



(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 I622043 B

(45) 公告日：中華民國 107 (2018) 年 04 月 21 日

(21) 申請案號：105117508

(22) 申請日：中華民國 105 (2016) 年 06 月 03 日

(51) Int. Cl. : **G10L21/0272 (2013.01)**(71) 申請人：瑞昱半導體股份有限公司 (中華民國) REALTEK SEMICONDUCTOR CORP. (TW)  
新竹科學園區創新二路二號

(72) 發明人：李明唐 LEE, MING-TANG (TW) ; 朱仲石 CHU, CHUNG-SHIH (TW)

(74) 代理人：吳豐任；戴俊彥

(56) 參考文獻：

TW I397057

TW 200627235A

CN 101622669B

CN 101957443B

CN 102075831B

審查人員：黃衍勳

申請專利範圍項數：10 項 圖式數：10 共 28 頁

(54) 名稱

聲源分離方法與裝置

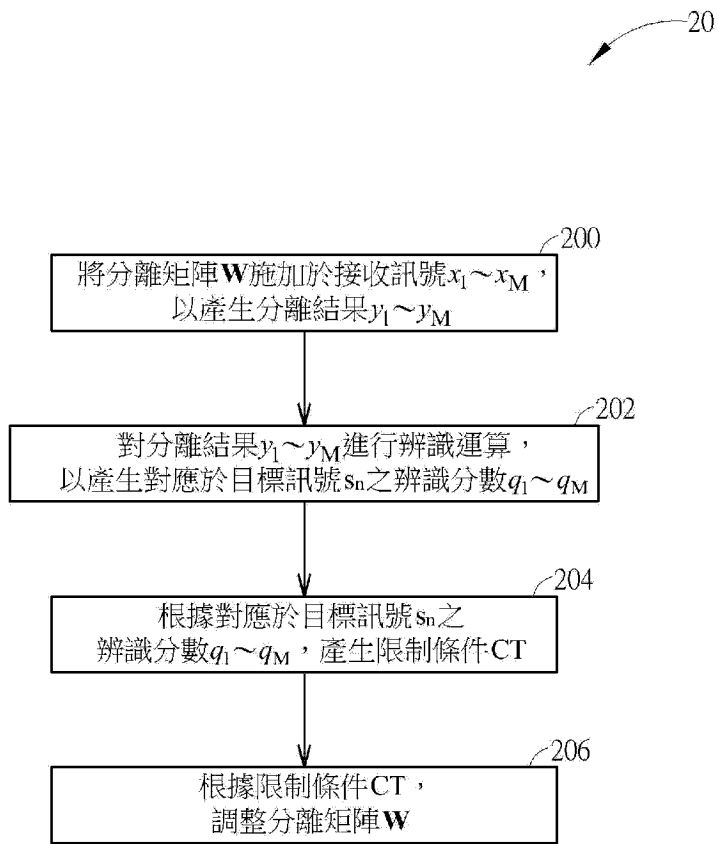
METHOD AND DEVICE OF AUDIO SOURCE SEPARATION

(57) 摘要

一種聲源分離方法，包含有將一分離矩陣施加於複數個接收訊號，以產生複數個分離結果；對該複數個分離結果進行一辨識運算，以產生複數個辨識分數；根據該複數個辨識分數，產生一限制條件；以及根據該限制條件，調整該分離矩陣；其中，調整後之該分離矩陣施加於該複數個接收訊號，以自該接收訊號中分離出更新的複數個分離結果。

A method of audio source separation includes applying a demixing matrix on a plurality of received signals to generate a plurality of separated results; performing a recognition operation on the plurality of separated results to generate a plurality of recognition scores; generating a constraint according to the plurality of recognition scores; and adjusting the demixing matrix according to the constraint; where the adjusted demixing matrix is applied to the plurality of received signals to generate a plurality of updated separated results from the plurality of received signals.

指定代表圖：



第2圖

## 【發明說明書】

【中文發明名稱】 聲源分離方法與裝置

【英文發明名稱】 Method and Device of Audio Source Separation

【技術領域】

【0001】 本發明係指一種聲源分離方法與裝置，尤指一種隨目標訊號的空間變化而據以調整的聲源分離方法與裝置。

【先前技術】

【0002】 語音輸入及語音辨識已廣泛地應用於手機等電子產品中，而多麥克風則常用於強化語音辨識的效果。在具有多麥克風之語音辨識系統中，可透過適應性波束形成器（Adaptive Beamformer）技術進行空間濾波以強化指定方向訊號，再對強化後的訊號進行語音辨識。適應性波束形成器需先對聲源進行來向（Direction-of-Arrival, DoA）估測來獲取或修正波束形成器之指向，然而，適應性波束形成器的缺點在於容易受到來向估測誤差的影響而導致指定方向失真。另外，習知技術中已發展出利用具有限制條件的盲訊號分離

（Constrained-Blind-Source-Separation, CBSS）方法來產生分離矩陣，其可從多麥克風陣列之接收訊號中分離出複數個聲源，並可同時解決習知盲訊號分離之分離聲源排列（Permutation）的問題。然而，習知盲訊號分離所需之限制條件無法針對目標訊號的空間變化而據以調整，而使分離目標訊號的效能降低。因此，習知技術實有改善之必要。

【發明內容】

【0003】 本發明之主要目的在於提供一種隨目標訊號的空間變化而據以調整

的聲源分離方法與裝置，以改善習知技術的缺點。

**【0004】** 本發明揭露一種聲源分離方法，用來自複數個接收訊號中進行聲源分離，該聲源分離方法包含有將一分離矩陣施加於該複數個接收訊號，以產生複數個分離結果；對該複數個分離結果進行一辨識運算，以產生複數個辨識分數，其中該複數個辨識分數相關於該複數個分離結果與一目標訊號之間的匹配程度；根據該複數個辨識分數，產生一限制條件，其中該限制條件為一空間限制或一遮罩限制；以及根據該限制條件，調整該分離矩陣；其中，調整後之該分離矩陣施加於該複數個接收訊號，以自該接收訊號中分離出更新之該複數個分離結果。

**【0005】** 本發明另揭露一種聲源分離裝置，用來自複數個接收訊號中進行聲源分離，該聲源分離裝置包含有一分離單元，用來將一分離矩陣施加於該複數個接收訊號，以產生複數個分離結果；一辨識單元，用來對該複數個分離結果進行一辨識運算，以產生複數個辨識分數，其中該複數個辨識分數相關於該複數個分離結果與一目標訊號之間的匹配程度；一限制產生器，用來根據該複數個辨識分數，產生一限制條件，其中該限制條件為一空間限制或一遮罩限制；以及一分離矩陣產生器，用來根據該限制條件，調整該分離矩陣；其中，調整後之該分離矩陣施加於該複數個接收訊號，以自該接收訊號中分離出更新之該複數個分離結果。

#### **【圖式簡單說明】**

#### **【0006】**

第1圖為本發明實施例一聲源分離裝置之示意圖。

第2圖為本發明實施例一聲源分離流程之示意圖。

第3圖為本發明實施例一限制產生器之示意圖。

第4圖為本發明實施例一更新控制器之示意圖。

第 2 頁，共 12 頁(發明說明書)

第5圖為本發明實施例一空間限制產生流程之示意圖。

第6圖為本發明實施例一限制產生器之示意圖。

第7圖為本發明實施例一更新控制器之示意圖。

第8圖為本發明實施例一遮罩限制產生流程之示意圖。

第9圖為本發明實施例一聲源分離裝置之示意圖。

第10圖為本發明實施例一辨識單元之示意圖。

### 【實施方式】

【0007】 請參考第1圖，第1圖為本發明實施例一聲源分離裝置1之示意圖。聲源分離裝置1可為一特殊應用積體電路（Application-Specific-Integrated-Circuit，ASIC），用來自接收訊號 $x_1 \sim x_M$ 中分離出聲源 $z_1 \sim z_M$ ，其中目標訊號 $s_1 \sim s_N$ 可為語音訊號而存在於聲源 $z_1 \sim z_M$ 之中。聲源 $z_1 \sim z_M$ 的型態有諸多種可能，舉例來說，聲源 $z_1 \sim z_M$ 可為背景噪聲、回音、干擾或是來自語者之語音等，目標訊號 $s_1 \sim s_N$ 於本發明之實施例可為一目標語者之語音或是特定之語音內容，故於一具有聲源 $z_1 \sim z_M$ 的環境當中，目標訊號 $s_1 \sim s_N$ 並非總是存在。為了更容易理解本發明，以下說明將假設以單一目標訊號 $s_n$ 為例。聲源分離裝置1可應用於語音辨識或語者辨識，其包含接收器 $R_1 \sim R_M$ 、一分離單元10、一辨識單元12、一限制產生器14以及一分離矩陣（Demixing Matrix）產生器16。接收器 $R_1 \sim R_M$ 可為麥克風，其可將所接收之接收訊號 $x_1 \sim x_M$ 輸入至分離單元10，其中接收訊號 $x_1 \sim x_M$ 可表示為一接收訊號組 $\mathbf{x}$ ，即 $\mathbf{x} = [x_1, \dots, x_M]^T$ 。分離單元10耦接於分離矩陣產生器16，分離單元10用來將接收訊號組 $\mathbf{x}$ 乘以分離矩陣產生器16所產生之一分離矩陣 $\mathbf{W}$ ，以產生一分離結果組 $\mathbf{y}$ ，分離結果組 $\mathbf{y}$ 包含分離結果 $y_1 \sim y_M$ ，即 $\mathbf{y} = [y_1, \dots, y_M]^T = \mathbf{W}\mathbf{x}$ ，其中分離結果 $y_1 \sim y_M$ 係為自接收訊號 $x_1 \sim x_M$ 中所分離出對應於聲源 $z_1 \sim z_M$ 之分離結果。辨識單元12用來對分離結果 $y_1 \sim y_M$ 分別進行一辨識運算，以產生對應於目標

