之一帶通濾波器陣列，來提取該等頻帶

信號分量，每一個分量都承載該原始輸入音訊信號的一單子頻帶，其振幅對應於在該特定子頻帶中該輸入音訊信號之該聲級的一振幅。

**【0020】** 該方法可以在頻域中執行。在這種情況下，該等頻帶分量被提取作為頻帶分格分量，並且每一個頻帶係一頻率分格。為了提取該等頻帶分格分量，該方法包含在頻域中把該輸入音訊信號轉換成頻帶分格分量，並且一頻帶分量的該聲級係一頻帶分格分量的一大小。在頻域中工作允許在運算時間與準確性之間取得良好的折衷，因此構成了該較佳實施例。以一種非限制性的方式，以下的描述將針對在頻域中執行的一方法，但是該描述的任何教導都可被套用到在時域中執行的一方法。

**【0021】** 可以使用許多不同的數學轉換來從諸如該輸入音訊信號的一時域信號中提取出頻帶分格分量。快速傅立葉轉換(FFT)係最有效且常用的轉換之一。在一較佳實施例中，使用了一種稱為短時傅立葉轉換(STFT)的變型，其中該FFT在該輸入音訊信號的連續分段上被執行。該方法的該等後續步驟然後被執行在每一個分段上，其連續地被儲存在一緩衝器中。可選擇地但係優選地，可以執行若干個額外的程序以改善該方法的結果。

**【0022】** 完全分開地把該處理施加在該輸入音訊信號的兩個連續分段上可能會導致被發送到該音訊呈現裝置之該輸出音訊信號中的該等兩個分段之間出現急劇的變化。這種急劇的變化可能會導致可聞的偽音，諸如失真，從而可能會降低該音訊品質。為了減輕分段的這一個缺點，可以實施重疊相加的技術，其中該輸入音訊信號的每一個緩衝分段(索引 $n$ )都被處理兩次：第一次係與前一個緩衝分段(索引 $n-1$ )一起而第二次係與後一個緩衝分段(索引 $n+1$ )一起。為了簡化下面的描述，該緩衝索引將無差別地被使用來指該緩衝分段以及由該重疊相加技術所產生的該緩衝器。

**【0023】** 在提取之後所獲得之該頻帶分格分量的該等大小最好可被正規化

以去除該提取的任何偏差。實際上，存在於常用函式庫中之許多的快速傅立葉轉換函數可能會引入偏差，特別是因為該轉換被施加於一音訊信號的短分段上(例如1024個樣本)。

【0024】 該正規化還可被使用來使得該等頻帶分格分量的該等大小適應於該揚聲器的該音量設置。實際上，該方法主要旨在減少由於該揚聲器播放該聲音而引起的可聞振動。由一揚聲器所引起的該振動在很大程度上取決於由該揚聲器所播放該聲音的該響度。由於一播放聲音的該響度取決於音訊內容的程度與取決於該設備之該等設置的程度相當，因此考慮到該音量設置可以更準確地估計該振動問題的可能性。在諸如電腦的許多設備上，音量調整係在該音訊處理之後被施加。結果，該音量設置不會出現在該輸入音訊信號中。因此，在一正規化過程中使用該音量設置可預期該音量設置對該聲音品質的影響。

【0025】 該等頻帶分格分量之該大小的該正規化可以遵循以下的公式：

$$|\tilde{S}(n, k, c)| = |S(n, k, c)| * Coeff * Volume$$

其中 $|S(n, k, c)|$ 是由該緩衝索引 $n$ 、該頻率分格 $k$ 、以及該聲道 $c$ 所定義之該等頻帶分格分量的該大小。 $Volume$ 係與該音量設置相對應的一實數。 $Coeff$ 係用於補償該偏差的一正規化係數，它取決於被使用於該轉換之函式的該軟體函式庫(例如，英特爾(Intel)的Integrated Performance Primitive IPP函式庫)。該正規化係數 $Coeff$ 還取決於該分析窗口。該係數 $Coeff$ 可例如遵循以下的公式：

$$Coeff = \frac{\alpha}{W}$$

其中 $\alpha$ 為一實數，而 $W$ 為取決於該分析窗口大小的一實數。 $\alpha$ 及 $W$ 可藉由把一750 Hz及0 dB滿量程的一正弦波作為該輸入音訊信號並把與750 Hz相對應之該頻率分格的該大小值限制為1來決定。

【0026】 也可以使用用於改善該等頻帶分量之該提取的其他技術。例如，

在把該輸入音訊信號轉換進入頻域之前，可以使用零值填充。在該分析窗口中的該信號資料被以零值填充，以便人為地增加轉換到頻域中的該頻率分辨率。

### *代表掩蔽頻率能量的指示符*

**【0027】** 一旦從該輸入音訊信號中提取出該等頻帶分量，至少一個代表掩蔽頻率能量的指示符被決定。該指示符係根據該裝置頻率響應及該等複數個提取出的頻帶分量來決定的。該等掩蔽頻率對應於高於一頻率臨界值的該等頻率，因此能夠掩蔽發生高於該頻率臨界值的振動。該指示符僅藉由使用高於該頻率臨界值的頻帶來決定。這意味著低於該頻率臨界值之該等頻帶分量的該等聲級不被使用來決定該代表掩蔽頻率能量的指示符。該頻率臨界值把該頻率範圍劃分為兩個子範圍：低於該頻率臨界值的一第一頻率子範圍，其會導致振動；以及高於該頻率臨界值的第二頻率子範圍，其會掩蓋或隱藏由該第一頻率子範圍的頻率所引起的該等振動。因此，該頻率臨界值的該值取決於該等設備的組配(該揚聲器的環境、等等)及該揚聲器的屬性(該揚聲器的頻率範圍、等等)。在大多數的情況下，該頻率臨界值最好高於6 kHz、最好高於8 kHz、以及最好高於10 kHz。為了說明的目的，以一非限制性的方式，將參考對應於高於6 kHz之相當高頻率的掩蔽頻率來進行該描述。應當注意，不一定要使用高於該頻率臨界值之所有的頻帶來決定一代表掩蔽頻率能量的指示符。例如，可以從高於8 kHz之一頻率臨界值但低於14 kHz的頻帶來決定一指示符。優選地，可從頻帶範圍在至少4 kHz高於該頻率臨界值上並且更優選地頻帶範圍在至少6 kHz高於該頻率臨界值上(例如12-20 kHz或10-16 kHz)來決定一代表掩蔽頻率能量的指示符。

**【0028】** 由該揚聲器所產生的該等寄生振動會產生諧波在整個可聞範圍上(例如，10 Hz到20 kHz)，但特別會存在於高頻中，即超過6或8 kHz。雖然在大多數的常聽聲音(音樂、人的聲音、等等)中都存在有低頻諧波，但在常聽聲音中通

常不包含高頻諧波。因此，寄生高頻諧波比起寄生低頻諧波更會引人注意。即使此類寄生高頻諧波顯示出比其基頻有低得多的聲級(通常至少低60 dB)，但它們仍可被高度地聽到，因為它們在聲音學上沒有被基頻或該等第一諧波所掩蓋。此外，當該設備的一組件振動時(特別是當其與其他組件接觸時)，會出現諧波間雜訊。這樣的一種諧波間雜訊更加會引人注意，因為它會產生非自然的聲音。

【0029】 被注意到的是，一些音樂曲調比其他更容易會出現由振動所產生之可聞的寄生聲音。例如，用一設備播放古典音樂可能導致會聽到可聞的振動，而用同一設備播放流行音樂時振動並不明顯。

【0030】 圖3及圖4圖示出該音訊內容在振動產生上的影響。圖3圖示出當一振動頻率被播放時具有可聞振動之第一播放音訊內容錄音的一頻譜30，以及圖4圖示出當該相同振動頻率在相同聲級下被播放時，不具有可聞振動之第二播放音訊內容錄音的一頻譜40。該第一播放音訊內容係一段容易產生振動偽音的古典音樂，而第二播放音訊內容係一段流行音樂。這兩個音訊內容由同一台膝上型電腦的同一揚聲器播放，並由同一麥克風來錄音。圖3及4還展示出了在一880 Hz振動頻率下之一示例性純音的頻譜35，其由同一台電腦單獨播放及錄音，並且具有與在880 Hz的振動頻率下該等播放內容之相同的聲級。從該示例性聲音的該頻譜35中可以看出，在該振動頻率下的音訊內容會產生振動，從而導致在較高頻率下發生寄生振動。

【0031】 在圖3中，在該等高頻31(高於6 kHz)內，該示例性聲音的該頻譜35非常接近第一播放音訊內容的該頻譜30。這意味著該頻譜30之該等高頻31的該大小基本上係由在該第一播放音訊內容內之該振動頻率處的該聲級所引起之振動而引起的。對於那些高頻31，該第一播放音訊內容具有其自己之低的其他大小，因此該等振動並未被隱藏在高頻31內。相反地，在圖4上，該示例性聲音之該頻譜35該等大小保持較低於該第二播放音訊內容之該頻譜40的該等大小。即

使振動係由該第二播放音訊內容在880 Hz的該振動頻率處所引起的，但由於該第二播放音訊內容掩蓋了該等高頻31的該等振動，因此無法聽到該等振動。因此，該等高頻31構成了掩蔽頻率。

【0032】 有可能的是可透過施加一增益函數，藉由降低針對該振動頻率之該播放內容大小來降低該振動頻率的影響。在圖5及6的該等實例中，對於兩種音訊內容，把-12 dB的增益施加到該振動頻率(880 Hz)的該大小。圖5展示出如圖3中該第一播放音訊內容的該頻譜30，以及在已經施加該增益之後該第一播放音訊內容的該頻譜32。如所預期的，該頻譜32相對於該原始頻譜30在該等較高頻率31內顯示出該大小的一顯著降低。而且，一聽者不再能夠感知到由該等振動所引起的該寄生雜訊。圖6展示出如圖4中的該第二播放音訊內容的該頻譜40，在已經施加該增益之後該第二播放音訊內容的該頻譜42。與圖5相反，兩個頻譜40、42之間沒有顯著差異，特別是對於包括在6 kHz與16 kHz之間的該等高頻而言。這是因為在該等高頻處之該聲級並非主要源自振動，而主要源自該第二播放音訊內容。與先前一樣，一聽者無法感知到由該等振動引起的寄生雜訊。

