



(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 I879690 B

(45)公告日：中華民國 114 (2025) 年 04 月 01 日

(21)申請案號：113139190

(22)申請日：中華民國 105 (2016) 年 02 月 22 日

(51)Int. Cl. : **G10L19/16 (2013.01)****G10L19/24 (2013.01)****G10L21/038 (2013.01)****G10L19/035 (2013.01)**

(30)優先權：2015/03/13 歐洲專利局

15159067.6

2015/03/16 美國

62/133,800

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

荷蘭

(72)發明人：維爾默斯 拉斯 VILLEMOS, LARS (DK)；普恩哈根 海克 PURNHAGEN, HEIKO

(DE)；伊斯坦德 皮爾 EKSTRAND, PER (SE)

(74)代理人：林志剛

(56)參考文獻：

TW 201246183A

TW 201438003A

US 8391371B2

審查人員：黃衍勳

申請專利範圍項數：8 項 圖式數：7 共 48 頁

(54)名稱

音訊處理單元、用於將經編碼的音訊位元流解碼之方法以及非暫態電腦可讀媒體

(57)摘要

實施例係關於音訊處理單元，其包括緩衝器、位元流負載去格式化器、以及解碼子系統。該緩衝器儲存經編碼的音訊位元流的至少一個區塊。該區塊包括填充元素，其起始於一識別符，其後接著填充資料。該填充資料包括第一旗標，識別是否對該經編碼的音訊位元流的該至少一個區塊的音訊內容執行頻譜帶複製處理的基本形式或頻譜帶複製處理的增強形式，以及若該第一旗標識別該頻譜帶複製處理的增強形式，則第二旗標識別是否致能或失能訊號自適應頻域超取樣。亦提出用於解碼一經編碼的音訊位元流的對應方法。

Embodiments relate to an audio processing unit that includes a buffer, bitstream payload deformatter, and a decoding subsystem. The buffer stores at least one block of an encoded audio bitstream. The block includes a fill element that begins with an identifier followed by fill data. The fill data includes a first flag identifying whether a base form of spectral band replication processing or an enhanced form of spectral band replication processing is to be performed on audio content of the at least one block of the encoded audio bitstream, and if the first flag identifies the enhanced form of spectral band replication processing, a second flag identifying whether signal adaptive frequency domain oversampling is enabled or disabled. A corresponding method for decoding an encoded audio bitstream is also provided.

指定代表圖：

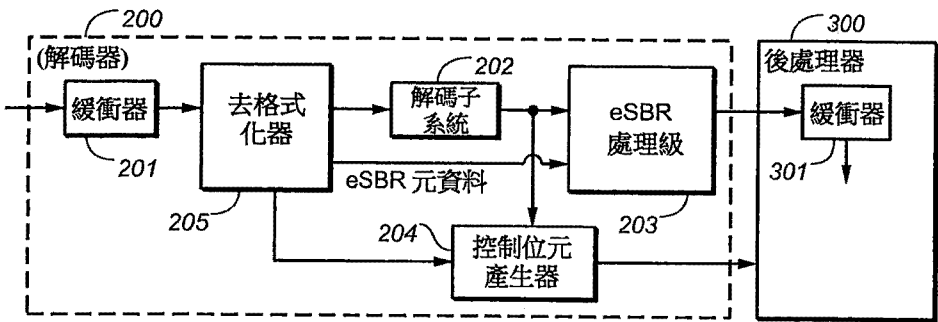


圖 3

符號簡單說明：

- 200:解碼器
- 201:緩衝器記憶體
- 202:音訊解碼子系統
- 203:eSBR 處理級
- 204:控制位元產生級
- 205:位元流負載去格式
化器(剖析器)
- 300:後處理器
- 301:緩衝器記憶體(緩
衝器)

I879690

【發明摘要】**【中文發明名稱】**

音訊處理單元、用於將經編碼的音訊位元流解碼之方法以及非暫態電腦可讀媒體

【英文發明名稱】

AUDIO PROCESSING UNIT, METHOD FOR DECODING AN
ENCODED AUDIO BITSTREAM, AND NON-TRANSITORY
COMPUTER READABLE MEDIUM

【中文】

實施例係關於音訊處理單元，其包括緩衝器、位元流負載去格式化器、以及解碼子系統。該緩衝器儲存經編碼的音訊位元流的至少一個區塊。該區塊包括填充元素，其起始於一識別符，其後接著填充資料。該填充資料包括第一旗標，識別是否對該經編碼的音訊位元流的該至少一個區塊的音訊內容執行頻譜帶複製處理的基本形式或頻譜帶複製處理的增強形式，以及若該第一旗標識別該頻譜帶複製處理的增強形式，則第二旗標識別是否致能或失能訊號自適應頻域超取樣。亦提出用於解碼一經編碼的音訊位元流的對應方法。

【 英文 】

Embodiments relate to an audio processing unit that includes a buffer, bitstream payload deformatter, and a decoding subsystem. The buffer stores at least one block of an encoded audio bitstream. The block includes a fill element that begins with an identifier followed by fill data. The fill data includes a first flag identifying whether a base form of spectral band replication processing or an enhanced form of spectral band replication processing is to be performed on audio content of the at least one block of the encoded audio bitstream, and if the first flag identifies the enhanced form of spectral band replication processing, a second flag identifying whether signal adaptive frequency domain oversampling is enabled or disabled. A corresponding method for decoding an encoded audio bitstream is also provided.

【代表圖】

【本案指定代表圖】：圖 3。

【本代表圖之符號簡單說明】：

200:解碼器

201:緩衝器記憶體

202:音訊解碼子系統

203:eSBR 處理級

204:控制位元產生級

205:位元流負載去格式化器(剖析器)

300:後處理器

301:緩衝器記憶體(緩衝器)

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：無

【發明說明書】

【中文發明名稱】

音訊處理單元、用於將經編碼的音訊位元流解碼之方法以及非暫態電腦可讀媒體

【英文發明名稱】

AUDIO PROCESSING UNIT, METHOD FOR DECODING AN
ENCODED AUDIO BITSTREAM, AND NON-TRANSITORY
COMPUTER READABLE MEDIUM

【技術領域】

本發明係關於音頻訊號處理。一些實施例係關於編碼及解碼音訊位元流(例如，具有 MPEG-4 AAC 格式之位元流)，其包括用於控制增強頻譜帶複製(eSBR)之元資料。其他實施例係關於藉由未被配置成執行 eSBR 處理以及忽略此種元資料的傳統解碼器解碼此種位元流，或者關於藉由回應於位元流產生 eSBR 控制資料來解碼不包括此種元資料的音訊位元流。

【先前技術】

典型音訊位元流包括指示音訊內容之一或多個通道的音訊資料(例如，經編碼的音訊資料)，以及指示音訊資料或音訊內容之至少一特徵的元資料二者。用於產生已編碼音訊位元流的一種公知格式是 MPEG-4 進階音訊編碼

(AAC)格式，其被描述於 MPEG 標準 ISO/IEC 14496-3:2009 中。在 MPEG-4 標準中，AAC 表示「進階音訊編碼(advanced audio coding)」以及 HE-AAC 表示「高效進階音訊編碼(high-efficiency advanced audio coding)」。

MPEG-4 AAC標準定義了幾種音訊規格(profile)，其決定兼容的編碼器及解碼器中存在哪些元件以及編碼工具。這些音訊規格的其中三種是(1)AAC規格、(2)HE-AAC規格及(3)HE-AAC v2規格。AAC規格包括AAC低複雜度(或“AAC-LC”)物件型式。AAC-LC物件係，藉由些許調整，對應於MPEG-2 AAC低複雜度規格，並且不包括頻譜帶複製(“SBR”)物件型式也不包括參數化立體聲(parametric stereo, “PS”)物件型式。HE-AAC規格是AAC規格的超集合，並且還包括SBR物件型式。HE-AAC v2規格是HE-AAC規格的超集合，並且還包括PS物件型式。

SBR物件型式包含頻譜帶複製工具，其係顯著提升感知音訊編解碼器之壓縮效率的重要編碼工具。SBR在接收方(例如，在解碼器中)重構音頻訊號的高頻分量。因此，編碼器僅需要編碼並傳輸低頻分量，允許低資料速率的較高音訊品質。SBR係依據由可用的有限頻寬訊號以及自編碼器獲得之控制資料複製諧波之序列，該諧波之序列事先被截斷以減少資料速率。音調分量及類噪聲分量的比率係由適應性逆濾波以及噪聲和正弦訊號之可選附加來維持。在MPEG-4 AAC標準中，SBR工具執行頻譜修補(patching)，其中將若干鄰接的正交鏡像濾波器(Quadrature Mirror

Filter, QMF)子帶從音頻訊號的傳輸的低頻帶部分複製到音頻訊號的高頻帶部分，該音頻訊號係產生於解碼器中。

頻譜修補對於某些音訊型式並不是理想的，例如具有相對低交越頻率的音樂內容。因此，需要用於改善頻譜帶複製的技術。

【發明內容】

第一類的實施例係關於音訊處理單元，其包括記憶體、位元流負載去格式化器(payload deformatter)、及解碼子系統。該記憶體被配置以儲存已編碼之音訊位元流(例如，MPEG-4 AAC位元流)的至少一個區塊。該位元流負載去格式化器被配置以解多工該經編碼的音訊區塊。該解碼子系統被配置以解碼該已編碼之音訊區塊的音訊內容。該經編碼的音訊區塊包括填充元素，其具有指示該填充元素之起始的標識符，以及包括在該標識符之後的填充資料。該填充資料包括至少一旗標，識別是否對該經編碼的音訊區塊的音訊內容執行增強形式頻譜帶複製處理(eSBR)。

