



(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 I446338 B

(45) 公告日：中華民國 103 (2014) 年 07 月 21 日

(21) 申請案號：100123209

(22) 申請日：中華民國 100 (2011) 年 06 月 30 日

(51) Int. Cl. : G10L19/02 (2013.01)

(30) 優先權：2010/07/01 美國

12/829,233

(71) 申請人：寶力康公司 (美國) POLYCOM, INC. (US)

美國

(72) 發明人：方金偉 FENG, JINWEI (US)；朱 彼得 CHU, PETER (US)

(74) 代理人：陳長文

(56) 參考文獻：

US 2008/0052068A1

Vinton, A SCALABLE AND PROGRESSIVE AUDIO CODEC, pp. 3277-3280,
0-7803-7041-4/01 (C) 2001 IEEE. Department of Electrical
Engineering, University of Washington. Box 352500. Seattle. WA
98195-2500, USA.

審查人員：涂淑惠

申請專利範圍項數：46 項 圖式數：16 共 0 頁

(54) 名稱

可擴縮音訊處理方法及裝置

SCALABLE AUDIO PROCESSING METHOD AND DEVICE

(57) 摘要

一種用於一處理裝置之可擴縮音訊編解碼器判定每一輸入音訊訊框之第一位元分配及第二位元分配。將第一位元分配給一第一頻率頻帶，且將第二位元分配給一第二頻率頻帶。該等分配係基於該兩個頻帶之間的能量比率在一逐個訊框基礎上進行。針對每一訊框，該編解碼器將兩個頻帶變換編碼成兩個變換係數集，然後基於該等位元分配封包化該兩個變換係數集。然後利用該處理裝置傳輸該等封包。另外，可依藉由功率位準及感知模型化所判定之重要性次序配置該等變換係數之頻率區。若發生位元剝除，假定已在該等頻帶之間分配位元且已按重要性排序該等變換係數之該等區，則在一接收裝置處之解碼器可產生適合品質之音訊。

A scalable audio codec for a processing device determines first and second bit allocations for each frame of input audio. First bits are allocated for a first frequency band, and second bits are allocated for a second frequency band. The allocations are made on a frame-by-frame basis based on the energy ratio between the two bands. For each frame, the codec transform codes both frequency bands into two sets of transform coefficients, which are then packetized based on the bit allocations. The packets are then transmitted with the processing device. Additionally, the frequency regions of the transform coefficients can be arranged in order of importance determined by power levels and perceptual modeling. Should bit stripping occur, the decoder at a receiving device can produce audio of suitable quality given that bits have been allocated between the bands and the regions of transform coefficients have been ordered by importance.

(無元件符號說明)

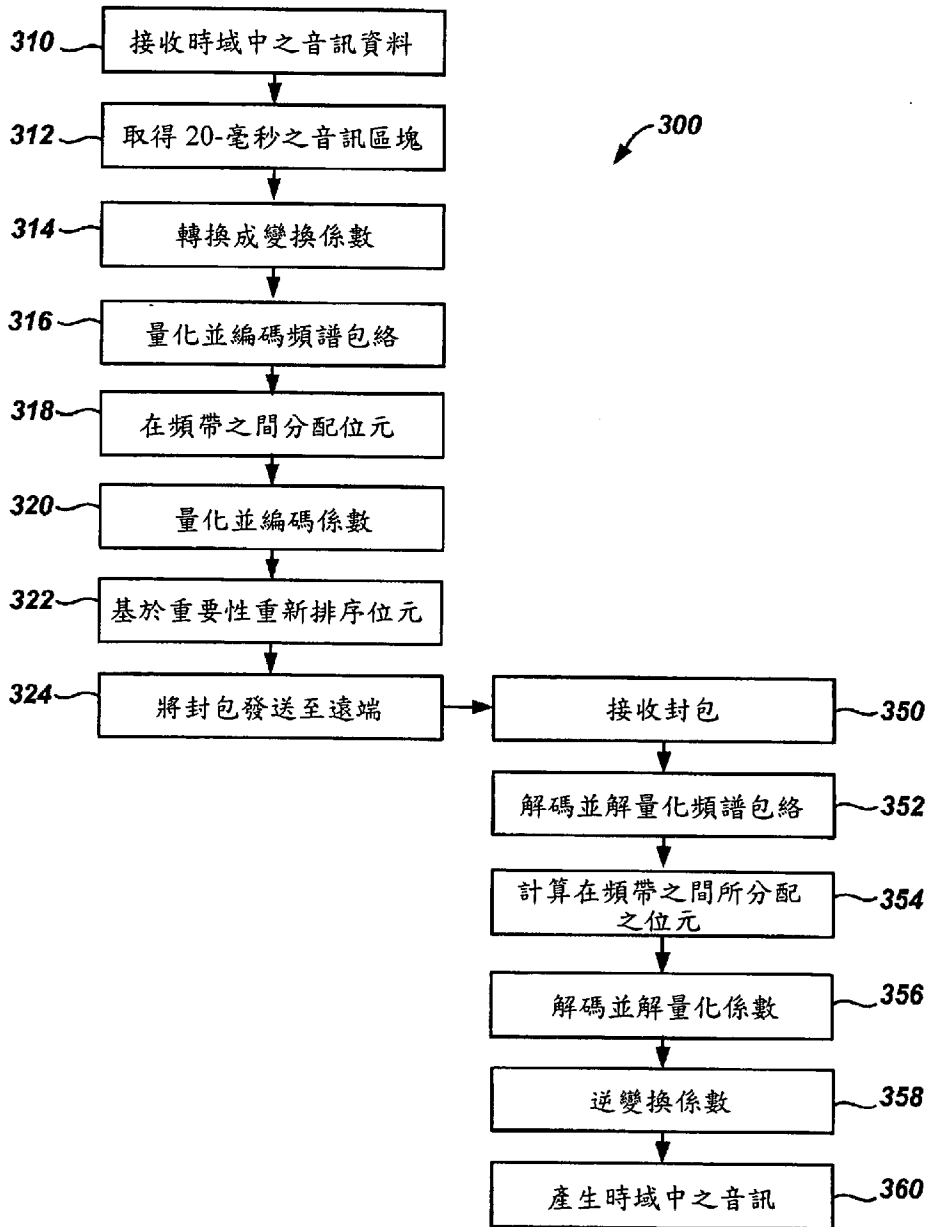


圖 3

103年1月2日 修正頁(本)
劃線

發明專利說明書

中文說明書替換頁(103年1月)

(本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

公告本

※ 申請案號：100123209

※ 申請日：100.6.30

※IPC 分類：G10L 19/02 (2013.01)

一、發明名稱：(中文/英文)

可擴縮音訊處理方法及裝置

SCALABLE AUDIO PROCESSING METHOD AND DEVICE

二、中文發明摘要：

一種用於一處理裝置之可擴縮音訊編解碼器判定每一輸入音訊訊框之第一位元分配及第二位元分配。將第一位元分配給一第一頻率頻帶，且將第二位元分配給一第二頻率頻帶。該等分配係基於該兩個頻帶之間的能量比率在一逐個訊框基礎上進行。針對每一訊框，該編解碼器將兩個頻帶變換編碼成兩個變換係數集，然後基於該等位元分配封包化該兩個變換係數集。然後利用該處理裝置傳輸該等封包。另外，可依藉由功率位準及感知模型化所判定之重要性次序配置該等變換係數之頻率區。若發生位元剝除，假定已在該等頻帶之間分配位元且已按重要性排序該等變換係數之該等區，則在一接收裝置處之解碼器可產生適合品質之音訊。

三、英文發明摘要：

A scalable audio codec for a processing device determines first and second bit allocations for each frame of input audio. First bits are allocated for a first frequency band, and second bits are allocated for a second frequency band. The allocations are made on a frame-by-frame basis based on the energy ratio between the two bands. For each frame, the codec transform codes both frequency bands into two sets of transform coefficients, which are then packetized based on the bit allocations. The packets are then transmitted with the processing device. Additionally, the frequency regions of the transform coefficients can be arranged in order of importance determined by power levels and perceptual modeling. Should bit stripping occur, the decoder at a receiving device can produce audio of suitable quality given that bits have been allocated between the bands and the regions of transform coefficients have been ordered by importance.

四、指定代表圖：

(一)本案指定代表圖為：第 (3) 圖。

(二)本代表圖之元件符號簡單說明：

(無元件符號說明)

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

(無)

六、發明說明：

【先前技術】

諸多種類型之系統使用音訊信號處理來形成音訊信號或根據此等信號再現聲音。通常，信號處理將音訊信號轉換為數位資料且編碼彼資料供在一網路上傳輸。然後，另外信號處理解碼該所傳輸之資料且將其轉換回至類比信號供再現為聲波。

存在各種用於編碼或解碼音訊信號之技術。(編碼或解碼一信號之一處理器或一處理模組通常稱作一編解碼器)。將音訊編解碼器用於會議中以減少必須自一近端傳輸至一遠端以表現音訊之資料量。舉例而言，用於音訊與視訊會議之音訊編解碼器壓縮高保真度音訊輸入，以便一形成之傳輸信號保留最佳品質但需要最少數目之位元。以此方式，具有音訊編解碼器之會議設備需要較少之儲存容量，且該設備傳輸音訊信號所用之通信頻道需要較少之頻寬。

音訊編解碼器可使用各種技術來編碼及解碼音訊供在一會議中自一個端點傳輸至另一端點。某些常用音訊編解碼器使用變換編碼技術來編碼及解碼在一網路上傳輸之音訊資料。一種類型之音訊編解碼器係Polycom's Siren編解碼器。Polycom's Siren編解碼器之一個版本係ITU-T(國際電信聯盟電信標準化組)推薦G.722.1 (Polycom Siren 7)。Siren 7係將信號最高編碼至7 kHz之一寬頻編解碼器。另一版本係ITU-T G.722.1.C (Polycom Siren 14)。Siren 14係

將信號最高編碼至14 kHz之一特級寬頻編解碼器。

Siren編解碼器係基於調變重疊變換(MLT)之音訊編解碼器。同樣，Siren編解碼器將一音訊信號自時域變換至一調變重疊變換(MLT)域。如所知曉，調變重疊變換(MLT)係用於變換編碼各種類型之信號之一餘弦調變濾波器組之一形式。一般而言，一重疊變換取得長度 L 之一音訊區塊，且將彼區塊變換成 M 個係數，條件係 $L > M$ 。對於此工作，在連續 L 至 M 個樣本區塊之間必定存在一重疊，以使得可使用連續之經變換係數區塊獲得一合成信號。

圖1A至圖1B簡要地展示一變換編碼編解碼器(諸如一Siren編解碼器)之特徵。一特定音訊編解碼器之實際細節取決於實施方案及所用編解碼器之類型。舉例而言，可在ITU-T推薦G.722.1 Annex C中找到Siren 14之已知細節，且可在ITU-T推薦G.722.1中找到Siren 7之已知細節，以引用方式將ITU-T推薦G.722.1 Annex C及ITU-T推薦G.722.1併入本文中。亦可在序號為11/550,629及11/550,682之美國專利申請案中找到關於音訊信號之變換編碼之額外細節，以引用方式將序號為11/550,629及11/550,682之美國專利申請案併入本文中。

在圖1A中圖解說明變換編碼編解碼器(例如Siren編解碼器)之一編碼器10。編碼器10接收已自一類比音訊信號轉換之一數位信號12。已以某一頻率取樣該類比音訊信號之振幅，且已將該振幅轉換為表現該振幅之一數字。典型取樣頻率係約8 kHz(亦即，每秒取樣8,000次)、16 kHz至196

kHz或兩者之間的某一值。在一項實例中，可以48 kHz或以約20個區塊或訊框每毫秒之其他速率取樣此數位信號12。

一變換20(其可係一離散餘弦變換(DCT))將數位信號12自時域轉換成具有變換係數之一頻域。舉例而言，變換20可針對每一音訊區塊或訊框產生960個變換係數之一頻譜。編碼器10在一正規化處理程序22中得出該等係數之平均能量位準(標準)。然後，編碼器10藉助一快速網格向量量化(FLVQ)演算法24或類似物量化該等係數以編碼一輸出信號14供分包及傳輸。

在圖1B中圖解說明變換編碼編解碼器(例如，Siren編解碼器)之一解碼器50。解碼器50取得自一網路接收之輸入信號52之傳入位元串流且根據其重新形成原始信號之一最佳估計。為進行此操作，解碼器50對輸入信號52執行一網格解碼(反FLVQ)60且使用一解量化處理程序62解量化該經解碼之變換係數。此外，然後可在各種頻率頻帶中校正變換係數之能量位準。最後，一逆變換64作為一反DCT操作且將信號自頻域轉換回成時域供作為一輸出信號54傳輸。

雖然此等音訊編解碼器有效，但音訊會議應用之不斷增加之需求及複雜性要求更多功能及增強之音訊編碼技術。舉例而言，音訊編解碼器必須在網路上操作，且各種條件(頻寬、接收器之不同連接速度)可動態地變化。一無線網路係其中一頻道之位元速率隨時間而變化之一項實例。因此，一無線網路中之一端點必須以不同位元速率發送一

位元串流以適應網路條件。

一MCU(多路控制單元)諸如Polycom's RMX系列及MGC系列產品之使用係其中可使用更多功能且增強之音訊編碼技術之另一實例。舉例而言，在一會議中，一MCU首先自一第一端點A接收一位元串流且然後需要以不同長度將位元串流發送至若干其他端點B、C、D、E、F...欲發送之不同位元串流將視該等端點中之每一者具有多少網路頻寬而定。舉例而言，一個端點B可以64 kbps(位元每秒)之音訊連接至該網路，而另一端點C可僅以8 kbps連接。

