



(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 I713866 B

(45)公告日：中華民國 109 (2020) 年 12 月 21 日

(21)申請案號：107124520

(22)申請日：中華民國 107 (2018) 年 07 月 16 日

(51)Int. Cl. : **H04S1/00 (2006.01)**(30)優先權：2017/07/14 歐洲專利局 EP17181488.2
2018/07/13 世界智慧財產權組織 PCT/EP2018/069140(71)申請人：弗勞恩霍夫爾協會(德國) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG
DER ANGEWANDTEN FORSCHUNG E. V. (DE)
德國(72)發明人：賀瑞 捷爾根 HERRE, JUERGEN (DE)；哈貝特斯 艾曼紐 HABETS, EMANUEL
(NL)

(74)代理人：邱珍元

(56)參考文獻：

TW 2016/14638A US 2013/0016842A1
US 2013/0142341A1 US 2013/0216070A1
WO 9741711AR.L. Bleidt et al., "Development of the MPEG-H TV Audio System for
ATSC 3.0", IEEE Trans. on Broadcasting, vol.63,no.1,pp.202-236,
March 2017.

審查人員：賴文能

申請專利範圍項數：47 項 圖式數：12 共 88 頁

(54)名稱

用於生成增強聲場描述的裝置與方法以及其計算機程式與記錄媒體

(57)摘要

一種用於生成一增強聲場描述的裝置，包括：一聲場生成器，用於生成至少一個聲場描述，該聲場描述指示關於至少一個參考位置的一聲場；以及一後設資料生成器，用於生成與該聲場的空間資訊有關的後設資料，其中，該至少一個聲場描述和該後設資料構成該增強聲場描述。該至少一個聲場描述在一某個情況下包括與該至少一個參考位置相關的一第一聲場描述以及與該參考位置不同的一另一參考位置相關的一第二聲場描述，以及與該空間資訊有關的該後設資料，該後設資料指示該參考位置和該另一參考位置或兩個參考位置之一以及在兩個參考位置之間延伸的一向量。

Apparatus for generating an enhanced sound field description, comprising: a sound field generator for generating at least one sound field description indicating a sound field with respect to at least one reference location; and a meta data generator for generating meta data relating to spatial information of the sound field, wherein the at least one sound field description and the meta data constitute the enhanced sound field description. The at least one sound field description comprises in a certain case a first sound field description related to the at least one reference location and a second sound field description related to a further reference location being different from the reference location, and the meta data relating to the spatial information indicates the reference location and the further reference location or one of both reference locations and a vector extending between both reference locations.

指定代表圖：

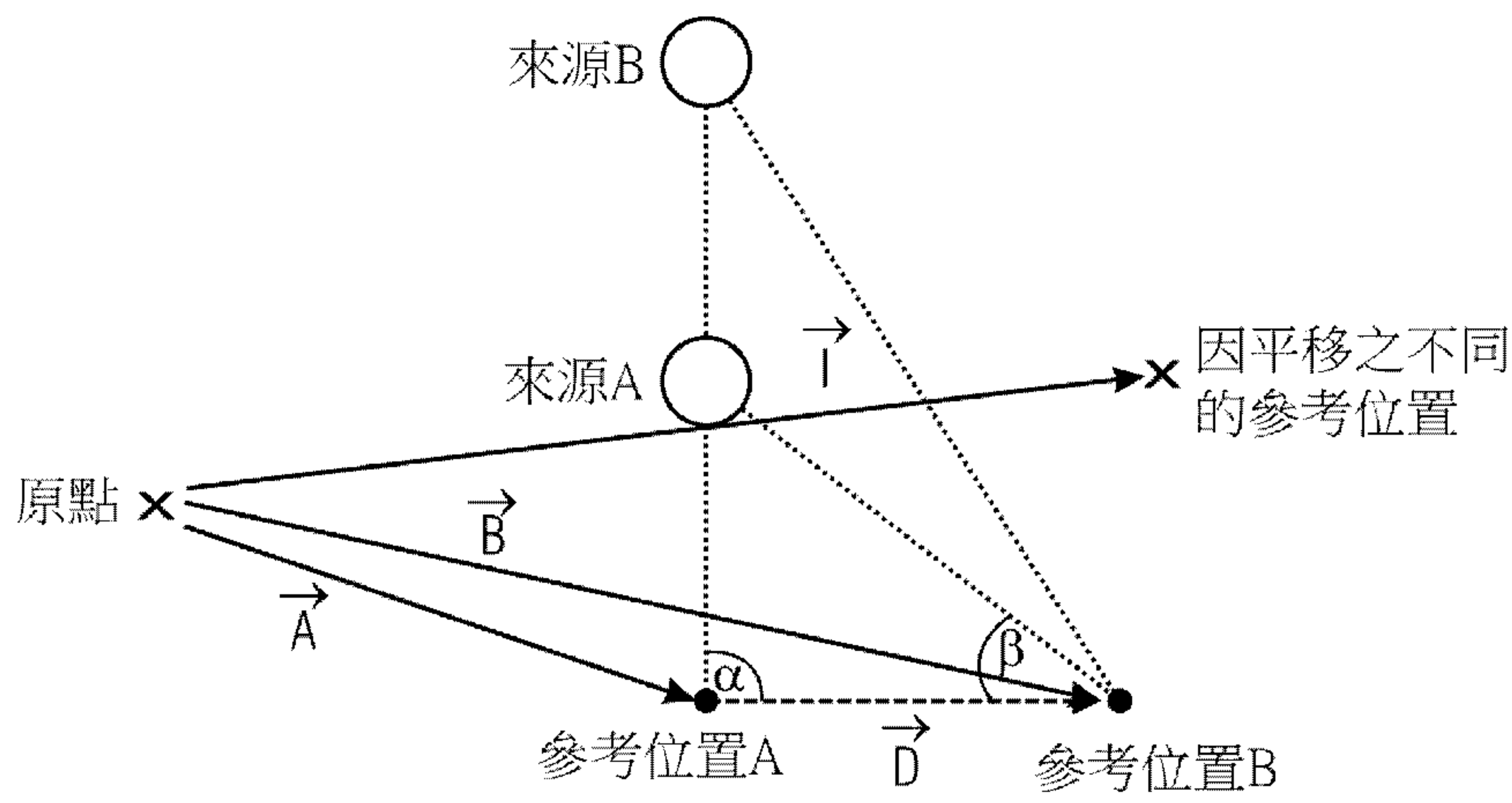


圖 4c



I713866

【發明摘要】

【中文發明名稱】用於生成增強聲場描述的裝置與方法以及其計算機程式與記錄媒體

【英文發明名稱】 APPARATUS AND METHOD FOR GENERATING AN ENHANCED SOUND FIELD DESCRIPTION, COMPUTER PROGRAM AND STORAGE MEDIUM

【中文】

一種用於生成一增強聲場描述的裝置，包括：一聲場生成器，用於生成至少一個聲場描述，該聲場描述指示關於至少一個參考位置的一聲場；以及一後設資料生成器，用於生成與該聲場的空間資訊有關的後設資料，其中，該至少一個聲場描述和該後設資料構成該增強聲場描述。該至少一個聲場描述在一某個情況下包括與該至少一個參考位置相關的一第一聲場描述以及與該參考位置不同的一另一參考位置相關的一第二聲場描述，以及與該空間資訊有關的該後設資料，該後設資料指示該參考位置和該另一參考位置或兩個參考位置之一以及在兩個參考位置之間延伸的一向量。

【英文】

Apparatus for generating an enhanced sound field description, comprising: a sound field generator for generating at least one sound field description indicating a

sound field with respect to at least one reference location; and a meta data generator for generating meta data relating to spatial information of the sound field, wherein the at least one sound field description and the meta data constitute the enhanced sound field description. The at least one sound field description comprises in a certain case a first sound field description related to the at least one reference location and a second sound field description related to a further reference location being different from the reference location, and the meta data relating to the spatial information indicates the reference location and the further reference location or one of both reference locations and a vector extending between both reference locations.

【指定代表圖】 圖4c

【代表圖之符號簡單說明】 無

【特徵化學式】

無

【發明說明書】

【中文發明名稱】用於生成增強聲場描述的裝置與方法以及其計算機程式與記錄媒體

【英文發明名稱】 APPARATUS AND METHOD FOR GENERATING AN ENHANCED SOUND FIELD DESCRIPTION, COMPUTER PROGRAM AND STORAGE MEDIUM

【技術領域】

【0001】 本揭露涉及音頻處理，並且特別是，涉及相對於諸如一麥克風或一虛擬麥克風位置的一參考位置所定義的聲場的音頻處理。

【先前技術】

【0002】 高保真度環繞聲信號(Ambisonics signals)包括該聲場的一截斷球形諧波分解。高保真度環繞聲有不同的風格。在文獻[31]，其揭露「傳統的」高保真度環繞聲，其現在被稱為「一階高保真度環繞聲」(FOA、First-Order Ambisonics)並且包括四個信號(亦即，一個全向信號和多達三個數字8字形方向信號)。最近的高保真度環繞聲變體被稱為「高階高保真度環繞聲」(HOA、Higher-Order Ambisonics高保真度環繞聲)，並且以攜帶更多信號為代價，提供增強的空間分辨率和更大的聆聽者最佳聆聽位置區域。通常，一完全定義的N階HOA表示係由 $(N+1)^2$ 個信號組成。

【0003】 與該高保真度環繞聲概念相關，該定向音頻編碼(DirAC、Directional Audio Coding)表示已經被設想為以一更緊湊的參數樣式，

以表示一FOA或HOA聲音場景。更具體地，該空間聲音場景係由一個(或多個)發送的音頻聲道表示，其表示該聲學場景的一降混混合以及在每個時間頻率(TF)區間中的該方向和擴散的相關聯的輔助資訊。有關定向音頻編碼(DirAC)的更多資訊可以在文獻[32,33]中找到。

【0004】 在文獻[32]中，DirAC可以和不同的麥克風系統及任意揚聲器設置一起被使用。該DirAC系統的該目的是使用多聲道/3D揚聲器系統盡可能精確地再現一既有聲學環境的該空間印象。在該所選擇的環境中，響應(連續聲音或脈衝響應)係由一全向麥克風(W)和一組麥克風所測量，其能夠測量聲音的該到達方向和聲音的該擴散。在文獻[34]中，一種常見的方法是應用與相應的笛卡爾坐標軸對齊的三個8字形麥克風(X，Y，Z)。一種方法是使用一聲場麥克風，它直接產生所有想要的響應。該W、X、Y和Z信號也可以從一組離散的全向麥克風計算出來。

【0005】 在DirAC中，該聲音信號首先將被分成頻率通道。根據每個頻率通道的時間測量該聲音方向和擴散。在傳輸時，一個或多個音頻通道以及分析的方向和擴散資料一起被發送。在合成時，施加到該揚聲器的該音頻可以是例如該全向通道W，或者每個揚聲器的該聲音可以被計算為W、X、Y和Z的一加權和，其形成具有針對每個揚聲器的一某個方向特性的一訊號。每個音頻通道被分成頻率通道，接著根據分析的擴散性將其可選擇的劃分為擴散串流和非擴散串流。利用一種技術，一擴散串流被再現，該技術產生一聲音場景的一擴散感知，在文獻[35-37]中，例如雙耳線索編碼(Binaural Cue Coding)中使用的該去相關(decorrelation)技術。利用一種技術(例如在文獻[38]中的VBAP)，非擴散聲音被再現，該技術其目的是根據該方向資料產生一類似點狀虛擬來源。

【0006】 三種具有一有限自由度的六個自由度(6DoF、six-degrees-of-freedom)的導航技術在文獻[39]中被提出。給定一單個高保真度環繞聲信號，一單個高保真度環繞聲信號被使用以下方法計算：1)模擬在一虛擬揚聲器陣列內的HOA回放和聆聽者移動、2)沿著平面波計算和平移，以及3)重新擴展關於該聆聽者的該聲場。

【0007】 此外，參考所描述的該DirAC技術，例如於2009年11月11-13日在日本宮城縣Zao舉行的International Workshop on the Principles and Applications of Spatial Hearing中由V. Pulkki等人著作的該出版物“Directional Audio Coding-Perception-Based Reproduction of Spatial Sound”。該參考文獻描述了定向音頻編碼作為相關聲場處理的一參考位置的一範例，特別是作為用於空間音頻處理的一感知激勵技術。在電話會議、在定向濾波和在虛擬聽覺環境，其具有空間聲音的捕獲、編碼和重新合成之應用。

【0008】 聲音場景的再現通常聚焦在揚聲器設置上，因為這是在私人場所的該典型再現，例如起居室和專業環境(亦即電影院)。此處，該場景與該再現幾何的該關係是靜態的，因為它伴隨著強迫該聆聽者在該正面方向上觀看的一個二維圖像。隨後，在生產時，該聲音和該視覺對象的該空間關係被定義並固定。

【0009】 在虛擬實境(VR、Virtual Reality)中，藉由允許該用戶在該場景中自由移動而明確地實現該沉浸。因此，有必要追蹤該用戶的移動並將該視覺和聽覺再現調整到該用戶的位置。通常，該用戶佩戴一頭戴式顯示器(HMD、Head-Mounted Display)和耳機。對耳機的一沉浸式體驗，該音頻必須被雙耳化。雙耳化是一種模擬人類頭部、耳朵和上部軀幹如何根據其方向和距離改變一聲源的該聲音。在文獻[1,2]中，這

是經由將該信號與頭部相關轉移函數(HRTF、Head-Related Transfer Functions)針對它們相對方向執行卷積運算來實現的。在文獻[3]中，雙耳化也使聲音看起來係來自該場景而不是來自頭部內。在文獻[4,5]中，已經成功解決的一常見情況是360°影片再現。此處，該用戶戴著一HMD或是手持一平板電腦或手機。藉由移動她/他的頭部或該設備，該用戶可以向任何方向環顧四周。這是一個三自由度(3DoF、three-Degrees-Of-Freedom)場景，因為用戶有三個移動度(俯仰、偏轉、滾動)。在視覺上，這藉由將該影片投影在該用戶周圍的一球體上來實現。在文獻[6]中，音頻通常用一空間麥克風記錄，例如，靠近該攝像機的一階高保真度環繞聲(FOA、First-Order Ambisonics)。在文獻[7]中，在高保真度環繞聲領域，該用戶的頭部旋轉以一直接的方式進行調整。接著例如該音頻被呈現給放置在該用戶周圍的虛擬揚聲器。這些虛擬揚聲器信號接著被雙耳化。

【0010】 現代VR應用程式允許六個自由度(6DoF、six-Degrees-Of-Freedom)。除了該頭部旋轉之外，該用戶可以四處移動，從而產生在三個空間維度上平移她/他的位置。該6DoF再現受到該步行區域的該整體尺寸的限制。在許多情況下，該區域相當小，例如一傳統的起居室。6DoF在VR遊戲中經常被遇到。這裡，該整個場景是由計算機生成的圖像(CGI、Computer-Generated Imagery)合成的。該音頻通常使用對象基礎的渲染而被生成，其中每個音頻對象係與距離相關的增益和基於該追蹤資料的來自該用戶的相對方向而被渲染。在文獻[8,9,10]中，經由人工混響(artificial reverberation)和衍射，可以增強真實性。

【0011】 關於錄製的內容，對於令人信服的視聽6DoF再現係存在一些明顯的挑戰。於文獻[11,12]中，在該空間平移領域中空間聲音操縱

的一早期例子是「聲學變焦」技術。此處，該聆聽者位置被虛擬地移動到該記錄的視覺場景中，類似於放大一圖像。該用戶選擇一個方向或圖像部分，接著可以從一平移點收聽。這必需要所有該到達方向(DoAs、Direction of Arrivals)係相對於原始的非縮放再現而改變。

【0012】 用於記錄內容的6DoF再現的方法係已經被提出，該記錄內容的6DoF再現已經使用空間分佈記錄位置。於文獻[13]中，對於影片，相機陣列可以被使用來生成光場渲染(light-field rendering)。對於音頻，一類似的設置採用分佈式麥克風陣列或高保真度環繞聲麥克風。於文獻[14]中，其已經表明，從這種記錄可以生成放置在任意位置的一「虛擬麥克風」的該信號。

【0013】 為了以一技術上方便的方式實現這種空間聲音修改，可以採用參數聲音處理或編碼技術(參見文獻[15]的概述)。於文獻[16]中，定向音頻編碼(DirAC、Directional Audio Coding)是一種受歡迎的方法，用於將該記錄轉換為一表示，該表示係由該聲音方向和擴散性的一音頻頻譜和參數輔助資訊所組成。它用於文獻[11]中的聲學變焦和文獻[14]中的虛擬麥克風的應用。

【0014】 這裡提出的方法能夠藉由一單個FOA(First-Order Ambisonics)麥克風的該記錄來實現6DoF再現。來自一單個空間位置的記錄已被用於3DoF再現或聲學變焦。但是，就發明人所知，到目前為止還沒有提出用於從這種資料進行交互式、完全6DoF再現的方法。經由整合關於在該記錄中該聲源的該距離的資訊，其可以實現該6DoF再現。該距離資訊被合併到DirAC的該參數表示中，使得該聆聽者的該改變視角係被正確映射。

【0015】 沒有任何該高保真度環繞聲聲場表示(無論是常規FOA或

HOA 高保真度環繞聲還是DirAC風格的參數聲場表示)能提供足夠的資訊，以允許6DoF應用程式所需的該聆聽者位置的一平移，因為在該聲音場景中的物件距離和絕對物件位置都不是以這些格式決定的。應該注意的是，該聆聽者位置的該移動可以轉移為該聲音場景在該相反方向上的一等效移位。

【0016】 當在6DoF中移動時的一典型問題如圖1b所示。讓我們假設使用高保真度環繞聲在位置A處描述該聲音場景。在這種情況下，來自來源A和來源B的聲音從相同方向到達，即它們具有相同的到達方向(DOA、Direction-Of-Arrival)。如果一個人移動到位置B處，來源A和來源B的該DOA是不同的。使用該聲場的一標準高保真度環繞聲描述，亦即沒有附加資訊，在給定位置A處的高保真度環繞聲信號時，其係不可能計算在位置B處的高保真度環繞聲信號。

【0017】 本揭露的一個目的是一方面提供一增強聲場描述或另一方面提供一改進的聲場描述的一生成，其允許一改進的、或靈活的、或有效的處理。

【0018】 該目的經由申請專利範圍第1項的用於生成一增強聲場描述的一裝置、申請專利範圍第8項的用於生成一修改聲場描述的一裝置、申請專利範圍第46項的生成一增強聲場描述的一方法，申請專利範圍第47項的生成一修改聲場描述的一方法，申請專利範圍第48項的一計算機程式或申請專利範圍第49項的一增強聲場描述來實現。

【0019】 本揭露基於以下發現：與一參考位置相關的典型聲場描述需要附加資訊，以便這些聲場描述可以被處理，使得與該原始參考位置無關但是與另一個參考位置的一修改聲場描述可以被計算。至此，與該聲場的空間資訊有關的後設資料(metadata)被生成，並且該後設資料與

該聲場描述一起對應於該增強聲場描述，其可以例如發送或存儲的。為了從該聲場描述和該後設資料生成一修改聲場描述，具體地，該後設資料與該聲場描述的空間資訊有關，使用該空間資訊、該聲場描述以及指示從一參考位置到一不同參考位置的一平移的一平移資訊，該修改聲場描述被計算。因此，由一聲場描述和與該聲場描述下的該聲場的空間資訊相關的後設資料所組成的該增強聲場描述，其係被處理，以獲得一修改聲場描述，該修改聲場描述係與由附加平移資訊定義的一不同參考位置相關的，例如，其可以在解碼器側被提供或被使用。

【0020】 然而，本揭露不僅涉及一編碼器/解碼器場景，但也可以被應用於一應用程式中，其中包含基本上在同一個位置發生的該增強聲場描述的該生成和該修改聲場描述的該生成。例如，該修改聲場描述可以是該修改聲場本身的一描述，或者實際上是在聲道信號中、雙耳信號的該修改聲場，或者再次是一參考位置相關的聲場，然而現在是相關到該新的或不同的參考位置而不是該原始的參考位置。例如，這樣的一應用將處在於一虛擬實境場景中，其中存在一聲場描述以及一後設資料，並且其中一聆聽者從給出該聲場的該參考位置移出並移動到不同的參考位置，並且其中，接著，在該虛擬區域中移動的該聆聽者的該聲場被計算，以對應於該聲場，但現在係在用戶移動到的該不同參考位置處。

【0021】 在一特定實施例中，該增強聲場描述具有與該(第一)參考位置相關的一第一聲場描述以及與不同於該(第一)參考位置的一另一(第二)參考位置相關的一第二聲場描述，並且後設資料具有關於該參考位置和該另一參考位置的資訊，諸如從一預定原點指向這些參考位置的向量。或者，該後設資料可以是指向該參考位置或該另一參考位置的一單個向量和在該兩個參考位置之間延伸的一向量，該兩個不同的聲場描

述與之相關。

【0022】 該聲場描述可以是非參數聲場描述，例如一階高保真度環繞聲描述或更高階的高保真度環繞聲描述。可選地或另外地，該聲場描述可以是DirAC描述或其他參數聲場描述，或者一個聲場描述例如可以是一參數聲場描述而另一個聲場描述可以是例如一非參數聲場描述。

【0023】 因此，該聲場描述可以針對每個聲場描述可生成該聲場的一DirAC描述，其具有一個或多個降混信號和個別的方向資料以及對不同時間頻率區間的可選擇的擴散資料。在此上下文中，該後設資料生成器被配置為針對兩個聲場描述生成幾何後設資料，使得可以該從後設資料中識別該參考位置和該附加參考位置。接著，可以從兩個聲場描述中擷取個別來源，並且為了生成一增強或修改的聲場描述而執行附加處理。

【0024】 高保真度環繞聲已成為虛擬、增強和混合實境應用環境中針對3D音頻最常用的格式之一。已經開發的各種各樣的音頻擷取和生產工具，它們係以高保真度環繞聲格式以生成一輸出信號。為了在交互式虛擬實境(VR)應用中呈現高保真度環繞聲編碼內容，該高保真度環繞聲格式係被轉換為用於再現的一雙耳信號或通道。在該上述應用中，該聆聽者通常能夠以交互方式改變在所呈現場景中他/她的方向至該程度，使得他/她能夠在該聲音場景中旋轉他/她的頭部，從而實現三個自由度(3DoF，即，俯仰、偏轉角及滾動)並且仍然可以體驗到一合適的音質。這是根據該頭部方向經由在渲染之前旋轉該聲音場景所實現，這可以以低計算複雜度而被實現並且是高保真度環繞聲表示的一優點。然而，在諸如VR的新興應用中，其期望允許該用戶在該聲音場景中自由移動而不僅僅是方向的改變(所謂的「六個自由度」或6DoF)。其結果係為，需要信號處理來改變該聲音場景的該角度(即，沿著x軸、y軸或z軸在該

聲音場景內虛擬地移動)。然而，高保真度環繞聲的一個主要缺點是該程序從該聲場中的一單個角度描述了該聲場。具體來說，它不包含該聲音場景中聲源的該實際位置的資訊，該實際位置的資訊其可允許移動該聲音場景(「平移」)，因為它是6DoF所需的。本揭露的描述提供了高保真度環繞聲的幾個擴展，以克服該問題並且還促進該平移，並因此實現真正的6DoF。

【0025】 一階高保真度環繞聲(FOA)錄音可以藉由耳機而被處理和再現。它們可以被旋轉以考慮該聆聽者頭部方向。然而，虛擬實境(VR)系統允許該聆聽者以六個自由度(6DoF)移動，即三個旋轉自由度加三個過渡自由度。此處，該聲源的該視角和距離係取決於該聆聽者的位置。一種技術以促進6DoF係被描述。特別地，一FOA記錄係被使用一參數模型來描述，該參數模型係基於該聆聽者的位置和關於到該來源的該距離的資訊來修改。該方法經由一聽力測試來評估，比較該聆聽者可以自由移動的一合成聲音場景的不同雙耳渲染。

