

發明專利說明書

94年11月23日修正替換頁

(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：94128376

※申請日期：94.8.19

※IPC 分類：G10L 19/00 (2006.01)

一、發明名稱：(中文/英文)

處理一組N個聲音信號之方法與裝置及與其相關聯之電腦程式

Method and Apparatus for Processing a Set of N Audio Signals, and Computer Program Associated Therewith

二、申請人：(共 1 人)

姓名或名稱：(中文/英文)

杜比實驗室特許公司/DOLBY LABORATORIES LICENSING CORPORATION

代表人：(中文/英文)

夏默 艾都亞德 A./SCHUMMER, EDUARD A.

住居所或營業所地址：(中文/英文)

美國加州舊金山市·帕翠洛街100號

100 Potrero Avenue, San Francisco, California 94103, USA

國籍：(中文/英文)

美國/USA

三、發明人：(共 2 人)

姓名：(中文/英文)

1. 薛費特 艾倫 J./SEEFELDT, ALAN JEFFREY

2. 文頓 馬克 S./VINTON, MARK STUART

國籍：(中文/英文)

1. 美國/USA

2. 紐西蘭/New Zealand

10年11月23日修正替換頁

四、聲明事項：

主張專利法第二十二條第二項 第一款或 第二款規定之事實，其事實發生日期為： 年 月 日。

申請前已向下列國家（地區）申請專利：

【格式請依：受理國家（地區）、申請日、申請案號 順序註記】

有主張專利法第二十七條第一項國際優先權：

1. 美國; 2004,08,25; 60/604,725

2. 美國; 2005,07,18; 60/700,137

3. 美國; 2005,08,05; 60/705,784

無主張專利法第二十七條第一項國際優先權：

主張專利法第二十九條第一項國內優先權：

【格式請依：申請日、申請案號 順序註記】

主張專利法第三十條生物材料：

須寄存生物材料者：

國內生物材料 【格式請依：寄存機構、日期、號碼 順序註記】

國外生物材料 【格式請依：寄存國家、機構、日期、號碼 順序註記】

不須寄存生物材料者：

所屬技術領域中具有通常知識者易於獲得時，不須寄存。

五、中文發明摘要：

N個聲音信號中的每一個被一唯一解相關濾波器特性濾波，該特性是在時域內的一因果線性非時變特性或其在頻域內的等效特性，且對於每一解相關濾波器特性，以一時變和頻變的方式，組合其輸入和輸出信號以提供一組N個處理的信號。該組解相關濾波器特性被設計，使得所有的該等輸入和輸出信號近似彼此解相關。該組N個聲音信號可藉由向上混合從M個聲音信號中被合成，其中M是一或更多且N大於M。

六、英文發明摘要：

Each of N audio signals are filtered with a unique decorrelating filter characteristic, the characteristic being a causal linear time-invariant characteristic in the time domain or the equivalent thereof in the frequency domain, and, for each decorrelating filter characteristic, combining, in a time and frequency varying manner, its input and output signals to provide a set of N processed signals. The set of decorrelation filter characteristics are designed so that all of the input and output signals are approximately mutually decorrelated. The set of N audio signals may be synthesized from M audio signals by upmixing, where M is one or more and N is greater than M.

七、指定代表圖：

(一)本案指定代表圖為：第(3)圖。

(二)本代表圖之元件符號簡單說明：

32...位元流解包器

40...乘法器

34...時域到頻域的轉換器或轉換函數

42...計算器或計算函數

36...向上混合器

44...乘法器

38...解相關濾波器

46...加法組合器或組合函數

八、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

九、發明說明：

【發明所屬之技術領域】

發明領域

本發明是關於音訊編碼器、解碼器和系統，關於相對應的方法，關於實施該等方法的電腦程式，以及關於由該等編碼器產生的一位元流(bitstream)。

【先前技術】

發明背景

一些(certain)最近提出的有限位元率(bit rate)編碼技術分析(analyze)一輸入多聲道信號以獲得(derive)一向下混合(downmix)的複合信號(composite signal)(一信號包含少於該輸入信號的聲道)和旁側資訊(side-information)，該旁側資訊包含原聲場(original sound field)的一參數模型(parametric model)。該旁側資訊和複合信號被發送(transmitted)到一解碼器，該解碼器在該複合信號上運用該參數模型以再產生(recreate)該原聲場的近似值。該等“空間編碼(spatial coding)”系統的主要目的是以非常有限的資料來再產生一多聲道聲場；因此其加強(enforce)了被用於模擬該原聲場的該參數模型的限制。該等空間編碼系統的細節被包含在多種文獻(documents)中，該等文獻包括在下文被引用在標題“合併參考(incorporation by reference)”下的文獻。

該等空間編碼系統通常採用參數以模擬該原聲場，如聲道間(interchannel)振幅差異、聲道間時間或相位差異和聲

道間交叉相關性(cross-correlation)。通常該等參數在每一正在被編碼的聲道之多頻譜帶(spectral bands)被評估(estimated)且隨著時間過去被動態評估。

第1a(編碼器)圖和第1b(解碼器)圖顯示了一典型的先前技術之空間編碼系統。利用一重疊(overlapped)離散頻率轉換(discrete frequency transform, DFT)，多輸入信號被轉換到頻域(frequency domain)。然後DFT頻譜被再分為接近聽覺臨界頻帶(ear's critical bands)的頻帶。對於每一頻帶，該聲道間振幅差異、聲道間時間或相位差異和聲道間相關性的一評估被計算出。該等評估被用於向下混合該等原輸入信號到一單音(monophonic)複合信號中。該複合信號隨同該等評估的空間參數被送入一解碼器，其中利用相同的重疊DFT和臨界頻帶間隔，該複合信號被轉換到頻域。然後該等空間參數被運用到其等相對應的頻帶以產生該原多聲道信號的近似值。

在該解碼器中，該等聲道間振幅及時間或相位差異的運用相對比較簡單，但修改向上混合的聲道以使其等聲道間相關性與該原多聲道信號的匹配更有挑戰性。通常，在該解碼器中僅運用振幅和時間或相位差異時，該等向上混合聲道產生的聲道間相關性大於該原信號的，且產生的音訊聽起來比該原信號更空間“衰弱的(collapsed)”或更不環繞的(ambient)。其經常可歸因於橫過頻率和/或時間求平均值以限制該旁側資訊的發送損失(transmission cost)。為了恢復該原聲道間相關性的感受(perception)，一些類型的解

相關必須在至少一些向上混合的聲道上被實現。在下文引用的 Breebarrrt 等人 AES 會議論文 6072 且序列號為 WO 03/090206 的國際申請案中，一種用於徵收(impose)兩個聲道間之一理想的(desired)聲道間相關性的技術被主張，該等

5 聲道已從一單獨向下混合的聲道中被向上混合。該向下混合的聲道首先穿過(run through)一解相關濾波器以產生一第二解相關信號。然後該二向上混合聲道中的每一個以該原向下混合信號和該解相關信號的線性組合(linear combination)被計算。該解相關濾波器被設計成一依頻延遲

10 (frequency dependent delay)濾波器，其中該延遲隨著頻率增加而減少。該濾波器具有理想特性，當減少暫態(transients)的時間差量(temporal dispersion)時，提供顯著的且聽得見的解相關。同樣，在該原信號中加入該解相關信號可不導致關於一固定延遲解相關濾波器之梳形濾波器(comb filter)

15 效應。

在 Breebarrrt 等人論文及申請案中的該技術被設計成僅用於兩個向上混合聲道，但該種技術被期望用於任意數目的向上混合聲道。本發明的各方面不僅為該更普遍的多聲道解相關問題提供一解決方案，也在頻域內提供一有效的

