



(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 I569260 B

(45)公告日：中華民國 106 (2017) 年 02 月 01 日

(21)申請案號：103141222

(22)申請日：中華民國 103 (2014) 年 11 月 27 日

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

G10L19/02 (2013.01)

(30)優先權：2013/11/27 歐洲專利局

13194664.2

(71)申請人：弗勞恩霍夫爾協會(德國) FRAUNHOFER-GESELLSCHAFT ZUR FORDERUNG
DER ANGEWANDTEN FORSCHUNG E. V. (DE)

德國

(72)發明人：帕露斯 喬尼 PAULUS, JOUNI (FI)；迪斯曲 薩斯洽 DISCH, SASCHA (DE)；福
契斯 哈拉德 FUCHS, HARALD (DE)；吉瑞爾 柏哈德 GRILL, BERNHARD
(DE)；賀穆斯 奧利薇 HELLMUTH, OLIVER (DE)；摩塔札 愛德瑞恩
MURTAZA, ADRIAN (RO)；萊德布奇 法科 RIDDERBUSCH, FALKO (DE)；泰
倫堤夫 黎恩 TERENCE, LEON (DE)

(74)代理人：惲軼群；陳文郎

(56)參考文獻：

WO 2009/131391A1

WO 2012/125855A1

WO 2012/146757A1

審查人員：涂淑惠

申請專利範圍項數：15 項 圖式數：12 共 73 頁

(54)名稱

用於在基於物件之音訊編碼系統中利用旁通音訊物件信號的通知響度估計之解碼器、編碼器及方法
DECODER, ENCODER AND METHOD FOR INFORMED LOUDNESS ESTIMATION EMPLOYING
BY-PASS AUDIO OBJECT SIGNALS IN OBJECT-BASED AUDIO CODING SYSTEMS

(57)摘要

本案提供一種用於產生包括一或多個音訊輸出聲道之一音訊輸出信號的解碼器。該解碼器包括一接收介面，該接收介面用於接收包括多個音訊物件信號之一音訊輸入信號、用於接收關於該等音訊物件信號之響度資訊，且用於接收指示該等音訊物件信號中的一或多者是否應放大或衰減的顯現資訊。此外，該解碼器包括用於產生該音訊輸出信號之該一或多個音訊輸出聲道的一信號處理器。該信號處理器經組配以取決於該響度資訊且取決於該顯現資訊而判定一響度補償值。此外，該信號處理器經組配以取決於該顯現資訊且取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道。一或多個旁通音訊物件信號係用於產生該音訊輸出信號。此外，提供一種編碼器。

A decoder for generating an audio output signal comprising one or more audio output channels is provided. The decoder comprises a receiving interface (110) for receiving an audio input signal comprising a plurality of audio object signals, for receiving loudness information on the audio object signals, and for receiving rendering information indicating whether one or more of the audio object signals shall be amplified or attenuated. Moreover, the decoder comprises a signal processor (120) for generating the one or more audio output channels of the audio output signal. The signal processor (120) is configured to determine a loudness compensation value depending on the loudness information and depending on the rendering

指定代表圖：

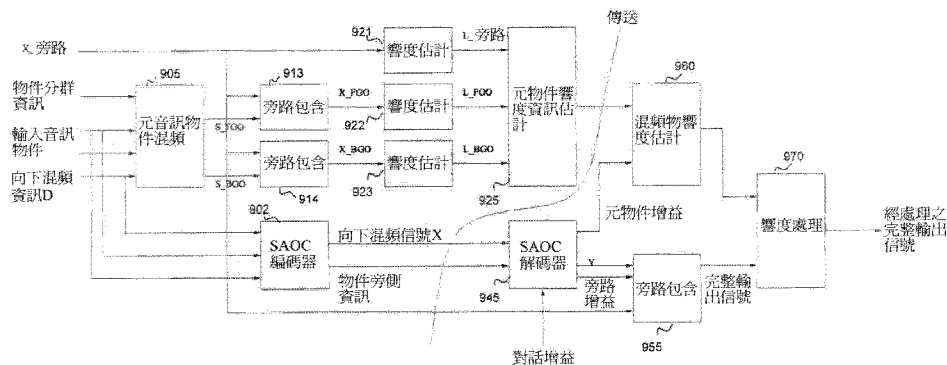


圖 9

符號簡單說明：

902 · · · SAOC 編碼
器

905 · · · 元音訊物件
混頻器

913 . . . 「旁路包
含」區塊

914 . . . 「旁路包
含」區塊

921 · · · 響度估計單
元

922 · · · 響度估計單元

923 · · · 響度估計單
元

925 · · · 元物件響度
資訊估計器

945 · · · SAOC 解碼
器

955 . . . 「旁路包
含」單元

960 · · · 混頻物響度
估計器

970 · · · 響度處理單元

發明摘要

公告本

※ 申請案號：103/14/222✓

※ 申請日：

103 11.27

※ IPC 分類：G10L 19/008 (2013.01)

19/02 (2013.01)

【發明名稱】(中文/英文)

用於在基於物件之音訊編碼系統中利用旁通音訊物件信號的通知響度估計之解碼器、編碼器及方法

DECODER, ENCODER AND METHOD FOR INFORMED LOUDNESS ESTIMATION EMPLOYING BY-PASS AUDIO OBJECT SIGNALS IN OBJECT-BASED AUDIO CODING SYSTEMS

【中文】

本案提供一種用於產生包括一或多個音訊輸出聲道之一音訊輸出信號的解碼器。該解碼器包括一接收介面，該接收介面用於接收包括多個音訊物件信號之一音訊輸入信號、用於接收關於該等音訊物件信號之響度資訊，且用於接收指示該等音訊物件信號中的一或多者是否應放大或衰減的顯現資訊。此外，該解碼器包括用於產生該音訊輸出信號之該一或多個音訊輸出聲道的一信號處理器。該信號處理器經組配以取決於該響度資訊且取決於該顯現資訊而判定一響度補償值。此外，該信號處理器經組配以取決於該顯現資訊且取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道。一或多個旁通音訊物件信號係用於產生該音訊輸出信號。此外，提供一種編碼器。

【英文】

A decoder for generating an audio output signal comprising one or more audio output channels is provided. The decoder comprises a receiving interface (110) for receiving an audio input signal comprising a plurality of audio object signals, for receiving loudness information on the audio object signals, and for receiving rendering information indicating whether one or more of the audio object signals shall be amplified or attenuated. Moreover, the decoder comprises a signal processor (120) for generating the one or more audio output channels of the audio output signal. The signal processor (120) is configured to determine a loudness compensation value depending on the loudness information and depending on the rendering information. Furthermore, the signal processor (120) is configured to generate the one or more audio output channels of the audio output signal from the audio input signal depending on the rendering information and depending on the loudness compensation value. One or more by-pass audio object signals are employed for generating the audio output signal. Moreover, an encoder is provided.

【代表圖】

【本案指定代表圖】：第（9）圖。

【本代表圖之符號簡單說明】：

902...SAOC編碼器

905...元音訊物件混頻器

913...「旁路包含」區塊

914...「旁路包含」區塊

921...響度估計單元

922...響度估計單元

923...響度估計單元

925...元物件響度資訊估計器

945...SAOC解碼器

955...「旁路包含」單元

960...混頻物響度估計器

970...響度處理單元

【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)

用於在基於物件之音訊編碼系統中利用旁通音訊物件信號的通知響度估計之解碼器、編碼器及方法

DECODER, ENCODER AND METHOD FOR INFORMED LOUDNESS ESTIMATION EMPLOYING BY-PASS AUDIO OBJECT SIGNALS IN OBJECT-BASED AUDIO CODING SYSTEMS

【技術領域】

[0001]本發明係關於音訊信號編碼、處理及解碼，且詳言之，係關於用於基於物件之音訊編碼系統中的通知響度估計之解碼器、編碼器及方法。

【先前技術】

[0002]最近，在音訊編碼[BCC、JSC、SAOC、SAOC1、SAOC2]及通知源分離[ISS1、ISS2、ISS3、ISS4、ISS5、ISS6]之領域中已提議用於包括多個音訊物件信號的音訊場景之具位元率效益的傳輸/儲存之參數技術。此等技術旨在基於描述所傳輸/儲存之音訊場景及/或該音訊場景中的源物件之額外旁側資訊來重建構所要輸出音訊場景或音訊源物件。此重建構使用通知源分離方案而發生於解碼器中。可組合重建構之物件以產生輸出音訊場景。取決於組合物件之方式，輸出場景之知覺響度可能改變。

[0003]在TV及無線電廣播中，各個節目之音軌的音量位準可基於諸如峰值信號位準或響度位準之各種態樣而加以

正規化。取決於信號之動態特性，具有相同峰值位準之兩個信號可能具有大不相同的感覺響度位準。現在在節目或頻道之間切換，信號響度的差異非常惱人，且已成為終端使用者對廣播的抱怨的主要來源。

[0004]在先前技術中，已提議使用基於知覺信號響度之測量將所有頻道上的所有節目類似地正規化至共同參考位準。在歐洲的一個此種推薦為EBU推薦R128 [EBU](後文稱為R128)。

[0005]該推薦稱「節目響度」，例如一個節目(或一個商業或某一其他有意義的節目實體)的平均響度，應等於指定位準(允許有小的偏差)。當愈來愈多的廣播台符合此推薦及所要求的正規化時，應最小化節目及頻道之間的平均響度差異。

[0006]可以若干方式進行響度估計存在用於估計音訊信號之知覺響度的若干數學模型。EBU推薦R128依賴於ITU-R BS.1770(後文稱為BS.1770)(見[ITU])中呈現的模型來進行響度估計。

[0007]如前所述，例如根據EBU推薦R128，節目響度(例如一個節目的平均響度)應等於指定位準，允許有小的偏差。然而，此導致進行音訊顯現時的顯著問題，迄今在先前技術中仍未解決。在解碼器側進行音訊顯現對所接收音訊輸入信號之總體/總響度具有顯著影響。然而，儘管進行場景顯現，但所接收音訊信號之總響度應保持相同。

[0008]當前，對於此問題不存在特定解碼器側解決方

案。

[0009] EP 2 146 522 A1 ([EP]) 係關於用於使用基於物件之元資料產生音訊輸出信號之概念。產生表示至少兩個不同音訊物件信號之重疊的至少一個音訊輸出信號，但其不提供對此問題的解決方案。

[0010] WO 2008/035275 A2 ([BRE]) 描述一種包括編碼器之音訊系統，該編碼器在編碼單元中編碼音訊物件，該編碼單元產生向下混頻音訊信號及表示多個音訊物件之參數資料。該向下混頻音訊信號及參數資料傳輸至包括解碼單元及顯現單元之解碼器，該解碼單元產生音訊物件之近似複本且該顯現單元自該等音訊物件產生輸出信號。此外，該解碼器含有用於產生發送至編碼器之編碼修改資料的處理器。該編碼器接著回應於該編碼修改資料而修改該等音訊物件之編碼，且尤其修改該參數資料。該方法允許由解碼器控制但完全或部分由編碼器進行對音訊物件之操縱。因此，可對實際上獨立的音訊物件而非對近似複本進行操縱，藉此提供改良的效能。

[0011] EP 2 146 522 A1 ([SCH]) 揭示一種用於產生表示至少兩個不同音訊物件之重疊的至少一個音訊輸出信號之設備，其包括用於處理音訊輸入信號以提供音訊輸入信號之物件表示的處理器，其中此物件表示可藉由原始物件之參數導引近似而使用物件向下混頻信號產生。物件操縱器使用關於個別音訊物件之基於音訊物件的元資料個別地操縱物件，以獲得經操縱之音訊物件。經操縱之音訊物件使用物件混

頻器混頻以取決於特定顯現設定而最終獲得具有一個若干個聲道信號的音訊輸出信號。

[0012] WO 2008/046531 A1([ENG])描述一種用於使用多個音訊物件產生經編碼物件信號之音訊物件寫碼器，其包含：向下混頻資訊產生器，其用於產生指示多個音訊物件至至少兩個向下混頻聲道之分配的向下混頻資訊；音訊物件參數產生器，其用於產生音訊物件之物件參數；以及輸出介面，其用於使用該向下混頻資訊及該等物件參數產生匯入之音訊輸出信號。音訊合成器使用向下混頻資訊用於產生可用於建立預定義音訊輸出組態之多個輸出聲道的輸出資料。

[0013] 將需要無延遲地具有輸出平均響度或平均響度改變的準確估計，且當節目不改變或顯現場景不改變時，平均響度估計亦應保持靜態。

【發明內容】

[0014] 本發明之目標係提供改良之音訊信號編碼、處理及解碼概念。本發明之目標藉由以下各者來解決：如請求項1之解碼器、如請求項9之編碼器、如請求項11之系統、如請求項12之方法、如請求項13之方法及如請求項15之電腦程式。

[0015] 提供用於估計基於物件之音訊編碼系統中的輸出之響度的通知方式。所提供之概念依賴於關於待提供至解碼器之音訊混頻物中的物件之響度的資訊。解碼器使用此資訊連同顯現資訊用於估計輸出信號之響度。舉例而

言，此允許接著估計預設向下混頻與顯現之輸出之間的響度差異。接著有可能補償該差異以無關於顯現資訊而獲得輸出中的大致恆定的響度。解碼器中的響度估計以全參數方式發生，且與基於信號之響度估計概念相比，其在計算上非常輕便且準確。

[0016]提供用於使用純參數概念獲得關於特定輸出場景之響度的資訊之概念，其接著允許在無解碼器中的明確的基於信號之響度估計的情況下進行響度處理。此外，描述藉由MPEG [SAOC]標準化的空間音訊物件寫碼(SAOC)之特定技術，但所提供之概念亦可結合其他音訊物件寫碼技術來使用。

[0017]提供用於產生包括一或多個音訊輸出聲道之音訊輸出信號的解碼器。該解碼器包括接收介面，該接收介面用於接收包括多個音訊物件信號之音訊輸入信號、用於接收關於該等音訊物件信號之響度資訊，且用於接收指示該等音訊物件信號中的一或多者是否應放大或衰減的顯現資訊。此外，該解碼器包括用於產生該音訊輸出信號之一或多個音訊輸出聲道的信號處理器。該信號處理器經組配以取決於該響度資訊，且取決於該顯現資訊而判定響度補償值。此外，該信號處理器經組配以取決於該顯現資訊且取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道。

[0018]根據一實施例，該信號處理器可經組配以取決於該顯現資訊且取決於該響度補償值而自該音訊輸入信號產

生該音訊輸出信號之該一或多個音訊輸出聲道，使得該音訊輸出信號之響度等於該音訊輸入信號之響度，或使得該音訊輸出信號之響度比經修改音訊信號之響度更接近於該音訊輸入信號之響度，該經修改音訊信號將自藉由根據該顯現資訊放大或衰減該音訊輸入信號之該等音訊物件信號而修改該音訊輸入信號來產生。

[0019]根據另一實施例，可將音訊輸入信號之音訊物件信號中之每一者指派至兩個或兩個以上群組中的僅一個群組，其中該兩個或兩個以上群組中之每一者可包括該音訊輸入信號之該等音訊物件信號中的一或多者。在此實施例中，該接收介面可經組配以接收該兩個或兩個以上群組中之每一群組的響度值作為響度資訊，其中該響度值指示該群組之該一或多個音訊物件信號之原始總響度。此外，該接收介面可經組配以接收針對該兩個或兩個以上群組中之至少一個群組藉由指示該群組之一或多個音訊物件信號之經修改總響度而指示是否應放大或衰減該群組之該一或多個音訊物件信號之顯現資訊。此外，在此實施例中，該信號處理器可經組配以取決於該兩個或兩個以上群組中之該至少一個群組中之每一者的經修改總響度，且取決於該兩個或兩個以上群組中之每一者的原始總響度而判定該響度補償值。此外，該信號處理器可經組配以取決於該兩個或兩個以上群組中之該至少一個群組中之每一者的經修改總響度，且取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之一或多個音訊輸出聲道。

