



(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 I505265 B

(45)公告日：中華民國 104 (2015) 年 10 月 21 日

(21)申請案號：101114952

(22)申請日：中華民國 98 (2009) 年 02 月 23 日

(51)Int. Cl. : G10L21/04 (2013.01)

(30)優先權：2008/03/10 美國 61/035,317
2009/02/17 世界智慧財產權組織 PCT/EP 2009/001108(71)申請人：弗勞恩霍夫應用研究促進協會(德國) FRAUNHOFER-GESELLSCHAFT ZUR
FORDERUNG DER ANGEWANDTEN FORSCHUNG E. V. (DE)
德國(72)發明人：迪施 薩沙 DISCH, SASCHA (DE)；納格爾 弗雷德里克 NAGEL, FREDERIK
(DE)；里特爾博謝 尼古拉斯 RETTELBACH, NIKOLAUS (DE)；馬特拉斯 馬
庫斯 MULTRUS, MARKUS (DE)；福克斯 紀堯姆 FUCHS, GUILLAUME (FR)

(74)代理人：劉緒倫

(56)參考文獻：

US 6266644B1 US 6766300B1
WO 02/084645A2

審查人員：黃衍勳

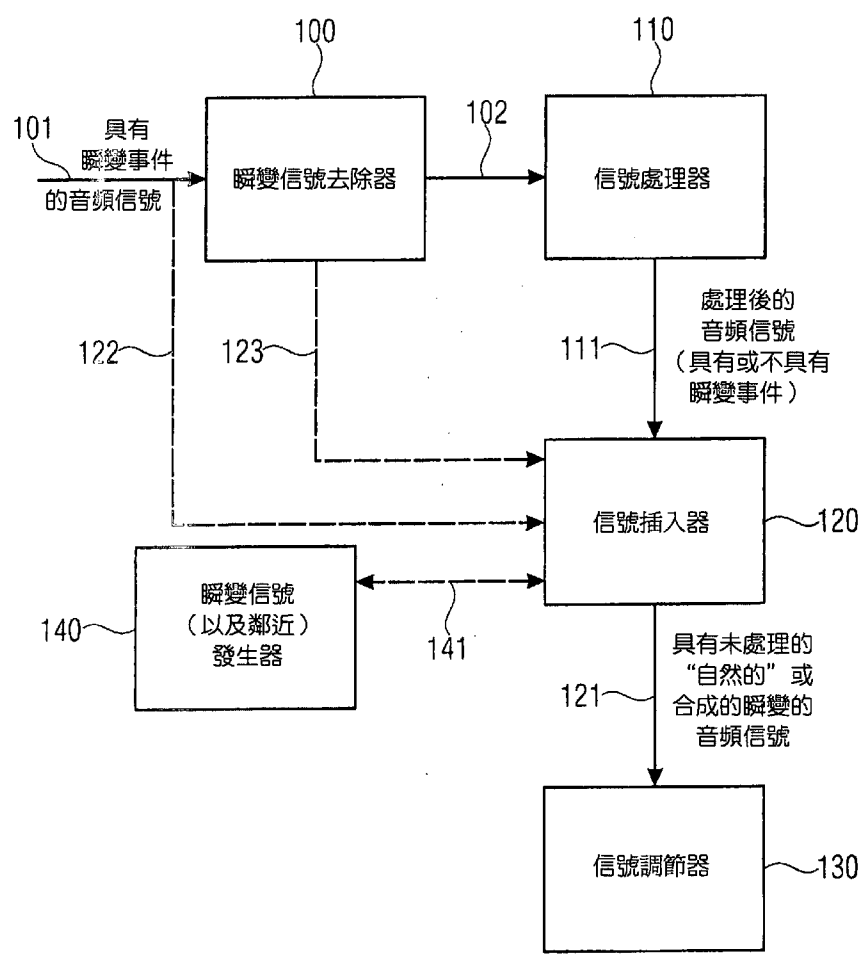
申請專利範圍項數：11 項 圖式數：9 共 52 頁

(54)名稱

操縱具有瞬變事件的音頻信號的設備和方法以及具有執行該方法之程式碼的電腦程式
DEVICE AND METHOD FOR MANIPULATING AN AUDIO SIGNAL HAVING A TRANSIENT
EVENT, AND A COMPUTER PROGRAM HAVING A PROGRAM CODE FOR PERFORMING THE
METHOD

(57)摘要

一種信號操縱器，用於操縱具有瞬變事件的音頻信號，可以包括：瞬變去除器(100)、信號處理器(110)以及信號插入器(120)，所述信號插入器(120)用於在信號位置將時間部分插入處理後的音頻信號中，使得受操縱音頻信號包括不受所述處理影響的瞬變事件，其中所述信號位置是在所述瞬變去除器進行處理之前去除瞬變事件的信號位置，從而瞬變事件的垂直相干性保持不變，而在信號處理器(110)中執行的任何處理都不會破壞瞬變的垂直相干性。



第一圖

- 100 . . . 瞬變信號去除器
- 101 . . . 輸入
- 102 . . . 輸出
- 110 . . . 信號處理器
- 111 . . . 信號處理器輸出
- 120 . . . 信號插入器
- 121 . . . 信號插入器輸出
- 130 . . . 信號調節器
- 140 . . . 瞬變信號發生器

發明專利說明書

公告本

(本說明書格式、順序，請勿任意更動，※記號部分請勿填寫)

※ 申請案號：101/114952

※ 申請日：98.2.23

※IPC 分類：G10L 21/04 (2013.01)

一、發明名稱：(中文/英文)

操縱具有瞬變事件的音頻信號的設備和方法以及具有執行該方法之程式碼的電腦程式 /

Device and Method for Manipulating an Audio Signal Having a Transient Event, and a Computer Program Having a Program Code for Performing the Method

二、中文發明摘要：

一種信號操縱器，用於操縱具有瞬變事件的音頻信號，可以包括：瞬變去除器（100）、信號處理器（110）以及信號插入器（120），所述信號插入器（120）用於在信號位置將時間部分插入處理後的音頻信號中，使得受操縱音頻信號包括不受所述處理影響的瞬變事件，其中所述信號位置是在所述瞬變去除器進行處理之前去除瞬變事件的信號位置，從而瞬變事件的垂直相干性保持不變，而在信號處理器（110）中執行的任何處理都不會破壞瞬變的垂直相干性。

三、英文發明摘要：

四、指定代表圖：

(一)本案指定代表圖為：第（ 一 ）圖。

(二)本代表圖之元件符號簡單說明：

瞬變信號去除器 100

輸入 101

輸出 102

信號處理器 110

信號處理器輸出 111

信號插入器 120

信號插入器輸出 121

信號調節器 130

瞬變信號發生器 140

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

六、發明說明：

【發明所屬之技術領域】

本發明涉及音頻信號處理，具體涉及在向包含瞬變事件的信號應用音頻效果的情況下的音頻信號操縱。

【先前技術】

已知操縱音頻信號使得改變再現速度，同時保持音高（pitch）不變。針對這樣的過程的已知方法是利用相位聲碼器（vocoder）或方法來實現的，如（音高同步的）疊加（overlap-add）、(P)SOLA，如在 J.L. Flanagan 和 R.M. Golden, The Bell System Technical Journal, November 1966, pp. 1349 to 1590；美國專利 6549884 Laroche, J. & Dolson, M.: Phase-vocoder pitch-shifting；Jean Laroche 和 Mark Dolson, New Phase-Vocoder Techniques for Pitch-Shifting, Harmonizing And Other Exotic Effects”，Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, Oct. 17-20, 1999；以及 Zölzer, U: DAFX: Digital Audio Effects；Wiley & Sons；Edition: 1(February 26, 2002)；pp. 201-298 中所描述的。

此外，可以使用這樣的方法（即，相位聲碼器或 (P)SOLA）對音頻信號進行轉換（transposition），其中這種轉換的具體問題是：轉換後的音頻信號與轉換之前的原始音頻信號具有相同的再現/重放長度，而音高發生改變。這是通過加速再現拉伸信號（stretched signal）而得到的，

其中執行加速再現的加速因數依賴於在時間上拉伸原始音頻信號的拉伸因數。在採用時間離散的信號表示時，該過程對應於：利用等於拉伸因數的因數對拉伸信號的下採樣（down-sampling）或對拉伸信號的抽取（decimation），其中採樣頻率保持不變。

在這樣的音頻信號操縱方面的具體挑戰是瞬變事件。瞬變事件是：在整個頻帶中或特定頻率範圍內信號的能量快速改變（即，快速增大或快速減小）的信號中的事件。具體瞬變（瞬變事件）的特有特徵（characteristic feature）是信號能量在頻譜中的分佈。典型地，在瞬變事件期間音頻信號的能量分佈在整個頻率上，而在非瞬變信號部分中，能量通常集中在音頻信號的低頻部分或特定頻帶中。這意味著，還稱作穩定或音調（tonal）信號部分的非瞬變信號部分具有非平坦的（non-flat）頻譜。換言之，信號的能量包含在很少數目的譜線/譜帶中，這些譜線/譜帶明顯高於音頻信號的雜訊基底（noise floor）。然而在瞬變部分，音頻信號的能量將分佈在許多不同頻帶上，具體地，將分佈在高頻部分，使得音頻信號的瞬變部分的頻譜會比較平坦，並且在任何事件下都會比音頻信號的音調部分的頻譜更為平坦。典型地，瞬變事件是時間上的強烈變化，這意味著當執行傅裏葉分解時信號將包括高次諧波（higher harmonic）。這些高次諧波的重要特徵是，這些高次諧波的相位有非常特殊的相互關係，使得所有這些正弦波的疊加（superposition）將導致信號能量的快速改變。

換言之，在頻譜上存在強相關（strong correlation）。

所有諧波之間的具體相位情況還可以稱作“垂直相干性（vertical coherence）”。該“垂直相干性”與信號的時間/頻率譜圖表示有關，在所述信號的時間/頻率譜圖表示中，水準方向對應於信號在時間上的演進，垂直尺度在頻率上描述了一個短時譜中譜分量的頻率（轉換頻率點（transform frequency bins））的相互依賴。

為了時間拉伸或縮短音頻信號而執行的典型處理步驟使得這種垂直相干性被破壞，這意味著當例如由相位聲碼器或任何其他方法對瞬變執行時間拉伸或縮短操作時，瞬變隨時間而“模糊（smear）”，所述相位聲碼器或任何其他方法執行基於頻率的處理，向音頻信號引入隨不同頻率係數而不同的相移。

當音頻信號處理方法破壞了瞬變的垂直相干性時，受操縱（manipulated）信號將會在穩定或非瞬變部分非常類似於原始信號，而在受操縱信號中瞬變部分將會品質降低。對瞬變的垂直相干性進行不受控制的操縱導致了瞬變的時間分散（temporal dispersion），這是因為：許多諧波分量對瞬變事件做貢獻，並且以不受控制的方式來改變所有這些分量的相位，不可避免地導致了這樣的偽像（artifact）。

然而，瞬變部分對於音頻信號的動態而言（如音樂信號或語言信號，其中在特定時刻能量的突然改變表示對受控信號的品質的大量主觀用戶印象）是尤為重要的。換言

之，典型地，音頻信號中的瞬變事件是語音信號的非常明顯的“重要事件”，其對主觀品質印象有超比例（over-proportional）的影響。受操縱的瞬變將使收聽者聽到失真的、迴響的並且不自然的聲音，在所述受操作瞬變中，垂直相關性被信號處理操作所破壞或相對於原始信號的瞬變部分而變差。

一些當前方法將瞬變周圍的時間拉伸到更高的程度，以便隨後在瞬變的持續時間期間不執行或僅執行小（minor）的時間拉伸。這樣的現有技術參考和專利描述了時間和/或音高操縱的方法。現有技術參考是：Laroche L., Dolson M.: Improved phase vocoder timescale modification of audio”, IEEE trans. Speech and Audio Processing, vol. 7, no. 3, pp. 323-332; Emmanuel Ravelli, Mark Sandler 和 Juan P. Bello: Fast implementation for non-linear time-scaling of stereo audio ; Proc. of the 8th Int. Conference on Digital Audio Effects (DAFx’ 05), Madrid, Spain, September 20-22, 2005 ; Duxbury, C. M. Davies 和 M. Sandler (2001, December) : Separation of transient information in musical audio using multiresolution analysis techniques. In proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01), Limerick, Ireland ; 以及 Röbel, A.: A NEW APPROACH TO TRANSIENT PROCESSING IN THE PHASE VOCODER ; Proc. of the 6th Int. Conference on Digital Audio Effect (DAFx-03), London, UK, September

8-11, 2003。

在相位聲碼器對音頻信號進行時間拉伸期間，時間分散使瞬變信號部分變得“模糊”，這是因為削弱了所謂的信號垂直相干性。使用所謂的疊加方法的方法，如(P)SOLA，可以產生瞬變聲音事件的干擾前回聲（pre-echo）和後回聲（post-echo）。通過瞬變環境中增大的時間拉伸，可以實際上解決這些問題；然而，如果要出現轉換，則在瞬變環境下轉換因數將不再是恒定的，即，所疊加的（可能是音調）信號分量的音高將改變並且將作為干擾而被感知。