【0026】 在進一步較佳的實施例中，該增強聲場描述由一輸出界面輸出，用於生成用於傳輸或存儲的一輸出信號，對一時間幀內，其中該輸出信號包括在該時間幀內從該聲場和該空間資訊所導出的一個或多個音頻信號。特別地，該聲場生成器在進一步的實施例中適於從該聲場導出方向資料，該方向資料指的是在一時間區段或一頻率區間內聲音的一到達方向，並且後設資料生成器被配置為導出該空間資訊，作為將一距離資訊與該方向資料相關聯的資料項目。

【0027】 特別地，在這樣的一實施例中，一輸出界面被配置為生成該輸出信號，使得該時間幀的資料項目係被鏈結到該不同頻率區間內的該方向資料。

【0028】 在一另一實施例中，該聲場生成器還被配置為生成該聲場的一時間幀的多個頻率區間內的一擴散資訊，其中，該後設資料生成器被配置為僅生成用於與一預定值不同的或者與無窮大不同的一頻率區間內的一距離資訊，或者當該擴散低於一預定或自適應閾值時，它根本生成該頻率區間內的一距離值。因此，對於具有一高擴散的時間/頻率區間，任何距離值一點也不被生成，或者生成由一解碼器以一某種方式解釋的一預定距離值。因此，確保對於具有一高擴散的時間/頻率區間，任何與距離相關的渲染係不被執行，因為一高擴散表示對於這樣的時間/頻區間，聲音不是來自一某個局部的來源，但來自任何方向，因此，無論該聲場是在該原始參考位置還是在不同的或新的參考位置處，其被感知都是相同的。

【0029】 關於聲場計算器，較佳實施例包括一平移界面，其用於提供該平移資訊或指示一預期聆聽者對該修改聲場的一旋轉的旋轉資訊，用於將該後設資料提供給該聲場計算器的一後設資料提供器和用於將該聲場描述提供給該聲場計算器的一聲場供應器，以及附加的用於輸出包括該修改聲場描述和修改後設資料的該修改聲場的一輸出界面，該修改後設資料使用該平移資訊從該後設資料被導出，或該輸出界面輸出多個揚聲器通道，每個揚聲器通道與一預定義的揚聲器位置相關，或者該輸出界面輸出該修改聲場的一雙耳表示。

【0030】 在一個實施例中，該聲場描述包括多個聲場分量。多個聲場分量包括一全向分量和至少一個方向分量。這樣的聲場描述例如是具有一全向分量和三個方向分量 X 、 Y 、 Z 的一階高保真度環繞聲聲場描述，或者這樣的一聲場是一個高階高保真度環繞聲描述，其包括該全向分量、相對於 X 、 Y 和 Z 方向的三個方向分量、以及另外的與 X 、 Y 、 Z 方向

之外的其他方向相關的方向分量。

【0031】 在一個實施例中，該裝置包括一分析器，用於分析該聲場分量，以針對不同的時間或頻率區間導出到達方向(DoA、direction of arrival)資訊。該裝置還具有一平移變換器，用於使用該DoA資訊和該後設資料以計算每個頻率或時間區間的修改DoA資訊，其中後設資料涉及一深度圖，該深度圖將一距離與包含在兩個聲場描述中的一來源相關聯的，其係經由例如相對於兩個不同參考位置和該距離/位置或該等參考位置而使用兩個角度的三角測量處理所獲得。這可以應用於一全頻帶表示或一時間幀的不同頻率區間。

【0032】 此外，該聲場計算器具有一距離補償器，用於使用一距離補償資訊計算該修改的聲場，該距離補償資訊係取決於使用該後設資料計算的該距離、並且取決於與該時間或頻率區間相關聯的一新距離，其中，該後設資料對於一來源的每個頻率或時間區間是相同的，針對每一個或一些時間/頻率區間，該來源係不相同，該新距離與該修改DoA資訊相關。

【0033】 在一個實施例中，該聲場計算器計算從該參考位置指向經由該聲場分析所獲得的一聲源的一第一向量。此外，該聲場計算器計算從該不同參考位置指向該聲源的一第二向量，並且此計算係使用該第一向量和該平移資訊所完成，其中該平移資訊定義從該參考位置到該不同參考位置的一平移向量。並且接著，使用該第二向量來計算從該不同參考位置到該聲源的一距離。

【0034】 此外，該聲場計算器被配置為除了該平移資訊之外還接收一旋轉資訊，該旋轉資訊指示該聆聽者的頭部在由俯仰、偏轉和滾動所給出的該三個旋轉方向之一的一旋轉。該聲場計算器接著被配置為執行

該旋轉變換，以使用該旋轉資訊俾旋轉針對一聲場的一修改的到達方向資料，其中，從經由該聲場描述和該平移資訊的一聲音分析所獲得的一到達方向資料，該修改的到達方向資料被導出。

【0035】 在一個實施例中，該聲場計算器被配置為經由一聲音分析來決定來自該聲場描述的來源信號、以及與該參考位置相關的該來源信號的方向。

【0036】 接著，計算與該不同參考位置相關的該聲源的新方向，並且此計算係使用該後設資料來完成，並且接著與該不同參考位置相關的該聲源的距離資訊被計算，接著使用該距離資訊和該聲源的該新方向來合成該修改聲場。

【0037】 在一個實施例中，經由將該聲源信號平移到由關於一再現設置的該新方向資訊所給出的一方向來執行一聲場合成，並且在執行該平移操作之前或執行該平移操作之後，使用該距離資訊完成該聲源信號的一縮放。

【0038】 在一另一實施例中，該聲源信號的一擴散部分被添加到該聲源信號的一直接部分，該直接部分在被添加到該擴散部分之前，藉由該距離資訊來修改該直接部分。

【0039】 特別地，執行該聲源合成較佳地以一頻譜表示，其中針對每個頻率區間來計算該新的方向資訊，其中針對每個頻率區間來計算該距離資訊，並且其中使用該頻率區間的該音頻信號對每個頻率區間的一直接合成是使用該頻率區間的一音頻信號來執行的，從該新方向資訊導出的針對該頻率區間的一平移增益和從該頻率區間的該距離資訊所導出的針對該頻率區間的一縮放因子係被執行。

【0040】 此外，使用從來自該頻率區間的該音頻信號所導出的一擴

散音頻信號以及使用由針對該頻率區間的該信號分析所導出的一擴散參數來執行一擴散合成，並且接著針對該時間或頻率區間，該直接信號和該擴散信號被組合以獲得一合成音頻信號，並且接著後針對其他時間/頻率區間，使用音頻信號，執行一頻率-時間轉換，以獲得一時域合成音頻信號，俾作為該修改聲場。

【0041】 因此，通常，該聲場計算器被配置為針對每個聲源合成與該不同參考位置相關的一聲場，例如，針對每個來源，使用該來源信號的該新方向來處理一來源信號，以獲得與該不同/新參考位置相關的該來源信號的一聲場描述。此外，在處理該來源信號之前或在使用該方向資訊處理該來源信號之後，該來源信號被修改。並且，最後，將該來源的該聲場描述相加在一起以獲得與該不同參考位置相關的該修改聲場。

【0042】 在進一步的實施例中，並且特別地，用於從該聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述，該聲場計算器使用關於該第一聲場的該空間資訊描述、使用關於該第二聲場描述的該空間資訊、以及使用指示一參考位置到一不同參考位置的一平移的該平移資訊來計算該修改的聲場。特別地，該後設資料可以例如是指向該聲場描述的該參考位置的一向量和從該相同原點指向該第二聲場描述的另一參考位置的另一向量。

【0043】 為了解決一平移資訊，經由對該第一和該第二聲場描述應用一來源分離、或波束成形、或一般任何種類的聲源分析來生成對象。接著，不管這些對象是寬頻帶對象還是各個時間/頻率區間的對象，計算所有對象的該到達方向資訊。接著，將從不同聲場描述中所擷取的對象彼此匹配，以便找到至少一個匹配對象，即在該第一和該第二聲場描述中都出現的一對象。舉例而言，經由使用該對象信號和/或到達方向資訊

或其他資訊的一相關或一致性計算來執行該匹配。

【0044】 因此，對於一匹配對象，該程序的該結果是存在與該參考位置有關的一第一DoA資訊以及與該另一參考位置有關的該第二DoA資訊。接著，使用關於該參考位置的該資訊或該關聯後設資料中所包括的該參考位置，基於三角測量，以計算該匹配對象的該位置，尤其是該匹配對象到該參考位置或該另一參考位置的該距離。

【0045】 該資訊，特別是該匹配對象的該位置資訊接著被用於基於該估計的位置和該期望的位置(即在平移之後)，使用一距離補償處理來修改每個匹配的對象。為了計算該新聆聽者位置的該新DoA資訊，來自兩個參考位置的該舊DoA資訊和該平移資訊被使用。基本上，因為每個匹配的對象都出現在兩個聲場描述中，可以對兩個個別的聲場描述執行該處理。然而，根據較佳實施例，具有最接近該平移之後的該新聆聽者位置的一參考位置的該聲場描述會被使用。

【0046】 接著，該新的DoA係用於計算與該不同參考位置相關的該匹配對象的一新聲場描述，即，該用戶已經移動到的參考位置。接著，為了也包含該不匹配的對象，也計算這些對象的聲場描述，但是使用舊的DoA資訊。並且最後，經由將所有個別的聲場描述加在一起來生成該修改的聲場。

【0047】 經由對該虛擬高保真度環繞聲信號應用一單個旋轉，可以實現任何方向變化。

【0048】 因此，該後設資料不用於直接提供一對象到一參考位置的該距離。相反地，該後設資料係被提供以用於識別兩個或更多個聲場描述中的每一個的該參考位置、並且基於例如三角測量處理步驟俾計算一參考位置和一某個匹配對象之間的該距離。

【0049】 本揭露的較佳實施例係隨後參照附圖描述。

【發明內容】

【0050】 依據本揭露之一特色，本揭露提出一種用於生成一增強聲場描述的裝置，其包含：一聲場生成器、及一後設資料生成器。該聲場生成器用於生成至少一個聲場描述，該聲場描述指示相對於至少一個參考位置的一聲場。該後設資料生成器用於生成與該聲場的空間資訊有關的後設資料。其中，該至少一個聲場描述和該後設資料構成該增強聲場描述。

【0051】 依據本揭露之另一特色，本揭露提出一種用於根據一聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述的裝置，包括一聲場計算器。該聲場計算器用於使用該空間資訊、該聲場描述和指示一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改的聲場。

【0052】 依據本揭露之又一特色，本揭露提出一種生成一增強聲場描述的方法，包括：生成至少一個聲場描述，該聲場描述指示關於至少一個參考位置的一聲場；以及生成與該聲場的空間資訊有關的後設資料；其中，該至少一個聲場描述和該後設資料構成該增強聲場描述。

【0053】 依據本揭露之再一特色，本揭露提出一種從一聲場描述和與該聲場描述的空間資訊有關的後設資料生成一修改聲場描述的方法，該方法包括：使用該空間資訊、該聲場描述和指示從一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改聲場。

【0054】 依據本揭露之更一特色，本揭露提出一種計算機程式，當其用於一計算機或處理器上運行時，執行本揭露前述的方法。

【0055】 依據本揭露之再一特色，本揭露提出一種增強的聲場描述，包括至少一個聲場描述及後設資料，該至少一個聲場描述指示關於至少一個參考位置的一聲場，該後設資料與該聲場的空間資訊有關。

【0056】 本揭露主題的各種目的、特徵、面向和優點將從以下對優選實施例的詳細描述以及附圖中變得更加明顯，附圖中相同的附圖標記表示相同的部件。

【0057】 所示實施例在附圖中以例子，而非限制的方式顯示，其中相同的參考標號表示相似的元件。

【圖式簡單說明】

【0058】

圖1a係用於生成一增強聲場描述的一裝置的一較佳實施例。

圖1b係說明本揭露的一示例性問題的一圖式。

圖2係用於生成一增強聲場描述的該裝置的一較佳實施方式。

圖3a係顯示包括音頻資料和音頻資料的輔助資訊的該增強聲場描述。

圖3b係顯示一增強聲場的一進一步說明，該增強聲場包括與空間資訊(例如每個聲場描述的幾何資訊)有關的音頻資料和後設資料。

圖4a係顯示針對生成一修改聲場描述的一裝置的一實現。

圖4b係顯示生成一修改聲場描述的一裝置的一另一實現。

圖4c係顯示具有一參考位置/地點A、一另一參考位置/地點B、以及由於平移的一不同參考位置的一場景。

圖5係顯示在一般意義上的空間音頻的該六個DoF再現。

圖6a係顯示用於實現一聲場計算器的一較佳實施例。

圖6b係顯示用於計算相對於一新的/不同的參考位置的一聲源的一新的DoA和一新距離的一較佳實施方式。

圖6c係顯示一6DoF再現的一較佳實施例，包括用於生成一增強聲場描述的一裝置，例如用於每個個別的聲場描述，和用於為該匹配的來源生成一修改的聲場描述的一裝置。

圖7係顯示用於選擇第一和第二聲場描述之一的一較佳實施例，其用於一寬頻帶或窄頻帶對象的一修改聲場的該計算。

圖8係顯示用於根據例如一單聲道信號的一音頻信號和到達方向資料生成一聲場描述的一示例性設備。

圖9係顯示該聲場計算器的一另一較佳實施例。

圖10係顯示用於生成一修改的聲場描述的該裝置的一較佳實施方式。

圖11係顯示用於生成一修改的聲場描述的一裝置的一另一較佳實施方式。

圖12a係顯示一現有技術的DirAC分析實現。

圖12b係顯示一現有技術的DirAC合成實現。

【實施方式】

【0059】 針對上述高保真度環繞聲/定向音頻編碼(Ambisonics/DirAC、Ambisonics/Directional Audio Coding)表示以致能六個自由度(6DoF、six-Degrees-Of-Freedom)應用程式，其有必要以一方式擴展這些表示，以提供平移處理所缺失的資訊。應注意的是，該擴展可以例如是1)將該對象的該距離或位置添加到該現有場景表示，

和/或2)添加能促進各個對象分離的該過程的資訊。

【0060】 此外，實施例的一目的是保留/重新使用現有(非參數或參數)高保真度環繞聲系統的該結構，以在這種意義上提供與這些表示/系統的向後兼容性：

- 該擴展表示可以轉換為現有的非擴展表示(例如，針對渲染)；以及
- 允許在使用該擴展表示時，重新使用現有的軟體和硬體實現。

【0061】 在下文中，幾種方法將被描述，即一種有限(但非常簡單)的方法和三種不同的擴展的高保真度環繞聲格式，以實現6DoF。

【0062】 使用兩個或更多個高保真度環繞聲信號來描述該聲音場景，每個高保真度環繞聲信號描述一不同位置處的該聲音場景，或者換句話說，從不同的角度描述聲音場景。其係假設該相對位置是已知的。在該聲音場景中的一期望位置處的一修改的高保真度環繞聲信號係從該輸入的高保真度環繞聲信號產生。一基於信號或基於參數的方法可被用於在該期望位置處生成一虛擬高保真度環繞聲信號。

【0063】 多點高保真度環繞聲表示的該概念適用於傳統和參數(DirAC風格)的高保真度環繞聲。

【0064】 在一基於信號的平移實施例中，使用以下步驟計算在一期望位置(即，平移之後)的一虛擬高保真度環繞聲信號：

1. 經由將來源分離應用於每個傳統的高保真度環繞聲信號來生成對象。
2. 針對每個傳統的高保真度環繞聲信號計算所有對象的該DOA。
3. 從一個傳統的高保真度環繞聲信號中擷取的該對象與從其他傳統的高保真度環繞聲信號中擷取的該對象相匹配。基於該相應

的DOA和/或該信號(例如，經由相關性/一致性)來執行該匹配。

4. 基於三角測量來估計該匹配對象的該位置。
5. 基於該估計的位置和該期望的位置(即，在平移之後)，使用一距離補償濾波器(distance-compensation filter)修改每個匹配的對象(單通道輸入)。
6. 針對每個匹配的對象計算在該期望的位置(即，平移後)的該到達方向(DOA)。這個到達方向由DOA'表示。
7. 為每個匹配的對象計算一高保真度環繞聲對象信號。該高保真度環繞聲對象信號被生成，使得該匹配對象具有一到達方向DOA'。
8. 為每個非匹配對象計算一高保真度環繞聲對象信號。該高保真度環繞聲對象信號被生成，使得該非匹配對象具有一到達方向DOA。
9. 經由將所有高保真度環繞聲對象信號加在一起來獲得該虛擬高保真度環繞聲信號。

【0065】 根據一另一實施例，在一基於參數的平移實施例中使用以下步驟計算在一期望位置(即，平移之後)的一虛擬高保真度環繞聲信號：

1. 一聲場模型被假設。該聲場可以被分解為一個或多個直接聲音分量和擴散聲音分量。該直接聲音分量包括信一號號和位置資訊(例如，在極坐標或笛卡爾坐標中)。或者，該聲場可以被分解為一個或多個直接/主要聲音分量和一個殘餘聲音分量(單聲道或多聲道)。
2. 使用該輸入的高保真度環繞聲信號，該假定聲場模型的該信號

分量和參數被估計。

3. 根據在該聲音場景中的該期望平移或期望位置，該信號分量和/或參數被修改。
4. 使用該修改的信號分量和修改的參數，該虛擬高保真度環繞聲信號被生成。

【0066】 生成多點高保真度環繞聲信號對於計算機生成及產生的內容，以及經由麥克風陣列或空間麥克風(例如，B格式麥克風)的自然記錄的該上下文來說是簡單的。在該實施例中，較佳地在步驟2之後執行一來源匹配或在步驟3之前執行一三角測量計算。此外，兩個實施例的一個或多個步驟也可以用在相應的其他實施例中。

【0067】 經由對該虛擬高保真度環繞聲信號應用一單個旋轉，可以實現方向的一改變。

【0068】 圖1a係顯示用於生成一增強聲場描述的一裝置，其包括一聲場(描述)生成器100，用於生成至少一個聲場描述，其指示關於至少一個參考位置的一聲場。此外，該裝置包括一後設資料生成器110，用於生成與該聲場的空間資訊有關的後設資料。該後設資料接收該聲場作為一輸入，或者替代地或另外地，接收關於聲源的分離資訊。

【0069】 該聲場描述生成器100和該後設資料生成器110兩者的輸出都構成該增強聲場描述。在一個實施例中，該聲場描述生成器100和該後設資料生成器110的該輸出都可以在一組合器120或輸出界面120內被組合，以獲得該增強聲場描述，其包括由該後設資料生成器110生成的該空間後設資料或該聲場的空間資訊。

【0070】 圖1b圖式說明了本揭露所解決的情況。例如，該位置A是該至少一個參考位置，並且一聲場係由來源A和來源B所生成，並且位

於該位置A的例如一某種實際或虛擬麥克風檢測來自來源A和來源B的該聲音。該聲音是來自該等發射聲音來源的該聲音的一疊加。這表示該聲場描述係由該聲場描述生成器所生成的。

【0071】 另外，經由某種實現，該後設資料生成器將導出關於來源A的一空間資訊和關於來源B的另一空間資訊，諸如這些來源到該參考位置的距離，諸如位置A。

【0072】 自然地，該參考位置另外也可以是位置B。接著，該實際或虛擬麥克風將被放置在位置B，並且該聲場描述將是一聲場，例如，由一階高保真度環繞聲分量、或更高階的高保真度環繞聲分量、或任何其他聲音分量所表示的，其具有該潛力以描述關於至少一個參考位置(即位置B)的一聲場。

【0073】 接著，該後設資料生成器可以生成關於該聲源的該資訊、聲音來源A到位置B的該距離或者來源B到位置B的該距離。當然，關於聲源的另外資訊可以是相對於一參考位置的該絕對或相對位置。該參考位置可以是一個一般坐標系統的該原點，或者可以位在與一個一般坐標系統的該原點有一個定義關係的位置處。

【0074】 其他後設資料可以是一個聲源的該絕對位置和相對於該第一聲源的另一個聲源的該相對位置等等。

【0075】 圖2係顯示用於生成一增強的聲場描述的一裝置，其中該聲場生成器包括用於該第一聲場的一聲場生成器250、用於該第二聲場的一聲場生成器260、以及用於一個或多個聲場的一任意數量的聲場生成器，該一個或多個聲場例如是一第三、第四等聲場。另外，該後設資料被配置為計算並向該組合器120轉發關於該第一聲場和該第二聲場的一資訊。所有這資訊被該組合器120所使用，以便生成該增強的聲場描

述。因此，該組合器120還被配置為一輸出界面，以生成該增強的聲場描述。

【0076】 圖3a係顯示作為一資料流的一增強聲場描述，其包括一第一聲場描述330、一第二聲場描述340以及與其相關聯的後設資料350，該後設資料350包括關於該第一聲場描述和該第二個聲場描述的資訊。例如，該第一聲場描述可以是一B格式描述、或一更高階描述、或任何其他描述，該任何其他描述允許決定以一全頻帶表示或一頻率選擇表示的聲源的一方向分佈。因此，例如該第一聲場描述330和該第二聲場描述340也可以是針對該不同參考位置的參數聲場描述，其具有例如一降混信號和不同時間/頻率區間的到達方向資料。

【0077】 然而，該第一和該第二聲場描述的該幾何資訊350對於包括在該第一聲場描述330中的所有來源或者對於在該第二聲場描述340中的該來源分別是相同的。因此，當示例性地在該第一聲場描述330中存在三個來源和關於該第一聲場描述存在一幾何資訊時，則該幾何資訊對於該第一聲場描述中的該三個來源是相同的。類似地，例如當在該第二聲場描述中存在五個來源時，則被包括在該後設資料350中的關於該第二聲場的該幾何資訊對於該第二聲場場描述中的所有來源是相同的。

【0078】 圖3b係顯示圖3a的該後設資料350的一示例性構造。在一個實施例中，該參考位置351可以被包括在該後設資料中。然而，該參考位置資訊351也可以被省略的情況下，這不一定是必要的。

【0079】 對於該第一聲場，所給出一第一幾何資訊，其可以例如是關於圖4c中所示的向量A的一資訊，該向量A從一原點指向與該第一聲場相關的該參考位置/地點A。

【0080】 例如，該第二幾何資訊可以是關於從該原點指向該第二參

考位置/地點B的該向量B的一資訊，該第二聲場描述係與之相關。

【0081】 A和B是兩個聲場描述的該參考位置或記錄位置。

【0082】 替代的幾何資訊例如可以是關於在參考位置A和該另一的參考位置B之間延伸的該向量D的一資訊和/或一原點和從該原點指向兩個點之一的一向量。因此，包括在該後設資料中的該幾何資訊可以包括向量A和向量D、或者可以包括向量B和向量D、或者可以包括向量A和向量B而沒有向量D、或者可以包括其他資訊，由此該參考位置A和該參考位置B可以在一某個三維坐標系中被識別。然而，該相同的考慮附加地被應用於一個二維聲音描述以及特別地在圖4c中說明了僅顯示該二維的情況。

【0083】 圖4a係顯示用於根據一聲場描述以及與該聲場描述的空間資訊有關的後設資料生成一修改聲場描述的一裝置的一較佳實施方式。特別地，該裝置包括一聲場計算器420，其使用後設資料、該聲場描述、以及平移資訊以生成該修改的聲場，該平移資訊係指示從一參考位置到一不同參考位置的一平移。

【0084】 在一個實施例中，該聲場計算器420連接到一輸入界面400，用於接收該增強的聲場描述，例如，參考圖1a或2討論，接著該輸入界面400一方面分離該聲場描述，亦即由圖1a的區塊100或圖2的區塊210所生成的。此外，該輸入界面400將該後設資料從該增強聲場描述分離，亦即圖3a的項目350或圖3b的可選的351和352至354。