20 實施。

【發明內容】

發明概要

本發明的一個方面提供處理一組 N 個聲音信號的技術，其藉由該等 N 個信號中的每一個以一唯一的解相關濾波

器特性來濾波，該特性是在時域內的一因果(causal)線性非時變(time-invariant)特性或其在頻域內的等效物，且對於每一解相關濾波器特性，以一時變和頻變方式，組合其輸入和輸出信號以提供一組N個處理的信號。該組合可以是一線性組合且可在接收到的參數幫助下運作。每一唯一的解相關濾波器特性可被選擇，使得每一濾波器特性的輸出信號與所有該N個聲音信號的相關性少於每一濾波器特性相對應的輸入信號與所有該N個信號的相關性，且使得每一輸出信號與所有其他輸出信號的相關性少於每一濾波器特性相對應的輸入信號與所有其他該N個信號的相關性。從而，每一唯一的解相關濾波器被選擇，使得每一濾波器的輸出信號與該N個聲音信號中的每一個近似解相關，且使得每一輸出信號與所有其他輸出信號近似解相關。該組N個聲音信號可從M個聲音信號中被合成，其中M是一或更多且N大於M，在這個情形中可有向上混合該M個聲音信號到N個聲音信號。

依據本發明的更多方面，描述該N個合成的聲音信號間的理想空間關係的參數可被接收到，在這個情形中該向上混合可在接收到的參數幫助下運作。該等接收到的參數可描述該N個合成的聲音信號間的理想空間關係，且該向上混合可在接收到的參數幫助下運作。

依據本發明的其他方面，每一解相關濾波器特性可由一具有多自由度(degree of freedom)的模型來定性。每一解相關濾波器特性可具有一頻變延遲形式的響應，其中該延

遲隨著頻率增加而單調(monotonically)減少。每一濾波器特性的脈衝響應可被一有限期間(finite duration)的正弦序列(sinusoidal sequence)指定，該序列的瞬間頻率單調減少，如在該序列的期間從 π 到零。例如，一雜訊序列(noise sequence)可被加入到該正弦序列的瞬間相位中，以在某些信號條件下減少聽得見的人工因素(audible artifact)。

依據本發明的再其他方面，描述該 N 個處理的信號中之理想空間關係的參數可被接收到，且組合度(degree of combining)可在接收到的參數幫助下運作。每一聲音信號可表示聲道，且幫助該組合運作之該等接收到的參數可以是關於聲道間交叉相關的參數。其他接收到的參數包括關於聲道間振幅差異和聲道間時間或相位差異中一個或更多個的參數。

例如，本發明運用於一空間編碼系統，其中在一編碼器中 N 個原聲音信號被向下混合到 M 個信號($M < N$)，且隨後利用在該編碼器中產生的旁側資訊，在一解碼器中被向上混合回 N 個信號。本發明的各方面不僅能運用在空間編碼系統中，如下文引用中被描述的系統，其中該多聲道向下混合是到(而該向上混合是從)一單個單音聲道，也能運用在該向下混合是到(而該向上混合是從)多聲道的系統，如在Mark Franklin Davis 於2005年2月28日提出申請且名為“Low Bit Rate Audio Encoding and Decoding in Which Multiple Channels Are Represented By Fewer Channels and Auxiliary Information”的國際申請案PCT/US2005/006359中

所揭露的系統。該PCT/US2005/006359申請案藉此全部被併入參考。

在該解碼器中，藉由運用在該旁側資訊中被發送的聲道間振幅和時間或相位差異，一第一組 N 個向上混合的信號從該 M 個向下混合的信號中被產生。下一步，藉由來自該第一組的該 N 個信號中的每一個用一唯一的解相關濾波器來濾波，一第二組 N 個向上混合的信號被產生。該等濾波器是“唯一的”，意思是有 N 個不同的解相關濾波器，每一信號對應一個。該組 N 個唯一的解相關濾波器被設計成產生 N 個彼此解相關的信號(見下文方程式3b)，該等信號也與該等濾波器輸入解相關(見下文方程式3a)。該等充分解相關(well-decorrelated)信號隨同未濾波的向上混合信號一起被使用，以從該解碼器產生輸出信號，該等輸出信號各自近似到該編碼器的每一輸入信號。每一近似值都以一線性組合被計算，該線性組合是來自該第一組向上混合信號的每一未濾波信號和來自該第二組向上混合信號的相對應的已濾波信號的線性組合。該線性組合的係數隨著時間和頻率而改變，且在由該編碼器產生的旁側資訊中被發送到該解碼器。為了在一些情形下有效實施該系統，較佳地，該 N 個解相關濾波器可被運用在頻域而不是時域。例如，該系統可藉由適當補零(zero-padding)和視窗化(windowing)如下文描述之用在該編碼器和該解碼器中的一DFT來實施。該等濾波器也可被運用在時域中。

圖式簡單說明

第1a和1b圖分別是一典型先前技術的空間編碼之編碼器和解碼器的簡化方塊圖。

第2圖是實施本發明各方面的一編碼器或編碼函數的一範例的簡化功能示意方塊圖。

5 第3圖是實施本發明各方面的一解碼器或解碼函數的一範例的簡化功能示意方塊圖。

第4圖是適合實施本發明各方面的一分析/合成視窗對的理想化敘述。

【實施方式】

10 較佳實施例之詳細說明

參考第2和3圖，該 N 個原聲音信號由 x_i 表示， $i=1\dots N$ 。在該編碼器中產生的該 M 個向下混合信號由 y_j 表示， $j=1\dots M$ 。經由該等聲道間振幅和時間或相位差異的運用而在該解碼器中產生的該第一組向上混合信號由 z_i 表示，
15 $i=1\dots N$ 。該解碼器中的該第二組向上混合信號由 \bar{z}_i 表示， $i=1\dots N$ 。該第二組經由該第一組和該等解相關濾波器的卷積(convolution)而計算得到。

$$\bar{z}_i = h_i * z_i, \quad (1)$$

其中是關於信號 i 的解相關濾波器之脈衝響應。最後，該等
20 原信號的近似值由 \hat{x}_i 表示， $i=1\dots N$ 。該等信號藉由以一時變和頻變之方式混合來自該等描述的第一和第二組的信號來計算：

$$\hat{X}_i[b,t] = \alpha_i[b,t]Z_i[b,t] + \beta_i[b,t]\bar{Z}[b,t], \quad (2)$$

其中 $Z_i[b,t]$ ， $\bar{Z}_i[b,t]$ 和 $\hat{X}_i[b,t]$ 分別是信號 z_i 、 \bar{z}_i 和 \hat{x}_i 在臨界頻帶

b 和時間塊(time block) t 的短時間(short-time)頻率表示。該等參數 $\alpha_i[b, t]$ 和 $\beta_i[b, t]$ 是在該編碼器產生的旁側資訊中被指定的時變和頻變混合係數。其等可由下文在標題“混合係數的計算”下被描述的內容來計算。

5 解相關濾波器的設計

該組解相關濾波器 $h_i (i=1\dots N)$ 被設計使得所有該等信號 z_i 和 \bar{z}_i 近似彼此解相關：

$$E\{z_i \bar{z}_j\} \cong 0 \quad i=1\dots N, j=1\dots N, \quad (3a)$$

$$E\{\bar{z}_i \bar{z}_j\} \cong 0 \quad i=1\dots N, j=1\dots N, i \neq j, \quad (3b)$$

- 10 其中 E 表示期望運算子。換句話說，每一唯一的解相關濾波器特性被選擇，使得每一濾波器特性的輸出信號 \bar{z}_i 與所有該等輸入聲音信號 z_i 的相關性少於每一濾波器特性相對應的輸入信號與所有該等輸入信號的相關性，且使得每一輸出信號 \bar{z}_i 與所有其他輸出信號的相關性少於每一濾波器特
- 15 性相對應的輸入信號 z_i 與所有其他該等輸入信號的相關性。如先前技術中眾所周知的，一簡單延遲可被使用作為一解相關濾波器，其中當該延遲增加時解相關效果變得更強。然而，當一信號由該解相關器濾波然後與該原信號相
- 20 加時，如方程式2中指定的，回音(echoes)(尤其是在更高頻率的回音)可被聽到。先前技術中也已知的一改良為一頻變延遲濾波器，其中該延遲隨著頻率從一最大延遲到零而線性減少。該濾波器中僅有的(only)自由參數(free parameter)是該最大延遲。由於該濾波器，高頻率沒有被重大延遲，從而消除感覺到的回音，而較低的頻率仍接收到重大延

