



(19)中華民國智慧財產局

(12)發明說明書公告本

(11)證書號數：TW I884423 B

(45)公告日：中華民國 114 (2025) 年 05 月 21 日

(21)申請案號：112106853

(22)申請日：中華民國 110 (2021) 年 07 月 29 日

(51)Int. Cl.：

*G10L19/008 (2013.01)**G10L19/012 (2013.01)**G10L19/16 (2013.01)**H04J3/17 (2006.01)*

(30)優先權：2020/07/30

歐洲專利局

20188707.2

2021/05/31

世界智慧財產權組織

PCT/EP2021/064576

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

德國

(72)發明人：福契斯 古拉米 FUCHS, GUILLAUME (FR)；塔瑪拉普 亞齊特 TAMARAPU, ARCHIT (IN)；尹申瑟 安德利亞 EICHENSEER, ANDREA (DE)；寇斯 斯里坎特 KORSE, SRIKANTH (IN)；多希拉 史蒂芬 DOEHLA, STEFAN (DE)；穆爾特 馬庫斯 MULTRUS, MARKUS (DE)

(74)代理人：劉法正；尹重君

(56)參考文獻：

TW 201911293A

CN 103180899A

CN 104318927A

US 2019/0385622A1

審查人員：施孝欣

申請專利範圍項數：15 項 圖式數：11 共 96 頁

(54)名稱

用以編碼音訊信號或用以解碼經編碼音訊場景之設備、方法及電腦程式

(57)摘要

本案揭示一種用於產生一經編碼音訊場景之設備，及一種用於解碼及/或處理一經編碼音訊場景之設備；以及相關方法及儲存指令之非暫時性儲存單元，該等指令在由一處理器執行時使該處理器執行一相關方法。

一種用於處理一經編碼音訊場景之設備可在一第一訊框中包含一第一聲場參數表示及一經編碼音訊信號，其中一第二訊框為一非作用訊框，該設備包含：

一活動偵測器，其用於偵測該第二訊框為該非作用訊框；

一合成信號合成器，其用於使用用於該第二訊框之參數描述來合成用於該第二訊框之一合成音訊信號；

一音訊解碼器，其用於解碼用於該第一訊框之該經編碼音訊信號；以及

一空間呈現器，其用於使用該第一聲場參數表示且使用用於該第二訊框之該合成音訊信號在空間上呈現用於該第一訊框之該音訊信號，或一轉碼器，其用於產生一元資料輔助輸出格式，該元資料輔助輸出格式包含用於該第一訊框之該音訊信號、用於該第一訊框之該第一聲場參數表示、用於該第二訊框之該合成音訊信號及用於該第二訊框之一第二聲場參數表示。

There are disclosed an apparatus for generating an encoded audio scene, and an apparatus for decoding and/or processing an encoded audio scene; as well as related methods and non-transitory storage units storing instructions which, when executed by a processor, cause the processor to perform a related method.

An apparatus (200) for processing an encoded audio scene (304) may comprise, in a first frame (346), a first soundfield parameter representation (316) and an encoded audio signal (346), wherein a second frame (348) is an inactive frame, the apparatus comprising:

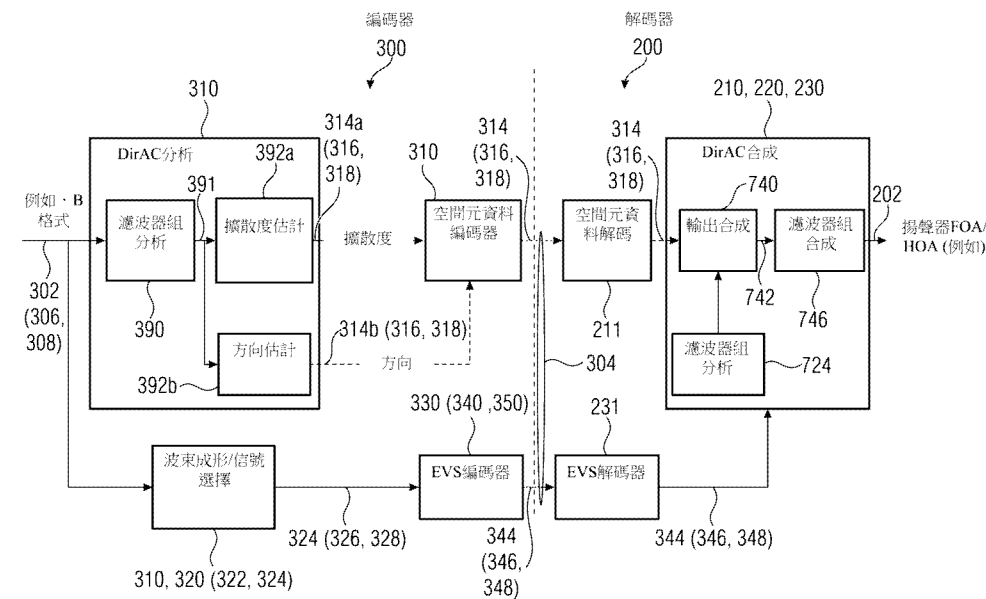
an activity detector (2200) for detecting that the second frame (348) is the inactive frame;

a synthetic signal synthesizer (210) for synthesizing a synthetic audio signal (228) for the second frame (308) using the parametric description (348) for the second frame (308);

an audio decoder (230) for decoding the encoded audio signal (346) for the first frame (306); and

a spatial renderer (240) for spatially rendering the audio signal (202) for the first frame (306) using the first soundfield parameter representation (316) and using the synthetic audio signal (228) for the second frame (308), or a transcoder for generating a meta data assisted output format comprising the audio signal (346) for the first frame (306), the first soundfield parameter representation (316) for the first frame (306), the synthetic audio signal (228) for the second frame (308), and a second soundfield parameter representation (318) for the second frame (308).

指定代表圖：



【圖2】

符號簡單說明：

- 200:解碼器/音訊解碼器設備
- 202:輸出聲道/最終輸出信號/輸出空間格式/空間音訊輸出信號/輸入信號/空間音訊輸入格式/音訊輸出格式/合成音訊信號
- 210:合成信號合成器/合成信號產生器/合成音訊合成器/第一部分/經恢復空間參數/舒適雜訊產生器/靜默插入描述符解碼器
- 211:空間元數據解碼
- 220:空間呈現器/空間呈現裝置
- 230:音訊解碼器/解碼裝置
- 231:EVS 解碼器
- 300:編碼器/音訊編碼器設備
- 302:輸入格式/輸入音訊信號/輸入版本/原始音訊輸入信號/輸入MASA 信號/輸入音訊

場景/B 格式輸入信號/
空間音訊輸入信號/多
聲道信號

304:經編碼音訊場景/
位元串流/參數表示

306:作用訊框/第一訊
框/輸入信號/作用階段

308:非作用訊框/第二
訊框/非作用階段

310:音訊場景分析器/
音訊信號分析器/

DirAC 分析區塊/聲場
參數產生器/場景音訊
分析器

314:聲場參數/作用空
間參數

314a:擴散度參數/擴散
度資訊/擴散度/DirAC
參數

314b:輸出/方向資訊/
到達方向/參數/DirAC
參數

316:作用空間參數/第
一聲場參數表示/第一
聲場參數/低位元速率
參數表示/DirAC 參數

318:非作用空間參數/
第二聲場參數表示/第
二聲場參數/低位元速
率參數表示/英特爾參
數/DirAC 參數

320:選擇器/區塊/話音
活動偵測器

322:第一偏差器

324:傳送聲道版本/傳
送聲道/降混版本/音訊
信號/降混信號/聲道信
號

326:傳送聲道/第一訊
框/經編碼音訊信號/聲

道信號/經寫碼音訊位
元串流/傳送聲道 SID/
降混版本/降混信號/作
用訊框

328:傳送聲道/第二訊
框/經編碼音訊信號/聲
道信號/降混信號/經寫
碼空間參數/參數描述/
非作用訊框/傳送聲道
SID/靜默資訊描述

330:音訊信號編碼器

340:傳送聲道編碼器/
區塊/傳送聲道編碼器
裝置

344:經編碼音訊信號/
編碼器音訊信號/傳送
聲道/經編碼版本/經編
碼資料

346:經編碼音訊信號/
第一訊框/作用訊框/經
編碼版本/經編碼資料/
作用階段

348:參數描述/第二訊
框/經編碼音訊信號/非
作用訊框/經編碼訊框/
經編碼版本/靜默參數
描述/傳送聲道 SID/非
作用階段/靜默插入描
述/靜默插入描述符訊
框

350:傳送聲道 SI 描述
符/區塊

390:濾波器組分析/濾
波器組分析區塊

391:輸出/頻率區間/頻
域資訊

392a:擴散度估計區塊/
擴散度分析區塊/階段

392b:方向估計區塊/方
向分析區塊/階段

- 724:濾波器組分析
- 740:混合區塊/混合/空間呈現器
- 742:混合信號
- 746:濾波器組合成區塊



I884423

【發明摘要】

【中文發明名稱】

用以編碼音訊信號或用以解碼經編碼音訊場景之設備、方法及電腦程式

【英文發明名稱】

APPARATUS, METHOD AND COMPUTER PROGRAM FOR ENCODING AN AUDIO SIGNAL OR FOR DECODING AN ENCODED AUDIO SCENE

【中文】

本案揭示一種用於產生一經編碼音訊場景之設備，及一種用於解碼及/或處理一經編碼音訊場景之設備；以及相關方法及儲存指令之非暫時性儲存單元，該等指令在由一處理器執行時使該處理器執行一相關方法。

一種用於處理一經編碼音訊場景之設備可在一第一訊框中包含一第一聲場參數表示及一經編碼音訊信號，其中一第二訊框為一非作用訊框，該設備包含：

一活動偵測器，其用於偵測該第二訊框為該非作用訊框；

一合成信號合成器，其用於使用用於該第二訊框之參數描述來合成用於該第二訊框之一合成音訊信號；

一音訊解碼器，其用於解碼用於該第一訊框之該經編碼音訊信號；

以及

一空間呈現器，其用於使用該第一聲場參數表示且使用用於該第二訊框之該合成音訊信號在空間上呈現用於該第一訊框之該音訊信號，或一轉碼器，其用於產生一元資料輔助輸出格式，該元資料輔助輸出格式包含用於該第一訊框之該音訊信號、用於該第一訊框之該第一聲場參數表示、用於該第二訊框之該合成音訊信號及用於該第二訊框之一第二聲場參數表示。

【英文】

There are disclosed an apparatus for generating an encoded audio scene, and an apparatus for decoding and/or processing an encoded audio scene; as well as related methods and non-transitory storage units storing instructions which, when executed by a processor, cause the processor to perform a related method.

An apparatus (200) for processing an encoded audio scene (304) may comprise, in a first frame (346), a first soundfield parameter representation (316) and an encoded audio signal (346), wherein a second frame (348) is an inactive frame, the apparatus comprising:

- an activity detector (2200) for detecting that the second frame (348) is the inactive frame;

- a synthetic signal synthesizer (210) for synthesizing a synthetic audio signal (228) for the second frame (308) using the parametric description (348) for the second frame (308);

- an audio decoder (230) for decoding the encoded audio signal (346) for the first frame (306); and

- a spatial renderer (240) for spatially rendering the audio signal (202) for the first frame (306) using the first soundfield parameter representation (316) and using the synthetic audio signal (228) for the second frame (308), or a transcoder for generating a meta data assisted output format comprising the audio signal (346) for the first frame (306), the first soundfield parameter representation (316) for the first frame (306), the synthetic audio signal (228) for the second frame (308), and a second soundfield parameter representation (318) for the second frame (308).

【指定代表圖】 圖2

【代表圖之符號簡單說明】

- 200: 解碼器/音訊解碼器設備
- 202: 輸出聲道/最終輸出信號/輸出空間格式/空間音訊輸出信號/輸入信號/空間音訊輸入格式/音訊輸出格式/合成音訊信號
- 210: 合成信號合成器/合成信號產生器/合成音訊合成器/第一部分/經恢復空間參數/舒適雜訊產生器/靜默插入描述符解碼器
- 211: 空間元數據解碼
- 220: 空間呈現器/空間呈現裝置
- 230: 音訊解碼器/解碼裝置
- 231: EVS 解碼器
- 300: 編碼器/音訊編碼器設備
- 302: 輸入格式/輸入音訊信號/輸入版本/原始音訊輸入信號/輸入 MASA 信號/輸入音訊場景/B 格式輸入信號/空間音訊輸入信號/多聲道信號
- 304: 經編碼音訊場景/位元串流/參數表示
- 306: 作用訊框/第一訊框/輸入信號/作用階段
- 308: 非作用訊框/第二訊框/非作用階段
- 310: 音訊場景分析器/音訊信號分析器/DirAC 分析區塊/聲場參數產生器/場景音訊分析器
- 314: 聲場參數/作用空間參數
- 314a: 擴散度參數/擴散度資訊/擴散度/DirAC 參數
- 314b: 輸出/方向資訊/到達方向/參數/DirAC 參數
- 316: 作用空間參數/第一聲場參數表示/第一聲場參數/低位元速率參數表示/DirAC 參數
- 318: 非作用空間參數/第二聲場參數表示/第二聲場參數/低位元速率參數表示/英特爾參數/DirAC 參數
- 320: 選擇器/區塊/話音活動偵測器
- 322: 第一偏差器
- 324: 傳送聲道版本/傳送聲道/降混版本/音訊信號/降混信號/聲道信號
- 326: 傳送聲道/第一訊框/經編碼音訊信號/聲道信號/經寫碼音訊位元串流/傳送聲道 SID/降混版本/降混信號/作用訊框
- 328: 傳送聲道/第二訊框/經編碼音訊信號/聲道信號/降混信號/經寫碼空間參數/參數描述/非作用訊框/傳送聲道 SID/靜默資訊描述
- 330: 音訊信號編碼器
- 340: 傳送聲道編碼器/區塊/傳送聲道編碼器裝置

- 344: 經編碼音訊信號/編碼器音訊信號/傳送聲道/經編碼版本/經編碼資料
- 346: 經編碼音訊信號/第一訊框/作用訊框/經編碼版本/經編碼資料/作用階段
- 348: 參數描述/第二訊框/經編碼音訊信號/非作用訊框/經編碼訊框/經編碼版本/靜默參數描述/傳送聲道 SID/非作用階段/靜默插入描述/靜默插入描述符訊框
- 350: 傳送聲道 SI 描述符/區塊
- 390: 濾波器組分析/濾波器組分析區塊
- 391: 輸出/頻率區間/頻域資訊
- 392a: 擴散度估計區塊/擴散度分析區塊/階段
- 392b: 方向估計區塊/方向分析區塊/階段
- 724: 濾波器組分析
- 740: 混合區塊/混合/空間呈現器
- 742: 混合信號
- 746: 濾波器組合成區塊

【特徵化學式】

無

【發明說明書】

【中文發明名稱】

用以編碼音訊信號或用以解碼經編碼音訊場景之設備、方法及電腦程式

【英文發明名稱】

APPARATUS, METHOD AND COMPUTER PROGRAM FOR ENCODING
AN AUDIO SIGNAL OR FOR DECODING AN ENCODED AUDIO SCENE

【技術領域】

發明領域

【0001】 本文尤其係指一種用於產生經編碼音訊場景之設備，且係指一種用於解碼及/或處理經編碼音訊場景之設備。本文亦指相關方法及儲存指令之非暫時性儲存單元，該等指令在由處理器執行時使處理器執行相關方法。

【0002】 本文論述關於音訊場景之不連續傳輸模式(DTX)及舒適雜訊產生(CNG)之方法，對於該等音訊場景，空間影像由方向音訊寫碼(DirAC)範式以參數方式寫碼或以元資料輔助空間音訊(MASA)格式傳輸。

【0003】 實施例係關於以參數方式寫碼之空間音訊之不連續傳輸，諸如DirAC 及 MASA 之 DTX 模式。

【0004】 本發明之實施例係關於有效傳輸及呈現例如藉由聲場麥克風捕捉之會話語音。因此所捕捉之音訊信號通常稱為三維(3D)音訊，此係由於聲音事件可在三維空間中局域化，其加強沉浸感且提高可懂度及使用者體驗二者。

【0005】 例如在三維中傳輸音訊場景需要處置通常引起大量資料傳輸之多個聲道。舉例而言，方向音訊寫碼(DirAC)技術[1]可用於降低較大原始資料速率。DirAC 被視為用於分析音訊場景且以參數方式表示該音訊場景之高效方法。藉助於每頻帶所量測之到達方向(DOA)及擴散度，在感知上促動及表示聲場。該

聲場係依據如下假定建構：在一個瞬間且對於一個臨界頻帶，聽覺系統之空間解析度限於解碼一個方向提示及另一耳間相干性提示。空間聲音隨後藉由使二個串流：非方向性擴散串流及方向性不擴散串流交叉漸進而在頻域中再現。

【0006】 此外，在典型的會話中，每一揚聲器在約百分之六十的時間內靜默。藉由區分含有語音(「作用訊框」)之音訊信號的訊框與僅含有背景雜訊或靜默(「非作用訊框」)之訊框，語音寫碼器可節省有效資料速率。非作用訊框通常被感知為攜載極少資訊或不攜載資訊，且語音寫碼器通常經組配以減小其用於此類訊框之位元速率，或甚至不傳輸資訊。在此情況下，寫碼器在所謂的不連續傳輸(DTX)模式中運行，其為在不存在話音輸入之情況下大幅度減小通信編解碼器之傳輸速率的高效方式。在此模式下，經判定為僅由背景雜訊組成之大部分訊框自傳輸丟棄且由解碼器中之一些舒適雜訊產生(CNG)替換。對於此等訊框，信號之極低速率參數表示係藉由定期但並非在每一訊框處發送之靜默插入描述符(SID)訊框傳送。此允許解碼器中之 CNG 產生類似於實際背景雜訊之人工雜訊。

【0007】 本發明之實施例係關於例如由聲場麥克風捕捉且可藉由寫碼方案基於 DirAC 範式及其類似者以參數方式寫碼的 3D 音訊場景之 DTX 系統且尤其是 SID 及 CNG。本發明實現對傳輸會話式沉浸式語音之位元速率需求的急劇減少。

【先前技術】

發明背景

【0008】 [1] V. Pulkki, M-V. Laitinen, J. Vilkkamo, J. Ahonen, T. Lokki, and T. Pihlajamäki, "Directional audio coding - perception-based reproduction of spatial sound", International Workshop on the Principles and Application on Spatial Hearing, Nov. 2009, Zao; Miyagi, Japan.

[2] 3GPP TS 26.194; Voice Activity Detector (VAD); - 3GPP technical specification Retrieved on 2009-06-17.

[3] 3GPP TS 26.449, "Codec for Enhanced Voice Services (EVS); Comfort Noise Generation (CNG) Aspects".

[4] 3GPP TS 26.450, "Codec for Enhanced Voice Services (EVS); Discontinuous Transmission (DTX)"

[5] A. Lombard, S. Wilde, E. Ravelli, S. Döhla, G. Fuchs and M. Dietz, "Frequency-domain Comfort Noise Generation for Discontinuous Transmission in EVS," *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, QLD, 2015, pp. 5893-5897, doi: 10.1109/ICASSP.2015.7179102.

[6] V. Pulkki, "Virtual source positioning using vector base amplitude panning", *J. Audio Eng. Soc.*, 45(6):456-466, June 1997.

[7] J. Ahonen and V. Pulkki, "Diffuseness estimation using temporal variation of intensity vectors", in *Workshop on Applications of Signal Processing to Audio and Acoustics WASPAA*, Mohonk Mountain House, New Paltz, 2009.

[8] T. Hirvonen, J. Ahonen, and V. Pulkki, "Perceptual compression methods for metadata in Directional Audio Coding applied to audiovisual teleconference", *AES 126th Convention 2009*, May 7-10, Munich, Germany.

[9] Vilkamo, Juha & Bäckström, Tom & Kuntz, Achim. (2013). Optimized Covariance Domain Framework for Time--Frequency Processing of Spatial Audio. *Journal of the Audio Engineering Society*. 61.

[10] M. Laitinen and V. Pulkki, "Converting 5.1 audio recordings to B-format for directional audio coding reproduction," *2011 IEEE International Conference on*

Acoustics, Speech and Signal Processing (ICASSP), Prague, 2011, pp. 61-64, doi: 10.1109/ICASSP.2011.5946328.

【發明內容】

發明概要

【0009】 根據一態樣，提供一種用於自具有第一訊框及第二訊框之音訊信號產生經編碼音訊場景的設備，其包含：

聲場參數產生器，其用於根據第一訊框中之音訊信號判定用於第一訊框之第一聲場參數表示，且根據第二訊框中之音訊信號判定用於第二訊框之第二聲場參數表示；

活動偵測器，其用於分析音訊信號以取決於音訊信號而判定第一訊框為作用訊框且第二訊框為非作用訊框；

音訊信號編碼器，其用於針對第一訊框為作用訊框而產生經編碼音訊信號，且針對第二訊框為非作用訊框而產生參數描述；以及

經編碼信號形成器，其用於藉由將用於第一訊框之第一聲場參數表示、用於第二訊框之第二聲場參數表示、用於第一訊框之經編碼音訊信號及用於第二訊框之參數描述組合在一起而構成經編碼音訊場景。

【0010】 該聲場參數產生器可經組配以產生第一聲場參數表示或第二聲場參數表示，使得第一聲場參數表示或第二聲場參數表示包含指示音訊信號相對於聽者位置之特性的參數。

【0011】 該第一聲場參數表示或該第二聲場參數表示可包含第一訊框中指示聲音相對於聽者位置之方向的一或多個方向參數，或第一訊框中指示相對於直接聲音之擴散聲音之一部分的一或多個擴散度參數，或第一訊框中指示直接聲音與擴散聲音之能量比的一或多個能量比參數，或第一訊框中之聲道間/環繞聲相干性參數。

【0012】 該聲場參數產生器可經組配以根據音訊信號之第一訊框或第二訊框判定多個個別聲源且針對每一聲源而判定參數描述。

【0013】 聲場產生器可經組配以將第一訊框或第二訊框分解成多個頻率區間，每一頻率區間表示個別聲源，且針對每一頻率區間而判定至少一個聲場參數，該聲場參數例示性地包含方向參數、到達方向參數、擴散度參數、能量比參數或表示由音訊信號之第一訊框表示之聲場相對於聽者位置之特性的任何參數。

【0014】 用於第一訊框及第二訊框之音訊信號可包含具有表示相對於聽者之聲場的多個分量之輸入格式，

其中聲場參數產生器經組配以例如使用多個分量之降混來計算用於第一訊框及第二訊框之一或多個傳送聲道，且分析輸入格式以判定與一或多個傳送聲道相關之第一參數表示，或

其中聲場參數產生器經組配以例如使用多個分量之降混來計算一或多個傳送聲道，且

其中活動偵測器經組配以分析自第二訊框中之音訊信號導出的一或多個傳送聲道。

【0015】 用於第一訊框或第二訊框之音訊信號可包含輸入格式，對於第一訊框及第二訊框中之每一訊框，該輸入格式具有與每一訊框相關聯之一或多個傳送聲道及元資料，

其中聲場參數產生器經組配以自第一訊框及第二訊框讀取元資料，且將第一訊框之元資料用作或處理為第一聲場參數表示且處理第二訊框之元資料以獲得第二聲場參數表示，其中獲得第二聲場參數表示之處理使得傳輸第二訊框之元資料所需的資訊單元之量相對於處理之前所需的量有所減少。

【0016】 該聲場參數產生器可經組配以處理第二訊框之元資料以減少元資

料中之資訊項目之數目或將元資料中之資訊項目再取樣至較低解析度，諸如時間解析度或頻率解析度，或相對於再量化之前的情形將第二訊框之元資料的資訊單元再量化至較粗略表示。

【0017】 該音訊信號編碼器可經組配以將用於非作用訊框之靜默資訊描述判定為參數描述，

其中靜默資訊描述例示性地包含用於第二訊框之諸如能量、功率或響度的振幅相關資訊及諸如頻譜塑形資訊之塑形資訊，或用於第二訊框之諸如能量、功率或響度之振幅相關資訊及用於第二訊框之線性預測寫碼 LPC 參數，或用於第二訊框之具有變化之相關聯頻率解析度的尺度參數，使得不同尺度參數係指具有不同寬度之頻帶。

