



- (21) 申請案號：110101276 (22) 申請日：中華民國 110 (2021) 年 01 月 13 日
- (51) Int. Cl. : **G01S7/41 (2006.01)** **G01S13/931 (2020.01)**
G01S13/86 (2006.01) **G01S13/89 (2006.01)**
G06K9/00 (2006.01) **G06K9/32 (2006.01)**
G06T7/62 (2017.01)
- (30) 優先權：2021/01/05 美國 17/142,199
- (71) 申請人：美商高通公司 (美國) QUALCOMM INCORPORATED (US)
美國
- (72) 發明人：烏尼克里許南 加雅克里許南 UNNIKRISHNAN, JAYAKRISHNAN (IN)；喬西阿福達 JOSHI, AVDHUT (IN)；艾卡瓦 席瓦母 AGARWAL, SHIVAM (IN)；Y 娜達拉簡 尤加 Y NADARAAJAN, YOGA (MY)；薩利米 阿米爾 SALIMI, AMIR (IR)；尼艾森 由爾斯 NIESEN, URS (US)；瓦德瑞 斯里席沙阿拉溫德 VADREVU, SREE SESA ARAVIND (IN)；薩德瓦 高譚 SACHDEVA, GAUTAM (US)
- (74) 代理人：李世章
- 申請實體審查：無 申請專利範圍項數：48 項 圖式數：19 共 123 頁

(54) 名稱

基於相機地圖及／或雷達資訊的物件尺寸估計

(57) 摘要

本發明提供了用於決定一或多個物件的一或多個尺寸的技術和系統。例如，可以獲取能夠辨識在圖像中偵測到的第一物件的邊界區域。亦可以獲取包括地圖點的地圖。地圖點對應於三維空間中的一或多個參考位置。辨識第一物件的邊界區域可以與地圖中包括的地圖點中的至少一個地圖點相關聯。使用邊界區域和至少一個地圖點，可以決定在圖像中偵測到的第一物件的估計三維位置和估計尺寸。在一些實例中，可以使用其他資訊來估計第一物件的估計三維位置和估計尺寸，例如雷達資訊及/或其他資訊。

Techniques and systems are provided for determining one or more sizes of one or more objects. For example, a bounding region identifying a first object detected in an image can be obtained. A map including map points can also be obtained. The map points correspond to one or more reference locations in a three-dimensional space. The bounding region identifying the first object can be associated with at least one map point of the map points included in the map. Using the bounding region and the at least one map point, an estimated three-dimensional position and an estimated size of the first object detected in the image can be determined. In some examples, other information can be used to estimate the estimated three-dimensional position and an estimated size of the first object, such as radar information and/or other information.

指定代表圖：

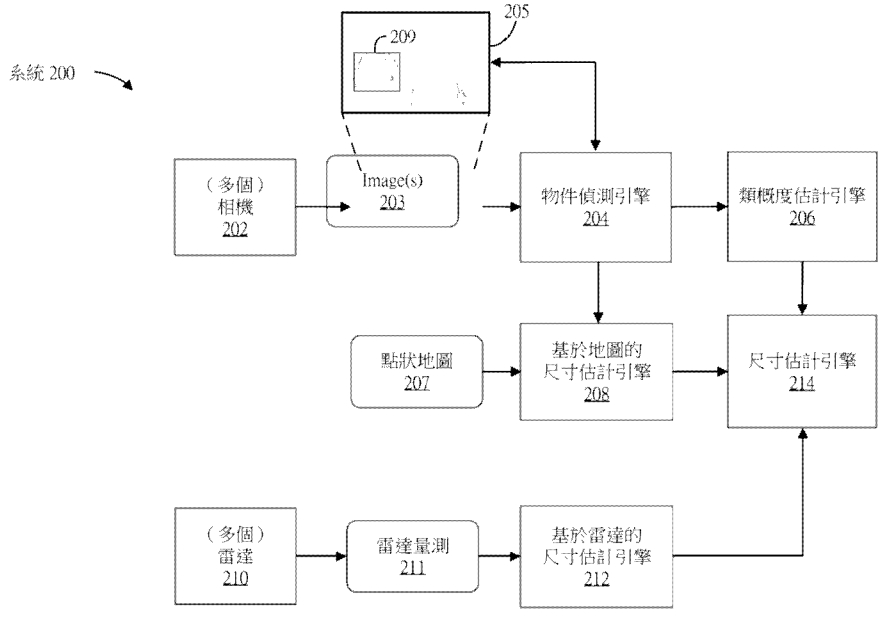


圖2

符號簡單說明：

- 200:系統
- 202:相機
- 203:圖像
- 204:物件偵測引擎
- 205:圖像
- 206:類概度估計引擎
- 207:點狀地圖
- 208:基於地圖的尺寸估計引擎
- 209:邊界區域
- 210:雷達
- 211:雷達量測值
- 212:基於雷達的尺寸估計引擎
- 214:尺寸估計引擎

【發明摘要】

【中文發明名稱】基於相機地圖及/或雷達資訊的物件尺寸估計

【英文發明名稱】 OBJECT SIZE ESTIMATION USING CAMERA MAP AND/OR
RADAR INFORMATION

【中文】

本發明提供了用於決定一或多個物件的一或多個尺寸的技術和系統。例如，可以獲取能夠辨識在圖像中偵測到的第一物件的邊界區域。亦可以獲取包括地圖點的地圖。地圖點對應於三維空間中的一或多個參考位置。辨識第一物件的邊界區域可以與地圖中包括的地圖點中的至少一個地圖點相關聯。使用邊界區域和至少一個地圖點，可以決定在圖像中偵測到的第一物件的估計三維位置和估計尺寸。在一些實例中，可以使用其他資訊來估計第一物件的估計三維位置和估計尺寸，例如雷達資訊及/或其他資訊。

【英文】

Techniques and systems are provided for determining one or more sizes of one or more objects. For example, a bounding region identifying a first object detected in an image can be obtained. A map including map points can also be obtained. The map points correspond to one or more reference locations in a three-dimensional space. The bounding region identifying the first object can be associated with at least one map point of the map points included in the map. Using the bounding region and the at least one map point, an estimated three-dimensional position and an estimated size of the first object detected in the image can be determined. In some examples, other

information can be used to estimate the estimated three-dimensional position and an estimated size of the first object, such as radar information and/or other information.

【指定代表圖】第（ 2 ）圖。

【代表圖之符號簡單說明】

2 0 0 : 系 統

2 0 2 : 相 機

2 0 3 : 圖 像

2 0 4 : 物 件 偵 測 引 擎

2 0 5 : 圖 像

2 0 6 : 類 概 度 估 計 引 擎

2 0 7 : 點 狀 地 圖

2 0 8 : 基 於 地 圖 的 尺 寸 估 計 引 擎

2 0 9 : 邊 界 區 域

2 1 0 : 雷 達

2 1 1 : 雷 達 量 測 值

2 1 2 : 基 於 雷 達 的 尺 寸 估 計 引 擎

2 1 4 : 尺 寸 估 計 引 擎

【特徵化學式】

無

【發明說明書】

【中文發明名稱】基於相機地圖及/或雷達資訊的物件尺寸估計

【英文發明名稱】 OBJECT SIZE ESTIMATION USING CAMERA MAP AND/OR
RADAR INFORMATION

【技術領域】

【0001】 本案大體而言係關於決定物件的尺寸（以及在某些情況下是位置），並且更具體地，係關於用於使用相機地圖、雷達資訊及/或其他資訊來決定物件的尺寸及/或位置的技術和系統。

【先前技術】

【0002】 物件偵測和追蹤可用於（例如，從數位圖像或視訊短片的視訊框中）辨識物件，並隨著時間追蹤該物件。物件偵測和追蹤可用於不同領域，包括自主駕駛、視訊分析、安全系統、機器人技術、航空等。在某些領域，物件可以決定環境中其他物件的位置，以便該物件可以準確地在環境中導航。此外，為了做出準確的運動規劃和軌跡規劃決策，物件應具有準確估計其他物件的尺寸的能力。但是，不同物件的尺寸可以變化，並且具有相同類別或分類的物件的尺寸亦可以變化。

【0003】 在物件需要能夠決定其他物件的位置和尺寸的領域的一個實例是經由（例如，自主車輛的）自主駕駛系統的自主駕駛。自主駕駛的關鍵要求（例如，在自主級別3及以上）是自主車輛偵測和追蹤該自主車輛周圍的其他車輛

的能力。儘管某些自主車輛能夠決定另一車輛的分類或類別，但即使在相同的分類或類別內，車輛的三維（3D）尺寸亦可能具有較大的差異。為了做出準確的運動規劃和軌跡規劃決策，自主駕駛系統應能夠準確估計道路上其他車輛的3D尺寸，尤其是長度。對於其他系統，例如機器人系統、航空系統（例如，無人駕駛飛行器）及/或具有類似需求的其他系統，類似的功能可能是有用的或必需的。

【發明內容】

【0004】 本文描述了用於使用各種類型的資訊（例如，相機地圖資訊、雷達資訊及/或其他資訊）來決定物件的尺寸（以及在某些情況下，位置及/或朝向）的系統和技術。根據一個說明性實例，提供了一種決定一或多個物件的一或多個尺寸的方法。該方法包括以下步驟：獲取辨識在圖像中偵測到的第一物件的邊界區域；獲取包括複數個地圖點的地圖，該複數個地圖點對應於三維空間中的一或多個參考位置；將辨識第一物件的邊界區域與地圖中包括的複數個地圖點中的至少一個地圖點相關聯；及使用邊界區域和至少一個地圖點，來決定在圖像中偵測到的第一物件的估計三維位置和估計尺寸。

【0005】 在另一實例中，提供了一種用於決定一或多個物件的一或多個尺寸的裝置，該裝置包括被配置為儲存一或多個圖像的記憶體，以及以電路系統實現並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且可以：獲取用於辨識圖像中偵測到的第一物件的邊界區域；獲取

包括複數個地圖點的地圖，該複數個地圖點對應於三維空間中的一或多個參考位置；將辨識第一物件的邊界區域與地圖中包括的複數個地圖點中的至少一個地圖點相關聯；及使用邊界區域和至少一個地圖點，來決定在圖像中偵測到的第一物件的估計三維位置和估計尺寸。

【0006】 在另一實例中，提供了一種非暫時性電腦可讀取媒體，其上儲存有指令，該等指令在由一或多個處理器執行時，使一或多個處理器執行以下操作：獲取辨識圖像中偵測到的第一物件的邊界區域；獲取包括複數個地圖點的地圖，該複數個地圖點對應於三維空間中的一或多個參考位置；將辨識第一物件的邊界區域與地圖中包括的複數個地圖點中的至少一個地圖點相關聯；及使用邊界區域和至少一個地圖點，來決定在圖像中偵測到的第一物件的估計三維位置和估計尺寸。

【0007】 在另一實例中，提供了一種用於決定一或多個物件的一或多個尺寸的裝置。該裝置包括：用於獲取辨識在圖像中偵測到的第一物件的邊界區域的構件；用於獲取包括複數個地圖點的地圖的構件，該複數個地圖點對應於三維空間中的一或多個參考位置；用於將辨識第一物件的邊界區域與地圖中包括的複數個地圖點中的至少一個地圖點相關聯的構件；及用於使用邊界區域和至少一個地圖點來決定在圖像中偵測到的第一物件的估計三維位置和估計尺寸的構件。

【0008】 在一些態樣，第一物件是車輛。在一些態樣，一或多個參考位置包括三維空間中道路上的車道。

【0009】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：將來自複數個地圖點的地圖點子集投影到圖像上；從該地圖點子集中決定至少一個地圖點是一或多個參考位置上離該邊界區域最近的地圖點；及基於決定至少一個地圖點是一或多個參考位置上離該邊界區域最近的地圖點，將該邊界區域與該至少一個地圖點相關聯。

【0010】 在一些態樣，使用校準資料和用於擷取圖像的相機的姿態，將地圖點子集投影到圖像上。在一些情況下，地圖點子集包括被包括在用於擷取圖像的相機的視野內、並且處於相機的選擇範圍內的地圖點。

【0011】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：決定至少一個地圖點與邊界區域的邊緣上的一或多個點相交。在此種態樣中，基於決定至少一個地圖點與邊界區域的邊緣上的一或多個點相交，該至少一個地圖點被決定為最近的地圖點。在一些情況下，邊界區域的邊緣包括邊界區域的底部邊緣。

【0012】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：決定邊界區域邊緣上的點；並且決定該至少一個地圖點與來自地圖點子集中的其他地圖點相比，最靠近邊界區域邊緣上的點。在此種態樣中，基於至少一個地圖點最靠近邊界區域邊緣上的點，將該至少一個地圖點決定為

最近的地圖點。在一些情況下，邊界區域的邊緣包括邊界區域的底部邊緣。

【0013】 在一些態樣，決定第一物件的估計三維位置和估計尺寸包括：獲取第一物件的寬度；決定第一物件的朝向在至少一個地圖點處平行於地圖的表面；基於第一物件的寬度和第一物件的朝向，決定代表第一物件的三維邊界框的頂點位置，其中該頂點位置對應於三維邊界框的最靠近用於擷取圖像的相機的拐角；基於該頂點位置，決定第一物件的估計三維位置；及基於頂點位置和邊界區域，決定代表第一物件的三維邊界框的長度和高度。

【0014】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：獲取第一物件的分類；及基於分類決定第一物件的寬度。

【0015】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：獲取圖像的第一物件的分類，該分類定義物件的類型；基於第一物件的分類，決定第一物件的最小尺寸和最大尺寸；及基於第一物件的最小尺寸和最大尺寸決定第一物件的估計尺寸。

【0016】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：基於第一物件和第二物件之間的距離決定分類的權重，第二物件包括用於擷取圖像的相機；基於該權重更新分類的累積權重，累積權重基於來自包括第一物件的多個圖像的多次偵測；基於分類的累積權重決定第一物件包括由分類定義的物件類型的概度性；基於所決定的概度性

和與一或多個其他分類相關聯的一或多個概度性，決定第一物件包括物件的類型；及基於由分類定義的物件類型，來決定第一物件的最小尺寸和最大尺寸。

【0017】 在一些態樣，第一物件是第一車輛，第二物件是第二車輛。在一些態樣，由分類定義的物件類型包括第一車輛的車輛類型。

【0018】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：獲取複數個雷達量測點，該複數個雷達量測點基於由第一物件反射的雷達信號；及基於複數個雷達量測來決定第一物件的附加估計尺寸。在一些態樣，使用第二物件上包括的複數個雷達獲取複數個雷達量測點，第二物件包括用於擷取圖像的相機。

【0019】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：決定雷達量測點沿著第一物件的朝向方向的最大擴展；及基於雷達量測點沿著第一物件的朝向方向的最大擴展，來決定第一物件的附加估計尺寸。

【0020】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：使用估計模型決定第一物件的尺寸，估計尺寸和附加估計尺寸被用作估計模型的輸入。在一些態樣，估計模型是卡爾曼濾波器。

【0021】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：從成像雷達獲取雷達圖像；及基於雷達圖像決定第一物件的附加估計尺寸。

【0022】 在一些態樣，上述方法、裝置和電腦可讀取媒體亦包括：使用估計模型決定第一物件的尺寸，估計尺寸和附加估計尺寸被用作估計模型的輸入。

【0023】 在一些實例中，提供了一種車輛，其包括用於決定一或多個物件的一或多個尺寸的元件。該車輛可以包括被配置為擷取複數個圖像的一或多個相機、被配置為儲存複數個圖像的記憶體，以及以電路系統實現並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並且可以執行上述任何技術。例如，一或多個過程被配置為並且能夠：獲取辨識在由一或多個相機中的相機擷取的圖像中偵測到的目標車輛的邊界區域；獲取包括複數個地圖點的地圖，該複數個地圖點對應於道路上車道的一或多條線；將辨識目標車輛的邊界區域與包括在地圖中的複數個地圖點中的至少一個地圖點相關聯；及使用邊界區域和至少一個地圖點來決定在圖像中偵測到的目標車輛的估計三維位置和估計尺寸。

【0024】 在一些態樣，該裝置是行動設備（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、擴展現實設備（例如，虛擬實境（VR）設備、增強現實（AR）設備或混合現實（MR）設備）、個人電腦、膝上型電腦、伺服器電腦、相機、車輛或計算設備或車輛的元件（例如，自主車輛）或其他設備，該裝置是以上該等設備的一部分及/或包括該等設備。在一些態樣，該裝置包括用於擷取一或多個圖像的一或多個相機。在一些態樣，

該裝置亦包括用於顯示一或多個圖像、通知及/或其他可顯示資料的顯示器。在一些態樣，上述裝置可以包括一或多個感測器（例如，一或多個慣性量測單元（IMU），例如一或多個陀螺儀、一或多個加速度計、其任意組合及/或其他感測器）。

【0025】 本概述不意欲辨識所主張保護的標的的關鍵或必要特徵，亦不意欲單獨用於決定所主張保護的標的的範疇。該標的應該經由參考本專利案的整個說明書的適當部分、任何或所有附圖，以及每個請求項來理解。

【0026】 經由參考以下說明書、請求項和附圖，前述內容連同其他特徵和實施例將變得更加顯而易見。

【圖式簡單說明】

【0027】 下文參考以下附圖詳細描述本案的說明性實施例：

【0028】 圖 1 是圖示根據一些實例的在道路上行駛的多個車輛的圖像；

【0029】 圖 2 是圖示根據一些實例的用於決定物件的尺寸、位置和朝向的系統的實例的方塊圖；

【0030】 圖 3 是圖示根據一些實例的從對應於環境中的參考位置的點地圖中決定點的子集的實例的圖；

【0031】 圖 4 是圖示根據一些實例的將地圖點子集從點地圖投影到圖像的圖像平面的實例的圖；

【0032】 圖 5 A 是圖示根據一些實例的表示來自追蹤物件的前置 (front - facing) 相機的圖像平面的圖像的實例的圖 ；

【0033】 圖 5 B 是圖示根據一些實例的表示來自追蹤物件的側面 (side - facing) 相機的圖像平面的圖像的實例的圖 ；

【0034】 圖 6 是圖示根據一些實例的將目標物件的邊界區域與來自地圖中地圖點子集的点相關聯的實例的圖 ；

【0035】 圖 7 是圖示根據一些實例的將目標物件的邊界區域與來自地圖中地圖點子集的点相關聯的實例的圖像 ；

【0036】 圖 8 是圖示根據一些實例的將目標物件的邊界區域與來自圖像的点相關聯的另一實例的圖像 ；

【0037】 圖 9 A 是圖示根據一些實例的將目標物件的三維 (3 D) 邊界框的拐角與目標物件的二維 (2 D) 邊界框的邊緣相關的幾何約束的實例的圖 ；

【0038】 圖 9 B 是圖示根據一些實例的將目標物件的三維 (3 D) 邊界框的拐角與目標物件的二維 (2 D) 邊界框的邊緣相關的幾何約束的另一實例的圖 ；

