



(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 I681386 B

(45)公告日：中華民國 109 (2020) 年 01 月 01 日

(21)申請案號：107135270

(22)申請日：中華民國 102 (2013) 年 12 月 05 日

(51)Int. Cl. : **G10L19/008 (2013.01)****H04S3/00 (2006.01)**

(30)優先權：2012/12/12 歐洲專利局

12306569.0

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

荷蘭

(72)發明人：克魯格 亞歷山大 KRUGER, ALEXANDER (DE)；科登 斯凡 KORDON, SVEN (DE)；波漢 約哈拿斯 BOEHM, JOHANNES (DE)

(74)代理人：林志剛

(56)參考文獻：

EP 2469742A2

US 2010/0329466A1

審查人員：黃衍勳

申請專利範圍項數：2 項 圖式數：6 共 37 頁

(54)名稱

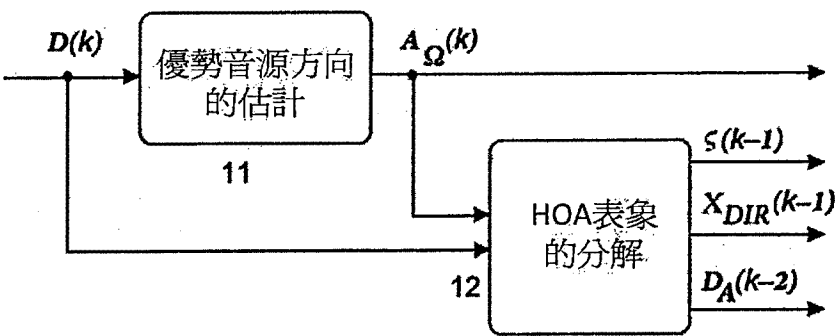
用於音場之高階保真立體音響表示的壓縮與解壓縮方法及裝置

(57)摘要

本發明改善了高階保真立體音響(HOA)表示的壓縮。上述高階保真立體音響(HOA)表示係在優勢音源的存在下被分析且其方向均經估計。隨後，HOA 表示係經分解為一些優勢方向訊號與一殘餘分量。為了於均勻取樣方向上得到一般平面波函數，此殘餘分量係經轉換為離散空間域，其係預測自優勢方向訊號。最終，預測誤差會被轉換回 HOA 域以及表示殘餘周圍 HOA 分量，其中係實施一位階降低，並緊接著進行優勢方向訊號與殘餘分量的感知編碼。

The invention improves HOA sound field representation compression. The HOA representation is analyzed for the presence of dominant sound sources and their directions are estimated. Then the HOA representation is decomposed into a number of dominant directional signals and a residual component. This residual component is transformed into the discrete spatial domain in order to obtain general plane wave functions at uniform sampling directions, which are predicted from the dominant directional signals. Finally, the prediction error is transformed back to the HOA domain and represents the residual ambient HOA component for which an order reduction is performed, followed by perceptual encoding of the dominant directional signals and the residual component.

指定代表圖：



符號簡單說明：

11 . . . 優勢音源方
向的估計

12 . . . HOA 表示的
分解

第一 A圖

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

用於音場之高階保真立體音響表示的壓縮與解壓縮方法及裝置

METHOD AND APPARATUS FOR COMPRESSING AND
DECOMPRESSING A HIGHER ORDER AMBISONICS
REPRESENTATION FOR A SOUND FIELD

【技術領域】

[0001] 本發明係關於一種用於音場之高階保真立體音響表示的壓縮與解壓縮方法及裝置。

【先前技術】

[0002] 高階保真立體音響(以下稱 HOA)提供一種表示三維聲音的方法。其他技術則為波場合成(Wave Field Synthesis, WFS)或以頻道為基礎的方法如 22.2。相較於以頻道為基礎的方法，HOA 表示的優點在於不需仰賴特殊揚聲器設置。然而，此項適用性是以解碼過程為代價，需在特別的揚聲器設置上回放 HOA 表示。相較於所需揚聲器之數量通常非常龐大的波場合成方法，HOA 亦可被提供予僅由少數揚聲器組成之設置。HOA 之另一優點在於相同的表示亦可在不作任何修改之下被應用於頭戴式耳機之雙耳演示技術(binaural rendering)。

[0003] HOA 係基於複諧平面波振福(complex harmonic

plane wave amplitudes)之空間密度之一表示而藉由截頭球諧展開。每一展開係數係為角頻率之一函數，其係可等效地藉由一時域函數表示。因此，不失一般性，完整 HOA 音場表示實際上可被假設為由 O 時域函數所組成，在此處 O 代表展開係數值。這些時域函數在後述會被相同地稱作為 HOA 係數序列。

[0004] HOA 表示的空間解析度係因展開之一最大位階 N 而改善。可惜，展開係數的數值 O 係以位階 N 而二次地成長，即 $O=(N+1)^2$ 。舉例來說，例如使用位階 $N=4$ 之典型 HOA 表示，需 $O=25$ HOA(展開)係數。根據上述考量，賦予所需採樣率 f_s 和每樣本之位元數 N_b ，即可由 $O \cdot f_s \cdot N_b$ 決定 HOA 訊號表示傳輸之總位元率。而以採用每樣本 $N_b=16$ 位元之採樣率 $f_s=48\text{kHz}$ 傳輸位階 $N=4$ 的 HOA 訊號表示會產生一 19.2Mbits/s 之位元率，其對於許多實際的應用(例如：串流)來說是非常高的，因此 HOA 表示的壓縮是極度被需要的。

【發明內容】

[0005] 已知方法相當罕見以 $N>1$ 壓縮 HOA 表示。其中之一採用感知進步聲訊編碼法(AAC)寫解碼器，進行直接編碼個別 HOA 係數序列，參見 E. Hellerud, I. Burnett, A. Solvang, U. Peter Svensson 合撰〈以 AAC 編碼高階保真立體音響〉，2008 年阿姆斯特丹第 124 次 AES 會議。然而，具有如此措施之固有問題是，從未聽到訊號的感知

編碼。重建之回放訊號，通常是由 HOA 係數序列加權合計而得。這是解壓縮 HOA 表示描繪在特別揚聲器設置時，有揭露感知編碼雜訊高度或然之原因所在。以更技術性而言，感知編碼雜訊表露之主要問題是，個別 HOA 係數序列間之高度交叉相關性。因為個別 HOA 係數序列內所編碼雜訊訊號，通常彼此不相關，會發生感知編碼雜訊之構成性重疊，同時，無雜訊 HOA 係數序列在重疊時取消。又一問題是，上述交叉相關性導致感知編碼器效率降低。為了將此兩者效應減至最小，EP 2469742 A2 擬議在感知編碼之前將 HOA 表示轉換成離散空間域內之等效表示。形式上，該等離散空間域係等同於複諧平面波震幅之空間密度的時域，其係於一些離散方向上取樣。該離散空間域訊號係因此以 o 習知時域訊號來表示，其可被解釋如來自取樣方向之一般平面波，且如果擴音器位在空間域轉換所假設之正確同樣方向，其亦相當於擴音器訊號。

[0006] 轉換成離散空間域，會減少個別空間域訊號間的交叉相關性。然而，交叉相關性並未完全消除。較高交叉相關性之例為方向性訊號，其方向落在空間域訊號涵蓋的相鄰方向之中間。

[0007] 上述方法之一主要缺點在於感知編碼訊號數為 $(N+1)^2$ ，且被壓縮 HOA 表示之資料率係以保真立體音響位階 N 呈二次方成長。

[0008] 為了降低感知編碼訊號數，歐洲專利申請案 EP 2665208 A1 提出將 HOA 表示解壓縮為優勢方向訊號之

一預定最大值以及一殘餘周圍分量。待感知編碼之訊號數的減少可經由降低殘餘周圍分量的位階數來達成。此方法背後的基礎原理在於當藉由一較低位階的 HOA 表示表示具有足夠準確性的殘餘時，相對優勢方向訊號保留一高空間解析度。

[0009] 只要滿足在音場上的假設，此方法便可運作的相當良好，即其係由少數優勢方向訊號(代表一般以完全位階 N 編碼的平面波函數)以及一不具方向性之殘餘周圍分量組成。然而，若接下來分解，該殘餘周圍分量仍包含一些優勢方向訊號，降階會導致誤差，其在表示接下來之解壓縮方面無疑地為可感知的。違反假設之 HOA 表示的典型例子就是以低於 N 的位階進行編碼之一般平面波。為了使音源表示更寬，此種位階低於 N 的一般平面波可由藝術創作 artistic creation 而產生，且易可藉由球形麥克風而與 HOA 音場表示的紀錄一併產生。在兩例子中，音場係以大量的高相關空間域訊號來表示(其解釋亦見「高階保真立體音響之空間解析度」一節)。

[0010] 本發明欲解決之一問題在於消除歐洲專利申請案 EP 2665208 A1 中所述流程衍生的缺點，因此也避免了上述其他引用之習知文件中的缺點。此問題係藉由申請專利範圍第 1 與 3 項所揭露之方法來解決。使用這些方法之相對應裝置係揭露於申請專利範圍第 2 與 4 項中。

[0011] 本發明改善了描述於歐洲專利申請案 EP 2665208 A1 中的 HOA 音場表示壓縮過程。首先，如同在

EP 2665208 A1 中，HOA 表示係對於優勢音源之存在而被分析，於其中該些方向係經估計的。以所知之優勢音源方向，HOA 表示係被分解為一些代表一般平面波之優勢方向訊號以及一殘餘分量。然而，取代直接降低此殘餘 HOA 分量之位階，其係經轉換為離散空間域以於代表殘餘 HOA 分量之均勻取樣方向上得到一般平面波函數。之後，自優勢方向訊號預測這些平面波函數。此操作之理由是在於部份殘餘 HOA 分量係可能與優勢方向訊號高度相關。該預測可以為一簡單者以便於僅產生小量的輔助資訊。在最簡單的例子中，該預設係由一適當之比例調整與延遲所組成。最後，預測誤差係被轉換回 HOA 域並被視作為殘餘周圍 HOA 分量，其中係執行一位階降低。

