



(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 I723805 B

(45)公告日：中華民國 110 (2021) 年 04 月 01 日

(21)申請案號：109108444

(22)申請日：中華民國 102 (2013) 年 07 月 12 日

(51)Int. Cl. : **H04S3/00 (2006.01)****H04S5/00 (2006.01)****G10L19/008 (2013.01)****G10L19/028 (2013.01)**

(30)優先權：2012/07/16 歐洲專利局

12305861.2

(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：波漢 約哈拿斯 BOEHM, JOHANNES (DE)；科登 斯凡 KORDON, SVEN

(DE)；克魯格 亞歷山德 KRUEGER, ALEXANDER (DE)；賈克斯 彼得 JAX, PETER (DE)

(74)代理人：林志剛

(56)參考文獻：

CN 101297353A

EP 2469741A1

US 2011/0305344A1

US 2012/0155653A1

US 2013/0010971A1

審查人員：李炳昌

申請專利範圍項數：3 項 圖式數：8 共 37 頁

(54)名稱

解碼高階立體音響 (HOA) 聲訊訊號之方法和設備及其電腦可讀取媒體

(57)摘要

一種編碼多通道 HOA 聲訊訊號以減少雜訊之方法，包括步驟為，使用逆適應 DSHT 令諸通道解相關(31)，逆適應 DSHT 包括旋轉操作(330)和逆 DSHT(310)，該旋轉操作旋轉 iDSHT 之空間抽樣柵格，以感知方式編碼(32)各解相關通道，編碼相關資訊(SI)，相關資訊包括界定該旋轉操作之參數，以及傳送或儲存以感知方式編碼之聲訊通道和編碼之相關資訊。

A method forencoding multi-channel HOA audio signals for noise reduction comprises steps of decorrelating (31) the channels using an inverse adaptive DSHT, the inverse adaptive DSHT comprising a rotation operation (330) and an inverse DSHT (310), with the rotation operation rotating the spatial sampling grid of the iDSHT, perceptually encoding (32) each of the decorrelated channels, encoding correlation information (SI), the correlation information comprising parameters defining said rotation operation, and transmitting or storing the perceptually encoded audio channels and the encoded correlation information.

指定代表圖：

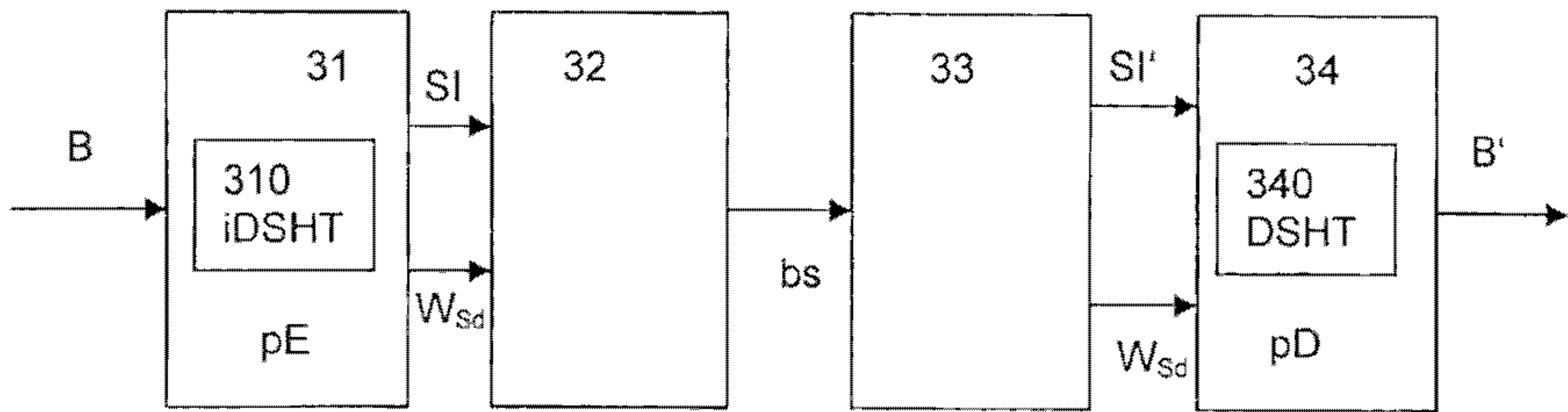


圖 3

符號簡單說明：

31:通道解相關步驟

32:各解相關通道以感知方式編碼步驟

33:接收資料解壓縮步驟

34:各通道以感知方式解碼步驟

310:逆 DSHT

340:解碼器內之構成方塊 DSHT

發明摘要

【發明名稱】(中文/英文)

解碼高階立體音響 (H O A) 聲訊訊號之方法和設備
及其電腦可讀取媒體

METHOD AND APPARATUS FOR DECODING HIGHER ORDER
AMBISONICS (HOA) AUDIO SIGNALS AND COMPUTER
READABLE MEDIUM THEREOF

【中文】

一種編碼多通道 HOA 聲訊訊號以減少雜訊之方法，包括步驟為，使用逆適應 DSHT 令諸通道解相關(31)，逆適應 DSHT 包括旋轉操作(330)和逆 DSHT(310)，該旋轉操作旋轉 iDSHT 之空間抽樣柵格，以感知方式編碼(32)各解相關通道，編碼相關資訊(SI)，相關資訊包括界定該旋轉操作之參數，以及傳送或儲存以感知方式編碼之聲訊通道和編碼之相關資訊。

【 英文 】

A method forencoding multi-channel HOA audio signals for noise reduction comprises steps of decorrelating (31) the channels using an inverse adaptive DSHT, the inverse adaptive DSHT comprising a rotation operation (330) and an inverse DSHT (310), with the rotation operation rotating the spatial sampling grid of the iDSHT, perceptually encoding (32) each of the decorrelated channels, encoding correlation information (SI), the correlation information comprising parameters defining said rotation operation, and transmitting or storing the perceptually encoded audio channels and the encoded correlation information.

【代表圖】

【本案指定代表圖】：第(3)圖。

【本代表圖之符號簡單說明】：

31：通道解相關步驟

32：各解相關通道以感知方式編碼步驟

33：接收資料解壓縮步驟

34：各通道以感知方式解碼步驟

310：逆 DSHT

340：解碼器內之構成方塊 DSHT

【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)

解碼高階立體音響 (H O A) 聲訊訊號之方法和設備
及其電腦可讀取媒體

METHOD AND APPARATUS FOR DECODING HIGHER ORDER
AMBISONICS (HOA) AUDIO SIGNALS AND COMPUTER
READABLE MEDIUM THEREOF

【技術領域】

[0001] 本發明係關於一種編碼多通道高階保真立體音響 (HOA) 聲訊訊號以減少雜訊之方法和裝置，以及對已減少雜訊的多通道 HOA 聲訊訊號解碼之方法和裝置。

【先前技術】

[0002] HOA 是一種多通道聲場表示法 [附註 4]，而 HOA 訊號為多通道聲訊訊號。多通道訊號表示法，尤其是 HOA 表示法，在特殊揚聲器設置上回放，需要特殊呈現，往往包含矩陣化操作。解碼後，保真立體音響訊號「被矩陣化」，即映射與例如揚聲器的實際空間位置相對應的新聲訊訊號。往往在單一通道之間存在有高度交互相關性。

[0003] 問題是曾經驗到在矩陣化操作後，編碼雜訊增加。在先前技術上，其原因未明。在以感知編碼器進行

壓縮之前，例如利用分立球諧函數轉換法 (DSHT)，將 HOA 訊號轉換到空間域時，也會發生此效應。

[0004] 用於 HOA 聲訊訊號表示法之通常壓縮方法，是對個別保真立體音響係數通道 [附註 7]，施加獨立的感知編碼器。詳言之，感知編碼器只考慮到在各個別單通道訊號內發生的雜訊罩覆效應進行編碼。然而，如此效應典型上為非線性。若將如此單通道矩陣化成新訊號，則容易發生雜訊未遮蔽。在以感知編碼器進行壓縮之前，利用分立球諧函數轉換法將 HOA 訊號轉換到空間域時，也會發生此效應 [附註 8]。

[0005] 此等多通道聲訊訊號表示法傳輸或儲存時，往往需要適當之多通道壓縮技術。通常，最後把 I 解碼訊號 $\hat{x}_i(l)$ ， $i = 1, \dots, I$ 矩陣化成 J 新訊號 $\hat{y}_j(l)$ ， $j = 1, \dots, J$ ，進行通道無關的感知解碼。矩陣化 (matrixing) 意指以加權方式添加或混合解碼之訊號 $\hat{x}_i(l)$ 。按照

$$\begin{aligned}\hat{\mathbf{x}}(l) &:= [\hat{x}_1(l) \dots \hat{x}_I(l)]^T \\ \hat{\mathbf{y}}(l) &:= [\hat{y}_1(l) \dots \hat{y}_J(l)]^T\end{aligned}$$

把全部訊號 $\hat{x}_i(l)$ ， $i = 1, \dots, I$ ，以及全部新訊號 $\hat{y}_j(l)$ ， $j = 1, \dots, J$ ，以向量配置。「矩陣化」源自事實上 $\hat{\mathbf{y}}(l)$ 是以數學方式從 $\hat{\mathbf{x}}(l)$ 通過矩陣操作所得：

$$\hat{\mathbf{y}}(l) = A\hat{\mathbf{x}}(l)$$

其中 A 指混合權值組成之混合矩陣。「混合」和「矩陣化」在此所用為同義字。使用混合/矩陣化之目的，是為任何特殊揚聲器設置用以呈現聲訊訊號。矩陣所依賴的特殊個別揚聲器設置，以及在操作當中矩陣化所用矩陣，在

感知編碼階段通常為未知的。

【發明內容】

[0006] 本發明記載適應性分立球諧函數轉換法(aDSHT)技術，把雜訊未遮蔽(unmask)效果（非所要）減到最小。又記載 aDSHT 如何整合到壓縮編碼器結構內。所述技術至少對 HOA 訊號特別有益。本發明之一優點是，減少要傳送的側資訊量。

[0007] 按照本發明一具體例，編碼多通道 HOA 聲訊訊號以減少雜訊之方法，包括步驟為，使用逆適應 DSHT 令通道解相關，逆適應 DSHT 包括旋轉操作和逆 DSHT(iDSHT)，以旋轉操作旋轉 iDSHT 之空間抽樣柵格，以感知方式編碼各解相關通道，編碼相關資訊，相關資訊包括界定該旋轉操作之參數，並傳送或儲存以感知方式編碼之聲訊通道和編碼之相關資訊。相關資訊包括所用 DSHT 柵格之至少一識別符，而旋轉資訊界定 DSHT 柵格之適應旋轉。

[0008] 按照本發明一具體例，解碼具有減少雜訊之已編碼多通道 HOA 聲訊訊號之方法，包括步驟為，接收已編碼多通道 HOA 聲訊訊號和通道相關資訊，解壓縮所接收資料，使用 DSHT 以感知方式解碼各通道，把以感知方式解碼之通道相關化，其中按照該相關資訊進行 DSHT 之空間抽樣柵格旋轉，並把相關的感知方式解碼通道矩陣化，其中獲得映射於揚聲器位置之可複製聲訊訊號。相關