【0033】 在這兩種情況下，在施加了該增益之後，都無法感知到來自振動的寄生雜訊。然而，施加在880 Hz的該增益已經修改了該音訊信號，該等兩個頻譜32、42在880 Hz處顯示出明顯的較小的大小33、43。結果，該第二播放音訊內容已被該增益被實質性地修改，即使沒有必要性：該寄生雜訊係不可被感知的，因為它被該高頻內容所掩蓋。

【0034】 確實，如圖3及4所示，由該振動所產生的該寄生雜訊係可聞的只有當引起該振動之該播放內容在一高頻範圍內具有足夠的能量來掩蓋該產生的寄生雜訊，從而掩蔽該振動時。因此，該發想係分析該等高頻頻帶分量的該能量，以決定是否一高頻寄生雜訊會被隱藏。如果該等高頻頻帶分量具有足夠的能量來掩蓋該高頻寄生雜訊，則因為該高頻寄生雜訊將不可感知，因此無需更改該音

訊信號。反之，如果該等高頻頻帶分量沒有足夠的能量來掩蓋該高頻寄生雜訊，則一增益衰減必須被施加以壓抑該高頻寄生雜訊，否則其將係可聞的。

**【0035】** 估計被包含在該等高頻中的該能量(即高於該頻率臨界值的能量)係評估該輸入音訊信號之隱藏能力的一個好方法。該高頻能量越高，該振動引起的雜訊將被掩蓋的可能性就越大。該音訊信號之該隱藏能力取決於兩個主要方面：

- 在將被播放之該輸入音訊信號的該等高頻中的能量，
- 在該等揚聲器之該裝置頻率響應中之該等高頻中的能量。

**【0036】** 該音訊呈現裝置的該裝置頻率響應係該音訊呈現裝置在響應於一激發時該輸出頻譜的該定量測量，並且被使用來表徵該音訊呈現裝置的該動態。它係該輸出相對於該輸入之一大小及相位的量度，並為頻率的一函數。估計一實體系統的該頻率響應通常涉及用一輸入信號激發該音訊呈現裝置、測量輸入及輸出時間歷史、並藉由諸如快速傅立葉轉換(FFT)的一程序把該等兩者進行比較。因此，可以使用諸如一掃描正弦波技術或帶有一濾波器組織一粉紅雜訊之類的傳統技術來預先決定該裝置的頻率響應。

**【0037】** 該代表掩蔽頻率能量的指示符最好係高於該頻率臨界值之頻帶分量之聲級總和的一函數。該代表掩蔽頻率能量的指示符也可以是該音訊信號輸入之該頻譜的一統計特性諸如該變異數，因此表示該頻譜之該平坦度，的一個函數。實際上，有很多方法可以建立一代表掩蔽頻率能量的指示符。下面的描述提出了根據頻帶分量之聲級總和來決定此一指示符的方法。

**【0038】** 優選地是，該指示符考慮到將由該等聽者聽到的聲音，因為它是將掩蓋該等音訊偽音之該等揚聲器的該音訊輸出。將由該等聽者聽到之該揚聲器之該音訊輸出的準確頻譜事先係未知的，因為它取決於該輸入音訊信號及該等揚聲器。對於每一個頻帶，該音訊輸出的頻譜係透過在一頻帶內該等揚聲器的

裝置頻率響應與該頻帶聲級之間的一乘積來估計的：

$$S_{out}(n, k, c) = |\tilde{S}(n, k, c)| * FR(k, c)$$

其中n為該緩衝索引、k為該頻帶索引、c為該聲道索引、FR(k, c)為針對該頻帶k及該聲道c之該設備的裝置頻率響應。該裝置頻率響應FR(k, c)更精確地係該聲道c發送至其之該揚聲器之頻率響應的聲級。

【0039】 然後，該高頻能量 $E_{HF}$ 係從針對該等高頻該音訊輸出 $S_{out}$ 之該估計值的平方和來得出的：

$$E_{HF}(n, c) = 10 * \log_{10} \left( \frac{1}{k_{max} - k_{min} + 1} * \sum_{k=k_{min}}^{k=k_{max}} (|\tilde{S}(n, k, c)| * FR(k, c))^2 \right)$$

其中n為該緩衝索引、k為該頻帶索引、c為該聲道索引、 $k_{min}$ 為被考慮用於決定該指示符之該等高頻帶之該較低索引以及 $k_{max}$ 為該等高頻帶之該較高索引。

【0040】 然後，該指示符 $Indic$ 對應到該聲道c中的該平均值，其中C代表該聲道總數：

$$Indic(n) = \frac{1}{C} * \sum_{d=1}^C E_{HF}(n, c)$$

【0041】 如果只有一個聲道c，則該指示符 $Indic$ 係該高頻能量估計 $E_{HF}$ 。

【0042】 對應於該等高頻頻帶之該較高索引 $k_{max}$ 的該頻率可以由人類可聞範圍之該較高極限來粗略地限定，並且例如可以在16 kHz至20 kHz之間做選擇。對應於該等高頻頻帶之該較高索引 $k_{max}$ 的該頻率也可以較低，以更好地限定將要考慮之該等振動的該頻率範圍。例如，如果800 Hz的一激發頻率產生低於12 kHz的振動，則該較高索引 $k_{max}$ 可被選擇來對應於在12 kHz處的該頻帶。

【0043】 該等高頻頻帶的該較低索引 $k_{min}$ 對應於定義被使用來決定該高頻能量指示符之該等頻率的該較低頻臨界值。如上所解釋的，該頻率臨界值優選地高於6 kHz、8 kHz、以及優選地高於10 kHz或更高。考慮以下幾點來選擇該頻率

臨界值的該值：

-該等振動導致出高階諧波，其頻率通常高於幾 kHz；

-通常，低於10 kHz，在電腦揚聲器或類似音訊呈現裝置的該裝置頻率響應中不會有重大的降低；

-大多數音訊內容具有低於10 kHz的能量，因此低於10 kHz之大部分的失真或振動聲音將被隱藏；

-如果該音訊內容在6 kHz至10 kHz之間沒有能量，則極不可能有10 kHz以上的能量。

【0044】藉由使用由在該設備附近錄音該聲音之一麥克風所產生的一麥克風信號，考慮環境聲音用來決定該代表掩蓋頻率能量的指示符係有可能的。該麥克風信號的內容當然源自於該音訊呈現裝置所呈現的該輸出音訊信號，而該輸出音訊信號源自於該輸入音訊信號。但是，該麥克風信號的部分內容也將與該輸出音訊信號無關，然後將與環境聲音相對應。可透過諸如使用一聲音迴音消除器的處理來隔離這個其他部分。諸如一中值濾波器之類的一濾波器也可被使用來改善該環境聲音的該隔離。如果該麥克風信號顯示該環境聲音包含足夠的高頻能量以掩蓋該振動引起的偽音，則可以修改該代表掩蓋頻率能量的指示符以反映該狀況。因此，如果該麥克風信號中與環境聲音相對應的部分具有足夠的高頻能量(例如，高於類似於被使用於該指示符的一頻率臨界值)，則可以把一新增項添加到該指示符以增加其值。相反的是，如果該麥克風信號中與環境聲音相對應的該部分具有一較低的高頻能量，則也可以降低該指示符。

### 校正因子

【0045】然後根據該指示符 $Indic$ 決定一校正因子(步驟16)。該校正因子旨在修改一校正臨界值，其決定必須如何校正每個一頻帶的該聲級。該校正因子可

被施加到一臨界值，方式係藉由乘以該臨界值，或者可被施加作為該臨界值的一指數。優選的是，對於被認為容易受到可聞振動聲音影響的輸入音訊信號把該校正因子被定義為達到1或更高的值，並且對於被認為不太可能受到可聞振動聲音影響的輸入音訊信號把校正因子定義為小於1。

【0046】該校正因子因此可對應於該指示符 $Indic$ ，但被加權並受邊界的限制。例如，該校正因子可被定義為該指示符的一線性函數：

$$Factor(n) = a \times Indic(n) + b$$

其中 $a$ 及 $b$ 為兩個實數。該校正因子應具有與該指示符之該趨勢相反的一趨勢：該指示符越高，則該校正因子越低，反之亦然。優選地，該校正因子會隨著該輸入音訊信號中該高頻能量的增加而減小，即隨著掩蔽頻率能量的增加而減小。因此，如果該指示符優選地隨著該高頻能量的增加而增加，則該數字 $a$ 因此最好為負值。因此，該校正因子係該代表掩蓋頻率能量的指示符之反向的一個函數，即該指示符的一反向加法。

【0047】該等數字 $a$ 及 $b$ 係固定的並事先被決定的，例如藉由至少兩個具有兩種不同音訊內容的輸入音訊信號：至少一個具有較低的高頻能量(如在圖3中所示的那一個)及至少另一個具有較高的高頻能量(如在圖4中所示的那一個)。 $a$ 及 $b$ 的值被搜尋以使得該校正因子在對應於具較低高頻能量之音訊內容的一有問題內容(即容易產生振動偽音)時高於1，並且該校正因子在對應於具較高高頻能量之音訊內容的一無問題內容時低於0.8。一線性回歸可被使用來決定 $a$ 及 $b$ 的值。此外，可進一步修改 $b$ 的值以保留該音訊內容的該音調，同時保持該等振動係小的。

【0048】通常， $a$ 在-0.1到-0.005之間， $b$ 在-200到100之間但更優選地為負數並且在-100到0之間。

【0049】該校正因子的該等值最好進一步會受到邊界限制，使之不管該指

示符值為何都有一可接受的校正。為此，把該等值限制在一下界與一上界之間的一範圍內：

$$Factor(n) = \min(\max(a \times Indic(n) + b, \text{下界}), \text{上界})$$

【0050】 為了舉一個例子，假設當該校正因子被線性地施加時，該校正因子可以在0.1與3之間。因此在這種情況下，該下限最好高於或等於0.1及該上界係低於或等於3。而且，該下界優選係低於或等於0.5及該上界係高於或等於1。但如果在該等邊界的套用之前該校正因子的計算係用分貝來算，該等邊界的該等值當然會不一樣。

【0051】 該校正因子係針對該輸入音訊信號的每一個分段來計算，這意味著該校正因子會隨著該輸入音訊內容的該等頻譜特性的變化而永久地變化。

【0052】 圖7展示了一校正因子隨時間推移而演變的一實例，該校正因子係從對振動偽音免疫之一音訊內容之一代表掩蔽頻率能量的指示符所決定的。在這裡該音訊內容係摘錄自在2009年發布之一首電子流行音樂主打歌。如在圖4之該實例中那樣，該音樂片段在該等高頻31中具有高的聲級。因此，該指示符係高的，並因此該校正因子係低的。該校正因子幾乎始終都在1以下，並且大多時在0.5到0.8之間振盪。

