



(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 I673707 B

(45)公告日：中華民國 108 (2019) 年 10 月 01 日

(21)申請案號：107116311

(22)申請日：中華民國 103 (2014) 年 07 月 16 日

(51)Int. Cl. : G10L19/00 (2013.01)

(30)優先權：2013/07/19 歐洲專利局 13306042.6

(71)申請人：瑞典商杜比國際公司(瑞典) DOLBY INTERNATIONAL AB (SE)
荷蘭

(72)發明人：波漢 約哈拿斯 BOEHM, JOHANNES (DE)

(74)代理人：林志剛

(56)參考文獻：

US 2012/0082319A1

F. Zotter, et al., "Energy-Preserving Ambisonic Decoding", ACTA
Acustica United with Acustica, Vol. 98, No. 1, January 2012,[http://docserver.ingentaconnect.com/deliver/connect/dav/16101928/
v98n1/s4.pdf?](http://docserver.ingentaconnect.com/deliver/connect/dav/16101928/v98n1/s4.pdf?)expires=1544790899&id=0000&titleid=75000347&checksum=BC764255A84A
828F769FA17ECAE541B7

審查人員：黃衍勳

申請專利範圍項數：13 項 圖式數：12 共 45 頁

(54)名稱

將以 L_1 個頻道為基礎之輸入聲音訊號產生至 L_2 個揚聲器頻道之方法及裝置，以及得到一能量保留混音矩陣之方法及裝置，用以將以輸入頻道為基礎之聲音訊號混音以用於 L_1 個聲音頻道至 L_2 個揚聲器頻道

(57)摘要

本發明係為一種產生多重頻道聲音之方法，將多頻道聲音內容混音係用於一特殊揚聲器安裝設定，然而，一客戶之聲音安裝設定極可能使用一不同揚聲器配置，本方法確保該訊號之空間訊號分量以相等響度播放。本發明揭示一種得到一能量保留混音矩陣(G)之方法，用以將 L_1 個輸入聲音頻道混音至 L_2 個輸出頻道，該方法包括以下步驟：得到(s711)一第一混音矩陣 \hat{G} ，在該第一混音矩陣 \hat{G} 上執行(s712)一奇異值分解以得到一奇異性矩陣 S ，處理(s713)該奇異性矩陣 S 以得到一處理過之奇異性矩陣 \hat{S} ，判定(s715)一縮放因子 a ，及根據 $G = a U \hat{S} V^T$ 以求出(s716)一改良式混音矩陣 G 。聽到任一揚聲器安裝設定上所播放之多頻道聲音之聲音、響度、音質及空間效果實際上與原揚聲器安裝設定者相等。

Multi-channel audio content is mixed for a particular loudspeaker setup. However, a consumer's audio setup is very likely to use a different placement of speakers. The present invention provides a method of rendering multi-channel audio that assures replay of the spatial signal components with equal loudness of the signal. A method for obtaining an energy preserving mixing matrix (G) for mixing L_1 input audio

channels to L_2 output channels comprises steps of obtaining (s711) a first mixing matrix \hat{G} , performing (s712) a singular value decomposition on the first mixing matrix \hat{G} to obtain a singularity matrix S , processing (s713) the singularity matrix S to obtain a processed singularity matrix \hat{S} , determining (s715) a scaling factor a , and calculating (s716) an improved mixing matrix G according to $\hat{G} = U S V^T$. The perceived sound, loudness, timbre and spatial impression of multi-channel audio replayed on an arbitrary loudspeaker setup practically equals that of the original speaker setup.

指定代表圖：

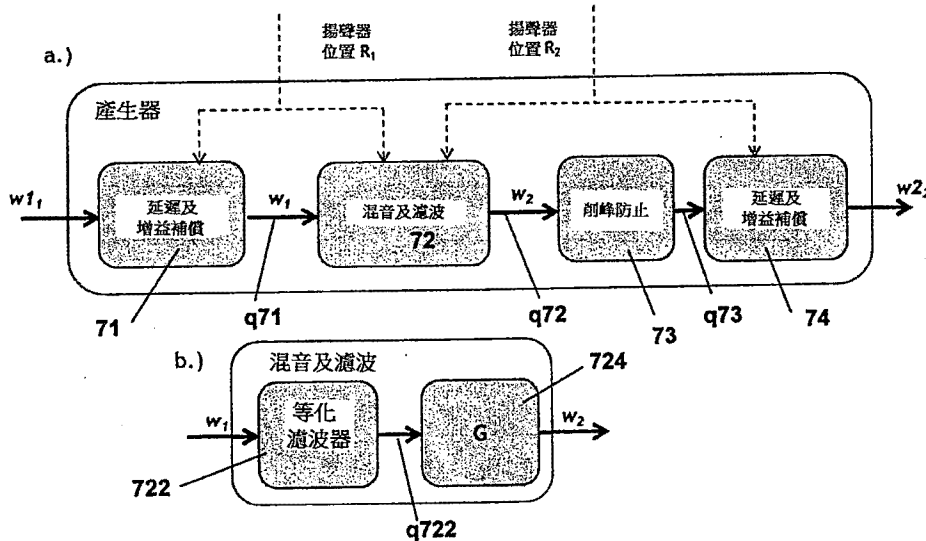


圖7

符號簡單說明：

- 71、74 . . . 延遲及增益補償區塊
- 72 . . . 混音及濾波區塊
- 73 . . . 削峰防止區塊
- 722 . . . 等化濾波器
- 724 . . . 能量保留混音矩陣
- q71 . . . 延遲及增益補償的輸入聲音訊號
- q72 . . . 再混音的聲音訊號
- q73 . . . 削峰的再混音聲音訊號
- q722 . . . 濾波的延遲及增益補償輸入聲音訊號
- W_1 . . . L_1 個數位聲音訊號
- W_2 . . . L_2 個輸出訊號
- w_{1_1} . . . 以 L_1 個頻道為基礎的輸入聲音訊號
- w_{2_2} . . . L_2 個揚聲器頻道

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

將以 L_1 個頻道為基礎之輸入聲音訊號產生至 L_2 個揚聲器頻道之方法及裝置，以及得到一能量保留混音矩陣之方法及裝置，用以將以輸入頻道為基礎之聲音訊號混音以用於 L_1 個聲音頻道至 L_2 個揚聲器頻道

Method and apparatus for rendering L_1 channel-based input audio signals to L_2 loudspeaker channels, and method and apparatus for obtaining an energy preserving mixing matrix for mixing input channel-based audio signals for L_1 audio channels to L_2 loudspeaker channels

【技術領域】

[0001] 本發明相關產生多重頻道聲音訊號的方法，及產生多重頻道聲音訊號的裝置，本發明特別相關產生多重頻道聲音訊號的方法及裝置，該訊號用於揚聲器頻道的 L_1 個頻道到不同的 L_2 個頻道。

【先前技術】

[0002] 以新立體(3D)頻道為基礎的聲音格式提供聲音混音以用於揚聲器頻道，其不僅環繞聆聽位置，亦包括定位在聆聽位置(甜蜜點)上方(高度)及下方的頻道，該混音適合這種揚聲器特殊定位，通用格式係 22.2(即 22 頻道)或 11.1(即 11 頻道)。

[0003] 圖 1 顯示不同揚聲器安裝設定中理想揚聲器位置的二範例：一 22 頻道揚聲器安裝設定(左)及一 12 頻道揚聲器安裝設定(右)，每一節點顯示一揚聲器的虛擬位置，與甜蜜點有不同距離的真實揚聲器位置係藉由增益及延遲補償映射到虛擬位置。用於以頻道為基礎的聲音的一產生器接收 L_1 個數位聲音訊號 W_1 ，並將輸出處理到 L_2 個輸出訊號 W_2 。圖 2 顯示一產生器整合成再製鏈。

[0004] 該產生器使用輸入揚聲器安裝設定的位置資訊及輸出揚聲器安裝設定的位置資訊作為輸入將處理鏈初始化，這顯示在圖 3 中，顯示二主要處理區塊：一混音與濾波區塊 31 及一延遲與增益補償區塊 32。

[0005] 揚聲器位置資訊可在笛卡兒(Cartesian)或球面座標中，可手動輸入用於輸出配置的位置 R_2 ，或透過以特殊測試訊號的麥克風量測或藉由其他任何方法得出。可藉由表格登錄像用於 5 頻道環繞音響的一指示器，使輸入配置的位置 R_1 隨內容產生。假定數個理想標準化揚聲器位置 [9]，該等位置亦可使用數個球面角位置以訊號顯示，假定一不變半徑用於輸入配置。

[0006] 令 $R_2 = [r2_1, r2_2, \dots, r2_{L_2}]$ 具有 $r2_l = [r2_l, \theta2_l, \phi2_l]^T = [r2_l, \hat{\Omega}_l]^T$ 係球面座標中輸出配置的位置，該座標系的原點係該甜蜜點（即聆聽位置）。 $r2_l$ 係該聆聽位置與一揚聲器 l 之間的距離，及 θ_l, ϕ_l 係相關球面角，其指明一揚聲器 l 相關該聆聽位置的空間方向。