因此，MCU以64 kbps將位元串流發送至一個端點B，以8 kbps將位元串流發送至另一端點C，且對於該等端點中之每一者亦如此。當前，MCU解碼來自第一端點A之位元串流，亦即將其轉換回至時域。然後，MCU針對每一單個端點B、C、D、E、F...進行編碼，以使得可將該等位元串流發送至該等端點。顯然，此方法需要諸多計算資源、引入信號延時，且由於所執行之轉碼使信號品質降級。

處理丟失之封包係其中可使用更多功能及增強之音訊編碼技術之另一區域。在視訊會議或VoIP電話聯絡中，舉例而言，以通常每一封包具有20毫秒音訊之封包發送經編碼音訊資訊。封包在傳輸期間可能丟失，且丟失之音訊封包導致所接收音訊中之間隙。對抗封包在網路中丟失之一種方法係多次傳輸封包(亦即，位元串流)，例如4次。丟失此等封包中之所有四個封包之機率甚低，因而減少了具有間隙之機率。

然而多次傳輸封包需要將網路頻寬增加到四倍。為使成本最小化，通常將同一20毫秒時域信號以較高位元速率（在一標準模式中，例如48 kbps）編碼並以一較低位元速率（例如，8 kbps）編碼。該較低（8 kbps）位元串流係傳輸多次之位元串流。如此，總需要頻寬係 $48+8*3=72$ kbps，而非將原始位元串流發送多次情形下之 $48*4=192$ kbps。由於遮蔽效應，當網路具有丟失封包時，在通話品質方面， $48+8*3$ 方案幾乎與 $48*4$ 方案一樣地執行。然而，以不同位元速率獨立地編碼同一20毫秒時域資料之此傳統方案需要計算資源。

最後，某些端點可不具有足夠的進行一全解碼之計算資源。舉例而言，一端點可具有一較慢信號處理器或該信號處理器可忙著做其他任務。若此係該情形，則解碼該端點所接收的位元串流之僅一部分可不產生有用音訊。如所習知，音訊品質取決於解碼器接收並解碼了多少個位元。

出於此等原因，存在對用於音訊及視訊會議中之可擴縮之一音訊編解碼器之需要。

【發明內容】

如在背景中所提及，音訊會議應用之不斷增加之需求及複雜性要求更多功能及增強之音訊編碼技術。具體而言，存在對用於音訊及視訊會議中之可擴縮之一音訊編解碼器之需要。

根據本發明，用於一處理裝置之一可擴縮音訊編解碼器判定每一輸入音訊訊框之第一位元分配及第二位元分配。

將第一位元分配給一第一頻率頻帶，且將第二位元分配給一第二頻率頻帶。該等分配係基於該兩個頻帶之間的能量比率在一逐個訊框基礎上進行。針對每一訊框，該編解碼器將兩個頻率頻帶變換成兩個變換係數集，基於該等位元分配將該兩個變換係數集量化且然後封包化。然後利用該處理裝置傳輸該等封包。另外，可按依功率位準及感知模型化所判定之重要性次序配置該等變換係數之頻率區。若發生位元剝除，假設已在該等頻帶之間分配位元且已按重要性排序該等變換係數之區，則在一接收裝置處之解碼器可產生適合品質之音訊。

該可擴縮音訊編解碼器對輸入音訊在一逐個訊框基礎上執行一動態位元分配。在一低頻率頻帶與一高頻率頻帶之間分配該訊框之總可用位元。在一個配置中，低頻率頻帶包括0至14 kHz，而高頻率頻帶包括14 kHz至22 kHz。給定訊框中之兩個頻帶之間的能量位準比率確定針對每一頻帶分配多少個可用位元。一般而言，意欲給低頻率頻帶分配較多可用位元。此在一逐個訊框基礎上之動態位元分配允許音訊編解碼器針對言語聲調之一致性感知來編碼及解碼所傳輸之音訊。換言之，即使在處理期間可發生之極低位元速率下，仍可將音訊視作全頻帶言語。此係由於始終獲得至少14 kHz之一頻寬。

該可擴縮音訊編解碼器將頻率頻寬擴展為至多全頻帶，亦即，22 kHz。整體地，該音訊編解碼器可自約10 kbps擴大為至多64 kbps。值10 kbps可不同且係針對一給定實施

方案之可接受編碼品質進行選擇。無論如何，所揭示之音訊編解碼器之編碼品質可與稱作 Siren 14 之音訊編解碼器之 22 kHz 版本之固定速率約相同。在 28 kbps 及以上之情形下，所揭示之音訊編解碼器與一 22 kHz 編解碼器相當。另外，在低於 28 kbps 下，所揭示之音訊編解碼器與一 14 kHz 編解碼器相當，乃因其在任一速率下皆具有至少 14 kHz 之頻寬。所揭示之音訊編解碼器可與眾不同地通過使用係真實語音信號之掃描音、白色雜訊之測試。然而，所揭示之音訊編解碼器需要僅係現有 Siren 14 音訊編解碼器當前所需要之約 1.5x 之計算資源及記憶體要求。

除位元分配之外，可擴縮音訊編解碼器基於該等頻率頻帶中之每一者中之每一區之重要性執行位元重新排序。舉例而言，一訊框之低頻率頻帶具有配置於複數個區中之變換係數。該音訊編解碼器判定此等區中之每一者之重要性且然後按重要性之次序將該等區與分配給該頻帶之位元封包化。判定該等區之重要性之一種方式係基於區之功率位準，從而按重要性次序以最高功率位準至最低功率位準配置彼等區。可基於使用周圍區之一加權來判定重要性之一感知模型擴張此判定。

藉助可擴縮音訊編解碼器來解碼封包利用了位元分配及根據重要性之經重新排序之頻率區。若出於某種原因，一所接收封包之位元串流之一部分被剝除，則音訊編解碼器可首先解碼該位元串流中之至少該較低頻率頻帶，其中較高頻率頻帶在一定程度上潛在地受到位元剝除。而且，由

於該頻帶之區針對重要性之排序，首先解碼具有較高功率位準之較重要位元，且該等較重要位元較不可能被剝除。

如上文所論述，本發明之可擴縮音訊編解碼器允許自編碼器所產生之一位元串流剝除位元，而解碼器仍可產生時域中之可理解音訊。出於此原因，可擴縮音訊編解碼器可用於若干應用中，下文將論述其等中之某些。

在一項實例中，可擴縮音訊編解碼器可用於其中一端點必須以不同位元速率發送一一位元串流以適應網路條件之一無線網路中。當使用一MCU時，該可擴縮音訊編解碼器可藉由剝除位元形成以不同位元速率發送至各個端點之位元串流，而不藉由習用做法。因此，該MCU可使用該可擴縮音訊編解碼器藉由自來自一第一端點之一64 kbps位元串流剝除位元來獲得用於一第二端點之一8 kbps位元串流，而仍維持有用音訊。

可擴縮音訊編解碼器之使用亦可在處理丟失封包時幫助節約計算資源。如前文所提及，處理丟失封包之習用解決方案係以高位元速率及低位元速率(例如，48 kbps及8 kbps)獨立地編碼同一20毫秒時域資料，以便可多次發送低品質(8 kbps)位元串流。然而，當使用可擴縮音訊編解碼器時，編解碼器僅需要編碼一次，乃因藉由自第一(高品質)位元串流剝除下位元來獲得第二(低品質)位元串流，而仍維持有用音訊。

最後，可擴縮音訊編解碼器在其中一端點可無足夠計算資源進行一全解碼之情形中有幫助。舉例而言，該端點可

具有一較低信號處理器，或該信號處理器可正忙於其他任務。在此情形中，使用可擴縮音訊編解碼器解碼該端點所接收之位元串流之一部分仍可產生有用音訊。

前述發明內容並不意欲概述本發明之每一潛在實施例或每一態樣。

【實施方式】

根據本發明之一音訊編解碼器係可擴縮的且在頻率頻帶之間分配可用位元。另外，該音訊編解碼器基於重要性來排序此等頻帶中之每一者之頻率區。若發生位元剝除，則首先將具有較重要性之彼等頻率區封包化於一位元串流中。以此方式，即使在發生位元剝除之情形下，亦將維持較有用之音訊。本文中揭示音訊編解碼器之此等及其他細節。

本發明之各種實施例可在諸如音訊會議、視訊會議及串流媒體(包括串流音樂或言語)之領域中找到有用應用。因此，本發明之一音訊處理裝置可包括：一音訊會議端點、一視訊會議端點、一音訊播放裝置、一個人音樂播放器、一電腦、一伺服器、一電信裝置、一蜂巢式電話、一個人數位助理、VoIP電話通信設備、呼叫中心設備、語音記錄設備、語音訊息接發設備等。舉例而言，特殊用途之音訊或視訊會議端點可受益於所揭示之技術。同樣，電腦或其他裝置可用於桌上會議或用於傳輸及接收數位音訊，且此等裝置亦可受益於所揭示之技術。

A. 會議端點

如上文所提及，本發明之一音訊處理裝置可包括一會議端點或終端機。圖2A示意性地展示一端點或終端機100之一實例。如所展示，會議終端機100可係在一網路125上之一傳輸器及一接收器兩者。亦如所展示，會議終端機100可具有視訊會議能力以及音訊能力。一般而言，終端機100具有一麥克風102及一揚聲器108且可具有各種其他輸入/輸出裝置，諸如一音訊相機103、顯示器109、鍵盤、滑鼠等。另外，終端機100具有一處理器160、記憶體162、轉換器電子器件164、及適合特定網路125之網路介面122/124。音訊編解碼器110根據適合於各個經網路化之終端機之一協定提供基於標準之會議。此等標準可完全以儲存於記憶體162且執行於處理器160上之軟體、在專用硬體上之軟體來執行，或使用其一組合來執行。

在一傳輸路徑中，轉換器電子器件164將麥克風102所拾取之類比輸入信號轉換成數位信號，且在終端機之處理器160上操作之音訊編解碼器110具有一編碼器200，編碼器200編碼該等數位音訊信號供經由一傳輸器介面122在網路125(諸如網際網路)上傳輸。若存在，具有一視訊編碼器170之一視訊編解碼器可針對視訊信號執行類似功能。

在一接收路徑中，終端機100具有耦合至音訊編解碼器110之一網路接收器介面124。一解碼器250解碼所接收之音訊信號，且轉換器電子器件164將數位信號轉換為類比信號供輸出至揚聲器108。若存在，具有一視訊解碼器172之一視訊編解碼器可針對視訊信號執行類似功能。

B. 音訊處理配置

圖 2B 展示一會議配置，其中一第一音訊處理裝置 100A (充當一傳輸器) 將經壓縮之音訊信號發送至一第二音訊處理裝置 100B (在此背景中充當一接收器)。傳輸器 100A 及接收器 100B 兩者皆具有一可擴縮音訊編解碼器 110，其類似於在 ITU G. 722.1 (Polycom Siren 7) 或 ITU G.722.1.C (Polycom Siren 14) 中所用地執行變換編碼。對於本論述，傳輸器及接收器 100A 至 100B 可係一音訊會議或視訊會議中之端點或終端機，雖然其等可係其他類型之裝置。

在操作期間，在傳輸器 100A 處之一麥克風 102 捕獲原音訊，且電子器件將彼音訊之區塊或訊框取樣。通常，音訊區塊或訊框橫跨 20 毫秒之輸入音訊。此時，音訊編解碼器 110 之一正向變換將每一音訊訊框轉換為一頻域變換係數組。使用該技術中所已知的技術，然後藉助一量化器 115 將此等變換係數量化並編碼。

一旦經編碼，傳輸器 100A 就使用其網路介面 120 以封包形式經由一網路 125 將該等經編碼之變換係數發送至接收器 100B。可使用任一適合網路，包括但不限於一 IP (網際網路協定) 網路、PSTN (公共交換電話網路)、ISDN (整合式服務數位網路) 或類似網路。對於此部分，所傳輸之封包可使用任何適合協定或標準。舉例而言，封包中之音訊資料可遵循一目錄，且組成一音訊訊框之所有八位元組皆可作為一單元附加至酬載。在 ITU-T 推薦 G.722.1 及 G.722.1C 中明確說明瞭音訊訊框及封包之額外細節，已將 ITU-T 推

薦 G.722.1 及 G.722.1C 併入本文中。

在接收器 100B 處，一網路介面 120 接收該等封包。在如下一反過程中，接收器 100B 使用編解碼器 110 之一解量化器 115 及一逆變換來解量化並解碼該等經編碼之變換係數。該逆變換將該等係數轉換回成時域以產生用於接收器之揚聲器 108 之輸出音訊。對於音訊及視訊會議，接收器 100B 及傳輸器 100A 可在一會議期間具有往復作用。

C. 音訊編解碼器操作

在理解了上文所提供之音訊編解碼器 110 及音訊處理裝置 100 之情形下，論述現在轉向音訊編解碼器 110 如何根據本發明編碼及解碼音訊。如在圖 3 中所展示，傳輸器 110A 處之音訊編解碼器 110 接收時域中之音訊資料(方塊 310)且取得一音訊區塊或音訊資料訊框(方塊 312)。

使用正向變換，音訊編解碼器 110 將音訊訊框轉換成頻域中之變換係數(方塊 314)。如上文所論述，音訊編解碼器 110 可使用 Polycom Siren 技術來執行此變換。然而，音訊編解碼器可係任一變換編解碼器，包括但不限於 MP3、MPEG AAC 等。

當變換該音訊訊框時，音訊編解碼器 110 亦針對該訊框量化並編碼頻譜包絡(方塊 316)。此包絡闡述正被編碼之音訊之振幅，雖然其不提供任何相細節。編碼包絡頻譜不需要大量位元，因而其可係容易實現的。然而，如下文將可見，若自傳輸剝除位元，則稍後在音訊解碼期間可使用頻譜包絡。

