



(21) 申請案號：112100208

(22) 申請日：中華民國 112 (2023) 年 01 月 04 日

(51) Int. Cl. : H04N5/225 (2006.01)

H04N5/14 (2006.01)

H04N7/15 (2006.01)

H04L65/1059(2022.01)

H04L65/10 (2022.01)

(30) 優先權：2022/03/04 美國

17/686,805

(71) 申請人：美商高通公司 (美國) QUALCOMM INCORPORATED (US)

美國

(72) 發明人：巴爾德 安什阿拜 BALDE, ANSH ABHAY (IN) ; 阿南德 桑托什 ANAND,

SANTOSH (IN)

(74) 代理人：李世章

申請實體審查：無 申請專利範圍項數：32 項 圖式數：8 共 86 頁

(54) 名稱

用於設備上的多個物件的視訊呼叫體驗

(57) 摘要

提供了用於視訊呼叫的系統、方法和電腦可讀取媒體。一種示例方法可包括：在第一設備與第二設備之間建立視訊呼叫；顯示第一相機饋送和第二相機饋送的預覽，該第一相機饋送包括由該第一設備的第一影像擷取裝置擷取的第一視訊訊框和由該第一設備的第二影像擷取裝置擷取的第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；接收對圖示在該預覽中的物件集合的選擇；及基於該第一相機饋送和該第二相機饋送來產生圖示該物件集合的單個訊框。

Systems, methods, and computer-readable media are provided for video calling. An example method can include establishing a video call between a first device and a second device; displaying a preview of a first camera feed and a second camera feed, the first camera feed including a first video frame captured by a first image capture device of the first device and a second video frame captured by a second image capture device of the first device, the first video frame and the second video frame being visually separated within the preview; receiving a selection of a set of subjects depicted in the preview; and generating, based on the first camera feed and the second camera feed, a single frame depicting the set of subjects.

指定代表圖：

符號簡單說明：

- 102:端點
- 202:通訊部件
- 204:計算部件
- 206:輸入裝置
- 208:輸出設備
- 210:相機感測器
- 212:儲存裝置
- 214:影像處理系統
- 216:訊框調節器
- 218:訊框分離器
- 220:視訊轉碼器
- 222:渲染系統

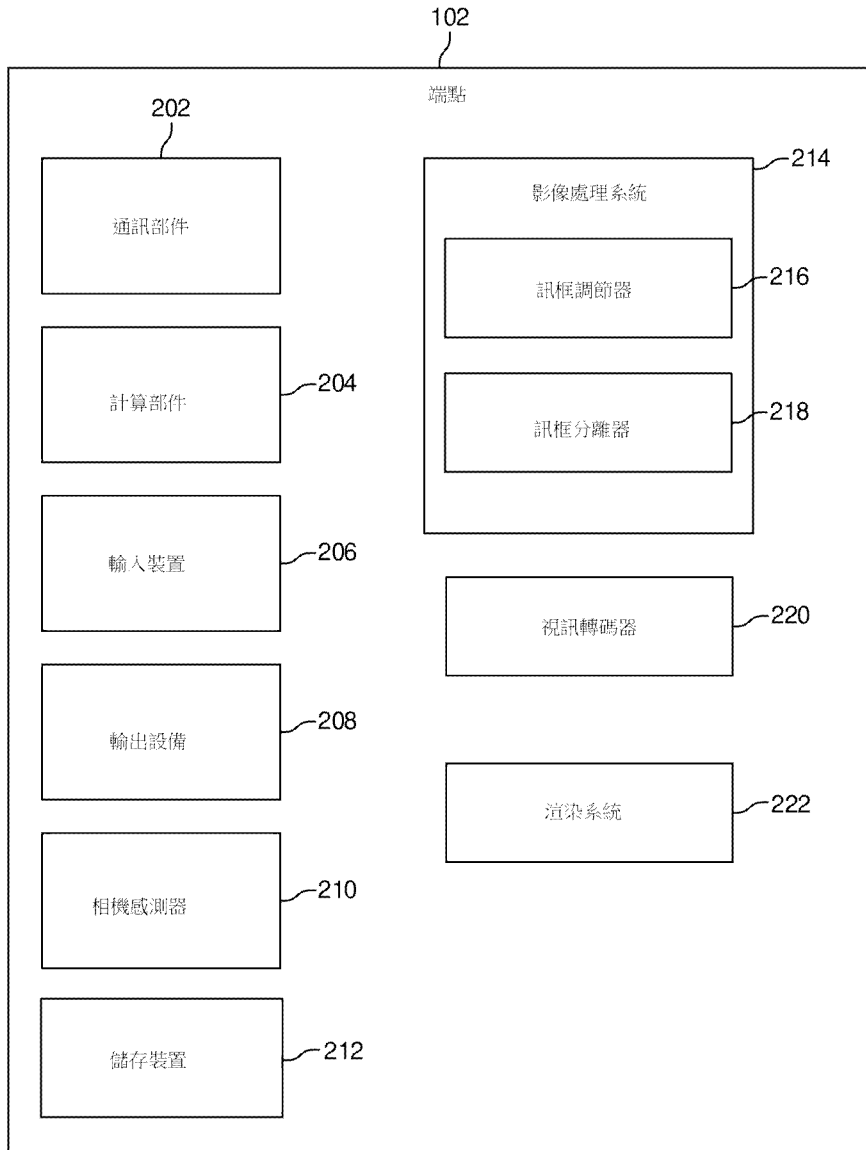


圖2

【發明摘要】

【中文發明名稱】用於設備上的多個物件的視訊呼叫體驗

【英文發明名稱】VIDEO CALLING EXPERIENCE FOR MULTIPLE SUBJECTS
ON A DEVICE

【中文】

提供了用於視訊呼叫的系統、方法和電腦可讀取媒體。一種示例方法可包括：在第一設備與第二設備之間建立視訊呼叫；顯示第一相機饋送和第二相機饋送的預覽，該第一相機饋送包括由該第一設備的第一影像擷取裝置擷取的第一視訊訊框和由該第一設備的第二影像擷取裝置擷取的第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；接收對圖示在該預覽中的物件集合的選擇；及基於該第一相機饋送和該第二相機饋送來產生圖示該物件集合的單個訊框。

【英文】

Systems, methods, and computer-readable media are provided for video calling. An example method can include establishing a video call between a first device and a second device; displaying a preview of a first camera feed and a second camera feed, the first camera feed including a first video frame captured by a first image capture device of the first device and a second video frame captured by a second image capture device of the first device, the first video frame and the second video frame being visually separated within the preview; receiving a selection of a set of subjects depicted in the preview; and generating, based on the first camera feed and the second camera feed, a single frame depicting the set of subjects.

【指定代表圖】第（ 2 ）圖。

【代表圖之符號簡單說明】

1 0 2 : 端 點

2 0 2 : 通 訊 部 件

2 0 4 : 計 算 部 件

2 0 6 : 輸 入 裝 置

2 0 8 : 輸 出 設 備

2 1 0 : 相 機 感 測 器

2 1 2 : 儲 存 裝 置

2 1 4 : 影 像 處 理 系 統

2 1 6 : 訊 框 調 節 器

2 1 8 : 訊 框 分 離 器

2 2 0 : 視 訊 轉 碼 器

2 2 2 : 渲 染 系 統

【特徵化學式】

無

【發明說明書】

【中文發明名稱】用於設備上的多個物件的視訊呼叫體驗

【英文發明名稱】VIDEO CALLING EXPERIENCE FOR MULTIPLE SUBJECTS
ON A DEVICE

【技術領域】

【0001】 本案整體上涉及視訊呼叫系統。例如，本案的各態樣涉及用於一個設備上的多個物件（subject）的視訊呼叫體驗。

【先前技術】

【0002】 電話系統已經成為遠端使用者進行通訊的普遍存在的工具。當今，電話系統通常支援許多共同特徵，諸如語音通訊、視訊通訊、呼叫等待、呼叫保持和會議等。通常，視訊呼叫的效能品質可取決於用於建立和進行視訊呼叫的網路連接的品質。然而，即使在更快和更可靠的通訊協定和技術變得越來越常見和可用的情況下，網路問題仍然普遍。例如，視訊呼叫的穩定性/連續性經常受到網路問題的阻礙。網路問題可能導致在呼叫（例如，兩個使用者之間的呼叫、會議呼叫等）期間丟棄訊框，此可能中斷呼叫並且迫使呼叫中的使用者等待問題解決以便繼續呼叫。

【0003】 除了連接/網路問題之外，在視訊呼叫期間使用的相機的放置可能引起問題，諸如例如呼叫的參與者在相機的視場（FOV）之外，並且因此從由相機擷取的視訊訊框中被排除，從而阻礙呼叫另一端的使用者看到該呼叫參與者。作為另一實例，取決於相機的放置，相機可擷取無意

被擷取並且被包括在於呼叫期間被發送的視訊訊框中的其他使用者或干擾。

【發明內容】

【0004】 揭示用於設備上的多個物件的視訊呼叫體驗的系統、方法和電腦可讀取媒體。根據至少一個實例，提供了一種用於設備上的多個物件的視訊呼叫體驗的方法。該方法可包括：在設備與遠端設備之間建立視訊呼叫；顯示第一相機饋送和第二相機饋送的預覽，該第一相機饋送包括由該設備的第一影像擷取裝置擷取的第一視訊訊框和由該設備的第二影像擷取裝置擷取的第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；接收對圖示在該預覽中的物件集合的選擇；及基於該第一相機饋送和該第二相機饋送來產生繪示該物件集合的單個訊框。

【0005】 根據至少一個實例，提供了一種裝置，用於該裝置上的多個物件的視訊呼叫體驗。該裝置可包括：記憶體，及耦合到該記憶體的一或多個處理器，該一或多個處理器被配置為：在設備與遠端設備之間建立視訊呼叫；顯示第一相機饋送和第二相機饋送的預覽，該第一相機饋送包括由該設備的第一影像擷取裝置擷取的第一視訊訊框和由該設備的第二影像擷取裝置擷取的第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；接收對圖示在該預覽中的物件集合的選擇；及基於該第一相

機饋送和該第二相機饋送來產生繪示該物件集合的單個訊框。

【0006】 根據至少一個實例，提供了另一種裝置，用於該裝置上的多個物件的視訊呼叫體驗。該裝置可包括用於以下的手段：在設備與遠端設備之間建立視訊呼叫；顯示第一相機饋送和第二相機饋送的預覽，該第一相機饋送包括由該設備的第一影像擷取裝置擷取的第一視訊訊框和由該設備的第二影像擷取裝置擷取的第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；接收對圖示在該預覽中的物件集合的選擇；及基於該第一相機饋送和該第二相機饋送來產生繪示該物件集合的單個訊框。

【0007】 根據至少一個實例，提供了一種用於裝置上的多個物件的視訊呼叫體驗的非暫時性電腦可讀取媒體。該非暫時性電腦可讀取媒體可包括指令，該等指令在由一或多個處理器執行時使得該一或多個處理器：在設備與遠端設備之間建立視訊呼叫；顯示第一相機饋送和第二相機饋送的預覽，該第一相機饋送包括由該設備的第一影像擷取裝置擷取的第一視訊訊框和由該設備的第二影像擷取裝置擷取的第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；接收對圖示在該預覽中的物件集合的選擇；及基於該第一相機饋送和該第二相機饋送來產生繪示該物件集合的單個訊框。

【0008】 在一些態樣，上文描述的該方法、非暫時性電腦可讀取媒體和裝置可從該第一影像擷取裝置獲得該第一視訊訊框及從該第二影像擷取裝置獲得該第二視訊訊框。

【0009】 在一些實例中，該第一視訊訊框和該第二視訊訊框繪示包括該物件集合的多個物件，及對該物件集合的該選擇可包括選擇該物件集合作為該多個物件中要被包括在該單個訊框中的子集的第一輸入及/或選擇該多個物件中要從該單個訊框中被排除的一或多個物件的第二輸入。在一些實例中，該一或多個物件可不同於該物件集合。

【0010】 在一些實例中，產生該單個訊框可包括基於該第一輸入及/或該第二輸入，從該單個訊框中排除該多個物件中的該一或多個物件；及向該遠端設備發出該單個訊框。在一些實例中，從該單個訊框中排除該一或多個物件可包括從該預覽、該第一視訊訊框、該第二視訊訊框及/或該單個訊框中移除該一或多個物件。

【0011】 在一些態樣，產生該單個訊框可包括將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框。在一些實例中，將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框可包括將該第一視訊訊框的至少該部分和該第二視訊訊框的至少該部分佈置成該單個訊框的相應訊框區域。在一些實例中，每個訊框區域可繪示來自該物件集合中的相應物件。

【0012】 在一些實例中，該第一視訊訊框可包括來自該物件集合的一或多個第一物件，及該第二視訊訊框可包括來自該物件集合的一或多個第二物件。在一些態樣，產生該單個訊框可包括基於與該第一視訊訊框相關聯的第一中繼資料來決定該一或多個第一物件之每一第一物件在該第一視訊訊框內的相應位置；基於與該第二視訊訊框相關聯的第二中繼資料來決定該一或多個第二物件之每一第二物件在該第二視訊訊框內的相應位置；及基於該一或多個第一物件之每一第一物件在該第一視訊訊框內的該相應位置及該一或多個第二物件之每一第二物件在該第二視訊訊框內的該相應位置，決定圖示該一或多個第一物件的該第一視訊訊框的第一部分和圖示該一或多個第二物件的該第二視訊訊框的第二部分。在一些實例中，該第一中繼資料可包括與該一或多個第一物件相關聯的座標。在一些實例中，該第二中繼資料可包括與該一或多個第二物件相關聯的座標。

【0013】 在一些實例中，產生該單個訊框可包括將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框。在一些情況下，將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框可包括將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分放置在該單個訊框的相應訊框區域內。在一些實例中，該相應訊框區域中的第一區域繪示該一或多

個第一物件，及該相應訊框區域中的第二區域繪示該一或多個第二物件。

【0014】 在一些實例中，產生該單個訊框可包括在該單個訊框內渲染該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分。在一些情況下，該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分在視覺上被分開。在一些情況下，該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分經由視覺標記在視覺上被分開。在一些實例中，該視覺標記可包括線、輪廓、方塊、高亮、標籤、顏色、陰影及/或視覺記號（*indicia*）。

【0015】 在一些態樣，上文描述的該方法、非暫時性電腦可讀取媒體和裝置可決定該一或多個物件在該第一影像擷取裝置的視場（*FOV*）內；及觸發該第一影像擷取裝置以擷取該第一視訊訊框。在一些實例中，該第一視訊訊框圖示該一或多個物件。

【0016】 在一些態樣，上文描述的該方法、非暫時性電腦可讀取媒體和裝置可基於由該第一影像擷取裝置擷取的一或多個視訊訊框來追蹤該一或多個物件；基於該追蹤，決定該一或多個物件不再位於該第一影像擷取裝置的該 *FOV* 內；及基於決定該一或多個物件不再位於該第一影像擷取裝置的該 *FOV* 內，決定停止從該第一影像擷取裝置向該遠端設備發送視訊資料。