【發明內容】

本發明的目的是為音頻信號操縱提供一種更高品質的構思。

利用依據申請專利範圍第 1 項所述的操縱音頻信號的設備、依據申請專利範圍第 12 項所述的產生音頻信號的設備、依據申請專利範圍第 13 項所述的操縱音頻信號的方法、依據申請專利範圍第 14 項所述的產生音頻信號的方法、依據申請專利範圍第 15 項所述的具有瞬變部分和輔助資訊的音頻信號、或者依據申請專利範圍第 16 項所述的電腦程式，實現了該目的。

為了解決在對瞬變部分的非受控處理中出現的品質問題，本發明保證根本不會以有害的方式對瞬變部分進行處理，即，在處理之前去除瞬變部分並且在處理之後將其

重新插入，或處理過瞬變部分，但是將其從處理過的信號中去除並替換成未處理過的瞬變事件。

優選地，插入處理過的信號中的瞬變部分是原始信號中相應瞬變部分的副本，使得受操縱信號由不包含瞬變事件的處理過的部分以及包含瞬變事件的未處理過的或不同地處理過的部分組成。例如，可以對原始瞬變進行抽取或任何類型的加權或參數化處理。然而，可選地，可以將瞬變部分替換成合成地產生的瞬變部分，以這樣的方式來合成所述合成地產生的瞬變部分，使得合成的瞬變部分在某些瞬變參數（如，在特定時刻的能量變化量，或描述瞬變事件特徵的任何其他量度）方面類似於原始瞬變部分。因此，甚至可以對原始音頻信號中的瞬變部分特徵化，可以在處理之前去除該瞬變，或將處理過的瞬變替換成合成瞬變，所述合成瞬變是根據瞬變參數資訊而合成地產生的。然而，出於效率原因，優選的是在操縱之前複製原始音頻信號的一部分，以及將該副本插入處理過的音頻信號中，這是因為該過程保證了處理過的信號中的瞬變部分與原始信號的瞬變相同。該過程將確保與處理之前的原始信號相比，在處理過的信號中保持了瞬變對聲音信號感知的特殊的高影響。因此，用於操縱音頻信號的任何類型的音頻信號處理都不會降低關於瞬變的主觀或客觀品質。

在優選實施例中，本申請提供了一種新方法，在這樣的處理的架構內，對瞬變聲音事件進行感知性良好的處理，否則將由於信號的分散而產生時間上的“模糊”。該優

選方法主要包括：在信號操縱之前去除瞬變聲音事件，以執行時間拉伸；隨後考慮到該拉伸，以精確的方式將未處理的瞬變信號部分添加到修改後的（拉伸後的）信號中。

【實施方式】

隨後參考附圖說明了本發明的優選實施例。

第一圖示出了操縱具有瞬變事件的音頻信號的優選設備。優選地，該設備包括瞬變信號去除器 100，瞬變信號去除器 100 具有用於具有瞬變事件的音頻信號的輸入 101。瞬變信號去除器的輸出 102 與信號處理器 110 連接。信號處理器輸出 111 與信號插入器 120 連接。信號插入器輸出 121 可以與諸如信號調節器（conditioner）130 之類的其他設備連接，其中在所述信號插入器輸出 121 上具有未處理的“自然的”或合成的瞬變的被操縱音頻信號是可用的，所述信號調節器 130 可以執行受操縱信號的任何其他處理，如為了帶寬擴展的目的而需要的下採樣/抽取，如結合第七圖 A 和第七圖 B 所討論的。

然而，如果按原樣使用在信號插入器 120 的輸出處得到的受操縱音頻信號，即，被儲存以進行進一步處理、被傳輸至接收機、或被傳輸至數位/類比轉換器，其中所述數位/類比轉換器最後與擴音器設備連接以最終產生表示受操縱音頻信號的聲音信號，則根本不能使用信號調節器 130。

在帶寬擴展的情況下，線 121 上的信號可以已經是高

頻段信號。那麼，信號處理器已經根據輸入的低頻段信號產生了高頻段信號，而且從音頻信號 101 提取的低頻段瞬變部分將會被置於高頻段的頻率範圍中，優選地，這是通過不干擾垂直相干性的信號處理來實現的，如抽取。在信號插入器之前執行這種抽取，以便將所抽取的瞬變部分插入塊 110 的輸出處的高頻段信號中。在該實施例中，信號調節器將執行高頻段信號的任何其他處理，如包絡整形、雜訊添加、反向濾波、或添加諧波等等，如在 MPEG4 頻帶複製（spectral band replication）中進行的。

優選地，信號插入器 120 經由線 123 接收來自去除器 100 的輔助資訊，以便根據將要插入 111 中的未處理信號來選擇正確的部分。

在實現具有設備 100、110、120、130 的實施例時，可以得到如結合第八圖 A 至第八圖 E 所討論的信號序列。然而，不一定要在信號處理器 110 中執行信號處理操作之前去除瞬變部分。在該實施例中，不需要瞬變信號去除器 100，信號插入器 120 確定要從輸出 111 上的處理信號中切除的信號部分，以及將該切除信號替換成如線 121 示意性所示的原始信號或如線 141 示意性所示的合成信號，其中該合成信號是可以從瞬變信號發生器 140 中產生的。為了能夠產生合適的瞬變，將信號插入器 120 配置為向瞬變信號發生器傳送瞬變描述參數。從而，如項目 141 所示的塊 140 與 120 之間的連接被示為雙向連接。如果在用於操縱的設備中提供特定的瞬變檢測器，那麼可以從該瞬變檢

測器（第一圖中未示出）向瞬變信號發生器 140 提供與瞬變有關的資訊。可以將瞬變信號發生器實現為具有可以直接使用的瞬變採樣或具有可以使用瞬變參數來加權的預先儲存的瞬變採樣，以實際產生/合成將由信號插入器 120 所使用的瞬變。

在一個實施例中，瞬變信號去除器 100 用於從音頻信號中去除第一時間部分，以得到瞬變減小的音頻信號，其中所述第一時間部分包括瞬變事件。

此外，優選地信號處理器用於處理瞬變減小的音頻信號，其中包括瞬變事件的第一時間部分被去除，或用於處理包括瞬變事件的音頻信號，以得到線 111 上的處理後的音頻信號。

優選地，信號插入器 120 用於：在第一時間部分被去除的信號位置，或在瞬變事件位於音頻信號中的信號位置，將第二時間部分插入處理後的音頻信號中，其中第二時間部分包括不受由信號處理器 110 執行的處理所影響的瞬變事件，從而得到輸出 121 處的已操縱音頻信號。

第二圖示出了瞬變信號去除器 100 的優選實施例。在音頻信號不包含與瞬變有關的任何輔助資訊/元資訊(meta information)的一個實施例中，瞬變信號去除器 100 包括瞬變檢測器 103、淡出(fade-out)/淡入(fade-in)計算器 104 以及第一部分去除器 105。在利用如隨後將參考第九圖來討論的編碼設備採集音頻信號中附到音頻信號的與瞬變有關的資訊的可選實施例中，瞬變信號去除器 100 包

括輔助資訊提取器 106，所述輔助資訊提取器 106 提取如線 107 所示附到音頻信號的輔助資訊。如線 107 所示，可以將與瞬變時間有關的資訊提供給淡出/淡入計算器 104。然而當音頻信號包括如元資訊時，不僅瞬變時間，(即出現瞬變事件的精確時間)，而且要從音頻信號排除的部分的開始/停止時間，(即音頻信號“第一部分”的開始時間和停止時間)，都是不需要的，而且也不需要淡出/淡入計算器 104，可以如線 108 所示將開始/停止時間資訊直接轉發給第一部分去除器 105。線 108 示出了選項，而且虛線所示的所有其他線也是可選的。

在第二圖中，優選地淡出/淡入計算器 104 輸出輔助資訊 109。該輔助資訊 109 與第一部分的開始/停止時間不同，這是因為考慮了第一圖的處理器 110 中的處理特性。此外，優選地將輸入音頻信號饋送至去除器 105。

優選地，淡出/淡入計算器 104 提供第一部分的開始/停止時間。這些時間根據瞬變時間計算而得，這樣第一部分去除器 105 不僅去除瞬變事件，還去除瞬變事件周圍的一些採樣。此外，優選的是，不僅利用時域矩形窗切除瞬變部分，還利用淡出部分和淡入部分執行提取。為了執行淡出或/淡入部分，可以應用相對於矩形濾波器而言具有平滑過渡 (smoother transition) 的任何種類的窗，如上升余弦窗，使得這種提取的頻率回應不如應用矩形窗時那樣成問題，儘管這也是選項。這種時域加窗操作輸出加窗操作的殘餘 (remainder)，即，不具有加窗部分 (windowed

portion) 的音頻信號。

在這種情況下可以使用任何瞬變抑制方法，包括在去除瞬變之後留下瞬變減小的或優選地完全非瞬變的殘留信號 (residual signal) 的瞬變抑制方法。與完全去除瞬變部分相比，其中在特定時間部分上將音頻信號設置為 0，瞬變抑制在以下情況下是有利的：由於這種被設置為 0 的部分對於音頻信號而言非常不自然，使得對音頻信號的進一步處理會受到被設置為 0 的部分的影響。

自然地，如結合第九圖所討論的，可以在編碼器側應用由瞬變檢測器 103 和淡出/淡入計算器 104 執行的所有計算，只要將這些計算的結果，如瞬變時間和/或第一部分的開始/停止時間，傳輸至信號操縱器，作為與音頻信號一起或與音頻信號分開的輔助資訊或元資訊，例如在要經由單獨傳輸通道來傳輸的單獨音頻元資料信號內。

第三圖 A 示出了第一圖的信號處理器 110 的優選實現。該實現包括頻率選擇分析器 112 以及後續連接的頻率選擇處理設備 113。實現頻率選擇處理設備 113，使得所述頻率選擇處理設備 113 對原始音頻信號的垂直相干性起到負面影響 (negative influence)。該處理的示例是，在時間上拉伸信號，或在時間上縮短信號，其中以頻率選擇的方式來應用這種拉伸或縮短，使得例如該處理向處理後的音頻信號引入了隨不同頻帶而不同的相移。

在相位聲碼器處理的情況下，在第三圖 B 中示出了一種優選的處理方式。通常，相位聲碼器包括：子帶/變換分

析器 114；隨後連接的處理器 115，用於對專案 114 所提供的多個輸出信號執行頻率選擇性處理；以及隨後的子帶/變換組合器 116，所述子帶/變換組合器 116 將由專案 115 處理的信號相組合以最終在輸出 117 處得到時域中的處理後的信號，由於子帶/變換組合器 116 執行對頻率選擇性信號的組合，使得只要處理後的信號 117 的帶寬大於由專案 115 與 116 之間的單個分支所表示的帶寬，那麼時域中的該處理後的信號就同樣是全帶寬信號或低通濾波後的信號。

隨後結合第五圖 A、第五圖 B、第五圖 C 和第六圖來討論相位聲碼器的其他細節。

隨後，在第四圖中討論並描述了第一圖的信號插入器 120 的優選實現。優選地，信號插入器包括用於計算第二時間部分的長度的計算器 122。在第一圖的信號處理器 110 進行信號處理之前已經去除了瞬變部分的實施例中，為了能夠計算第二時間部分的長度，需要所去除的第一部分的長度以及時間拉伸因數（或時間縮短因數），以便在項目 122 中計算第二時間部分的長度。如結合第一圖和第二圖所討論的，可以從外部來輸入這些資料項目。例如，通過將第一部分的長度乘以拉伸因數來計算第二時間部分的長度。

將第二時間部分的長度轉發給計算器 123，以計算音頻信號中的第二時間部分的第一邊界和第二邊界。具體地，可以將計算器 133 實現為：在不具有在輸出 124 處供

應的瞬變事件的處理後的音頻信號與具有瞬變事件的音頻信號之間執行互相關處理，所述具有瞬變事件的音頻信號提供如在輸入 125 處供應的第二部分。優選地，計算器 123 受另外的控制輸入 126 的控制，使得與稍後將討論的瞬變事件的負移位相比，第二時間部分內瞬變事件的正移位是優選的。

將第二時間部分的第一邊界和第二邊界提供給提取器 127。優選地，提取器 127 切除該部分，即，從輸入 125 處提供的原始音頻信號中切除第二時間部分。因為使用隨後的交叉衰減器 (cross-fader) 128，所以使用矩形濾波器進行切除。在交叉衰減器 128 中，通過對開始部分將權重從 0 增大到 1，和/或在結束部分中將權重從 1 減小到 0，對第二時間部分的開始部分以及第二時間部分的停止部分進行加權，使得在該交叉衰減區域內，處理後的信號的結束部分與所提取的信號的開始部分在相加時產生有用的信號。在提取之後，針對第二時間部分的結束以及處理後的音頻信號的開始，在交叉衰減器 128 中執行類似的處理。交叉衰減保證了不出現時域偽像，否則當不具有瞬變部分的已處理音頻信號的邊界未與第二時間部分邊界完美地匹配在一起時，所述時域偽像將作為滴答聲偽像 (clicking artifact) 被感知。