遲，因此保持該解相關效果。如本發明的一方面，由具有
 更多自由度(degrees of freedom)的一模型來定性的一解相
 關濾波器特性被提出。特別的，該濾波器可具有一單調減
 少的瞬間頻率函數(instantaneous frequency function)，理論
 5 上，其可採用(take on)一無限變化的形式。每一濾波器的脈
 衝響應被一有限期間的正弦序列指定，該序列的瞬間頻率
 單調減少，例如，在該序列的期間從 π 到零。其意味著對於
 奈奎斯特頻率(Nyquist frequency)該延遲等於零，而對於DC
 該延遲等於該序列的長度。在其一般形式中，每一濾波器
 10 的脈衝響應可由以下方程式給出：

$$h_i[n] = A_i \sqrt{|\omega_i(n)|} \cos(\phi_i(n)), \quad n=0 \dots L_i-1, \quad (4a)$$

$$\phi_i(t) = \int \omega_i(t) dt + \phi_0, \quad (4b)$$

其中 $\omega_i(t)$ 是該單調減少的瞬間頻率函數， $\omega_i(t)$ 是該瞬間頻
 率的第一導數(derivative)， $\phi_i(t)$ 是由該瞬間頻率的積分
 15 (integral)加上某一初始相位 ϕ_0 而得出的瞬間相位，而 L_i 是該
 濾波器的長度。該乘法項(multiplicative term) $\sqrt{|\omega_i(t)|}$ 是必要
 的以使 $h_i[n]$ 的頻率響應橫過所有頻率都近似平坦的(flat)，
 且該濾波器振幅 A_i 被選擇以使幅度(magnitude)頻率響應近
 似一(unity)。其相當於選擇 A_i 以使以下保持：

$$20 \quad \sum_{n=0}^{L_i-1} h_i^2[n] = 1. \quad (4c)$$

該函數 $\omega_i(t)$ 的一有效參數化由以下方程式給出：

$$\omega_i(t) = \pi \left(1 - \frac{t}{L_i} \right)^{\alpha_i}, \quad (5)$$

其中該參數 α_i 控制該瞬間頻率在該序列的期間多快減少到零。可使用方程式5將該延遲 t 解為一角頻率 (radian frequency) ω 的函數：

$$t_i(\omega) = L_i \left(1 - \left(\frac{\omega}{\pi} \right)^{\frac{1}{\alpha_i}} \right) \quad (6)$$

5 需要注意的是，當 $\alpha_i=0$ ，對所有的 ω ， $t_i(\omega)=L_i$ ；換句話說，該濾波器變成具有長度 L_i 的一純(pure)延遲。當 $\alpha_i=\infty$ ，對所有的 ω ， $t_i(\omega)=0$ ：該濾波器僅是一脈衝。為了聽覺解相關目的，已發現設定 α_i 在1和10之間的某處產生最佳聲音效果 (sounding results)。然而，由於方程式4a中的濾波器脈衝響應 $h_i[n]$ 有一像啁啾聲(chirp-like)序列的形式，利用該濾波器對脈衝聲音信號濾波，有時會導致在已濾波的信號內且在原暫態位置上之聽得見的“啁啾聲”人工因素。該效果的可聞度(audibility)隨著 α_i 增加而減少，但藉由加入一雜訊序列到該濾波器之正弦序列的瞬間相位中該效果可更進一步減少。其可藉由加入一雜訊項到濾波器響應的瞬間相位中被實現：

$$h_i[n] = A_i \sqrt{|\omega_i(n)|} \cos(\phi_i(n) + N_i[n]), \quad n=0 \dots L_i-1 \quad (7)$$

使該雜訊序列 $N_i[n]$ 等於具有一是 π 一小部分的方差 (variance)的白高斯雜訊(white Gaussian noise)，已足夠使該脈衝響應聽起來更像雜訊而不像啁啾聲，而由 $\omega_i(t)$ 指定的頻率和延遲間的理想關係仍然被大量保持。在方程式7中，具有如方程式5中指定的 $\omega_i(t)$ 的濾波器有四個自由參數： L_i ， α_i ， ϕ_0 和 $N_i[n]$ 。橫過所有該等濾波器 $h_i[n](i=1 \dots N)$ ，藉

由選擇彼此十分不同的該等參數，方程式3中的理想解相關條件可被滿足(met)。

混合係數的計算

該時變和頻變混合係數 $\alpha_i[b,t]$ 和 $\beta_i[b,t]$ 可從成對的該等原信號 x_i 之間的每一頻帶(per-band)相關性在該編碼器中被產生。特別的，在頻帶 b 和時間 t 上的信號 i 和 j 之間(其中“ i ”是該等信號 $1\dots N$ 中的任一個而“ j ”是該等信號 $1\dots N$ 中的其他任一個)的標準化(normalized)相關性可由以下方程式給出：

$$C_{ij}[b,t] = \frac{|E_{\tau}\{X_i[b,\tau]X_j^*[b,\tau]\}|}{\sqrt{E_{\tau}\{|X_i[b,\tau]\}^2}E_{\tau}\{|X_j[b,\tau]\}^2}}, \quad (8)$$

其中該期望值 E 是遍及時間 τ 且在時間 t 的附近區域中被實現。給定(3)中的條件和額外的限制 $\alpha_i^2[b,t] + \beta_i^2[b,t] = 1$ ，可顯示成對且每一個近似一輸入信號的該等解碼器輸出信號 \hat{x}_i 和 \hat{x}_j 間的標準化(normalized)相關性由以下方程式給出：

$$\hat{C}_{ij}[b,t] \equiv \alpha_i[b,t]\alpha_j[b,t]。 \quad (9)$$

本發明的一方面是識別該 N 個值 $\alpha_i[b,t]$ 不足以對所有 i 和 j 再產生值 $C_{ij}[b,t]$ ，但其等可被選擇，使得對於相對所有其它信號 j 的一特定信號 i ， $\hat{C}_{ij}[b,t] \equiv C_{ij}[b,t]$ 。本發明的另一方面是識別可選擇信號 i 作為在頻帶 b 和時間 t 時的最主要(the most dominant)信號。該主要信號被定義為在 $i=1\dots N$ 中 $E_{\tau}\{|X_i[b,\tau]\}^2$ 最強的信號。將該主要信號的索引標示為 d ，則該等參數 $\alpha_i[b,t]$ 由以下方程式給出：

$$\alpha_i[b,t]=1, \quad i=d,$$

$$\alpha_i[b,t]=C_{di}[b,t], \quad i \neq d. \quad (9)$$

該等參數 $\alpha_i[b,t]$ 被送入該空間編碼系統的旁側資訊中。在該解碼器中，該等參數 $\beta_i[b,t]$ 可由以下方程式被計算出：

$$5 \quad \beta_i[b,t]=\sqrt{1-\alpha_i^2[b,t]}. \quad (10)$$

為了減少該旁側資訊的發送損失，可僅發送參數 $\alpha_i[b,t]$ 給主要聲道和第二最主要聲道。然後給所有其它聲道的 $\alpha_i[b,t]$ 值被設定為該第二最主要聲道的值。作為另一近似值，對於所有聲道該參數 $\alpha_i[b,t]$ 可被設定為同樣的值。在該

10 種情形下，該主要聲道和該第二最主要聲道之間的標準化相關性的平方根(square root)可被使用。

在頻域中該等解相關濾波器的實施

具有適當分析和合成視窗選擇的一重疊DFT可被應用以有效實施本發明的各方面。第4圖描述了一合適的

15 (suitable)分析/合成視窗對的一範例。第4圖顯示了重疊DFT的分析和合成視窗，該等視窗用於在頻域中運用解相關。重疊錐形(tapered)視窗被需要以在重建的信號中最小化人工因素(artifact)。

該分析視窗被設計使得對於選擇的重疊間距(overlap spacing)該等重疊分析視窗之和等於一(unity)。例如，可選擇一凱撒貝索衍生(Kaiser-Bessel-Derived, KBD)視窗的平方。根據該分析視窗，如果對該等重疊DFT沒有做出任何修改，則可在沒有合成視窗的情形下理想地(perfectly)合成一分析的信號。為了透過頻域內的乘法實現與該等解相關

