



(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 201423728 A

(43)公開日：中華民國 103 (2014) 年 06 月 16 日

(21)申請案號：102136011

(22)申請日：中華民國 102 (2013) 年 10 月 04 日

(51)Int. Cl. : **G10L19/008 (2013.01)**

(30)優先權：2012/10/05 美國 61/710,128
2013/05/13 歐洲專利局 13167485.5

(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

紐倫堡大學(德國) FRIEDRICH-ALEXANDER-UNIVERSITAET ERLANGEN-
NUERNBERG (DE)

德國

(72)發明人：迪斯曲 薩斯洽 DISCH, SASCHA (DE)；福契斯 哈拉德 FUCHS, HARALD
(DE)；帕露斯 喬尼 PAULUS, JOUNI (FI)；泰倫堤夫 黎恩 THERENTIV, LEON
(DE)；賀穆斯 奧利薇 HELLMUTH, OLIVER (DE)；希瑞 喬根 HERRE,
JUERGEN (DE)

(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：18 項 圖式數：18 共 80 頁

(54)名稱

用於反向相容多重解析度空間音訊物件編碼之編碼器、解碼器及方法

ENCODER, DECODER AND METHODS FOR BACKWARD COMPATIBLE MULTI-RESOLUTION
SPATIAL-AUDIO-OBJECT-CODING

(57)摘要

提供一種用於產生一未經混合音訊信號之解碼器，該未經混合音訊信號包含多個未經混合音訊通道。此外，提供一種編碼器及一種經編碼之音訊信號。該解碼器包含一未混合資訊判定器，該未混合資訊判定器用於藉由接收關於至少一音訊物件信號的第一參數旁側資訊及關於該至少一音訊物件信號之第二參數旁側資訊來判定未混合資訊，其中該第二參數旁側資訊的頻率解析度高於該第一參數旁側資訊的頻率解析度。此外，該解碼器包含一未混合模組，該未混合模組用於將該未混合資訊應用於指示至少一音訊物件信號之一下降混合的一下降混合信號，以獲得包含該多個未經混合音訊通道的一未經混合音訊信號。該未混合資訊判定器經組配以藉由修改該第一參數資訊及該第二參數資訊以獲得經修改參數資訊來判定該未混合資訊，使得該經修改參數資訊具有高於該第一頻率解析度的一頻率解析度。

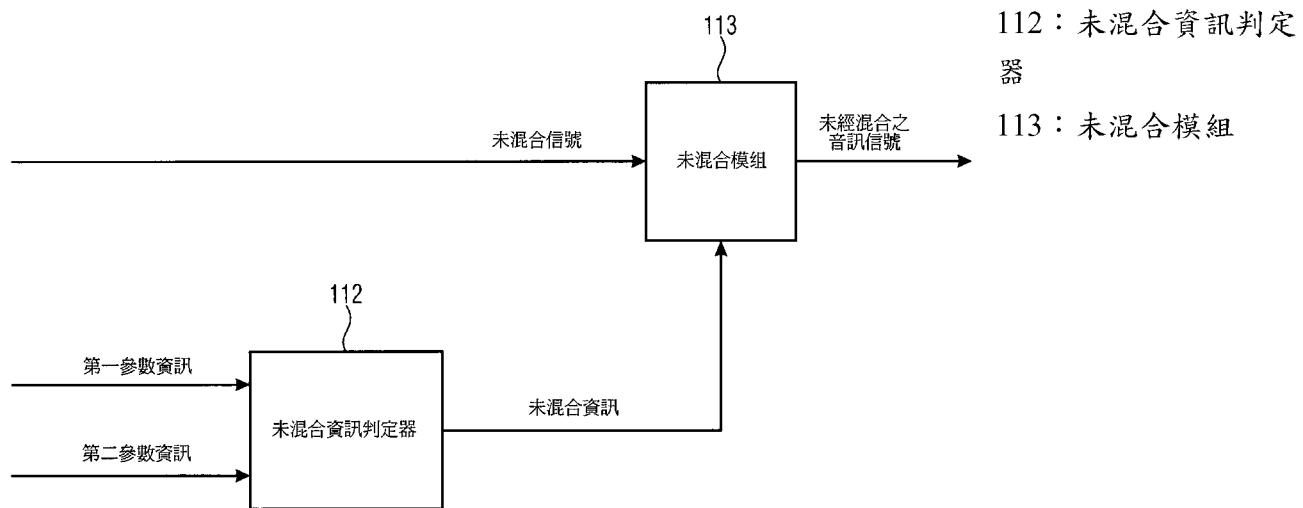


圖1a



(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 201423728 A

(43)公開日：中華民國 103 (2014) 年 06 月 16 日

(21)申請案號：102136011

(22)申請日：中華民國 102 (2013) 年 10 月 04 日

(51)Int. Cl. : **G10L19/008 (2013.01)**

(30)優先權：2012/10/05 美國 61/710,128
2013/05/13 歐洲專利局 13167485.5

(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

紐倫堡大學(德國) FRIEDRICH-ALEXANDER-UNIVERSITAET ERLANGEN-
NUERNBERG (DE)

德國

(72)發明人：迪斯曲 薩斯洽 DISCH, SASCHA (DE)；福契斯 哈拉德 FUCHS, HARALD
(DE)；帕露斯 喬尼 PAULUS, JOUNI (FI)；泰倫堤夫 黎恩 THERENTIV, LEON
(DE)；賀穆斯 奧利薇 HELLMUTH, OLIVER (DE)；希瑞 喬根 HERRE,
JUERGEN (DE)

(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：18 項 圖式數：18 共 80 頁

(54)名稱

用於反向相容多重解析度空間音訊物件編碼之編碼器、解碼器及方法

ENCODER, DECODER AND METHODS FOR BACKWARD COMPATIBLE MULTI-RESOLUTION
SPATIAL-AUDIO-OBJECT-CODING

(57)摘要

提供一種用於產生一未經混合音訊信號之解碼器，該未經混合音訊信號包含多個未經混合音訊通道。此外，提供一種編碼器及一種經編碼之音訊信號。該解碼器包含一未混合資訊判定器，該未混合資訊判定器用於藉由接收關於至少一音訊物件信號的第一參數旁側資訊及關於該至少一音訊物件信號之第二參數旁側資訊來判定未混合資訊，其中該第二參數旁側資訊的頻率解析度高於該第一參數旁側資訊的頻率解析度。此外，該解碼器包含一未混合模組，該未混合模組用於將該未混合資訊應用於指示至少一音訊物件信號之一下降混合的一下降混合信號，以獲得包含該多個未經混合音訊通道的一未經混合音訊信號。該未混合資訊判定器經組配以藉由修改該第一參數資訊及該第二參數資訊以獲得經修改參數資訊來判定該未混合資訊，使得該經修改參數資訊具有高於該第一頻率解析度的一頻率解析度。

發明摘要

※ 申請案號：102136011

※ 申請日：102.10.04

※IPC 分類：G10L 19/08 (2013.01)

【發明名稱】(中文/英文)

用於反向相容多重解析度空間音訊物件編碼之編碼器、解碼器及方法
ENCODER, DECODER AND METHODS FOR BACKWARD COMPATIBLE
MULTI-RESOLUTION SPATIAL-AUDIO-OBJECT-CODING

【中文】

提供一種用於產生一未經混合音訊信號之解碼器，該未經混合音訊信號包含多個未經混合音訊通道。此外，提供一種編碼器及一種經編碼之音訊信號。該解碼器包含一未混合資訊判定器，該未混合資訊判定器用於藉由接收關於至少一音訊物件信號的第一參數旁側資訊及關於該至少一音訊物件信號之第二參數旁側資訊來判定未混合資訊，其中該第二參數旁側資訊的頻率解析度高於該第一參數旁側資訊的頻率解析度。此外，該解碼器包含一未混合模組，該未混合模組用於將該未混合資訊應用於指示至少一音訊物件信號之一下降混合的一下降混合信號，以獲得包含該多個未經混合音訊通道的一未經混合音訊信號。該未混合資訊判定器經組配以藉由修改該第一參數資訊及該第二參數資訊以獲得經修改參數資訊來判定該未混合資訊，使得該經修改參數資訊具有高於該第一頻率解析度的一頻率解析度。

【英文】

A decoder for generating an un-mixed audio signal comprising a plurality of un-mixed audio channels is provided. Moreover, an encoder and an encoded audio signal is provided. The decoder comprises an un-mixing-information determiner for determining un-mixing information by receiving first parametric side information on the at least one audio object signal and second parametric side information on the at least one audio object signal, wherein the frequency resolution of the second parametric side information is higher than the frequency resolution of the first parametric side information. Moreover, the decoder comprises an un-mix module for applying the un-mixing information on a downmix signal, indicating a downmix of at least one audio object signal, to obtain an un-mixed audio signal comprising the plurality of un-mixed audio channels. The un-mixing-information determiner is configured to determine the un-mixing information by modifying the first parametric information and the second parametric information to obtain modified parametric information, such that the modified parametric information has a frequency resolution which is higher than the first frequency resolution.

【代表圖】

【本案指定代表圖】：第（1a）圖。

【本代表圖之符號簡單說明】：

112...未混合資訊判定器

113...未混合模組

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：

(無)

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)

用於反向相容多重解析度空間音訊物件編碼之編碼器、解碼器及方法

ENCODER, DECODER AND METHODS FOR BACKWARD
COMPATIBLE MULTI-RESOLUTION
SPATIAL-AUDIO-OBJECT-CODING

【技術領域】

發明領域

[0001]本發明係關於音訊信號編碼、音訊信號解碼及音訊信號處理，且詳言之，係關於用於反向相容多重解析度空間音訊物件編碼(SAOC)之編碼器、解碼器及方法。

【先前技術】

發明背景

[0002]在現代數位音訊系統中，主要傾向係允許在接收器側上對所傳輸內容之音訊物件相關修改。此等修改包括音訊信號之所選擇部分的增益修改及/或在經由空間分佈之揚聲器之多通道播放狀況下專用音訊物件的空間重定位。此情形可藉由將音訊內容之不同部分個別地遞送至不同揚聲器來達成。

[0003]換言之，在音訊處理、音訊傳輸及音訊儲存之技術中，存在增加之需要以允許關於物件導向式音訊內容播放的使用者互動且亦存在需求以利用多通道播放之擴展可

能性以個別地呈現音訊內容或其部分以便改良聽取印象。藉由此情形，多通道音訊內容之使用帶來對使用者之顯著改良。舉例而言，可獲得三維聽取印象，其帶來娛樂應用中改良之使用者滿意度。然而，多通道音訊內容亦在專業環境中(例如，在電話會議應用)中係有用的，此係因為發話人可懂度可藉由使用多通道音訊播放來改良。另一可能應用係將音樂片段給予給收聽者以個別地調整不同部分(亦稱為「音訊物件」)或曲目(諸如，人聲部分或不同樂器)的播放層級及/或空間位置。使用者出於個人口味、為了易於轉譯來自音樂片段之一或多個部分、教育用途、伴唱機、排演等原因而可執行此調整。

[0004](例如)呈脈碼調變(PCM)資料或甚至經壓縮音訊格式之形式的所有數位多通道或多物件音訊內容的直接離散傳輸需要極高位元率。然而，亦需要以位元率有效方式來傳輸並儲存音訊資料。因此，希望接受音訊品質與位元率要求之間的合理取捨以便避免由多通道/多物件應用引起的額外資源負載。

[0005]最近，在音訊編碼之領域中，多通道/多物件音訊信號之位元率有效傳輸/儲存的參數技術已藉由(例如)動畫專家組(MPEG)及其他者引入。一個實例係作為通道導向式方法[MPS, BCC]的 MPEG 環繞立體聲(MPS)，或作為物件導向式方法[JSC, SAOC, SAOC1, SAOC2]的 MPEG 空間音訊物件編碼(SAOC)。另一物件導向式方法稱為「消息源分離」[ISS1, ISS2, ISS3, ISS4, ISS5, ISS6]。此等技術旨在

在通道/物件及描述所傳輸/所儲存音訊場景及/或音訊場景中音訊源物件的額外旁側資訊之下降混合基礎上重建所要輸出音訊場景或所要音訊源物件。

[0006]此等系統中通道/物件相關旁側資訊的估計及應用以時間頻率選擇性方式來進行。因此，此等系統使用時間頻率變換，諸如離散傅立葉變換(DFT)、短時間傅立葉變換(STFT)或類似於正交鏡相濾波器(QMF)組的濾波器組等。此等系統之基本原理使用 MPEG SAOC 的實例描繪於圖 4 中。

[0007]在 STFT 之狀況下，時間維度藉由時間區塊數字來表示，且頻譜維度藉由頻譜係數(「區間」)編號來捕獲。在 QMF 之狀況下，時間維度藉由時槽編號來表示，且頻譜維度藉由子頻帶編號來捕獲。若 QMF 之頻譜解析度藉由第二濾波器級之後續應用來改良，則整個濾波器組被稱為混合式 QMF 且精細解析度子頻帶被稱為混合式子頻帶。

[0008]如上文已提及，在 SAOC 中，通用處理以時間頻率選擇性方式進行，且可在每一頻帶內描述如下：

- N 個輸入音訊物件信號 $s_1 \dots s_N$ 使用由元素 $d_{1,1} \dots d_{N,P}$ 組成之下降混合矩陣下降混合至 P 個通道 $x_1 \dots x_P$ 作為編碼器處理的部分。此外，編碼器提取旁側資訊，旁側資訊描述輸入音訊物件之特性(旁側資訊估計器(SIE)模組)。對於 MPEG SAOC，物件功率關於彼此之關係係此旁側資訊之最基本形式。

- 傳輸/儲存下降混合信號及旁側資訊。為此目的，下

降混合音訊信號可(例如)使用諸如 MPEG-1/2 層 II 或 III(又名 mp3)、MPEG-2/4 進階音訊編碼(AAC)等熟知感知音訊編碼器經壓縮。

- 在接收端，解碼器概念上試圖使用所傳輸之旁側資訊自(經解碼)下降混合信號來復原原始物件信號(「物件分離」)。此等經近似物件信號 $\hat{s}_1 \dots \hat{s}_N$ 接著使用由圖 4 中之係數 $r_{1,1} \dots r_{N,M}$ 描述的呈現矩陣經混合成藉由 M 個音訊輸出通道 $\hat{y}_1 \dots \hat{y}_M$ 表示的目標場景。所要目標場景在極端狀況下可係來自混合物之僅一個源信號的呈現(源分離情境)，但亦係由所傳輸之物件組成的其他任意聲學場景。舉例而言，輸出可係單通道、2 通道立體或 5.1 多通道目標場景。

[0009]基於時間頻率之系統可利用具有靜態時間及頻率解析度的時間頻率(t/f)變換。選定某固定之 t/f 解析度柵格通常涉及時間解析度與頻率解析度之間的取捨。

[0010]可關於音訊信號混合物中之典型物件信號的實例來論證固定之 t/f 解析度之效應。舉例而言，音調聲音之頻譜展現具有基本頻率及若干泛音之諧波相關的結構。此等信號之能量集中於某頻率區處。對於此等信號，所利用 t/f 表示之高頻率解析度對於分離窄頻音調頻譜區與信號混合物係有益的。相反，類似於鼓聲之瞬變信號常常具有獨特時間結構：實質能量僅存在歷時短時間週期，且分佈於廣泛範圍之頻率上。對於此等信號，所利用之 t/f 表示之高時間解析度對於分離瞬變信號部分與信號混合物係有利的。

[0011]自標準 SAOC 表示獲得之頻率解析度限於參數頻帶之數目，從而在標準 SAOC 中具有最大值 28。參數頻帶獲得自由 64 頻帶 QMF 分析組成的混合式 QMF 組，其中最低頻帶上之額外混合式濾波級將此等頻帶進一步劃分成至多 4 個複合子頻帶。所獲得之頻帶經分組成參數頻帶，從而模擬人類聽覺系統的關鍵頻帶解析度。分組允許將所需要之旁側資訊資料速率減小至在實際應用中可經有效地進行處置的大小。

[0012]當前音訊物件編碼方案給予 SAOC 處理之僅時間頻率選擇性的有限可變性。舉例而言，MPEG SAOC [SAOC] [SAOC1] [SAOC2]限於可藉由使用所謂混合式鏡面濾波器組(混合式 QMF)及其後續至參數頻帶之分組獲得的時間頻率解析度。因此，標準 SAOC 中之物件復原常常遭受混合式 QMF 之粗糙頻率解析度，從而導致來自其他音訊物件之可聽經調變串擾(例如，語音之雙重講話人爲效應或音樂中之聽覺粗糙人爲效應)。

[0013]給定合理地低之資料速率，現有系統產生合理之分離品質。主要問題係音調聲音之完全分離的不足頻率解析度。此情形經展現爲其他物件的包圍一物件之音調分量的「光暈」。此情形感知地觀測爲粗糙度或類似於聲碼器的人爲效應。此光暈之有害效應可藉由增加參數頻率解析度來減小。注意到，等於或高於 512 個頻帶(以 44.1 kHz 取樣速率)之解析度足夠產生測試信號中感知上顯著改良的分離。此高參數解析度情況下之問題係需要之旁側資訊的量

大幅度地增加至不切實際的量。此外，與現有標準 SAOC 系統之相容性將失去。

[0014]因此高度重視的是，是否可提供教示如何克服目前技術之上述約束的概念。

【發明內容】

發明概要

[0015]本發明之目標係提供用於音訊物件編碼之此等改良之概念。本發明之目標藉由以下各者來解決：如技術方案 1 之解碼器、如技術方案 9 之編碼器、如技術方案 14 之經編碼音訊信號、如技術方案 15 之系統、如技術方法 16 之用於解碼之方法、如技術方案 17 之用於編碼之方法，及如技術方案 18 之電腦程式。

[0016]與目前技術 SAOC 相反，本發明之實施例提供一種頻譜參數化，使得

- 發源於標準 SAOC 編碼器之 SAOC 參數位元串流仍可以與藉由標準解碼器獲得之感知品質相當之感知品質藉由增強型解碼器來解碼，
- 增強型 SAOC 參數位元串流可以與藉由標準 SAOC 位元串流可獲得之品質相當之品質藉由標準 SAOC 解碼器解碼，
- 增強型 SAOC 參數位元串流可藉由增強型解碼器以最佳品質解碼，
- 增強型 SAOC 解碼器(例如)視可用之計算資源可動態地調整增強層級，

- 標準及增強型 SAOC 參數位元串流(例如)在多點控制單元(MCU)情境下可混合成一個共同位元串流，該共同位元串流可以藉由解碼器提供之品質藉由標準解碼器或增強型解碼器解碼，且
- 額外參數化為緊湊的。

[0017]對於上文提及之性質，較佳的是具有參數化，該參數化藉由標準 SAOC 解碼器理解，而且亦允許以較高頻率解析度高效地遞送資訊。基本時間頻率表示之解析度判定增強之最大效能。本發明此處界定一種用於以一緊湊且允許反向相容解碼的方式遞送增強型高頻率資訊的方法。

[0018]增強型 SAOC 感知品質可(例如)藉由動態地調適濾波器組或變換之時間/頻率解析度來獲得，該濾波器組或變換用以估計或用以合成對輸入音訊物件之特定性質的音訊物件提示。舉例而言，若音訊物件在某時間跨度期間係準穩定的，則對粗糙時間解析度及精細頻率解析度有益地執行參數估計及合成。若音訊物件含有瞬變或在某時間跨度期間係非穩定的，則使用精細時間解析度及粗糙頻率解析度有利地進行參數估計及合成。藉此，濾波器組或變換之動態調適允許

- 準穩定信號之頻譜分離中的高頻率選擇性以便避免物件間串擾，以及
- 物件開始或瞬變事件的高時間精度以便使前回聲及後回聲最小化。

[0019]同時，傳統 SAOC 品質可藉由將標準 SAOC 資

料映射於藉由本發明之反向相容信號自適應性變換提供的時間頻率柵格上而獲得，該變換取決於描述物件信號特性的旁側資訊。

[0020]使用一個共同變換能夠解碼標準 SAOC 資料及增強型 SAOC 資料兩者啓用涵蓋標準 SAOC 資料與新穎增強型 SAOC 資料之混合的應用之直接反向相容性。其亦允許優於標準品質之時間頻率選擇性增強。

