



(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 I558228 B

(45)公告日：中華民國 105 (2016) 年 11 月 11 日

(21)申請案號：101145071

(22)申請日：中華民國 101 (2012) 年 11 月 30 日

(51)Int. Cl. : **H04R5/027 (2006.01)**

(30)優先權：2011/12/02 歐洲專利局 11191828.0
2012/04/12 美國 13/445,560

(71)申請人：弗勞恩霍夫爾協會(德國) FRAUNHOFER-GESELLSCHAFT ZUR FORDERUNG
DER ANGEWANDTEN FORSCHUNG E. V. (DE)
德國

(72)發明人：戴加多 紀凡尼 DEL GALDO, GIOVANNI (IT)；錫蓋特 奧莉薇 THIERGART,
OLIVER (DE)；庫奇 法比恩 KUCH, FABIAN (DE)；哈貝斯 艾曼紐 HABETS,
EMANUEL (NL)；克拉休 亞歷山大 CRACIUN, ALEXANDRA (RO)

(74)代理人：邱珍元

(56)參考文獻：

CN 101779476A

Giovanni Del Galdo, "Generating virtual microphone signals using
geometrical information gathered by distributed arrays", Hands-free
Speech Communication and Microphone Arrays (HSCMA), 2011 Joint
Workshop on IEEE, 30 May 2011, pages 185-190

審查人員：范美華

申請專利範圍項數：11 項 圖式數：27 共 82 頁

(54)名稱

依據空間能量密度定位麥克風之設備及方法

APPARATUS AND METHOD FOR MICROPHONE POSITIONING BASED ON A SPATIAL POWER
DENSITY

(57)摘要

提供一種定位麥克風之設備。此設備包含一空間能量分佈決定器(10)及一空間資訊估計器(20)。空間能量分佈決定器(10)係適合於依據聲音源資訊(表示位於環境中之一個或多個聲音源之一個或多個能量數值及一個或多個位置數值)來決定一空間能量密度(表示關於一環境之複數個位置之能量數值)。空間資訊估計器(20)係適合於依據空間能量密度來估計聲源空間資訊。

An apparatus for microphone positioning is provided. The apparatus comprises a spatial power distribution determiner (10) and a spatial information estimator (20). The spatial power distribution determiner (10) is adapted to determine a spatial power density indicating power values for a plurality of locations of an environment based on sound source information indicating one or more power values and one or more position values of one or more sound sources located in the environment. The spatial information estimator (20) is adapted to estimate acoustic spatial information based on the spatial power density.

指定代表圖：

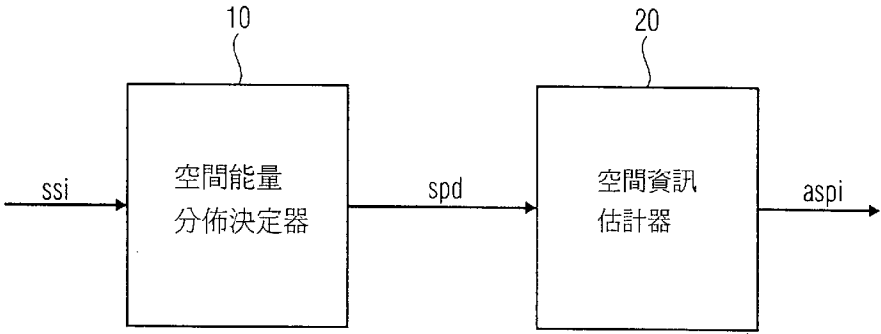


圖 1

符號簡單說明：

aspi . . . 聲源空間資訊

spd . . . 空間能量密度

ssi . . . 聲音源資訊

10 . . . 空間能量分佈決定器

20 . . . 空間資訊估計器

公告本

發明摘要

※ 申請案號：101145071

※ 申請日：101 11 20

※ IPC 分類：H04R 5/027 (2006.01)

【發明名稱】 依據空間能量密度定位麥克風之設備及方法

APPARATUS AND METHOD FOR MICROPHONE

POSITIONING BASED ON A SPATIAL POWER DENSITY

【中文】

提供一種定位麥克風之設備。此設備包含一空間能量分佈決定器（10）及一空間資訊估計器（20）。空間能量分佈決定器（10）係適合於依據聲音源資訊（表示位於環境中之一個或多個聲音源之一個或多個能量數值及一個或多個位置數值）來決定一空間能量密度（表示關於一環境之複數個位置之能量數值）。空間資訊估計器（20）係適合於依據空間能量密度來估計聲源空間資訊。

【英文】

An apparatus for microphone positioning is provided. The apparatus comprises a spatial power distribution determiner (10) and a spatial information estimator (20). The spatial power distribution determiner (10) is adapted to determine a spatial power density indicating power values for a plurality of locations of an environment based on sound source information indicating one or more power values and one or more position values of one or more sound sources located in the environment. The spatial information estimator (20) is adapted to estimate acoustic spatial information based on the spatial power density.

【代表圖】

【本案指定代表圖】：圖 1。

【本代表圖之符號簡單說明】：

aspi：聲源空間資訊

spd：空間能量密度

ssi：聲音源資訊

10：空間能量分佈決定器

20：空間資訊估計器

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：

無

發明專利說明書

【發明名稱】 依據空間能量密度定位麥克風之設備及方法
APPARATUS AND METHOD FOR MICROPHONE
POSITIONING BASED ON A SPATIAL POWER DENSITY

【技術領域】

【0001】 本發明是有關於音源信號處理，且特別是有關於一種自動定位麥克風之設備及方法。

【先前技術】

【0002】 音源信號處理變得越來越重要。更特別是，空間聲音錄製係被應用於多種應用中。空間聲音錄製著眼於利用多重麥克風之幫助來擷取聲場，以於再生側使聆聽者察覺到如於錄製場地（location）之聲音影像。

【0003】 空間聲音錄製之標準方法通常涉及到隔開的全向麥克風（例如，AB 立體音響），一致指向性麥克風（例如，在強度立體音響中），或例如在高傳真身歷聲響複製（Ambisonics）中之更多先進麥克風（例如一B-格式麥克風）。例如，請參考文獻：

[1] Michael A. Gerzon。多聲道廣播及視頻中之高傳真身歷聲響複製（Ambisonics in multichannel broadcasting and video）。J. 音源工程學會, 33（11）:859-871, 1985 年。

【0004】 一種空間麥克風（譬如指向性麥克風、麥克風陣列等）係能夠錄製空間聲。專業術語「空間麥克風」表示定向地選擇性獲得空間聲之任何設備（例如指向性麥克風、麥克風陣列等）。

【0005】 對聲音再生而言，既存的非參數方法直接從錄製的麥克風信號推導出期望的音源回放信號。這些方法之一項主要缺點係為空間影像錄製總是相對於所使用之空間麥克風。

【0006】 在多數應用中，將一空間麥克風置放在期望位置（舉例而言，其可能是靠近一個或多個聲音源之一位置）中是不可能或可行的。於此情況下，將多重空間麥克風置放在更進一步遠離有效聲音源且仍然能夠依期望擷取聲音場景將是更有利的。

【0007】 某些應用採用兩個以上的真實空間麥克風。吾人應注意到

專業術語「真實空間麥克風」表示實體上存在之期望的麥克風型式或麥克風組合（例如一指向性麥克風，如使用於共同立體聲麥克風中之一對指向性麥克風，但亦是一麥克風陣列）。

【0008】 對每個真實空間麥克風而言，到達方向（DOA）可在時頻域中被估計。藉由使用由真實空間麥克風所蒐集資訊，與它們的相對位置之知識一起，其可能計算出實際上被置於一環境中之任意位置（隨意）之一空間麥克風之輸出信號。在下文中，這種空間麥克風被稱為「虛擬空間麥克風」。

【0009】 在這種應用中，需要手動輸入一個或多個虛擬麥克風之位置及方位。然而，如果將自動決定一個或多個虛擬麥克風之一最佳位置及/或方位的話，則那是可讚賞的。

【0010】 如果一設備及方法將是可得到以決定何處放置一虛擬麥克風，何處放置一物理麥克風或決定一最佳聆聽位置，則這將是有利的。此外，如何將一麥克風置放在一最佳方位中將是有利的。專業術語「麥克風定位」及「地位資訊」相關於如何決定一麥克風或一聆聽者之一適當位置與如何決定一麥克風或一聆聽者之一適當方位。

【發明內容】

【0011】 本發明之目的係用以提供關於麥克風定位之改善概念。本發明之目的係藉由依據申請專利範圍第 1 項之設備，藉由依據申請專利範圍第 17 項之方法以及藉由依據申請專利範圍第 18 項之電腦程式而達成。

【0012】 提供一種用於決定最佳麥克風或聆聽位置之設備。此設備包含一空間能量分佈決定器及一空間資訊估計器。空間能量分佈決定器係適合於依據聲音源資訊（表示位於環境中之一個或多個聲音源之一個或多個能量數值及一個或多個場地數值）來決定一空間能量密度（表示一環境之複數個位置之能量數值）。空間資訊估計器係適合於依據空間能量密度來估計聲源空間資訊。

【0013】 以下，專業術語「虛擬麥克風」一般將表示任何型式之麥克風。更特別是，專業術語「虛擬麥克風」相關於用以決定定位資訊之虛擬空間或非空間麥克風兩者，以及相關於用以決定定位資訊之實體上存在的空間或非空間麥克風。

【0014】 空間資訊估計器係適合於依據由空間能量分佈決定器所決定之空間能量密度，決定一環境中之一最佳虛擬麥克風位置或一最佳虛擬麥克風方位。空間能量密度係依據聲音源之能量數值及對應的位置資訊而由空間能量分佈決定器所決定。

【0015】 提供一種決定一個或多個麥克風之一最佳位置及/或方位之自動方式，用以說明聲音場景（譬如一個或多個虛擬麥克風）。

【0016】 在某些實施例中，空間能量分佈決定器可能適合於利用由一有效度量標準（significance metric）所提供之可選擇的資訊，有效度量標準譬如表示關於 ESS 位置之估計之可靠度之測定。

【0017】 例如，在某些實施例中，聲音之擴散 Psi 可被使用作為有效度量標準。項目（1-Psi）接著可簡單地乘以來源能量數值，同時計算空間能量分佈，以使在空間能量分佈之測定上擴散聲音的貢獻將少於直達聲音。

【0018】 提出概念之一項重要優點係為它們可獨立於房間條件被應用，且並不需要任何關於揚聲器及/或物理聲音源之數目或位置之先驗資訊。藉此，此系統是自立更生的，且可以適合於藉由只使用聲音分析之任何種類之方案。依據習知技術，一先驗資訊必須是可得到的，用以決定一個或多個麥克風之一最佳位置及/或方位。這不是限制此應用，就是必須利用預估方式進行，如此將限制其精確度。藉由採用上述之實施例，這並非是需要的。虛擬麥克風（或複數個虛擬麥克風）之位置係藉由完成一半濾度場景分析，然後依據目標應用之需求改變它而被計算出。

【0019】 不像用以估計虛擬麥克風之一最佳位置及/或方位之其他方法，提出的方法並不需要考量幾何場景之任何資訊。例如，不需要關於有效聲音源之數目（例如，學術研討會中之參與者之數目）之一先驗資訊，也不需要關於有效聲音源之相對位置（例如，一學術研討會房間中之參與者之配置）之任何資訊。對於聲音之資訊係只從有效聲音源之特性推導出，其被稱為說明聲音場景之「有效聲音源」（ESS）。ESS 仿效一空間聲場景，其乃因為一個或多個 ESS 係於一某個時間瞬間或在一某個時頻箱（time-frequency bin）中是有效的。以下，專業術語「物理源」係用於說明來自聲音場景（例如一揚聲器）之一實際來源，而專業術語有效聲音源（ESS）（亦被稱為「聲音源」）係用於說明一種在單一時間或時頻箱中是有效的聲

音事件。每個 ESS 係以一位置及以一能量為其特徵。這種資訊允許建構一空間能量分佈（例如一空間能量密度），其允許決定虛擬麥克風之最佳位置或方位。

【0020】 ESS 之參數可譬如藉由採用關於用以於一可配置的虛擬位置產生一虛擬麥克風之一音源輸出信號之設備而說明於下的概念被獲得。聲音事件位置估計係關於用以產生一虛擬麥克風之一音源輸出信號之設備而被說明於下，更特別是參考圖 15-17 被說明。那裡所說明之概念可被採用以決定一有效聲音源之位置。傳播補償係關於用以產生一虛擬麥克風之一音源輸出信號之設備而說明於下，更特別是參考圖 17-20 被說明。那裡所說明的概念可被採用以決定一有效聲音源之能量。

【0021】 依據一實施例，空間資訊估計器可包含一聲音場景中心估計器，用以估計環境中之一聲音場景之一中心之一位置。空間資訊估計器可更包含一麥克風位置計算器，用以依據聲音場景之中心之位置計算出一麥克風之一位置以作為聲源空間資訊。

【0022】 在另一實施例中，麥克風位置計算器可能適合於計算麥克風之位置，其中麥克風係為一虛擬麥克風。

【0023】 此外，依據另一實施例，聲音場景中心估計器可能適合於計算空間能量密度之一重心，用以估計聲音場景之中心。

【0024】 在更進一步的實施例中，聲音場景中心估計器可能設計成用以依據空間能量密度來決定一能量延遲分佈圖，並依據關於環境中之複數個場地之每一個之能量延遲分佈圖來決定一均方根延遲。聲音場景中心估計器可能設計成用以決定複數個場地之一個場地之位置以作為聲音場景之中心，其具有複數個場地之均方根延遲之最小均方根延遲。

【0025】 在另一實施例中，聲音場景中心估計器可能適合於實施圓圈積分，用以估計聲音場景之中心，其中聲音場景中心估計器可能適合於藉由應用譬如下述公式而藉由以一圓圈捲積空間能量密度來實施圓圈積分

$$g(x, y) = \Gamma(x, y) * C_{(r_0)}(x, y)$$

其中 $\Gamma(x, y)$ 係為空間能量密度，且其中 $C_{(r_0)}(x, y)$ 表示一圓圈，用以當環境係為一種二維環境時，決定關於環境之複數個場地之每一個之一圓圈積分值。

【0026】 或者，聲音場景中心估計器可能適合於藉由應用例如下述公式而藉由以一球體捲積空間能量密度來實施圓圈積分

$$g(x, y, z) = \Gamma(x, y, z) * C_{(r,0)}(x, y, z)$$

其中， $\Gamma(x, y, z)$ 係為空間能量密度，且其中 $C_{(r,0)}(x, y, z)$ 表示一球體，用以當環境係為三維環境時，決定關於環境之複數個場地之每一個之一圓圈積分值。

【0027】 此外，依據一實施例，聲音場景中心估計器可能適合於決定環境之複數個場地之每一個之圓圈積分值之一最大值，用以估計聲音場景之中心。

【0028】 在更進一步的實施例中，麥克風位置計算器可能適合於決定經由環境中之聲音場景之中心之複數條線之一條寬度最寬的線。經由聲音場景之中心之複數條線之每一條可具有一能量寬度，其中寬度最寬的線可能是經由具有最大能量寬度之聲音場景之中心之複數條線之線。

【0029】 依據一實施例，複數條線之一條考量線之能量寬度可能表示在考量線上之一段之一最大長度，以使限制此段之此段之第一點，且以使限制此段之此段之一不同的第二點，兩者都具有由空間能量密度（可能大於或等於一預定能量數值）所表示之一能量數值。麥克風位置計算器可能適合於決定麥克風之位置，以使經由聲音場景之中心與麥克風之位置之一第一線可能垂直於寬度最寬的線。

【0030】 在一實施例中，麥克風位置計算器可能設計成用以將一奇異值分解應用至一具有複數行之矩陣。此矩陣之這些行表示在相對於聲音場景之中心之環境中的多個場地之位置。此外，此矩陣之這些行只表示具有由空間能量密度（大於一預定閾值）所表示之多個能量數值之這些場地之位置，或此矩陣之這些行只表示具有由空間能量密度（大於或等於一預定閾值）所表示之多個能量數值之這些場地之位置。

【0031】 依據另一實施例，空間資訊估計器可包含一方位決定器，用以依據空間能量密度來決定麥克風之一方位。方位決定器可能適合於決定麥克風之方位，以使麥克風係朝向聲音場景之中心被配向。方位決定器可能設計成用以藉由應用下述公式來決定關於複數個方向 ϕ 之每一個之一積分值 $f(\phi)$

$$f(\varphi) = \int_0^{r_{\max}} \Gamma(r \cos(\varphi), r \sin(\varphi)) \cdot r \, dr,$$

其中 $r_{\max}$ 定義距離麥克風之一最大距離，且其中方位決定器係被設計成用以依據決定的積分值 $f(\varphi)$ 來決定麥克風之方位。

【0032】 在另一實施例中，空間能量分佈決定器可能適合於藉由應用下述公式而為一時頻箱 (k, n) 決定關於環境之複數個場地之空間能量密度

$$\Gamma(x, y, k, n) = \sum_{i=1}^N \text{power } i(k, n) \cdot g(\gamma_i, x - x_{\text{ESS}i}, y - y_{\text{ESS}i}, k, n),$$

當環境係為一種二維環境時，或藉由應用下述公式

$$\Gamma(x, y, z, k, n) = \sum_{i=1}^N \text{power } i(k, n) \cdot g(\gamma_i, x - x_{\text{ESS}i}, y - y_{\text{ESS}i}, z - z_{\text{ESS}i}, k, n)$$

當環境係為一種三維環境時，

其中 k 表示頻率指數及 n 表示時間指數，其中 x 、 y 、 z 表示複數個場地之其中一個之座標，其中能量 $i(k, n)$ 表示位於時頻箱 (k, n) 之第 i 個聲音源之能量數值，其中 $x_{\text{ESS}i}$ 、 $y_{\text{ESS}i}$ 、 $z_{\text{ESS}i}$ 表示第 i 個聲音源之座標，其中 γ_i 係為一純量值，其可能表示一每個有效聲音源之位置估計是如何可靠的指示因子，且其中 g 係為取決於 x 、 y 、 z 、 $x_{\text{ESS}i}$ 、 $y_{\text{ESS}i}$ 、 $z_{\text{ESS}i}$ 、 k 、 n 及 γ_i 之一函數。

【圖式簡單說明】

【0033】

以下參考附圖說明本發明之實施例，其中：

圖 1 顯示一種依據一實施例之定位麥克風之設備。

圖 2 說明一種依據另一實施例之定位麥克風之設備。

圖 3 顯示依據一實施例之定位麥克風之設備之輸入及輸出。

圖 4a-4c 顯示一種定位麥克風之設備之複數個應用方案。

圖 5 說明依據一實施例之空間能量分佈決定器 21。

圖 6a 顯示用於建構函數 g 之差量函數。

圖 6b 說明用於建構函數 g 之分佈函數。

圖 7 顯示依據一實施例之空間資訊估計器。

圖 8 顯示依據更進一步的實施例之一空間資訊估計器。

圖 9 顯示依據描述更多細節之另一實施例之麥克風位置/方位計算器

44。

圖 10a-10c 說明依據一實施例之依據投射之能量寬度之最佳化。

圖 11 顯示依據另一實施例之空間資訊估計器，其中空間資訊估計器更包含一方位決定器。

圖 12 顯示一種依據一實施例之用於產生一音源輸出信號之設備。

圖 13 顯示依據一實施例之用於產生一音源輸出信號之設備及方法之輸入及輸出。

圖 14 顯示依據一實施例之用於產生一音源輸出信號之設備之基本構造，其包含一聲音事件位置估計器及一資訊計算模組。

圖 15 顯示一示範方案，於其中真實空間麥克風係被描繪為均一線性陣列 (Uniform Linear Arrays)，每個有 3 個麥克風。

圖 16 說明在 3D 中之兩個空間麥克風，用於在 3D 空間中估計到達方向。

圖 17 顯示一幾何形狀，於此目前時頻箱 (k,n) 之一等向性點狀聲音源係位於一位置 $P_{IPLS}(k,n)$ 。

圖 18 說明依據一實施例之資訊計算模組。

圖 19 說明依據另一實施例之資訊計算模組。

圖 20 顯示兩個真實空間麥克風，一定位的聲音事件及一虛擬空間麥克風之位置。

圖 21 顯示依據一實施例之如何相對於一虛擬麥克風獲得到達方向。

圖 22 說明一種從依據一實施例之虛擬麥克風之觀點而言，用來推導出聲音之 DOA 之可能方式。

圖 23 顯示依據一實施例之包含一擴散計算單元之一資訊計算方塊。

圖 24 說明依據一實施例之一擴散計算單元。

圖 25 顯示一方案，於此聲音事件位置估計並不可能。

圖 26 顯示兩個真實空間麥克風、一定位的聲音事件及一虛擬麥克風之位置。

圖 27a-27c 顯示多個方案，於此兩個麥克風陣列接收直達聲音，被牆壁回響之聲音及擴散聲音。

【實施方式】

【0034】 圖 1 顯示依據一實施例之一種定位麥克風之設備。此設備包含一空間能量分佈決定器 10 及一空間資訊估計器 20。空間能量分佈決定器 10 係適合於決定一空間能量密度 spd ，其依據聲音源資訊 ssi 指示在環境中之複數個位置之能量數值，其中聲音源資訊 ssi 指示位於環境中之一個或多個有效聲音源（EES）之一個或多個能量數值及一個或多個位置數值。空間資訊估計器 20 適合於依據空間能量密度估計聲源空間資訊 $aspi$ 。

【0035】 圖 2 顯示依據另一實施例之定位麥克風之設備。此設備包含一空間能量分佈決定器 21，用於依據有效聲音源資訊決定一個指示環境之複數個位置之能量數值之空間能量密度（SPD）（亦以空間能量分佈表示），其中有效聲音源資訊指示部署於環境中之一個或多個有效聲音源之一個或多個核心數值及位置數值。此設備更包含一空間資訊估計器 22，用於依據空間能量密度估計一虛擬麥克風（VM）之一位置及/或方位。

【0036】 圖 3 顯示依據一實施例之定位麥克風之設備之輸入及輸出。至此設備之輸入 91 、 92 、... $9N$ 包含能量，例如，聲場壓力平方及位置之絕對值，例如，2D 或 3D 迪卡兒（Cartesian）座標。有效聲音源（ESS）描述聲音場景（聲場）。

【0037】 有效聲音源可能例如等於如關於設備之以下所述之瞬間點狀聲音源（IPLS），用於在一可配置的虛擬位置產生一虛擬麥克風之一音源輸出信號。

【0038】 於輸出，傳回一個或多個虛擬麥克風位置及場地。以下，專業術語「物理源」係用於說明來自聲音場景之實際來源，例如是一揚聲器，而專業術語有效聲音源（ESS），（亦以「聲音源」表示），係用於說明一聲音事件，其在單一時間或時頻箱（time-frequency bin）是有效的，如亦用來作 IPLS，其相關於用於在一可配置的虛擬位置產生一虛擬麥克風之一音源輸出信號之設備而說明於下。

【0039】 此外，吾人應注意到，專業術語「聲音源」涵蓋物理源與有效聲音源兩者。

【0040】 依據圖 2 之本實施例之設備之輸入 $91, 92, \dots, 9N$ 包含 N 個有效聲音源之位置及對應能量之資訊， N 個有效聲音源係在一時間實例或一時頻箱之內被定位的，如下所述之關於用於在一可配置的虛擬位置產生一虛擬麥克風之一音源輸出信號之設備，其係揭露於下列文獻：

[20] Giovanni Del Galdo, Oliver Thiergart, Tobias Weller, 以及 E. A. P. Habets。藉由使用由配置的陣列所蒐集之幾何資訊產生虛擬麥克風信號 (Generating virtual microphone signals using geometrical information gathered by distributed arrays)。於免持語音通信和麥克風陣列的第三次聯合研討會 (HSCMA '11) 英國愛丁堡，2011 年 5 月。

【0041】 舉例而言，此種資訊可被包含於設備之資訊計算模組之圖 14 之輸出 106，此設備用於在一可配置的虛擬位置產生一虛擬麥克風之一音源輸出信號，可配置的虛擬位置是當應用短時間傅立葉轉換 (STFT) 時，關於 $1, 2, \dots, N$ 之不同頻率箱 (frequency bin) 被考量於以下。

【0042】 關於定位麥克風之設備，不同的操作模式可在某個時間間隔期間變成有效，每個表示用於為一個或多個虛擬麥克風定位及定方位之各種方案。可將一定位麥克風之設備使用於複數個應用方案：

【0043】 於一第一應用方案中， N 個全向虛擬麥克風可能置於聲音場景之內 (參見圖 4a)。因此，於本應用方案中，一些虛擬麥克風覆蓋整個聲音場景。

【0044】 於一第二應用方案中，單一虛擬麥克風係定位於聲音場景之聲源中心。舉例而言，將全向虛擬麥克風、心型指向虛擬麥克風或一虛擬空間麥克風 (例如 B-格式麥克風) 安置成使所有參與者被最佳化地擷取 (圖 4b)。