訊號 $s_n$ 相似度之辨識分數（Recognition Scores） $q_1 \sim q_M$ ，並將辨識分數 $q_1 \sim q_M$ 輸入至限制產生器14，其中辨識分數 $q_m$ 越高代表分離結果 $y_m$ 與目標訊號 $s_n$ 之間的匹配程度（或相似程度）越高。限制產生器14可根據辨識分數 $q_1 \sim q_M$ 產生一限制條件CT，用來做為指向空間中一特定方向的一控制訊號，並將限制條件CT輸入至分離矩陣產生器16。分離矩陣產生器16即可根據限制條件CT產生新的分離矩陣 $\mathbf{W}$ （即調整分離矩陣 $\mathbf{W}$ ），調整後的分離矩陣 $\mathbf{W}$ 即可施加於接收訊號 $x_1 \sim x_M$ ，以分離出聲源 $z_1 \sim z_M$ 。上述分離矩陣產生器16，依據一實施例，可透過具有限制條件的盲訊號分離（Constrained-Blind-Source-Separation, CBSS）方法來產生該分離矩陣 $\mathbf{W}$ 。

**【0008】** 辨識單元12可包含一特徵擷取器26、一參考模型訓練器22及一匹配器（Matcher）24，如第10圖所示。特徵擷取器20可根據分離結果 $y_1 \sim y_M$ 分別產生一組特徵訊號 $b_1 \sim b_M$ ，以語音辨識之應用為例，特徵擷取器20所擷取之特徵可為梅爾倒頻譜係數（Mel-Frequency-Cepstral-Coefficients, MFCC）。當一訓練旗標FG指示辨識單元12處於一訓練階段時，特徵擷取器20自擷取分離結果 $y_1 \sim y_M$ 相關於目標訊號 $s_n$ 的特徵，並輸入至參考模型訓練器22，以產生目標訊號 $s_n$ 的參考模型；而當訓練旗標FG指示辨識單元12處於一測試階段時，匹配器24將自分離結果 $y_1 \sim y_M$ 擷取出之特徵與參考模型進行比對，以產生對應之辨識分數 $q_1 \sim q_M$ 。換言之，參考模型訓練器22可以預先訓練出對應於目標訊號 $s_n$ 的一參考模型，接著，匹配器24再將該參考模型與特徵擷取器20輸出的特徵訊號 $b_1 \sim b_M$ 分別進行比對，比較其相似程度，並輸出辨識分數 $q_1 \sim q_M$ 。其餘關於辨識單元12之操作細節為本領域具通常知識者所熟知，故不在此贅述。

**【0009】** 簡而言之，因辨識分數 $q_1 \sim q_M$ 會隨目標訊號相對於接收器 $R_1 \sim R_M$ 的

第4頁，共12頁(發明說明書)

空間特性不斷地改變，聲源分離裝置1可根據辨識單元12於不同時間所產生的辨識分數 $q_1 \sim q_M$ ，產生不同的限制條件CT以做為指向某空間方向的一控制訊號，並根據更新後的限制條件CT調整分離矩陣 $\mathbf{W}$ ，以分離聲源 $z_1 \sim z_M$ ，並獲得更新的分離結果 $y_1 \sim y_M$ 。如此一來，聲源分離裝置1所產生的限制條件CT及分離矩陣 $\mathbf{W}$ 皆可針對目標訊號的空間變化而據以調整，進而增進目標訊號分離的效能。關於聲源分離裝置1的操作流程可歸納成為一聲源分離流程20，如第2圖所示，聲源分離流程20包含以下步驟：

【0010】 步驟200：將分離矩陣 $\mathbf{W}$ 施加於接收訊號 $x_1 \sim x_M$ ，以產生分離結果 $y_1 \sim y_M$ 。

【0011】 步驟202：對分離結果 $y_1 \sim y_M$ 進行辨識運算，以產生對應於目標訊號 $s_n$ 之辨識分數 $q_1 \sim q_M$ 。

【0012】 步驟204：根據對應於目標訊號 $s_n$ 之辨識分數 $q_1 \sim q_M$ ，產生限制條件CT。

【0013】 步驟206：根據限制條件CT，調整分離矩陣 $\mathbf{W}$ 。

【0014】 於一實施例中，限制產生器14可產生限制條件CT為一空間限制（Spatial Constraint） $\mathbf{c}$ ，而分離矩陣產生器16可根據空間限制 $\mathbf{c}$ 產生新的分離矩陣 $\mathbf{W}$ 。空間限制 $\mathbf{c}$ 可用來限制分離矩陣 $\mathbf{W}$ 對空間中一特定方向的響應，即限制分離矩陣 $\mathbf{W}$ 於該特定方向具有空間濾波（Spatial Filtering）效能。分離矩陣產生器16根據空間限制 $\mathbf{c}$ 產生分離矩陣 $\mathbf{W}$ 的方式並未有所限，例如，分離矩陣產生器16可產生分離矩陣 $\mathbf{W}$ 使其滿足 $\mathbf{w}_m^H \mathbf{c} = c_1$ ，其中 $c_1$ 可為任意常數，而 $\mathbf{w}_m^H$ 代表分離矩陣 $\mathbf{W}$ 之一列向量（即分離矩陣 $\mathbf{W}$ 可表示為 $\mathbf{W}^H = \begin{bmatrix} \mathbf{w}_1^H \\ \vdots \\ \mathbf{w}_M^H \end{bmatrix}$ ）。

【0015】 關於限制產生器產生空間限制 $\mathbf{c}$ 的具體實現方式，請參考第3圖及第4圖，第3圖及第4圖分別為本發明實施例一限制產生器34及一更新控制器342之示意圖。限制產生器34可根據分離矩陣 $\mathbf{W}$ 及辨識分數 $q_1 \sim q_M$ 產生空間限制 $\mathbf{c}$ ，其包含更新控制器342、一反矩陣轉換單元30及一平均單元36，更新控制器342包含一映射（Mapping）單元40、一正規化（Normalization）單元42、一最大值選取器44及一權重組合單元46。反矩陣轉換單元30可耦接於分離矩陣產生器16以接收分離矩陣 $\mathbf{W}$ ，並將分離矩陣 $\mathbf{W}$ 進行一反矩陣運算，以產生一估測混合矩陣（Estimated Mixing Matrix） $\mathbf{W}^{-1}$ ，更新控制器342根據估測混合矩陣 $\mathbf{W}^{-1}$ 及辨識分數 $q_1 \sim q_M$ 產生一更新速率 $\alpha$ 及一更新係數 $\mathbf{c}_{\text{update}}$ ，而平均單元36根據更新速率 $\alpha$ 及更新係數 $\mathbf{c}_{\text{update}}$ 產生空間限制 $\mathbf{c}$ 。