隨後，參考第五圖 A、第五圖 B、第五圖 C 和第六圖來說明在相位聲碼器的情況下信號處理器 110 的優選實現。

在下文中，參考第五圖和第六圖說明了根據本發明的聲碼器的優選實現。第五圖 A 示出了相位聲碼器的濾波器組實現，其中在輸入 500 處饋入音頻信號，在輸出 510 處得到音頻信號。具體地，第五圖 A 所示的示意性濾波器組中的每個通道包括帶通濾波器 501 和下游（downstream）振盪器 502。利用組合器將來自每個通道的所有振盪器的輸出信號相組合，例如，將所述組合器實現為加法器並且由 503 表示，以得到輸出信號。實現每個濾波器 501，使得濾波器 501 一方面提供幅度信號，另一方面提供頻率信號。幅度信號和頻率信號是時間信號，說明了濾波器 501 中的幅度隨時間的演進，頻率信號表示由濾波器 501 濾波的信號的頻率的演進。

在第五圖 B 中示出了濾波器 501 的示意性設置。可以如第五圖 B 所示來設置第五圖 A 的每個濾波器，然而其中僅供應至兩個輸入混頻器（mixer）551 和加法器 552 的頻率 f_i 隨通道的不同而不同。由低通 553 對混頻器輸出信號進行低通濾波，其中，這些低通信號與在本地振盪器頻率（LO 頻率）所產生的情況下不同，它們是 90° 異相（out of phase）的。上面的低通濾波器 553 提供正交信號 554，而下面的濾波器 553 提供同相信號 555。將這兩個信號（即，I 和 Q）供應至座標變換器 556，所述座標變換器 556 根據矩形表示產生量值（magnitude）相位表示。在輸出 557 處隨時間分別輸出第五圖 A 的量值信號或幅度信號。將相位信號供應至相位展開器（unwrapper）558。在

元件 558 的輸出處，不再存在總是位於 0 至 360° 之間的相位值，而是出現線性增大的相位值。將這種“展開的”相位值供應至相位/頻率轉換器 559，例如可以將所述相位/頻率轉換器 559 實現為簡單的相位差形成器，所述相位差形成器從當前時間點的相位減去先前時間點的相位以得到當前時間點的頻率值。將該頻率值加上濾波器通道 i 的恒定頻率值 f_i ，以在輸出 560 處得到時變頻率值。輸出 560 處的頻率值具有直流分量 $=f_i$ 和交流分量 $=$ 濾波器通道中信號的當前頻率偏離平均頻率 f_i 的頻率偏差 (frequency deviation)。

因此，如第五圖 A 和第五圖 B 所示，相位聲碼器實現了譜資訊與時間資訊的分離。分別地，譜資訊在特定通道中或在為每個通道提供頻率的直流部分的頻率 f_i 中，而時間資訊分別包含在隨時間變化的頻率偏差或量值中。

第五圖 C 示出了根據本發明的、針對帶寬增大而執行的操縱，具體是在聲碼器中，以及在第五圖 A 中以虛線繪製的所示電路位置處執行的操縱。

例如，對於時間縮放，可以對每個通道中的幅度信號 $A(t)$ 或每個信號中的信號頻率 $f(t)$ 進行抽取或插值。出於轉換的目的，由於其對本發明是有用的，因而執行插值，即信號 $A(t)$ 和 $f(t)$ 的時間擴展或延展 (temporal extension or spreading)，以得到延展信號 $A'(t)$ 和 $f'(t)$ ，其中在帶寬擴展情況下該插值受延展因數的控制。通過相位變數 (variation) 的插值，即，加法器 552 加上恒定頻率之前

的值，第五圖 A 中每個獨立振盪器 502 的頻率不變。然而，總體音頻信號的時間變化減慢，即，以因數 2 減慢。得到的結果是具有原始音高（即原始基波（fundamental wave）以及其諧波）的時間延展音調。

通過執行如第五圖 C 所示的信號處理，其中在第五圖 A 的每個濾波器頻段通道中執行這樣的處理，以及通過然後在抽取器中對得到的時間信號進行抽取，音頻信號縮回（shrink back）其原始持續時間，而所有頻率同時加倍。這使得由因數 2 進行音高轉換，然而其中得到了與原始音頻信號具有相同長度（即，相同數目的採樣）的音頻信號。

作為對第五圖 A 所示的濾波器組實現的備選，還可以如第六圖所示來使用相位聲碼器的變換實現。這裏，將音頻信號 100 饋送至 FFT 處理器，或更普遍地饋送至短時傅裏葉變換（Short-Time-Fourier-Transform）處理器 600，作為時間採樣的序列。第六圖中示意性地實現了 FFT 處理器 600，以對音頻信號執行時間加窗（time window），從而隨後通過 FFT 計算譜的量值和相位，其中針對與強交疊的音頻信號塊有關的連續譜來執行該計算。

在極端情況下，可以對於每個新的音頻信號採樣來計算新的譜，其中還可以例如僅針對每 20 個新的採樣來計算新的譜。優選地，這種兩個譜之間的採樣的距離 a 是由控制器 602 給出的。控制器 602 還用於供給 IFFT 處理器 604，所述 IFFT 處理器 604 用於執行交疊操作。具體地，將 IFFT 處理器 604 實現為：通過根據修改後的譜的量值

和相位為每個譜執行一個 IFFT 來執行逆短時傅裏葉變換，以便然後執行疊加操作，其中根據所述疊加操作得到結果時間信號。疊加操作消除了分析加窗的影響。

在利用 IFFT 處理器 604 來處理兩個譜時，利用這兩個譜之間的距離 b 來實現時間信號的延展，所述距離 b 大於在產生 FFT 譜時譜之間的距離 a 。基本思想是，利用比分析 FFT 相隔更遠的逆 FFT 來延展音頻信號。因此，與原始音頻信號相比，合成音頻信號的時間變化出現得更為緩慢。

然而，在塊 606 中沒有相位重縮放的情況下，這將導致偽像。例如，在考慮單個頻率點時，其中針對該頻率點以 45° 間隔實現連續相位值，這意味著該濾波器組內的信號在相位上以 $1/8$ 週期的速率增大，即，每個時間間隔增大 45° ，這裏所述時間間隔是連續 FFT 之間的時間間隔。如果現在使逆 FFT 彼此相隔更遠，則這意味著跨越更長的時間間隔出現 45° 相位增大。這意味著，由於相移，後續疊加過程中出現失配，導致了不期望的信號抵消 (cancellation)。為了消除這種偽像，以實際上相同的因數來重縮放相位，其中利用該因數對音頻信號進行時間延展。從而每個 FFT 譜值的相位以因數 b/a 而增大，使得消除這種失配。

在第五圖 C 所示實施例中，針對第五圖 A 的濾波器組實現中的一個信號振盪器，通過幅度/頻率控制信號的插值來實現延展，而利用兩個 IFFT 之間的距離大於兩個 FFT

譜之間的距離來實現第六圖中的擴展，即， b 大於 a ，然而，其中為了防止偽像，根據 b/a 來執行相位重縮放。

關於相位聲碼器的詳細描述，參考以下文獻：

“The phase Vocoder: A tutorial”, Mark Dolson, Computer Music Journal, vol. 10, no.4, pp. 14—27, 1986, 或
 “New phase Vocoder techniques for pitch-shifting, harmonizing and other exotic effects”, L. Laroche und M. Dolson, Proceedings 1999 IEEE Workshop on applications of signal processing to audio and acoustics, New Paltz, New York, October 17-20, 1999, pages 91 to 94; “New approached to transient processing interphase vocoder”, A. Röbel, Proceeding of the 6th international conference on digital audio effects (DAFx-03), London, UK, September 8-11, 2003, pages DAFx-1 to DAFx-6; “Phase-locked Vocoder”, Meller Puckette, Proceedings 1995, IEEE ASSP, Conference on applications of signal processing to audio and acoustics, 或美國專利申請號 6,549,884.

可選地，其他信號延展方法是可用的，例如，“音高同步疊加”方法。音高同步疊加（簡稱 PSOLA）是一種合成方法，在該方法中語言信號的記錄位於資料庫中。只要這些信號是週期信號，就為其提供與基頻（音高）有關的資訊並且標記每個週期的開始。在合成中，利用窗函數以特定的環境來切除這些週期，並將它們添加到要合成的信號中合適的位置：根據所期望的基頻是高於還是低於資料

庫條目的基頻，相應地比原始更密集或更稀疏地組合它們。為了調整可聽的持續時間，該週期可以被省略或雙倍輸出。該方法還稱作 TD-PSOLA，其中 TD 代表時域，並強調方法在時域中操作。另外的發展是多頻段再合成疊加（multiband resynthesis overlap add）方法，簡稱 MBROLA。這裏通過預處理使資料庫中的片段達到統一的基頻，並將諧波的相位位置歸一化（normalize）。這樣，在從一個片段到另一片段的瞬變的合成中，產生更少的感知性干擾，並且所實現的語言品質更高。

在另外的備選方案中，在延展之前已經對音頻信號進行帶通濾波，使得延展和抽取後的信號已經包含期望的部分，並且可以省略隨後的帶通濾波。這樣，設置帶通濾波器，使得帶通濾波器的輸出信號中仍然包含可能在帶寬擴展之後已經濾除的音頻信號部分。從而帶通濾波器包含了在延展和抽取之後的音頻信號中並未包含的頻率範圍。具有該頻率範圍的信號是形成合成高頻信號的所需信號。

如第一圖所示的信號操縱器還可以額外包括信號調節器 130，用於對線 121 上具有未處理的“自然的”或合成的瞬變的音頻信號進行進一步處理。該信號調節器可以是帶寬擴展應用中的信號抽取器，所述信號抽取器在其輸出處產生高頻段信號，然後通過使用要與 HFR（高頻重建）資料流程一起傳輸的高頻（HF）參數來進一步調節（adapt）所述高頻段信號，以使其非常類似原始高頻段信號的特性。

第七圖 A 和第七圖 B 示出了帶寬擴展方案，有利地，該方案可以使用第七圖 B 的帶寬擴展編碼器 720 內的信號調節器的輸出信號。將音頻信號饋送至輸入 700 處的低通/高通組合中。低通/高通組合一方面包括低通 (LP)，產生音頻信號 700 的低通濾波版本，如第七圖 A 中的 703 所示。採用音頻編碼器 704 對該低通濾波後的音頻信號進行編碼。例如，音頻編碼器是 MP3 編碼器 (MPEG1 層 3) 或 AAC 編碼器，還稱作 MP4 編碼器，如在 MPEG4 標準中描述的。在編碼器 704 中可以使用提供頻段受限音頻信號 703 的透明 (transparent) 表示或有利地為感知性透明表示的備選音頻編碼器，以分別產生完全編碼的或感知性編碼的、(優選為感知性透明編碼的音頻信號 705。

濾波器 702 的高通部分 (表示為“HP”) 在輸出 706 處輸出音頻信號的上頻段 (upper band)。將音頻信號的高通部分，即，也表示為 HF 部分的上頻段或 HF 頻段，供應至用於計算不同參數的參數計算器 707。例如，這些參數是在相對粗糙解析度下上頻段 706 的譜包絡，例如，分別針對每個心理聲學 (psychoacoustic) 頻率組或針對 Bark 尺度 (scale) 上每個 Bark 頻段的尺度因數的表示。參數計算器 707 可以計算的另外的參數是上頻段中的雜訊基底，其每頻段能量可以優選地與該頻段中包絡的能量有關。參數計算器 707 可以計算的其他參數包括針對上頻段的每個局部 (partial) 頻段的音調測量 (tonality measure)，其指示譜能量如何在頻段中分佈，即，譜能量是否相對均

勻地分佈在頻段中（其中，那麼該頻段中存在非音調信號），或該頻段中的能量是否相對強烈地集中在頻段中的特定位置（其中，那麼相反，該頻段存在音調信號）。

其他參數包括：對上頻段中在其高度和其頻率方面相對強烈地突出的峰值的顯式（explicitly）編碼，在未對上頻段中顯著的正弦部分進行這種顯式編碼的重建中，帶寬擴展構思只會非常基本地或根本不恢復相同的信號。

在任何情況下，參數計算器 707 用於僅產生針對上頻段的參數 708，其中，可以對所述參數 708 執行類似的熵減小步驟，因為還可以在音頻編碼器 704 中針對量化的頻譜值來執行這些步驟，例如差分編碼、預測或霍夫曼編碼等。然後將參數表示 708 和音頻信號 705 供應至用於提供輸出輔助資料流程 710 的資料流程格式器 709，典型地，所述輸出輔助資料流程 710 是具有特定格式的位元流，如在 MPEG4 標準中標準化的格式。

因為尤其適於本發明，所以以下參考第七圖 B 對解碼器側進行說明。資料流程 710 進入資料流程解釋器（interpreter）711，所述資料流程解釋器 711 用於將與帶寬擴展有關的參數部分 708 與音頻信號部分 705 分開。利用參數解碼器 712 對參數部分 708 進行解碼，以得到解碼後的參數 713。與此並行地，利用音頻解碼器 714 對音頻信號部分 705 進行解碼，以得到音頻信號。