第二類的實施例係關於用於解碼已編碼之音訊位元流的方法。該方法包括接收已編碼之音訊位元流的至少一個區塊、解多工該已編碼之音訊位元流的該至少一個區塊的至少某些部分、以及解碼該已編碼之音訊位元流的該至少一個區塊的至少某些部分。該已編碼之音訊位元流的該至少一個區塊包括填充元素，其具有指示該填充元素之起始的標識符，以及包括在該標識符之後的填充資料。該填充

資料包括至少一旗標，識別是否對該經編碼的音訊位元流的該至少一個區塊的音訊內容執行增強形式頻譜帶複製處理(eSBR)。

其他類的實施例係關於編碼及轉碼音訊位元流，該音訊位元流包含識別是否將執行增強頻譜帶複製(eSBR)處理的元資料。

【圖式簡單說明】

[圖 1]是系統之實施例的方塊圖，該系統被配置以執行本發明方法之實施例。

[圖 2]是編碼器的方塊圖，該編碼器是本發明音訊處理單元的實施例。

[圖 3]是包括解碼器之系統的方塊圖，該解碼器是本發明音訊處理單元的實施例，並且可選地有與其耦合的後處理器。

[圖 4]是解碼器的方塊圖，該解碼器是本發明音訊處理單元的實施例。

[圖 5]是解碼器的方塊圖，該解碼器是本發明音訊處理單元的另一實施例。

[圖 6]是本發明音訊處理單元之另一實施例的方塊圖。

[圖 7]是 MPEG-4 AAC 位元流之區塊的圖，包括該位元流被分割而成的區段。

【實施方式】

符號和術語

在整個本揭示內容中，包括在申請專利範圍中，對 (“on”)訊號或資料執行操作(例如，濾波、縮放、轉換、或施加增益至訊號或資料)的描述在廣義上用於表示直接對該訊號或資料執行操作，或對信號或資料之經處理後的版本(例如，對在操作執行之前已經過初步過濾或預處理的信號的版本)執行操作。

在整個本揭示內容中，包括在申請專利範圍中，「音訊處理單元」的表述在廣義上用於表示配置來處理音訊資料的系統、裝置或設備。音訊處理單元的範例包括但不限於編碼器(例如，轉碼器)、解碼器、編解碼器(codecs)、預處理系統、後處理系統、及位元流處理系統(有時被稱為位元流處理工具)。幾乎所有的消費性電子，例如行動電話、電視、膝上型電腦、及平板電腦，包含一音訊處理單元。

在整個本揭示內容中，包括在申請專利範圍中，「耦合」或「被耦合」的術語在廣義上用於指直接或間接連接之其中之一者。因此，若第一裝置耦合於第二裝置，該連接可能經由直接連接、或經由透過其他裝置或連接的間接連接。此外，整合進入其他元件或與其他元件整合的元件亦為彼此耦合。

本發明之實施例的詳細說明

MPEG-4 AAC 標準考量經編碼的 MPEG-4 AAC 位元流包括元資料，其指示解碼器將施用(若有將施用)來解碼位元流的音訊內容的 SBR 處理的各種類型，及/或其控制此 SBR 處理，及/或其指示將被採用來解碼位元流之音訊內容的至少一個 SBR 工具的至少一特徵或參數。本文中，使用“SBR 元資料”表述來表示在 MPEG-4 AAC 標準中描述或提及的此種類型的元資料。

MPEG-4 AAC 位元流的頂層是資料區塊的序列(“raw_data_block”元素)，各個資料區塊為包含音訊資料(典型用於 1024 或 960 個採樣的時間週期)及相關資訊及/或其他資料的資料的區段(本文中稱為“區塊(block)”)。本文中，使用“區塊”術語來表示 MPEG-4 AAC 位元流的區段，其包含決定或指示一個(但不超過一個)“raw_data_block”元素的音訊資料(及相應的元資料和可選地其他相關資料)。

MPEG-4 AAC 位元流的各個區塊可包括一些語法元素(各個語法元素亦在位元流中被具體化為資料的區段)。在 MPEG-4 AAC 標準中定義了此語法元素的七種類型。每個語法元素是由資料元素“id_syn_ele”的不同值來識別。語法元素的範例包括“single_channel_element()”、“channel_pair_element()”、及“fill_element()”。單聲道元素為一容器，包括單音訊通道的音訊資料(單聲道音頻訊號)。雙聲道元素包括兩個音訊通道的音訊資料(即，立體聲音頻訊號)。

填充元素為一資訊的容器，該資訊包括識別符(例如，上述元素“id_syn_ele”之值)緊接著資料(其被稱為“填充資料”)。填充元素歷來被用以調整將在固定速率通道上被傳輸的位元流的瞬時位元率。藉由將適當數量的填充資料加進各個區塊，可以達到固定資料速率。

依據本發明之實施例，填充資料可包括一或多個擴充負載(extension payload)，其擴充能在位元流中傳輸的資料的類型(例如，元資料)。接收具有包含新資料類型的填充資料的位元流的解碼器，可任選地被接收位元流的裝置(例如，解碼器)用來擴充該裝置的功能性。因此，如本領域之技術人員可理解的，填充元素為資料結構的特殊類型，且不同於典型用以傳輸音訊資料(例如，包含通道資料的音訊負載)的資料結構。

在本發明的某些實施例中，用以識別填充元素的識別符可由一三位元最高有效位元傳輸在先之無正負號整數(“uimsbf”)組成，其具有 0x6 的值。在一區塊中，可能出現相同類型之語法元素的多個實例(例如，多個填充元素)。

用於編碼音訊位元流之另一標準為 MPEG 聯合語音及音訊編碼(Unified Speech and Audio Coding, USAC)標準(ISO/IEC 23003-3:2012)。MPEG USAC 標準描述使用頻譜帶複製處理(包括 MPEG-4 AAC 標準中所述之 SBR 處理，且亦包括頻譜帶複製處理的其他增強形式)之音訊內容的編碼及解碼。此處理應用了 MPEG-4 AAC 標準中所描述

之 SBR 工具之集合的擴充及增強版的頻譜帶複製工具(在本文中有時被稱為“增強的 SBR 工具”或“eSBR 工具”)。因此，eSBR(如 USAC 標準中所定義)為 SBR(如 MPEG-4 AAC 標準中所定義)之改良。

本文中，使用“增強的 SBR 處理”(或“eSBR 處理”)之表述來表示使用在 MPEG-4 AAC 標準中未描述或提及的至少一個 eSBR 工具(例如，在 MPEG USAC 標準中描述或提及的至少一個 eSBR 工具)的頻譜帶複製處理。此種 eSBR 工具的範例為諧波移調(harmonic transposition)、QMF-修補(QMF-patching)額外預處理或“預平坦化(pre-flattening)”、及子帶間樣本時間包絡成型(Temporal Envelope Shaping)或“inter-TES”。

依據 MPEG USAC 標準所產生的位元流(在本文中有時被稱為“USAC 位元流”)包括經編碼的音訊內容，且典型地包括將由解碼器施用來解碼 USAC 位元流之音訊內容的頻譜帶複製處理的各個類型的元資料、及/或控制此頻譜帶複製處理及/或表示將被採用來解碼 USAC 位元流之音訊內容的至少一個 SBR 工具及/或 eSBR 工具之至少一個特徵或參數的元資料。

本文中，使用“增強的 SBR 元資料”(或“eSBR 元資料”)之表述來表示指示將由解碼器施用來解碼已編碼之音訊位元流(例如，USAC 位元流)之音訊內容的頻譜帶複製處理的各個類型的元資料、及/或控制此頻譜帶複製處理的元資料、及/或指示將被採用來解碼此音訊內容、但未

在 MPEG-4 AAC 標準中被描述或提及的至少一個 SBR 工具及/或 eSBR 工具之至少一個特徵或參數的元資料。eSBR 元資料之一範例為在 MPEG USAC 標準中被描述或提及但未在 MPEG-4 AAC 標準中被描述或提及的元資料(指示頻譜帶複製處理、或用於控制頻譜帶複製處理)。因此，eSBR 元資料在本文中表示非 SBR 元資料的元資料，而 SBR 元資料在本文中表示非 eSBR 元資料的元資料。

USAC 位元流可包括 SBR 元資料及 eSBR 元資料二者。更具體地，USAC 位元流可包括控制解碼器之 eSBR 處理效能的 eSBR 元資料、及控制解碼器之 SBR 處理效能的 SBR 元資料。依據本發明的典型實施例，eSBR 元資料(例如，eSBR 特定配置資料)係包含在(依據本發明)MPEG-4 AAC 位元流中(例如，在 SBR 負載之末端的 `sbr_extension()` 容器中)。

在使用 eSBR 工具集(包含至少一個 eSBR 工具)解碼一經編碼的位元流的期間，由解碼器執行 eSBR 處理，依據在編碼過程中被截斷之諧波序列的複製來重新產生音頻訊號的高頻帶。此種 eSBR 處理，典型地調整所產生的高頻帶的頻譜包絡，並施用反向濾波、及增加噪聲和正弦分量以重新建立原始音頻訊號的頻譜特性。

依據本發明的典型實施例，在經編碼的音訊位元流(例如，MPEG-4 AAC 位元流)之一或多個元資料區段中包含 eSBR 元資料(例如，包含係 eSBR 元資料的少數控制位元)，該經編碼的音訊位元流亦包含經編碼的音訊資料於

其他區段(音訊資料區段)中。典型地，位元流之每個區段的至少一個此種元資料區段係(或包括)一填充元素(包含一識別符，指示該填充元素的起始)，且 eSBR 元資料係包含在填充元素中、在識別符之後。

