



(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 I497485 B

(45)公告日：中華民國 104 (2015) 年 08 月 21 日

(21)申請案號：101147783

(22)申請日：中華民國 94 (2005) 年 08 月 12 日

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

(30)優先權：2004/08/25 美國

60/604,836

(71)申請人：杜比實驗室特許公司(美國)DOLBY LABORATORIES LICENSING CORPORATION
(US)

美國

(72)發明人：文頓 馬克 S VINTON, MARK STUART (NZ)；薛費特 艾倫 J SEEFELDT, ALAN
JEFFREY (US)

(74)代理人：惲軼群；陳文郎

(56)參考文獻：

TW 200537436A

US 5812971

Todd, Craig C.; Davidson, G. A.; Davis, M. F.; Fielder, Louis D.;
Link, B. D.; Vernon, S. "AC-3: Flexible Perceptual Coding for Audio
Transmission and Storage," AES Convention 96 (February 1994),
Dolby Laboratories, San Francisco, CA, 1994/02/01 Page number
3796.

審查人員：涂淑惠

申請專利範圍項數：10 項 圖式數：2 共 26 頁

(54)名稱

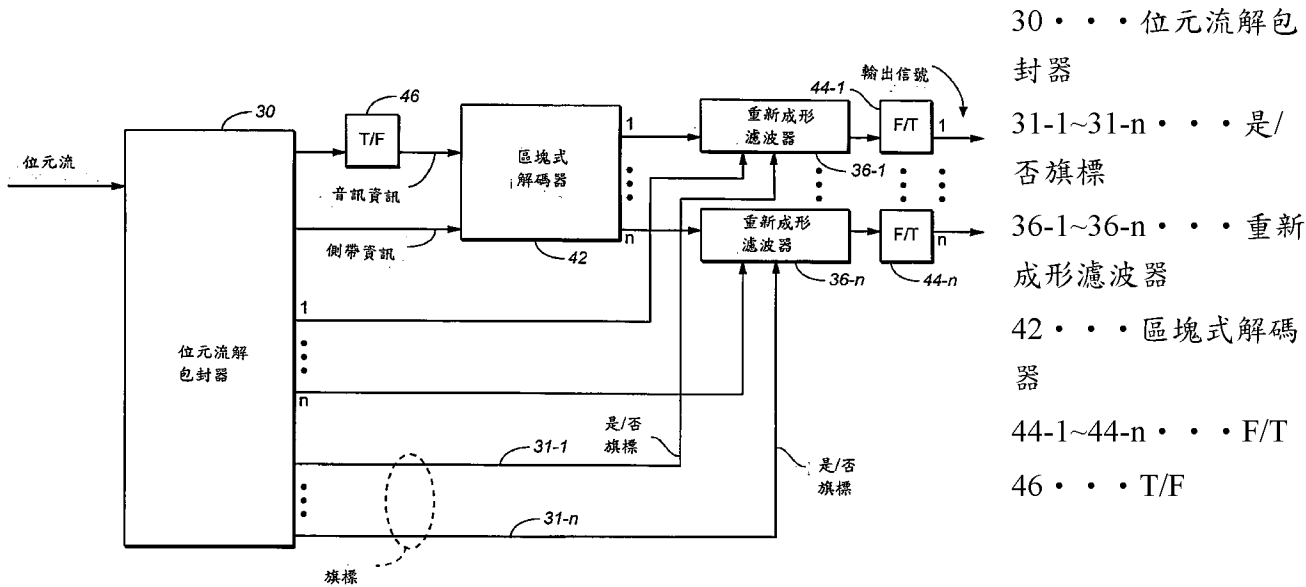
用以重塑經合成輸出音訊信號之時域包絡以更接近輸入音訊信號之時域包絡的方法

METHOD FOR RESHAPING THE TEMPORAL ENVELOPE OF SYNTHESIZED OUTPUT AUDIO
SIGNAL TO APPROXIMATE MORE CLOSELY THE TEMPORAL ENVELOPE OF INPUT AUDIO
SIGNAL

(57)摘要

某些型式之參數式空間編碼編碼器使用聲道間振幅差異、聲道間時間差異、與聲道間一致性或相關性以建立一多聲道聲場之一參數式模型，其被一解碼器使用以構建該原始聲場之近似。然而，此種參數式模型不會重建該聲場之聲道的原始時域包絡，其對一些音訊信號被發現是極端重要的。本發明提供在一空間編碼系統一個或多個被解碼之聲道的時域包絡之重新成形以較佳地媒配一個或多個時域包絡。

Certain types of parametric spatial coding encoders use interchannel amplitude differences, interchannel time differences, and interchannel coherence or correlation to build a parametric model of a multichannel soundfield that is used by a decoder to construct an approximation of the original soundfield. However, such a parametric model does not reconstruct the original temporal envelope of the soundfield's channels, which has been found to be extremely important for some audio signals. The present invention provides for the reshaping the temporal envelope of one or more of the decoded channels in a spatial coding system to better match one or more original temporal envelopes.