【0085】 此外，一平移界面410從一聆聽者獲得該平移資訊和/或附加或分離的旋轉資訊。該平移界面410的一實現可以是一頭部追蹤單元，其不僅追蹤在一虛擬實境環境中的一頭部的該旋轉，而且還追蹤該頭部從一個位置(亦即圖1b中的位置A)到另一個位置(亦即圖1b中的位置B)

的一平移。

【0086】 圖4b係顯示與圖1a類似的另一種實現方式，但與一編碼器/解碼器方案無關，但與一個一般方案有關，其中該後設資料供應係由一後設資料提供器402所指示的，由一聲場提供器404指示的該聲場供應在沒有一某個輸入界面的情況下完成，其分隔一編碼或增強聲場描述的，但舉例而言，在一虛擬實境應用中的一實際方案中全部完成。然而，本揭露不限於虛擬實境應用，而是還可以在任何其他應用中實現，其中，與一參考位置相關的聲場的該空間音頻處理是有用的，以將與一第一參考位置相關的一聲場轉換至與一不同的第二參考位置相關的另一聲場。

【0087】 該聲場計算器420接著生成該修改聲場描述、或者生成一(虛擬)揚聲器表示、或者生成諸如用於一耳機再現的一雙聲道表示的一個雙耳表示。因此，該聲場計算器420可以生成一修改聲場描述，以作為該修改聲場，其基本上與該原始聲場描述相同，但是現在係相對於一新的參考位置。在一替代實施例中，可以為諸如5.1方案的一預定揚聲器設置或具有更多揚聲器的一揚聲器設置生成一虛擬或實際揚聲器表示，並且特別地，具有揚聲器的一個三維佈置而不是僅一個二維佈置，亦即一揚聲器安排，其揚聲器相對於該用戶位置而被升高。對虛擬實境應用特別有用的其他應用係用於雙耳再現的應用，即用於可應用於該虛擬實境用戶頭部的一耳機。

【0088】 示例性地，隨後描述的圖6係顯示一種情況，其中一DirAC合成器僅在諸如全向或壓力分量的一下混合分量上操作，而在關於圖12b所示的一另一替代實施例中，該DirAC合成器在整個聲場資料上操作，即在圖12b中的該實施例中，該完整分量表示係具有全向分量 w 和三個方向分量 x 、 y 、 z 的一場描述。

【0089】 圖4c係顯示作為本揭露較佳實施例基本的該場景。該圖式係顯示一第一參考位置/地點A、一第二參考位置/地點B、和兩個不同的聲源A和B、以及一平移向量I。

【0090】 聲源A和B兩者都被包括在與參考位置A相關的該聲場描述和與參考位置B相關的該第二聲場描述中。

【0091】 為了計算來源A的該距離，例如，到該第一參考位置或到該第二參考位置，與A和B相關的兩個不同聲場描述都經歷一來源分離程序，接著，經由這些不同聲音分離程序所獲得的該來源的一匹配被獲得。例如，這將導致來源A。針對該第一聲場描述以及該第二聲場描述，在來源分離演算法中，來源A被找到。當從與參考位置A相關的該第一聲場描述獲得時，來源A的該到達方向資訊將是角度 α 。另外，但現在相同來源A的到達方向資訊係從與該另一參考位置B相關的該第二聲場描述獲得，該相同來源A的該到達方向資訊將是角度 β 。

【0092】 現在，基於已知的或可計算的距離D，其例如可從該聲場描述的該後設資料獲得或可計算，並且基於該兩個角度 α 和 β ，由來源A定義的該三角形，該參考位置A和該參考位置B是被完全定義的。因此，從來源A到該參考位置A的該距離或從來源A到該參考位置B的該距離或來源A的該一般位置係為可以計算，亦即，例如經由三角測量處理操作，從該原點指向來源A的該實際位置的該向量可以被計算。該位置或距離兩者都代表一距離或一位置的資訊。

【0093】 接著可以對每個匹配的來源執行該相同的程序，亦即也可對來源B執行該相同的程序。

【0094】 因此，每個匹配的來源的一距離/位置資訊被計算，接著，可以以該距離/位置是完全已知的或者係例如由附加後設資料所給出的

情況處理每個匹配的來源。然而，僅需要該第一聲場描述和該第二聲場描述的該幾何資訊，而不是每個個別來源的任何距離/深度資訊。

【0095】 圖8係顯示用於執行與該DirAC合成器不同的一合成的另一實現。例如，當一聲場分析器為每個來源信號生成一分離的單聲道信號S和一原始到達方向時，並且根據該平移資訊，當一新的到達方向被計算時，則例如圖8的該高保真度環繞聲信號生成器430將被用於生成針對該聲源信號的一聲場描述，即該單聲道信號S，但是針對該新的到達方向(DoA)資料，其由一水平角度 θ 或一仰角 θ 和一方位角 φ 組成的。接著，由圖4b的該聲場計算器420執行的一程序將生成例如一個一階高保真度環繞聲聲場表示，其針對具有該新的到達方向的每個聲源，並且然後，可以使用一縮放因子來執行每個聲源的一進一步修改，該縮放因子取決於該聲場到該新參考位置的該距離，並且然後，來自個別來源的所有聲場可以再一次例如在一高保真度環繞聲表示被相關至一某個的新參考位置相互疊加以最終獲得該修改聲場。

【0096】 當解釋由圖6的一DirAC分析器422、422a、422b處理的每個時間/頻率區間係表示一某個的(頻寬受限)聲源時，則該高保真度環繞聲信號生成器430可以被使用，以取代該DirAC合成器425、425a、425b，來為每個時間/頻率區間生成一完整的高保真度環繞聲表示，其使用該降混信號或壓力信號或該時間/頻率區間的全向分量，以作為圖8的該「單聲道信號S」。接著，針對W、X、Y、Z分量中的每一個，頻率-時間轉換器中的一單獨頻率-時間轉換將導致一聲場描述，其係不同於圖4c中所示的聲場描述。

【0097】 該場景係從該麥克風的該視點(PoV、Point of View)而被記錄，該位置用作為該參考坐標系的該原點。該場景必須從該聆聽者

的該視點(PoV)中再次生成，該聆聽者在6DoF中被追蹤，參見圖5。這裡係顯示一單個聲源用於圖式說明，該關係適用於每個時間頻率區間。

【0098】 圖5係顯示空間音頻的該6DoF再現。一聲源藉由一麥克風而被記錄，其具有在該距離 d_r 的該到達方向 $\mathbf{r}_r(\text{DoA } \mathbf{r}_r)$ ，該距離 d_r 及該到達方向 $\mathbf{r}_r$ 係相對於該麥克風位置和方向(黑線和弧線)的。它必須相對於具有該到達方向 $\mathbf{r}_1(\text{DoA } \mathbf{r}_1)$ 和距離 d_1 (虛線)的該移動聆聽者而被再現。這必須考慮該聆聽者平移 l 和旋轉 o (點線)。該DOA被表示為指向該來源的一單位長度的向量。

【0099】 在坐標 $\mathbf{d}_r \in \mathbb{R}^3$ 的該聲源是由表示為該單位向量 $\mathbf{r}_r = \mathbf{d}_r / \|\mathbf{d}_r\|$ 的該到達方向(DoA)所記錄的。該到達方向(DoA)可以經由分析該記錄而被估計。它來自該距離 $d_r = \|\mathbf{d}_r\|$ 。假設該資訊可以從每個來源的該後設資料導出，或者通常從圖3b的項目352、353、354導出，並且可以表示為具有從該記錄位置到該距離(例如以米為單位給出)的任何方向 $\mathbf{r}$ 的距離描述，例如經由使用關於兩個不同參考位置和距離/位置或該參考位置的兩個角度的三角測量處理獲得的距離描述。

【0100】 該聆聽者在6DoF中被追蹤。在一給定的時間，相對於該麥克風，他位於 $\mathbf{l} \in \mathbb{R}^3$ 的一位置，並且相對於該麥克風的坐標系統具有一旋轉 $\mathbf{o} \in \mathbb{R}^3$ 。該記錄位置係被選為該坐標系統的原點，俾簡化符號。

【0101】 因此，該聲音必須以一不同的距離 d_1 再現，導致一音量改變，並且一不同的該到達方向 $\mathbf{r}_1(\text{DoA } \mathbf{r}_1)$ 是平移和隨後旋轉的結果。

【0102】 如以下分段所解釋，一種用於藉由基於一參數表示的專用變換從該聆聽者的角度獲得一虛擬信號的方法被概述。

【0103】 所提出的該方法可參見文獻[16]，其係基於針對參數空間聲音編碼的該基本DirAC方法。假設在該分析頻譜的每個時頻實例中存

在一個主要直接來源，並且這些可以獨立地處理。使用短時傅立葉變換 (STFT、Short Time Fourier transform) 將該記錄變換為一時頻表示。該時間幀索引用 n 表示，該頻率索引用 k 表示。該變換後的記錄接著被分析，以估計該複數頻譜 $P(k, n)$ 的每個時間頻率區間的方向 $r_r(k, n)$ 和擴散率 $\psi(k, n)$ 。在該合成中，該信號被分成一直接和擴散部分。此處，經由根據該揚聲器位置而平移該直接部分並添加該擴散部分，來計算揚聲器信號。

【0104】 參見圖6c，根據在6DoF中的該聆聽者觀點之轉換一個一階高保真度環繞聲(FOA)信號的方法可以被分為五個步驟。

【0105】 圖6c係顯示一6DoF再現的方法。B格式的該記錄FOA信號係由一DirAC編碼器處理，該編碼器計算該複數頻譜的每個時頻區間的方向和擴散值。經由該聆聽者的追蹤位置並根據針對每個來源(例如經由三角測量計算所導出)的一距離圖中給所出的該距離資訊，該方向向量接著被變換。根據該頭部旋轉，該所得的方向向量接著被旋轉。最後，在該DirAC解碼器中，用於8+4個虛擬揚聲器聲道的信號被合成。接著它們將被雙耳化。

【0106】 在該實施例中，該輸入信號係在該DirAC編碼器422中被分析，該距離資訊係從該距離圖 $m(r)$ 被添加，該距離圖 $m(r)$ 為每個(匹配的)來源給出一距離，接著該聆聽者被追蹤的平移和旋轉係在該新穎變換423和424中被應用。該DirAC解碼器425高保真度環繞聲合成用於8+4個虛擬揚聲器的信號，其另外被雙耳化427以用於耳機回放。應注意的是，由於在該平移之後該聲音場景的該旋轉是一獨立操作，所以它可以另外地應用在該雙耳渲染器中。為6DoF所轉換的該唯一參數是該方向向量。經由模型定義，該擴散部分係被假設為等向性和同質性的，因

此保持不變。

【0107】 該DirAC編碼器的輸入是B格式表示的一FOA聲音信號。其由四個通道組成，即該全向聲壓和三個一階空間梯度，它們在某個的假設下與粒子速度成比例。參見文獻[18]，該信號以參數方式被編碼。該參數從該複數聲壓 $P(k, n)$ 所導出，它是該已被變換的全向信號、以及對應於該已被變換的梯度信號的該複數粒子速度向量 $\mathbf{U}(k, n) = [U_X(k, n), U_Y(k, n), U_Z(k, n)]^T$ 。

【0108】 該DirAC表示係由在每個時間頻率區間處的該聲波的該信號 $P(k, n)$ 、該擴散 $\psi(k, n)$ 和方向 $\mathbf{r}(k, n)$ 所組成。為了得到後者，如文獻[18]所示，首先，該有效聲音強度向量 $\mathbf{I}_a(k, n)$ 被計算，其係為壓力向量與該速度向量的共軛複數(由 $(\cdot)^*$ 表示)的乘積的實部(由 $\text{Re}(\cdot)$ 表示)：

$$\mathbf{I}_a(k, n) = \frac{1}{2} \text{Re}(\mathbf{P}(k, n) \mathbf{U}^*(k, n))。 \quad (\text{公式1})$$

【0109】 如文獻[18]所示，從該向量的該變異係數，該擴散性被估計為：

$$\psi(k, n) = \sqrt{1 - \frac{\|E\{\mathbf{I}_a(k, n)\}\|}{E\{\|\mathbf{I}_a(k, n)\|\}}}, \quad (\text{公式2})$$

其中E表示沿時間幀的期望運算子，實現為移動平均。

【0110】 由於打算使用一方向基礎距離圖來操縱該聲音，該距離圖具有針對每個(匹配的)來源到該參考位置的一距離，因此該方向估計的該變異數在一可選的實施例中應該較低。由於幀通常很短，但情況並非總是如此。因此，一移動平均被應用以獲得一平滑的方向估計 $\bar{\mathbf{I}}_a(k, n)$ 。在一個實施例中，該信號的該直接部分的該DoA接著被計算，以作為在該相反方向的單位長度向量：

$$\mathbf{r}_r(k, n) = - \frac{\bar{\mathbf{I}}_a(k, n)}{\|\bar{\mathbf{I}}_a(k, n)\|}。 \quad (\text{公式3})$$

【0111】 由於該方向被編碼為針對每個時間頻率區間的單位長度的一個三維向量，因此可以直接對該距離資訊進行積分。該方向向量與其對應的地圖項目相乘，使得該向量長度表示該相應聲源 $\mathbf{d}_r(k, n)$ 的該距離：

$$\begin{aligned} \mathbf{d}_r(k, n) &= \mathbf{r}_r(k, n) d_r(k, n) \\ &= \mathbf{r}_r(k, n) m(\mathbf{r}_r(k, n)), \end{aligned} \quad (\text{公式4})$$

其中 $\mathbf{d}_r(k, n)$ 是從該麥克風的該記錄位置指向在時間 n 和頻率區間 k 活躍的該聲源的一向量。

【0112】 該聆聽者位置由當前處理幀的該追蹤系統給出為 $\mathbf{l}(n)$ 。參考圖6b，利用來源位置的該向量表示，可以減去該追蹤位置向量 $\mathbf{l}(n)$ 以生成具有長度 $d_1(k, n) = \|\mathbf{d}_1(k, n)\|$ 的該新的平移方向向量 $\mathbf{d}_1(k, n)$ 。從該聆聽者的PoV到該聲源的該距離係被導出，並且該DoA在一單一步驟中被調整：

$$\mathbf{d}_1(k, n) = \mathbf{d}_r(k, n) - \mathbf{l}(n)。 \quad (\text{公式5})$$

【0113】 真實再現的一個重要觀點是該距離衰減。於文獻[19]中，該衰減假設是聲源和聆聽者之間該距離的一函數。該方向向量的該長度用於編碼該衰減或放大，以用於再現。到該記錄位置的該距離係根據該距離圖以 $\mathbf{d}_r(k, n)$ 編碼，並且要再現的該距離以 $\mathbf{d}_1(k, n)$ 編碼。如果將該向量正規化為單位長度，接著乘以舊的及新的距離的該比率，則可以看出該所需長度可經由將 $\mathbf{d}_1(k, n)$ 除以該原始向量的該長度而被給出：

$$\mathbf{d}_v(k, n) = \frac{\mathbf{d}_1(k, n)}{\|\mathbf{d}_1(k, n)\|} \frac{\|\mathbf{d}_1(k, n)\|}{\|\mathbf{d}_r(k, n)\|} = \frac{\mathbf{d}_1(k, n)}{\|\mathbf{d}_r(k, n)\|}。 \quad (\text{公式6})$$

【0114】 該聆聽者的方向的該更改將被應用於以下步驟。該追蹤所

給出的該方向可以被寫為由該俯仰、偏轉和滾動組成的向量 $\mathbf{o}(n) = [o_X(n), o_Z(n), o_Y(n)]^T$ ，其係相對於作為該原點的該記錄位置。該來源方向根據該聆聽者方向而被旋轉，這是使用2D旋轉矩陣實現的：

$$\mathbf{d}_p(k, n) = \mathbf{R}_Y(o_Y(n))\mathbf{R}_Z(o_Z(n))\mathbf{R}_X(o_X(n))\mathbf{d}_v(k, n)。$$
 (公式7)

【0115】 該聆聽者的該結果DoA接著由被正規化為單位長度的該向量所給出：

$$\mathbf{r}_p(k, n) = \frac{\mathbf{d}_p(k, n)}{\|\mathbf{d}_p(k, n)\|}。$$
 (公式8)

【0116】 該經變換的方向向量、該擴散和該複數頻譜係用於合成一均勻分佈的8+4虛擬揚聲器設置的信號。八個虛擬揚聲器位於該聆聽者平面(高度0°)上的45°方位角步階，並且四個位於以上45°仰角的90°交叉陣列中。於文獻[16]中，對於每個揚聲器通道i，該合成被分成一直接和擴散部分，其中 $1 \leq i \leq I$ 、 $I=12$ 是揚聲器的該數量：

$$Y_i(k, n) = Y_{i,S}(k, n) + Y_{i,D}(k, n)。$$
 (公式9)

【0117】 於文獻[20]中，對於該直接部分，邊緣衰落幅度平移(EFAP、Edge Fading Amplitude Panning)係在給定該虛擬揚聲器幾何結構的情況下被應用於從該正確方向再現該聲音。給定DoA向量 $\mathbf{r}_p(k, n)$ ，這為每個虛擬揚聲器通道i提供一平移增益 $G_i(\mathbf{r})$ 。每個DoA的該距離相關增益是從該方向向量 $\mathbf{d}_p(k, n)$ 的該結果長度所導出的。通道i的該直接合成變為：

$$Y_{i,S}(k, n) = \sqrt{1 - \psi(k, n)}P(k, n)$$

$$G_i(\mathbf{r}_p(k, n))(\|\mathbf{d}_p(k, n)\|)^{-\gamma}，$$
 (公式10)

其中指數 γ 是一個調整因子，於文獻[19]中其通常設置為約1。應注意

的是，當 $\gamma = 0$ 時，該距離相關增益被關閉。

【0118】 該壓力 $P(k, n)$ 用於生成 I 去相關信號 $\tilde{P}_i(k, n)$ 。這些去相關的信號作為擴散分量被添加到該個別揚聲器通道。這遵循文獻[16]中的該標準方法：

$$Y_{i,D}(k, n) = \sqrt{\psi(k, n)} \frac{1}{\sqrt{I}} \tilde{P}_i(k, n)。$$
 (公式11)

【0119】 每個通道的該擴散和直接部分被加在一起，並且經由一逆 STFT 將該信號轉換回時域。取決於該揚聲器位置，這些通道時域信號針對左耳和右耳而與 HRTF 進行卷積運算，以產生雙耳信號。

【0120】 圖 6a 係顯示用於使用該空間資訊、和該第一和第二聲場描述以及該平移資訊以計算該修改的聲場的一另一較佳實施例，該平移資訊指示一參考位置到一不同參考位置的一平移，例如，如關於圖 4c 或圖 5 中所討論的向量 l 。

【0121】 圖 6a 係顯示指示一聲音分離的一應用的區塊 700，或者通常，其係與圖 4c 的參考位置 A 有關的該第一聲場描述和與圖 4c 的參考位置 B 有關的該第二聲場描述的聲音分析程序。

【0122】 該程序將導致具有一個或多個擷取對象的一個第一群組，並且另外，具有一個或多個擷取對象的一個第二群組。

【0123】 在區塊 702 內使用這些群組來計算所有分離的來源的該到達方向資訊，亦即，對於擷取來源的該第一群組和一個或多個擷取來源的該第二群組。

【0124】 在其他實施例中，步驟 700 和 702 在一單個程序內實現，一方面提供該來源的該信號，另一方面提供該來源的該 DoA 資訊。對於參數程序也是如此，例如 DirAC 等時間/頻率選擇程序，其中該來源信號是

在一時間/頻率區間中的該B格式表示的該信號、或該時間/頻率區間的該壓力信號或全向信號、並且該DoA資訊是該特定區間的該DoA參數。

【0125】 接著，在步驟704中，在該第一群組的該來源和該第二群組的該來源之間執行一來源匹配，並且該來源匹配的結果是匹配的來源。

【0126】 這些匹配的來源係被用於使用該新的DoA和該新的距離俾為每個匹配的對象計算一聲場，如區塊710所示。此外，該匹配對象的該到達方向資訊係在區塊706中被使用，其亦即針對來源A每個對象有兩個，例如圖4c的 α 和 β ，以便計算該匹配對象的該位置，或者可選地或另外地，使用例如三角測量操作的該匹配對象的該距離。

【0127】 區塊706的結果是每個匹配對象的該位置，或者可選地或另外地，一匹配對象與該第一或該第二參考位置A、B中的一個的該距離，例如在圖4c中所示的那樣。

【0128】 另外，較佳的是不僅使用區塊708中的平移資訊而且還使用旋轉資訊，來計算該匹配對象的該新的到達方向資訊和該新距離。

【0129】 雖然已經概述了該匹配對象的該位置被輸入到步驟708，但是要強調的是，其僅用於計算一匹配對象的該新到達方向資訊、該匹配對象的該實際位置，或者換句話說，該匹配對象的該距離對於計算相對於一聆聽者已移動到的一新(不同)參考位置的該新到達方向不是必需的，例如該距離不是必需的。

【0130】 然而，為了使該來源信號適應該新的情況，該距離是必要的。因此，當該來源或聲音對象到該新參考位置的該距離變得更短時，將算出小於1的一縮放因子。然而，當該距離變得更高時，將算出高於1的一比例因子，例如，如關於圖6b所討論的。因此，儘管在圖6a中針對

一實施例圖式說明，但不一定是該匹配對象的顯式位置、並且接著計算該匹配對象的該距離、並且接著針對每個匹配對象使用該新的到達方向和該新的距離計算該聲場的情況。相反，僅一匹配對象到兩個參考位置中的一個參考位置的該距離通常是足夠的，接著，使用該新的DoA和該新距離以計算每個匹配對象的一聲場。

【0131】 另外，區塊714係顯示使用由區塊702獲得的舊DoA資訊，計算該非匹配對象的聲場。此外，在區塊712中組合由區塊710中獲得的該匹配對象的該聲場以及由區塊714獲得的該非匹配對象的該聲場，以便獲得該修改的聲場描述，該修改的聲場描述例如可以是諸如一個一階高保真度環繞聲描述、一個更高階的高保真度環繞聲描述的一高保真度環繞聲描述，或者，與某個揚聲器設置相關的一揚聲器聲道描述，當然，其對於區塊710和區塊714是相同的，使得可以在區塊712中執行簡單的逐聲道添加。

【0132】 圖6b係顯示該聲場計算器420的一較佳實施方式。在區塊1102中，對於每個來源的一來源分離和一到達方向或一般方向資訊的計算係被執行。接著，在區塊1104中，該到達方向向量被乘以該距離資訊向量，亦即，從該原始參考位置到該聲源的該向量，亦即例如，從圖5的項目520到項目510的該向量。接著，在區塊1106中，該平移資訊，亦即，從圖5的項目520到項目500的該向量被考量，以便計算該新的平移方向向量，該新的平移方向向量是從該聆聽者位置500到該聲源的位置510的該向量。接著，由 $\mathbf{d}_v$ 所指示的具有該正確長度的該新到達方向向量的在區塊1108中被計算。該向量係指向與 $\mathbf{d}_r$ 該相同的方向，但具有一不同的長度，因為該向量的該長度反映了該聲源510被記錄在具有一某個的聲量的該原始聲場中的該事實，因此， $\mathbf{d}_v$ 的該長度或多或少表示該

響度變化。這是藉由將向量 $\mathbf{d}_l$ 除以該記錄距離 $\mathbf{d}_r$ 來獲得的，亦即從該麥克風520到該聲源510的向量 $\mathbf{d}_r$ 的該長度來獲得的。如上所述，經由三角測量計算，從該麥克風520到該聲源510的該向量 $\mathbf{d}_r$ 的該長度可以被導出。當該麥克風處於該第一聲場描述的該參考位置時，則使用從該第一聲場描述的該參考位置到該聲源的該距離。然而，當該麥克風處於該第二聲場描述的該另一參考位置時，則使用從該第二聲場描述的該另一參考位置到該聲源的該距離。