資訊包括所用 DSHT 柵格之至少一識別符，和界定 DSHT 柵格適應性旋轉之旋轉資訊。

[0009] 用以解碼多通道 HOA 聲訊訊號之裝置記載於申請專利範圍第 3 項。

[0010] 在一面向中，電腦可讀式媒體具有可執行指令，促成電腦進行包括上述步驟之編碼方法，或進行包括上述步驟之解碼方法。

[0011] 本發明有利實施例，揭載於申請專利範圍附屬項、以下說明和附圖中。

【圖式簡單說明】

[0012]

第 1 圖表示對 M 個係數方塊進行比率壓縮之已知編碼器和解碼器；

第 2 圖表示使用習知 DSHT（分立球諧函數轉換）和習知逆 DSHT 把 HOA 訊號轉換入空間域所用編碼器和解碼器；

第 3 圖使用適應 DSHT 和適應逆 DSHT 把 HOA 訊號轉換入空間域之編碼器和解碼器；

第 4 圖表示測試訊號；

第 5 圖表示編碼器和解碼器構成方塊內所用電碼簿之球面抽樣位置例；

第 6 圖表示訊號適應 DSHT 構成方塊（ pE 和 pD ）；

第 7 圖為本發明第一實施例；

第 8 圖為本發明第二實施例。

【實施方式】

[0013] 茲參見附圖說明本發明實施例。

[0014] 第 2 圖表示已知系統，使用逆 DSHT 把 HOA 訊號轉換入空間域內。訊號經使用 iDSHT 21、比率壓縮 E1/解壓縮 D1，進行轉換，並使用 DSHT 24 再轉換成係數域 S24。與此不同的是，第 3 圖表示本發明系統：已知解決方法的 DSHT 處理方塊被以控制適應 DSHT 之處理方塊 31,32 取代。側資訊 SI 是在位元流 bs 內發送。

[0015] 下述為界定和說明未遮蔽的數學模式。假設指定分立時間多通道訊號，包含 I 通道 $x_i(m)$, $i = 1, \dots, I$ ，其中 m 指時間樣本索引。個別訊號可為實數值或複數值。把 M 樣本圖幅在時間樣本索引 $m_{\text{START}} + 1$ 起頭，假設其中個別訊號為固定的。相對應樣本依據下式被配置在矩陣 $\mathbf{X} \in \mathbb{C}^{I \times M}$ ，

$$\mathbf{X} := [\mathbf{x}(m_{\text{START}} + 1), \dots, \mathbf{x}(m_{\text{START}} + M)] \quad (1)$$

$$\text{其中 } \mathbf{x}(l) := [x_1(l), \dots, x_I(l)]^T \quad (2)$$

$(\cdot)^T$ 指轉置。相對應實驗相關矩陣得自下式：

$$\Sigma_{\mathbf{X}} := \mathbf{X} \mathbf{X}^H \quad (3)$$

其中 $(\cdot)^H$ 指聯合複數共軛和轉置。

[0016] 現假設把多通道圖幅編碼，因而在重建時引進編碼錯誤雜訊。因此，重見圖幅樣本之矩陣以 $\hat{\mathbf{X}}$ 註明，是根據下式由真樣本矩陣 $\mathbf{X}$ 和編碼雜訊組份 $\mathbf{E}$ 組

成：

$$\hat{\mathbf{X}} = \mathbf{X} + \mathbf{E} \quad (4)$$

$$\text{其中 } \mathbf{E} := [\mathbf{e}(m_{\text{START}} + 1), \dots, \mathbf{e}(m_{\text{START}} + L)] \quad (5)$$

$$\text{而 } \mathbf{e}(m) := [e_1(m), \dots, e_I(m)]^T \quad (6)$$

[0017] 由於假設各通道已單獨編碼，對 $i = 1, \dots, I$ 而言，可假設編碼雜訊訊號 $e_i(m)$ 彼此獨立。利用此性能和假設，即雜訊訊號是零平均，雜訊訊號的經驗相關矩陣由如下式對角線矩陣所給出：

$$\Sigma_{\mathbf{E}} = \text{diag}(\sigma_{e_1}^2, \dots, \sigma_{e_I}^2) \quad (7)$$

其中 $\text{diag}(\sigma_{e_1}^2, \dots, \sigma_{e_I}^2)$ 指在其對角線上有經驗雜訊訊號功率之對角線矩陣：

$$\sigma_{e_i}^2 = \frac{1}{M} \sum_{m=m_{\text{START}}+1}^{m_{\text{START}}+M} |e_i(m)|^2 \quad (8)$$

又一基本假設是，進行編碼使對各通道滿足訊雜比 (SNR)。不失一般通則，假設對各通道之預定 SNR 相等，即：

$$\text{SNR}_x = \frac{\sigma_{x_i}^2}{\sigma_{e_i}^2} \text{ for all } i = 1, \dots, I \quad (9)$$

$$\text{其中 } \sigma_{x_i}^2 := \frac{1}{M} \sum_{m=m_{\text{START}}+1}^{m_{\text{START}}+M} |x_i(m)|^2 \quad (10)$$

[0018] 茲考慮把重建訊號矩陣化成 J 新訊號 $y_j(m)$ ， $j = 1, \dots, J$ 。不引進任何編碼錯誤，矩陣化訊號之樣本矩陣可如此表示：

$$\mathbf{Y} = \mathbf{A}\mathbf{X} \quad (11)$$

其中 $\mathbf{A} \in \mathbb{C}^{J \times I}$ 指混合矩陣，而其中

$$\mathbf{Y} := [\mathbf{y}(m_{\text{START}} + 1), \dots, \mathbf{y}(m_{\text{START}} + M)] \quad (12)$$

$$\text{而 } \mathbf{y}(m) := [y_1(m), \dots, y_J(m)]^T \quad (13)$$

[0019] 然而由於編碼雜訊，矩陣化訊號之樣本矩陣為：

$$\hat{\mathbf{Y}} = \mathbf{Y} + \mathbf{N} \quad (14)$$

$\mathbf{N}$ 係含矩陣化雜訊訊號的樣本之矩陣，可表達為：

$$\mathbf{N} = \mathbf{A}\mathbf{E} \quad (15)$$

$$\mathbf{N} = [\mathbf{n}(m_{\text{START}} + 1) \dots \mathbf{n}(m_{\text{START}} + M)] \quad (16)$$

$$\text{其中 } \mathbf{n}(m) := [n_1(m) \dots n_J(m)]^T \quad (17)$$

係時間樣本索引 m 時，全部矩陣化雜訊訊號之向量。

[0020] 利用式(11)，矩陣化無雜訊訊號之經驗相關矩陣，可以下式表示：

$$\Sigma_{\mathbf{Y}} = \mathbf{A}\Sigma_{\mathbf{X}}\mathbf{A}^H \quad (18)$$

[0021] 因此，即 $\Sigma_{\mathbf{Y}}$ 對角線上的第 j 個元件的第 j 個的矩陣化無雜訊訊號之經驗冪可寫成：

$$\sigma_{y_j}^2 = \mathbf{a}_j^H \Sigma_{\mathbf{X}} \mathbf{a}_j \quad (19)$$

其中 $\mathbf{a}_j$ 是 $\mathbf{A}^H$ 的第 j 列，按照

$$\mathbf{A}^H = [\mathbf{a}_1, \dots, \mathbf{a}_J] \quad (20)$$

[0022] 同理，由式(15)可把矩陣化雜訊訊號之經驗相關矩陣改寫成：

$$\Sigma_{\mathbf{N}} = \mathbf{A}\Sigma_{\mathbf{E}}\mathbf{A}^H \quad (21)$$

[0023] 即 $\Sigma_{\mathbf{N}}$ 對角線上之第 j 個元件的第 j 個矩陣化雜訊訊號之經驗冪如下式：

$$\sigma_{n_j}^2 = \mathbf{a}_j^H \Sigma_{\mathbf{E}} \mathbf{a}_j \quad (22)$$

[0024] 因此，矩陣化訊號的經驗 SNR 可界定為：

$$\text{SNR}_{y_j} := \frac{\sigma_{y_j}^2}{\sigma_{n_j}^2} \quad (23)$$

使用式(19)和(22)可改寫成：

$$\text{SNR}_{y_j} = \frac{\mathbf{a}_j^H \boldsymbol{\Sigma}_X \mathbf{a}_j}{\mathbf{a}_j^H \boldsymbol{\Sigma}_E \mathbf{a}_j} \quad (24)$$

[0025] 利用 $\boldsymbol{\Sigma}_X$ 分解成其對角線和非對角線組份，即：

$$\boldsymbol{\Sigma}_X = \text{diag}(\sigma_{x_1}^2, \dots, \sigma_{x_I}^2) + \boldsymbol{\Sigma}_{X,NG} \quad (25)$$

$$\boldsymbol{\Sigma}_{X,NG} := \boldsymbol{\Sigma}_X - \text{diag}(\sigma_{x_1}^2, \dots, \sigma_{x_I}^2) \quad (26)$$

並利用性質：

$$\text{diag}(\sigma_{x_1}^2, \dots, \sigma_{x_I}^2) = \text{SNR}_x \cdot \text{diag}(\sigma_{e_1}^2, \dots, \sigma_{e_I}^2) \quad (27)$$

由假設(7)和(9)，全部通道的 SNR 常數 (SNR_x) 結果，最後為矩陣化訊號的經驗 SNR 得所需表現：

$$\text{SNR}_{y_j} = \frac{\mathbf{a}_j^H \text{diag}(\sigma_{x_1}^2, \dots, \sigma_{x_I}^2) \mathbf{a}_j}{\mathbf{a}_j^H \boldsymbol{\Sigma}_E \mathbf{a}_j} + \frac{\mathbf{a}_j^H \boldsymbol{\Sigma}_{X,NG} \mathbf{a}_j}{\mathbf{a}_j^H \boldsymbol{\Sigma}_E \mathbf{a}_j} \quad (28)$$

$$\text{SNR}_{y_j} = \text{SNR}_x \left(1 + \frac{\mathbf{a}_j^H \boldsymbol{\Sigma}_{X,NG} \mathbf{a}_j}{\mathbf{a}_j^H \text{diag}(\sigma_{x_1}^2, \dots, \sigma_{x_I}^2) \mathbf{a}_j} \right) \quad (29)$$

[0026] 由此表現方式可見此 SNR 是由預定 SNR, SNR_x 乘以視訊號相關矩陣 $\boldsymbol{\Sigma}_X$ 之對角線和非對角線分量而定之項所得。具體而言，如果訊號 $x_i(m)$ 彼此不相關，使 $\boldsymbol{\Sigma}_{X,NG}$ 變成零矩陣，則矩陣化訊號之經驗 SNR 等於預定 SNR，即

$$\text{SNR}_{y_j} = \text{SNR}_x \quad \text{對全部 } j = 1, \dots, J, \quad \text{若 } \boldsymbol{\Sigma}_{X,NG} = \mathbf{0}_{I \times I} \quad (30)$$

