



(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 I498883 B

(45)公告日：中華民國 104 (2015) 年 09 月 01 日

(21)申請案號：101150176

(22)申請日：中華民國 94 (2005) 年 03 月 01 日

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

(30)優先權：2004/03/01	美國	60/549,368
2004/06/14	美國	60/579,974
2004/07/14	美國	60/588,256

(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：戴維斯 馬克 F DAVIS, MARK FRANKLIN (US)

(74)代理人：惲軼群；陳文郎

(56)參考文獻：

TW	374152	TW	489295
TW	552580	TW	556153
US	554334	US	2002/0147582A1
US	2004/0204930A1		

審查人員：涂淑惠

申請專利範圍項數：4 項 圖式數：9 共 61 頁

(54)名稱

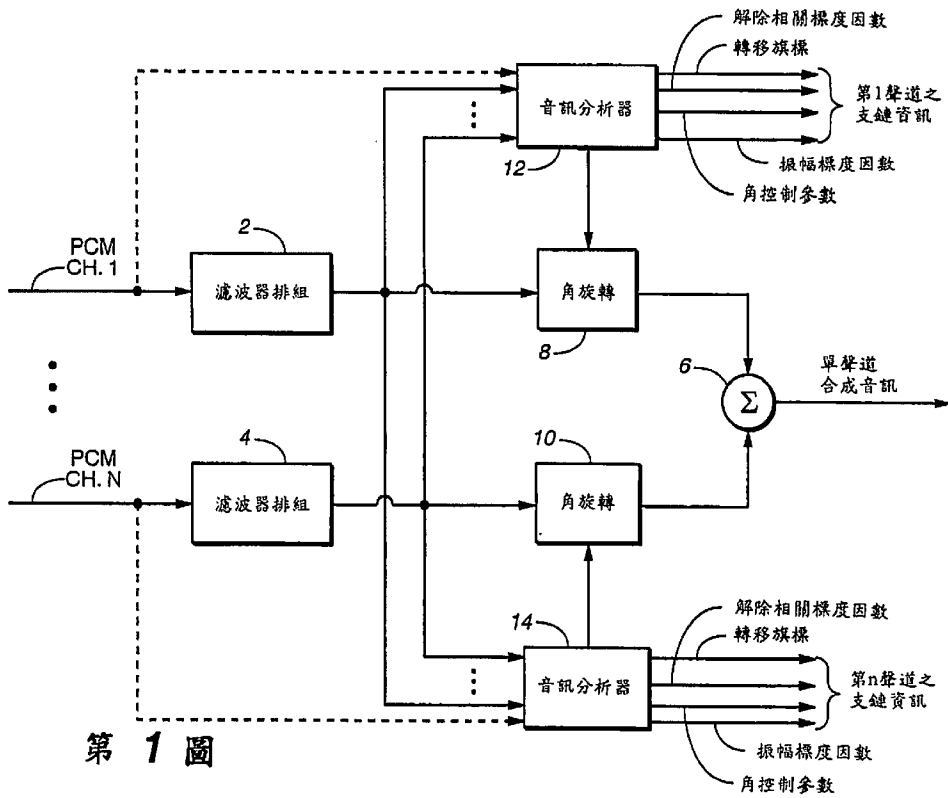
用以解碼代表N個音訊聲道之M個經編碼音訊聲道的方法

METHOD FOR DECODING M ENCODED AUDIO CHANNELS REPRESENTING N AUDIO CHANNELS

(57)摘要

多聲道之音訊被組合為單聲道合成信號或具有相關輔助資訊之多聲道的音訊，多聲道音訊由該輔助資訊被重建，包括多聲道音訊對單聲道音訊信號之改良的向下混頻或由單聲道音訊聲道或由多聲道音訊被導出多聲道音訊之改良的解除相關。所揭示之本發明的層面在音訊編碼器、解碼器、編碼/解碼系統、向下混頻器、向上混頻器與解除相關器中為有用的。

Multiple channels of audio are combined either to a monophonic composite signal or to multiple channels of audio along with related auxiliary information from which multiple channels of audio are reconstructed, including improved downmixing of multiple audio channels to a monophonic audio signal or to multiple audio channels and improved decorrelation of multiple audio channels derived from a monophonic audio channel or from multiple audio channels. Aspects of the disclosed invention are usable in audio encoders, decoders, encode/decode systems, downmixers, upmixers, and decorrelators.



第 1 圖

- 2 . . . 濾波器排組
- 4 . . . 濾波器排組
- 6 . . . 加法組合器
- 6' . . . 向下混頻矩陣
- 8 . . . 旋轉角
- 10 . . . 旋轉角
- 12 . . . 音訊分析器
- 14 . . . 音訊分析器

公告本

双面影印

發明摘要

※ 申請案號：101150176 (由94106045分割)

※ 申請日：中華民國94年3月1日

※ IPC 分類：G10L 19/02

(2006.01)

【發明名稱】(中文/英文)

用以解碼代表 N 個音訊聲道之 M 個經編碼音訊聲道的方法 / Method for Decoding M Encoded Audio Channels Representing N Audio Channels

【中文】

多聲道之音訊被組合為單聲道合成信號或具有相關輔助資訊之多聲道的音訊，多聲道音訊由該輔助資訊被重建，包括多聲道音訊對單聲道音訊信號之改良的向下混頻或由單聲道音訊聲道或由多聲道音訊被導出多聲道音訊之改良的解除相關。所揭示之本發明的層面在音訊編碼器、解碼器、編碼/解碼系統、向下混頻器、向上混頻器與解除相關器中為有用的。

【英文】

Multiple channels of audio are combined either to a monophonic composite signal or to multiple channels of audio along with related auxiliary information from which multiple channels of audio are reconstructed, including improved downmixing of multiple audio channels to a monophonic audio signal or to multiple audio channels and improved decorrelation of multiple audio channels derived from a monophonic audio channel or from multiple audio channels. Aspects of the disclosed invention are usable in audio encoders, decoders, encode/decode systems, downmixers, upmixers, and decorrelators.

【代表圖】

【本案指定代表圖】：第（ 1 ）圖。

【本代表圖之符號簡單說明】：

2…濾波器排組

4…濾波器排組

6…加法組合器

6' …向下混頻矩陣

8…旋轉角

10…旋轉角

12…音訊分析器

14…音訊分析器

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

用以解碼代表 N 個音訊聲道之 M 個經編碼音訊聲道的方法 / Method for Decoding M Encoded Audio Channels Representing N Audio Channels

【技術領域】

發明領域

【0001】 本發明係大致有關於音訊信號處理。更明確地說，本發明之層面係有關於一編碼器(或編碼處理)、一解碼器(或解碼處理)與一編碼/解碼系統(或編碼/解碼處理)，用於具有非常低之位元率的音訊信號，其中數個音訊聲道用一合成的單聲道(單音)音訊聲道與輔助資訊(支鏈)被呈現。替選的是，數個音訊聲道用數個音訊聲道與支鏈資訊被呈現。本發明之層面亦有關於多聲道對合成單聲道之向下混頻器(或向下混頻處理)、單聲道對多聲道(向上混頻處理)及多聲道對多聲道解除相關器(或解除相關處理)。本發明之其他層面係有關於多聲道對多聲道向上混頻器(或向上混頻處理)與一解除相關器(解除相關處理)。

【先前技術】

發明背景

【0002】 在 AC-3 數位音訊編碼與解碼系統中，在該系統變得對位元渴望時可選擇性地被組合或被「耦合」於高頻率。AC-3 之細節在本技藝中為相當習知的一例如見先進電視系統委員會 2001 年 8 月 20 日之 ATSC 標準 A52/A：數位音訊壓縮標準(AC-3)，修訂版 A。A/52A 文件可在全球資訊網 <http://www.atsc.org/standards.html> 取得。該 A/52A 文件之整體在此被納為參考。

【0003】 高於 AC-3 系統組合之隨選聲道的頻率被稱為「耦合」頻率。高於耦合頻率，被耦合之聲道被組合為「耦合」或合成聲道。該編碼器在每一聲道為高於該耦合頻率之每一子帶產生「耦合座標」(振幅標度因數)。該耦合座標表示每一被耦合之子帶的原始能量對在合成聲道中對應的子帶

之能量的比值。被耦合之子帶的相位極性在該聲道與一個或更多其他被耦合之聲道被組合以減少相位外信號成份抵銷前被逆轉。該合成及就每一子帶基準包括該等耦合座標與該聲道之相位是否被逆轉之支鏈資訊被送至解碼器。在實務上，AC-3 系統之商業實施例所運用之耦合頻率具有約 10kHz 至約 3500Hz 之範圍。美國專利第 5,583,962, 5,633,981, 5,727,119, 5,909,664 與 6,021,386 號包括關於組合多音訊聲道為一合成聲道與輔助或支鏈資訊及由其恢復為近似原始多聲道的教習。每一該等專利之整體在此被納為參考。

【發明內容】

發明概要

【0004】 本發明之層面可被視為對 AC-3 編碼與解碼之「耦合」技術，及也對其中音訊多聲道被組合為單聲道合成信號或具有相關輔助資訊之單聲道音訊與多聲道音訊由此被重建的其他技術之改良。本發明之層面亦可被視為對多聲道音訊向下混頻單聲道音訊信號或多聲道音訊及對由單聲道音訊聲道或由多聲道音訊被導出之多聲道音訊解除相關技術的改良。

【0005】 本發明之層面可在 $M:1:N$ 空間音訊編碼技術(其中 N 為音訊聲道數)或 $M:1:N$ 空間音訊編碼技術(其中 M 為被編碼之音訊聲道數及 N 為被解碼之音訊聲道數)中被運用，其藉由在提供改善的相位補償、解除相關機制與信號相依之可變的時間常數之各事項來對聲道耦合改良。本發明之層面亦可 $N:x:N$ 與 $M:x:N$ 空間音訊編碼技術中被運用，其中 x 可為 1 或大於 1。目標包括藉由在向下混頻前調整聲道間相位移位來減少耦合抵銷之人工物，及藉由恢復相位角與解碼器之解除相關程度來改善被再生之信號的空間維度。本發明之層面在實務實施例中被實施時應允許連續而非隨選之聲道耦合及比起如在 AC-3 系統中較低的耦合頻率而降低所需要之資料率。

【圖式簡單說明】

圖式簡單說明

【0006】 第 1 圖為一理想化方塊圖，顯示實施本發明之層面的 $N:1$ 編碼配置原理功能或裝置。

【0007】 第 2 圖為一理想化方塊圖，顯示實施本發明之層面的 $1:N$

解碼配置原理功能或裝置。

【0008】 第 3 圖顯示沿著一(垂直)頻率軸之 bin 與子帶及沿著一(水平)時間軸之區塊與訊框的簡化概念的組織例子。此圖並非依比例畫出。

【0009】 第 4 圖為一混合流程圖與功能性方塊圖之性質，顯示實施本發明之層面的編碼配置之功能的編碼步驟或裝置。

【0010】 第 5 圖為一混合流程圖與功能性方塊圖之性質，顯示實施本發明之層面的解碼配置之功能的解碼步驟或裝置。

【0011】 第 6 圖為一理想化方塊圖，顯示實施本發明之層面的一第一 $N : x$ 編碼配置之原理功能或裝置。

【0012】 第 7 圖為一理想化方塊圖，顯示實施本發明之層面的 $x : M$ 解碼配置之原理功能或裝置。

【0013】 第 8 圖為一理想化方塊圖，顯示實施本發明之層面的一第一替選 $x : M$ 解碼配置之原理功能或裝置。

【0014】 第 9 圖為一理想化方塊圖，顯示實施本發明之層面的一第二替選 $x : M$ 解碼配置之原理功能或裝置。

【實施方式】

【0015】 較佳實施例之詳細說明

【0016】 基本的 $N : 1$ 編碼器

【0017】 參照第 1 圖，實施本發明之層面之 $N : 1$ 編碼器功能或裝置被顯示。該圖為實施本發明之層面的基本編碼器之功能或結構例子。實施本發明之層面之其他功能或結構配置可被運用，包括下面被描述之替選的及/或等值功能或結構。

【0018】 二個或更多的音訊輸入聲道被施用至該編碼器。雖然在原理上本發明之層面可用類比、數位或混合式類比/數位實施例被實作，此處所揭示之例子為數位實施例。因而，該等輸入信號可為時間樣本，其可為已由類比音訊信號被導出。該等時間樣本可被編碼為線性脈波碼調變(PCM)信號。每一線性 PCM 音訊輸入聲道用具有如 512 點視窗化遞送離散傅立葉變換(DFT)(如用快速傅立葉(FFT)施作)之同相位與正交輸出的一濾波器排組功能或裝置被處理。該濾波器排組可被視為時間域對頻率域變換。

【0019】 第 1 圖分別顯示被施用至一濾波器排組功能與裝置(濾波器排組 2)之一第一 PCM 聲道輸入(聲道 1)與被施用至另一濾波器排組功能與裝置(濾波器排組 4)之一第二 PCM 聲道輸入(聲道 n)。其有 n 個輸入聲道，其中 n 為等於 2 或以上之整個正整數。因而其亦有 n 個濾波器排組，每一個接收 n 個輸入聲道的獨一個。為了呈現簡單，第 1 圖僅顯示二輸入聲道 1 與 n。

【0020】 當一濾波器排組用一 FFT 被施作時，輸入時間域信號被分段為連續的區塊，且經常在重疊的區塊中被處理。該等 FET 之離散頻率輸出(變換係數)被稱為 bin，每一個具有一複數分別以其實數部與虛數部對應於同相位與正交成份。連續變換 bin 可被分組為近似於人耳之關鍵帶寬的子帶，且編碼器所產生之大多數支鏈資訊如將被描述地以每一子帶之基準被計算與被傳輸以使處理資源最小化及降低位元率。多連續時間域區塊可被組成訊框，以各區塊值對每一區塊被平均或被組合或累積以使支鏈資料率最小化。在此處被描述之例子中，每一濾波器排組被 FFT 施作、連續的變換 bin 被組成子帶、區塊被組成訊框、及支鏈資料以每訊框一次之基準被傳送。替選的是支鏈資料以多於每訊框一次基準被傳送(例如每區塊一次)。例如見第 3 圖與其此後之描述。明顯的是，在支鏈資訊被傳送之頻率與所要求的位元率間有取捨。

【0021】 本發明之層面的適當施作在 48 kHz 抽樣率被運用時可運用約 32 毫秒之固定長度的訊框，每一訊框具有約每個 5.3 毫秒間隔之 6 個區塊(例如運用具有約 10.6 毫秒長度及 50%重疊之區塊)。然而，既非運用固定長度訊框亦非其被分割為固定數目之區塊的這類時機在假設此處所描述之資訊以每訊框基準被傳送係以約 20 至 40 毫秒被傳送時對實施本發明之層面為關鍵的。訊框可為任意大小且其大小可動態地變化。可變的區塊長度可在如上述的 AC-3 系統中被運用。其被了解此處係對「訊框」與「區塊」被提到。

【0022】 實務上，若合成單聲道或多聲道信號，或合成單聲道或多聲道信號與離散低頻率聲道例如用下面描述之感覺編碼器被編碼，運用與在感覺編碼器被運用相同的訊框與區塊組配為方便的。此外，若該編碼器運

用可變的區塊長度使得隨時間不同由一區塊長度切換為另一種時，若此處所描述之一個或更多支鏈資訊在此區塊切換發生時被更新，其會為所欲的。為了在區塊切換發生時使更新支鏈資訊的資料費用增加最小化，被更新之支鏈資訊的頻率解析度可被降低。

【0023】 第 3 圖顯示沿著一(垂直)頻率軸之 bin 與子帶及沿著一(水平)時間軸之區塊與訊框的簡化概念的組織例子。當 bin 被分為近似關鍵頻帶之子帶時，最低頻率之子帶具有最少 bin(如 1 個)，且每子帶之 bin 的數目隨著頻率漸增而增加。

【0024】 回到第 1 圖，由每一聲道之各濾波器排組(在此例中為濾波器排組 2 與 4)產生的每一 n 個時間域輸入聲道的一頻率域版本利用加法組合功能與裝置(加法組合器 6)被加在一起(向下混頻)成為單聲道(mono)合成音訊信號。

【0025】 該向下混頻可被施用至該等輸入音訊信號之整個頻寬，或備選地其可被限制於高於某一特定「耦合」頻率，因此向下混頻處理之人工物可在中至低頻率變得更可聽到的。在這類情形中，該等聲道可在低於該耦合頻率離散地被輸送。此策略可為所欲的，就算處理人工物並非問題所在，原因在於藉由將變換 bin 組成為類似關鍵頻帶(大小大略與頻率成比例)所構建之中/低頻率子帶在低頻率具有小數目之變換 bin(在非常低頻率為 1 bin)且以少數或比傳送具有支鏈資訊之向下混頻的單聲道音訊信號少之位元直接地被編碼。在本發明之層面的實際實施例中，低到如 2300Hz 之耦合頻率被發現為適合的。然而，該耦合頻率並非關鍵的，且較低的耦合頻率，甚至是在被施用於編碼器之音訊信號頻帶底部的耦合頻率就某些應用，特別是非常低位元率為重要者為可接受的。

【0026】 在向下混頻前，本發明之一層面為要改善聲道相位之彼此相對的對準角，以降低該等聲道被組合時不同相位信號成份之抵銷及提供改善的單聲道合成聲道。此可藉由將一些聲道之一些或全部變換 bin 隨著時間可控制地移位「絕對角」而被完成。例如，代表高於一耦合頻率之音訊的全部變換 bin(因而定義所論及之頻帶)在當一聲道被用作為基準時，除了該參考聲道外的所有聲道，或在每一聲道，於必要時被隨著時間可控制地移

位。

【0027】 一 bin 之「絕對角」可採用為用一濾波器排組被產生之每一複數值變換 bin 的振幅與角度呈現之角度。Bin 在一聲道之絕對角可控制的移位利用角旋轉功能與裝置(旋轉角)被實施。旋轉角 8 可在濾波器排組 2 之輸出施用至加法組合器 6 所提供之向下混頻加總前處理該輸出，而旋轉角 10 可在濾波器排組 4 之輸出施用至加法組合器 6 所提供之向下混頻加總前處理該輸出。其將被了解，在某些信號條件下，對一時期(在此處描述之例子中為一訊框之時期)而言，特定的變換 bin 可不需要角旋轉。在低於耦合頻率下，該聲道資訊可離散地被編碼(第 1 圖中未畫出)。

【0028】 原則上，聲道之相位角彼此對齊可在所論及之整個頻帶的每一區塊利用其絕對相位角之負數將每一變換 bin 或子帶相位移位被完成。雖然此實質上避免不同相位信號成份之抵銷，其易於致使人造物為可聽到的，特別是若該單聲道合成信號以隔離被聆聽時。因而，其欲藉由最多僅如使向下混頻處理中不同處理抵銷最小化與使解碼器重新構成之多聲道信號的空間影像崩潰最小化所必要地將一聲道之 bin 的絕對角移位。用於決定此角移位之一較佳的技術在下面被描述。

【0029】 能量常規化如下面進一步描述地亦可在編碼器中以每一 bin 之基準被實施。亦如下面進一步描述地能量常規化亦可以每一子帶之基準(在解碼器內)被實施以確保單聲道合成信號之能量等於該等歸因聲道之能量和。

