



(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 I825492 B

(45)公告日：中華民國 112 (2023) 年 12 月 11 日

(21)申請案號：110137741

(22)申請日：中華民國 110 (2021) 年 10 月 12 日

(51)Int. Cl. : **G10L19/008 (2013.01)****G10L19/02 (2013.01)**

(30)優先權：2020/10/13	歐洲專利局	20201633.3
2020/12/18	歐洲專利局	20215651.9
2021/07/07	歐洲專利局	21184367.7

(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：艾肯賽爾 安德烈亞 EICHENSEER, ANDREA (DE)；寇爾斯 斯里坎特 KORSE, SRIKANTH (IN)；拜耶 史丹芬 BAYER, STEFAN (AT)；庫奇 法比恩 KUECH, FABIAN (DE)；錫蓋特 奧莉薇 THIERGART, OLIVER (DE)；夫杰斯 貴勞美 FUCHS, GUILLAUME (FR)；韋克貝克 多米尼克 WECKBECKER, DOMINIK (DE)；賀瑞 捷爾根 HERRE, JUERGEN (DE)；木翠斯 馬庫斯 MULTRUS, MARKUS (DE)

(74)代理人：邱珍元

(56)參考文獻：

TW	201207846A	TW	201503112A
CN	101199121B	CN	102176311B
JP	6268180B2	US	2013/0013323A1
US	2015/0049872A1	US	2015/0081312A1
WO	2019/068638A1		

審查人員：廖天佑

申請專利範圍項數：32 項 圖式數：12 共 84 頁

(54)名稱

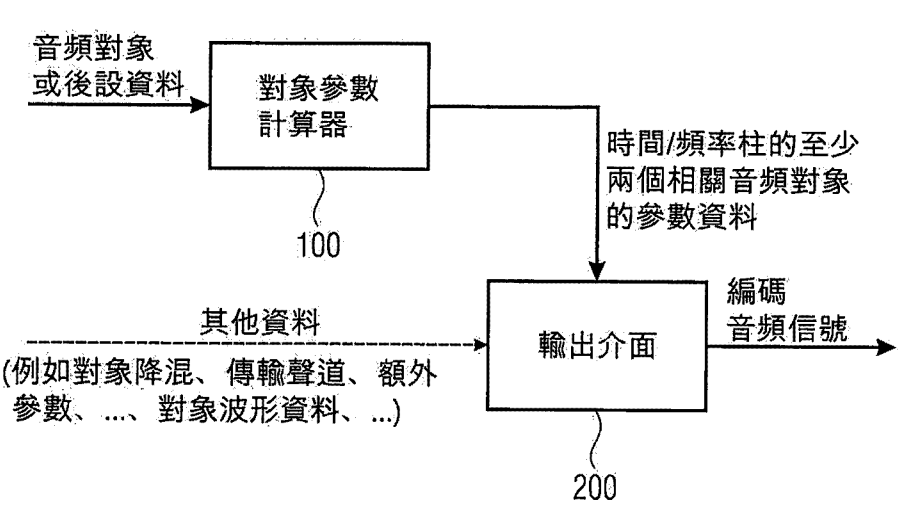
對多個音頻對象進行編碼的設備和方法、使用兩個以上之相關音頻對象進行解碼的設備和方法、電腦程式及資料結構產品

(57)摘要

一種用於編碼複數個音頻對象的設備，包含：一對象參數計算器，其被配置為針對與一時間幀相關的多個頻率柱中的一個以上之頻率柱計算至少兩個相關音頻對象的參數資料，其中該至少兩個相關音頻對象的數量低於該多個音頻對象的總數量，以及一輸出介面，用於輸出一編碼音頻信號，其包括關於一個以上之頻率柱的至少兩個相關音頻對象的參數資料的資訊。

Apparatus for encoding a plurality of audio objects, comprising: an object parameter calculator configured for calculating, for one or more frequency bins of a plurality of frequency bins related to a time frame, parameter data for at least two relevant audio objects, wherein a number of the at least two relevant audio objects is lower than a total number of the plurality of audio objects, and an output interface for outputting an encoded audio signal comprising information on the parameter data for the at least two relevant audio objects for the one or more frequency bins.

指定代表圖：



符號簡單說明：
100:對象參數計算器
200:輸出介面

圖 1a



I825492

【發明摘要】

【中文發明名稱】 對多個音頻對象進行編碼的設備和方法、使用兩個以上之相關音頻對象進行解碼的設備和方法、電腦程式及資料結構產品

【英文發明名稱】 APPARATUS AND METHOD FOR ENCODING A PLURALITY OF AUDIO OBJECTS, APPARATUS AND METHOD FOR DECODING USING TWO OR MORE RELEVANT AUDIO OBJECTS, COMPUTER PROGRAM AND DATA STRUCTURE PRODUCT

【中文】

一種用於編碼複數個音頻對象的設備，包含：一對象參數計算器，其被配置為針對與一時間幀相關的多個頻率柱中的一個以上之頻率柱計算至少兩個相關音頻對象的參數資料，其中該至少兩個相關音頻對象的數量低於該多個音頻對象的總數量，以及一輸出介面，用於輸出一編碼音頻信號，其包括關於一個以上之頻率柱的至少兩個相關音頻對象的參數資料的資訊。

【英文】

Apparatus for encoding a plurality of audio objects, comprising: an object parameter calculator configured for calculating, for one or more frequency bins of a plurality of frequency bins related to a time frame, parameter data for at least two relevant audio objects, wherein a number of the at least two relevant audio objects is lower than a total number of the plurality of audio objects, and an output interface for outputting an encoded audio signal comprising information on the parameter data for the at least two relevant audio objects for the one or more frequency bins.

【指定代表圖】 圖1a

【代表圖之符號簡單說明】

100: 對象參數計算器

200: 輸出介面

【特徵化學式】

無

【發明說明書】

【中文發明名稱】對多個音頻對象進行編碼的設備和方法、使用兩個以上之相關音頻對象進行解碼的設備和方法、電腦程式及資料結構產品

【英文發明名稱】 APPARATUS AND METHOD FOR ENCODING A PLURALITY OF AUDIO OBJECTS, APPARATUS AND METHOD FOR DECODING USING TWO OR MORE RELEVANT AUDIO OBJECTS, COMPUTER PROGRAM AND DATA STRUCTURE PRODUCT

【技術領域】

【0001】 本發明關於對音頻信號(如音頻對象)進行編碼，以及對編碼音頻信號(如編碼音頻對象)進行解碼。

【先前技術】

【0002】 導論

【0003】 本說明書描述了一種使用定向音頻編碼(DirAC)，以低位元率對基於對象的音頻內容進行編碼和解碼的參數化方法。所呈現的實施例作為3GPP沉浸式語音和音頻服務(IVAS)編解碼器的一部分運行，並且其中提供了對具有後設資料的獨立流(ISM)模式的低位元率的一種有利替代方案，這是一種離散編碼方法。

【0004】 習知技術

【0005】 對象的離散編碼

【0006】 對基於對象的音頻內容進行編碼的最直接的方法是單獨編碼並連同相應的後設資料來傳輸對象，這種方法的主要缺點是隨著對象數量的增加對該等對象進行編碼所需的位元消耗過高，此問題的一個簡單解決方案是採用“參數化方法”，其中一些相關參數係從輸入信號中進行計算、量化並與合適的降混信號一起傳輸，該降混信號組合了多個對象波形。

【0007】 空間音頻對象編碼(SAOC)

【0008】 空間音頻對象編碼[SAOC_STD, SAOC_AES]是一種參數化方法，其中編碼器基於某些降混矩陣D計算出一降混信號和一組參數，並將兩者傳輸到解碼器。該等參數代表所有個別對象的心理聲學相關屬性和關係。在解碼器處，使用渲染矩陣R將該降混信號渲染到特定的揚聲器佈局。

【0009】 SAOC的主要參數是大小為N×N的對象共變異數矩陣E，其中N是指對象的數量，此參數作為對象級別差異(OLD)和可選的對象間共變異數(IOC)傳輸到解碼器。

【0010】 矩陣E的各個元素 $e_{i,j}$ 由下式給出：

$$e_{i,j} = \sqrt{OLD_i OLD_j IOC_{i,j}}$$

【0011】 對象級別差(OLD)定義為

$$OLD_i^{l,m} = \frac{nr g_{i,i}^{l,m}}{NRG^{l,m}},$$

【0012】 其中 $nr g_{i,i}^{l,m}$ 和絕對對象能量(NRG)被描述為

$$nr g_{i,i}^{l,m} = \frac{\sum_{n \in l} \sum_{k \in m} x_i^{n,k} (x_i^{n,k})^*}{\sum_{n \in l} \sum_{k \in m} 1} + \varepsilon$$

以及

$$NRG^{l,m} = \max_i (nr g_{i,i}^{l,m}),$$

【0013】 其中i和j分別是對象 x_i 和 x_j 的對象索引，n表示時間索引，k表示頻率索引，l表示一組時間索引，m表示一組頻率索引， ε 是一個加性常數，以避免分母為零，例如， $\varepsilon = 10$ 。

【0014】 輸入對象(IOC)的相似性度量可以例如由互相關給出：

$$IOC_{i,j}^{l,m} = Re \left\{ \frac{nr g_{i,j}^{l,m}}{\sqrt{nr g_{i,i}^{l,m} nr g_{j,j}^{l,m}}} \right\}$$

【0015】 大小為 $N_{dmx} \times N$ 的降混矩陣 D 由元素 $d_{i,j}$ 定義，其中 i 表示降混信號的聲道索引， j 表示對象索引。對於一立體聲降混($N_{dmx} = 2$)， $d_{i,j}$ 由參數DMG和DCLD計算為

$$d_{0,j} = 10^{0.05DMG_j} \sqrt{\frac{10^{0.1DCLD_j}}{1 + 10^{0.1DCLD_j}}}, \quad d_{1,j} = 10^{0.05DMG_j} \sqrt{\frac{1}{1 + 10^{0.1DCLD_j}}},$$

【0016】 其中 DMG_i 和 $DCLD_i$ 由下式給出：

$$DMG_i = 10 \log_{10}(d_{1,i}^2 + d_{2,i}^2 + \varepsilon)$$

$$DCLD_i = 10 \log_{10} \left(\frac{d_{1,i}^2 + \varepsilon}{d_{2,i}^2 + \varepsilon} \right)$$

【0017】 對於單聲道降混($N_{dmx} = 1$)的情況， $d_{i,j}$ 僅從DMG參數計算為

$$d_{0,j} = 10^{0.05DMG_j},$$

其中

$$DMG_i = 10 \log_{10}(d_{1,i}^2 + \varepsilon)$$

【0018】 空間音頻對象編碼-3D (SAOC-3D)

【0019】 空間音頻對象編碼-3D的音頻再現(SAOC-3D)[MPEGH_AES、MPEGH_IEEE、MPEGH_STD、SAOC_3D_PAT]是上述MPEG SAOC技術的延伸，該技術以非常高效的位元率的方式壓縮和渲染聲道和對象信號。

【0020】 與SAOC的主要區別在於：

- 雖然原始SAOC最多僅支援兩個降混聲道，但SAOC-3D可以將多對象輸入映射到任意數量的降混聲道(以及相關的輔助資訊)。

- 與使用環繞音訊(MPEG Surround)作為多聲道輸出處理器的典型SAOC相比，直接渲染到多聲道輸出。
- 捨棄了一些工具，例如殘量編碼工具。

【0021】 儘管存在這些差異，但從參數角度來看，SAOC-3D與SAOC是相同的。SAOC-3D解碼器係類似於SAOC解碼器，可接收多聲道降混 X 、共變異數矩陣 E 、渲染矩陣 R 、和降混矩陣 D 。

【0022】 渲染矩陣 R 由輸入聲道和輸入對象進行定義，並分別從格式轉換器(聲道)和對象渲染器(對象)接收。

【0023】 降混矩陣 D 由元素 $d_{i,j}$ 進行定義，其中 i 表示降混信號的聲道索引， j 表示對象索引，並根據降混增益(DMG)計算得出：

$$d_{i,j} = 10^{0.05 DMG_{i,j}},$$

其中

$$DMG_{i,j} = 10 \log_{10}(d_{i,j}^2 + \varepsilon).$$

【0024】 大小為 $N_{\text{out}} \times N_{\text{out}}$ 的輸出共變異數矩陣 C 定義為：

$$C = R E R^*$$

【0025】 相關方案

【0026】 存在其他幾種本質上與上述SAOC相似但略有不同的方案：

- 對象的雙耳線索編碼(BCC)已在例如[BCC2001]中進行描述，並且其是SAOC技術的前身。
- 聯合對象編碼(JOC)和高級聯合對象編碼(A-JOC)執行與SAOC類似的功能，同時在解碼器側提供大致分離的對象，而無需將其渲染到特定的輸出揚聲器佈局[JOC_AES、AC4_AES]。該技術將昇混矩陣的元素從降混傳輸到分離的對象以作為參數(並非OLD)。

【0027】 定向音頻編碼(DirAC)

【0028】 另一種參數化方法是定向音頻編碼，定向音頻編碼(DirAC)[Pulkki2009]是空間聲音的感知驅動再現，其假設在某一時刻和一個臨界

頻帶，人類聽覺系統的空間解析度僅限於解碼一個方向線索和一個聽覺間相關性的線索。

【0029】 基於這些假設，DirAC通過交叉衰落兩個串流(一非定向擴散流和一定向非擴散流)來表示一個頻段中的空間聲音，DirAC處理分兩個階段執行：如圖12a和12b所示的分析階段和合成階段。

【0030】 在DirAC分析階段，將B格式的一階重合麥克風視為輸入，並在頻域中分析聲音的擴散和到達方向。

【0031】 在DirAC合成階段，聲音被分為兩個串流，包括非擴散流和擴散流，非擴散流使用幅度平移再現為點源，這可以通過使用向量基幅度平移(VBAP)[Pulkki1997]來完成，擴散流負責包圍的感覺，並通過將相互去相關的信號傳送到揚聲器而產生。

【0032】 圖12a中的分析階段包括一頻帶濾波器1000、一能量估計器1001、一強度估計器1002、時間平均元件999a和999b、一擴散計算器1003、及一方向計算器1004，計算的空間參數是對每個時間/頻率磚的0到1之間的擴散值，以及由方塊1004生成的每個時間/頻率磚的到達方向參數。在圖12a中，方向參數包括一方位角和一仰角，其指示聲音相對於參考或收聽位置的到達方向，特別是相對於麥克風所在的位置，從該位置收集輸入到頻帶濾波器1000的四個分量信號。在圖12a的圖示中，這些分量信號是一階環繞聲分量，包括一全向分量W、一方向分量X、另一方向分量Y和另一方向分量Z。

【0033】 圖12b所示的DirAC合成階段包括一頻帶濾波器1005，用於生成B格式麥克風信號W、X、Y、Z的時間/頻率表示，個別時間/頻率磚的對應信號被輸入到一虛擬麥克風階段1006，其為每個聲道生成一虛擬麥克風信號。特別地，為了產生虛擬麥克風信號，例如對於中央聲道，一虛擬麥克風被指向中央聲道的方向，且產生的信號是中央聲道的對應的分量信號。然後，通過一直接信號分支1015和一擴散信號分支1014處理該信號，兩個分支都包括相應的增益調節器或放大器，其由從方塊1007、1008中的原始擴散參數導出的擴散值所控制，並且在方塊1009、1010中進一步處理，以獲得一定的麥克風補償。

【0034】 直接信號分支1015中的分量信號也使用從由一方位角和一仰角組成的方向參數導出的一增益參數進行增益調整。特別地，這些角度被輸入到VBAP(向量基幅度平移)增益表1011，其結果針對每個聲道被輸入到一揚聲器增益平均階段1012以及一再歸一化器1013，然後將所得增益參數轉發到在直接信號分支1015中的放大器或增益調節器。在去相關器1016的輸出處生成的擴散信號和直接信號或非擴散流在組合器1017中組合，然後在另一個組合器1018中添加其他子頻帶，例如，其可以是一個合成濾波器組，因此，可以生成針對某個揚聲器的揚聲器信號，並且針對某個揚聲器設置中的其他揚聲器1019的其他聲道可以執行相同的流程。

【0035】 圖12b顯示DirAC合成的高品質版本，其中合成器接收所有B格式信號，從中為每個揚聲器方向計算虛擬麥克風信號。所使用的方向圖案通常是偶極。然後根據關於分支1016和1015所討論的後設資料以非線性方式修改虛擬麥克風信號。圖12b中未示出DirAC的低位元率版本，但是，在這種低位元率版本中，僅傳輸單個聲道的音頻。處理的不同之處在於，所有虛擬麥克風信號都將被接收到的這個單一音頻聲道所取代。虛擬麥克風信號被分為兩個串流，包括一擴散流和一非擴散流，其係分別進行處理。通過使用向量基振幅平移(VBAP)將非擴散聲音再現為點源。在平移中，單聲道聲音信號在與揚聲器特定的增益因子相乘後應用於揚聲器的子集。增益因子是使用揚聲器設置和指定平移方向的資訊計算的。在低位元率版本中，輸入信號被簡單地平移到後設資料暗示的方向。在高品質版本中，每個虛擬麥克風信號都乘以相應的增益因子，以便產生與平移相同的效果，但不太容易出現任何非線性偽物。

【0036】 擴散聲音合成的目的是創造環繞聽者的聲音感知。在低位元率版本中，通過去相關輸入信號並從每個揚聲器再現，來再現擴散流。在高品質版本中，擴散流的虛擬麥克風信號已經存在一定程度的不相關性，只需對其進行輕度去相關即可。

【0037】 DirAC參數，亦稱為空間後設資料，由擴散度和方向元組組成，在球面坐標中由兩個角度表示，即方位角和仰角。如果分析和合成階段都在解碼器側運行，則DirAC參數的時頻解析度可以選擇為與用於DirAC分析和

合成的濾波器組相同，即每個時隙的不同參數集和音頻信號的濾波器組表示的頻率柱。

【0038】 目前已經付出一些努力來減少後設資料的大小，使DirAC範式能夠用於空間音頻編碼和電話會議場景[Hirvonen2009]。

【0039】 在專利申請號[WO2019068638]中，介紹了一種基於DirAC的通用空間音頻編碼系統，與專為B格式(一階環繞聲格式)輸入設計的典型DirAC相比，該系統可以接受一階或更高階的環繞聲、多聲道或基於對象的音頻輸入，還允許混合式輸入信號。所有信號類型都以單獨或組合的方式有效地編碼和傳輸，前者在渲染器(解碼器側)結合不同的表示，而後者使用DirAC域中不同音頻表示的編碼器側組合。

【發明內容】

【0040】 與DirAC框架的兼容性

【0041】 本實施例建立在專利申請號[WO2019068638]中提出的針對任意輸入類型的統一框架之上，並且類似於專利申請號[WO2020249815]對多聲道內容所做的工作，旨在消除無法有效應用DirAC參數(方向和擴散)到對象輸入的問題。事實上，根本不需要擴散參數，但發現每個時間/頻率單元的單個方向線索不足以再現高品質的對象內容。因此，本實施例提出在每個時間/頻率單元採用多個方向線索，並且因此引入在對象輸入的情況下代替典型DirAC參數的適應參數集。

【0042】 低位元率的彈性系統

【0043】 與DirAC相比，DirAC從聽者的角度使用基於場景的表示，而SAOC和SAOC-3D則是基於聲道和對象的內容而設計的，其中參數描述了聲道/對象之間的關係。為了對對象輸入使用基於場景的表示並因此與DirAC渲染器兼容，同時確保有效表示和高品質再現，需要一組經過調整的參數以允許信令多個方向線索。

【0044】 本實施例的一個重要目的是找到一種以低位元率和對越來越多的對象具有良好可擴展性的有效編碼對象輸入的方法。對每個對象信號進行離

散編碼不能提供這樣的可擴展性：每個增加的對象都會導致整體位元率的顯著上升。如果增加的對象數量超過允許的位元率，這將直接導致輸出信號的明顯衰減；這種衰減是支持本實施例的又一個論據。

【0045】 本發明的一個目的是提供一種改進的對多個音頻對象進行編碼或對編碼的音頻信號進行解碼的概念。

【0046】 本目的通過請求項1的編碼設備、請求項18的解碼器、請求項28的編碼方法、請求項29的解碼方法、請求項30的電腦程式或請求項31的編碼音頻信號來實現。

【0047】 在本發明的一實施態樣中，本發明基於以下發現：對於多個頻率柱中的一個以上之頻率柱，定義至少兩個相關音頻對象，並且與該至少兩個相關音頻對象相關的參數資料係包含在編碼器側並用於解碼器側以獲得高品質但高效的音頻編碼/解碼概念。

【0048】 根據本發明的另一實施態樣，本發明基於以下發現：執行適合於與每個對象相關聯的方向資訊的特定降混，使得具有關聯方向資訊的每個對象對整個對象有效，亦即，對於時間幀中的所有頻率柱，其用於將此對象降混到數個傳輸聲道中，例如，方向資訊的使用相當於將傳輸聲道生成為具有某些可調節特性的虛擬麥克風信號。

【0049】 在解碼器側，執行依賴於共變異數合成的特定合成，在特定實施例中，共變異數合成特別適用於不受去相關器引入的偽物影響的高品質共變異數合成。在其他實施例中，使用依賴於與標準共變異數合成相關的特定改良的進階共變異數合成，以便提高音頻品質及/或減少計算共變異數合成中使用的混合矩陣所需的計算量。

【0050】 然而，即使在更經典的合成中，音頻渲染是通過基於傳輸的選擇資訊顯式決定時間/頻率柱內的個別貢獻來完成的，音頻品質相對於習知技術的對象編碼方法或聲道降混方法而言是優越的。在這種情況下，每個時間/頻率柱都有一個對象標識資訊，並且在進行音頻渲染時，即在計算每個對象的方向貢獻時，使用該對象標識來查找與該對象資訊關聯的方向，以決定每個時間/頻率柱的各個輸出聲道的增益值。因此，當時間/頻率柱中只有一個相關對象時，則

根據對象ID和關聯對象的方向資訊的“碼本”，僅決定每個時間/頻率柱中該單個對象的增益值。

【0051】 然而，當時間/頻率柱中有超過1個相關對象時，則計算每個相關對象的增益值，以便將傳輸聲道的相應時間/頻率柱分配到相應的輸出聲道中，該輸出聲道係通過用戶提供的輸出格式，例如某個聲道格式是立體聲格式、5.1格式等。無論增益值是否用於共變異數合成的目的，即用於應用混合矩陣的目的將傳輸聲道混合到輸出聲道中，或者無論增益值是否用於通過將增益值乘以一個以上之傳輸聲道的相應時間/頻率柱來顯式決定時間/頻率柱中每個對象的單獨貢獻，且接著在相應的時間/頻率柱中總結每個輸出聲道的貢獻，其可能通過增加擴散信號分量來增強，然而，由於通過決定每個頻率柱的一個以上之相關對象而提供的靈活性，可以提高輸出音頻的品質。

【0052】 本決定操作是非常可行的，因為對於時間/頻率柱僅一個以上之對象ID必須與每個對象的方向資訊一起被編碼並傳輸到解碼器，然而這也是非常可行的，這是因為對於一個幀，所有頻率柱只有一個方向資訊。

【0053】 因此，無論是使用較佳增強共變異數合成還是使用每個對象的明顯傳輸聲道貢獻的組合來進行合成，都可獲得高效和高品質的對象降混，其係較佳地通過使用特定對象方向相關降混來改良，此降混依賴於降混權重，其係將傳輸聲道的生成反映為虛擬麥克風信號。

【0054】 與每個時間/頻率柱的兩個以上之相關對象相關的實施態樣可以較佳地與執行對象的特定方向相關降混到傳輸聲道中的實施態樣相結合。然而，這兩個實施態樣也可以彼此獨立地應用。此外，雖然在某些實施例中每個時間/頻率柱執行具有兩個以上之相關對象的共變異數合成，但是也可以通過僅傳輸每個時間/頻率柱的單個對象標識來執行進階共變異數合成和進階傳輸聲道到輸出聲道的昇混。

【0055】 此外，無論每個時間/頻率柱包括單個還是多個相關對象，也可以通過計算標準或增強共變異數合成中的混合矩陣來執行昇混，或者可以通過對時間/頻率柱的貢獻的單獨決定來執行昇混，該決定基於用於從方向“碼本”擷取特定方向資訊以決定對應貢獻的增益值的對象標識。在每個時間/頻率柱有兩

個以上之相關對象的情況下，接著將其加總以獲得每個時間/頻率柱的全部貢獻，然後，該加總步驟的輸出等效於混合矩陣應用的輸出，並且執行最終濾波器組處理以便為相應的輸出格式生成時域輸出聲道信號。