第 2 圖

發明專利說明書

公告本

雙面影印

(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：101147783 (由94122500分割)

※申請日期：94.8.12

※IPC 分類：G10L 19/00 (2013.01)

一、發明名稱：(中文/英文)

用以重塑經合成輸出音訊信號之時域包絡以更接近輸入音訊信號之時域包絡的方法 / Method for Reshaping the Temporal Envelope of Synthesized Output Audio Signal to Approximate More Closely the Temporal Envelope of Input Audio Signal

二、中文發明摘要：

某些型式之參數式空間編碼編碼器使用聲道間振幅差異、聲道間時間差異、與聲道間一致性或相關性以建立一多聲道聲場之一參數式模型，其被一解碼器使用以構建該原始聲場之近似。然而，此種參數式模型不會重建該聲場之聲道的原始時域包絡，其對一些音訊信號被發現是極端重要的。本發明提供在一空間編碼系統一個或多個被解碼之聲道的時域包絡之重新成形以較佳地媒配一個或多個時域包絡。

三、英文發明摘要：

Certain types of parametric spatial coding encoders use interchannel amplitude differences, interchannel time differences, and interchannel coherence or correlation to build a parametric model of a multichannel soundfield that is used by a decoder to construct an approximation of the original soundfield. However, such a parametric model does not reconstruct the original temporal envelope of the soundfield's channels, which has been found to be extremely important for some audio signals. The present invention provides for the reshaping the temporal envelope of one or more of the decoded channels in a spatial coding system to better match one or more original temporal envelopes.

四、指定代表圖：

(一)本案指定代表圖為：第 (2) 圖。

(二)本代表圖之元件符號簡單說明：

30…位元流解包封器

31-1~31-n…是／否旗標

36-1~36-n…重新成形濾波器

42…區塊式解碼器

44-1~44-n…F/T

46…T/F

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】

發明領域

[0001]本發明係有關於音訊解碼器，其中之音訊資訊在被編碼時具有被區塊率限制之時域包絡解析度，其包括感覺式與參數式音訊編碼器、解碼器與系統，亦有關於對應之方法、有關於用於施作這類方法之電腦、及有關於此類編碼器所產生之位元流。

【先前技術】

發明背景

[0002]很多縮減位元率之音訊編碼技術被稱為「區塊式」，因其包括的處理程序將正在被編碼之一或多個音訊信號的每一個分割為時間區塊，且不會比區塊率更頻繁地更新與該被編碼之音訊相關聯的至少一些側帶資訊。結果為其音訊資訊在被解碼時具有被區塊率限制之時域包絡解析度。後果為隨著時間被解碼之音訊信號的詳細構造不會就小於編碼技術之粒度(典型上為每一區塊8至50毫秒之範圍內)的期間而被保留。

[0003]此類區塊式音訊編碼技術不僅包括其中離散聲道一般透過編碼／解碼處理被保留之被習知為AC-3，AAC與各種MPEG形成之完備建立的感覺式編碼技術，亦包括最近被引進之有限位元率編碼技術，有時被稱為「立體聲提示編碼」與「參數式立體聲編碼」，其中多輸入聲道透過編碼／解碼處理被向下混頻為單聲場或由之被向上混頻。此

類編碼系統之細節被包含於各種文件中，包括下文由「參考文獻」之標題下被引述者。做為此類編碼系統中使用單聲道之後果，該被重建之輸出信號就特定區塊有必要彼此為振幅被刻度的版本，該等各種輸出信號有必要具有實質上相同細微之包絡構造。

[0004]雖然所有區塊式編碼技術可由其被解碼之音訊信號的改善時域包絡解析度受益，對此改善之需求在整個編碼／解碼處理中未保留離散聲道之區塊式編碼技術中特別大。如掌聲之某些型式的輸入信號對此類系統是特別有問題的，致使被再生之感覺空間影像變窄或潰散。

【發明內容】

發明概要

[0005]依照本發明之一第一層面，一種用於音訊信號編碼之方法被提供，其中一個或多個音訊信號被編碼成為一位元流，包含音訊資訊及與該音訊資訊相關且在將該位元流解碼為有用之側帶資訊，該編碼包括處理，其將每一該等一個或多個音訊信號分割為時間區塊並以不會比該區塊率頻繁地更新至少一些側帶資訊，使得該音訊資訊在被解碼時具有被該區塊率限制之一時域包絡解析度。比較在至少一音訊信號之時域包絡與每一至少此一音訊信號的被估計之解碼後的重建之時域包絡間被執行，此被估計之重建運用至少一些音訊資訊與至少一些側帶資訊，比較之結果的呈現用於改善至少一些音訊資訊在被解碼後之時域包絡解析度為有用的。

[0006]依照本發明之一另一層面，一種用於音訊信號編碼之方法被提供，其中一個或多個音訊信號被編碼成為一位元流，包含音訊資訊及與該音訊資訊相關且在將該位元流解碼為有用之側帶資訊，該位元流被接收及該音訊資訊使用該側帶資訊被解碼以提供一個或多個輸出音訊信號，該等編碼與解碼包括處理，其分別將一個或多個音訊信號與該被解碼之位元流的每一個分割成為時間區塊，該編碼以不會比該區塊率頻繁地更新至少一些側帶資訊，使得該音訊資訊在被解碼時具有被該區塊率限制之一時域包絡解析度。比較在至少一音訊信號之時域包絡與每一至少此一音訊信號的被估計之解碼後的重建之時域包絡間被執行，此被估計之重建運用至少一些音訊資訊與至少一些側帶資訊，比較之結果的呈現用於改善至少一些音訊資訊在被解碼後的時域包絡解析度為有用的。輸出至少一些該等呈現被執行，及將該位元流解碼被執行，該解碼運用該音訊資訊、該側帶資訊與該等被輸出之呈現。

[0007]依照本發明之一進一步層面，一種用於音訊信號解碼之方法被提供，其中一個或多個音訊信號已被編碼成為一位元流包含音訊資訊及與該音訊資訊相關且在將該位元流解碼為有用之側帶資訊，該編碼包括處理，其將每一該等一個或多個音訊信號分割為時間區塊並以不會比該區塊率頻繁地更新至少一些側帶資訊，使得該音訊資訊在被解碼時具有被該區塊率限制之一時域包絡解析度。該編碼進一步包含比較在至少一音訊信號之時域包絡與每一至少

