



(19)中華民國智慧財產局

(12)發明說明書公開本 (11)公開編號：TW 201036464 A1

(43)公開日：中華民國 99 (2010) 年 10 月 01 日

(21)申請案號：098132269

(22)申請日：中華民國 98 (2009) 年 09 月 24 日

(51)Int. Cl. : **H04S5/00 (2006.01)**

(30)優先權：2008/10/07 美國 61/103,303
2009/05/15 歐洲專利局 09006598.8

(71)申請人：弗勞恩霍夫爾協會 (德國) FRAUNHOFER-GESELLSCHAFT ZUR FORDERUNG
DER ANGEWANDTEN FORSCHUNG E. V. (DE)

德國

杜比瑞典公司 (瑞典) DOLBY SWEDEN AB (SE)
瑞典

飛利浦國際公司知識產權及標準部 (荷蘭) PHILIPS INTERNATIONAL B. V.

INTELLECTUAL PROPERTY & STANDARDS (NL)
荷蘭

(72)發明人：穆迪特 哈瑞 MUNDT, HARALD (DE)；泰倫堤夫 雷尼德 TERENTIEV, LEONID (DE)；費曲 寇尼利亞 FALCH, CORNELIA (DE)；希爾波特 喬漢娜 HILPERT, JOHANNES (DE)；賀穆斯 奧利薇 HELLMUTH, OLIVER (DE)；普洛格史提珍 PLOGSTIES, JAN (DE)；維爾摩斯 萊斯 VILLEMOES, LARS (SE)；布瑞巴特 傑洛恩 BREEBAART, JEROEN (NL)；科沛恩 傑洛恩 KOPPENS, JEROEN (NL)；艾德加德 喬納斯 ENGDEGARD, JONAS (SE)

(74)代理人：惲軼群；陳文郎

申請實體審查：有 申請專利範圍項數：11 項 圖式數：6 共 56 頁

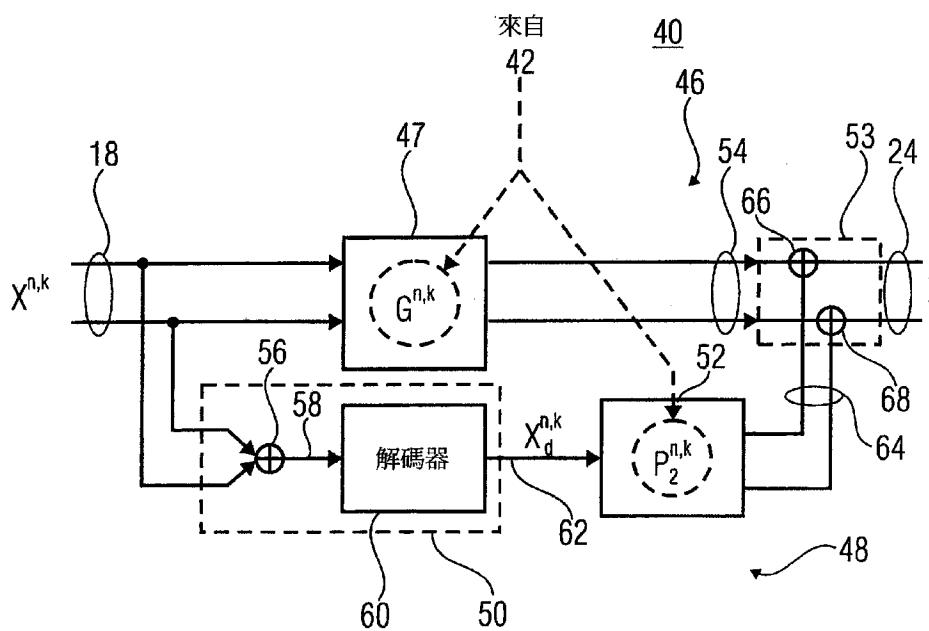
(54)名稱

多聲道音訊信號之雙耳演示技術

BINAURAL RENDERING OF A MULTI-CHANNEL AUDIO SIGNAL

(57)摘要

本文描述將一多聲道音訊信號雙耳演示為一雙耳輸出信號。該多聲道音訊信號包含多個音訊信號被降混的一立體聲降混信號及旁側資訊，且該旁測資訊包含一指示出對於每一音訊信號，該各自音訊信號已分別混合至該立體聲降混信號的第一聲道及第二聲道中的程度的降混資訊，及該等音訊信號的目標位準資訊，及描述在該等音訊信號之音訊信號對之間的相似性的目標內互相關資訊。基於第一演示指示，自該立體聲降混信號之該第一及第二聲道運算出一初步雙耳信號。產生一經去相關的信號作為對該立體聲降混信號之該第一及第二聲道的一單降混的一感知等效物，但然是與該單降混去相關。根據第二演示指示，自該經去相關信號運算出一校正雙耳信號，且該初步雙耳信號與該校正雙耳信號相混合，以獲得該雙耳輸出信號。



- 18：降混信號
- 24：雙耳輸出信號
- 40：降混預處理單元
- 42：SAOC 參數處理單元
- 46：乾式路徑
- 47：乾式演示單元
- 48：濕式路徑
- 50：去相關信號產生器
- 52：濕式演示單元
- 53：混合階段
- 54：初步雙耳輸出信號
- 56：相加器
- 58：多聲道降混
- 60：去相關器
- 62：經去相關信號
- 64：校正雙耳輸出信號
- 66：相加器
- 68：相加器



(19)中華民國智慧財產局

(12)發明說明書公開本 (11)公開編號：TW 201036464 A1

(43)公開日：中華民國 99 (2010) 年 10 月 01 日

(21)申請案號：098132269

(22)申請日：中華民國 98 (2009) 年 09 月 24 日

(51)Int. Cl. : **H04S5/00 (2006.01)**

(30)優先權：2008/10/07 美國 61/103,303
2009/05/15 歐洲專利局 09006598.8

(71)申請人：弗勞恩霍夫爾協會 (德國) FRAUNHOFER-GESELLSCHAFT ZUR FORDERUNG
DER ANGEWANDTEN FORSCHUNG E. V. (DE)

德國

杜比瑞典公司 (瑞典) DOLBY SWEDEN AB (SE)
瑞典

飛利浦國際公司知識產權及標準部 (荷蘭) PHILIPS INTERNATIONAL B. V.

INTELLECTUAL PROPERTY & STANDARDS (NL)
荷蘭

(72)發明人：穆迪特 哈瑞 MUNDT, HARALD (DE)；泰倫堤夫 雷尼德 TERENTIEV, LEONID (DE)；費曲 寇尼利亞 FALCH, CORNELIA (DE)；希爾波特 喬漢娜 HILPERT, JOHANNES (DE)；賀穆斯 奧利薇 HELLMUTH, OLIVER (DE)；普洛格史提珍 PLOGSTIES, JAN (DE)；維爾摩斯 萊斯 VILLEMOES, LARS (SE)；布瑞巴特 傑洛恩 BREEBAART, JEROEN (NL)；科沛恩 傑洛恩 KOPPENS, JEROEN (NL)；艾德加德 喬納斯 ENGDEGARD, JONAS (SE)

(74)代理人：惲軼群；陳文郎

申請實體審查：有 申請專利範圍項數：11 項 圖式數：6 共 56 頁

(54)名稱

多聲道音訊信號之雙耳演示技術

BINAURAL RENDERING OF A MULTI-CHANNEL AUDIO SIGNAL

(57)摘要

本文描述將一多聲道音訊信號雙耳演示為一雙耳輸出信號。該多聲道音訊信號包含多個音訊信號被降混的一立體聲降混信號及旁側資訊，且該旁測資訊包含一指示出對於每一音訊信號，該各自音訊信號已分別混合至該立體聲降混信號的第一聲道及第二聲道中的程度的降混資訊，及該等音訊信號的目標位準資訊，及描述在該等音訊信號之音訊信號對之間的相似性的目標內互相關資訊。基於第一演示指示，自該立體聲降混信號之該第一及第二聲道運算出一初步雙耳信號。產生一經去相關的信號作為對該立體聲降混信號之該第一及第二聲道的一單降混的一感知等效物，但然而是與該單降混去相關。根據第二演示指示，自該經去相關信號運算出一校正雙耳信號，且該初步雙耳信號與該校正雙耳信號相混合，以獲得該雙耳輸出信號。

六、發明說明：

描述

【發明所屬之技術領域】

本應用有關於多聲道音訊信號的雙耳演示。

【先前技術】

許多音訊編碼演算法已遭提出，以有效地編碼或壓縮單一聲道的音訊資料，即單音訊信號。使用心理聲學，音訊樣本予以適當地調節、量化或甚至設為零，以將不相關性從例如 PCM 經編碼音訊信號中移除。也執行冗餘的移除。

更進一步地，在立體聲音訊信號的左聲道與右聲道之間的類似性已予以使用，以有效地編碼/壓縮立體聲音訊信號。

然而，即將的應用引起對音訊編碼演算法的進一步需求。例如，在電信會議、電腦遊戲、音樂性能等中，部分地或甚至完全不相關聯的多個音訊信號必須並行地予以發送。為了保持用以編碼此等音訊信號所需要的位元率足夠低，以與低位元率的發送應用相容，近來已提出將多個輸入音訊信號降混為一降混信號(諸如一立體聲或甚至單降混信號)的音訊編解碼器。例如，MPEG 環繞標準以該標準所指示的方式將該等輸入聲道降混為降混信號。該降混藉由所謂的 OTT⁻¹ 及 TTT⁻¹ 方塊的使用來執行，該等 OTT⁻¹ 及 TTT⁻¹ 方塊分別用以將二個信號降混為一個信號且將三個信號降混為二個信號。為了降混多於三個的信號，此等

方塊的一階層結構予以使用。除了輸出單降混信號，每一 OTT^{-1} 方塊輸出在二個輸入聲道之間的聲道位準差、及表示在二個輸入聲道之間的相干性或互相關性的聲道內相干性參數/互相關性參數。該等參數與 MPEG 環繞資料流中的 MPEG 環繞編碼器的降混信號一起輸出。類似地，每一 TTT^{-1} 方塊發送能夠從產生的立體聲降混信號中恢復該等三個輸入聲道的聲道預測係數。該等聲道預測係數也作為 MPEG 環繞資料流中的旁側資訊予以發送。該 MPEG 環繞解碼器藉由經發送的旁側資訊的使用升混該降混信號，且恢復輸入至該 MPEG 環繞編碼器中的原始聲道。

然而，不幸的是，MPEG 環繞不能滿足許多應用的所有需要。例如，該 MPEG 環繞解碼器專用於升混該 MPEG 環繞編碼器的降混信號，使得 MPEG 環繞編碼器的輸入聲道恢復成原先的樣子。換句話說，該 MPEG 環繞資料流專用於藉由已用以編碼的揚聲器組態的使用，或藉由像立體聲的典型組態來播放。

然而，根據一些應用，如果揚聲器的組態可在解碼器端自由地改變，將是有利的。

為了處理後者的需要，當前設計了空間音訊目標編碼(SAOC)標準。每一聲道作為一個別的目標來對待，且將所有的目標降混為一降混信號。也就是說，該等目標以彼此獨立，不依附於任何特定的揚聲器組態，但能夠任意地將(虛擬的)揚聲器定位於解碼器端的音訊信號來處理。該等個別的目標可包含個別的聲源，例如樂器或聲道。不同於

MPEG 環繞解碼器，SAOC 解碼器可自由地個別升混該降混信號，以在任何揚聲器組態上重播該等個別的目標。為了使 SAOC 解碼器能夠恢復已編碼於 SAOC 資料流中的個別目標，目標位準差和對於一起形成一立體聲(或多聲道)信號之目標的目標內互相關參數作為 SAOC 位元流中的旁側資訊予以發送。除此之外，SAOC 解碼器/轉碼器被提供揭示如何將個別目標降混為降混信號的資訊。因而，在解碼器端，藉由使用受使用者控制的演示資訊來恢復該等個別 SAOC 聲道，且在任何揚聲器組態上演示此等信號，是可能的。

然而，雖然上述的編解碼器(即 MPEG 環繞及 SAOC)能夠在具有多於二個揚聲器的揚聲器組態上發送及演示多聲道音訊內容，但是以耳機作為音訊再生系統的需求日益增加，迫使此等編解碼器也必須能夠在耳機上演示音訊內容。對比於揚聲器的播放，藉由耳機所再現的立體聲音訊內容在其頭部內側予以感知。在某些實體位置處，不存在從聲源至耳膜的聲學路徑的影響，致使空間影像因為判定一聲音源的所感知的方位、高度及距離的提示本質上缺失了或極其不準確，而聽起來不自然。因而，為了解決在耳機上由於不準確或缺少聲源定位提示所導致的不自然的聲音階段，已經提出各種技術來模擬一虛擬的揚聲器裝備。概念是將聲源定位的提示加至每一揚聲器信號上。如果空間聲學特性包括於此等量測資料中，那麼這透過由所謂的頭部相關轉換函數(HRTF)或雙耳空間脈衝響應(BRIR)來過

濾音訊信號而獲實現。然而，由上述的函數來過濾每一揚聲器信號將使在解碼器/再生端，必須有一顯著較高量的運算能力。特別的是，在“虛擬”揚聲器位置上演示多聲道音訊信號必須予以首先執行，接著，其中所獲得的每一揚聲器信號從而由各自的轉換函數或脈衝響應來過濾，以獲得雙耳輸出信號的左聲道及右聲道。更糟糕的是：由於為了實現虛擬揚聲器信號，一相當大量的合成去相關信號將必須混合至該等升混信號中，以補償在原始不相關音訊輸入信號之間的相關性(該相關性由將該等音訊輸入信號降混為降混信號而產生)，所獲得的雙耳輸出信號從而將具有一差的音訊品質。

在目前的 SAOC 編解碼器版本中，旁側資訊內的 SAOC 參數允許使用者使用原則上包括耳機的任何播放裝備，來交互地空間演示音訊目標。對耳機的雙耳演示允許使用頭部相關轉換函數(HRTF)參數，來在 3D 空間中空間控制虛擬的目標位置。例如，在 SAOC 中的雙耳演示可藉由限制此種情況為單降混的 SAOC 情況(其中將輸入信號均等地混合至單聲道中)，而予以實現。不幸的是，單降混迫使所有音訊信號必須混合為一共同的單降混信號，使得最大程度地失去在原始音訊信號之間的原始相關性特性，因而雙耳演示輸出信號的演示品質不是最佳的。