[0020]在特定實施例中，該兩個或兩個以上群組中之至少一個群組可包括該等音訊物件信號中的兩者或兩者以上。

[0021]此外，提供一種編碼器。該編碼器包括基於物件之編碼單元，該基於物件之編碼單元用於編碼多個音訊物件信號以獲得包括該多個音訊物件信號之經編碼音訊信號。此外，該編碼器包括用於編碼關於該等音訊物件信號之響度資訊的物件響度編碼單元。該響度資訊包括一或多個響度值，其中該一或多個響度值中之每一者取決於該等音訊物件信號中的一或多者。

[0022]根據一實施例，可將經編碼音訊信號之音訊物件信號中之每一者指派至兩個或兩個以上群組中的僅一個群組，其中該兩個或兩個以上群組中之每一者包括該經編碼音訊信號之該等音訊物件信號中的一或多者。該物件響度編碼單元可經組配以藉由判定該兩個或兩個以上群組中之每一群組的響度值而判定該響度資訊之該一或多個響度值，其中該群組之該響度值指示該群組之該一或多個音訊物件信號之原始總響度。

[0023]此外，提供一種系統。該系統包括根據上述實施例中的一者之編碼器，該編碼器用於編碼多個音訊物件信號以獲得包括該多個音訊物件信號之經編碼音訊信號且用於編碼關於該等音訊物件信號之響度資訊。此外，該系統包括根據上述實施例中的一者之解碼器，該解碼器用於產生包括一或多個音訊輸出聲道之音訊輸出信號。該解碼器

經組配以接收作為音訊輸入信號之經編碼音訊信號及響度資訊。此外，該解碼器經組配以進一步接收顯現資訊。此外，該解碼器經組配以取決於該響度資訊且取決於該顯現資訊而判定響度補償值。此外，該解碼器經組配以取決於該顯現資訊且取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道。

[0024] 此外，提供一種用於產生包括一或多個音訊輸出聲道之音訊輸出信號的方法。該方法包括：

- 接收包括多個音訊物件信號之音訊輸入信號。
- 接收關於該等音訊物件信號之響度資訊。
- 接收指示該等音訊物件信號中的一或多者是否應放大或衰減之顯現資訊。
- 取決於該響度資訊且取決於該顯現資訊而判定響度補償值。及：
- 取決於該顯現資訊且取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道。

[0025] 此外，提供一種用於編碼之方法。該方法包括：

- 編碼包括多個音訊物件信號之音訊輸入信號。及：
- 編碼關於該等音訊物件信號之響度資訊，其中該響度資訊包括一或多個響度值，其中該一或多個響度值中之每一者取決於該等音訊物件信號中的一或多者。

[0026] 此外，提供一種用於在電腦或信號處理器上執行時實施上述方法的電腦程式。

[0027] 在附屬請求項中提供較佳實施例。

【圖式簡單說明】

[0028] 下文中，參考諸圖更詳細地描述本發明之實施例，其中：

圖 1 例示根據一實施例之用於產生包括一或多個音訊輸出聲道之音訊輸出信號的解碼器，

圖 2 例示根據一實施例之編碼器，

圖 3 例示根據一實施例之系統，

圖 4 例示包括 SAOC 編碼器及 SAOC 解碼器之空間音訊物件寫碼系統，

圖 5 例示包括旁側資訊解碼器、物件分隔器及顯現器之 SAOC 解碼器，

圖 6 例示響度改變時之輸出信號響度估計行為，

圖 7 描繪根據一實施例之通知響度估計，其例示根據一實施例之編碼器及解碼器的組件，

圖 8 例示根據另一實施例之編碼器，

圖 9 例示根據一實施例之關於包括旁路聲道之 SAOC 對話增強的編碼器及解碼器，

圖 10 描繪使用用於以參數方式估計響度改變之所提供概念來量測響度改變及其結果的第一圖解，

圖 11 描繪使用用於以參數方式估計響度改變之所提供概念來量測響度改變及其結果的第二圖解，以及

圖 12 例示用於進行響度補償之另一實施例。

【實施方式】

[0029] 在詳細描述較佳實施例之前，描述響度估計、空間音訊物件寫碼(SAOC)及對話增強(DE)。

[0030] 首先，描述響度估計。

[0031] 如之前所述，EBU recommendation R128依賴於ITU-R BS.1770中呈現的模型來進行響度估計。此量測將用作實例，但下文所描述之概念亦可應用於其他響度量測。

[0032] 根據BS.1770之響度估計操作相對簡單，且其係基於以下主要步驟[ITU]：

- 用 K 過濾器(傾斜型濾波器與高通濾波器之組合)對輸入信號 x_i (或多聲道信號情況下之信號)進行濾波以獲得信號 y_i 。

- 計算信號 y_i 之均方能量 z_i 。

- 在多聲道信號之情況下，應用聲道加權 G_i ，且對經加權信號進行求和。接著將信號響度定義為

$$L = c + 10 \log_{10} \sum_i G_i z_i ,$$

其中常數值 $c = -0.691$ 。接著將輸出表達為「LKFS」之單位(響度，經 K 加權，相對於全標度)，其標度類似於分貝標度。

[0033] 在上式中， G_i 對於聲道中之一些可例如等於1，而 G_i 對於一些其他聲道可例如為1.41。舉例而言，若考慮左聲道、右聲道、中央聲道、左環繞聲道及右環繞聲道，則各別權重 G_i 對於左聲道、右聲道及中央聲道可例如為1，且對於左環繞聲道及右環繞聲道可例如為1.41，見[ITU]。

[0034] 可以看出，響度值 L 與信號能量之對數緊密相

關。

[0035] 下文中，描述空間音訊物件寫碼。

[0036] 基於物件之音訊寫碼概念允許在鏈之解碼器側的極大靈活性。基於物件之音訊寫碼觀念之一實例為空間音訊物件寫碼(SAOC)。

[0037] 圖4例示包括空間音訊物件寫碼(SAOC)編碼器410及SAOC解碼器420之SAOC系統。

[0038] SAOC編碼器410接收 N 個音訊物件信號 $S_1, \dots, S_N$ 作為輸入。此外，SAOC編碼器410進一步接收指令「混頻資訊 $\mathbf{D}$ 」：應如何組合此等物件以獲得包括 M 個向下混頻聲道 $X_1, \dots, X_M$ 之向下混頻信號。SAOC編碼器410自物件及向下混頻過程提取一些旁側資訊，且傳輸此旁側資訊及/或將其與向下混頻信號一起儲存。

[0039] SAOC系統之主要特性為包括向下混頻聲道 $X_1, \dots, X_M$ 之向下混頻信號 X 形成語義上有意義的信號。換言之，有可能聽取向向下混頻信號。舉例而言，若接收器不具有SAOC解碼器功能性，則接收器可仍然始終將向下混頻信號提供為輸出。

[0040] 圖5例示包括旁側資訊解碼器510、物件分隔器520及顯現器530之SAOC解碼器。圖5例示之SAOC解碼器例如自SAOC編碼器接收向下混頻信號及旁側資訊。向下混頻信號可視為包括音訊物件信號之音訊輸入信號，此係因為音訊物件信號在向下混頻信號內混頻(音訊物件信號在向下混頻信號之一或多個向下混頻聲道內混頻)。

[0041] SAOC解碼器可例如接著試圖(虛擬地)重建構原始物件，例如藉由使用物件分離器520，例如使用經解碼旁側資訊。此等(虛擬)物件重建構 $\hat{s}_1, \dots, \hat{s}_N$ (例如，重建構之音訊物件信號)接著基於顯現資訊(例如顯現矩陣 $\mathbf{R}$)加以組合以產生音訊輸出信號 $\mathbf{Y}$ 之 K 個音訊輸出聲道 $Y_1, \dots, Y_K$ 。

[0042]通常，在SAOC中，舉例而言，音訊物件信號例如藉由使用自SAOC編碼器傳輸至SAOC解碼器之協方差資訊(例如信號協方差矩陣 $\mathbf{E}$)而加以重建構。

[0043]舉例而言，可使用下式來在解碼器側重建構音訊物件信號：

$$\mathbf{S} = \mathbf{G}\mathbf{X} \quad \text{其中 } \mathbf{G} \approx \mathbf{E}\mathbf{D}^H(\mathbf{D}\mathbf{E}\mathbf{D}^H)^{-1}$$

其中

N 音訊物件信號之數目

$N_{\text{樣本}}$ 音訊物件信號之所考慮樣本的數目

M 向下混頻聲道之數目，

$\mathbf{X}$ 向下混頻音訊信號，大小為 $M \times N_{\text{樣本}}$ ，

$\mathbf{D}$ 向下混頻矩陣，大小為 $M \times N$

$\mathbf{E}$ 定義為 $\mathbf{E} = \mathbf{X}\mathbf{X}^H$ 之信號協方差矩陣，大小為 $N \times N$

$\mathbf{S}$ 以參數重建構之 N 個音訊物件信號，大小為 $N \times N_{\text{樣本}}$

$(\cdot)^H$ 自伴(赫密特)算子，其表示 $(\cdot)$ 之共軛轉置

[0044]接著，可將顯現矩陣 $\mathbf{R}$ 應用於經重建構之音訊物件信號 $\mathbf{S}$ 上以獲得音訊輸出信號 $\mathbf{Y}$ 之音訊輸出聲道，例如根

據下式：

$$\mathbf{Y} = \mathbf{R}\mathbf{S}$$

其中

[0045]

K 音訊輸出信號 $\mathbf{Y}$ 之音訊輸出聲道 $Y_1, \dots, Y_K$ 之數目。

$\mathbf{R}$ 大小為 $K \times N$ 之顯現矩陣

$\mathbf{Y}$ 包括 K 個音訊輸出聲道之音訊輸出信號，大小為 $K \times N_{\text{樣本}}$

[0046] 在圖5中，例如由物件分離器520進行物件重建構之過程加上了表述「虛擬」或「任擇」，因為其可能不必發生，但所要功能性可藉由在參數域中組合重建構與顯現步驟(亦即，組合方程式)而獲得。

[0047] 換言之，替代首先使用混頻資訊 $\mathbf{D}$ 及協方差資訊 $\mathbf{E}$ 重建構音訊物件信號且接著將顯現資訊 $\mathbf{R}$ 應用於經重建構之音訊物件信號以獲得音訊輸出聲道 $Y_1, \dots, Y_K$ ，可將兩個步驟在單一步驟中進行，以使得直接自向下混頻聲道出生音訊輸出聲道 $Y_1, \dots, Y_K$ 。

[0048] 舉例而言，可使用下式：

$$[0049] \quad \mathbf{Y} = \mathbf{R}\mathbf{G}\mathbf{X} \quad \text{其中 } \mathbf{G} \approx \mathbf{E} \mathbf{D}^H (\mathbf{D} \mathbf{E} \mathbf{D}^H)^{-1}。$$

[0050] 原理上，顯現資訊 $\mathbf{R}$ 可請求原始音訊物件信號之任何組合。然而，在實務中，物件重建構可包括重建構錯誤，且所請求之輸出場景可能未必能達到。作為涵蓋許多實際情況之大致一般規則，所請求之輸出場景與向下混頻

信號差異愈大，可聞之重建構錯誤將愈大。

[0051] 下文中，描述對話增強(DE)。可例如使用SAOC技術來實現場景。應注意，即使名稱「對話增強」提示集中於以對話為導向的信號，相同原理亦可用於其他信號類型。

[0052] 在DE場景中，系統中的自由度自一般情況受限。

[0053] 舉例而言，音訊物件信號 $s_1, \dots, s_N = \mathbf{S}$ 分群(且可能混頻)為兩個元物件：前景物件(FGO) $\mathbf{s}_{FGO}$ 及背景物件(BGO) $\mathbf{s}_{BGO}$ 。

[0054] 此外，輸出場景 $Y_1, \dots, Y_K = \mathbf{Y}$ 類似於向下混頻信號 $X_1, \dots, X_M = \mathbf{X}$ 。更特定言之，兩個信號具有相同維度，亦即 $K = M$ ，且最終使用者可僅控制兩個元物件FGO及BGO之相對混頻位準。更確切而言，藉由用某一純量權重來對FGO與BGO進行混頻而獲得向下混頻信號

$$\mathbf{X} = h_{FGO} \mathbf{S}_{FGO} + h_{BGO} \mathbf{S}_{BGO},$$

且類似地藉由對 FGO 與 BGO 之純量加權而獲得輸出場景：

$$[0055] \quad \mathbf{Y} = g_{FGO} \mathbf{S}_{FGO} + g_{BGO} \mathbf{S}_{BGO}.$$

[0056] 取決於混頻權重之相對值，FGO與BGO之間的平衡可能改變。舉例而言，對於設定

$$\begin{cases} g_{FGO} > h_{FGO} \\ g_{BGO} = h_{BGO} \end{cases}$$

[0057] 有可能在混頻物中增大FGO之相對含量。若FGO為對話，則此設定提供對話增強功能性。

[0058] 作為使用情況實例，BGO可為體育場噪音及體育

賽事期間之其他背景聲音，且FGO為解說員的語音。DE功能性允許最終使用者相對於背景放大或衰減解說員之位準。

[0059]實施例係基於以下發現：在廣播場景中利用SAOC技術(或類似技術)允許向最終使用者提供擴展之信號操縱功能性。提供比僅改變聲道及調整播放音量更多之功能性。

[0060]上文簡要描述使用DE技術之一個可能性。若為用於SAOC之向下混頻信號的廣播信號例如根據R128而在位準上正規化，則在不施加(SAOC)處理(或顯現描述與向下混頻描述相同)時，不同節目具有類似的平均響度。然而，當施加某一(SAOC)處理時，輸出信號不同於預設向下混頻信號，且輸出信號之響度可不同於預設向下混頻信號之響度。自最終使用者之觀點觀之，此可導致以下情境：聲道或節目之間的輸出信號響度可再次具有不合需要的跳躍或差異。換言之，藉由廣播台施加之正規化之益處部分地失去。

[0061]此問題並非SAOC或DE場景所獨有，而在允許最終使用者與內容互動之其他音訊寫碼概念中亦可能發生。然而，在許多情況下，若輸出信號具有與預設向下混頻不同之響度，其並不造成任何傷害。

[0062]如上所述，音訊輸入信號節目之總響度應等於指定位準，允許有小的偏差。然而，如已經概述的，此在進行音訊顯現時導致顯著問題，因為顯現對所接收音訊輸入

信號之總體/總響度可具有顯著影響。然而，儘管進行場景顯現，但所接收音訊信號之總響度應保持相同。

[0063]一種方法將為在播放時估計信號之響度，且藉由適當的時間整合概念，估計在某一時間之後可能會聚至真實的平均響度。然而，會聚所需之時間自最終使用者觀點觀之係有問題的。當響度估計改變時，即使當無改變施加至信號時，響度改變補償亦應起作用且改變其行為。此將導致輸出信號具有時變平均響度，其在感覺上相當惱人。

[0064]圖6例示響度改變時的輸出信號響度估計行為。尤其描繪基於信號之輸出信號響度估計，其例示剛剛描述之解決方案之效果。估計相當緩慢地接近正確估計。替代基於信號之輸出信號響度估計，即刻正確地判定輸出信號響度之通知輸出信號響度估計將為較佳的。

[0065]詳言之，在圖6中，使用者輸入(例如，對話物件之位準)藉由增大值而在時刻 T 改變。真實輸出信號位準及(相應地)響度在相同時刻改變。當利用某一時間整合時間來自輸出信號進行輸出信號響度估計，估計將逐漸改變，且在某一延遲之後達到正確值。在此延遲期間，估計值不斷改變，且不能可靠地用於進一步處理輸出信號，例如，用於響度位準校正。

[0066]如已經描述的，將需要無延遲地具有輸出平均響度或平均響度改變之準確估計，且當節目不改變或顯現場景不改變時，平均響度估計亦應保持靜態。換言之，當應用某一響度改變補償時，補償參數應僅在節目改變或存在