【0018】 該音訊信號編碼器可經組配以使用時域或頻域編碼模式針對第一訊框而編碼音訊信號，經編碼音訊信號包含例如經編碼時域樣本、經編碼頻譜域樣本、經編碼 LPC 域樣本及自音訊信號之分量獲得或自一或多個傳送聲道獲得的旁側資訊，該一或多個傳送聲道係例如藉由降混操作自音訊信號之分量導出。

【0019】 該音訊信號可包含輸入格式，該輸入格式為第一階立體混響 (Ambisonics) 格式、高階立體混響格式、與諸如 5.1 或 7.1 或 7.1+4 之給定揚聲器設置或表示一或若干個不同音訊物件之一或多個音訊聲道相關聯的多聲道格式，該一或若干個不同音訊物件位於如由包括於相關聯元資料中之資訊所指示的空間中，或包含為元資料相關聯之空間音訊表示的輸入格式，

其中聲場參數產生器經組配用於判定第一聲場參數表示及第二聲場表示，使得參數表示相對於所界定聽者位置之聲場，或

其中音訊信號包含如由真實麥克風或虛擬麥克風拾取之麥克風信號或例如呈第一階立體混響格式或高階立體混響格式之以合成方式產生之麥克風信號。

【0020】該活動偵測器可經組配以用於偵測第二訊框及第二訊框之後的一個或多個訊框上之不活動階段，且

其中音訊信號編碼器經組配以僅針對另一第三訊框而產生用於非作用訊框之另一參數描述，就訊框之時間序列而言，該另一第三訊框與第二訊框相隔至少一個訊框，且

其中聲場參數產生器經組配以用於僅針對訊框而判定另一聲場參數表示，音訊信號編碼器已針對訊框判定參數描述，或

其中活動偵測器經組配以用於判定包含第二訊框及該第二訊框之後的八個訊框的非作用階段，且其中音訊信號編碼器經組配以用於僅在每第八個訊框處產生用於非作用訊框之參數描述，且其中聲場參數產生器經組配以用於針對每一第八個非作用訊框而產生聲場參數表示，或

其中聲場參數產生器經組配以用於針對每一非作用訊框而產生聲場參數表示，甚至在音訊信號編碼器不產生用於非作用訊框之參數描述時亦如此，或

其中聲場參數產生器經組配以用於判定相比音訊信號編碼器產生用於一或多個非作用訊框之參數描述而具有較高訊框率之參數表示。

【0021】該聲場參數產生器可經組配以用於使用頻帶中之一或多個方向的空間參數及對應於一個方向分量與總能量之比率的頻帶中之相關聯能量比來判定用於第二訊框之第二聲場參數表示，或

判定指示擴散聲音或直接聲音之比率的擴散度參數，或

使用與第一訊框中之量化相比較更粗略量化方案判定方向資訊，或

使用方向隨時間或頻率之平均值以獲得較粗略時間或頻率解析度，或

判定用於一或多個非作用訊框之聲場參數表示，該一或多個非作用訊框具有與用於作用訊框之第一聲場參數表示中相同的頻率解析度，且相對於用於非作用訊框之聲場參數表示中之方向資訊，具有低於用於作用訊框之時間發生率

的時間發生率，或

判定具有擴散度參數之第二聲場參數表示，其中該擴散度參數係以與作用訊框相同之時間或頻率解析度但以一較粗略量化傳輸，或

用第一數目個位元量化用於第二聲場表示之擴散度參數，且其中僅傳輸每一量化索引之第二數目個位元，位元之該第二數目小於位元之第一數目，或

若音訊信號具有對應於定位於空間域中之聲道的輸入聲道，則針對第二聲場參數表示而判定聲道間相干性，或若音訊信號具有對應於定位於空間域中之聲道的輸入聲道，則判定聲道間聲級差，或

判定環繞聲相干性，其經界定為由音訊信號表示之聲場中相干的擴散能量之比率。

【0022】 根據一態樣，提供一種用於處理經編碼音訊場景之設備，該經編碼音訊場景在第一訊框中包含第一聲場參數表示及經編碼音訊信號，其中第二訊框為非作用訊框，設備包含：

活動偵測器，其用於偵測第二訊框為非作用訊框；

合成信號合成器，其用於使用用於第二訊框之參數描述來合成用於第二訊框之合成音訊信號；

音訊解碼器，其用於解碼用於第一訊框之經編碼音訊信號；以及

空間呈現器，其用於使用第一聲場參數表示且使用用於第二訊框之合成音訊信號在空間上呈現用於第一訊框之音訊信號，或轉碼器，其用於產生元資料輔助輸出格式，該元資料輔助輸出格式包含用於第一訊框之音訊信號、用於第一訊框之第一聲場參數表示、用於第二訊框之合成音訊信號及用於第二訊框之第二聲場參數表示。

【0023】 該經編碼音訊場景可包含用於第二訊框之第二聲場參數描述，且其中設備包含用於自第二聲場參數表示導出一或多個聲場參數之聲場參數處理

器，且其中空間呈現器經組配以將用於第二訊框之一或多個聲場參數用於第二訊框之合成音訊信號之呈現。

【0024】 該設備可包含用於導出用於第二訊框之一或多個聲場參數的參數處理器，

其中參數處理器經組配以儲存用於第一訊框之聲場參數表示且使用用於第一訊框之所儲存之第一聲場參數表示來合成用於第二訊框之一或多個聲場參數，其中第二訊框在時間上在第一訊框之後，或

其中參數處理器經組配以儲存用於在時間上出現於第二訊框之前或在時間上出現於第二訊框之後的若干訊框之一或多個聲場參數表示，以使用用於若干訊框之一或多個聲場參數表示中的至少二個聲場參數表示進行外推或內插，以判定用於第二訊框之一或多個聲場參數，且

其中空間呈現器經組配將用於第二訊框之一或多個聲場參數用於第二訊框之合成音訊信號之呈現。

【0025】 該參數處理器可經組配以進行外推或內插以判定用於第二訊框之一或多個聲場參數時使用在時間上出現於第二訊框之前或之後的至少二個聲場參數表示中所包括的方向執行抖動。

【0026】 該經編碼音訊場景可包含用於第一訊框之一或多個傳送聲道，

其中合成信號產生器經組配以產生用於第二訊框之一或多個傳送聲道作為合成音訊信號，且

其中空間呈現器經組配以在空間上呈現用於第二訊框之一或多個傳送聲道。

【0027】 該合成信號產生器可經組配以針對第二訊框而產生用於與空間呈現器之音訊輸出格式相關的個別分量之多個合成分量音訊信號作為合成音訊信號。

【0028】 該合成信號產生器可經組配以至少針對與音訊輸出格式相關之至少二個個別分量之子集中的每一者而產生個別合成分量音訊信號，

其中第一個別合成分量音訊信號與第二個別合成分量音訊信號去相關，且其中空間呈現器經組配以使用第一個別合成分量音訊信號與第二個別合成分量音訊信號之組合來呈現音訊輸出格式之分量。

【0029】 該空間呈現器可經組配以應用協方差法。

【0030】 該空間呈現器可經組配以不使用任何去相關器處理或控制去相關器處理，使得僅使用藉由如由協方差法指示之去相關器處理產生的一定量之去相關信號來產生音訊輸出格式之分量。

【0031】 該合成信號產生器為舒適雜訊產生器。

【0032】 該合成信號產生器可包含雜訊產生器且第一個別合成分量音訊信號係由雜訊產生器之第一取樣產生，且第二個別合成分量音訊信號係由雜訊產生器之第二取樣產生，其中第二取樣不同於第一取樣。

【0033】 該雜訊產生器可包含雜訊表，且其中第一個別合成分量音訊信號係藉由獲取雜訊表之第一部分而產生，且其中第二個別合成分量音訊信號係藉由獲取雜訊表之第二部分而產生，其中雜訊表之第二部分不同於雜訊表之第一部分，或

其中雜訊產生器包含偽雜訊產生器，且其中第一個別合成分量音訊信號係藉由使用偽雜訊產生器之第一種子而產生，且其中第二個別合成分量音訊信號係使用偽雜訊產生器之第二種子而產生。

【0034】 該經編碼音訊場景可包含用於第一訊框之二個或更多個傳送聲道，且

其中合成信號產生器包含雜訊產生器且經組配以使用用於第二訊框之參數描述藉由對雜訊產生器進行取樣而產生第一傳送聲道及藉由對雜訊產生器進行

取樣而產生第二傳送聲道，其中如藉由對雜訊產生器進行取樣而判定之第一傳送聲道及第二傳送聲道係使用用於第二訊框之相同參數描述進行加權。

【0035】 該空間呈現器可經組配以

使用直接信號與由去相關器在第一聲場參數表示之控制下自直接信號產生之擴散信號的混合，在用於第一訊框之第一模式下操作，且

使用第一合成分量信號與第二合成分量信號之混合，在用於第二訊框之第二模式下操作，其中第一合成分量信號及第二合成分量信號係由合成信號合成器藉由雜訊處理或偽雜訊處理之不同實現而產生。

【0036】 該空間呈現器可經組配以根據藉由參數處理器針對第二訊框導出的擴散度參數、能量分佈參數或相干性參數而控制第二模式下之混合。

【0037】 該合成信號產生器可經組配以使用用於第二訊框之參數描述來產生用於第一訊框之合成音訊信號，且

其中空間呈現器經組配以在空間呈現之前或之後執行用於第一訊框之音訊信號與用於第一訊框之合成音訊信號的加權組合，其中在該加權組合中，用於第一訊框之合成音訊信號的強度相對於用於第二訊框之合成音訊信號的強度有所減小。

【0038】 參數處理器可經組配以針對第二非作用訊框而判定環繞聲相干性，該環繞聲相干性經界定為由第二訊框表示之聲場中相干的擴散能量之比率，其中空間呈現器經組配以用於基於聲音相干性在第二訊框中之直接信號與擴散信號之間重分佈能量，其中自待重分佈至方向分量之擴散能量移除聲音環繞相干分量之能量，且其中在再現空間中平移方向分量。

【0039】 該設備可包含輸出介面，該輸出介面用於將由空間呈現器產生之音訊輸出格式轉換成經轉碼輸出格式，諸如包含專用於待置放於預定位置處之揚聲器的數個輸出聲道的輸出格式，或包含 FOA 或 HOA 資料之經轉碼輸出格

式，或

其中，替代空間呈現器，提供轉碼器以用於產生元資料輔助輸出格式，該元資料輔助輸出格式包含用於第一訊框之音訊信號、用於第一訊框之第一聲場參數及用於第二訊框之合成音訊信號及用於第二訊框之第二聲場參數表示。

【0040】 該活動偵測器可經組配以用於偵測第二訊框為非作用訊框。

【0041】 根據一態樣，提供一種自具有第一訊框及第二訊框之音訊信號產生經編碼音訊場景的方法，其包含：

根據第一訊框中之音訊信號判定用於第一訊框之第一聲場參數表示，且根據第二訊框中之音訊信號判定用於第二訊框之第二聲場參數表示；

分析音訊信號以取決於音訊信號而判定第一訊框為作用訊框且第二訊框為非作用訊框；

針對為作用訊框之第一訊框而產生經編碼音訊信號且針對為非作用訊框之第二訊框而產生參數描述；以及

藉由將用於第一訊框之第一聲場參數表示、用於第二訊框之第二聲場參數表示、用於第一訊框之經編碼音訊信號及用於第二訊框之參數描述組合在一起而構成經編碼音訊場景。

【0042】 根據一態樣，提供一種處理經編碼音訊場景之方法，該經編碼音訊場景在第一訊框中包含第一聲場參數表示及經編碼音訊信號，其中第二訊框為非作用訊框，方法包含：

偵測第二訊框為非作用訊框且提供用於第二訊框之參數描述；

使用用於第二訊框之參數描述來合成用於第二訊框之合成音訊信號；

解碼用於第一訊框之經編碼音訊信號；以及

使用第一聲場參數表示且使用用於第二訊框之合成音訊信號在空間上呈現用於第一訊框之音訊信號，或產生元資料輔助輸出格式，該元資料輔助輸出格

式包含用於第一訊框之音訊信號、用於第一訊框之第一聲場參數表示、用於第二訊框之合成音訊信號及用於第二訊框之第二聲場參數表示。

【0043】 該方法可包含提供該用於第二訊框之參數描述。

【0044】 根據一態樣，提供一種經編碼音訊場景，其包含：

用於第一訊框之第一聲場參數表示；

用於第二訊框之第二聲場參數表示；

用於第一訊框之經編碼音訊信號；以及

用於第二訊框之參數描述。

【0045】 根據一態樣，提供一種電腦程式，其用於在電腦或處理器上運行時執行以上或以下之方法。

【圖式簡單說明】

【0046】 圖 1 (其劃分為圖 1a 及圖 1b) 展示根據先前技術之可用於根據實例分析及合成的實例。

【0047】 圖 2 展示根據實例之解碼器及編碼器的實例。

【0048】 圖 3 展示根據實例之編碼器的實例。

【0049】 圖 4 及 5 展示組件之實例。

【0050】 圖 5 展示根據實施之組件的實例。

【0051】 圖 6 至 11 展示解碼器之實例。

【實施方式】

較佳實施例之詳細說明

【0052】 首先，提供已知範式(DTX、DirAC、MASA 等)之一些論述，其中一些技術之描述可至少在一些情況下實施於本發明之實例中。

DTX

【0053】 舒適雜訊產生器通常用於語音之不連續傳輸(DTX)。在此類模式

中，語音首先由話音活動偵測器(VAD)分類於活動及非作用訊框中。VAD 之實例可見於[2]中。基於 VAD 結果，僅以標稱位元速率寫碼及傳輸活動語音訊框。在僅存在背景雜訊之長停頓期間，位元速率降低或調零，且背景雜訊以章節及參數方式寫碼。隨後顯著降低平均位元速率。雜訊係在解碼器側處之非作用訊框期間由舒適雜訊產生器(CNG)產生。舉例而言，語音寫碼器 AMR-WB [2]及 3GPP EVS [3,4]二者有可能在 DTX 模式中運行。高效 CNG 之實例在[5]中給出。

【0054】 本發明之實施例以一方式擴展此原理，該方式使得其將相同原理應用於具有聲音事件之空間定位的沉浸式會話語音。

DirAC

【0055】 DirAC 係在感知上促動空間聲音之重現。假定在一個瞬間且對於一個臨界頻帶，聽覺系統之空間解析度限於解碼一個方向提示及另一耳間相干性提示。

【0056】 基於此等假定，DirAC 藉由使二個串流：非方向性擴散串流及方向性不擴散串流交叉漸進而表示一個頻帶中的空間聲音。DirAC 處理係在二個階段中執行：如圖 1 (圖 1a 展示合成，圖 1b 展示分析)中所描繪之分析及合成。

【0057】 在 DirAC 分析階段中，呈 B 格式之第一階重合麥克風被視為輸入且在頻域中分析擴散度及聲音之到達方向。

【0058】 在 DirAC 合成階段中，聲音被分成二個串流，不擴散串流及擴散串流。使用振幅平移將不擴散串流再現為點源，振幅平移可藉由使用向量基礎振幅平移(VBAP)[6]進行。擴散串流大體負責包封之感覺且藉由將彼此去相關信號傳送至揚聲器而產生。

【0059】 在下文中亦稱為空間元資料或 DirAC 元資料之 DirAC 參數由擴散度及方向之元組組成。方向可藉由二個角，方位角及仰角以球面座標表示，而擴散度可為介於 0 與 1 之間的標量因數。

【0060】 已進行一些工作以用於減小元資料之大小，以使 DirAC 範式能夠用於空間音訊寫碼及電話會議情境中[8]。

【0061】 據本發明人瞭解，未曾圍繞參數空間音訊寫碼解碼器建構或提議且甚至很少基於 DirAC 範式建構或提議 DTX 系統。此為本發明之實施例之主題。

MASA

【0062】 元資料輔助空間音訊(MASA)係自 DirAC 原理導出之空間音訊格式，該 DirAC 原理可直接根據原始麥克風信號計算且傳送至音訊寫碼解碼器而無需經過如立體混響之中間格式。可由例如頻帶中之方向參數及/或例如頻帶中之能量比參數(例如，指示定向聲音能量之比例)組成之參數集亦可用作音訊寫碼解碼器或呈現器之空間元資料。此等參數可根據麥克風陣列捕捉之音訊信號估計；舉例而言，單聲道或立體聲信號可自麥克風陣列信號產生以與空間元資料一起傳送。單聲道或立體聲信號可例如藉由類似於 3GPPEVS 之核心寫碼器或其衍生物進行編碼。解碼器可將音訊信號解碼至頻帶中之聲音中且對其進行處理(使用所傳輸空間元資料)以獲得空間輸出，該空間輸出可為雙耳輸出、揚聲器多聲道信號或呈立體混響格式之多聲道信號。

動機

【0063】 沉浸式語音通信係一個新的研究領域且極少系統存在，此外無設計用於此類應用之 DTX 系統。

【0064】 然而，可簡單地組合現有解決方案。可例如對每一個別多聲道信號獨立地應用 DTX。此極簡方法面臨若干問題。為此，需要分離地傳輸與低位元速率通信約束不相容且因此幾乎不與 DTX 相容之每一個別聲道，DTX 經設計以用於低位元速率通信情況。此外，隨後需要使跨越聲道之 VAD 決策同步以避免不尋常事件及未遮蔽效應，且亦需要充分利用 DTX 系統之位元速率降低。實際上，為中斷傳輸及從中獲利，需要確保所有聲道上之話音活動決策同步。

【0065】 當藉由一或多個舒適雜訊產生器在非作用訊框期間產生遺失背景雜訊時，另一問題出現在接收器側上。對於沉浸式通信，尤其當直接將 DTX 應用於個別聲道時，每一聲道需要一個產生器。若通常對隨機雜訊取樣之此等產生器被獨立地使用，則聲道之間的相干性將為零或接近零，且可在感知上偏離原始聲音景觀。另一方面，若僅使用一個產生器且將所得舒適雜訊複製至所有輸出聲道，則相干性將極高且沉浸感將大幅度降低。

【0066】 此等問題可藉由以下操作解決：不將 DTX 直接應用於系統之輸入或輸出聲道，而是替代地在如 DirAC 之參數空間音訊寫碼方案之後將 DTX 應用於所得傳送聲道上，該等傳送聲道通常為原始多聲道信號之降混或減少版本。在此情況下，有必要界定如何將非作用訊框參數化且接著藉由 DTX 系統空間化。此並非無足輕重的且為本發明之實施例的主題。空間影像必須在活動與非作用訊框之間一致，且必須在感知上儘可能忠實於原始背景雜訊。

【0067】 圖 3 展示根據實例之編碼器 300。編碼器 300 可根據音訊信號 302 產生經編碼音訊場景 304。

【0068】 音訊信號 304 (位元串流)或音訊場景 304 (以及下文所揭示之其他音訊信號)可被劃分成訊框(例如，其可為訊框序列)。訊框可與時槽相關聯，該等時槽可隨後彼此界定(在一些實例中，先前態樣可與後續訊框重疊)。對於每一訊框，時域(TD)或頻域(FD)中之值可寫入位元串流 304 中。在 TD 中，可針對每一樣本(具有例如離散樣本序列之每一訊框)提供值。在 FD 中，可針對每一頻率區間提供值。如稍後將解釋，可將每一訊框分類(例如，藉由活動偵測器)為作用訊框 306 (例如，非空訊框)或非作用訊框 308 (例如，空訊框，或靜默訊框，或僅雜訊訊框)。亦可結合作用訊框 306 及非作用訊框 308 提供不同參數(例如，作用空間參數 316 或非作用空間參數 318) (在無資料之情況下，元件符號 319 展示未提供資料)。

【0069】音訊信號 302 可為例如多聲道音訊信號(例如，具有二個聲道或更多)。音訊信號 302 可為例如立體聲音訊信號。音訊信號 302 可例如為例如呈 A 格式或 B 格式之立體混響信號。音訊信號 302 可具有例如元資料輔助空間音訊 (MASA)格式。音訊信號 302 可具有輸入格式，該輸入格式為第一階立體混響格式、高階立體混響格式、與諸如 5.1 或 7.1 或 7.1+4 之給定揚聲器設置或表示一或若干個不同音訊物件之一或多個音訊聲道相關聯的多聲道格式，該一或若干個不同音訊物件位於如由包括於相關聯元資料中之資訊所指示的空間中，或具有為元資料相關聯之空間音訊表示的輸入格式。音訊信號 302 可包含如由真實麥克風或虛擬麥克風拾取之麥克風信號。音訊信號 302 可包含以合成方式產生之麥克風信號(例如，呈第一階立體混響格式或高階立體混響格式)。

【0070】音訊場景 304 可包含以下各者中之至少一者或組合：

用於第一訊框 306 之第一聲場參數表示(例如，作用空間參數) 316；

用於第二訊框 308 之第二聲場參數表示(例如，非作用空間參數) 318；

用於第一訊框 306 之經編碼音訊信號 346；以及

用於第二訊框 308 之參數描述 348(在一些實例中，非作用空間參數 318 可包括於參數描述 348 中，但參數描述 348 亦可包括並非空間參數之其他參數)。

【0071】作用訊框 306 (第一訊框)可為含有語音(或在一些實例中，亦不同於純雜訊之其他音訊聲音)之彼等訊框。非作用訊框 308 (第二訊框)可理解為不包含語音(或在一些實例中，亦不同於純雜訊之其他音訊聲音)且可理解為含有獨特雜訊的彼等訊框。

【0072】可提供音訊場景分析器(聲場參數產生器) 310 例如以產生音訊信號 302 之傳送聲道版本 324 (在 326 及 328 當中細分)。此處，吾等可指每一第一訊框 306 之一或多個傳送聲道 326 及/或每一第二訊框 308 之一或多個傳送聲道 328 (一或多個傳送聲道 328 可理解為提供例如靜默或雜訊之參數描述)。一或多

個傳送聲道 324 (326、328)可為輸入格式 302 之降混版本。一般而言，若輸入音訊信號 302 為立體聲聲道，則傳送聲道 326、328 中之每一者可為例如一個單聲道。若輸入音訊信號 302 具有二個以上聲道，則輸入音訊信號 302 之降混版本 324 之聲道可少於輸入音訊信號 302，但在一些實例中，仍具有一個以上聲道(例如，若輸入音訊信號 302 具有四個聲道，則降混版本 324 可具有一個、二個或三個聲道)。

【0073】 音訊信號分析器 310 可另外或替代性地提供用 314 指示之聲場參數(空間參數)。特定言之，聲場參數 314 可包括與第一訊框 306 相關聯之作用空間參數(第一空間參數或第一空間參數表示) 316，及與第二訊框 308 相關聯之非作用空間參數(第二空間參數或第二空間參數表示) 318。每一作用空間參數 314 (316、318)可包含(例如，為)指示音訊信號(302)相對於聽者位置之空間特性的參數。在一些其他實例中，作用空間參數 314 (316、318)可包含(例如，為)至少部分指示音訊信號 302 相對於揚聲器位置之特性的參數。在一些實例中，作用空間參數 314 (316、318)可包含(例如，為)可至少部分地包含如自信號源獲取之音訊信號之特性。

【0074】 舉例而言，空間參數 314 (316、318)可包括擴散度參數：例如指示相對於第一訊框 306 及/或第二訊框 308 中之聲音的擴散信號比之一或多個擴散度參數，或指示第一訊框 306 及/或第二訊框 308 中之直接聲音與擴散聲音之能量比之一或多個能量比參數，或第一訊框 306 及/或第二訊框 308 中之聲道間/環繞聲相干性參數，或第一訊框 306 及/或第二訊框 308 中之一或多個相干擴散功率比，或第一訊框 306 及/或第二訊框 308 中之一或多個信號擴散比。

【0075】 在實例中，一或多個作用空間參數(第一聲場參數表示) 316 及/或一或多個非作用空間參數 318 (第二聲場參數表示)可在其完整聲道版本或其子集，如高階立體混響輸入信號之第一階分量中自輸入信號 302 獲得。