其中 $\mathbf{0}_{I \times I}$ 指零矩陣，有 I 行和 I 列。意即若 $x_i(m)$ 相關，

矩陣化訊號之經驗 SNR 可能偏離預定 SNR。在最壞情況， SNR_{y_j} 還遠低於 SNR_x 。此現象在此稱為矩陣化時雜訊未遮蔽。

[0027] 下一段簡略介紹高階立體保真音響(HOA)，並界定待處理的訊號（資料率壓縮）。

[0028] HOA 是根據假定無聲源的所關注緊密區域內的聲場之描述。在此情況下，關注區域（在球面座標）內，於時間 t 和位置 $\mathbf{x}=[r,\theta,\phi]^T$ 的聲壓 $p(t,\mathbf{x})$ 之空間時間行為，在物理上完全由單相波方程式決定。可見聲壓相對於時間之傅里葉轉換式，即：

$$P(\omega,\mathbf{x}) = \mathcal{F}_t \{ p(t,\mathbf{x}) \} \quad (31)$$

其中 ω 指角頻率（而 $\mathcal{F}_t \{ \}$ 相當於 $\int_{-\infty}^{\infty} p(t,\mathbf{x})e^{-\omega t}dt$ ），可按照[附註 10]展成球諧函數(SH)系列：

$$P(k c_s, \mathbf{x}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m(k) j_n(kr) Y_n^m(\theta, \phi) \quad (32)$$

[0029] 在方程式(32)內， c_s 指聲速，而 $k=\frac{\omega}{c_s}$ 為角波數。又， $j_n(\cdot)$ 表示 n 階第一種球面 Bessel 函數，和 $Y_n^m(\cdot)$ 指 n 階和 m 度之球諧函數(SH)。關於聲場之完整資訊實際上含在聲場係數 $A_n^m(k)$ 內。

[0030] 須知 SH 一般而言是複數值函數。然而，利用其適當線性組合，可得實數值函數，並相對於此等函數進行展開。

[0031] 相對於方程式(32)內壓力聲場說明，聲場可界定為：

$$D(k, c_s, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n B_n^m(k) Y_n^m(\Omega) \quad (33)$$

其中聲場或頻幅密度 [附註 9] $D(k, c_s, \Omega)$ 視角波數和角方向 $\Omega = [\theta, \phi]^T$ 而定。源場可包含遠場/近場、分立/連續源 [附註 1]。聲場係數 B_n^m 與聲場係數 A_n^m 有關 [附註 1]：

$$A_n^m = \begin{cases} 4\pi i^n B_n^m & \text{對遠場} \\ -i k h_n^{(2)}(kr_s) B_n^m & \text{對近場} \end{cases} \quad (34)$$

其中 $h_n^{(2)}$ 是第二種球面 Hankel 函數，而 r_s 為與原點之源距離。（使用正頻率和第二種球面 Hankel 函數為入射波，關係到 e^{-ikr} 。）

[0032] HOA 界域內之訊號可在頻率域或時間域內，以聲場或聲場係數之逆傅里葉轉換式表示。以下說明假設使用聲場係數之時間域表示法為有限數：

$$b_n^m = i\mathcal{F}_t \{ B_n^m \} \quad (35)$$

式(33)內之有限序列在 $n=N$ 截止。截止相當於空間帶寬限制。係數（或 HOA 通道）數為：

$$O_{3D} = (N+1)^2 \text{ 對 3D 而言} \quad (36)$$

或 $O_{2D} = 2N+1$ 只為 2D 說明。對稍後以揚聲器複製而言，係數 b_n^m 包括一時間樣本 m 之聲訊資訊。可儲存或再傳送，因此為資料率壓縮之標的。係數之單一時間樣本可以有 O_{3D} 元件之向量 $b(m)$ 表示：

$$b(m) := [b_0^0(m), b_1^{-1}(m), b_1^0(m), b_1^1(m), b_2^{-2}(m), \dots, b_N^N(m)]^T \quad (37)$$

而 M 時間樣本以矩陣 B 表示：

$$B := [b(m_{\text{START}} + 1), b(m_{\text{START}} + 2), \dots, b(m_{\text{START}} + M)] \quad (38)$$

[0033] 聲場之二維度表示法可藉圓諧函數展開推演。此可見於上述概述之特殊情況，使用固定傾角 $\theta = \frac{\pi}{2}$ 、係數之不同加權，和縮小到 O_{2D} 係數 ($m = \pm n$) 的集合。因此，以下考慮全部可應用於 2D 表示法。則球體需以圓面取代。

[0034] 以下說明從 HOA 係數域轉換至以通道為基本之空間域，或反之。方程式(33)可就單位球體，為 l 分立空間樣本位置 $\Omega_l = [\theta_l, \phi_l]^T$ ，使用時間域 HOA 係數改寫：

$$d_{\Omega_l} := \sum_{n=0}^N \sum_{m=-n}^n b_n^m Y_n^m(\Omega_l) \quad (39)$$

[0035] 假設 $L_{sd} = (N+1)^2$ 球面樣本位置 Ω_l ，可為 HOA 資料區塊 B ，以向量記號改寫：

$$W = \Psi_l B \quad (40)$$

其中 $W := [w(m_{START} + 1), w(m_{START} + 2), \dots, w(m_{START} + M)]$ 而 $w(m) = [d_{\Omega_1}(m), \dots, d_{\Omega_{L_{sd}}}(m)]^T$ 代表 L_{sd} 多通道訊號之單一時間樣本，而矩陣 $\Psi_l = [y_1, \dots, y_{L_{sd}}]^H$ 其中向量 $y_l = [Y_0^0(\Omega_l), Y_1^{-1}(\Omega_l), \dots, Y_N^N(\Omega_l)]^T$ 。若很有規律選擇球面樣本位置，則矩陣 Ψ_f 存在，而

$$\Psi_f \Psi_l = I \quad (41)$$

其中 I 為 $O_{3D} \times O_{3D}$ 識別矩陣。則相對應轉換成方程式(40)，可界定為：

$$B = \Psi_f W \quad (42)$$

[0036] 方程式(42)把 L_{sd} 球面訊號轉換成係數域，可改寫成順向轉換：

$$B = DSHT\{W\} \quad (43)$$

其中 $DSHT\{\}$ 指分立球諧函數轉換。轉換 O_{3D} 係數訊號相對應逆轉換為空間域，以形成 L_{sd} 通道為基本之訊號，而方程式(40)變成：

$$W = iDSHT\{B\} \quad (44)$$

[0037] 此項分立球諧函數轉換之定義，足夠在此考慮有關 HOA 資料之資料率壓縮，因為可以指定之係數 B 開始，且唯有 $B = DSHT\{iDSHT\{B\}\}$ 的情況有益。分立球諧函數轉換更嚴格之定義可查[附註 2]。為 $DSHT$ 推演此等位置之適當球面樣本位置和程序，可查[附註 3,4,5,6]。抽樣柵格之實施例如第 5 圖所示。

[0038] 具體而言，第 5 圖表示編碼器和解碼器構成方塊 pE, pD 所用電碼簿之球面抽樣位置例，即在第 5a 圖中 $L_{sd}=4$ ，第 5b 圖中 $L_{sd}=9$ ，第 5c 圖中 $L_{sd}=16$ ，而在第 5d 圖， $L_{sd}=25$ 。

[0039] 以下說明高階立體保真音響係數資料率壓縮和雜訊未遮蔽。首先，界定測試訊號以強調某些性能，用於下述。

[0040] 位於方向 Ω_{s_1} 之單一遠場源，以 M 分立時間樣本之向量 $g = [g(m), \dots, g(M)]^T$ 表示，可以 HOA 係數方塊代表，利用編碼：

$$B_g = yg^T \quad (45)$$

其中矩陣 B_g 類比方程式(38)，且編碼向量 $y = [Y_0^{0*}(\Omega_{s_1}), Y_1^{-1*}(\Omega_{s_1}), \dots, Y_N^{N*}(\Omega_{s_1})]^T$ 由在方向 $\Omega_{s_1} = [\theta_{s_1}, \phi_{s_1}]^T$ 評估的共軛

複合球諧函數組成（若使用即時加值 SH，共軛沒有效果）。測試訊號 B_g 可視為 HOA 訊號之最單純情況。更複雜訊號包含許多此等訊號疊置。

[0041] 關於 HOA 通道直接壓縮，以下顯示當 HOA 係數通道被壓縮時，何以會發生雜訊未遮蔽。HOA 資料 B 實際方塊的 O_{3D} 係數通道之直接壓縮和解壓縮，會類比方程式(4)引進編碼雜訊 E ：

$$\hat{B} = B + E \quad (46)$$

[0042] 假設常數 SNR_{B_g} 一如方程式(9)。欲經揚聲器重播此訊號，訊號需經描繪。此過程可由下式說明：

$$\hat{W} = A\hat{B} \quad (47)$$

其中解碼矩陣 $A \in \mathbb{C}^{L \times O_{3D}}$ （和 $A^H = [a_1, \dots, a_L]$ ）而矩陣 $\hat{W} \in \mathbb{C}^{L \times M}$ ，保有 L 擴音器訊號之 M 時間樣本。此類比方程式(14)。應用上述所述考量，揚場器通道 l 之 SNR 可載明為（類比方程式(29)）：

$$SNR_{w_l} = SNR_{B_g} \left(1 + \frac{a_l^H \Sigma_{B,NG} a_l}{a_l^H \text{diag}(\sigma_{B_1}^2, \dots, \sigma_{B_{O_{3D}}}^2) a_l} \right) \quad (48)$$

其中 $\sigma_{B_0}^2$ 係第 0 個對角線元件，而 $\Sigma_{B,NG}$ 保持下式之非對角線元件：

$$\Sigma_B = B B^H \quad (49)$$

[0043] 由於無法影響解碼矩陣 A ，因為希望能夠解碼至任意揚聲器佈置，矩陣 Σ_B 需變成對角線，以獲得 $SNR_{w_l} = SNR_{B_g}$ 。由方程式(45)和(49)，（ $B = B_g$ ） $\Sigma_B = y g^H g y^H = c y y^H$ 變成非對角線，有一定標量值 $c = g^T g$ 。

與 SNR_{B_g} 相較，在揚聲器通道之訊雜比 SNR_{w_l} 降低。但因在編碼階段，往往既不知源訊號 g ，又不知揚聲器佈置，係數通道之直接損耗壓縮，會導致失控的未遮蔽效應，尤其是對低資料率。

[0044] 以下說明使用 DSHT 後，當 HOA 係數在空間域內壓縮時，為何發生雜訊未遮蔽。

[0045] HOA 係數資料 B 之現時方塊，如方程式(40)所示，於使用球諧函數轉換式壓縮之前，轉換成空間域：

$$W_{sd} = \Psi_i B \quad (50)$$

其中逆轉換矩陣 Ψ_i 涉及 $L_{sd} \geq 0_{3D}$ 空間樣本位置，和空間訊號矩陣 $W_{SH} \in \mathbb{C}^{L_{sd} \times M}$ 。此等經壓縮和解壓縮，並增加量化雜訊（類比方程式(4)）：

$$\widehat{W}_{sd} = W_{sd} + E \quad (51)$$

其中編碼雜訊組份 E 係按照方程式(5)。再假設 SNR，則 SNR_{sd} 是所有空間通道一定。訊號轉換為係數域方程式(42)，使用轉換矩陣 Ψ_f ，具有方程式(41)性能： $\Psi_f \Psi_i = I$ 。係數 $\widehat{B}$ 之新方塊變成：

$$\widehat{B} = \Psi_f \widehat{W}_{sd} \quad (52)$$

[0046] 此訊號描繪至 L 揚聲器訊號 $\widehat{W} \in \mathbb{C}^{L \times M}$ ，應用解碼矩陣 A_D ： $\widehat{W} = A_D \widehat{B}$ 。此可用方程式(52)和 $A = A_D \Psi_f$ 改寫：

$$\widehat{W} = A \widehat{W}_{sd} \quad (53)$$

[0047] 於此， A 變成混合矩陣，其 $A \in \mathbb{C}^{L \times L_{sd}}$ 。方程式(53)應看做類比方程式(14)。再應用上述全部考量，擴音

