



(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 I792006 B

(45)公告日：中華民國 112 (2023) 年 02 月 11 日

(21)申請案號：109120318

(22)申請日：中華民國 109 (2020) 年 06 月 15 日

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

(30)優先權：2019/06/14 歐洲專利局 19180385.7

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

德國

(72)發明人：布泰翁 亞歷山大 BOUTHEON, ALEXANDRE (FR)；福契斯 古拉米 FUCHS, GUILLAUME (FR)；穆爾特斯 馬庫斯 MULTRUS, MARKUS (DE)；庫奇 法比恩 KUECH, FABIAN (DE)；錫蓋特 奧莉薇 THIERGART, OLIVER (DE)；拜爾 史蒂芬 BAYER, STEFAN (AT)；迪斯曲 薩斯洽 DISCH, SASCHA (DE)；赫爾 汝根 HERRE, JUERGEN (DE)

(74)代理人：劉哲郎

(56)參考文獻：

TW I395204

TW I569260

TW 201423729A

TW 201521469A

CN 101411214A

US 2014/0321652A1

審查人員：施孝欣

申請專利範圍項數：45 項 圖式數：11 共 149 頁

(54)名稱

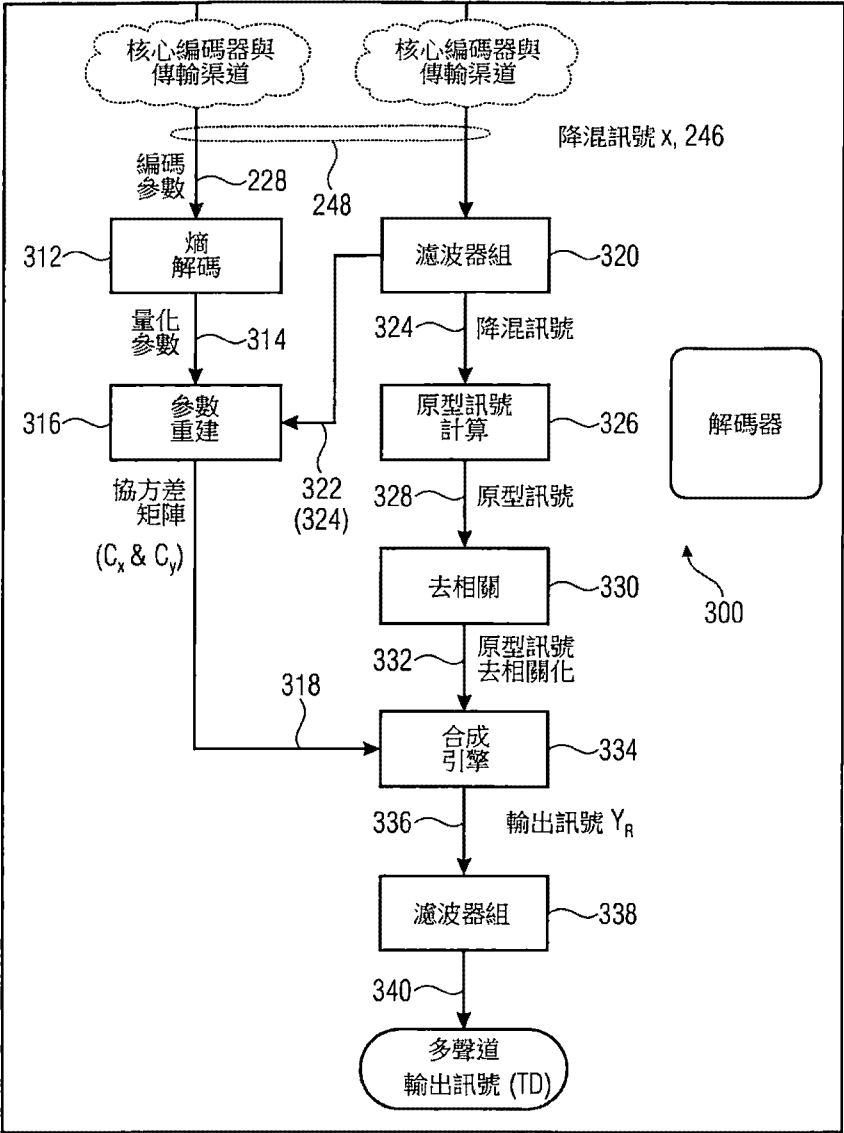
音訊合成器、訊號產生方法及儲存單元

(57)摘要

編碼及解碼技術的幾個示例被揭露。特別地，一種用於從一降混訊號(246、x)產生一合成訊號(336、340、 y_R)的音訊合成器(300)，包括：一輸入介面(312)，用於接收該降混訊號(246、x)，該降混訊號(246、x)具有一降混聲道數及旁側資訊(228)，該旁側資訊(228)包括一原始訊號(212、y)的聲道位準及相關資訊(314、 ξ 、 χ)，該原始訊號(212、y)具有一原始聲道數；及一合成處理器(404)，用於根據至少一個混合規則，使用以下內容產生該合成訊號(336、340、 y_R)：該原始訊號(212、y)的聲道位準及相關資訊(220、314、 ξ 、 χ)；及與該降混訊號(324、246、x)相關聯的協方差資訊(C_x)。

There are disclosed several examples of encoding and decoding technique. In particular, an audio synthesizer (300) for generating a synthesis signal (336, 340, y_R) from a downmix signal (246, x), comprises: an input interface (312) for receiving the downmix signal (246, x), the downmix signal (246, x) having a number of downmix channels and side information (228), the side information (228) including channel level and correlation information (314, ξ , χ) of an original signal (212, y), the original signal (212, y) having a number of original channels; and a synthesis processor (404) for generating, according to at least one mixing rule, the synthesis signal (336, 340, y_R) using: channel level and correlation information (220, 314, ξ , χ) of the original signal (212, y); and covariance information (C_x) associated with the downmix signal (324, 246, x).

指定代表圖：



解碼器模組的概觀

【第3b圖】

符號簡單說明：

- 228:旁側資訊
- 246:降混訊號
- 248:位元流
- 300:解碼器
- 312:熵解碼器
- 314:量化參數
- 316:參數重建模組
- 318:參數
- 320:濾波器組
- 322:降混訊號的一版本
- 324:降混訊號的頻域版本
- 326:原型訊號計算器
- 328:原型訊號
- 330:去相關模組
- 332:原型訊號
- 334:合成引擎
- 336:合成訊號
- 338:濾波器組
- 340:合成訊號
- C_x:協方差矩陣
- C_y:協方差矩陣
- Y_R:合成訊號
- x:降混訊號



公告本

I792006

【發明摘要】

【中文發明名稱】 音訊合成器、訊號產生方法及儲存單元

【英文發明名稱】 AUDIO SYNTHESIZER, SIGNAL GENERATION METHOD,
AND STORAGE UNIT

【中文】

編碼及解碼技術的幾個示例被揭露。特別地，一種用於從一降混訊號(246、 x)產生一合成訊號(336、340、 y_R)的音訊合成器(300)，包括：一輸入介面(312)，用於接收該降混訊號(246、 x)，該降混訊號(246、 x)具有一降混聲道數及旁側資訊(228)，該旁側資訊(228)包括一原始訊號(212、 y)的聲道位準及相關資訊(314、 ξ 、 χ)，該原始訊號(212、 y)具有一原始聲道數；及一合成處理器(404)，用於根據至少一個混合規則，使用以下內容產生該合成訊號(336、340、 y_R)：該原始訊號(212、 y)的聲道位準及相關資訊(220、314、 ξ 、 χ)；及與該降混訊號(324、246、 x)相關聯的協方差資訊(C_x)。

【英文】

There are disclosed several examples of encoding and decoding technique. In particular, an audio synthesizer (300) for generating a synthesis signal (336, 340, y_R) from a downmix signal (246, x), comprises: an input interface (312) for receiving the downmix signal (246, x), the downmix signal (246, x) having a number of downmix channels and side information (228), the side information (228) including channel level and correlation information (314, ξ , χ) of an original signal (212, y), the original signal (212, y) having a number of original channels; and a synthesis processor (404)

for generating, according to at least one mixing rule, the synthesis signal (336, 340, y_R) using: channel level and correlation information (220, 314, ξ , χ) of the original signal (212, y); and covariance information (C_x) associated with the downmix signal (324, 246, x).

【指定代表圖】 第 (3b) 圖。

【代表圖之符號簡單說明】

228:旁側資訊

246:降混訊號

248:位元流

300: 解碼器

312:熵解碼器

314:量化參數

316: 參數重建模組

318:參數

320:濾波器組

322:降混訊號的一版本

324:降混訊號的頻域版本

326:原型訊號計算器

328:原型訊號

330:去相關模組

332:原型訊號

334:合成引擎

336:合成訊號

338:濾波器組

340:合成訊號

C_x :協方差矩陣

C_y :協方差矩陣

Y_R :合成訊號

x :降混訊號

【發明說明書】

【中文發明名稱】 音訊合成器、訊號產生方法及儲存單元

【英文發明名稱】 AUDIO SYNTHESIZER, SIGNAL GENERATION METHOD,
AND STORAGE UNIT

【技術領域】

【0001】 1. 簡介

【0002】 在此，有編碼及解碼技術的幾個示例被公開。特別地，一種發明針對在低位元率編碼及解碼多聲道音訊內容，譬如使用DirAC框架。這種方法可以在使用低位元率的同時獲得一高品質的輸出。這可以被用於許多應用程式，包括藝術作品、通訊及虛擬實境。

【先前技術】

【0003】 1.1 先前技術

【0004】 本節簡要地描述先前技術。

【0005】 1.1.1 多聲道(Multichannel)內容的離散編碼(Discrete Coding)

【0006】 編碼及傳輸多聲道內容的最直接方法是直接量化及編碼多聲道音訊訊號的波形，而無需任何事先處理或假設。儘管該方法在理論上可以完美地工作，但存在一個主要缺點，即編碼該多聲道內容所需的位元消耗。因此，將被描述的其他方法(以及所提出的發明)是所謂的“參數方法”，因為它們使用元參數(meta-parameters)以描述及發送該多聲道音訊訊號而不是原始音訊多聲道訊號本身。

【0007】 1.1.2 MPEG環繞(MPEG Surround)

【0008】 MPEG環繞是在2006年被完成的ISO/MPEG標準，用於多聲道聲音的參數編碼 [1]。此方法主要依賴於兩參數集：

- 該聲道間同調度(Interchannel coherences，ICC)，它描述在一給定多聲道音訊訊號的每個聲道之間的同調度(coherence)。
- 該聲道位準差(Channel Level Difference，CLD)，對應於多聲道音訊訊號的兩個輸入聲道之間的位準差(level difference)。

【0009】 MPEG環繞的一種特殊性是使用所謂的“樹狀結構(tree-structures)”，這些結構允許“通過單個輸出聲道描述兩個輸入聲道(describe two inputs channels by means of a single output channels)”(引用自 [1])。

【0010】 作為一示例，以下可以找到使用MPEG環繞的一個5.1多聲道音訊訊號的編碼器方案。在此圖上，六個輸入聲道(在圖上被標記為“L”、“L_s”、“R”、“R_s”、“C”及“LFE”)通過一樹狀結構元件(在圖上被標記為“R_OTT”)被依次處理。這些樹狀結構元件中的每一個將產生一參數集如前面提到的數個ICC(ICC_s)及數個CLD(CLD_s)及一殘餘訊號(residual signal)，該殘餘訊號將通過另一樹狀結構被再次處理並產生另一參數集。一旦到達該樹的末端，先前被計算的不同參數被傳輸到該解碼器，像降混訊號一樣。這些元素由該解碼器使用以產生一輸出多聲道訊號，該解碼器處理基本上是由該編碼器使用的逆樹狀結構。

【0011】 MPEG環繞的主要優勢取決於此結構及在前面被提到的參數的使用。然而，MPEG環繞的缺點之一是由於該樹狀結構缺乏靈活性。同樣由於處理的特殊性，在某些特定項目上可能會發生品質惡化(quality degradation)。

【0012】 除了其他之外，參見第7圖，顯示從[1]被抽取的用於一個5.1訊號的一個MPEG環繞編碼器的一概觀。

【0013】 1.2 定向音訊編碼(Directional Audio Coding)

【0014】 定向音訊編碼(被縮寫為“DirAC”) [2] 也是一種再現空間音訊的參數方法，它是由在芬蘭的阿爾托(Aalto)大學的維爾·普爾基(Ville Pulkki)開發的。DirAC依靠一頻帶處理，該頻帶處理使用兩參數集來描述空間聲音：

- 該到達方向(DOA)，這是一個角度，以度為單位，描述在一音訊訊號中占主導地位的聲音(predominant sound)的到達方向。
- 擴散度(Diffuseness)，這是介於0與1之間的一值，用於描述該聲音有多“擴散(diffuse)”。如果該值為0，則該聲音是非擴散的，並且可以被同化為來自一精確角度的一點狀源；如果該值為1，則該聲音是完全擴散的，並且被假定為來自“每一個(every)”角度。

【0015】 為了合成該數個輸出訊號，DirAC假定其被分解為一擴散及非擴散部分，該擴散聲音合成旨在產生對一周圍聲音的感知，而直接聲音合成則旨在產生占主導地位的聲音。

【0016】 鑒於DirAC提供高品質的輸出，但它有一個主要缺點：它不適用於多聲道音訊訊號。因此，該DOA及擴散參數不太適合描述一多聲道音訊輸入，因此，輸出品質受到影響。

【0017】 1.3 雙耳提示編碼(Binaural Cue Coding)

【0018】 雙耳提示編碼(BCC) [3] 是由克裡斯托夫·法爾(Christof Faller)開發的一種參數化方法。此方法依賴於一類似的參數集如同那些被描述用於MPEG環繞(請參見 1.1.2)，即：

- 該聲道間位準差(Interchannel Level Difference, ICLD)，其是在該多聲道輸入訊號的兩個聲道之間的能量比的一測量(measure)。
- 該聲道間時差(ICTD)，其是在該多聲道輸入訊號的兩個聲道之間的該延遲的一測量。
- 該聲道間相關聯(ICC)，其是在該多聲道輸入訊號的兩個聲道之間的關聯的一測量。

【0019】 與稍後將被描述的新穎發明相比，該BCC方法就發送的參數的計算而言具有非常相似的特性，但是它缺乏被發送的參數的靈活性及可縮放性。

【0020】 1.4 MPEG空間音訊對象編碼(Spatial Audio Object Coding)

【0021】 空間音訊對象編碼 [4] 將在此被簡單提及。這是用於對所謂的音訊對象進行編碼的MPEG標準，這在一定程度上與多聲道訊號有關。它使用與MPEG環繞類似的諸多參數。

【發明內容】

【0022】 1.5 先前技術的誘因/缺點

【0023】 1.5.1 誘因

【0024】 1.5.1.1 使用DirAC框架(framework)

【0025】 本發明必須被提到的一個方面是當前發明必須適合於該DirAC框架。儘管如此，之前也提到過DirAC的參數不適用於一多聲道音訊訊號。有關此主題應給予更多解釋。

【0026】 該原始的DirAC處理使用麥克風訊號或歧義訊號(ambisonics signals)。從這些訊號，計算諸多參數，即到達方向(DOA)及擴散度。

【0027】 為了將DirAC與多聲道音訊訊號一起使用，被嘗試的第一種方法是使用一種由維爾·普爾基(Ville Pulkki)提出的方法，將該多聲道訊號轉換為歧義內容，如在 [5] 所述。然後，一旦這些歧義訊號從該多聲道音訊訊號中被導出，就可以使用DOA及擴散進行常規的DirAC處理。首次嘗試的結果是被輸出的多聲道訊號的品質及空間特徵惡化，且無法滿足目標應用程式的要求。

【0028】 因此，此新穎發明背後的主要動機是使用一參數集，該參數集有效地描述該多聲道訊號，並且還使用該DirAC框架，進一步的解釋將在1.1.2節中給出。

【0029】 1.5.1.2 提供在低位元率運作的一系統

【0030】 本發明的目標及目的之一是提出一種允許低位元率應用的方法。這需要找到最佳資料集以描述在編碼器與解碼器之間的多聲道內容。這還需要就傳輸參數的數量及輸出品質而言找到最佳的權衡。

【0031】 1.5.1.3 提供一靈活的系統

【0032】 本發明的另一個重要目標是提出一種靈活的系統，該系統可以接受旨在任何揚聲器設置上被再現的任何多聲道音訊格式。取決於輸入設置，輸出品質不應受到損害。

【0033】 1.5.2 先前技術的缺點

【0034】 在前面提到的先前技術的幾個缺點在下表中被列出。

缺點	先前技術關切的	評論
不合適的位元率	多聲道內容的離散編碼	對於我們的要求及目標應用而言，多聲道內容的直接編碼導致位元率太高。

不適當的參數/描述符	舊版 DirAC	舊版 DirAC 方法使用擴散及 DOA 作為描述參數，事實證明這些參數不太適合描述一多聲道音訊訊號
該方法缺乏靈活性	MPEG 環繞 BCC	MPEG 環繞及 BCC 對於目標應用程式的要求不夠靈活

【0035】 2. 發明敘述

【0036】 2.1 發明內容

【0037】 根據一個方面，提供一種音訊合成器(編碼器)，用於從一降混訊號產生一合成訊號，該合成訊號具有一合成聲道數，該音訊合成器包括：

一輸入介面，被配置用於接收該降混訊號，該降混訊號具有一降混聲道數及旁側資訊，該旁側資訊包括一原始訊號的聲道位準及相關資訊，該原始訊號具有一原始聲道數；及

一合成處理器，被配置用於根據至少一個混合規則使用以下內容產生該合成訊號：

該原始訊號的聲道位準及相關資訊；及

與該降混訊號相關聯的協方差資訊。

【0038】 該音訊合成器可以包括：

一原型訊號計算器，被配置用於從該降混訊號計算一原型訊號，該原型訊號具有該合成聲道數；

一混合規則計算器，被配置用於使用以下內容計算至少一個混合規則：

該原始訊號的該聲道位準及相關資訊；及

與該降混訊號相關聯的該協方差資訊；

其中該合成處理器被配置用於使用該原型訊號及該至少一個混合規則產生該合成訊號。

【0039】該音訊合成器可以被配置成重建該原始訊號的一目標協方差資訊。

【0040】該音訊合成器可以被配置成重建適應於該合成訊號的該聲道數的該目標協方差資訊。

【0041】該音訊合成器可以被配置成通過將數個原始聲道群組指派給數個單一合成聲道，以重建適應於該合成訊號的該聲道數的該協方差資訊，或者反之亦然，以便該重建目標協方差資訊被通報給該合成訊號的該聲道數。

【0042】該音訊合成器可以被配置成通過產生針對該些原始聲道數的該目標協方差資訊並且後續應用一降混規則或一升混規則以及一能量補償，以得出針對該數個合成聲道的該目標協方差，以重建適應於該合成訊號的該聲道數的該協方差資訊。

【0043】該音訊合成器可以被配置成基於該原始協方差資訊的一估計版本重建該協方差資訊的該目標版本，其中該原始協方差資訊的該估計版本被通報給該合成聲道數或該原始聲道數。

【0044】該音訊合成器可以被配置成從與該降混訊號相關聯的協方差資訊獲得該原始協方差資訊的該估計版本。

【0045】該音訊合成器可以被配置成通過將一估算規則應用於與該降混訊號相關聯的該協方差資訊，該估算規則是或被關聯到用於計算該原型訊號的一原型規則，以獲得該原始協方差資訊的該估計版本。

【0046】該音訊合成器可以被配置成針對至少一個聲道對，將該原始協方差資訊(C_y)的該估計版本($\widehat{C}_y$)正規化為該聲道對中的該數個聲道的該數個位準的該數個平方根。

【0047】該音訊合成器可以被配置成以該原始協方差資訊的正規化估計版本理解一矩陣。

【0048】該音訊合成器可以被配置成通過插入在該位元流的該旁側資訊中所獲得的數個元來完成該矩陣。

【0049】該音訊合成器可以被配置成通過按形成該聲道對的該數個聲道的該數個位準的該平方根來縮放該原始協方差資訊的該估計版本，將該矩陣進行去正規化。

【0050】該音訊合成器可以被配置成在該降混訊號的該旁側資訊之中進行檢索，該音訊合成器還被配置成由來自以下兩者的該原始聲道位準及相關資訊的一估計版本重建該協方差資訊的該目標版本：

用於至少一個第一聲道或聲道對的協方差資訊；及

用於至少一個第二聲道或聲道對的聲道位準及相關資訊。

【0051】該音訊合成器可以被配置成偏好該聲道位準及相關資訊描述從該位元流的該旁側資訊所獲得的該聲道或聲道對，而不是針對相同聲道或聲道對從該降混訊號被重建的該協方差資訊。

【0052】該原始協方差資訊的該重建目標版本可被理解為描述在一聲道對之間的一能量關係至少部分地是基於被關聯到該聲道對中的每個聲道的數個位準。

【0053】該音訊合成器可以被配置成獲得該降混訊號的一頻域版本，該降混訊號的該頻域版本被劃分為數個頻帶或數個頻帶群組，其中不同聲道位準及相關資訊與不同頻帶或頻帶群組相關聯，

其中該音訊合成器被配置成針對不同頻帶或頻帶群組進行不同操作，以獲得針對不同頻帶或頻帶群組的不同混合規則。

【0054】該降混訊號被劃分為數個時隙，其中不同的聲道位準及相關資訊與不同時隙相關聯，並且該音訊合成器被配置成針對不同時隙進行不同操作，以獲得針對不同時隙的不同混合規則。

【0055】該降混訊號被劃分為數個訊框，並且每個訊框被劃分為數個時隙，其中當在一個訊框中的暫態的存在及位置被發訊表明(signaled)為在一個暫態時隙中，該音訊合成器被配置成：

將該當前的聲道位準及相關資訊與該暫態時隙及/或該訊框的暫態時隙後續的數個時隙相關聯；及

將該暫態時隙以前的該訊框的時隙與該以前的時隙的該聲道位準及相關資訊相關聯。

【0056】該音訊合成器可以被配置成選擇一原型規則，該原型規則被配置用於在該合成聲道數的基礎上計算一原型訊號。

【0057】該音訊合成器可以被配置成在數個預存原型規則之中選擇該原型規則。

【0058】該音訊合成器可以被配置成在一手動選擇的基礎上定義一原型規則。

【0059】該原型規則可以基於或包括一矩陣，該矩陣具備一第一維度及一第二維度，其中該第一維度與該降混聲道數相關聯，並且該第二維度與該合成聲道數相關聯。

【0060】該音訊合成器可以被配置成操作在等於或低於160千位元/秒的一位元率。

【0061】該音訊合成器還可以包括一熵解碼器，用於獲得具備該旁側資訊的該降混訊號。

【0062】該音訊合成器還包括一去相關模組，以減少在不同聲道之間的相關量。

【0063】該原型訊號可以被直接提供給該合成處理器，沒有進行去相關。

【0064】該原始訊號的該聲道位準及相關資訊、該至少一個混合規則及與該降混訊號相關聯的該協方差資訊中的至少一者為一矩陣形式。

【0065】該旁側資訊包括該數個原始聲道的一標識；

【0066】其中該音訊合成器還可以被配置用於使用該原始訊號的該聲道位準及相關資訊、與該降混訊號相關聯的一協方差資訊、該數個原始聲道的該標識，及該數個合成聲道的一標識中的至少一者來計算該至少一個混合規則。

【0067】該音訊合成器可以被配置成通過奇異值分解來計算至少一個混合規則。

【0068】該降混訊號可以被劃分為數個訊框，該音訊合成器被配置成使用針對一在前的訊框所獲得的具備一參數的一線性組合、一被估計或被重建的值或一混合矩陣來平滑一被接收的參數、一被估計或被重建的值或一混合矩陣。

【0069】 該音訊合成器可以被配置成當在一個訊框中的一暫態的存在及/或位置被發訊表明時，停用該被接收的參數、該被估計或被重建的值或該混合矩陣的平滑。

【0070】 該降混訊號可以被劃分為數個訊框，並且該數個訊框被劃分為數個時隙，其中該原始訊號的該聲道位準及相關資訊是以一逐訊框的方式從該位元流的該旁側資訊所獲得，該音訊合成器被配置成針對一當前的訊框使用一混合矩陣(或混合規則)，通過按沿著該當前的訊框的該數個後續時隙增加的一係數針對現在的訊框所計算而縮放該混合矩陣(或混合規則)，及通過將被用於該先前的訊框的該混合矩陣(或混合規則)添加在按沿著該當前的訊框的該數個後續時隙的一減少係數被縮放的一版本中，來獲得該混合規則。

【0071】 該合成聲道數可以大於該原始聲道數。該合成聲道數可以小於該原始聲道數。該合成聲道數及該原始聲道數可以大於該降混聲道數。

【0072】 該合成聲道數、該原始聲道數及該降混聲道數中的至少一個或全部為一複數(a plural number)。

【0073】 該至少一個混合規則可以包括一第一混合矩陣及一第二混合矩陣，該音訊合成器包括：

一第一路徑，包括：

一第一混合矩陣塊，被配置用於根據從以下內容計算出的該第一混合矩陣來合成該合成訊號的一第一分量：

與該合成訊號相關聯的一協方差矩陣，該協方差矩陣是從該聲道位準及相關資訊被重建；及

與該降混訊號相關聯的一協方差矩陣，

一第二路徑，用於合成該合成訊號的一第二分量，該第二分量是一殘餘分量，該第二路徑包括：

一原型訊號塊，被配置用於將該降混訊號從該降混聲道數升混到該合成聲道數；

一去相關器，被配置用於將該被升混的原型訊號進行去相關；

一第二混合矩陣塊，被配置用於根據來自該降混訊號的該去相關版本的一第二混合矩陣來合成該合成訊號的該第二分量，該第二混合矩陣為一殘餘混合矩陣，

其中該音訊合成器被配置成從以下內容估計該第二混合矩陣：

由該第一混合矩陣塊提供的一殘餘協方差矩陣；及

從與該降混訊號相關聯的該協方差矩陣獲得的該數個去相關原型訊號的該協方差矩陣的一估計，

其中該音訊合成器還包括一加法器塊，用於將該合成訊號的該第一分量與該合成訊號的該第二分量進行求和。

【0074】根據一個方面，提供一種音訊合成器，用於從具有一降混聲道數的一降混訊號產生一合成訊號，該合成訊號具有一合成聲道數，該降混訊號是具有一原始聲道數的一原始訊號的一降混版本，該音訊合成器包括：

一第一路徑，包括：

一第一混合矩陣塊，被配置用於根據從以下計算出的一第一混合矩陣以合成該合成訊號的一第一分量：

被關聯到該合成訊號的一協方差矩陣；及

被關聯到該降混訊號的一協方差矩陣；

一第二路徑，用於合成該合成訊號的一第二分量，其中該第二分量是一殘餘分量，該第二路徑包括：

一原型訊號塊，被配置用於將該降混訊號從該降混聲道數升混到該合成聲道數；

一去相關器，被配置用於對該被升混的原型訊號(613c)進行去相關；

一第二混合矩陣塊，被配置用於根據來自該降混訊號的該去相關版本的一第二混合矩陣以合成該合成訊號的該第二分量，該第二混合矩陣是一殘餘混合矩陣，

其中該音訊合成器被配置成從以下內容計算該第二混合矩陣：

由該第一混合矩陣塊提供的該殘餘協方差矩陣；及

從被關聯到該降混訊號的該協方差矩陣獲得的該數個去相關的原型訊號的該協方差矩陣的一估計，

其中該音訊合成器還包括一加法器塊，用於將該合成訊號的該第一分量與該合成訊號的該第二分量進行求和。

【0075】 通過從被關聯到該合成訊號的該協方差矩陣減去通過將該第一混合矩陣應用於被關聯到該降混訊號的該協方差矩陣所獲得的一矩陣，來獲得該殘餘協方差矩陣。

【0076】 該音訊合成器可以被配置成從以下內容定義該第二混合矩陣：

一第二矩陣，其通過分解被關聯到該合成訊號的該剩餘協方差矩陣而被獲得；

一第一矩陣，其是從該數個去相關的原型訊號的該協方差矩陣的該估計被獲得的一對角矩陣的逆矩陣或正則化逆矩陣。

【0077】 可以通過將該平方根函數應用於該數個去相關的原型訊號的該協方差矩陣的數個主對角元素，來獲得該對角矩陣。

【0078】 可以通過將奇異值分解應用於被關聯到該合成訊號的該殘餘協方差矩陣，來獲得該第二矩陣。

【0079】 該音訊合成器可以被配置成通過將該第二矩陣與從該數個去相關的原型訊號的該協方差矩陣的該估計及一第三矩陣所獲得的該對角矩陣的逆矩陣或正則化逆矩陣進行相乘，來定義該第二混合矩陣。

【0080】 該音訊合成器可以被配置成通過將奇異值分解應用於從該數個去相關的原型訊號的該協方差矩陣的一正規化(normalized)版本所獲得的一矩陣，其中該正規化是對該殘餘協方差矩陣及該對角矩陣及該第二矩陣的主對角線進行，來獲得該第三矩陣。

【0081】 該音訊合成器可以被配置成從一第二矩陣及該第二矩陣的逆矩陣或正則化逆矩陣來定義該第一混合矩陣，

其中通過分解被關聯到該降混訊號的該協方差矩陣來獲得該第二矩陣，及通過分解被關聯到該降混訊號的該重建目標協方差矩陣來獲得該第二矩陣。

【0082】 該音訊合成器可以被配置成從應用於被關聯到該降混訊號的該協方差矩陣所獲得的該矩陣的該數個對角元估計該數個去相關的原型訊號的該協方差矩陣，在該原型塊處被使用的該原型規則用於將該降混訊號從該降混聲道數升混到該合成聲道數。

【0083】 該數個頻帶被彼此聚合為數個聚合頻帶群組，其中關於該數個聚合頻帶群組的資訊被提供在該位元流的旁側資訊中，其中該原始訊號的該聲道

位準及相關資訊按每頻帶群組被提供，以便針對相同聚合頻帶群組的不同頻帶計算相同的至少一個混合矩陣。

【0084】 根據一個方面，提供一種音訊編碼器，用於從一原始訊號產生一降混訊號，該原始訊號具有數個原始聲道，該降混訊號具有一降混聲道數，該音訊編碼器包括：

一參數估計器，被配置成估計該原始訊號的聲道位準及相關資訊，及一位元流寫入器，用於將該降混訊號編碼成一位元流，使得該降混訊號被編碼在該位元流中，以便具有旁側資訊，該旁側資訊包括該原始訊號的聲道位準及相關資訊。

【0085】 該音訊編碼器可以被配置成提供該原始訊號的該聲道位準及相關資訊作為數個正規化值。

【0086】 被編碼在該旁側資訊中的該原始訊號的該聲道位準及相關資訊至少表示被關聯到該數個原始聲道的總數的聲道位準資訊。

【0087】 被編碼在該旁側資訊中的該原始訊號的該聲道位準及相關資訊至少表示相關資訊，該相關資訊描述在至少一個不同原始聲道對之間的數個能量關係，但小於該數個原始聲道的總數。

【0088】 該原始訊號的該聲道位準及相關資訊包括至少一個同調度值，該同調度值描述在一原始聲道對中的兩個聲道之間的同調度。

【0089】 該同調度值可以被正規化。該同調度值可以是

$$\xi_{i,j} = \frac{C_{y_{i,j}}}{\sqrt{C_{y_{i,i}} \cdot C_{y_{j,j}}}}$$

其中 $C_{y_{i,j}}$ 是在該數個聲道 i 與 j 之間的一協方差， $C_{y_{i,i}}$ 與 $C_{y_{j,j}}$ 分別是被關聯到該

第 15 頁，共 119 頁(發明說明書)

數個聲道*i*與*j*的數個位準。

【0090】 該原始訊號的該聲道位準及相關資訊包括至少一個聲道間位準差(ICLD)。

【0091】 該至少一個ICLD可以被提供作為一對數值。該至少一個ICLD可以是

$$\chi_i = 10 \cdot \log_{10} \left(\frac{P_i}{P_{\text{dmx},i}} \right)$$

其中

χ_i 是針對聲道*i*的該聲道間位準差，

P_i 是當前聲道*i*的該功率，

$P_{\text{dmx},i}$ 是該降混訊號的該協方差資訊的該數個值的一線性組合。

【0092】 該音訊編碼器可以被配置成在狀態資訊的基礎上選擇是否編碼或不編碼該原始訊號的該聲道位準及相關資訊的至少一部分，以便在有效載荷相對較低的情況下，在該旁側資訊中包括一增加數量的聲道位準及相關資訊。

【0093】 該音訊編碼器可以被配置成在關於該數個聲道的衡量指標的基礎上選擇該原始訊號的該聲道位準及相關資訊的哪一部分要被編碼在該旁側資訊中，以便在該旁側資訊中包括被關聯到更敏感的衡量指標的聲道位準及相關資訊。

【0094】 該原始訊號的該聲道位準及相關資訊可以為一矩陣的數個元的形式。

【0095】該矩陣可以是一對稱矩陣或一厄米特矩陣，其中該聲道位準及相關資訊的數個元被提供針對該矩陣的在對角線中的全部或少於總數的該數個元及/或針對少於該矩陣的該數個非對角元的一半。

【0096】該位元流寫入器被配置成將至少一個聲道的標識進行編碼。

【0097】該原始訊號或其一處理版本可以被劃分為數個相等時間長度的後續訊框。

【0098】該音訊編碼器可以被配置成將針對每個訊框特定的該原始訊號的聲道位準及相關資訊編碼在該旁側資訊中。

【0099】該音訊編碼器可以被配置成將被共同關聯於數個連續訊框的該原始訊號的相同聲道位準及相關資訊編碼在該旁側資訊中。

【0100】該音訊編碼器可以被配置成選擇一連續訊框數，以使該原始訊號的相同聲道位準及相關資訊被選擇，使得：

一相對較高的位元率或較高的有效載荷蘊含該連續訊框數的一增加，以使與該原始訊號的相同聲道位準及相關資訊相關聯，反之亦然。

【0101】該音訊編碼器可以被配置成減少該連續訊框數，以使該原始訊號的相同聲道位準及相關資訊與一暫態的檢測相關聯。

【0102】每個訊框可以被細分為一整數個的連續時隙(an integer number of consecutive slots)。

【0103】該音訊編碼器可以被配置成針對每個時隙估計該聲道位準及相關資訊，並且在該旁側資訊中編碼針對不同時隙被估計的該聲道位準及相關資訊的總和或平均值或另一預定線性組合。

【0104】 該音訊編碼器可以被配置成對該訊框的時域版本進行一暫態分析，以決定在該訊框內的一暫態的發生。

【0105】 該音訊編碼器可以被配置成決定該暫態已經發生在該訊框的哪個時隙中，及：

將被關聯到該暫態已經發生的該時隙及/或在該訊框中的後續時隙的該原始訊號的該聲道位準及相關資訊進行編碼，

沒有將被關聯到在該暫態以前的該數個時隙的該原始訊號的該聲道位準及相關資訊進行編碼。

【0106】 該音訊編碼器可以被配置成在該旁側資訊中以訊號表明該暫態發生在該訊框的一個時隙中。

【0107】 該音訊編碼器可以被配置成在該旁側資訊中以訊號表明該暫態已經發生在該訊框的哪個時隙中。

【0108】 該音訊編碼器可以被配置成估計被關聯到該訊框的多個時隙的該原始訊號的聲道位準及相關資訊，並對它們求和或對它們取平均或將它們線性地組合，以獲得被關聯到該訊框的聲道位準及相關資訊。