因而，本發明的目的是提供用以雙耳演示一多聲道音訊信號的一方案，使得雙耳演示的結果獲得改良，同時避免對由原始音訊信號組成降混信號的自由度的限制。

此目的藉由根據申請專利範圍第 1 項所述之裝置及根據申請專利範圍第 10 項所述之方法來實現。

【發明內容】

本發明以下的基本觀點之一是，自一立體聲降混信號開始雙耳演示一多聲道音訊信號，較自一單降混音訊信號開始雙耳演示多聲道音訊信號更加有利，因為由於極少的目標存在於立體聲降混信號中的事實，在個別音訊信號之間的去相關量被更佳地保存，且因為在編碼器端在立體聲降混信號的二個聲道之間選擇的可能性，使不同降混聲道中的音訊信號之間的相關性特性能夠予以部分地保存。換句話說，由於編碼器的降混，目標內相干性被退化，這在解碼端必須考量，在解碼端雙耳輸出信號的聲道內相干性對於虛擬聲源寬度的感知是一重要的測量，而使用立體聲降混代替單降混降低了退化量，使得藉由雙耳演示立體聲降混信號來恢復/產生適當量的聲道內相干性，能實現更佳的品質。

本申請案的另一主要觀點是，前述 ICC(ICC=聲道內相干性)控制可藉由一去相關信號來實現，該去相關信號形成對立體聲降混信號之降混聲道的一單降混的一感知等效物，然而是與該單降混去相關。因而，立體聲降混信號代替一單降混信號的使用保存了該等音訊信號的一些相關性特性，而這些特性在使用一單降混信號時會失去，雙耳演示可基於表示第一及第二降混聲道二者的一經去相關信號，從而與單獨地去相關每一立體聲降混聲道相比，減少

了去相關或合成信號處理量。

圖式簡單說明

參照圖式，其等更詳細地描述本申請案的較佳實施例，其中該等圖式為：

第 1 圖顯示可供本發明之該等實施例實施的一 SOAC 編碼器/解碼器安排的一方塊圖；

第 2 圖顯示一單音訊信號的一頻譜表示的一示意及說明圖；

第 3 圖顯示根據本發明之一實施例之能夠雙耳演示的一音訊解碼器的一方塊圖；

第 4 圖顯示根據本發明之一實施例之第 3 圖的降混預處理方塊的一方塊圖；

第 5 圖顯示根據第一替代方式，由第 3 圖之 SAOC 參數處理單元 42 所執行的步驟的一流程圖；

第 6 圖顯示說明該等收聽測試結果的一圖形。

【實施方式】

在以下更詳細地描述本發明之實施例前，先說明 SAOC 編解碼器及一 SAOC 位元流中所發送的 SAOC 參數，以使能夠更容易理解下面所更詳細描述的特定實施例。

第 1 圖顯示一 SAOC 編碼器 10 及一 SAOC 解碼器 12 的一大致安排。該 SAOC 編碼器 10 接收作為輸入的 N 個目標，即音訊信號 14_1 至 14_N 。特別的是，編碼器 10 包含一降混器 16，該降混器 16 接收該等降混信號 14_1 至 14_N 且將其等降混為一降混信號 18。在第 1 圖中，該降混信號示

範地顯示為一立體聲降混信號。然而，該編碼器 10 及解碼器 12 也可能以一單模式來操作，在這種情況下，該降混信號將是一單降混信號。然而，下面的描述專注於立體聲降混的情況。立體聲降混信號 18 的聲道被表示為 LO 及 RO。

為了使 SAOC 解碼器 12 能夠恢復個別目標 14_1 至 14_N ，降混器 16 向 SAOC 解碼器 12 提供包括 SAOC 參數的旁側資訊，該等 SAOC 參數包括目標位準差(OLD)、目標內互相關參數(IOC)、降混增益值(DMG)及降混聲道位準差(DCLD)。包括該等 SAOC 參數的旁側資訊 20，與該降混信號 18 一起形成由 SAOC 解碼器 12 所接收的 SAOC 輸出資料流 21。

該 SAOC 解碼器 12 包含接收降混信號 18 及旁側資訊 20 的一升混器 22，以藉由輸入至 SAOC 解碼器 12 的演示資訊 26 及 HRTF 參數 27 所指定的演示，來在任何使用者所選定的聲道組 24_1 至 24_M 上恢復及演示該等音訊信號 14_1 及 14_N ，其意思在下面予以更詳細地描述。下面的描述專注於雙耳演示，其中 $M'=2$ ，且輸出信號特別地專用於耳機的再現，儘管解碼器 12 也能夠根據使用者輸入 26 中的指令而在其他(非雙耳)揚聲器組態上演示。

該等音訊信號 14_1 至 14_N 可以任何編碼域(例如以時域或頻譜域)輸入至降混器 16 中。在實例中，音訊信號 14_1 至 14_N 以時域(諸如 PCM 編碼)輸入至降混器 16 中，降混器 16 使用諸如一混合 QMF 組的一濾波器組，例如具有對於最低頻帶尼奎斯特濾波器擴展以增加其頻率解析度的一組

複指數調變濾波器，以將該等信號轉換至頻譜域中，其中該等音訊信號在一特定的濾波器組解析度下，表現於與不同頻譜部分相關聯的多個子帶中。如果該等音訊信號 14_1 至 14_N 已在降混器 16 所期望的表示中，那麼同樣地不必執行頻譜分解。

第 2 圖顯示在上述的頻譜域中的一音訊信號。如所見的，音訊信號表示為多個子帶信號。每一子帶信號 30_1 至 30_P 由一序列的子帶值組成，該序列子帶值由小方框 32 指出。如所見的，該等子帶信號 30_1 至 30_P 的子帶值 32 於時間上互相同步，使得對於每一個連續濾波器組的時槽 34，每一子帶 30_1 至 30_P 恰好包含一子帶值 32。如頻率軸 35 所繪示，該等子帶信號 30_1 至 30_P 與不同的頻率區域相關聯，且如時間軸 37 所繪示，該等濾波器組的時槽 34 於時間中連續配置。

如上所述，降混器 16 運算來自輸入音訊信號 14_1 至 14_N 的 SAOC 參數。降混器 16 以一時間/頻率解析度來執行此運算，該時間/頻率解析度可相對於由濾波器組的時槽 34 及子帶分解所判定之原始的時間/頻率解析度而降低某一量，其中此量可藉由各自的語法元素 `bsFrameLength` 及 `bsFreqRes`，在旁側資訊 20 中發信至解碼器側。例如，連續濾波器組的時槽 34 的群組可分別形成一音框 36。換句話說，音訊信號可分割為例如在時間中交疊或在時間中相鄰的音框。在這種情況下，`bsFrameLength` 可定義每個音框之時槽 38 參數的數目，即供諸如 OLD 及 IOC 之 SAOC 參數

於一 SAOC 音框 36 中被運算的時間單元，且 bsFreqRes 可定義 SAOC 參數被運算的處理頻帶的數目，即頻域被細分割且該等 SAOC 參數被判定及發送之頻帶的數目。藉由此方式，每一音框分割為在第 2 圖中由虛線所示範表示的時間/頻率瓦片 39。

該降混器 16 根據下面的公式計算 SAOC 參數。特別的是，降混器 16 對每一目標 i 運算目標位準差，為

$$OLD_i = \frac{\sum_{n} \sum_{k \in m} x_i^{n,k} x_i^{n,k*}}{\max_j \left(\sum_n \sum_{k \in m} x_j^{n,k} x_j^{n,k*} \right)}$$

其中和及指數 n 及 k 分別貫穿所有濾波器組的時槽 34，及屬於某一時間/頻率瓦片 39 之所有濾波器組的子帶 30。因而，一音訊信號或目標 i 之所有子帶值 x_i 的能量被加總，且以所有目標或音訊信號中的瓦片最高能量值正規化。

而且，SAOC 降混器 16 能夠運算不同輸入目標 14_1 至 14_N 對之相對應時間/頻率瓦片的一相似性測量。雖然 SAOC 降混器 16 可運算在所有的輸入目標 14_1 至 14_N 對之間的相似性測量，但是降混器 16 也可抑制相似性測量的發信或限制相似性測量的運算為形成一共同立體聲聲道的左聲道或右聲道的音訊目標 14_1 至 14_N 。在任何情況下，該相似性測量被稱為目標內互相關參數 $IOC_{i,j}$ 。該運算如下

$$IOC_{i,j} = IOC_{j,i} = \operatorname{Re} \left\{ \frac{\sum_{n} \sum_{k \in m} x_i^{n,k} x_j^{n,k*}}{\sqrt{\sum_n \sum_{k \in m} x_i^{n,k} x_i^{n,k*} \sum_n \sum_{k \in m} x_j^{n,k} x_j^{n,k*}}} \right\}$$

其中增益指數 n 及 k 貫穿屬於某一時間/頻率瓦片 39 的所有子帶值，且 i 及 j 表示音訊目標 14_1 至 14_N 的某一對。

降混器 16 藉由用於每一目標 14_1 至 14_N 之增益因素的使用，降混該等目標 14_1 至 14_N 。

在一立體降混信號的情況(此情況在第 1 圖中予以示範地表示)下，一增益因素 $D_{1,i}$ 用於目標 i，且接著對所有經此等增益放大的目標計算總和，以獲得左降混聲道 L0，且增益因素 $D_{2,i}$ 用於目標 i，且接著對該等經增益放大的目標計算總和，以獲得右降混聲道 R0。因而，因數 $D_{1,i}$ 及 $D_{2,i}$ 形成大小為 $2 \times N$ 的一降混矩陣 D，其中

$$D = \begin{pmatrix} D_{1,1} & \cdots & D_{1,N} \\ D_{2,1} & \cdots & D_{2,N} \end{pmatrix} \text{ 及 } \begin{pmatrix} L_0 \\ R_0 \end{pmatrix} = D \cdot \begin{pmatrix} Obj_1 \\ \vdots \\ Obj_N \end{pmatrix}.$$

此降混指示藉由降混增益 DMG_i 發信至解碼器側，且在一立體聲降混信號的情況下，藉由降混聲道位準差 $DCLD_i$ 而發信至解碼器側。

該等降混增益予以計算，根據：

$$DMG_i = 10 \log_{10} (D_{1,i}^2 + D_{2,i}^2 + \epsilon)$$

其中 ϵ 是低於最大信號輸入之諸如 10^{-9} 或 96dB 的一小數目。

對於 $DCLD_s$ 使用下面的公式：

$$DCLD_1 = 10 \log_{10} \left(\frac{D_{1,i}^2}{D_{2,i}^2} \right).$$

降混器 16 產生立體聲降混信號，根據：

$$\begin{pmatrix} LO \\ RO \end{pmatrix} = \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \cdot \begin{pmatrix} Obj_1 \\ \vdots \\ Obj_N \end{pmatrix}$$

因而，在上述的公式中，參數 OLD 及 IOC 是該等音訊信號的一函數，且參數 DMG 及 DCLD 是 D 的一函數。同時，應注意的是 D 可在時間中變化。

在雙耳演示(在此所描述的解碼器操作模式)的情況下，輸出信號自然地包含二個聲道，即 $M'=2$ 。然而，上述的演示資訊 26 指示的是如何將該等輸入信號 14_1 至 14_N 分散至虛擬的揚聲器位置 1 至 M 上，其中 M 可高於 2。因而，該演示資訊可包含指示如何將該等輸入目標 obj_i 分散至虛擬的揚聲器位置 j 上，以獲得虛擬揚聲器信號 vs_j 的一演示矩陣 M ，其中 j 在 1 與 M 之間，且 i 在 1 與 N 之間，其中

$$\begin{pmatrix} vs_1 \\ \vdots \\ vs_M \end{pmatrix} = M \cdot \begin{pmatrix} Obj_1 \\ \vdots \\ Obj_N \end{pmatrix}$$

該演示資訊可以任何方式由使用者提供或輸入。更有可能的是，演示資訊 26 包含於 SAOC 流 21 自身的旁側資訊中。當然，可允許該演示資訊隨時間變化。例如，時間解析度可等於音框解析度，即可為每一音框 36 來定義 M 。即使頻率上的 M 變化也是可能的。例如，可為每一瓦片 39 來定義 M 。下面，例如 $M_{ren}^{l,m}$ 將用於表示 M ，其中 m 表示頻帶且 l 表示參數時間片段 38。

最後，在下面中，將提及 HRTF 27。此等 HRTF 描述如何將一虛擬揚聲器信號 j 分別在左耳及右耳上顯示，使得雙耳提示獲得保存。換句話說，對於每一虛擬揚聲器位

置 j ，存在二個 HRTF，即一個對應於左耳，且另一個對應於右耳。如下面更詳細的描述，可能的是，解碼器被提供 HRTF 參數 27，該等 HRTF 參數 27 包含對於每一虛擬揚聲器位置 j ，描述在由雙耳所接收的信號之間且來自於同一聲源 j 的一相移偏移量 Φ_j ，及分別對應於右耳及左耳，描述由於收聽者的頭部而產生雙耳衰減的二個振幅放大/衰減 $P_{i,R}$ 及 $P_{i,L}$ 。該 HRTF 參數 27 可是關於時間的常數，而在可能等於該 SAOC 參數解析度的某一頻率解析度(即每個頻帶)下來定義。在下面中，HRTF 參數以 Φ_j^m 、 $P_{j,R}^m$ 及 $P_{j,L}^m$ 所給定，其中 m 表示頻帶。

第 3 圖更詳細地顯示第 1 圖中的 SAOC 解碼器 12。如所示，解碼器 12 包含一降混預處理單元 40 及一 SAOC 參數處理單元 42。該降混預處理單元 40 受組配用以接收該立體聲降混信號 18，且將其轉換為雙耳輸出信號 24。該降混預處理單元 40 以一由 SAOC 參數處理單元 42 所控制的方式來執行此轉換。特別的是，該 SAOC 參數處理單元 42 向降混預處理單元 40 提供一演示指示資訊 44，該演示指示資訊 44 是該 SAOC 參數處理單元 42 由 SAOC 旁側資訊 20 及演示資訊 26 推導出的。

