



(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 I856980 B

(45)公告日：中華民國 113 (2024) 年 10 月 01 日

(21)申請案號：108136436

(22)申請日：中華民國 108 (2019) 年 10 月 08 日

(51)Int. Cl. : **G10L19/008 (2013.01)****G06F13/16 (2006.01)**

(30)優先權：2018/10/08 美國

62/742,729

(71)申請人：美商杜拜研究特許公司(美國) DOLBY LABORATORIES LICENSING CORPORATION (US)

美國

瑞典商都比國際公司(瑞典) DOLBY INTERNATIONAL AB (SE)

瑞典

(72)發明人：布魯恩 史蒂芬 BRUHN, STEFAN (DE)；艾克特 麥可 ECKERT, MICHAEL (US)；托瑞斯 瓊恩 菲立克斯 TORRES, JUAN FELIX (US)；伯朗 史蒂芬妮 BROWN, STEPHANIE (AU)；麥格拉斯 大衛 S MCGRATH, DAVID S. (AU)

(74)代理人：陳長文

(56)參考文獻：

US 2012/0054664A1

US 2016/0099001A1

US 2017/0076735A1

審查人員：林坤隆

申請專利範圍項數：9 項 圖式數：7 共 42 頁

(54)名稱

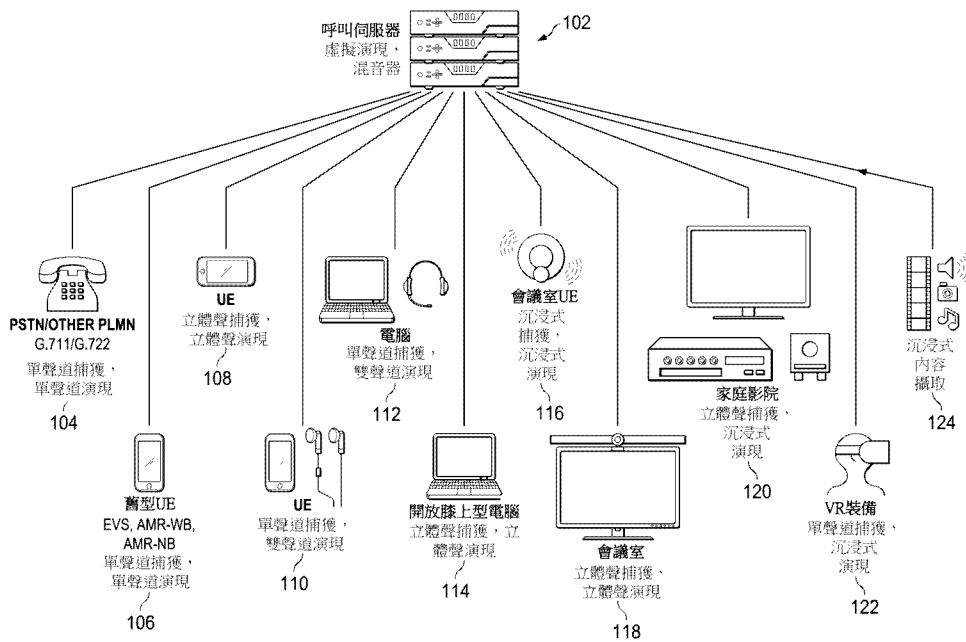
用於音頻信號處理至減少數量之音頻格式之系統、方法及裝置

(57)摘要

所揭示實施例實現將藉由各種捕獲器件捕獲之各種格式中之音頻信號轉變至可藉由一音頻編解碼器(例如，一沉浸式語音及音頻服務(IVAS)編解碼器)處理之有限數量之格式。在一項實施例中，音頻器件之一簡化單元接收藉由耦合至該音頻器件之一或多個音頻捕獲器件捕獲之一音頻信號。該簡化單元判定該音頻信號是否在該音頻器件之一編碼單元支援/不支援之一格式中。基於該判定，該簡化單元將該音頻信號轉變至該編碼單元支援之一格式。在一項實施例中，若該簡化單元判定該音頻信號係在一空間格式中，則該簡化單元可將該音頻信號轉變至該編碼支援之一空間「夾層」格式。

The disclosed embodiments enable converting audio signals captured in various formats by various capture devices into a limited number of formats that can be processed by an audio codec (e.g., an Immersive Voice and Audio Services (IVAS) codec). In an embodiment, a simplification unit of the audio device receives an audio signal captured by one or more audio capture devices coupled to the audio device. The simplification unit determines whether the audio signal is in a format that is supported/not supported by an encoding unit of the audio device. Based on the determining, the simplification unit, converts the audio signal into a format that is supported by the encoding unit. In an embodiment, if the simplification unit determines that the audio signal is in a spatial format, the simplification unit can convert the audio signal into a spatial “mezzanine” format supported by the encoding.

指定代表圖：



【圖1】

符號簡單說明：

102:呼叫伺服器

104:公用交換電話網路(PSTN)/其他公用陸地行動網路(PLMN)器件、器件、公用交換電話網路(PSTN)/公用陸地行動網路(PLMN)電話

106:舊型使用者設備

108:使用者設備

110:使用者器件

112:電腦器件

114:膝上型電腦

116:會議室使用設備

118:會議室系統

120:家庭影院

122:虛擬實境(VR)裝備

124:沉浸式內容攝取



I856980

【發明摘要】

公告本

【中文發明名稱】

用於音頻信號處理至減少數量之音頻格式之系統、方法及裝置

【英文發明名稱】

SYSTEM, METHOD AND APPARATUS FOR AUDIO SIGNAL
PROCESSING INTO A REDUCED NUMBER OF AUDIO FORMATS

【中文】

所揭示實施例實現將藉由各種捕獲器件捕獲之各種格式中之音頻信號轉變至可藉由一音頻編解碼器(例如，一沉浸式語音及音頻服務(IVAS)編解碼器)處理之有限數量之格式。在一項實施例中，音頻器件之一簡化單元接收藉由耦合至該音頻器件之一或多個音頻捕獲器件捕獲之一音頻信號。該簡化單元判定該音頻信號是否在該音頻器件之一編碼單元支援/不支援之一格式中。基於該判定，該簡化單元將該音頻信號轉變至該編碼單元支援之一格式。在一項實施例中，若該簡化單元判定該音頻信號係在一空間格式中，則該簡化單元可將該音頻信號轉變至該編碼支援之一空間「夾層」格式。

【英文】

The disclosed embodiments enable converting audio signals captured in various formats by various capture devices into a limited number of formats that can be processed by an audio codec (e.g., an Immersive Voice and Audio Services (IVAS) codec). In an embodiment, a simplification

unit of the audio device receives an audio signal captured by one or more audio capture devices coupled to the audio device. The simplification unit determines whether the audio signal is in a format that is supported/not supported by an encoding unit of the audio device. Based on the determining, the simplification unit, converts the audio signal into a format that is supported by the encoding unit. In an embodiment, if the simplification unit determines that the audio signal is in a spatial format, the simplification unit can convert the audio signal into a spatial “mezzanine” format supported by the encoding.

【指定代表圖】

圖1

【代表圖之符號簡單說明】

- | | |
|-----|--|
| 102 | 呼叫伺服器 |
| 104 | 公用交換電話網路(PSTN)/其他公用陸地行動網路(PLMN)
器件、器件、公用交換電話網路(PSTN)/公用陸地行動網
路(PLMN)電話 |
| 106 | 舊型使用者設備 |
| 108 | 使用者設備 |
| 110 | 使用者器件 |
| 112 | 電腦器件 |
| 114 | 膝上型電腦 |
| 116 | 會議室使用設備 |
| 118 | 會議室系統 |

- 120 家庭影院
- 122 虛擬實境(VR)裝備
- 124 沉浸式內容攝取

【發明說明書】

【中文發明名稱】

用於音頻信號處理至減少數量之音頻格式之系統、方法及裝置

【英文發明名稱】

SYSTEM, METHOD AND APPARATUS FOR AUDIO SIGNAL
PROCESSING INTO A REDUCED NUMBER OF AUDIO FORMATS

【技術領域】

【0001】 本發明之實施例大體上係關於音頻信號處理，且更明確言之係關於經捕獲音頻信號之分配。

【先前技術】

【0002】 語音及視訊編碼器/解碼器(「編解碼器」)標準開發最近集中於開發用於沉浸式語音及音頻服務(IVAS)之一編解碼器。預期IVAS將支援一系列服務能力，諸如關於單聲道至立體聲至完全沉浸式音頻編碼、解碼及演現之操作。一合適IVAS編解碼器亦提供針對不同傳輸條件下之封包丟失及延遲抖動之高誤差穩健性。IVAS旨在由廣泛範圍之器件、端點及網路節點支援，包含(但不限於)行動及智慧型電話、電子平板電腦、個人電腦、會議電話、會議室、虛擬實境及擴增實境器件、家庭影院器件及其他合適器件。因為此等器件、端點及網路節點可具有用於聲音捕獲及演現之各種聲介面，所以一IVAS編解碼器解決其中捕獲及演現一音頻信號之所有不同方式可能不切實際。

【發明內容】

【0003】 所揭示實施例能夠將藉由各種捕獲器件捕獲之各種格式中之音頻信號轉變至可藉由一編解碼器(例如，一IVAS編解碼器)處理之有限數量之格式。

【0004】 在一些實施例中，建置於一音頻器件中之一簡化單元接收一音頻信號。該音頻信號可為藉由與該音頻器件耦合之一或多個音頻捕獲器件捕獲之一信號。例如，該音頻信號可為不同位置處之人之間的一視訊會議之一音頻。該簡化單元判定該音頻信號是否在該音頻器件之一編碼單元(通常被稱為一「編碼器」)不支援之一格式中。例如，簡化單元可判定音頻信號是否在一單聲道、立體聲或一標準或專有空間格式中。基於判定音頻信號在編碼單元不支援之一格式中，簡化單元將音頻信號轉變至編碼單元支援之一格式。例如，若簡化單元判定音頻信號係在一專有空間格式中，則簡化單元可將音頻信號轉變至編碼單元支援之一空間「夾層」格式。簡化單元將該經轉變音頻信號傳送至編碼單元。