【0045】 於一第三應用方案中，將一空間麥克風置於聲音場景之「外部」。舉例而言，安置一虛擬立體聲麥克風以獲得一寬闊空間影像，如圖 4c 所示。

【0046】 於一第四應用方案，估計虛擬麥克風之最佳方位，同時將虛擬麥克風安置在一固定位置 (預先決定的位置)，譬如虛擬麥克風之位置

及方向性可能是預定的，而只有方位是自動地被計算的。

【0047】 吾人應注意到，所有上述應用可能包含暫時的適應性。例如，當揚聲器於房間中移動時，虛擬點麥克風的位置/方位跟隨此揚聲器。

【0048】 於圖 2 及 3，可選擇的資訊係由一有效度量標準 (significant metric) 13 提供，有效度量標準 13 譬如表示用於估計 ESS 位置之可靠度之測定。舉例而言，這種度量標準可從到達方向估計器之變異（當如所說明地使用兩個以上的麥克風陣列時）推導出，如說明於以下之在一可配置的虛擬位置產生一虛擬麥克風之一音源輸出信號之設備；或從擴散參數推導出，請參照下列文獻的計算：

[29] Ville Pulkki。利用定向音源編碼之空間聲再生 (Spatial sound reproduction with directional audio coding)。J.音源工程學會，55(6):503-516, 2007 年 6 月。

【0049】 度量標準可相對於所有輸入 $91, \dots, 9N$ （譬如，可能使用關於所有輸入之度量標準之一固定數值）被表示，或可關於每個輸入 $91, \dots, 9N$ 不同地被定義。圖 2 之設備之輸出 15、16 可能包含一個或多個虛擬麥克風之位置及/或方位。依據應用，可能產生複數個虛擬麥克風之輸出（位置及方位），每個對應於一特定虛擬麥克風。

【0050】 圖 5 顯示依據一實施例之一空間能量分佈決定器 21。空間能量分佈決定器包含一空間能量分佈主處理單元 31 及一空間能量分佈後處理單元 32。空間能量分佈決定器 21 係適合於決定（更正確地說是計算）一種以 $\Gamma(x, y, z, k, n)$ 表示於以下之變化空間能量密度 (SPD)， $\Gamma(x, y, z, k, n)$ 表示於空間中之關於每個時頻箱 (k, n) 之某個點（例如 (x, y, z) ）被定位的之能量。SPD 係藉由對在有效聲音源 $91, \dots, 9N$ 之位置之能量數值（被輸入至空間能量分佈決定器 21 中）進行積分而產生。

【0051】 關於一時頻箱 (k, n) SPD 之計算可能依據公式而完成

$$\Gamma(x, y, z, k, n) = \sum_{i=1}^N \text{power}_i(k, n) \cdot g(\gamma_i, x - x_{\text{ESS}_i}, y - y_{\text{ESS}_i}, z - z_{\text{ESS}_i}, k, n), \quad (1)$$

其中， (x, y, z) 表示系統之座標，而 x_{ESS_i} 、 y_{ESS_i} 、 z_{ESS_i} 係為有效聲音源 i 之座標。有效度量標準 103 γ_i 表示一種每個有效聲音源之位置估計是多麼可靠之

指示因子。默認情況下，有效度量標準可能等於 1。於此，吾人應注意到能量 i ($power_i$) 及座標 x_{ESSi} 、 y_{ESSi} 及 z_{ESSi} 對應至圖 3 之輸入 $9i$ 。此外，吾人應注意到，為表示法之簡化之便， (k,n) 延伸將不會被寫入於下。然而，下述公式仍然取決於特別考量的時頻箱 (k,n) 。

【0052】 由空間能量分佈主處理單元 31(例如於圖 5 中)產生之 SPD 可能更進一步的被空間能量分佈後處理單元 32 (SPD 及暫時積分模組之後處理) 處理，並且例如，藉由採用一自回歸濾波器而於時間上積分。為了更強健以抵抗聲音場景離群值 (亦即，由不正確的位置估計所導致的)，任何種類之後處理濾波器都可能被應用於 SPD 上。這一種後處理濾波器可能譬如是一低通濾波器或一形態 (侵蝕、膨脹) 濾波器。

【0053】 當計算一個或多個虛擬麥克風之位置及/或方位時，可能採用一種可選擇的參數，其取決於 SPD。此種參數可能參考例如，置放虛擬麥克風 (VM) 之房間之被禁止的及/或較佳的區域，或者可能參考 SPD，選擇滿足某些預先決定的規則之特定 SPD 範圍。

【0054】 如於公式 (1) 可瞭解到， g 為在空間中之有效度量標準 γ (更正確地說是 γ_i) 之函數，其在默認情況下具有等於 1 之數值。否則， γ 可能用於考慮不同貢獻。舉例而言，如果 σ^2 是位置估計之變異，則例如 γ

可能設定為 $\gamma = \frac{1}{\sigma^2}$ 。

【0055】 或者，可能採用在麥克風陣列計算之平均擴散 Ψ ，導致 $\gamma = 1 - \Psi$ 。

【0056】 藉此，可能選擇 γ ，以使其減少更不可靠的估計及增加更可靠的估計。

【0057】 對於建構函數 g 存在複數個可能性。實際上特別有用之兩個例子是：

$$g(\gamma, s) = \frac{1}{(2\pi)^{3/2} |\Sigma_\gamma|^{1/2}} e^{-\frac{1}{2}(s-\mu)' \Sigma_\gamma^{-1} (s-\mu)} \quad (2)$$

$$g(\gamma, x, y, z) = \gamma \cdot \delta(x) \cdot \delta(y) \cdot \delta(z) \quad (3)$$

【0058】 於第一函數， $\delta(x)$ 、 $\delta(y)$ 及 $\delta(z)$ 表示差量函數（參見顯示差量函數之圖 6a）。於第二函數， $s = [x \ y \ z]^T$ ， $\mu = [\mu_x \ \mu_y \ \mu_z]^T$ 是平均向量，而 Σ_γ 是高斯分佈函數 g （參見顯示分佈函數之圖 6b）之協方差矩陣（covariance matrix）。協方差矩陣係使用下述公式計算：

$$\Sigma_\gamma = E[(s - \mu)(s - \mu)^T]; \quad (4)$$

其取決於關於 $\gamma = \frac{1}{\sigma^2}$ 之方案 γ 之選擇，即考慮到譬如關於 1D 情況：

$$\sigma^2 = E[(x - \mu_x)^2]. \quad (5)$$

【0059】 於公式（3）可瞭解到，函數 g 可被在由輸入 91...9N 所得之有效聲音源位置周圍之一分佈函數描述，於此例如有效度量標準係為一高斯分佈之變異之反函數。如果一聲音源位置之估計具有一高可靠度，則根據分佈將會是相當狹小，而一更不可靠的估計將對應至一高變異，且將因此是一寬廣分佈，參見譬如顯示一種 1D 例子之圖 6b。

【0060】 圖 7 顯示依據一實施例之一空間資訊估計器 22。空間資訊估計器包含一聲音場景中心估計器 41，用於估計環境中之聲音場景之中心之位置。再者，空間資訊估計器包含一麥克風位置計算器 42，用於基於聲音場景之中心之位置計算麥克風之位置作為聲源空間資訊。

【0061】 圖 8 顯示依據更進一步的實施例之空間資訊估計器 22。空間資訊估計器包含一虛擬麥克風位置計算器 44，其適合於計算一虛擬麥克風之位置，且更進一步適合於決定一虛擬麥克風之方位。虛擬麥克風位置計算器 44 因此亦被稱為麥克風位置/方位計算器 44。

【0062】 圖 8 之空間資訊估計器 22 使用以前產生之 SPD 23 作為輸入。其依據目標應用回傳一個或多個虛擬麥克風之位置 15 及方位 16 作為輸出。第一處理方塊（聲音場景中心估計器 41）提供聲音場景中心之估計。方塊 41 之輸出 43，例如聲音場景中心之位置，接著被提供至第二處理方塊（虛擬麥克風位置/方位計算器 44）作為輸入。虛擬麥克風位置/方位計算器 44 依據目標應用執行一個或多個虛擬麥克風之最終位置 15 及方位 16 之實際估計。

【0063】 聲音場景中心估計器 41 提供聲音場景中心之估計。聲音場景中心估計器 41 之輸出接著被供至麥克風位置/方位計算器 44 作為輸

入。麥克風位置/方位計算器 44 依據表示目標應用的特色之操作模式執行一個或多個虛擬麥克風之最終位置 15 及/或方位 16 之實際估計。

【0064】 現在將更詳細地說明聲音場景中心估計器之實施例。為了獲得聲音場景之中心，存在有數個可能的概念。

【0065】 依據一第一實施例之第一概念，聲音場景之中心係藉由計算 SPD $\Gamma(x, y, z)$ 之重心而獲得。 $\Gamma(x, y, z)$ 之數值可能是 s ，被解釋成為於空間中之點 (x, y, z) 之既存質量。

【0066】 依據一第二實施例之第二概念，應該可發現具有通道之最小時間分散之空間中之位置。這是藉由考慮到均方根 (RMS) 延遲擴展而達成。首先，關於空間中之每個點 $p = (x_0, y_0)$ ，依據 SPD $\Gamma(x, y, z)$ 計算一能量延遲分佈圖 (PDP) $A_p(\tau)$ ，例如使用

$$A_p(\tau) = \iint_{x, y} \Gamma(x, y) \cdot \delta(t - \tau) dy dx$$

於此 $\tau = \sqrt{(x - x_0)^2 + (y - y_0)^2} / c$

【0067】 從 $A_p(\tau)$ ，接著使用下述的方程式計算 RMS 延遲：

$$\tau_{\text{RMS}, p} = \sqrt{\frac{\int_0^\infty (\tau - \bar{\tau}_s)^2 A_p(\tau) d\tau}{\int_0^\infty A_p(\tau) d\tau}},$$

於此 $\bar{\tau}_s$ 表示 $A_p(\tau)$ 之平均延遲。平均延遲 $\tau_{\text{RMS}, p}$ 最小之位置將表示聲音場景之中心。

【0068】 依據一第三實施例之第三概念，其可能被採用作為聲音場景中心估計之替代方案，提出一種「圓圈整合 (circle-integration)」。舉例而言，於 2D 狀況下，SPD $\Gamma(x, y)$ 係依據下述公式而以一圓圈 $C(r, o)$ 纏繞：

$$g(x, y) = \Gamma(x, y) * C_{(r, o)}(x, y),$$

其中 r 是圓圈之半徑，而 o 定義圓圈之中心。半徑 r 可能是常數或可隨著於點 (x, y) 之能量數值來變化。舉例而言，在點 (x, y) 之高能量可能對應至大半徑，而低能量可能對應至小半徑。對能量之額外依賴關係亦是可能的。一個這種例子將會是在使用二元高斯函數以建構函數 $g(x, y)$ 之前，利用二元高斯函數來纏繞圓圈。依據這一種實施例，二元高斯函數之協方差矩

陣變成取決於位置 (x,y) 之能量，亦即，高能量對應至低變異，而低能量對應至高變異。

【0069】 一旦計算出 $g(x,y)$ ，聲音場景之中心就可能依據下述公式來決定：

$$\hat{p}_{\text{center}} = \arg \max_{x,y} g(x,y).$$

【0070】 於更進一步的實施例中，此種概念係藉由採用具有一球體之 $\Gamma(x,y,z)$ 之 3D 卷積而延伸至 3D，依此類推。

【0071】 圖 9 顯示依據描述更多細節之另一實施例之麥克風位置/方位計算器 44。聲音場景之中心 43 與 SPD 23 一起被提供至麥克風位置/方位計算器 44 作為輸入。於麥克風位置/方位計算器 44 中，關於聲音場景之中心 43 之資訊，可依據目標應用所需要之操作被複製至輸出，並被直接地使用作為一虛擬麥克風之位置（其譬如在適合圖 4b 之應用方案時，關於使一個虛擬麥克風安置於聲源聲音場景之中心之方案）。或者，關於聲音場景之中心 43 之資訊可被使用作為麥克風位置/方位計算器 44 之內部之一修改參數。

【0072】 可能應用不同概念來計算一麥克風位置，例如：

依據投射之能量寬度之最佳化，

依據主成分分析之最佳化。

【0073】 關於假設之示範目的，可能依據圖 4c 之應用方案（關於在側場景之外部之一空間麥克風之方案）來計算麥克風之位置。然而，這些說明係同等適合於任何其他應用方案。

【0074】 以前列舉之依據實施例之用於估計虛擬麥克風之位置之概念，現在將更詳細說明於下。

【0075】 依據投射之能量寬度之最佳化定義一組 M 條等間距隔開的線，其通過聲音場景之中心。關於在例如 2D 方案中之每條線，SPD $\Gamma(x,y)$ 係正交地投影於它們之上並被加總。

【0076】 圖 10a 至 10c 顯示依據投射之能量寬度之最佳化。於圖 10a 中，投影的能量函數 P_{proj} 係為每一條 $\ell_1, \dots, \ell_i, \dots, \ell_M$ 線被計算。然後計算函數之相對應的寬度，參見圖 10b。舉例而言，寬度可能被定義為 -3 dB

寬度，其係對應至使距離段之最左邊及最右邊點對應至一預定能量位準（譬如高於-3 dB 之能量位準）之距離。然後，識別出具有最廣寬度之線，且將虛擬麥克風置於與其正交之方向。可能設定虛擬麥克風之方位以使其指向聲音場景之中心，如下一章節所說明。利用此種方法，可獲得兩個可能的虛擬麥克風（VM）位置，因為可將 VM 安置於正或負正交方向上。

【0077】 可能譬如依據幾何考量以及虛擬麥克風之開口角度計算出 VM 所安置之距離。此乃以圖 10c 表示。VM 安置之距離隨著目標應用所特定之操作模式變化。這意味著建構三角形，以使圖 10c 之寬度 i 表示三角形之一側，而重心 COG 係為此側之中點。藉由採正交線於 COG 並界定其為 VM 開口角度 α 之平分線，找到三角形之第三頂點。然後，平分線之長度給予在 VM 位置與聲音場景之中心之間之距離。

【0078】 依據另一實施例，基於投射之能量所說明之最佳化概念可能延伸至 3D。於此情況下， $M2$ 個等間距隔開的平面（於方位角及俯仰方向）係被定義以取代 M 條線。然後，寬度係對應至圓圈之直徑，其包含投射之能量之最大部分。最終位置是藉由置放 VM 於最大圓圈直徑之正交平面表面而獲得。依據一實施例，從聲音場景之中心至 VM 位置之距離可能再被計算，類似於 2D 情況，其使用幾何考量以及由操作模式而特別指定之開口角度。

【0079】 依據另一實施例，採用依據一主成分分析之最佳化。依據一種像主成分分析之處理之最佳化，直接使用從 SPD 可得到的資訊。首先，將 SPD $\Gamma(x,y,z)$ 量化，且將一閾值-選擇性濾波器應用於量化之資料集上。藉此，捨棄具有小於某個閾值之能階之所有點。後來，剩下的點 $h_i = [h_{x,i}, h_{y,i}, h_{z,i}]^T$ 是以平均值為中心（亦即，以平均值為中心的點表示第 i 個有效來源減去聲音場景中心之座標之座標），且然後被重組於一資料矩陣 H 如下：

$$H = \begin{pmatrix} h_{x,0} & h_{x,1} & h_{x,2} & \dots & h_{x,N} \\ h_{y,0} & h_{y,1} & h_{y,2} & \dots & h_{y,N} \\ h_{z,0} & h_{z,1} & h_{z,2} & \dots & h_{z,N} \end{pmatrix},$$

於此 N 定義在取閾值之後之點之數目。然後，奇異值分解（SVD）被應用至 H ，以使其被因式分解成下述乘積：

$$H = U \cdot \Sigma \cdot V^T$$

【0080】 U 之第一行表示主分量，其具有資料集之最高變異性。U 之第二行係垂直於第一行並表示我們要置放 VM 之方向。寬度係隱式地由矩陣 Σ 中之第一奇異值所得。知道寬度以及方向後，我們可計算 VM 之位置及方位，如上面參考圖 10a 至 10c 所說明之依據投射之能量寬度之最佳化方法。

【0081】 於另一實施例中，將這些方法應用至 2D 問題，其是直截了當的，因為僅需要從方程式及計算中忽略/移除 z 軸分量。

【0082】 關於其他應用，例如圖 4a 之應用方案（複數個虛擬麥克風覆蓋整個聲音場景），可能採用一種不同概念，例如一疊代最佳化機構。於一第一步驟，識別出具有 SPD 之最大值之位置。藉此，標示出總數 N 個虛擬麥克風之第一 VM 之位置。於此之後，圍繞此位置（亦即，至多達某個距離）之所有能量被移除離開 SPD。重複先前步驟，直到找到 N 個虛擬麥克風之所有位置為止。在未定義 N 的狀況下，則繼續重複先前步驟，直到 SPD 之最大值小於某個閾值為止。

【0083】 圖 11 顯示另一實施例，其中一空間資訊估計器 22 更包含一方位決定器 45。方位決定器 45 係適合於依據空間能量密度 23 決定麥克風之一（適當的）方位 16。

【0084】 以下，將說明方位估計。依據投射之能量寬度以及主成份分析之最佳化方法隱式地計算虛擬麥克風之方位 16，因為將虛擬麥克風假設為被定方位至聲音場景之中心。

【0085】 然而，關於某些其他應用方案，可能譬如於一應用方案中適合於詳細地計算方位，其中估計虛擬麥克風之最佳方位，其中虛擬麥克風係位於一固定位置。於此情況下，應決定方位以使虛擬麥克風拾起聲音場景中之大部分的能量。

【0086】 依據一實施例，為了決定一虛擬麥克風之方位，首先取樣可能的方向 φ ，且執行於這些方向之每個上之遍及能量之積分。獲得 φ 之下述函數：

$$f(\varphi) = \int_0^{r_{\max}} \Gamma(r \cos(\varphi), r \sin(\varphi)) \cdot r \, dr,$$

於此 $r_{\max}$ 被定義為離開 VM 之最大距離，並控制 VM 之拾音模式。然後，VM 最終方位 ϕ 被計算為：

$$\hat{\phi} = \arg \max_{\phi} \int_{-\pi}^{\pi} f(\varphi) \cdot w_{\phi}(\varphi) d\varphi,$$

於此 $w_{\phi}(\varphi)$ 係為基於 VM 之輸入特徵之加權函數。例如， $w_{\phi}(\varphi)$ 可能是定義來自方向 φ 之能量是如何被縮放之函數，假設某個觀察方向 ϕ 及 VM 之特定拾音模式的話。

【0087】 以下將說明用於產生一音源輸出信號以模擬一虛擬麥克風於一環境中之一可配置的虛擬位置之錄音之設備。依據上述實施例之其中一個之一定位麥克風之設備可能被採用以為產生音源輸出信號之設備決定虛擬位置。

【0088】 圖 12 顯示用於產生一音源輸出信號以模擬一虛擬麥克風於一環境中之一可配置的虛擬位置 posVmic 之錄音之設備。此設備包含一聲音事件位置估計器 110 及一資訊計算模組 120。聲音事件位置估計器 110 接收來自一第一真實空間麥克風之一第一方向資訊 di1 及來自一第二真實空間麥克風之一第二方向資訊 di2 。聲音事件位置估計器 110 適合於估計一聲音源位置 ssp ，其指示一聲音源於環境中之一位置，聲音源放射一聲波，其中聲音事件位置估計器 110 適合於依據一第一方向資訊 di1 及依據一第二方向資訊 di2 來估計聲音源位置 ssp ，其中第一方向資訊 di1 是由位於環境中之一第一真實麥克風位置 pos1mic 之一第一真實空間麥克風提供，而第二方向資訊 di2 是由位於環境中之一第二真實麥克風位置之一第二真實空間麥克風提供。資訊計算模組 120 適合於依據一第一錄製之音源輸入信號 is1 （由第一真實空間麥克風錄製）、依據第一真實麥克風位置 pos1mic 及依據虛擬麥克風之虛擬位置 posVmic 來產生音源輸出信號。資訊計算模組 120 包含一傳播補償器，其適合於產生一第一變化音源信號，其所採取的是藉由修改第一錄製之音源輸入信號 is1 ，其所採取的是藉由補償在於第一真實空間麥克風由聲音源發出之聲波之到達與於虛擬麥克風之聲波之到達之間之一第一延遲或振幅衰減，其所採取的是藉由調整第一錄製之音源輸入信號 is1 之一振幅數值、一幅度數值或一相位數值，用來獲得音源輸出信號。

【0089】 圖 13 顯示依據一實施例之設備及方法之輸入及輸出。來自兩個以上的真實空間麥克風 111、112、...、11N 之資訊係回饋至設備/被方法處理。此種資訊包含被真實空間麥克風拾音之音源信號與來自真實空間麥克風之方向資訊，例如到達方向 (DOA) 估計。音源信號與方向資訊 (例如到達方向估計) 可能以一時頻域表示。舉例而言，如果期望 2D 幾何重建並選擇一傳統 STFT (短時間傅立葉轉換) 範圍以供信號之表現用，則 DOA 可能依據 k 及 n (亦即頻率及時間指數) 被表示為方位角角度。

【0090】 於實施例中，在空間中以及描述虛擬麥克風之位置之聲音事件定位的可能依據共同座標系統中之實際及虛擬空間麥克風之位置及方位來實施。此種資訊可能以圖 13 中之輸入 121...12N 及輸入 104 表示。輸入 104 可能另外指定虛擬空間麥克風之特徵，例如，其位置及拾音模式，如以下所將探討的。如果虛擬空間麥克風包含多重虛擬感測器，則可能考量它們的位置及相對應的不同的拾音模式。

【0091】 當期望時，設備之輸出或一對應的方法可能是一個或多個聲音信號 105，其可能已被一空間麥克風拾音，而空間麥克風是如由 104 而特別指定地被定義且安置。此外，設備 (更正確地說是方法) 可能提供對應的空間側資訊 106 (可能藉由採用虛擬空間麥克風來估計) 作為輸出。

【0092】 圖 14 顯示依據一實施例之設備，其包含兩個主處理單元、一聲音事件位置估計器 201 及一資訊計算模組 202。聲音事件位置估計器 201 可能基於包含於輸入 111...11N 之 DOA 及基於真實空間麥克風之位置及方位之知識 (DOA 已被計算出) 來實現幾何重建。聲音事件位置估計器 205 之輸出包含聲音源之位置估計 (2D 或 3D)，於此聲音事件為每個時間及頻率箱產生。第二處理方塊 202 係為一資訊計算模組。依據圖 14 之本實施例，第二處理方塊 202 計算一虛擬麥克風信號及空間側資訊。其因此亦以虛擬麥克風信號及側資訊計算方塊 202 表示。虛擬麥克風信號及側資訊計算方塊 202 使用聲音事件之位置 205 來處理包含於真實空間麥克風 111...11N 之音源信號，以輸出虛擬麥克風音源信號 105。如果需要的話，方塊 202 亦可能計算對應於虛擬空間麥克風之空間側資訊 106。以下實施例說明方塊 201 及 202 可能如何操作之可能性。

【0093】 以下，將更詳細地說明依據一實施例之一聲音事件位置估

計器之位置估計。

【0094】 依據問題 (2D 或 3D) 之維數及空間麥克風之數目，對於位置估計可能有數個解決方法。

【0095】 如果在 2D (最簡單的可能狀況) 中存在有兩個空間麥克風，則可能採用簡單的三角測量。圖 15 顯示一示範方案，於其中真實空間麥克風被描繪為均一線性陣列 (ULA)，每個有 3 個麥克風。被表示為方位角角度 $a_1(k,n)$ 及 $a_2(k,n)$ 之 DOA，係為時頻箱 (k,n) 被計算。這是藉由採用一適當的 DOA 估計器 (例如 ESPRIT 或 (root) MUSIC) 至轉換成為時頻域之壓力信號而達成，其中，DOA 估計器為 ESPRIT 或 (root) MUSIC 係分別揭露於下列文獻

[13] R. Roy, A. Paulraj 以及 T. Kailath, "藉由子空間旋轉法之到達方向估計-ESPRIT", 在 1986 年 4 月之美國加利福尼亞州史丹佛大學之語音與信號處理 (ICASSP) 之 IEEE 國際學術研討會中 (IEEE 音響國際會議，語音和信號處理 (ICASSP)，史丹福，美國加州，1986 年 4 月)；及

[14] R. Schmidt, "多重發射器場地及信號參數估計 (Multiple emitter location and signal parameter estimation)", 關於天線與傳播之 IEEE 學報，第 34 卷，第 3 號，第 276-280 頁，1986 年。