某一使用者互動時才改變。

[0067]所要行為例示於圖6之最下圖解(通知輸出信號響度估計)中。輸出信號響度估計應在使用者輸入改變時即刻改變。

[0068]圖2例示根據一實施例之編碼器。

[0069]該編碼器包括基於物件之編碼單元210，其用於編碼多個音訊物件信號以獲得包括該多個音訊物件信號之經編碼音訊信號。

[0070]此外，該編碼器包括物件響度編碼單元220，其用於編碼關於該等音訊物件信號之響度資訊。該響度資訊包括一或多個響度值，其中該一或多個響度值中之每一者取決於該等音訊物件信號中的一或多者。

[0071]根據一實施例，該經編碼音訊信號之該等音訊物件信號中之每一者被指派至兩個或兩個以上群組中之僅一個群組，其中該兩個或兩個以上群組中之每一者包括該經編碼音訊信號之該等音訊物件信號中的一或多者。該物件響度編碼單元220經組配以藉由判定該兩個或兩個以上群組中之每一群組的響度值而判定該響度資訊之一或多個響度值，其中該群組之該響度值指示該群組之該等一或多個音訊物件信號的原始總響度。

[0072]圖1例示根據一實施例之用於產生包括一或多個音訊輸出聲道之音訊輸出信號的解碼器。

[0073]該解碼器包括接收介面110，其用於接收包括多個音訊物件信號之音訊輸入信號、用於接收關於該等音訊

物件信號之響度資訊，且用於接收指示該等音訊物件信號中的一或多者應放大或衰減的顯現資訊。

[0074]此外，該解碼器包括信號處理器120，其用於產生音訊輸出信號之一或多個音訊輸出聲道。信號處理器120經組配以取決於響度資訊且取決於顯現資訊而判定響度補償值。此外，信號處理器120經組配以取決於顯現資訊且取決於響度補償值而自音訊輸入信號產生音訊輸出信號之一或多個音訊輸出聲道。

[0075]根據一實施例，信號處理器110經組配以取決於顯現資訊且取決於響度補償值而自音訊輸入信號產生音訊輸出信號之一或多個音訊輸出聲道，使得音訊輸出信號之響度等於音訊輸入信號之響度，或使得音訊輸出信號之響度比經修改音訊信號之響度更接近於音訊輸入信號之響度，該經修改音訊信號將藉由根據顯現資訊修改或衰減音訊輸入信號之音訊物件信號而自修改該音訊輸入信號產生。

[0076]根據另一實施例，音訊輸入信號之音訊物件信號中之每一者被指派至兩個或兩個以上群組中之僅一個群組，其中該兩個或兩個以上群組中之每一者包括音訊輸入信號之音訊物件信號中的一或多者。

[0077]在此實施例中，接收介面110經組配以接收兩個或兩個以上群組中之每一群組的響度值作為響度資訊，其中該響度值指示該群組之一或多個音訊物件信號的原始總響度。此外，接收介面110經組配以接收對於兩個或兩個以

上群組中之至少一個群組藉由指示該群組之一或多個音訊物件信號的經修改總響度而指示該群組之一或多個音訊物件信號是否應放大或衰減的顯現資訊。此外，在此實施例中，信號處理器120經組配以取決於兩個或兩個以上群組中之該至少一個群組中之每一者的經修改總響度且取決於兩個或兩個以上群組中之每一者的原始總響度而判定響度補償值。此外，信號處理器120經組配以取決於該兩個或兩個以上群組中之該至少一個群組中之每一者的經修改總響度且取決於該響度補償值而自該音訊輸入信號產生音訊輸出信號之一或多個音訊輸出聲道。

[0078]在特定實施例中，兩個或兩個以上群組中之至少一個群組包括音訊物件信號中之兩者或兩者以上。

[0079]音訊物件信號 i 之能量 e_i 與音訊物件信號 i 之響度 L_i 之間根據下式存在直接關係：

$$L_i = c + 10 \log_{10} e_i, \quad e_i = 10^{(L_i - c)/10}$$

其中 c 為常數值。

[0080]實施例係基於以下發現：音訊輸入信號之不同音訊物件信號可具有不同響度且因此具有不同能量。舉例而言，若使用者想要增大音訊物件信號之響度，可相應地調整顯現資訊，且此音訊物件信號之響度增大增大此音訊物件之能量。此將導致音訊輸出信號之響度增大。為保持總響度恆定，必須進行響度補償。換言之，將必須調整自將顯現資訊應用於音訊輸入信號而產生之經修改音訊信號。然而，音訊物件信號中之一者之放大對經修改音訊信號之

總響度的確切影響取決於經放大音訊物件信號(例如響度增大之音訊物件信號)之原始響度。若此物件之原始響度對應於相當低的能量，則對音訊輸入信號之總響度的影響將微乎其微。然而，若此物件之原始響度對應於相當高的能量，則對音訊輸入信號之總響度的影響將為顯著的。

[0081]可考慮兩個實例。在兩個實例中，音訊輸入信號包括兩個音訊物件信號，且在兩個實例，藉由應用顯現資訊，音訊物件信號中之第一者的能量增大50%。

[0082]在第一實例中，第一音訊物件信號貢獻20%且第二音訊物件信號貢獻80%的音訊輸入信號總能量。然而，在第二實例中，第一音訊物件，第一音訊物件信號貢獻40%且第二音訊物件信號貢獻60%的音訊輸入信號總能量。在兩個實例中，此等貢獻可自關於音訊物件信號之響度資訊導出，因為響度與能量之間存在直接關係。

[0083]在第一實例中，第一音訊物件的能量之50%增大導致藉由將顯現資訊應用於音訊輸入信號而產生之經修改音訊信號具有為音訊輸入信號的能量之 $1.5 \times 20 \% + 80 \% = 110\%$ 的總能量。

[0084]在第二實例，第一音訊物件之能量的50%增大導致藉由將顯現資訊應用於音訊輸入信號而產生之經修改音訊信號具有為音訊輸入信號的能量之 $1.5 \times 40 \% + 60 \% = 120\%$ 的總能量。

[0085]因此，在將顯現資訊應用於音訊輸入信號之後，在第一實例中，經修改音訊信號之總能量僅須減小

9%(10/110)以在音訊輸入信號與音訊輸出信號兩者中獲得相等能量，而在第二實例中，經修改音訊信號之總能量須減小17% (20/120)。為此，可計算響度補償值。

[0086]舉例而言，響度補償值可為應用於音訊輸出信號之所有音訊輸出聲道的純量。

[0087]根據一實施例，信號處理器經組配以藉由根據顯現資訊修改或衰減音訊輸入信號之音訊物件信號而修改音訊輸入信號來產生經修改音訊信號。此外，信號處理器經組配以藉由將響度補償值應用於經修改音訊信號而產生音訊輸出信號，使得音訊輸出信號之響度等於音訊輸入信號之響度，或使得音訊輸出信號之響度比經修改音訊信號之響度更接近於音訊輸入信號之響度。

[0088]舉例而言，在上文第一實例中，響度補償值 lcv 可例如設定為值 $lcv = 10/11$ ，且可將放大因數10/11應用於根據顯現資訊顯現音訊輸入聲道而產生之所有聲道。

[0089]因此，舉例而言，在上文第二實例中，響度補償值 lcv 可例如設定為值 $lcv = 10/12 = 5/6$ ，且可將放大因數5/6應用於根據顯現資訊顯現音訊輸入聲道而產生之所有聲道。

[0090]在其他實施例中，音訊物件信號中之每一者可被指派至多個群組中之一者，且可針對該等群組中之每一者傳輸指示該群組之音訊物件信號的總響度值之響度值。如顯現資訊指定該等群組中之一者的能量衰減或放大，例如，如上所述放大50%，則可計算總能量增加，且可如上

所述判定響度補償值。

[0091]舉例而言，根據一實施例，音訊輸入信號之音訊物件信號中之每一者被指派至作為兩個或兩個以上群組的僅兩個群組中之僅一個群組。音訊輸入信號之音訊物件信號中之每一者被指派至該僅兩個群組之前景物件群組或該僅兩個群組之背景物件群組。接收介面110經組配以接收前景物件群組之一或多個音訊物件信號的原始總響度。此外，接收介面110經組配以接收背景物件群組之一或多個音訊物件信號的原始總響度。此外，接收介面110經組配以接收針對該僅兩個群組中之至少一個群組藉由指示該群組之一或多個音訊物件信號的經修改總響度而指示該至少一個群組中之每一者的一或多個音訊物件信號是否應放大或衰減的顯現資訊。

[0092]在此實施例中，信號處理器120經組配以取決於該至少一個群組中之每一者的經修改總響度、取決於前景物件群組之一或多個音訊物件信號的原始總響度且取決於背景物件群組之一或多個音訊物件信號的原始總響度而判定響度補償值。此外，信號處理器120經組配以取決於該至少一個群組中之每一者的經修改總響度且取決於該響度補償值而自音訊輸入信號產生音訊輸出信號之一或多個音訊輸出聲道。

[0093]根據一些實施例，音訊物件信號中之每一者被指派至三個或三個以上群組中之一者，且接收介面可經組配以接收三個或三個以上群組中之每一者的響度值，該響度

值指示該群組之音訊物件信號的總響度。

[0094]根據一實施例，為判定兩個或兩個以上音訊物件信號之總響度值，舉例而言，針對每一音訊物件信號判定對應於響度值之能量值，對所有響度值之能量值求和以獲得能量總和，且將對應於該能量總和之響度值判定為兩個或兩個以上音訊物件信號之總響度值。舉例而言，可使用下式

$$L_i = c + 10 \log_{10} e_i, \quad e_i = 10^{(L_i - c)/10}。$$

[0095]在一些實施例中，針對音訊物件信號中之每一者傳輸響度值，或將音訊物件信號中之每一者指派至一個或兩個或兩個以上群組，其中對於該等群組中之每一者，傳輸一響度值。

[0096]然而，在一些實施例中，對於一或多個音訊物件信號或對於包括音訊物件信號之群組中的一或多者，不傳輸響度值。替代地，解碼器可例如假定未傳輸其響度值之此等音訊物件信號或音訊物件信號群組具有預定義響度值。舉例而言，解碼器可使所有進一步判定基於此預定義響度值。

[0097]根據一實施例，接收介面110經組配以接收包括一或多個向下混頻聲道之向下混頻信號作為音訊輸入信號，其中該一或多個向下混頻聲道包括音訊物件信號，且其中該等音訊物件信號之數目小於該一或多個向下混頻聲道之數目。接收介面110經組配以接收指示音訊物件信號如何在一或多個向下混頻聲道內混頻之向下混頻資訊。此

外，信號處理器120經組配以取決於該向下混頻資訊、取決於該顯現資訊且取決於該響度補償值而自該音訊輸入信號產生音訊輸出信號之一或多個音訊輸出聲道。在特定實施例中，舉例而言，信號處理器120可經組配以取決於向下混頻資訊而計算響度補償值。

[0098]舉例而言，向下混頻資訊可為向下混頻矩陣。在諸實施例中，解碼器可為SAOC解碼器。在此等實施例中，舉例而言，接收介面110可進一步經組配以接收協方差資訊，例如，如上文所述之協方差矩陣。

[0099]關於指示音訊物件信號中的一或多者是否應放大或衰減之顯現資訊，應注意，舉例而言，指示音訊物件信號中的一或多者應如何放大或衰減之資訊為顯現資訊。舉例而言，顯現矩陣 $\mathbf{R}$ ，例如，SAOC之顯現矩陣為顯現資訊。

[0100]圖 3 例示根據一實施例之系統。

[0101]該系統包括根據上述實施例中之一者的編碼器310，其用於編碼多個音訊物件信號以獲得包括該多個音訊物件信號之經編碼音訊信號。

[0102]此外，該系統包括根據上述實施例中之一者的解碼器320，其用於產生包括一或多個音訊輸出聲道之音訊輸出信號。該解碼器經組配以接收作為音訊輸入信號之經編碼音訊信號及響度資訊。此外，解碼器320經組配以進一步接收顯現資訊。此外，解碼器320經組配以取決於響度資訊且取決於顯現資訊而判定響度補償值。此外，解碼器

320 經組配以取決於該顯現資訊且取決於該響度補償值而自音訊輸入信號產生音訊輸出信號之一或多個音訊輸出聲道。

[0103]圖 7 例示根據一實施例之通知響度估計。在傳送串流 730 之左方，例示基於物件之音訊寫碼編碼器之組件。詳言之，例示基於物件之編碼單元 710(「基於物件之音訊編碼器」)及物件響度編碼單元 720(「物件響度估計」)。

[0104]傳送串流 730 自身包括響度資訊 **L**、向下混頻資訊 **D** 及基於物件之音訊編碼器 710 之輸出 **B**。

[0105]在傳送串流 730 之右方，例示基於物件之音訊寫碼解碼器之信號處理器的組件。未例示解碼器之接收介面。描繪輸出響度估計器 740 及基於物件之音訊解碼單元 750。輸出響度估計器 740 可經組配以判定響度補償值。基於物件之音訊解碼單元 750 可經組配以藉由應用顯現資訊 **R** 而自輸入至解碼器之音訊信號判定經修改音訊信號。圖 7 中未展示將響度補償值應用於經修改音訊信號以補償由顯現引起的總響度改變。

[0106]至編碼器之輸入至少係由輸入物件 **S** 組成。系統例如藉由物件響度編碼單元 720 估計每一物件之響度(或某一其他響度相關資訊，諸如物件能量)，且傳輸及/或儲存資訊 **L**。(亦有可能將物件之響度提供為至系統之輸入，且可省略系統內之估計步驟)。

[0107]在圖 7 之實施例中，解碼器至少接收物件響度資訊及(例如)描述物件至輸出信號之混頻之顯現資訊 **R**。基於

此等資訊，例如，輸出響度估計器 740 估計輸出信號之響度且將此資訊提供為其輸出。

[0108]向下混頻資訊 **D** 可提供為顯現資訊，在此情況下，響度估計提供向下混頻信號響度之估計。亦有可能將向下混頻資訊提供為至物件響度估計之輸入，且隨物件響度資訊傳輸及/或儲存該資訊。輸出響度估計可接著同時估計向下混頻信號之響度及所顯現之輸出，且將此等兩個值或其差異提供為輸出響度資訊。差異值(或其倒數)描述應應用於所顯現輸出信號上以使得其響度。類似於向下混頻信號之響度的所需補償。物件響度資訊可另外包含關於各個物件之間的相關係數的資訊，且此相關資訊可用於輸出響度估計中以實現更準確的估計。

[0109]下文中，描述用於對話增強應用之較佳實施例。

[0110]在對話增強應用中，如上文所述，將輸入音訊物件信號分群且部分地向下混頻以形成兩個元物件 FGO 及 BGO，其可接著以普通方式求和以獲得最終向下混頻信號。

[0111]遵循 SAOC [SAOC]之描述，將 N 個輸入物件信號表示為大小為 $N \times N$ 樣本之矩陣 **S**，且將向下混頻資訊表示為大小為 $M \times N$ 之矩陣 **D**。向下混頻信號可接著獲得為 $\mathbf{X} = \mathbf{DS}$ 。

[0112]該向下混頻資訊 **D** 現對於元物件可分成兩個部分

$$\mathbf{D} = \mathbf{D}_{FGO} + \mathbf{D}_{BGO} \text{。}$$

[0113]由於矩陣 $\mathbf{D}$ 之每一行對應於原始音訊物件信號，可藉由將對應於其他元物件之行設定為零而獲得兩個分量向下混頻矩陣(假定無原始物件可存在於兩個元物件上)。換言之，對應於元物件 BGO 之行在 $\mathbf{D}_{FGO}$ 中設定為零，且反之亦然。