【0076】設備 300 可包括活動偵測器 320。活動偵測器 320 可分析輸入音訊信號(在其輸入版本 302 中或在其降混版本 324 中)，以取決於音訊信號(302 或 324)而判定訊框為作用訊框 306 抑或為非作用訊框 308，從而對訊框執行分類。如自圖 3 可見，可將活動偵測器 320 假定為控制(例如，經由控制件 321)第一偏差器 322 及第二偏差器 322a。第一偏差器 322 可在作用空間參數 316 (第一聲場參數表示)與非作用空間參數 318 (第二聲場參數表示)之間進行選擇。因此，活動偵測器 320 可決定是否輸出作用空間參數 316 或非作用空間參數 318 (例如，在位元串流 304 中發信)。同一控制件 321 可控制第二偏差器 322a，該第二偏移器可在輸出傳送聲道 324 中之第一訊框 326 (306)或傳送聲道 326 中之第二訊框 328 (308)(例如，參數描述)之間進行選擇。第一偏差器 322 及第二偏差器 322a 之活動彼此協調：當輸出作用空間參數 316 時，隨後亦輸出第一訊框 306 之傳送聲道 326，且當輸出非作用空間參數 318 時，隨後輸出第一訊框 306 傳送聲道之傳送聲道 328。此係因為作用空間參數 316 (第一聲場參數表示)描述第一訊框 306 之空間特性，而非作用空間參數 318 (第二聲場參數表示)描述第二訊框 308 之空間特性。

【0077】活動偵測器 320 可因此基本上決定輸出第一訊框 306 (326、346)及其相關參數(316)及第二訊框 308 (328、348)及其相關參數(318)當中的哪一者。活動偵測器 320 亦可控制在位元串流中對一些發信之編碼，其發訊框為作用抑或非作用的(可使用其他技術)。

【0078】活動偵測器 320 可對輸入音訊信號 302 之每一訊框 306/308 執行處理(例如，藉由量測訊框中，例如音訊信號之特定訊框之全部或至少多個頻率區間中之能量)，且可將特定訊框分類為第一訊框 306 或第二訊框 308。一般而言，活動偵測器 320 可決定一個單一完整訊框之一個單一分類結果，而不區分同一訊框之不同頻率區間與不同樣本。舉例而言，一個分類結果可為「語音」(其

將相當於由作用空間參數 316 在空間上描述之第一訊框 306、326、346)或「靜默」(其將相當於由非作用空間參數 318 在空間上描述之第二訊框 308、328、348)。因此，根據由活動偵測器 320 施加之分類，偏差器 322 及 322a 可執行其交換，且其結果原則上對於經分類訊框之所有頻率區間(及樣本)有效。

【0079】 設備 300 可包括音訊信號編碼器 330。音訊信號編碼器 330 可產生經編碼音訊信號 344。詳言之，音訊信號編碼器 330 可針對第一訊框(306、326)提供例如由傳送聲道編碼器 340 產生的經編碼音訊信號 346，該傳送聲道編碼器可為音訊信號編碼器 330 之部分。經編碼音訊信號 344 可為或包括靜默之參數描述 348 (例如，雜訊之參數描述)，且可由可為音訊信號編碼器 330 之部分的傳送聲道 SI 描述符 350 產生。所產生之第二訊框 348 可對應於原始音訊輸入信號 302 之至少一個第二訊框 308 及對應於降混信號 324 之至少一個第二訊框 328，且可由非作用空間參數 318 (第二聲場參數表示)在空間上描述。值得注意的是，經編碼音訊信號 344 (無論 346 或 348)亦可在傳送聲道中(且可因此為降混信號 324)。經編碼音訊信號 344 (無論 346 或 348)可經壓縮，以便減小其大小。

【0080】 設備 300 可包括經編碼信號形成器 370。經編碼信號形成器 370 可寫入至少經編碼音訊場景 304 之經編碼版本。經編碼信號形成器 370 可藉由將用於第一訊框 306 之第一(作用)聲場參數表示 316、用於第二訊框 308 之第二(非作用)聲場參數表示 318、用於第一訊框 306 之經編碼音訊信號 346 及用於第二訊框 308 之參數描述 348 組合在一起而進行操作。因此，音訊場景 304 可為位元串流，其可經傳輸或儲存(或二者)且由通用解碼器使用以用於產生待輸出之音訊信號，該音訊信號為原始輸入信號 302 之複本。在音訊場景(位元串流) 304 中，可因此獲得「第一訊框」/「第二訊框」之序列，以允許輸入信號 306 之再現。

【0081】 圖 2 展示編碼器 300 及解碼器 200 之實例。在一些實例中，編碼器 300 可與圖 3 之編碼器相同(或為其變體) (在一些其他實例中，其可為不同實

施例)。編碼器 300 可輸入有音訊信號 302(其可例如呈 B 格式)且可具有第一訊框 306(其可為例如作用訊框)及第二訊框 308(其可為例如非作用訊框)。音訊信號 302 可在選擇器 320 (其可包括與偏差器 322 及 322a 相關聯之音訊)中內部的選擇之後作為信號 324 (例如，作為用於第一訊框之經編碼音訊信號 326，及用於第二訊框之經編碼音訊信號 328 或參數表示)提供至音訊信號編碼器 330。值得注意的是，區塊 320 亦可具有將來自輸入信號 302 (306、308)之降混形成至傳送聲道 324 (326、328)上的能力。基本上，區塊 320 (波束成形/信號選擇區塊)可理解為包括圖 3 之活動偵測器 320 之功能，但圖 3 中由區塊 310 執行之一些其他功能(諸如產生空間參數 316 及 318)可由圖 2 之「DirAC 分析區塊」310 執行。因此，聲道信號 324 (326、328)可為原始信號 302 之降混版本。然而，在一些情況下，以下情況亦可為可能的：不對信號 302 執行降混，且信號 324 僅為第一訊框與第二訊框之間的選擇。音訊信號編碼器 330 可包括區塊 340 及 350 中之至少一者，如上文所解釋。音訊信號編碼器 330 之輸出端可針對第一訊框 346 或針對第二訊框 348 輸出編碼器音訊信號 344。圖 2 並不展示經編碼信號形成器 370，儘管其可存在。

【0082】 如所展示，區塊 310 可包括 DirAC 分析區塊(或更一般而言，聲場參數產生器 310)。區塊 310 (聲場參數產生器)可包括濾波器組分析 390。濾波器組分析 390 可將輸入信號 302 之每一訊框細分為多個頻率區間，該等頻率區間可為濾波器組分析 390 之輸出 391。擴散度估計區塊 392a 可例如為由濾波器組分析 390 輸出的多個頻率區間 391 中之每一頻率區間提供擴散度參數 314a (其可為用於作用訊框 306 之一或多個作用空間參數 316 的一個擴散度參數或用於非作用訊框 308 之一或多個非作用空間參數 318 的一個擴散度參數)。聲場參數產生器 310 可包括方向估計區塊 392b，該方向估計區塊之輸出 314b 可為例如用於由濾波器組分析 390 輸出之多個頻率區間 391 中之每一頻率區間的方向參數(其可

為用於作用訊框 306 之一或多個作用空間參數 316 之一個方向參數或用於非作用訊框 308 之一或多個非作用空間參數 318 之一個方向參數)。

【0083】圖 4 展示區塊 310 (聲場參數產生器)之實例。聲場參數產生器 310 可與圖 2 之聲場參數產生器相同及/或可與圖 3 之聲場參數產生器相同或至少實施區塊 310 之功能，不管圖 3 之區塊 310 亦能夠執行輸入信號 302 之降混之事實，同時此事實並未展示(或未實施)於圖 4 之聲場參數產生器 310 中。

【0084】圖 4 之聲場參數產生器 310 可包括濾波器組分析區塊 390 (其可與圖 2 之濾波器組分析區塊 390 相同)。濾波器組分析區塊 390 可為每一訊框及每一波束(頻率塊)提供頻域資訊 391。頻域資訊 391 可提供至可為圖 3 中展示之彼等者的擴散度分析區塊 392a 及/或方向分析區塊 392b。擴散度分析區塊 392a 及/或方向分析區塊 392b 可提供擴散度資訊 314a 及/或方向資訊 314b。此等資訊可經提供以用於每一第一訊框 306 (346)及用於每一第二訊框 308 (348)。綜合地，由區塊 392a 及 392b 提供之資訊被視為聲場參數 314，該等聲場參數包含第一聲場參數 316 (作用空間參數)及第二聲場參數 318 (非作用空間參數)二者。可將作用空間參數 316 提供至作用空間元資料編碼器 396，且可將非作用空間參數 318 提供至非作用空間元資料編碼器 398。所得為可經編碼於位元串流 304 中(例如，經由編碼器信號形成器 370)且經儲存以供隨後由解碼器播放的第一聲場參數表示及第二聲場參數表示(316、318，用 314 綜合指示)。無論作用空間元資料編碼器 396 或非作用空間參數 318 將編碼訊框，此可藉由諸如圖 3 中之控制件 321 的控制(偏差器 322 未展示於圖 2 中)來控制，例如由活動偵測器操作之徹底分類。(應注意，在一些實例中，編碼器 396、398 亦可執行量化)。

【0085】圖 5 展示可能聲場參數產生器 310 之另一實例，其可替代圖 4 之聲場參數產生器且其亦可實施於圖 2 及圖 3 之實例中。在此實例中，輸入音訊信號 302 可能已呈 MASA 格式，其中空間參數已為例如用於多個頻率區間中之每

一頻率區間之輸入音訊信號 302 之部分(例如，作為空間元資料)。因此，無需具有擴散度分析區塊及/或方向區塊，而是其可由 MASA 讀取器 390M 取代。MASA 讀取器 390M 可讀取音訊信號 302 中之特定資料欄位，該欄位已含有諸如一或多個作用空間參數 316 及一或多個非作用空間參數 318 之資訊(根據信號 302 之訊框為第一訊框 306 抑或第二訊框 308 之事實)。可編碼於信號 302 中之參數(且其可由 MASA 讀取器 390M 讀取)之實例可包括方向、能量比、環繞聲相干性、散佈相干性等中之至少一者。在 MASA 讀取器 390M 之下游，作用空間元資料編碼器 396 (例如，如圖 4 中之一者)及非作用空間元資料編碼器 398 (例如，如圖 4 中之一者)可經提供以分別輸出第一聲場參數表示 316 及第二聲場參數表示 318。若輸入音訊信號 302 為 MASA 信號，則活動偵測器 320 可實施為讀取輸入 MASA 信號 302 中之經判定資料欄位且基於資料欄位中經編碼之值而分類為作用訊框 306 或非作用訊框 308 的元件。圖 5 之實例可針對已編碼於空間資訊中之音訊信號 302 而一般化，該空間資訊可經編碼為作用空間參數 316 或非作用空間參數 318。

【0086】 本發明之實施例應用於例如圖 2 中所示之空間音訊寫碼系統，其中描繪基於 DirAC 之空間音訊編碼器及解碼器。其論述如下。

【0087】 編碼器 300 可通常分析呈 B 格式之空間音訊場景。替代地，DirAC 分析可經調整以分析不同音訊格式，如音訊物件或多聲道信號或任何空間音訊格式之組合。

【0088】 DirAC 分析(例如如在階段 392a、392b 中之任一者執行)可自輸入音訊場景 302 (輸入信號)提取參數表示 304。每個時間頻率單位量測之到達方向(DOA) 314b 及/或擴散度 314a 形成一或多個參數 316、318。DirAC 分析(例如如在階段 392a、392b 中之任一者執行)後可接著空間元資料編碼器(例如，396 及/或 398)，該空間元資料編碼器可量化及/或編碼 DirAC 參數以獲得低位元速率參

數表示(在諸圖中，低位元速率參數表示 316、318 係以空間元資料編碼器 396 及/或 398 上游之參數表示的相同元件符號指示)。

【0089】 連同參數 316 及/或 318 一起，可藉由習知音訊核心編碼器寫碼自一或多個不同源(例如，不同麥克風)或一或多個音訊輸入信號(例如，多聲道信號之不同分量) 302 導出之降混信號 324 (326) (例如，以供傳輸及/或以供儲存)。在較佳實施例中，EVS 音訊寫碼器(例如 330，圖 2)可較佳地用於寫碼降混信號 324 (326、328)，但本發明之實施例不限於此核心編碼器且可應用於任何音訊核心編碼器。降混信號 324 (326、328)可由例如亦被稱作傳送聲道之不同聲道組成：信號 324 可取決於目標位元速率而為例如或包含構成 B 格式信號、立體聲對或單音降混的四個係數信號。經寫碼空間參數 328 及經寫碼音訊位元串流 326 可在經由通信聲道傳輸(或儲存)之前進行多工。

【0090】 在解碼器(參見下文)中，傳送聲道 344 係由核心解碼器解碼，而 DirAC 元資料(例如，空間參數 316、318)可在與經解碼傳送聲道一起傳送至 DirAC 合成之前先經解碼。DirAC 合成使用經解碼元資料以用於控制直接聲音串流之再現及其與擴散聲音串流之混合。再現聲場可再現於任意揚聲器佈局上或可以任意次序以立體混響格式(HOA/FOA)產生。

DirAC 參數估計

【0091】 此處解釋用於估計空間參數 316、318 (例如，擴散度 314a、方向 314b)之非限制性技術。提供 B 格式之實例。

【0092】 在每一頻帶中(例如，如自濾波器組分析 390 獲得)，可估計聲音之到達方向 314a 連同聲音之擴散度 314b。根據輸入 B 格式分量 $w^i(n), x^i(n), y^i(n), z^i(n)$ 之時間頻率分析，壓力及速度向量可判定為：

$$P^i(n, k) = W^i(n, k)$$

$$U^i(n, k) = X^i(n, k)\mathbf{e}_x + Y^i(n, k)\mathbf{e}_y + Z^i(n, k)\mathbf{e}_z$$

其中 i 係輸入 302 之指數，且 k 及 n 係時間頻率塊之時間及頻率索引，且 e_x, e_y, e_z 表示笛卡兒單元向量。在一些實例中，可需要 $P(n, k)$ 及 $U(n, k)$ 以經由例如強度向量之計算來計算 DirAC 參數(316、318)，即 DOA 314a 及擴散度 314a：

$$I(k, n) = \frac{1}{2} \Re\{P(k, n) \cdot \overline{U(n, k)}\},$$

其中 $\overline{(\cdot)}$ 指示複共軛。組合式聲場之擴散度係藉由下式給出：

$$\psi(k, n) = 1 - \frac{\|E\{I(k, n)\}\|}{cE\{E(k, n)\}}$$

其中 $E\{\cdot\}$ 表示時間平均算子， c 表示聲音之速度且聲場能量 $E(k, n)$ 由下式給出：

$$E(n, k) = \frac{\rho_0}{4} \|U(n, k)\|^2 + \frac{1}{\rho_0 c^2} |P(n, k)|^2$$

【0093】 聲場之擴散度經界定為聲音強度與能量密度之間的比率，該比率介於 0 與 1 之間。

【0094】 到達方向(DOA)係藉助於單位向量表達，該單位向量方向 (n, k) 經界定為：

$$\text{方向}(n, k) = -\frac{I(n, k)}{\|I(n, k)\|}$$

【0095】 到達方向 314b 可由 B 格式輸入信號 302 之能量分析(例如，在 392b 處)判定且可經界定為強度向量之相對方向。方向經界定於笛卡兒座標中但可例如容易地在由單位半徑、方位角及仰角界定之球面座標中變換。

【0096】 就傳輸而言，參數 314a、314b (316、318)需要經由位元串流(例如，304)傳輸至接收器側(例如，解碼器側)。對於經由負載量有限之網路的更穩固傳輸，低位元速率位元串流係較佳或甚至必要的，其可藉由設計 DirAC 參數 314a、314b (316、318)之高效寫碼方案來達成。其可藉由對不同頻帶及/或時間單位上

的參數求平均值而使用例如諸如頻帶分組、預測、量化及熵寫碼之技術。在解碼器處，在網路中未出現錯誤之情況下，可針對每一時間/頻率單位(k,n)解碼所傳輸參數。然而，若網路條件並不足夠好以保證恰當封包傳輸，則封包可能在傳輸期間丟失。本發明之實施例旨在提供對後一情況的解決方案。

解碼器

【0097】 圖 6 展示解碼器設備 200 之實例。該解碼器設備可為用於處理經編碼音訊場景(304)之設備，該經編碼音訊場景在第一訊框(346)中包含第一聲場參數表示(316)及經編碼音訊信號(346)，其中第二訊框(348)為非作用訊框。解碼器設備 200 可包含以下各者中之至少一者：

活動偵測器(2200)，其用於偵測第二訊框(348)為非作用訊框且用於提供用於第二訊框(308)之參數描述(328)；

合成信號合成器(210)，其用於使用用於第二訊框(308)之參數描述(348)合成用於第二訊框(308)之合成音訊信號(228)；

音訊解碼器(230)，其用於解碼用於第一訊框(306)之經編碼音訊信號(346)；以及

空間呈現器(240)，其用於使用第一聲場參數表示(316)且使用用於第二訊框(308)之合成音訊信號(228)在空間上呈現用於第一訊框(306)之音訊信號(202)。

【0098】 值得注意的是，活動偵測器(2200)可發出命令 221'，該命令可判定輸入訊框被分類為作用訊框 346 抑或非作用訊框 348。活動偵測器 2200 可例如根據資訊 221 判定輸入訊框之分類，該資訊係發信抑或根據所獲得訊框之長度判定。

【0099】 合成信號合成器(210)可例如使用自參數表示 348 獲得之資訊(例如，參數資訊)例如產生雜訊 228。空間呈現器 220 可以一方式產生輸出信號 202，該方式經由非作用空間參數 318 處理非作用訊框 228 (自經編碼訊框 348 獲得)，

以獲得人類聽者具有雜訊之來源的 3D 空間印象。

【0100】 應注意，在圖 6 中，標號 314、316、318、344、346、348 與圖 3 之標號相同，此係因為其由於係自位元串流 304 獲得而對應。儘管如此，可存在一些輕微差異(例如，歸因於量化)。

【0101】 圖 6 亦展示控制 221'，其可控制偏差器 224'，使得可例如經由由活動偵測器 220 操作之分類而選擇信號 226 (由合成信號合成器 210 輸出)或音訊信號 228 (由音訊解碼器 230 輸出)。值得注意的是，信號 224 (226 或 228)仍可為降混信號，其可提供至空間呈現器 220 以使得空間呈現器經由作用非作用空間參數 314 (316、318)產生輸出信號 202。在一些實例中，信號 224 (226 或 228)仍可升混以使得信號 224 之聲道的數目相對於經編碼版本 344 (346、348)增大。在一些實例中，儘管經升混，但信號 224 之聲道之數目可小於輸出信號 202 之聲道之數目。

【0102】 在下文中，提供解碼器設備 200 之其他實例。圖 7 至圖 10 展示可體現解碼器設備 200 之解碼器設備 700、800、900、1000 的實例。

【0103】 即使在圖 7 至圖 10 中一些元件展示為在空間呈現器 220 內部，但其在一些實例中仍可處於空間呈現器 220 外部。舉例而言，合成器 210 可部分抑或完全在空間呈現器 220 外部。

【0104】 在彼等實例中，可包括參數處理器 275 (其可在空間呈現器 220 內部或外部)。儘管未展示，參數處理器 275 亦可被視為存在於圖 6 之解碼器中。

【0105】 圖 7 至圖 10 中之任一者的參數處理器 275 可包括例如用於提供非作用訊框之非作用空間參數解碼器 278，該等非作用訊框可為英特爾參數 318 (例如，如自位元串流 304 中之發信獲得)及/或區塊 279 (「恢復未經傳輸之訊框解碼器中之空間參數」)，該區塊提供未在位元串流 304 中讀取但例如藉由外推獲得(例如，恢復、重建構、外推、推斷等)或以合成方式產生之非作用空間參數。

【0106】 因此，第二聲場參數表示亦可為所產生之參數 219，該參數不存在於位元串流 304 中。如稍後將解釋，經恢復(經重建構、經外推、經推斷等)之空間參數 219 可例如經由「維持策略」至「方向策略之外推」及/或經由「方向之抖動」而獲得(參見下文)。因此，參數處理器 275 可外推或以任何方式自先前訊框獲得空間參數 219。如圖 6 至圖 9 中可見，切換器 275'可在如在位元串流 304 中發信之非作用空間參數 318 與經恢復空間參數 219 之間選擇。如上文所解釋，靜默訊框 348 (SID)之編碼(以及非作用空間參數 318 之編碼)以比第一訊框 346 之編碼低的位元速率更新：非作用空間參數 318 相對於作用空間參數 316 以較低頻率更新，且一些策略係由參數處理器 275 (1075)執行以用於恢復用於非傳輸非作用訊框之非發信空間參數 219。因此，切換器 275'可在經發信非作用空間參數 318 與非經發信(但恢復或以其他方式重建構)非作用空間參數 219 之間選擇。在一些情況下，參數處理器 275'可針對在第二訊框之前出現或在時間上在第二訊框之後出現的若干訊框儲存一或多個聲場參數表示 318，以外推(或內插)用於第二訊框之聲場參數 219。一般而言，空間呈現器 220 可將用於第二訊框 219 之一或多個聲場參數 318 用於第二訊框 308 之合成音訊信號 202 之呈現。另外或替代地，參數處理器 275 可儲存用於作用空間參數之聲場參數表示 316 (圖 10 中所示)且使用所儲存之第一聲場參數表示 316 (作用訊框)合成用於第二訊框(非作用訊框)之聲場參數 219 以產生經恢復之空間參數 319。如圖 10 中所示(且亦可實施於圖 6 至圖 9 中之任一者中)，亦有可能亦包括作用空間參數解碼器 276，作用空間參數 316 可藉以該作用空間參數解碼器自位元串流 304 獲得。此可在外推或內插時執行抖動，其中方向包括於在時間上在第二訊框(308)之前或之後出現的至少二個聲場參數表示中，以判定用於第二訊框(308)之一或多個聲場參數。

【0107】 合成信號合成器 210 可在空間呈現器 220 內部，或可在其外部，

或在一些情況下，該合成信號合成器可具有內部部分及外部部分。合成合成器 210 可對傳送聲道 228 之降混聲道(其小於輸出聲道)進行操作(此處應注意， M 為降混聲道之數目且 N 為輸出聲道之數目)。合成信號產生器 210 (合成信號合成器之其他名稱)可針對第二訊框產生用於與空間呈現器之外格式相關的個別分量之多個合成分量音訊信號(在傳送信號之聲道中之至少一者中或在輸出音訊格式之至少一個個別分量中)作為合成音訊信號。在一些情況下，該多個合成分量音訊信號可在降混信號 228 之聲道中，且在一些情況下，其可在空間呈現之內部聲道中之一者中。

【0108】圖 7 展示其中自合成音訊信號 228 獲得之至少 K 個聲道 228a (例如，在其版本 228b 中，其在濾波器組分析 720 下游)可去相關的實例。舉例而言，當合成合成器 210 在合成音訊信號 228 之 M 個聲道中之至少一者中產生合成音訊信號 228 時，獲得此情形。此相關處理 730 可在濾波器組分析區塊 720 下游應用於信號 228b (或其分量中之至少一者或一些)，使得可獲得至少 K 個聲道(其中 $K \geq M$ 及/或 $K \leq N$ ，其中 N 為輸出聲道之數目)。隨後，可將 K 個去相關聲道 228a 及/或信號 228b 之 M 個聲道提供至區塊 740 以用於產生混合增益/矩陣，該混合增益/矩陣可經由空間參數 218、219 (參見上文)提供混合信號 742。混合信號 742 可經受濾波器組合成區塊 746，以在 N 個輸出聲道 202 中獲得輸出信號。基本上，圖 7 之元件符號 228a 可為自個別合成分量音訊信號 228b 去相關之個別合成分量音訊信號，使得空間呈現器(及區塊 740)利用分量 228a 與分量 228b 之組合。圖 8 展示全部聲道 228 產生於 K 個聲道中之實例。

【0109】此外，在圖 7 中，應用於 K 個去相關聲道 228b 之去相關器 730 在濾波器組分析區塊 720 下游。此可例如針對擴散場執行。在一些情況下，信號 228b 之 M 個聲道在反饋分析區塊 720 下游且可經提供至區塊 744，從而產生混合增益/矩陣。協方差法可用於例如藉由將聲道 228b 按與不同聲道之間的協方

