



(21)申請案號：100103234

(22)申請日：中華民國 100 (2011) 年 01 月 28 日

(51)Int. Cl. : **G10L19/00 (2013.01)**

(30)優先權：2010/01/29 美國 12/696,788

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

(72)發明人：朱 彼得 CHU, PETER (US) ; 涂 赫民 TU, ZHEMIN (US)

(74)代理人：陳長文

(56)參考文獻：

US 6029126 US 6973184B1

US 2002/0007273A1 US 2002/0089602A1

審查人員：黃衍勳

申請專利範圍項數：27 項 圖式數：11 共 0 頁

(54)名稱

藉由變換內插之音訊封包損失隱蔽

AUDIO PACKET LOSS CONCEALMENT BY TRANSFORM INTERPOLATION

(57)摘要

本發明提供一種用於一音訊或視訊會議之音訊處理，在音訊處理中，一終端機接收具有用於重新建構已經歷變換編碼之一音訊信號的變換係數之音訊封包。當接收該等封包時，該終端機判定是否存在任意丟失的封包並用來自先前及隨後的良好訊框之變換係數作為內插值。為以內插值取代丟失係數，該終端機用一第一權重來加權來自該先前的良好訊框之第一係數，用一第二權重來加權來自該隨後的良好訊框之第二係數，且將此等經加權的係數加總在一起以插入於該等丟失封包中。該等權重可基於音訊頻率及/或所涉及的丟失封包之數目。自此內插，該終端機藉由逆變換該等係數產生一輸出音訊信號。

In audio processing for an audio or video conference, a terminal receives audio packets having transform coefficients for reconstructing an audio signal that has undergone transform coding. When receiving the packets, the terminal determines whether there are any missing packets and interpolates transform coefficients from the preceding and following good frames. To interpolate the missing coefficients, the terminal weights first coefficients from the preceding good frame with a first weighting, weights second coefficients from the following good frame with a second weighting, and sums these weighted coefficients together for insertion into the missing packets. The weightings can be based on the audio frequency and/or the number of missing packets involved. From this interpolation, the terminal produces an output audio signal by inverse transforming the coefficients.

(無元件符號說明)

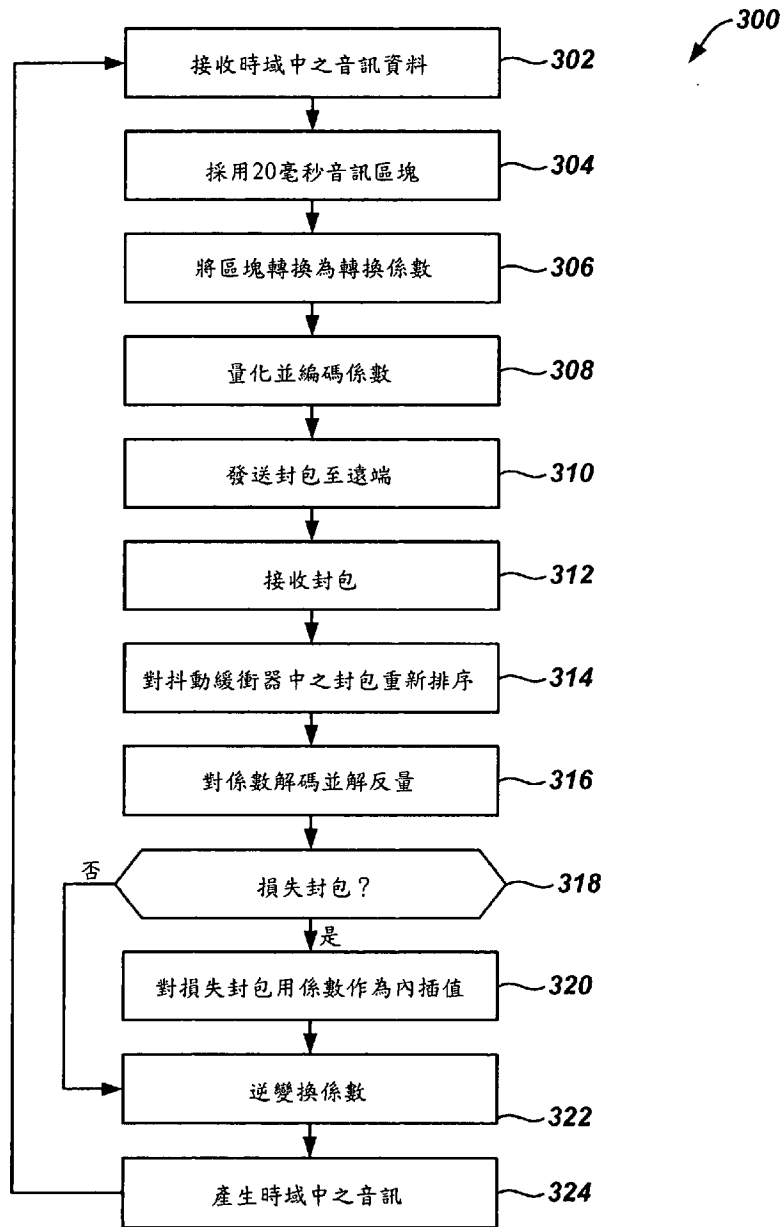


圖 4

發明專利說明書

公告本

(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)

※申請案號：100103234

※申請日：100.1.28

※IPC 分類：G10L 19/04 ~~2006-01~~
(2013.01)

一、發明名稱：(中文/英文)

藉由變換內插之音訊封包損失隱蔽

AUDIO PACKET LOSS CONCEALMENT BY TRANSFORM
INTERPOLATION

二、中文發明摘要：

本發明提供一種用於一音訊或視訊會議之音訊處理，在音訊處理中，一終端機接收具有用於重新建構已經歷變換編碼之一音訊信號的變換係數之音訊封包。當接收該等封包時，該終端機判定是否存在任意丟失的封包並用來自先前及隨後的良好訊框之變換係數作為內插值。為以內插值取代丟失係數，該終端機用一第一權重來加權來自該先前的良好訊框之第一係數，用一第二權重來加權來自該隨後的良好訊框之第二係數，且將此等經加權的係數加總在一起以插入於該等丟失封包中。該等權重可基於音訊頻率及/或所涉及的丟失封包之數目。自此內插，該終端機藉由逆變換該等係數產生一輸出音訊信號。

三、英文發明摘要：

In audio processing for an audio or video conference, a terminal receives audio packets having transform coefficients for reconstructing an audio signal that has undergone transform coding. When receiving the packets, the terminal determines whether there are any missing packets and interpolates transform coefficients from the preceding and following good frames. To interpolate the missing coefficients, the terminal weights first coefficients from the preceding good frame with a first weighting, weights second coefficients from the following good frame with a second weighting, and sums these weighted coefficients together for insertion into the missing packets. The weightings can be based on the audio frequency and/or the number of missing packets involved. From this interpolation, the terminal produces an output audio signal by inverse transforming the coefficients.

四、指定代表圖：

(一)本案指定代表圖為：第(4)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】

許多類型的系統使用音訊信號處理以產生音訊信號或自此等信號重製聲音。通常，信號處理將音訊信號轉換為數位資料並編碼該資料以經由一網路傳輸。接著，信號處理解碼該資料並將其轉換回類比信號以重製為聲波。

存在用於編碼或解碼音訊信號之各種方式。(編碼並解碼一信號之一處理器或一處理模組一般係稱為一編碼解碼器)。舉例而言，用於音訊及視訊會議之音訊處理使用音訊編碼解碼器以壓縮高保真度音訊輸入，使得用於傳輸之一所得信號保留最佳品質但需要最少數目的位元。依此方式，具有該音訊編碼解碼器之會議裝備需要較小的儲存容量，且該裝備用於傳輸該音訊信號之通信頻道需要較小的頻寬。

名為「7 kHz audio-coding within 64 kbit/s」之ITU-T(國際電信聯盟電信標準領域)提議G.722 (1988)描述64千位元/秒內7 kHz音訊編碼之一方法，該提議藉此以引用方式併入。ISDN(整合服務數位網路)線具有以64千位元/秒傳輸資料之能力。此方法基本上將通過使用一ISDN線之一電話網路之音訊頻寬自3 kHz增加至7 kHz。所感知的音訊品質得以改良。雖然此方法使高品質音訊可通過現有的電話網路得到，但是其通常需要來自一電話公司之ISDN服務，該ISDN服務比一常規的窄頻帶電話服務更昂貴。