【0039】 圖 1 0 是圖示根據一些實例的由各種雷達感測器接收的雷達量測的縱向分量相對於時間繪製的實例的曲線圖 ；

【0040】 圖 1 1 是圖示根據一些實例的使用圖 1 0 所示的雷達量測決定的目標物件的縱向末端的估計的實例的曲線圖 ；

【0041】 圖 1 2 A 和 圖 1 2 B 是 根 據 一 些 實 例 的 雷 達 圖 像 的 說 明 性 實 例 。

【0042】 圖 1 3 是 圖 示 根 據 一 些 實 例 的 深 度 學 習 網 路 的 實 例 的 方 塊 圖 。

【0043】 圖 1 4 是 圖 示 根 據 一 些 實 例 的 迴 旋 神 經 網 路 的 實 例 的 方 塊 圖 。

【0044】 圖 1 5 是 圖 示 根 據 一 些 實 例 的 C i f a r - 1 0 神 經 網 路 的 實 例 的 圖 。

【0045】 圖 1 6 A - 圖 1 6 C 是 圖 示 根 據 一 些 實 例 的 單 次 物 件 偵 測 器 的 實 例 的 圖 。

【0046】 圖 1 7 A - 圖 1 7 C 是 圖 示 根 據 一 些 實 例 的 你 僅 需 看 一 次 (Y O L O) 偵 測 器 的 實 例 的 圖 。

【0047】 圖 1 8 是 圖 示 根 據 一 些 實 例 的 使 用 此 處 描 述 的 技 術 來 決 定 一 或 多 個 物 件 的 一 或 多 個 尺 寸 的 過 程 的 實 例 的 流 程 圖 ； 及

【0048】 圖 1 9 是 根 據 一 些 實 例 的 可 用 於 實 現 此 處 描 述 的 技 術 的 一 些 態 樣 的 示 例 性 計 算 設 備 的 方 塊 圖 。

【實施方式】

【0049】 下 文 提 供 了 本 案 的 某 些 態 樣 和 實 施 例 。 對 熟 習 此 項 技 術 者 而 言 顯 而 易 見 的 是 ， 該 等 態 樣 和 實 施 例 中 的 一 些 可 以 獨 立 應 用 ， 並 且 其 中 一 些 可 以 組 合 應 用 。 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 的 實 施 例 的 透 徹 理 解 。 然 而 ， 顯 而 易 見 的 是 ， 沒 有 該 等 具

體細節亦可以實施各種實施例。附圖和描述並不意欲限制性。

【0050】 隨後的描述僅提供示例性實施例，並不意欲限制本案的範疇、適用性或配置。相反，隨後對示例性實施例的描述將為熟習此項技術者提供實現示例性實施例的賦能描述。應當理解，在不脫離所附請求項中闡述的本案的精神和範疇的情況下，可以對元件的功能和佈置進行各種改變。

【0051】 物件偵測和追蹤可用於辨識物件並隨時間追蹤物件。例如，可以獲取物件的圖像，並且可以對圖像執行物件偵測以偵測圖像中的一或多個物件。在一些情況下，偵測到的物件可以被分類到物件類別中，並且可以產生邊界區域來辨識物件在圖像中的位置。各種類型的系統可用於物件偵測，包括基於神經網路的物件偵測器。

【0052】 物件偵測和追蹤可用於自主駕駛系統、視訊分析、安全系統、機器人系統、航空系統等系統。在此種系統中，在環境中追蹤其他物件（被稱為目標物件）的物件（被稱為追蹤物件）可以決定其他物件的位置和尺寸。決定環境中目標物件的位置和尺寸允許追蹤物件經由做出智慧運動規劃和軌跡規劃決策，來精確地在環境中導航。然而，很難精確辨識目標物件的尺寸。例如，由於具有相同類別或分類的物件的尺寸不同，基於目標物件的分類或類別來決定目標物件的尺寸可能是困難的。

【0053】 在追蹤物件需要能夠決定目標物件的位置和尺寸的領域的一個實例是經由自主駕駛系統的自主駕駛（例如，自主車輛）。自主駕駛系統的一個重要目標是自主車輛能夠偵測和追蹤自主車輛周圍的其他車輛。對於更高的自主級別，例如自主級別3及以上，情況尤其如此。例如，自主級別0需要駕駛員的完全控制，因為車輛沒有自主駕駛系統，而自主級別1涉及基本的協助特徵，如巡航控制，在此種情況下，車輛駕駛員完全控制車輛。自主級別2是指半自主駕駛，車輛可以執行直線行駛、停留在特定車道、控制與車輛前方的其他車輛的距離，或擁有一些其他功能等功能。自主級別3、4和5包括更多的自主。例如，自主級別3指的是在某些情況下可以接管所有駕駛功能的車載自主駕駛系統，在此種情況下，若需要，駕駛員隨時準備接管。自主級別4指的是完全自主的體驗，即使在複雜的駕駛情況下（例如，在高速公路和繁忙的城市交通中），亦不需要使用者的幫助。利用自主級別4，人仍然可以留在方向盤後面的駕駛座上。在自主級別4下運行的車輛可以通訊並通知其他車輛即將進行的操作（例如，車輛正在變道、轉彎、停車等）。自主級別5車輛是可在所有條件下自主運行的完全自主的自動駕駛車輛。車輛不需要人操作員採取任何行動。

【0054】 圖1是圖示包括在道路上行駛的眾多車輛的環境的圖像100。車輛包括追蹤車輛102、目標車輛104、目標車輛106和目標車輛108。追蹤車輛102是以特定的自主級

別運行的自主車輛。追蹤車輛 102 可以追蹤目標車輛 104、106 和 108，以便導航環境。例如，追蹤車輛 102 可以決定目標車輛 104 的位置和尺寸，以決定何時減速、加速、變道及 / 或執行一些其他功能。儘管車輛 102 被稱為追蹤車輛 102，並且車輛 104、106 和 108 相對於圖 1 被稱為目標車輛，但是若當車輛 104、106 和 108 正在追蹤其他車輛時，車輛 104、106 和 108 亦可以被稱為追蹤車輛，其中其他車輛成為目標車輛。

【0055】 儘管一些自主車輛可能（例如，基於物件偵測和分類）能夠決定另一車輛的分類或類別，但是車輛的三維（3D）尺寸甚至在相同的分類或類別中亦可能具有很大的差異。例如，「卡車」的車輛類別可以包括許多不同形狀和尺寸的卡車，包括小型卡車、中型卡車和大型卡車。實際上，一些卡車，如半掛卡車和移動卡車，比小型卡車大很多倍。精確估計道路上其他車輛的 3D 尺寸（包括長度）是自主車輛的自主駕駛系統的重要特徵，以能夠做出精確的運動規劃和軌跡規劃決策。

【0056】 其他類型的系統亦可以受益於精確決定目標物件的尺寸及 / 或位置。例如對物件執行操作的機器人系統可能需要能夠精確地估計物件的 3D 尺寸。在一個說明性實例中，用於製造設備的機器人設備需要知道設備及其元件的尺寸、維度和位置，以便構建設備。在另一個說明性實例中，航空系統（例如，無人駕駛飛行器等）可以受益於飛行路徑中物件的精確決定，使得航空系統可以精確地在物件

周圍導航。亦存在許多其他需要能夠辨識物件尺寸和位置的系統的實例。

【0057】 本文描述了用於使用各種類型的資訊（例如相機資訊、地圖資訊、雷達資訊及/或其他資訊）來決定物件的尺寸和位置的系統、裝置、方法（亦稱為過程）和電腦可讀取媒體（統稱為「系統和技術」）。系統可以包括在追蹤一或多個其他物件（稱為目標物件）的物件（稱為追蹤物件）中，並且該等技術可以由該物件執行。在一些實例中，基於地圖的逆透視映射（IPM）技術可以由追蹤物件來執行，以決定估計的三維（3D）位置及/或朝向，以及在圖像或多個圖像中偵測到的目標物件的估計尺寸。（多個）圖像可以由位於追蹤物件中或追蹤物件上的一或多個相機擷取。基於地圖的IPM技術可以使用點狀地圖。點狀地圖可以包括對應於3D空間中的一或多個參考位置的複數個地圖點。來自點狀地圖的複數個地圖點的子集可以被投影到其中偵測到目標物件的圖像上。使用基於地圖的位置關聯函數，可以將對應於偵測到的目標物件的邊界區域（例如，邊界框、邊界橢圓或其他邊界區域）與來自被投影的地圖點子集的地圖點相關聯。一旦地圖上的地圖點與被偵測到的目標物件的邊界區域相關聯，則邊界區域和地圖點可用於決定目標物件的估計3D位置（及/或朝向）和估計尺寸。例如，給定物件的寬度（例如，基於物件的給定分類或類別）和目標物件的朝向，可以決定目標物件的長度和高度連同表示目標物件的3D邊界框的拐角的位置。

置。3 D 邊界框的拐角的位置提供了目標物件的 3 D 位置。基於地圖的 IP M 技術可以在其中偵測到目標物件的單個圖像或多個圖像上執行。

【0058】 在一些實例中，對於在一或多個圖像中偵測到的給定目標物件，可由追蹤物件執行概度估計，以追蹤目標物件包括各種分類（在此亦稱為類）的概度性。例如，概度濾波器可用於經由組合來自經由從一或多個多相機的偵測所提供的類標籤的資訊，來在目標物件的追蹤歷史上追蹤各種類的概度性。例如，給定輸入圖像，可以執行物件偵測技術來偵測目標物件，並為目標物件提供類（或在某些情況下多個類）。可以使用物件偵測來處理物件的多個圖像，從而有可能為目標物件決定多個類。可以向為目標物件決定的各種類之每一者類提供權重。在一些情況下，權重可以基於目標物件和包括用於擷取一或多個圖像的一或多個相機的追蹤物件之間的距離。在某些情況下，特定類的權重亦可以基於由物件偵測所輸出的置信度得分。置信度得分指示目標物件屬於給定類的置信度（例如，概率）。偵測到的目標物件的（來自多個類）的類的最佳估計被決定為在目標物件的偵測歷史上累積的具有最高概度性的類。由概度估計所估計的物件類可用於界定目標物件的尺寸（例如，長度、寬度及 / 或高度）的上限和下限。當估計目標物件的尺寸時，可以使用該上限和下限。

【0059】 在一些實例中，追蹤物件可以使用雷達資訊來估計目標物件的尺寸。例如，在週期性時刻，可以在最近的

雷達訊框中從位於追蹤物件上的多個雷達感測器獲取目標物件反射的雷達信號。可以對來自多個雷達感測器的雷達信號的雷達量測值進行聯合處理，並且可以使用沿著目標物件的朝向方向的雷達量測點的最大縱向擴展，來決定目標物件的尺寸（例如，長度或其他尺寸或維度）的暫態估計。在一些實施方式中，來自基於雷達的尺寸估計的尺寸（例如，長度）的最佳估計被決定為一定數量的尺寸估計中給定的最大暫態估計。

【0060】 基於雷達的尺寸估計可以單獨使用，或者與基於地圖的尺寸估計及/或概度估計結合使用，以決定物件的尺寸。例如，在一些情況下，估計模型可以考慮基於地圖的尺寸估計、基於雷達的尺寸估計及/或概度估計，來決定目標物件的最終估計尺寸。在一些實施方式中，估計模型可以包括充當卡爾曼濾波器的估計框架（例如，貝氏估計框架或其他估計模型框架）。

【0061】 圖 2 是圖示用於決定環境中物件的尺寸及/或位置的系統 200 的實例的方塊圖。系統 200 可以包括在追蹤一或多個目標物件的追蹤物件中。如前述，追蹤物件是指追蹤一或多個被稱為目標物件的其他物件的物件。在一個說明性實例中，系統 200 可以包括自主車輛中包括的自主駕駛系統（作為追蹤物件的實例）。在另一個說明性實例中，系統 200 可以包括機器人設備或系統中包括的自主導航系統。儘管為了說明的目的在此處使用自主駕駛系統和自主車輛來描述實例，但是一般技術者將會理解，在此描述的

系統 200 和相關技術可以被包括在用於決定物件的尺寸及/或位置的任何其他系統或設備中，並且由該等系統或設備來執行。

【0062】 系統 200 可用於使用二維（2D）邊界區域偵測和來自基於相機的物件偵測的相應物件類型分類、使用來自雷達的點偵測、使用雷達圖像、使用其組合及/或使用其他資訊，來估計環境中的物件尺寸。在一些情況下，除了使用 2D 邊界區域偵測、類型分類及/或雷達點偵測之外或作為其替代，系統 200 可以使用來自成像雷達的物件偵測來估計環境中物件的尺寸。在一個說明性實例中，系統 200 可以使用 2D 邊界框偵測和來自相機的相應的車輛類型分類、來自雷達的點偵測以及可選地來自成像雷達的物件偵測，來估計在道路上偵測到的目標車輛的尺寸。如下文更詳細描述的，系統 200 可以應用以下一或多個的任意組合：基於相機的物件類型概度濾波器、用於物件（例如，車輛或其他物件）維度估計的相機地圖融合技術、基於雷達的長度估計技術及/或基於成像雷達的物件偵測，並且可以使用由基於地圖的尺寸決定、基於雷達的尺寸估計及/或成像雷達偵測所提供的量測來實現估計模型，以追蹤物件的尺寸（例如，長度及/或其他尺寸維度）的最佳估計。

【0063】 系統 200 包括各種元件，包括一或多個相機 202、物件偵測引擎 204、分類（類）概度估計引擎 206、基於地圖的尺寸估計引擎 208、一或多個雷達 210、基於雷達的尺寸估計引擎 212 和尺寸估計引擎 214。系統 200 的元

件可以包括軟體、硬體或兩者。例如，在一些實施方式中，系統 200 的元件可以包括及/或可以使用電子電路或其他電子硬體來實現，電子電路或其他電子硬體可以包括一或多個可程式設計電子電路（例如，微處理器、圖形處理單元（GPU）、數位信號處理器（DSP）、中央處理單元（CPU）及/或其他合適的電子電路），及/或可以包括及/或使用電腦軟體、韌體或其任意組合來實現，以執行此處描述之各種操作。該軟體及/或韌體可以包括儲存在電腦可讀取儲存媒體上並可由實現系統 200 的計算設備中的一或多個處理器執行之一或多個指令。

【0064】 儘管系統 200 被示為包括某些元件，但是一般技術者將會理解，系統 200 可以包括比圖 2 所示的元件更多或更少的元件。例如，系統 200 可以包括計算設備或物件，或者是該計算設備或物件的一部分，該計算設備或物件包括一或多個輸入設備和一或多個輸出設備（未圖示）。在一些實施方式中，系統 200 亦可以包括計算設備或可以是計算設備的一部分，該計算設備包括一或多個記憶體設備（例如，一或多個隨機存取記憶體（RAM）元件、唯讀記憶體（ROM）元件、快取記憶體元件、緩衝元件、資料庫元件及/或其他記憶體設備）、與一或多個記憶體設備通訊及/或電連接到一或多個記憶體設備的一或多個處理設備（例如，一或多個 CPU、GPU 及/或其他處理設備）、用於執行無線通訊之一或多個無線介面（例如，包括一或多個收發機和用於每個無線介面的基頻處理器）、用於經由

一或多個硬佈線連接執行通訊的一或多個有線介面（例如，諸如通用序列匯流排（USB）輸入的序列介面、照明連接器及/或其他有線介面），及/或圖2中未圖示的其他元件。

【0065】 如前述，系統200可以由計算設備或其他物件實現及/或包含在計算設備或其他物件中。在一些情況下，可以使用多個計算設備來實現系統200。例如，用於實現系統200的計算設備可以包括作為設備或物件的一部分的電腦或多台電腦，例如車輛、機器人設備、監控系統及/或具有執行本文所描述的技術的資源能力的任何其他計算設備或物件。在一些實施方式中，系統200可以與一或多個軟體應用程式整合（例如，整合到軟體中，作為一或多個外掛程式被添加，作為一或多個庫函數被包括，或者以其他方式與一或多個軟體應用程式整合），例如自主駕駛或導航軟體應用程式或軟體應用程式套件。一或多個軟體應用程式可以安裝在實現系統200的計算設備或物件上。

【0066】 系統200的一或多個相機202可以擷取一或多個圖像203。在一些情況下，一或多個相機202可以包括多個相機。例如，包括系統200的自主車輛可以在車輛的前面具有相機或多個相機，在車輛的後面具有相機或多個相機，在車輛的每一側具有相機或多個相機，及/或其他相機。在另一個實例中，包括系統200的機器人設備可以在機器人設備的各個部分上包括多個相機。在另一個實例

中，包括系統 200 的航空設備可以在航空設備的不同部分上包括多個相機。

【0067】 一或多個圖像 203 可以包括靜止圖像或視訊訊框。一或多個圖像 203 中的每一個包含場景的圖像。圖 2 中圖示圖像 205 的實例。圖像 205 圖示由包括多個目標車輛的追蹤車輛的相機擷取的圖像的實例。當擷取視訊訊框時，視訊訊框可以是一或多個視訊序列的一部分。在一些情況下，由一或多個相機 202 擷取的圖像可被儲存在儲存設備（未圖示）中，並且一或多個圖像 203 可以被從儲存設備中取得或以其他方式獲取。一或多個圖像 203 可以是由可選地具有深度圖的圖元（或體素）組成的點陣圖像、由向量或多邊形組成的向量圖像，或其組合。圖像 203 可以包括沿著一或多個平面（例如，水平或 x 方向上的平面和垂直或 y 方向上的平面）的場景的一或多個二維表示，或者該場景的一或多個三維表示。

【0068】 物件偵測引擎 204 可以獲取並處理一或多個圖像 203，以偵測及 / 或追蹤一或多個圖像 203 中的一或多個物件。物件偵測引擎 204 可以輸出物件作為偵測和追蹤的物件。物件偵測引擎 204 可以決定在圖像中偵測到的每個物件的分類（稱為類）或類別，並且亦可以產生用於辨識圖像之每一者物件的邊界區域（例如，辨識圖像 205 中的目標車輛的邊界區域 209）。例如，可以在圖像中偵測物件，並且物件偵測引擎 204 可以為偵測到的物件輸出邊界區域和類標籤（亦稱為類別標籤）。系統 200 的其他元件可以

使用邊界區域來辨識包括偵測到的物件的圖像的區域。在一些情況下，邊界區域的維度（例如，寬度及/或高度、對角線的長度，例如從左下角到右上角、從左上角到右上角，或其他維度）亦可以由物件偵測引擎 204 輸出。分配給偵測到的物件的邊界區域可以包括邊界框、邊界圓、邊界橢圓或表示偵測到的物件的任何其他合適形狀的區域。儘管為了說明的目的在此使用邊界框描述了實例，但是在此描述的技術和系統亦可以使用其他合適形狀的邊界區域以應用。與偵測到的物件相關聯的邊界框可以具有矩形、正方形或其他合適的形狀。在一些情況下，物件偵測引擎 204 可以輸出偵測到的物件的多個類、連同指示該物件屬於每個類的置信度的置信度得分（例如，該物件是汽車的置信度得分為 0.85，該物件是卡車的置信度得分為 0.14，以及該物件是摩托車的置信度得分為 0.01）。