【0053】 圖8展示了一校正因子隨時間推移而演變的一實例，該校正因子係從易受振動汙染之一音訊內容之一代表掩蔽頻率能量的指示符所決定的。在這裡該音訊內容係一段可聽見偽音之一段古典音樂的摘錄。如在圖5之該實例中那樣，該音樂片段在該等高頻31中具有低的聲級。因此，該指示符係低的，並因此該校正因子係高的。該校正因子幾乎始終都高於1，並且大多時在1到1.2之間振盪。

### 第一聲級臨界值

【0054】該校正因子被使用來校正第一聲級臨界值，其獨立於該輸入音訊信號的內容。複數個頻帶中每一個頻帶之一組預定的第一聲級臨界值係可用的。該等第一聲級臨界值與硬體有關，即它們取決於該設備，並且獨立於任何的音訊內容。如以上所解釋的，當特定的振動頻率被播放時，無論該音訊內容如何，通常都會出現諸如振動、失真或嗡嗡聲等之類的偽音。偽音可能以與該振動頻率非常不同的頻率來出現。

【0055】如在圖3及4中所示，以一振動頻率(此處為880 Hz)播放一純音會產生諸如失真的偽音。除了對應於在880 Hz處之該基頻的該峰值36之外，該頻譜還顯示出諧波37，並且其他偽音也特別地出現在該等高頻31上，這正是由該設備所播放之該振動頻率處之該聲音所引起的該等振動及其他失真所引起的。該等偽音的呈現也取決於一給定裝置在該振動頻率處正被播放之該聲音的響度。該響度越高，該等偽音越強。

【0056】每一個頻帶的第一聲級臨界值指出用於該頻帶的最高聲級，該聲級可以由該設備播放而沒有像是振動的明顯偽音。與一頻帶相關聯之預定的第一聲級臨界值對應於在與該頻帶相對應之一激發頻率處一頻帶分量的一聲級，當以該頻帶分量聲級在該激發頻率處被激發時，不會產生由該音訊呈現裝置所引起之聲音的一估計振動聲級。

【0057】可藉一由自動校準程序或藉由一手動校準程序(手動調諧)來決定該等第一聲級臨界值。對於複數個頻率中的每一個頻率，一純音調(正弦波)係由該音訊呈現裝置來播放。對於手動調諧，一操作員傾聽該呈現的音訊，並藉由改變該聲級來決定相應的第一聲級臨界值，在該臨界值以下聽不到該等振動及失真。該操作員還可以傾聽一些有問題的內容，並可以使用一頻譜分析儀為每一次聽到的振動識別出有問題的頻率。藉由例如使用一等化器來改變該有問題的頻率的該聲級，可以決定該第一聲級臨界值，在該第一聲級臨界值以下聽不到該等

振動及失真。

**【0058】** 在該自動校準程序中，該激發音訊信號在該設備的該揚聲器上被播放，並且所產生的聲音由一麥克風錄音。該激發音訊信號包含旨在揭露該設備之該頻率相關行為的刺激。通常，該激發音訊信號包含一系列激發，該激發由在複數個激發頻率處並具有變化的聲級的頻帶分量所組成的。更具體地說，該等刺激是在許多激發頻率(即音調)處孤立的純正弦音。該系列的刺激可以識別出針對每一個激發頻率下該裝置的振動。因此，該激發音訊信號的目的是識別出該等頻率與該等聲級，高於其振動會出現。該等刺激可以以各種方式被佈置在該激發音訊信號內。例如，可以選擇一掃描音調，即在頻率中做線性變化的一音調，例如在10 Hz至20 kHz之間。

**【0059】** 也可以使用一系列具有變化聲級之固定頻率的音調。該激發音訊信號可以包含具有一隨時間變化的聲級之複數個時間隔離的頻帶分量。

**【0060】** 必須對將被測試之該等激發頻率做出選擇。例如，可以選擇與普通音符相對應的頻率。該測試範圍通常可以在50 Hz至20 kHz之間擴展。

**【0061】** 當播放該激發音訊信號時由該音訊呈現裝置所呈現的該聲音被錄音，以及針對每一個激發頻率，以該等刺激的該聲級做變化，一代表掩蔽頻率能量的指示符從其被構建。可以如上所述地構建該指示符。在圖9中，曲線50展示出此一指示符作為該激發音訊信號聲級之一函數的一變化實例。該指示符的值已被平滑化以便於該展示，其中平滑常數為幾秒鐘。該指示符會隨著該激發音訊信號聲級的增加而增加，並且該等振動需要一特定的聲級才會出現。

**【0062】** 然後把該指示符與一校準臨界值進行比較，該校準臨界值對應於當沒有振動發生時該指示符的一平均值(具有一範圍)。該校準臨界值的該值由在圖9中的直線51來表示。在該所示實例中，該校準臨界值為-85 dB(當然，該值取決於該指示符被計算的方式)。在該激發音訊信號的聲級較低時，即低於-20 dB的

情況下，該指示符的值低於-85 dB的校準臨界值，其意味著不會發生明顯的振動。在該激發音訊信號的該聲級較高時，即高於-20 dB的情況下，該指示符的值高於-85 dB的校準臨界值，其意味著開始發生明顯的振動。該指示符值的曲線50與該校準臨界值的直線51相交之點52給出了針對該激發頻率的該第一聲級臨界值，在該所示實例中為-20 dB。藉由對每一個激發頻率進行相同的處理，可以把一初步第一臨界值與每一個激發頻率相關聯。

**【0063】** 圖10展示了一組第一聲級臨界值的實例，該組第一聲級臨界值係針對與該等頻率分格之該等頻率相對應之複數個激發頻率來決定的，並且係一由校準程序所產生的。該曲線55在700 Hz以下及在1700 Hz以上主要係平坦的，其中該等第一聲級臨界值為0 dB，並且包含在700 Hz至1700 Hz之間的一負峰56，其中該等第一聲級臨界值顯著地低於零，甚至低於-17 dB。這表明該被測裝置對大約在700 Hz至1700 Hz之間的頻率敏感，因此在那些頻率內的聲級內容可能會引起振動。相反地，對於低於700 Hz及高於1700 Hz的頻率，該受測裝置將不會聽見振動。

### *第二聲級臨界值*

**【0064】** 該校正因子被使用來校正該等第一聲級臨界值，以獲得第二聲級臨界值。如同該等第一聲級臨界值，該第二聲級臨界值定義了可以在沒有振動或偽音的情況下播放之每一個頻帶的一最高聲級，但這一次還考慮了該輸入音訊信號的該內容。該等第一聲級臨界值獨立於該輸入音訊信號的內容，但該校正係數係取決於該輸入音訊信號。源自該校正的該等第二聲級臨界值透過該校正因子也依賴於該輸入音訊信號，並透過該等第一聲級臨界值係也與硬體相關的。

**【0065】** 對每一個頻帶決定一第二聲級臨界值(步驟20)，方式係用該校正因子修改(步驟17)與該頻帶相關聯之該第一聲級臨界值。該校正因子被施加到該

等第一聲級臨界值，例如藉由把該等第一聲級臨界值乘以該校正因子(以一種線性的方式)：

$$\text{第二臨界值}(n, k, c) = \text{Factor}(n) \times \text{第一臨界值}(k, c)$$

其中 $n$ 為該緩衝索引、 $k$ 為該頻代索引、 $c$ 為該聲道索引。

該校正因子也可被施加作為該第一聲級臨界值的一指數：

$$\text{第二臨界值}(n, k, c) = \text{第一臨界值}(k, c)^{\text{Factor}(n)}$$

**【0066】** 優選地，對於被使用來決定該代表掩蔽頻率能量的指示符之該範圍內的該等頻率(例如，掩蔽頻率高於該頻率臨界值，通常高於6 kHz、8 kHz、或10 kHz)，該校正因子被設置為1，並且該等第二聲級臨界值等於該等第一聲級臨界值。實際上，最好不要因在被使用來定義該校正因子之該等頻率分量上施加一校正因子而引入一偏差。

**【0067】** 該等第二聲級臨界值對應於源自該等振動之偽音係不可聞的該最高聲級，因為不是：

- 該振動所引發的聲音被該音訊內容所隱藏，就是
- 沒有振動發生。

**【0068】** 因為該第一聲級臨界值可被使用於該等掩蔽頻率，所以可以僅針對低於該頻率臨界值(例如8 kHz或10 kHz)的頻率來決定該等第二聲級臨界值。

### 比較方式

**【0069】** 在一比較步驟中，把每一個頻帶分量的該聲級與該第二聲級臨界值進行比較(步驟30)，並由此決定一增益。優選地是，針對每一個聲道 $c$ 獨立地進行該比較。一頻帶分量的該聲級可以是該音訊輸出 $S_{out}$ 之一估計的聲級，該估計值考慮了該設備的裝置頻率響應、或該正規化後的聲級 $\tilde{S}(n, k, c)$ 、或該頻帶分量的原始聲級。

【0070】可以把該增益決定為在該第二聲級臨界值與該頻帶分量之一聲級間的一比值，例如：

$$\text{增益}(n, k, c) = \frac{\text{第二臨界值}(n, k, c)}{|\tilde{S}(n, k, c)|}$$

其中n為該緩衝索引、f為該頻率、c為該聲道索引。該決定的增益可以重新群組化為一增益向量。

### 校正

【0071】一旦已針對每一個頻帶分量決定了該增益，一校正被執行，其中把該增益施加到該等頻帶分量(步驟40)。有幾種可能的方法可以施加該校正。一增益向量可直接被施加到每一個頻帶分量的該聲級，通常透過頻率掩蔽。也可以施加一峰值濾波器(在時間或頻率上)。

【0072】為了改善該校正的該等結果，尤其要降低處理偽音的該等影響，最好可後續處理將被施加的該等降低增益。對於每一個聲道，確定該輸入音訊信號之該等頻帶分量的該局部最小值。在兩個連續局部最小值之間的頻帶(例如，頻率分格)被分組成為頻率分組 $B_i$ 。可以根據每一個頻帶的該中心頻率對該等頻帶進行分組。在圖11上，展示了前(即較低頻率的)六個分組 $B_1$ 、 $B_2$ 、 $B_3$ 、 $B_4$ 、 $B_5$ 、 $B_6$ ，但較高頻率的其他分組 $B_i$ 存在，但因彼此之間過於靠近以致於無法被清晰地顯示。在該輸入音訊信號的之該頻譜60的局部最小值之間定義了六個分組 $B_1$ 、 $B_2$ 、 $B_3$ 、 $B_4$ 、 $B_5$ 、 $B_6$ 。例如，五個頻率分格被分組到該第五分組 $B_5$ 中。它們藉由該頻譜60上的五個小圓圈來表示。圖11還展示出用於每一個頻帶的該等第二聲級臨界值 62。圖11還展示出了用於每一個頻帶之該等降低增益 63的該等值(虛線)，其係由在該第二聲級臨界值與用於每一個頻帶之該頻帶分量的一聲級之間的差所引起的，即，該頻譜60的該等值。