【0030】 每一輸入聲道具有與其相關之一音訊分析器功能與裝置(音訊分析器)用於為此聲道產生支鏈資訊及用於在其被施用於向下混頻加法 6 前控制被施用於該聲道之角旋轉的數量或角度。聲道 1 與 n 之濾波器排組輸出分別被施用於音訊分析器 12 與音訊分析器 14。音訊分析器 12 為聲道 1 產生支鏈資訊或角旋轉的數量。音訊分析器 14 為聲道 n 產生支鏈資訊或角旋轉的數量。其將被此處所稱之「角」係指相位角。

【0031】 用一音訊分析器為每一聲道產生之每一聲道的支鏈資訊可包括：

【0032】 一振幅標度因數(振幅 SF)

【0033】 一角度控制參數，

【0034】 一解除相關標度因數(解除相關 SF)，及

【0035】 一暫態旗標。

【0036】 此支鏈資訊可被特徵化爲「空間參數」表示該等聲道之空間性質及/或表示與空間處理相關之信號特徵，如暫態。在每一情形中，該支鏈資訊於用於單一子帶(暫態旗標除外，其施用於一聲道內之所有子帶)且可如下面描述之例子地就每訊框或就相關編碼器中之一區塊切換發生被更新一次。編碼器中特定聲道之角旋轉可被採用作爲極性逆轉後之角控制參數。

【0037】 若一參考聲道被運用，此聲道可不需要一音訊分析器，或替選地可需要一音訊分析器，其僅產生振幅標度因數支鏈資訊。若一振幅標度因數可用一解碼器由其他非參考聲道之振幅標度因數以充分的精確度被導出，便沒必要傳送該標度因數。若在編碼器之能量常規化確保在任一子帶內所有聲道之標度因數平方和如下面描述地實質等於 1，則在該解碼器中導出該參考聲道之振幅標度因數的近似值爲可能的。該被導出之振幅標度因數近似值會因在所再生之多聲道音訊中造成影像位移結果的振幅標度因數之相對粗略數量化所致具有誤差的結果。然而在低資料率環境中，此類人工物比起使用該等位元來傳送該參考聲道之振幅標度因數是比較能接受的。不過在某些情形中，其可能欲爲至少產生振幅標度因數支鏈資訊之參考聲道運用一音訊分析器。

【0038】 第 1 圖以虛線顯示由 PCM 時間域輸入至聲道中之音訊分析器的備選輸入。此輸入可被音訊分析器使用以偵測一時期(在此處描述之例中爲一區塊或一訊框之期間)上的暫態及在響應一暫態下產生一暫態指標(如一位元之「暫態旗標」)。或替選地如下面描述者，一暫態可在頻率域中被偵測，音訊分析器在此情形中不須接收一時間域輸入。

【0039】 全部聲道(或除了參考聲道外之全部聲道)所用的單聲道合成信號與支鏈資訊可被儲存、傳輸、或儲存且傳輸至一解碼功能與裝置(解碼器)。除了基本的儲存、傳輸、或儲存且傳輸外，各種音訊信號與各種支鏈資訊可被多工及被封裝爲一個或更多的位元流適用於儲存、傳輸、或儲存且傳輸媒體。該單聲道合成音訊可在儲存、傳輸、或儲存且傳輸前被施

用於一資料率降低的編碼功能與裝置，例如為一感覺編碼器，或被施用於一感覺編碼器與一熵編碼器(如算術或赫夫曼(Huffman)編碼器)(有時被稱為「無損失」編碼器)。同時如上面提及者，該等單聲道合成音訊與相關的支鏈資訊可僅為高於某一頻率(耦合頻率)之音訊頻率由多輸入聲道被導出。在此情形中，在每一多輸入聲道中低於耦合頻率之音訊頻率可被儲存、傳輸、或儲存且傳輸作為離散的聲道，或可用非此處所描述的一些方式被組合或處理。這類離散或否則被組合之聲道亦可被施用於一資料率降低的編碼功能與裝置，例如為一感覺編碼器，或被施用於一感覺編碼器與一熵編碼器。該等單聲道合成音訊與離散多聲道音訊可都被施用於一整合的感覺編碼或感覺及熵編碼功能與裝置。該等各種支鏈資訊可被承載於否則未被使用或資訊隱藏式地在該音訊資訊之被編碼的形式內。

【0040】 基本的 1 : N 與 1 : M 解碼器

【0041】 參照第 2 圖，實施本發明之層面之一解碼器功能與裝置(解碼器)被顯示。此圖為實施本發明之層面的基本解碼器之功能或構造的例子。實作本發明之層面之其他功能或構造配置可被運用，包括下面被描述之替選的及/或功能或構造配置。

【0042】 該解碼器為所有聲道或除了參考聲道之所有聲道接收單聲道合成音訊信號與支鏈資訊。必要時，該等單聲道合成音訊信號與相關的支鏈資訊被解除多工、解除封包及/或解碼。解碼可運用一檢查表，其目標為要以此處被描述之本發明的位元率降低技術來由該單聲道合成音訊聲道導出數個各別的音訊聲道近似於被施用於第 1 圖之編碼器的各音訊聲道。

【0043】 當然，吾人可選擇不恢復被施用至編碼器之所有聲道或僅使用單聲道合成信號。替選的是，除了被施用至編碼器之聲道外可藉由實施本發明之層面的 2002 年 2 月 7 日申請、2002 年 8 月 15 日申請之指定給美國的國際專利申請案第 PCT/US 02/03619 號及其結果所得之 2003 年 8 月 5 日申請的美國申請案 S.N. 10/467,213 號與 2003 年 8 月 6 日申請、2004 年 3 月 4 日申請之指定給美國的國際專利申請案第 WO 2004/019656 號及其結果所得之 2005 年 1 月 27 日申請的美國申請案 S.N. 10/522,515 號而依據本發明之層面由一解碼器之輸出被導出。該等申請案之整體被納於此處做為參

考。用實施本發明之層面的解碼器所恢復之聲道在所述且被採納之申請案的相關聲道多工技術中特別有用之處不僅在於具有有用的聲道間振幅關係也具有有用的聲道間相位關係。另一替選做法為運用矩陣解碼器以導出額外的聲道。本發明之層面的聲道間振幅與相位保存使得實施本發明之層面的解碼器之輸出聲道特別適用於振幅與相位敏感的矩陣解碼器。例如，若本發明之層面在 $N:1:N$ 系統中被實施(其中 $N=2$)，被解碼器恢復之二聲道可被施用至一 $2:M$ 有作用的矩陣解碼器。很多有用的矩陣解碼器為本技藝相當習知的，包括“Pro Logic”與“Pro Logic II”解碼器(“Pro Logic 為杜比實驗室發照公司的註冊商標)及在下列一個或更多美國專利與公告之國際申請案(每一個指定給美國)所揭示之主題事項實施層面的矩陣解碼器：4,799,260；4,941,177；5,046,098；5,274,740；5,400,433；5,625,696；5,644,640；5,504,819；5,428,687；5,172,415；WO 01/41504；WO 01/41505；以及 WO 02/19768，其整體被納於此處做為參考。

【0044】 再參照第 2 圖，該被接收之單聲道合成音訊聲道被施用至數個信號路徑，各被恢復之多聲道音訊由此被導出。每一聲道導出之路徑包括一振幅調整功能與裝置(調整振幅)與一角旋轉功能與裝置(角旋轉)，其順序為二者均可。

【0045】 該調整振幅對單聲道合成信號施用增益或損失，使得在某些信號狀況下由其被導出之輸出聲道的相對輸出振幅(或能量)類似在編碼器的輸入聲道者。替選的是，在某些信號狀況下當「隨機化」角變異如接著被描述地被施加時，一可控制數量之「隨機化」振幅變異亦可被施加至被恢復之聲道的振幅以改善其針對其他被恢復之聲道的解除相關。

【0046】 該等角旋轉施用相位旋轉，使得在某些信號狀況下由單聲道合成信號被導出之輸出聲道的相對相位角類似編碼器之輸入聲道者。較佳的是，在某些信號狀況下，一可控制數量之「隨機化」角變異亦可被施加至被恢復之聲道的角以改善其針對其他被恢復之聲道的解除相關。

【0047】 如下面進一步被討論者，「隨機化」角振幅變異不僅包括虛擬隨機與真正隨機變異，亦包括確定產生之變異，其具有降低聲道間交叉相關之效果。

【0048】 概念上，調整振幅與角旋轉為特定聲道比例調整單聲道合成音訊 DFT 係數而為該聲道得到重建之變換 bin 的值。

【0049】 每一聲道之調整振幅可至少用被恢復之支鏈標度因數為特定聲道，在參考聲道的情形，由該被恢復之支鏈標度因數為該參考聲道；或在其他非參考聲道的情形，由該被恢復之支鏈標度因數被導出的振幅標度因數被控制。替選的是，為強化該等恢復之聲道的解除相關，該調整振幅亦可用為一特定聲道由該被恢復之支鏈標度因數與為該特定聲道的被恢復之支鏈暫態旗標被導出之一隨機化振幅標度因數參數被控制。每一聲道之角旋轉可至少用該被恢復之支鏈角控制參數(在此情形中，解碼器中之角旋轉實質上可不進行編碼器中之角旋轉所提供的角旋轉)被控制。為強化該等恢復之聲道的解除相關，角旋轉亦可用為特定聲道由該被恢復之支鏈解除相關標度因數與該被恢復之支鏈暫態旗標被導出的隨機化角控制參數被控制。一聲道之隨機化控制參數與若有被運用之一聲道的隨機化振幅標度因數可用一可控制的解除相關器功能與裝置(可控制的解除相關器)由該聲道之該被恢復之解除相關標度因數與該聲道之該被恢復之暫態旗標被導出。

【0050】 參照第 2 圖之例子，該該被恢復之單聲道合成音訊被施用至一第一聲道音訊恢復路徑 22，其導出該聲道 1 音訊及被施用至一第二聲道音訊恢復路徑 24，其導出該聲道 n 音訊。音訊路徑 22 包括一調整振幅 26、一角旋轉 28、及若 PCM 輸出為所欲時之逆濾波器排組功能與裝置(逆功能與裝置)30。類似地，音訊路徑 24 包括一調整振幅 32、一角旋轉 34、及若 PCM 輸出為所欲時之逆濾波器排組功能與裝置(逆功能與裝置)36。就如第 1 圖之情形，為了呈現簡單起見，只有二聲道被顯示，其將被了解聲道可多於二個。

【0051】 第一聲道(聲道 1)之該被恢復之支鏈資訊如上述相關基本編碼器所述地可包括一振幅標度因數、一角控制參數、一解除相關標度因數與一暫態旗標。振幅標度因數被施用至調整振幅 26。暫態旗標與解除相關標度因數被施用至一可控制的解除相關器 38，其在對此響應下產生一隨機化角控制參數。該一位元之暫態旗標的狀態如下面進一步解釋地隨機化角

解除相關的二多重模式之一。該角控制參數與隨機化角控制參數用一加法組合器或組合功能 40 被加在一起而為角旋轉 28 提供一控制信號。替選的是，可控制的解除相關器 38 在除了產生一隨機化角控制參數外亦可在響應暫態旗標與解除相關標度因數下產生一隨機化振幅標度因數。該振幅標度因數可與一隨機化振幅標度因數用一加法組合器或組合功能(未畫出)被相加而為調整振幅 26 提供控制信號。

【0052】 類似地，第二聲道(聲道 n)之該被恢復之支鏈資訊如上述相關基本編碼器所述地可包括一振幅標度因數、一角控制參數、一解除相關標度因數與一暫態旗標。振幅標度因數被施用至振幅 32。暫態旗標與解除相關標度因數被施用至一可控制的解除相關器 42，其在對此響應下產生一隨機化角控制參數。如聲道 1 者，該一位元之暫態旗標的狀態如下面進一步解釋地隨機化角解除相關的二多重模式之一。該角控制參數與隨機化角控制參數用一加法組合器或組合功能 44 被加在一起而為角旋轉 34 提供一控制信號。替選地如配合聲道 1 所描述的是，可控制的解除相關器 42 在除了產生一隨機化角控制參數外亦可在響應暫態旗標與解除相關標度因數下產生一隨機化振幅標度因數。該振幅標度因數可與一隨機化振幅標度因數用一加法組合器或組合功能(未畫出)被相加而為調整振幅 32 提供控制信號。

【0053】 雖然剛所描述之一處理或拓樸就了解是有用的，基本上相同的結果可用達成相同或類似結果之替選的處理或拓樸被獲得。例如，調整振幅 26(32)與角旋轉 28(34)之順序可被逆轉及/或其有一個以上的角旋轉——一個響應角控制參數及另一個響應隨機化角控制參數。角旋轉亦可被視為下面第 5 圖描述之例子中的三個而非一或二個功能與裝置。若一隨機化振幅標度因數被運用，其可有多於一個之調整振幅——一個響應振幅標度因數及另一個響應隨機化振幅調整振幅。由於人耳對振幅相對於相位之較敏感，若一隨機化振幅調整振幅被運用，其可能欲將其效應相對於隨機化角控制參數之效應比例調整，使得其對振幅之效應小於隨機化角控制參數對相位角之效應。至於另一替選的處理或拓樸，該解除相關標度因數可被用以控制隨機化相位角移位對基本相位角移位之比值，及若如此被運用之隨

機化振幅移位對基本振幅移位之比值(即在每一情形中之可變的交叉衰減)。

【0054】 若一參考聲道如上面相關基本編碼器所討論地被運用，該聲道用之角旋轉、可控制的解除相關器與加法組合器可被省略，如此該參考聲道之支鏈資訊可僅包括振幅標度因數(或替選地，若該支鏈資訊就該參考聲道不含有振幅標度因數，其可在編碼器中之能量常規化確保一子帶內整個聲道的標度因數平方和為 1 時由其他聲道之振幅標度因數被導出)。一調整振幅就該參考聲道被提供且其就該參考聲道用被接收或被導出之振幅標度因數被控制。每當該參考聲道之振幅標度因數由支鏈被導出或在解碼器被導出，該被恢復之參考聲道為該合成單聲道的振幅標度調整後之形式。由於其是其他聲道旋轉之基準，其不需角旋轉。

【0055】 雖然調整該被恢復之聲道的相對振幅可提供最緩和程度之解除相關，若單獨被使用，振幅調整可能形成實質上缺乏很多信號狀況之空間化或成像的再生音場(如「潰散的」音場)。振幅調整可能影響耳之內部聲音位準差異，其為耳朵所運用之心理上聲響方向性清晰之一。因而，依據本發明之層面，某些角度調整技術可視信號狀況被運用以提供額外的解除相關。參照表 1，其提供了解複式角度調整技術或依據本發明之層面被運用的作業模式為有用的。其他在下面配合第 8 與 9 圖之例子被描述的解除相關技術可除了或取代第 1 圖之技術外被運用。

【0056】 在實務上，施用角旋轉與振幅變更可形成圓圈迴旋(亦被習知為循環或週期性迴旋)。雖然一般而言欲避免圓圈迴旋，其可在本發明之層面之低成本施作被容忍，特別是其中向下混頻為單聲道或多聲道僅在如高於 1500Hz 之音訊頻帶部分發生之情形(此情形中之圓圈迴旋的可聽到之效應為最小的)。替選的是，圓圈迴旋可用任一適合的技術被避免或最小化，例如包括零填入之適當使用。使用零填入之一方法為變換所提出之頻率域變異(角旋轉與調整振幅)為時間域、將之視窗化(用任意的視窗)、用零填入，然後變換回頻率域並乘以將被處理之音訊(該音訊不須被視窗化)的頻率域形式。

【0057】 表 1 角度調整解除相關技術

	技術1	技術2	技術3
--	-----	-----	-----

信號型式 (典型例)	頻譜上為靜態來源	複數之連續信號	複數之脈衝性信號 (暫態)
對解除相關 之效應	將低頻率及穩定狀態之信號成份解除相關	將非脈衝性之複數信號成份解除相關	將脈衝性高頻率信號成份解除相關
對訊框中出現之暫態的效應	以縮短的時間常數操作	不操作	操作
完成事項	在一聲道中bin角之緩慢移位(逐一訊框地)	在一聲道中以逐一訊框的基準對技術1之角移位添加一隨機化之角移位	在一聲道中以逐一子帶的基準對技術1之角移位添加迅速變化(逐一區塊)之隨機化角移位
加以控制或比例調整	基本移位程度用角控制參數被控制	額外移位程度直接用解除相關SF被比例調整; 整個子帶用同一調整, 每一訊框更新調整因數	額外移位程度間接用解除相關SF被比例調整; 整個子帶用同一調整, 每一訊框更新調整因數
角移位之頻率解析度	子帶(在每一子帶中被施用至所有bin之同一或被內插之移位值)	bin(被施用至每一bin之不同的隨機化移位值)	子帶(在每一子帶被施用至所有bin之同一隨機化移位值; 在聲道中被施用至每一子帶之不同的隨機化移位值)
時間解析度	訊框(每訊框更新移位值)	隨機化移位值維持相同不變	區塊(每區塊更新隨機化移位值)

【0058】 就例如為高音管音調之頻譜上實質為靜態的信號而言，一第一技術(技術 1)相對於每一其他該被恢復之聲道的角恢復該被接收之單聲道合成信號的角為類似聲道的原始角相對於該編碼器之輸入的其他聲道之角(受限於頻率與時間顆粒度及受限於數量化)。相位角差異為有用的，特別是用於提供低於約 1500Hz 之低頻率信號成份，此處耳朵會遵循該音訊信號之各別的週期。較佳的是，技術 1 在所有信號狀況下操作以提供基本的角移位。

【0059】 就高於約 1500Hz 之高頻率信號成份而言，耳朵不遵循聲音之各別週期，而是代之對波形包線響應(以關鍵頻帶為基準)。因此，高於約 1500Hz 之解除相關最好是用信號包線之差異而非相位角差異被提供。僅依照技術 1 施用相位角差異不會變更信號包線差異到足以將高頻率信號解除相關。該等第二與第三技術(技術 2 與技術 3)在某些信號狀況下添加可控制數量之隨機化角變異至技術 1 所決定之角而致使造成可控制數量之包線變異，此可強化解除相關。

【0060】 相位角之隨機化變化為造成信號包線之隨機化變化的一種

所欲之方法。一特定的包線係為在一子帶內頻譜成份之振幅與相位的特定組合之相互作用的結果。雖然改變一子帶內頻譜成份之振幅，大的振幅改變被要求以獲得在包線內重大的改變，由於人耳對頻譜振幅之變異為敏感的，故此非所欲的。對照之下，改變頻譜成份之相位角對包線的影響比起改變頻譜成份之振幅者較大—頻譜成份不再以相同方式對齊，所以定義該包線之強化與減除在不同時間發生而改變該包線。雖然人耳對包線有一些敏感，人耳對相位是相對上為聾的，故整體的音響品質維持實質上類似的。不過就一些信號狀況而言，頻譜成份之振幅以及相位的隨機化在假設此振幅隨機化不會造成不欲有之可聽到的人工物下可提供強化的信號包線隨機化。

【0061】 較佳的是，一可控制程度技術 2 或技術 3 在某些信號狀況下與技術 1 一起操作。暫態旗標選擇技術 2(視暫態旗標是以訊框或區塊率被傳送，訊框或區塊中未出現暫態)或技術 3(訊框或區塊中有出現暫態)。因而，其有多種操作模式，依暫態是否出現而定。替選的是，此外在某些信號狀況下，一可控制程度的振幅隨機化亦與尋求要恢復原始聲道振幅之調整振幅一起操作。