[0007] 延遲及增益補償

該等距離係用以得出數個延遲及增益 g_l ，其藉由放大或衰減元素應用到揚聲器饋入，及一延遲線具有 d_l 單元樣本延遲步驟。首先，判定一揚聲器與該甜蜜點之間的最大距離： $r_{2_{max}} = \max([r_{2_1}, \dots, r_{2_{L_2}}])$ 。

用於各揚聲器，藉由：

$$d_l = [(r_{2_{max}} - r_{2_l})f_s/c + 0.5], \quad (1)$$

求出該延遲， f_s 係取樣率， c 係聲音速度(在攝氏溫度 20 度 $c \cong 343\text{m/s}$ (公尺/秒))，及 $[x + 0.5]$ 指明捨入下一整數。

$$\text{藉由 } g_l = \frac{r_{2_l}}{r_{2_{max}}} \quad (2)$$

判定揚聲器增益 g_l ，延遲及增益補償建造區塊的任務係為衰減及延遲比其他揚聲器更靠近聆聽者的揚聲器，使此等較靠近的揚聲器不主導聽到的聲音方向，因此該等揚聲器係設置在如圖 1 所示的虛擬球面上，混音及濾波區塊 31 茲可使用數個虛擬揚聲器位置 $\widehat{\mathbf{R}}_2 = [\widehat{r}_{2_1}, \widehat{r}_{2_2}, \dots, \widehat{r}_{2_{L_2}}]$ 搭配 $\widehat{\mathbf{r}}_{2_l} = [r_{2_{max}}, \widehat{\boldsymbol{\Omega}}_l^T]^T$ 具有一不變的揚聲器距離。

[0008] 混音及濾波

在一初始化相位中，使用輸入及理想化輸出配置 R_1 ， $\widehat{\mathbf{R}}_2$ 的揚聲器位置以得出一 $L_2 \times L_1$ 混音矩陣 \mathbf{G} ，在產生過程期間，此混音矩陣應用到輸入訊號以得出揚聲器輸出訊號。如圖 4 所示，存在二通用措施，在圖 4(a) 所示的第一措施中，混音矩陣係依存於聲音頻率，及藉由：

$$\mathbf{W}_2 = \mathbf{G}\mathbf{W}_1, \quad (3)$$

得出輸出，其中 $\mathbf{W}_1 \in \mathcal{R}^{L_1 \times \tau}$ ， $\mathbf{W}_2 \in \mathcal{R}^{L_2 \times \tau}$ 在矩陣表示法中表示 L_1, L_2 個聲音頻道的輸入及輸出訊號及 τ 時間樣本，

最優的方法係向量基振幅平移 (VBAP) [1]。在其他措施中，如圖 4b) 所示，混音矩陣成為頻率依存的 ($G(f)$)。接著，需要足夠解析度的一濾波器堆積，及根據公式 (3) 將一混音矩陣應用到每一頻帶樣本。

[0009] 用於後者措施的數個範例係 [2]、[3] 及 [4]，為得出該混音矩陣，使用以下措施：如圖 5 所示，一虛擬麥克風陣列放置在該甜蜜點四周，將接收輸入配置 (原方向，左手側) 來的聲音的麥克風訊號 M_1 ，與接收所想揚聲器配置 (右手側) 來的聲音的麥克風訊號 M_2 作比較。令 $\mathcal{M}_1 \in \mathcal{R}^{M \times \tau}$ 表示 M 個麥克風訊號接收該輸入配置發射出的聲音，及 $\mathcal{M}_2 \in \mathcal{R}^{M \times \tau}$ 係來自輸出配置的聲音的 M 個麥克風訊號，可由以下公式得出

$$\mathcal{M}_1 = H_{M,L_1} W_1 \quad (4)$$

$$\text{及 } \mathcal{M}_2 = H_{M,L_2} W_2 \quad (5)$$

具有 $H_{M,L_1} \in \mathbb{C}^{M \times L_1}$ ， $H_{M,L_2} \in \mathbb{C}^{M \times L_2}$ 係自由場中理想聲音發射的複轉移函數，假定球面波或平面波發射，該等轉移函數係頻率依存的，選取相關一濾波器堆積的一中頻 f_m ，可使用公式 (3) 使公式 (4) 與公式 (5) 成為等式，用於每一 f_m ，需要解出以下公式以得到 $G(f_m)$ ：

$$H_{M,L_1} W_1 = H_{M,L_2} G W_1 \quad (6)$$

可得出依存於輸入訊號及使用 H_{M,L_2} 的偽逆矩陣的一解為：

$$G = H_{M,L_2}^+ H_{M,L_1} \quad (7)$$

通常這措施所產生結果並不令人滿意，[2] 及 [5] 並提

出更複雜途徑以解出公式(6)以用於 G 。

[0010] 此外，有一完全不同的訊號適應產生方式，其中將進來聲音內容的方向訊號像聲音物件般加以萃取及產生，將剩餘訊號平移及解關連到輸出揚聲器。按計算複雜度而言，這類聲音產生昂貴許多，且常無法避免人工產物，在此並不使用訊號適應產生，提及只為要完整說明。

[0011] 有一問題在於用戶在家的安裝設定，極可能因客廳的實地限制而使用不同的揚聲器放置，揚聲器數亦會不同。因此產生器的任務為使以頻道為基礎的聲音訊號調適到新安裝設定，以便聽到的聲音、響度、音質及空間效果盡可能接近在其原揚聲器安裝設定像混音室中播放時以原頻道為基礎的聲音。

【發明內容】

[0012] 本發明涉及一種產生聲音訊號的方法，其確保空間訊號分量的播放(即再製)具有該訊號的相等響度(如在原安裝設定中)，後者意指從原混音中的一方向聽到的一方向訊號在產生到新揚聲器安裝設定時亦以相等響度聽到。此外，提供數個濾波器，其將該等輸入訊號等化用以再製一音質盡可能接近在聆聽原安裝設定時所聽到的。

[0013] 在一方面，本發明相關一種將以 L_1 個頻道為基礎的輸入聲音訊號產生到 L_2 個揚聲器頻道的方法，其中 L_1 不同於 L_2 ，如申請專利範圍第 1 項所揭示。根據本發明的一對應裝置揭示在申請專利範圍第 8 項中。

[0014] 在一方面，本發明相關一種得到一能量保留混音矩陣 \mathbf{G} 的方法，用以將以輸入頻道為基礎的聲音訊號混音以用於 L_1 個聲音頻道到 L_2 個揚聲器頻道，如申請專利範圍第 7 項所揭示。根據本發明的一對應裝置揭示在申請專利範圍第 14 項中。

[0015] 在一方面，本發明相關一種電腦可讀媒體，具有數個可執行指令，用以令一電腦執行如申請專利範圍第 1 項所述的方法，或如申請專利範圍第 7 項所述的方法。

[0016] 一種得到一能量保留混音矩陣 \mathbf{G} 的方法，用以將以輸入頻道為基礎聲音訊號混音以用於 L_1 個聲音頻道到 L_2 個揚聲器頻道，該方法包括以下步驟：得到一第一混音矩陣 $\hat{\mathbf{G}}$ ，在該第一混音矩陣 $\hat{\mathbf{G}}$ 上執行一奇異值分解以得到一奇異性矩陣 \mathbf{S} ，處理該奇異性矩陣 \mathbf{S} 以得到一處理過的奇異性矩陣 $\hat{\mathbf{S}}$ 具有 s_m 非零對角元素，根據 $a = \sqrt{\frac{L_1}{s_m}}$ (用於 $L_2 \leq L_1$) 或 $a = \sqrt{\frac{L_2}{s_m}}$ (用於 $L_2 > L_1$) 判定一縮放因子 a ，及根據 $\mathbf{G} = a \mathbf{U} \hat{\mathbf{S}} \mathbf{V}^T$ 求出一混音矩陣 \mathbf{G} 。結果，聽到在任一揚聲器安裝設定上播放的多重頻道聲音的聲音、響度、音質及空間效果盡可能接近以頻道為基礎的聲音，仿如以頻道為基礎的原本聲音在其原本揚聲器安裝設定上播放般。

[0017] 以下參考附圖及後附申請專利範圍的說明，將更了解本發明進一步的目的、特徵及有利點。

能量保留混音矩陣 \mathbf{G} 將以輸入頻道為基礎的聲音訊號混音以用於 L_1 個聲音頻道到 L_2 個揚聲器頻道，該方法包括以下步驟：從虛擬來源位置或方向 $\widehat{\mathbf{R}}_1$ 及目標揚聲器位置或方向 $\widehat{\mathbf{R}}_2$ 得到 s711 一第一混音矩陣 $\widehat{\mathbf{G}}$ ，其中使用一平移方法，根據 $\widehat{\mathbf{G}} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ 在第一混音矩陣 $\widehat{\mathbf{G}}$ 上執行 s712 一奇異值分解，其中 $\mathbf{U} \in \mathcal{R}^{L_2 \times L_2}$ 及 $\mathbf{V} \in \mathcal{R}^{L_1 \times L_1}$ 係正交矩陣，及 $\mathbf{S} \in \mathcal{R}^{L_1 \times L_2}$ 係一奇異性矩陣且具有 s 個第一對角元素係 \mathbf{G} 的奇異值以遞減順序，及 \mathbf{S} 的其他所有元素皆為零，處理 s713 奇異性矩陣 \mathbf{S} ，其中得到一量化奇異性矩陣 $\widehat{\mathbf{S}}$ 具有數個對角元素高於一臨界值設成一，及數個對角元素低於一臨界值設成零，判定 s714 一對角元素數目 s_m ，