【0016】 詳細來說，估測混合矩陣 $\mathbf{W}^{-1}$ 可代表一混合矩陣（Mixing Matrix） $\mathbf{H}$ 之估測值，混合矩陣 $\mathbf{H}$ 用來表示聲源 $z_1 \sim z_M$ 與接收訊號 $x_1 \sim x_M$ 之間的對應關係，即 $\mathbf{x} = \mathbf{H}\mathbf{z}$ ，其中 $\mathbf{z} = [z_1, \dots, z_M]^T$ 。混合矩陣 $\mathbf{H}$ 包含指向向量（Steering Vector） $\mathbf{h}_1 \sim \mathbf{h}_M$ ，即 $\mathbf{H} = [\mathbf{h}_1 \dots \mathbf{h}_M]$ ，換句話說，估測混合矩陣 $\mathbf{W}^{-1}$ 包含估測指向向量（Estimated Steering Vector） $\hat{\mathbf{h}}_1 \sim \hat{\mathbf{h}}_M$ ，且可表示為 $\mathbf{W}^{-1} = [\hat{\mathbf{h}}_1 \dots \hat{\mathbf{h}}_M]$ 。另外，更新控制器342可根據辨識分數 $q_1 \sim q_M$ 產生權重值 $\omega_1 \sim \omega_M$ ，並產生更新係數 $\mathbf{c}_{\text{update}}$ 為 $\mathbf{c}_{\text{update}} = \sum_{m=1}^M \omega_m \hat{\mathbf{h}}_m$ 。另外，更新控制器342可透過映射單元40對辨識分數 $q_1 \sim q_M$ 進行一映射運算，即將辨識分數 $q_1 \sim q_M$ 以線性或非線性的方式，映射至0到1之間的數值，以產生對應於辨識分數 $q_1 \sim q_M$ 之映射值 $\tilde{q}_1 \sim \tilde{q}_M$ （映射值 $\tilde{q}_1 \sim \tilde{q}_M$ 皆介於0到1之間），並透過正規化單元42對映射值 $\tilde{q}_1 \sim \tilde{q}_M$ 進行一正規化運算，以產生權重值 $\omega_1 \sim \omega_M$ （即 $\omega_m = \tilde{q}_m / \sum_{n=1}^M \tilde{q}_n$ ）。另一方面，更新控制器342可透過最大值選取器44產生更新速率 $\alpha$ 為映射值 $\tilde{q}_1 \sim \tilde{q}_M$ 之一最大值，即 $\alpha = \max_m \tilde{q}_m$ 。如此一來，更新控



制器342即可輸出更新速率  $\alpha$  及更新係數  $\mathbf{c}_{\text{update}}$  至平均單元36，而平均單元36可計算空間限制  $\mathbf{c}$  為  $\mathbf{c} = (1 - \alpha)\mathbf{c} + \alpha\mathbf{c}_{\text{update}}$ 。限制產生器34將空間限制  $\mathbf{c}$  輸出至分離矩陣產生器16，分離矩陣產生器16即可根據空間限制  $\mathbf{c}$  產生新的分離矩陣  $\mathbf{W}$ ，以分離出聲源  $z_1 \sim z_M$ 。

【0017】 關於限制產生器34的操作流程可歸納成為一空間限制產生流程50，如第5圖所示，空間限制產生流程50包含以下步驟：

【0018】 步驟500：對分離矩陣  $\mathbf{W}$  進行反矩陣運算，以產生估測混合矩陣  $\mathbf{W}^{-1}$ ，其中估測混合矩陣  $\mathbf{W}^{-1}$  包含估測指向向量  $\hat{\mathbf{h}}_1 \sim \hat{\mathbf{h}}_M$ 。

【0019】 步驟502：根據辨識分數  $q_1 \sim q_M$ ，產生權重值  $\omega_1 \sim \omega_M$ 。

【0020】 步驟504：根據辨識分數  $q_1 \sim q_M$ ，產生更新速率  $\alpha$ 。

【0021】 步驟506：根據權重值  $\omega_1 \sim \omega_M$  及估測指向向量  $\hat{\mathbf{h}}_1 \sim \hat{\mathbf{h}}_M$ ，產生更新係數  $\mathbf{c}_{\text{update}}$ 。

【0022】 步驟508：根據更新速率  $\alpha$  及更新係數  $\mathbf{c}_{\text{update}}$ ，產生空間限制  $\mathbf{c}$ 。

【0023】 於另一實施例中，限制產生器14可產生限制條件CT為一遮罩限制 (Mask Constraint)  $\Lambda$ ，而分離矩陣產生器16可根據遮罩限制  $\Lambda$  產生新的分離矩陣  $\mathbf{W}$ 。遮罩限制  $\Lambda$  可用來限制分離矩陣  $\mathbf{W}$  對目標訊號的響應，即對該目標訊號產生遮罩效能。分離矩陣產生器16根據遮罩限制  $\Lambda$  產生分離矩陣  $\mathbf{W}$  的方式並未有所限，例如，分離矩陣產生器16可利用一遞迴 (Recursive) 演算法 (如牛頓法(Newton Method)、梯度法(Gradient Method)等) 估算聲源  $z_1 \sim z_M$  與接收訊號  $x_1 \sim x_M$  之間之混合矩陣  $\mathbf{H}$  之估測值，而利用遮罩限制  $\Lambda$  限制遞迴演算法中每一次迭代之變化量，換句話說，第  $k+1$  次迭代所計算出混合矩陣之估測值  $\hat{\mathbf{H}}_{k+1}$  可表示為  $\hat{\mathbf{H}}_{k+1} = \hat{\mathbf{H}}_k + \Delta\mathbf{H} \cdot \Lambda$ ，其中，分離矩陣產生器16可產生分離矩陣  $\mathbf{W}$  為  $\mathbf{W} = \hat{\mathbf{H}}_{k+1}^{-1}$ ，其

中  $\Delta\mathbf{H}$  可因不同遞迴演算法而異。另外，遮罩限制  $\Lambda$  可為一對角矩陣 (Diagonal Matrix)，其可對聲源  $z_1 \sim z_M$  中第  $n^*$  個聲源  $z_{n^*}$  進行遮罩運作 (聲源  $z_{n^*}$  即為目標訊號  $s_n$ )，具體來說，限制產生器 14 可令遮罩限制  $\Lambda$  之第  $n^*$  個對角元素為介於 0 到 1 之間之一特定值  $G$ ，而其餘對角元素的值為  $1 - G$ ，即遮罩限制  $\Lambda$  之第  $i$  個對角元素  $[\Lambda]_{i,i}$  可表示為  $[\Lambda]_{i,i} = \begin{cases} G, & i = n^* \\ 1 - G, & i \neq n^* \end{cases}$ 。

**【0024】** 關於限制產生器產生遮罩限制  $\Lambda$  的具體實現方式，請參考第 6 圖及第 7 圖，第 6 圖及第 7 圖分別為本發明實施例一限制產生器 64 及一更新控制器 642 之示意圖。限制產生器 64 可根據分離結果  $y_1 \sim y_M$  及辨識分數  $q_1 \sim q_M$  產生遮罩限制  $\Lambda$ ，其包含更新控制器 642、一能量單元 60、一加權能量產生器 62、一參考能量產生器 68 及一遮罩產生器 66，其中更新控制器 642 包含一映射單元 70、一正規化單元 72 及一轉換單元 74。能量單元 60 接收分離結果  $y_1 \sim y_M$  以計算對應於分離結果  $y_1 \sim y_M$  (亦對應於聲源  $z_1 \sim z_M$ ) 之聲源能量  $P_1 \sim P_M$ ，更新控制器 642 根據辨識分數  $q_1 \sim q_M$  產生權重值  $\omega_1 \sim \omega_M$  及權重值  $\beta_1 \sim \beta_M$ ，加權能量產生器 62 根據權重值  $\omega_1 \sim \omega_M$  及聲源能量  $P_1 \sim P_M$  產生一加權能量  $P_{wei}$ ，參考能量產生器 68 根據權重值  $\beta_1 \sim \beta_M$  及聲源能量  $P_1 \sim P_M$  產生一參考能量  $P_{ref}$ ，遮罩產生器 66 根據權重值  $\omega_1 \sim \omega_M$ 、加權能量  $P_{wei}$  及參考能量  $P_{ref}$  產生遮罩限制  $\Lambda$ 。

**【0025】** 詳細來說，加權能量產生器 62 可產生加權能量  $P_{wei}$  為  $P_{wei} = \sum_{m=1}^M \omega_m P_m$ ，而參考能量產生器 68 可產生參考能量  $P_{ref}$  為  $P_{ref} = \sum_{m=1}^M \beta_m P_m$ ，更新控制器 642 所包含之映射單元 70 及正規化單元 72 可與映射單元 40 及正規化單元 42 相同，故關於其操作細節不再贅述。另外，轉換單元 74 可將權重值  $\omega_1 \sim \omega_M$  轉換為權重值  $\beta_1 \sim \beta_M$ ，轉換單元 74 不限於特定方式產生權重值  $\beta_1 \sim \beta_M$ ，例如，轉換單元 74 可產生權重值

$\beta_m$  為  $\beta_m = 1 - \omega_m$ ，而不在此限。

【0026】 另一方面，遮罩產生器66可根據加權能量 $P_{wei}$ 及參考能量 $P_{ref}$ 產生遮罩限制 $\Lambda$ 所需之特定值 $G$ ，例如，遮罩產生器66可計算特定值 $G$ 為  $G = \begin{cases} 1, & P_{wei} > \gamma P_{ref} \\ 0, & P_{wei} \leq \gamma P_{ref} \end{cases}$ ，