【0109】 該原始訊號可以被轉換為一頻域訊號，其中該音訊編碼器被配置成將該原始訊號的該聲道位準及相關資訊以一逐頻帶的方式編碼在該旁側資訊中。

【0110】 該音訊編碼器可以被配置成將該原始訊號的一頻帶數聚合成一更為減少的頻帶數(a more reduced number of bands)，以便將該原始訊號的該聲道位準及相關資訊以一逐聚合頻帶的方式編碼在該旁側資訊中。

【0111】 該音訊編碼器可以被配置在檢測到該訊框中的一暫態的情況下進一步聚合該數個頻帶，使得：

該頻帶數被減少；及/或

至少一個頻帶的寬度通過與另一頻帶聚合而被增加。

【0112】 該音訊編碼器還可以被配置成在該位元流中編碼一個頻帶的至少一個聲道位準及相關資訊作為相對於一先前被編碼的聲道位準及相關資訊的一增量。

【0113】 該音訊編碼器可以被配置成相對於由該估計器估計的該聲道位準及相關資訊在該位元流的該旁側資訊中編碼該聲道位準及相關資訊的一不完整版本。

【0114】 該音訊編碼器可以被配置成在由該估計器估計的整體的聲道位準及相關資訊之中適應性地選擇要被編碼在該位元流該旁側資訊中的被選擇資訊，使得由該估計器估計的聲道位準及/或相關資訊的剩餘未被選擇資訊不被編碼。

【0115】 該音訊編碼器可以被配置成從被選擇的聲道位準及相關資訊重建該聲道位準及相關資訊，從而在該解碼器處模擬未被選擇的聲道位準及相關資訊的估計，並且計算在以下內容之間的錯誤資訊：

由該編碼器估計的該未被選擇的聲道位準及相關資訊；及

通過在該解碼器處模擬未被選擇的聲道位準及相關資訊的估計而被重建的

該未被選擇的聲道位準及相關資訊；及

以便在該被計算的錯誤資訊的基礎上進行區分：

可適當重建的聲道位準及相關資訊；與

不可適當重建的聲道位準及相關資訊，

以便決定：

選擇在該位元流的該旁側資訊中要被編碼的該不可適當重建的聲道位準及相關資訊；及

不選擇該可適當重建的聲道位準及相關資訊，從而避免在該位元流的該旁側資訊中編碼該可適當重建的聲道位準及相關資訊。

【0116】 該聲道位準及相關資訊可以根據一預定順序被索引，其中該編碼器被配置成在該位元流的該旁側資訊中以訊號表明被關聯到該預定排序的數個索引，該數個索引指示該聲道位準及相關資訊中的哪一個被編碼。該數個索引通過一位元映像被提供。該數個索引根據將一個一維索引關聯於一矩陣的數個元的一組合編號系統而被定義。

【0117】 該音訊編碼器可以被配置成在以下內容中進行一選擇：

該聲道位準及相關資訊的一適應條款，在該適應條款中，被關聯到該預定順序的數個索引被編碼在該位元流的該旁側資訊中；及

該聲道位準及相關資訊的一固定條款，使得該被編碼的聲道位準及相關資訊是預定的，並且根據一預定的固定順序被排序，沒有一索引條款。

【0118】 該音訊編碼器可以被配置成在該位元流的該旁側資訊中以訊號表明該聲道位準及相關資訊是否根據該適應條款或根據該固定條款被提供。

【0119】 該音訊編碼器還可以被配置成在該位元流中編碼當前的聲道位準及相關資訊作為相對於該先前的聲道位準及相關資訊的增量。

【0120】 該音訊編碼器還可以被配置成根據一靜態降混產生該降混訊號。

【0121】 根據一個方面，提供一種用於從一降混訊號產生一合成訊號的方法，該合成訊號具有一合成聲道數，該方法包括：

接收一降混訊號及旁側資訊，該降混訊號具有一降混聲道數，該旁側資訊包括：

一原始訊號的聲道位準及相關資訊，該原始訊號具有一原始聲道數；
使用該原始訊號的聲道位準及相關資訊及被關聯到該降混訊號的協方差資訊產生該合成訊號。

【0122】 該方法可以包括：

從該降混訊號計算一原型訊號，該原型訊號具有該合成聲道數；
使用該原始訊號的聲道位準及相關資訊及被關聯到該降混訊號的協方差資訊計算一混合規則；及
使用該原型訊號及該混合規則產生該合成訊號。

【0123】 根據一個方面，提供一種用於從一原始訊號產生一降混訊號的方法，該原始訊號具有一原始聲道數，該降混訊號具有一降混聲道數，該方法包括：

估計該原始訊號的聲道位準及相關資訊，
將該降混訊號編碼成一位元流，使得該降混訊號被編碼在該位元流中，以便具有旁側資訊，該旁側資訊包括該原始訊號的聲道位準及相關資訊。

【0124】 根據一個方面，提供一種用於從一降混訊號產生一合成訊號的方法，該降混訊號具有一降混聲道數，該合成訊號具有一合成聲道數，該降混訊號為具有一原始聲道數的一原始訊號的一降混版本，該方法包括以下階段：

一第一階段，包括：

根據從以下內容計算出的一第一混合矩陣合成該合成訊號的一第一分量：

被關聯到該合成訊號的一協方差矩陣；及

被關聯到該降混訊號的一協方差矩陣，

一第二階段，用於合成該合成訊號的一第二分量，其中該第二分量是一殘餘分量，該第二階段包括：

一原型訊號步驟，將該降混訊號從該降混聲道數升混到該合成聲道數；

一去相關器步驟，將該被升混的原型訊號進行去相關化；

一第二混合矩陣步驟，根據來自該降混訊號的該去相關版本的一第二混合矩陣合成該合成訊號的該第二分量，該第二混合矩陣是一殘餘混合矩陣，

其中該方法從以下內容計算出該第二混合矩陣：

由該第一混合矩陣步驟提供的該殘餘協方差矩陣；及

從被關聯到該降混訊號的該協方差矩陣獲得的該數個去相關的原型訊號的該協方差矩陣的一估計，

其中該方法還包括一加法器步驟，將該合成訊號的該第一分量與該合成訊號的該第二分量求和，從而獲得該合成訊號。

【0125】 根據一個方面，提供一種用於從一降混訊號產生一合成訊號的音訊合成器，所述合成訊號具有一合成通道數，該合成通道數大於一或大於二，該音訊合成器包括：以下內容中的至少一者：

一輸入介面，被配置用於接收該降混訊號，所述降混訊號具有至少一個縮混聲道及旁側資訊，該旁側資訊包括以下內容中的至少一者：

一原始訊號的聲道位準及相關資訊，該原始訊號具有一原始聲道數，
該原始聲道數大於一或大於二；

一部件，諸如一原型訊號計算器[譬如“原型訊號計算”]，被配置用於從該降混訊號計算一原型訊號，該原型訊號具有該合成通道數；

一部件，諸如一混合規則計算器[譬如“參數重建”]，被配置用於使用該原始訊號的聲道位準及相關資訊、與該降混訊號相關聯的協方差資訊計算一個(或多個)混合規則；及

一部件，諸如一合成處理器[譬如“合成引擎”]，被配置用於使用該原型訊號及該混合規則產生該合成訊號。

【0126】 該合成聲道數可以大於該原始聲道數。替代地，該合成聲道數可以小於該原始聲道數。

【0127】 該音訊合成器(特別是，在某些方面，該混合規則計算器)可以被配置成重建該原始聲道位準及相關資訊的一目標版本。

【0128】 該音訊合成器(特別是，在某些方面，該混合規則計算器)可以被配置成重建該原始聲道位準及相關資訊的一目標版本，該相關資訊適應於該合成訊號的該聲道數。

【0129】 該音訊合成器(特別是，在某些方面，該混合規則計算器)可以被配置成重建該原始聲道位準及相關資訊的一目標版本，該相關資訊基於該原始聲道位準及相關資訊的一估計版本。

【0130】 該音訊合成器(特別是，在某些方面，該混合規則計算器)可以被配置成從與該降混訊號相關聯的協方差資訊獲得該原始聲道位準及相關資訊的該估計版本。

【0131】該音訊合成器(特別是，在某些方面，該混合規則計算器)可以被配置成針對該原型訊號，通過將與該原型訊號計算器所使用的一原型規則相關聯的一估計規則應用於與該降混訊號相關聯的該協方差資訊，獲得該原始聲道位準及相關資訊的該估計版本。

【0132】該音頻合成器(尤其是，在某些方面，該混合規則計算器)可以被配置成在該降混訊號的旁側資訊之中檢索以下兩者：

與該降混訊號相關聯的協方差資訊，描述在該降混訊號中的一第一聲道的位準或在一聲道對之間的一能量關係；及

該原始訊號的聲道位準及相關資訊，描述在該原始訊號中的一第一聲道的位準或在一聲道對之間的一能量關係，

以便通過使用以下內容中的至少一者來重建該原始聲道位準及相關資訊的該目標版本：

針對至少一個第一聲道或聲道對的該原始聲道的協方差資訊；及

描述該至少一個第一聲道或聲道對的該聲道位準及相關資訊。

【0133】該音頻合成器(尤其是，在某些方面，該混合規則計算器)可以被配置成偏好該聲道位準及相關資訊描述該聲道或聲道對，而不是針對相同聲道或聲道對的該原始聲道的該協方差資訊。

【0134】該原始聲道位準及相關資訊的該重建目標版本描述在一聲道對之間的一能量關係至少部分地是基於被關聯到該聲道對中的每個聲道的數個位準。

【0135】該降混訊號可以被劃分為數個頻帶或數個頻帶群組：不同聲道位準及相關資訊可以與不同頻帶或頻帶群組相關聯；該音訊合成器(該原型訊號計

算器，尤其是，在某些方面，該混合規則計算器及該合成處理器中的至少一個)被配置成針對不同頻帶或頻帶群組進行不同操作，以獲得針對不同頻帶或頻帶群組的不同混合規則。

【0136】 該降混訊號可以被劃分為數個時隙，其中不同的聲道位準及相關資訊與不同時隙相關聯，並且該音訊合成器的至少一個構件(譬如該原型訊號計算器、該混合規則計算器、該合成處理器或該合成器的其他元件)被配置成針對不同時隙進行不同操作，以獲得針對不同時隙的不同混合規則。

【0137】 該音訊合成器(譬如該原型訊號計算器)可以被配置成選擇一原型規則，該原型規則被配置用於在該合成聲道數的基礎上計算一原型訊號。

【0138】 該音訊合成器(譬如該原型訊號計算器)可以被配置成在數個預存原型規則之中選擇該原型規則。

【0139】 該音訊合成器(譬如該原型訊號計算器)可以被配置成在一手動選擇的基礎上定義一原型規則。

【0140】 該原型規則(譬如該原型訊號計算器)可以包括一矩陣，該矩陣具備一第一維度及一第二維度，其中該第一維度與該降混聲道數相關聯，並且該第二維度與該合成聲道數相關聯。

【0141】 該音訊合成器(譬如該原型訊號計算器)可以被配置成操作在等於或低於160千位元/秒的一位元率。

【0142】 該旁側資訊可以包括該數個原始聲道的一標識[譬如L、R、C等]。

【0143】 該音訊合成器(尤其是，在某些方面，該混合規則計算器)可以被配置用於使用該原始訊號的該聲道位準及相關資訊、與該降混訊號相關聯

的一協方差資訊，及該數個原始聲道的該標識，及該數個合成聲道的一標識來計算[譬如“參數重建”]一混合規則[譬如混合矩陣]。

【0144】該音訊合成器可以針對該合成訊號選擇[譬如通過選擇諸如手動選擇，或者通過預選擇，或者自動地譬如通過識別揚聲器數]多個聲道，一聲道數無關於在該旁側資訊中的該原始聲道的該聲道位準及相關資訊中的至少一者。

【0145】在一些示例中，該音訊合成器可以針對不同的選擇來選擇不同的原型規則。該混合規則計算器可以被配置為計算該混合規則。

【0146】根據一個方面，提供一種用於從一降混訊號產生一合成訊號的方法，該合成訊號具有一合成聲道數，該合成聲道數大於一或大於二，該方法包括：

接收該降混訊號，該降混訊號具有至少一個降混聲道及旁側資訊，該旁側資訊包括：

一原始訊號的聲道位準及相關資訊，該原始訊號具有一原始聲道數，該原始聲道數大於一或大於二；

從該降混訊號計算一原型訊號，該原型訊號具有該合成訊號數；

使用該原始訊號的該聲道位準及相關資訊、與該降混訊號相關聯的協方差資訊來計算一混合規則；及

使用該原型訊號及該混合規則[譬如一規則]產生該合成訊號。

【0147】根據一個方面，提供一種音訊編碼器，用於從一原始訊號[譬如y]產生一降混訊號，該原始訊號具有至少兩個聲道，該降混訊號具有至少一個降混聲道，該音訊編碼器包括以下的至少一個：

一參數估計器，被配置用於估計該原始訊號的聲道位準及相關資訊，
一位元流寫入器，用於將該降混訊號編碼成一位元流，使得該降混訊號被
編碼在該位元流中，以便具有旁側資訊，該旁側資訊包括該原始訊號的聲
道位準及相關資訊。

【0148】 被編碼在該旁側資訊中的該原始訊號的該聲道位準及相關資訊
表示被關聯到小於該原始訊號的該數個聲道的總數的聲道位準資訊。

【0149】 被編碼在該旁側資訊中的該原始訊號的該聲道位準及相關資訊
表示相關資訊，該相關資訊描述在該原始聲道中的至少一個不同聲道對之間的
數個能量關係，但小於該數個原始訊號的該數個聲道的總數。

【0150】 該原始訊號的聲道位準及相關資訊可以包括至少一個同調度
值，該同調度值描述一聲道對中的兩個聲道之間的同調度。

【0151】 該原始訊號的聲道位準及相關資訊可以包括在一聲道對的兩個
聲道之間的至少一個聲道間位準差(ICLD)。

【0152】 該音訊編碼器可以被配置成在狀態資訊的基礎上選擇是否編碼
或不編碼該原始訊號的該聲道位準及相關資訊的至少一部分，以便在有效載荷
相對較低的情況下，在該旁側資訊中包括一增加數量的聲道位準及相關資訊。

【0153】 該音訊編碼器可以被配置成在關於該數個聲道的衡量指標的基
礎上選擇該原始訊號的該聲道位準及相關資訊的哪一部分要被編碼在該旁側資
訊中，以便在該旁側資訊中包括被關聯到更敏感的衡量指標[譬如衡量指標為被
關聯到感知上更顯著的協方差]的聲道位準及相關資訊。

【0154】 該原始訊號的該聲道位準及相關資訊可以為一矩陣的形式。

【0155】 該位元流寫入器被配置成將至少一個聲道的標識進行編碼。

【0156】 根據一個方面，提供一種從一原始訊號產生一降混訊號的方法，該原始訊號具有至少兩個聲道，該降混訊號具有至少一個降混聲道。

【0157】 該方法可以包括：

估計該原始訊號的聲道位準及相關資訊，

將該降混訊號編碼成一位元流，使得該降混訊號被編碼在該位元流中，以便具有旁側資訊，該旁側資訊包括原始訊號的聲道位準及相關資訊。

【0158】 該音訊編碼器可以與該解碼器無關(agnostic to the decoder)。該音訊合成器可以與該解碼器無關。

【0159】 根據一個方面，提供一種系統，包括如上或以下的該音訊合成器及如上或以下的一音訊編碼器。

【0160】 根據一個方面，提供一種儲存指令的非暫時性儲存單元，當該指令由一處理器執行時致使該處理器進行一種如上或如下的方法。

【圖式簡單說明】

【0161】 3. 示例

【0162】 3.1 圖式

〔第1圖〕：顯示根據本發明的一處理的一簡化概圖。

〔第2a圖〕：顯示根據本發明的一音訊編碼器。

〔第2b圖〕：顯示根據本發明的音訊編碼器的另一視圖。

〔第2c圖〕：顯示根據本發明的音訊編碼器的另一視圖。

〔第2d圖〕：顯示根據本發明的音訊編碼器的另一視圖。

〔第3a圖〕：顯示根據本發明的一音訊合成器(解碼器)。

- 〔第3b圖〕：顯示根據本發明的音訊合成器(解碼器)的另一視圖。
- 〔第3c圖〕：顯示根據本發明的音訊合成器(解碼器)的另一視圖。
- 〔第4a圖〕：顯示協方差合成的一示例。
- 〔第4b圖〕：顯示協方差合成的另一示例。
- 〔第4c圖〕：顯示協方差合成的另一示例。
- 〔第4d圖〕：顯示協方差合成的另一示例。
- 〔第5圖〕：顯示根據本發明的用於一音訊編碼器的濾波器組的一示例。
- 〔第6a圖〕：顯示根據本發明的一音訊編碼器的運作的一示例。
- 〔第6b圖〕：顯示根據本發明的一音訊編碼器的運作的另一示例。
- 〔第6c圖〕：顯示根據本發明的一音訊編碼器的運作的另一示例。
- 〔第7圖〕：顯示先前技術的一示例。
- 〔第8a圖〕：顯示根據本發明的如何獲得協方差資訊的一示例。
- 〔第8b圖〕：顯示根據本發明的如何獲得協方差資訊的另一示例。
- 〔第8c圖〕：顯示根據本發明的如何獲得協方差資訊的另一示例。
- 〔第9a圖〕：顯示諸多聲道間同調矩陣的一示例。
- 〔第9b圖〕：顯示諸多聲道間同調矩陣的另一示例。
- 〔第9c圖〕：顯示諸多聲道間同調矩陣的另一示例。
- 〔第9d圖〕：顯示諸多聲道間同調矩陣的另一示例。
- 〔第10a圖〕：顯示諸多訊框的一示例。
- 〔第10b圖〕：顯示諸多訊框的另一示例。
- 〔第11圖〕：顯示由該解碼器使用於獲得一混合矩陣的一方案。

【實施方式】

【0163】 3.2 關於本發明的諸多概念

【0164】 將被顯示的是諸多示例基於該編碼器對一訊號(signal)212進行降混並對該解碼器提供聲道位準及相關資訊(channel level and correlation information)220。該解碼器可以從該聲道位準及相關資訊220產生一混合規則(mixing rule)(譬如混合矩陣)。對於產生該混合規則的重要資訊可以包括該原始訊號212的協方差資訊(covariance information)(譬如一協方差矩陣 C_y)及該降混訊號的協方差資訊(譬如一協方差矩陣 C_x)。雖然該協方差矩陣 C_x 可以由該解碼器通過分析該降混訊號直接估計，但是該原始訊號212的協方差矩陣 C_y 容易由該解碼器估計。該原始訊號212的該協方差矩陣 C_y 通常是一對稱矩陣(譬如在一5聲道原始訊號212的情況下為一5x5矩陣)：雖然該矩陣在該對角處展示每個聲道的位準，但它在數個非對角元(non-diagonal entries)處的該數個聲道之間展示諸多協方差。該矩陣是對角矩陣，因為在數個通用聲道 i 與 j 之間的該協方差與在 j 與 i 之間的該協方差相同。因此，為了對該解碼器提供整個協方差資訊，有必要對該解碼器以訊號表明(to signal to the decoder)在該數個對角元處的5個位準及在該數個非對角元處的10個協方差。然而，將被顯示的是，減少要被編碼的資訊量是可行的。

【0165】 此外，將被顯示的是，在某些情況下，可以提供數個正規化的值，代替該數個位準及數個協方差。例如：可以提供指示數個能量值的數個聲道間同調度值(ICCs，也以 $\xi_{i,j}$ 被指示)及數個聲道間位準差(ICLDs，也以 χ_i 被指示)。該ICCs可以是例如提供的數個相關值，而不是該矩陣 C_y 的該數個非對角元的該協方差。相關資訊的一示例可以為 $\xi_{i,j} = \frac{c_{y_{i,j}}}{\sqrt{c_{y_{i,i}} \cdot c_{y_{j,j}}}}$ 的形式。在某些示例中，僅對該 $\xi_{i,j}$ 的一部分進行實際編碼。

【0166】 以此方式，產生一ICC矩陣。該ICC矩陣的該數個對角元原則上將相等為1，因此不必在該位元流中對它們進行編碼。然而，已被理解的是，該編碼器對該解碼器提供ICLDs是可行的，譬如以 $\chi_i = 10 \cdot \log_{10} \left(\frac{P_i}{P_{\text{dmx},i}} \right)$ 的形式(也參見下文)。在某些示例中，所有該 χ_i 都被實際編碼。

【0167】 第9a至9d圖顯示一ICC矩陣900的諸多示例，其中數個對角線值“d”可以是數個ICLD χ_i ，而數個非對角線值以902、904、905、906、907(請參見下文)被指示，這可以是數個ICC $\xi_{i,j}$ 。

【0168】 在本文件中，在數個矩陣之間的乘積通過不帶一符號的方式被指示。譬如在矩陣A與矩陣B之間的乘積通過AB被指示。一矩陣的共軛轉置以一星號(*)被指示。

【0169】 當參考該對角線時，它是指主對角線(main diagonal)。

【0170】 3.3 本發明

【0171】 第1圖顯示具有一編碼器側及一解碼器側的一音訊系統100。該編碼器側可以由一編碼器200實施，並且可以獲得廣告音訊訊號212，譬如從一音訊感測器單元(譬如麥克風)，或者可以從一儲存單元或從一遠程單元(譬如經由一無線電傳輸)。該解碼器側可以由一音訊解碼器(音訊合成器)300實施，這可以將音訊內容提供給一音訊再現單元(譬如揚聲器)。該編碼器200及該解碼器300可以彼此通訊，譬如通過一通訊頻道，這可以是有線的或無線的(譬如通過射頻波、光或超音波等)。該編碼器及/或該解碼器因此可以包括或被連接到數個通訊單元(譬如天線、收發器等)，用於將該被編碼的位元流248從該編碼器200傳送到該解碼器300。在一些情況下，該編碼器200可以將該被編碼的位元流248儲存在一儲存單元(譬如RAM記憶體、FLASH記憶體等)中，以供將來使用。類似地，該解

碼器300可以讀取被儲存在一儲存單元中的該位元流248。在某些示例中，該編碼器200及該解碼器300可以是相同的裝置：在已經對位元流248進行編碼及保存後，該裝置可能需要讀取它以回放音訊內容。

【0172】 第2a、2b、2c及2d圖顯示諸多編碼器200的諸多示例。在某些示例中，第2a及2b及2c及2d圖的編碼器可以相同，並且僅因一個及/或另一幅圖中缺少某些要素而彼此不同。

【0173】 該音訊編碼器200可以被配置用於從一原始訊號212(具有至少兩個(譬如三個或更多個)聲道的該原始訊號212及具有至少一個降混聲道的該降混訊號246)產生一降混訊號246。

【0174】 該音訊編碼器200可以包括一參數估計器(parameter estimator)218，該參數估計器218被配置成估計該原始訊號212的聲道位準及相關資訊220。該音訊編碼器200可以包括一位元流寫入器(bitstream writer)226，用於將該降混訊號246編碼成一位元流248。因此，該降混訊號246以這樣的方式在該位元流248中被編碼，使得它具有旁側資訊228，該旁側資訊228包括原始訊號212的聲道位準及相關資訊。

【0175】 特別地，在某些示例中，該輸入訊號212可以被理解為一時域音訊訊號(time domain audio signal)，諸如例如諸多音訊樣本的一時間序列(a temporal sequence of audio samples)。該原始訊號212具有至少兩個聲道，該至少兩個聲道可以例如對應於不同的麥克風(譬如用於一立體聲音訊位置，或是然而，一多聲道音訊位置)，或者例如對應於一音訊再現單元的不同揚聲器位置。該輸入訊號212可以在一降混器計算塊244處被降混以獲得該原始訊號212的一降混版本246(也表示為x)。該原始訊號212的此降混版本也被稱為降混訊號246。

該降混訊號246具有至少一個降混聲道。該降混訊號246具有比該原始訊號212更少的數個聲道。該降混訊號212可以存在時域中。

【0176】 該降混訊號246由該位元流寫入器226(譬如包括一熵編碼器或一多工器或核心編碼器)在位元流248中被編碼，用於將一位元流儲存或傳送到一接收器(譬如與該解碼器側相關聯)。該編碼器200可以包括一參數估計器(或參數估計塊)218。該參數估計器218可以估計與該原始訊號212相關聯的聲道位準及相關資訊220。該聲道位準及相關資訊220可以在位元流248中被編碼為旁側資訊228。在諸多示例中，聲道位準及相關資訊220由該位元流寫入器226編碼。在諸多示例中，即使第2b圖未在該降混計算塊235的下游顯示該位元流寫入器226，該位元流寫入器226可以逕為存在。在第2c圖中，顯示該位元流寫入器226可以包括一核心編碼器247，以對該降混訊號246進行編碼，以便獲得該降混訊號246的一編碼版本(coded version)。第2c圖還顯示的是，該位元流寫入器226可以包括一多工器249，該多工器249在該位元流228中對該被編碼的降混訊號246及在該旁側資訊228中的聲道位準及相關資訊220(譬如作為被編碼的參數)兩者進行編碼。

【0177】 如第2b圖所示(在第2a及2c圖中缺少的)，該原始訊號212可以被處理(譬如通過濾波器組214，見下文)，以獲得該原始訊號212的一頻域版本(frequency domain version)216。

【0178】 參數估計的一示例被顯示在第6c圖中，其中一參數估計器218定義諸多參數 $\xi_{i,j}$ 及 χ_i (譬如諸多正規化的參數)，以後續被編碼在該位元流中。數個協方差估計器502及504分別對於要被編碼的降混訊號246及該輸入訊號212估計該協方差 C_x 及 C_y 。然後，在ICLD塊506，數個ICLD參數 χ_i 被計算並被提供到該

位元流寫入器246。在該協方差對同調度塊(covariance-to-coherence block)510處，數個 $\text{ICC}\xi_{i,j}$ (412)被獲得。在塊250處，僅一些ICC被選擇要被編碼。

【0179】一參數量化塊(parameter quantization block)222(第2b圖)可以允許獲得處於一量化版本(quantized version)224的該聲道位準及相關資訊220。

【0180】該原始訊號212的該聲道位準及相關資訊220通常可以包括關於該原始訊號212的一聲道的能量(或位準)的資訊。附加地或替代地，該原始訊號212的該聲道位準及相關資訊220可以包括在數個聲道對之間的相關資訊，諸如在兩個不同聲道之間的關聯。該聲道位準及相關資訊可以包括與協方差矩陣 C_y 相關聯的資訊(譬如以其正規化形式，諸如該相關聯或數個ICC)，其中每一列及每一行都與該原始訊號212的一特定聲道相關聯，並且通過該矩陣 C_y 的該數個對角元素及該相關資訊以描述該數個聲道位準，並且通過該矩陣 C_y 的數個非對角元以描述該相關資訊。該矩陣 C_y 可以是一對稱矩陣(即它等於其轉置矩陣)或一厄米特矩陣(即它等於其共軛轉置)。 C_y 通常是正半定的(positive semidefinite)。在某些示例中，該相關聯可以由該協方差替代(並且由協方差資訊替代該相關資訊)。已被理解的是，在該位元流248的該旁側資訊228中編碼與少於該原始訊號212的該數個聲道的總數相關聯的資訊是可行的。例如：不必提供關於所有聲道或所有聲道對的一聲道位準及相關資訊。例如：關於在該降混訊號212的數個聲道對之間的該相關聯的一減少的資訊集可以僅在該位元流248中被編碼，而該剩餘資訊可以在該解碼器側被估計。通常，將比 C_y 的對角元更少的元素進行編碼是可行的，並且將比 C_y 對角線之外的該數個元素更少的元素進行編碼是可行的。

【0181】例如：該聲道位準及相關資訊可以包括該原始訊號212的一協方差矩陣 C_y (該原始訊號的聲道位準及相關資訊220)及/或該降混訊號246的該協方

差矩陣 C_x (該降混訊號的協方差資訊)的數個元，譬如以正規化形式。例如：該協方差矩陣可以將每一行及每一列與每個聲道相關聯，以表示在不同聲道之間的數個協方差，並且在該矩陣的對角線上表示每個聲道的該位準。在某些示例中，作為編碼在該旁側資訊228中的該原始訊號212的該聲道位準及相關資訊220可以僅包括聲道位準資訊(譬如僅該相關聯矩陣 C_y 的對角線的數個值)或僅包括相關資訊(例如僅該相關聯矩陣 C_y 的對角線外部的數個值)。同樣應用於該降混訊號的該協方差資訊。

【0182】 如後續將被顯示的，該聲道位準及相關資訊220可以包括至少一個同調度值($\xi_{i,j}$)，描述在一聲道對(a couple of channels) i 、 j 中的兩個聲道 i 與 j 之間的同調度。附加地或替代地，該聲道位準及相關資訊220可以包括至少一個聲道間位準差， $ICLD(\chi_i)$ 。特別地，定義具有數個ICLD值或數個ICC值的一矩陣是可行的。因此，以上關於該矩陣 C_y 及 C_x 的數個元素的該傳輸的諸多示例可以被通用化(generalized)，用於要被編碼(譬如被傳輸)的其他值，用於實施該聲道位準及相關資訊220及/或該降混聲道的同調度資訊。

【0183】 該輸入訊號212可以被細分為數個訊框(a plurality of frames)。不同的訊框可以具有例如相同的時間長度(譬如每個訊框可以在經過一訊框的時間期間由在時域中的相同數量的樣本建構)。因此，不同的訊框通常具有相等的時間長度。在該位元流248中，降混訊號246(其可以是一時域訊號)可以用一逐訊框的方式(或者在任何情況下，將其細分為數個訊框可以由解碼器決定)被編碼。如在該位元流248中被編碼作為旁側資訊228那樣，該聲道位準及相關資訊220可以與每個訊框相關聯(譬如可以為每個訊框或者為數個連續的訊框提供該聲道位準及相關資訊220的該數個參數)。據此，對於該降混訊號246的每個訊框，一被關聯

的旁側資訊228(譬如數個參數)可以被編碼在該位元流248的該旁側資訊228中。在一些情況下，數個連續的訊框可以與如在該位元流248的該旁側資訊228中被編碼的相同的聲道位準及相關資訊220(譬如數個相同的參數)相關聯。據此，一個參數可以導致被共同地相關聯於數個連續的訊框。在某些示例中，當兩個連續的訊框具有相似的屬性時，或者當該位元率需要被降低(譬如由於減少有效載荷的必要性)時，這可能發生。例如：

在高有效載荷(payload)的情況下，增加與相同特定參數相關聯的數個連續的訊框的數量，以便減少被寫入該位元流的位元數量；

在有效載荷較低的情況下，減少與相同特定參數相關聯的數個連續的訊框的數量，以便提高該混合品質。

在其他情況下，當位元率被減少時，增加與相同特定參數相關聯的數個連續的訊框的數量，以便減少被寫入該位元流的位元數量，反之亦然。

【0184】 在某些情況下，可行的是使用在一當前的訊框以前的具備數個參數(或數個被重建的或被估計的值，諸如數個協方差)的數個線性組合以平滑數個參數(或數個被重建的或被估計的值，諸如數個協方差)，譬如通過加法、平均等。

【0185】 在某些示例中，一訊框可以在數個後續時隙(a plurality of subsequent slots)之間被劃分。第10a圖顯示一訊框920(被細分為四個連續的時隙921至924)，第10b圖顯示訊框930(細分為四個連續的時隙931至934)。不同時隙的時間長度可以相同。如果該訊框的長度是20毫秒(ms)及1.25 ms的時隙大小，則在一訊框中有16個時隙($20/1.25=16$)。

【0186】 該時隙細分可以在諸多濾波器組(例如214)中被進行，如下所討論的。

【0187】 在一個示例中，濾波器組是一複雜調變的低延遲濾波器組 (CLDFB)，該訊框的大小為20 ms，該時隙的大小為1.25 ms，導致每訊框16個濾波器組以及每個時隙的數個頻帶的一數量取決於輸入取樣頻率以及該數個頻帶具有的一寬度為400赫茲(Hz)。因此，譬如對於48千赫(kHz)的一輸入取樣頻率，在諸多樣本中的訊框的長度為960，該時隙長度為60個樣本，每時隙的濾波器組樣本的數量也是60。

【0188】			
取樣頻率/kHz	訊框的長度/樣本	時隙長度/樣本	濾波器組頻帶的數量
48	960	60	60
32	640	40	40
16	320	20	20
8	160	10	10

【0189】 即使每個訊框(以及每個時隙)可以在時域中被編碼，一逐頻帶的分析也可以被執行。在諸多示例中，對於每個訊框(或時隙)分析數個頻帶。例如：該濾波器組可以被應用於該時間訊號，並且所得的子頻帶訊號可以被分析。在某些示例中，該聲道位準及相關資訊220還以一逐頻帶的方式被提供。例如：對於該輸入訊號212或降混訊號246的每個頻帶，一相關聯的聲道位準及相關資訊220(譬如C_y或ICC矩陣)可以被提供。在某些示例中，該數個頻帶的數量可以基於該訊號及/或被請求的位元率或當前有效載荷上的測量的屬性被修改。在某些示例中，被需要的時隙越多，被使用的頻帶越少，以維持一相似的位元率。

【0190】 由於該時隙的大小小於該訊框的大小(在時間長度上)，因此在一訊框內檢測到該原始訊號212中的暫態的情況下，該數個時隙可以適時地被使用：該編碼器(尤其是該濾波器組214)可以識別該暫態的存在，以訊號表明其在

該位元流中的存在，並且在該位元流248的該旁側資訊228中指示在該訊框的哪個時隙中已經發生暫態。此外，被編碼在該位元流248的該旁側資訊228中的該聲道位準及相關資訊220的該數個參數可以因而僅與該暫態後續的數個時隙及/或該暫態已經發生的時隙相關聯。因此，該解碼器將決定該暫態的存在，並且將聲道位準及相關資訊220僅與該暫態後續的數個時隙及/或該暫態已經發生的時隙相關聯(對於該暫態以前的該數個時隙，該解碼器將使用該先前的訊框的聲道位準及相關資訊220)。在第10a圖中，沒有暫態已經發生，並且在該旁側資訊228中被編碼的該數個參數220因此可以被理解為與整個訊框920相關聯。在第10b圖中，該暫態已經發生在時隙932處：因此，在該旁側資訊228中被編碼的該數個參數220將引用該數個時隙932、933及934，而與該時隙931相關聯的該數個參數將被假定為與在該訊框930以前的訊框相同。

【0191】 鑑於以上內容，對於每個訊框(或時隙)及每個頻帶，與該原始訊號212有關的一特定聲道位準及相關資訊220可以被定義。例如：該協方差矩陣 C_y 的數個元素(譬如數個協方差及/或數個位準)可以針對每個頻帶被估計。