其在量化奇異性矩陣 $\widehat{\mathbf{S}}$ 中設成一，根據 $a = \sqrt{\frac{L_1}{s_m}}$ 用於 ($L_2 \leq L_1$) 或 $a = \sqrt{\frac{L_2}{s_m}}$

用於 ($L_2 > L_1$) 判定 s715 一縮放因子 a ，及根據 $\mathbf{G} = a\mathbf{U}\widehat{\mathbf{S}}\mathbf{V}^T$ 求出 s716 一混音矩陣 \mathbf{G} 。該等步驟係執行在一或多個處理元件如數個微處理器、一 GPU(圖形處理單元)的脈絡等中。

[0021] 圖 7 根據本發明的一實施例顯示一產生架構，在根據本發明的產生器過程或產生架構中，使用一外加“增益及延遲補償”區塊 71 用以預先處理不同的輸入安裝設定如球面、柱面或直角輸入安裝設定。此外，使用一修改版“混音及濾波”區塊 72，其能保留原響度。在一實施例中，“混音及濾波”區塊 72 包括一等化濾波器 722。以下相關圖 7b)及圖 8 詳細說明“混音及濾波”區塊 72，一削峰

防止區塊 73 防止由於修正的混音矩陣會發生的訊號溢流。圖 8 顯示該混音及濾波區塊中一等化濾波器 722 的結構，該等化濾波器原則上係一濾波堆積，具有 L_1 個濾波器 EF_1, \dots, EF_{L_1} ，各輸入頻道有一個。以下將說明該等濾波器的設計及特徵。

[0022] 新產生器解決以下至少一問題：

首先，可將以新立體聲音頻道為基礎的內容混音以用於一球面、直角或柱面的揚聲器安裝設定，此資訊需要隨同表格登錄的一指數傳送，以訊號顯示該輸入配置(其假定一不變的揚聲器半徑)能用以算出真實輸入揚聲器位置。或者，完整輸入揚聲器位置座標可隨同作為媒介資料的內容傳送，為使用獨立於該混音類型的數個混音矩陣，提供一增益及延遲補償用於該輸入配置。

[0023] 第二，提供一能量保留混音矩陣 G ，傳統上，混音矩陣並非能量保留的，相較於使用一播放系統的相同校正時在混音室中的內容響度，能量保留確保該內容在產生後具有相同響度(參閱附錄及 [6]、[7]、[8])。這亦確保如 22 頻道輸入或 10 頻道輸入具有相等‘響度，K-加權、相對全刻度’(LKFS)內容響度在產生後似乎同樣大聲。本發明的一有利點在於允許產生能量(及響度)保留、頻率獨立的混音矩陣。請注意，相同原則亦可用於頻率依存混音矩陣，但其並非是想要的。一頻率獨立混音矩陣按計算複雜度而言是有利的，但可能常有再混音後音質會改變的缺點，為避開這種混音後音質不匹配的情形，在一實

施例中，在混音前將在各輸入揚聲器頻道應用數個簡單濾波器，這就是等化濾波器 722，以下將提出一種設計這類濾波器的方法。

[0024] 能量保留產生法具有一缺點，即峰值聲音訊號分量可能會訊號超載，一外加削峰防止區塊 73 防止超載，這可簡單理解為一飽和，較複雜地說，此區塊係用於峰值聲音的一動態處理器，此建構區塊包含在本發明的一實施例中。

[0025] 以下相關輸入增益及延遲補償 71。

若輸入配置係藉由一表格登錄加上混音室資訊以訊號表示，像直角、柱面或球面配置，則配置座標係自特別預備的表格(如 RAM(隨機存取記憶體)) 讀取為球面座標，若該等座標係直接傳送則轉換到球面座標。令 $R_1 = [r1_1, r1_2, \dots, r1_{L_1}]$ 具有 $r1_l = [r1_l, \theta1_l, \phi1_l]^T = [r1_l, \Omega_l^T]^T$ 係此輸入配置的位置。在一第一步驟中，偵測到最大半徑： $r1_{max} = \max([r1_1, \dots, r1_{L_2}])$ ，因用於此建構區塊只對相對差異感興趣，因此該等半徑係藉由 $r2_{max}$ 縮放的 $r1_l$ ，其可得自該輸出配置的增益及延遲補償初始化：

$$\widehat{r1}_l = r1_l \frac{r2_{max}}{r1_{max}} \quad (8)$$

以 $\widehat{r1}_{max} = r2_{max}$ 求出每一揚聲器的延遲定位點數 \check{d}_l 及增益值 \check{g}_l 如下：

$$\check{d}_l = [(r2_{max} - \widehat{r1}_l) f_s / c + 0.5], \quad (9)$$

f_s 係取樣頻率、 c 係聲音速度(在攝氏 20 度溫度 $c \cong 343m/s$ (公尺/秒))，及 $[x + 0.5]$ 指明捨入下個整數。

$$\text{藉由 } \check{g}_l = \frac{\widehat{r}_{1l}}{\widehat{r}_{1\max}} \quad (10)$$

判定揚聲器增益 \check{g}_l ，該混音及濾波區塊茲可使用數個虛擬揚聲器位置 $\widehat{\mathbf{R}}_1 = [\widehat{\mathbf{r}}_{11}, \widehat{\mathbf{r}}_{12}, \dots, \widehat{\mathbf{r}}_{1L_1}]$ 搭配 $\widehat{\mathbf{r}}_{1l} = [\widehat{r}_{1\max}, \boldsymbol{\Omega}_l^T]^T$ 具有一不變的揚聲器距離。

[0026] 以下將說明混音矩陣設計。

首先，討論揚聲器訊號的能量及聽到的響度。圖 7a 以方塊圖顯示定義該等描述變數，必須將 L_1 個揚聲器訊號處理到 L_2 個訊號(通常 $L_2 \leq L_1$)，揚聲器饋入訊號 W_1 的播放理想上聽到的響度應與聆聽混音室中以最適揚聲器安裝設定播放的響度相等。令 W_1 係 L_1 個揚聲器頻道(列)及 τ 樣本(行)的一矩陣。

[0027] τ 時間樣本區塊的訊號 W_1 的能量定義如下：

$$E_{w_1} = \|W_1\|_{fro}^2 = \sum_{i=1}^{\tau} \sum_{l=1}^{L_1} W_{1l,i}^2 = \sum_{i=1}^{\tau} \mathbf{w}_{1t}^T \mathbf{w}_{1t} \quad (11)$$

在此 $W_{l,i}$ 係 W_1 的矩陣元素， l 表示揚聲器指數， i 表示樣本指數， $\|\cdot\|_{fro}$ 表示 Frobenius(弗羅貝尼烏斯)矩陣範數， \mathbf{w}_{1t} 係 W_1 的第 t 行向量，及 $[\]^T$ 表示向量或矩陣轉置。

[0028] 此能量 E_w 提供以頻道為基礎的聲音的一最佳聲音量測估算，如 [6]、[7]、[8] 中所定義，其中 K-濾波器抑制頻率低於 200Hz(赫)。 W_1 的混音提供數個訊號 W_2 ，該訊號能量在混音後成為：

$$E_{w_2} = \|W_2\|_{fro}^2 = \sum_{i=1}^{\tau} \sum_{l=1}^{L_2} W_{2l,i}^2 \quad (12)$$

其中 L_2 係新的揚聲器數目，具有 $L_2 \leq L_1$ 。

假定由一混音矩陣 \mathbf{G} 執行該產生過程，從 W_1 得出數

個訊號 W_2 如下：

$$W_2 = GW_1 \quad (13)$$

評估 E_{W_2} 及使用 $W_1 = [w_{1,1}, \dots, w_{1,t}, \dots, w_{1,\tau}]$ 的行向量分解搭配以

$$w_{1,t} = [w_{1,t,1}, \dots, w_{1,t,l}, \dots, w_{1,t,L}]^T, \text{ 接著得到}$$

$$E_{W_2} = \sum_{i=1}^{\tau} \sum_{l=1}^L W_{2,l,i}^2 = \sum_{i=1}^{\tau} [Gw_{1,t}]^T M w_{1,t} = \sum_{i=1}^{\tau} w_{1,t}^T G^T G w_{1,t} \quad (14)$$