此一音訊信號的被估計之解碼後的重建之時域包絡間被執行，此被估計之重建運用至少一些音訊資訊與至少一些側帶資訊，比較之結果的呈現用於改善至少一些音訊資訊在被解碼後之時域包絡解析度為有用的，及該編碼進一步包括輸出至少一些呈現。接收該位元流及將之解碼被執行，該解碼運用該音訊資訊、該側帶資訊與該等被輸出之呈現。

[0008]本發明之其他層面包括裝置裝置被適應於執行上述之方法；一電腦程式被儲存於電腦可讀取之媒體用於致使一電腦執行上述之方法；用上述方法被產生之一位元流；以及用被適應於執行上述方法之裝置被產生的一位元流。

圖式簡單說明

[0009]第1圖為實施本發明之層面的編碼器或編碼功能之示意功能方塊圖。

15 [0010]第2圖為實施本發明之層面的解碼器或解碼功能之示意功能方塊圖。

【實施方式】

較佳實施例之詳細說明

[0011]第1圖顯示本發明之層面可被運用的一解碼器或編碼環境之一例。如PCM信號之數個音訊輸入信號，各類比音訊信號之時間樣本1至n被施用至各時域對頻域變換器或變換功能(T/F)2-1至2-n。該等音訊信號例如可呈現如左、中、右等之空間方向。每一T/F例如可藉由將該等輸入音訊樣本分割成為區塊、將該等區塊作成窗、重疊該等區塊

、利用離散頻率變換(DFT)將每一該等窗與重疊區塊變換為頻域以及例如使用等值長方形波帶(ERB)刻度將結果所得之頻譜分割成為例如21個波帶的刺激人耳之關鍵波而被施作。此類DFT處理在本技藝中為相當習知的。其他之時域對頻域變換參數與技術可被運用。既非該等特定參數也非該等特定技術對本發明為緊要的。然而為解釋容易之目的，下列描述假設此一DTF變換技術被運用。

[0012]T/F 2-1至2-n之頻域輸出每一個為一組頻譜係數。這些組可分別用 $Y[k]_1$ 至 $Y[k]_n$ 被表示。所有這些組可被施用至區塊式編碼器或編碼器功能(區塊式編碼器)4。該區塊式編碼器例如可為上面獨立地被提及之任一習知的區塊式編碼器或有時為與包括上述這些編碼器之變形的任何未來區塊式編碼器成組合。雖然本發明之層面在配合於音訊資訊與解碼之際未保留離散聲道的區塊式編碼器使用為特別有益，本發明之層面在配合任何虛擬區塊式編碼器為有用的。

[0013]典型的區塊式編碼器4之輸出的特徵為「音訊資訊」與「側帶資訊」。音訊資訊可包含代表如在AC-3，AAC與其他者的區塊式編碼系統中可能的多個單聲道之資料，或其可僅包含利用將多輸入聲道向下混頻所導出之單聲道(如前述之二元提示編碼與參數式立體聲編碼系統，在二元提示編碼器或參數式立體聲編碼系統中被向下混頻之聲道亦可例如以AAC或一些其他適合之編碼感覺式地被編碼)。其亦可包含如2004年7月14日由Davis等人申請之美國專

利申請案第 S.N. 60/588,256 號的 "Low Bit Rate Audio Encoding And Decoding in Which Multiple Channels are Represented By Monophonic Channel and Auxiliary Information" 所揭示地利用將多輸入聲道向下混頻被導出之單聲道或多聲道。該側帶資訊可包含與該音訊資訊有關且在將之解碼為有用的資料。在各種向下混頻編碼系統之情形中，該側帶資訊可包含例如聲道間振幅差異、聲道間時間或相位差異與聲道間相關性之空間參數。

[0014] 然後來自區塊式編碼器 4 之音訊資訊與側帶資訊可被施用至各頻域對時域變換器或變換功能 (F/T) 6，其每一個一般執行上述 T/F 之反向功能，即逆 FFT，隨後為作窗與重疊添加。來自 F/T 6 之時域資訊被施用至位元流包封器或包封功能 (位元流包封器) 10，其提供被編碼之位元流輸出。或者，若該編碼器將提供代表頻域資訊之位元流，F/T 6 可被省略。

[0015] 來自區塊式編碼器 4 之頻域音訊資訊與側帶資訊亦被施用至解碼估計器或估計功能 (解碼估計器) 14。解碼估計器 14 可刺激至少一部分解碼器或解碼功能，被設計以將位元流包封器 10 所提供之已編碼之位元流加以解碼。此解碼器或解碼功能之例在下面相關的第 2 圖中被描述。解碼估計器 14 可提供多組之頻譜係數 $X[k]_1$ 至 $X[k]_n$ ，其近似被期待將在該解碼器或解碼功能被獲得之對應的輸入音訊信號之多組頻譜係數 $Y[k]_1$ 至 $Y[k]_n$ 。或者，其可就比所有輸入音訊信號較少地、就比該等輸入音訊信號之所有時間區塊較

少地、及／或就比所有頻帶較少地提供這些頻譜係數(即其不會提供所有之頻譜係數)。此例如可在若僅要改善代表應該是比其他者較重要之聲道的輸入信號時發生。作為另一例，此例如可在僅要改善其中人耳對時間性波形包絡之細微細節較敏感的較低頻率之部分的信號時會發生。

[0016]T/F 2-1至2-n之每一頻域輸出，該等各組頻譜係數 $Y[k]_1$ 至 $Y[k]_n$ ，每一個亦被施用至各比較裝置或功能(比較)12-1至12-n。這些組與各比較12-1至12-n中被估計之頻譜係數 $X[k]_1$ 至 $X[k]_n$ 的對應之時間區塊的對應之組被比較。在每一比較12-1至12-n中之比較結果每一個被施用至一濾波計算器或計算功能(濾波計算)15-1至15-n。此資訊就每一濾波計算為每一個功能區塊定義一濾波係數應為足夠的，此濾波器在被施用至一輸入信號之解碼後的重建時會形成具有改善解析度之一時域包絡的信號之結果。換言之，該濾波器會使該信號重新成形，使得其更近似地複製原始信號之時域包絡。該被改善之解析度為比該區塊率更細的解析度。較佳濾波器之進一步細節在下面被設立。