圖 1 是示例性的音訊處理鏈(音訊資料處理系統)之方塊圖，其中該系統之一或多個元件可依據本發明之實施例而被配置。該系統包括以下元件，耦合在一起如圖所示：編碼器 1、傳遞子系統 2、解碼器 3、及後處理單元 4。在所示系統的變型中，省略該等元件的其中一或多個，或者包含額外的音訊資料處理單元。

在一些實施方式中，編碼器 1(其可選地包括預處理單元)被配置成接受包含音訊內容的 PCM(時域)樣本作為輸入，並輸出表示音訊內容的經編碼的音訊位元流(具有符合 MPEG-4 AAC 標準的格式)。表示音訊內容的位元流資料在本文中有時被稱為“音訊資料”或“經編碼的音訊資料”。若依據本發明之典型實施例來配置編碼器，則自該編碼器輸出的音訊位元流包括 eSBR 元資料(並且典型地亦包括其他元資料)以及音訊資料。

自編碼器 1 輸出的一或多個經編碼的音訊位元流可被判斷提示(assert)至經編碼的音訊傳遞子系統 2。子系統 2 被配置成儲存及/或傳遞自編碼器 1 輸出的各個經編碼的位元流。自編碼器 1 輸出的經編碼的位元流可由子系統 2 儲存(例如，以 DVD 或藍光光碟的形式)，或由子系統 2 傳輸(其可實現傳輸鏈結或網路)、或由子系統 2 儲存並且

傳輸。

解碼器 3 被配置成解碼經編碼的 MPEG-4 AAC 音訊位元流(由編碼器 1 所產生)，其經由子系統 2 接收。在某些實施例中，解碼器 3 被配置成從位元流的各區塊抽取 eSBR 元資料，並解碼該位元流(包括藉由使用被抽取的 eSBR 元資料來執行 eSBR 處理)以產生經解碼的音訊資料(例如，經解碼的 PCM 音訊樣本的串流)。在某些實施例中，解碼器 3 被配置成從位元流抽取 SBR 元資料(但忽略位元流中所包含的 eSBR 元資料)，並解碼該位元流(包括藉由使用被抽取的 SBR 元資料來執行 SBR 處理)以產生經解碼的音訊資料(例如，經解碼的 PCM 音訊樣本的串流)。典型地，解碼器 3 包括緩衝器，該緩衝器儲存(例如，以非暫態的方式)從子系統 2 接收的經編碼的音訊位元流的區段。

圖 1 的後處理單元 4 被配置成接受來自解碼器 3 的經解碼的音訊資料的串流(例如，經解碼的 PCM 音訊樣本)，並對其執行後處理。後處理單元 4 亦可被配置成呈現經後處理的音訊內容(或從解碼器 3 接收的經解碼的音訊)用於由一或多個揚聲器播放。

圖 2 是編碼器(100)的方塊圖，該編碼器為本發明之音訊處理單元的實施例。編碼器 100 的任何組件或元件可被實現為硬體、軟體、或硬體與軟體之組合中的一或多個處理過程及/或一或多個電路(例如，ASICs、FPGAs、或其他積體電路)。編碼器 100 包括編碼器 105、填充器/格式

化器級 107、元資料產生級 106、及緩衝器記憶體 109，如圖所示連接。典型地，編碼器 100 亦包括其他處理元件(未示出)。編碼器 100 被配置成將輸入音訊位元流轉換成經編碼的輸出 MPEG-4 AAC 位元流。

元資料產生器 106 被耦合且被配置成產生(及/或通過級 107)元資料(包括 eSBR 元資料及 SBR 元資料)，該元資料將被級 107 包含在待被輸出自編碼器 100 的經編碼的位元流中。

編碼器 105 被耦合且被配置成編碼輸入音訊資料(例如，藉由對其執行壓縮)，並且將該產生的經編碼的音訊判斷提示至級 107，用於包含在待被輸出自級 107 的經編碼的位元流中。

級 107 被配置成將來自編碼器 105 的經編碼的音訊以及來自產生器 106 的元資料(包括 eSBR 元資料及 SBR 元資料)多工以產生待被輸出自級 107 的經編碼的位元流，較佳地使得該經編碼的位元流具有如本發明之其中一個實施例所指定的格式。

緩衝器記憶體 109 被配置成儲存(例如，以非暫態的方式)輸出自級 107 的經編碼的音訊位元流的至少一個區塊，且該經編碼的音訊位元流的一序列的區塊將接著被判斷提示自緩衝器記憶體 109 作為自編碼器 100 至傳遞系統的輸出。

圖 3 是包括解碼器(200)之系統的方塊圖，該解碼器為本發明之音訊處理單元的實施例，並且可選地亦有耦合

至其的後處理器(300)。解碼器 200 及後處理器 300 的任何組件或元件可被實現為硬體、軟體、或硬體與軟體之組合中的一或多個處理過程及/或一或多個電路(例如，ASICs、FPGAs、或其他積體電路)。解碼器 200 包含緩衝器記憶體 201、位元流負載去格式化器(剖析器)205、音訊解碼子系統 202(有時被稱為“核心”解碼級或“核心”解碼子系統)、eSBR 處理級 203、及控制位元產生級 204，連接如圖示。典型地，解碼器 200 亦包括其他處理元件(未示出)。

緩衝器記憶體(緩衝器)201 儲存(例如，以非暫態的方式)由解碼器 200 所接收的經編碼的 MPEG-4 AAC 音訊位元流的至少一個區塊。在解碼器 200 的操作中，位元流的一序列的區塊由緩衝器 201 被判斷提示至去格式化器 205。

在圖 3 實施例(或者將被描述的圖 4 實施例)的變型中，不是解碼器的 APU(例如，圖 6 的 APU 500)包括緩衝器記憶體(例如，等同於緩衝器 201 的緩衝器記憶體)，其儲存(例如，以非暫態的方式)由圖 3 或圖 4 的緩衝器 201 所接收之相同形式的經編碼的音訊位元流(例如，MPEG-4 AAC 音訊位元流)的至少一個區塊(即，包括 eSBR 元資料的經編碼的音訊位元流)。

再次參照圖 3，去格式化器 205 被耦合且被配置成將位元流的各個區塊解多工以從其抽取 SBR 元資料(包括經量化的包絡資料)及 eSBR 元資料(以及通常還包括其他元

資料)，用以至少將該 eSBR 元資料及該 SBR 元資料判斷提示至 eSBR 處理級 203，並且典型地亦將其他抽取出的元資料判斷提示至解碼子系統 202(以及可選地亦判斷提示至控制位元產生器 204)。去格式化器 205 亦被耦合且被配置成從位元流的各個區塊抽取音訊資料，並將該被抽取出的音訊資料判斷提示至解碼子系統(解碼級)202。

圖 3 的系統可選地亦包括後處理器 300。後處理器 300 包括緩衝器記憶體(緩衝器)301 以及其他處理元件(未示出)，其包括耦合至緩衝器 301 的至少一個處理元件。緩衝器 301 儲存(例如，以非暫態的方式)由後處理器 300 接收自解碼器 200 的經解碼的音訊資料地至少一個區塊(或框(frame))。後處理器 300 的處理元件被耦合且被配置成接收且適應性地處理輸出自緩衝器 301 的經解碼的音訊的一序列區塊(或框)，其使用自解碼子系統 202(及/或去格式化器 205)輸出的元資料及/或自解碼器 200 的級 204 輸出的控制位元。

解碼器 200 的音訊解碼子系統 202 被配置成解碼由剖析器 205 所抽取的音訊資料(此種解碼可被稱為“核心”解碼操作)以產生經解碼的音訊資料，並且判斷提示該經解碼的音訊資料至 eSBR 處理級 203。解碼係在頻域中執行，並且通常包括反量化其後接著頻譜處理。典型地，子系統 202 中的處理的最終級對經解碼的頻域音訊資料施用頻域至時域轉換，使得子系統的輸出為時域經解碼的資料。級 203 被配置成對經解碼的音訊資料施用由 SBR 元

資料及 eSBR 元資料(由剖析器 205 抽取)所指示的 SBR 工具及 eSBR 工具(即，使用 SBR 及 eSBR 元資料對解碼子系統 202 之輸出執行 SBR 及 eSBR 處理)，以產生經完全解碼的音訊資料，其自解碼器 200 輸出(例如，至後處理器 300)。典型地，解碼器 200 包括一記憶體(可由子系統 202 以及級 203 存取)，該記憶體儲存輸出自去格式化工器 205 的經去格式化的音訊資料及元資料，並且級 203 被配置成存取在 SBR 及 eSBR 處理期間所需要的音訊資料及元資料(包括 SBR 元資料及 eSBR 元資料)。級 203 中的 SBR 處理及 eSBR 處理可被視為對核心解碼子系統 202 之輸出的後處理。可選地，解碼器 200 亦包括一最終升混(upmixing)子系統(其可施用 MPEG-4 AAC 標準中所定義的參數化立體聲(“PS”)工具，使用由去格式化工器 205 所抽取的 PS 元資料及/或在子系統 204 中所產生的控制位元)，其被耦合且被配置成對級 203 之輸出執行升混，以產生經完全解碼、升混的音訊，其自解碼器 200 輸出。替代地，後處理器 300 被配置成對解碼器 200 之輸出執行升混(例如，使用由去格式化工器 205 所抽取的 PS 元資料及/或在子系統 204 中所產生的控制位元)。

回應於由去格式化工器 205 所抽取的元資料，控制位元產生器 204 可產生控制資料，且該控制資料可在解碼器 200 內(例如，在最終升混子系統中)被使用及/或被判斷提示作為解碼器 200 之輸出(例如，至後處理器 300，用於在後處理中使用)。回應於從輸出位元流所抽取的元資料(以