【0073】 然後，對於每一個分組 $B_i$ ，在與該分組 $B_i$ 之該等頻帶相關聯之該等降低增益中決定出該最低之對應降低增益。然後，該最低的降低增益被相關聯到該分組 $B_i$ 之所有的頻帶，

$$\text{增益}(n, B_i, c) = \min_{j \in B_i} (\text{增益}(n, j, c))$$

其中 $n$ 為該緩衝索引、 $B_i$ 為該頻帶分組、以及 $c$ 為該聲道。

【0074】 在圖11的該實例中，該第五組 $B_5$ 的該等五個頻率分格分別與降低增益0、-10、-8、0及0相關聯。實際上，僅兩個頻率分格的振幅高於該第二聲級臨界值，因此與非零的降低增益相關聯。結果，該第五組 $B_5$ 的該等五個頻率分格現在具有-10 dB的一共同降低增益。

【0075】 可選擇地，可藉由使用可能具有不對稱行為的遞迴、中值或均值濾波器在每一個頻帶上進行濾波來平滑該等降低增益。例如，可以在施加之前對該頻率掩蔽進行平滑處理，以避免增益泵激及/或失真效應。一種一階低通濾波器可被使用，其具有一非常小的侵蝕或無侵蝕以及有一釋放時間例如為100 ms至300 ms。該濾波涉及用於當前緩衝的降低增益以及前一個緩衝中的降低增益。例如，該低通濾波器可以是：

$$\text{增益}(n, k, c) = (1 - \alpha) * \text{增益}(n - 1, k, c) + \alpha * \text{增益}_{\text{理論的}}(n, k, c)$$

其中 $n$ 為該緩衝索引、 $k$ 為該頻分格索引、 $c$ 為該聲道索引、 $\alpha$ 係定義該釋放的一調整參數、以及 $\text{增益}_{\text{理論的}}$ 係無需平滑即可獲得的該降低增益。

【0076】 一旦該等降低增益已被施加到該等頻帶分量的該聲級，該等不同經校正的頻帶分量可被轉換(步驟50)為一輸出音訊信號，例如如果FFT被使用來從該輸入音訊信號中提取出該等複數個頻帶分量，則可藉由使用一逆FFT把其轉換到時域中。然後，由該音訊呈現裝置呈現該輸出音訊信號(步驟60)：聲音由該設備來播放。

【0077】 如果該接收到的輸入音訊信號係一多聲道輸入音訊信號，則該多

聲道輸入音訊信號的每一個聲道被分配給諸如一揚聲器之個別的音訊呈現裝置，各種方法可以被使用。整個方法可為每一個聲道被獨立地複製，每一個聲道都有其自己的音訊輸入信號、裝置頻率響應以及第一聲級臨界值。然後，與一聲道相關聯之該所得的輸出音訊信號由相關聯到該聲道的該音訊呈現裝置來播放。

**【0078】** 也可能在聲道之間僅部分地複製一多聲道輸入音訊信號的該處理。這對於實現空間效果諸如3D空間化效果、經耳或雙耳特別地有用。例如，可以對幾個聲道，通常是右聲道及左聲道，施加相同的校正。為了這麼做，選擇在該等聲道之間該最低的降低增益。也可以首先針對每一個聲道決定初步降低增益，然後，透過在該聲道之該初步降低增益與各個聲道之初步降低增益中一最小的初步降低增益之間的一加權平均值來決定用於一聲道之每一個提取出頻帶的降低增益。該加權可例如由在0至1之間的一控制參數來控制，以調整該等不同的聲道可被組合到什麼程度。

**【0079】** 可以針對相同的輸入音訊信號決定複數個代表掩蓋頻率能量的指示符。可以使用不同的頻率臨界值，通常對於每一個指示符使用一不同的頻率臨界值。更具體地說，數個子範圍可以在該輸入音訊信號的該頻譜中被定義，並且一不同之代表掩蔽頻率能量的指示符可從不同的頻率子範圍的頻帶被決定出。該輸入音訊信號的該頻率範圍可被劃分為子範圍，方式係根據定義一第一子範圍及一第二子範圍之不同的劃分模式，該第二子範圍由作為一下界的一頻率臨界值來定義。該第一子範圍決定與一代表掩蔽頻率能量的指示符相關聯的該等頻帶；而該第二子範圍決定該代表掩蔽頻率能量的指示符係會從其被決定出的該等頻帶。該第一子範圍包含具有比第二子範圍較低頻率的頻帶。當一激發頻率引起在多個頻率子範圍處的振動或兩個激發頻率引起例如在兩個不同頻率子範圍處的振動時，這會特別地有用。該頻率範圍的該劃分不需要為全部，某些頻帶

可能不是第一子範圍也不是第二子範圍的一部分。同樣，一子範圍不一定是連續的。例如，一第一子範圍可延伸從200 Hz至600 Hz以及從1000至4000 Hz。

**【0080】** 作為一實例，一第一劃分模式可以包括基於在8 kHz處的一第一頻率臨界值把0-20 kHz的頻率範圍劃分為一第一子範圍0-8 kHz以及一第二子範圍8 kHz -20 kHz。一第一代表掩蔽頻率能量的指示符係從該第二子範圍8 kHz -20 kHz，即高於該第一該頻率臨界值8 kHz內之頻帶來決定的，並且與在該第一子範圍0-8 kHz內的該等頻帶相關聯。一第二種劃分模式可包括把0-20 kHz的該頻率範圍劃分為8 kHz-12 kHz的一第一子範圍以及12 kHz-20 kHz的一第二子範圍。該選擇的第一子範圍係該可能第一子範圍(0-12 kHz)的一部分。一第二代表掩蔽頻率能量的指示符係從在該第二子範圍12 kHz -20 kHz，即高於一第二頻率臨界值12 kHz內之頻帶來決定的，並且與在該第一子範圍8-12 kHz內的該等頻帶相關聯。

**【0081】** 該等劃分模式被預先決定，例如基於在一個校準過程中該等揚聲器之一先前測量的行為。與一頻帶相關聯的該校正因子係從與該頻帶相關聯之該等複數個代表掩蔽頻率能量的指示符中的至少一個來決定的。

**【0082】** 某些指示符可能不被使用。優選地，一校正因子係從每一個代表掩蔽頻率能量的指示符來決定的，從而產生複數個與該等頻率子範圍的頻帶相關聯的校正因子。

**【0083】** 在上面的該實例中，在該等兩個第一子範圍之間並沒有重疊，因此每一個頻帶僅與一代表掩蔽頻率能量的指示符相關聯，即使它與用於在該子範圍0-8kHz內之頻帶及在該子範圍8 kHz -12 kHz之頻帶的指示符不同。其結果是，當一校正因子係從每一個代表掩蔽頻率能量的指示符來決定時，一第一校正因子與在該子範圍0-8 kHz內之頻帶相關聯以及一第二校正因子與在該子範圍8 kHz-12 kHz內之頻帶相關聯。

【0084】與一頻帶相關聯的該校正因子被使用來決定用於該頻帶的一降低增益。如果一激發頻率帶分量透過一校正因子的使用被降低並且如果該頻帶係在另一個指示符的該第二頻率子範圍中(該第二頻率子範圍決定該其他指示符會從其被決定出的該等頻帶)，則該其他指示符必須在該激發頻帶之潛在降低增益施加之後被計算。使用前面的實例，其中該第一指示符係從8 kHz到20 kHz之間的頻帶被決定的而該第二指示符係從12 kHz到20 kHz之間的頻帶被決定的，並且被使用於在8 kHz到12 kHz之間的頻帶。該第二降低增益首先被計算並被施加到在8 kHz到12 kHz之間所有的頻帶，在基於8 kHz至20 kHz間的頻帶來計算該第一指示符之前。

【0085】該劃分模式還可以導致重疊的子範圍，從而導致一頻帶與幾個指示符相關聯。例如，在上面的實例中除了該等第一及第二劃分模式之外，一第三劃分模式可包含把該頻率範圍劃分為一第一子範圍0-9 kHz及一第二子範圍10 kHz-20 kHz，基於一第三頻率臨界值10 kHz。一第三代表掩蔽頻率能量的指示符可由在該第二子範圍10 kHz-20 kHz內的頻帶來決定，即高於該第三頻率為臨界值10 kHz，並與在該第一子範圍0-9 kHz內的頻帶相關聯。結果，在0-8 kHz之間的頻帶與該第一指示符及該第三指示符相關聯、在8 kHz與9 kHz之間的頻帶與該第二指示符及該第三指示符相關聯、以及在9 kHz與12 kHz之間的頻帶與該第二個指示符相關聯。藉由選擇用於決定一校正因子之該等指示符中之一個，可以只為與數個指示符相關聯之一頻帶決定一個校正因子。例如，可以藉由以下的公式決定索引為k之一頻帶的一校正因子：

$$Factor(n, k) = a * \max(Indic_1(n, k), Indic_2(n, k)) + b$$

【0086】其中 $Indic_1(n, k)$ 及 $Indic_2(n, k)$ 係與索引k之該頻帶相關聯之兩個代表掩蔽頻率能量的指示符。

【0087】還可以為每一個代表掩蔽頻率能量的指示符決定一校正因子，然

後選擇其中之一以計算該第二聲級臨界值。按照上述的實例，一第三校正因子被決定並且與在該子範圍0 - 9 kHz內的頻帶相關聯。其結果是，在0-8 kHz之間的頻帶與該第一校正因子及該第三校正因子相關聯、在8k Hz至9 kHz之間的頻帶與該第二校正因子及該第三校正因子相關聯、以及在9k Hz至12 kHz之間的頻帶與該第二校正因子相關聯。應被注意的是，對於12 kHz以上的頻帶，該校正因子被認為等於1，因此不需要被計算。該第一聲級臨界值將被使用來決定該降低增益。