[0017]雖然第1圖之例子原則上顯示在頻域中之比較與濾波器計算，該比較與濾波器計算可在時域中被執行。不論是在頻域或時域中被執行，每一時間區塊只有一個濾波器組構被決定(雖然相同之濾波器組構可被施用至某些數個連續之重建區塊)。雖然一濾波器組構原則上可用逐一頻帶基準(如ERB尺度之每一頻帶)被決定，如此做會需要傳送大量之側帶資訊位元，其會破壞本發明之優點(即以位元率

之低提高來改善時域包絡解析度)。

[0018]在每一比較12-1至12-n中的比較量測為每一個被施用至一決定裝置或功能(決定)16-1至16-n。每一決策針對一臨界值比較該比較量測。一比較量測可採用各種形式且為非關鍵的。例如，每一個對應之係數值的差異之絕對值可被計算且該等差異被加總以提供一單一數值，其值表示該等信號波形在一時間區塊之際彼此差異的程度。此數值可與一臨界值被比較，使得若其超過該臨界，一「是」指標被提供至對應之濾波器計算。在沒有「是」指標時，該等濾波器計算可就該區塊被禁止，或者若被計算，其不會被該濾波器計算輸出。每一是／否資訊構成一旗標，其亦可被施用至位元流包封器10用於在該位元流中納入(因而，其可有數個旗標，每一輸入信號有一個，且每一此類旗標可用一位元被呈現)。

[0019]替選的是，每一決策16-1至16-n可由各濾波器計算14-1至14-n接收資訊以取代或添加至來自各比較12-1至12-n之資訊。各決策16可運用該等被計算之濾波器特徵(如其平均數或其尖峰振幅)作為進行決策或協助決策之基準。

如上述者，每一濾波器計算14-1至14-n提供比較結果之呈現，其可構成一濾波器之係數，此濾波器在被施用至一輸入信號之解碼後重建時會形成具有改善解析度之時域包絡的信號之結果。若該頻譜估計之頻譜係數 $X[k]_1$ 至 $X[k]_n$ 為不完整的(在解碼估計器比所有輸入音訊信號較少地、比該等輸入音訊信號之所有時間區塊較少地、及／或比所有

頻帶較少地提供頻譜係數之情形中)，其可能就所有時間區塊、頻帶與輸入信號不為每一比較12-1至12-n之輸出。讀者應注意到 $X[k]_1$ 至 $X[k]_n$ 係指被重建之輸出，而 $Y[k]_1$ 至 $Y[k]_n$ 係指輸入。

5 [0020]每一濾波器計算14-1至14-n之輸出可被施用至位元流組合器10。雖然濾波器資訊可與位元流分離地被傳送，較佳地是其作為部分之位元流與部分之側帶資訊被傳送。當本發明之層面被施用至現存之區塊式編碼系統時，被本發明之層面提供的額外資訊可被插入被欲承載輔助資
10 訊之此類系統的部分位元流中。

[0021]在實際實施例中，不僅是音訊資訊，而且是該輸入與該等濾波器係數也將以某種方法被數量化或被編碼。然而，為了呈現上簡單起見且由於此類細節為相當習知的及未有助於了解本發明，故無數量化及解除數量化在圖中
15 被顯示。

頻域中之維納濾波器設計

[0022]每一濾波器計算裝置或功能14-1至14-n的較佳特徵在於代表獲得一信號聲道之原始時域包絡的較精準再生所需之時域中乘法性變化的頻域中之一FIR濾波器。此濾
20 波器問題可被公式化為一最小平方問題，此經常被稱為維納濾波器設計。例如，見X Rong Li之Probability, Random Signals, and Statistics, 紐約CRC出版社，1999年出版，第423頁。施用維納濾波器技術具有之優點為減少運送重新成形之濾波器資訊至解碼器所需的額外位元。維納濾波器

之慣常的施用典型在時域上被設計及被施用。

[0023]該頻域最小平方濾波器設計問題可被定義為如下：在已知原始信號 $Y[k]$ 之DFT頻譜呈現與此原始聲道 $X[k]$ 之近似的DFT頻譜呈現下，計算會使公式1最小化之一組濾
5 波器係數(a_m)。注意， $Y[k]$ 與 $X[k]$ 為複數值，因而 a_m 一般亦將為複數。

$$\min_{a_m} \left[E \left| Y[k] - \sum_{m=0}^{M-1} a_m X[k-m] \right|^2 \right] \quad (1)$$

[0024]此處 k 為頻譜指標， E 為期望值運算元，及 M 為被設計之濾波器的長度。

$$\min_{a_m} \left[E \left| \bar{Y}_k - \bar{A}^T \bar{X}_k \right|^2 \right] \quad (2)$$

此處

$$\bar{Y}_k = [Y[k]]$$

$$\bar{X}_k^T = [X[k] \quad X[k-1] \quad \cdots \quad X[k-M+1]]$$

及

$$\bar{A}^T = [a_0 \quad a_1 \quad \cdots \quad a_{M-1}]$$

[0025]因而，利用公式2針對每一該等濾波器係數對0
設定部分微分，顯示對以公式3所得之該最小化問題為簡單的。

$$\bar{A} = \bar{R}_{xx}^{-1} \bar{R}_{xy} \quad (3)$$

20 此處

$$\bar{R}_{XX} = \begin{bmatrix} E(X_K X_K^*) & E(X_K X_{K-1}^*) & \cdots & E(X_K X_{K-M+1}^*) \\ E(X_{K-1} X_K^*) & E(X_{K-1} X_{K-1}^*) & \cdots & E(X_{K-1} X_{K-M+2}^*) \\ \cdots & \cdots & \ddots & \cdots \\ E(X_{K-M+1} X_K^*) & E(X_{K-M+1} X_{K-1}^*) & \cdots & E(X_{K-M+1} X_{K-M+1}^*) \end{bmatrix}$$