【0192】 如果在數個訊框被共同相關聯於相同參數的同時發生一暫態的檢測，則減少被共同相關聯於相同參數的該數個訊框的數量是可行的，從而增加該混合品質。

【0193】 第10a圖顯示該訊框920(在此被指示為“正常訊框”)，在該原始訊號212中為其定義八個頻帶(在縱坐標顯示八個頻帶1...8，而在橫坐標顯示該數個時隙921至924)。該聲道位準及相關資訊220的該數個參數可以在理論上以一逐頻帶的方式(譬如對於每個原始頻帶將存在一個協方差矩陣)在該位元流248的該旁側資訊228中被編碼。然而，為了減少旁側資訊228的數量，該編碼器可以聚合

多個原始頻帶(譬如數個連續頻帶)，以獲得由多個原始頻帶形成的至少一個聚合頻帶(aggregated band)。例如：在第10a圖中，八個原始頻帶被分組以獲得四個聚合頻帶(聚合頻帶1與原始頻帶1相關聯；聚合頻帶2與原始頻帶2相關聯；聚合頻帶3將原始頻帶3及5分組；聚合頻帶4將原始頻帶5...8分組)。協方差、相關聯、ICC等的矩陣可以與該數個聚合頻帶中的每一個相關聯。在某些示例中，在該位元流248的該旁側資訊228中被編碼的是從與每個聚合頻帶相關聯的該數個參數的總和(或平均值或另一線性組合)獲得的數個參數。因此，該位元流248的該旁側資訊228的大小被進一步降低。在下文中，“聚合頻帶(aggregated band)”也被稱為“參數頻帶(parameter band)”，因為它意指被用於決定該數個參數220的那些頻帶。

【0194】 第10b圖顯示其中發生一暫態訊框931(被細分為四個連續的時隙931至934，或為另一個整數)。在此，該暫態發生在第二時隙932(“暫態時隙(transient slot)”)中。在這種情況下，該解碼器可以決定僅將該聲道位準及相關資訊220的該數個參數引用到該暫態時隙932及/或後續時隙933及934。先前時隙931的聲道位準及相關資訊220將不被提供：已被理解的是，該時隙931的聲道位準及相關資訊在原則上將與該數個時隙的該聲道位準及相關資訊特別不同，但是可能會更類似在訊框930以前的訊框的聲道位準及相關資訊。因此，該解碼器將在該訊框930以前的訊框的聲道位準及相關資訊應用於該時隙931，並且訊框930的聲道位準及相關資訊僅應用於時隙932、933及934。

【0195】 由於具備該暫態的該時隙931的存在及位置可以被以訊號表明(譬如261中，如稍後所示)在該位元流248的該旁側資訊228中，因此一種技術已經被開發以避免或減小該旁側資訊228的大小增加：在數個聚合頻帶之間的分組可

以被更改：例如：該聚合頻帶1將原始頻帶1及2分組，該聚合頻帶2將原始頻帶3...8分組。因此，相對於第10a圖的情況，該數個頻帶的數量被進一步降低，並且將僅為兩個聚合頻帶提供該數個參數。

【0196】 第6a圖顯示該參數估計塊(參數估計器)218能夠檢索(retrieving)一定數量的聲道位準及相關資訊220。

【0197】 第6a圖顯示該參數估計器218能夠檢索一定數量的參數(聲道位準及相關資訊220)，這可以是第9a至9d圖的該矩陣900的該數個ICC。

【0198】 但是，實際上僅有一部分估計參數被提交到該位元流寫入器226，以對該旁側資訊228進行編碼。這是因為該編碼器200可以被配置成選擇(在第1至5圖中未被顯示的一決定塊250處)是否對該原始訊號212的該聲道位準及相關資訊220的至少一部分進行編碼。

【0199】 這在第6a圖中被圖解說明作為數個開關254s，這些開關受控於來自該決定塊250的一選擇(命令)254。如果該塊參數估計218的該數個輸出220中的每一個是第9c圖的該矩陣900的一ICC，則不是由該參數估計塊218估計的該整體的數個參數實際上未被編碼在該位元流248的該旁側資訊228中：特別是，雖然該數個元908(在該數個聲道之間的數個ICC：R與L；C與L；C與R；RS與CS)實際上被編碼，但該數個元907未被編碼(即，該決定塊250，可以是與第6c圖的那個相同，可以被視為已經打開用於該數個未被編碼的元907的開關254s，但是已經關閉用於在該位元流248的該旁側資訊228中要被編碼的該數個元908的開關254s。要被注意的是，在數個參數已被選擇要被編碼的資訊254'(數個元908)可以被編碼(譬如作為一位元映像(bitmap)或數個元(entries)908被編碼的其他資訊)。實際上，該資訊254'(例如可以是一ICC映像(ICC map))可以包括該數個被編碼的

元908的該數個索引(在第9d圖中被示意)。該資訊254'可以是一位元映像的形式：譬如該資訊254'可以由一固定長度的欄位構成，每個位置根據一預定順序與一索引相關聯，每個位元的值提供的資訊有關與該索引相關聯的參數是否被實際提供。

【0200】 通常，該決定塊250例如可以選擇是否對該聲道位準及相關資訊220的至少一部分進行編碼(即，決定該矩陣900的一元是否要被編碼)，例如：在狀態資訊252的基礎上。該狀態資訊252可以是基於一有效載荷狀態(payload status)：例如：在一傳輸為高度負載的情況下，將有可能減少要在該位元流248中要被編碼的該旁側資訊228的數量。例如：並且參考第9c圖：

在高有效載荷的情況下，減少被實際寫入該位元流248的該旁側資訊228中的該矩陣900的數個元908的數量；

在有效載荷較低的情況下，減少被實際寫入該位元流248的該旁側資訊228中的該矩陣900的數個元908的數量。

【0201】 替代地或附加地，衡量指標252可以被評估以決定哪些參數220要被編碼在該旁側資訊228中(譬如該矩陣900的哪些元被指定為數個被編碼的元908，以及哪些元要被丟棄)。在這種情況下，可能僅在該位元流中編碼該數個參數220(與數個更敏感的衡量指標相關聯，譬如與感知上更重要的協方差相關聯的衡量指標可以與要被選擇作為數個被編碼的元908的數個元相關聯)。

【0202】 要被注意的是，可以對於每個訊框(或者在降取樣的情況下對於多個訊框)及對於每個頻帶重複此過程。

【0203】 因此，除了該數個狀態衡量指標等之外，該決定塊250還可以由參數估計器218通過第6a圖中的命令251被控制。

【0204】 在某些示例中(譬如第6b圖)，該音訊編碼器可以進一步被配置成在該位元流248中將當前的(current)聲道位準及相關資訊220t編碼作為相對於先前的(previous)聲道位準及相關資訊220(t-1)的增量220k。由此位元流寫入器226在該旁側資訊228中編碼的內容可以是與相對於一先前的訊框的當前的訊框(或時隙)相關聯的一增量220k。這在第6b圖中顯示。一當前的聲道位準及相關資訊220t被提供到一儲存元件(storage element)270，使得該儲存元件270儲存用於後續的訊框的當前的聲道位準及相關資訊220t的值。同時，可以將當前的聲道位準及相關資訊220t與先前獲得的聲道位準及相關資訊220(t-1)進行比較。(這在第6b圖中被顯示為該減法器273)。因此，可以由該減法器273獲得一減法結果220 Δ 。該差220 Δ 可以在該縮放器220s處被使用，以獲得在先前的聲道位準及相關資訊220(t-1)與當前的聲道位準及相關資訊220t之間的一相對增量220k。例如：如果當前的聲道位準及相關資訊220t比先前的聲道位準及相關資訊220(t-1)大10%，則由該位元流寫入器226在該旁側資訊228中編碼的該增量220將指示該10%的增量的資訊。在某些示例中，代替提供該相對增量220k，可以簡單地對該差220 Δ 進行編碼。

【0205】 在如上及如下所討論的諸如ICC及ICLD的參數之中，要被實際編碼的參數的選擇可以調適應特定情況。例如：在某些示例中：

針對一個第一訊框，僅第9c圖的數個ICC 908被選擇在該位元流248的該旁側資訊228中要被編碼，而該數個ICC 907在該位元流248的該旁側資訊228中未被編碼；

針對一第二訊框，不同的ICC被選擇要被編碼，而不同的未被選擇的ICC未被編碼。

【0206】對於數個時隙及數個頻帶(及對於不同的參數，諸如數個ICLD)，可能同樣是有效的。因此，該編碼器(特別是塊250)可以決定哪個參數要被編碼及哪個參數不被編碼，因而使得要被編碼的參數的選擇適應於特定情況(譬如狀態、選擇...)。一“重要性特徵(feature for importance)”可以因此被分析，以便選擇哪個參數要被編碼及哪個參數不被編碼。該重要性特徵可以是例如與由該解碼器進行的數個操作的模擬中獲得的結果相關聯的一衡量指標(a metrics)。例如：該編碼器可以模擬該解碼器對該未編碼的協方差參數907的重建，並且該重要性的特徵可以是指示在該未編碼的協方差參數907與推測由該解碼器重建的相同參數之間的該絕對誤差的一衡量指標。通過測量不同模擬場景中的誤差(譬如每個模擬場景與某些被編碼的協方差參數908的傳輸及影響未被編碼的協方差參數907的重建的誤差的測量相關聯)，決定受到錯誤影響最小的模擬場景(譬如該模擬場景中有關重建中的所有錯誤的衡量指標)是可行的，以便基於該受影響最小的模擬場景將要被編碼的協方差參數908與不被編碼的協方差參數907區分開。在該受影響最小的場景的情況下，該未被選擇的參數907是最易於重建的參數，而該被選擇的參數908傾向於與該誤差相關聯的衡量指標最大的參數。

【0207】相同的內容可以被進行，通過模擬該解碼器的重建或估計該協方差，或者通過模擬混合特性或混合結果，而不是模擬像是ICC及ICLD的參數。值得注意的是，該模擬可以對於每個訊框或每個時隙進行，並且可以對於每個頻帶或聚合頻帶進行。

【0208】一個示例可以是從該位元流248的該旁側資訊228中被編碼的該數個參數開始，使用公式(4)或(6)(請參見下文)進行模擬該協方差的重建。

【0209】更通常地，從被選擇的聲道位準及相關資訊重建聲道位準及相關資訊是可行的，從而在該解碼器(300)處模擬未被選擇的聲道位準及相關資訊(220、Cy)的該估計，並且計算在以下內容之間的錯誤資訊：

由該編碼器估計的該未被選擇的聲道位準及相關資訊(220)；及
通過在該解碼器(300)處模擬未被編碼的聲道位準及相關資訊(220)的該估計而被重建的該未被選擇的聲道位準及相關資訊；及
以便在該被計算出的錯誤資訊的基礎上進行區分：

可適當重建的聲道位準及相關資訊；從

不可適當重建的聲道位準及相關資訊，

以便決定：

選擇在該位元流(248)的該旁側資訊(228)中要被編碼的該不可適當重建的聲道位準及相關資訊；及

未選擇該可適當重建的聲道位準及相關資訊，從而避免在該位元流(248)的該旁側資訊(228)中編碼該可適當重建的聲道位準及相關資訊。

一般而言，該編碼器可以模擬該解碼器的任何操作，並根據該模擬結果評估一錯誤衡量指標(an error metrics)。

【0210】在某些示例中，該重要性的特徵與被關聯到該錯誤的一衡量指標的評估可以不同(或包括其他衡量指標不同)。在某些情況下，該重要性的特徵可以與一手動選擇(a manual selection)相關聯，或基於奠基在心理聲學標準的一重要性。例如：即使沒有一模擬(simulation)，最重要的聲道對也可以被選擇要被編碼(908)。

【0211】現在，提供一些額外的討論用於解釋該編碼器如何以訊號表明實際上在該位元流248的該旁側資訊220中編碼哪些參數908。

【0212】參考第9d圖，在一ICC矩陣900的對角線上的數個參數與有序索引1...10(該順序是預定的且該解碼器已知)相關聯。在第9c圖中，顯示該數個被選擇的要被編碼的參數908是用於分別由索引1、2、5、10進行索引的數個對L-R、L-C、R-C、LS-RS的數個ICC。因此，在該位元流248的該旁側資訊228中，還將提供數個索引1、2、5、10的一指示(譬如在第6a圖的資訊254'中)。據此，借助於由該編碼器在該旁側資訊228中提供的關於該數個索引1、2、5、10的資訊，該解碼器將理解的是，在該位元流248的該旁側資訊228中被提供的四個ICC是L-R、L-C、R-C、LS-RS。可以例如通過將一位元映像中的每個位元的位置與預定的位置相關聯來提供該數個索引。例如：為了以訊號表明該數個索引1、2、5、10，可以寫入“1100100001”(在該旁側資訊228的欄位254'中)，因為第一、第二、第五及第十位元是指數個索引1、2、5、10(其他可能性可由技術人員支配)。這是所謂的一維索引(one-dimensional index)，但是其他索引策略也是可能的。例如：一種組合數技術，根據該組合數技術(在該旁側資訊228的該欄位254'中)對一數字N進行編碼，該數字N明確地與一特定的聲道對相關聯(另請參見https://en.wikipedia.org/wiki/Combinatorial_number_system)。當該位元映像引用數個ICC時，也可以被稱為一ICC映像。

【0213】要被注意的是，在某些情況下，一非適應(固定)的參數條款被使用。這意謂著，在第6a圖的示例中，在該數個要被編碼的參數之中的選擇254是固定的，並且不需要在欄位254'中指示該數個被選擇的參數。第9b圖顯示該數個參數的固定條款的一示例：所選擇的數個ICC為L-C、L-LS、R-C、C-RS，並且

無需進行以訊號表明它們的索引，因為該解碼器已經知道在該位元流248的該旁側資訊228中被編碼哪些ICC。

【0214】 然而，在某些情況下，該編碼器可以在該數個參數的一固定條款與該數個參數的一適應條款(adaptive provision)之間進行一選擇。該編碼器可以在該位元流248的該旁側資訊228中以訊號表明該選擇，以便該解碼器可以知道哪些參數被實際編碼。

【0215】 在某些情況下，至少一些參數可以被提供而不進行修改：例如：該數個ICDL可以在任何情況下被編碼，而無需在一位元映像中指示它們；及該數個ICC可能要接受一適應條款(adaptive provision)。

【0216】 該數個解釋涉及每個訊框(frame)、時隙(slot)或頻帶(band)。對於一後續的(subsequent)訊框、時隙或頻帶，不同的參數908被提供給該解碼器，將不同的索引與該後續的訊框、時隙或頻帶相關聯；並且可以進行不同的選擇(譬如固定的與適應的)。第5圖顯示該編碼器200的一濾波器組214的一示例，其可以被用於處理該原始訊號212，以獲得該頻域訊號216。從第5圖可以看出，該時域(TD)訊號212可以通過該暫態分析塊258(暫態檢測器)被分析。此外，由濾波器263(可以實現例如一傅立葉濾波器、一短傅立葉濾波器、一正交鏡等)提供在多個頻帶中的該輸入訊號212的一頻域(FD)版本264的一轉換。該輸入訊號212的該頻域版本264可以被分析，例如在頻帶分析塊267處，頻帶分析塊267可以決定(命令268)要在分區分組塊265處被進行的一特定頻帶分組(a particular grouping of the bands)。此後，該FD訊號216將是一聚合頻帶數量減少的一訊號。該數個頻帶的聚合已經如上關於第10a圖及第10b圖進行說明。該分區分組塊(partition grouping

block)267還可以由該暫態分析塊258進行該暫態分析而被調節。如上所述，在暫態的情況下，有可能進一步減少該數個聚合頻帶的數量：因此，關於該暫態的資訊260可以調節該分區分組。附加地或替代地，關於該暫態的資訊261被編碼在該位元流248的該旁側資訊228中。當該資訊261被編碼在該旁側資訊228中，該資訊261可以包括譬如指示該暫態是否已經發生的一旗標(flag)(諸如：“1”，意謂著“在該訊框中存在暫態(transient)”與“0”，意謂著：“在該訊框中沒有暫態”)及/或該暫態在該訊框中的位置的一指示(諸如指示該暫態在哪個時隙中已被觀察到的一欄位)。在某些示例中，當該資訊261指示在該訊框中沒有暫態(“0”)時，沒有該暫態的位置的指示被編碼在該旁側資訊228中，以減小該位元流248的大小。資訊261也被稱為“暫態參數(transient parameter)”，並且如第2d及6b圖所示，被編碼為該位元流246的該旁側資訊228中。

【0217】 在某些示例中，在塊265處的該分區分組還可以由外部資訊260’進行調節，諸如關於該傳輸的狀態的資訊(譬如與該傳輸相關聯的測量、錯誤率等)。例如：有效載荷越高(或該錯誤率越大)，該聚合就越大(傾向較少的聚合頻帶是較寬的)，從而具有較少量的旁側資訊228要被編碼在該位元流248中。在某些示例中，該資訊260’可以類似於第6a圖的資訊或衡量指標252。

【0218】 通常對於每個頻帶/時隙組合發送數個參數是不可行的，但是數個濾波器組樣本在許多時隙及許多頻帶上都被分組，以減少每訊框發送的參數集的數量。沿著該頻率軸，將該數個頻帶群組為數個參數頻帶會在數個參數頻帶中使用一非恆定的劃分，其中在數個參數頻帶中的該頻帶數量不是恆定的，而是嘗試遵循一心理聽覺激勵的參數頻帶解析度(a psychoacoustically motivated parameter band resolution)，即，在數個較低頻帶處，該數個參數頻帶僅包含一個

或少量的濾波器組頻帶，並且對於數個較高的參數頻帶，將較大(且穩定增加的)數量的濾波器組頻帶群組為一個參數頻帶。

【0219】 因此，譬如對於一輸入取樣率為48 kHz且該參數頻帶的數量設為14的情況，該跟隨向量 grp_{14} 描述該數個濾波器組索引，這些濾波器組索引給出用於該參數頻帶的該頻帶邊界(索引從0開始)：

$$grp_{14} = [0, 1, 2, 3, 4, 5, 6, 8, 10, 13, 16, 20, 28, 40, 60]$$

參數頻帶 j 包含該數個濾波器組頻帶 $[grp_{14}[j], grp_{14}[j + 1]]$

【0220】 注意的是，通過將該頻帶簡單截斷，以48 kHz進行分組的頻帶也可以被直接用於其他可能的取樣率，因為該分組都遵循一心理聽覺激勵的頻率刻度(psychoacoustically motivated frequency scale)並且具有與每個取樣頻率的頻帶數量相對應的某些頻帶邊界(表1)。

【0221】 如果一訊框是非暫態的，或者沒有暫態處理被實現，則沿該時間軸的分組將遍歷在一訊框中的所有時隙，以便每參數頻帶可用一個參數集。

【0222】 儘管如此，該參數集的數量還是很大，但是該時間解析度可以低於該數個20 ms的訊框(平均40 ms)。因此，為了進一步減少每訊框發送的參數集的數量，僅該數個參數頻帶的一子集被使用於決定及編碼用於在該位元流中發送給該解碼器的該數個參數。該數個子集是固定的，並且對於該編碼器及解碼器都是已知的。在該位元流中發送的該特定子集被以訊號表明在該位元流中的一欄位，以指示該解碼器傳輸的參數屬於數個參數頻帶的哪個子集，並且該解碼器然後以該數個被傳輸的參數(數個ICC、數個ICLD)替換用於該數個參數的此子集，並且對於不在當前子集中的所有參數頻帶保持來自該數個先前的訊框的數個參數(數個ICC、數個ICLD)。

【0223】 在一示例中，該數個參數頻帶可以被分為兩個子集，該兩個子集大致包含一半的全部參數頻帶及針對該數個較低的參數頻帶的連續子集及針對該數個較高的參數頻帶的一個連續子集。由於我們有兩個子集，用於以訊號表明子集的該位元流欄位是一個位元，並且用於48 kHz及14個參數頻帶的該數個子集的一示例是：

$$s_{14} = [1,1,1,1,1,1,1,0,0,0,0,0,0,0]$$

其中 $s_{14}[j]$ 指示屬於參數頻帶 j 的哪個子集。

【0224】 要被注意的是，該降混訊號246實際上可以在該位元流248中被編碼為在該時域中的一訊號：簡單地，該後續的參數估計器218將在該頻域中估計該數個參數220(譬如 $\xi_{i,j}$ 及/或 χ_i)(並且該解碼器300將使用該數個參數220用於準備該混合規則(譬如混合矩陣)403，這將被解釋如下。

【0225】 第2d圖顯示一編碼器200的一示例，該編碼器200可以是前述諸多編碼器中的一個或可以包括先前討論的諸多編碼器的諸多元素。一TD輸入訊號212被輸入到該編碼器，並且輸出一位元流248，該位元流248包括降混訊號246(譬如被該核心編碼器247編碼的)及在該旁側資訊228中被編碼的關聯與位準資訊220。

【0226】 從第2d圖可以看出，可以包括一濾波器組(filterbank)214(在第5圖中提供濾波器組的一示例)。在一塊263中提供一頻域(frequency domain, FD)轉換(頻域DMX)，以獲得一FD訊號264，該FD訊號264為該輸入訊號212的該FD版本。獲得數個頻帶中的FD訊號264(也用X表示)。該頻帶/時隙分組塊265(其可以實施為第5圖的該分組塊265)可以被提供，以獲得在數個聚合頻帶中的該FD訊號216。在某些示例中，該FD訊號216可以是在較少的頻帶中的該FD訊號264的一

版本。後續，該訊號216可以被提供給該參數估計器218，其包括數個協方差估計塊(covariance estimation blocks)502、504(在此被顯示為一個單一的塊)，以及在下流的一參數估計及編碼塊(a parameter estimation and coding block)506、510(元件502、504、506及510的實施例在第6c圖中被顯示)。該參數估計編碼塊506、510還可以提供在該位元流248的該旁側資訊228中要被編碼的該數個參數220。一暫態檢測器(transient detector)258(其可以實施為第5圖的該暫態分析塊258)可以找出該暫態及/或在一訊框內的一暫態的位置(譬如在哪個時隙中已經識別一暫態)。因此，關於該暫態(譬如暫態參數)的資訊261可以被提供給該參數估計器218(譬如決定哪些參數要被編碼)。該暫態檢測器258還可以提供資訊或命令(268)給該塊265，以便通過考慮在該暫態在該訊框中的存在及/或位置以執行分組(grouping)。

【0227】 第3a、3b、3c圖顯示諸多音訊解碼器300(也稱為音訊合成器)的諸多示例。在諸多示例中，第3a、3b、3c圖的該數個解碼器可以是相同的解碼器，只是為了避免不同的要素而具備一些差異。在諸多示例中，該解碼器300可以與第1及4圖的解碼器相同。在諸多示例中，該解碼器300也可以是與該編碼器200相同的裝置。

【0228】 該解碼器300可以被配置用於從在TD(246)或FD(314)中的一降混訊號(downmix signal) x 產生一合成訊號(synthesis signal)(336、340、 y_R)。該音訊合成器300可以包括一輸入介面(input interface)312，該輸入介面被配置用於接收該降混訊號246(譬如被該編碼器200編碼的相同降混訊號)及旁側資訊(side information)228(譬如在該位元流248中被編碼的)。如上所述，該旁側資訊228可以包括，如上所述，一原始訊號(其可以是在該編碼器側的該原始輸入訊號212、

y)的聲道位準及相關資訊(220、314)，諸如 ξ 、 χ 等或其元素中的一個(如下所述)。
在某些示例中，由該解碼器300獲得該ICC矩陣900的對角線之外的所有ICLD(χ)
及一些元(但非全部)906或908(數個ICC或數個 ξ 值)。

【0229】該解碼器300可以被配置(譬如通過一原型訊號計算器或原型訊號計算模組326)，用於從該降混訊號(324、246、 x)計算一原型訊號328，該原型訊號328具有該合成訊號336的該聲道數(大於一個)。

【0230】該解碼器300可以被配置(譬如通過一混合規則計算器402)，用於使用以下內容的至少一者進行計算一混合規則403：

該原始訊號(212、 y)的該聲道位準及相關資訊(例如314、 C_y 、 ξ 、 χ 或其元素)；
及

與該降混訊號(324、246、 x)相關聯的協方差資訊(譬如 C_x 或其元素)。

【0231】該解碼器300可以包括一合成處理器404，該合成處理器404被配置用於使用該原型訊號328及該混合規則403以產生該合成訊號(336、340、 y_R)。

【0232】該合成處理器404及該混合規則計算器402可以被收集在一個合成引擎(synthesis engine)334中。在某些示例中，該混合規則計算器402可以在該合成引擎334的外部。在某些示例中，第3a圖的該混合規則計算器402與第3b圖的該參數重建模組316可以被整合。

【0233】該合成訊號(336、340、 y_R)的該合成聲道數大於1(在某些情況下大於2或大於3)，並且可以大於、小於或等於該原始訊號(212、 y)的該原始聲道數，該原始聲道數也大於1(在某些情況下大於2或大於3)。該降混訊號(246、216、 x)的該聲道數至少為一或兩個，並且小於該原始訊號(212、 y)的該原始聲道數及該合成訊號(336、340、 y_R)的該合成聲道數。

【0234】 該輸入介面312可以讀取一被編碼的位元流248(譬如由該編碼器200編碼的相同位元流248)。該輸入介面312可以是或包括一位元流讀取器(bitstream reader)及/或一熵解碼器(entropy decoder)。如上所述，該位元流248可以如上所述對該降混訊號(246、 x)及旁側資訊228進行編碼。該旁側資訊228可以例如包含該原始聲道位準及相關資訊220，以被該參數估計器218或該參數估計器218下游的任何元素(譬如參數量化塊222等)輸出的形式。該旁側資訊228可以包含數個編碼值(encoded values)或數個索引值(indexed values)或兩者。即使在第3b圖中未針對該降混訊號(346、 x)顯示該輸入介面312，該輸入介面312也可以如第3a圖所示被應用於該降混訊號。在某些示例中，該輸入介面312可以量化從該位元流248獲得的數個參數。

【0235】 因此，解碼器300可以獲得該降混訊號(246、 x)，該降混訊號(246、 x)可以是在時域中。如上所述，該降混訊號246可以被劃分為數個訊框(frames)及/或數個時隙(slots)(請參見上文)。在諸多示例中，一濾波器組(filterbank)320可以轉換在時域中的該降混訊號246以獲得在頻域中的該降混訊號246的一版本324。如上所述，該降混訊號246的該頻域版本324的該數個頻帶可以被分組為數個頻帶群組(groups of bands)。在諸多示例中，可以執行針對在該濾波器組214處被進行的相同分組(grouping)(請參見上文)。用於該分組的該數個參數(譬如哪些頻帶及/或多少頻帶要被分組...)可以例如基於該分區分組器265或該頻帶分析塊267的信令(signalling)，該信令被編碼在該旁側資訊228中。

【0236】 該解碼器300可以包括一原型訊號計算器326。該原型訊號計算器326可以從該降混訊號(譬如該數個版本324、246、 x 中的一個)計算一原型訊號328，譬如通過應用一原型規則(譬如一矩陣 Q)。該原型規則可以通過具備一第一

維度及一第二維度的一原型矩陣(Q)被實施，其中該第一維度與該降混聲道數相關聯，該第二維度與該合成聲道數相關聯。因此，該原型訊號具有最終要被產生的該合成訊號340的該聲道數。

【0237】 該原型訊號計算器326可以將所謂的升混應用於該降混訊號(324、246、x)，在某種意義上，它只是在一被增加的聲道數中產生該降混訊號(324、246、x)的一版本(要被產生的該合成訊號的該聲道數)，但無需施加過多的“智能(intelligence)”。在諸多示例中，該原型訊號計算器326可以簡單地將一固定的預定原型矩陣(在本文件中被標識為“Q”)應用於該降混訊號246的該FD版本324。在諸多示例中，該原型訊號計算器326可以將不同的原型矩陣應用於不同的頻帶。例如在特定的降混聲道數及特定的合成聲道數的基礎上，可以在數個預存原型規則中選擇該原型規則(Q)。

【0238】 該原型訊號328可以在一去相關模組(decorrelation module)330處去相關化，以獲得該原型訊號328的一去相關版本332。然而，在某些示例中，有利地，該去相關模組330是不存在的，因為本發明已被證明是足夠有效以允許其迴避。

【0239】 該原型訊號(以其版本328、332中的任何一個)可以被輸入到該合成引擎334(並且特別是該合成處理器404)。在此，對該原型訊號(328、332)進行處理以獲得該合成訊號(336、 y_R)。該合成引擎334(並且特別是該合成處理器404)可以應用一混合規則403(在某些示例中，討論如下)，該混合規則是兩個，譬如一個用於該合成訊號的一主要分量，一個用於一殘餘分量)。該混合規則403可以例如通過一矩陣被實施。該矩陣403可以例如由該混合規則計算器402基於該原始訊號(212、y)的該聲道位準及相關資訊(314，諸如 ξ 、 χ 或其元素)而被產生。

【0240】 由該合成引擎334(特別是由該合成處理器404)輸出的該合成訊號336可以是可選地在一濾波器組338處被濾波。附加地或替代地，該合成訊號336可以在該濾波器組338處被轉換成時域。因此，合成訊號336的版本340(在時域中或在濾波後)可以用於音訊再現(譬如通過數個揚聲器)。

【0241】 為了獲得該混合規則(譬如混合矩陣)403，該原始訊號的聲道位準及相關資訊(譬如 C_y 、 C_{yR} 等)及與該降混訊號相關聯的協方差資訊(譬如 C_x)可以被提供給該混合規則計算器402。為了這個目標，利用該編碼器200在該旁側資訊228中編碼該聲道位準及相關資訊220是可行的。

【0242】 然而，在某些情況下，為了減少在該位元流248中被編碼的資訊的數量，不是所有參數都由該編碼器200編碼(譬如不是該原始訊號212的整個聲道位準及相關資訊及/或不是該降混訊號246的整個協方差資訊)。因此，一些參數318將在該參數重建模組316處被估計。

【0243】 該參數重建模組316可以例如被饋送以下內容中的至少一個：
該降混訊號246(x)的一版本322，其可以是例如該降混訊號246的一濾波版本(filtered version)或一FD版本；及
該旁側資訊228(包括聲道位準及相關資訊228)。

【0244】 該旁側資訊228可以包括與該原始訊號(212、y)的關聯矩陣 C_y 相關聯的資訊(作為該輸入訊號的位準暨相關資訊)：然而，在某些情況下，並非該關聯矩陣 C_y 的所有元素都被實際編碼。因此，估計及重建技術已經被開發用於重建該相關聯矩陣 C_y 的一版本(C_{yR})(譬如通過獲得的一估計版本 $\widehat{C_y}$ 的諸多中間步驟)。

【0245】 被提供給該模組316的該數個參數314可以由該熵解碼器312(輸入介面)獲得並且可以例如被量化。

【0246】 第3c圖顯示一解碼器300的一示例，該解碼器可以是第1至3b圖的諸多解碼器中的一個的一實施例。在此，該解碼器300包括由該解多工器表示的一輸入介面312。該解碼器300輸出一合成訊號340，該合成訊號例如可以在TD中(訊號340)，要被諸多揚聲器回放或在FD中(訊號336)。第3c圖的該解碼器300可以包括一核心解碼器(core decoder)347，該核心解碼器347也可以是該輸入介面312的一部分。該核心解碼器347因此可以提供該降混訊號 x 、246。一濾波器組320可以將該降混訊號246從TD轉換為FD。該降混訊號 x 、246的該FD版本以324被指示。該FD降混訊號324可以被提供給一協方差合成塊388。該協方差合成塊388可以在FD中提供該合成訊號336(Y)。一逆濾波器組(inverse filterbank)338可以轉換在其TD版本340的該音訊訊號314。該FD降混訊號324可以被提供給一頻帶/時隙分組塊(band/slot grouping block)380。該頻帶/時隙分組塊380可以進行在編碼器中由第5及2d圖的該分區分組塊265已經進行的相同操作。在該編碼器中，作為第5及2d圖的該降混訊號216的該數個頻帶已經在少數頻帶(具備較寬的寬度)中被分組或被聚合，並且該數個參數220(數個ICC，數個ICLD)已與數個聚合頻帶群組相關聯，現在有必要以相同的方式聚合該被解碼的降混訊號，將每個聚合頻帶給一相關參數。因此，標號385意指已經已被聚合後的該降混訊號 X_B 。要被注意的是，該濾波器提供未聚合的FD表徵(unaggregated FD representation)，以便能夠如在該編碼器中的相同方式在該解碼器(380)中將該頻帶/時隙進行分組以處理該數個參數，進行作為該編碼器在該頻帶/時隙上的相同聚合，以提供該被聚合的降混 X_B 。

【0247】該頻帶/時隙分組塊380還可以在一訊框中的不同時隙上聚合，使得該訊號385也以類似於該編碼器的時隙尺寸被聚合。該頻帶/時隙分組塊380還可以接收在該位元流248的該旁側資訊228中被編碼的資訊261，該資訊261指示暫態的存在，並且視情況還指示該暫態在該訊框內的位置。

【0248】在協方差估計塊384處，該降混訊號246(324)的協方差 C_x 被估計。該協方差 C_y 在該協方差計算塊386處被獲得，譬如通過利用公式(4)至(8)可用於此目的。第3c圖顯示一“多聲道參數(multichannel parameter)”，其可以是例如該數個參數220(數個ICC及數個ICLD)。然後將該數個協方差 C_y 及 C_x 提供給該協方差合成塊388，以合成該合成訊號388。在某些示例中，該數個塊384、386及388在一起實施時，該參數重建316及該混合都將被計算402，並且該合成處理器404將如上文及下文所討論的。

【0249】 4 討論(Discussion)

【0250】 4.1 概述(Overview)

【0251】本示例的新穎方法尤其旨在以低位元率(意謂著等於或低於160 kbits/sec)進行多聲道內容的編碼及解碼，同時保持一音質盡可能接近該原始訊號並保存該多聲道訊號的諸多空間特性。該新穎方法的一種功能還在於適合前面提到的該DirAC框架(architecture)。該輸出訊號可以在與該輸入212相同的揚聲器設置上被渲染(rendered)，也可以在不同的揚聲器設置上被渲染(就揚聲器而言，可以更大或更小)。同樣，該輸出訊號可以使用雙耳渲染(binaural rendering)在揚聲器上被渲染。

【0252】當前部分將提供對本發明以及組成本發明的不同模組的深入描述。

【0253】 該被提議的系統由兩個主要部分組成：

- 該編碼器200，其從該輸入訊號212中導出數個必要參數220，對它們進行量化(在222處)並對它們進行編碼(在226處)。該編碼器200還可以計算將在該位元流248中被編碼的該降混訊號246(並且可以被發送到該解碼器300)。
- 該解碼器300，其使用該數個被編碼的(譬如被發送的)參數及一降混訊號246，以便產生品質盡可能接近該原始訊號212的一多聲道輸出。

【0254】 第1圖顯示根據一示例提議的新穎方法的一概述。請注意，某些示例將僅使用在總體圖式中所示的該數個構造塊的一子集，並取決於應用場景捨棄某些處理塊。

【0255】 本發明的輸入212(y)是在時域或時頻(time-frequency)域中的一多聲道音訊訊號212(也被稱為“多聲道流(multichannel stream)”) (譬如訊號216)，例如：一組音訊訊號由一組揚聲器產生或意謂著要被播放。