【0096】 於圖 15 中，顯示兩個真實空間麥克風 (於此是兩個真實空間麥克風陣列 410、420)。兩個估計的 DOA $a_1(k,n)$ 及 $a_2(k,n)$ 係以兩條線表示，一第一線 430 表示 DOA $a_1(k,n)$ ，而一第二線 440 表示 DOA $a_2(k,n)$ 。經由知道每個陣列之位置及方位之簡單幾何計算，三角測量是可能的。

【0097】 當兩條線 (第一線 430、第二線 440) 正好平行時，三角測量會失敗。然而，在實際應用中，這是非常不可能的。然而，並非所有三角測量結果都會對應至考慮空間中之聲音事件的一物理或可行位置。舉例而言，聲音事件之估計位置可能是太遠離假設空間或甚至在假設空間外部，藉以表示 DOA 可能並未對應至任何可利用所使用的模型而實體上被演奏之聲音事件。這種結果可能由感測器噪音或太強大的房間混響所導致。因此，依據一實施例，這種不希望得到的結果被標記，以使資訊計算模組 202 可適當地處理它們。

【0098】 圖 16 說明一種方案，於此一聲音事件之位置係在 3D 空間中被估計。採用了適當的空間麥克風，譬如，一平面或 3D 麥克風陣列。在圖 16 中，顯示了一第一空間麥克風 510（譬如一第一 3D 麥克風陣列）以及一第二空間麥克風 520（譬如一第一 3D 麥克風陣列）。3D 空間中之 DOA 可能譬如被表示成方位角及仰角。可能採用單位向量 530、540 以表示 DOA。兩條線 550、560 係依據 DOA 被投影。在 3D 中，即使利用很可靠的估計，依據 DOA 所投影的兩條線 550、560 可能不會相交。然而，三角測量仍然可譬如藉由選擇連接兩條線之最小段之中點來實現。

【0099】 類似於 2D 情況，三角測量可能失敗或可能產生關於方向之某些組合之不能實行的結果，其接著亦可被標記成例如圖 14 之資訊計算模組 202。

【0100】 如果存在兩個以上的空間麥克風，則數個解決方法是可能的。舉例而言，上面所說明之三角測量可以為所有真實空間麥克風對（如果 $N = 3$ ，1 與 2，1 與 3，以及 2 與 3）而實現。所產生之位置然後可被平均（沿著 x 及 y ，且如果考量 3D 的話，沿著 z ）。

【0101】 或者，可能使用更多複合概念。舉例而言，可能應用機率方法，揭露於下列文獻：

[15] J. Michael Steele, "平面上之隨機樣品之最佳三角測量 (Optimal Triangulation of Random Samples in the Plane)", 概率的史冊, 第 10 卷, 第 3 號 (1982 年 8 月), 第 548-553 頁。

【0102】 依據一實施例，聲場可能在時頻域中被分析，舉例而言，經由一短時間傅立葉轉換 (STFT) 而獲得，於其中 k 及 n 分別表示頻率指標 k 及時間指標 n 。對某個 k 及 n 而言，位於一任意位置 p_v 之複合壓力 $P_v(k, n)$ 係塑造成由一窄帶等向性點狀源 (narrow-band isotropic point-like source) 所放射之單一球面波，例如藉由採用下述公式：

$$P_v(k, n) = P_{IPLS}(k, n) \cdot \gamma(k, P_{IPLS}(k, n), p_v), \quad (1)$$

於此， $P_{IPLS}(k, n)$ 係為於其位置 $P_{IPLS}(k, n)$ 而由 IPLS 所發出之信號。複合因子 $\gamma(k, P_{IPLS}, p_v)$ 表示從 $P_{IPLS}(k, n)$ 至 p_v 之傳播，例如，其導入適當的相位及大小修正。於此，可能應用下述假設，在每個時頻箱中，只有一個 IPLS 是有效的。然而，位於不同位置之多重窄帶 IPLS 於單一時間瞬間

亦可能是有效的。

【0103】 每個 IPLS 不是仿效直達聲音就是仿效一不同的房間回響。其位置 $P_{\text{IPLS}}(k,n)$ 理想上可能分別對應至位於房間內部之一真實聲源，或位於外部之一鏡像聲源。因此，位置 $P_{\text{IPLS}}(k,n)$ 亦可能表示一聲音事件之位置。

【0104】 請注意專業術語「真實聲源」表示真實存在於錄製環境中之真實聲源，例如揚聲器或樂器。反之，關於「聲源」或「聲音事件」或「IPLS」，我們表示有效的聲源，其於某些時間瞬間或於某些時頻箱是有效的，其中聲源譬如可表示真實聲源或鏡像源。

【0105】 圖 27a-27b 顯示定位聲源之麥克風陣列。定位的聲源依據它們的本質可具有不同的物理演奏。當麥克風陣列接收直達聲音時，它們可能能夠定位一真實聲源（例如揚聲器）之位置。當麥克風陣列接收回響時，它們可能定位一鏡像源之位置。鏡像源亦是聲源。

【0106】 圖 27a 顯示一種方案，於此兩個麥克風陣列 151 及 152 接收來自一真實聲源 153（一真實存在的聲源）之直達聲音。

【0107】 圖 27b 顯示一種方案，於此兩個麥克風陣列 161、162 接收反響聲音，其中聲音已被一牆壁反響。因為回響，麥克風陣列 161、162 定位此位置，於此聲音似乎來自位於一鏡像源 165 之一位置（其不同於揚聲器 163 之位置）。

【0108】 圖 27a 之真實聲源 153 與鏡像源 165 兩者都是聲源。

【0109】 圖 27c 顯示一種方案，於此兩個麥克風陣列 171、172 接收擴散聲音，且並不能夠定位一聲源。

【0110】 雖然這種單波模型只有對輕度混響環境而言是正確的，但鑑於源信號滿足 W-分離正交（WDO）條件，亦即，時頻重疊足夠小。這對於語音信號而言通常是真實的，請參照下列文獻：

[12] S. Rickard 及 Z. Yilmaz, "針對言語之接近 W-分離正交 (On the approximate W-disjoint orthogonality of speech)", 於語音及信號處理, 2002. ICASSP 2002. 2002 年 4 月之 IEEE 國際學術研討會, 第 1 卷。

【0111】 然而，此模型亦提供一良好估計給其他環境，且因此亦適合於那些環境。

【0112】 以下，說明依據一實施例之位置 $P_{IPLS}(k,n)$ 之估計。在某個時頻箱中之一有效 IPLS 之位置 $P_{IPLS}(k,n)$ ，從而是一時頻箱中之一聲音事件之估計，係基於在至少兩個不同的觀察點所測得之聲音之到達方向 (DOA) 而經由三角測量被估計。

【0113】 圖 17 顯示一種幾何形狀，於此目前時頻槽 (time-frequency slot) (k,n) 之 IPLS 係位於未知位置 $P_{IPLS}(k,n)$ 中。爲了決定需要的 DOA 資訊，採用具有一已知的幾何形狀、位置及方位之兩個真實空間麥克風 (於此是兩個麥克風陣列)，其係分別被設置於位置 610 及 620。向量 p_1 及 p_2 分別指向位置 610、620。陣列方位係由單位向量 c_1 及 c_2 所界定。聲音之 DOA 係藉由使用一 DOA 估計演算法 (例如如由 DirAC 分析 (參見文獻[2]、[3]) 所揭露者) 而在每個 (k,n) 之位置 610 及 620 中被決定。由此，相關於麥克風陣列之一觀點之一第一觀點單位向量 $e_1^{POV}(k,n)$ 及一第二觀點單位向量 $e_2^{POV}(k,n)$ (兩者未顯示於圖 17 中)，可能被提供作爲 DirAC 分析之輸出。舉例而言，當在 2D 中操作時，第一觀點單位向量變成：

$$e_1^{POV}(k,n) = \begin{bmatrix} \cos(\varphi_1(k,n)) \\ \sin(\varphi_1(k,n)) \end{bmatrix}, \quad (2)$$

【0114】 於此， $\varphi_1(k,n)$ 表示於第一麥克風陣列所估計之 DOA 之方位角，如圖 17 所示。相關於原點之全域座標系統之相對應的 DOA 單位向量 $e_1(k,n) = R_1 \cdot e_1^{POV}(k,n)$ ，可能藉由應用下述公式而被計算出：

$$e_2(k,n) = R_2 \cdot e_2^{POV}(k,n), \quad (3)$$

於此， R 係爲座標變換矩陣，例如，

$$R_1 = \begin{bmatrix} c_{1,x} & -c_{1,y} \\ c_{1,y} & c_{1,x} \end{bmatrix}, \quad (4)$$

當在 2D 中操作且 $c_1 = [c_{1,x}, c_{1,y}]^T$ 時。爲了執行三角測量，方向向量 $d_1(k,n)$

及 $d_2(k,n)$ 可能被計算成：

$$\begin{aligned} d_1(k, n) &= d_1(k, n) e_1(k, n), \\ d_2(k, n) &= d_2(k, n) e_2(k, n), \end{aligned} \quad (5)$$

於此， $d_1(k, n) = \|d_1(k, n)\|$ and $d_2(k, n) = \|d_2(k, n)\|$ 係為在 IPLS 與兩個麥克風陣列之間的未知距離。下述方程式

$$p_1 + d_1(k, n) = p_2 + d_2(k, n) \quad (6)$$

可能解出 $d_1(k, n)$ 。最後，IPLS 之位置 $P_{\text{IPLS}}(k, n)$ 係由下述方程式得到

$$p_{\text{IPLS}}(k, n) = d_1(k, n) e_1(k, n) + p_1. \quad (7)$$

【0115】 在另一實施例中，方程式 (6) 可能解出 $d_2(k, n)$ ，而 $P_{\text{IPLS}}(k, n)$ 係採用 $d_2(k, n)$ 而類似地被計算出。

【0116】 當在 2D 中操作時，方程式 (6) 總是提供一種解決方法，除非 $e_1(k, n)$ 及 $e_2(k, n)$ 是平行的。然而，當使用兩個以上的麥克風陣列時或當在 3D 中操作時，無法在方向向量 d 並未相交時獲得一種解決方法。依據一實施例，於此情況下，最靠近所有方向向量 d 之點係被計算出，且結果可被使用作為 IPLS 之位置。

【0117】 在一實施例中，所有觀察點 p_1 、 p_2 、... 應被設置，以使由 IPLS 所發出之聲音落入相同的暫時區塊 n 中。當任兩個觀察點之間的距離 Δ 小於下述方程式時，這種需求可能簡單地被滿足

$$\Delta_{\max} = c \frac{n_{\text{FFT}}(1 - R)}{f_s}, \quad (8)$$

於此， n_{FFT} 係為 STFT 視窗長度， $0 \leq R < 1$ 指明在連續時框 (time frames) 之間的重疊，而 f_s 係為取樣頻率。舉例而言，對於一種於 48 kHz 下具有 50 % 重疊 ($R = 0.5$) 之 1024 點 STFT 而言，在陣列之間用以滿足上述需求之最大間距係為 $\Delta = 3.65$ m。

【0118】 以下將更詳細說明依據一實施例之一資訊計算模組 202 (例如一虛擬麥克風信號及側資訊計算模組)。

【0119】 圖 18 顯示依據一實施例之一種資訊計算模組 202 之圖表概觀。資訊計算單元包含一傳播補償器 500、一組合器 510 以及一頻譜加權單元 520。資訊計算模組 202 接收由一聲音事件位置估計器所估計之聲源位置估計 ssp ，一個或多個音源輸入信號係由一個或多個真實空間麥克風、一個或多個真實空間麥克風之位置 $posRealMic$ 以及虛擬麥克風之虛擬位置 $posVmic$ 所記錄。其輸出表示虛擬麥克風之一音源信號之一音源輸出信號 os 。

【0120】 圖 19 顯示依據另一實施例之一種資訊計算模組。圖 19 之資訊計算模組包含一傳播補償器 500、一組合器 510 以及一頻譜加權單元 520。傳播補償器 500 包含一傳播參數計算模組 501 及一傳播補償模組 504。組合器 510 包含一組合因子計算模組 502 及一組合模組 505。頻譜加權單元 520 包含一頻譜權重計算單元 503、一頻譜加權應用模組 506 以及一空間側資訊計算模組 507。

【0121】 爲了計算虛擬麥克風之音源信號，幾何資訊（例如真實空間麥克風之位置及方位 $121...12N$ ，虛擬空間麥克風之位置、方位及特徵 104，以及聲音事件之位置估計 205）係饋入到資訊計算模組 202，更特別是，饋入到傳播補償器 500 之傳播參數計算模組 501，饋入到組合器 510 之組合因子計算模組 502 以及饋入到頻譜加權單元 520 之頻譜權重計算單元 503。傳播參數計算模組 501、組合因子計算模組 502 與頻譜權重計算單元 503 計算使用於在傳播補償模組 504、組合模組 505 與頻譜加權應用模組 506 中之音源信號 $111...11N$ 之修改中的參數。

【0122】 在資訊計算模組 202 中，音源信號 $111...11N$ 最初可能被修正以補償由聲音事件位置與真實空間麥克風之間的不同傳播長度所得到的效果。然後，可合成這些信號以例如改善信噪比（SNR）。最後，所產生之信號接著可被頻譜加權以將虛擬麥克風之方向拾音模式與任何距離相依增益函數納入考量。這三個步驟係更詳細討論於下。

【0123】 現在更詳細說明傳播補償。在圖 20 之上部中，顯示兩個真實空間麥克風（一第一麥克風陣列 910 及一第二麥克風陣列 920），關於時頻箱 (k,n) 之一定位聲音事件 930 之位置，以及虛擬空間麥克風 940 之位置。

【0124】 圖 20 之下部說明一暫時軸線。假設一聲音事件係於時間 t_0 被發出，然後傳輸至真實及虛擬空間麥克風。抵達之時間延遲與振幅隨著距離改變，俾能使傳播長度更遠，振幅更弱且抵達之時間延遲更長。

【0125】 位於兩個真實陣列之信號是可比較的，只有如果它們之間的相對延遲 Dt_{12} 是小的話。否則，兩個信號之其中一個必須暫時被再對齊以補償相對延遲 Dt_{12} ，且儘可能地按比例縮放以補償不同的衰減。

【0126】 補償在抵達虛擬麥克風與抵達真實麥克風陣列（抵達其中一個真實空間麥克風）之間的延遲，改變了獨立於聲音事件之定位之延遲，使得其大部分應用而言是多餘的。

【0127】 回到圖 19，傳播參數計算模組 501 係適合於計算出待為每個真實空間麥克風及為每個聲音事件被校正之延遲。如果需要的話，其亦計算出待被認為補償不同的振幅降低之增益因子。

【0128】 傳播補償模組 504 係被設計成用於使用此種資訊來相應地修改音源信號。如果這些信號要平移小量的時間（相較於濾波器組（filter bank）之時窗），則簡單的相位旋轉就足夠了。如果延遲是較大的，則需要更多複合的實施例。

【0129】 傳播補償模組 504 之輸出係為表示在原始時頻域中之變化音源信號。

【0130】 以下，將參考圖 17 說明依據一實施例之關於一虛擬麥克風之傳播補償之一特定估計，圖 17 特別顯示一第一真實空間麥克風之位置 610 與一第二真實空間麥克風之位置 620。

【0131】 於現在說明的本實施例中，假設至少一第一錄製音源輸入信號，例如真實空間麥克風（例如麥克風陣列）之至少一者之一壓力信號（譬如一第一真實空間麥克風之壓力信號）是可得到的。我們將考量的麥克風稱為參考麥克風，將其位置稱為參考位置 $pref$ 並將其壓力信號稱為參考壓力信號 $P_{ref}(k,n)$ 。然而，傳播補償不僅可能只相關於一個壓力信號，而且相關於複數個或所有真實空間麥克風之壓力信號被處理。

【0132】 在由 IPLS 所發出之壓力信號 $P_{IPLS}(k,n)$ 及位於 $pref$ 之一參考麥克風之一參考壓力信號 $P_{ref}(k,n)$ 之間的關係，可以下述公式（9）表示：

$$P_{\text{ref}}(k, n) = P_{\text{IPLS}}(k, n) \cdot \gamma(k, p_{\text{IPLS}}, p_{\text{ref}}), \quad (9)$$

【0133】 一般而言，複合因子 $\gamma(k, p_a, p_b)$ 表示由一球面波從其在 p_a 中之原點至 p_b 之傳播所產生之相位旋轉及振幅衰減。然而，實際測試指出只考慮到 γ 中之振幅衰減相較於亦考慮到相位旋轉，會導致具有顯著較少假象 (artifacts) 之虛擬麥克風信號之似真實的印象。

【0134】 可以在空間中的某個點被測量之聲音能量強烈地取決於距離聲源，在圖 6 中距離聲源之位置 P_{IPLS} 之距離 r 。在多數情況中，這種依存性可藉由使用熟知之物理原理（譬如，一個點源之遠場中的聲音壓力之 $1/r$ 衰減）而被仿效成具有足夠精度。當已知一參考麥克風（譬如第一真實麥克風）與聲源之距離時，且當亦已知虛擬麥克風與聲源之距離時，則位於虛擬麥克風之位置之聲音能量可以從參考麥克風（例如第一真實空間麥克風）之信號及能量被估計出。這表示虛擬麥克風之輸出信號可藉由施加適當增益至參考壓力信號而被獲得。

【0135】 假設第一真實空間麥克風係為參考麥克風，則 $p_{\text{ref}} = p_1$ 。在圖 17 中，虛擬麥克風係位於 p_v 中。因為詳細知道圖 17 中之幾何形狀，所以可以容易決定在參考麥克風（在圖 17 中：第一真實空間麥克風）與 IPLS 之間的距離 $d_1(k, n) = \|d_1(k, n)\|$ ，以及在虛擬麥克風與 IPLS 之間的距離 $s(k, n) = \|s(k, n)\|$ ，亦即

$$s(k, n) = \|s(k, n)\| = \|p_1 + d_1(k, n) - p_v\|. \quad (10)$$

【0136】 位於虛擬麥克風之位置之聲音壓力 $P_v(k, n)$ 係藉由結合公式 (1) 及 (9) 而計算出，藉以導致

$$P_v(k, n) = \frac{\gamma(k, p_{\text{IPLS}}, p_v)}{\gamma(k, p_{\text{IPLS}}, p_{\text{ref}})} P_{\text{ref}}(k, n). \quad (11)$$

【0137】 如上所述，在某些實施例中，因子 γ 可能只考量到由於傳播之振幅衰減。例如假設聲音壓力隨著 $1/r$ 減少，則

$$P_v(k, n) = \frac{d_1(k, n)}{s(k, n)} P_{\text{ref}}(k, n). \quad (12)$$

【0138】 當公式 (1) 中之模型維持時，例如，當只有直達聲音出現

時，則公式 (12) 可以正確地重建大小資訊。然而，在純擴散聲場的情況下，例如，當並未滿足模型假設時，提出的方法在使虛擬麥克風移動遠離感測器陣列之位置時產生一信號之暗示反混響 (dereverberation)。事實上，如上所述，在擴散聲場中，我們期望大部分的 IPLS 被定位靠近兩個感測器陣列。因此，當使虛擬麥克風移動遠離這些位置時，我們很可能增加圖 17 中之距離 $s = \|s\|$ 。因此，參考壓力之大小係在依據公式 (11) 應用加權時被減少。同樣地，當使虛擬麥克風移動接近一真實聲源時，對應於直達聲音之時頻箱將被放大，以使整體音源信號將被察覺較少擴散。藉由調整公式 (12) 中之規則，吾人可以隨意地控制直達聲音放大及擴散聲音抑制。

【0139】 藉由針對第一真實空間麥克風之錄製的音源輸入信號（例如壓力信號）執行傳播補償，獲得了一第一變化音源信號。

【0140】 在實施例中，一第二變化音源信號可能藉由針對第二真實空間麥克風之一錄製的第二音源輸入信號（第二壓力信號）執行傳播補償而獲得。

【0141】 在其他實施例中，更遠的音源信號可能藉由針對更遠的真實空間麥克風之錄製的更遠的音源輸入信號（更遠的壓力信號）執行傳播補償而獲得。

【0142】 現在，更詳細說明依據一實施例之結合圖 19 中的方塊 502 及 505。假設來自複數個不同真實空間麥克風之兩個以上的音源信號已被修改以補償不同傳播路徑，用以獲得兩個以上的變化音源信號。一旦來自不同的真實空間麥克風之音源信號已被修改以補償不同傳播路徑，它們就可能被結合以改善音源品質。藉此，舉例而言，可以增加 SNR 或可以減少餘響。

【0143】 對於此組合之可能的解決方法包含：

- 加權平均，例如，考慮到 SNR，或至虛擬麥克風之距離，或由真實空間麥克風所估計之擴散。可能採用傳統的解決方法，譬如，最大比率合成 (MRC) 或等增益合成 (EQC)，或者

- 某些或所有的變化音源信號之線性組合以獲得一組合信號。變化音源信號可能在線性組合中被加權以獲得組合信號，或者

- 選擇，例如，只使用一個信號，舉例而言，依據 SNR 或距離或擴散

【0144】 如果適合的話，模組 502 之任務係用以計算供合成用之參數，其在模組 505 中被實現。

【0145】 現在，更詳細說明依據實施例之頻譜加權。對於這一點，參考圖 19 之方塊 503 及 506。於這個最後步驟，由此組合或輸入音源信號之傳播補償所產生的音源信號係依據虛擬空間麥克風之空間特徵（如由輸入 104 特別指定）及/或依據重建的幾何形狀（在方塊 205 中所提供）而在時頻域中被加權。

【0146】 對每個時頻箱而言，幾何上的重建允許我們容易地獲得相對於虛擬麥克風之 DOA，如圖 21 所示。再者，亦可輕易計算出在虛擬麥克風與聲音事件之位置之間的距離。

【0147】 接著考慮到期望的虛擬麥克風之型式，計算出關於時頻箱之權重。

【0148】 在指向性麥克風的情況下，頻譜權重可能依據一預定拾音模式而被計算。舉例而言，依據一實施例，一心型麥克風可具有一由下述函數 $g(\theta)$ 所界定之拾音模式，

$$g(\theta) = 0.5 + 0.5 \cos(\theta),$$

於此， θ 係為在虛擬空間麥克風之注視方向與從虛擬麥克風之角度來看的聲音之 DOA 之間的角度。

【0149】 另一種可能性係為藝術（非物理）的衰減函數。在某些應用上，吾人可能期望以一個大於一個特性自由場傳播之因子來抑制聲音事件遠離虛擬麥克風。為了這個目的，某些實施例導入一額外加權函數，其取決於虛擬麥克風與聲音事件之間的距離。在一實施例中，應該只有拾音在距離虛擬麥克風之某段距離（例如以公尺計）之內的聲音事件。

【0150】 相關於虛擬麥克風指向性，可以為虛擬麥克風應用任意的指向性模式。這樣做，吾人可以例如將一來源與一複合聲音場景分離。

【0151】 因為聲音之 DOA 可以在虛擬麥克風之位置 p_v 中被計算出，亦即

$$\varphi_v(k, n) = \arccos\left(\frac{s \cdot c_v}{\|s\|}\right), \quad (13)$$

於此， $\mathbf{cv}$ 係為描述虛擬麥克風之方位之一單位向量，可實現供虛擬麥克風用之任意的指向性。舉例而言，假設 $P_v(k,n)$ 表示組合信號或傳播補償變化音源信號，則此公式：

$$\tilde{P}_v(k,n) = P_v(k,n) [1 + \cos(\varphi_v(k,n))] \quad (14)$$

計算出一虛擬麥克風之輸出具有心形指向性。依此方式可潛在地被產生之方向模式係取決於位置估計之精度。

【0152】 在多個實施例中，一個或多個真實、非空間麥克風（譬如，一全向麥克風或例如一心型（cardioid）之一指向性麥克風）係被設置於除了真實空間麥克風以外之聲音場景，用以更進一步改善圖 8 中之虛擬麥克風音源信號 105 之聲音品質。這些麥克風並非用以蒐集任何幾何資訊，而是只提供一個更清晰的音源信號。這些麥克風可能被置於比空間麥克風更接近聲源。於此情況下，依據一實施例，真實、非空間麥克風之音源信號與它們的位置係只回饋至圖 19 之傳播補償模組 504 以供處理，以取代真實空間麥克風之音源信號。傳播補償接著相關於一個或多個非空間麥克風之位置而為非空間麥克風之一個或多個錄製的音源信號作處理。藉此，藉由使用額外非空間麥克風來實現一實施例。

【0153】 在更進一步的實施例中，實現了虛擬麥克風之空間側資訊之計算。為了計算麥克風之空間側資訊 106，圖 19 之資訊計算模組 202 包含一空間側資訊計算模組 507，其適合於接收聲源之位置 205 與虛擬麥克風之位置、方位及特徵 104 作為輸入。在某些實施例中，依據需要被計算之空間側資訊 106，虛擬麥克風音源信號 105 亦可被納入考量作為輸入至空間側資訊計算模組 507。