及

$$\bar{R}_{YX}^T = \begin{bmatrix} E(Y_K X_K^*) & E(Y_K X_{K-1}^*) & \cdots & E(Y_K X_{K-M+1}^*) \end{bmatrix}$$

[0026]公式3定義最適濾波器係數之計算，其使一特定

- 5 聲道之原始頻譜 $Y[k]$ 與重建頻譜 $X[k]$ 間的誤差最小化。一般而言，一組濾波器係數就每一輸入信號之每一時間區塊被計算。

- [0027]在本發明之層面的實際實施例中，第12階之維納濾波器被運用，雖然本發明不限於使用此大小之維納濾波
10 器。此實際實施例運用符合DFT之頻域中的處理。後果為該等之維納濾波器係數為複數且每一濾波器需要傳輸24個實數。為有效率地運送此濾波器資訊至解碼器，向量數量化(VQ)可被用以將每一濾波器之係數編碼。一編碼簿可被運用使得僅有一指標須被傳送至解碼器以運送該第12階複
15 數濾波器資訊。在一實際實施例中，具有24維度及16,536個登入值之一VQ表編碼簿被發現為有用的。本發明不限於使用向量數量化或使用編碼簿。

[0028]雖然上面之描述假設使用一DFT以估計其頻譜內容及設計該維納濾波器，一般而言任何變換可被使用。

- 20 [0029]第2圖顯示本發明之層面被運用之一解碼器或解碼處理環境的一例。此種解碼器或解碼處理可適於如相關

第1圖之例子所描述地配合一編碼器或編碼處理的作業。如第1圖之配置所產生之一編碼後位元流被信號傳輸或儲存
的任何適合之模式被接收且被施用至一位元流解包封器30
，其如必要地將該位元流解開包封以分離來自側帶資訊之
5 被編碼的音訊資訊與該等是／否旗標(若被包括於該位元
流中)。該側帶資訊較佳地包括一組濾波器係數用於改善被
施用至第1圖之編碼配置的一個或多個輸入信號之每一的
重建。

[0030]在此例中，其被假設有一再生信號對應於每一輸
10 入每一及時域包絡重新成形濾波器資訊為每一再生信號被
提供，雖然此如上述地不須為此情形。因而，1至n組之濾
波器係數側帶資訊為每一輸入信號被施用至各重新成形濾
波器36-1至36-n，其作業在下面被解釋。每一濾波器亦可接
收各是／否旗標31-1至31-n，指出該濾波器是否在一特定時
15 間區塊應為有源的。

[0031]來自位元流解包封器30之側帶資訊亦可包括其
他資訊，例如為在二元提示編碼或參數立體聲系統情形中
之聲道間振幅差異、聲道間時間或相位差異與聲道間交叉
相關。區塊式解碼器42接收來自解包封器30之側帶資訊
20 與解包封器30之時域對頻域變換後的音訊資訊。來自解包
封器30之音訊資訊經由一時域對頻域輸出器或輸出功能
(T/F)被施用，其可能與第1圖之任一頻域輸出器或輸出功能
2-1至2-n相同。

[0032]區塊式解碼器42提供一個或多個輸出，其每一個

為第1圖中對應的輸入信號之一近似。雖然某些輸入信號可不具有對應的信號，第2圖之例顯示輸出1至n，其每一個為第1圖中對應的各輸入信號1-n之一近似。在此例中，解碼器之每一該等輸出信號1至n被施用至各重新成形濾波器36-1至36-n，其每一個可被施作成為一FIR濾波器。每一FIR濾波器之係數以區塊基礎被其重建輸出將被改善之相關的特定輸入聲道之各濾波器資訊加以控制。在時域中之乘法性包絡重新成形藉由用每一濾波器36-1至36-n中的一區塊式解碼器輸出迴旋每一FIR濾波器而被較佳地達成。因而，依照本發明之層面的時域包絡重新成形採用該時間頻率雙重性的益處一時域中之迴旋等值於頻域中之乘法性，反之亦然。然後每一被解碼及被濾波之輸出信號被施用至各頻域對時域變換器或變換功能(F/T)44-1至44-n，其每一個在作窗及重疊添加後執行上述T/F之反函數，即反F/T。替選的是，適合之時域重新成形濾波器可在每一頻域對時域變換器後被運用。例如，第n階多項式曲線之n個多項式係數可取代FIR濾波器係數與在時域中被乘法所施用之曲線被傳送作為側帶資訊。雖然運用維納濾波器技術以運送重新成形濾波器資訊至解碼器為較佳，其他之頻域與時域技術可如Truman與Vinton在2002年3月28日申請之美國專利申請案第S.N.10/113,858號且在2003年10月2日公布之US 2003/0187663 A1號的”Broadband Frequency Translation for High Frequency Regeneration”中所設立者被運用。該申請案之整體在此處被納入作為參考。

施作

[0033]本發明可用硬體或軟體或二者之組合(如可程式邏輯陣列)被施作。除非有特別明說，被包括作為部分之本發明的法則並非固守於任何特定之電腦或其他裝置。特別是，各種通用機器可以具有依照此處之教習被寫成的程式被使用，或其可更方便地構建更專業之裝置(如積體電路)以執行所要求之方法步驟。因而，本發明可被施作成為在每一個包含至少一處理器、至少一資料儲存系統(包括依電性與非依電性記憶體及／或儲存元件)、至少一輸入裝置或埠與至少一輸出裝置或埠之一個或多個可程式電腦系統上執行的一個或多個電腦程式。程式碼被施用至輸入資料以執行此處所描述之功能並產生輸出資訊。該輸出資訊以習知方式被施用至一個或多個輸出裝置。