【0256】 該處理的第一部分是該編碼部分；從該多聲道音訊訊號，將計算出一個所謂的“降混(down-mix)”訊號246(請參見4.2.6)連同(along with)一參數集或旁側資訊228(請參見4.2.2及4.2.3)，其是從在時域或頻域中的該輸入訊號212被導出的。這些參數將被編碼(請參見4.2.5)，並視情況被發送到該解碼器300。

【0257】 然後可以將該降混訊號246及該編碼參數228發送到一核心編碼器及一傳輸渠道(transmission canal)，該傳輸渠道鏈接該過程的該編碼器側與該解碼器側。

【0258】 在該解碼器側，該降混訊號被處理(4.3.3及4.3.4)且該數個被傳送的參數被解碼(請參見4.3.2)該數個被解碼的參數將被用於使用協方差合成(請參見4.3.5)進行該輸出訊號的合成，這將導致在時域中最終的多聲道輸出訊號。

【0259】 在詳細介紹之前，需要建立一些一般特徵，該一般特徵中的至少一個是有效的：

- 該處理可以與任何揚聲器設置一起使用。請記住，當增加揚聲器的數量時，該處理的複雜性以及對該數個被傳輸的參數進行編碼所需的位元也會增加。
- 整個處理可以在一訊框的基礎上完成，即，該輸入訊號212可以被劃分成被獨立處理的數個訊框。在該編碼器側，每個訊框將產生一參數集，該些參數將被傳送到該解碼器側以被處理。
- 一訊框也可以被劃分為數個時隙；這些時隙然後呈現出無法以一訊框比例(frame scale)獲得的諸多統計屬性。一訊框可以被劃分為例如八個時隙，並且每個時隙的長度將等於訊框長度的1/8。

【0260】 4.2 編碼器

【0261】 該編碼器的目的是抽取數個適當的參數220以描述該多聲道訊號212，對它們進行量化(在222處)，對它們進行編碼(在226處)作為旁側資訊228，然後視情況將它們發送到該解碼器側。在此將詳細描述該數個參數220以及如何計算它們。

【0262】 該編碼器200的一更詳細的方案可以在第2a至2d圖中找到。此概述突顯出該編碼器的兩個主要輸出228及246。

【0263】 該編碼器200的該第一輸出是從該多聲道音訊輸入212計算出的該降混訊號228；該降混訊號228是在比該原始內容(212)更少的聲道上的該原始多聲道流(訊號)的一表徵(representation)。有關它的計算的更多資訊，請參見第4.2.6節。

【0264】 該編碼器200的該第二輸出是被表示為在該位元流248中的旁側資訊228的該數個被編碼的參數220；這些參數220是本示例的一關鍵點：它們是將被用於在該解碼器側有效描述該多聲道訊號的諸多參數。這些參數220提供在位元流248中對它們進行編碼所需的品質及位元數量之間的一良好權衡。在該編碼器側，該參數計算可以被分成幾個步驟完成；該過程將在頻域中被描述，但也可以在時域中進行。該數個參數220首先從該多聲道輸入訊號212被估計，然後它們在該量化器222處被量化，然後它們可以被轉換為一數位的位元流248作為旁側資訊228。有關這些步驟的更多資訊，請參見第4.2.2、4.2.3及4.2.5節。

【0265】 4.2.1 濾波器組及分區分組(Filter bank & Partition Grouping)

【0266】 針對該編碼器側(譬如濾波器組214)或該解碼器側(譬如濾波器組320及/或338)討論濾波器組。

【0267】 本發明可以在處理期間的各個點處使用諸多濾波器組。這些濾波器組可以將一訊號從時域轉換到頻域(所謂的聚合頻帶或參數頻帶)，在這種情況下稱為“分析濾波器組(analysis filter bank)”，也可以從頻率轉換到時域(例如338)，在這種情況下稱為“合成濾波器組(synthesis filter bank)”。

【0268】 該濾波器組的選擇必須符合所需的性能及最佳化要求，但是其餘的處理可以獨立於一特定選擇的濾波器組而被進行。例如：使用基於正交鏡濾

波器的一濾波器組(a filter bank based on quadrature mirror filters)或一基於短時傅立葉變換的濾波器組(Short-Time Fourier transform based filter bank)。

【0269】 參照第5圖，該編碼器200的該濾波器組214的輸出將是在一定數量的頻帶(266相對於264)上表示的在頻域中的一訊號216。對於所有頻帶(264)進行其餘處理可以被理解為提供一更好的品質及一更好的頻率解析度，但是還需要更重要的位元率以傳輸所有資訊。因此，連同該濾波器組處理一所謂的“分區分組(partition grouping)”(265)，其對應於將某些頻率分組在一起，以便在一較小的頻帶群組表示資訊266。

【0270】 例如：該濾波器263的該輸出264(第5圖)可以被表示在128個頻帶，並且在265處的分區分組可以導致一訊號266(216)僅具備20個頻帶。有幾種將數個頻帶分組在一起的方法，一種有意義的方法可以是例如嘗試近似成等效矩形頻寬(equivalent rectangular bandwidth)。該等效矩形頻寬是一種心理聽覺激勵的頻帶劃分(a type of psychoacoustically motivated band division)，其試圖模型化(model)人類聽覺系統如何處理音訊事件，即，目的是以適合人類聽覺的方式對該濾波器組進行分組。

【0271】 4.2.2 參數估計(譬如估計器218)

【0272】 方面1：使用諸多協方差矩陣描述及合成多聲道內容

【0273】 在218處的參數估計是本發明的要點之一；它們在該解碼器側被用於合成該輸出的多聲道音訊訊號。那些參數220(被編碼為旁側資訊228)已被選擇，因為它們有效地描述該多聲道輸入流(訊號)212，並且它們不需要傳輸大量資料。這些參數220在該編碼器側被計算，並且稍後與在該解碼器側的該合成引擎被共同使用以計算該輸出訊號。

【0274】 在此，該數個協方差矩陣可以在該多聲道音訊訊號與該降混訊號的該數個聲道之間被計算。意即：

C_y ：該多聲道流(訊號)的協方差矩陣，及/或

C_x ：該降混流(訊號)246的協方差矩陣

【0275】 該處理可以在一參數頻帶的基礎上進行，因此，一個參數頻帶與另一個參數頻帶無關，並且可以在不損失概括性的情況下對於一給定的參數頻帶描述諸多公式。

【0276】 對於一給定的參數頻帶，該數個協方差矩陣被定義如下：

$$\begin{aligned} C_y &= \Re\{Y_B Y_B^*\} \\ C_x &= \Re\{X_B X_B^*\} \end{aligned} \quad (1)$$

其中

- $\Re$ 表示該實部運算符。
- 除了實部，它可以是導致一實際值具有與衍生自(譬如絕對值)的複數值的一關係的任何其他運算。
- $*$ 表示該共軛轉置運算符。
- B 表示在數個頻帶的原始數量與該數個被分組的頻帶之間的關係(有關分區分組，請參見4.2.1)。
- Y 及 X 分別是在頻域中的該原始多聲道訊號212及該降混訊號246。

【0277】 C_y (或其元素，或從 C_y 或從其元素獲得的諸多值)也被指示作為該原始訊號212的聲道位準及相關資訊(channel level and correlation information)。

C_x (或其元素，或從 C_y 或從其元素獲得的諸多值)，也被指示作為與該降混訊號212相關聯的協方差資訊。

【0278】對於一給定的訊框(及頻帶)，僅一個或兩個協方差矩陣 C_y 及/或 C_x ，可以譬如被估計器塊218輸出。該過程是基於時隙(slot-based)而不是基於訊框(frame-based)，關於在一給定時隙與對於整個訊框的數個矩陣之間的關係，可以採用不同的實現方式。例如：可以為在一訊框內的每個時隙計算該(數個)協方差矩陣並對它們求和(sum them)，以便為一個訊框輸出的該數個矩陣。注意的是，用於計算該數個協方差矩陣的定義是數學上的定義，但是如果希望獲得具備諸多特定特性的一輸出訊號，則事先計算或至少修改那些矩陣也是可行的。

【0279】如上所述，該(諸多)矩陣的所有元素 C_y 及/或 C_x 不必實際被編碼在該位元流248的該旁側資訊228中。對於 C_x ，從通過應用公式(1)被編碼的該降混訊號246簡單地估計它是可行的，並且因此該編碼器200可以容易地避免短暫的(tout-court)，對 C_x (或者更通常地，關於與該降混訊號相關聯的協方差資訊)的任何元素進行編碼。對於 C_y (或針對與該原始訊號相關聯的該聲道位準及相關資訊)，使用以下討論的技術在該解碼器側估計 C_y 的數個元素中的至少一個是可行的。

【0280】方面2a：傳輸該數個協方差矩陣及/或能量以描述及重建一多聲道音訊訊號

【0281】如前所述，數個協方差矩陣被用於該合成。將那些協方差矩陣(或它的一子集)從該編碼器直接傳送到該解碼器是可行的。

【0282】在某些示例中，該矩陣 C_x 不一定必需被傳送，由於可以使用該降混訊號246在該解碼器側再次計算矩陣，但是取決於應用情景，此矩陣可能需作為一被發送的參數。

【0283】 從一實現的觀點來看，那些矩陣 C_x 、 C_y 中的所有值並非必須被編碼或被傳送，譬如以便滿足關於位元率的某些特定要求。該數個未被傳送的值可以在該解碼器側被估計(請參見4.3.2)。

【0284】 方面2b：傳輸聲道間同調度及聲道間位準差以描述及重建一多聲道訊號

【0285】 根據該數個協方差矩陣 C_x 、 C_y ，一組備用參數可以被定義，並被用於在該解碼器側重建該多聲道訊號212。這些參數可以是，例如：該聲道間同調度(ICC)及/或聲道間位準差(ICLD)。

【0286】 該聲道間同調度描述在該多聲道流的每個聲道之間的該同調度。該參數可以從該協方差矩陣 C_y 被導出，並按以下方式計算(對於一給定的參數頻帶及兩個給定的聲道*i*及*j*)：

$$\xi_{i,j} = \frac{c_{y_{i,j}}}{\sqrt{c_{y_{i,i}} \cdot c_{y_{j,j}}}} \quad (2)$$

其中

- $\xi_{i,j}$ 在該輸入訊號212的數個聲道*i*與*j*之間的該ICC
- $c_{y_{i,j}}$ 在該輸入訊號212的數個聲道*i*與*j*之間的該多聲道訊號的先前被定義在公式(1)中的該協方差矩陣中的該數個值

【0287】 該數個ICC值可以在該多聲道訊號的每個聲道之間被計算，隨著該多聲道訊號大小的增長，這可能導致大量資料。實際上，一組被減少的ICC可以被編碼及/或被發送。在某些示例中，必須根據該性能要求來定義被編碼及/或被傳送的該數個值。

【0288】 例如：當處理由一5.1(或5.0)定義的揚聲器設置，如ITU薦議“ITU-R BS. 2159-4”，則選擇僅發送四個ICC是可行的。這四個ICC可以是在以下內容之間的一個：

- 中央及右聲道
- 中央及左聲道
- 左與左環繞道
- 右與右環繞道

【0289】 通常，從ICC矩陣中選擇的ICC的索引由ICC映像描述。

【0290】 通常，對於每個揚聲器設置，可以選擇平均給出最佳品質的一組固定的ICC，以被編碼及/或被傳送到該解碼器。該ICC數及那些ICC要被發送可以取決於該揚聲器設置及/或可用的總位元率，並且在該編碼器及該解碼器上均可用，而無需在該位元流248中傳輸該ICC映像。換句話說，譬如取決於該揚聲器設置及/或該總位元率，一組固定的ICC及/或一相應的固定的ICC映像可以被使用。

【0291】 此固定的組可能不適用於特定材料，並且在某些情況下，使用一組固定的ICC產生比所有材料的平均品質明顯差的品質。為了在另一個示例中針對每個訊框(或時隙)克服這一點，可以基於某個ICC的重要性的特徵來估計一組最佳的ICC及一對應的ICC映像。然後，將被用於當前的訊框的該ICC映像與被量化的ICC一起在該位元流248中明確地編碼及/或傳送。

【0292】 例如：可以通過使用來自公式(1)的降混協方差 C_x 產生協方差 $\widehat{C}_y$ 的估計類似於使用來自4.3.2的公式(4)及(6)的該解碼器來產生該ICC矩陣 $\widehat{\xi}_{i,j}$ 的估計，來決定一ICC重要性的特徵。取決於所選擇的特徵，該特徵針對每個ICC或

在該協方差矩陣中用於每個頻帶的對應的元進行計算，對於那些參數將在該當前的訊框中被發送並對於所有頻帶進行組合。然後，該被組合的特徵矩陣被用於決定數個最重要的ICC，從而決定要被使用的該組ICC及要被發送的該ICC映像。

【0293】 例如：一ICC的重要性的特徵是在該被估計的協方差 $\widehat{C}_y$ 與該實際的協方差 C_y 的數個元之間的絕對誤差，而該被組合的特徵矩陣是在當前的訊框中要在所有頻帶上被傳送的每個ICC的絕對誤差之總和。從該被組合的特徵矩陣中，該 n 個元被選擇，其中該被求和的絕對誤差是最高的， n 是要針對揚聲器/位元率組合被發送的ICC數，並從這些元建構該ICC映像。

【0294】 此外，在如第6b圖所示的另一個示例中，為了避免在數個訊框之間的ICC映像改變太多，對於前一個參數訊框的該所選的ICC映像中的每個元，該特徵矩陣可以被強調，例如：在該協方差的該絕對誤差的情況，通過將一係數 $> 1(220k)$ 應用於該先前的訊框的該ICC映像的該數個元。

【0295】 此外，在另一示例中，在該位元流248的該旁側資訊228中被發送的一旗標可以指示在當前的訊框中是否使用該固定的ICC映像或該最佳的ICC映像，並且如果該旗標指示該固定的組，則不在該位元流248中傳送該ICC映像。

【0296】 最佳ICC映像例如被編碼及/或被發送作為一位元映像(譬如該ICC映像可以實施第6a圖的資訊254')。

【0297】 用於傳送該ICC映像的另一個示例是將該索引傳送到所有可能的ICC映像的一表中，其中該索引本身是例如被附加地熵編碼的。例如：該所有可能的ICC映像的表沒有被儲存在記憶體中，但是由該索引指示的該ICC映像從該索引被直接計算。

【0298】 可以與該ICC共同被發送的一第二參數(或單獨的)是該數個ICLD。“ICLD”代表聲道間位準差(Inter-channel level difference)，並且它描述在該輸入的多聲道訊號212的每個聲道之間的能量關係。該ICLD沒有唯一的定義；此值的重要方面是它描述在該多聲道流內的諸多能量比。

作為一示例，來自數個ICLD的轉換 C_y 可以被獲得如下：

$$\chi_i = 10 \cdot \log_{10} \left(\frac{P_i}{P_{\text{dmx},i}} \right) \quad (3)$$

其中：

- χ_i 用於聲道 i 的 ICLD。
- P_i 當前聲道 i 的功率，可以從 C_y 的對角線： $P_i = C_{y,i,i}$ 中抽取。
- $P_{\text{dmx},i}$ 取決於該聲道 i ，但將始終是在 C_x 的數個值的一線性組合，它還取決於該原始揚聲器設置。

【0299】 在諸多示例中， $P_{\text{dmx},i}$ 並非每個聲道都相同，而是取決於與該降混矩陣(也是用於該解碼器的該原型矩陣)相關的一映像(mapping)，這通常在公式(3)下的諸多要點中的一個被提到。取決於是否僅將該聲道 i 降混到該數個降混聲道中的一個降混聲道或在它們之中的一個以上。換句話說，在該降混矩陣中存在一非零元素的情況下， $P_{\text{dmx},i}$ 可能為或包括 C_x 的所有對角元素的總和，因此公式(3)可以重寫為：

$$\chi_i = 10 \cdot \log_{10} \left(\frac{P_i}{P_{\text{dmx},i}} \right)$$

$$P_{\text{dmx},i} = \alpha_i \sum_j C_{x,j,j}, j \in \{Q_{j,i} \neq 0\}$$

$$P_i = C_{y,i,i}$$

其中， α_i 是與一聲道對該降混的該預期能量貢獻相關的一加權因子，此加權因子對於一特定的輸入揚聲器配置是固定的，並且在編碼器及解碼器處都是已知的。該矩陣Q的概念將在下面被提供。在文件的最後部分還提供 α_i 及數個矩陣Q的一些值。

【0300】 在一實現為每個輸入聲道i定義一映像的情況下，其中該映像索引是該降混的該聲道j，該輸入聲道i僅被混到其中，或者如果該映像索引大於該降混聲道數。因此，我們有一映像索引(mapping index) $m_{ICLD,i}$ ，用於以如下方式決定 $P_{dmx,i}$ ：

$$P_{dmx,i} = \begin{cases} \alpha_i C_x m_{ICLD,i}, m_{ICLD,i} \leq n_{DMX} \\ \alpha_i \sum_{j=1}^{n_{DMX}} C_x j, m_{ICLD,i} > n_{DMX} \end{cases}$$

【0301】 4.2.3 參數量化(Parameter Quantization)

【0302】 為了獲得數個量化參數224，該參數220的量化的諸多示例可以例如由第2b及4圖的該參數量化模組222進行。

【0303】 一旦該參數集220被計算出，意謂著該數個協方差矩陣 $\{C_x, C_y\}$ 或該數個ICC及數個ICLD $\{\xi, \chi\}$ ，它們被量化。該量化器的選擇可以在品質與要被傳輸的資料量之間進行一權衡，但是關於該被使用的量化器是沒有限制的。

【0304】 作為一示例，在使用該數個ICC及數個ICLD的情況下；針對該數個ICC，一非線性量化器可以在間隔 $[-1,1]$ 包含10個量化步階(quantization steps)，而針對該數個ICLD，另一個非線性量化器可以在間隔 $[-30,30]$ 包含20個量化步階。

【0305】 同樣，作為一實現最佳化方案，選擇對數個要被傳送的參數進行降取樣是可行的，意謂該數個被量化參數224被使用在一列中的兩個或更多個訊框。

【0306】 在一方面，在當前的訊框中被發送的參數的子集由在該位元流中的一參數訊框索引以訊號表明。

【0307】 4.2.4 暫態處理、降取樣參數

【0308】 下文討論的某些示例可以理解為被顯示在第5圖中，其又可以是第1及2d圖的塊214的一示例。

【0309】 在降取樣的參數集的情況下(譬如在第5圖中的塊265處獲得的)，即，用於數個參數頻帶的一子集的一參數集220可以被用於一個以上的被處理的訊框，出現在一個以上的子集中的數個暫態就本地化及同調度(localization and coherence)而言是無法被保留的。因此，在這樣的一訊框中發送所有頻帶的諸多參數可能是有利的。這種特殊類型的參數訊框可以例如通過在該位元流中的一旗標被發訊表明。

【0310】 在一方面，在258處的一暫態檢測被用於檢測在該訊號212中的這樣的數個暫態。該暫態在當前的訊框中的位置也可以被檢測。時間粒度(granularity)可以有利地鏈接到所使用的濾波器組214的時間粒度，從而每個暫態位置可以對應於該濾波器組214的一個時隙或數個時隙的一群組。然後，基於該暫態位置來選擇用於計算該數個協方差矩陣 $\mathbf{C}_y$ 及 $\mathbf{C}_x$ 的諸多時隙，例如僅使用從包含該暫態的時隙到當前的訊框結束。

【0311】 該暫態檢測器(或暫態分析塊258)可以是還被用於該降混訊號212進行編碼的一暫態檢測器，例如：一IVAS核心編碼器的時域暫態檢測器。因此，第5圖的示例還可以在該降混計算塊244的上游被應用。

【0312】 在一個示例中，使用一個位元對一暫態的發生進行編碼(諸如：“1”，意謂“在該訊框中存在暫態”與“0”，意謂“在該訊框中沒有暫態”)，如果另外檢測到一暫態，則該暫態的位置被編碼及/或作為在該位元流248中的被編碼的欄位261(關於該暫態的資訊)被發送，以允許在該解碼器300中進行一類似的處理。

【0313】 如果檢測到一暫態並且進行所有頻帶的傳輸(譬如以訊號表明)，則使用該正常的分區分組發送該參數220可能會導致該傳輸參數220作為在該位元流248中的旁側資訊228所需的資料速率的一尖峰。此外，該時間解析度比該頻率解析度更重要。因此，在塊265處，將用於這樣的一訊框的該分區分組改變為具有更少的頻帶以發送(譬如從在該訊號版本264中的許多頻帶到在該訊號版本266中的較少頻帶)可能是有利的。一個示例採用這種不同的分區分組，例如通過將所有頻帶上的兩個相鄰頻帶群組合為該數個參數的一正常的降取樣因子為2。一般而言，一暫態的發生暗示該數個協方差矩陣本身可以被預期為在該暫態之前及之後有極大的不同。為了避免在該暫態以前的數個時隙出現諸多偽影(artifacts)，可以僅考慮該暫態時隙本身以及後續的所有時隙，直到該訊框結束為止。這也基於假設，即，該訊號事先足夠穩定，並且有可能使用資訊及諸多混合規則，這些資訊及混合規則是針對先前的訊框導出的，也適用於該暫態以前的諸多時隙。

【0314】 總而言之，該編碼器可以被配置成決定該暫態已經發生在該訊框的哪個時隙中，並且對與該暫態已經發生的時隙及/或在該訊框中的後續時隙相

關聯的該原始訊號(212、y)的該聲道位準及相關資訊(220)進行編碼，而無需對與該暫態以前的時隙相關聯的該原始訊號(212、y)的聲道位準及相關資訊(220)進行編碼。

【0315】 類似地，當在一個訊框中的該暫態的存在及位置被以訊號表明(261)時，該解碼器可以(譬如在塊380處)：

將當前的聲道位準及相關資訊(220)與已經發生暫態的時隙及/或在該訊框中的後續時隙相關聯；及

將在已經發生該暫態的時隙以前的訊框的時隙與先前時隙的聲道位準及相關資訊(220)相關聯。

【0316】 該暫態的另一個重要方面是，在決定當前的訊框中存在暫態的情況下，不再對當前的訊框執行平滑操作。在一暫態的情況下，沒有對 C_y 及 C_x 進行平滑(smoothing)，但是來自當前的訊框的 C_{yR} 及 C_x 被用於該數個混合矩陣的計算。

【0317】 4.2.5 熵編碼(Entropy Coding)

【0318】 該熵編碼模組(位元流寫入器)226可以是最後的編碼器的模組；它的目的是將先前獲得的量化值轉換為一個二進制位元流，其也將被稱為“旁側資訊(side information)”。

【0319】 用於對該數個值進行編碼的方法可以例如是霍夫曼編碼(Huffmann coding) [6] 或差量編碼(delta coding)。該編碼方法不是至關重要的，將只會影響最終的位元率。一個人員應該取決於他想要達到的位元率來調適該編碼方法。

【0320】 幾種實現最佳化方案可以被執行以減小該位元流248的大小。作為一示例，一切換機制(**switching mechanism**)可以被實現，該切換機制取決於從一位元流大小的觀點來看哪個更有效以從一編碼方案切換到另一編碼方案。

【0321】 例如：這些參數可以沿一個訊框的頻率軸被進行差量編碼，並且由一範圍編碼器(**range coder**)對所得的增量索引熵的序列進行編碼。

【0322】 同樣，在該參數降取樣的情況下，也作為一示例，一種機制可以被實現，以每個訊框僅發送該數個參數頻帶的一子集，以便連續發送資料。

【0323】 這兩個示例需要數個訊號化位元(**signalization bits**)，以在該編碼器側以訊號表明該解碼器的特定處理方面。

【0324】 4.2.6 降混計算(**Down-mix Computation**)

【0325】 該處理的該降混部分244可以是簡單的，但是在某些示例中是至關重要的。在本發明中被使用的降混可以是一被動的(**passive**)降混，這意謂著在處理期間它的計算方式保持相同，並且在一給定時間與訊號或其特徵無關。然而，已經理解的是，在244處的降混計算可以被擴展到一主動的(**active**)降混計算(例如在 [7] 中所描述的)。

【0326】 該降混訊號246可以在兩個不同的位置被計算：

- 第一次在該編碼器側進行該參數估計(請參閱4.2.2)，因為它可能需要(在某些示例中)計算該協方差矩陣 C_x 。
- 第二次在該編碼器側，在該編碼器200與該解碼器300之間(在時域中)，該降混訊號246被編碼及/或被傳送到該解碼器300，並且被用於模組334處的該合成的一基礎。

【0327】 作為一示例，對於一5.1輸入的一立體聲降混，該降混訊號可以如下方式計算：

- 該降混的左聲道是該左聲道、該左環繞道及該中央聲道的總和。

【0328】 該降混的右聲道是該右聲道、該右環繞道及該中央聲道的總和。或者，在一5.1輸入為一單音降混(monophonic down-mix)的情況下，該降混訊號被計算為該多聲道流中的每個聲道的總和。

【0329】 在諸多示例中，該降混訊號246的每個聲道可以被獲得而作為該原始訊號212的該數個聲道的一線性組合，例如具備諸多常數參數，從而實現一被動降混(passive downmix)。

【0330】 根據該處理的需要，該降混訊號的計算可以被擴展並被適用於其他揚聲器設置。

【0331】 方面3：使用一被動降混及一低延遲濾波器組的低延遲處理

【0332】 本發明可以通過使用一被動降混例如先前針對一5.1輸入所描述的降混及一低延遲濾波器組來提供低延遲處理。使用這兩個元素，有可能在該編碼器200與該解碼器300之間實現低於5毫秒的延遲。

【0333】 4.3 解碼器(Decoder)

【0334】 該解碼器的目的是通過使用該被編碼的(譬如被傳送的)降混訊號(246、324)及該被編碼的旁側資訊228，在一給定的揚聲器設置上合成該音訊輸出訊號(336、340、 y_R)。該解碼器300可以在如被用於該輸入(212、 y)的揚聲器設置的相同揚聲器設置上或在不同的揚聲器設置上渲染該被輸出的音訊訊號(334、240、 y_R)。在不失一般性的前提下，將假定該輸入及輸出揚聲器設置是相

同(但在諸多示例中，它們可能不同)。在此部分中，將描述可以構成該解碼器300的不同模組。

【0335】 第3a及3b圖描繪可能的解碼器處理的一詳細概述。重要而要注意的是，取決於一給定應用的需要及要求，在第3b圖中的該數個模組中的至少一些(特別是具有虛線邊框的模組，例如320、330、338)可以被丟棄。該解碼器300可以輸入(譬如接收)來自該編碼器200的兩組資料：

- 具備數個被編碼的參數的該旁側資訊228(如4.2.2中所述)
- 該降混訊號(246、y)可以是在時域中(如4.2.6中所述)。

【0336】 該數個被編碼的參數228可能需要首先被解碼(譬如通過該輸入單元312)，譬如以先前被使用的該逆編碼方法。一旦完成此步驟，就可以重建用於該合成的相關參數，例如該數個協方差矩陣。並行地，可以通過幾個模組處理該降混訊號(246、x)：首先可以使用一分析濾波器組320(請參見4.2.1)以獲得該降混訊號246的一頻域版本324。然後，可以計算該原型訊號328(請參見4.3.3)，並且可以執行一附加的去相關步驟(在330處)(請參見4.3.4)。該合成的一關鍵點是該合成引擎334，其使用協方差矩陣(譬如在塊316處被重建)及該原型訊號(328或332)作為輸入，並且產生該最終訊號336作為一輸出(參見4.3.5)。最終，在一合成濾波器組338處的一最後步驟可以被完成(譬如如果該分析濾波器組320先前被使用)，則在時域中產生該輸出訊號340。

【0337】 4.3.1 熵解碼(Entropy Decoding)(譬如塊312)

【0338】 在塊312(輸入介面)處的該熵解碼可以允許獲得先前在4.2.3中獲得的該量化參數314。該位元流248的該解碼可以被理解為一直截了當的操作；可以根據在4.2.5中使用的該編碼方法讀取該位元流248，然後對它進行解碼。

【0339】 從一實現方案的觀點來看，該位元流248可以包含數個信令位元(signaling bits)，該些信令位元不是資料，但該些信令位元是指示在該編碼器側的進行處理的某些特殊性。

【0340】 例如：在該編碼器200具有在幾種編碼方法之間切換的可能性的情況下，所使用的兩個第一位元可以指示已經使用哪種編碼方法。接下來的位元也可以被用來描述當前正在傳送哪些參數頻帶。

【0341】 可以被編碼在該位元流248的該旁側資訊中的其他資訊可以包括一旗標，該旗標指示一暫態及指示在一訊框的哪個時隙中已經發生一暫態的欄位(field)261。

【0342】 4.3.2 參數重建

【0343】 參數重建可以例如由塊316及/或該混合規則計算器402進行。

【0344】 此參數重建的一目標是從該降混訊號246及/或從旁側資訊228(或以它被該量化參數314表示的版本)重建該數個協方差矩陣 C_x 及 C_y (或更通常地，與該降混訊號246相關聯的協方差資訊及該原始訊號的位準暨相關資訊)。這些協方差矩陣 C_x 及 C_y 對於該合成可能是必需的，因為它們是有效描述該多聲道訊號246的矩陣。

【0345】 在模組316處的該參數重建可以是一兩步驟過程：

首先，該矩陣 C_x (或更通常地，與該降混訊號246相關聯的該協方差資訊)是從該降混訊號246被重新計算(在與該降混訊號246相關聯的該協方差資訊實際上被編碼在位元流248的該旁側資訊228中的情況下，可以避免此步驟)；及

然後，該矩陣 C_y (或更通常地，該原始訊號212的該位準暨相關資訊)可以被恢復，譬如至少部分地使用該數個被傳送的參數及 C_x 或更通常地與該降混訊號246相關聯的該協方差資訊(在該原始訊號212的該位準暨相關資訊實際上被編碼在位元流248的該旁側資訊228中的情況下，可以避免此步驟)。

【0346】 注意的是，在某些示例中，對於每個訊框，使用具備當前的訊框以前的一被重建的協方差矩陣的一線性組合是可行的，譬如通過加法、平均等，以平滑當前的訊框的該協方差矩陣 C_x 。例如：在第 t 訊框，要被用於公式(4)的最終協方差可以考慮為先前的訊框重建的該目標協方差，譬如

$$C_{x,t} = C_{x,t} + C_{x,t-1}。$$

然而，在決定當前的訊框中的一暫態存在的情況下，不再對當前的訊框執行平滑操作。在一暫態的情況下，不使用當前的訊框進行任何平滑 C_x 。

【0347】 該過程的一概述可以在下面被找到。

【0348】 注意：至於該編碼器，在此的處理可以針對每個頻帶在一參數頻帶的基礎上被獨立完成，為了清楚起見，將僅針對一個特定頻帶描述該處理，並對標記法進行相應調適。

【0349】 方面4a：在該數個協方差矩陣被傳送的情況下重建數個參數

【0350】 對於此方面，假設在該旁側資訊228(與該降混訊號246相關聯的協方差矩陣及該原始訊號212的聲道位準及相關資訊)中的被編碼(譬如被傳送)的數個參數是該數個協方差矩陣(或它的一子集)，如在方面2a中所定義。然而，在某些示例中，與該降混訊號246相關聯的該協方差矩陣及/或該原始訊號212的該聲道位準及相關資訊可以由其他資訊來實施。

【0351】 如果完整的協方差矩陣 C_x 及 C_y 被編碼(譬如被傳送)，則在塊318處沒有進一步的處理要做(因此在這樣的示例中塊318可以被避免)。如果僅那些矩陣中的至少一個矩陣的一子集被編碼(譬如被傳送)，則必須估計該數個缺失值。如在該合成引擎334中(或更具體地在該合成處理器404中)被使用的最終協方差矩陣將在該解碼器側由該數個被編碼的(譬如被傳送的)值228及該數個被估計的值組成。例如：如果僅該矩陣 C_y 的一些元素被編碼在該位元流248的該旁側資訊228中，則 C_y 的剩餘元素在此被估計。

【0352】 對於該降混訊號246的該協方差矩陣 C_x ，通過在該解碼器側使用該降混訊號246以計算該數個缺失值並應用公式(1)是可行的。

【0353】 在一方面，其中一暫態的發生及位置被傳送或被編碼，如同在該編碼器側使用相同的數個時隙用於計算該降混訊號246的該協方差矩陣 C_x 。

【0354】 對於該協方差矩陣 C_y ，可以按以下方式以第一估算計算數個缺失值：

$$\widehat{C}_y = Q C_x Q^* \quad (4)$$

其中：

- $\widehat{C}_y$ 該原始訊號212的該協方差矩陣的一估計(這是該原始聲道位準及相關資訊的估計版本的示例)
- Q 所謂的原型矩陣(原型規則、估計規則)，它描述在該降混訊號與該原始訊號之間的關係(請參見4.3.3)(這是原型規則的一示例)
- C_x 該降混訊號的該協方差矩陣(這是該降混訊號212的協方差資訊的示例)
- $*$ 標示該共軛轉置

【0355】一旦這些步驟被完成後，該協方差矩陣將再次被獲得，並可以被用於最終合成。

【0356】方面4b：在該數個ICC及該ICLD被傳送的情況下重建數個參數

【0357】對於此方面，可以假設在旁側資訊228中的該數個被編碼的(譬如被傳送的)參數是在方面2b中被定義的該數個ICC及數個ICLD(或它們的一子集)。

【0358】在此情況下，可能首先需要重新計算該協方差矩陣 C_x 。這可以使用在該解碼器側的該降混訊號212並應用公式(1)來完成。

【0359】在一方面，其中一暫態的發生及位置被傳送，如同該編碼器中使用相同時隙用於計算該降混訊號的該協方差矩陣 C_x 。然後，該協方差矩陣 C_y 可以從該數個ICC及數個ICLD被重新計算；此操作可以被進行如下：

【0360】該多聲道輸入的每個聲道的能量(也被稱為位準)可以被獲得。使用傳輸的聲道間位準差及以下公式得出這些能量

$$P_i = P_{dmx,i} \cdot 10^{\frac{X_i}{10}} \quad (5)$$

其中

$$P_{dmx,i} = \alpha_i \sum_j C_{x,j,j}, j \in \{Q_{j,i} \neq 0\}$$

$$P_i = C_{y_{i,i}}$$

其中， α_i 是關於一聲道對該降混的預期能量貢獻的加權因子，此加權因子對於某些輸入的揚聲器配置是固定的，並且在編碼器及解碼器處均為已知。在一實現為對於每個輸入的聲道 i 定義一映像的情況下，其中該映像索引是該降混的該聲

道 j ，僅將該輸入聲道 i 混到其中，或者如果該映像索引大於該降混聲道數。因此，我們有一個映像索引， $m_{ICLD,i}$ 其被用於利用以下方式決定 $P_{dmx,i}$ ：

$$P_{dmx,i} = \begin{cases} \alpha_i C_x m_{ICLD,i}, m_{ICLD,i} \leq n_{DMX} \\ \alpha_i \sum_{j=1}^{n_{DMX}} C_{x,j}, m_{ICLD,i} > n_{DMX} \end{cases}$$