20

濾波器的卷積，該分析視窗也必須被補零(zero-padded)。如果沒有補零的話，圓形卷積(circular convolution)而不是正常卷積(normal convolution)發生。如果最大的解相關濾波器長度假設是 L_{max} ，則在該分析視窗至少 L_{max} 後的補零是需要的。然而，該等聲道間振幅及時間和相位差異也被運用在頻域，且該等修改導致該分析視窗之前和之後的卷積洩漏(convolutional leakage)。因此，額外的補零被加到該分析視窗的主瓣(main lobe)的之前和之後。最後，橫過該分析視窗之主瓣和該 L_{max} 長度之補零是一(unity)的一合成視窗被使用。然而，在該區域外，為了消除該合成的音訊中的波形干擾(glitch)，該合成視窗衰減至零。本發明的各方面包括該等分析/合成視窗的組態和補零的使用。

一組合適的視窗參數列於以下：

	DFT長度：	2048
15	分析視窗主瓣長度(AWML)：	1024
	跳躍尺寸(hop size, HS)：	512
	超前補零(leading zero-pad, ZPlead)：	256
	滯後補零(lagging zero-pad, ZPlag)：	768
	合成視窗衰減(SWT)：	128
20	L_{max} ：	640

儘管該等視窗參數已被認為是合適的，但該等特定值對本發明而言並不是臨界值(critical)。

假設 $Z_i[k, t]$ 是信號 z_i 在面元(bin) k 和時間塊(time block) t 上的重疊DFT，且 $H_i[k]$ 是解相關濾波器 h_i 的DFT，則信號 \bar{z}_i

的重疊DFT可由以下方程式計算出：

$$\bar{Z}_i[k,t] = H_i[k]Z_i[k,t], \quad (11)$$

其中利用該論述的分析視窗， $Z_i[k,t]$ 已由該等向下混合信號
 $y_j(j=1\dots M)$ 的重疊DFT被計算出。假設 k_{bBegin} 和 k_{bEnd} 是與頻帶
 5 b 有關的開始(beginning)和結束(ending)面元索引，方程式(2)
 可被實施為：

$$\hat{X}_i[k,t] = \alpha[b,t]Z_i[k,t] + \beta[b,t]H_i[k]Z_i[k,t], \quad k_{bBegin} \leq k \leq k_{bEnd} \quad (12)$$

然後，藉由在每一區塊上實現反轉DFT及利用上述的合成
 視窗重疊和相加產生的(resulting)時域段(segment)，該等信
 10 號 \hat{x}_i 從 $\hat{X}_i[k,t]$ 中被合成。

參考第2圖，其中實施本發明各方面之編碼器的一簡化
 範例被顯示，該等輸入信號 x_i (複數聲音輸入信號如PCM信
 號，個別類比聲音信號的時間採樣，從1到n)被運用到個別
 時域到頻域的轉換器(converter)或轉換函數(“T/F”)22。為了
 15 描述簡便，僅顯示了一個T/F方塊，需要理解的是該等從1
 到N的輸入信號中的每一個對應一個T/F方塊。例如，該等
 輸入聲音信號可表示空間方向如左、中央、右等等。每一
 T/F可被實施，例如，藉由將該等輸入音訊採樣分成區塊、
 視窗化該等區塊、重疊該等區塊、轉換每一視窗化的和重
 20 疊的區塊到頻域，及將產生的頻譜分割為頻帶，該轉換是
 藉由計算一離散頻率轉換(DFT)，該等頻帶模擬聽覺臨界頻
 帶，例如，使用如等效矩形頻帶(equivalent-rectangular
 band, ERB)尺度的二十一頻帶。該種DFT方法在先前技術
 中是眾所周知的。其他時域到頻域的轉換參數及技術也可

被使用。對本發明而言，既不是該等特定參數也不是該特定技術是有決定性的(critical)。然而，為了簡單說明的目的，此處的描述假設該DFT轉換技術被使用。

T/F 22的頻域輸出每一個是一組頻譜係數。所有該等組
5 可被運用在一向下混合器或向下混合函數(“向下混合”)24
中。該向下混合器或向下混合函數可以是如在引用的空間
編碼刊物(publication)的不同個中所描述的，或是如在以上
引用的Davis等人的國際專利申請案中所描述的。該向下混
10 合24的輸出(在該等引用的空間編碼系統情形下是一單個
聲道yj，或在該引用的Davis等人文獻中是多聲道yj)可利用
任何合適的編碼方式如AAC、AC-3等來感知(perceptually)
編碼。闡明合適的感知編碼(perceptual coding)系統之細節
的刊物被包括在下文標題“合併參考(incorporation by
15 10 reference)”之下。該向下混合24的該(等)輸出(不管是否是感
知編碼)可被定性為“聲音資訊(audio information)”。該聲音
資訊可由一頻域到時域的轉換器或轉換函數(“F/T”)26轉換
回該時域，每一F/T通常實現一上述T/F的反轉函數(inverse
function)，也就是一反轉DFT，且接著執行視窗化和重疊相
20 加(overlap-add)。來自F/T 26的時域資訊被運用在一提供一
編碼的位元流(bitstream)輸出的位元流封包器(packer)或封
包函數(“位元流封包器”)28中。

由T/F 22產生的該等頻譜係數組也被運用於一空間參
數計算器(calculator)或計算函數30，該計算函數計算可包含
“空間參數”的“旁側資訊”，該等空間參數如在該等引用的

空間編碼刊物的不同個中所述之聲道間振幅差異、聲道間時間或相位差異及聲道間交叉相關性。該空間參數旁側資訊被運用於可將該等空間參數包括在該位元流中的該位元流封包器28。

- 5 由T/F 22產生的該等頻譜係數組也被運用於一交叉相關因數(factor)計算器或計算函數(“計算交叉相關因數”)32，該計算函數計算如上述的該等交叉相關因數 $\alpha_i[b,t]$ 。該等交叉相關因數被運用於可將該等交叉相關因數包括在該位元流中的該位元流封包器28。該等交叉相關因數
- 10 數也可被定性為“旁側資訊”。旁側資訊是解碼聲音資訊的有用資訊。

在實際實施例中，不僅該聲音資訊，該旁側資訊和該等交叉相關因數也可能以某種最小化其等發送損失的方式被量化(quantized)或被編碼。然而，為了表達簡便的目的，

15 以及因為該等細節是眾所周知的且對理解本發明沒有幫助，所以在圖式中沒有顯示量化和分離量化(de-quantizing)。

參考第3圖，其中顯示了實施本發明各方面的一解碼器的一簡化範例，一位元流(例如，由關於第2圖中描述之類型的一編碼器產生的)被運用到提供該空間資訊旁側資訊、該交叉相關旁側資訊($\alpha_i[b,t]$)和該聲音資訊的一位元流解包器(unpacker)32。該聲音資訊被運用到可與第2圖中該等轉換器22之一相同的一時域到頻域的轉換器或轉換函數(“T/F”)34。該頻域聲音資訊被運用於一向上混合器36，該

向上混合器36在其也接收的空間參數旁側資訊的幫助下運作(operate)。該向上混合器可如在引用的空間編碼刊物中之不同個所描述的來運作，或者在該聲音資訊正在多聲道中被傳輸的情形下，可如在Davis等人之該國際申請案中所描述

5 述的來運作。該等向上混合器輸出是以上提到的複數信號 z_i 。每一向上混合信號 z_i 被運用到上述具有一特性 h_i 的一唯一解相關濾波器38。為了表達簡單，僅顯示了一單個濾波器，但需要理解的是每一向上混合信號有一個別和唯一的濾波器。該等解相關濾波器的輸出是上述複數信號 \bar{z}_i 。該

10 等交叉相關因數 $\alpha_i[b,t]$ 被運用到一乘法器40，且在該乘法器40中該等交叉相關因數 $\alpha_i[b,t]$ 分別乘上述之向上混合信號 z_i 。該等交叉相關因數 $\alpha_i[b,t]$ 也被運用到一從上述之該等交叉相關因數 $\alpha_i[b,t]$ 獲得(derive)該等交叉相關因數 $\beta_i[b,t]$ 的計算器或計算函數(“計算 $\beta_i[b,t]$ ”)42。該等交叉相關因數