建議用於電信中之一最近的方法係名為「Low-

complexity coding at 24 and 32 kbit/s for hands-free operation in system with low frame loss」之ITU-T提議 G.722.1 (2005)，該提議藉此以引用方式併入本文中。此提議描述提供以比該 G.722 低很多之 24 千位元/秒或 32 千位元/秒之一位元速率操作之 50 Hz 至 7 kHz 之一音訊頻寬之一數位寬頻編碼器演算法。在此資料速率下，具有使用常規的類比電話線之一常規的數據機之一電話可傳輸寬頻音訊信號。因此，只要兩端處的電話機可執行如 G.722.1 中所描述的編碼/解碼，大部分現有的電話網路可支援寬頻對話。

一些常用音訊編碼解碼器使用變換編碼技術來編碼並解碼經由一網路傳輸之音訊資料。舉例而言，ITU-T 提議 G.719 (Polycom® Siren™22) 以及 G.722.1.C (Polycom® Siren14™) (兩者以引用方式併入本文中) 使用熟知的調變重疊變換 (MLT) 編碼來壓縮該音訊以用於傳輸。如已知，該調變重疊變換 (MLT) 係用於各種類型的信號之變換編碼之一餘弦調變過濾器組的一種形式。

一般而言，一重疊變換採用長度為 L 之一音訊區塊並將該區塊變換為 M 個係數，其中條件為 $L > M$ 。為使此產生作用， $L - M$ 個樣本之連續區塊之間必須存在一重疊，使得可使用經變換係數之連續區塊得到一合成信號。

對於一調變重疊變換 (MLT)，該音訊區塊之該長度 L 等於係數之數目 M ，因此該重疊係 M 。因此，用於直接 (分析) 變換之 MLT 基函數係由下式給出：

$$p_a(n,k) = h_a(n) \sqrt{\frac{2}{M}} \cos \left[\left(n + \frac{M+1}{2} \right) \left(k + \frac{1}{2} \right) \frac{\pi}{M} \right] \dots \dots \dots (1)$$

類似地，用於逆(合成)變換之MLT基函數係由下式給出：

$$p_s(n,k) = h_s(n) \sqrt{\frac{2}{M}} \cos \left[\left(n + \frac{M+1}{2} \right) \left(k + \frac{1}{2} \right) \frac{\pi}{M} \right] \dots \dots \dots (2)$$

在此等方程式中，M係該區塊大小，頻率指數 k 自0變化至 $M-1$ ，且時間指數 n 自0變化至 $2M-1$ 。最後， $h_a(n) = h_s(n) = -\sin \left[\left(n + \frac{1}{2} \right) \frac{\pi}{2M} \right]$ 係所使用的理想重新建構窗。

自此等基函數如下判定MLT係數。直接變換矩陣 P_a 係第 n 列及第 k 行中之項目為 $p_a(n,k)$ 之一矩陣。類似地，逆變換矩陣 P_s 係具有項目 $p_s(n,k)$ 之一矩陣。對於一輸入信號 $x(n)$ 之 $2M$ 個輸入樣本之一區塊 x ，藉由 $\bar{X} = P_a^T x$ 計算變換係數之其對應的向量 \bar{X} 。繼而，對於經處理的變換係數之一向量 \bar{Y} ，經重新建構的 $2M$ 個樣本向量 y 係由 $y = P_s \bar{Y}$ 給出。最後，用 M 樣本重疊將該等經重新建構的 y 向量彼此疊加以產生用於輸出之經重新建構的信號 $y(n)$ 。

圖1展示一典型的音訊或視訊會議配置，其中用作為一傳輸器之一第一終端機10A將經壓縮的音訊信號發送至在此背景中用作為一接收器之一第二終端機10B。該傳輸器10A與該接收器10B兩者具有執行變換編碼(諸如G.722.1.C(Polycom® Siren14™)或G.719(Polycom® Siren™22)中所使用之變換編碼)之一音訊編碼解碼器16。

該傳輸器10A處之一麥克風12通常跨越20毫秒將源音訊及電子取樣源音訊擷取於音訊區塊14中。此刻，該音訊編

碼解碼器 16 之變換將該等音訊區塊 14 轉換為若干頻域變換係數組。每一變換係數具有一量值且可為正或負。接著使用此項技術中已知的技術對此等係數量化(18)、編碼，並且經由一網路 20(諸如網際網路)發送至該接收器。

在該接收器 10B 處，一逆程序解碼並解量化(19)該等經編碼的係數。最後，該接收器 10B 處之該音訊編碼解碼器 16 對該等係數執行一逆變換以將其等轉換回時域中，以產生在該接收器之揚聲器 13 處最終重放之輸出音訊區塊 14。

音訊封包損失係經由該等網路(諸如網際網路)進行視訊會議及音訊會議中之一常見問題。如已知，音訊封包代表小的音訊片段。當該傳輸器 10A 經由該網際網路 20 將該等變換係數之封包發送至該接收器 10B 時，一些封包可在傳輸期間損失。在產生輸出音訊之後，該等損失封包將在該揚聲器 13 輸出之聲音中產生靜音間隙。因此，該接收器 10B 較佳地用已由已自該傳輸器 10A 接收的該等封包合成之一些形式的音訊填充此等間隙。

如圖 1 中所示，該接收器 10B 具有偵測損失封包之一損失封包偵測模組 15。接著，當輸出音訊時，一音訊轉發器 17 填充由此等損失封包導致的間隙。該音訊轉發器 17 所使用之一現有技術藉由於時域中持續轉發在該封包損失之前發送的最近音訊片段來簡單地填充此等間隙。雖然轉發音訊以填充間隙之現有技術係有效，但是其可於所得音訊中產生蜂鳴及機械假訊，且使用者易於發現此等假訊令人反感。此外，若損失大於 5% 的封包，則當前技術產生逐漸

減少之易聽懂的音訊。

因此，需要一種以產生較佳的音訊品質且避免蜂鳴及機械假訊之方式處置在經由網際網路進行會議時之損失音訊封包的技術。

【發明內容】

本文所揭示的音訊處理技術可用於音訊或視訊會議。在該等處理技術中，一終端機接收具有用於重新建構已經歷變換編碼之一音訊信號的變換係數之音訊封包。當接收該等封包時，該終端機判定是否存在任意丟失的封包並用來自先前及隨後的良好訊框之變換係數作為內插值以用於插入作為該等丟失封包之係數。為以內插值取代該等丟失係數，舉例而言，該終端機用一第一權重來加權來自該先前良好的訊框之第一係數，用一第二權重來加權來自該隨後良好的訊框之第二係數，且將此等經加權的係數加總在一起以插入於該等丟失封包中。該等權重可基於音訊頻率及/或所涉及之丟失封包的數目。自此內插，該終端機藉由逆變換該等係數產生一輸出音訊信號。

前述內容並非意欲概述本發明之每一個可能的實施例或每一態樣。

【實施方式】

圖2A展示一音訊處理配置，其中用作為一傳輸器之一第一終端機100A將經壓縮的音訊信號發送至在此背景中用作為一接收器之一第二終端機100B。該傳輸器100A與該接收器100B兩者具有執行變換編碼(諸如G.722.1.C(Polycom®

Siren14™)或G.719(Polycom®Siren™22)中所使用之變換編碼)之一音訊編碼解碼器110。對於當前討論，該傳輸器100A及該接收器100B可為一音訊或視訊會議中之端點，但是其等可為其他類型的音訊裝置。

在操作期間，該傳輸器100A處之一麥克風102擷取源音訊，且源音訊之電子取樣區塊或訊框通常跨越20毫秒。(討論同時參考展示根據本發明之一損失封包處置技術300之圖3中之流程圖)。此刻，音訊編碼解碼器110之變換將每一音訊區塊轉換為一組頻域變換係數。為此，該音訊編碼解碼器110接收時域中之音訊資料(方塊302)，採用一20毫秒的音訊區塊或訊框(方塊304)，並將該區塊轉換為變換係數(方塊306)。每一變換係數具有一量值且可為正或負。

接著使用此項技術中已知的技術，用一量化器115量化此等變換係數並編碼(方塊308)，且該傳輸器100A經由一網路125(諸如一IP(網際網路協定)網路、PSTN(公眾交換電話網路)、ISDN(整合服務數位網路)或類似物)將經編碼的變換係數以封包發送至該接收器100B(方塊310)。該等封包可使用任意適合的協定或標準。舉例而言，音訊資料可遵循一內容表，且包括一音訊訊框之所有八位元組可作為一單元附加至有效負載。舉例而言，ITU-T提議G.719及G.722.1C(其等已併入本文中)指定該等音訊訊框之細節。