【圖式簡單說明】

【0056】 以下將結合附圖說明本發明的較佳實施例，其中：

圖1a是根據第一實施態樣之音頻編碼器的實施，其中每個時間/頻率柱具有至少兩個相關對象；

圖1b是根據第一實施態樣之編碼器的實施，其具有依賴於方向的對象降混；

圖2是根據第二實施態樣之編碼器的較佳實施；

圖3是根據第一實施態樣之編碼器的較佳實施；

圖4是根據第一及第二實施態樣之解碼器的較佳實施；

圖5是如圖4所示之共變異數合成處理的一較佳實施；

圖6a是根據第一實施態樣之解碼器的實施；

圖6b是根據第二實施態樣之解碼器；

圖7a是一流程圖，用於說明根據第一實施態樣之參數資訊的決定流程；

圖7b是參數資料的進一步決定流程的較佳實施；

圖8a顯示高解析度濾波器組時間/頻率表示；

圖8b顯示根據第一和第二實施態樣之較佳實施的幀J的相關輔助資訊的傳輸；

圖8c顯示一“方向碼本”，其係包含於編碼音頻信號中；

圖9a顯示根據第二實施態樣之較佳編碼方法；

圖9b顯示根據第二實施態樣之靜態降混的實施；

圖9c顯示根據第二實施態樣之動態降混的實施；

圖9d顯示第二實施態樣的另一個實施例；

圖10a是一流程圖，顯示第一實施態樣的解碼器側的較佳實施的流程圖；

圖10b顯示如圖10a所示之輸出聲道計算的較佳實施，其係根據具有每個輸出聲道的貢獻的加總和的實施例；

圖10c顯示根據第一實施態樣為多個相關對象決定功率值的較佳方法；

圖10d顯示如圖10a所示之輸出聲道的計算的實施例，其係使用依賴於混合矩陣的計算和應用的共變異數合成；

圖11顯示用於時間/頻率柱的混合矩陣的進階計算的幾個實施例；

圖12a顯示習知技術的DirAC編碼器；以及

圖12b顯示習知技術的DirAC解碼器。

【實施方式】

【0057】 圖1a顯示一種用於編碼多個音頻對象的設備，其係在輸入處接收音頻對象本身、及/或音頻對象的後設資料。編碼器包括一對象參數計算器100，其提供時間/頻率柱的至少兩個相關音頻對象的參數資料，並且該資料被轉發到輸出介面200。具體地，對象參數計算器針對與時間幀相關的多個頻率柱中的一個以上之頻率柱，計算至少兩個相關音頻對象的參數資料，其中，具體地，至少兩個相關音頻對象的數量小於多個音頻對象的總數，因此，對象參數計算器100實際上執行一選擇並且不是簡單地將所有對象指示為相關。在較佳實施例中，該選擇是通過相關性的方式來完成的，並且相關性是通過與幅度相關的度量來決定的，例如幅度、功率、響度或通過將幅度提高到與1不同的功率(較佳是大於1)而獲得的另一度量。然後，如果一定數量的相關對象可用於時間/頻率柱，則選擇具有最相關特徵的對象，即在所有對象中具有最高功率的對象，並且這些所選對象的資料是包含在參數資料中。

【0058】 輸出介面200被配置為輸出一編碼音頻信號，該編碼音頻信號包括關於一個以上之頻率柱的至少兩個相關音頻對象的參數資料的資訊。根據本實施，輸出介面可以接收其他資料並將其輸入到編碼音頻信號中，例如對象降混或表示對象降混的一個以上之傳輸聲道、或是在混合表示中的額外參數或對象波形資料，其中幾個對象是降混，或其他對象在單獨的表示中。在這種情況下，對象被直接導入或“複製”到相應的傳輸聲道中。

【0059】 圖1b顯示根據第二實施態樣的用於編碼多個音頻對象的設備的較佳實施，其中音頻對象與指示關於該多個音頻對象的方向資訊，即是對各對象分別提供一個方向資訊，或是若一組對象關聯至同一方向資訊時，對該組對象提供一個方向資訊。音頻對象被輸入到一降混器400，用於對多個音頻對象進行降混以獲得一個以上之傳輸聲道。此外，提供一傳輸聲道編碼器300，其對該一個以上之傳輸聲道進行編碼以獲得一個以上之編碼傳輸聲道，然後將其輸入到一輸出介面200，具體而言，降混器400連接到一對象方向資訊提供器110，其係在輸入處接收可以從中導出對象後設資料的任何資料，並輸出被降混器400實際使用的方向資訊。從對象方向資訊提供器110轉發到降混器400的方向資訊較佳地是一去量化的方向資訊，即是後續在解碼器側可用的相同方向資訊。為此，對象方向資訊提供器110被配置為導出或提取或擷取非量化對象後設資料，然後量化對象後設資料以導出表示一量化索引的量化對象後設資料，在較佳實施例中，該量化對象後設資料係在“其他資料”之中提供給如圖1b所示的輸出介面200。此外，對象方向資訊提供器110被配置為對量化的對象方向資訊進行去量化以獲得從方塊110轉發到降混器400的實際方向資訊。

【0060】 較佳地，輸出介面200被配置為額外地接收音頻對象的參數資料、對象波形資料、每個時間/頻率柱的單個或多個相關對象的一個以上之標識、以及如前所述的量化方向資料。

【0061】 接著，進一步說明其他實施例，其提出一種用於編碼音頻對象信號的參數化方法，該方法允許以低位元率進行有效傳輸，同時在消費者側進行高品質再現。基於考慮每個關鍵頻帶和時刻(時間/頻率磚)的一個方向線索的DirAC原理，為輸入信號的時間/頻率表示的每個這種時間/頻率磚決定一最主要對象。由於經證明這對於對象輸入是不夠的，因此為每個時間/頻率磚決定一個額外的第二主要對象，並基於這兩個對象，計算功率比以決定兩個對象中的每一個對所考慮的時間/頻率磚的影響。注意：為每個時間/頻率單元考慮兩個以上最主要對象也是可以想像的，尤其是對於越來越多的輸入對象，為簡單起見，以下描述主要基於每個時間/頻率單元的兩個主要對象。

【0062】 因此，傳輸到解碼器的參數輔助資訊包括：

- 為每個時間/頻率磚(或參數頻帶)的相關(主要)對象的子集進行計算的功率比。
- 表示每個時間/頻率磚(或參數頻)的相關對象的子集的對象索引。
- 與對象索引相關聯並為每個幀提供的方向資訊(其中每個時域幀包括多個參數頻帶。且每個參數頻帶包括多個時間/頻率磚)。

【0063】 通過與音頻對象信號相關聯的輸入後設資料檔案使方向資訊成為可用，例如，可以基於幀來指定後設資料。除輔助資訊之外，組合輸入對象信號的降混信號也被傳輸到解碼器。

【0064】 在渲染階段，傳輸的方向資訊(通過對象索引導出)用於將傳輸的降混信號(或更一般地說：傳輸聲道)平移到適當的方向，降混信號根據傳輸的功率比分配到兩個相關的對象方向，其係用作為加權因子。對解碼的降混信號的時間/頻率表示的每個時間/頻率磚進行上述處理。

【0065】 本章節概述了編碼器側的處理，然後是參數和降混計算的詳細說明。音頻編碼器接收一個以上之音頻對象信號，每個音頻對象信號係相關聯到描述對象屬性的後設資料檔案。在本實施例中，關聯後設資料檔案中描述的對象屬性對應於以幀為基礎提供的方向資訊，其中一幀對應20毫秒。每個幀都由一個幀編號標識，該編號也包含在後設資料檔案中。方向資訊以方位角和仰角資訊的形式給出，其中方位角的值取自(-180, 180]度，仰角的值取自[-90, 90]度，後設資料中提供的其他屬性可能包括距離、展開、增益；在本實施例中不考慮這些特性。

【0066】 後設資料檔案中提供的資訊與實際音頻對象檔案一起使用以創建一組參數，該組參數傳輸到解碼器並用於渲染最終音頻輸出檔案。更具體地說，編碼器估算每個給定時間/頻率磚的主要對象子集的參數，即功率比，主要對象的子集由對象索引表示，這些索引也用於識別對象方向，這些參數與傳輸聲道和方向後設資料一起傳輸到解碼器。

【0067】 圖2顯示編碼器的概略圖，其中傳輸聲道包括從輸入對象檔案和輸入後設資料中提供的方向資訊計算出的降混信號，傳輸聲道的數量總是小於輸入對象檔案的數量。在一實施例的編碼器中，編碼音頻信號由編碼傳輸聲道

表示，且編碼參數輔助資訊由編碼對象索引、編碼功率比和編碼方向資訊指示。編碼傳輸聲道和編碼參數輔助資訊一起形成由一多工器220輸出的位元流。特別地，編碼器包括接收輸入對象音頻檔案的濾波器組102。此外，對象後設資料檔案被提供給一提取方向資訊方塊110a，方塊110a的輸出被輸入到量化方向資訊方塊110b，其係將方向資訊輸出到執行降混計算的降混器400。此外，量化的方向資訊(即量化索引)從方塊110b轉發到編碼方向資訊方塊202，其較佳地執行某種熵編碼以便進一步降低所需的位元率。

【0068】 此外，濾波器組102的輸出被輸入到信號功率計算方塊104中，而信號功率計算方塊104的輸出被輸入到對象選擇方塊106中，且另外被輸入到功率比計算方塊108中，功率比計算方塊108還連接到對象選擇方塊106，以便計算功率比，即僅所選對象的組合值。在方塊210中，其係對計算出的功率比或組合值進行量化和編碼。正如稍後將概述的，功率比是較佳的，以便節省一個功率資料項目的傳輸。然而，在不需要這種節省的其他實施例中，可以在對象選擇器106的選擇下將實際信號功率或由方塊104決定的信號功率導出的其他值，輸入到量化器和編碼器中，而不是功率比。然後，不需要功率比計算108，且對象選擇106確保僅相關參數資料(即相關對象的功率相關資料)被輸入到方塊210中，以用於量化和編碼的目的。

【0069】 比較圖1a和圖2，圖1a的對象參數計算器100較佳地包括方塊102、104、110a、110b、106、108，且圖1a的輸出介面方塊200較佳地包括方塊202、210、220。

【0070】 此外，圖2中的核心編碼器300對應於圖1b的傳輸聲道編碼器300，降混計算方塊400對應於圖1b的降混器400，且圖1b的對象方向資訊提供器110對應於圖2的方塊110a、110b。此外，圖1b的輸出介面200較佳地以與圖1a的輸出介面200相同的方式實現，且其包括圖2的方塊202、210、220。

【0071】 圖3顯示一種編碼器之變化例，其中降混計算是可選的並且不依賴於輸入後設資料。在這個變化例中，輸入音頻檔案可以直接饋送到核心編碼器，核心編碼器從輸入音頻檔案創建傳輸聲道，因此傳輸聲道的數量對應於輸

入對象檔案的數量；如果輸入對象的數量為1或2，這種情況特別有趣。對於更多數量的對象，仍將使用降混信號來減少要傳輸的資料量。

【0072】 如圖3所示，其中與圖2所示的相似的參考符號表示相似的功能，這不僅對圖2和圖3成立，而且對本說明書中描述的所有其他圖式同樣成立。與圖2不同，圖3在沒有任何方向資訊的情況下執行降混計算400，因此，降混計算可以是例如使用預先已知的降混矩陣的靜態降混，或者可以是不依賴於與包括在輸入對象音頻檔案中的對象相關聯的任何方向資訊的能量相關的降混。然而，方向資訊在方塊110a中被提取，並在方塊110b中被量化，而且量化的值被轉發到方向資訊編碼器202，以便在編碼音頻信號中具有編碼方向資訊，例如二進制編碼音頻信號形成的位元流。

【0073】 在輸入音頻對象檔案的數量不是太多的情況下、或者在具有足夠的可用傳輸頻寬的情況下，還可以省去降混計算方塊400，使得輸入音頻對象檔案直接表示核心編碼器進行編碼的傳輸聲道。在這種實施中，方塊104、104、106、108、210也不是必需的。然而，較佳實施會導致一混合實施，其中一些對象被直接導入傳輸聲道，而其他對象被降混到一個以上之傳輸聲道。在這種情況下，為了生成在編碼傳輸聲道內直接具有一個以上之對象以及由圖2或圖3中的任一者的降混器400生成的一個以上之傳輸聲道的位元流，則需要圖3中所示的所有方塊。

【0074】 參數計算

【0075】 時域音頻信號(包括所有輸入對象信號)使用濾波器組轉換到時域/頻域，例如：複雜低延遲濾波器組(*complex low-delay filterbank*, CLDFB)分析濾波器將20毫秒的幀(對應於在48 kHz採樣率下的960個樣本)轉換為大小為16x60的時間/頻率磚，其具有16個時隙和60個頻段。對於每個時間/頻率單位，瞬時信號功率計算如下

$$P_i(k,n) = |X_i(k,n)|^2,$$

【0076】 其中，k表示頻帶索引，n表示時隙索引，i表示對象索引。由於就最終位元率而言，每個時間/頻率磚的傳輸參數的耗費非常大，因此採用分組

的方式以便計算減少數量的時間/頻率磚的參數，例如：16個時隙可以組合為一個時隙，60個頻段可以根據心理聲學標度分為11個頻段，此方式將16x60的初始尺寸減少到1x11，其對應於11個所謂的參數帶。瞬時信號功率值根據分組求和，得到降維後的信號功率：

$$P_i(l, m) = \sum_{k=0}^T \sum_{n=B_S}^{n=B_E} P_i(k, n),$$

其中，T在本例中對應為15， B_S 和 B_E 定義參數帶邊界。

【0077】 為了決定要為其計算參數的最主要對象的子集，所有N個輸入音頻對象的瞬時信號功率值按降序排序。在本實施例中，我們決定兩個最主要對象，並將範圍從0到N-1的相應對象索引儲存為要傳輸的參數的一部分。此外，計算將兩個主要對象信號相互關聯的功率比：

$$PR_1(l, m) = \frac{P_1(l, m)}{(P_1(l, m) + P_2(l, m))}$$

$$PR_2(l, m) = \frac{P_2(l, m)}{(P_1(l, m) + P_2(l, m))} = 1 - PR_1(l, m)$$

【0078】 或者在不限於兩個對象的更一般的表達式中：

$$PR_i(l, m) = \frac{P_i(l, m)}{\sum_{j=1}^S P_j(l, m)}$$

其中，在本文中，S表示要考慮的主要對象的數量，並且：

$$\sum_{i=1}^S PR_i(l, m) = 1$$

【0079】 在兩個主要對象的情況下，兩個對象中的每一個對象的功率比為0.5，其意味著兩個對象在相應的參數帶內同等存在，而功率比為1和0表示兩個對象其中之一不存在。這些功率比儲存為要傳輸的參數的第二部分。由於功率比之和為1，因此傳輸S-1的值就足以取代S。

【0080】 除了每個參數帶的對象索引和功率比的值之外，還必須傳輸從輸入後設資料檔案中提取的每個對象的方向資訊。由於資訊最初是在幀的基礎上提供的，因此對每一幀都進行了處理(其中，在上述示例中，每一幀包括11個參數帶或總共16x60個時間/頻率磚)，因此，對象索引間接表示對象方向。注意：由於功率比之和為1，每個參數帶傳輸的功率比的數量可以減1；例如：在考慮2個相關對象的情況下，傳輸1個功率比的值就足夠了。

【0081】 方向資訊和功率比的值都被量化並與對象索引組合以形成參數輔助資訊，然後將此參數輔助資訊編碼，並與編碼的傳輸聲道/降混信號一起混合到最終的位元流表示中。例如，通過使用每個值3位元對功率比進行量化，可以實現輸出品質和消耗的位元率之間的良好權衡。在一實際示例中，方向資訊可以以5度的角解析度提供，並且隨後對每個方位角的值以7位元進行量化、並對每個仰角的值以6位元進行量化。

【0082】 降混計算

【0083】 所有輸入音頻對象信號被組合成包括一個以上之傳輸聲道的一降混信號，其中傳輸聲道的數量小於輸入對象信號的數量。注意：在本實施例中，僅當只有一個輸入對象時才會出現單個傳輸聲道，這意味著跳過降混計算。

【0084】 如果降混包括兩個傳輸聲道，則該立體聲降混可以例如被計算為一虛擬心形麥克風信號，虛擬心形麥克風信號是通過應用後設資料檔案中為每一幀提供的方向資訊來決定的(在此假設所有的仰角值都為零)：

$$w_L = 0.5 + 0.5 * \cos(\text{azimuth} - \pi/2)$$

$$w_R = 0.5 + 0.5 * \cos(\text{azimuth} + \pi/2)$$

【0085】 其中，虛擬心形位於90°和-90°，兩個傳輸聲道(左和右)中的每一個的個別權重因此被決定並應用於相應的音頻對象信號：

$$DMX_L = \sum_{j \in N} x_j * w_L$$

$$DMX_R = \sum_{j \in N} x_j * w_R$$

【0086】 在本實施例中，N是輸入對象的數量，其係大於或等於2。如果為每一幀更新虛擬心形權重，則採用適應方向資訊的動態降混。另一種可能方式是採用固定降混，其係假設每個對象都位於靜態位置，例如，該靜態位置可以對應於對象的初始方向，接著導致靜態虛擬心形權重，其對於所有幀都相同。

【0087】 如果目標比特率允許，可以想像多於兩個的傳輸信道。在三個傳輸通道的情況下，心形指向可以均勻排列，例如，在0°、120°和-120°。如果使用四個傳輸通道，則第四個心形指向上方或四個心形可以再次以均勻的方式水平佈置。如果對象位置例如僅是一個半球的一部分，則該佈置也可以針對對象位置進行調整。產生的下混信號由核心編碼器處理，並與編碼的參數輔助信息一起轉化為比特流表示。

【0088】 或者，輸入對象信號可以被饋送到核心編碼器而不被組合成降混信號。在這種情況下，產生的傳輸聲道的數量對應於輸入對象信號的數量。通常而言，會給出與總位元率相關的最大傳輸聲道數量，然後僅當輸入對象信號的數量超過傳輸聲道的最大數量時才會採用降混信號。

【0089】 圖6a顯示用於解碼一編碼音頻信號(如圖1a、圖2或圖3的輸出信號)的解碼器，該信號包括用於多個音頻對象的一個以上之傳輸聲道和方向資訊。此外，編碼音頻信號包括針對時間幀的一個以上之頻率柱的至少兩個相關音頻對象的參數資料，其中至少兩個相關對象的數量低於多個音頻對象的總數。特別地，解碼器包括一輸入介面，用於以在時間幀中具有多個頻率柱的頻譜表示提供一個以上之傳輸聲道，這表示信號從輸入介面方塊600轉發到音頻渲染器方塊700。特別地，音頻渲染器700被配置用於使用包括在編碼音頻信號中的方向資訊，將一個以上之傳輸聲道渲染成多個音頻聲道，音頻聲道的數量較佳是立體聲輸出格式的兩個聲道，或者具更高數量之輸出格式的兩個以上的聲道，例如3聲道、5聲道、5.1聲道等。特別地，音頻渲染器700被配置為針對該一個以上之頻率柱中的每一個，根據與至少兩個相關音頻對象中的一第一相關

音頻對象相關聯的第一方向資訊和根據與至少兩個相關音頻對象中的一第二相關音頻對象相關聯的第二方向資訊，計算來自一個以上之傳輸聲道的貢獻。特別地，多個音頻對象的方向資訊包括與第一對象相關聯的第一方向資訊和與第二對象相關聯的第二方向資訊。

【0090】 圖8b顯示一幀的參數資料，在一較佳實施例中，其包括多個音頻對象的方向資訊810、以及另外由方塊812表示的特定數量的參數帶中的每一個的功率比、以及較佳地由方塊814表示的每個參數帶的兩個以上的對象索引。特別地，在圖8c中更詳細地顯示多個音頻對象的方向資訊810。圖8c顯示一表格，其第一列具有從1到N的某個對象ID，其中N是多個音頻對象的數量，此外，表格的第二列具有每個對象的方向資訊，其係較佳為方位角值和仰角值，或者在二維情況下，僅具有方位角值，這顯示於欄位818處。因此，圖8c顯示包括在輸入到圖6a的輸入介面600的編碼音頻信號中的“方向碼本”。來自欄位818的方向資訊與來自欄位816的某個對象ID具有唯一相關聯，並且對一幀中的“整個”對象皆有效，即對一幀中的所有頻帶皆有效。因此，不管頻率柱的數量是高解析度表示中的時間/頻率磚、還是較低解析度表示中的時間/參數帶，對於每個對象標識，只有單個方向資訊將被輸入介面傳輸和使用。

【0091】 在本實施例中，圖8a顯示由圖2或圖3的濾波器組102生成的時間/頻率表示，其中該濾波器組被實現為之前討論的複合低延遲濾波器組 (CLDFB)。對於如前面關於圖8b和8c所討論的方式所獲得的方向資訊的幀，濾波器組生成如圖8a所示之從0到15的16個時隙和從0到59的60個頻帶，因此，一個時隙和一個頻帶表示一個時間/頻率磚802或804。然而，為了降低輔助資訊的位元率，較佳將高解析度表示轉換為如圖8b所示的低解析度表示，如圖8b中的欄位812所示，其中僅存在單個時間柱、並且其中60個頻帶被轉換為11個參數頻帶。因此，如圖10c所示，高解析度表示由時隙索引n和頻帶索引k指示，而低解析度表示由分組的時隙索引m和參數頻帶索引l給出。然而，在本說明書中，時間/頻率柱可以包括圖8a所示的高解析度時間/頻率磚802、804，或由在圖10c中的方塊731c的輸入處的分組的時隙索引和參數頻帶索引標識的低解析度時間/頻率單元。

【0092】 在如圖6a所示的實施例中，音頻渲染器700被配置為對於一個以上之頻率柱中的每一個，從根據與至少兩個相關音頻對象中的一第一相關音頻對象相關聯的第一方向資訊並且根據與至少兩個相關音頻對象中的一第二相關音頻對象相關聯的第二方向資訊的一個以上之傳輸聲道中，計算一貢獻。在如圖8b所示的實施例中，方塊814具有參數帶中每個相關對象的對象索引，即具有兩個以上之對象索引，使得每個時間頻率柱存在兩個貢獻。

【0093】 以下將參考圖10a進行說明，貢獻的計算可以通過混合矩陣間接完成，其中每個相關對象的增益值被決定並用於計算混合矩陣。或者，如圖10b所示，可以使用增益值再次顯式計算貢獻，然後在特定時間/頻率柱中按每個輸出聲道對顯式計算的貢獻求和。因此，無論貢獻是顯式計算還是隱式計算所得，音頻渲染器仍然使用方向資訊將一個以上之傳輸聲道渲染成數個音頻聲道，從而對於一個以上之頻率柱中的每一個，根據與至少兩個相關音頻對象中的第一相關音頻對象相關聯的第一方向資訊以及根據與至少兩個相關音頻對象中的第二相關音頻對象相關聯的第二方向資訊，將來自一個以上之傳輸聲道的貢獻包含在該數個音頻聲道中。

【0094】 圖6b顯示一種用於解碼一編碼音頻信號的解碼器的第二實施態樣，該編碼音頻信號包括多個音頻對象的一個以上之傳輸聲道和方向資訊、以及一時間幀的一個以上之頻率柱的音頻對象的參數資料。同樣地，解碼器包括接收編碼音頻信號的一輸入介面600，並且解碼器包括一音頻渲染器700，用於使用方向資訊將一個以上之傳輸聲道渲染成數個音頻聲道。特別地，音頻渲染器被配置為根據多個頻率柱中的每個頻率柱的一個以上之音頻對象、以及與頻率柱中的相關之一個以上之音頻對象相關聯的方向資訊，計算出一直接響應資訊。該直接響應資訊較佳包括用於一共變異數合成或一進階共變異數合成、或用於從一個以上之傳輸聲道的貢獻的顯式計算的增益值。

【0095】 較佳地，音頻渲染器被配置為使用時間/頻帶中的一個以上之相關音頻對象的直接響應資訊、並使用數個音頻聲道的資訊來計算一共變異數合成資訊。此外，共變異數合成信息(較佳是混合矩陣)被應用於一個以上之傳輸聲道以獲得數個音頻聲道。在另一實施方式中，直接響應資訊是每一個音頻對

象的直接響應向量，共變異數合成資訊是共變異數合成矩陣，並且音頻渲染器被配置為在應用共變異數合成資訊時按頻率柱執行一矩陣運算。

【0096】 此外，音頻渲染器700被配置為在直接響應資訊的計算中導出一個以上之音頻對象的一直接響應向量，並為一個以上之音頻對象計算來自各該直接響應向量的一共變異數矩陣。此外，在共變異數合成資訊的計算中，計算一目標共變異數矩陣。然而，不是使用目標共變異數矩陣，而是使用目標共變異數矩陣的相關資訊，即一個以上之最主要對象的直接響應矩陣或向量，以及由功率比的應用所決定的直接功率的對角矩陣(表示為E)。

【0097】 因此，目標共變異數資訊不一定是一顯式目標共變異數矩陣，而是從一個音頻對象的共變異數矩陣或一時間/頻率柱中更多音頻對象的共變異數矩陣中導出，從時間/頻率柱中的相應的一個或多個音頻對象的功率資訊中導出，以及從用於一個以上之時間/頻率柱的一個或多個傳輸聲道中導出的功率資訊中導出。

【0098】 位元流表示由解碼器讀取，並且編碼傳輸聲道和包含在其中的編碼參數輔助資訊可用於進一步處理。參數輔助資訊包括：

- 如量化方位角和仰角值的方向資訊(對於每一幀)
- 表示相關對象之子集的對象索引(對於每個參數帶)
- 將相關對象相互關聯的量化功率比(對於每個參數帶)