【0062】 技術 2 適用於複數連續信號，其如大量管弦提琴，在諧振合弦是很豐富的。技術 3 適用於複數脈衝性或暫態信號，如鼓掌聲與響板等(技術 2 在鼓掌中夾雜爆裂聲使其不適用於此類信號)。如下面進一步解釋者，為了使可聽到的人工物最少，技術 2 與技術 3 其中不同的時間與頻率解析度用於施用隨機化角度異—當暫態未出現時技術 2 被選擇，而當暫態出現時技術 3 被選擇。

【0063】 技術 1 緩慢地(逐一訊框)移位在一聲道中之 bin 角。此基本移位程度用角控制參數被控制(若參數為 0 便無移位)。如下面進一步解釋者，同一或被內插之參數被施用至子帶中之所有 bin 且該參數在每訊框被更新。後果為每一聲道之每一子帶可針對其他聲道具有一相位移位，提供在低頻率(低於 1500Hz)之一程度的解除相關。就此類信號狀況而言，再生之聲道會展現惱人的不穩定之梳濾波器效應。在掌聲之情形中，由於所有聲道在一訊框期間傾向具有相同振幅，基本上無解除相關藉由調整該被恢復

之聲道的相對振幅被提供。

【0064】 技術 2 在暫態未出現時操作。技術 2 在一聲道中以逐一 bin 之基準(每一 bin 具有不同的隨機化移位)添加不隨時間變化之一隨機化角移位至技術 1 之角移位，致使該等聲道之包線彼此不同而提供聲道間之複數信號的解除相關。對時間維持隨機化相位角值為固定係可避免區塊或訊框人工物，此可能是由 bin 相位角之區塊對區塊或訊框對訊框變更所致之結果。雖然此技術在暫態未出現時是非常有用的解除相關，其可能暫時污損一暫態(形成經常被稱為「前置雜訊」)之結果，而後暫態污損被提供暫態遮蔽。技術 2 提供之添加移位的程度用解除相關標度因數直接被比例調整(若標度因數為 0 便無添加的移位)。理想上，被添加至基本角移位(技術 1)之隨機化相位角數量用解除相關標度因數被控制，其方式為避免可聽到的信號清晰人工物。雖然不同的添加隨機化角移位值被施用至每一 bin 及此移位值未改變，相同的比例調整被施用至整個子帶且該比例調整在每一訊框被更新。

【0065】 技術 3 在訊框或區塊中有暫態出現時操作，視暫態旗標被傳送之比率而定。其以對子帶中所有 bin 為相同之一獨一隨機化角度值逐一區塊地移位一聲道中每一子帶中的所有 bin，不僅致使訊框的信號中之包線亦致使振幅與相位針對其他聲道隨著區塊而改變。此減少訊框間之穩定狀態信號的類似性並提供聲道之解除相關而實質地不致有「前置雜訊」人工物。當二個或更多聲道在其由擴音器至聽者的途徑上以聲響混頻時，雖然人耳不直接於高頻率對純粹角度變化響應，相位差異會造成振幅變化(梳濾波器效應)，其可能是可聽到且討厭的，這些可用技術 3 粉碎。信號之脈衝性特徵使可能否則會發生之區塊率人工物最小化。因而，技術 3 在一聲道中以逐一子帶之基準添加迅速變化(逐一區塊地)隨機化角移位至技術 1 之相位移位。添加移位之程度如下面描述地用解除相關標度因數間接地被比例調整(若標度因數為 0 便無添加移位)。相同的比例調整被施用至整個子帶且該比例調整在每一訊框被更新。

【0066】 雖然角度調整已被特徵化為三種技術，但此為語意上的問題，且其亦可被特徵化為二種技術：(1)技術 1 為可變程度(可能為 0)之技術

2的組合，及(2)技術1為可變程度(可能為0)之技術3的組合。為了方便呈現，該等技術被視為三種技術。

【0067】 多模式解除相關技術之層面與其修改可在提供例如用向下混頻由一個或更多音訊聲道被導出之音訊信號的解除相關中被運用，就算此類音訊聲道並非由依據本發明之層面之編碼器被導出亦然。這類配置在被施用至單聲道合成音訊時有時被稱為「虛擬立體聲」功能與裝置。任何適合的功能與裝置(向上混頻器)可被運用以由單聲道音訊或多聲道音訊導出多重信號。一旦此類多聲道音訊用一向上混頻器被導出，其一個或更多可針對一個或更多其他被導出之音訊信號藉由施用此處所描述之多模式解除相關技術被導出。在此應用中，該等解除相關技術被施用之每一被導出的音訊聲道可藉由偵測該被導出之音訊聲道本身中之暫態而由一操作模式切換至另一個。替選的是，有暫態出現之技術(技術3)的操作可被簡化，以在暫態出現時以提供頻譜成份之相位角的無移位。

【0068】 支鏈資訊

【0069】 如上述者，該支鏈資訊可包括：一振幅標度因數、一角控制參數一解除相關標度因數與一暫態旗標。實施本發明之層面之此支鏈資訊可彙整如下列表2。典型上，該支鏈資訊可每一訊框被更新一次。

【0070】 表2 一聲道之支鏈資訊特徵

支鏈資訊	範圍	代表 (一種量測)	數量化 水準	主要目的
子帶角 控制參 數	$0 \rightarrow +2\pi$	一聲道中之子帶 每一bin與一參 考聲道中對應的 bin之角度間差 異在整個子帶時 間上之平滑平均 數	6位元(64個水準)	為聲道中之每一 bin提供基本的 角旋轉
子帶解除 相關因 數	$0 \rightarrow 1$ 子帶解除相關 標度因數與聲 道間角一致性 因數均為高值	一聲道中之子帶 對時間之信號特 徵之頻譜穩定性 (頻譜穩定性因 數)與一子帶的 角對一致性(聲 道間角一致性因 數)	3位元(8個水準)	對被添加至基本 角旋轉之隨機化 角移位(若例被 運轉時)亦對被添 加之隨機化振地 整至基本因數比 因幅標度及備選 對調整響程
支鏈資訊	範圍	代表 (一種量測)	數量化 水準	主要目的

子帶振幅標度因數	0 至 31(全為整振數), 0 為最高幅, 31 為最低幅	一聲道中之子帶的對所有子帶的振幅	5 位元(32 個水準)顆粒度為 1.5dB, 故範圍為 $31 \times 1.5\text{dB} = 46.5\text{dB}$ 以上 最終值=off	對一聲道中之一子帶的 bin 之振幅比例調整
暫態旗標	1, 0 (真/假) (極性為任意)	在訊框或區塊中暫態之出現	1 位元(2 水準)	決定要運用那一種添加隨機化角移位或角移位及振幅移位的技術

【0071】 在每一情形中，一聲道之支鏈資訊施用至單一子帶(暫態旗標除外，其施用至所有子帶)且每一訊框被更新一次。雖然所表示之時間解析度(每一訊框一次)、頻率解析度(子帶)、數值範圍與數量化水準已被發現在低位元率與績效間提供有用的績效及有用的折衷，這些時間與頻率解析度、數值範圍與數量化水準並非關鍵的，且其他的解析數、範圍與水準可在實施本發明之層面的被運用。例如，該暫態旗標可每一區塊被更新一次而支鏈資料費用的增加僅為最小的，如此做的優點為切換技術 2 至技術 3 可更精確，反之亦然。此外如上述者，支鏈資訊可根據相關編碼器之區塊切換的出現被更新。

【0072】 其將被指出，上述的技術 2(見表 1)提供 bin 頻率解析度而非子帶頻率解析度(即不同的虛擬隨機相位角移位被施用至每一 bin 而非每一子帶)，就算同一子帶解除相關標度因數被施用至一子帶之所有 bin 亦然。其亦將被指出，上述的技術 3(見表 1)提供區塊頻率解析度(即不同的隨機化相位角移位被施用至每一區塊而非每一訊框)，就算同一子帶解除相關標度因數被施用至一子帶之所有區塊亦然。大於支鏈資訊之解析度的此類解析度為可能的，原因在於該隨機化相位角移位可在一解碼器被產生且不須在編碼器中被知道(就算該編碼器亦施用一隨機化相位角移位至被編碼之單聲道合成信號，此情形亦然，此為下面被描述之一替選方式)。換言之，沒有必要傳送具有 bin 或區塊顆粒度之支鏈資訊，就該等解除相關技術運的暫態偵測器而被強化，以提供比訊框率甚至是比區塊率更精細的時間解析度。此補充性的暫態偵測器可偵測在該解碼器所接收之單聲道或多聲道合成音訊信號中的暫態之發生，且此偵測資訊被轉送至每一可控制的解除相關器(如第 2 圖之 38, 42)。然後在接收其暫態旗標之際，該可控制的解除相關

器於接收該解碼器之局部偵測資訊指示時由技術 2 切換為技術 3。因而，時間解析度之實質改善在不提高支鏈位元率(縱然是空間精確度降低)為可能的(該編碼器在其向下混頻前偵測每一輸入聲道中之暫態，而解碼器中之偵測在向下混頻後完成)。

【0073】 作為對逐一訊框基準傳送支鏈資訊的替選方式，支鏈資訊至少可就高度動態的信號在每一區塊被更新。如上述者，在每一區塊更新暫態旗標形成支鏈資料費用增加很小之結果。為了不實質地提高支鏈資料率地完成其他支鏈資訊的時間解析度之此提高，一區塊浮動點差別編碼可被使用。例如，連續的變換區塊可對一訊框以 6 個一組被收集。完整支鏈資訊可為第一區塊中之每一子帶聲道被傳送。在後續的 5 個區塊中，僅有差分值被送出，每一個為目前區塊之振幅與角度及來自前一區塊之同等值間的差。此就如高音管音調之靜態信號形成非常低資料率之結果。就較為動態的區塊而言需要較大範圍之差異值但較不精準。所以就每一個 5 差異值之群組而言，一指數可使用 3 位元首先被傳送，然後差異值被數量化為例如 2 位元之精確度。此配置以大約為 2 之因子降低平均最壞情形的支鏈資料率。進一步之降低可藉由如上面討論地為一參考聲道省略支鏈資料(由於其他聲道被導出)及例如使用算術編碼被獲得。此外或替選地，整個頻率之差別編碼可藉由例如子帶角度或振幅之差異被運用。

【0074】 不論支鏈資訊是以逐一訊框基準或更明頻繁地被傳送，在一訊框中的各區塊內插支鏈值為有用的。對時間之線性內插可如下面描述地以對頻率之線性內插被運用。

【0075】 本發明之層面之適合的施作運用處理步驟或裝置，其如接著被設立地施作各處理步驟。雖然下列編碼與解碼步驟可用電腦軟體指令序列以下面列出之步驟順序被實施，其將被了解等值或類似結果可在考慮某些數量由較早者被導出下以其他方式之順序的步驟被獲得。例如多線之電腦軟體指令序列可被運用，使得某些步驟序列並行地被實施。替選的是，所描述之步驟可被施作為實施所欲功能之裝置，該等各種裝置具有如此後被描述之功能性的相互關係。

【0076】 編碼

【0077】 該編碼器或編碼功能可在一訊框導出支鏈資訊前收集一訊框之資料份量，並將該訊框之音訊聲道向下混頻為一單聲道音訊聲道(以上述第 1 圖之方式，或以下面描述之第 6 圖的方式變為多聲道音訊)。藉由如此做，支鏈資訊可首先被傳送至一解碼器，允許解碼器在接收單聲道或多聲道音訊資訊之際立刻開始解碼。編碼處理之步驟(編碼步驟)可如下列地被描述。針對編碼步驟參照第 4 圖，其為混合式流程圖與功能方塊圖之性質。透過步驟 419，第 4 圖顯示一聲道用之編碼步驟。步驟 420 與 421 施用至所有多聲道，其被組合以提供一合成單聲道信號輸出或一起被做成矩陣以如下面相關第 6 圖描述地提供多聲道。

【0078】 步驟 401 偵測暫態

【0079】 a. 實施一輸入音訊聲道中之 PCM 值的暫態偵測。

【0080】 b. 若一暫態在該聲道之一訊框的任一區塊出現，設定一個 1 位元之暫態旗標為真。

【0081】 有關步驟 401 之註解：

【0082】 該暫態旗標形成一部分之支鏈資訊且如下面描述地亦在步驟 411 中被使用。在解碼器中比區塊率細之暫態解析度可改善解碼器績效。雖然如上面討論地，一區塊率而非訊框率暫態旗標可用位元率最緩和之增加形成一部分之支鏈資訊，類似但空間精確度降低之結果可藉由偵測在解碼器中被接收之單聲道合成信號中的暫態發生而不致提高支鏈位元率地完成。

【0083】 每一訊框之每一聲道有一暫態旗標，其原因為在時間域被導出，有必要施用至此聲道之所有子帶。該暫態偵測可以類似 AC-3 編碼器中所運用之方式被實施，用於控制何時要在長與短音訊區塊間切換之決策，但具有較高的敏感度及就其中一區塊之暫態旗標為真的任一訊框其暫態旗標為真(AC-3 編碼器以區塊之基準偵測暫態)。特別是參見上述 A/52A 文件之第 8.2.2 節。第 8.2.2 節描述之偵測暫態的敏感度可藉由添加一敏感度因數 F 至其中被設立之公式而被提高。A/52A 文件之第 8.2.2 節在下面設立，敏感度因數已被加入(如下面被再製之第 8.2.2 節被修正以表示其低通濾波器為一種串接二階(cascaded biquad)直接型式 II 之 IIR 濾波器而非公布之

A/52A 文件中的「型式 I」；第 8.2.2 節在較早之 A/52 文件中被修正。雖然並非關鍵的，0.2 之敏感度因數已被發現是為本發明之層面之實施例的一適合之值。

【0084】 替選的是，在美國專利第 5,394,473 號所描述之類似的暫態偵測技術可被運用。該“473 專利更詳細地描述 A/52A 文件之暫態偵測器的層面。A/52A 文件與“473 專利二者均以整體被納於此處做為參考。

【0085】 另一替選的是，暫態可在頻率域而非時間域中被偵測。在此情形中，步驟 401 可被省略，及在頻率域中被運用之一替選的步驟在下面被描述。

【0086】 步驟 402 視窗化與 DFT

【0087】 將 PCM 時間樣本之重疊區塊乘以一時間視窗並經由用一 FFT 所施作之一 DFT 將之變換為複數頻率值。

【0088】 步驟 403 變換複數值為振幅與角度

【0089】 使用標準複數操作變換每一頻率域複數變換 bin 值(a+bj)為振幅與角度呈現：

【0090】 a. 振幅=square_root(a^2+b^2)

【0091】 b. 角度=arctan(b/a)

【0092】 有關步驟 403 之註解：

【0093】 一些下列步驟可使用一 bin 之能量被定義為上述振幅之平方(即能量=(a^2+b^2))而作為一替選做法。

【0094】 步驟 404 計算子帶能量

【0095】 a. 藉由將每一子帶內之 bin 能量值相加(對整個頻率加總)而計算每一區塊之子帶能量。

【0096】 b. 藉由平均或累積一訊框中之所有區塊(對整個時間之平均/累積)而計算每一訊框之子帶能量。

【0097】 c. 若該編碼器之聲道耦合頻率低於約 1000Hz，施用于子帶訊框平均後或訊框累積後之能量至一時間平滑器，其對低於此頻率且高於該耦合頻率之所有子帶操作。

【0098】 有關步驟 404c 之註解：

【0099】 在低頻率子帶提供訊框間平滑之時間平滑會是有用的。爲了避免在子帶界限之 bin 值間人工物所造成的不連續，由包容且高於該耦合頻率的最低頻率子帶(平滑在此處具有顯著效果)一直到其中該時間平滑效果爲可測量的(但爲聽不到的，雖然是幾乎可聽到)較高頻率子帶施用一種漸進降低之時間平滑爲有用的。對最低頻率範圍子帶(此處若子帶爲關鍵頻帶，其爲單一之 bin)爲適合的時間常數例如爲在 50 至 100 毫秒之範圍內。該漸進降低之時間平滑可持續至包容約 1000Hz 之一子帶，此處該時間常數例如可爲約 10 毫秒。

【00100】 雖然一第一階之平滑器爲適合的，該平滑器可爲一個二階段平滑器，其具有一可變的時間常數縮短其在響應一暫態下的攻擊與延遲時間(此種二階段平滑器可爲美國專利第 3,846,719 與 4,922,535 號所描述之類比二階段平滑器的數位等值物，其每一專利之整體被納於此處做爲參考)。該穩定狀態之時間常數可依據頻率被比例調整且亦可在響應一暫態下爲可變的。替選的是，此平滑可在步驟 412 中被施用。

【00101】 步驟 405 計算 bin 量之和

【00102】 a. 計算每一子帶之每區塊 bin 量(步驟 403)的和(整個頻率之加總)。

【00103】 b. 藉由對一訊框中整個區塊平均或累積步驟 405a 之量來計算每一子帶之每訊框 bin 量的和(對時間之平均/累積)。這些和被用以計算下面步驟 410 之聲道間角度一致性因數。

【00104】 c. 若編碼器之耦合頻率低於約 1000Hz，施用于子帶訊框平均後或累積後之量至一時間平滑器，其對低於此頻率且高於該耦合頻率之所有子帶操作。

【00105】 有關步驟 405c 之註解：

【00106】 見有關步驟 404c 之註解，除了步驟 405c 之情形外，該時間平滑可替選地被實施作爲步驟 410 之一部分。

【00107】 步驟 406 計算相對聲道間 bin 相位角度

【00108】 藉由將步驟 403 之 bin 角度減掉參考聲道(例如爲第一聲道)之對應的 bin 角度計算每一區塊之每一變換 bin 的相對聲道間 bin 相位角

度。其結果(如此處之其他角度加法或減法)藉由加或減 2 釘直至其結果落在所欲的-釘至+釘的範圍內為止(即 modulo (釘, -釘)運算)。

【00109】 步驟 407 計算聲道間子帶相位角度

【00110】 為每一聲道如下列地計算一訊框率振幅加權平均之聲道間相位角度：

【00111】 a. 為每一 bin，由步驟 403 之量與步驟 406 之相對子帶間 bin 相位角度構建一複數。

【00112】 b. 對整個每一子帶將步驟 407a 所構建之複數相加(對整個頻率相加)。

【00113】 有關步驟 407b 之註解：

【00114】 例如，若一子帶具有二 bin 且該等 bin 之一具有 $1+1j$ 之複數值及另一具有 $2+2j$ 之複數值，其複數和為 $3+3j$ 。

【00115】 c. 對每一訊框之整個區塊為步驟 407b 之每一子帶平均或累積每一區塊複數和(對整個時間平均或累積)。

【00116】 d. 若該編碼器之耦合頻率低於約 1000Hz，施用該子帶訊框平均或累積後之複數值至一時間平滑器，其對低於此頻率且高於該耦合頻率之所有子帶操作。

【00117】 有關步驟 407d 之註解：

【00118】 見有關步驟 404c 之註解，除了步驟 407d 之情形外，該時間平滑可替選地被實施為步驟 407c 或 410 之一部分。

【00119】 e. 如每一步驟 403 地計算步驟 407d 之複數結果的量。

【00120】 有關步驟 407e 之註解：

【00121】 此量在下面的步驟 410a 被使用。在步驟 407b 所給予之簡單例中， $3+3j$ 之量被作 $\text{square_root}(9+9)=4.24$ 。