差互補之值相關聯的值縮放而減少去相關器 730 之問題。

【0110】圖 8 展示在頻域中之合成信號合成器 210 之實例。協方差法可用於圖 8 之合成合成器 210 (810)，值得注意的是，合成音訊合成器 210 (810)在 K 個聲道中提供其輸出 228c，同時將在 M 個聲道中提供傳送聲道 228 (其中 $K \geq M$)。

【0111】圖 9 展示解碼器 900 (解碼器 200 之實施例)之實例，其可理解為利用圖 8 之解碼器 800 及圖 7 之解碼器 700 的混合技術。如此處可見，合成信號合成器 210 包括在降混信號 228 之 M 個聲道中產生合成音訊信號 228 的第一部分 210 (710)。信號 228 可輸入至濾波器組分析區塊 730，該濾波器組分析區塊可提供輸出 228b，其中多個濾波器帶區別於彼此。此時，可使聲道 228b 去相關以在 K 個聲道中獲得去相關信號 228a。同時， M 個聲道中之濾波器組分析的輸出 228b 經提供至區塊 740，以用於產生可提供混合信號 742 之混合版本的混合增益矩陣。混合信號 742 可考慮非作用空間參數 318 及/或用於非作用訊框 219 之經恢復(經重建構)空間參數。應注意，去相關器 730 之輸出 228a 亦可在加法器 920 處加至合成信號合成器 210 之第二部分 810 之輸出 228d，該合成信號合成器在 K 個聲道中提供合成信號 228d。在加法區塊 920 處，可將信號 228d 加總至去相關信號 228a 以將經加總信號 228e 提供至混合區塊 740。因此，有可能藉由使用分量 228b 與分量 228e 之組合來呈現最終輸出信號 202，該分量 228e 考慮了去相關分量 228a 及所產生分量 228d 二者。圖 8 及圖 7 之分量 228b、228a、228d、228e (存在)可理解為例如合成信號 228 之擴散及非擴散分量。詳言之，參考圖 9 之解碼器 900，基本上，信號 228e 之低頻帶可自傳送聲道 710 獲得(且自 228a 獲得)且信號 228e 之高頻帶可在合成器 810 中產生(且在聲道 228d 中)，低頻帶及高頻帶在加法器 920 處之相加准許在信號 228e 中具有其兩者。

【0112】值得注意的是，在以上圖 7 至圖 10 中，未展示用於作用訊框之傳送聲道解碼器。

【0113】 圖 10 展示解碼器 1000 (解碼器 200 之實施例)之實例，其中展示音訊解碼器 230 (其提供經解碼聲道 226)及合成信號合成器 210 (此處被視為劃分成第一外部部分 710 與第二內部部分 810)二者。展示出切換器 224'，其可類似於圖 6 之切換器(例如，受由活動偵測器 220 提供之控制或命令 221'控制)。基本上，有可能在將經解碼音訊場景 226 提供至空間呈現器 220 之模式與提供合成音訊信號 228 之另一模式之間進行選擇。降混信號 224 (226、228)在通常少於輸出信號 202 之 N 個輸出聲道的 M 個聲道中。

【0114】 信號 224 (226、228)可輸入至濾波器組分析區塊 720。濾波器組分析 720 之輸出 228b (在多個頻率區間中)可輸入至升混加法區塊 750 上，該輸出亦可由合成信號合成器 210 之第二部分 810 提供的信號 228d 輸入。升混加法區塊 750 之輸出 228f 可輸入至相關器處理 730。去相關器處理 730 之輸出 228a 可與升混加法區塊 750 之輸出 228f 一起提供至區塊 740，以用於產生混合增益及矩陣。升混加法區塊 750 可例如將聲道之數目自 M 增大至 K (且在一些情況下，其可將該等聲道例如以恆定係數倍增)且可將 K 個聲道與由合成信號合成器 210 (例如，第二內部部分 810)產生之 K 個聲道 228d 相加。為呈現第一(作用)訊框，混合區塊 740 可考慮如提供於位元串流 304 中之作用空間參數 316 中之至少一者，如以外推或其他方式獲得之經恢復(經重建構)空間參數 210 (參見上文)。

【0115】 在一些實例中，濾波器組分析區塊 720 之輸出可在 M 個聲道中，但可考慮不同頻帶。對於第一訊框(及如位於圖 10 中之切換器 224' 及切換器 222')，經解碼信號 226 (在至少二個聲道中)可提供至濾波器組分析 720，且可因此經由 K 個雜訊聲道 228d (合成信號聲道)在升混加法區塊 750 處加權以在 K 個聲道中獲得信號 228f。應記住， $K \geq M$ 且可包含例如擴散聲道及定向聲道。特定言之，擴散聲道可由去相關器 730 去相關以獲得去相關信號 228a。因此，經解碼音訊信號 224 可(例如，在區塊 750 處)藉由合成音訊信號 228d 加權，該合

成音訊信號可掩蔽作用訊框與非作用訊框(第一訊框與第二訊框)之間的轉變。隨後，合成信號合成器 210 之第二部分 810 不僅用於作用訊框且用於非作用訊框。

【0116】圖 11 展示解碼器 200 之另一實例，其可包含第一訊框(346)、第一聲場參數表示(316)及經編碼音訊信號(346)，其中第二訊框(348)係非作用訊框，設備包含：活動偵測器(220)，其用於偵測第二訊框(348)係非作用訊框且用於提供用於第二訊框(308)之參數描述(328)；合成信號合成器(210)，其用於使用用於第二訊框(308)之參數描述(348)合成用於第二訊框(308)之合成音訊信號(228)；音訊解碼器(230)，其用於解碼用於第一訊框(306)之經編碼音訊信號(346)；以及空間呈現器(240)，其用於使用第一聲場參數表示(316)且使用用於第二訊框(308)之合成音訊信號(228)在空間上呈現用於第一訊框(306)之音訊信號(202)，或轉碼器，其用於產生元資料輔助輸出格式，該元資料輔助輸出格式包含用於第一訊框(306)之音訊信號(346)、用於第一訊框(306)之第一聲場參數表示(316)、用於第二訊框(308)之合成音訊信號(228)及用於第二訊框(308)之第二聲場參數表示(318)。

【0117】在上文實例中參考合成信號合成器 210，如上文所解釋，其可包含(或甚至為)雜訊產生器(例如，舒適雜訊產生器)。在實例中，合成信號產生器(210)可包含雜訊產生器且第一個別合成分量音訊信號係由雜訊產生器之第一取樣產生，且第二個別合成分量音訊信號係由雜訊產生器之第二取樣產生，其中第二取樣不同於第一取樣。

【0118】另外或替代地，雜訊產生器包含雜訊表，且其中第一個別合成分量音訊信號係藉由獲取雜訊表之第一部分而產生，且其中第二個別合成分量音訊信號係藉由獲取雜訊表之第二部分而產生，其中雜訊表之第二部分不同於雜訊表之第一部分。

【0119】在實例中，雜訊產生器包含偽雜訊產生器，且其中第一個別合成

分量音訊信號係藉由使用偽雜訊產生器之第一種子而產生，且其中第二個別合成分量音訊信號係使用偽雜訊產生器之第二種子而產生。

【0120】 一般而言，在圖 6、圖 7、圖 9、圖 10 及圖 11 之實例中，空間呈現器 220 可使用直接信號與由去相關器(730)根據在第一聲場參數表示(316)之控制下的直接信號產生之擴散信號的混合，以用於第一訊框(306)之第一模式進行操作，且使用第一合成分量信號與第二合成分量信號之混合，以用於第二訊框(308)之第二模式進行操作，其中第一合成分量信號及第二合成分量信號係由合成信號合成器(210)藉由雜訊程序或偽雜訊程序之不同實現而產生。

【0121】 如上文所解釋，空間呈現器(220)可經組配以在第二模式下藉由參數處理器利用針對第二訊框(308)導出的擴散度參數、能量分佈參數或相干性參數控制混合(740)。

【0122】 以上實例亦關於一種自具有第一訊框(306)及第二訊框(308)之音訊信號產生經編碼音訊場景的方法，其包含：根據第一訊框(306)中之音訊信號判定用於第一訊框(306)之第一聲場參數表示(316)，且根據第二訊框(308)中之音訊信號判定用於第二訊框(308)之第二聲場參數表示(318)；分析音訊信號以取決於音訊信號而判定第一訊框(306)為作用訊框且第二訊框(308)為非作用訊框；針對第一訊框(306)為作用訊框而產生經編碼音訊信號，且針對第二訊框(308)為非作用訊框而產生參數描述(348)；以及藉由將用於第一訊框(306)之第一聲場參數表示(316)、用於第二訊框(308)之第二聲場參數表示(318)、用於第一訊框(306)之經編碼音訊信號及用於第二訊框(308)之參數描述(348)組合在一起而構成經編碼音訊場景。

【0123】 以上實例亦關於一種處理經編碼音訊場景之方法，該經編碼音訊場景在第一訊框(306)中包含第一聲場參數表示(316)及經編碼音訊信號，其中第二訊框(308)為非作用訊框，該方法包含：偵測第二訊框(308)為非作用訊框且用

於提供用於第二訊框(308)之參數描述(348)；使用用於第二訊框(308)之參數描述(348)合成用於第二訊框(308)之合成音訊信號(228)；解碼用於第一訊框(306)之經編碼音訊信號；以及使用第一聲場參數表示(316)且使用用於第二訊框(308)之合成音訊信號(228)在空間上呈現用於第一訊框(306)之音訊信號，或產生元資料輔助輸出格式，該元資料輔助輸出格式包含用於第一訊框(306)之音訊信號、用於第一訊框(306)之第一聲場參數表示(316)、用於第二訊框(308)之合成音訊信號(228)及用於第二訊框(308)之第二聲場參數表示(318)。

【0124】 亦提供經編碼音訊場景(304)，其包含：用於第一訊框(306)之第一聲場參數表示(316)；用於第二訊框(308)之第二聲場參數表示(318)；用於第一訊框(306)之經編碼音訊信號；以及用於第二訊框(308)之參數描述(348)。

【0125】 在以上實例中，有可能針對每一頻帶(次頻帶)傳輸空間參數 316 及/或 318。

【0126】 根據一些實例，此靜默參數描述 348 可含有此部分參數 318，該部分參數可因此為 SID 348 之一部分。

【0127】 用於非作用訊框之空間參數 318 對於每一頻率次頻帶(或頻帶或頻率)可為有效的。

【0128】 上文所論述之在作用階段 346 期間傳輸或編碼且及在 SID 348 中之空間參數 316 及/或 318 可具有不同頻率解析度，且另外或替代地，上文所論述之在作用階段 346 期間傳輸或編碼且及在 SID 348 中之空間參數 316 及/或 318 可具有不同時間解析度，且另外或替代地，上文所論述之在作用階段 346 期間傳輸或編碼且及在 SID 348 中之空間參數 316 及/或 318 可具有不同量化解析度。

【0129】 應注意，解碼裝置及編碼裝置可為如 CELP 或 DCX 或頻寬擴展模組之裝置。

【0130】 亦有可能利用基於 MDCT 之寫碼方案(修改型離散餘弦轉換)。

【0131】 在解碼器設備 200 之本實例中(在其任一實施例中，例如圖 6 至圖 11 之彼等實施例)，有可能用轉碼器取代音訊解碼器 230 及空間呈現器 240，以用於產生元資料輔助輸出格式，該元資料輔助輸出格式包含用於第一訊框之音訊信號、用於第一訊框之第一聲場參數表示、用於第二訊框之合成音訊信號及用於第二訊框之第二聲場參數表示。

論述

【0132】 本發明之實施例提出一種將 DTX 擴展至參數空間音訊寫碼的方式。因此提議將習知 DTX/CNG 應用於降混/傳送聲道(例如，324、224)且在解碼器側處藉由空間參數(稱為後方空間 SID)，例如 316、318 及在非作用訊框(例如，308、328、348、228)上之空間呈現來擴展該降混/傳送聲道。為恢復非作用訊框(例如，308、328、348、228)之空間影像，用經特定設計且與沉浸式背景雜訊相關之一些空間參數(空間 SID) 319 (或 219)修正傳送聲道 SID 326、226。本發明之實施例(下文論述及/或上文論述)覆蓋至少二個態樣：

擴展傳送聲道 SID 以用於空間呈現。為此，用例如自 DirAC 範式或 MASA 格式導出之空間參數 318 修正描述符。如擴散度 314a 及/或一或多個到達方向 314b 及/或聲道間/環繞聲相干性及/或能量比之參數 318 中之至少一者可連同傳送聲道 SID 328 (348)一起傳輸。在某些情況下且在某些假定下，可捨棄一些參數 318。舉例而言，若假定背景雜訊完全擴散，則吾等可捨棄隨後無意義的方向 314b 之傳輸。

藉由在空間中呈現傳送聲道 CNG 而在接收器側對非作用訊框進行空間化：DirAC 合成原理或其衍生物中之一者可由背景雜訊之空間 SID 描述符內的最終傳輸之空間參數 318 導引。至少存在二個選項，其甚至可以組合：可僅針對傳送聲道 228 產生傳送聲道舒適雜訊產生(此為圖 7 之情況，其中舒適雜訊 228 係由合成信號合成器 710 產生)；亦可針對傳送聲道以及呈現器中用於升混之額外

聲道產生傳送聲道 CNG (此為圖 9 之情況，其中一些舒適雜訊 228 係由合成信號合成器第一部分 710 產生，但其他一些舒適雜訊 228d 係由合成信號合成器第二部分 810 產生)。在最新情況下，例如用不同種子對隨機雜訊 228d 取樣之 CNG 第二部分 710 可自動地使所產生之聲道 228d 去相關且最小化去相關器 730 之採用，該去相關器可為典型偽聲源。此外，亦可在作用訊框中使用 CNG (如圖 10 中所示)，但在一些實例中，其中平滑化作用與非作用階段(訊框)之間的轉變之強度減小，且亦掩蔽來自傳送聲道寫碼器及參數 DirAC 範式之最終偽聲。

【0133】 圖 3 描繪編碼器設備 300 之實施例的概述。在編碼器側，信號可由 DirAC 分析來分析。DirAC 可分析如 B 格式或第一階立體混響(FOA)之信號。然而，亦有可能將原理擴展至高階立體混響(HOA)，且甚至擴展至與如[10]中所提出之如 5.1 或 7.1 或 7.1+4 之給定揚聲器設置相關聯之多聲道信號。輸入格式 302 亦可為表示藉由包括於相關聯元資料中之資訊而定位於空間中之一或若干不同音訊物件的個別音訊聲道。替代地，輸入格式 302 可為元資料相關聯空間音訊(MASA)。在此情況下，空間參數及傳送聲道直接傳送至編碼器設備 300。可隨後跳過音訊場景分析(例如如圖 5 中所示)，且僅必須針對空間參數 318 之非作用集合或針對空間參數 316、318 之作用及非作用集合二者執行最終空間參數(再)量化及再取樣。

【0134】 可針對作用及非作用訊框 306、308 二者進行音訊場景分析且產生二組空間參數 316、318。在作用訊框 308 之情況下產生第一組空間參數 316，且在非作用訊框 308 之情況下產生另一組空間參數(318)。有可能不具有非作用空間參數，但在本發明之較佳實施例中，相比於作用空間參數 316，非作用空間參數 318 較少及/或經較粗略量化。此後，可獲得二個版本之空間參數(亦稱作 DirAC 元資料)。重要的是，本發明之實施例可主要係關於自聽者視角之音訊場景的空間表示。因此，考慮如 DirAC 參數 318、316 之空間參數，包括一或若干

個方向以及最終擴散度因數或一或多個能量比。不同於聲道間參數，來自聽者視角之此等空間參數具有不可知聲音捕捉及再現系統之較大優勢。此參數化並非特定針對於任何特定麥克風陣列或揚聲器佈局。

【0135】 話音活動偵測器(或更一般而言，活動偵測器) 320 可隨後應用於由音訊場景分析器產生之輸入信號 302 及/或傳送聲道 326。傳送聲道小於輸入聲道之數目；通常為單聲道降混、立體聲降混、A 格式或第一階立體混響信號。基於 VAD 決策，處理程序下之當前訊框經界定為作用(306、326)或非作用(308、328)。在作用訊框(306、326)之情況下，執行傳送聲道之習知語音或音訊編碼。所得碼資料隨後與作用空間參數 316 組合。在非作用訊框(308、328)之情況下，傳送聲道 324 之靜默資訊描述 328 通常在非作用階段期間以規則訊框間隔，例如每隔 8 個作用訊框(306、326、346)以章節方式產生。傳送聲道 SID (328、348) 可隨後在具有非作用空間參數之多工器(經編碼信號形成器) 370 中經修正。在非作用空間參數 318 為空值之情況下，隨後僅傳輸傳送聲道 SID 348。總 SID 通常可為極低位元速率描述，其例如低至 2.4 或 4.25 kbps。在非作用階段中，平均位元速率甚至更低，此係由於大部分時間未進行傳輸且不發送資料。

【0136】 在本發明之較佳實施例中，傳送聲道 SID 348 具有 2.4 kbps 之大小，且包括空間參數之總 SID 具有 4.25 kbps 之大小。對於 DirAC 具有如 FOA 之多聲道信號作為輸入，非作用空間參數之計算在圖 4 中予以描述，該多聲道信號可直接自高階立體混響(HOA)導出，對於 MASA 輸入格式，非作用空間參數之計算在圖 5 中予以描述。如前所述，可與作用空間參數 316 並行地導出非作用空間參數 318，從而平均化及/或再量化已寫碼之作用空間參數 318。在如 FOA 之多聲道信號作為輸入格式 302 之情況下，多聲道信號 302 之濾波器組分析可在計算各時間及頻率塊之空間參數、方向及擴散度之前執行。元資料編碼器 396、398 可在應用量化器及寫碼經量化參數之前平均化不同頻帶及/或時槽上之參數

316、318。其他非作用空間元資料編碼器可繼承在作用空間元資料編碼器中導出之經量化參數中之一些以將其直接用於非作用空間參數中或將其再量化。在 MASA 格式之情況下(例如圖 5)，首先可讀取輸入元資料且以給定時間頻率及位元度解析度提供元資料編碼器 396、398。一或多個元資料編碼器 396、398 隨後將進一步藉由以下操作進行處理：最終轉換一些參數，調適其解析度(亦即，降低例如對其進行平均化之解析度)以及在例如藉由熵寫碼方案對其寫碼之前再量化該等參數。

【0137】 如例如圖 6 中所描繪，在解碼器側首先藉由偵測所傳輸封包(例如，訊框)之大小或藉由偵測封包之非傳輸來恢復 VAD 資訊 221 (例如，訊框被分類為作用抑或非作用)。在作用訊框 348 中，解碼器在作用中模式運行，且傳送聲道寫碼器有效負載以及作用空間參數經解碼。空間呈現器 220 (DirAC 合成) 隨後使用呈輸出空間格式之經解碼空間參數 316、318 對經解碼傳送聲道進行升混/空間化。在非作用訊框中，舒適雜訊可藉由傳送聲道 CNG 部分 810 (例如在圖 10 中)產生於傳送聲道中。CNG 由傳送聲道 SID 導引以用於通常調整能量及頻譜形狀(經由例如應用於頻域中之比例因數或應用於時域合成濾波器之線性預測寫碼係數)。一或多個舒適雜訊 228d、228a 等隨後在此時由非作用空間參數 318 導引之空間呈現器(DirAC 合成) 740 中呈現/空間化。輸出空間格式 202 可為雙耳信號(2 個聲道)、給定揚聲器佈局之多聲道或呈立體混響格式之多聲道信號。在替代性實施例中，輸出格式可為元資料輔助空間音訊(MASA)，其意謂經解碼傳送聲道或傳送聲道舒適雜訊連同作用或非作用空間參數分別直接輸出以用於由外部裝置呈現。

非作用空間參數之編碼及解碼

【0138】 非作用空間參數 318 可由頻帶中之多個方向及對應於一個方向分量與總能量之比率的頻帶中相關聯能量比中之一者組成。在一個方向之情況

下，如在較佳實施例中，能量比可由擴散度替換，該擴散度與能量之比互補且隨後遵循參數之原始 **DirAC** 集合。由於一或多個方向分量一般預期為與非作用訊框中之擴散部分相比較不相關，因此其亦可諸如在作用訊框中使用較粗略量化方案及/或藉由對隨時間推移之方向或頻率求平均值以獲得較粗略時間及/或頻率解析度而在較少位元上傳輸。在較佳實施例中，可針對作用訊框每 20 ms 而非 5 ms，但使用 5 個非均一頻帶之相同頻率解析度發送方向。

【0139】 在較佳實施例中，擴散度 314a 可以與作用訊框中相同的時間/頻率但以較少位元傳輸，從而迫使實現最小量化索引。舉例而言，若擴散度 314a 在作用訊框中之 4 個位元上經量化，則其隨後僅在 2 個位元上傳輸，從而避免原始索引自 0 至 3 之傳輸。經解碼索引隨後將添加偏移量+4。

【0140】 在一些實例中，亦有可能完全避免發送方向 314b 或替代地避免發送擴散度 314a 且在解碼器處將其替換為預設值或估計值。

【0141】 此外，若輸入聲道對應於定位於空間域之聲道，則吾等可考慮傳輸聲道間相干性。聲道間聲級差亦為方向之替代方案。

【0142】 更相關的是發送環繞聲相干性，該環繞聲相干性被界定為在聲場中相干的擴散能量之比。可例如藉由在直接信號與擴散信號之間重分佈能量而在空間呈現器(**DirAC** 合成)處利用該環繞聲相干性。環繞相干分量之能量自待重分佈至方向分量之擴散能量移除，該等方向分量隨後將在空間中更均一地平移。

【0143】 自然地，對於非作用空間參數，可考慮先前所列之參數之任何組合。出於節省位元之目的，亦可設想在非作用階段中不發送任何參數。

【0144】 非作用空間元資料編碼器之例示性偽程式碼在下文給出：

```
bistream = inactive_spatial_metadata_encoder (
    azimuth, /* i:來自作用空間元資料編碼器之方位角值*/
    elevation, /* i:來自作用空間元資料編碼器之仰角值*/
```

```

diffuseness_index, /*i/o: 來自作用空間元資料編碼器之擴散度索引*/
metadata_sid_bits /*i 分配至非作用空間元資料(空間 SID)之位元*/
)
{
    /*發信 2D*/
    not_in_2D = 0;
    for ( b = start_band; b < nbands; b++ )
    {
        for ( m = 0; m < nblocks; m++ )
        {
            not_in_2D += elevation[b][m];
        }
    }

    write_next_indice( bistream, (not_in_2D > 0 ), 1 ); /*2D 旗標*/

    /*計數所需位元*/
    bits_dir = 0;
    bits_diff = 0;
    for ( b = start_band; b < nbands; b++ )
    {
        diffuseness_index[b] = max( diffuseness_index[b], 4 );
        bits_diff += get_bits_diffuseness(diffuseness_index[b] - 4,
DIRAC_DIFFUSE_LEVELS - 4);
        if ( not_in_2D == 0 )

```

```

{
    bits_dir += get_bits_azimuth(diffuseness_index[b]);
}
else
{
    bits_dir += get_bits_spherical(diffuseness_index[b]);
}
}

/*藉由增加擴散度索引減少位元需求*/
bits_delta = metadata_sid_bits - 1 - bits_diff - bits_dir;
while ( ( bits_delta < 0 ) && (not_in_2D > 0 ) )
{
    for ( b = nbands - 1; b >= start_band && ( bits_delta < 0 ); b-- )
    {
        if ( diffuseness_index[b] < ( DIRAC_DIFFUSE_LEVELS - 1 ) )
        {
            bits_delta += get_bits_spherical(diffuseness_index[b]);
            diffuseness_index[b]++;
            bits_delta -= get_bits_spherical(diffuseness_index[b]);
        }
    }
}
}