[0029] 在一實施例中，接著得到響度保留如下。

$$\text{若：} E_1 = E_2 \quad (15)$$

則原訊號混音的響度保留在新產生的訊號中。

由公式(14)明顯看出混音矩陣 M 需是正交的及

$$G^T G = I \quad (16)$$

其中 I 係 $L_1 \times L_1$ 單元矩陣。

[0030] 根據本發明的一實施例，可得到一最適產生矩陣(亦稱為混音矩陣或解碼矩陣)如下。

步驟 1：藉由使用平移方向得出一傳統混音矩陣 \hat{G} ，將來自原揚聲器組的一單個揚聲器 l_1 看做是一音源，將由新揚聲器安裝設定的 L_2 個揚聲器再製。較佳的平移方法係 VBAP(向量基振幅平移)[1]或用於定頻的穩健平移[2](即用於此步驟可使用一已知技術)。為判定混音矩陣 \hat{G} ，使用修正的揚聲器位置 \hat{R}_2 ， \hat{R}_1 ， \hat{R}_2 用於輸出配置及 \hat{R}_1 用於虛擬來源方向。

步驟 2：使用緊奇異值分解，該混音矩陣表示為三個矩陣的乘積：

$$\hat{G} = U S V^T \quad (17)$$

$U \in \mathcal{R}^{L_2 \times L_2}$ 及 $V \in \mathcal{R}^{L_1 \times L_1}$ 係正交矩陣，及 $S \in \mathcal{R}^{L_1 \times L_2}$ 具有 s 個第一對角元素(該等奇異值在遞減順序中)，具有 $s \leq L_2$ 。其他矩陣元素係零。請注意，用於 $L_2 \leq L_1$ 的情形保持如此，(再混音 $L_2 = L_1$ ，向下混音 $L_2 < L_1$)，用於向上混音的情形($L_2 > L_1$)， L_2 在此區段需由 L_1 取代。

步驟 3：由 S 形成一新矩陣 \hat{S} ，其中該等對角元素由一值一取代，但極低值的奇異值 $s_{\&} \ll s_{max}$ 則由零取代。通常在 -10dB (分貝)... -30dB 或更小的範圍中選取一臨界值(如 -20dB 係一典型值)，由於將發生二群組對角元素：具較大值元素及具相當較小值元素，因此由實際範例中的實際數據明顯看出該臨界值。該臨界值係在此二群組之中用以區別。

[0031] 用於大部分的揚聲器設定，非零對角元素數 s_m 係 $s_m = L_2$ ，但用於一些設定變成較低且 $s_m < L_2$ ，其意指 $L_2 - s_m$ 個揚聲器將不用以播放內容；只因該等揚聲器未有任何聲音資訊，因此仍無聲。

[0032] 令 s_m 表示將由一取代的最終奇異值，接著藉由：

$$G = a U \hat{S} V^T \quad (18)$$

判定混音矩陣 G 具有該縮放因子。

$$a = \sqrt{\frac{L_1}{s_m}} \text{ 用於 } (L_2 \leq L_1) \quad (19)$$

或，分別地

$$a = \sqrt{\frac{L_2}{s_m}} \text{ 用於 } (L_2 > L_1)。 \quad (19')$$

該縮放因子係得自： $G^T G = a^2 V \hat{S}^2 V^T = a^2 V V^T$ ，其中 $V V^T$

具有 s_m 本徵值等於一，其意指 $|\mathbf{V}\mathbf{V}^T|_{fro} = \sqrt{s_m}$ 。因此，將 L_1 個訊號簡單向下混音到 s_m 訊號將減少能量，除非 $s_m=L_1$ (換言之：當輸出揚聲器數匹配輸入揚聲器數)。利用 $|\mathbf{I}_{L_1}|_{fro} = \sqrt{L_1}$ ，一縮放因子 $a = \sqrt{\frac{L_1}{s_m}}$ 補償向下混音期間的能量損失。

[0033] 作為一範例，以下說明一奇異性矩陣的處理，例如，根據公式(17)： $\hat{\mathbf{G}} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ ，使用緊奇異值分解將一初始(傳統)混音矩陣分解，奇異性矩陣 \mathbf{S} 係在該形式的對角矩陣

$$\begin{bmatrix} s_1 & \dots & 0 \\ 0 & s_2 & \vdots \\ \vdots & \ddots & 0 \\ 0 & \dots & s_L \end{bmatrix} \text{ 有 } s_1 \geq s_2 \geq \dots \geq s_L \text{ (即 } s_1 = s_{\max}) \text{。}$$

接著藉由將該等係數 $s_1 \geq s_2 \geq \dots \geq s_L$ 設成 1 或 0 以處理該奇異矩陣，依各係數是否高於一臨界值如 $0.06 * s_{\max}$ 而定，此類似於該等係數的一相對量化，該等臨界值示範為 0.06，但亦可(當以分貝表示時)在 -10dB 或更低的範圍中。

[0034] 用於一情形，具有如 $L=5$ 及如只有 s_1 及 s_2 係高於臨界值，及 s_3 , s_4 及 s_5 係低於臨界值，作為結果的處理過(或"量化")的奇異性矩陣 $\hat{\mathbf{S}}$ 係

$$\hat{\mathbf{S}} = \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ 因此其非零對角係數的數目 } s_m \text{ 係二。}$$

[0035] 以下將說明等化濾波器 722。

當在不同 3D(立體)安裝設定之間混音時，特別當從立

體安裝設定混音到 2D(平面)安裝設定時，音質會改變，例如用於 3D 到 2D，原來自上方的一聲音今只使用平面上的數個揚聲器再製，等化濾波器的工作係將此音質不匹配減到最小及將能量保留最大化。如圖 7b 所示，個別的濾波器 F_l 在應用該混音矩陣前應用到輸入配置的 L_1 個頻道的各頻道，以下將顯示理論上的推演，及說明得出該等濾波器的頻率響應。

[0036] 使用根據圖 7 的一模型及公式(4)及(5)，為求方便，在此皆再列出二公式：

$$\mathcal{M}_1 = H_{M,L_1} W_1 \quad (20)$$

及

$$\mathcal{M}_2 = H_{M,L_2} W_2 \quad (21)$$

利用 $H_{M,L_1} \in \mathbb{C}^{M \times L_1}$ ， $H_{M,L_2} \in \mathbb{C}^{M \times L_2}$ 係假定球面波或平面波輻射的自由場中理想聲音輻射的複變換函數。此等矩陣係頻率函數，並可使用位置資訊 $\widehat{R}_2, \widehat{R}_1$ 求出。定義 $W_2 = \widetilde{G} W_1$ ，其中 \widetilde{G} 係一頻率函數。不用公式(4)及(5)，如在先前技術段落提及，茲將等化該等能量。並因想要等化以用於輸入配置中揚聲器方向的聲音，因此可解決在一時間(在 L_1 個上的迴路)為各輸入揚聲器的考量。

[0037] 在該等虛擬麥克風測量以用於輸入安裝設定的能量，若只有一揚聲器 l 有作用，則係由

$$|\mathcal{M}_{1,l}|_{fro}^2 = |h_{M,l} w_{1l}|_{fro}^2 \quad (22)$$

提供，有 $h_{M,l}$ 代表 H_{M,L_1} 的第 l 行，及 w_{1l} 代表 W_1 的一列，即揚聲器 l 的時間訊號具有 τ 個樣本。將 Frobenius

(弗羅貝尼烏斯) 範數模擬重寫到公式(11)，可進一步將公式(22)求值到：

$$|\mathcal{M}_{1,l}|_{fro}^2 = \sum_{i=1}^r \mathbf{w}_{1l}^T \mathbf{w}_{1l} \mathbf{h}_{M,l}^H \mathbf{h}_{M,l} = E_{wl} \mathbf{h}_{M,l}^H \mathbf{h}_{M,l} \quad (23)$$

其中 $()^H$ 係共軛複轉置(Hermitian transposed(埃爾米特轉置))，及 E_{wl} 係揚聲器訊號 l 的能量，向量 $\mathbf{h}_{M,l}$ 係由數個複指數所複合(參閱公式(31)、(32))，及一元素與其共軛複形相乘等於一，因此 $\mathbf{h}_{M,l}^H \mathbf{h}_{M,l} = L_1$ ：

$$|\mathcal{M}_{1,l}|_{fro}^2 = E_{wl} L_1 \quad (24)$$

混音後在虛擬麥克風的量測係由 $\mathcal{M}_2 = \mathbf{H}_{M,L_2} \tilde{\mathbf{G}} \mathbf{W}_1$ 提供。若只有一揚聲器有作用，則可重寫成：

$$\mathcal{M}_{2,l} = \mathbf{H}_{M,L_2} \tilde{\mathbf{g}}_l \mathbf{w}_{1l} \quad (25)$$