【0017】 在一些態樣，上文描述的該方法、非暫時性電腦可讀取媒體和裝置可回應於決定至少一個物件在該第二影

像擷取裝置的 F O V 內，經由該第二影像擷取裝置獲得圖示該至少一個物件的第三視訊訊框；及向該遠端設備發出該第三視訊訊框。

【0018】 在一些態樣，上文描述的該裝置可包括一或多個感測器。在一些實例中，上文描述的該裝置可包括行動電話、會議系統、電話系統、可穿戴設備、顯示裝置、行動電腦、頭戴式顯示器及/或相機。

【0019】 在一些態樣，上文描述的該裝置中的一或多個是、可是其一部分，或者可包括物聯網路（I o T）設備及/或擴展現實（X R）設備（例如，虛擬實境（V R）設備、增強現實（A R）設備或混合現實（M R）設備）。在一些實例中，該裝置可包括或作為車輛、行動設備（例如，行動電話或所謂的「智慧型電話」或其他行動設備）、可穿戴設備、個人電腦、膝上型電腦、平板電腦、伺服器電腦、機器人設備或系統、航空系統或其他設備的一部分。在一些態樣，該裝置包括用於擷取一或多個影像的影像感測器（例如，相機）或多個影像感測器（例如，多個相機）。在一些態樣，該裝置包括用於顯示一或多個影像、通知及/或其他可顯示資料的一或多個顯示器。在一些態樣，該裝置包括一或多個揚聲器、一或多個發光設備及/或一或多個麥克風。在一些態樣，上文描述的該裝置可包括一或多個感測器。例如，該一或多個感測器可包括光感測器、音訊感測器、運動感測器、溫度感測器、濕度感測器、影像感測器、加速度計、陀螺儀、壓力感測器、觸摸感測器和磁

力計中的至少一個。在一些情況下，該一或多個感測器可被用於決定該裝置的位置、該裝置的狀態，及/或用於其他目的。

【0020】 本發明內容既非意欲標識所主張的標的的關鍵或本質特徵，亦非意欲單獨用於決定所主張的標的的範圍。應經由參考本專利的整個說明書的適當部分、任何或所有附圖和每個請求項來理解標的。

【0021】 在參考以下說明書、申請專利範圍和附圖時，前述連同其他特徵和態樣將變得更加顯而易見。

【圖式簡單說明】

【0022】 下文參考以下附圖詳細描述本案的說明性實例：

【0023】 圖 1 A 是圖示根據本案的一些實例的使用者之間的示例視訊呼叫的圖；

【0024】 圖 1 B 是圖示根據本案的一些實例的使用者之間的示例視訊會議的圖；

【0025】 圖 2 圖示了根據本案的一些實例的可實施視訊呼叫體驗的端點的實例；

【0026】 圖 3 是圖示根據本案的一些實例的用於提供示例視訊呼叫體驗的示例流水線的圖；

【0027】 圖 4 圖示了根據本案的一些實例的繪示渲染在接收者的端點處的合併的相機訊框的示例視訊呼叫介面；

【0028】 圖 5 圖示了根據本案的一些實例的繪示渲染在接收者的端點處的合併的相機訊框的另一示例視訊呼叫介面；

【0029】 圖 6 A 圖示了根據本案的一些實例的圖示由發送者的端點產生的示例預覽的示例使用者介面；

【0030】 圖 6 B 圖示了根據本案的一些實例的圖示從發送者的端點接收的合併的訊框的示例使用者介面；

【0031】 圖 7 是圖示根據本案的一些實例的用於產生視訊呼叫的視訊訊框的示例程序的流程圖；及

【0032】 圖 8 圖示了根據本案的一些實例的示例計算設備架構。

【實施方式】

【0033】 下文提供本案的某些態樣。該等態樣中的一些可獨立地應用，並且其中的一些可組合應用，此對於本領域技藝人士是顯而易見的。在以下描述中，出於解釋的目的，闡述了具體細節以便提供對本案各態樣的透徹理解。然而，顯而易見的是，可在沒有該等具體細節的情況下實踐各個態樣。附圖和說明並不意欲是限制性的。

【0034】 隨後的描述僅提供示例性態樣，並不意欲限制本案的範圍、適用性或配置。相反，示例態樣的隨後描述將向本領域技藝人士提供用於實施示例態樣的使能描述。應當理解，在不脫離所附請求項中闡述的本案的精神和範圍的情況下，可對元件的功能和佈置進行各種改變。

【0035】 如先前所解釋，視訊呼叫的品質、可靠性、效能等可能受到諸如網路問題、相機放置、相機能力等許多因素的負面影響。在一個說明性實例中，在使用者 A 與使用者 B 之間的視訊呼叫期間，使用者 C 走進使用者 B 的房間/空間

中並且想要成為使用者 A 與使用者 B 之間的視訊呼叫對話的一部分。在該實例中，使用者 B 隨後需要經由確保使用者 C 在使用者 B 的設備上的特定相機（例如，前置相機）的視場（FOV）內來容納使用者 C，使得使用者 B 和使用者 C 兩者皆可被來自使用者 B 的設備的相機錄製的視訊擷取。替代地，使用者 B 可打開（例如，為了擷取、記錄及 / 或傳輸的目的而啟動及 / 或利用）或轉換到使用者 B 的設備的附加相機（例如，後置相機），在此種情況下，只有使用者 C 可是視訊的一部分並且在視訊呼叫的另一端被使用者 A 看到。此可能對使用者 A 的視訊呼叫體驗不利，因為使用者 A 不能與使用者 B 和使用者 C 兩者進行預期的對話，使得使用者 B 和使用者 C 兩者同時對使用者 A 可見（例如，在視訊呼叫期間在同一視訊訊框內或同時顯示的視訊訊框內）。

【0036】 作為另一實例，若在使用者 A 與使用者 B 之間的視訊呼叫期間，使用者 A 想要四處移動，而使用者 A 的設備保持在靜止位置，則使用者 A 可能不再處於使用者 A 的設備上的特定相機的 FOV 內。因此，使用者 A 將不會在發出給使用者 B 的記錄的視訊訊框中被擷取，並且使用者 B 將不能在來自呼叫的視訊中看到使用者 A。當使用者 A 從使用者 A 的設備的第一相機的 FOV（例如，前置相機的 FOV）移動到第二相機的 FOV（例如，使用者 A 的設備的後置相機的 FOV）時，使用者 A 需要手動地將相機切換到用於視訊呼叫的第二相機，使得使用者 A 被來自第二相機的視訊訊框擷取。替代地，使用者 A 可返回到第一相機的 FOV 內的區域，

以由來自使用者 A 的設備上的第一相機的記錄的視訊訊框擷取。例如，若使用者 A 完全移出使用者 A 的設備上的所有相機（例如，前置相機和後置相機）的 FOV，但是想要繼續對話（例如，使用耳機），則使用者 A 將能夠繼續音訊對話，但是將不會在由使用者 A 的設備記錄的視訊訊框中被擷取，並且因此將不會對視訊呼叫的另一端上的任何使用者（例如，使用者 B）可見。

【0037】 本文描述了用於使用同一設備參與視訊呼叫的多個物件的改進的視訊呼叫體驗的系統、裝置、程序（亦稱為方法）和電腦可讀取媒體（本文統稱為「系統和技術」）。在一些實例中，本文所描述的系統和技術可在視訊呼叫期間開啟（及/或保持開啟）由多個使用者共用的設備的多個相機（例如，前置相機和後置相機），以允許相機（例如，前置相機和後置相機）從同一設備擷取視訊呼叫上的兩個使用者。在一些實例中，該設備可追蹤該使用者中的任何使用者並且進行任何調整，以確保即使該使用者中的一或多個使用者在視訊呼叫期間移動，兩個使用者亦被記錄的視訊訊框擷取到。

【0038】 在一些實例中，本文描述的系統和技術可使用較低功率的面部追蹤模式（例如，相對於在使用較高功率資源及/或實施較高功率/複雜度軟體部件的設備處可用的較高功率的面部追蹤模式）來在視訊呼叫期間追蹤使用者。說明如下：在使用者 C 在視訊呼叫期間走進使用者 B 的房間/空間並且想要參與視訊呼叫的先前實例中，系統可擷取使

用者 B 和使用 者 C 兩者的訊框，將圖示使用者 B 和使用者 C 的訊框合併或組合成單個訊框，並且向視訊呼叫的另一端上的使用者 A 的設備發出該單個訊框。例如，系統可佈置使用者 B 和使用者 C 的訊框，使得一個訊框（例如，擷取使用者 B 的訊框）在另一個訊框（例如，擷取使用者 C 的訊框）的上方。系統隨後可將所得的訊框發送給使用者 A 的設備（例如，作為視訊呼叫的一部分）。在一些實例中，系統可將訊框拼接成單個訊框以容納從同一設備參與視訊呼叫的兩個物件（例如，使用者 B 和使用者 C）。

【0039】 在一些情況下，若在使用者 B 的設備的第一相機（例如，後置相機）的 FOV 中存在兩個物件，則系統可將訊框與第一相機分離，並且將其作為與在使用者 B 的設備的第二相機（例如，前置相機）的 FOV 中擷取物件的訊框縫合在一起的分離的訊框來發出。例如，若使用者 C 和使用者 D 在設備的第一相機（例如，後置相機）的 FOV 內移動，而使用者 B 在第二相機（例如，前置相機）的 FOV 內，則使用者 B 通常需要將視訊從第二相機切換到第一相機以示出使用者 C 和使用者 D。在該示例場景中，若使用者 C 和使用者 D 相距很遠（例如，使得使用者 C 和 D 不在特定相機的 FOV 內），則使用者 C 和使用者 D 的多個面部特徵可能不被擷取並且顯示在視訊呼叫中。然而，本文描述的系統和技術可將描述使用者 C 和使用者 D 的訊框分離為分開的字元，如本文進一步描述的。

【0040】 在一些情況下，使用設備參與視訊呼叫的使用者可調用（例如，觸發、請求、發起等）相機切換視訊呼叫使用者介面（UI）。設備的多個相機（例如，前置相機和後置相機）可被打開，並且可將所擷取的訊框發出到設備的處理器（例如，數位訊號處理器等），隨後可將其發出到顯示器以供發送者預覽該訊框（例如，使用相機切換呼叫UI）。在一些實例中，設備可以分離的方式（例如，由線、字元、物件、形狀等分離）發出來自兩個相機的饋送。該系統可實現訊框分離器以分離所擷取的使用者（或使用者的面部）並且在預覽UI上顯示上述各者。

【0041】 在一些實例中，發送者（例如，向視訊呼叫中的其他使用者發出視訊饋送的（多個）使用者）可向系統提供使用者輸入，該使用者輸入指示發送者想要在給定訊框中包括哪些使用者或面對哪些使用者的選擇及/或想要在預覽UI的特定側或區域上發送預覽。在一些情況下，系統可輸出（例如，顯示、播放音訊，或以其他方式輸出）要求發送者確認輸入的使用者或面部的許可訊息。隨後可將對應於所選擇使用者或面部的影像資料提供給編碼器以進行編碼，並且經由視訊呼叫流水線將其發送到接收者。接收者可經由與發送者相同的設備接收描述發送者及參與視訊呼叫的附加使用者的視訊訊框。

【0042】 在一些實例中，可使用調節器元件來調節第一相機（例如，前置相機）和第二相機（例如，後置相機）饋送。例如，調節器元件可檢查物件的訊框並且經由僅在存

在物件時發出相機饋送來減少編碼器工作負載。在一些情況下，相機饋送可包括時間戳記、緩衝器資訊及/或中繼資料。在一些實例中，中繼資料可包括面部座標及/或面部座標中繼資料標籤。面部座標可用於分離視訊呼叫中的面部。在使用視訊饋送之前，系統可調用使用者輸入流水線以決定使用者想要從使用者的設備的相機上的使用者可用FOV發出什麼面部。該資訊可被中繼到調節器和訊框分離器。調節器可經由消除使用者不想包括在呼叫中的面部來調節視訊呼叫，並且訊框分離器可根據包括的面部的數目將多個訊框重構為單個訊框。在一些實例中，可在單個訊框內佈置或重新排序兩個面部，可佈置三個面部，使得一個面部在第一區域（例如，頂部區域）上，而其他兩個面部在一或多個其他區域（例如，底部區域或任何其他佈置）中，等等。

【0043】 在一些情況下，本文描述的系統和技術可實現物件追蹤模式。例如，系統可僅在使用者存在於使用者設備的任一/任何相機（例如，諸如前置相機的第一相機、諸如後置相機的第二相機等）的FOV中時發出訊框。這可包括物件追蹤。在系統偵測到面部特徵時，當使用者在另一相機（例如，諸如後置相機的第二相機）的FOV內移動時，相機可停止從一個相機（例如，諸如前置相機的第一相機）發出訊框。因此，當（多個）使用者在相機的FOV內時，系統可發出從該相機擷取的訊框，而當（多個）使用者在

不同相機的FOV內移動時，系統可發出由不同相機擷取的訊框。

【0044】 圖1A圖示了使用者112與114之間的示例視訊呼叫100。視訊呼叫100可包括及/或支援一或多個呼叫特徵，諸如例如音訊、視訊、呼叫保持、呼叫等待、會議、互動式語音回應（IVR）、雙音多頻訊號傳遞（DTMF）、螢幕共用、邊聊天（side chat）及/或任何其他呼叫特徵。

【0045】 使用者112可經由端點102在視訊呼叫100上與使用者114通訊，並且使用者114可經由端點104與使用者112通訊。端點102和104中的每一個可包括任何計算設備或電話系統，諸如例如智慧型電話、智慧可穿戴設備（例如，智慧手錶、智慧護目鏡）、智慧電視、遊戲系統、平板電腦、IP（網際網路協定）電話、膝上型電腦、車載電話、臺式電話、IoT（物聯網）設備等。

【0046】 端點102和104可支援視訊和語音/音訊。在一些情況下，端點102和104亦可支援其他特徵，諸如例如但不限於面部偵測/辨識、對象偵測/辨識、追蹤、DTMF、自動話語處理（ASR）、關鍵字偵測、活動說話者偵測及/或任何其他呼叫及/或影像處理功能性。在一些實例中，端點102和104可支援一或多個其他呼叫及/或多媒體特徵，例如例如呼叫會議、呼叫等待、語音郵件及/或任何其他呼叫或多媒體特徵。

【0047】 端點102和104可經由網路110建立視訊呼叫100。網路110可表示一或多個公共及/或私人網路絡。網

路 110 可包括例如公共交換電信網路 (PTSN)、無線網路 (例如蜂巢網路、無線區域網路、衛星通訊網路、微波網路等)、網際網路協定 (IP) 網路、呼叫提供商網路、託管呼叫及 / 或會議伺服的雲網路、綜合伺服數位網路 (ISDN)、虛擬網路、電話伺服供應商網路、陸線網路及 / 或任何其他類型的網路。