【0005】 所揭示實施例之一優點在於，可藉由將可能較大數量之音頻捕獲格式減少至有限數量之格式(例如，單聲道、立體聲及空間)而降低一編解碼器(例如，一IVAS編解碼器)之複雜性。因此，可將該編解碼器部署於各種器件上，不考慮該等器件之音頻捕獲能力。

【0006】 此等及其他態樣、特徵及實施例可被表示為用於執行一功能之方法、裝置、系統、組件、程式產品、方式或步驟及以其他方式表示。

【0007】 在一些實施方案中，一音頻器件之一簡化單元接收一第一格式中之一音頻信號。該第一格式係該音頻器件支援之多個音頻格式之一集合中之一者。該簡化單元判定音頻器件之一編碼器是否支援第一格式。

根據該編碼器不支援第一格式，簡化單元將音頻信號轉變至編碼器支援之一第二格式。該第二格式係第一格式之一替代表示。簡化單元將第二格式中之音頻信號傳送至編碼器。編碼器編碼音頻信號。音頻器件儲存該經編碼音頻信號或將該經編碼音頻信號傳輸至一或多個其他器件。

【0008】 將音頻信號轉變至第二格式可包含產生用於音頻信號之後設資料。該後設資料可包含音頻信號之一部分之一表示。編碼音頻信號可包含將第二格式中之音頻信號編碼至一第二器件支援之一輸送格式。音頻器件可藉由傳輸包括第二格式不支援之音頻信號之一部分之一表示之後設資料而傳輸該經編碼音頻信號。

【0009】 在一些實施方案中，藉由簡化單元判定音頻信號是否在第一格式中可包含判定音頻捕獲器件之一數量及用於捕獲音頻信號之各捕獲器件之一對應位置。一或多個其他器件之各者可經組態以自第二格式重現音頻信號。一或多個其他器件之至少一者可能無法自第一格式重現音頻信號。

【0010】 第二格式可將音頻信號表示為一音頻場景中之音頻物件之一數量，兩者皆依靠用於攜載空間資訊之音頻通道之一數量。第二格式可包含用於攜載空間資訊之一進一步部分之後設資料。第一格式及第二格式皆可為空間音頻格式。第二格式可為一空間音頻格式且第一格式可為與後設資料相關聯之一單聲道格式或與後設資料相關聯之一立體聲格式。音頻器件支援之多個音頻格式之集合可包含多個空間音頻格式。第二格式可為第一格式之一替代表示且其進一步特徵在於實現可比程度之體驗品質。

【0011】 在一些實施方案中，一音頻器件之一演現單元接收一第一格式中之一音頻信號。該演現單元判定該音頻器件是否能夠重現該第一格

式中之該音頻信號。回應於判定音頻器件無法重現第一格式中之音頻信號，演現單元調適音頻信號以在第一第二格式中可用。演現單元傳送第二格式中之音頻信號以用於演現。

【0012】 在一些實施方案中，藉由演現單元將音頻信號轉變至第二格式可包含使用包含用於編碼之一第四格式不支援之音頻信號之一部分之一表示之後設資料連同一第三格式中之音頻信號。此處，在簡化單元之背景內容中該第三格式對應於術語「第一格式」，該「第一格式」係編碼器側處支援之多個音頻格式之一集合中之一者。在簡化單元之背景內容中該第四格式對應於術語「第二格式」，該「第二格式」係編碼器支援之一格式且係第三格式之一替代表示。在本說明書中之此處及別處，術語第一、第二、第三及第四係用於識別且並不一定指示一特定順序。

【0013】 一解碼單元接收一輸送格式中之音頻信號。該解碼單元將該輸送格式中之音頻信號解碼至第一格式，且將第一格式中之音頻信號傳送至演現單元。在一些實施方案中，調適音頻信號以在第二格式中可用可包含調適解碼以產生第二格式中之經接收音頻。在一些實施方案中，多個器件之各者經組態以重現第二格式中之音頻信號。多個器件之一或多者無法重現第一格式中之音頻信號。

【0014】 在一些實施方案中，一簡化單元自一聲預處理單元接收多個格式中之音頻信號。該簡化單元自一器件接收該器件之屬性，該等屬性包含該器件支援之一或多個音頻格式之指示。該一或多個音頻格式包含一單聲道格式、一立體聲格式或一空間格式之至少一者。簡化單元將音頻信號轉變至作為一或多個音頻格式之一替代表示之一攝取格式。簡化單元將該經轉變之音頻信號提供至一編碼單元以進行下游處理。聲預處理單元、

簡化單元及該編碼單元之各者可包含一或多個電腦處理器。

【0015】 在一些實施方案中，一編碼系統包含：一捕獲單元，其經組態以捕獲一音頻信號；一聲預處理單元，其經組態以執行包括預處理該音頻信號之操作；一編碼器；及一簡化單元。該簡化單元經組態以執行以下操作。簡化單元自該聲預處理單元接收一第一格式中之一音頻信號。該第一格式係該編碼器支援之多個音頻格式之一集合中之一者。簡化單元判定編碼器是否支援第一格式。回應於判定編碼器不支援第一格式，簡化單元將音頻信號轉變至編碼器支援之一第二格式。簡化單元將該第二格式中之音頻信號傳送至編碼器。編碼器經組態以執行包含以下項之操作：編碼音頻信號；及儲存該經編碼音頻信號或將該經編碼音頻信號傳輸至另一器件之至少一者。

【0016】 在一些實施方案中，將音頻信號轉變至第二格式包含產生用於音頻信號之後設資料。該後設資料可包含第二格式不支援之音頻信號之一部分之一表示。編碼器之操作可進一步包含藉由傳輸包含第二格式不支援之音頻信號之一部分之一表示之後設資料而傳輸經編碼音頻信號。

【0017】 在一些實施方案中，第二格式將音頻信號表示為一音頻場景中之物件之一數量及用於攜載空間資訊之通道之一數量。在一些實施方案中，預處理音頻信號可包含執行雜訊消除、執行回波消除、減少音頻信號之通道之一數量、增加音頻信號之音頻通道之該數量或產生聲後設資料之一或多者。

【0018】 在一些實施方案中，一解碼系統包含一解碼器、一演現單元及一重播單元。該解碼器經組態以執行包含(例如)將一音頻信號自一輸送格式解碼至一第一格式之操作。該演現單元經組態以執行以下操作。演

現單元接收該第一格式中之音頻信號。演現單元判定一音頻器件是否能夠重現一第二格式中之音頻信號。該第二格式實現比第一格式使用更多輸出器件。回應於判定該音頻器件能夠重現第二格式中之音頻信號，演現單元將音頻信號轉變至第二格式。演現單元演現第二格式中之音頻信號。重播單元經組態以執行包含起始在一揚聲器系統上播放經演現音頻信號之操作。

【0019】 在一些實施方案中，將音頻信號轉變至第二格式可包含使用包含用於編碼之一第四格式不支援之音頻信號之一部分之一表示之後設資料連同一第三格式中之音頻信號。此處，在簡化單元之背景內容中該第三格式對應於術語「第一格式」，該「第一格式」係編碼器側處支援之多個音頻格式之一集合中之一者。在簡化單元之背景內容中該第四格式對應於術語「第二格式」，該「第二格式」係編碼器支援之一格式且係第三格式之一替代表示。

【0020】 在一些實施方案中，解碼器之操作可進一步包含接收一輸送格式中之音頻信號及將第一格式中之音頻信號傳送至演現單元。

【0021】 將自包含技術方案之以下描述明白此等及其他態樣、特徵及實施例。

【圖式簡單說明】

【0022】

在圖式中，為便於描述，展示示意性元件(諸如表示器件、單元、指令塊及資料元素之彼等)之特定配置或排序。然而，熟習此項技術者應理解，圖式中之示意性元件之特定排序或配置並不意欲暗示需要一特定處理順序或序列或程序分離。此外，在一圖式中包含一示意性元件並不意欲暗

示在所有實施例中需要此元件或藉由此元件表示之特徵可能不包含於一些實施例中之其他元件中或結合一些實施例中之其他元件。

此外，在圖式中，在使用連接元件(諸如實線或虛線或箭頭)來繪示兩個或兩個以上其他示意性元件之間或中間之一連接、關係或關聯之情況下，不存在任何此等連接元件並不意欲暗示無連接、關係或關聯可存在。換言之，在圖式中未展示元件之間的一些連接、關係或關聯以免模糊本發明。另外，為便於圖解說明，使用一單個連接元件來表示元件之間的多個連接、關係或關聯。例如，在一連接元件表示信號、資料或指令之通信之情況下，熟習此項技術者應理解，此元件表示如實現該通信可能需要之一或多個信號路徑。

圖1繪示根據本發明之一些實施例之IVAS系統可支援之各種器件。

圖2A係根據本發明之一些實施例之用於將經捕獲音頻信號轉換至準備用於編碼之一格式之一系統的一方塊圖。

圖2B係根據本發明之一些實施例之用於將經捕獲音頻轉換回至一合適重播格式之一系統的一方塊圖。

圖3係根據本發明之一些實施例之用於將一音頻信號轉換至一編碼單元支援之一格式之例示性動作的一流程圖。

圖4係根據本發明之一些實施例之用於判定一音頻信號是否在編碼單元支援之一格式中之例示性動作的一流程圖。

圖5係根據本發明之一些實施例之用於將一音頻信號轉換至一合適重播格式之例示性動作的一流程圖。

圖6係根據本發明之一些實施例之用於將一音頻信號轉換至一可用重播格式之例示性動作的另一流程圖。

圖7係根據本發明之一些實施例之用於實施參考圖1至圖6所描述之特徵之一硬體架構的一方塊圖。

【實施方式】

【0023】

相關申請案之交叉參考

本申請案主張於2018年10月8日申請之美國臨時專利申請案第62/742,729號之優先權利，該案之全文以引用的方式併入。

【0024】 在以下描述中，出於解釋目的，闡述數種具體細節以提供對本發明之一透徹理解。然而，將明白，可在沒有此等具體細節之情況下實踐本發明。

【0025】 現將詳細參考實施例，其等之實例係在附圖中進行繪示。在以下詳細描述中，闡述數種具體細節以提供對各項所描述實施例之一透徹理解。然而，一般技術者將明白，可在不具有此等具體細節之情況下實踐各項所描述實施例。在其他例項中，未詳細描述熟知方法、程序、組件及電路以免不必要地模糊實施例之態樣。以下描述可各彼此獨立使用或與其他特徵之任何組合一起使用之若干特徵。