$\tilde{\mathbf{g}}_l$ 係 $\tilde{\mathbf{G}}$ 的第 l 行。將 $\tilde{\mathbf{G}}$ 定義成可分解成相關揚聲器 l 的一頻率依存部分，及由公式(24)得出混音矩陣 \mathbf{G} ：

$$\tilde{\mathbf{G}}(f) = \text{diag}(\mathbf{b}(f)) \mathbf{G} \quad (26)$$

\mathbf{b} 作為 L_1 複元素的一頻率依存向量，及 (f) 表示頻率依存性，其在以下已略過以求簡化。利用此方式，公式(25)變成：

$$\mathcal{M}_{2,l} = \mathbf{H}_{M,L_2} b_l \mathbf{g} \mathbf{w}_{1l} \quad (27)$$

其中 \mathbf{g} 係 \mathbf{G} 的第 l 行，及 b_l 係 \mathbf{b} 的第 l 個元素。使用以上 Frobenius 範數的相同考量，在該等虛擬麥克風的能量成為：

$$|\mathcal{M}_{2,l}|_{fro}^2 = E_{wl} (\mathbf{H}_{M,L_2} b_l \mathbf{g})^H (\mathbf{H}_{M,L_2} b_l \mathbf{g}) \quad (28)$$

其可求值到：

$$|\mathcal{M}_{2,l}|_{fro}^2 = E_{wl} b_l^2 \mathbf{g}^T \mathbf{H}_{M,L_2}^H \mathbf{H}_{M,L_2} \mathbf{g} \quad (29)$$

茲可分別根據公式(24)及公式(29)將該等能量建立方程，及解出 b_l 以用於各頻率 f ：

$$b_l = \sqrt{\frac{L_1}{\mathbf{g}^T \mathbf{H}_{M,L_2}^H \mathbf{H}_{M,L_2} \mathbf{g}}} \quad (30)$$

公式(30)的 b_l 係頻率依存增益因子或縮放因子，及由於 b_l 及 $\mathbf{H}_{M,L_2}^H \mathbf{H}_{M,L_2}$ 係頻率依存的，因此可作為等化濾波器 722 的係數使用以用於各頻帶。

[0038] 以下說明用於等化濾波器 722 的實用濾波器設計。以下將虛擬麥克風陣列半徑及轉移函數列入考量。為匹配人類最佳感知音質效果，選取 0.09 公尺的一麥克風半徑 r_M (人腦的平均直徑係大約 0.18 公尺)，在環繞原點(甜蜜點，聆聽位置)在一球面或半徑 r_M 上放置 $M \gg L_1$ 個虛擬麥克風，在[11]可找出合適的位置，在(該座標系統的)原點加入一額外虛擬麥克風。

使用一平面波或球面波模型設計轉移矩陣 $\mathbf{H}_{M,L_2} \in \mathbb{C}^{M \times L_2}$ ，為稍後由於該等增益及延遲補償階段，可忽略該等振幅衰減效應。令 $h_{m,l}$ 係該轉移矩陣 $\mathbf{H}_{M,L}$ 的一抽象矩陣元素，以用於從揚聲器 l 到麥克風 m 的自由場轉移函數(其亦指明該等矩陣的行及列指數)。藉由

$$h_{m,l} = e^{ikr_m \cos(\gamma_{l,m})} \quad (31)$$

提供該平面波轉移函數， i 係想像單元， r_m 係麥克風位置的半徑 (r_M 或零以用於原位置)，及 $\cos(\gamma_{l,m}) = \cos \theta_l \cos \theta_m + \sin \theta_l \sin \theta_m \cos(\phi_l - \phi_m)$ 係揚聲器 l 與麥克風 m 位置球面角的餘弦，藉由 $k = \frac{2\pi f}{c}$ 提供頻率依存性，

將該奇異值分解應用到以傳統方式得出的一混音矩陣。

[0044] 在一實施例中，根據公式(19)或公式(19')，藉由

$\sqrt{\frac{L_1}{s_m}}$ (用於 $L_1 \geq L_2$) 的一因子，或藉由 $\sqrt{\frac{L_2}{s_m}}$ (用於 $L_1 < L_2$) 的一因子，將該矩陣縮放。

可藉由使用各種不同平移方法如 VBAP(向量基振幅平移)或穩健平移得出傳統矩陣，此外，傳統矩陣亦使用理想化的輸入及輸出揚聲器位置(球面投射，參閱上述說明)。因此，在一方面，本發明相關一種濾波方法，在應用該混音矩陣前將該 L_1 個輸入頻道濾波，在一實施例中，在一延遲及增益補償區塊 71 中，將使用不同揚聲器位置的數個輸入訊號映射到一球面投射。

[0045] 在一實施例中，從藉由上述方法求出的頻率響應得出數個等化濾波。

[0046] 在一實施例中，由以下數個建構及處理區塊組裝成一元件，用以將 L_1 個頻道以頻道為基礎的聲音內容，產生到 L_2 個頻道以頻道為基礎的聲音內容：

- 數個輸入(及輸出)增益及延遲補償區塊 71, 74，目的為將該等輸入及輸出揚聲器位置映射到一虛擬球面，上述混音矩陣可應用需要此類球面結構；

- 數個等化濾波器 722，係由上述方法得出，在輸入增益及延遲補償後用以將 L_1 個頻道濾波；

- 一混音單元 72，用以藉由應用上述方法得出的能量保留混音矩陣 724 將 L_1 個輸入頻道混音到 L_2 個輸出頻

道，等化濾波器 722 可為混音單元 72 的一部分，或可為一分開模組；

- 一訊號溢流偵測及削峰防止區塊 73，用以防止訊號超載到 L_2 個頻道的訊號；及

- 一輸出增益及延遲校正區塊。

[0047] 在一實施例中，一種得到一能量保留混音矩陣 G 用以將 L_1 個輸入聲音頻道混音到 L_2 個輸出頻道的方法，包括以下步驟：得到 s711 一第一混音矩陣 \hat{G} ，在第一混音矩陣 \hat{G} 上執行 s712 一奇異值分解以得到一奇異性矩陣 S ，處理 s713 奇異性矩陣 S 以得到一處理過的奇異性矩陣 \hat{S} ，判定 s715 一縮放因子 a ，及根據 $G = aU\hat{S}V^T$ 求出 s716 一改良式混音矩陣 G 。一有利點在於在任一揚聲器安裝設定上播放的多重頻道聲音，實際上與原揚聲器安裝設定所聽到的聲音、響度、音質及空間效果相等。

[0048] 最後，參考圖 11 說明一示範過程用以設定一播放位階，一粉紅雜訊測試用以藉由調整揚聲器放大率 G_l 以位階調整各揚聲器的聲壓位階。在混音及呈現地點中的聲壓位階 (SPL) 調整及在混音室中的內容響度位階調整，能在節目或項目之間切換時令聽到的響度保持不變。如在 [8] 中所述，設定各揚聲器饋入的放大器增益 G_l ，以便具有 $-18 \text{ dBFS}_{\text{rms}}$ (全刻度分貝均方根) 的一數位全頻粉紅雜訊輸入造成 $78 \pm 5 \text{ dBA}$ (A 加權分貝) 的一聲壓位階。

[0049] 關於內容響度位階校正，若依此方式設定混音設備及呈現地點的播放位階，則項目或節目之間的切換

可能不用進一步的位階調整。用於以頻道為基礎的內容，若在混音地點將內容調到令人愉悅的一響度位階，就可簡單達到此目的。用於此令人愉悅的聆聽位階的參考可以是整個項目本身或一錨訊號的響度。

若係整個項目本身，則若內容係儲存為一檔案，則這情形對於“短形式內容”係有用的。除了藉由聆聽來調整以外，根據 EBU R128 建議，響度量測(LUFS)中的一響度量測[6]亦可用以響度調整該內容。LUFS 的另一名稱係源自 ITU-R BS.1770 建議的‘響度，K-加權，相對全刻度’ [7] (1LUFS=1LKFS)，不幸的是，[6]支援用於安裝設定的內容只達到 5 頻道環繞音響，而 22 頻道檔案的響度量測，其中全部 22 頻道皆由一的相等頻道有效值進行因數分解，可與聽到的響度相關連，但尚未藉由全面清單測試得到證據或證明。

若此訊號係一錨訊號如對話，則相關該訊號選取位階，這對於“長形式內容”的影片聲音、現場錄音及廣播等有用。茲延伸該愉悅聆聽位階的一額外要求係口說文字的可理解性，同樣在藉由聆聽的調整之外，亦可相關一響度量測將該內容正規化，如在 ATSC A/85[8]中所界定，將該內容的第一部分識別為錨部分，接著求出如[7]中定義的一量測，或判定此等訊號及一增益因子以達到目標響度，使用該增益因子來將整個項目加以縮放。不幸的是，支援的最大頻道數同樣限制到 5 個。

[0050] 圖 12 顯示如在 EBU R128 [2] 及 ATSC A/85

[4]中使用的 ITU-R BS.1770 [3]響度量測。[2]提議將全內容項目測得的響度以增益調整到 -23 dBLKFS，在[4]中只測量錨訊號響度並以增益調整該內容，使該等錨部分達到 -24 dBLKFS 的目標響度。