根據該實現，可以經由第一輸出 715 輸出音頻信號 100。在輸出 715 處，然後可以得到具有小帶寬從而具有

低品質的音頻信號。然而，為了提高品質，執行本發明的帶寬擴展 720，以分別在輸出側得到具有擴展或高帶寬從而具有高品質的音頻信號 712。

根據 WO 98/57436 已知，在編碼器側對音頻信號執行頻段限制，並利用高品質的音頻編碼器僅對音頻信號的低頻段進行編碼。然而，僅非常粗糙地（即，利用再現上頻段的譜包絡的一組參數）描述上頻段的特徵。然後，在解碼器側合成上頻段。為此，提出諧波轉換，其中，將解碼後的音頻信號的下頻段供應至濾波器組。下頻段的濾波器組通道與上頻段的濾波器組通道連接，或“拼湊（patch）”下頻段的濾波器組通道，對每個拼湊的帶通信號進行包絡調節。這裏屬於特定分析濾波器組的合成濾波器組接收下頻段中的音頻信號的帶通信號，並接收下頻段的包絡調節後的帶通信號，該信號在上頻段中諧波地（harmonically）被拼湊。合成濾波器組的輸出信號是在其帶寬方面被擴展的音頻信號，以很低的資料速率從編碼器側向解碼器側傳輸該音頻信號。具體地，濾波器組領域中的濾波器組計算以及拼湊可能變得需要很大的計算量。

這裏所提出的方法解決了所提出的問題。與現有方法相比，本方法的新穎之處在於，從要操縱的信號中去除包含瞬變的加窗部分，以及還從原始信號中額外選擇出第二加窗部分（通常與第一部分不同），其中還可以將所述第二加窗部分重新插入受操縱信號中，以便在瞬變的環境下盡可能多地保留時間包絡。選擇所述第二部分，使得該第

二部分會精確適合被時間拉伸操作所改變的凹處 (recess)。通過計算所得到的凹處的邊沿與原始瞬變部分的邊沿的最大互相關，來執行所述精確適合。

因此，瞬變的主觀音頻品質不再被分散 (dispersion) 或回聲效應削弱。

為了選擇合適部分，例如，可以通過在合適的時間段上進行能量的移動質心 (moving centroid) 計算，來精確地確定瞬變的位置。

第一部分的大小與時間拉伸因數一起確定了第二部分的所需大小。優選地，將選擇該大小，使得第二部分容納多於一個的瞬變，只有在彼此緊鄰的瞬變之間的時間間隔低於人類感知獨立時間事件的閾值的情況下，所述第二部分才會用於重新插入。

根據最大互相關對瞬變的最優適合可能需要相對於該瞬變原始位置的微小時間偏移。然而，由於存在時間前掩蔽 (pre-masking) 效應以及特別是後掩蔽 (post-masking) 效應，重新插入的瞬變的位置不需要與原始位置精確匹配。由於後掩蔽動作的擴展週期，所以瞬變在正時間方向上的移位是優選的。

通過插入原始信號部分，在隨後的抽取步驟改變採樣速率的情況下，其音色 (timbre) 或音高將發生改變。然而這通常被瞬變自身通過心理聲學時間掩蔽機制所掩蔽。具體地，如果出現以整數因數進行的拉伸，則音色只會發生微小改變，因為在瞬變環境外部只會佔用每第 n 個

(n =拉伸因數) 諧波。

使用新的方法，有效防止了在通過時間拉伸和轉換方法處理瞬變的過程中產生的偽像（分散、前回聲和後回聲）。避免了對疊加的（可能是音調）信號部分的品質的潛在削弱。

本方法適於其中音頻信號的再現速度或它們的音高將發生改變的任何音頻應用。

隨後，將根據第八圖 A 至第八圖 E 來討論優選實施例。第八圖 A 示出了音頻信號的表示，然而與直向前（straight forward）時域音頻採樣序列不同，第八圖 A 示出了能量包絡表示，所述能量包絡表示例如是通過對時域採樣圖例中的每個音頻採樣求平方而得到的。具體地，第八圖 A 示出了具有瞬變事件 801 的音頻信號 800，其中瞬變事件的特徵在於能量隨時間的急劇增大或減小。自然地，瞬變還可以是：當能量保持在特定高度時，該能量的急劇升高；或當能量在下降之前已經在特定高度保持了特定時間時，該能量的急劇降低。例如，瞬變的具體形式是，掌聲或由打擊工具產生的任何其他音調。此外，瞬變是工具的快速擊打，其開始大聲播放音調，即，在特定閾值級別以上特定閾值時間以下將聲音能量提供到特定頻帶中或多個頻帶中。自然地，其他能量波動，如第八圖 A 中的音頻信號 800 的能量波動 802 未被檢測為瞬變。瞬變檢測器是現有技術中已知的，並且在文獻中被廣泛描述，其依賴於許多不同的演算法，所述演算法可以包括：頻率選擇

性處理，以及將頻率選擇性處理的結果與閾值相比較，以及隨後確定是否存在瞬變。

第八圖 B 示出了加窗瞬變。從利用所示窗形狀加權的信號中減去實線限定的區域。在處理之後，再次添加由虛線標記的區域。具體地，必須從音頻信號 800 中切除在特定瞬變時間 803 出現的瞬變。穩妥起見，不僅要從原始信號中切除瞬變，還要切除一些相鄰/鄰近採樣。從而，確定第一時間部分 804，其中第一時間部分從開始時刻 805 延伸至停止時刻 806。通常，選擇第一時間部分 804，使得瞬變時間 803 包含在第一時間部分 804 內。第八圖 C 示出了拉伸之前沒有瞬變的信號。從緩慢衰落 (slowly-decaying) 的邊沿 807 和 808 可以看出，不僅通過矩形濾波器/加窗器 (windower) 來切除第一時間部分，還執行加窗以使音頻信號具有緩慢衰落的邊沿或側邊 (flank)。

重要的是，第八圖 C 示出了第一圖的線 102 上的音頻信號，即，在瞬變信號去除之後的音頻信號。緩慢衰落/升高的側邊 807、808 提供了由第四圖的交叉衰減器 128 使用的淡入或淡出區域。第八圖 D 示出了第八圖 C 的信號，然而是以拉伸後的狀態示出的，即，在信號處理器 110 進行處理之後。因此，第八圖 D 中的信號是第一圖的線 111 上的信號。由於拉伸操作使得第一部分 804 變得更長。因此，第八圖 D 的第一部分 804 被拉伸到了第二時間部分 809，所述第二時間部分 809 具有第二時間部分起始時刻

810 和第二時間部分停止時刻 811。通過拉伸信號，還拉伸了側邊 807、808，從而拉伸了側邊 807'、808' 的時間長度。如第四圖的計算器 122 所執行的，當對第二時間部分的長度進行計算時，說明了該拉伸。

如第八圖 B 中的虛線所示，一旦確定了第二時間部分的長度，就從第八圖 A 所示的原始音頻信號中切除與第二時間部分的長度相對應的部分。這樣，第二時間部分 809 進入了第八圖 E。如所述的，第二時間部分的起始時刻 812（即，原始音頻信號中第二時間部分 809 的第一邊界）與第二時間部分的停止時刻 813（即，原始音頻信號中第二時間部分的第二邊界）不必須相對於瞬變事件時間 803、803' 而對稱以使瞬變 801 精確位於與其在原始信號中相同的時刻上。相反，第八圖 B 的時刻 812、813 可以有微小變化，使得原始信號中這些邊界上的信號形狀之間的互相關結果盡可能地與拉伸後的信號中相應的部分相類似。從而，可以將瞬變 803 的實際位置移出第二時間部分的中央，直到如第八圖 E 中由參考數字 803' 所指示的特定程度為止，參考數字 803' 指示相對於第二時間部分的特定時間，其偏離了相對於第八圖 B 中的第二時間部分的對應時間 803。如結合第四圖所述，瞬變相對於時間 803 向時間 803' 的正位移是優選的，這歸因於比前掩蔽效應更為顯著（pronounced）的後掩蔽效應。第八圖 E 還示出了交迭（crossover）/過渡區域 813a、813b，在所述交迭/過渡區域 813a、813b 中，交叉衰減器 128 提供不具有瞬變的拉

伸信號與包括瞬變的原始信號副本之間的交叉衰減器。

如第四圖所示，用於計算第二時間部分 122 的長度的計算器被配置為接收第一時間部分的長度以及拉伸因數。可選地，計算器 122 還可以接收與鄰近瞬變包含在同一個第一時間部分中的容許性（allowability）有關的資訊。因此，根據該容許性，計算器可以獨立地確定第一時間部分 804 的長度，然後根據拉伸/縮短因數來計算第二時間部分 809 的長度。

如以上所述，信號插入器的功能在於，該信號插入器從原始信號中去除針對第八圖 E 的間隙（gap）的合適區域（其在拉伸後的信號內被擴大），並使用互相關計算使該合適區域（即，第二時間部分）適合處理過的信號以確定時刻 812 和 813，以及優選地還在交叉衰減區域 813a 和 813b 中執行交叉衰減操作。

第九圖示出了用於產生音頻信號的輔助資訊的設備，當在編碼器側執行瞬變檢測，並且計算出關於該瞬變檢測的輔助資訊並將其傳輸至然後將表示解碼器側的信號操縱器時，該設備可以用在本發明的情況下。這樣，應用與第二圖中的瞬變檢測器 103 相類似的瞬變檢測器來分析包含瞬變事件的音頻信號。瞬變檢測器計算瞬變時間，即，第一圖中的時間 803，並且將該瞬變時間轉發至元資料計算器 104'，可以將所述元資料計算器 104' 構造為類似於第二圖中的淡出/淡入計算器 104'。通常，元資料計算器 104' 可以計算要轉發至信號輸出介面 900 的元資料，其

中該元資料可以包括：針對瞬變去除的邊界，即，針對第一時間部分的邊界，即，第八圖 B 中的邊界 805 和 806，或如第八圖 B 中 812、813 所示的針對瞬變插入（第二時間部分）的邊界，或瞬變事件時刻 803 或甚至 803'。即使在後一種情況下，信號操縱器將能夠根據瞬變事件時刻 803 來確定所有所需資料，即，第一時間部分資料、第二時間部分資料等。

將如專案 104' 所產生的元資料轉發至信號輸出介面，使得信號輸出介面產生信號，即，用於傳輸或儲存的輸出信號。輸出信號可以僅包括元資料或可以包括元資料和音頻信號，其中，在後一種情況下，元資料將表示音頻信號的輔助資訊。這樣，可以經由線 901 將音頻信號轉發至信號輸出介面 900。可以將信號輸出介面 900 所產生的輸出信號儲存在任何類型的儲存介質上，或經由任何種類的傳輸通道傳輸至信號操縱器或需要瞬變資訊的任何其他設備。

將注意的是，儘管以方框圖的形式描述了本發明，其中方框表示實際的或邏輯的硬體元件，然而還可以通過電腦實現的方法來實現本發明。在後一種情況下，方框表示相應的方法步驟，其中這些步驟代表由相應的邏輯或物理硬體模組所執行的功能。

所述實施例僅僅是為了說明本發明的原理。應理解，對這裏所述的佈置和細節的修改和改變對於本領域技術人員而言顯而易見的。因此，意圖在於，僅受限於所附申

請專利範圍的範圍，而不受限於這裏以對實施例的描述和解釋的方式而表現的特定細節。

取決於本發明方法的特定實現要求，可以採用硬體或軟體的形式來實現本發明的方法。可以使用數位儲存介質來執行所述實現，所述數位儲存介質具體可以是磁片、儲存有電可讀控制信號的 DVD 或 CD，它們與可編程電腦系統協作以執行本發明的方法。通常，因而可以將本發明實現為電腦程式產品，具有儲存在機器可讀載體上的程式碼，用於當電腦程式產品在電腦上運行時執行本發明的方法。換言之，本發明的方法從而是具有程式碼的電腦程式，所述程式碼用於當所述電腦程式在電腦上運行時執行本發明的方法中至少一個方法。本發明的元資料信號可以儲存在任何機器可讀的儲存介質上，如數位儲存介質。