器通道 l 之 SNR 可類似方程式 (29)，由下式載明：

$$SNR_{w_l} = SNR_{s_d} \left(1 + \frac{\mathbf{a}_l^H \boldsymbol{\Sigma}_{W_{s_d}, NG} \mathbf{a}_l}{\mathbf{a}_l^H \text{diag}(\sigma_{s_{d_1}}^2, \dots, \sigma_{s_{d_{L_{sd}}}}^2) \mathbf{a}_l} \right) \quad (54)$$

其中 $\sigma_{s_{d_l}}^2$ 係第 l 個對角線元件，而 $\boldsymbol{\Sigma}_{W_{s_d}, NG}$ 保持非對角線元件，如下式：

$$\boldsymbol{\Sigma}_{W_{s_d}} = W_{s_d} W_{s_d}^H \quad (55)$$

[0048] 因為無法影響 A_D （如果能夠描繪於任何揚聲器佈置），故對 A 無任何影響， $\boldsymbol{\Sigma}_{W_{s_d}}$ 需變成接近對角線，以保持所需 SNR：使用方程式 (45) 之簡單測試訊號 ($B = B_g$)，則 $\boldsymbol{\Sigma}_{W_{s_d}}$ 變成：

$$\boldsymbol{\Sigma}_{W_{s_d}} = c \boldsymbol{\Psi}_i \mathbf{y} \mathbf{y}^H \boldsymbol{\Psi}_i^H \quad (56)$$

其中常數 $c = \mathbf{g}^T \mathbf{g}$ 。使用固定球諧函數轉換 ($\boldsymbol{\Psi}_i, \boldsymbol{\Psi}_f \text{fixed}$)， $\boldsymbol{\Sigma}_{W_{s_d}}$ 只有在很罕見甚至更壞情況成為對角線，已如上述，

則此項 $\frac{\mathbf{a}_l^H \boldsymbol{\Sigma}_{W_{s_d}, NG} \mathbf{a}_l}{\mathbf{a}_l^H \text{diag}(\sigma_{s_{d_1}}^2, \dots, \sigma_{s_{d_{L_{sd}}}}^2) \mathbf{a}_l}$ 視係數訊號空間性能而定。因此，HOA 係數在球面域內之低率損耗壓縮，會導致 SNR 降低，以及失控之未遮蔽效果。

[0049] 本發明基本概念是使用適應 DSHT(aDSHT) 把雜訊未遮蔽效果減到最小，該適應 DSHT 係由 DSHT 相對於 HOA 輸入訊號的空間性能有關的空間抽樣柵格之轉動，和 DSHT 本身所構成。

[0050] 以下說明訊號適應 DSHT(aDSHT)，其具有配合 HOA 係數 O_{3D} 數量的許多球面位置 L_{sd} ，見方程式 (36)。首先選擇預設球面樣本柵格，一如習知非適應

DSHT。對 M 時間樣本區塊而言，旋轉球面樣本柵格，使下式所示項之對數最小化：

$$\sum_{l=1}^{L_{sd}} \sum_{j=1}^{L_{sd}} |\Sigma_{W_{sd} l, j}| - \sum (\sigma_{sd_1}^2, \dots, \sigma_{sd_{L_{sd}}}^2) \quad (57)$$

其中 $|\Sigma_{W_{sd} l, j}|$ 是 $\Sigma_{W_{sd}}$ 諸元件（矩陣列索引 l 和行索引 j ）之絕對值，而 $\sigma_{sd_l}^2$ 是 $\Sigma_{W_{sd}}$ 之對角線元件。此等於把方程式

(54) 之項 $\frac{\mathbf{a}_l^H \Sigma_{W_{sd}, NG} \mathbf{a}_l}{\mathbf{a}_l^H \text{diag}(\sigma_{sd_1}^2, \dots, \sigma_{sd_{L_{sd}}}^2) \mathbf{a}_l}$ 最小化。選擇之預設球面

抽樣柵格視 HOA 階而定，即 HOA 係數 O_{3D} 數量。所選擇型式之球面抽樣柵格隱然已知用於解碼，或可由所接收訊號，例如從 HOA 階或 HOA 係數之數量加以推導出。

[0051] 視覺上，此過程相當於 DSHT 球面抽樣柵格旋轉，其方式是單一空間樣本位置匹配最強源方向，如第 4 圖所示。使用方程式(45)之簡單測試訊號 ($\mathbf{B} = \mathbf{B}_g$)，可見方程式(55)之項 $\mathbf{W}_{sd}$ 變成向量 $\in \mathbb{C}^{L_{sd} \times 1}$ ，所有元件除了一個以外，都接近零。因此， $\Sigma_{W_{sd}}$ 變成接近對角線，可保持所需 SNR SNR_{sd} 。

[0052] 第 4 圖表示被轉換至空間域的測試訊號 $\mathbf{B}_g$ 。在第 4a 圖內使用預設抽樣柵格，而在第 4b 圖內使用 aDSHT 之旋轉柵格。空間通道之相關 $\Sigma_{W_{sd}}$ 值（以 dB 計），在相對應樣本位置周圍，以 Voronoi 分格之顏色/灰色變異表示。空間結構之各分格代表抽樣點，分格之明/暗代表訊號強度。由第 4b 圖可見，已發現最強源方向，

並旋轉抽樣柵格，使其一側（即單一空間樣本位置）匹配最強源方向。此側以白色表示（相當於強源方向），而其他側均暗色（相當於低源方向）。在第 4a 圖，即旋轉之前，無側面匹配最強源方向，有若干側面多少呈灰色，意即在個別抽樣點接到相當可觀（但非最大）強度之聲訊訊號。

[0053] 以下說明壓縮編碼器和解碼器內所用 aDSHT 之主要構成方塊。

[0054] 編碼器和解碼器構成方塊 pE 和 pD 細節，如第 6 圖所示。二種方塊擁有 DSHT 基礎之球面抽樣位置柵格之同樣電碼簿。起先，按照共同電碼簿，使用係數 O_{3D} 數選擇模組 pE 內 $L_{sd} = O_{3D}$ 位置之基礎柵格。 L_{sd} 必須傳送至方塊 pD，以啟動選擇同樣基礎之抽樣位置柵格，如第 3 圖所示。基礎抽樣柵格以矩陣 $\mathbf{\Omega}_{DSHT} = [\mathbf{\Omega}_1, \dots, \mathbf{\Omega}_{L_{sd}}]$ 說明，其中 $\mathbf{\Omega}_i = [\theta_i, \phi_i]^T$ 界定在單位球體上之位置。如上所述，第 5 圖表示基礎柵格之實施例。

[0055] 輸入到旋轉尋找方塊（構成方塊「找到最佳旋轉」）320 的是係數矩陣 B 。構成方塊負責旋轉基礎抽樣柵格，使方程式(57)的值最小。旋轉是以「軸角度」表示法表示，而與此旋轉有關之壓縮軸 ψ_{rot} 和旋轉角度 ϕ_{rot} 輸出至此構成方塊，做為側資訊 SI。旋轉軸 ψ_{rot} 可以藉由從原點至單位球體上位置之單位向量加以說明。於球面座標內，可由藉由兩個角度來結合： $\psi_{rot} = [\theta_{axis}, \phi_{axis}]^T$ ，具有不需傳送之一個隱涵的相關半徑。藉由使用訊號通知重新

使用先前使用的值以建立側資訊 SI 的特殊逃逸圖型，對三個角度 $\theta_{axis}, \phi_{axis}, \varphi_{rot}$ 進行量化和熵編碼。

[0056] 構成方塊 'Build Ψ_i ' 330 解碼旋轉軸和角度成為 $\hat{\psi}_{rot}$ 和 $\hat{\varphi}_{rot}$ ，並將此旋轉應用至基礎抽樣柵格 $\mathfrak{D}_{DSHT}$ ，以得到旋轉柵格 $\hat{\mathfrak{D}}_{DSHT} = [\hat{\Omega}_1, \dots, \hat{\Omega}_{L_{sd}}]$ 。輸出 $iDSHT$ 矩陣 $\Psi_i = [\mathbf{y}_1, \dots, \mathbf{y}_{L_{sd}}]$ ，係由向量 $\mathbf{y}_l = [Y_0^0(\hat{\Omega}_l), Y_1^{-1}(\hat{\Omega}_l), \dots, Y_N^N(\hat{\Omega}_l)]^T$ 推演得到。

[0057] 在構成方塊 'iDSHT' 310 內，HOA 係數資料 B 之實際方塊，利用 $\mathbf{W}_{sd} = \Psi_i B$ 轉換入空間域。

[0058] pD 之構成方塊 'Build Ψ_f ' 350 接收並解碼旋轉軸和角度成為 $\hat{\psi}_{rot}$ 和 $\hat{\varphi}_{rot}$ ，並應用此旋轉於基礎抽樣柵格 $\mathfrak{D}_{DSHT}$ ，以推演出旋轉柵格 $\hat{\mathfrak{D}}_{DSHT} = [\hat{\Omega}_1, \dots, \hat{\Omega}_{L_{sd}}]$ 。 $iDSHT$ 矩陣 $\Psi_i = [\mathbf{y}_1, \dots, \mathbf{y}_{L_{sd}}]$ 是以向量 $\mathbf{y}_l = [Y_0^0(\hat{\Omega}_l), Y_1^{-1}(\hat{\Omega}_l), \dots, Y_N^N(\hat{\Omega}_l)]^T$ 推演得到，而 DSHT 矩陣 $\Psi_f = \Psi_i^{-1}$ 是在解碼側計算。

[0059] 在解碼器 34 之構成方塊 'DSHT' 340 內，空間域資料 $\hat{\mathbf{W}}_{sd}$ 之實際方塊轉換回到係數域資料方塊 $\hat{B} = \Psi_f \hat{\mathbf{W}}_{sd}$ 。

[0060] 以下說明諸有益實施例，其含有壓縮編解碼器之總體構造。第一實施例可用單一 aDSHT。第二實施例使用頻帶中的複數 aDSHT。

[0061] 第 7 圖表示編碼器和解碼器二者之第一（基礎）實施例。具有 O_{3D} 係數通道 $b(m)$ 的索引 m 之 HOA 時間樣本，先儲存於緩衝器 71 內，形成 M 個樣本之方塊和時間索引 μ 。在上述構成方塊 pE72 內使用適應 iDSHT

將 $B(\mu)$ 轉換為空間域。空間訊號方塊 $W_{sd}(\mu)$ 輸入至 L_{sd} 聲訊壓縮單聲道編碼器 73 (像 AAC 或 MPEG-1 層 3(mp3) 編碼器) 或單一 AAC 多通道編碼器 (L_{sd} 通道)。位元流 S73 由具有整合側資訊 SI 的複數編碼器位元流圖幅之多工圖幅，或者整合有側資訊 SI (較佳作為輔助資料) 之單一多通道位元流構成。