【0099】 所有處理均以逐幀方式完成，其中每一幀包含一個或多個子幀，例如，一個幀可以由四個子幀組成，在這種情況下，一個子幀的持續時間為5毫秒。圖4顯示解碼器的簡單概略圖。

【0100】 圖4顯示實現第一和第二實施態樣的音頻解碼器。如圖6a和圖6b所示的輸入介面600包括一解多功器602、一核心解碼器604、用於解碼對象索引的一解碼器608、用於解碼和去量化功率比的一解碼器610、以及用於解碼和去量化的方向資訊的一解碼器612。此外，輸入介面包括一濾波器組606，用於提供時間/頻率表示中的傳輸聲道。

【0101】 音頻渲染器700包括一直接響應計算器704、由例如一使用者介面接收的輸出配置所控制的一原型矩陣提供器702、一共變異數合成方塊706、

以及一合成濾波器組708，以便最終提供一輸出音頻檔案，其包含聲道輸出格式的數個音頻聲道。

【0102】 因此，項目602、604、606、608、610、612較佳包括在如圖6a和圖6b所示的輸入介面中，並且圖4所示的項目702、704、706、708是如圖6a或圖6b所示的音頻渲染器(以參考符號700表示)的一部分。

【0103】 編碼的參數輔助資訊被解碼，並且重新獲得量化的功率比值、量化的方位角和仰角值(方向資訊)以及對象索引。未傳輸的一個功率比值是通過利用所有功率比值總和為1的事實來獲得的，其解析度 (l, m) 對應於在編碼器側採用的時間/頻率磚分組。在使用更精細的時間/頻率解析度 (k, n) 的進一步處理步驟期間，參數帶的參數對於包含在該參數帶中的所有時間/頻率磚有效，其對應於一擴展處理使得 $(l, m) \rightarrow (k, n)$ 。

【0104】 編碼傳輸聲道由核心解碼器解碼，使用濾波器組(與編碼器中使用的濾波器組匹配)，因此得到的解碼音頻信號的每一幀都被轉換為時間/頻率表示，其解析度通常更精細於(但至少等於)用於參數輔助資訊的解析度。

【0105】 輸出信號渲染/合成

【0106】 以下描述適用於一幀的音頻信號； T 表示轉置運算符：

【0107】 使用解碼傳輸聲道 $x = x(k, n) = [X_1(k, n), X_2(k, n)]^T$ ，即是時頻表示的音頻信號(在這種情況下包括兩個傳輸聲道)和參數輔助資訊，推導出每個子幀(或降低計算複雜度的幀)的混合矩陣 M 來合成時頻輸出信號 $y = y(k, n) = [Y_1(k, n), Y_2(k, n), Y_3(k, n), \dots]^T$ ，其包含數個輸出聲道(例如5.1、7.1、7.1+4等)：

【0108】 •對於所有(輸入)對象，使用傳輸對象方向，決定所謂的直接響應值，其描述要用於輸出聲道的平移增益。這些直接響應值特定於目標佈局，即揚聲器的數量和位置(提供作為輸出配置的一部分)。平移方法的示例包括向量基幅度平移 (VBAP) [Pulkki1997]和邊緣衰落幅度平移(EFAP) [Borß2014]，每個對象都有一個與其相關聯的直接響應值 dr_i (包含與揚聲器一樣多的元素)的向量，這些向量每幀計算一次。注意：如果對象位置對應於揚聲器位置，則向量包含

該揚聲器的值為1，所有其他值均為0；如果對象位於兩個(或三個)揚聲器之間，則對應的非零向量元素數為2(或3)。

【0109】 • 實際合成步驟(在本實施例中共變異數合成 [Vilkamo2013])包括以下子步驟(參見圖5所示)：

- o 對於每個參數帶，對象索引(描述分組到該參數帶的時間/頻率磚內的輸入對象中的主要對象的子集)用於提取進一步處理所需的向量 $\mathbf{dr}_i$ 的子集，例如，由於只考慮 2 個相關對象，因此需要與這 2 個相關對象相關聯的 2 個向量 $\mathbf{dr}_i$ 。
- o 接著，為每個相關對象從直接響應值 $\mathbf{dr}_i$ 計算大小為輸出聲道×輸出聲道的共變異數矩陣 C_i ：

$$C_i = \mathbf{dr}_i * \mathbf{dr}_i^T$$

- o 對於每個時間/頻率磚(在參數帶內)，決定音頻信號功率 $P(k, n)$ ，在兩個傳輸聲道的情況下，第一個聲道的信號功率係加到第二個聲道的信號功率；對於該信號功率，每個功率比值都相乘，因此為每個相關/主要對象 i 產生一個直接功率值：

$$DP_i(k, n) = PR_i(k, n) * P(k, n)$$

- o 對於每個頻帶 k ，通過對(子)幀內的所有時隙 n 求和以及對所有相關對象求和，來獲得大小為輸出聲道×輸出聲道的最終目標共變異數矩陣 C_Y ：

$$C_Y = \sum_n \sum_i DP_i(k, n) C_i$$

【0110】 圖5顯示在如圖4所示之方塊706中執行的共變異數合成步驟的詳細概述。特別地，圖5所示的實施例包括一信號功率計算方塊721、一直接功率計算方塊722、一共變異數矩陣計算方塊723、一目標共變異數矩陣計算方塊724、一輸入共變異數矩陣計算方塊726、一混合矩陣計算方塊725和一渲染方塊727，如圖5所示，渲染方塊727另外包括圖4所示之濾波器組方塊708，使得

方塊727的輸出信號較佳對應於時域輸出信號。然而，當方塊708不包括在圖5的渲染方塊中，則結果會是對應音頻聲道的譜域表示。

【0111】（以下步驟是習知技術[Vilkamo2013]的一部分，添加於此以為釐清。）

- o 對於每個(子)幀和每個頻帶，從解碼音頻信號計算大小為傳輸聲道 $\times$ 傳輸聲道的一輸入共變異數矩陣 $C_x = xx^T$ 。可選地，可以僅使用主對角線的條目，在這種情況下，其他非零條目被設置為零。
- o 定義了大小為輸出聲道 $\times$ 輸出聲道的原型矩陣，其描述了傳輸聲道到輸出聲道(提供作為輸出配置的一部分)的映射，其數量由目標輸出格式(例如，目標揚聲器佈局)給出。這個原型矩陣可以是靜態的，也可以是逐幀變化的。示例：如果僅傳輸單個傳輸聲道，則該傳輸聲道映射到每個輸出聲道；如果傳輸兩個傳輸聲道，則左(第一)聲道被映射到位於 $(+0^\circ, +180^\circ)$ 範圍內的所有輸出聲道，即“左”聲道，右(第二)聲道對應地映射到位於 $(-0^\circ, -180^\circ)$ 範圍內的所有輸出聲道，即“右”聲道。(注意： 0° 表示聽者前方的位置，正角表示聽者左側的位置，負角表示聽者右側的位置，如果採用不同的規定，則角度的符號需要進行相應調整。)
- o 使用輸入共變異數矩陣 C_x 、目標共變異數矩陣 C_r 和原型矩陣，計算每個(子)幀和每個頻帶的混合矩陣[Vilkamo2013]，例如，可以對每個(子)幀得到 60 個混合矩陣。
- o 混合矩陣在(子)幀之間(例如線性地)內插，對應於時間平滑。
- o 最後，輸出聲道 y 係以逐頻段合成，其通過將最終的混合矩陣 M (每個都是大小為輸出聲道 $\times$ 傳輸聲道)的集合，乘以解碼傳輸聲道 x 的時間/頻率表示的相應頻段：

$$y = Mx$$

【0112】 請注意，我們沒有使用[Vilkamo2013]中描述的殘差信號 r 。

【0113】 使用濾波器組將輸出信號 y 轉換回時域表示 $y(t)$ 。

【0114】 優化共變異數合成

【0115】 由於本實施例所示的如何計算輸入共變異數矩陣 C_x 和目標共變異數矩陣 C_y ，可以達成[Vilkamo2013]所揭露之共變異數合成的最優混合矩陣計算的某些優化，這導致混合矩陣計算的計算複雜度的顯著降低。請注意，在本節中，阿達馬運算子(Hadamard operator)“ $\circ$ ”表示對矩陣進行逐元素運算，即不遵循如矩陣乘法等規則，而是逐個元素進行相應運算。該運算子表示相應的運算不是對整個矩陣進行，而是對每個元素分別進行，例如，矩陣A和矩陣B的相乘不對應於矩陣乘法 $AB=C$ ，而是對應於逐元素運算 $a_{ij} \times b_{ij} = c_{ij}$ 。

【0116】 SVD(.)表示奇異值分解，[Vilkamo2013]中作為Matlab函數(列表1)呈現的演算法如下(習知技術)：

輸入：大小為 $m \times m$ 的矩陣 C_x ，包括輸入信號的共變異數

輸入：大小為 $n \times n$ 的矩陣 C_y ，包括輸入信號的目標共變異數

輸入：大小為 $n \times m$ 的矩陣 Q ，原型矩陣

輸入：標量 α ， S_x 的正則化因子([Vilkamo2013]建議 $\alpha=0.2$)

輸入：標量 β ， G_y 的正則化因子([Vilkamo2013]建議 $\beta=0.001$)

輸入：布林值a，表示是否應執行能量補償來取代計算殘量共變異數 C_r

輸出：大小為 $n \times m$ 的矩陣M，最佳混合矩陣

輸出：大小為 $n \times n$ 的矩陣 C_r ，包含殘量共變異數

```

%  $C_y$ 的分解([Vilkamo2013]，方程式(3))
1  $U_{C_y}, S_{C_y}, V_{C_y} \leftarrow \text{SVD}(C_y)$  %  $n \times n$ 矩陣的 SVD
2  $K_y \leftarrow U_{C_y} S_{C_y}^{\circ 1/2}$ 
%  $C_x$ 的分解([Vilkamo2013]，方程式(3))
3  $U_{C_x}, S_{C_x}, V_{C_x} \leftarrow \text{SVD}(C_x)$  %  $m \times m$ 矩陣的 SVD
4  $K_x \leftarrow U_{C_x} S_{C_x}^{\circ 1/2}$ 
%  $K_x$ 的SVD ([Vilkamo2013]，章節3.2，
 $V_x = I$ )
5  $U_x \leftarrow U_{C_x}$ 

```

```

6   $S_x \leftarrow S_{C_x}^{\circ 1/2}$ 
   %  $S_x$ 的正則化([Vilkamo2013]，章節3.2，
    $V_x = I$ )
7   $s_{x_{\max}} \leftarrow \max(\text{diag}(S_x))$ 
8   $s'_{x_{i,i}} \leftarrow \max(s_{x_{i,i}}, \alpha s_{x_{\max}}), i = 1, \dots, m$ 
   % 制定正則化 $K_x'^{-1}$ ([Vilkamo2013]，章節 3.2， $V_x = I$ )
9   $K_x'^{-1} \leftarrow S_x'^{\circ -1} U_x^H$ 
   % 制定正則化矩陣 $G_{\hat{v}}$ 
10  $C_{\hat{v}} \leftarrow Q C_x Q^H$  [Vilkamo2013]，方程式(5)、
   (2)
   % 以下正則化步驟僅在[Vilkamo2013]的列表1中找到，但未在本文中說明
11  $c_{\hat{v}_{\max}} \leftarrow \max(\text{diag}(C_{\hat{v}}))$ 
12  $c'_{\hat{v}_{i,i}} \leftarrow \max(c_{\hat{v}_{i,i}}, \beta c_{\hat{v}_{\max}}), i = 1, \dots, n$ 
13  $g_{\hat{v}_{i,i}} \leftarrow \sqrt{\frac{c_{y_{i,i}}}{c'_{\hat{v}_{i,i}}}}, i = 1, \dots, n$  % [Vilkamo2013]，方程式(7)
   % 制定最優 $P$  ([Vilkamo2013]，章節 3.1)
14  $U, S, V \leftarrow \text{SVD}(K_x^H Q^H G_{\hat{v}}^H K_v)$  %  $m \times n$ 矩陣的 SVD
   % [Vilkamo2013]，方程式(3)，" $\Lambda$ "是來自[Vilkamo2013]章節 3.3 的擴展單位矩陣
15  $P \leftarrow V \Lambda U^H$ 
   % 制定最優 $M_{\text{opt}}$ 
16  $M_{\text{opt}} \leftarrow K_v P K_x'^{-1}$  % [Vilkamo2013]，方程式
   (11)
   % 制定殘量共變異數矩陣 $C_r$ 
17  $C_{\hat{v}} \leftarrow M_{\text{opt}} C_x M_{\text{opt}}^H$ 
18  $C_r \leftarrow C_Y - C_{\hat{v}}$  % [Vilkamo2013]，方程式
   (15)

```

```
% 能量補償
19 if  $\alpha$  then
     $g_{i,i} \leftarrow \sqrt{\frac{c_{y_{i,i}}}{c_{\hat{y}_{i,i}}}}, i = 1, \dots, n$  % [Vilkamo2013]，方程式
    % [Vilkamo2013]，方程式
     $M \leftarrow M_{opt} G$  % [Vilkamo2013]，方程式
    % [Vilkamo2013]，方程式
21
22 else
     $M \leftarrow M_{opt}$ 
23
```

【0117】 如上一節所述，只有 C_x 的主對角元素是可選的，所有其他條目都設置為零。在這種情況下， C_x 是一個對角矩陣和一個滿足[Vilkamo2013]的方程式(3)的有效分解，其是

$$K_x = C_x^{\circ 1/2}$$

且不再需要來自習知技術之演算法的第3行的SVD。

【0118】 考慮從上一節中的直接響應 dr_i 和直接功率(或直接能量)生成目標共變異數的公式

$$C_i = dr_i * dr_i^T$$

$$DP_i(k,n) = PR_i(k,n) * P(k,n)$$

$$C_Y = \sum_n \sum_i DP_i(k,n) C_i ,$$

【0119】 最後一個公式可以重新排列並寫成

$$C_Y = \sum_i C_i \sum_n DP_i(k,n) .$$

【0120】 如果現在定義

$$E_i = \sum_n DP_i(k,n)$$

則可以得到

$$C_Y = \sum_i C_i E_i ,$$

【0121】 可以很容易得知，如果將直接響應排列在用於k個最主要對象的一個直接響應矩陣 $R = [dr_1 \cdots dr_k]$ 中，並創建一個直接功率的對角矩陣，如E，其中 $e_{i,i} = E_i$ ，而 C_Y 也可以表示為

$$C_Y = RER^H$$

【0122】 並且滿足[Vilkamo2013]的方程式(3)的 C_Y 的有效分解，如由下式：

$$C_Y = RE^{\circ 1/2}$$

【0123】 因此，不再需要來自習知技術之演算法的第1行的SVD。

【0124】 這可以導出本實施例中的共變異數合成的優化算法，其還考慮到一直被使用的能量補償選項，因此不需要殘差目標共變異數 C_r ：

- 輸入: 大小為 $m \times m$ 的對角矩陣 C_x ，包括具m個聲道的輸入信號的共變異數
- 輸入: 大小為 $n \times k$ 的矩陣 R ，包括對k個主要對象的直接響應
- 輸入: 對角矩陣E，包括對主要對象的目標功率
- 輸入: 大小為 $n \times m$ 的矩陣 Q ，原型矩陣
- 輸入: 標量 α ， S_x 的正則化因子([Vilkamo2013]建議 $\alpha=0.2$)
- 輸入: 標量 β ， $G_{\hat{x}}$ 的正則化因子([Vilkamo2013]建議 $\beta=0.001$)
- 輸出: 大小為 $n \times m$ 的矩陣M，最佳混合矩陣

```
% C_Y的分解(進步性)
1  K_Y ← RE^{\circ 1/2}
   % C_x的分解(進步性)
2  K_x ← C_x^{\circ 1/2}
```

% S_x 的正則化，(進步性， K_x 是一對角矩陣，所以本步驟可以簡化)

3 $S_x \leftarrow K_x$

4 $s_{x_{\max}} \leftarrow \max(\text{diag}(S_x))$

5 $s'_{x_{i,i}} \leftarrow \max(s_{x_{i,i}}, \alpha s_{x_{\max}}), i = 1, \dots, m$

% 制定正則化 $K_x'^{-1}$ (進步性，亦可簡化)

6 $K_x'^{-1} \leftarrow S_x'^{-1}$

% 制定正則化矩陣 $G_{\hat{y}}$

11 $C_{\hat{y}} \leftarrow Q C_x Q^H$ ilkamo2013]，方程式(5)、(2)

12 $C_y = R E R^H$

% 以下正則化步驟僅在[Vilkamo2013]的列表1中找到，但未在本文中說明

12 $c_{\hat{y}_{\max}} \leftarrow \max(\text{diag}(C_{\hat{y}}))$

13 $c'_{\hat{y}_{i,i}} \leftarrow \max(c_{\hat{y}_{i,i}}, \beta c_{\hat{y}_{\max}}), i = 1, \dots, n$

13 $g_{\hat{y}_{i,i}} \leftarrow \sqrt{\frac{c_{y_{i,i}}}{c'_{\hat{y}_{i,i}}}}, i = 1, \dots, n$ % [Vilkamo2013]，方程式(7)

% 制定最優 P ([Vilkamo2013]，章節 3.1)

14 $U, S, V \leftarrow \text{SVD}(K_x^H Q^H G_{\hat{y}}^H K_y)$ % $m \times n$ 矩陣的 SVD

% 在此不需使用 Λ (進步性)

15 $P \leftarrow V U^H$

% 制定最優 M_{opt}

16 $M_{\text{opt}} \leftarrow K_y P K_x'^{-1}$ % [Vilkamo2013]，方程式 (11)

% 能量補償

17 $C_{\hat{y}} \leftarrow M_{\text{opt}} C_x M_{\text{opt}}^H$

18

$$g_{i,i} \leftarrow \sqrt{\frac{c_{y_{i,i}}}{c_{\hat{y}_{i,i}}}}, i = 1, \dots, n$$

% [Vilkamo2013], 方程式

(17)

19

$$M \leftarrow M_{opt} G$$

% [Vilkamo2013], 方程式

(17)

【0125】 仔細比較習知技術之演算法和本發明之演算法，發現前者需要大小分別為 $m \times m$ 、 $n \times n$ 和 $m \times n$ 的三個矩陣的SVD，其中 m 是降混聲道的數量， n 是對象渲染到的輸出聲道的數量。

【0126】 本發明之演算法只需要大小為 $m \times k$ 的一個矩陣的SVD，其中 k 是主要對象的數量。此外，由於 k 通常遠小於 n ，因此該矩陣小於習知技術之演算法的相應矩陣。

【0127】 對於 $m \times n$ 矩陣[Golub2013]，標準SVD實施的複雜性大致為 $O(c_1 m^2 n + c_2 n^3)$ ，其中 c_1 和 c_2 是常數，其取決於所使用的演算法，因此，與習知技術之演算法相比，本發明之演算法能夠達到計算複雜度的顯著降低。

【0128】 隨後，關於第一實施態樣的編碼器側的較佳實施例將參照圖7a、7b進行討論，此外，關於第一實施態樣的編碼器側的較佳實施例將參照圖9a至9d進行討論。

【0129】 圖7a顯示如圖1a所示的對象參數計算器100的一較佳實施方式。在方塊120中，音頻對象被轉換成頻譜表示，這由圖2或圖3的濾波器組102實現。然後，在方塊122中，例如在圖2或圖3所示的方塊104中計算選擇資訊，為此，可以使用幅度相關度量，例如幅度本身、功率、能量或通過將幅度提高到功率而獲得的任何其他幅度相關的度量，其中功率不等於1；方塊122的結果是一個選擇資訊的集合，其對應時間/頻率柱中的每個對象。接著，在方塊124中，導出每個時間/頻率柱的對象ID；在第一實施態樣，導出每個時間/頻率柱的兩個或更多個對象ID；在第二實施態樣，每個時間/頻率柱的對象ID的數量甚至可以僅為單個對象ID，以便從方塊122提供的資訊中，在方塊124中識別出最重要或最強或最相關的對象，方塊124輸出關於參數資料的資訊，並且包括最相關的一個或多個對象的單個或多個索引。

【0130】 在每個時間/頻率柱具有兩個或更多相關對象的情況下，方塊126的功能是用來計算表徵時間/頻率柱中的對象的幅度相關度量，這種幅度相關的測量可以相同於在方塊122中已經計算的選擇資訊，或者較佳地，組合值是使用方塊102已經計算的資訊來計算的，如方塊122和方塊126之間的虛線所示，並且接著在方塊126中計算與幅度相關的量度或一個以上之組合值，並將其轉發到量化器和編碼器方塊212，以便將輔助資訊中的編碼幅度相關或編碼組合值作為附加參數輔助資訊。在圖2或圖3的實施例中，這些是“編碼功率比”，其係與“編碼對象索引”一起包含在位元流中。在每個頻率柱只有一個對象ID的情況下，時間頻率柱中最相關對象的索引便足以執行解碼器端渲染，而功率比計算和量化編碼則不是必需的。

【0131】 圖7b顯示選擇資訊的計算的一較佳實施方式。如方塊123所示，為每個對象和每個時間/頻率柱計算信號功率作為選擇資訊。然後，方塊125說明圖7a的方塊124的一較佳實施方式，其中，具有最高功率的單個或較佳為兩個或更多個對象的對象ID被提取和輸出。此外，方塊127說明圖7a的方塊126的一較佳實施方式，其中，在兩個或更多相關對象的情況下，如方塊127所示計算一功率比，其中針對與由方塊125找到的對象ID對應的所有提取對象的功率相關的提取對象ID，計算功率比。這個過程是有利的，因為只需要傳輸比時間/頻率柱的對象數量少一個的組合值的數量，因為如同實施例的說明，在這個過程中存在解碼器已知的規則，即所有對象的功率比必須加起來為1。較佳地，圖7a的方塊120、122、124、126及/或圖7b的方塊123、125、127的功能由圖1a的對象參數計算器100實現，而圖7a的方塊212的功能由圖1a的輸出介面200實現。

【0132】 隨後，藉由幾個實施例來更詳細地解釋如圖1b所示的第二實施態樣的用於編碼的設備。在步驟110a中，從輸入信號中提取方向資訊(如圖12a所示)，或者通過讀取或解析包括在後設資料部分或後設資料檔案中的後設資料資訊來提取方向資訊。在步驟110b中，每幀和每音頻對象的方向資訊被量化，並且每幀每對象的量化索引被轉發到一編碼器或一輸出介面，例如圖1b的輸出介面200。在步驟110c中，方向量化索引被去量化，以取得一去量化值，其亦可以在某些實施方式中由方塊110b直接輸出。然後，基於去量化的方向索引，方

塊422基於某個虛擬麥克風設置計算每個傳輸聲道和每個對象的權重，該虛擬麥克風設置可以包括佈置在相同位置並具有不同方向的兩個虛擬麥克風信號，或者可以是具有相對於參考位置或方向(如虛擬聽者的位置或方向)的兩個不同位置的設置，具有兩個虛擬麥克風信號的設置將導致每個對象的兩個傳輸聲道的權重。

【0133】 在生成三個傳輸聲道的情況下，虛擬麥克風設置可以被認為包括來自佈置在相同位置並具有不同方向、或相對於參考位置或方向的三個不同位置的麥克風的三個虛擬麥克風信號，其中該參考位置或方向可以是虛擬聽者的位置或方向。

【0134】 再者，可以基於虛擬麥克風設置生成四個傳輸聲道，其係從佈置在相同位置並具有不同方向的麥克風、或從佈置在相對於參考位置或參考方向的四個不同位置的四個虛擬麥克風信號，生成四個虛擬麥克風信號，其中參考位置或方向可以是虛擬聽者位置或虛擬聽者方向。

【0135】 另外，為了計算每個對象和每個傳輸聲道 w_L 和 w_R 的權重，例如兩個聲道，虛擬麥克風信號是從以下麥克風導出的信號，如虛擬一階麥克風、或虛擬心形麥克風、或虛擬八字形麥克風、或偶極麥克風、或雙向麥克風、或虛擬定麥克風、或虛擬亞心形麥克風、或虛擬單向麥克風、或虛擬超心形麥克風、或虛擬全向麥克風。

【0136】 在這種情況下，需注意者，為了計算權重，不需要放置任何實際麥克風。相反地，計算權重的規則根據虛擬麥克風設置而變化，即虛擬麥克風的位置和虛擬麥克風的特性。

【0137】 在圖9a的方塊404中，將權重應用於對象，以便對於每個對象，在權重不為0的情況下獲得對象對某個傳輸聲道的貢獻。因此，方塊404 接收對象發出信號以作為輸入；然後，在方塊406中，將每個傳輸聲道的貢獻相加，從而例如將來自第一傳輸聲道的對象的貢獻加在一起、並且將來自第二傳輸聲道的對象的貢獻加在一起，以此類推；然後，如方塊406所示，方塊406的輸出例如是時域中的傳輸聲道。

【0138】 較佳地，輸入到方塊404的對象信號是具有全頻帶資訊的時域對象信號，且在時域中執行方塊404中的應用和方塊406中的求和。然而，在其他實施例中，這些步驟也可以在頻譜域中執行。

【0139】 圖9b顯示實現靜態降混的另一實施例。為此，在方塊130中提取一第一幀的一方向資訊，並且如方塊403a所示根據第一幀計算權重，然後，對於方塊408中指示的其他幀，權重保持原樣，以便實現靜態降混。