[0021]所提供實施例不限於任何特定時間頻率變換，而是可藉由提供足夠高之頻率解析度的任何變換應用。文獻描述藉由切換式時間頻率解析度至基於離散傅立葉變換 (DFT) 之濾波器組的應用。在此方法中，時域信號經再分成亦可重疊之較短區塊。每一較短區塊中之信號藉由開窗函數來加權(正常地在中間具有大值，且在兩個末端處遞減至零)。最終，經加權信號藉由所選擇之變換(此處，藉由 DFT 之應用)變換至頻域。

[0022]提供一種用於產生一未經混合音訊信號之解碼器，該未經混合音訊信號包含多個未經混合音訊通道。解碼器包含一未混合資訊判定器，該未混合資訊判定器用於藉由接收關於至少一音訊物件信號的第一參數旁側資訊及關於該至少一音訊物件信號之第二參數旁側資訊來判定未混合資訊，其中該第二參數旁側資訊的頻率解析度高於該第一參數旁側資訊的頻率解析度。此外，該解碼器包含一未混合模組，該未混合模組用於將該未混合資訊應用於指示至少一音訊物件信號之一下降混合的一下降混合信號，

以獲得包含該多個未經混合音訊通道的一未經混合音訊信號。該未混合資訊判定器經組配以藉由修改該第一參數資訊及該第二參數資訊以獲得經修改參數資訊來判定該未混合資訊，使得該經修改參數資訊具有高於該第一頻率解析度的一頻率解析度。

[0023]此外，提供一種用於編碼一或多個輸入音訊物件信號的編碼器。編碼器包含下降混合單元，該下降混合單元用於下降混合一或多個輸入音訊物件信號以獲得一或多個下降混合信號。此外，編碼器包含參數旁側資訊產生器，該參數旁側資訊產生器用於產生關於至少一音訊物件信號之第一參數旁側資訊及關於至少一音訊物件信號的第二參數旁側資訊，使得第二參數旁側資訊之頻率解析度高於第一參數旁側資訊的頻率解析度。

[0024]此外，提供一種經編碼音訊信號。經編碼音訊信號包含一指示一或多個輸入音訊物件信號之下降混合的下降混合部分，以及包含關於至少一音訊物件信號之第一參數旁側資訊及關於至少一音訊物件信號之第二參數旁側資訊的參數旁側資訊部分。第二參數旁側資訊的頻率解析度高於第一參數旁側資訊之頻率解析度。

[0025]此外，提供一種系統。系統包含一如上文所描述之編碼器及如上文所描述的解碼器。編碼器經組配以藉由以下操作來編碼一或多個輸入音訊物件信號：獲得指示一或多個輸入音訊物件信號之下降混合的一或多個下降混合信號，獲得關於至少一音訊物件信號之第一參數旁側資訊

及獲得關於至少一音訊物件信號的第二參數旁側資訊，其中第二參數旁側資訊的頻率解析度高於第一參數旁側資訊之頻率解析度。解碼器經組配以基於一或多個下降混合信號且基於第一參數旁側資訊及第二參數旁側資訊來產生未經混合音訊信號。

[0026] 編碼器經組配以藉由以下操作來編碼一或多個輸入音訊物件信號：獲得指示一或多個輸入音訊物件信號之下降混合的一或多個下降混合信號，獲得關於至少一音訊物件信號之第一參數旁側資訊及獲得關於至少一音訊物件信號的第二參數旁側資訊，其中第二參數旁側資訊的頻率解析度高於第一參數旁側資訊之頻率解析度。解碼器經組配以基於一或多個下降混合信號且基於第一參數旁側資訊及第二參數旁側資訊來產生音訊輸出信號。

[0027] 此外，提供一種用於產生一未經混合音訊信號之方法，該未經混合音訊信號包含多個未經混合音訊通道。該方法包含：

- 藉由接收關於至少一音訊物件信號之第一參數旁側資訊及關於該至少一音訊物件信號之第二參數旁側資訊來判定未混合資訊，其中該第二參數旁側資訊之頻率解析度高於該第一參數旁側資訊的頻率解析度。以及：
- 將該未混合資訊應用於一指示至少一音訊物件信號之一下降混合的下降混合信號，以獲得包含該多個未經混合音訊通道的一未經混合音訊信號。

[0028] 判定該未混合資訊包含修改該第一參數資訊及

該第二參數資訊以獲得經修改之參數資訊，使得該經修改之參數資訊具有高於該第一頻率解析度的一頻率解析度。

[0029]此外，提供一種用於編碼一或多個輸入音訊物件信號之方法。該方法包含：

- 下降混合該一或多個輸入音訊物件信號以獲得一或多個下降混合信號。以及：
- 產生關於至少一音訊物件信號之第一參數旁側資訊及關於該至少一音訊物件信號之第二參數旁側資訊，使得該第二參數旁側資訊之頻率解析度高於該第一參數旁側資訊的頻率解析度。

[0030]此外，提供一種當在一電腦或信號處理器上執行時用於實施上述方法中之一者之電腦程式。

[0031]較佳實施例將提供於附屬請求項中。

【圖式簡單說明】

[0032]在以下內容中，參看諸圖更詳細地描述本發明之實施例，其中：

圖 1a 說明根據一實施例之解碼器，

圖 1b 說明根據另一實施例之解碼器，

圖 2a 說明根據一實施例之編碼器，

圖 2b 說明根據另一實施例之編碼器，

圖 2c 說明根據一實施例之經編碼的音訊信號，

圖 3 說明根據一實施例之系統，

圖 4 展示 SAOC 系統之概念概述的示意性方塊圖，

圖 5 展示單通道音訊信號之時間頻譜表示的示意性且

例示性圖，

圖 6 展示 SAOC 編碼器內旁側資訊之時間頻率選擇性計算的示意性方塊圖，

圖 7 說明根據實施例之反向相容表示，

圖 8 說明根據一實施例之真實參數值與低解析度平均值之間的差異曲線，

圖 9 描繪根據一實施例之提供具有增強之反向相容位元串流之增強型編碼器的高層級說明，

圖 10 說明根據實施編碼器之參數路徑的特定實施例之編碼器的方塊圖，

圖 11 描繪根據一實施例之能夠解碼標準位元串流及增強型位元串流兩者的增強型解碼器之高層級方塊圖，

圖 12 說明增強型 PSI 解碼單元之實施例的方塊圖，

圖 13 描繪根據一實施例之藉由增強型 SAOC 解碼器解碼標準 SAOC 位元串流的方塊圖，

圖 14 描繪根據一實施例之解碼器的主要功能方塊，

圖 15 說明音調及雜訊信號，且詳言之，高解析度功率頻譜及對應粗略重建，

圖 16 說明兩個實例信號的修改，詳言之，實例信號之校正因數，

圖 17 說明針對兩個實例信號的原始校正因數及基於降階線性預測的近似，且

圖 18 說明對粗略重建應用經模型化校正因數的結果。

【實施方式】

較佳實施例之詳細說明

[0033]在描述本發明之實施例之前，提供關於目前技術 SAOC 系統的更多背景內容。

[0034]圖 4 展示 SAOC 編碼器 10 及 SAOC 解碼器 12 之通用配置。SAOC 編碼器 10 接收 N 個物件(亦即，音訊信號 s_1 至 s_N)作為輸入。詳言之，編碼器 10 包含下降混合器 16，其接收音訊信號 s_1 至 s_N 且將前述音訊信號下降混合成下降混合信號 18。或者，下降混合可經外部地提供(「藝術家風格下降混合」)，且系統估計額外旁側資訊以使得所提供之下降混合與所計算的下降混合匹配。在圖 4 中，下降混合信號展示為 P 通道信號。因此，任何單($P=1$)、立體($P=2$)或多通道($P>2$)下降混合信號組配係可想到的。

[0035]在立體下降混合之狀況下，下降混合信號 18 之通道標註為 $L0$ 及 $R0$ ，在單下降混合狀況下，通道簡單地標註為 $L0$ 。為了使 SAOC 解碼器 12 能夠恢復個別物件 s_1 至 s_N ，旁側資訊估計器 17 向 SAOC 解碼器 12 提供包括 SAOC 參數的旁側資訊。舉例而言，在立體下降混合之狀況下，SAOC 參數包含物件層級差(OLD)、物件間相關(IOC)(物件間交叉相關參數)、下降混合增益值(DMG)及下降混合通道層級差(DCLD)。包括 SAOC 參數之旁側資訊 20 連同下降混合信號 18 一起形成由 SAOC 解碼器 12 接收到的 SAOC 輸出資料串流。

[0036]SAOC 解碼器 12 包含向上混合器，其接收下降混合信號 18 以及旁側資訊 20 以便將音訊信號 s_1 及 s_N 恢復

並呈現於通道 $\hat{y}_I$ 至 $\hat{y}_M$ 的任何使用者所選擇集合上，其中呈現藉由呈現輸入至 SAOC 解碼器 12 中的資訊 26 來規定。

[0037] 音訊信號 s_I 至 s_N 可在任何編碼域中(諸如，在時間或頻譜域中)輸入至編碼器 10 中。在音訊信號 s_I 至 s_N 在時域中饋入至編碼器 10 中(諸如，經 PCM 編碼)的狀況下，編碼器 10 可使用諸如混合式 QMF 組的濾波器組，以便以特定濾波器組解析度將信號傳送至頻譜域中，其中以與不同頻譜部分相關聯的若干子頻帶表示音訊信號。若音訊信號 s_I 至 s_N 已經呈編碼器 10 預期到之表示，則編碼器 10 不必執行頻譜分解。

[0038] 圖 5 展示剛剛提及之頻譜域中的音訊信號。如可看出，音訊信號表示為多個子頻帶信號。每一子頻帶信號 30_I 至 30_K 由藉由小框 32 指示之子頻帶值的時間序列組成。如可看出，子頻帶信號 30_I 至 30_K 之子頻帶值 32 在時間上經彼此同步，使得對於連續濾波器組時槽 34 中的每一者，每一子頻帶 30_I 至 30_K 包含準確之一個子頻帶值 32。如由頻率軸 36 所說明，子頻帶信號 30_I 至 30_K 係與不同頻率區相關聯，且如由時間軸 38 所說明，濾波器組時槽 34 在時間上經連續地配置。

[0039] 如上文所概述，圖 4 之旁側資訊提取器 17 自輸入音訊信號 s_I 至 s_N 計算 SAOC 參數。根據當前實施之 SAOC 標準，編碼器 10 以時間/頻率解析度執行此計算，該解析度相對於如由濾波器組時槽 34 及子頻帶分解判定之原始時間/頻率解析度可被減低某量，其中此某量在旁側資訊 20

內信號傳輸至解碼器側。連續濾波器組時槽 34 之數個群組可形成 SAOC 訊框 41。又，SAOC 訊框 41 內參數頻帶的數目在旁側資訊 20 內予以輸送。因此，時間/頻率域經劃分成藉由虛線 42 例證於圖 5 中的時間/頻率平鋪塊。在圖 5 中，參數頻帶在各種所描繪之 SAOC 訊框 41 中以同一方式分佈，使得獲得時間/頻率平鋪塊之規則配置。然而，一般而言，視各別 SAOC 訊框 41 中對頻譜解析度的不同需要，參數頻帶可在一個 SAOC 訊框 41 與下一 SAOC 訊框 41 之間發生變化。此外，SAOC 訊框 41 之長度又可發生變化。因而，時間/頻率平鋪塊之配置可係不規則的。然而，特定 SAOC 訊框 41 內之時間/頻率平鋪塊通常具有相同持續時間且在時間方向上對準，亦即，該 SAOC 訊框 41 中之所有 t/f 平鋪塊於給定 SAOC 訊框 41 之起點處開始，且於該 SAOC 訊框 41 之終點處結束。

[0040]描繪於圖 4 中之旁側資訊提取器 17 根據以下方程式計算 SAOC 參數。詳言之，旁側資訊提取器 17 針對每一物件 i 計算物件層級差如下

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

[0041]其中 l 及索引 n 及 k 分別遍歷所有時間索引 34，且屬於某時間/頻率平鋪塊 42 的所有頻率索引 30 由 SAOC 訊框(或處理時槽)之索引 l 及參數頻帶的索引 m 提及，且 $x_i^{n,k*}$ 係 $x_i^{n,k}$ 的複共軛。藉此，音訊信號或物件 i 之所有子頻帶值

x_i 的能量經總計並正規化為彼平鋪塊之在所有物件或音訊信號中的最高能量值。

[0042] 另外，SAOC 旁側資訊提取器 17 能夠計算數對不同輸入物件 s_I 至 s_N 之對應時間/頻率平鋪塊的類似性量測。儘管 SAOC 旁側資訊提取器 17 可計算所有成對輸入物件 s_I 至 s_N 之間的類似性量測，但 SAOC 旁側資訊提取器 17 亦可抑制類似性量測的信號傳輸或將類似性量測的計算約束至形成常見立體通道之左側或右側通道的音訊物件 s_I 至 s_N 。在任何狀況下，類似性量測被稱作物件間交叉相關參數 $IOC_{i,j}^{l,m}$ 。計算係如下

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

[0043] 其中索引 n 及 k 再次遍歷屬於某時間/頻率平鋪塊 42 之所有子頻帶值， i 及 j 標明某對音訊物件 s_I 至 s_N ，且 $\text{Re}\{\}$ 標明保留複數值自變數之僅實部(亦即，捨棄虛部)的運算。

[0044] 圖 4 之下降混合器 16 藉由使用應用至每一物件 s_I 至 s_N 的增益因數來下降混合物件 s_I 至 s_N 。亦即，增益因數 d_i 應用至物件 i ，且接著所有因此經加權之物件 s_I 至 s_N 經總計以獲得單下降混合信號，其例證於圖 4 中(在 $P=1$ 情況下)。在描繪於圖 4(在 $P=2$ 情況下)中之雙通道下降混合信號的另一實例狀況下，增益因數 $d_{l,i}$ 應用至物件 i ，且接著所有此等增益放大之物件經總計以便獲得左側下降混合

通道 $L0$ ，且增益因數 $d_{2,i}$ 應用至物件 i ，且接著因此增益放大之物件經總計以便獲得右側下降混合通道 $R0$ 。在多通道下降混合 ($P>2$) 之狀況下將應用類似於以上處理的處理。

[0045] 此下降混合規定借助於下降混合增益 DMG_i 信號傳輸至解碼器側，且在立體下降混合信號狀況下借助於下降混合通道層級差 $DCLD_i$ 而信號傳輸至解碼器側。

[0046] 下降混合增益根據以下方程式計算：

$$DMG_i = 20 \log_{10}(d_i + \varepsilon) \quad , \text{ (單下降混合) } ,$$

$$DMG_i = 10 \log_{10}(d_{1,i}^2 + d_{2,i}^2 + \varepsilon) \quad , \text{ (立體下降混合) } ,$$

其中 ε 係諸如 10^{-9} 的小數字。

對於 $DCLD$ ，應用以下方程式：

$$DCLD_i = 20 \log_{10} \left(\frac{d_{1,i}}{d_{2,i} + \varepsilon} \right) .$$

[0047] 在正常模式中，下降混合器 16 根據以下方程式產生下降混合信號：

$$(L0) = (d_i) \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix} .$$

[0048] 對於單下降混合，或

$$\begin{pmatrix} L0 \\ R0 \end{pmatrix} = \begin{pmatrix} d_{1,i} \\ d_{2,i} \end{pmatrix} \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix}$$

分別對於立體下降混合。

[0049] 因此，在上述方程式中，參數 OLD 及 IOC 係音訊信號之函數，且參數 DMG 及 $DCLD$ 係下降混合係數 d 的函數。順便地，請注意， d 在時間及頻率上可係可變的。

[0050]因此，在正常模式中，下降混合器 16 在無偏好情況下(亦即，在等同地處置所有物件 s_1 至 s_N 情況下)混合所有物件 s_1 至 s_N 。

[0051]在解碼器側，向上混合器執行下降混合程序的反轉程序及一個計算步驟中(亦即，在雙通道下降混合狀況下)藉由矩陣 $\mathbf{R}$ (在文獻資料中有時亦稱作 $\mathbf{A}$)表示之「呈現資訊」26 的實施

$$\begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_M \end{pmatrix} = \mathbf{R} \mathbf{E} \mathbf{D}^* (\mathbf{D} \mathbf{E} \mathbf{D}^*)^{-1} \begin{pmatrix} L0 \\ R0 \end{pmatrix},$$

[0052]其中矩陣 $\mathbf{E}$ 係參數 OLD 及 IOC 之函數，且矩陣 $\mathbf{D}$ 含有如下下降混合係數

$$\mathbf{D} = \begin{pmatrix} d_{1,1} & \cdots & d_{1,N} \\ \vdots & \ddots & \vdots \\ d_{P,1} & \cdots & d_{P,N} \end{pmatrix},$$

[0053]且其中 $\mathbf{D}^*$ 標明 $\mathbf{D}$ 之複數轉置。矩陣 $\mathbf{E}$ 係音訊物件 s_1 至 s_N 之所估計協方差矩陣。在當前 SAOC 實施中，所估計協方差矩陣 $\mathbf{E}$ 的計算通常以 SAOC 參數之頻譜/時間解析度(亦即，對於每一 (l,m))執行，使得所估計協方差矩陣可被撰寫為 $\mathbf{E}^{l,m}$ 。所估計協方差矩陣 $\mathbf{E}^{l,m}$ 具有大小 $N \times N$ ，其中其係數界定如下

$$e_{i,j}^{l,m} = \sqrt{OLD_i^{l,m} OLD_j^{l,m}} IOC_{i,j}^{l,m}。$$

[0054]因此，由於 $OLD_i^{l,m} = OLD_j^{l,m}$ 且 $IOC_{i,j}^{l,m} = 1$ (對於 $i=j$)，因此具有下式之矩陣 $\mathbf{E}^{l,m}$ 沿其對角線具有物件層級差，亦即 $e_{i,j}^{l,m} = OLD_i^{l,m}$ (對於 $i=j$)：

$$E^{l,m} = \begin{pmatrix} e_{1,1}^{l,m} & \cdots & e_{1,N}^{l,m} \\ \vdots & \ddots & \vdots \\ e_{N,1}^{l,m} & \cdots & e_{N,N}^{l,m} \end{pmatrix}。$$

在其對角線外部，所估計協方差矩陣 E 具有藉由物件間交叉相關量測 $IOC_{i,j}^{l,m}$ 進行加權的分別表示物件 i 及 j 之物件層級差之幾何平均值的矩陣係數。

[0055] 圖 6 顯示關於作為 SAOC 編碼器 10 之部分的旁側資訊估計器(SIE)之實施的實施之一個可能原理。SAOC 編碼器 10 包含混合器 16 及旁側資訊估計器(SIE)17。SIE 概念上由兩個模組組成：計算每一信號之基於短時間的 t/f 表示(例如，STFT 或 QMF)的一個模組 45。所計算之短時間 t/f 表示饋入至第二模組 46(t/f 選擇性旁側資訊估計模組 (t/f-SIE))中。t/f-SIE 模組 46 計算每一 t/f 平鋪塊的旁側資訊。在當前 SAOC 實施中，時間/頻率變換對於所有音訊物件 s_1 至 s_N 係固定且等同的。此外，SAOC 參數在 SAOC 訊框上進行判定，該等 SAOC 參數對於所有音訊物件係相同的且對於所有音訊物件 s_1 至 s_N 具有相同時間/頻率解析度，因此不管在一些狀況下對精細時間解析度或在其他狀況下對精細頻譜解析度的物件特定需要。

[0056] 在以下內容中，描述本發明之實施例。

[0057] 圖 1a 說明根據實施例之用於產生未經混合音訊信號的解碼器，該未經混合音訊信號包含多個未經混合音訊信號。

[0058] 解碼器包含未混合資訊判定器 112，其用於藉由接收關於至少一音訊物件信號的第一參數旁側資訊及關於

至少一音訊物件信號之第二參數旁側資訊來判定未混合資訊，其中第二參數旁側資訊的頻率解析度高於第一參數旁側資訊的頻率解析度。

[0059]此外，解碼器包含未混合模組 113，其用於將未混合資訊應用於指示至少一音訊物件信號之下降混合的下降混合信號，以獲得包含多個未經混合音訊通道的未經混合音訊信號。