出於藝術考量，內容必須藉由在混音工作室聆聽加以調整，可使用響度量測作為一支援及用以顯示並未超過一明確界定的響度。

根據公式(11)的能量 E_w 提供聽到的如一錨訊號響度的一公平估算以用於超過 200Hz(赫)的頻率。因 K-濾波器抑制頻率低於 200Hz [5]，因此 E_w 與該響度量測大致成正比。

[0051] 雖然本發明的基本新穎特徵如應用到其數個較佳實施例所顯示、說明及指出者，但應瞭解，不背離本發明的精神，熟諳此藝者在所述裝置及方法中，在揭示構件的形式及細節中，及在其操作中，可作出不同的省略、添加和變動。明白地希望為達成相同結果，在實質相同方式中執行實質相同功能的該等元件的所有組合皆涵蓋在本發明的範圍內。亦完全希望及考慮到從一說明實施例到另一說明實施例的元件替換，應瞭解本發明單純藉由範例加以說明，不背離本發明的範圍即可作出細節修改。

[0052] 在本說明書及(在適當處)後附申請專利範圍及附圖中揭示的各特徵可獨立地提供或在任何適當組合中提供，數個特徵在適當處可實施在硬體、軟體或兩者的組合中，數個連接方式在可應用處可實施作為無線連接或有線

連接，不必是直接或專用的連接。

[0053] 後附申請專利範圍中出現的參考數字符號係只在繪示方式，不應在後附申請專利範圍的範疇中具有任何限制效果。

【符號說明】

[0054]

31、72：混音及濾波區塊(單元)

32、71、74：延遲及增益補償區塊

73：削峰防止區塊(單元)

722：等化濾波器

724：能量保留混音矩陣

EF_1, \dots, EF_{L1} ：濾波器

G ：能量保留混音矩陣

$G(f)$ ：頻率依存混音矩陣

G_l ：揚聲器放大率

q71：延遲及增益補償之輸入聲音訊號

q72：再混音聲音訊號

q73：削峰之再混音聲音訊號

q722：濾波之延遲及增益補償輸入聲音訊號

R_1 、 R_2 ：揚聲器位置

S ：奇異矩陣

s60、s61、s622、s624、s63、s64、s711、s712、s713、

s714、s715、s716：步驟

s710：方法

U、V：矩陣

W_1 ： L_1 個數位聲音訊號

W_2 ： L_2 個輸出訊號

w_{1_1} ：以 L_1 個頻道為基礎的輸入聲音訊號

w_{2_2} ： L_2 個揚聲器頻道

參考文獻

[1] Pulkki, V., “使用向量基振幅平移的虛擬音源定位 (Virtual Sound Source Positioning Using Vector Base Amplitude Panning)”, 聲音工程協會期刊, 第 45 期, 456-466 頁(1997 年 6 月)。

[2] Poletti, M., “用於非統一揚聲器布局的穩健 2D 環繞音響再製 (Robust two-dimensional surround sound reproduction for non-uniform loudspeaker layouts)”, 聲音工程協會期刊, 第 55(7/8)期; 598-610 頁, 2007 年 7/8 月。

[3] O. Kirkeby 及 P. A. Nelson, “平面波音場的再製 (Reproduction of plane wave sound fields)”, 音響學協會期刊, 第 94(5)期, 2992-3000 頁(1993 年)。

[4] Fazi, F.; Yamada, T; Kamdar, S.; Nelson P.A.; Otto, P., “基於虛擬麥克風陣列的環繞音響平移技術 (Surround Sound Panning Technique Based on a Virtual Microphone Array)”, AES(進階加密標準)協定: 128 (2010 年 5 月)文件編號: 8119。

[5] Shin, M.; Fazi, F.; Seo, J.; Nelson, P.A., “有效率立體 (3D) 音場再製 (Efficient 3-D Sound Field Reproduction)”, AES 協定: 130 (2011 年 5 月)文件編號: 8404。

[6] EBU (歐洲廣播聯盟電子工程師學會) 技術建議書 R128, “聲音訊號的響度正規化及允許最大位準

(Loudness Normalization and Permitted Maximum Level of Audio Signals)”，日內瓦，2010。

[<http://tech.ebu.ch/docs/r/r128.pdf>]

[7] ITU-R(國際通信聯盟無線通信部門)建議書 BS.1770-2，“測量聲音程式設計響度及真正峰值聲音位階的演算法 (Algorithms to measure audio programme loudness and true-peak audio level)”，日內瓦，2011年。

[<http://tech.ebu.ch/docs/r/r128.pdf>]

[8] ATSC(美國廣播電視標準)A/85，“建立及維護數位電視聲音響度的技術 (Techniques for Establishing and Maintaining Audio Loudness for Digital Television)”，進階電視系統委員會，華盛頓，2011年7月25日。

[9] ITU-R BS 775-1 建議書(1994年)。

[10] Hamasaki, K.; Nishiguchi T.; Okumura, R.; Nakayama, Y.; Ando, A., “超高畫質電視(UHDTV)的 22.2 多重頻道聲音系統 (A 22.2 multichannel sound system for ultrahigh-definition TV (UHDTV)”，SMPTE(電影及電視協會)動態影像期刊，44-49 頁，2008年4月。

[11] Jorg Fliege 及 Ulrike Maier，計算球面的求體積公式的二階段方法 (A two-stage approach for computing cubature formulae for the sphere)，Fachereich Mathematil, Universitat Dortmund，技術論文，1999年，可於網址 <http://www.personal.soton.ac.uk/jf1w07/nodes/nodes.html> 找到節點數及論文。

I673707

發明摘要

【發明名稱】(中文/英文)

將以 L_1 個頻道為基礎之輸入聲音訊號產生至 L_2 個揚聲器頻道之方法及裝置，以及得到一能量保留混音矩陣之方法及裝置，用以將以輸入頻道為基礎之聲音訊號混音以用於 L_1 個聲音頻道至 L_2 個揚聲器頻道

Method and apparatus for rendering L_1 channel-based input audio signals to L_2 loudspeaker channels, and method and apparatus for obtaining an energy preserving mixing matrix for mixing input channel-based audio signals for L_1 audio channels to L_2 loudspeaker channels

【中文】

本發明係為一種產生多重頻道聲音之方法，將多頻道聲音內容混音係用於一特殊揚聲器安裝設定，然而，一客戶之聲音安裝設定極可能使用一不同揚聲器配置，本方法確保該訊號之空間訊號分量以相等響度播放。本發明揭示一種得到一能量保留混音矩陣 (G) 之方法，用以將 L_1 個輸入聲音頻道混音至 L_2 個輸出頻道，該方法包括以下步驟：得到 (s711) 一第一混音矩陣 \hat{G} ，在該第一混音矩陣 \hat{G} 上執行 (s712) 一奇異值分解以得到一奇異性矩陣 S ，處理 (s713) 該奇異性矩陣 S 以得到一處理過之奇異性矩陣 \hat{S} ，判定 (s715) 一縮放因子 a ，及根據 $G = aU\hat{S}V^T$ 以求出 (s716) 一改良式混音矩陣 G 。聽到任一揚聲器安裝設定上所播放之多頻道聲音之聲音、響度、音質及空間效果實際上與原揚聲器安裝設定者相等。

【 英文 】

Multi-channel audio content is mixed for a particular loudspeaker setup. However, a consumer's audio setup is very likely to use a different placement of speakers. The present invention provides a method of rendering multi-channel audio that assures replay of the spatial signal components with equal loudness of the signal. A method for obtaining an energy preserving mixing matrix (\mathbf{G}) for mixing L_1 input audio channels to L_2 output channels comprises steps of obtaining (s711) a first mixing matrix $\hat{\mathbf{G}}$, performing (s712) a singular value decomposition on the first mixing matrix $\hat{\mathbf{G}}$ to obtain a singularity matrix \mathbf{S} , processing (s713) the singularity matrix \mathbf{S} to obtain a processed singularity matrix $\hat{\mathbf{S}}$, determining (s715) a scaling factor a , and calculating (s716) an improved mixing matrix \mathbf{G} according to $\hat{\mathbf{G}} = \mathbf{U} \mathbf{S} \mathbf{V}^T$. The perceived sound, loudness, timbre and spatial impression of multi-channel audio replayed on an arbitrary loudspeaker setup practically equals that of the original speaker setup.