[0034]每一此程式可用任一所欲之電腦語言(包括機器語言、組合語言、或高階程序、邏輯或物件導向程式語言)以與電腦系統通訊。在任一情形中，該語言可為已編譯或已解譯之語言。

[0035]每一此電腦程式較佳地被儲存或被下載至為通用或特殊目的可程式電腦可讀取之儲存媒體或裝置(如固態記憶體或媒體，或磁性或光學媒體)，用於在被該電腦系統讀取時組構及操作該電腦以執行此處所描述之程序。本發明性之系統亦可被考慮將被施作成為電腦可讀取之媒體、以電腦程式被組構，此處如此被組構之儲存媒體致使一電腦系統以特定與預定方式操作以執行此處被描述之功能

。本發明之數個實施例已被描述。不過，其將被了各種修改可不偏離本發明之精神與領域地被做成。例如，此處所描述之一些步驟可為順序上獨立的，且因而可與此所描述不同之順序被執行。

5 參考文獻

[0036]下列之專利、專利申請案與公報的每一個整體在此被納入作為參考。

AC-3

2001年8月20日之先進電視系統委員會的ATSC標準
10 A52/A：Digital Audio Compression Standard(AC-3)，修正版A。該A52/A文件可在全球資訊網之
<http://www.atsc.org/standards.html>取得。

1995年8月IEEE Trans Consumer Electronics第41卷第3期，Steve Vernon之”Design and Implementation of AC-3
15 Coders”。

1993年10月第95屆AES年會之Audio Engineering Society再版3774，Mark Davis的”The AC-3 Multichannel Coder”。

1992年10月第93屆AES年會Audio Engineering Society
20 再版3365，Bosi等人的”High Quality Low-Rate Audio Transform Coding for Transmission and Multimedia Applications”。

美國專利第5,583,962；5,632,005；5,633,981；5,727,119與6,021,386號。

AAC

1996年ISO/IEC IS-14496(第三部,音訊)之ISO/IEC JTC1/SC29, "Information technology-very low bitrate audio-visual coding"。

- 5 1997年(1) ISO/ICE 13818-7"MPEG-2 advanced audio coding, ACC"國際標準。

1996年第101屆AES年會論文集之M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson與Y. Oikawa的
10 "ISO/IEC MPEG-2 Advanced Audio Coding"。

1997年10月AES期刊第45卷第10期第789-814頁之M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson與Y. Oikawa的"ISO/IEC MPEG-2 Advanced Audio Coding"。

- 15 1999年意大利佛羅倫斯AES第17屆高品質音訊編碼國際研討會論文集之Karlheinz Brandenburg的"MP3與ACC explained"。

1998年Audio Eng. Soc.期刊第46卷第3期第164-177頁之G. A. Soulodre等人的"Subjective Evaluation of
20 State-of-the-Art Two-Channel, Audio Coders"。

MPEG強度立體聲

美國專利第5,323,396; 5,539,829; 5,606,618與5,621,885號。

美國專利申請案公報US 2001/0044713號(已出版)。

空間與參數式編碼

2004年7月14日 Davis等人申請之美國臨時專利申請案第 S. N. 60/588,256號的”Low Bit Rate Audio encoding and Decoding in Which Multiple Channels and Represented By Monophonic Channel and Auxiliary Information”。

2003年2月6日出版之美國專利申請案公報 US 2003/0026441號。

2003年2月20日出版之美國專利申請案公報 US 2003/0035553號。

2003年11月27日出版之美國專利申請案公報 US 2003/0219130號。(Baumgarte與Faller)。

2003年3月音訊工程協會論文第5852號。

2003年10月30日出版之國際專利申請案 WO 03/090206號。

2003年10月30日出版之國際專利申請案 WO 03/090207號。

2003年10月30日出版之國際專利申請案 WO 03/090208號。

2003年1月22日出版之國際專利申請案 WO 03/007656號。

2003年12月25日出版之 Baumgarte等人的美國專利申請案公報 US 2003/0236583 A1 號之 ”Hybrid Multi-Channel/Cue Coding/Decoding of Audio Signals”(申請案第 S. N. 10/246,570號)。

2002年5月慕尼黑第112屆音訊工程協會年會論文第5574號之Faller等人的”Binaural Cue Coding Applied to Stereo and Multi-Channel Audio Compression”。

2002年5月慕尼黑第112屆音訊工程協會年會論文第5575號之”Why Binaural Cue Coding is Better than Intensity Stereo Coding”。

2002年5月洛杉磯第113屆音訊工程協會年會論文第5706號之”Design and Evaluation of Binaural Cue Coding Schemes”。

2001年10月紐約New Paltz, IEEE Workshop on Applications of Processing to Audio and Acoustics之Faller等人的”Efficient Representation of Spatial Audio Using Perceptual Parameterization”。

2002年5月佛羅里達之奧倫多ICASSP 2002第II-1801-II-1804頁之Baumgarte等人的”Estimation of Auditory Spatial Cues for Binaural Cue Coding”。

2002年5月佛羅里達之奧倫多ICASSP 2002第II-1841-II-1844頁之Faller等人的”Binaural Cue Coding: A Navel and Efficient Representation of Spatial Audio”。

2004年5月柏林第116屆音訊工程協會年會論文第6072號之Breebaart等人的”High-quality parametric spatial audio coding at low bitrates”。