【0140】 圖9c顯示另一種實施，其係計算動態降混。為此，方塊132提取每一幀的方向資訊，並且如方塊403b所示為每一幀更新權重。然後，在方塊405，更新的權重被應用於該等幀以實現逐幀變化的動態降混。在圖9b和9c所顯示的數個極端情況之間的其他實施也是可行的，例如，其中僅對每第二個、每第三個、或每第n個幀更新權重，及/或隨著時間的推移執行權重的平滑，以便為了根據方向資訊進行降混時，天線特性不會經常變化太大。圖9d顯示由圖1b的對象方向資訊提供器110控制的降混器400的另一實施方式。在方塊410中，降混器被配置為分析一幀中所有對象的方向資訊，並且在方塊112中，為了計算立體聲示例的權重 w_L 和 w_R 之目的之麥克風被放置在與分析結果一致，其中麥克風的放置是指麥克風的位置及/或麥克風的指向性。在方塊414中，類似於關於圖9b的方塊408所討論的靜態降混，麥克風被留給其他幀，或者根據以上關於圖9c的方塊405所討論的內容來更新麥克風，以便獲得圖9d的方塊414的功能。關於方塊412的功能，可以放置麥克風以便獲得良好的分離，使得第一虛擬麥克風“對”到第一組對象、並且第二虛擬麥克風“對”到與第一組對象不同的第二組對象，兩組對象的不同之處較佳在於，一組的任何對象盡可能不包括在另一組中。或者，方塊410的分析可以通過其他參數來增強，並且其設置也可以通過其他參數來控制。

【0141】 隨後，根據第一或第二實施態樣並且關於如圖6a和圖6b所討論的解碼器的較佳實施方式，將參考圖10a、10b、10c、10d 和11分別說明如下。

【0142】 在方塊613中，輸入介面600被配置為擷取與對象ID相關聯的個體對象方向資訊。該過程對應於圖4或5的方塊612的功能性、並且產生如關於圖8b(特別是圖8c)所示和討論的“用於一幀的碼本”。

【0143】 此外，在方塊609中，擷取每個時間/頻率柱的一個以上之對象ID，而不管該些資料對於低解析度參數帶或高解析度頻率塊是否可用。對應於圖4中的方塊608的過程的方塊609的結果是一個以上之相關對象的時間/頻率柱中的特定ID。然後，在方塊611中，從“一幀的碼本”，即從圖8c所示的示例表中，擷取每個時間/頻率柱的特定的一個以上之ID的特定對象方向資訊。接著，在方塊704中，針對各個輸出聲道的一個以上之相關對象計算增益值，如由每個時間/頻率柱計算的輸出格式所支配。然後，在方塊730或706、708中，計算輸出聲道。輸出聲道的計算功能可以在如圖10b所示的一個以上之傳輸聲道的貢獻的顯式計算內完成，或者可以通過如圖10d或11所示之傳輸聲道貢獻的間接計算和使用來完成。圖10b顯示其中在與圖4的功能相對應的方塊610中擷取功率值或功率比的功能，然後，將這些功率值應用於如方塊733和735所示的每個相關對象的各個傳輸聲道。此外，除了由方塊704決定的增益值之外，這些功率值還被應用到各個傳輸聲道，使得方塊733、735導致傳輸聲道(例如傳輸聲道ch1、ch2, ...)的對象特定貢獻，接著，在方塊737中，這些明確計算的聲道傳輸貢獻針對每時間/頻率柱每個輸出聲道加總在一起。

【0144】 然後，根據本實施方式，可以提供一擴散信號計算器741，其係為每個輸出聲道ch1、ch2、...等，生成在相應的時間/頻率柱中的擴散信號，並且將擴散信號的組合和方塊737的貢獻結果進行組合，以便獲得每個時間/頻率柱中的完整聲道貢獻。當共變異數合成另外依賴於擴散信號時，該信號對應於圖4的濾波器組708的輸入。然而，當共變異數合成706不依賴於擴散信號、而僅依賴於沒有任何去相關器的處理時，則至少每個時間/頻率柱的輸出信號的能量對應於在圖10b的方塊739的輸出的聲道貢獻的能量。此外，在不使用擴散信號計算器741的情況下，方塊739的結果對應於方塊706的結果，其中每個時間/頻率柱具有完整的聲道貢獻，可以為每個輸出聲道ch1、ch2單獨轉換，以便最終獲得具有時域輸出聲道的輸出音頻檔案，其可以儲存、或轉發到揚聲器或任何類型的渲染裝置。

【0145】 圖10c顯示如圖10b或4的方塊610的功能的一較佳實施方式。在步驟610a中，針對某個時間/頻率柱擷取一個或數個組合的(功率)值。在方塊

610b中，基於所有組合值必須加總為一的計算規則，計算時間/頻率柱中的其他相關對象的對應之其他值。

【0146】 然後，結果將較佳是低解析度表示，其中每個分組的時隙索引和每個參數帶索引具有兩個功率比，這代表低時間/頻率解析度。在方塊610c中，時間/頻率解析度可以擴展到高時間/頻率解析度，使得具有高解析度時隙索引 n 和高解析度頻帶索引 k 的時間/頻率磚的功率值，此擴展可以包括直接使用一個和相同的低解析度索引，其用於分組時隙內的相應時隙、和參數頻帶內的相應頻帶。

【0147】 圖10d顯示用於計算圖4的方塊706中的共變異數合成資訊的功能的較佳實施方式，該功能由混合矩陣725表示，該混合矩陣725用於將兩個或更多個輸入傳輸聲道混合成兩個或更多個輸出信號。因此，例如，當有兩個傳輸聲道和六個輸出聲道時，每個單獨的時間/頻率柱的混合矩陣的大小將為六行和兩列。在對應於圖5中的方塊723的功能的圖10d中的方塊723中，接收每個時間/頻率柱中每個對象的增益值或直接響應值，並計算共變異數矩陣。在方塊722中，接收功率值或比率、並計算時間/頻率柱中每個對象的直接功率值，並且圖10d中的方塊722對應於圖5的方塊722。

【0148】 方塊721和722的結果都被輸入到目標共變異數矩陣計算器724中。另外或替代地，目標共變異數矩陣 C_y 的顯式計算不是必需的。取而代之的是，將目標共變異數矩陣中包含的相關資訊，即矩陣 R 中指示的直接響應值資訊和矩陣 E 中指示的兩個或多個相關對象的直接功率值，輸入到方塊725a中以計算每個時間/頻率柱的混合矩陣。此外，混合矩陣725a接收關於原型矩陣 Q 和從對應於圖5的方塊726的方塊726中所示的兩個或更多傳輸聲道導出的輸入共變異數矩陣 C_x 的資訊。每個時間/頻率柱和每幀的混合矩陣可以經受如方塊725b所示的時間平滑，並且在對應於圖5的渲染方塊的至少一部分的方塊727中，混合矩陣以非平滑或平滑的形式應用於傳輸相應的時間/頻率柱中的聲道，以獲得時間/頻率柱中的完整聲道貢獻，該貢獻基本上類似於前面關於圖10b在方塊739的輸出處所討論的相應完整貢獻。因此，圖10b說明了傳輸聲道貢獻的顯式計算的實施方式，而圖10d說明了針對每個時間/頻率柱和每個時間頻率柱中的每個

相關對象的傳輸聲道貢獻的隱式計算的過程，經由目標共變異數矩陣 C_y 或經由直接引入混合矩陣計算方塊725a中的方塊723和722的相關資訊 R 和 E 。

【0149】 隨後，圖11顯示出了用於共變異數合成的較佳優化演算法，其中圖11中顯示出的所有步驟是在圖4的共變異數合成706內、或在混合矩陣計算方塊725(如圖5)或725a(如圖10d)內計算。在步驟751中，計算第一分解結果 K_y 。由於如圖10d所示，矩陣 R 中包含的增益值資訊和來自兩個或多個相關對象的資訊，特別是矩陣 ER 中包含的直接功率資訊可以直接使用、無需顯式計算共變異數矩陣，因此可以很容易地計算出該分解結果。因此，方塊751中的第一分解結果可以直接計算並且無需太多功夫，因為不再需要特定的奇異值分解。

【0150】 在步驟752中，計算第二分解結果為 K_x 。這個分解結果也可以在沒有顯式奇異值分解的情況下計算，因為輸入共變異數矩陣被視為對角矩陣，其中非對角元素被忽略。

【0151】 然後，在步驟753中，根據第一正則化參數 α 計算第一正則化結果，並且在步驟754中，根據第二正則化參數 β 計算第二正則化結果。在較佳實施方式中，令 K_x 為對角矩陣，第一正則化結果的計算753相對於習知技術是簡化的，因為 S_x 的計算只是參數變化而不是像習知技術那樣的分解方式。

【0152】 進一步地，對於步驟754中的第二正則化結果的計算，第一步只是另外對參數重命名，而不是如習知技術中的與矩陣 U_x^{HS} 相乘。

【0153】 此外，在步驟755中，計算歸一化矩陣 G_y ，並且基於步驟755，在步驟756中基於 K_x 和原型矩陣 Q 以及方塊751獲得的 K_y 的資訊，計算么正矩陣 P 。由於這裡不需要任何矩陣 Λ ，因此相對於習知技術可以簡化么正矩陣 P 的計算。

【0154】 然後，在步驟757，計算沒有能量補償的混合矩陣 M_{opt} ，為此，使用么正矩陣 P 、方塊754的結果和方塊751的結果。然後，在方塊758中，使用補償矩陣 G 執行能量補償。執行能量補償使得從去相關器導出的任何殘餘信號都不是必需的。然而，代替執行能量補償，在本實施方式中將添加具有足夠大的能量以填充混合矩陣 M_{opt} 留下的能量間隙，而沒有能量資訊的殘餘信號。然

而，為了本發明的目的，不依賴去相關信號以避免去相關器引入的任何偽物，但是較佳的是如步驟758中所示的能量補償。

【0155】 因此，共變異數合成的優化演算法在步驟751、752、753、754中以及在步驟756中為止矩陣P的計算提供了優勢。需要強調的是，優化演算法甚至提供優於先前演算法的優勢，其中僅步驟755、752、753、754、756中的一個或這些步驟的子組被實施，但相應的其他步驟如習知技術中那樣實施。原因是改進不相互依賴，而是可以相互獨立應用。然而，實施的改進越多，就實施的複雜性而言，該過程就越好。因此，圖11實施例的完整實施是較佳的，因為其提供了最大量的複雜性降低，但即使當根據優化演算法僅實施步驟751、752、753、754、756之一時，其他步驟與習知技術相同，在沒有任何品質惡化的情況下獲得複雜度的降低。

【0156】 本發明的實施例也可以被認為是通過混合三個高斯噪音源來為立體聲信號生成柔和噪音的過程，其一是針對每個聲道和第三個公共噪音源，以創建相關的背景噪音，或者附加地或單獨地控制混合與SID幀一起傳輸的相關值的噪音源。

【0157】 需注意者，以上所述和下面討論的所有替代方案或實施態樣、以及由後續請求項定義的所有實施態樣都可以單獨使用，即，除了預期的替代方案、目標或獨立請求項之外，不與任何其他替代方案或目標或獨立請求項組合。然而，在其他實施例中，兩個或更多個替代方案或實施態樣或獨立請求項可以彼此組合，並且在其他實施例中，所有實施態樣或替代方案和所有獨立請求項可以彼此組合。

【0158】 本發明編碼的信號可以儲存在數位儲存媒體或非暫時性儲存媒體上，或者可以在傳輸媒體上傳輸，如無線傳輸媒體或有線傳輸媒體(如網際網路)。

【0159】 儘管已經在設備的說明中描述了一些實施態樣，但很明顯地，這些實施態樣也代表了相應方法的描述，其中方塊或裝置對應於方法步驟或方法步驟的特徵。類似地，在方法步驟的說明中描述的實施態樣也表示相應設備的相應方塊或項目或特徵的描述。

【0160】 根據某些實施要求，本發明的實施例可以利用硬體或軟體來實現，該實現可以使用數位儲存媒體來執行，例如磁碟、DVD、CD、ROM、PROM、EPROM、EEPROM或FLASH記憶體，其具有儲存在其上的電子可讀控制信號，其配合(或可配合)可編程計算機系統運作，從而執行相應的方法。

【0161】 根據本發明的一些實施例包括具有電子可讀控制信號的資料載體，其能夠配合可編程計算機系統運作，從而執行本說明書所述的方法其中之一。

【0162】 一般而言，本發明的實施例可以實現為具有程式碼的電腦程式產品，當電腦程式產品在電腦上運行時，該程式碼可操作用於執行該等方法其中之一，程式碼可以例如儲存在機器可讀載體上。

【0163】 其他實施例包括用於執行本說明書所描述的該等方法其中之一的電腦程式，其儲存在機器可讀載體或非暫時性儲存媒體上。

【0164】 換句話說，本發明之方法的實施例因此是具有程式碼的電腦程式，當該電腦程式在電腦上運行時，用於執行所描述的該等方法其中之一。

【0165】 因此，本發明之方法的另一實施例是一資料載體(或數位儲存媒體、或電腦可讀媒體)，其上記錄有用於執行本說明書所述的該等方法其中之一的電腦程式。

【0166】 因此，本發明之方法的另一實施例是一資料流或信號序列，其表示用於執行所描述的該等方法其中之一的電腦程式，該資料流或信號序列可以例如被配置為經由資料通訊連接(例如經由網際網路)來傳輸。

【0167】 另一個實施例包括一處理裝置，例如電腦或可編程邏輯裝置，其被配置為或適合於執行所描述的該等方法其中之一。

【0168】 另一實施例包括其上安裝有用於執行所描述的該等方法其中之一的電腦程式的電腦。

【0169】 在一些實施例中，可編程邏輯裝置(例如現場可編程閘極陣列)可用於執行所述方法的一些或全部功能。在一些實施例中，現場可編程閘極陣列可以與微處理器協作以執行所述的方法其中之一。通常，這些方法較佳由任何硬體設備執行。

【0170】 上述實施例僅用於說明本發明的原理。應當理解，對本領域技術人員而言，這裡描述的各種修改和變化的配置及其細節將是顯而易見的。因此，其意圖是僅受限於後續的申請專利範圍，而不是受限於通過本說明書之實施例的描述和解釋所呈現的具體細節。

實施態樣(彼此獨立使用、或與所有其他實施態樣一起使用、或僅是其他實施態樣的一個子組)

【0171】 一種設備、方法或電腦程式，包括以下一個或多個特徵：

【0172】 關於新穎性實施態樣的創造性示例：

- 多波想法與對象編碼相結合(每個T/F磚使用超過一個以上的方向提示)
- 盡可能接近DirAC範例的對象編碼方法，允許在IVAS中使用任何類型的輸入類型(目前尚未涵蓋對象內容)

【0173】 關於參數化(編碼器)的創造性示例：

- 對於每個T/F磚：此T/F磚中的n個最相關對象的選擇資訊加上這n個最相關對象的貢獻之間的功率比
- 對於每一幀，對於每個對象：一個方向

關於渲染(解碼器)的創造性示例：

- 從傳輸的對象索引和方向資訊以及目標輸出佈局中獲取每個相關對象的直接響應值
- 從直接響應中獲取共變異數矩陣
- 根據每個相關對象的降混信號功率和傳輸功率比計算直接功率
- 從直接功率和共變異數矩陣中獲取最終目標共變異數矩陣
- 僅使用輸入共變異數矩陣的對角元素
- 優化共變異數合成

【0174】 關於與SAOC差異的一些旁注：

- 考慮n個主要對象而不是所有對象
→ 功率比因此與OLD相關，但計算方式不同
- SAOC不使用編碼器的方向 -> 僅在解碼器(渲染矩陣)導入的方向資訊
→ SAOC-3D解碼器接收用於渲染矩陣的對象後設資料

- SAOC採用降混矩陣並傳輸降混增益
- 在本發明的實施例中不考慮擴散

【0175】 以下總結本發明的進一步示例。

【0176】 1. 一種用於對多個音頻對象和指示關於該多個音頻對象之一方向資訊之一後設資料進行編碼的設備，包含：

一降混器(400)，用於降混該多個音頻對象以獲得一個以上之傳輸聲道；

一傳輸聲道編碼器(300)，用於對該一個以上之傳輸聲道進行編碼以獲得一個以上之編碼傳輸聲道；以及

一輸出介面(200)，用於輸出包括該一個以上之編碼傳輸聲道的一編碼音頻信號，

其中，該降混器(400)被配置為響應於關於該多個音頻對象之該方向資訊對該多個音頻對象進行降混。

【0177】 2. 如示例1所述之設備，其中該降混器(400)被配置為生成兩個傳輸聲道以作為兩個虛擬麥克風信號，該兩個虛擬麥克風信號被安排在相同的位置並具有不同的方向，或者在相對於一參考位置或方向(例如一虛擬聽者位置或方向)的兩個不同位置，或

生成三個傳輸聲道以作為三個虛擬麥克風信號，該三個虛擬麥克風信號被安排在相同的位置並具有不同的方向，或者在相對於一參考位置或方向(例如一虛擬聽者位置或方向)的三個不同位置，或

生成四個傳輸聲道以作為四個虛擬麥克風信號，該四個虛擬麥克風信號被安排在相同的位置並具有不同的方向，或者在相對於一參考位置或方向(例如一虛擬聽者位置或方向)的四個不同位置，或

其中，該等虛擬麥克風信號為虛擬第一階麥克風信號、或虛擬心形麥克風信號、或虛擬八字形或偶極或雙向麥克風信號、或虛擬定向麥克風信號、或虛擬亞心形麥克風信號、或虛擬單向麥克風信號、或虛擬超心形麥克風信號、或虛擬全向麥克風信號。

【0178】 如示例1或2所述之設備，其中該降混器(400)被配置為

使用對應之該音頻對象的該方向資訊為該多個音頻對象中的各該音頻對象導出(402)各該傳輸聲道的一加權資訊；

使用一特定傳輸聲道的該音頻對象的該加權資訊對相應之該音頻對象進行加權(404)，以獲得該特定傳輸聲道的一對象貢獻，以及組合(406)來自該多個音頻對象的該特定傳輸聲道的對象貢獻，以獲得該特定傳輸聲道。

【0179】 4. 如以上示例中任一所述之設備，

其中，該降混器(400)被配置為將該一個以上之傳輸聲道計算為一個以上之虛擬麥克風信號，該等虛擬麥克風信號被安排在相同的位置並具有不同的方向、或在相對於一參考位置或方向(例如一虛擬聽者位置或方向)的不同位置，其與該方向資訊相關，

其中，該等不同的位置或方向位於或朝向一中心線的左側、以及位於或朝向該中心線的右側，或者其中該等不同的位置或方向均等或不均等地分佈到水平位置或方向(例如相對於該中心線的+90度或-90度、或相對於該中心線的-120度、0度和+120度)，或者其中該等不同的位置或方向包括相對於一虛擬聽者所處位置的一水平面的至少一個朝上或朝下的位置或方向，其中關於該多個音頻對象的該方向資訊係相關於該虛擬聽者位置、或該參考位置、或該方向。

【0180】 5. 如以上示例中任一所述之設備，更包含：

一參數處理器(110)，用於量化指示關於該多個音頻對象的該方向資訊的該後設資料，以獲得該多個音頻對象的量化方向項目，

其中，該降混器(400)被配置為響應於該量化方向項目作為該方向資訊進行操作，以及

其中，該輸出介面(200)被配置為將該量化方向項目的資訊導入該編碼音頻信號中。

【0181】 6. 如以上示例中任一所述之設備，

其中，該降混器(400)被配置為執行關於該多個音頻對象的該方向資訊的一分析，並且根據該分析的一結果放置用於該傳輸聲道之生成的一個以上之虛擬麥克風。

7. 如以上示例中任一所述之設備，

其中，該降混器(400)被配置為使用在多個時間幀上靜態的一降混規則來進行降混(408)，或

其中，該方向資訊在多個時間幀上是可變的，並且其中該降混器(400)被配置為使用在多個時間幀上可變的一降混規則來進行降混(405)。

【0182】 8. 如以上示例中任一所述之設備，

其中，該降混器(400)被配置為使用對該多個音頻對象的樣本以逐個樣本加權和組合的方式，在一時域中進行降混。

【0183】 9. 如以上示例中任一所述之設備，更包含：

一對象參數計算器(100)，其被配置為針對與一時間幀相關的多個頻率柱中的一個以上之頻率柱計算至少兩個相關音頻對象的參數資料，其中該至少兩個相關音頻對象的數量低於該多個音頻對象的總數量，以及

其中，該輸出介面(200)被配置為將關於該一個以上之頻率柱的該至少兩個相關音頻對象的該參數資料的資訊導入該編碼音頻信號中。

【0184】 10. 如示例9所述之設備，其中該對象參數計算器(100)被配置為將該多個音頻對象中的各該音頻對象轉換(120)為具有該多個頻率柱的一頻譜表示，

針對該一個以上之頻率柱，從各該音頻對象計算(122)一選擇資訊，及基於該選擇資訊導出(124)對象標識以作為指示該至少兩個相關音頻對象的該參數資料，以及

其中，該輸出介面(200)被配置為將該對象標識的資訊導入該編碼音頻信號中。

【0185】 11. 如示例9或10所述之設備，其中該對象參數計算器(100)被配置為量化和編碼(212)一個以上之幅度相關量度或者從該一個以上之頻率柱中的該等相關音頻對象的該幅度相關量度中導出的一個以上之組合數值，以及

其中，該輸出介面(200)被配置為將量化的該一個以上之幅度相關量度或量化的該一個以上之組合數值導入該編碼音頻信號中。

【0186】 12. 如示例10或11所述之設備，

其中，該選擇資訊是與幅度相關的量度(例如一幅度值、一功率值或一響度值)、或提高到與該音頻對象之功率不同的功率的幅度，以及

其中，該對象參數計算器(100)被配置為計算(127)一組合數值(例如一相關音頻對象的一幅度相關量度和該相關音頻對象的兩個以上之幅度相關量度之和的比率)，以及

其中，該輸出介面(200)被配置為將該組合數值的資訊導入該編碼音頻信號中，其中該編碼音頻信號中的該組合數值之資訊項目的數量係大於等於1、且小於該一個以上之頻率柱的等相關音頻對象的數量。

【0187】 13. 如示例10至12中任一所述之設備，

其中，該對象參數計算器(100)被配置為基於該一個以上之頻率柱中的該多個音頻對象的該選擇資訊的一順序來選擇該對象標識。

【0188】 14. 如示例10至13中任一所述之設備，其中，該對象參數計算器(100)被配置為

計算(122)一信號功率以作為該選擇資訊，

分別針對各該頻率柱，導出(124)對應之該一個以上之頻率柱中具有該等最大信號功率值的該兩個以上之音頻對象的該對象標識，

計算(126)具有該最大信號功率值的該兩個以上之音頻對象的該信號功率之和與具有導出之該對象標識的該等音頻對象中的至少一個的該信號功率之間的功率比，以作為該參數資料，及

量化和編碼(212)該功率比，以及

其中，該輸出介面(200)被配置為將量化和編碼之該功率比導入該編碼音頻信號中。

【0189】 15. 如示例10至14中任一所述之設備，其中該輸出介面(200)被配置為將下列資訊導入該編碼音頻信號；一個以上之編碼傳輸聲道；作為該參數資料，該時間幀中的該多個頻率柱中的該一個以上之頻率柱中的各該頻率柱的

該等相關音頻對象的兩個以上之編碼對象標識，以及一個以上之編碼組合數值或編碼幅度相關量度；以及該時間幀中的各該音頻對象的量化和編碼方向資料，該方向資料對於該一個以上之頻率柱的所有該等頻率柱是恆定的。

【0190】 16. 如示例9至15中任一所述之設備，其中該對象參數計算器(100)被配置為計算該一個以上之頻率柱中的至少一最主要對象及一第二主要對象的該參數資料，或

其中，該多個音頻對象的數量為三個以上，該多個音頻對象包括一第一音頻對象、一第二音頻對象、及一第三音頻對象，以及

其中，該對象參數計算器(100)被配置為僅以一第一音頻對象群組(例如該第一音頻對象和該第二音頻對象)作為該相關音頻對象來計算該一個以上之頻率柱中的一第一頻率柱，以及僅以一第二音頻對象群組(例如該第二音頻對象和該第三音頻對象、或是該第一音頻對象和該第三音頻對象)作為該相關音頻對象來計算該一個以上之頻率柱中的一第二頻率柱，其中該第一音頻對象群組與該第二音頻對象群組之間至少有一個群組成員是不同的。

【0191】 17. 如示例9至16中任一所述之設備，其中該對象參數計算器(100)被配置為

計算具有一第一時間或頻率解析度的一原始參數資料，並將該原始參數資料組合到具有低於該第一時間或頻率解析度的一第二時間或頻率解析度的一組合參數資料，以及計算關於具有該第二時間或頻率解析度的該組合參數資料的該至少兩個相關音頻對象的該參數資料，或

決定具有與在該多個音頻對象的一時間或頻率分解中使用的一第一時間或頻率解析度不同的一第二時間或頻率解析度的參數頻帶，以及計算用於具有該第二時間或頻率解析度的該參數頻帶的該至少兩個相關音頻對象的該參數資料。