[0114]此等新向下混頻矩陣描述可自輸入物件獲得兩個元物件之方式，即：

$$\mathbf{S}_{FGO} = \mathbf{D}_{FGO} \mathbf{S} \quad \text{且} \quad \mathbf{S}_{BGO} = \mathbf{D}_{BGO} \mathbf{S} ,$$

且實際向下混頻簡化為

$$[0115] \quad \mathbf{X} = \mathbf{S}_{FGO} + \mathbf{S}_{BGO} .$$

[0116]亦可認為物件(例如，SAOC)解碼器試圖重建構元物件：

$$\tilde{\mathbf{S}}_{FGO} \approx \mathbf{S}_{FGO} \quad \text{且} \quad \tilde{\mathbf{S}}_{BGO} \approx \mathbf{S}_{BGO} ,$$

且 DE 特定顯現可寫為此等兩個元物件重建構之組合：

$$[0117] \quad \mathbf{Y} = g_{FGO} \mathbf{S}_{FGO} + g_{BGO} \mathbf{S}_{BGO} \approx g_{FGO} \tilde{\mathbf{S}}_{FGO} + g_{BGO} \tilde{\mathbf{S}}_{BGO} .$$

[0118]物件響度估計將兩個元物件 $\mathbf{S}_{FGO}$ 及 $\mathbf{S}_{BGO}$ 接收為輸入，且估計其中每一者之響度： L_{FGO} 為 $\mathbf{S}_{FGO}$ 之(總/總體)響度，且 L_{BGO} 為 $\mathbf{S}_{BGO}$ 之(總/總體)響度。傳輸及/或儲存此等響度值。

[0119]替代地，使用該等元物件中之一者(例如，FGO)作為參考，有可能將此等兩個物件之響度差異計算為(例如)

$$[0120] \quad \Delta L_{FGO} = L_{BGO} - L_{FGO} .$$

[0121]接著傳輸及/或儲存此單一值。

[0122]圖 8 例示根據另一實施例之編碼器。圖 8 之編碼器包括物件向下混頻器 811 及物件旁側資訊估計器 812。此

外，圖 8 之編碼器進一步包括物件響度編碼單元 820。此外，圖 8 之編碼器包括元音訊物件混頻器 805。

[0123]圖 8 之編碼器使用中間音訊元物件作為至物件響度估計之輸入。在諸實施例中，圖 8 之編碼器可經組配以產生兩個音訊元物件。在其他實施例中，圖 8 之編碼器可經組配以產生三個或三個以上音訊元物件。

[0124]所提供之概念尤其提供編碼器可例如估計所有輸入物件之平均響度的新特徵。該等物件可例如在所傳輸之向下混頻信號內混頻。此外，所提供之概念提供可將物件響度及向下混頻資訊例如包含在所傳輸之物件寫碼旁側資訊中之新特徵。

[0125]舉例而言，解碼器可使用物件寫碼旁側資訊來(虛擬地)分隔物件，且使用顯現資訊重新組合該等物件。

[0126]此外，所提供之概念提供可使用向下混頻資訊來估計預設向下混頻信號之響度、可使用顯現資訊及所接收物件響度來估計輸出信號之平均響度及/或可自此等兩個值估計響度改變之新特徵。或者，可使用向下混頻及顯現資訊來自預設向下混頻估計響度改變，此為所提供之概念的另一新特徵。

[0127]此外，所提供之概念提供可修改解碼器輸出以補償響度改變以使得經修改信號之平均響度匹配預設向下混頻之平均響度的新特徵。

[0128]圖 9 中例示與 SAOC-DE 有關之特定實施例。系統接收輸入音訊物件信號、向下混頻資訊，以及物件至元

物件之分群資訊。基於此等資訊，元音訊物件混頻器 905 形成兩個元物件 s_{FGO} 及 s_{BGO} 。有可能用 SAOC 處理之信號部分並不組成整個信號。舉例而言，在 5.1 聲道組態中，SAOC 可部署於聲道之子集中，如部署於前聲道上(左、右及中央)，而其他聲道(左環繞、右環繞及低頻效果)繞開(略過)SAOC 且如此遞送。未經 SAOC 處理之此等聲道標示為 x_{BYPASS} 。需要提供可能之旁路聲道以使編碼器更準確地估計響度資訊。

[0129]可以多種方式處置旁路聲道。

[0130]舉例而言，旁路聲道可例如形成獨立元物件。此允許界定顯現以使得所有三個元物件獨立地進行按比例調整。

[0131]或者，舉例而言，旁路聲道可例如與其他兩個元物件中之一者組合。該元物件之顯現設定亦控制旁路聲道部分。舉例而言，在對話增強場景中，組合旁路聲道與背景元物件係有意義的： $x_{BGO} = s_{BGO} + x_{BYPASS}$ 。

[0132]或者，舉例而言，可例如忽略旁路聲道。

[0133]根據諸實施例，編碼器之基於物件之編碼單元 210 經組配以接收音訊物件信號，其中該等音訊物件信號中之每一者被指派至僅兩個群組中之僅一者，其中該僅兩個群組中之每一者包括該等音訊物件信號中的一或多者。此外，基於物件之編碼單元 210 經組配以向下混頻該僅兩個群組包含之音訊物件信號以獲得包括一或多個向下混頻音訊聲道之向下混頻信號作為經編碼音訊信號，其中該一或

多個向下混頻聲道之數目小於該僅兩個群組所包含之音訊物件信號之數目。物件響度編碼單元 220 經指派以接收一或多個其他旁路音訊物件信號，其中該一或多個其他旁路音訊物件信號中之每一者被指派至第三群組，其中該一或多個其他旁路音訊物件信號中之每一者不被第一群組包括且不被第二群組包括，其中基於物件之編碼單元 210 經組配以不在向下混頻信號內向下混頻該一或多個其他旁路音訊物件信號。

[0134]在一實施例中，物件響度編碼單元 220 經組配以判定響度資訊之第一響度值、第二響度值及第三響度值，該第一響度值指示第一群組的一或多個音訊物件信號之總響度，該第二響度值指示第二群組的一或多個音訊物件信號之總響度，且該第三響度值指示第三群組的一或多個其他旁路音訊物件信號之總響度。在另一實施例中，物件響度編碼單元 220 經組配以判定響度資訊之第一響度值及第二響度值，該第一響度值指示第一群組的一或多個音訊物件信號之總響度，且該第二響度值指示第二群組的一或多個音訊物件信號及第三群組的一或多個其他旁路音訊物件信號之總響度。

[0135]根據一實施例，解碼器之接收介面 110 經組配以接收向下混頻信號。此外，接收介面 110 經組配以接收一或多個其他旁路音訊物件信號，其中該一或多個其他旁路音訊物件信號未在該向下混頻信號內混頻。此外，接收介面 110 經組配以接收響度資訊，該響度資訊指示關於在向

下混頻信號內混頻之音訊物件信號之響度的資訊且指示關於未在向下混頻信號內混頻之一或多個其他旁路音訊物件信號之響度的資訊。此外，信號處理器 120 經組配以取決於關於在向下混頻信號內混頻之音訊物件信號之響度的資訊且取決於關於未在向下混頻信號內混頻之一或多個其他旁路音訊物件信號之響度的資訊而判定響度補償值。

[0136]圖 9 例示根據一實施例與 SAOC-DE 有關的編碼器及解碼器，其包括旁路聲道。圖 9 之編碼器尤其包括 SAOC 編碼器 902。

[0137]在圖 9 之實施例中，旁路聲道與另一元物件之可能組合發生於兩個「旁路包含」區塊 913、914 中，從而產生具有來自所包含之旁路聲道的所界定部分之元物件 $\mathbf{x}_{FGO}$ 及 $\mathbf{x}_{BGO}$ 。

[0138]在響度估計單元 921、922、923 中估計此等元物件兩者的知覺響度 L_{BYPASS} 、 L_{FGO} 及 L_{BGO} 。此響度資訊接著在元物件響度資訊估計器 925 中變換為適當編碼，且接著傳輸及/或儲存。

[0139]實際 SAOC 編碼器及解碼器按照預期操作，從而自物件提取物件旁側資訊、建立向下混頻信號 $\mathbf{x}$ ，且將該資訊傳輸及/或儲存至解碼器。可能旁路聲道隨其他資訊一起傳輸及/或儲存至解碼器。

[0140]SAOC-DE 解碼器 945 接收增益值「對話增益」作為使用者輸入。基於此輸入及所接收之向下混頻資訊，SAOC 解碼器 945 判定顯現資訊。SAOC 解碼器 945 接著產

生所顯現之輸出場景作為信號 $\mathbf{Y}$ 。此外，其產生應應用於可能旁路信號 $\mathbf{x}_{\text{BYPASS}}$ 之增益因數(及延遲值)。

[0141]「旁路包含」單元 955 連同所顯現之輸出場景及旁路信號一起接收此資訊，且建立完整輸出場景信號。SAOC 解碼器 945 亦產生一組元物件增益值，此等增益值之量取決於元物件分群及所要響度資訊形式。

[0142]增益值提供至亦自編碼器接收元物件響度資訊之混頻物響度估計器 960。

[0143]混頻物響度估計器 960 接著能夠判定所要響度資訊，其包含但不限於向下混頻信號之響度、所顯現輸出場景之響度及/或向下混頻信號與所顯現輸出場景之間的響度差異。

[0144]在一些實施例中，響度資訊自身即為足夠的，而在其他實施例中，需要取決於所判定之響度資訊而處理完整輸出。此處理可例如為補償向下混頻信號與所顯現輸出場景之間的任何可能響度差異。例如由響度處理單元 970 進行之此處理將在廣播場景中有意義，因為其將減少所感知信號響度的改變而無關於使用者互動(輸入「對話增益」之設定)。

[0145]此特定實施例中之響度相關處理包括多個新特徵。尤其是，FGO、BGO 及可能旁路聲道在最終聲道組態中預先混頻，以使得可藉由簡單地將兩個預先混頻的信號相加在一起而進行向下混頻(例如，向下混頻矩陣係數 1)，其構成新特徵。此外，作為另一新特徵，估計 FGO 與 BGO

之平均響度，且計算差異。此外，物件混頻於所傳輸之向下混頻信號中。此外，作為另一新特徵，將響度差異資訊包含至所傳輸之旁側資訊。(新)此外，解碼器使用該旁側資訊用於(虛擬地)分隔物件，且使用基於向下混頻資訊及使用者輸入修改增益之顯現資訊重新組合該等物件。此外，作為另一新特徵，解碼器使用修改增益及所傳輸之響度資訊來估計與預設向下混頻相比之系統輸出的平均響度之改變。

[0146]在下文中，提供實施例之正式描述。

[0147]假定物件響度值特徵類似於能量值之對數(當對物件求和時)，即響度值必須變換至線性域，在該處相加，且最終變換回至對數域。現將呈現經由 BS.1770 之定義來促進此操作(為簡單起見，聲道之數目設定為 1，但相同原理可適用於多聲道信號(在聲道上適當求和))。

[0148]具有均方能量 e_i 之第 i 個 K 濾波信號 z_i 的響度定義為

$$L_i = c + 10 \log_{10} e_i,$$

其中 c 為偏移常數。舉例而言， c 可為 -0.691。由此，可藉由下式判定信號之能量

$$e_i = 10^{(L_i - c)/10}。$$

[0149] N 個不相關信號之能量總和 $z_{SUM} = \sum_{i=1}^N z_i$ 由此為

$$e_{SUM} = \sum_{i=1}^N e_i = \sum_{i=1}^N 10^{(L_i - c)/10}，$$

且此總和信號之響度由此為

$$L_{SUM} = c + 10 \log_{10} e_{SUM} = c + 10 \log_{10} \sum_{i=1}^N 10^{(L_i - c)/10} \text{。}$$

[0150]若信號不相關，則在對總和信號之能量求近似時必須考慮相關係數 $C_{i,j}$ ：

$$e_{SUM} = \sum_{i=1}^N \sum_{j=1}^N e_{i,j} \text{，}$$

其中第 i 個物件與第 j 個物件之間的交叉能量 $e_{i,j}$ 定義為

$$\begin{aligned} e_{i,j} &= C_{i,j} \sqrt{e_i e_j} \\ &= C_{i,j} \sqrt{10^{(L_i - c)/10} 10^{(L_j - c)/10}} \\ &= C_{i,j} \sqrt{10^{(L_i + L_j - 2c)/10}} \text{，} \end{aligned}$$

其中 $-1 \leq C_{i,j} \leq 1$ 為兩個物件 i 與 j 之間的相關係數。當兩個物件不相關時，相關係數等於 0，且當兩個物件相同時，相關係數等於 1。

[0151]利用待在混頻過程中應用於信號上的混頻權重 g_i 來進一步擴展該模型，即 $z_{SUM} = \sum_{i=1}^N g_i z_i$ ，總和信號之能量將為

$$e_{SUM} = \sum_{i=1}^N \sum_{j=1}^N g_i g_j e_{i,j} \text{，}$$

且可如前所述自此獲得混頻信號之響度，其中

$$L_{SUM} = c + 10 \log_{10} e_{SUM} \text{。}$$

[0152]兩個信號的響度之間的差可估計為

$$\Delta L(i, j) = L_i - L_j \text{。}$$

[0153]若現在使用如前所述之響度定義，則此可寫為

$$\begin{aligned}\Delta L(i, j) &= L_i - L_j \\ &= (c + 10 \log_{10} e_i) - (c + 10 \log_{10} e_j) \\ &= 10 \log_{10} \frac{e_i}{e_j},\end{aligned}$$

可觀察到其為信號能量之函數。若現在需要估計具有可能不同的混頻權重 g_i 及 h_i 的兩個混頻物之間的響度差異

$$z_A = \sum_{i=1}^N g_i z_i \quad \text{且} \quad z_B = \sum_{i=1}^N h_i z_i$$

則此可用下式進行估計

$$\begin{aligned}\Delta L(A, B) &= 10 \log_{10} \frac{e_A}{e_B} \\ &= 10 \log_{10} \frac{\sum_{i=1}^N \sum_{j=1}^N g_i g_j e_{i,j}}{\sum_{i=1}^N \sum_{j=1}^N h_i h_j e_{i,j}} \\ &= 10 \log_{10} \frac{\sum_{i=1}^N \sum_{j=1}^N g_i g_j C_{i,j} \sqrt{10^{(L_i + L_j - 2c)/10}}}{\sum_{i=1}^N \sum_{j=1}^N h_i h_j C_{i,j} \sqrt{10^{(L_i + L_j - 2c)/10}}} \\ &= 10 \log_{10} \frac{\sum_{i=1}^N \sum_{j=1}^N g_i g_j C_{i,j} \sqrt{10^{(L_i + L_j)/10}}}{\sum_{i=1}^N \sum_{j=1}^N h_i h_j C_{i,j} \sqrt{10^{(L_i + L_j)/10}}}.\end{aligned}$$

[0154]在物件不相關的情況下 ($C_{i,j} = 0, \forall i \neq j$ 且 $C_{i,j} = 1, \forall i = j$)，差異估計變為

$$\begin{aligned}\Delta L(A, B) &= 10 \log_{10} \frac{\sum_{i=1}^N g_i^2 10^{(L_i - c)/10}}{\sum_{i=1}^N h_i^2 10^{(L_i - c)/10}} \\ &= 10 \log_{10} \frac{\sum_{i=1}^N g_i^2 10^{L_i/10}}{\sum_{i=1}^N h_i^2 10^{L_i/10}}.\end{aligned}$$

[0155]在下文中，考慮差分編碼。

[0156]有可能將每物件響度值編碼為來自所選參考物件的響度的差異：

$$K_i = L_i - L_{REF},$$

其中 L_{REF} 為參考物件之響度。此編碼在不需要絕對響度值作為結果的情況下係有益的，因為現在有必要傳輸的值少了一個，且響度差異估計可寫為

$$\Delta L(A, B) = 10 \log_{10} \frac{\sum_{i=1}^N \sum_{j=1}^N g_i g_j C_{i,j} \sqrt{10^{(K_i + K_j)/10}}}{\sum_{i=1}^N \sum_{j=1}^N h_i h_j C_{i,j} \sqrt{10^{(K_i + K_j)/10}}},$$