【0133】 當如圖5所示，該再現距離大於該記錄距離時，則 $\mathbf{d}_v$ 的該長度將低於1。這將導致該聲源510的一衰減，其針對在該新的聆聽者位置處的該再現。然而，當該再現距離 $\mathbf{d}_l$ 小於該記錄距離時，由區塊1108計算的 $\mathbf{d}_v$ 的該長度將大於1，並且一相應的縮放因子將導致該聲源的一放大。

【0134】 在圖6a中，項目710指示使用該新的到達方向資訊和該新距離來計算每個匹配對象的該聲場。然而，基本上，為了計算每個匹配對象的該聲場，從一個或多個擷取來源的該第一群組或一個或多個擷取來源的該第二群組獲得的對象信號通常可以被使用。然而，在一個實施例中，圖7中所示的一特定選擇被執行，以便決定在區塊710中使用哪個聲場描述來執行該聲場計算。在區塊720中，從該新聆聽者位置到該第一聲場描述的該第一參考位置的該第一距離被決定。關於圖4c，這是在該區別參考位置和參考位置A之間的該距離。

【0135】 此外，在步驟722中，從該新聆聽者位置到該第二聲場描述的該第二參考位置的該第二距離被決定。在圖4c的該實施例中，這是在該不同參考位置(由於平移)和參考位置B之間的該距離。

【0136】 看起來，從該不同參考位置到參考位置B的該距離低於從

該不同參考位置到該參考位置A的差。因此，這將在步驟724中被決定。並且，在步驟726中，從該具有較小距離的該聲場描述所導出的該群組中選擇該對象信號。因此，為了渲染對應於圖4c中所顯示的匹配來源之來源A和B，將使用從該第二聲場描述以導出該聲源信號，該第二聲場描述係與該另一參考位置B相關的。

【0137】 然而，在其他實施例中，在該平移指向從該原點到一不同參考位置的情況下，例如，在圖4c圖式中從該原點到圖4c的該左邊，該較小距離將從該另外的參考位置到該參考位置A，接著，該第一聲場描述將被使用，以在圖6b的區塊710中最終計算的每個匹配對象的該聲場。再次，該選擇將經由圖7中所示的該程序來執行。

【0138】 圖9係顯示一另一個較佳實施例。在步驟740中，關於該第一聲場描述的一聲場分析被執行，例如，以圖6c的區塊422中所示的一DirAC分析的形式的一參數聲場分析。

【0139】 這產生一第一組參數，例如對於每個時間/頻率區間，其中每組參數包括一DoA參數，並且包括一可選的擴散參數。

【0140】 在步驟741中，對該第二聲場描述執行一聲場分析，並且再次，如在區塊740中那樣執行一DirAC分析，並且例如，如關於圖6c的區塊422所討論的那樣。

【0141】 這產生一第二組參數，例如針對時間/頻率區間的參數。

【0142】 接著，在區塊746中，使用來自第一時間/頻率區間的該對應DoA參數和來自該第二組參數的該相同時間/頻率區間的該DoA參數，可以決定每個參數對的一位置。這將產生每個參數對的一位置。然而，針對相應時間/頻率區間，該擴散在該第一組參數和/或該第二組參數中越低，該位置將越有用。

【0143】 因此，其較佳的是僅進一步使用來自該時間/頻率區間的該位置，其在該第一組參數和第該二組參數兩者中都產生一非常低的擴散。

【0144】 另外，其較佳地還對該時間/頻率區間中的該對應信號執行一相關(correlation)，也由區塊740和區塊741輸出。

【0145】 因此，例如，在圖6a中的步驟704的該「來源匹配」可以被完全避免，並且可以藉由基於該擴散參數之匹配來源/匹配時間/頻率區間的一決定來代替，或者例如，使用來自該B格式分量或者使用圖6c的區塊422輸出的該壓力信號或目標信號的該時間/頻率區間中的相應信號，該匹配可以被另外執行。

【0146】 在任何情況下，區塊46將產生某些(選定的)時間/頻率區間的某些位置，其對應於在圖6a的區塊704中所找到的「匹配對象」。

【0147】 接著，在區塊748中，計算由區塊746所獲得的該位置和/或例如由一帽子跟踪器(hat tracker)所獲得的該相應平移/旋轉的修改參數和/或信號，並且區塊748的該輸出表示針對不同時間/頻率區間的修改參數和/或修改信號。

【0148】 因此，區塊748可以對應於區塊424的該平移變換423和旋轉變換，用於計算修改的參數之目的，並且，例如，修改信號的該計算將被圖6c的區塊425所執行，該執行係較佳地在考慮從該相應時間/頻率區間的該位置所導出的某個比例因子的情況下發生。

【0149】 最後，在區塊750中使用該修改的資料，該聲場描述的一合成被執行。例如，這可以經由使用該第一或該第二聲場描述的一DirAC合成來完成，或者可以經由如在區塊425所示的高保真度環繞聲信號生成器來執行，並且該結果將是用於傳輸/存儲/渲染的該新聲場描

述。

【0150】 圖10係顯示該聲場計算器420的一另一較佳實施方式。至少部分的圖10所示該程序係分別對每個匹配來源執行。該區塊1120決定一匹配來源的該距離，例如經由三角測量計算來決定。

【0151】 基於該聲場描述，一全波段到達方向或一每波段到達方向在1100中被決定。這些到達方向資訊係表示該聲場的該到達方向資料。基於該到達方向資料，一平移變換在區塊1110中被執行。至此，區塊1120針對每一匹配來源計算該距離。基於該資料，區塊1110生成該聲場的該新的到達方向資料，在這個實現中，其僅取決於從該參考位置到該不同參考位置的該平移。至此，區塊1110接收例如經由在一虛擬實境實現的上下文中的一追蹤而生成的該平移資訊。

【0152】 較佳地或替代地，也使用一旋轉資料。至此，區塊1130使用該旋轉資訊執行一旋轉變換。當平移和旋轉兩者被執行時，則在計算已經包括來自該平移的該資訊以及來自區塊1120的該來源距離的該聲場的該新DoA之後，其較佳地執行該旋轉變換。

【0153】 接著，在區塊1140中，該新的聲場描述被生成。至此，該原始聲場描述可以被使用，或者，經由一來源分離算法從該聲場描述分離的來源信號可以被使用，或者任何其他應用可以被使用。基本上，該新的聲場描述可以是例如由該高保真度環繞聲生成器430獲得的或由一DirAC合成器425生成的一方向聲場描述，或者可以是在該隨後的雙耳渲染中從一虛擬揚聲器表示所生成的一雙耳表示。

【0154】 較佳地，如圖10所示，每個到達方向的該距離也被用於生成該新的聲場描述，以使一某個的聲源的該音量或響度適應該新的位置，即該新的或不同的參考位置。

【0155】 儘管圖10係顯示一情況，其係在該平移變換之後執行該旋轉變換的，但是應注意，該順序可以是不同的。特別地，該旋轉變換可以被應用於由區塊1100所生成的該聲場的該DoA，並且然後，由於一對象從該參考位置到該不同參考位置的該平移引起的，該附加平移變換係被應用。

【0156】 一旦該聲場的該DoA被區塊1100決定了，該距離資訊就使用區塊1120而從該後設資料中被擷取，該距離資訊然後在區塊1140中被使用，以生成該新的聲場描述，用來計算一改變的距離，並且因此改變了該某個的來源相對於一某個的參考位置的響度。基本上，其可以說在該距離變大的情況下，該特定聲源信號被衰減，而當該距離變短時，則該聲源信號被放大。自然地，取決於該距離的該某個的聲源的該衰減或放大與該距離變化成比例地進行，但是，在其他實施例中，可以以非常粗略的增量將較不復雜的操作應用於該聲源信號的這種放大或衰減。與任何距離變化完全被忽略的一情況相比，即使是這一種不太複雜的實現也能提供更好的結果。

【0157】 圖11係顯示該聲場計算器的一另一個較佳實施例。在區塊1200中，來自該聲場的該個別來源係被決定，例如，每個頻帶或全頻帶。當每幀和頻帶的一決定被執行時，則這可以經由一DirAC分析來完成。如果一全頻帶或子頻帶的一決定被執行時，則可以經由任何類型的全頻帶或子頻帶來源分離演算法來完成。

【0158】 在區塊1210中，例如藉由頭部追蹤，一聆聽者的一平移和/或一旋轉係被決定。

【0159】 在區塊1220中，藉由使用該後設資料並且例如藉由使用該三角測量計算的該後設資料，每個來源的一舊距離係被決定。因此，

每個頻帶被認為是一某個的來源(假設該擴散係低於一某個的閾值)，並且接著，針對具有一低擴散值的每個時間/頻率區間的一某個的距離係被決定。

【0160】 接著，在區塊1230中，每個來源的一新距離例如經由每個頻帶的一向量計算而被獲得，例如，在圖6b的上下文中討論。

【0161】 此外，如區塊1240所示，例如，經由在一DirAC分析中被獲得的一DoA計算或例如經由一來源分離演算法中的一到達方向或方向資訊分析，每個來源的一舊方向被決定。

【0162】 接著，在區塊1250中，例如經由執行每個頻帶或全頻帶的一向量計算，每個來源的一新方向被決定。

【0163】 接著，在區塊1260，一新的聲場為該平移和旋轉的聆聽者而被生成。例如，這可以經由在該DirAC合成中縮放每個通道的該直接部分來完成。取決於該具體實施方式，除了在區塊1260中執行該距離修改，新增地或替代地，在區塊1270a、1270b或1270c中該距離修改可被完成。

【0164】 例如，當決定該聲場僅具有一單個來源時，則該距離修改可以在區塊1270a中已經被執行。

【0165】 或者，當個別來源信號被區塊1200計算時，該實際新聲場在區塊1260中被生成之前，針對各個來源，該距離修改可以在區塊1270b中被執行。

【0166】 另外，例如，當在區塊1260中的該聲場生成不渲染為一揚聲器設置信號或一雙耳信號，而是另一聲場描述時，例如，使用一高保真度環繞聲編碼器或計算器430，則該距離修改也可以是在區塊1260中該生成之後被執行，這意指在區塊1270c中。根據該實施方式，一距離

修改也可以被分配給多個修改器，以便最終一某個的聲源處於一某個的響度，該某個的響度由該原始距離與該新距離之間的該差異所引導，該原始距離係該聲源和該參考位置之間的距離、該新距離係該聲源和該不同參考位置之間的距離。

【0167】 圖12a係顯示最初公開的一DirAC分析器，例如，在2009年IWPASH的早先引用的參考文獻“Directional Audio Coding”中。

【0168】 該DirAC分析器包括一組帶通濾波器1310、一能量分析器1320、一強度分析器1330、一時間平均區塊1340和一擴散度計算器1350以及該方向計算器1360。

【0169】 在DirAC中，分析和合成兩者都在頻率域中進行。有幾種方法可以將該聲音分成頻帶，每種方法都在不同的屬性中。該最常用的頻率變換包括短時傅立葉變換(STFT、Short Time Fourier Transform)和正交鏡像濾波器組(QMF、Quadrature Mirror Filter)。除了這些之外，還可以完全自由地設計具有任意濾波器的一濾波器組，該濾波器組可針對任何特定目的進行優化。方向分析的該目標是在每個頻帶中估計聲音的該到達方向，以及聲音是否同時從一個或多個方向到達的一估計。原則上，這可以用許多技術來執行，然而，聲場的該能量分析已經被發現是合適的，這在圖12a中圖式說明。當從一單個位置捕獲一維、二維或三維中的該壓力信號和速度信號時，該能量分析可以被執行。在第一階B格式信號中，該全向信號被稱為W信號，其已經被該2的平方根縮小。該聲壓可以被估計為 $P = \sqrt{2} * W$ ，在該STFT域中被表示。

【0170】 該X、Y和Z通道具有沿該笛卡爾軸定向的一偶極子的該方向樣型，其一起形成一向量 $U=[X, Y, Z]$ 。該向量估計該聲場速度向量，並且也在STFT域中被表示。該聲場的該能量E被計算。可以經由定

向麥克風的重合定位或者一使用緊密間隔的全向麥克風組，B格式信號的該捕獲可被獲得。在一些應用中，該麥克風信號可以在一計算域中被形成，亦即模擬。

【0171】 該聲音方向被定義為該強度向量 I 的該相反方向。在該發送的后設資料中，該方向被表示為對應的角度方位角和仰角值。還使用該強度向量和該能量的一期望運算子來計算聲場的該擴散。該公式的該結果是在0和1之間的一實數值，其特徵在於該聲音能量是從一單個方向到達(擴散是零)還是從所有方向(擴散是1)。在該完整的3D或更小尺寸的速度資訊可用的情況下，該程序是適當的。

【0172】 圖12b係顯示一DirAC合成，再一次，其具有一組帶通濾波器1370、一虛擬麥克風區塊1400、一直接/擴散合成器塊1450、以及一某個的揚聲器設置或一虛擬預期揚聲器設置1460。另外，一擴散-增益變換器1380、一基於向量的振幅平移(VBAP、Vector Based Amplitude Panning)增益表區塊1390、一麥克風補償區塊1420、一揚聲器增益平均區塊1430和用於其他通道的一分配器1440係被使用。

【0173】 在利用揚聲器的這種DirAC合成中，圖12b中所示該高品質版本的DirAC合成接收所有B格式信號，為此針對該揚聲器設置1460的每個揚聲器方向，一虛擬麥克風信號被計算。該所使用的定向樣型通常是一偶極子。接著，取決於該後設資料，該虛擬麥克風信號以非線性方式被修改。DirAC的該低位元率版本未在圖12b中顯示，然而，在這種情況下，如圖6所示，僅一個音頻聲道被發送。處理的該不同之處在於所有虛擬麥克風信號都將被所接收的該單個音頻通道所取代。該虛擬麥克風信號被分成兩個串流：該擴散串流和該非擴散串流，它們被分開處理。

【0174】藉由使用向量基礎幅度平移(VBAP、Vector Base Amplitude Panning)，該非擴散聲音將被再現為點來源。在平移中，在與揚聲器特定的增益因子相乘之後，一單聲道聲音信號將被應用於揚聲器的一子集。使用一揚聲器設置的該資訊和特定的平移方向，該增益因子被計算。在低位元率版本中，該輸入信號簡單地平移到該後設資料所隱含的該方向。在該高品質版本中，每個虛擬麥克風信號與該相應的增益因子相乘，這會產生與平移相同的該效果，然而它不太容易出現任何非線性人工瑕疵。

【0175】在許多情況下，該方向後設資料受到突然的時間變化的影響。為了避免人工瑕疵，VBAP計算的揚聲器的該增益因子經由與頻率相關的時間常數的時間積分而被平滑，該時間常數等於每個頻帶處的約50個循環週期。這有效地消除了人工瑕疵，然而，在大多數情況下，方向的該變化不會被認為比沒有平均的更慢。

【0176】該擴散聲音的該合成的該目的是創造圍繞該聆聽者的聲音感知。在該低位元率版本中，經由去相關該輸入信號並將其從每個揚聲器再現，該擴散串流被再現。在該高品質版本中，擴散串流的該虛擬麥克風信號在某種程度上已經不連貫，並且它們需要僅稍微去相關。與該低位元率版本相比，這種方法為環繞聲混響和環境聲音提供了更好的空間品質。

【0177】對於使用耳機進行該DirAC合成，DirAC針對該非擴散串流採用圍繞該聆聽者的一某個數量的虛擬揚聲器規劃、以及針對該擴散串流採用一某個數量的揚聲器規劃。該虛擬揚聲器被實現為輸入信號與具有一測量的頭部相關轉移函數(HRTF、Head-Related Transfer Functions)的捲積運算。

【0178】 儘管已經在一裝置的上下文中描述了一些觀點，但是顯然這些觀點也表示該對應方法的一描述，其中一區塊或設備對應於一方法步驟或一方法步驟的一特徵。類似地，在一方法步驟的上下文中描述的觀點還表示一對應裝置的一對應區塊或項目或特徵的一描述。

【0179】 本揭露的增強聲場描述可以存儲在一數位存儲媒體或非暫時性存儲媒體上，或者可以在一傳輸媒體上被傳輸，諸如一無線傳輸媒體、或諸如網際網路的一有線傳輸媒體。

【0180】 根據某些實現要求，本揭露的實施例可以用硬體或軟體實現。該實現可以使用一數位存儲媒體來執行，例如一軟性磁碟、一DVD、一CD、一ROM、一PROM、一EPROM、一EEPROM或一FLASH記憶體，其上存儲有電子可讀取控制信號，它們配合(或者能夠與)一可編程計算機系統協作，以便執行該相應的方法。

【0181】 根據本揭露的一些實施例，其包括具有一電子可讀取控制信號的非暫時性資料載體，其能夠與一可編程計算機系統協作，從而執行本文所述的該方法之一。

【0182】 通常，本揭露的實施例可以實現為具有一程式編碼的計算機程式產品，該程式編碼可操作用於在該計算機程式產品在一計算機上運行時執行這些方法之一。該程式編碼可以例如存儲在一機器可讀載體上。

【0183】 其他實施例包括用於執行存儲在一機器可讀載體上的本文所述方法之一的該計算機程式。

【0184】 換句話說，本揭露方法的一實施例因此是具有一程式編碼的一計算機程式，當該計算機程式在一計算機上運行時，該程式編碼用於執行本文所述的該方法之一。

【0185】 因此，本揭露方法的一另一實施例是一資料載體(或一數位存儲媒體、或一計算機可讀媒體)，其包括記錄在其上的用於執行本文所述方法之一的該計算機程式。

【0186】 因此，本揭露方法的一另一實施例是表示用於執行本文所述方法之一的計算機程式的一資料串流或一信號序列。該資料串流或該信號序列可以例如被配置為經由一資料通信連接傳輸，例如經由該網際網路。

【0187】 一另一實施例包括一處理裝置，例如一計算機或一可編程邏輯設備，其被配置為或適於執行本文所述的該方法之一。

【0188】 另一實施例包括一計算機，其上安裝有用於執行本文所述方法之一的該計算機程式。

【0189】 在一些實施例中，一可編輯邏輯器件(例如現場可編輯閘陣列)可用於執行本文所述方法的一些或全部該功能。在一些實施例中，該現場可編輯閘陣列可以與一微處理器協作，以便執行本文描述的方法之一。通常，該方法較佳地由任何硬體設備執行。

【0190】 在借助附圖對本揭露的實施例進行詳細說明之前，應當注意的是，在不同的附圖中，相同的、功能上相同的和相等的元件、物件和/或結構被提供有相同的附圖標記，使得不同實施例中的這些元件的描述是可互換和/或相互適用的。

【0191】 儘管已經在一設備的上下文中描述了一些觀點，但是應當理解，所述觀點還表示對應方法的一描述，使得一設備的一區塊或一結構組件也應被理解為一對應的方法步驟、或作為一方法步驟的一個特徵。經由類推，已經結合一方法步驟或作為一方法步驟描述的觀點也表示一對應設備的一對應區塊或細節或特徵的一描述。

【0192】 上述實施例僅代表了本揭露該原理的一說明。應理解，本領域其他技術人員將理解本文所述的佈置和細節的任何修改和變化。上述實施例僅係為了方便說明而舉例而已，本揭露所主張之權利範圍自應以申請專利範圍該為準，而非僅限於上述實施例。

【0193】 參考文獻

- [1] Liitola, T., *Headphone sound externalization*, Ph.D. thesis, Helsinki University of Technology. Department of Electrical and Communications Engineering Laboratory of Acoustics and Audio Signal Processing., 2006.
- [2] Blauert, J., *Spatial Hearing - Revised Edition: The Psychophysics of Human Sound Localization*, The MIT Press, 1996, ISBN 0262024136.
- [3] Zhang, W., Samarasinghe, P. N., Chen, H., and Abhayapala, T. D., "Surround by Sound: A Re-view of Spatial Audio Recording and Reproduction," *Applied Sciences*, 7(5), p. 532, 2017.
- [4] Bates, E. and Boland, F., "Spatial Music, Virtual Reality, and 360 Media," in *Audio Eng. Soc. Int. Conf. on Audio for Virtual and Augmented Reality*, Los Angeles, CA, U.S.A., 2016.
- [5] Anderson, R., Gallup, D., Barron, J. T., Kontkanen, J., Snavely, N., Esteban, C. H., Agarwal, S., and Seitz, S. M., "Jump: Virtual Reality Video," *ACM Transactions on Graphics*, 35(6), p. 198, 2016.

- [6] Merimaa, J., *Analysis, Synthesis, and Perception of Spatial Sound: Binaural Localization Modeling and Multichannel Loudspeaker Reproduction*, Ph.D. thesis, Helsinki University of Technology, 2006.
- [7] Kronlachner, M. and Zotter, F., “Spatial Transformations for the Enhancement of Ambisonics Recordings,” in *2nd International Conference on Spatial Audio*, Erlangen, Germany, 2014.
- [8] Tsingos, N., Gallo, E., and Drettakis, G., “Perceptual Audio Rendering of Complex Virtual Environments,” *ACM Transactions on Graphics*, 23(3), pp. 249–258, 2004.
- [9] Taylor, M., Chandak, A., Mo, Q., Lauterbach, C., Schissler, C., and Manocha, D., “Guided multi-view ray tracing for fast auralization,” *IEEE Trans. Visualization & Comp. Graphics*, 18, pp. 1797– 1810, 2012.
- [10] Rungta, A., Schissler, C., Rewkowski, N., Mehra, R., and Manocha, D., “Diffraction Kernels for Interactive Sound Propagation in Dynamic Environments,” *IEEE Trans. Visualization & Comp. Graphics*, 24(4), pp. 1613–1622, 2018.
- [11] Thiergart, O., Kowalczyk, K., and Habets, E. A. P., “An Acoustical Zoom based on Informed Spatial Filtering,” in *Int. Workshop on Acoustic Signal Enhancement*, pp. 109–113, 2014.
- [12] Khaddour, H., Schimmel, J., and Rund, F., “A Novel Combined System of Direction Estimation and Sound Zooming of Multiple Speakers,” *Radioengineering*, 24(2), 2015.

- [13] Ziegler, M., Keinert, J., Holzer, N., Wolf, T., Jaschke, T., op het Veld, R., Zakeri, F. S., and Foessel, S., “Immersive Virtual Reality for Live-Action Video using Camera Arrays,” in *IBC*, Amsterdam, Netherlands, 2017.
- [14] Thiergart, O., Galdo, G. D., Taseska, M., and Habets, E. A. P., “Geometry-Based Spatial Sound Acquisition using Distributed Microphone Arrays,” *IEEE Trans. Audio, Speech, Language Process.*, 21(12), pp. 2583–2594, 2013.
- [15] Kowalczyk, K., Thiergart, O., Taseska, M., Del Galdo, G., Pulkki, V., and Habets, E. A. P., “Parametric Spatial Sound Processing: A Flexible and Efficient Solution to Sound Scene Acquisition, Modification, and Reproduction,” *IEEE Signal Process. Mag.*, 32(2), pp. 31–42, 2015.
- [16] Pulkki, V., “Spatial Sound Reproduction with Directional Audio Coding,” *J. Audio Eng. Soc.*, 55(6), pp. 503–516, 2007.
- [17] International Telecommunication Union, “ITU-R BS.1534-3, Method for the subjective assessment of intermediate quality level of audio systems,” 2015.
- [18] Thiergart, O., Del Galdo, G., Kuech, F., and Prus, M., “Three-Dimensional Sound Field Analysis with Directional Audio Coding Based on Signal Adaptive Parameter Estimators,” in *Audio Eng. Soc. Conv. Spatial Audio: Sense the Sound of Space*, 2010.
- [19] Kuttruff, H., *Room Acoustics*, Taylor & Francis, 4 edition, 2000.
- [20] Borß, C., “A polygon-based panning method for 3D loudspeaker setups,” in *Audio Eng. Soc. Conv.*, pp. 343–352, Los Angeles, CA, USA, 2014.