【0048】 圖 1B 圖示了使用者 112-116 之間的示例會議呼叫 120。會議呼叫 120 可包括及 / 或支援一或多個呼叫特徵，諸如例如音訊、視訊、呼叫保持、語音訊息傳遞、呼叫等待、IVR、DTMF、螢幕共用、呼叫記錄，及 / 或本文描述的任何其他呼叫特徵。在一些實例中，會議呼叫 120 可包括及 / 或支援其他特徵，諸如例如內容共用、即時訊息收發、桌面或螢幕共用、聊天、在場共用、對象偵測 / 辨識、面部偵測 / 辨識、追蹤、影像處理及 / 或其他類型的媒體和通訊。

【0049】 使用者 112-116 可使用其各自的端點 102-106 在會議呼叫 120 上彼此通訊。端點 102-106 中的每一個可包括任何計算設備或電話系統，並且可支援如先前關於圖 1A 所描述的多個呼叫特徵。端點 102-106 可經由網路 110 連接到會議呼叫 120。在一些實例中，會議呼叫 120 可由會議伺服 122 託管和管理。

【0050】 會議伺服 122 可管理會議呼叫 120 的各個態樣，諸如內容、通訊、資料、狀態、設置、功能性、路由、橋接等。在一些實例中，會議伺服 122 可託管併發會議、持

久會議及會議的任何其他類型或組合。會議伺服器122可在任何時間或時間段（例如，一小時、一天、一周、一個月等）託管一或多個會議，其中相應的參與者經由其端點從不同的地理位置、網路等連接。在一些情況下，會議呼叫120中的參與者的數目及/或類型可變化並且可動態地改變。例如，會議呼叫120中的參與者的數目可在會議呼叫120期間改變，因為參與者可動態地加入或離開會議。

【0051】 會議伺服器122可包括一或多個伺服器、橋接器、伺服器應用、雲伺服器、路由器、會議橋接器、閘道、多點控制單元、會議應用等。此外，會議伺服器122的基礎設施可在不同部署中變化。例如，會議伺服器122可經由組織或企業的內部（on-premises）會議基礎設施，作為託管在一或多個雲計算環境或資料中心上的基於雲的伺服器被部署在包括內部會議基礎設施和基於雲的伺服器等的混合基礎設施中。在一些情況下，會議伺服器122可是基於雲的會議伺服器或基礎設施。

【0052】 會議伺服器122可支援不同的會議使用者客戶端/端點和技術。例如，會議伺服器122可支援SIP（對話啟動協定）端點、H.323視訊會議端點、視訊參與者、僅音訊參與者、VoIP（網際網路協定語音）端點、PSTN端點等。在一些實例中，端點102-106可執行使使用者112-116能夠參與由會議伺服器122託管的會議呼叫120的特定軟體。例如，使用者112-116可使用在端點102-106處執行的會議應用來加入和參與由會議伺服器122託管的會

議呼叫 120。會議應用可充當端點 102-106 處的會議使用者客戶端。會議應用可是專門為會議伺服器 122 所託管的會議而配置的本端使用者客戶端應用、具有用於 web 會議的特定功能的 web 瀏覽器或瀏覽器外掛程式/元件/應用、支援一或多個會議技術或協定的使用者客戶端應用，或者適合於會議的任何其他軟體應用。

【0053】 在一些實例中，端點 102-106 及/或端點 102-106 上的會議應用可包括用於會議的各種工具和能力（包括軟體及/或硬體），諸如網路能力、視訊能力、音訊能力、壓縮能力、NAT/防火牆穿越能力、一或多個解碼器等。可由端點 102-106 及/或端點 102-106 上的會議應用用來參與、建立或建立、管理或以其他方式支援託管在會議伺服器 122 上的會議呼叫 120 的技術的非限制性實例包括 SIP、H.263、H.264、H.264 高級簡檔、H.264 SVC（SSL VPN 使用者客戶端）、H.239、H.320、H.323、SIP、VoIP、G.711、G.722、G.729、T.120、VP8、RTP、TCP/IP、HD 視訊會議、遠端訊框緩衝協定、即時通訊協定等。

【0054】 圖 2 圖示了可實現本文所描述的視訊呼叫態樣的端點 102 的實例。端點 102 可包括能夠建立視訊呼叫的任何計算設備或電話系統。例如，端點 102 可包括智慧型電話、智慧可穿戴設備（例如，智慧手錶、智慧護目鏡等）、智慧電視、遊戲系統，平板電腦、IP 電話、膝上型電腦、車載電話、臺式電話、IoT 設備、XR 設備或能夠參與視訊呼

叫的任何其他通訊設備。端點102可支援視訊和語音/音訊，如本文進一步描述的。

【0055】 在圖2所示的實例中，端點102可包括一或多個通訊部件202、一或多個計算部件204、一或多個輸入裝置206、一或多個輸出設備208、相機感測器210、記憶體212、影像處理系統214、訊框調節器216、訊框分離器218、視訊解碼器（例如，視訊轉碼器及/或視訊解碼器）220和渲染系統222。儘管端點102被示為包括某些組件，但是本領域的一般技藝人士將理解，端點102可包括比圖2所示的元件更多或更少的（及/或不同的）組件。例如，在一些情況下，端點102可包括一或多個記憶體設備（例如，RAM、ROM、快取記憶體等）、一或多個相機及/或圖2中未圖示的任何其他硬體或處理設備。以下參考圖8描述可由端點102實現的計算設備和硬體元件的說明性實例。

【0056】 在一些實例中，一或多個通訊部件202可包括用於在蜂巢網路上通訊的蜂巢天線、用於在無線區域網路（WLAN）上通訊的WIFI天線、用於將端點102連接到資料網路（例如，有線及/或無線網路）的網路介面，及/或用於將端點102連接到電話陸上線路的介面元件。在一些情況下，一或多個通訊部件202亦可包括其他天線或通訊介面，諸如例如藍芽天線、GPS天線及/或用於發出及/或接收有線及/或無線信號的任何其他合適的硬體元件。端點102可使用一或多個通訊部件202來建立和參與視訊呼叫

(例如, 100、120)。端點102亦可使用一或多個通訊部件202來發出和接收來自其他設備的資料。

【0057】 一或多個計算部件204可包括中央處理單元(CPU)、圖形處理單元(GPU)、數位訊號處理器(DSP)、影像信號處理器(ISP)、特殊應用積體電路(ASIC)、控制器設備及/或任何其他處理設備。一或多個計算部件204可執行各種操作, 並且可管理/控制端點102的其他元件, 包括一或多個通訊部件202、一或多個輸入裝置206、一或多個輸出設備208、相機感測器210、儲存裝置212等。計算部件204可實施影像處理系統214、訊框調節器216、訊框分離器218、視訊轉碼器220及/或渲染系統222。在一些實例中, 計算部件204亦可實現一或多個其他處理引擎。

【0058】 一或多個輸入裝置206可包括小鍵盤、觸控式螢幕、麥克風、影像感測器、控制器、鍵盤、定點設備及/或能夠以任何輸入形式(例如, 機械運動、音訊、視覺等)接收使用者輸入的任何其他輸入裝置。使用者可使用一或多個輸入裝置206來管理呼叫(例如, 發起呼叫、終止呼叫、保持呼叫、啟用端點102的一或多個元件、啟動一或多個特徵/功能等), 與端點102互動, 提供輸入, 及啟動及/或管理一或多個特徵及/或控制, 諸如訊框/物件選擇特徵、訊框佈置/配置特徵、追蹤特徵、物件偵測特徵、保持特徵、靜音功能、記錄功能、音量控制、端點設置、呼叫設置等。

【0059】 一或多個輸出設備 208 可包括顯示器。在一些實例中，一或多個輸出設備 208 亦可包括（多個）揚聲器、投影儀及 / 或能夠輸出資料的任何元件。例如，在一些情況下，一或多個輸出設備 208 可包括能夠顯示內容和接收使用者輸入的觸控式螢幕。

【0060】 端點 102 可實現相機感測器 210 以擷取影像 / 訊框。例如，端點 102 可實現相機感測器 210 以擷取視訊呼叫的視訊訊框。在一些實例中，相機感測器 210 可包括一或多個前置相機感測器和一或多個後置 / 後方相機感測器。在其他實例中，相機感測器 210 可包括任何其他多相機佈置 / 配置。在一些情況下，端點 102 可包括一或多個附加感測器，諸如例如但不限於慣性量測單元（IMU）、運動偵測感測器、光感測器、音訊感測器、加速度計、陀螺儀、磁力計、高度計、傾斜感測器、光偵測和測距（LIDAR）感測器、無線電偵測和測距（RADAR）感測器、接近度感測器及 / 或任何其他感測器。在一些情況下，相機感測器 210 及 / 或附加感測器可用於偵測物件、偵測物件的位置、偵測一或多個條件（例如，光、運動等）、擷取資料（例如，影像資料、音訊等）、量測物件或環境的一或多個特性（例如，取向、形狀、大小、狀態等）、收集特定類型的量測等。

【0061】 儲存裝置 212 可包括用於儲存諸如影像資料、檔、軟體、視訊、文字資料、訊息、音訊資料、設備 / 元件簡檔、使用者簡檔、設置、使用者輸入、網路資料、日誌

等資料的任何（多個）儲存裝置。此外，儲存裝置 212 可儲存來自端點 102 的任何元件的資料。例如，儲存裝置 212 可儲存來自一或多個通訊部件 202、一或多個計算部件 204、一或多個輸入裝置 206、一或多個輸出設備 208、相機感測器 210、影像處理系統 214、訊框調節器 216、訊框分離器 218、視訊轉碼器 220 和渲染系統 222 的資料。

【0062】 影像處理系統 214 可執行任何影像處理任務。在一些實例中，影像處理系統 214 可處理由相機感測器 210 擷取的訊框。在一些情況下，影像處理系統 214 可執行物件偵測以偵測圖示在（多個）訊框中的一或多個目標（例如，面部、使用者、動物、背景、結構等）、物件追蹤以追蹤圖示在（多個）訊框中的一或多個目標、影像風格化、縮小/放大、色度鍵控、影像分割、濾波、去馬賽克、去噪、內插、模糊、剪切、色彩校正、影像增強等。

【0063】 訊框調節器 216 可經由消除由相機感測器 210 擷取的訊框及/或由相機感測器 210 擷取的訊框的部分來調節視訊呼叫。例如，訊框調節器 216 可移除由相機感測器 210 擷取的一或多個訊框中圖示的一或多個面部。在一些實例中，端點 102 可向使用者呈現在由相機感測器 210 擷取的訊框中圖示的物件（例如，使用者、面部等）的預覽。使用者可從預覽中選擇使用者想要從被發送到參與與端點 102 的視訊呼叫的另一端點的訊框中排除（及/或包括）的任何物件。隨後，訊框調節器 216 可從由相機感測器 210 擷取的

訊框中移除使用者不想包括在發出到另一端點的訊框中的任何物件。

【0064】 在一些情況下，訊框調節器 216 可使用中繼資料來決定由相機感測器 210 擷取的一或多個訊框內的一或多個物件的位置，並且使用一或多個物件的位置來從被發出到另一端點的訊框中移除使用者想要排除的任何物件。在一些實例中，中繼資料可包括由相機感測器 210 擷取的（多個）訊框中圖示的任何面部的座標。

【0065】 訊框分離器 218 可將由相機感測器 210 擷取的多個訊框重構為單個訊框。例如，訊框分離器 218 可將來自相機感測器 210 的訊框合併/組合為單個訊框以用於傳輸到另一端點（例如，在視訊呼叫期間）。在一些實例中，訊框分離器 218 可將訊框的部分/區域佈置成單個訊框，並且管理單個訊框內的部分/區域的佈置。例如，訊框分離器 218 可將擷取面部的訊框的區域與擷取另一訊框的另一訊框的另一區域合併/組合為圖示兩個面部的單個訊框。訊框分離器 218 可根據任何佈置將面部佈置在單個訊框內。為了說明，訊框分離器 218 可佈置該面部以看起來在單個訊框（例如，合併的/組合的訊框）內並排、在單個訊框內的頂部和底部、畫中畫，或任何其他佈置/配置。

【0066】 視訊解碼器 220 可對由相機感測器 210 擷取的訊框進行視訊解碼操作。例如，視訊解碼器 220 可進行視訊編碼、視訊解碼或者視訊編碼和視訊編碼兩者。渲染系統 222 可在端點 102 的顯示裝置上渲染資料。例如，渲染系統

2.2.2 可渲染由相機感測器 210 擷取的訊框、從另一個端點接收的訊框、由相機感測器 210 擷取的訊框的預覽、使用者介面，及/或如本文進一步描述的任何其他資料。

【0067】 在一些實例中，端點 102 可打開（及/或保持打開）相機感測器 210（例如，可打開前置相機感測器和後置相機感測器）。在一些情況下，相機感測器 210 可在視訊呼叫期間由多個使用者使用/共用，以允許相機感測器 210 在來自同一設備（例如，端點 102）的視訊呼叫上擷取多個使用者。在一些實例中，端點 102 可追蹤任何使用者並且進行任何調整，以確保即使一或多個使用者在視訊呼叫期間移動，亦可經由由相機感測器 210 擷取的視訊訊框來圖示使用者。

【0068】 在一些實例中，端點 102 可使用較低功率面部追蹤模式（例如，相對於在端點 102 處可用的使用較高功率資源及/或實現較高功率/複雜度軟體元件的較高功率面部追蹤模式）來在視訊呼叫期間追蹤使用者。為了說明，若第一使用者正在使用端點 102 與另一個端點上的第二使用者進行視訊呼叫，並且在視訊呼叫期間走進第一使用者的房間/空間的第三使用者想要參與視訊呼叫，則端點 102 可擷取第一使用者和第三使用者的訊框（例如，經由相機感測器 210），將圖示第一使用者和第三使用者的訊框合併/組合為單個訊框，及向第二使用者的端點發出該單個訊框。隨後，第二使用者可在由端點 102 發送的單個訊框內看到第一使用者和第三使用者。例如，端點 102 可取得第

一使用者和第三使用者的訊框，並且將該訊框佈置成一個訊框在另一個訊框的上方或旁邊（或任何其他佈置），並且將所得的訊框發送到第二使用者的端點（例如，作為視訊呼叫的一部分）。在一些實例中，端點 102 可將第一使用者和第三使用者的訊框拼接成單個訊框，以容納參與來自同一設備的視訊呼叫的兩個使用者。

【0069】 在一些情況下，若在相機感測器 210 之一（例如，後置/後方相機感測器）的 FOV 記憶體在多個使用者，則端點 102 可將訊框與該相機感測器分開，並且將其作為與擷取另一相機感測器 210（例如，前置相機感測器）的 FOV 內的使用者的訊框拼接在一起的分開訊框來發出。例如，若第三使用者和第四使用者在端點 102 的後置相機感測器的 FOV 內移動，而第一使用者在端點 102 的前置相機感測器的 FOV 內，則第一使用者通常需要將視訊從前置相機感測器切換到後置相機感測器，以在視訊呼叫期間被發送的訊框內示出第三使用者和第四使用者。在該示例場景中，端點 102 可將圖示第三使用者和第四使用者的訊框分離為分開的字元，如本文進一步描述的。