在該接收器100B處，一介面120接收該等封包(方塊312)。當發送該等封包時，該傳輸器100A產生包含於已發送的每一封包中之一序號。如已知，封包可經由該網路

125自該傳輸器100A通過不同路線至該接收器100B，且該等封包可在不同時間到達該接收器100B。因此，該等封包到達之順序可能隨機。

為處置此不同時間的到達(稱為「抖動」)，該接收器100B具有耦合至該接收器之介面120之一抖動緩衝器130。通常，該抖動緩衝器130每次保持四個或四個以上封包。因此，該接收器100B基於該等封包之序號將該抖動緩衝器130中之該等封包重新排序(方塊314)。

雖然該等封包可不按順序到達該接收器100B，但是損失封包處置器140適當地將該抖動緩衝器130中之該等封包重新排序，且基於該序列偵測任意損失(丟失)的封包。當該抖動緩衝器130中之該等封包之序號中存在間隙時，宣告一損失封包。舉例而言，若該處置器140在該抖動緩衝器130中發現序號005、006、007、011，則該處置器140可宣告封包008、009、010損失。事實上，此等封包可能實際上並未損失，而是該等封包僅可能係延遲到達。然而，由於延時及緩衝器長度約束，該接收器100B丟棄遲於某一臨限值到達之任意封包。

在隨後的一逆程序中，該接收器100B解碼並解量化該等經編碼的變換係數(方塊316)。若該處置器140已偵測到損失封包(決定318)，則該損失封包處置器140知道丟失封包間隙之前及之後的良好封包。變換合成器150使用此技術導出或以內插值取代該等損失封包之丟失的變換係數，因此新變換係數可取代來自該等損失封包之丟失的係數(方

塊320)。(在當前實例中，該音訊編碼解碼器使用MLT編碼，使得該等變換係數在本文可被稱為MLT係數)。在此階段，該接收器100B處之音訊編碼解碼器110對該等係數執行一逆變換，且將其等轉換回時域中以對該接收器之揚聲器產生輸出音訊(方塊322至方塊324)。

如在以上程序中可知，該損失封包處置器140將該基於變換之編碼解碼器110之損失封包作為變換係數之一損失組處置，而非偵測損失封包及持續轉發已接收音訊之先前片段來填充該間隙。該變換合成器150接著用自相鄰封包導出之經合成的變換係數來替代來自該等損失封包之變換係數的該損失組。接著，可使用該等係數之一逆變換而於該接收器100B處產生並輸出不具有來自損失封包之音訊間隙之一完整的音訊信號。

圖2B更詳細地示意性展示一會議端點或終端機100。如所示，該會議終端機100可為該IP網路125上之一傳輸器與接收器兩者。亦如所示，該會議終端機100可具有視訊會議能力以及音訊能力。一般而言，該終端機100具有一麥克風102及一揚聲器104，且可具有各種其他輸入/輸出裝置(諸如視訊相機106、顯示器108、鍵盤、滑鼠等)。此外，該終端機100具有一處理器160、記憶體162、轉換器電子器件164及適合該特定網路125之網路介面122/124。該音訊編碼解碼器110提供根據網路終端機之一適合協定之基於標準的會議。此等標準可完全以儲存於記憶體162中且在該處理器160、專用硬體上執行或使用其等之一組

合執行之軟體實施。

在一傳輸路徑中，轉換器電子器件164將該麥克風102所拾取之類比輸入信號轉換為數位信號，且在該終端機之處理器160上操作之該音訊編碼解碼器110具有編碼該等數位音訊信號之一編碼器200，以經由一傳輸器介面122在該網路125(諸如網際網路)上傳輸。若存在具有一視訊編碼器170之一視訊編碼解碼器，則其可對視訊信號執行類似功能。

在一接收路徑中，該終端機100具有耦合至該音訊編碼解碼器110之一網路接收器介面124。一解碼器250解碼已接收的信號，且轉換器電子器件164將該等數位信號轉換為類比信號以輸出至該揚聲器104。若存在具有一視訊解碼器172之一視訊編碼解碼器，則其可對視訊信號執行類似功能。

圖3A至圖3B簡要展示一變換編碼的編碼解碼器(諸如一Siren編碼解碼器)之特徵。一特定音訊編碼解碼器之實際細節取決於所使用的編碼解碼器之實施方案及類型。可在ITU-T提議G.722.1 Annex C中找到Siren14™之已知細節，且可在ITU-T提議G.719(2008)之「Low-complexity,full-band audio coding for high-quality,conversational applications」中找到Siren™22之已知細節，該兩者已以引用的方式併入本文中。亦可在美國專利申請案第11/550,629號及11/550,682號中找到關於音訊信號之變換編碼的額外細節，該等專利申請案係以引用的方式併入本

文中。

圖3A中圖解說明用於一變換編碼的編碼解碼器(例如，一Siren編碼解碼器)之一編碼器200。該編碼器200接收已自一類比音訊信號轉換之一數位信號202。舉例而言，可能已經以48 kHz或其他速率在約20毫秒區塊或訊框中取樣此數位信號202。可為一離散餘弦變換(DCT)之一變換204將來自時域之該數位信號202轉換為具有變換係數之一頻域。舉例而言，該變換204可對每一音訊區塊或訊框產生960個變換係數之一頻譜。該編碼器200在一正規化程序206中發現該等係數之平均能量位準(規範)。接著，該編碼器202用一快速點陣向量量化(FLVQ)演算法208或類似物量化該等係數以編碼一輸出信號210用於封裝及傳輸。

圖3B中圖解說明用於該變換編碼的編碼解碼器(例如，Siren編碼解碼器)之一解碼器250。該解碼器250採用自一網路接收之輸入信號252之傳入位元流，且自該傳入位元流重新產生初始信號之一最佳估計。為此，該解碼器250對該輸入信號252執行一點陣解碼(逆FLVQ)254，且使用一解量化程序256解量化該等經解碼的變換係數。而且，接著可在各種頻帶中修正該等變換係數之能量位準。

此刻，該變換合成器258可對丟失封包用係數作為內插值。最後，一逆變換260作為一逆DCT操作並將該信號自頻域轉換回時域中以作為一輸出信號262傳輸。如可知，該變換合成器258有助於填充可由該等丟失封包引起之任意間隙。然而，該解碼器250之所有現有功能及演算法仍

相同。

在對以上所提供的該終端機100及該音訊編碼解碼器110有所瞭解情況下，討論現在轉向該音訊編碼解碼器100如何藉由使用來自經由該網路接收之相鄰訊框、區塊或封包組來對丟失封包用變換係數作為內插值。(在MLT係數方面呈現隨後的討論，但是所揭示的內插程序可同等應用於其他形式的變換編碼之其他變換係數)。

如圖5中利用圖表所展示，用於在損失封包中用變換係數作為內插值之程序400涉及將一內插規則應用於(方塊410)來自先前良好的訊框、區塊或封包組(即，無損失封包)(方塊402)及來自隨後良好的訊框、區塊或封包組(方塊404)之變換係數。因此，該內插規則(方塊410)判定一給定組中損失之封包數目並因此從來自該等良好組之該等變換係數取得(方塊402/方塊404)。接著，該程序400對該等損失封包用新變換係數作為內插值以插入於該給定組中(方塊412)。最後，該程序400執行一逆變換(方塊414)並合成音訊組用於輸出(方塊416)。

圖5更詳細地利用圖表展示該內插程序之內插規則500。如前所述，該內插規則500係依據一訊框、音訊區塊或封包組中之損失封包之數目。實際訊框大小(位元/八位元組)取決於所使用的變換編碼演算法、位元速率、訊框長度及取樣速率。舉例而言，對於一48千位元/秒位元速率、一32 kHz取樣速率及一20毫秒訊框長度下之G.722.1 Annex C，該訊框大小將為960位元/120八位元組。對於G.719，

該訊框係20毫秒，該取樣速率係48 kHz，且該位元速率可於任意20毫秒訊框邊界處在32千位元/秒與128千位元/秒之間變化。RFC 5404中指定G.719之有效負載格式。

一般而言，已損失之一給定封包可具有一或多個音訊訊框(例如，20毫秒)、可僅包括一訊框之一部分、可具有一或多個音訊頻道之一或多個訊框、可在一或多個不同位元速率下具有一或多個訊框、且可具有熟習此項技術者已知並與所使用的特定變換編碼演算法及有效負載格式相關之其他複雜性。然而，用於對該等丟失封包以內插值取代丟失變換係數之該內插規則500可調適於一給定實施方案中之特定變換編碼及有效負載格式。