[0012] 有利的是，自該殘餘 HOA 分量中減去可預測之訊號的效果係用以降低其總功率以及優勢方向訊號的殘餘量，而且，在此方法中，亦降低了因位階降低而導致的分解誤差。

[0013] 原則上，本發明之壓縮方法係適於壓縮用於一音場之一高階保真立體音響表示(以 HOA 來表示)，該方法包含步驟：

- 自 HOA 係數之一目前時間框估計優勢音源方向；
- 基於該 HOA 係數以及基於該優勢音源方向分解該 HOA 表示為時域中之優勢方向訊號與一殘餘 HOA 分量，其中為了在代表該殘餘 HOA 分量之均勻採樣方向上得到平面波函數，將該殘餘 HOA 分量轉換為分離空間域，且

其中該平面波函數係自該優勢方向訊號預測而得，因而提供描述該預測之參數，而對應之預測誤差係被轉換回該 HOA 域；

- 降低該殘餘 HOA 分量之目前位階至一較低位階，產生一降階殘餘 HOA 分量；

- 解相關該降階殘餘 HOA 分量以得到對應之殘餘 HOA 分量時域訊號；

- 感知編碼該優勢方向訊號以及該殘餘 HOA 分量時域訊號以便提供壓縮之優勢方向訊號以及壓縮之殘餘 HOA 分量時域訊號。

[0014] 原則上，本發明之壓縮裝置係適於壓縮用於一音場之一高階保真立體音響表示(以 HOA 來表示)，該裝置包含：

- 用以自 HOA 係數之一目前時間訊框估計優勢音源方向之機構；

- 用以基於該 HOA 係數以及基於該優勢音源方向分解該 HOA 表示為時域中之優勢方向訊號與一殘餘 HOA 分量之機構，其中為了在代表該殘餘 HOA 分量之均勻採樣方向上得到平面波函數，將該殘餘 HOA 分量轉換為分離空間域，且其中該平面波函數係自該優勢方向訊號預測而得，因而提供描述該預測之參數，而對應之預測誤差係被轉換回該 HOA 域；

- 用以降低該殘餘 HOA 分量之目前位階至一較低位階，產生一降階殘餘 HOA 分量之機構；

- 用以解相關該降階殘餘 HOA 分量以得到對應之殘餘 HOA 分量時域訊號之機構；

- 用以感知編碼該優勢方向訊號以及該殘餘 HOA 分量時域訊號以便提供壓縮之優勢方向訊號以及壓縮之殘餘 HOA 分量時域訊號之機構。

[0015] 原則上，本發明之解壓縮方法係適於解壓縮根據上述壓縮方法所壓縮之一高階保真立體音響表示，該解壓縮方法包含步驟：

- 感知解碼該壓縮之優勢方向訊號以及該壓縮之殘餘分量訊號以便提供解壓縮之優勢方向訊號與於空間域中代表該殘餘 HOA 分量之解壓縮之時域訊號；

- 互相關該解壓縮之時域訊號以得到一對應之降階殘餘 HOA 分量；

- 延伸該降階殘餘 HOA 分量的位階至原位階以便提供一對應之解壓縮殘餘 HOA 分量；

- 使用該解壓縮之優勢方向訊號、該原位階解壓縮之殘餘 HOA 分量、該估計之優勢音源方向與描述該預測之該參數，組成 HOA 係數之一對應之壓縮與再組成框。

[0016] 原則上，本發明之解壓縮裝置係適於解壓縮根據上述壓縮方法所壓縮之一高階保真立體音響表示，該解壓縮裝置包含：

- 用以感知解碼該壓縮之優勢方向訊號以及該壓縮之殘餘分量訊號以便提供解壓縮之優勢方向訊號與於空間域中代表該殘餘 HOA 分量之解壓縮之時域訊號之機構；

- 用以互相關該解壓縮之時域訊號以得到一對應之降階殘餘 HOA 分量之機構；
- 用以延伸該降階殘餘 HOA 分量的位階至原位階以便提供一對應之解壓縮的殘餘 HOA 分量之機構；
- 用以使用該解壓縮之優勢方向訊號、該原位階解壓縮之殘餘 HOA 分量、該估計之優勢音源方向與描述該預測之該參數組成 HOA 係數之一對應的解壓縮與再組成框之機構。

[0017] 本發明之其他有利實施例係個別揭露於附屬項中。

【圖式簡單說明】

[0018] 本發明之範例性實施例係參考附圖一併說明，該些附圖係繪示如：

第一 A 圖顯示壓縮步驟 1：將 HOA 訊號轉為一些優勢方向訊號、一殘餘周圍 HOA 分量與輔助資訊之解壓縮；

第一 B 圖顯示壓縮步驟 2：對周圍 HOA 分量之位階降低與解相關以及兩分量的感知編碼；

第二 A 圖顯示解壓縮步驟 1：時域信號的感知解碼、代表殘餘周圍 HOA 分量之訊號的互相關與位階延伸；

第二 B 圖顯示解壓縮步驟 2：總 HOA 表示的組成；

第三圖顯示高階保真立體音響解壓縮；

第四圖顯示高階保真立體音響壓縮；以及

第五圖顯示球面座標系統。

第六圖顯示對於不同位階值 N 之標準化函數 $v_N(\theta)$ 。

【實施方式】

[0019] 壓縮處理

[0020] 根據本發明之壓縮處理包含分別描述於第一 A 圖與第一 B 圖中之兩個連續步驟。個別訊號的確切定義係描述於「保真立體音響(HOA)分解與再組成細說」一節中。使用一以訊框方式之流程，其係用於以長度 B 之 HOA 係數序列之非重疊輸入框 $\mathbf{d}(k)$ 的壓縮。其中 k 代表框指數。該些框係相對於具體說明於式(42)中之 HOA 係數序列而被定義為：

$$\mathbf{d}(k) := [\mathbf{d}((kB + 1)T_s) \quad \mathbf{d}((kB + 2)T_s) \quad \dots \quad \mathbf{d}((kB + B)T_s)] \quad (1)$$

其中， T_s 代表取樣週期。

[0021] 在第一 A 圖中，HOA 係數序列之一訊框 $\mathbf{d}(k)$ 係經輸入至一優勢音源方向估計步驟或階段，其係於優勢方向訊號的存在下分析 HOA 表示，且其中該些方向係經估計的。上述方向估計可藉由例如歐洲專利申請案 EP 2665208 A1 所描述的流程來處理。所估計之方向可以 $\hat{\mathbf{n}}_{\text{DOM},1}(k), \dots, \hat{\mathbf{n}}_{\text{DOM},D}(k)$ 來表示，在此處， D 代表方向估計的最大值。他們可經假設而被配置於一矩陣中為 $\mathbf{A}_{\hat{\mathbf{n}}}(k)$ 如：

$$\mathbf{A}_{\hat{\mathbf{n}}}(k) := [\hat{\mathbf{n}}_{\text{DOM},1}(k) \quad \dots \quad \hat{\mathbf{n}}_{\text{DOM},D}(k)] \quad (2)$$

[0022] 暗自假設的是，該些方向估計可藉由將其分配至來自先前框之方向估計而被合適地安排。因此，一個別方向估計之暫時性序列係經假設為描述一優勢音源的方向軌道。具體地來說，若第 d 個優勢音源假定不為積極

者，則可能藉由分配一無效值給 $\hat{\mathbf{n}}_{\text{DOM},d}(k)$ 以將此指出。然後，使用在 $\mathbf{a}_d(k)$ 中之該些估計方向，HOA 表示係於一分解步驟或階段 12 中分解為一些最大值 $\mathcal{D}$ 優勢方向訊號 $\mathbf{x}_{\text{DIR}}(k-1)$ ，一些描述自優勢方向訊號預測該殘餘 HOA 分量之該空間域訊號的參數 $\zeta(k-1)$ ，以及一代表預測誤差之周圍 HOA 分量 $\mathbf{d}_A(k-2)$ 。此分解之細述將提供於「HOA 分解」一節中。

[0023] 在第一 B 圖中，係顯示方向訊號 $\mathbf{x}_{\text{DIR}}(k-1)$ 與殘餘周圍 HOA 分量 $\mathbf{d}_A(k-2)$ 的感知編碼。方向訊號 $\mathbf{x}_{\text{DIR}}(k-1)$ 係為常見之可單獨使用任何已知之感知壓縮技術來進行壓縮的時域訊號。殘餘 HOA 域分量 $\mathbf{d}_A(k-2)$ 係經由兩連續步驟或階段來完成。在一位階降低步驟或階段 13 中，至保真立體音響位階 N_{RED} 的降低係經完成，例如 $N_{\text{RED}} = 1$ ，而產生周圍 HOA 分量 $\mathbf{d}_{A,\text{RED}}(k-2)$ 。該等位階降低係藉由抑制 $\mathbf{d}_A(k-2)$ 僅僅 N_{RED} HOA 係數以及降低其他者來完成。在解碼器之一側，如下方解釋，對於省略值，相對應的零值係經附加上去。

[0024] 必須注意的是，相較於歐洲專利申請案 EP 2665208 A1 中的方法，由於總功率以及殘餘周圍 HOA 分量之方向性的殘餘量較小，一般可挑選較小之降低位階 N_{RED} 。因此，該位階降低相較於 EP 2665208 A1 造成較小的誤差。

[0025] 在後續解相關步驟或階段 14 中，代表位階降低之周圍 HOA 分量 $\mathbf{d}_{A,\text{RED}}(k-2)$ 的 HOA 係數序列係經解相關

以得到時域訊號 $W_{A,RED}(k-2)$ ，其係輸入至(一排)平行之以任何已知的感知壓縮技術操作的感知編碼器或壓縮器 15。上述解相關係經實施以於表示 HOA 表示緊接其解壓縮時避免感知編碼雜訊表露(其解釋請見歐洲專利申請案 EP 12305860.4)。大抵之解相關可使用描述於 EP 2469742 A2 中之一球諧轉換將 $D_{A,RED}(k-2)$ 轉換為在空間域中之 O_{RED} 等效訊號來達成。