[0062] 在一實施例中，亦如第 7 圖所示之個別壓縮解碼器構成區塊包含：把位元流解多工成為 L_{sd} 位元流加側資訊 SI 並把位元流饋送至 L_{sd} 單聲道解碼器；解碼至具有 M 樣本之 L_{sd} 空間聲訊通道，以形成方塊 $\hat{W}_{sd}(\mu)$ (在第 7 圖的方塊 74 內兼含在 L_{sd} 單聲道解碼器內之解多工和解碼)；並把 $\hat{W}_{sd}(\mu)$ 和側資訊 SI 饋送至訊號適應 DSHT 解碼構成方塊 pD。

[0063] 在另一實施例中，個別壓縮解碼器構成方塊包括：例如從儲存器接收位元流；並將之解碼成 L_{sd} 多通道訊號 $\hat{W}_{sd}(\mu)$ ；把側資訊 SI 解封裝並饋送該多通道訊號 $\hat{W}_{sd}(\mu)$ 和該側資訊 SI 至訊號適應 DSHT 解碼構成方塊 pD。在此實施例中，側資訊之解封裝和在 L_{sd} 單聲道解碼器內解碼係被包含在第 7 圖之方塊 74 內。

[0064] 在訊號適應 DSHT 解碼構成方塊 pD 內，使用具有側資訊 SI 的適應 DSHT， $\hat{W}_{sd}(\mu)$ 轉換至係數域，以形成 HOA 訊號 $B(\mu)$ 方塊，其係被儲存於緩衝器內，有待解幅以形成係數之時間訊號 $b(m)$ 。

[0065] $\hat{W}_{sd}(\mu)$ 被使用具有在 pD 內的 SI 之適應 DSHT

轉換為係數域，以形成 HOA 訊號 $B(\mu)$ 之方塊，這些信號係被儲存於緩衝器內以待解幅。經解幅後，它們形成係數之時間訊號 $b(m)$ 。

[0066] 上述第一實施例在某些條件下，會有二缺點：第一，由於空間訊號分佈變更，從方塊 μ 至 $\mu+1$ 會有組塊假影。第二，在同一時間會有超過一個的強訊號，使得 aDSHT 之解相關效果相當小。在頻率域內操作的第二實施例係針對此二缺點加以改進。aDSHT 應用於標度因數頻帶資料，其組合複數頻帶資料。利用時間頻率轉換 (TFT) 與覆層添加 (OLA) 處理的疊合方塊，來避免組塊假影。可以藉由使用本發明在 J 譜帶內，傳送 SI_j 資料率，在增加額外負擔的代價下，卻可達成改進的解相關。

[0067] 第二實施例有些細節如第 8 圖所示，說明如下：訊號 $b(m)$ 之各係數通道受到時間頻率轉換 (TFT)。廣用 TFT 之一例為修正餘弦轉換 (MDCT)。在 TFT 成幅中，建構成 50% 的疊合方塊（方塊索引 μ ），而 TFT 指方塊轉換。在譜帶化中，TFT 頻率帶被組合以形成 J 新譜帶和有關訊號 $B_j(\mu) \in \mathbb{C}^{O_{3D} \times K_j}$ ，其中 K_j 指帶 j 內頻率係數之數量。對各個這些譜帶，有一處理方塊 pE_j ，其建立訊號 $W_{j_{sd}}(\mu) \in \mathbb{C}^{L_{sd} \times K_j}$ 和側資訊 SI_j 。譜帶可匹配有損聲訊壓縮法之譜帶（像 AAC/mp3 標度因數帶），或具有較粗之顆粒性。在後一情況，「無 TFT 方塊之通道無關有損聲訊壓縮」方塊需把譜帶化重新配置。處理方塊作用像頻率域內之 L_{sd} 多通道聲訊編碼器，把一恆定位元率分配到各聲訊

通道。位元流在位元流封裝中格式化。

[0068] 解碼器接收並儲存部份位元流，將其解封裝並饋送聲訊資料至多通道聲訊解碼器（「無 TFT 之通道無關聲訊解碼」），以及側資訊 S_{ij} 饋送至 pD_j 。聲訊解碼器（「無 TFT 之通道無關聲訊解碼」）解碼聲訊資訊，格式化 J 譜帶訊號 $\vec{W}_{Jsd}(\mu)$ ，作為至 pD_j 的輸入，此等訊號在此轉換至 HOA 係數域，以形成 $\hat{B}_j(\mu)$ 。在「解頻帶化」中， J 個譜帶重新組群，以匹配 TFT 之帶化。它們在 $iTFT \& OLA$ 內，以方塊疊合覆層添加處理加以轉換至時間域。該輸出經解幅，以製作訊號 $\hat{b}(m)$ 。

[0069] 本發明係基於發現通道間之交叉相關造成 SNR 之提高。感知編碼器只會考慮發生在每個個別單一通道訊號內的編碼雜訊未遮蔽。然而，此等效應典型上為非線性。因此，當此等單通道矩陣化為新訊號時，可能發生雜訊未遮蔽。此即矩陣化操作後，何以編碼雜訊會增加之原因。

[0070] 本發明提出利用使不需要的雜訊未遮蔽效應最小化的適應分立球諧函數轉換 (aDSHT)，來對多數通道解相關。aDSHT 係整合在壓縮編碼器和解碼器構造內。

[0071] 因為它包含針對 HOA 輸入訊號之空間性能來調整 DSHT 的空間抽樣柵格的旋轉操作，所以它是適應的。aDSHT 包括適應旋轉和實際習知 DSHT。實際習知 DSHT 是一種矩陣，可按先前技術構成。將適應旋轉應用至該矩陣，導致通道間的相關性最小化，所以導致矩陣化

後之 SNR 增加的最小化。在一實施例中，旋轉軸和角度係由自動化搜尋操作找出。在另一實施例中，旋轉軸和角度是以分析方式找出。旋轉軸和角度經編碼和傳送，以使得能在解碼後和矩陣化之前進行重新相關，其中使用逆適應 DSHT(iaDSHT)。

[0072] 適應 DSHT 與其他轉換相較，尤其與 Karhunen-Loève 轉換(KLT)相較，有其特別優點。aDSHT 之一特點是，其旋轉 aDSHT 之空間抽樣柵格。為了正確解碼，需要旋轉資訊，其包括旋轉軸和旋轉角度。旋轉軸和旋轉角度被以側資訊 SI 傳送。旋轉軸亦可以藉二角度表達。諸如 KLT 等其他轉換也適用於旋轉和鏡映座標系統，但不能移動抽樣點。又，諸如 KLT 等之其他轉換需要轉換矩陣，以供正確解碼，使得轉換矩陣之係數需當作側資訊 SI 加以傳送。因此，由於此等轉換矩陣之係數遠較 aDSHT 的旋轉軸和旋轉角度有更多的資料，所以使用 aDSHT 之一優良效果是降低了待傳送的側資訊 SI 的量。aDSHT 之另一優點是由於空間適應性，其提供在聲訊訊號內之改進連續性。諸如 KLT 等的其他轉換，則容易造成訊號不連續，這通常為妨礙其用途之問題所在。此問題也被使用 aDSHT 所解決。

[0073] 在一實施例中，進行時間頻率轉換(TFT)和譜帶化，而 aDSHT/iaDSHT 單獨應用於各譜帶。

[0074] 在一實施例中，一種編碼多通道 HOA 聲訊訊號以減少雜訊之方法包括步驟為：使用逆適應 DSHT 令通

道解相關(31)，逆適應 DSHT 包括旋轉操作(330)和逆 DSHT(310)，該旋轉操作旋轉 iDSHT 之空間抽樣柵格；以感知方式編碼(32)各解相關通道；編碼旋轉資訊(SI)，該旋轉資訊包括界定該旋轉操作之參數；以及傳送或儲存以感知方式編碼之聲訊通道和編碼之旋轉資訊。

[0075] 一實施例另外包括傳送或儲存所用球面 DSHT 柵格索引（即 DSHT 抽樣柵格型式，例如其階）。

[0076] 在一具體例中，逆適應 DSHT 包括步驟為，選擇初始預設球面抽樣柵格；測定最強源方向；為 M 時間樣本方塊，旋轉球面抽樣柵格，使單一空間抽樣位置匹配最強源方向。

[0077] 在一具體例中，旋轉球面樣本柵格，使此項

$$\sum_{l=1}^{L_{sd}} \sum_{j=1}^{L_{sd}} |\Sigma_{W_{sd,l,j}}| - \sum (\sigma_{sd,1}^2, \dots, \sigma_{sd,L_{sd}}^2)$$

之對數減到最少，其中 $|\Sigma_{W_{sd,l,j}}|$ 是 $\Sigma_{W_{sd}}$ 諸元件（具有矩陣列索引 l 和行索引 j ）之絕對值，而 $\sigma_{sd,l}^2$ 是 $\Sigma_{W_{sd}}$ 之對角線元件。如上所述， $\Sigma_{W_{sd}}$ 是按照 $\Sigma_{W_{sd}} = W_{sd} W_{sd}^H$ 計算，其中 $W_{sd} = \Psi_1 B$ 是旋轉抽樣柵格的逆轉換矩陣 Ψ_1 和輸入訊號方塊 B 之乘積，而 W_{sd}^H 是其聯合複數共軛。

[0078] 在一實施例中，一種解碼具有被編碼以減少雜訊的多通道 HOA 聲訊訊號之方法包括步驟為，接收所編碼多通道 HOA 聲訊訊號、球面 DSHT 柵格索引和通道旋轉資訊(SI)；把所接收資料解壓縮(33)；使用適應 DSHT 以感知方式解碼(34)；把以感知方式解碼之通道相關化，

其中按照該旋轉資訊(SI)進行適應 DSHT 的空間抽樣柵格之旋轉；以及把相關的感知方式解碼之通道矩陣化，其中獲得映射於揚聲器位置之可複製聲訊訊號。球面 DSHT 柵格索引是抽樣柵格之獨特識別符，故容許解碼器在旋轉之前，重建抽樣柵格。柵格本身（即柵格點之座標）不需傳送、儲存或接收。

[0079] 在一實施例中，適應 DSHT 包括步驟為：為適應 DSHT 選擇初始預設抽樣柵格；為 M 時間樣本方塊，按照該相關資訊旋轉球面抽樣柵格。

[0080] 在一實施例中，相關資訊係具有二或三分量之空間向量 ψ_{rot} 。

[0081] 在一實施例中，相關資訊係包括二角度之空間向量 ($\psi_{rot} = [\theta_{axis}, \phi_{axis}]^T$)。

[0082] 在一實施例中，該等角度被量化並以特殊逃逸圖型進行熵編碼，該圖型發訊重新使用先前使用數值，以製作側資訊(SI)。

[0083] 在一實施例中，一種編碼多通道 HOA 聲訊訊號以減少雜訊之裝置，包括：解相關器，使用逆適應 DSHT 把諸通道解相關，逆適應 DSHT 包括旋轉操作和逆 DSHT(iDSHT)，該旋轉操作旋轉 iDSHT 之空間抽樣柵格；感知編碼器(E)，以感知方式編碼各解相關通道；側資訊編碼器，供編碼旋轉資訊，旋轉資訊包括界定該旋轉操作之參數；和界面，供傳送或儲存以感知方式編碼之聲訊通道和所編碼旋轉資訊。

[0084] 在一實施例中，編碼裝置包括轉換機構，供進行逆適應 DSHT，轉換機構具有處理器，以選擇初始預設球面抽樣柵格，決定最強源方向，並為 M 時間樣本方塊，旋轉球面抽樣柵格，使單一空間抽樣位置匹配最強源方向。