【00122】 f. 計算步驟 403 之複數結果的角度。

【00123】 有關步驟 407f 之註解：

【00124】 在步驟 407b 所給予之簡單例中， $3+3j$ 之角度為 $\arctan(3/3)=45$ 度=釘/4。此子帶角度被信號相依式地求時間平滑(見步驟 413)及被數量化(見步驟 414)以如下列般地產生子帶角控制參數支鏈資訊。

【00125】 步驟 408 計算 bin 頻譜穩定度因數

【00126】 就每一 bin 而言，計算 0 至 1 範圍之一 bin 頻譜穩定度因數如下：

【00127】 a. 令 x_m = 在步驟 403 所計算之目前區塊的 bin 量。

【00128】 b. 令 y_m = 在前一個區塊的對應之 bin 量。

【00129】 c. 若 $x_m > y_m$ 則 bin 動態振幅因數 $= (y_m/x_m)^2$ ，

【00130】 d. 否則，若 $y_m < x_m$ ，bin 動態振幅因數 $= (x_m/y_m)^2$ ，

【00131】 e. 否則，若 $y_m = x_m$ ，則 bin 振幅因數 = 1。

【00132】 有關步驟 408 之註解：

【00133】 「頻譜穩定度」為頻譜成份(如頻譜係數或 bin 值)隨時間變化程度之量度。bin 動態振幅因數為 1 表示在某一特定期間不隨時間變化。

【00134】 替選的是，步驟 408 可查對三個連續區塊。若編碼器之該耦合頻率低於約 1000Hz，步驟 408 可查對多於三個連續區塊。連續區塊之數目可考慮頻率而變化，使得該數目隨著子帶頻率範圍減小而逐漸增加。

【00135】 作為進一步替選做法，bin 能量可取代 bin 量被使用。

【00136】 而作為再進一步替選做法，步驟 408 可運用如下列步驟 409 後之註解所描述的一「事件決策」偵測技術。

【00137】 步驟 409 計算子帶頻譜穩定度因數

【00138】 藉由如下列地對整個各區塊的每一子帶形成 bin 頻譜穩定度因數的一振幅加權平均數計算尺度 0 至 1 之一訊框率子帶頻譜穩定度因數：

【00139】 a. 就每一 bin，計算步驟 408 之 bin 頻譜穩定度因數與步驟 403 之 bin 量的乘積。

【00140】 b. 將每一子帶之乘積相加(對整個頻率之相加)。

【00141】 c. 將一訊框內所有區塊中步驟 409b 之和平均或累積(對整個時間之平均/累積)。

【00142】 d. 若該編碼器之耦合頻率低於約 1000Hz，施用該子帶訊框平均或累積後之和至一時間平滑器，其對低於此頻率且高於該耦合頻率之所有子帶操作。

【00143】 e. 將步驟 409c 或步驟 409d 之結果適當地除以子帶內之

bin 量(步驟 403)

【00144】 有關步驟 409e 之註解：

【00145】 步驟 409a 之量相乘與步驟 409e 之量相加提供振幅加權。步驟 408 之輸出與絕對振幅無關，且若未被振幅加權可能致使步驟 409 之輸出被很小的振幅控制，此為非欲的。

【00146】 f. 藉由將該範圍由 $\{0.5...1\}$ 映射至 $\{0...1\}$ 而把結果比例調整以獲得頻譜穩定度因數。此可利用將結果乘以 2 減 1，並將小於 0 的值限制為 0 而被做成。

【00147】 有關步驟 409f 之註解：

【00148】 步驟 409f 在確保因子帶頻譜穩定度因數為 0 的聲道雜訊為有用的。

【00149】 有關步驟 408 與 409 之註解：

【00150】 步驟 408 與 409 之目標為要測量頻譜穩定度—在一聲道中一子帶的頻譜成份隨時間之改變。替選的是，如國際專利公報中 WO 02/097792 A1 號(指定給美國)所描述之「事件決策」感應層面可被運用以測量頻譜穩定度而取代剛剛相關步驟 408 與 409 所描述之做法。2003 年 11 月 20 日美國專利 S.N. 10/478,538 號即為 PCT 公報 WO 02/097792 A1。該等 PCT 公報與美國專利整體被納於此做為參考。依據這些被納入之參考案，每一 bin 之複數 FFT 的量被計算及被常規化(例如最大之量被設定為 1)。然後在連續區塊中對應的 bin 之量(以 dB 表示)被減除(忽略其正負號)、bin 間之差被相加、且該和若超過一臨界值，該區塊界限被視為一音響事件的界限。替選的是，由區塊至區塊的振幅變化亦可與頻譜量變化被考慮(利用注意所需要的常規化之量)。

【00151】 若所納入之事件感應應用的層面被運用以測量頻譜穩定度，常規化可不需要且頻譜量變化(若常規化被省略，量之變化不會被測量)較佳地以一子帶基準被考慮。取代上述之步驟 408 的是，每一子帶之對應 bin 間的頻譜量之分貝差可依據該等應用之教習被加總。然後代表由區塊至區塊之頻譜變化程度的每一這些和可被比例調整，使得其結果為頻譜穩定度因素為 0 至 1，其中 1 表示最高穩定度，即就某一特定 bin，由區塊至區

塊的變化為 0 dB。0 之值表示最低穩定度，可被指定為大於或等於例如為 12 dB 之一適當值。這些結果之一 bin 頻譜穩定度因數可以與步驟 409 運用剛剛所描述之事件決策技術所獲得之一 bin 頻譜穩定度因數時，步驟 409 之變換一 bin 頻譜穩定度因數可被使用做為一暫態之指標。例如，若步驟 409 所產生之值的範圍為 0 至 1，當其子帶頻譜穩定度因數為一小值時(如 0.1)，一暫態可被視為是出現的，表示實質上的頻譜不穩定。

【00152】其將被了解用步驟 408 與用剛所描述之步驟 408 的替選做法所產生之 bin 頻譜穩定因數每一均一致性地提供一某一程度為可變的臨界值，其係根據由區塊至區塊之相對變化而定。備選的是，利用特別提供該臨界值移位響應例如一訊框之多暫態或數個較小暫態中之一個大暫態(如來自高於中度至低度位準掌聲之大聲的暫態)來補充此一致性為有用的。在後者之情形中，一事件偵測器可起始地辨識每一掌聲為一事件，但一大聲的暫態(如鼓聲)使其欲將該門檻值移位，使得僅有該鼓聲被辨識為一事件。

【00153】替選的是，一隨機度量尺可被運用(例如，美國專利 Re 36,714 所描述者，其整體被納入此處做為參考)，而取代對時間所量測之頻譜穩定度。

【00154】步驟 410 計算聲道間角度一致性因數

【00155】a. 將步驟 407e 之複數和的量除以步驟 405 之量的和。結果之「原始」角度一致性因數為範圍 0 至 1 之數值。

【00156】b. 計算一校正因素：令 n = 對上述二步驟之二數量的歸因之子帶的整個數值(換言之， n 為該子帶中 bin 的數目)。若 n 小於 2，該角度一致性因數為 1 並前進至步驟 411 與 413。

【00157】c. 令 r = 期望隨機變異數 = $1/n$ ，將 r 由步驟 410b 之結果減除。

【00158】d. 將步驟 410c 之結果除以 $(1-r)$ 而常規化。其結果之最大值為 1，將其最小值如所需地限制為 0。

【00159】有關步驟 410 之註解：

【00160】聲道間角度一致性因數為一子帶內之聲道相位角在一訊框期間有多類似之一量度。若子帶之所有 bin 聲道角皆相同，該子帶間角度一

致性因數為 1.0；而若該等聲道間角為隨機散佈，該值趨近於 0。

【00161】該子帶間角度一致性因數表示聲道間是否有虛幻影像。若該一致性為低的，則欲將該等聲道解除相關。一高值表示融合的影響。影像融合係與其他信號特徵獨立無關。

【00162】其將被注意到，子帶間角度一致性因數雖然為一角度參數，其係由二量間接地被決定。若聲道間角均相同，複數值相加再取得其量與取得其量再相加之結果相同，故其商為 1。其聲道間角為散佈的，則複數值相加(即具有不同角度之向量相加)會有至少部份相抵消之結果，故和之量小於 1，且其商小於 1。

【00163】下列為具有二 bin 之子帶的簡單例子：

【00164】假設二複數 bin 值為 $3+4j$ 與 $6+8j$ (二者之角度相同：角度 = $\arctan(\text{虛數}/\text{實數})$)，故角度 $1 = \arctan(4/3)$ 及角度 $2 = \arctan(8/6) = \arctan(4/3)$ 。複數值相加，和 = $9+12j$ ，其量 $\text{square_root}(81+144)=15$ 。

【00165】而量之和為 $(3+4j)$ 之量 + $(6+8j)$ 之量 = $5+10=15$ 。其商因此為 $15/15=1$ (在 $1/n$ 常規化前，在常規化後亦為 1)(常規化後之一致性 = $(1-0.5)/(1-0.5)=1.0$)。

【00166】若上面 bin 之一具有不同之角度，如第二個之複數值為 $6-8j$ ，其具有相同之量，15。其複數和現在為 $9-4j$ ，具有之量為 $\text{square_root}(81+16)=9.85$ ，故其一致性(常規化前)商 = $9.85/15=0.66$ 。為常規化，減掉 $1/n=1/2$ 並除以 $1-1/n$ (常規化後之一致性 = $(0.66-0.5)/(1-0.5)=0.32$)。

【00167】雖然上述用於決定子帶角度一致性因數已被發現為有用的，但其並非關鍵的。其他合適的技術可被運用。例如，吾人可使用標準公式來計算標準差。在任何情形其均欲運用振幅加權以使小信號對所計算之一致性值的影響最小化。

【00168】此外，子帶角度一致性因數之替選的導出作法可使用能量(該等量之平方)取代量。此可藉由將步驟 403 之量在其被施用至步驟 405 與 407 前將之平方而完成。

【00169】步驟 411 導出子帶解除相關標度因數

【00170】為每一子帶導出一訊框率解除相關標度因數如下：

【00171】 a. 令 x =步驟 409f 之訊框率頻譜穩定度因素。

【00172】 b. 令 y =步驟 410e 之訊框率角度一致性因數。

【00173】 c. 則該訊框率子帶解除相關標度因數 $= (1-x) * (1-y)$ ，介於 0 與 1 間之數。

【00174】 有關步驟 411 之註解：

【00175】 該子帶解除相關標度因數為一聲道之一子帶中時間上的信號特徵(頻譜穩定度因數)與一聲道 bin 角度同一子帶針對一參考聲道之對應的 bin 之一致性(聲道間角度一致性因數)的函數。該子帶解除相關標度因數只有在該頻譜穩定度因數與該聲道間角度一致性因數二者均低時為高的。

【00176】 如上面解釋者，該解除相關標度因數控制在編碼器中被角度一致性因數之包線解除相關的程度。對時間展現頻譜穩定度因數的信號較佳地不利用變更其包線而被解除相關(不管在其他聲道發生什麼)，因其會產生可聽到的人工物之結果，即信號之波段或顫音。

【00177】 步驟 412 導出子帶振幅標度因數

【00178】 由步驟 404 之子帶訊框能量及由所有其他聲道之子帶訊框能量值(如可由對應於步驟 404 或其等值步驟可得到者)。導出訊框率子帶振幅標度因數如下：

【00179】 a. 就每一子帶，對整個所有輸入聲道之每一訊框加總其能量值。

【00180】 b. 每一訊框將每一子帶能量(來自步驟 404)除以整個所有輸入聲道之能量值(來自步驟 412a)以創立範圍 0 至 1 的值。

【00181】 c. 在- 至 0 之範圍內變換每一比值為 dB。

【00182】 d. 除以標度因數顆粒度(其例如可被設定為 1.5dB)、改變符號以得到非負值、限制為一最大值(例如 31，即 5 位元之精準度)、及取最近之整數以創立數量化的值。這些值為訊框子帶標度因數且被輸送作為該支鏈資訊之一部份。

【00183】 e. 若該編碼器之耦合頻率低於約 1000Hz，施用該子帶訊框平均或累積後之和至一時間平滑器，其對低於此頻率且高於該耦合頻率之所有子帶操作。

【00184】 有關步驟 412e 之註解：

【00185】 見有關步驟 404c 之註解，除了步驟 412e 之情形外，其無該時間平滑可替選地被實施之適合的後續步驟。

【00186】 有關步驟 412 之註解：

【00187】 雖然此處所指出之顆粒度(解析度)與數量化精確度被發現為有用的，其並非關鍵的，且其他的值可提供可接受之結果。

【00188】 替選的是，吾人可使用振幅取代能量以產生該等振幅標度因數。若使用振幅，吾人會使用 $\text{dB}=20 * \log(\text{振幅比})$ ；而若使能量，吾人經由 $\text{dB}=10 * \log(\text{能量比})$ 將之變換為 dB，此處振幅比= $\text{square_root}(\text{能量比})$ 。

【00189】 步驟 413 信號相依之時間平滑聲道間的子帶相位角度

【00190】 施用信號相依之時間平滑至訊框率聲道間角度(在步驟 407f 被導出)：

【00191】 a. 令 v =步驟 409d 之子帶頻譜穩定度因數。

【00192】 b. 令 w =對應的步驟 410e 之頻譜穩定度因數。

【00193】 c. 令 $x=(1-v)*w$ ，此為介於 0 與 1 間之值，若頻譜穩定度因數為低且角度一致性因數為高的，其為高的。

【00194】 d. 令 $y=1-x$ ，若頻譜穩定度因數為高且角度一致性因數為低的， y 為高的。

【00195】 e. 令 $z=y\text{exp}$ ，此處 exp 為一常數(可為=0.1)， z 亦在 0 至 1 的範圍內，但向 1 偏斜，對應於一緩慢的時間常數。

【00196】 f. 若聲道之暫態旗標(步驟 401)被設定，設定 $z=0$ ，對應於在暫態出現之一快速的時間常數。

【00197】 g. 計算 $\text{lim}=(0.1 * w)$ ，此為 z 之最大可允許的值，此範圍為 0.9(若角度一致性因數為高的)至 1.0(若角度一致性因數為低的(0))。

【00198】 h. 如所需地用 lim 限制 z ：若 $z > m$ 則 $z=\text{lim}$ 。

【00199】 i. 用 z 之值與為每一子帶所維持之角度的一進行中之平滑值來平滑步驟 407f 之子帶角度。若 A =步驟 407f 之角度及 RSA =前一區塊之進行中的平滑後角度，與 NewRSA 為進行中的平滑後角度的新值，則 $\text{NewRSA}=\text{RSA} * z + A * (1-z)$ 。 RSA 之值在處理隨後之區塊前因之被設定等

於 NewRSA。NewRSA 為步驟 413 之信號相依的時間平滑後的角度輸出。

【00200】有關步驟 413 之註解：

【00201】當一暫態被偵測，子帶角度更新時間常數被設定為 0，允許快速的子帶角度變化。此為所欲的，原因在於其允許正常的角度更新機制使用一範圍之相當緩慢的時間常數，使靜態或等靜態信號之際的影像漂動最小化，而快速變化之信號以快速時間常數被處理。

【00202】雖然其他的平滑技術與參數為可使用的，施作步驟 413 之一第一階平滑器已被發現為有用的。若被施作為一第一階平滑器/低通濾波器，該變數 z 對應於前途係數(有時記為 $ff0$)，而 $1-z$ 對應於回授係數(有時記為 $fb1$)。

【00203】步驟 414 數量化平滑聲道間子帶相位角度

【00204】將步驟 413i 中導出之平滑聲道間子帶相位角度數量化以獲得角控制參數：

【00205】a. 若該值小於 0，加上 2 釘，使得將被數量化之所有角度值為 0 至 2 釘之範圍內。

【00206】b. 除以角度顆粒度(解析度，其可為 2 釘/64 徑度值)並取其整數。其最大值可在 63 被設定，對應於 6 位元之數量化。

【00207】有關步驟 414 之註解：

【00208】該數量化後之值被視為非負之整數，故將該角度數量化之一簡易的方法被映射至非負之浮點數字(若小於 0 則加上 2 釘，使其範圍為 0 至 2 釘)、用顆粒度(解析度)調整，並取整數值。類似地，將該整數解除數量化(其或可簡單查表被完成)可藉由利用該角度顆粒度因數之倒數調整、變換非負整數為非負浮點角度(再次地以 0 至 2 釘為範圍)被完成，此後其可再被常規化為範圍±釘以便進一步使用。雖然該子帶角控制參數之此數量化已被發現為有用的，此數量化為非關鍵的且其他的數量化可提供可接受之結果。

【00209】步驟 415 子帶解除相關支鏈之數量化

【00210】藉由乘以 7.49 並取其最近的整數而將步驟 411 之子帶解除相關支鏈數量化為例如 8 等級(3 位元)。這些數量化後之值為部分的支鏈資訊。

【00211】 有關步驟 415 之註解：

【00212】 雖然該子帶角控制參數之此數量化已被發現為有用的，此數量化為非關鍵的且其他的數量化可提供可接受之結果。

【00213】 步驟 416 子帶角控制參數解除數量化

【00214】 將子帶角控制參數數量化(見步驟 414)以在向下混頻前使用。

【00215】 有關步驟 416 之註解：

【00216】 使用編碼器中之數量化後的值有助於維持編碼器與解碼器間之同步化。

【00217】 步驟 417 在整個區塊分散訊框解除數量化後之角控制參數

【00218】 爲了準備向下混頻，將步驟 416 之整個時間每一訊框解除數量化一次的角控制參數分散至訊框內每一區塊之子帶。

【00219】 有關步驟 417 之註解：

【00220】 同一訊框值可被指定給訊框中之每一區塊。替選的是，在一訊框中整個所有區塊內插子帶角控制參數可爲有用的。對時間之線性內插可以如下面描述之對頻率線性內插的方式被運用。

【00221】 步驟 418 內插區塊子帶角控制參數至 bin

【00222】 對整個頻率爲每一聲道分散該等區塊子帶角控制參數至 bin，較佳地爲使用下面描述之線性內插。

【00223】 有關步驟 418 之註解：

【00224】 若對頻率之線性內插被運用。步驟 418 使通過一子帶界限由 bin 至 bin 之相位角度變化最小化而使混疊的人工物最小化。子帶角度係彼此獨立地被計算，每一個代表對整個子帶之平均。因而，由一子帶至下一個可能有大變化。若一子帶之淨角度值被施用至該子帶之所有 bin(一種「長方形」子帶分配)，由一子帶至鄰近子帶之整個相位變化在二 bin 間發生。若其有強的信號成份於此，其可能有嚴重的可能可聽到的混疊。線性內插在子帶中之所有 bin 散佈相位角度變化，使任一對 bin 間的變化爲最小，例如使得在一子帶低端部的角度與該子帶高端部的角度偶配，而又維持整體

平均數與某一特定被計算之子帶角度相同。換言之，取代長方形之子帶分配的是該子帶角度分配可為梯形。

【00225】 例如，最低被耦合之子帶具有一 bin 及 20 度之子帶角，下一個子帶具有三 bin 及 40 度之子帶角，及第三個子帶具有五 bin 及 100 度之子帶角。在沒有內插下，假設該第一個 bin(一子帶)以 20 度被移位、下三個 bin(另一子帶)以 40 度被移位、下五個 bin(再一子帶)以 100 度被移位。在此例中由 bin 4 至 bin 5 有 60 度之最大變化。在有線性內插下，該第一 bin 仍被移位 20 度；下三個 bin 被移位約 30，40 與 50 度；及接著五個 bin 被移位約 67，83，100，117 與 133 度。平均子帶角度移位相同，但最大的 bin 對 bin 變化被降低為 17 度。