這些符號與4.2.3中的該參數估計中被使用的符號相同。

【0361】 這些能量可以被用來正規化(normalize)該被估計的 C_y 。在不是所有的ICC都從該編碼器側被傳送的情況下，可以針對該數個未被傳送的值計算 C_y 的一估計。該被估計的協方差矩陣 $\widehat{C_y}$ 可以使用公式(4)以該原型矩陣 Q 及該協方差矩陣 C_x 被獲得。

【0362】 該協方差矩陣的此估計導致該ICC矩陣的一估計，為此，該索引 (i,j) 的項可以由下式給出：

$$\widehat{\xi}_{i,j} = \frac{\widehat{C_{y_{i,j}}}}{\sqrt{\widehat{C_{y_{i,i}}} \widehat{C_{y_{j,j}}}}} \quad (6)$$

因此，“重建(reconstructed)”矩陣可以被定義如下：

$$\xi_{R,i,j} = \begin{cases} \xi_{i,j} & \text{若 } (i,j) \in \{\text{被傳送的指標}\} \\ \widehat{\xi}_{i,j} & \text{或者否則} \end{cases} \quad (7)$$

其中：

- 該下標 R 指示該重建矩陣(其是該原始位準暨相關資訊的重建版本的一示例)

- 該集合體(ensemble){被傳送的指標(transmitted indices)}對應於在該旁側資訊 228 中已經被解碼(譬如從該編碼器被傳送到該解碼器)的所有該 (i, j) 對。

【0363】在諸多示例中，通過 $(by)\widehat{\xi_{i,j}}$ 不如該被編碼的值 $\xi_{i,j}$ 準確，因此 $\xi_{i,j}$ 可能比 $\widehat{\xi_{i,j}}$ 更可取。

【0364】最後，由此被重建的ICC矩陣，該被重建的協方差矩陣可以被推論 C_{y_R} 。此矩陣可以通過將公式(5)中獲得的能量應用於該被重建的ICC矩陣而被獲得，因此可以得到該數個指標 (i, j) 如下：

$$C_{y_{R_{i,j}}} = \xi_{R_{i,j}} \cdot \sqrt{P_i \cdot P_j} \quad (8)$$

在完整的ICC矩陣被傳送的情況下，僅需要公式 (5)及 (8)。前面的段落描述一種重建該缺失參數的方法，其他方法可以被使用，並且所提出的方法不是唯一的。從使用一 5.1 訊號的方面 1b 的示例中，可被注意的是，該數個未被傳送的值是在該解碼器側需要被估計的數個值。

【0365】現在可以得到該數個協方差矩陣 C_x 及 C_{y_R} 。重要的是要詮釋該重建矩陣 C_{y_R} 可以是該輸入訊號212的該協方差矩陣 C_y 的一估計。本發明的權衡可以是使在該解碼器側的該協方差矩陣的該估計與該原始的足夠接近，但也要傳送盡可能少的參數。這些矩陣對於4.3.5中描述的最終合成可能是必備的。

【0366】注意的是，在某些示例中，對於每個訊框，可以使用與在當前的訊框以前的一被重建的協方差矩陣的一線性組合以平滑該當前的訊框的該被重建的協方差矩陣，例如通過加法、平均等。例如：在第t訊框，要用於該合成的該最終協方差可以考慮為該先前的訊框重建的該目標協方差，譬如

$$C_{y,t} = C_{y_R,t} + C_{y_R,t-1}$$

然而，在一暫態的情況下，沒有平滑被完成，並且用於該當前的訊框的 C_{y_R} 被用於該混合矩陣的計算。

【0367】 還應注意的是，在某些示例中，對於每個訊框，該數個降混聲道 C_x 的該未平滑的協方差矩陣用被於參數重建，而如第4.2.3節所述的一平滑的協方差矩陣 $C_{x,t}$ 被用於該合成。

【0368】 第8a圖在該解碼器300處恢復用於獲得該數個協方差矩陣 C_x 及 C_{y_R} 的操作(譬如在塊386或316...處被進行的)。在第8a圖的數個塊中，還在括號之間指示特定的塊所採用的公式。可以看出，通過公式(1)，該協方差估計器384允許達成該降混訊號324(或它的降頻版本385)的該協方差 C_x 。通過使用公式(4)及適當類型的規則 Q ，該第一協方差估計器塊384'允許達成該協方差 C_y 的第一估計 $\widehat{C}_y$ 。後續，通過應用公式(6)，一協方差對同調度塊(covariance-to-coherence block)390獲得該數個同調度 $\hat{\xi}$ 。後續，一ICC替換塊(ICC replacement block)392通過採用公式(7)，在該數個被估計的ICC($\hat{\xi}$) 及在該位元流348的該旁側資訊228中被以訊號表明的該ICC)之間進行選擇。然後將所選擇的數個同調度 ξ_R 輸入到一能量施加塊(energy application block)394，該能量施加塊394根據該ICLD(χ_i)施加能量。然後，該目標協方差矩陣 C_{y_R} 被提供給第3a圖的該混合器規則計算器402或該協方差合成塊388，或第3c圖的該混合器規則計算器或第3b圖的一合成引擎344。

【0369】 4.3.3 原型訊號計算(塊326)

【0370】 該原型訊號模組326的一目的是以能夠被合成引擎334使用的方式成形該降混訊號212(或它的頻域版本324)(請參見4.3.5)。該原型訊號模組326

可以進行該降混訊號的一升混(upmixing)。該原型訊號模組326可以通過將該降混訊號212(或324)乘以所謂的原型矩陣Q以完成該原型訊號328的計算：

$$Y_p = XQ \quad (9)$$

其中

- Q為該原型矩陣(其是原型規則的一示例)
- X為該降混訊號(212或324)
- Y_p 為該原型訊號(328)。

【0371】 建立該原型矩陣的方式可能是與處理相依的 (processing-dependent)，並且可以被定義為滿足應用程式的要求。唯一的限制可能是該原型訊號328的聲道數必須與該期望的輸出聲道數相同；這直接限制該原型矩陣的大小。例如：Q可以是一矩陣，該矩陣具有的列數是該降混訊號(212、324)的聲道數，以及行數是最終合成輸出訊號(332、340)的聲道數。

【0372】 作為一示例，在5.1或5.0訊號的情況下，該原型矩陣可以被建立如下：

$$Q = \begin{pmatrix} 1 & 0 & \sqrt{2} & 1 & 0 \\ 0 & 1 & \sqrt{2} & 0 & 1 \end{pmatrix}$$

【0373】 注意的是，該原型矩陣可以是預定的並且是固定的。例如：對於所有訊框，Q可以是相同的，但是對於不同的頻帶可以不同。此外，對於在該降混訊號的聲道數與該合成訊號的聲道數之間的不同關係，存在數個不同的Q。例如：在特定的降混聲道數及特定的合成聲道數的基礎上，Q可以從數個預存的Q中被選擇。

【0374】 方面5：在該輸出揚聲器設置與該輸入揚聲器設置不同的情況下重建數個參數：

【0375】 被提出的發明的一種應用是在一揚聲器設置上產生與該原始訊號212不同的一輸出訊號336或340(譬如意調著具有更多或更少數量的揚聲器)。

【0376】 為此，必須相應地修改該原型矩陣。在這種情況下，通過公式(9)獲得的原型訊號將包含如同該輸出揚聲器設置的許多聲道。例如：如果我們有5個聲道的訊號作為一輸入(在訊號212的一側)，並且想要獲得一7聲道的訊號作為一輸出(在訊號336的一側)，則該原型訊號將已經包含7聲道。

【0377】 這樣一來，在公式(4)中的該協方差矩陣的估計仍然成立，並且仍將被用於估計在該輸入訊號212中不存在的該數個聲道的該數個協方差參數。

【0378】 在該編碼器與該解碼器之間的該數個被傳送的參數228仍然是相關的，且公式(7)仍然可以被使用。更精確地，該數個被編碼(譬如被傳送)的參數必須被指派給在幾何學上盡可能接近該原始設置的該數個聲道對。基本上，需要進行一調適操作(adaptation operation)。

【0379】 例如：如果在該編碼器側估計在右側的一個揚聲器與左側的一個揚聲器之間的一ICC值，則可以將此值指派給具有相同左與右位置的輸出設置的該聲道對；在幾何形狀不同的情況下，此值可以被指派給位置與該原始揚聲器盡可能接近的該揚聲器對。

【0380】 然後，一旦獲得用於該新輸出設置的該目標協方差矩陣 C_y ，其餘的處理就保持不變。

【0381】 因此，為了使該目標協方差矩陣(C_{yR})適應於該合成聲道數，可行的是：

使用一原型矩陣 Q ，其從該降混聲道數轉換為該合成聲道數；這可以通過調適公式(9)，使該原型訊號具有該合成聲道數；

調適公式(4)，從而以合成聲道數估計 $\widehat{C}_y$ ；

保持公式(5)至(8)，其可因此獲得原始聲道數；

但將數個原始聲道群組(譬如數個原始聲道對)指派到單個合成聲道上(譬如根據幾何形狀選擇分配)，反之亦然。

【0382】 在第8b圖中提供一個示例，其是第8a圖的版本，其中指示一些矩陣及向量的聲道數。當在392處將該數個ICC(從該位元流348的該旁側資訊228被獲得)應用於該ICC矩陣時，將數個原始聲道群組(譬如數對原始聲道)移到單個合成聲道上(就幾何形狀來選擇分配)，反之亦然。

【0383】 對於不同於該輸入聲道數的數個輸出聲道產生一目標協方差矩陣的另一種可能性是，首先對於該輸入聲道數(譬如該輸入訊號212的原始聲道數)產生該目標協方差矩陣，然後使此第一目標協方差矩陣適應於該合成聲道數，獲得與該輸出聲道數對應的一第二目標協方差矩陣。這可以通過應用一升混規則或降混規則被完成，譬如將包含用於對該輸出聲道的某些輸入(原始)聲道的組合的數個因子的一矩陣應用於第一目標協方差矩陣 C_{y_R} ，然後在第二步驟中將此矩陣 C_{y_R} 應用於該數個被傳送的輸入聲道功率(數個ICLD)並取得用於該輸出(合成)聲道數的一聲道功率向量，並根據向量調整該第一目標協方差矩陣，以獲得具備所需合成聲道數的一第二目標協方差矩陣。該被調整的第二目標協方差矩陣現在可以被使用在該合成中。在第8c圖中提供其一示例，第8c圖是第8a圖的一版本，其中該數個塊390至394操作進行重建該目標協方差矩陣 C_{y_R} 以具有該原始訊號212的該原始聲道數。在那之後，在塊395處，一原型訊號 Q_N (以轉換

為該合成聲道數)及該向量ICLD可以被施加。值得注意的是，第8c圖的塊386與第8a圖的塊386相同，除了以下事實：在第8c圖中，該重建目標協方差的聲道數與該輸入訊號212的原始聲道數完全相同(且在第8a圖中，為了通常性，該重建目標協方差具有該合成聲道數)。

【0384】 4.3.4 去相關(Decorrelation)

【0385】 該去相關模組330的目的是減少在該原型訊號的每個聲道之間的相關性的數量。高度相關的揚聲器訊號可能會導致諸多幻覺源(phantom sources)，並降級該輸出多聲道訊號的品質及空間特性。此步驟是可選的，並且可以根據該應用程式需求而被執行或不執行。在本發明中，去相關在該合成引擎之前被使用。作為一示例，一全通頻率去相關器可以被使用。

【0386】 關於MPEG環繞(MPEG Surround)的注意事項：

【0387】 在根據先前技術的MPEG環繞中，使用所謂的“混合矩陣(Mix-matrices)”(在標準中被標示 M_1 及 M_2)。該矩陣 M_1 控制如何將該諸多可用的降混訊號輸入到該諸多去相關器。 M_2 矩陣描述直接的訊號及去相關的訊號應如何被組合以產生該輸出訊號。

【0388】 儘管可能與在4.3.3中被定義的該原型矩陣以及在本節中被描述的去相關器的用法相似，但重要的是要注意：

- 該原型矩陣Q的功能與在MPEG環繞中被使用的矩陣完全不同，此矩陣的要點是產生該原型訊號。該原型訊號的目的是要被輸入到該合成引擎中。
- 該原型矩陣無意為該諸多去相關器準備該諸多降混訊號，並且可以取決於該需求及目標應用進行調適。譬如該原型矩陣可以對於一輸出揚聲器設置大於該輸入揚聲器設置產生一原型訊號。

- 在所提出的發明中，該諸多去相關器的使用不是強制性的；該處理過程依賴在該合成引擎內的該協方差矩陣的使用(請參見5.1)。
- 所提出的發明沒有通過組合一直接訊號及一去相關訊號來產生該輸出訊號。
- M_1 及 M_2 的計算高度取決於樹狀結構，從該結構的觀點來看，這些矩陣的不同係數視情況而定(case-dependent)。在所提出的發明中不是這種情況，該處理與該降混計算無關(請參見5.2)，並且在概念上，所提出的處理旨在考慮在每個聲道之間的關係，而不是僅考慮諸多聲道對，因為可以使用一樹狀結構被完成。

【0389】 因此，本發明不同於根據先前技術的MPEG環繞。

【0390】 4.3.5 合成引擎(Synthesis Engine)、矩陣計算

【0391】 該解碼器的最後一步包括該合成引擎334或合成處理器402(如果需要，還包括一合成濾波器組338)。該合成引擎334的一目的是相對於某些約束產生最終的輸出訊號336。該合成引擎334可以計算一輸出訊號336，該輸出訊號336的特性受到該諸多輸入參數的約束。在本發明中，除了該原型訊號328(或332)之外，該合成引擎338的該輸入參數318是該數個協方差矩陣 C_x 及 C_y 。由於輸出訊號的特性應盡可能接近於由 C_y 定義的目標協方差矩陣，因此 C_{yR} 尤其被稱為目標協方差矩陣(它將被顯示該目標協方差矩陣的一估計版本及預建版本)。

【0392】 可以被使用的該合成引擎334不是唯一的，作為一示例，一先前技術的協方差合成可以被使用 [8]，其通過引用併入本文。可以被使用的另一種合成引擎333將是在 [2] 的DirAC處理中被描述的該合成引擎。

【0393】該合成引擎334的該輸出訊號可能需要通過該合成濾波器組338進行其他處理。

【0394】作為一最終結果，該輸出多聲道訊號340在時域中被獲得。

【0395】方面6：使用該“協方差合成”的高品質輸出訊號

【0396】如上所述，所使用的合成引擎334不是唯一的，並且使用該數個被傳送的參數或它的一子集的任何引擎可以被使用。然而，本發明的一方面可以提供諸多高品質的輸出訊號336，譬如通過使用該協方差合成[8]。

【0397】該合成方法旨在計算一輸出訊號336，該輸出訊號336的諸多特性由該協方差矩陣 C_{y_R} 定義。為此，計算諸多所謂的最佳混合矩陣(optimal mixing matrices)，這些矩陣會將該原型訊號328混合到該最終輸出訊號336中，從一數學觀點來看，在給定一目標協方差矩陣 C_{y_R} 的情況下提供最佳結果。

【0398】該混合矩陣 M 是將經由該關係 $y_R = Mx_P$ 將該原型訊號 x_P 轉換為該輸出訊號 y_R (336)的矩陣。

【0399】該混合矩陣也可以是將經由該關係 $y_R = Mx$ 將該降混訊號 x 轉換為該輸出訊號的一矩陣。從此關係，我們還可以推論 $C_{y_R} = MC_x M^*$ 。

【0400】在被呈現的處理 C_{y_R} 及 C_x 中，並且在某些示例中可能是已知的(因為它們分別是該降混訊號246的該目標協方差矩陣 C_{y_R} 及該協方差矩陣 C_x)。

【0401】從一數學觀點來看，一種解決方案是通過 $M = K_y P K_x^{-1}$ 給定的，其中 K_y 及 K_x^{-1} 是通過對 C_x 及 C_{y_R} 進行奇異值分解(singular value decomposition)所獲得的所有矩陣。對於 P 而言，它在此是開放參數，但是相對於由該原型矩陣 Q 所支配的約束，可以找到一最佳解決方案(從傾聽者的一感知角度來看)。在此說明的數學證明可在 [8] 中被找到。

【0402】該合成引擎334提供高品質的輸出336，因為該方法被設計為提供對輸出訊號問題的重建的最佳數學解決方案。

【0403】用較少的數學術語，對瞭解協方差矩陣表示在一多聲道音訊訊號的不同聲道之間的諸多能量關係非常重要。用於該原始多聲道訊號212的該矩陣 C_y 及用於該降混多聲道訊號246的矩陣 C_x 。這些矩陣的每個值都反映該多聲道流的兩個聲道之間的能量關係。

【0404】因此，該協方差合成背後的哲理是產生一訊號，該訊號的特性由該目標協方差矩陣 C_{y_R} 驅動。此矩陣 C_{y_R} 被計算的方式是描述該原始輸入訊號212(或在不同於該輸入訊號的情況下，我們想要獲得該輸出訊號)。然後，具有這些元素，該協方差合成將最佳地混合該原型訊號，以便產生該最終的輸出訊號。

【0405】在另一方面，用於一時隙的合成的該混合矩陣是該當前的訊框的該混合矩陣 M 與該先前的訊框的該混合矩陣 M_p 的一組合，以確保一平滑的合成，例如基於當前的訊框內的該時隙索引的一線性內插(linear interpolation)。

【0406】在另一方面，其中一暫態的發生及位置被傳送，在該暫態位置之前，將先前的混合矩陣 M_p 用於所有時隙，並且將該混合矩陣 M 用於包含該暫態位置的時隙及在該當前的訊框中的所有後續時隙。注意的是，在某些示例中，對於每個訊框或時隙，可以使用具備用於一先前的訊框或時隙的一混合矩陣的一線性組合以平滑該當前的訊框或時隙的該混合矩陣，例如通過加法、平均等。讓我們假設，對於一當前的訊框 t ，該輸出訊號的該數個時隙 b 及 i 通過

$Y_{s,i} = M_{s,i}X_{s,i}$ 被獲得，其中 $M_{s,i}$ 是用於該先前的訊框的該混合矩陣 $M_{t-1,i}$ 的一組

合，並且 $M_{t,i}$ 是用於該當前的訊框所計算的混合矩陣，例如在它們之間的線性插值：

$$M_{s,i} = \left(1 - \frac{s}{n_s}\right) M_{t-1,i} + \frac{s}{n_s} M_{t,i}$$

其中， n_s 是在一訊框中的該時隙數(例如 16)，且 $t-1$ 及 t 指示先前的訊框及當前的訊框。更通常地，與每個時隙相關的混合矩陣 $M_{s,i}$ 可以被獲得，通過沿著一當前的訊框 t 的數個後續時隙以一增加係數縮放為該當前的訊框所計算的該混合矩陣 $M_{t,i}$ ，及通過沿著該當前的訊框 t 的數個後續時隙加上以一減少係數被縮放的該混合矩陣 $M_{t-1,i}$ 。該數個係數可以是線性的。

【0407】 可被提供的是，在一暫態(譬如在資訊261中被發訊表明)的情況下該當前混合矩陣及過去混合矩陣不被組合，而是先前的直到包含該暫態的時槽以及當前的用於包含該暫態的時槽及所有後續的時槽，直到該訊框結束為止。

$$Y_{s,i} = \begin{cases} M_{t-1,i} X_{s,i}, & s < s_t \\ M_{t,i} X_{s,i}, & s \geq s_t \end{cases}$$

其中 s 是該時隙索引， i 是該頻帶索引， t 及 $t-1$ 指示當前的訊框及先前的訊框，並且 s_t 是包含暫態的時隙。

【0408】 與先前技術文件[8]的差異

【0409】 同樣重要的是要注意，所提出的發明超出在 [8] 中被提出的方法的範圍。顯著的差異尤其是：

- 該目標協方差矩陣 C_{y_R} 是在所提出的處理的該編碼器側被計算。
- 該目標協方差矩陣 C_{y_R} 也可以用不同的方式被計算(在所提出的發明中，該協方差矩陣不是一擴散直接的部份的和)。

- 該處理不是針對每個頻帶單獨進行，而是針對數個參數頻帶進行分組(如在 4.2.1中所述)。
- 從一更全域的看法：該協方差合成在此只是整個過程的一個塊，並且必須與在解碼器側的所有其他元件一起使用。

【0410】 4.4 偏好方面作為一列表

【0411】 以下諸多方面中的至少一個可以表徵(characterize)本發明：

1. 在該編碼器側
 - a. 輸入一多聲道音訊訊號246。
 - b. 使用一濾波器組214將該訊號212從時域轉換到頻域(216)
 - c. 在塊244處計算該降混訊號246
 - d. 從該原始訊號212及/或該降混訊號246，估計一第一參數集以描述該多聲道流(訊號)246：數個協方差矩陣 C_x 及/或 C_y
 - e. 傳送及/或編碼該數個協方差矩陣 C_x 及/或 C_y 直接或計算該數個ICC及/或數個ICLD並傳送它們
 - f. 使用一適當的編碼方案在該位元流248中編碼該數個被傳送的參數228
 - g. 在時域中計算該降混訊號246
 - h. 在時域中傳送該旁側資訊(即該數個參數)及該降混訊號246
2. 在該解碼器側
 - a. 對包含該旁側資訊228及該降混訊號246的該位元流248進行解碼
 - b. (可選的)將該濾波器組320應用於該降混訊號246，以便獲得在頻域中的該降混訊號246的一版本324

- c. 從數個先前被解碼的參數228及降混訊號246重建該協方差矩陣 C_x 及 C_{y_R}
- d. 從該降混訊號246計算該原型訊號328(324)
- e. (可選的)將該原型訊號進行去相關(在塊330處)
- f. 使用 C_x 及 C_{y_R} 將該合成引擎334應用於該原型訊號作為被重建的
- g. (可選的)將該合成濾波器組338應用於該協方差合成334的該輸出336
- h. 獲得該輸出多聲道訊號340

【0412】 4.5 協方差合成(Covariance synthesis)

【0413】 在本節中，討論可以在第1至3d圖的系統中被實現的一些技術。

然而，這些技術也可以被獨立實現：例如：在某些示例中，不需要如針對第8a至8c圖及公式(1)至(8)中所實行的該協方差計算。因此，在某些示例中，當提及 C_{y_R} (重建目標協方差)時，也可以由 C_y 替代(其也可以被直接提供，而無需重建)。儘管如此，此節的技術可以有利地與上述技術一起使用。

【0414】 現在參考第4a至4d圖。在此，討論協方差合成塊388a至388d的諸多示例。數個塊388至388d可以實施為例如第3c圖的塊388，以進行協方差合成。數個塊388a至388d可以例如是第3a圖的該合成引擎334的該合成處理器404及該混合規則計算器402及/或參數重建塊316的該合成處理器404及該混合規則計算器402中的一部分。在第4a至4d圖中，該降混訊號324在頻域FD中(即，在該濾波器組320的下游)，並且用X指示，而該合成訊號336也在FD中，並且用Y指示，然而，在時域中概括這些結果是可行的。注意的是，第4a至4d圖的該數個協方差合成塊388a至388d中的每一個可以被稱為單一個頻帶(譬如一旦在380中被分

解)，並且該數個協方差矩陣 C_x 及 C_{y_R} (或其他被重建的資訊)因此可以與一個特定的頻帶相關聯。例如：該協方差合成可以以一逐訊框的方式被進行，並且在那種情況下，數個協方差矩陣 C_x 及 C_{y_R} (或其他被重建的資訊)是與單一個訊框(或數個連續的訊框)相關聯：因此，該協方差合成可以以一逐訊框的方式或以一逐多訊框(multiple-frame-by-multiple-frame)的方式進行。

【0415】 在第4a圖中，該協方差合成塊388a可以由一個能量補償的最佳混合塊600a及缺少相關器塊構成。基本上，單一個混合矩陣 M 被找到，並且被附加執行的唯一重要操作是一能量補償混合矩陣 M' 的計算。

【0416】 第4b圖顯示受[8]啟發的一協方差合成塊388b。該協方差合成塊388b可以允許獲得該合成訊號336作為具有一第一主要分量336M及一第二殘餘分量336R的一合成訊號。儘管該主要分量336M可以在一最佳的主要分量混合矩陣600b處被獲得，譬如通過從該數個協方差矩陣 C_x 及 C_{y_R} 中找出一混合矩陣 M_M ，且不使用諸多去相關器，但是該殘餘分量336R可以用另一種方式獲得。 M_R 原則上應滿足該關係 $C_{y_R} = MC_x M^*$ 。通常，所獲得的混合矩陣不能完全滿足該要求，並且可以用 $C_r = C_{y_R} - MC_x M^*$ 找到一殘餘目標協方差。可以看出，該降混訊號324可以被導出到一路徑610b上(該路徑610b可以被稱為第二路徑，該第二路徑與一第一路徑610b'平行，該第一路徑610b'包括塊600b)。該降混訊號324的一原型版本613b(用 Y_{pR} 表示)可以在原型訊號塊(升混塊)612b處被獲得。例如：可以使用諸如公式(9)的公式，即

$$Y_{pR} = XQ$$

【0417】 在本文件中提供Q(原型矩陣或升混矩陣)的諸多示例。在塊612b的下游，呈現一去相關器614b，以便對該原型訊號613b進行去相關，以獲得一去相關訊號615b(也用 $\hat{Y}$ 指示)。在塊616b處，從去相關訊號615b，估計該去相關訊號 $\hat{Y}$ (615b)的該協方差矩陣 $C_{\hat{Y}}$ 。通過使用該去相關訊號 $\hat{Y}$ 的該協方差矩陣 $C_{\hat{Y}}$ 作為主要分量混合的 C_x 的等效值及 C_r 作為另一個最佳混合塊中的該目標協方差的，可以在一最佳殘餘分量混合矩陣塊(optimal residual component mixing matrix block)618b處獲得該合成訊號336的該殘餘分量336R。該最佳殘餘分量混合矩陣塊618b可以用這樣的方式被實現：產生一混合矩陣 M_R ，以便混合該去相關訊號615b，並獲得該合成訊號336的該殘餘分量336R(針對一特定頻帶)。在加法器塊620b處，該殘餘分量336R被加到該主要分量336M上(因此該數個路徑610b及610b'在加法器塊620b處被聯結在一起)。

【0418】 第4c圖顯示替代第4b圖的協方差合成388b的協方差合成388c的一示例。該協方差合成塊388c允許獲得該合成訊號336作為具有一第一主要分量336M'及一第二殘餘分量336R'的一訊號Y。儘管該主要分量336M'可以在一最佳主要分量混合矩陣600c處被獲得，譬如通過從該數個協方差矩陣 C_x 及 C_{y_R} (或 C_y 其他資訊220)中找出一混合矩陣 M_M ，且不使用諸多相關器，但是可以用另一種方式得到該殘餘分量336R'。該降混訊號324可以被導出到一路徑610c上(該路徑610c可以被稱為第二路徑，該第二路徑與一第一路徑610c'平行，該第一路徑610c'包括塊600c)。通過應用該原型矩陣Q(譬如以一聲道數即該合成聲道數將該降混訊號234升混到該降混訊號234的一版本613c上的一矩陣)，該降混訊號324的一原型版本613c可在降混塊(升混塊)612c處被獲得。例如：可以使用諸如公式(9)的一公式。本文件提供Q的諸多示例。在塊612c的下游，可以提供一去相關器

614c。在某些示例中，該第一路徑沒有去相關器，而該第二路徑具有一去相關器。

【0419】 該去相關器614c可以提供一去相關訊號615c(也用 $\hat{Y}$ 指示)。然而，與在第4b圖的該協方差合成塊388b中被使用的技術相反，在第4c圖的該協方差合成塊388c中，不從去相關訊號615c($\hat{Y}$)估計去相關訊號615c的協方差矩陣 $C_{\hat{Y}}$ 。相反，該去相關訊號615c的協方差矩陣 $C_{\hat{Y}}$ 是從以下位置所獲得的(在塊616c處)：

該降混訊號324的該協方差矩陣 C_x (譬如如在第3c圖的塊384處及/或使用公式(1)被估計的)；及

該原型矩陣 Q 。

【0420】 通過使用從該降混訊號324的該協方差矩陣 C_x 估計出的該協方差矩陣 $C_{\hat{Y}}$ 作為主要分量混合矩陣的 C_x 及 C_r 作為目標協方差矩陣的的等效物，在最佳殘餘分量混合矩陣塊618c處獲得該合成訊號336的該殘餘分量336R'。該最佳殘餘分量混合矩陣塊618c可以用產生一殘餘分量混合矩陣 M_R 的方式被實現，以便通過根據殘餘分量混合矩陣 M_R 混合該去相關訊號615c以獲得該殘餘分量336R'。在加法器塊620c處，該殘餘分量336R'被加到該主要分量336M'，以便獲得該合成訊號336(該數個路徑610c及610c'因此在加法器塊620c處被聯接在一起)。

【0421】 在某些示例中，該殘餘分量336R或336R'不總是或不需被計算(並且該路徑610b或610c不總是被使用)。在某些示例中，雖然對於某些頻帶執行該協方差合成而不計算該殘餘訊號336R或336R'，但是對於相同訊框的其他頻帶，還考慮該殘餘訊號336R或336R'以處理該協方差合成。第4d圖顯示該協方差合成塊388d的一示例，其可以是該協方差合成塊388b或388c的一特定情況：在此，

一頻帶選擇器630可以選擇或取消選擇(以開關631表示的方式)該殘餘訊號336R或336R'的計算。例如：該路徑610b或610c可以由選擇器630針對某些頻帶選擇性地啟用，而對於其他頻帶停用。特別地，該路徑610b或610c可以針對超過一預定閾值(譬如一固定閾值)的數個頻帶而被停用，該預定閾值(譬如最大值)可以是區分人耳對相位不敏感的數個頻帶(頻率高於閾值的數個頻帶)及人耳對相位敏感的數個頻帶(頻率低於閾值的數個頻帶)，因此不會為頻率低於閾值的該數個頻帶計算該殘餘分量336R或336R'，並針對頻率高於閾值的數個頻帶計算該殘餘分量336R或336R'。

【0422】 第4d圖的示例還可以通過用第4a圖的塊600a替換塊600b或600c，並且用第4b圖的協方差合成塊388b或第4c圖的協方差合成塊388c替換該塊610b或610c來獲得。

【0423】 在此提供關於如何在塊338、402(或404)、600a、600b、600c等處獲得該混合規則(矩陣)的一些指示。如上所述，有許多獲得混合矩陣的方法，但是這裡將更詳細地討論其中一些。

【0424】 特別地，首先，參考第4b圖的該協方差合成塊388b。在最佳主要分量混合矩陣塊600c處，例如：可以從以下公式獲得該合成訊號336的該主要分量336M的該混合矩陣M：

該原始訊號212的該協方差矩陣 C_y (C_y 可以使用上面討論的公式(6)至(8)中的至少一些被估計，例如參見第8圖；它可以是所謂的“目標版本(target version)”形式 C_{y_R} ，譬如根據公式(8)估算的值)；及
該降混訊號246、324的協方差矩陣 C_x (C_y 可以使用例如使用公式(1)被估計)。

【0425】 例如：如[8]所提議的，根據以下的因式分解，它被承認以分解數個協方差矩陣 C_x 及 C_y ，它們是厄米特(Hermitian)矩陣及正半定矩陣：

$$\begin{aligned} C_x &= K_x K_x^* \\ C_y &= K_y K_y^* \end{aligned}$$

K_x 及 K_y 可以例如通過從 C_x 及 C_y 應用兩次奇異值分解(SVD)而被獲得。例如：

C_x 的SVD可以提供數個奇異向量(譬如數個左奇異向量)的一矩陣 U_{Cx} ；及數個奇異值的一對角矩陣 S_{Cx} ；

因此， K_x 可以通過將 U_{Cx} 乘以一對角矩陣而被獲得，該對角矩陣在它的數個元中具有 S_{Cx} 的該數個相應的元中的數個值的數個平方根。

【0426】 此外，關於 C_y 的SVD可以提供：

【0427】 數個奇異向量(譬如數個右奇異向量)的一矩陣 V_{Cy} ；及

【0428】 數個奇異值的一對角矩陣 S_{Cy}

【0429】 因此， K_y 可以通過將 U_{Cy} 乘以一對角矩陣被獲得，該對角矩陣在它的數個元中具有 S_{Cy} 的數個對應的元中的數個值的數個平方根。

【0430】 然後，獲得一主要分量混合矩陣 M_M 是可行的，當將其應用於該降混訊號324時，將允許獲得該合成訊號336的該主要分量336M。該主要分量混合矩陣 M_M 可以被獲得如下：

$$M_M = K_y P K_x^{-1}$$

【0431】 如果 K_x 是一不可逆矩陣，則可以用已知技術獲得一正則化逆矩陣(regularized inverse matrix)，並用 K_x^{-1} 代替。

【0432】 該參數P通常是開放的(free)，但是它可以被最佳化。為了得出P，可以將SVD應用於：

C_x (該降混訊號324的協方差矩陣)；及

$C_{\hat{y}}$ (該原型訊號613b的協方差矩陣)。

【0433】 一旦執行該數個SVD，就有可能獲得P，如

$$P = V\Lambda U^*$$

【0434】 Λ 是一個矩陣，其具有的列數(rows)與該合成聲道數相同，而行數(columns)與該降混聲道數相同。 Λ 是在它的第一個正方形塊中的一標識，並在該數個其餘的元中以零完成。現在說明V及U 如何從 C_x 及 $C_{\hat{y}}$ 被獲得，V及U 是從一SVD獲得的數個奇異向量的數個矩陣：

$$USV^* = K_x^* Q^* G_{\hat{y}}^* K_y$$

【0435】 S是通常通過SVD獲得的數個奇異值的該對角矩陣。 $G_{\hat{y}}$ 是一對角矩陣，其將該原型訊號 $\hat{y}$ (615b)的每聲道能量正規化為該合成訊號y的能量。為了獲得 $G_{\hat{y}}$ ，首先需要計算 $C_{\hat{y}} = Q C_x Q^*$ ，即該原型訊號 $\hat{y}$ 的協方差矩陣(614b)。然後，為了從 $C_{\hat{y}}$ 得出 $G_{\hat{y}}$ ，將 $C_{\hat{y}}$ 的數個對角線值正規化為 C_y 的數個對應的對角的值，從而提供 $G_{\hat{y}}$ 。一個示例是 $G_{\hat{y}}$ 的數個對角元被計算為 $g_{\hat{y}ii} = \sqrt{\frac{c_{yii}}{c_{\hat{y}ii}}}$ ，其中 c_{yii} 是 C_y 的該數個對角元的數個值及 $c_{\hat{y}ii}$ 是 $C_{\hat{y}}$ 的該數個對角元的數個值。

【0436】 一旦獲得 $M_M = K_y P K_x^{-1}$ ，該殘餘分量的該協方差矩陣 C_r 可從

$$C_r = C_y - M_M C_x M_M^*$$

【0437】 一旦獲得 C_r ，就有可能獲得用於混合該去相關訊號615b以獲得該殘餘訊號336R的一混合矩陣，其中在一相同最佳混合 C_r 具有與該主要最佳混合的 C_{y_R} 相同的作用的情況，該數個去相關原型 $C_{\hat{y}}$ 的該協方差的作用為該輸入訊號協方差 C_x 具有該主要最佳混合。