如所示，先前良好的訊框或組510之變換係數(此處展示為MLT係數)被稱為 $MLT_A(i)$ ，且隨後良好的訊框或組530之MLT係數被稱為 $MLT_B(i)$ 。若該音訊編碼解碼器使用Siren™22，則指數(i)處於自0至959的範圍。對該等丟失封包之所內插之MLT係數540的絕對值之一般內插規則520係基於應用於該先前及隨後MLT係數510/530之權重512/532判定，如下所示：

$$|MLT_{Interpolated}(i)| = Weight_A * |MLT_A(i)| + Weight_B * |MLT_B(i)|$$

在該一般內插規則中，以相等的概率將該丟失訊框或組之該等所內插之MLT係數 $MLT_{Interpolated}(i)$ 540之正負號522隨機設定為正或負。此隨機可有助於自此等經重新建構封包產生之音訊聽起來更自然且不太機械化。

在依此方式內插該等MLT係數540之後，該變換合成器

(150；圖 2A) 填充該等丟失封包之間隙，該接收器 (100B) 處之該音訊編碼解碼器 (110；圖 2A) 可接著完成其之合成操作以重新建構輸出信號。舉例而言，該音訊編碼解碼器 (110) 使用已知的技術以採用包含已接收之良好的 MLT 係數以及在需要處填充的所內插之 MLT 係數之經處理的變換係數之一向量 \bar{y} 。該編碼解碼器 (110) 自此向量 \bar{y} 重新建構由 $y = P_s \bar{y}$ 給出之一 $2M$ 樣本向量 y 。最後，隨著處理繼續，該合成器 (150) 採用該等經重新建構的 y 向量並將其等與 M 取樣重疊疊加以產生一經重新建構的信號 $y(n)$ 用於在該接收器 (100B) 處輸出。

隨著丟失封包之數目發生變化，該內插規則 500 對該先前 MLT 係數 510 及隨後 MLT 係數 530 應用不同權重 512/532 以判定該等所內插之 MLT 係數 540。以下係用於基於丟失封包之數目及其他參數判定兩個權重因數 $Weight_A$ 及 $Weight_B$ 之特定規則。

1. 單一損失封包

如圖 7A 中所圖表展示，該損失封包處置器 (140；圖 2A) 可偵測一主題訊框或封包組 620 中之一單一損失封包。若損失一單一封包，該處置器 (140) 基於關於該丟失封包之音訊頻率 (例如，該丟失封包之前的音訊之當前頻率)，將權重因數 ($Weight_A$ 、 $Weight_B$) 用於以內插值取代該損失封包之丟失的 MLT 係數。如以下圖表中所示，可相對於當前音訊之一 1 kHz 頻率判定先前訊框或組 610A 中之對應封包之該權重因數 ($Weight_A$) 及隨後訊框或組 610B 中之對應封包之該權

重因數 ($Weight_B$)，如下所示：

頻率	$Weight_A$	$Weight_B$
小於 1 kHz	0.75	0.0
大於 1 kHz	0.5	0.5

2. 兩個損失封包

如圖 7B 中所圖表展示，該損失封包處置器 (140) 可偵測一主題訊框或組 622 中之兩個損失封包。在此情況中，該處置器 (140) 將權重因數 ($Weight_A$ 、 $Weight_B$) 用於用 MLT 係數作為內插值以用於該先前訊框或組 610A 及隨後訊框或組 610B 之對應封包中之丟失封包，如下所示：

損失封包	$Weight_A$	$Weight_B$
第一(較舊)封包	0.9	0.0
最後的(較新)封包	0.0	0.9

若每一個封包包括一音訊訊框(例如，20 毫秒)，則圖 7B 之每一組 610A 至 610B 及 622 將基本上包含若干封包(即，若干訊框)，使得額外封包實際上不可能存在於如圖 7A 中描繪之該等組 610A 至 610B 及 622 中。

3. 三至六個損失封包

如圖 7C 中所圖表展示，該損失封包處置器 (140) 可偵測一主題訊框或組 624 中之三個至六個損失封包(圖 7C 中展示三個)。三個至六個丟失封包可表示在一給定時間間隔中損失之多達 25% 的封包。在此情況中，該處置器 (140) 將權重因數 ($Weight_A$ 、 $Weight_B$) 用於用 MLT 係數作為內插值以用於

該先前訊框或組 610A 及隨後訊框或組 610B 之對應封包中之丟失封包，如下所示：

損失封包	$Weight_A$	$Weight_B$
第一(較舊)封包	0.9	0.0
一或多個中間封包	0.4	0.4
最後的(較新)封包	0.0	0.9

圖 7A 至圖 7C 之該等圖表中之封包及訊框或組之配置係意謂闡釋性。如前所述，一些編碼技術可使用包括一特定長度(例如，20 毫秒)的音訊之訊框。而且，一些技術可對每一音訊訊框(例如，20 毫秒)使用一封包。然而，取決於實施方案，一給定封包可具有一或多個音訊訊框(例如，20 毫秒)之資訊或可具有一音訊訊框(例如，20 毫秒)之唯一的一部分之資訊。

為定義用於以內插值取代丟失變換係數之權重因數，上述參數使用頻率位準、在一訊框中丟失之封包數目及一丟失封包在丟失封包之一給定組中的位置。可使用此等內插參數之任一者或任意組合定義該等權重因數。用於用變換係數作為內插值之上文所揭示之該等權重因數($Weight_A$ 、 $Weight_B$)、頻率臨限值及內插參數係闡釋性。據信，在一會議期間，此等權重因數、臨限值及參數在填充來自丟失封包之間隙時產生最佳的主觀音訊品質。然而，此等因數、臨限值及參數可對於一特定實施方案而不同，可擴展至闡釋性地呈現之範圍以外，且可取決於所使用的設備類型、

所涉及的音訊類型(即，音樂、語音等)、所應用的變換編碼類型及其他考量。

無論如何，當對基於變換之音訊編碼解碼器隱蔽損失的音訊封包時，所揭示的音訊處理技術產生比先前技術解決方案品質更佳之聲音。特定言之，即使損失25%的封包，該所揭示的技術仍可產生比當前技術更易聽懂之音訊。音訊封包損失通常發生在視訊會議應用中，因此改良此等條件期間之品質對改良總體視訊會議體驗係重要的。然而，重要的是，隱蔽封包損失所採取之步驟無需操作用於隱蔽該損失之終端機處之過多的處理或儲存資源。藉由對先前及隨後良好的訊框中之變換係數應用加權，該等所揭示的技術可減小所需要的處理及儲存資源。

雖然已在音訊或視訊會議方面進行描述，但是本發明之教示在涉及包含串流音樂及話音之串流媒體之其他領域中係有用。因此，本發明之教示可應用於除一音訊會議端點及一視訊會議端點之外之其他音訊處理裝置，包含一音訊重放裝置、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、一個人數位助理等。舉例而言，專用的音訊或視訊會議端點可受益於該等所揭示的技術。同樣地，電腦或其他裝置可用於桌面會議中或用於數位音訊之傳輸及接收，且此等裝置亦可受益於該等所揭示的技術。

本發明之該等技術可在電子電路、電腦硬體、韌體、軟體或此等之任意組合中實施。舉例而言，該等所揭示的技

術可實施為用於導致一可程式化控制裝置執行該等所揭示的技術之儲存於一程式儲存裝置上的指令。適用於有形地體現程式指令及資訊之程式儲存裝置包含所有形式的非揮發性記憶體，舉例而言，包含半導體記憶體裝置(諸如可擦除且可程式化唯讀記憶體(EPROM)、電可擦除且可程式化唯讀記憶體(EEPROM)及快閃記憶體裝置)；磁碟(諸如內部硬碟及抽換式磁碟)；磁光碟；及唯讀光碟(CD-ROM)。前述之任一者可由特定應用積體電路(ASIC)(特定應用積體電路)補充或併入其中。

對較佳及其他實施例之以上描述並非意欲限制或約束申請人所設想之發明概念之範疇或適用性。作為揭示本文所含有之該等發明概念的交換，該等申請人期望由隨附申請專利範圍提供之所有專利權利。因此，意欲使隨附申請專利範圍在最大程度上包含處於以下申請專利範圍或其相等物之範疇內之所有修改及變更。