[0060]未混合資訊判定器 112 經組配以藉由修改第一參數資訊及第二參數資訊以獲得經修改參數資訊來判定未混合資訊，使得經修改參數資訊具有高於第一頻率解析度的頻率解析度。

[0061]圖 1b 說明根據另一實施例之用於產生未經混合音訊信號的解碼器，該未經混合音訊信號包含多個未經混合音訊信號。圖 1b 之解碼器此外包含第一變換單元 111，其用於變換在時域中表示之下降混合輸入以獲得在時間頻率域中表示的下降混合信號。此外，圖 1b 之解碼器包含第二變換單元 114，其用於將未經混合音訊信號自時間頻率域變換至時域。

[0062]圖 2a 說明根據實施例的用於編碼一或多個輸入音訊物件信號的編碼器。

[0063]編碼器包含下降混合單元 91，其用於下降混合一或多個輸入音訊物件信號以獲得一或多個下降混合信號。

[0064]此外，編碼器包含參數旁側資訊產生器 93，其

用於產生關於至少一音訊物件信號之第一參數旁側資訊及關於至少一音訊物件信號的第二參數旁側資訊，使得第二參數旁側資訊之頻率解析度高於第一參數旁側資訊的頻率解析度。

[0065]圖 2b 說明根據另一實施例的用於編碼一或多個輸入音訊物件信號的編碼器。圖 2b 之編碼器進一步包含變換單元 92，其用於將一或多個輸入音訊物件信號自時域變換至時間頻率域以獲得一或多個經變換音訊物件信號。在圖 2b 之實施例中，參數旁側資訊產生器 93 經組配以基於一或多個經變換音訊物件信號產生第一參數旁側資訊及第二參數旁側資訊。

[0066]圖 2c 說明根據一實施例的經編碼音訊信號。經編碼音訊信號包含一指示一或多個輸入音訊物件信號之下降混合的下降混合部分 51，以及包含關於至少一音訊物件信號之第一參數旁側資訊及關於至少一音訊物件信號之第二參數旁側資訊的參數旁側資訊部分 52。第二參數旁側資訊的頻率解析度高於第一參數旁側資訊之頻率解析度。

[0067]圖 3 說明根據實施例的系統。系統包含如上文所描述之編碼器 61 及如上文所描述的解碼器 62。

[0068]編碼器 61 經組配以藉由以下操作來編碼一或多個輸入音訊物件信號：獲得指示一或多個輸入音訊物件信號之下降混合的一或多個下降混合信號，獲得關於至少一音訊物件信號之第一參數旁側資訊及獲得關於至少一音訊物件信號的第二參數旁側資訊，其中第二參數旁側資訊的

頻率解析度高於第一參數旁側資訊之頻率解析度。

[0069]解碼器 62 經組配以基於一或多個下降混合信號且基於第一參數旁側資訊及第二參數旁側資訊來產生未經混合音訊信號。

[0070]在以下內容中，描述使用反向相容頻率解析度改良的增強型 SAOC。

[0071]圖 7 說明根據實施例的反向相容表示。待表示之信號性質(例如，功率頻譜包絡 71)在頻率上發生變化。頻率軸經分割成參數頻帶，且單一集合之信號描述符針對每一子頻帶進行指派。使用信號描述符而非遞送每一頻率區間的描述分離地允許所要求之旁側資訊之量的節省而無感知品質的顯著損失。在標準 SAOC 中，每一頻帶之單一描述符係逐區間描述符的平均值 72、73、74。如可理解，此情形可引入資訊損失，資訊損失的量值取決於信號性質。在圖 7 中，頻帶 $k-1$ 及 k 具有相當大之誤差，而在頻帶 $k+1$ 中，誤差要小得多。

[0072]圖 8 說明根據實施例的真實參數值與低解析度平均值之間的差異曲線 81(例如，在標準 SAOC 參數化中失去之精細結構資訊)。描述一種用於以有效方式參數化並傳輸平均值 72、73、74(例如，標準 SAOC 描述符)與真實精細解析度值之間的差異曲線 81 從而允許近似解碼器中的精細解析度結構之方法。

[0073]請注意，將增強資訊添加至混合物中之單一物件不僅改良彼特定物件的所得品質，而且改良共用近似空間

位置且具有某頻譜重疊之所有物件的品質。

[0074]在以下內容中，描述藉由增強型編碼器(詳言之，增強型 SAOC 編碼器)進行之反向相容增強型 SAOC 編碼，該增強型 SAOC 編碼器產生含有反向相容旁側資訊部分及額外增強的位元串流。所添加資訊可以一方式插入至標準 SAOC 位元串流中，使得陳舊之標準相容解碼器在增強型解碼器使用所添加資料同時簡單地忽略所添加資料。現有標準 SAOC 解碼器可解碼參數旁側資訊(PSI)之反向相容部分，且產生物件之重建，同時增強型 SAOC 解碼器使用的添加資訊在多數狀況下改良重建的感知品質。另外，若增強型 SAOC 解碼器正在有限資源上執行，則增強可被忽略，且仍獲得基本品質重建。請注意，使用僅標準 SAOC 相容 PSI 自標準 SAOC 及增強型 SAOC 解碼器的重建不同，但斷定為感知上極其類似的(差異係關於如增強型 SAOC 解碼器情況下解碼標準 SAOC 位元串流上的類似本質)。

[0075]圖 9 描繪根據實施例之增強型編碼器的高層級說明，該增強型編碼器提供具有增強的反向相容位元串流。

[0076]編碼器包含下降混合單元 91，其用於下降混合多個音訊物件信號以獲得一或多個下降混合信號。舉例而言，音訊物件信號(例如，個別(音訊)物件)由下降混合單元 91 使用以產生下降混合信號。此情形可在時域、頻域中發生，或甚至可使用外部提供之下降混合。

[0077]在 PSI 路徑中，(音訊)物件信號由變換單元 92

自時域變換至頻域、時間頻率域或頻譜域(舉例而言，藉由包含一或多個 t/f 變換單元 921、922 的變換單元 92)。

[0078]此外，編碼器包含用於產生參數旁側資訊的參數旁側資訊產生器 93。在圖 9 之實施例中，參數旁側資訊產生器 93 可(例如)包含 PSI 提取單元 94 及 PSI 分裂器 95。根據此實施例，在頻域中，PSI 藉由 PSI 提取單元 94 來提取。PSI 分裂器 95 接著進行以將 PSI 分裂成兩個部分：可藉由任何標準相容 SAOC 解碼器解碼的標準頻率解析度部分，及增強型頻率解析度部分。後者可「隱藏」於位元串流元素中，使得此等位元串流元素將由標準解碼器忽略，但由增強型解碼器利用。

[0079]圖 10 說明根據實施上文描述之編碼器之參數路徑的特定實施例之編碼器的方塊圖。粗黑功能區塊(102、105、106、107、108、109)指示本發明處理的主要組件。詳言之，圖 10 說明產生反向相容位元串流與針對能力更強解碼器之增強的二級編碼的方塊圖。編碼器經組配以產生可藉由兩個解碼器型式解碼的 PSI。圖 9 之變換單元 92 藉由圖 10 中的瞬變偵測單元 101、藉由產生窗序列單元 102 且藉由 t/f 分析單元 103 來實施。圖 10 中之其他單元 104、105、106、107、108、109 實施參數旁側資訊產生器 93(例如，單元 104、105、106、107、108、109 可實施 PSI 提取單元 94 及 PSI 分裂器 95 之組合的功能性)。

[0080]首先，將信號再分成分析訊框，該等分析訊框接著經變換至頻域。多個分析訊框接著分組成固定長度參數

訊框，例如，在標準 SAOC 中，16 及 32 分析訊框的長度係常見的。假定，信號性質在參數訊框期間保持準穩定，且可因此藉由僅一個參數集來特徵化。若信號特性在參數訊框內改變，則遭受模型化誤差，且將較長參數訊框再分成準穩定之假定再次經實現的部分將係有益的。爲了此目的，需要瞬變偵測。

[0081]在一實施例中，視包含一或多個輸入音訊物件信號中之至少一者之信號值的信號變換區塊的窗長度而定，變換單元 92 經組配以將一或多個輸入音訊物件信號自時域變換至時間頻率域。變換單元 92 包含瞬變偵測單元 101，其用於判定指示瞬變是否存在於至少一音訊物件信號中之一或多者中的瞬變偵測結果，其中瞬變指示至少一音訊物件信號中之一或多者中的信號改變。此外，變換單元 92 進一步包含窗序列單元 102，其用於視瞬變偵測結果來判定窗長度。

[0082]舉例而言，瞬變可藉由瞬變偵測單元 101 自所有輸入物件分離地偵測，且在瞬變事件存在於物件中之僅一者中時，該位置宣告爲全域瞬變位置。瞬變位置之資訊用於建構適當開窗序列。建構可係基於(例如)以下邏輯：

- 設定預設窗長度，亦即，預設信號變換區塊的長度(例如，2048 個樣本)。
- 設定具有 50%重疊之對應於 4 個預設窗的參數訊框長度(例如，4096 個樣本)。參數訊框將多個窗分組在一起且單一集合之信號描述符用於整個區塊而非分離地具有針

對每一窗的描述符。此情形允許減小 PSI 的量。

- 若尚未偵測到瞬變，則使用預設窗及全參數訊框長度。
- 若偵測到瞬變，則調適開窗以在瞬變之位置處提供更好時間解析度。

[0083]產生窗序列單元 102 建構開窗序列。同時，產生窗序列單元 102 亦自一或多個分析窗產生參數子訊框。每一子集作為整體進行分析，且針對每一子區塊傳輸僅一個集合之 PSI 參數。為了提供標準 SAOC 相容 PSI，所界定參數區塊長度用作主參數區塊長度，且彼區塊內之可能之所定位瞬變界定參數子集。

[0084]所建構窗序列經輸出以供 t/f 分析單元 103 進行之輸入音訊信號的時間頻率分析，且在 PSI 之增強型 SAOC 增強部分中進行傳輸。

[0085]PSI 由以下各者組成：物件層級差(OLD)集合，物件間相關(IOC)，及用以自編碼器中之個別物件產生下降混合信號的下降混合矩陣 D 的資訊。每一參數集係與參數邊界相關聯，該邊界界定參數關聯至的時間區。

[0086]每一分析窗之頻譜資料由 PSI 估計單元 104 用於估計標準 SAOC 部分的 PSI。此藉由將頻譜區間分組成標準 SAOC 之參數頻帶且估計頻帶中 IOC、OLD 及絕對物件能量(NRG)來進行。鬆散地遵循標準 SAOC 之記數法，參數化平鋪塊中兩個物件頻譜 $S_i(f,n)$ 及 $S_j(f,n)$ 的正規化乘積界定為：

$$nrg_{i,j}(b) = \frac{\sum_{n=0}^{N-1} \sum_{f=0}^{F_n-1} \mathbf{K}(b, f, n) \mathbf{S}_i(f, n) \mathbf{S}_j^*(f, n)}{\sum_{n=0}^{N-1} \sum_{f=0}^{F_n-1} \mathbf{K}(b, f, n)},$$

其中矩陣 $\mathbf{K}(b, f, n)$: $\mathbb{R}^{B \times F_n \times N}$ 藉由下式界定自訊框 n 中 F_n t/f 表示區間至 B 參數頻帶中的映射

$$\mathbf{K}(b, f, n) = \begin{cases} 1, & \text{若 } f \in b \\ 0, & \text{在其他情況下} \end{cases}。$$

[0087] 頻譜解析度在單一參數區塊內之訊框之間可發生變化，因此映射矩陣將資料轉換成共同解析度基準。此參數化平鋪塊中之最大物件能量經定為最大物件能量 $NRG(b) = \max_i(nrg_{i,i}(b))$ 。具有此值，接著將 OLD 界定為正規化物件能量

$$OLD_i(b) = \frac{nrg_{i,i}(b)}{NRG(b)}。$$

且最終 IOC 可獲得自交叉功率如下

$$IOC_{i,j}(b) = \text{Re} \left\{ \frac{nrg_{i,j}(b)}{\sqrt{nrg_{i,i}(b) nrg_{j,j}(b)}} \right\}$$

[0088] 此情形推斷出位元串流之標準 SAOC 相容部分的估計。

[0089] 粗糙功率頻譜重建單元 105 經組配以將 OLD 及 NRG 用於在參數分析區塊中重建頻譜包絡的粗略估計。包絡以用於該區塊中之最高頻率解析度來建構。

[0090] 每一分析窗之原始頻譜由功率頻譜估計單元 106 用於計算彼窗中的功率頻譜。

[0091]所獲得之功率頻譜藉由頻率解析度調適單元 107 經轉換成常見高頻率解析度表示。此情形可(例如)藉由對功率頻譜值進行內插來進行。接著，平均功率頻譜概況藉由對參數區塊內之頻譜進行平均來計算。此情形粗略地對應於省略參數頻帶聚合的 OLD 估計。所獲得之頻譜概況被視為精細解析度 OLD。

[0092]編碼器進一步包含德耳塔估計單元 108，其用於藉由使至少一音訊物件信號中之一者的多個 OLD 中之每一者除以至少一音訊物件信號的該一者之功率頻譜重建的值來估計多個校正因數以獲得第二參數旁側資訊，其中該多個 OLD 具有相較於該功率頻譜重建較高的頻率解析度。

[0093]在實施例中，德耳塔估計單元 108 經組配以視至少一音訊物件信號而基於多個參數值來估計多個校正因數以獲得第二參數旁側資訊。例如，德耳塔估計單元 108 可經組配以(例如)藉由用粗略功率頻譜重建除以精細解析度 OLD 來估計校正因數「德耳塔」。結果，此情形針對每一頻率區間提供(例如，倍增)校正因數，該校正因數在給定粗略頻譜情況下可用於近似精細解析度 OLD。

[0094]最終，德耳塔模型化單元 109 經組配而以有效方式模型化所估計校正因數以供傳輸。使用線性預測係數(LPC)進行模型化的一個可能性稍後在下文進行描述。

[0095]有效地，增強型 SAOC 修改由將開窗序列資訊及用於傳輸「德耳塔」的參數添加至位元串流組成。

[0096]在以下內容中，描述增強型解碼器。

[0097]圖 11 描繪根據實施例之增強型解碼器的高層級方塊圖，該解碼器能夠解碼標準位元串流及增強型位元串流兩者。詳言之，圖 11 說明能夠解碼標準位元串流以及包括頻率解析度增強之位元串流兩者的增強型解碼器之操作方塊圖。

[0098]輸入下降混合信號藉由 t/f 變換單元 111 變換至頻域。

[0099]所估計之未混合矩陣藉由未混合單元 110 應用於經變換下降混合信號以產生未混合輸出。

[0100]另外，去相關路徑經包括以允許未混合中物件的更好空間控制。去相關單元 119 對經變換下降混合信號進行去相關，且將去相關之結果饋入至未混合單元 110 中。未混合單元 110 將去相關結果用於產生未混合輸出。

[0101]未混合輸出接著藉由 f/t 變換單元 114 經變換回至時域。

[0102]參數處理路徑可採用標準解析度 PSI 作為輸入，在該狀況下，藉由標準 PSI 解碼單元 115 產生的經解碼 PSI 藉由頻率解析度轉換單元 116 調適至用於 t/f 變換中的頻率解析度。

[0103]替代性輸入組合 PSI 之標準頻率解析度部分與增強型頻率解析度部分，且計算包括增強型頻率解析度資訊。更詳細地，增強型 PSI 解碼單元 117 產生展現增強型頻率解析度的經解碼 PSI。

[0104]未混合矩陣產生器 118 基於接收自頻率解析度

轉換單元 116 或接收自增強型 PSI 解碼單元 117 的經解碼 PSI 產生未混合矩陣。未混合矩陣產生器 118 亦可基於呈現資訊(例如，基於呈現矩陣)產生未混合矩陣。未混合單元 110 經組配以藉由將由未混合矩陣產生器 118 產生的此未混合矩陣應用於經變換下降混合信號來產生未混合輸出。

[0105]圖 12 說明方塊圖，該方塊圖說明圖 11 之增強型 PSI 解碼單元 117 之實施例。

[0106]第一參數資訊包含多個第一參數值，其中第二參數資訊包含多個第二參數值。未混合資訊判定器 112 包含一頻率解析度轉換子單元 122 及一組合器 124。頻率解析度轉換單元 112 經組配以(例如)藉由複製第一參數值而產生額外參數值，其中第一參數值及額外參數值一起形成多個第一經處理參數值。組合器 124 經組配以組合第一經處理參數值與第二參數值以獲得多個經修改參數值作為經修改參數資訊。

[0107]根據實施例，標準頻率解析度部分藉由解碼子單元 121 解碼，且藉由頻率解析度轉換子單元 122 轉換成由增強部分使用的頻率解析度。藉由增強型 PSI 解碼子單元 123 產生之經解碼增強部分藉由組合器 124 與經轉換標準解析度部分組合。

[0108]在以下內容中，更詳細地描述可能實施情況下的兩個解碼模式。

[0109]首先，描述藉由增強型解碼器解碼標準 SAOC 位元串流：

[0110]增強型 SAOC 解碼器經設計，使得其能夠以良好品質解碼來自標準 SAOC 編碼器的位元串流。解碼限於僅參數重建，且忽略可能之殘餘串流。

[0111]圖 13 描繪根據一實施例之說明解碼程序的藉由增強型 SAOC 解碼器解碼標準 SAOC 位元串流的方塊圖。粗黑功能區塊(131、132、133、135)指示本發明處理的主要部分。

[0112]未混合矩陣計算器 131、時間內插器 132 及窗頻率解析度調適單元 133 實施圖 1 之標準 PSI 解碼單元 115、頻率解析度轉換單元 116 及未混合矩陣產生器 118 的功能性。窗序列產生器 134 及 t/f 分析模組 135 實施圖 11 的 t/f 變換單元 111。

[0113]正常地，基礎時間/頻率表示的頻率區間經分組成參數頻帶。頻帶之間隔類似於人類聽覺系統中關鍵頻帶的間隔。此外，多個 t/f 表示訊框可經分組成參數訊框。此等操作中之兩者以模型化不準確性為代價提供所需要旁側資訊之量的減小。

[0114]如在 SAOC 標準中所描述，OLD 及 IOC 用以計算未混合矩陣 $\mathbf{G} = \mathbf{E}\mathbf{D}^*\mathbf{J}$ ，其中 E 元素界定為近似物件交叉相關矩陣的 $\mathbf{E}(i,j) = \text{IOC}_{i,j} \sqrt{\text{OLD}_i \text{OLD}_j}$ ， i 及 j 係物件索引， $\mathbf{J} \approx (\mathbf{D}\mathbf{E}\mathbf{D}^*)^{-1}$ 。未混合矩陣計算器 131 可進行以計算未混合矩陣。

[0115]按照標準 SAOC，未混合矩陣接著藉由時間內插器 132 在參數訊框上自先前訊框的未混合矩陣起進行線性

內插直至達到所估計值所在之參數邊界。此導致每一時間/頻率分析窗及參數頻帶的未混合矩陣。

[0116] 未混合矩陣之參數頻帶頻率解析度藉由窗頻率解析度調適單元 133 擴展至彼分析窗中之時間/頻率表示的解析度。當時間訊框中參數頻帶 b 之經內插未混合矩陣界定為 $\mathbf{G}(b)$ 時，同一未混合係數用於彼參數頻帶內部的所有頻率區間。

[0117] 窗序列產生器 134 經組配以使用來自 PSI 之參數集範圍資訊以判定用於分析輸入下降混合音訊信號的恰當開窗序列。主要要求為，當在 PSI 中存自參數集邊界時，連續分析窗之間的交叉點應與參數集邊界匹配。開窗亦判定每一窗內資料的頻率解析度(如較早所描述，用於未混合資料擴展中)。

[0118] 經開窗資料接著使用(例如)以下各者之恰當時間頻率變換藉由 t/f 分析模組 135 變換成頻域表示：離散傅立葉變換(DFT)、複合經修改離散餘弦變換(CMDCT)或奇數堆疊離散傅立葉變換(ODFT)。