【0154】 空間側資訊計算模組 507 之輸出係為虛擬麥克風之側資訊 106。從虛擬麥克風的角度來看，這種側資訊可以例如是關於每個時頻箱 (k,n) 之 DOA 或聲音之擴散。另一種可能的側資訊可以例如是有效聲音強度向量 $I_a(k,n)$ ，其將已在虛擬麥克風之位置中被測量。現在將說明可以如何推導出這些參數。

【0155】 依據一實施例，實現了關於虛擬空間麥克風之 DOA 估計。

基於虛擬麥克風之一位置向量以及基於如圖 22 所顯示的聲音事件之一位置向量，資訊計算模組 120 係適合於估計虛擬麥克風之到達方向以作為空間側資訊。

【0156】 圖 22 說明一種從虛擬麥克風的角度推導出聲音之 DOA 之可能方式。由圖 19 中之方塊 205 所提供之聲音事件之位置可以利用一位置向量 $r(k,n)$ (聲音事件之位置向量) 來作關於每個時頻箱 (k,n) 之描述。同樣地，提供作為圖 19 中之輸入 104 之虛擬麥克風之位置可以利用一位置向量 $s(k,n)$ (虛擬麥克風之位置向量) 來作描述。虛擬麥克風之注視方向可藉由一向量 $v(k,n)$ 作描述。相對於虛擬麥克風之 DOA 由 $a(k,n)$ 得到。其表示在 v 與聲音傳播路徑 $h(k,n)$ 之間的角度。 $h(k,n)$ 可以藉由採用下述公式而計算出：

$$h(k,n) = s(k,n) - r(k,n)。$$

【0157】 現在可例如經由 $h(k,n)$ 及 $v(k,n)$ 之內積 (dot product) 之定義計算關於每個 (k,n) 之期望 DOA $a(k,n)$ ，亦即

$$a(k,n) = \arccos (h(k,n) \cdot v(k,n) / (\|h(k,n)\| \|v(k,n)\|))。$$

【0158】 在另一實施例中，基於虛擬麥克風之一位置向量以及基於如由圖 22 所顯示的聲音事件之一位置向量，資訊計算模組 120 可能適合於估計位於虛擬麥克風之有效聲音強度以作為空間側資訊。

【0159】 從如上所界定的 DOA $a(k,n)$ ，我們可推導出位於虛擬麥克風之位置之有效聲音強度 $I_a(k,n)$ 。對於這一點，假設圖 19 中之虛擬麥克風音源信號 105 對應至一全向麥克風之輸出，例如，我們假設虛擬麥克風為一全向麥克風。此外，假設圖 22 中之注視方向 v 平行於座標系統之 x 軸線。因為期望的有效聲音強度向量 $I_a(k,n)$ 描述了經由虛擬麥克風之位置之淨能量流，所以我們可例如依據下述公式計算出 $I_a(k,n)$ ：

$$I_a(k,n) = - (1/2 \rho) |P_v(k,n)|^2 * [\cos a(k,n), \sin a(k,n)]^T,$$

於此 $[]^T$ 表示一轉置向量， ρ 係為空氣密度，而 $P_v(k,n)$ 係為由虛擬空間麥克風所測量之聲音壓力，例如，圖 19 中之方塊 506 之輸出 105。

【0160】 如果有效強度向量應該被計算表示在一般的座標系統中但仍然位於虛擬麥克風之位置，則可應用下述公式：

$$I_a(k, n) = (1/2 \rho) |P_v(k, n)|^2 h(k, n) / \|h(k, n)\|.$$

【0161】 聲音之擴散表示如何在一既定時頻槽中擴散此聲場（參見，譬如[2]）。擴散係以一數值 ψ 表示，其中 $0 \leq \psi \leq 1$ 。1 之擴散表示一聲場之總聲場能量是完全地擴散。這種資訊例如在空間聲之再生上是重要的。傳統上，擴散係於置放有一麥克風陣列之空間中之特定點被計算出。

【0162】 依據一實施例，擴散可能作為一額外參數被計算成為虛擬麥克風（VM）所產生之側資訊，其可將虛擬麥克風隨意置於聲音場景中之一任意位置。藉此，一種又計算出位於一虛擬麥克風之一虛擬位置之除音源信號以外的擴散之設備可以被看作是一虛擬 DirAC 前端，其乃因為可能為聲音場景中之一任意點產生一 DirAC 流動（亦即一音源信號、到達方向以及擴散）。DirAC 流可能更進一步被處理、儲存、傳輸以及在任意的多揚聲器裝備上回放。於此情況下，聆聽者體驗此聲音場景，猶如他或她係身在由虛擬麥克風而特別指定之位置且係朝由其方位所決定之方向注視。

【0163】 圖 23 顯示依據一實施例之一種包含用以計算位於虛擬麥克風之擴散之一擴散計算單元 801 之資訊計算方塊。資訊計算方塊 202 係適合於接收輸入 111 至 11N，除了圖 14 之輸入以外，其亦包含位於真實空間麥克風之擴散。使 $\psi(SM1)$ 至 $\psi(SMN)$ 表示這些數值。這些額外輸入係回饋至資訊計算模組 202。擴散計算單元 801 之輸出 103 係為於虛擬麥克風之位置所計算之擴散參數。

【0164】 一實施例之一擴散計算單元 801 係顯示於圖 24 中，藉以提供更詳細說明。依據一實施例，估計出位於 N 個空間麥克風之每一個之直達及擴散聲音之能量。然後，藉由使用關於 IPLS 之位置之資訊，以及關於空間以及虛擬麥克風之位置之資訊，獲得了位於虛擬麥克風之位置之這些能量之 N 個估計。最後，可結合估計以改善估計精度，且可輕易計算出位於虛擬麥克風之擴散參數。

【0165】 使 $E_{dir}^{(SM1)}$ 至 $E_{dir}^{(SMN)}$ 以及 $E_{diff}^{(SM1)}$ 至 $E_{diff}^{(SMN)}$ 表示關於由能量分析單元 810 所計算之 N 個空間麥克風之直達及擴散聲音之能量之估計。如果 P_i 係為複合壓力信號且 ψ_i 係為關於第 i 個空間麥克風之擴散，則可譬如依據下述公式計算出能量：

$$E_{\text{dir}}^{(\text{SM}i)} = (1 - \Psi_i) \cdot |P_i|^2$$

$$E_{\text{diff}}^{(\text{SM}i)} = \Psi_i \cdot |P_i|^2$$

【0166】 擴散聲音之能量在所有位置中應該是相等的，因此，位於虛擬麥克風之擴散聲音能量 $E_{\text{diff}}^{(\text{VM})}$ 之估計可以譬如依據下述公式，只藉由例如在一擴散組合單元 820 中平均 $E_{\text{diff}}^{(\text{SM}1)}$ 至 $E_{\text{diff}}^{(\text{SM}N)}$ 而被計算出：

$$E_{\text{diff}}^{(\text{VM})} = \frac{1}{N} \sum_{i=1}^N E_{\text{diff}}^{(\text{SM}i)}$$

【0167】 估計 $E_{\text{diff}}^{(\text{SM}1)}$ 至 $E_{\text{diff}}^{(\text{SM}N)}$ 之一種更有效的組合可以藉由考慮到估計器之變異數（例如藉由考慮到 SNR）而被實現。

【0168】 直達聲音之能量取決於由於傳播到達來源之距離。因此， $E_{\text{dir}}^{(\text{SM}1)}$ 至 $E_{\text{dir}}^{(\text{SM}N)}$ 可能被修正以將此納入考量。這可能例如藉由一直達聲音傳播調整單元 830 而實現。舉例而言，如果假設指向聲場之能量衰減在距離平方是 1，則可能依據下述公式計算出關於位於第 i 個空間麥克風之虛擬麥克風之直達聲音之估計：

$$E_{\text{dir},i}^{(\text{VM})} = \left(\frac{\text{distance SM}i - \text{IPLS}}{\text{distance VM} - \text{IPLS}} \right)^2 E_{\text{dir}}^{(\text{SM}i)}$$

【0169】 類似於擴散組合單元 820，於不同的空間麥克風所獲得之直達聲音能量之估計可以例如藉由一直達聲音組合單元 840 而合成。結果係為 $E_{\text{dir}}^{(\text{VM})}$ ，例如，關於位於虛擬麥克風之直達聲音能量之估計。位於虛擬麥克風 $\Psi(\text{VM})$ 之擴散可能例如依據下述公式而譬如藉由一擴散子計算器 850 而被計算出：

$$\Psi^{(\text{VM})} = \frac{E_{\text{diff}}^{(\text{VM})}}{E_{\text{diff}}^{(\text{VM})} + E_{\text{dir}}^{(\text{VM})}}$$

【0170】 如上所述，在某些情況下，藉由一聲音事件位置估計器而被實現之聲音事件位置估計會失敗，例如，在一錯誤到達方向估計的情況下。圖 25 顯示這種方案。在這些情況下，不管於不同空間麥克風所估計以

及如被接收為輸入 111 至 11N 之擴散參數為何，關於虛擬麥克風 103 之擴散可能被設定到 1（亦即，完全擴散），其乃因為沒有空間相干再生（spatially coherent reproduction）是可能的。

【0171】 此外，可能考量位於 N 個空間麥克風之 DOA 估計之可靠度。這可能例如從 DOA 估計器之變異數或 SNR 的觀點被表示。這種資訊可能藉由擴散子計算器 850 被納入考量，俾能使 VM 擴散 103 可以在 DOA 估計是不可靠的情況下被人為地增加。事實上，結果是位置估計 205 亦將是不可靠的。

【0172】 圖 26 顯示一設備 991，用以依據一實施例之產生一虛擬輸出信號。用以產生一虛擬輸出信號之設備 991 包含一種依據上述實施例之其中一個（其包含一麥克風位置計算器 993）而用於定位麥克風之設備 992。再者，用以產生一虛擬輸出信號之設備包含一種用以依據上述實施例之其中一個來產生一音源輸出信號之設備 994。由用以產生一音源輸出信號之設備 994 所產生之輸出信號係為虛擬輸出信號 vos。用以定位麥克風之設備 991 之麥克風位置計算器 992 係被設計成用以計算一麥克風之位置以作為一計算出的麥克風位置 cmp。用以產生一音源輸出信號之設備 994 係被設計成用以模擬位於計算出的麥克風位置（由用於定位麥克風之設備 992 所計算出）之一虛擬麥克風之一錄製。藉此，用於定位麥克風之設備 992 計算關於用以產生一音源輸出信號之設備 994 之虛擬麥克風之虛擬位置。

【0173】 雖然已在設備之上下文中說明某些實施樣態，但顯然這些方面亦表示相對應的方法之說明，於此之方塊或裝置係對應至一方法步驟或一方法步驟之特徵。依此類推，在一方法步驟之上下文中所說明的實施樣態亦表示一對應方塊或項目之說明或對應設備之特徵。

【0174】 所發明的分解信號可儲存在一數位儲存媒體上，或可在一傳輸媒介上傳輸，例如一無線傳輸媒介或一有線傳輸媒介，例如網際網路。

【0175】 依據某些實施需求，本發明之實施例可在硬體或軟體中實施。實施例可藉由使用一數位儲存媒體（譬如軟碟、DVD、CD、ROM、PROM、EPROM、EEPROM 或快閃記憶體）而執行，數位儲存媒體具有儲存於其上之電子可讀取的控制信號，其與一可程式化電腦系統合作（或能夠合作），以執行各種方法。

【0176】 依據本發明之某些實施例包含一非暫時性資料載體，其具有能夠與一可程式化電腦系統合作之電子可讀取的控制信號，以執行於此所說明之其中一種方法。

【0177】 一般而言，本發明之實施例可被實施作為具有一程式碼之一電腦程式產品，當電腦程式產品在一電腦上運行時，程式碼對於執行其中一種方法是有效的。程式碼可譬如被儲存在一機器可讀取的載體上。

【0178】 其他實施例包含儲存在一機器可讀取的載體上之電腦程式，用以執行於此所說明之其中一種方法。

【0179】 換言之，本發明方法之一實施例因此係為一種具有一程式碼之電腦程式，用於在電腦程式在一電腦上運行時，執行於此所說明之其中一種方法。

【0180】 因此，本發明方法之更進一步的實施例係為資料載體（或數位儲存媒體或電腦可讀取的媒體），其包含記錄於其上之電腦程式，用於以執行於此所說明之其中一種方法。

【0181】 因此，本發明方法之更進一步的實施例係為表示電腦程式之一資料流或一序列之信號，用以執行於此所說明之其中一種方法。資料流或此序列之信號可譬如具體形成以經由一資料通訊連接（譬如經由網際網路）來傳輸。

【0182】 更進一步的實施例包含一處理機構（譬如電腦或可程式化邏輯元件），其具體形成或適合於執行於此所說明之其中一種方法。

【0183】 更進一步的實施例包含上面裝設有電腦程式之電腦，用於執行於此所說明之其中一種方法。

【0184】 在某些實施例中，可能使用一可程式化邏輯元件（譬如現場可程式化閘陣列）以執行於此所說明之方法之某些或所有功能。在某些實施例中，一現場可程式化閘陣列可協同一微處理器，以便執行於此所說明之其中一種方法。一般而言，這些方法最好是藉由任何硬體設備而執行。

【0185】 上述實施例僅為了本發明之原理而說明。吾人理解到，其他熟習本項技藝者將明白於此所說明之配置與細節之修改及變化。因此，意圖只受限於即將到來的專利申請專利範圍之範疇，而不受限於經由實施例之記述及說明於此所提出的特定細節。

【0186】 文獻：

【0187】 [1] Michael A. Gerzon。多聲道廣播及視頻中之高傳真身歷聲響複製 (Ambisonics in multichannel broadcasting and video)。J. 音源工程學會, 33 (11) :859-871, 1985 年。

【0188】 [2] V. Pulkki, "空間聲再生及立體聲升混中之音源編碼定向音源編碼 (Directional audio coding in spatial sound reproduction and stereo upmixing)", 於 AES 第 28 屆國際會議論文集, 第 251-258 頁, 瑞典 Piteå, 2006 年 6 月 30 日-7 月 2 日。

【0189】 [3] V. Pulkki, "利用定向音源編碼之空間聲再生 (Spatial sound reproduction with directional audio coding)", J. 音源工程學會, 第 55 卷, 第 6 號, 第 503-516 頁, 2007 年 6 月。

【0190】 [4] C. Faller。關於空間音源編碼器之麥克風前端 (Microphone front-ends for spatial audio coders)。於 AES 第 125 屆國際公約的程序, 舊金山, 2008 年 10 月。

【0191】 [5] M. Kallinger, H. Ochsenfeld, G. Del Galdo, F. Küch, D. Mahne, R. Schultz-Amling. 及 O. Thiergart, "一種用以定向音源編碼之空間濾波方法 (A spatial filtering approach for directional audio coding)", 於音源工程協會大會 126, 德國慕尼黑, 2009 年 5 月。

【0192】 [6] R. Schultz-Amling, F. Küch, O. Thiergart, 及 M. Kallinger, "基於一參數聲場表現之聲學 1 縮放 (Acoustical zooming based on a parametric sound field representation)", 於音源工程協會大會 128, 英國倫敦, 2010 年 5 月。

【0193】 [7] J. Herre, C. Falch, D. Mahne, G. Del Galdo, M. Kallinger, 及 O. Thiergart, "結合空間音源對象編碼及 DirAC 技術之互動線上討論 (Interactive teleconferencing combining spatial audio object coding and DirAC technology)", 於音源工程協會大會 128, 英國倫敦, 2010 年 5 月。

【0194】 [8] E. G. Williams, 傅立葉音響：聲音輻射及近場聲音全像術 (Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography), 科學出版社, 1999 年。

【0195】 [9] A.Kuntz 及 R.Rabenstein, "從圓形測量的外推波場的限

制 (Limitations in the extrapolation of wave fields from circular measurements) " , 於第 15 屆歐洲信號處理會議 (EUSIPCO 2007) , 2007 年。

【0196】 [10] A. Walther 及 C. Faller, "藉由使用 b-格式錄製之隔開的麥克風陣列之線性模擬 (Linear simulation of spaced microphone arrays using b-format recordings) " , 於音源工程協會大會 128 , 英國倫敦, 2010 年 5 月。

【0197】 [11] US61/287,596 : 一種用以將一第一參數空間音源信號轉換成一第二參數空間音源信號之設備及方法 (An Apparatus and a Method for Converting a First Parametric Spatial Audio Signal into a Second Parametric Spatial Audio Signal) 。

【0198】 [12] S. Rickard 及 Z. Yilmaz, "針對言語之接近 W-分離正交 (On the approximate W-disjoint orthogonality of speech) " , 於語音及信號處理, 2002. ICASSP 2002. 2002 年 4 月之 IEEE 國際學術研討會, 第 1 卷。

【0199】 [13] R. Roy, A. Paulraj 以及 T. Kailath, "藉由子空間旋轉法之到達方向估計-ESPRIT", 在 1986 年 4 月之美國加利福尼亞州史丹佛大學之語音與信號處理 (ICASSP) 之 IEEE 國際學術研討會中 (IEEE 音響國際會議, 語音和信號處理 (ICASSP), 史丹福, 美國加州, 1986 年 4 月) 。

【0200】 [14] R. Schmidt, "多重發射器場地及信號參數估計 (Multiple emitter location and signal parameter estimation) " , 關於天線與傳播之 IEEE 學報, 第 34 卷, 第 3 號, 第 276-280 頁, 1986 年。

【0201】 [15] J. Michael Steele, "平面上之隨機樣品之最佳三角測量 (Optimal Triangulation of Random Samples in the Plane) " , 概率的史冊, 第 10 卷, 第 3 號 (1982 年 8 月), 第 548-553 頁。

【0202】 [16] F. J. Fahy, 聲音強度 (Sound Intensity), 埃塞克斯 : Elsevier 科學出版有限公司, 1989 年。

【0203】 [17] R. Schultz-Amling, F. Küch, M. Kallinger, G. Del Galdo, T. Ahonen 及 V. Pulkki, "藉由使用定向音源編碼之空間音源之分析及再生之平面麥克風陣列處理 (Planar microphone array processing for the analysis and reproduction of spatial audio using directional audio coding) " , 於音源工程協會大會 124 , 荷蘭阿姆斯特丹, 2008 年 5 月。

【0204】 [18] M. Kallinger, F. Küch, R. Schultz-Amling, G. Del Galdo, T. Ahonen 及 V. Pulkki, "使用麥克風陣列提高方向估計以供定向音源編碼用 (Enhanced direction estimation using microphone arrays for directional audio coding)" ; 於免持語音通信和麥克風陣列, 2008. HSCMA 2008, May 2008 年 5 月, 第 45-48 頁。

【0205】 [19] R. K. Furness, "高傳真身歷聲響複製-概要 (Ambisonics - An overview)" , 於 AES 第 8 屆國際學術研討會, 1990 年 4 月, 第 81-189 頁。

【0206】 [20] Giovanni Del Galdo, Oliver Thiergart, Tobias Weller, 以及 E. A. P. Habets。藉由使用由配置的陣列所蒐集之幾何資訊產生虛擬麥克風信號 (Generating virtual microphone signals using geometrical information gathered by distributed arrays)。於免持語音通信和麥克風陣列的第三次聯合研討會 (HSCMA '11) 英國愛丁堡, 2011 年 5 月。

【0207】 [21] Ville Pulkki。利用定向音源編碼之空間聲再生 (Spatial sound reproduction with directional audio coding)。J. 音源工程學會, 55 (6) :503-516, 2007 年 6 月。

【符號說明】

【0208】

- 10、21：空間能量分佈決定器
- 103：VM 擴散/虛擬麥克風/輸出
- 104：輸入/位置、方位及特徵
- 105：輸出/虛擬麥克風音源信號
- 106：輸出/空間側資訊
- 11：定位麥克風之設備
- 110、201：聲音事件位置估計器
- 111...11N：輸入/音源信號/真實空間麥克風
- 120：資訊計算模組
- 121...12N：輸入
- 13：有效度量標準
- 15：位置/輸出

- 151、152、161、162、171、172：麥克風陣列
- 153：真實聲源
- 16：方位
- 163：揚聲器
- 165：鏡像源
- 20、22：空間資訊估計器
- 202：(側) 資訊計算方塊/模組
- 205：位置/位置估計/聲音事件位置估計器
- 23、spd：空間能量密度 (SPD)
- 31：空間能量分佈主處理單元
- 32：空間能量分佈後處理單元
- 41：聲音場景中心估計器
- 410、420：真實空間麥克風陣列
- 42、993：麥克風位置計算器
- 43：輸出/聲音場景之中心
- 430：第一線
- 44：(虛擬) 麥克風位置/方位計算器
- 440：第二線
- 45：方位決定器
- 500：傳播補償器
- 501：傳播參數計算模組
- 502：組合因子計算模組/模組
- 503：頻譜權重計算單元
- 504：傳播補償模組
- 505：組合模組/模組
- 506：頻譜加權應用模組
- 507：空間側資訊計算模組
- 510：第一空間麥克風/組合器
- 520：第二空間麥克風/頻譜加權單元
- 530、540：單位向量

550、560：線
 610、620、 $P_{IPLS}(k,n)$ 、posRealMic：位置
 801：擴散計算單元
 810：能量分析單元
 820：擴散組合單元
 830：直達聲音傳播調整單元
 840：直達聲音組合單元
 850：擴散子計算器
 91...9N：輸入/有效聲音源
 910：第一麥克風陣列
 920：第二麥克風陣列
 930：定位聲音事件
 940：虛擬（空間）麥克風
 991、994：設備
 992：麥克風位置計算器/定位麥克風之設備
 $a(k,n)$ 、 $a_1(k,n)$ 、 $a_2(k,n)$ ：方位角角度
 aspi：聲源空間資訊
 cmp：麥克風位置
 COG：聲音場景之重心
 dCOG-VM：在 COG 及 VM 之間的距離
 di1：第一方向資訊
 di2：第二方向資訊
 Dt12：相對延遲
 $h(k,n)$ ：聲音傳播路徑
 is1：錄製之音源輸入信號
 os：音源輸出信號
 posImic：第一真實麥克風位置
 posVmic：虛擬位置
 r：位置向量
 ssi：聲音源資訊

ssp：聲音源位置/聲源位置估計

t0：時間

vos：虛擬輸出信號

α ：VM 開口角度

(k,n)：時頻箱/聲音事件之位置

申請專利範圍

1、一種用於計算麥克風之位置之設備，包含：

一空間能量密度決定器（10；21），用以依據表示位於一環境中之一個或多個聲音源之一個或多個能量數值及一個或多個位置數值之聲音源資訊，來決定一個表示該環境之複數個場地之能量數值之空間能量密度；以及

一空間資訊估計器（20；22），用以依據該空間能量密度來估計麥克風之位置，

其中，該空間資訊估計器（20；22）包含一聲音場景中心估計器（41），用以估計該環境中之一聲音場景之一中心之一位置，

其中，該空間資訊估計器（20；22）更包含一麥克風位置計算器（42；44），用以依據該聲音場景之該中心之該位置決定出該麥克風之該位置，

其中該空間資訊估計器（20；22）包含一方位決定器（45），用以決定該麥克風之一方位，其中該方位決定器（45）係適合於依據該空間能量密度決定該麥克風之該方位，

其中該空間能量密度決定器（10；21）係適合於藉由應用下述公式來決定該空間能量密度

$$\Gamma(x,y,k,n) = \sum_{i=1}^N \text{power } i(k,n) \cdot g(\gamma_i, x - x_{ESSi}, y - y_{ESSi}, k, n),$$

用以為一時頻箱(k,n)計算出關於該環境中之該複數個場地之該些能量數值 x,y,k,n ，其中 k 表示該頻率指數及 n 表示該時間指數，其中 N 表示該些聲音源之一些，其中 $x、y$ 表示該複數個場地之其中一個之座標，其中能量 $i(k,n)$ 表示位於一時頻箱(k,n)之第 i 個有效聲音源之

該能量數值，其中 x_{ESSi} 、 y_{ESSi} 表示該第 i 個聲音源之座標，其中 γ_i 係為一純量值且其中 g 係為取決於 x 、 y 、 x_{ESSi} 、 y_{ESSi} 、 k 、 n 及 γ_i 之一函數，當該環境係為一種二維環境時，或者

其中該空間能量密度決定器(10；21)係適合於藉由應用下述公式來決定該空間能量密度

$$\Gamma(x,y,z,k,n) = \sum_{i=1}^N \text{power } i(k,n) \cdot g(\gamma_i, x - x_{ESSi}, y - y_{ESSi}, z - z_{ESSi}, k, n)$$