當在一網路(諸如網際網路)上通信時，頻寬可改變，封包可丟失，且連接速率可不同。為慮及此等挑戰，本發明之音訊編解碼器110係可擴縮的。以此方式，在稍後予以更詳細闡述之一過程中音訊編解碼器110在至少兩個頻率頻帶之間分配可用位元(方塊318)。編解碼器之編碼器200量化並編碼所分配之頻率頻帶中之每一者中之變換係數(方塊320)且然後基於區之重要性重新排序每一頻率區之位元(方塊322)。從頭到尾，整個編碼過程可僅引入約20毫秒之一延遲。

下文所更詳細闡述之判定一位元重要性改良了在位元出於若干原因被剝除之情形下可在遠端再現之音訊品質。在重新排序該等位元之後，將位元分包供發送至遠端。最後，將該等封包傳輸至遠端，以便可處理下一訊框(方塊324)。

在遠端，接收器100B接收該等封包，根據已知技術處置該等封包。編解碼器之解碼器250然後解碼並解量化頻譜包絡(方塊352)且判定在頻率頻帶之間所分配之位元(方塊354)。稍後將提供解碼器250如何判定在頻率頻帶之間之位元分配之細節。在知曉位元分配之情形下，解碼器250然後解碼並解量化該等變換係數(方塊356)且對每一頻帶中之係數執行一逆變換(方塊358)。最終，解碼器250將音訊轉換回成時域以產生用於接收器之揚聲器之輸出音訊(方塊360)。

D. 編碼技術

如上文所提及，所揭示之音訊編解碼器110係可擴縮的且使用變換編碼來將音訊編碼於分配給至少兩個頻率頻帶之位元中。在圖4整個流程圖中展示可擴縮音訊編解碼器110所執行之編碼技術之細節。最初，音訊編解碼器110獲得一輸入音訊訊框(方塊402)且使用此項技術中所習知之一調變重疊變換技術來將該訊框轉換成變換係數(方塊404)。如所已知，此等變換係數中之每一者皆具有一量值且可係正或負。音訊編解碼器110亦如前文所提及量化並編碼該頻譜包絡[0 Hz至22 kHz](方塊406)。

此時，音訊編解碼器110在至少兩個頻率頻帶之間分配該訊框之位元(方塊408)。此位元分配係當音訊編解碼器110編碼所接收之音訊資料時動態地在一逐個訊框基礎上來判定。在該兩個頻帶之間選擇一劃分頻率，以便將第一數目個可用位元分配給低於該劃分頻率之一低頻率區，且將剩餘位元分配給高於該劃分頻率之一較高頻率區。

在針對頻帶判定位元分配之後，音訊編解碼器110以該等經正規化之係數之各別分配位元將該等經正規化之係數編碼於低頻率頻帶及高頻率頻帶兩者中(方塊410)。然後，音訊編解碼器110判定此兩個頻率頻帶中之每一頻率區之重要性(方塊412)且基於所判定之重要性排序該等頻率區(方塊414)。

如前文所提及，音訊編解碼器110可類似於Siren編解碼器且可將音訊信號自時域變換成具有MLT係數之頻域。(簡明起見，本發明針對此一MLT變換來提及變換係數，

雖然可使用其他類型之變換，諸如FFT(快速傅立葉變換)及DCT(離散餘弦變換)等)。

在該取樣速率下，MLT變換產生約960個MLT係數(亦即，每25 Hz一個係數)。此等係數根據具有0、1、2...之索引之遞增順序配置於頻率區中。舉例而言，一第一區0涵蓋頻率範圍[0至500 Hz]，下一區1涵蓋[500至1000 Hz]，且以此類推。可擴縮音訊編解碼器110並不簡單地如習用方式所做按遞增順序發送該等頻率區，而是在整個音訊之背景中判定該等區之重要性，且然後基於較高重要性至較低重要性來重新排序該等區。在該兩個頻率頻帶中進行基於重要性之此重新配置。

可以諸多方式進行對每一頻率區之重要性之判定。在一項實施方案中，編碼器200基於經量化信號功率頻譜來判定區之重要性。在此情形中，具有較高功率之區具有較高重要性。在另一實施方案中，可使用一感知模型來判定該等區之重要性。該感知模型遮蔽人們感知不到之外來音訊、雜訊及類似物。稍後更詳細地論述此等技術中之每一者。

在基於重要性之排序之後，首先封包化最重要之區，後跟一重要性較小一點之區，後跟較不重要區，以此類推(方塊416)。最後，可在網路上將經排序及經封包化之區發送至遠端(方塊420)。在發送該等封包中，無需發送關於排序變換係數之區之編索引資訊。而是，可在解碼器中基於自位元串流解碼出之頻譜包絡來計算編索引資訊。

若發生位元剝除，則朝向該終端之彼等經封包化之位元可被剝除。由於該等區已經排序，因而在最重要區中之係數已被首先封包化。因此，最後經封包化之較不重要區在發生位元剝除之情形下較可能被剝除。

在遠端，解碼器250解碼並變換所接收之資料，該所接收之資料已反映最初由傳輸器100A給出之經排序之重要性。以此方式，當接收器100B解碼該等封包且產生時域中之音訊時，該接收器之音訊編解碼器110實際上將接收到並處理該輸入音訊中之較重要係數區之機會增加。如所預期，在會議期間，頻寬、計算能力及其他資源之改變可改變，從而使得音訊丟失、未被編碼等。

在已將音訊分配於頻率頻帶之間的位元中且針對重要性排序之後，音訊編解碼器110可增加在遠端將處理較有用音訊之機會。鑒於所有此原因，當出於某種原因而存在降低之音訊品質時，即使自位元串流剝除位元(亦即，部分位元串流)，音訊編解碼器110仍可產生一有用音訊信號。

1. 位元分配

如前文所提及，本發明之可擴縮音訊編解碼器110在兩個頻率頻帶之間分配可用位元。如在圖4B中所展示，該音訊編解碼器(110)在一特定頻率(例如48 kHz)下將一音訊信號430取樣及數位化於每一者約為20毫秒之連續訊框F1、F2、F3等中。(實際上，該等訊框可重疊)。因此，每一訊框F1、F2、F3等具有約960個樣本($48 \text{ kHz} \times 0.02 \text{ s} = 960$)。音訊編解碼器(110)然後將每一訊框F1、F2、F3等自時域

變換為頻域。對於一給定訊框，舉例而言，該變換如在圖4C中所展示產生一MLT係數組。針對該訊框存在約960個MLT係數(亦即，每25 Hz一個MLT係數)。由於22 kHz之編碼頻寬，因而可忽略表現在約22 kHz以上之頻率之MLT變換係數。

自0至22 kHz之頻域中之變換係數組必須經編碼，以便可將該經編碼資訊封包化且在一網路上傳輸。在一個配置中，音訊編解碼器(110)經組態以便以一最大速率(其可係64 kbps)編碼該全頻帶音訊信號。然而，如本文中所闡述，該音訊編解碼器(110)分配可用位元用於在兩個頻率頻帶之間編碼訊框。

為分配該等位元，音訊編解碼器110可在一第一頻帶[0至12 kHz]與一第二頻帶[12 kHz至22 kHz]之間劃分總可用位元。兩個頻帶之間的12 kHz之劃分頻率可主要基於言語聲調改變及主觀測試來選擇。對於一給定實施方案可使用其他劃分頻率。

基於兩個頻帶之間的能量比率來分割該等總可用位元。在一項實例中，可存在用於在兩個頻帶之間分割之四個可能方式。舉例而言，可如下劃分64 kbps之該等總可用位元：

表 1

四模式位元分配實例

模式	<12 kHz之信號 之分配	>12 kHz之信號 之分配	總可用頻寬(kbps)
0	48	16	64
1	44	20	64
2	40	24	64
3	36	28	64

在傳輸至遠端之資訊中表現此四個可能性需要編碼器 (200) 在傳輸之位元串流中使用 2 個位元。遠端解碼器 (250) 可使用來自此等所傳輸位元之資訊在接收到給定訊框時判定該給定訊框之位元分配。在知曉位元分配之情形下，解碼器 (250) 然後可基於此所判定之位元分配來解碼該信號。

在圖 4C 中所展示之另一配置中，該音訊編解碼器 (110) 經組態以藉由在一第一頻帶 (LoBand) 440 [0 至 14 kHz] 與一第二頻帶 (HiBand) 450 [14 kHz 至 22 kHz] 之間劃分總可用位元來分配該等位元。雖然可端視實施方案而使用其他值，但由言語/音樂、嘈雜/乾淨、男聲/女聲等看來，基於主觀收聽品質，14 kHz 之劃分頻率可係較佳的。在 14 kHz 處將信號分割成 HiBand 與 LoBand 亦使可擴縮音訊編解碼器 110 與現有 Siren 14 音訊編解碼器相當。

在此配置中，可以八 (8) 個可能分割模式在一逐個訊框基礎上分割該等訊框。該八個模式 (bit_split_mode) 係基於兩個頻帶 440/450 之間的能量比率。此處，將低頻率頻帶 (LoBand) 之能量或功率值標示為 LoBandsPower，而將高頻

率頻帶(HiBand)之能量或功率值標示為HiBandsPower。如下判定一給定訊框之特定模式(bit_split_mode)：

若 $(\text{HiBandsPower} > (\text{LoBandsPower} * 4.0))$ ，

則 $\text{bit_split_mode} = 7$ ；

否則，若 $(\text{HiBandsPower} > (\text{LoBandsPower} * 3.0))$ ，

則 $\text{bit_split_mode} = 6$ ；

否則，若 $(\text{HiBandsPower} > (\text{LoBandsPower} * 2.0))$ ，

則 $\text{bit_split_mode} = 5$ ；

否則，若 $(\text{HiBandsPower} > (\text{LoBandsPower} * 1.0))$ ，

則 $\text{bit_split_mode} = 4$ ；

否則，若 $(\text{HiBandsPower} > (\text{LoBandsPower} * 0.5))$ ，

則 $\text{bit_split_mode} = 3$ ；

否則，若 $(\text{HiBandsPower} > (\text{LoBandsPower} * 0.01))$

則 $\text{bit_split_mode} = 2$ ；

否則，若 $(\text{HiBandsPower} > (\text{LoBandsPower} * 0.001))$

則 $\text{bit_split_mode} = 1$ ；

否則 $\text{bit_split_mode} = 0$ ；

此處，低頻率頻帶之功率值(LoBandsPower)係按照 $\sum_i \text{quantized_region_power}[i]$ 來計算，其中區索引 $i=0、1、2、\dots、25$ 。(由於每一區之頻寬係500-Hz，因而對應頻率範圍係0 Hz至12,500 Hz)。可使用如可用於現有Siren編解碼器之一預界定表來量化每一區之功率以獲得 $\text{quantized_region_power}[i]$ 之值。對於此部分，類似地計算高頻率頻帶之功率值(HiBandsPower)，但使用自13 kHz至22 kHz之頻率範

圍。因此，在此位元分配技術中該劃分頻率實際上係13 kHz，雖然信號頻譜係在14 kHz處分割。進行此操作以通過一掃描正弦波測試。

然後如上文所提及，基於根據頻帶之功率值之能量比率所判定之bit_split_mode來計算兩個頻率頻帶440/450之位元分配。特定而言，HiBand頻率頻帶獲得總可用64 kbps之 $(16 + 4 * \text{bit_split_mode})$ kbps，而LoBand頻率頻帶獲得總64 kbps之剩餘位元。此分解為以下針對8個模式之分配：

表 2

八模式位元分配實例

模式	<14 kHz之信號 之分配	> 14 kHz之信號 之分配	總可用頻寬(kbps)
0	48	16	64
1	44	20	64
2	40	24	64
3	36	28	64
4	32	32	64
5	28	36	64
6	24	40	64
7	20	44	64

在傳輸至遠端之資訊中表現此八個可能性需要傳輸編解碼器(110)在位元串流中使用3個位元。遠端解碼器(250)可使用來自此3個位元之所指示之位元分配，且可基於此位元分配解碼該給定訊框。

圖4D用圖表表示該八個可能模式(0-7)之位元分配460。由於該等訊框具有20毫秒之音訊，因而64 kbps之最大位元

速率對應於每一訊框之總1280個可用位元(亦即, 64,000 bps \times 0.02 s)。同樣, 所用模式取決於兩個頻率頻帶之功率值474與475之能量比率。各個比率值470亦以圖表形式繪示於圖4D中。

因此, 若HiBand之功率值475大於LoBand之功率值474之四倍, 則所判定之bit_split_mode將係「7」。此對應於針對LoBand之20 kbps(或400個位元)之一第一位元分配464且對應於針對可用64 kbps(或1280個位元)之HiBand之44 kbps(或880個位元)之一第二位元分配465。作為另一實例, 若HiBand之功率值464大於LoBand之功率值465之一半但小於LoBand之功率值464之一倍, 則所判定之bit_split_mode將係「3」。此對應於針對LoBand之36 kbps(或720個位元)之第一位元分配464且對應於針對可用64 kbps(或1280個位元)之HiBand之28 kbps(或560個位元)之第二位元分配465。

如自此兩個可能位元分配形式可見, 判定如何在兩個頻率頻帶之間分配位元可取決於一給定實施方案之細節之數目, 且此等位元分配方案意欲係實例性。甚至可以想像在位元分配中可涉及多於兩個頻率頻帶以進一步細化一給定音訊信號之位元分配。因此, 在給出本發明之教示之情形下, 本發明之整個位元分配及音訊編碼/解碼可經擴張而涵蓋多於兩個頻率頻帶及更多或更少之分割模式。