[0119] 最終，未混合單元 136 將每訊框每頻率區間之未混合矩陣應用於下降混合信號 $\mathbf{x}$ 之頻譜表示以獲得參數呈現 $\mathbf{y}$ 。輸出通道 j 係下降混合通道的線性組合 $\mathbf{y}_j = \sum_i \mathbf{G}_{j,i} \mathbf{x}_i$ 。

[0120] 對於大多數用途，可藉由此程序獲得之品質與藉由標準 SAOC 解碼器獲得之結果係感知上難區分的。

[0121] 請注意，以上文字描述個別物件之重建，但在標準 SAOC 中，呈現包括於未混合矩陣中，亦即，呈現包括

於參數內插中。作為線性運算，運算之次序並無重要關係，但差係值得一提的。

[0122]在以下內容中，描述藉由增強型解碼器解碼增強型 SAOC 位元串流。

[0123]增強型 SAOC 解碼器之主要功能性已在解碼標準 SAOC 位元串流中較早進行了描述。此章節將詳述 PSI 中所引入之增強型 SAOC 增強可用於獲得更好感知品質的方式。

[0124]圖 14 描繪根據實施例之說明解碼頻率解析度增強的主要功能區塊。粗黑功能區塊(141、142、143)指示本發明處理的主要部分。頻帶上值擴展單元 141、德爾塔函數恢復單元 142、德爾塔應用單元 143、未混合矩陣計算器 131、時間內插器 132 及窗頻率解析度調適單元 133 實施圖 11 之增強型 PSI 解碼單元 117 及未混合矩陣產生器 118 的功能性。

[0125]圖 14 之解碼器包含未混合資訊判定器 112。其中，未混合資訊判定器 112 包含德爾塔函數恢復單元 142 及德爾塔應用單元 143。第一參數資訊視至少一音訊物件信號而包含多個參數值(例如，物件層級差值)。第二參數資訊包含校正因數參數化。德爾塔函數恢復單元 142 經組配以反轉該校正因數參數化以獲得德爾塔函數。該德爾塔應用單元 143 經組配以對參數值(例如，對物件層級差值)應用該德爾塔函數以判定未混合資訊。在實施例中，校正因數參數化包含多個線性預測係數，且德爾塔函數恢復單元 142

經組配以藉由視多個線性預測係數而產生多個校正因數來反轉校正因數參數化，且經組配以基於多個校正因數產生德爾塔函數。

[0126]舉例而言，首先，頻帶上值擴展單元 141 針對每一參數頻帶調適 OLD 及 IOC 值至用於增強中的頻率解析度(例如，至 1024 個區間)。此情形藉由複製對應於參數頻帶上之頻率區間上的值來進行。此導致新 OLD $OLD_i^{enh}(f) = \mathbf{K}(f, b)OLD_i(b)$ 及 IOC $IOC_{i,j}^{enh}(f) = \mathbf{K}(f, b)IOC_{i,j}(b)$ 。 $\mathbf{K}(f, b)$ 係界定頻率區間 f 至參數頻帶 b 之指派的核心矩陣。

[0127]與以上情形並行，德爾塔函數恢復單元 142 反轉校正因數參數化以獲得具有與經擴展 OLD 及 IOC 相同之大小的德爾塔函數 $\mathbf{C}_i^{rec}(f)$ 。

[0128]接著，德爾塔應用單元 143 對經擴展 OLD 值應用德爾塔，且所獲得之精細解析度 OLD 值藉由 $OLD_i^{fine}(f) = \mathbf{C}_i^{rec}(f)OLD_i^{enh}(f)$ 來獲得。

[0129]在特定實施例中，未混合矩陣之計算可(例如)藉由未混合矩陣計算器 131 如同解碼標準 SAOC 位元串流一般來進行： $\mathbf{G}(f) = \mathbf{E}(f)\mathbf{D}^*(f)\mathbf{J}(f)$ ，其中 $\mathbf{E}_{i,j}(f) = IOC_{i,j}^{enh}(f)\sqrt{OLD_i^{fine}(f)OLD_j^{fine}(f)}$ 且 $\mathbf{J}(f) \approx (\mathbf{D}(f)\mathbf{E}(f)\mathbf{D}^*(f))^{-1}$ 。若想要，則呈現矩陣可被倍增至未混合矩陣 $\mathbf{G}(f)$ 中。時間內插器 132 進行之時間內插遵照標準 SAOC。

[0130]由於每一窗中之頻率解析度可不同於(低於)標稱高頻率解析度，因此窗頻率解析度調適單元 133 需要調適未混合矩陣以與來自音訊之頻譜資料的解析度匹配以允

許應用該頻率解析度。此情形可(例如)在頻率軸上對係數重新取樣達正確解析度來進行。或若解析度係整數倍，則自高解析度資料對對應於較低解析度中一個頻率區間的索引簡單地進行平均 $\mathbf{G}^{low}(b) = 1/\|b\| \sum_{f \in b} \mathbf{G}(f)$ 。

[0131]來自位元串流之開窗序列資訊可用以獲得與用於編碼器中之分析完全互補的時間頻率分析，或開窗序列可如同在標準 SAOC 位元串流解碼中進行一般基於參數邊界來建構。為此，可使用窗序列產生器 134。

[0132]下降混合音訊之時間頻率分析接著藉由 t/f 分析模組 135 使用給定窗進行。

[0133]最終，經時間內插且頻譜(可能)調適之未混合矩陣藉由未混合單元 136 應用於輸入音訊的時間頻率表示上，且輸出通道 j 可作為輸入通道的線性組合而獲得 $\mathbf{Y}_j(f) = \sum_i \mathbf{G}_{j,i}^{low}(f) \mathbf{X}_i(f)$ 。

[0134]在以下內容中，描述實施例的特定態樣。

[0135]在實施例中，圖 10 之德爾塔模型化單元 109 經組配以藉由進行線性預測而自多個校正因數(德爾塔)判定線性預測係數。

[0136]現在，描述根據此實施例的校正因數(德爾塔)之估計程序及使用線性預測係數(LPC)的可能模型化替代例。

[0137]首先，描述根據實施例之德爾塔估計。

[0138]對估計之輸入由在參數區塊上及來自基於 OLD 及 NRG 參數之功率頻譜概況之粗糙重建的所估計精細解析度功率頻譜概況組成。精細功率頻譜概況以如下方式進

行計算。 $S_i(f, n)$ 係第*i*物件之複合頻譜，其中*f*係頻率區間索引，且 $0 \leq n \leq N-1$ 係長度為*N*之模型化區塊中的時間窗索引。精細解析度功率頻譜則係

$$P_i(f) = \frac{1}{N} \sum_{n=0}^{N-1} S_i(f, n) S_i^*(f, n)$$

粗糙重建藉由下式自(去量化)OLD及NRG計算

$$Z_i(f) = K(f, b) OLD_i(b) NRG_i(b)$$

其中 $K(f, b)$ 係界定頻率區間*f*至參數頻帶*b*之指派的核
心矩陣。

[0139]具有不同頻譜性質的兩個信號在此章節中將用作實例：第一信號係具有實際上平坦頻譜的(粉紅)雜訊(忽略頻譜傾斜)，且第二信號係來自樂器鐵琴(glockenspiel)之具有高音調(亦即，尖峰)頻譜的音調。

[0140]圖 15 說明音調信號及雜訊信號的功率頻譜。其高解析度功率頻譜(「原始」)及基於 OLD 及 NRG 的對應粗略重建(「重建」)。詳言之，圖 15 說明兩個信號的精細功率頻譜及粗糙功率頻譜。更特定而言，展示原始音調信號 151 及原始雜訊信號 152 的功率頻譜以及音調信號 153 及雜訊信號 154 的經重建功率頻譜。請注意，在以下諸圖中，對於信號 153 及 154，寧可草繪縮放因數(經重建之功率頻譜參數)而非充分重建的信號。

[0141]可快速地注意到，精細值與粗糙值之間的平均差在雜訊信號之狀況下係相當小的，而音調信號中的差係極大的。此等差引起所有物件之參數重建的感知降級。

[0142]校正因數藉由使精細解析度曲線除以粗糙重建曲線來獲得：

$$\mathbf{C}_i(f) = \mathbf{P}_i(f) / \mathbf{Z}_i(f)。$$

此情形允許恢復可應用於粗略重建上以獲得精細解析度曲線的倍增因數：

$$\mathbf{P}_i^{rec}(f) = \mathbf{Z}_i(f) \mathbf{C}_i(f)。$$

[0143]圖 16 說明兩個實例信號的修改，詳言之，實例信號之校正因數。詳言之，展示音調信號 151 及雜訊信號 152 的校正因數。

[0144]在以下內容中，描述德爾塔模型化。

[0145]校正曲線 $\mathbf{C}$ 在頻率軸線上指派至一或多個模型化區塊中。本質替代例係使用與用於標準 SAOC PSI 之參數頻帶定義相同的參數頻帶定義。模型化接著以如下步驟對於每一區塊分離地進行：

1. 藉由離散傅立葉逆變換(IDFT)將頻譜校正因數 $\mathbf{C}$ 變換至時域自動相關序列。

當模型化區塊之長度係奇數時，待變換之偽頻譜界定如下

$$\mathbf{R}(l) = \begin{cases} \mathbf{C}(l), & \text{其中 } 0 \leq l \leq N-1 \\ \mathbf{C}(2N-2-l), & \text{其中 } N \leq l \leq 2N-3 \end{cases}。$$

當模型化區塊為偶數時，偽頻譜界定如下

$$\mathbf{R}(l) = \begin{cases} \mathbf{C}(l), & \text{其中 } 0 \leq l \leq N-1 \\ \mathbf{C}(2N-1-l), & \text{其中 } N \leq l \leq 2N-2 \end{cases}。$$

變換結果則係 $r(t) = \text{IDFT}(\mathbf{R}(l))$ 。

2. 結果經截斷為前半部分：
$$\begin{cases} 0 \leq t \leq N-2, & \text{當 } N \text{ 為奇數時} \\ 0 \leq t \leq N-1, & \text{當 } N \text{ 為偶數時} \end{cases}$$

3. 列文遜-杜賓遞歸應用於自動相關序列 $r(t)$ 以獲取反射係數 $\mathbf{k}$ 及模型化殘餘變數 $\mathbf{e}$ 從而使模型階數增加。

4. 可選：基於模型化殘餘變數 $\mathbf{e}$ ，省略整個模型化(由於未獲得增益)或選擇恰當階數。

5. 模型參數經量化用於傳輸。

[0146] 若德爾塔應針對每一 t - f 平鋪塊(界定頻率範圍之標準參數頻帶及界定時間範圍的參數區塊)獨立地進行傳輸，則有可能作出一決策。該決策可基於(例如)以下各項進行：

- 檢查德爾塔模型化殘餘能量。若模型化殘餘能量並未超出某臨限值，則不傳輸增強資訊。

- 量測精細解析度模型化參數描述、德爾塔模型化或音訊物件信號之功率頻譜包絡的「尖銳度」/不平坦度。視所量測值而定，描述精細頻譜解析度之德爾塔模型化參數經傳輸或不傳輸，或根本視音訊物件信號之功率頻譜包絡的不平坦度進行計算)。恰當量測係(例如)頻譜波峰因數、頻譜平坦度量測，或最小值與最大值比率。

- 獲得重建的感知品質。編碼器在有增強且無增強情況下計算呈現重建，且判定每一增強的品質增益。接著，定位模型化複雜性與品質增益之間的恰當平衡之點，且傳輸所指示之增強。舉例而言，感知上經加權之失真與信號比率或增強型感知量測可用於決策。該決策可針對每一(粗

糙)參數頻帶分離地進行(亦即，局部品質最佳化)，但亦在考慮鄰近頻帶以解決由時間頻率係數之時間及頻率變數操控引起的信號失真情況下進行(亦即，全域品質最佳化)。

[0147]現在，描述德爾塔重建及應用。

[0148]校正曲線之重建遵循如下步驟：

1. 所接收反射係數 $\mathbf{k}$ (長度 $L-1$ 之向量) 在偽碼語法中經解量化且變換成具有長度 L 的 IIR 濾波係數 $\mathbf{a}$ (其中函數 $\mathbf{X} = \text{diag}(\mathbf{x})$ 輸出矩陣 X ，其中 X 之對角線元素係 x ，且 X 之所有非對角線元素係零)：

$\mathbf{A} = \text{diag}(\mathbf{k})$

for ii=1 to L

for l=1 to ii-1

$\mathbf{A}(l, ii) = \mathbf{A}(l, ii-1) + \mathbf{k}(ii) * \mathbf{A}(ii-1, ii-1)$

end

end

$\mathbf{a} = [1; \mathbf{A}(1 \text{ to } \text{end}, \text{end})]$

2. 所得濾波器 $\mathbf{a}$ 之頻率回應 $\mathbf{h}(n)$ 藉由

$\mathbf{h}(n) = 1 / \sum_{l=0}^{L-1} \mathbf{a}(l) \exp(-in2\pi / N)$ 來計算，其中 i 標明虛數單位

$i = \sqrt{-1}$ 。

3. 校正函數重建藉由 $\mathbf{C}^{rec}(n) = \mathbf{h}(n)\mathbf{h}^*(n)$ 自此獲得。

4. 回應經正規化以具有單位平均值，使得模型化區塊

之總體能量並不改變 $\mathbf{C}^{rec}(n) = \mathbf{C}^{raw}(n) / \sum_{n=0}^{N-1} \mathbf{C}^{raw}(n)$ 。

5. 校正因數應用於 OLD 上，其已擴展至精細解析度 $OLD_i^{fine}(f) = C_i^{rec}(f)K(f,b)OLD_i(b)$ 。請注意，在絕對能量中可被忽略，此係由於絕對能量在進一步計算中將被消除。

[0149] 圖 17 說明針對實例信號兩者的原始校正因數及基於降階 LPC 的近似(在模型化之後)。詳言之，展示音調信號 151、原始雜訊信號 152 的原始校正因數，以及音調信號 153 及雜訊信號 154 的經重建校正因數估計。

[0150] 圖 18 說明對在圖 15 中說明之對粗略重建應用經模型化校正因數的結果。詳言之，展示原始音調信號 151 及原始雜訊信號 152 的功率頻譜以及音調信號 153 及雜訊信號 154 的經重建功率頻譜估計。此等曲線現可在應用模型化校正因數之後替代 OLD 用於以下計算(詳言之，重建之精細解析度功率頻譜)中。此處，絕對能量資訊經包括以使得比較為更明顯，但同一原理在無絕對能量資訊情況下亦起作用。

[0151] 本發明方法及裝置使用具有高頻率解析度之濾波器組或時間頻率變換且提供額外資訊之有效參數化來減輕先前技術 SAOC 處理的前述缺陷。此外，有可能以一方式傳輸此額外資訊，使得標準 SAOC 解碼器可解碼與使用符合標準 SAOC 編碼器獲得之品質相當之可獲得品質解碼資訊的反向相容部分，且仍允許增強型解碼器將額外資訊用於更好感知品質。最重要地，額外資訊可以極其緊湊之方式表示以供有效傳輸或儲存。

[0152] 所呈現之本發明方法可應用於任何 SAOC 方

案。其可與任何當前且又將來之音訊格式組合。本發明方法藉由頻譜旁側資訊之雙層級表示而允許 SAOC 應用中增強之感知音訊品質。

[0153]同一理念亦可在藉由通道層級差(CLD)替換 OLD 概念時結合 MPEG 環繞立體聲來使用。

[0154]提供如上文所描述的音訊編碼器或音訊編碼方法或相關電腦程式。此外，提供如上文所描述的音訊編碼器或音訊解碼方法或相關電腦程式。此外，提供上文所描述之經編碼音訊信號或儲存有經編碼音訊信號的儲存媒體。

[0155]儘管一些態樣已在裝置之內容脈絡中予以描述，但清楚的是，此等態樣亦表示對應方法之描述，其中區塊或器件對應於方法步驟或方法步驟的特徵。類似地，在方法步驟之內容脈絡中描述的態樣亦表示對應裝置之對應區塊或項目或特徵的描述。

[0156]本發明分解信號可儲存於數位儲存媒體上或可在諸如無線傳輸媒體或有線傳輸媒體(諸如，網際網路)之傳輸媒體上進行傳輸。

[0157]視某些實施要求而定，本發明之實施例可以硬體或以軟體來實施。實施可使用具有上面儲存之電可讀控制信號的數位儲存媒體(例如，軟性磁碟、DVD、CD、ROM、PROM、EPROM、EEPROM 或快閃記憶體)來執行，該數位儲存媒體與可程式化電腦系統協作(或能夠與可程式化電腦系統協作)，使得各別方法被執行。

[0158]根據本發明之一些實施例包含具有電可讀控制信號的非暫時性資料載體，該非暫時性資料載體能夠與可程式化電腦系統協作，使得本文中所描述之方法中的一者被執行。

[0159]通常，本發明之實施例可實施為具有程式碼之電腦程式產品，程式碼係可操作的從而在電腦程式產品在電腦上執行時執行方法中的一者。程式碼可(例如)儲存於機器可讀載體上。

[0160]其他實施例包含儲存於機器可讀載體上之用於執行本文中描述之方法中的一者的電腦程式。

[0161]換言之，本發明方法之實施例因此係具有程式碼的電腦程式，該程式碼用於在電腦程式在電腦上執行時執行本文中所描述之方法中的一者。

[0162]本發明方法之其他實施例因此係資料載體(或數位儲存媒體或電腦可讀媒體)，該資料載體包含記錄於其上之用於執行本文中所描述之方法中的一者的電腦程式。

[0163]本發明方法之其他實施例因此係資料串流或信號序列，其表示用於執行本文中所描述之方法中之一者的電腦程式。資料串流或信號序列可(例如)經組配以經由資料通信連接(例如，經由網際網路)進行傳送。

[0164]其他實施例包含經組配或調適以執行本文中所描述之方法中之一者的處理構件，例如電腦或可程式化邏輯器件。

[0165]其他實施例包含上面安裝有用於執行本文中所

描述之方法中之一者的電腦程式之電腦。

[0166] 在一些實施例中，可程式化邏輯器件(例如，場可程式化閘陣列)可用以執行本文中所描述之方法之功能性的一些或全部。在一些實施例中，場可程式化閘陣列可與微處理器協作以便執行本文中所描述之方法中的一者。通常，方法較佳由任何硬體裝置執行。

[0167] 上述實施例係僅說明本發明之原理。應理解，本文中所描述之配置及細節的修改及變化對於熟習此項技術者將係顯而易見的。因此，意圖是僅由迫近之專利申請專利範圍之範疇限制，而非藉由借助於本文中之實施例之描述及解釋呈現的特定細節限制。

References

- [BCC] C. Faller and F. Baumgarte, "Binaural Cue Coding - Part II: Schemes and applications, " IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003.
- [JSC] C. Faller, "Parametric Joint-Coding of Audio Sources", 120th AES Convention, Paris, 2006.
- [SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: "From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio", 22nd Regional UK AES Conference, Cambridge, UK, April 2007.
- [SAOC2] J. Engdegård, B. Resch, C. Falch, O. Hellmuth, J.

Hilpert, A. Hölzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen: " Spatial Audio Object Coding (SAOC) - The Upcoming MPEG Standard on Parametric Object Based Audio Coding", 124th AES Convention, Amsterdam, 2008.

[SAOC] ISO/IEC, "MPEG audio technologies - Part 2: Spatial Audio Object Coding (SAOC), " ISO/IEC JTC1/SC29/WG11 (MPEG) International Standard 23003-2:2010.

[AAC] M .Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, "ISO/IEC MPEG-2 Advanced Audio Coding", J. Audio Eng. Soc, vol 45, no 10, pp. 789-814, 1997.

[ISS1] M. Parvaix and L. Girin: "Informed Source Separation of underdetermined instantaneous Stereo Mixtures using Source Index Embedding", IEEE ICASSP, 2010.

[ISS2] M. Parvaix, L. Girin, J.-M. Brossier: "A watermarking-based method for informed source separation of audio signals with a single sensor", IEEE Transactions on Audio, Speech and Language Processing, 2010.