及可選地亦回應於控制資料)，級 204 可產生(及判斷提示至後處理器 300)控制位元，其指示從 eSBR 處理級 203 輸出的經解碼的音訊資料應進行特定類型的後處理。在一些實施方式中，解碼器 200 被配置成從輸入位元流將由去格式化器 205 所抽取的元資料判斷提示至後處理器 300，且後處理器 300 被配置成使用該元資料，對輸出自解碼器 200 的經解碼的音訊資料執行後處理。

圖 4 是音訊處理單元(“APU”)(210)的方塊圖，該音訊處理單元是本發明之音訊處理單元的另一實施例。APU 210 是傳統的解碼器，其並未被配置來執行 eSBR 處理。APU 210 的任何組件或元件可被實現為硬體、軟體、或硬體與軟體之組合中的一或多個處理過程及/或一或多個電路(例如，ASICs、FPGAs、或其他積體電路)。APU 210 包含緩衝器記憶體 201、位元流負載去格式化器(剖析器)215、音訊解碼子系統 202(有時被稱為“核心”解碼級或“核心”解碼子系統)、及 SBR 處理級 213，如圖所示連接。典型地，APU 210 亦包括其他處理元件(未示出)。

APU 210 的元件 201 及 202 與解碼器 200(圖 3 的)的相同編號的元件相同，並且它們的上述說明將不再重複。在 APU 210 的操作中，由 APU 210 所接收之經編碼的音訊位元流(MPEG-4 AAC 位元流)的一序列區塊係從緩衝器 201 被判斷提示至去格式化器 215。

去格式化器 215 被耦合且被配置成將位元流的各個區塊解多工以抽取 SBR 元資料(包括經量化的包絡資料)，以

及典型地亦從其抽取其他的元資料，但忽略根據本發明之其他實施例之可能包含在位元流中的 eSBR 元資料。去格式化器 215 被配置成將至少 SBR 元資料判斷提示至 SBR 處理級 213。去格式化器 215 亦被耦合且被配置成從位元流的各個區塊抽取音訊資料，並將該抽取的音訊資料判斷提示至解碼子系統(解碼級)202。

解碼器 200 的音訊解碼子系統 202 被配置成解碼由去格式化器 215 所抽取的音訊資料(此種解碼可被稱為“核心”解碼操作)，以產生經解碼的音訊資料，並且將該經解碼的音訊資料判斷提示至 SBR 處理級 213。該解碼係在頻域中執行。典型地，子系統 202 中的處理的最終級對經解碼的頻域音訊資料施用頻域至時域轉換，使得子系統的輸出為時域經解碼的資料。級 213 被配置成對經解碼的音訊資料施用由 SBR 元資料(由去格式化器 215 所抽取)所指示的 SBR 工具(但不施用 eSBR 工具)(即，使用 SBR 元資料對解碼子系統 202 之輸出執行 SBR 處理)，以產生經完全解碼的音訊資料，其自 APU 210 輸出(例如，至後處理器 300)。典型地，APU 210 包括一記憶體(可由子系統 202 以及級 213 存取)，該記憶體儲存輸出自去格式化器 215 的經去格式化的音訊資料及元資料，且級 213 被配置成存取在 SBR 處理期間所需要的音訊資料及元資料(包括 SBR 元資料)。級 213 中的 SBR 處理可被視為對核心解碼子系統 202 之輸出的後處理。可選地，APU 210 亦包括一最終升混子系統(其可施用在 MPEG-4 AAC 標準中所定義的參數

化立體聲(“PS”)工具，使用由去格式化器 215 所抽取的 PS 元資料)，其被耦合且被配置成對級 213 之輸出執行升混，以產生經完全解碼、升混的音訊，其自 APU 210 輸出。替代地，一後處理器被配置成對 APU 210 之輸出執行升混(例如，使用由去格式化器 215 所抽取的 PS 元資料及/或在 APU 210 中所產生的控制位元)。

編碼器 100、解碼器 200、及 APU 210 的各種實施方式係被配置成執行本發明方法的不同實施例。

依據某些實施例，經編碼的音訊位元流(例如，MPEG-4 AAC 位元流)中包含 eSBR 元資料(例如，包含係 eSBR 元資料的少量控制位元)，使得傳統的解碼器(其不被配置成剖析 eSBR 元資料，或不被配置成使用該 eSBR 元資料所屬的任何 eSBR 工具)可以忽略 eSBR 元資料，但仍然盡可能的不使用 eSBR 元資料或該 eSBR 元資料所屬的任何 eSBR 工具來解碼該位元流，通常在解碼音訊品質上無任何重大損失。然而，被配置成剖析位元流以識別 eSBR 元資料以及回應該 eSBR 元資料而使用至少一個 eSBR 工具的 eSBR 解碼器，將享受到使用至少一個這種 eSBR 工具的好處。因此，本發明之實施例提供一種用於以向後兼容的方式有效傳輸增強頻譜帶複製(eSBR)控制資料或元資料的機制。

典型地，位元流中的 eSBR 元資料表示下列 eSBR 工具(其描述於 MPEG USAC 標準中，且其在產生位元流的期間可能或可能不被編碼器所施用)之其中一或多者(例

如，表示其中一或多者之至少一個特徵或參數)：

- 諧波移調；
- QMF-修補)額外預處理(預平坦化)；及
- 子帶間樣本時間包絡成型或“inter-TES”。

例如，位元流中所包括的 eSBR 元資料可表示參數的值(描述於 MPEG USAC 標準中及本揭示內容中)：

harmonicSBR[ch]、sbrPatchingMode[ch]、
sbrOversamplingFlag[ch]、sbrPitchInBins[ch]、
sbrPitchInBins[ch]、bs_interTes、
bs_temp_shape[ch][env]、
bs_inter_temp_shape_mode[ch][env]、及
bs_sbr_preprocessing。

本文中，符號 $X[ch]$ ，其中 X 為某一參數，表示該參數屬於待解碼之經編碼的位元流的音訊內容的聲道(“ch”)。為了簡化，有時候省略[ch]的表述，並假定相關參數屬於音訊內容的聲道。

本文中，符號 $X[ch][env]$ ，其中 X 為某一參數，表示該參數屬於待解碼之經編碼的位元流的音訊內容的聲道(“ch”)的 SBR 包絡(“env”)。為了簡化，有時候省略[env]及[ch]的表述，並假定相關參數屬於音訊內容的聲道 SBR 包絡。

如所述，MPEG USAC 標準考慮到，USAC 位元流包括 eSBR 元資料，其控制由解碼器所執行之 eSBR 處理的效能。該 eSBR 元資料包括下列一位元的元資料參數：

harmonicSBR；bs_interTES；及 bs_pvc。

參數“harmonicSBR”表示針對 SBR 使用諧波修補(諧波移調)。具體地，harmonicSBR = 0 表示非諧波、頻譜修補，如 MPEG-4 AAC 標準第 4.6.18.6.3 節中所述；以及 harmonicSBR = 1 表示諧波 SBR 修補(具有 eSBR 中使用的形式，如 MPEG USAC 標準第 7.5.3 或 7.5.4 節中所述)。依據非 eSBR 頻譜帶複製(即，並非是 eSBR 的 SBR)，不使用諧波 SBR 修補。經由此揭示內容，頻譜修補被稱為頻譜帶複製的基本形式，而諧波移調被稱為頻譜帶複製的增強形式。

參數“bs_interTES”的值表示使用 eSBR 的 inter-TES 工具。

參數“bs_pvc”的值表示使用 eSBR 的 PVC 工具。

在將經編碼的位元流解碼的期間，在(針對該位元流所指示的音訊內容的各個聲道“ch”)解碼的 eSBR 處理級期間的諧波移調的效能係由下列 eSBR 元資料參數所控制：
sbrPatchingMode[ch]；sbrOversamplingFlag[ch]；
sbrPitchInBinsFlag[ch]；及 sbrPitchInBins[ch]。

值“sbrPatchingMode[ch]”表示 eSBR 中所使用的移調器(transposer)類型：sbrPatchingMode[ch] = 1 表示非諧波修補，如 MPEG-4 AAC 標準第 4.6.18.6.3 節中所述；sbrPatchingMode[ch] = 0 表示諧波 SBR 修補，如 MPEG USAC 標準第 7.5.3 或 7.5.4 節中所述。

值“sbrOversamplingFlag[ch]”表示在 eSBR 中使用訊

號自適應頻域超取樣，結合基於 DFT 的諧波 SBR 修補，如 MPEG USAC 標準第 7.5.3 節中所述。此旗標控制在移調器中所使用的 DFT 的大小：1 表示允許訊號自適應頻域超取樣，如 MPEG USAC 標準第 7.5.3.1 節中所述；0 表示禁止訊號自適應頻域超取樣，如 MPEG USAC 標準第 7.5.3.1 節中所述。

值“sbrPitchInBinsFlag[ch]”控制 sbrPitchInBins[ch]參數的解譯：1 表示 sbrPitchInBins[ch]中的值係有效的(valid)且大於零；0 表示 sbrPitchInBins[ch]的值被設定為零。

值“sbrPitchInBins[ch]”控制在 SBR 諧波移調器中，交叉乘積(cross product)項的增加。值 sbrPitchInBins[ch]為在範圍[0,127]中的整數值，並且表示作用於核心編碼器之採樣頻率的 1536 線 DFT 的頻槽(frequency bins)中所測量的距離。