[0026] 另可選擇地，可使用如歐洲專利申請案 EP 12305861.2 所提出之一適合的球諧轉換，在此處，取樣方向之網格係被轉動以達到一最佳可能的解相關效果。。再一可選擇之解相關技術係為在歐洲專利申請案 EP 12305860.4 中所描述的 Karhunen-Loève 轉換(KLT)。值得注意的是，對於最後兩種型態的解相關，一些種類之輔助資訊(以 $\alpha(k-2)$ 表示)係為了於一 HOA 解壓縮階段使解相關的逆轉成為可行而被提供。

[0027] 在一實施例中，為了改善編碼效率，所有時域訊號 $X_{DIR}(k-1)$ 與 $W_{A,RED}(k-2)$ 的感知壓縮係為共同實施的。

[0028] 感知編碼的輸出係為壓縮之方向訊號 $\bar{X}_{DIR}(k-1)$ 以及壓縮之周圍時域訊號 $\bar{W}_{A,RED}(k-2)$ 。

[0029] 解壓縮處理

[0030] 解壓縮處理係如第二 A 圖與第二 B 圖所示。與壓縮一樣，其係包含有兩連續步驟。在第二 A 圖中，在一感知解碼或解壓縮步驟或階段 21 中係實施方向訊號 $\bar{X}_{DIR}(k-1)$ 以及代表殘餘周圍 HOA 分量 $\bar{W}_{A,RED}(k-2)$ 的時域訊號

之一感知解壓縮。為了提供位階 N_{RED} 之殘餘分量 HOA 表示 $\hat{\mathbf{D}}_{\text{A,RED}}(k-2)$ ，所致之以感知方式解壓縮的時域訊號 $\bar{\mathbf{W}}_{\text{A,RED}}(k-2)$ 係於一互相關步驟或階段 22 中進行互相關。視情況地，該互相關係可如兩個在步驟/階段 14 描述之可選擇的流程所述以一相反的方式來完成，且其係使用基於已使用之解相關方法的傳送或儲存的參數 $\alpha(k-2)$ 。之後，於位階延伸步驟或階段 23 中，從 $\hat{\mathbf{D}}_{\text{A,RED}}(k-2)$ ，位階 N 之一適當的 HOA 表示 $\hat{\mathbf{D}}_{\text{A}}(k-2)$ 係藉由位階延伸來估計。該位階延伸係藉附加對應“零”值列至 $\hat{\mathbf{D}}_{\text{A,RED}}(k-2)$ 來達成，因此假設該 HOA 係數相對於較高位階具有零值。

[0031] 在第二 B 圖中，於一組成步驟或階段 24 中，總 HOA 表示不但從解壓縮之優勢方向訊號 $\hat{\mathbf{X}}_{\text{DIR}}(k-1)$ 與對應之方向 $\mathbf{A}_{\hat{\mathbf{n}}}(k)$ 以及預測參數 $\zeta(k-1)$ ，也從殘餘周圍 HOA 分量 $\hat{\mathbf{D}}_{\text{A}}(k-2)$ ，再組成而產生解壓縮與再組成之 HOA 係數的訊框 $\hat{\mathbf{D}}(k-2)$ 。

[0032] 假設為了改善編碼效率而共同實施所有時域訊號 $\mathbf{X}_{\text{DIR}}(k-1)$ 與 $\mathbf{W}_{\text{A,RED}}(k-2)$ 的感知壓縮，壓縮之方向訊號 $\bar{\mathbf{X}}_{\text{DIR}}(k-1)$ 以及壓縮之時域訊號 $\bar{\mathbf{W}}_{\text{A,RED}}(k-2)$ 的感知解壓縮也會對應地共同實施。

[0033] 上述再組成之細述將提供於「HOA 再組成」一節中。

[0034] HOA 分解

[0035] 用以說明實施 HOA 分解之操作的一方塊圖係如第三圖所示。該操作係概述如下：首先，平滑化優勢方

向訊號 $\mathbf{x}_{\text{DIR}}(k-1)$ 係經計算並輸出予感知壓縮。然後，介於優勢方向訊號之 HOA 表示 $\mathbf{D}_{\text{DIR}}(k-1)$ 與原 HOA 表示間 $\mathbf{D}(k-1)$ 的殘餘係以一些 O 方向訊號 $\tilde{\mathbf{x}}_{\text{GRID,DIR}}(k-1)$ 來表示，其可被視作為來自均勻分散方向的一般平面波。這些方向訊號係自優勢方向訊號預測而得，在此處，該些預測參數 $\boldsymbol{\zeta}(k-1)$ 係經輸出。最後，介於原 HOA 表示 $\mathbf{D}(k-2)$ 與 HOA 表示與優勢方向訊號之 HOA 表示 $\mathbf{D}_{\text{DIR}}(k-1)$ 間的殘餘 $\mathbf{D}_A(k-2)$ 以及來自均勻分散方向之預測方向訊號的 HOA 表示係經計算並輸出。

[0036] 在進入細節前，要提到的是，連續框間之方向改變，會導致方向性訊號中斷。因此，對於重疊框之個別訊號的即時估計係優先計算，其具有一長度 $2B$ 。接著，使用適當窗函數，連續重疊框之結果係使用適當窗函數進行平滑化。然而，每一次平滑化處理會導致一單框的潛候期。

[0037] 計算即時優勢方向訊號

[0038] 在步驟或階段 30 中，自在 $\mathbf{A}_n(k)$ 中之估計音源方向，對於 HOA 表示序列之一目前訊框 $\mathbf{D}(k)$ ，即時優勢方向訊號的計算係基於如 M. Poletti 於 J. Audio Eng. Soc., 53(11), pages 1004-1025, 2005 發表之 "基於球諧之三維環繞音響 (Three-Dimensional Surround Sound Systems Based on Spherical Harmonics)" 中的模態匹配。具體地來說，這些方向訊號係經調查哪一個 HOA 表示導致所給 HOA 訊號之最佳近似值。

[0039] 再者，不失一般性地，一積極優勢音源之每一方向估計 $\hat{\mathbf{n}}_{\text{DOM},d}(k)$ 係經假設藉由包含有一傾斜角 $\theta_{\text{DOM},d}(k) \in [0, \pi]$ 與一方位角 $\phi_{\text{DOM},d}(k) \in [0, 2\pi]$ (請見第五圖 for illustration) 之一向量根據

$$\hat{\mathbf{n}}_{\text{DOM},d}(k) := \left(\hat{\theta}_{\text{DOM},d}(k), \hat{\phi}_{\text{DOM},d}(k) \right)^T \quad (3)$$

而可被明確地說明。

[0040] 首先，基於積極優勢音源之方向估計的模態矩陣根據

$$\mathbf{E}_{\text{ACT}}(k) := \begin{bmatrix} \mathbf{S}_{\text{DOM},d_{\text{ACT},1}(k)}(k) & \mathbf{S}_{\text{DOM},d_{\text{ACT},2}(k)}(k) & \dots & \mathbf{S}_{\text{DOM},d_{\text{ACT},D_{\text{ACT}}(k)}(k)}(k) \end{bmatrix} \in \mathbb{R}^{O \times D_{\text{ACT}}(k)} \quad (4)$$

與

$$\mathbf{S}_{\text{DOM},d}(k) := \left[S_0^0(\hat{\mathbf{n}}_{\text{DOM},d}(k)), S_1^{-1}(\hat{\mathbf{n}}_{\text{DOM},d}(k)), S_1^0(\hat{\mathbf{n}}_{\text{DOM},d}(k)), \dots, S_N^N(\hat{\mathbf{n}}_{\text{DOM},d}(k)) \right]^T \in \mathbb{R}^O \quad (5)$$

來計算。

[0041] 在式(4)中， $D_{\text{ACT}}(k)$ 代表對於第 k 框之積極方向的數目，而 $d_{\text{ACT},j}(k)$ 、 $1 \leq j \leq D_{\text{ACT}}(k)$ 表示其指數。 $S_n^m(\cdot)$ 代表實值球諧函數，其係於「實值球諧函數的定義」一節中說明。

[0042] 其次，對於定義如下之第 $(k-1)$ 框以及第 k 框，

$$\tilde{\mathbf{X}}_{\text{DIR}}(k) := [\tilde{\mathbf{x}}_{\text{DIR}}(k, 1) \quad \tilde{\mathbf{x}}_{\text{DIR}}(k, 2) \quad \dots \quad \tilde{\mathbf{x}}_{\text{DIR}}(k, 2B)] \quad (6)$$

與

$$\tilde{\mathbf{x}}_{\text{DIR}}(k, l) := [\tilde{x}_{\text{DIR},1}(k, l), \tilde{x}_{\text{DIR},2}(k, l), \dots, \tilde{x}_{\text{DIR},D}(k, l)]^T \in \mathbb{R}^D, 1 \leq l \leq 2B \quad (7)$$

計算包含所有優勢方向訊號之即時估計的矩陣 $\tilde{\mathbf{X}}_{\text{DIR}}(k) \in \mathbb{R}^{D \times 2B}$ ，且此係經由兩個步驟來完成。在第一個步驟中，將

對應消極方向之這些列中的方向訊號樣本被設置為零，
即：

$$\tilde{x}_{\text{DIR},d}(k,l) = 0 \quad \forall 1 \leq l \leq 2B, \quad \text{if } d \notin \mathcal{M}_{\text{ACT}}(k) \quad (8)$$

在此處， $\mathcal{M}_{\text{ACT}}(k)$ 表示一組積極方向。在第二個步驟中，將對應積極方向的方向訊號樣本根據

$$\tilde{\mathbf{X}}_{\text{DIR,ACT}}(k) := \begin{bmatrix} \tilde{x}_{\text{DIR},d_{\text{ACT},1}(k)}(k,1) & \dots & \tilde{x}_{\text{DIR},d_{\text{ACT},1}(k)}(k,2B) \\ \vdots & \ddots & \vdots \\ \tilde{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k)}}(k,1) & \dots & \tilde{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k)}}(k,2B) \end{bmatrix} \quad (9)$$