[ISS3] A. Liutkus and J. Pinel and R. Badeau and L. Girin

and G. Richard: "Informed source separation through spectrogram coding and data embedding", Signal Processing Journal, 2011.

[ISS4] A. Ozerov, A. Liutkus, R. Badeau, G. Richard: "Informed source separation: source coding meets source separation", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2011.

[ISS5] S. Zhang and L. Girin: "An Informed Source Separation System for Speech Signals", INTERSPEECH, 2011.

[ISS6] L. Girin and J. Pinel: "Informed Audio Source Separation from Compressed Linear Stereo Mixtures", AES 42nd International Conference: Semantic Audio, 2011.

[ISS7] A. Nesbit, E. Vincent, and M. D. Plumbley: "Benchmarking flexible adaptive time-frequency transforms for underdetermined audio source separation", IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 37-40, 2009.

【符號說明】

10...SAOC 編碼器

16...下降混合器

12...SAOC 解碼器

17...旁側資訊估計器

18...下降混合信號	101...瞬變偵測單元
20...旁側資訊	102...產生窗序列單元
26...資訊	103...t/f 分析單元
30 _l 至 30 _k ...子頻帶信號	104...PSI 估計單元
32...小框/子頻帶值	105...粗糙功率頻譜重建單元
34...濾波器組時槽	106...功率頻譜估計單元
36...頻率軸	107...頻率解析度調適單元
38...時間軸	108...德耳塔估計單元
41...SAOC 訊框	109...德耳塔模型化單元
42...虛線	111...第一變換單元
45...模組	112...未混合資訊判定器
46...第二模組	113...未混合模組
51...下降混合部分	114...第二變換單元
52...參數旁側資訊部分	114...f/t 變換單元
61...編碼器	115...標準 PSI 解碼單元
62...解碼器	116...頻率解析度轉換單元
71...功率頻譜包絡	117...增強型 PSI 解碼單元
72、73、74...平均值	118...未混合矩陣產生器
81...差異曲線	119...去相關單元
91...下降混合單元	121...解碼子單元
92...變換單元	122...頻率解析度轉換子單元
93...參數旁側資訊產生器	123...增強型 PSI 解碼子單元
94...PSI 提取單元	124...組合器
95...PSI 分裂器	131...未混合矩陣計算器

132...時間內插器	151...原始音調信號
133...窗頻率解析度調適單元	152...原始雜訊信號
134...窗序列產生器	153...音調信號
135...t/f 分析模組	154...雜訊信號
136...未混合單元	921、922...t/f 變換單元
141...頻帶上值擴展單元	s_1 至 s_N ...音訊信號
142...德爾塔函數恢復單元	$\hat{s}_1$ 及 $\hat{s}_N$...音訊信號
143...德爾塔應用單元	$\hat{y}_1$ 至 $\hat{y}_M$...通道

申請專利範圍

1. 一種用於產生一未經混合音訊信號之解碼器，該未經混合音訊信號包含多個未經混合之音訊通道，其中該解碼器包含：

一未混合資訊判定器，其用於藉由接收關於至少一音訊物件信號之第一參數旁側資訊及關於該至少一音訊物件信號之第二參數旁側資訊來判定未混合資訊，其中該第二參數旁側資訊之頻率解析度高於該第一參數旁側資訊的頻率解析度，以及

一未混合模組，其用於將該未混合資訊應用於指示至少一音訊物件信號之一下降混合的一下降混合信號，以獲得包含該多個未經混合音訊通道的一未經混合音訊信號，

其中該未混合資訊判定器經組配以藉由修改該第一參數資訊及該第二參數資訊以獲得經修改參數資訊來判定該未混合資訊，使得該經修改參數資訊具有一高於該第一頻率解析度的頻率解析度。

2. 如請求項 1 之解碼器，

其中該解碼器進一步包含一第一變換單元，該第一變換單元用於變換一表示於一時域中之下降混合輸入，以獲得表示於一時間頻率域中的該下降混合信號，且

其中該解碼器包含一第二變換單元，該第二變換單

元用於將該未經混合音訊信號自該時間頻率域變換至該時域。

3. 如請求項 1 或 2 之解碼器，其中該未混合資訊判定器經組配以藉由組合該第一參數資訊與該第二參數資訊以獲得該經修改參數資訊來判定該未混合資訊，使得該經修改參數資訊具有一等於該第二頻率解析度的頻率解析度。

4. 如前述請求項中任一項之解碼器，

其中該第一參數資訊包含多個第一參數值，其中該第二參數資訊包含多個第二參數值，

其中該未混合資訊判定器包含一頻率解析度轉換子單元及一組合器，

其中該頻率解析度轉換單元經組配以產生額外參數值，其中該等第一參數值及該等額外參數值一起形成多個第一經處理參數值，且

其中該組合器經組配以組合該等第一經處理參數值與該等第二參數值以獲得多個經修改參數值作為該經修改參數資訊。

5. 如前述請求項中任一項之解碼器，

其中該未混合資訊判定器包含一德爾塔函數恢復單元及一德爾塔應用單元，

其中該第一參數資訊視該至少一音訊物件信號而包含多個參數值，且其中該第二參數資訊包含一校正因數參數化，

其中該德爾塔函數恢復單元經組配以反轉該校正因數參數化以獲得一德爾塔函數，且

其中該德爾塔應用單元經組配以對該等參數值應用該德爾塔函數以判定該未混合資訊。

6. 如請求項 5 之解碼器，

其中該校正因數參數化包含多個線性預測係數，

其中該德爾塔函數恢復單元經組配以藉由視該多個線性預測係數產生多個校正因數來反轉該校正因數參數化，且

其中該德爾塔函數恢復單元經組配以基於該多個校正因數產生該德爾塔函數。

7. 如前述請求項中任一項之解碼器，

其中該解碼器進一步包含一未混合矩陣產生器，該未混合矩陣產生器用於視該第一參數旁側資訊、視該第二參數旁側資訊且視呈現資訊而產生一未混合矩陣，且

其中該未混合模組經組配以對該經變換下降混合應用該未混合矩陣以獲得該未經混合音訊信號。

8. 如前述請求項中任一項之解碼器，

其中該未混合模組包含一去相關單元及一未混合單元，

其中該去相關單元經組配以對該經變換下降混合進行去相關以獲得一去相關結果，

且其中該未混合單元經組配以使用該去相關結果來獲得該未經混合音訊信號。

9. 一種用於編碼一或多個輸入音訊物件信號之編碼器，其包含：

一下降混合單元，其用於下降混合該一或多個輸入音訊物件信號以獲得一或多個下降混合信號，以及

一參數旁側資訊產生器，其用於產生關於至少一音訊物件信號之第一參數旁側資訊及關於該至少一音訊物件信號之第二參數旁側資訊，使得該第二參數旁側資訊之頻率解析度高於該第一參數旁側資訊的頻率解析度。

10. 如請求項 9 之編碼器，

其中該編碼器進一步包含一變換單元，該變換單元用於將該一或多個輸入音訊物件信號自一時域變換至一時間頻率域以獲得一或多個經變換音訊物件信號，且

其中該參數旁側資訊產生器經組配以基於該一或多個經變換音訊物件信號產生該第一參數旁側資訊及該第二參數旁側資訊。

11. 如請求項 10 之編碼器，

其中該變換單元經組配以視包含該一或多個輸入音訊物件信號中之至少一者之信號值的一信號變換區塊的一窗長度而將該一或多個輸入音訊物件信號自該時域變換至該時間頻率域，

其中該變換單元包含一瞬變偵測單元，該瞬變偵測單元用於判定一瞬變偵測結果從而指示一瞬變是否存在於該至少一音訊物件信號中的一或多者中，其中一瞬

變指示該至少一音訊物件信號中之一或多者中的一信號改變，且

其中該變換單元進一步包含一窗序列單元，該窗序列單元用於視該瞬變偵測結果而判定該窗長度。

12. 如請求項 9 至 11 中任一項之編碼器，其中該編碼器進一步包含一德爾塔估計單元，該德爾塔估計單元用於視該至少一音訊物件信號基於多個參數值而估計多個校正因數以獲得該第二參數旁側資訊。

13. 如請求項 12 之編碼器，其中該編碼器進一步包含一德爾塔模型化單元，該德爾塔模型化單元用於藉由進行一線性預測而自該多個校正因數來判定線性預測係數。

14. 一種經編碼音訊信號，其包含：

一下降混合部分，其指示一或多個輸入音訊物件信號的一下降混合，

一參數旁側資訊部分，其包含關於至少一音訊物件信號之第一參數旁側資訊及關於該至少一音訊物件信號的第二參數旁側資訊，其中該第二參數旁側資訊之頻率解析度高於該第一參數旁側資訊的頻率解析度。

15. 一種系統，其包含：

如請求項 9 至 13 中任一項之編碼器，其用於藉由以下操作來編碼一或多個輸入音訊物件信號：獲得指示一或多個輸入音訊物件信號之一下降混合的一或多個下降混合信號，獲得關於該至少一音訊物件信號之第一參數旁側資訊及獲得關於該至少一音訊物件信號的第

二參數旁側資訊，其中該第二參數旁側資訊的頻率解析度高於該第一參數旁側資訊之頻率解析度，以及

如請求項 1 至 8 中任一項之解碼器，其用於基於該一或多個下降混合信號且基於該第一參數旁側資訊及該第二參數旁側資訊來產生一未經混合音訊信號。

16. 一種用於產生一未經混合音訊信號之方法，該未經混合音訊信號包含多個未經混合音訊通道，其中該方法包含：

藉由接收關於至少一音訊物件信號之第一參數旁側資訊及關於該至少一音訊物件信號之第二參數旁側資訊來判定未混合資訊，其中該第二參數旁側資訊之頻率解析度高於該第一參數旁側資訊的頻率解析度，以及