第 4 圖更詳細地顯示根據本發明之一實施例的降混預處理單元 40。特別的是，根據第 4 圖，該降混預處理單元 40 包含並行連接於輸入(此處接收立體聲降混信號 18，即 $X^{n,k}$)與單元 40 之一輸出(此處輸出雙耳輸出信號 $\hat{X}^{n,k}$)之間的二個路徑，即稱為乾式路徑 46(供一乾式演示單元串列連

接)的一路徑及一濕式路徑 48(供一去相關信號產生器 50 及一濕式演示單元 52 串列連接)，其中一混合階段 53 將二路徑 46 及 48 的輸出相混合以獲得最終的結果，即雙耳輸出信號 24。

如下面將更詳細的描述，該乾式演示單元 47 受組配以由立體聲降混信號 18 運算出一初步雙耳輸出信號 54，其中該初步雙耳輸出信號 54 表示該乾式演示路徑 46 的輸出。該乾式演示單元 47 基於由該 SAOC 參數處理單元 42 所提供的一乾式演示指定來執行其運算。在下面所描述的特定實施例中，該演示指定由一乾式演示矩陣 $G^{n,k}$ 來定義。上述的提供在第 4 圖中藉由一虛線箭頭來說明。

該去相關信號產生器 50 受組配以透過降混由該立體聲降混信號 18 產生一經去相關信號 $X_d^{n,k}$ ，使得其是對該立體聲降混信號 18 之右及左聲道的單降混的一感知等效物，然而是對單降混去相關。如第 4 圖所示，該經去相關產生器 50 可包含一相加器 56，其用以在例如比率 1:1 下或在例如某一其他的固定比率下，對該立體聲降混信號 18 的左及右聲道求和，以獲得各自的單降混 58，該相加器 56 之後是一去相關器 60，其用以產生前述的經去相關信號 $X_d^{n,k}$ 。該去相關器 60 可例如包含一或多個延遲級，以由經延遲版本或該單降混 58 之經延遲版本的一加權和或甚至關於該單降混 58 與單降混之一個(多個)經延遲版本的一加權和，形成該經去相關信號 $X_d^{n,k}$ 。當然，對於去相關器 60 存在許多的替代方式。實際上，分別由去相關器 60 及去相關信號產生

器 50 所執行的去相關趨於降低在藉由上述相對應於目標內互相關的公式測量時，該經去相關信號 62 與該單降混 58 之間的聲道內相干性，以在對於目標位準差藉由上述公式來測量時實質上維持其等的目標位準差。

該濕式演示單元 52 受組配以由該經去相關信號 62 運算出一校正雙耳輸出信號 64，從而所獲得之校正的雙耳輸出信號 64 表示該濕式演示路徑 48 的輸出。該濕式演示單元 52 使其運算基於一濕式演示指示，該濕式演示指示依據由乾式演示單元 47 所使用的乾式演示指示而定，如下所述。因此，在第 4 圖中表示為 $P_2^{n,k}$ 的濕式演示指示從 SAOC 參數處理單元 42 中獲得，如第 4 圖中由虛線箭頭所指出的。

該混合階段 53 將乾式及濕式演示路徑 46 及 48 的雙耳輸出信號 54 及 64 二者相混合，以獲得最終的雙耳輸出信號 24。如第 4 圖所示，該混合階段 53 受組配以將該等雙耳輸出信號 54 及 56 的左及右聲道個別地相混合，且因此可分別包含用以對其等左聲道求和的一相加器 66，及用以對其等右聲道求和的一相加器 68。

在描述完 SAOC 解碼器 12 的結構及降混預處理單元 40 的內部結構之後，下面來描述其等的功能。特別的是，下面所描述的詳細實施例對於 SAOC 參數處理單元 42 呈現出不同的替代方式，來推導出演示指示資訊 44，從而控制雙耳輸出信號 24 的聲道內相干性。換句話說，該 SAOC 參數處理單元 42 不僅運算出該演示指示資訊 44，還同時控制混合率，藉由該混合率，將初步及校正雙耳信號 55 及 64

混合為最終的雙耳輸出信號 24。

根據第一替代方式，該 SAOC 參數處理單元 42 受組配以控制上述的混合率，如第 5 圖所示。特別的是，在步驟 80 中，該初步雙耳輸出信號 54 的一實際雙耳聲道內的相干性值藉由單元 42 來判定或評估。在步驟 82 中，SAOC 參數處理單元 42 判定一目標雙耳聲道內相干性值。從而基於此等經判定的聲道內相干性值，在步驟 84 中，該 SAOC 參數處理單元 42 設定上述的混合率。特別的是，步驟 84 可包含，該 SAOC 參數處理單元 42 基於分別在步驟 80 及 82 中所判定出的聲道內相干性值，分別適當地運算出由乾式演示單元 42 所使用的乾式演示指示，及由濕式演示單元 52 所使用的濕式演示指示。

在下面中，上述的替代方式將在一數學的基礎上來描述。該等替代方式在 SAOC 參數處理單元 42 判定演示指示資訊 44 方面相互不同，該演示指示資訊 44 包括固有地控制乾式與濕式演示路徑 46 與 48 之間之混合率的乾式演示指示及濕式演示指示。根據第 5 圖所述之第一替代方式，該 SAOC 參數處理單元 42 判定一目標雙耳聲道內的相干性值。如下面將更詳細的描述，單元 42 可基於一目標相干性矩陣 $F = A \cdot E \cdot A^*$ 的成分來執行此判定，其中“*”表示共軛轉置， A 是一目標雙耳演示矩陣，該目標雙耳演示矩陣使該等目標/音訊信號 1...N 分別相關於雙耳輸出信號 24 及初步雙耳輸出信號 54 的右聲道及左聲道，且由演示資訊 26 及 HRTF 參數 27 推導出，且 E 是一矩陣，該矩陣的係數由

$IOC_{ij}^{l,m}$ 及目標位準差 $OLD_i^{l,m}$ 推導出。該運算可執行於該等 SAOC 參數的空間/時間解析度中，即對於每一 (l,m) 。然而，更可能的是，在各自的結果之間內插的一較低的解析度中執行該運算。後者的陳述對於下面提出的後續運算也是適合的。

因為目標雙耳演示矩陣 A 使輸入目標 $1 \dots N$ 分別相關於該雙耳輸出信號 24 及初步雙耳輸出信號 54 的左聲道與右聲道，所以其大小為 $2 \times N$ ，即

$$A = \begin{pmatrix} a_{11} & \dots & a_{1N} \\ a_{21} & \dots & a_{2N} \end{pmatrix}$$

上述矩陣 E 的大小為 $N \times N$ ，其中其等係數定義為

$$e_{ij} = \sqrt{OLD_i \cdot OLD_j} \cdot \max(IOC_{ij}, 0)$$

因而，該矩陣 E 為

$$E = \begin{pmatrix} e_{11} & \cdots & e_{1N} \\ \vdots & \ddots & \vdots \\ e_{N1} & \cdots & e_{NN} \end{pmatrix}$$

具有沿著其對角線的目標位準差，即

$$e_{ii} = OLD_i$$

因為對於 $i=j$ ， $IOC_{ij}=1$ ，而矩陣 E 具有在其對角線外的矩陣係數，該等矩陣係數表示分別由目標內互相關測量 IOC_{ij} 加權（大於 0 時設為 IOC_{ij} ，否則係數設為 0）之目標 i 及 j 的目標位準差的幾何平均值。

與此進行比較，下面所描述的第二及第三替代方式藉由找出方程式之最小平方意義上的最佳匹配，力求獲得該

等演示矩陣，該方程式藉由乾式演示矩陣 G 將立體聲降混信號 18 映射於初步雙耳輸出信號 54 上，以使目標演示方程式經由矩陣 A 將該等輸入目標映射於該“目標”雙耳輸出信號 24 上，其中該第二及第三替代方式在最佳匹配形成方面及濕式演示矩陣選擇方面相互不同。

為了能夠更容易地理解下面的替代，在數學上重新描述上述的第 3 及 4 圖的描述。如上所述，立體聲降混信號 18 $X^{n,k}$ 與該等 SAOC 參數 20 及使用者所定義的演示資訊 26 一起到達 SAOC 解碼器 12。而且，SAOC 解碼器 12 及 SAOC 參數處理單元 42 分別如箭頭所指示，對一 HRTF 資料庫 27 進行存取。該等經發送的 SAOC 參數包含對於所有 N 個目標 i,j 的目標位準差 $OLD_i^{l,m}$ 、目標內互相關值 $IOC_{ij}^{l,m}$ 、降混增益 $DMG_i^{l,m}$ 及降混聲道的位準差 $DCLD_i^{l,m}$ ，其中“ l,m ”表示各自的時間/頻譜瓦片 39，其中 l 表示時間且 m 表示頻率。對於所有的虛擬揚聲器位置或虛擬空間聲源位置 q ，對於左(L)及右(R)雙耳聲道及對於所有的頻帶 m ，HRTF 參數 27 示範地假設以 $P_{q,L}^m$, $P_{q,R}^m$ 及 Φ_q^m 給定出。

降混預處理單元 40 受組配以運算雙耳輸出 $\hat{X}^{n,k}$ ，如由立體聲降混 $X^{n,k}$ 及經去相關單降混信號 $X_d^{n,k}$ 來運算出，為

$$\hat{X}^{n,k} = G^{n,k} X^{n,k} + P_2^{n,k} X_d^{n,k}$$

該經去相關信號 $X_d^{n,k}$ 感知地等效於該立體聲降混信號 18 的左及右降混聲道的和 58，但對其予以最大地去相關，根據

$$X_d^{n,k} = \text{decorrFunction}\left(\begin{pmatrix} 1 & 1 \end{pmatrix} X^{n,k}\right)$$

參照第 4 圖，該去相關信號產生器 50 執行上述公式的 decorrFunction 函數。

而且，還如上所述，該降混預處理單元 40 包含二並行的路徑 46 及 48。因此，上述的方程式基於二個相依於時間/頻率的矩陣，即對於乾式路徑的 $G^{l,m}$ 及對於濕式路徑的 $P_2^{l,m}$ 。

如第 4 圖所示，在濕式路徑上的去相關可藉由左及右降混聲道的和來實施，該和傳送至產生一信號 62 的一去相關器 60 中，該信號 62 感知地等效於其輸入 58，但對該輸入 58 予以最大地去相關。

上述矩陣的元素藉由 SAOC 預處理單元 42 來運算。還如上所述，上述矩陣的元素可在該等 SAOC 參數的時間/頻率解析度下(即對於每一時槽 l 及每一處理頻帶 m)運算。從而所獲得的矩陣元素可在頻率上擴展且在時間上被內插，產生對應於所有濾波器組的時槽 n 及頻率子帶 k 而定義的矩陣 $E^{n,k}$ 及 $P_2^{l,m}$ 。然而，如上，也有一些替代方式。例如，可去除內插，使得在上面的方程式中，指數 n,k 可有效地由“ l,m ”替代。而且，上述矩陣之元素的運算甚至可在內插於解析度 l,m 或 n,k 上之一經降低的時間/頻率解析度下執行。因而，同樣，雖然在下面中，指數 l,m 指示，該等矩陣計算對應於每一瓦片 39 來執行，該計算可在某一較低的解析度下執行，當由降混預處理單元 40 施以該等各自矩陣時，可將該等演示矩陣內插直至一最終的解析度，諸如下至個別子帶值 32 的 QMF 時間/頻率解析度。

根據上述的第一替代方式，乾式演示矩陣 $\mathbf{G}^{l,m}$ 個別地對應於左及右降混聲道而運算出，使得

$$\mathbf{G}^{l,m} = \begin{pmatrix} P_L^{l,m,1} \cos(\beta^{l,m} + \alpha^{l,m}) \exp(j \frac{\phi^{l,m,1}}{2}) & P_L^{l,m,2} \cos(\beta^{l,m} + \alpha^{l,m}) \exp(j \frac{\phi^{l,m,2}}{2}) \\ P_R^{l,m,1} \cos(\beta^{l,m} - \alpha^{l,m}) \exp(-j \frac{\phi^{l,m,1}}{2}) & P_R^{l,m,2} \cos(\beta^{l,m} - \alpha^{l,m}) \exp(-j \frac{\phi^{l,m,2}}{2}) \end{pmatrix}$$

該等相對應的增益 $P_L^{l,m,x}$ 、 $P_R^{l,m,x}$ 及相位差 $\phi^{l,m,x}$ 定義為

$$P_L^{l,m,x} = \sqrt{\frac{f_{11}^{l,m,x}}{V^{l,m,x}}} , \quad P_R^{l,m,x} = \sqrt{\frac{f_{22}^{l,m,x}}{V^{l,m,x}}} ,$$

$$\phi^{l,m,x} = \begin{cases} \arg(f_{12}^{l,m,x}) & \text{if } 0 \leq m \leq \text{const}_1 \wedge \frac{|f_{12}^{l,m,x}|}{\sqrt{f_{11}^{l,m,x} f_{22}^{l,m,x}}} \geq \text{const}_2 \\ 0 & \text{else} \end{cases}$$

其中 const_1 可是例如 11，且 const_2 可是 0.6。該指標 x 表示左或右降混聲道，且因此假設為 1 或 2。

大體上來說，上面的條件於一較高頻譜範圍與一較低頻譜範圍間有區別，且特別地僅(可能)滿足於較低的頻譜範圍。此外或可選擇地，該條件依據該實際雙耳聲道內相干性值與目標雙耳聲道內相干性值之其中一者是否與相干性臨界值具有一預定的關係而定，即僅在該相干性超過該臨界值時，(可能)滿足該情況。如上所述的個別子條件可藉由一及運算來結合。

純量 $V^{l,m,x}$ 運算為

$$V^{l,m,x} = \mathbf{D}^{l,m,x} \mathbf{E}^{l,m} (\mathbf{D}^{l,m,x}) + \varepsilon .$$

應注意的是 ε 可與上述定義降混增益之 ε 相同或不同。該矩陣 \mathbf{E} 在上面已經介紹過。指標 (l,m) 僅表示上面已提及之矩陣運算的時間/頻率的相依性。而且，該等矩陣 $\mathbf{D}^{l,m,x}$ 也已在上面針對於降混增益及降混聲道之位準差的定