【00226】 備選的是，由子帶至子帶之子帶變化配合如步驟 417 之此處所描述的此與其他步驟亦可以類似的內插方式被處理。然而，在由一子帶至下一個子帶之振幅傾向於更自然之連續性，其可能不必要如此做。

【00227】 步驟 419 為聲道施用角旋轉為 bin 變換值

【00228】 如下列般地对 bin 變換值施用相位角旋轉：

【00229】 a. 令 x =如步驟 418 所計算之此 bin 的 bin 角度。

【00230】 b. 令 $y=-x$ ；

【00231】 c. 以角度 y 計算 z ，即一單位量複數相位旋轉標度因數， $z=\cos y+ \sin yj$ 。

【00232】 d. 將 bin 值($a+bj$)乘以 z 。

【00233】 有關步驟 419 之註解：

【00234】 被施用至該編碼器之相位角旋轉為由子帶角控制參數被導出之角度的倒數。

【00235】 在向下混頻(步驟 420)前於一編碼器或編碼處理中如此處所描述之相位角度調整具有數個好處：(1)其使被加為單聲道合成信號或被矩陣化為多聲道的聲道之抵銷為最小，(2)其使對能量常規化(步驟 421)之依賴為最小，及(3)其預先補償解碼器反相位角旋轉而減少混疊。

【00236】 該等相位校正因數可藉由由該子帶之每一變換 bin 值的角度減除每一子帶相位校正值而將編碼器移位。此係等值於將每一複數 bin 值乘

以量為 1.0 之複數與等於該相位校正值之負數的一角度。注意，就量為 1 之複數而言，角度 A 等於 $\cos A + j \sin A$ 。後者之數量以 A=此子帶之負相位校正為每一聲道之每一子帶被計算一次，然後乘以每一 bin 信號值以實現相位被移位之 bin 值。

【00237】該相位移位為圓圈形，造成圓形迴旋(如上述者)。雖然圓形迴旋就一些連續信號可為溫和的，其可能某些連續的複數信號(如高音管)創造激烈的頻譜成份，或不同的相位角度就不同的子帶被使用可能造成暫態之模糊。後果為，避免圓形迴旋之適合的技術可被運用，或暫態旗標可被運用，使得例如當暫態旗標為真，該角度計算結果可被蓋掉，且一聲道中之所有子帶可使用如 0 或隨機化之值的同一相位校正因數。

【00238】步驟 420 向下混頻

【00239】藉由將整個聲道的對應之複數變換 bin 相加而向下混頻為單聲道或以如下面描述之第 6 圖例子的方式藉由將輸入聲道作成矩陣而向下混頻為多聲道。

【00240】有關步驟 420 之註解：

【00241】在編碼器中，一旦所有聲道之變換 bin 已被相位移位，該等聲道被逐一 bin 地相加以創造單聲道合成音訊信號。替選的是，該等聲道可被施用至一被動或主動矩陣，其提供簡單相加為一聲道(如第 1 圖之 N:1 編碼)或成為多聲道。該等矩陣係數可為實數或複數(實數與虛數)。

【00242】步驟 421 常規化

【00243】為避免隔離的 bin 之抵消及過度強調同相位信號，如下列般地單聲道合成之每一 bin 的振幅常規化以具有實質上該等歸因能量之和相等的能量：

【00244】a. 令 $x = \text{bin 能量所有聲道之和(即步驟 403 所計算之 bin 量的平方)}$ 。

【00245】b. 令 $y = \text{單聲道合成之對應的 bin 之能量(如步驟 403 所計算者)}$ 。

【00246】c. 令 $z = \text{標度因數} = \text{square_root}(x/y)$ ，若 $x=0$ 則 $y=0$ ，且 z 被設定為 1。

【00247】 d. 限制 z 為例如 100 之最大值。若 z 起始地大於 100(意即來自向下混頻之強烈的抵消)，將例如為 $0.01 * \text{square_root}(x)$ 之任意值加至該單聲道合成 bin 之實數部與虛數部，此將確保其夠大以用下列步驟被常規化。

【00248】 e. 用 z 乘以該複數單聲道合成 bin 值。

【00249】 有關步驟 421 之註解：

【00250】 雖然一般係欲就編碼與解碼使用相同的相位因數，甚至一子帶相位校正值之最適選擇會造成該子帶內一個或更多可聽的頻譜成份在編碼向下混頻過程之際，因步驟 419 之相位移係以子帶而非 bin 基準被實施而被抵消。在此情形中，編碼器中隔離的 bin 之一不同的相位因數可其若被偵測到這些 bin 之能量和小於此頻率之各別聲道 bin 的能量和很多時可被使用。一般而言，其沒必要施用被隔離之一校正因素至該解碼器，因此被隔離之 bin 對整體影像品質之影響通常為很小。若多聲道而非單聲道被運用，類似的常規化可被施用。

【00251】 步驟 422 組合及封包為位元流

【00252】 每一聲道之振幅標度因數、角控制參數、解除相關標度因數與暫態旗標的支鏈資訊以及普通的單聲道合成音訊或矩陣多聲道如可能所欲地被多工及被封包為適用於該等儲存、傳輸、或儲存且傳輸媒體之一個或更多的位元流。

【00253】 有關步驟 422 之註解：

【00254】 該等單聲道合成音訊或多聲道音訊可在封包前被施用至一資料率編碼功能與裝置，例如為一可感覺的編碼器或至一可感覺的編碼器與一熵編碼器(如算術或赫夫曼編碼器)(有時被稱為「無損失」編碼器)。同時如上述者，單聲道合成音訊(或多聲道音訊)與相關的支鏈資訊可僅就高於某種頻率(一「耦合」頻率)之音訊頻率由多輸入聲道被導出。在此情形中，在每一該等多輸入聲道中低於該耦合頻率之音訊頻率可被儲存、傳輸、或儲存且傳輸為離散的聲道，或以非此處所描述之一些方式被組合或被處理。離散或否則被組合之聲道亦被施用至一資料率編碼功能與裝置，例如為一可感覺的編碼器或至一可感覺的編碼器與一熵編碼器。該等單聲道合

成音訊(或多聲道音訊)與離散的多聲道音訊全部可在封包前被施用至一整合的感覺編碼或感覺與熵編碼功能與裝置。

【00255】 解碼

【00256】 解碼處理之步驟(「解碼步驟」)可如下列般地被描述。針對解碼步驟係參照一混合式流程圖與功能方塊圖性質之第 5 圖。為簡單起見該圖係顯示為一聲道之支鏈資訊成份的導出，其被了解該等支鏈資訊成份必須就每一聲道被獲得，除非該聲道為如別處被解釋之此類成份的一參考聲道。

【00257】 步驟 501 將支鏈資訊解除封包及解碼

【00258】 為每一聲道(在第 5 圖中被顯示之一聲道)之每一訊框如所需地將支鏈資料成份(振幅標度因數、角控制參數、解除相關標度因數與暫態旗標)解除封包及解碼。查表可被用以將振幅標度因數、角控制參數與解除相關標度因數解碼。

【00259】 有關步驟 501 之註解：

【00260】 如上面解釋者，若一參考聲道被運用，該參考聲道之支鏈資料不包括角控制參數與解除相關標度因數。

【00261】 步驟 502 將單聲道合成或多聲道音訊信號解除封包及解碼

【00262】 為單聲道合成或多聲道音訊信號之每一變換 bin 如所需地將單聲道合成或多聲道音訊信號解除封包及解碼以提供 DFT 係數。

【00263】 有關步驟 502 之註解：

【00264】 步驟 501 與 502 可被視為部分之單一解除封包及解碼步驟。步驟 502 可包括一被動或主動矩陣。

【00265】 步驟 503 對整個所有區塊分散角控制參數

【00266】 區塊子帶角控制參數值由解除數量化後之訊框子帶角控制參數值被導出。

【00267】 有關步驟 503 之註解：

【00268】 步驟 503 可藉由分散同一參數值至訊框中每一區塊而被施作。

【00269】 步驟 504 對整個所有區塊分散子帶解除相關標度因數

【00270】 區塊子帶解除相關標度因數值由解除數量化後之訊框子帶解除相關標度因數值被導出。

【00271】 有關步驟 504 之註解：

【00272】 步驟 504 可藉由分散同一標度因數值至訊框中每一區塊而被施作。

【00273】 步驟 505 加入隨機化相位角度偏差(技術 3)

【00274】 依照上述之技術 3，當暫態旗標表示有暫態時，將步驟 503 所提供之區塊子帶角控制參數加入解除相關標度因數所調整之一隨機化偏差值(此調整可在此步驟中間接地被設立)。

【00275】 a. 令 y =區塊子帶解除相關標度因數。

【00276】 b. 令 $z=y\exp$ ，其中 \exp 為例如 5 之常數， z 亦將為在 0 至 1 之範圍，但向 0 偏斜，除非該解除相關標度因數值為高的，否則反映隨機化變異數朝向低水準之偏差。

【00277】 c. 令 x =介於+1 與-1 間之一隨機化數字，為每一區塊之每一子帶分離地被選擇。

【00278】 d. 然後被加到該區塊子帶角控制參數以依據技術 3 加入隨機化角度偏差值之值為 $x * \pi * z$ 。

【00279】 有關步驟 505 之註解：

【00280】 如一般熟習本技藝者將了解者，用於被解除相關標度因數調整之「隨機化」角度(或，若振幅亦被調整，則為隨機化振幅)可不僅包括虛擬隨機或真實隨機之變異數，亦包括確定被產生之變異數，其在被施用至相位角度或至相位角度與至振幅時，具有降低聲道間交叉相關之效果。此類「隨機化」變異數可用很多方法被獲得。例如，具有各式種子值之虛擬隨機數產生器可被運用。替選的是，真實隨機數可使用硬體隨機數產生器被產生。因此，僅約 1 度之一隨機化角度解析度將為足夠的，具有二或三位小數點(如 0.84 或 0.844)之隨機化數字表可被運用。

【00281】 雖然步驟 505 之非線性間接調整已被發現為有用的，但其為非關鍵的，其他適合的調整可被運用—特別是就指數而言之其他值可被運用

以獲得類似之結果。

【00282】 當子帶解除相關標度因數值為 1，由-釘至+釘全範圍的角度被加入(在此情形中步驟 503 所產生之區塊子帶角控制參數值被不相關地提供)。隨著子帶解除相關標度因數朝 0 減小，該隨機化角度偏差亦朝 0 減小，致使步驟 505 之輸出朝步驟 503 所產生之子帶角控制參數值移動。

【00283】 若所欲時，上述的編碼器在向下混頻前依照技術 3 亦加入一調整後之隨機化偏差到被施用至一聲道的角度移位。如此做可改善解碼器中之混疊抵消。其亦可有益於改善編碼器與解碼器之同步性。

【00284】 步驟 506 對整個頻率線性內插

【00285】 由解碼器步驟 503 之區塊子帶角度導出 bin 角度，對此隨機化偏差在暫態旗標表示一暫態時已被步驟 505 加入。

【00286】 有關步驟 506 之註解：

【00287】 bin 角度可由子帶角度用如上述有關步驟 418 所描述的對整個頻率之線性內插被導出。

【00288】 步驟 507 加入隨機化相位角度偏差(技術 2)

【00289】 依照上述之技術 2，當暫態旗標未表示有暫態時為每一 bin 對步驟 503 所提供之一訊框中的所有區塊子帶角控制參數(步驟 505 只在暫態旗標表示有暫態時操作)加入該解除相關標度因數所調整之不同的隨機化偏差值(該調整可在此步驟於此直接被設立)：

【00290】 a. 令 y =區塊子帶解除相關標度因數。

【00291】 b. 令 x =介於+1 與-1 間之一隨機化數字，為每一訊框之每一 bin 分別被選擇。

【00292】 c. 然後被加到該區塊子帶角控制參數以依據技術 3 加入隨機化角度偏差值之值為 $x * \pi * z$ 。

【00293】 有關步驟 507 之註解：

【00294】 見對隨機化角度偏差之有關步驟 505 之註解。

【00295】 雖然步驟 507 之直接調整已被發現為有用的，但其為非關鍵的，其他適合的調整可被運用。

【00296】 為使時間不連續性最小化，為每一声道之每一 bin 的獨一之

隨機化角度值較佳地不隨時間變化。所有 bin 之隨機化角度值用以訊框率被更新之同一子帶解除相關標度因數被調整。因而，當子帶解除相關標度因數值為 1，由-釘至+釘之全範圍的隨機角度被加入(在此情形中，由解除數量化之訊框子帶角度值被導出的區塊子帶角度值不相關地被提供)。隨著子帶解除相關標度因數值朝 0 消失，該隨機化角度值亦朝 0 消失。不像步驟 504 者，此步驟 507 之調整可為子帶解除相關標度因數值之直接函數。例如，0.5 之子帶解除相關標度因數以 0.5 成比例地降低每一隨機角度變異數。

【00297】 然後調整後之隨機化角度值由解碼器步驟 506 被加入 bin 角度。解除相關標度因數值以每一訊框被更新一次。在該訊框之暫態旗標出現中此步驟被跳越以避免暫態的前置雜訊人工物。

【00298】 若所欲時，上述的編碼器在向下混頻前依照技術 3 亦加入一調整後之隨機化偏差到被施用至一聲道的角度移位。如此做可改善解碼器中之混疊抵消。其亦可有益於改善編碼器與解碼器之同步性。

【00299】 步驟 508 常規化振幅標度因數

【00300】 對整個常規化振幅標度因數，使得其平方和為 1。

【00301】 有關步驟 508 之註解：

【00302】 例如，若二聲道具有之解除數量化標度因數為-3.0dB ($=2 * 1.5\text{dB}$ 之顆粒度)(0.70795)，該平方和為 1.002。將其每一個除以 1.002 之平方根 1.001，得到二個 0.7072(-3.01dB)之二值。

【00303】 步驟 509 昇高步驟標度因數水準(備選的)

【00304】 備選地，當暫態旗標表示無暫態時，依子帶解除相關標度因數水準施用稍微的昇高至子帶標度因數水準：以小的因數乘以每一常規化後之子帶振幅標度因數(如 $1+0.2 * \text{子帶解除相關標度因數}$)。當暫態旗標為真，跳越此步驟。

【00305】 有關步驟 509 之註解：

【00306】 由於解碼器解除相關步驟 507 可形成最後逆濾波器排組處理之稍微降低的水準結果，此步驟可為有用的。

【00307】 步驟 510 對整個 bin 分散子帶振幅值

【00308】 步驟 510 可藉由分散同一子帶振幅標度因數值至該子帶之

每一 bin 而被施作。

【00309】 步驟 510a 加入隨機化振幅偏差(備選的)

【00310】 備選地，依子帶解除相關標度因數水準與暫態旗標施用一隨機化變異數至隨機化子帶振幅標度因數。在暫態不出現時以逐一 bin 基準(隨 bin 不同)地加入不隨時間變化之一隨機化振幅標度因數，及在暫態出現(在訊框或區塊中)時，加入以逐一區塊基準(隨區塊不同)變化及隨子帶變化(對一子帶所有 bin 為同一移位；隨子帶不同)之一隨機化振幅標度因數。步驟 510a 在圖中未被畫出。

【00311】 有關步驟 510a 之註解：

【00312】 雖然隨機化振幅移位被加入之程度可用解除相關標度因數被控制，咸信一特定標度因數值應該會比由相同標度因數值結果所得的對應之隨機化相位移位造成較小的振幅移位以避免可聽到的人工物。

【00313】 步驟 511 向上混頻

【00314】 a. 就每一輸出聲道之每一 bin，由解碼器步驟 508 之振幅與解碼器步驟 507 之 bin 角度構建一複數向上混頻標度因數。

【00315】 b. 就每一輸出聲道，將複數 bin 值乘以複數向上混頻標度因數以產生該聲道之每一 bin 的向上混頻後之複數輸出 bin 值。

【00316】 步驟 512 實施逆 DFY(備選的)

【00317】 備選地，對每一輸出聲道之 bin 實施逆 DFT 變換以得到多聲道輸出 PCM 值。如相當習知者，配合此逆 DFT 變換，時間樣本之各別區塊被作成視窗，且相鄰區塊被相疊及被加在一起以重新構建最終連續的時間輸出 PCM 音訊信號。

【00318】 有關步驟 512 之註解：

【00319】 依據本發明之解碼器不會提供 PCM 輸出。在解碼器處理僅在高於某一特定頻率被運用及離散的 MDCT 係數就低於此頻率之每一聲道被傳送的情形中，其可能欲變換該解碼器向上混頻步驟 511a 與 511b 導出之 DFT 係數為 MDCT 係數，使得其與較低頻率之離散 MDCT 係數可被組合及重新被數量化，以提供例如與如一標準 AC-3 SP/DIF 位元流之具有大量被安裝使用者之編碼系統相容的位元流，用於施用至逆變換可被實施之一

外部裝置。逆 DFT 變換可被施用至輸出聲道之一以提供 PCM 輸出。

【00320】 A/52A 文件之 8.2.2 節

【00321】 以敏感度因數“F”被

【00322】 加入之 8.2.2 暫態偵測

【00323】 暫態在全帶寬聲道被偵測以決定何時要切換至短長度音訊區塊以改善前置回聲績效。該等信號之高通濾波後的版本就由一子區塊時間段至下一個之能量提高被檢查。子區塊在不同的時間標度被檢查。若一暫態在聲道之一音訊區塊的第二半部被偵測，此聲道切換為短區塊。被區塊切換之一聲道係使用 D45 指數策略[即其資料具有較粗的頻率解析度以降低時間解析度增加所致之資料費用]。

【00324】 該暫態偵測器被用以決定何時要由長變換區塊(長度 512)變換為短區塊(長度 256)。其對每一音訊區塊之 512 樣本操作。此以二回合被完成，以每一回合處理 256 個樣本。暫態偵測被分為四個步驟：(1)高通濾波、(2)區塊分段為子多聲道、(3)在每一子區塊分段內之尖峰偵測、及(4)臨界值比較。該暫態偵測器為每一全帶寬聲道輸出一旗標 $blksw[n]$ ，其在被設定為“1”時表示在對應的聲道之 512 長度輸入區塊的第二半部有一暫態出現。

【00325】 (1) 高通濾波：該高通濾波器被施作為具有 8kHz 切斷之一串接雙線組直接型式 II 之 IIR 濾波器。

【00326】 (2) 區塊分段：256 個高通濾波後之樣本的區塊被分為階層樹，其中第一層代表 256 長度之區塊，第二層為兩個長度 128 之分段，及第三層四個長度 64 之分段。

【00327】 (3) 尖峰偵測：具有最大之樣本就該階層樹之每一層的每一分段被定出。單一層之尖峰如下列般地被指出：

【00328】 $P[j][k]=\max(x(n))$

【00329】 $n=(512 \text{ 毕}(k-1)/2^j), (512 \text{ 毕}(k-1)/2^j)+1, \dots, (512 \text{ 毕}k/2^j)-1$

【00330】 及 $k=1, \dots, 2^{(j-1)}$ ；

【00331】 其中 $x(n)=256$ 長度區塊中之第 n 樣本

【00332】 $j=1, 2, 3$ 為該階層之層數

【00333】 k =第 j 層內之分段數

【00334】 注意， $P[j][0]$ ，(即 $k=0$)被定義為在目之樹即刻之前被計算的樹之第 j 層的最後一分段的尖峰。例如，先行樹中之 $P[3][4]$ 為目前樹中之 $P[3][0]$ 。