【代表圖】

【本案指定代表圖】：第(7)圖。

【本代表圖之符號簡單說明】：

71、74：延遲及增益補償區塊

72：混音及濾波區塊

73：削峰防止區塊

722：等化濾波器

724：能量保留混音矩陣

q71：延遲及增益補償的輸入聲音訊號

q72：再混音的聲音訊號

q73：削峰的再混音聲音訊號

q722：濾波的延遲及增益補償輸入聲音訊號

W_1 ： L_1 個數位聲音訊號

W_2 ： L_2 個輸出訊號

w_{1_1} ：以 L_1 個頻道為基礎的輸入聲音訊號

w_{2_2} ： L_2 個揚聲器頻道

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：無

申請專利範圍

1. 一種將以 L_1 個頻道為基礎之輸入聲音訊號產生至 L_2 個揚聲器頻道之方法，其中 L_1 不同於 L_2 ，該方法包括以下步驟：

- 判定該等 L_1 個輸入聲音訊號之一混音類型，其中該混音類型明確界定被用於定義揚聲器位置的一座標系並且其中可能之混音類型包含球面、柱面及直角中之至少一者；

- 根據該判定之混音類型在該等 L_1 個輸入聲音訊號上執行一第一延遲及增益補償，其中得到一延遲及增益補償之輸入聲音訊號，具有 L_1 個頻道及具有一界定混音類型；

- 將該延遲及增益補償之輸入聲音訊號混音以用於 L_2 個聲音頻道，其中得到一再混音之聲音訊號以用於 L_2 個聲音頻道；

- 將該再混音之聲音訊號削峰，其中得到一削峰之再混音聲音訊號以用於 L_2 個聲音頻道；以及

- 在該削峰之再混音聲音訊號上執行一第二延遲及增益補償以用於 L_2 個聲音頻道，其中得到 L_2 個揚聲器頻道；

其中，該混音使用藉由以下步驟所得到的一能量保留混音矩陣 G ：

- 使用一平移方法，從數個虛擬來源方向 \widehat{R}_1 及數個目標揚聲器方向 \widehat{R}_2 得到一第一混音矩陣 \widehat{G} ；

- 根據 $\widehat{G} = USV^T$ 在該第一混音矩陣 \widehat{G} 上執行一奇異值分解，其中 $U \in \mathcal{R}^{L_2 \times L_2}$ 與 $V \in \mathcal{R}^{L_1 \times L_1}$ 係正交矩陣，及

$S \in \mathcal{R}^{L_1 \times L_2}$ 係一奇異性矩陣及具有 s 個第一對角元素，係 G 在遞減順序中之奇異值，及 S 之所有其他元素係零；

- 處理該奇異性矩陣 S ，其中利用高於一臨界值之數個對角元素設成一，及低於一臨界值之數個對角元素設成零，得到一量化奇異性矩陣 \hat{S} ；

- 判定該量化奇異性矩陣 \hat{S} 中設成一之對角元素之數目 s_m ；

- 根據 $a = \sqrt{\frac{L_1}{s_m}}$ 用於 $(L_2 \leq L_1)$ 或 $a = \sqrt{\frac{L_2}{s_m}}$ 用於 $(L_2 > L_1)$ 判定一縮放因子 a ；以及

- 根據 $G = a U \hat{S} V^T$ 計算出該能量保留混音矩陣 G 。

2. 根據申請專利範圍第1項所述之方法，更包括一濾波步驟，將具有 L_1 個頻道之延遲及增益補償之輸入聲音訊號濾波，

其中，得到一濾波之延遲及增益補償輸入聲音訊號，並且

其中，該混音使用該濾波之延遲及增益補償輸入聲音訊號。

3. 根據申請專利範圍第2項所述之方法，其中具有 L_1 個頻道之延遲及增益補償輸入聲音訊號之濾波使用一等化濾波器，具有不同類型濾波器以用於該等頻道，其中至少一頻道使用一高通濾波器，及至少一頻道使用一低通濾波器。

4. 根據申請專利範圍第1項所述之方法，其中該界定之混音類型係球面。

5. 根據申請專利範圍第1項所述之方法，其中將該輸入訊號最適化以用於 L_1 個正規揚聲器位置，及將該產生最適化以用於 L_2 個任意揚聲器位置，其中該等任意揚聲器位置中之至少一者不同於該等正規揚聲器位置。

6. 一種用於產生一能量保留混音矩陣 \mathbf{G} 之電腦所實施的方法，該能量保留混音矩陣 \mathbf{G} 用以將以輸入頻道為基礎之聲音訊號混音以用於 L_1 個聲音頻道至 L_2 個揚聲器頻道，該方法包括以下由該電腦所執行的步驟：

- 從數個虛擬來源方向 $\widehat{\mathbf{R}}_1$ 及數個目標揚聲器方向 $\widehat{\mathbf{R}}_2$ 得到一第一混音矩陣 $\widehat{\mathbf{G}}$ ，其中使用一平移方法；

- 根據 $\widehat{\mathbf{G}} = \mathbf{U} \mathbf{S} \mathbf{V}^T$ 在該第一混音矩陣 $\widehat{\mathbf{G}}$ 上執行一奇異值分解，其中 $\mathbf{U} \in \mathcal{R}^{L_2 \times L_2}$ 與 $\mathbf{V} \in \mathcal{R}^{L_1 \times L_1}$ 係正交矩陣，及 $\mathbf{S} \in \mathcal{R}^{L_1 \times L_2}$ 係一奇異性矩陣及具有 s 個第一對角元素，係 \mathbf{G} 在遞減順序中之奇異值，及 \mathbf{S} 之所有其他元素係零；

- 處理奇異性矩陣 \mathbf{S} ，其中利用高於一臨界值之數個對角元素設成一，及低於一臨界值之數個對角元素設成零，得到一量化奇異性矩陣 $\widehat{\mathbf{S}}$ ；

- 判定該量化奇異性矩陣 $\widehat{\mathbf{S}}$ 中設成一之對角元素之數目 s_m ；

- 根據 $a = \sqrt{\frac{L_1}{s_m}}$ 用於 $(L_2 \leq L_1)$ 或 $a = \sqrt{\frac{L_2}{s_m}}$ 用於 $(L_2 > L_1)$ ，

判定一縮放因子 a ；以及

- 根據 $\mathbf{G} = a \mathbf{U} \widehat{\mathbf{S}} \mathbf{V}^T$ 計算出該能量保留混音矩陣 \mathbf{G} 。

7. 一種將以 L_1 個頻道為基礎之輸入聲音訊號產生至 L_2 個揚聲器頻道之裝置，其中 L_1 不同於 L_2 ，該裝置包括至

少一個包括至少各個以下單元中之一者的處理器：

- 判定單元，用以判定 L_1 個輸入聲音訊號之一混音類型，其中該混音類型明確界定被用於定義揚聲器位置的一座標系並且其中可能之混音類型包含球面、柱面及直角中之至少一者；

- 第一延遲及增益補償單元，用以根據該判定之混音類型在 L_1 個輸入聲音訊號上執行一第一延遲及增益補償，其中得到一延遲及增益補償之輸入聲音訊號，具有 L_1 個頻道及具有一界定之混音類型；

- 混音單元，用以將該延遲及增益補償之輸入聲音訊號混音以用於 L_2 個聲音頻道，其中得到一再混音之聲音訊號以用於 L_2 個聲音頻道；

- 削峰單元，用以將該再混音之聲音訊號削峰，其中得到一削峰之再混音聲音訊號以用於 L_2 個聲音頻道；以及

- 第二延遲及增益補償單元，用以在該削峰之再混音聲音訊號上執行一第二延遲及增益補償以用於 L_2 個聲音頻道，其中得到 L_2 個揚聲器頻道；

其中，將該延遲及增益補償之輸入聲音訊號混音以用於 L_2 個聲音頻道的該混音單元係使用藉由一混音矩陣產生單元所得到的一能量保留混音矩陣 G ，其中該混音矩陣產生單元包括一或更多個處理器用以實施以下模組：

- 第一計算模組，用於使用一平移方法，從數個虛擬來源方向 \widehat{R}_1 及數個目標揚聲器方向 \widehat{R}_2 得到一第一混音矩陣 \widehat{G} ；

- 奇異值分解模組，用於根據 $\hat{G} = USV^T$ 在該第一混音矩陣 \hat{G} 上執行一奇異值分解，其中 $U \in \mathcal{R}^{L_2 \times L_2}$ 與 $V \in \mathcal{R}^{L_1 \times L_1}$ 係正交矩陣，及 $S \in \mathcal{R}^{L_1 \times L_2}$ 係一奇異性矩陣及具有 s 個第一對角元素，係 G 在遞減順序中之奇異值，及 S 之所有其他元素係零；

- 處理模組，用於處理該奇異性矩陣 S ，其中利用高於一臨界值之數個對角元素設成一，及低於一臨界值之數個對角元素設成零，得到一量化奇異性矩陣 \hat{S} ；

- 計數模組，用於判定該量化奇異性矩陣 \hat{S} 中設成一之對角元素之數目 s_m ；

- 第二計算模組，用於根據 $a = \sqrt{\frac{L_1}{s_m}}$ 用於 ($L_2 \leq L_1$) 或 $a = \sqrt{\frac{L_2}{s_m}}$ 用於 ($L_2 > L_1$) 判定一縮放因子 a ；以及

- 第三計算模組，用於根據 $G = aU\hat{S}V^T$ 計算出一混音矩陣 G 。

8. 根據申請專利範圍第7項所述之裝置，更包括一等化濾波器，用以將具有 L_1 個頻道之延遲及增益補償之輸入聲音訊號濾波，其中得到一濾波之延遲及增益補償輸入聲音訊號。