義而提及，使得 $D^{l,m,1}$ 相對應於上述之 D_1 ，且 $D^{l,m,2}$ 相對應於上述之 D_2 。

然而，為了更容易理解 SAOC 參數處理單元 42 如何由所接收之 SAOC 參數推導出乾式產生矩陣 $G^{l,m}$ ，在聲道降混矩陣 $D^{l,m,x}$ 與降混指示之間的相對應性再次被表示，但是以相反方向，該降混指示包含降混增益 $D^{l,m}$ 及 $DCLD_i^{l,m}$ 。特別的是，大小為 $1 \times N$ 之聲道降混矩陣 $D^{l,m,x}$ 的元素 $d_i^{l,m,x}$ ，即 $D^{l,m,x} = (d_1^{l,m,x}, \dots, d_N^{l,m,x})$ ，給出為

$$d_i^{l,m,1} = 10 \frac{DMG_i^{l,m}}{20} \sqrt{\frac{\tilde{d}_i^{l,m}}{1 + \tilde{d}_i^{l,m}}}, \quad d_i^{l,m,2} = 10 \frac{DMG_i^{l,m}}{20} \sqrt{\frac{1}{1 + \tilde{d}_i^{l,m}}}$$

其中元素 $\tilde{d}_i^{l,m}$ 定義為

$$\tilde{d}_i^{l,m} = 10^{\frac{DCLD_i^{l,m}}{10}}.$$

在上面 $G^{l,m}$ 的方程式中，增益 $P_L^{l,m,x}$ 與 $P_R^{l,m,x}$ 及相位差 $\phi^{l,m,x}$ 依據一聲道- x 個別的目標協方差矩陣 $F^{l,m,x}$ 的係數 f_{uv} 而定，該聲道- x 個別的目標協方差矩陣 $F^{l,m,x}$ 將依次如下面更詳細的描述，依據大小為 $N \times N$ 之一矩陣 $E^{l,m,x}$ 而定，該矩陣 $E^{l,m,x}$ 的元素 $e_{ij}^{l,m,x}$ 被運算為

$$e_{ij}^{l,m,x} = e_{ij}^{l,m} \left(\frac{d_i^{l,m,x}}{d_i^{l,m,1} + d_i^{l,m,2}} \right) \left(\frac{d_j^{l,m,x}}{d_j^{l,m,1} + d_j^{l,m,2}} \right)$$

如上所述，給出大小為 $N \times N$ 的矩陣 $E^{l,m}$ 的元素 $e_{ij}^{l,m}$ ，為

$$e_{ij}^{l,m} = \sqrt{OLD_i^{l,m} \cdot OLD_j^{l,m}} \cdot \max(IOC_{ij}^{l,m}, 0).$$

具有元素 $f_{uv}^{l,m,x}$ ，大小為 2×2 的上述目標協方差矩陣 $F^{l,m,x}$ 相似於上面所指出的協方差矩陣 F ，其給出為

$$F^{l,m,x} = A^{l,m} E^{l,m,x} (A^{l,m})^* ,$$

其中“*”相對應於共軛轉置。

目標雙耳演示矩陣 $A^{l,m}$ 由所有 N_{HRTF} 虛擬揚聲器位置 q 的 HRTF 參數 Φ_q^m 、 $P_{q,R}^m$ 與 $P_{q,L}^m$ 及演示矩陣 $M_{ren}^{l,m}$ 推導出，且其大小為 $2 \times N$ 。其等元素 $a_{ui}^{l,m}$ 定義在所有目標 i 與雙耳輸出信號之間所期望的關係，為

$$a_{1,i}^{l,m} = \sum_{q=0}^{N_{\text{HRTF}}-1} m_{q,i}^{l,m} P_{q,L}^m \exp\left(j \frac{\phi_q^m}{2}\right), \quad a_{2,i}^{l,m} = \sum_{q=0}^{N_{\text{HRTF}}-1} m_{q,i}^{l,m} P_{q,R}^m \exp\left(-j \frac{\phi_q^m}{2}\right) .$$

具有元素 $m_{qi}^{l,m}$ 的演示矩陣 $M_{ren}^{l,m}$ 使每一音訊目標 i 相關於由 HRTF 所表示的一虛擬揚聲器 q 。

濕式升混矩陣 $P_2^{l,m}$ 基於矩陣 $G^{l,m}$ 來計算，為

$$\mathbf{P}_2^{l,m} = \begin{pmatrix} P_L^{l,m} \sin(\beta^{l,m} + \alpha^{l,m}) \exp\left(j \frac{\arg(c_{12}^{l,m})}{2}\right) \\ P_R^{l,m} \sin(\beta^{l,m} - \alpha^{l,m}) \exp\left(-j \frac{\arg(c_{12}^{l,m})}{2}\right) \end{pmatrix}$$

該等增益 $P_L^{l,m}$ 及 $P_R^{l,m}$ 定義為

$$P_L^{l,m} = \sqrt{\frac{c_{11}^{l,m}}{V^{l,m}}}, \quad P_R^{l,m} = \sqrt{\frac{c_{22}^{l,m}}{V^{l,m}}} .$$

乾式雙耳信號 54 之具有元素 $c_{u,v}^{l,m}$ 的 2×2 的協方差矩陣 $C^{l,m}$ 遭評估為

$$C^{l,m} = \tilde{G}^{l,m} D^{l,m} E^{l,m} (D^{l,m})^* (\tilde{G}^{l,m})^*$$

其中

$$\tilde{G}^{l,m} = \begin{pmatrix} P_L^{l,m,1} \exp\left(j \frac{\phi^{l,m,1}}{2}\right) & P_L^{l,m,2} \exp\left(j \frac{\phi^{l,m,2}}{2}\right) \\ P_R^{l,m,1} \exp\left(-j \frac{\phi^{l,m,1}}{2}\right) & P_R^{l,m,2} \exp\left(-j \frac{\phi^{l,m,2}}{2}\right) \end{pmatrix}$$

計算純量 $V^{l,m}$ ，為

$$V^{l,m} = W^{l,m} E^{l,m} (W^{l,m})^* + \varepsilon .$$

給出大小為 $1 \times N$ 之濕式單降混矩陣 $W^{l,m}$ 的元素 $w_i^{l,m}$ ，

為

$$w_i^{l,m} = d_i^{l,m,1} + d_i^{l,m,2} \circ$$

給出大小為 $2 \times N$ 之立體聲降混矩陣 $D^{l,m}$ 的元素 $d_{x,i}^{l,m}$ ，為

$$d_{x,i}^{l,m} = d_i^{l,m,x} \circ$$

在上述的 $G^{l,m}$ 方程式中， $\alpha^{l,m}$ 及 $\beta^{l,m}$ 表示專用於 ICC 控制的旋轉角。特別的是，旋轉角 $\alpha^{l,m}$ 控制乾式及濕式雙耳信號的混合，以將雙耳輸出 24 的 ICC 調整至雙耳目標的 ICC。在設定旋轉角時，乾式雙耳信號 54 的 ICC 應予以考慮，該乾式雙耳信號 54 的 ICC 依據音訊內容及立體聲降混矩陣 D 而定，典型地小於 1.0 且大於目標 ICC。這與一單降混為基式雙耳演示形成對比，其中該乾式雙耳信號的 ICC 總是等於 1.0。

該等旋轉角 $\alpha^{l,m}$ 及 $\beta^{l,m}$ 控制乾式及濕式雙耳信號的混合。該乾式雙耳演示立體聲降混 54 的 ICC $\rho_c^{l,m}$ 在步驟 80 中評估為

$$\rho_c^{l,m} = \min\left(\frac{|c_{12}^{l,m}|}{\sqrt{c_{11}^{l,m} c_{22}^{l,m}}}, 1\right) \circ$$

整體的雙耳目標 ICC $\rho_c^{l,m}$ 在步驟 82 中評估為或判定是

$$\rho_T^{l,m} = \min\left(\frac{|f_{12}^{l,m}|}{\sqrt{f_{11}^{l,m} f_{22}^{l,m}}}, 1\right) \circ$$

該等用以使濕式信號之能量最小化的旋轉角 $\alpha^{l,m}$ 及 $\beta^{l,m}$ 在步驟 84 中被設定為

$$\alpha^{l,m} = \frac{1}{2} (\arccos(\rho_T^{l,m}) - \arccos(\rho_c^{l,m})) \circ$$

$$\beta^{l,m} = \arctan\left(\tan(\alpha^{l,m}) \frac{P_R^{l,m} - P_L^{l,m}}{P_L^{l,m} + P_R^{l,m}}\right)。$$

因而，根據上述對用以產生雙耳輸出信號 24 之 SAOC 解碼器 12 之功能的數學描述，該 SAOC 參數處理單元 42 在判定實際雙耳 ICC 中，藉由上述 $\rho_C^{l,m}$ 的方程式及上述輔助方程式的使用來計算 $\rho_C^{l,m}$ 。類似地，SAOC 參數處理單元 42 在判定步驟 82 中之目標雙耳 ICC 中，藉由上面所示方程式及輔助方程式來運算 $\rho_C^{l,m}$ 。在此基礎上，SAOC 參數處理單元 42 在步驟 84 中判定該等旋轉角，從而設定在乾式與濕式演示路徑之間的混合率。根據此等旋轉角，SAOC 參數處理單元 42 建立該等乾式及濕式演示矩陣或升混參數 $G^{l,m}$ 及 $P_2^{l,m}$ ，其等依次在解析度 n,k 下由降混預處理單元 40 使
用，以由立體聲降混 18 推導出雙耳輸出信號 24。

應注意的是上述的第一替代方式可在某些方面上變化。例如，上述聲道內相位差 $\Phi_C^{l,m}$ 的方程式可改變至一定程度，使第二子條件可將該乾式經雙耳演示立體聲降混的實際 ICC 與 const_2 (而不是由聲道的個別協方差矩陣 $F^{l,m,x}$ 所判定的 ICC)進行比較，使得在此方程式中， $\frac{|f_{12}^{l,m,x}|}{\sqrt{f_{11}^{l,m,x} f_{22}^{l,m,x}}}$ 部分將由

項目 $\frac{|c_{12}^{l,m}|}{\sqrt{c_{11}^{l,m} c_{22}^{l,m}}}$ 替代。

而且，應注意的是，根據所選擇的符號，在上面的一些方程式中，當諸如 ε 的一純量常量加至一矩陣使得此常數加至各自矩陣的每一係數中時，可省略全為 1 的矩陣。

具有較高目標擷取可能的乾式演示矩陣的另一產生方

式是基於該等左及右降混聲道的一聯合處理。為了簡明，省略該子帶指標對，此精神的目的在於最小平方意義上的最佳匹配

$$\hat{X} = G X$$

到目標演示

$$Y = AS .$$

這產生物標協方差矩陣：

$$YY^* = ASS^*A^*$$

其中複數值的目標雙耳演示矩陣 A 在一先前的公式中給出，且矩陣 S 按列包含原始目標的子帶信號。

該最小平方的匹配由二階資訊來運算，該二階資訊由經傳達的目標及降混資料推導出。也就是，執行下面的替代

$$XX^* \leftrightarrow DED^* ,$$

$$YX^* \leftrightarrow AED^* ,$$

$$YY^* \leftrightarrow AEA^* .$$

為了致動該等替代，請回想 SAOC 目標參數典型地載有目標功率資訊(OLD)及(選定的)目標內互相關(IOC)。由此等參數，推導出 $N \times N$ 的目標協方差矩陣 E ，該目標協方差矩陣 E 表示 SS^* 的一近似值，即 $E \approx SS^*$ ，從而產生 $YY^* = AEA^*$ 。

而且， $X = DS$ 及降混協方差矩陣變成：

$$XX^* = DSS^*D^* ,$$

其可再次藉由 $XX^* = DED^*$ 從 E 中推導出。

乾式演示矩陣 G 藉由解出最小平方的問題而獲得

$$\min \{ \text{norm}\{ Y-X \} \}.$$

$$G = G_0 = YX^* (XX^*)^{-1}$$

其中 YX^* 被運算為 $YX^* = AED^*$ 。

因而，乾式演示單元 42 藉由 2×2 的升混矩陣 G 的使用，藉由 $\hat{X} = GX$ ，由降混信號 X 判定雙耳輸出信號 \hat{X} ，且該 SAOC 參數處理單元藉由上面公式的使用將 G 判定為

$$G = AED^* (DED^*)^{-1},$$

給出複數值的乾式演示矩陣，複數值濕式演示矩陣 P (以前表示為 P_2) 藉由考慮遺漏的協方差誤差矩陣而在該 SAOC 參數處理單元 42 中運算

$$\Delta R = YY^* - G_0 XX^* G_0^*.$$

可顯示出的是，此矩陣是正的，且透過選擇相對應於 ΔR 之最大特徵值 λ 的一單元規範特徵向量 u 及調節其，給出 P 的一較佳選擇，根據

$$P = \sqrt{\frac{\lambda}{V}} u,$$

其中，純量 V 如上來運算，即 $V = WE(W)^* + \varepsilon$ 。

換句話說，因為濕式路徑被安置，以校正所獲得之乾式解的相關性， $\Delta R = AEA^* - G_0 DED^* G_0^*$ 表示遺漏的協方差誤差矩陣，即分別地 $YY^* = \hat{X} \hat{X}^* + \Delta R$ 或 $\Delta R = YY^* - \hat{X} \hat{X}^*$ ，且因而該 SAOC 參數處理單元 42 保留 P ，使得 $PP^* = \Delta R$ ，對此的一解透過選擇上述的單元規範特徵向量 u 而給出。

用以產生乾式及濕式演示矩陣的一第三方法表示出基

於經約束複數預測的提示對演示參數的一評估，且將恢復正確的複數協方差結構的優點與對於改良目標擷取之降混聲道的聯合處理的利益相結合。由此方法所提供之一附加機會是，在許多情況下能夠完全地省略濕式升混，從而為一具有較低運算複雜性的雙耳演示版本作好準備。如依據該第二替代方式，下面所呈現的第三替代方式基於左及右降混聲道的一聯合處理。

本原理的目的在於最小平方意義上的最佳匹配

$$\hat{\mathbf{X}} = \mathbf{G}\mathbf{X}$$

到正確複數協方差之約束下的目標演示 $\mathbf{Y} = \mathbf{AS}$

$$\mathbf{GXX}^* \mathbf{G}^* + \mathbf{VPP}^* = \hat{\mathbf{Y}}\hat{\mathbf{Y}}^*.$$