```

```

/*寫入擴散度索引*/

for ( b = start_band; b < nbands; b++ )
{
    Write_diffuseness(bitstream,          diffuseness_index[b]-          4,
DIRAC_DIFFUSE_LEVELS - 4);
}

/*計算且追蹤每帶之平均方向*/

for ( b = start_band; b < nbands; b++ )
{
    set_zero( avg_direction_vector, 3 );
    for ( m = 0; m < nblocks; m++ )
    {
        /*計算平均方向*/
        azimuth_elevation_to_direction_vector(azimuth[b][m], elevation[b][m],
direction_vector );
        v_add( avg_direction_vector, direction_vector, avg_direction_vector, 3 );
    }
    direction_vector_to_azimuth_elevation(          avg_direction_vector,
&avg_azimuth[b], &avg_elevation[b] );

    /*量化平均方向*/

    if ( not_in_2D > 0 )
    {

```

```

        Code_and_write_spherical_angles(bitsream,          avg_elevation[b],
avg_azimuth[b],          get_bits_spherical(diffuseness_index[b]));
    }
    else
    {
        Code_and_write_azimuth          (bitsream,          avg_azimuth[b],
get_bits_azimuth(diffuseness_index[b]));
    }
}

```

```

For(i=0; i<delta_bits; i++)
{
    Write_next_bit ( bitstream, 0); /*用值 0 填充位元*/
}
}

```

【0145】 非作用空間元資料解碼器之例示性偽程式碼在下文給出：

```
[diffuseness, azimuth, elevation] = inactive_spatial_metadata_decoder(bitstream)
```

```
/*讀取 2D 發信*/
```

```
not_in_2D = read_next_bit(bitstream);
```

```
/*解碼擴散度*/
```

```
for ( b = start_band; b < nbands; b++ )
```

```
{
```

```

        diffuseness_index[b]      =      read_diffuseness_index(      bitstream,
DIFFUSE_LEVELS - 4 ) + 4;

        diffuseness_avg = diffuseness_reconstructions[diffuseness_index[b]];

        for ( m = 0; m < nblocks; m++ )

                diffuseness[b][m] = diffusenessavg;
}

```

/*解碼器 DOA*/

```

if (not_in_2D > 0)
{
        for ( b = start_band; b < nbands; b++ )
        {
                bits_spherical = get_bits_spherical(diffuseness_index[b]);
                spherical_index = Read_spherical_index( bitstream, bits_spherical);
                azimuth_avg = decode_azimuth(spherical_index, bits_spherical);
                elevation_avg = decode_elevation(spherical_index, bits_spherical);

                for ( m = 0; m < nblocks; m++ )
                {
                        elevation[b][m] *= 0.9f;

                        elevation[b][m] += 0.1f * elevation_avg;

                        azimuth[b][m] *= 0.9f;

                        azimuth[b][m] += 0.1f * azimuth_avg;

                }

```

```

    }
}
else
{
    for ( b = start_band; b < nbands; b++ )
    {
        bits_azimuth = get_bits_azimuth(diffuseness_index[b]);
        azimuth_index = Read_azimuth_index( bitstream, bits_azimuth);
        azimuth_avg = decode_azimuth(diffuseness_index, bits_azimuth);
        for ( m = 0; m < nblocks; m++ )
        {
            elevation[b][m] *= 0.9f;
            azimuth[b][m] *= 0.9f;
            azimuth[b][m] += 0.1f * azimuth_avg;
        }
    }
}

```

在解碼器側非傳輸之情況下恢復空間參數

【0146】 在 SID 處於非作用階段期間之情況下，空間參數可經完全或部分解碼且隨後用於後續 DirAC 合成。

【0147】 在無資料傳輸之情況下或在無空間參數 318 與該 348 之傳送聲道一起傳輸之情況下，可能需要復原空間參數 219。此可藉由考慮過去接收之參數 (例如，316 及 7 或 318) 以合成方式產生遺失參數 219 (例如圖 7 至圖 10) 而達成。不穩定空間影像可在感知上令人不適，尤其對於視為穩定且並不快速放出的背

景雜訊。另一方面，絕對恆定之空間影像可被感知為不自然的。可應用不同策略：

維持策略

【0148】認為空間影像必須隨時間推移而相對穩定通常係安全的，其可針對 DirAC 參數，即在訊框之間不會改變很多的 DOA 及擴散度而進行轉譯。出於此原因，簡單但有效之方法為保持最後接收之空間參數 316 及/或 318 作為經恢復空間參數 219。其至少對於具有長期特性之擴散度為極穩固之方法。然而，對於方向，可設想不同策略，如下所列。

方向之外推：

【0149】替代地或另外，可設想估計音訊場景中之聲音事件的軌跡且接著嘗試外推所估計軌跡。在聲音事件作為點源在空間中良好局部化的情況下尤其相關，該點源藉由低擴散度在 DirAC 模型中反映。所估計軌跡可自過去方向之觀測結果計算且擬合此等點當中之曲線，此可演進內插抑或平滑化。亦可採用回歸分析。接著可藉由評估經擬合曲線超出所觀測資料之範圍(例如，包括先前參數 316 及/或 318)執行參數 219 之外推。然而，此方法可導致對於非作用訊框 348 相關性較低，其中背景雜訊係無用的且預期大部分擴散。

方向之抖動：

【0150】當聲音事件更為擴散(其特別係對於背景雜訊之情況)時，方向不大具有意義且可被視為隨機程序之實現。抖動可接著藉由在將隨機雜訊用於非傳輸訊框之前將隨機雜訊注入至先前方向而幫助使所呈現聲場愈加自然且愈加令人愉快。所注入之雜訊及其方差可隨擴散度變化。舉例而言，方位角及仰角中所注入之雜訊的方差 σ_{azi} 及 σ_{ele} 可遵循擴散度 Ψ 之簡單模型函數，如下：

$$\sigma_{azi} = 65\Psi^{3.5} + \sigma_{ele}$$

$$\sigma_{ele} = 33.25\Psi + 1.25$$

舒適雜訊產生及空間化(解碼器側)

【0151】 現論述上文提供之一些實例。

【0152】 在第一實施例中，舒適雜訊產生器 210 (710)在如圖 7 中所描繪之核心解碼器中進行。所得舒適雜訊注入於傳送聲道中且接著藉助於所傳輸非作用空間參數 318 或在非傳輸情況下使用如先前所描述推導出之空間參數 219 在 DirAC 合成中空間化。時空可接著例如藉由產生二個串流來實現如較早所描述之方式，該二個串流係衍生自經解碼傳送聲道之方向性及非方向性串流，且在非作用訊框之情況下來自傳送聲道舒適雜訊。接著將二個串流升混且在區塊 740 處取決於空間參數 318 而混合在一起。

【0153】 替代地，舒適雜訊或其部分可直接在濾波器組域中之 DirAC 合成內產生。實際上，DirAC 可藉助於傳送聲道 224、空間參數 318、316、319 及一些去相關器(例如，730)來控制經復原場景之相干性。去相關器 730 可減小合成聲場之相干性。隨後在頭戴式耳機再現之情況下以更多寬度、深度、擴散、混響或外化感知空間影像。然而，去相關器常常傾向於典型的可聽偽聲，且希望減少其使用。此可例如藉由所謂的協方差合成方法[5]藉由利用傳送聲道之已存在的非相干分量來達成。然而，此方法可具有限制，尤其在單音傳送聲道之情況下。

【0154】 如果舒適雜訊係由隨機噪聲產生，則針對每一輸出聲道或其至少一子集產生專用舒適雜訊係有利的。更具體言之，有利於不僅對傳送聲道而且對空間呈現器(DirAC 合成) 220 中(及在混合區塊 740 中)所使用之中間音訊聲道應用舒適雜訊產生。擴散場之去相關隨後將藉由使用不同雜訊產生器，而非使用可降低偽聲之量且亦降低總複雜度之去相關器 730 而直接給出。實際上，按照定義，隨機雜訊之不同實現係去相關的。圖 8 及圖 9 示出藉由完全或部分地在空間呈現器 220 內產生舒適雜訊達成此情形之二種方式。在圖 8 中，CN 如[5]

中所描述在頻域中進行，其可直接藉由空間呈現器之濾波器組域而產生，從而避免濾波器組分析 720 及去相關器 730 二者。此處，舒適雜訊針對其而產生的聲道之數目 K 等於或大於傳送聲道之數目 M ，且低於或等於輸出聲道之數目 N 。在最簡單的情況下， $K=N$ 。

【0155】圖 9 示出在呈現器中包括舒適雜訊產生 810 之另一替代方案。舒適雜訊產生在空間呈現器 220 內部(710 處)與外部(810 處)之間分離。將呈現器 220 內之舒適雜訊 228d 添加(在加法器 920 處)至最終去相關器輸出 228a。舉例而言，低頻帶可在與核心寫碼器中相同之域外產生，以便能夠容易地更新所需記憶體。另一方面，舒適雜訊產生可為高頻率而直接在呈現器中執行。

【0156】此外，舒適雜訊產生亦可在作用訊框 346 期間應用。代替在作用訊框 346 期間完全斷開舒適雜訊產生，其可藉由減小其強度而保持在作用中。其隨後用於掩蔽作用與非作用訊框之間的過渡，亦掩蔽核心寫碼器及參數空間音訊模型二者的偽聲及缺陷。此在[11]中針對單音語音寫碼提出。相同原理可擴展至空間語音寫碼。圖 10 示出實施方式。此時空間呈現器 220 中之舒適雜訊產生開啟作用階段與非作用階段二者。在非作用階段 348 中，其與在傳送聲道中執行的舒適雜訊產生互補。在呈現器中，在等於或大於 M 個傳送聲道的 K 個聲道上達成舒適雜訊，旨在減少去相關器之使用。空間呈現器 220 中之舒適雜訊產生被加至傳送聲道之升混版本 228f，該升混版本可藉由 M 個聲道至 K 個聲道之簡單複製來達成。

態樣

【0157】對於編碼器：

1. 一種用以編碼具有多個聲道或一或若干個音訊聲道之空間音訊格式與描述音訊場景之元資料的音訊編碼器設備(300)，其包含以下各者中之至少一者：

- a. 空間音訊輸入信號(302)之場景音訊分析器(310)，其經組配以產生描述含有一或若干個傳送聲道的輸入信號(202)之空間影像及降混版本(326)的第一組或第一組及第二組空間參數(318、319)，傳送聲道之數目小於輸入聲道之數目；
 - b. 傳送聲道編碼器裝置(340)，其經組配以在作用階段(306)藉由編碼含有傳送聲道的經降混信號(326)來產生經編碼資料(346)；
 - c. 傳送聲道靜默插入描述符(350)，其在非作用階段(308)中產生傳送聲道(328)之背景雜訊的靜默插入描述(348)；
 - d. 多工器(370)，其用於在作用階段(306)期間將第一組空間參數(318)與經編碼資料(344)組合成位元串流(304)，且用於在非作用階段(308)期間不發送資料或用於發送靜默插入描述(348)，或組合發送靜默插入描述(348)及第二組空間參數(318)。
2. 如 1 之音訊編碼器，其中場景音訊分析器(310)遵循方向音訊寫碼(DirAC)原理。
 3. 如 1 之音訊編碼器，其中場景音訊分析器(310)解譯輸入元資料以及一或若干個傳送聲道(348)。
 4. 如 1 之音訊編碼器，其中場景音訊分析器(310)自輸入元資料導出一或二組參數(316、318)且自一或若干個輸入音訊聲道導出傳送聲道。
 5. 如 1 之音訊編碼器，其中空間參數為一或若干個到達方向(DOA)(314b)，或擴散度(314a)，或一或若干個相干性。
 6. 如 1 之音訊編碼器，其中針對不同頻率次頻帶導出空間參數。
 7. 如 1 之音訊編碼器，其中傳送聲道編碼裝置遵循 CELP 原理，或為基於 MDCT 之寫碼方案或二個方案之切換組合。
 8. 如 1 之音訊編碼器，其中作用階段(306)及非作用階段(308)由對傳送聲道執行之話音活動偵測器(320)判定。

9. 如 1 之音訊編碼器，其中第一組及第二組空間參數(316、318)在時間或頻率解析度，或量化解析度，或參數之性質方面不同。

10. 如 1 之音訊編碼器，其中空間音訊輸入格式(202)呈立體混響格式或 B 格式，或為與給定揚聲器設置相關聯之多聲道信號，或自麥克風陣列導出之多聲道信號，或一組個別音訊聲道以及元資料，或元資料輔助空間音訊(MASA)。

11. 如 1 之音訊編碼器，其中空間音訊輸入格式由二個以上音訊聲道組成。

12. 如 1 之音訊編碼器，其中傳送聲道之數目為 1、2 或 4 (可選擇其他數目)。

對於解碼器：

【0158】 1. 一種用於解碼位元串流(304)之音訊解碼器設備(200)，以便自空間音訊輸出信號(202)產生位元串流，該位元串流(304)包含繼之以至少非作用階段(308)之至少作用階段(306)，其中位元串流已在其中編碼至少靜默插入描述符訊框 SID (348)，該靜默插入描述符訊框描述傳送/降混聲道(228)之背景雜訊特性及/或空間影像資訊，該音訊解碼器設備(200)包含以下各者中之至少一者：

a. 靜默插入描述符解碼器(210)，其經組配以解碼靜默 SID (348)，以便重建構傳送/降混聲道(228)中之背景雜訊；

b. 解碼裝置(230)，其經組配以在作用階段(306)期間重建構來自位元串流(304)之傳送/降混聲道(226)；

c. 空間呈現裝置(220)，其經組配以在作用階段(306)期間重建構(740)來自經解碼傳送/降混聲道(224)之空間輸出信號(202)及所傳輸空間參數(316)，且在非作用階段(308)期間重建構來自傳送/降混聲道(228)中經重建構背景雜訊之空間輸出信號。

2. 如 1 之音訊解碼器，其中在作用階段中傳輸之空間參數(316)由擴散度或到達方向或相干性組成。
3. 如 1 之音訊解碼器，其中空間參數(316、318)藉由頻率次頻帶傳輸。
4. 如 1 之音訊解碼器，其中靜默插入描述(348)除傳送/降混聲道(228)之背景雜訊特性外，還含有空間參數(318)。
5. 如 4 之音訊解碼器，其中在 SID (348)中傳輸之參數(318)可由擴散度或到達方向或相干性組成。
6. 根據 4 之音訊解碼器，其中在 SID (348)中傳輸之空間參數(318)由頻率次頻帶傳輸。
7. 如 4 之音訊解碼器，其中在作用階段(346)期間及在 SID (348)中傳輸或編碼之空間參數(316、318)具有不同頻率解析度或時間解析度或量化解析度。
8. 如 1 之音訊解碼器，其中空間呈現器(220)可由以下各者構成：
 - a. 去相關器(730)，其用於獲得一或多個經解碼傳送/降混聲道(226)及/或經重建構背景雜訊(228)之去相關版本(228b)；
 - b. 升混器，其用於自一或多個經解碼傳送/降混聲道(226)或經重建構背景雜訊(228)及其去相關版本(228b)且自空間參數(348)導出輸出信號。
9. 如 8 之音訊解碼器，其中空間呈現器之升混器包括
 - a. 至少二個雜訊產生器(710、810)，其用於產生具有靜默描述符(448)中描述之特性及/或由應用於作用階段(346)中之雜訊估計給出的至少二個去相關背景雜訊(228、228a、228d)。
10. 如 9 之音訊解碼器，其中考慮到作用階段中傳輸之空間參數及/或包括於 SID 中之空間參數，升混器中之所產生去相關背景雜訊與經解碼傳送聲道或傳送聲道中之經重建構背景雜訊混合。
11. 如前述態樣中之一者的音訊解碼器，其中解碼裝置包含如 CELP 之語音

寫碼器或如 TCX 之通用音訊寫碼器或頻寬擴展模組。

圖式之其他表徵

【0159】 圖 1：來自[1]之 DirAC 分析及合成。

圖 2：低位元速率 3D 音訊寫碼器中之 DirAC 分析及合成的詳細方塊圖。

圖 3：解碼器之方塊圖。

圖 4：DirAC 模式中之音訊場景分析器的方塊圖。

圖 5：用於 MASA 輸入格式之音訊場景分析器之方塊圖。

圖 6：解碼器之方塊圖。

圖 7：空間呈現器(DirAC 合成)之方塊圖，其中傳送聲道中之 CNG 在呈現器外部。

圖 8：空間呈現器(DirAC 合成)之方塊圖，其中在呈現器之濾波器組域中針對 K 個聲道直接執行 CNG， $K \geq M$ 個傳送聲道。

圖 9：空間呈現器(DirAC 合成)之方塊圖，其中在空間呈現器外部及內部二者中執行 CNG。

圖 10：空間呈現器(DirAC 合成)之方塊圖，其中在空間呈現器外部及內部二者中執行且亦針對作用及非作用訊框二者接通 CNG。

優勢

【0160】 本發明之實施例允許以高效方式將 DTX 擴展至參數空間音訊寫碼。甚至對於非作用訊框，其可藉由高感知保真度對背景雜訊進行復原，對於此，可為節省通信頻寬而中斷傳輸。

【0161】 為此，傳送聲道之 SID 藉由與描述背景雜訊之空間影像相關的非作用空間參數擴展。所產生之舒適雜訊在藉由呈現器(DirAC 合成)空間化之前應用於傳送聲道中。替代地，對於品質之改良，CNG 可應用於比呈現內之傳送聲道更多的聲道。其允許複雜度降低且減少去相關器偽聲之煩擾。

其他態樣

【0162】 此處應提及，可個別地使用如之前所論述的所有替代方案或態樣及如在以下態樣中由獨立態樣定義之所有態樣，亦即，不具有除預期替代方案、物件或獨立態樣外的任何其他替代方案或物件。然而，在其他實施例中，替代方案或態樣或獨立態樣中之二者或更多者可彼此組合，且在其他實施例中，所有態樣或替代方案及所有獨立態樣可彼此組合。

【0163】 本發明之經編碼信號可儲存於數位儲存媒體或非暫時性儲存媒體上，或可在傳輸媒體，諸如無線傳輸媒體或諸如網際網路之有線傳輸媒體上傳輸。

【0164】 儘管已在設備之上下文中描述一些態樣，但顯然，此等態樣亦表示對應方法之描述，其中區塊或裝置對應於方法步驟或方法步驟之特徵。類似地，方法步驟之上下文中所描述的態樣亦表示對應設備之對應區塊或項目或特徵的描述。

【0165】 取決於某些實施要求，本發明之實施例可在硬體或軟體中實施。可使用其上儲存有與可程式電腦系統協作(或能夠協作)之電子可讀控制信號的數位儲存媒體，例如軟碟、DVD、CD、ROM、PROM、EPROM、EEPROM 或快閃記憶體執行實施方式，使得執行各別方法。

【0166】 根據本發明之一些實施例包含具有電子可讀控制信號之資料載體，該資料載體能夠與可程式電腦系統協作，使得執行本文中所描述之方法中的一者。

【0167】 大體而言，本發明之實施例可實施為具有程式碼之電腦程式產品，當電腦程式產品在電腦上運行時，程式碼操作性地用於執行方法中之一者。程式碼可例如儲存於機器可讀載體上。

【0168】 其他實施例包含用於執行本文中描述的方法中之一者之電腦程

式，該電腦程式儲存於機器可讀載體或非暫時性儲存媒體上。

【0169】換言之，因此，本發明方法之實施例為具有程式碼之電腦程式，當電腦程式在電腦上運行時，該程式碼用於執行本文中所描述之方法中的一者。

【0170】因此，本發明方法之另一實施例為資料載體(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的用於執行本文中所描述之方法中的一者的電腦程式。

【0171】因此，本發明方法之另一實施例為表示用於執行本文中所描述之方法中的一者之電腦程式之資料串流或信號序列。資料串流或信號序列可例如經組配以經由資料通信連接，例如經由網際網路而傳送。

【0172】另一實施例包含處理構件，例如經組配或經調適以執行本文中所描述之方法中的一者的電腦或可程式邏輯裝置。

【0173】另一實施例包含電腦，該電腦上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。

【0174】在一些實施例中，可程式邏輯裝置(例如，現場可程式閘陣列)可用以執行本文中所描述之方法的功能中之一些或全部。在一些實施例中，現場可程式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。一般而言，方法較佳地由任何硬體設備執行。

【0175】上述實施例僅說明本發明之原理。應理解，對本文中所描述之配置及細節的修改及變化將對熟習此項技術者顯而易見。因此，意圖為僅受到接下來之專利態樣之範疇限制，而不受到藉由本文中之實施例之描述及解釋所呈現的特定細節限制。

【0176】用於第一組實施例及第二組實施例之隨後定義之態樣可經組合，使得一組實施例之某些特徵可包括於另一組實施例中。

【0177】由前文討論可以識出，本發明係可以包含但不限於下列實例的多

種形式體現：

實例 1：一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景。

實例 2：如實例 1 之設備，其中該聲場參數產生器經組配以產生該第一聲場參數表示或該第二聲場參數表示，使得該第一聲場參數表示或該第二聲場參數表示包含指示該音訊信號相對於一聽者位置之一特性之一參數。

實例 3：如實例 1 或 2 之設備，其中該第一聲場參數表示或該第二聲場參數表示包含指示該第一訊框中相對於一聽者位置之聲音的一方向的一或多個方向參數，或指示該第一訊框中相對於一直接聲音之一擴散聲音之一部分的一或多個擴散度參數，或指示該第一訊框中一直接聲音與一擴散聲音之一能量比的一或多個能量比參數，或該第一訊框中之一聲道間/環繞相干性參數。

實例 4：如前述實例中任一者之設備，

其中該聲場參數產生器經組配以從該音訊信號之該第一訊框或該第二訊框判定多個個別聲源且針對每一聲源判定一參數描述。

實例 5：如實例 4 之設備，

其中該聲場產生器經組配以將該第一訊框或該第二訊框分解成多個頻率區間，每一頻率區間表示一個別聲源，且針對每一頻率區間判定至少一個聲場參數，該聲場參數例示性地包含一方向參數、一到達方向參數、一擴散度參數、一能量比參數或表示由該音訊信號之該第一訊框表示之聲場相對於一聽者位置之一特性的任何參數。

實例 6：如前述實例中任一者之設備，其中針對該第一訊框及該第二訊框之該音訊信號包含一輸入格式，該輸入格式具有表示相對於一聽者之一聲場的多個分量，

其中該聲場參數產生器經組配以例如使用該等多個分量之一降混來計算針對該第一訊框及該第二訊框之一或多個傳送聲道，且分析該輸入格式以判定與該一或多個傳送聲道相關之第一參數表示，或

其中該聲場參數產生器經組配以例如使用該等多個分量之一降混來計算一或多個傳送聲道，且

其中該活動偵測器經組配以分析自該第二訊框中之該音訊信號導出的該一或多個傳送聲道。

實例 7：如實例 1 至 5 中任一者之設備，

其中針對該第一訊框或該第二訊框之該音訊信號包含一輸入格式，對於該第一訊框及該第二訊框中之每一訊框，該輸入格式具有與每一訊框相關聯之一或多個傳送聲道及元資料，

其中該聲場參數產生器經組配以自該第一訊框及該第二訊框讀取該元資料，且將該第一訊框之該元資料用作或處理為該第一聲場參數表示且處理該第二訊框之該元資料以獲得該第二聲場參數表示，其中獲得該第二聲場參數表示之該處理使得傳輸該第二訊框之該元資料所需的資訊單元之一量相對於該處理

之前所需的一量有所減少。

實例 8：如實例 7 之設備，

其中該聲場參數產生器經組配以處理該第二訊框之該元資料以減少該元資料中之資訊項目之一數目或將該元資料中之該等資訊項目再取樣至一較低解析度，諸如一時間解析度或一頻率解析度，或相對於再量化之前的一情形將該第二訊框之該元資料的該等資訊單元再量化至一較粗略表示。

實例 9：如前述實例中任一者之設備，

其中該音訊信號編碼器經組配以將用於該非作用訊框之一靜默資訊描述判定為該參數描述，

其中該靜默資訊描述例示性地包含針對該第二訊框之諸如一能量、一功率或一響度的一振幅相關資訊及諸如一頻譜塑形資訊之一塑形資訊，或針對該第二訊框之諸如一能量、一功率或一響度之一振幅相關資訊及針對該第二訊框之線性預測寫碼 LPC 參數，或針對該第二訊框之具有一變化之相關聯頻率解析度的尺度參數，使得不同尺度參數係指具有不同寬度之頻帶。

