



(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 I748636 B

(45)公告日：中華民國 110 (2021) 年 12 月 01 日

(21)申請案號：109130726

(22)申請日：中華民國 104 (2015) 年 06 月 26 日

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

(30)優先權：2014/06/27 歐洲專利局 14306027.5

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

(72)發明人：克魯格 亞歷山德 KRUEGER, ALEXANDER (DE)；科登 斯凡 KORDON, SVEN (DE)

(74)代理人：林志剛

(56)參考文獻：

US 2012/0155653A1

US 2013/0216070A1

審查人員：吳鴻鎮

申請專利範圍項數：3 項 圖式數：6 共 40 頁

(54)名稱

用於解碼聲音或聲場的高階保真立體音響 (HOA) 表示的方法

(57)摘要

本發明係為一種編碼之高階保真立體音響(HOA)資料框表示，其包含非差分增益值係與一 HOA 資料框表示之資料框特定者之聲道信號相關聯，當壓縮一 HOA 資料框表示時，於進行知覺編碼(16)前，應用一增益控制(15、151)用於各聲道信號，依一差分方式傳送增益值作為邊資訊。然而，用以開始此類串流壓縮 HOA 資料框表示之解碼，需要絕對增益值，其應以最小位元數編碼。用以判定此類最低整數位元數，將 HOA 資料框表示(C(k))於空間域中呈現至位於一單位球面上之虛擬揚聲器信號，隨後進行 HOA 資料框表示(C(k))之正規化，接著將最低整數位元數設成

$$\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil + 1) \rceil$$

When compressing an HOA data frame representation, a gain control (15, 151) is applied for each channel signal before it is perceptually encoded (16). The gain values are transferred in a differential manner as side information. However, for starting decoding of such streamed compressed HOA data frame representation absolute gain values are required, which should be coded with a minimum number of bits. For determining such lowest integer number (β_e) of bits the HOA data frame representation (C(k)) is rendered in spatial domain to virtual loudspeaker signals lying on a unit sphere, followed by normalisation of the HOA data frame representation (C(k)). Then the lowest integer number of bits is set to

$$\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil + 1) \rceil.$$

指定代表圖：

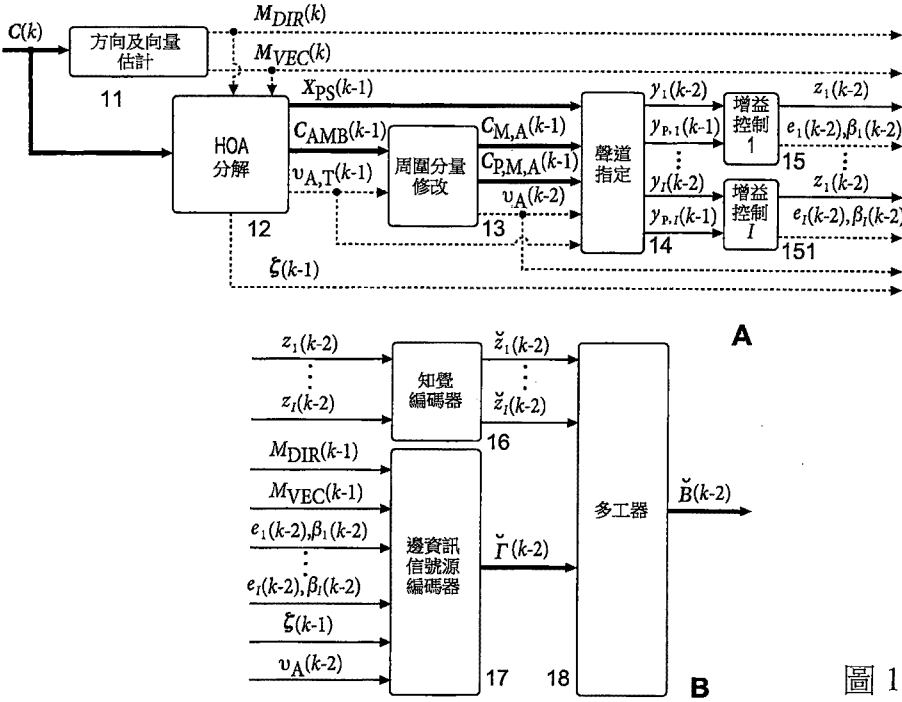


圖 1

符號簡單說明：

11:方向及向量估計處理步驟

12:HOA 分解處理步驟

13:周圍分量修改處理步驟

14:聲道指定步驟

15,151:增益控制處理步驟

16:知覺編碼器步驟

17:邊資訊信號源編碼器步驟

18:多工器

$\tilde{B}(k-2)$:輸出訊框

$C(k)$:初始訊框

$C_{AMB}(k-1)$:周圍 HOA 分量之訊框

$C_{M,A}(k-1)$:修改周圍 HOA 分量

$C_{P,M,A}(k-1)$:暫預測修改周圍 HOA 分量

$e_1(k-2), \dots, e_I(k-2)$:指數

$\beta_1(k-2), \dots, \beta_I(k-2)$:異常旗標

$M_{DIR}(k), M_{VEC}$

$(k), M_{DIR}(k-1), M_{VEC}$

$(k-1)$:元組集

$v_{A,T}(k-1)$:目標指定向量

$v_A(k-2)$:最終指定向量

$X_{PS}(k-1)$:所有主要聲音信號框

$y_1(k-2), \dots, y_I(k-2)$:信號框

$y_{P,1}(k-1), \dots, y_{P,I}(k-1)$:預測信號框

$z_1(k-2), \dots, z_I(k-2)$:信號

$\check{z}_1(k-2), \dots,$

$\check{z}_I(k-2)$:編碼信號

$\check{r}(k-2)$:編碼邊資
訊

$\zeta(k-1)$:預測參數

I748636

發明摘要

【發明名稱】(中文/英文)

用於解碼聲音或聲場的高階保真立體音響(HOA)表示的方法

METHOD FOR DECODING A HIGHER ORDER AMBISONICS (HOA)
REPRESENTATION OF A SOUND OR SOUNDFIELD

【中文】

本發明係為一種編碼之高階保真立體音響(HOA)資料框表示，其包含非差分增益值係與一 HOA 資料框表示之資料框特定者之聲道信號相關聯，當壓縮一 HOA 資料框表示時，於進行知覺編碼(16)前，應用一增益控制(15、151)用於各聲道信號，依一差分方式傳送增益值作為邊資訊。然而，用以開始此類串流壓縮 HOA 資料框表示之解碼，需要絕對增益值，其應以最小位元數編碼。用以判定此類最低整數位元數，將 HOA 資料框表示($C(k)$)於空間域中呈現至位於一單位球面上之虛擬揚聲器信號，隨後進行 HOA 資料框表示($C(k)$)之正規化，接著將最低整數位元數設成 $\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil + 1) \rceil$ 。

【 英文 】

When compressing an HOA data frame representation, a gain control (15, 151) is applied for each channel signal before it is perceptually encoded (16). The gain values are transferred in a differential manner as side information. However, for starting decoding of such streamed compressed HOA data frame representation absolute gain values are required, which should be coded with a minimum number of bits. For determining such lowest integer number (β_e) of bits the HOA data frame representation ($C(k)$) is rendered in spatial domain to virtual loudspeaker signals lying on a unit sphere, followed by normalisation of the HOA data frame representation ($C(k)$). Then the lowest integer number of bits is set to $\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil + 1) \rceil$.

【代表圖】

【本案指定代表圖】：第(1)圖。

【本代表圖之符號簡單說明】：

11：方向及向量估計處理步驟

12：HOA 分解處理步驟

13：周圍分量修改處理步驟

14：聲道指定步驟

15, 151：增益控制處理步驟

16：知覺編碼器步驟

17：邊資訊信號源編碼器步驟

18：多工器

$\tilde{B}(k-2)$ ：輸出訊框

$C(k)$ ：初始訊框

$C_{AMB}(k-1)$ ：周圍 HOA 分量之訊框

$C_{M,A}(k-1)$ ：修改周圍 HOA 分量

$C_{P,M,A}(k-1)$ ：暫預測修改周圍 HOA 分量

$e_1(k-2), \dots, e_I(k-2)$ ：指數

$\beta_1(k-2), \dots, \beta_I(k-2)$ ：異常旗標

$\mathcal{M}_{DIR}(k), \mathcal{M}_{VEC}(k), \mathcal{M}_{DIR}(k-1), \mathcal{M}_{VEC}(k-1)$ ：元組集

$v_{A,T}(k-1)$ ：目標指定向量

$v_A(k-2)$ ：最終指定向量

$X_{PS}(k-1)$ ：所有主要聲音信號框

$y_1(k-2), \dots, y_I(k-2)$ ：信號框

$y_{P,1}(k-1), \dots, y_{P,I}(k-1)$ ：預測信號框

$z_1(k-2), \dots, z_I(k-2)$ ：信號

$\tilde{z}_1(k-2), \dots, \tilde{z}_I(k-2)$ ：編碼信號

$\tilde{I}(k-2)$ ：編碼邊資訊

$\zeta(k-1)$ ：預測參數

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：無

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

用於解碼聲音或聲場的高階保真立體音響(HOA)表示的方法

METHOD FOR DECODING A HIGHER ORDER AMBISONICS (HOA)
REPRESENTATION OF A SOUND OR SOUNDFIELD

【技術領域】

【0001】本發明相關編碼高階保真立體音響(HOA)資料框表示，其包含非差分增益值係與高階保真立體音響資料框表示的資訊框特定者的聲道信號相關聯。

【先前技術】

【0002】高階保真立體音響(HOA)提供一可能性用以表示立體聲，其他技術係波場合成(WFS)或基於聲道的措施像”22.2”，對照到基於聲道的方法，HOA 表示提供不受特定揚聲器設置支配的優勢，然而，此彈性係以解碼過程作為代價，其要求在一特定揚聲器設置上回播 HOA 表示。相較於 WFS 措施，其中通常需要極大數量的揚聲器，HOA 亦可呈現到僅由極少揚聲器組成的設置。HOA 的另一優勢在於亦可利用相同表示，不用任何修改用於耳機的雙聲道呈現。

【0003】HOA 係基於複合平面諧波振幅藉由截斷球諧

函數(SH)展開的空間密度表示，各展開係數係一角頻率函數，其可等效地由一時域函數表示。因此，不失一般性，完整的 HOA 聲場表示實際上可理解為由 O 個時域函數組成，其中 O 表示展開係數的數目。以下此等時域函數將等效地稱為 HOA 係數序列或稱為 HOA 聲道。