2004年5月柏林第116屆音訊工程協會年會論文第6060號之Baumgarte等人的”Audio Coder Enhancement using

Scalable Binaural Cue Coding with Equalized Mixed”。

2004年5月柏林第116屆音訊工程協會年會論文第6073號之Schuijers等人的”Low Complexity Parametric Stereo Coding”。

5 2004年5月柏林第116屆音訊工程協會年會論文第6074號之Engdegard等人的”Synthetic Ambience in Parametric Stereo Coding”。

其他

10 1998年9月22日Herre之美國專利第5,812,971號的”Enhanced Joint Stereo Coding Method Using Temporal Envelope Shaping”。

1994年阿姆斯特丹音訊工程協會第96屆年會論文第3799號(再版)之Herre等人的”Intensity Stereo Coding”。

【主要元件符號說明】

1至n…音訊信號樣本	16-1~16-n…決策裝置
2-1至2-n…時域對頻域變換器	30…位元流解包封器
(T/F)	31-1~31-n…是／否旗標
4…區塊式編碼器	36-1~36-n…重新成形濾波器
6…頻域對時域變換器(F/T)	42…區塊式解碼器
10…位元流包封器	44-1~44-n…F/T
12-1~12-n…比較裝置	46…T/F
14…解碼估計器	
15-1~15-n…濾波計算器	

七、申請專利範圍：

1. 一種用以重塑至少一輸出音訊信號之時域包絡以更接近輸入音訊信號之時域包絡的方法，其包含下列步驟：

a)接收包含一向下混頻音訊信號及包絡側帶資訊之一空間編碼位元串流，其中：

該向下混頻音訊信號包含經分割為時間槽的一音訊信號，且該包絡側帶資訊係與該音訊信號相關，且可用在重塑該至少一輸出音訊信號之該時域包絡，

該包絡側帶資訊包含藉由比較該輸入音訊信號之一包絡及從該向下混頻音訊信號所導出的一音訊信號之一包絡所取得的資訊，該向下混頻音訊信號係至少從該輸入音訊信號進行向下混頻，

b)解碼該空間編碼位元串流，以生成該至少一輸出音訊信號，以及

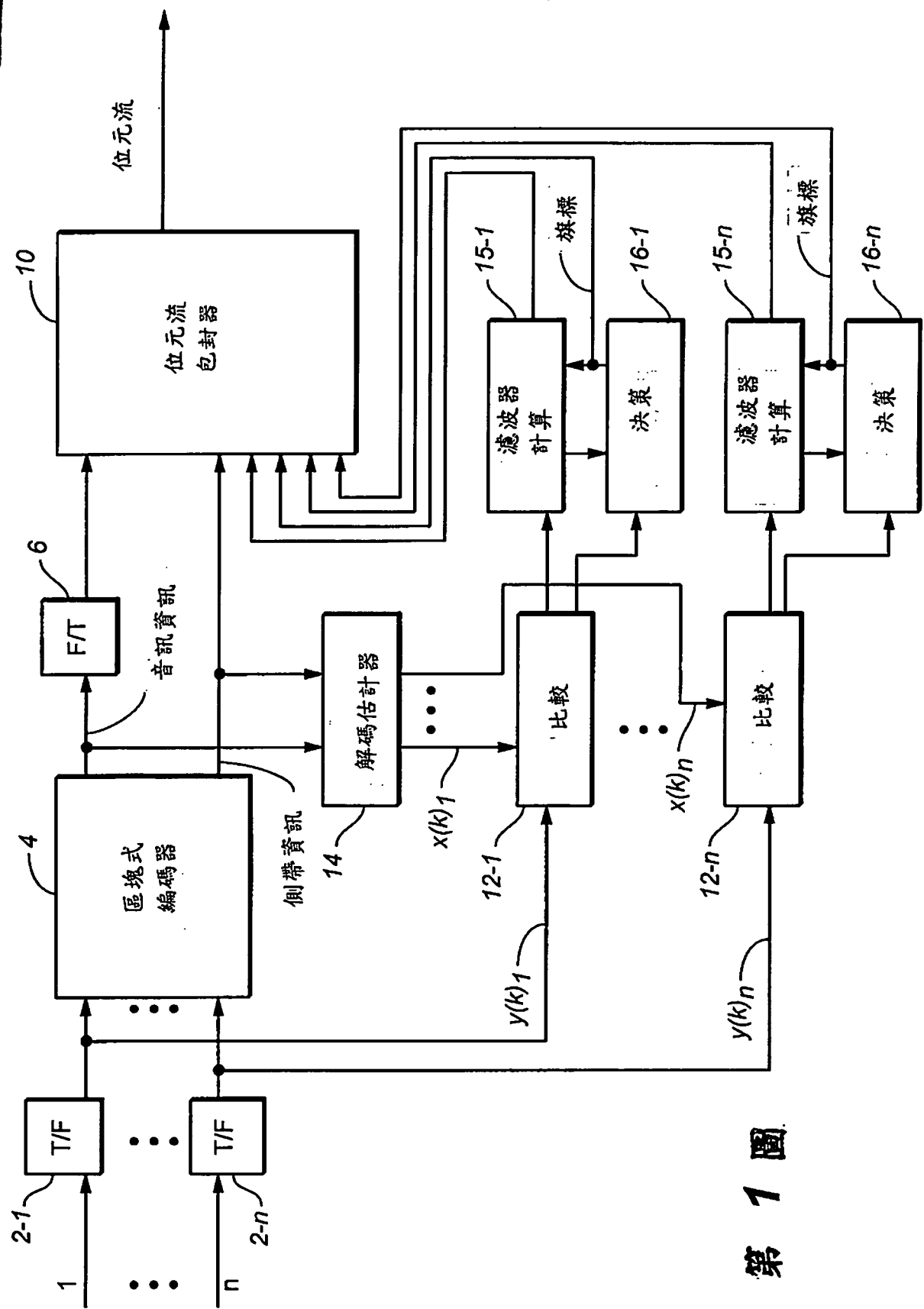
c)根據該包絡側帶資訊重塑該至少一輸出音訊信號之該時域包絡，以更接近該輸入音訊信號之該時域包絡。