用以為一時頻箱(k,n)計算出關於該環境中之該複數個場地之該些能量數值 x,y,z,k,n ，其中 k 表示該頻率指數及 n 表示該時間指數，其中 N 表示該些聲音源之一些，其中 x 、 y 、 z 表示該複數個場地之其中一個之座標，其中能量 $i(k,n)$ 表示位於一時頻箱(k,n)之第 i 個聲音源之該能量數值，其中 x_{ESSi} 、 y_{ESSi} 、 z_{ESSi} 表示該第 i 個聲音源之座標，其中 γ_i 係為一純量值，且其中 g 係為取決於 x 、 y 、 z 、 x_{ESSi} 、 y_{ESSi} 、 z_{ESSi} 、 k 、 n 及 γ_i 之一函數，當該環境係為一種三維環境時。

2、如申請專利範圍第 1 項所述之設備，其中該方位決定器(45)係適合於決定該麥克風之該方位，以使該麥克風係朝向該聲音場景之該中心被配向。

3、如申請專利範圍第 1 項所述之設備，其中該麥克風位置計算器(42；44)係適合於計算該麥克風之該位置，其中該麥克風係為一虛擬空間麥克風。

4、如申請專利範圍第 1 項所述之設備，其中該聲音場景中心估計器(41)係適合於計算該空間能量密度之一重心，用以估計該聲音場景之該中心。

5、如申請專利範圍第 1 項所述之設備，

其中，該聲音場景中心估計器(41)係被設計成用以依據該空間能量密度來決定一能量延遲分佈圖，並依據關於該環境中之複數個場地之每一個之該能量延遲分佈圖來決定一均方根延遲，且

其中該聲音場景中心估計器（41）係被設計成用以決定該複數個場地之該場地以作為該聲音場景之該中心，其具有該複數個場地之該些均方根延遲之一最小均方根延遲。

6、如申請專利範圍第 1 項所述之設備，

其中該麥克風位置計算器（42；44）係適合於決定經由該環境中之該聲音場景之該中心之複數條線之一條寬度最寬的線，

其中經由該聲音場景之該中心之該複數條線之每一條係與一能量寬度相關，且其中該寬度最寬的線係被定義為經由具有一最大能量寬度之該聲音場景之該中心之該複數條線之該條線，

其中該麥克風位置計算器（42；44）係適合於決定該麥克風之該位置，以使經由該聲音場景之該中心與該麥克風之該位置之一第二線係垂直於該寬度最寬的線。

7、如申請專利範圍第 6 項所述之設備，其中該複數條線之一條考量線之能量寬度表示在該考量線上之一段之一最大寬度，以使限制該段之該段之一第一點，且以使限制該段之該段之一不同的第二點，兩者都具有由該空間能量密度所表示之一能量數值，該空間能量密度大於或等於一預定能量數值。

8、如申請專利範圍第 1 項所述之設備，其中該麥克風位置計算器（42；44）係被設計成用以將一奇異值分解應用至一具有複數行之矩陣，其中該矩陣之該些行表示在相對於該聲音場景之該中心之該環境中的複數個場地之位置，且

其中該矩陣之該些行只表示具有由大於一預定閾值之該空間能量密度

所表示之複數個能量數值之該些場地之位置，或該矩陣之該些行只表示具有由大於或等於一預定閾值之該空間能量密度所表示之複數個能量數值之該些場地之位置。

9、一種用以產生一虛擬輸出信號之設備（991），包含：

一種如申請專利範圍第 1 項之定位麥克風之設備（992），其中該定位麥克風之設備之該麥克風位置計算器（993）係被設計成用以計算一麥克風之該位置以作為一計算出的麥克風位置，以及

一設備（994），用以產生一音源輸出信號以作為該虛擬輸出信號，用以模擬位於該計算出的麥克風位置之一虛擬麥克風之一錄製，其中用以產生一音源輸出信號之該設備包含：

一聲音事件位置估計器（110），用以估計表示該環境中之一聲音源之一位置之一聲音源位置，該聲音源發出一聲波，其中該聲音事件位置估計器（110）係適合於依據一第一方向資訊，以及依據一第二方向資訊來估計該聲音源位置，該第一方向資訊由位於一環境中之一第一真實麥克風位置之一第一真實空間麥克風所提供，該第二方向資訊由位於該環境中之一第二真實麥克風位置之一第二真實空間麥克風所提供；以及

一資訊計算模組（120），用以依據一個由該第一真實空間麥克風所錄製之第一錄製之音源輸入信號、依據該第一真實麥克風位置以及依據該計算出的麥克風位置來產生該音源輸出信號。

10、一種用於計算麥克風之位置之方法，包含：

依據表示位於一環境中之一個或多個聲音源之一個或多個能量數值及一個或多個位置數值之聲音源資訊，來決定一個表示關於該環境之複數個場地之能量數值之空間能量密度；

依據該空間能量密度來計算該麥克風之該位置；及

決定該麥克風之方位，

其中，依據該空間能量密度來計算該麥克風之該位置之步驟，係藉由計算該環境中之該聲音場景之該中心，以及藉由依據該聲音場景之該中心的位置決定麥克風的位置來執行，

其中該空間能量密度決定器（10；21）係適合於藉由應用下述公式來決定該空間能量密度

$$\Gamma(x,y,k,n) = \sum_{i=1}^N \text{power } i(k,n) \cdot g(\gamma_i, x - x_{ESSi}, y - y_{ESSi}, k, n),$$

用以為一時頻箱(k,n)計算出關於該環境中之該複數個場地之該些能量數值 x,y,k,n，其中 k 表示該頻率指數及 n 表示該時間指數，其中 N 表示該些聲音源之一些，其中 x、y 表示該複數個場地之其中一個之座標，其中能量 i(k,n)表示位於一時頻箱(k,n)之第 i 個有效聲音源之該能量數值，其中 x_{ESSi}、y_{ESSi} 表示該第 i 個聲音源之座標，其中 γ_i 係為一純量值且其中 g 係為取決於 x、y、x_{ESSi}、y_{ESSi}、k、n 及 γ_i 之一函數，當該環境係為一種二維環境時，或者

其中該空間能量密度決定器(10；21)係適合於藉由應用下述公式來決定該空間能量密度

$$\Gamma(x,y,z,k,n) = \sum_{i=1}^N \text{power } i(k,n) \cdot g(\gamma_i, x - x_{ESSi}, y - y_{ESSi}, z - z_{ESSi}, k, n)$$

用以為一時頻箱(k,n)計算出關於該環境中之該複數個場地之該些能量數值 x,y,z,k,n，其中 k 表示該頻率指數及 n 表示該時間指數，其中 N 表示該些聲音源之一些，其中 x、y、z 表示該複數個場地之其中一個之座標，其中能量 i(k,n)表示位於一時頻箱(k,n)之第 i 個聲音源之該

能量數值，其中 x_{ESSi} 、 y_{ESSi} 、 z_{ESSi} 表示該第 i 個聲音源之座標，其中 γ_i 係為一純量值，且其中 g 係為取決於 x 、 y 、 z 、 x_{ESSi} 、 y_{ESSi} 、 z_{ESSi} 、 k 、 n 及 γ_i 之一函數，當該環境係為一種三維環境時。

- 11、一種用於執行如申請專利範圍第 10 項所述之方法之電腦程式，其中該電腦程式係在一電腦或處理器上執行。

圖式

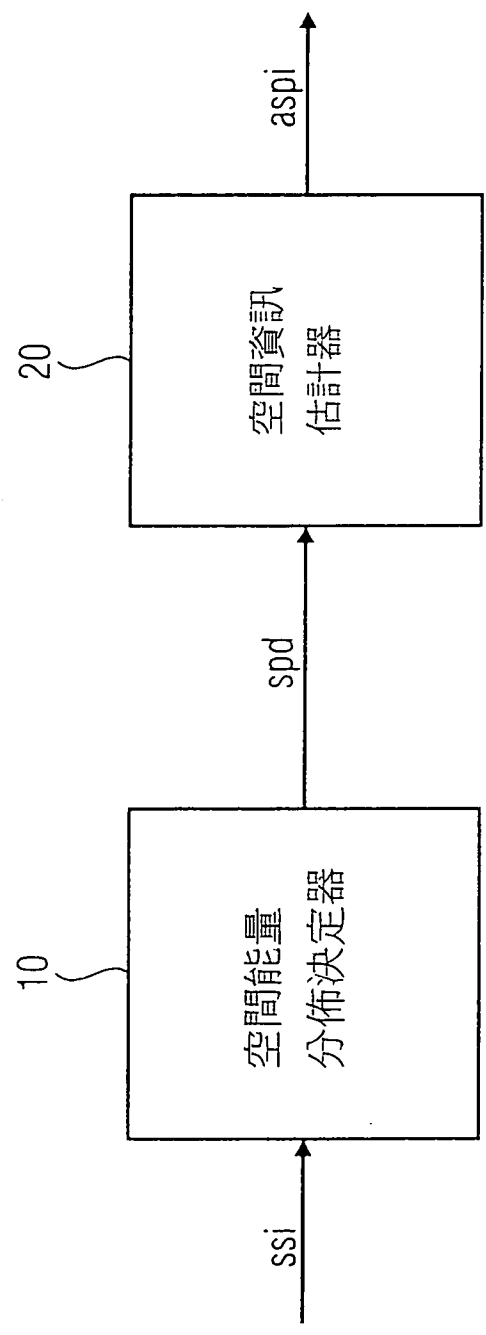


圖 1

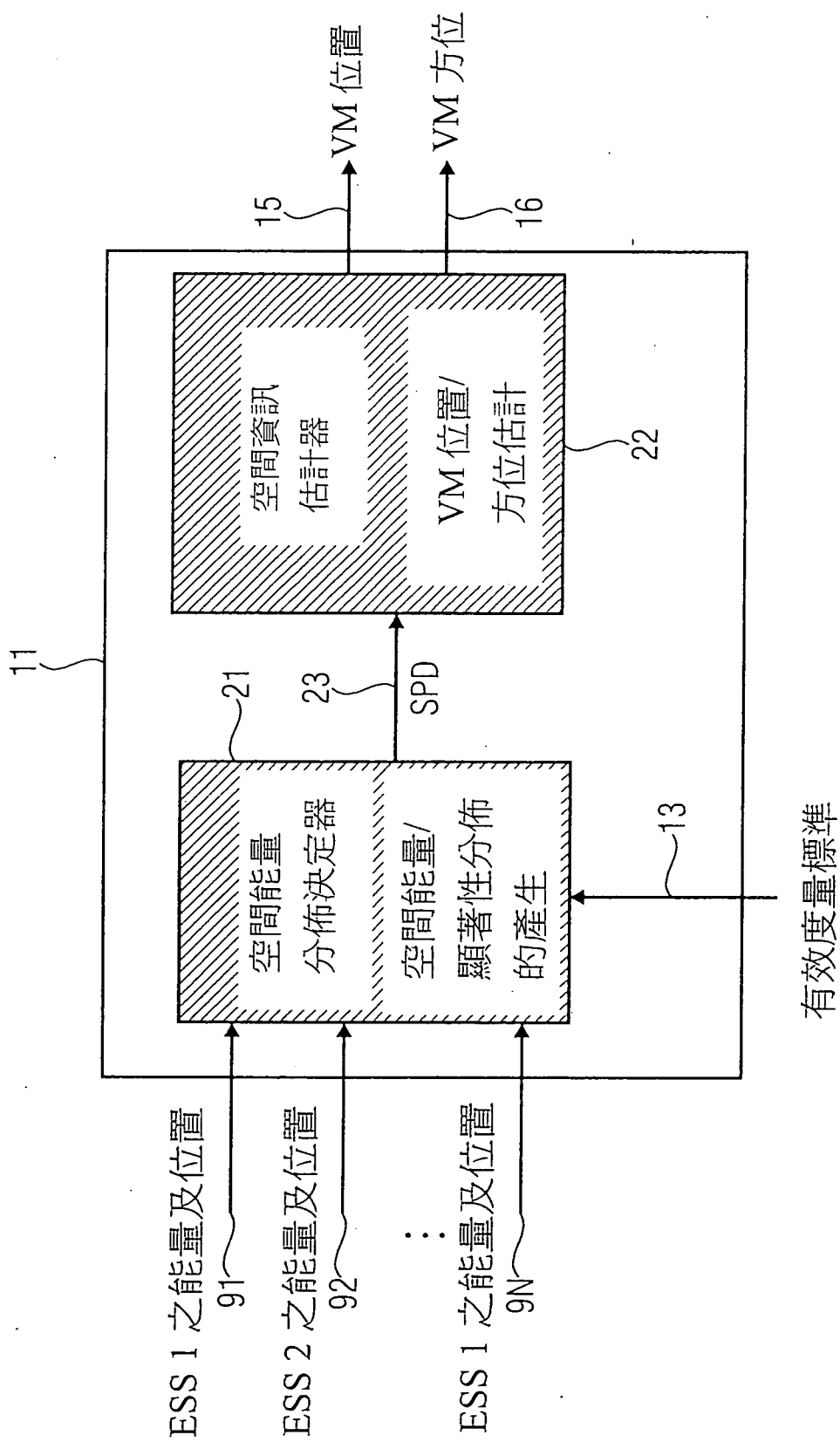


圖 2

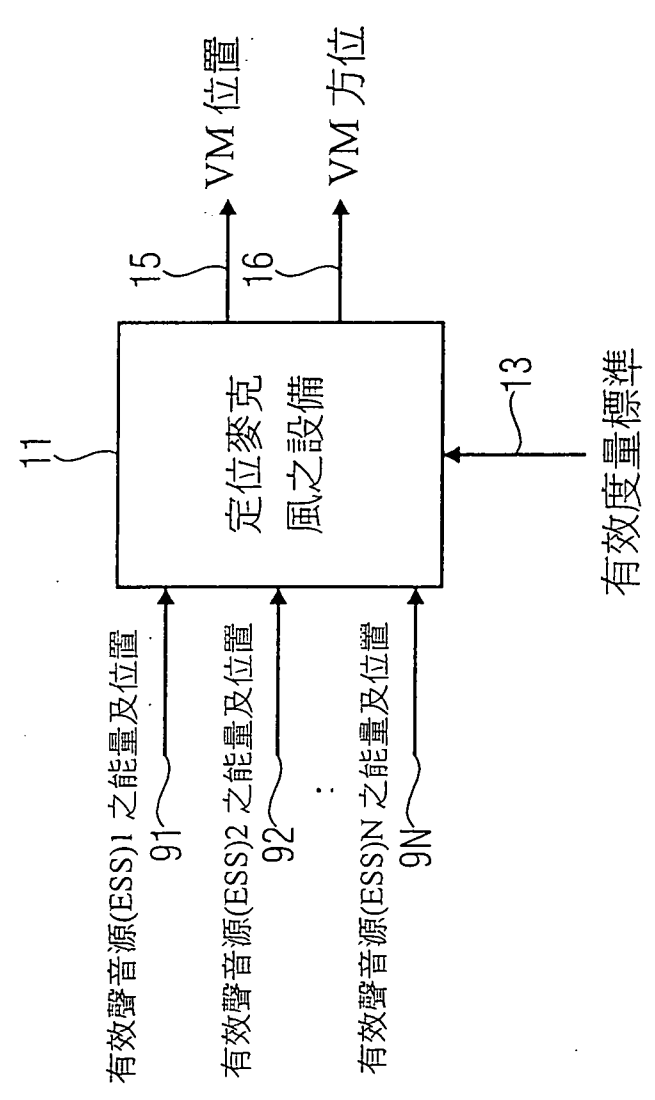


圖 3

全向虛擬麥克風

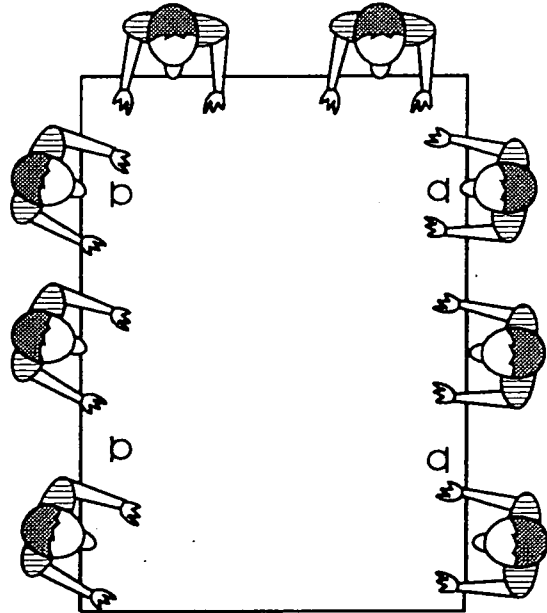


圖 4a

α 全向虛擬麥克風

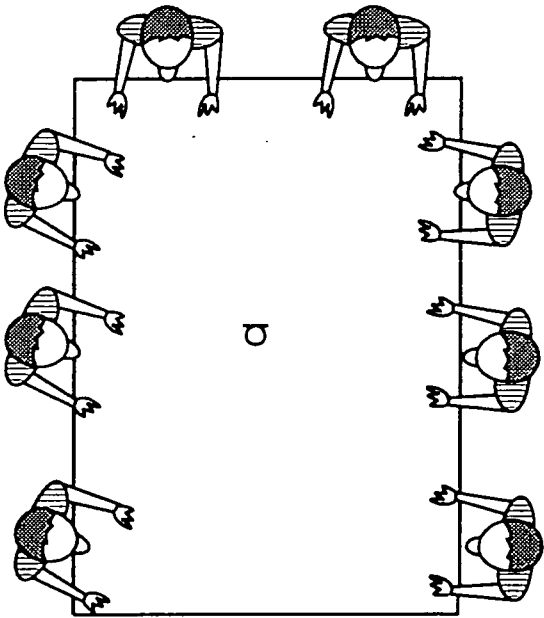


圖 4b

虛擬立體聲麥克風

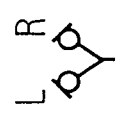
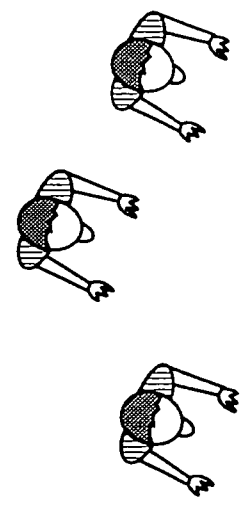
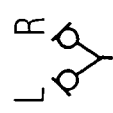