【00335】 (4) 臨界值比較：該臨界值比較器之第一階段檢查在目前的區塊中是否有顯著的信號位準。此藉由比較目前區塊之整體尖峰值 $P[1][1]$ 與一「靜默的臨界值」被完成。若 $P[1][1]$ 低於此臨界值，則長區塊被迫使用。該靜默的臨界值為 $100/32768$ 。該比較器之下一階段為檢查該階層樹之每一層上相鄰分段的相對尖峰水準。若一特定層之任二相鄰分段的尖峰比超過此層之預先定義的臨界值被設定以表示在目前 256 長度之區塊中一暫態之出現。該等比值如下列地被比較：

【00336】 $\text{mag}(P[j][k] \times T[j]) > (F * \text{mag}(P[j][(k-1)]))$

【00337】 [注意該“F”敏感度因數]

【00338】 其中： $T[j]$ 為第 j 層被預先定義之臨界值，定義如下：

【00339】 $T[1]=0.1$

【00340】 $T[2]=0.075$

【00341】 $T[s]=0.05$

【00342】 若此不等式對任一層上任二分段尖峰為真，則一暫態就該 512 長度之輸入區塊的第一半部被指示。此處理之第二回合決定暫態在該 512 長度之輸入區塊的第二半部中出現。

【00343】 $N:M$ 編碼

【00344】 本發明之層面不限於相關第 1 圖所描述之 $N:1$ 編碼。更一般言之，本發明之層面可應用於以第 6 圖之方式(即 $N:M$ 編碼)變換任何數目之輸入聲道(n 輸入聲道)為任何數目之輸出聲道(m 輸出聲道)。由於在很多普通應用中，輸入聲道之數目 n 大於輸出聲道之數目 m ，第 6 圖之 $N:M$ 編碼配置將被稱為「向下混頻」以方便描述。

【00345】 參照第 6 圖之細節，取代如第 1 圖之配置中的加法組合器 6 將角旋轉 8 與角旋轉 10 之輸入相加的是，這些輸出可被施用至一向下混頻矩陣功能與裝置 6'(向下混頻矩陣)。向下混頻矩陣 6'可為一被動或主動矩

陣，其提供簡單的加為一聲道(如第 1 圖之 $N:1$ 編碼)或為多聲道。該等矩陣係數可為實數或複數(實數與虛數)。第 6 圖之其他功能與裝置與第 1 圖之配置相同，且其帶有相同的元件編號。

【00346】 向下混頻矩陣 6' 可提供一混合式頻率相依的函數，使得其例如提供頻率範圍為 f_1 至 f_2 之 mf_1-f_2 聲道及頻率範圍為 f_2 至 f_3 之 mf_2-f_3 聲道。例如在低於如 1000Hz 之一耦合頻率，向下混頻矩陣 6' 可提供二聲道，及在高於如 1000Hz 之一耦合頻率，向下混頻矩陣 6' 可提供一聲道。藉由運用低於該耦合頻率之二聲道，較佳的頻譜逼真度可被獲得，特別是若該等二聲道代表二水平方向(以配合人耳之水平性)為然。

【00347】 雖然第 6 圖顯示與第 1 圖配置就每一聲道產生相同的支鏈資訊，在一個或更多聲道被向下混頻矩陣 6' 之輸出提供時省略該等支鏈資訊之一為可能的。在一些情形中，可接受的結果只在振幅標度因數支鏈資訊被第 6 圖配置提供時可被獲得。有關支鏈選項之進一步細節在下面配合相關第 7, 8, 9 圖被討論。

【00348】 如剛剛上述者，向下混頻矩陣 6' 所提供之多聲道不必比輸入聲道之數目 n 小。當如第 6 圖之編碼器的目的為減少傳輸或儲存所用之位元數目時，其可能向下混頻矩陣 6' 所提供之多聲道比輸入聲道之數目 n 小。然而第 6 圖之配置亦可被用作為一「向上混頻器」。在此情形中，其可能有應用，其中向下混頻矩陣 6' 所提供之多聲道不必比輸入聲道之數目 n 大。

【00349】 $M:N$ 解碼

【00350】 第 2 圖之更一般化的形式在第 7 圖中被顯示，其中一向上混頻矩陣功能與裝置(或向上混頻矩陣)20 接收第 6 圖之配置所產生之 1 至 m 聲道。該向上混頻矩陣 20 可為一被動矩陣。其可為第 6 圖配置之向下混頻矩陣 6' 的共軛換位(即補數)。替選的是，該向上混頻矩陣 20 可為一主動矩陣——可變矩陣組合之一被動矩陣。若一主動矩陣解碼器被運用，在其放鬆狀態中，其可為該向下混頻矩陣之複數共軛或其可與該向下混頻矩陣為獨立的。該支鏈資訊可被施用為如第 7 圖顯示者以控制該調整振幅與角旋轉功能與裝置。在此情形中，該向上混頻矩陣(若為一主動矩陣)與該支鏈資

訊獨立地操作及僅對被施用至此之聲道響應。替選的是，一些或全部支鏈資訊可被施用至該主動矩陣以協助其操作。在此情形，一個或二個調整振幅與角旋轉功能與裝置可被省略。第 7 圖之解碼器例可如上述相關第 2 與 5 圖般地在某些信號狀況下運用施用一程度之隨機化振幅變異數的替選做法。

【00351】 當向上混頻矩陣 20 為一主動矩陣時，第 7 圖之配置的特徵在於為一「混合式矩陣解碼器」用於在一「混合式矩陣編碼器/解碼器系統」中操作。「混合式」在此文意中係指該解碼器可由其輸入音訊信號導出控制資訊之某些量度(即該主動矩陣對被施用至此之聲道中被編碼的頻譜資訊響應)，及由頻譜參數支鏈資訊導出控制資訊之進一步量度。用於混合式矩陣解碼器之適合的主動矩陣解碼器如上述很多有用的矩陣解碼器為本技藝相當習知的，包括“Pro Logic”與“Pro Logic II”解碼器(“Pro Logic 為杜比實驗室發照公司的註冊商標)及在下列一個或更多美國專利與公告之國際申請案(每一個指定給美國)所揭示之主題事項實施層面的矩陣解碼器：4,799,260；4,941,177；5,046,098；5,274,740；5,400,433；5,625,696；5,644,640；5,504,819；5,428,687；5,172,415；WO 01/41504；WO 01/41505；以及 WO 02/19768。第 7 圖之其他元件與第 2 圖之配置中者相同，且帶有相同的元件編號。

【00352】 替選的解除相關

【00353】 第 8 與 9 圖顯示一般化之第 7 圖的解碼器。特別是第 8 圖之配置與第 9 圖之配置顯示第 2 與 7 圖之解除相關技術的替選做法。在第 8 圖中，各別的解除相關器功能與裝置(解除相關器)46 與 48 為在 PCM 域內，每一個在其聲道的各別逆濾波器排組 30 與 36 後。在第 9 圖中，各別的解除相關器功能與裝置(解除相關器)50 與 52 為在頻率域內，每一個在其聲道的各別逆濾波器排組 30 與 36 前。在第 8 圖與第 9 圖配置二者中，每一解除相關器(46, 48, 50, 52)具有獨一的特徵，使得其輸出針對彼此相互地被解除相關。其解除相關標度因數例如可被用以控制在每一聲道中解除相關對未解除相關信號之比值。替選的是其暫態旗標亦可被用以如下面被解釋地移動該解除相關器之操作模式。在第 8 圖與第 9 圖配置二者中，每一解除相關器可為一施洛德式(Schroeder-type)的混響器，具有其本身獨特的特

徵，其中其混響程度用其解除相關標度因數被控制(例如藉由控制該解除相關輸出形成該解除相關輸入與輸出之一部分線性組合的程度被施作)。替選的是，其他可控制的解除相關技術可獨自地或彼此組合地或與該施洛德式混響器被運用。施洛德式混響器為相當習知的，且可由二期刊論文追蹤其起源：IRE Transactions on Audio, 1961 年 AU-9 期，pp.209-214，M. R. Schroeder 與 B. F. Logan 之“‘Colorless’ Artificial Reverberation”與 A.E.S.期刊 1962 年 7 月，第 10 卷第 2 期，pp.219-223，M. R. Schroeder 之“Natural Sounding Artificial Reverberation”。

【00354】當解除相關器 46 與 48 如在第 8 圖配置中地於 PCM 域中操作時，需要單一(即寬帶)的解除相關標度因數。此可用任一數種方法被獲得。例如單一的解除相關標度因數可在第 1 圖或第 7 圖之編碼器中被產生。替選的是，若第 1 圖或第 7 圖之編碼器以子帶為基準產生解除相關標度因數，該等解除相關標度因數可在振幅或電力上於第 1 圖或第 7 圖之編碼器或第 8 圖之解碼器中被相加。

【00355】當解除相關器 50 與 52 如第 9 圖配置中在頻率域操作時，其可為每一子帶或多群組之子帶接收一解除相關標度因數，且附隨地為該等子帶或多群組之子帶提供解除相關之一相稱的程度。

【00356】第 8 圖之解除相關器 46 與 48 及第 9 圖之解除相關器 50 與 52 可備選地接收該暫態旗標。在第 8 圖之 PCM 域解除相關器中，該暫態旗標可被運用以移動各別解除相關器之操作模式。例如，該解除相關器可在暫態未出現時操作成一施洛德式混響器，但在此接收之際就短的后續期間(如 1 至 10 毫秒)操作成固定的延遲。每一聲道可具有預設之固定的延遲或該延遲可在響應一短期間內之數個暫態下被改變。在第 9 圖之頻率域解除相關器中，該暫態旗標亦可被運用以移動各別解除相關器之操作模式。然而在此情形中，一暫態旗標之接收例如可觸發其中該旗標發生之聲道中振幅的短(數毫秒)增加。

【00357】如上述者，當除了支鏈資訊外有二個或更多的聲道被傳送時，減少支鏈參數之數目為可接受的。例如，僅傳送振幅標度因數為可接受的，在此情形中，解碼器中之解除相關與角度功能與裝置可被省略(在此

情形，第 7，8 與 9 圖縮減為同一配置)。

【00358】 替選的是，只有振幅標度因數、解除相關標度因數與備選的暫態旗標可被傳送。在此情形，任一第 7，8 或 9 圖配置可被運用(省略其每一中之角旋轉 28 與 34)。

【00359】 至於另一替選做法為只有振幅標度因數與角控制參數被傳送。在此情形，任一第 7，8 或 9 圖配置可被運用(省略第 7 圖之解除相關器 38 與 42 及第 8 與 9 圖之 46，48，50，52)。

【00360】 如在第 1 與 2 圖者，第 6-9 圖之配置欲顯示任何數目之輸入與輸出聲道，雖然為了呈現簡單起見只有二聲道被顯示。

【00361】 混合式單聲道/立體聲編碼與解碼

【00362】 如配合上述相關第 1，2 與 6 至 9 圖之例子的描述，本發明之層面就改善低位元率編碼/解碼系統之績效亦為有用的，其中離散的二聲道(立體聲，其可已由多於二聲道被向下混頻)輸入音訊信號在二聲道例如用感覺式編碼被編碼、傳輸或儲存、解碼及再生為低於一耦合頻率 f_m 之一離散的立體聲音訊信號與一般為高於該頻率 f_m 之一單聲道(mono)音訊信號(換言之，在高於該 f_m 頻率，二聲道中實質上無立體聲聲道隔離—其二者基本上承載相同的音訊資訊)。藉由在高於該耦合頻率 f_m 組合該等立體聲輸入聲道，需要被傳輸或儲存之位元較少。藉由運用適合的耦合頻率，被產生之混合式單聲/立體聲信號可依音訊材料與聆聽者之感覺性而定地提供可接受的績效。如上述配合相關第 1 與 6 圖之例子的描述，低至 2300Hz 甚至是 1000Hz 的一耦合或暫態頻率可為適當的，但該耦合頻率並非為關鍵的。耦合頻率之另一可能的選擇為 4kHz。其他的頻率可在位元節省與聆聽者接受度間提供有用的平衡，且特定耦合頻率之選擇對本發明並非為關鍵的。該耦合可為可變的，若為可變的，其例如可直接或間接地依輸入信號特徵而定。

【00363】 雖然此一系統為大多數的音樂材料與大多數聆聽者提供可接受之結果，假設該等改善為可向後計算且不提供被設計來接收該等混合式單聲/立體聲信號之退化或不可用的解碼器「繼承物」的已安裝基礎時，其可能欲改善此一系統之績效。這類改善例如可包括額外的再生聲道，如

「環繞音效」聲道。雖然環繞音效聲道可利用一主動矩陣解碼器由一個二聲道立體聲信號被導出，很多此類解碼器運用帶寬控制電路，其僅在被施用至此的信號對整個該等信號之帶寬為立體聲時可適當地操作—當混合式單聲/立體聲信號被施用至此時此類解碼器在一些信號狀況下未適當地操作。

【00364】 例如，在一 2:5(二聲道進、五聲道出)之矩陣解碼器中其提供代表左前、前中、右前、左(後面/側面)環繞與右(後面/側面)環繞方向輸出，並在基本上同一信號被施用至其輸入時操縱其輸出至前中，高於該頻率 f_m 之一凌越的信號(此處即一混合式單聲/立體聲系統中之單聲道信號)可致使所有的信號成份(包括可瞬間出現之低於頻率 f_m 者)被該前中輸出再生。此矩陣解碼器特徵會在該凌越的信號由高於 f_m 移位至低於 f_m 時形成突然的信號位置移位之結果，反之亦然。

【00365】 運用寬帶控制電路之主動矩陣解碼器的例子包括 Dolby Pro Logic 與 Dolby Pro Logic II 解碼器。“Dolby”與“Pro Dolby”為 Dolby 實驗室發照公司之註冊商標。Pro Logic 解碼器之層面在美國專利第 4,799,260 與 4,941,177 號被揭示，其每一個整體被納於此處做為參考。Pro Logic II 解碼器之層面被揭示於 2000 年 3 月 22 日申請之美國專利審理中案件第 S. N. 09/532,711 號且在 2001 年 6 月 7 日被公告為 WO 01/41504 的 Fosgate 之題目為“Method for Deriving at Least Three Audio Signal from Two Input Audio Signal”與 2003 年 2 月 25 日申請之美國專利審理中案件第 S. N. 10/362,786 號且在 2004 年 7 月 1 日被公告為 US 2004/0125960 A1 的 Fosgate 等人之題目為“Method for Apparatus for Audio Matrix Decoding”。每一該等申請案之整體被納於此處做為參考。Dolby Pro Logic 與 Pro Logic II 解碼器之操作的一些層面例如在 Dolby 實驗室之網頁(www.dolby.com)可取得之論文：Roger Dressler 之“Dolby Surround Pro Logic Decoder Principles of Operation”與 Jim Hilson 之“Mixing with Dolby Pro Logic II Technology”中被解釋。其他的主動矩陣解碼器被習知，其運用寬帶控制電路與導出來自一個二聲道立體聲輸入之多於二輸出聲道。

【00366】 本發明之層面不受限於使用 Dolby Pro Logic 或 Dolby Pro II

矩陣解碼器。替選的是，該主動矩陣解碼器可如為在 Davis 之國際專利申請案 PCT/US02/03619，題目為“Auido Channel Translation”，且指定給美國在 2002 年 8 月 15 日被公告為 WO 02/063925 A2 及 Davis 之國際專利申請案 PCT/US2003/024570，題目為“Auido Channel Spatial Translation”，且指定給美國在 2004 年 3 月 4 日被公告為 WO 2004/019656 A2 被描述的多頻帶主動矩陣解碼器。每一該等國際專利申請案之整體被納於此處做為參考。雖然，由於其多頻帶控制，此主動矩陣解碼器在一繼承單聲/立體聲解碼器被使用時不會遭受該凌越的信號由高於 f_m 移位至低於 f_m (反之亦然)的突然信號位置移位之問題(不論是否有凌越信號成份高於頻率 f_m ，該多頻帶主動矩陣解碼器正常地就低於頻率 f_m 之信號成份操作)，此種多頻帶主動矩陣解碼器在其輸入為如上述之單聲/立體聲信號時不提供高於該頻率 f_m 之聲道相乘。

【00367】 放大低位元率混合式立體/單聲編碼/解碼描述(如剛所描述之系統或類似的系統)，使得高於頻率 f_m 之單聲道音訊資訊被放大而近似該原始立體聲音訊資訊會為有用的，至少在被施用至一主動矩陣解碼器(特別是運用寬帶控制電路者)時到達形成被放大之二聲道音訊的結果之程度，致使該矩陣解碼器實質地或更幾近地操作成就好像該原始寬頻帶立體聲音訊資訊被施用至此。

【00368】 如將被描述者，本發明之層面亦可被運用以改善在一混合式單聲/立體聲解碼器中向下混頻為單聲道。此改善後之向下混頻不論在上述之放大是否被運用及不論一主動矩陣解碼器是否在一混合式單聲/立體聲解碼器之輸出被運用，於改善一混合式單聲/立體聲的再生輸出為有用的。

【00369】 其將被了解本發明之其他變形與修改之施作對熟習本技藝者將為明白的，及本發明不受限於所描述之這些特定的實施例。其因而企圖以本發明涵蓋任何與所有修改、變形或等值事項，其落在此處所揭示之基本的基礎原理之真實精神與領域。

【符號說明】

【00370】

2...濾波器排組

6...加法組合器

4...濾波器排組

6'...向下混頻矩陣

- 8...旋轉角
- 10...旋轉角
- 12...音訊分析器
- 14...音訊分析器
- 20...解除相關矩陣
- 22...第一聲道音訊恢復路徑
- 24...第二聲道音訊恢復路徑
- 26...調整振幅
- 28...角旋轉
- 30...逆濾波器排組
- 32...調整振幅
- 34...角旋轉
- 36...逆濾波器排組
- 38...可控制的解除相關器
- 40...加法組合器
- 42...解除相關器
- 44...加法組合器
- 46...解除相關器
- 48...解除相關器
- 50...解除相關器
- 52...解除相關器