15 $\beta_i[b,t]$ 被運用到乘法器44，且在該乘法器44中該等交叉相關因數 $\beta_i[b,t]$ 分別乘上述之已解相關濾波的向上混合信號 \bar{z}_i 。乘法器40和44的輸出在一加法組合器(an additive combiner)或組合函數(“+”)46中相加，以產生複數輸出信號 \hat{x}_i ，且每一輸出信號近似於一相對應的輸入信號 x_i 。

20 實施

本發明可在硬體或軟體或二者的組合中被實施(如可程式邏輯陣列)。除非另外被指定，否則被包括作為本發明之一部分的演算法不是本來(inherently)就與任何特定電腦或其他裝置有關的。尤其是，多種普遍用途的機器可與依

據此處教示而編寫的程式一起被使用，或可更方便地構成更多特定裝置(specialized apparatus)(如積體電路)以實現所要求的方法步驟。因此，本發明可在一或更多電腦程式中被實施，該等電腦程式在一或更多可程式電腦系統上執行，且每一電腦系統包含至少一處理器、至少一資料儲存系統(包括依電性和非依電性記憶體和/或儲存元件)、至少一輸入裝置或埠及至少一輸出裝置或埠。程式碼被運用到輸入資料以實現此處描述的函數(function)及產生輸出資訊。該輸出資訊以已知的方式(in known fashion)被運用到一或更多輸出設備。

每一該種程式可在任何理想的電腦語言(包括機器、組合(assembly)或高階程序(high level procedural)、邏輯或物件導向(object oriented)程式語言)中實施以與一電腦系統相通訊。在任何情形下，該語言可以是一編譯(compiled)或解釋(interpreted)語言。

每一該種電腦程式被較佳地儲存在或下載到一儲存媒體或裝置(例如固態記憶體或媒體，磁或光媒體)，該媒體或裝置可由一普遍或特定用途的可程式電腦讀取，且在該儲存媒體或裝置由該電腦系統讀取以實現此處描述的該程序時，該電腦程式用於設置和運作該電腦。本發明系統也可考慮以根據一電腦程式被設置的一電腦可讀取儲存媒體來實施，其中如此被設置的該儲存媒體引起一電腦系統在一特定和預設方式下運作以實現此處描述的函數。本發明的很多實施例已被描述。然而，需要理解的是對實施例做出

的多種修改仍沒有脫離本發明的精神和範圍。例如，此處描述的一些步驟(steps)可以是沒有順序性的，因此可以以不同於被描述的順序來被實現。

合併參考

- 5 以下專利、專利申請案和刊物(publication)藉此全部被併入參考。

AC-3

ATSC 標準(Standard) A52/A: 數位聲音壓縮標準
(Digital Audio Compression Standard (AC-3))，第A版
10 (Revision A)，先進電視系統委員會(Advanced Television Systems Committee)，2001年8月20日。該A52/A文獻可在全球資訊網(world wide web)<http://www.atsc.org/standards.html>上找到。

Steve Vernon的“AC-3編碼器的設計和實現(Design and
15 Implementation of AC-3 Coders)”於電氣和電子工程師協會之消費性電子會刊(*IEEE Trans. Consumer Electronics*)第41期第3卷，1995年8月。

Mark Davis 的“AC-3多通道編碼器(The AC-3
Multichannel Coder)”於音訊工程協會預印文獻(Audio
20 Engineering Society Preprint)3374，第95次AES會議，1993年10月。

Bosi 等人的“用於傳播和多媒體應用的高品質、低速率聲音轉換碼(High Quality, Low-Rate Audio Transform Coding for Transmission and Multimedia Applications)”於音

訊工程協會預印文獻3365，第93次AES會議，1992年10月。

美國專利號5,583,962、5,632,005、5,633,981、5,727,119
與6,021,386。

AAC

- 5 ISO/IEC JTC1/SC29的“訊息技術—極低位元率的聲音-
視訊編碼 (Information Technology –Very Low Bitrate
Audio-Visual Coding)”於 ISO/IEC IS-14496 (第3部分，聲
音)，1996年。

- 10 ISO/IEC 13818-7的“MPEG-2先進的聲音編碼(MPEG-2
Advanced Audio Coding, AAC)”，國際標準，1997年。

- 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先進的聲音編碼 (ISO/IEC
MPEG-2 Advanced Audio Coding)”於第101屆AES會議公報
15 (*Proc. of the 101st AES-Convention*)，1996年。

- 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先進的聲音編碼 (ISO/IEC
MPEG-2 Advanced Audio Coding)”於AES期刊(*Journal of*
20 *the AES*)，第45期第10卷第789-814頁，1997年十月。

Karlheinz Brandenburg的“MP3與AAC的說明(MP3
and AAC explained)”於高品質聲音編碼的第17屆國際會議
之會議公報(*Proc. of the AES 17th International Conference*
on High Quality Audio Coding)，佛羅倫斯，義大利，1999

年。

G.A. Soulodre等人的“最先進的雙通道聲音編解碼器之主觀評估 (Subjective Evaluation of State-of-the-Art Two-Channel Audio Codecs)”於聲音工程協會 (*J. Audio Eng. Soc.*)第46期第3卷的第164-177頁，1998年3月。

MPEG強度立體聲(Intensity Stereo)

美國專利號 US 5,323,396、5,539,829、5,606,618及5,621,855。

美國已公開的專利申請案 US 2001/0044713。

10 空間和參數編碼

Mark Franklin Davis在2005年2月28日提出申請的國際申請案 PCT/US2005/006359且名稱為“在由較少聲道和輔助資訊表示的多聲道中之低位元率音訊編碼和解碼 (Low Bit Rate Audio Encoding and Decoding in Which Multiple Channels Are Represented By Fewer Channels and Auxiliary Information)”。

美國已公開的專利申請案 US 2003/0026441，公開於2003年2月6日。

美國已公開的專利申請案 US 2003/0035553，公開於2003年2月20日。

美國已公開的專利申請案 US 2003/0219130 (Baumgarte & Faller)，公開於2003年11月27日。

音訊工程協會文獻5852，2003年3月

公開的國際專利申請案 WO 03/090207且公開於2003年

10月30日。

公開的國際專利申請案WO 03/090208且公開於2003年
10月30日。

公開的國際專利申請案WO 03/007656且公開於2003年
5 1月22日。

公開的國際專利申請案WO 03/090206且公開於2003年
10月30日。

Baumgarte等人且公開號為US2003/0236583 A1的美國
專利申請案“聲音信號的混合多聲道/提示編碼/解碼(Hybrid
10 Multi-Channel/Cue Coding/Decoding of Audio Signals)”此案
於2003年12月25日公開，且申請號為10/246,570。

Faller等人的“應用立體提示編碼到立體和多聲道音訊
壓縮 (Binaural Cue Coding Applied to Stereo and
Multi-Channel Audio Compression)”於音訊工程協會文獻
15 5574，第112次會議，慕尼黑，2002年5月。

Baumgarte等人的“為何立體提示編碼優於強度立體編
碼(Why Binaural Cue Coding is Better than Intensity Stereo
Coding)”於音訊工程協會文獻5575，第112次會議，慕尼黑，
2002年5月。

20 Baumgarte等人的“設計和評估立體提示編碼方案
(Design and Evaluation of Binaural Cue Coding Schemes)”於
音訊工程協會文獻5706，第113次會議，洛杉磯，2002年10
月。

Faller等人的“利用感知參數有效表示空間音訊

(Efficient Representation of Spatial Audio Using Perceptual Parameterization)”於電氣和電子工程師協會關於應用信號處理到音訊和聲學的2001年討論會，紐柏茲(New Paltz)，紐約，2001年10月，第199-202頁。

- 5 Baumgarte等人的“立體提示編碼之聽覺空間提示的評估(Estimation of Auditory Spatial Cues for Binaural Cue Coding)”於2002年ICASSP公報(Proc. ICASSP 2002)，佛羅里達州的奧蘭多市，2002年5月，第II-1801-1804頁。