【圖式簡單說明】

圖1圖解說明具有一傳輸器及一接收器並使用根據先前技術之損失封包技術之一會議配置。

圖2A圖解說明具有一傳輸器及一接收器並使用根據本發明之損失封包技術之一會議配置。

圖2B更詳細地圖解說明一會議終端機。

圖3A至圖3B各自展示一變換編碼的編碼解碼器之一編碼器及解碼器。

圖4係根據本發明之一編碼、解碼及損失封包處置技術

之一流程圖。

圖5利用圖表展示根據本發明之用於在損失封包中用變換係數作為內插值之一程序。

圖6利用圖表展示內插程序之一內插規則。

圖7A至圖7C利用圖表展示用於對丟失封包用變換係數作為內插值之權重。

【主要元件符號說明】

10A	傳輸器/第一終端機
10B	接收器/第二終端機
12	麥克風
13	揚聲器
14	音訊區塊
15	損失封包偵測
16	編碼解碼器
17	音訊轉發器
18	量化器
19	解量化器
20	網際網路
100A	傳輸器/第一終端機
100B	接收器/第二終端機
102	麥克風
104	揚聲器
106	視訊相機
108	顯示器

110	編碼解碼器
115	量化器
120	介面
122	傳輸器介面
124	接收器介面
125	網際網路
130	抖動緩衝器
140	損失封包處置器
150	變換合成器
160	處理器
162	記憶體
164	類比轉數位轉換器
170	視訊編碼器
175	視訊解碼器
200	編碼器
210	輸出信號
250	解碼器
510	先前良好的訊框
512	權重 A
520	主題訊框
522	隨機正負號
530	隨後良好的訊框
532	權重 B
540	所內插之調變重疊變換(MLT)係數

610A	先前的訊框或組
610B	隨後的訊框或組
620	主題訊框或封包組
622	主題訊框或組
624	主題訊框或組

七、申請專利範圍：

1. 一種音訊處理方法，其包括：

102年9月27日修正本

經由一網路在一音訊處理裝置處接收若干封包組，每一組具有該等封包之一或多者，每一封包具有頻域的數個變換係數，其係用於重新建構已經歷變換編碼之時域中之一音訊信號；

判定已接收組之一給定組中之一或多個丟失封包，該一或多個丟失封包定序在具有一給定序列之該給定組中；

對定序在該給定組之前的一第一組中之一或多個第一封包之第一變換係數應用一第一權重，該一或多個第一封包在該第一組中具有一第一序列，該第一序列對應到在該給定組中之該一或多個丟失封包之該給定序列；

對定序在該給定組之後的一第二組中之一或多個第二封包之第二變換係數應用一第二權重，該一或多個第二封包在該第二組中具有一第二序列，該第二序列對應到在該給定組中之該一或多個丟失封包(520)之該給定序列；

藉由加總對應之第一經加權的變換係數與對應之第二經加權的變換係數而用變換係數作為內插值；

將該等所內插之變換係數插入於取代該一或多個丟失封包中之該給定組中；及

藉由對該等變換係數執行一逆變換而對該音訊處理裝置產生一輸出音訊信號。

2. 如請求項1之方法，其中該音訊處理裝置係選自由一音訊會議端點、一視訊會議端點、一音訊重放裝置、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。
3. 如請求項1之方法，其中該網路包括一網際網路協定網路。
4. 如請求項1之方法，其中該等變換係數包括一調變重疊變換之係數。
5. 如請求項1之方法，其中每一組具有一個封包，且其中該一個封包包括一輸入音訊訊框。
6. 如請求項1之方法，其中接收包括：解碼該等封包。
7. 如請求項6之方法，其中接收包括：解量化該等已解碼的封包。
8. 如請求項1之方法，其中判定該一或多個丟失封包包括：對在一緩衝器中接收的封包定序及找到該定序中之間隙。
9. 如請求項1之方法，其中用該等變換係數作為內插值包括：對經加總的該等第一經加權的變換係數及該等第二經加權的變換係數指派一隨機的正號或負號。
10. 如請求項1之方法，其中應用於該等第一變換係數及該等第二變換係數之該第一權重及該第二權重係基於該第一及該第二變換係數之多個頻率。
11. 如請求項10之方法，其中針對在一臨限值以下之該第一及該第二變換係數之該等頻率之每一者，該第一權重強

調該等第一變換係數，且該第二權重解除強調該等第二變換係數。

12. 如請求項11之方法，其中該臨限值係1 kHz。
13. 如請求項11之方法，其中該等第一變換係數係以75%加權，且其中該等第二變換係數係以零加權。
14. 如請求項10之方法，其中針對在一臨限值以上之該第一及該第二變換係數之該等頻率之每一者，該第一權重及該第二權重同等強調該等第一變換係數及該等第二變換係數。
15. 如請求項14之方法，其中該等第一變換係數與該等第二變換係數兩者皆係以50%加權。
16. 如請求項1之方法，其中應用於該等第一變換係數及該等第二變換係數之該第一權重及該第二權重係基於該等丟失封包之一數目。
17. 如請求項16之方法，其中若在該給定組中丟失該等封包之一者，則
 針對在一臨限值以下之該第一及該第二變換係數之每一頻率，該第一權重強調該等第一變換係數且該第二權重解除強調該等第二變換係數；及
 針對在一臨限值以上之該第一及該第二變換係數之每一頻率，該第一權重及該第二權重同等強調該等第一變換係數及該等第二變換係數。
18. 如請求項16之方法，其中若在該給定組中丟失該等封包之兩者，則

該第一權重對該兩個封包之前一者強調該等第一變換係數，且對該兩個封包之後一者解除強調該等第一變換係數；及

該第二權重對該前一個封包解除強調該等第二變換係數，且對該後一個封包強調該等第二變換係數。

19. 如請求項18之方法，其中該等經強調的係數係以90%加權，且其中該等經解除強調的係數係以零加權。

20. 如請求項16之方法，其中若在該給定組中丟失三個或三個以上封包，則

該第一權重對該等封包之一第一者強調該等第一變換係數，且對該等封包之一最後一者解除強調該等第一變換係數；

該第一權重及該第二權重對該等封包之一或多個中間封包同等強調該等第一變換係數及該等第二變換係數；及

該第二權重對該等封包之該第一者解除強調該等第二變換係數，且對該等封包之該最後一者強調該等第二變換係數。

21. 如請求項20之方法，其中該等經強調的係數係以90%加權，其中該等經解除強調的係數係以零加權，且其中該等經同等強調的係數係以40%加權。

22. 一種程式儲存裝置，其具有儲存於其上之用於導致一可程式化控制裝置執行如請求項1至21之任一項之一音訊處理方法之指令。

23. 一種音訊處理裝置，其包括：

一音訊輸出介面；

一網路介面，其與至少一網路通信並接收音訊封包組，每一組具有該等封包之一或多者，每一封包具有頻域之變換係數；

記憶體，其與該網路介面通信並儲存已接收的該等封包；

一處理單元，其與該記憶體及該音訊輸出介面通信，該處理單元經一音訊解碼器程式化，該音訊解碼器係經組態以執行如請求項1至21之任一項之一音訊處理方法。

24. 如請求項23之裝置，其中該裝置包括一會議端點。
25. 如請求項23之裝置，其進一步包括可通信地耦合至該音訊輸出介面之一揚聲器。
26. 如請求項23之裝置，其進一步包括一音訊輸入介面及可通信地耦合至該音訊輸入介面之一麥克風。
27. 如請求項26之裝置，其中該處理單元係與該音訊輸入介面通信，且係經一音訊編碼器程式化，該音訊編碼器係經組態以：