申請專利範圍

1. 一種用以解碼代表 N 個音訊聲道之 M 個經編碼音訊聲道的方法，其用以解碼該 M 個經編碼音訊聲道以及具一或多個的空間參數之一空間參數組，其中 N 為兩個或多個且其中該等空間參數之一或多個係經差分編碼，該方法包含：

a) 接收該 M 個經編碼音訊聲道和該空間參數組，

b) 將一差分解碼程序施用至該等一或多個經差分編碼的空間參數，

c) 從該 M 個經編碼聲道導出 N 個音訊信號，其中每個音訊信號係分成複數個頻帶，其中每個頻帶包含一或多個頻譜成分，及

d) 從該 N 個音訊信號和該等空間參數產生一多聲道輸出信號，其中

M 為兩個或多個，

該 N 個音訊信號之至少一者為從該等 M 個經編碼音訊聲道之至少兩個的一加權組合所導出之一相關信號，

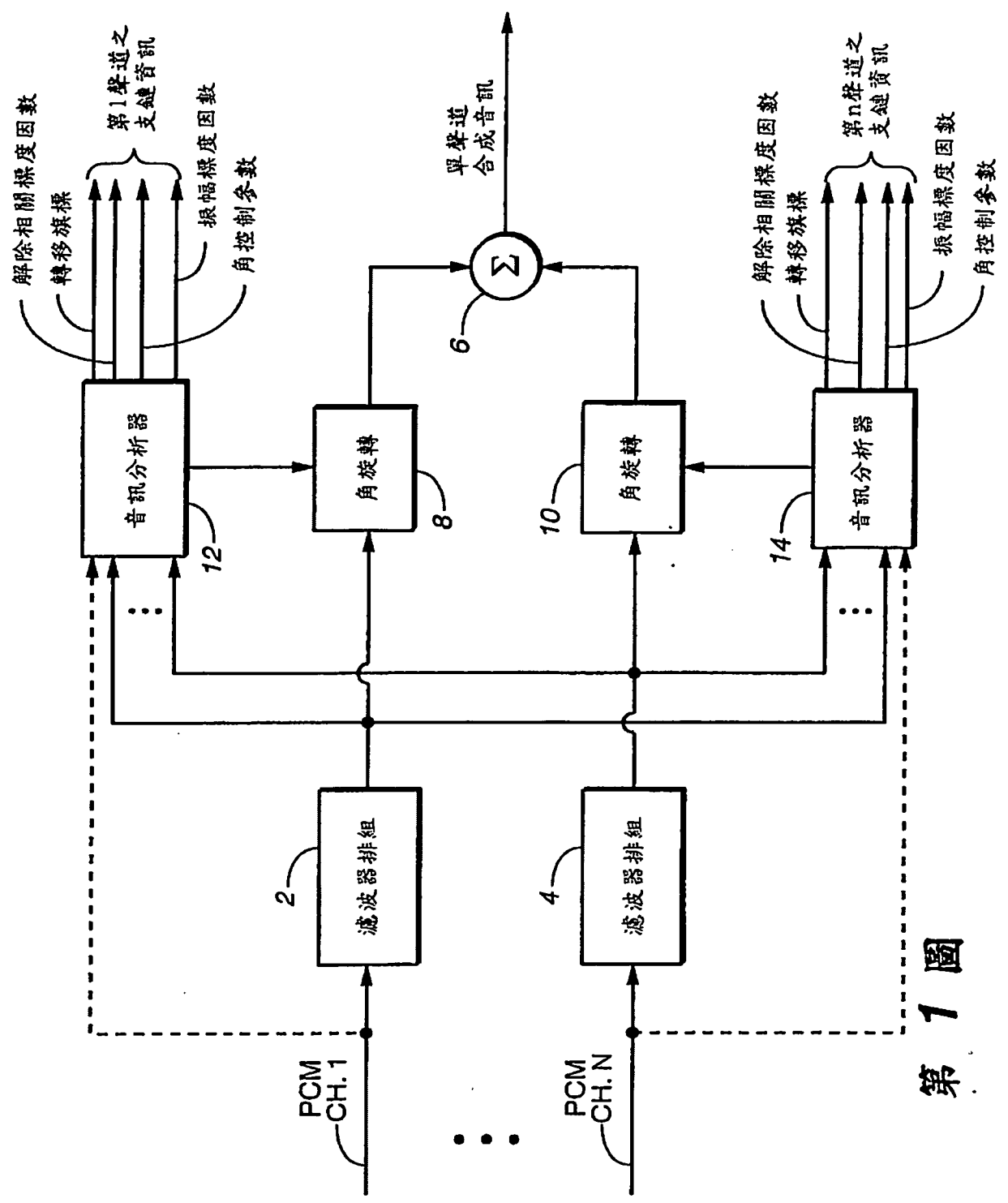
該空間參數組包括指示與一相關信號混和之一不相關信號之量之一第一參數，且步驟 d) 包括從該至少一相關信號導出至少一不相關信號，且響應於該等空間參數之一者或多者來控制在該多聲道輸出信號之至少一聲道中之該至少一相關信號對該至少一不相關信號的比例，其中該控制係至少部分根據該第一參數。

2. 如請求項 1 所述之方法，其中運用該差分解碼程序橫跨時間部份。

3. 如請求項 1 所述之方法，其中運用該差分解碼程序橫跨頻率部份。

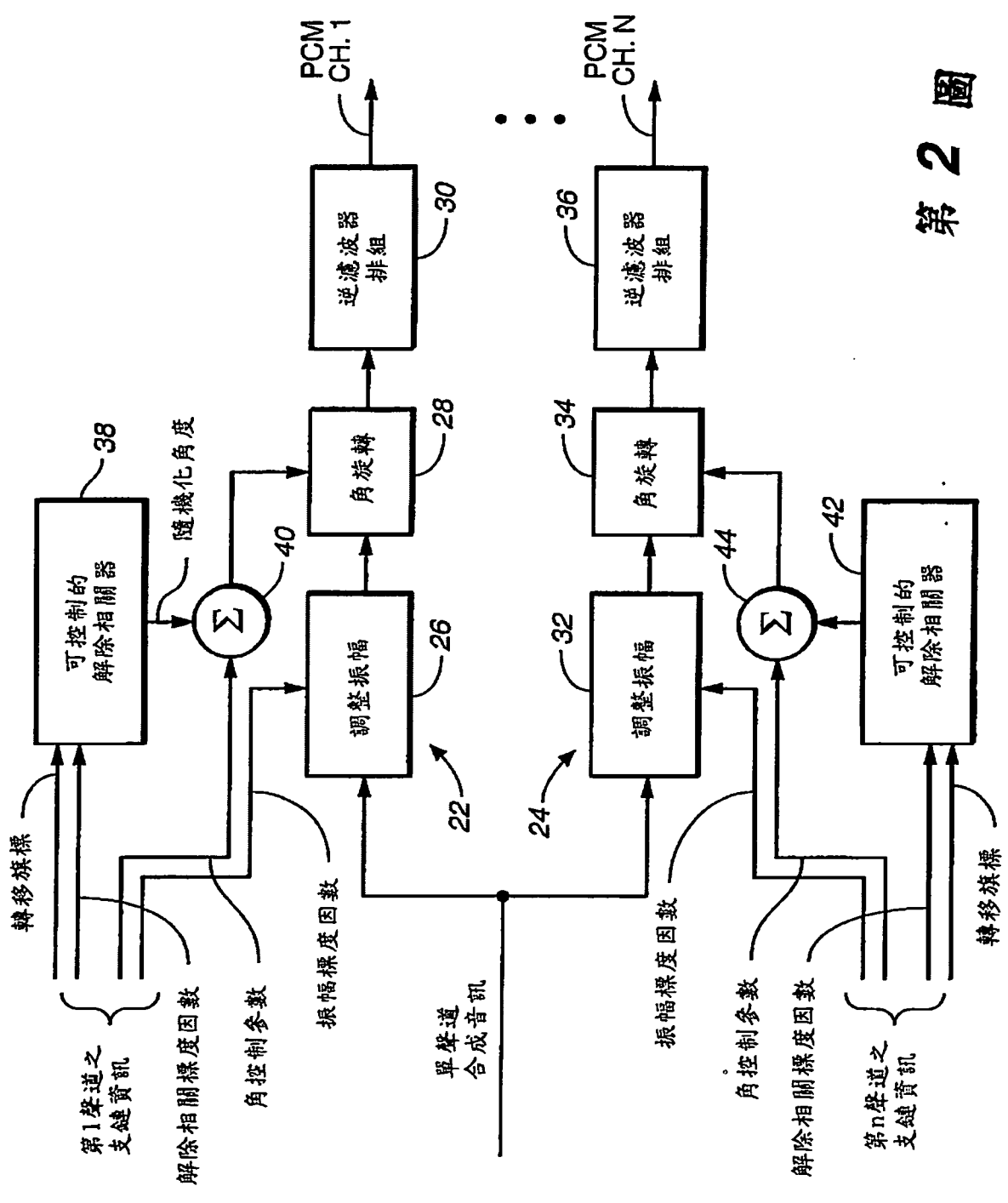
4. 如請求項 1 所述之方法，其中運用該差分解碼程序橫跨時間和頻率兩者。

03年11月3日 修正頁(本)

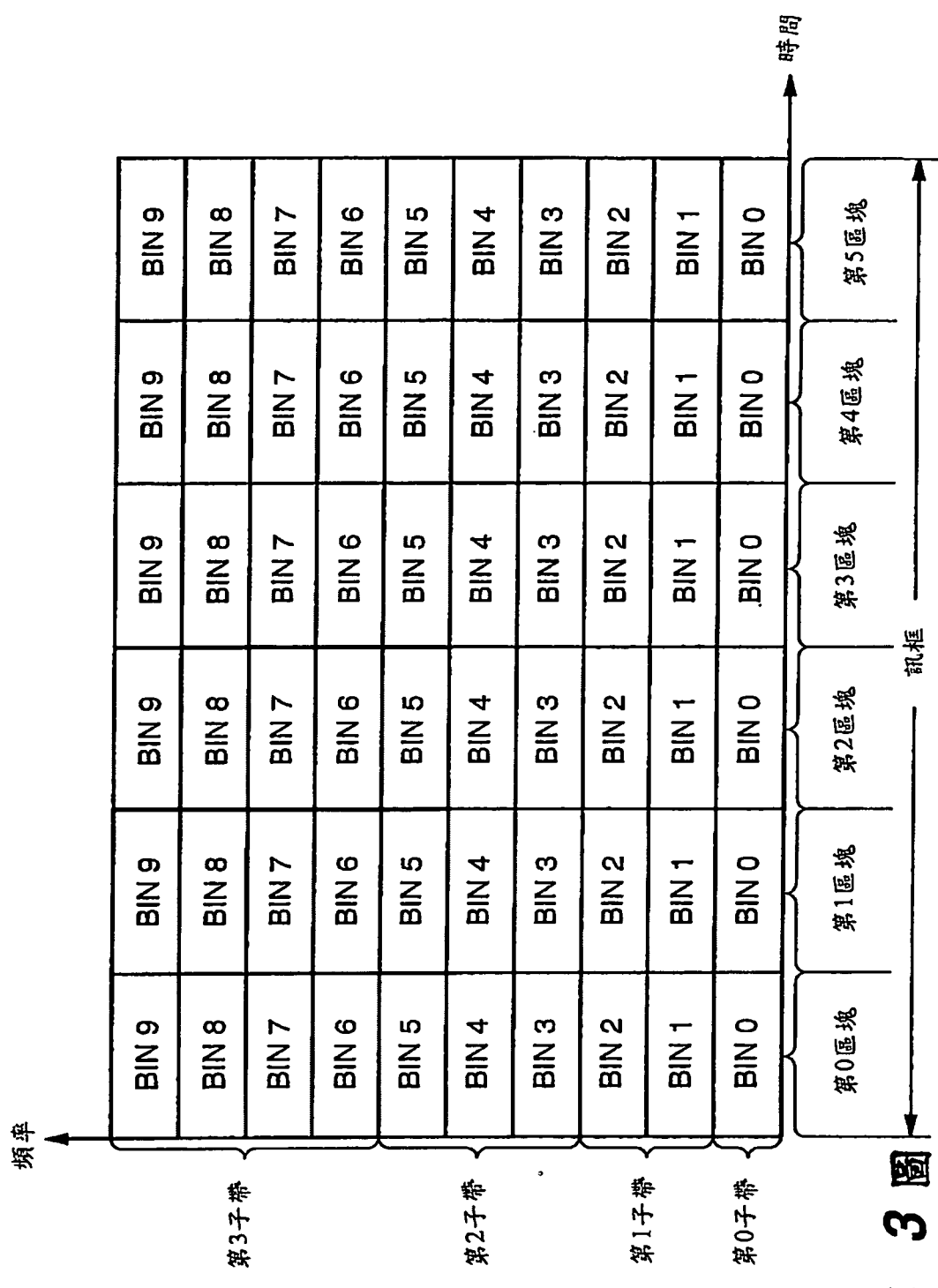


第 1 圖

83年11月3日修正頁(本)



第 2 圖

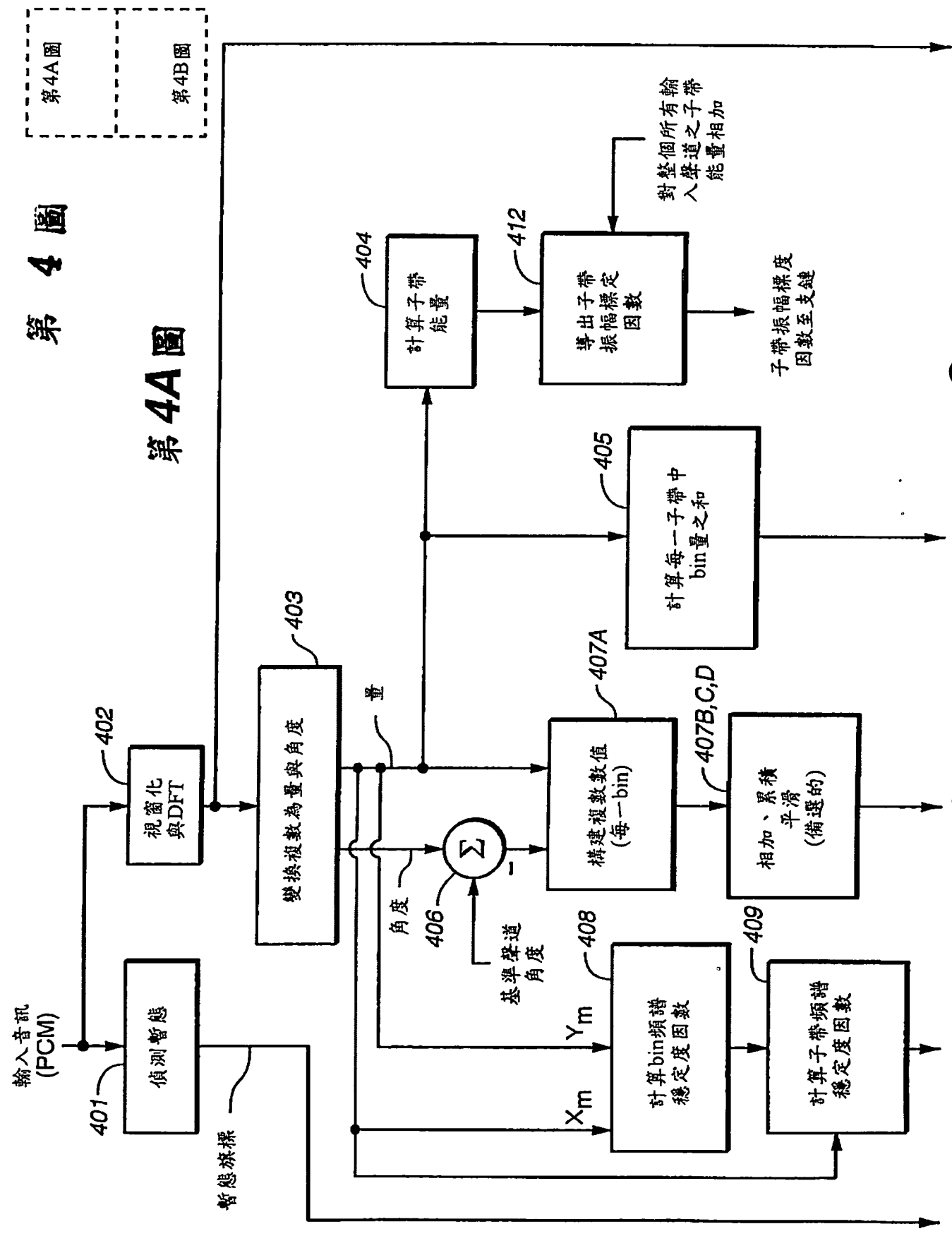


第 3 圖

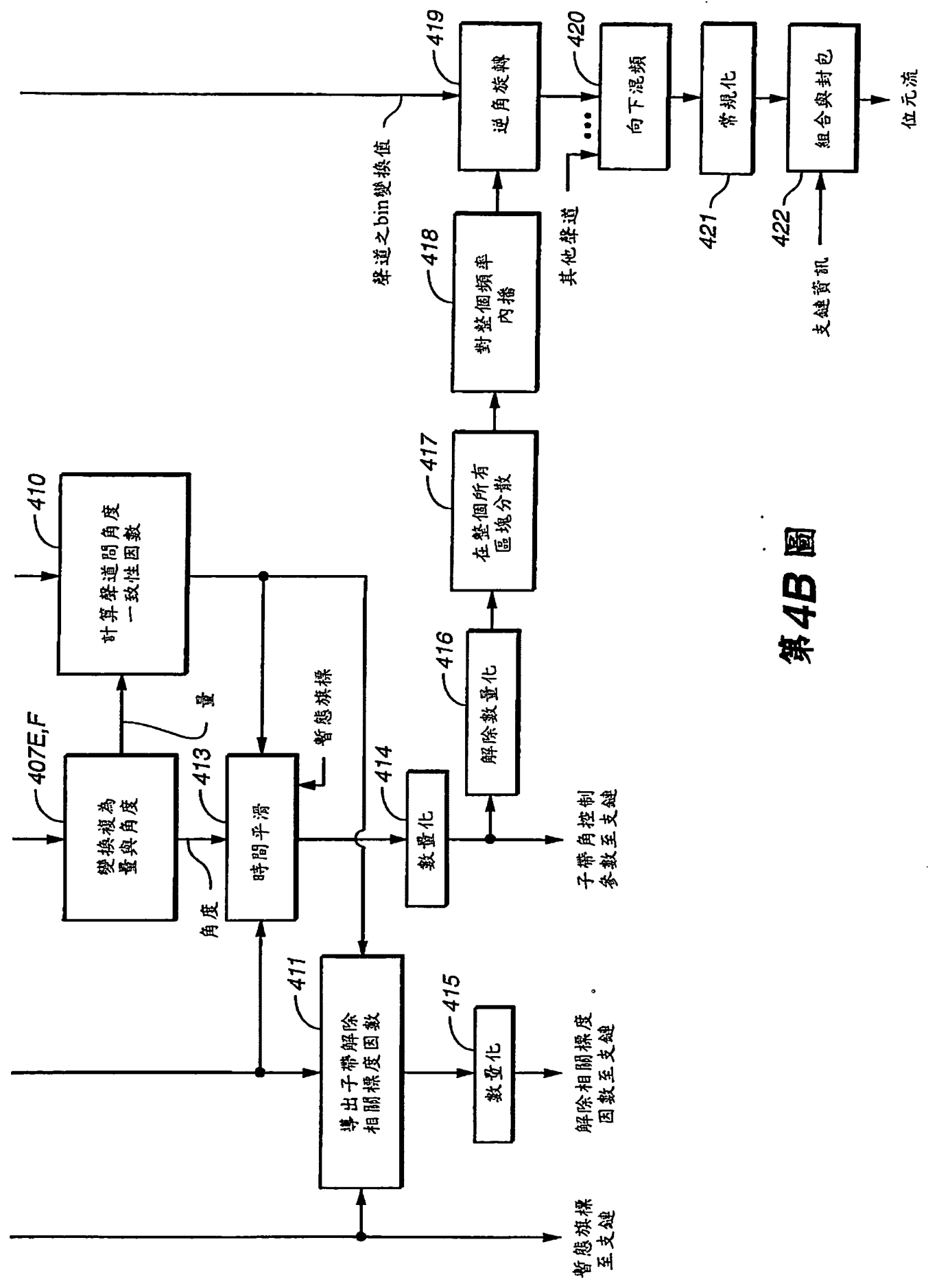
63年11月3日修正頁(本)