【0069】 物件偵測引擎 204 可以執行任何合適的物件偵測及/或分類技術。在一些情況下，物件偵測引擎 204 可以使用基於機器學習的物件偵測器，例如使用一或多個神經網路。例如，基於深度學習的物件偵測器可以用於偵測和分類一或多個圖像 203 中的物件。在一個說明性實例中，基於 C i f a r - 1 0 神經網路的偵測器可用於執行物件分類以對物件進行分類。在某些情況下，可以訓練 C i f a r - 1 0 偵測器僅對某些物件進行分類，例如僅對車輛進行分類。下文參照圖 15 描述 C i f a r - 1 0 偵測器的更多細節。

【0070】 基於深度學習的偵測器的另一個說明性實例是快速單次（single-shot）物件偵測器（SSD），其包括神經網路並且可以應用於多個物件類別。SSD模型的特徵是使用多尺度迴旋邊界框輸出，該等輸出附接到神經網路頂部的多個特徵映射。此種表現允許SSD有效地建模不同的邊界框形狀。已經證明，給定相同的VGG-16基本架構，SSD在精度和速度態樣兩者皆媲美其最先進的物件偵測器配對物。在K. Simonyan和A. Zisserman的「Very deep convolutional networks for large-scale image recognition」(CoRR, abs/1409.1556, 2014)中，更詳細地描述了SSD深度學習偵測器，其全部內容為了所有目的經由引用的方式結合於此。以下將參照圖16A-圖16C對SSD偵測器的更多細節進行描述。

【0071】 可用於偵測和分類一或多個圖像203中的物件的基於深度學習的偵測器的另一個說明性實例包括你僅需看一次（YOLO）偵測器。YOLO偵測器在Titan X上執行時，以40-90fps的速度、78.6%的mAP處理圖像（基於VOC 2007）。YOLO深度學習偵測器在J. Redmon、S. Divvala、R. Girshick和A. Farhadi的「You only look once: Unified, real-time object detection」(arXiv preprint arXiv:1506.02640, 2015)中進行了詳細描述，在此為了所有目的經由引用將其全部內容併入本文。下文參照圖17A-圖17C描述YOLO偵測器的進一步細節。儘管Cifar-10、SSD和YOLO偵測器被提供作

為基於深度學習的物件偵測器的說明性實例，但是一般技術者將理解，任何其他合適的物件偵測和分類可以由物件偵測引擎 204 來執行。

【0072】 對於從中偵測到一或多個目標物件的給定圖像，類概度估計引擎 206 可以獲取邊界區域和為一或多個目標物件決定的一或多個類。例如，對於在圖像中偵測到的目標物件，類概度估計引擎 206 可以獲取由物件偵測引擎 204 為目標物件決定的邊界區域和類（或多個類）。類概度估計引擎 206 可以使用每個目標物件的（多個）類和邊界區域，來決定目標物件是特定類的物件的概度性。在一些情況下，類概度估計引擎 206 可以包括概度濾波器。概度濾波器可用於經由基於從一或多個相機 202 提供的圖像的處理，組合來自由物件偵測引擎 204 為特定目標物件輸出的各種類標籤的資訊，來在特定目標物件的追蹤歷史上追蹤特定目標物件包括各種類的概度性。例如，可以處理來自一或多個相機的目標物件的多個圖像，並且可以在多個圖像的每一個中偵測目標物件。目標物件的追蹤歷史可以包括自圖像中該目標的第一次偵測以來、所有圖像中的該目標物件的偵測。為每個圖像中的目標物件的每次偵測決定類，導致可以跨多個圖像為目標物件決定多個類的概度性。多個類可以由概度濾波器處理，以決定目標物件是特定類的物件的概度性。基於該概度性，類概度估計引擎 206 可以為目標物件決定最可能的類。

【0073】 類概度估計引擎 206 可以跨多個圖像決定用於為目標物件決定的各種類中每個類的權重。更高的權重被決定以用於與被認為更確定的物件偵測結果相關聯的類。物件偵測結果的確定性以及由該物件偵測結果產生的關聯類可基於目標物件與包括系統 200 的追蹤物件之間的距離，在此種情況下，分配給該類的權重可以基於該距離。在某些情況下，用於決定類權重的距離可以基於圖像中偵測到的目標物件與用於擷取該圖像的追蹤物件上的相機之間的距離。在某些情況下，類的權重亦可以基於物件偵測所輸出的置信度得分，該置信度得分指示目標物件屬於給定類的置信度（例如，概率）。

【0074】 在一個說明性實例中，對於每個類 c ，術語 cw_c 表示對於給定的追蹤物件已接收到的所有被觀測的類的累積權重。每當類概度估計引擎 206 接收到類 c 的新偵測 y 時，類 c 的累積權重 cw_c 可以被更新以包括分配給新偵測的權重 w_y 。在一個實例中，以下公式可用於更新類 c 的累積權重 cw_c ：

$$cw_c = cw_c + w_y$$

方程式（1），

【0075】 其中 w_y 可被決定如下：

$$w_y = \frac{A}{B + \left\| x_{\text{目標}} - x_{\text{追蹤}} \right\|}$$

方程式（2）。

【0076】 在方程式（2）中，項 A 和 B 是正常數， $x_{\text{目標}}$ 是偵測時目標物件的估計位置， $x_{\text{追蹤}}$ 是追蹤物件的估計位置（或追

蹤物件上用於擷取目標物件的圖像的相機的估計位置)。可以將正常數 A 和 B 設置為任何合適的值。在一個說明性實例中，假設位置以米來表示，則 A 等於 50 米 (m)，而 B 等於 5 m。使用車輛作為說明性實例，當追蹤車輛 (有時稱為自我 (ego) 車輛) 正在追蹤目標車輛時， $x_{\text{目標}}$ 是在圖像中偵測到目標車輛時目標車輛的估計位置，並且 $x_{\text{追蹤}}$ 是在圖像中偵測到目標車輛時追蹤 (自我) 車輛的估計位置 (或用於擷取目標車輛的圖像的相機的估計位置)。從方程式 (2) 可以看出，對於給定的目標物件的偵測，目標物件與追蹤物件 (或其相機) 之間的距離越大，由該目標物件的偵測決定的類的權重 w_y 越大。如方程式 (1) 所示，分配給由目標物件的給定偵測所決定的類 c 的權重 w_y 被添加到類 c 的累積權重 cw_c 中。可以使用方程式 (1) 和 (2) 或使用另一種合適的技術，來更新為目標物件決定的每個不同的類。

【0077】 類 c 的概度性可被決定為以下比率：

$$L_c = \frac{cw_c}{\sum_i cw_i}$$

方程式 (3) ，

【0078】 其中 cw_i 表示在所分析的目標物件的所有偵測中不同類 i 的累積權重。在一個說明性實例中，累積權重 cw_1 可被保持用於第一類 (例如，「汽車」類)，累積權重 cw_2 可被保持用於第二類 (例如，「卡車」類)，並且累積權重 cw_3 可被保持用於第三類 (例如，「摩托車」類)，其中目標物件在其追蹤歷史中被分類。在此種實例中，對於其中目標物件被偵測和分類的當前圖像，「汽車」類的累積權重 cw_1 的

當前值可以為 70，「卡車」類的累積權重 cw_2 的當前值可以為 25，「摩托車」類的累積權重 cw_3 的當前值可以為 5。使用該等示例性值和方程式 (3)，目標物件是「汽車」的概度性被決定為 $\frac{70}{70+25+5}$ (對應於值 0.7)。類似地，目標物件是「卡車」的概度性 L_c 被決定為 $\frac{25}{70+25+5}$ (對應於值 0.25)，並且目標物件是「摩托車」的概度性 L_c 被決定為 $\frac{5}{70+25+5}$ (對應於值 0.05)。

【0079】 對於給定的圖像或圖像群組，目標物件的類的最佳估計由在目標物件偵測的整個追蹤歷史中累積的概度性最高的類提供。例如，繼續上文的實例，車輛的類的最佳估計可以是「汽車」類，因為在「汽車」、「卡車」和「摩托車」類中，「汽車」類在目標車輛的追蹤歷史中具有最高概度性。由類概度估計引擎 206 估計的物件類可用於界定目標物件的尺寸 (例如，長度、寬度及 / 或高度) 的上限和下限。例如，「汽車」類的長度的上限和下限可以分別基於所有已知汽車的最小長度和最大長度。如下文更詳細描述的，當估計目標物件的尺寸時，尺寸估計引擎 214 可以使用該上限和下限。

【0080】 基於地圖的尺寸估計引擎 208 可以執行相機地圖融合，以決定在圖像中偵測到的目標物件的估計三維 (3D) 位置 (及 / 或朝向) 和估計尺寸 (例如，一或多個維度)。基於地圖的尺寸估計引擎 208 可以使用從物件偵測引擎 204 輸出的點狀地圖 207 和邊界框 (或其他邊界區域) 來決定與邊界框相關聯的目標物件的估計 3D 位置和尺寸。在出

於說明目的而使用的自主車輛的一個實例中，物件偵測引擎 204 可以向基於地圖的尺寸估計引擎 208 輸出圖像中與 2D 軸對準的邊界框，其中邊界框與圖像中偵測到的車輛相鄰。在圖 2 的圖像 205 中圖示與 2D 軸對準的邊界框 209 的實例，該邊界框與車輛相鄰。軸對準意味著邊界框的邊緣平行於圖像的邊緣（例如，如圖 5B 所示，邊界框 550 的邊緣平行於圖像 510 的邊緣）。在一些情況下，物件偵測引擎 204 亦可以輸出被偵測到的車輛的類（例如，「汽車」、「卡車」、「摩托車」等）。

【0081】 由於來自物件偵測的 2D 邊界框是車輛的投影，因此邊界框不提供物件的 3D 維度的直接量測。基於地圖的尺寸估計引擎 208 可以執行基於地圖的逆透視映射（IPM）技術，以決定在圖像中被偵測到的目標物件的估計 3D 位置（及 / 或朝向）和估計尺寸。例如，基於地圖的 IPM 技術可用於決定目標物件的 6 自由度（6-DOF）姿態，包括 3D 位置和 3D 朝向。例如，6-DOF 姿態可以包括 3D 旋轉向量（包括沿橫軸俯仰、沿縱軸滾動和沿法向軸偏航的角度）和 3D 平移向量（包括沿水平（x）方向、垂直（y）方向和深度（z）方向的平移）。相對於相機的俯仰、滾動和偏航可以被概念化為偏航是相機相對於地面的水平旋轉（例如，相對於水平軸從左向右），俯仰是相機相對於地面的垂直旋轉（例如，相對於水平軸上下），而滾動是相機相對於地平線的並排旋轉（例如，相對於水平軸並排）。可以使用點狀地圖 207 從基於地圖的位置關聯來決定 3D 朝向，並且

可以使用下文描述的基於地圖的IPM技術來決定3D位置。基於地圖的IPM技術可以將來自點狀地圖207的資訊與2D邊界框相結合，以獲取所偵測到的車輛的3D位置（及/或朝向）和尺寸估計。目標物件可被建模為立方體或3D朝向的邊界框。用於擷取目標物件的圖像的相機在現實世界中的位置（或姿態）亦是已知的。在一些情況下，使用點狀地圖207、被偵測到的目標物件的2D邊界框以及現實世界中的相機位置，可以決定表示物件的3D邊界框的位置和物件的尺寸的估計。

【0082】 點狀地圖207可包括與3D空間中的一或多個參考位置對應的複數個地圖點。在某些情況下，點狀地圖207可以被稱為高清晰度（HD）地圖。在使用自主車輛作為物件的說明性實例的一個實例中，點狀地圖207的點限定與道路相關的靜止實體參考位置，例如道路車道及/或其他資料。例如，點狀地圖207可以將道路上的車道表示為相連的點集。兩個地圖點之間限定線段，其中多個線段限定車道的不同線（例如，車道的邊界線和中心線）。線段可以組成使用地圖點限定的分段線性曲線。例如，連接的點集（或線段）可以表示道路上車道的中心線和邊界線，此舉允許自主車輛決定其在道路上的位置以及目標物件在道路上的位置。在某些情況下，可以為世界上的不同區域維護不同的點狀地圖（例如，紐約市的點狀地圖，三藩市的點狀地圖，新奧爾良的點狀地圖，等等）。在一些實例中，不同的點狀地圖可以被包括在單獨的資料檔案中（例如，

Geo-JavaScript Object Notation (GeoJSON) 檔案、ShapeFiles、comma-separated values (CSV) 檔案及 / 或其他檔案)。

【0083】 基於地圖的尺寸估計引擎 208 可以執行基於地圖的位置關聯函數，以將所偵測到的物件的邊界框（或其他類型的邊界區域）與點狀地圖 207 的地圖點（亦稱為地圖路標點）相關聯。例如，基於地圖的尺寸估計引擎 208 可以將來自點狀地圖 207 的複數個地圖點子集投影到從中偵測到目標物件的圖像上。使用基於地圖的位置關聯函數，可以將所偵測到的目標物件的邊界框與來自投影的地圖點子集中的地圖點相關聯。下文提供了使用車輛作為說明性實例來描述基於地圖的位置關聯函數的細節。一旦所偵測到的目標物件的邊界框與特定地圖點相關聯，該邊界框、地圖點以及與地圖點相關聯的車道的已知朝向可用於決定目標物件的估計 3D 位置及 / 或朝向（例如 6-DoF 姿態）和估計尺寸。例如，給定物件的寬度（例如，基於物件的給定分類或類別）和目標物件的朝向（例如，基於車道朝向），可以決定目標物件的長度和高度連同表示目標物件的 3D 邊界框的拐角的位置。

【0084】 如前述，對於自主車輛，點狀地圖 207 可以包括表示現實世界中道路上的車道（作為「參考位置」的實例）的複數個點，其中兩個地圖點之間限定線段。由追蹤車輛實現的系統 200 的基於地圖的尺寸估計引擎 208 可以使用點狀地圖 207 偵測並追蹤追蹤車輛周圍的目標車輛。追蹤

車輛中包括的一或多個相機 202 中的相機可以擷取目標車輛的圖像，並且可以使用物件偵測引擎 204 在擷取的圖像中偵測目標車輛。基於地圖的尺寸估計引擎 208 可以將來自點狀地圖 207 的點的子集投影到相機圖像，該等點的子集位於相機的視野內。隨後，基於地圖的尺寸估計引擎 208 可以應用基於地圖的位置關聯函數，以將點的子集中的點與用於辨識在圖像中被偵測到的目標車輛的邊界框相關聯。當對於自主駕駛執行時，基於地圖的位置關聯函數可被稱為基於地圖的車道車輛關聯 (Map LVA)，其可用於將所偵測到的目標車輛的每個 2D 邊界框與來自點狀地圖 207 的被投影到圖像上的點的子集中的點相關聯。

【0085】 基於地圖的車道車輛關聯 (例如，Map LVA) 可以幫助清除來自其他車道的偵測，並且可以幫助關聯引擎 (未圖示) 以防止交叉關聯。關聯引擎負責將新的一組偵測分配 (或關聯) 到現有的一組軌跡。例如，如下所述，關聯引擎可用於將來自物件偵測引擎 204 的基於相機的物件偵測與來自系統 200 的其他元件 (例如，基於雷達的尺寸估計引擎 212 及 / 或基於雷達的尺寸估計引擎 212) 的估計相關聯。軌跡可以被保持用於每個目標物件 (例如，每個目標車輛或其他物件)。例如，在多目標追蹤場景中，單獨的估計 (或軌跡) 被保持用於每個目標物件。在一些情況下，一或多個觀測值與現有軌跡不相關聯 (稱為無關聯事件)，在此種情況下，一或多個觀測值可以與新軌跡相關聯。例如，若新的目標物件 (例如，目標車輛) 進入

由追蹤物件（例如，追蹤車輛）擷取的場景，則可能發生此種無關聯事件。交叉關聯是關聯引擎的故障事件之一，其中從車輛 A 發出的觀測值被錯誤地關聯到被追蹤的車輛 B。清理偵測 / 觀測有助於防止交叉關聯。在一個說明性實例中，可以經由讓關聯引擎不考慮（經由基於地圖的位置關聯函數，例如 `Map LVA`）與包含被追蹤車輛的車道不同的車道相關聯的偵測，來完成清理。在一些實例中，關聯引擎可以忽略來自點狀地圖 207 外部的偵測（例如，來自護欄、車道障礙物及 / 或除點狀地圖 207 中的位置以外的其他位置）。例如，關聯引擎可以刪減（或移除）基於地圖的車道車輛關聯（例如，`Map LVA`）已經失敗的偵測。在由系統 200（例如，由基於地圖的尺寸估計引擎 208、尺寸估計引擎 214 等）執行的後續過程中，可以不考慮刪減或移除的偵測。

【0086】 由關聯引擎執行的關聯可以使用任何合適的技術來執行，例如使用加權二分圖匹配的形式。在一個實例中，來自單個成像雷達訊框的偵測可以與追蹤物件（例如，追蹤車輛或其他物件）當前保持的軌跡相關聯。例如，二分圖可以包括在二分圖的一部分上形成節點的軌跡，以及在二分圖的另一部分中來自一個感測器訊框的偵測。每個軌跡節點可以經由邊緣連接到每個偵測節點，並且可以為每個邊緣分配權重，其中權重與軌跡的當前估計和偵測之間的不一致程度成比例。在一個說明性實例中，在成像雷達訊框的情況下，權重可以與軌跡的平均位置的當前估計和

偵測到的物件（例如，偵測到的目標車輛）的位置之間的平方距離成比例。經由加權二分圖上的最小權重匹配，提供偵測與軌跡的最佳關聯。

【0087】 如前述，基於地圖的尺寸估計引擎 208 可以將來自點狀地圖 207 的點的子集投影到所擷取的圖像。在一些實例中，來自點狀地圖 207 的點的子集可以包括在相機視野內並且在相機的選擇範圍內（例如，在 50 米、75 米、100 米、150 米或其他距離內）的所有地圖點。選擇範圍在本文中可以被稱為點子集選擇範圍。圖 3 是圖示從與環境中的參考位置相對應的點地圖中決定點的子集的實例的圖。圖 3 中的實例從俯視圖（或「鳥瞰圖」）中圖示高速公路的三個車道，包括左車道 322、中間車道 324 和右車道 326。每個車道用一條中心線和兩條邊界線圖示，其中中間車道 324 與左車道 322 和右車道 326 共享邊界線。在中間車道 324 中顯示了車輛 320。如本文所述，車輛上的一或多個相機可以擷取車輛周圍環境的圖像。圖 3 中圖示點子集選擇範圍 328，其圖示相對於相機的範圍，從該範圍可以從點狀地圖 207 選擇點的子集以投影到由相機擷取的圖像上。如前述，點狀地圖 207 可以包括代表車道線的點（或路標點）。例如，每條線可以由多個點限定。限定車道 322、324 和 326 的中心線和邊界線的在點子集選擇範圍 328 內的所有點可以被選擇以投影到由車輛 320 上的相機擷取的圖像上。在一些實施方式中，可以基於給定的應用（例如，