【0004】HOA 表示的空間解析度係利用展開的成長最大階 N 得以提升，不幸地，展開係數的數目 O 隨著階 N 成二次方成長，尤其 $O = (N + 1)^2$ 。例如，使用階 $N = 4$ 的典型 HOA 表示需要 $O = 25$ 的 HOA(展開)係數。已知一期望單聲道取樣率 f_s 及每樣本的位元數 N_b ，用於 HOA 表示傳輸的總位元率係由 $O \cdot f_s \cdot N_b$ 判定，利用每樣本 $N_b = 16$ 位元，以 $f_s = 48\text{kHz}$ (千赫)的取樣率，傳輸階 $N = 4$ 的 HOA 表示造成 19.2 百萬位元秒的位元率，其用於許多實際應用如串流係極高位元率。因此高度期望 HOA 表示的壓縮。

【0005】HOA 聲場表示的壓縮先前曾揭示在歐洲專利號 EP2665208 A1、EP2743922 A1、EP2800401 A1 中，請參考 2014 年一月所頒佈 ISO/IEC JTC1/SC29/WG11，N14264，MPEG-H 立體聲的 WD1-HOA 內文。此等措施的共同點在於，其執行聲場分析並將已知 HOA 表示分解成方向分量及殘餘周圍分量。最終壓縮表示一方面係假設由數個量化信號組成，由方向信號及向量為基信號的知覺編碼以及周圍 HOA 分量的相關係數序列形成該等量化信號，另一方面，最終壓縮表示包括量化信號相關的額外邊資訊，其係 HOA 表示從其壓縮版本重建所需。

【0006】在傳遞到知覺編碼器前，要求此等中間時域信號具有值範圍 $[-1,1]$ 內的最大振幅，其係從目前可用知覺編碼器的實施引發的要求，為在壓縮 HOA 表示時滿足此要求，在知覺編碼器前面，使用一增益控制處理單元(參閱歐洲專利號 EP2824661 A1 及上述 ISO/IEC JTC1/SC29/ WG11 N14264 文件)，其平順地減弱或增大輸入信號。假設作為結果的信號修改係不可逆且係逐訊框應用，其中尤其假設連續框之間信號振幅的變化係'2'的乘幂。為促成此信號修改在 HOA 解壓縮器中的反轉，在總邊資訊中包括對應的正規化邊資訊，此正規化邊資訊可由底數'2'的指數組成，該等指數描述二連續框之間的相對振幅變化。由於連續框之間更可能發生小振幅變化而非較大振幅變化，因此根據上述 ISO/IEC JTC1/SC29/WG11 N14264 文件，使用遊程碼編碼此等指數。

【發明內容】

【0007】使用差分編碼振幅變化用以在 HOA 解壓縮中重建原始信號振幅係可行的，例如若單一檔案係從頭到尾不用任何時序跳躍以解壓縮，然而，為促進隨機存取，在編碼表示(其通常係一位元流)中必須存在獨立存取單位，為要允許解壓縮從一期望位置(或至少在其附近)開始，不用管先前訊框來的資訊。此一獨立存取單位必須包含增益控制處理單元從第一訊框直到目前訊框造成的總絕對振幅變化(即非差分增益值)，假設二連續框之間的振幅

變化係'2'的乘冪，亦藉由底數'2'的指數描述總絕對振幅變化即足夠。用於此指數的有效率編碼，在增益控制處理單元的應用前知道信號的潛在最大增益係必要的。然而，此知識係高度依賴待壓縮 HOA 表示的值範圍相關的限制規格，可惜 MPEG-H 立體聲文件 ISO/IEC JTC1/SC29/WG11 N14264 的確只提供格式描述用於輸入 HOA 表示，無設定值範圍相關的任何限制。

【0008】待由本發明解決的難題係提供非差分增益值表示所需的最低整數位元數，解決此難題係藉由後附申請專利範圍第 1 項中揭示的編碼高階保真立體音響(HOA)資料框表示。

【0009】在後附申請專利範圍的各別依附項中揭示本發明有利的附加實施例。

【0010】在應用 HOA 壓縮器內的增益控制處理單元前，本發明建立輸入 HOA 表示的值範圍與信號的潛在最大增益之間的相互關係，基於該相互關係，判定所需位元總數 - 用於一輸入 HOA 表示的值範圍的已知規格 - 以用於底數'2'的指數的有效率編碼，用以在一存取單位內描述修改信號由增益控制處理單元從第一訊框直到目前訊框造成的總絕對振幅變化(即非差分增益值)。

【0011】另外，一旦固定指數編碼所需位元總數的計算規則，本發明即使用一處理用以證實一已知 HOA 表示是否滿足所需值範圍限制，以便正確地壓縮該 HOA 表示。

【圖式簡單說明】

【0012】以下將參考附圖以描述本發明的示範實施例，圖中：

圖 1 顯示 HOA 壓縮器；

圖 2 顯示 HOA 解壓縮器；

圖 3 顯示定標值 K 用於虛擬方向 $\Omega_j^{(N)}$, $1 \leq j \leq O$ 以用於 HOA 階 $N = 1, \dots, 29$ ；

圖 4 顯示反模式矩陣 Ψ^{-1} 的歐幾里德範數用於虛擬方向 $\Omega_{\text{MIN},d}$, $d = 1, \dots, O_{\text{MIN}}$ 以用於 HOA 階 $N_{\text{MIN}} = 1, \dots, 9$ ；

圖 5 顯示虛擬揚聲器信號的最大允許量 γ_{dB} 的判定，在位置 $\Omega_j^{(N)}$, $1 \leq j \leq O$ ，其中 $O = (N + 1)^2$ ；

圖 6 顯示球面坐標系。

【實施方式】

【0013】即若未明確說明，以下實施例係可運用在任何組合或子組合中。

【0014】以下提出 HOA 壓縮及解壓縮的原理，為要提供發生上述問題的較詳細相關情境，此說明的基礎係 MPEG-H 立體聲文件 ISO/IEC JTC1/SC29/WG11 N14264 中所述處理，亦請參閱歐洲專利號 EP2665208 A1、EP2800401 A1 及 EP2743922 A1。在 N14264 中，'方向分量'係延伸到一'主要聲音分量'，作為方向分量，假設主要聲音分量係部分由方向信號表示，意指該等信號係具有對應方向的單聲道信號，假設其從該對應方向撞擊聆聽者，

連同一些預測參數用以從方向信號中預測部分的原始 HOA 表示。此外，亦假設主要聲音分量由‘向量為基信號’表示，意指該等信號係具有一對應向量的單聲道信號，該向量定義向量為基信號的方向分布。

【0015】

HOA 壓縮

圖 1 繪示歐洲專利號 EP2800401 A1 所揭示 HOA 壓縮器的整體架構，其具有一空間 HOA 編碼部分如圖 1A 繪示及一知覺及信號源編碼部分如圖 1B 繪示。空間 HOA 編碼器提供第一壓縮 HOA 表示，由 I 個信號連同描述如何產生其 HOA 表示的邊資訊組成，在將二編碼表示進行多工前，在知覺及邊資訊信號源編碼器中，將 I 個信號進行知覺編碼，並使邊資訊受信號源編碼。

【0016】

空間 HOA 編碼

在第一步驟中，將原始 HOA 表示的目前第 k 訊框 $C(k)$ 輸入到一方向及向量估計處理步驟或級 11，假設其提供元組集 $\mathcal{M}_{\text{DIR}}(k)$ 及 $\mathcal{M}_{\text{VEC}}(k)$ 。元組集 $\mathcal{M}_{\text{DIR}}(k)$ 係由元組組成，其第一元素表示方向信號索引及第二元素表示各別量化方向，元組集 $\mathcal{M}_{\text{VEC}}(k)$ 係由元組組成，其第一元素指出向量為基信號索引及第二元素表示定義信號方向分布的向量，即如何計算向量為基信號的 HOA 表示。

【0017】使用元組集 $\mathcal{M}_{\text{DIR}}(k)$ 及 $\mathcal{M}_{\text{VEC}}(k)$ 兩者，在一 HOA 分解步驟或級 12 中，將初始 HOA 訊框 $C(k)$ 分解成所有主

要聲音(即方向及向量為基)信號的訊框 $\mathbf{X}_{PS}(k-1)$ 及周圍 HOA 分量的訊框 $\mathbf{C}_{AMB}(k-1)$ 。請注意一訊框的延遲，其係由於交疊加處理，為要避免區塊效應。此外，為豐富主要聲音 HOA 分量，假設 HOA 分解步驟/級 12 輸出一些預測參數 $\zeta(k-1)$ ，描述如何從方向信號中預測部分的原始 HOA 表示。此外，假設待提供一目標指定向量 $\mathbf{v}_{A,T}(k-1)$ 到 I 個可用聲道，該向量含有 HOA 分解處理步驟或級 12 中所判定主要聲音信號的指定有關的資訊。可假設受影響的聲道被佔用，意指該等聲道不可在各別時間框中用以傳送周圍 HOA 分量的任何係數序列。

【0018】在周圍分量修改處理步驟或級 13 中，根據目標指定向量 $\mathbf{v}_{A,T}(k-1)$ 提供的資訊以修改周圍 HOA 分量的訊框 $\mathbf{C}_{AMB}(k-1)$ ，尤其(在其他方面之中)取決於哪些聲道係可用且未由主要聲音信號佔用的有關資訊(包含在目標指定向量 $\mathbf{v}_{A,T}(k-1)$ 中)，判定周圍 HOA 分量的哪些係數序列待傳輸在已知 I 個聲道中。此外，若選擇的係數序列索引在連續框之間有變化，則執行係數序列的淡入及淡出。

【0019】此外，假設總是選擇周圍 HOA 分量 $\mathbf{C}_{AMB}(k-2)$ 的第一 O_{MIN} 個係數序列待知覺編碼及傳輸，其中 $O_{MIN} = (N_{MIN} + 1)^2$ ， $N_{MIN} \leq N$ 通常係比原始 HOA 表示的階小的階。為將此等 HOA 係數序列去相關，可將其在步驟/級 13 中變換到一些預設方向 $\boldsymbol{\Omega}_{MIN,d}$ ， $d = 1, \dots, O_{MIN}$ 撞擊來的方向信號(即一般平面波函數)。

【0020】配合修改的周圍 HOA 分量 $\mathbf{C}_{M,A}(k-1)$ ，在步

驟/級 13 中計算一暫預測修改周圍 HOA 分量 $\mathbf{C}_{P,M,A}(k-1)$ ，並使用在增益控制處理步驟或級 15、151 中，為要允許一合理預見，其中周圍 HOA 分量修改有關的資訊係與聲道指定步驟或級 14 中所有可能信號類型指定到可用聲道直接相關。假設該指定有關的最終資訊係包含在最終指定向量 $\mathbf{v}_A(k-2)$ 中，為在步驟/級 13 中計算此向量，因此利用目標指定向量 $\mathbf{v}_{A,T}(k-1)$ 中包含的資訊。