【0192】 18. 一種用於解碼一編碼音頻信號的解碼器，該編碼音頻信號包括多個音頻對象的一個以上之傳輸聲道和方向資訊、及一時間幀的一個以上之頻率柱的一音頻對象的一參數資料，該解碼器包含：

一輸入介面(600)，用於以在該時間幀中具有該多個頻率柱的一頻譜表示來提供該一個以上之傳輸聲道；以及

一音頻渲染器(700)，用於使用該方向資訊將該一個以上之傳輸聲道渲染成數個音頻聲道，

其中，該音頻渲染器(700)被配置為根據該多個頻率柱的各該頻率柱的該一個以上之音頻對象、以及與該頻率柱的相關之該一個以上之音頻對象相關聯的該方向資訊(810)來計算一直接響應資訊(704)。

【0193】 19. 如示例18所述之解碼器，

其中，該音頻渲染器(700)被配置為使用該直接響應資訊和該數個音頻聲道的資訊(702)來計算(706)一共變異數合成資訊，並且將該共變異數合成資訊應用(727)於該一個以上之傳輸聲道以獲得該數個音頻聲道，或

其中，該直接響應資訊(704)是各該一個以上之音頻對象的一直接響應向量，並且其中該共變異數合成資訊是一共變異數合成矩陣，並且其中該音頻渲染器(700)被配置為應用(727)該共變異數合成資訊對每一頻率柱執行一矩陣運算。

【0194】 20. 如示例18或19所述之解碼器，其中該音頻渲染器(700)被配置為

在該直接響應資訊(704)的計算中，導出該一個以上之音頻對象的一直接響應向量，並為該一個以上之音頻對象從各該直接響應向量計算一共變異數矩陣，

在該共變異數合成資訊的計算中從以下導出(724)一目標共變異數資訊：該一個音頻對象的該共變異數矩陣或該多個音頻對象的該等共變異數矩陣，相應之該一個以上之音頻對象的一功率資訊，以及從該一個以上之傳輸聲道導出的一功率資訊。

【0195】 21. 如示例20所述之解碼器，其中該音頻渲染器(700)被配置為在該直接響應資訊的計算中，導出該一個以上之音頻對象的一直接響應向量，並為各該一個以上之音頻對象從各該直接響應向量計算(723)一共變異數矩陣，

從該傳輸聲道導出(726)一輸入共變異數資訊，以及
從該目標共變異數資訊、該輸入共變異數資訊和關於該數個音頻聲道之資訊導出(725a、725b)一混合資訊，以及
將該混合資訊應用(727)到該時間幀中的各該頻率柱的該等傳輸聲道。

【0196】 22. 如示例21所述之解碼器，其中將該混合資訊應用到該時間幀中的各該頻率柱的結果轉換(708)到一時域中以獲得該時域中的該數個音頻聲道。

【0197】 23. 如示例18至22中任一所述之解碼器，其中該音頻渲染器(700)被配置為

在從該等傳輸聲道導出的一輸入共變異數矩陣的一分解(752)中，僅使用該輸入共變異數矩陣的主對角元素，或
使用該等對象或該等傳輸聲道的一直接響應矩陣和一功率矩陣，執行一目標共變異數矩陣的一分解(751)，或
通過取該輸入共變異數矩陣的各該主對角元素的根來執行(752)該輸入共變異數矩陣的一分解，或
計算(753)已分解之該輸入共變異數矩陣的一正規化逆矩陣，或
執行(756)一奇異值分解以在沒有一擴展單位矩陣的情況下計算用於一能量補償的一最佳矩陣。

【0198】 24. 如示例18至23中任一所述之解碼器，其中該一個以上之音頻對象的該參數資料包括至少兩個相關音頻對象的一參數資料，其中該至少兩個相關音頻對象的數量少於該多個音頻對象的總數，以及

其中，該音頻渲染器(700)被配置為對於該一個以上之頻率柱中的每一個，根據與該至少兩個相關音頻對象的一第一相關音頻對象的一第一方向資訊以及與該至少兩個相關音頻對象的一第二相關音頻對象的一第二方向資訊，從該一個以上之傳輸聲道中計算一貢獻。

【0199】 25. 如示例24所述之解碼器，

【0200】 其中，該音頻渲染器(700)被配置為對於該一個以上之頻率柱忽略與該至少兩個相關音頻對象不同的一音頻對象的一方向資訊。

【0201】 26. 如示例24或25所述之解碼器，

其中，該編碼音頻信號包括各該相關音頻對象的一幅度相關度量或與該參數資料中的至少兩個相關音頻對象相關的一組合值，以及

其中，該音頻渲染器(700)被配置為根據與該至少兩個相關音頻對象的一第一相關音頻對象相關聯的一第一方向資訊以及與該至少兩個相關音頻對象的一第二相關音頻對象相關聯的一第二方向資訊，將來自該一個以上之傳輸聲道的一貢獻考慮在內以進行操作，或者根據該幅度相關度量或該組合值來決定該一個以上之傳輸聲道的一定量貢獻。

【0202】 27. 如示例26所述之解碼器，其中該編碼信號包括該參數資料中的該組合值，以及

其中，該音頻渲染器(700)被配置為使用該等相關音頻對象其中之一的該組合值和該相關音頻對象的該方向資訊來決定該一個以上之傳輸聲道的該貢獻，以及

其中，該音頻渲染器(700)被配置為使用從該一個以上之頻率柱中的該等相關音頻對象其中之另一的該組合值和該另一相關音頻對象的該方向資訊所導出的一值來決定該一個以上之傳輸聲道的該貢獻。音頻對象。

【0203】 28. 如示例24至27中任一所述之解碼器，其中該音頻渲染器(700)被配置為

從該多個頻率柱中的各該頻率柱的該等相關音頻對象和與等頻率柱中的該等相關音頻對象相關聯的該方向資訊，計算該直接響應資訊(704)。

【0204】 29. 如示例28所述之解碼器，

其中，該音頻渲染器(700)被配置為使用一擴散資訊(如包括在該後設資料中的一擴散參數)或一去相關規則來決定(741)該多個頻率柱中的各該頻率柱的一擴散信號，並且組合該擴散信號與由該直接響應資訊所決定之一直接響應，以獲得用於該數個聲道其中之一聲道的一頻譜域渲染信號。

【0205】 30. 一種用於對多個音頻對象和指示關於該多個音頻對象之一方向資訊之一後設資料進行編碼的方法，包括：

降混該多個音頻對象以獲得一個以上之傳輸聲道；

編碼該一個以上之傳輸聲道以獲得一個以上之編碼傳輸聲道；以及輸出包括該一個以上之編碼傳輸聲道的一編碼音頻信號，其中，該降混之步驟包括對應於該多個音頻對象的該方向資訊對該多個音頻對象進行降混。

【0206】 31. 一種用於解碼一編碼音頻信號的方法，該編碼音頻信號包括多個音頻對象的一個以上之傳輸聲道和方向資訊、及一時間幀的一個以上之頻率柱的一音頻對象的一參數資料，該方法包括：

以在該時間幀中具有該多個頻率柱的一頻譜表示來提供該一個以上之傳輸聲道；以及

使用該方向資訊將該一個以上之傳輸聲道音頻渲染成數個音頻聲道，

其中，該音頻渲染之步驟包括根據該多個頻率柱的各該頻率柱的該一個以上之音頻對象、以及與該頻率柱的相關之該一個以上之音頻對象相關聯的該方向資訊來計算一直接響應資訊。

【0207】 32. 一種電腦程式，當其運行於一電腦或一處理器時，用以執行如示例30所述之方法或如示例31所述之方法。

【0208】 參考書目或參考文獻

[Pulkki2009] V. Pulkki, M-V. Laitinen, J. Vilamo, J. Ahonen, T. Lokki, and T. Pihlajamäki, “Directional audio coding perception-based reproduction of spatial sound”, International Workshop on the Principles and Application on Spatial Hearing, Nov. 2009, Zao; Miyagi, Japan.

[SAOC_STD] ISO/IEC, “MPEG audio technologies Part 2: Spatial Audio Object Coding (SAOC).” ISO/IEC JTC1/SC29/WG11 (MPEG) International Standard 23003-2.

[SAOC_AES] J. Herre, H. Purnhagen, J. Koppens, O. Hellmuth, J. Engdegård, J. Hilpert, L. Villemoes, L. Terentiv, C. Falch, A. Hölzer, M. L. Valero, B. Resch, H. Mundt H, and H. Oh, “MPEG spatial audio object coding—the ISO/MPEG standard for efficient coding of interactive audio scenes,” *J. AES*, vol. 60, no. 9, pp. 655–673, Sep. 2012.

[MPEGH_AES] J. Herre, J. Hilpert, A. Kuntz, and J. Plogsties, “MPEG-H audio—the new standard for universal spatial/3D audio coding,” in *Proc. 137th AES Conv.*, Los Angeles, CA, USA, 2014.

[MPEGH_IEEE] J. Herre, J. Hilpert, A. Kuntz, and J. Plogsties, “MPEG-H 3D Audio—The New Standard for Coding of Immersive Spatial Audio“, IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, VOL. 9, NO. 5, AUGUST 2015

[MPEGH_STD] *Text of ISO/MPEG 23008–3/DIS 3D Audio, Sapporo*, ISO/IEC JTC1/SC29/WG11 N14747, Jul. 2014.

[SAOC_3D_PAT] APPARATUS AND METHOD FOR ENHANCED SPATIAL AUDIO OBJECT CODING, WO 2015/011024 A1

[Pulkki1997] V. Pulkki, “Virtual sound source positioning using vector base amplitude panning,” *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466, Jun. 1997.

[DELAUNAY] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, “The quickhull algorithm for convex hulls,” in *Proc. ACM Trans. Math. Software (TOMS)*, New York, NY, USA, Dec. 1996, vol. 22, pp. 469–483.

[Hirvonen2009] T. Hirvonen, J. Ahonen, and V. Pulkki, “Perceptual compression methods for metadata in Directional Audio Coding applied to audiovisual teleconference”, AES 126th Convention 2009, May 7–10, Munich, Germany.

[Borß2014] C. Borß, “A Polygon-Based Panning Method for 3D Loudspeaker Setups”, AES 137th Convention 2014, October 9 -12, Los Angeles, USA.

[WO2019068638] Apparatus, method and computer program for encoding, decoding, scene processing and other procedures related to DirAC based spatial audio coding, 2018

[WO2020249815] PARAMETER ENCODING AND DECODING FOR MULTICHANNEL AUDIO USING DirAC, 2019

[BCC2001] C. Faller, F. Baumgarte: “Efficient representation of spatial audio using perceptual parametrization”, Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575).

[JOC_AES] Heiko Purnhagen; Toni Hirvonen; Lars Villemoes; Jonas Samuelsson; Janusz Klejsa: “Immersive Audio Delivery Using Joint Object Coding”, 140th AES Convention, Paper Number: 9587, Paris, May 2016.

[AC4_AES] K. Kjörling, J. Rödén, M. Wolters, J. Riedmiller, A. Biswas, P. Ekstrand, A. Gröschel, P. Hedelin, T. Hirvonen, H. Hörich, J. Klejsa, J. Koppens, K. Krauss, H-M. Lehtonen, K. Linzmeier, H. Muesch, H. Mundt, S. Norcross, J. Popp, H. Purnhagen, J. Samuelsson, M. Schug, L. Sehlström, R. Thesing, L. Villemoes, and M. Vinton: “AC-4 – The Next Generation Audio Codec”, 140th AES Convention, Paper Number: 9491, Paris, May 2016.

[Vilkamo2013] J. Vilkamo, T. Bäckström, A. Kuntz, “Optimized covariance domain framework for time-frequency processing of spatial audio”, Journal of the Audio Engineering Society, 2013.

[Golub2013] Gene H. Golub and Charles F. Van Loan, “Matrix Computations”, Johns Hopkins University Press, 4th edition, 2013.

【符號說明】

【0209】

100: 對象參數計算器

102: 濾波器組、方塊

104: 信號功率計算方塊、方塊

106: 對象選擇方塊、對象選擇器、對象選擇、方塊

108: 功率比計算方塊、功率比計算、方塊

110: 對象方向資訊提供器、方塊、參數處理器

110a: 提取方向資訊方塊、方塊、步驟

110b: 量化方向資訊方塊、方塊、步驟

110c: 步驟

120: 方塊、轉換

122: 方塊、計算

123: 方塊

124: 方塊、導出
125: 方塊
126: 方塊、計算
127: 方塊、計算
130: 方塊
132: 方塊
200: 輸出介面、輸出介面方塊
202: 編碼方向資訊方塊、方塊、方向資訊編碼器
210: 方塊
212: 量化器和編碼器方塊、方塊、量化和編碼
220: 多工器、方塊
300: 傳輸聲道編碼器、核心編碼器
400: 降混器、降混計算方塊、降混計算
402: 導出
403a: 方塊
403b: 方塊
404: 方塊、加權
405: 方塊、降混
406: 方塊、組合
408: 方塊、降混
410: 方塊
412: 方塊
414: 方塊
600: 輸入介面方塊、輸入介面
602: 解多功器、項目
604: 核心解碼器、項目
606: 濾波器組、項目
608: 解碼器、項目、方塊

- 609: 方塊
- 610: 解碼器、項目、方塊
- 610a: 步驟
- 610b: 方塊
- 610c: 方塊
- 611: 方塊
- 612: 解碼器、項目、方塊
- 613: 方塊
- 700: 音頻渲染器方塊、音頻渲染器
- 702: 原型矩陣提供器、項目、音頻聲道的資訊
- 704: 直接響應計算器、項目、方塊、直接響應資訊
- 706: 共變異數合成方塊、項目、方塊、共變異數合成、計算
- 708: 合成濾波器組、項目、濾波器組方塊、方塊、濾波器組、轉換
- 721: 信號功率計算方塊、方塊
- 722: 直接功率計算方塊、方塊
- 723: 共變異數矩陣計算方塊、方塊、計算
- 724: 目標共變異數矩陣計算方塊、目標共變異數矩陣計算器、導出
- 725: 混合矩陣計算方塊、混合矩陣
- 725a: 方塊、混合矩陣、混合矩陣計算方塊、導出
- 725b: 方塊、導出
- 726: 輸入共變異數矩陣計算方塊、方塊、導出
- 727: 渲染方塊、方塊、應用
- 730: 方塊
- 733: 方塊
- 735: 方塊
- 737: 方塊
- 739: 方塊
- 741: 擴散信號計算器、決定

751: 步驟、方塊、分解
752: 步驟、分解、執行
753: 步驟、計算
754: 步驟、方塊
755: 步驟
756: 步驟、執行
757: 步驟
758: 方塊、步驟
810: 方向資訊
812: 方塊、欄位
814: 方塊
816: 欄位
818: 欄位
999a: 時間平均元件
999b: 時間平均元件
1000: 頻帶濾波器
1001: 能量估計器
1002: 強度估計器
1003: 擴散計算器
1004: 方向計算器、方塊
1005: 頻帶濾波器
1006: 虛擬麥克風階段
1007: 方塊
1008: 方塊
1009: 方塊
1010: 方塊
1011: VBAP(向量基幅度平移)增益表
1012: 揚聲器增益平均階段

- 1013: 再歸一化器
- 1014: 擴散信號分支
- 1015: 直接信號分支、分支
- 1016: 去相關器、分支
- 1017: 組合器
- 1018: 組合器
- 1019: 揚聲器

【發明申請專利範圍】

【請求項1】 一種用於編碼多個音頻對象的設備，包含：

一對象參數計算器，其被配置為針對與一時間幀相關的多個頻率柱中的一個以上之頻率柱計算至少兩個相關音頻對象的參數資料，其中該至少兩個相關音頻對象的數量低於該多個音頻對象的總數量，以及
一輸出介面，用於輸出一編碼音頻信號，其包括關於一個以上之頻率柱的至少兩個相關音頻對象的參數資料的資訊。

【請求項2】 如請求項1所述之設備，其中該對象參數計算器被配置為將該多個音頻對象中的各該音頻對象轉換為具有該多個頻率柱的一頻譜表示，針對該一個以上之頻率柱，從各該音頻對象計算一選擇資訊，及基於該選擇資訊導出對象標識以作為指示該至少兩個相關音頻對象的該參數資料，以及

其中，該輸出介面被配置為將該對象標識的資訊導入該編碼音頻信號中。

【請求項3】 如請求項1所述之設備，其中該對象參數計算器被配置為量化和編碼一個以上之幅度相關量度或者從該一個以上之頻率柱中的該等相關音頻對象的該幅度相關量度中導出的一個以上之組合數值，以及

其中，該輸出介面被配置為將量化的該一個以上之幅度相關量度或量化的該一個以上之組合數值導入該編碼音頻信號中。

【請求項4】 如請求項2所述之設備，

其中該選擇資訊是與幅度相關的量度(其可為一幅度值、一功率值或一響度值)、或提高到與該音頻對象之功率不同的功率的幅度，以及

其中，該對象參數計算器被配置為計算一組合數值(其可為一相關音頻對象的一幅度相關量度和該相關音頻對象的兩個以上之幅度相關量度之和的比率)，以及

其中，該輸出介面被配置為將該組合數值的資訊導入該編碼音頻信號中，其中該編碼音頻信號中的該組合數值之資訊項目的數量係大於等於1、且小於該一個以上之頻率柱的等相關音頻對象的數量。

【請求項5】 如請求項2所述之設備，

其中，該對象參數計算器被配置為基於該一個以上之頻率柱中的該多個音頻對象的該選擇資訊的一順序來選擇該對象標識。

【請求項6】 如請求項2所述之設備，其中該對象參數計算器被配置為計算一信號功率以作為該選擇資訊，

分別針對各該頻率柱，導出對應之該一個以上之頻率柱中具有該等最大信號功率值的該兩個以上之音頻對象的該對象標識，

計算具有該最大信號功率值的該兩個以上之音頻對象的該信號功率之和與具有導出之該對象標識的該等音頻對象中的至少一個的該信號功率之間的功率比，以作為該參數資料，及

量化和編碼該功率比，以及

其中，該輸出介面被配置為將量化和編碼之該功率比導入該編碼音頻信號中。

【請求項7】 如請求項1所述之設備，其中該輸出介面被配置為將下列資訊導入該編碼音頻信號，

一個以上之編碼傳輸聲道；

作為該參數資料，該時間幀中的該多個頻率柱中的該一個以上之頻率柱中的各該頻率柱的該等相關音頻對象的兩個以上之編碼對象標識，以及一個以上之編碼組合數值或編碼幅度相關量度；以及

該時間幀中的各該音頻對象的量化和編碼方向資料，該方向資料對於該一個以上之頻率柱的所有該等頻率柱是恆定的。

【請求項8】 如請求項1所述之設備，其中該對象參數計算器被配置為計算該一個以上之頻率柱中的至少一最主要對象及一第二主要對象的該參數資料，其中該最主要對象及該第二主要對象代表相關的對象，或

其中，該多個音頻對象的數量為三個以上，該多個音頻對象包括一第一音頻對象、一第二音頻對象、及一第三音頻對象，以及

其中，該對象參數計算器被配置為僅以一第一音頻對象群組(其可為該第一音頻對象和該第二音頻對象)作為該相關音頻對象來計算該一個以上之頻率柱中的一第一頻率柱，以及僅以一第二音頻對象群組(其可為該第二音頻對象和該第三音頻對象、或是該第一音頻對象和該第三音頻對象)作為該相關音頻對象來計算該

一個以上之頻率柱中的一第二頻率柱，其中該第一音頻對象群組與該第二音頻對象群組之間至少有一個群組成員是不同的。

【請求項9】 如請求項1所述之設備，其中該對象參數計算器被配置為計算具有一第一時間或頻率解析度的一原始參數資料，並將該原始參數資料組合到具有低於該第一時間或頻率解析度的一第二時間或頻率解析度的一組合參數資料，以及計算關於具有該第二時間或頻率解析度的該組合參數資料的該至少兩個相關音頻對象的該參數資料，或

決定具有與在該多個音頻對象的一時間或頻率分解中使用的一第一時間或頻率解析度不同的一第二時間或頻率解析度的參數頻帶，以及計算用於具有該第二時間或頻率解析度的該參數頻帶的該至少兩個相關音頻對象的該參數資料。

【請求項10】 如請求項1所述之設備，其中該多個音頻對象包括指示關於該多個音頻對象之一方向資訊之一後設資料，以及

其中，該設備更包含：

一降混器，用於降混該多個音頻對象以獲得一個以上之傳輸聲道，其中該降混器被配置為響應於關於該多個音頻對象之該方向資訊對該多個音頻對象進行降混；以及

一傳輸聲道編碼器，用於對該一個以上之傳輸聲道進行編碼以獲得一個以上之編碼傳輸聲道；

其中，該輸出介面被配置為將該一個以上之傳輸聲道導入該編碼音頻信號。

【請求項11】 如請求項10所述之設備，其中該降混器被配置為生成兩個傳輸聲道以作為兩個虛擬麥克風信號，該兩個虛擬麥克風信號被安排在相同的位置並具有不同的方向，或者在相對於一參考位置或方向(其可為一虛擬聽者位置或方向)的兩個不同位置，或

生成三個傳輸聲道以作為三個虛擬麥克風信號，該三個虛擬麥克風信號被安排在相同的位置並具有不同的方向，或者在相對於一參考位置或方向(其可為一虛擬聽者位置或方向)的三個不同位置，或

生成四個傳輸聲道以作為四個虛擬麥克風信號，該四個虛擬麥克風信號被安排在相同的位置並具有不同的方向，或者在相對於一參考位置或方向(其可為一虛擬聽者位置或方向)的四個不同位置，或

其中，該等虛擬麥克風信號為虛擬第一階麥克風信號、或虛擬心形麥克風信號、或虛擬八字形或偶極或雙向麥克風信號、或虛擬定向麥克風信號、或虛擬亞心形麥克風信號、或虛擬單向麥克風信號、或虛擬超心形麥克風信號、或虛擬全向麥克風信號。

【請求項12】 如請求項10所述之設備，其中該降混器被配置為使用對應之該音頻對象的該方向資訊為該多個音頻對象中的各該音頻對象導出各該傳輸聲道的一加權資訊；

使用一特定傳輸聲道的該音頻對象的該加權資訊對相應之該音頻對象進行加權，以獲得該特定傳輸聲道的一對象貢獻，以及

組合來自該多個音頻對象的該特定傳輸聲道的對象貢獻，以獲得該特定傳輸聲道。

【請求項13】 如請求項10所述之設備，

其中，該降混器被配置為將該一個以上之傳輸聲道計算為一個以上之虛擬麥克風信號，該等虛擬麥克風信號被安排在相同的位置並具有不同的方向、或在相對於一參考位置或方向(其可為一虛擬聽者位置或方向)的不同位置，其與該方向資訊相關，

其中，該等不同的位置或方向位於或朝向一中心線的左側、以及位於或朝向該中心線的右側，或者其中該等不同的位置或方向均等或不均等地分佈到水平位置或方向(其可為相對於該中心線的+90度或-90度、或相對於該中心線的-120度、0度和+120度)，或者其中該等不同的位置或方向包括相對於一虛擬聽者所處位置的一水平面的至少一個朝上或朝下的位置或方向，其中關於該多個音頻對象的該方向資訊係相關於該虛擬聽者位置、或該參考位置、或該方向。

【請求項14】 如請求項10所述之設備，更包含：

一參數處理器，用於量化指示關於該多個音頻對象的該方向資訊的該後設資料，以獲得該多個音頻對象的量化方向項目，

其中，該降混器被配置為響應於該量化方向項目作為該方向資訊進行操作，以及

其中，該輸出介面被配置為將該量化方向項目的資訊導入該編碼音頻信號中。

【請求項15】 如請求項10所述之設備，

其中，該降混器被配置為執行關於該多個音頻對象的該方向資訊的一分析，並且根據該分析的一結果放置用於該傳輸聲道之生成的一個以上之虛擬麥克風。

【請求項16】 如請求項10所述之設備，

其中，該降混器被配置為使用在多個時間幀上靜態的一降混規則來進行降混，或

其中，該方向資訊在多個時間幀上是可變的，並且其中該降混器被配置為使用在多個時間幀上可變的一降混規則來進行降混。

【請求項17】 如請求項10所述之設備，其中，該降混器被配置為使用對該多個音頻對象的樣本以逐個樣本加權和組合的方式，在一時域中進行降混。

【請求項18】 一種用於解碼一編碼音頻信號的解碼器，該編碼音頻信號包括多個音頻對象的一個以上之傳輸聲道和方向資訊、及一時間幀的一個以上之頻率柱的至少兩個相關音頻對象的一參數資料，該至少兩個相關音頻對象的數量少於該多個音頻對象的總數，該解碼器包含：

一輸入介面，用於以在該時間幀中具有該多個頻率柱的一頻譜表示來提供該一個以上之傳輸聲道；以及

一音頻渲染器，用於使用該方向資訊將該一個以上之傳輸聲道渲染成數個音頻聲道，藉以被認定為根據與該至少兩個相關音頻對象的一第一相關音頻對象的一第一方向資訊以及與該至少兩個相關音頻對象的一第二相關音頻對象的一第二方向資訊，從該一個以上之傳輸聲道中計算一貢獻，或