【0026】 如本文中所使用，術語「包含」及其變體應被解讀為意謂「包含(但不限於)」之開放式術語。術語「或」應被解讀為「及/或」，除非上下文另有明確規定。術語「基於」應被解讀為「至少部分基於」。

【0027】 圖1繪示IVAS系統可支援之各種器件。在一些實施方案中，此等器件透過呼叫伺服器102通信，該呼叫伺服器102可自(例如)藉由PSTN/其他PLMN器件104繪示之一公用交換電話網路(PSTN)或一公用陸地行動網路(PLMN)器件接收音頻信號。此器件可使用G.711及/或G.722

標準用於音頻(話音)壓縮及解壓縮。一器件104通常僅能夠捕獲及演現單聲道音頻。IVAS系統經啟用以亦支援舊型使用者設備106。該等舊型器件可包含增強型語音服務(EVS)器件、自適應多速率寬頻(AMR-WB)話音至音頻寫碼標準支援器件、自適應多速率窄頻(AMR-NB)支援器件及其他合適器件。此等器件通常僅演現及捕獲單聲道中之音頻。

【0028】 IVAS系統亦經啟用以支援捕獲及演現各種格式(包含先進音頻格式)中之音頻信號之使用者設備。例如，IVAS系統經啟用以支援立體聲捕獲及演現器件(例如，使用者設備108、膝上型電腦114及會議室系統118)、單聲道捕獲及雙聲道演現器件(例如，使用者器件110及電腦器件112)、沉浸式捕獲及演現器件(例如，會議室使用設備116)、立體聲捕獲及沉浸式演現器件(例如，家庭影院120)、單聲道捕獲及沉浸式演現(例如，虛擬實境(VR)裝備122)、沉浸式內容攝取124及其他合適器件。為直接支援所有此等格式，用於IVAS系統之編解碼器將需要非常複雜且昂貴的安裝。因此，將需要用於在編碼階段之前簡化編解碼器之一系統。

【0029】 儘管以下描述集中於一IVAS系統及編解碼器，然所揭示實施例可應用於用於任何音頻系統之任何編解碼器，其中一優點在於，將較大數量之音頻捕獲格式減少至一較小數量以降低音頻編解碼器之複雜性或用於任何其他所要原因。

【0030】 圖2A係根據本發明之一些實施例之用於將經捕獲音頻信號轉換至準備用於編碼之一格式之一系統200的一方塊圖。捕獲單元210自一或多個捕獲器件(例如，麥克風)接收一音頻信號。例如，捕獲單元210可自一個麥克風接收一音頻信號(例如，單聲道信號)、自兩個麥克風接收一音頻信號(例如，立體聲信號)、自三個麥克風或自另一數量及組態之音

頻捕獲器件接收一音頻信號。捕獲單元210可包含藉由一或多個第三方之客製化，其中該等客製化可特定於所使用之捕獲器件。

【0031】 在一些實施方案中，用一個麥克風捕獲一單聲道音頻信號。例如，可用如圖1中所繪示之PSTN/PLMN電話104、舊型使用者設備106、具有一免提耳機之使用者器件110、具有一經連接耳機之電腦器件112及虛擬實境裝備122捕獲該單聲道信號。

【0032】 在一些實施方案中，捕獲單元210接收使用各種錄製/麥克風技術捕獲之立體聲音頻。例如，可藉由使用者設備108、膝上型電腦114、會議室系統118及家庭影院120捕獲立體聲音頻。在一實例中，用相同位置處之以約90度或更大之一擴展角放置之兩個指向性麥克風捕獲立體聲音頻。立體聲效應由通道間層級差所引起。在另一實例中，立體聲音頻係藉由兩個空間移位之麥克風捕獲。在一些實施方案中，該等空間移位之麥克風係全向麥克風。此組態中之立體聲效應由通道間層級差及通道間時間差所引起。麥克風之間的距離對經感知立體聲寬度具有相當大影響。在又另一實例中，用具有17厘米位移及110度之一擴展角之兩個指向性麥克風捕獲音頻。此系統通常被稱為Office de Radiodiffusion Télévision Française (「ORTF」)立體聲麥克風系統。又另一立體聲捕獲系統包含具有不同特性之兩個麥克風，該兩個麥克風經配置使得一個麥克風信號係中間信號且另一個麥克風信號係旁側信號。此配置通常被稱為中間-旁側(M/S)錄製。來自M/S之信號之立體聲效應通常建立在通道間層級差上。

【0033】 在一些實施方案中，捕獲單元210接收使用多麥克風技術捕獲之音頻。在此等實施方案中，音頻之捕獲涉及三個或三個以上麥克風之一配置。通常需要此配置用於捕獲空間音頻且此配置亦可有效地執行環

境雜訊抑制。在麥克風數量增加時，可藉由麥克風捕獲之一空間場景之細節數量亦增加。在一些例項中，當麥克風數量增加時，亦改良經捕獲場景之準確度。例如，以免提模式操作之圖1之各種使用者設備(UE)可利用多個麥克風以產生一單聲道、立體聲或空間音頻信號。此外，具有多個麥克風之一開放膝上型電腦114可用於產生一立體聲捕獲。一些製造商發行具有兩至四個微機電系統(「MEMS」)麥克風之膝上型電腦，從而容許立體聲捕獲。例如，可在會議室使用者設備116中實施多麥克風沉浸式音頻捕獲。

【0034】 經捕獲音頻通常在被攝取至一語音或音頻編解碼器中之前經歷一預處理階段。因此，聲預處理單元220自捕獲單元210接收一音頻信號。在一些實施方案中，聲預處理單元220執行雜訊及回波消除處理、通道降混及升混(例如，減少或增加音頻通道之一數量)及/或任何種類之空間處理。聲預處理單元220之音頻信號輸出通常適用於編碼及傳輸至其他器件。在一些實施方案中，聲預處理單元220之特定設計係由一器件製造商執行，此係因為該特定設計取決於藉由一特定器件之音頻捕獲之細節。然而，由相關聲介面規範設定之要求可對此等設計設定限制，且確保滿足特定品質要求。執行聲預處理之一目的係產生一IVSA編解碼器支援之一或多個不同種類之音頻信號或音頻輸入格式以實現各種IVAS目標使用案例或服務層級。取決於與此等使用案例相關聯之特定IVAS服務要求，可能需要一IVAS編解碼器來支援單聲道、立體聲及空間格式。

【0035】 通常，當單聲道格式係唯一可用格式(例如，基於捕獲器件之類型，例如，若發送器件之捕獲能力受限)時，使用單聲道格式。對於立體聲音頻信號，聲預處理單元220將經捕獲信號轉變至滿足特定慣例(例

如，通道排序左-右慣例)之一正規化表示。對於M/S立體聲捕獲，此程序可涉及(例如)一矩陣操作，使得使用左-右慣例表示信號。在預處理之後，立體聲信號滿足特定慣例(例如，左-右慣例)。然而，移除關於特定立體聲捕獲器件之資訊(例如，麥克風數量及組態)。

【0036】 對於空間格式，在聲預處理之後獲得之空間輸入信號或特定空間音頻格式之種類可取決於發送器件類型及發送器件用於捕獲音頻之能力。同時，IVAS服務需求可能需要之空間音頻格式包含低解析度空間、高解析度空間、後設資料輔助之空間音頻(MASA)格式，及高階環境立體聲(「HOA」)輸送格式(HTF)或甚至進一步空間音頻格式。因此，具有空間音頻能力之一發送器件之聲預處理單元220必須準備提供滿足此等要求之適當格式中之一空間音頻信號。

【0037】 低解析度空間格式包含空間WXY、一階環境立體聲(「FOA」)及其他格式。空間WXY格式係關於其中省略高度分量(Z)之三通道一階平面B格式音頻表示。此格式對於其中空間解析度要求並非很高且其中空間高度分量可被視為不相關之位元率高效沉浸式電話學及沉浸式會議情景係有用的。該格式對於會議電話特別有用，此係因為其使接收客戶端能夠執行在具有多個參與者之一會議室中捕獲之會議場景之沉浸式演現。同樣地，該格式適用於在一虛擬會議室中空間安排會議參與者之會議伺服器。相比之下，FOA含有高度分量(Z)作為第4分量信號。FOA表示係與低速率VR應用有關。

【0038】 高解析度空間格式包含基於通道、物件及場景之空間格式。取決於所涉及之音頻分量信號之數量，此等格式之各者容許以實際上無限制之解析度表示空間音頻。然而，出於各種原因(例如，位元率限制

及複雜性限制)，相對較少分量信號(例如，十二個)存在實際限制。進一步空間格式包含或可依靠MASA或HTF格式。

【0039】 要求支援IVAS之一器件以支援上文所論述之大量及各種音頻輸入格式可導致在複雜性、記憶體佔用面積、實施方案測試及維護方面之巨大成本。然而，並非所有器件將具有支援所有音頻格式之能力或受益於支援所有音頻格式。例如，可具有僅支援立體聲但不支援空間捕獲之IVAS啟用器件。其他器件可僅支援低解析度空間輸入，而進一步類別之器件可僅支援HOA捕獲。因此，不同器件將僅利用音頻格式之特定子集。因此，若IVAS編解碼器必須支援所有音頻格式之直接寫碼，則IVAS編解碼器將變得不必要地複雜及昂貴。