因而，它的目的在於找出 \mathbf{G} 及 \mathbf{P} 的解，使得

- 1) $\hat{\mathbf{Y}}\hat{\mathbf{Y}}^* = \mathbf{YY}^*$ (是對 2) 中公式的約束)；及
- 2) $\min\{\text{norm}\{\mathbf{Y} - \hat{\mathbf{Y}}\}\}$ ，如其在第二替代方式中所要求的一樣。

由於拉格朗日乘數的理論，由此推斷出存在一自伴隨矩陣 $\mathbf{M} = \mathbf{M}^*$ ，使得

$$\mathbf{MP} = \mathbf{0} \text{，且}$$

$$\mathbf{MGXX}^* = \mathbf{YX}^*.$$

在一般的情況下，其中 \mathbf{YX}^* 及 \mathbf{XX}^* 二者是非奇異的，它得自於 \mathbf{M} 為非奇異的第二方程式，且從而 $\mathbf{P} = \mathbf{0}$ 是對第一方程式的唯一解。這是不具濕式演示的解。設定 $\mathbf{K} = \mathbf{M}^{-1}$ ，可看出的是，相對應的乾式升混由下給出

$$\mathbf{G} = \mathbf{KG}_0$$

其中 \mathbf{G}_0 是上面針對第二替代方式所推導出的預測解，且該自伴隨矩陣 \mathbf{K} 解決

$$\mathbf{K}\mathbf{G}_0\mathbf{X}\mathbf{X}^*\mathbf{G}_0^*\mathbf{K}^* = \mathbf{Y}\mathbf{Y}^*.$$

如果其唯一為正且因此矩陣 $\mathbf{G}_0\mathbf{X}\mathbf{X}^*\mathbf{G}_0^*$ 的自伴隨矩陣的平方根由 \mathbf{Q} 表示，那麼該解可寫為

$$\mathbf{K} = \mathbf{Q}^{-1}(\mathbf{Q}\mathbf{Y}\mathbf{Y}^*\mathbf{Q})^{1/2}\mathbf{Q}^{-1}.$$

因而，SAOC 參數處理單元 42 判定 \mathbf{G} 為 $\mathbf{K}\mathbf{G}_0 = \mathbf{Q}^{-1}(\mathbf{Q}\mathbf{Y}\mathbf{Y}^*\mathbf{Q})^{1/2}\mathbf{Q}^{-1}\mathbf{G}_0 = (\mathbf{G}_0\mathbf{D}\mathbf{E}\mathbf{D}^*\mathbf{G}_0^*)^{-1}(\mathbf{G}_0\mathbf{D}\mathbf{E}\mathbf{D}^*\mathbf{G}_0^* \mathbf{A}\mathbf{E}\mathbf{A}^*\mathbf{G}_0\mathbf{D}\mathbf{E}\mathbf{D}^*\mathbf{G}_0^*)^{1/2}(\mathbf{G}_0\mathbf{D}\mathbf{E}\mathbf{D}^*\mathbf{G}_0^*)^{-1}\mathbf{G}_0$ ，其中 $\mathbf{G}_0 = \mathbf{A}\mathbf{E}\mathbf{D}^*(\mathbf{D}\mathbf{E}^*)^{-1}$ 。

對於內部平方根，將大體上有四個自伴隨解，且導致 $\hat{\mathbf{X}}$ 至 \mathbf{Y} 之最佳匹配的解被予以選擇。

實際上，必須例如藉由對所有乾式演示矩陣係數之絕對平方值的和限制條件，將乾式演示矩陣 $\mathbf{G} = \mathbf{K}\mathbf{G}_0$ 限制為一最大大小，這可表示為

$$\text{trace}(\mathbf{G}\mathbf{G}^*) \leq g_{\max}.$$

如果解違背了此限制條件，那麼取決於界限的解將予以替代。這透過將約束條件

$$\text{trace}(\mathbf{G}\mathbf{G}^*) = g_{\max}$$

加至該等先前的約束條件中及重新推導出拉格朗日方程式來實現。其結果是，先前的方程式

$$\mathbf{M}\mathbf{G}\mathbf{X}\mathbf{X}^* = \mathbf{Y}\mathbf{X}^*$$

必須由

$$\mathbf{M}\mathbf{G}\mathbf{X}\mathbf{X}^* + \mu\mathbf{I} = \mathbf{Y}\mathbf{X}^*$$

來替代。其中 μ 是一附加的中間複數參數，且 I 是 2×2 的單位矩陣。可產生一具有非零濕式演示 P 的解。特別的是，濕式升混矩陣的解可藉由 $PP^* = (YY^* - GXX^*G^*) / V = (AEA^* - GDED^*G^*) / V$ 來找出，其中 P 的選擇較佳地基於上述針對於第二替代方式之特徵值的考慮，且 V 是 $WEW^{*++\varepsilon}$ 。 P 稍後的判定也藉由 SAOC 參數處理單元 42 來完成。

從而判定出的矩陣 G 及 P 接著由該等濕式及乾式演示單元使用，如先前所述。

如果需要一低複雜性的版本，那麼下一步驟是代替，即使此解即是不具有濕式演示的解。實現此的一較佳方法是，將複數協方差的限制減少為僅在對角線上匹配，使得正確的信號功率仍能在右及左聲道中實現，但互協方差處於未知的狀態。

關於第一替代方式，收聽測試的個人被引導至一聲學隔離的收聽室中，該收聽室被設計為允許高品質的收聽。該結果在下面予以描述。

播放透過使用耳機(具有 Lake-People 式數位/類比轉換器的 STAX SR Lambda Pro 耳機及 STAX SRM 監測器)來完成。該測試方法在用於空間音訊驗證測試的標準程序之後，基於對於中等品質音訊之主觀估計的“隱藏參考和基準的多刺激”(MUSHRA)方法。

總共 5 位收聽者參與了所執行的每一項測試。所有個體可被認為是有經驗的收聽者。根據 MUSHRA 方法學，該

等收聽者被指導去相對於參照比較所有的測試條件。該等測試條件自動地隨機賦予每一測試項目及每一收聽者。該等主觀的響應藉由一電腦為基的 MUSHRA 程式，按從 0 至 100 的一刻度範圍來記錄。允許在該等待測項目之間瞬間轉換。該等 MUSHRA 測試已經予以導入，以評估該 MPEG SAOC 系統之所期望的立體聲至雙耳處理的感知性能。

為了評估所期望之系統相較於單聲道至雙耳性能的一感知品質增益，由該單聲道至雙耳系統處理的項目也包括於該測試中。該等相對應的單聲道及立體聲降混信號在每聲道每秒 80 kbit 下予以 AAC 編碼。

隨著 HRTF 資料庫“KEMAR_MIT_COMPACT”予以使用。參考條件透過考慮所期望之演示的適當加權的 HRTF 脈衝響應，雙耳過濾目標而產生。該基準條件是低通過濾參考條件(在 3.5 kHz 下)。

表格 1 包含該等經測試項目的列表。

表格 1 – 收聽測試的音訊專案

收聽項目	接近.單聲道/立體聲目標	目標角 目標增益(dB)
disco1	10/0	[-30, 0, -20, 40, 5,-5, 120, 0, -20, -40]
disco2		[-3, -3, -3, -3, -3, -3, -3, -3,-3] [-30, 0, -20, 40, 5, -5, 120, 0, -20, -40] [-12, -12, 3, 3, -12, -12, 3, -12, 3, -12]
coffee1	6/0	[10, -20, 25, -35, 0, 120]
coffee2		[0, -3, 0, 0, 0, 0] [10, -20, 25, -35, 0, 120] [3, -20, -15, -15, 3, 3]
pop2	1/5	[0, 30, -30, -90, 90, 0, 0, -120, 120,

		-45, 45] [4, -6, -6, 4, 4, -6, -6, -6, -6, -16, -16]
--	--	--

五個不同的場景已經予以測試，其等是演示來自 3 個不同目標聲源庫的(單聲道或立體聲)目標的結果。三個不同的降混矩陣已用於 SAOC 編碼器中，參見表格 2。

表格 2 – 降混類型

降混類型	單聲道	立體聲	雙重單聲道
Matlab 符號	<code>dmx1=ones(1,N);</code>	<code>dmx2=zeros(2,N);</code> <code>dmx2(1,1:2:N)=1;</code> <code>smx2(2,2:2:N)=1;</code>	<code>dmx3=ones(2,N);</code>

該等升混表示的品質評估測試已經定義於表格 3 中。

表格 3 – 收聽測試條件

測試條件	降混類型	核心編碼器
x-1-b	單聲道	AAC@80kbps
x-2-b	立體聲	AAC@160kbps
x-2-b_Dual/Mono	雙重單聲道	AAC@160kbps
5222	立體聲	AAC@160kbps
5222_Dual/Mono	雙重單聲道	AAC@160kbps

該“5522”系統使用立體聲降混預處理器，如於 2008 年 7 月在德國漢諾威舉行的第 85 屆運動圖像專家組(MPEG)會議中提出的“ISO/IEC CD 23003-2:200x Spatial Audio Object Coding (SAOC)”，文件號第 N10045 號之 ISO/IEC JTC 1/SC 29/WG 11 (MPEG) 中所描述，該立體聲降混預處理器具有複數值的雙耳目標演示矩陣 $A^{l,m}$ 作為一輸入。也就是說，沒有 ICC 控制予以執行。非正式的收聽測試已經顯示，藉由對於上方頻帶採用 $A^{l,m}$ 的振幅，而不是使所有頻帶為複數值，改良了性能。改良的“5522”系統已經用於測試中。

在第 6 圖中可找到證明該等所獲得之收聽測試結果的圖形的一簡單回顧。此等描繪顯示，關於所有收聽者每一項目的平均 MUSHRA 分級，及關於所有經評估項目與相關的 95% 可信區間的統計平均值。應注意的是，隱藏參考的資料因為所有的個體已經正確地識別出，而在該等 MUSHRA 描繪中予以省略。

下面的觀察可基於該等收聽測試的結果予以作出：

- “x-2-b_DualMono”的表現比的上“5522”。
- “x-2-b_DualMono”的表現明顯優於“5222_DualMono”。
- “x-2-b_DualMono”的表現比的上“x-1-b”。
- 根據上面第一替代方式所實施的“x-2-b”與所有其他條件相比，具有稍微較佳的表現。
- 項目“disco1”在該等結果中沒有顯示出太多變化，因此可能不是適當的。

因而，在 SAOC 中立體聲降混信號的雙耳演示的一概念已在上面予以描述，來滿足不同降混矩陣的需要。特別的是，雙重單似降混的品質相同於真實的單降混，此已在一收聽測試中驗證。從立體聲降混與單降混進行比較所獲得的品質改良，也可從該收聽測試中看出。上述實施例的基本處理方塊是立體聲降混的乾式雙耳演示，及與一去相關濕式雙耳信號相混合(以二者方塊的一適當結合)。

- 特別的是，濕式雙耳信號使用具有單降混輸入的一去相關器來運算，使得左及右功率及 IPD 與在

該乾式雙耳信號中相同。

- 濕式及乾式雙耳信號的混合藉由目標 ICC 及乾式雙耳信號的 ICC 來控制，使得其典型地與單降混為基式雙耳演示相比需要較少的去相關，從而產生較高的總的聲音品質。
- 而且，上面的實施例可以一穩定的方式，對應於單聲道/立體聲降混輸入與單聲道/立體聲/雙耳輸出的任何結合而予以簡單地修改。

換句話說，上面描述了提供用於由聲道內相干性控制來解碼及雙耳演示立體聲降混為基式 SAOC 位元流的信號處理架構和方法的實施例。單或立體聲降混輸入與單、立體聲或雙耳輸出的所有組合可作為所描述之立體聲降混為基式的概念的特殊情況來處理。立體聲降混為基式概念的品質結果顯示出，其典型地與單降混為基式的概念相比品質更佳，此已在上述的 MUSHRA 收聽測試中獲驗證。

在 2008 年 7 月，德國漢諾威舉行的第 85 屆 MPEG 會議中提出的“ISO/IEC CD 23003-2:200x Spatial Audio Object Coding (SAOC)”，檔號第 N10045 號，空間音訊目標編碼(SAOC) ISO/IEC JTC 1/SC 29/WG 11 (MPEG) 中，多個音訊目標被降混為一單聲道或立體聲信號。此信號予以編碼，且與旁側資訊(SAOC 參數)一起發送至 SAOC 解碼器。該等上面的實施例，使雙耳輸出信號的聲道內相干性(ICC)成為感知虛擬聲源寬度的一重要測量，且由於編碼器降混，品質降低的或甚至損壞的，(幾乎)完全地予以修正。

輸入系統的是立體聲降混、SAOC 參數、空間演示資訊及一 HRTF 資料庫。輸出是雙耳信號。輸入及輸出二者典型地藉由諸如 MPEG 環繞混合 QMF 濾波器組(ISO/IEC 23003-1:2007，資訊技術-MPEG 音訊技術-第一部分：具有充分低的帶內混疊的 MPEG 環繞)的一過取樣複數調變分析濾波器組，在解碼器轉換域中給出。該雙耳輸出信號藉由該合成濾波器組，轉換回 PCM 時間域。換句話說，該系統從而是一可能的單降混為基式雙耳演示對於立體聲降混信號的一擴展。對於雙重單降混信號，系統的輸出與此單降混為基式系統是相同的。因而，該系統可藉由以一穩定的方式設定該等演示參數，而來處理單/立體聲降混輸入與單/立體聲/雙耳輸出的任何結合。

再換句話說，該等上面的實施例由 ICC 控制來執行立體聲降混為基式 SAOC 位元流的雙耳演示及解碼。與一單降混為基式雙耳演示進行比較，該等實施例可在兩個方面利用該立體聲降混：

- 在不同降混聲道中之目標之間的相關特性獲得部分地保存
- 因為在一降混聲道中存在較少的目標，目標的擷取獲得改良

因而，在 SAOC 中立體聲降混信號的雙耳演示的一概念已在上面予以描述，來滿足不同降混矩陣的需要。特別的是，雙重單似降混的品質與真實單降混相同，此已在一收聽測試中獲驗證。從立體聲降混與單降混進行比較可獲