Faller等人的“立體提示編碼：新的和有效的空間聲音表示法(Binaural Cue Coding：A Novel and Efficient Representation of Spatial Audio)”於2002年ICASSP公報(Proc. ICASSP 2002)，佛羅里達州的奧蘭多市，2002年5月。

Breebaart等人的“於低位元率的高品質參數空間聲音編碼(High-quality Parametric Spatial Audio Coding at Low Bitrates)”於音訊工程協會文獻6072，第116次會議，柏林，2004年5月。

Baumgarte等人的“使用可攀登的具有平衡混音的立體提示編碼的聲音編碼器之增強(Audio Coder Enhancement Using Scalable Binaural Cue Coding with Equalized Mixing)”於音訊工程協會文獻6060，第116次會議，柏林，2004年5月。

Schuijers等人的“低複雜度參數立體編碼(Low Complexity Parametric Stereo Coding)”於音訊工程協會文獻6073，第116次會議，柏林，2004年5月。

Engdegard等人的“以參數立體編碼合成環境(Synthetic Ambience in Parametric Stereo Coding)”於音訊工程協會文獻6074，第116次會議，柏林，2004年5月。

其他

- 5 Herre等人且專利號為5,812,971的美國專利“使用時間波包形狀的增強聯合立體編碼法(Enhanced Joint Stereo Coding Method Using Temporal Envelope Shaping)”於1998年9月22日取得專利。

- 10 Herre等人的“強度立體編碼(Intensity Stereo Coding)”於音訊工程協會預印文獻(Audio Engineering Society Preprint)3799，第96次會議，阿姆斯特丹，1994。

- 15 Truman等人且公開號為US2003/0187663 A1的美國專利申請案“高頻表示法的多頻帶轉換(Broadband Frequency Translation of High Frequency Regeneration)”此案於2003年10月2日公開，且申請號為10/113,858。

【圖式簡單說明】

第1a和1b圖分別是一典型先前技術的空間編碼之編碼器和解碼器的簡化方塊圖。

- 20 第2圖是實施本發明各方面的一編碼器或編碼函數的一範例的簡化功能示意方塊圖。

第3圖是實施本發明各方面的一解碼器或解碼函數的一範例的簡化功能示意方塊圖。

第4圖是適合實施本發明各方面的一分析/合成視窗對的理想化敘述。

【主要元件符號說明】

- 22…時域到頻域的轉換器或轉換函數
- 24…向下混合器或向下混合函數
- 26…頻域到時域的轉換器或轉換函數
- 28…位元流封包器或封包函數
- 30…空間參數計算器或計算函數
- 32…交叉相關因數計算器或計算函數、位元流解包器
- 34…時域到頻域的轉換器或轉換函數
- 36…向上混合器
- 38…解相關濾波器
- 40…乘法器
- 42…計算器或計算函數
- 44…乘法器
- 46…加法組合器或組合函數

101年11月23日修正本 3376 號申請案申請專利範圍修正本

12-11-23

十、申請專利範圍：

1. 一種用於處理一組 N 個聲音信號的方法，該方法包含下列步驟：以一唯一解相關濾波器特性來濾波該 N 個信號中的每一個，該特性是在時域中的一因果線性非時變特性或其於頻域中的等效特性，以及，對於每一解相關濾波器特性，以一時變和頻變方式組合其輸入和輸出信號以提供一組 N 個處理的信號，其中該組 N 個聲音信號從 M 個聲音信號中被合成，其中 M 是一或更多且 N 大於 M ，且更進一步包含向上混合該 M 個聲音信號到 N 個聲音信號。
5
2. 如申請專利範圍第1項所述之方法，其中每一唯一解相關濾波器特性被選擇，使得每一濾波器特性的輸出信號與所有該等 N 個聲音信號的相關性少於每一濾波器特性相對應的輸入信號與所有該等 N 個聲音信號的相關性，且使得每一輸出信號與隔一輸出信號的相關性少於每一濾波器特性相對應之輸入信號與該等 N 個聲音信號中隔一信號的相關性。
10
3. 如申請專利範圍第1項所述之方法，更進一步包含接收描述該 N 個合成的聲音信號中之理想空間關係的參數，且其中該向上混合在接收到的參數幫助下運作。
15
4. 如申請專利範圍第2項所述之方法，更進一步包含接收描述該 N 個合成的聲音信號中之理想空間關係的參數，且其中該向上混合在接收到的參數幫助下運作。
20
5. 如申請專利範圍第14項中任一項所述之方法，其中每一

解相關濾波器特性由一具有多自由度的模型來定性。

6. 如申請專利範圍第7項中任一項所述之方法，其中每一解相關濾波器特性具有一頻變延遲形式的響應，其中該延遲隨著頻率增加而單調減少。
- 5 7. 如申請專利範圍第1-4項中任一項所述之方法，其中每一解相關濾波器特性具有一頻變延遲形式的響應，其中該延遲隨著頻率增加而單調減少。
8. 如申請專利範圍第2項所述之方法，其中每一濾波器特性的脈衝響應被一有限期間的正弦序列指定，該序列的瞬間頻率單調減少。
- 10 9. 如申請專利範圍第8項所述之方法，其中一雜訊序列被加入到該正弦序列的瞬間相位中。
10. 如申請專利範圍第1項中任一項所述之方法，其中該組合是一線性組合。
- 15 11. 如申請專利範圍第1項中任一項所述之方法，其中該組合的組合度在接收到的參數幫助下運作。
12. 如申請專利範圍第1項中任一項所述之方法，更進一步包含接收描述該N個處理的信號間之理想空間關係的參數，且其中該組合的組合度在接收到的參數幫助下運作。
- 20 13. 如申請專利範圍第13項或第14項所述之方法，其中該等N個聲音訊號中的每一個代表聲道，且幫助該組合運作之該等接收到的參數是有關聲道間交叉相關性的參數。
14. 如申請專利範圍第15項所述之方法，其中其他接收到的

參數包括有關聲道間振幅差異和聲道間時間或相位差異中一個或更多個的參數。

15. 一種適合實現申請專利範圍1-4項中任一項之方法的裝置。

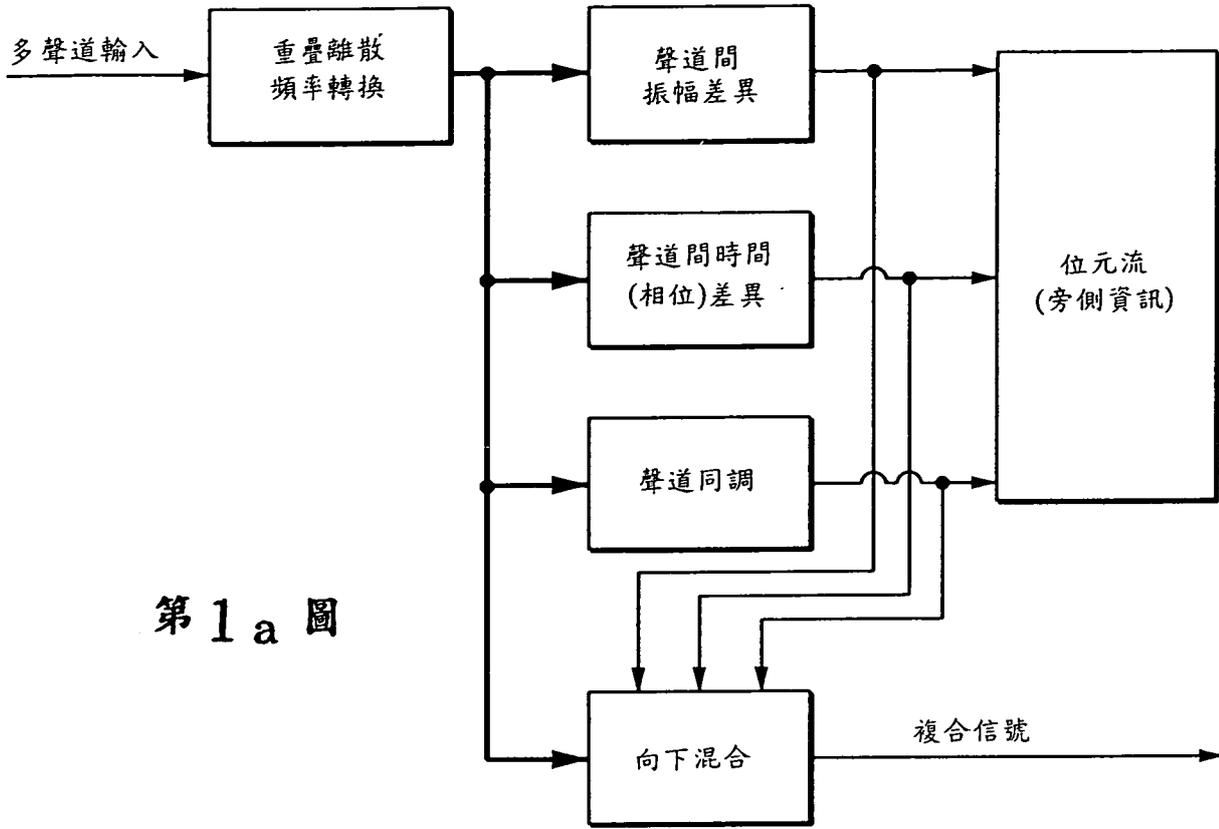
5 16. 一種儲存在一電腦可讀式媒體上的電腦程式，用於使一電腦實現申請專利範圍1-4項中任一項之方法。