9. 根據申請專利範圍第8項所述之裝置，其中該等化濾波器包括不同類型濾波器，其係用於該等頻道，其中至少一頻道使用一高通濾波器，及至少一頻道使用一低通濾波器。

10. 根據申請專利範圍第7項所述之裝置，其中該界定之混音類型係球面。

11. 根據申請專利範圍第7項所述之裝置，其中將該等輸入訊號最適化以用於 L_1 個正規揚聲器位置，及將該產生最適化以用於 L_2 個任意揚聲器位置，其中該等任意揚聲器位置中之至少一者係不同於該等正規揚聲器位置。

12. 一種得到一能量保留混音矩陣 G 之裝置，用以將以輸入頻道為基礎之聲音訊號混音以用於 L_1 個聲音頻道至 L_2 個揚聲器頻道，該裝置包括至少一個處理器，該處理器包括至少一個處理元件用以實施模組：

- 第一計算模組，用於從數個虛擬來源方向 \widehat{R}_1 及數個目標揚聲器方向 \widehat{R}_2 得到一第一混音矩陣 \widehat{G} ，其中一平移方法被使用；

- 奇異值分解模組，用於根據 $\widehat{G} = USV^T$ 在該第一混音矩陣 \widehat{G} 上執行一奇異值分解，其中 $U \in \mathcal{R}^{L_2 \times L_2}$ 與 $V \in \mathcal{R}^{L_1 \times L_1}$ 係正交矩陣，及 $S \in \mathcal{R}^{L_1 \times L_2}$ 係一奇異性矩陣及具有 s 個第一對角元素，係 G 在遞減順序中之奇異值，及 S 之所有其他元素係零；

- 處理模組，其處理該奇異性矩陣 S ，其中利用高於一臨界值之數個對角元素設成一，及低於一臨界值之數個對角元素設成零，得到一量化奇異性矩陣 \widehat{S} ；

- 計數模組，用於判定該量化奇異性矩陣 \widehat{S} 中設成一之對角元素之數目 s_m ；

- 第二計算模組，用於根據 $a = \sqrt{\frac{L_1}{s_m}}$ 用於 $(L_2 \leq L_1)$ 或 $a = \sqrt{\frac{L_2}{s_m}}$ 用於 $(L_2 > L_1)$ 判定一縮放因子 a ；以及

- 第三計算模組，用於根據 $G = aU\widehat{S}V^T$ 計算出該能量

保留混音矩陣 G 。

13. 一種非暫態的電腦可讀取儲存媒體，其具有儲存於其上之指令，當該指令在一電腦上被執行時，導致該電腦實行根據申請專利範圍第1項之方法。

【圖式簡單說明】

[0018] 以下將參考附圖說明本發明的數個示範實施例，圖中：

圖 1 係二揚聲器安裝設定的範例；

圖 2 係一已知普遍結構，用以產生內容以用於一新揚聲器安裝設定；

圖 3 係一般習知結構用於以頻道為基礎的聲音產生；

圖 4 係將 L_1 個頻道混音到 L_2 個輸出頻道的二途徑：
a.) 一頻率獨立混音矩陣 G ，及 b.) 一頻率依存混音矩陣 $G(f)$ ；

圖 5 係一虛擬麥克風陣列，用以將原安裝設定發射出的聲音(輸入配置)和所要的輸出配置作出較；

圖 6a)係以流程圖根據本發明繪示將以 L_1 個頻道為基礎的輸入聲音訊號產生到 L_2 個揚聲器頻道的方法；

圖 6b)係以流程圖根據本發明繪示得到一能量保留混音矩陣 G 的方法；

圖 7 係根據本發明一實施例的產生架構；

圖 8 係混音及濾波區塊中一濾波器實施例的結構；

圖 9 係用於 5 頻道再混音的示範頻率響應；

圖 10A 及 10B 係用於 22 頻道再混音的示範頻率響應；

圖 11 以圖繪示調整各揚聲器的聲壓位階；及

圖 12 係如在 EBU R128 及 ATSC A/85 中使用的 ITU-R BS.1770 響度量測。

【實施方式】

[0019] 圖 6a)根據本發明的一實施例以流程圖顯示將以 L_1 個頻道為基礎的輸入聲音訊號產生到 L_2 個揚聲器頻道的方法，將以 L_1 個頻道為基礎的輸入聲音訊號 w_{11} 產生到 L_2 個揚聲器頻道的方法，其中 L_1 個與 L_2 個不同，包括以下步驟：判定 s_{60} L_1 個輸入聲音訊號的一混音類型，根據判定的混音類型在 L_1 個輸入聲音訊號上執行一第一延遲及增益補償 s_{61} ，其中得到一延遲及增益補償的輸入聲音訊號，具有 L_1 個頻道及具有一限定混音類型，將該延遲及增益補償的輸入聲音訊號混音 s_{62} 以用於 L_2 個聲音頻道，其中得到一再混音聲音訊號以用 L_2 個聲音頻道，將該再混音聲音訊號削峰 s_{63} ，其中得到一削峰的再混音聲音訊號以用於 L_2 個聲音頻道，及在該再混音聲音訊號上執行一第二延遲及增益補償 s_{64} 以用於 L_2 個聲音頻道，其中得到至 L_2 個揚聲器頻道 w_{22} 。可能的混音類型包含球面、柱面及直角(或更普遍係立體)中的至少一者，在一實施例中，該方法包括又一濾波步驟，在一等化濾波器中將具有 L_1 個頻道的延遲及增益補償輸入聲音訊號 q_{71} 濾波，其中得到一濾波過的延遲及增益補償輸入聲音訊號。雖然該等化濾波原則上獨立於一能量保留混音矩陣的使用，及可不用該等能量保留混音矩陣，但兩者結合使用特別有利。

[0020] 圖 6b)根據本發明的一實施例以流程圖顯示得到一能量保留混音矩陣 G 的方法，方法 s_{710} 用以得到一

第 107116311 號

民國 108 年 3 月 15 日修正

f 係頻率及 c 係聲音速度，藉由

$$h_{m,l} = e^{-ikr_{l,m}} \quad (32)$$

提供該球面波轉移函數，以 $r_{l,m}$ 為揚聲器 l 到麥克風 m 的距離。

[0039] 使用在 F_N 離散頻率上的一迴路及在所有輸入配置揚聲器 L_1 個上的一迴路，求出該濾波器的頻率響應 $B_{resp} \in \mathbb{C}^{L_1 \times F_N}$ ：

根據以上在 5.2 “最適產生矩陣的設計” 中的說明求出 **G**：

用於 (f=0; f=f+f 個步驟； f< F_N f 個步驟) /* 在頻率上的迴路 */

$$k=2*\pi*f/342;$$

根據公式(31)或公式(32)求出 $H_{M,L_2}(f)$

$$\tilde{H} = H_{M,L_2}^H H_{M,L_2}$$

用於 (l=1; l++; l<= L_1) /* 在輸入頻道上的迴路 */

$$g = G(:, 1)$$

$$B_{resp}(l, f) = \sqrt{\frac{L_1}{g^T \tilde{H} g}}$$

結束

結束

可使用一標準技術從頻率響應 $B_{resp}(l, f)$ 得出該等濾波響應。通常，可能得出位階等於或小於 64 的一 FIR(有限脈衝響應)濾波器設計，或使用串聯雙四角形區域的 IIR(無限脈衝響應)濾波器設計，計算複雜性甚至更小。圖

9、10A 及 10B 顯示數個設計範例。

[0040] 在圖 9 中，顯示濾波器的頻率響應範例以用於五頻道 ITU 安裝設定 [9](L,R,C,Ls,Rs)到 ± 30 度 2 頻道立體音響的再混音，及作為示範結果的 2×5 混音矩陣 G 。使用 [2] 根據段落 5.2 得出混音矩陣以用於 500Hz (赫)，使用一平面波模型用於該轉移函數。如所示，該等濾波器中的二者(上列，用於該等頻道中的二者)原則上具有低通(LP)特性，及該等濾波器中的三者(下列，用於其餘三頻道)原則上具有高通(HP)特性，因該等濾波器一起形成一等化濾波器(或等化濾波堆積)，因此希望該等濾波器不具有理想的 HP 或 LP 特性。通常，並非全部濾波器具有大致相同特性，以便利用至少一 LP 或至少一 HP 濾波器以用於不同頻道。

[0041] 在圖 10A 及 10B 中，顯示數個濾波器的示範響應，以用於 22.2 NHK(日本放送協會)安裝設定的 22 頻道 [10]到 ITU 5 頻道環繞立響 [9]的再混音，及作為結果的一 5×22 混音矩陣。

[0042] 本發明可用以利用任意定義的 L_1 個揚聲器位置以調整以聲音頻道為基礎的內容，使能播放到 L_2 個真實的揚聲器位置。

[0043] 在一方面，本發明相關一種產生 L_1 個頻道到 L_2 個頻道以頻道為基礎的聲音的方法，其中使用一響度及能量保留混音矩陣，如以上在”最適產生矩陣的設計”段落中所述，該矩陣係由奇異值分解以得出，在一實施例中，