[0085] 在一實施例中，一種多媒體 HOA 聲訊訊號減少雜訊之解碼裝置包括：界面機構，供接收所編碼多通道 HOA 聲訊訊號、球面 DSHT 柵格索引和通道旋轉資訊；解壓縮模組，把所接收資料解壓縮；感知解碼器，使用 DSHT 以感知方式解碼各通道；相關器，使感知方式解碼之通道相關化，其中按照該旋轉資訊，進行旋轉 DSHT 之空間抽樣柵格；以及混合器，把已相關的感知方式解碼之通道矩陣化，其中獲得映射在揚聲器位置之可複製聲訊訊號。

[0086] 在一具體例中，解碼裝置包括處理器，為適應 DSHT 選擇初始預設球面抽樣柵格，並為 M 時間樣本之方塊，按照該相關資訊，旋轉球面抽樣柵格。

[0087] 在全部實施例中，減少雜訊至少關係到避免編碼雜訊未遮蔽效應。

[0088] 聲訊訊號之感知編碼意指適於人員感知的聲訊之編碼。應注意，以感知方式編碼聲訊訊號時，通常不是對寬頻聲訊訊號樣本進行量化，而是針對與人類感知有關之個別頻帶進行量化。因此，訊號功率與量化雜訊之比可在個別頻帶之間加以改變。

[0089] 上述技術可當作是對使用 Karhunen-Loève 轉換(KLT)的解相關作改進之替代方案。

[0090] 本發明已就較佳實施例圖示、說明，並舉出基本新穎特點，須知技術專家均可就所述裝置和方法、所揭示機件形式和細節及其操作，進行各種省略、置換、變更，不違本發明之精神。凡以實質上同樣方式，進行實質上同樣功用，以達成同樣結果的此等元件之組合，均在本發明範圍內。由一具體例之元件置換另一件，亦完全在意圖和設想之內。

[0091] 須知本發明純就實施例加以說明，可進行細部修飾，不違本發明範圍。

[0092] 說明書和（適當時）申請專利範圍及附圖之各特點，可單獨或以任何適當組合方式提供。諸特點可視適當情形在硬體、軟體，或二者組合方式實施。連接可視應用情形，實施無線連接或有線連接，不一定直接或專用。申請專利範圍內出現之參考數字只供說明，對申請專利範圍無限制效用。

附註文獻

[1] T.D. Abhayapala. Generalized framework for spherical microphone arrays: Spatial and frequency decomposition. In Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), (accepted) Vol. X, pp. , April 2008, Las Vegas, USA.

[2] James R. Driscoll and Dennis M. Healy Jr. Computing

fourier transforms and convolutions on the 2-sphere. *Advances in Applied Mathematics*, 15:202–250, 1994.

[3] Jörg Fliege. Integration nodes for the sphere, <http://www.personal.soton.ac.uk/jflw07/nodes/nodes.html>

[4] Jörg Fliege and Ulrike Maier. A two-stage approach for computing cubature formulae for the sphere. Technical Report, Fachbereich Mathematik, Universität Dortmund, 1999.

[5] R. H. Hardin and N. J. A. Sloane. Webpage: Spherical designs, spherical t-designs. <http://www2.research.att.com/~njas/sphdesigns>

[6] R. H. Hardin and N. J. A. Sloane. McLaren's improved snub cube and other new spherical designs in three dimensions. *Discrete and Computational Geometry*, 15:429–441, 1996.

[7] Erik Hellerud, Ian Burnett, Audun Solvang, and U. Peter Svensson. Encoding higher order Ambisonics with AAC. In 124th AES Convention, Amsterdam, May 2008.

[8] Peter Jax, Jan-Mark Batke, Johannes Boehm, and Sven Kordon. Perceptual coding of HOA signals in spatial domain. European patent application EP2469741A1 (PD100051).

[9] Boaz Rafaely. Plane-wave decomposition of the sound

field on a sphere by spherical convolution. J. Acoust. Soc. Am., 4(116):2149–2157, October 2004.

[10] Earl G. Williams. Fourier Acoustics, volume 93 of Applied Mathematical Sciences. Academic Press, 1999.

【符號說明】

[0093]