17. 一種用於處理一組 N 個聲音信號的裝置，其包含：

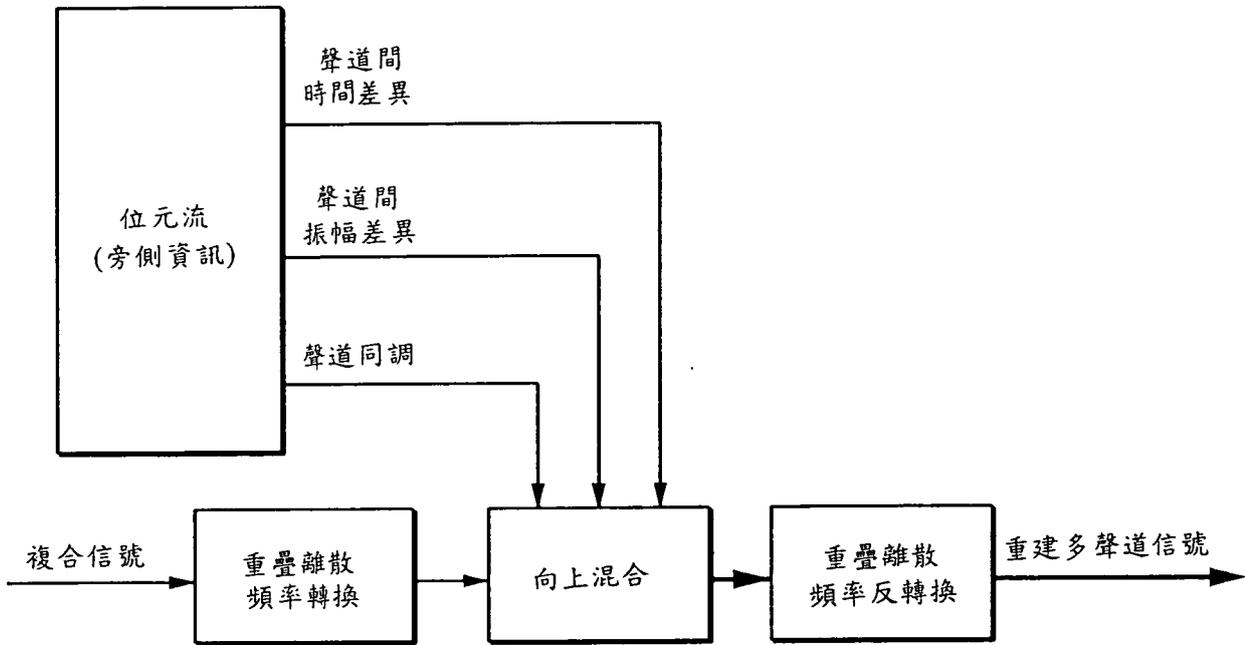
10 使該等 N 個聲音信號中的每一個以一唯一解相關濾波器特性來濾波的裝置，該特性是在時域內的一因果線性非時變特性或其在頻域中的等效特性，

對於每一解相關濾波器特性，以時變和頻變方式組合其輸入和輸出信號以提供一組 N 個處理的信號之裝置，以及

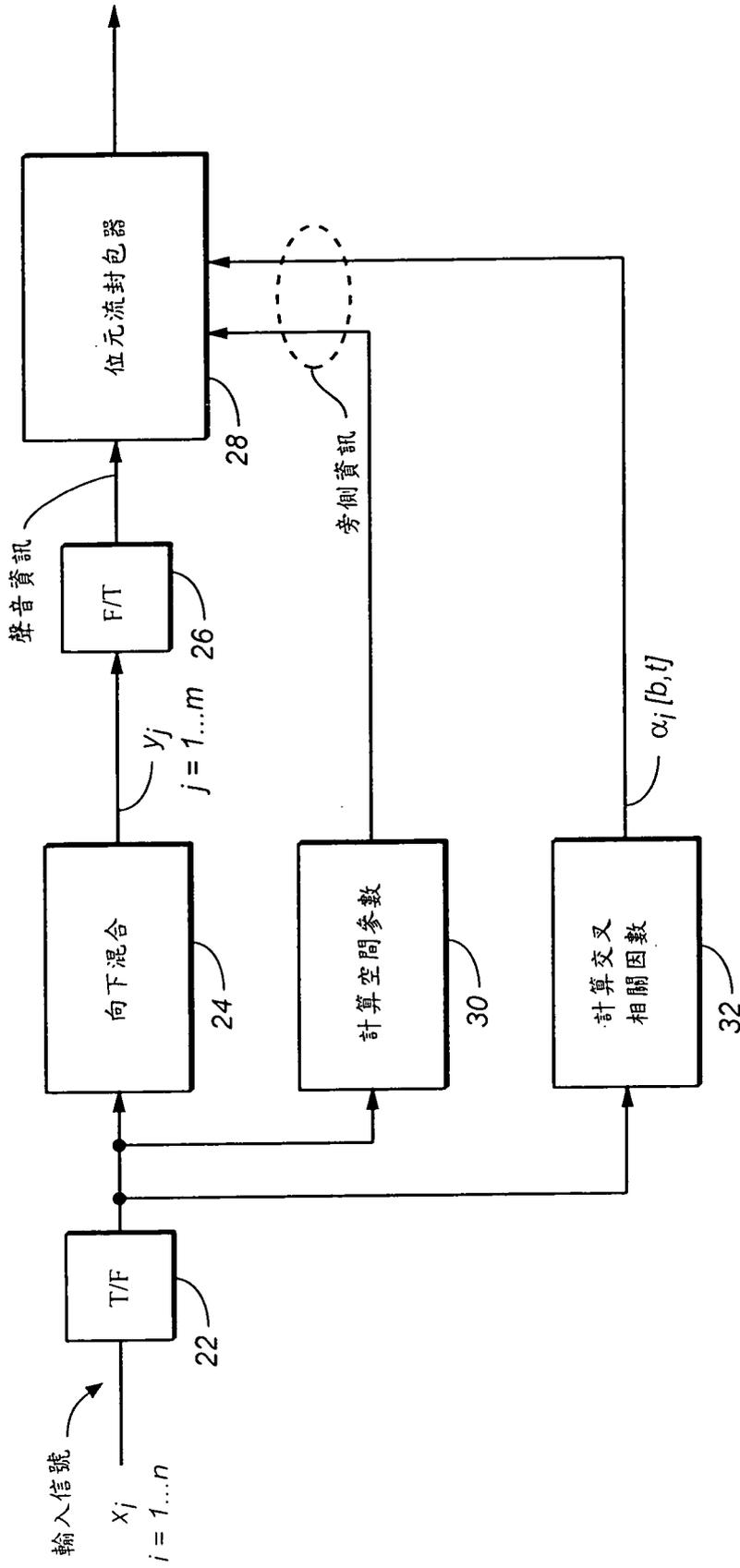
15 其中該組 N 個聲音信號從 M 個聲音信號中被合成，其中 M 是一或更多且 N 大於 M ，且更進一步包含向上混合該 M 個聲音信號到 N 個聲音信號。