【0021】步驟/級 14 中的聲道指定利用指定向量 $\mathbf{v}_A(k-2)$ 提供的資訊，將包含在訊框 $\mathbf{X}_{PS}(k-2)$ 中及包含在訊框 $\mathbf{C}_{M,A}(k-2)$ 中的適當信號指定到 I 個可用聲道，得出信號框 $\mathbf{y}_i(k-2)$ ， $i = 1, \dots, I$ 。另外，亦將包括在訊框 $\mathbf{X}_{PS}(k-1)$ 中及訊框 $\mathbf{C}_{P,AMB}(k-1)$ 中的適當信號指定到 I 個可用聲道，得出預測信號框 $\mathbf{y}_{P,i}(k-1)$ ， $i = 1, \dots, I$ 。

【0022】最後藉由增益控制 15、151 處理信號框 $\mathbf{y}_i(k-2)$ ， $i = 1, \dots, I$ 中的每一者，結果造成指數 $e_i(k-2)$ 及異常旗標 $\beta_i(k-2)$ ， $i = 1, \dots, I$ 及信號 $\mathbf{z}_i(k-2)$ ， $i = 1, \dots, I$ ，其中平順地修改信號增益，如用以達成適合知覺編碼器步驟或級 16 的值範圍。步驟/級 16 輸出對應的編碼信號框 $\tilde{\mathbf{z}}_i(k-2)$ ， $i = 1, \dots, I$ ，預測信號框 $\mathbf{y}_{P,i}(k-1)$ ， $i = 1, \dots, I$ 允許一種預見，為要避免連續區塊之間的嚴重增益變化。在邊資訊信號源編碼器步驟或級 17 中，將邊資訊資料 $\mathcal{M}_{DIR}(k-1)$ 、 $\mathcal{M}_{VEC}(k-1)$ 、 $e_i(k-2)$ 、 $\beta_i(k-2)$ 、 $\zeta(k-1)$ 及 $\mathbf{v}_A(k-2)$ 進行信號源編碼，結果造成編碼邊資訊框 $\tilde{\mathbf{r}}(k-2)$ ，在一多工器 18 中，將訊框 $(k-2)$ 的編碼信號 $\tilde{\mathbf{z}}_i(k-2)$ 與用於此訊框的編碼

邊資訊資料 $\tilde{\mathbf{I}}(k-2)$ 合併，結果造成輸出訊框 $\tilde{\mathbf{B}}(k-2)$ 。在一空間 HOA 解碼器中，假設步驟/級 15、151 中的增益修改係藉由使用指數 $e_i(k-2)$ 及異常旗標 $\beta_i(k-2)$ ， $i = 1, \dots, I$ 組成的增益控制邊資訊來回復。

【0023】

HOA 解壓縮

圖 2 繪示歐洲專利號 EP2800401 A1 揭露的 HOA 解壓縮器的整體架構，係由 HOA 壓縮器組件的相等類似者依相反次序配置所組成，及包括一知覺及信號源解碼部分如圖 2A 繪示及一空間 HOA 解碼部分如圖 2B 繪示。

【0024】在知覺及信號源解碼部分(表示一知覺及邊資訊信號源解碼器)中，一解多工步驟或級 21 接收位元流來的輸入訊框 $\tilde{\mathbf{B}}(k)$ ，及提供 I 個信號的知覺編碼表示 $\tilde{\mathbf{z}}_i(k)$ ， $i = 1, \dots, I$ ，及編碼邊資訊資料 $\tilde{\mathbf{I}}(k)$ ，描述如何產生其一 HOA 表示。在一知覺解碼器步驟或級 22 中，將 $\tilde{\mathbf{z}}_i(k)$ 信號知覺解碼，結果造成解碼信號 $\hat{\mathbf{z}}_i(k)$ ， $i = 1, \dots, I$ ，在一邊資訊信號源解碼器步驟或級 23 中，將編碼邊資訊資料 $\tilde{\mathbf{I}}(k)$ 解碼，結果造成資料集 $\mathcal{M}_{\text{DIR}}(k+1)$ 、 $\mathcal{M}_{\text{VEC}}(k+1)$ 、指數 $e_i(k)$ 、異常旗標 $\beta_i(k)$ 、預測參數 $\zeta(k+1)$ ，及一指定向量 $\mathbf{v}_{\text{AMB,ASSIGN}}(k)$ 。關於 $\mathbf{v}_A$ 與 $\mathbf{v}_{\text{AMB,ASSIGN}}$ 之間的差異，請參閱上述 MPEG 文件 N14264。

【0025】

空間 HOA 解碼

在空間 HOA 解碼部分中，將各知覺解碼信號 $\hat{\mathbf{z}}_i(k)$ ， $i =$

1, ..., I 連同其關聯增益校正指數 $e_i(k)$ 及增益校正異常旗標 $\beta_i(k)$ 一起輸入到一逆增益控制處理步驟或級 24、241。第 i 個逆增益控制處理步驟/級提供一增益校正信號框 $\hat{y}_i(k)$ 。

【0026】將所有 I 個增益校正信號框 $\hat{y}_i(k)$, $i = 1, \dots, I$ 連同指定向量 $\mathbf{v}_{\text{AMB,ASSIGN}}(k)$ 及元組集 $\mathcal{M}_{\text{DIR}}(k+1)$ 及 $\mathcal{M}_{\text{VEC}}(k+1)$ 饋到一聲道重指定步驟或級 25，請參閱上述元組集 $\mathcal{M}_{\text{DIR}}(k+1)$ 及 $\mathcal{M}_{\text{VEC}}(k+1)$ 的定義。指定向量 $\mathbf{v}_{\text{AMB,ASSIGN}}(k)$ 係由 I 個分量組成，該等分量指出各傳輸聲道是否包含周圍 HOA 分量的一係數序列及包含哪一者。在聲道重指定步驟/級 25 中，將增益校正信號框 $\hat{y}_i(k)$ 重分配，為要重建所有主要聲音信號(即所有方向及向量為基信號)的訊框 $\hat{\mathbf{X}}_{\text{PS}}(k)$ 及周圍 HOA 分量的一中間表示的訊框 $\mathbf{C}_{\text{I,AMB}}(k)$ 。此外，提供在第 k 訊框有效的周圍 HOA 分量的係數序列索引集 $\mathcal{J}_{\text{AMB,ACT}}(k)$ ，及周圍 HOA 分量的係數索引的資料集 $\mathcal{J}_{\text{E}}(k-1)$ 、 $\mathcal{J}_{\text{D}}(k-1)$ 及 $\mathcal{J}_{\text{U}}(k-1)$ ，其必須係賦能、去能及在第 $(k-1)$ 訊框保持有效。

【0027】在一主要音合成步驟或級 26 中，使用元組集 $\mathcal{M}_{\text{DIR}}(k+1)$ 、預測參數集 $\zeta(k+1)$ 、元組集 $\mathcal{M}_{\text{VEC}}(k+1)$ 及資料集 $\mathcal{J}_{\text{E}}(k-1)$ 、 $\mathcal{J}_{\text{D}}(k-1)$ 及 $\mathcal{J}_{\text{U}}(k-1)$ ，從所有主要聲音信號的訊框 $\hat{\mathbf{X}}_{\text{PS}}(k)$ 中計算出主要聲音分量 $\hat{\mathbf{C}}_{\text{PS}}(k-1)$ 的 HOA 表示。

【0028】在一環音聲合成步驟或級 27 中，使用周圍 HOA 分量的係數序列的索引集 $\mathcal{J}_{\text{AMB,ACT}}(k)$ (其係現用在第 k 訊框)，從周圍 HOA 分量的中間表示的訊框 $\mathbf{C}_{\text{I,AMB}}(k)$ 中產生周圍 HOA 分量框 $\hat{\mathbf{C}}_{\text{AMB}}(k-1)$ 。由於與主要聲音 HOA 分量的同

步化，因此引入一訊框的延遲。

【0029】最後在一 HOA 組成步驟或級 28 中，將周圍 HOA 分量框 $\hat{\mathbf{c}}_{\text{AMB}}(k-1)$ 與主要聲音 HOA 分量的訊框 $\hat{\mathbf{c}}_{\text{PS}}(k-1)$ 疊合，以便提供解碼 HOA 訊框 $\hat{\mathbf{c}}(k-1)$ 。

【0030】之後，空間 HOA 解碼器從 I 個信號及邊資訊中產生重建 HOA 表示，若在編碼端將周圍 HOA 分量變換到方向信號，則在步驟/級 27 中在解碼器端反轉該變換。

【0031】信號的潛在最大增益在 HOA 壓縮器內的增益控制處理步驟/級 15、151 前係高度依賴輸入 HOA 表示的值範圍，因此，首先定義一有意義值範圍用於輸入 HOA 表示，隨後在進入增益控制處理步驟/級前，在信號的潛在最大增益上作出斷定。

【0032】

輸入 HOA 表示的正規化

用以使用本發明的處理，在那之前要實施(總)輸入 HOA 表示信號的正規化，執行一逐訊框處理以用於 HOA 壓縮，其中相關段落高階保真立體音響基本原理中在方程(54)中規定的時間連續 HOA 係數序列的向量 $\mathbf{c}(t)$ ，將原始輸入 HOA 表示的第 k 訊框 $\mathbf{C}(k)$ 定義為

$$\mathbf{C}(k) := [\mathbf{c}((kL+1)T_s) \quad \mathbf{c}((kL+2)T_s) \quad \dots \quad \mathbf{c}((k+1)L T_s)] \in \mathbb{R}^{O \times L}, \quad (1)$$

其中 k 表示訊框索引， L 表示訊框長度(依樣本)， $O = (N+1)^2$ 表示 HOA 係數序列數目，及 T_s 指出取樣期間。

【0033】如在歐洲專利號 EP2824661 A1 中提及，由

於此等時域函數並非在呈現後由揚聲器所播放的信號，因此一 HOA 表示的有意義正規化自實際觀點看來，並非藉由在個別 HOA 係數序列 $c_n^m(t)$ 的值範圍上強加限制所達成。反而，更便利的是考慮‘等效空間域表示’，其係以 HOA 表示呈現到 O 個虛擬揚聲器信號 $w_j(t)$, $1 \leq j \leq O$ 所得到。假設各別虛擬揚聲器位置係藉由一球面坐標系表達，其中假設各位置位在單位球面上及具有半徑‘1’。因此，位置係可由階依存方向 $\boldsymbol{\Omega}_j^{(N)} = (\theta_j^{(N)}, \phi_j^{(N)})$, $1 \leq j \leq O$ 等效地表達，其中 $\theta_j^{(N)}$ 及 $\phi_j^{(N)}$ 分別表示斜度及方位角(亦請參閱圖 6 及其用於球面坐標系定義的說明)。此等方向應儘可能均勻地分布在單位球面上，用於特定方向的計算，請參閱如 J. Fliege 及 U. Maier 於 1999 年在多特蒙德大學數學系發表的技術報告，”計算球體體積公式之二階段方法(A two-stage approach for computing cubature formulae for the sphere)” ，網址在 <http://www.mathematik.uni-dortmund.de/lxx/research/projects/fliege/nodes/nodes.html>。此等位置通常係依賴‘均勻分布在球面上’的定義類型，因此，並非不明確的。