自主駕駛、機器人技術、視訊分析或其他應用) 調整該範圍以滿足不同的需求。

【0088】 圖 4 是圖示將地圖點的子集從點地圖投影到由相機 430 擷取的圖像的圖像平面的實例的圖。相機 430 可以是車輛或其他設備或物件的一部分。點狀地圖 207 具有與其相關聯的方向，其中點狀地圖 207 中的點按順序排列。在某些情況下，序列的方向與追蹤物件（例如，追蹤車輛或其他物件）的移動方向相同。對於當前正在觀測或處理的點狀地圖 207 的序列中的給定點（被稱為「當前地圖點」），將在點序列中排列的後續點稱為點狀地圖 207 中的「下一地圖點」。在圖 4 中，圖示具有當前地圖點 435 和下一地圖點 436 的基本路線 434。基本路線 434 是道路上的特定車道的中心線。為了執行映射，基於地圖的尺寸估計引擎 208 可以在將來自點狀地圖 207 的點的投影投影到圖像上之後，檢查當前地圖點 435 和下一地圖點 436 是否在圖像內。若不在，則基於地圖的尺寸估計引擎 208 可以使用相機的視野（或視域）來查看從相機延伸的（相機視野的）底部平面 432 與基本路線 434 相交的位置，並使用交點（圖示為交點 437）作為圖像上的點。

【0089】 圖 5 A 是圖示表示來自追蹤物件的前置相機的圖像平面的圖像 505 的實例的圖。圖 5 A 所示的圖像 505 由追蹤車輛的前置相機擷取。道路上的多條車道中心線被圖示，包括車道中心線 541 和車道中心線 544 等。投影到圖像 505 上的地圖點（其限定車道中心線）圖示為白點。如前

述，投影到圖像 505 上的車道中心線的地圖點可以包括在前置相機的視野內並且在點子集選擇範圍內的車道中心線的所有地圖點。例如，地圖點 542、地圖點 543，以及地圖點 542 和 543 之間的其他地圖點限定車道中心線 541。類似地，地圖點 545、地圖點 546 以及地圖點 545 和 546 之間的其他地圖點限定車道中心線 544。所偵測到的目標物件的邊界框 540 被圖示為包括沿車道中心線 541 和沿車道中心線 544 的某些地圖點。

【0090】 圖 5B 是圖示表示來自追蹤物件的側向相機的圖像平面的圖像 510 的實例的圖。圖像 510 由追蹤車輛的側向相機擷取。與圖 5A 類似，圖 5B 中圖示道路上的車道中心線，包括車道中心線 551、車道中心線 552 和車道中心線 553。限定車道中心線並投影到圖像 510 上的地圖點被圖示為白點。投影到圖像 510 上的地圖點包括車道中心線的地圖點，該等地圖點位於側向相機的視野內並且在點子集選擇範圍內。所偵測到的目標物件的邊界框 550 被圖示為包括沿車道中心線 551、552 和 553 的某些地圖點。

【0091】 可以使用相機的姿態（已知）並使用校準資料來執行圖像上的點投影。如前述，相機的姿態是已知的。例如，相機的姿態可以註冊為具有每個偵測（或者在某些情況下具有偵測的某個子集）的中繼資料，並且可以在需要時從中繼資料中查詢。校準資料可以包括一組變換（例如，表示為矩陣，例如 3×4 矩陣或具有其他尺寸的矩陣），其將 3D 點從一個參考訊框變換到另一個參考訊框（例如，使

用針孔相機模型)。在某些情況下，校準資料亦可以包括固有的相機參數，例如焦距和主點。固有的相機參數可用於計算相機座標系中給定3D點的圖元位置。校準資料可以在系統200之外單獨保持。

【0092】 使用相機姿態和校準資料的點到圖像投影可以使用任何合適的技術來執行。在一個說明性實例中，可以執行兩步過程，其中第一步包括將點變換到相機訊框（對應於相機姿態），第二步包括將點投影到圖像上（在圖元域中）。例如，給定世界訊框中的待投影點，表示為 P_w ，可以使用 T_{cw} （將點從世界訊框變換到相機訊框的 3×4 矩陣）將點 P_w 變換到相機訊框： $P_c = T_{cw} * P_w$ 。為了將變換後的點 P_c 投影到圖像（圖元域）上，被投影的點可以由 $Proj_c = K P_c^n$ 提供，其中 K 是相機的固有矩陣，其來自校準資料，並且 P_c^n 是正規化座標中的 P_c 。例如，以下內容可用於定義正規化點 - $P_c^n = \begin{bmatrix} \frac{P_{c,x}}{P_{c,z}}, \frac{P_{c,y}}{P_{c,z}}, 1 \end{bmatrix}$ ，其中 $P_c = [P_{c,x}, P_{c,y}, P_{c,z}]$ 來自以上步驟1。

【0093】 一旦在用於偵測目標物件的相機圖像中投影地圖點，就可以選擇與目標物件的邊界框關聯的地圖點。資訊可以被提供用於由基於地圖的位置關聯函數（例如，MAPLVA）決定的每個關聯，包括用於關聯的地圖中心線（中心線路標點）的地圖點（來自點狀地圖207）、沿著點狀地圖207的方向的距離（例如，正平行距離，包括目標物件的邊界框上的點與一或多個投影點之間在車道朝向的方向（或點狀地圖中表示的其他位置）上的平行距離，如下文參照圖6所述），以及從地圖中心線到被偵測物件的邊界

框底部邊緣中心的垂直距離。可以使用最近的地圖點之前的地圖點進行關聯，從而導致正平行距離。在某些情況下，僅可以選擇與被偵測到的物件具有正距離的地圖點，以與被偵測到的物件的邊界框相關聯。例如，即使距離被偵測到的物件的邊界框上的點（例如，邊界框底部邊緣上的中點）最近的地圖點具有負平行距離，亦可能不會選擇該點進行關聯，並且距離邊界框上的點最近的地圖點之前的地圖點可以被選擇以與邊界框關聯（導致平行正距離）。可以基於包括系統 200 的追蹤物件上的相機的位置來執行一組不同的操作。例如，與追蹤車輛側面的側面相機相比，可以對追蹤車輛的前後車輛上的前後相機執行一組不同的操作。基於對圖像中物件的偵測而產生的邊界框可以與投影到圖像上的地圖點一起使用。例如，邊界框與地圖點或參考位置（例如，道路上的車道線）上的其他點的交點可用於選擇地圖點或其他點以與目標物件的邊界框關聯。在另一個實例中，從邊界框上的點到各種地圖點的距離可以用於選擇地圖點以與邊界框相關聯。下文出於說明目的提供了使用自主駕駛應用的實例。

【0094】 在一個說明性實例中，當在由自主車輛的自主駕駛系統的側面相機接收的圖像中偵測到目標車輛時，可以基於邊界框上的點與道路的車道上的線的各種地圖點（例如，車道中心線的地圖點）之間的距離來選擇用於與目標車輛的邊界框關聯的地圖點。例如，邊界框的底部邊緣可以由基於地圖的尺寸估計引擎 208 獲取。基於地圖的尺寸

估計引擎 208 可以計算底部邊緣的中點，並且可以比較從中點到一或多個車道中心線上的地圖點的距離。例如，與邊界框的底部邊緣具有最短距離的車道中心線（因此亦是最接近的車道中心線）可以決定為與目標車輛相關聯的車道（對應於車輛所在的車道）。邊界框內和最近的中心線上的所有地圖點可以被選擇作為候選關聯點。底部邊緣的中點與候選關聯點之間的距離可以被決定。例如，可以計算出中點與投影點之間的平行（沿車道朝向的方向）距離（例如，作為圖像平面中的 2D 距離）。候選關聯點可被過濾掉，以僅包括與中點的平行距離大於 0（正平行距離）的地圖點，以及中點位於當前地圖點和追蹤車輛的移動方向上在當前地圖點之後出現的後續地圖點之間的地圖點。在一些實例中，若在過濾之後剩餘多個候選關聯點，則基於地圖的尺寸估計引擎 208 可以選擇屬於在邊界框的頂部和底部邊緣之間通過的中心線的點。所選地圖點提供了車輛所在的車道的線段（其中線段在兩個地圖點之間）。

【0095】 圖 6 是圖示將在圖像中偵測到的目標物件的邊界框 631 與來自投影到圖像上的地圖點中的地圖點 635 相關聯的實例的圖。圖 6 中的實例圖示當圖像被追蹤車輛的側面相機 630 擷取時的情況。在圖 6 的實例中，追蹤車輛的移動方向是從左到右，如箭頭所示。圖示兩條車道中心線，包括車道中心線 633 和車道中心線 634。從側向相機 630 投影到道路平面的射線的交點 638（對應於邊界框 631 的底邊緣的中點），以及交點 638 和車道中心線 634 之間的垂直距離

被圖示。地圖點 635 和交點 638 之間的平行距離 639 被圖示。由於平行距離 639 大於 0 並且交點 638 位於地圖點 635 和在追蹤車輛的移動方向上在地圖點 635 之後出現的後續地圖點 636 之間，所以選擇地圖點 635 與邊界框 631 相關聯。

【0096】 圖 7 是圖示將在圖像中偵測到的目標物件的邊界框 750 與來自投影到圖像 710 上的地圖點中的地圖點 754 相關聯的實例的圖像 710。圖像 710 由追蹤車輛的側面相機擷取。圖示道路上車道的中心線，包括車道中心線 751、車道中心線 752 和車道中心線 753。限定車道中心線並被投影到圖像 710 上的地圖點被圖示為白點。圖示偵測到的目標物件的邊界框 750。基於地圖的尺寸估計引擎 208 可以獲取邊界框 750 的底部邊緣，並且可以決定底部邊緣的中點 756。基於地圖的尺寸估計引擎 208 可以比較從中點 756 到車道中心線 751、752 和 753 上的地圖點的距離。基於地圖的尺寸估計引擎 208 可以選擇地圖點 754 作為將與邊界框 750 相關聯的點，因為地圖點 754 和中點 756 之間的平行距離 739 大於 0，並且中點 756 在地圖點 754 和在追蹤車輛的移動方向上（在圖 7 中從左到右）在地圖點 754 之後出現的後續地圖點之間。

【0097】 在另一個說明性實例中，當在由自主駕駛系統的前置或後置相機擷取的圖像中偵測到目標車輛時，可以基於邊界框與車道中心線的交點，來選擇用於與目標車輛的邊界框相關聯的地圖點。例如，基於地圖的尺寸估計引擎

208 可以獲取目標車輛的邊界框，並且可以決定邊界框的底部邊緣。可以經由辨識出底部邊緣與來自被投影到圖像平面上的點狀地圖 207 中的點的任何車道中心線相交的位置，來決定用於關聯的地圖點。若上述技術未能辨識與邊界框相關聯的點（例如，底部邊緣不與來自投影點的車道中心線相交），則底部邊緣可以被延伸，使得其與圖像中的所有車道中心線相交。邊界的底部邊緣被選擇為要延伸的邊緣，因為從前向或後向的圖像中產生的邊界框的底部邊緣將位於路面上（由於車輛的底部在道路上，並且由於邊界框與車輛接界或顯示車輛的輪廓）。基於地圖的尺寸估計引擎 208 隨後可以獲取與車道中心線相關聯的車道邊界點（限定車道邊界線的點，如圖 3 所示）。經由決定包圍邊界框底部邊緣中點的兩個車道邊界點，可以獲取車輛所處的車道。兩個車道邊界點提供了車輛所在的車道的線段。車道的中心線上決定車輛位於的點可以被決定為將與邊界框相關聯的地圖點。

【0098】 圖 8 是圖示將圖像中偵測到的目標物件的邊界框 850 與來自圖像 810 的點 856 相關聯的實例的圖像 810。圖 8 中的圖像 810 由追蹤車輛的前置相機擷取。圖示道路上車道的中心線，包括車道中心線 851、車道中心線 852 和車道中心線 853。限定車道中心線並被投影到圖像 710 上的地圖點被圖示為白點。圖示偵測到的目標物件的邊界框 850。基於地圖的尺寸估計引擎 208 可以獲取邊界框 850 的底部邊緣，並且可以決定底部邊緣與車道中心線 851 相交。點

856 是邊界框 850 的底部邊緣與車道中心線 851 相交的交點。

【0099】 如前述，基於地圖的位置關聯函數（例如，MapLVA）的結果是來自點狀地圖 207 的與目標物件（例如，目標車輛或其他目標物件）的 2D 邊界框相關聯的關聯地圖點。關聯地圖點限定目標物件的朝向。例如，使用自主車輛作為說明性實例，相關聯的地圖點指示目標車輛所在的車道，並且可以假設目標車輛的朝向平行於車道（車輛的縱向或前方指向平行於車道的方向）。由點狀地圖 207 限定的各種車道的朝向是已知的，因此目標車輛的朝向可以被假設為具有與被決定在其中的車道相同的朝向。

【0100】 由物件偵測引擎 204 分配給目標物件的偵測的類或類別標籤亦提供了目標物件的維度的有力的先例。在一個說明性實例中，車輛的「汽車」的類或類別標籤提供了車輛的一或多個可能維度的實例。例如，給定決定的類，可以假定車輛的寬度。此種假定是合理的，因為在同一類別或類中的車輛的寬度幾乎沒有變化，而長度（有時是高度）可以在更大程度上變化。

【0101】 給定來自物件偵測引擎 204 的 2D 邊界框、從基於地圖的位置關聯函數（其定義目標車輛的方向）決定的關聯地圖點，以及基於所決定的類的目標車輛（或其他目標物件）的假定寬度，基於地圖的尺寸估計引擎 208 可以執行基於地圖的逆透視映射（IPM）技術，來決定與來自物件偵測引擎 204 的 2D 邊界框相對應的物件的 3D 朝向邊界

框 (OBB)。3D OBB 表示目標車輛 (或其他目標物件) 的尺寸。

【0102】 例如，假設目標物件 (例如，目標車輛或其他物件) 的 2D 邊界框 (2D BB) 與 3D OBB 的投影相切 (指示 3D OBB 的拐角 (例如，四個或更多個拐角) 的投影位於 2D BB 的邊緣上)，目標物件的朝向 (例如，目標車輛的朝向) 平行於點狀地圖 207 中的車道 (或其他位置) 的朝向，點狀地圖 207 包含來自基於地圖的位置關聯函數的關聯地圖點 (指示目標物件和與地圖點關聯的車道對準或朝向)，並且目標物件的底部 (例如，目標車輛的底部) 與關聯地圖點處的切面對準 (指示 3D OBB 的底面位於穿過關聯地圖點的平面上，該平面與關聯地圖點處的路面相切；相切，因為道面在關聯地圖點處可以是彎曲的，並且可以假設 3D OBB 的底面是平坦的)，可以得到一組線性方程式，並且可以針對需要估計的參數進行求解。兩個說明性的場景 (稱為場景 1 和場景 2) 被提供以指示何者參數需要被估計。可以執行方位角檢查來決定何者場景適用。方位角檢查的實例如下：若 $Azm_{左} Azm_{右} < 0$ ，場景 1 適用，否則，場景 2 適用，其中 $Azm_{左} = (\hat{h} \times d_l) \cdot \hat{n}$ 並且 $Azm_{右} = (\hat{h} \times d_r) \cdot \hat{n}$ 。向量 d_l 是與地圖平面和 camera-bb2d-left-edge 平面 (其是 3D 中包含相機中心以及圖像中 2D 偵測的左邊緣的平面) 之間的交線相關聯的方向向量，並且向量 d_r 是與地圖平面和 camera-bb2d-right-edge 平面 (其是 3D 中包含

相機中心以及圖像中 2 D 偵測的右邊緣的平面) 之間的交線相關聯的方向向量。

【0103】 在場景 1 中，基於地圖的尺寸估計引擎 2 0 8 可以經由求解以下線性方程組來估計 3 D O B B 的位置、高度和長度：

$$\begin{bmatrix} d_b.n_2 & \hat{h}.n_2 & \hat{n}.n_2 \\ d_b.n_1 & 0 & \hat{n}.n_1 \\ d_b.n_t & \hat{h}.n_t & \hat{n}.n_t \end{bmatrix} \begin{bmatrix} \lambda \\ l \\ h \end{bmatrix} = \begin{bmatrix} -(p_b - p_2).n_2 \\ -(p_b - p_1 + w\hat{b}).n_1 \\ -(p_b - p_t + w\hat{b}).n_t \end{bmatrix}$$

方程式 (4) 。

【0104】 在方程式 (4) 中， p_1 、 p_2 、 n_1 、 n_2 是由進一步檢查左方位角 $Azm_{\text{左}}$ 來決定的，其可以總結如下：

若 $Azm_{\text{左}} < 0$ ，則 $p_1 = p_l$ 、 $n_1 = n_l$ 、 $p_2 = p_r$ 、 $n_2 = n_r$ 、 $\hat{b} = \hat{n} \times \hat{h}$ ，否則， $p_1 = p_r$ 、 $n_1 = n_r$ 、 $p_2 = p_l$ 、 $n_2 = n_l$ 、 $\hat{b} = -\hat{n} \times \hat{h}$ 。

方程式 (5)

【0105】 在方程式 (4) 和 (5) 中， p_l 和 n_l 是定義穿過相機中心和目標物件的 2 D B B 的左邊緣的 3 D 平面的點 (p_l) 和法向量 (n_l) 。類似地， p_r 和 n_r 是定義穿過相機中心和目標物件的 2 D B B 的右邊緣的平面的點 (p_r) 和法向量 (n_r) 。項 p_t 和 n_t 是定義穿過相機中心和 2 D B B 的頂部邊緣的平面的點 (p_t) 和法向量 (n_t) 。項 p_b 和 d_b 是任何點 - 方向 - 向量對 (p_b 是點， d_b 是方向向量) ，其定義了 `camera-bb2d-bottom-edge` 平面 (其是 3 D 中包含圖像中的相機中心以及 2 D 偵測的底部邊緣的平面) 和來自基於地圖的位置關聯函數的關聯地圖點處的切面 (例如，在關聯地圖點處與路面相切的平面) 的交線。項 \hat{h} 表示沿著前進

方向的單位向量，項 \hat{n} 表示沿關聯地圖點處的道路法線的單位向量。向量 \hat{b} 是沿 3D OBB 的橫向維度（寬度）（例如，在道路平面上）的 3D 向量，並且與前進向量 \hat{h} 正交，法向量 \hat{n} 沿著 3D OBB 的高度，前進向量 \hat{h} 沿著 3D OBB 的長度。上述線性方程式（4）中的未知數包括：標量 λ ，其是此種標量，使得 $p_b + \lambda d_b$ 定義最接近追蹤物件（例如，追蹤車輛）的目標物件的 3D OBB（例如，偵測到的目標車輛的 3D OBB）的拐角（或頂點）；長度 l ；和目標物件的高度 h 。