**【0088】** 只有一個第二聲級臨界值將被使用來決定被施加到一頻帶的該降低增益。例如，可以基於該等指示符之該等各別值來預先決定將使用哪一個代表掩蔽頻率能量的指示符。否則，一校正因子從該等複數個與該頻帶相關聯的校正因子中被選擇出。優選的是，在該等預定複數個校正因子中選擇一校正因子使得在該等第二聲級臨界值中產出一最低的第二聲級臨界值，其中該等第二聲級臨界值係藉由使用該等複數個校正因子來決定的。在上面的實例中，對於低於8 kHz之一給定的頻帶，如果該第一校正因子導致一第二聲級臨界值為-10 dB，而該第三校正因子導致一第二聲級臨界值為-20 dB，則該第三校正因子將成為該選擇的校正因子，並且該第二聲級臨界值將會是-20 dB。通常，該選擇被執行在該等校正因子的值上。

**【0089】** 儘管已經針對某些較佳的實施例描述了本發明，但是顯然本發明絕不侷限於此，並且本發明包含所描述之構件及其組合之所有技術的等效物。特別地是，對於本領域的習知技藝者將顯而易見的是，在不脫離所附請求項所限定之本發明範圍的情況下，可以進行各種的改變及修改。

### **【符號說明】**

#### **【0090】**

S10~S60、S15~S17、S20:方塊

30、32、35、40、42、60: 頻譜

31:高頻

33、43:大小

36:峰值

37:諧波

50、55:曲線

51:直線

52:點

56:負峰

62:第二聲級臨界值

63:降低增益

## 【發明申請專利範圍】

【請求項1】 一種音訊呈現之方法，該呈現藉由包含至少一個音訊呈現裝置之設備進行，該方法包含以下步驟：

- a) 接收一輸入音訊信號，
- b) 從該輸入音訊信號中提取出複數個頻帶分量，每一個頻帶分量具有與一頻帶相關聯的一聲級，
- c) 從該等複數個提取出的頻帶分量中決定代表掩蔽頻率能量的至少一個指示符，掩蔽頻率對應於高於一頻率臨界值之頻帶，代表掩蔽頻率能量的該指示符係僅從高於該頻率臨界值的頻帶決定，
- d) 從代表掩蔽頻率能量的至少一個指示符中決定至少一個校正因子，
- e) 對於每一個頻帶，藉由使用該至少一個校正因子修改與該頻帶相關聯之一預定第一聲級臨界值來決定一第二聲級臨界值，其中該第一聲級臨界值係獨立於該輸入音訊信號的內容，該第一聲級臨界值指出針對該頻帶的一最高聲級，其可以在沒有明顯偽音的情況下由該設備播放，以及其中該第二聲級臨界值取決於該輸入音訊信號，
- f) 對於每一個頻帶，從在該頻帶之該提取出的頻帶分量的一聲級和與該頻帶相關聯之該第二聲級臨界值之間的一比較決定一降低增益，以及
- g) 把該降低增益施加到該等頻帶分量的該等聲級，
- h) 把不同的該等頻帶分量轉換為一輸出音訊信號，
- i) 藉由該音訊呈現裝置對該輸出音訊信號進行音訊呈現。

【請求項2】 如請求項1之方法，其中該音訊呈現裝置具有一預定的裝置頻率響應，以及代表掩蔽頻率能量的該指示符也是從該裝置頻率響應決定。

【請求項3】 如請求項1之方法，其中該頻率臨界值高於或等於6 kHz。

【請求項4】 如請求項1之方法，其中代表掩蔽頻率能量的複數個指示符被

決定，代表掩蔽頻率能量的該等指示符係從對應於由不同頻率臨界值所限定之不同頻率子範圍的頻帶決定。

**【請求項5】** 如請求項4之方法，其中數個校正因子係從代表掩蔽頻率能量的該等複數個指示符決定，並且使用不同的校正因子來決定不同頻帶的第二聲級臨界值，該校正因子係取決於該頻帶而使用。

**【請求項6】** 如請求項3之方法，其中至少一個頻帶被相關聯到代表掩蔽頻率能量的若干個指示符，並且被使用來決定與該頻帶相關聯之該第二聲級臨界值的該校正因子係從與該頻帶相關聯之代表掩蔽頻率能量的該等若干個指示符中之至少一者決定。

**【請求項7】** 如請求項1之方法，其中該等頻帶分量被提取作為數個頻帶信號分量，並且一頻帶信號分量的聲級係該頻帶信號分量的一振幅，或者

其中該等頻帶分量被提取作為頻帶分格分量，並且每一個頻帶係一頻率分格，並且其中步驟b)包含把該輸入音訊信號在一頻域中轉換成數個頻帶分格分量，其中一頻帶分格分量的聲級係一頻帶分格分量的一大小，並且其中步驟h)包含把頻帶分格分量轉換成一時域。

**【請求項8】** 如請求項1之方法，其中在步驟c)中所決定之代表掩蔽頻率能量的該指示符係高於該頻率臨界值之頻帶分量之一聲級總和的一函數，或是諸如變異數之頻譜之統計特性的一函數。

**【請求項9】** 如請求項1之方法，其中在該輸入音訊信號中之該掩蔽頻率能量越高，則該校正因子越低；反之，在該輸入音訊信號中之該掩蔽頻率能量越低，則該校正因子越高。

**【請求項10】** 如請求項1之方法，其中每一個第一聲級臨界值被包含在0到-60 dB之間，並且係在接收該輸入音訊信號之前由一自動校準過程或一手動校準過程產生。

【請求項11】如請求項1之方法，其中該輸入音訊信號係一多聲道輸入音訊信號，每一個聲道與一音訊呈現裝置相關聯，並且至少該等步驟b)至e)係針對每一個聲道獨立執行。

【請求項12】如請求項1之方法，其中該等頻帶分量被提取作為頻帶分格分量，且每一個頻帶為一頻率分格，並且其中步驟g)包含：

-g1)對於每一個頻帶，從在該頻帶之該提取出的頻帶分量的一聲級和與該頻帶相關聯之該第二聲級臨界值之間的一比較決定一初始增益，

-g2)在該等頻帶分量中之頻率的聲級上檢測複數個局部最小值，

-g3)把位於兩個局部最小值之間的頻帶分組在一起以形成複數個頻率分組，

-g4)對於每一個分組，決定該分組的一降低增益，該降低增益對應於該分組之該等頻帶之該等初始增益當中一最低的初始增益，

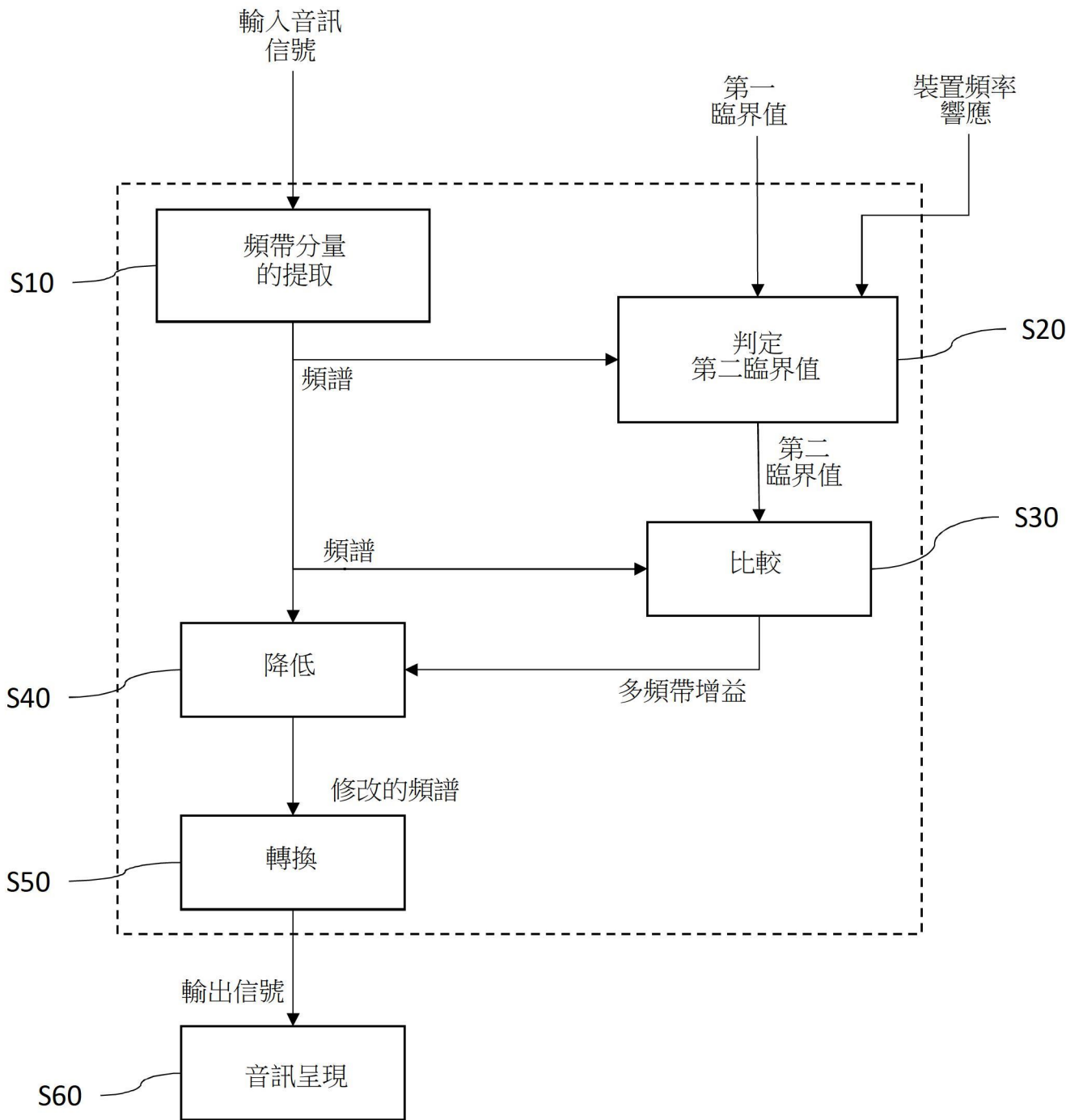
g5)把該分組之該降低增益影響到該分組之每一個提取出的頻帶。

【請求項13】如請求項1之方法，其中當該輸出音訊信號由該音訊呈現裝置呈現時，一麥克風在該設備附近對一麥克風信號錄音，並且該麥克風信號被使用來修改該指示符做為對應於該設備附近之環境聲音之該麥克風信號之一部分的一函數。