- [21] Rummukainen, O., Schlecht, S., Plinge, A., and Habets, E. A. P., "Evaluating Binaural Reproduction Systems from Behavioral Patterns in a Virtual Reality – A Case Study with Impaired Binaural Cues and Tracking Latency," in *Audio Eng. Soc. Conv. 143*, New York, NY, USA, 2017.
- [22] Engelke, U., Darcy, D. P., Mulliken, G. H., Bosse, S., Martini, M. G., Arndt, S., Antons, J.-N., Chan, K. Y., Ramzan, N., and Brunnström, K., "Psychophysiology-Based QoE Assessment: A Survey," *IEEE Selected Topics in Signal Processing*, 11(1), pp. 6–21, 2017.
- [23] Schlecht, S. J. and Habets, E. A. P., "Sign-Agnostic Matrix Design for Spatial Artificial Reverberation with Feedback Delay Networks," in *Proc. Audio Eng. Soc. Conf.*, pp. 1–10– accepted, Tokyo, Japan, 2018
- [31] M. A. Gerzon, "Periphony: With-height sound reproduction," *J. Acoust. Soc. Am.*, vol. 21,110. 1, pp. 2-10, 1973.
- [32] V. Pulkki, "Directional audio coding in spatial sound reproduction and stereo upmixing," in *Proc. of the 28th AES International Conference*, 2006.
- [33] —, "Spatial sound reproduction with directional audio coding," *Journal Audio Eng. Soc.*, vol. 55, no. 6, pp. 503-516, Jun. 2007.
- [34] C. G. and G. M., "Coincident microphone simulation covering three dimensional space and yielding various directional outputs," U.S. Patent 4 042 779, 1977.

- [35] C. Faller and F. Baumgarte, "Binaural cue coding - part ii: Schemes and applications," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, Nov. 2003.
- [36] C. Faller, "Parametric multichannel audio coding: Synthesis of coherence cues," *IEEE Trans. Speech Audio Process.*, vol. 14, no. 1, Jan. 2006 .
- [37] H. P. J. E. E. Schuijers, J. Breebaart, "Low complexity parametric stereo coding," in *Proc. of the 116th AES Convention*, Berlin, Germany, 2004.
- [38] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *J. Acoust. Soc. Am.*, vol. 45, no. 6, pp. 456-466, Jun. 1997.
- [39] J. G. Tylka and E. Y. Choueiri, "Comparison of techniques for binaural navigation of higher- order ambisonics sound fields," in *Proc. of the AES International Conference on Audio for Virtual and Augmented Reality*, New York, Sep. 2016.

【符號說明】

【0194】

100：聲場描述生成器

110：後設資料生成器

120：組合器

- 120：輸出界面
- 250、260：聲場生成器
- 330：第一聲場描述
- 340：第二聲場描述
- 350：後設資料
- 350：幾何資訊
- 351：參考點
- 420：聲場計算器
- 400：輸入界面
- 100、210：區塊
- 350、351、352至354：項目
- 410：平移界面
- 402：後設資料提供器
- 404：聲場提供器
- 430：高保真度環繞聲信號生成器
- 422、422a、422b：DirAC分析器
- 425、425a、425b：DirAC合成器
- 352、353、354：項目
- 422：DirAC編碼器
- 423、424：新穎變換
- 425：DirAC解碼器
- 427：雙耳化
- 700、702、710、706、708、714、712：區塊
- 700、702、704、708、722、724、726、741：步驟

1102、1104、1106、1108：區塊

520、510、710：項目

510：聲源的位置

510：聲源

520：麥克風

422、746、741、46、704、748、424、425、750：區塊

423：平移變換

1120、1110、1130、1140、1100：區塊

425：DirAC合成器

1200、1210、1220、1230、1240、1250、1260：區塊

1270a、1270b、1270c：區塊

430：高保真度環繞聲編碼器或計算器

1310：一組帶通濾波器

1320：能量分析器

1330：強度分析器

1340：時間平均區塊

1350：擴散度計算器

1360：方向計算器

1370：一組帶通濾波器

1400：虛擬麥克風區塊

1450：直接/擴散合成器塊

1460：揚聲器設置或虛擬預期揚聲器設置

1380：擴散-增益變換器

1390：基於向量的振幅平移增益表區塊

- 1420：麥克風補償區塊
- 1430：揚聲器增益平均區塊
- 1440：分配器
- 1460：揚聲器設置

【發明申請專利範圍】

【第1項】 一種用於生成一增強聲場描述的裝置，該裝置包括：

一聲場生成器，用於生成一第一聲場描述以及一第二聲場描述，該第一聲場描述指示相對於一第一參考位置的一聲場，該第二聲場描述指示相對於一第二參考位置的該聲場，該第二參考位置與該第一參考位置不相同；以及

一後設資料生成器，用於生成與該聲場的空間資訊有關的後設資料；

其中，該後設資料生成器被配置為決定該第一聲場描述之一第一幾何描述與該第二聲場描述之一第二幾何描述其中之至少一，以作為該後設資料；以及

其中，該第一聲場描述、該第二聲場描述和包含該第一幾何描述與該第二幾何描述其中之至少一的該後設資料構成該增強聲場描述。

【第2項】 如申請專利範圍第1項之該裝置，

其中，該後設資料生成器被配置為決定該第一參考位置與該第二參考位置其中之一為該第一幾何描述和決定該第一參考位置與該第二參考位置其中之一為該第二幾何描述、或者決定該第一參考位置與該第二參考位置之間的一距離或該第一參考位置與該第二參考位置之間的一位置向量，以作為該後設資料。

【第3項】 如申請專利範圍第1項之該裝置，其中，該第一聲場描述是一第一高保真度環繞聲(Ambisonics)描述，並且其中該第二聲場描述是一第二高保真度環繞聲描述，或者其中該第一聲場描述和該第二聲場描述是一高保真度環繞聲描述或一定向音頻編碼(DirAC、Directional Audio Coding)描述之一。

【第4項】 如申請專利範圍第1項之該裝置，其中，該第一幾何描述是關於從一預定原點指向該第一聲場描述的該第一參考位置的一第一向量的一資訊；以及

其中，該第二幾何描述是關於從該預定原點指向作為該第二聲場描述的該第二參考位置的一第二向量的一資訊。

【第5項】 如申請專利範圍第1項之該裝置，其中，該第一幾何描述是關於從一預定原點指向該第一聲場描述的該第一參考位置的一第一向量和從該預定原點指向該第二聲場描述的該第二參考位置的一第二向量之一的資訊，以及該第二幾何描述包含關於該第一參考位置和該第二參考位置之間的該向量的一資訊。

【第6項】 如申請專利範圍第1項之該裝置，其中，該第一參考位置和該第二的參考位置之一是表示該第一幾何描述的一預定原點，並且其中該第二幾何描述包括關於該預定原點與該第一參考位置和該第二參考位置其中之另一之間的一向量的一資訊。

【第7項】 如申請專利範圍第1項之該裝置，其中，該聲場生成器被配置為使用一真實麥克風設備或經由使用一虛擬麥克風技術的一聲音合成來生成該第一聲場描述或該第二聲場描述。

【第8項】 一種用於根據一聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述的裝置，包括：

一聲場計算器，用於使用該空間資訊、該聲場描述和指示一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改的聲場，

其中，該聲場計算器被配置為：

接收與該參考位置有關的一第一聲場描述並接收與不同於該參考位置的一另一參考位置有關的一第二聲場描述；

執行對該第一聲場描述的一來源分離以獲得一個或多個擷取對象的一第一群組和對該第二聲場描述的一來源分離以獲得一個或多個擷取對象的一第二群組；

計算該擷取對象的到達方向(DoA)資料；

匹配該第一聲場描述的擷取對象的該第一群組和該第二聲場描述的擷取對象的該第二群組，以便得到一個或多個匹配對象；

估計該一個或多個匹配對象其中之一匹配對象的一位置；以及

基於該一個或多個匹配對象其中之一的該匹配對象的該估計位置和該不同參考位置，來修改該一個或多個匹配對象其中之一的該匹配對象。

【第9項】 如申請專利範圍第8項之該裝置，其中，該聲場計算器被配置為：

計算該一個或多個匹配對象的到達方向資料；以及

使用針對該不同參考位置的該計算的到達方向資料，來決定該一個或多個匹配對象中的每個匹配對象的聲場描述資料。

【第10項】 如申請專利範圍第8項之該裝置，其中，該聲場計算器被配置為計算一個或多個非匹配對象的到達方向資料，以及使用針對該一個或多個非匹配對象所獲得的該到達方向資料，來計算一個或多個該非匹配對象的一聲場資料。

【第11項】 如申請專利範圍第10項之該裝置，其中，該聲場計算器被配置為經由組合該一個或多個匹配對象的該聲場描述資料和該一個或多個非匹配對象的該聲場描述資料，來計算該修改的聲場描述。

【第12項】 一種用於根據一聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述的裝置，包括：

一聲場計算器，用於使用該空間資訊、該聲場描述和指示一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改的聲場，

其中，該聲場計算器被配置為：

假設一個聲場模型；

估計假設之該聲場模型的信號分量和/或參數；

基於該平移資訊或基於該不同的參考位置，來修改該信號分量和/或該參數；以及

使用該修改的信號分量和/或該修改的參數，來生成該修改的聲場描述。

【第13項】 如申請專利範圍第12項之該裝置，其中，該聲場計算器被配置為使用假設之該聲場模型的一分解將其分解為一個或多個直接聲音分量和一個或多個擴散聲音分量、或經由使用假設之該聲場模型的一分解將其分解為一個或多個直接/主要聲音分量和一殘餘聲音分量，來估計該信號分量和/或該參數，其中該殘餘聲音分量可以是一單聲道信號或一多聲道信號。

【第14項】 如申請專利範圍第8或12項之該裝置，其中，該聲場計算器被配置為應用該修改的聲場描述的一旋轉。

【第15項】 如申請專利範圍第8或12項之該裝置，其中，該修改的聲場描述是一高保真度環繞聲聲場描述。

【第16項】 一種用於根據一聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述的裝置，包括：

一聲場計算器，用於使用該空間資訊、該聲場描述和指示一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改的聲場，

其中，該聲場計算器被配置為：

接收作為該聲場描述的一第一聲場描述和一第二聲場描述；

在該第一聲場描述和該第二聲場描述上執行一來源分離，以擷取該第一聲場描述和該第二聲場描述的來源以及擷取的該等來源的到達方向(DoA)資料；

使用該DoA資料和該平移資訊，針對擷取的該等來源的各該來源，計算關於不同之該參考位置的修改的DoA資料；以及

處理擷取的該等來源和該修改的DoA資料，以獲得該修改的聲場描述。

【第17項】 如申請專利範圍第16項之該裝置，其中，該聲場計算器被配置為對每個聲場描述個別地執行該來源分離並為每個聲場描述擷取一環境信號、一漫射信號、或一殘留信號。

【第18項】 如申請專利範圍第8、12或16項之該裝置，其更包含：一平移界面，用於提供該平移資訊或一旋轉資訊，以指示一預定聆聽者對該修改的聲場描述的一旋轉；

一後設資料提供器，用於將該後設資料提供給該聲場計算器；

一聲場提供器，用於將該聲場描述提供給該聲場計算器；以及

一輸出界面，用於輸出包括該修改聲場描述和該修改後設資料的一修改聲場，該修改後設資料是使用該平移資訊從該後設資料導出的，或者用於輸出多個揚聲器聲道，每個揚聲器聲道與一預定義的揚聲器位置相關，或者用於輸出該修改聲場的一雙耳表示。

【第19項】 一種用於根據一聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述的裝置，包括：

一聲場計算器，用於使用該空間資訊、該聲場描述和指示一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改的聲場，

其中，該聲場計算器包括：

一聲場分析器，用於分析該聲場分量，以針對不同的頻率區間導出到達方向資訊；

一平移變換器，用於使用該到達方向資訊和後設資料，以計算每個頻率區間的修改的到達方向資訊，該後設資料包括將一距離資訊與由該頻率區間表示的一來源相關聯的一深度圖；以及

一距離補償器，用於使用一距離補償資訊計算該修改聲場描述，該距離補償資訊取決於該來源的該深度圖提供的該距離、以及與該頻率區間相關聯的一新距離，該新距離與該修改到達方向資訊相關。

【第20項】一種用於根據一聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述的裝置，包括：

一聲場計算器，用於使用該空間資訊、該聲場描述和指示一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改的聲場，

其中，該聲場計算器針對一個或多個來源被配置為：

用於計算從該參考位置指向一聲源的一第一向量，其係由該聲場描述所描述之一聲場的一分析所獲得；

用於計算一第二向量，該第二向量使用該第一向量和該平移資訊從該不同參考位置指向該聲源，該平移資訊係定義為從該參考位置到該不同參考位置的一平移向量；以及

用於計算一距離修改值，其使用該不同參考位置、該聲源的一位置和該第二向量，或是使用從該不同參考位置到該聲源的該位置的一距離和該第二向量。

【第21項】 如申請專利範圍第8、12、16、19及20項之該裝置，其中，經由將一到達方向的單位向量乘以該後設資料中所包括的一距離來計算一第一向量；或

其中，經由從該第一向量中減去一平移向量來計算一第二向量，其中該平移資訊定義從該參考位置到該不同參考位置的該平移向量；或者

其中，經由將該第二向量除以該第一向量的一範數來計算一距離修改值。

【第22項】 如申請專利範圍第8、12、16、19及20項之該裝置，其中，該聲場計算器被配置為除了該平移資訊之外還接收一旋轉資訊；以及

其中該聲場計算器被配置為使用該旋轉資訊執行一旋轉變換，以旋轉由該聲場描述所描述之該聲場的一到達方向資料，其中該到達方向資料係從經由該聲場描述的一聲音分析所獲得的一到達方向資料以及使用該平移資訊導出的。

【第23項】 一種用於根據一聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述的裝置，包括：

一聲場計算器，用於使用該空間資訊、該聲場描述和指示一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改的聲場，

其中，該聲場計算器被配置為：

經由一聲場分析，從該聲場描述和該來源的方向決定來源；

針對一個來源，使用該後設資料，以決定該來源與該參考位置的一距離；

決定與該不同參考位置相關的該來源的一新方向，其使用該來源的該方向和該平移資訊；

決定與該不同參考位置相關的該來源的一新距離資訊；以及
生成該修改聲場描述，其係使用該來源的該新方向、該新距離資訊、以及該聲場描述或對應於從該聲場描述所導出的該來源的來源信號。

【第24項】 一種用於根據一聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述的裝置，包括：

一聲場計算器，用於使用該空間資訊、該聲場描述和指示一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改的聲場，

其中，該聲場計算器被配置為：

經由一聲音分析，以決定來自該聲場描述的來源信號和與該參考位置相關的該等來源信號的方向；

使用該平移資訊，以計算與該不同參考位置相關的該等來源信號的新方向；

計算與該不同參考位置相關的該等來源信號的距離資訊；以及

使用該距離資訊、該等來源信號和該等新方向，以合成該修改聲場描述。

【第25項】 如申請專利範圍第24項之該裝置，其中，該聲場計算器配置為：

經由將該等來源信號其中之一來源信號平移到與一重放設置相關的該新方向所給出的一方向，來合成該修改聲場描述；以及

經由在執行該平移之前或者在執行該平移之後，使用該距離資訊，俾縮放該來源信號。

【第26項】 如申請專利範圍第24項之該裝置，其中，該聲場計算器被配置為將一擴散信號添加到該來源信號的一直接部分，該直接部分在被添加到該擴散信號之前被該距離資訊修改。

【第27項】 如申請專利範圍第24項之該裝置，其中，該聲場計算器被配置為：

執行該聲場描述的一時頻轉換以及計算一時間幀的多個頻率區間的一到達方向；

計算各該頻率區間的該新方向；

計算各該頻率區間的該距離資訊；以及

使用針對一頻率區間的一音頻信號、從該頻率區間的該新方向資訊導出的針對該頻率區間的一平移增益、和從該對應來源信號的該距離資訊所導出的針對該頻率區間的一縮放因子，以執行對每個頻率區間的一直接合成，以便獲得一直接部分。

【第28項】 如申請專利範圍第27項之該裝置，其中，該聲場計算器配置為：

使用從針對該頻率區間的該音頻信號導出的一擴散音頻信號、並使用針對該頻率區間的該聲音分析導出的一擴散參數來執行一擴散合成，以便獲得一擴散部分，並且組合該直接部分和該擴散部分以獲得該頻率區間的一合成音頻信號；以及

使用針對一時間幀的該頻率區間的該合成音頻信號，執行一頻率-時間轉換，以獲得一時域合成音頻信號，俾作為該修改聲場描述。

【第29項】 如申請專利範圍第24項之該裝置，其中，該聲場計算器被配置為針對各該來源信號合成與該不同參考位置相關的一聲場，該合成包括：

對於各該來源信號，使用針對該來源信號的該新方向以處理該來源信號，俾獲得與該不同參考位置相關的該來源信號的的一聲場描述；

在處理該來源信號之前修改該來源信號、或使用該方向資訊修改與該不同參考位置相關的該來源信號的該聲場描述；以及

添加該等來源信號至該聲場描述以獲得與該不同參考位置相關的該修改聲場描述。

【第30項】 如申請專利範圍第24項之該裝置，其中，該聲音分析被配置為經由一來源分離演算法決定該來源信號，並從該聲場描述中減去至少一些該來源信號，以獲得該擴散信號。

【第31項】 一種用於根據一聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述的裝置，包括：

一聲場計算器，用於使用該空間資訊、該聲場描述和指示一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改的聲場，

其中，該聲場計算器被配置為：

匹配對該第一聲場描述的來源的一第一群組和該第二聲場描述的來源的一第二群組，以獲得一個或多個匹配來源；

針對該一個或多個匹配來源的各該匹配來源決定一到達方向資訊；

使用該到達方向資訊和該聲場描述的該後設資料，來決定一來源到該不同參考位置的一距離；以及

使用該來源到該不同參考位置的該距離，來決定一縮放因子。

【第32項】 一種用於根據一聲場描述和與該聲場描述的空間資訊有關的後設資料以生成一修改的聲場描述的裝置，包括：

一聲場計算器，用於使用該空間資訊、該聲場描述和指示一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改的聲場，

其中，該聲場計算器被配置為：

對該聲場描述執行一聲音分析，以獲得一個或多個擷取對象的一第一群組或一個或多個擷取對象的一到達方向資訊以及與一另一個參考位置有關的一另一個聲場描述，使用該後設資料獲得一個或多個擷取對象的一第二群組或一個或多個擷取對象的一到達方向資訊；

使用來自該群組的對象信號或來自該群組的對象的到達方向資訊，以匹配該第一群組和該第二群組，俾找到至少一個匹配對象；

計算該匹配對象的一估計位置，其使用該後設資料、和經由執行該聲音分析獲得該第一群組該匹配對象之所獲得的一第一到達方向資訊、以及經由執行該聲音分析獲得該第二群組該匹配對象之所獲得的一第二到達方向資訊；以及

基於該估計的位置和該平移資訊，對該匹配的對象應用一距離補償處理。

【第33項】 如申請專利範圍第32項之該裝置，其中，該聲音分析是一DirAC分析，產生每個時間頻頻區間的一信號、和每個時間頻頻區間的該到達方向值、和一可選的擴散值；

其中，對每個個別的時間頻頻區間執行該匹配，以決定至少一個區間作為該匹配對象；以及

其中，該至少一個區間的該估計位置被計算。

【第34項】 如申請專利範圍第32項之該裝置，其中，該聲音分析是一寬頻帶來源分析，產生一擷取的寬頻帶信號、和針對該擷取的寬頻帶信號的一到達方向資訊、以及一可選的擴散信號；

其中，針對來自該第一群組的一擷取的寬頻帶信號和來自該第二群組的一擷取的寬頻帶信號的每對，使用一相關性度量來對於該等擷取的寬頻帶信號執行該匹配，以獲得至少一匹配的寬頻帶對象；以及

其中，該至少一個匹配的寬頻帶對象的該估計位置被計算。

【第35項】 如申請專利範圍第32項之該裝置，其中，該聲場計算器被配置為使用一三角測量計算來執行該至少一匹配對象的該位置的該計算，該三角測量計算接收從該聲場描述所導出的該第一到達方向資訊、和從該另一聲場描述所導出的該第二到達方向資訊，以及包括資訊的該後設資料，從其可以導出關於該參考位置和該另一參考位置之間的一向量的一資訊。

【第36項】 如申請專利範圍第8或35項之該裝置，其中，針對與該不同參考位置相關的該匹配對象的一聲場描述的一計算，該聲場計算器被配置為選擇與該聲場描述有關的資訊或者與該另一聲場描述有關的資訊，其取決於該參考位置或該另一參考位置到由該平移資訊所決定的一聆聽者位置的一距離。

【第37項】 如申請專利範圍第36項之該裝置，其中，該聲場計算器被配置為：

計算該參考位置和該不同參考位置之間的一第一距離；

計算該另一參考位置與該不同參考位置之間的一第二距離；以及

選擇具有距該第一距離和該第二距離之該較小距離的一參考位置的該聲場描述。

【第38項】 如申請專利範圍第32項之該裝置，其中，該聲場計算器被配置為替每個匹配的對象使用從該聲場描述和該另一聲場描述所導出的該到達方向資訊之一，計算一新的到達方向資訊。

【第39項】 如申請專利範圍第37項之該裝置，其中，該聲場計算器被配置為替每個匹配的對象使用該新的到達方向資訊計算一聲場描述，並使用該匹配的對象的該估計的位置來縮放該來源信號或該來源信號的該號聲場描述。

【第40項】 如申請專利範圍第32項之該裝置，其中，該聲場計算器被配置為從該第一群組計算每個非匹配對象的一聲場描述，並使用該相應的到達方向資訊計算來自該第二群組的每個非匹配對象的一聲場描述。

【第41項】 如申請專利範圍第40項之該裝置，其中，該聲場計算器被配置為經由組合該匹配對象的該聲場描述和該非匹配對象的該聲場描述，來計算該修改的聲場描述。

【第42項】 如申請專利範圍第32項之該裝置，其中，該聲場計算器被配置為決定關於該第一群組和該第二群組中的至少一個的一擴散信號的一資訊；以及

其中，該聲場計算器被配置為將該擴散信號添加到一匹配對象的一聲場描述或一非匹配對象的一聲場描述。

【第43項】 一種生成一增強聲場描述的方法，包括：

生成一第一聲場描述以及一第二聲場描述，該第一聲場描述指示關於一第一參考位置的一聲場，該第二聲場描述指示關於一第二參考位置的一聲場，該第二參考位置與該第一參考位置不相同；以及

生成與該聲場的空間資訊有關的後設資料；

其中，生成該後設資料的步驟包括決定該第一聲場描述之一第一幾何描述與該第二聲場描述之一第二幾何描述其中之至少一，以作為該後設資料；以及

其中，該第一聲場描述、該第二聲場描述和包含該第一幾何描述與該第二幾何描述其中之至少一的該後設資料構成該增強聲場描述。

【第44項】一種從一聲場描述和與該聲場描述的空間資訊有關的後設資料生成一修改聲場描述的方法，該方法包括：

使用該空間資訊、該聲場描述和指示從一參考位置到一不同參考位置的一平移的一平移資訊，來計算該修改聲場描述，

其中該計算步驟包括接收與該參考位置有關的一第一聲場描述並接收與不同於該參考位置的一另一參考位置有關的一第二聲場描述；執行對該第一聲場描述的一來源分離以獲得一個或多個擷取對象的一第一群組和對該第二聲場描述的一來源分離以獲得一個或多個擷取對象的一第二群組；計算該擷取對象的到達方向(DoA)資料；匹配該第一聲場描述的擷取對象的該第一群組和該第二聲場描述的擷取對象的該第二群組，以便得到一個或多個匹配對象；估計該一個或多個匹配對象其中之一一匹配對象的一位置；以及基於該一個或多個匹配對象其中之一的該匹配對象的該估計位置和該不同參考位置，來修改該一個或多個匹配對象其中之一的該匹配對象；或