【圖式簡單說明】

第一圖示出了本發明的用於操縱具有瞬變的音頻信號的設備或方法的優選實施例；

第二圖示出了第一圖的瞬變信號去除器的優選實現；

第三圖 A 示出了第一圖的信號處理器的優選實現；

第三圖 B 示出了實現第一圖的信號處理器的另外優選實施例；

第四圖示出了第一圖的信號插入器的優選實現；

第五圖 A 示出了在第一圖的信號處理器中使用的聲碼器的實現的概圖；

第五圖 B 示出了第一圖的信號處理器的一部分（分析）的實現；

第五圖 C 示出了第一圖的信號處理器的其他部分（拉伸）；

第六圖示出了在第一圖的信號處理器中使用的相位聲碼器的變換實現；

第七圖 A 示出了帶寬擴展處理方案的編碼器側；

第七圖 B 示出了帶寬擴展方案的解碼器側；

第八圖 A 示出了具有瞬變事件的音頻輸入信號的能量表示；

第八圖 B 示出了具有加窗瞬變（windowed transient）的第八圖 A 的信號；

第八圖 C 示出了拉伸之前沒有瞬變部分的信號；

第八圖 D 示出了拉伸之後第八圖 C 的信號；以及

第八圖 E 示出了在插入了原始信號的相應部分之後的受操縱信號。

第九圖示出了用於針對音頻信號產生輔助資訊的設備。

【主要元件符號說明】

瞬變信號去除器 100

輸入 101

輸出 102

瞬變檢測器 103

淡出/淡入計算器 104

第一部分去除器 105

輔助資訊提取器 106

信號處理器 110

信號處理器輸出 111

頻率選擇分析器 112

頻率選擇處理設備 113

子帶/變換分析器 114

處理器 115

子帶/變換組合器 116

信號插入器 120

信號插入器輸出 121

計算器 122、123

提取器 127

在交叉衰減器 128
 信號調節器 130
 瞬變信號發生器 140
 輸入 500
 帶通濾波器 501
 下游振盪器 502
 加法器 503
 輸出 510
 輸入混頻器 551
 加法器 552
 低通 553
 正交信號 554
 同相信號 555
 座標變換器 556
 輸出 557
 相位展開器 558
 相位/頻率轉換器 559
 輸出 560
 FFT 處理器 600
 控制器 602
 IFFT 處理器 604
 輸入 700
 編碼器 704
 參數計算器 707

資料流程格式器 709

資料流程解釋器 711

參數解碼器 712

參數 713

音頻解碼器 714

帶寬擴展編碼器 720

音頻信號 800

瞬變事件 801

能量波動 802

信號輸出介面 900

七、申請專利範圍：

1、一種用於操縱具有瞬變事件（801）的音頻信號的設備，包括：

信號處理器（110），用於處理瞬變減小的音頻信號，或用於處理包括瞬變事件（803）的音頻信號，以得到處理後的音頻信號，在所述瞬變減小的音頻信號中，包括瞬變事件（801）的第一時間部分（804）被去除了；

信號插入器（120），用於在信號位置處將第二時間部分（809）插入處理後的音頻信號中，所述信號位置是第一部分被去除的信號位置或瞬變事件在處理後的音頻信號中所處的信號位置，其中第二時間部分（809）包括不受信號處理器（110）執行的處理的影響的瞬變事件（801），以得到受操縱的音頻信號，

其中，所述信號處理器（110）執行對瞬變減小的音頻信號的拉伸，使第一時間部分（804）被拉伸至第二時間部分（809），第二時間部分（809）在時間上比第一時間部分（804）長，以及

所述信號插入器（120）被配置為：複製包括瞬變事件的音頻信號的部分（809）以及瞬變事件之前或之後的信號部分，使得所述瞬變事件之前或之後的信號部分與所述第一部分一共具有第二部分（809）的持續時間；以及在處理後的音頻信號中插入未修改的副本，或插入其中僅起始部分（813）或結尾部分（813b）被修改過的、包括瞬變的信號的副本。

2、依據申請專利範圍第 1 項所述的設備，還包括：瞬變信號去除器（100），用於從音頻信號中去除第一時間部分（804），以得到瞬變減小的音頻信號，所述第一時間部分（804）包括瞬變事件（801）。

3、依據申請專利範圍第 1 或 2 項所述的設備，其中，所述信號處理器（110）被配置為以基於頻率的方式（112，113）來處理瞬變減小的音頻信號，使得該處理向瞬變減小的音頻信號中引入隨不同的譜分量而有所不同的相移。

4、依據申請專利範圍第 1 項所述的設備，其中，所述信號插入器（120）被配置為通過複製至少第一時間部分（804）來產生第二時間部分，使得第二時間部分至少包括來自具有瞬變事件的音頻信號的第一時間部分的副本。

5、依據申請專利範圍第 1 項所述的設備，其中，所述信號插入器（120）被配置為確定第二部分（809），使得所述第二部分在第二時間部分的起始或結尾處與處理後的音頻信號具有交疊，以及所述信號插入器（120）被配置為在處理後的音頻信號與第二時間部分之間的邊界處執行交叉衰減（128）。

6、依據申請專利範圍第 1 項所述的設備，其中，所述信號處理器包括聲碼器、相位聲碼器、或(P)SOLA 處理器。

7、依據申請專利範圍第 1 項所述的設備，還包括信號調節器（130），用於通過對受操縱音頻信號的時間離散

版本進行抽取或插值來調節所述受操縱音頻信號。

8、依據申請專利範圍第 1 項所述的設備，其中，所述信號插入器（120）被配置為：

確定（122）要從具有瞬變事件的音頻信號複製的第二時間部分（809）的時間長度，

優選地通過找到最大互相關計算來確定（123）第二時間部分的起始時刻或第二時間部分的停止時刻，使得優選地第二時間部分的邊界盡可能地與處理後的音頻信號的相應邊界相匹配，

其中，受操縱音頻信號中瞬變事件的時間位置（803'）與音頻信號中瞬變事件的時間位置（803）一致，或與音頻信號中瞬變事件的時間位置（803）偏離小於心理聲學可承受程度的時間差，所述心理聲學可承受程度由瞬變事件的前掩蔽或後掩蔽來確定。

9、依據申請專利範圍第 1 項所述的設備，還包括瞬變檢測器（103），用於檢測音頻信號中的瞬變事件，或

還包括輔助資訊提取器（106），用於提取並解釋與音頻信號相關聯的輔助資訊，所述輔助資訊指示瞬變事件的時間位置（803），或指示第一時間部分或第二時間部分的起始時刻或停止時刻。

10、一種操縱具有瞬變事件（801）的音頻信號的方法，包括：

處理（110）瞬變減小的音頻信號，或處理包括瞬變事件（803）的音頻信號，以得到處理後的音頻信號，在

所述瞬變減小的音頻信號中，包括瞬變事件（801）的第一時間部分（804）被去除了；

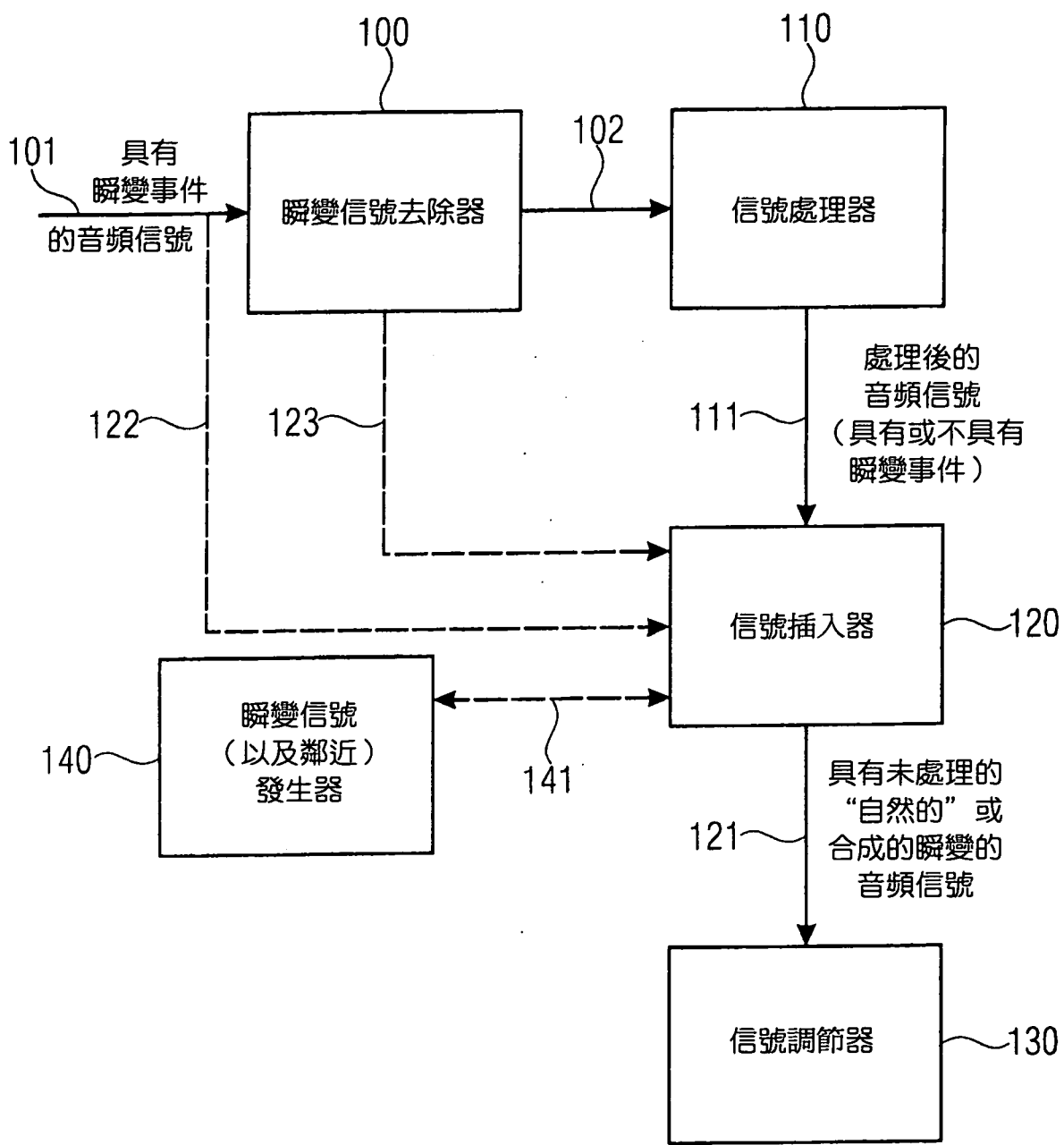
在信號位置處將第二時間部分（809）插入（120）處理後的音頻信號中，所述信號位置是第一部分被去除的信號位置，或瞬變事件在處理後的音頻信號中所處的信號位置，其中第二時間部分（809）包括不受所述處理影響的瞬變事件（801），以得到受操縱的音頻信號，

其中，處理（110）信號的步驟包括執行對瞬變減小的音頻信號的拉伸，使第一時間部分（804）被拉伸至第二時間部分（809），第二時間部分（809）在時間上比第一時間部分（804）長，以及

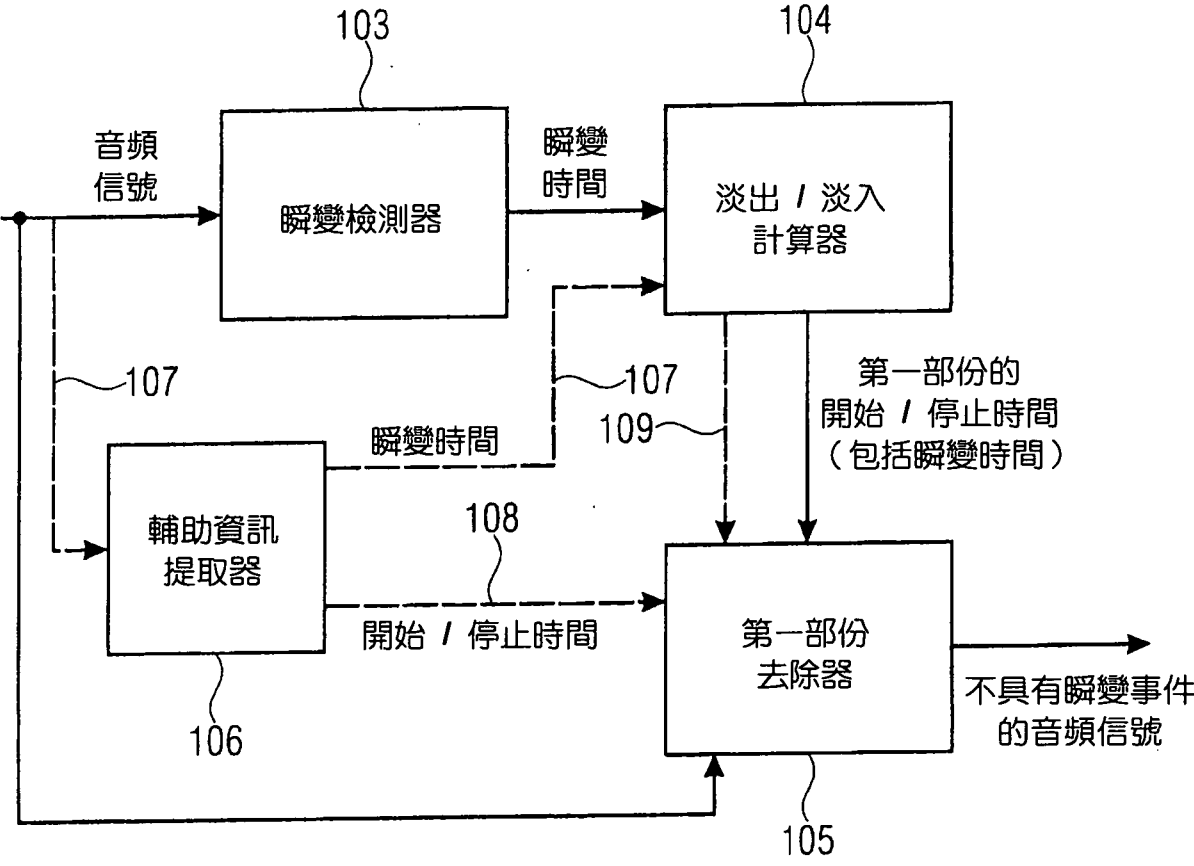
插入（120）步驟包括：複製包括瞬變事件的音頻信號的部分（809）以及瞬變事件之前或之後的信號部分，使得所述瞬變事件之前或之後的信號部分與所述第一部分一共具有第二部分（809）的持續時間；以及在處理後的音頻信號中插入未修改的副本，或插入其中僅起始部分（813）或結尾部分（813b）被修改過的、包括瞬變的信號的副本。