之一矩陣配置而得。接著，此矩陣經計算以將誤差的歐幾裏德範數 (Euclidean norm) 減到最小

$$\mathbf{E}_{\text{ACT}}(k) \tilde{\mathbf{X}}_{\text{DIR,ACT}}(k) - [\mathbf{D}(k-1) \quad \mathbf{D}(k)] \quad (10)$$

由下式得到答案：

$$\tilde{\mathbf{X}}_{\text{DIR,ACT}}(k) = [\mathbf{E}_{\text{ACT}}^T(k) \mathbf{E}_{\text{ACT}}(k)]^{-1} \mathbf{E}_{\text{ACT}}^T(k) [\mathbf{D}(k-1) \quad \mathbf{D}(k)] \quad (11)$$

[0043] 瞬時平滑

[0044] 對於步驟或階段 31，因為其他類型的訊號可以一完全相似的方法來完成，故上述平滑係僅針對方向訊號 $\tilde{\mathbf{X}}_{\text{DIR}}(k)$ 進行解釋。該些方向訊號 $\tilde{x}_{\text{DIR},d}(k,l)$ ， $1 \leq d \leq D$ 的 (其樣本係可根據式 (6) 包含於矩陣 $\tilde{\mathbf{X}}_{\text{DIR}}(k)$ 中) 估計可藉由一適當窗函數 $w(l)$ 開窗：

$$\tilde{x}_{\text{DIR,WIN},d}(k,l) := \tilde{x}_{\text{DIR},d}(k,l) \cdot w(l), \quad 1 \leq l \leq 2B \quad (12)$$

此窗函數必然滿足在重疊區域中使移動之窗 (假設為 B 樣本之移動) 合計等於 1 之條件：

$$w(l) + w(B+l) = 1 \quad \forall 1 \leq l \leq B \quad (13)$$

[0045] 窗函數之例，係使用下式界定之周期性 Hamming 窗賦予：

$$w(l) := 0.5 \left[1 - \cos\left(\frac{2\pi(l-1)}{2B}\right) \right] \quad \text{for } 1 \leq l \leq 2B \quad (14)$$

[0046] 對於第 $(k-1)$ 框之平滑化方向訊號係藉由開窗之即時估計的適當重疊根據下式計算而得：

$$x_{\text{DIR},d}((k-1)B+l) = \tilde{x}_{\text{DIR},\text{WIN},d}(k-1, B+l) + \tilde{x}_{\text{DIR},\text{WIN},d}(k, l) \quad (15)$$

[0047] 對於第 $(k-1)$ 框之所有平滑化方向訊號的樣本係以矩陣

$$\mathbf{X}_{\text{DIR}}(k-1) := [x_{\text{DIR}}((k-1)B+1) \quad x_{\text{DIR}}((k-1)B+2) \quad \dots \quad x_{\text{DIR}}((k-1)B+B)] \in \mathbb{R}^{\mathcal{D} \times B} \quad (16)$$

$$\text{與 } \mathbf{x}_{\text{DIR}}(l) = [x_{\text{DIR},1}(l), x_{\text{DIR},2}(l), \dots, x_{\text{DIR},\mathcal{D}}(l)]^T \in \mathbb{R}^{\mathcal{D}} \quad (17) \text{ 配置。}$$

[0048] 平滑化優勢方向訊號 $x_{\text{DIR},d}(l)$ 係預期為一連續性訊號，其係可連續地被輸入至感知編碼器。

[0049] 計算平滑化優勢方向訊號之 HOA 表示

[0050] 自 $\mathbf{X}_{\text{DIR}}(k-1)$ 與 $\mathbf{A}_{\hat{n}}(k)$ ，為了照對於 HOA 組成實施之相同運算，平滑化優勢方向訊號之 HOA 表示係於步驟或階段 32 中依據該些連續性訊號 $x_{\text{DIR},d}(l)$ 來計算。因為連續框之間方向估計的改變可導致一中斷，再一次計算長度 $2B$ 之重疊框的即時 HOA 表示經計算並將連續重疊框的結果使用一適當的窗函數而平滑化處理。因此，HOA 表示 $\mathbf{D}_{\text{DIR}}(k-1)$ 可藉由下式而得

$$\mathbf{D}_{\text{DIR}}(k-1) = \mathbf{E}_{\text{ACT}}(k) \mathbf{X}_{\text{DIR},\text{ACT},\text{WIN1}}(k-1) + \mathbf{E}_{\text{ACT}}(k-1) \mathbf{X}_{\text{DIR},\text{ACT},\text{WIN2}}(k-1) \quad (18)$$

$$\begin{aligned} &\text{在此處， } \mathbf{X}_{\text{DIR},\text{ACT},\text{WIN1}}(k-1) := \\ &\begin{bmatrix} x_{\text{DIR},d_{\text{ACT},1}(k)}((k-1)B+1) \cdot w(1) & \dots & x_{\text{DIR},d_{\text{ACT},1}(k)}(kB) \cdot w(B) \\ x_{\text{DIR},d_{\text{ACT},2}(k)}((k-1)B+1) \cdot w(1) & & x_{\text{DIR},d_{\text{ACT},2}(k)}(kB) \cdot w(B) \\ \vdots & \ddots & \vdots \\ x_{\text{DIR},d_{\text{ACT},\mathcal{D}_{\text{ACT}}}(k)}((k-1)B+1) \cdot w(1) & \dots & x_{\text{DIR},d_{\text{ACT},\mathcal{D}_{\text{ACT}}}(k)}(kB) \cdot w(B) \end{bmatrix} \quad (19) \end{aligned}$$

以及 $\mathbf{X}_{\text{DIR},\text{ACT},\text{WIN2}}(k-1) :=$

$$\begin{bmatrix} x_{\text{DIR},d_{\text{ACT},1}(k-1)}((k-1)B+1) \cdot w(B+1) & \dots & x_{\text{DIR},d_{\text{ACT},1}(k-1)}(kB) \cdot w(2B) \\ x_{\text{DIR},d_{\text{ACT},2}(k-1)}((k-1)B+1) \cdot w(B+1) & & x_{\text{DIR},d_{\text{ACT},2}(k-1)}(kB) \cdot w(2B) \\ \vdots & \ddots & \vdots \\ x_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k-1)}(k-1)}((k-1)B+1) \cdot w(B+1) & \dots & x_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k-1)}(k-1)}(kB) \cdot w(2B) \end{bmatrix} \quad (20)$$

)

[0051] 藉由均勻網格上之方向訊號表示殘餘 HOA 表示

[0052] 自 $\mathbf{D}_{\text{DIR}}(k-1)$ 與 $\mathbf{D}(k-1)$ (即藉由延遲框 381 延遲之 $\mathbf{D}(k)$)，藉由一均勻網格上之方向訊號的一殘餘 HOA 表示係於步驟或階段 33 中進行計算。此運算的目的係在於得到來自固定、近乎均勻分散之方向 (亦稱作為網格方向) $\hat{\mathbf{n}}_{\text{GRID},o}$ 、 $1 \leq o \leq O$ 的方向訊號 (即一般平面波函數) 以表示該殘餘 $[\mathbf{D}(k-2) \ \mathbf{D}(k-1)] - [\mathbf{D}_{\text{DIR}}(k-2) \ \mathbf{D}_{\text{DIR}}(k-1)]$ 。

[0053] 首先，相對於網格方向，模態矩陣 $\mathbf{E}_{\text{GRID}}$ 係計算如：

$$\mathbf{E}_{\text{GRID}} := [\mathbf{S}_{\text{GRID},1} \ \mathbf{S}_{\text{GRID},2} \ \dots \ \mathbf{S}_{\text{GRID},O}] \in \mathbb{R}^{O \times O} \quad (21)$$

與

$$\mathbf{S}_{\text{GRID},o} := [S_0^o(\hat{\mathbf{n}}_{\text{GRID},o}), S_1^{-1}(\hat{\mathbf{n}}_{\text{GRID},o}), S_1^0(\hat{\mathbf{n}}_{\text{GRID},o}), \dots, S_N^N(\hat{\mathbf{n}}_{\text{GRID},o})]^T \in \mathbb{R}^O \quad (22)$$

由於在整個壓縮過程中網格方向係固定的，網格方向 $\mathbf{E}_{\text{GRID}}$ 僅需計算一次即可。

[0054] 個別網格上之方向訊號係可得到如：

$$\tilde{\mathbf{X}}_{\text{GRID},\text{DIR}}(k-1) = \mathbf{E}_{\text{GRID}}^{-1}([\mathbf{D}(k-2) \ \mathbf{D}(k-1)] - [\mathbf{D}_{\text{DIR}}(k-2) \ \mathbf{D}_{\text{DIR}}(k-1)]) \quad (23)$$

[0055] 自優勢方向訊號預測均勻網格上之方向訊號

[0056] 自 $\tilde{\mathbf{X}}_{\text{GRID},\text{DIR}}(k-1)$ 與 $\mathbf{X}_{\text{DIR}}(k-1)$ ，均勻網格上之方向訊號係於步驟或階段 34 中被預測。由來自方向訊號之網

格方向 $\hat{\Omega}_{\text{GRID},o}$ 、 $1 \leq o \leq O$ 組成之均勻網格上之方向訊號的預測為了平滑化目的而係基於兩連續框，即(長度 $2B$ 之)網格訊號 $\tilde{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$ 的延伸框係自平滑化優勢方向訊號的延伸框來預測

$$\tilde{\mathbf{X}}_{\text{DIR,EXT}}(k-1) := [\mathbf{X}_{\text{DIR}}(k-3) \quad \mathbf{X}_{\text{DIR}}(k-2) \quad \mathbf{X}_{\text{DIR}}(k-1)] \quad (24)$$