其中該計算步驟包括假設一個聲場模型；估計假設之該聲場模型的信號分量和/或參數；基於該平移資訊或基於該不同的參考位置，來修改該信號分量和/或該參數；以及使用該修改的信號分量和/或該修改的參數，來生成該修改的聲場描述；或

其中該計算步驟包括接收作為該聲場描述的一第一聲場描述和一第二聲場描述；在該第一聲場描述和該第二聲場描述上執行一來源分離，以擷取該第一聲場描述和該第二聲場描述的來源以及擷取的該等來源的到達方向(DoA)資料；使用該DoA資料和該平移資訊，針對擷

取的該等來源的各該來源，計算關於不同之該參考位置的修改的DoA資料；以及處理擷取的該等來源和該修改的DoA資料，以獲得該修改的聲場描述；或

其中該計算步驟包括分析該聲場分量，以針對不同的頻率區間導出到達方向資訊；使用該到達方向資訊和後設資料，以計算每個頻率區間的修改的到達方向資訊，該後設資料包括將一距離資訊與由該頻率區間表示的一來源相關聯的一深度圖；以及使用一距離補償資訊計算該修改聲場描述，該距離補償資訊取決於該來源的該深度圖提供的該距離、以及與該頻率區間相關聯的一新距離，該新距離與該修改到達方向資訊相關；或

其中該計算步驟包括，針對一個或多個來源，計算從該參考位置指向一聲源的一第一向量，其係由該聲場描述所描述之一聲場的一分析所獲得；計算一第二向量，該第二向量使用該第一向量和該平移資訊從該不同參考位置指向該聲源，該平移資訊係定義為從該參考位置到該不同參考位置的一平移向量；以及計算一距離修改值，其使用該不同參考位置、該聲源的一位置和該第二向量，或是使用從該不同參考位置到該聲源的該位置的一距離和該第二向量；或

其中該計算步驟包括經由一聲場分析，從該聲場描述和該來源的方向決定來源；針對一個來源，使用該後設資料，以決定該來源與該參考位置的一距離；決定與該不同參考位置相關的該來源的一新方向，其使用該來源的該方向和該平移資訊；決定與該不同參考位置相關的該來源的一新距離資訊；以及生成該修改聲場描述，其係使用該來源的該新方向、該新距離資訊、以及該聲場描述或對應於從該聲場描述所導出的該來源的來源信號；或

其中該計算步驟包括經由一聲音分析，以決定來自該聲場描述的來源信號和與該參考位置相關的該等來源信號的方向；使用該平移資訊，以計算與該不同參考位置相關的該等來源信號的新方向；計算與該不同參考位置相關的該等來源信號的距離資訊；以及使用該距離資訊、該等來源信號和該等新方向，以合成該修改聲場描述；或

其中該計算步驟包括匹配對該第一聲場描述的來源的一第一群組和該第二聲場描述的來源的一第二群組，以獲得一個或多個匹配來源；針對該一個或多個匹配來源的各該匹配來源決定一到達方向資訊；使用該到達方向資訊和該聲場描述的該後設資料，來決定一來源到該不同參考位置的一距離；以及使用該來源到該不同參考位置的該距離，來決定一縮放因子；或

其中該計算步驟包括對該聲場描述執行一聲音分析，以獲得一個或多個擷取對象的一第一群組或一個或多個擷取對象的一到達方向資訊以及與一另一個參考位置有關的一另一個聲場描述，使用該後設資料獲得一個或多個擷取對象的一第二群組或一個或多個擷取對象的一到達方向資訊；使用來自該群組的對象信號或來自該群組的對象的到達方向資訊，以匹配該第一群組和該第二群組，俾找到至少一個匹配對象；計算該匹配對象的一估計位置，其使用該後設資料、和經由執行該聲音分析獲得該第一群組該匹配對象之所獲得的一第一到達方向資訊、以及經由執行該聲音分析獲得該第二群組該匹配對象之所獲得的一第二到達方向資訊；以及基於該估計的位置和該平移資訊，對該匹配的對象應用一距離補償處理。