【0034】定義值範圍用於虛擬揚聲器信號比定義值範圍用於 HOA 係數序列有利，係因可直覺地將用於前者的值範圍同等地設成區間 $[-1,1]$ ，如用於傳統揚聲器信號假設 PCM 表示的情況。此導致一空間均勻分布量化誤差，以便量化有利地應用在相關實際聆聽的一領域中。在此相關情況中，一重要方面係可選擇每樣本的位元數係如通常

用於傳統揚聲器信號時一樣低，即 16，其增加效率，優於 HOA 係數序列的直接量化，其中通常要求每樣本較高位元數(如 24 或甚至 32)。

【0035】為詳細說明空間域中的正規化過程，將所有虛擬揚聲器信號彙總在一向量中作為 $\mathbf{w}(t):=$

$$[w_1(t) \quad \dots \quad w_O(t)]^T, \quad (2)$$

其中 $(\cdot)^T$ 表示換位，相關虛擬方向 $\boldsymbol{\Omega}_j^{(N)}$, $1 \leq j \leq O$ 的模式矩陣由 $\boldsymbol{\Psi}$ 表示，其係由 $\boldsymbol{\Psi} := [S_1 \quad \dots \quad S_O] \in \mathbb{R}^{O \times O}$ (3)

定義，具有 $S_j :=$ (4)

$$[s_0^0(\boldsymbol{\Omega}_j^{(N)}) \quad s_1^{-1}(\boldsymbol{\Omega}_j^{(N)}) \quad s_1^0(\boldsymbol{\Omega}_j^{(N)}) \quad s_1^1(\boldsymbol{\Omega}_j^{(N)}) \quad \dots \quad s_N^{N-1}(\boldsymbol{\Omega}_j^{(N)}) \quad s_N^N(\boldsymbol{\Omega}_j^{(N)})]^T,$$

可將呈現過程公式化為一矩陣乘法 $\mathbf{w}(t) = (\boldsymbol{\Psi})^{-1} \cdot \mathbf{c}(t)$. (5)

使用此等定義，有關虛擬揚聲器信號的合理要求係：

$$\|\mathbf{w}(lT_s)\|_\infty = \max_{1 \leq j \leq O} |w_j(lT_s)| \leq 1 \quad \forall l, \quad (6)$$

其意指要求各虛擬揚聲器信號的幅度位在 $[-1, 1]$ 的範圍內，時間 t 的一時間瞬間係由該等 HOA 資料框的樣本值的一樣本索引 l 與一樣本期間 T_s 表示。

【0036】揚聲器信號的總功率因此滿足條件

$$\|\mathbf{w}(lT_s)\|_2^2 = \sum_{j=1}^O |w_j(lT_s)|^2 \leq O \quad \forall l. \quad (7)$$

在圖 1A 的輸入 $\mathbf{c}(k)$ 的上游實施 HOA 資料框表示的呈現及正規化。

【0037】

增益控制前用於信號值範圍的結果

假設執行輸入 HOA 表示的正規化係根據段落輸入 HOA 表示的正規化中的說明，以下考慮信號 y_i , $i = 1, \dots, I$ 的

值範圍，該等信號係輸入到 HOA 壓縮器中的增益控制處理單元 15、151。此等信號係藉由將以下中的一或多者指定到 I 個可用聲道所產生：HOA 係數序列，或主要聲音信號 $x_{PS,d}$, $d = 1, \dots, D$ ，及 / 或周圍 HOA 分量 $c_{AMB,n}$, $n = 1, \dots, O$ (空間變換應用到其一部分) 中的特定係數序列。因此在方程 (6) 的正規化假說下，必須分析所述此等不同信號類型的可能值範圍。由於所有信號種類係從原始 HOA 係數序列在中間計算，因此要看一下其可能值範圍。圖 1A 及圖 2B 中未繪示 I 個聲道中只包含一或多個 HOA 係數序列的情況，即在此類情況中不需 HOA 分解、周圍分量修改及對應的合成區塊。

【0038】

用於 HOA 表示的值範圍的結果

從虛擬揚聲器信號中得到時間連續 HOA 表示係藉由

$$c(t) = \Psi w(t), \quad (8)$$

其係方程 (5) 中操作的逆操作，因此使用方式 (8) 及 (7)，將所有 HOA 係數序列的總功率定界限如下：

$$\|c(lT_s)\|_2^2 \leq \|\Psi\|_2^2 \cdot \|w(lT_s)\|_2^2 \leq \|\Psi\|_2^2 \cdot O, \quad (9)$$

在球諧函數的 N3D 正規化的假說下，可藉由 $\|\Psi\|_2^2 = K \cdot O$,

$$(10a)$$

寫出模式矩陣的平方歐幾里德範數，其中 $K = \frac{\|\Psi\|_2^2}{O}$

$$(10b)$$

表示模式矩陣的平方歐幾里德範數與 HOA 係數序列數目 O 之間的比率，此比率係依賴特定 HOA 階 N 及特定虛

擬揚聲器方向 $\boldsymbol{\Omega}_j^{(N)}$, $1 \leq j \leq O$ ，其可藉由將各別參數表附加到比率來表達如下：

$$K = K(N, \boldsymbol{\Omega}_1^{(N)}, \dots, \boldsymbol{\Omega}_O^{(N)}) \quad (10c)$$

【0039】圖 3 係根據上述 Fliege 等人文章用於 HOA 階 $N = 1, \dots, 29$ 以顯示 K 的值用於虛擬方向 $\boldsymbol{\Omega}_j^{(N)}$, $1 \leq j \leq O$ 。

【0040】結合所有先前爭議及考量，提供一上限用於 HOA 係數序列數量如下：

$$\|\mathbf{c}(lT_s)\|_\infty \leq \|\mathbf{c}(lT_s)\|_2 \leq \sqrt{K} \cdot O, \quad (11)$$

其中第一不等式直接由範數定義形成。

【0041】重要的是，要注意到方程(6)中的條件隱含方程(11)中的條件，但反過來卻不然，即方程(11)不隱含方程(6)。另一重要方面係，在近乎均勻分布虛擬揚聲器位置的假說下，模式矩陣 $\boldsymbol{\Psi}$ 的行向量(其表示相關虛擬揚聲器位置的模式向量)幾乎互為正交，及各具有 $N+1$ 的歐幾里德範數。此特性意指空間變換幾乎保留歐幾里德範圍，但一乘法常數除外，即

$$\|\mathbf{c}(lT_s)\|_2 \approx (N+1)\|\mathbf{w}(lT_s)\|_2. \quad (12)$$

真範數 $\|\mathbf{c}(lT_s)\|_2$ 越不同於方程(12)中的近似，越違反相關模式向量的正交假說。

【0042】

用於主要聲音信號的值範圍的結果

主要聲音信號的兩類型(方向及向量為基)的共同點在於，其對 HOA 表示的貢獻係利用 $N+1$ 的歐幾里德範數由單一向量 $\mathbf{v}_1 \in \mathbb{R}^O$ 描述，即 $\|\mathbf{v}_1\|_2 = N+1$.

$$(13)$$

若為方向信號，此向量對應到相關一特定信號源方向 $\boldsymbol{\Omega}_{S,1}$ 的模式向量，即

$$\boldsymbol{v}_1 = \boldsymbol{S}(\boldsymbol{\Omega}_{S,1}) \quad (14)$$

$$:= [S_0^0(\boldsymbol{\Omega}_{S,1}) \ S_1^{-1}(\boldsymbol{\Omega}_{S,1}) \ S_1^0(\boldsymbol{\Omega}_{S,1}) \ S_1^1(\boldsymbol{\Omega}_{S,1}) \ \dots \ S_N^{N-1}(\boldsymbol{\Omega}_{S,1}) \ S_N^N(\boldsymbol{\Omega}_{S,1})]^T \quad (15)$$

藉由一 HOA 表示，此向量描述進入信號源方向 $\boldsymbol{\Omega}_{S,1}$ 的一方向束。在向量為基信號的情況中，未限制向量 $\boldsymbol{v}_1$ 係相關任何方向的模式向量，及因此可描述單聲道向量為基信號的較一般方向分布。

【0043】以下考量 D 個主要聲音信號 $\boldsymbol{x}_d(t)$, $d = 1, \dots, D$ 的一般情形，該等信號可集中在向量 $\boldsymbol{x}(t)$ 中係根據

$$\boldsymbol{x}(t) = [x_1(t) \ x_2(t) \ \dots \ x_D(t)]^T. \quad (16)$$

必須基於矩陣

$$\boldsymbol{V} := [\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ \dots \ \boldsymbol{v}_D] \quad (17)$$

以判定此等信號，該矩陣係由表示單聲道主要聲音信號 $\boldsymbol{x}_d(t)$, $d = 1, \dots, D$ 的方向分布的所有向量 $\boldsymbol{v}_d$, $d = 1, \dots, D$ 形成。

【0044】用於主要聲音信號 $\boldsymbol{x}(t)$ 的有意義萃取，將以下限制寫成公式：

- a) 得到各主要聲音信號作為原始 HOA 表示的係數序列的線性組合，即

$$\boldsymbol{x}(t) = \boldsymbol{A} \cdot \boldsymbol{c}(t), \quad (18)$$

其中 $\boldsymbol{A} \in \mathbb{R}^{D \times O}$ 表示混合矩陣。

- b) 應選擇混合矩陣 $\boldsymbol{A}$ ，使其歐幾里德範數不超過值‘1’，即

$$\|A\|_2 \stackrel{!}{\leq} 1, \quad (19)$$

並使原始 HOA 表示與主要聲音信號者之間殘餘的平方歐幾里德範數(或等效地指乘冪)不大於原始 HOA 表示的平方歐幾里德範數(或等效地指乘冪)，即

$$\|c(t) - V \cdot x(t)\|_2^2 \stackrel{!}{\leq} \|c(t)\|_2^2 \quad (20)$$

藉由將方程(18)插入方程(20)中，可看出

方程(20)係同等於限制

$$\|I - V \cdot A\|_2 \stackrel{!}{\leq} 1, \quad (21)$$

其中 I 表示身份矩陣。

從方程(18)中及方程(19)中的限制，及從歐幾里德矩陣及向量範數的相容性，使用方程(18)、(19)及(11)，由

$$\|x(lT_s)\|_\infty \leq \|x(lT_s)\|_2 \quad (22)$$

$$\leq \|A\|_2 \|c(lT_s)\|_2 \quad (23)$$

$$\leq \sqrt{K} \cdot O, \quad (24)$$

找出一上限用於主要聲音信號的幅度。因此，確保主要聲音信號保持在原始 HOA 係數序列相同的範圍中(比較方程(11))，即

$$\boxed{\|x(lT_s)\|_\infty \leq \sqrt{K} \cdot O} \quad (25)$$

【 0045 】

範例用於混合矩陣的選擇

得到如何判定混合矩陣滿足限制(20)的範例係藉由計算主要聲音信號，使萃取後殘餘的歐幾里德範數減到最小，即

$$x(t) = \operatorname{argmin}_{x(t)} \|V \cdot x(t) - c(t)\|_2 \quad (26)$$

方程(26)中最小化問題的解係由 $\mathbf{x}(t) = \mathbf{V}^+ \mathbf{c}(t)$, (27)