圖 4c

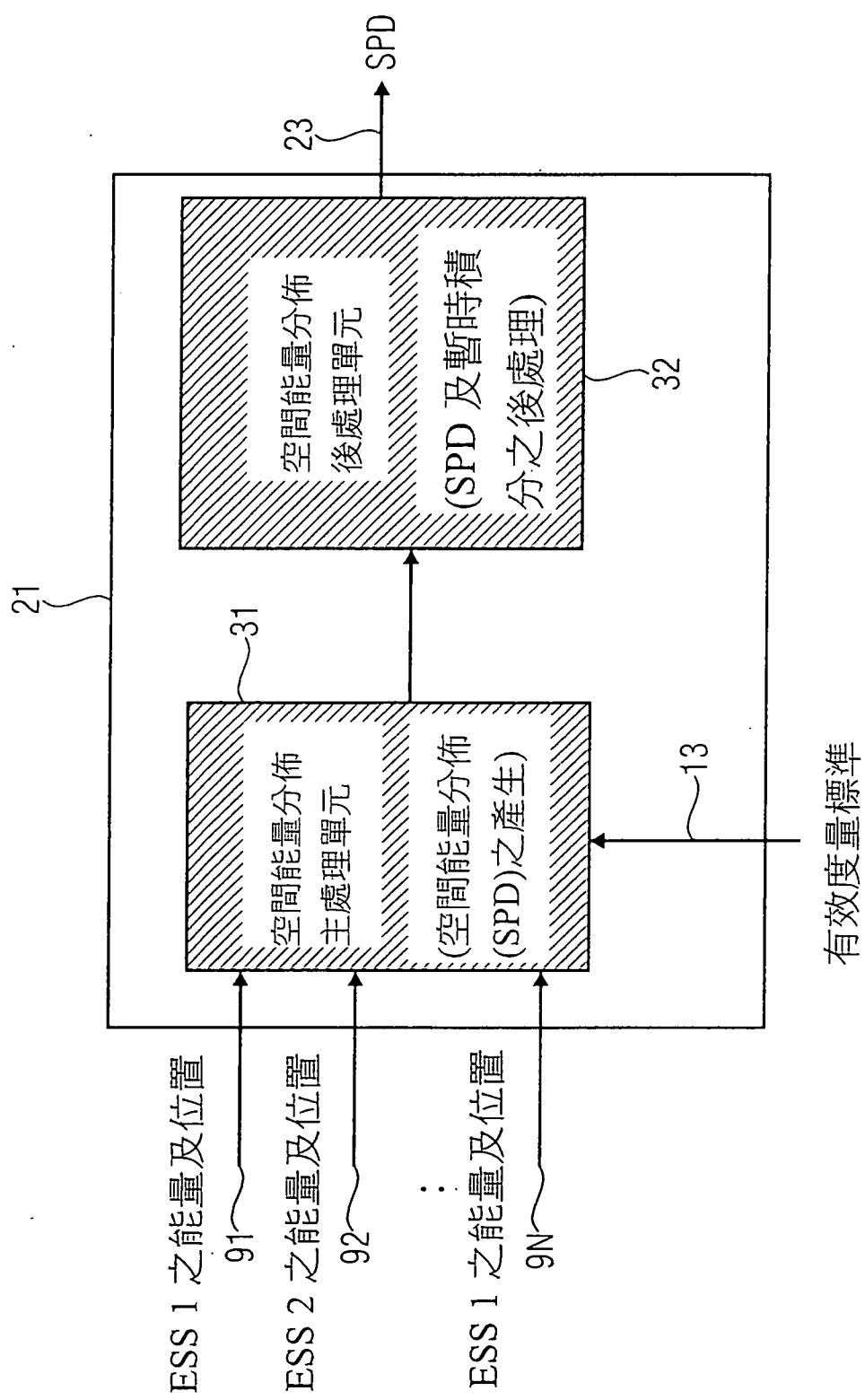


圖 5

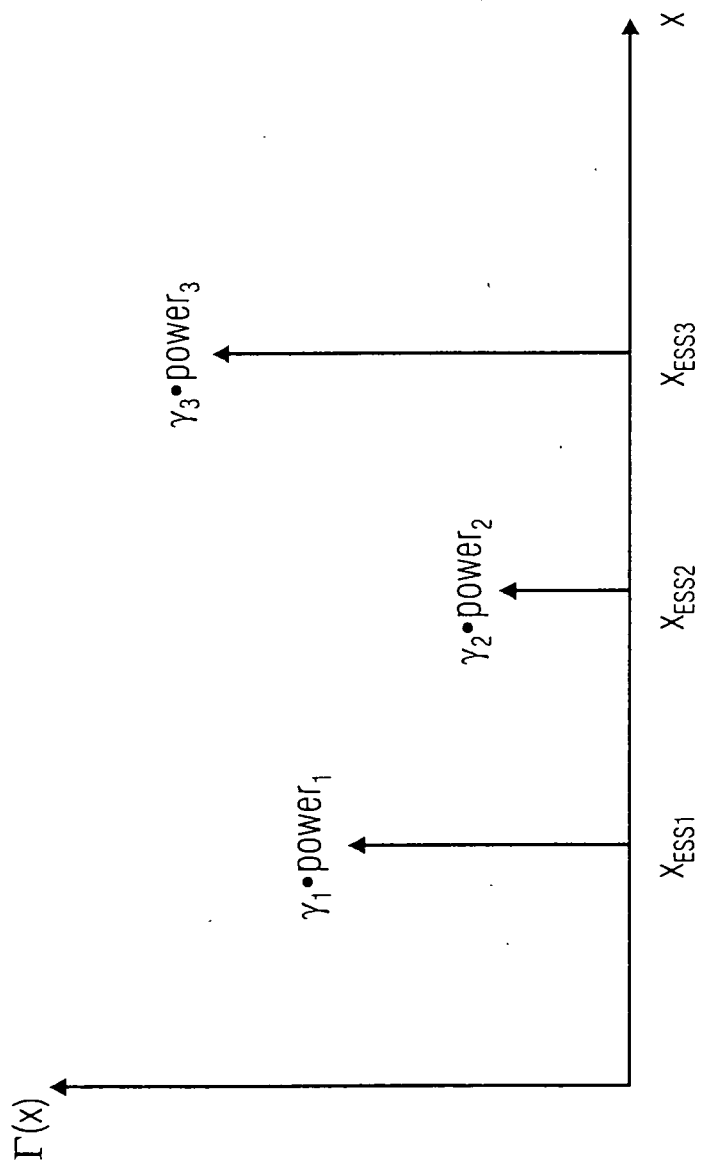


圖 6a

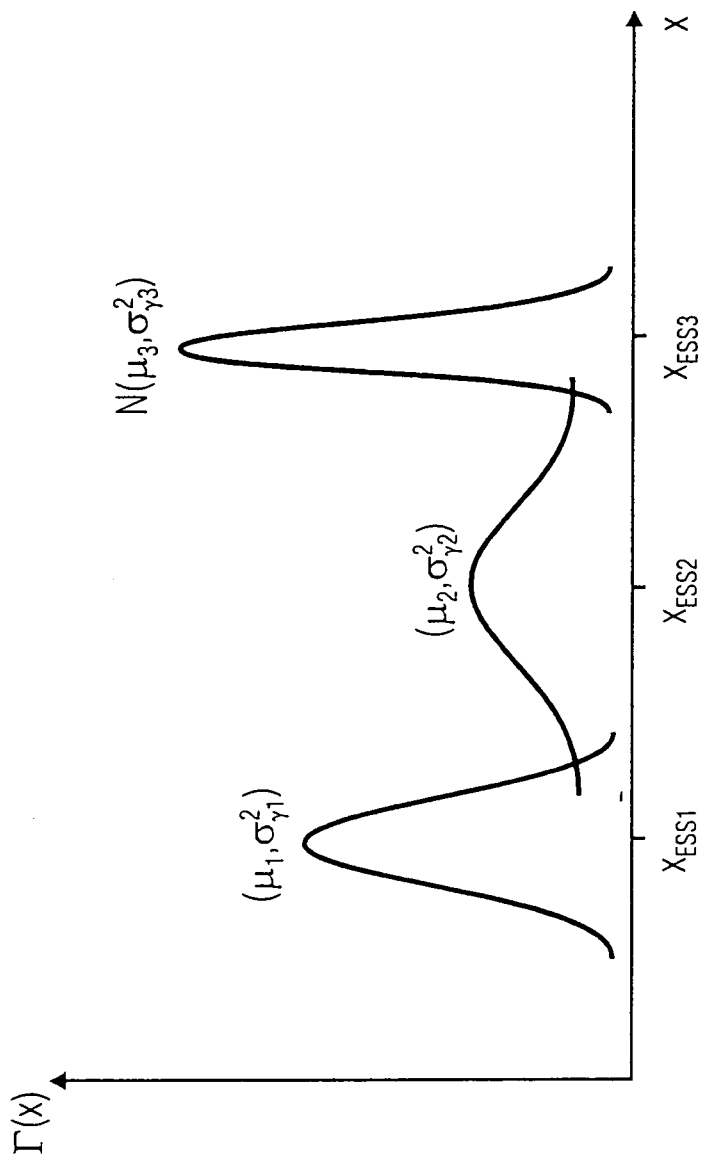


圖 6b

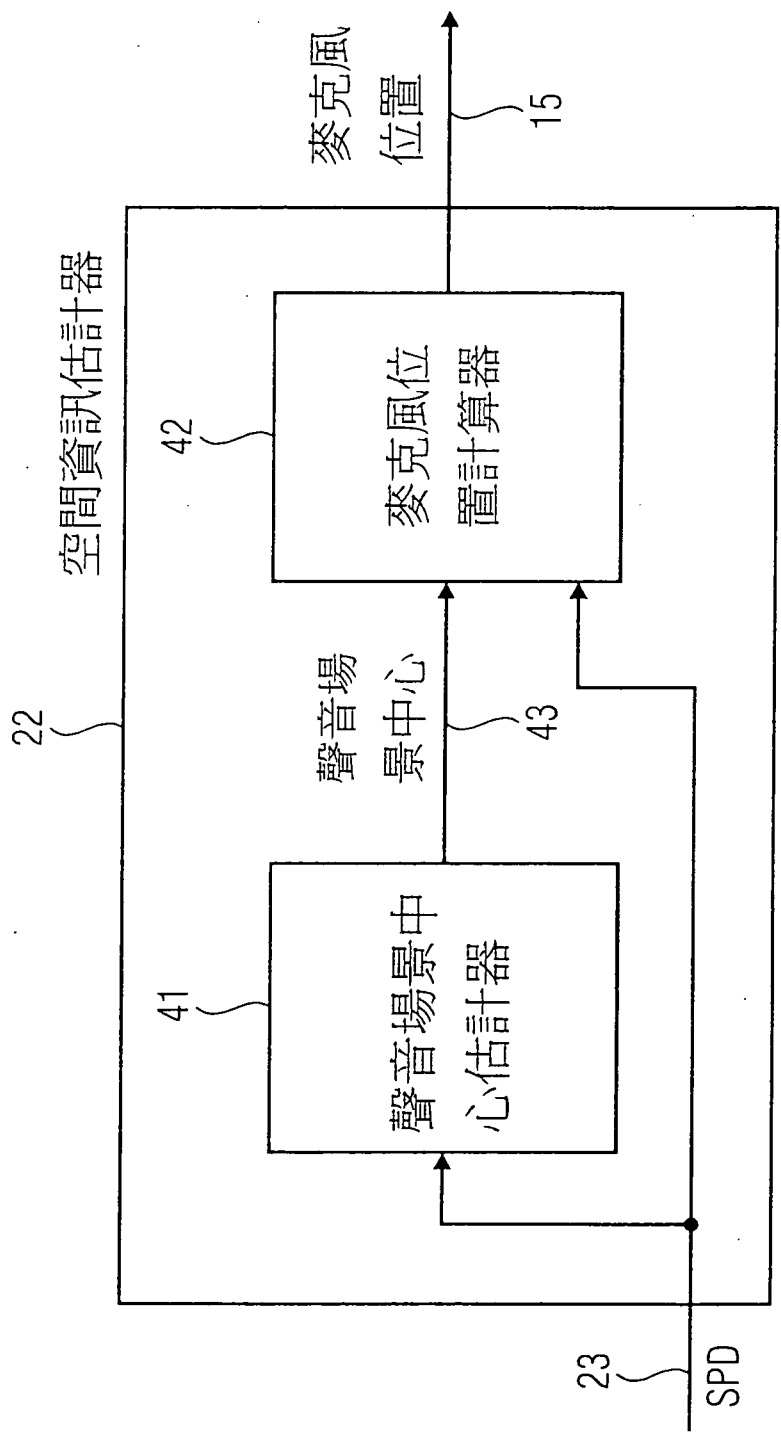


圖 7

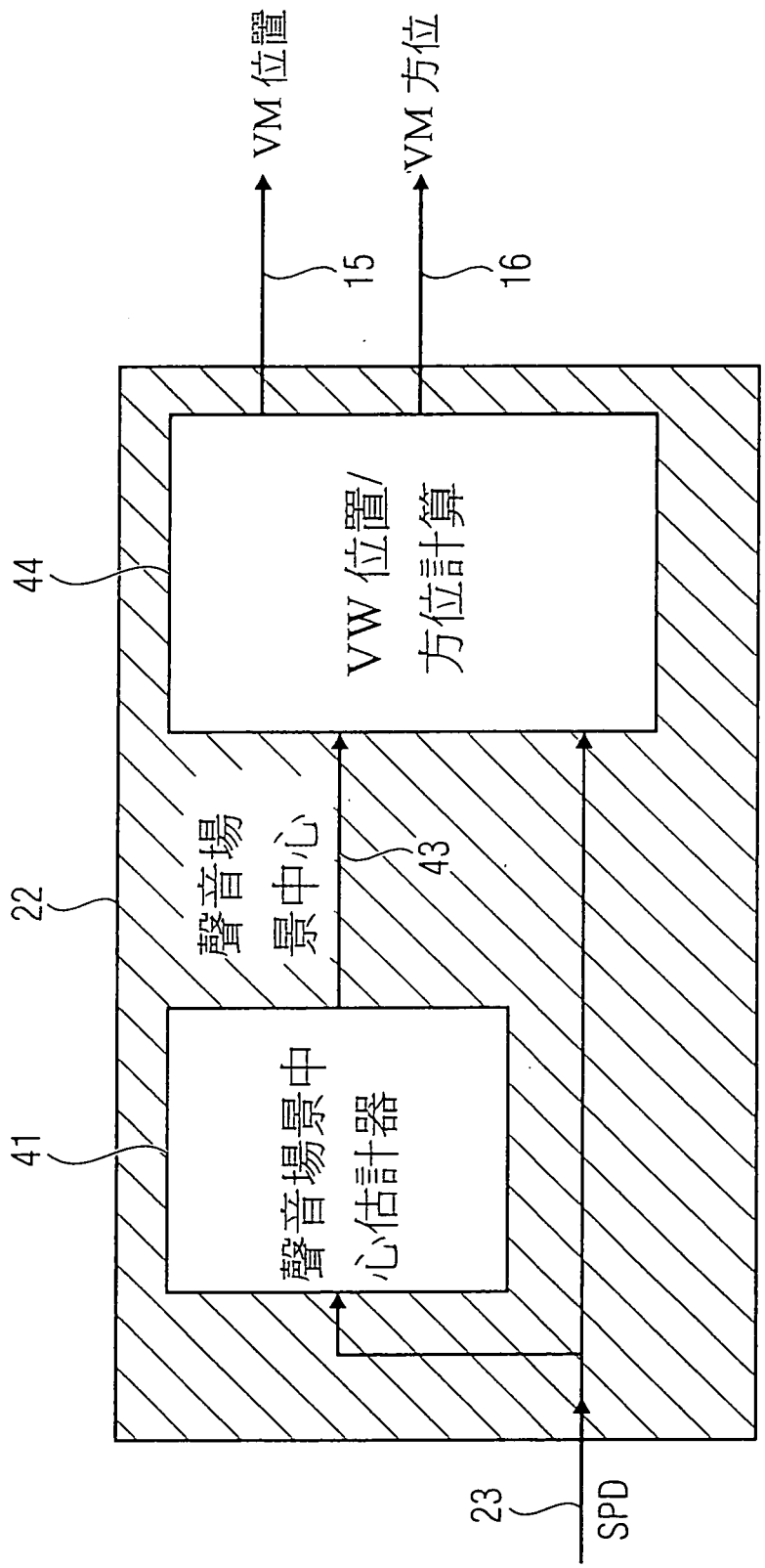


圖 8

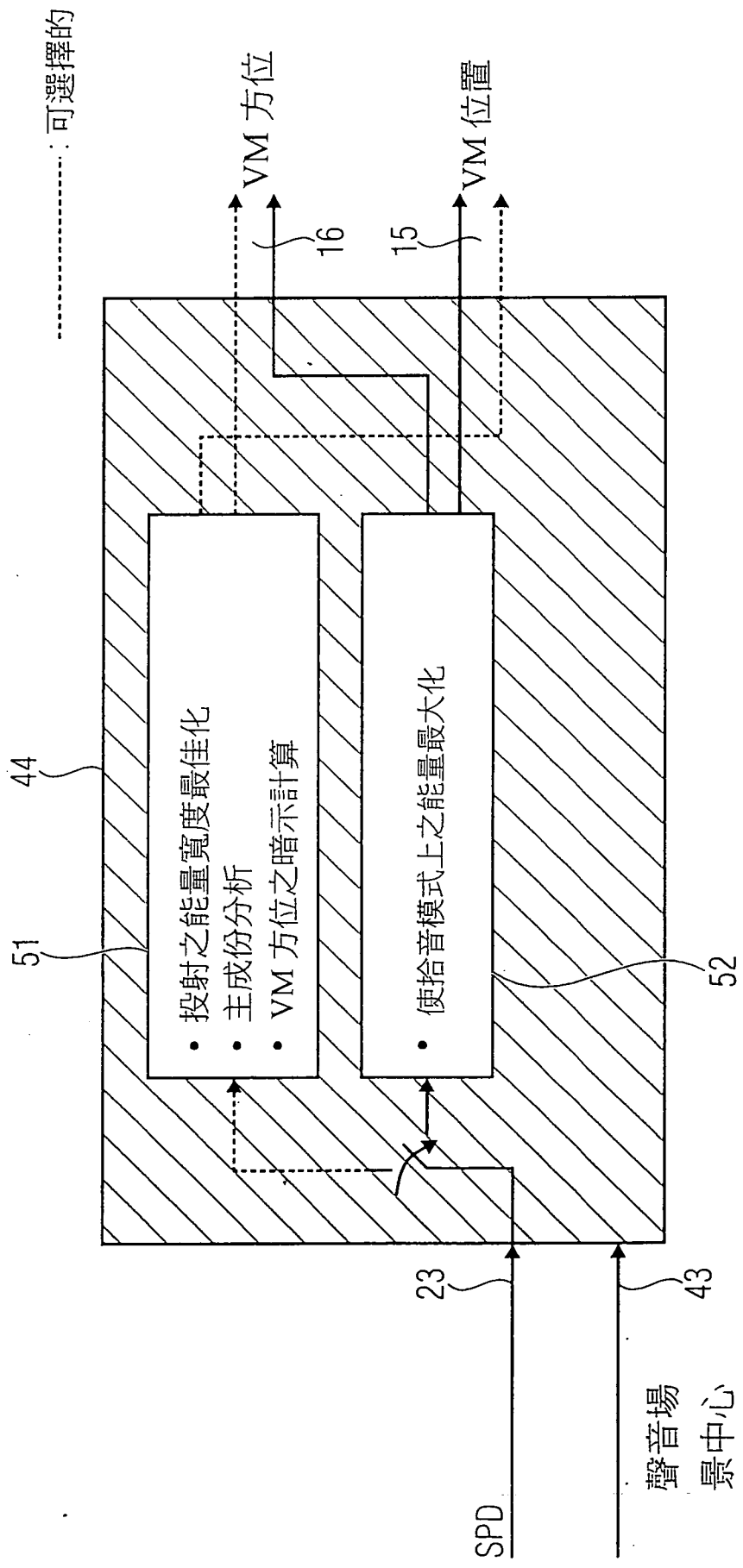


圖 9

COG : 聲音場景之重心

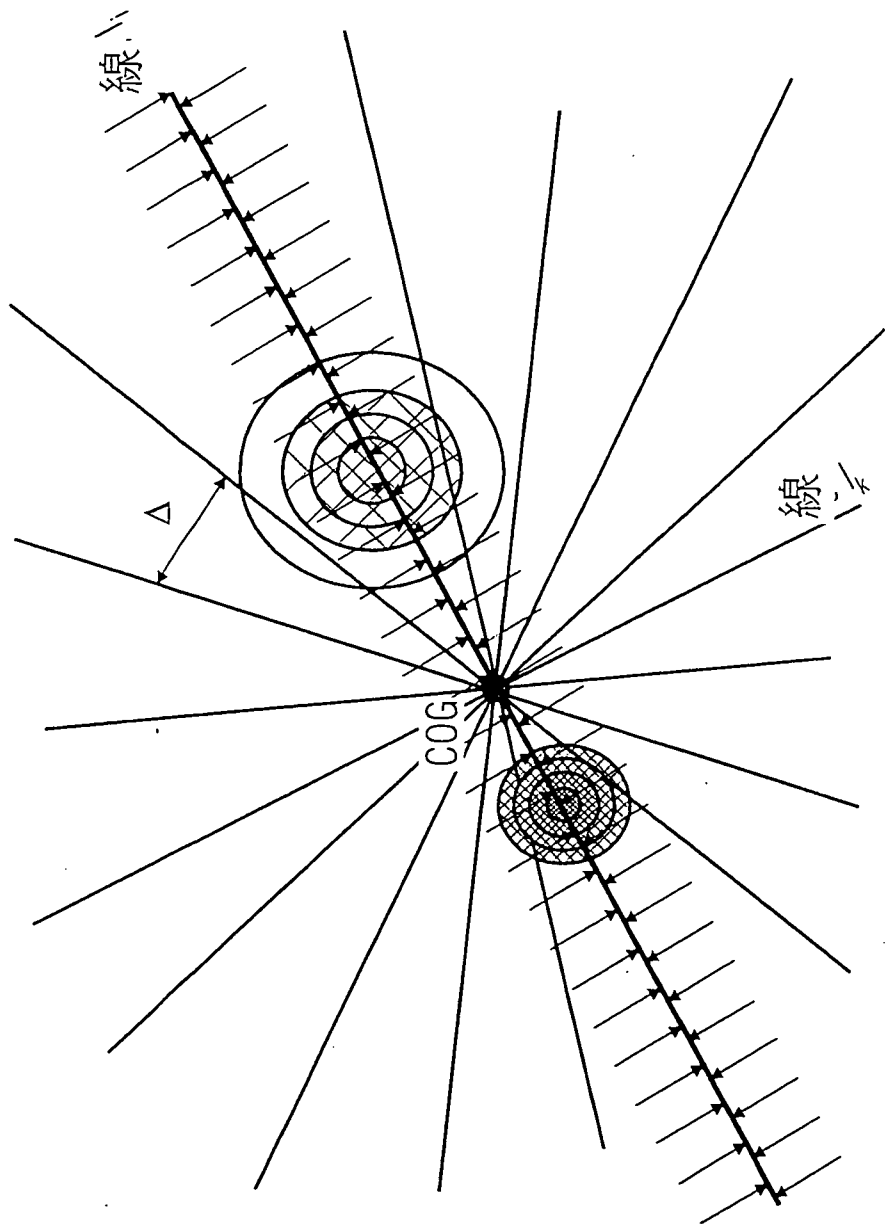