【0040】 為解決此問題，圖2A之系統200包含一簡化單元230。聲預處理單元220將音頻信號傳送至簡化單元230。在一些實施方案中，聲預處理單元220產生連同音頻信號一起傳送至簡化單元230之聲後設資料。該聲後設資料可包含與音頻信號有關之資料(例如，格式後設資料，諸如單聲道、立體聲、空間)。聲後設資料亦可包含雜訊消除資料及(例如)與捕獲單元210之物理或幾何性質有關之其他合適資料。

【0041】 簡化單元230將一器件支援之各種輸入格式轉變至一減少之通用編解碼器攝取格式集合。例如，IVAS編解碼器可支援三種攝取格式：單聲道、立體聲及空間。雖然單聲道及立體聲格式係類似或相同於如藉由聲預處理單元產生之各自格式，但空間格式可為一「夾層」格式。一夾層格式係可準確地表示自聲預處理單元220獲得且在上文所論述之任何空間音頻信號之一格式。此包含以基於任何通道、物件及場景之格式(或其等之組合)表示之空間音頻。在一些實施方案中，夾層格式可將音頻信

號表示為一音頻場景中之物件之一數量及用於攜載用於該音頻場景之空間資訊之通道之一數量。另外，夾層格式可表示MASA、HTF或其他空間音頻格式。一合適空間夾層格式可將空間音頻表示為m個物件及第n階HOA (「mObj+HOAn」)，其中m及n係包含零之低整數。

【0042】 圖3之程序300繪示用於將音頻資料自一第一格式轉換至一第二格式之例示性動作。在302，簡化單元230 (例如)自聲預處理單元220接收一音頻信號。如上文所論述，自聲預處理單元220接收之該音頻信號可為已執行雜訊及回波消除處理以及執行通道降混及升混處理(例如，減少或增加音頻通道之一數量)之一信號。在一些實施方案中，簡化單元230接收聲後設資料連同音頻信號。聲後設資料可包含格式指示及如上文所論述之其他資訊。

【0043】 在304，簡化單元230判定音頻信號是否在音頻器件之一編碼單元240支援或不支援之一第一格式中。例如，如圖2A中所展示，音頻格式偵測單元232可分析自聲預處理單元220接收之音頻信號且識別該音頻信號之一格式。若音頻格式偵測單元232判定音頻信號係在一單聲道格式或一立體聲格式中，則簡化單元230將信號傳遞至編碼單元240。然而，若音頻格式偵測單元232判定信號係在一空間格式中，則音頻格式偵測單元232將音頻信號傳遞至轉換單元234。在一些實施方案中，音頻格式偵測單元232可使用聲後設資料以判定音頻信號之格式。

【0044】 在一些實施方案中，簡化單元230藉由判定用於捕獲音頻信號之音頻捕獲器件(例如，麥克風)之一數量、組態或位置而判定音頻信號是否在第一格式中。例如，若音頻格式偵測單元232判定音頻信號係藉由一單個捕獲器件(例如，單個麥克風)捕獲，則音頻格式偵測單元232可

判定該音頻信號係一單聲道信號。若音頻格式偵測單元232判定音頻信號係藉由彼此成一特定角度之兩個捕獲器件捕獲，則音頻格式偵測單元232可判定該信號係一立體聲信號。

【0045】 圖4係根據本發明之一些實施例之用於判定一音頻信號是否在編碼單元支援之一格式中之例示性動作的一流程圖。在402，簡化單元230存取音頻信號。例如，音頻格式偵測單元232可接收音頻信號作為輸入。在404，簡化單元230判定音頻器件之聲捕獲組態，例如，用於捕獲音頻信號之麥克風之一數量及麥克風之位置組態。例如，音頻格式偵測單元232可分析音頻信號且判定三個麥克風定位於一空間內之不同位置處。在一些實施方案中，音頻格式偵測單元232可使用聲後設資料以判定聲捕獲組態。即，聲預處理單元220可產生指示各捕獲器件之位置及捕獲器件之數量之聲後設資料。後設資料亦可含有經偵測音頻性質之描述，諸如一聲源之方向或指向性。在406，簡化單元230比較聲捕獲組態與一或多個經儲存聲捕獲組態。例如，經儲存聲捕獲組態可包含各麥克風之一數量及位置以識別一特定組態(例如，單聲道、立體聲或空間)。簡化單元230比較該等聲捕獲組態之各者與音頻信號之聲捕獲組態。

【0046】 在408，簡化單元230判定聲捕獲組態是否匹配與一空間格式相關聯之一經儲存聲捕獲組態。例如，簡化單元230可判定用於捕獲音頻信號之麥克風之一數量及麥克風在一空間中之位置。簡化單元230可比較該資料與用於空間格式之經儲存已知組態。若簡化單元230判定不與一空間格式匹配(此可為音頻格式係單聲道或立體聲之一指示)，則程序400移至412，其中簡化單元230將音頻信號傳送至一編碼單元240。然而，若簡化單元230將音頻格式識別為屬於空間格式集合，則程序400移至410，

其中簡化單元230將音頻信號轉變至一夾層格式。

【0047】 返回參考圖3，在306，簡化單元230根據判定音頻信號係在編碼單元不支援之一格式中而將音頻信號轉變至編碼單元支援之一第二格式。例如，轉換單元234可將音頻信號轉換至一夾層格式。該夾層格式準確地表示最初以任何基於通道、物件及場景之格式(或其等之組合)表示之一空間音頻信號。另外，夾層格式可表示MASA、HTF或另一合適格式。例如，可用作空間夾層格式之一格式可將音頻表示為m個物件及第n階HOA (「mObj+HOAn」，其中m及n係包含零之低整數。夾層格式可因此需要表示具有可捕獲音頻信號之顯式性質之波形(信號)及後設資料之音頻。

【0048】 在一些實施方案中，轉換單元234在將音頻信號轉變至第二格式時產生用於音頻信號之後設資料。該後設資料可與在第二格式中之音頻信號之一部分相關聯，例如，物件後設資料包含一或多個物件之位置。另一實例係其中使用一組專有捕獲器件捕獲音頻及其中編碼單元及/或夾層格式不支援或有效地表示該等器件之數量及組態。在此等情況中，轉換單元234可產生後設資料。該後設資料可包含轉換後設資料或聲後設資料之至少一者。該轉換後設資料可包含與編碼程序及/或夾層格式不支援之格式之一部分相關聯之一後設資料子集。例如，當在經組態以特別輸出藉由專有組態捕獲之音頻之一系統上重播音頻信號時，轉換後設資料可包含用於捕獲(例如，麥克風)組態之器件設定及/或用於輸出器件(例如，揚聲器)組態之器件設定。源自於聲預處理單元220及/或轉換單元234之後設資料亦可包含聲後設資料，該聲後設資料描述特定音頻信號性質，諸如經捕獲聲音所來自之一空間方向、聲音之一指向性或一擴散度。在此實例

中，可判定音頻係空間的，在空間格式中，但經表示為具有額外後設資料之一單聲道或一立體聲信號。在此情況中，該等單聲道或立體聲信號及該後設資料係經傳播至編碼器240。

【0049】 在308，簡化單元230將第二格式中之音頻信號傳送至編碼單元。如圖2A中所繪示，若音頻格式偵測單元232判定音頻係在一單聲道或立體聲格式中，則音頻格式偵測單元232將音頻信號傳送至編碼單元。然而，若音頻格式偵測單元232判定音頻信號係在一空間格式中，則音頻格式偵測單元232將音頻信號傳送至轉換單元234。轉換單元234在將空間音頻轉換至(例如)夾層格式之後，將音頻信號傳送至編碼單元240。在一些實施方案中，除了音頻信號之外，轉換單元234亦將轉換後設資料及聲後設資料傳送至編碼單元240。

【0050】 編碼單元240接收第二格式(例如，夾層格式)中之音頻信號且將第二格式中之音頻信號編碼至一輸送格式。編碼單元240將經編碼音頻信號傳播至某一發送實體，該發送實體將經編碼音頻信號傳輸至一第二器件。在一些實施方案中，編碼單元240或後續實體儲存經編碼音頻信號以用於稍後傳輸。編碼單元240可接收單聲道、立體聲或夾層格式中之音頻信號且編碼該等信號以用於音頻輸送。若音頻信號係在夾層格式中且編碼單元自簡化單元230接收轉換後設資料及/或聲後設資料，則編碼單元將轉換後設資料及/或聲後設資料傳送至第二器件。在一些實施方案中，編碼單元240將轉換後設資料及/或聲後設資料編碼至第二器件可接收並解碼之一特定信號。編碼單元接著將經編碼音頻信號輸出至待輸送至一或多個其他器件之音頻輸送。因此，(例如，圖1中之器件之)各器件能夠編碼第二格式(例如，夾層格式)中之音頻信號，但該等器件通常無法編碼第一格

式中之音頻信號。

【0051】 在一實施例中，編碼單元240 (例如，先前描述之IVAS編解碼器)對藉由簡化階段提供之單聲道、立體聲或空間音頻信號進行操作。依靠可基於協商之IVAS服務層級、發送及接收側器件能力及可用位元率之一或多者之一編解碼器模式選擇來進行編碼。

【0052】 舉例而言，服務層級可包含IVAS立體聲電話學、IVAS沉浸式會議、IVAS使用者產生之VR串流化或另一合適服務層級。可對選擇IVAS編解碼器操作之一合適模式所針對之一特定IVAS服務層級指派一特定音頻格式(單聲道、立體聲、空間)。

【0053】 此外，可回應於發送及接收側器件能力來選擇IVAS編解碼器操作模式。例如，取決於發送器件能力，編碼單元240可能無法存取(例如)一空間攝取信號，此係因為編碼單元240僅被提供一單聲道或一立體聲信號。另外，一端至端能力交換或一對應編解碼器模式請求可指示接收端具有特定演現限制，從而無需編碼及傳輸一空間音頻信號或反之亦然。在另一實例中，另一器件可請求空間音頻。