提供，其中 $(\cdot)^+$ 指出莫耳-潘若斯 (Moore-Penrose) 偽逆。藉由比較方程(27)與方程(18)，在此範例中，隨後發生混合矩陣等於矩陣 $\mathbf{V}$ 的莫耳-潘若斯 (Moore-Penrose) 偽逆，即 $\mathbf{A} = \mathbf{V}^+$ 。

然而，仍必須選擇矩陣 $\mathbf{V}$ 滿足限制(19)，即 $\|\mathbf{V}^+\|_2 \leq 1$. (28)

若只是方向信號，其中矩陣 $\mathbf{V}$ 係模式矩陣相關一些來源信號方向 $\boldsymbol{\Omega}_{S,d}$, $d = 1, \dots, D$ ，即 $\mathbf{V} = [\mathbf{s}(\boldsymbol{\Omega}_{S,1}) \ \mathbf{s}(\boldsymbol{\Omega}_{S,2}) \ \dots \ \mathbf{s}(\boldsymbol{\Omega}_{S,D})]$, (29)

則藉由選擇來源信號方向 $\boldsymbol{\Omega}_{S,d}$, $d = 1, \dots, D$ 可滿足限制(28)，使任二鄰近方向的距離不會太小。

【0046】

結果用於周圍 HOA 分量的係數序列的值範圍

計算周圍 HOA 分量係藉由從原始 HOA 表示中減去主要聲音信號的 HOA 表示，即 $\mathbf{c}_{\text{AMB}}(t) = \mathbf{c}(t) - \mathbf{V} \cdot \mathbf{x}(t)$. (30)

若根據準則(20)以判定主要聲音信號 $\mathbf{x}(t)$ 的向量，可推斷如下

$$\|\mathbf{c}_{\text{AMB}}(lT_S)\|_\infty \leq \|\mathbf{c}_{\text{AMB}}(lT_S)\|_2 \quad (31)$$

$$\stackrel{(30)}{=} \|\mathbf{c}(lT_S) - \mathbf{V} \cdot \mathbf{x}(lT_S)\|_2 \quad (32)$$

$$\stackrel{(20)}{\leq} \|\mathbf{c}(lT_S)\|_2 \quad (33)$$

$$\stackrel{(11)}{=} \sqrt{K} \cdot O \quad (34)$$

【0047】

周圍 HOA 分量的空間變換係數序列的值範圍

在歐洲專利號 EP2743922 A1 所揭露 HOA 壓縮處理中及在上述 MPEG 文件 N14264 中的另一方面係，總是選擇周圍 HOA 分量的第一 O_{MIN} 個係數序列指定到傳輸聲道，其中 $O_{\text{MIN}} = (N_{\text{MIN}} + 1)^2$ ， $N_{\text{MIN}} \leq N$ 通常係較小階，小於原始 HOA 表示的階。為使此等 HOA 係數序列去相關，可將此等係數序列變換到一些預設方向 $\Omega_{\text{MIN},d}$ ， $d = 1, \dots, O_{\text{MIN}}$ 撞擊來的虛擬揚聲器信號(類似於段落輸入 HOA 表示的正規化中所述概念)。

【0048】定義周圍 HOA 分量的所有係數序列的向量具有階索引 $n \leq N_{\text{MIN}}$ (以 $\mathbf{c}_{\text{AMB},\text{MIN}}(t)$)及相關虛擬方向 $\Omega_{\text{MIN},d}$ ， $d = 1, \dots, O_{\text{MIN}}$ 的模式矩陣(以 Ψ_{MIN})，得到所有虛擬揚聲器信號的向量(定義以) $\mathbf{w}_{\text{MIN}}(t)$ 如下：

$$\mathbf{w}_{\text{MIN}}(t) = \Psi_{\text{MIN}}^{-1} \cdot \mathbf{c}_{\text{AMB},\text{MIN}}(t) \quad (35)$$

因此，使用歐幾里德矩陣及向量範數的相容性，

$$\|\mathbf{w}_{\text{MIN}}(lT_s)\|_{\infty} \leq \|\mathbf{w}_{\text{MIN}}(lT_s)\|_2 \quad (36)$$

$$\stackrel{(35)}{\leq} \|\Psi_{\text{MIN}}^{-1}\|_2 \cdot \|\mathbf{c}_{\text{AMB},\text{MIN}}(lT_s)\|_2 \quad (37)$$

$$\stackrel{(34)}{\leq} \|\Psi_{\text{MIN}}^{-1}\|_2 \cdot \sqrt{K} \cdot O \quad (38)$$

【0049】在上述 MPEG 文件 N14264 中，係根據上述 Fliege 等人文章以選擇虛擬方向 $\Omega_{\text{MIN},d}$ ， $d = 1, \dots, O_{\text{MIN}}$ ，在圖 4 中繪示模式矩陣 Ψ_{MIN} 的反矩陣的各別歐幾里德範數以用於階 $N_{\text{MIN}} = 1, \dots, 9$ ，可看出

$$\|\Psi_{\text{MIN}}^{-1}\|_2 < 1 \text{ 用於 } N_{\text{MIN}} = 1, \dots, 9 \quad (39)$$

然而，通常此不保持用於 $N_{\text{MIN}} > 9$ ，其中 $\|\Psi_{\text{MIN}}^{-1}\|_2$ 的值通常係遠大於‘1’。然而，至少用於 $1 \leq N_{\text{MIN}} \leq 9$ ，虛擬揚聲器

信號的幅度係定界限如下

$$\|w_{\text{MIN}}(IT_S)\|_{\infty} \leq \sqrt{K} \cdot O \quad \text{用於 } 1 \leq N_{\text{MIN}} \leq 9 \quad (40)$$

【0050】藉由限制輸入 HOA 表示以滿足條件(6)，其要求由此 HOA 表示產生的虛擬揚聲器信號的振幅不超過值 '1'，在以下條件下可保證信號的振幅在增益控制前不會超過值 $\sqrt{K} \cdot O$ (參閱方程(25)、(34)及(40))：

- 係根據方程/限制(18)、(19)及(20)以計算所有主要聲音信號 $x(t)$ 的向量；
- 若使用上述 Fliege 等人文章中定義的該等虛擬揚聲器位置時，最小階 N_{MIN} (其判定周圍 HOA 分量中應用空間變換的第一係數序列數目 O_{MIN}) 必須低於 '9'。

【0051】另外尚可推論出，信號的振幅在增益控制前不會超過值 $\sqrt{K_{\text{MAX}}} \cdot O$ 以用於任一階 N 直到感興趣最大階 N_{MAX} ，即 $1 \leq N \leq N_{\text{MAX}}$ ，其中 $K_{\text{MAX}} = \max_{1 \leq N \leq N_{\text{MAX}}} K(N, \boldsymbol{\Omega}_1^{(N)}, \dots, \boldsymbol{\Omega}_O^{(N)})$ (41a)

尤其，從圖 3 可推論出，若假設係根據 Fliege 等人文章中的分配以選擇虛擬揚聲器方向 $\boldsymbol{\Omega}_j^{(N)}, 1 \leq j \leq O$ 用於初始空間變換，及若額外假設感興趣最大階係 $N_{\text{MAX}} = 29$ (如在 MPEG 文件 N14264 中)，則由於此特殊情況中 $\sqrt{K_{\text{MAX}}} < 1.5$ ，信號的振幅在增益控制前不會超過值 $1.5 \cdot O$ ，即可選擇 $\sqrt{K_{\text{MAX}}} = 1.5$ 。

【0052】 K_{MAX} 係依賴感興趣最大階 N_{MAX} 及虛擬揚聲器方向 $\boldsymbol{\Omega}_j^{(N)}, 1 \leq j \leq O$ ，其可表達如下

$$K_{\text{MAX}} = K_{\text{MAX}} \left(\left\{ \Omega_1^{(N)}, \dots, \Omega_0^{(N)} \mid 1 \leq N \leq N_{\text{MAX}} \right\} \right) . \quad (41b)$$

因此，由增益控制為確保信號在知覺編碼前位在區間 $[-1,1]$ 內應用的最小增益係由 $2^{e_{\text{MIN}}}$ 提供，其中

$$e_{\text{MIN}} = -\lceil \log_2(\sqrt{K_{\text{MAX}}} \cdot 0) \rceil < 0 \quad (41c)$$

若信號的振幅在增益控制前太小，在 MPEG 文件 N14264 中揭示，可能平順地以高達 $2^{e_{\text{MAX}}}$ 的一因子增大信號，其中 $e_{\text{MAX}} \geq 0$ 係傳送作為編碼 HOA 表示內的邊資訊。

【0053】因此，底數‘2’的各指數(於存取單位內描述一修改信號由增益控制處理單元從第一訊框直到目前訊框造成的總絕對振幅變化)可假設區間 $[e_{\text{MIN}}, e_{\text{MAX}}]$ 內的任一整數值。因此，編碼所需(最低整數)位元數 β_e 係提供如下

$$\beta_e = \lceil \log_2(|e_{\text{MIN}}| + e_{\text{MAX}} + 1) \rceil = \lceil \log_2(\lceil \log_2(\sqrt{K_{\text{MAX}}} \cdot 0) \rceil + e_{\text{MAX}} + 1) \rceil . \quad (42)$$

若信號的振幅在增益控制前不會太小，可簡化方程 (42)：

$$\beta_e = \lceil \log_2(|e_{\text{MIN}}| + 1) \rceil = \lceil \log_2(\lceil \log_2(\sqrt{K_{\text{MAX}}} \cdot 0) \rceil + 1) \rceil \quad (42a)$$