其中，該音頻渲染器被配置為該多個頻率柱的各該頻率柱，根據與該至少兩個相關音頻對象的一第一相關音頻對象的一第一方向資訊以及與該至少兩個相關音頻對象的一第二相關音頻對象的一第二方向資訊，從該一個以上之傳輸聲道中計算一貢獻。

【請求項19】 如請求項18所述之解碼器，

其中，該音頻渲染器被配置為對於該一個以上之頻率柱忽略與該至少兩個相關音頻對象不同的一音頻對象的一方向資訊。

【請求項20】 如請求項18所述之解碼器，

其中該編碼音頻信號包括各該相關音頻對象的一幅度相關度量或與該參數資料中的至少兩個相關音頻對象相關的一組合值，以及

其中，該音頻渲染器被配置為根據該幅度相關度量或該組合值來決定該一個以上之傳輸聲道的一定量貢獻。

【請求項21】 如請求項20所述之解碼器，其中該編碼信號包括該參數資料中的該組合值，以及

其中，該音頻渲染器被配置為使用該等相關音頻對象其中之一的該組合值和該相關音頻對象的該方向資訊來決定該一個以上之傳輸聲道的該貢獻，以及

其中，該音頻渲染器被配置為使用從該一個以上之頻率柱中的該等相關音頻對象其中之另一的該組合值和該另一相關音頻對象的該方向資訊所導出的一值來決定該一個以上之傳輸聲道的該貢獻。

【請求項22】 如請求項18所述之解碼器，其中該音頻渲染器被配置為從該多個頻率柱中的各該頻率柱的該等相關音頻對象和與等頻率柱中的該等相關音頻對象相關聯的該方向資訊，計算一直接響應資訊。

【請求項23】 如請求項22所述之解碼器，

其中，該音頻渲染器被配置為使用一擴散資訊(其可為包括在該後設資料中的一擴散參數)或一去相關規則來決定該多個頻率柱中的各該頻率柱的一擴散信號，並且組合該擴散信號與由該直接響應資訊所決定之一直接響應，以獲得用於該數個聲道其中之一聲道的一頻譜域渲染信號，或

使用該直接響應資訊和該數個音頻聲道的資訊來計算一合成資訊，並且將該共變異數合成資訊應用於該一個以上之傳輸聲道以獲得該數個音頻聲道，或

其中，該直接響應資訊是各該一個以上之音頻對象的一直接響應向量，並且其中該共變異數合成資訊是一共變異數合成矩陣，並且其中該音頻渲染器被配置為利用該共變異數合成資訊對每一頻率柱執行一矩陣運算。

【請求項24】 如請求項22所述之解碼器，其中該音頻渲染器被配置為

在該直接響應資訊的計算中，導出該一個以上之音頻對象的一直接響應向量，並為該一個以上之音頻對象從各該直接響應向量計算一共變異數矩陣，在該共變異數合成資訊的計算中從以下推導出一目標共變異數資訊：該一個音頻對象的該共變異數矩陣或該多個音頻對象的該等共變異數矩陣，相應之該一個以上之音頻對象的一功率資訊，以及從該一個以上之傳輸聲道導出的一功率資訊。

【請求項25】如請求項24所述之解碼器，其中該音頻渲染器被配置為在該直接響應資訊的計算中，導出該一個以上之音頻對象的一直接響應向量，並為各該一個以上之音頻對象從各該直接響應向量計算一共變異數矩陣，從該傳輸聲道導出一輸入共變異數資訊，以及從該目標共變異數資訊、該輸入共變異數資訊和關於該數個音頻聲道之資訊導出一混合資訊，以及將該混合資訊應用到該時間幀中的各該頻率柱的該等傳輸聲道。

【請求項26】如請求項25所述之解碼器，其中將該混合資訊應用到該時間幀中的各該頻率柱的結果轉換到一時域中以獲得該時域中的該數個音頻聲道。

【請求項27】如請求項22所述之解碼器，其中該音頻渲染器被配置為在從該等傳輸聲道導出的一輸入共變異數矩陣的一分解中，僅使用該輸入共變異數矩陣的主對角元素，或使用該等對象或該等傳輸聲道的一直接響應矩陣和一功率矩陣，執行一目標共變異數矩陣的一分解，或通過取該輸入共變異數矩陣的各該主對角元素的根來執行該輸入共變異數矩陣的一分解，或計算已分解之該輸入共變異數矩陣的一正規化逆矩陣，或執行一奇異值分解以在沒有一擴展單位矩陣的情況下計算用於一能量補償的一最佳矩陣。

【請求項28】一種用於對多個音頻對象進行編碼的方法，包括：

針對與一時間幀相關的多個頻率柱中的一個以上之頻率柱計算至少兩個相關音頻對象的參數資料，其中該至少兩個相關音頻對象的數量低於該多個音頻對象的總數量，以及

輸出一編碼音頻信號，其包括關於該一個以上之頻率柱的該至少兩個相關音頻對象的參數資料的資訊。

【請求項29】一種用於解碼一編碼音頻信號的方法，該編碼音頻信號包括多個音頻對象的一個以上之傳輸聲道和方向資訊、及一時間幀的一個以上之頻率柱的至少兩個相關音頻對象的一參數資料，該至少兩個相關音頻對象的數量少於該多個音頻對象的總數，該方法包括：

以在該時間幀中具有該多個頻率柱的一頻譜表示來提供該一個以上之傳輸聲道；以及

使用該方向資訊將該一個以上之傳輸聲道音頻渲染成數個音頻聲道，

其中，該音頻渲染之步驟包括為該多個頻率柱的各該頻率柱，根據與該至少兩個相關音頻對象的一第一相關音頻對象的一第一方向資訊以及與該至少兩個相關音頻對象的一第二相關音頻對象的一第二方向資訊，從該一個以上之傳輸聲道中計算一貢獻，或是被認定為根據與該至少兩個相關音頻對象的一第一相關音頻對象的一第一方向資訊以及與該至少兩個相關音頻對象的一第二相關音頻對象的一第二方向資訊，從該一個以上之傳輸聲道中計算一貢獻。

【請求項30】一種對多個音頻對象進行編碼或解碼的電腦程式，當其運行於一電腦或一處理器時，用以執行如請求項28所述之方法或如請求項29所述之方法。

【請求項31】一種編碼音頻信號之資料結構產品，其中該編碼音頻信號代表多個音頻對象，其中該編碼音頻信號包含針對與一時間幀相關的多個頻率柱中的一個以上之頻率柱的至少兩個相關音頻對象的參數資料的資訊，其中該至少兩個相關音頻對象的數量低於該多個音頻對象的總數量。

【請求項32】如請求項31所述之編碼音頻信號之資料結構產品，其中該編碼音頻信號更包含：

一個以上之編碼傳輸聲道；

作為該參數資料，該時間幀中的該多個頻率柱中的該一個以上之頻率柱中的各該頻率柱的該等相關音頻對象的兩個以上之編碼對象標識，以及一個以上之編碼組合數值或編碼幅度相關量度；以及
該時間幀中的各該音頻對象的量化和編碼方向資料，該方向資料對於該一個以上之頻率柱的所有該等頻率柱是恆定的。