或在物件不相關的情況下

$$\Delta L(A, B) = 10 \log_{10} \frac{\sum_{i=1}^N g_i^2 10^{K_i/10}}{\sum_{i=1}^N h_i^2 10^{K_i/10}}.$$

[0157]在下文中，考慮對話增強場景。

[0158]再次考慮對話增強的應用場景。在解碼器中界定顯現資訊之自由度僅限於改變兩個元物件之位準。此外，假定兩個元物件不相關，即 $C_{FGO, BGO} = 0$ 。若元物件之向下混

頻權重為 h_{FGO} 及 h_{BGO} ，且以增益 f_{FGO} 及 f_{BGO} 來顯現該等元物件，則相對於預設向下混頻之輸出響度為

$$\begin{aligned}\Delta L(A, B) &= 10 \log_{10} \frac{f_{FGO}^2 10^{(L_{FGO}-\epsilon)/10} + f_{BGO}^2 10^{(L_{BGO}-\epsilon)/10}}{h_{FGO}^2 10^{(L_{FGO}-\epsilon)/10} + h_{BGO}^2 10^{(L_{BGO}-\epsilon)/10}} \\ &= 10 \log_{10} \frac{f_{FGO}^2 10^{L_{FGO}/10} + f_{BGO}^2 10^{L_{BGO}/10}}{h_{FGO}^2 10^{L_{FGO}/10} + h_{BGO}^2 10^{L_{BGO}/10}}.\end{aligned}$$

[0159]若需要在輸出與預設向下混頻中具有相同響度，
則此由此亦需要補償。

[0160] $\Delta L(A, B)$ 可認為係響度補償值，其可由解碼器之信號處理器 120 傳輸。 $\Delta L(A, B)$ 亦可命名為響度改變值，且因此實際補償值可為倒數值。或者，其亦可使用「響度補償因數」來命名？因此，本文件上文中提及之響度補償值 lcv 將對應於下文之值 g_{Δ} 。

[0161]舉例而言， $g_{\Delta} = 10^{-\Delta L(A, B)/20}$ $1/\Delta L(A, B)$ 可應用為藉由將顯現資訊應用於音訊輸入信號而產生的經修改音訊信號之每一聲道上的放大因數。 g_{Δ} 之此等式在線性域中有效。在對數域中，該等式將不同，諸如 $1/\Delta L(A, B)$ ，且相應地應用。

[0162]若向下混頻過程經簡化而使得兩個元物件可以單位權重混頻以獲得向下混頻信號，即， $h_{FGO} = h_{BGO} = 1$ ，且此等兩個物件之顯現增益標示為 g_{FGO} 及 g_{BGO} 。此將響度改變等式簡化為

$$\begin{aligned}\Delta L(A, B) &= 10 \log_{10} \frac{g_{FGO}^2 10^{(L_{FGO}-\epsilon)/10} + g_{BGO}^2 10^{(L_{BGO}-\epsilon)/10}}{10^{(L_{FGO}-\epsilon)/10} + 10^{(L_{BGO}-\epsilon)/10}} \\ &= 10 \log_{10} \frac{g_{FGO}^2 10^{L_{FGO}/10} + g_{BGO}^2 10^{L_{BGO}/10}}{10^{L_{FGO}/10} + 10^{L_{BGO}/10}}.\end{aligned}$$

[0163]再次， $\Delta L(A, B)$ 可認為係藉由信號處理器 120 判定之響度補償值。

[0164]大體而言， g_{FGO} 可認為係前景物件 FGO (前景物件群組)之顯現增益，且 g_{BGO} 可認為係背景物件 BGO (背景物件群組)之顯現增益。

[0165]如前所述，有可能傳輸響度差異而非絕對響度。將參考響度定義為 FGO 元物件之響度 $L_{REF} = L_{FGO}$ ，即， $K_{FGO} = L_{FGO} - L_{REF} = 0$ 且 $K_{BGO} = L_{BGO} - L_{REF} = L_{BGO} - L_{FGO}$ 。現在，響度改變為

$$\Delta L(A, B) = 10 \log_{10} \frac{g_{FGO}^2 + g_{BGO}^2 10^{K_{BGO}/10}}{1 + 10^{K_{BGO}/10}}。$$

[0166]如同 SAOC-DE 中之情況，亦可能兩個元物件不具有個別按比例調整因數，但該等物件中之一者未經修改，而另一者衰減以獲得該等物件之間的正確混頻比率。在此顯現設定中，輸出響度將低於預設混頻物響度，且響度改變為

$$\Delta L(A, B) = 10 \log_{10} \frac{\hat{g}_{FGO}^2 + \hat{g}_{BGO}^2 10^{K_{BGO}/10}}{1 + 10^{K_{BGO}/10}}，$$

其中

$$\hat{g}_{FGO} = \begin{cases} 1 & \text{若, if } g_{FGO} \geq g_{BGO} \\ \frac{g_{FGO}}{g_{BGO}} & \text{若, if } g_{FGO} < g_{BGO} \end{cases}，\text{ 且 } \hat{g}_{BGO} = \begin{cases} \frac{g_{BGO}}{g_{FGO}} & \text{若 } g_{BGO} < g_{FGO} \\ 1 & \text{若 } g_{BGO} \geq g_{FGO} \end{cases}。$$

[0167]形式已經相當簡單，且對於所使用之響度量測來說相當不可知。唯一的現實要求為，響度值應在指數域中

求和。有可能傳輸/儲存信號能量值而非響度值，因為兩者具有緊密關聯。

[0168]在上式中之每一者中， $\Delta L(A, B)$ 可認為係響度補償值，其可由解碼器之信號處理器 120 傳輸。

5 [0169]在下文中，考慮實例情況。經由兩個實例信號說明所提供概念之準確性。兩個信號皆具有 5.1 向下混頻，環繞及左聲道自 SAOC 處理略過。

[0170]使用兩個主要方法：一者為具有三個元物件 FGO、BGO 及旁路聲道之(「3 項」)，例如，

$$10 \quad \mathbf{X} = \mathbf{X}_{FGO} + \mathbf{X}_{BGO} + \mathbf{X}_{BYPASS},$$

[0171]且另一者為具有兩個元物件之(「2 項」)，例如：

$$[0172] \quad \mathbf{X} = \mathbf{X}_{FGO} + \mathbf{X}_{BGO}。$$

[0173]在 2 項方法中，旁路聲道可例如與 BGO 混頻在一起以進行元物件響度估計。估計兩個(或所有三個)物件之響度以及向下混頻信號之響度，且將值儲存。

[0174]顯現指令對於兩個方法分別具有以下形式

$$\mathbf{Y} = \hat{g}_{FGO} \mathbf{X}_{FGO} + \hat{g}_{BGO} \mathbf{X}_{BGO} + \hat{g}_{BGO} \mathbf{X}_{BYPASS}$$

及

$$\mathbf{Y} = \hat{g}_{FGO} \mathbf{X}_{FGO} + \hat{g}_{BGO} \mathbf{X}_{BGO}。$$

20 [0175]增益值係例如根據下式而判定：

$$\hat{g}_{FGO} = \begin{cases} 1 & \text{若 } g_{FGO} > 1 \\ g_{FGO} & \text{否則} \end{cases}, \text{ 且 } \hat{g}_{BGO} = \begin{cases} 1/g_{FGO} & \text{若 } g_{FGO} > 1 \\ 1 & \text{否則} \end{cases},$$

其中 FGO 增益 g_{FGO} 在 -24 至 +24 dB 之間變化。

[0176]顯現輸出場景，量測響度，且計算自向下混頻信號響度之衰減。

[0177]此結果用具有圓圈標記之藍線顯示於圖 10 及圖 11 中。圖 10 描繪量測響度改變及使用所提供概念用於以純
5 參數方式估計響度改變之結果的第一圖解，且圖 11 描繪其第二圖解。

[0178]接下來，使用所儲存之元物件響度值及向下混頻以及顯現資訊來以參數方式估計自向下混頻之衰減。使用
10 三個元物件之響度估計用具有方形標記之綠線例示，且使用兩個元物件之響度估計用具有星形標記之紅線例示。

[0179]自該等圖可看出，2 項及 3 項方法提供實際上相同之結果，且其皆相當良好地近似於量測值。

[0180]所提供之概念展現多個優點。舉例而言，所提供之概念允許自形成混頻物之分量信號的響度估計混頻物信號之響度。此舉之益處為可一次性估計分量信號響度，且
15 可針對任何混頻物以參數方式獲得混頻物之響度估計，而無需實際基於信號之響度估計。此在需要各種混頻物之響度估計的總體系統之計算效率上提供相當大的改良。舉例而言，當最終使用者改變顯現設定時，輸出之響度估計立
20 即可用。

[0181]在一些應用中，諸如當符合 EBU R128 推薦時，整個節目之平均響度係重要的。若基於所接收之信號進行接收器中之響度估計(例如，在廣播場景中)，則估計僅在已接收到整個節目之後才會聚至平均響度。由此，對響度之

任何補償將具有錯誤或展現時間變化。當根據提議估計分量物件之響度且傳輸響度資訊時，有可能在接收器中無延遲地估計平均混頻物響度。

[0182]若需要輸出信號之平均響度保持(大體)恆定而無關於顯現資訊之改變，則出於此原因，所提供之概念允許判定補償因數。解碼器中為此而需之計算使其計算複雜性可忽略，且因此有可能將功能性添加至任何解碼器。

[0183]存在以下情況：輸出之絕對響度位準並不重要，但重要性在於自參考場景判定響度之改變。在此等情況下，物件之絕對位準並不重要，但其相對位準係重要的。此允許將物件中之一者定義為參考物件且相對於此參考物件之響度表示其他物件之響度。考慮到響度資訊之傳送及/或儲存，此舉具有一些益處。

[0184]首先，不必傳送參考響度位準。在兩個元物件之應用情況中，此將待傳輸之資料量減半。第二益處係關於響度值之可能量化及表示。由於物件之絕對位準可為幾乎任何情況，因此絕對響度值亦可幾乎為任何情況。另一方面，假定相對響度值具有 0 平均值及圍繞平均值之形式相當良好的分佈。表示之間的差異允許以準確度潛在較高且與經量化表示使用相同數目的位元的方式來界定相對表示之量化柵格。

[0185]圖 12 例示用於進行響度補償之另一實施例。在圖 12 中，可進行響度補償以例如補償響度損失。處於此目的，例如，可使用來自 **DE_control_info** 之值

DE_loudness_diff_dialogue (= K_{FGO}) 及
DE_loudness_diff_background (= K_{BGO})。此處，
DE_control_info 可指定高級清潔音訊「對話增強」(DE)控制資訊

5 [0186]響度補償係藉由將增益值「 g 」應用於 SAOC-DE 輸出信號及旁路聲道(在多聲道信號之情況下)而達成。

[0187]在圖 12 之實施例中，如下進行此操作：

使用有限對話修改增益值 m_G 來判定前景物件(FGO，例如，對話)及背景物件(BGO，例如，周圍環境)之有效增益。
 10 此操作由「增益對映」區塊1220進行，「增益對映」區塊1220 產生增益值 m_{FGO} 及 m_{BGO} 。

[0188]「輸出響度估計器」區塊 1230 使用響度資訊 K_{FGO} 及 K_{BGO} 及有效增益值 m_{FGO} 及 m_{BGO} 來估計與預設向下混頻情況相比之此可能響度改變。接著將該改變對映至「響度補償因數」，該響度補償因數應用於輸出聲道上以用於產生
 15 最終「輸出信號」。

[0189]應用以下步驟來進行響度補償：

- 自 SAOC-DE 解碼器(如在條款 12.8 「SAOC-DE」

[DE]的修改範圍控制中所定義)接收有限增益值 m_G ，且判定
 20 所應用之 FGO/BGO 增益：

$$\begin{cases} m_{FGO} = m_G, \text{ 且 } m_{BGO} = 1 & \text{若 } m_G \leq 1 \\ m_{FGO} = 1, \text{ 且 } m_{BGO} = m_G^{-1} & \text{若 } m_G > 1 \end{cases}.$$

- 獲得元物件響度資訊 K_{FGO} 及 K_{BGO} 。

- 利用下式計算與預設向下混頻相比的輸出響度之改變

$$\Delta L = 10 \log_{10} \frac{m_{FGO}^2 10^{K_{FGO}/10} + m_{BGO}^2 10^{K_{BGO}/10}}{10^{K_{FGO}/10} + 10^{K_{BGO}/10}}。$$

- 計算響度補償增益 $g_{\Delta} = 10^{-0.05\Delta L}$ 。

$$\mathbf{g} = \begin{bmatrix} g_1 \\ \vdots \\ g_N \end{bmatrix}$$

- 計算按比例調整因數

$$g_i = \begin{cases} g_{\Delta} & \text{若聲道 } i \text{ 屬 SAOC-DE 輸出} \\ m_{BGO} g_{\Delta} & \text{若聲道 } i \text{ 為旁路聲道} \end{cases}, \text{ 且 } N \text{ 為輸出聲道之}$$

總數目。在圖 12 中，增益調整分成兩個步驟：在與「SAOC-DE 輸出聲道」組合之前用 m_{BGO} 調整可能「旁路聲道」之增益，且接著將共同增益 g_{Δ} 應用於所有經組合聲道上。此僅為增益調整操作之一種可能重排序，而此處 $\mathbf{g}$ 將兩個增益調整步驟組合成一個增益調整。

將按比例調整值 $\mathbf{g}$ 應用於由「SAOC-DE 輸出聲道」 $\mathbf{Y}_{SAOC}$ 及可能時間對準之「旁路聲道」組成之音訊聲道 $\mathbf{Y}_{FULL}$ 上，
 $\mathbf{Y}_{BYPASS} : \mathbf{Y}_{FULL} = \mathbf{Y}_{SAOC} \cup \mathbf{Y}_{BYPASS}$

[0190] 將按比例調整值 $\mathbf{g}$ 應用於音訊聲道 $\mathbf{Y}_{FULL}$ 上由增益調整單元 1240 進行。

[0191] 如上文所計算之 ΔL 可認為係響度補償值。大體而言， m_{FGO} 指示前景物件 FGO (前景物件群組) 之顯現增益，且 m_{BGO} 指示背景物件 BGO (背景物件群組) 之顯現增益。

[0192]儘管在設備之上下文中已描述了一些態樣，但清楚的是，此等態樣亦表示對應方法之描述，其中一區塊或裝置對應於一方法步驟或一方法步驟之一特徵。類似地，方法步驟之上下文中所描述之態樣亦表示對應設備之對應
5 區塊或項目或特徵的描述。

[0193]發明性分解信號可儲存在數位儲存媒體上或可在諸如無線傳輸媒體或有線傳輸媒體之傳輸媒體(諸如網際網路)上傳輸。

[0194]取決於某些實施要求，本發明之實施例可在硬體
10 中或軟體中實施。可使用上面儲存有電子可讀控制信號之數位儲存媒體(例如，軟磁碟、DVD、CD、ROM、PROM、EPROM、EEPROM 或快閃記憶體)來進行該實施方案，該數位儲存媒體與可規劃電腦系統協作(或能夠與之協作)，使得進行個別地方法。

[0195]根據本發明之一些實施例包括具有電子可讀控制信號的資料載體，其能夠與可規劃電腦系統協作，使得
15 進行本文所述方法中之一者。

[0196]通常，本發明之實施例可實施為具有程式代碼之電腦程式產品，當該電腦程式產品在電腦上運行時，該程
20 式代碼操作以用於進行該等方法中之一者。該程式代碼可例如儲存於機器可讀載體上。

[0197]其他實施例包括儲存於機器可讀載體上之用於進行本文所述方法中之一者的電腦程式。

[0198]換言之，本發明之方法的實施例因此為具有程式代碼之電腦程式，當該電腦程式在電腦上運行時，該程式代碼用於進行本文所述之方法中的一者。

[0199]本發明之方法的另一實施例因此為資料載體(或數位儲存媒體，或電腦可讀媒體)，其上面記錄有用於進行本文所述方法中之一者的電腦程式。

[0200]本發明之方法的另一實施例因此為表示用於進行本文所述方法中之一者的電腦程式的資料串流或信號序列。該資料串流或信號序列可例如經組配來經由資料通信連接(例如經由網際網路)傳送。

[0201]另一實施例包括一種處理構件，例如電腦或可規劃邏輯裝置，其經組配來或適於進行本文所述方法中的一者。

[0202]另一實施例包括一種電腦，其上面安裝有用於進行本文所述方法中之一者的電腦程式。

[0203]在一些實施例中，一種可規劃邏輯裝置(例如，現場可規劃門陣列)可用以進行本文所述方法之功能性中的一些或全部。在一些實施例中，現場可規劃門陣列可與微處理器協作，以便進行本文所述方法中之一者。通常，該等方法較佳由任何硬體設備進行。

[0204]上文所述之實施例僅例示本發明之原理。應理解，本文描述之佈置及細節之修改及改變對於熟習此項技術者而言將顯而易見。因此，本發明僅欲受所附申請專利範圍的範疇限制，而不受為描述及解釋本文之實施例而呈現之特定細節之限

制。

參考文獻

[BCC] C. Faller and F. Baumgarte, 「Binaural Cue Coding - Part II: Schemes and applications,」 IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003.