【0054】 在一些實施方案中，一端至端能力交換不能完全解決遠端器件能力。例如，編碼點可能不具有關於解碼單元(有時被稱為一解碼器)是否將為一單個單聲道揚聲器、立體聲揚聲器或其是否將經雙聲道演現之資訊。實際演現情景可在一服務會話期間改變。例如，若經連接重播設備改變，則演現情景可改變。在一實例中，可能不存在端至端能力交換，此係因為在IVAS編碼會話期間未連接阱(sink)器件。此可針對語音郵件服務或在(使用者產生之)虛擬實境內容串流化服務中發生。其中接收器件能力未知或歸因於模糊度而無法解決之另一實例係需要支援多個端點之一單個

編碼器。例如，在一IVAS會議或虛擬實境內容分配中，一端點可使用一耳機且另一端點可向立體聲揚聲器演現。

【0055】 解決此問題之一方式係假定最小可能接收器件能力及選擇一對應IVAS編解碼器操作模式(在特定情況中，其可為單聲道)。解決此問題之另一方式係需要IVAS解碼器(即使編碼器係在支援空間或立體聲音頻之一模式中操作)推導可在具有相對較低音頻能力之器件上演現之一經解碼音頻信號。即，編碼為一空間音頻信號之一信號亦應可針對立體聲演現及單聲道演現兩者來解碼。同樣地，編碼為立體聲之一信號亦應可針對單聲道演現來解碼。

【0056】 例如，在IVAS會議中，一呼叫伺服器應僅需要執行一單一編碼且發送相同編碼至多個端點，該多個端點中之一些可為雙聲道的且一些可為立體聲的。因此，一單一雙通道編碼可支援在(例如)具有立體聲揚聲器之膝上型電腦114及會議室系統118上之演現及在使用者器件110及虛擬實境裝備122上之具有雙聲道呈現之沉浸式演現兩者。因此，一單一編碼可同時支援兩個結果。因此，一意涵在於，雙通道編碼支援藉由一單一編碼之立體聲揚聲器播出及雙聲道演現播出兩者。

【0057】 另一實例涉及高品質單聲道提取。系統可支援自一經編碼空間或立體聲音頻信號提取一高品質單聲道信號。在一些實施方案中，可提取一增強型語音服務(「EVS」)編解碼器位元串流以(例如)使用標準EVS解碼器進行單聲道解碼。

【0058】 替代性地或除了服務層級及器件能力之外，可用位元率係可控制編解碼器模式選擇之另一參數。在一些實施方案中，位元率需求隨著可在接收端處提供之體驗品質及隨著音頻信號之分量之相關聯數量而增

加。在最低端位元率下，僅單聲道音頻演現係可能的。EVS編解碼器提供低至每秒5.9千位元之單聲道操作。隨著位元率增加，可達成較高品質服務。然而，編碼品質(「QoE」)仍歸因於僅單聲道操作及演現而受限。對於(習知)雙通道立體聲，次高層級之QoE係可能的。然而，系統需要高於最低單聲道位元率之一位元率以提供有用品質，此係因為現有兩個音頻信號分量待傳輸。空間聲音體驗需要高於立體聲之QoE。在位元率範圍之較低端處，可用可被稱為「空間立體聲」之空間信號之一雙聲道表示來實現此體驗。空間立體聲依靠至編碼器(例如，編碼單元240)中之空間音頻信號攝取之編碼器側雙聲道預演現(具有適當標頭相關傳送功能(「HRTF」))且因其僅由兩個音頻分量信號組成而有可能為最緊湊空間表示。因為空間立體聲攜載更多感知資訊，所以達成一足夠品質所需之位元率有可能高於一習知立體聲信號所需之位元率。然而，空間立體聲表示在客製化接收端處之演現方面可能會有限制。此等限制可包含對耳機演現、對使用一組預選定HRTF或對無需標頭追蹤之演現之限制。藉由用於編碼一空間格式中之音頻信號之一編解碼器模式實現較高位元率下之甚至更高QoE，該空間格式並不依靠編碼器中之雙聲道預演現而是表示經攝取之空間夾層格式。取決於位元率，可調整該格式之所表示音頻分量信號之數量。例如，此可導致在自如上文所論述之空間WXY至高解析度空間音頻格式之範圍內之一更有力或較不有力之空間表示。此取決於可用位元率實現低至高空間解析度且提供解決大範圍之演現情景(包含使用標頭追蹤之雙聲道)之靈活性。此模式被稱為「通用空間」模式。

【0059】 在一些實施方案中，IVAS編解碼器以EVS編解碼器之位元率(即，在每秒5.9千位元至128千位元之一範圍中)操作。對於使用在頻寬

限制環境中之傳輸之低速率立體聲操作，可需要低至13.2 kbp之位元率。此要求可能經受使用一特定IVAS編解碼器之技術可行性，且可能仍實現有吸引力之IVAS服務操作。對於使用在頻寬限制環境中之傳輸之低速率立體聲操作，實現空間演現及同時立體聲演現之最低位元率可能低至每秒24.4千位元。對於通用空間模式中之操作，低空間解析度(空間WXY、FOA)有可能低至每秒24.4千位元，然而，在此空間解析度下，可如同空間立體聲操作模式一樣達成音頻品質。

【0060】 現參考圖2B，一接收器件接收包含經編碼音頻信號之一音頻輸送串流。該接收器件之解碼單元250接收(例如，在如藉由一編碼器編碼之一輸送格式中之)經編碼音頻信號且將其解碼。在一些實施方案中，解碼單元250接收在以下四種模式之一者中編碼之音頻信號：單聲道、(習知)立體聲、空間立體聲或通用空間。解碼單元250將音頻信號傳送至演現單元260。演現單元260自解碼單元250接收音頻信號以演現音頻信號。值得注意的是，通常無需恢復被攝取至簡化單元230中之原始第一空間音頻格式。此實現一IVAS解碼器實施方案之解碼器複雜性及/或記憶體佔用面積之顯著節省。

【0061】 圖5係根據本發明之一些實施例之用於將一音頻信號轉換至一可用重播格式之例示性動作的一流程圖。在502，演現單元260接收一第一格式中之一音頻信號。例如，演現單元260可接收以下格式中之該音頻信號：單聲道、習知立體聲、空間立體聲、通用空間。在一些實施方案中，模式選擇單元262接收音頻信號。模式選擇單元262識別音頻信號之格式。若模式選擇單元262判定重播組態支援音頻信號之格式，則模式選擇單元262將音頻信號傳送至演現器264。然而，若模式選擇單元判定

不支援音頻信號，則模式選擇單元執行進一步處理。在一些實施方案中，模式選擇單元262選擇一不同解碼單元。

【0062】 在504，演現單元260判定音頻器件是否能夠重現重播組態支援之一第二格式中之音頻信號。例如，演現單元260可(例如，基於揚聲器及/或其他輸出器件之數量及其等與經解碼音頻相關聯之組態及/或後設資料)判定音頻信號係在空間立體聲格式中，但音頻器件能夠僅重播單聲道中之經接收音頻。在一些實施方案中，並非系統中之所有器件(例如，如圖1中所繪示)能夠重現第一格式中之音頻信號，但所有器件能夠重現一第二格式中之音頻信號。

【0063】 在506，演現單元260基於判定輸出器件能夠重現第二格式中之音頻信號而調適音頻解碼以產生第二格式中之一信號。作為一替代例，演現單元260 (例如，模式選擇單元262或演現器264)可使用後設資料(例如，聲後設資料、轉換後設資料或聲後設資料與轉換後設資料之一組合)以將該音頻信號調適至第二格式。在508，演現單元260傳送經支援之第一格式或經支援之第二格式中之音頻信號以用於音頻輸出(例如，傳送至與一揚聲器系統介接之一驅動器)。

【0064】 在一些實施方案中，演現單元260藉由使用包含第二格式不支援之音頻信號之一部分之一表示之後設資料連同第一格式中之音頻信號而將音頻信號轉變至第二格式。例如，若接收一單聲道格式中之音頻信號且後設資料包含空間格式資訊，則演現單元可使用後設資料將該單聲道格式中之音頻信號轉變至一空間格式。

【0065】 圖6係根據本發明之一些實施例之用於將一音頻信號轉換至一可用重播格式之例示性動作的另一方塊圖。在602，演現單元260接

收一第一格式中之一音頻信號。例如，演現單元260可接收一單聲道、習知立體聲、空間立體聲或通用空間格式中之該音頻信號。在一些實施方案中，模式選擇單元262接收音頻信號。在604，演現單元260擷取音頻器件之音頻輸出能力(例如，音頻重播能力)。例如，演現單元260可擷取揚聲器之一數量、該等揚聲器之位置組態及/或可用於重播之其他重播器件之組態。在一些實施方案中，模式選擇單元262執行該擷取操作。

【0066】 在606，演現單元260比較第一格式之音頻性質與音頻器件之輸出能力。例如，模式選擇單元262可(例如，基於聲後設資料、轉換後設資料或聲後設資料與轉換後設資料之一組合)判定音頻信號係在一空間立體聲格式中且音頻器件能夠經由一立體聲揚聲器系統僅重播習知立體聲格式中之音頻信號(例如，基於揚聲器及其他輸出器件組態)。演現單元260可比較第一格式之音頻性質與音頻器件之輸出能力。在608，演現單元260判定音頻器件之輸出能力是否匹配第一格式之音頻輸出性質。若音頻器件之輸出能力與第一格式之音頻性質不匹配，則程序600移至610，其中演現單元260(例如，模式選擇單元262)執行獲得至一第二格式之音頻信號之動作。例如，演現單元260可調適解碼單元250以解碼第二格式中之經接收音頻或演現單元可使用聲後設資料、轉換後設資料或聲後設資料與轉換後設資料之一組合以將音頻自空間立體聲格式轉換至經支援之第二格式(在給定實例中，其係習知立體聲)。若音頻器件之輸出能力匹配第一格式之音頻輸出性質，或在轉換操作610之後，則程序600移至612，其中演現單元260(例如，使用演現器264)將現確保支援之音頻信號傳送至輸出器件。