其中比例 $\gamma$ 可視實際情況而調整，另外，遮罩產生器66可計算特定值 $G$ 為  $G = P_{wei} / P_{ref}$  或  $G = P_{wei} / (P_{ref} + P_{wei})$ ，而不在此限。另外，遮罩產生器66可根據權重值 $\omega_1 \sim \omega_M$ （即根據辨識分數 $q_1 \sim q_M$ ）決定目標訊號的目標指標 $n^*$ ，例如，遮罩產生器66可決定目標指標 $n^*$ 為對應於權重值 $\omega_1 \sim \omega_M$ 中一最大權重值之一權重指標，即決定目標指標 $n^*$ 為  $n^* = \arg \max_m \omega_m$ 。如此一來，遮罩產生器66取得特定值 $G$

以及目標指標 $n^*$ 後，即可產生遮罩限制 $\Lambda$ 為  $[\Lambda]_{i,i} = \begin{cases} G, & i = n^* \\ 1 - G, & i \neq n^* \end{cases}$ 。限制產生器64

將遮罩限制 $\Lambda$ 輸出至分離矩陣產生器16，分離矩陣產生器16即可遮罩限制 $\Lambda$ 產生新的分離矩陣 $W$ ，以分離出聲源 $z_1 \sim z_M$ 。

【0027】 關於限制產生器64的操作流程可歸納成為一遮罩限制產生流程80，如第8圖所示，遮罩限制產生流程80包含以下步驟：

【0028】 步驟800：根據分離結果 $y_1 \sim y_M$ 計算對應於聲源 $z_1 \sim z_M$ 之聲源能量 $P_1 \sim P_M$ 。

【0029】 步驟802：根據辨識分數 $q_1 \sim q_M$ ，產生權重值 $\omega_1 \sim \omega_M$ 及權重值 $\beta_1 \sim \beta_M$ 。

【0030】 步驟804：根據聲源能量 $P_1 \sim P_M$ 及權重值 $\omega_1 \sim \omega_M$ ，產生加權能量 $P_{wei}$ 。

【0031】 步驟806：根據聲源能量 $P_1 \sim P_M$ 及權重值 $\beta_1 \sim \beta_M$ ，產生參考能量 $P_{ref}$ 。

【0032】 步驟808：根據加權能量 $P_{wei}$ 及參考能量 $P_{ref}$ ，產生特定值 $G$ 。

【0033】 步驟810：根據權重值 $\omega_1 \sim \omega_M$ ，決定目標指標 $n^*$ 。

【0034】 步驟812：根據特定值 $G$ 及目標指標 $n^*$ ，產生遮罩限制 $\Lambda$ 。

【0035】 另一方面，聲源分離裝置不限於以特殊應用積體電路來實現，請參考第9圖，第9圖為本發明實施例一聲源分離裝置90之示意圖，聲源分離裝置90包含一處理單元902及一儲存單元904。前述聲源分離流程20、空間限制產生流程50、遮罩限制產生流程80可編譯成一程式碼908並儲存於儲存單元904中，以指示處理單元902執行流程20、50、80。其中，處理單元902可為一數位訊號處理器（DSP），而不在此限。儲存單元904可為一非揮發性記憶體（Non-Volatile Memory，例如，一電子抹除式可複寫唯讀記憶體（Electrically Erasable Programmable Read Only Memory, EEPROM）或一快閃記憶體（Flash Memory）），而不在此限。

【0036】 此外，為了易於理解，於上述實施例中，聲源 $z$ 、目標訊號 $s$ 、接收器 $R$ 或者其他種類的輸出訊號（如聲源能量 $P$ 、辨識分數 $q$ 、分離結果 $y$ 等）皆用 $M$ 來表示其為多數個，但本發明並不限制相互之間的數目必然相同，亦可有不同數目的應用。例如：假設接收器 $R$ 數目有2個，聲源 $z$ 數目可能為4個，目標訊號 $s$ 數目可能為1個。

【0037】 綜上所述，本發明可根據辨識分數更新限制條件，並根據更新後的限制條件調整分離矩陣，其可針對目標訊號的空間變化而據以調整，進而增進聲源分離的效能。

以上所述僅為本發明之較佳實施例，凡依本發明申請專利範圍所做之均等變化與修飾，皆應屬本發明之涵蓋範圍。

【符號說明】

## 【0038】

1	聲源分離裝置
20、50、80	流程
200~206、500~508、800~812	步驟
10	分離單元
12	辨識單元
14	限制產生器
16	分離矩陣產生器
22	參考模型訓練器
24	匹配器
26	特徵擷取器
34、64	限制產生器
342、642	更新控制器
30	反矩陣轉換單元
36	平均單元
40	映射單元
42	正規化單元
44	最大值選取器
46	權重組合單元
60	能量單元
62	加權能量產生器
68	參考能量產生器
66	遮罩產生器
70	映射單元

72	正規化單元
74	轉換單元
902	處理單元
904	儲存單元
908	程式碼
<b>c</b>	空間限制
<b>CT</b>	限制條件
$\mathbf{c}_{\text{update}}$	更新係數
$P_1 \sim P_M$	聲源能量
$P_{\text{wei}}$	加權能量
$P_{\text{ref}}$	參考能量
$q_1 \sim q_M$	辨識分數
$b_1 \sim b_M$	特徵訊號
$R_1 \sim R_M$	接收器
$x_1 \sim x_M$	接收訊號
<b>x</b>	接收訊號組
$y_1 \sim y_M$	分離結果
<b>W</b>	分離矩陣
$\alpha$	更新速率
$\beta_1 \sim \beta_M, \omega_1 \sim \omega_M$	權重值
<b><math>\Lambda</math></b>	遮罩限制

## 【發明申請專利範圍】

【第1項】 一種聲源分離方法，用來自複數個接收訊號中進行聲源分離，該聲源分離方法包含有：

將一分離矩陣施加於該複數個接收訊號，以產生複數個分離結果；

對該複數個分離結果進行一辨識運算，以產生複數個辨識分數，其中該複數個辨識分數相對於該複數個分離結果與一目標訊號之間的匹配程度；

根據該複數個辨識分數，產生一限制條件，其中該限制條件為一空間限制或一遮罩限制；以及

根據該限制條件，調整該分離矩陣；

其中，調整後之該分離矩陣施加於該複數個接收訊號，以自該接收訊號中分離出更新之該複數個分離結果。

【第2項】 如請求項1所述之聲源分離方法，其中對該複數個分離結果進行該辨識運算，以產生該複數個辨識分數的步驟包含有：

建立對應於該目標訊號之一參考模型；

擷取該分離結果之特徵；以及

將該複數個分離結果之特徵與該參考模型進行比對，以產生該複數個辨識分數。

【第3項】 如請求項1所述之聲源分離方法，其中根據該複數個辨識分數，產生該空間限制的步驟包含有：

根據該複數個辨識分數，產生複數個第一權重值；

根據該複數個辨識分數，產生一更新速率；

第 1 頁，共 4 頁(發明申請專利範圍)

根據該分離矩陣及該複數個第一權重值，產生一更新係數；以及  
根據該更新係數及該更新速率，產生該空間限制。

【第4項】 如請求項3所述之聲源分離方法，其中根據該複數個辨識分數，產生該複數個第一權重值的步驟包含有：

對該複數個辨識分數進行一映射運算，以取得複數個映射值；以及  
對該複數個映射值進行一正規化運算，以取得該複數個第一權重值。

【第5項】 如請求項4所述之聲源分離方法，其中根據該複數個辨識分數，產生該更新速率的步驟包含有：

取得該更新速率為該複數個映射值之一最大值。

【第6項】 如請求項3所述之聲源分離方法，其中根據該分離矩陣及該複數個第一權重值，產生該更新係數的步驟包含有：

對該分離矩陣進行一反矩陣運算，以產生複數個估測指向向量；以及  
根據該複數個估測指向向量及該複數個第一權重值，產生該更新係數。

【第7項】 如請求項3所述之聲源分離方法，其中根據該更新係數及該更新速率，產生該空間限制的步驟包含有：

執行  $\mathbf{c} = (1 - \alpha)\mathbf{c} + \alpha\mathbf{c}_{\text{update}}$  ；

其中  $\mathbf{c}$  代表該空間限制， $\alpha$  代表該更新速率， $\mathbf{c}_{\text{update}}$  代表該更新係數。