可在增益控制步驟/級 15,...,151 的輸入計算此位元數 β_e 。

【0054】使用此位元數 β_e 用於指數，確保可捕捉到 HOA 壓縮器增益控制處理單元 15,...,151 造成的所有可能絕對振幅變化，允許在壓縮表示內的一些預設登錄點開始解壓縮。

【0055】當 HOA 解壓縮器中開始壓縮 HOA 表示的解壓縮時，依增益控制步驟/級 15,...,151 中實施處理的相反

方式，為應用一正確增益控制，在逆增益控制步驟或級 24,...,241 中使用非差分增益值(表示總絕對振幅變化，係指定到邊資訊用於一些資料框且從解多工器 21 中由接收的資料流 $\tilde{B}$ 中所接收)。

【0056】

進一步實施例

當實施如段落 HOA 壓縮、空間 HOA 編碼、HOA 分解及空間 HOA 解碼中所述特殊 HOA 壓縮/分解系統時，用於指數編碼的位元總數 β_e 必須根據方程(42)依一定標因子 $K_{\text{MAX,DES}}$ 設定，該定標因子本身係依賴待壓縮 HOA 表示的一期望最大階 $N_{\text{MAX,DES}}$ 及特定虛擬揚聲器方向 $\boldsymbol{\Omega}_{\text{DES},1}^{(N)}, \dots, \boldsymbol{\Omega}_{\text{DES},O}^{(N)}$, $1 \leq N \leq N_{\text{MAX}}$ 。

【0057】例如，當根據 Fliege 等人文章以假設 $N_{\text{MAX,DES}} = 29$ 及選擇虛擬揚聲器方向時，合理選擇會是 $\sqrt{K_{\text{MAX,DES}}} = 1.5$ 。在該情形中，保證正確壓縮用於階 N 的 HOA 表示， $1 \leq N \leq N_{\text{MAX}}$ ，其係根據段落輸入 HOA 表示的正規化，使用相同虛擬揚聲器方向 $\boldsymbol{\Omega}_{\text{DES},1}^{(N)}, \dots, \boldsymbol{\Omega}_{\text{DES},O}^{(N)}$ 進行正規化。然而，在以下情形中無法提供此保證：若一 HOA 表示(用於效率理由)亦同等地依 PCM 格式由虛擬揚聲器信號表示，但其中選擇虛擬揚聲器的方向 $\boldsymbol{\Omega}_j^{(N)}, 1 \leq j \leq O$ 係與在系統設計階段假設的虛擬揚聲器方向 $\boldsymbol{\Omega}_{\text{DES},1}^{(N)}, \dots, \boldsymbol{\Omega}_{\text{DES},O}^{(N)}$ 不同。

【0058】由於虛擬揚聲器位置的此不同選擇，即使此等虛擬揚聲器信號的振幅位在區間 $[1,1[$ 內，仍不再能保證

信號的振幅在增益控制前不會超過值 $\sqrt{K_{\text{MAX,DES}}} \cdot O$ ，及因此無法保證此 HOA 表示具有適當正規化用於根據 MPEG 文件 N14264 中所述處理的壓縮。

【0059】在此情況中，有利的是具有一系統，其基於虛擬揚聲器位置的知識，提供虛擬揚聲器信號的最大允許振幅以確保各別 HOA 表示適用根據 MPEG 文件 N14264 中所述處理的壓縮。在圖 5 中繪示此一系統，其採取虛擬揚聲器位置 $\Omega_j^{(N)}$, $1 \leq j \leq O$ 作為輸入，其中 $O = (N+1)^2$, $N \in \mathbb{N}_0$ ，及提供虛擬揚聲器信號的最大允許振幅 γ_{dB} (用分貝測量) 作為輸出。在步驟或級 51 中，係根據方程 (3) 以計算相關虛擬揚聲器位置的模式矩陣 Ψ ，在一隨後步驟或級 52 中，計算模式矩陣的歐幾里德範數 $\|\Psi\|_2$ ，在第三步驟或級 53 中，將振幅 γ 計算為 '1' 及虛擬揚聲器位置數與 $K_{\text{MAX,DES}}$ 的平方根的乘積與模式矩陣的歐幾里德範數之間的商數中的最小值，即

$$\gamma = \min \left(1, \frac{\sqrt{O} \cdot \sqrt{K_{\text{MAX,DES}}}}{\|\Psi\|_2} \right) \quad (43)$$

$$\text{得到分貝值係藉由 } \gamma_{\text{dB}} = 20 \log_{10}(\gamma) \quad (44)$$

【0060】用於說明：由以上導算可看出，若 HOA 係數序列的數量不超過值 $\sqrt{K_{\text{MAX,DES}}} \cdot O$ ，亦即，若

$$\|c(lT_s)\|_{\infty} \leq \sqrt{K_{\text{MAX,DES}}} \cdot O, \quad (45)$$

則所有信號在增益控制處理單元 15、151 前將因此不超過此值，其係用於適當 HOA 壓縮的要求。

【0061】從方程 (9) 中發現到 HOA 係數序列的數量係定界限如下

$$\|c(lT_s)\|_{\infty} \leq \|c(lT_s)\|_2 \leq \|\Psi\|_2 \cdot \|w(lT_s)\|_2 \quad (46)$$

因此，若 γ 係根據方程(43)設定及依 PCM 格式的虛擬揚聲器信號滿足

$$\|w(lT_s)\|_{\infty} \leq \gamma \quad (47)$$

$$\text{則由方程(7)推論出 } \|w(lT_s)\|_2 \leq \gamma \cdot \sqrt{O} \quad (48)$$

及滿足要求(45)，意即方程(6)中的最大量值‘1’係由方程(47)中的最大量值 γ 取代。

【0062】

高階保真立體音響的基本原理

高階保真立體音響(HOA)係基於感興趣緊密區內的聲場描述，其係假設為無音源。在該情形中，由同質波方程完全實體判定感興趣區內在時間 t 及位置 $\mathbf{x}$ 的聲壓 $p(t, \mathbf{x})$ 的時空反應。以下假設一球面坐標系，如圖 6 所示，在使用的坐標系中， x 軸指向前方位置， y 軸指向左方，及 z 軸指向上方。由一半徑 $r > 0$ (即到坐標原點的距離)、一斜角 $\theta \in [0, \pi]$ (自極軸 z (!) 測得) 及一方位角 $\phi \in [0, 2\pi]$ (在 $x-y$ 平面中自 x 軸反時鐘方向測得) 表示一空間位置 $\mathbf{x} = (r, \theta, \phi)^T$ 。另外， $(\cdot)^T$ 表示換位。

【0063】接著，可由”傅立葉聲學”教科書顯示，聲壓相關時間的傅立葉變換係由 $\mathcal{F}_t(\cdot)$ 表示，即

$$P(\omega, \mathbf{x}) = \mathcal{F}_t(p(t, \mathbf{x})) = \int_{-\infty}^{\infty} p(t, \mathbf{x}) e^{-i\omega t} dt \quad (49)$$

ω 表示角頻率及 i 表示虛數單位，根據

$$P(\omega = kc_s, r, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n A_n^m(k) j_n(kr) S_n^m(\theta, \phi) \quad (50)$$

可展開成球諧函數的級數。

其中， c_s 表示音速及 k 表示角波數，其係按照 $k = \frac{\omega}{c_s}$ 相關角度頻率 ω 。另外，

$j_n(\cdot)$ 表示第一類的球面 Bessel 函數，及 $S_n^m(\theta, \phi)$ 表示 n 階及 m 次的實數值球諧函數，其係定義在段落實數值球諧函數的定義中。展開係數 $A_n^m(k)$ 只取決於角波數 k ，請注意，已暗示地假設聲壓係空間上受頻帶限制。因此，在一上限 N 相關階索引 n 截斷該等級數，該上限稱為 HOA 表示的階。

【0064】若聲場係由從角度元組 (θ, ϕ) 規定的所有可能方向抵達的無限個不同角頻率 ω 的平面諧波疊加來表示，則可顯示(請參閱 B. Rafaely 的文章，”球體上之聲場藉由球面卷積之平面波分解(Plane-wave decomposition of the sound field on a sphere by spherical convolution)，美國聲學學會期刊，第 4(116)期，第 2149-2157 頁，2004 年 10 月)，各別平面波複合振幅函數 $C(\omega, \theta, \phi)$ 係可由以下球諧函數展開來表達：

$$C(\omega = kc_s, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n C_n^m(k) S_n^m(\theta, \phi) \quad (51)$$

其中展開係數 $C_n^m(k)$ 係相關展開係數 $A_n^m(k)$ 如下

$$A_n^m(k) = i^n C_n^m(k). \quad (52)$$

假設個別係數 $C_n^m(k = \omega/c_s)$ 係角頻率 ω 的函數，逆傅立葉變換(由 $\mathcal{F}^{-1}(\cdot)$ 表示)的應用提供時域函數

$$c_n^m(t) = \mathcal{F}_t^{-1}(C_n^m(\omega/c_s)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_n^m\left(\frac{\omega}{c_s}\right) e^{i\omega t} d\omega \quad (53)$$

以用於各 n 階及 m 次。此等時域函數在此稱為連續時間 HOA 係數序列，其可集中在單一向量 $\mathbf{c}(t)$ 中如下

$$\mathbf{c}(t) = \begin{bmatrix} c_0^0(t) & c_1^{-1}(t) & c_1^0(t) & c_1^1(t) & c_2^{-2}(t) & c_2^{-1}(t) & c_2^0(t) & c_2^1(t) & c_2^2(t) & \dots & c_N^{N-1}(t) & c_N^N(t) \end{bmatrix}^T \quad (54)$$

向量 $\mathbf{c}(t)$ 內的一 HOA 係數序列 $c_n^m(t)$ 的位置索引係由 $n(n+1)+1+m$ 提供。向量 $\mathbf{c}(t)$ 中的元素總數係由 $O = (N+1)^2$ 提供。

最終保真立體音響格式係使用一取樣頻率 f_s 以提供 $\mathbf{c}(t)$ 的取樣版本如下

$$\{\mathbf{c}(lT_s)\}_{l \in \mathbb{N}} = \{\mathbf{c}(T_s), \mathbf{c}(2T_s), \mathbf{c}(3T_s), \mathbf{c}(4T_s), \dots\} \quad (55)$$

其中 $T_s = 1/f_s$ 表示取樣期間， $\mathbf{c}(lT_s)$ 的元素在此稱為分離時間 HOA 係數序列，其係可顯示總為實數值。此特性明顯亦保持用於連續時間版本 $c_n^m(t)$ 。