11、一種具有程式碼的電腦程式，當所述電腦程式運行在電腦上時，所述程式碼執行依據申請專利範圍第 10 項所述的方法。

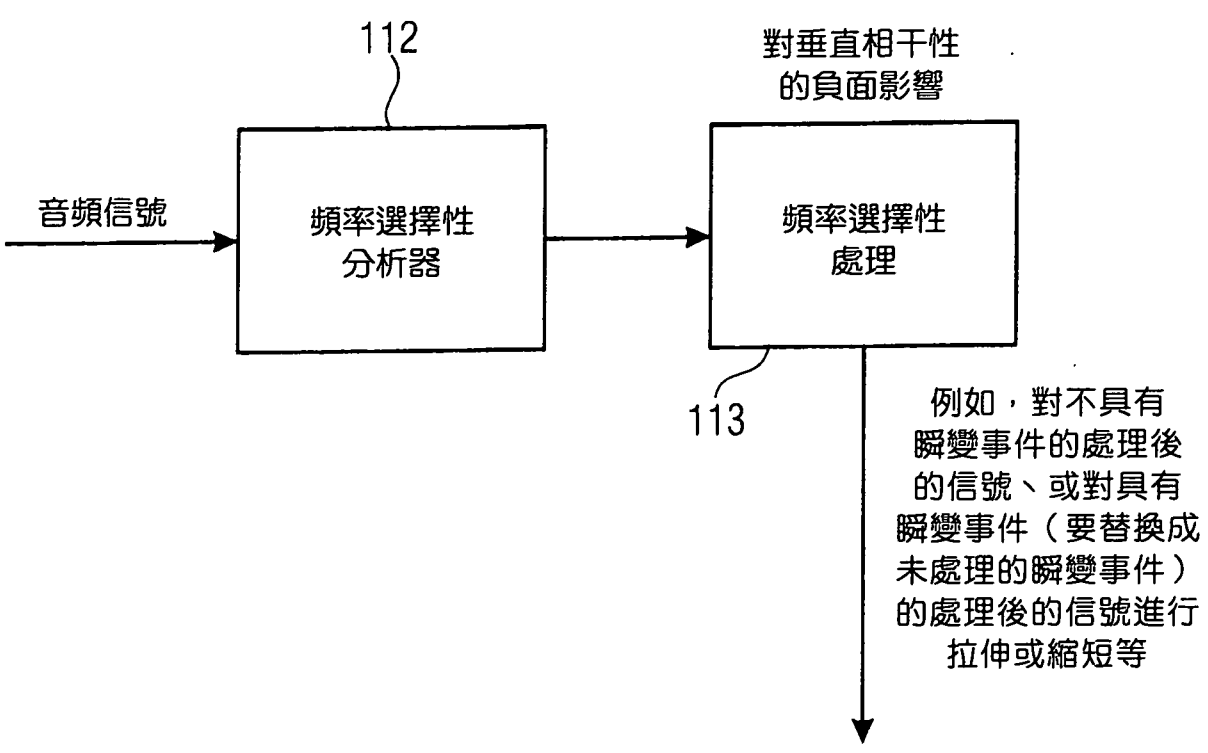
八、圖式：



第一圖

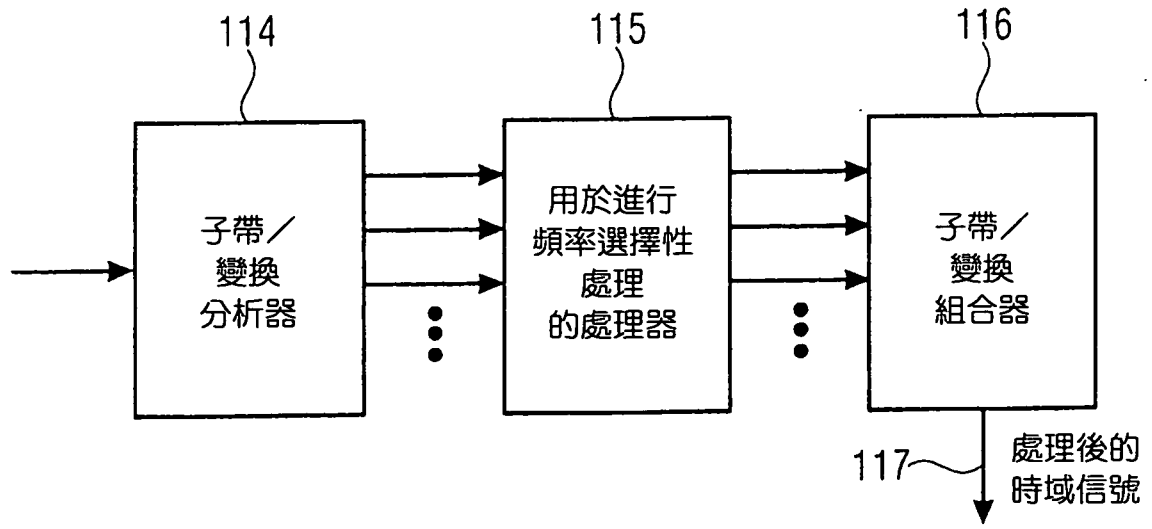


第二圖

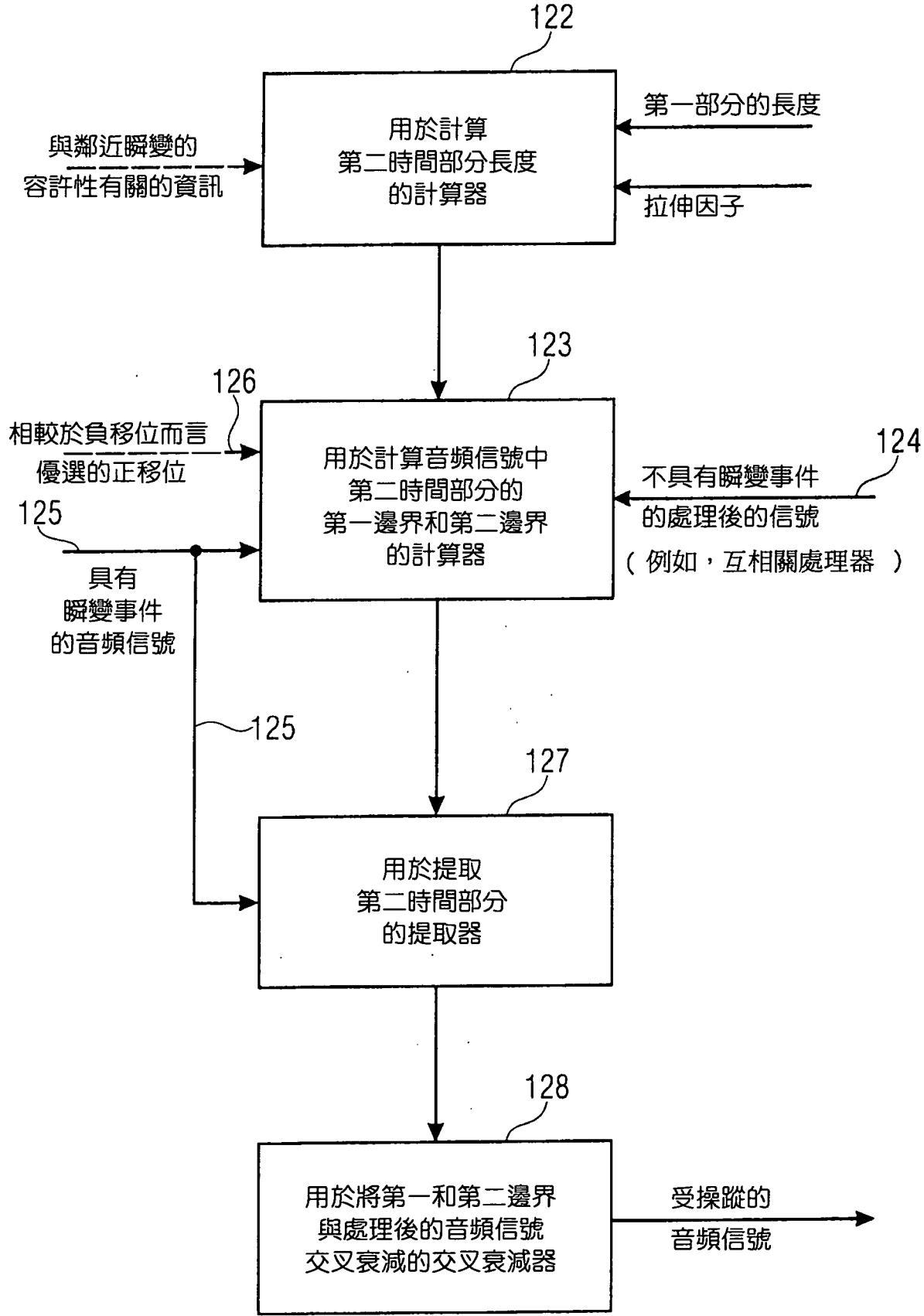


第三圖 A

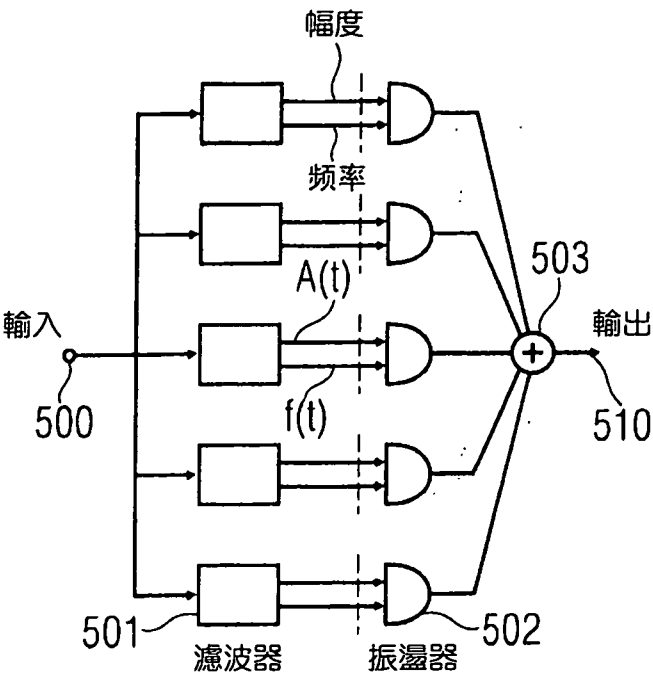
例如，相位聲碼器處理



第三圖 B

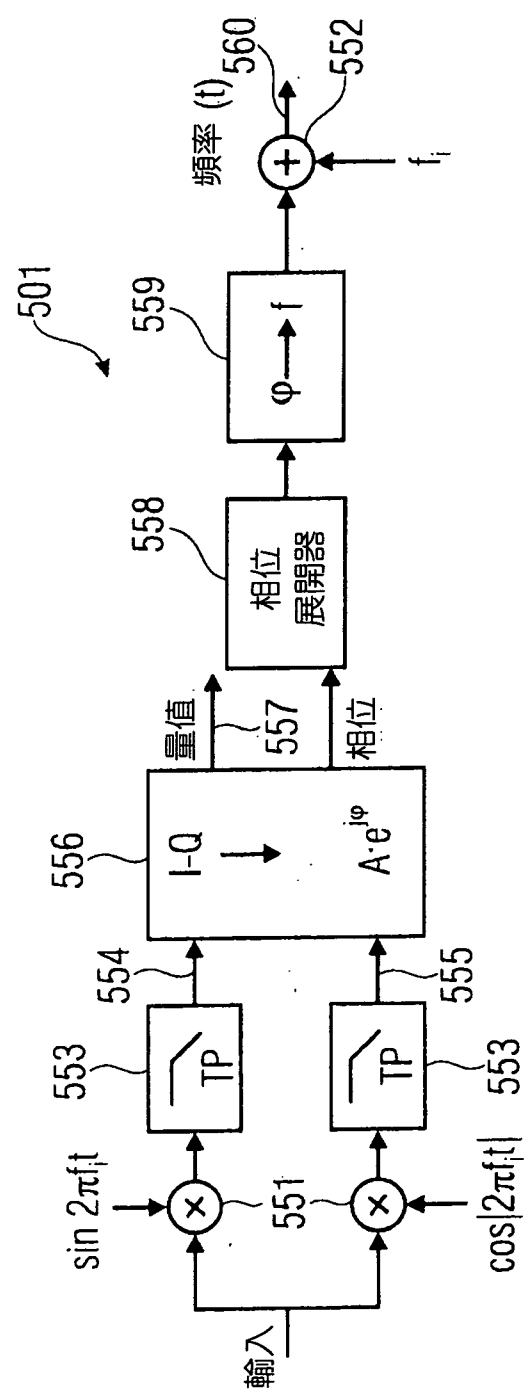