【發明圖式】

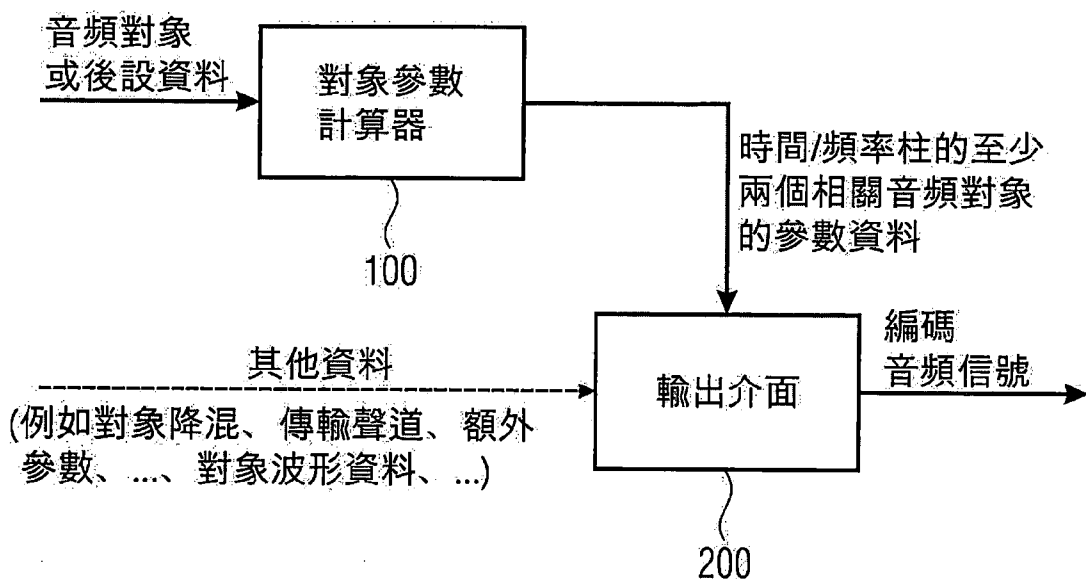


圖 1a

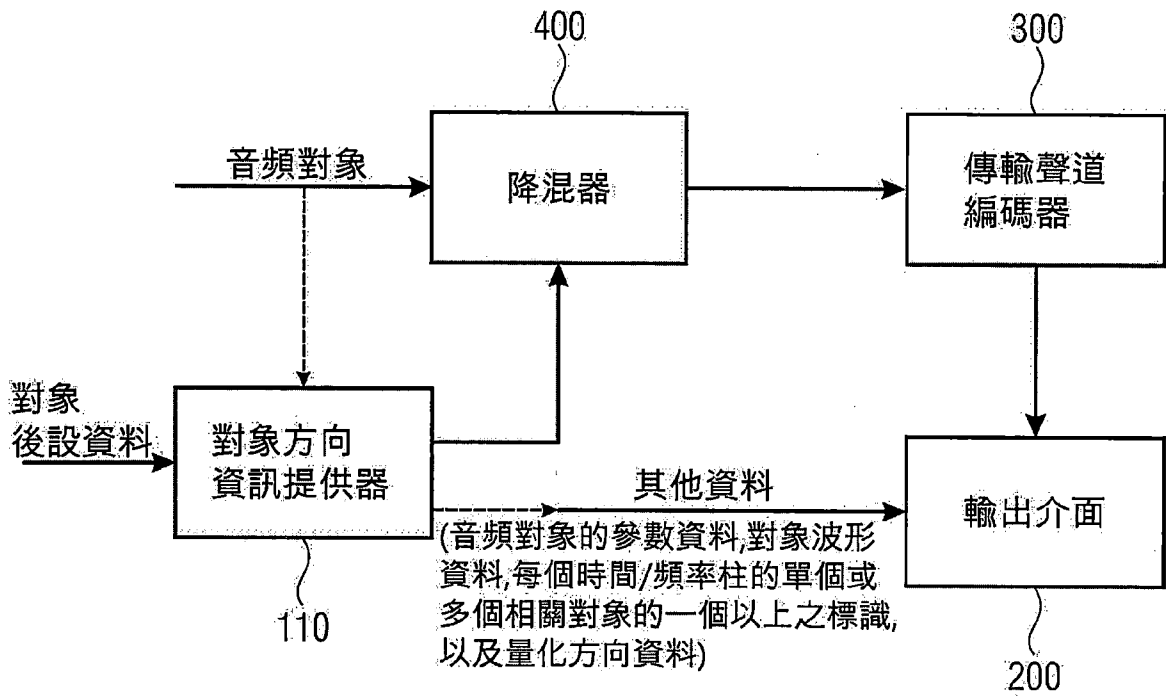


圖 1b

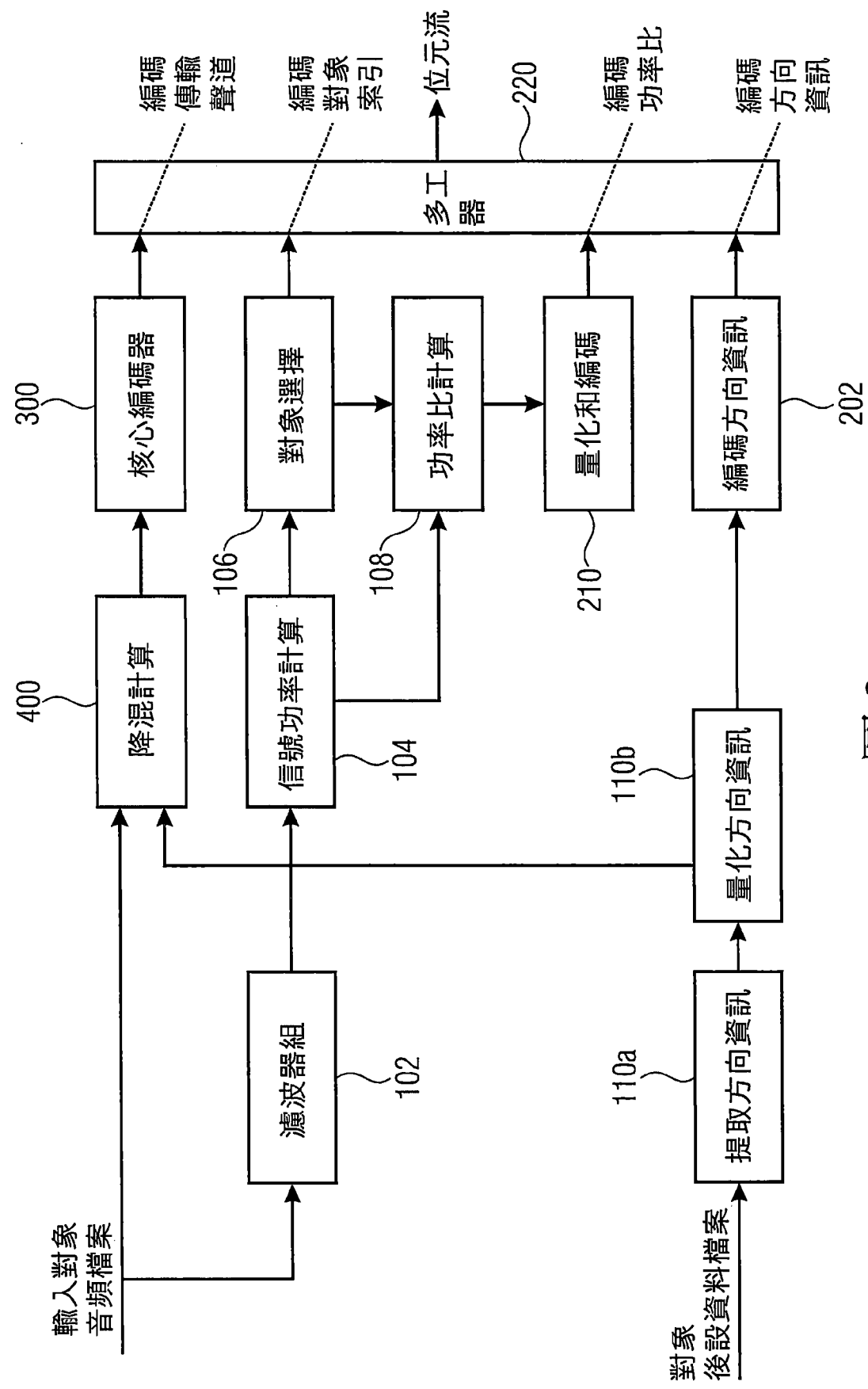


圖 2

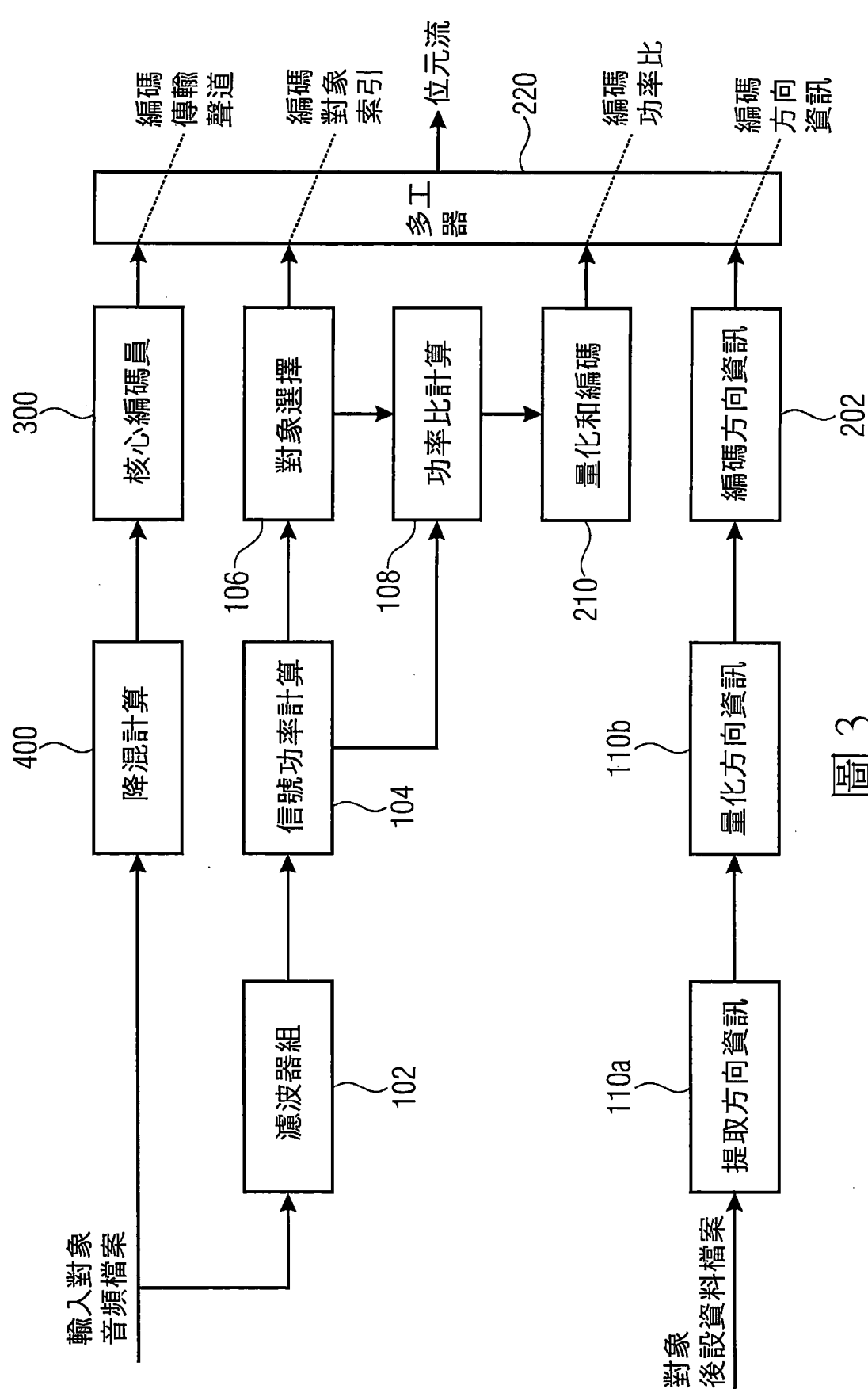


圖 3

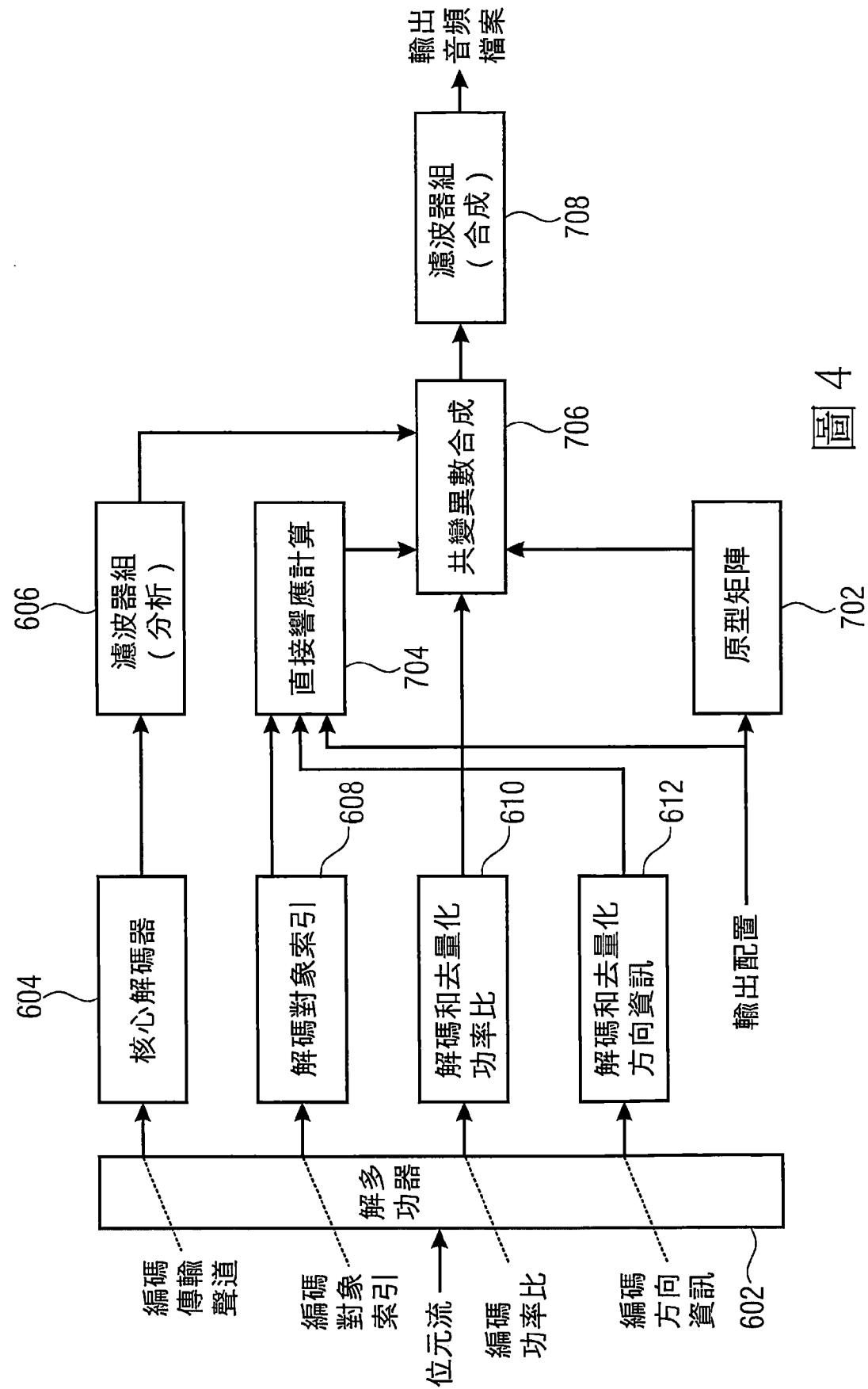


圖 4

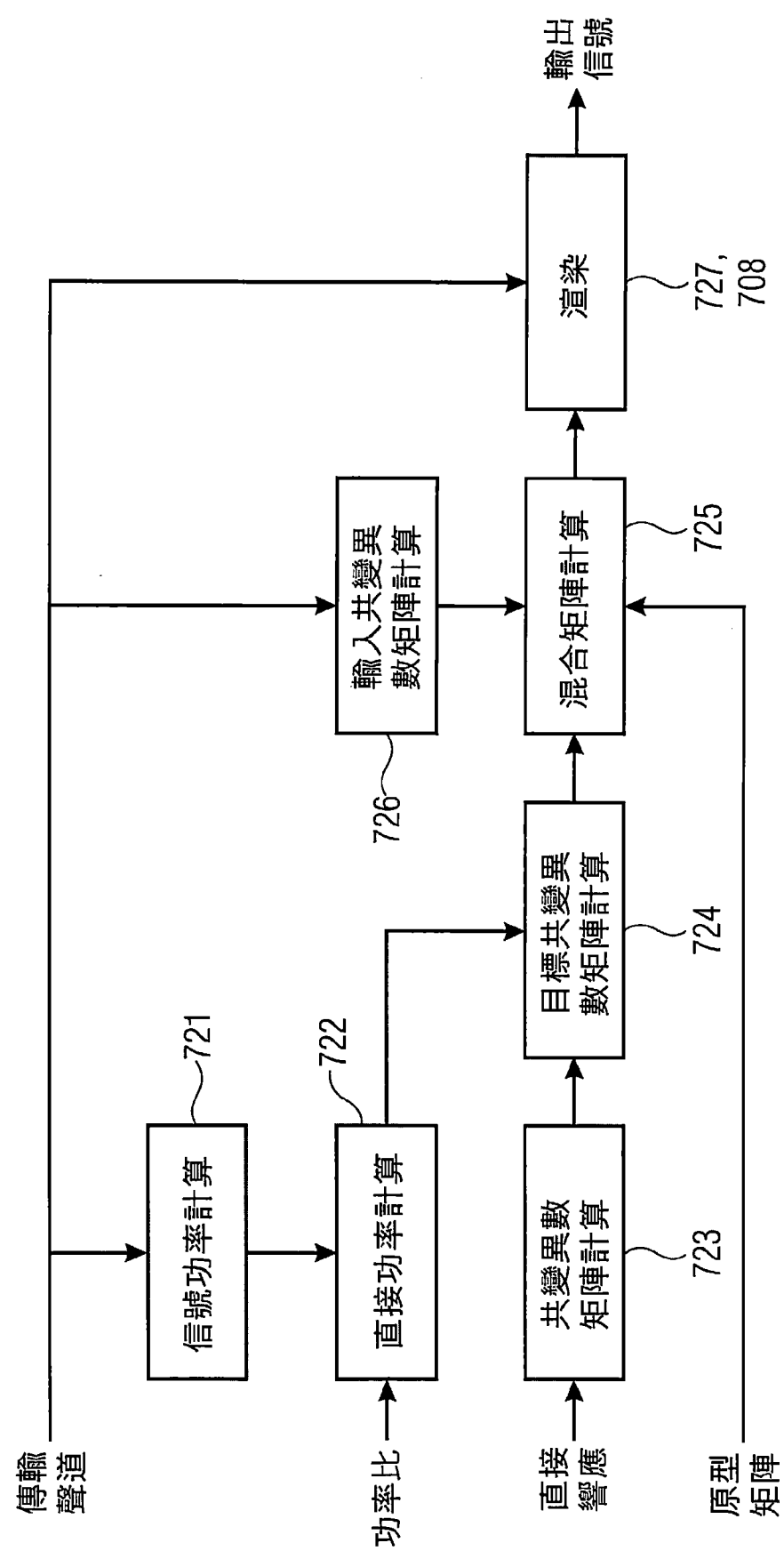


圖 5

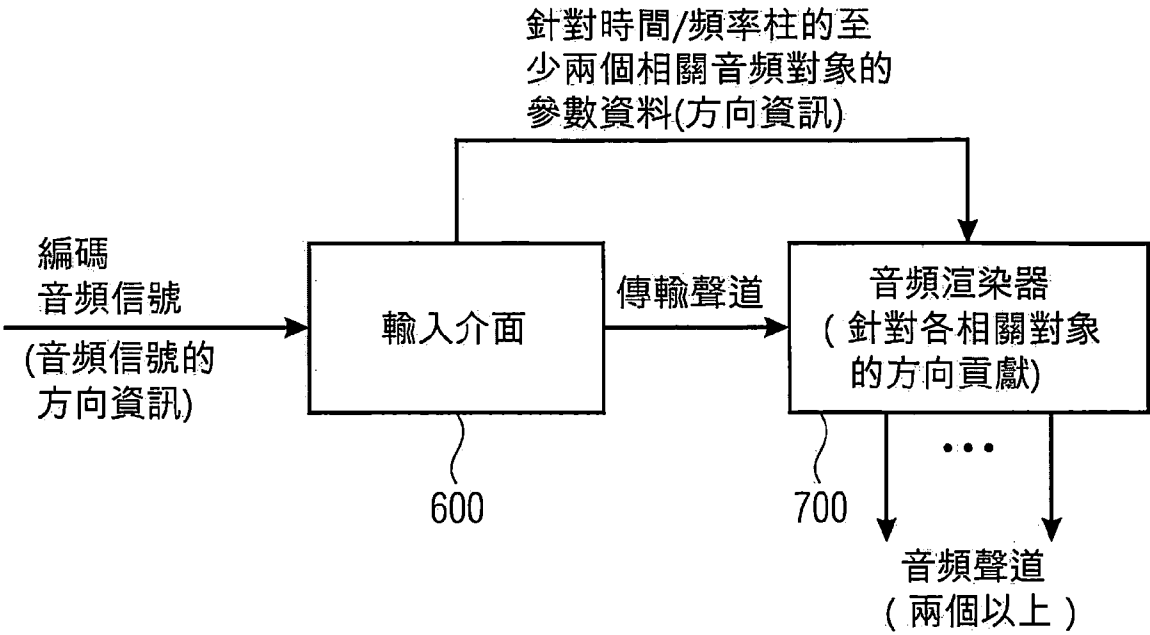


圖 6a

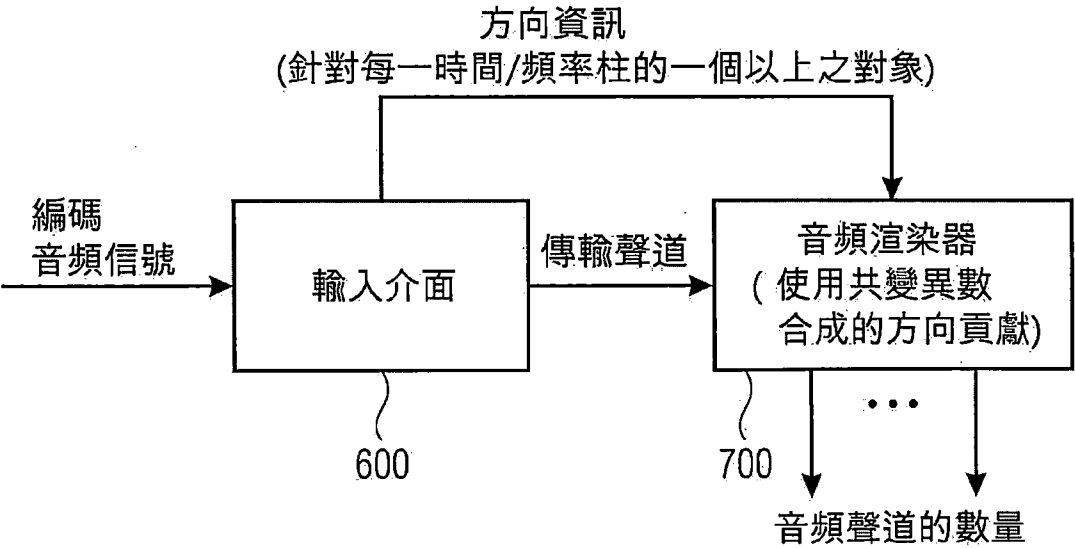


圖 6b

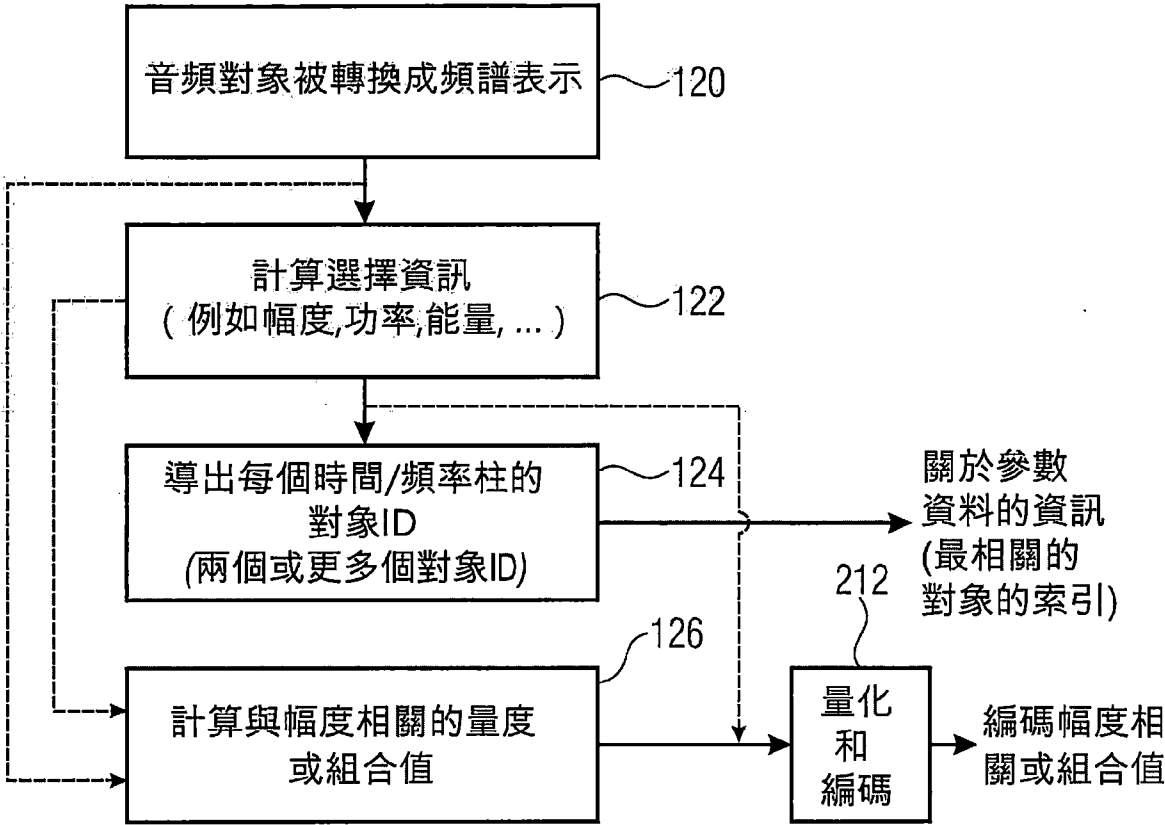


圖 7a

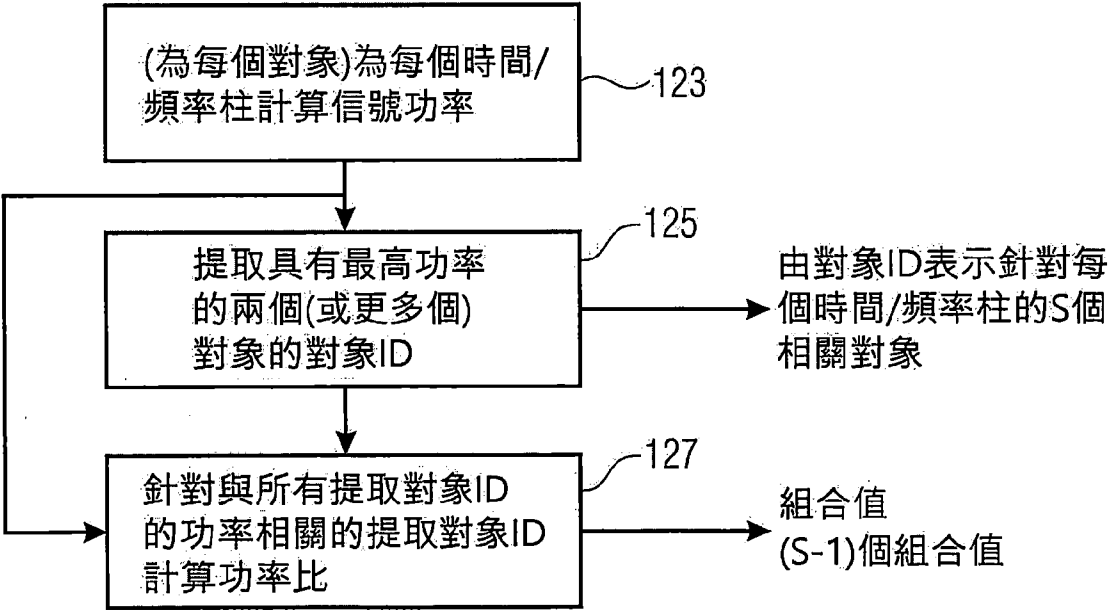


圖 7b

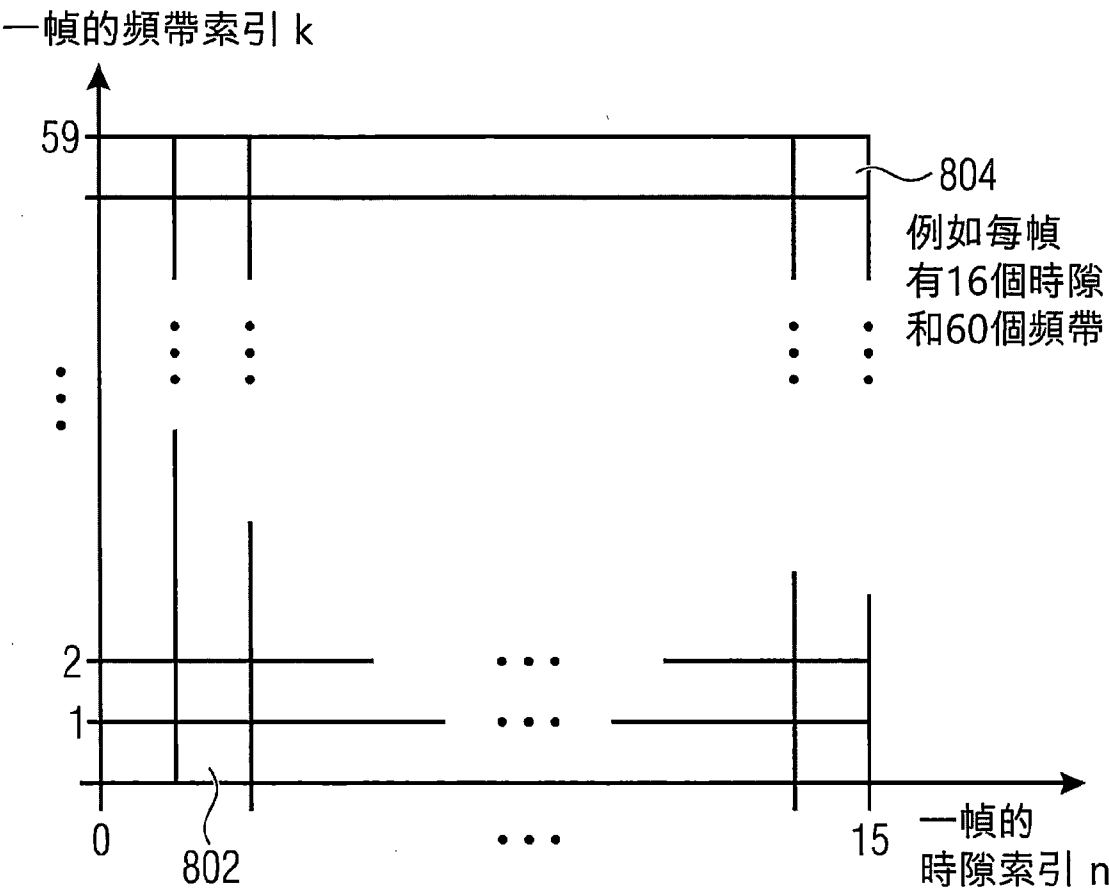


圖 8a

(高解析度濾波器組時間/頻率檔案)

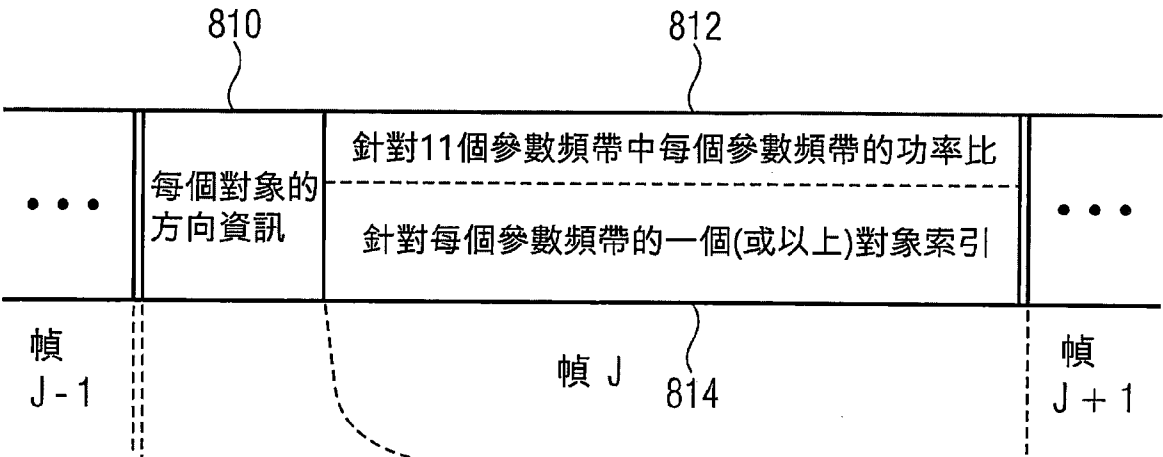


圖 8b
(針對每幀的參數資料)