【0067】 圖7展示適用於實施本發明之實例性實施例之一實例性系

統700的一方塊圖。如所展示，系統700包含一中央處理單元(CPU) 701，該中央處理單元701能夠根據儲存於(例如)一唯讀記憶體(ROM) 702中之一程式或自(例如)一儲存單元708載入至一隨機存取記憶體(RAM) 703之一程式執行各種程序。在RAM 703中，亦視需要儲存在CPU 701執行各種程序時所需之資料。CPU 701、ROM 702及RAM 703係經由一匯流排704彼此連接。一輸入/輸出(I/O)介面705亦連接至匯流排704。

【0068】 以下組件連接至I/O介面705：一輸入單元706，其可包含一鍵盤、一滑鼠或類似者；一輸出單元707，其可包含一顯示器(諸如一液晶顯示器(LCD))及一或多個揚聲器；儲存單元708，其包含一硬碟或另一合適儲存器件；及一通信單元709，其包含一網路介面卡，諸如一網路卡(例如，有線或無線)。

【0069】 在一些實施方案中，輸入單元706包含不同位置中之一或多個麥克風(取決於主機器件)，從而實現各種格式(例如，單聲道、立體聲、空間、沉浸式及其他合適格式)中之音頻信號的捕獲。

【0070】 在一些實施方案中，輸出單元707包含具有各種數量之揚聲器之系統。如圖1中所繪示，輸出單元707(取決於主機器件之能力)可演現各種格式(例如，單聲道、立體聲、沉浸式、雙聲道及其他合適格式)中之音頻信號。

【0071】 通信單元709經組態以(例如，經由一網路)與其他器件通信。一驅動器710亦視需要連接至I/O介面705。一可移除媒體711(諸如一磁碟、一光學磁碟、一磁光碟、一快閃隨身碟或另一合適可移除媒體)安裝於驅動器710上，使得自其讀取之一電腦程式視需要安裝至儲存單元708中。熟習此項技術者將理解，儘管系統700被描述為包含上述組件，

但在實際應用中，可添加、移除及/或替換此等組件中之一些且所有此等修改或變更全部落在本發明之範疇內。

【0072】 根據本發明之實例性實施例，上文所描述之程序可實施為電腦軟體程式或在一電腦可讀儲存媒體上實施。例如，本發明之實施例包含包括有形地體現於一機器可讀媒體上之一電腦程式之一電腦程式產品，該電腦程式包含用於執行方法之程式碼。在此等實施例中，電腦程式可經由通信單元709自網路下載並安裝，及/或自可移除媒體711安裝。

【0073】 通常，本發明之各種實例性實施例可實施於硬體或專用電路(例如，控制電路)、軟體、邏輯或其等之任何組合中。例如，簡化單元230及上文所論述之其他單元可藉由控制電路(例如，一CPU連同圖7之其他組件)執行，因此，控制電路可執行本發明中所描述之動作。一些態樣可實施於硬體中，而其他態樣可實施於可藉由一控制器、微處理器或其他運算器件(例如，控制電路)執行之韌體或軟體中。雖然本發明之實例性實施例之各項態樣被繪示及描述為方塊圖、流程圖或使用某一其他圖形表示來繪示及描述，但將瞭解，作為非限制性實例，本文中所描述之該等方塊、裝置、系統、技術或方法可實施於硬體、軟體、韌體、專用電路或邏輯、通用硬體或控制器或其他運算器件或其等之某一組合中。

【0074】 此外，流程圖中所展示之各種方塊可被視為方法步驟及/或被視為由電腦程式碼之操作所引起之操作，及/或被視為經建構以實行(若干)相關聯功能之複數個經耦合邏輯電路元件。例如，本發明之實施例包含包括有形地體現於一機器可讀媒體上之一電腦程式之一電腦程式產品，該電腦程式含有經組態以實行如上文所描述之方法之程式碼。

【0075】 在本發明之背景內容中，一機器可讀媒體可為可含有或儲

存一程式以供一指令執行系統、裝置或器件使用或結合該指令執行系統、裝置或器件使用之任何有形媒體。該機器可讀媒體可為一機器可讀信號媒體或一機器可讀儲存媒體。一機器可讀媒體可為非暫時性的且可包含(但不限於)一電子、磁性、光學、電磁、紅外或半導體系統、裝置或器件或前述項之任何合適組合。機器可讀儲存媒體之更特定實例將包含具有一或多個導線之一電連接、一可攜式電腦磁片、一硬碟、一隨機存取記憶體(RAM)、一唯讀記憶體(ROM)、一可擦除可程式化唯讀記憶體(EPROM或快閃記憶體)、一光纖、一可攜式光碟唯讀記憶體(CD-ROM)、一光學儲存器件、一磁性儲存器件或前述項之任何合適組合。

【0076】 用於實行本發明之方法之電腦程式碼可用一或多個程式設計語言之任何組合撰寫。此等電腦程式碼可經提供至一通用電腦、專用電腦或具有控制電路之其他可程式化資料處理裝置之一處理器，使得程式碼在藉由電腦或其他可程式化資料處理裝置之處理器執行時，引起實施流程圖及/或方塊圖中所指定之功能/操作。程式碼可完全在一電腦上、部分在該電腦上、作為一獨立軟體封裝、部分在該電腦上且部分在一遠端電腦上或完全在該遠端電腦或伺服器上執行，或分佈遍及一或多個遠端電腦及/或伺服器。

【符號說明】

【0077】

- 102 呼叫伺服器
- 104 公用交換電話網路(PSTN)/其他公用陸地行動網路(PLMN)
器件、器件、公用交換電話網路(PSTN)/公用陸地行動網路(PLMN)電話

106	舊型使用者設備
108	使用者設備
110	使用者器件
112	電腦器件
114	膝上型電腦
116	會議室使用設備
118	會議室系統
120	家庭影院
122	虛擬實境(VR)裝備
124	沉浸式內容攝取
200	系統
210	捕獲單元
220	聲預處理單元
230	簡化單元
232	音頻格式偵測單元
234	轉換單元
240	編碼單元/編碼器
250	解碼單元
260	演現單元
262	模式選擇單元
264	演現器
300	程序
302	動作

304	動作
306	動作
308	動作
400	程序
402	動作
404	動作
406	動作
408	動作
410	動作
412	動作
502	動作
504	動作
506	動作
508	動作
600	程序
602	動作
604	動作
606	動作
608	動作
610	動作/轉換操作
612	動作
700	系統
701	中央處理單元(CPU)

- 702 唯讀記憶體(ROM)
- 703 隨機存取記憶體(RAM)
- 704 匯流排
- 705 輸入/輸出(I/O)介面
- 706 輸入單元
- 707 輸出單元
- 708 儲存單元
- 709 通信單元
- 710 驅動器
- 711 可移除媒體

【發明申請專利範圍】

【第1項】

一種用於音頻信號處理之方法，其包括：

藉由一發送器件中之一簡化單元自一聲預處理單元接收複數個格式中之音頻信號；

藉由該簡化單元自一接收器件接收該接收器件之屬性(attribute)，該等屬性包含該接收器件支援之一或多個音頻格式之指示，該一或多個音頻格式包含一單聲道格式、一立體聲格式或一空間格式之至少一者；