【0438】然而，已被理解的是，與第4b圖的技術相比，第4c圖的技術具有一些優點。在某些示例中，第4c圖的技術與第4c圖的技術相同，至少用於計算該主要矩陣並用於產生該合成訊號的該主要分量。相反，第4c圖的技術與第4b圖的技術的區別在於該殘餘混合矩陣的計算，並且更一般而言，用於產生該合成訊號的該殘餘分量。現在參考第11圖結合第4c圖用於計算該殘餘混合矩陣。在第4c圖的示例中，在頻域中的一去相關器614c被使用，其確保該原型訊號613c的去相關，但是保留該原型訊號613b本身的能量。

【0439】此外，在第4c圖的示例中，我們可以假設(至少通過近似)該去相關訊號615c的該數個去相關聲道是互不同調的，因此該數個去相關訊號的該協方差矩陣的所有非對角元都是零。通過這兩個假設，我們可以簡單通過在 C_x 上應用Q以估計該去相關原型的該協方差，而僅採用該協方差的該主對角線(即該原型訊號的能量)。從該去相關訊號615b著手，第4c圖的技術要比第4b圖的示例進行估計的效率更高，其中我們需要進行與已經對 C_x 進行的相同的頻帶/時隙聚合。因此，在第4c圖的示例中，我們可以簡單地應用已經聚合的 C_x 的一矩陣乘法。因此，對於相同的聚合頻帶群組的所有頻帶計算相同的混合矩陣。

【0440】因此，可以在710處使用以下內容估計該去相關訊號的該協方差711(C_y)：

$$P_{\text{decorr}} = \text{diag}(QC_xQ^*)$$

作為具備所有非對角元被設置為零的一矩陣的主對角線，其被用於作為輸入訊號協方差 C_y 。在諸多示例中 C_x 被平滑以用於進行該合成訊號的該主要分量336M'的合成，該技術可以被使用根據 C_x 被用於計算 P_{decorr} 為非平滑的 C_x 。

【0441】現在，一原型矩陣 Q_R 應該被使用。然而，已經被注意到的是，對於該殘餘訊號， Q_R 是單位矩陣(identity matrix)。 $C_{\hat{y}}$ (對角矩陣)及 Q_R (恆等矩陣)的屬性知識可進一步簡化該混合矩陣的計算(至少可以省略一個SVD)，請參見以下技術及Matlab清單(Listing)。

【0442】首先，類似於第4b圖的示例，該輸入訊號212的該殘餘目標協方差矩陣 C_r (Hermitian、正半定的)可以被分解為 $C_r = K_r K_r^*$ 。可以通過SVD(702)獲得矩陣 K_r ：該SVD 702用於 C_r 產生：

數個奇異向量(譬如數個左奇異向量)的一矩陣 U_{Cr} ；

數個奇異值的一對角矩陣 S_{Cr} ；

因此 K_r 通過在對角矩陣中將 U_{Cr} 乘以一對角矩陣被獲得(在706中)，該對角矩陣在它的數個元中具有在 S_{Cr} 的數個對應的元中的數個值的數個平方根(後者已在704處被獲得)。

【0443】在此點上，從理論上講，這次可以將另一個SVD應用於該去相關原型的該協方差 $\hat{y}$ 。

【0444】然而，在此示例中(第4c圖)，為了減少計算量，已選擇不同的路徑。從 $P_{\text{decorr}} = \text{diag}(Q C_x Q^*)$ 估計的 $C_{\hat{y}}$ 是一對角矩陣，因此不需要SVD(一對角矩陣的SVD給出數個奇異值作為對角元素的一排序向量，而左與右奇異向量僅指示該排序的索引)。通過計算(在712處)在 $C_{\hat{y}}$ 的對角線的該數個元處的每個值的平方根，獲得一對角矩陣 $\hat{R}_y$ 。該對角矩陣 $\hat{R}_y$ 是使得 $\hat{R}_y \hat{R}_y^* = C_{\hat{y}}$ ，具備優點是為了獲得 $\hat{R}_y$ 不需要SVD。從該數個去相關訊號 $C_{\hat{y}}$ 的該對角協方差，計算該去相關訊號615c的一估計協方差矩陣 $\widehat{C}_y$ 。但是由於該原型矩陣是 Q_R (即同質性矩陣)，因

此可以直接使用 $\mathbf{C}_{\hat{\mathbf{y}}}$ 於公式化 $\widehat{\mathbf{G}}_{\mathbf{y}}$ 作為 $\widehat{\mathbf{g}}_{\mathbf{y}_{11}} = \sqrt{\frac{c_{r_{11}}}{\widehat{c}_{\mathbf{y}_{11}}}}$ ，其中 $c_{r_{11}}$ 是 $\mathbf{C}_r$ 的數個對角元的數個值及 $\widehat{c}_{\mathbf{y}_{11}}$ 是 $\mathbf{C}_{\hat{\mathbf{y}}}$ 的數個對角元的數個值。 $\mathbf{G}_{\hat{\mathbf{y}}}$ 是一對角矩陣(在722處獲得)，其將該去相關訊號 $\hat{\mathbf{y}}$ (615b)的每聲道能量正規化為該合成訊號 $\mathbf{y}$ 的期望能量。

【0445】此時，有可能(在734處)將 $\widehat{\mathbf{K}}_{\mathbf{y}}$ 乘以 $\widehat{\mathbf{G}}_{\mathbf{y}}$ (也稱為乘法734的結果735)。 $\widehat{\mathbf{K}}_{\mathbf{y}}$ 然後(736)，將 $\mathbf{K}_r$ 乘以 $\widehat{\mathbf{K}}_{\mathbf{y}}$ 得到 $\mathbf{K}'_{\mathbf{y}}$ (即 $\mathbf{K}'_{\mathbf{y}} = \mathbf{K}_r \widehat{\mathbf{K}}_{\mathbf{y}}$)。從 $\mathbf{K}'_{\mathbf{y}}$ ，可以執行一SVD(738)，以便獲得一左奇異向量矩陣 $\mathbf{U}$ 及一右奇異向量矩陣 $\mathbf{V}$ 。通過將 $\mathbf{V}$ 及 $\mathbf{U}^*$ 相乘(740)，獲得一矩陣 $\mathbf{P}$ 。 $\mathbf{P} = \mathbf{V}\mathbf{U}^H$ 最後(742)，可以通過應用以下內容獲得該殘餘訊號的該混合矩陣 $\mathbf{M}_R$ ：

$$\mathbf{M}_R = \mathbf{K}_r \mathbf{P} \widehat{\mathbf{K}}_{\mathbf{y}}^{-1}$$

其中 $\widehat{\mathbf{K}}_{\mathbf{y}}^{-1}$ (在745處被獲得)可以由該正則化逆的進行。 $\mathbf{M}_R$ 因此可以在塊618c處被使用於該殘餘混合。

【0446】這裡提供用於執行如上所述的協方差合成的一Matlab代碼 (code)。注意的是，代碼中的星號(*)表示乘法，而頂點(^)表示厄米特矩陣。

```
%Compute residual mixing matrix
function [M] = ComputeMixingMatrixResidual(C_hat_y,Cr,reg_sx,reg_ghat)

EPS_ = single(1e-15); %Epsilon to avoid divisions by zero
num_outputs = size(Cr,1);
%Decomposition of Cy
[U_Cr, S_Cr] = svd(Cr);
Kr = U_Cr*sqrt(S_Cr);

%SVD of a diagonal matrix is the diagonal elements ordered,
%we can skip the ordering and get Kx directly form Cx
K_hat_y=sqrt(diag(C_hat_y));

limit=max(K_hat_y)*reg_sx+EPS_;
S_hat_y_reg_diag=max(K_hat_y,limit);
```

```

%Formulate regularized Kx
K_hat_y_reg_inverse=1./S_hat_y_reg_diag;

% Formulate normalization matrix G hat
% Q is the identity matrix in case of the residual/diffuse part so
%  $Q * C_x * Q' = C_x$ 
Cy_hat_diag = diag(C_hat_y);
limit = max(Cy_hat_diag)*reg_ghat+EPS_;
Cy_hat_diag = max(Cy_hat_diag,limit);
G_hat = sqrt(diag(Cr)./Cy_hat_diag);

%Formulate optimal P
%Kx, G_hat are diagonal matrixes, Q is I...
K_hat_y=K_hat_y.*G_hat;
for k =1:num_outputs
    Ky_dash(k,:)=Kr(k,:)*K_hat_y(k);
end
[U,~,V] = svd(Ky_dash);
P=V*U';

%Formulate M
M=Kr*P;
for k = 1:num_outputs
    M(:,k)=M(:,k)*K_hat_y_reg_inverse(k);
end

end

```

【0447】 在此提供關於第4b及4c圖的協方差合成的討論。在某些示例中，對於每個頻帶可以考慮兩種合成方式，對於某些頻帶通常使用高於人耳對相位不敏感的一特定頻率的頻帶包括來自第4b圖的該剩餘路徑的完全合成，以達到將一能量補償應用在該聲道中的所需能量。

【0448】 因此，同樣在第4b圖的示例中，對於低於某個(固定的、解碼器已知的)頻帶邊界(閾值)的數個頻帶，可以執行根據第4b圖的完全合成(譬如在第4d圖的情況下)。在第4b圖的示例中，該去相關訊號615b的該協方差 $C_{\hat{y}}$ 是從該去相關訊號615b本身被導出的。相反，在第4c圖的示例中，在頻域中的一去相關器

614c被使用，其確保該原型訊號613c的去相關，但是保留該原型訊號613b本身的能量。

【0449】 進一步的考量：

- 在第4b及4c圖兩者的示例中：在該第一路徑(610b'、610c')處，通過依賴該原始訊號212的該協方差 C_y 及該降混訊號324的該協方差 C_x 以產生一混合矩陣 M_M (在塊600b、600c處)；
- 在第4b及4c圖兩者的示例中：在該第二路徑(610b、610c)處，有一去相關器(614b、614c)，並且產生一混合矩陣 M_R (在塊618b、618c處)，這應當考慮該去相關訊號(616b、616c)的該協方差 $C_{\hat{y}}$ ；但是
 - 在第4b圖的示例中，使用該去相關訊號(616b、616c)作為直觀地計算該去相關訊號(616b、616c)的該協方差 $C_{\hat{y}}$ ，並且在該原始聲道y的能量中被加權。
 - 在第4c圖的示例中，通過從該矩陣 C_x 估計並以直觀的方式反算該去相關訊號(616b、616c)的該協方差，並且在原始聲道y的能量中被加權。

【0450】 注意的是，該協方差矩陣(C_{y_R})可以是上面討論的該重建目標矩陣(譬如從被寫在該位元流248的該旁側資訊228中的該聲道位準及相關資訊220所獲得)，並且因此可以被認為與該原始訊號212的該協方差相關聯。無論如何，因為它將被用於該合成訊號336，所以該協方差矩陣(C_{y_R})也可以被認為是與該合成訊號相關聯的協方差。同樣應用於該剩餘協方差矩陣 C_r ，其也可以被理解為與該合成訊號相關聯的殘餘協方差矩陣(C_r)，而該主要協方差矩陣也可以被理解為與該合成訊號相關聯的主要協方差矩陣。

【0451】 5. 優勢(Advantages)**【0452】 5.1 減少對去相關的使用及該合成引擎的最佳化使用**

【0453】 給定所提出的技術，以及被用於處理的數個參數以及這些參數與該合成引擎334組合的方式，說明對該音訊訊號的強烈去相關的需求(譬如在它的版本328中)被降低，甚至在缺乏該去相關模組330的情況下，如果未被去除，也可以減小該去相關的影響(譬如空間特性的偽影或劣化或訊號品質的劣化)。

【0454】 更精確地，如前所述，該處理的該去相關部分330是可選的。實際上，該合成引擎334通過使用該目標協方差矩陣 C_y (或它的一子集)以對該訊號328進行去相關，並確保構成該輸出訊號336的數個聲道在它們之間被適當地去相關。 C_y 在該協方差矩陣中的數個值表示我們的多聲道音訊訊號的不同聲道之間的能量關係，這就是為什麼它用作合成的一目標的原因。

【0455】 此外，與該合成引擎334組合的該數個被編碼(譬如被傳送的)參數228(譬如在它們的版本314或318中)可以確保一高品質輸出336，其給定的事實為該合成引擎334使用該目標協方差矩陣 C_y ，以便重現一輸出多聲道訊號336，該輸出多聲道訊號336的空間特性及聲音品質與輸入訊號212盡可能接近。

【0456】 5.2 降混不可知處理(Down-mix agnostically processing)

【0457】 給定所提出的技術，以及該原型訊號328被計算的方式及它們如何與該合成引擎334一起使用，在此說明的是，所提出的解碼器與在該編碼器處被計算的該降混訊號212的方式無關。

【0458】 這意謂著，所提出的發明在該解碼器300處可以獨立於在該編碼器處計算該降混訊號246的方式被執行，並且該訊號336(或340)的該輸出品質不依賴於一特定的降混方法。

【0459】 5.3 數個參數的可縮放性(Scalability)

【0460】 給定所提出的技術，以及該數個參數(28、314、318)被計算的方式及它們與該合成引擎334一起使用的方式，以及它們在該解碼器側被估算的方式，這說明的是被用於描述該數個多聲道音訊訊號的該數個參數在數量及用途上都是可縮放的。

【0461】 通常，僅在該編碼器側被估計的該數個參數的一子集(譬如 C_y 及/或 C_x 的一子集，譬如其諸多元素)被編碼(譬如被傳送)：這允許減少由該處理所使用的諸多位元率。因此，給定該數個未被傳送的參數在該解碼器側被重建的事實，該數個被編碼(譬如被傳送)的參數(譬如 C_y 及/或 C_x 的元素)的數量可以是可縮放的。這給出機會就輸出品質及位元率以縮放整個處理過程，被傳送的參數越多，輸出品質越好，反之亦然。

【0462】 而且，那些參數(譬如 C_y 及/或 C_x 或其元素)在目的上是可縮放的，這意謂著它們可以由用戶輸入而被控制，以便修改該輸出多聲道訊號的特性。此外，可以針對每個頻帶計算那些參數，並且因此允許一可縮放的頻率解析度。

【0463】 譬如可以決定要以該輸出訊號(336、340)取消一個揚聲器，因此可以直接在該解碼器側操縱該數個參數，以實現這樣的一轉換(transformation)。

【0464】 5.4 輸出設置的靈活性(Flexibility of the output setup)

【0465】 給定所提出的技術，以及所使用的合成引擎334及該數個參數(譬如 C_y 及/或 C_x 或其元素)的靈活性，在此說明的是，所提出的發明允許涉及該輸出設置的一廣泛的渲染可能性(large spectrum of rendering possibilities)。

【0466】 更準確地說，該輸出設置不必與該輸入設置相同。操縱被饋入該合成引擎的該重建目標協方差矩陣是可行的，以便在一揚聲器設置上產生一輸

出訊號340，該揚聲器設置大於或小於或僅具備一幾何形狀不同於原始的揚聲器設置。這是可能的，因為要被傳送的數個參數以及所提出的系統與該降混訊號無關(請參見5.2)。

【0467】 由於這些原因，從該數個輸出揚聲器設置的觀點解釋所提出的發明是靈活的。

【0468】 5.5 數個原型矩陣的某些示例

【0469】 在此，下面的表已經針對5.1，但是LFE被排除在外，此後我們也將LFE包括在該處理中(只有用於關係LFE/C的一個ICC及用於LFE的ICLD僅在最低參數頻帶中被發送並且對於在該解碼器側處的該合成中所有其他頻帶分別設定為1及0)。聲道命名及諸多順序遵循ISO/IEC 23091-3“資訊技術-編碼獨立代碼點-第3部分：音訊”中的數個CICP，Q始終被用於作為在該解碼器中的原型矩陣及在該編碼器中的降混矩陣。5.1 (CICP6)。 α_i 要被用於計算該數個ICLD。

$$Q = \begin{pmatrix} 1 & 0 & \sqrt{2} & \sqrt{2} & 1 & 0 \\ 0 & 1 & \sqrt{2} & \sqrt{2} & 0 & 1 \end{pmatrix}$$

$$\alpha_i = [0.4444 \quad 0.4444 \quad 0.2 \quad 0.2 \quad 0.4444 \quad 0.4444]$$

7.1 (CICP12)

$$Q = \begin{pmatrix} 1 & 0 & \sqrt{2} & \sqrt{2} & 1 & 0 & 1 & 0 \\ 0 & 1 & \sqrt{2} & \sqrt{2} & 0 & 1 & 0 & 1 \end{pmatrix}$$

$$\alpha_i = [0.2857 \quad 0.2857 \quad 0.5714 \quad 0.5714 \quad 0.2857 \quad 0.2857 \quad 0.2857 \quad 0.2857]$$

5.1+4 (CICP16)

$$Q = \begin{pmatrix} 1 & 0 & \sqrt{2} & \sqrt{2} & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & \sqrt{2} & \sqrt{2} & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

$$\alpha_i =$$

$$[0.1818 \quad 0.1818 \quad 0.3636 \quad 0.3636 \quad 0.1818 \quad 0.1818 \quad 0.1818 \quad 0.1818 \quad 0.1818 \quad 0.1818]$$

7.1+4 (CICP19)

$$Q = \begin{pmatrix} 1 & 0 & \sqrt{2} & \sqrt{2} & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & \sqrt{2} & \sqrt{2} & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

α_i

= [0.1538 0.1538 0.3077 0.3077 0.1538 0.1538 0.1538 0.1538 0.1538 0.1538 0.1538 0.1538]

【0470】 6. 方法

【0471】 儘管以上技術主要被討論為構件或功能裝置，但是本發明也可以被實現為方法。以上討論的塊及元件也可以被理解為方法的步驟及/或階段。

【0472】 例如：提供一種用於從一降混訊號產生一合成訊號的解碼方法，該合成訊號具有一合成聲道數，該方法包括：

接收一降混訊號(246、x)，該降混訊號(246、x)具有一降混聲道數，及旁側資訊(228)，該旁側資訊(228)包括：

一原始訊號(212、y)的聲道位準及相關資訊(220)，該原始訊號(212、y)具有一原始聲道數；

使用該原始訊號(212、y)的該聲道位準及相關資訊(220)以及與該訊號(246、x)相關聯的協方差資訊(C_x)來產生該合成訊號。

【0473】 該解碼方法可以包括以下步驟中的至少一個：

從該降混訊號(246、x)計算一原型訊號，該原型訊號具有該合成聲道數；

使用該原始訊號的該聲道位準及相關資訊(212、y)以及與該降混訊號(246、x)相關聯的協方差資訊來計算一混合規則；及

使用該原型訊號及該混合規則來產生該合成訊號。

【0474】 還提供一種解碼方法，用於從具有一降混聲道數的一降混訊號(324、x)產生一合成訊號(336)，該降混訊號(336)具有一合成聲道數，該降混訊號(324、x)為具有一原始聲道數的一原始訊號(212)的一降混版本，該方法包括以下階段：

一第一階段(610c')，包括：

根據從以下內容計算出的一第一混合矩陣(M_M)合成該合成訊號的一第一分量(336M')：

與該合成訊號相關聯的一協方差矩陣(C_{y_R})(譬如該原始訊號的該協方差的該重建目標版本)；及

與該降混訊號(324)相關聯的一協方差矩陣(C_x)。

一第二階段(610c)，用於合成該合成訊號的一第二分量(336R')，其中該第二分量(336R')是一殘餘分量，該第二階段(610c)包括：

一原型訊號步驟(612c)，將該降混訊號(324)從該降混聲道數升混到該合成聲道數；

一去相關器步驟(614c)，將該被升混的原型訊號(613c)進行去相關；

一第二混合矩陣步驟(618c)，根據來自該降混訊號(324)的該去相關版本(615c)的一第二混合矩陣(M_R)合成該合成訊號的該第二分量(336R')，該第二混合矩陣(M_R)是一殘餘混合矩陣，

其中，該方法從以下內容計算該第二混合矩陣(M_R)：

由該第一混合矩陣步驟(600c)提供的該殘餘協方差矩陣(C_r)；及

從與該降混訊號(324)相關聯的該協方差矩陣(C_x)獲得的該被數個去相關的原型訊號(C_y)的該協方差矩陣的一估計，

其中該方法還包括一加法器步驟(620c)，將該合成訊號的該第一分量(336M')與該合成訊號的該第二分量(336R')相加，從而獲得該合成訊號(336)。

【0475】 此外，提供一種編碼方法，用於從一原始訊號(212、y)產生一降混訊號(246、x)，該原始訊號(212、y)具有一原始聲道數，該降混訊號(246、x)具有一降混聲道數，該方法包括：

估計(218)該原始訊號(212、y)的聲道位準及相關資訊(220)，
將該降混訊號(246、x)編碼(226)成一位元流(248)，使得該降混訊號(246、x)在該位元流(248)中被編碼，以便具有旁側資訊(228)，該旁側資訊(228)包括該原始訊號(12、y)的聲道位準及相關資訊(220)。

【0476】 這些方法可以在以上討論的任何編碼器及解碼器中被實現。

【0477】 7. 儲存單元(Storage units)

【0478】 此外，本發明可以在儲存諸多指令的一非暫時性儲存單元中被實現，該些指令在由該處理器執行時致使該處理器執行如上所述的一方法。

【0479】 此外，本發明可以在儲存諸多指令的一非暫時性儲存單元中被實現，該些指令在由該處理器執行時致使該處理器控制該編碼器或該解碼器的該諸多功能中的至少一者。

【0480】 該儲存單元可以例如是該編碼器200或該解碼器300的一部分。

【0481】 8. 其他方面

【0482】 儘管一些方面已經在一裝置的上下文中被描述，但是明顯的是，這些方面也代表該對應方法的一描述，其中一塊或裝置對應於一方法步驟或一方法步驟的一特徵。類似地，在一方法步驟的上下文中被描述的諸多方面也表

示一對應裝置的一相應塊或項目或特徵的一描述。方法步驟中的一些或全部可以由(或使用)一硬體裝置像是例如一微處理器、一可程式化電腦或一電子電路執行。在一些方面，這樣的一種裝置可以執行一些最重要的方法步驟中的一個或多個。

【0483】 取決於某些實現需求，本發明的諸多方面可以用硬體或軟體來實施。該實現可以使用一數位儲存介質被進行，例如軟性磁碟、一DVD、一CD、一ROM、一PROM、一EPROM、一EEPROM或一FLASH記憶體，其上儲存諸多電子可讀控制訊號，這些訊號與可程式化電腦系統協作(或能夠協作，使得該相應的方法被進行。因此，該數位儲存介質可以是電腦可讀的。

【0484】 根據本發明的一些方面包括一資料載體，該資料載體具有諸多電子可讀控制訊號，這些訊號能夠與一可程式化電腦系統協作，使得本文所述的方法之一被進行。

【0485】 一般而言，本發明的諸多方面可以被實現為具備一程式代碼的一電腦程式產品，當該電腦程式產品在一電腦上運行時，該程式代碼可操作於進行方法之一。該程式代碼可以例如被儲存在一機器可讀載體上。

【0486】 其他方面包括被儲存在機器可讀載體上的用於執行本文描述的諸多方法之一的該電腦程式。

【0487】 換句話說，因此，本發明方法的一方面是一種電腦程式，該電腦程式具有一程式代碼，當該電腦程式在一電腦上運行時，該程式代碼用於進行本文描述的諸多方法之一。

【0488】 因此，本發明方法的另一方面是一種資料載體(或一種數位儲存介質或一種電腦可讀介質)，包括被記錄在其上的該電腦程式，該電腦程式用於進

行本文描述的諸多方法之一。該資料載體、該數位儲存介質或該記錄介質通常是有形的及/或非暫時性的。

【0489】 因此，本發明方法的另一方面是一資料流或一訊號序列，代表用於執行本文描述的諸多方法之一的該電腦程式。該資料流或該訊號序列可以例如被配置成經由一資料通訊連接，例如經由網際網路。

【0490】 另一方面包括一處理裝置，例如一電腦或一可程式化邏輯裝置，被配置為或適應進行本文描述的諸多方法之一。

【0491】 另一方面包括一種電腦，該電腦上已安裝該電腦程式，用於進行本文描述的諸多方法之一。

【0492】 根據本發明的另一方面包括一種裝置或一種系統，該裝置或系統被配置成將用於進行本文描述的諸多方法之一的一電腦程式(譬如電子地或光學地)轉移到一接收器。該接收器可以是例如一電腦、一行動裝置、一記憶裝置或類似物。該裝置或系統可以例如包括一檔案伺服器，用於將該電腦程式轉移到該接收器。

【0493】 在一些方面，一可程式化邏輯裝置(例如一可程式化邏輯陣列)可以被用於進行本文描述的諸多方法的諸多功能中的一些或全部。在一些方面，一可程式化邏輯陣列可以與一微處理器協作，以便執行本文描述的諸多方法之一。通常，該方法較佳地由任何硬體裝置進行。

【0494】 本文描述的裝置可以使用一硬體設備或使用一電腦，或使用一硬體設備及一電腦的一組合來實現。

【0495】 本文描述的方法可以使用一硬體設備或使用一電腦，或使用一硬體設備及一電腦的一組合來執行。

【0496】 如上所述的諸多方面僅是對本發明的諸多原理的說明。應當理解的是，本文描述的佈置及細節的修改及變化對於所屬技術領域中具有通常知識者將是顯而易見的。因此，本發明的意向僅由即將來臨的專利請求項的範圍限制，而不受本文的各方面的描述及解釋所呈現的具體細節的限制。

【0497】 9. 參考書目

- [1] J. Herre, K. Kjörling, J. Breebart, C. Faller, S. Disch, H. Purnhagen, J. Koppens, J. Hilpert, J. Rödén, W. Oomen, K. Linzmeier and K. S. Chong, “MPEG Surround—The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding,” *Audio English Society*, vol. 56, no. 11, pp. 932-955, 2008.
- [2] V. Pulkki, “Spatial Sound Reproduction with Directional Audio Coding,” *Audio English Society*, vol. 55, no. 6, pp. 503-516, 2007.
- [3] C. Faller and F. Baumgarte, “Binaural Cue Coding - Part II: Schemes and Applications,” *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 6, pp. 520-531, 2003.
- [4] O. Hellmuth, H. Purnhagen, J. Koppens, J. Herre, J. Engdegård, J. Hilpert, L. Villemoes, L. Terentiv, C. Falch, A. Hölzer, M. L. Valero, B. Resch, H. Mundt and H.-O. Oh, “MPEG Spatial Audio Object Coding – The ISO/MPEG Standard for Efficient Coding of Interactive Audio Scenes,” in *AES*, San Fransisco, 2010.
- [5] L. Mikko-Ville and V. Pulkki, “Converting 5.1. Audio Recordings to B-Format for Directional Audio Coding Reproduction,” in *ICASSP*, Prague, 2011.
- [6] D. A. Huffman, “A Method for the Construction of Minimum-Redundancy Codes,” *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098-1101, 1952.
- [7] A. Karapetyan, F. Fleischmann and J. Plogsties, “Active Multichannel Audio Downmix,” in *145th Audio Engineering Society*, New York, 2018.
- [8] J. Vilkamo, T. Bäckström and A. Kuntz, “Optimized Covariance Domain Framework for Time-Frequency Processing of Spatial Audio,” *Journal of the Audio Engineering Society*, vol. 61, no. 6, pp. 403-411, 2013.

【符號說明】

1~10:索引順序

100:音訊系統

200:編碼器

212:原始訊號

214:濾波器組

216: 頻域訊號

218:參數估計器

220: 聲道位準及相關資訊

220k:增量

220s:縮放器

220t:當前的聲道位準及相關資訊

220(t-1):先前的聲道位準及相關資訊

220 Δ :差

222:參數量化塊

224:量化版本

226:位元流寫入器

228: 旁側資訊

230:核心編碼器與傳輸渠道

235:降混計算塊

244:降混器計算塊

246:降混訊號

247:核心編碼器

248:位元流

249:多工器

250:決定塊

251:命令

252:狀態資訊

254:命令

254':資訊

254s:開關

258:暫態分析塊

260:資訊

260':外部資訊

261:資訊

263:濾波器

264:頻域版本

265:分區分組塊

267:頻帶分析塊

268:命令

270:儲存元件

273:減法器

300: 解碼器

312:熵解碼器/輸入介面

314:量化參數

316: 參數重建模組

318:參數

320:濾波器組

322:降混訊號的一版本

324:降混訊號的頻域版本

326:原型訊號計算器

328:原型訊號

330:去相關模組

332:原型訊號

334:合成引擎

336:合成訊號

336M:主要分量

336M':主要分量

336R:殘餘分量

336R':殘餘分量

338:濾波器組

340:合成訊號

347:核心解碼器

380:頻帶/時隙分組塊

384:協方差估計塊

384':第一協方差估計器塊

385:降頻訊號

386:塊

388:協方差合成塊

388a:協方差合成塊

388b:協方差合成塊

388c:協方差合成塊

388d:協方差合成塊

390:協方差對同調度塊

392:ICC替換塊

394:能量施加塊

395:塊

402:混合規則計算器

403:混合規則

404:合成處理器

502:協方差估計器

504:協方差估計器

506:ICLD塊

508:訊號

510:協方差對同調度塊

512:訊號

600a:合成處理器

600b:合成處理器

600c:第一混合矩陣塊

610b:第二路徑

610b':第一路徑

610c: 第二路徑

610c':第一路徑

612b:升混塊

612c:升混塊

613b:原型訊號

613c:原型訊號

614b:去相關模組

614c:去相關模組

615b:去相關訊號

615c:去相關訊號

616b:去相關訊號

616c:去相關訊號

618b:最佳殘餘分量混合矩陣塊

618c:最佳殘餘分量混合矩陣塊

620b:加法器塊

620c:加法器塊

630:選擇器

631:開關

702:奇異值分解(SVD)

704:平方根

706: 乘法

710:估計

711:協方差

712:平方根

722:正規化/正則化

734:乘法

735:乘法結果

736:乘法

738:SVD

740:乘法

742:乘法

745: 逆的/正則化逆的

900:矩陣

902:非對角線值

904:非對角線值

905:非對角線值

906:非對角線值

907:非對角線值

908: 聲道間同調度(ICC)

920:訊框

- 921:時隙
- 922:時隙
- 923:時隙
- 924:時隙
- 930:訊框
- 931:時隙
- 932:時隙
- 933:時隙
- 934:時隙
- C:ICC
- C_r :矩陣
- C_x :協方差矩陣
- C_y :協方差矩陣
- $C_{\hat{y}}$:協方差矩陣
- C_{y_R} :原始協方差的重建目標版本
- $\widehat{C_y}$:估計協方差矩陣
- $\widehat{C_y'}$:矩陣
- M_R :混合矩陣
- I:單位矩陣
- K_r :矩陣
- K_y' :矩陣
- $\hat{K}_y$: 對角矩陣

$\widehat{K}_y$:矩陣

$\widehat{K}_y^{-1}$:矩陣

L:ICC

LS:ICC

P:矩陣

Q:原型規則

Q_N :原型訊號

Q_R :原型矩陣

R:ICC

RS:ICC

S_{Cr} :對角矩陣

U:左奇異向量矩陣

U_{Cr} :奇異向量矩陣

V:右奇異向量矩陣

X:降混訊號

X_B :降混訊號

Y:合成訊號

Y_B :訊號

Y_M :原型訊號

Y_R :合成訊號

$\widehat{Y}$:去相關訊號

ξ : 同調度

$\hat{\xi}$:同調度

ξ_R :同調度

χ :參數

χ_i :聲道間位準差(ICLD)

d:對角線值

f:頻率

t:訊框

x:降混訊號

y:原始訊號

L: 輸入聲道

L_S : 輸入聲道

R: 輸入聲道

R_S : 輸入聲道

C: 輸入聲道

LFE:輸入聲道

R_OTT:樹狀結構元件

ICC: 聲道間同調度

CLD: 聲道位準差

M:輸出訊號

res: 殘餘訊號

【發明申請專利範圍】

【請求項1】 一種音訊合成器(300)，用於從一降混訊號(246、 x)產生一合成訊號(336、340、 y_R)，該合成訊號(336、340、 y_R)具有一合成聲道數，該音訊合成器(300)包括：

一輸入介面(312)，被配置用於接收該降混訊號(246、 x)，該降混訊號(246、 x)具有一降混聲道數及旁側資訊(228)，該旁側資訊(228)包括一原始訊號(212、 y)的聲道位準及相關資訊(314、 ξ 、 χ)，該原始訊號(212、 y)具有一原始聲道數；及

一合成處理器(404)，被配置用於根據至少一個混合規則使用以下內容產生該合成訊號(336、340、 y_R)：

該原始訊號(212、 y)的聲道位準及相關資訊(220、314、 ξ 、 χ)；及
與該降混訊號(324、246、 x)相關聯的協方差資訊(C_x)；

該音訊合成器被配置成重建(386)該原始訊號的一目標協方差資訊(C_y)；

該音訊合成器被配置成基於該原始協方差資訊(C_y)的一估計版本($\widehat{C_y}$)重建該協方差資訊(C_y)的該目標版本(C_{y_R})，其中該原始協方差資訊(C_y)的該估計版本($\widehat{C_y}$)被通報給該合成聲道數或該原始聲道數；

該音訊合成器被配置成從與該降混訊號(324、246、 x)相關聯的協方差資訊(C_x)獲得該原始協方差資訊的該估計版本($\widehat{C_y}$)；

該音訊合成器被配置成在該降混訊號(324、246、 x)的該旁側資訊(228)、聲道位準及相關資訊(ξ 、 χ)之中進行檢索，

該音訊合成器還被配置成由來自以下兩者的該原始聲道位準及相關資訊(220)的一估計版本($\widehat{C_y}$)重建該協方差資訊(C_y)的該目標版本(C_{y_R})：

用於至少一個第一聲道或聲道對的協方差資訊(C_x)；及

用於至少一個第二聲道或聲道對的聲道位準及相關資訊(ξ 、 χ)。

【請求項2】 如請求項 1 所述的音訊合成器(300)，包括：

一原型訊號計算器(326)，被配置用於從該降混訊號(324、246、 x)計算

一原型訊號(328)，該原型訊號(328)具有該合成聲道數；

一混合規則計算器(402)，被配置用於使用以下內容計算至少一個混合規則(403)：

該原始訊號(212、 y)的該聲道位準及相關資訊(314、 ξ 、 χ)；及

與該降混訊號(324、246、 x)相關聯的該協方差資訊(C_x)；

其中該合成處理器(404)被配置用於使用該原型訊號(328)及該至少一個混合規則(403)產生該合成訊號(336、340、 y_R)。

【請求項3】 如請求項 2 所述的音訊合成器，被配置成重建適應於該合成訊號(336、340、 y_R)的該聲道數的該目標協方差資訊(C_y)。

【請求項4】 如請求項 3 所述的音訊合成器，被配置成通過將數個原始聲道群組指派給數個單一合成聲道，以重建適應於該合成訊號(336、340、 y_R)的該聲道數的該協方差資訊(C_y)，或者反之亦然，以便該重建目標協方差資訊(C_{y_R})被通報給該合成訊號(336、340、 y_R)的該聲道數。