[0057] 首先，包含在 $\tilde{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$ 中之每一網格訊號 $\tilde{x}_{\text{GRID,DIR},o}(k-1,l)$ 、 $1 \leq o \leq O$ 係分配給包含在 $\tilde{\mathbf{X}}_{\text{DIR,EXT}}(k-1)$ 中之一優勢方向訊號 $\tilde{x}_{\text{DIR,EXT},d}(k-1,l)$ 、 $1 \leq d \leq D$ 。此分配係基於網格訊號與所有優勢方向訊號間標準化交叉相關函數的計算。具體地來說，該等優勢方向訊號係分配給網格訊號，其係提供標準化交叉相關函數的最高值。該分配的結果可藉由一分配函數 $f_{\mathcal{A},k-1}: \{1, \dots, O\} \rightarrow \{1, \dots, D\}$ 分配第 o 個網格訊號給第 $f_{\mathcal{A},k-1}(o)$ 個優勢方向訊號而以公式表示。

[0058] 其次，每一網格訊號 $\tilde{x}_{\text{GRID,DIR},o}(k-1,l)$ 係預測自經分配的優勢方向訊號 $\tilde{x}_{\text{DIR,EXT},f_{\mathcal{A},k-1}(o)}(k-1,l)$ 。該預測網格訊號 $\hat{\tilde{x}}_{\text{GRID,DIR},o}(k-1,l)$ 係藉由自經分配之優勢方向訊號 $\tilde{x}_{\text{DIR,EXT},f_{\mathcal{A},k-1}(o)}(k-1,l)$ 之延遲以及比例調整而計算如下

$$\hat{\tilde{x}}_{\text{GRID,DIR},o}(k-1,l) = K_o(k-1) \cdot \tilde{x}_{\text{DIR,EXT},f_{\mathcal{A},k-1}(o)}(k-1, l - \Delta_o(k-1)) \quad (25)$$

在此處， $K_o(k-1)$ 代表比例因數而 $\Delta_o(k-1)$ 代表樣本延遲。這些參數係經選擇以降低預測誤差。

[0059] 若預測誤差的功率大於該網格訊號本身之總功率，則該預測係被認為為失敗的。然後，個別預測參數可被設定為任何無效值。

[0060] 值得注意的是，其他種型態的預測也是可能

的。舉例來說，代替計算一全頻帶比例因數，亦可判斷感知位向之頻率頻帶的比例因數。然而，此種運算改善了在輔助資訊之一增加量成本方面的預測。

[0061] 所有預測參數可被配置於參數矩陣中如：

$$\zeta(k-1) := \begin{bmatrix} f_{\mathcal{A},k-1}(1) & K_1(k-1) & \Delta_1(k-1) \\ f_{\mathcal{A},k-1}(2) & K_2(k-1) & \Delta_2(k-1) \\ \vdots & \vdots & \vdots \\ f_{\mathcal{A},k-1}(O) & K_O(k-1) & \Delta_O(k-1) \end{bmatrix} \quad (26)$$

[0062] 所有預測訊號 $\hat{x}_{\text{GRID,DIR},o}(k-1, l)$ 、 $1 \leq o \leq O$ ，係假設為配置於矩陣 $\hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$ 中。

[0063] 計算均勻網格上之預測方向訊號的 HOA 表示

[0064] 自 $\hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$ 根據

$$\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-1) = \mathbf{E}_{\text{GRID}} \hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1) \quad (27)$$

於步驟或階段 35 中計算預測網格訊號的 HOA 表示。

[0065] 計算殘餘周圍音場分量的 HOA 表示

[0066] 自 $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2)$ (其係 $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-1)$ 之一暫時性平滑化形式 (在步驟/階段 36))、自 $\mathbf{D}(k-2)$ (其係 $\mathbf{D}(k)$ 之一雙框延遲形式 (延遲 381 與 383))、以及自 $\mathbf{D}_{\text{DIR}}(k-2)$ (其係 $\mathbf{D}_{\text{DIR}}(k-1)$ 之一框延遲形式 (延遲 382))，殘餘周圍音場分量的 HOA 表示係藉由

$$\mathbf{D}_A(k-2) = \mathbf{D}(k-2) - \hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2) - \mathbf{D}_{\text{DIR}}(k-2) \quad (28)$$

於步驟或階段 37 中進行計算。

[0067] HOA 再組成

[0068] 在詳細描述第四圖中個別步驟或階段的詳細流程之前，先提供一總結。相對於均勻分散方向之方向訊號 $\hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$ 係使用預測參數 $\zeta(k-1)$ 而預測自解碼之優勢方

向訊號 $\hat{\mathbf{X}}_{\text{DIR}}(k-1)$ 。接著，總 HOA 表示 $\hat{\mathbf{D}}(k-2)$ 係由優勢方向訊號之 HOA 表示 $\hat{\mathbf{D}}_{\text{DIR}}(k-2)$ 、預測方向訊號之 HOA 表示 $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2)$ 以及殘餘周圍 HOA 分量 $\hat{\mathbf{D}}_{\text{A}}(k-2)$ 所組成。

[0069] 計算優勢方向訊號之 HOA 表示

[0070] $\mathbf{A}_{\hat{\mathbf{n}}}(k)$ 與 $\hat{\mathbf{X}}_{\text{DIR}}(k-1)$ 係經輸入至一步驟或階段 41 中以判斷優勢方向訊號之一 HOA 表示。在自方向估計 $\mathbf{A}_{\hat{\mathbf{n}}}(k)$ 與 $\mathbf{A}_{\hat{\mathbf{n}}}(k-1)$ 計算模態矩陣 $\mathbf{E}_{\text{ACT}}(k)$ 與 $\mathbf{E}_{\text{ACT}}(k-1)$ 之後，基於對於第 k 框與第 $(k-1)$ 框之積極音源的方向估計，優勢方向訊號之 HOA 表示 $\hat{\mathbf{D}}_{\text{DIR}}(k-1)$ 係藉由下式而得：

$$\hat{\mathbf{D}}_{\text{DIR}}(k-1) = \mathbf{E}_{\text{ACT}}(k) \mathbf{X}_{\text{DIR,ACT,WIN1}}(k-1) + \mathbf{E}_{\text{ACT}}(k-1) \mathbf{X}_{\text{DIR,ACT,WIN2}}(k-1) \quad (29)$$

在此處， $\mathbf{X}_{\text{DIR,ACT,WIN1}}(k-1) :=$

$$\begin{bmatrix} \hat{x}_{\text{DIR},d_{\text{ACT},1}}(k)((k-1)B+1) \cdot w(1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},1}}(kB) \cdot w(B) \\ \hat{x}_{\text{DIR},d_{\text{ACT},2}}(k)((k-1)B+1) \cdot w(1) & & \hat{x}_{\text{DIR},d_{\text{ACT},2}}(kB) \cdot w(B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}}(k)}((k-1)B+1) \cdot w(1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}}(k)}(kB) \cdot w(B) \end{bmatrix} \quad (30)$$

以及 $\mathbf{X}_{\text{DIR,ACT,WIN2}}(k-1) :=$

$$\begin{bmatrix} \hat{x}_{\text{DIR},d_{\text{ACT},1}}(k-1)((k-1)B+1) \cdot w(B+1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},1}}(kB) \cdot w(2B) \\ \hat{x}_{\text{DIR},d_{\text{ACT},2}}(k-1)((k-1)B+1) \cdot w(B+1) & & \hat{x}_{\text{DIR},d_{\text{ACT},2}}(kB) \cdot w(2B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}}(k-1)}((k-1)B+1) \cdot w(B+1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}}(k-1)}(kB) \cdot w(2B) \end{bmatrix} \quad (31)$$

)。

[0071] 自優勢方向訊號預測均勻網格上之方向訊號

[0072] $\hat{\mathbf{z}}(k-1)$ 與 $\hat{\mathbf{X}}_{\text{DIR}}(k-1)$ 係經輸入至一步驟或階段 43 中以自優勢方向訊號預測均勻網格上之方向訊號。均勻網格上之預測方向訊號的延伸框係由元素 $\hat{x}_{\text{GRID,DIR},o}(k-1,l)$ 根據

$$\hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1) = \begin{bmatrix} \hat{x}_{\text{GRID,DIR},1}(k-1,1) & \dots & \hat{x}_{\text{GRID,DIR},1}(k-1,2B) \\ \hat{x}_{\text{GRID,DIR},2}(k-1,1) & & \hat{x}_{\text{GRID,DIR},2}(k-1,2B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{GRID,DIR},O}(k-1,1) & \dots & \hat{x}_{\text{GRID,DIR},O}(k-1,2B) \end{bmatrix} \quad (32)$$

所組成，且其係藉由

$$\hat{x}_{\text{GRID,DIR},o}(k-1, l) = K_o(k-1) \cdot \hat{x}_{\text{DIR},f_{\mathcal{A},k-1}(o)}((k-1)B + l - \Delta_o(k-1)) \quad (33)$$

預測自優勢方向訊號。

[0073] 計算均勻網格上之預測方向訊號的 HOA 表示

[0074] 在用以計算均勻網格上之預測方向訊號之 HOA 表示的一步驟或階段 44 中，該預測網格方向訊號之 HOA 表示係藉由下式而得：

$$\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-1) = \mathbf{E}_{\text{GRID}} \hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1) \quad (34)$$

在此處， $\mathbf{E}_{\text{GRID}}$ 代表相對於該預測網格方向之模態矩陣(定義請見式(21))。

[0075] 組成 HOA 音場表示

[0076] 自 $\hat{\mathbf{D}}_{\text{DIR}}(k-2)$ (即藉由框延遲 42 延遲之 $\hat{\mathbf{D}}_{\text{DIR}}(k-1)$)， $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2)$ (其係步驟或階段 45 中 $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-1)$ 之一暫時性平滑化形式)與 $\hat{\mathbf{D}}_{\text{A}}(k-2)$ ，總 HOA 音場表示係最終於一步驟或階段 46 中組成如：

$$\hat{\mathbf{D}}(k-2) = \hat{\mathbf{D}}_{\text{DIR}}(k-2) + \hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2) + \hat{\mathbf{D}}_{\text{A}}(k-2) \quad (35)$$