藉由該簡化單元將該等音頻信號轉變至作為該一或多個音頻格式之一替代表示之一攝取(ingest)格式；

藉由該簡化單元將該經轉變之音頻信號提供至一編碼單元以進行下游(downstream)處理；及

藉由該編碼單元將該攝取格式音頻信號編碼為該接收器件可解碼之一輸送格式中之一經編碼音頻信號，

其中當該一或多個音頻格式包含該空間格式時，該攝取格式對應至一夾層(mezzanine)格式，

其中將該等音頻信號轉變至該攝取格式包含將該等音頻信號轉變至該夾層格式，該夾層格式為該空間格式之一替代表示。

【第2項】

如請求項1之方法，其進一步包括：

傳輸該經編碼音頻信號至該接收器件。

【第3項】

如請求項1之方法，其中將該音頻信號轉變至該攝取格式包括產生用

於該音頻信號之後設資料(metadata)，其中該後設資料包括該音頻信號之一部分之一表示。

【第4項】

如請求項3之方法，其進一步包括藉由傳輸包括該音頻信號之一部分之一表示之該後設資料而傳輸該經編碼音頻信號。

【第5項】

如請求項1之方法，其中該攝取格式將該音頻信號表示為一音頻場景中之音頻物件之一數量，兩者皆依靠用於攜載空間資訊之音頻通道之一數量。

【第6項】

如請求項5之方法，其中該攝取格式進一步包括用於攜載空間資訊之一進一步部分之後設資料。

【第7項】

如請求項1之方法，其中該攝取格式進一步特徵在於實現一可比程度之體驗品質。

【第8項】

一種用於音頻信號處理之裝置，其包括：

一或多個電腦處理器；及

一或多個非暫時性儲存媒體，其或其等儲存指令，該等指令在藉由該一或多個電腦處理器執行時引起該一或多個電腦處理器執行如請求項1之操作。

【第9項】

一種編碼系統，其包括：

一捕獲單元，其經組態以捕獲一音頻信號；

一聲預處理單元，其經組態以執行包括預處理該音頻信號之操作；