【0065】

實數值球譜函數的定義

實數值球譜函數 $S_n^m(\theta, \phi)$ (假設 SN3D 正規化，係根據 J. Daniel 於 2001 年 6 月在巴黎大學發表的博士論文，名稱為”聲場之表示，應用至多媒體環境中複合聲音場景之傳輸及再製 (Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimedia)”，章節 3.1) 係提供如下

$$S_n^m(\theta, \phi) = \sqrt{(2n+1) \frac{(n-|m|)!}{(n+|m|)!}} P_{n,|m|}(\cos\theta) \text{trg}_m(\phi) \quad (56)$$

具有

$$\text{trg}_m(\phi) = \begin{cases} \sqrt{2}\cos(m\phi) & m > 0 \\ 1 & m = 0 \\ -\sqrt{2}\sin(m\phi) & m < 0 \end{cases} \quad (57)$$

相關 Legendre 函數 $P_{n,m}(x)$ 係定義為

$$P_{n,m}(x) = (1-x^2)^{m/2} \frac{d^m}{dx^m} P_n(x), m \geq 0 \quad (58)$$

具有 Legendre 多項式 $P_n(x)$ ，及不像在 E.G. Williams 的文章(傅立葉聲學 (*Fourier Acoustics*)，應用數學科學期刊，第 93 期，學術出版品，1999 年)中，並無 Condon-Shortley 相位項 $(-1)^m$ 。

【0066】實施本發明處理係可藉由單一處理器或電子電路，或藉由並聯操作或在本發明處理的不同部分操作的數個處理器或電子電路。

【0067】用以操作該處理器或該等處理器的指令可儲存在一或多個記憶體中。

【符號說明】

【0068】

圖 1

11：方向及向量估計處理步驟

12：HOA 分解處理步驟

13：周圍分量修改處理步驟

14：聲道指定步驟

15, 151：增益控制處理步驟

16：知覺編碼器步驟

17：邊資訊信號源編碼器步驟

18：多工器

$\tilde{B}(k-2)$ ：輸出訊框

$C(k)$ ：初始訊框

$C_{AMB}(k-1)$ ：周圍 HOA 分量的訊框

$C_{M,A}(k-1)$ ：修改周圍 HOA 分量

$C_{P,M,A}(k-1)$ ：暫預測修改周圍 HOA 分量

$e_1(k-2), \dots, e_I(k-2)$ ：指數

$\beta_1(k-2), \dots, \beta_I(k-2)$ ：異常旗標

$\mathcal{M}_{DIR}(k), \mathcal{M}_{VEC}(k)$ ：元組集

$\mathcal{M}_{DIR}(k-1), \mathcal{M}_{VEC}(k-1)$

$v_{A,T}(k-1)$ ：目標指定向量

$v_A(k-2)$ ：最終指定向量

$X_{PS}(k-1)$ ：所有主要聲音信號框

$y_1(k-2), \dots, y_I(k-2)$ ：信號框

$y_{P,1}(k-1), \dots, y_{P,I}(k-1)$ ：預測信號框

$z_1(k-2), \dots, z_I(k-2)$ ：信號

$\check{z}_1(k-2), \dots, \check{z}_I(k-2)$ ：編碼信號

$\tilde{r}(k-2)$ ：編碼邊資訊

$\zeta(k-1)$ ：預測參數

圖 2

21：解多工步驟

22：知覺解碼器步驟

23：邊資訊信號源解碼器步驟

24, 241：逆增益控制處理步驟

25：聲道重指定步驟

26：主要音合成步驟

27：環音聲合成步驟

28：HOA 組成步驟

$\check{B}(k)$ ：輸入訊框

$\hat{C}_{\text{AMB}}(k-1)$ ：周圍 HOA 分量訊框

$\hat{C}(k-1)$ ：解碼 HOA 訊框

$C_{\text{L,AMB}}(k)$ ：周圍 HOA 分量的中間表示訊框

$\hat{C}_{\text{PS}}(k-1)$ ：主要聲音 HOA 分量訊框

$e_1(k), \dots, e_I(k)$ ：增益校正指數

$\beta_1(k), \dots, \beta_I(k)$ ：增益校正異常旗標

$\mathcal{M}_{\text{DIR}}(k+1), \mathcal{M}_{\text{VEC}}(k+1)$ ：元組集

$v_{\text{AMB,ASSIGN}}(k)$ ：指定向量

$\hat{X}_{\text{PS}}(k)$ ：所有主要聲音信號框

$\hat{y}_1(k), \dots, \hat{y}_I(k)$ ：增益校正信號框

$\check{z}_1(k), \dots, \check{z}_I(k)$ ： I 個信號的知覺編碼表示

$\hat{z}_i(k), \dots, \hat{z}_i(k)$ ：解碼信號

$\check{I}(k)$ ：編碼邊資訊資料

$\zeta(k+1)$ ：預測參數

$J_{\text{AMB,ACT}}(k)$ ：周圍 HOA 分量的係數序列索引，在第 k 框中有效

$J_{\text{E}}(k-1), J_{\text{D}}(k-1)$ ：資料集

$J_{\text{U}}(k-1)$ ：

圖 3

K ：比率

N ：HOA 階

圖 4

N_{MIN} ：最小階

$\|\Psi_{\text{MIN}}^{-1}\|_2$ ：模式矩陣的反矩陣的歐幾里德範數

圖 5

51：計算模式矩陣

52：計算歐幾里德範數

53：計算增益

$\Omega_1^{(N)}, \dots, \Omega_0^{(N)}$ ：虛擬揚聲器的方向

Ψ ：模式矩陣

$\|\Psi\|_2$ ：模式矩陣的歐幾里德範數

γ_{dB} ：分貝值

圖 6

x, y, z ：坐標軸

r ：半徑

θ ：斜角

ϕ ：方位角

申請專利範圍

【請求項1】一種用於解碼聲音或聲場的壓縮的高階保真立體音響(HOA)聲音表示的方法，該方法包括：

接收含有該壓縮的HOA表示的位元流，其中，位元串流包含關於對應於該壓縮的HOA表示的HOA係數之一個數目的資訊；以及

基於最低整數位元數 β_e 解碼該壓縮的HOA表示，其中該最低整數位元數 β_e 基於 $\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil + e_{MAX} + 1) \rceil$ 確定，

其中 $K_{MAX} = \max_{1 \leq N \leq N_{MAX}} K(N, \Omega_1^{(N)}, \dots, \Omega_0^{(N)})$ ， N 係該壓縮的HOA表示的一個階， N_{MAX} 係該壓縮的HOA表示的感興趣最大階， $\Omega_1^{(N)}, \dots, \Omega_0^{(N)}$ 係虛擬揚聲器之方向， $O = (N+1)^2$ 係該等HOA係數序列之一個數目，及 K 係模式矩陣之平方歐基里德範數 $\|\Psi\|_2^2$ 與 O 間之比率，

其中 $e_{MAX} > 0$ ，並且

其中 $\sqrt{K_{MAX}} = 1.5$ 。

【請求項2】一種用於根據請求項1之方法執行解碼的非暫態儲存媒體。

【請求項3】一種用於解碼聲音或聲場的壓縮的高階保真立體音響(HOA)聲音表示的設備，該設備包括：

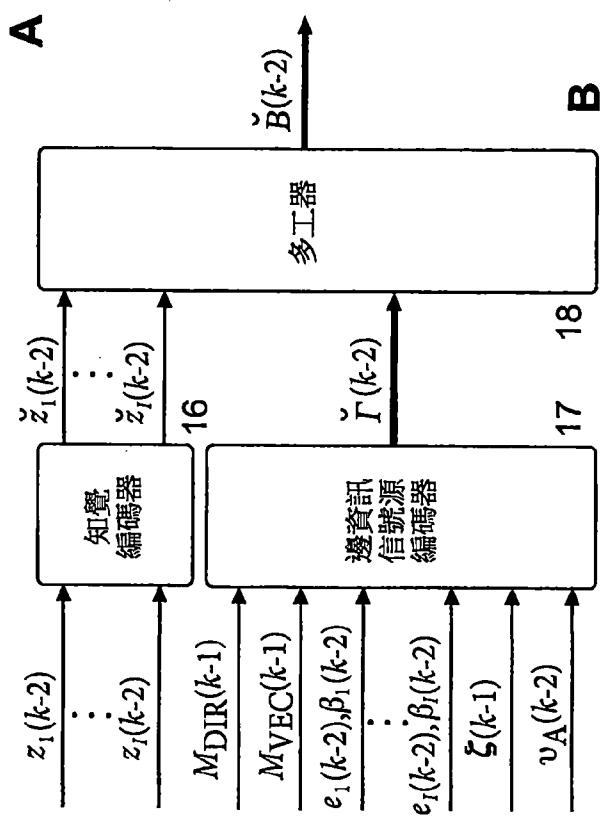
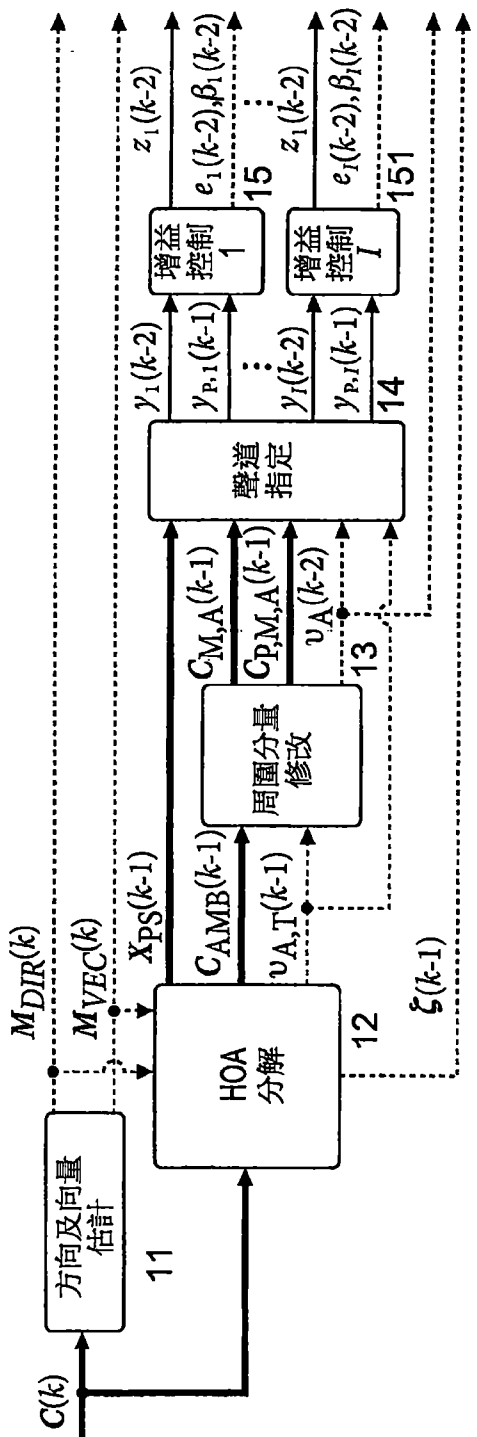
處理器，架構用於接收含有該壓縮的HOA表示的位元流，其中，位元串流包含關於對應於該壓縮的HOA表示的HOA係數之一個數目的資訊，並且其中，該處理器更被架構用於基於最低整數位元數 β_e 解碼該壓縮的HOA表示，