將一音訊信號之時域樣本之訊框變換為頻域變換係數；

量化該等變換係數；及

編碼該等經量化的變換係數。

八、圖式：

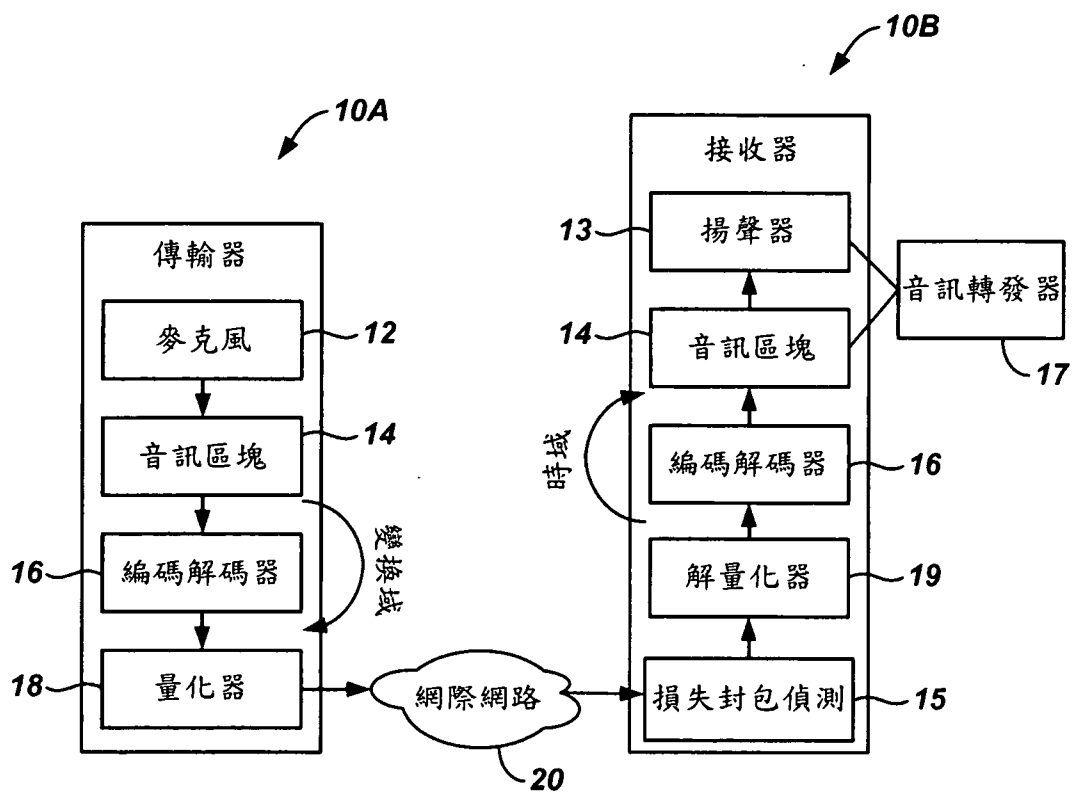


圖 1

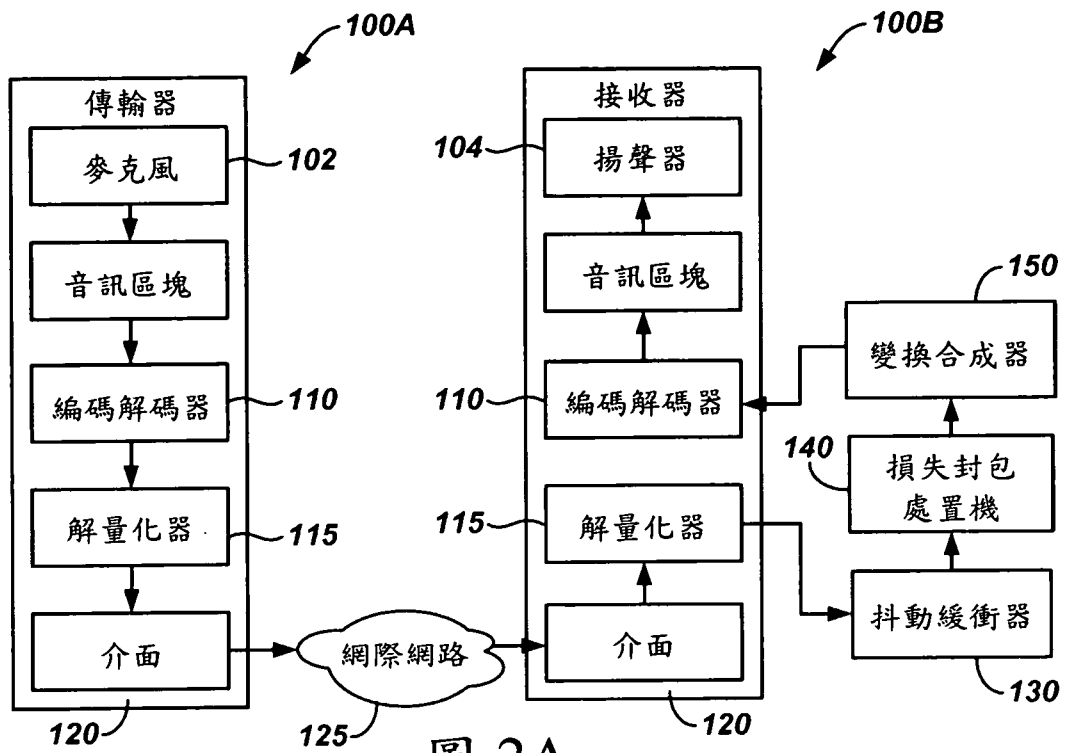


圖 2A

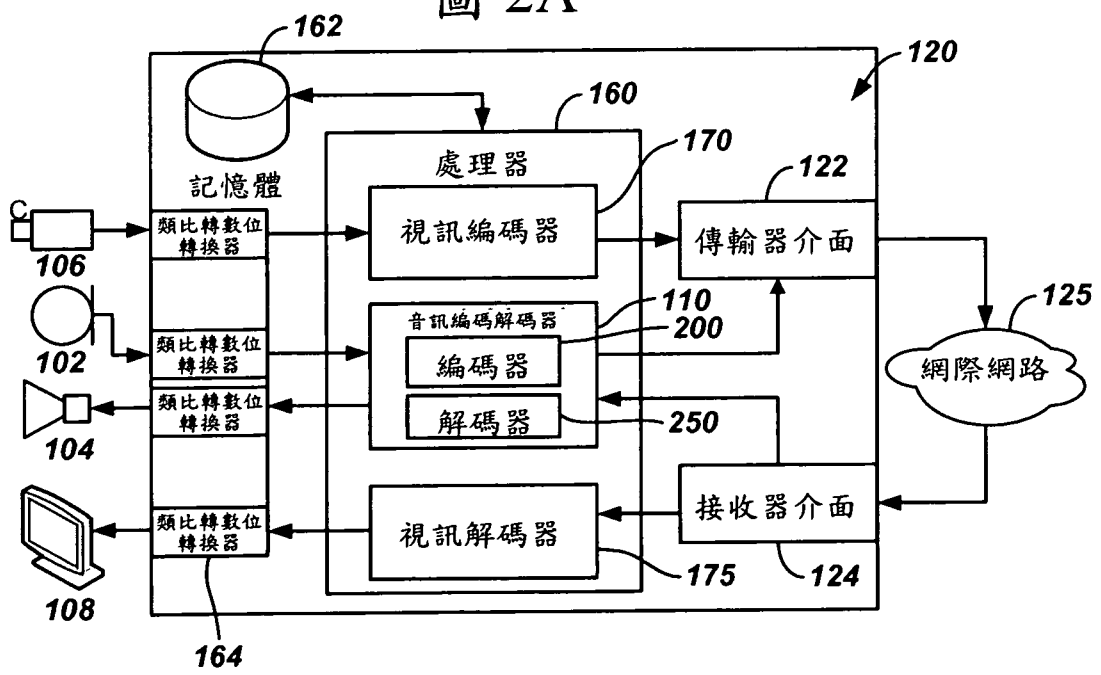