在 MPEG-4 AAC 位元流指示其聲道未耦合的 SBR 雙聲道(而不是單一 SBR 聲道)的情形中，該位元流指示上述語法的兩個實例(用於諧波或非諧波移調)，一個實例用於 sbr_channel_pair_element()的一個聲道。

eSBR 工具的諧波移調通常改善了在相對低交越頻率的經解碼的音樂訊號的品質。諧波移調應在解碼器中經由基於 DFT 或基於 QMF 的諧波移調而被實施。非諧波移調(即，傳統的頻譜修補或複製)通常改善了語音訊號。因此，決定哪種類型的移調較佳用於編碼特定音訊內容的起

始點，係依據語音/音樂偵測來選擇移調方法，諧波移調用於音樂內容，而頻譜修補用於語音內容。

eSBR 處理期間的預平坦化的效能係由被稱為“bs_sbr_preprocessing”的一位元的 eSBR 元資料參數的值所控制，這個意思是依據此單一位元值來執行或不執行預平坦化。當使用 SBR QMF 修補演算法(如 MPEG-4 AAC 標準第 4.6.18.6.3 節中所述)時，可執行預平坦化的步驟(當由“bs_sbr_preprocessing”參數指示時)，努力避免被輸入至後續包絡調整器(該包絡調整器執行該 eSBR 處理的其他級)的高頻訊號的頻譜包絡的形狀中的不連續。預平坦化通常改善了後續包絡調整器級的操作，產生被視為是更穩定的高頻帶訊號。

在解碼器中的 eSBR 處理期間，子帶間樣本時間包絡成型(“inter-TES”工具)的效能係由下列針對將被解碼的 USAC 位元流的音訊內容的各個聲道(“ch”)的 SBR 包絡(“env”)的 eSBR 元資料參數所控制：

bs_temp_shape[ch][env]；及

bs_inter_temp_shape_mode[ch][env]。

inter-TES 工具處理在包絡調整器之後的 QMF 子帶樣本。此處理步驟利用比包絡調整器更精細的時間粒度來將較高頻帶的時間包絡整型。藉由將一增益因子施加至 SBR 包絡中的各個 QMF 子帶樣本，inter-TES 將 QMF 子帶樣本之間的時間包絡整型。

參數“bs_temp_shape[ch][env]”為一旗標，其發出使用

inter-TES 的訊號。參數

“bs_inter_temp_shape_mode[ch][env]”表示(如 MPEG USAC 標準中所定義)inter-TES 中參數 γ 的值。

依據本發明的某些實施例，針對包括在 MPEG-4 AAC 位元流中的整體位元率要求，表示上述 eSBR 工具(諧波移調、預平坦化、及 inter-TES)的 eSBR 元資料被預期是每秒幾百個位元的量級，因為只有執行 eSBR 處理所需的差分控制資料被傳輸。傳統的解碼器可忽略此資訊，因為它以向後兼容的方式被包括(將於稍後說明)。因此，由於某些原因，對與包含 eSBR 元資料相關的位元率的不利影響是可以忽略的，該些原因包括下列：

- 位元率損失(bitrate penalty)(由於包含該 eSBR 元資料所造成)是總位元率的一非常小的部分，因為只有執行 eSBR 處理所需要的差分控制資料被傳輸(而不是 SBR 控制資料的聯播)；

- SBR 相關控制資訊的調整(tuning)通常不依賴移調(transposition)的細節；以及

- inter-TES 工具(在 eSBR 處理期間採用)執行經移調的訊號的單端(single ended)後處理。

因此，本發明的實施例提供了以向後兼容的方式高效傳輸增強頻譜帶複製(eSBR)控制資料或元資料的機制。此種 eSBR 控制資料的高效傳輸降低了採用本發明之態樣的解碼器、編碼器、及轉碼器中的記憶體需求，同時對於位元率沒有明顯的不利影響。此外，亦降低了與依據本發明

之實施例執行 eSBR 相關連的複雜度和處理要求，因為 SBR 資料只需要被處理一次，而不是聯播，這可以是若 eSBR 被當成是 MPEG-4 AAC 中一完全獨立的物件，而不是以向後兼容的方式被集成到 MPEG-4 AAC 編碼解器中的情形。

接著，參考圖 7，將說明依據本發明之某些實施例的 MPEG-4 AAC 位元流之區塊(“raw_data_block”)的元素，該 MPEG-4 AAC 位元流中包括 eSBR 元資料。圖 7 為 MPEG-4 AAC 位元流之一區塊(“raw_data_block”)的示圖，顯示其之一些區段。

MPEG-4 AAC 位元流之一區塊可包括至少一個“single_channel_element()”(例如，圖 7 中所示之單聲道元素)，及/或至少一個“channel_pair_element()”(儘管其可能存在，但在圖 7 中未明確示出)，其包括用於音訊節目之音訊資料。該區塊亦可包括一些“fill_elements”(例如，圖 7 的填充元素 1 及/或填充元素 2)，其包括關於該節目的資料(例如，元資料)。各個“single_channel_element()”包括一識別符(例如，圖 7 的“ID1”)，其指示單聲道元素的起始，並可包括指示多聲道音訊節目之一不同聲道的音訊資料。各個“channel_pair_element”包括一識別符(圖 7 中未示出)，其指示雙聲道元素的起始，並可包括指示該節目之兩個聲道的音訊資料。

MPEG-4 AAC 位元流之一 fill_element(在本文中稱為填充元素)包括一識別符(圖 7 的“ID2”)，其指示填充元素

的起始，且填充資料在該識別符之後。識別符 ID2 可由一三位元最高有效位元傳輸在先之無正負號整數(“uimsbf”)組成，其具有 0x6 的值。填充資料可包括一 `extension_payload()` 元素(在本文中有時被稱為擴充負載)，其語法示於 MPEG-4 AAC 標準之表 4.57 中。存在數種擴充負載的類型，且透過“`extension_type`”參數而被識別，該參數為一四位元最高有效位元傳輸在先之無正負號整數(“uimsbf”)。

填充資料(例如，其之擴充負載)可包括標頭或識別符(例如，圖 7 的“標頭 1”)，其指示表示 SBR 物件之填充資料的區段(即，該標頭初始化一“SBR 物件”類型，在 MPEG-4 AAC 標準中稱為 `sbr_extension_data()`)。例如，頻譜帶複製(SBR)擴充負載被標示為值‘1101’或‘1110’，用於在標頭中的 `extension_type` 欄位，其中識別符‘1101’識別具有 SBR 資料的擴充負載，而‘1110’識別具有 SBR 資料的擴充負載使用循環冗餘檢測(CRC)以驗證該 SBR 資料之正確性。

當標頭(例如，`extension_type` 欄位)初始化一 SBR 物件類型時，SBR 元資料(在本文中有時被稱為“頻譜帶複製資料”，且在 MPEG-4 AAC 標準中被稱為 `sbr_data()`)跟在該標頭之後，且至少一個頻譜帶複製擴充元素(例如，圖 7 之填充元素 1 的“SBR 擴充元素”)可跟在該 SBR 元資料之後。此一頻譜帶複製擴充元素(該位元流之一區段)在 MPEG-4 AAC 標準中被稱為“`sbr_extension()`”容器。頻譜

帶複製擴充元素可選地包括一標頭(例如，圖 7 之填充元素 1 的“SBR 擴充標頭”)。

MPEG-4 AAC 標準考慮到，一頻譜帶複製擴充元素可包括用於一節目的音訊資料的 PS(參數化立體聲)資料。MPEG-4 AAC 標準考慮到，當填充元素的標頭(例如，其之擴充負載)初始化一 SBR 物件類型(如圖 7 之“標頭 1”一樣)且該填充元素的頻譜帶複製擴充元素包括 PS 資料時，該填充元素(例如，其之擴充負載)包括頻譜帶複製資料，以及“bs_extension_id”參數，該參數值(即，bs_extension_id = 2)指示 PS 資料係包含在該填充元素的頻譜帶複製擴充元素中。

依據本發明之一些實施例，eSBR 元資料(例如，指示是否對該區塊的音訊內容執行增強頻譜帶複製(eSBR)處理的旗標)係包含在填充元素的頻譜帶複製擴充元素中。例如，圖 7 的填充元素 1 中指示此一旗標，其中該旗標出現在填充元素 1 的“SBR 擴充元素”的標頭(填充元素 1 的“SBR 擴充標頭”)之後。可選地，此一旗標及額外的 eSBR 元資料亦包括在頻譜帶複製擴充元素中，其在頻譜帶複製擴充元素的標頭之後(例如，在圖 7 中的填充元素的 SBR 擴充元素中，在該 SBR 擴充標頭之後)。依據本發明之一些實施例，包括 eSBR 元資料的填充元素亦包括“bs_extension_id”參數，該參數值(例如，bs_extension_id = 3)指示 eSBR 元資料係包含在該填充元素中，並指示將對該相關區塊的音訊內容執行 eSBR 處理。

依據本發明之一些實施例，eSBR元資料係包含在MPEG-4 AAC位元流的填充元素(例如，圖7的填充元素2)中，而不是在該填充元素的頻譜帶複製擴充元素(SBR擴充元素)中。這是因為包含具有SBR資料或具有CRC之SBR資料的extension_payload()的填充元素並不包含任何其他擴充類型的任何其他擴充負載。因此，在eSBR元資料係保存其自己的擴充負載的實施例中，使用一單獨的填充元素來儲存該eSBR元資料。此一填充元素包括一識別符(例如，圖7的“ID2”)，其指示填充元素的起始，且填充資料在該識別符之後。該填充資料包括一extension_payload()元素(在本文中有時被稱為擴充負載)，其語法顯示在MPEG-4 AAC標準的表4.57中。該填充資料(例如，其之擴充負載)包括一標頭(例如，圖7之填充元素2的“標頭2”)，其表示一eSBR物件(即，該標頭初始化一增強頻譜帶複製(eSBR)物件類型)，且該填充資料(例如，其之擴充負載)包括eSBR元資料在該標頭之後。例如，圖7的填充元素2包括此一標頭(“標頭2”)且亦包括在該標頭之後的eSBR元資料(即，在填充元素2中的“旗標”，其表示是否對該區塊的音訊內容執行增強頻譜帶複製(eSBR)處理。可選地，在標頭2之後，額外的eSBR元資料亦包含在圖7之填充元素2的填充資料中。在本段中所描述的實施例中，該標頭(例如，圖7的標頭2)具有一識別值，該識別值不是MPEG-4 AAC標準之表4.57中所定義的常規值之其中一者，反而是表示一eSBR擴充負載(使得該標頭的extension_type欄位指示該填充資料包括eSBR元資料)。