其中該最低整數位元數 β_e 基於 $\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{\text{MAX}}} \cdot O) \rceil + e_{\text{MAX}} + 1) \rceil$ 確定，

其中 $K_{\text{MAX}} = \max_{1 \leq N \leq N_{\text{MAX}}} K(N, \Omega_1^{(N)}, \dots, \Omega_O^{(N)})$ ， N 係該壓縮的 HOA 表示的一個階， N_{MAX} 係該壓縮的 HOA 表示的感興趣最大階， $\Omega_1^{(N)}, \dots, \Omega_O^{(N)}$ 係虛擬揚聲器之方向， $O = (N + 1)^2$ 係該等 HOA 係數序列之一個數目，及 K 係模式矩陣之平方歐基里德範數 $\|\Psi\|_2^2$ 與 O 間之比率，

其中 $e_{\text{MAX}} > 0$ ，

其中 $\sqrt{K_{\text{MAX}}} = 1.5$ 。



圖式

圖 1

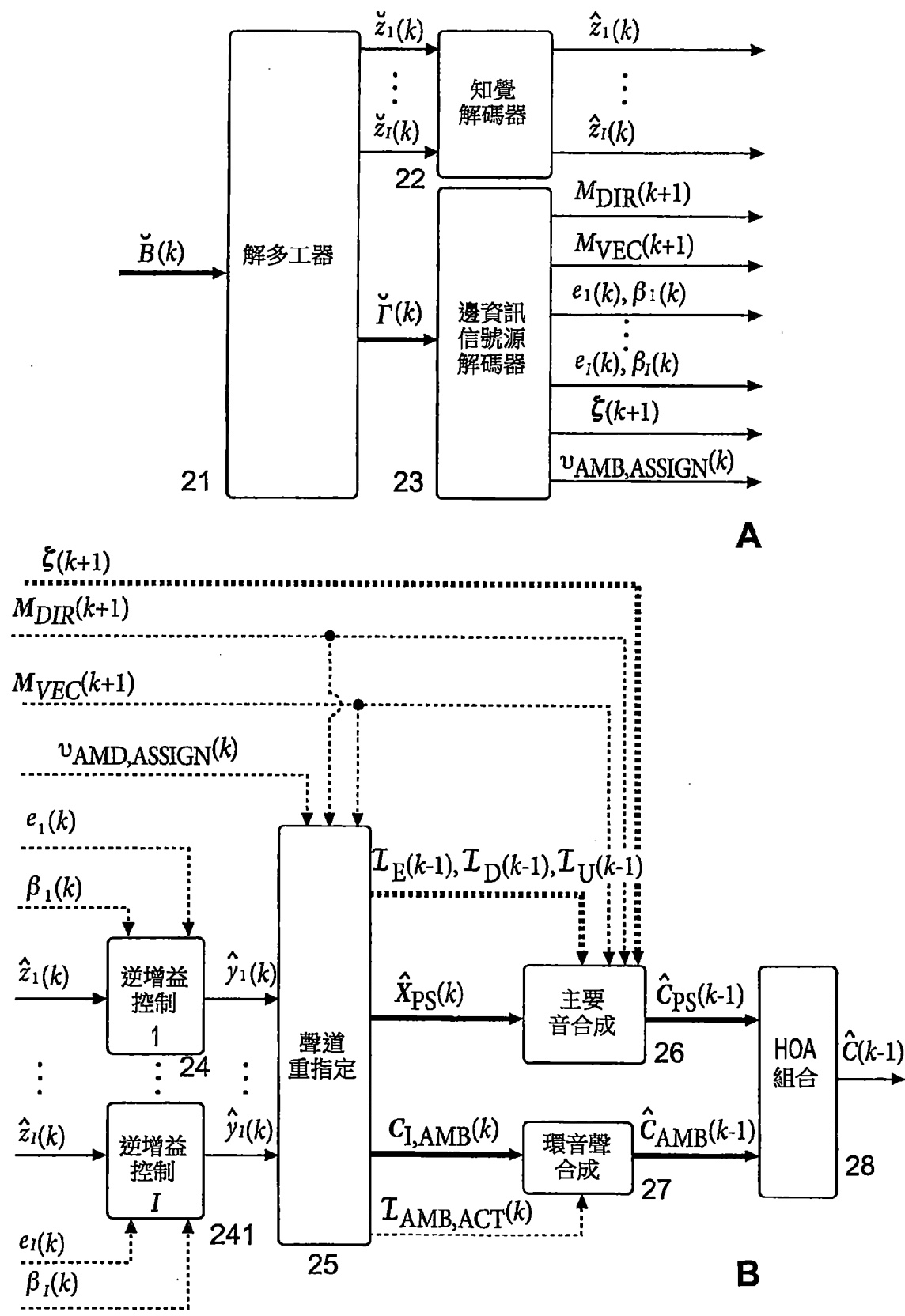


圖 2

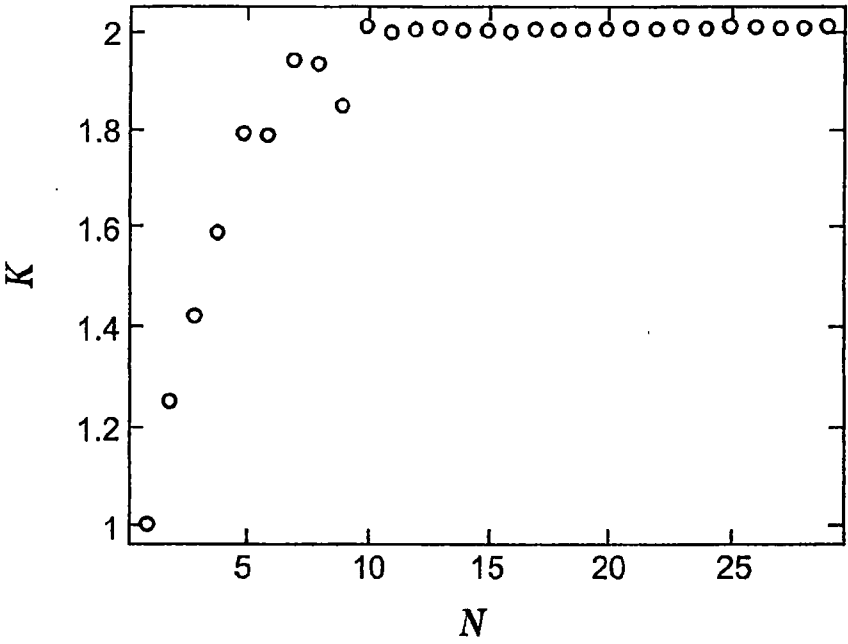


圖 3

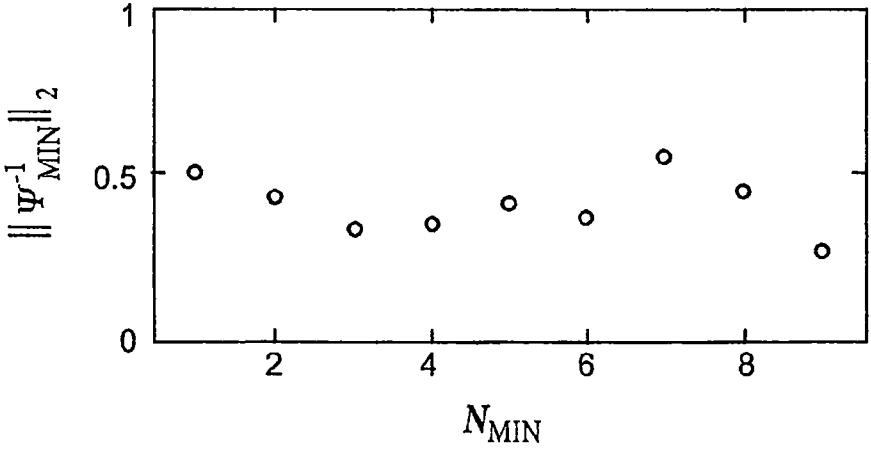


圖 4

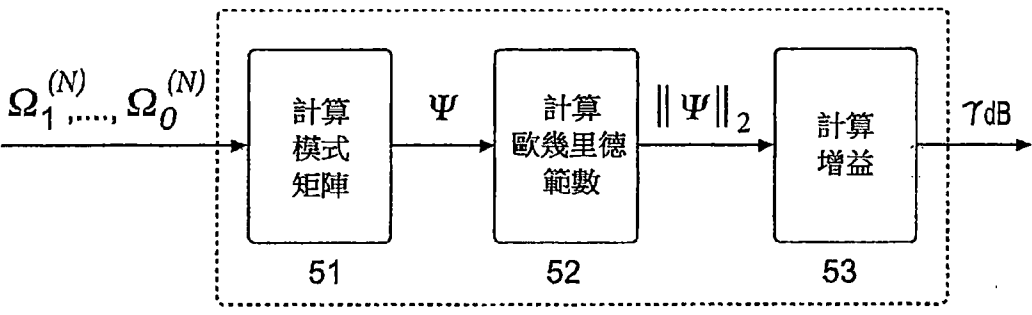


圖 5

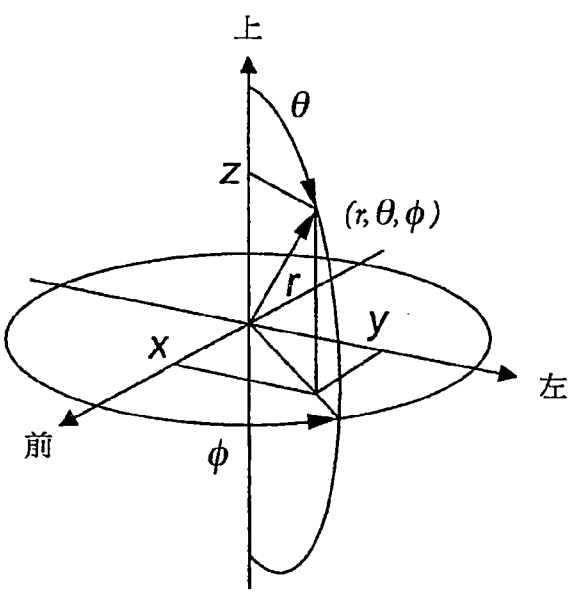


圖 6