一編碼器；及

一簡化單元，其經組態以執行包括以下項之操作：

自該聲預處理單元接收複數個格式之一者中之一音頻信號；

自一接收器件接收該接收器件之屬性，該等屬性包含該接收器件支援之一或多個音頻格式之指示，該一或多個音頻格式包含一單聲道格式、一立體聲格式或一空間格式之至少一者；

將該音頻信號轉變至作為該一或多個音頻格式之一替代表示之一攝取格式；及

將該攝取格式中之該音頻信號傳送至該編碼器，

其中該編碼器經組態以執行包括以下項之操作：

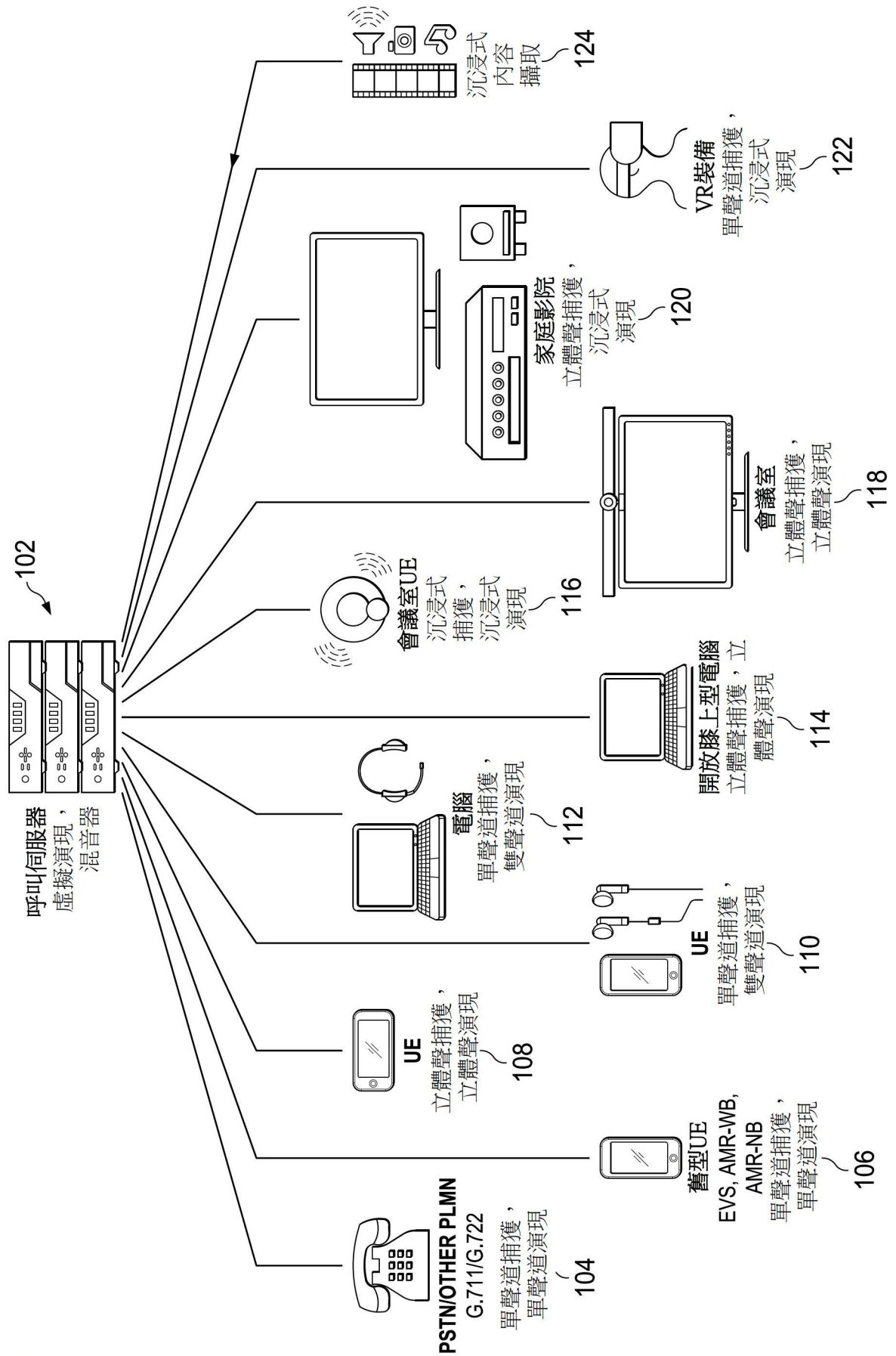
將該攝取格式音頻信號編碼為該接收器件可解碼之一輸送格式中之一經編碼音頻信號；及

將該經編碼音頻信號儲存或傳輸至另一器件，

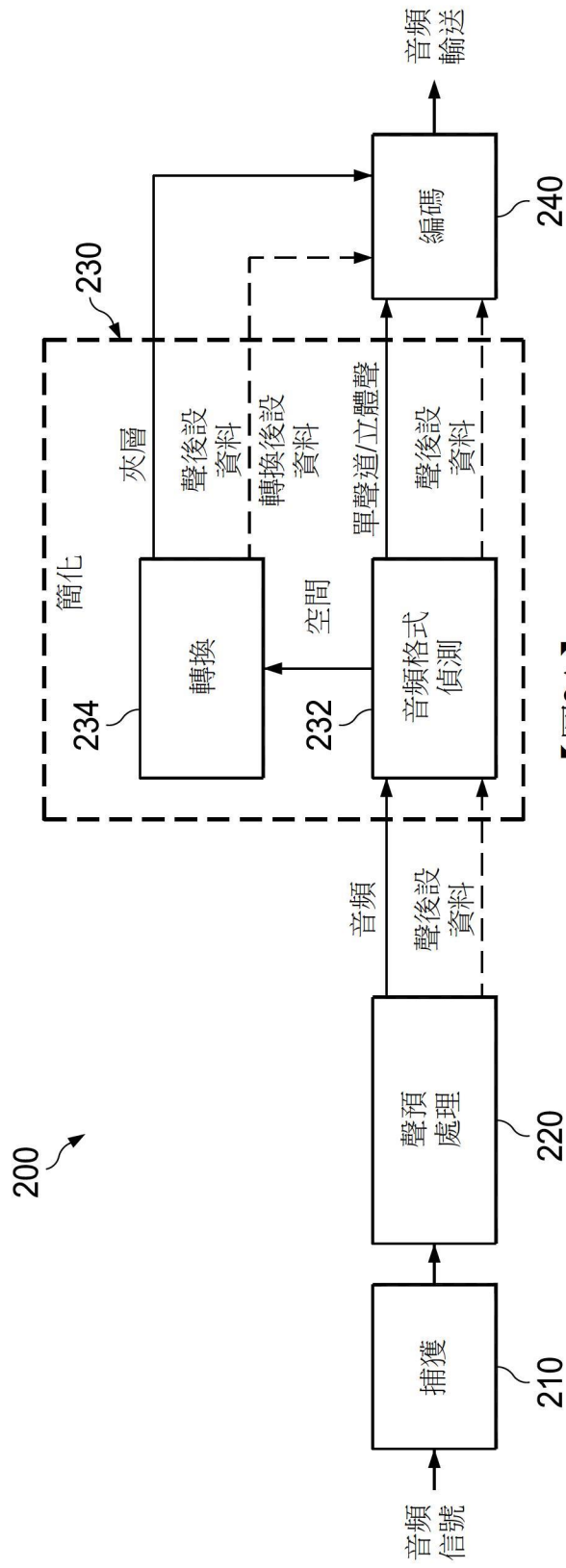
其中當該一或多個音頻格式包含該空間格式時，該攝取格式對應至一夾層格式，

其中將該等音頻信號轉變至該攝取格式包含將該等音頻信號轉變至該夾層格式，該夾層格式為該空間格式之一替代表示。

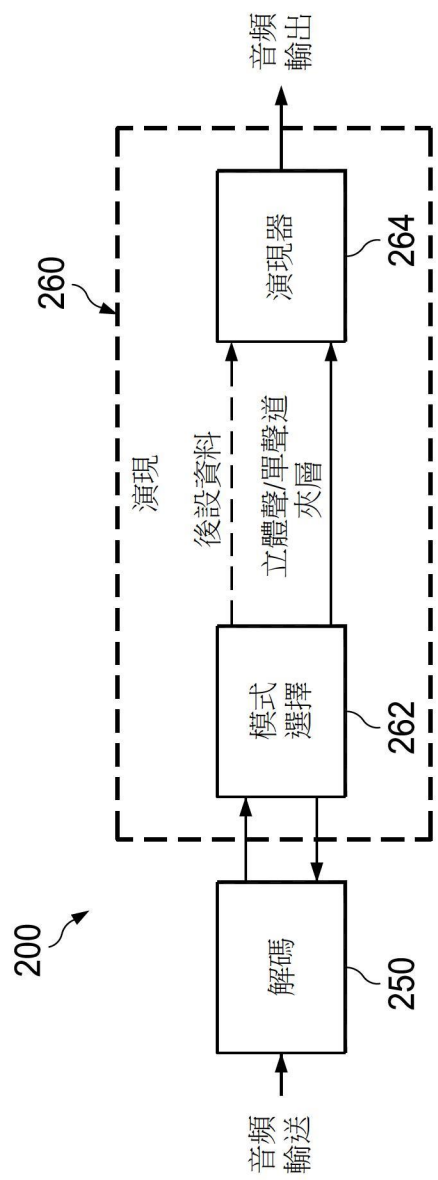
【發明圖式】



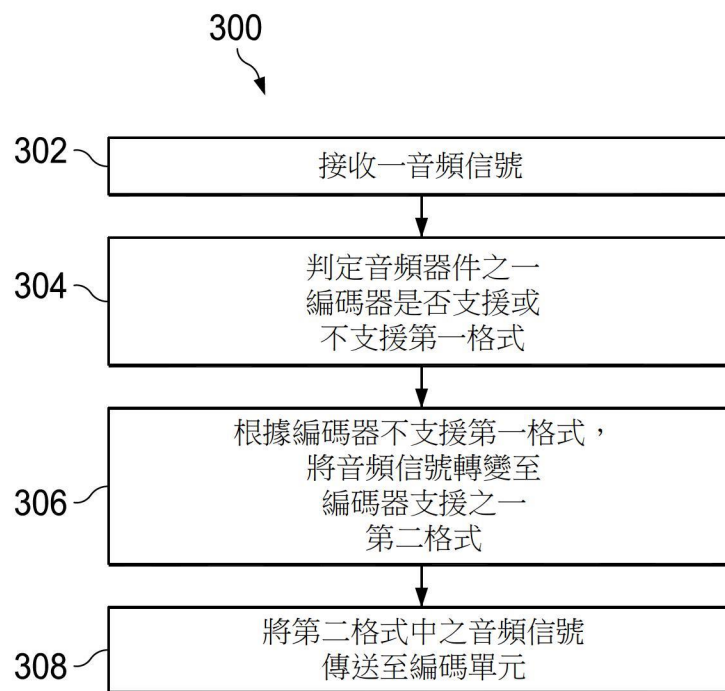
【圖1】



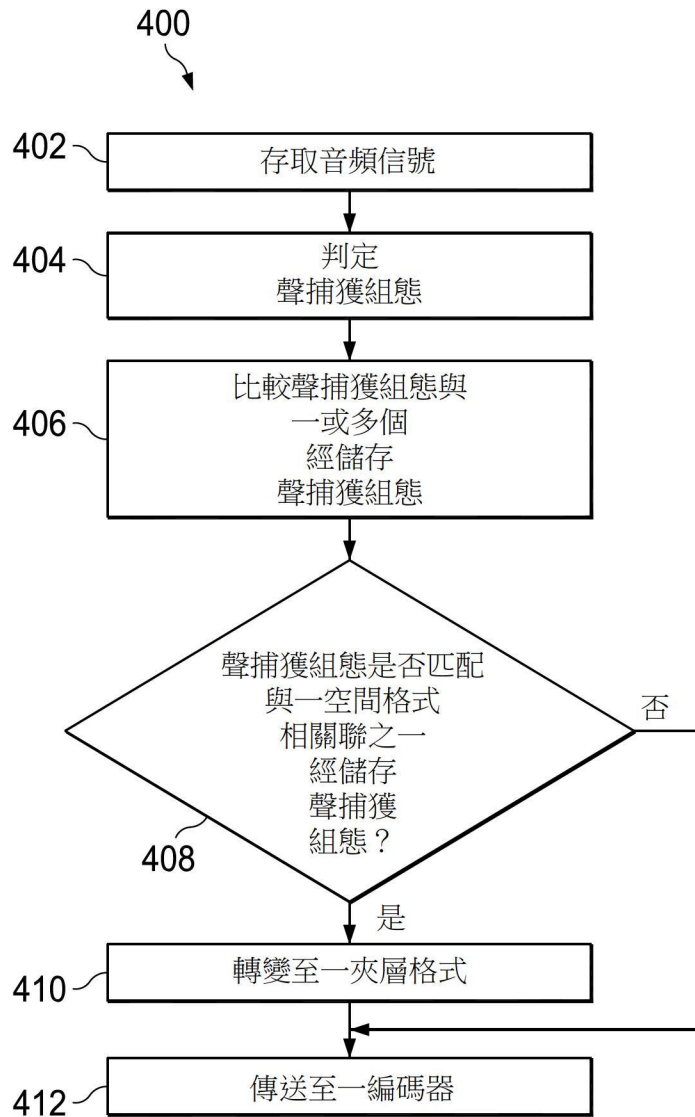
【圖2A】



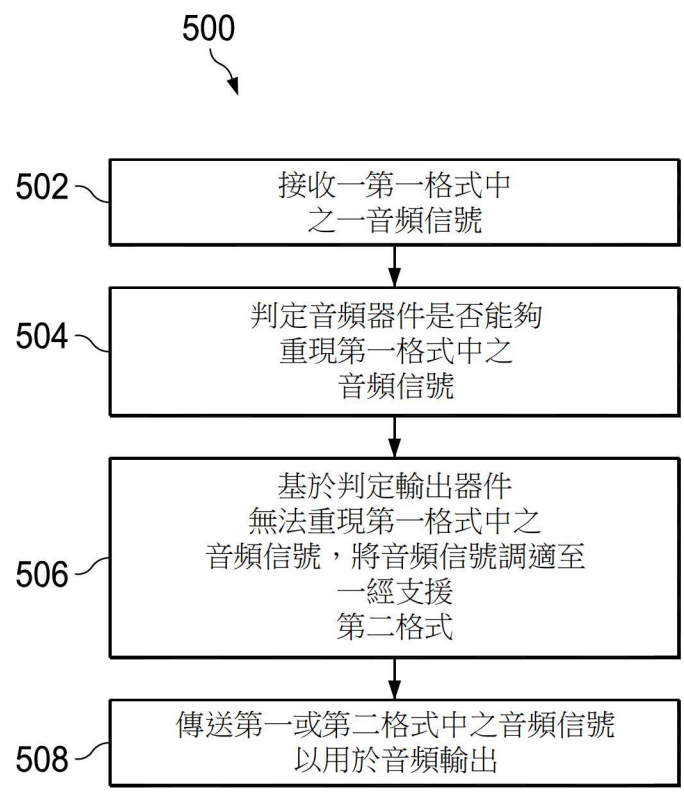
【圖2B】



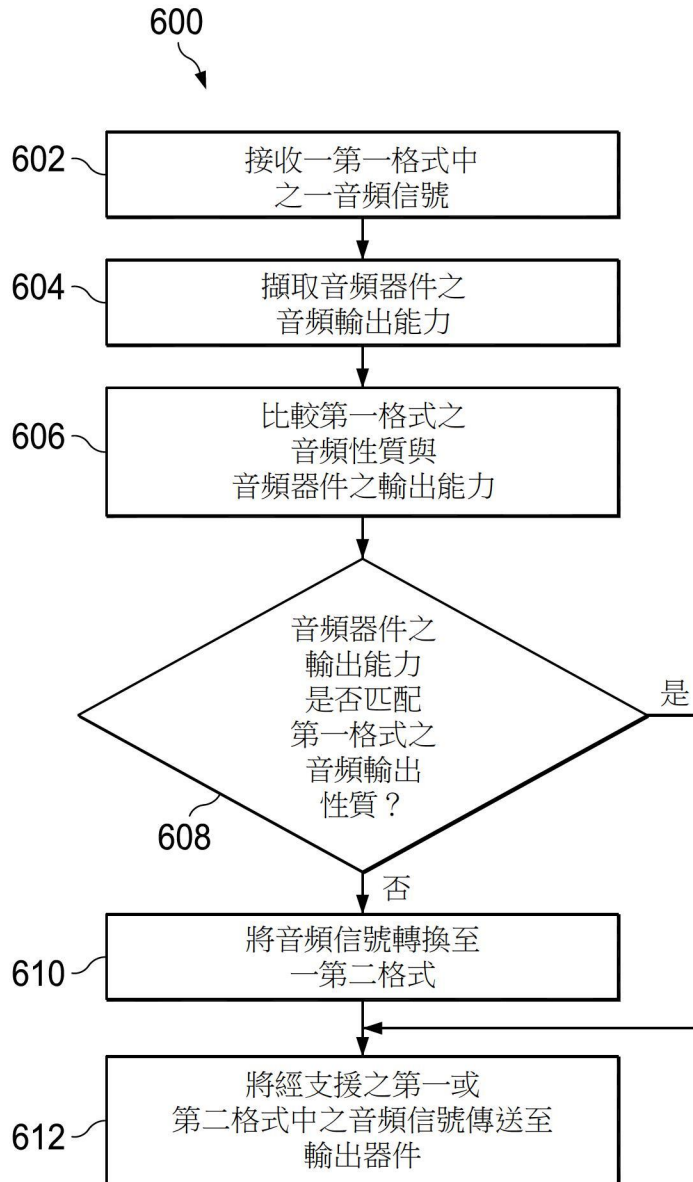
【圖3】



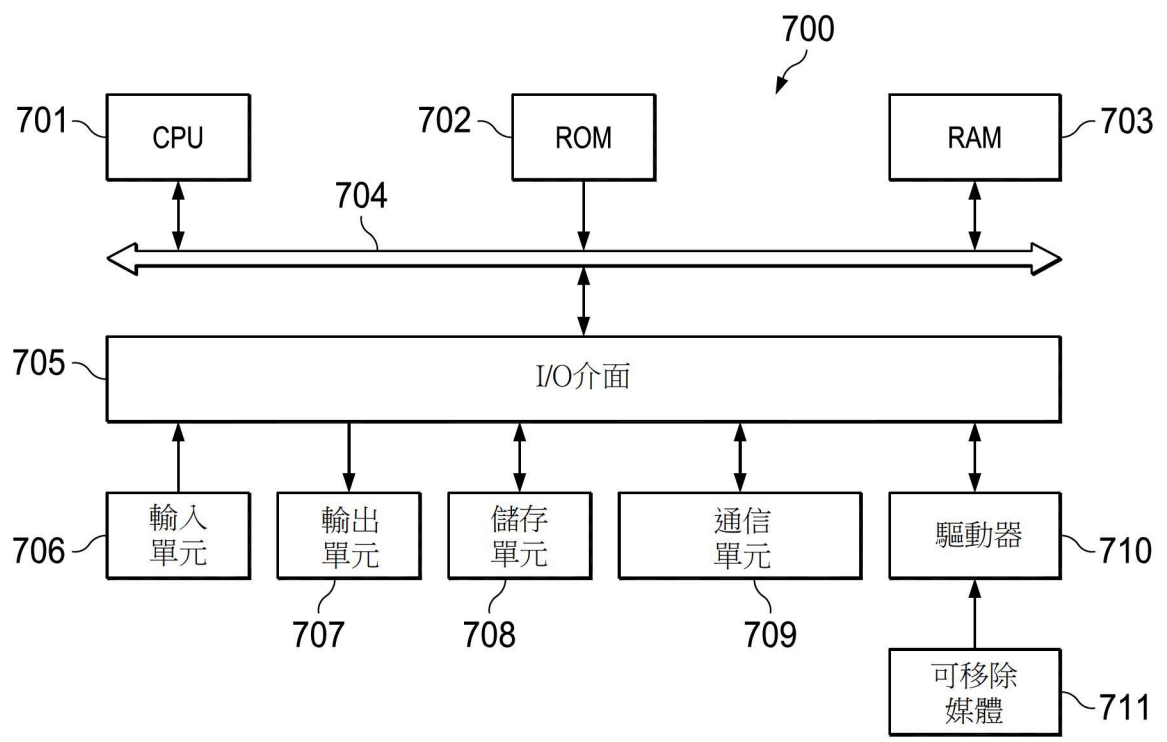
【圖4】



【圖5】



【圖6】



【圖7】