圖 2B

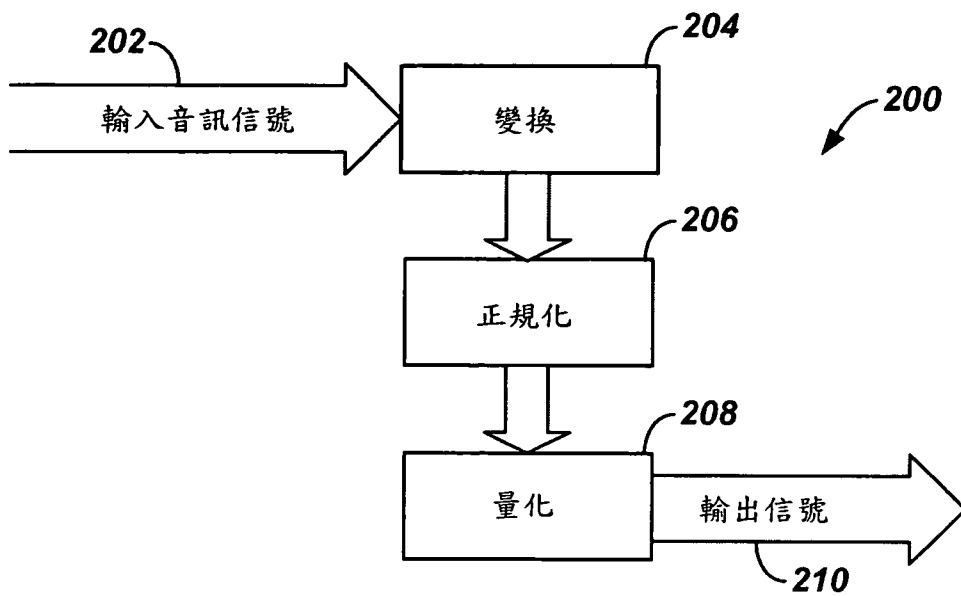


圖 3A

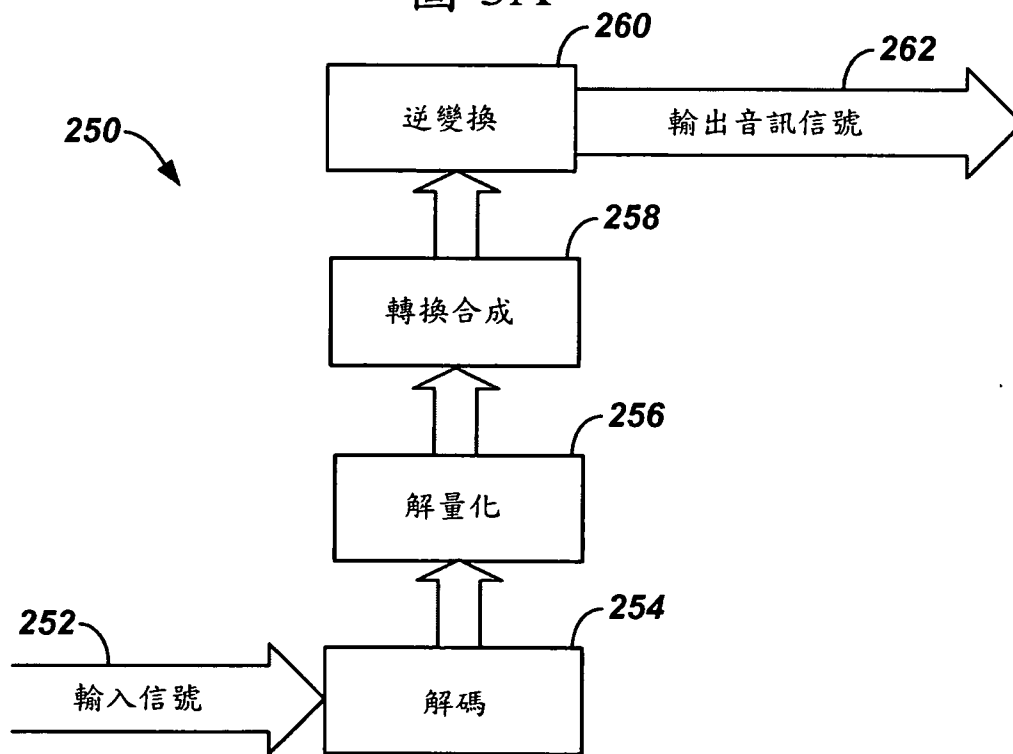


圖 3B

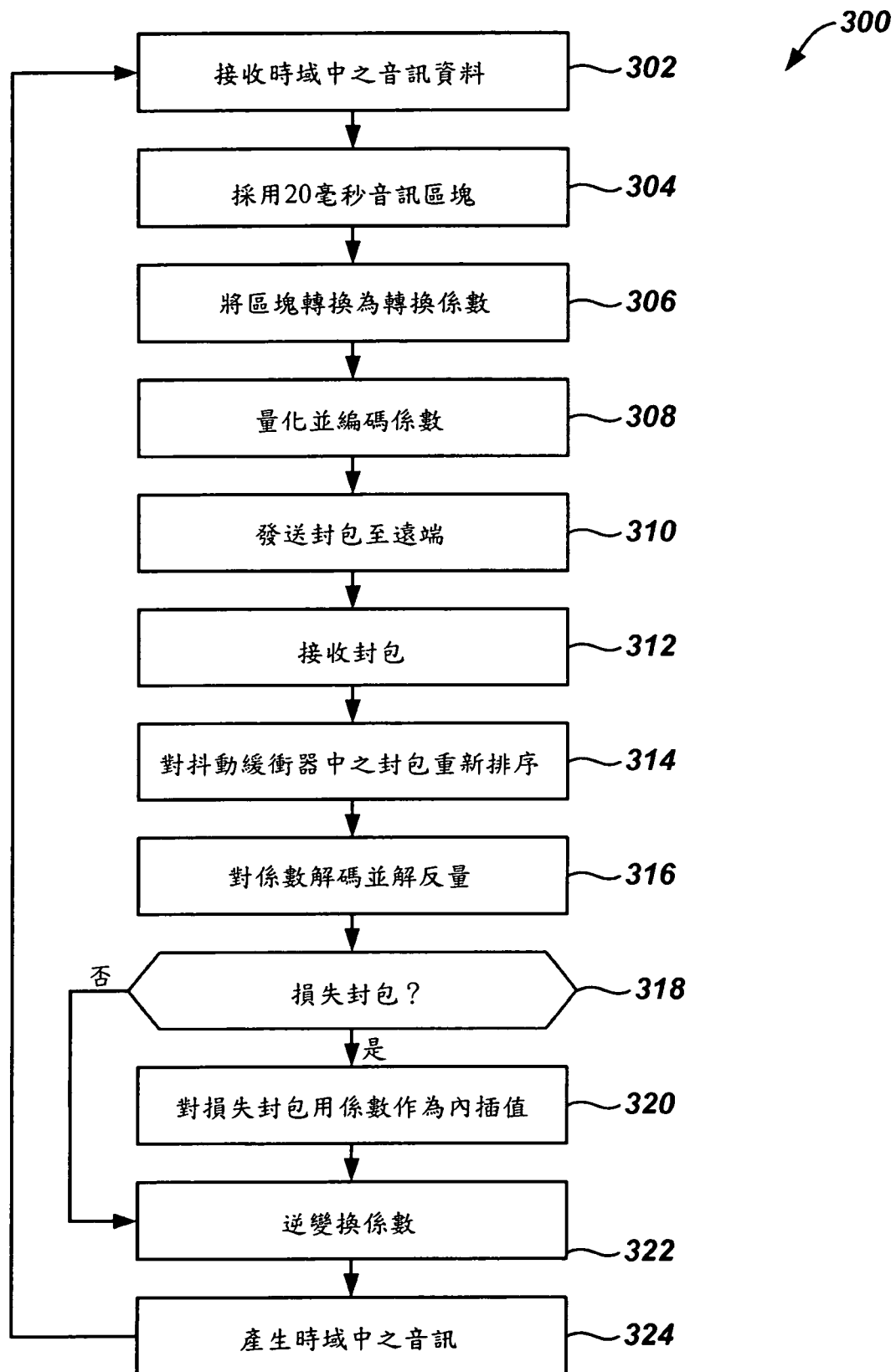


圖 4

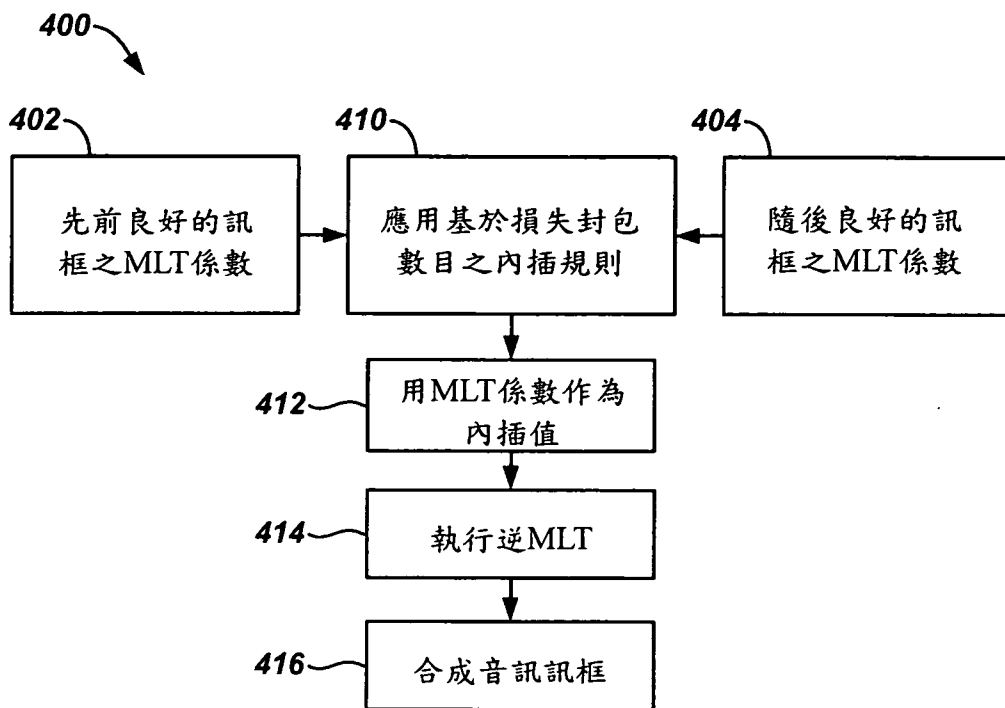