2. 重新排序

如上文所提及, 除位元分配之外, 所揭示音訊編解碼器

(110)重新排序在較重要區中之係數，以便首先將其封包化。以此方式，當由於通信問題位元自位元串流剝除時較少可能移除該等較重要區。舉例而言，圖5A展示進入一位元串流500中之區之一習用封包化次序。如前文所提及，每一區具有針對一對應頻率範圍之變換係數。如所展示，在此習用配置中，針對頻率範圍[0至500 Hz]之第一區「0」首先被封包化。其次封包化涵蓋[500至1000Hz]之下一區「1」，且重複此過程，直至將最後一個區封包化為止。結果係具有按頻率區0、1、2、...、N之遞增順序配置之區之習用位元串流500。

藉由判定區之重要性且然後首先將最重要區封包化於位元串流中，本發明之音訊編解碼器110產生如圖5B中所展示的一位元串流510。此處，首先封包化最重要區(與其頻率範圍無關)，後跟第二最重要區。重複此過程，直至將最不重要區封包化為止。

如在圖5C中所展示，出於某些原因，位元可自位元串流510剝除。舉例而言，位元可在傳輸位元串流或接收位元串流時被漏掉。然而，仍可對剩餘位元串流進行解碼直至已保留之彼等位元。由於已基於重要性排序該等位元，因而針對最不重要區之位元520在發生位元剝除時係最可能被剝除之位元。最後，如在圖5C中所證明，即使在所重新排序之位元串流510上發生位元剝除，仍可保留整體音訊品質。

3. 用於判定重要性之功率頻譜技術

如前文所提及，一種用於判定經編碼音訊中之區之重要性之技術使用該等區之功率信號來排序該等區。如在圖6A中所展示，所揭示音訊編解碼器(110)使用的一功率頻譜模型600計算每一區(亦即，區0[0至500 Hz]、區1[500至1000 Hz]等)之信號功率(方塊602)。進行此操作之一種方法係，對於音訊編解碼器(110)，計算給定區中之變換係數中之每一者之平方之和，且使用此值代表給定區之信號功率。

在將給定頻率頻帶之音訊轉換成變換係數(舉例而言，如在圖4之方塊410處所進行)之後，音訊編解碼器(110)計算每一區中之係數之平方。對於當前變換，每一區涵蓋500 Hz且具有各自涵蓋25 Hz之20個變換係數。在給定區中之此20個變換係數中之每一者之平方之和產生此區之功率頻譜。此係針對所討論頻帶中之每一區來進行，以計算該所討論頻帶中之區中之每一者之一功率頻譜值。

一旦計算出該等區之信號功率(方塊602)，就將其量化(方塊603)。然後，模型600以功率遞減順序將該等區排序，在每一頻帶中以最高功率區開始且以最低功率區結束(方塊604)。最後，音訊編解碼器(110)藉由以所判定之次序將該等係數之位元封包化來完成模型600(方塊606)。

最後，音訊編解碼器(110)已基於與其他區相比之一區之信號功率判定該區之重要性。在此情形中，具有較高功率之區具有較高重要性。若在傳輸過程中出於某種原因最後經封包化之區被剝除，則具有較大功率信號之彼等區已被首先封包化且較可能含有將不被剝除之有用音訊。

4. 用於判定重要性之感知技術

如前文所提及，用於判定在經編碼信號中之一區之重要性之另一技術使用一感知模型 650--在圖 6B 中展示其一實例。首先，感知模型 650 計算兩個頻帶中之每一者中之每一區之信號功率，其可以與上文所闡述之方式極其相同之方式來進行(方塊 652)，且然後模型 650 量化該信號功率(方塊 653)。

模型 650 然後界定每一區之一經修改區功率值(亦即 `modified_region_power`)(方塊 654)。經修改區功率值係基於一經加權和，其中當考量一給定區之重要性時慮及周圍區之效應。因此，感知模型 650 利用一個區中之信號功率可遮蔽另一區中之量化雜訊且當該等區在頻譜上接近時此遮蔽效應較大之事實。因此，可按如下界定一給定區之經修改區功率值(亦即，`modified_region_power(region_index)`)：

$$\text{SUM}(\text{權}[\text{region_index}, r] * \text{quantized_region_power}(r)) ;$$

其中 $r=[0\dots 43]$ ，

其中 `quantized_region_power(r)` 係該區之經計算信號功率；及

其中 `權[region_index, r]` 係隨著頻譜距離 $|\text{region_index}-r|$ 增加而下降之一固定函數。

因此，若如下界定加權函數，則感知模型 650 還原至圖 6A 之模型：

當 $r=\text{region_index}$ 時， $\text{權}(\text{region_index}, r)=1$

當 $r \neq \text{region_index}$ 時， $\text{權}(\text{region_index}, r)=0$

在如上文所略述地計算經修改區功率值之後，感知模型 650 基於該等經修改區功率值以遞減順序將該等區排序(方塊 656)。如上文所提及，由於已進行加權，因而一個區中之信號功率可遮蔽另一區中之量化雜訊，尤其當該等區在頻譜上彼此接近時。音訊編解碼器(110)然後藉由按所判定之次序封包化該等區之位元來完成模型 650(方塊 658)。

5. 封包化

如上文所論述，所揭示之音訊編解碼器(110)編碼該等位元且將其封包化，以使得可將用於低頻率頻帶及高頻率頻帶之特定位元分配細節發送至遠端解碼器(250)。此外，將頻譜包絡連同所分配的用於該兩個經封包化之頻率頻帶中之變換係數之位元一起封包化。下表展示如何將位元封包化(自第一位元至最後位元)於欲自近端傳輸至遠端之一給定訊框之一位元串流中。

表 3

封包化實例

分割模式	LoBand 頻率		HiBand 頻率	
用於 split_mode 之 3 個位元(總共 8 個模式)	以上升之區 次序用於包 絡之位元	所分配的用於 所重新排序的 正規化係數之 位元	以上升之區 次序用於包 絡之位元	所分配的用於 所重新排序的 正規化係數之 位元

如可見，首先針對該訊框封包化指示(該八個可能模式之)特定位元分配之三(3)個位元。然後，藉由首先將用於低頻率頻帶(LoBand)之頻譜包絡之位元封包化來封包化此頻帶。通常，包絡無需編碼諸多位元，乃因其包括振幅資

訊而非相。在將包絡之位元封包化之後，將用於低頻率頻帶(LoBand)之正規化係數之所分配之特定數目個位元封包化。用於頻譜包絡之位元簡單地基於其典型遞增順序封包化。然而，所分配之用於低頻率頻帶(LoBand)係數之位元如其已經重新排序地根據重要性封包化，如前文所略述。

最後，可見，藉由首先封包化用於高頻率頻帶(HiBand)之頻譜包絡之位元且然後以同樣方式封包化所分配的用於HiBand頻率頻帶之正規化係數之特定數目個位元來封包化此頻帶。

E. 解碼技術

如前文在圖2A中所提及，所揭示音訊編解碼器110之解碼器250在接收到封包時解碼位元，以便音訊編解碼器110可將該等係數變換回至時域以產生輸出音訊。在圖7中更詳細地展示此過程。

最初，接收器(例如，圖2B之100B)接收該位元串流中之封包且使用已知技術處置該等封包(方塊702)。當發送該等封包時，舉例而言，傳輸器100A形成序號，該等序號包括於所發送之封包中。如所已知，封包可在網路125上經由不同路線自傳輸器100A傳遞至接收器100B，且該等封包可在不同時間到達接收器100B。因此，封包到達之次序可係隨機的。為處置此不同到達時間(稱作「抖動」)，接收器100B具有耦合至該接收器之介面120之一抖動緩衝器(未展示)。通常，抖動緩衝器一次容納四個或四個以上封包。因此，接收器100B基於封包之序號在抖動緩衝器中重

新排序封包。

使用位元串流中之前三個位元(例如，圖5B之520)，解碼器250解碼用於正被處置之給定訊框之位元分配之封包(方塊704)。如前文所提及，端視組態，在一項實施方案中可存在8個可能位元分配。在知曉所用分割(如前三個位元所指示)之情形下，解碼器250然後針對分配給每一頻帶之位元之數目解碼。

以低頻開始，解碼器250解碼並解量化該訊框之低頻率頻帶(LoBand)之頻譜包絡(方塊706)。然後，解碼器250解碼並解量化低頻率頻帶之係數，只要位元已被接收且未被剝除。因此，解碼器250經歷一反覆過程且判定是否還有位元剩下(決定710)。只要存在位元，解碼器250就解碼低頻率頻帶中之區之正規化係數(方塊712)並計算當前係數值(方塊714)。對於該計算，解碼器250按照如下計算變換係數：係數=包絡*normalized_coeff，其中將頻譜包絡之值乘以正規化係數之值(方塊714)。此操作繼續，直至針對低頻率頻帶將所有位元解碼且將其乘以頻譜包絡值為止。

由於已根據頻率區之重要性排序該等位元，因而解碼器250可能首先解碼位元串流中之最重要區，而無論該位元串流是否有位元剝除。解碼器250然後解碼第二最重要區，且以此類推。解碼器250繼續，直至所有位元用完為止(決定710)。

當對所有位元操作完時(由於位元剝除，其實際上可並非所有彼等經原始編碼之位元)，用雜訊填充可能已剝除

之彼等最不重要區以完成此低頻率頻帶中之信號之剩餘部分。

若該位元串流已被剝除位元，則所剝除之位元之係數資訊已丟失。然而，解碼器 250 已接收到並解碼低頻率頻帶之頻譜包絡。因此，解碼器 250 至少知曉該信號之振幅，但不知曉其相。為填充雜訊，解碼器 250 在所剝除之位元中針對已知振幅填充相資訊。

為填充雜訊，解碼器 250 計算缺乏位元之任何剩餘區之係數(方塊 716)。按照頻譜包絡之值乘以一雜訊填充值來計算剩餘區之此等係數。此雜訊填充值可係用於填充由於位元剝除導致丟失之缺失區之係數之一隨機值。藉由用雜訊填充，解碼器 250 最終可將該位元串流視作全頻帶，即使在一極低之位元速率下，諸如 10 kbps。

在處置低頻率頻帶之後，解碼器 250 對高頻率頻帶 (HiBand) 重複整個過程(方塊 720)。因此，解碼器 250 解碼並解量化 HiBand 之頻譜包絡，解碼位元之正規化係數，計算位元之當前係數值，且計算缺乏位元之剩餘區之雜訊填充係數(若被剝除)。

既然解碼器 250 已判定在 LoBand 及 HiBand 兩者中之所有區之變換係數，且知曉根據頻譜包絡得出之區之次序，解碼器 250 對變換係數執行一逆變換以將訊框轉換為時域(方塊 722)。最後，音訊編解碼器可在時域中產生音訊(方塊 724)。

F. 音訊丟失封包恢復

如本文中所揭示，可擴縮音訊編解碼器 110 可用於當已發生位元剝除時處置音訊。另外，可擴縮音訊編解碼器 110 亦可用於幫助丟失封包之恢復。為對抗封包丟失，一普通方法係藉由簡單地重複先前已接收之已經處理供輸出之音訊來填充由丟失之封包所致之間隙。雖然此方法減少由缺失之音訊間隙所致的失真，但其並不避免失真。舉例而言，對於超過百分之五之封包丟失率，由重複先前所發送之音訊所導致之人為產物變得顯著。

本發明之可擴縮音訊編解碼器 110 可藉由使一音訊訊框之高品質版本與低品質版本在連續封包中交錯來對抗封包丟失。由於其係可擴縮的，因而音訊編解碼器 110 可減少計算成本，乃因無需在不同品質下將音訊訊框編碼兩次。而是，簡單地藉由自己由可擴縮音訊編解碼器 110 所產生之高品質版本剝除位元來獲得低品質版本。

圖 8 展示在傳輸器 100A 處之所揭示之音訊編解碼器 110 如何可使音訊訊框之高品質版本與低品質版本交錯而不必將該音訊編碼兩次。在以下論述中，參考一「訊框」，該訊框可意指本文中所闡述之約 20 毫秒之一音訊區塊。然而，該交錯過程可適用於傳輸封包、變換係數區、位元之集合或類似物。另外，雖然該論述係參考 32k bps 之一最小恆定位元速率及 8 kbps 之一較低品質速率，但音訊編解碼器 110 所用之交錯技術可適用於其他位元速率。

通常，所揭示之音訊編解碼器 110 可使用 32 kbps 之一最小恆定位元速率來達成不降級之音訊品質。由於封包各自

具有20毫秒之音訊，因而此最小位元速率對應於每一封包640個位元。然而，該位元速率可偶爾降低至8 kbps(或160個位元每一封包)而具有可忽略之主觀失真。由於用640個位元編碼之封包看似遮蔽了由僅用160個位元編碼之彼等偶然封包所致的編碼失真，此係可能的。

在此過程中，傳輸器100A處之音訊編解碼器110在32 kbps之一最小位元速率之情形下，使用每一20毫秒封包640個位元來編碼一當前20毫秒之音訊訊框。為處理封包之潛在丟失，音訊編解碼器110針對每一未來訊框使用較低品質160個位元編碼N個數目之未來音訊訊框。然而音訊編解碼器110不必將訊框編碼兩次，而是藉由自較高品質版本剝除位元來形成較低品質之未來訊框。由於可引入某種傳輸音訊延遲，因而可編碼之可能低品質訊框之數目可受到限制，舉例而言，限制為 $N=4$ ，而無需向傳輸器100A添加額外之音訊延遲。

在此階段，傳輸器100A然後將高品質位元及低品質位元組合進一單個封包中，且將該封包發送至接收器100B。如在圖8中所展示，舉例而言，以32 kbps之最小恆定位元速率編碼一第一音訊訊框810a。亦以32 kbps之最小恆定位元速率編碼一第二音訊訊框810b，但亦在160個位元之低品質下編碼一第二音訊訊框810b。如本文中所提及，此較低品質版本814b實際上係藉由自己已經編碼之較高品質版本812b剝除位元來達成。考慮到所揭示之音訊編解碼器110將區之重要性進行排序，將較高品質版本812b位元剝除為