2. 如申請專利範圍第1項所述之方法，其中該等時間槽之粒度與關於該向下混頻音訊信號之一包絡的一輸入音訊信號之一包絡有關。
3. 如申請專利範圍第1項所述之方法，其中解碼該位元串流包括控制多個重塑濾波器與該包絡側帶資訊之解相關程序。

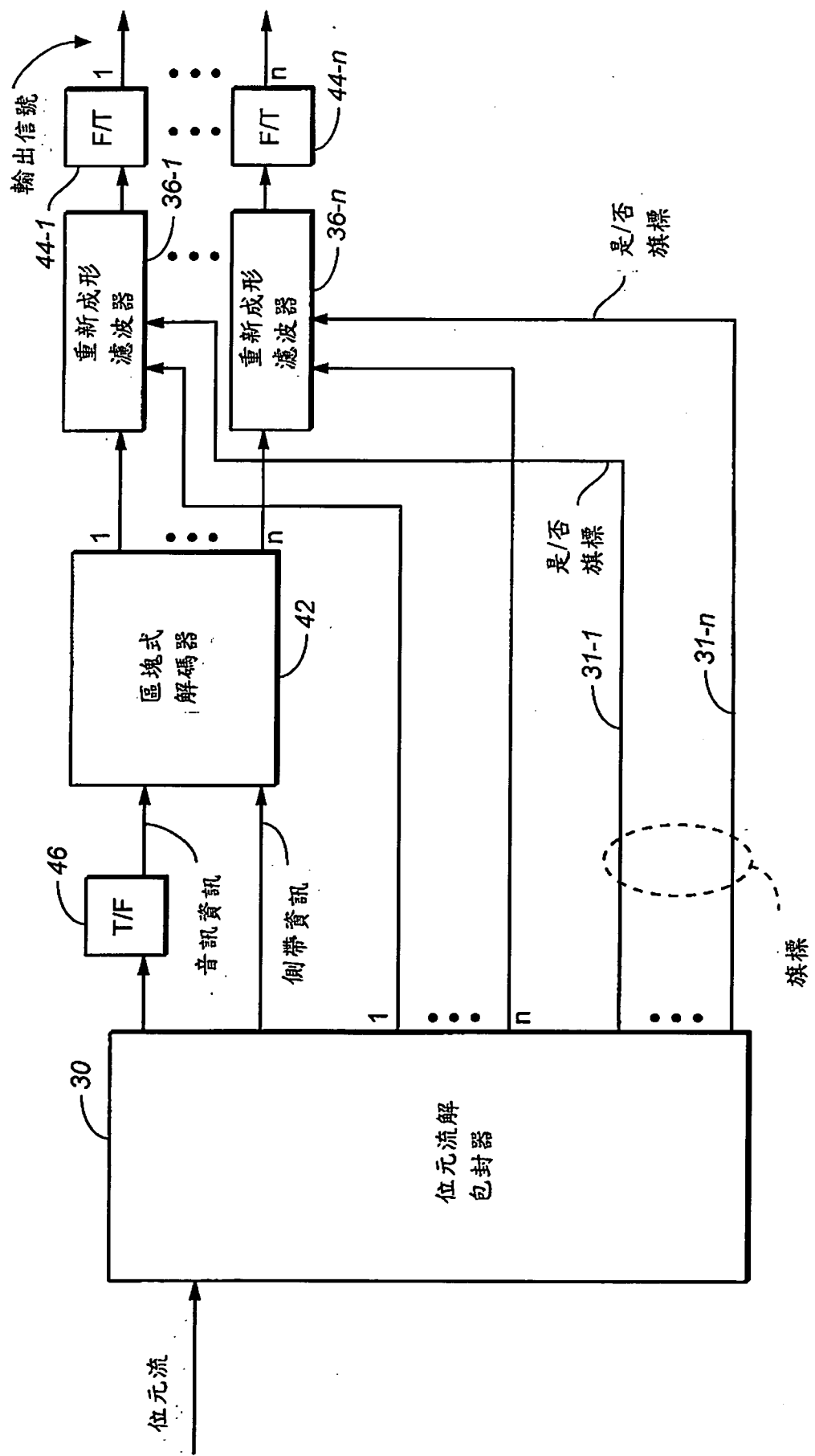
4. 如申請專利範圍第1項所述之方法，其中一輸入音訊信號之一包絡與從該向下混頻音訊信號所導出之一信號的一包絡之比較包含於經比較的該等包絡之該側帶資訊中的一表示。
5. 如申請專利範圍第1項所述之方法，其中該空間編碼位元串流包含從一或多個輸入音訊信號所構成之一位元串流。
6. 如申請專利範圍第5項所述之方法，其中該一或多個輸入音訊信號包含一多聲道聲場。
7. 如申請專利範圍第1項所述之方法，其中該一音訊信號分割成的該等時間槽各包含一時間區塊。
8. 如申請專利範圍第7項所述之方法，其中該時間槽之粒度包含由上述時間區塊之一比率所限制的一解析度。
9. 如申請專利範圍第7項所述之方法，其中上述時間區塊具有限制該包絡側帶資訊之一解析度的一區塊率。
10. 如申請專利範圍第1項所述之方法，其中從該向下混頻音訊信號所導出的該信號包含該輸入音訊信號的一估計解碼重建。

101147783

雙面影印



第 1 圖



第 2 圖