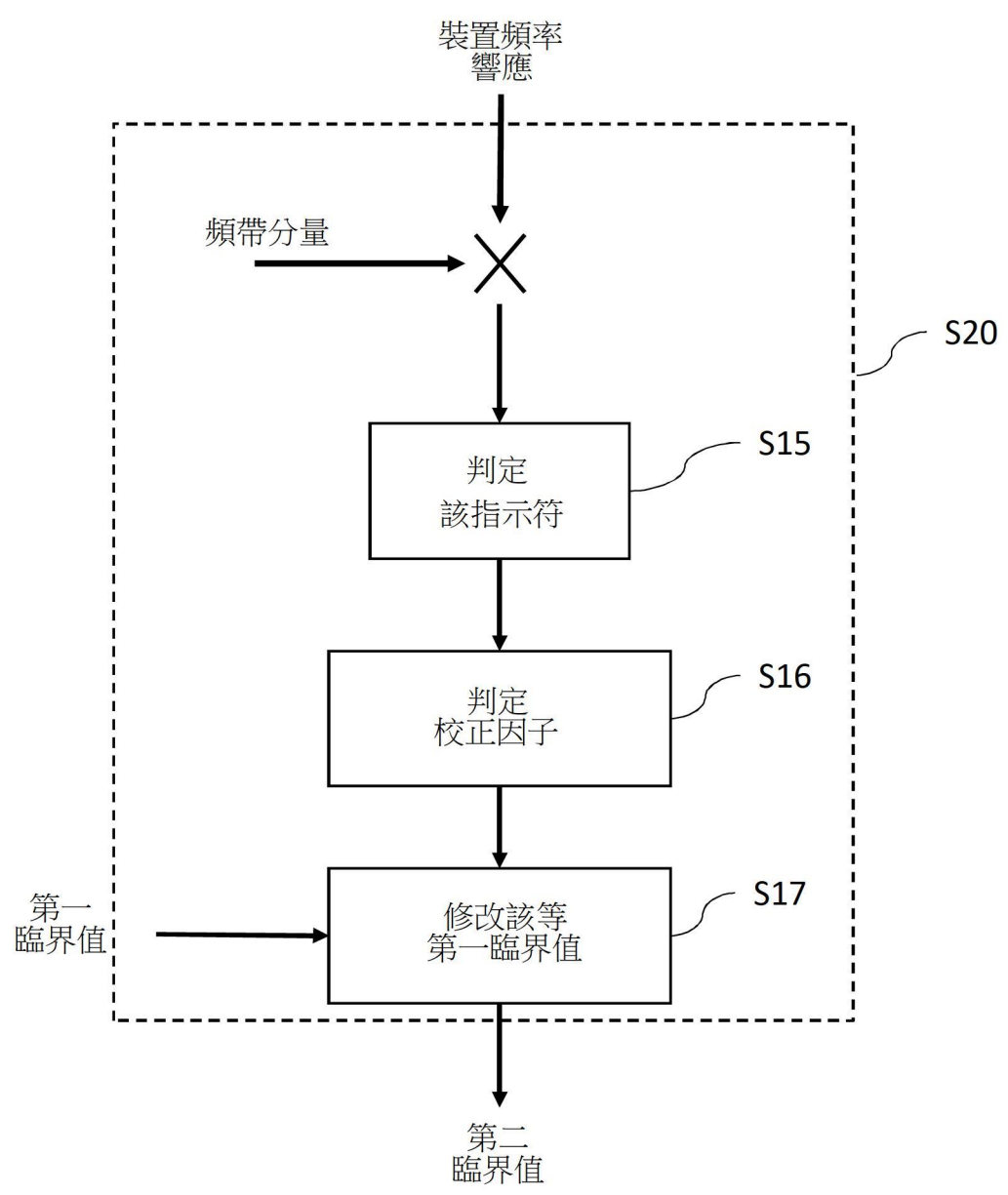
【請求項14】一種電子設備，其包含有至少一個音訊呈現裝置及一處理單元，該設備被組配來執行如請求項1至13中任一項之方法的該等步驟。