較低品質版本814b實際上可保留音訊之某一有用品質，即使係在此較低品質版本814b之情形下。

為產生一第一經編碼封包820a，將第一音訊訊框810a之高品質版本812a與第二音訊訊框810b之較低品質版本814b組合。此經編碼封包820a可併入上文所揭示的用於低頻率頻帶分割及高頻率頻帶分割之位元分配及重新排序技術，且此等技術可適用於較高及低品質版本812a/814b中之一者或兩者。因此，舉例而言，經編碼封包820a可包括一位元分割分配之一指示、針對該訊框之高品質版本812a之一低頻率頻帶之一第一頻譜包絡、按低頻率頻帶之經排序區重要性之第一變換係數、針對該訊框之高品質版本812a之一高頻率頻帶之一第二頻譜包絡及按高頻率頻帶之經排序區重要性之第二變換係數。然後，此可簡單地後跟下一訊框之低品質版本814b，而不慮及位元分配及類似物。另一選擇係，下一訊框之低品質版本814b可包括頻譜包絡及兩個頻帶頻率係數。

貫穿該編碼過程重複：較高品質編碼、位元剝除為一較低品質及與毗鄰音訊訊框組合。因此，舉例而言，產生一第二經編碼封包820b，其包括與第三音訊訊框810c之較低音訊版本814c(亦即，經位元剝除版本)組合之第二音訊訊框810b之高品質版本810b。

在接收端，接收器100B接收所傳輸之封包820。若一封包係好的(亦即，被接收到)，則接收器之音訊編解碼器110解碼表現當前20毫秒音訊之640個位元且將其提供出接收

器之揚聲器。舉例而言，在接收器110B處所接收到之第一經編碼封包820a可係好的，因而接收器110B解碼封包820a中之第一訊框810a之較高品質版本812a以產生一第一經解碼音訊訊框830a。所接收到之第二經編碼封包820b可亦係好的。因此，接收器110B解碼在此封包820b中之第二訊框810b之較高品質版本812b以產生一第二經解碼音訊訊框830b。

若一封包係壞的或遺失的，則接收器之音訊編解碼器110使用所接收之上一個好封包中所含有之當前訊框之較低品質版本(160個位元之經編碼資料)來恢復該遺失音訊。如所展示，舉例而言，第三經編碼封包820c在傳輸期間被丟失。並不如習用方式所做用另一訊框之音訊填充該間隙，在接收器100B處之音訊編解碼器110使用自先前經編碼封包820b(其係好的)獲得之遺失訊框810c之較低品質音訊版本814c。然後可使用此較低品質音訊來重新建構遺失之第三經編碼音訊訊框830c。以此方式，針對遺失封包820c之訊框，可使用實際遺失之音訊，雖然係以一較低品質。然而，預期此較低品質由於遮蔽而不會造成大量可察覺之失真。

已闡述將本發明之可擴縮音訊編解碼器與一會議端點或終端機一起使用。然而，所揭示之可擴縮音訊編解碼器可用於各種會議組件中，諸如端點、終端機、路由器、會議橋及其他。在此等組件中之每一者中，所揭示之可擴縮音訊編解碼器可節約頻寬、計算及記憶體資源。同樣，所揭

示之音訊編解碼器可在較低延時及較少人為產物方面改良音訊品質。

本發明之技術可實施於數位電子電路中或電腦硬體、韌體、軟體中或此等之組合中。用於實踐所揭示技術之設備可實施於有形地體現於一機器可讀儲存裝置中供一可程式化處理器執行之一電腦程式產品中，可藉由一可程式化處理器來執行所揭示技術之方法步驟，該可程式化處理器藉由操作輸入資料並產生輸出來執行一程式指令以執行所揭示技術之功能。合適之處理器包括(舉例而言)通用及專用微處理器兩者。一般而言，一處理器將自一唯讀記憶體及/或一隨機存取記憶體接收指令及資料。一般而言，一電腦將包括用於儲存資料檔案之一或多個大量儲存裝置；此等裝置包括：磁碟(例如，內部硬磁碟及可抽換式磁碟)；一磁光碟；及光碟。適合於有形地體現電腦程式指令及資料之儲存裝置包含所有形式之非揮發性記憶體，其包括：(舉例而言)半導體記憶體裝置(例如，EPROM、EEPROM及快閃記憶體裝置)；磁碟(例如，內部硬磁碟及可抽換式磁碟)；磁光碟；及CD-ROM磁碟。前述者中之任一者皆可由ASIC(專用積體電路)進行補充或併入於ASIC中。

前述對較佳及其他實施例之說明並不意欲限制或限定申請者所構想的本發明之概念之範疇或適用性。作為揭示本文中所含有之發明性概念之交換，申請者期望隨附申請專利範圍所提供之所有專利權利。因此，希望隨附申請專利範圍最大程度地包括歸屬於以下申請專利範圍之範疇或其

等效內容內之所有修改及變化形式。

【圖式簡單說明】

圖1A展示一變換編碼編解碼器之一編碼器。

圖1B展示一變換編碼編解碼器之一解碼器。

圖2A圖解說明用於使用根據本發明之編碼及解碼技術之一音訊處理裝置，諸如一會議終端機。

圖2B圖解說明具有用於使用根據本發明之編碼及解碼技術之一傳輸器及一接收器之一會議配置。

圖3係根據本發明之一音訊編碼技術之一流程圖。

圖4A係更詳細地展示編碼技術之一流程圖。

圖4B展示經取樣為若干訊框之一類比音訊信號。

圖4C展示經自時域中之一經取樣訊框變換之頻域中之一變換係數組。

圖4D展示用於將變換係數編碼於兩個頻率頻帶中之八個分配可用位元模式。

圖5A至圖5C展示基於重要性排序經編碼音訊中之區之實例。

圖6A係展示用於判定經編碼音訊中之區之重要性之一功率頻譜技術之一流程圖。

圖6B係展示用於判定經編碼音訊中之區之重要性之一感知技術之一流程圖。

圖7係更詳細地展示解碼技術之一流程圖。

圖8展示用於使用所揭示之可擴縮音訊編解碼器處理音訊封包丟失之一技術。

【主要元件符號說明】

10	編碼器
12	數位信號
14	輸出信號
20	變換
22	正規化處理程序
24	演算法
50	解碼器
52	輸入信號
54	輸出信號
60	網格解碼
62	解量化處理程序
64	逆變換
100	端點或終端機
100A	第一音訊處理裝置
100B	第二音訊處理裝置
102	麥克風
103	音訊相機
108	揚聲器
109	顯示器
110	音訊編解碼器
115	量化器
120	量化器
122	網路介面

124	網路介面
125	網路
160	處理器
162	記憶體
164	轉換器電子器件
170	編碼器
172	解碼器
200	編碼器
250	解碼器

七、申請專利範圍：

103年/月/日 修正頁(本)
劃線

1. 一種用於一處理裝置之可擴縮音訊處理方法，其包含：

將輸入音訊之若干訊框自一時域變換編碼成一頻域中之若干變換係數；

針對每一訊框，將一編碼位元速率之總可用位元分配為一第一位元分配及一第二位元分配，該第一位元分配經分配給該訊框之該等變換係數之一第一組，該第二位元分配經分配給該訊框之該等變換係數之一第二組；

針對每一訊框，將該等變換係數之該第一組及該第二組與對應之該第一位元分配及該第二位元分配封包化至一封包中；及

利用該處理裝置傳輸該等封包。

2. 如請求項 1 之方法，其中針對該輸入音訊逐訊框地進行分配該第一位元分配及該第二位元分配。

3. 如請求項 1 之方法，其中將該編碼位元速率之該等總可用位元分配為該第一位元分配及該第二位元分配包含：

計算該等變換係數之該第一組及該第二組之一能量比率；及

基於該所計算之比率來分配該訊框之該第一位元分配及該第二位元分配。

4. 如請求項 1 之方法，其中將該等變換係數之該第一組及該第二組中之每一者配置於頻率區中，且其中封包化該等變換係數之該第一組及該第二組中之每一變化係數包含：

判定該等頻率區之重要性；

基於該所判定之重要性排序該等頻率區；及

按照排序封包化該等頻率區。

5. 如請求項4之方法，其中判定重要性及排序該等頻率區包含：

判定該等頻率區中之每一者之一功率位準；及

自最大功率位準至最小功率位準排序該等區。

6. 如請求項5之方法，其中判定該功率位準進一步包含：
使用基於該等頻率區之間的頻譜距離之一固定函數來加權該等頻率區之該等功率位準。

7. 如請求項1之方法，其中封包化包含：封包化該第一位元分配及該第二位元分配之一指示。

8. 如請求項1之方法，其中封包化包含：封包化該等變換係數之該第一組及該第二組兩者之頻譜包絡。

9. 如請求項1之方法，其中封包化包含：在封包化一第一頻率頻帶及一第二頻率頻帶中之一較高頻率頻帶之前針對該等變換係數之該第一組及該第二組封包化該第一頻率頻帶及該第二頻率頻帶中之一較低頻率頻帶。

10. 如請求項1之方法，其中針對每一訊框變換編碼及封包化包含：

藉由以一第一位元速率變換編碼該訊框而產生該訊框之一第一版本；

藉由將該第一版本剝除為低於該第一位元速率之一第二位元速率而產生該訊框之一第二版本；及

將該訊框之該第一版本連同該等訊框中之一前一訊框之第二版本一起封包化至該封包中。

11. 如請求項1之方法，其中該等變換係數之該第一組係位於約0 kHz至約12 kHz之一第一頻率頻帶中，且其中該等變換係數之該第二組係位於約12 kHz至約22 kHz之一第二頻率頻帶中。
12. 如請求項1之方法，其中該等變換係數之該第一組係位於約0 Hz至約12,500 Hz之一第一頻率頻帶中，且其中該等變換係數之該第二組係位於約13 kHz至約22 kHz之一第二頻率頻帶中。
13. 如請求項1之方法，其中該第一位元分配及該第二位元分配總共為64 kbps之該編碼位元速率之該等總可用位元。
14. 如請求項1之方法，其中該等變換係數包含一調變重疊變換之若干係數。
15. 一種可程式化儲存裝置，其上儲存有程式指令用於致使一可程式化控制裝置執行根據請求項1至請求項14中之任一項之一可擴縮音訊處理方法。
16. 一種處理裝置，其包含：
 - 一網路介面；
 - 一處理器，其以通信方式耦合至該網路介面且獲得輸入音訊，該處理器經組態以：
 - 將一時域中之該輸入音訊之若干訊框變換編碼成一頻域中之若干變換係數；

針對每一訊框，將一編碼位元速率之總可用位元分配為一第一位元分配及一第二位元分配，該第一位元分配經分配給該訊框之該等變換係數之一第一組，該第二位元分配經分配給該訊框之該等變換係數之一第二組；

針對每一訊框，將該等變換係數之該第一組及該第二組與對應之該第一位元分配及該第二位元分配封包化至一封包中；及

藉助該網路介面傳輸該等封包。

17. 如請求項16之裝置，其中該處理裝置係選自由一音訊會議端點、一視訊會議端點、一音訊播放裝置、一個人音樂播放器、一電腦、一伺服器、一電信裝置、一蜂巢式電話及一個人數位助理組成之群組。

18. 如請求項16之裝置，其中該處理器經組態以針對該輸入音訊逐訊框地進行分配該第一位元分配及該第二位元分配。

19. 如請求項16之裝置，其中為將該編碼位元速率之該等總可用位元分配為該第一位元分配及該第二位元分配，該處理器經組態以：

計算該等變換係數之該第一組及該第二組之一能量比率；及

基於該所計算之比率來分配該訊框之該第一位元分配及該第二位元分配。

20. 如請求項16之裝置，其中將該等變換係數之該第一組及該第二組中之每一變換係數配置於頻率區中，且其中為

封包化該等變換係數之該第一組及該第二組中之每一變換係數，該處理器經組態以：

判定該等頻率區之重要性；

基於該所判定之重要性排序該等頻率區；及

按照排序封包化該等頻率區。

21. 如請求項20之裝置，其中為判定重要性及排序該等頻率區，該處理器經組態以：

判定該等頻率區中之每一者之一功率位準；及

自最大功率位準至最小功率位準排序該等區。

22. 如請求項21之裝置，其中為判定該功率位準，該處理器經組態以使用基於該等頻率區之間的頻譜距離之一固定函數來加權該等頻率區之該等功率位準。

23. 如請求項16之裝置，其中為進行封包化，該處理器經組態以封包化該第一位元分配及該第二位元分配之一指示。

24. 如請求項16之裝置，其中為進行封包化，該處理器經組態以封包化該等變換係數之該第一組及該第二組兩者之頻譜包絡。

25. 如請求項16之裝置，其中為進行封包化，該處理器經組態以在封包化一第一頻率頻帶及一第二頻率頻帶中之一較高頻率頻帶之前針對該等變換係數之該第一組及該第二組封包化該第一頻率頻帶及該第二頻率頻帶中之一較低頻率頻帶。