31：通道解相關步驟

32：各解相關通道以感知方式編碼步驟

33：接收資料解壓縮步驟

34：各通道以感知方式解碼步驟

71：緩衝器方塊

72：pE 方塊

73：單編碼器方塊

74：單解碼器方塊

75：pD 方塊

76：緩衝器方塊

310：逆 DSHT

320：找到最佳旋轉方塊

330：旋轉操作方塊

340：解碼器內之構成方塊 DSHT

350：pD 之構成方塊 Ψ_f

申請專利範圍

【請求項 1】一種用以解碼高階立體音響（HOA）聲訊訊號的方法，該方法包含：

根據感知解碼，解壓縮該 HOA 聲訊訊號，以決定對應於該 HOA 聲訊訊號的至少一 HOA 表示法；

根據球面抽樣柵格的旋轉，決定旋轉轉換；

根據該旋轉轉換與該 HOA 表示法，決定旋轉 HOA 表示法；及

描繪該旋轉 HOA 表示法，以輸出至一揚聲器設置。

【請求項 2】一種用以解碼高階立體音響（HOA）聲訊訊號的設備，該設備包含：

解碼器，被配置用以：

根據感知解碼，解壓縮該 HOA 聲訊訊號，以決定對應於該 HOA 聲訊訊號的 HOA 表示法；

根據球面抽樣柵格的旋轉，決定旋轉轉換；

根據該旋轉轉換與該 HOA 表示法，決定旋轉 HOA 表示法；及

描繪器，被配置用以描繪該旋轉 HOA 表示法，以輸出至一揚聲器設置。

【請求項 3】一種非暫態電腦可讀取媒體，包括指令，當所述指令為處理器所執行時，用以執行如請求項 1 之方法。

圖 式

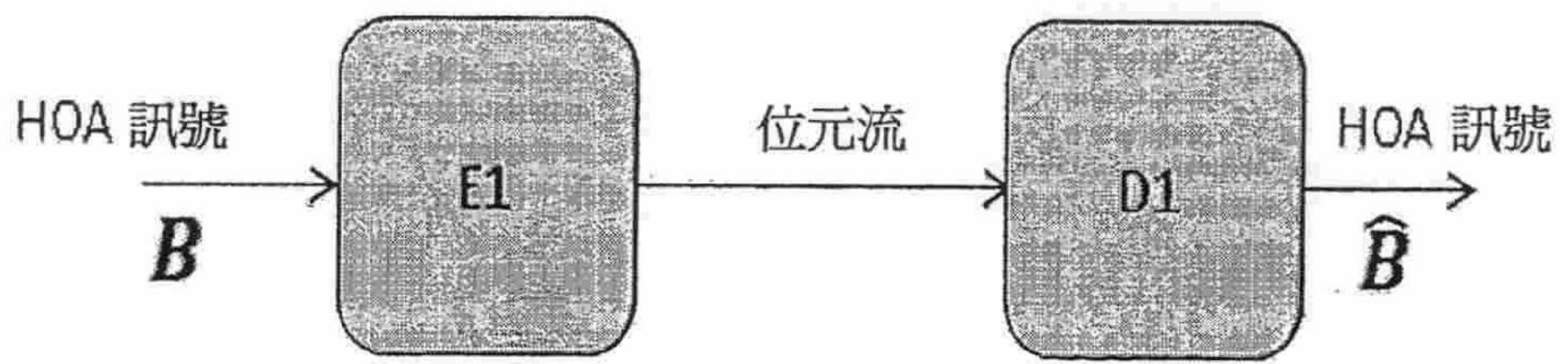


圖 1

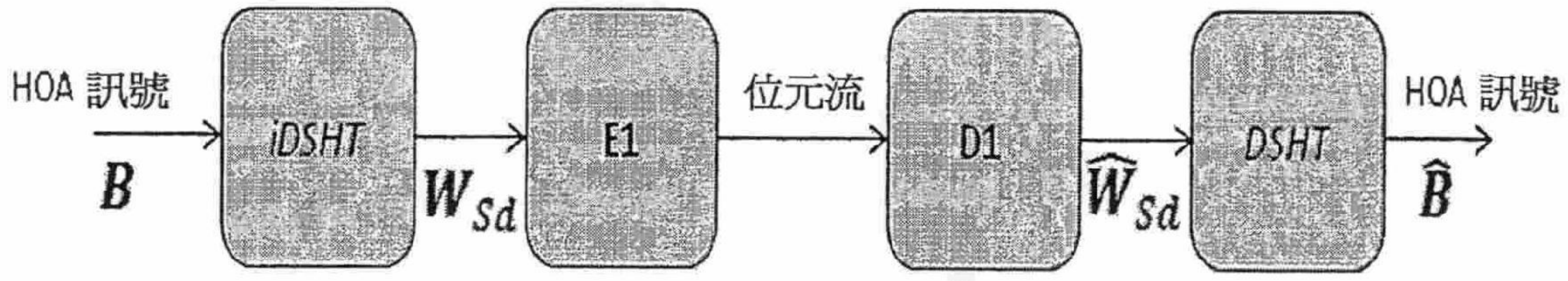


圖 2

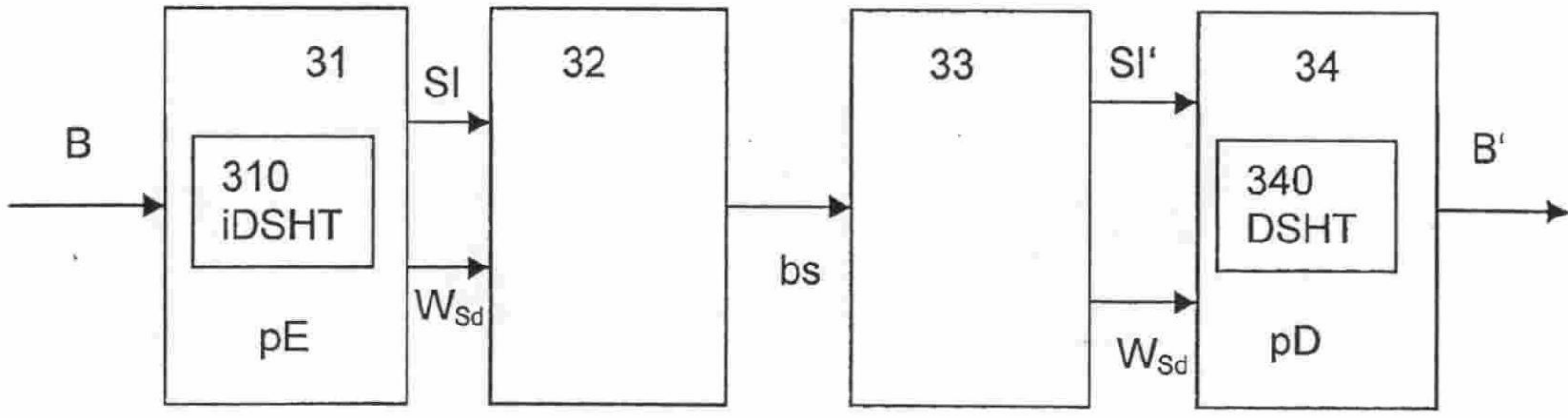
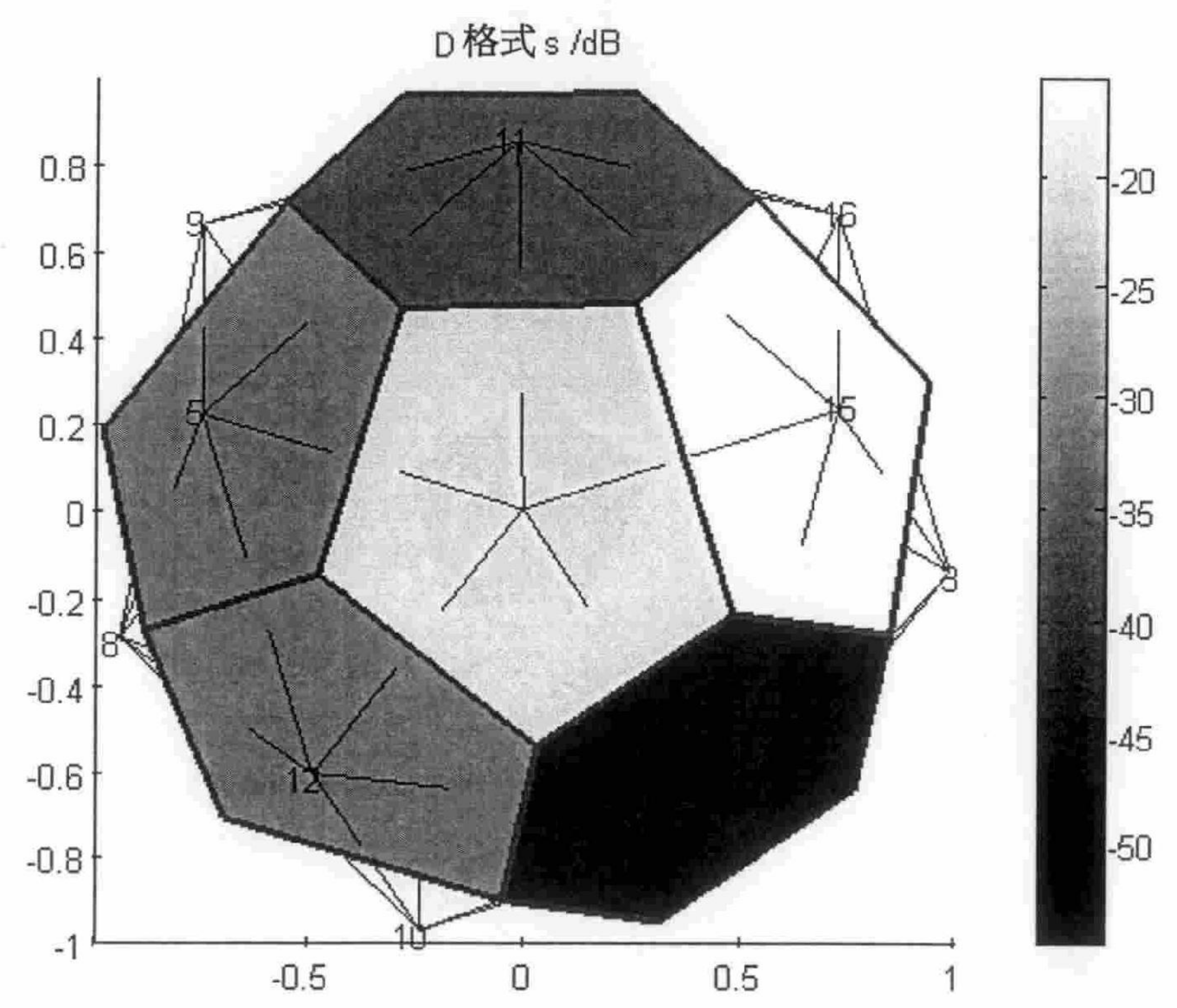
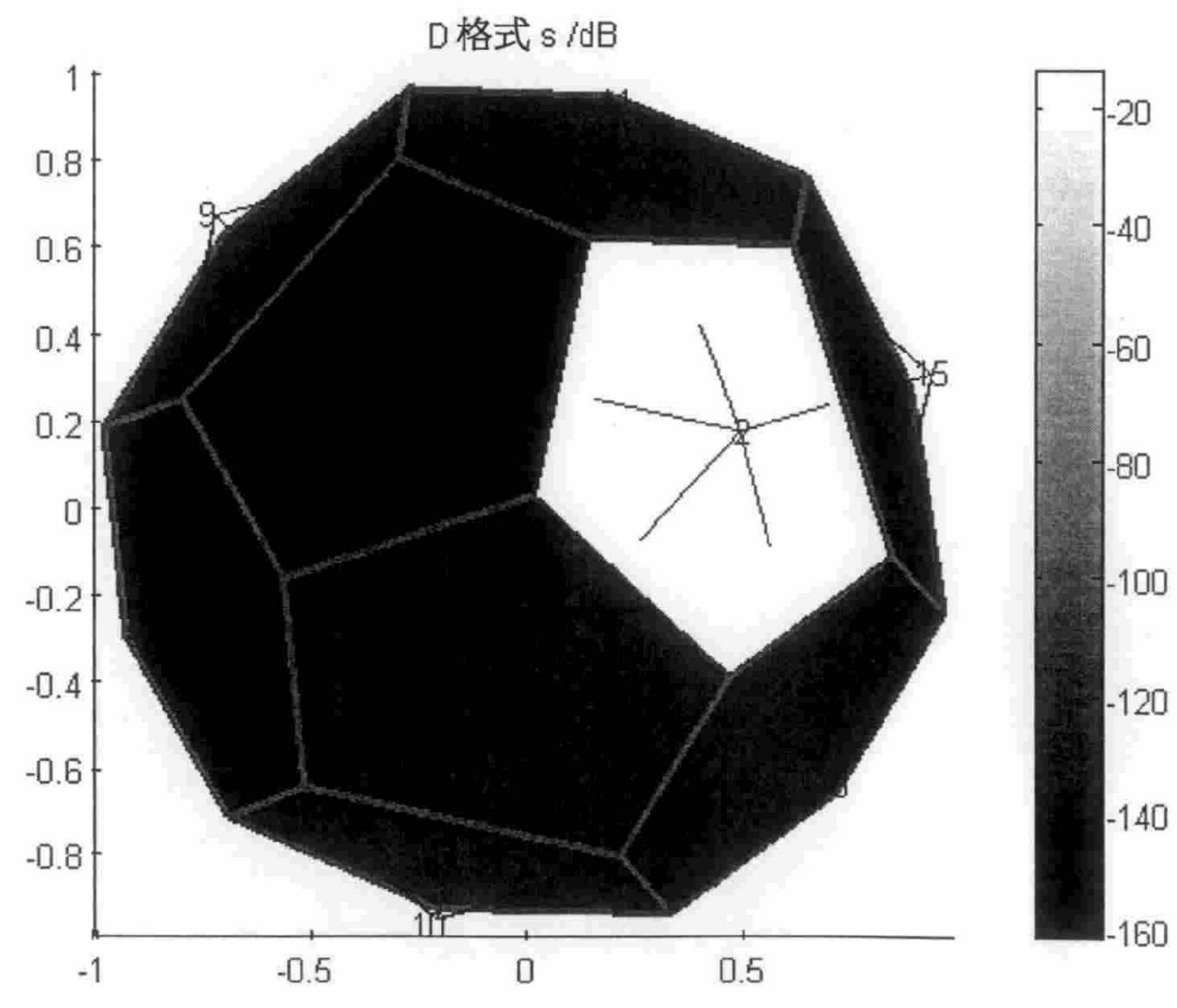


圖 3



a)



b)