實例 10：如前述實例中任一者之設備，

其中該音訊信號編碼器經組配以使用一時域或頻域編碼模式針對該第一訊框而編碼該音訊信號，該經編碼音訊信號包含例如經編碼時域樣本、經編碼頻譜域樣本、經編碼 LPC 域樣本及自該音訊信號之分量獲得或自一或多個傳送聲道獲得的旁側資訊，該一或多個傳送聲道係例如藉由一降混操作自該音訊信號之該等分量導出。

實例 11：如前述實例中任一者之設備，

其中該音訊信號包含一輸入格式，該輸入格式為一第一階立體混響 (Ambisonics) 格式、一高階立體混響格式、與諸如 5.1 或 7.1 或 7.1+4 之一給定揚聲器設置相關聯的一多聲道格式，或表示一或若干個不同音訊物件之一或多個

音訊聲道，該一或若干個不同音訊物件位於如由包括於相關聯元資料中之資訊所指示的一空間中，或為一元資料相關聯之空間音訊表示的一輸入格式，

其中該聲場參數產生器經組配以用於判定該第一聲場參數表示及第二聲場表示，使得該等參數表示相對於一所界定聽者位置之一聲場，或

其中該音訊信號包含如由真實麥克風或一虛擬麥克風拾取之一麥克風信號或例如呈一第一階立體混響格式或一高階立體混響格式之一合成產生麥克風信號。

實例 12：如前述實例中任一者之設備，

其中該活動偵測器經組配以用於偵測該第二訊框及該第二訊框之後的一或多個訊框上之一不活動階段，且

其中該音訊信號編碼器經組配以僅針對一另一第三訊框而產生針對一非作用訊框之一另一參數描述，就一訊框之時間序列而言，該另一第三訊框與該第二訊框相隔至少一個訊框，且

其中該聲場參數產生器經組配以用於僅針對一訊框而判定一另一聲場參數表示，該音訊信號編碼器已針對該訊框判定一參數描述，或

其中該活動偵測器經組配以用於判定包含該第二訊框及該第二訊框之後的八個訊框的一非作用階段，且其中該音訊信號編碼器經組配以用於僅在每第八個訊框處產生針對一非作用訊框之一參數描述，且其中該聲場參數產生器經組配以用於針對每一第八個非作用訊框而產生一聲場參數表示，或

其中該聲場參數產生器經組配以用於針對每一非作用訊框而產生一聲場參數表示，甚至在該音訊信號編碼器不產生針對一非作用訊框之一參數描述時亦如此，或

其中該聲場參數產生器經組配以用於判定相比該音訊信號編碼器產生用於一或多個非作用訊框之該參數描述而具有一較高訊框率的一參數表示。

實例 13：如前述實例中任一者之設備，

其中該聲場參數產生器經組配以用於使用頻帶中之一或多個方向的空間參數及對應於一個方向分量與一總能量之一比率的頻帶中之相關聯能量比來判定該第二訊框之該第二聲場參數表示，或

判定指示擴散聲音或直接聲音之一比率的一擴散度參數，或

使用與該第一訊框中之一量化相比較粗略之一量化方案判定一方向資訊，
或

使用一方向隨時間或頻率之一平均值以獲得一較粗略時間或頻率解析度，
或

判定針對一或多個非作用訊框之一聲場參數表示，該一或多個非作用訊框具有與針對一作用訊框之該第一聲場參數表示中相同的頻率解析度，且相對於針對該非作用訊框之該聲場參數表示中之一方向資訊，具有低於針對作用訊框之時間發生率之一時間發生率，或

判定具有一擴散度參數之該第二聲場參數表示，其中該擴散度參數係以與作用訊框相同之時間或頻率解析度但以一較粗略量化傳輸，或

用第一數目個位元量化針對該第二聲場表示之一擴散度參數，且其中僅傳輸每一量化索引之第二數目個位元，位元之該第二數目小於位元之該第一數目，或

若該音訊信號具有對應於定位於一空間域中之聲道的輸入聲道，則針對該第二聲場參數表示而判定一聲道間相干性，或若該音訊信號具有對應於定位於該空間域中之聲道的輸入聲道，則判定聲道間聲級差，或

判定一環繞聲相干性，其經界定為由該音訊信號表示之一聲場中相干的擴散能量之一比率。

實例 14：一種用於處理一經編碼音訊場景之設備，該經編碼音訊場景在一

第一訊框中包含一第一聲場參數表示及一經編碼音訊信號，其中一第二訊框為一非作用訊框，該設備包含：

- 一活動偵測器，其用於偵測該第二訊框為該非作用訊框；

- 一合成信號合成器，其用於使用針對該第二訊框之參數描述來合成針對該第二訊框之一合成音訊信號；

- 一音訊解碼器，其用於解碼針對該第一訊框之該經編碼音訊信號；以及

- 一空間呈現器，其用於使用該第一聲場參數表示且使用針對該第二訊框之該合成音訊信號在空間上呈現針對該第一訊框之該音訊信號，或

- 一轉碼器，其用於產生一元資料輔助輸出格式，該元資料輔助輸出格式包含針對該第一訊框之該音訊信號、針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該合成音訊信號及針對該第二訊框之一第二聲場參數表示。

實例 15：如實例 14 之設備，其中該經編碼音訊場景包含針對該第二訊框之一第二聲場參數描述，且其中該設備包含用於自該第二聲場參數表示導出一或多個聲場參數之一聲場參數處理器，且其中該空間呈現器經組配以將針對該第二訊框之該一或多個聲場參數用於該第二訊框之該合成音訊信號之呈現。

實例 16：如實例 14 之設備，其包含用於導出針對該第二訊框之一或多個聲場參數的一參數處理器，

- 其中該參數處理器經組配以儲存針對該第一訊框之該聲場參數表示且使用針對該第一訊框之所儲存之第一聲場參數表示來合成針對該第二訊框之一或多個聲場參數，其中該第二訊框在時間上在該第一訊框之後，或

- 其中該參數處理器經組配以儲存在時間上出現於該第二訊框之前或在時間上出現於該第二訊框之後的若干訊框之一或多個聲場參數表示，以使用針對若干訊框之該一或多個聲場參數表示中的至少二個聲場參數表示進行外推或內插，以判定針對該第二訊框之該一或多個聲場參數，且

其中該空間呈現器經組配以將針對該第二訊框之該一或多個聲場參數用於該第二訊框之該合成音訊信號之該呈現。

實例 17：如實例 16 之設備，

其中該參數處理器經組配以在進行外推或內插以判定針對該第二訊框之該一或多個聲場參數時，使用在時間上出現於該第二訊框之前或之後的該至少二個聲場參數表示中所包括的方向執行一抖動。

實例 18：如實例 14 至 17 中任一者之設備，

其中該經編碼音訊場景包含針對該第一訊框之一或多個傳送聲道，

其中合成信號產生器經組配以產生針對該第二訊框之一或多個傳送聲道作為該合成音訊信號，且

其中該空間呈現器經組配以在空間上呈現針對該第二訊框之該一或多個傳送聲道。

實例 19：如實例 14 至 18 中任一者之設備，

其中該合成信號產生器經組配以針對該第二訊框而產生針對與該空間呈現器之一音訊輸出格式相關的個別分量之多個合成分量音訊信號作為該合成音訊信號。

實例 20：如實例 19 之設備，其中該合成信號產生器經組配以至少針對與該音訊輸出格式相關之至少二個個別分量之一子集中的每一者而產生一個別合成分量音訊信號，

其中一第一個別合成分量音訊信號與一第二個別合成分量音訊信號去相關，且

其中該空間呈現器經組配以使用該第一個別合成分量音訊信號與該第二個別合成分量音訊信號之一組合來呈現該音訊輸出格式之一分量。

實例 21：如實例 20 之設備，

其中該空間呈現器經組配以應用一協方差法。

實例 22：如實例 21 之設備，

其中該空間呈現器經組配以不使用任何去相關器處理或控制一去相關器處理，使得僅使用藉由如由該協方差法指示之該去相關器處理產生的一定量之去相關信號來產生該音訊輸出格式之一分量。

實例 23：如實例 14 至 22 中任一者之設備，其中該合成信號產生器為一舒適雜訊產生器。

實例 24：如實例 20 至 23 中任一者之設備，其中該合成信號產生器包含一雜訊產生器且該第一個別合成分量音訊信號係由該雜訊產生器之一第一取樣產生，且該第二個別合成分量音訊信號係由該雜訊產生器之一第二取樣產生，其中該第二取樣不同於該第一取樣。

實例 25：如實例 24 之設備，其中該雜訊產生器包含一雜訊表，且其中該第一個別合成分量音訊信號係藉由取該雜訊表之一第一部分而產生，且其中該第二個別合成分量音訊信號係藉由取該雜訊表之一第二部分而產生，其中該雜訊表之該第二部分不同於該雜訊表之該第一部分，或

其中該雜訊產生器包含一偽雜訊產生器，且其中該第一個別合成分量音訊信號係藉由使用該偽雜訊產生器之一第一種子而產生，且其中該第二個別合成分量音訊信號係使用該偽雜訊產生器之一第二種子而產生。

實例 26：如實例 14 至 25 中任一者之設備，

其中該經編碼音訊場景包含針對該第一訊框之二個或更多個傳送聲道，且

其中該合成信號產生器包含一雜訊產生器且經組配以使用針對該第二訊框之該參數描述，藉由對該雜訊產生器進行取樣而產生一第一傳送聲道及藉由對該雜訊產生器進行取樣而產生一第二傳送聲道，其中如藉由對該雜訊產生器進行取樣而判定之該第一傳送聲道及該第二傳送聲道係使用針對該第二訊框之相

同參數描述進行加權。

實例 27：如實例 14 至 26 中任一者之設備，其中該空間呈現器經組配以使用一直接信號與由一去相關器在該第一聲場參數表示之一控制下自該直接信號產生之一擴散信號的一混合，在針對該第一訊框之一第一模式下操作，且

使用一第一合成分量信號與第二合成分量信號之一混合，在針對該第二訊框之一第二模式下操作，其中該第一合成分量信號及該第二合成分量信號係由該合成信號合成器藉由一雜訊處理或一偽雜訊處理之不同實現而產生。

實例 28：如實例 27 之設備，其中該空間呈現器經組配以透過藉由一參數處理器針對該第二訊框導出的一擴散度參數、一能量分佈參數或一相干性參數而控制該第二模式下之該混合。

實例 29：如實例 14 至 28 中任一者之設備，

其中該合成信號產生器經組配以使用針對該第二訊框之該參數描述來產生針對該第一訊框之一合成音訊信號，且

其中該空間呈現器經組配以在空間呈現之前或之後執行針對該第一訊框之該音訊信號與針對該第一訊框之該合成音訊信號的一加權組合，其中在該加權組合中，針對該第一訊框之該合成音訊信號的一強度相對於針對該第二訊框之該合成音訊信號的一強度有所減小。

實例 30：如實例 14 至 29 中任一者之設備，

其中一參數處理器經組配以針對第二非作用訊框而判定一環繞聲相干性，該環繞聲相干性經界定為由該第二訊框表示之一聲場中相干的擴散能量之一比率，其中該空間呈現器經組配以用於基於聲音相干性重分佈該第二訊框中之直接信號與擴散信號之間的一能量，其中自待重分佈至方向分量之擴散能量移除聲音環繞相干分量之一能量，且其中在一再現空間中平移該等方向分量。

實例 31：如實例 14 至 18 中任一者之設備，其進一步包含一輸出介面，該輸出介面用於將由該空間呈現器產生之一音訊輸出格式轉換成一經轉碼輸出格式，諸如包含專用於待置放於預定位置處之揚聲器的數個輸出聲道的一輸出格式，或包含 FOA 或 HOA 資料之一經轉碼輸出格式，或

其中，替代該空間呈現器，提供該轉碼器以用於產生該元資料輔助輸出格式，該元資料輔助輸出格式包含針對該第一訊框之該音訊信號、針對該第一訊框之第一聲場參數及針對該第二訊框之該合成音訊信號及針對該第二訊框之一第二聲場參數表示。

實例 32：如實例 14 至 31 中任一者之設備，其中該活動偵測器經組配以用於偵測該第二訊框為該非作用訊框。

實例 33：一種自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的方法，其包含：

從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景。

實例 34：一種處理一經編碼音訊場景之方法，該經編碼音訊場景在一第一訊框中包含一第一聲場參數表示及一經編碼音訊信號，其中一第二訊框為一非作用訊框，該方法包含：

偵測該第二訊框為該非作用訊框；

使用針對該第二訊框之參數描述來合成針對該第二訊框之一合成音訊信號；

解碼針對該第一訊框之該經編碼音訊信號；以及

使用該第一聲場參數表示且使用針對該第二訊框之該合成音訊信號在空間上呈現針對該第一訊框之該音訊信號，或產生一元資料輔助輸出格式，該元資料輔助輸出格式包含針對該第一訊框之該音訊信號、針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該合成音訊信號及針對該第二訊框之一第二聲場參數表示。

實例 35：如實例 33 之方法，其進一步包含提供針對該第二訊框之一參數描述。

實例 36：一種經編碼音訊場景，其包含：

針對一第一訊框之一第一聲場參數表示；

針對一第二訊框之一第二聲場參數表示；

針對該第一訊框之一經編碼音訊信號；以及

針對該第二訊框之一參數描述。

實例 37：一種電腦程式，其用於在一電腦或處理器上運行時執行如實例 33 或實例 34 之方法。

【符號說明】

【0178】

200: 解碼器/音訊解碼器設備

202: 輸出聲道/最終輸出信號/輸出空間格式/空間音訊輸出信號/輸入信號/空間音訊輸入格式/音訊輸出格式/合成音訊信號

210: 合成信號合成器/合成信號產生器/合成音訊合成器/第一部分/經恢復空間參

數/舒適雜訊產生器/靜默插入描述符解碼器

211: 空間元資料解碼

218,319: 空間參數

219: 所產生參數/經恢復空間參數/非發信非作用空間參數/聲場參數

220: 空間呈現器/空間呈現裝置

221: VAD 資訊

221': 命令/控制

222',275': 切換器

224': 偏差器/切換器

224: 經解碼音訊信號/傳送/降混聲道/降混信號

226: 經解碼聲道/經解碼音訊場景/降混信號/經解碼信號/傳送聲道 SID/傳送/降混聲道

228: 合成音訊信號/非作用訊框/降混信號/舒適雜訊/經重建構背景雜訊/去相關背景雜訊/傳送/降混聲道

228a: 去相關聲道/合成分量音訊信號/去相關信號/去相關分量/舒適雜訊/去相關器輸出/去相關背景雜訊

228b: 合成分量音訊信號/分量/去相關聲道/輸出/去相關版本

228c: 輸出

228d: 輸出/所產生分量/雜訊聲道/合成音訊信號/舒適雜訊/隨機雜訊/去相關背景雜訊

228e: 經加總信號/分量

228f: 輸出/信號/升混版本

230: 音訊解碼器/解碼裝置

231: EVS 解碼器

- 240: 空間呈現器
- 275,1075: 參數處理器
- 276: 作用空間參數解碼器
- 278: 非作用空間參數解碼器
- 279,744: 區塊
- 300: 編碼器/音訊編碼器設備
- 302: 輸入格式/輸入音訊信號/輸入版本/原始音訊輸入信號/輸入 MASA 信號/輸入音訊場景/B 格式輸入信號/空間音訊輸入信號/多聲道信號
- 304: 經編碼音訊場景/位元串流/參數表示
- 306: 作用訊框/第一訊框/輸入信號/作用階段
- 308: 非作用訊框/第二訊框/非作用階段
- 310: 音訊場景分析器/音訊信號分析器/DirAC 分析區塊/聲場參數產生器/場景音訊分析器
- 314: 聲場參數/作用空間參數
- 314a: 擴散度參數/擴散度資訊/擴散度/DirAC 參數
- 314b: 輸出/方向資訊/到達方向/參數/DirAC 參數
- 316: 作用空間參數/第一聲場參數表示/第一聲場參數/低位元速率參數表示/DirAC 參數
- 318: 非作用空間參數/第二聲場參數表示/第二聲場參數/低位元速率參數表示/英特爾參數/DirAC 參數
- 320: 選擇器/區塊/話音活動偵測器
- 321: 控制件
- 322: 第一偏差器
- 322a: 第二偏差器

- 324: 傳送聲道版本/傳送聲道/降混版本/音訊信號/降混信號/聲道信號
- 326: 傳送聲道/第一訊框/經編碼音訊信號/聲道信號/經寫碼音訊位元串流/傳送聲道 SID/降混版本/降混信號/作用訊框
- 328: 傳送聲道/第二訊框/經編碼音訊信號/聲道信號/降混信號/經寫碼空間參數/參數描述/非作用訊框/傳送聲道 SID/靜默資訊描述
- 330: 音訊信號編碼器
- 340: 傳送聲道編碼器/區塊/傳送聲道編碼器裝置
- 344: 經編碼音訊信號/編碼器音訊信號/傳送聲道/經編碼版本/經編碼資料
- 346: 經編碼音訊信號/第一訊框/作用訊框/經編碼版本/經編碼資料/作用階段
- 348: 參數描述/第二訊框/經編碼音訊信號/非作用訊框/經編碼訊框/經編碼版本/靜默參數描述/傳送聲道 SID/非作用階段/靜默插入描述/靜默插入描述符訊框
- 350: 傳送聲道 SI 描述符/區塊
- 370: 經編碼信號形成器/多工器
- 390: 濾波器組分析/濾波器組分析區塊
- 390M: MASA 讀取器
- 391: 輸出/頻率區間/頻域資訊
- 392a: 擴散度估計區塊/擴散度分析區塊/階段
- 392b: 方向估計區塊/方向分析區塊/階段
- 396: 作用空間元資料編碼器
- 398: 非作用空間元資料編碼器
- 448: 靜默描述符
- 700,800,900,1000: 解碼器/解碼器設備
- 710: 傳送聲道/第一外部部分/合成信號合成器第一部分/合成信號合成器/CNG 第二部分/舒適雜訊產生器/合成信號產生器

720: 濾波器組分析/濾波器組分析區塊/反饋分析區塊

724: 濾波器組分析

730: 相關處理/去相關器/相關器處理/去相關器處理

740: 混合區塊/混合/空間呈現器

742: 混合信號

746: 濾波器組合成區塊

750: 升混加法區塊

810: 合成信號合成器/第二內部部分/合成信號合成器第二部分/傳送聲道 CNG 部分/舒適雜訊產生/雜訊產生器

920: 加法器/加法區塊

2200: 活動偵測器

【發明申請專利範圍】

【請求項1】 一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中該第一聲場參數表示或該第二聲場參數表示包含指示該第一訊框中相對於一聽者位置之聲音的一方向的一或多個方向參數，或指示該第一訊框中相對於一直接聲音之一擴散聲音之一部分的一或多個擴散度參數，或指示該第一訊框中一直接聲音與一擴散聲音之一能量比的一或多個能量比參數，或該第一訊框中之一聲道間/環繞相干性參數。

【請求項2】 如請求項 1 之設備，其中該聲場參數產生器經組配以產生該第一聲場參數表示或該第二聲場參數表示，使得該第一聲場參數表示或該第二聲場參數表示包含指示該音訊信號相對於一聽者位置之一特性之一參數。

【請求項3】 一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一

訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中該聲場參數產生器經組配以從該音訊信號之該第一訊框或該第二訊框判定多個個別聲源且針對每一聲源判定一參數描述。

【請求項4】 一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中該聲場產生器經組配以將該第一訊框或該第二訊框分解成多個頻率區

間，且針對每一頻率區間判定至少一個聲場參數，該聲場參數例示性地包含一方向參數、一到達方向參數、一擴散度參數、一能量比參數或表示由該音訊信號之該第一訊框表示之聲場相對於一聽者位置之一特性的任何參數。

【請求項5】 一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中該聲場參數產生器經組配以從該音訊信號之該第一訊框或該第二訊框判定多個個別聲源且針對每一聲源判定一參數描述，其中該聲場產生器經組配以將該第一訊框或該第二訊框分解成多個頻率區間，每一頻率區間表示一個別聲源，且針對每一頻率區間判定至少一個聲場參數，該聲場參數例示性地包含一方向參數、一到達方向參數、一擴散度參數、一能量比參數或表示由該音訊信號之該第一訊框表示之聲場相對於一聽者位置之一特性的任何參數。

【請求項6】 一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一

訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中針對該第一訊框及該第二訊框之該音訊信號包含一輸入格式，該輸入格式具有表示相對於一聽者之一聲場的多個分量，

其中該聲場參數產生器經組配以例如使用該等多個分量之一降混來計算針對該第一訊框及該第二訊框之一或多個傳送聲道，且分析該輸入格式以判定與該一或多個傳送聲道相關之第一參數表示，或

其中該聲場參數產生器經組配以例如使用該等多個分量之一降混來計算一或多個傳送聲道，且

其中該活動偵測器經組配以分析自該第二訊框中之該音訊信號導出的該一或多個傳送聲道。

【請求項7】 一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一

訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中針對該第一訊框或該第二訊框之該音訊信號包含一輸入格式，對於該第一訊框及該第二訊框中之每一訊框，該輸入格式具有與每一訊框相關聯之一或多個傳送聲道及元資料，

其中該聲場參數產生器經組配以自該第一訊框及該第二訊框讀取該元資料，且將該第一訊框之該元資料用作或處理為該第一聲場參數表示且處理該第二訊框之該元資料以獲得該第二聲場參數表示，其中獲得該第二聲場參數表示之該處理使得傳輸該第二訊框之該元資料所需的資訊單元之一量相對於該處理之前所需的一量有所減少。

【請求項8】 如請求項 7 之設備，

其中該聲場參數產生器經組配以處理該第二訊框之該元資料以減少該元資料中之資訊項目之一數目或將該元資料中之該等資訊項目再取樣至一較低解析度，諸如一時間解析度或一頻率解析度，或相對於再量化之前的一情形將該第二訊框之該元資料的該等資訊單元再量化至一較粗略表示。

【請求項9】 一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中該音訊信號編碼器經組配以將用於該非作用訊框之一靜默資訊描述判定為該參數描述，

其中該靜默資訊描述例示性地包含針對該第二訊框之諸如一能量、一功率或一響度之一振幅相關資訊及諸如一頻譜塑形資訊之一塑形資訊，或針對該第二訊框之諸如一能量、一功率或一響度之一振幅相關資訊及針對該第二訊框之線性預測寫碼 LPC 參數，或針對該第二訊框之具有一變化之相關聯頻率解析度的尺度參數，使得不同尺度參數係指具有不同寬度之頻帶。

【請求項10】一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表

示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中該音訊信號編碼器經組配以使用一時域或頻域編碼模式針對該第一訊框而編碼該音訊信號，該經編碼音訊信號包含例如經編碼時域樣本、經編碼頻譜域樣本、經編碼 LPC 域樣本及自該音訊信號之分量獲得或自一或多個傳送聲道獲得的旁側資訊，該一或多個傳送聲道係例如藉由一降混操作自該音訊信號之該等分量導出。

【請求項11】一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中該音訊信號包含一輸入格式，該輸入格式為一第一階立體混響(Ambisonics)格式、一高階立體混響格式、與諸如 5.1 或 7.1 或 7.1+4 之一給定揚聲器設置相關聯的一多聲道格式，或表示一或若干個不同音訊物件之一或多個音訊聲道，該一或若干個不同音訊物件位於如由包括於相關聯元資料中之資訊所指示的一空間中，或為一元資料相關聯之空間音訊表示的一輸入格式，

其中該聲場參數產生器經組配以用於判定該第一聲場參數表示及第二聲場表示，使得該等參數表示相對於一所界定聽者位置之一聲場，或

其中該音訊信號包含如由真實麥克風或一虛擬麥克風拾取之一麥克風信號或例如呈一第一階立體混響格式或一高階立體混響格式之一合成產生麥克風信號。

【請求項12】一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中該活動偵測器經組配以用於偵測該第二訊框及該第二訊框之後的一或多個訊框上之一不活動階段，且