26. 如請求項16之裝置，其中為針對每一訊框進行變換編碼

及封包化，該處理器經組態以：

藉由以一第一位元速率變換編碼該訊框而產生該訊框之一第一版本；

藉由將該第一版本剝除為低於該第一位元速率之一第二位元速率而產生該訊框之一第二版本；及

將該訊框之該第一版本連同該等訊框中之一前一訊框之第二版本一起封包化至該封包中。

27. 如請求項 16 之裝置，其中該等變換係數之該第一組係位於約 0 kHz 至約 12 kHz 之一第一頻率頻帶中，且其中該等變換係數之該第二組係位於約 12 kHz 至約 22 kHz 之一第二頻率頻帶中。
28. 如請求項 16 之裝置，其中該等變換係數之該第一組係位於約 0 Hz 至約 12,500 Hz 之一第一頻率頻帶中，且其中該等變換係數之該第二組係位於約 13 kHz 至約 22 kHz 之一第二頻率頻帶中。
29. 如請求項 16 之裝置，其中該第一位元分配及該第二位元分配總共為 64 kbps 之該編碼位元速率之該總可用位元。
30. 如請求項 16 之裝置，其中該等變換係數包含一調變重疊變換之若干係數。
31. 一種用於一處理裝置之音訊處理方法，其包含：

接收輸入音訊訊框之若干封包，該等封包中之每一者具有在一頻域中之若干變換係數；

判定該等封包中之每一者中之該等訊框之第一位元分配及第二位元分配，該等第一位元分配中之每一者分配

給該封包中之該訊框之該等變換係數之一第一組，該等第二位元分配中之每一者分配給該封包中之該訊框之該等變換係數之一第二組；

將該等封包中之該等訊框中之每一者之該等變換係數之該第一組及該第二組逆變換編碼成輸出音訊；

依據該等封包中之該等訊框中之每一者之該等第一位元分配及該等第二位元分配判定是否有位元遺失；及

將音訊填充至經判定為遺失的該等位元中之任一者中。

32. 如請求項31之方法，其中接收該等封包包含接收該等訊框之該等變換係數之該第一組及該第二組中之每一者之一頻譜包絡，且其中填充音訊包含利用該頻譜包括按比例調整一音訊信號。
33. 如請求項31之方法，其中基於該等變換係數之該第一組及該第二組之一所計算之能量比率來分配該訊框之該等第一位元分配及該等第二位元分配。
34. 如請求項31之方法，其中該等變換係數之該第一組及該第二組中之每一變換係數係配置於基於頻率區之一經判定之重要性而經排序及封包化之該等頻率區中。
35. 如請求項34之方法，其中該等頻率區之該經判定之重要性之該排序係基於該等頻率區之一最大功率位準至一最小功率位準。
36. 如請求項35之方法，其中使用基於該等頻率區之間的頻譜距離之一固定函數來加權該等功率位準。

37. 如請求項31之方法，其中判定該等封包中之每一者中之該等訊框之第一位元分配及第二位元分配包含自該等封包獲得該等第一位元分配及該等第二位元分配之一指示。
38. 如請求項31之方法，其中該等封包之每一者具有該等變換係數之該第一組及該第二組兩者之頻譜包絡。
39. 如請求項31之方法，其中在一第一頻率頻帶及一第二頻率頻帶中之一較高頻率頻帶之前，該等封包之每一者針對該等變換係數之該第一組及該第二組具有該第一頻率頻帶及該第二頻率頻帶中之一較低頻率頻帶。
40. 如請求項31之方法，其中針對每一訊框，該等封包之每一者包含以一第一位元速率變換編碼之該訊框之一第一版本且包含由低於該第一位元速率之一第二版本之先前訊框之一第一版本剝除之該先前訊框之一第二版本。
41. 如請求項31之方法，其中該等變換係數之該第一組係位於約0 kHz至約12 kHz之一第一頻率頻帶中，且其中該等變換係數之該第二組係位於約12 kHz至約22 kHz之一第二頻率頻帶中。
42. 如請求項31之方法，其中該等變換係數之該第一組係位於約0 Hz至約12,500 Hz之一第一頻率頻帶中，且其中該等變換係數之該第二組係位於約13 kHz至約22 kHz之一第二頻率頻帶中。
43. 如請求項31之方法，其中該等第一位元分配及該等第二位元分配總共為64 kbps之該編碼位元速率之該等總可用

位元。

44. 如請求項31之方法，其中該等變換係數包含一調變重疊變換之若干係數。

45. 一種用於一處理裝置之音訊處理方法，其包含：

藉由以一第一位元速率變換編碼連續輸入音訊訊框中之每一者來產生該等連續訊框之第一版本；

藉由將該等第一版本中之每一者剝除為低於該第一位元速率之一第二位元速率而產生該等連續訊框中之每一者之第二版本；

將該等連續訊框之該等第一版本中之每一者連同該等連續訊框中之前一訊框之該第二版本封包化至若干封包中；及

利用該處理裝置傳輸該等封包。

46. 一種用於一處理裝置之音訊處理方法，其包含：

接收連續輸入音訊訊框之封包，該等封包中之每一者具有該等連續訊框中之一者之一第一版本且具有該等連續訊框中之前一訊框之一第二版本，該等第一版本中之每一者包括以一第一位元速率變換編碼之該一個訊框，該等第二版本中之每一者包括經剝除為低於該第一位元速率之一第二位元速率之該前一訊框之該第一版本；