【0106】 在場景 2 中，基於地圖的尺寸估計引擎 208 可以經由求解以下線性方程組來估計 3D OBB 的位置、高度和寬度：

$$\begin{bmatrix} d_b \cdot n_r & 0 & \hat{n} \cdot n_r \\ d_b \cdot n_l & \hat{b} \cdot n_l & \hat{n} \cdot n_l \\ d_b \cdot n_t & \hat{b} \cdot n_t & \hat{n} \cdot n_t \end{bmatrix} \begin{bmatrix} \lambda \\ w \\ h \end{bmatrix} = \begin{bmatrix} -(p_b - p_r) \cdot n_r \\ -(p_b - p_l) \cdot n_l \\ -(p_b - p_t + l\hat{h}) \cdot n_t \end{bmatrix}$$

方程式（6）。

【0107】 在方程式（6）中， p_l 和 n_l 是定義穿過相機中心和目標物件的 2D BB 的左邊緣的平面的點（ p_l ）和法向量（ n_l ）。類似地， p_r 和 n_r 是定義穿過相機中心和目標物件的 2D BB 的右邊緣的平面的點（ p_r ）和法向量（ n_r ）。項 p_t 和 n_t 是定義穿過相機中心和 2D BB 的頂部邊緣的平面的點（ p_t ）和法向量（ n_t ）。項 p_b 和 d_b 是任何點 - 方向 - 向量對（ p_b 是點， d_b 是方向向量），其定義了 camera-bb2d-bottom-edge 平面和（來自基於地圖的位置關聯函數）的關聯地圖點處的切面的交線。項 \hat{h} 表示沿著前進方向的單位向量，項 \hat{n} 表示沿關聯地圖點處的道路法

線的單位向量。上述線性方程式（4）中的未知數包括：標量 λ ，其是此種標量，使得 $p_b + \lambda d_b$ 定義最接近追蹤物件（例如，追蹤車輛）的目標物件的3D OBB的拐角（或頂點）、目標物件的寬度 w 和高度 h 。

【0108】 圖9A（場景1的代表）是圖示將目標物件的3D朝向邊界框（3D OBB）960的拐角（或頂點）與目標物件的2D邊界框（2D BB）950的邊緣相關聯的幾何約束的實例的圖。方程式（4）和（5）適用於圖9A所示的場景。如圖所示，點和法向量對 p_r 、 n_r 定義了穿過相機中心和目標物件的2D BB 950的右邊緣的平面，因此3D OBB 960的最右邊的拐角或頂點（如圖9A所示）位於2D BB 950的右邊緣上。點和法向量對 p_l 、 n_l 定義了穿過相機中心和2D BB 950的左邊緣的平面，指示3D OBB 960的最左邊的拐角位於2D BB 950的左邊緣。點和法向量對 p_t 、 n_t 定義了穿過相機中心和2D BB 950的頂部邊緣的平面，表示3D OBB 960的最上面的拐角位於2D BB 950的頂部邊緣。一旦使用方程式（4）決定未知標量 λ ，公式 $p_b + \lambda d_b$ 定義了3D OBB 960的最接近追蹤物件（例如，追蹤車輛）的拐角的位置，該拐角是圖9A中的最底部的拐角。最接近追蹤物件的3D OBB 960的拐角的位置提供了3D OBB 960的3D位置（如前述，6 DoF姿態的3D朝向由基於地圖的位置關聯提供）。3D OBB 960的最底部的拐角的位置固有限定了與2D BB 950的其他邊緣相鄰的其他三個拐角的位置。

【0109】 圖 9 B 是表示場景 2 的圖，圖示將目標物件的 3 D 朝向邊界框（3 D O B B）9 8 0 的拐角與目標物件的 2 D 邊界框 9 7 0 的邊緣相關聯的幾何約束的另一實例。方程式（6）適用於圖 9 B 中所示的場景，並且圖 9 B 中所示的項與上文關於圖 9 A 描述的彼等項相似。

【0110】 在一些實施方式中，可以進行可行性評估以決定合適的場景。一旦基於地圖的尺寸估計引擎 2 0 8 決定了適當的場景，基於地圖的尺寸估計引擎 2 0 8 就可以應用對應於所決定的場景的線性方程。經由求解所決定的線性方程組，基於地圖的尺寸估計引擎 2 0 8 可以決定三個未知值的估計，包括標量 λ （其提供目標車輛或其他目標物件的位置）、目標車輛或其他目標物件的長度 l ，以及目標車輛或其他目標物件的高度 h 。

【0111】 基於雷達的尺寸估計引擎 2 1 2 可以從一或多個雷達 2 1 0 獲取雷達量測值 2 1 1，並且可以使用雷達量測值 2 1 1 來估計目標物件的尺寸。一或多個雷達 2 1 0 中的雷達可以包括具有射頻（R F）源的設備或系統，該 R F 源發送 R F 信號（例如，高頻電磁波的脈衝），該等 R F 信號可以從目標物件反射回 R F 信號源。被反射的 R F 信號可以由雷達設備或系統的 R F 接收器接收。被反射的 R F 信號可用於決定目標物件的尺寸。一或多個雷達 2 1 0 可以包括位於追蹤物件上不同位置的多個雷達。例如，使用自主追蹤車輛作為說明性實例，追蹤車輛可以具有位於車輛前部、拐角、側面及 / 或後部的一或多個雷達。由追蹤物件上的所有感測器接收的

被反射的 R F 信號可以被基於雷達的尺寸估計引擎 2 1 2 評估和使用，以估計從其反射信號的目標物件的尺寸（例如，長度及 / 或其他維度）。

【0112】 例如，在週期性時刻，從目標物件反射的雷達信號可以在最近的雷達訊框中從位於追蹤物件上的多個雷達感測器獲取。每個雷達可以以固定頻率（例如，每 5 0 毫秒（ m s ）一次或其他頻率）量測其周圍環境。此處使用的術語頻率指的是量測速率（連續量測例子或訊框之間時間差的倒數。雷達在單個量測例子中獲取的一組量測值是雷達訊框。來自每個雷達的量測值會產生不同的雷達訊框。可以對來自多個雷達感測器的雷達信號的雷達量測值進行聯合處理，並且可以使用沿著目標物件的朝向方向的雷達量測點的最大縱向擴展來決定目標物件的尺寸（例如，長度或其他尺寸或維度）的暫態估計。物件（如車輛或其他移動物件）的縱向是物件向前移動的方向。在一個實例中，如下文參考圖 1 0 和圖 1 1 所圖示和解釋的，基於雷達的尺寸估計引擎 2 1 2 可以使用兩個末端之間間隔作為追蹤物件的長度的估計。在一些實施方式中，暫態尺寸（例如，長度）估計被寫入資料結構，該資料結構可以追蹤獲取的用於目標物件的前 K 個最大估計（其中 K 是固定的數值）。在此種情況下，在任何時間點，來自基於雷達的尺寸估計的尺寸（例如，長度）的最佳估計是獲取的第 K 個最長的暫態估計。

【0113】 使用自主車輛作為說明性實例，在追蹤車輛上運行的基於雷達的尺寸估計引擎 2 1 2 可以根據點雷達偵測來估計目標車輛的長度。然而，由於可能無法從目標車輛的整個範圍獲取 R F 信號反射，因此經由點雷達偵測來估計目標車輛的長度可能具有挑戰性。為了最大限度地覆蓋來自車輛表面的點，可以聯合考慮追蹤車輛上所有雷達的 ping。例如，在週期性時刻，在最近的雷達訊框中從追蹤車輛上的所有雷達感測器獲取的從目標車輛反射的雷達信號被聯合考慮，並且沿著追蹤車輛的朝向方向（稱為縱向方向）的點的最大縱向擴展被用於得到車輛長度的暫態估計。最大縱向擴展表示物件兩個縱向末端之間（車輛前後之間）的間隔，從而提供了物件長度的估計值。例如，若所有的點皆投影到平行於縱向方向的線上，則投影中的極值點就是端點（該等極值點之間的距離提供了長度的估計值）。縱向方向在上文被定義為被追蹤物件的朝向方向。可以經由追蹤物件（例如，目標車輛）來知道該朝向。若朝向未知，則可以使用點狀地圖 2 0 7、使用點狀地圖 2 0 7 中定義的相對於物件的位置的朝向（例如，如前述，被追蹤車輛位置處的車道的朝向），來決定該朝向。

【0114】 圖 1 0 是圖示由各種雷達感測器接收的雷達量測的縱向分量（沿 y 軸）相對於時間（沿 x 軸）繪製的實例的曲線圖 1 0 0 0。雷達量測是相對於時間繪製的、由追蹤車輛前部、後部和拐角處的不同雷達接收到的來自目標車輛的 R F 信號反射。圖 1 0 中的量測資料跨越 7 秒的間隔，在該間

隔期間，追蹤車輛在道路上超過（或經過）目標車輛。黑點代表地面真實位置（表示為地面真值 $y - \text{minimum}(\text{gt_y_min})$ 和地面真值 $y - \text{maximum}(\text{gt_y_max})$ ），對應於目標車輛前後端的實際位置，並且被圖示以供參考。每種顏色的點代表不同雷達的量測。例如，紅點代表追蹤車輛前部遠端雷達（表示為 ESR ）的雷達量測。如圖所示，不同的雷達在不同的時間接收從目標物件反射的 RF 信號。例如，從大約 89 秒到 93 秒，追蹤車輛前部的雷達（紅點，表示為 ESR ）和右前角的雷達（深藍色點，表示為 SRR_FR ）為目標車輛提供雷達量測值，此情形是當追蹤車輛超過目標車輛時，目標車輛在追蹤車輛的前方和旁邊。從大約 91.5 秒到 97 秒，追蹤車輛右後角的雷達（亮深藍色點，表示為 SRR_RR ）為目標車輛提供雷達量測。從大約 94 秒到 97 秒，追蹤車輛後部的雷達（綠點，表示為 ESR_R ）為目標車輛提供雷達量測。

【0115】 從曲線圖 1000 中所示的量測可以看出，當與由黑點表示的地面真實位置進行比較時，雷達最初（從大約 89 秒到 93 秒）可靠地觀測目標車輛的後端，並且隨後（從大約 93 秒到 97 秒）可靠地觀測目標車輛的前端。亦可以看到，當目標車輛在追蹤車輛的前面時，從目標車輛的後面連貫獲取雷達 ping 。類似地，可以看到，當目標車輛在追蹤車輛的後面時，從目標車輛的前部連貫地獲取雷達 ping 。

【0116】 圖 11 是圖示使用圖 10 的曲線圖 1000 中所示的雷達量測決定的目標車輛的縱向末端的估計的實例的曲線圖

1100。術語「`max_in_win`」是沿縱向的時間訊窗中的最大值，術語「`min_in_win`」是沿縱向的時間訊窗中的最小值。縱向末端是指給定的時間點的最大或最小經度的量測（例如，給定訊框中最大或最小的縱向值）。紅點（最大縱向末端）代表目標車輛前端的估計值，綠點（最小縱向末端）代表目標車輛後端的估計值。在任何給定時刻，圖中兩個縱向末端（代表車輛的前部和後部）之間的間隔為基於雷達的尺寸估計引擎 212 提供目標車輛的長度的估計。

【0117】 可以經由批量估計來計算目標車輛縱向末端的估計值。如此處所使用的，批是在此種時間訊窗中的觀測結果，該時間訊窗被處理以獲取長度或其他尺寸的估計。例如，單批可以包括在某個時間訊窗（例如，50 毫秒的時間訊窗）內從所有感測器獲取的所有量測值的集合。當目標車輛在追蹤車輛後面時，目標車輛前端的估計是準確的，當目標車輛在追蹤車輛前面時，目標車輛後端的估計是準確的。當目標車輛縱向靠近追蹤車輛時，兩種估計皆是可靠的。

【0118】 在某些情況下，即使有多個雷達，點雷達量測亦可能是稀疏的（數量少且空間密度低）。由於點雷達偵測的稀疏性，車輛的末端可能無法被觀測到，並且暫態長度估計可能低估了車輛或其他物件的長度。如前述，暫態估計可以被寫入資料結構（例如，量測值陣列），該資料結構可以追蹤目標車輛曾經獲取的前 K 個最大估計，其中 K 是

固定的數值。在任何時間點，基於雷達的長度估計的長度（或其他維度）的最佳估計是獲取的第 K 個最大暫態估計。例如， $K = 1$ 將是最大估計， $K = 5$ 將是第 5 個最大估計。任何數值皆可以用作選擇的估計，如 $K = 3$ 。在一些實例中，基於雷達的長度估計所產生的估計的可靠性可以基於目標車輛相對於不同雷達的視野的位置來量化。例如，目標車輛離追蹤車輛（或自我車輛）越近，尺寸估計越可靠。

【0119】 儘管上述基於雷達的估計的實例描述了基於縱向量測（對應於物件的縱向方向）來決定物件的長度，但是可以使用物件的橫向量測（對應於物件的橫向方向，例如垂直於車輛或其他移動物件的前進方向的方向）執行類似的技術來決定物件的寬度。

【0120】 在一些實例中，追蹤物件（例如，追蹤車輛或其他物件）可以包括成像雷達感測器。雷達圖像可以具有任何合適的頻率，例如毫米頻帶或微波頻帶的頻率。圖 12 A 和圖 12 B 圖示雷達圖像的說明性實例。可用於位置和尺寸/維度估計的雷達圖像的實例包括 10 GHz 圖像、30 GHz 圖像、60 GHz 圖像、100 GHz 圖像、300 GHz 圖像及/或具有任何其他合適高頻的雷達圖像。雷達圖像可以是毫米波雷達圖像，其是具有從第一波長大小（例如，1 毫米）到第二波長大小（例如，10 毫米）的短波長及/或落入第一頻率（例如，30 GHz）和第二頻率（例如，300 GHz）之間的頻帶或頻譜範圍的雷達圖像。毫米波雷達圖像有時被稱為毫米頻帶、極高頻（EHF）或超高頻（VHF）。可

以替代地或附加地使用毫米頻帶之外的其他無線電頻率和波長，例如在300兆赫（MHz）和30 GHz之間的微波區域中的頻帶。在某些情況下，雷達圖像可以直接從雷達系統接收。在一些情況下，可以從系統200中包括的儲存設備或記憶體中取得雷達圖像，或者在包括系統200的物件或設備中包括雷達圖像。在一些情況下，可以從系統200外部或包括系統200的物件或設備外部的儲存設備或記憶體中取得雷達圖像。

【0121】 雷達系統可以包括天線陣列，每個天線包括接收器或與接收器耦合。在一些實施方式中，雷達系統可以具有能夠傳輸射頻（RF）信號的單個傳輸器，該RF信號從環境中的一或多個物件（例如，目標物件）反射。在此種實施方式中，天線陣列的天線和接收器接收來自傳輸器的反射RF信號，每個天線和接收器接收不同版本的反射信號。每個天線和接收器皆可以記錄接收到的反射信號的幅度和相位等資料。在其他實施方式中，天線陣列的每個天線可以包括傳輸器或者與傳輸器耦合，在此種情況下，為陣列之每一者天線提供接收器-傳輸器對。對於給定的接收器-傳輸器對，傳輸器可以傳輸RF信號，該RF信號從環境中的一或多個物件（例如，面部）反射，並且接收器可以接收反射的RF信號。

【0122】 在一些實例中，雷達系統可以被實現為追蹤物件（例如，追蹤車輛或其他追蹤物件）上的一或多個多千兆（multi-gigabit）無線電。例如，使用高頻帶（例如，

10 GHz、30 GHz、60 GHz、100 GHz、300 GHz 或其他合適的高頻)的多千兆技術(例如,多千兆WLAN技術)可以被實現用於許多計算設備(例如,行動設備、自主車輛等)中的無線通訊。多千兆無線電可以在雷達模式下操作,用於擷取附近物件反射的傳輸信號。在一些實施方式中,一或多個多千兆無線電可以用於產生雷達圖像。在一個說明性實例中,一或多個多千兆無線電可以包括一或多個60 GHz WLAN無線電。在此種實例中,多千兆無線電可以包括天線陣列(連同接收器和傳輸器,或者接收器-傳輸器對)。

【0123】 雷達圖像的每個圖元可以對應於天線陣列(例如,一維陣列、二維陣列或其他維度)中的天線(和接收器或接收器-傳輸器對)。利用一維陣列,獲取的「圖像」是反射強度圖像,其軸沿著目標的 `range_to_target` 和 `azimuth_angle`。在另一個說明性實例中,天線陣列可以包括 32×32 個天線的陣列,在此種情況下,雷達系統總共包括 1024 個天線。此種雷達系統產生的圖像將包括 32×32 圖元的二維陣列,每個圖元對應於天線。因此,圖像總共將有 1024 個圖元。因此,圖像的寬度和高度-以及每邊的圖元或體素的數量-是陣列中天線數量的函數。至少如此處所論述的,術語「天線」應該被理解為僅代表天線(用於至少一個接收器、傳輸器、收發機,或相應地包括在陣列中或耦合到陣列的其組合),或者可以代表整個接收器、

傳輸器或收發機。如此，天線陣列可以是接收器、傳輸器、收發機或其組合的陣列。

【0124】 在某些情況下，來自雷達系統天線陣列的天線（和接收器）可以將信號分類到不同的距離倉（bin） n 中，該等倉對應於不同的距離範圍。例如，每個天線（和接收器）可以根據相對於傳輸脈衝的到達時間將接收到的RF信號返回分類到一組倉 n 中。時間間隔與到反射RF波的（多個）物件的往返距離成比例。經由檢查倉中的接收信號強度，天線（和接收器）可以在不同倉 n （倉對應於不同範圍）之間分類返回信號。此舉可以在掃描所需的方位角和仰角時執行。具有許多距離倉允許更精確的距離決定。短持續時間的脈衝可以被偵測並映射到少量的距離倉（例如，僅一個或兩個距離倉），而更長的脈衝持續時間、寬度及/或傳輸功率允許更大量的信號能量被傳輸，並且允許接收器整合能量的時間更長，從而導致更長的偵測距離。當接收的信號被分類到距離倉中時，可以為每個距離倉 n 產生雷達圖像。

【0125】 成像雷達感測器可以提供目標車輛的位置（及/或朝向）和維度的暫態估計。在一些實例中，使用基於圖像的處理及/或深度神經網路，當在成像雷達圖像中偵測到目標車輛時，系統200亦可以產生車輛長度的估計。例如，每次成像雷達進行量測時，該成像雷達皆會產生成像雷達周圍場景的圖像。該圖像是灰階圖像，一個軸沿距離，另一個軸沿方位角。圖像中圖元的強度是在特定距離和方位