在第一類的實施例中，本發明為一音訊處理單元(例如，解碼器)，包含：

記憶體(例如，圖 3 或 4 的緩衝器 201)，被配置成儲存經編碼的音訊位元流的至少一個區塊(例如，MPEG-4 AAC 位元流的至少一個區塊)；

位元流負載去格式化器(例如，圖 3 的元件 205 或圖 4 的元件 215)，被耦合至該記憶體，且被配置成將該位元流的該區塊的至少一部分解多工；以及

解碼子系統(例如，圖 3 的元件 202 及 203、或圖 4 的元件 202 及 213)，被耦合且被配置成將該位元流之該區塊的音訊內容的至少一部分解碼，其中該區塊包括：

填充元素，其包括指示該填充元素之起始的識別符(例如，“id_syn_ele”識別符具有MPEG-4 AAC標準之表4.85的值0x6)，且填充資料在該識別符之後，其中該填充資料包括：

至少一旗標，識別是否對區塊的音訊內容執行增強形式頻譜帶複製處理(eSBR)(例如，使用該區塊中所包含的頻譜帶複製資料及eSBR元資料)。

該旗標為 eSBR 元資料，且該旗標的一範例為 sbrPatchingMode 旗標。該旗標的另一範例為 harmonicSBR 旗標。這兩個旗標皆指示是否對該區塊的音訊資料執行頻譜帶複製的基本形式或是頻譜複製的增強形式。頻譜複製的基本形式是頻譜修補，而頻譜複製的增強形式為諧波移調。

在某些實施例中，該填充資料亦包括額外的 eSBR 元

資料(即，除了該旗標之外的 eSBR 元資料)。

該記憶體可以是緩衝器記憶體(例如，圖 4 之緩衝器 201 的實施方式)，其儲存(例如，以非暫態的方式)該經編碼的音訊位元流的至少一個區塊。

據估計，在包括 eSBR 元資料(表示這些 eSBR 工具)的 MPEG-4 AAC 位元流的解碼期間，由 eSBR 解碼器所執行的 eSBR 處理(使用 eSBR 諧波移調、預平坦化、及 inter_TES 工具)的效能的複雜度可係如下(用於利用指示的參數的典型解碼)：

- 諧波移調(16 kbps, 14400/28800 Hz)
 - 基於 DFT: 3.68 WMOPS(每秒加權百萬次操作數)；
 - 基於 QMF: 0.98 WMOPS；
- QMF 修補預處理(預平坦化): 0.1 WMOPS；及
- 子帶間樣本時間包絡成型(inter-TES): 最多 0.16 WMOPS。

已知的是，針對瞬變(transients)，基於 DFT 的置換通常比基於 QMF 的置換執行得更好。

依據本發明之一些實施例，包含 eSBR 元資料的(經編碼的音訊位元流的)填充元素亦包含一參數(例如，“bs_extension_id”參數)，該參數值(例如，bs_extension_id = 3)發出 eSBR 元資料係包含在填充元素中的信號以及發出將對相關區塊的音訊內容執行 eSBR 處理的信號，及/或一參數(例如，相同的“bs_extension_id”參數)，該參數值(例如，bs_extension_id = 2)發出該填充元素之 sbr_extension()容器包括 PS 資料的信號。例如，如下面表 1

中所示，此種具有值 `bs_extension_id = 2` 的參數可發出該填充元素之 `sbr_extension()` 容器包括 PS 資料的訊號，且此種具有值 `bs_extension_id = 3` 的參數可發出該填充元素之 `sbr_extension()` 容器包括 eSBR 元資料的訊號：

表 1

bs_extension_id	含義	
0	保留(Reserved)	
1	保留(Reserved)	
2	EXTENSION_ID_PS	
3	EXTENSION_ID_ESBR	

依據本發明之一些實施例，包括 eSBR 元資料及/或 PS 資料之各個頻譜帶複製擴充元素的語法係如下面表 2 中所示(其中“`sbr_extension()`”表示一容器，該容器為頻譜帶複製擴充元素，“`bs_extension_id`”係如上面表 1 中所述，“`ps_data`”表示 PS 資料，以及“`esbr_data`”表示 eSBR 元資料)：

表 2

sbr_extension(bs_extension_id, num_bits_left)		
{		
switch (bs_extension_id) {		
case EXTENSION_ID_PS:		
num_bits_left -= ps_data();		註 1
break;		
case EXTENSION_ID_ESBR:		
num_bits_left -= esbr_data();		註 2
break;		
default:		
bs_fill_bits;	註 3	
num_bits_left = 0;		
break;		
}		
}		
註1: ps_data()返回讀取的位元數。		
註2: esbr_data()返回讀取的位元數。		
註3: 參數bs_fill_bits包含N個位元，其中 N = num_bits_left。		

在一示例性實施例中，在上面表 2 所提及的 esbr_data()指示以下元資料參數的值：

- 1. 上述一位元的元資料參數 “harmonicSBR”；“bs_interTES”；及“bs_sbr_preprocessing”之各者；
- 2. 針對待解碼之經編碼的位元流的音訊內容的各個聲道(“ch”)，上述參數之各者：“sbrPatchingMode[ch]”；“sbrOversamplingFlag[ch]”；“sbrPitchInBinsFlag[ch]”；及“sbrPitchInBins[ch]”；以及

3. 針對待解碼之經編碼的位元流的音訊內容的各個聲道(“ch”)的各個 SBR 包絡(“env”), 上述參數之各者: “bs_temp_shape[ch][env]”; 及 “bs_inter_temp_shape_mode[ch][env]”。

例如, 在某些實施例中, esbr_data()可具有表 3 中所示的語法, 以指示這些元資料參數:

表 3

esbr_data()		
{		
harmonicSBR;	1	
bs_interTes;	1	
bs_sbr_preprocessing;	1	
if (harmonicSBR) {		
if (sbrPatchingMode[0] == 0) {	1	
sbrOversamplingFlag[0];	1	
if (sbrPitchInBinsFlag[0])	1	
sbrPitchInBins[0];	7	
Else		
sbrPitchInBins[0] = 0;		
} else {		
sbrOversamplingFlag[0] = 0;		
sbrPitchInBins[0] = 0;		
}		
}		
if (bs_interTes) {		
/* 執行ch和env 的迴圈*/		
bs_temp_shape[ch][env];	1	
if (bs_temp_shape[ch][env]) {		
bs_inter_temp_shape_mode[ch][env];	2	
}		
}		
}		

在表 3 中，在中間行的數字表示在左邊行之對應參數的位元數。

在某些實施例中，本發明為一種方法，包括將音訊資料編碼以產生經編碼的位元流(例如，MPEG-4 AAC 位元流)的步驟，該步驟包括藉由將 eSBR 元資料包括在該經編碼的位元流的至少一個區塊的至少一個區段中，以及將音訊資料包括在該區塊的至少一個其他區段中。在典型的實施例中，該方法包括在該經編碼的位元流的各個區塊中將該音訊資料與該 eSBR 元資料多工的步驟。在 eSBR 解碼器中的經編碼的位元流的典型的解碼中，解碼器從該位元流抽取 eSBR 元資料(包括藉由剖析及解多工 eSBR 元資料及音訊資料)，並使用該 eSBR 元資料來處理該音訊資料，以產生經解碼的音訊資料的串流。

本發明的另一態樣為 eSBR 解碼器，其被配置成，在不包括 eSBR 元資料之經編碼的音訊位元流(例如，MPEG-4 AAC 位元流)的解碼期間，執行 eSBR 處理(例如，使用被稱為諧波移調、預平坦化、或 inter_TES 之 eSBR 工具的其中至少一者)。將參照圖 5 來描述此種解碼器的一範例。

圖 5 之 eSBR 解碼器(400)包括緩衝器記憶體 201(其等同於圖 3 及 4 的記憶體 201)、位元流負載去格式化器 215(其等同於圖 4 的去格式化器 215)、音訊解碼子系統 202(有時被稱為“核心”解碼級或“核心”解碼子系統，且其等同於圖 3 的核心解碼子系統 202)、eSBR 控制資料產生子系統 401、及 eSBR 處理級 203(其等同於圖 3 的級 203)，連接如圖示。典型地，解碼器 400 亦包括其他處理

元件(未示出)。

在解碼器 400 的操作中，由解碼器 400 所接收之經編碼的音訊位元流(MPEG-4 AAC 位元流)的一序列區塊係從緩衝器 201 被判斷提示至去格式化器 215。

去格式化器 215 被耦合且被配置成將位元流的各個區塊解多工以抽取 SBR 元資料(包括經量化的包絡資料)，以及通常亦從其抽取其他的元資料。去格式化器 215 被配置成將至少該 SBR 元資料判斷提示至 eSBR 處理級 203。去格式化器 215 亦被耦合且被配置成從該位元流的各個區塊抽取音訊資料，並將該抽取的音訊資料判斷提示至解碼子系統(解碼級)202。