【0070】 在一些情況下，使用端點 102 參與視訊呼叫的使用者可調用（例如，觸發、請求、發起等）相機切換視訊呼叫使用者介面（UI）。端點 102 的相機感測器 210 可被打開，並且可將擷取的訊框發出到計算部件 204（例如，發出到諸如數位訊號處理器的處理器），計算部件 204 隨後被發出到顯示器以供使用者預覽訊框。在一些實例中，

端點 102 可以分離的方式（例如，由線、字元、物件、形狀等分離）發出來自相機感測器 210 的饋送。端點 102 可實現訊框分離器 218 以分離所擷取的面部並且在預覽 UI 上示出上述各者。

【0071】 在一些實例中，發送者（例如，在視訊呼叫上向（多個）其他使用者發出視訊饋送的使用者）可選擇發送者想要發出 UI 的特定側或區域上的預覽的面部，並且接收請求發送者確認輸入面部的許可訊息。隨後，所選擇的面部可被發出到視訊解碼器 220，以經由視訊呼叫流水線被攜帶到接收者。接收者可經由端點 102 接收圖示發送者及參與視訊呼叫的附加使用者的視訊呼叫訊框。

【0072】 在一些實例中，可使用訊框調節器 216 來調節相機感測器饋送，訊框調節器 216 可檢查物件的訊框並且經由僅在物件存在時發送相機感測器饋送來減少視訊解碼器工作量。在一些情況下，相機感測器饋送可包括時間戳記、緩衝器資訊及/或中繼資料。在一些實例中，中繼資料可包括目標（例如，面部、使用者等）的座標及/或目標的座標中繼資料標籤。座標可被用於分離視訊呼叫中的目標（例如，使用者、面部等）。在使用視訊饋送之前，端點 102 可調用使用者輸入流水線以決定使用者想要從相機感測器 210 的 FOV 內的目標發出什麼目標。該資訊可被中繼到訊框調節器 216 和訊框分離器 218。訊框調節器 216 可消除端點 102 的使用者不希望包括在視訊呼叫中的目標（例如，使用者、面部等），諸如未被使用者選擇來包括或者被使

用者選擇來排除的目標。訊框分離器 218 可根據包括的目標的數目將訊框重新構造成單個訊框。

【0073】 在一些情況下，端點 102 可實現物件追蹤模式。例如，端點 102 可僅在使用者出現在任何相機感測器 210 的 FOV 中時發出訊框。這可包括物件追蹤。在端點 102 偵測到面部特徵時，當使用者在另一相機感測器（例如，後置/後方相機感測器）的 FOV 內移動時，端點 102 可停止從一個相機感測器（例如，前置相機感測器）發送訊框。因此，當（多個）使用者在一個相機感測器的 FOV 內時，端點 102 可發出從該相機感測器擷取的訊框，並且當（多個）使用者在不同相機感測器的 FOV 內移動時，端點 102 可發出由不同相機感測器擷取的訊框。

【0074】 圖 3 是圖示根據本案的一些實例的用於提供示例視訊呼叫體驗的示例流水線 300 的圖。在該實例中，流水線 300 包括預覽流水線 330、訊框管理流水線 340 和視訊呼叫流水線。一般來說，預覽流水線 330 可向使用者呈現針對視訊呼叫擷取的訊框和預覽訊框中所圖示的目標（例如，使用者、面部等）的預覽，訊框管理流水線 340 可移除預覽訊框中使用者不希望包括在視訊呼叫訊框中的任何目標，及視訊呼叫流水線 350 可管理視訊呼叫通信期和資料傳送、編碼和發出視訊流、解碼和播放所接收的視訊流等。

【0075】 在預覽流水線 330 中，端點 102 的相機感測器 210 可產生第一相機訊框 302 和第二相機訊框 304。在一些實例

中，第一相機訊框 302 可是來自第一相機感測器的相機饋送的一部分，並且第二相機訊框 304 可是來自第二相機感測器的不同相機饋送的一部分。例如，第一相機訊框 302 可是來自位於端點 102 前面的前置相機感測器的相機饋送的一部分，而第二相機訊框 304 可是來自位於端點 102 後面的後置/後方相機感測器的相機饋送的一部分。為了簡化和解釋的目的，圖 3 僅圖示來自兩個相機感測器的兩個相機訊框或饋送。然而，在一些情況下，相機感測器 210 可產生比圖 3 所示更多或更少的相機訊框/饋送。

【0076】 相機感測器 210 可向處理器 306 提供第一相機訊框 302 和第二相機訊框 304，處理器 306 可產生/渲染包括第一相機訊框 302 和第二相機訊框 304 的預覽(例如，圖示)的使用者介面 308。處理器 306 可向顯示器提供具有用於呈現給端點 102 的使用者的預覽的使用者介面 308。處理器 306 可表示一或多個處理器，諸如例如一或多個 DSP、CPU、ISP 等。在一些情況下，處理器 306 可將第一相機訊框 302 和第二相機訊框 304 拼接在一起成為預覽，並且將該預覽(例如，拼接的訊框)發出到顯示器以用於呈現。在一些實例中，處理器 306 可準備/配置第一相機訊框 302 和第二相機訊框 304 以經由諸如線、方塊、邊界等視覺標記來分離/分開地呈現。這可向使用者指示正在顯示的預覽的不同部分對應於不同的訊框(及預覽的哪些部分對應於不同的訊框)。在一些情況下，處理器 306 可與訊框分離

器 3 1 0 通訊以在預覽內分離第一相機訊框 3 0 2 和第二相機訊框 3 0 4，如先前所解釋。

【0077】 在一些實例中，端點 1 0 2 的使用者可選擇預覽中的特定訊框（及 / 或由預覽中的特定訊框圖示的特定目標（例如，使用者、面部等）），使用者想要將該特定訊框包括在及 / 或排除在發出到視訊呼叫上的另一端點的視訊中。例如，若預覽包括圖示第一面部的第一訊框、圖示第二面部的第二訊框、及圖示第三面部的第三訊框，則使用者可提供指示第一訊框中圖示的第一面部和第二訊框中圖示的第二面部應當被包括在發出到視訊呼叫上的另一端點的視訊中、並且應當排除 / 移除第三訊框中圖示的第三面部的輸入選擇（例如，經由使用者介面 3 0 8）。

【0078】 在訊框管理流水線 3 4 0 中，訊框調節器 3 1 2 可向訊框分離器 3 1 0 提供正被發出到另一端點的任何訊框（例如，圖示特定或選定目標的訊框）。訊框調節器 3 1 2 可基於使用者的輸入選擇來包括和排除任何訊框（及 / 或所圖示的目標）。為了說明，在先前實例中，若使用者選擇保持在第一和第二訊框中圖示的第一面部和第二面部、但不包括在第三訊框中圖示的第三面部，則訊框調節器 3 1 2 可移除圖示第三面部的第三訊框，並且將第一訊框和第二訊框提供給訊框分離器 3 1 0。若使用者沒有選擇從預覽中排除任何訊框（及 / 或所圖示的目標，諸如面部），則訊框調節器 3 1 2 可包括預覽中的所有訊框並且將其發送到訊框分離器 3 1 0。在一些情況下，使用者介面 3 0 8 可呈現許可訊息，

該許可訊息請求來自使用者的輸入訊框（及/或所圖示的目標）的確認以包括在被發出到另一端點（例如，由使用者選擇的訊框及/或所圖示的目標）的流中。

【0079】 在一些實例中，訊框調節器 312 可基於與相機訊框相關聯的中繼資料來辨識特定相機訊框中所圖示的目標的位置。中繼資料可包括辨識訊框內目標位置的座標。訊框調節器 312 可根據（多個）使用者選擇使用座標來辨識並且保持或移除訊框中所圖示的特定目標。例如，若使用者選擇保持第一相機訊框 302 中圖示的第一面部和第二相機訊框 304 中圖示的第二面部、但不包括第三相機訊框中圖示的一或多個第三面部，則訊框調節器 312 可使用與第一、第二和第三相機訊框相關聯的中繼資料來決定第一、第二和第三相機訊框內的第一面部、第二面部和一或多個第三面部的座標。中繼資料可分別包括第一相機訊框 302、第二相機訊框 304 和第三相機訊框內的第一面部、第二面部和一或多個第三面部的座標。訊框調節器 312 可使用座標來移除第三相機訊框中的一或多個第三面部，並且保持第一和第二面部被包括在發出到另一端點的流中。訊框調節器 312 可向訊框分離器 310 提供圖示第一面部的第一相機訊框 302 和圖示第二面部的第二相機訊框 304。

【0080】 在其他實例中，訊框調節器 312 可執行物件偵測以決定相機訊框中所圖示的面部的位置。例如，訊框調節器 312 可處理相機訊框以偵測相機訊框中圖示的面部及其在相機訊框內的位置。

【0081】 訊框調節器 3 1 2 可在對第一相機訊框 3 0 2 和第二相機訊框 3 0 4 進行或不進行任何處理及 / 或改變的情況下提供第一相機訊框 3 0 2 和第二相機訊框 3 0 4。例如，在一些情況下，訊框調節器 3 1 2 可裁剪第一相機訊框 3 0 2 以包括第一面部並且排除第一相機訊框 3 0 2 中第一面部周圍的一或多個區域，及裁剪第二相機訊框 3 0 4 以包括第二面部並且排除第二相機訊框 3 0 4 中第二面部周圍的一或多個區域。

【0082】 訊框分離器 3 1 0 可使用來自訊框調節器 3 1 2 的訊框（例如，被設置為包括在流中並且排除由訊框調節器 3 1 2 移除的任何訊框的訊框（若有的話））來產生包括來自訊框調節器 3 1 2 的訊框的單個訊框。例如，若訊框分離器 3 1 0 從訊框調節器 3 1 2 接收到第一相機訊框 3 0 2 和第二相機訊框 3 0 4，則訊框分離器 3 1 0 可將第一相機訊框 3 0 2 和第二相機訊框 3 0 4 重構（例如，合併 / 組合、排列等）為圖示包括在第一相機訊框 3 0 2 和第二相機訊框 3 0 4 中的目標（例如，使用者、面部等）的單個訊框。在一些情況下，訊框分離器 3 1 0 可在由訊框分離器 3 1 0 產生的單個訊框內分離第一相機訊框 3 0 2 和第二相機訊框 3 0 4。例如，訊框分離器 3 1 0 可在單個訊框內配置 / 佈置 / 圖示第一相機訊框 3 0 2 和第二相機訊框 3 0 4，以便經由單個訊框內的視覺標記來表現為分離 / 分離。視覺標記可包括例如但不限於線、方塊、邊界及 / 或任何其他視覺標記。

【0083】 在一些實例中，為了產生單個訊框，訊框分離器 3 1 0 可基於與相機訊框相關聯的中繼資料來辨識特定相機

訊框中所圖示的目標的位置。如前所解釋的，中繼資料可包括辨識訊框內目標的位置的座標。訊框分離器 310 可使用座標來辨識相機訊框中所圖示的特定目標的位置並且將目標佈置在單個訊框內。例如，若訊框分離器 310 從訊框調節器 312 接收圖示第一面部的第一相機訊框（例如，第一相機訊框 302）和圖示第二面部的第二相機訊框（例如，第二相機訊框 304），則訊框分離器 310 可使用被包括在與第一和第二相機訊框相關聯的中繼資料中的面部座標來辨識第一和第二相機訊框中的面部的位置。訊框分離器 310 可使用座標來知道面部在相機訊框內的位置，以便在由訊框分離器 310 產生的單個訊框內分離及 / 或佈置面部。

【0084】 在其他實例中，訊框分離器 310 可進行物件偵測以決定相機訊框中所圖示的目標的位置。例如，訊框調節器 312 可處理相機訊框以偵測相機訊框中圖示的面部及其在相機訊框內的位置。

【0085】 訊框分離器 310 可使用第一相機訊框 302 和第二相機訊框 304 來產生具有或不具有第一處理及 / 或修改第一相機訊框 302 和第二相機訊框 304 的單個訊框。例如，在一些情況下，訊框分離器 310 可裁剪第一相機訊框 302 以包括第一面部並且排除第一相機訊框 302 中第一面部周圍的一或多個區域，及裁剪第二相機訊框 304 以包括第二面部並且排除第二相機訊框 304 中第二面部周圍的一或多個區域。隨後，訊框分離器 310 可使用經裁剪的相機訊框來從第一和第二相機訊框產生繪示目標的單個訊框。

【0086】 訊框分離器 310 可將所產生的單個訊框提供給視訊呼叫流水線 350。在視訊呼叫流水線 350 中，編碼器 314 可對單個訊框進行編碼，並且將經編碼訊框提供到視訊電話流水線 316 以供傳輸到呼叫上的（多個）其他端點。編碼器 314 可由端點 102 託管。例如，編碼器 314 可是端點 102 上的視訊解碼器 220 的一部分。視訊電話流水線 316 可管理來自端點 102 的視訊呼叫資料傳送（例如，經編碼訊框/流），管理來自端點 102 的資料連接（例如，視訊呼叫連接），管理控制或訊號傳遞平面，管理媒體平面，管理視訊呼叫控制，建立視訊呼叫，管理/實施視訊電話協定及/或執行任何其他視訊電話操作。

【0087】 視訊電話流水線 316 可將經編碼訊框發送到視訊呼叫上的（多個）其他端點。視訊呼叫上的（多個）其他端點可使用解碼器 318 來解碼經編碼訊框以用於在（多個）其他端點處呈現。在一些情況下，解碼器 318 可向視訊呼叫應用程式設計介面（API）320 提供經解碼訊框。視訊呼叫 API 320 可向在（多個）其他端點處的視訊呼叫應用提供經解碼訊框以用於渲染。視訊呼叫 API 320 可產生對視訊呼叫應用的調用（例如，請求、回應等）。在一些情況下，視訊呼叫 API 320 可整合一或多個視訊呼叫工具，諸如視訊聊天、螢幕共用、記錄、工作流整合、效果、篩檢程式、訊息傳遞等。

【0088】 在一些情況下，端點 102 可實現物件追蹤模式以追蹤使用者在端點 102 的相機感測器（例如，相機感測器

210) 的視場 (FOV) 內的存在。在物件追蹤模式中，當在相機感測器的FOV內沒有使用者時 (例如，當使用者從相機感測器的FOV內的區域移動到相機感測器的FOV外的區域時)，端點102可停止從相機感測器擷取及/或發出訊框，並且當端點102決定一或多個使用者在相機感測器的FOV內時 (例如，當端點102決定一或多個使用者在相機感測器的FOV內時)，端點102可開始從相機感測器擷取及/或發出訊框。例如，端點102可分析由相機感測器210擷取的相機訊框以決定在相機訊框中是否圖示了任何使用者。端點102可偵測相機訊框內的面部特徵並且決定 (多個) 使用者存在於擷取該相機訊框的相機感測器的FOV內。若在由相機感測器擷取的相機訊框內沒有偵測到使用者，則端點102可停止將來自該相機感測器的訊框包括在視訊呼叫饋送中，直到在該相機感測器的FOV內偵測到 (多個) 使用者。