[0077] 高階保真立體音響之基本原理

[0078] 高階保真立體音響係基於在一緊密關注區域(compact area of interest，且其係經假設不具有音源)中一音場的描述。在該例中，音壓 $p(t, x)$ 於時間 t 以及在關注區域中位置 x 的時空行為係實質上完全地藉由同質波動方程式(homogeneous wave equation)來偵測。後續係基於如第五圖所示之一球面座標系統。 x 軸係指向前方的位置， y 軸指向左側，以及 z 軸指向頂端。在空間中之一位置 $x =$

$(r, \theta, \phi)^T$ 係藉由一半徑 $r > 0$ 來表示 (即至座標原點的距離)，一量測自極軸 z 之傾斜角 $\theta \in [0, \pi]$ 以及一自 x 軸在 x - y 平面以逆時針方向量測之方位角 $\phi \in [0, 2\pi[$ 。 $(\cdot)^T$ 代表轉移。

[0079] 相對於以 $\mathcal{F}_t(\cdot)$ 代表之時間之音壓的傅里葉轉換 (可見於由 Earl G. Williams 著教科書《傅里葉聲學》，列於應用算術科學第 93 卷，學術出版社，1999 年)，即

$$P(\omega, x) = \mathcal{F}_t(p(t, x)) = \int_{-\infty}^{\infty} p(t, x) e^{-i\omega t} dt \quad (36)$$
 以 ω 代表角頻率與 i 代表虛擬單位，可根據下式被展開成一系列球諧 (Spherical Harmonics)

$$P(\omega = kc_s, r, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n A_n^m(k) j_n(kr) S_n^m(\theta, \phi) \quad (37)$$

其中 c_s 代表音速以及 k 代表角波數，其係藉由 $k = \frac{\omega}{c_s}$ 而與角頻率 ω 相關， $j_n(\cdot)$ 代表第一階之球貝塞爾 (Bessel) 函數，以及 $S_n^m(\theta, \phi)$ 代表 n 階與 m 度之實值球諧函數，其係定義於「實值球諧函數之定義」一節中。展開係數 $A_n^m(k)$ 係僅基於角波數 k 。必須注意的是，其係經暗自假設該音壓為空間的有限頻寬。因此，該系列係於一較高的限度 N 相對於位階指數 n 而被截短，其係稱作為 HOA 表示的位階。

[0080] 若該音場係藉由不同角頻率 ω 之諧平面波之一無限數值之一重疊來表示且係來自藉由角組合 (angle tuple) (θ, ϕ) 之所有可能方向，其可知的是 (請見 B. Rafaely 在〈聲場使用球形褶合在球體上之平面波分解〉所述，美國音響學會會刊第 4 卷第 116 期，2149-2157 頁，2004 年) 平面波複振幅函數 $D(\omega, \theta, \phi)$ 可藉由球諧展開來表示

$$D(\omega = kc_s, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n D_n^m(k) S_n^m(\theta, \phi) \quad (38)$$

其中，展開係數 $D_n^m(k)$ 藉由係與展開係數 $A_n^m(k)$ by $A_n^m(k) = 4\pi i^n D_n^m(k)$ (39) 相關。

[0081] 將個別係數 $D_n^m(k = \omega/c_s)$ 假設為角頻率 ω 的函數，逆傅里葉轉換 (以 $\mathcal{F}_t^{-1}(\cdot)$ 表示) 的應用係提供如下時域函數

$$d_n^m(t) = \mathcal{F}_t^{-1} \left(D_n^m \left(\frac{\omega}{c_s} \right) \right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} D_n^m \left(\frac{\omega}{c_s} \right) e^{i\omega t} d\omega \quad (40)$$

予於每一 n 階以及 m 度，其係可被收集於一單一向量中

$$\mathbf{d}(t) = \begin{bmatrix} d_0^0(t) & d_1^{-1}(t) & d_1^0(t) & d_1^1(t) & d_2^{-2}(t) & d_2^{-1}(t) & d_2^0(t) & d_2^1(t) & d_2^2(t) & \dots \\ d_N^{N-1}(t) & d_N^N(t) & & & & & & & & \end{bmatrix}^T \quad (41)$$

[0082] 在向量 $\mathbf{d}(t)$ 中之一時域函數 $d_n^m(t)$ 的位置指數係經由 $n(n+1)+1+m$ 而定。

[0083] 最終保真立體音響格式使用一取樣頻率 f_s 提供 $\mathbf{d}(t)$ 之樣本形式如

$$\{\mathbf{d}(lT_s)\}_{l \in \mathbb{N}} = \{\mathbf{d}(T_s), \mathbf{d}(2T_s), \mathbf{d}(3T_s), \mathbf{d}(4T_s), \dots\} \quad (42)$$

其中， $T_s = 1/f_s$ 代表取樣週期。 $\mathbf{d}(lT_s)$ 的元素亦稱作為保真立體音響係數。值得注意的是，時域訊號 $d_n^m(t)$ 以及因此保真立體音響係數為實值。

[0084] 實值球諧函數之定義

[0085] 實值球諧函數 $S_n^m(\theta, \phi)$ 係由下式而定

$$S_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}} P_{n,|m|}(\cos \theta) \text{trg}_m(\phi) \quad (43)$$

$$\text{與 } \text{trg}_m(\phi) = \begin{cases} \sqrt{2}\cos(m\phi) & m > 0 \\ 1 & m = 0 \\ -\sqrt{2}\sin(m\phi) & m < 0 \end{cases} \quad (44) \text{ 而定。}$$

[0086] 相關連之勒讓德 (Legendre) 函數係以勒讓德多項式 $P_n(x)$ 而定義為

$$P_{n,m}(x) = (1-x^2)^{m/2} \frac{d^m}{dx^m} P_n(x), m \geq 0 \quad (45)$$

以及，不若在上述所指之 E.G. Williams 教科書，不具有 Condon-Shortley 相位 $(-1)^m$ 。

[0087] 高階保真立體音響之空間解析度

[0088] 來自一方向 $\boldsymbol{\Omega}_0 = (\theta_0, \phi_0)^T$ 之一般平面波函數 $x(t)$ 係藉由下式而表示於 HOA 中：

$$d_n^m(t) = x(t) S_n^m(\boldsymbol{\Omega}_0), \quad 0 \leq n \leq N, |m| \leq n \quad (46)$$

[0089] 平面波振幅 $d(t, \boldsymbol{\Omega}) := \mathcal{F}_t^{-1}(D(\omega, \boldsymbol{\Omega}))$ 之相對應的空間密度係 given by

$$d(t, \boldsymbol{\Omega}) = \sum_{n=0}^N \sum_{m=-n}^n d_n^m(t) S_n^m(\boldsymbol{\Omega}) \quad (47)$$

$$= x(t) \underbrace{\left[\sum_{n=0}^N \sum_{m=-n}^n S_n^m(\boldsymbol{\Omega}_0) S_n^m(\boldsymbol{\Omega}) \right]}_{v_N(\boldsymbol{\theta})} \quad (48)$$

由式(48)可知，其係一般平面波函數 $x(t)$ 與一空間分散函數 $v_N(\boldsymbol{\theta})$ 的產物，且可僅依據具有下述性質之介於 $\boldsymbol{\Omega}$ 與 $\boldsymbol{\Omega}_0$ 間的角度 $\boldsymbol{\theta}$ ：

$$\cos \boldsymbol{\theta} = \cos \theta \cos \theta_0 + \cos(\phi - \phi_0) \sin \theta \sin \theta_0 \quad (49)$$

[0090] 如預期，在一無限位階數的限度中，即 $N \rightarrow \infty$ ，空間分散函數轉為一狄拉克 $\delta(\cdot)$ ，即

$$\lim_{N \rightarrow \infty} v_N(\boldsymbol{\theta}) = \frac{\delta(\boldsymbol{\theta})}{2\pi} \quad (50)$$

然而，在有限位階數 N 的例子中，來自方向 $\boldsymbol{\Omega}_0$ 之一般平面波的貢獻係被模糊而至相鄰之方向，其中該模糊的程度會隨著一增加的位階而減少。對於不同位階值 N 之標準化函數 $v_N(\boldsymbol{\theta})$ 係繪示如第六圖。

[0091] 必須指明的是，平面波振幅之空間密度之時域行為的任一方向 $\boldsymbol{\Omega}$ 係為其於任何其他方向上之行為的倍數。具體的來說，對於一些固定方向 $\boldsymbol{\Omega}_1$ 與 $\boldsymbol{\Omega}_2$ 之函數 $d(t, \boldsymbol{\Omega}_1)$ 與 $d(t, \boldsymbol{\Omega}_2)$ 係相對於時間 t 而彼此高度相關。

[0092] 離散空間領域

[0093] 若平面波振幅之空間密度係以一些 O 空間方向 $\boldsymbol{\Omega}_o$ 、 $1 \leq o \leq O$ (其係近乎均勻地分散在單位球體上) 離散，得到 O 方向訊號 $d(t, \boldsymbol{\Omega}_o)$ 。收集這些訊號為一向量：

$$\mathbf{d}_{\text{SPAT}}(t) := [d(t, \boldsymbol{\Omega}_1) \ \dots \ d(t, \boldsymbol{\Omega}_O)]^T \quad (51)$$

其可使用式(47)驗證此向量可藉由如 $\mathbf{d}_{\text{SPAT}}(t) = \boldsymbol{\Psi}^H \mathbf{d}(t)$ (52) 之一簡單矩陣乘法而自定義於式(41)中之連續保真立體音響表示來計算，在此處， $(\cdot)^H$ 代表共同轉移與結合，而 $\boldsymbol{\Psi}$ 代表由 $\boldsymbol{\Psi} := [S_1 \ \dots \ S_O]$ (53) 與 $S_o := [S_o^0(\boldsymbol{\Omega}_o) \ S_o^{-1}(\boldsymbol{\Omega}_o) \ S_o^0(\boldsymbol{\Omega}_o) \ S_o^1(\boldsymbol{\Omega}_o) \ \dots \ S_o^{N-1}(\boldsymbol{\Omega}_o) \ S_o^N(\boldsymbol{\Omega}_o)]$ (54)) 所定義之模態矩陣。