將該未混合資訊應用於一指示至少一音訊物件信號之一下降混合的下降混合信號，以獲得包含該多個未經混合音訊通道的一未經混合音訊信號，

其中判定該未混合資訊包含修改該第一參數資訊及該第二參數資訊以獲得經修改之參數資訊，使得該經修改之參數資訊具有高於該第一頻率解析度的一頻率解析度。

17. 一種用於編碼一或多個輸入音訊物件信號之方法，其包含：

下降混合該一或多個輸入音訊物件信號以獲得一或多個下降混合信號，以及

產生關於至少一音訊物件信號之第一參數旁側資

訊及關於該至少一音訊物件信號之第二參數旁側資訊，使得該第二參數旁側資訊之頻率解析度高於該第一參數旁側資訊的頻率解析度。

18. 一種當在一電腦或信號處理器上執行時用於實施如請求項 16 或 17 之方法之電腦程式。

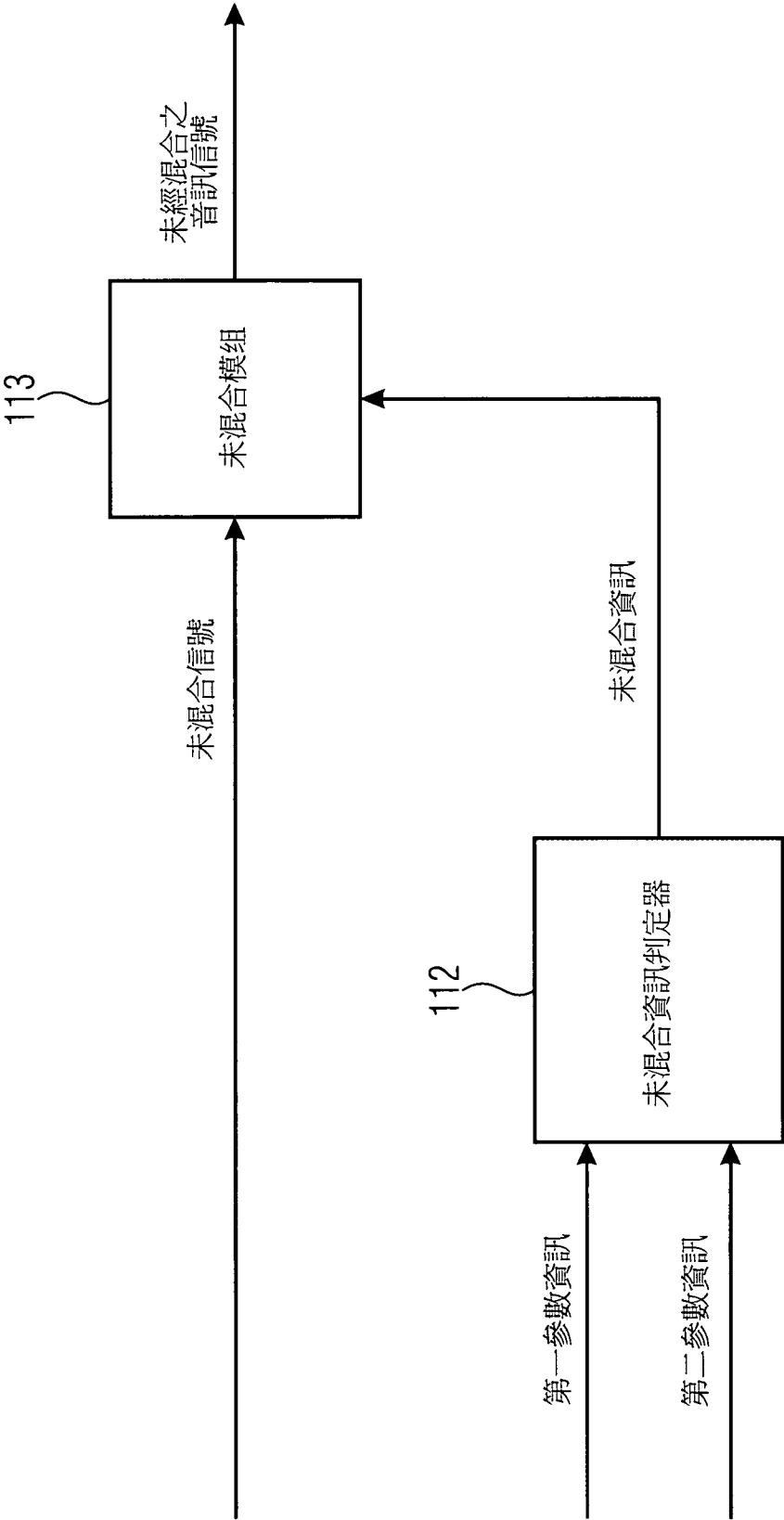


圖1a

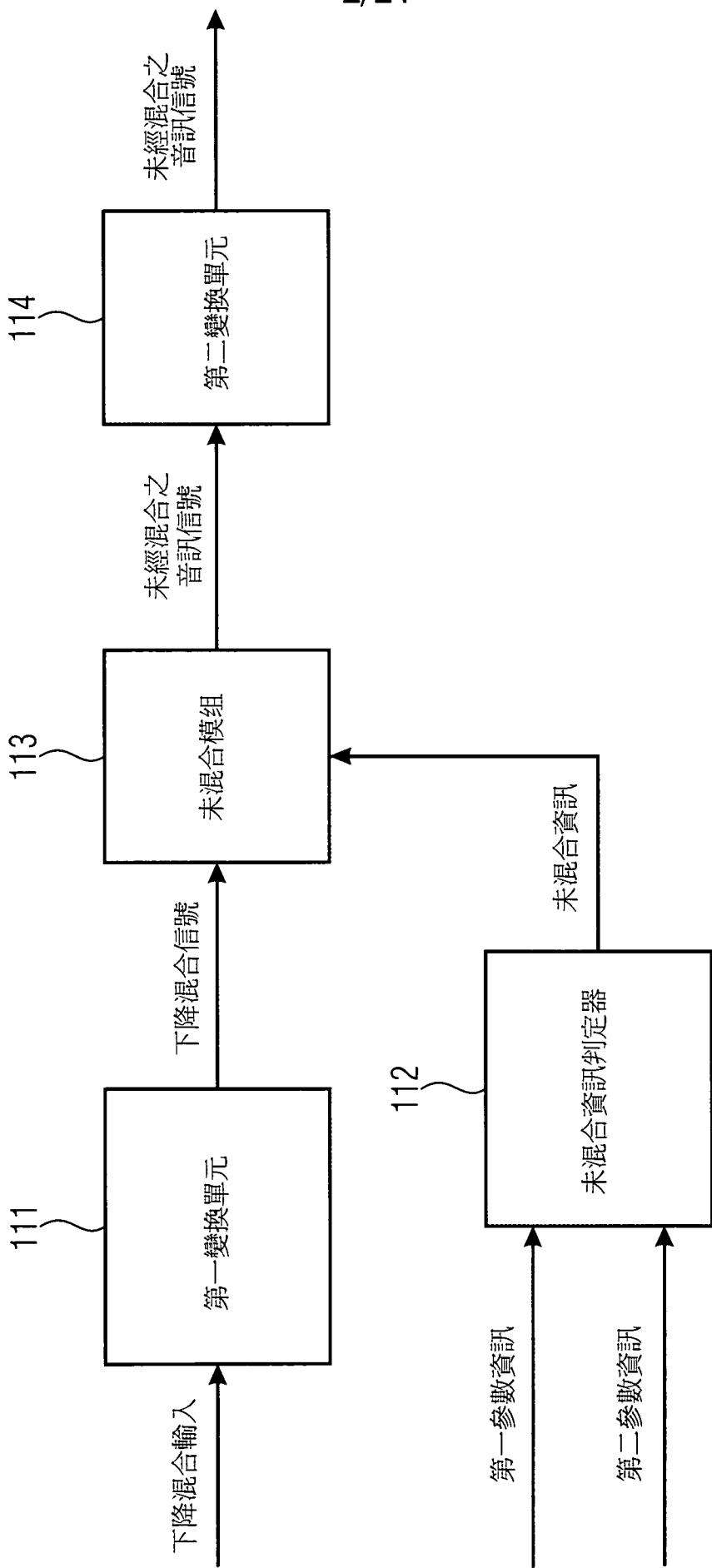


圖1b

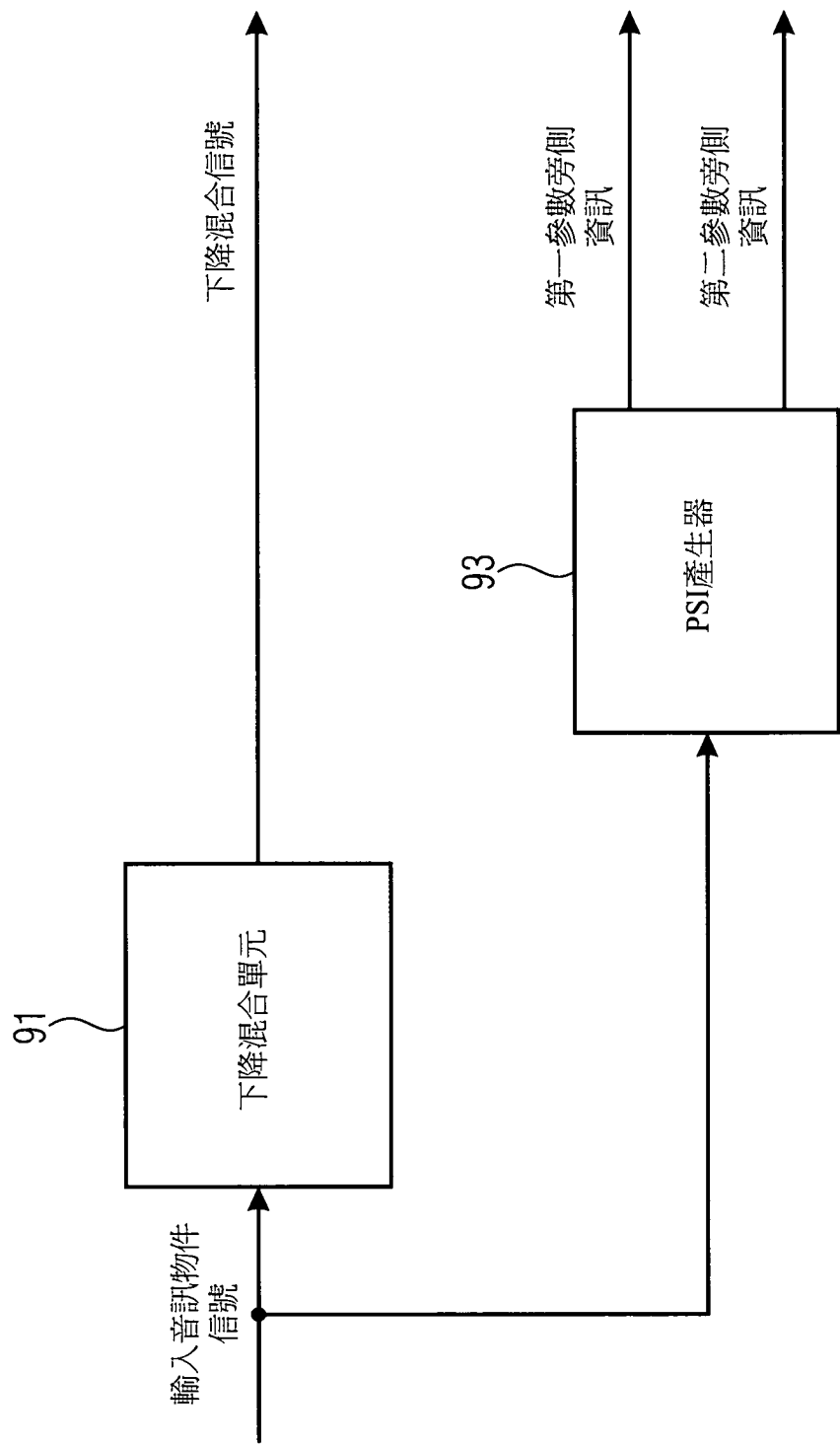


圖2a

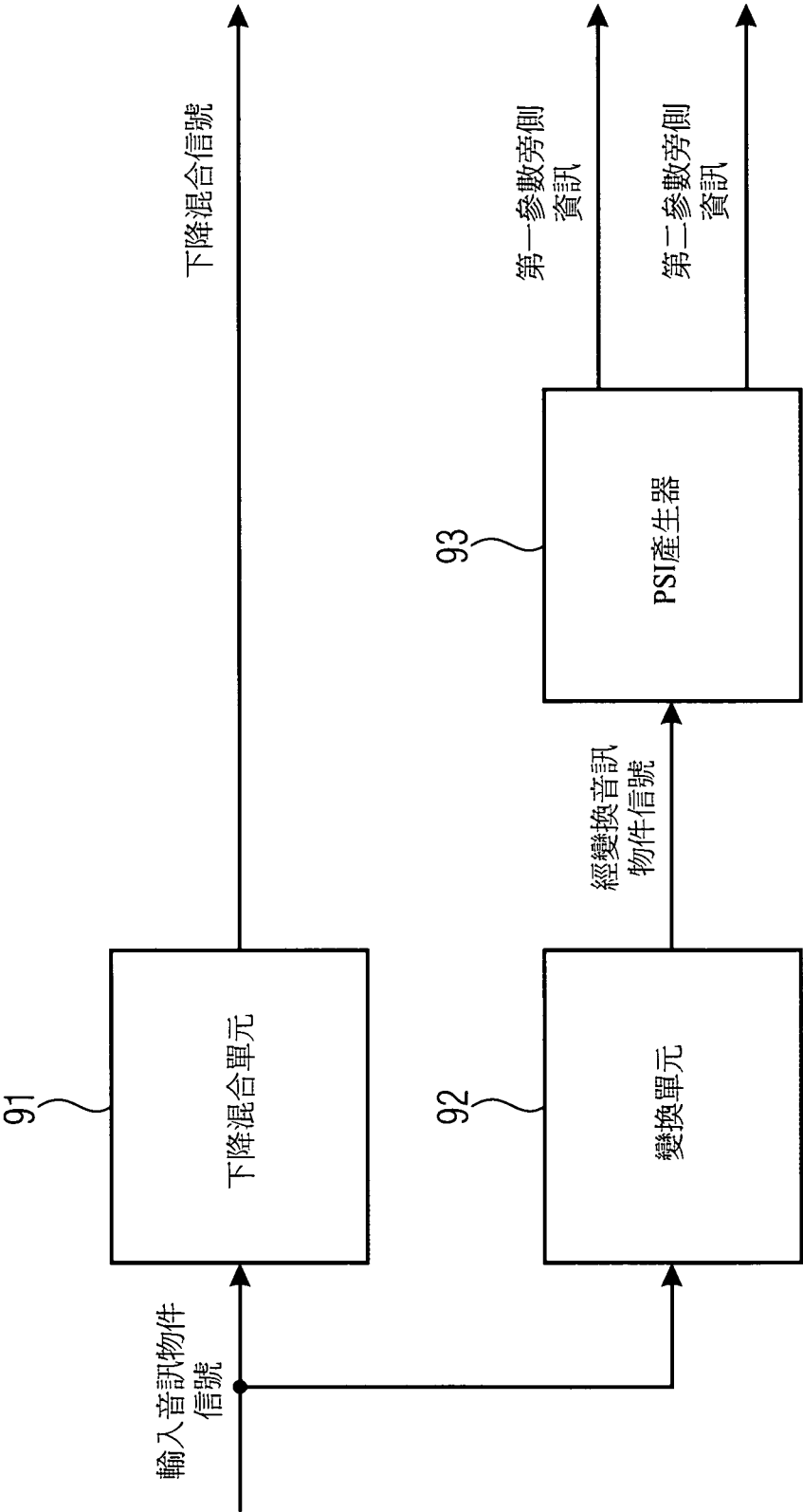
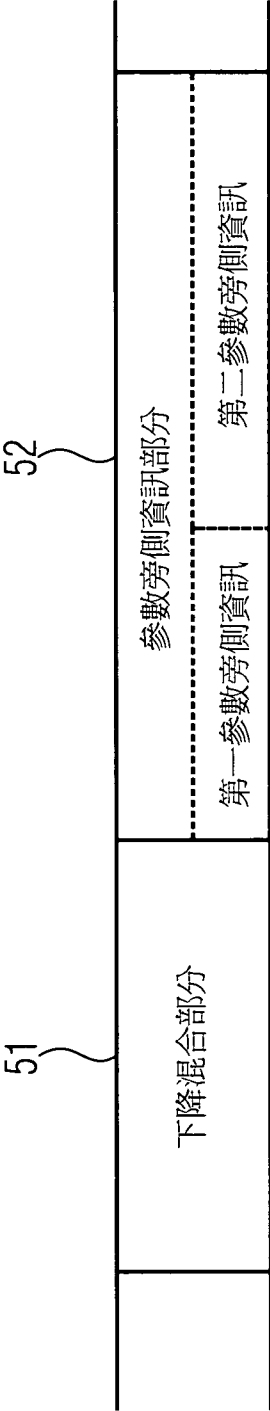