【請求項15】一種非暫時性電腦可讀媒體，其包含有存儲在其上的數個程式指令，當該等程式指令由一電腦讀取時，致使該電腦執行如請求項1至13中任一項之方法的該等步驟。

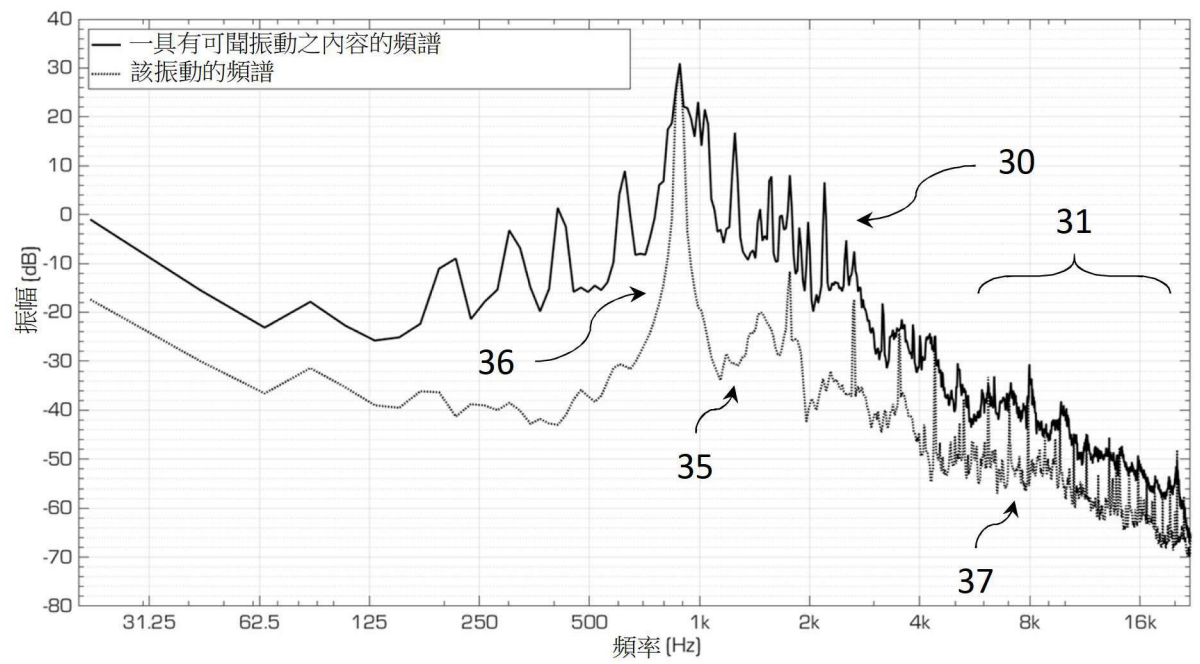
【發明圖式】



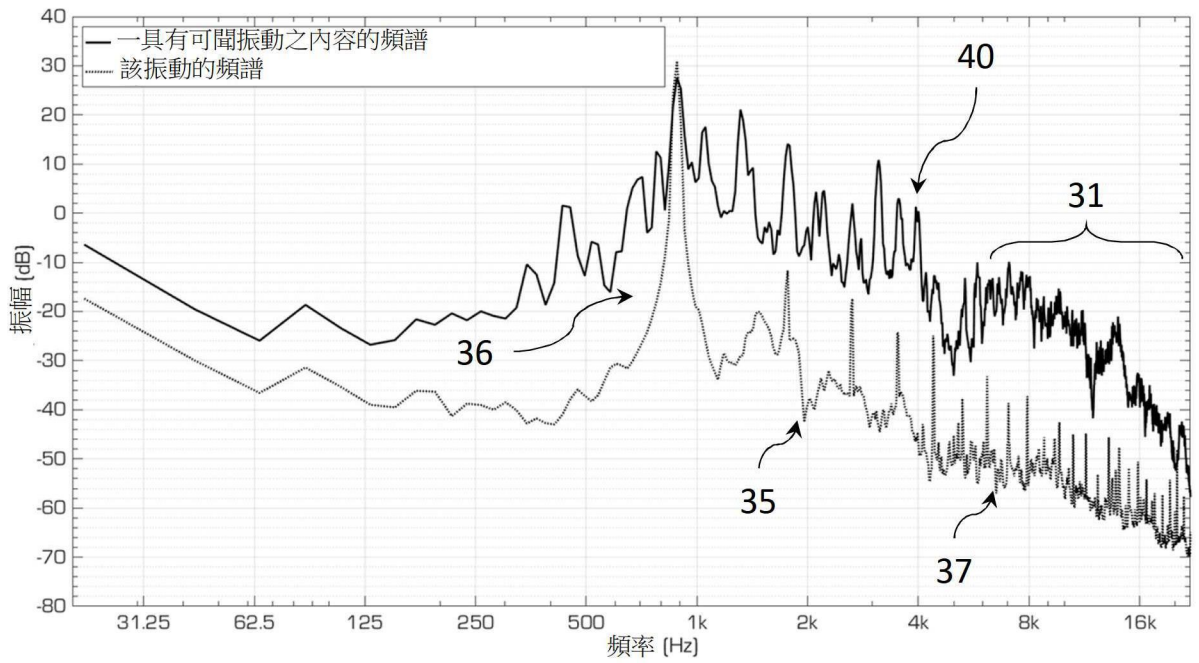
【圖1】



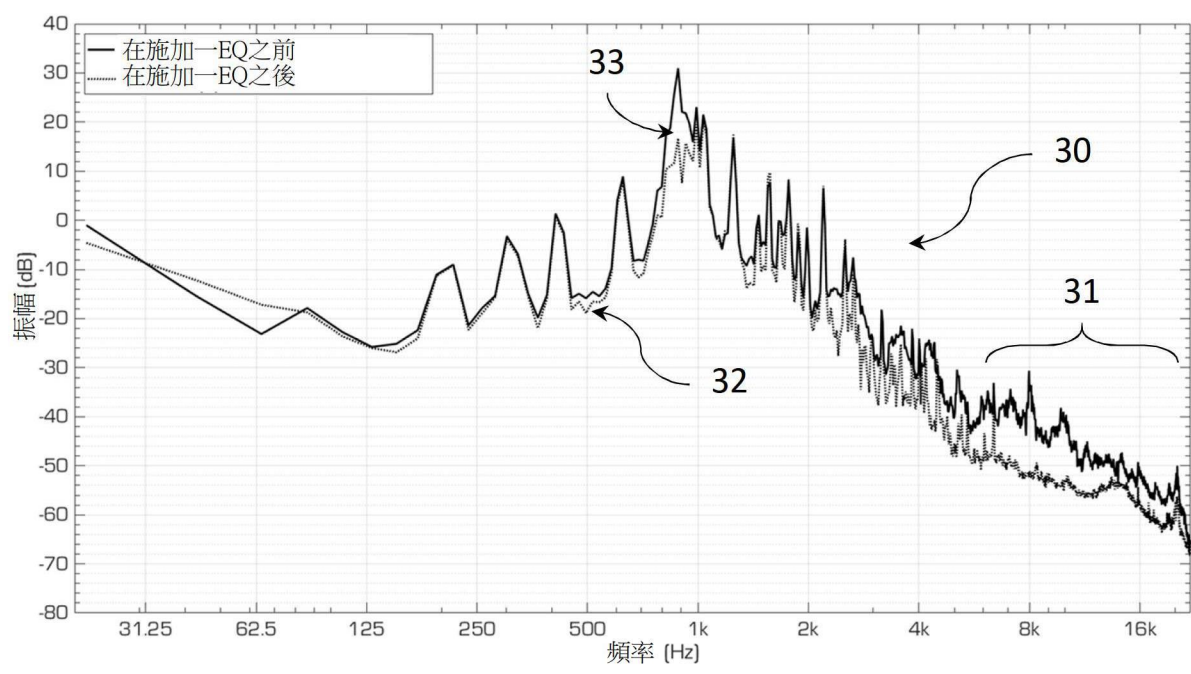
【圖2】



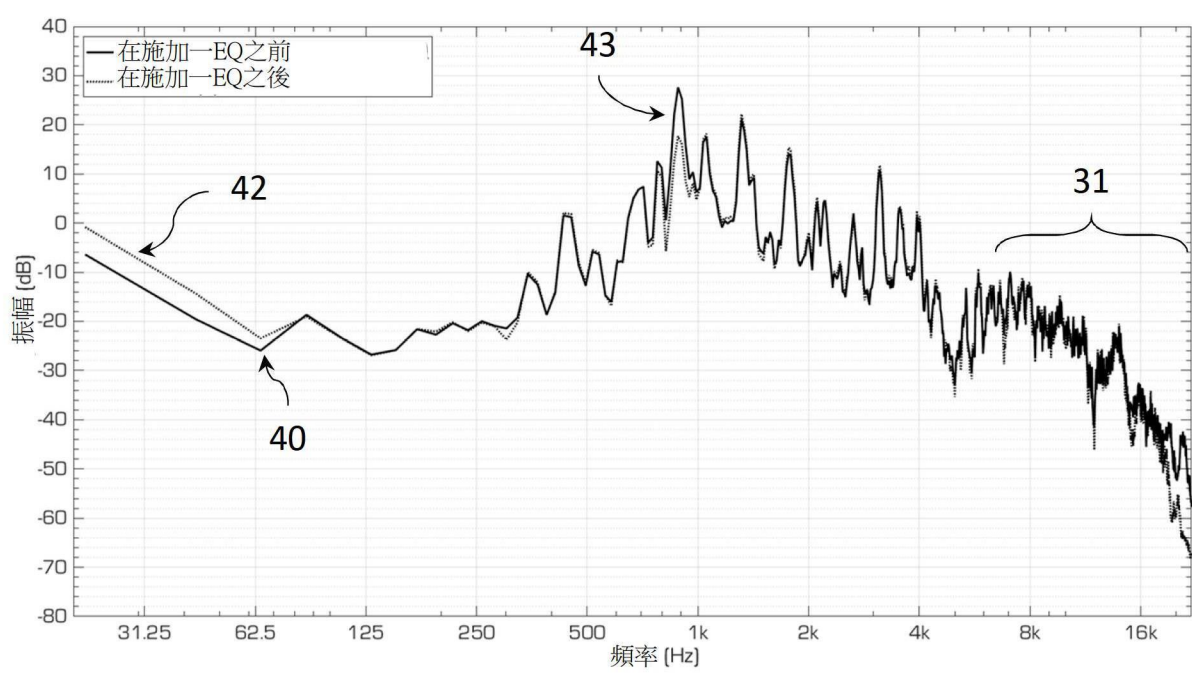