[0094] 由於方向 $\boldsymbol{\Omega}_o$ 係近乎均勻地分散於單位球體上，模態矩陣一般來說為可逆的。因此，該連續性保真立體音響表示係可藉由

$$\mathbf{d}(t) = \boldsymbol{\Psi}^{-H} \mathbf{d}_{\text{SPAT}}(t) \quad (55)$$

而自方向訊號 $d(t, \boldsymbol{\Omega}_o)$ 來計算。

[0095] 該些式均構成保真立體音響表示與空間域間之一轉換以及一逆轉換。在此應用中，這些轉換可稱作為球諧函數轉換以及逆球諧函數轉換。

[0096] 由於方向 $\boldsymbol{\Omega}_o$ 係近乎均勻地分散在單位球體上， $\boldsymbol{\Psi}^H \approx \boldsymbol{\Psi}^{-1}$ (56)，其證明了在式(52)中以 $\boldsymbol{\Psi}^{-1}$ 代替 $\boldsymbol{\Psi}^H$ 的使用。

[0097] 有利地，所有提及之關係亦對離散時間領域 (discrete-time domain) 有效。

[0098] 在編碼之一側和在解碼之一側一樣，該些發

明流程可藉由單一處理器或電路，或藉由數個並聯運作以及/或在發明流程之不同部份上運作之處理器或電路來完成。

[0099] 本發明可被應用於處理對應之聲音訊號，其係可於一家庭環境中之一喇叭設置上或於一劇院之一喇叭設置上表示或演示。

【符號說明】

[0100]

11：優勢音源方向的估計

12：HOA 表示的分解

13：位階降低

14：解相關

15：感知編碼

21：感知解碼

22：互相關

23：位階延伸

24：HOA 表示的組成

30：計算即時性方向訊號

31：實施暫時性平滑化

32：計算平滑化優勢方向訊號之 HOA 表示

33：藉由均勻網格上之方向訊號表示殘餘 HOA 分量

34：自優勢方向訊號預測均勻網格上之方向訊號

35：計算均勻網格上之預測方向訊號之 HOA 表示

- 36：實施暫時性平滑化
- 37：計算殘餘周圍音場分量之 HOA 表示
- 381：框延遲
- 382：框延遲
- 383：框延遲
- 41：計算優勢方向訊號之 HOA 表示
- 42：框延遲
- 43：自優勢方向訊號預測均勻網格上之方向訊號
- 44：計算均勻網格上之預測方向訊號之 HOA 表示
- 45：實施暫時性平滑化
- 46：組成總 HOA 音場表示

I681386

發明摘要

【發明名稱】(中文/英文)

用於音場之高階保真立體音響表示的壓縮與解壓縮方法
及裝置

METHOD AND APPARATUS FOR COMPRESSING AND
DECOMPRESSING A HIGHER ORDER AMBISONICS
REPRESENTATION FOR A SOUND FIELD

【中文】

本發明改善了高階保真立體音響（HOA）表示的壓縮。上述高階保真立體音響（HOA）表示係在優勢音源的存在下被分析且其方向均經估計。隨後，HOA 表示係經分解為一些優勢方向訊號與一殘餘分量。為了於均勻取樣方向上得到一般平面波函數，此殘餘分量係經轉換為離散空間域，其係預測自優勢方向訊號。最終，預測誤差會被轉換回 HOA 域以及表示殘餘周圍 HOA 分量，其中係實施一位階降低，並緊接著進行優勢方向訊號與殘餘分量的感知編碼。

【 英文 】

The invention improves HOA sound field representation compression. The HOA representation is analyzed for the presence of dominant sound sources and their directions are estimated. Then the HOA representation is decomposed into a number of dominant directional signals and a residual component. This residual component is transformed into the discrete spatial domain in order to obtain general plane wave functions at uniform sampling directions, which are predicted from the dominant directional signals. Finally, the prediction error is transformed back to the HOA domain and represents the residual ambient HOA component for which an order reduction is performed, followed by perceptual encoding of the dominant directional signals and the residual component.

【代表圖】

【本案指定代表圖】：第(1A)圖。

【本代表圖之符號簡單說明】：

11：優勢音源方向的估計

12：HOA 表示的分解

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：無

申請專利範圍

1. 一種用於音場之高階保真立體音響（HOA）表示的解壓縮方法，該方法包含：

感知地解碼經壓縮的優勢方向訊號與經壓縮的殘餘分量訊號以便提供經解壓縮的優勢方向訊號與代表空間域中之該殘餘 HOA 分量之經解壓縮的時間域訊號；

重新關聯該經解壓縮的時間域訊號以得到一對應的降階殘餘 HOA 分量；

基於該對應的降階殘餘 HOA 分量判定經解壓縮的殘餘 HOA 分量；

至少基於參數判定預測的方向訊號；

基於該等經解壓縮的優勢方向訊號、該等預測的方向信號以及該等經解壓縮的殘餘 HOA 分量判定 HOA 音場表示，以及

其中該參數有關於在當前框中的數個有效方向訊號。

2. 一種用於解壓縮高階保真立體音響（HOA）表示的設備，該設備包含：

解碼器，其感知地解碼經壓縮的優勢方向訊號與經壓縮的殘餘分量訊號以便提供經解壓縮的優勢方向訊號與代表空間域中之該殘餘 HOA 分量之經解壓縮的時間域訊號；

再關聯器，其重新關聯該經解壓縮的時間域訊號以得到一對應的降階殘餘 HOA 分量；

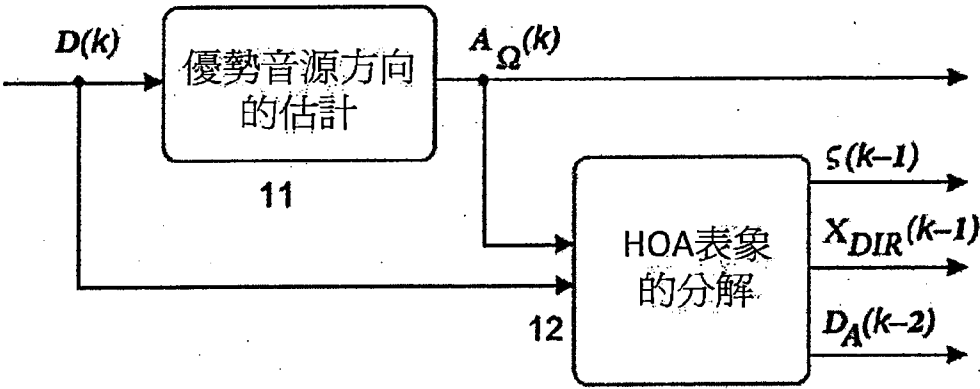
處理器，其組態以基於該對應的降階殘餘 HOA 分量

判定經解壓縮的殘餘 HOA 分量，該處理器進一步組態以至少基於參數判定預測的方向訊號；

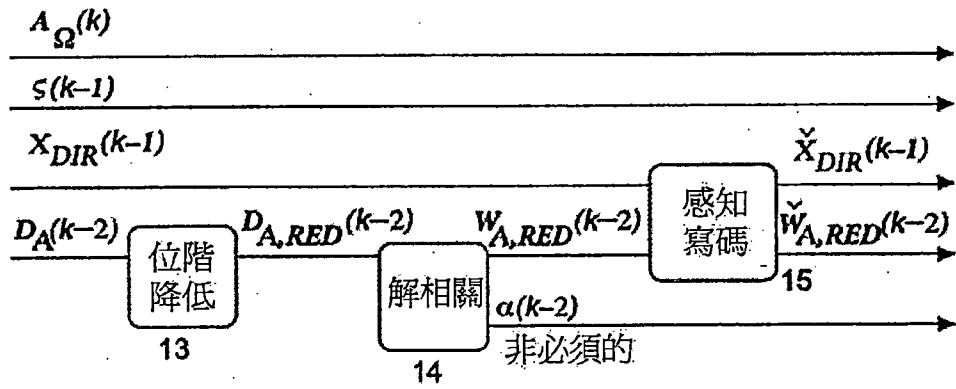
其中該處理器係進一步組態以基於該等經解壓縮的優勢方向訊號、該等預測的方向信號以及該等經解壓縮的殘餘 HOA 分量判定 HOA 音場表示，以及