【0089】 例如，端點102可處理來自相機感測器210的相機訊框，並且基於在相機訊框中偵測到的任何面部特徵來決定任何使用者是否在相機感測器的FOV內。若使用者從端點102的後置/後方相機感測器的FOV內的區域移動，則端點102可停止在視訊呼叫上將相機訊框從後置/後方相機感測器發出到另一端點104。若使用者在端點102的前置相機感測器的FOV內移動，則端點102可使用前置相機感測器來擷取使用者的相機訊框，並且將使用者的相機訊框包括在發出到另一端點104的流中。若使用者移動回到後置/

後方相機感測器的FOV內的區域，則端點102可再次將由後置/後方相機感測器擷取的相機訊框（及圖示使用者）包括在發出到另一端點104的流中。

【0090】 圖4圖示了圖示在接收者的端點104處渲染的合併的相機訊框402的示例視訊呼叫介面400。端點104可從發送者的端點102接收合併的相機訊框402。如前所描述，端點102可產生（例如，經由訊框分離器310和訊框調節器312）合併的相機訊框402。

【0091】 如圖4所示，合併的相機訊框402包括圖示第一使用者的第一相機訊框404和圖示第二使用者的第二相機訊框406。端點104處的使用者因此可在視訊呼叫介面400內看到第一使用者和第二使用者。第一使用者和第二使用者可是端點102的使用者。在一些實例中，端點102可使用端點102的第一相機感測器來擷取圖示第一使用者的第一相機訊框404，並且使用端點102的第二相機感測器來擷取圖示第二使用者的第二相機訊框406。因此，可從同一設備擷取第一使用者和第二使用者的相機訊框，並且第一使用者和第二使用者可從同一設備（例如，端點102）參與與端點104的使用者的視訊呼叫。

【0092】 圖示第一使用者的第一相機訊框404和圖示第二使用者的第二相機訊框406可以任何方式被佈置在合併的相機訊框402內。例如，在圖4中，圖示第一使用者的第一相機訊框404被佈置在圖示第二使用者的第二相機訊框

406 上方。然而，此種佈置僅僅是出於解釋目的而提供的一個說明性實例。其他示例可包括任何其他佈置。

【0093】 在一些實例中，視訊呼叫介面 400 亦可在端點 104 處顯示圖示使用者的相機訊框 408。因此，端點 104 處的使用者可查看合併的相機訊框 402 內的第一使用者和第二使用者，並且亦可查看視訊呼叫介面 400 內的她自己/他自己的相機訊框。端點 104 處的使用者的相機訊框 408 可以任何方式被佈置在視訊呼叫介面 400 內。例如，可在合併的相機訊框 402 旁邊、在合併的相機訊框 402 的區域內，或根據任何其他佈置來顯示相機訊框 408。此外，相機訊框 408 可被顯示為由端點 104 的相機感測器擷取或根據修改的配置來顯示，該修改的配置諸如為修改的大小、修改的背景、修改的形狀、裁剪的配置、將相機訊框 408 標識為接收者（例如，端點 104 處的使用者）的訊框的視覺標記（例如，框、高亮、線、標籤、輪廓等）及/或任何其他配置。

【0094】 圖 5 圖示了圖示在接收者端點 104 處渲染的合併的相機訊框的另一示例視訊呼叫介面 500。端點 104 可從發送者的端點 102 接收合併的相機訊框。如前所描述，端點 102 可產生（例如，經由訊框分離器 310 和訊框調節器 312）合併的相機訊框。

【0095】 在該實例中，合併的相機訊框包括圖示發送者的端點 102 處的第一使用者的第一訊框 502、圖示發送者的端點 102 處的第二使用者的第二訊框 504、及圖示發送者的端

點 102 處的第三使用者的第三訊框 506。第一訊框 502 可由發送者的端點 102 的第一相機感測器擷取。第二訊框 504 可由發送者的端點 102 的第二相機感測器擷取。第三訊框 506 可由發送者的端點 102 的第二相機感測器或發送者的端點的第三相機感測器擷取。例如，在一些情況下，第一訊框 502 可由發送者的端點 102 的前置相機感測器擷取，並且第二訊框 504 和第三訊框 506 兩者可由發送者的端點 102 的同一相機感測器（諸如後置/後方相機感測器）擷取。

【0096】 在一些實例中，發送者的端點 102 的相機感測器可擷取圖示第二使用者和第三使用者兩者的相機訊框。發送者的端點 102 可（例如，經由訊框分離器 310）分離所擷取的相機訊框內的第二使用者和第三使用者，以便表現為圖示第二使用者和第三使用者的分開的訊框。例如，發送者的端點 102 可將圖示的第二使用者和第三使用者佈置在相機訊框中、在合併的相機訊框內，並且在合併的訊框中包括視覺標記（例如，線、方塊、高亮、輪廓、陰影、邊緣/邊界、標籤等），該視覺標記在合併的訊框內在視覺上劃分/分離第二使用者和第三使用者，以便表現為圖示第二使用者和第三使用者的分開的訊框。發送者的端點 102 處的訊框分離器 310 可使用與圖示第二使用者和第三使用者的相機訊框相關聯的中繼資料中所包括的座標來標識第二使用者和第三使用者在相機訊框內的位置。發送者的端點 102 處的訊框分離器 310 可使用所決定的第二使用者和第三使用者的位置來分離合併的相機訊框內的第二使用者和

第三使用者。在其他情況下，訊框分離器 310 可使用物件偵測來決定第二使用者和第三使用者的位置，並且基於所決定的位置來分離第二使用者和第三使用者。

【0097】 訊框分離器 310 可類似地使用第一使用者的位置來將第一使用者與合併的相機訊框內的第二使用者和第三使用者分離，如圖 5 所示。如前所描述，訊框分離器 310 可基於相關聯的中繼資料中的座標或經由物件偵測來決定第一使用者的位置。

【0098】 如圖 5 所示，本文描述的系統和技術可允許多個使用者參與來自同一設備的視訊呼叫，並且在視訊呼叫的另一端（例如，在端點 104 處）是可見的。此外，本文所描述的系統和技術可允許從同一相機感測器記錄多個使用者並且將其顯示在合併的訊框內的視訊呼叫的另一端（例如，在端點 104 處），如同多個使用者是從不同相機感測器記錄的及 / 或如同多個使用者正在參與來自分開的設備的視訊呼叫一樣。

【0099】 在一些實例中，視訊呼叫介面 500 亦可顯示圖示端點 104 處的使用者的相機訊框 508。因此，端點 104 處的使用者可查看合併的相機訊框（例如，訊框 502、訊框 504、訊框 506）內的第一使用者、第二使用者和第三使用者，並且亦可查看視訊呼叫介面 500 內的她自己 / 他自己的相機訊框。端點 104 處的使用者的相機訊框 508 可以任何方式被佈置在視訊呼叫介面 500 內。例如，可顯示相機訊框 508：在圖示第一、第二和第三使用者的合併的相機訊框旁邊；

在合併的相機訊框的區域內；或根據任何其他佈置。此外，相機訊框 508 可被顯示為由端點 104 的相機感測器擷取或根據修改的配置來顯示，該修改的配置諸如為修改的大小、修改的背景、修改的形狀、裁剪的配置、將相機訊框 508 標識為接收者（例如，端點 104 處的使用者）的訊框的視覺標記（例如，框、高亮、線、標籤、輪廓等）及 / 或任何其他配置。

【0100】 圖 6 A 圖示了圖示由端點 102 產生的示例預覽的示例使用者介面 600。在該實例中，預覽包括圖示視訊呼叫的另一端上的使用者（諸如端點 104 處的使用者）的相機訊框 602。端點 102 可從另一端點（例如，端點 104）接收相機訊框 602，並且將相機訊框 602 包括在預覽中。

【0101】 預覽亦可包括圖示端點 102 處的多個使用者的合併的訊框。合併的訊框可包括圖示端點 102 處的一或多個使用者的第一訊框 604 和圖示端點 102 處的一或多個不同使用者的第二訊框 606。在該實例中，第一訊框 604 包括由端點 102 處的第一相機感測器（諸如後置 / 後方相機感測器）擷取的多個使用者的訊框。第二訊框 606 包括由端點 102 處的第二相機感測器（諸如前置相機感測器）擷取的使用者的訊框。端點 102 處的使用者可選擇預覽中的合併的訊框（例如，第一訊框 604 和第二訊框 606）內的要被包括在發送到（多個）其他端點的合併的訊框中或從其排除的任何使用者 / 面部。端點 102 處的訊框調節器 312 可移除被選

擇用於排除的任何使用者/面部，並且訊框分離器 310 可產生包括被選擇用於包括的使用者/面部的合併的訊框。

【0102】 例如，若使用者從預覽中選擇第一訊框 604 內的第一面部以用於排除，則訊框調節器 312 可移除第一訊框 604 內的第一面部，並且將未被選擇用於排除的第二面部和第三面部保持在第一訊框 604 內。隨後，訊框分離器 310 可產生繪示第二面部和第三面部的合併的訊框。端點 102 可將圖示第二面部和第三面部的合併的訊框發出到另一端點以用於向視訊呼叫上的（多個）其他使用者顯示。

【0103】 圖 6B 圖示了圖示從端點 102 接收的合併的訊框的示例使用者介面 650。在該實例中，渲染在端點 104 處的合併的訊框包括圖示被選擇用於包括在合併的訊框中的使用者（例如，經由圖 6A 中所示和先前描述的預覽）的訊框 606，圖示被選擇用於包括在合併的訊框中的另一使用者（例如，經由圖 6A 中所示的預覽）的訊框 652，及被選擇用於包括在合併的訊框中的另一使用者（例如，經由圖 6A 中所示和先前描述的預覽）。

【0104】 然而，如圖 6B 所示，合併的訊框不包括圖 6A 所示的預覽的第一訊框 604 中所圖示的使用者之一。本文中，端點 102 處的使用者選擇該特定使用者用於從合併的訊框中排除（或者不選擇該使用者用於包括）。因此，端點 102 處的訊框調節器 312 從提供給端點 102 處的訊框分離器 310 的訊框中移除該特定使用者。訊框分離器 310 產生合併的訊框而不包括該特定使用者（及/或描述該特定使用

者的訊框)。因此，即使該特定使用者最初被圖示在與由端點102的相機感測器擷取的訊框606、652和654中所圖示的使用者相同的相機訊框中，該特定使用者在渲染在端點104處的使用者介面650中的合併的訊框內不可見。以此方式，端點102處的使用者可選擇哪些使用者應當被圖示或從被發送到端點104的合併的訊框中移除，並且端點102可從被發送到端點104的合併的訊框中移除被選擇用於排除（或不被選擇用於包括在其中）的任何使用者，即使移除的使用者中的任一個最初出現在（例如，由同一相機感測器擷取的）同一相機訊框中及/或即使任何移除的使用者最初出現在從同一端點擷取的分開的訊框中。

【0105】 圖7是圖示用於產生視訊呼叫的視訊訊框的示例程序700的流程圖。在方塊702處，程序700可包括在第一端點（例如，端點102）與第二端點（例如，端點104）之間建立視訊呼叫。

【0106】 在方塊704處，程序700可包括顯示第一相機饋送和第二相機饋送的預覽。在一些實例中，該第一相機饋送可包括由第一端點的第一影像擷取裝置擷取的第一視訊訊框和由第一端點的第二影像擷取裝置擷取的第二視訊訊框。在一些實例中，該第一視訊訊框和該第二視訊訊框可在該預覽內在視覺上被分開。

【0107】 在一些態樣，程序700可包括從該第一影像擷取裝置獲得該第一視訊訊框及從該第二影像擷取裝置獲得該第二視訊訊框。

【0108】 在方塊 706 處，程序 700 可包括接收對圖示在該預覽中的物件集合的選擇。在一些情況下，對該物件集合的該選擇包括選擇該物件集合作為該多個物件中要被包括在該單個訊框中的子集的第一輸入及 / 或選擇該多個物件中要從該單個訊框中被排除的一或多個物件的第二輸入。

【0109】 在方塊 708 處，程序 700 可包括基於該第一相機饋送和該第二相機饋送來產生繪示該物件集合的單個訊框。在一些態樣，產生該單個訊框可包括將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框。

【0110】 在一些實例中，將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框可包括將該第一視訊訊框的至少該部分和該第二視訊訊框的至少該部分佈置成該單個訊框的相應訊框區域。在一些實例中，每個訊框區域可圖示來自該物件集合的相應物件。

【0111】 在一些實例中，該第一視訊訊框和該第二視訊訊框圖示包括該物件集合的多個物件。在一些情況下，對該物件集合的該選擇包括選擇該物件集合要被包括在該單個訊框中的第一輸入及 / 或選擇該多個物件中要從該單個訊框中被排除的一或多個物件的第二輸入。該一或多個物件可不同於該物件集合。

【0112】 在一些情況下，產生該單個訊框可包括基於該第一輸入及 / 或該第二輸入，從該單個訊框中排除該多個物件中的該一或多個物件。在一些實例中，程序 700 可包括向

該第二端點發出該單個訊框。在一些情況下，從該單個訊框中排除該一或多個物件可包括從該預覽、該第一視訊訊框、該第二視訊訊框及/或該單個訊框中移除該一或多個物件。

【0113】 在一些情況下，該第一視訊訊框可包括來自該物件集合的一或多個第一物件，及該第二視訊訊框可包括來自該物件集合的一或多個第二物件。在一些態樣，產生該單個訊框可包括基於與該第一視訊訊框相關聯的第一中繼資料來決定該一或多個第一物件之每一第一物件在該第一視訊訊框內的相應位置；基於與該第二視訊訊框相關聯的第二中繼資料來決定該一或多個第二物件之每一第二物件在該第二視訊訊框內的相應位置；及基於該一或多個第一物件之每一第一物件在該第一視訊訊框內的該相應位置及該一或多個第二物件之每一第二物件在該第二視訊訊框內的該相應位置，決定圖示該一或多個第一物件的該第一視訊訊框的第一部分和圖示該一或多個第二物件的該第二視訊訊框的第二部分。

【0114】 在一些實例中，該第一中繼資料可包括與該一或多個第一物件相關聯的座標，及該第二中繼資料可包括與該一或多個第二物件相關聯的座標。在一些情況下，產生該單個訊框可包括將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框。在一些態樣，將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框可包括將該第一視訊訊框的

該第一部分和該第二視訊訊框的該第二部分放置在該單個訊框的相應訊框區域內。在一些實例中，該相應訊框區域中的第一區域繪示該一或多個第一物件，及該相應訊框區域中的第二區域繪示該一或多個第二物件。

【0115】 在一些態樣，產生該單個訊框可包括在該單個訊框內渲染及/或圖示該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分。在一些情況下，該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分在視覺上被分開。在一些實例中，該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分經由視覺標記在視覺上被分開。在一些實例中，該視覺標記可包括線、輪廓、方塊、高亮、標籤、顏色、陰影及/或不同的視覺記號。