解碼器 400 的音訊解碼子系統 202 被配置成解碼由去格式化器 215 所抽取的音訊資料(此種解碼可被稱為“核心”解碼操作)以產生經解碼的音訊資料，並且將該經解碼的音訊資料判斷提示至 eSBR 處理級 203。該解碼係在頻域中執行。典型地，子系統 202 中的處理的最終級對經解碼的頻域音訊資料施用頻域至時域轉換，使得子系統之輸出為時域經解碼的音訊資料。級 203 被配置成對經解碼的音訊資料施用由 SBR 元資料(由去格式化器 215 所抽取)以及子系統 401 中產生的 eSBR 元資料所指示的 SBR 工具(及 eSBR 工具)(即，使用 SBR 及 eSBR 元資料對解碼子系統 202 之輸出執行 SBR 及 eSBR 處理)，以產生經完全解碼的音訊資料，其自解碼器 400 輸出。典型地，解碼器 400 包括一記憶體(可由子系統 202 以及級 203 存取)，該

記憶體儲存自去格式化器 215(及可選地亦自子系統 401)輸出的經去格式化的音訊資料及元資料，並且級 203 被配置成存取在 SBR 及 eSBR 處理期間所需要的音訊資料及元資料。級 203 中的 SBR 處理可被視為對解碼子系統 202 之輸出的後處理。可選地，解碼器 400 亦包括一最終升混子系統(其可施用在 MPEG-4 AAC 標準中所定義的參數化立體聲(“PS”)工具，使用由去格式化器 215 所抽取的 PS 元資料)，其被耦合且被配置成對級 203 之輸出執行升混，以產生經完全解碼、升混的音訊，其自 APU 210 輸出。

圖 5 的控制資料產生子系統 401 被耦合且被配置成偵測待解碼之經編碼的音訊位元流的至少一個屬性，並回應該偵測步驟的至少一個結果來產生 eSBR 控制資料(依據本發明之其他實施例，其可以是或可包括經編碼的音訊位元流中所包含的任何類型的 eSBR 元資料)。該 eSBR 控制資料被判斷提示至級 203，用以當偵測到該位元流之一特定屬性(或屬性的組合)時觸發個別 eSBR 工具或 eSBR 工具的組合的應用，及/或用以控制此 eSBR 工具的應用。例如，為了控制使用諧波移調之 eSBR 處理的效能，控制資料產生子系統 401 的某些實施例可包括：音樂偵測器(例如，傳統音樂偵測器的簡易版本)，用於回應偵測到該位元流表示或非表示音樂而設定 `sbrPatchingMode[ch]` 參數(以及判斷提示該設定的參數至級 203)；瞬變偵測器，用於回應偵測到該位元流所指示的音訊內容中存在或不存在

瞬變而設定 `sbrOversamplingFlag[ch]` 參數 (以及判斷提示該設定的參數至級 203)；及/或音高 (pitch) 偵測器，用於回應偵測到該位元流所指示的音訊內容的音高而設定 `sbrPitchInBinsFlag[ch]` 及 `sbrPitchInBins[ch]` 參數 (以及判斷提示該設定的參數至級 203)。本發明的其他態樣為由本段落以及前一段落中所述之本發明解碼器的任何實施例所執行的音訊位元流解碼方法。

本發明的態樣包括編碼或解碼方法，具有本發明 APU、系統或裝置之任何實施例被配置 (例如，被編程) 以執行的類型。本發明的其他態樣包括系統或裝置，其被配置 (例如，被編程) 以執行本發明方法的任何實施例，以及電腦可讀取媒體 (例如，光碟)，其儲存程式碼 (例如，以非暫態的方式) 用於執行本發明方法或其步驟的任何實施例。例如，本發明系統可以是或可包括可編程通用處理器、數位訊號處理器、或微處理器，其以軟體或韌體編程及/或另外被配置以對資料執行任何的多種操作，其包括本發明方法或其步驟的實施例。此種通用處理器可以是或可包括電腦系統，其包括輸入裝置、記憶體、及處理電路，被編程 (及/或另外被配置) 以回應被判斷提示至其的資料而執行本發明方法 (或其步驟) 的實施例。

本發明之實施例可在硬體、韌體、或軟體、或兩者之組合 (例如，可編程邏輯陣列) 中實現。除非另有規定，否則被包括作為本發明之一部分的演算法或處理並非固有地與任何特定電腦或其他裝置相關。尤其是，各種通用機器

可以與依據本文之教示所編寫的程式碼一起使用，或者可以更方便的建構更專用的設備(例如，積體電路)來執行所需的方法步驟。因此，可在一或多個電腦程式中實施本發明，該一或多個電腦程式執行在一或多個可編程的電腦系統上(例如，圖 1 的任何元件、或圖 2 的編碼器 100(或其元件)、或圖 3 的解碼器 200(或其元件)、或圖 4 的解碼器 210(或其元件)、或圖 5 的解碼器 400(或其元件)的實施方式)，該一或多個可編程的電腦系統各包含至少一個處理器、至少一個資料儲存系統(包括揮發性或非揮發性記憶體及/或儲存元件)、至少一個輸入裝置或埠、及至少一個輸出裝置或埠。程式碼被應用到輸入資料，用以執行本文所述之功能，並產生輸出資訊。該輸出資訊以已知的方式被應用至一或多個輸出裝置。

每個此種程式可以以任何期望的電腦語言(包括機器語言、組合語言、或高階程序語言、邏輯語言、或物件導向程式語言)來實施，用以與電腦系統通訊。在任何情況下，該語言可以是編譯語言或是解釋語言。

例如，當由電腦軟體指令序列來實施時，本發明之實施例的各種功能及步驟可以由在適當的數位訊號處理硬體中運行的多緒軟體指令序列來實施，在此情況下，實施例的各種裝置、步驟、及功能可對應於軟體指令的部分。

每個此種電腦程式較佳地被儲存在或下載至通用或專用可編程的電腦可讀的儲存媒體或裝置(例如，固態記憶體或媒體、或磁或光學媒體)，用於當該儲存媒體或裝置

由該電腦系統讀取以執行本文所述的程序時，配置及操作該電腦。本發明系統亦可實施作為電腦可讀取儲存媒體，其配置有(即，儲存)電腦程式，其中如此配置的儲存媒體使得電腦系統以特定且預定的方式操作以執行本文所述的功能。

已描述了本發明之多個實施例。然而，將被理解的是，可在不悖離本發明之精神和範圍的前提下作出各種修改。按照上述教示，本發明的許多修改和變型是可能的。應當理解的是，在所附申請專利範圍的範圍內，可以有別於本文所具體描述之方式實施本發明。包含在以下申請專利範圍中的任何標號僅用於說明的目的，不應當用於以任何方式解釋或限制申請專利範圍。

【符號說明】

1:編碼器

2:傳遞子系統

3:解碼器

4:後處理單元

100:編碼器

105:編碼器

106:元資料產生器級

107:填充器/格式化器級

109:緩衝器記憶體

200:解碼器

201:緩衝器記憶體
202:解碼子系統
203:eSBR 處理級
204:控制位元產生器級
205:位元流負載去格式化器(剖析器)
210:音訊處理單元(APU)
213:SBR 處理級
215:位元流負載去格式化器(剖析器)
300:後處理器
301:緩衝器記憶體(緩衝器)
400:eSBR 解碼器
401:eSBR 控制資料產生子系統
500:音訊處理單元(APU)

【發明申請專利範圍】

【請求項 1】一種音訊處理單元，包含：

緩衝器，被配置成儲存經編碼的音訊位元流的至少一個區塊；

位元流負載去格式化器，耦合至該緩衝器，且被配置成將該經編碼的音訊位元流的該至少一個區塊的至少一部分解多工；以及

解碼子系統，耦合至該位元流負載去格式化器，且被配置成將該經編碼的音訊位元流的該至少一個區塊的至少一部分解碼，其中該經編碼的音訊位元流的該至少一個區塊包括：

填充元素，具有指示該填充元素之起始的識別符，以及在該識別符之後的填充資料，其中該填充資料包括：

至少一旗標，識別是否對該經編碼的音訊位元流的該至少一個區塊的音訊內容執行頻譜帶複製的基本形式或頻譜帶複製的增強形式，其中該頻譜帶複製的基本形式包括頻譜修補，該頻譜帶複製的增強形式包括諧波移調，該旗標的一個值指示應對該音訊內容執行所述頻譜帶複製的增強形式，而該旗標的另一個值指示應對該音訊內容執行所述頻譜帶複製的基本形式而非所述諧波移調。

【請求項 2】如請求項 1 之音訊處理單元，其中該填充資料還包括增強頻譜帶複製元資料。

【請求項 3】如請求項 2 之音訊處理單元，其中該增

強頻譜帶複製元資料被包含在填充元素的擴充負載中。

【請求項 4】如請求項 2 之音訊處理單元，其中該增強頻譜帶複製元資料包括定義主頻帶表的一或多個參數。

【請求項 5】如請求項 2 之音訊處理單元，其中該增強頻譜帶複製元資料包括包絡比例因數或雜訊底限比例因數。

【請求項 6】一種用於將經編碼的音訊位元流解碼之方法，該方法包含：

接收經編碼的音訊位元流的至少一個區塊；

將該經編碼的音訊位元流的該至少一個區塊的至少一部分解多工；以及

將該經編碼的音訊位元流的該至少一個區塊的至少一部分解碼，其中該經編碼的音訊位元流的該至少一個區塊包括：

填充元素，具有指示該填充元素之起始的識別符，以及在該識別符之後的填充資料，其中該填充資料包括：

至少一旗標，識別是否對該經編碼的音訊位元流的該至少一個區塊的音訊內容執行頻譜帶複製的基本形式或頻譜帶複製的增強形式，其中該頻譜帶複製的基本形式包括頻譜修補，該頻譜帶複製的增強形式包括諧波移調，該旗標的一個值指示應對該音訊內容執行所述頻譜帶複製的增強形式，而該旗標的另一個值指示應對該音訊內容執行所述頻譜帶複製的基本形式而非所述諧波移調。