第四圖

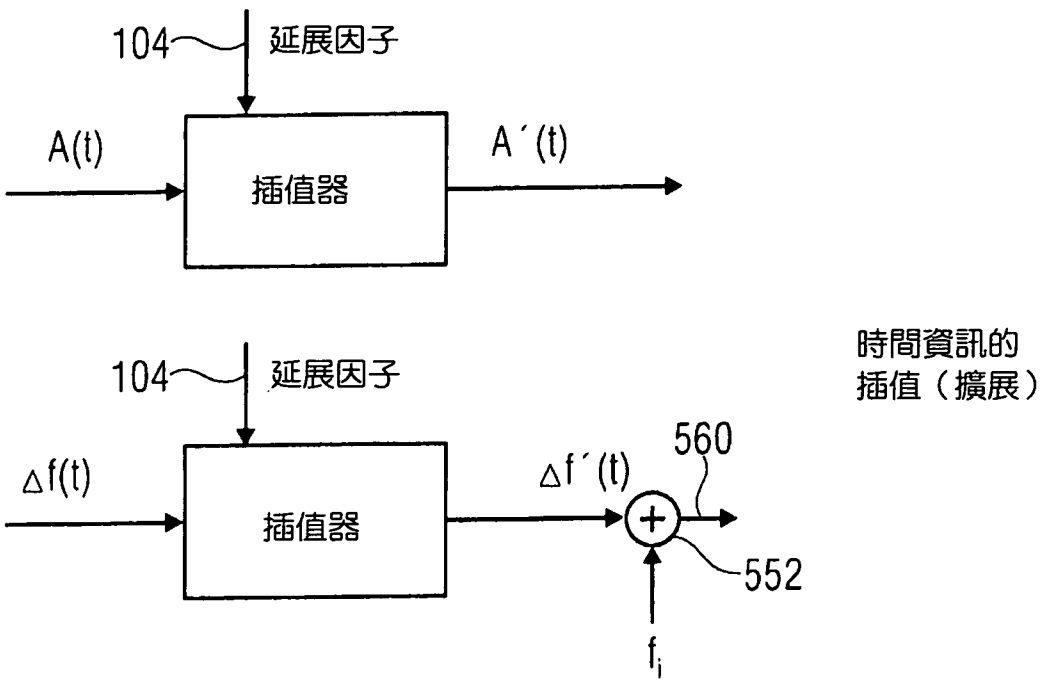


譜資訊與時間資訊
的分離

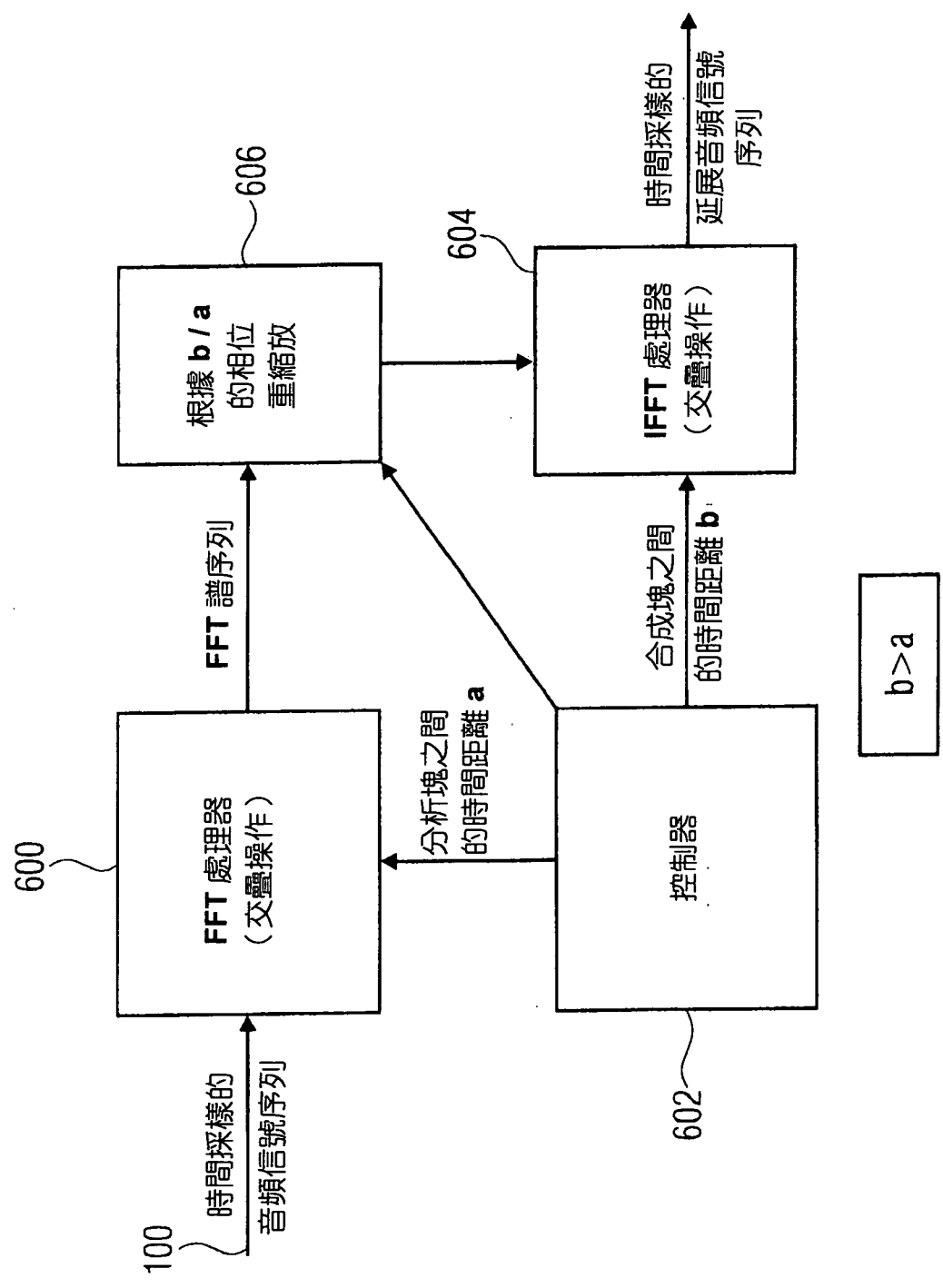
第五圖 A
(濾波器組實現)



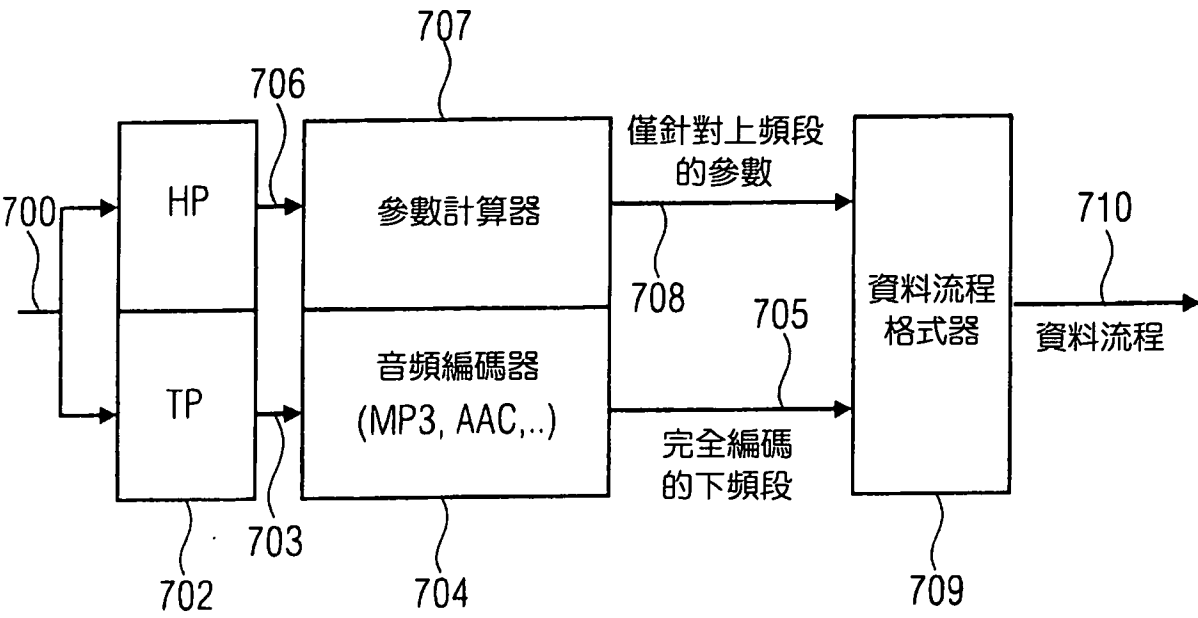
第五圖 B



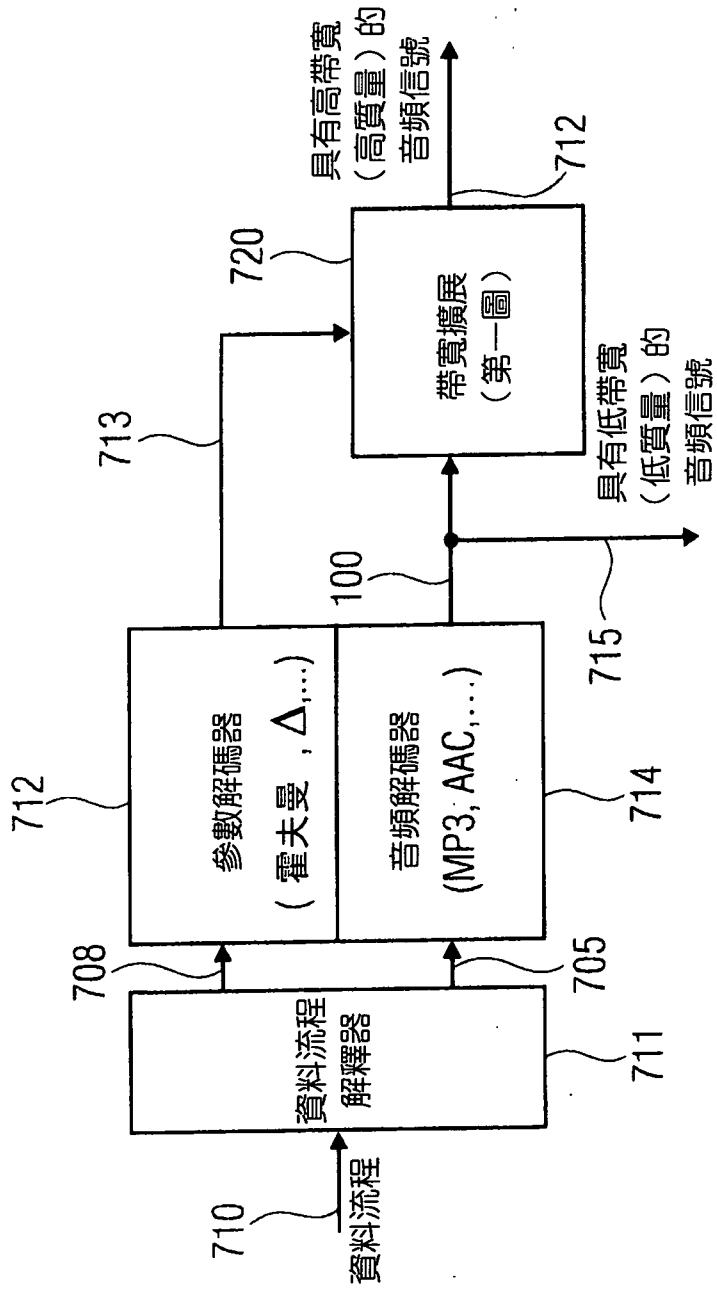
第五圖 C



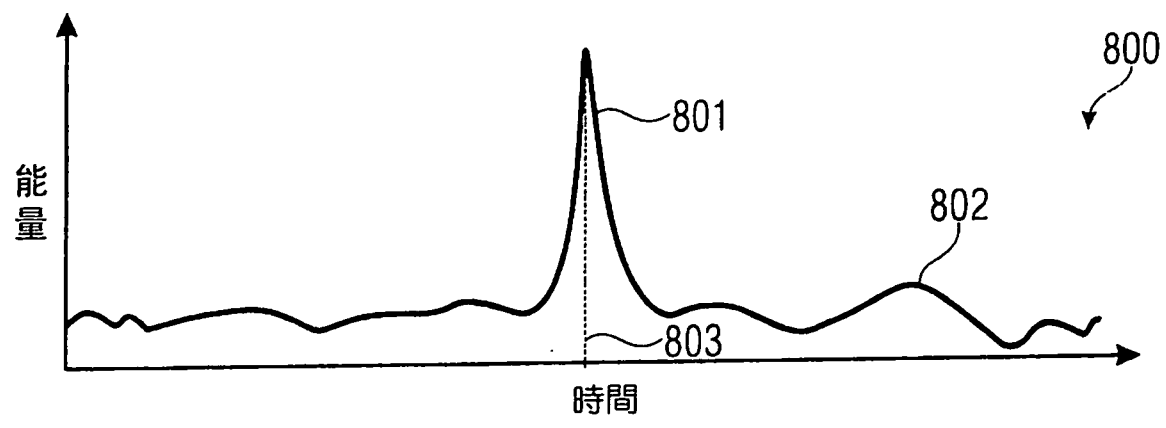
第六圖
(變換實現)