從反射器返回的強度。被訓練成在雷達圖像上執行物件偵測的深度學習網路可以是由物件偵測引擎 204 使用的物件偵測網路的修改，但被訓練成在雷達圖像上偵測目標物件（例如，車輛）的位置和限度。與相機圖像中的偵測不同，使用雷達圖像的偵測可以在俯視圖（「鳥瞰圖」）中獲取，此舉意味著車輛的長度和寬度被保留在圖像中。因此，使用雷達圖像偵測的副產品是對車輛長度的估計。

【0126】 尺寸估計引擎 214 可以使用來自類概度估計引擎 206、基於地圖的尺寸估計引擎 208 及 / 或基於雷達的尺寸估計引擎 212 的結果的任意組合，來決定目標物件的尺寸（例如，長度及 / 或高度）以及位置及 / 或朝向。例如，在一些情況下，尺寸估計引擎 214 可以應用考慮了來自基於地圖的尺寸估計引擎 208 的基於地圖的尺寸估計、來自基於雷達的尺寸估計引擎 212 的基於雷達的尺寸估計及 / 或來自類概度估計引擎 206 的概度估計的估計模型，來決定目標物件的最終估計尺寸（例如，目標車輛的長度）。在一些實施方式中，估計模型可以包括作為卡爾曼濾波器操作的估計框架（例如，貝氏估計框架或其他估計模型框架）。

【0127】 例如，再次出於說明的目的而使用自主車輛為例，目標車輛的長度（或其他估計維度）的最終估計可以經由順序貝氏估計框架獲取，該框架可以被解釋為簡並的卡爾曼濾波框架，其中表示物件（例如，車輛）長度的狀態被建模為靜態的並且不隨時間變化。例如，由於物件（例如，車輛）的長度是固定的，所以不存在與狀態相關聯的

動力學、不存在狀態轉換、不存在狀態進化等。長度 X 可以假設為高斯隨機變數，其先驗分佈的平均值等於（例如，由類概度估計引擎 206 決定的）該類被追蹤車輛中的車輛的標準長度（或其他估計維度），以及其方差由該類被追蹤車輛的典型長度方差提供。當從上述基於地圖的尺寸估計引擎 208、基於雷達的尺寸估計引擎 212 及 / 或基於雷達圖像的尺寸估計中的一或多個的任意組合接收到新的長度量測值 Y_i 時，可以使用貝氏估計來順序地細化長度估計（或其他估計維度）。該等量測值 Y_i 可以被建模為受高斯雜訊（被表示為 N_i ）擾動的長度的獨立估計值，如下所示：

$$Y_i = X + N_i$$

方程式（7）。

【0128】 來自不同量測值 Y_i 的雜訊的標準差被設置為等於量測值中誤差的標準差。在一些實例中，基於目標車輛的類（由類概度估計引擎 206 決定）對量測值 Y_i 執行異常值拒絕，使得僅可行長度被提供作為貝氏濾波器的輸入。例如，如前述，由類概度估計引擎 206 估計的物件類可用於定義目標物件的尺寸（例如，長度、寬度及 / 或高度）的上限和下限。上限和下限分別表示給定類或類別的目標物件的最大尺寸和最小尺寸。尺寸估計引擎 214 可以拒絕超出上限和下限的任何估計長度（例如，小於下限或大於上限），並且不將其輸入到貝氏濾波器。

【0129】 在處理量測值 Y_i 之後，目標車輛的長度（或其他維度，例如高度）的最佳估計值 \hat{X}_i 可被遞迴計算為：

$$\hat{X}_i = \widehat{X}_{i-1} + \frac{\sigma_{t-1}^2}{\sigma_{t-1}^2 + \sigma_i^2} (Y_i - \widehat{X}_{i-1})$$

方程式 (8) ，

【0130】 其中 σ_{t-1} 是長度的當前估計值 \hat{X}_i 的標準差， σ_i 表示量測值 Y_i 中雜訊 N_i 的標準差。估計值的標準差被更新為：

$$\sigma_t^2 = \frac{\sigma_{t-1}^2 \sigma_w^2}{\sigma_{t-1}^2 + \sigma_w^2}$$

【0131】 儘管長度被用作可由尺寸估計引擎 2 1 4 估計的目標物件的維度的實例，但是相同的方案亦可被用於過濾從基於地圖的尺寸估計引擎 2 0 8 獲取的目標物件（例如，目標車輛）的寬度和高度估計值。在某些情況下，對於某些物件（例如車輛），該等物件的高度和寬度在同一類物件的不同模型之間不會有很大的變化（例如，同一車輛類型的不同模型的寬度和高度有時會有很小的變化）。在此種情況下，尺寸估計引擎 2 1 4 可以基於由類概度估計引擎 2 0 6 辨識的最可能的類，將目標物件（例如，目標車輛或其他物件）的寬度及 / 或高度預測為常數。

【0132】 在多目標追蹤場景中，追蹤物件同時追蹤多個目標物件。例如，自主車輛在行駛時可能需要追蹤道路上的多輛其他車輛。在此種多目標追蹤場景中，尺寸估計引擎 2 1 4 可以在關聯引擎（未圖示）之前，該關聯引擎可以將來自物件偵測引擎 2 0 4 的基於相機的物件偵測、來自基於雷達的尺寸估計引擎 2 1 2 的基於雷達的估計及 / 或基於成像雷達的估計關聯到正確的目標軌跡（其中如前述，軌跡指的是用於每個偵測到的目標物件所保持的估計值）。

【0133】 將不同的資訊源用於不同的尺寸及 / 或位置估計，例如利用基於圖像的物件偵測和基於雷達的偵測的結果的基於地圖的估計，允許不同的估計相互補充。例如，基於圖像的物件偵測可以為兩個不同的物件提供相同或相似的偵測結果，該兩個不同的物件由於偏航角的微小差異而在長度上有很大差異。基於雷達的估計可以補充基於地圖的估計。此外，雷達感測器，作為基於相機的尺寸估計方法的補充解決方案，不太容易受到變化的天氣條件（例如雨或雪）的影響，並且不受日照暴露量變化的影響，從而提高了整體系統穩健性。利用多個資訊源可以幫助系統 200 獲取高度精確的尺寸和位置估計，從而允許包括系統 200 的追蹤物件對追蹤物件周圍的環境進行更精確的估計。例如，使用此處描述的技術，自主車輛可以對可用於機動的道路空間進行更精確的估計，例如執行變道、超車、停車、加速等。在另一個實例中，用於製造的機器人設備可以使用此處描述的技術來更準確地辨識機器人設備可用的可用空間和部件。許多其他技術領域可以受益於此處提供的精確尺寸和位置（及 / 或朝向）估計技術。

【0134】 如前述，針對物件偵測引擎 204，使用基於神經網路的偵測器的各種模型可被用於偵測圖像中的物件。可由物件偵測器使用的神經網路的說明性實例包括迴旋神經網路（CNN）、自動編碼器、深度信念網路（DBN）、遞迴神經網路（RNN）、產生對抗網路（GAN）或任何其他合適的神經網路。

【0135】 圖 13 是可由物件偵測引擎 204 使用的深度學習神經網路 1300 的說明性實例。輸入層 1320 包括輸入資料。在一個說明性實例中，輸入層 1320 可以包括表示輸入視訊框的圖元的資料。神經網路 1300 包括多個隱藏層 1322a、1322b 至 1322n。隱藏層 1322a、1322b 至 1322n 包括「n」個隱藏層，其中「n」是大於或等於 1 的整數。可以使隱藏層的數量包括和給定的應用所需的層一樣多的層。神經網路 1300 亦包括輸出層 1324，其提供由隱藏層 1322a、1322b 至 1322n 執行的處理而產生的輸出。在一個說明性實例中，輸出層 1324 可以為輸入視訊框中的物件提供分類。該分類可以包括辨識物件類型（例如，人、狗、貓或其他物件）的類。

【0136】 神經網路 1300 是互連節點的多層神經網路。每個節點可以代表一條資訊。與節點相關聯的資訊在不同的層之間共享，並且每個層在處理資訊時保留資訊。在一些情況下，神經網路 1300 可以包括前饋網路，在此種情況下，沒有回饋連接，其中網路的輸出被回饋到其自身。在一些情況下，神經網路 1300 可以包括遞迴神經網路，其可以具有允許資訊在讀取輸入時跨節點傳送的迴路。

【0137】 可以經由各層之間的節點到節點互連來在節點之間交換資訊。輸入層 1320 的節點可以啟用第一隱藏層 1322a 中的一組節點。例如，如圖所示，輸入層 1320 的每個輸入節點連接到第一隱藏層 1322a 的每個節點。隱藏層 1322a、1322b 至 1322n 的節點可以經由對該等資訊應用

啟用函數，來變換每個輸入節點的資訊。隨後，從變換中得到的資訊可以被傳遞到並可以啟用下一個隱藏層 1 3 2 2 b 的節點，該節點可以執行其自己的指定函數。示例性函數包括迴旋、上取樣、資料變換及 / 或任何其他合適的函數。隨後，隱藏層 1 3 2 2 b 的輸出可以啟用下一隱藏層的節點，依此類推。最後一個隱藏層 1 3 2 2 n 的輸出可以啟用輸出層 1 3 2 4 的一或多個節點，並在此處提供輸出。在一些情況下，儘管神經網路 1 3 0 0 中的節點（例如，節點 1 3 2 6）被示為具有多條輸出線，但是節點具有單一輸出，並且被圖示為從節點輸出的所有線表示相同的輸出值。

【0138】 在一些情況下，每個節點或節點之間的互連可以具有權重，該權重是從神經網路 1 3 0 0 的訓練中匯出的一組參數。一旦神經網路 1 3 0 0 被訓練，神經網路 1 3 0 0 可以被稱為被訓練過的神經網路，其可以用於對一或多個物件進行分類。例如，節點之間的互連可以表示關於互連節點學習的一條資訊。互連可以具有可調的數值權重，該數值權重可以被調節（例如，基於訓練資料集），從而使得神經網路 1 3 0 0 能夠適應輸入，並且能夠隨著越來越多的資料被處理而學習。

【0139】 神經網路 1 3 0 0 被預先訓練以使用不同的隱藏層 1 3 2 2 a、1 3 2 2 b 至 1 3 2 2 n 來處理來自輸入層 1 3 2 0 中資料的特徵，以便經由輸出層 1 3 2 4 提供輸出。在其中使用神經網路 1 3 0 0 來辨識圖像中的物件的實例中，可以使用包括圖像和標籤的訓練資料來訓練神經網路 1 3 0 0。例如，可以將

訓練圖像輸入到網路中，每個訓練圖像具有指示每個圖像中的一或多個物件的類的標籤（一般，向網路指示物件是什麼以及該等物件具有什麼特徵）。在一個說明性實例中，訓練圖像可以包括數值2的圖像，在此種情況下，圖像的標籤可以是 $[0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$ 。

【0140】 在一些情況下，神經網路1300可以使用稱為反向傳播的訓練過程來調整節點的權重。反向傳播可以包括正向傳遞、損失函數、反向傳遞和權重更新。對一次訓練反覆運算執行正向傳遞、損失函數、反向傳遞和參數更新。對於每組訓練圖像，該過程可以重複一定次數的反覆運算，直到神經網路1300被訓練得足夠好，使得各層的權重被精確調節。

【0141】 對於辨識圖像中的物件的實例，正向傳遞可以包括經由神經網路1300傳遞訓練圖像。在神經網路1300被訓練之前，權重最初是隨機化的。圖像可以包括例如代表圖像圖元的數值陣列。陣列之每一者數值可以包括從0到255之間的值，其描述陣列中該位置的圖元強度。在一個實例中，該陣列可以包括 $28 \times 28 \times 3$ 的數值陣列，具有28行28列圖元以及3種顏色分量（例如紅色、綠色和藍色，或者亮度和兩種色度分量等）。

【0142】 對於神經網路1300的第一次訓練反覆運算，由於在初始化時隨機選擇了權重，因此輸出可能包括不對任何特定類給予偏好的值。例如，若輸出是具有物件包括不同類的概率的向量，則每個不同類的概率值可以相等或者至

少非常相似（例如，對於十個可能的類，每個類可以具有 0.1 的概率值）。利用初始權重，神經網路 1300 不能決定低級別特徵，因此不能準確決定物件的分類可能是什麼。損失函數可用於分析輸出中的誤差。可以使用任何合適的損失函數定義。損失函數的一個實例包括均方誤差（MSE）。MSE 被定義為 $E_{總} = \sum \frac{1}{2}(\text{目標} - \text{輸出})^2$ ，其計算實際答案減去預測（輸出）答案的平方的一半的總和。損失可被設置為等於 $E_{總}$ 的值。

【0143】 對於第一個訓練圖像，損失（或誤差）會很高，因為實際值將與預測輸出相差很大。訓練的目標是最小化損失量，以使得預測的輸出與訓練標籤相同。神經網路 1300 可以經由決定何者輸入（權重）對網路的損失貢獻最大來執行反向傳遞，並且可以調整權重，使得損失減小並最終最小化。

【0144】 可以計算損失相對於權重的導數（表示為 dL/dW ，其中 W 是特定層的權重），以決定對網路損失貢獻最大的權重。在計算導數之後，可以經由更新濾波器的所有權重來執行權重更新。例如，可以更新權重，使其沿梯度的相反朝向變化。權重更新可以表示為 $w = w_i - \eta \frac{dL}{dw}$ ，其中 w 表示權重， w_i 表示初始權重， η 表示學習率。學習率可被設置為任何合適的值，高學習率包括較大的權重更新，較低的值表示較小的權重更新。

【0145】 神經網路 1300 可以包括任何合適的深度網路。一個實例包括迴旋神經網路（CNN），其包括輸入層和輸出

層，在輸入層和輸出層之間有多個隱藏層。下文參照圖 14 描述 CNN 的實例。CNN 的隱藏層包括一系列迴旋層、非線性層、池化層（用於下取樣）和全連接層。神經網路 1300 可以包括除了 CNN 之外的任何其他深度網路，例如自動編碼器、深度信念網路（DBN）、遞迴神經網路（RNN）等。

【0146】 圖 14 是迴旋神經網路 1400（CNN 1400）的說明性實例。CNN 1400 的輸入層 1420 包括表示圖像的資料。例如，資料可以包括表示圖像圖元的數值陣列，陣列之每一者數值包括從 0 到 255 之間的值，用於描述陣列中該位置的圖元強度。使用前面的實例，該陣列可以包括 $28 \times 28 \times 3$ 的數值陣列，具有 28 行 28 列的圖元和 3 個顏色分量（例如，紅色、綠色和藍色，或者亮度和兩個色度分量等）。圖像可以經由迴旋隱藏層 1422a、可選的非線性啟用層、池化隱藏層 1422b 和全連接隱藏層 1422c 傳遞，以在輸出層 1424 得到輸出。儘管圖 14 中僅圖示每個隱藏層中的一個，但是一般技術者將會理解，CNN 1400 中可以包括多個迴旋隱藏層、非線性層、池化隱藏層及 / 或全連接層。如前述，輸出可以指示物件的單個類，或者可以包括最佳描述圖像中的物件的類的概率。

【0147】 CNN 1400 的第一層是迴旋隱藏層 1422a。迴旋隱藏層 1422a 分析輸入層 1420 的圖像資料。迴旋隱藏層 1422a 的每個節點連接到被稱為感受野的輸入圖像的節點（圖元）區域。迴旋隱藏層 1422a 可以被認為是一或多個濾波器（每個濾波器對應於不同的啟用或特徵映射），其

中濾波器的每個迴旋反覆運算是迴旋隱藏層 1 4 2 2 a 的節點或神經元。例如，濾波器在每次迴旋反覆運算中覆蓋的輸入圖像的區域將是濾波器的感受野。在一個說明性實例中，若輸入圖像包括 28×28 陣列，並且每個濾波器（和相應的感受野）是 5×5 陣列，則在迴旋隱藏層 1 4 2 2 a 中將有 24×24 個節點。節點和該節點的感受野之間的每個連接學習權重，並且在某些情況下學習整體偏置，使得每個節點學習分析其在輸入圖像中的特定局部感受野。隱藏層 1 4 2 2 a 的每個節點將具有相同的權重和偏置（被稱為共享權重和共享偏置）。例如，濾波器具有權重（數值）陣列和具有與輸入相同的深度。對於視訊訊框實例，濾波器的深度為 3（根據輸入圖像的三種顏色分量）。濾波器陣列的說明性實例尺寸是 $5 \times 5 \times 3$ ，對應於節點的感受野的尺寸。

【0148】 迴旋隱藏層 1 4 2 2 a 的迴旋特性是由於迴旋層的每個節點被應用於其相應的感受野。例如，迴旋隱藏層 1 4 2 2 a 的濾波器可以從輸入圖像陣列的左上角開始，並且可以圍繞輸入圖像進行迴旋。如前述，濾波器的每個迴旋反覆運算可以被認為是迴旋隱藏層 1 4 2 2 a 的節點或神經元。在每次迴旋反覆運算中，濾波器的值與相應數量的圖像原始圖元值相乘（例如， 5×5 濾波器陣列與輸入圖像陣列左上角的輸入圖元值的 5×5 陣列相乘）。每次迴旋反覆運算的乘積可以相加在一起，以獲取該反覆運算或節點的總和。接下來，根據迴旋隱藏層 1 4 2 2 a 中的下一個節點的感受野，在輸入圖像中的下一個位置繼續該過程。例如，濾波器可

以向下一個感受野移動步長量。步長量可以設置為1或其他合適的量。例如，若步長量設置為1，則每次迴旋反覆運算時，濾波器將向右移動1個圖元。在輸入量的每個唯一位置處理濾波器會產生代表該位置的濾波結果的數值，從而為迴旋隱藏層1422a的每個節點決定總和值。

【0149】 從輸入層到迴旋隱藏層1422a的映射被稱為啟用映射（或特徵映射）。啟用映射包括每個節點的值，該值代表每個位置處的輸入量的過濾結果。啟用映射可以包括此種陣列，該陣列包括由輸入量上的濾波器的每次反覆運算產生的各種總和值。例如，若對 28×28 輸入圖像的每個圖元（步長量為1）應用 5×5 濾波器，啟用映射將包括 24×24 陣列。迴旋隱藏層1422a可以包括幾個啟用映射，以便辨識圖像中的多個特徵。圖14所示的實例包括三個啟用映射。使用三個啟用映射，迴旋隱藏層1422a可以偵測三種不同類型的特徵，每個特徵在整個圖像上皆是可偵測的。

【0150】 在一些實例中，可以在迴旋隱藏層1422a之後應用非線性隱藏層。非線性層可用於將非線性引入一直在計算線性運算的系統。非線性層的一個說明性實例是整流線性單元（ReLU）層。ReLU層可以將函數 $f(x) = \max(0, x)$ 應用於輸入量中的所有值，此舉會將所有負啟用更改為0。ReLU因此可以在不影響迴旋隱藏層1422a的感受野的情況下增加CNN 1400的非線性特性。