【0116】 在一些態樣，程序700可包括決定來自該多個物件的該一或多個物件在該第一影像擷取裝置的FOV內，及觸發該第一影像擷取裝置以擷取該第一視訊訊框。在一些實例中，該第一視訊訊框圖示該一或多個物件。

【0117】 在一些態樣，程序700可包括基於由該第一影像擷取裝置擷取的一或多個視訊訊框來追蹤該一或多個物件；基於該追蹤，決定該一或多個物件不再位於該第一影像擷取裝置的該FOV內；及基於決定該一或多個物件不再位於該第一影像擷取裝置的該FOV內，決定停止從該第一影像擷取裝置向該第二端點發送視訊資料（例如，視訊訊框）。

【0118】 在一些態樣，程序700可包括回應於決定至少一個物件在該第二影像擷取裝置的FOV內，經由該第二影像擷取裝置獲得圖示該至少一個物件的第三視訊訊框；及向該第二端點發出該第三視訊訊框。

【0119】 圖8圖示了可實現本文所描述之各種技術之示例計算設備之示例計算設備架構800。例如，計算設備架構800可實現圖2所示之端點102之至少一些部分。計算設備架構800之部件被示出為使用諸如匯流排之連接裝置805彼此電通訊。示例計算設備架構800包括處理單元（CPU或處理器）810，及將包括諸如唯讀記憶體（ROM）820和隨機存取記憶體（RAM）825之計算設備記憶體815之各種計算設備部件耦合到處理器810之計算設備連接裝置805。

【0120】 計算設備架構800可包括直接與處理器810連接、緊鄰處理器810或作為處理器810之一部分整合之高速記憶體之快取記憶體。計算設備架構800可將資料從記憶體815及/或儲存裝置830複製到快取記憶體812，以用於由處理器810快速存取。以此方式，快取記憶體可提供避免處理器810在等待資料時延遲之效能提升。該等和其他模組可控制或被配置為控制處理器810執行各種動作。亦可使用其他計算設備記憶體815。記憶體815可包括具有不同效能特性之多個不同類型之記憶體。處理器810可包括任何通用處理器，及儲存在儲存裝置830中並且被配置為控制處理器810之硬體或軟體伺服，及其中軟體指令被併

入到處理器設計中的專用處理器。處理器 810 可是包含多個核或處理器、匯流排、記憶體控制器、快取記憶體等的自包含系統。多核處理器可是對稱的或不對稱的。

【0121】 為了使使用者能夠與計算設備架構 800 互動，輸入裝置 845 可表示任何數目的輸入機制，諸如用於語音的麥克風、用於手勢或圖形輸入的觸敏螢幕、鍵盤、滑鼠、運動輸入、語音等。輸出設備 835 亦可是本領域技藝人士已知的多種輸出機制中的一或多個，諸如顯示器、投影儀、電視、揚聲器設備。在一些實例中，多模式計算設備可使使用者能夠提供多種類型的輸入以與計算設備架構 800 通訊。通訊介面 840 通常可治理和管理使用者輸入和計算設備輸出。對於在任何特定硬體佈置上操作沒有限制，因此，隨著開發改進的硬體或韌體佈置，此處的基本特徵可容易地替換為改進的硬體或韌體佈置。

【0122】 儲存裝置 830 是非揮發性記憶體，並且可是硬碟或可儲存可由電腦存取的資料的其他類型的電腦可讀取媒體，諸如磁帶盒、快閃記憶卡、固態儲存裝置、數位多功能光碟、盒式磁帶、隨機存取記憶體（RAM）825、唯讀記憶體（ROM）820 及其混合。儲存裝置 830 可包括用於控制處理器 810 的軟體、代碼、韌體等。可設想其他硬體或軟體模組。儲存裝置 830 可被連接到計算設備連接裝置 805。在一個態樣，執行特定功能的硬體模組可包括儲存在電腦可讀取媒體中的軟體部件與必要的硬體部件（諸如

處理器 810、連接裝置 805、輸出設備 835 等) 連接以執行功能。

【0123】 術語「電腦可讀取媒體」包括但不限於可攜式或非可攜式儲存裝置、光儲存裝置及能夠儲存、包含或攜帶(多個)指令及/或資料的各種其他媒體。電腦可讀取媒體可包括其中可儲存資料並且不包括無線地傳播或經由有線連接傳播的載波及/或暫時性電子信號的非暫時性媒體。非暫時性媒體的實例可包括但不限於磁碟或磁帶、諸如壓縮光碟(CD)或數位多功能光碟(DVD)的光學儲存媒體、快閃記憶體、記憶體或記憶體設備。電腦可讀取媒體可具有儲存在其上的代碼及/或機器可執行指令,該等代碼/指令可表示程序、函數、副程式、程式、常式、子常式、模組、套裝軟體、類或者指令、資料結構或程式語句的任何組合。程式碼片段可經由傳遞及/或接收資訊、資料、引數、參數或記憶體內容而被耦合到另一程式碼片段或硬體電路。資訊、引數、參數、資料等可經由包括記憶體共享、訊息傳遞、符記傳遞、網路傳輸等任何合適的方式被傳遞、轉發或發送。

【0124】 在一些態樣,電腦可讀儲存裝置、媒體和記憶體可包括包含位元串流等的電纜或無線信號。然而,當提及時,非暫時性電腦可讀取儲存媒體明確地排除了諸如能量、載波信號、電磁波和信號本身的媒體。

【0125】 在以上描述中提供了具體細節以提供對本文提供的態樣和示例的透徹理解。然而,本領域的技藝人士將瞭

解，可在沒有該等特定細節的情況下實踐該態樣。為了解釋的清楚，在一些實例中，本技術可被呈現為包括單獨的功能方塊，該等單獨的功能方塊包括設備、設備部件、體現於軟體中的方法的步驟或常式，或者硬體和軟體的組合。除了在附圖中示出及/或本文描述的彼等之外，可使用附加的部件。例如，電路、系統、網路、程序和其他部件可被示為方塊圖形式的部件，以免在不必要的細節中混淆各態樣。在其他實例中，可在沒有不必要細節的情況下示出眾所周知的電路、程序、演算法、結構和技術，以便避免混淆各態樣。

【0126】 以上可將各態樣描述為被圖示為流程圖、流程圖表、資料串流圖、結構圖或方塊圖的程序或方法。儘管流程圖可將操作描述為順序程序，但許多操作可並行或者同時被執行。此外，可重新安排操作的順序。當程序的操作完成時，該程序終止，但可能包括未包括在圖中的附加步驟。程序可對應於方法、函數、程序、子常式、副程式等。當程序對應於函數時，其終止可對應於函數返回到調用函數或主函數。

【0127】 可使用電腦可執行指令來實現根據上述實例的程序和方法，該電腦可執行指令儲存在電腦可讀取媒體中或可從電腦可讀取媒體獲得。此種指令可包括，例如，使得或以其他方式配置通用電腦、專用電腦或處理設備來執行特定功能或功能組的指令和資料。所使用的部分電腦資源可經由網路存取。電腦可執行指令可是例如二進位元、中

間格式指令，諸如組合語言、韌體、原始程式碼。可用於儲存指令、所使用的資訊及/或在根據所描述的實例的方法期間建立的資訊的電腦可讀取媒體的實例包括磁碟或光碟、快閃記憶體、以非揮發性記憶體提供的USB設備、網路儲存裝置等等。

【0128】 實現根據該等揭示的程序和方法的設備可包括硬體、軟體、韌體、中介軟體、微代碼、硬體描述語言或其任何組合，並且可採用多種形式因素中的任何一種。當以軟體、韌體、中介軟體或微代碼實現時，執行必要任務（例如，電腦程式產品）的程式碼或程式碼片段可被儲存在電腦可讀或機器可讀取媒體中。（多個）處理器可執行必要的任務。形式因素的典型實例包括膝上型電腦、智慧型電話、行動電話、平板設備或其他小形式因素個人電腦、個人數位助理、機架式設備、獨立設備等。本文描述的功能性亦可在周邊設備或插入卡中實現。作為進一步的實例，此種功能性亦可在電路板上在單個設備中執行的不同晶片或不同程序之中實現。

【0129】 指令、用於傳遞此種指令的媒體、用於執行其計算資源及用於支援此種計算資源的其他結構是用於提供本方案中描述的功能的示例部件。

【0130】 在前文的描述中，參考本案的具體態樣描述了其各態樣，但是本領域技藝人士將認識到本案不限於此。因此，儘管本文已經詳細描述了本案的說明性態樣，但是應當理解，本發明的概念可以其他方式以各種方式實現和採

用，並且所附請求項意欲被解釋為包括除了由現有技術限制之外的該等變化。上述應用的各種特徵和態樣可單獨或聯合使用。此外，在不脫離本說明書的更寬的精神和範圍的情況下，可在本文描述的彼等之外的任何數目的環境和應用中利用各態樣。因此，說明書和附圖被認為是說明性的而不是限制性的。出於說明的目的，方法被以特定的次序描述。應當瞭解，在備選態樣，該方法可與所描述的次序不同的次序執行。

【0131】 本領域一般技藝人士將理解，在不脫離本說明書的範圍的情況下，本文使用的小於（「 $<$ 」）和大於（「 $>$ 」）符號或術語可分別用小於或等於（「 \leq 」）和大於或等於（「 \geq 」）符號代替。

【0132】 在部件被描述為「被配置為」執行某些操作的情況下，此種配置可例如經由設計電子電路或其他硬體來執行操作、經由對可程式設計電子電路（例如，微處理器或其他合適的電子電路）進行程式設計以執行操作或其任何組合來實現。

【0133】 短語「耦合到」是指任一部件直接地或間接地實體連接到另一部件，及/或任一部件與另一部件直接地或間接地通訊（例如，經由有線或無線連接及/或其他合適的通訊介面連接到另一部件）。

【0134】 在本文中敘述集合「中的至少一個」及/或集合「中的一或多個」的請求項語言或其他語言表示該集合中的一個成員或該集合中的多個成員（以任何組合）滿足該請求

項。例如，敘述「A和B中的至少一個」或「A或B中的至少一個」的請求項語言意指A、B或A和B。在另一實例中，敘述「A、B和C中的至少一個」或「A、B或C中的至少一個」的請求項語言意指A、B、C，或A和B，或A和C，或B和C，或A和B和C。語言集合「中的至少一個」及/或集合中的「一或多個」並不將該集合限制為該集合中所列的專案。例如，敘述「A和B中的至少一個」或「A或B中的至少一個」的請求項語言可意指A、B，或A和B，並且可另外包括未在A和B的集合中列出的專案。

【0135】 結合本文所揭示的示例而描述的各種說明性邏輯區塊、模組、電路和演算法步驟可被實施為電子硬體、電腦軟體、韌體或其組合。為了清楚地說明硬體和軟體的此種可互換性，上文已就其功能性大體描述了各種說明性元件、方塊、模組、電路和步驟。將此種功能性實施為硬體亦是軟體取決於施加在整個系統上的特定的應用和設計約束。技藝人士可針對每個特定應用以不同的方式實施所描述的功能，但是此種實施方式決定不應被解釋為導致偏離本案的範圍。

【0136】 本文所描述的技術亦可以電子硬體、電腦軟體、韌體或其任何組合來實施。此種技術可在多種設備中的任一者中實現，諸如通用電腦、無線通訊設備手持機，或者具有包括無線通訊設備手持機和其他設備中的應用在內的多個用途的積體電路設備。描述為模組或部件的任何特徵可在整合邏輯裝置中一起實施，或者單獨地作為離散但可

交互操作的邏輯裝置來實施。若以軟體實施，則該等技術可至少部分地經由包括程式碼的電腦可讀取資料儲存媒體實施，該程式碼包括指令，當該指令被執行時，進行上文描述的方法、演算法及/或操作中的一或多個。電腦可讀取資料儲存媒體可形成電腦程式產品的一部分，其可包括封裝材料。電腦可讀取媒體可包括記憶體或資料儲存媒體，諸如隨機存取記憶體（RAM）（諸如同步動態隨機存取記憶體（SDRAM））、唯讀記憶體（ROM）、非揮發性隨機存取記憶體（NVRAM）、電子可抹除可程式設計唯讀記憶體（EEPROM）、快閃記憶體、磁或光資料儲存媒體等。附加地或替代地，可至少部分地由電腦可讀通訊媒體來實現該等技術，該電腦可讀通訊媒體以指令或資料結構的形式攜帶或傳送程式碼，並且該程式碼可由電腦（例如傳播的信號或波）存取、讀取及/或執行。

【0137】 程式碼可由處理器執行，該處理器可包括一或多個處理器，諸如一或多個數位訊號處理器（DSP）、通用微處理器、特殊應用積體電路（ASIC）、現場可程式設計邏輯陣列（FPGA）或其他等效的整合或離散邏輯電路系統。此種處理器可被配置為執行本案中所描述的技術中的任一個。通用處理器可是微處理器；但在替代方案中，處理器可是任何一般處理器、控制器、微控制器或狀態機。處理器亦可實現為計算設備的組合，例如，DSP和微處理器的組合、多個微處理器、與DSP核心結合的一或多個微處理器，或者任何其他此類配置。因此，如本文所使用的

術語「處理器」可代表任何前述結構、前述結構的任何組合，或適用於實施本文描述的技術的任何其他結構或裝置。

【0138】 本案的說明性實例包括：

【0139】 態樣 1。一種用於處理視訊呼叫的裝置，包括：記憶體；及一或多個處理器，耦合到該記憶體，該一或多個處理器被配置為：在該裝置與遠端設備之間建立視訊呼叫；顯示第一相機饋送和第二相機饋送的預覽，該第一相機饋送包括由該裝置的第一影像擷取裝置擷取的第一視訊訊框和由該裝置的第二影像擷取裝置擷取的第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；接收對圖示在該預覽中的物件集合的選擇；及基於該第一相機饋送和該第二相機饋送來產生繪示該物件集合的單個訊框。

【0140】 態樣 2。根據態樣 1 的裝置，其中該第一視訊訊框和該第二視訊訊框繪示包括該物件集合的多個物件，及其中對該物件集合的該選擇包括第一輸入和第二輸入中的至少一個，該第一輸入選擇該物件集合作為該多個物件中要被包括在該單個訊框中的子集，該第二輸入選擇該多個物件中要從該單個訊框中被排除的一或多個物件，該一或多個物件不同於該物件集合。

【0141】 態樣 3。根據態樣 2 的裝置，其中為了產生該單個訊框，該一或多個處理器亦被配置為：基於該第一輸入和該第二輸入中的至少一個，從該單個訊框中排除該多個物件中的該一或多個物件；及向該遠端設備發出該單個訊框。

【0142】 態樣 4。根據態樣 3 的裝置，其中為了從該單個訊框中排除該一或多個物件，該一或多個處理器被配置為從該預覽、該第一視訊訊框、該第二視訊訊框和該單個訊框中的至少一個中移除該一或多個物件。

【0143】 態樣 5。根據態樣 1 至 4 中任一項的裝置，其中為了產生該單個訊框，該一或多個處理器亦被配置為：將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框。