【圖3】



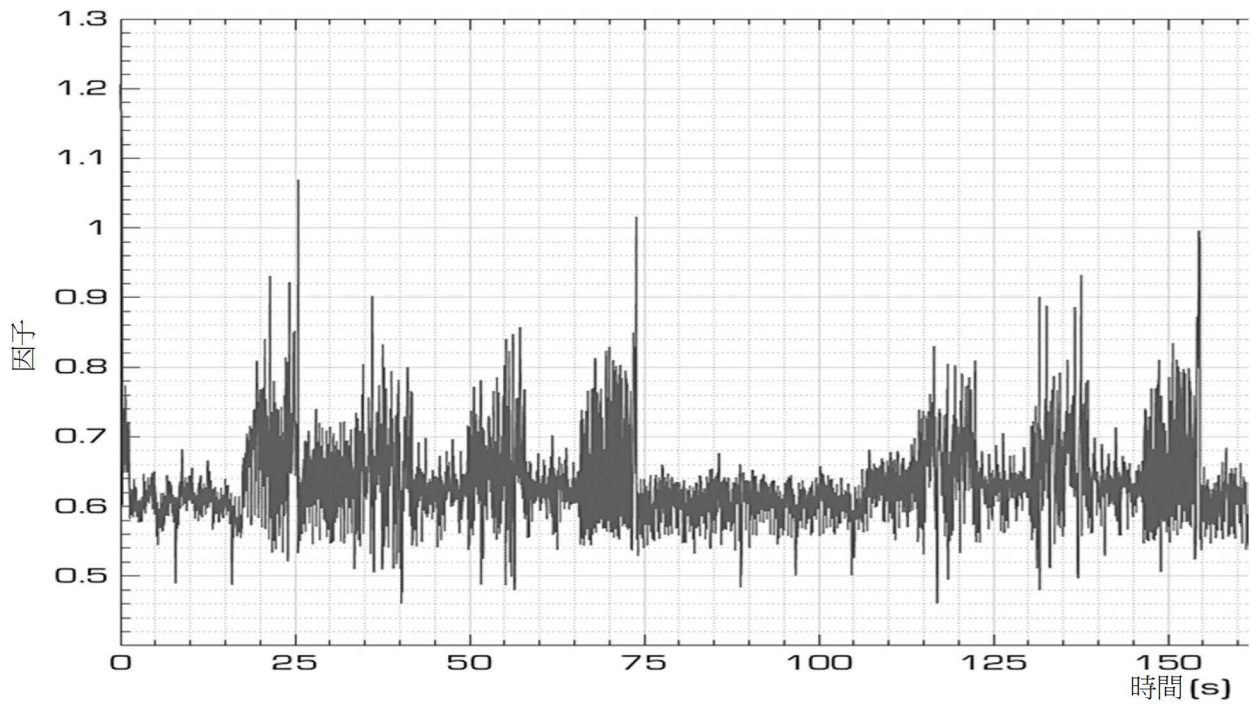
【圖4】



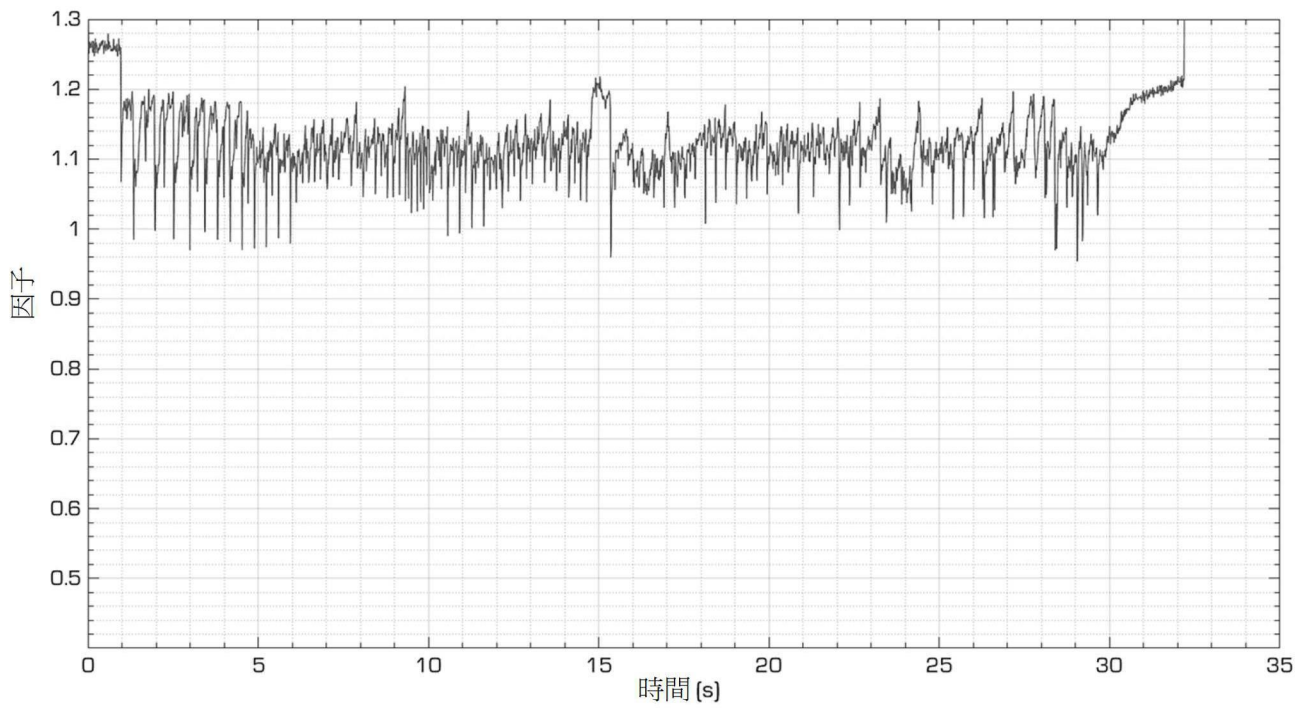
【圖5】



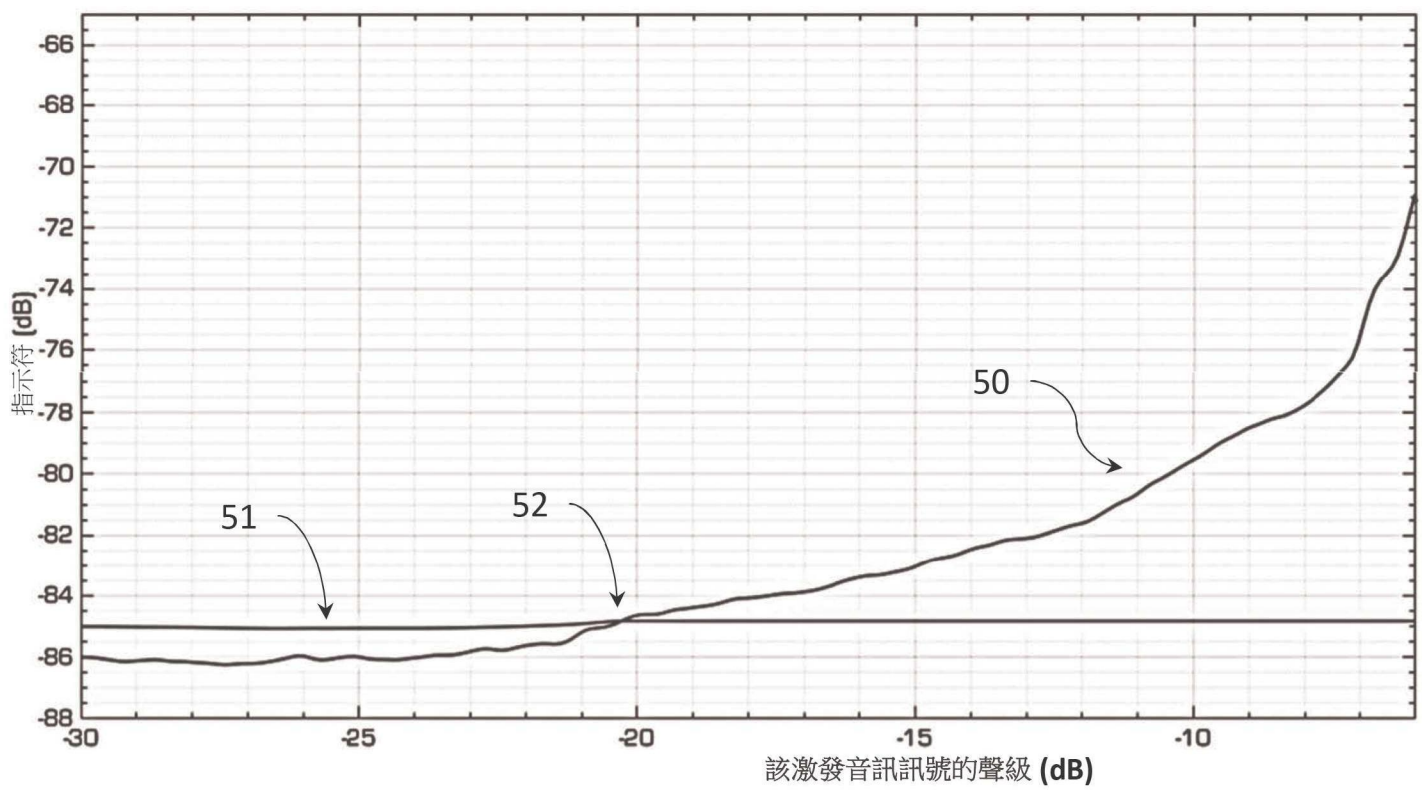
【圖6】



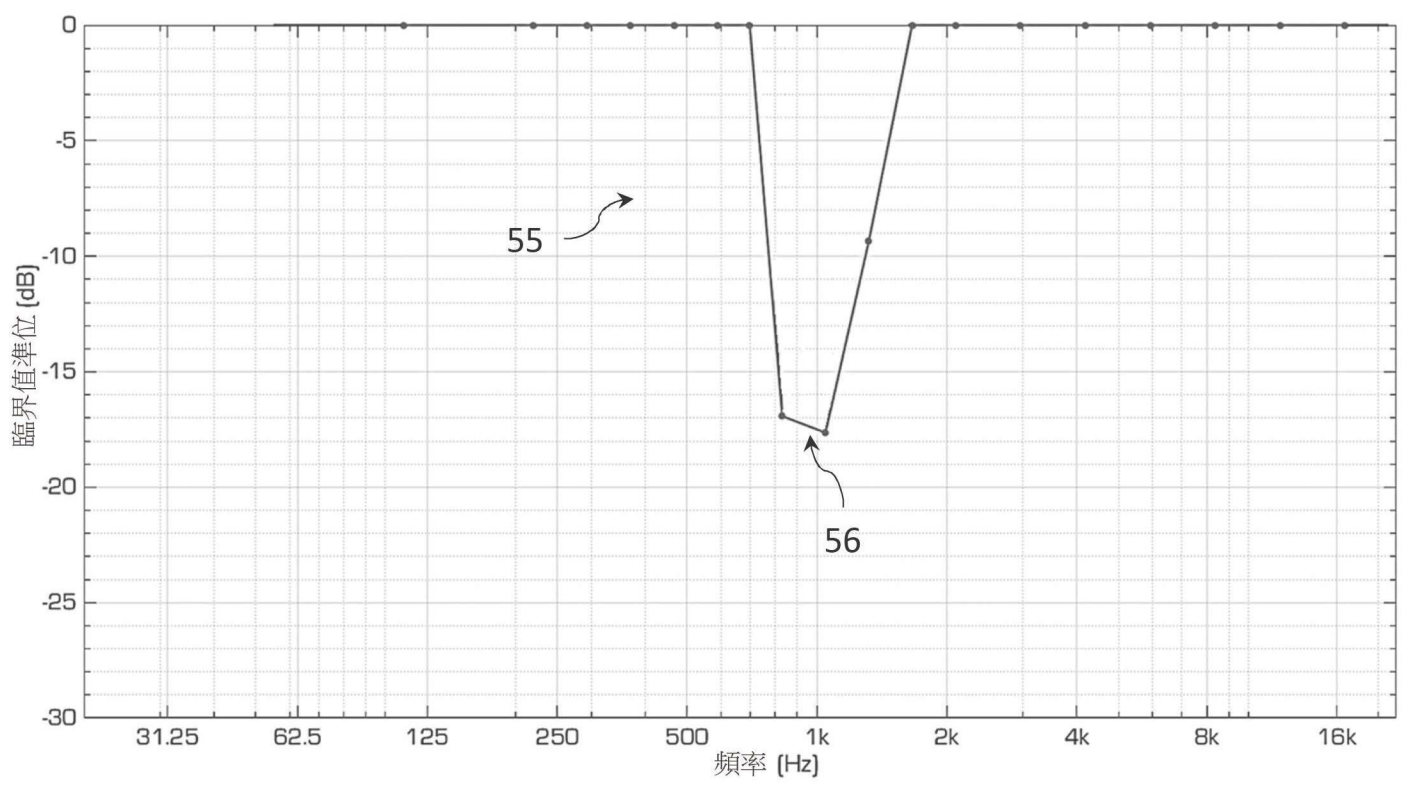
【圖7】



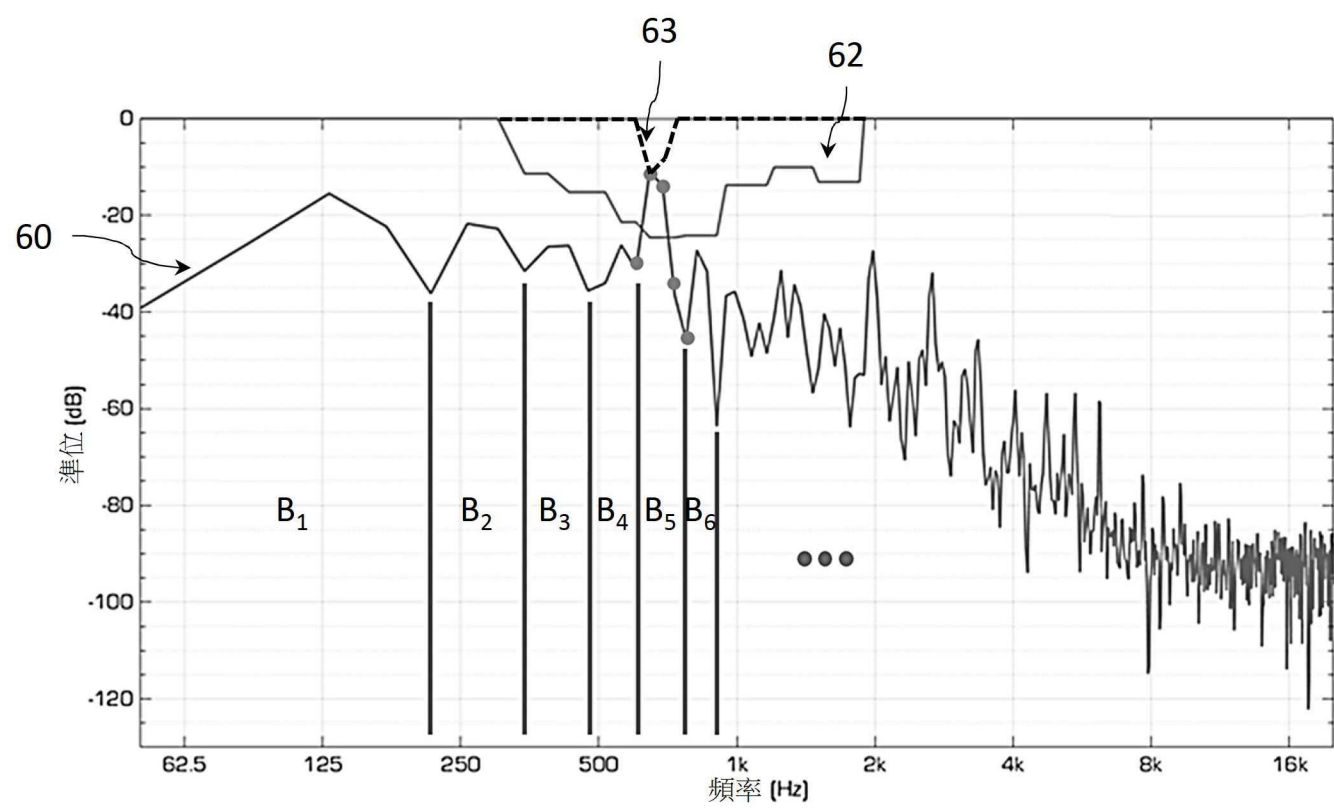
【圖8】



【圖9】



【圖10】



【圖11】