【第8項】 如請求項1所述之聲源分離方法，其中根據該複數個辨識分數，產生該遮罩限制的步驟包含有：



根據該複數個辨識分數，產生複數個第一權重值；

根據該複數個第一權重值，產生複數個第二權重值；

根據該分離結果，產生複數個聲源能量；

根據該複數個聲源能量及該複數個第一權重值，產生一加權能量；

根據該複數個聲源能量及該複數個第二權重值，產生一參考能量；以及

根據該加權能量、該參考能量及該複數個第一權重值，產生該遮罩限制。

**【第9項】** 如請求項8所述之聲源分離方法，其中根據該加權能量、該參考能量及該複數個第一權重值，產生該遮罩限制的步驟包含有：

根據該加權能量及該參考能量，產生一特定值；

決定該目標指標為對應於該複數個第一權重值中一最大權重值之一權重指標；以及

根據該特定值及該目標指標，產生該遮罩限制。

**【第10項】** 一種聲源分離裝置，用來自複數個接收訊號中進行聲源分離，該聲源分離裝置包含有：

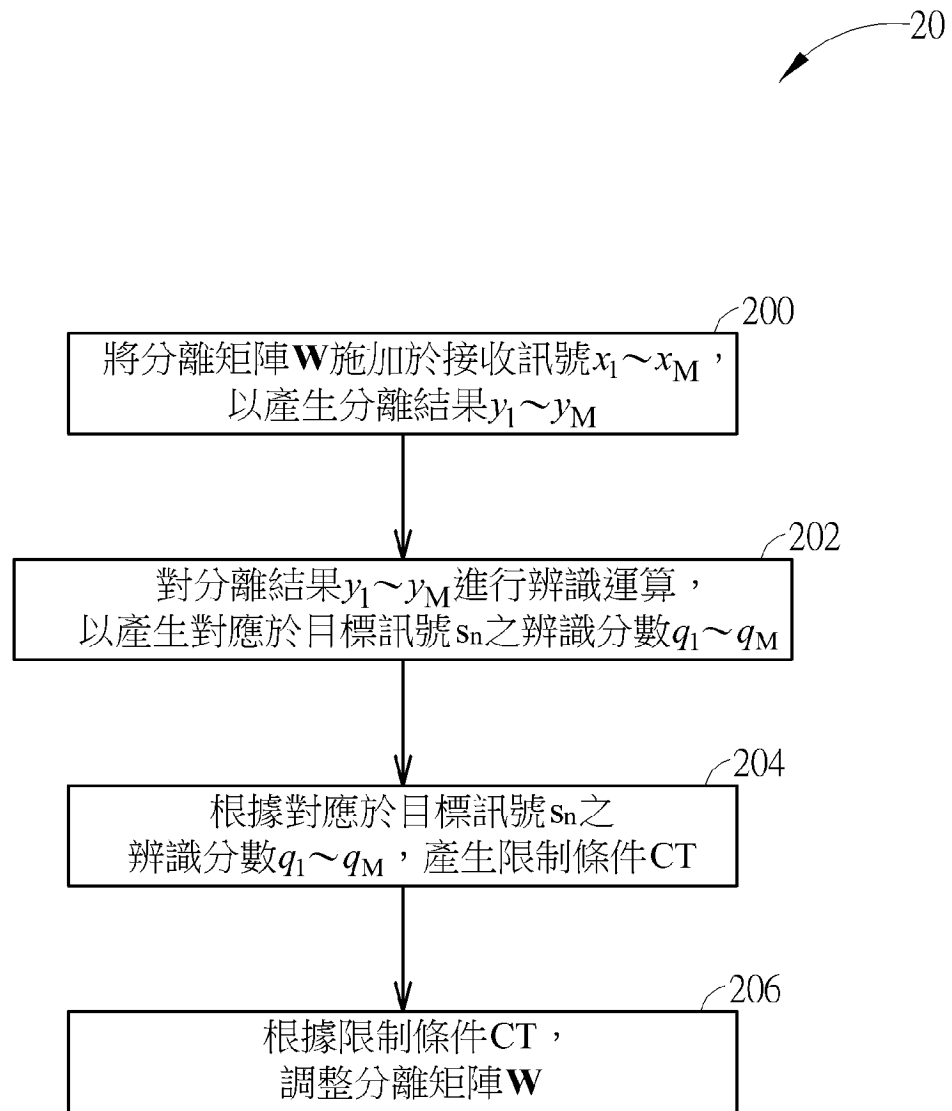
一分離單元，用來將一分離矩陣施加於該複數個接收訊號，以產生複數個分離結果；

一辨識單元，用來對該複數個分離結果進行一辨識運算，以產生複數個辨識分數，其中該複數個辨識分數相關於該複數個分離結果與一目標訊號之間的匹配程度；

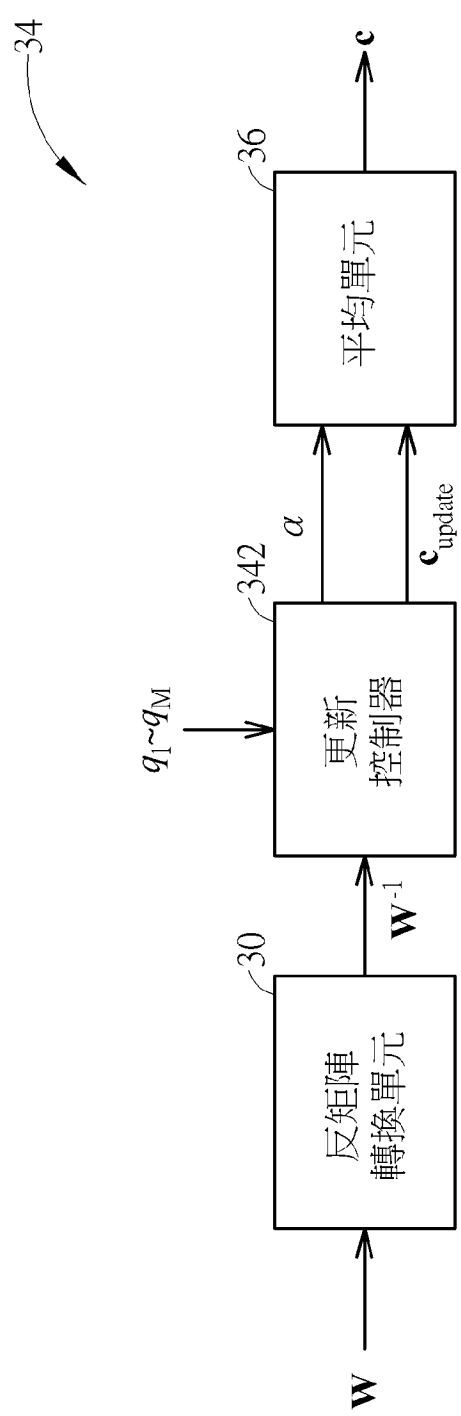
一限制產生器，用來根據該複數個辨識分數，產生一限制條件，其中該限制條件為一空間限制或一遮罩限制；以及

一分離矩陣產生器，用來根據該限制條件，調整該分離矩陣；

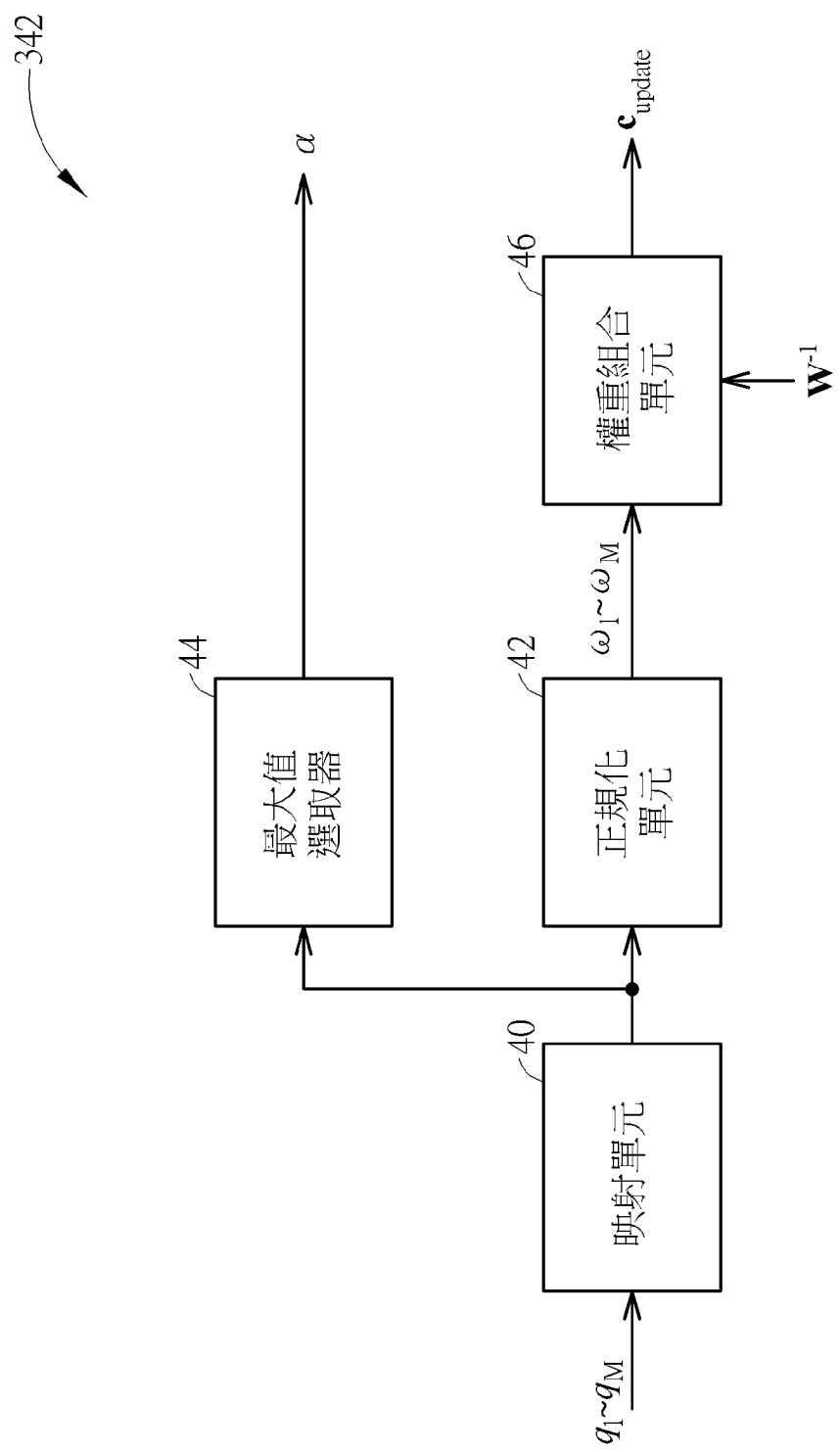
其中，調整後之該分離矩陣施加於該複數個接收訊號，以自該接收訊號中分離出更新之該複數個分離結果。