圖 10a

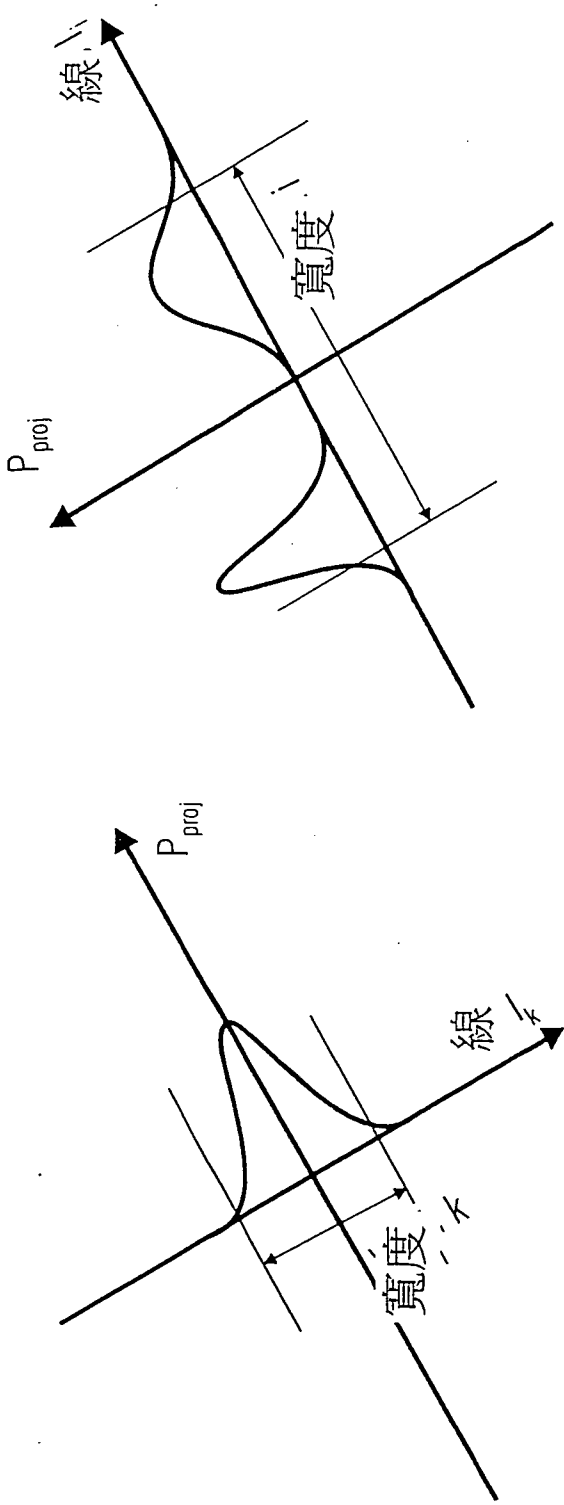


圖 10b

α : 虛擬麥克風之開口角度
 $d_{\text{COG-VM}}$: 在 COG 及 VM 之間的距離

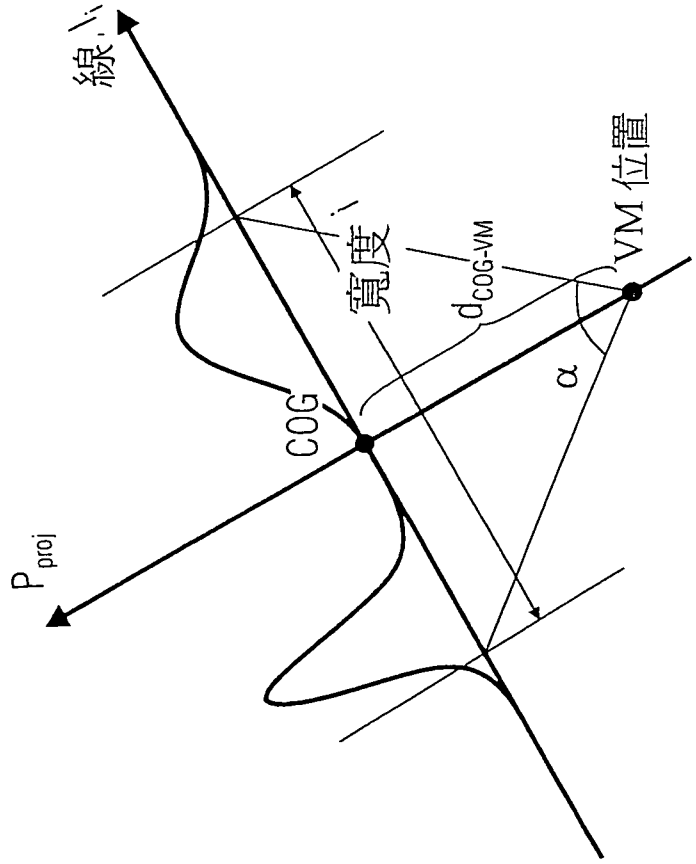


圖 10c

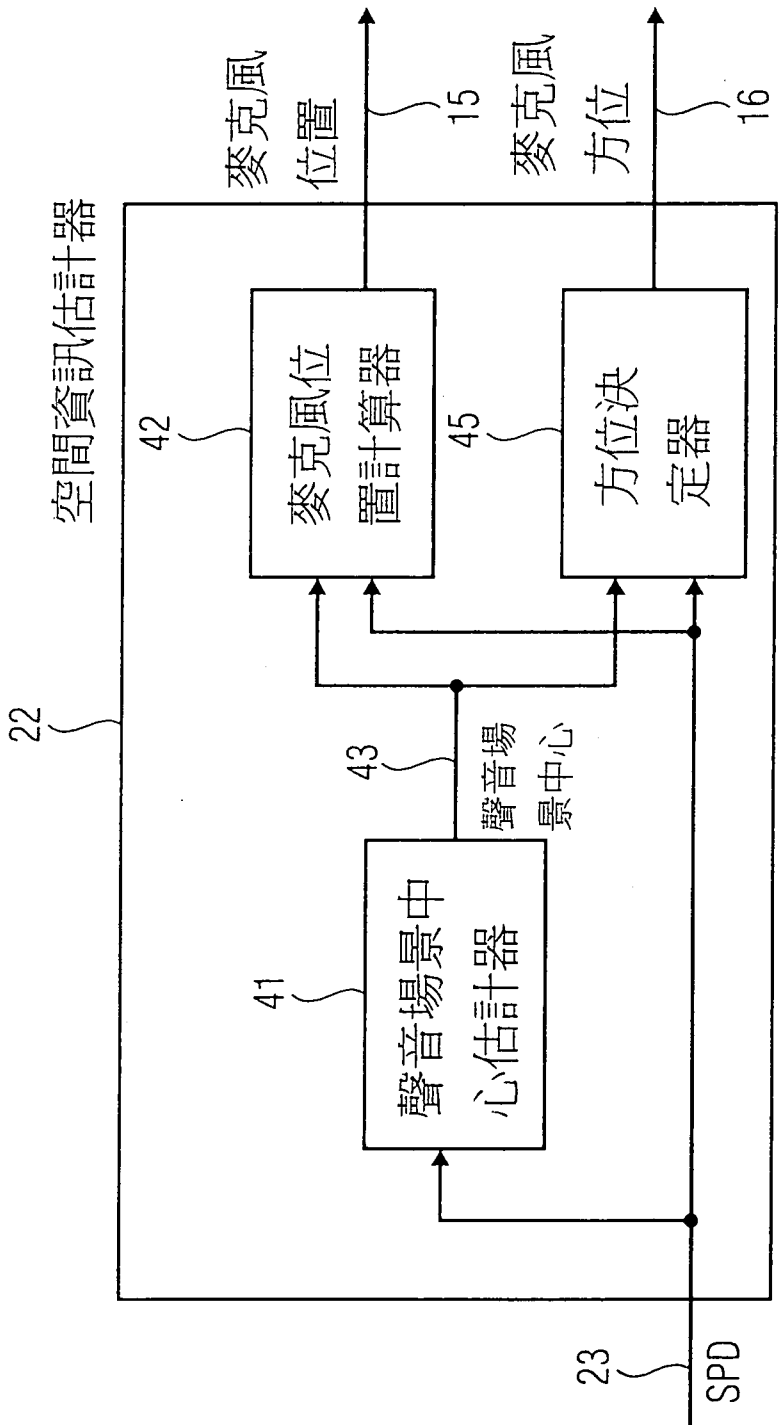


圖 11

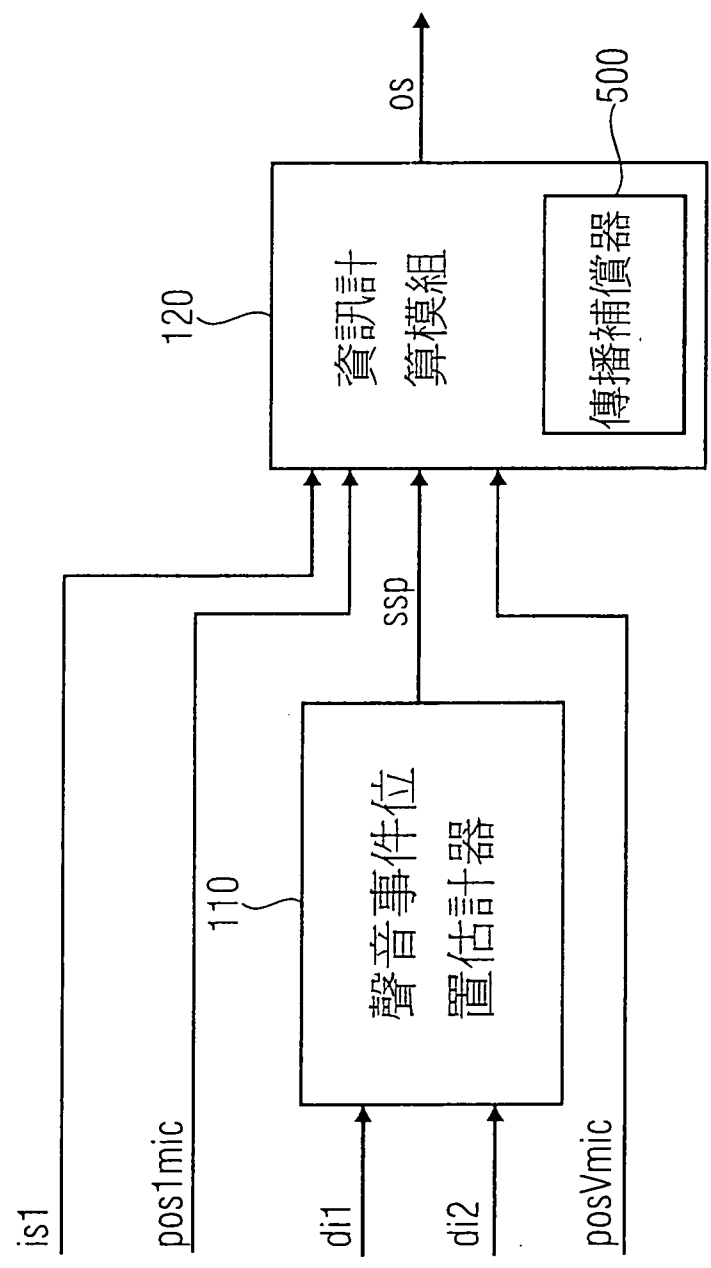


圖 12

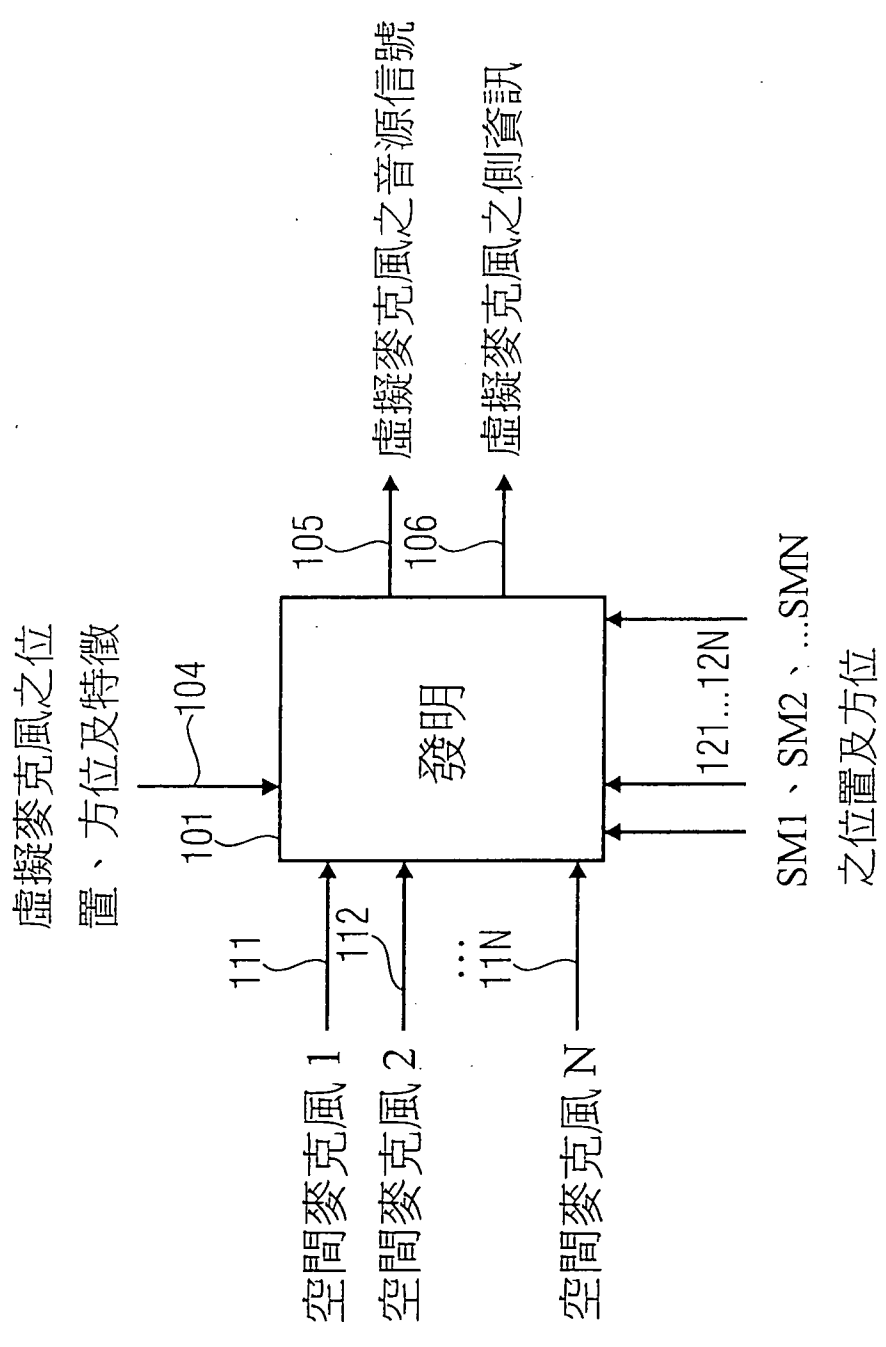
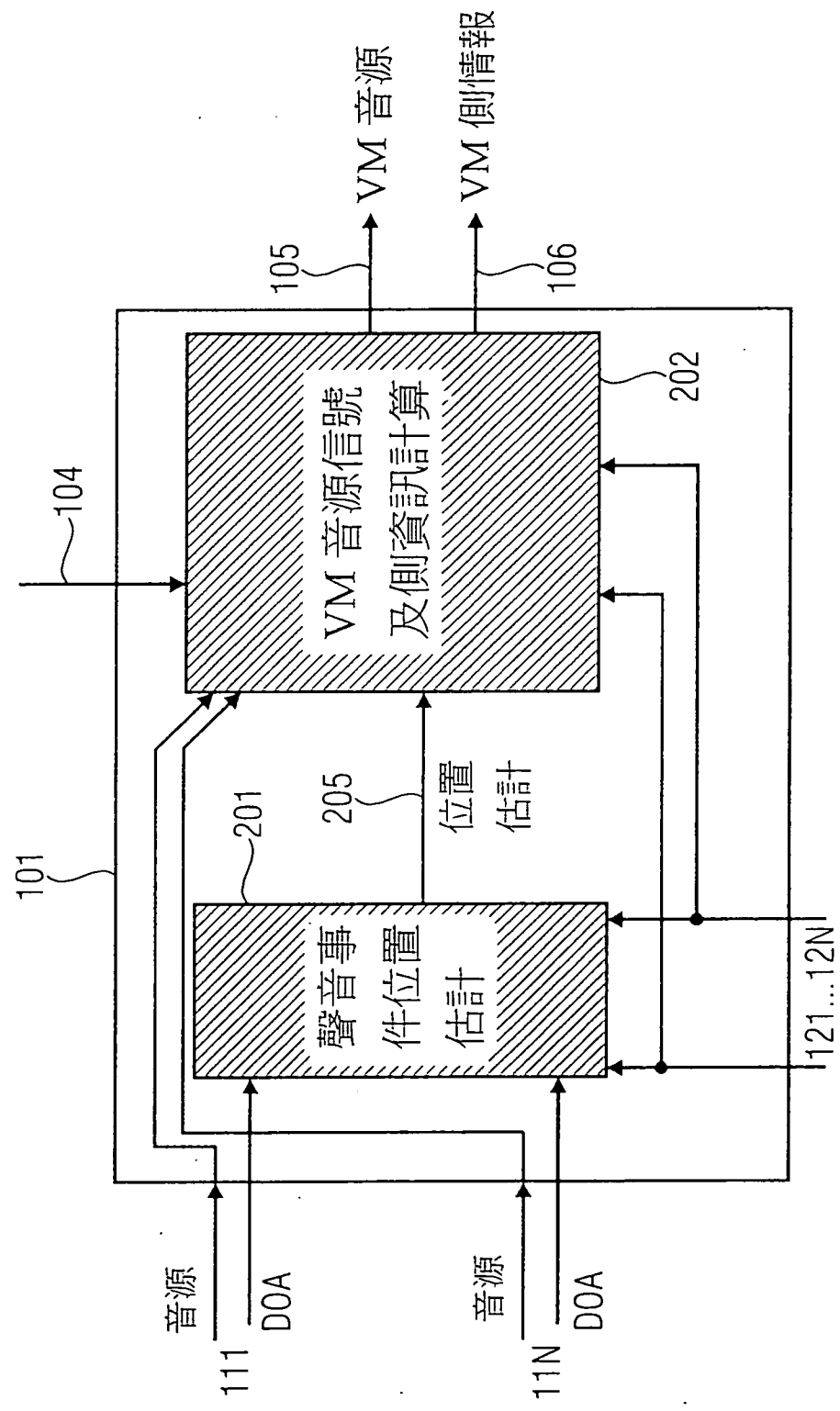