圖2b



第2c圖

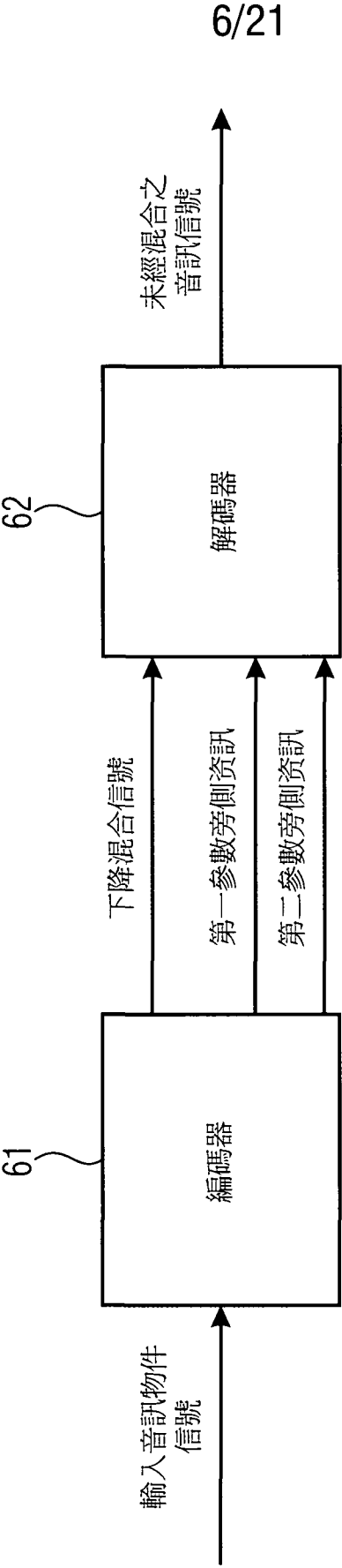


圖3

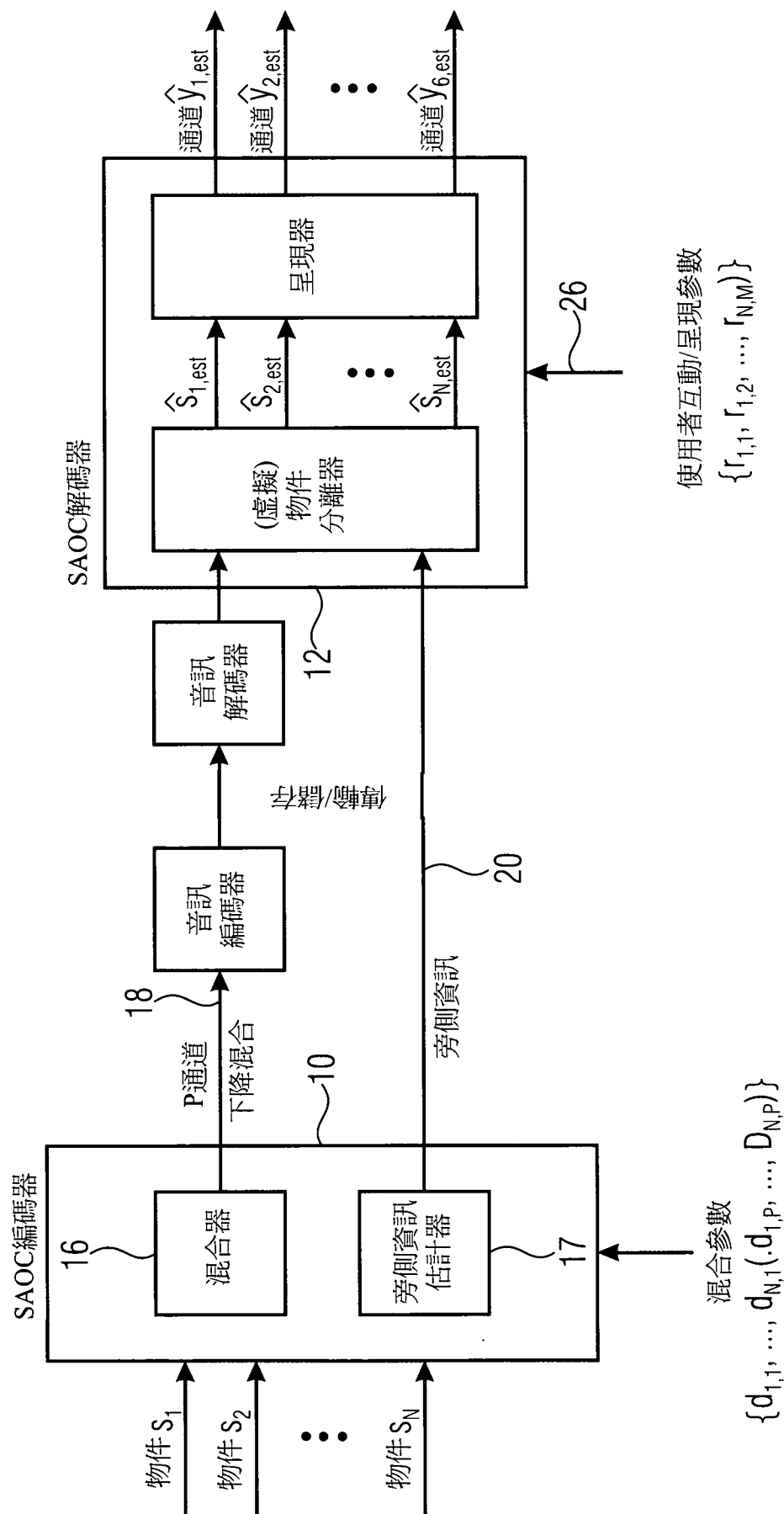


圖4

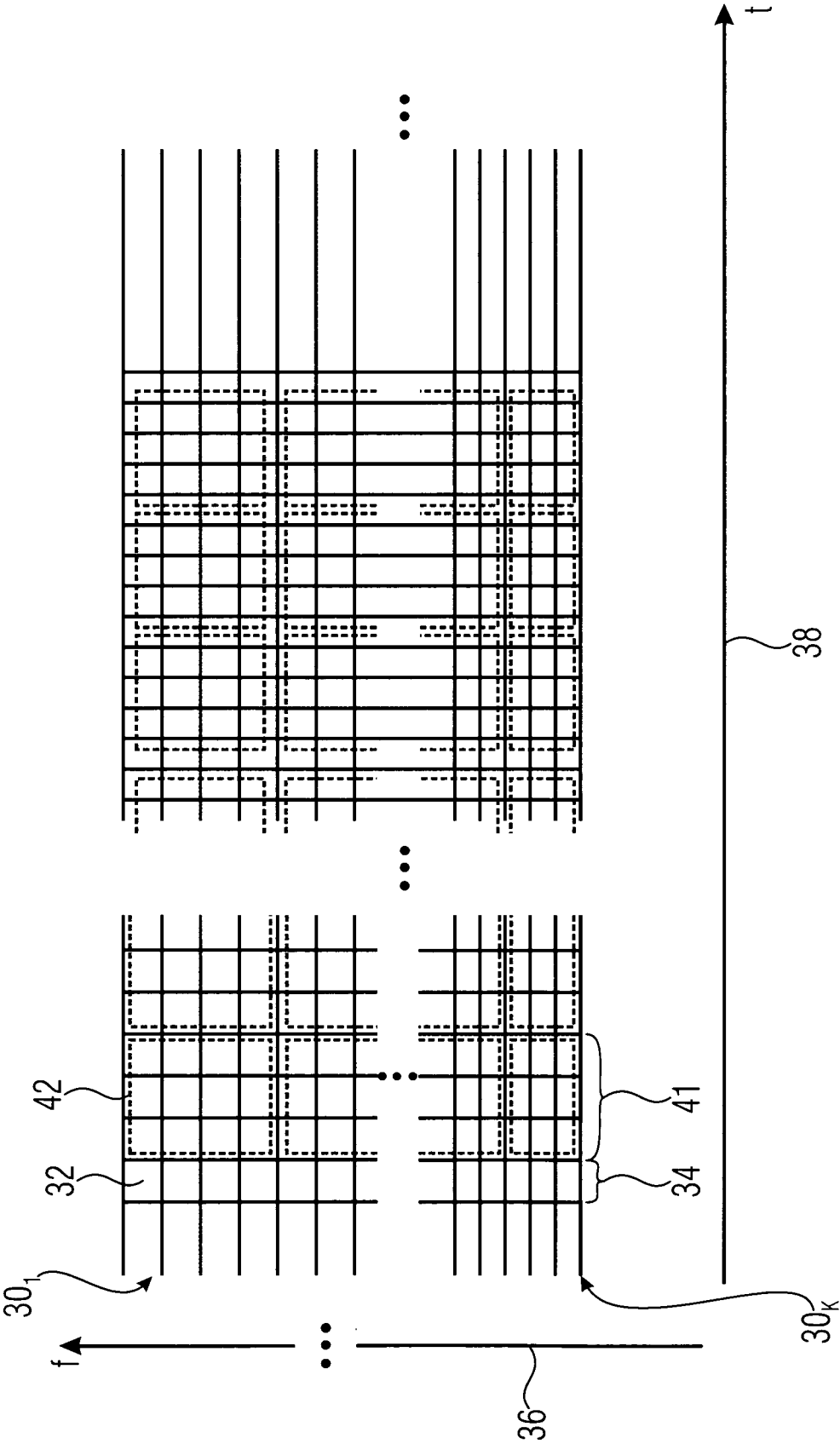


圖5

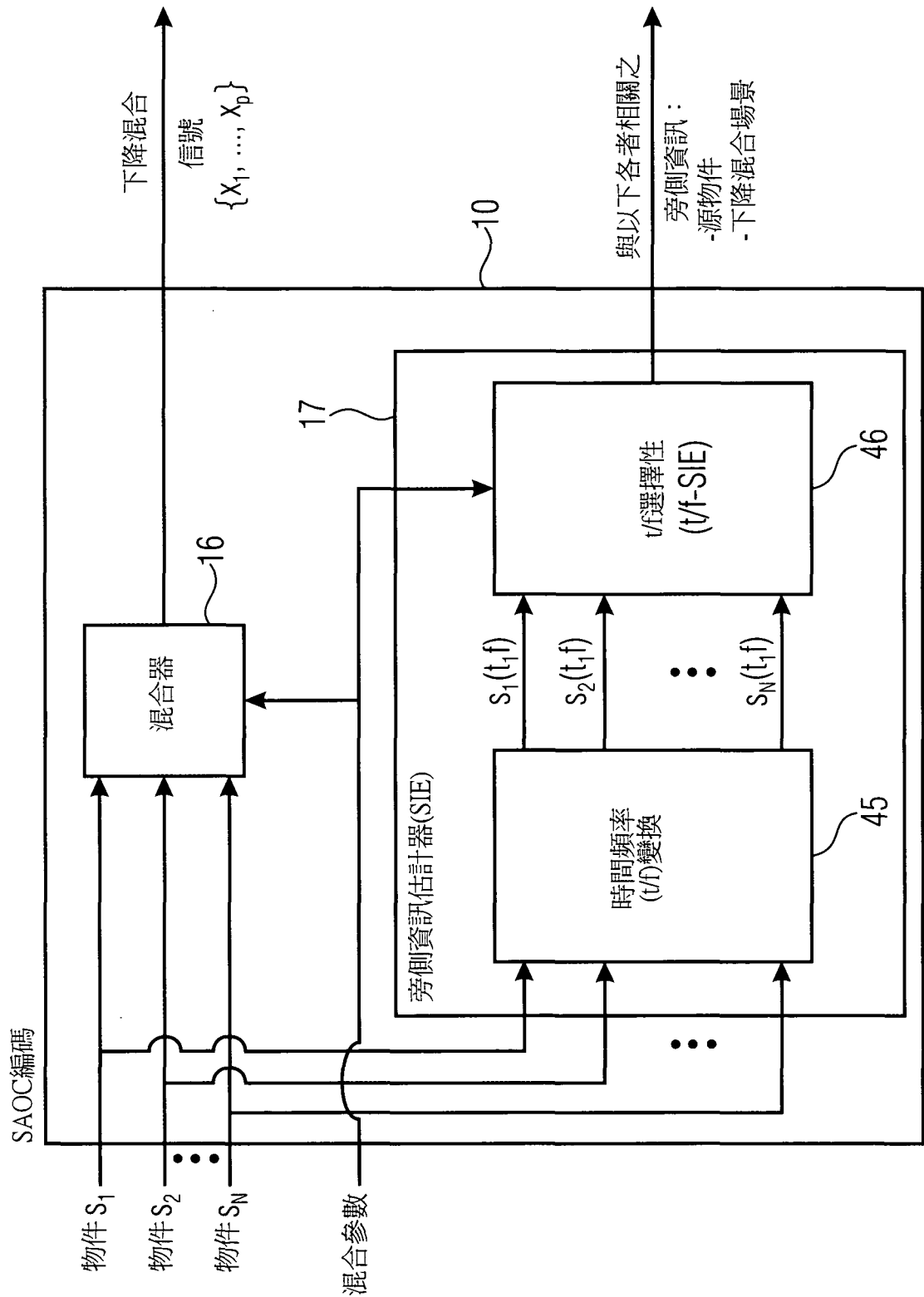


圖6

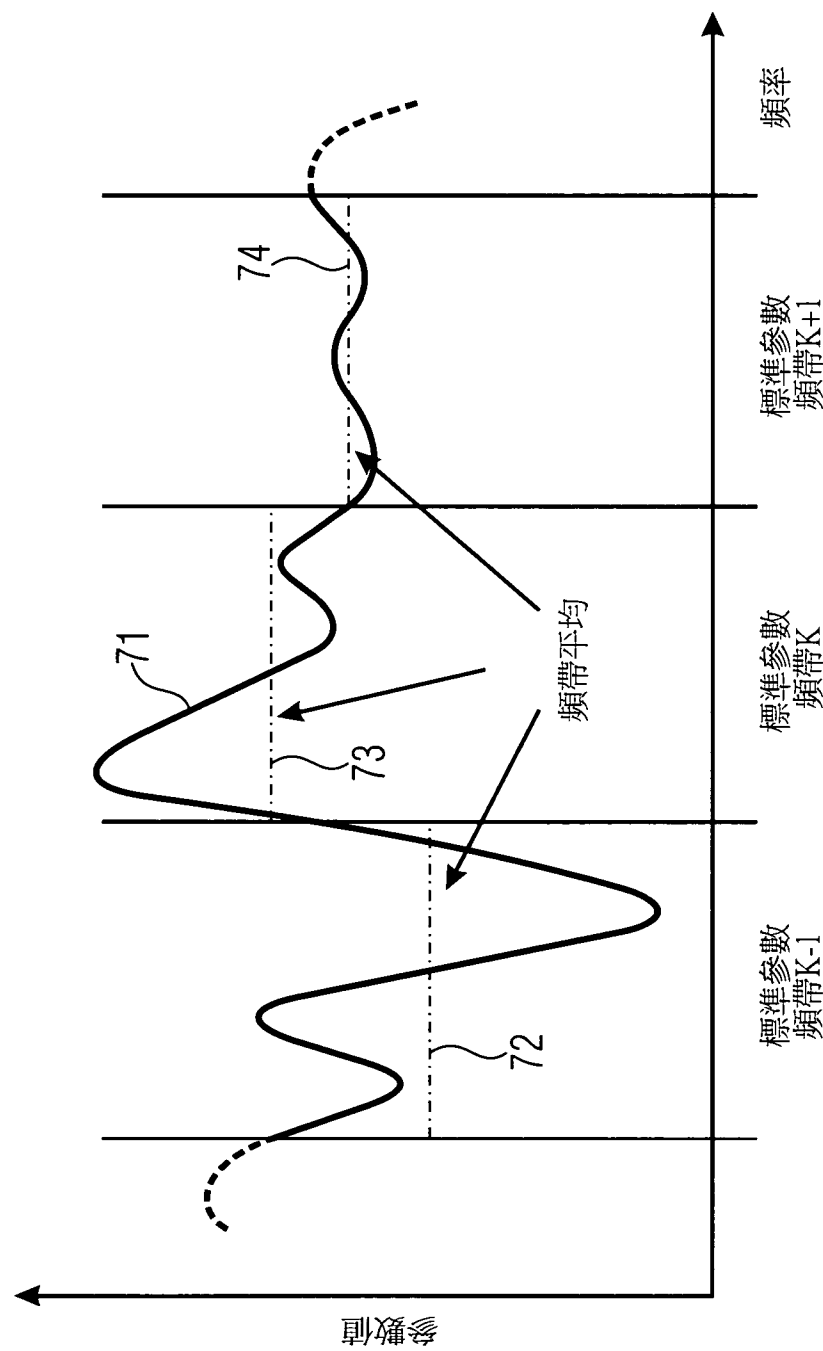


圖7

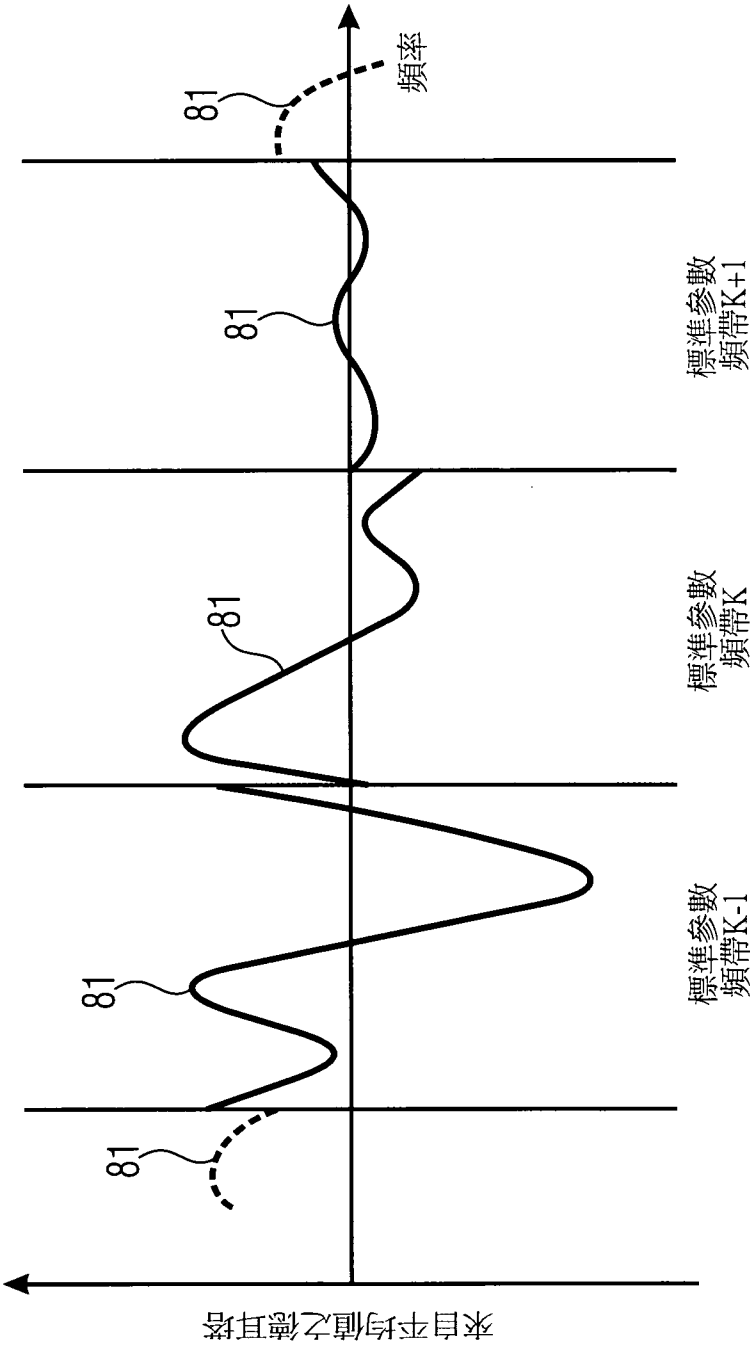


圖8

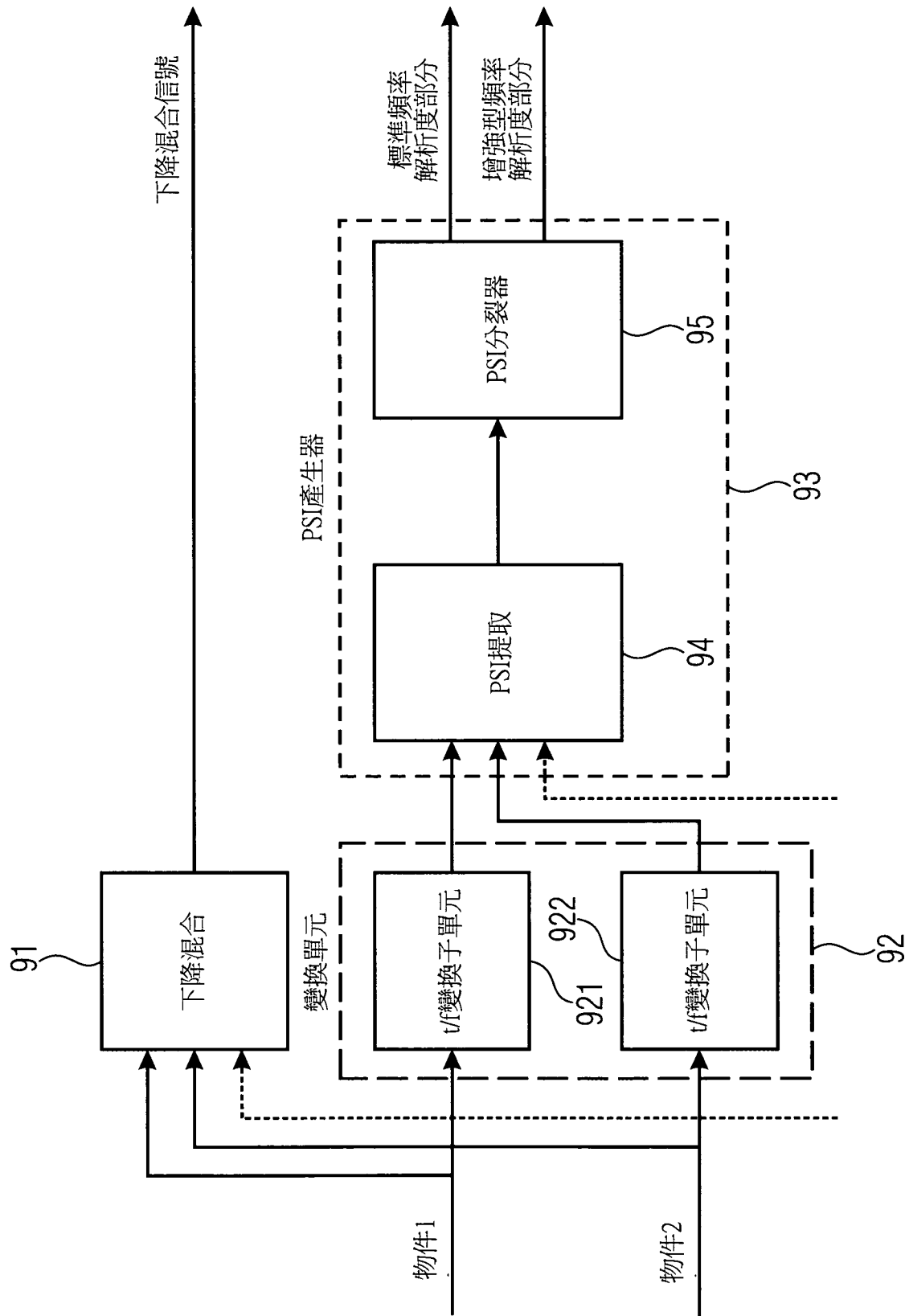


圖9

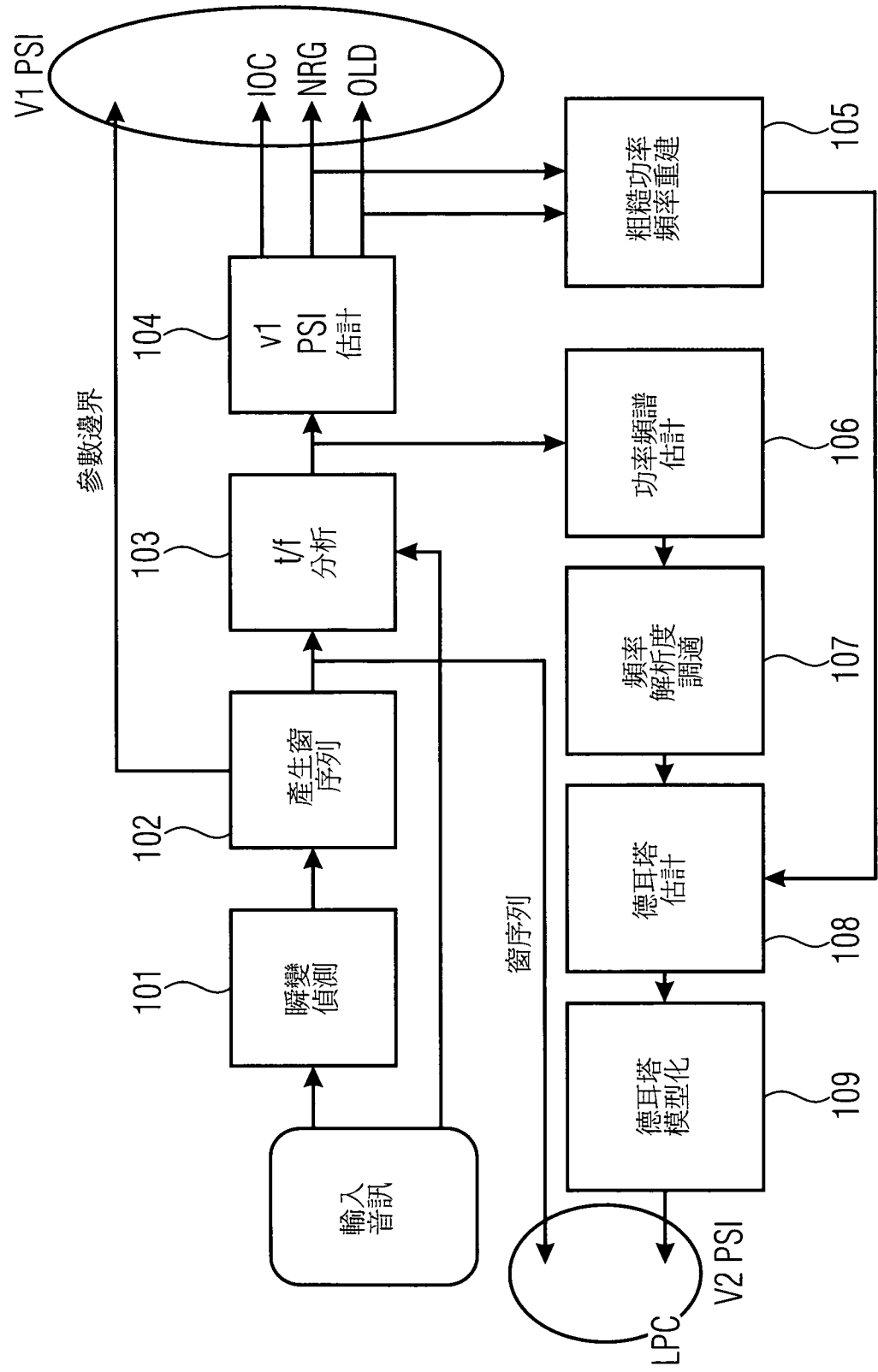


圖10

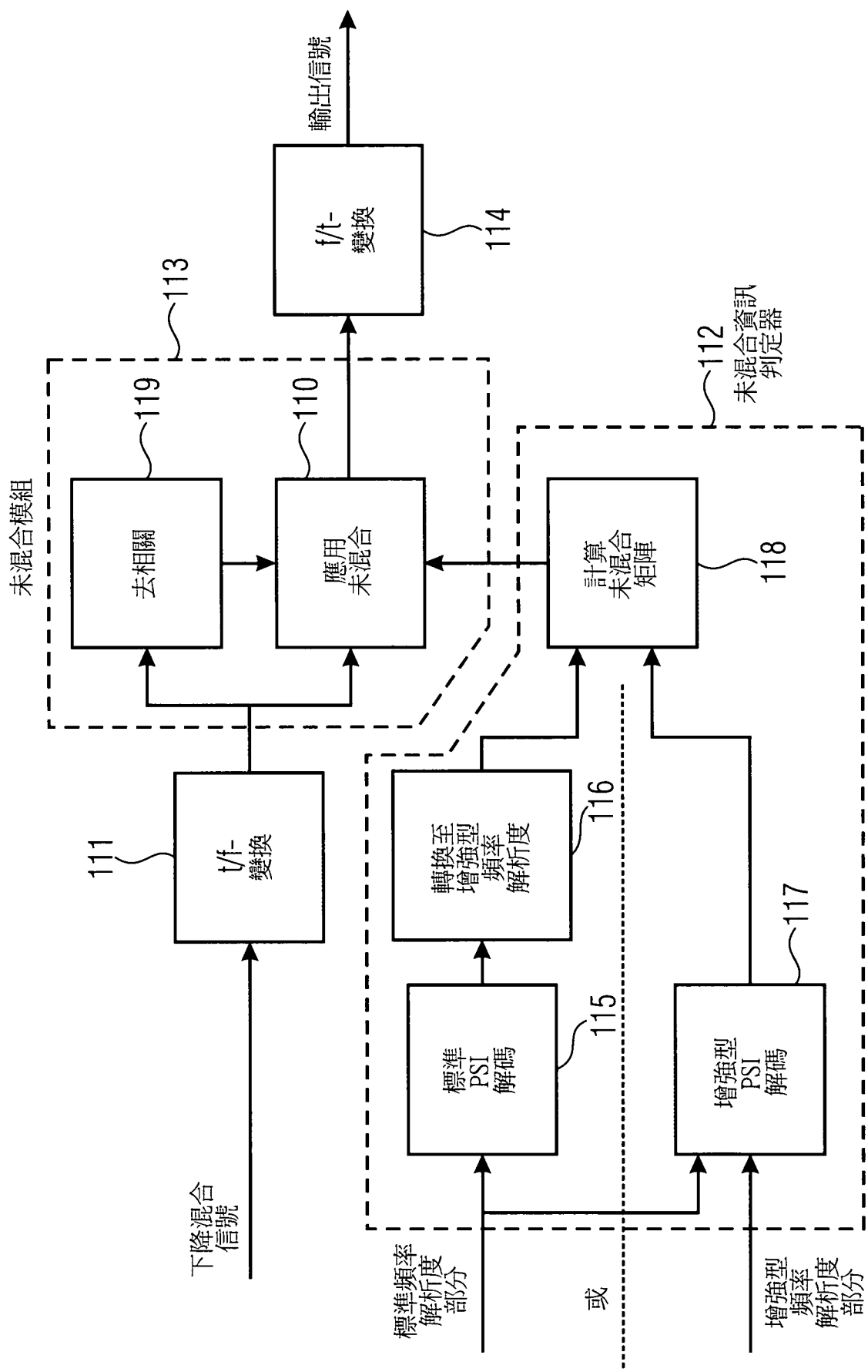


圖11

.....