解碼該等封包中之每一者；

偵測所接收之該等封包中之一者之一封包錯誤；

藉由使用該一個封包之一遺失訊框之該第二版本而自所接收之該等封包之前一封包重現該一個封包之該遺失

修正頁(本)
劃線
103年1月21日

訊框；及

利用該等訊框之該第一版本及該經重現之遺失訊框而
產生輸出音訊。

八、圖式：

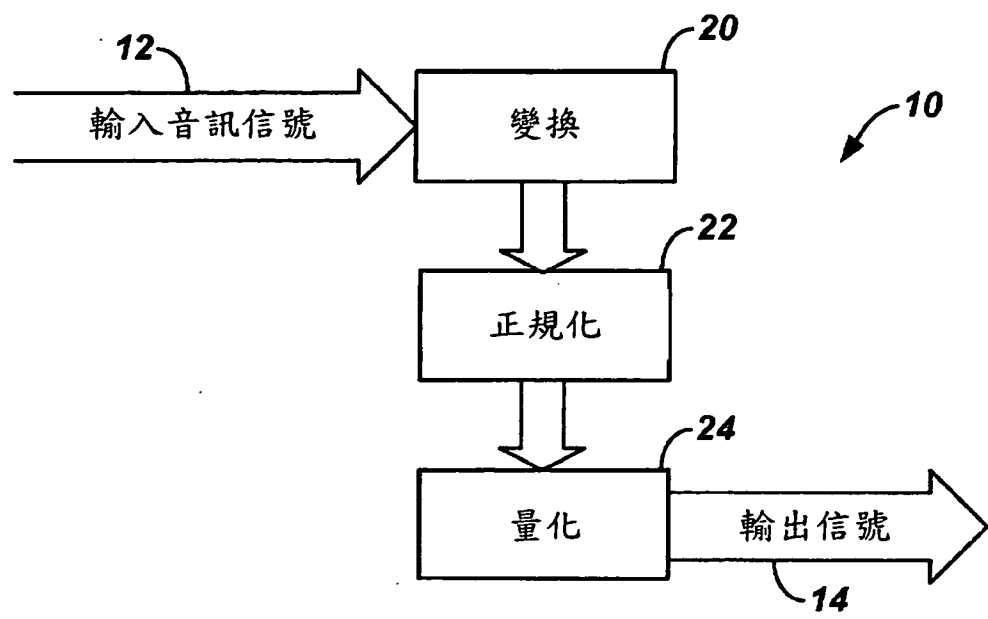


圖 1A

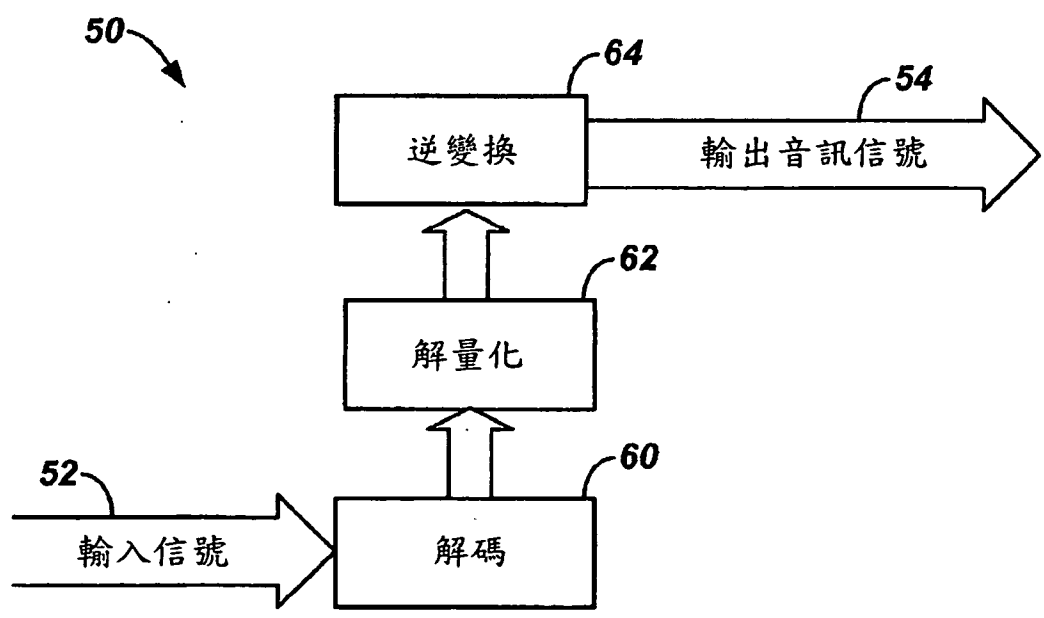


圖 1B

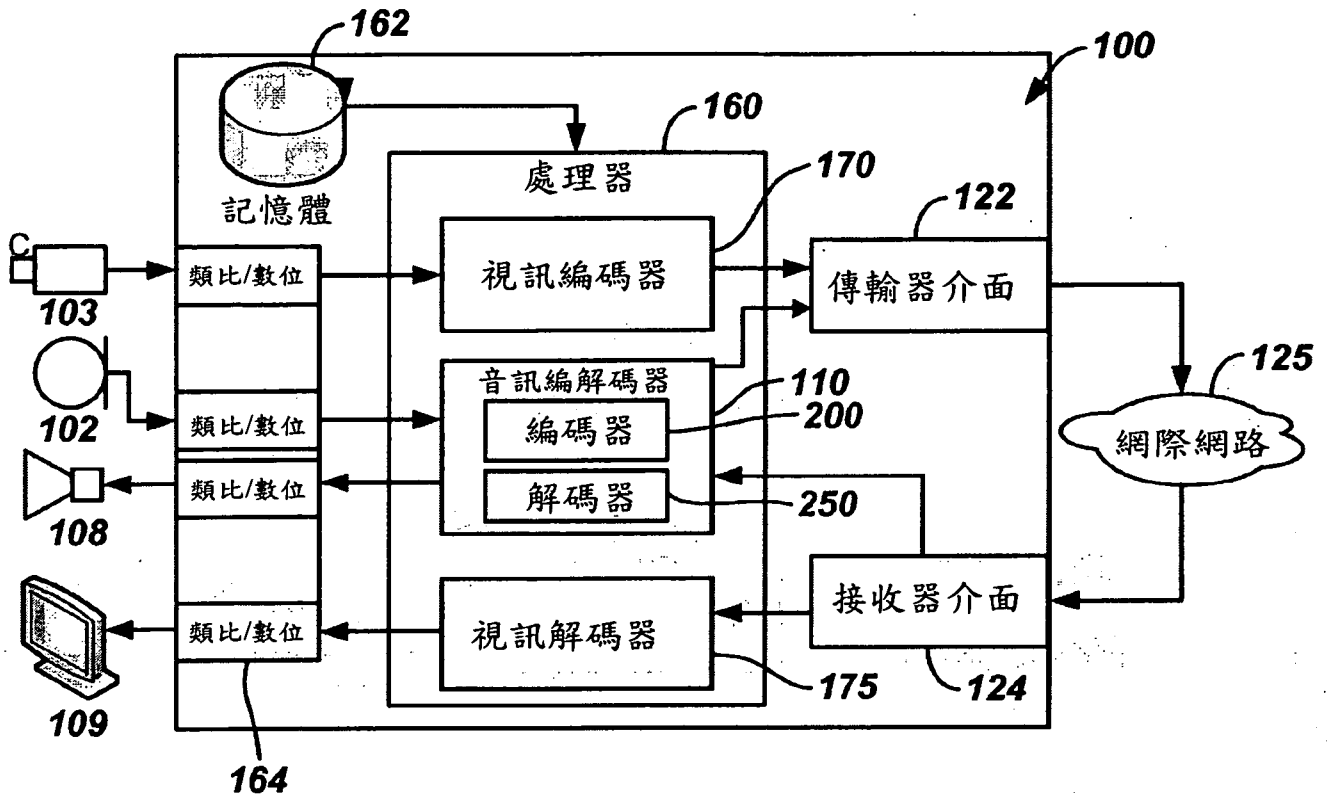


圖 2A

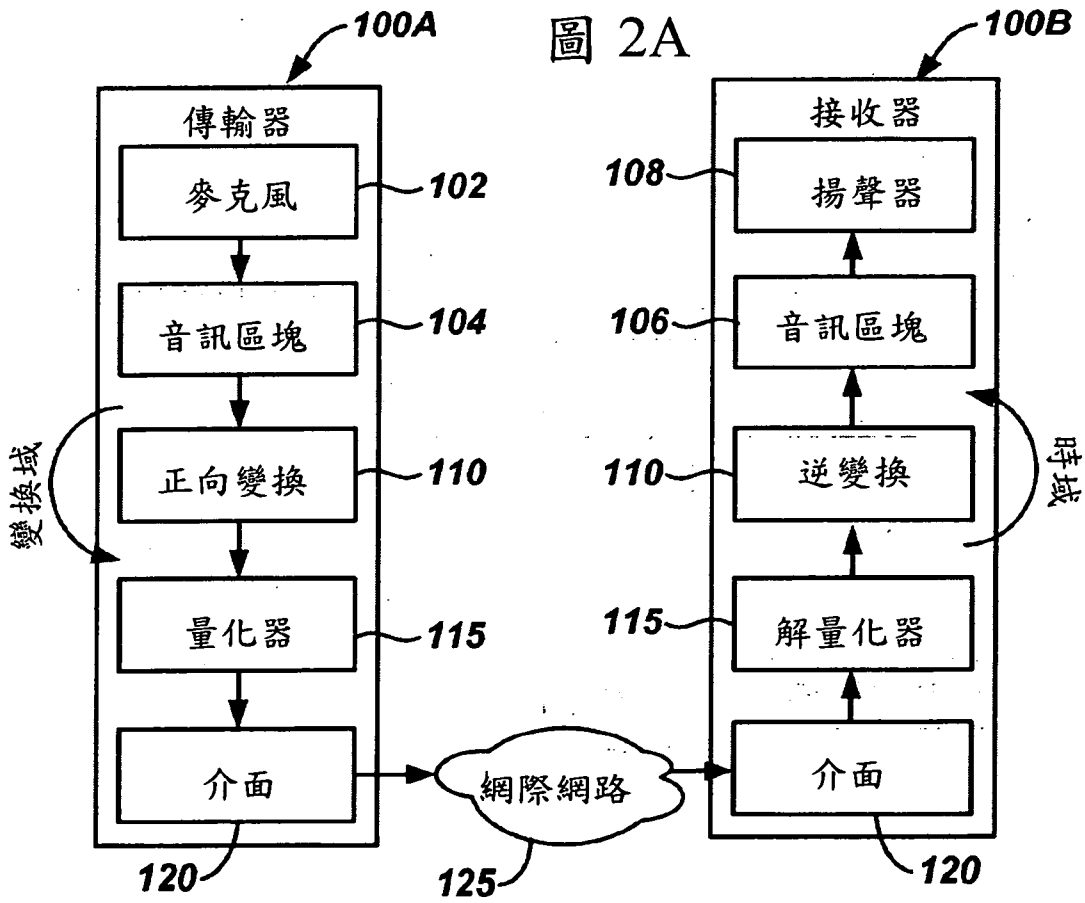


圖 2B

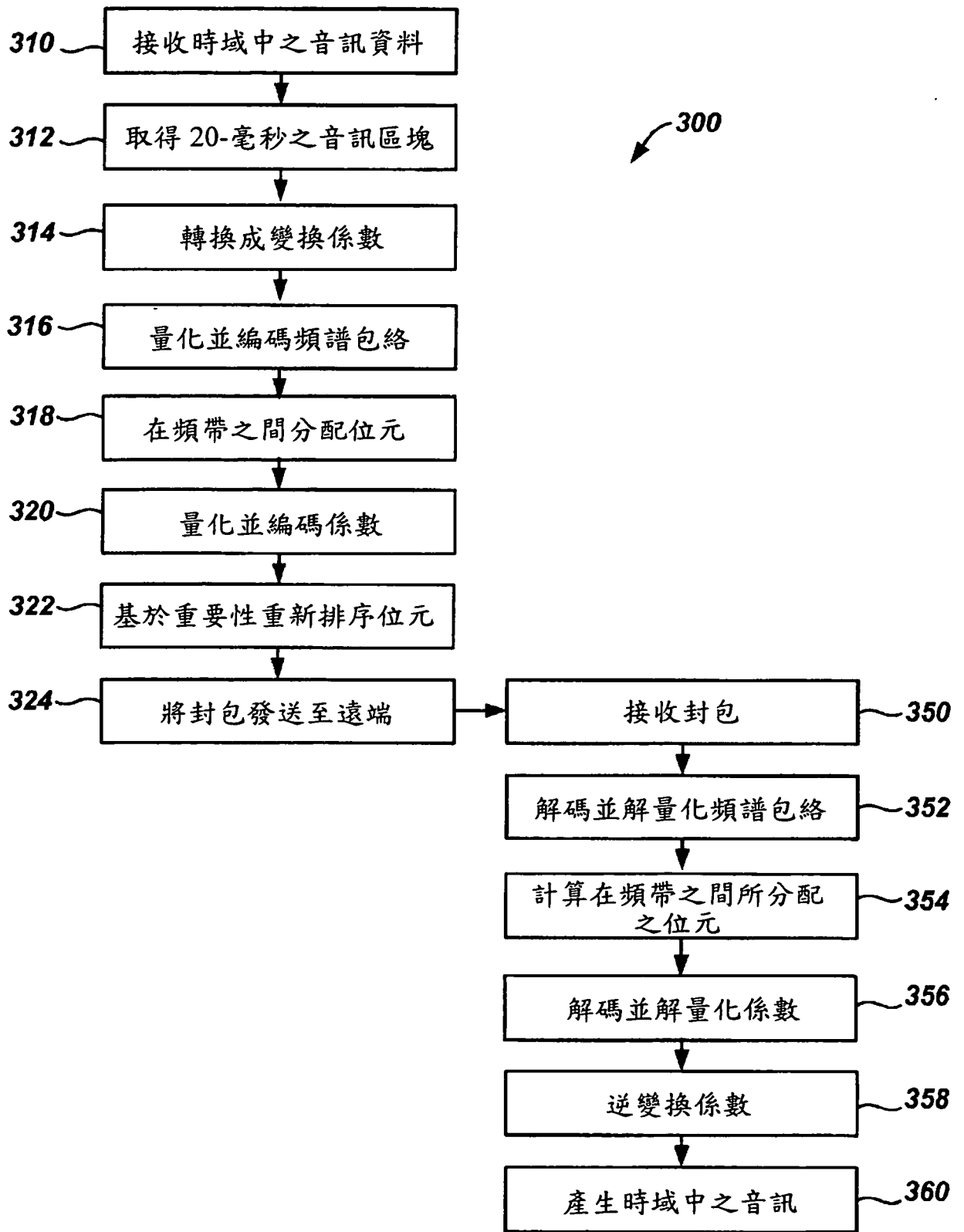


圖 3

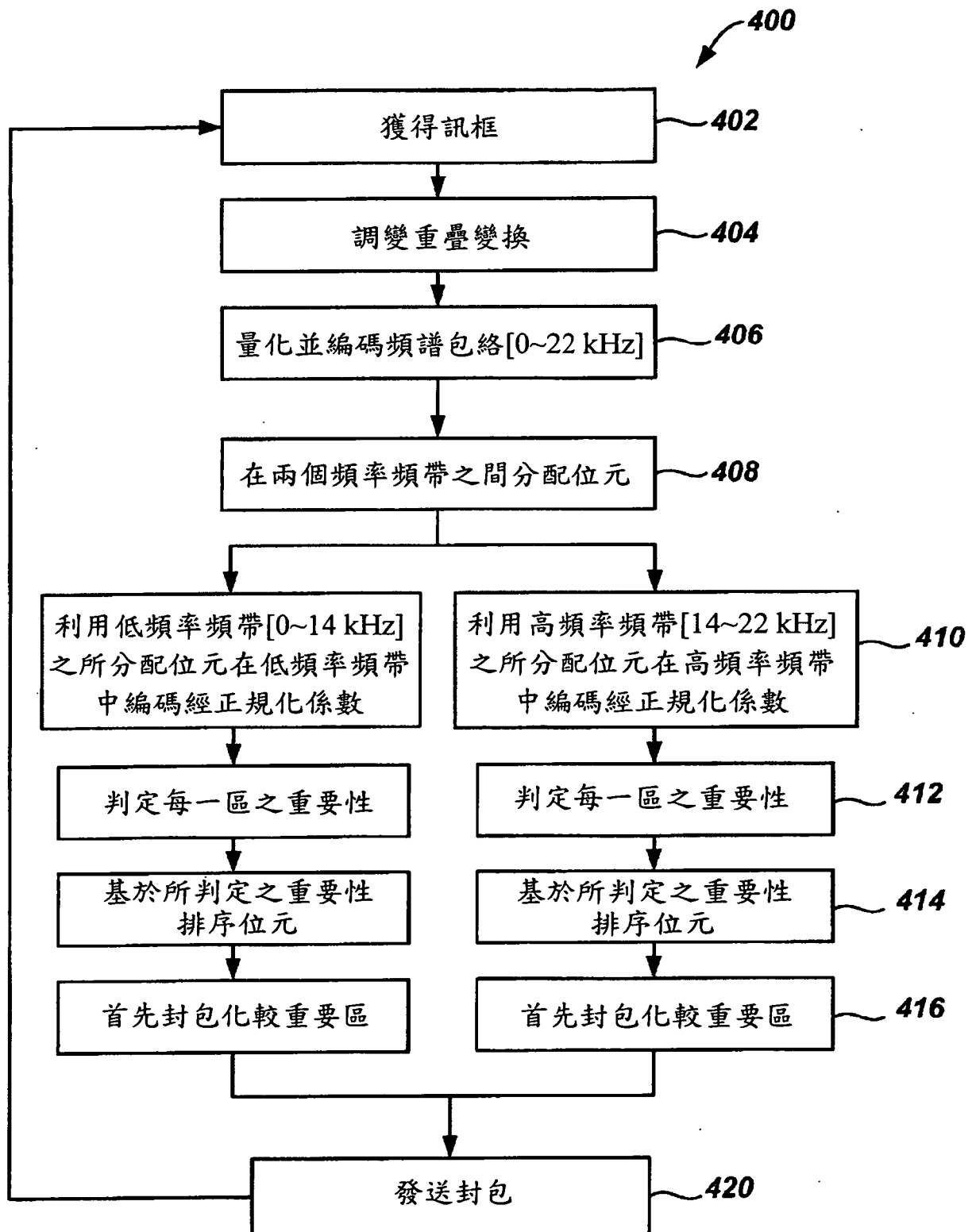


圖 4A

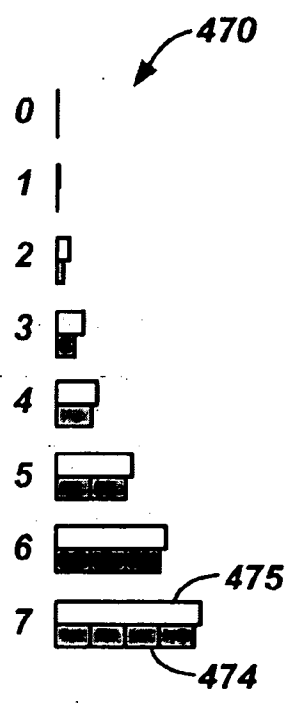
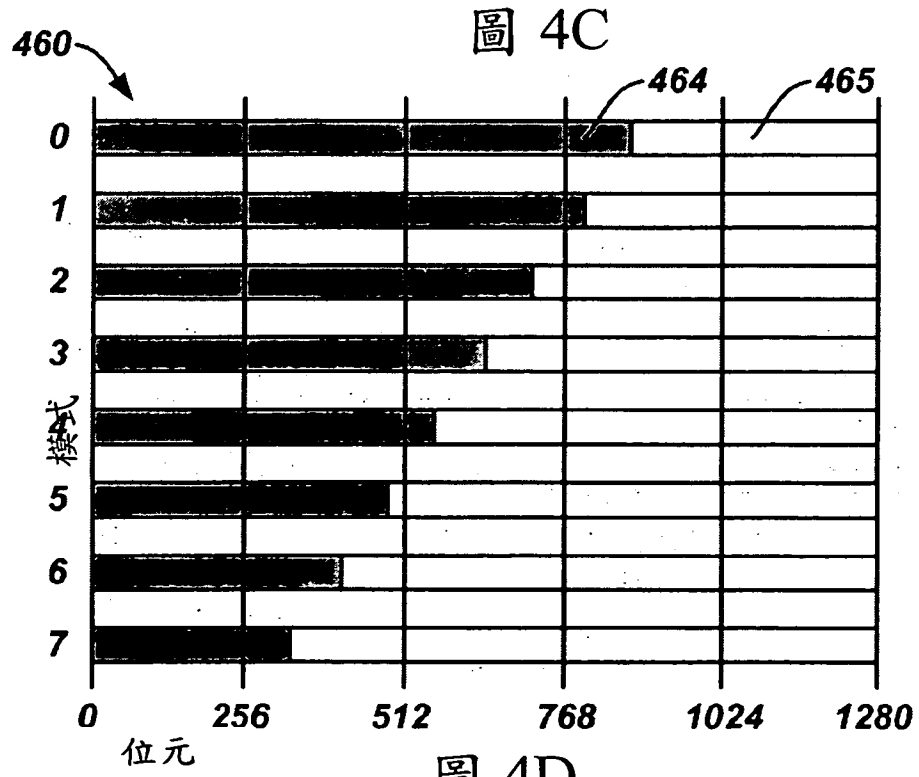
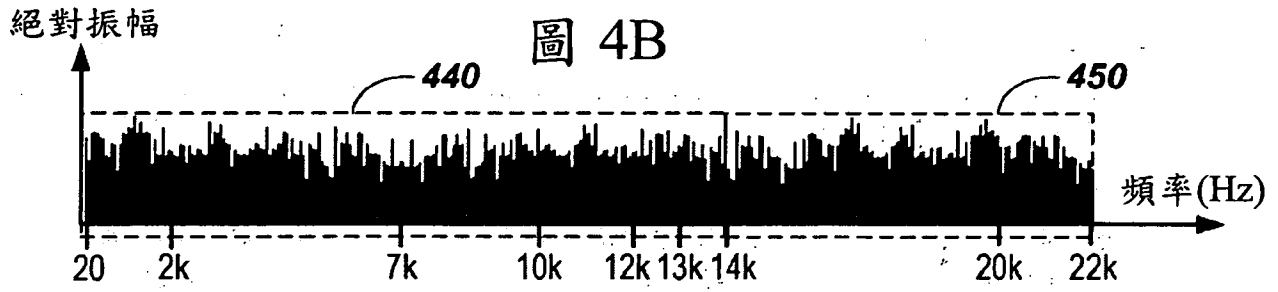
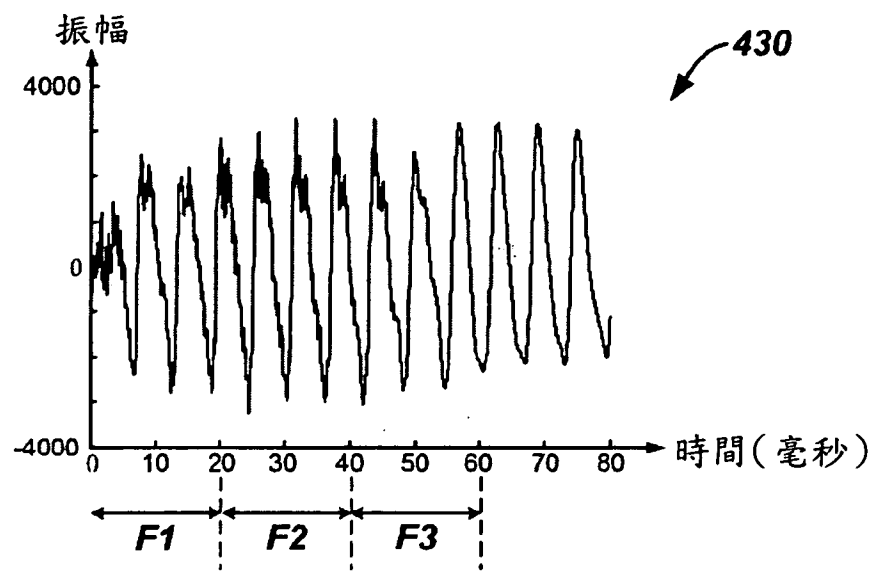