[EBU] EBU Recommendation R 128 「Loudness normalization and permitted maximum level of audio signals」, Geneva, 2011.

[JSC] C. Faller, 「Parametric Joint-Coding of Audio Sources」, 120th AES Convention, Paris, 2006.

[ISS1] M. Parvaix and L. Girin: 「Informed Source Separation of underdetermined instantaneous Stereo Mixtures using Source Index Embedding」, IEEE ICASSP, 2010.

[ISS2] M. Parvaix, L. Girin, J.-M. Brossier: 「A watermarking-based method for informed source separation of audio signals with a single sensor」, IEEE Transactions on Audio, Speech and Language Processing, 2010.

[ISS3] A. Liutkus and J. Pinel and R. Badeau and L. Girin and G. Richard: 「Informed source separation through spectrogram coding and data embedding」, Signal Processing Journal, 2011.

[ISS4] A. Ozerov, A. Liutkus, R. Badeau, G. Richard: 「Informed source separation: source coding meets source separation」, IEEE Workshop on Applications of Signal

Processing to Audio and Acoustics, 2011.

[ISS5] S. Zhang and L. Girin: 「An Informed Source Separation System for Speech Signals」, INTERSPEECH, 2011.

[ISS6] L. Girin and J. Pintel: 「Informed Audio Source Separation from Compressed Linear Stereo Mixtures」, AES 42nd International Conference: Semantic Audio, 2011.

[ITU] International Telecommunication Union: 「Recommendation ITU-R BS.1770-3 – Algorithms to measure audio programme loudness and true-peak audio level」, Geneva, 2012.

[SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: 「From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio」, 22nd Regional UK AES Conference, Cambridge, UK, April 2007.

[SAOC2] J. Engdegård, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hölzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen: 「Spatial Audio Object Coding (SAOC) – The Upcoming MPEG Standard on Parametric Object Based Audio Coding」, 124th AES Convention, Amsterdam 2008.

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

[EP] EP 2146522 A1: S. Schreiner, W. Fiesel, M. Neusinger, O. Hellmuth, R. Sperschneider, 「Apparatus and method for generating audio output signals using object based metadata」, 2010.

[DE] ISO/IEC, 「MPEG audio technologies – Part 2: Spatial Audio Object Coding (SAOC) – Amendment 3, Dialogue Enhancement,」 ISO/IEC 23003-2:2010/DAM 3, Dialogue Enhancement.

[BRE] WO 2008/035275 A2.

[SCH] EP 2 146 522 A1.

[ENG] WO 2008/046531 A1.

【符號說明】

110...接收介面	720...物件響度編碼單元
120...信號處理器	730...傳送串流
210...基於物件之編碼單元	740...輸出響度估計器
220...物件響度編碼單元	750...基於物件之音訊解碼單
310...編碼器	元
320...解碼器	805...元音訊物件混頻器
410...SAOC編碼器	811...物件向下混頻器
420...SAOC解碼器	812...物件旁側資訊估計器
510...旁側資訊解碼器	820...物件響度編碼單元
520...物件分離器	902...SAOC編碼器
530...顯現器	905...元音訊物件混頻器
710...基於物件之音訊編碼器	913...「旁路包含」區塊

914...「旁路包含」區塊

960...混頻物響度估計器

921...響度估計單元

970...響度處理單元

922...響度估計單元

1220...「增益對映」區塊

923...響度估計單元

1230...「輸出響度估計器」區

925...元物件響度資訊估計器

塊

945...SAOC解碼器

1240...增益調整單元

955...「旁路包含」單元

申請專利範圍

1. 一種用於產生包括一或多個音訊輸出聲道之音訊輸出信號的解碼器，其中該解碼器包括：

一接收介面，其用於接收包含多個音訊物件信號之一音訊輸入信號、用於接收關於該等音訊物件信號之響度資訊，及用於接收指示該等音訊物件信號中的一或多者是否應放大或衰減的顯現資訊，及

一信號處理器，其用於產生該音訊輸出信號之該一或多個音訊輸出聲道，

其中該接收介面經組配以接收包括一或多個降混頻聲道之一降混頻信號作為該音訊輸入信號，其中該一或多個降混頻聲道包括該等音訊物件信號，以及其中該一或多個向下混頻聲道之數目小於該等音訊物件信號之數目，

其中該接收介面經組配以接收關於該音訊輸入信號的降混頻資訊，該降混頻資訊指示該等音訊物件信號將如何混合於該一或多個降混頻聲道內，其中該接收介面經組配以接收一或多個其他旁路音訊物件信號，其中該一或多個其他旁路音訊物件信號不混合於該降混頻信號內，

其中該接收介面經組配以接收該響度資訊，該響度資訊指示和要混合在該降混頻信號內之該等音訊物件信號之該響度有關的資訊且指示和未混合在該降混頻

信號內之該一或多個其他旁路音訊物件信號之該響度有關的資訊，

其中該信號處理器經組配以取決於和在該降混頻信號內混合之該等音訊物件信號之該響度有關的該資訊且取決於和未在該降混頻信號內混合之該一或多個其他旁路音訊物件信號之該響度有關的該資訊而判定一響度補償值，且

其中該信號處理器經組配以取決於該降混頻資訊、取決於該顯現資訊及取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道。

2. 如請求項 1 之解碼器，其中該信號處理器經組配以取決於該顯現資訊及取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道，使得該音訊輸出信號之一響度等於該音訊輸入信號之一響度，或使得該音訊輸出信號之該響度比一經修改音訊信號之一響度更接近於該音訊輸入信號之響度，該經修改音訊信號會藉由根據該顯現資訊放大或衰減該音訊輸入信號之該等音訊物件信號而修改該音訊輸入信號所導致。

3. 如請求項 2 之解碼器，

其中該信號處理器經組配以藉由根據該顯現資訊放大或衰減該音訊輸入信號之該等音訊物件信號而修改該音訊輸入信號來產生該經修改音訊信號，且

其中該信號處理器經組配以藉由將該響度補償值施用於該經修改音訊信號上而產生該音訊輸出信號，使得該音訊輸出信號之該響度等於該音訊輸入信號之該響度，或使得該音訊輸出信號之該響度比該經修改音訊信號之該響度更接近於該音訊輸入信號之該響度。

4. 如請求項 1 之解碼器，

其中該音訊輸入信號之該等音訊物件信號中之每一者被指派至兩個或兩個以上群組中的僅一個群組，其中該兩個或兩個以上群組中之每一者包含該音訊輸入信號之該等音訊物件信號中的一或多者，

其中該接收介面經組配以接收該兩個或兩個以上群組中之每一群組的一響度值作為該響度資訊，

其中該信號處理器經組配以取決於該兩個或兩個以上群組中之每一者的該響度值而判定該響度補償值，且

其中該信號處理器經組配以取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道。

5. 如請求項 1 之解碼器，其中該兩個或兩個以上群組中之至少一個群組包括兩個或更多個該等音訊物件信號。

6. 如請求項 1 之解碼器，

其中該音訊輸入信號之該等音訊物件信號中之每一者被指派至作為該兩個或兩個以上群組的僅兩個群組中之僅一個群組，

其中該音訊輸入信號之該等音訊物件信號中之每一者被指派至該僅兩個群組中之一前景物件群組或該僅兩個群組中之一背景物件群組，

其中該接收介面經組配以接收該前景物件群組之該響度值，

其中該接收介面經組配以接收該背景物件群組之該響度值，

其中該信號處理器經組配以取決於該前景物件群組之該響度值且取決於該背景物件群組之該響度值而判定該響度補償值，且

其中該信號處理器經組配以取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道。

7. 如請求項 6 之解碼器，

其中該信號處理器經組配以根據下式判定一響度補償值 ΔL

$$\Delta L = 10 \log_{10} \frac{m_{FGO}^2 10^{K_{FGO}/10} + m_{BGO}^2 10^{K_{BGO}/10}}{10^{K_{FGO}/10} + 10^{K_{BGO}/10}}$$

其中 K_{FGO} 指示該前景物件群組之該響度值，

其中 K_{BGO} 指示該背景物件群組之該響度值，

其中 m_{FGO} 指示該前景物件群組之一重現增益，且

其中 m_{BGO} 指示該背景物件群組之一重現增益。

8. 如請求項 6 之解碼器，

其中該信號處理器經組配以根據下式判定一響度

補償值 ΔL

$$\Delta L = 10 \log_{10} \frac{g_{FGO}^2 10^{L_{FGO}/10} + g_{BGO}^2 10^{L_{BGO}/10}}{10^{L_{FGO}/10} + 10^{L_{BGO}/10}}$$

其中 L_{FGO} 指示該前景物件群組之該響度值，

其中 L_{BGO} 指示該背景物件群組之該響度值，

5 其中 g_{FGO} 指示該前景物件群組之一重現增益，且

其中 g_{BGO} 指示該背景物件群組之一重現增益。

9. 一種用以編碼多個音訊物件信號之編碼器，其包括：

一基於物件之編碼單元，其用於編碼該等多個音訊物件信號以獲得包含該多個音訊物件信號之一經編碼音訊信號，及

一物件響度編碼單元，其用於編碼關於該等音訊物件信號之響度資訊，

其中該響度資訊包括一或多個響度值，其中該一或多個響度值中之每一者取決於該等音訊物件信號中的一或多者，

其中該基於物件之編碼單元經組配以接收該等音訊物件信號，其中該等音訊物件信號中之每一者被指派至兩個或兩個以上群組中的僅一者，其中該兩個或兩個以上群組中之每一者包括該等音訊物件信號中的一或多者，

其中該基於物件之編碼單元經組配以降混頻由該兩個或兩個以上群組構成之該等音訊物件信號以獲得包括一或多個降混頻音訊聲道之一降混頻信號作為該

經編碼音訊信號，其中該一或多個降混頻聲道之數目小於該兩個或兩個以上群組包括之該等音訊物件信號之數目，

其中該物件響度編碼單元經指派以接收一或多個其他旁路音訊物件信號，其中該一或多個其他旁路音訊物件信號中之每一者被指派至一第三群組，其中該一或多個其他旁路音訊物件信號中之每一者不被該第一群組包括且不被該第二群組包括，其中該基於物件之編碼單元經組配以不在該降混頻信號內降混頻該一或多個其他旁路音訊物件信號，且

其中該物件響度編碼單元經組配以判定該響度資訊之一第一響度值、一第二響度值及一第三響度值，該第一響度值指示該第一群組之該一或多個音訊物件信號之一總響度，該第二響度值指示該第二群組之該一或多個音訊物件信號之一總響度，且該第三響度值指示該第三群組之該一或多個其他旁路音訊物件信號之一總響度，或經組配以判定該響度資訊之一第一響度值及一第二響度值，該第一響度值指示該第一群組之該一或多個音訊物件信號之一總響度，且該第二響度值指示該第二群組之該一或多個音訊物件信號及該第三群組之該一或多個其他旁路音訊物件信號之一總響度。

10. 如請求項 9 之編碼器，

其中該兩個或兩個以上群組為僅兩個群組，

其中該等音訊物件信號中之每一者被指派至該僅

兩個群組中之僅一者，其中該僅兩個群組中之每一者包括該等音訊物件信號中的一或多者，

其中該基於物件之編碼單元經組配以降混頻由該僅兩個群組構成之該等音訊物件信號以獲得包括一或多個降混頻音訊聲道之一降混頻信號作為該經編碼音訊信號，其中該一或多個降混頻聲道之數目小於該僅兩個群組包括之該等音訊物件信號之數目。

11. 一種用以編碼多個音訊物件信號以獲得經編碼音訊信號及用以自該經編碼音訊信號來產生音訊輸出信號之系統，其包括：

如請求項9或10之編碼器，其用於編碼該等多個音訊物件信號以獲得包括該等多個音訊物件信號之該經編碼音訊信號，及

如請求項1至8中任一項之解碼器，其用於產生包括一或多個音訊輸出聲道之該音訊輸出信號，

其中該解碼器經組配以接收該經編碼音訊信號作為一音訊輸入信號且接收該響度資訊

其中該解碼器經組配以進一步接收顯現資訊，

其中該解碼器經組配以取決於該響度資訊且取決於該顯現資訊而判定一響度補償值，且

其中該解碼器經組配以取決於該顯現資訊且取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道。

12. 一種用於產生包括一或多個音訊輸出聲道之音訊輸出信號的方法，其中該方法包括：

接收包括多個音訊物件信號之一音訊輸入信號，

接收響度資訊，該響度資訊指示和在該降混頻信號內混合之該等音訊物件信號之該響度有關的資訊且指示和未在該降混頻信號內混合之該一或多個其他旁路音訊物件信號之該響度有關的資訊，及

接收指示該等音訊物件信號中的一或多者是否應放大或衰減之顯現資訊，

接收包括一或多個降混頻聲道之一降混頻信號作為該音訊輸入信號，其中該一或多個降混頻聲道包括該等音訊物件信號，且其中該一或多個降混頻聲道之數目小於該等音訊物件信號之數目，

接收關於該音訊輸入信號的降混頻資訊，其指示該等音訊物件信號如何混合於該一或多個降混頻聲道內，

接收一或多個其他旁路音訊物件信號，其中該一或多個其他旁路音訊物件信號未在該降混頻信號內混頻，

取決於和在該降混頻信號內混合之該等音訊物件信號之該響度有關的該資訊且取決於和未在該降混頻信號內混合之該一或多個其他旁路音訊物件信號之該響度有關的該資訊而判定一響度補償值，及

取決於該降混頻資訊、取決於該顯現資訊且取決於該響度補償值而自該音訊輸入信號產生該音訊輸出信號之該一或多個音訊輸出聲道。

13. 一種用於編碼之方法，其包括：

編碼包括多個音訊物件信號之一音訊輸入信號，及
編碼關於該等音訊物件信號之響度資訊，其中該響
度資訊包括一或多個響度值，其中該一或多個響度值中
之每一者取決於該等音訊物件信號中的一或多者，

其中該等音訊物件信號中之每一者被指派至兩個
或兩個以上群組中之僅一者，其中該兩個或兩個以上群
組中之每一者包括該等音訊物件信號中的一或多者，

其中編碼關於該等音訊物件信號之該響度資訊係
藉由降混頻以該僅兩個或兩個以上群組所構成之該等
音訊物件信號，以獲得包括一或多個降混頻音訊聲道之
一降混頻信號作為該經編碼音訊信號而進行，其中該一
或多個降混頻聲道之數目小於該兩個或兩個以上群組
包括之該等音訊物件信號之數目，