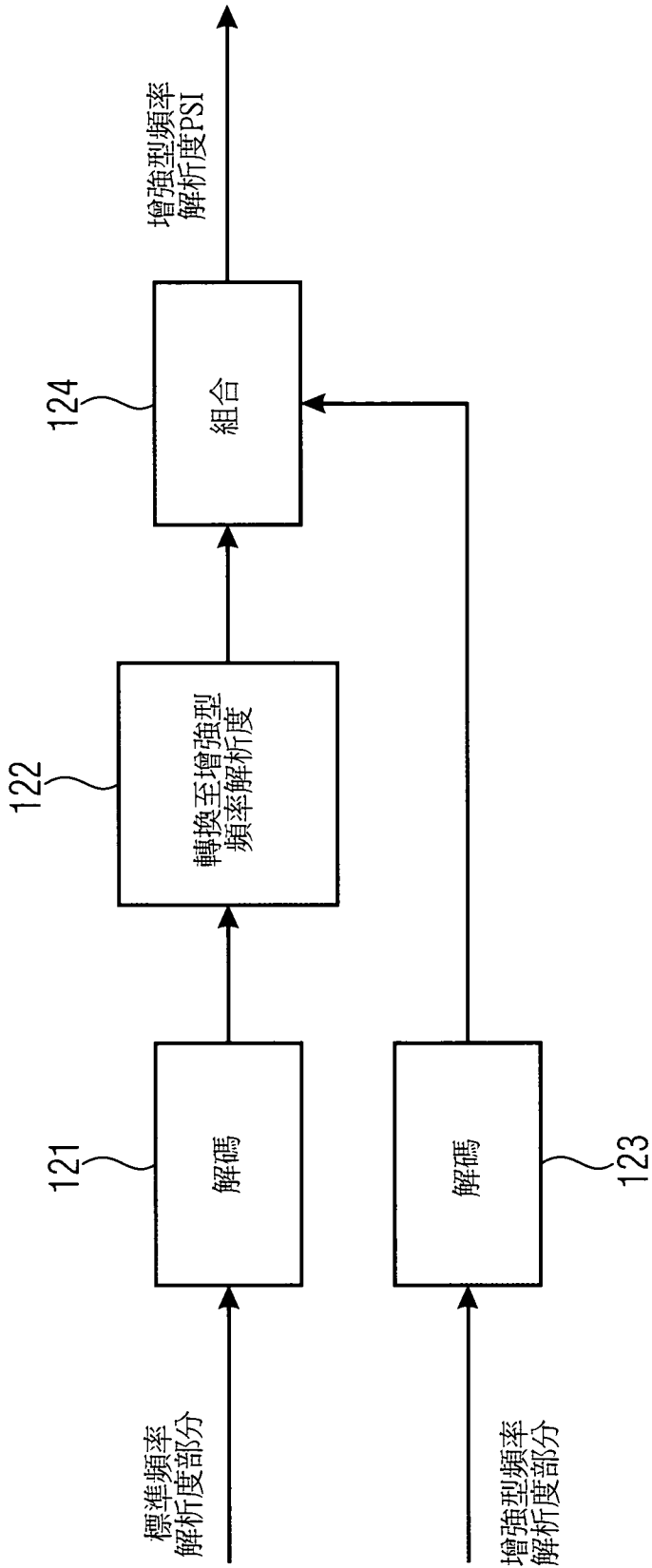


圖12

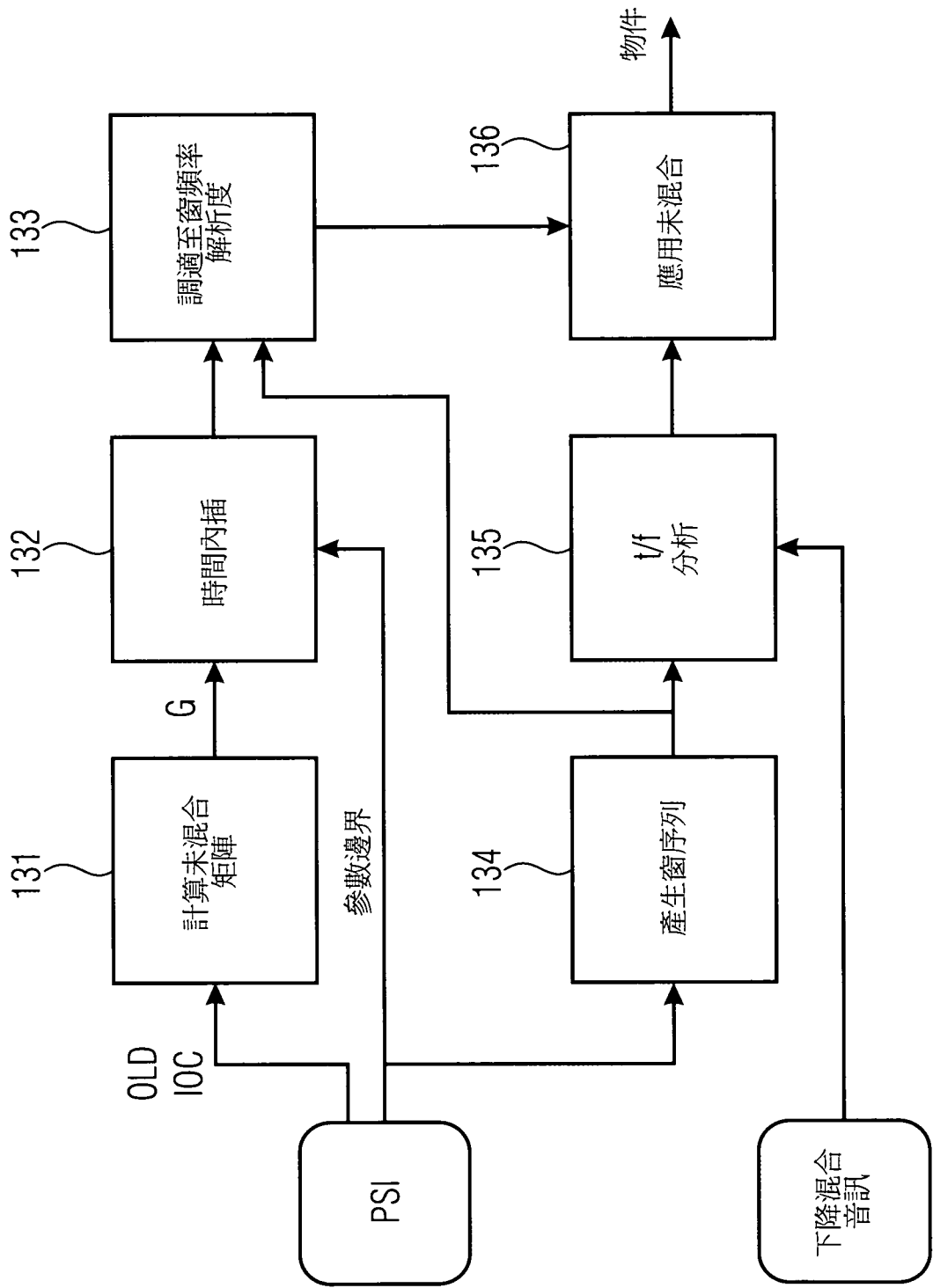
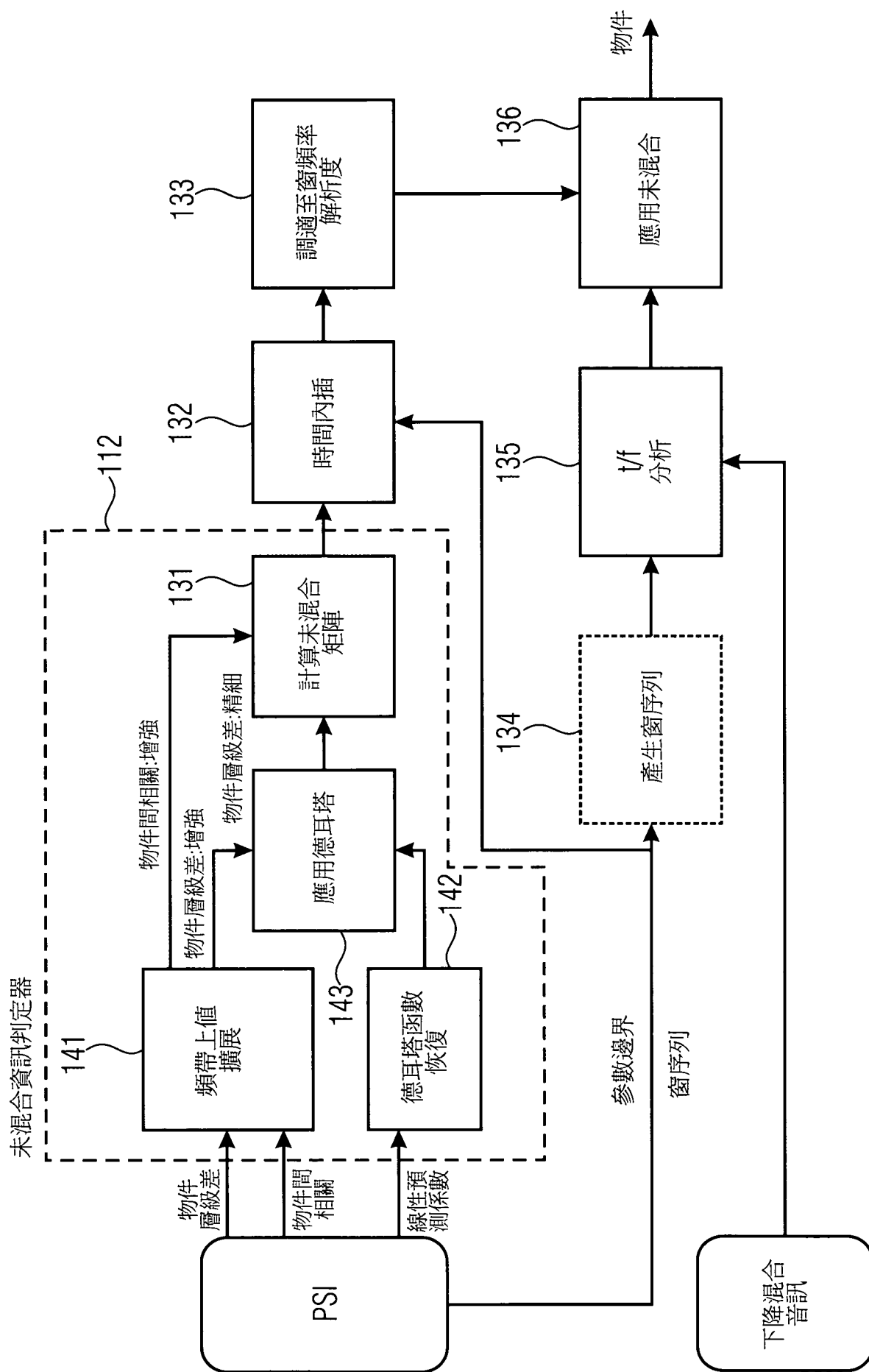


圖13

...



14回

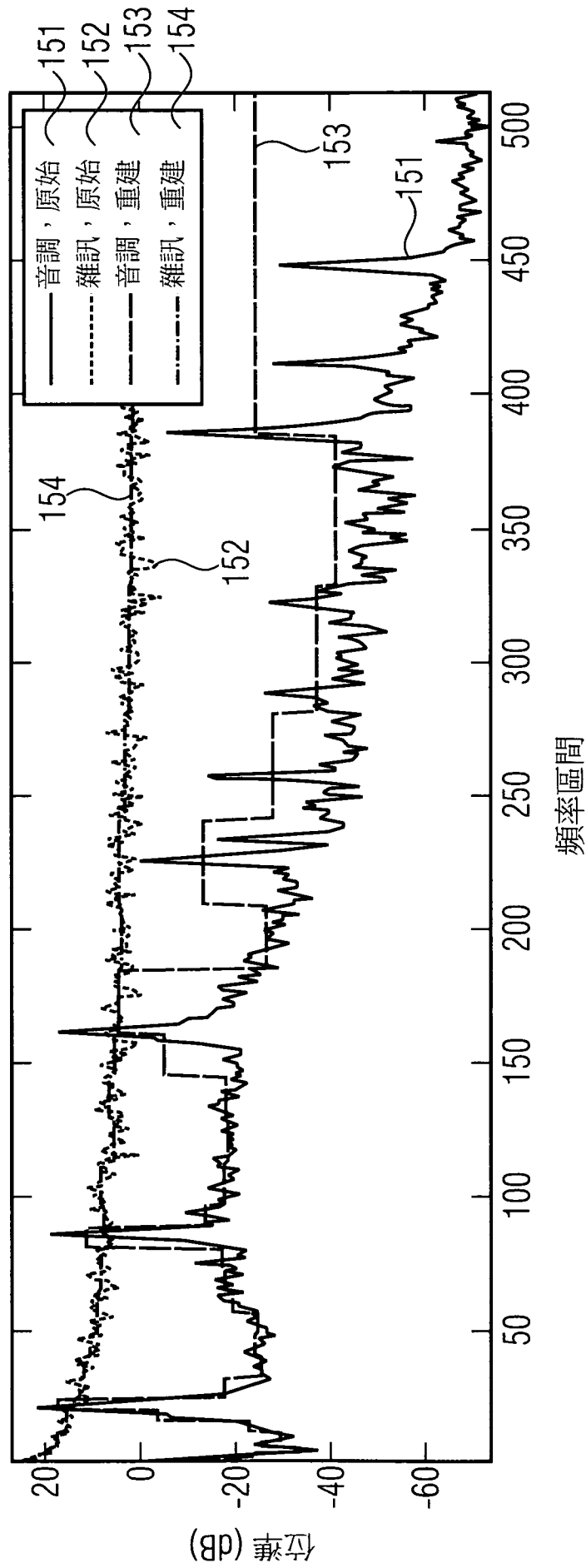


圖15

19/21

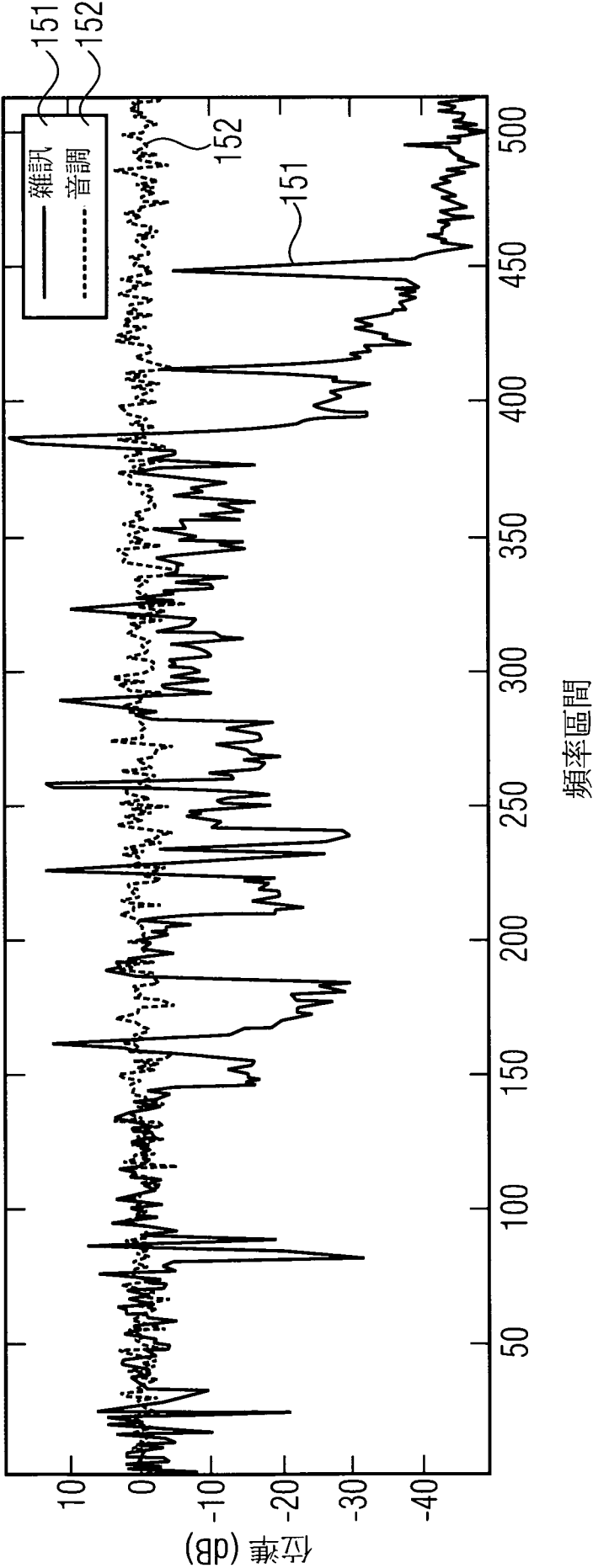


圖16

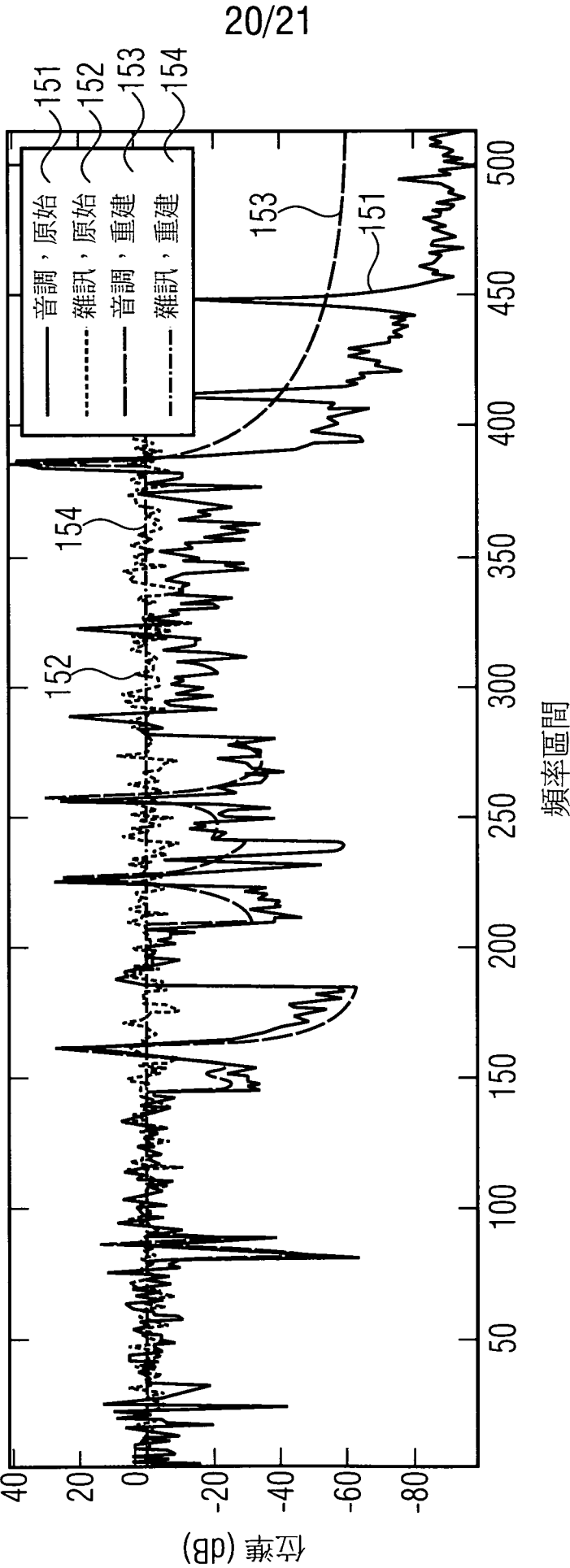


圖17

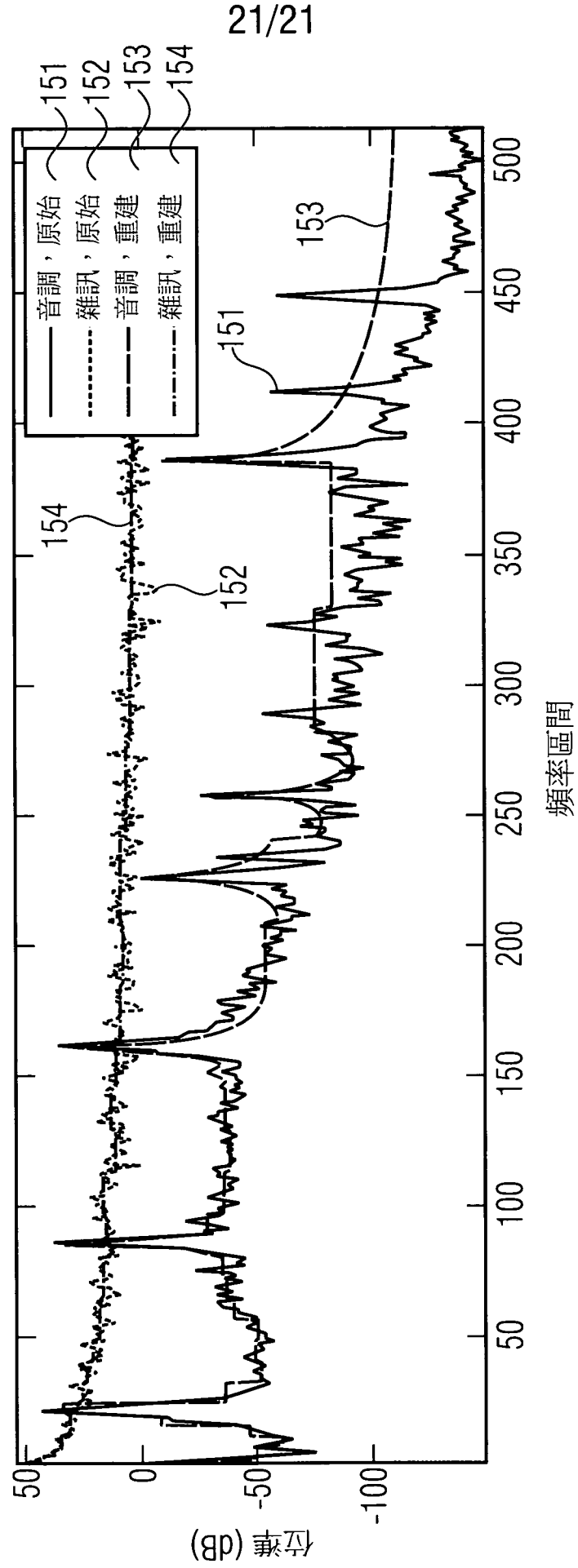


圖18