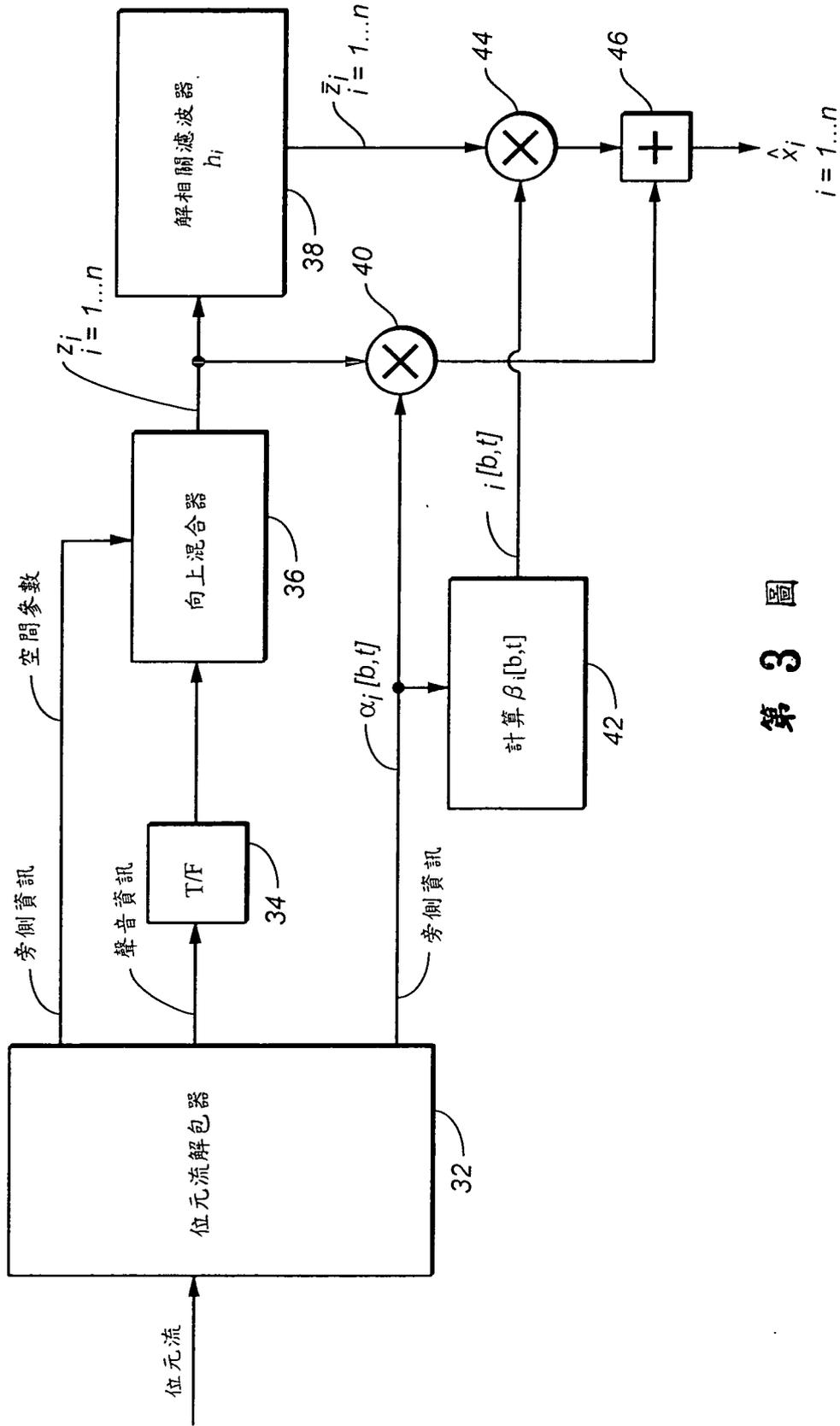
第 1 a 圖



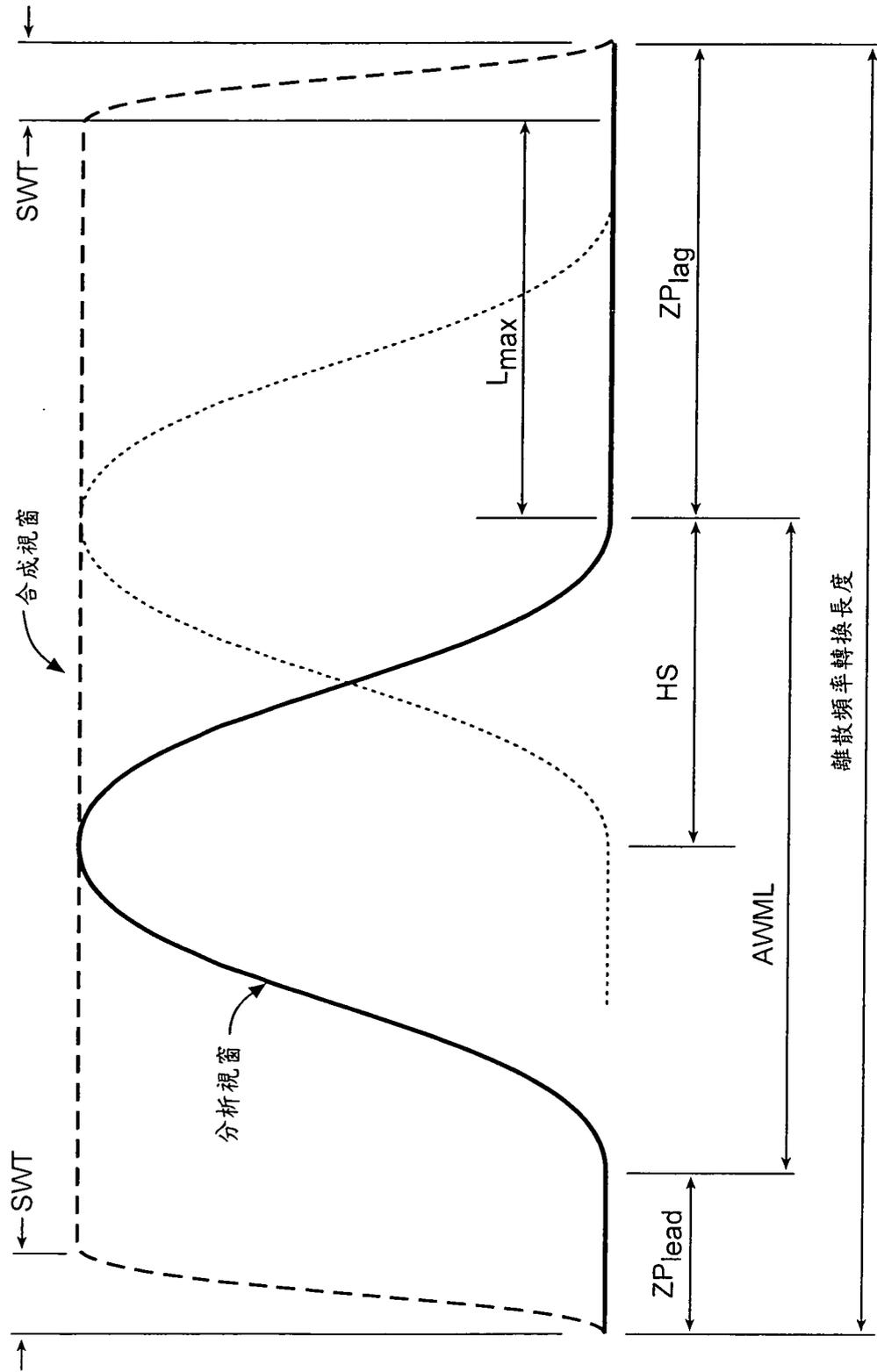
第 1 b 圖



第 2 圖



第 3 圖



第 4 圖