其中一或多個其他旁路音訊物件信號中之每一者
被指派至一第三群組，其中該一或多個其他旁路音訊物
件信號中之每一者不被該第一群組包括且不被該第二
群組包括，

其中編碼關於該等音訊物件信號之該響度資訊係
藉由不在該降混頻信號內降混該一或多個其他旁路音
訊物件信號而進行，且

其中編碼關於該等音訊物件信號之該響度資訊係
藉由判定該響度資訊之一第一響度值、一第二響度值及
一第三響度值而進行，該第一響度值指示該第一群組之

該一或多個音訊物件信號之一總響度，該第二響度值指示該第二群組之該一或多個音訊物件信號之一總響度，且該第三響度值指示該第三群組之該一或多個其他旁路音訊物件信號之一總響度，或經組配以判定該響度資訊之一第一響度值及一第二響度值，該第一響度值指示該第一群組之該一或多個音訊物件信號之一總響度，且該第二響度值指示該第二群組之該一或多個音訊物件信號及該第三群組之該一或多個其他旁路音訊物件信號之一總響度。

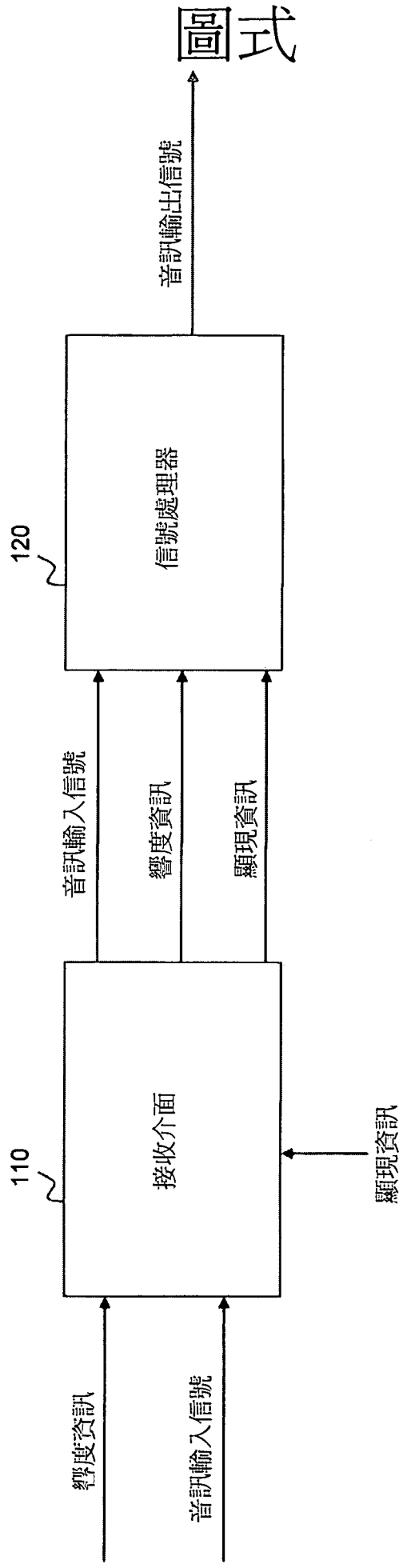
14. 如請求項 13 之方法，

其中該兩個或兩個以上群組為僅兩個群組，

其中該等音訊物件信號中之每一者被指派至該僅兩個群組中之僅一者，其中該僅兩個群組中之每一者包括該等音訊物件信號中的一或多者，

其中編碼關於該等音訊物件信號之該響度資訊係藉由降混頻由該僅兩個群組構成之該等音訊物件信號以獲得包括一或多個降混頻音訊聲道之一降混頻信號作為該經編碼音訊信號而進行，其中該一或多個向下混頻聲道之數目小於該僅兩個群組包括之該等音訊物件信號之數目。