其中該參數有關於在當前框中的數個有效方向訊號。

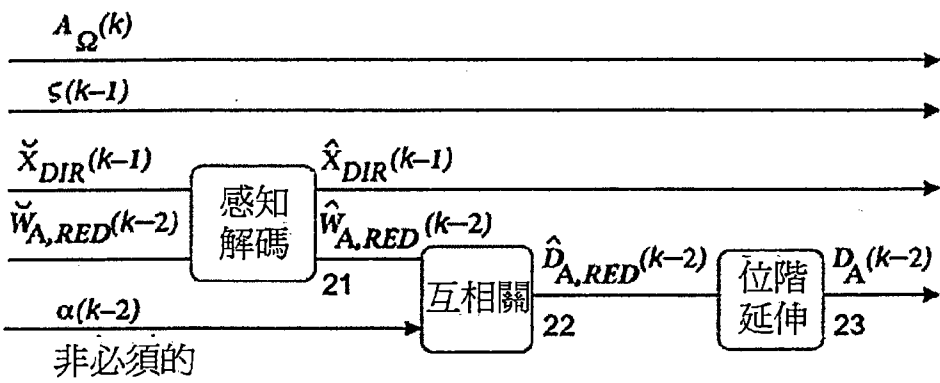
圖式



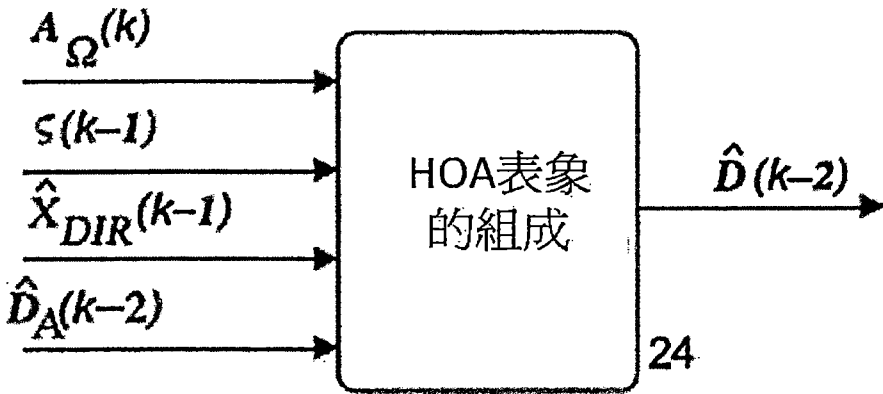
第一 A圖



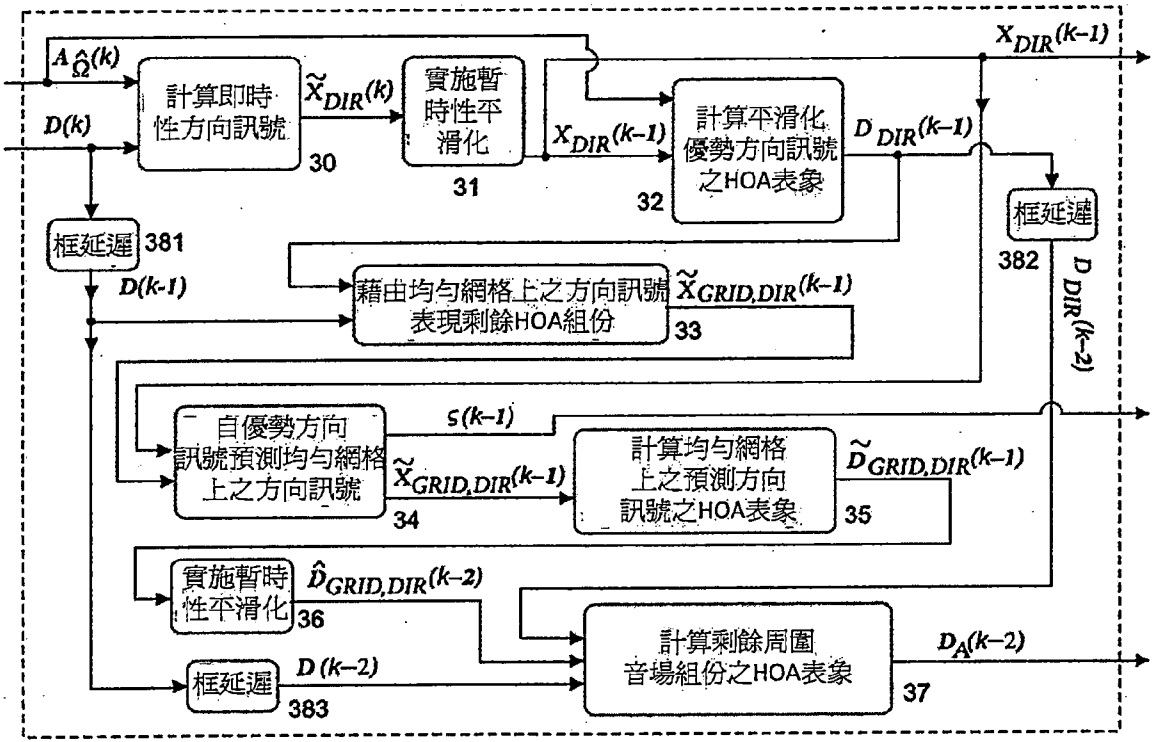
第一 B圖



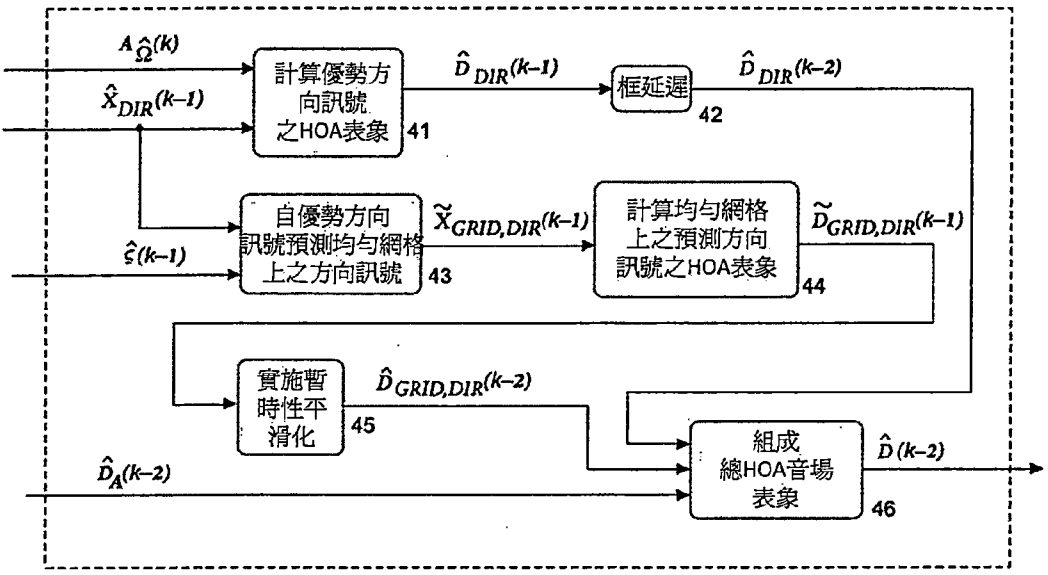
第二 A圖



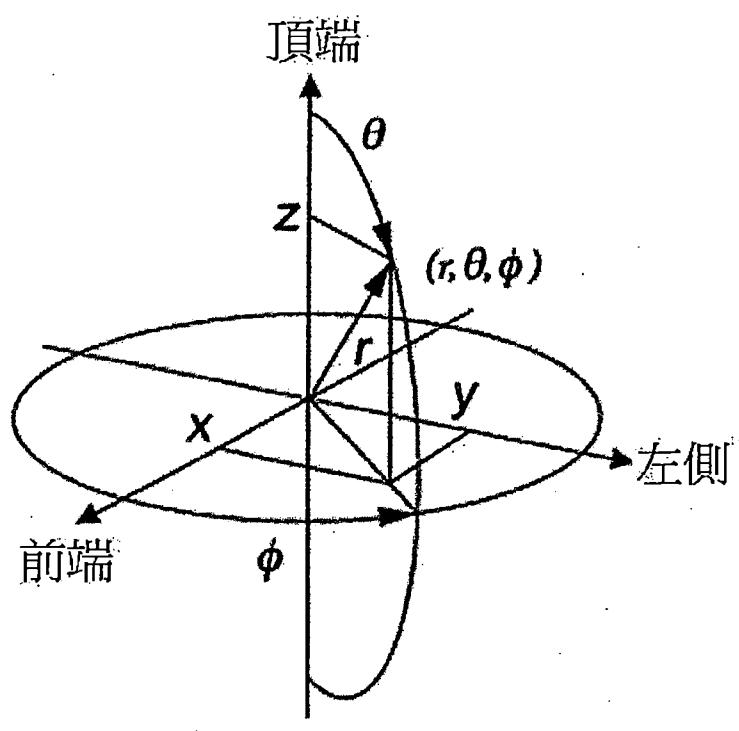
第二 B圖



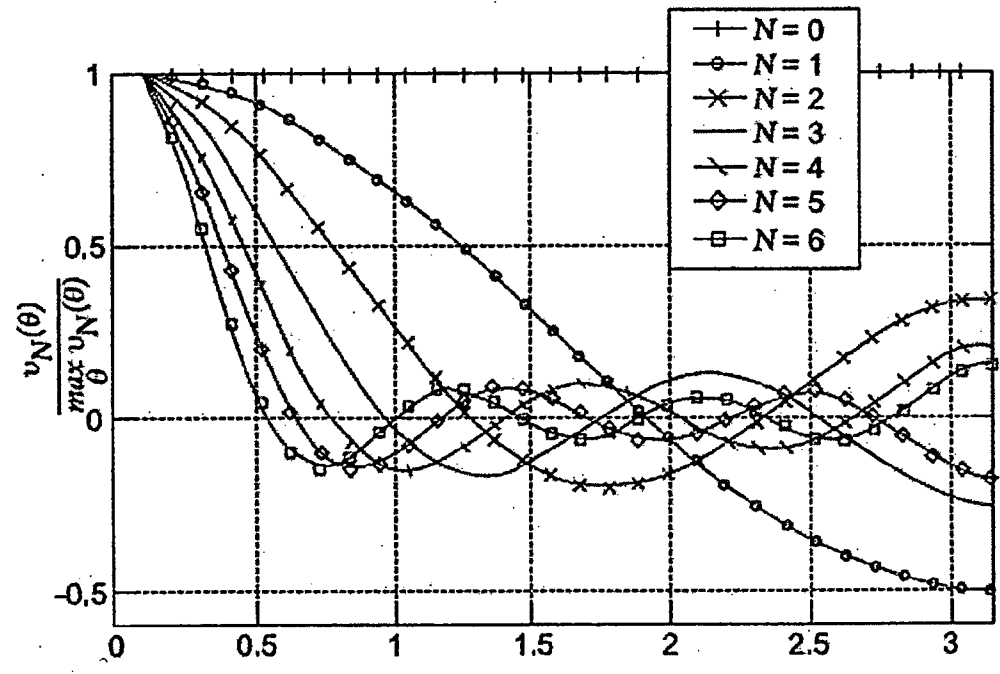
第三圖



第四圖



第五圖



第六圖