【0144】 態樣 6。根據態樣 5 的裝置，其中為了將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框，該一或多個處理器被配置為：將該第一視訊訊框的至少該部分和該第二視訊訊框的至少該部分佈置成該單個訊框的相應訊框區域，每個訊框區域圖示來自該物件集合的相應物件。

【0145】 態樣 7。根據態樣 1 至 6 中任一項的裝置，其中該第一視訊訊框包括來自該物件集合的一或多個第一物件，及該第二視訊訊框包括來自該物件集合的一或多個第二物件，及其中為了產生該單個訊框，該一或多個處理器亦被配置為：基於與該第一視訊訊框相關聯的第一中繼資料來決定該一或多個第一物件之每一第一物件在該第一視訊訊框內的相應位置，該第一中繼資料包括與該一或多個第一物件相關聯的座標；基於與該第二視訊訊框相關聯的第二中繼資料來決定該一或多個第二物件之每一第二物件在該第二視訊訊框內的相應位置，該第二中繼資料包括與該一

或多個第二物件相關聯的座標；及基於該一或多個第一物件之每一第一物件在該第一視訊訊框內的該相應位置及該一或多個第二物件之每一第二物件在該第二視訊訊框內的該相應位置，決定圖示該一或多個第一物件的該第一視訊訊框的第一部分和圖示該一或多個第二物件的該第二視訊訊框的第二部分。

【0146】 態樣 8。根據態樣 7 的裝置，其中為了產生該單個訊框，該一或多個處理器亦被配置為：將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框。

【0147】 態樣 9。根據態樣 8 的裝置，其中為了將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框，該一或多個處理器被配置為：將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分放置在該單個訊框的相應訊框區域內，其中該相應訊框區域中的第一區域繪示該一或多個第一物件，及該相應訊框區域中的第二區域繪示該一或多個第二物件。

【0148】 態樣 10。根據態樣 1 至 9 中任一項的裝置，其中為了產生該單個訊框，該一或多個處理器亦被配置為：在該單個訊框內渲染該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分，其中該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分在視覺上被分開。

【0149】 態樣 11。根據態樣 10 的裝置，其中該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分經由

視覺標記在視覺上被分開，該視覺標記包括線、輪廓、方塊、高亮、標籤、顏色、陰影和視覺記號中的至少一個。

【0150】 態樣 12。根據態樣 1 至 11 中任一項的裝置，其中該一或多個處理器被配置為：從該第一影像擷取裝置獲得該第一視訊訊框，及從該第二影像擷取裝置獲得該第二視訊訊框。

【0151】 態樣 13。根據態樣 12 的裝置，亦包括該第一影像擷取裝置和該第二影像擷取裝置中的至少一個。

【0152】 態樣 14。根據態樣 1 至 13 中任一項的裝置，其中該一或多個處理器被配置為：決定該一或多個物件在該第一影像擷取裝置的視場（FOV）內；及觸發該第一影像擷取裝置以擷取該第一視訊訊框，其中該第一視訊訊框圖示該一或多個物件。

【0153】 態樣 15。根據態樣 14 的裝置，其中該一或多個處理器被配置為：基於由該第一影像擷取裝置擷取的一或多個視訊訊框來追蹤該一或多個物件；基於該追蹤，決定該一或多個物件不再位於該第一影像擷取裝置的該 FOV 內；及基於決定該一或多個物件不再位於該第一影像擷取裝置的該 FOV 內，決定停止從該第一影像擷取裝置向該遠端設備發送視訊資料。

【0154】 態樣 16。根據態樣 15 的裝置，其中該一或多個處理器被配置為：回應於決定至少一個物件在該第二影像擷取裝置的 FOV 內，經由該第二影像擷取裝置獲得圖示該至

少一個物件的第三視訊訊框；及向該遠端設備發出該第三視訊訊框。

【0155】 態樣 17。根據態樣 1 至 16 中任一項的裝置，其中該裝置包括行動設備。

【0156】 態樣 18。一種用於處理視訊呼叫的方法，該方法包括：在設備與遠端設備之間建立視訊呼叫；顯示第一相機饋送和第二相機饋送的預覽，該第一相機饋送包括由該設備的第一影像擷取裝置擷取的第一視訊訊框和由該設備的第二影像擷取裝置擷取的第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；接收對圖示在該預覽中的物件集合的選擇；及基於該第一相機饋送和該第二相機饋送來產生繪示該物件集合的單個訊框。

【0157】 態樣 19。根據態樣 18 的方法，其中該第一視訊訊框和該第二視訊訊框圖示包括該物件集合的多個物件，並且其中對該物件集合的該選擇包括第一輸入和第二輸入中的至少一個，該第一輸入選擇該物件集合作為該多個物件中要被包括在該單個訊框中的子集，該第二輸入選擇該多個物件中要從該單個訊框中被排除的一或多個物件，該一或多個物件不同於該物件集合。

【0158】 態樣 20。根據態樣 19 的方法，其中產生該單個訊框亦包括：基於該第一輸入和該第二輸入中的至少一個，從該單個訊框中排除該多個物件中的該一或多個物件；及向該遠端設備發出該單個訊框。

【0159】 態樣 2 1。根據態樣 2 0 的方法，其中從該單個訊框中排除該一或多個物件包括從該預覽、該第一視訊訊框、該第二視訊訊框和該單個訊框中的至少一個中移除該一或多個物件。

【0160】 態樣 2 2。根據態樣 1 8 至 2 1 中任一項的方法，其中產生該單個訊框亦包括：將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框。

【0161】 態樣 2 3。根據態樣 2 2 的方法，其中將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框包括：將該第一視訊訊框的至少該部分和該第二視訊訊框的至少該部分佈置成該單個訊框的相應訊框區域，每個訊框區域圖示來自該物件集合的相應物件。

【0162】 態樣 2 4。根據態樣 1 8 至 2 3 中任一項的方法，其中該第一視訊訊框包括來自該物件集合的一或多個第一物件，及該第二視訊訊框包括來自該物件集合的一或多個第二物件，並且其中產生該單個訊框亦包括：基於與該第一視訊訊框相關聯的第一中繼資料來決定該一或多個第一物件之每一第一物件在該第一視訊訊框內的相應位置，該第一中繼資料包括與該一或多個第一物件相關聯的座標；基於與該第二視訊訊框相關聯的第二中繼資料來決定該一或多個第二物件之每一第二物件在該第二視訊訊框內的相應位置，該第二中繼資料以包括與該一或多個第二物件相關聯的座標；及基於該一或多個第一物件之每一第一物件在該第一視訊訊框內的該相應位置及該一或多個第二物件之

每一第二物件在該第二視訊訊框內的該相應位置，決定圖示該一或多個第一物件的該第一視訊訊框的第一部分和圖示該一或多個第二物件的該第二視訊訊框的第二部分。

【0163】 態樣 2 5。根據態樣 2 4 的方法，其中產生該單個訊框亦包括：將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框。

【0164】 態樣 2 6。根據態樣 2 5 的方法，其中將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框包括：將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分放置在該單個訊框的相應訊框區域內，其中該相應訊框區域中的第一區域繪示該一或多個第一物件，及該相應訊框區域中的第二區域繪示該一或多個第二物件。

【0165】 態樣 2 7。根據態樣 1 8 至 2 6 中任一項的方法，其中產生該單個訊框亦包括：在該單個訊框內渲染該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分，其中該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分在視覺上被分開。

【0166】 態樣 2 8。根據態樣 2 7 的方法，其中該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分經由視覺標記在視覺上被分開，該視覺標記包括線、輪廓、方塊、高亮、標籤、顏色、陰影和視覺記號中的至少一個。

【0167】 態樣 29。根據態樣 18 至 28 中任一項的方法，亦包括：從該第一影像擷取裝置獲得該第一視訊訊框，及從該第二影像擷取裝置獲得該第二視訊訊框。

【0168】 態樣 30。根據態樣 18 至 29 中任一項的方法，亦包括：決定該一或多個物件在該第一影像擷取裝置的視場（FOV）內；及觸發該第一影像擷取裝置以擷取該第一視訊訊框，其中該第一視訊訊框圖示該一或多個物件。

【0169】 態樣 31。根據態樣 30 的方法，亦包括：基於由該第一影像擷取裝置擷取的一或多個視訊訊框來追蹤該一或多個物件；基於該追蹤，決定該一或多個物件不再位於該第一影像擷取裝置的該 FOV 內；及基於決定該一或多個物件不再位於該第一影像擷取裝置的該 FOV 內，決定停止從該第一影像擷取裝置向該遠端設備發送視訊資料。

【0170】 態樣 32。根據態樣 31 的方法，亦包括：回應於決定至少一個物件在該第二影像擷取裝置的 FOV 內，經由該第二影像擷取裝置獲得圖示該至少一個物件的第三視訊訊框；及向該遠端設備發出該第三視訊訊框。

【0171】 態樣 33。包含指令的至少一種非暫時性電腦可讀取媒體，該指令在由一或多個處理器執行時，使該一或多個處理器執行態樣 18 至 32 中任一項的方法。

【0172】 態樣 34。一種裝置，包括用於執行態樣 18 至 32 中任一項的方法的手段。

【0173】 態樣 34。根據態樣 34 的裝置，其中該裝置包括行動設備。

【符號說明】**【0174】**

1 0 0 : 視 訊 通 話
1 0 2 : 端 點
1 0 4 : 端 點
1 0 6 : 端 點
1 1 0 : 網 路
1 1 2 : 使 用 者
1 1 4 : 使 用 者
1 1 6 : 使 用 者
1 2 0 : 會 議 呼 叫
1 2 2 : 會 議 伺 服
2 0 2 : 通 訊 部 件
2 0 4 : 計 算 部 件
2 0 6 : 輸 入 裝 置
2 0 8 : 輸 出 設 備
2 1 0 : 相 機 感 測 器
2 1 2 : 儲 存 裝 置
2 1 4 : 影 像 處 理 系 統
2 1 6 : 訊 框 調 節 器
2 1 8 : 訊 框 分 離 器
2 2 0 : 視 訊 轉 碼 器
2 2 2 : 渲 染 系 統
3 0 0 : 流 水 線

- 3 0 2 : 第一相機訊框
- 3 0 4 : 第二相機訊框
- 3 0 6 : 處理器
- 3 0 8 : 使用者介面
- 3 1 0 : 訊框分離器
- 3 1 2 : 訊框調節器
- 3 1 4 : 編碼器
- 3 1 6 : 視訊電話流水線
- 3 1 8 : 解碼器
- 3 2 0 : 視訊呼叫 A P I
- 3 3 0 : 預覽流水線
- 3 4 0 : 訊框管理流水線
- 3 5 0 : 視訊呼叫流水線
- 4 0 0 : 視訊呼叫介面
- 4 0 2 : 相機訊框
- 4 0 4 : 第一相機訊框
- 4 0 6 : 第二相機訊框
- 4 0 8 : 相機訊框
- 5 0 0 : 視訊呼叫介面
- 5 0 2 : 第一訊框
- 5 0 4 : 第二訊框
- 5 0 6 : 第三訊框
- 5 0 8 : 相機訊框
- 6 0 0 : 使用者介面

- 6 0 2 : 相 機 訊 框
- 6 0 4 : 第 一 訊 框
- 6 0 6 : 第 二 訊 框
- 6 5 0 : 使 用 者 介 面
- 6 5 2 : 訊 框
- 6 5 4 : 訊 框
- 7 0 0 : 程 序
- 7 0 2 : 步 驟
- 7 0 4 : 步 驟
- 7 0 6 : 步 驟
- 7 0 8 : 步 驟
- 8 0 0 : 計 算 設 備 架 構
- 8 0 5 : 計 算 設 備 連 接 裝 置
- 8 1 0 : 處 理 單 元
- 8 1 2 : 快 取 記 憶 體
- 8 1 5 : 計 算 設 備 記 憶 體
- 8 2 0 : 唯 讀 記 憶 體 (R O M)
- 8 2 5 : 隨 機 存 取 記 憶 體 (R A M)
- 8 3 0 : 儲 存 裝 置
- 8 3 5 : 輸 出 設 備
- 8 4 0 : 通 訊 介 面
- 8 4 5 : 輸 入 裝 置

【生物材料寄存】

國內寄存資訊(請依寄存機構、日期、號碼順序註記)

無

國外寄存資訊(請依寄存國家、機構、日期、號碼順序註記)

無

【發明申請專利範圍】

【請求項 1】 一種用於處理視訊呼叫的裝置，包括：

一記憶體；及

一或多個處理器，耦合到該記憶體，該一或多個處理器被配置為：

在該裝置與一遠端設備之間建立一視訊呼叫；

顯示一第一相機饋送和一第二相機饋送的一預覽，該第一相機饋送包括由該裝置的一第一影像擷取裝置擷取的一第一視訊訊框和由該裝置的一第二影像擷取裝置擷取的一第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；

接收對圖示在該預覽中的一物件集合的選擇；及

基於該第一相機饋送和該第二相機饋送來產生繪示該物件集合的一單個訊框。

【請求項 2】 根據請求項 1 之裝置，其中該第一視訊訊框和該第二視訊訊框圖示包括該物件集合的多個物件，及其中對該物件集合的該選擇包括一第一輸入和一第二輸入中的至少一個，該第一輸入選擇該物件集合作為該多個物件中要被包括在該單個訊框中的一子集，該第二輸入選擇該多個物件中要從該單個訊框中被排除的一或多個物件，該一或多個物件不同於該物件集合。

【請求項 3】 根據請求項 2 之裝置，其中為了產生該單個訊框，該一或多個處理器亦被配置為：

基於該第一輸入和該第二輸入中的至少一個，從該單

個訊框中排除該多個物件中的該一或多個物件；及
向該遠端設備發出該單個訊框。

【請求項4】 根據請求項3之裝置，其中為了從該單個訊框中排除該一或多個物件，該一或多個處理器被配置為從該預覽、該第一視訊訊框、該第二視訊訊框和該單個訊框中的至少一個中移除該一或多個物件。

【請求項5】 根據請求項1之裝置，其中為了產生該單個訊框，該一或多個處理器亦被配置為：

將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框。

【請求項6】 根據請求項5之裝置，其中為了將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框，該一或多個處理器被配置為：

將該第一視訊訊框的至少該部分和該第二視訊訊框的至少該部分佈置成該單個訊框的相應訊框區域，每個訊框區域圖示來自該物件集合的一相應物件。

【請求項7】 根據請求項1之裝置，其中該第一視訊訊框包括來自該物件集合的一或多個第一物件，及該第二視訊訊框包括來自該物件集合的一或多個第二物件，並且其中為了產生該單個訊框，該一或多個處理器亦被配置為：

基於與該第一視訊訊框相關聯的第一中繼資料來決定該一或多個第一物件之每一第一物件在該第一視訊訊框內的一相應位置，該第一中繼資料包括與該一或多個第

一物件相關聯的座標；

基於與該第二視訊訊框相關聯的第二中繼資料來決定該一或多個第二物件之每一第二物件在該第二視訊訊框內的一相應位置，該第二中繼資料包括與該一或多個第二物件相關聯的座標；及