【請求項5】 如請求項 4 所述的音訊合成器，被配置成通過產生針對該些原始聲道數的該目標協方差資訊並且後續應用一降混規則或一升混規則以及一能量補償，以得出針對該數個合成聲道的該目標協方差，以重建適應於該合成訊號(336、340、 y_R)的該聲道數的該協方差資訊(C_y)。

【請求項6】 如請求項 1 所述的音訊合成器，被配置成通過將一估算規則(Q)應用於與該降混訊號(324、246、 x)相關聯的該協方差資訊(C_x)，該估算規則(Q)是或被關聯到用於計算該原型訊號(326)的一原型規則，以獲

得該原始協方差資訊(220)的該估計版本($\widehat{C}_y$)。

【請求項7】 如請求項 1 或 6 所述的音訊合成器，被配置成針對至少一個聲道對，將該原始協方差資訊(C_y)的該估計版本($\widehat{C}_y$)正規化為該聲道對中的該數個聲道的該數個位準的該數個平方根。

【請求項8】 如請求項 7 所述的音訊合成器，被配置成以該原始協方差資訊(C_y)的正規化估計版本($\widehat{C}_y$)理解一矩陣。

【請求項9】 如請求項 8 所述的音訊合成器，被配置成通過插入在該位元流(248)的該旁側資訊(228)中所獲得的數個元(908)來完成該矩陣。

【請求項10】 如請求項 7 所述的音訊合成器，被配置成通過按形成該聲道對的該數個聲道的該數個位準的該平方根來縮放該原始協方差資訊(C_y)的該估計版本($\widehat{C}_y$)，將該矩陣進行去正規化。

【請求項11】 如請求項 1 所述的音訊合成器，被配置成偏好該聲道位準及相關資訊(ξ 、 χ)描述從該位元流(248)的該旁側資訊(228)所獲得的該聲道或聲道對，而不是針對相同聲道或聲道對從該降混訊號(324、246、 x)被重建的該協方差資訊(C_y)。

【請求項12】 如請求項 1 所述的音訊合成器，其中該原始協方差資訊(C_y)的該重建目標版本(C_{yR})描述在一聲道對之間的一能量關係，或是至少部分地基於被關聯到該聲道對中的每個聲道的數個位準。

【請求項13】 如請求項 1 所述的音訊合成器，被配置成獲得該降混訊號(246、 x)的一頻域版本(324)，該降混訊號(246、 x)的該頻域版本(324)被劃分為數個頻帶或數個頻帶群組，其中不同聲道位準及相關資訊(220)與不同頻帶或頻帶群組相關聯，

其中該音訊合成器被配置成針對不同頻帶或頻帶群組進行不同操作，以獲得針對不同頻帶或頻帶群組的不同混合規則。

【請求項14】 如請求項 1 所述的音訊合成器，其中該降混訊號(324、246、x)被劃分為數個時隙，其中不同的聲道位準及相關資訊(220)與不同時隙相關聯，並且該音訊合成器被配置成針對不同時隙進行不同操作，以獲得針對不同時隙的不同混合規則(403)。

【請求項15】 如請求項 1 所述的音訊合成器，其中該降混訊號(324、246、x)被劃分為數個訊框，並且每個訊框被劃分為數個時隙，其中當在一個訊框中的暫態的存在及位置被發訊表明為在一個暫態時隙中，該音訊合成器被配置成：

將該當前的聲道位準及相關資訊(220)與該暫態時隙及/或該訊框的暫態時隙後續的數個時隙相關聯；及

將該暫態時隙以前的該訊框的時隙與該以前的時隙的該聲道位準及相關資訊(220)相關聯。

【請求項16】 如請求項 1 所述的音訊合成器，被配置成選擇一原型規則(Q)，該原型規則(Q)被配置用於在該合成聲道數的基礎上計算一原型訊號(328)。

【請求項17】 如請求項 16 所述的音訊合成器，被配置成在數個預存原型規則之中選擇該原型規則(Q)。

【請求項18】 如請求項 1 所述的音訊合成器，被配置成在一手動選擇的基礎上定義一原型規則(Q)。

【請求項19】 如請求項 17 或 18 所述的音訊合成器，其中該原型規則包括一矩陣(Q)，該矩陣(Q)具備一第一維度及一第二維度，其中該第一維度與該降混聲道數相關聯，並且該第二維度與該合成聲道數相關聯。

【請求項20】 如請求項 1 所述的音訊合成器，被配置成操作在等於或低於 160 千位元/秒的一位元率。

【請求項21】 如請求項 1 所述的音訊合成器，還包括一熵解碼器(312)，用於獲得具備該旁側資訊(314)的該降混訊號(246、x)。

【請求項22】 如請求項 1 所述的音訊合成器，還包括一去相關模組(614b、614c、330)，以減少在不同聲道之間的相關量。

【請求項23】 如請求項 1 所述的音訊合成器，其中該原型訊號(328)被直接提供給該合成處理器(600a、600b、404)，沒有進行去相關。

【請求項24】 如請求項 1 所述的音訊合成器，其中該原始訊號(212、y)的該聲道位準及相關資訊(ξ 、 χ)、該至少一個混合規則(403)及與該降混訊號(246、x)相關聯的該協方差資訊(C_x)中的至少一者為一矩陣形式。

【請求項25】 如請求項 1 所述的音訊合成器，被配置成通過奇異值分解來計算至少一個混合規則。

【請求項26】 如請求項 1 所述的音訊合成器，其中該降混訊號被劃分為數個訊框，該音訊合成器被配置成使用針對一在前的訊框所獲得的具備一參數的一線性組合、一被估計或被重建的值或一混合矩陣來平滑一被接收的參數、一被估計或被重建的值或一混合矩陣。

【請求項27】 如請求項 26 所述的音訊合成器，被配置成當在一個訊框中的一暫態的存在及/或位置被發訊表明(261)時，停用該被接收的參數、該被估計或被重建的值或該混合矩陣的平滑。

【請求項28】 如請求項 1 所述的音訊合成器，其中該合成聲道數大於該原始聲道數。

【請求項29】 如請求項 1 所述的音訊合成器，其中該合成聲道數小於該原始聲道數。

【請求項30】 如請求項 1 所述的音訊合成器，其中該合成聲道數、該原始聲道數及該降混聲道數中的至少一個為一複數。

【請求項31】 如請求項 1 所述的音訊合成器，其中該至少一個混合規則包括一第一混合矩陣(M_M)及一第二混合矩陣(M_R)，該音訊合成器包括：

一第一路徑(610c')，包括：

一第一混合矩陣塊(600c)，被配置用於根據從以下內容計算出的該第一混合矩陣(M_M)來合成該合成訊號的一第一分量(336M')：

與該合成訊號(212)相關聯的一協方差矩陣(C_{y_R})，該協方差矩陣(C_{y_R})是從該聲道位準及相關資訊(220)被重建；及

與該降混訊號(324)相關聯的一協方差矩陣(C_x)，

一第二路徑(610c)，用於合成該合成訊號的一第二分量(336R')，該第二分量(336R')是一殘餘分量，該第二路徑(610c)包括：

一原型訊號塊(612c)，被配置用於將該降混訊號(324)從該降混聲道數升混到該合成聲道數；

一去相關器(614c)，被配置用於將該被升混的原型訊號(613c)進行去相關；

一第二混合矩陣塊(618c)，被配置用於根據來自該降混訊號(324)的該去相關版本(615c)的一第二混合矩陣(M_R)來合成該合成訊號的該第二分量(336R')，該第二混合矩陣(M_R)為一殘餘混合矩陣，其中該音訊合成器(300)被配置成從以下內容估計(618c)該第二混合矩陣(M_R)：

由該第一混合矩陣塊(600c)提供的一殘餘協方差矩陣(C_r)；及

從與該降混訊號(324)相關聯的該協方差矩陣(C_x)獲得的該數個去相關原型訊號(C_y)的該協方差矩陣的一估計，

其中該音訊合成器(300)還包括一加法器塊(620c)，用於將該合成訊號的該第一分量(336M')與該合成訊號的該第二分量(336R')進行求和。

【請求項32】 一種音訊合成器(300)，用於從具有一降混聲道數的一降混訊號(324、x)產生一合成訊號(336)，該合成訊號(336)具有一合成聲道數，該降混訊號(324、x)是具有一原始聲道數的一原始訊號(212)的一降混版本，該音訊合成器(300)包括：

一第一路徑(610c')，包括：

一第一混合矩陣塊(600c)，被配置用於根據從以下計算出的一第一混合矩陣(M_M)以合成該合成訊號的一第一分量(336M')：

被關聯到該合成訊號(212)的一協方差矩陣(C_{y_R})；及

被關聯到該降混訊號(324)的一協方差矩陣(C_x)；

一第二路徑(610c)，用於合成該合成訊號的一第二分量(336R')，其中該第二分量(336R')是一殘餘分量，該第二路徑(610c)包括：

一原型訊號塊(612c)，被配置用於將該降混訊號(324)從該降混聲道數升混到該合成聲道數；

一去相關器(614c)，被配置用於對該被升混的原型訊號(613c)進行去相關；

一第二混合矩陣塊(618c)，被配置用於根據來自該降混訊號(324)的該去相關版本(615c)的一第二混合矩陣(M_R)以合成該合成訊號的該第二分量(336R')，該第二混合矩陣(M_R)是一殘餘混合矩陣，其中該音訊合成器(300)被配置成從以下內容計算(618c)該第二混合矩陣(M_R)：

由該第一混合矩陣塊(600c)提供的該殘餘協方差矩陣(C_r)；及

從被關聯到該降混訊號(324)的該協方差矩陣(C_x)獲得的該數個去相關的原型訊號(C_y)的該協方差矩陣的一估計，

其中該音訊合成器(300)還包括一加法器塊(620c)，用於將該合成訊號的

該第一分量(336M')與該合成訊號的該第二分量(336R')進行求和。

【請求項33】 如請求項 32 所述的音訊合成器，其中通過從被關聯到該合成訊號(212)的該協方差矩陣(C_{y_R})減去通過將該第一混合矩陣(M_M)應用於被關聯到該降混訊號(324)的該協方差矩陣(C_x)所獲得的一矩陣，來獲得該殘餘協方差矩陣(C_r)。

【請求項34】 如請求項 32 或 33 所述的音訊合成器，被配置成從以下內容定義該第二混合矩陣(M_R)：

一第二矩陣(K_r)，其通過分解被關聯到該合成訊號的該剩餘協方差矩陣(C_r)而被獲得；

一第一矩陣(K_y^{-1})，其是從該數個去相關的原型訊號(C_y)的該協方差矩陣的該估計(711)被獲得的一對角矩陣($\hat{K}_y$)的逆矩陣或正則化逆矩陣。

【請求項35】 如請求項 34 所述的音訊合成器，其中通過將該平方根函數(712)應用於該數個去相關的原型訊號(C_y)的該協方差矩陣的數個主對角元素，來獲得該對角矩陣($\hat{K}_y$)。

【請求項36】 如請求項 34 所述的音訊合成器，其中通過將奇異值分解(702)應用於被關聯到該合成訊號的該殘餘協方差矩陣(C_r)，來獲得該第二矩陣(K_r)。

【請求項37】 如請求項 34 所述的音訊合成器，被配置成通過將該第二矩陣(K_r)與從該數個去相關的原型訊號(C_y)的該協方差矩陣的該估計及一第三矩陣(P)所獲得的該對角矩陣($\hat{K}_y$)的逆矩陣($\hat{K}_y^{-1}$)或正則化逆矩陣進行相乘(742)，來定義該第二混合矩陣(M_R)。

【請求項38】 如請求項 37 所述的音訊合成器，被配置成通過將奇異值分解(738)應用於從該數個去相關的原型訊號(C_y)的該協方差矩陣的一正規化版本($\widehat{G}_y$)所獲得的一矩陣(K'_y)，其中該正規化是對該殘餘協方差矩陣

(C_r)、該對角矩陣($\hat{K}_y$)及該第二矩陣(K_r)的主對角線進行，來獲得該第三矩陣(P)。

【請求項39】 如請求項 32 所述的音訊合成器，被配置成從一第二矩陣及該第二矩陣的逆矩陣或正則化逆矩陣來定義該第一混合矩陣(M_M)，

其中通過分解被關聯到該降混訊號的該協方差矩陣來獲得該第二矩陣，
及

通過分解被關聯到該降混訊號的該重建目標協方差矩陣來獲得該第二矩陣。

【請求項40】 如請求項 32 所述的音訊合成器，被配置成從應用於被關聯到該降混訊號(324)的該協方差矩陣(C_x)所獲得的該矩陣的該數個對角元估計該數個去相關的原型訊號(C_y)的該協方差矩陣，在該原型塊(612c)處被使用的該原型規則(Q)用於將該降混訊號(324)從該降混聲道數升混到該合成聲道數。

【請求項41】 一種音訊合成器(300)，用於從一降混訊號(246、x)產生一合成訊號(336、340、 y_R)，該合成訊號(336、340、 y_R)具有一合成聲道數，該音訊合成器(300)包括：

一輸入介面(312)，被配置用於接收該降混訊號(246、x)，該降混訊號(246、x)具有一降混聲道數及旁側資訊(228)，該旁側資訊(228)包括一原始訊號(212、y)的聲道位準及相關資訊(314、 ξ 、 χ)，該原始訊號(212、y)具有一原始聲道數；及

一合成處理器(404)，被配置用於根據至少一個混合規則使用以下內容產生該合成訊號(336、340、 y_R)：

該原始訊號(212、y)的聲道位準及相關資訊(220、314、 ξ 、 χ)；及
與該降混訊號(324、246、x)相關聯的協方差資訊(C_x)；

其中該音訊合成器與該解碼器無關。

【請求項42】 一種音訊合成器(300)，用於從一降混訊號(246、 x)產生一合成訊號(336、340、 y_R)，該合成訊號(336、340、 y_R)具有一合成聲道數，該音訊合成器(300)包括：

一輸入介面(312)，被配置用於接收該降混訊號(246、 x)，該降混訊號(246、 x)具有一降混聲道數及旁側資訊(228)，該旁側資訊(228)包括一原始訊號(212、 y)的聲道位準及相關資訊(314、 ξ 、 χ)，該原始訊號(212、 y)具有一原始聲道數；及

一合成處理器(404)，被配置用於根據至少一個混合規則使用以下內容產生該合成訊號(336、340、 y_R)：

該原始訊號(212、 y)的聲道位準及相關資訊(220、314、 ξ 、 χ)；及

與該降混訊號(324、246、 x)相關聯的協方差資訊(C_x)；

其中數個頻帶被彼此聚合為數個聚合頻帶群組，其中關於該數個聚合頻帶群組的資訊被提供在該位元流(248)的旁側資訊(228)中，其中該原始訊號(212、 y)的該聲道位準及相關資訊(220、 ξ 、 χ)按每頻帶群組被提供，以便針對相同聚合頻帶群組的不同頻帶計算相同的至少一個混合矩陣。

【請求項43】 一種用於從一降混訊號產生一合成訊號的方法，該合成訊號具有一合成聲道數，該方法包括：

接收一降混訊號(246、 x)及旁側資訊(228)，該降混訊號(246、 x)具有一降混聲道數，該旁側資訊(228)包括：

一原始訊號(212、 y)的聲道位準及相關資訊(220)，該原始訊號(212、 y)具有一原始聲道數；

使用該原始訊號(212、 y)的聲道位準及相關資訊(220)及被關聯到該降

混訊號(246、x)的協方差資訊(C_x)產生該合成訊號。

【請求項44】 如請求項 43 所述的方法，該方法包括：

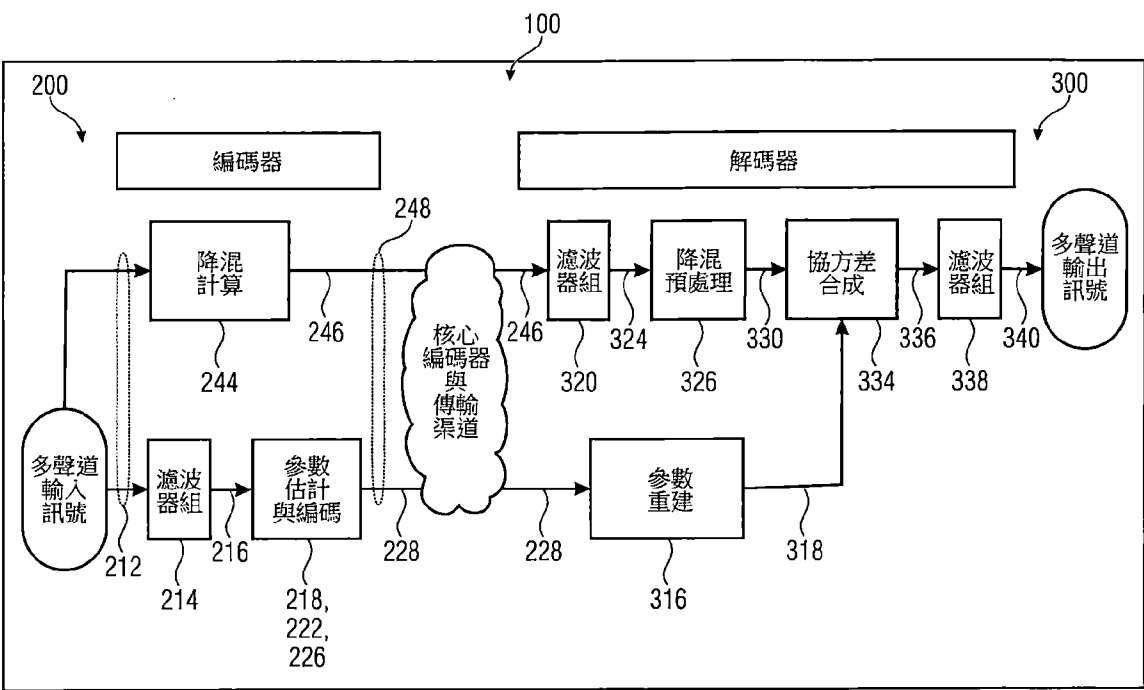
從該降混訊號(246、x)計算一原型訊號，該原型訊號具有該合成聲道數；

使用該原始訊號的聲道位準及相關資訊(212、y)及被關聯到該降混訊號(246、x)的協方差資訊計算一混合規則；及

使用該原型訊號及該混合規則產生該合成訊號。

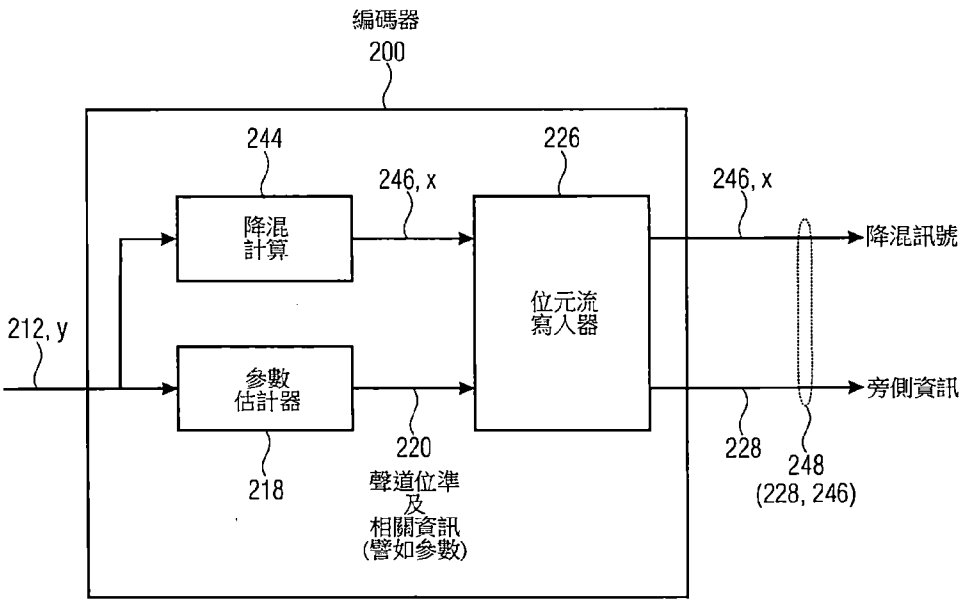
【請求項45】 一種儲存指令的非暫時性儲存單元，當該指令由一處理器執行時致使該處理器進行一種根據請求項 43 所述的方法。