其中該音訊信號編碼器經組配以僅針對一另一第三訊框而產生針對一非作用訊框之一另一參數描述，就一訊框之時間序列而言，該另一第三訊框與該第二訊框相隔至少一個訊框，且

其中該聲場參數產生器經組配以用於僅針對一訊框而判定一另一聲場參數表示，該音訊信號編碼器已針對該訊框判定一參數描述，或

其中該活動偵測器經組配以用於判定包含該第二訊框及該第二訊框之後的八個訊框的一非作用階段，且其中該音訊信號編碼器經組配以用於僅在每第八個訊框處產生針對一非作用訊框之一參數描述，且其中該聲場參數產生器經組配以用於針對每一第八個非作用訊框而產生一聲場參數表示，或

其中該聲場參數產生器經組配以用於針對每一非作用訊框而產生一聲場參數表示，甚至在該音訊信號編碼器不產生針對一非作用訊框之一參數描述時亦如此，或

其中該聲場參數產生器經組配以用於判定相比該音訊信號編碼器產生用於一或多個非作用訊框之該參數描述而具有一較高訊框率的一參數表示。

【請求項13】一種用於自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的設備，其包含：

一聲場參數產生器，其用於從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

一活動偵測器，其用於分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

一音訊信號編碼器，其用於產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

一經編碼信號形成器，其用於藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

其中該聲場參數產生器經組配以用於判定該第二訊框之該第二聲場參數表示，此判定作業進一步係

藉由使用頻帶中之一或多個方向的空間參數及對應於一個方向分量與

一總能量之一比率的頻帶中之相關聯能量比，或

要判定指示擴散聲音或直接聲音之一比率的一擴散度參數，或

要使用與該第一訊框中之一量化相比較粗略之一量化方案判定一方向資訊，或

藉由使用一方向隨時間或頻率之一平均值以獲得一較粗略時間或頻率解析度，或

要判定針對一或多個非作用訊框之一聲場參數表示，該一或多個非作用訊框具有與針對一作用訊框之該第一聲場參數表示中相同的頻率解析度，且相對於針對該非作用訊框之該聲場參數表示中之一方向資訊，具有低於針對作用訊框之時間發生率之一時間發生率，或

要判定具有一擴散度參數之該第二聲場參數表示，其中該擴散度參數係以與作用訊框相同之時間或頻率解析度但以一較粗略量化傳輸，或

要用第一數目個位元量化針對該第二聲場表示之一擴散度參數，且其中僅傳輸每一量化索引之第二數目個位元，位元之該第二數目小於位元之該第一數目，或

要在該音訊信號具有對應於定位於一空間域中之聲道的輸入聲道時則針對該第二聲場參數表示而判定一聲道間相干性，或在該音訊信號具有對應於定位於該空間域中之聲道的輸入聲道時則判定聲道間聲級差，或

要判定一環繞聲相干性，其經界定為由該音訊信號表示之一聲場中相干的擴散能量之一比率。

【請求項14】一種自具有一第一訊框及一第二訊框之一音訊信號產生一經編碼音訊場景的方法，其包含：

從該第一訊框中之該音訊信號判定針對該第一訊框之一第一聲場參數表示且從該第二訊框中之該音訊信號判定針對該第二訊框之一第二聲場參數表示；

分析該音訊信號以取決於該音訊信號而判定該第一訊框為一作用訊框且該第二訊框為一非作用訊框；

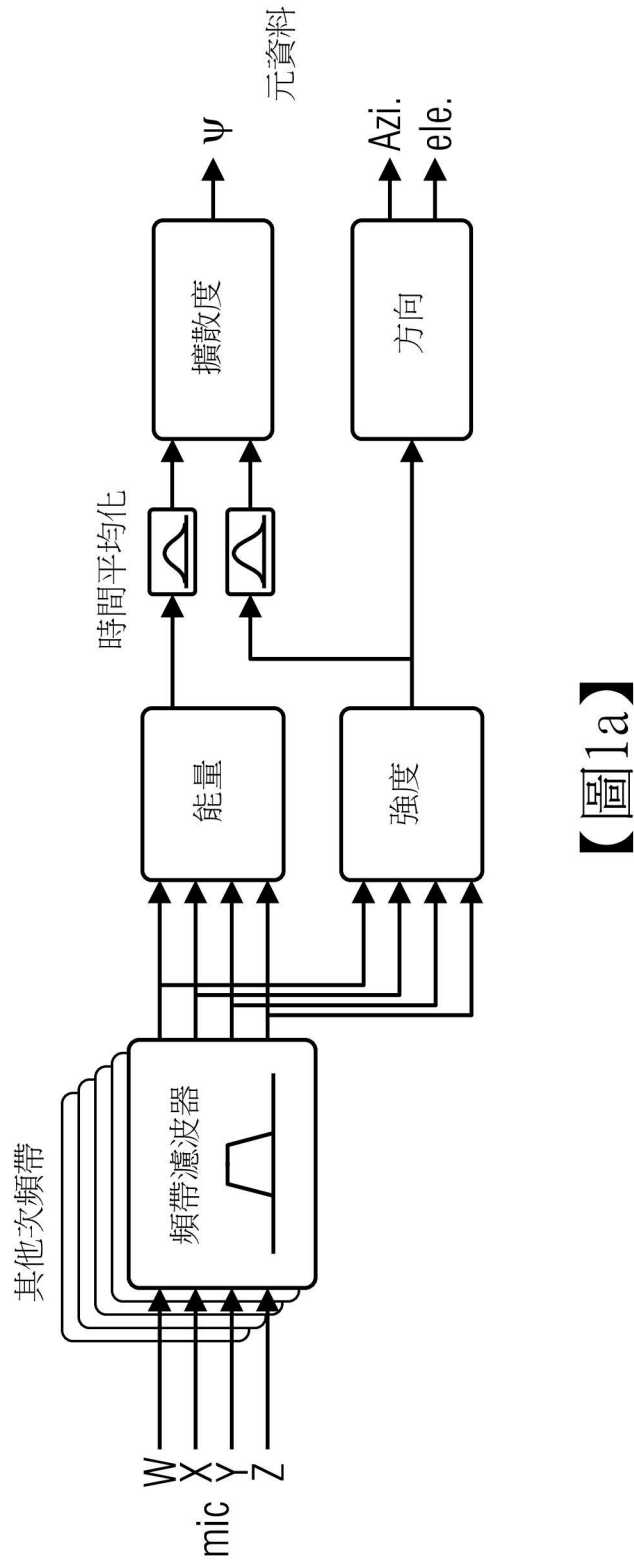
產生針對為該作用訊框之該第一訊框之一經編碼音訊信號且產生針對為該非作用訊框之該第二訊框之一參數描述；以及

藉由將針對該第一訊框之該第一聲場參數表示、針對該第二訊框之該第二聲場參數表示、針對該第一訊框之該經編碼音訊信號及針對該第二訊框之該參數描述組合在一起而構成該經編碼音訊場景，

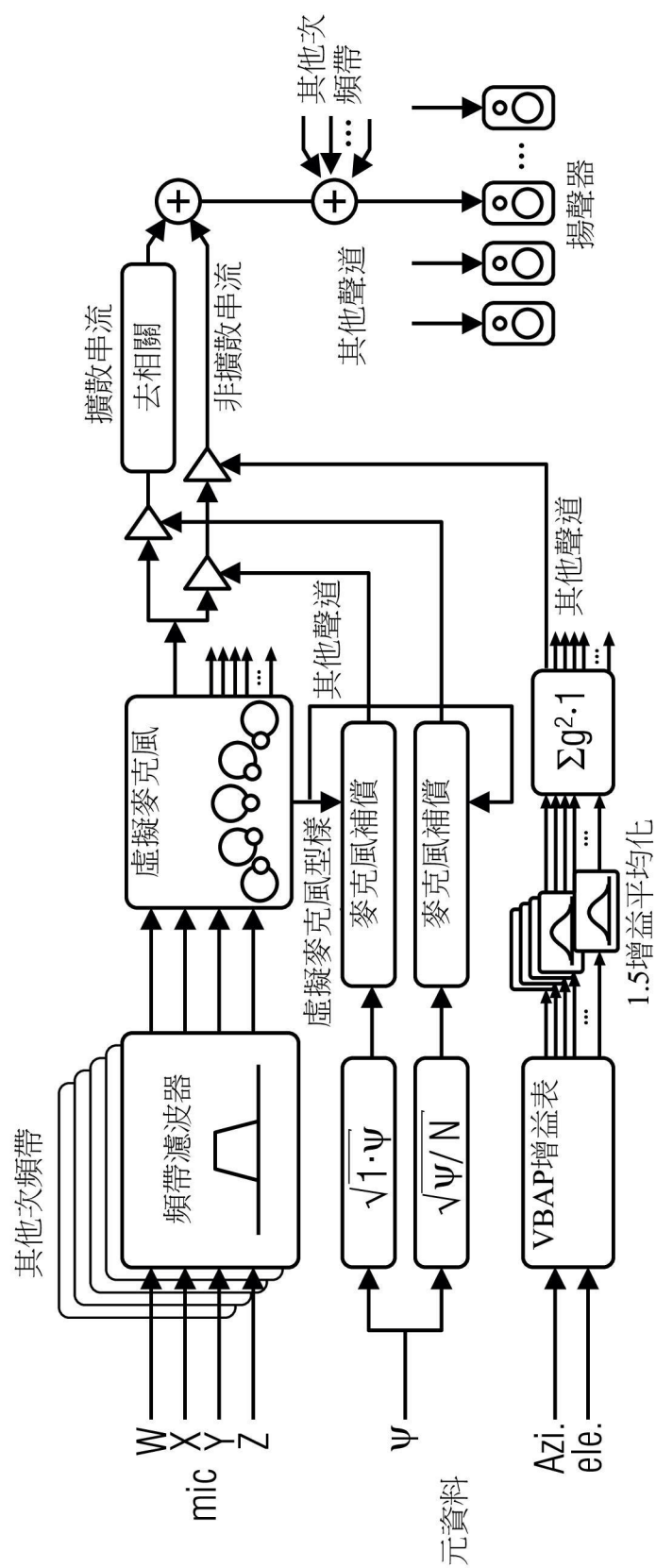
其中該第一聲場參數表示或該第二聲場參數表示包含指示該第一訊框中相對於一聽者位置之聲音的一方向的一或多個方向參數，或指示該第一訊框中相對於一直接聲音之一擴散聲音之一部分的一或多個擴散度參數，或指示該第一訊框中一直接聲音與一擴散聲音之一能量比的一或多個能量比參數，或該第一訊框中之一聲道間/環繞相干性參數。