得的品質改良，也可從收聽測試中看出。上述實施例的基本處理方塊是乾式雙耳演示立體聲降混，及與一去相關濕式雙耳信號相混合(以二者方塊的一適當結合)。特別的是，該濕式雙耳信號透過使用有單降混輸入的一去相關器來運算，使得左及右功率及 IPD 與乾式雙耳信號中相同。濕式及乾式雙耳信號的混合受該目標 ICC 及單降混為基式雙耳演示來控制，從而產生較高的總的聲音品質。而且，上面的實施例可以一穩定的方式，對應於單/立體聲降混輸入與單/立體聲/雙耳輸出的任何結合予以簡單地修改。根據該等實施例，該立體聲降混信號 $X^{n,k}$ 與該等 SAOC 參數、使用者所定義的演示資訊及一 HRTF 資料庫一起作為輸入。該等經發送的 SAOC 參數是所有 N 個目標 i,j 的 $OLD_i^{l,m}$ (目標位準差)、 $IOC_{ij}^{l,m}$ (目標內互相關)、 $DMG_i^{l,m}$ (降混增益)及 $DCLD_i^{l,m}$ (降混聲道位準差)。該等 HRTF 參數以所有 HRTF 資料庫指標 q 的 $P_{q,L}^m$ 、 $P_{q,R}^m$ 及 ϕ_q^m 而給定，該指標 q 與某一空間聲源的位置相關聯。

最後，應注意的是，雖然在上面的描述中，術語“聲道內相干性”及“目標內互相關”以“相干性”為一個術語且“互相關”為另一個術語中，而予以不同地解讀，但是後面的術語可交換性地分別用作對於聲道內與目標內的類似性的測量。

根據一實際的實施，發明的雙耳演示概念可實施於硬體或軟體中。因而，本發明也相關於一電腦程式，該電腦程式可儲存於諸如一 CD、一磁碟、DVD、一記憶體條、一

記憶體卡或一記憶體晶片的一電腦可讀媒體中。本發明因而也是具有一程式碼的一電腦程式，該程式碼在於一電腦上執行時，執行結合於上面圖式所述之編碼、轉換或解碼的發明方法。

儘管此發明已經根據多個較佳實施例而獲描述，但是仍具有屬於此發明之範圍內的變更、置換及等效物。還應注意的是，實施本發明之方法及組成具有許多可選擇的方式。因而其打算將後面的附加申請專利範圍解讀為包括屬於本發明之真正精神及範圍內的所有此等變更、置換及等效物。

另外，應注意的是，在流程圖中所指示的所有步驟藉由分別在解碼器中的各自裝置來實施，該等實施的一裝置可包含執行於一CPU上的副程式、一ASIC的電路部分等。一相似的描述對於在該等方塊圖中該等方塊功能是真實的。

參考文獻：

2008年7月德國漢諾威舉行的第85屆MPEG會議中提出的ISO/IEC JTC 1/SC 29/WG 11 (MPEG)，第N10045號文件，“ISO/IEC CD 23003-2:200x Spatial Audio Object Coding (SAOC)”

1999年10月EBU技術介紹：“MUSHRA-EBU Method for Subjective Listening Tests of Intermediate Audio Quality”，文件號第B/ AIM022號

ISO/IEC 23003-1:2007, Information technology –

MPEG audio technologies – Part 1: MPEG Surround

2007 年 7 月在美國聖何塞提出的 ISO/IEC JTC1/SC29/WG11 (MPEG)，第 N9099 號文件：“Final Spatial Audio Object Coding Evaluation Procedures and Criterion”

Jeroen 、 Breebaart 、 Christof Faller: Spatial Audio Processing. MPEG Surround and Other Applications. Wiley & Sons, 2007.

2006 年在韓國首爾，Jeroen 、 Breebaart 等人提出的： Multi-Channel goes Mobile : MPEG Surround Binaural Rendering. AES 29th International Conference

【圖式簡單說明】

第 1 圖顯示可供本發明之該等實施例實施的一 SOAC 編碼器 / 解碼器安排的一方塊圖；

第 2 圖顯示一單音訊信號的一頻譜表示的一示意及說明圖；

第 3 圖顯示根據本發明之一實施例之能夠雙耳演示的一音訊解碼器的一方塊圖；

第 4 圖顯示根據本發明之一實施例之第 3 圖的降混預處理方塊的一方塊圖；

第 5 圖顯示根據第一替代方式，由第 3 圖之 SAOC 參數處理單元 42 所執行的步驟的一流程圖；

第 6 圖顯示說明該等收聽測試結果的一圖形。

【主要元件符號說明】

10... SAOC 編碼器	141~14N... 音訊信號
12... SAOC 解碼器	16... 降混器

18...降混信號	46...乾式路徑
20...旁側資訊	47...乾式演示單元
21... SAOC輸出資料流	48...濕式路徑
22...升混	50...去相關信號產生器
24...雙耳輸出信號	52...濕式演示單元
241~24 M'...聲道組	53...混合階段
26...演示資訊	54...初步雙耳輸出信號
27... HRTF參數	56...相加器
301~30P...子帶信號	58...單降混
32...子帶值	60...去相關器
34...時槽	62...經去相關信號
35...頻率軸	64...校正雙耳輸出信號
36...音框	66/68...相加器
37...時間軸	80~84...步驟
38...參數時間片段	L0...聲道
39...時間/頻率瓦片	R0...聲道
40...降混預處理單元	$X_d^{n,k}$...經去相關信號
42... SAOC參數處理單元	$P_2^{n,k}$...濕式演示指示
44...演示指示資訊	$G^{n,k}$...乾式演示指示

201036464

發明專利說明書

(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： 98132269

※申請日： 98.9.24 ※IPC 分類： A06F 5/00 (2006.01)

一、發明名稱：(中文/英文)

多聲道音訊信號之雙耳演示技術

BINAURAL RENDERING OF A MULTI-CHANNEL AUDIO SIGNAL

二、中文發明摘要：

本文描述將一多聲道音訊信號雙耳演示為一雙耳輸出信號。該多聲道音訊信號包含多個音訊信號被降混的一立體聲降混信號及旁側資訊，且該旁測資訊包含一指示出對於每一音訊信號，該各自音訊信號已分別混合至該立體聲降混信號的第一聲道及第二聲道中的程度的降混資訊，及該等音訊信號的目標位準資訊，及描述在該等音訊信號之音訊信號對之間的相似性的目標內互相關資訊。基於一第一演示指示，自該立體聲降混信號之該第一及第二聲道運算出一初步雙耳信號。產生一經去相關的信號作為對該立體聲降混信號之該第一及第二聲道的一單降混的一感知等效物，但然而是與該單降混去相關。根據一第二演示指示，自該經去相關信號運算出一校正雙耳信號，且該初步雙耳信號與該校正雙耳信號相混合，以獲得該雙耳輸出信號。

三、英文發明摘要：

Binaural rendering a multi-channel audio signal into a binaural output signal is described. The multi-channel audio signal comprises a stereo downmix signal into which a plurality of audio signals are downmixed, and side information comprising a downmix information indicating, for each audio signal, to what extent the respective audio signal has been mixed into a first channel and a second channel of the stereo downmix signal, respectively, as well as object level information of the plurality of audio signals and inter-object cross correlation information describing similarities between pairs of audio signals of the plurality of audio signals. Based on a first rendering prescription, a preliminary binaural signal is computed from the first and second channels of the stereo downmix signal. A decorrelated signal is generated as a perceptual equivalent to a mono downmix of the first and second channels of the stereo downmix signal being, however, decorrelated to the mono downmix. Depending on a second rendering prescription, a corrective binaural signal is computed from the decorrelated signal and the preliminary binaural signal is mixed with the corrective binaural signal to obtain the binaural output signal.

七、申請專利範圍：

1. 一種用於將一多聲道音訊信號雙耳演示為一雙耳輸出信號的設備，該多聲道音訊信號包含多個音訊信號被降混的一立體聲降混信號，且包含旁側資訊，該旁測資訊包含一指示出對於每一音訊信號，該各自音訊信號已分別混合至該立體聲降混信號的第一聲道及一第二聲道中的程度的降混資訊，及該等音訊信號的目標位準資訊，及描述在該等音訊信號之音訊信號對之間的類似性的目標內互相關資訊，該設備包含：

用於基於第一演示指示自該立體聲降混信號之該第一及第二聲道來運算出一初步雙耳信號的裝置，該第一演示指示依據該目標內互相關資訊、該目標位準資訊、該降混資訊、使每一音訊信號相關於一虛擬揚聲器位置的演示資訊及頭部相關轉換函數(HRTF)參數而定；

用於產生一經去相關信號的裝置，該經去相關信號作為對該立體聲降混信號之該第一及第二聲道的一單降混的一感知等效物，且然而是與該單降混去相關；

用於根據一第二演示指示自該經去相關信號運算出一校正雙耳信號的裝置，該第二演示指示依據該目標內互相關資訊、該目標位準資訊、該降混資訊、該演示資訊及該等HRTF參數而定；及

用於將該初步雙耳信號與該校正雙耳信號相混合，以獲得該雙耳輸出信號的裝置。

2. 如申請專利範圍第1項所述之設備，其中用於產生該經

去相關信號的該裝置受組配以用以對該立體聲降混信號的該第一及第二聲道求和，且用以對該和去相關以獲得該經去相關信號。

3. 如申請專利範圍第 1 或 2 項所述之設備，更包含：

用於評估該初步雙耳信號的一實際雙耳聲道內相干性值的裝置；

用於判定一目標雙耳聲道內相干性值的裝置；及

用於基於該實際雙耳聲道內相干性值及該目標雙耳聲道內相干性值，設定一混合率的裝置，該混合率判定該雙耳輸出信號分別由於該立體聲降混信號的該第一及第二聲道由用於運算出該初步雙耳信號的該裝置來處理，及該立體聲降混信號的該第一及第二聲道由用於產生該經去相關信號的該裝置及用於運算出該校正雙耳信號的該裝置來處理，而受到的影響的程度。

4. 如申請專利範圍第 3 項所述之設備，其中用於設定該混合率的該裝置受組配以基於該實際雙耳聲道內相干性值及該目標雙耳聲道內相干性值，透過設定該第一演示指示及該第二演示指示來設定該混合率。

5. 如申請專利範圍第 3 或 4 項所述之設備，其中用於判定該目標雙耳聲道內相干性值的該裝置受組配，以基於一目標協方差矩陣 $F = A E A^*$ 的成分來執行該判定，其中“ $*$ ”表示共軛轉置， A 是使該等音訊信號分別與該雙耳輸出信號的該第一及第二聲道相關的一目標雙耳演示矩陣，且惟獨由該演示資訊及該等 HRTF 參數來決定，

且 E 是惟獨由該目標內互相關資訊及該目標位準資訊來決定的一矩陣。

6. 如申請專利範圍第 5 項所述之設備，其中用於運算出該初步雙耳信號的該裝置受組配以執行該運算，使得

$$\hat{X}_1 = G \cdot X$$

其中 X 是一 2×1 的向量，其元素相對應於該立體聲降混信號的該第一及第二聲道， \hat{X}_1 是一 2×1 的向量，其元素相對應於該初步雙耳信號的該第一及第二聲道， G 是表示該第一演示指示且具有一 2×2 的大小的第一演示矩陣，即

$$G = \begin{pmatrix} P_L^1 \cos(\beta + \alpha) \exp(j \frac{\phi^1}{2}) & P_L^2 \cos(\beta + \alpha) \exp(j \frac{\phi^2}{2}) \\ P_R^2 \cos(\beta - \alpha) \exp(-j \frac{\phi^1}{2}) & P_R^2 \cos(\beta - \alpha) \exp(-j \frac{\phi^2}{2}) \end{pmatrix}$$

其中， $x \in \{1, 2\}$ ，

$$P_L^x = \sqrt{\frac{f_{11}^x}{V^x}}, \quad P_R^x = \sqrt{\frac{f_{22}^x}{V^x}},$$

$$\phi^x = \begin{cases} \arg(f_{12}^x) & \text{如果符合第一條件} \\ 0 & \text{其他} \end{cases}$$

其中 f_{11}^x 、 f_{12}^x 及 f_{22}^x 是大小為 2×2 的子目標協方差矩陣 F^x 的係數，即 $F^x = A E^x A^*$ ，

其中 $e_{ij}^x = e_{ij} \left(\frac{d_i^x}{d_i^1 + d_i^2} \right) \left(\frac{d_j^x}{d_j^1 + d_j^2} \right)$ 是 $N \times N$ 矩陣 E^x 的係數， N 是

音訊信號的數目， e_{ij} 是大小為 $N \times N$ 之該矩陣 E 的係數，且 d_i^x 惟獨由該降混資訊來決定，其中 d_i^1 指示音訊信號 i 已混合至該立體聲降混信號之該第一聲道中的程度，且 d_i^2 定義音訊信號 i 已混合至該立體聲輸出信號之該第二

聲道中的程度，

其中 V^x 是一純量，即 $V^x = D^x E (D^x)^* + \varepsilon$ ，且 D^x 是一 $1 \times N$ 的矩陣，其係數是 d_i^x ，

其中用於運算出一校正雙耳輸出信號的該裝置受組配以執行該運算，使得

$$\hat{X}_2 = P_2 \cdot X_d$$

其中 X_d 是該經去相關信號， \hat{X}_2 是一 2×1 的向量，其元素相對應於該校正雙耳信號的第一及第二聲道，且 P_2 是表示該第二演示指示且具有一 2×2 大小的一第二演示矩陣，即

$$P_2 = \begin{pmatrix} P_L \sin(\beta + \alpha) \exp(j \frac{\arg(c_{12})}{2}) \\ P_R \sin(\beta - \alpha) \exp(-j \frac{\arg(c_{12})}{2}) \end{pmatrix}$$