【0151】 可以在迴旋隱藏層1422a之後（並且當使用時，在非線性隱藏層之後）應用池化隱藏層1422b。池化隱藏

層 1 4 2 2 b 用於簡化來自迴旋隱藏層 1 4 2 2 a 的輸出中的資訊。例如，池化隱藏層 1 4 2 2 b 可以取得從迴旋隱藏層 1 4 2 2 a 輸出的每個啟用映射，並使用池化函數產生壓縮啟用映射（或特徵映射）。最大池化是由池化隱藏層執行的函數的一個實例。池化隱藏層 1 4 2 2 a 可以使用其他形式的池化函數，例如平均池化、L 2 範數池化或其他合適的池化函數。池化函數（例如，最大池化濾波器、L 2 範數濾波器或其他合適的池化濾波器）被應用於迴旋隱藏層 1 4 2 2 a 中包括的每個啟用映射。在圖 1 4 所示的實例中，三個池化濾波器用於迴旋隱藏層 1 4 2 2 a 中的三個啟用映射。

【0152】 在一些實例中，可以經由將具有步長量（例如，等於濾波器的維度，例如步長量 2）的最大池化濾波器（例如，具有 2×2 的尺寸）應用於從迴旋隱藏層 1 4 2 2 a 輸出的啟用映射，來使用最大池化。最大池化濾波器的輸出包括濾波器迴旋的每個子區域中的最大數量。以 2×2 濾波器為例，池化層之每一者單元可以匯總上一層中 2×2 個節點的區域（每個節點是啟用映射中的一個值）。例如，啟用映射中的四個值（節點）將在濾波器的每次反覆運算中由 2×2 最大池化濾波器進行分析，四個值中的最大值作為「max」值輸出。若此種最大池化濾波器被應用於來自具有 24×24 節點維度的迴旋隱藏層 1 4 2 2 a 的啟用濾波器，則來自池化隱藏層 1 4 2 2 b 的輸出將是 12×12 節點的陣列。

【0153】 在一些實例中，亦可以使用 L 2 範數池化濾波器。L 2 範數池化濾波器包括計算啟用映射的 2×2 區域（或其他

合適的區域) 中的值的平方和的平方根(而不是像在最大池化中一般計算最大值), 並且使用所計算的值作為輸出。

【0154】 直觀地, 池化函數(例如, 最大池化、L2 範數池化或其他池化函數) 決定是否能在圖像區域的任何地方找到給定特徵, 並丟棄精確的位置資訊。此舉可以在不影響特徵偵測結果的情況下完成, 因為一旦發現特徵, 特徵的精確位置就不如相對於其他特徵的近似位置重要。最大池化(以及其他池化方法) 的好處是被池化的特徵更少, 從而減少了 CNN 1400 後續層所需的參數數量。

【0155】 網路中的連接的最後一層是全連接層, 其將池化隱藏層 1422b 的每個節點連接到輸出層 1424 之每一者輸出節點。使用上述實例, 輸入層包括對輸入圖像的圖元強度進行編碼的 28×28 個節點; 基於對三個啟用映射應用 5×5 個局部感受野(用於濾波器), 迴旋隱藏層 1422a 包括 $3 \times 24 \times 24$ 個隱藏特徵節點; 並且, 基於對三個特徵映射中的每一個的 2×2 個區域應用最大池化濾波器, 池化層 1422b 包括具有 $3 \times 12 \times 12$ 個隱藏特徵節點的層。對該實例進行擴展, 輸出層 1424 可以包括十個輸出節點。在此種實例中, $3 \times 12 \times 12$ 池化隱藏層 1422b 的每個節點連接到輸出層 1424 的每個節點。

【0156】 全連接層 1422c 可以獲取先前池化層 1422b 的輸出(其應當代表高級別特徵的啟用映射), 並決定與特定類最相關的特徵。例如, 全連接層 1422c 可以決定與特定類最相關的高級別特徵, 並且可以包括高級別特徵的權

重(節點)。可以計算全連接層 1422c 和池化隱藏層 1422b 的權重之間的乘積，以獲取不同類的概率。例如，若 CNN 1400 正被用於預測視訊框中的物件是人，則在表示人的高級別特徵(例如，有兩條腿，在物件的頂部存在臉，在臉的左上和右上有兩隻眼睛，在臉的中間有一個鼻子，臉的底部有一張嘴，及/或人常見的其他特徵)的啟用映射中將出現高值。

【0157】 在一些實例中，來自輸出層 1424 的輸出可以包括 M 維向量(在以前的實例中， $M=10$)，其中 M 可以包括在對圖像中的物件進行分類時程式必須選擇從中選擇的類的數量。亦可以提供其他示例性輸出。N 維向量之每一者數值皆可以表示物件屬於某一類的概率。在一個說明性實例中，若表示十個不同類的物件的 10 維輸出向量是 $[0.05 \ 0.8 \ 0 \ 0.15 \ 0 \ 0 \ 0 \ 0]$ ，則該向量指示圖像是第三類物件(例如，狗)的概率為 5%，該圖像是第四類物件(例如，人)的概率為 80%，該圖像是第六類物件(例如，袋鼠)的概率為 15%。類的概率可以被認為是該物件是該類的一部分的置信度。

【0158】 可由物件偵測引擎 204 使用的基於神經網路的偵測器的一個具體實例是基於 Cifar-10 神經網路的偵測器。圖 15 是圖示 Cifar-10 神經網路 1500 的實例的圖。在某些情況下，可以訓練 Cifar-10 神經網路僅對人和車進行分類。如圖所示，Cifar-10 神經網路 1500 包括各種迴旋層 (Conv1 層 1502、Conv2/Relu2 層 1508 和

Conv3/Relu3 層 1514)、多個池化層 (Pool1/Relu1 層 1504、Pool2 層 1510 和 Pool3 層 1516) 以及混合在其中的整流線性單元層。亦提供了正規化層 Norm1 1506 和 Norm2 1512。最後一層是 ip1 層 1518。

【0159】 可被物件偵測引擎 204 用來偵測或分類圖像中的物件的另一種基於深度學習的偵測器包括 SSD 偵測器，其是可應用於多個物件類別或類的快速單次物件偵測器。SSD 模型使用附接到神經網路頂部的多個特徵映射的多尺度迴旋邊界框輸出。此種表示允許 SSD 有效地建模不同的框形狀。圖 16A 包括圖像，圖 16B 和圖 16C 包括圖示 SSD 偵測器 (具有 VGG 深度網路基礎模型) 如何操作的圖。例如，SSD 將物件與具有不同縱橫比的預設框相匹配 (在圖 16B 和圖 16C 中顯示為虛線矩形)。特徵映射的每個元素皆有多個與之關聯的預設框。任何預設框，其與真實值框的交並比 (intersection-over-union) 超過閾值 (例如，0.4、0.5、0.6 或其他合適的閾值)，則被認為與該物件匹配。例如，8×8 框中的兩個 (圖 16B 中的框 1602 和框 1604) 與貓匹配，4×4 框中的一個 (圖 16C 中的框 1606) 與狗匹配。SSD 有多個特徵映射，每個特徵映射負責不同比例的物件，從而使其能夠辨識大範圍比例的物件。例如，圖 16B 的 8×8 特徵映射中的框比圖 16C 的 4×4 特徵映射中的框小。在一個說明性實例中，SSD 偵測器總共可以具有六個特徵映射。

【0160】 對於每個單元之每一者預設框，SSD神經網路輸出長度為 c 的概率向量，其中 c 是類的數量，表示包含每個類的物件的框的概率。在某些情況下，會包含背景類，指示框中沒有物件。SSD網路亦輸出（用於每個單元之每一者預設框的）偏移向量，該向量具有包含使預設框與底層物件的邊界框匹配所需的預測偏移的四個條目。該等向量以 (c_x, c_y, w, h) 格式提供，其中 c_x 表示中心 x ， c_y 表示中心 y ， w 表示寬度偏移， h 表示高度偏移。只有當預設框中實際包含物件時，向量才有意義。對於圖16A中所示的圖像，除了三個匹配的框（兩個用於貓，一個用於狗）之外，所有概率標籤將指示背景類。

【0161】 另一種基於深度學習的偵測器可由物件偵測引擎204用來偵測或分類圖像中的物件，包括「你僅需看一次」（YOLO）偵測器，YOLO偵測器是SSD物件偵測系統的替代。圖17A包括圖像，圖17B和圖17C包括圖示YOLO偵測器如何操作的圖。YOLO偵測器可以將單個神經網路應用於完整圖像。如圖所示，YOLO網路將圖像分成多個區域，並預測每個區域的邊界框和概率。該等邊界框由預測的概率加權。例如，如圖17A所示，YOLO偵測器將圖像分成 13×13 個單元的網格。每個單元負責預測五個邊界框。置信度分數被提供以指示預測的邊界框實際上包圍物件的確定程度。該分數不包括可能在框中的物件的分類，但是指示框的形狀是否合適。預測的邊界框如圖17B所示。置信度分數較高的框具有較粗的邊界。

【0162】 每個單元亦預測每個邊界框的類。例如，所有可能的類的概率分佈被提供。可以偵測任意數量的類，例如自行車、狗、貓、人、汽車或其他合適的物件類。邊界框的置信度分數和類預測被組合成最終分數，該最終分數指示該邊界框包含特定類型的物件的概率。例如，圖 17 B 中圖像左側帶有粗邊框的灰色框是 85% 決定其包含物件類「狗」。此處有 169 (13 x 13) 個網格單元，每個單元預測 5 個邊界框，因此總共有 1745 個邊界框。許多邊界框將具有非常低的分數，在此種情況下，僅保留最終分數高於閾值（例如，高於 30% 概率、40% 概率、50% 概率或其他合適閾值）的框。圖 17 C 圖示具有最終預測的邊界框和類的圖像，包括狗、自行車和汽車。如圖所示，從產生的總共 1745 個邊界框中，僅有圖 17 C 所示的三個邊界框被保留，因為其最終分數最好。

【0163】 圖 18 是圖示使用本文描述的物件驗證技術經由使用雷達圖像執行物件驗證的過程 1800 的實例的流程圖。在方塊 1802 處，過程 1800 包括獲取辨識在圖像中偵測到的第一物件的邊界區域。第一物件可以包括任何被追蹤的物件。例如，在一些實例中，第一物件是車輛。在一些實例中，物件是機器人。在一些實例中，物件是飛機。

【0164】 在方塊 1804 處，過程 1800 包括獲取包括複數個地圖點的地圖。複數個地圖點對應於三維空間中的一或多個參考位置。例如，地圖可以包括來自圖 2 的點狀地圖

207。在一些態樣，一或多個參考位置包括三維空間中道路上的車道，如以上提供的說明性實例中所述。

【0165】 在方塊1806處，過程1800包括將辨識第一物件的邊界區域與地圖中包括的複數個地圖點中的至少一個地圖點相關聯。在一些實例中，過程1800包括將來自複數個地圖點的地圖點子集投影到圖像上，以及從地圖點子集中決定至少一個地圖點是一或多個參考位置上離邊界區域最近的地圖點。過程1800可以包括基於決定至少一個地圖點是一或多個參考位置上離邊界區域最近的地圖點，將該邊界區域與該至少一個地圖點相關聯。如前述，在一些情況下，僅有具有正平行距離的地圖點可以與物件的邊界區域相關聯。在一些實例中，使用校準資料和用於擷取圖像的相機的姿態將地圖點子集投影到圖像上。在一些情況下，地圖點子集包括被包括在用於擷取圖像的相機的視野內並且處於相機的選擇範圍內的地圖點。

【0166】 在一些實施方式中，過程1800包括決定至少一個地圖點與邊界區域邊緣上的一或多個點相交。在此種實施方式中，基於決定至少一個地圖點與邊界區域的邊緣上的一或多個點相交，該至少一個地圖點被決定為最近的地圖點。在一些情況下，如前述，邊界區域的邊緣包括邊界區域的底部邊緣。

【0167】 在一些實施方式中，過程1800包括決定邊界區域的邊緣上的點，以及決定與來自地圖點子集的其他地圖點相比、至少一個地圖點最靠近邊界區域的邊緣上的點。在

此種實施方式中，基於至少一個地圖點最靠近邊界區域的邊緣上的點，該至少一個地圖點被決定為最近的地圖點。在一些情況下，如前述，邊界區域的邊緣包括邊界區域的底部邊緣。

【0168】 在方塊 1808 處，過程 1800 包括使用邊界區域和至少一個地圖點來決定圖像中偵測到的第一物件的估計三維位置和估計尺寸。

【0169】 在一些實例中，決定第一物件的估計三維位置和估計尺寸包括獲取第一物件的寬度，決定第一物件的朝向在至少一個地圖點處平行於地圖的表面，以及基於第一物件的寬度和第一物件的朝向決定代表第一物件的三維邊界框的頂點位置。頂點位置對應於最靠近用於擷取圖像的相機的三維邊界框的拐角（例如，如圖 9A 及 / 或圖 9B 所示）。過程 1800 可以包括基於頂點位置決定第一物件的估計三維位置，以及基於頂點位置和邊界區域決定代表第一物件的三維邊界框的長度和高度。

【0170】 在一些實例中，過程 1800 包括獲取第一物件的分類，以及基於該分類決定第一物件的寬度。

【0171】 在一些實例中，過程 1800 包括獲取圖像的第一物件的分類，其中該分類定義物件的類型。過程 1800 可以包括基於第一物件的分類來決定第一物件的最小尺寸和最大尺寸，以及可以基於第一物件的最小尺寸和最大尺寸來決定第一物件的估計尺寸。例如，如前述，由概度估計來估計的物件的物件類可用於定義目標物件的尺寸（例如，長

度、寬度及/或高度)的上限和下限,其可由估計模型(例如,卡爾曼濾波器)用於決定物件的估計尺寸。

【0172】 在一些實例中,過程1800包括基於第一物件和第二物件之間的距離來決定分類的權重,其中第二物件包括用於擷取圖像的相機。過程1800可以包括基於權重更新分類的累積權重。累積權重基於來自包括第一物件的多個圖像的多次偵測。過程1800可以包括基於分類的累積權重來決定第一物件包括由分類定義的物件類型的概度性,以及可以基於所決定的概度性和與一或多個其他分類相關聯的一或多個概度性來決定第一物件包括該物件類型。過程1800可以包括基於由分類定義的物件類型來決定第一物件的最小尺寸和最大尺寸。上文關於方程式(1)-(3)描述了說明性的實例。在一些實例中,第一物件是第一車輛,第二物件是第二車輛。在一些態樣,由分類定義的物件類型包括第一車輛的車輛類型。

【0173】 在一些實例中,過程1800包括獲取複數個雷達量測點,其中複數個雷達量測點是基於由第一物件反射的雷達信號。過程1800可以包括基於複數個雷達量測點來決定第一物件的附加估計尺寸。在一些實例中,使用包括在第二物件上的複數個雷達獲取複數個雷達量測點,其中第二物件包括用於擷取圖像的相機。在一些實施方式中,過程1800包括決定雷達量測點沿著第一物件的朝向方向的最大擴展,以及基於雷達量測點沿著第一物件的朝向方向的最大擴展來決定第一物件的附加估計尺寸。在一些實例

中，過程 1800 包括使用估計模型來決定第一物件的尺寸，在此種情況下，估計尺寸和附加估計尺寸被用作估計模型的輸入。在一些實施方式中，估計模型是卡爾曼濾波器。

【0174】 在一些態樣，過程 1800 包括從成像雷達獲取雷達圖像，以及基於雷達圖像決定第一物件的附加估計尺寸。在一些實例中，過程 1800 包括使用估計模型來決定第一物件的尺寸，在此種情況下，估計尺寸和附加估計尺寸被用作估計模型的輸入。在一些實施方式中，估計模型是卡爾曼濾波器。

【0175】 在一些實例中，此處描述的過程（例如，過程 1800 及 / 或此處描述的其他過程）可以由計算設備或裝置來執行，諸如實現圖 2 的系統 200 的計算設備。在一些實例中，計算設備可以包括圖 19 的計算設備 1900 的架構。在一個實例中，過程 1800 可以由實現系統 200 的計算設備 1900 來執行。計算設備可以包括任何合適的設備，例如自主車輛、機器人設備、行動設備（例如，行動電話）、臺式計算設備、平板計算設備、可穿戴設備及 / 或具有能夠執行此處描述的過程（包括過程 1800）的資源能力的任何其他計算設備。在一些情況下，計算設備或裝置可以包括各種元件，例如一或多個輸入設備、一或多個輸出設備、一或多個處理器、一或多個微處理器、一或多個微型電腦、一或多個相機、一或多個感測器及 / 或被配置為執行此處描述的過程步驟的其他元件。在一些實例中，計算設備可以包括顯示器、被配置為傳送及 / 或接收資料的網路介面、其任意組合

及/或其他元件。網路介面可以被配置為傳送及/或接收基於網際網路協定（IP）的資料或其他類型的資料。

【0176】 計算設備的元件可以用電路系統實現。例如，元件可以包括及/或可以使用電子電路或其他電子硬體來實現，電子電路或其他電子硬體可以包括一或多個可程式設計電子電路（例如，微處理器、圖形處理單元（GPU）、數位信號處理器（DSP）、中央處理單元（CPU）及/或其他合適的電子電路），及/或可以包括及/或使用電腦軟體、韌體或其任意組合來實現，以執行本文描述的重種操作。

【0177】 過程1800被圖示為邏輯流程圖，其操作表示可以以硬體、電腦指令或其組合來實現的一系列操作。在電腦指令的上下文中，操作表示儲存在一或多個電腦可讀取儲存媒體上的電腦可執行指令，該等電腦可執行指令當由一或多個處理器執行時，執行所述操作。通常，電腦可執行指令包括執行特定功能或實現特定資料類型的常式、程式、物件、元件、資料結構等。該等操作被描述的順序並不意欲被解釋為限制，並且任何數量的所述操作可以以任何順序組合及/或並行組合以執行過程。

【0178】 另外，此處描述的過程可以在配置有可執行指令的一或多個電腦系統的控制下執行，並且可以經由硬體或其組合實現為在一或多個處理器上共同執行的代碼（例如，可執行指令、一或多個電腦程式或一或多個應用程式）。如前述，代碼可以儲存在電腦可讀取或機器可讀取儲存媒體上，例如，以包括可由一或多個處理器執行的複

數個指令的電腦程式的形式。電腦可讀取或機器可讀取儲存媒體可以是非暫時性的。

【0179】 圖 19 圖示可以實現此處描述的各種技術的示例性計算設備的示例性計算設備架構 1900。例如，計算設備架構 1900 可以實現圖 2 所示的系統 200。計算設備架構 1900 的元件被示為使用諸如匯流排的連接 1905 彼此電通訊。示例性計算設備架構 1900 包括處理單元（CPU 或處理器）1910 和計算設備連接 1905，計算設備連接 1905 將包括計算設備記憶體 1915（例如唯讀記憶體（ROM）1920 和隨機存取記憶體（RAM）1925）在內的各種計算設備元件耦合到處理器 1910。