【請求項15】一種電腦程式，其用於在一電腦或處理器上運行時執行如請求項 14 之方法。

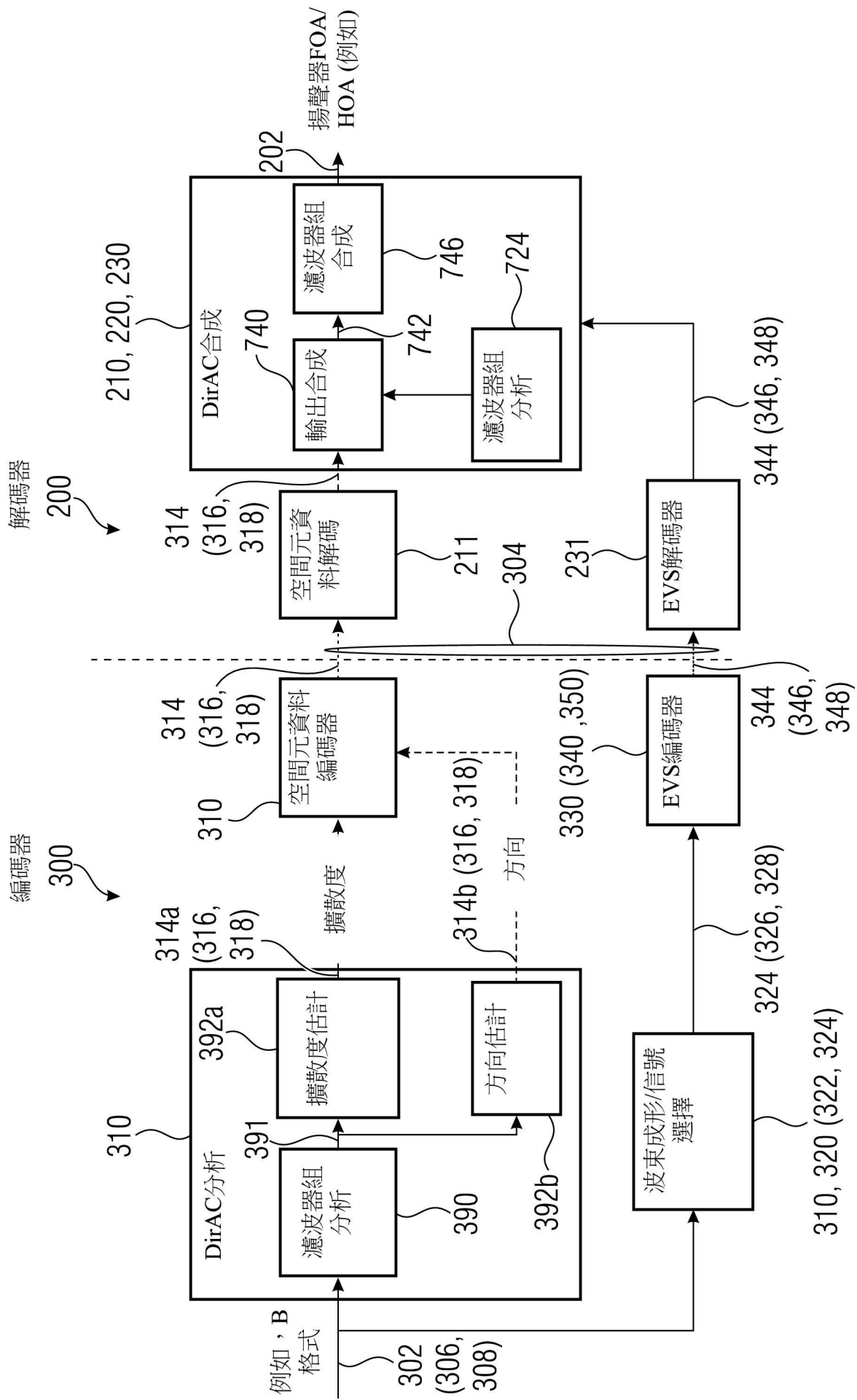
【發明圖式】



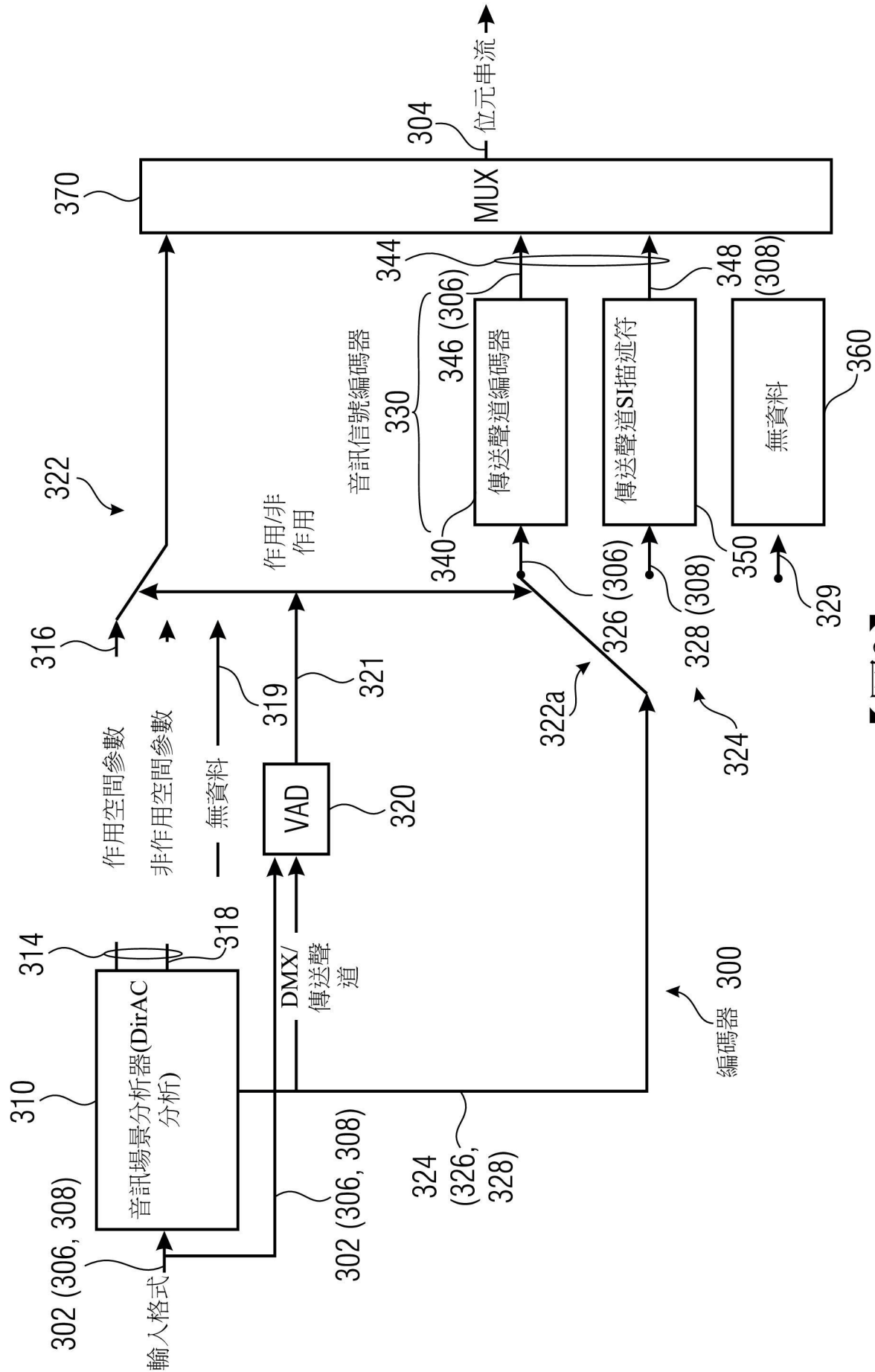
【圖1a】



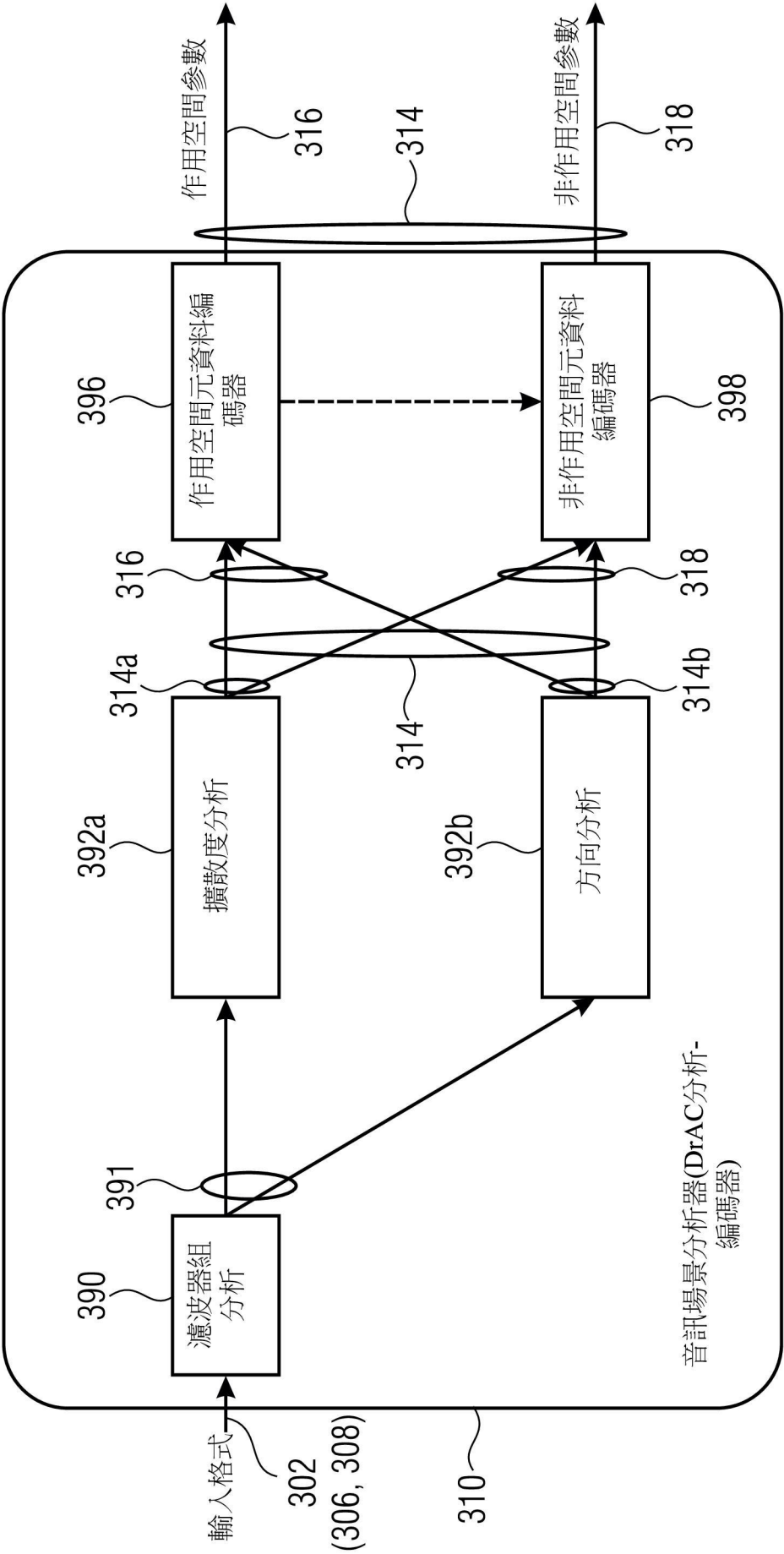
【圖 16】



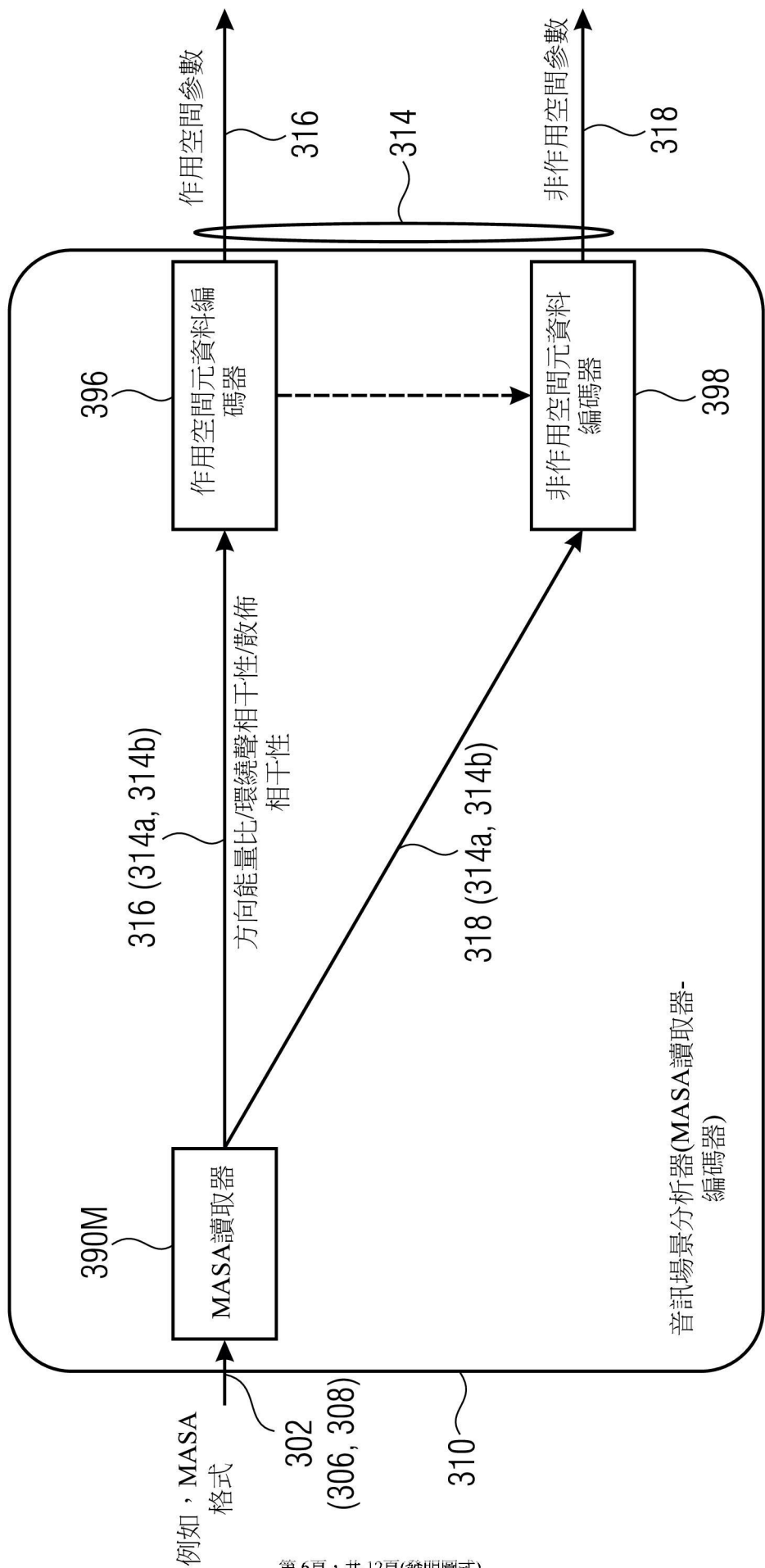
【圖2】



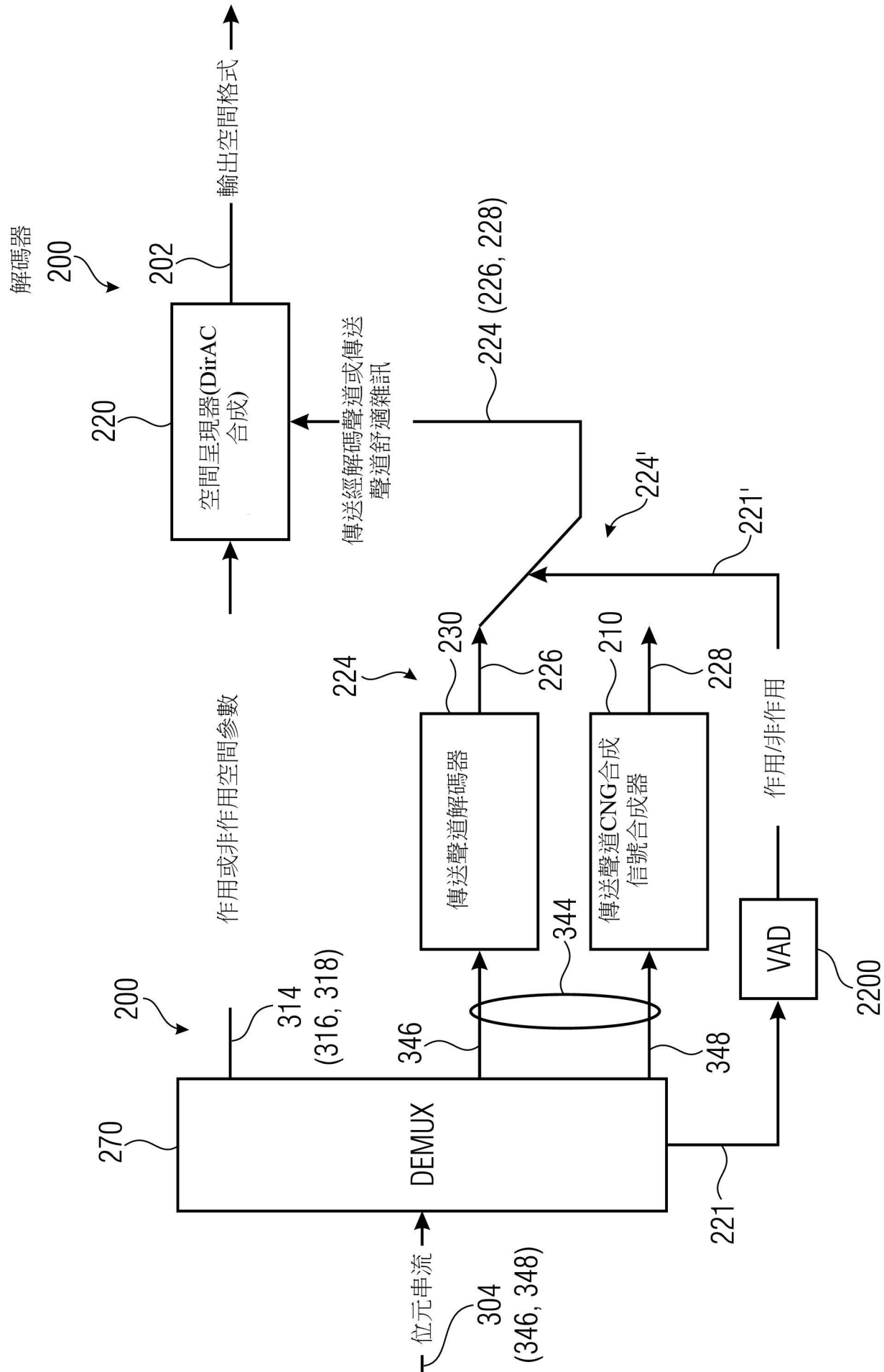
【圖3】



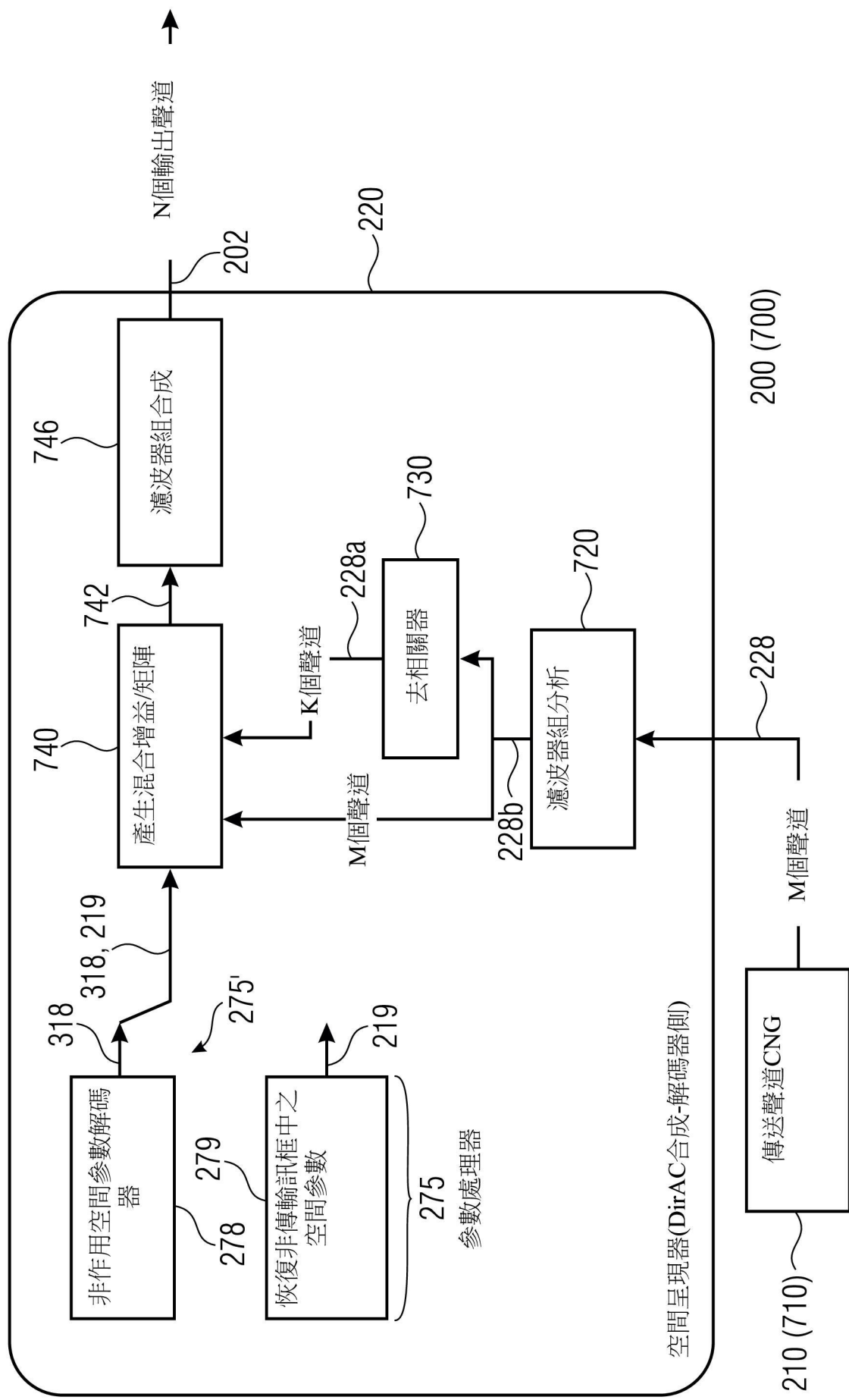
【圖4】



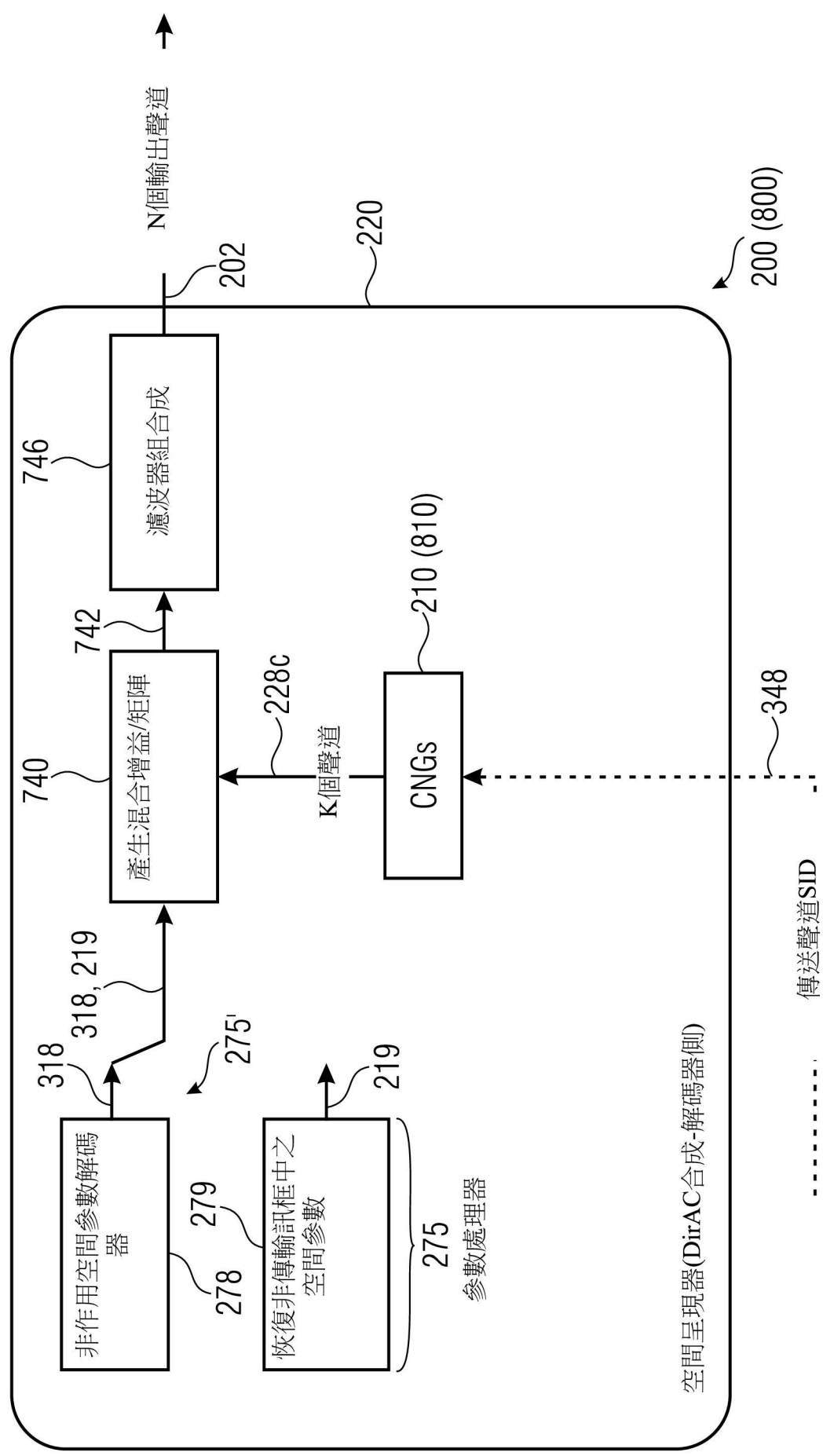
【圖5】



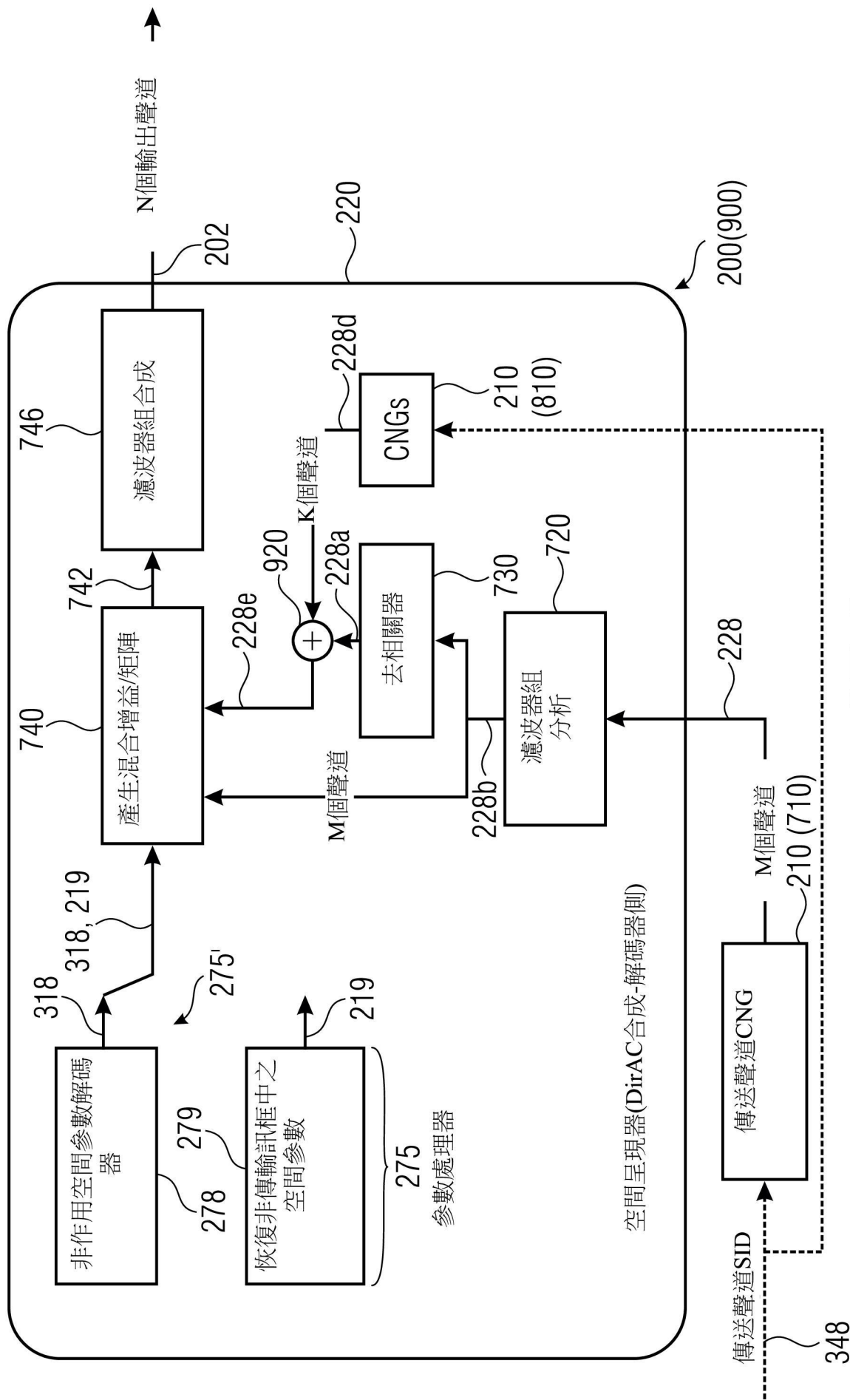
【圖6】



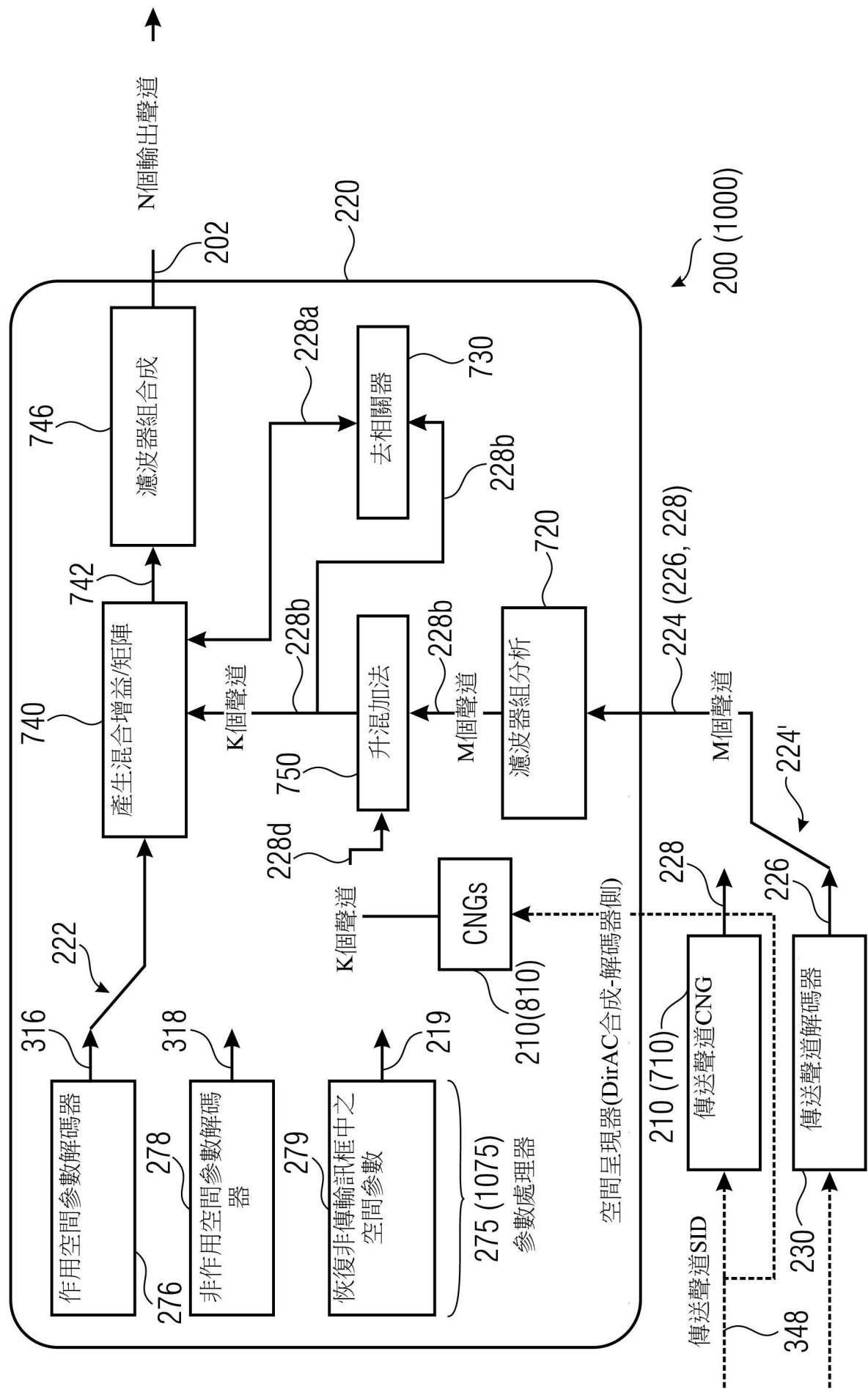
【圖7】



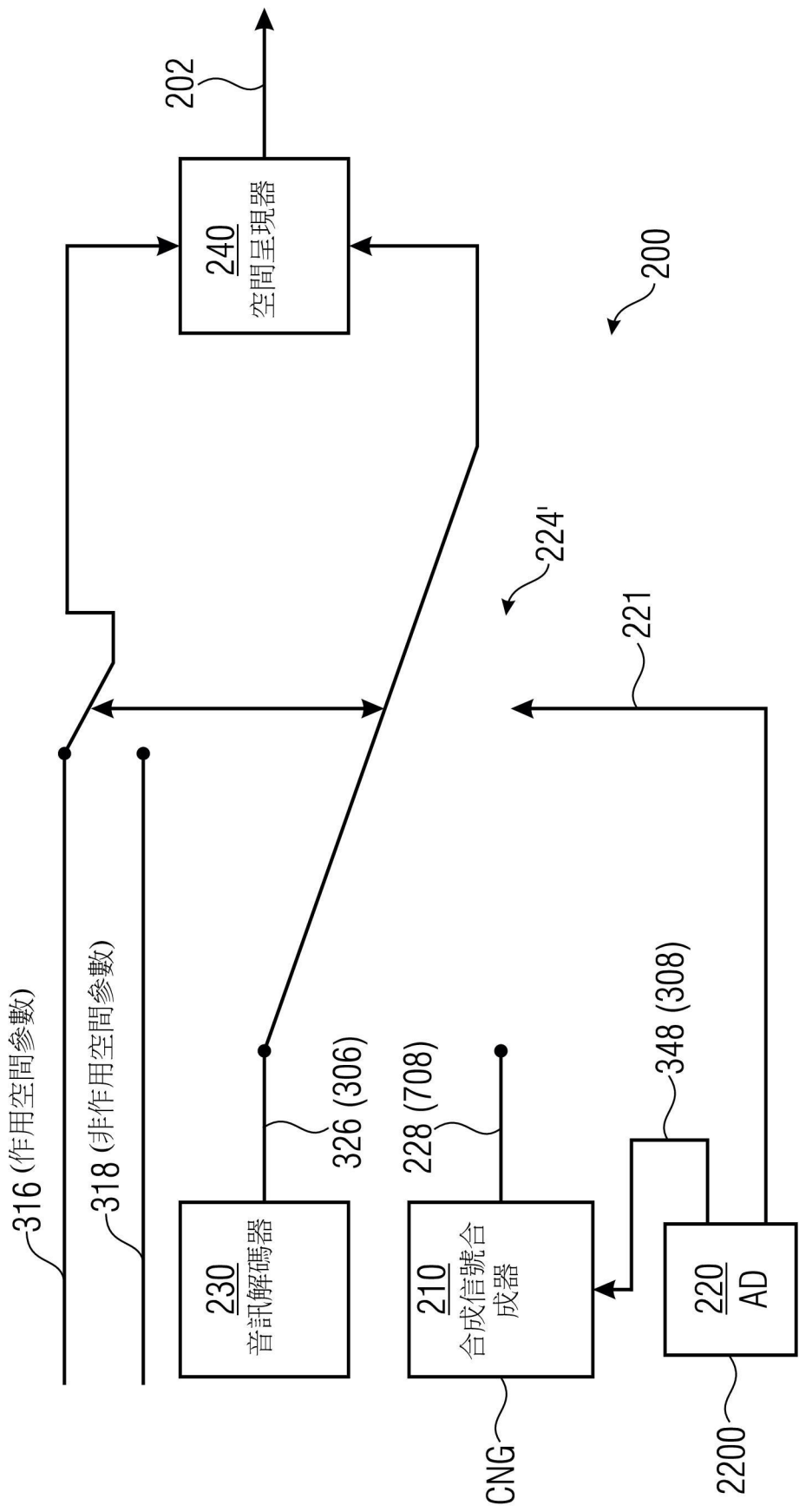
【圖8】



【圖9】



【圖10】



【圖11】