其中增益 P_L 及 P_R 定義為

$$P_L = \sqrt{\frac{c_{11}}{V}}, P_R = \sqrt{\frac{c_{22}}{V}}$$

其中 c_{11} 及 c_{22} 是該初步雙耳信號的一 2×2 協方差矩陣

C 的係數，即

$$C = \tilde{G} D E D^* \tilde{G}^*$$

其中 V 是一純量，即 $V = W E W^* + \varepsilon$ ， W 是大小為 $1 \times N$ 的一單降混矩陣，其係數惟獨由 d_i^x 來決定， $D = \begin{pmatrix} D^1 \\ D^2 \end{pmatrix}$ ，且 \tilde{G} 為

$$\tilde{G}^{l,m} = \begin{pmatrix} P_L^1 \exp(j \frac{\phi^1}{2}) & P_L^{l,m,2} \exp(j \frac{\phi^2}{2}) \\ P_R^1 \exp(-j \frac{\phi^1}{2}) & P_R^2 \exp(-j \frac{\phi^2}{2}) \end{pmatrix}$$

其中用於評估該實際雙耳聲道內相干性值的該裝置受組配以判定該實際雙耳聲道內相干性值為

$$\rho_C = \min\left(\frac{|c_{12}|}{\sqrt{c_{11}c_{22}}}, 1\right)$$

其中用於判定該目標雙耳聲道內相干性值的該裝置受組配以判定該目標雙耳聲道內相干性值為

$$\rho_T = \min\left(\frac{|f_{12}|}{\sqrt{f_{11}f_{22}}}, 1\right), \text{ 且}$$

其中用於設定該混合率的該裝置受組配以判定旋轉角 α 及 β ，根據

$$\alpha = \frac{1}{2}(\arccos(\rho_T) - \arccos(\rho_C)) ,$$

$$\beta = \arctan\left(\tan(\alpha)\frac{P_R - P_L}{P_L + P_R}\right) ,$$

其中 ϵ 表示用於避免分別被 0 除的一較小常數。

7. 如申請專利範圍第 1 項所述之設備，其中用於運算出該初步雙耳信號的該裝置受組配以執行該運算，使得

$$\hat{X}_1 = \mathbf{G} \cdot X$$

其中 X 是一 2×1 的向量，其元素相對應於該立體聲降混信號的該第一及第二聲道， \hat{X}_1 是一 2×1 向量，其元素相對應於該初步雙耳信號的該第一及第二聲道， \mathbf{G} 是表示該第一演示指示且具有一 2×2 大小的第一演示矩陣，即

$$\mathbf{G} = AED^*(DED^*)^{-1} ,$$

其中 E 是惟獨由該目標內互相關資訊及該目標位準資訊來決定的一矩陣；

D 是一 $2 \times N$ 的矩陣，其係數 d_{ij} 惟獨由該降混資訊來決定，其中 d_{1j} 指示音訊信號 j 已混合至該立體聲降混信號之該第一聲道中的該程度，且 d_{2j} 定義音訊信號 j 已混合至該立體聲輸出信號之該第二聲道中的程度；

A 是使該等音訊信號分別與該雙耳輸出信號之該等第一及第二聲道相關的一目標雙耳演示矩陣，且其惟獨由該演示資訊及該等 HRTF 參數來判定，

其中用於運算出一校正雙耳輸出信號的該裝置受組配以執行該運算，使得

$$\hat{X}_2 = P \cdot X_d$$

其中 X_d 是該經去相關信號， \hat{X}_2 是一 2×1 的向量，其的該等係數相對應於該校正雙耳信號的第一及第二聲道，且 P 是表示該第二演示指示且具有一 2×2 大小的第一二演示矩陣，且被決定以使得 $PP^* = \Delta R$ ，其中 $\Delta R = AEA^* - G_0 DED^* G_0^*$ ，而 $G_0 = G$ 。

8. 如申請專利範圍第 1 項所述之設備，其中用於運算出該初步雙耳信號的該裝置受組配以執行該運算，使得

$$\hat{X}_1 = G \cdot X$$

其中 X 是一 2×1 的向量，其元素相對應於該立體聲降混信號的該第一及第二聲道， \hat{X}_1 是一 2×1 的向量，其元素相對應於該初步雙耳信號的該第一及第二聲道， G 是表示該第一演示指示且具有一 2×2 大小的第一演示矩陣，即

$$G = (G_0 DED^* G_0^*)^{-1} (G_0 DED^* G_0^* AEA^* G_0 DED^* G_0^*)^{1/2} (G_0 DED^* G_0^*)$$

\mathbf{G}_0^{-1} 其中 $\mathbf{G}_0 = \mathbf{A} \mathbf{E} \mathbf{D}^* (\mathbf{D} \mathbf{E} \mathbf{D}^*)^{-1}$

其中 \mathbf{E} 是惟獨由該目標內互相關資訊及該目標位準資訊來決定出的一矩陣；

\mathbf{D} 是一 $2 \times N$ 的矩陣，其係數 d_{ij} 惟獨由該降混資訊來決定出，其中 d_{1j} 指示音訊信號 j 已混合至該立體聲降混信號之該第一聲道中的該程度，且 d_{2j} 定義音訊信號 j 已混合至該立體聲輸出信號之該第二聲道中的程度；

\mathbf{A} 是使該等音訊信號分別與該雙耳輸出信號之該等第一及第二聲道相關的一目標雙耳演示矩陣，且其惟獨由該演示資訊及該等 HRTF 參數來判定，

其中用於運算出一校正雙耳輸出信號的該裝置受組配以執行該運算，使得

$$\hat{\mathbf{X}}_2 = \mathbf{P} \cdot \mathbf{X}_d$$

其中 \mathbf{X}_d 是該經去相關信號， $\hat{\mathbf{X}}_2$ 是一 2×1 的向量，其元素相對應於該校正雙耳信號的第一及第二聲道，且 \mathbf{P} 是表示該第二演示指示且具有一 2×2 大小的一第二演示矩陣，且其被決定以使得 $\mathbf{P} \mathbf{P}^* = (\mathbf{A} \mathbf{E} \mathbf{A}^* - \mathbf{G} \mathbf{D} \mathbf{E} \mathbf{D}^* \mathbf{G}^*) / V$ ，其中 V 是一純量。

9. 如先前申請專利範圍之任何一項所述之設備，其中該降混資訊是時間相依，且該目標位準資訊及該目標內互相關資訊是時間及頻率相依。

10. 一種用於將一多聲道音訊信號雙耳演示為一雙耳輸出信號的方法，該多聲道音訊信號包含多個音訊信號被降

混的一立體聲降混信號，且包含旁側資訊，該旁測資訊包含一指示出對於每一音訊信號，該各自音訊信號已分別混合至該立體聲降混信號的第一聲道及一第二聲道中的程度的降混資訊，及該等音訊信號的目標位準資訊，及描述在該等音訊信號之音訊信號對之間的類似性的目標內互相關資訊，該方法包含以下步驟：

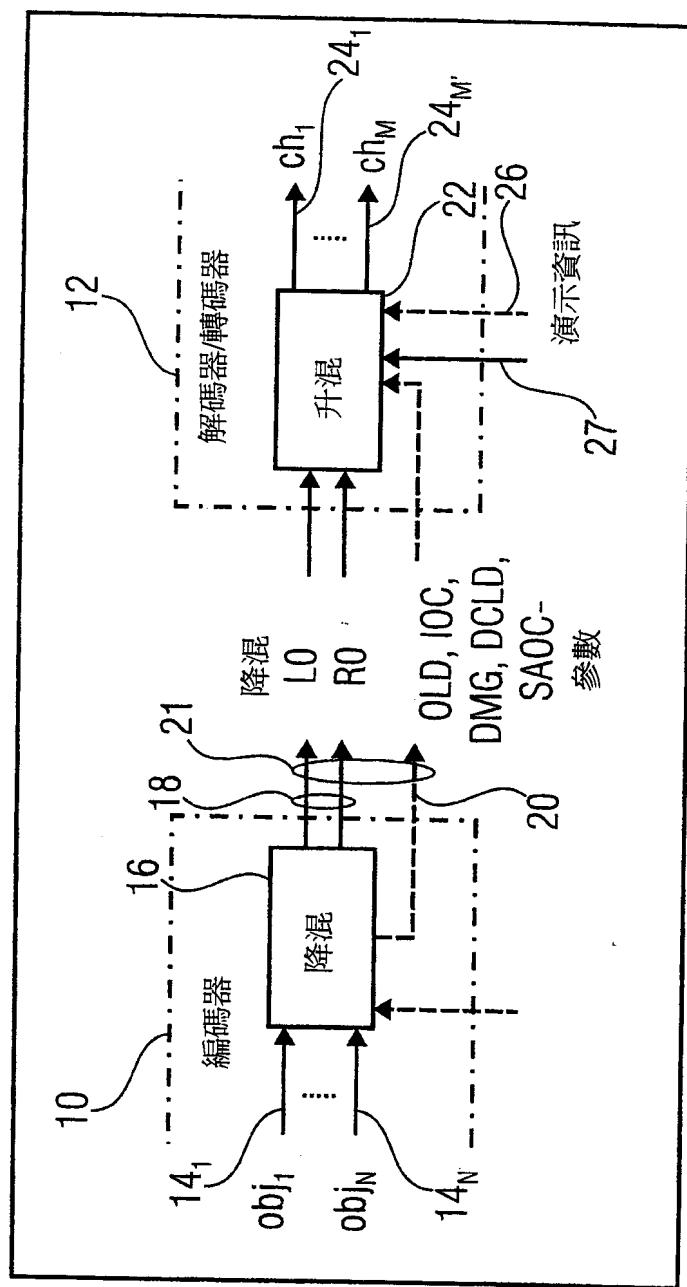
○ 基於一第一演示指示自該立體聲降混信號之該第一及第二聲道來運算出一初步雙耳信號，該第一演示指示依據該目標內互相關資訊、該目標位準資訊、該降混資訊、使每一音訊信號相關於一虛擬揚聲器位置的演示資訊及頭部相關轉換函數(HRTF)參數而定；

○ 產生一經去相關信號，該經去相關信號作為對該立體聲降混信號之該第一及第二聲道的一單降混的一感知等效物，且然而是與該單降混去相關；

○ 根據一第二演示指示自該經去相關信號運算出一校正雙耳信號，該第二演示指示依據該目標內互相關資訊、該目標位準資訊、該降混資訊、該演示資訊及該等HRTF參數而定；及

將該初步雙耳信號與該校正雙耳信號相混合，以獲得該雙耳輸出信號。

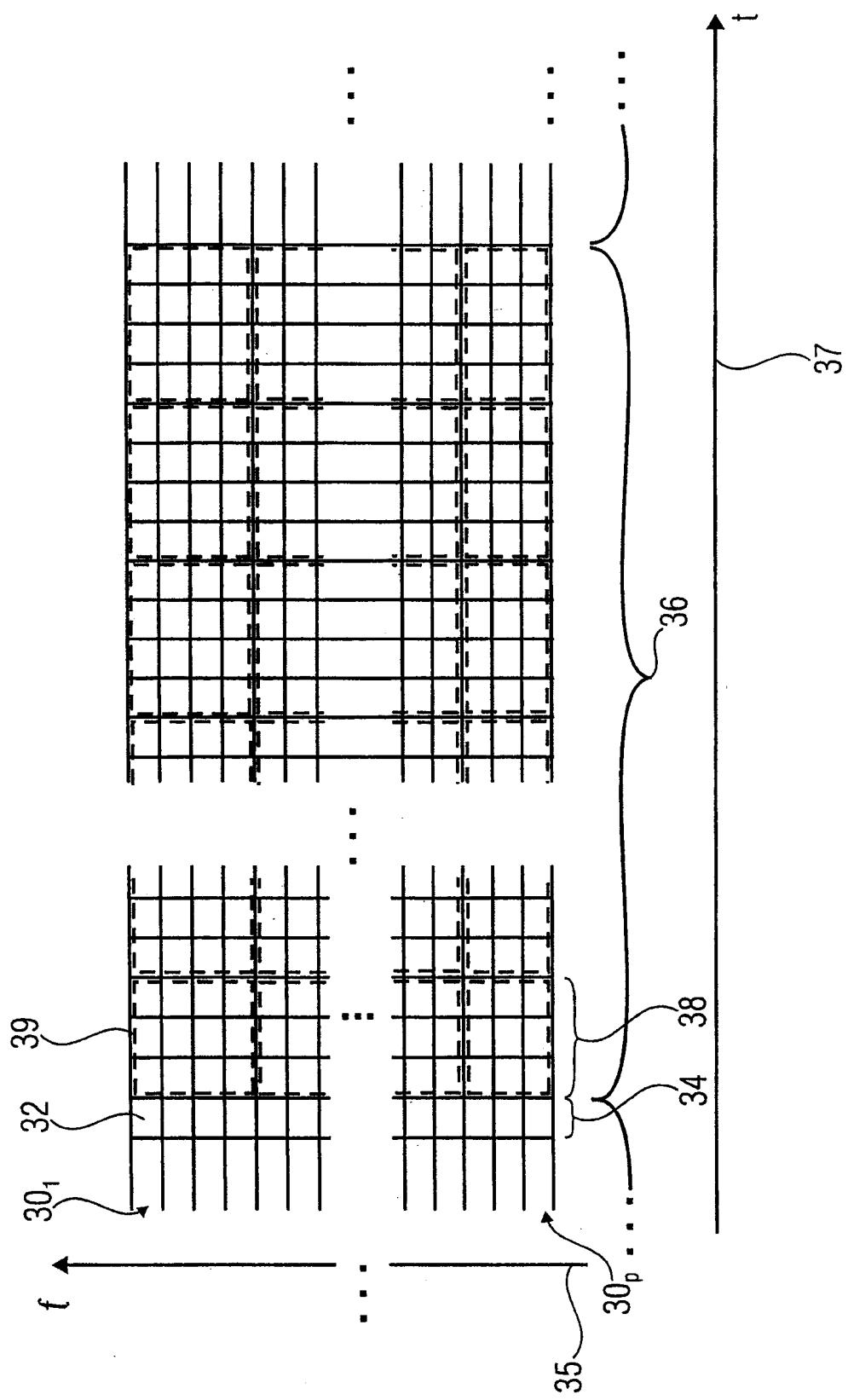
11. 一種具有複數指令的電腦程式，該等指令在於一電腦上執行時，用於執行根據申請專利範圍第 10 項所述的方法。