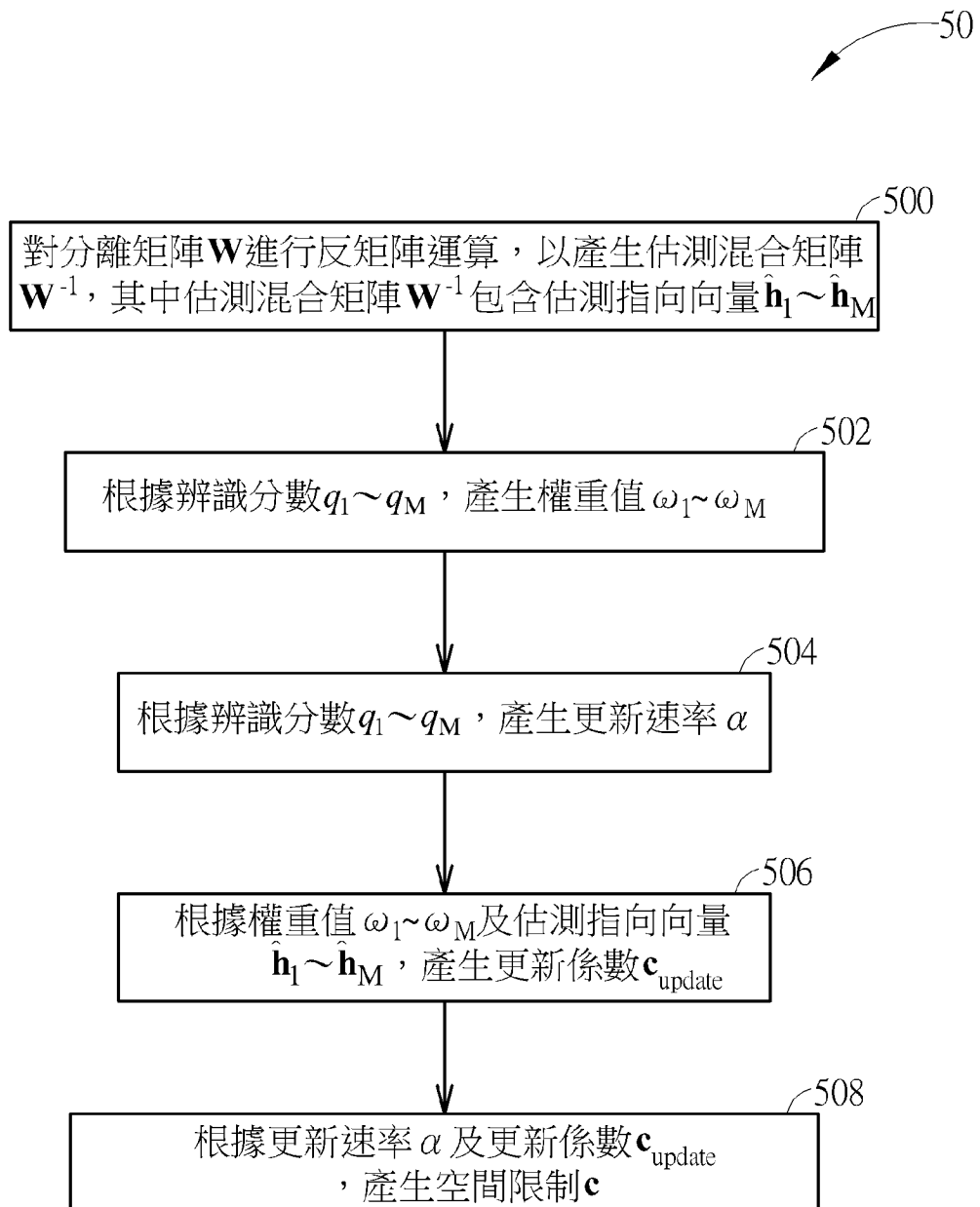
第2圖



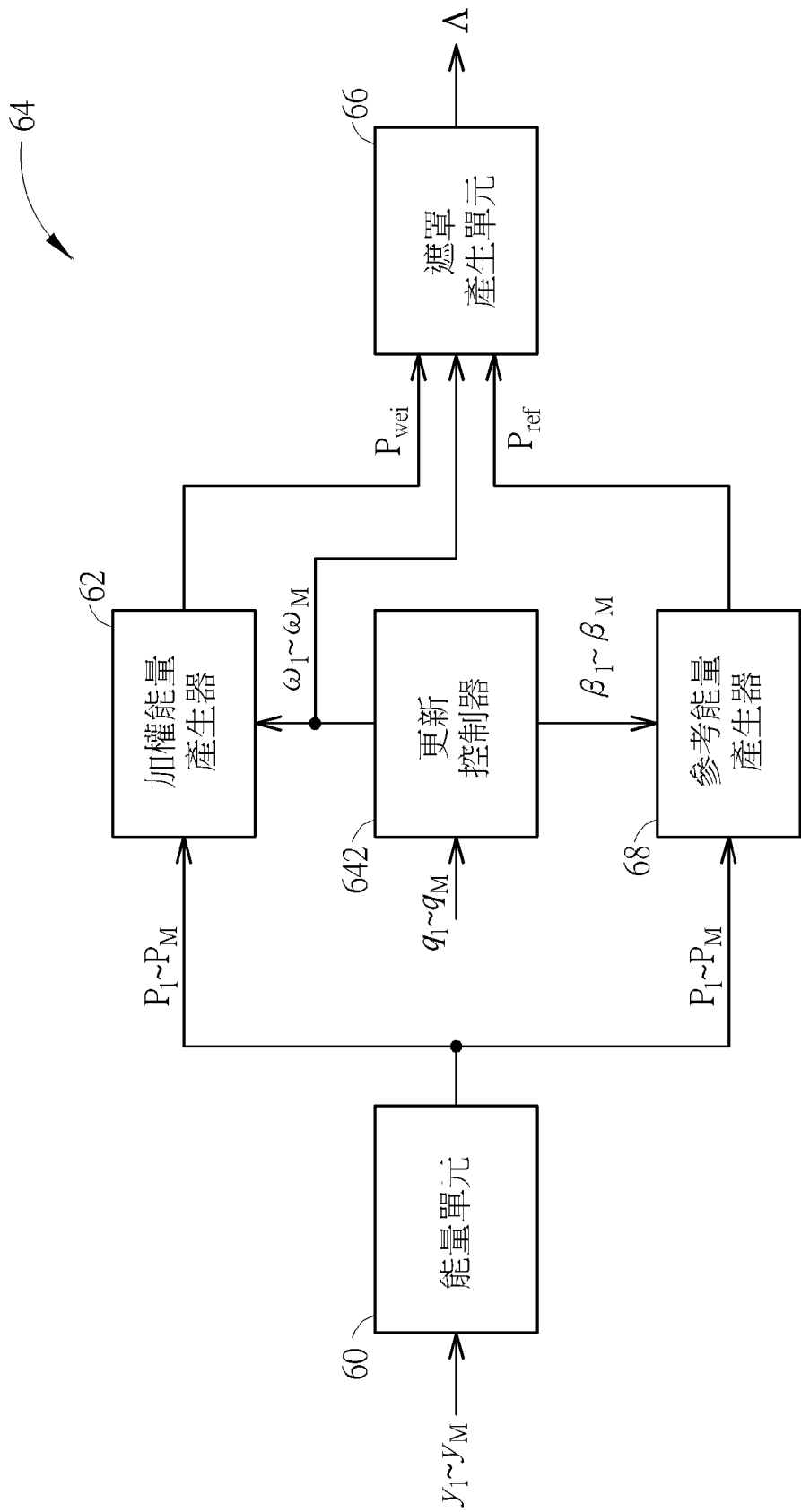
第3圖



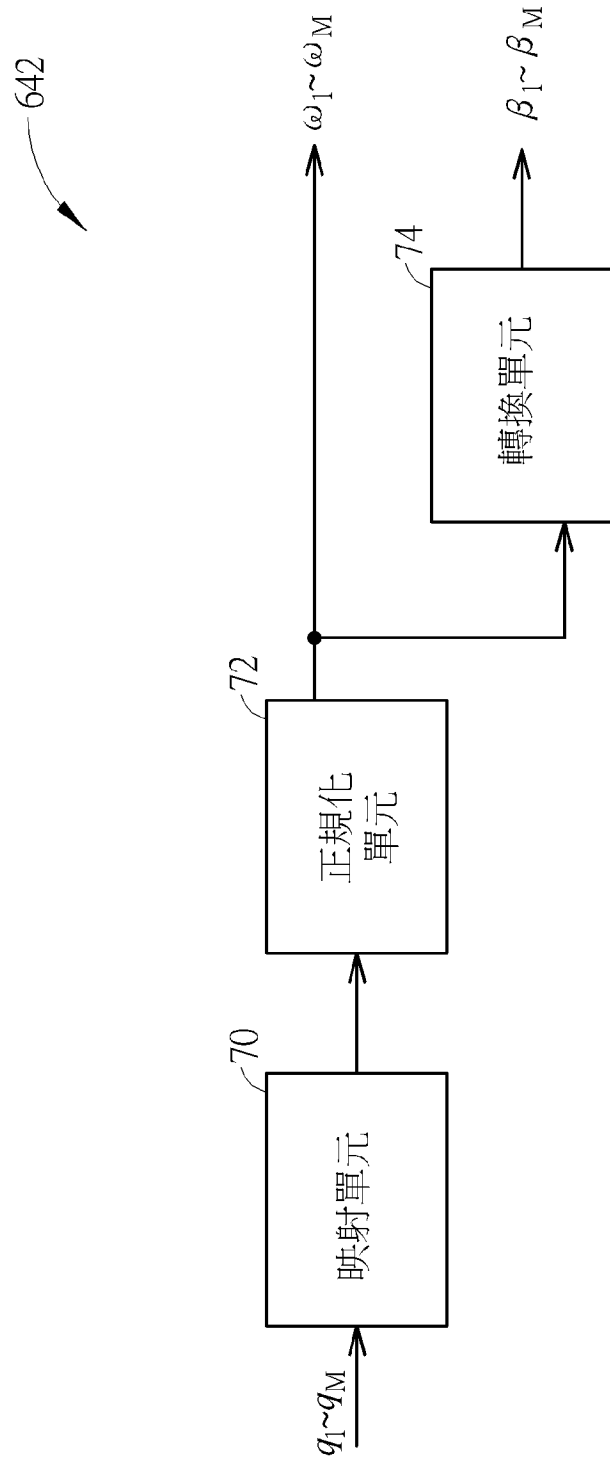
第4圖



第5圖

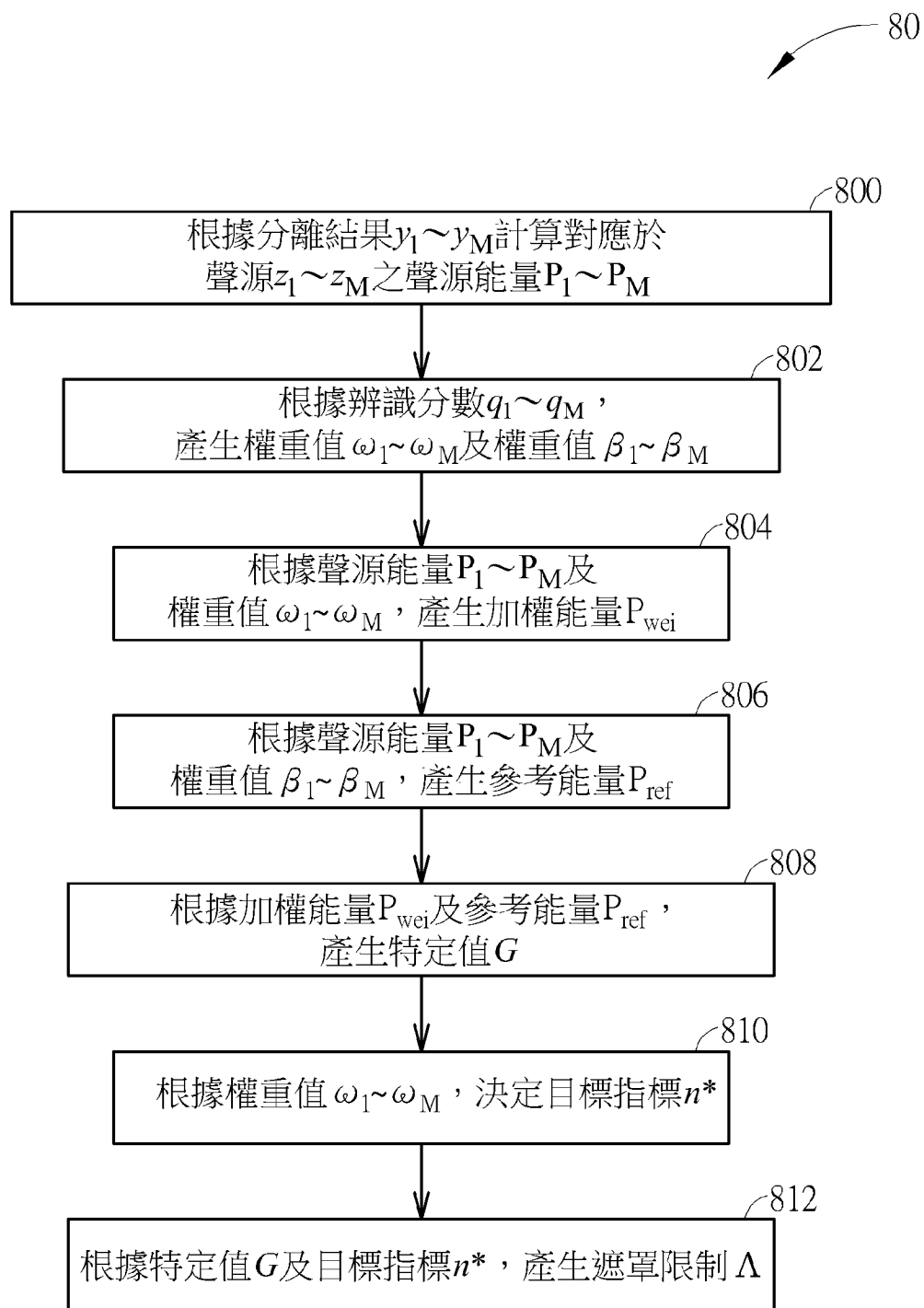


第6圖

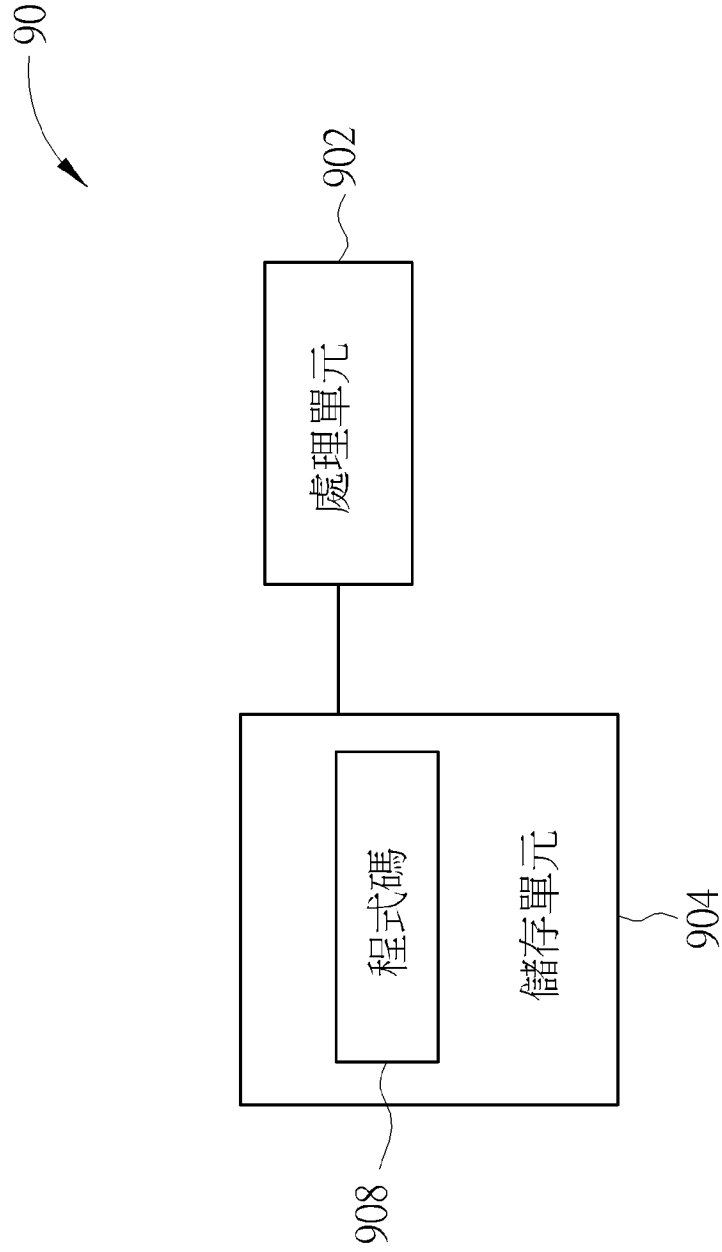


第7圖

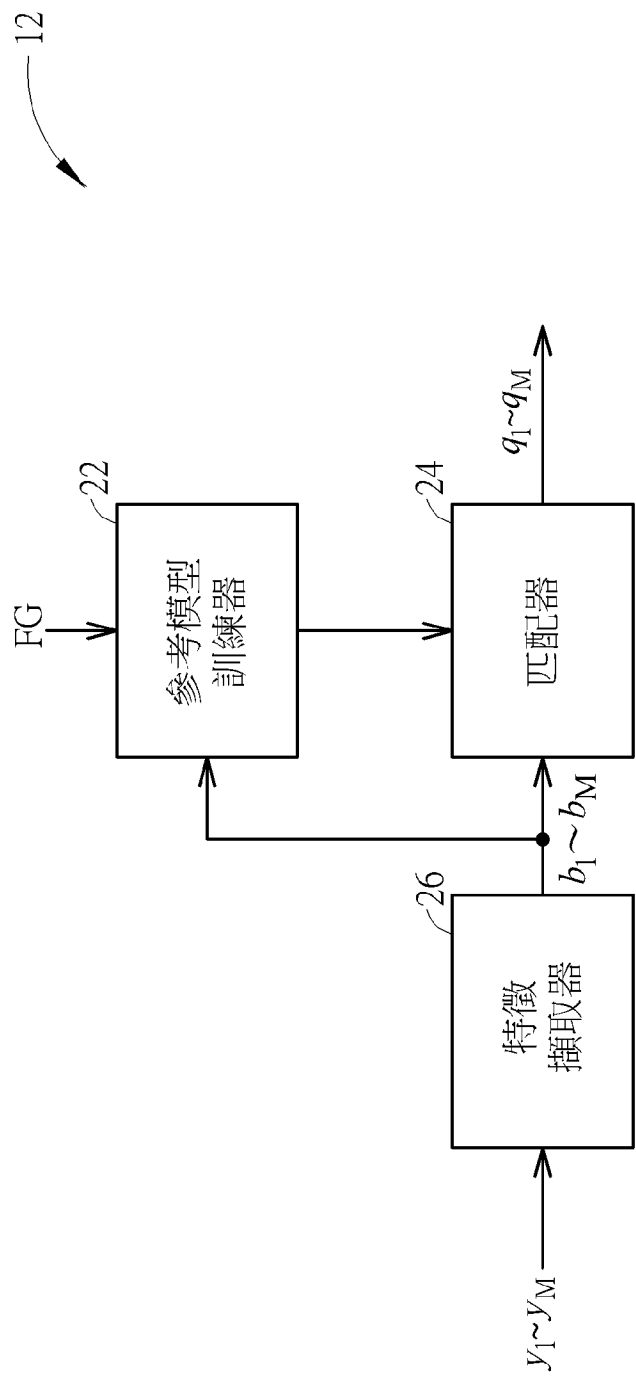




第8圖



第9圖



第10圖

【指定代表圖】第（ 2 ）圖。

【代表圖之符號簡單說明】

20

聲源分離流程

200~206

步驟

【特徵化學式】

無



## 【發明摘要】

【中文發明名稱】 聲源分離方法與裝置