圖 4

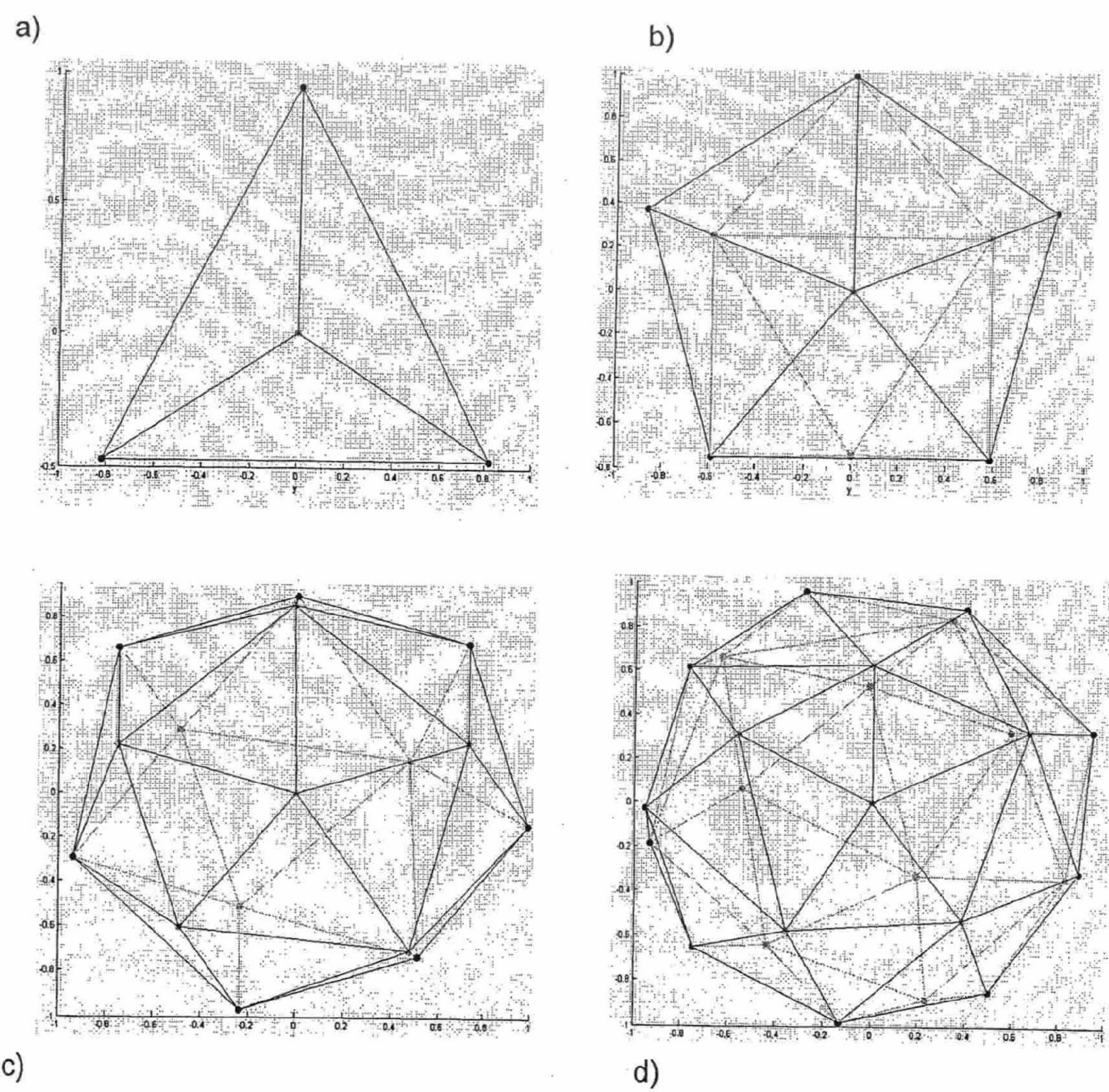


圖 5

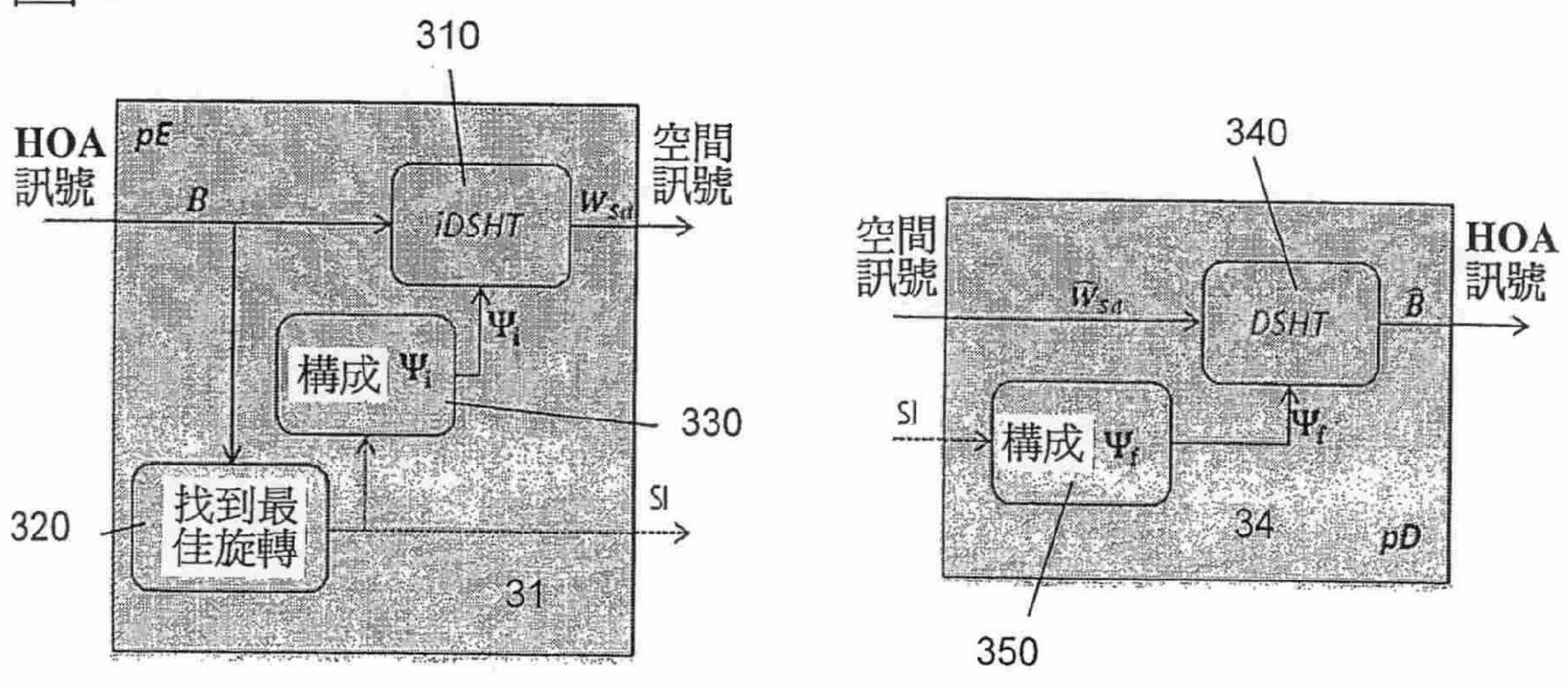


圖 6

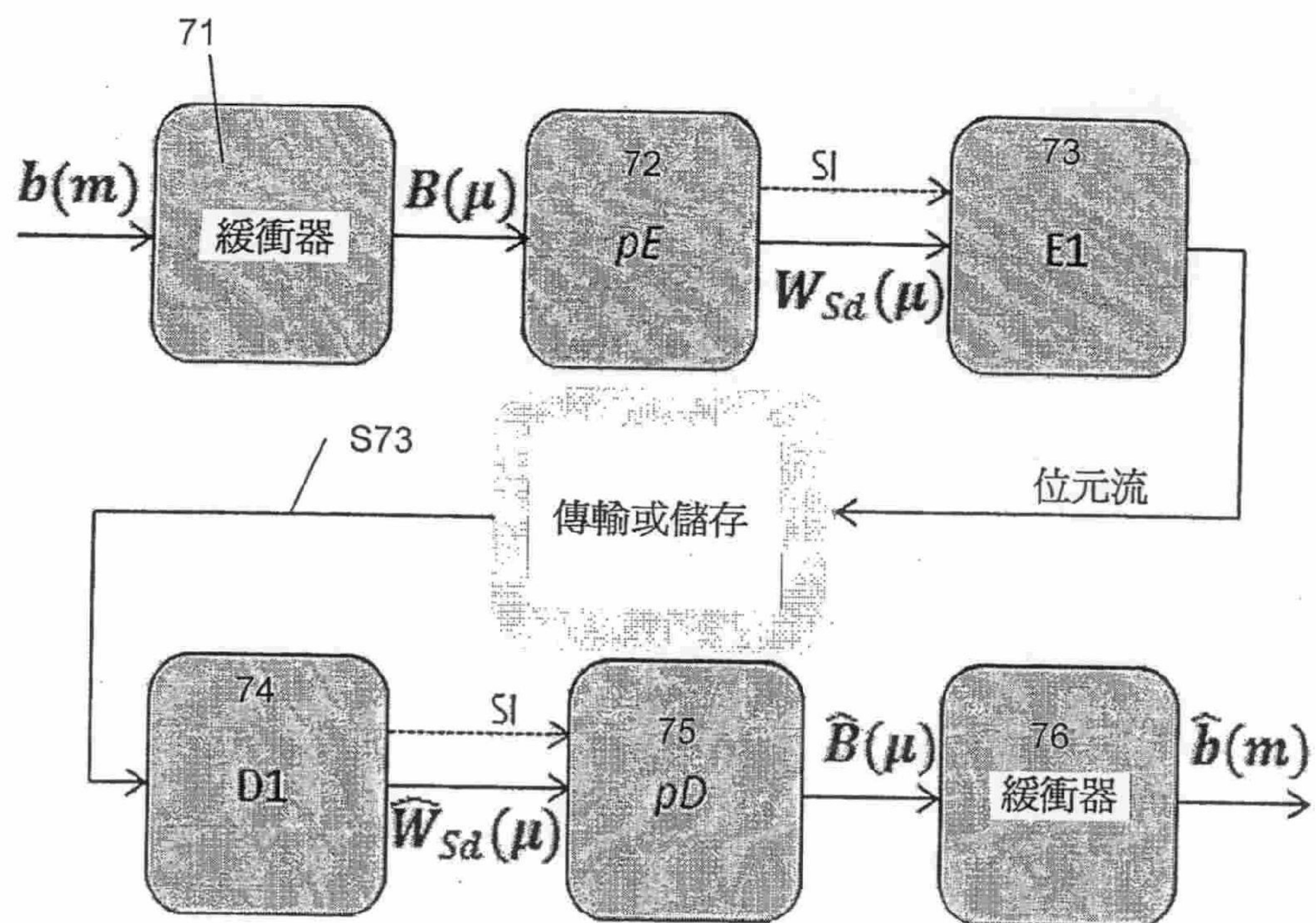


圖 7

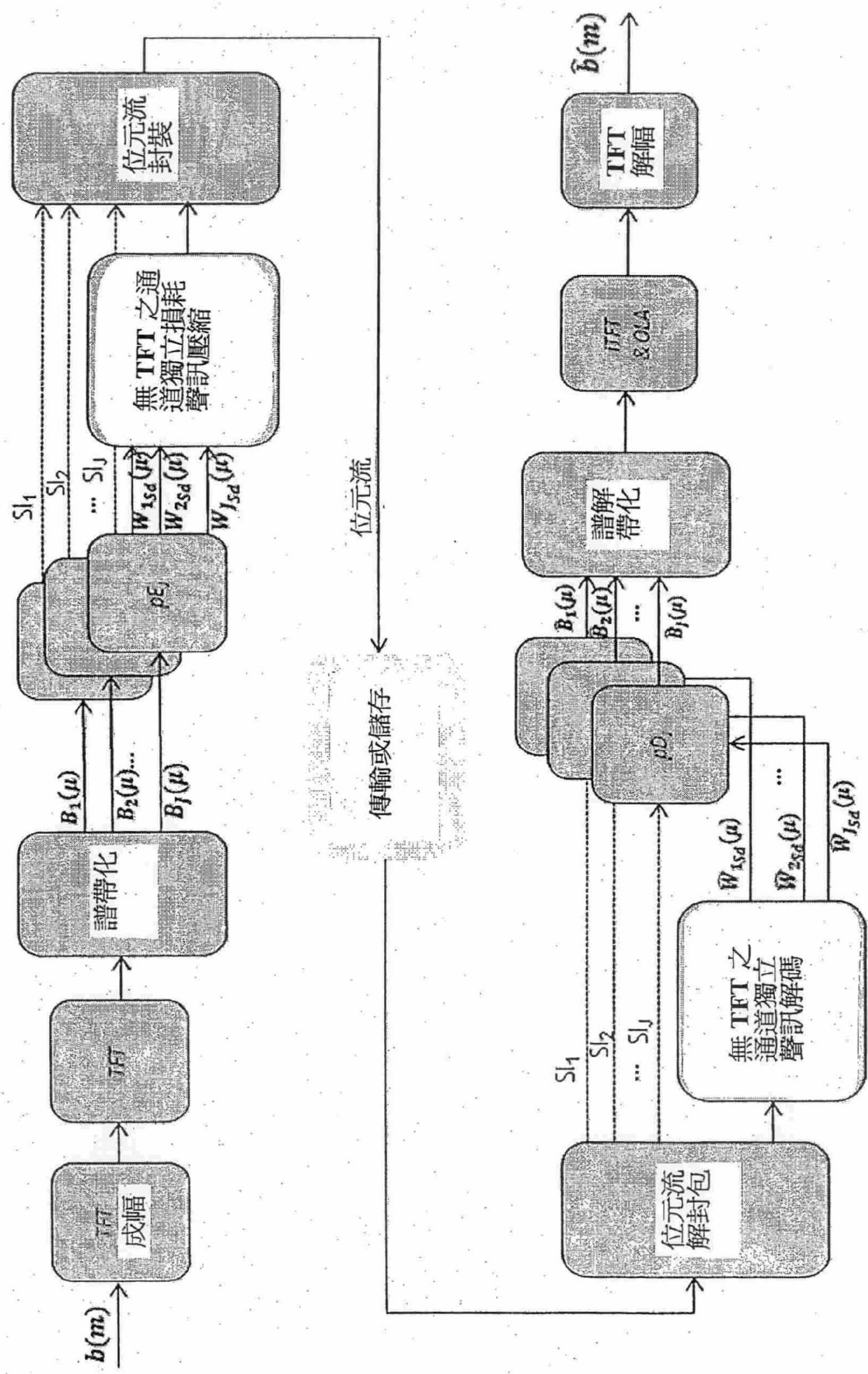


圖 8