15. 一種用於在一電腦或信號處理器上執行時實施如請求項 12 至 14 中任一項所述之方法的電腦程式。



圖式

圖 1

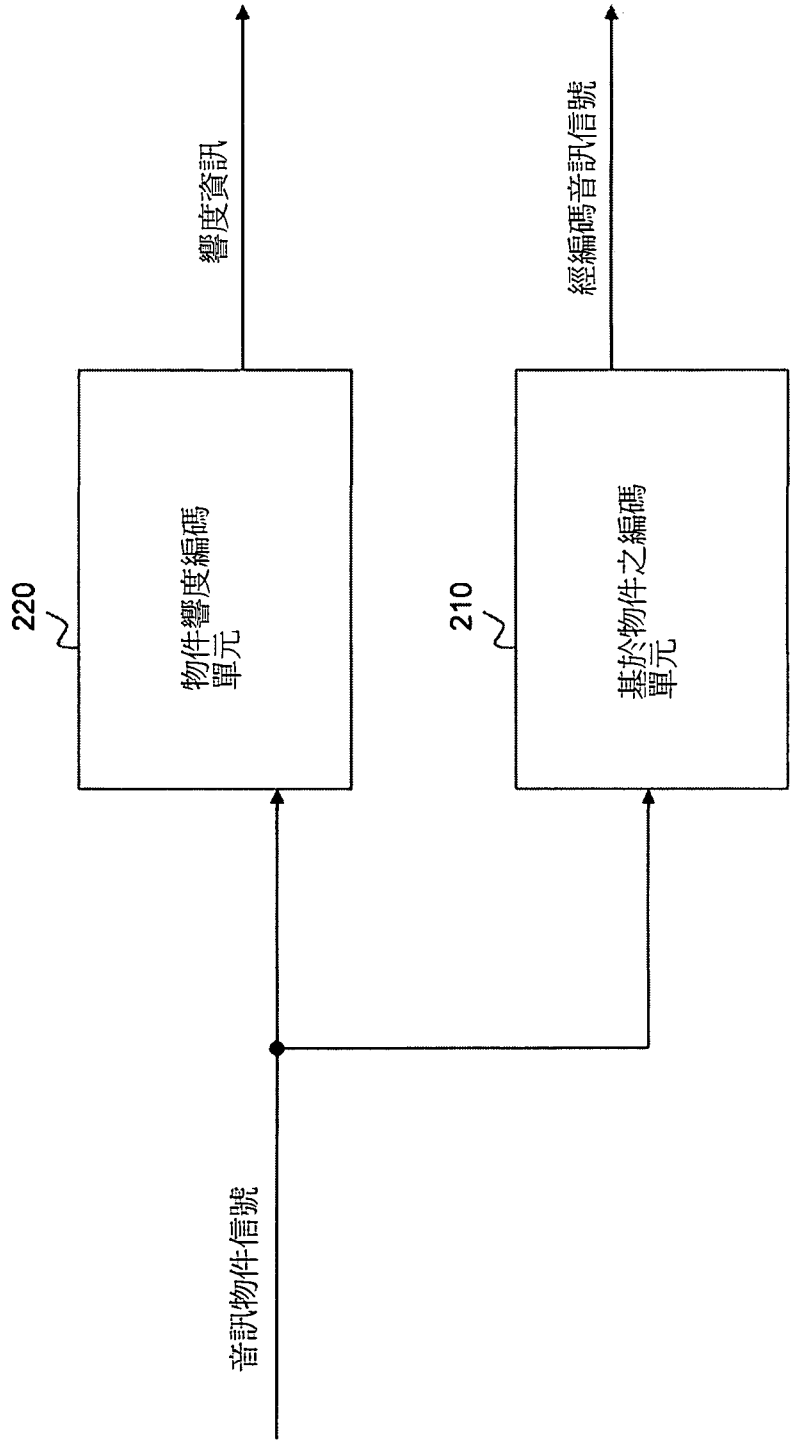


圖 2

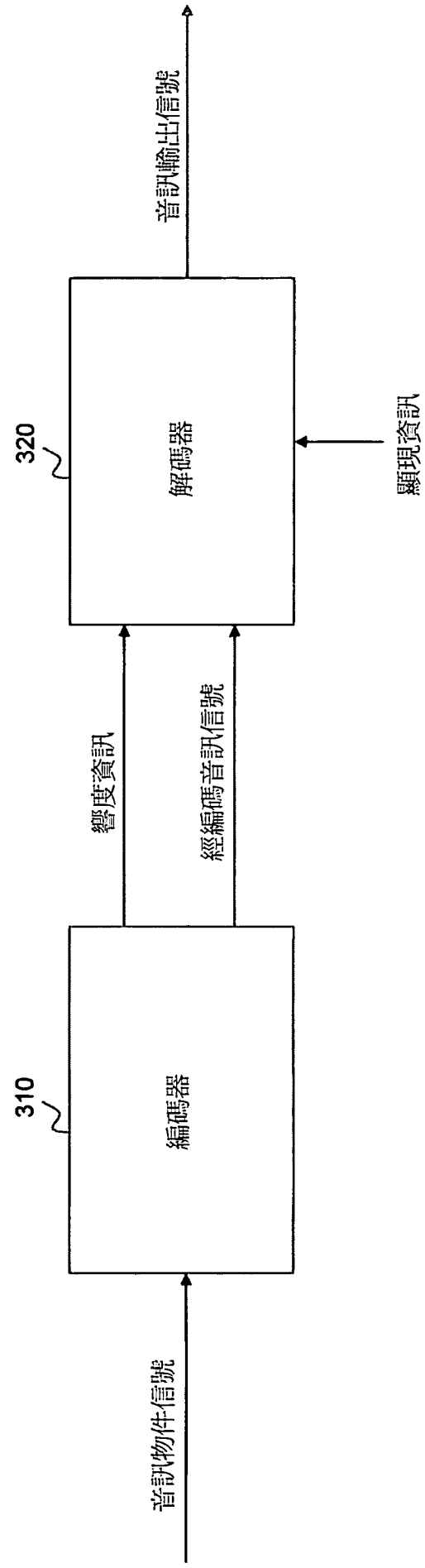


圖 3

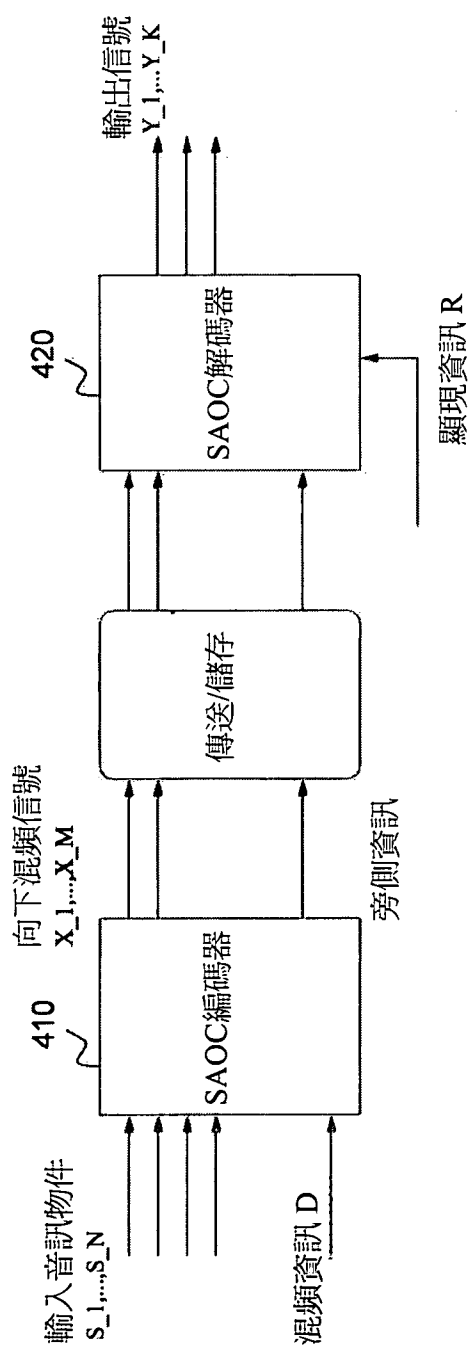


圖 4

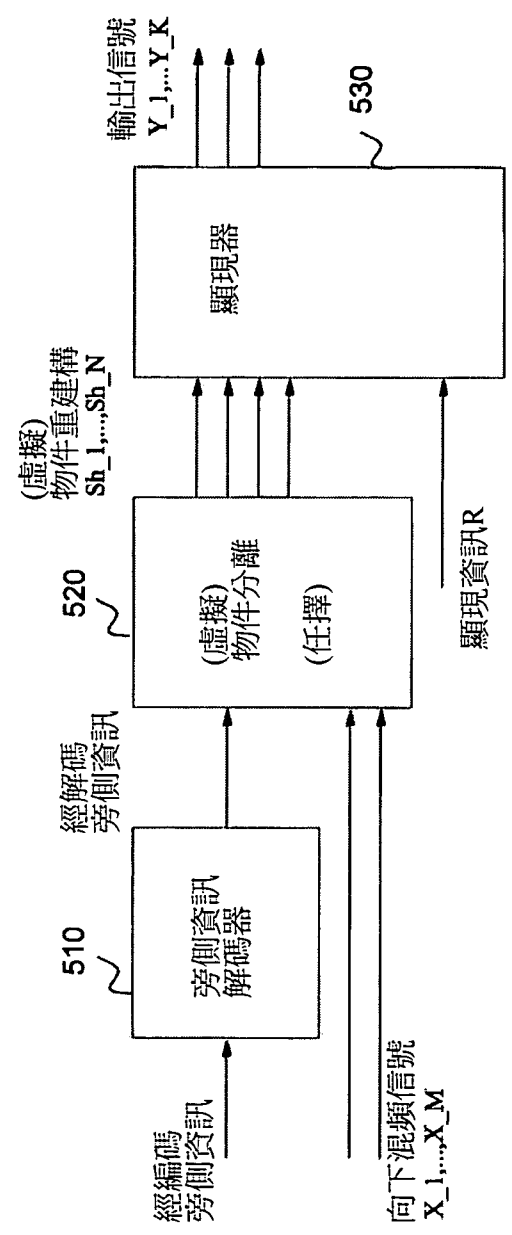


圖 5

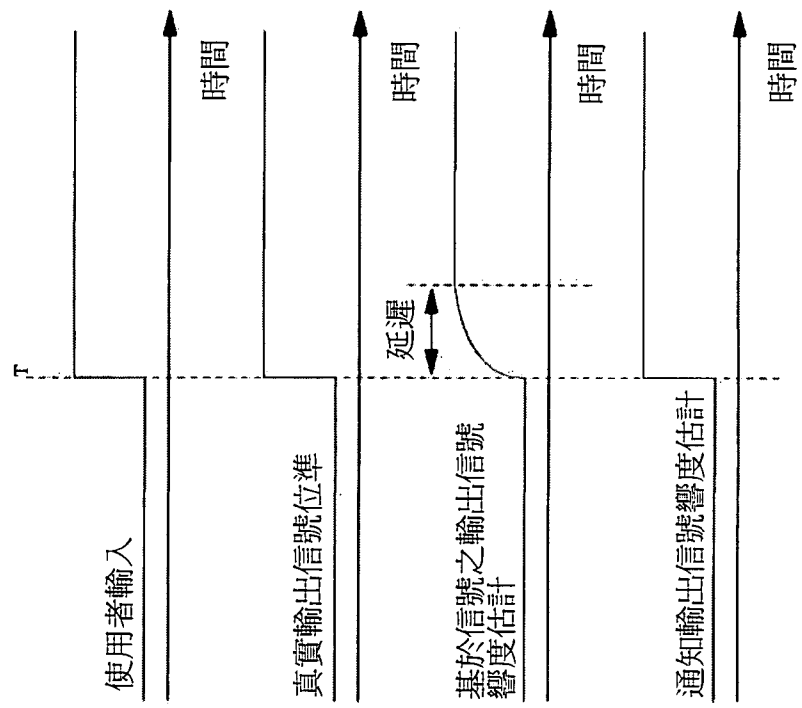


圖 6

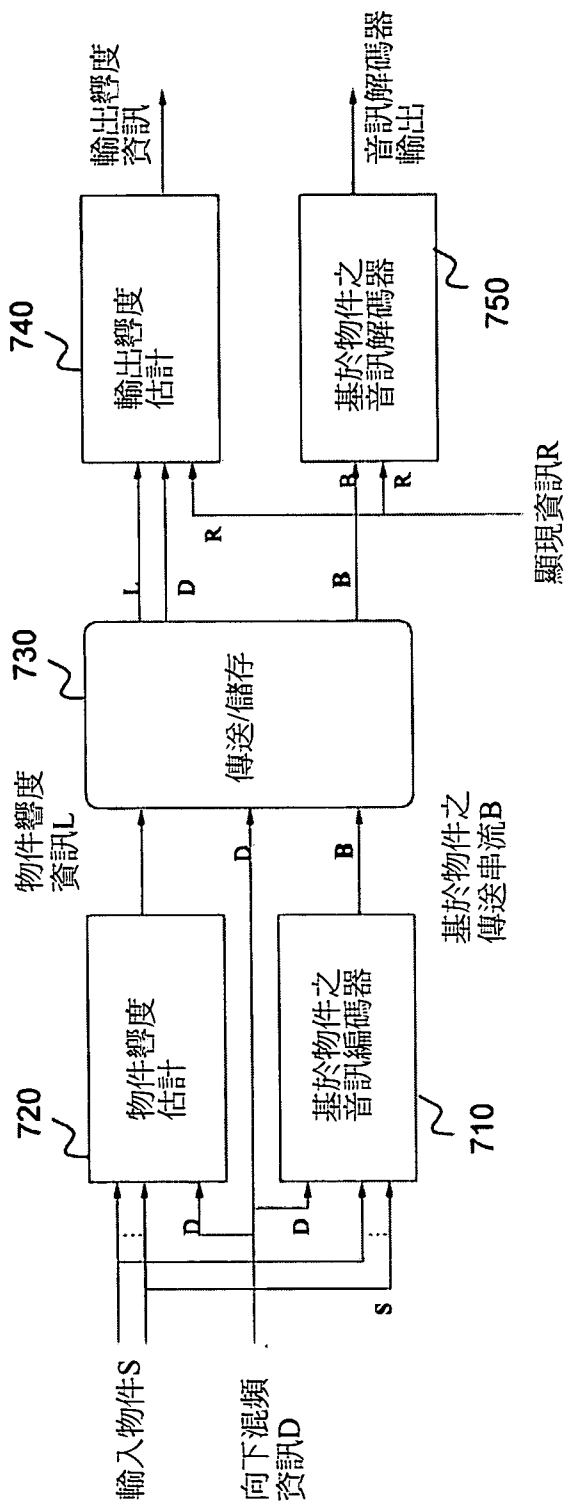


圖 7

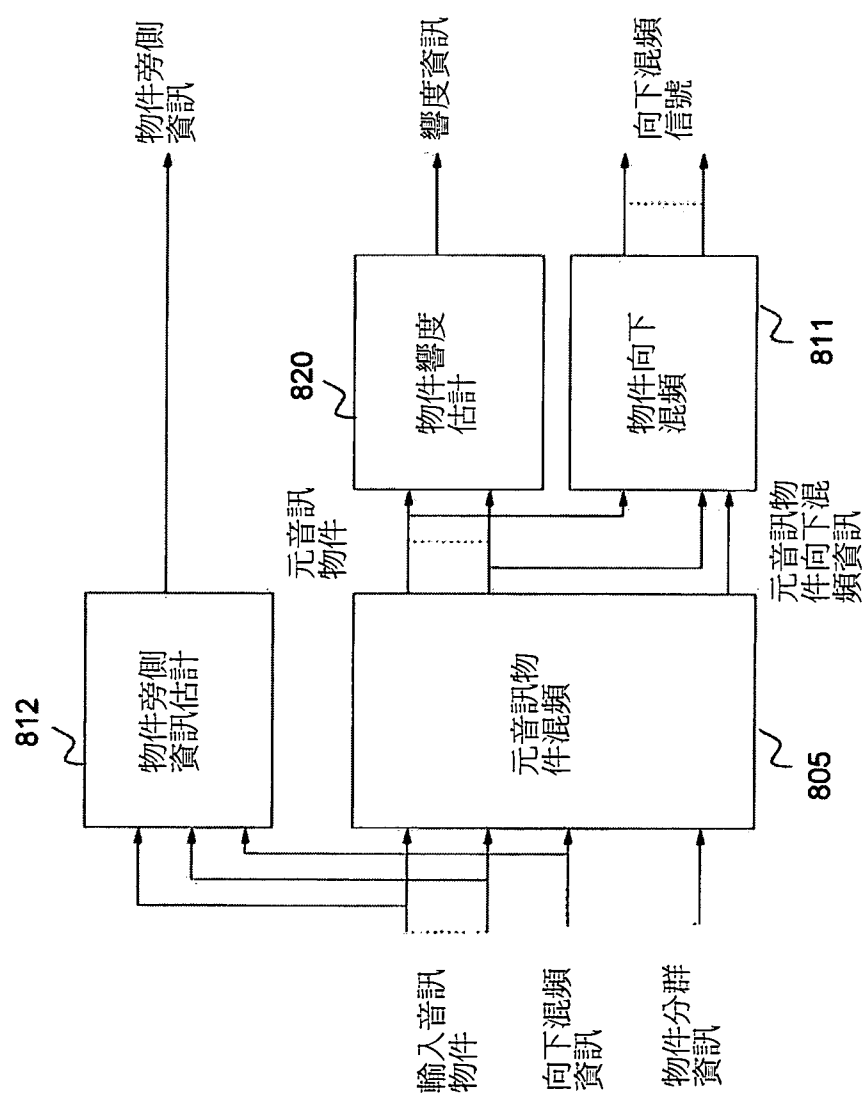


圖 8

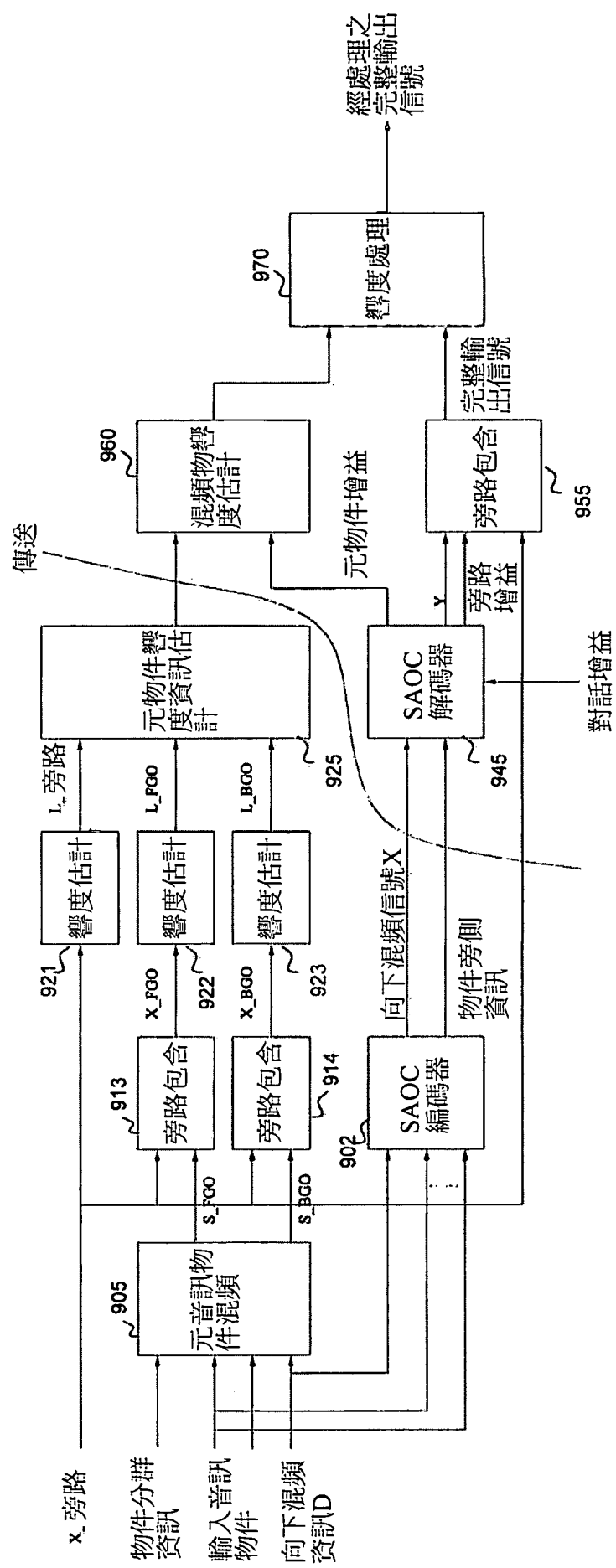


圖 9

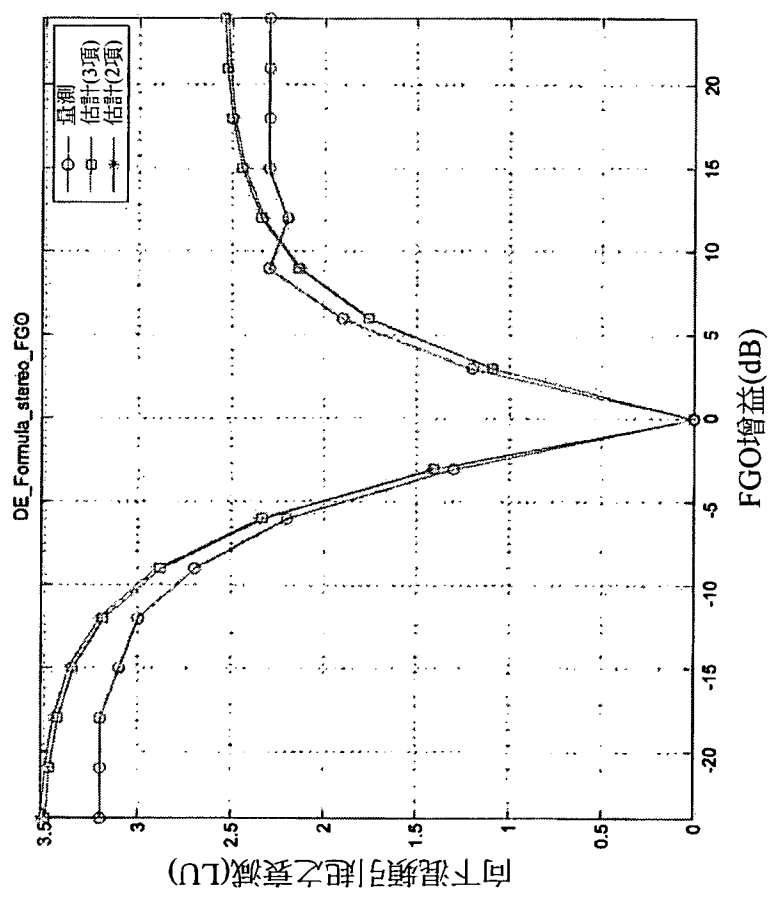


圖 10

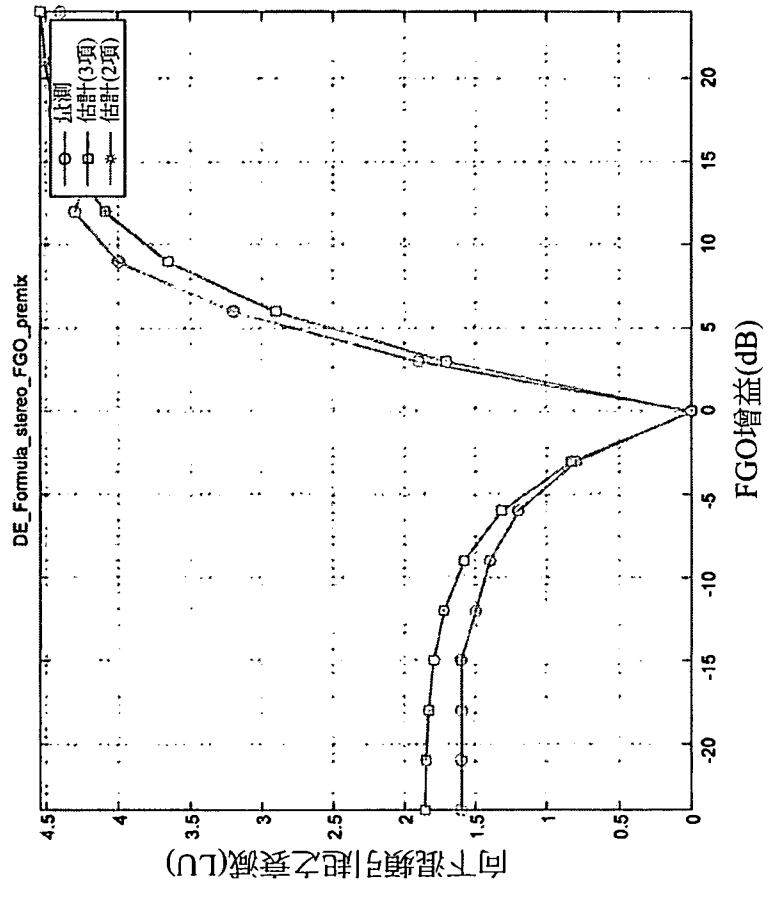


圖 11

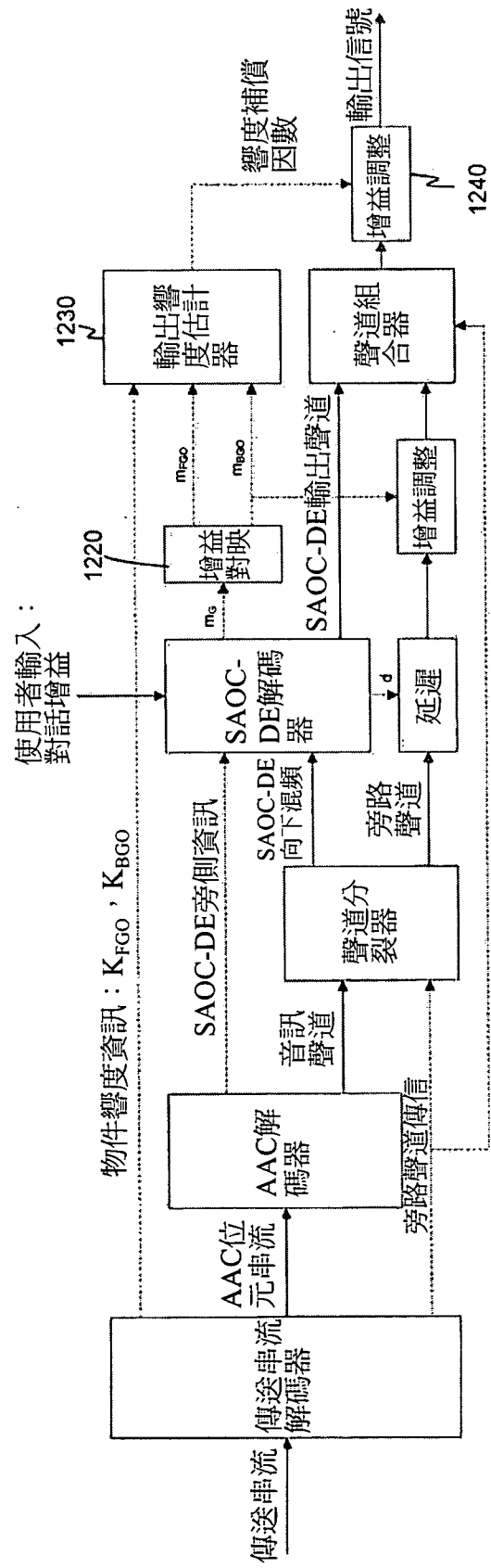


圖 12