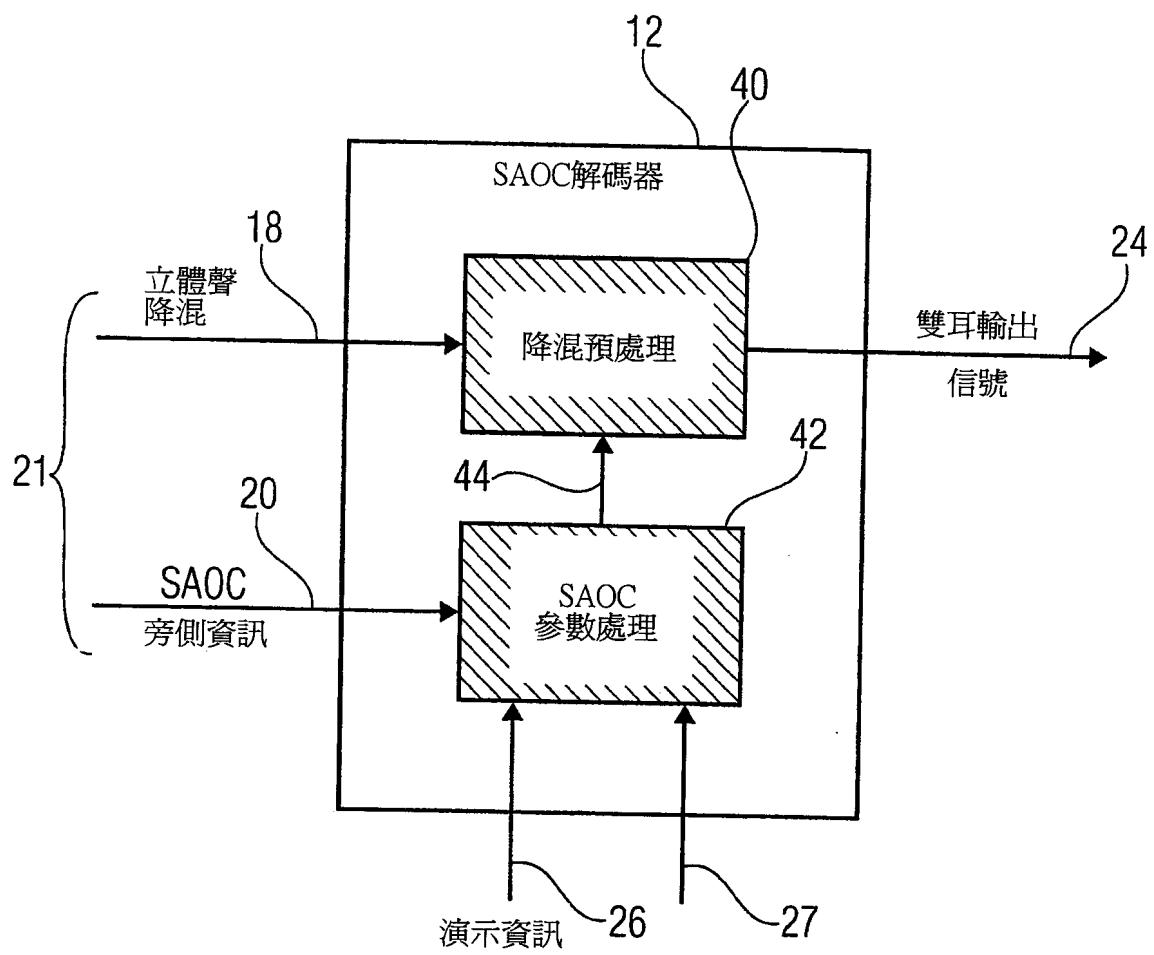
第1圖

201036464

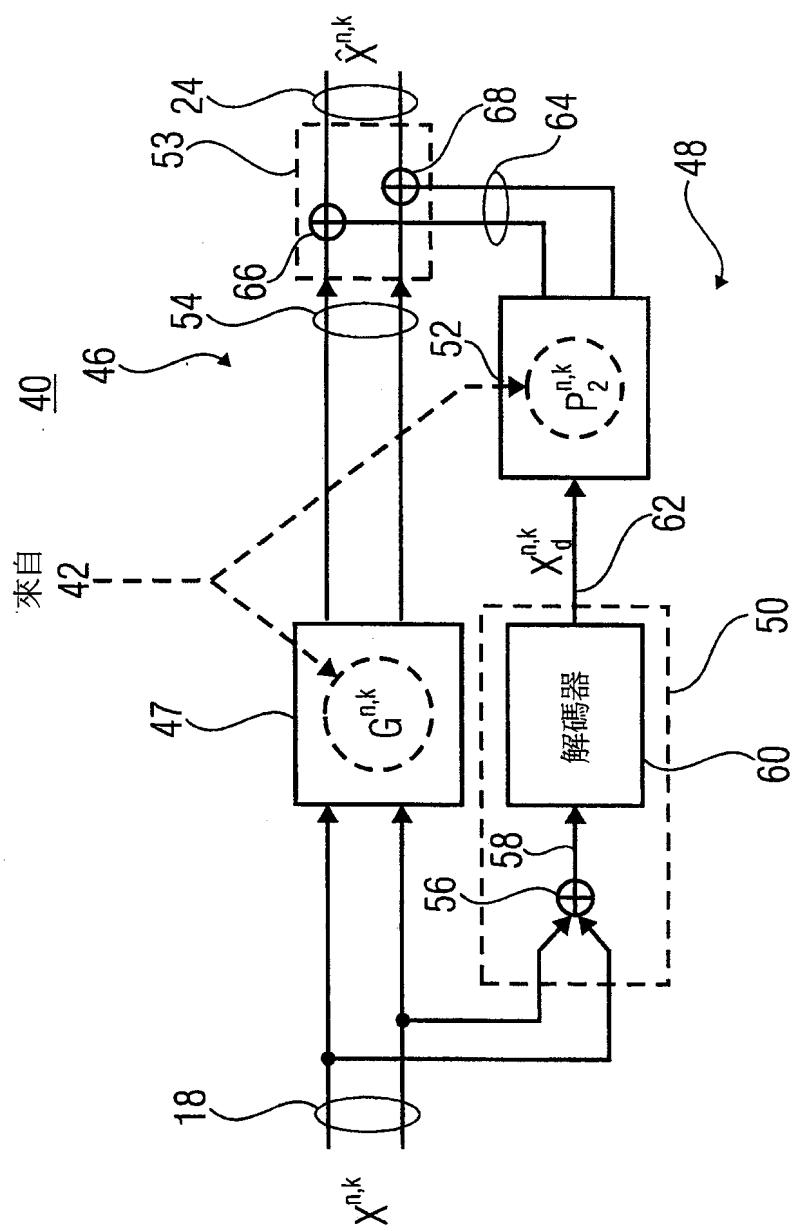
2/6



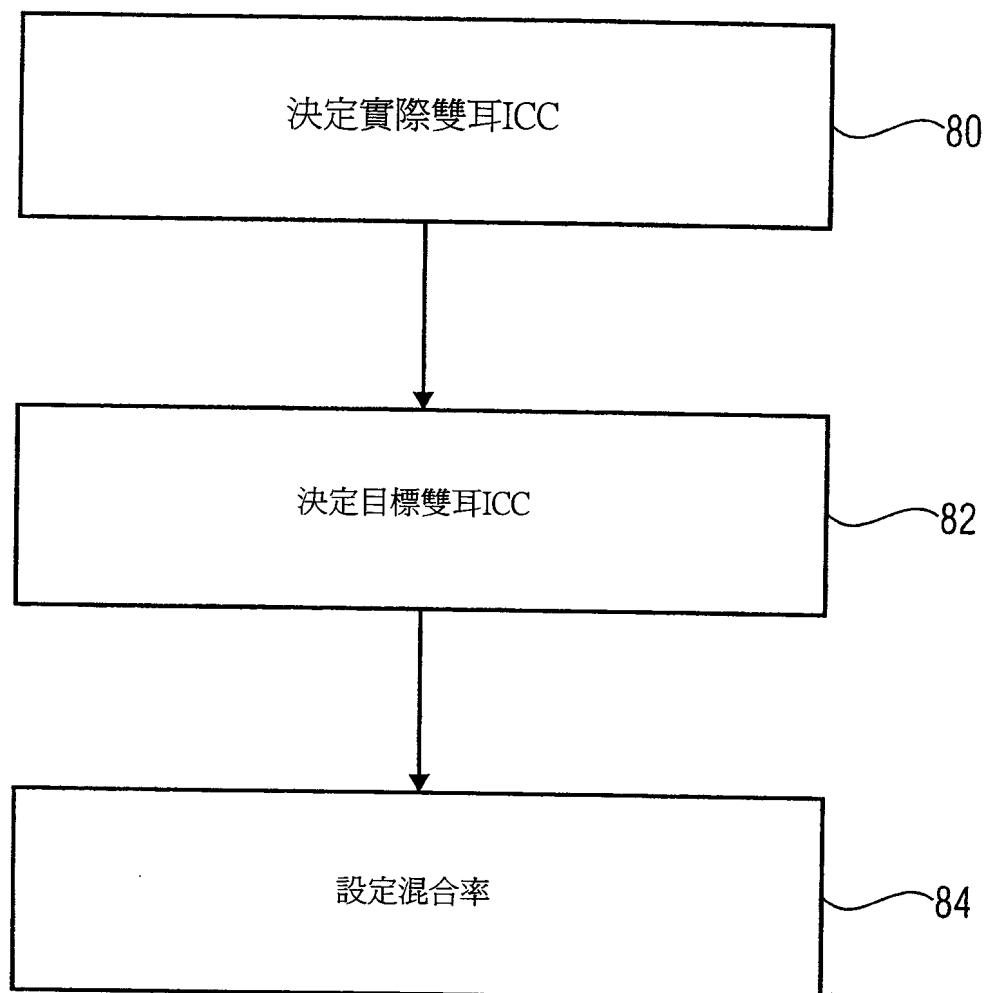
第2圖



第3圖



第4圖



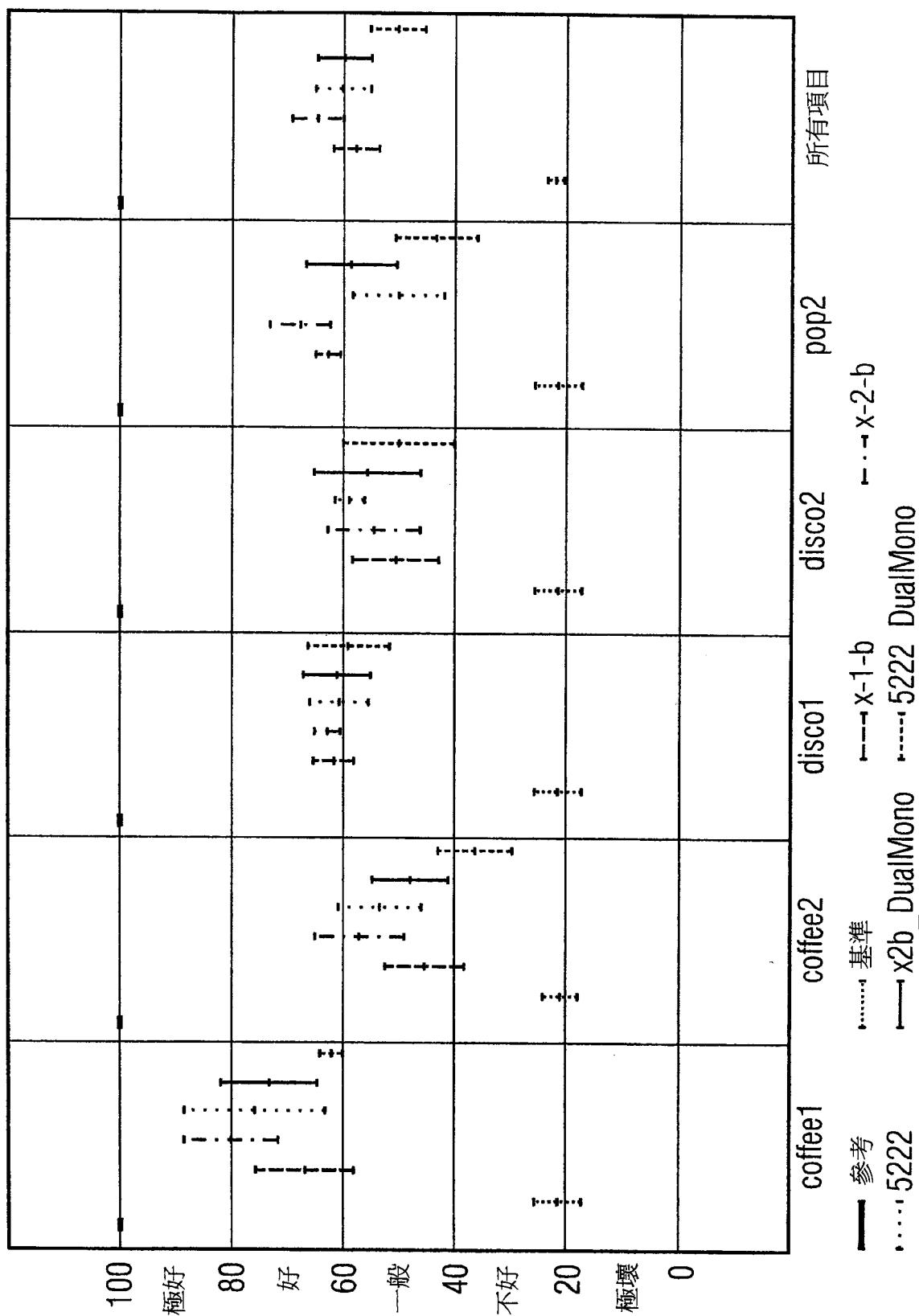
第5圖

201036464

6/6

第6圖

平均及95%可信區間



四、指定代表圖：

(一)本案指定代表圖為：第（ 4 ）圖。

(二)本代表圖之元件符號簡單說明：

18... 降混信號	53... 混合階段
24... 雙耳輸出信號	54... 初步雙耳輸出信號
40... 降混預處理單元	56... 相加器
42... SAOC 參數處理單元	58... 多聲道降混
46... 乾式路徑	60... 去相關器
47... 乾式演示單元	62... 經去相關信號
48... 濕式路徑	64... 校正雙耳輸出信號
50... 去相關信號產生器	66/68... 相加器
52... 濕式演示單元	

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：