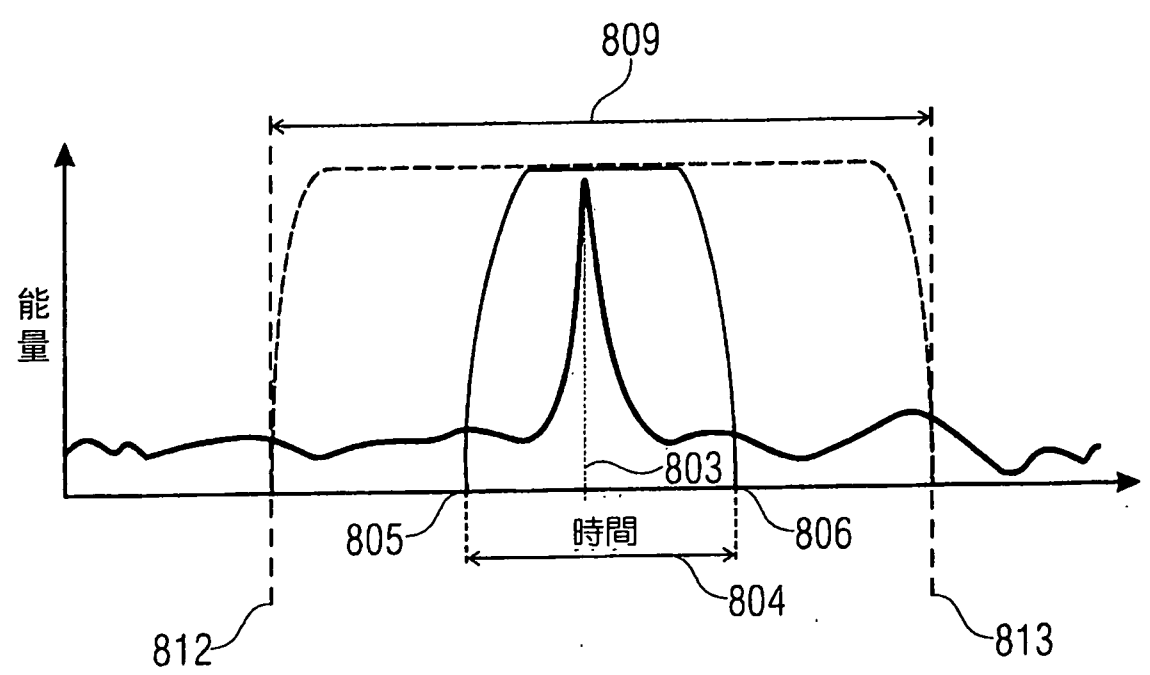
第七圖 A
(編碼器側)



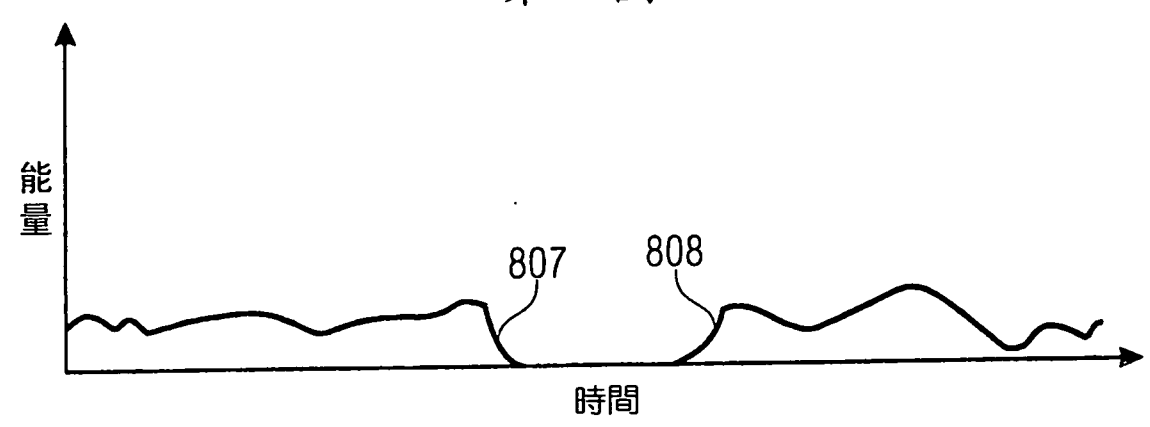
第七圖 B
(解碼器側)



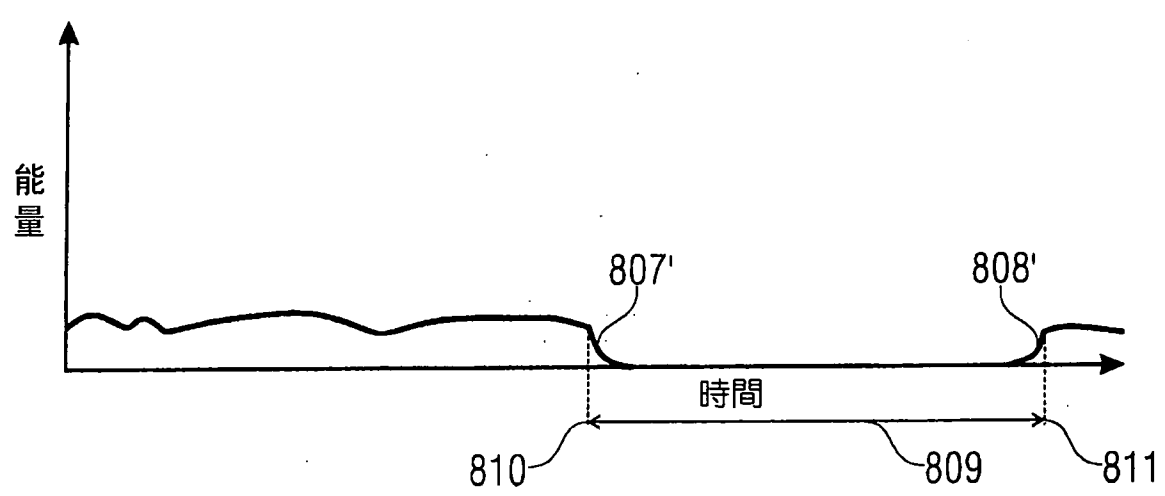
第八圖 A



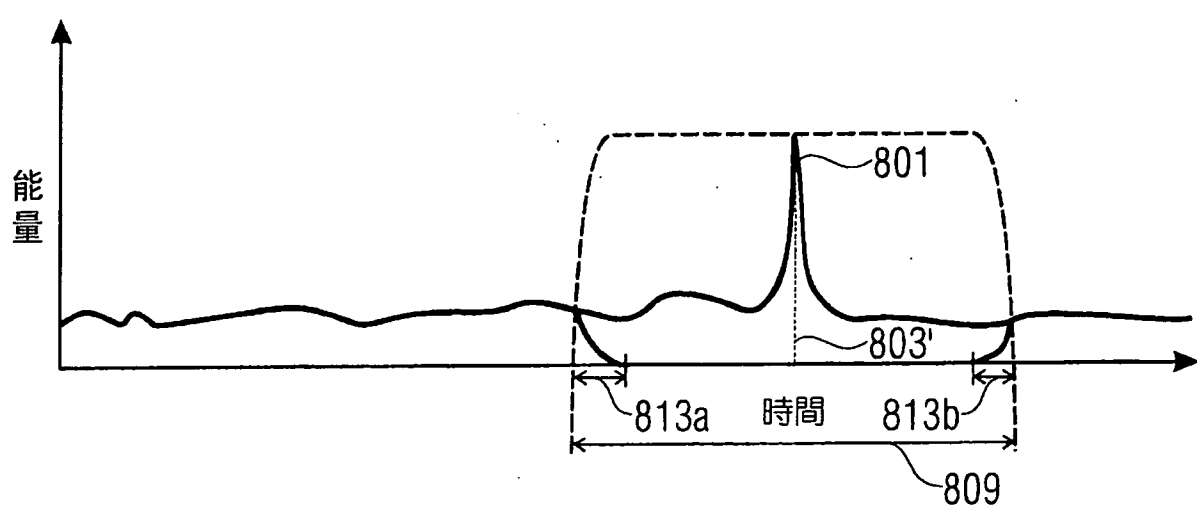
第八圖 B



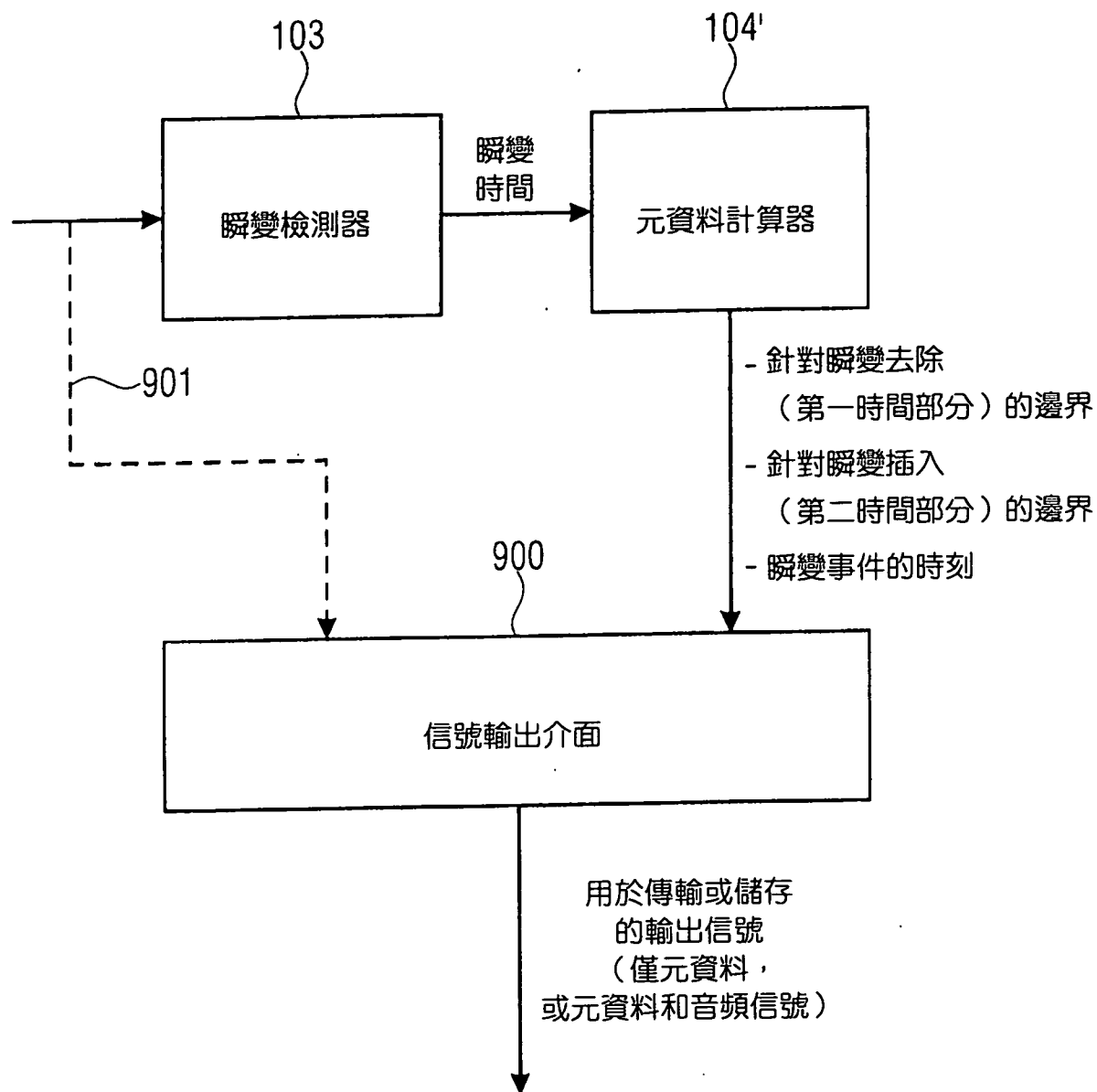
第八圖 C



第八圖 D



第八圖 E



第九圖