圖 4D

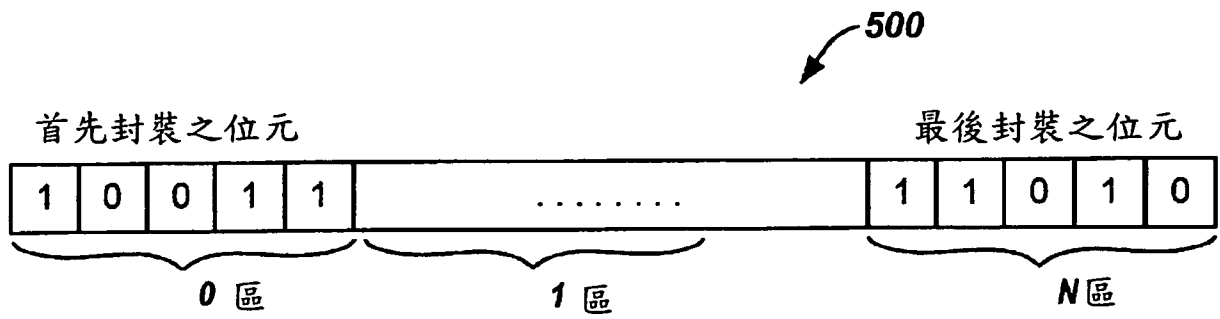


圖 5A

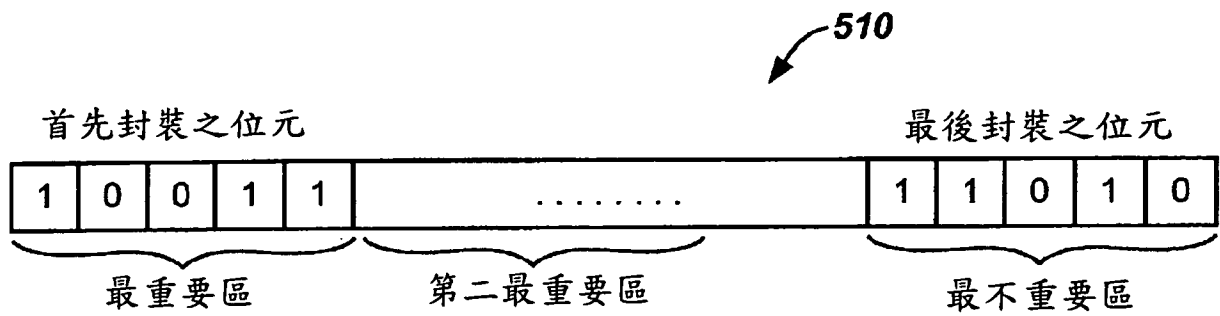


圖 5B

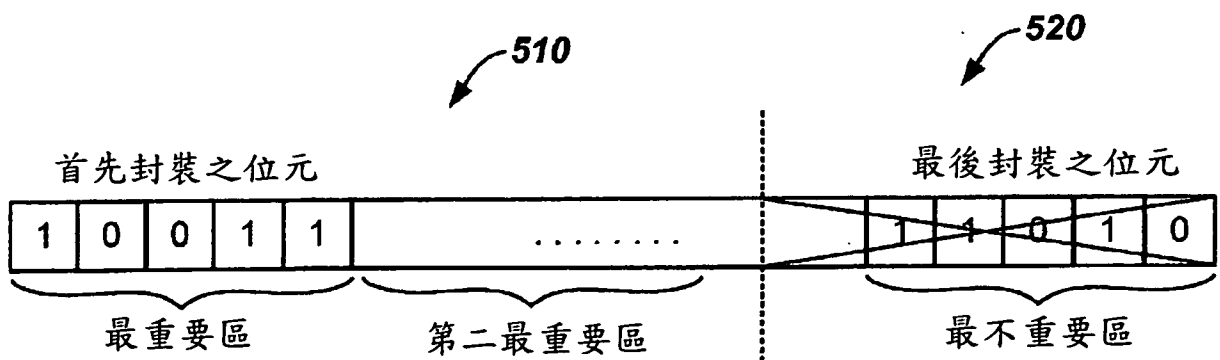


圖 5C

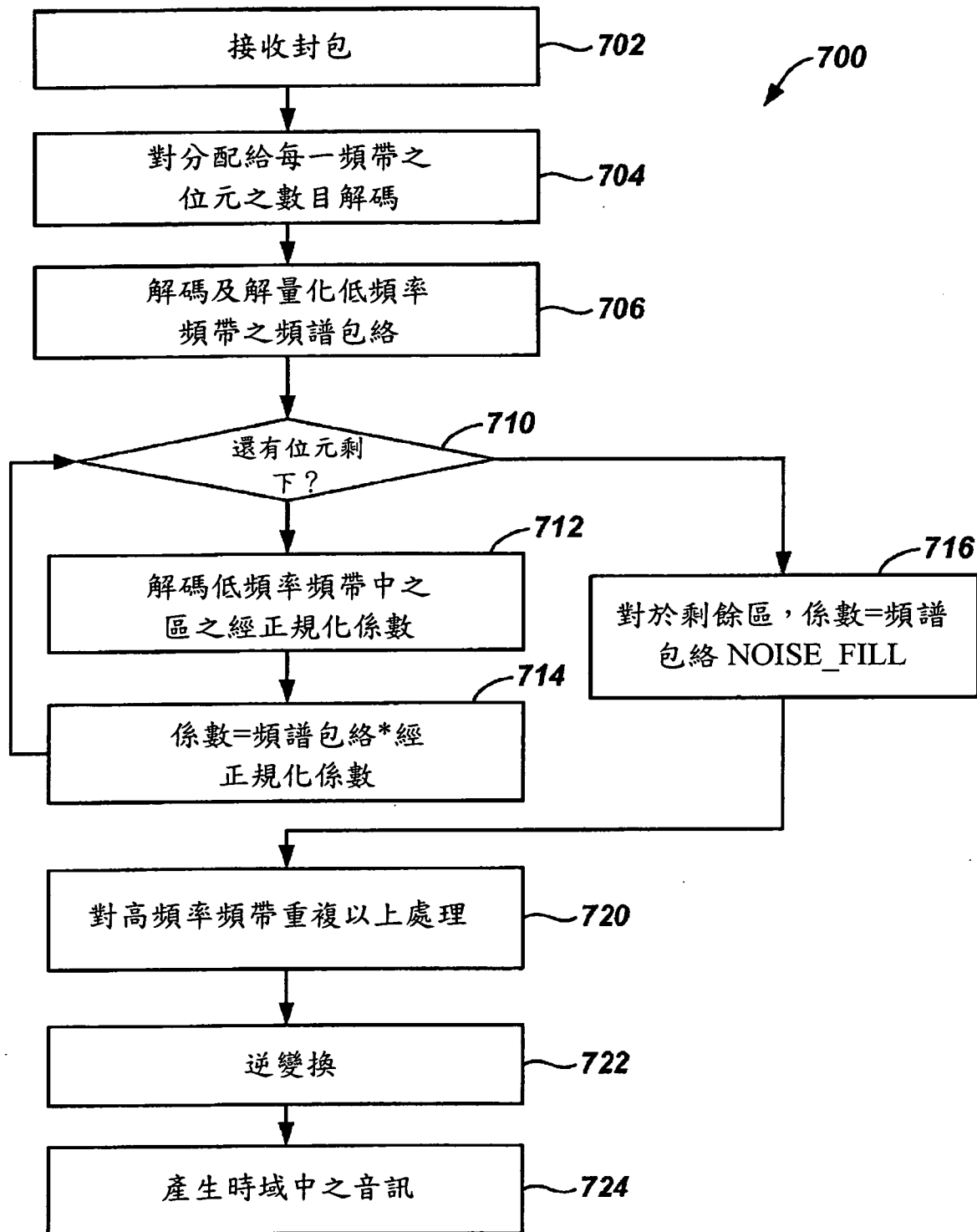


圖 7

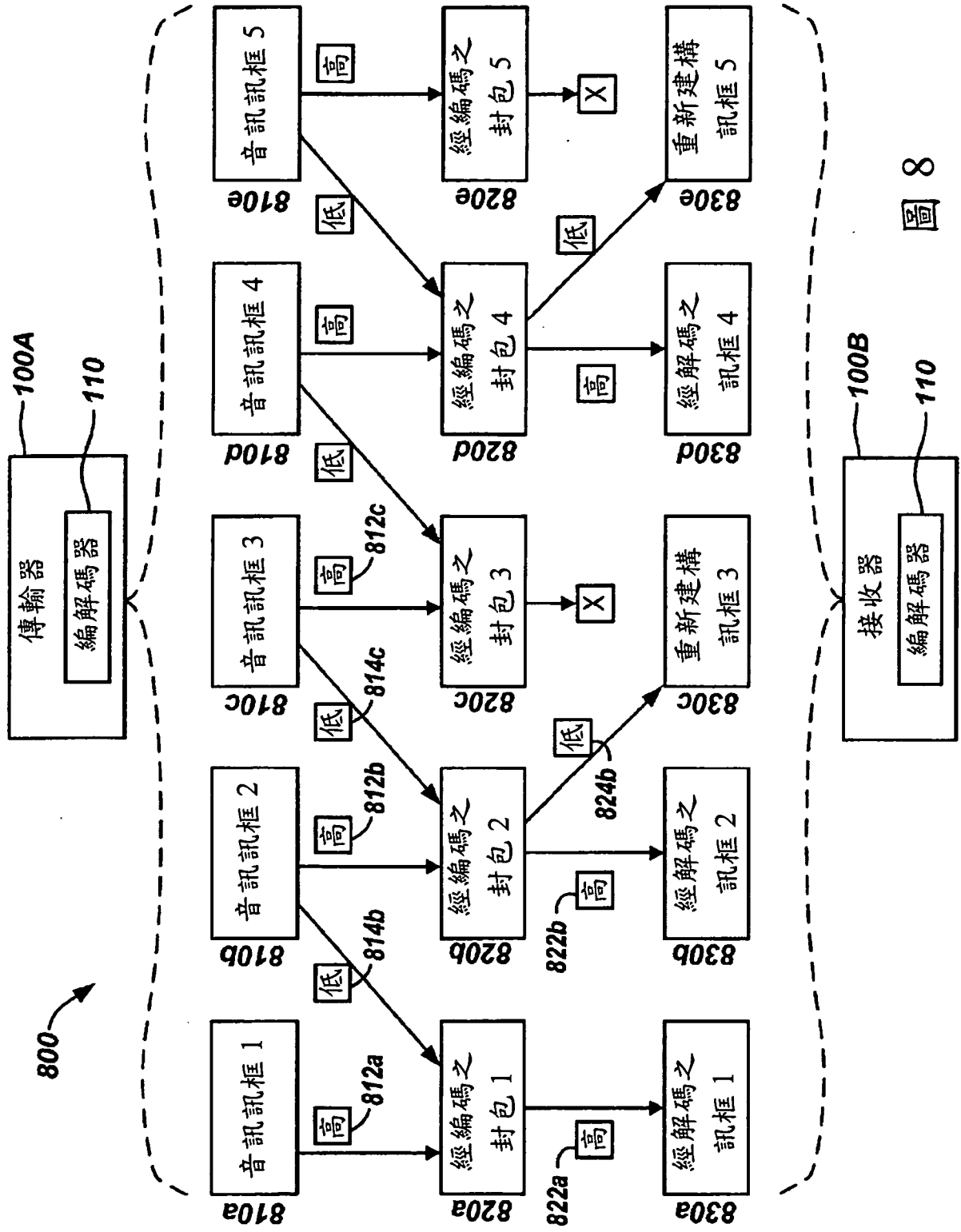


圖 8