第 4 圖

第 4A 圖



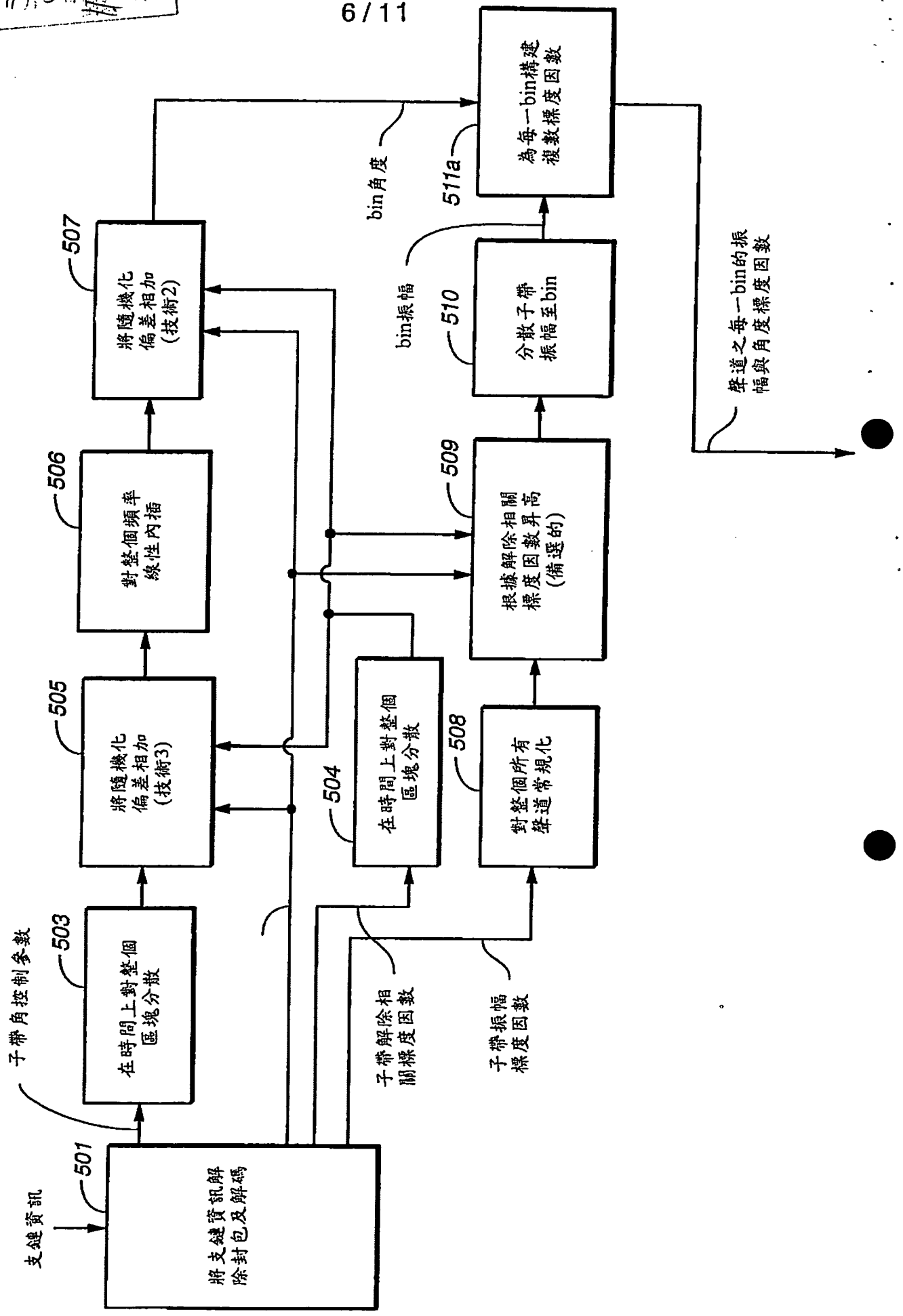
03年11月3日 修正 頁(末)

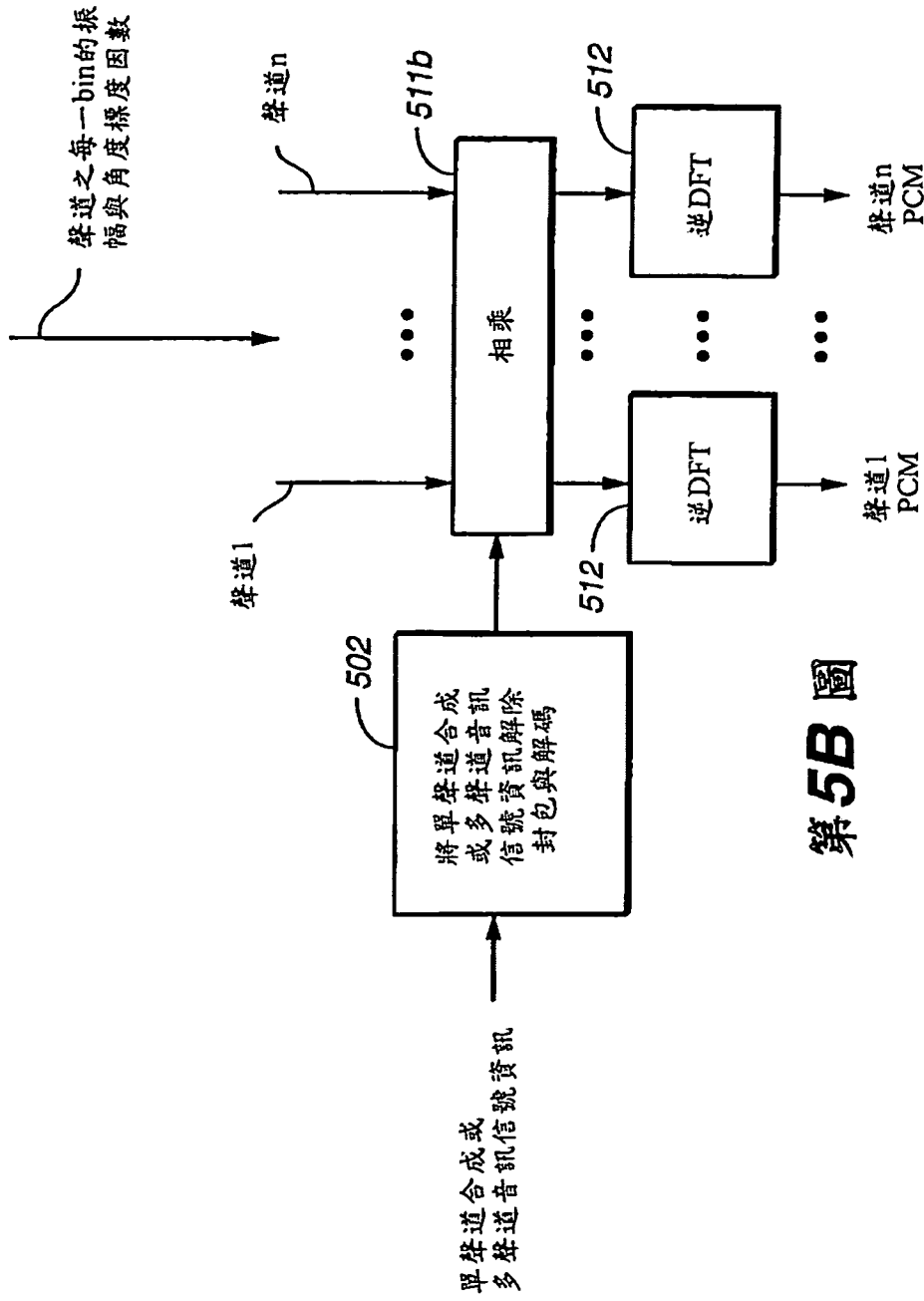


第4B圖

63年11月3日修正頁(本)

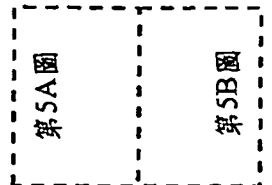
第5A圖

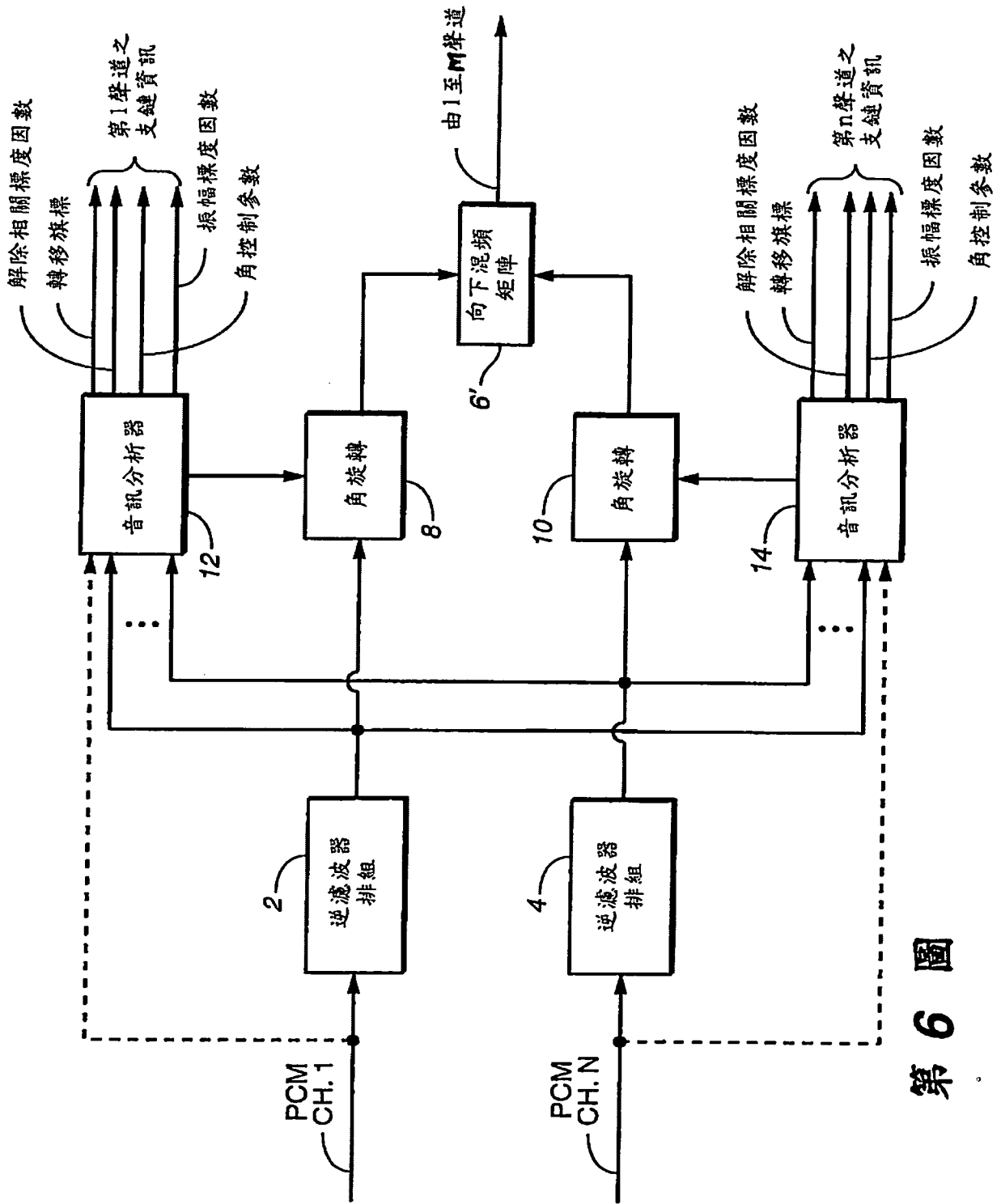




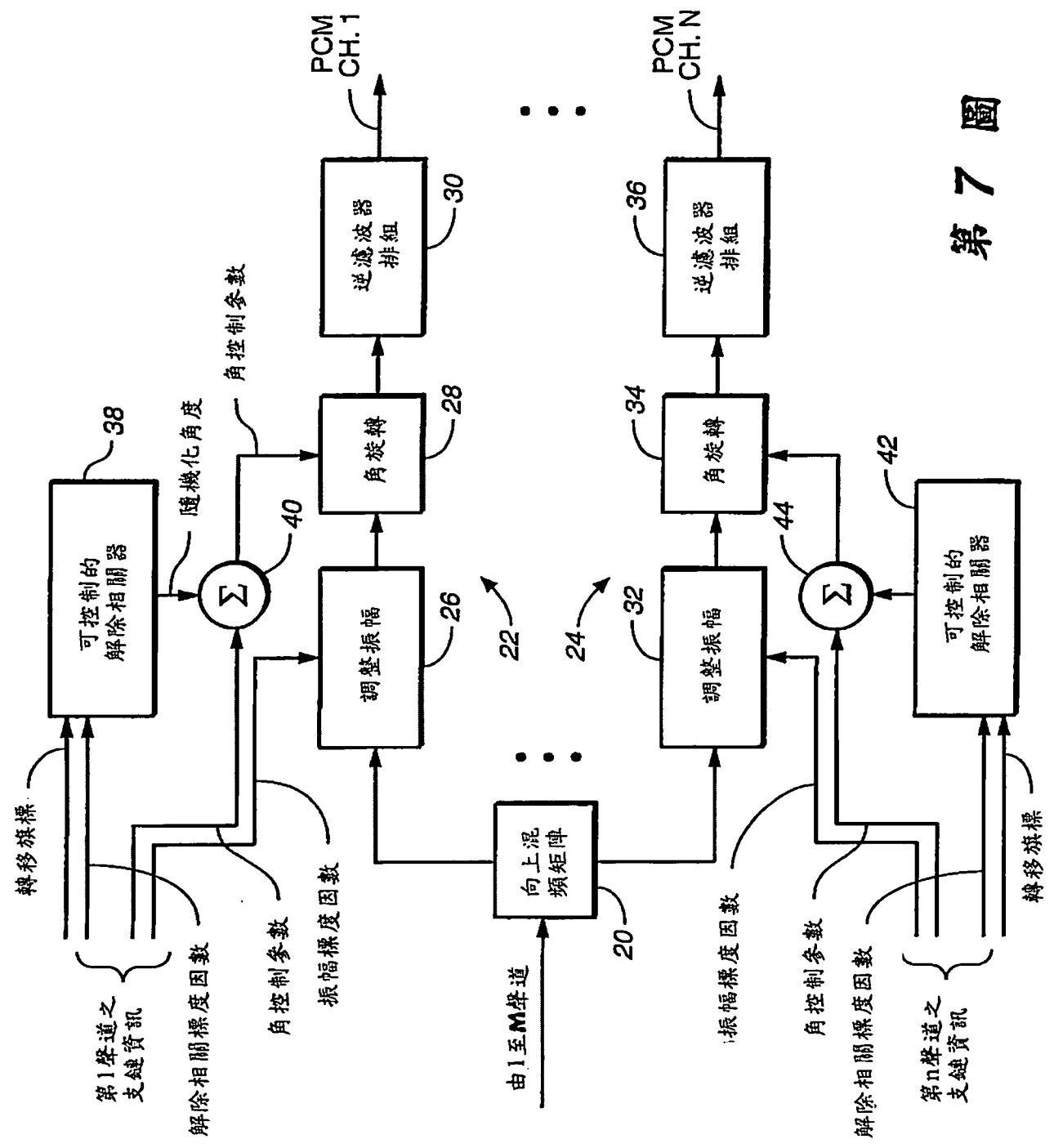
第5B圖

第5圖

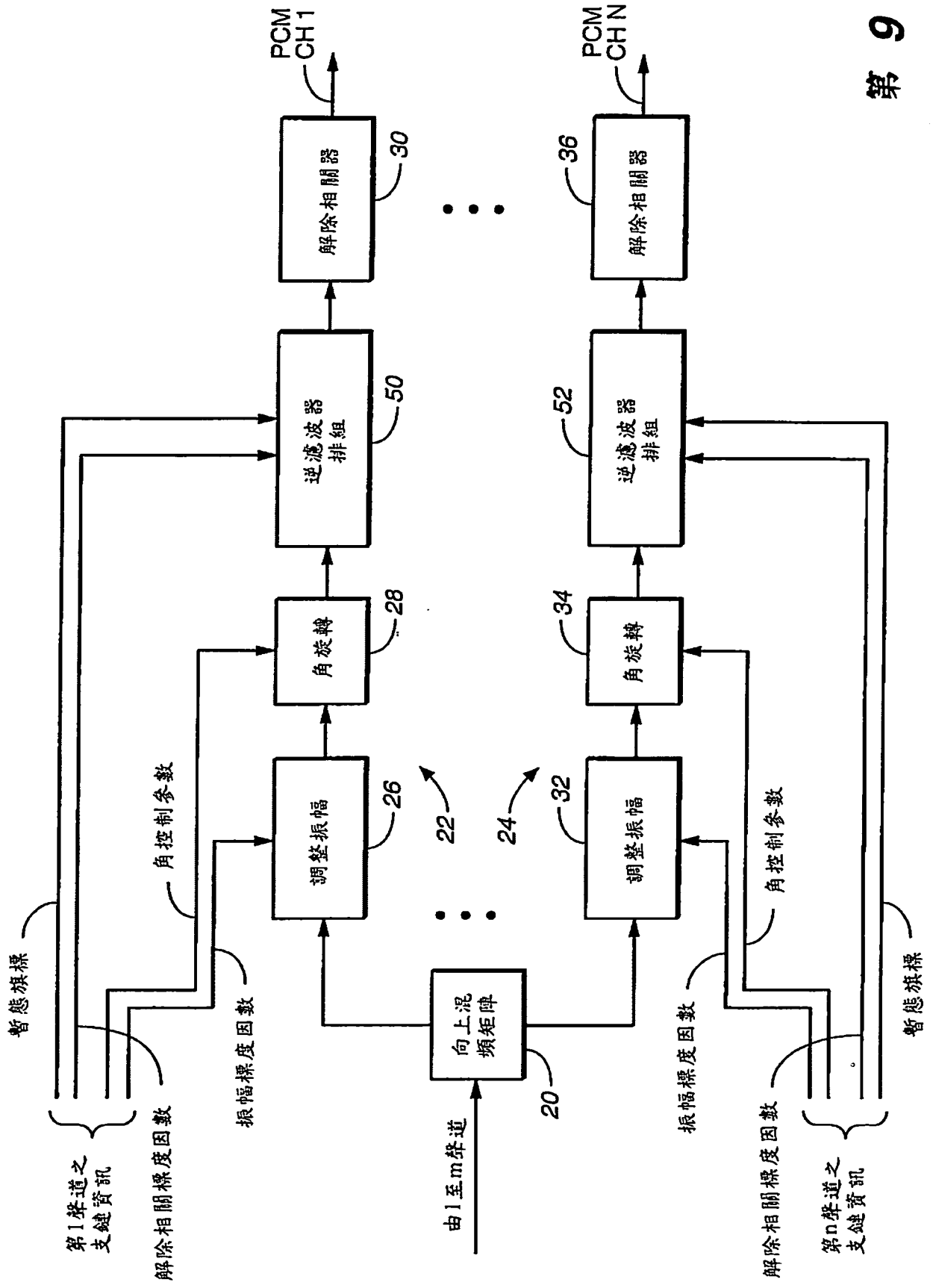




第 6 圖



第7圖



第 9 圖