【發明圖式】

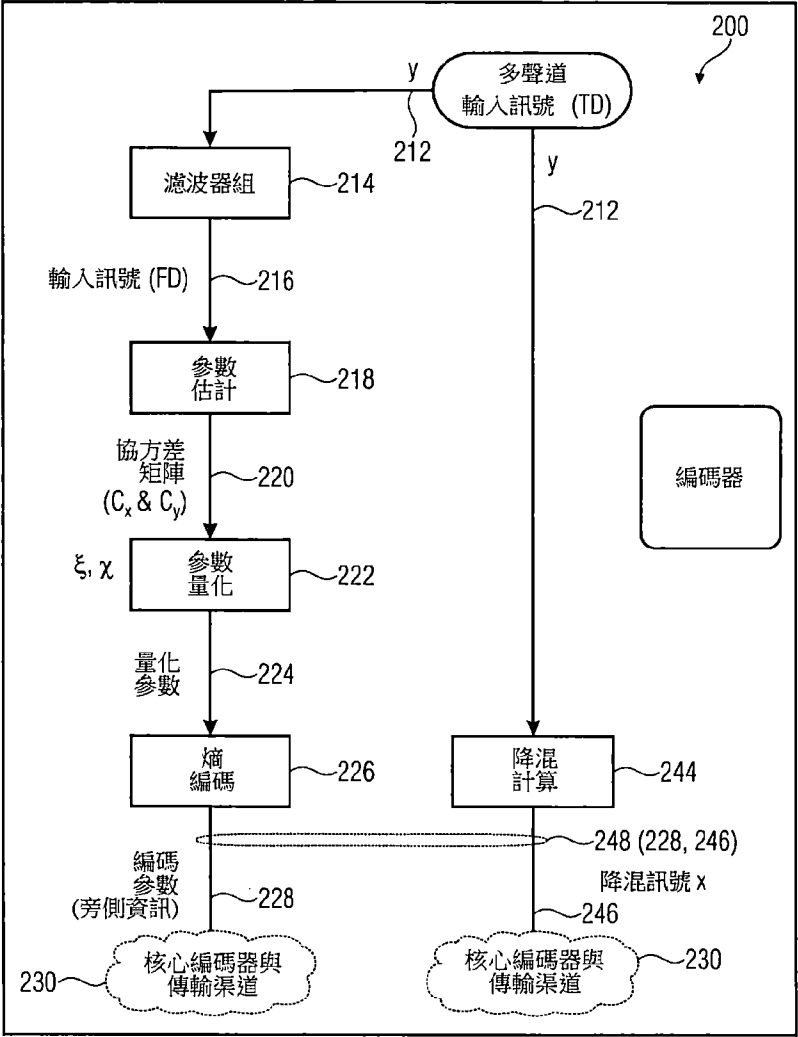


整體處理的簡化概觀

【第1圖】

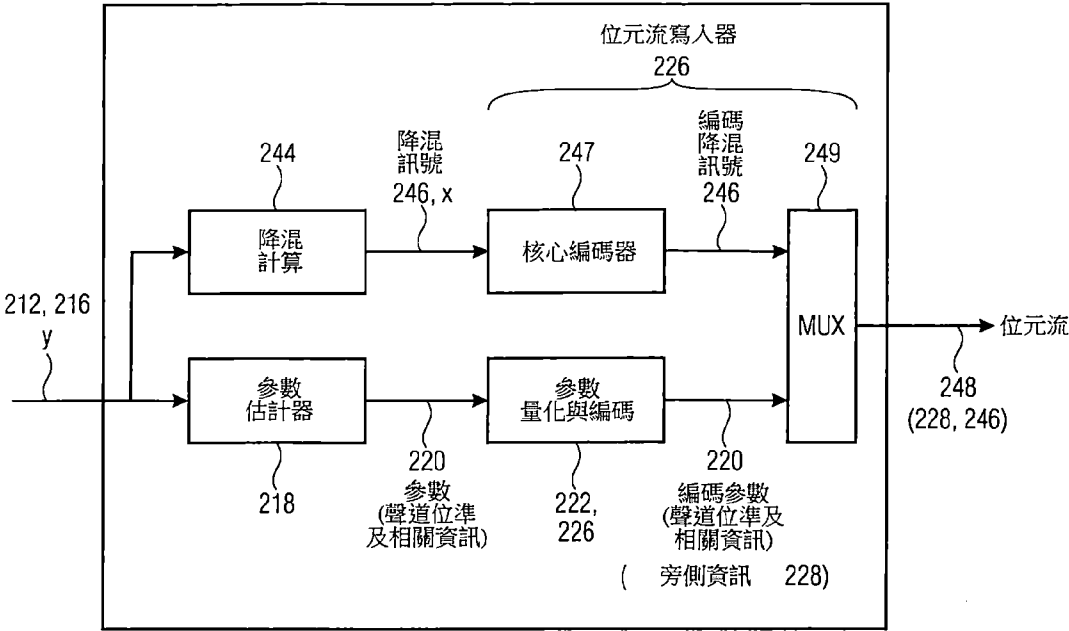


【第2a圖】

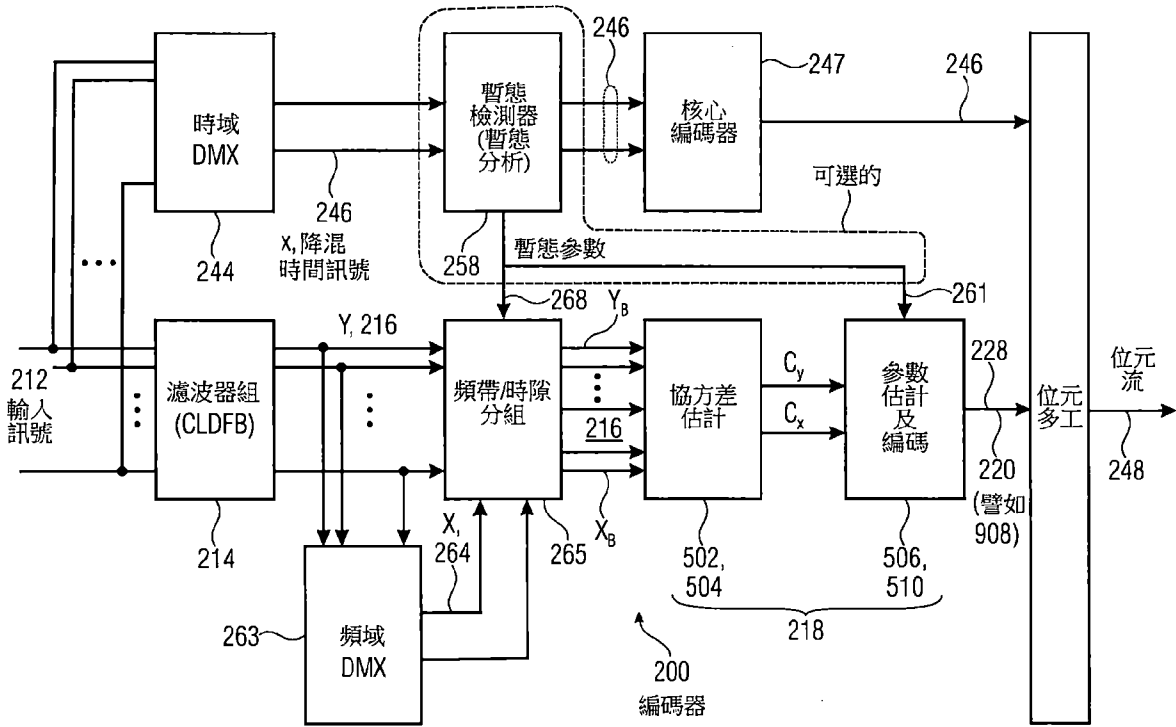


編碼器的詳細概觀

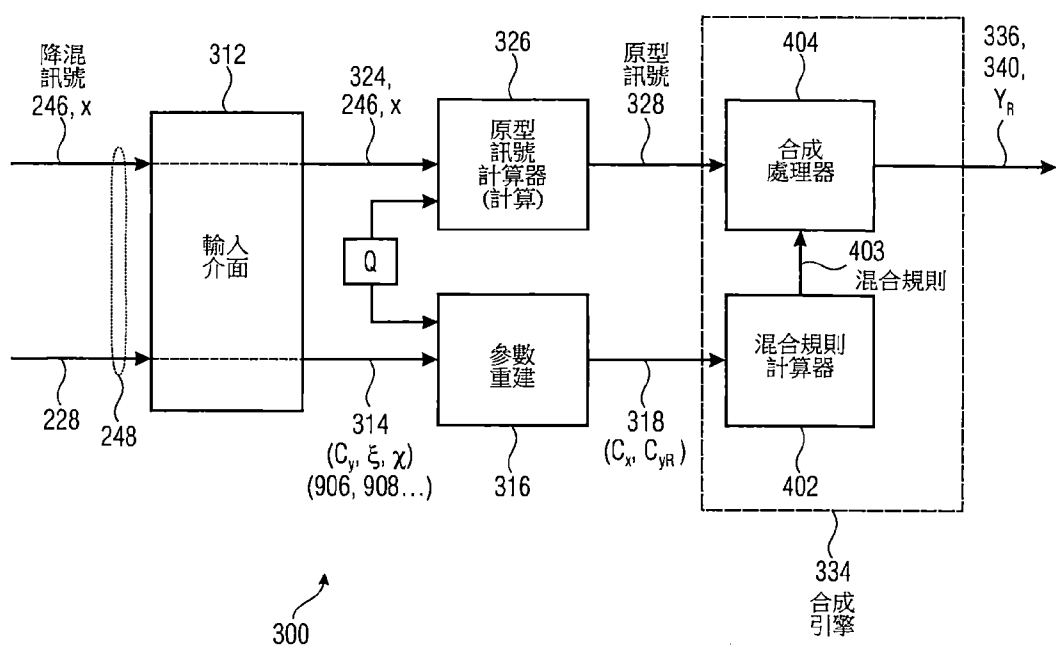
【第2b圖】



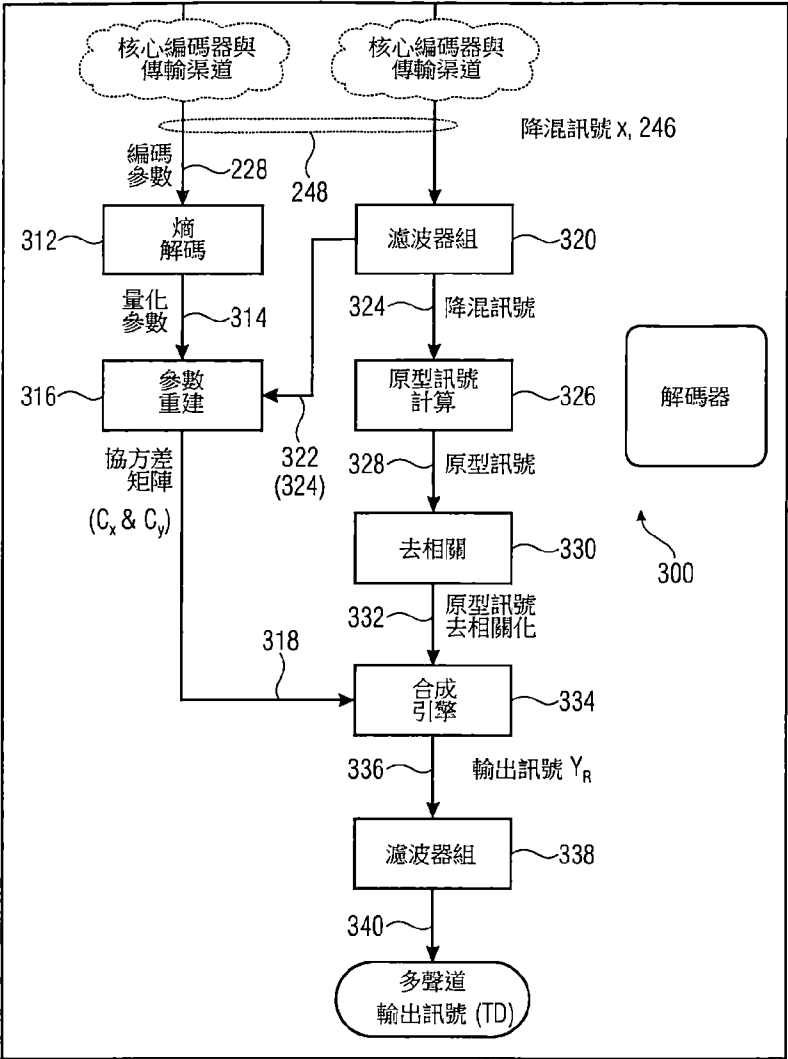
【第2c圖】



【第2d圖】

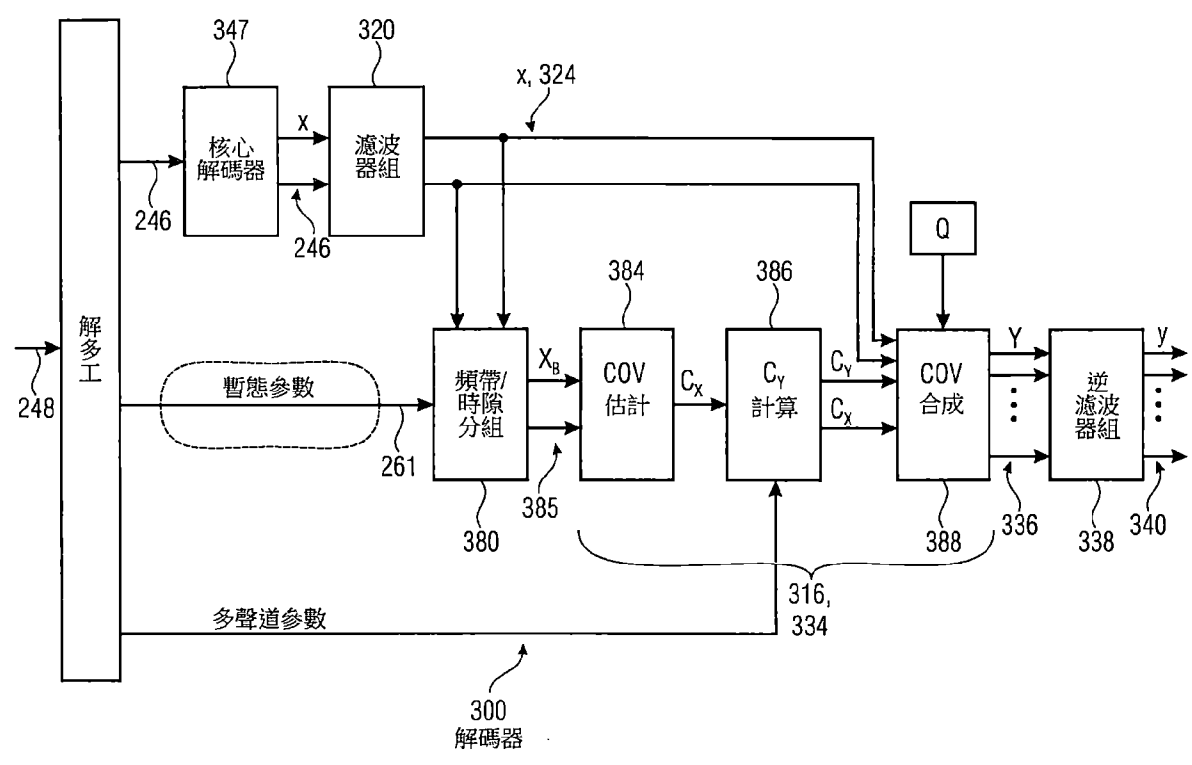


【第3a圖】

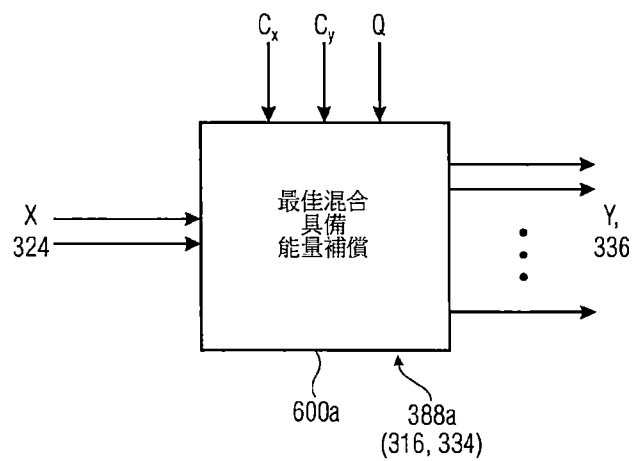


解碼器模組的概觀

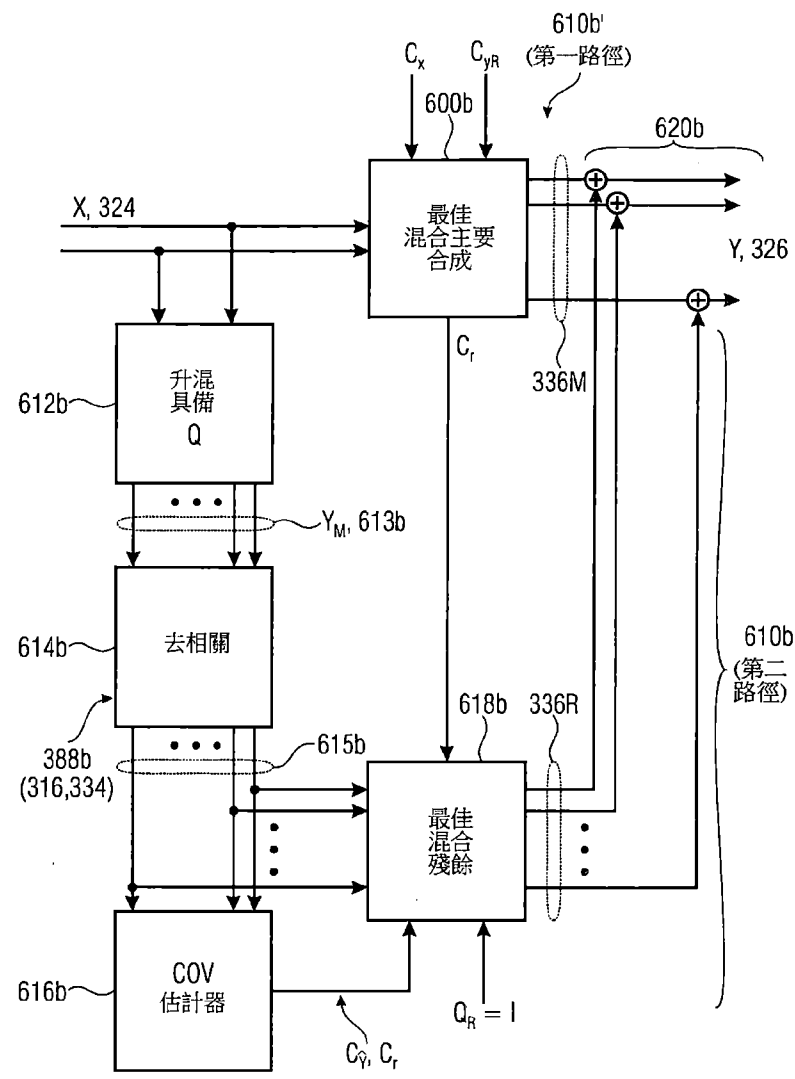
【第3b圖】



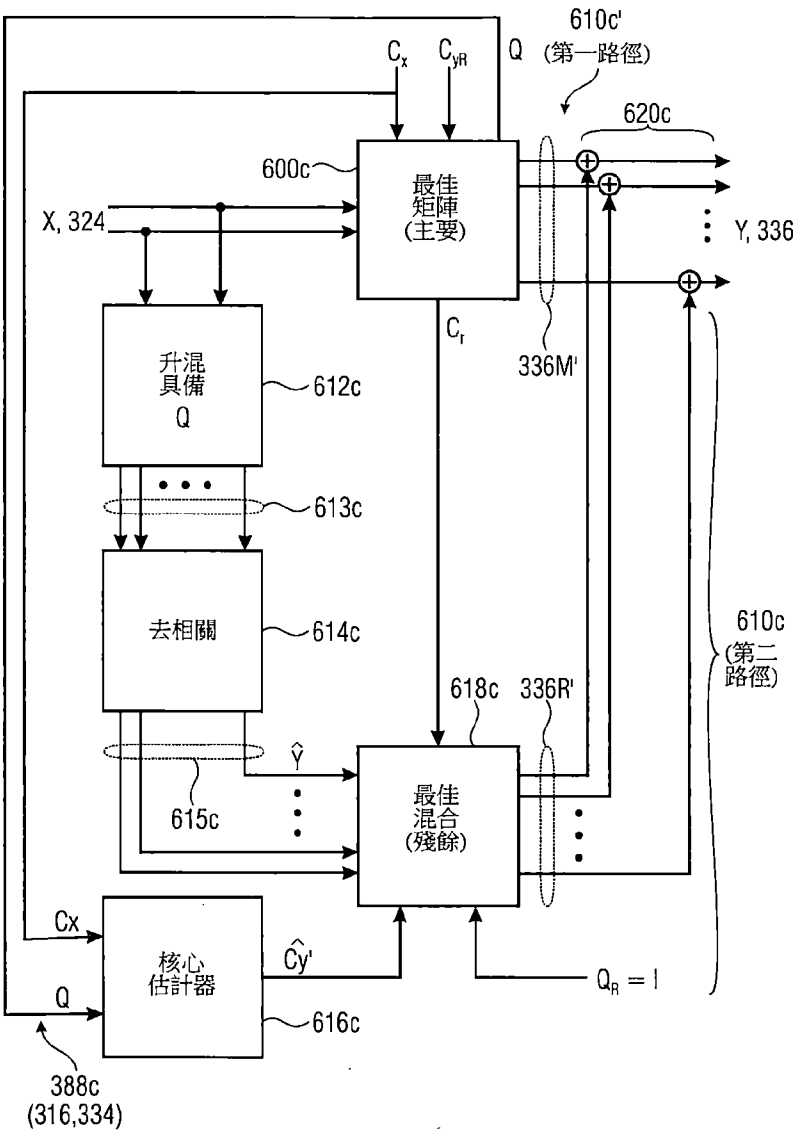
【第3c圖】



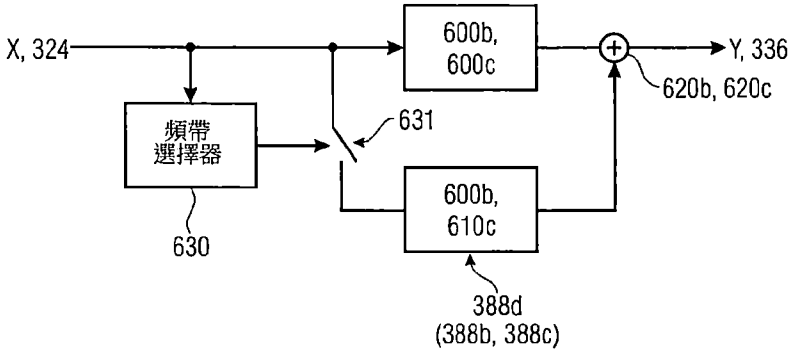
【第4a圖】



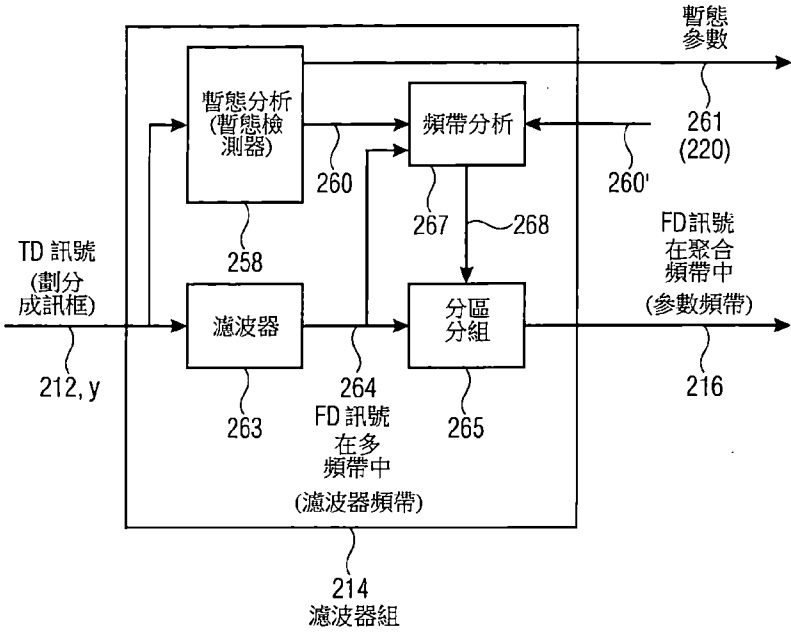
【第4b圖】



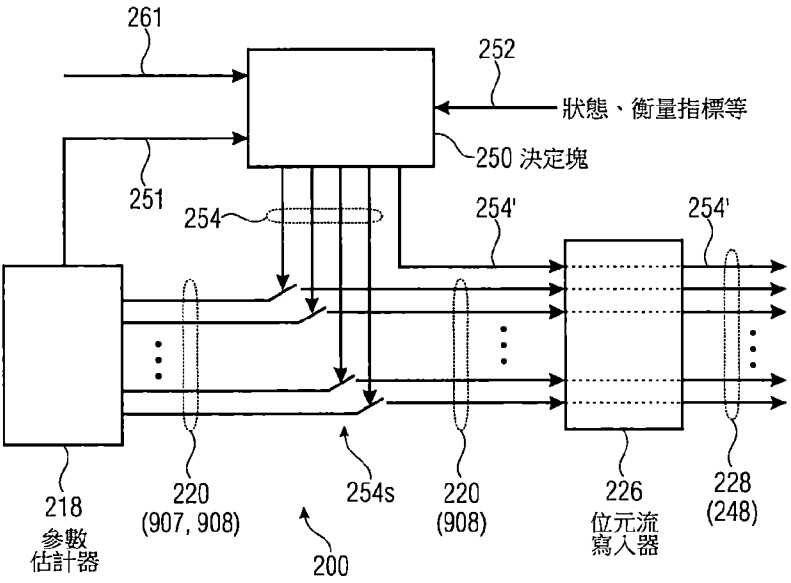
【第4c圖】



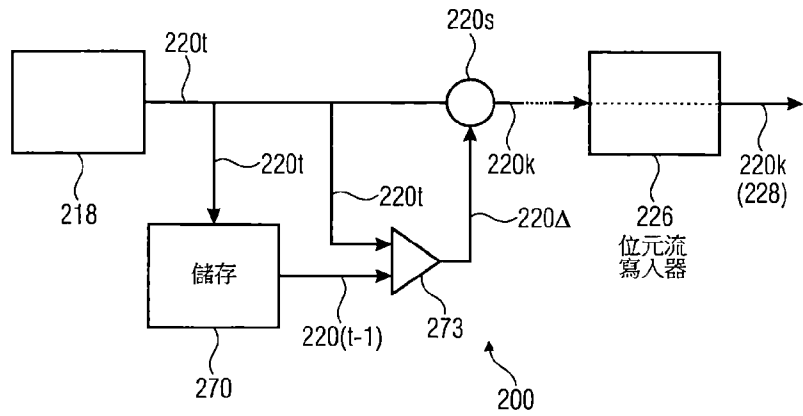
【第4d圖】



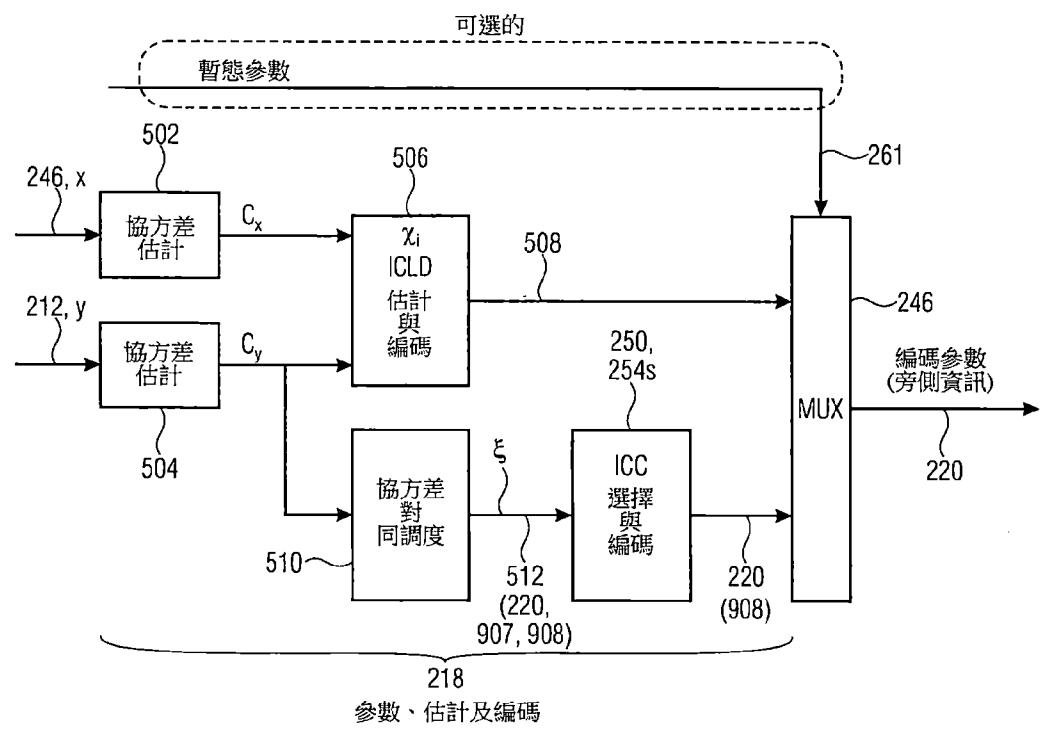
【第5圖】



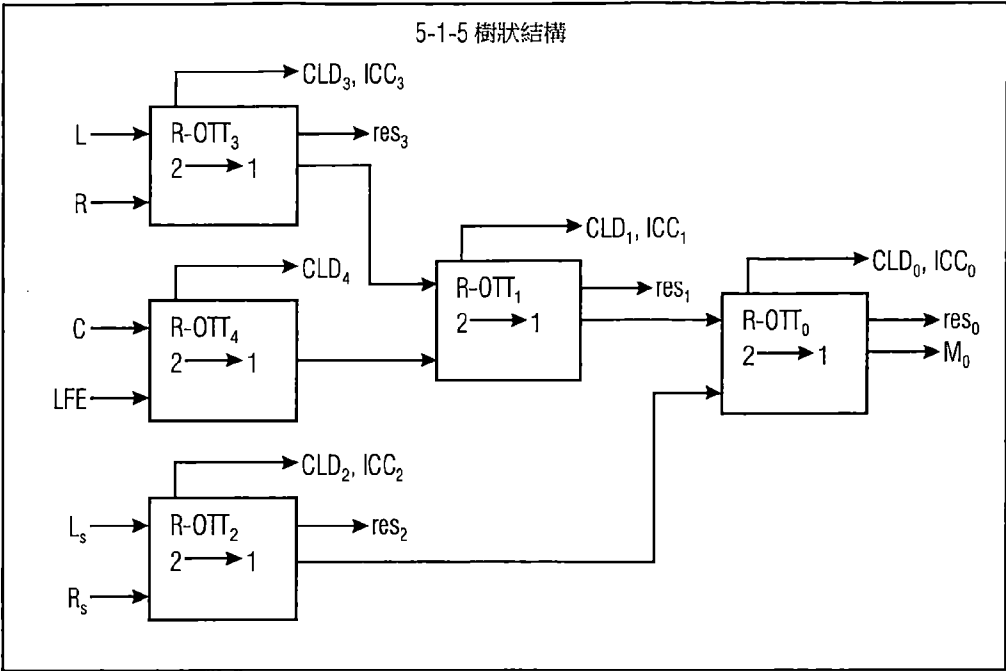
【第6a圖】



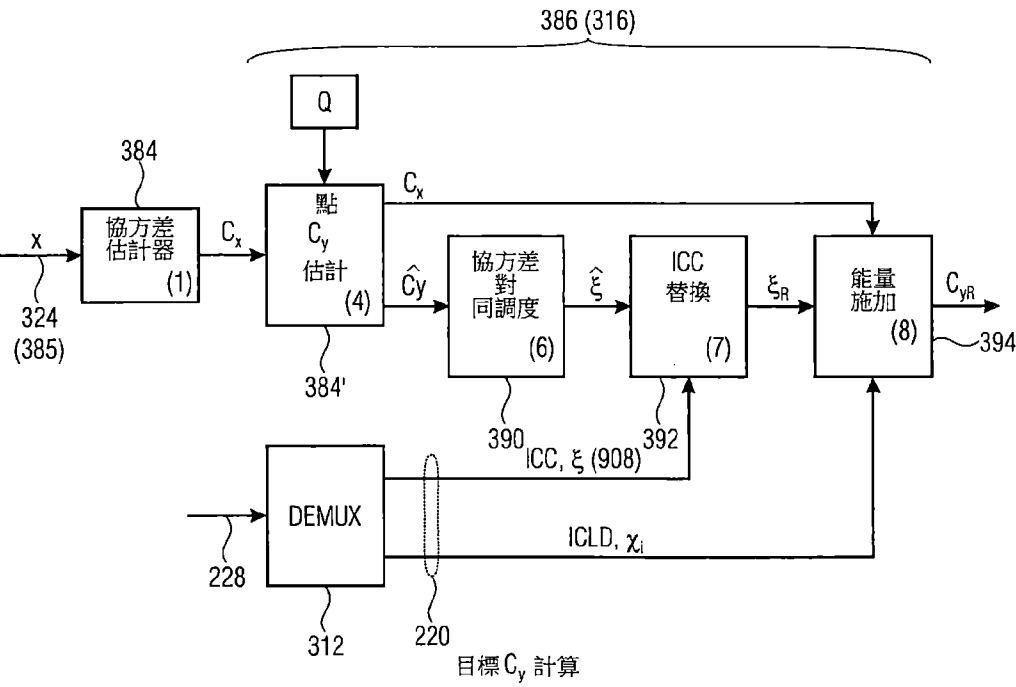
【第6b圖】



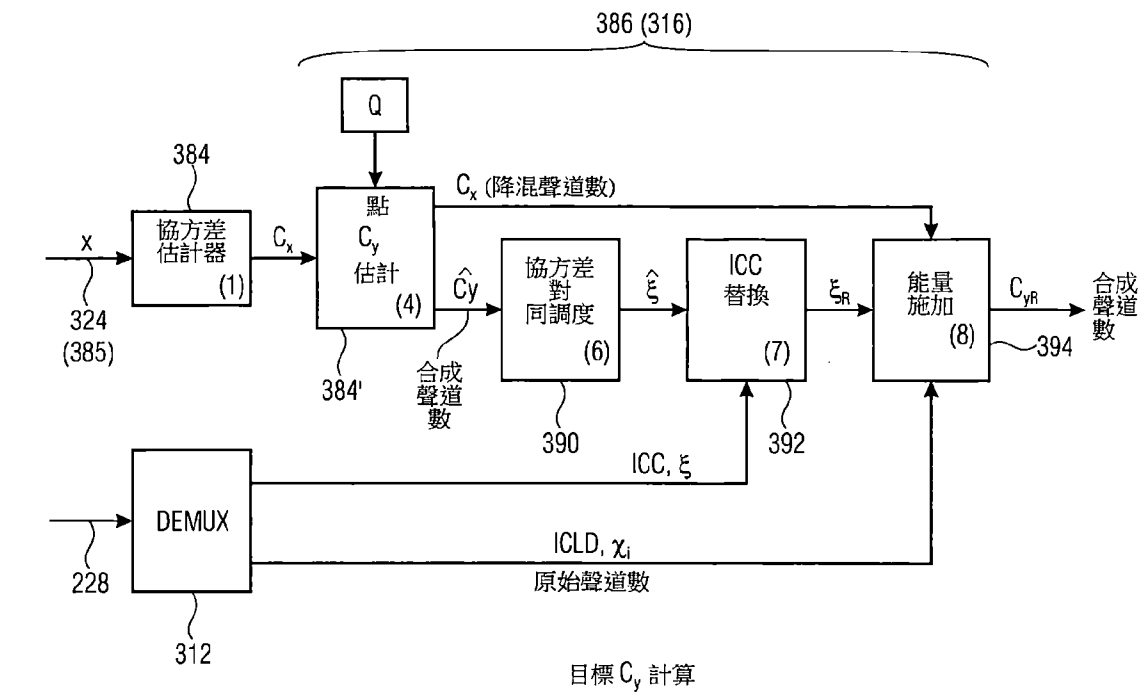
【第6c圖】



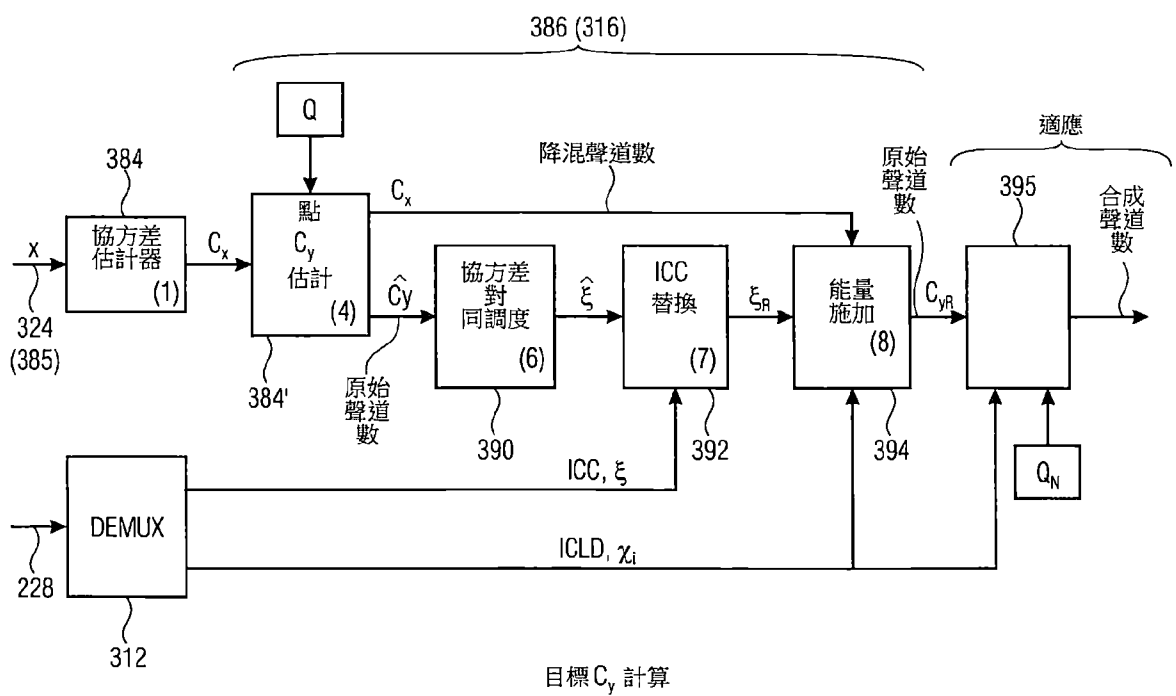
【第7圖】



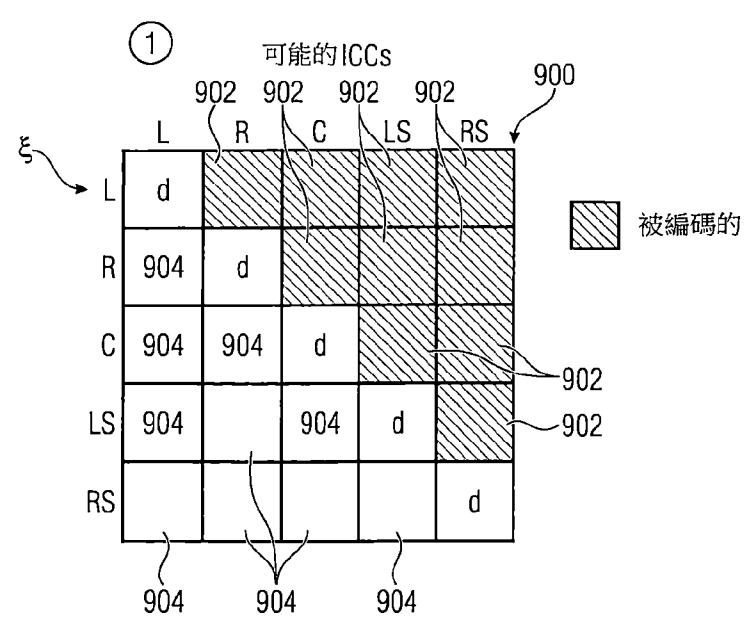
【第8a圖】



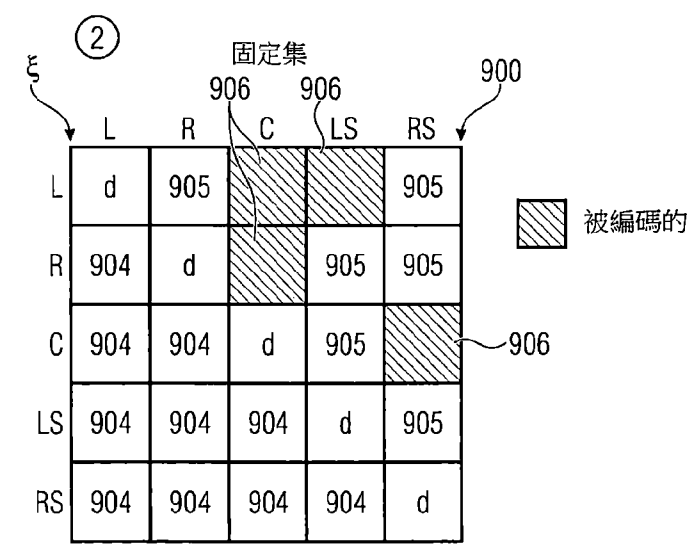
【第8b圖】



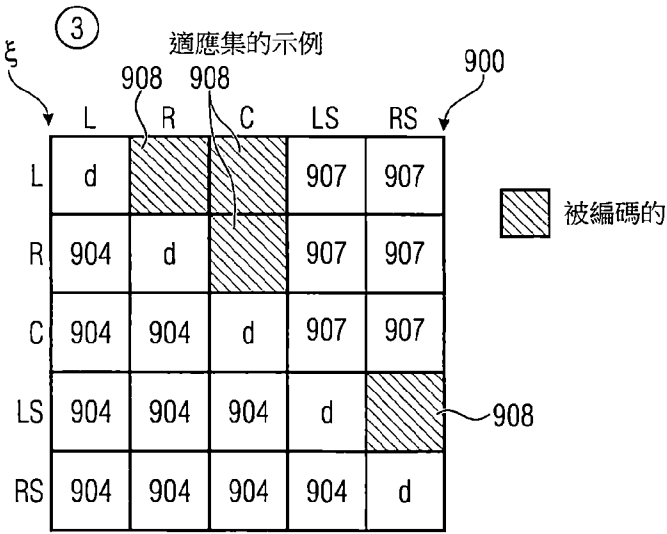
【第8c圖】



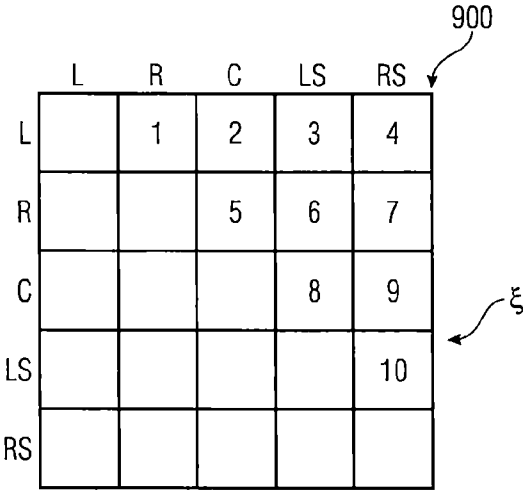
【第9a圖】



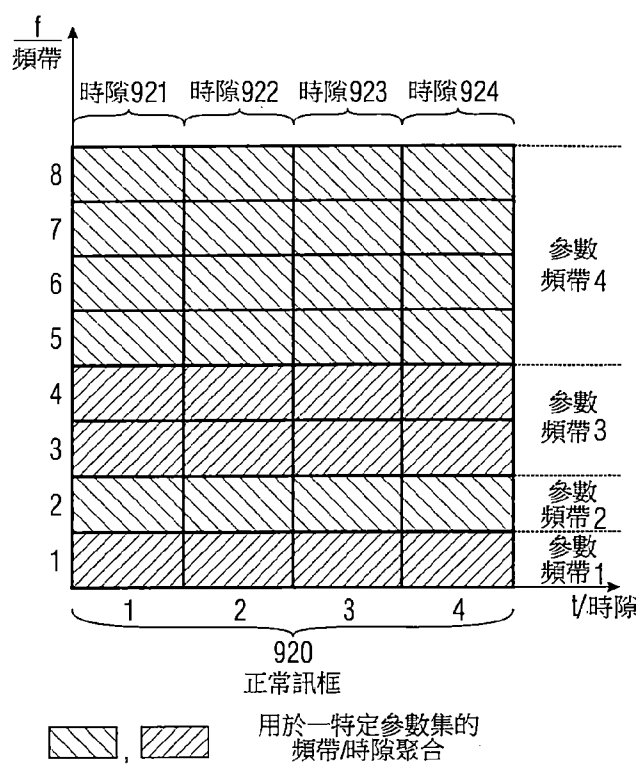
【第9b圖】



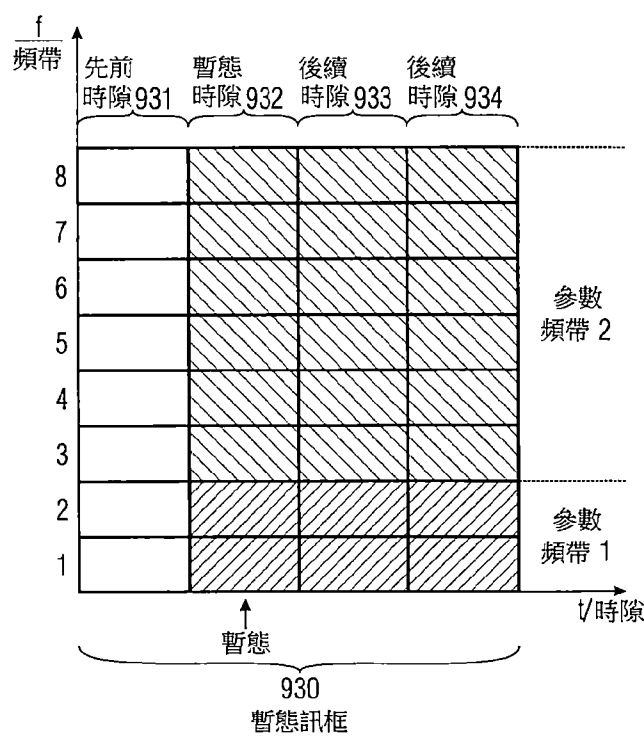
【第9c圖】



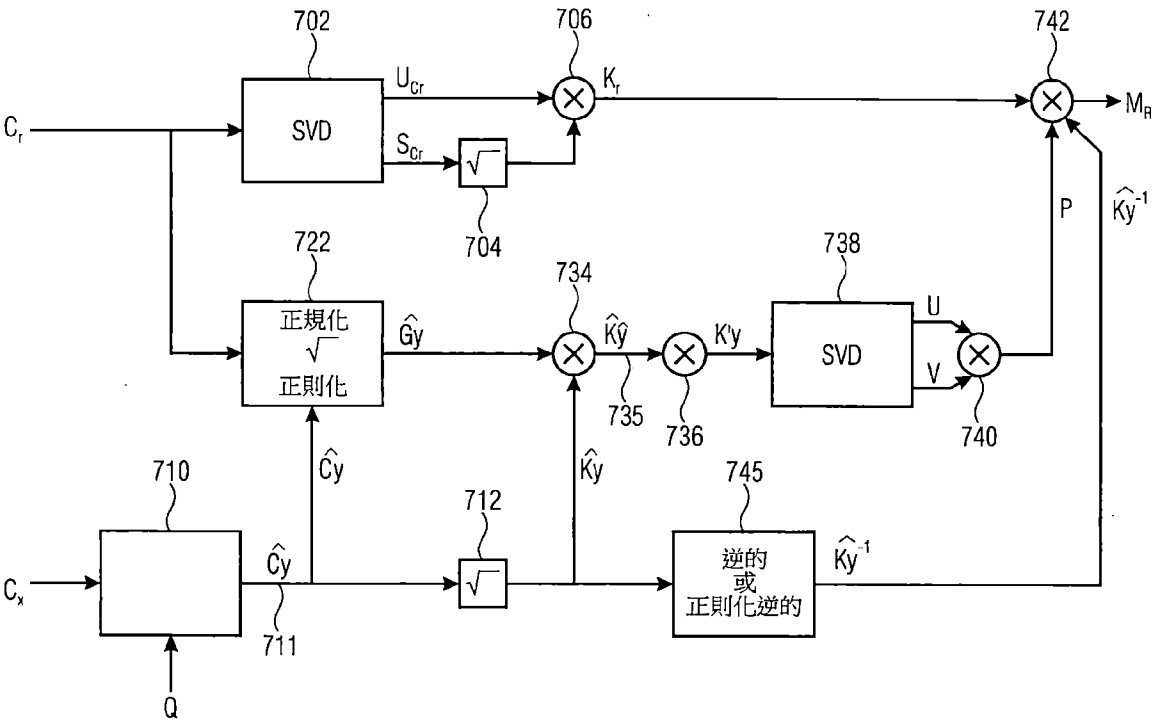
【第9d圖】



【第10a圖】



【第10b圖】



【第11圖】