對象 ID	方向資訊
1	方位角/仰角
2	方位角/仰角
3	方位角/仰角
4	
5	
⋮	⋮
N-1	
N	

圖 8c

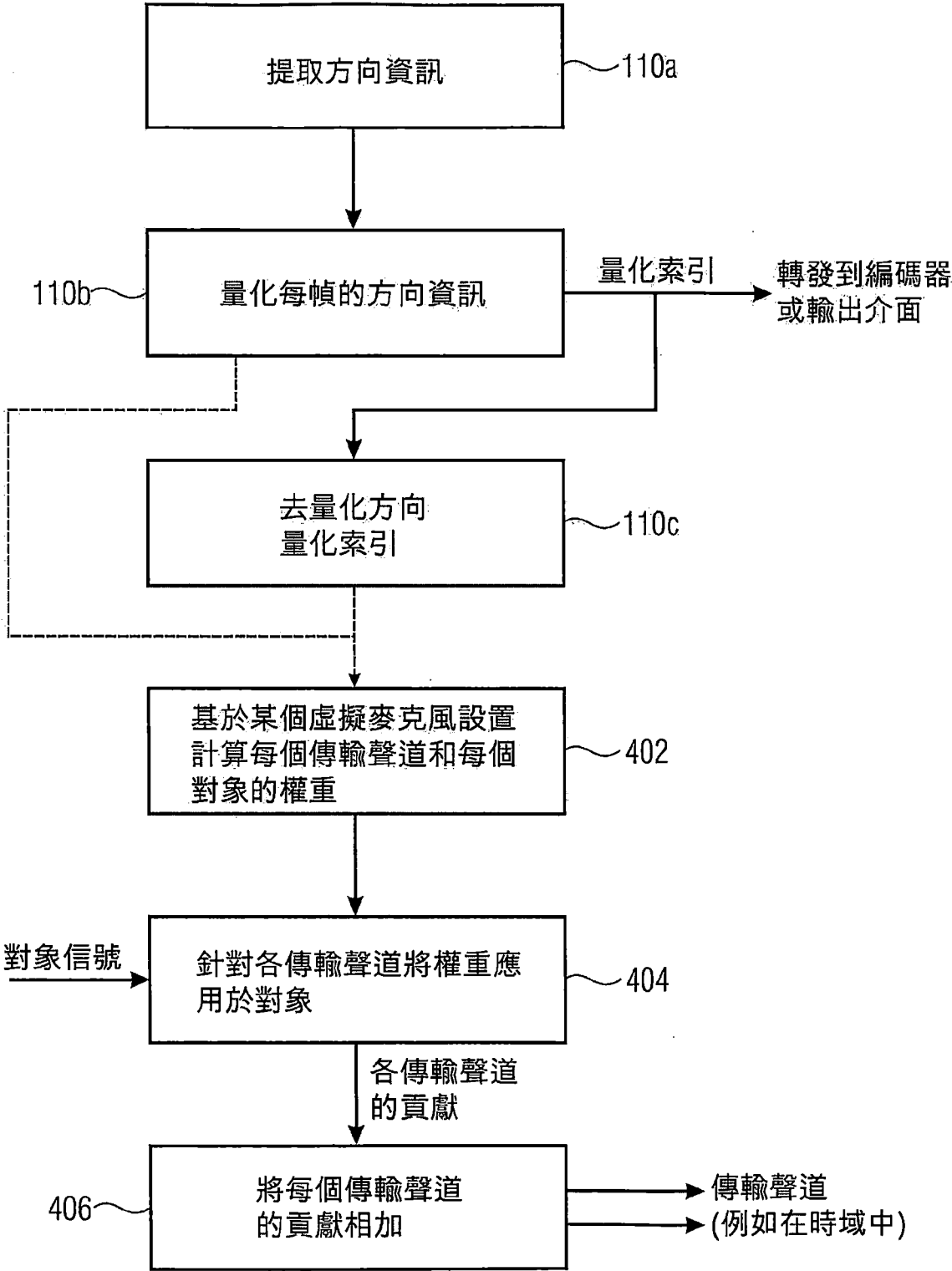


圖 9a

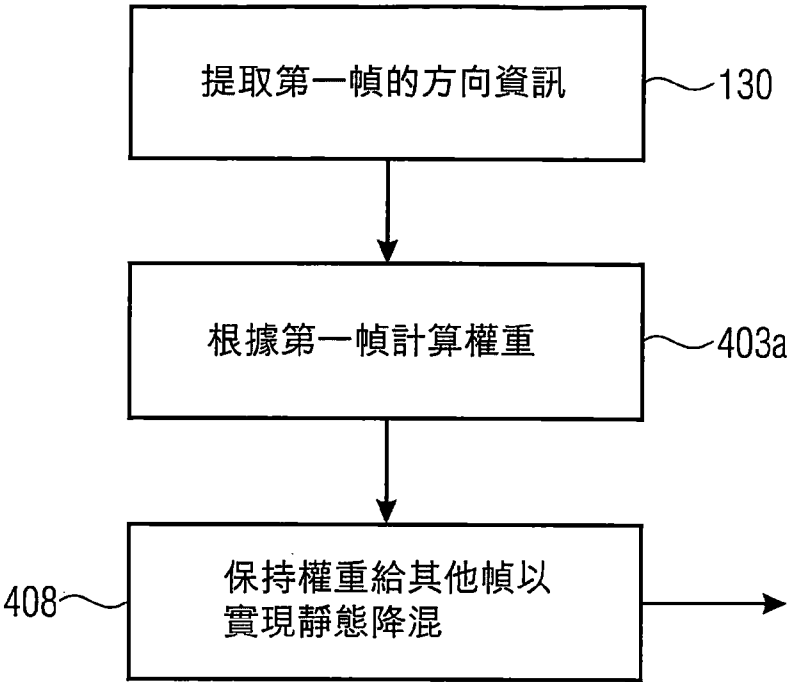


圖 9b

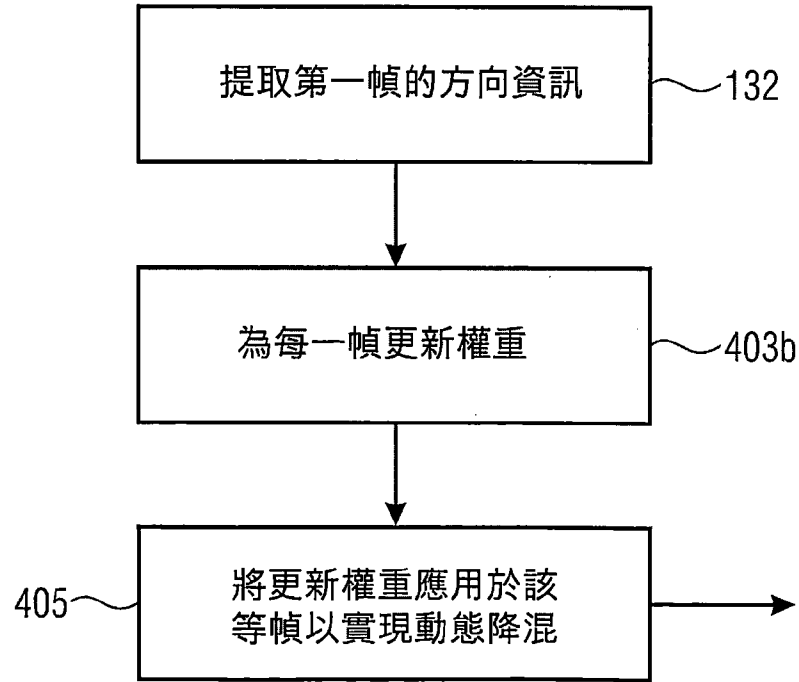


圖 9c

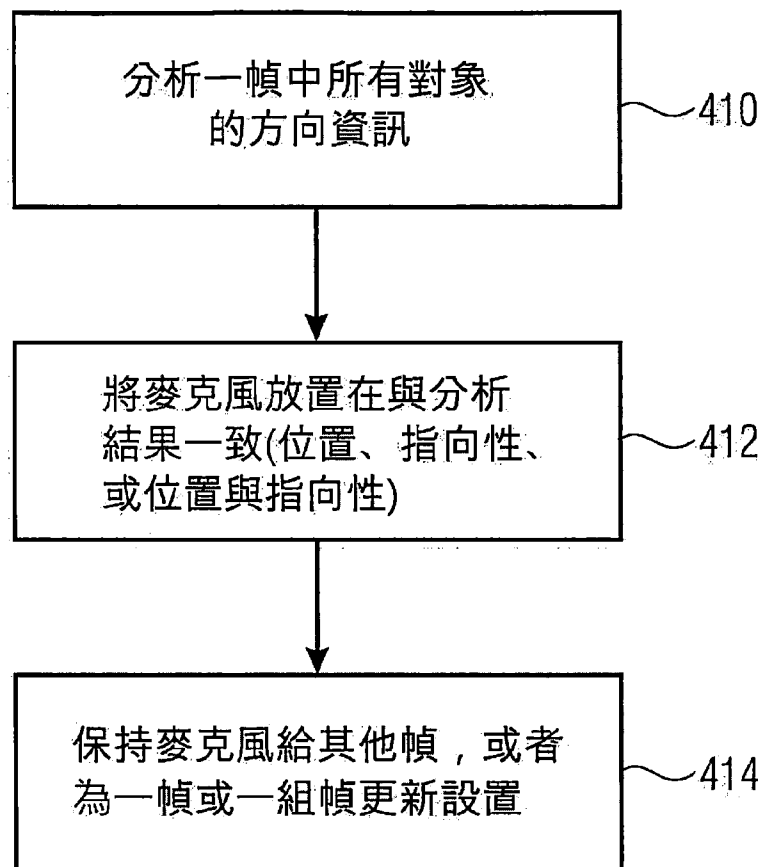


圖 9d

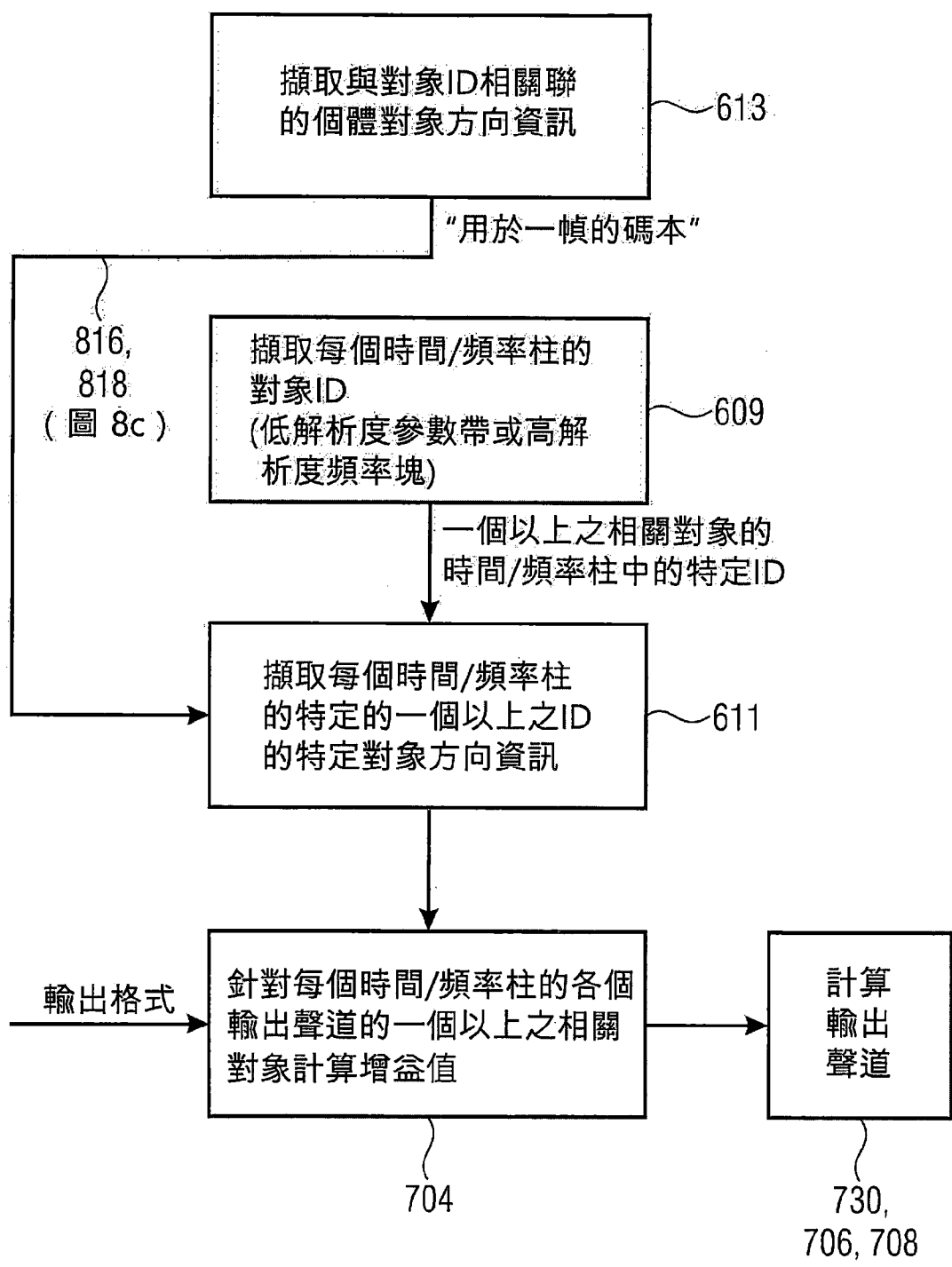


圖 10a

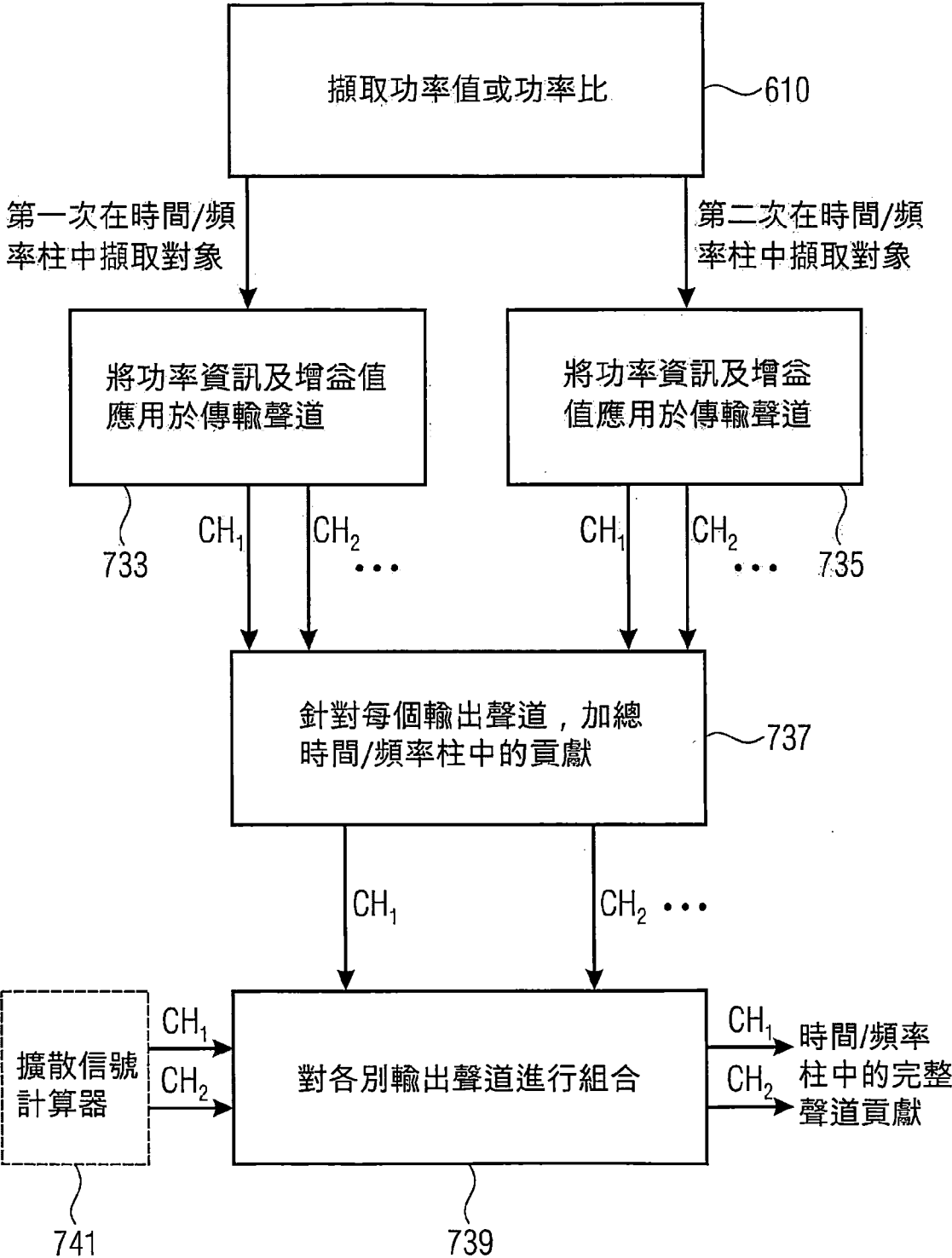


圖 10b

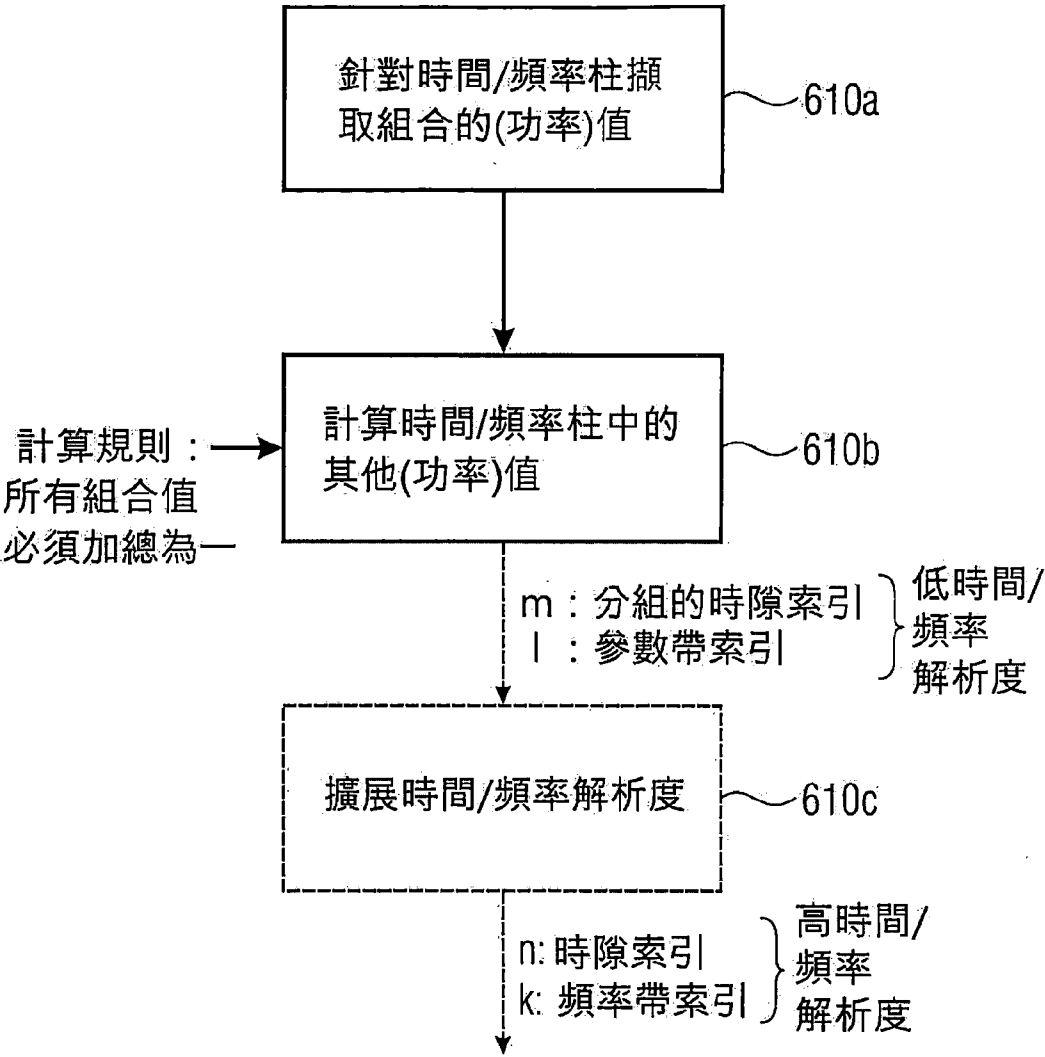


圖 10c

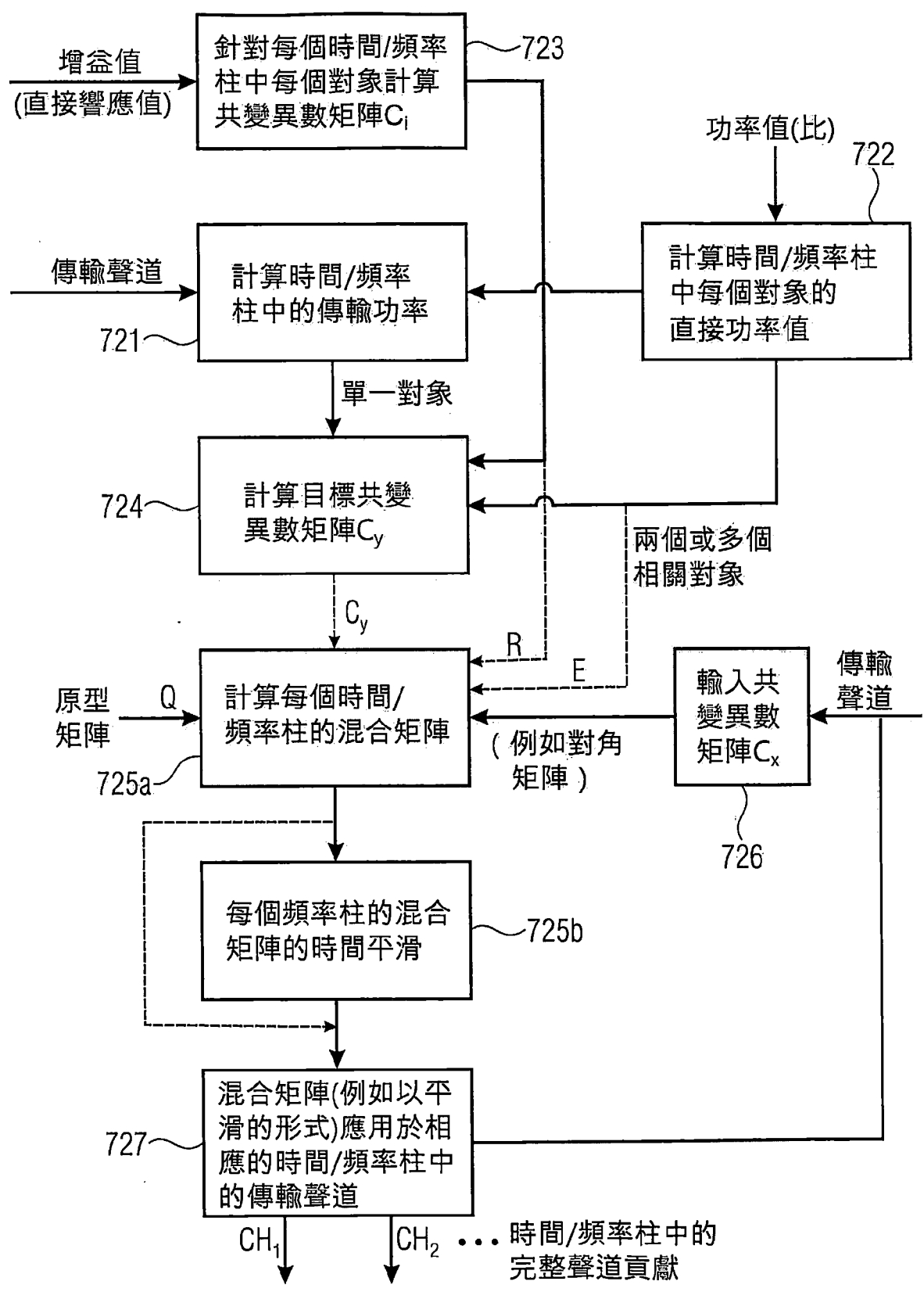


圖 10d

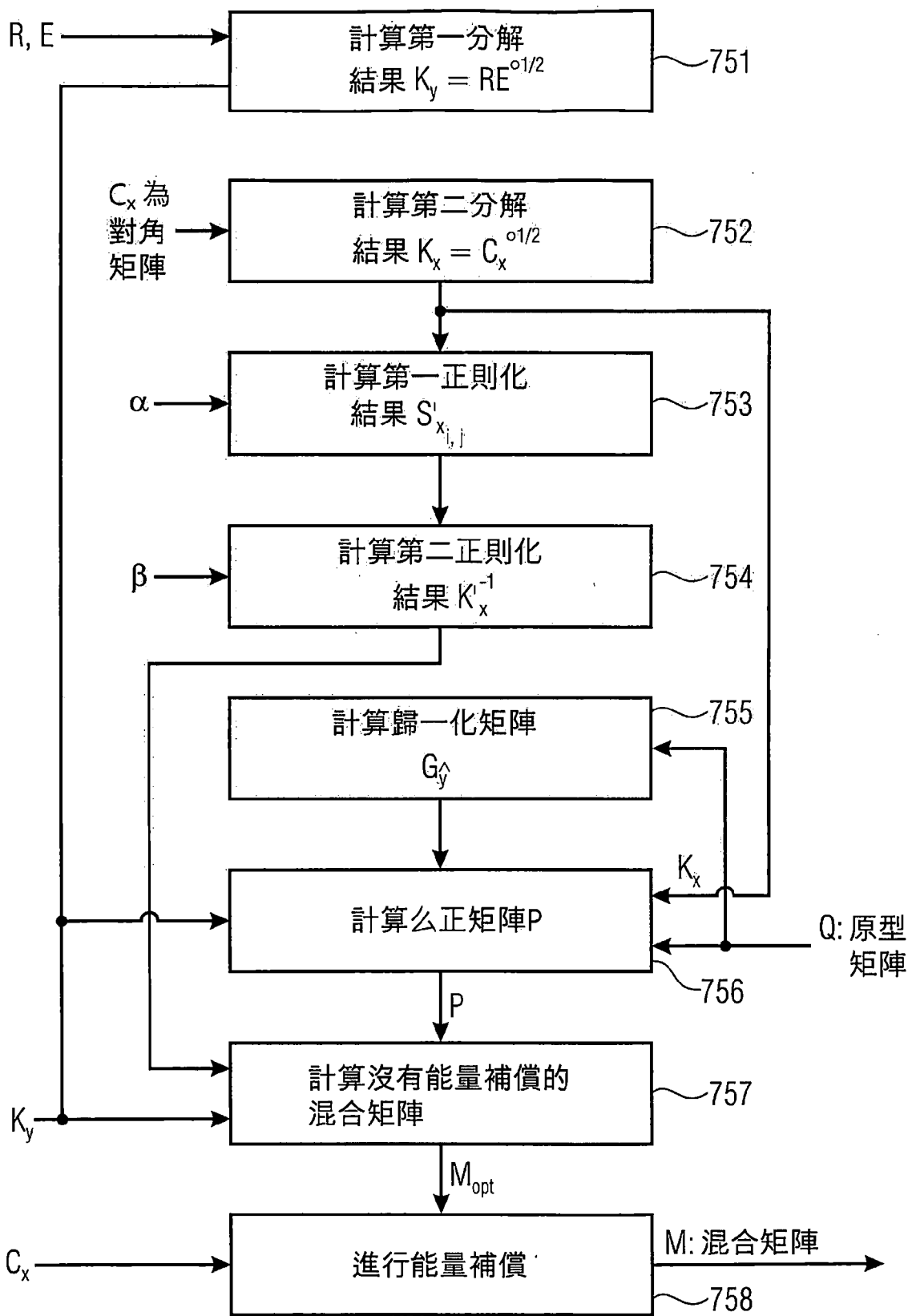


圖 11

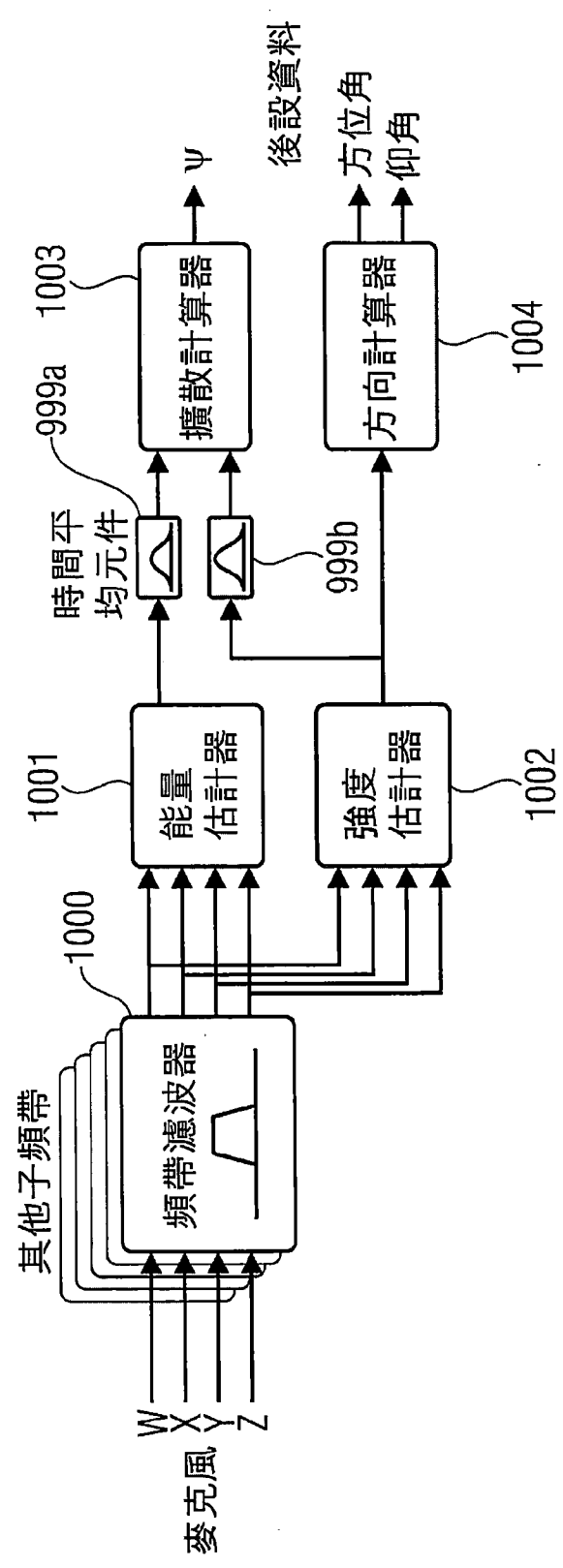


圖 12a
(先前技術)

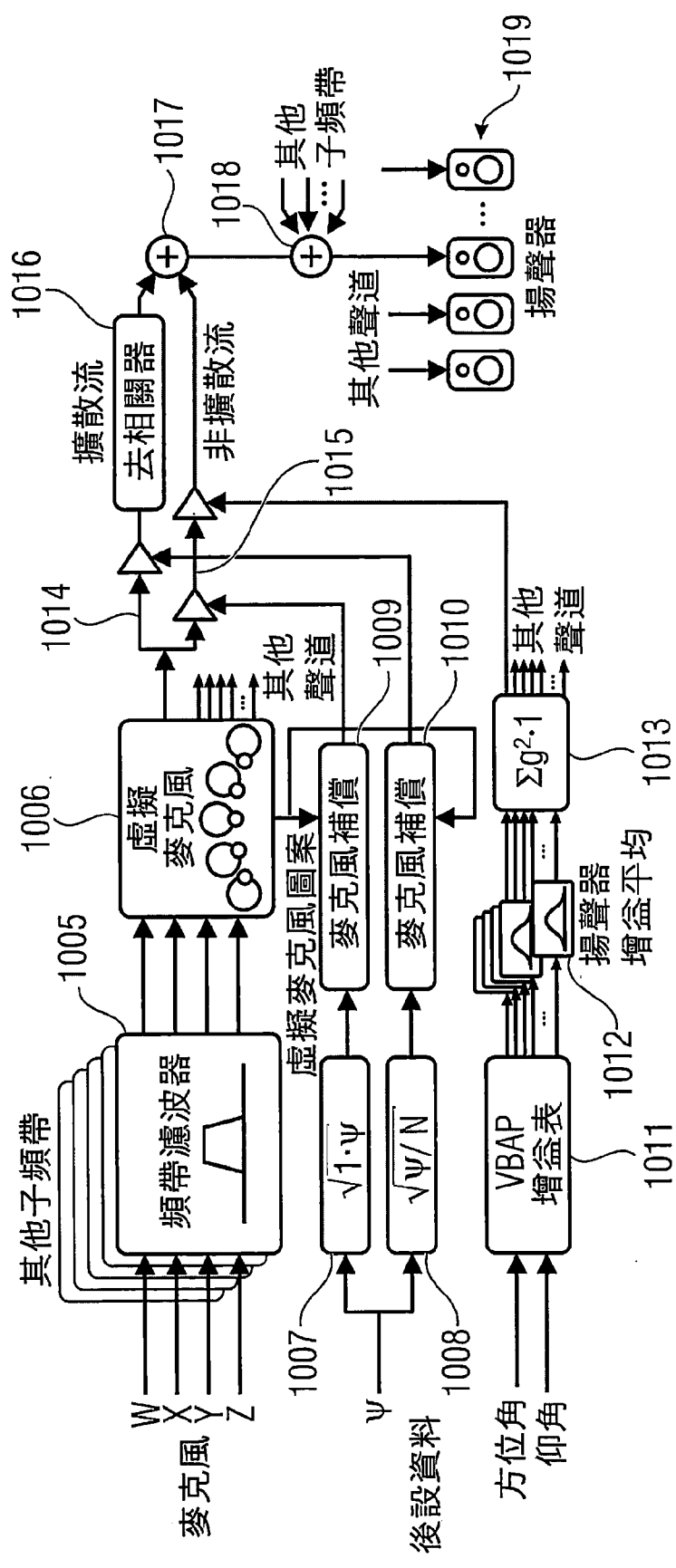


圖 12b
(先前技術)