【英文發明名稱】 Method and Device of Audio Source Separation

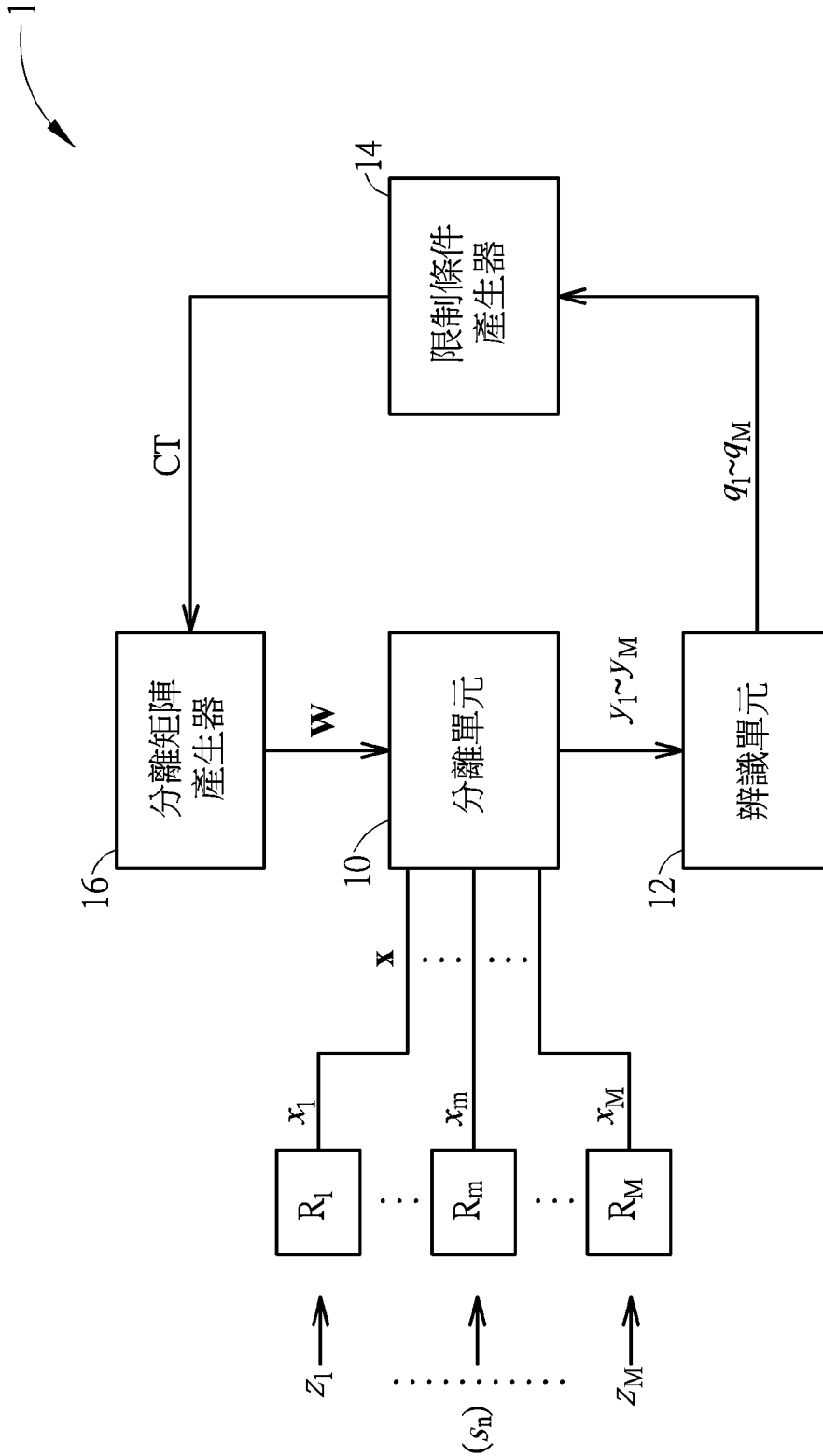
### 【中文】

一種聲源分離方法，包含有將一分離矩陣施加於複數個接收訊號，以產生複數個分離結果；對該複數個分離結果進行一辨識運算，以產生複數個辨識分數；根據該複數個辨識分數，產生一限制條件；以及根據該限制條件，調整該分離矩陣；其中，調整後之該分離矩陣施加於該複數個接收訊號，以自該接收訊號中分離出更新的複數個分離結果。

### 【英文】

A method of audio source separation includes applying a demixing matrix on a plurality of received signals to generate a plurality of separated results; performing a recognition operation on the plurality of separated results to generate a plurality of recognition scores; generating a constraint according to the plurality of recognition scores; and adjusting the demixing matrix according to the constraint; where the adjusted demixing matrix is applied to the plurality of received signals to generate a plurality of updated separated results from the plurality of received signals.

【發明圖式】



第1圖