基於該一或多個第一物件之每一第一物件在該第一視訊訊框內的該相應位置及該一或多個第二物件之每一第二物件在該第二視訊訊框內的該相應位置，決定圖示該一或多個第一物件的該第一視訊訊框的一第一部分和圖示該一或多個第二物件的該第二視訊訊框的一第二部分。

【請求項 8】 根據請求項 7 之裝置，其中為了產生該單個訊框，該一或多個處理器亦被配置為：

將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框。

【請求項 9】 根據請求項 8 之裝置，其中為了將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框，該一或多個處理器被配置為：

將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分放置在該單個訊框的相應訊框區域內，其中該相應訊框區域中的一第一區域繪示該一或多個第一物件，並且該相應訊框區域中的一第二區域繪示該一或多個第二物件。

【請求項 10】 根據請求項 1 之裝置，其中為了產生該單個

訊框，該一或多個處理器亦被配置為：

在該單個訊框內渲染該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分，其中該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分在視覺上被分開。

【請求項 11】 根據請求項 10 之裝置，其中該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分經由視覺標記在視覺上被分開，該視覺標記包括線、輪廓、方塊、高亮、標籤、顏色、陰影和視覺記號中的至少一個。

【請求項 12】 根據請求項 1 之裝置，其中該一或多個處理器被配置為：

從該第一影像擷取裝置獲得該第一視訊訊框，及從該第二影像擷取裝置獲得該第二視訊訊框。

【請求項 13】 根據請求項 12 之裝置，亦包括該第一影像擷取裝置和該第二影像擷取裝置中的至少一個。

【請求項 14】 根據請求項 1 之裝置，其中該一或多個處理器被配置為：

決定該一或多個物件在該第一影像擷取裝置的一視場（FOV）內；及

觸發該第一影像擷取裝置以擷取該第一視訊訊框，其中該第一視訊訊框圖示該一或多個物件。

【請求項 15】 根據請求項 14 之裝置，其中該一或多個處理器被配置為：

基於由該第一影像擷取裝置擷取的一或多個視訊訊框來追蹤該一或多個物件；

基於該追蹤，決定該一或多個物件不再位於該第一影像擷取裝置的該 FOV 內；及

基於該決定該一或多個物件不再位於該第一影像擷取裝置的該 FOV 內，決定停止從該第一影像擷取裝置向該遠端設備發送視訊資料。

【請求項 16】 根據請求項 15 之裝置，其中該一或多個處理器被配置為：

回應於決定至少一個物件在該第二影像擷取裝置的一 FOV 內，經由該第二影像擷取裝置獲得圖示該至少一個物件的一第三視訊訊框；及

向該遠端設備發出該第三視訊訊框。

【請求項 17】 根據請求項 1 之裝置，其中該裝置包括一行動設備。

【請求項 18】 一種用於處理視訊呼叫的方法，該方法包括：

在一設備與一遠端設備之間建立一視訊呼叫；

顯示一第一相機饋送和一第二相機饋送的一預覽，該第一相機饋送包括由該設備的一第一影像擷取裝置擷取的一第一視訊訊框和由該設備的一第二影像擷取裝置擷取的一第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；

接收對圖示在該預覽中的一物件集合的選擇；及

基於該第一相機饋送和該第二相機饋送來產生繪示該物件集合的單個訊框。

【請求項 19】根據請求項 18 之方法，其中該第一視訊訊框和該第二視訊訊框圖示包括該物件集合的多個物件，並且其中對該物件集合的該選擇包括一第一輸入和一第二輸入中的至少一個，該第一輸入選擇該物件集合作為該多個物件中要被包括在該單個訊框中的一子集，該第二輸入選擇該多個物件中要從該單個訊框中被排除的一或多個物件，該一或多個物件不同於該物件集合。

【請求項 20】根據請求項 19 之方法，其中產生該單個訊框亦包括：

基於該第一輸入和該第二輸入中的至少一個，從該單個訊框中排除該多個物件中的該一或多個物件；及
向該遠端設備發出該單個訊框。

【請求項 21】根據請求項 20 之方法，其中從該單個訊框中排除該一或多個物件亦包括從該預覽、該第一視訊訊框、該第二視訊訊框和該單個訊框中的至少一個中移除該一或多個物件。

【請求項 22】根據請求項 18 之方法，其中產生該單個訊框亦包括：

將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合成該單個訊框。

【請求項 23】根據請求項 22 之方法，其中將該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分組合

成該單個訊框包括：

將該第一視訊訊框的至少該部分和該第二視訊訊框的至少該部分佈置成該單個訊框的相應訊框區域，每個訊框區域圖示來自該物件集合的一相應物件。

【請求項 24】 根據請求項 18 之方法，其中該第一視訊訊框包括來自該物件集合的一或多個第一物件，並且該第二視訊訊框包括來自該物件集合的一或多個第二物件，並且其中產生該單個訊框亦包括：

基於與該第一視訊訊框相關聯的第一中繼資料來決定該一或多個第一物件之每一第一物件在該第一視訊訊框內的一相應位置，該第一中繼資料包括與該一或多個第一物件相關聯的座標；

基於與該第二視訊訊框相關聯的第二中繼資料來決定該一或多個第二物件之每一第二物件在該第二視訊訊框內的一相應位置，該第二中繼資料包括與該一或多個第二物件相關聯的座標；及

基於該一或多個第一物件之每一第一物件在該第一視訊訊框內的該相應位置及該一或多個第二物件之每一第二物件在該第二視訊訊框內的該相應位置，決定圖示該一或多個第一物件的該第一視訊訊框的一第一部分和圖示該一或多個第二物件的該第二視訊訊框的一第二部分。

【請求項 25】 根據請求項 24 之方法，其中產生該單個訊框亦包括：

將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框。

【請求項 26】根據請求項 25 之方法，其中將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分組合成該單個訊框包括：

將該第一視訊訊框的該第一部分和該第二視訊訊框的該第二部分放置在該單個訊框的相應訊框區域內，其中該相應訊框區域中的一第一區域繪示該一或多個第一物件，並且該相應訊框區域中的一第二區域繪示該一或多個第二物件。

【請求項 27】根據請求項 18 之方法，其中產生該單個訊框亦包括：

在該單個訊框內渲染該第一視訊訊框的至少一部分和該第二視訊訊框的至少一部分，其中該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分在視覺上被分開。

【請求項 28】根據請求項 27 之方法，其中該第一視訊訊框的該至少一部分和該第二視訊訊框的該至少一部分經由視覺標記在視覺上被分開，該視覺標記包括線、輪廓、方塊、高亮、標籤、顏色、陰影和視覺記號中的至少一個。

【請求項 29】根據請求項 18 之方法，亦包括：

從該第一影像擷取裝置獲得該第一視訊訊框，及從該第二影像擷取裝置獲得該第二視訊訊框。

【請求項30】根據請求項18之方法，亦包括：

決定該一或多個物件在該第一影像擷取裝置的一視場（FOV）內；及

觸發該第一影像擷取裝置以擷取該第一視訊訊框，其中該第一視訊訊框圖示該一或多個物件。

【請求項31】根據請求項30之方法，亦包括：

基於由該第一影像擷取裝置擷取的一或多個視訊訊框來追蹤該一或多個物件；

基於該追蹤，決定該一或多個物件不再位於該第一影像擷取裝置的該FOV內；及

基於該決定該一或多個物件不再位於該第一影像擷取裝置的該FOV內，決定停止從該第一影像擷取裝置向該遠端設備發送視訊資料。

【請求項32】包含指令的至少一種非暫時性電腦可讀取媒體，該等指令在由一或多個處理器執行時，使該一或多個處理器：

在一設備與一遠端設備之間建立一視訊呼叫；

顯示一第一相機饋送和一第二相機饋送的一預覽，該第一相機饋送包括由該設備的一第一影像擷取裝置擷取的一第一視訊訊框和由該設備的一第二影像擷取裝置擷取的一第二視訊訊框，該第一視訊訊框和該第二視訊訊框在該預覽內在視覺上被分開；

接收對圖示在該預覽中的一物件集合的選擇；及

基於該第一相機饋送和該第二相機饋送來產生繪示該

物件集合的單個訊框。

300 ↗

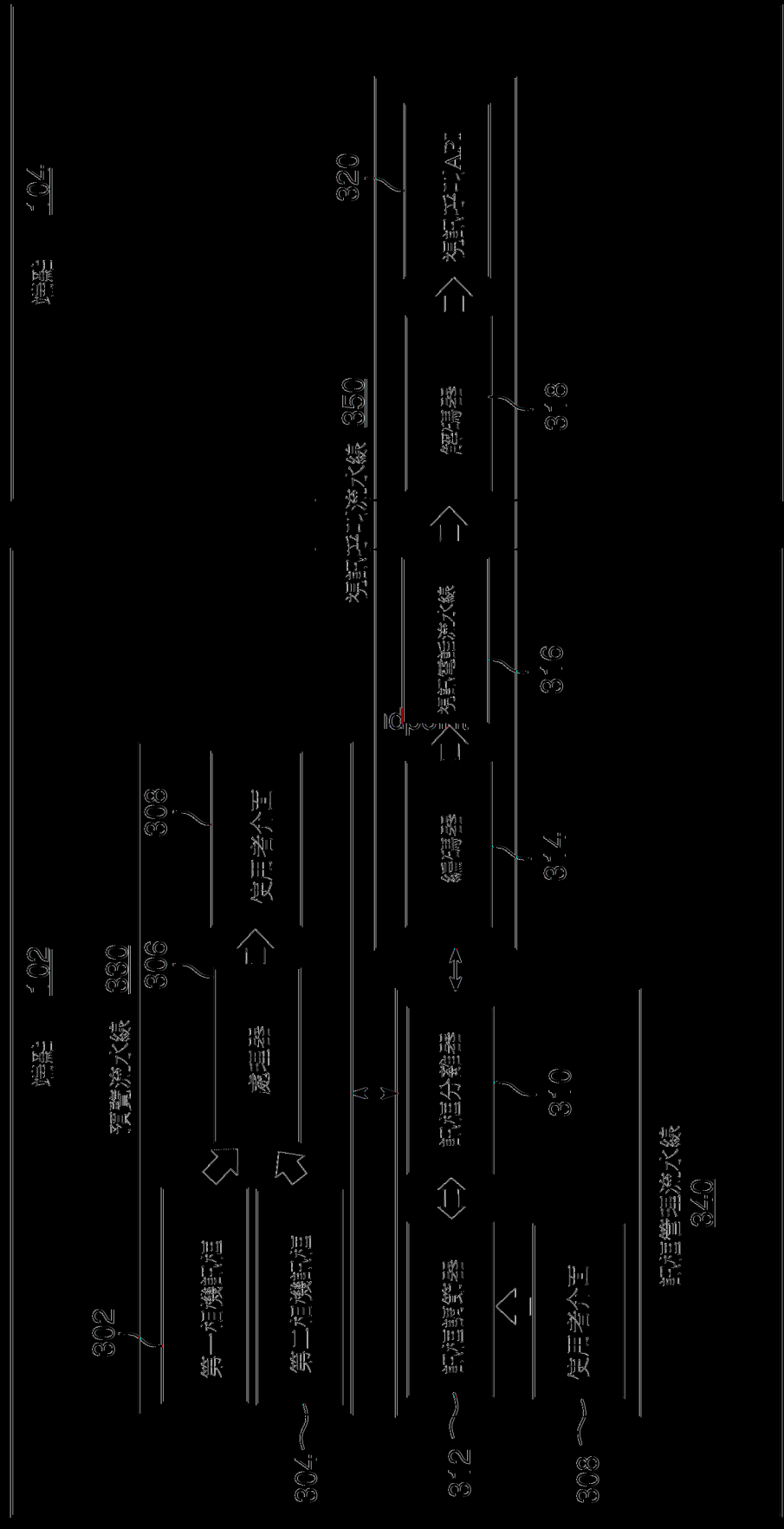




圖6A

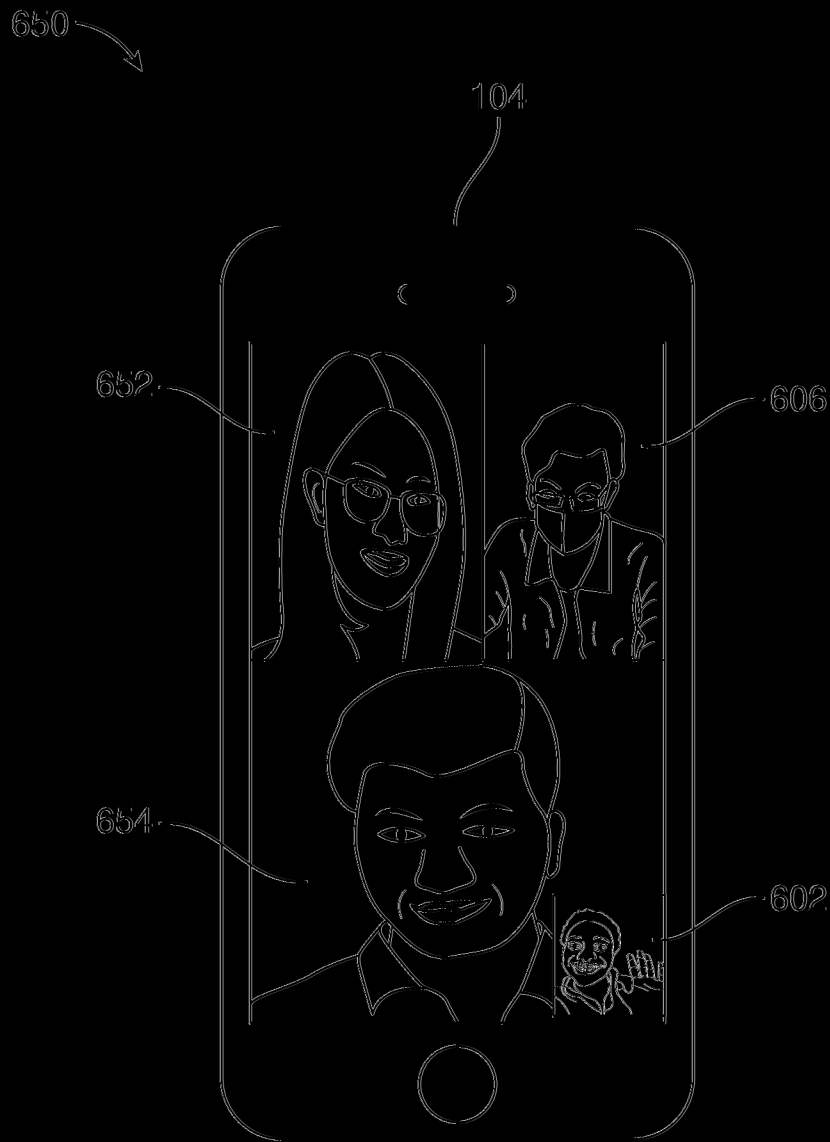


圖6B