圖 13



SM1、SM2、...SMN
之位置及方位

圖 14

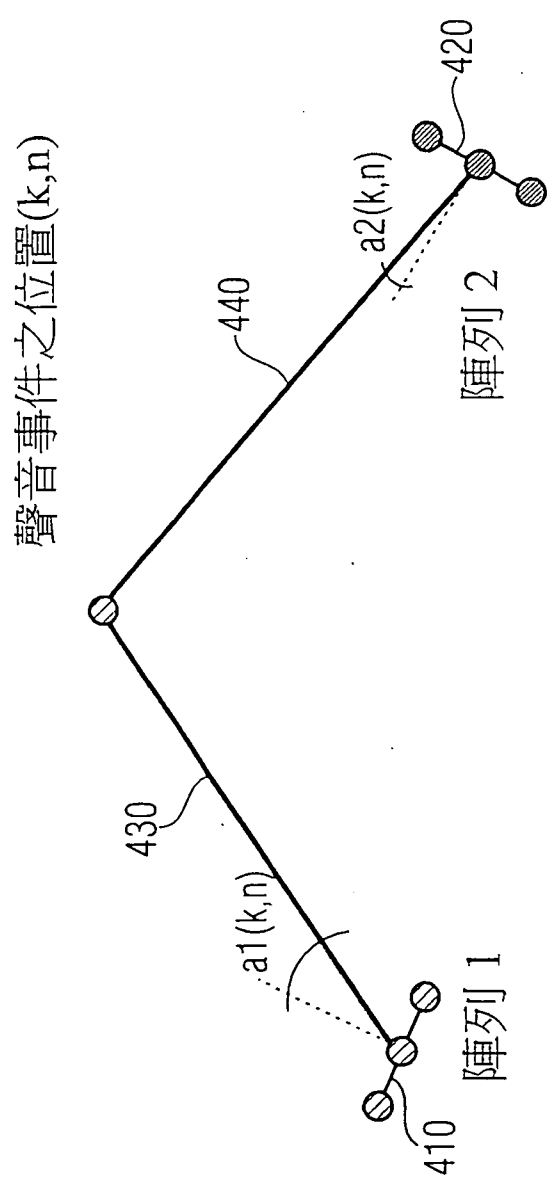


圖 15

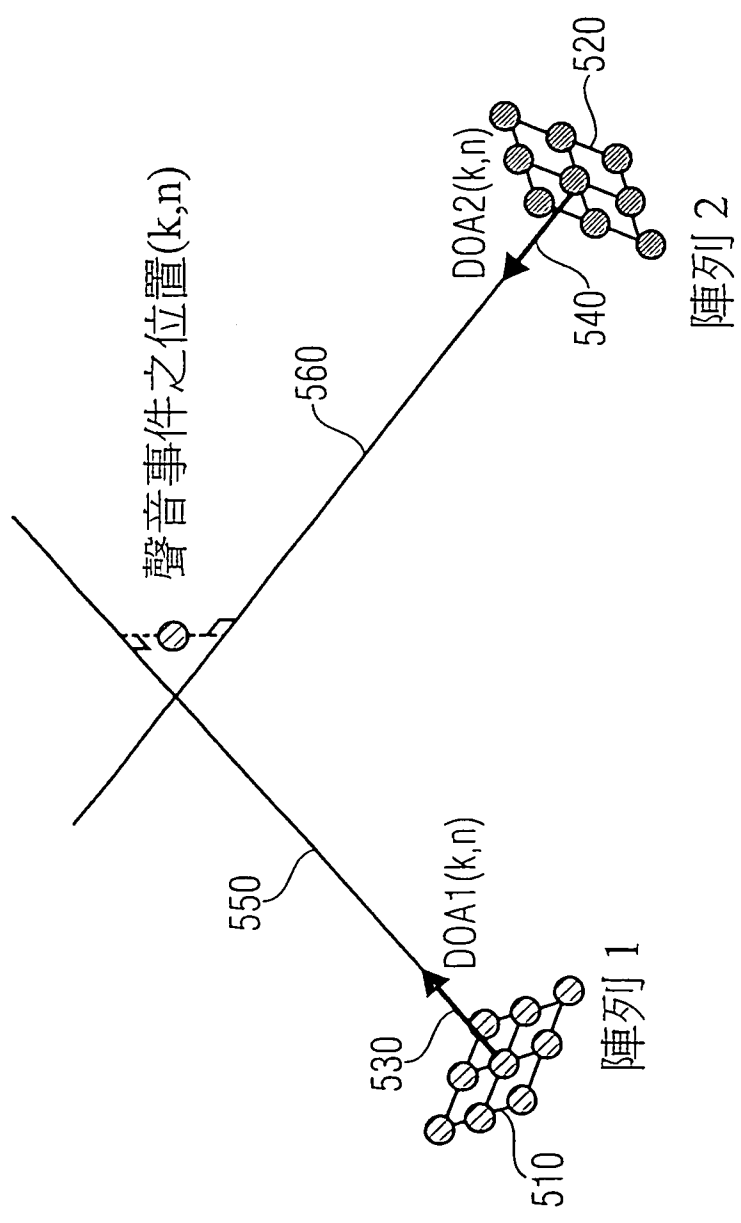


圖 16

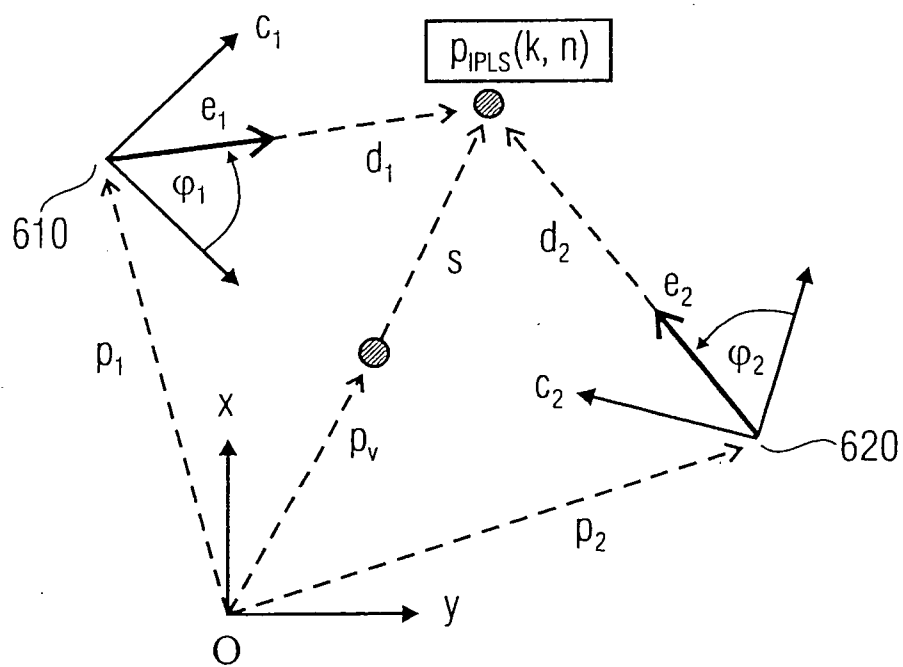


圖 17

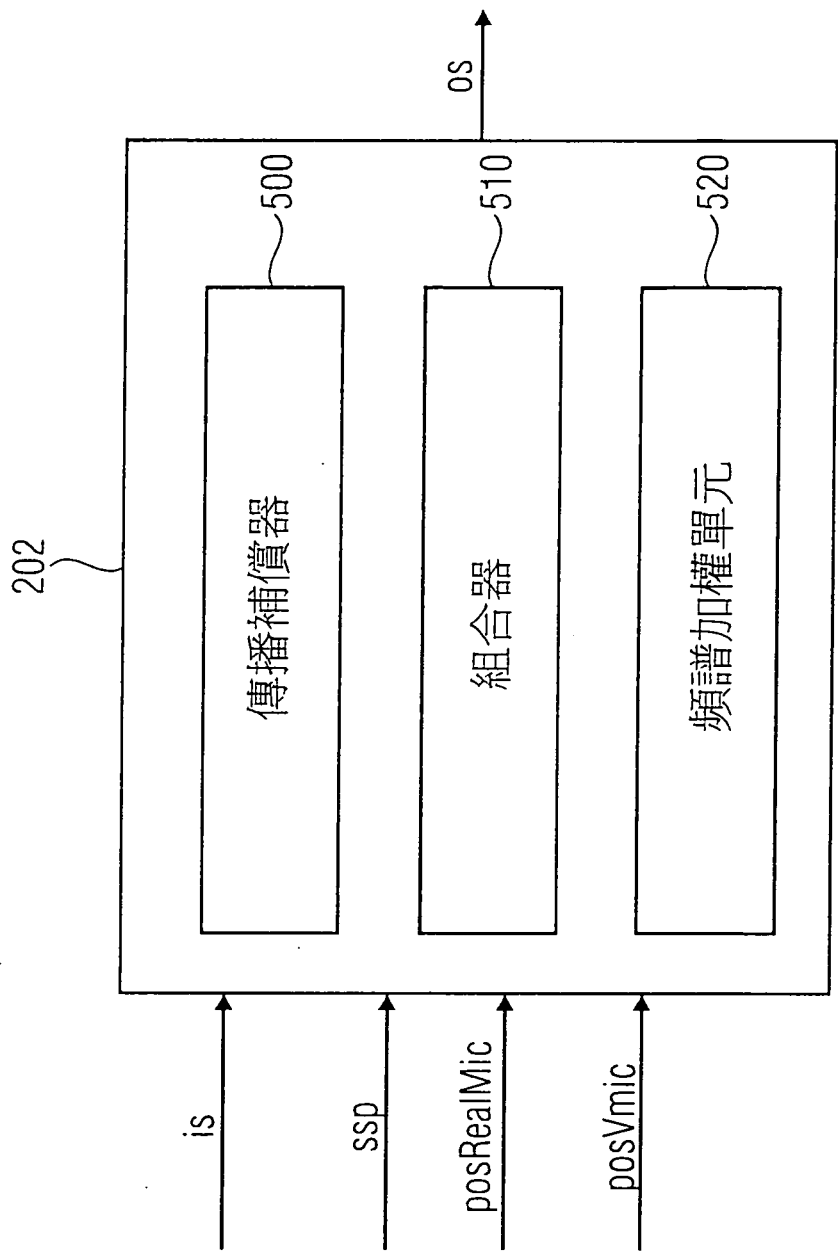


圖 18

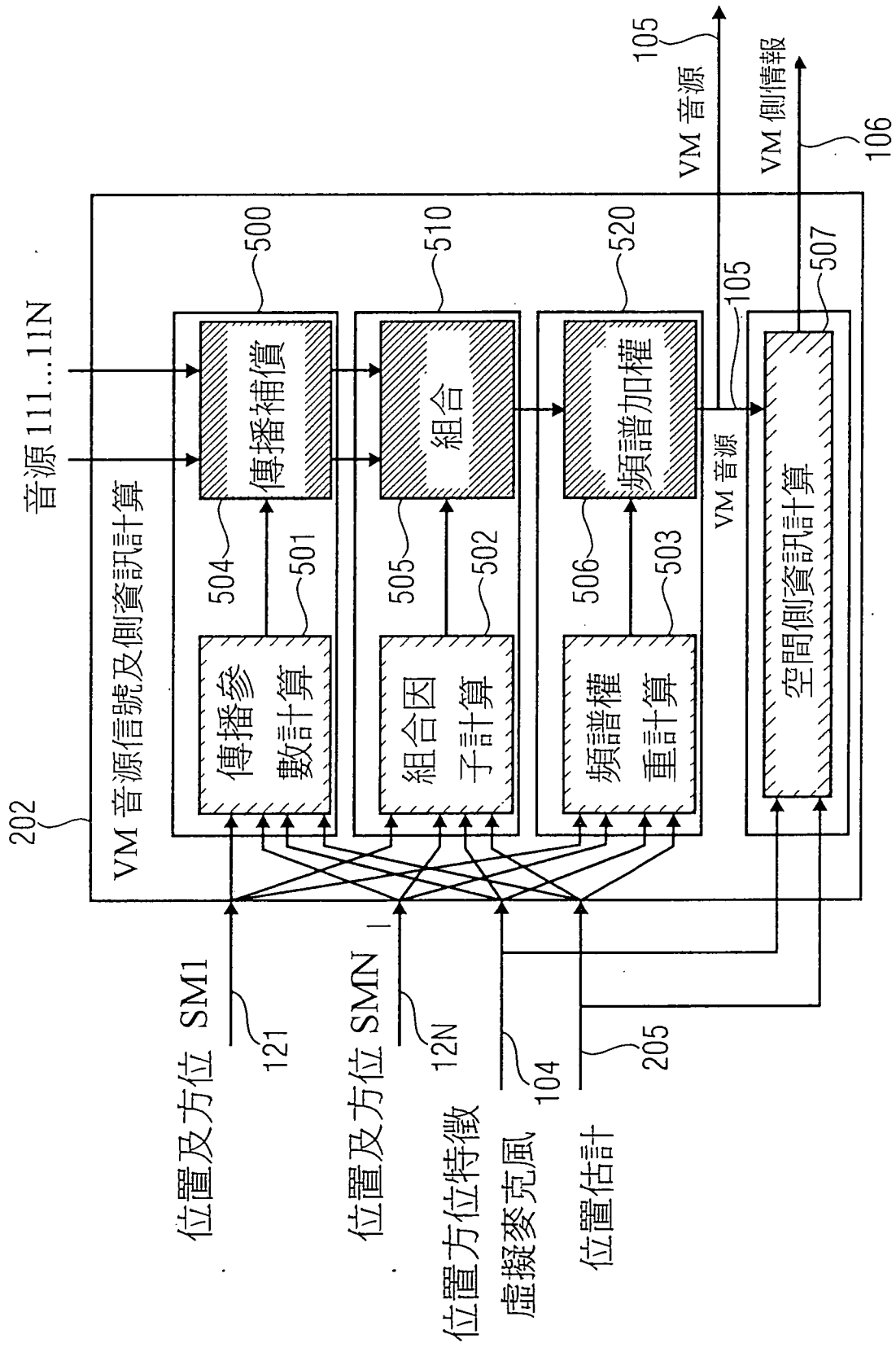


圖 19

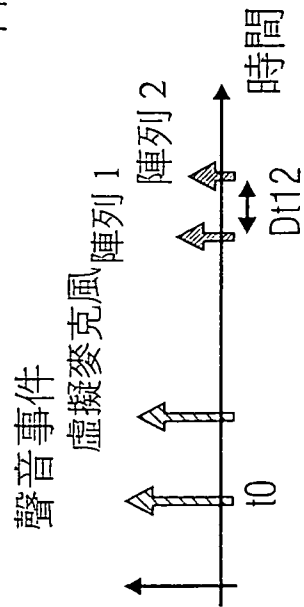
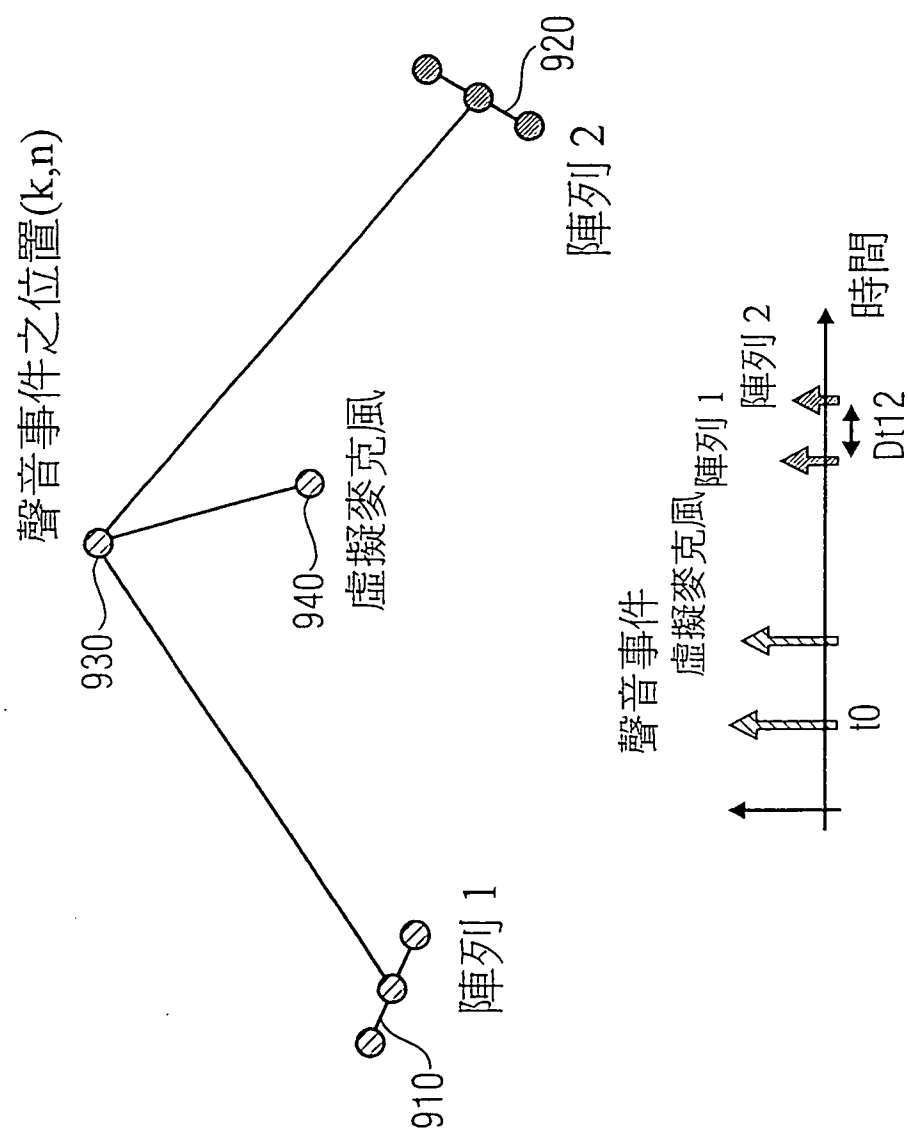


圖 20

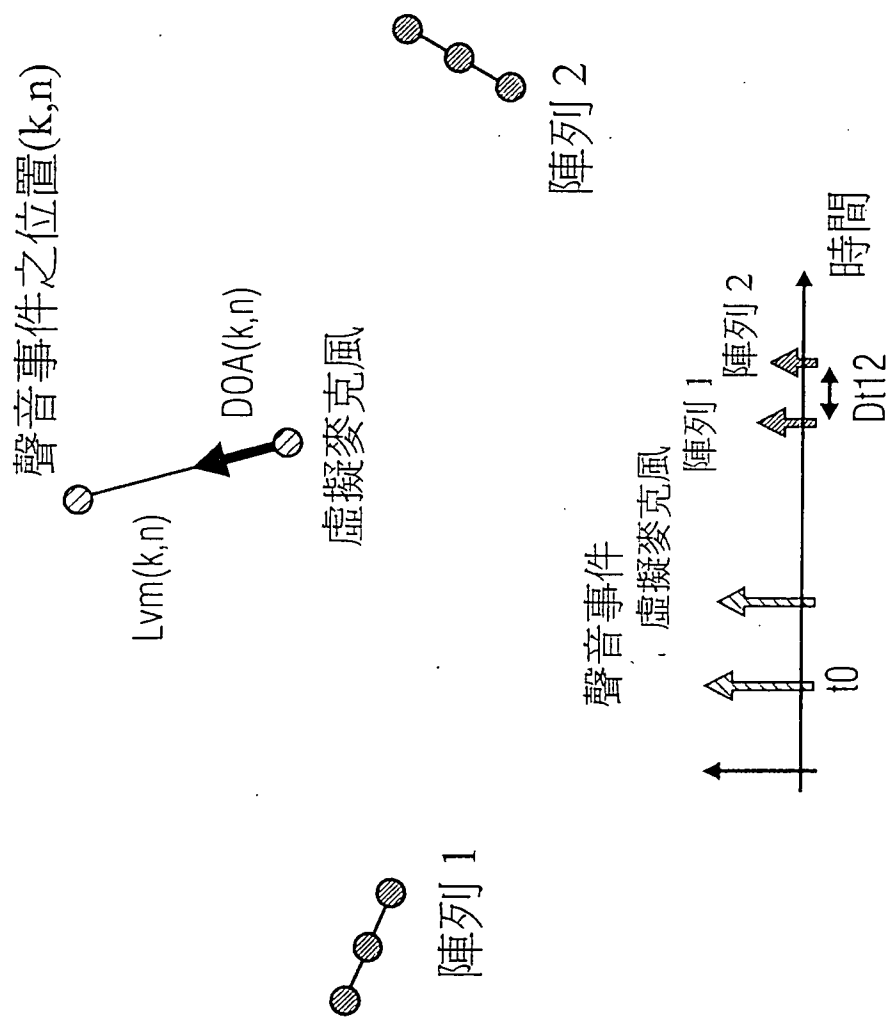


圖 21

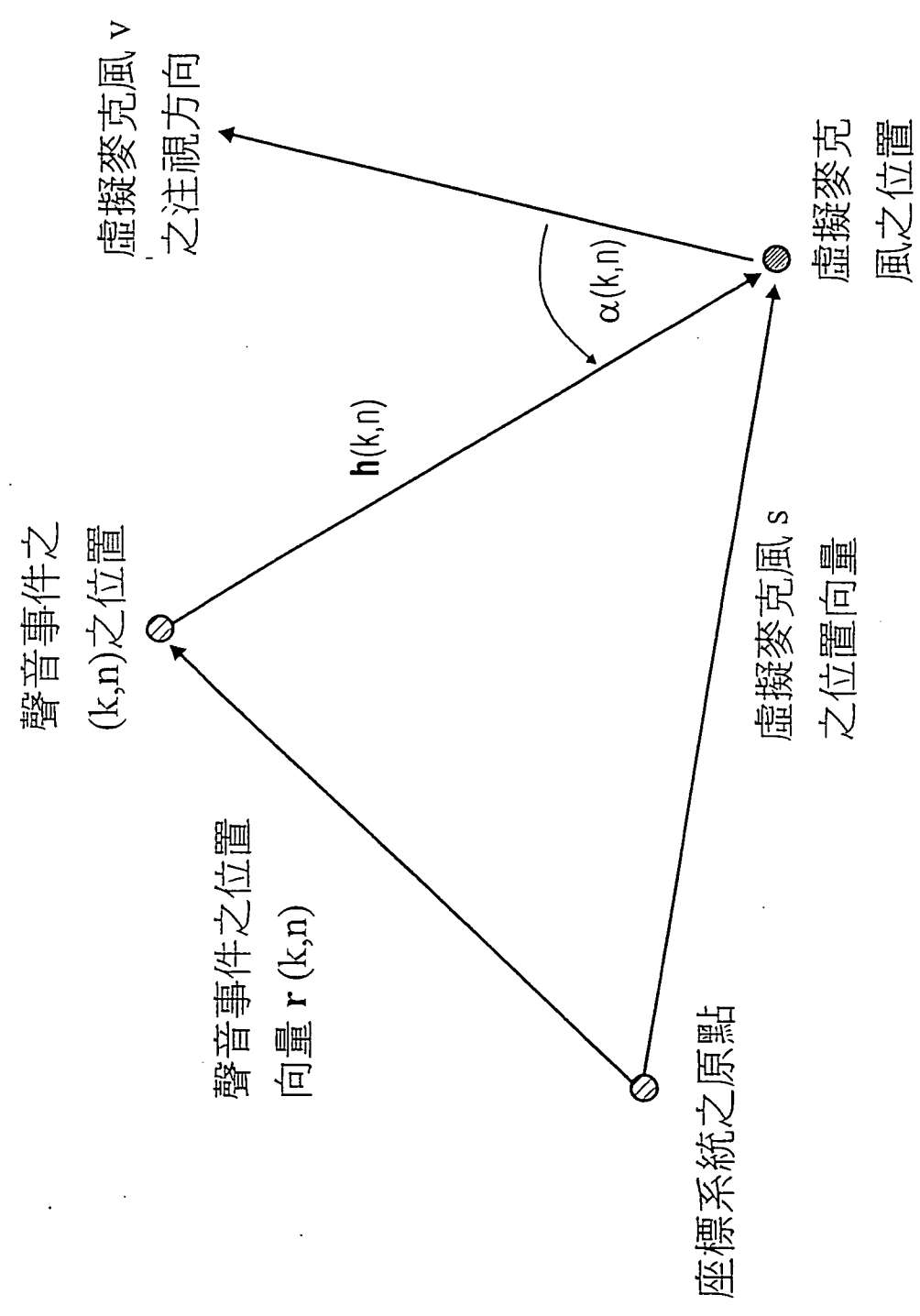


圖 22

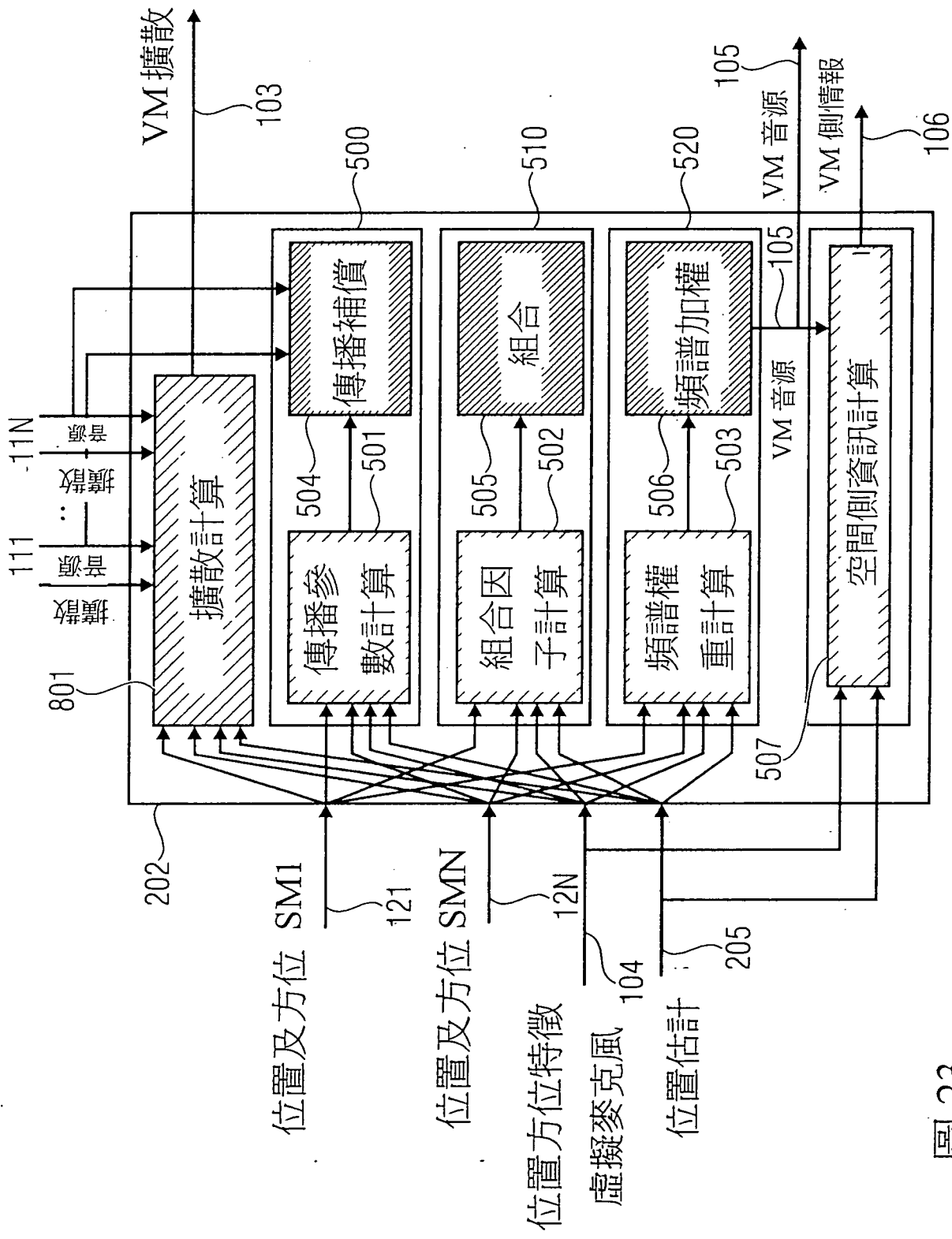


圖 23

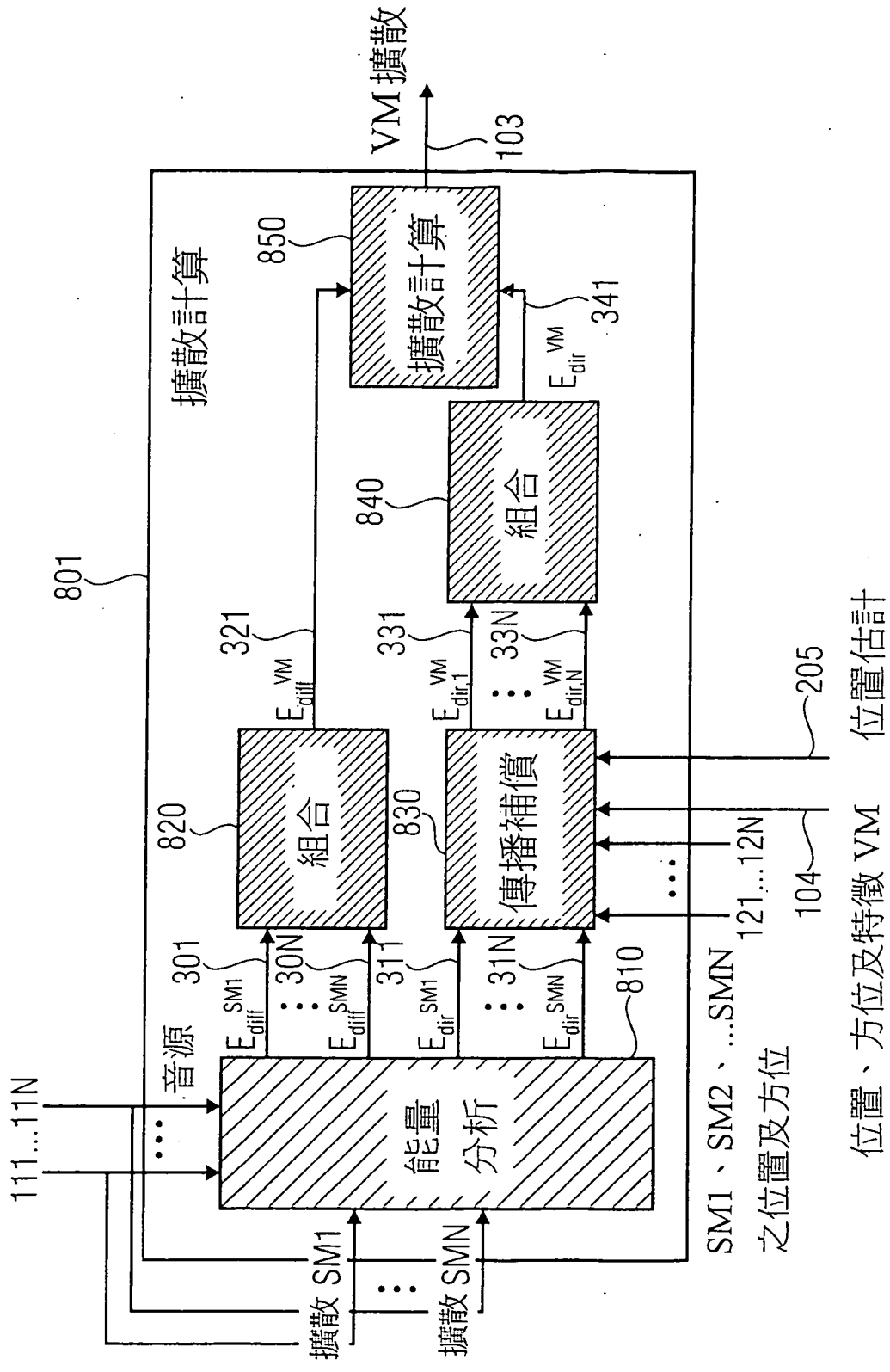


圖 24

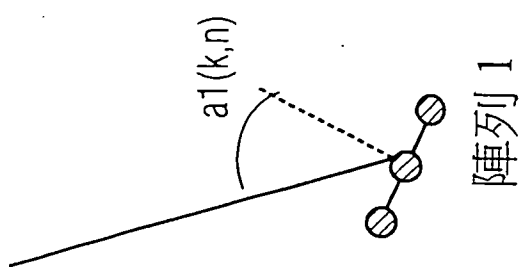
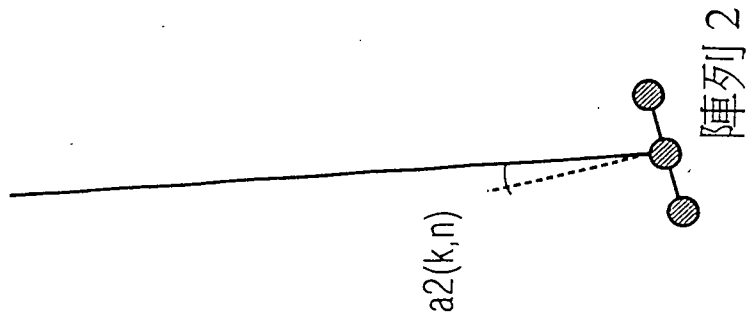


圖 25

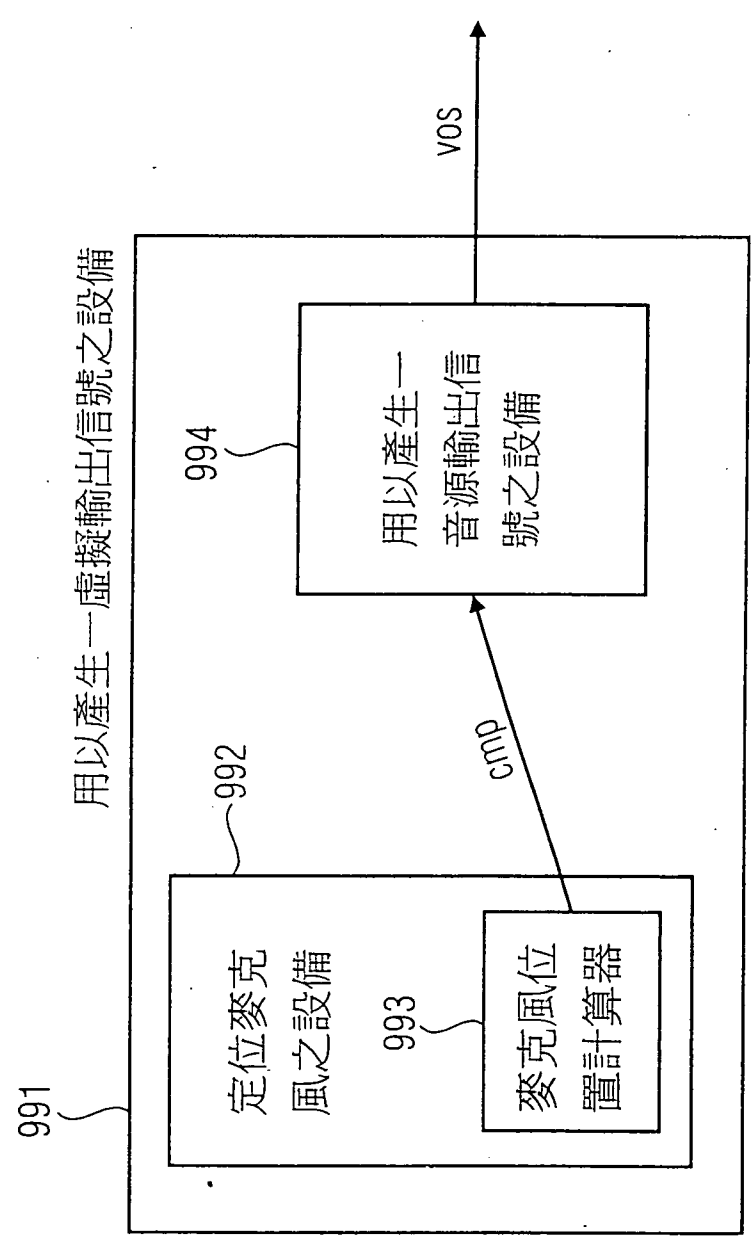
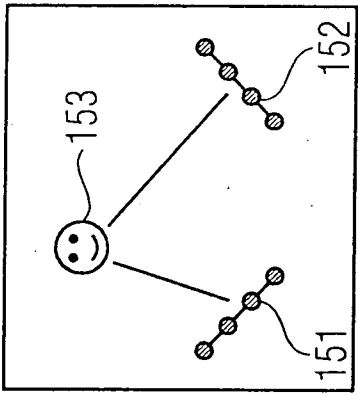
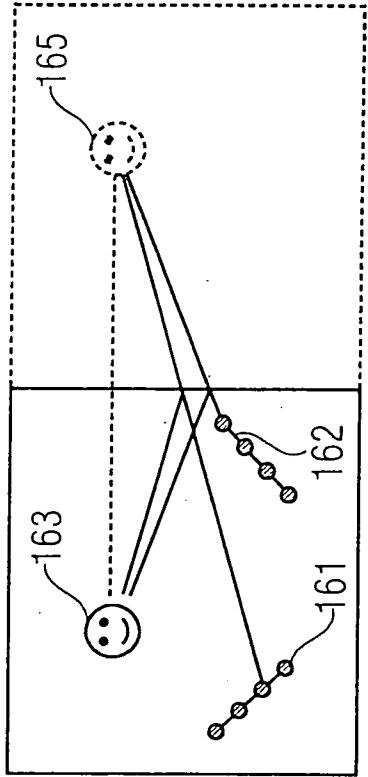


圖 26



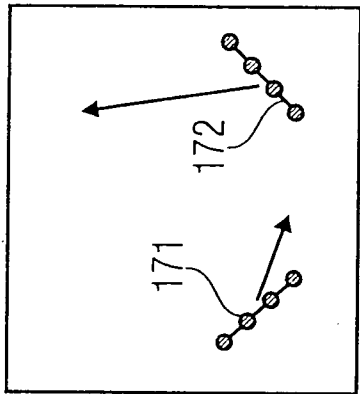
直達聲音

圖 27a



回響

圖 27b



擴散聲音及噪音

圖 27c