【請求項 7】如請求項 6 之方法，其中該識別符是最高有效位元傳輸的三位元無正負號整數並且具有 0x6 之值。

【請求項 8】如請求項 6 之方法，其中該填充資料還包括增強頻譜帶複製元資料。

【發明圖式】

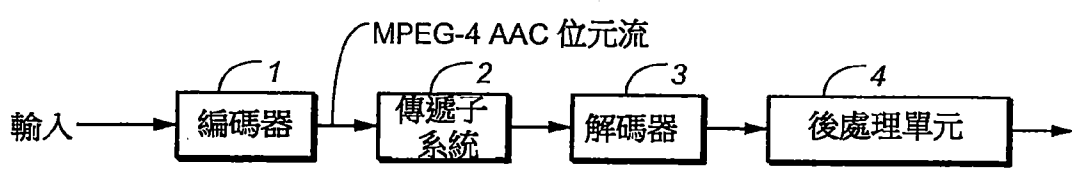


圖 1

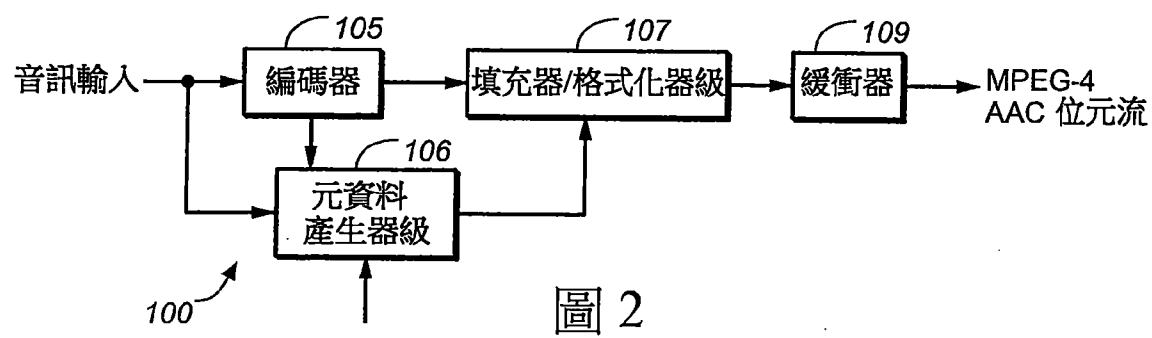


圖 2

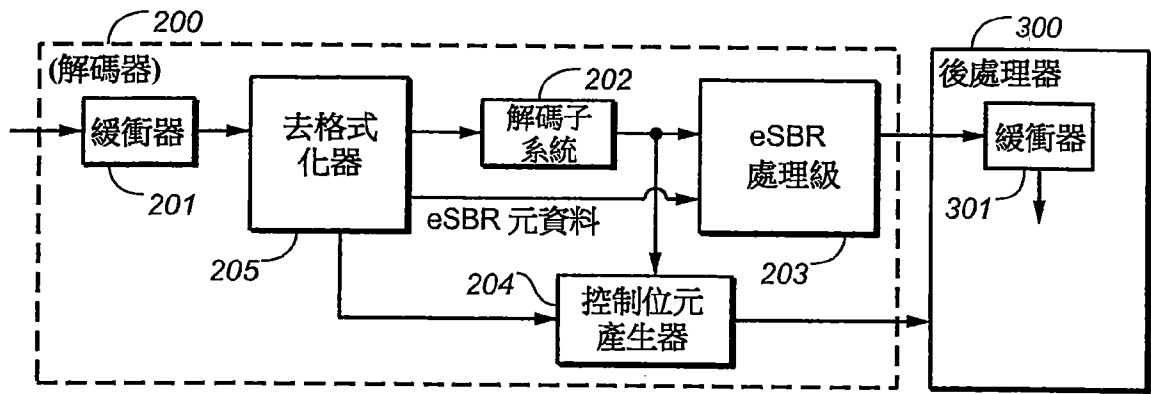


圖 3

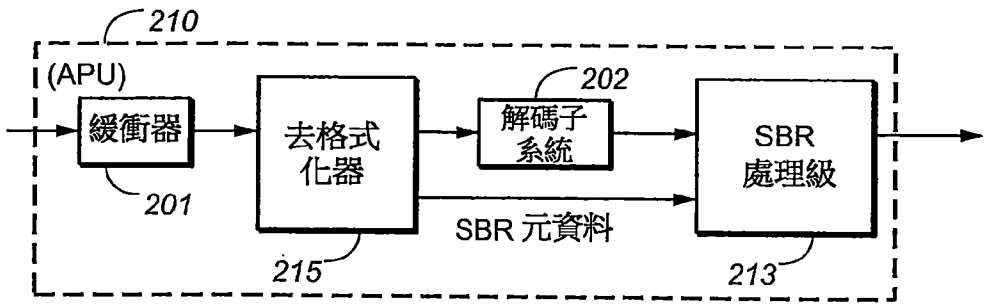


圖 4

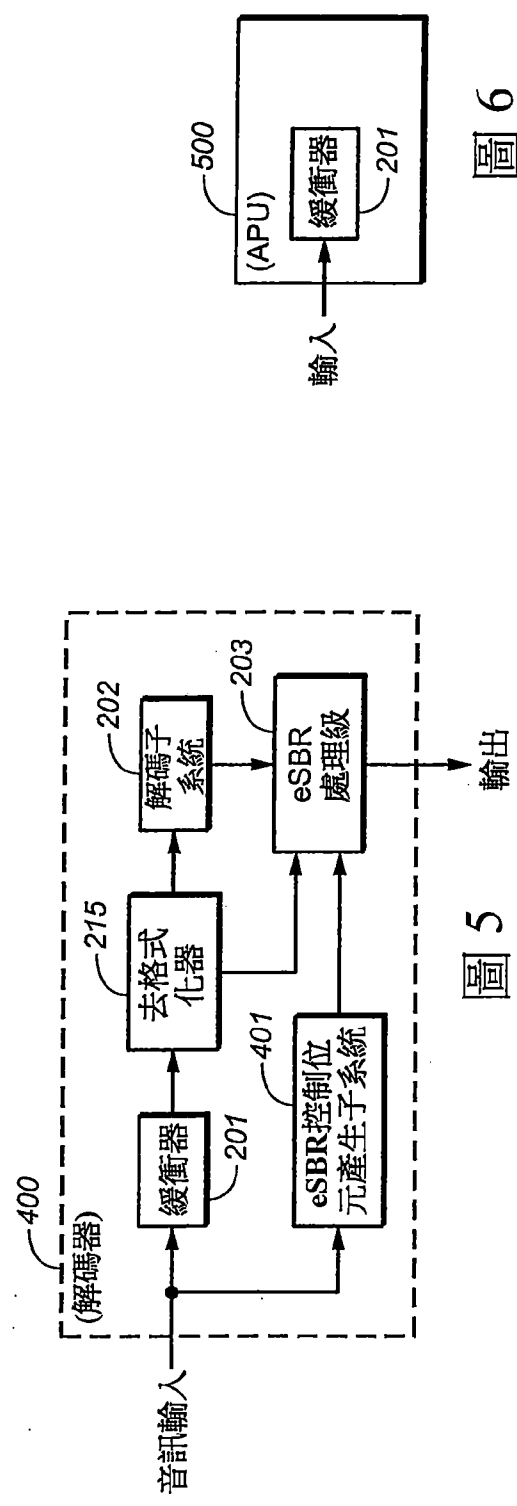


圖 5 輸出

圖 6

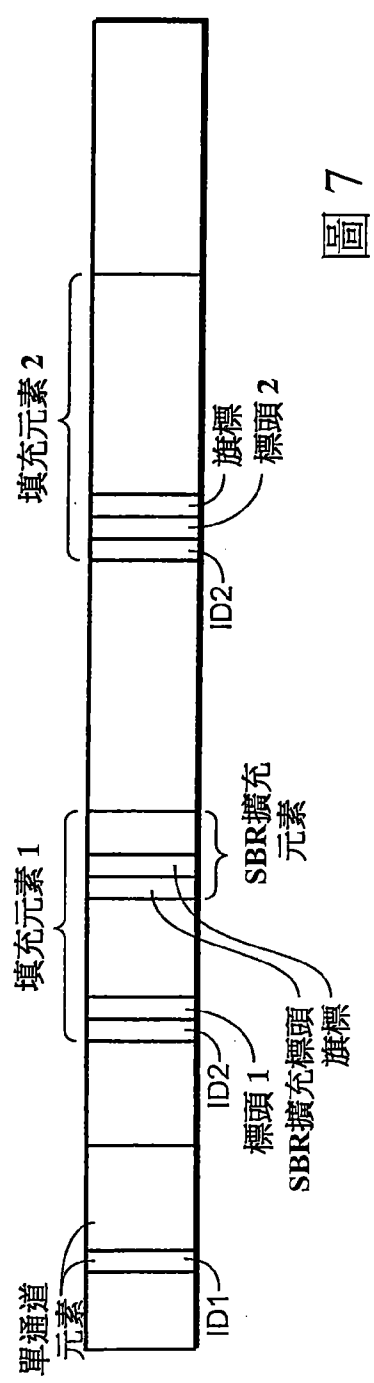


圖 7