【第45項】一種計算機程式，當其用於一計算機或處理器上運行時，執行申請專利範圍第43或44項的方法。

【第46項】 一種儲存媒體，其係儲存有一計算機程式，當該計算機程式用於一計算機或處理器上運行時，執行申請專利範圍第43項的方法。

【第47項】 一種儲存媒體，其係儲存有一計算機程式，當該計算機程式用於一計算機或處理器上運行時，執行申請專利範圍第44項的方法。

【發明圖式】

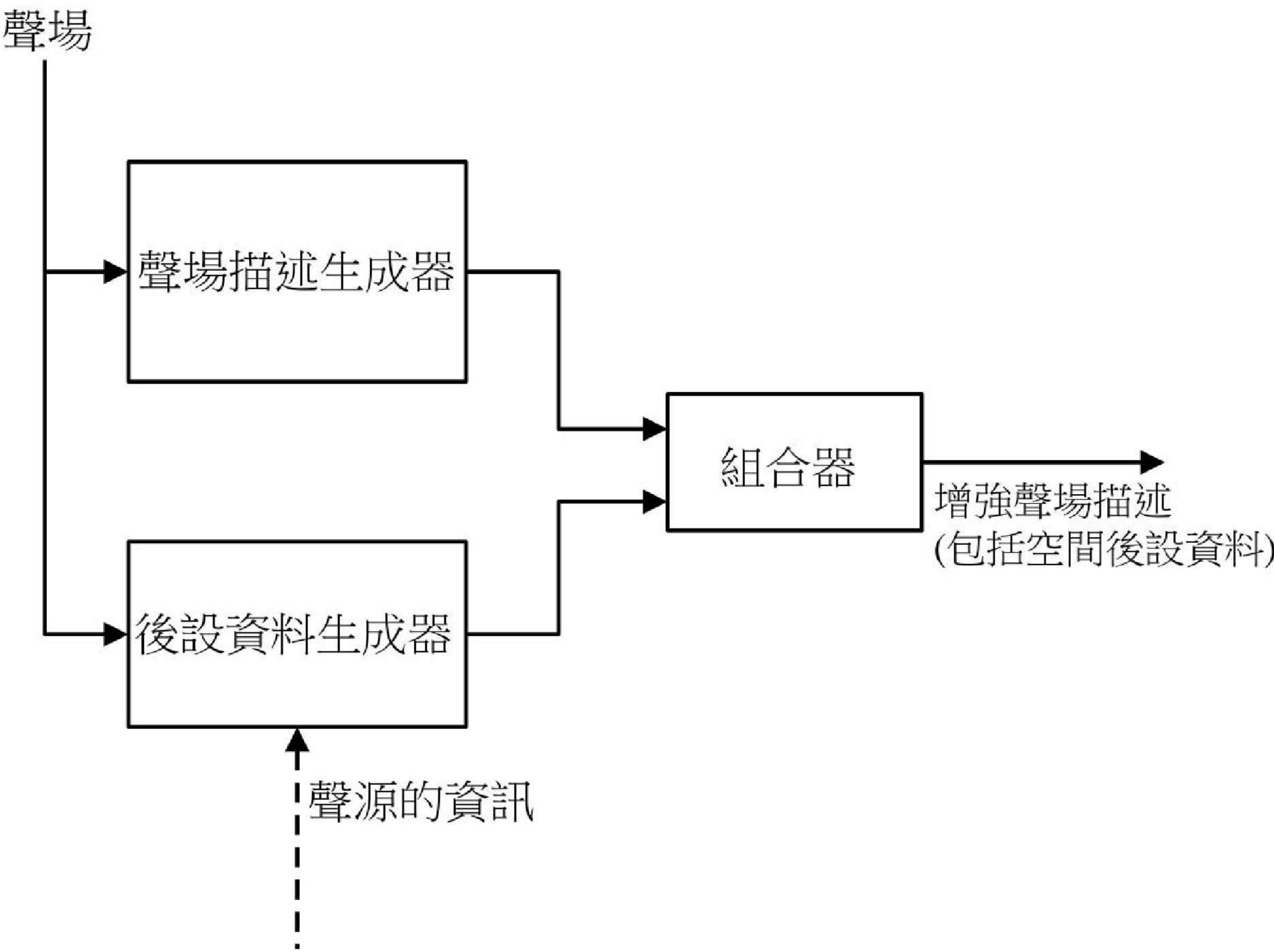


圖 1a

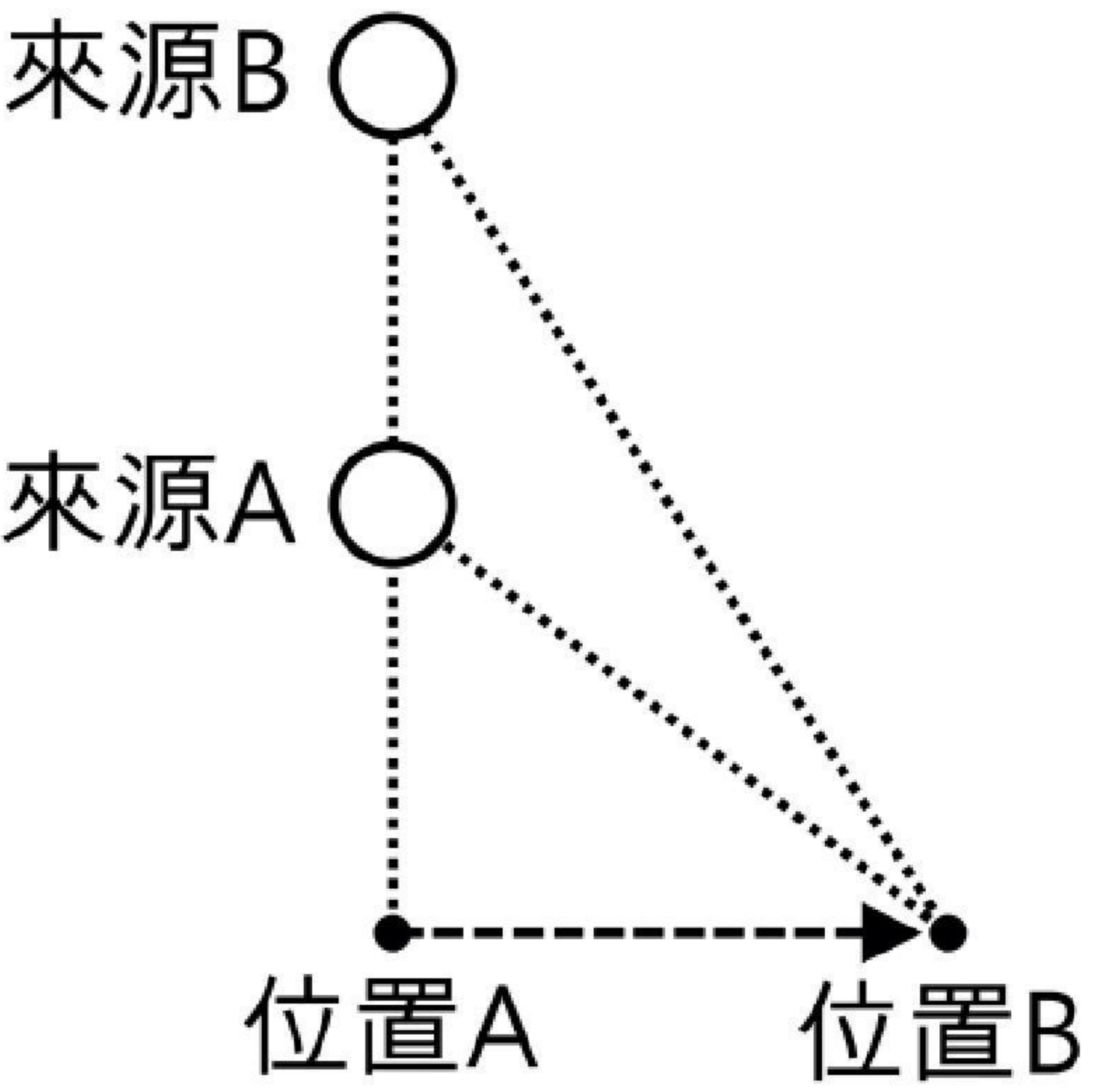


圖 1b

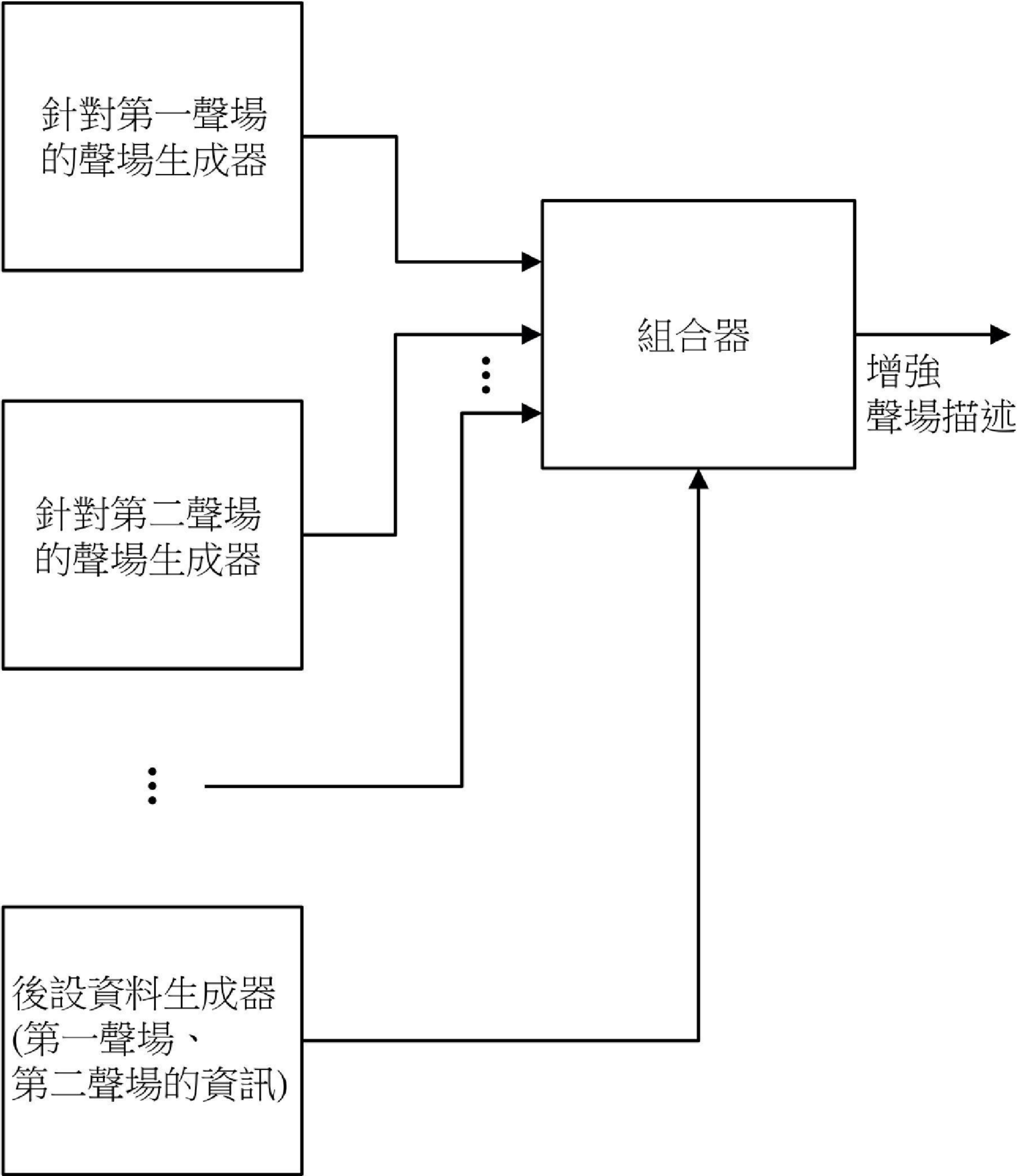


圖 2

	第一聲場描述	第二聲場描述	第一聲場及 第二聲場的資訊
--	--------	--------	------------------

圖 3a

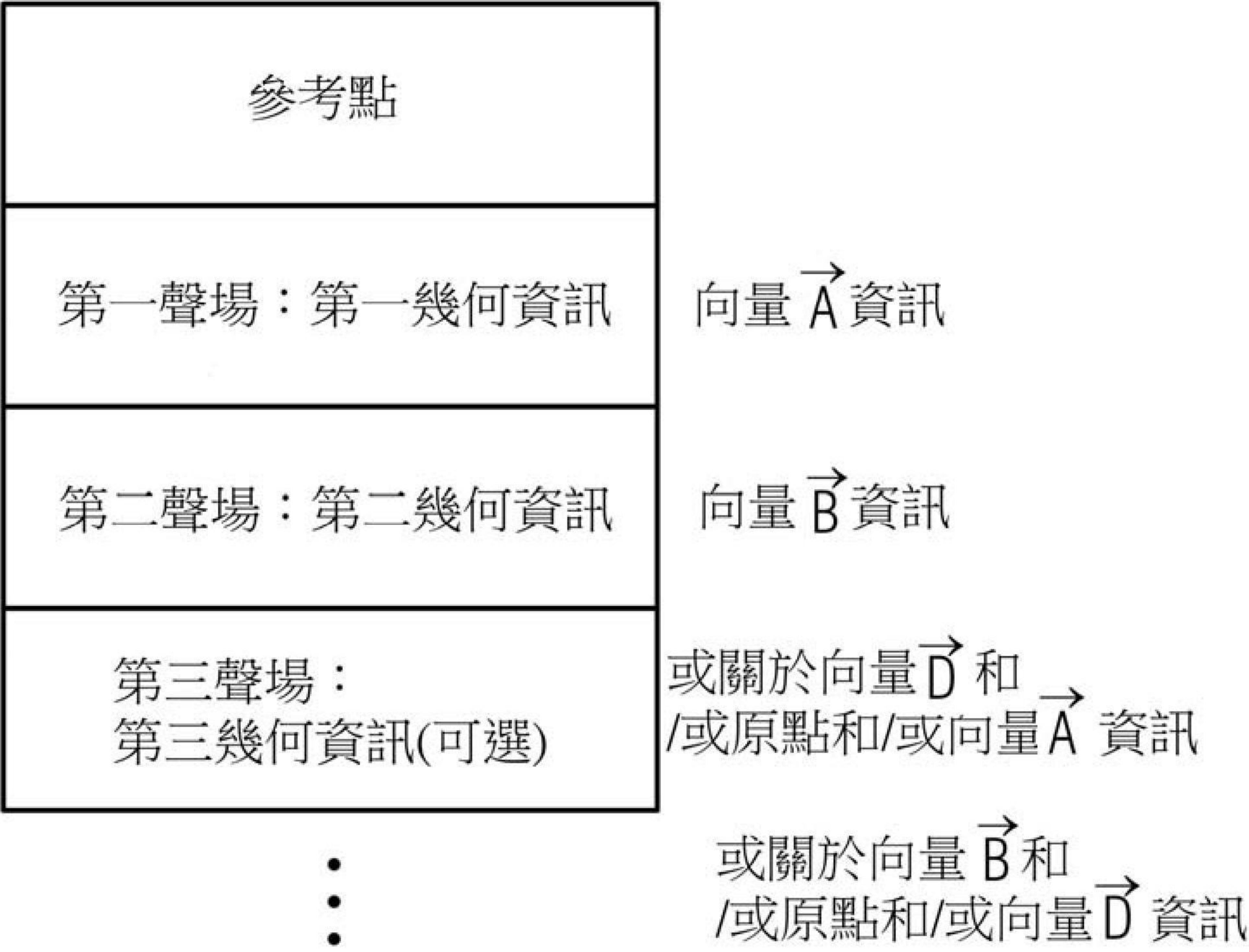


圖 3b

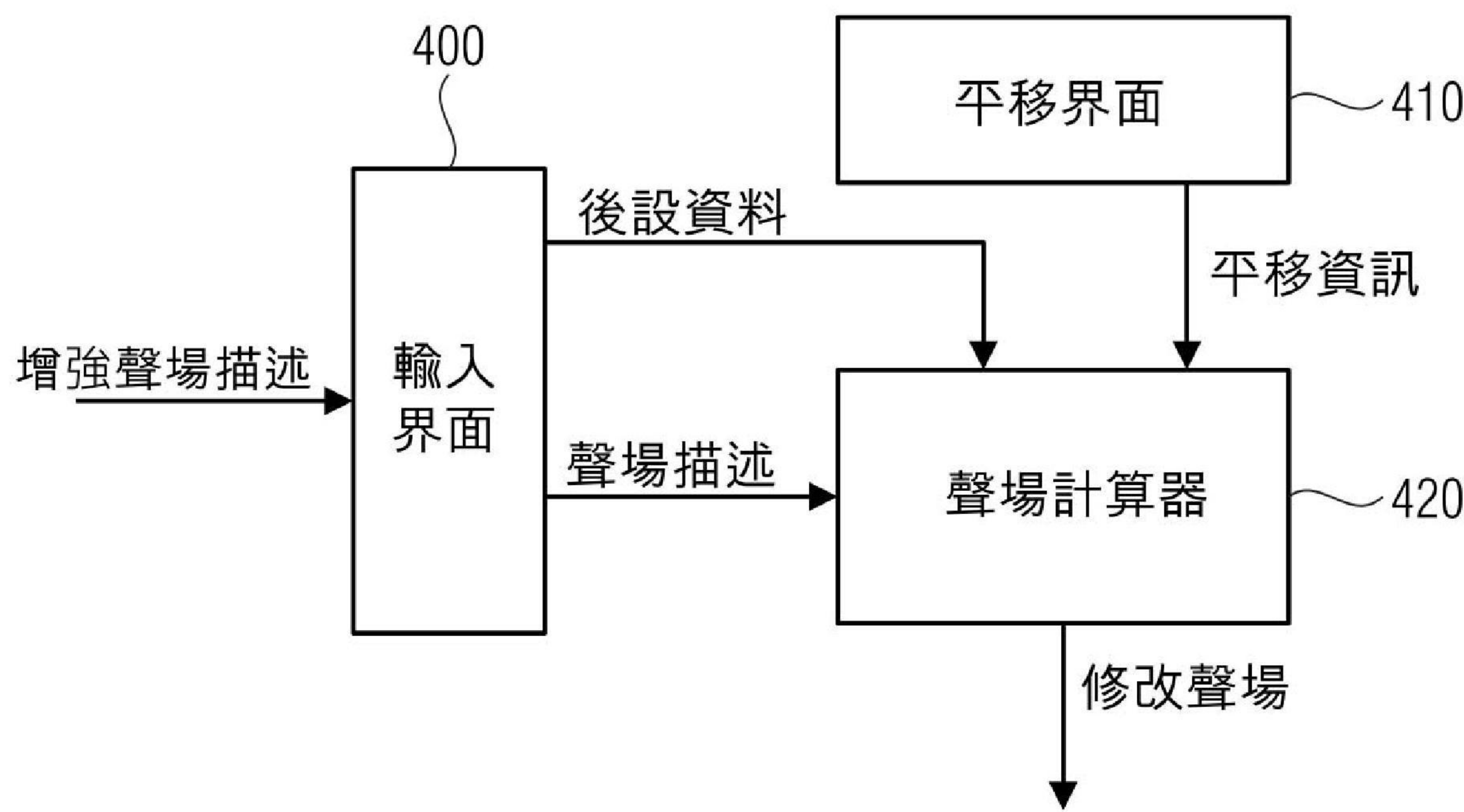


圖 4a

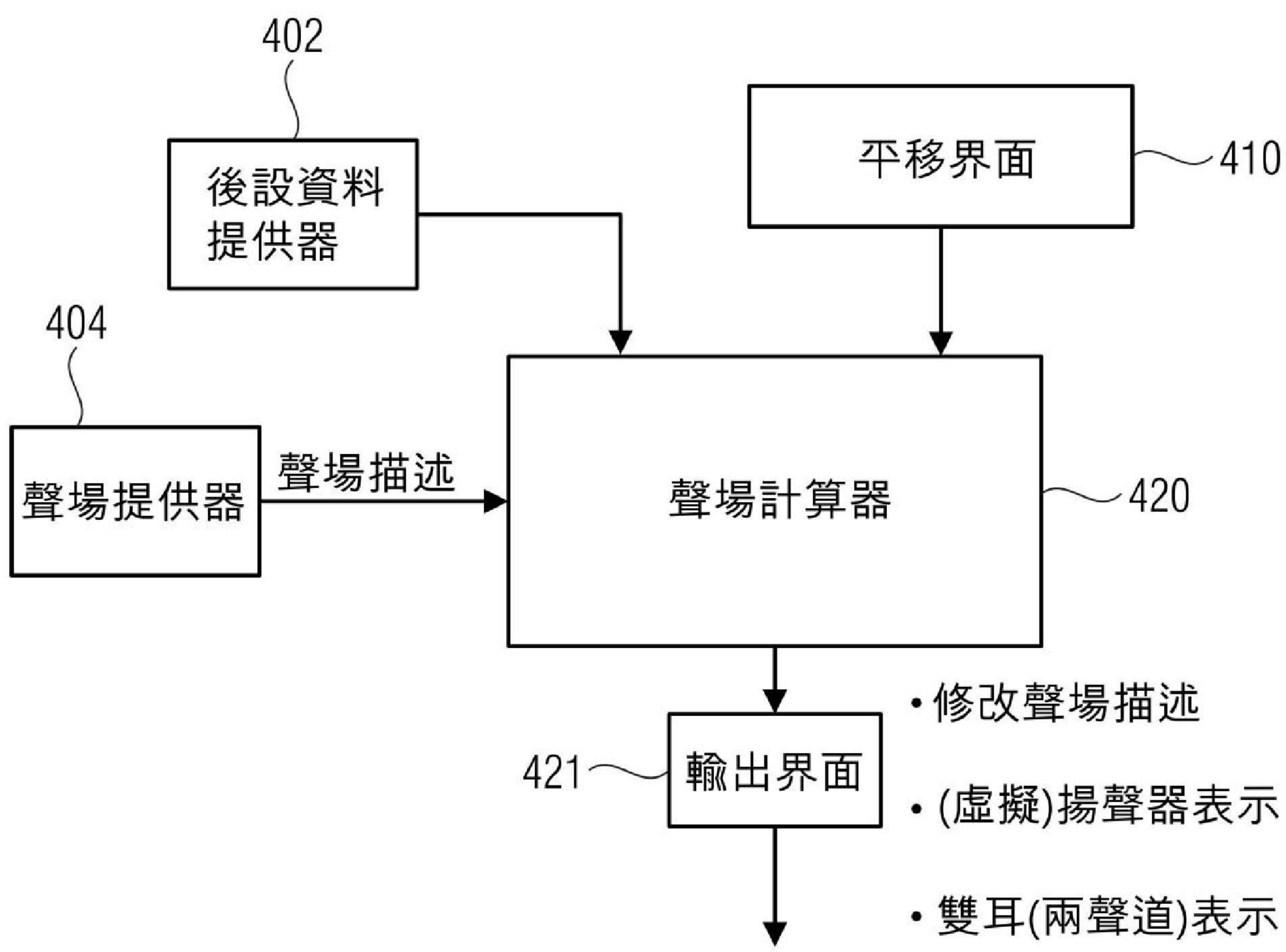


圖 4b

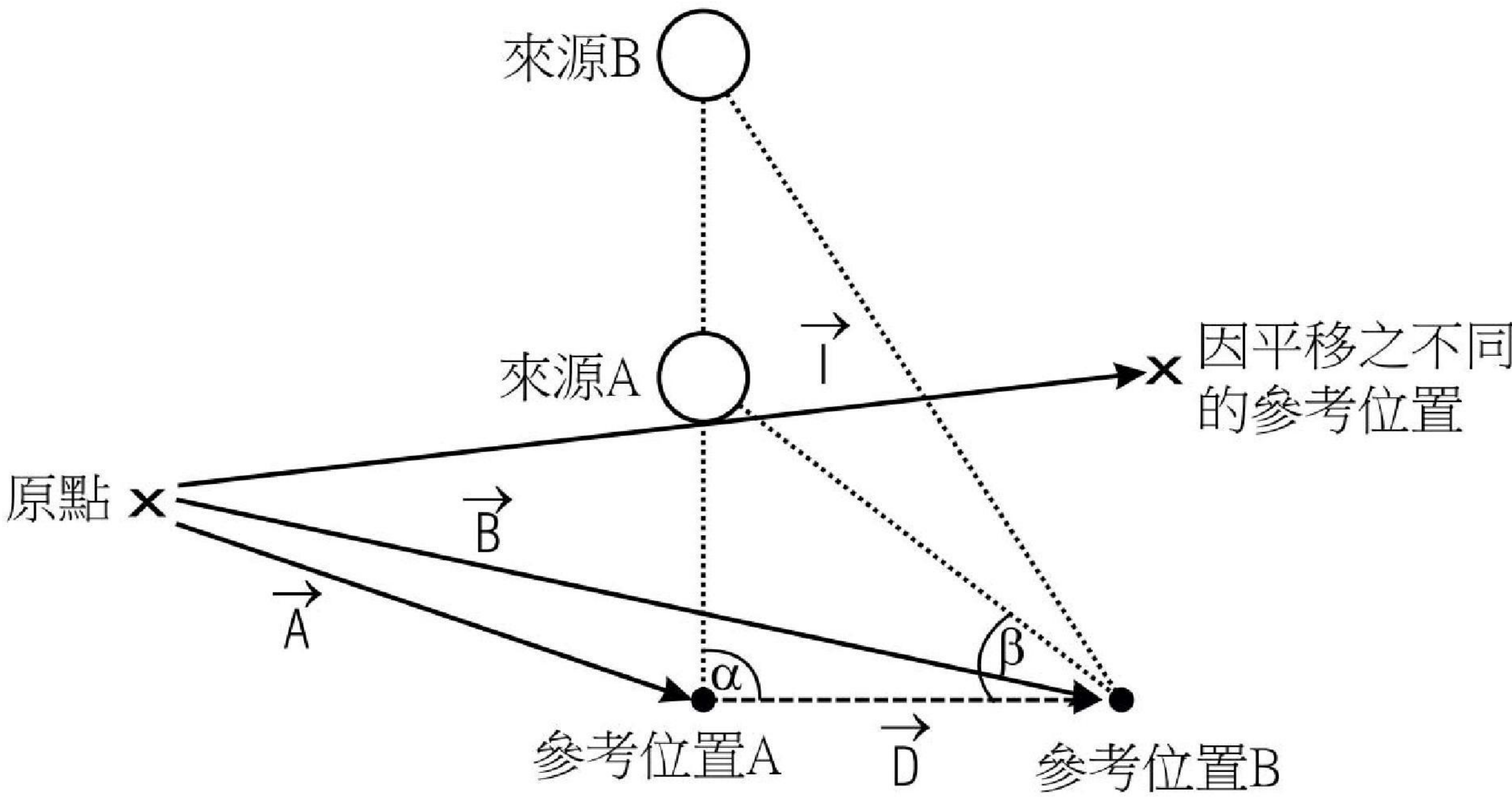


圖 4c

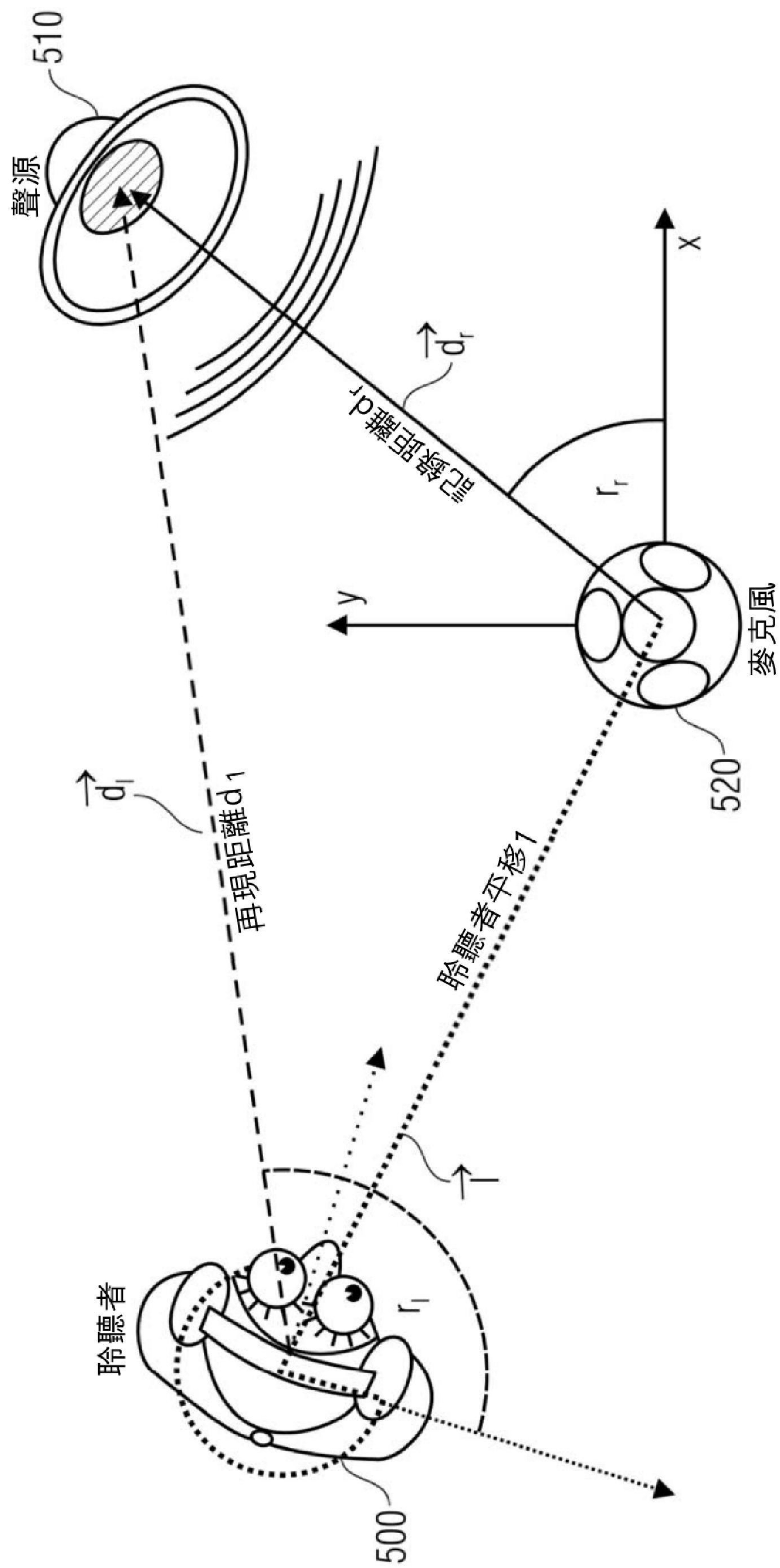


圖 5

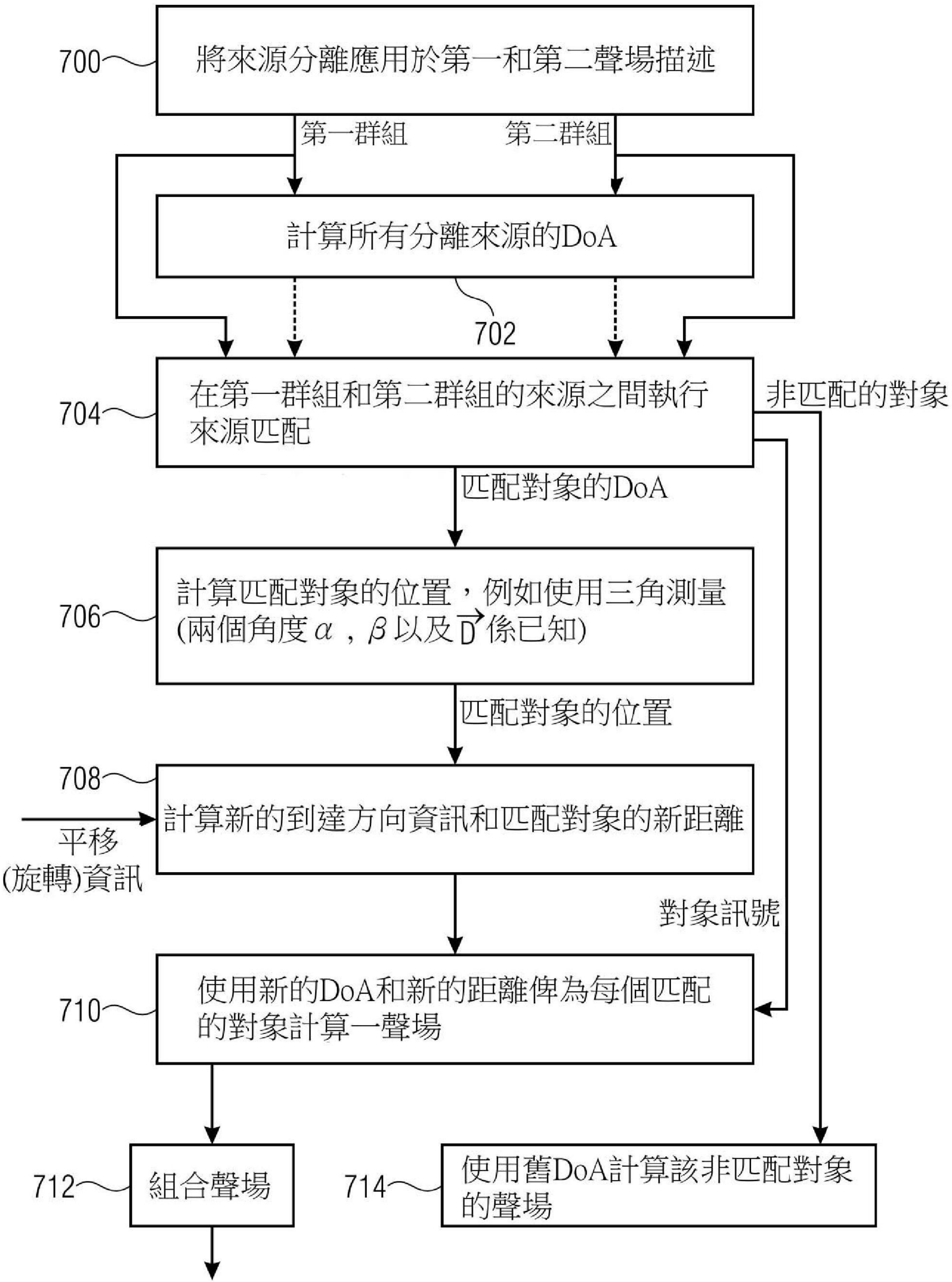


圖 6a