圖 5

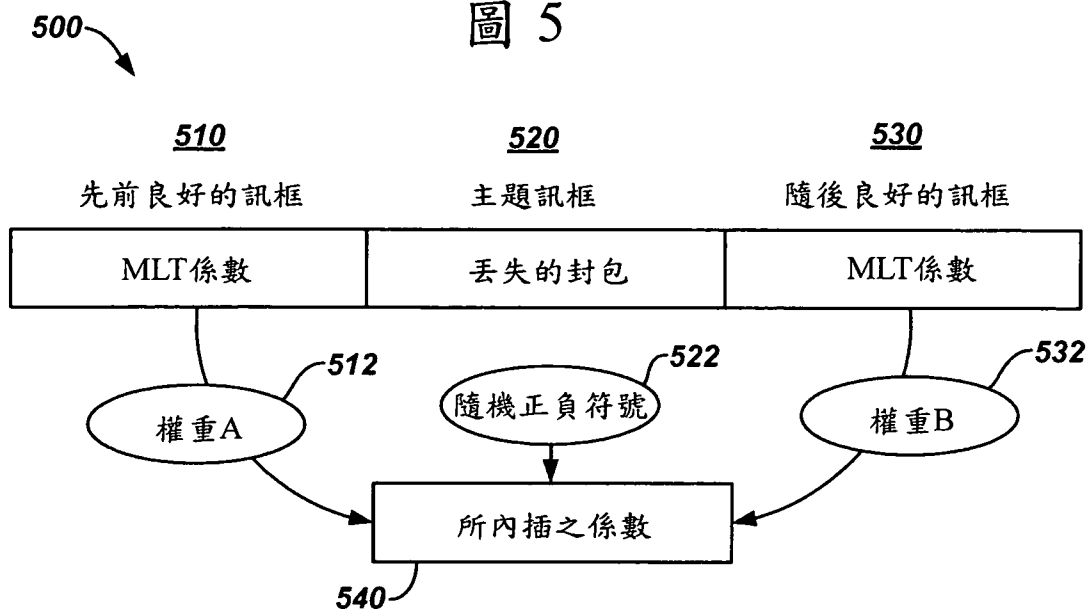


圖 6

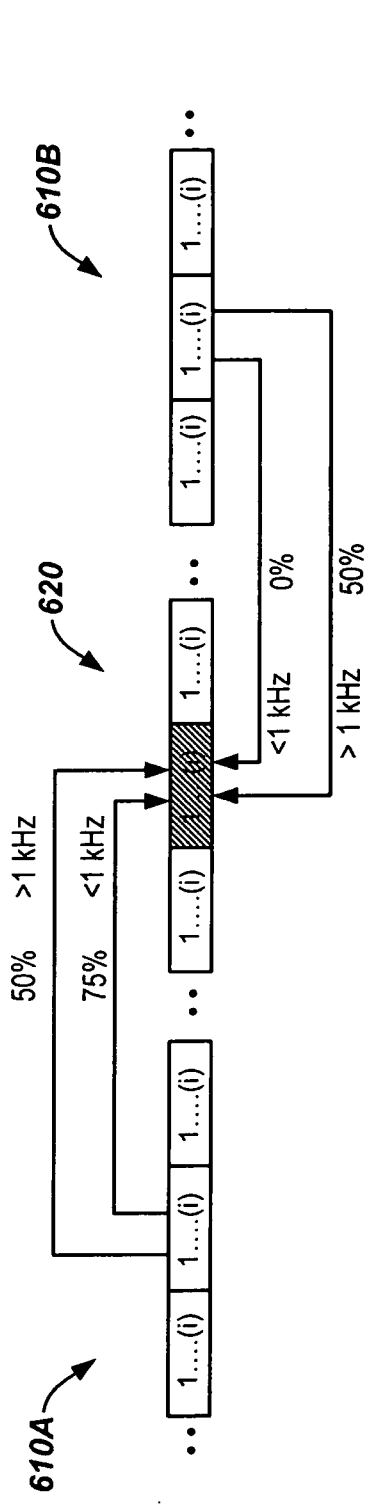


圖 7A

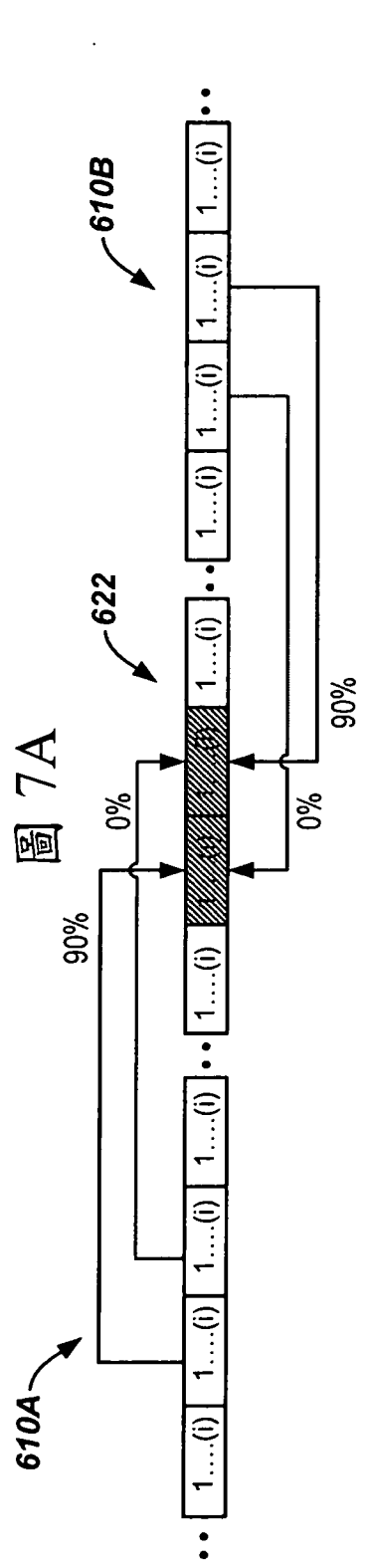


圖 7B

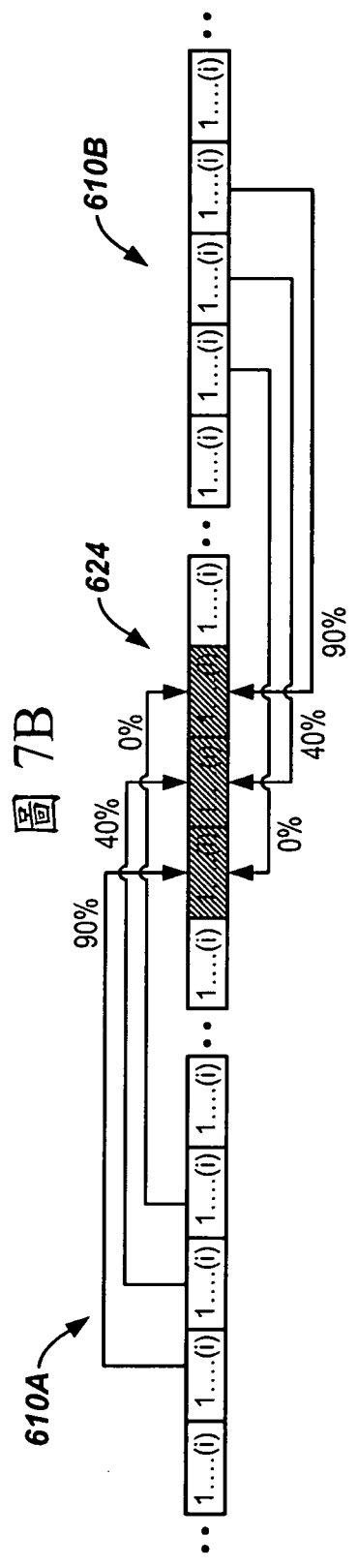


圖 7C