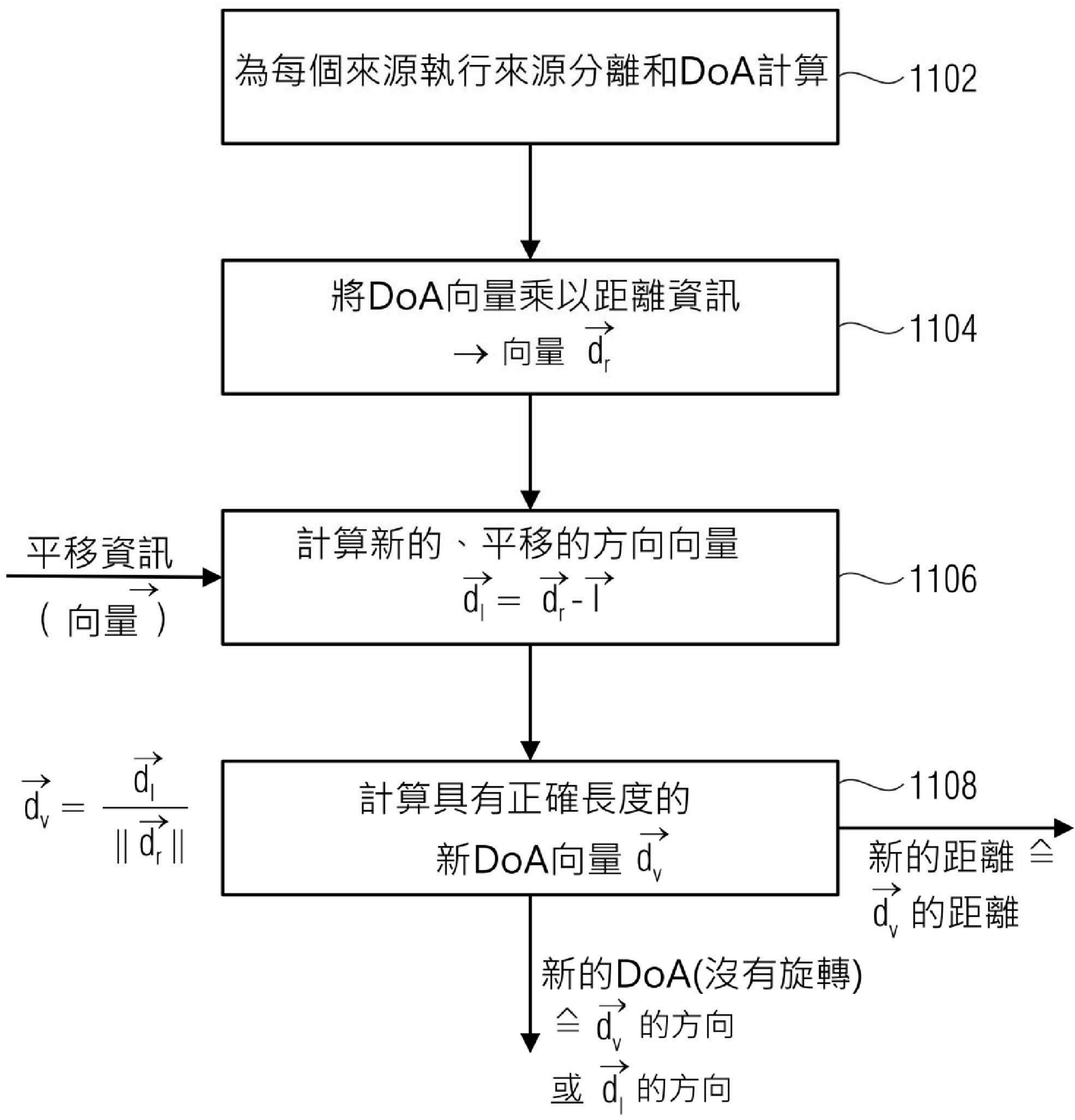


圖 6b

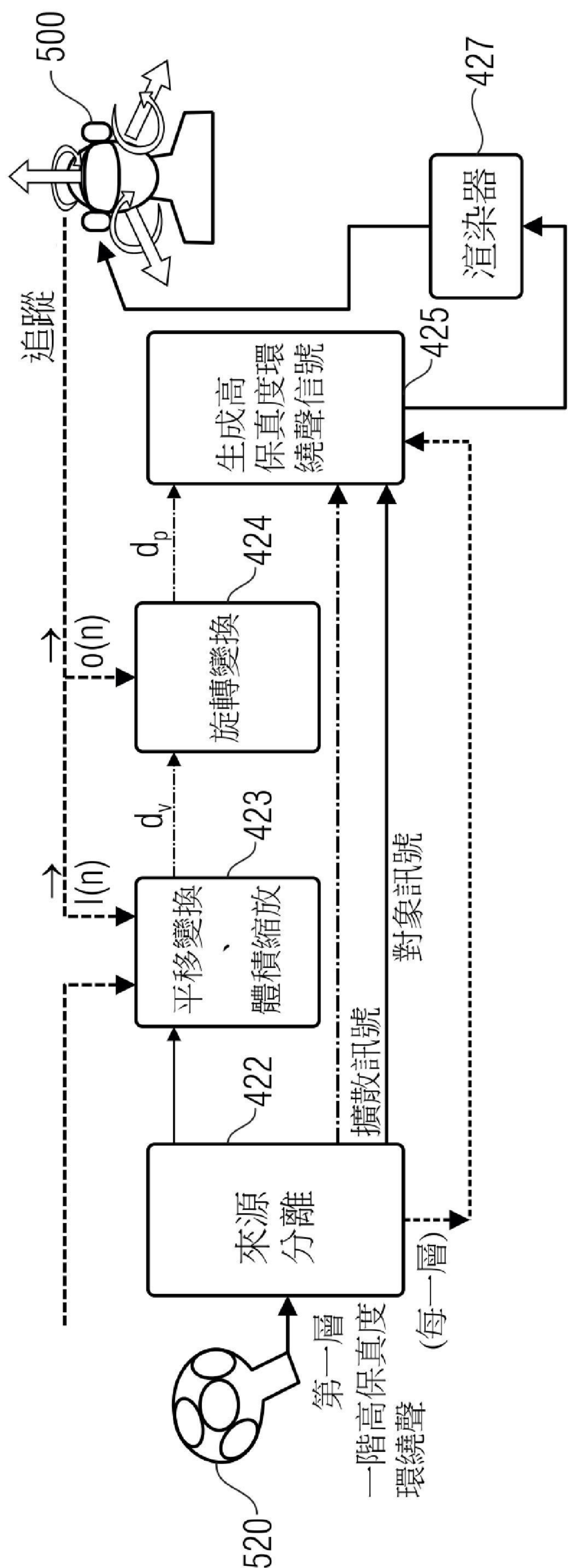


圖 6c

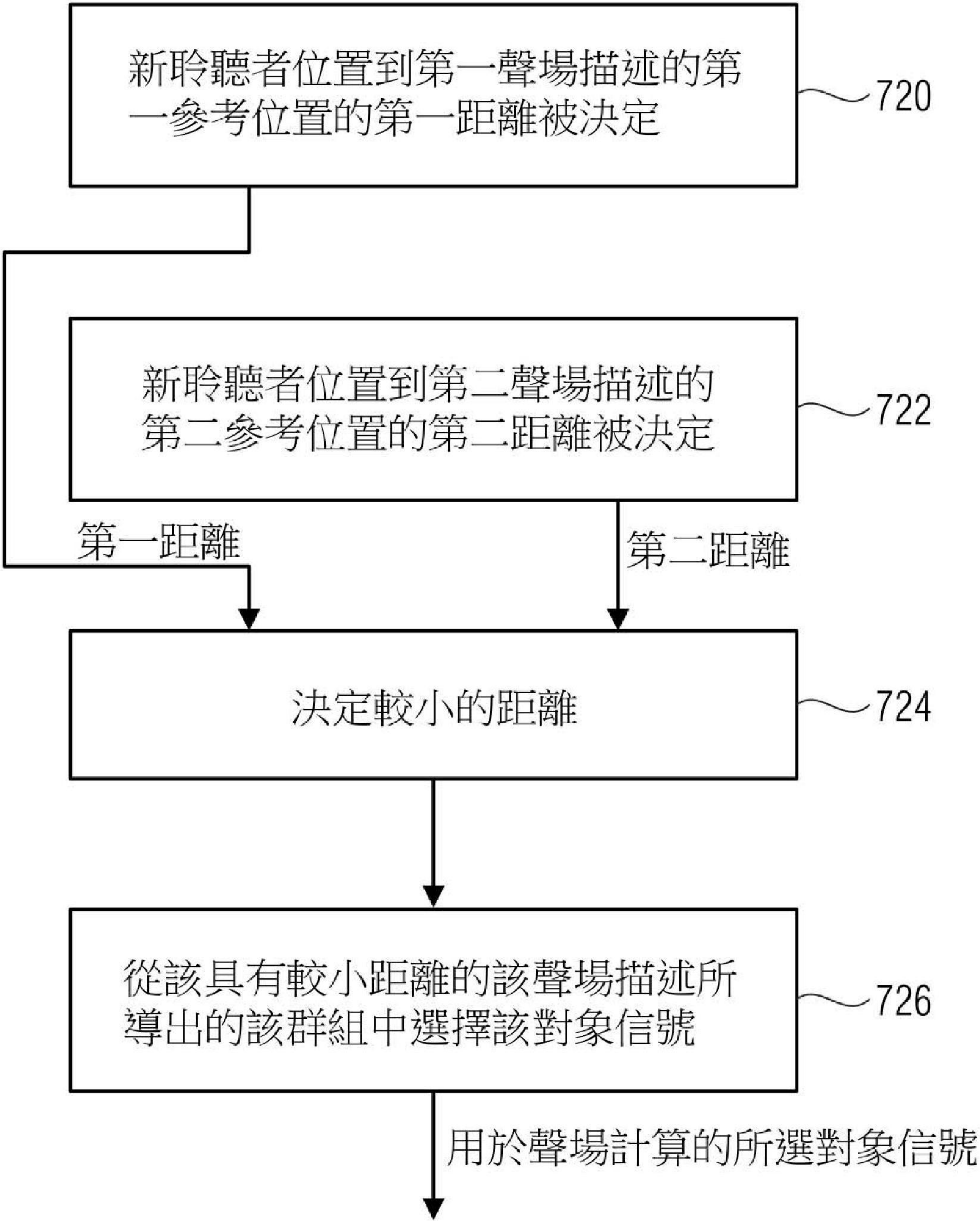
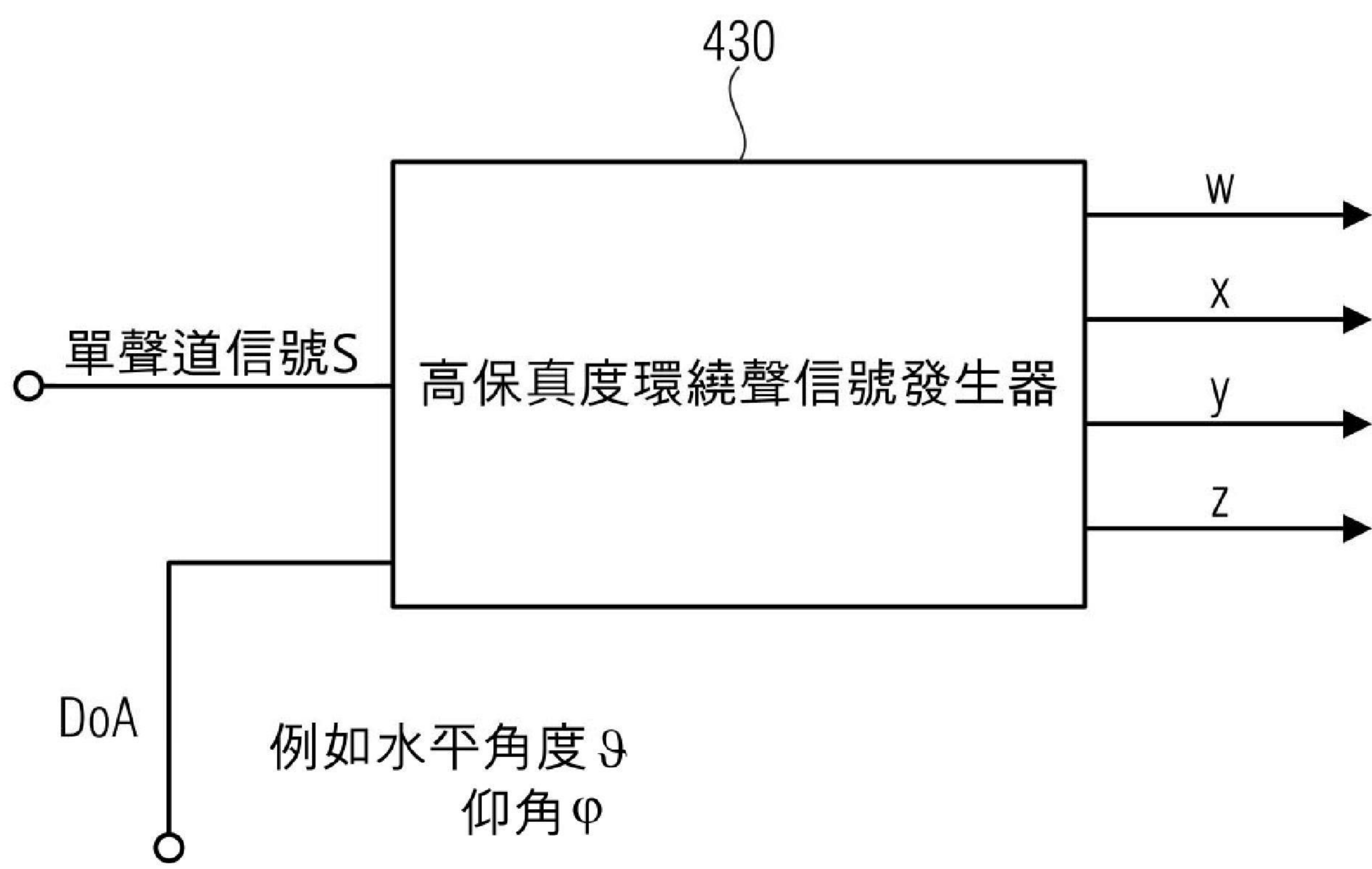


圖 7



全向分量 $W = S \cdot \sqrt{\frac{1}{2}}$

方向分量 $\begin{cases} X = S \cdot \cos \vartheta \cos \varphi \\ Y = S \cdot \sin \vartheta \cos \varphi \\ Z = S \cdot \sin \varphi \end{cases}$

圖 8

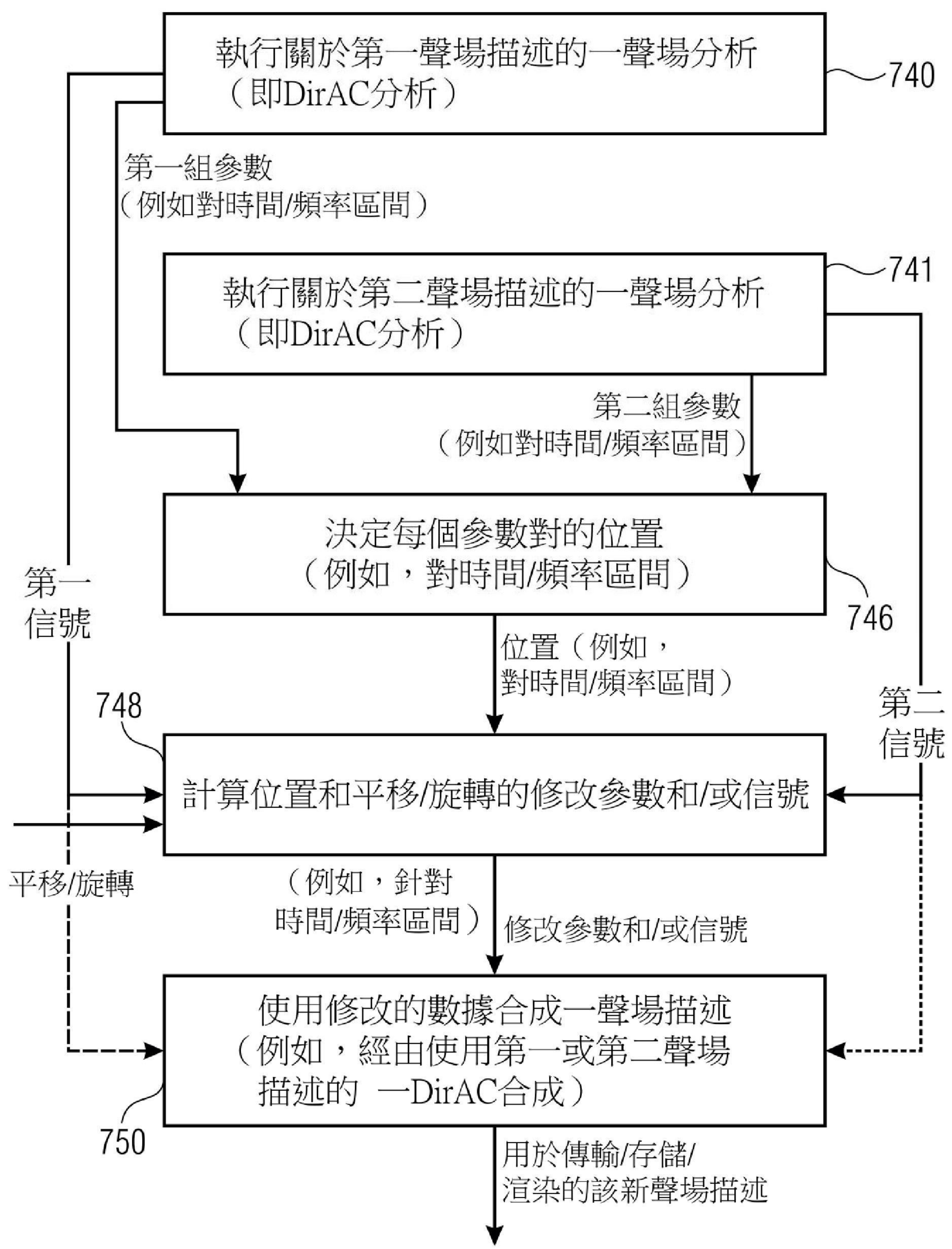


圖 9

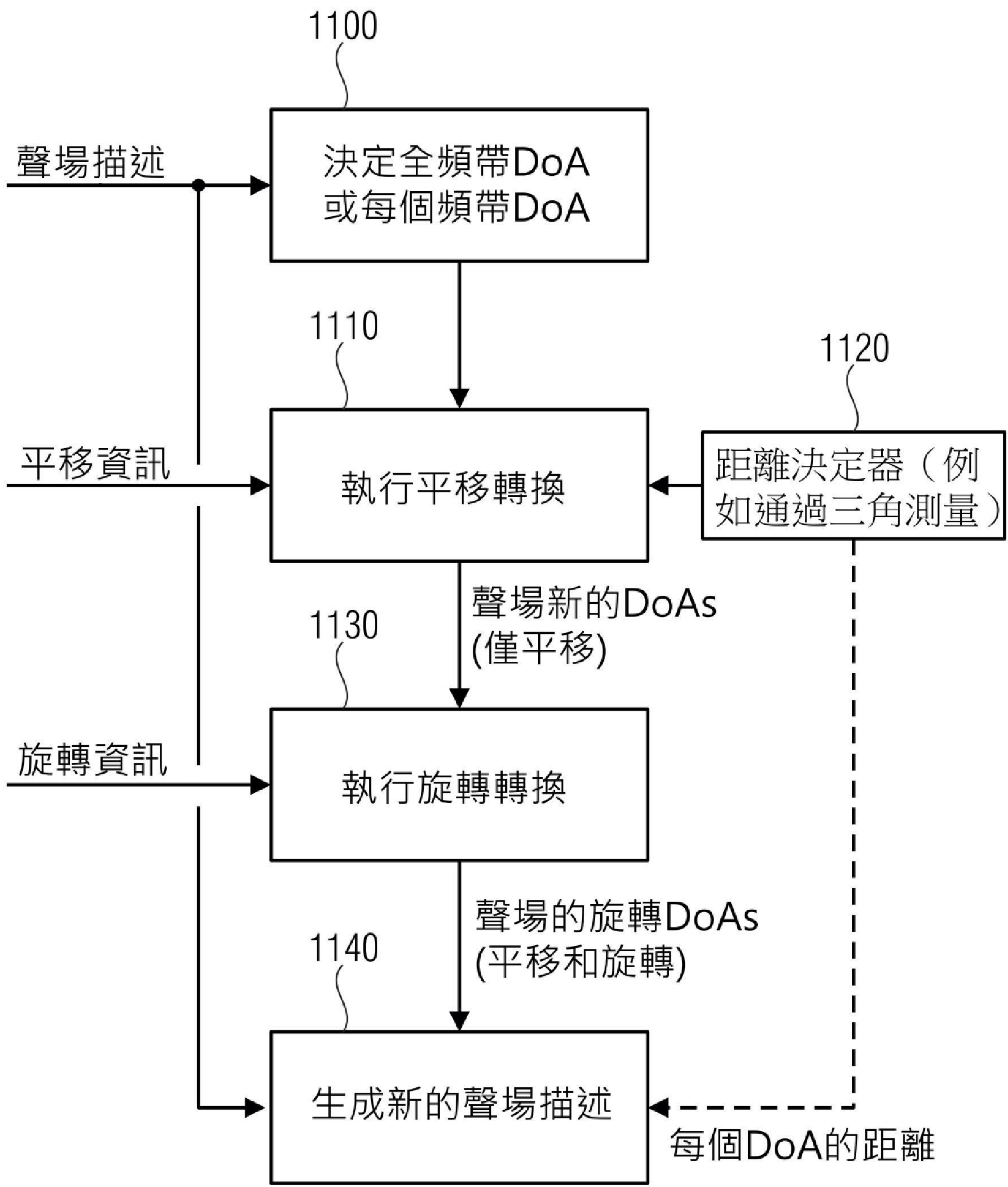
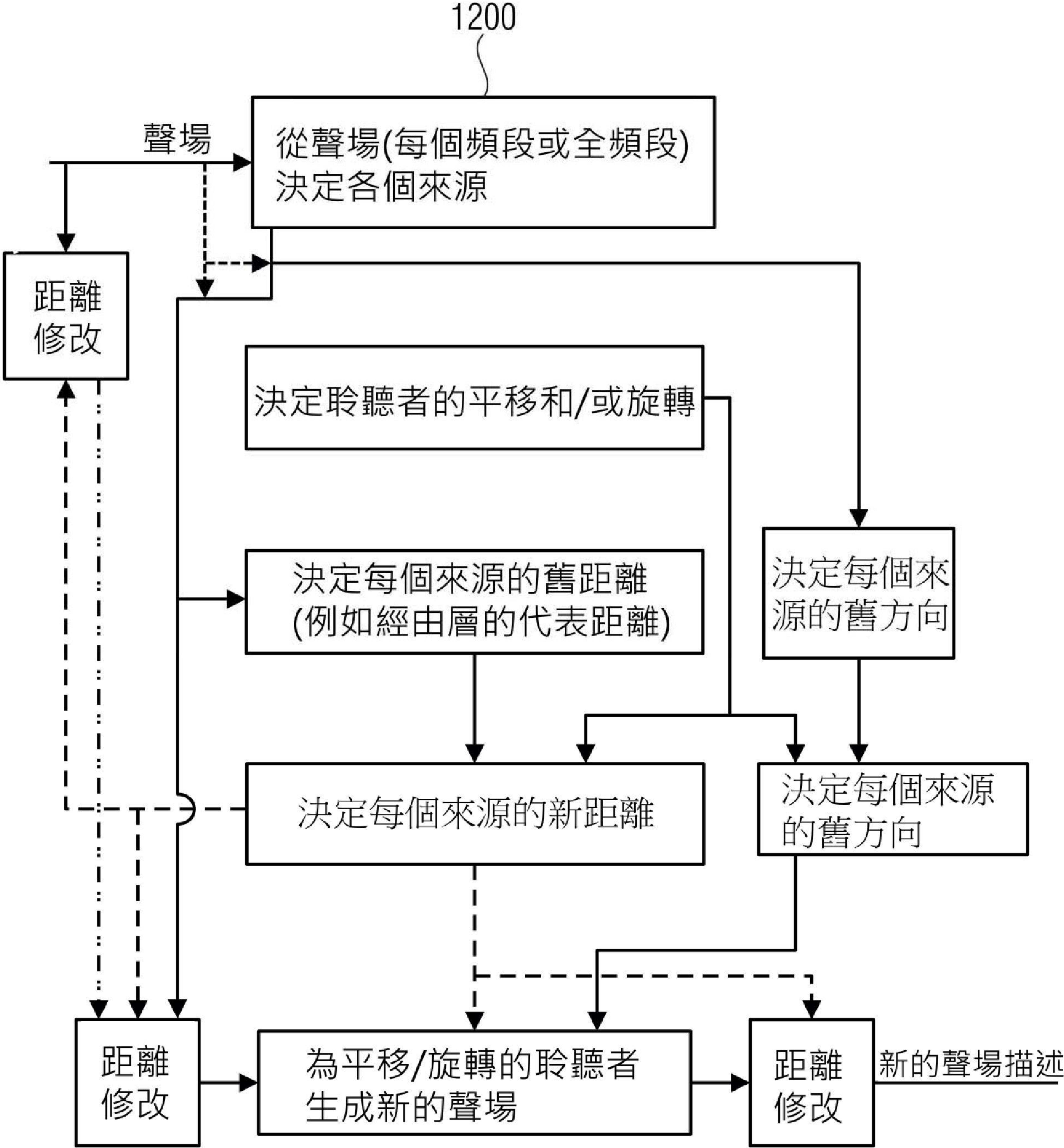
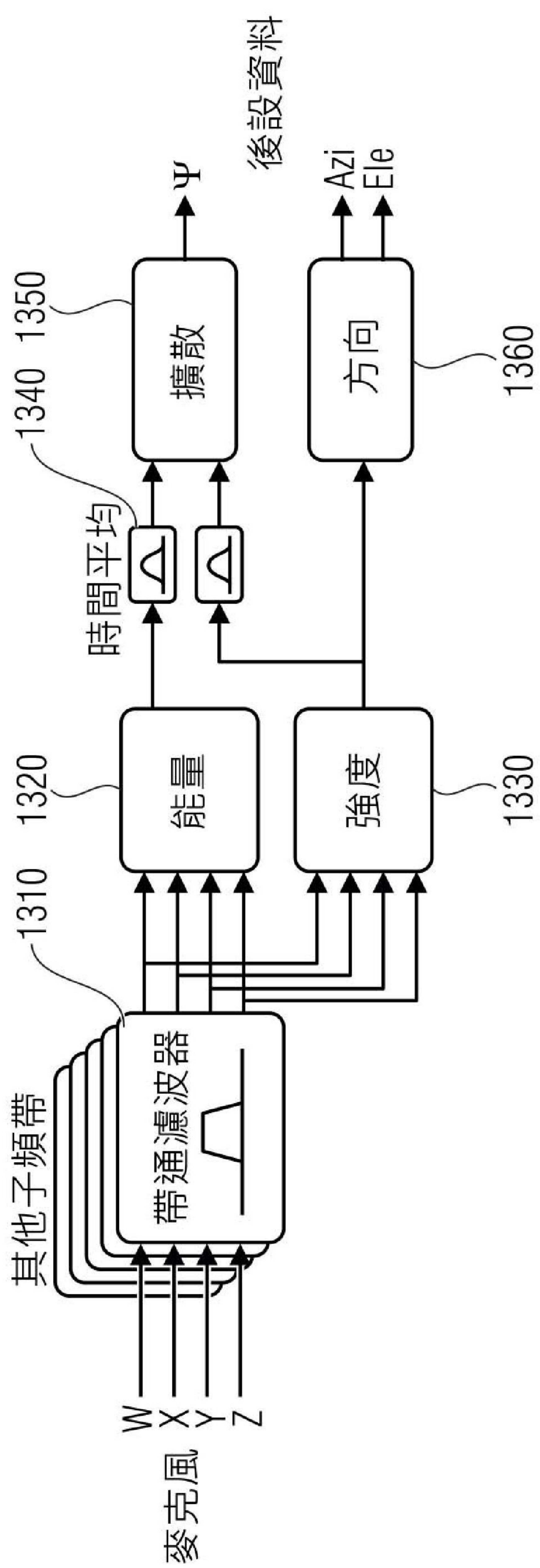


圖 10



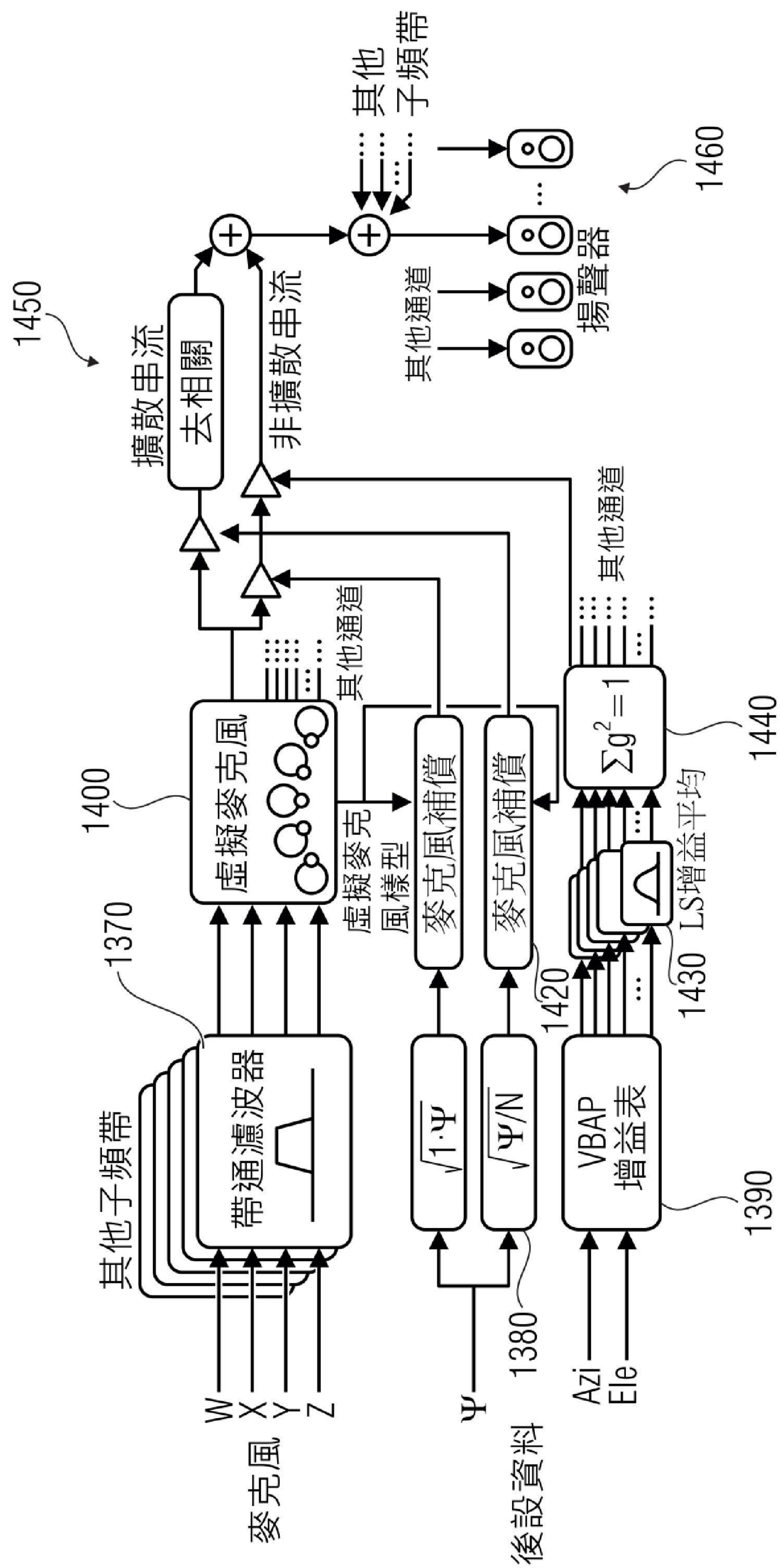
例如對原始聲場、或對來自第1群組和第2群組的分離來源、或對個別新的高保真度環繞聲信號發生器信號、或對來自所有來源的增加的高保真度環繞聲信號發生器信號的距離修改

圖 11



DirAC 分析

圖 12a



DirAC 分析

圖 12b