【0180】 計算設備架構 1900 可以包括與處理器 1910 直接連接、緊密地連接或整合為處理器 1910 的一部分的高速記憶體的快取記憶體。計算設備架構 1900 可以將資料從記憶體 1915 及 / 或儲存設備 1930 複製到快取記憶體 1912，以供處理器 1910 快速存取。如此，快取記憶體可以提供效能提升，從而避免處理器 1910 在等待資料時的延遲。該等和其他模組可以控制或被配置為控制處理器 1910 執行各種動作。亦可以使用其他計算設備記憶體 1915。記憶體 1915 可以包括具有不同效能特性的多種不同類型的記憶體。處理器 1910 可以包括任何通用處理器以及硬體或軟體服務，例如儲存在儲存設備 1930 中的服務 1 1932、服務 2 1934 和服務 3 1936，其被配置為控制處理器 1910 以及將軟體指令併入處理器設計中的專用處理器。處理器 1910 可

以是包含多個核心或處理器、匯流排、記憶體控制器、快取記憶體等的獨立系統。多核處理器可以是對稱的，亦可以是非對稱的。

【0181】 為了實現使用者與計算設備架構 1900 的互動，輸入設備 1945 可以表示任何數量的輸入機制，例如用於語音的麥克風、用於手勢或圖形輸入的觸控式螢幕、鍵盤、滑鼠、運動輸入、語音等等。輸出設備 1935 亦可以是熟習此項技術者已知的多種輸出機制中的一或多個，例如顯示器、投影儀、電視、揚聲器設備等。在一些情況下，多模式計算設備可以使使用者能夠提供多種類型的輸入以與計算設備架構 1900 通訊。通訊介面 1940 通常可以管理使用者輸入和計算設備輸出。在任何特定的硬體配置上操作沒有任何限制，因此隨著硬體或韌體配置的發展，此處的基本特徵可以容易地被改良的硬體或韌體配置所替代。

【0182】 儲存設備 1930 是非揮發性記憶體，並且可以是硬碟或其他類型的電腦可讀取媒體，其可以儲存可由電腦存取的資料，例如磁帶、快閃記憶卡、固態記憶體設備、數位多功能光碟、盒式磁帶、隨機存取記憶體 (RAM) 1925、唯讀記憶體 (ROM) 1920 及其組合。儲存設備 1930 可以包括用於控制處理器 1910 的服務 1932、1934、1936。可以考慮其他硬體或軟體模組。儲存設備 1930 可以連接到計算設備連接 1905。在一個態樣，執行特定功能的硬體模組可以包括儲存在電腦可讀取媒體中的軟體元件，該軟體

元件與諸如處理器 1910、連接 1905、輸出設備 1935 等必要的硬體元件結合以執行該功能。

【0183】 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設備、光學儲存設備以及能夠儲存、包含或攜帶指令及/或資料的各種其他媒體。電腦可讀取媒體可以包括其中可以儲存資料的非暫時性媒體，並且不包括以無線方式或經由有線連接傳播的載波及/或暫時性電子信號。非暫時性媒體的實例可以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。電腦可讀取媒體上可以儲存代碼及/或機器可執行指令，該等指令可以代表程序、函數、副程式、程式、常式、子常式、模組、套裝軟體、軟體元件或指令、資料結構或程式語句的任意組合。程式碼片段可以經由傳遞及/或接收資訊、資料、實際參數（argument）、形式參數（parameter）或記憶體內容而耦合到另一個程式碼片段或硬體電路。資訊、實際參數、形式參數、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而被傳遞、轉發或傳輸。

【0184】 在一些實施例中，電腦可讀取儲存設備、媒體和記憶體可以包括包含位元串流等的電纜或無線信號。然而，當被提及時，非暫時性電腦可讀取儲存媒體明確地排除諸如能量、載波信號、電磁波和信號本身之類的媒體。

【0185】 在以上描述中提供了具體細節，以提供對本文所提供的實施例和實例的透徹理解。然而，一般技術者將理解，實施例可以在沒有該等具體細節的情況下實施。為了解釋清楚，在一些情況下，本技術可以被呈現為包括單獨的功能方塊，該等功能方塊包括設備、設備元件、以軟體或硬體和軟體的組合實現的方法中的步驟或常式。除了圖中所示及/或本文所述的彼等之外，可以使用附加的元件。例如，電路、系統、網路、過程和其他元件可以以方塊圖形式圖示為元件，以避免在不必要的細節上使實施例模糊。在其他情況下，可以在沒有不必要的細節的情況下圖示公知的電路、過程、演算法、結構和技術，以避免使實施例模糊。

【0186】 各個實施例可以在上文中被描述為過程或方法，其被圖示為流程圖、流程示意圖、資料流程圖、結構圖或方塊圖。儘管流程圖可以將操作描述為連續過程，但是許多操作可以並行或併發地執行。此外，可以重新安排操作的順序。當過程的操作完成時，該過程被終止，但是可以有圖中沒有包括的附加步驟。過程可以對應於方法、函數、程序、子常式、副程式等。當過程對應於函數時，過程的終止可以對應於該函數返回到調用函數或主函數。

【0187】 根據上述實例的過程和方法可以使用儲存在電腦可讀取媒體中或可從電腦可讀取媒體獲得的電腦可執行指令來實現。此種指令可以包括例如可致使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

群組的指令和資料。使用的部分電腦資源可以經由網路存取。電腦可執行指令可以是，例如二進位檔案、中間格式指令（例如組合語言、韌體、原始程式碼等）。根據所描述的實例，可用於儲存在方法期間的指令、所使用的資訊及/或所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、配備有非揮發性記憶體的USB設備、網路儲存設備等。

【0188】 根據該等揭示實現過程和方法的設備可以包括硬體、軟體、韌體、中間軟體、微碼、硬體描述語言或其任意組合，並且可以採用各種形式因素中的任何一種。當以軟體、韌體、中間軟體或微代碼實現時，執行必要任務的程式碼或程式碼片段（例如，電腦程式產品）可以儲存在電腦可讀取或機器可讀取媒體中。（多個）處理器可以執行必要的任務。形式因素的典型實例包括筆記型電腦、智慧手機、行動電話、平板設備或其他小型個人電腦、個人數位助理、機架式設備、獨立設備等。此處描述的功能亦可以體現在周邊設備或外掛程式卡中。作為另一個實例，此種功能亦可以在單個設備中執行的不同晶片或者不同過程之間在電路板上實現。

【0189】 指令、用於傳達此種指令的媒體、用於執行該等指令的計算資源以及用於支援此種計算資源的其他結構是用於提供本文中描述的功能的示例性手段。

【0190】 在前面的描述中，參照本案的具體實施例描述了本案的各態樣，但是熟習此項技術者將認識到本案不限於

此。因此，儘管此處已經詳細描述了本案的說明性實施例，但是應當理解，發明構思可以以其他方式不同地體現和應用，並且所附請求項意欲被解釋為包括此種變化，除非受到現有技術的限制。上述申請案的各種特徵和態樣可以單獨或共同使用。此外，在不脫離本說明書的更廣泛的精神和範疇的情況下，實施例可以在除了此處描述的環境和應用之外的任何數量的環境和應用中使用。因此，說明書和附圖被認為是說明性的，而不是限制性的。為了說明起見，方法是按特定順序描述的。應當理解，在替代實施例中，該等方法可以以不同於所描述的順序來執行。

【0191】 一般技術者將會理解，在不脫離本說明書的範疇的情況下，此處使用的小於（「 $<$ 」）和大於（「 $>$ 」）符號或術語可以分別用小於或等於（「 \leq 」）和大於或等於（「 \geq 」）符號來代替。

【0192】 在元件被描述為被「配置為」執行某些操作的情況下，此種配置可以例如經由設計電子電路或其他硬體以執行操作、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意組合來實現。

【0193】 術語「耦合到」是指直接或間接地被實體連接到另一元件的任何元件，及/或直接或間接地與另一元件通訊（例如，經由有線或無線連接及/或其他合適的通訊介面被連接到另一元件）的任何元件。

【0194】 引述集合中的「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的一個成員或該集合的多個成員（以任意組合）滿足該請求項。例如，引述「A和B中的至少一個」的請求項語言是指A、B或A和B。在另一個實例中，引述「A、B和C中的至少一個」的請求項語言是指A、B、C或A和B，或A和C，或B和C，或A和B和C。集合中的「至少一個」及/或集合中的「一或多個」的語言並不將該集合限制為該集合中所列的項。例如，引述「A和B中的至少一個」的請求項語言可以表示A、B或A和B，並且可以另外包括未在A和B的集合中列出的項。

【0195】 結合本文揭示的實施例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現為電子硬體、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，上文已經大體上從功能性態樣對各種說明性的元件、方塊、模組、電路和步驟進行了描述。此種功能性用硬體還是軟體實現取決於特定的應用和對整體系統施加的設計約束。熟習此項技術者可以針對每個特定應用以不同的方式實現所描述的功能，但是此種實現決策不應被解釋為導致脫離本案的範疇。

【0196】 此處描述的技術亦可以在電子硬體、電腦軟體、韌體或其任意組合中實現。此種技術可以在多種設備中的任何一種中實現，例如通用電腦、無線通訊設備手機或具有多種用途的積體電路元件，包括在無線通訊設備手機和

其他設備中的應用。被描述為模組或元件的任何特徵可以在整合邏輯設備中一起實現，或者作為個別但可交互操作的邏輯設備單獨實現。若以軟體實現，該等技術可以至少部分地由包括程式碼的電腦可讀取資料儲存媒體來實現，該程式碼包括當被執行時執行一或多個上述方法的指令。電腦可讀取資料儲存媒體可以形成電腦程式產品的一部分，其可以包括封裝材料。電腦可讀取媒體可以包括記憶體或資料儲存媒體，例如隨機存取記憶體（RAM），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，該等技術可以至少部分地由電腦可讀取通訊媒體來實現，該電腦可讀取通訊媒體攜帶或傳送指令形式或資料結構形式的程式碼，並且可以由電腦存取、讀取及/或執行，例如傳播的信號或波。

【0197】 程式碼可以由處理器執行，該處理器可以包括一或多個處理器，例如一或多個數位信號處理器（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個別邏輯電路系統。此種處理器可以被配置為執行本案中描述的任何技術。通用處理器可以是微處理器；但是可選地，處理器可以是任何習知的處理器、控制器、微控制器或狀態機。處理器亦可以被實現為計算設備的組合，例如，DSP和微

處理器、複數個微處理器、與 D P S 核心配合使用的一或多個微處理器的組合或者任何其他此種配置。因此，此處使用的術語「處理器」可以指任何前述結構、前述結構的任何組合，或者適合於實現此處描述的技術的任何其他結構或裝置。

【符號說明】

【0198】

1 0 0 : 圖 像

1 0 2 : 追 蹤 車 輛

1 0 4 : 目 標 車 輛

1 0 6 : 目 標 車 輛

1 0 8 : 目 標 車 輛

2 0 0 : 系 統

2 0 2 : 相 機

2 0 3 : 圖 像

2 0 4 : 物 件 偵 測 引 擎

2 0 5 : 圖 像

2 0 6 : 類 概 度 估 計 引 擎

2 0 7 : 點 狀 地 圖

2 0 8 : 基 於 地 圖 的 尺 寸 估 計 引 擎

2 0 9 : 邊 界 區 域

2 1 0 : 雷 達

2 1 1 : 雷 達 量 測 值

2 1 2 : 基 於 雷 達 的 尺 寸 估 計 引 擎

- 2 1 4 : 尺 寸 估 計 引 擎
- 3 2 0 : 車 輛
- 3 2 2 : 左 車 道
- 3 2 4 : 中 間 車 道
- 3 2 6 : 右 車 道
- 3 2 8 : 點 子 集 選 擇 範 圍
- 4 3 0 : 相 機
- 4 3 2 : 底 部 平 面
- 4 3 4 : 基 本 路 線
- 4 3 5 : 當 前 地 圖 點
- 4 3 6 : 下 一 地 圖 點
- 4 3 7 : 交 點
- 5 0 5 : 圖 像
- 5 1 0 : 圖 像
- 5 4 0 : 邊 界 框
- 5 4 1 : 車 道 中 心 線
- 5 4 2 : 地 圖 點
- 5 4 3 : 地 圖 點
- 5 4 4 : 車 道 中 心 線
- 5 4 5 : 地 圖 點
- 5 4 6 : 地 圖 點
- 5 5 0 : 邊 界 框
- 5 5 1 : 車 道 中 心 線
- 5 5 2 : 車 道 中 心 線

5 5 3 : 車 道 中 心 線

6 3 0 : 側 面 相 機

6 3 1 : 邊 界 框

6 3 3 : 車 道 中 心 線

6 3 4 : 車 道 中 心 線

6 3 5 : 地 圖 點

6 3 6 : 後 續 地 圖 點

6 3 8 : 交 點

6 3 9 : 平 行 距 離

7 1 0 : 圖 像

7 3 9 : 平 行 距 離

7 5 0 : 邊 界 框

7 5 1 : 車 道 中 心 線

7 5 2 : 車 道 中 心 線

7 5 3 : 車 道 中 心 線

7 5 4 : 地 圖 點

7 5 6 : 中 點

8 1 0 : 圖 像

8 5 0 : 邊 界 框

8 5 1 : 車 道 中 心 線

8 5 2 : 車 道 中 心 線

8 5 3 : 車 道 中 心 線

8 5 6 : 點

9 5 0 : 2 D 邊 界 框 (2 D B B)

960: 3D 朝向邊界框 (3D OBB)

970: 2D 邊界框

980: 3D 朝向邊界框 (3D OBB)

1000: 曲線圖

1100: 曲線圖

1300: 深度學習神經網路

1320: 輸入層

1322a: 隱藏層

1322b: 隱藏層

1322n: 隱藏層

1324: 輸出層

1326: 節點

1400: 迴旋神經網路

1420: 輸入層

1422a: 迴旋隱藏層

1422b: 池化隱藏層

1422c: 全連接隱藏層

1424: 輸出層

1500: Cifar-10 神經網路

1502: Conv1 層

1504: Pool1 / ReLU1 層

1506: 正規化層 Norm1

1508: Conv2 / ReLU2 層

1510: Pool2 層

1 5 1 2 : N o r m 2
1 5 1 4 : C o n v 3 / R e l u 3 層
1 5 1 6 : P o o l 3 層
1 5 1 8 : i p 1 層
1 6 0 2 : 框
1 6 0 4 : 框
1 6 0 6 : 框
1 8 0 0 : 過 程
1 8 0 2 : 方 塊
1 8 0 4 : 方 塊
1 8 0 6 : 方 塊
1 8 0 8 : 方 塊
1 9 0 0 : 計 算 設 備
1 9 0 5 : 連 接
1 9 1 0 : 處 理 器
1 9 1 2 : 快 取 記 憶 體
1 9 1 5 : 計 算 設 備 記 憶 體
1 9 2 0 : 唯 讀 記 憶 體 (R O M)
1 9 2 5 : 隨 機 存 取 記 憶 體 (R A M)
1 9 3 0 : 儲 存 設 備
1 9 3 2 : 服 務 1
1 9 3 4 : 服 務 2
1 9 3 5 : 輸 出 設 備
1 9 3 6 : 服 務 3

1940: 通訊介面

1945: 輸入設備

【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】 一種決定一或多個物件的一或多個尺寸的方法，該方法包括以下步驟：

獲取辨識在一圖像中偵測到的一第一物件的一邊界區域；

獲取包括複數個地圖點的一地圖，該複數個地圖點對應於一三維空間中的一或多個參考位置；

將辨識該第一物件的該邊界區域與該地圖中包括的該複數個地圖點中的至少一個地圖點相關聯；及

使用該邊界區域和該至少一個地圖點，來決定在該圖像中偵測到的該第一物件的一估計三維位置和一估計尺寸。

【請求項 2】 根據請求項 1 之方法，其中該第一物件是一車輛。

【請求項 3】 根據請求項 2 之方法，其中該一或多個參考位置包括該三維空間中一道路上的車道。

【請求項 4】 根據請求項 1 之方法，亦包括以下步驟：

將來自該複數個地圖點的一地圖點子集投影到該圖像上；

從該地圖點子集中決定該至少一個地圖點是該一或多個參考位置上離該邊界區域最近的一地圖點；及

基於決定該至少一個地圖點是該一或多個參考位置上離該邊界區域最近的該地圖點，將該邊界區域與該至少一個地圖點相關聯。

【請求項5】 根據請求項4之方法，其中使用校準資料和用於擷取該圖像的一相機的一姿態，將該地圖點子集投影到該圖像上。

【請求項6】 根據請求項4之方法，其中該地圖點子集包括被包括在用於擷取該圖像的一相機的一視野內、並且處於該相機的一選擇範圍內的地圖點。

【請求項7】 根據請求項4之方法，亦包括以下步驟：

決定該至少一個地圖點與該邊界區域的一邊緣上的一或多個點相交；

其中基於決定該至少一個地圖點與該邊界區域的一邊緣上的一或多個點相交，該至少一個地圖點被決定為該最近的地圖點。

【請求項8】 根據請求項7之方法，其中該邊界區域的該邊緣包括該邊界區域的一底部邊緣。

【請求項9】 根據請求項4之方法，亦包括以下步驟：

決定該邊界區域的一邊緣上的一點；及

決定與來自該地圖點子集的其他地圖點相比、該至少一個地圖點最靠近該邊界區域的該邊緣上的該點；

其中基於該至少一個地圖點最靠近該邊界區域的該邊緣上的該點，該至少一個地圖點被決定為該最近的地圖點。

【請求項10】 根據請求項9之方法，其中該邊界區域的該邊緣包括該邊界區域的一底部邊緣。

【請求項11】 根據請求項1之方法，其中決定該第一物件

的該估計三維位置和該估計尺寸之步驟包括以下步驟：

獲取該第一物件的一寬度；

決定該第一物件的一朝向在該至少一個地圖點處平行於該地圖的一表面；

基於該第一物件的該寬度和該第一物件的該朝向，決定代表該第一物件的一三維邊界框的一頂點位置，其中該頂點位置對應於該三維邊界框的最靠近用於擷取該圖像的一相機的一拐角；

基於該頂點位置來決定該第一物件的該估計三維位置；及

基於該頂點位置和該邊界區域，來決定代表該第一物件的該三維邊界框的一長度和高度。

【請求項 12】 根據請求項 11 之方法，亦包括以下步驟：

獲取該第一物件的一分類；及

基於該分類決定該第一物件的該寬度。

【請求項 13】 根據請求項 1 之方法，亦包括以下步驟：

獲取該圖像的該第一物件的一分類，該分類定義物件的一類型；

基於該第一物件的該分類，來決定該第一物件的一最小尺寸和一最大尺寸；及

基於該第一物件的該最小尺寸和該最大尺寸，來決定該第一物件的該估計尺寸。

【請求項 14】 根據請求項 13 之方法，亦包括以下步驟：

基於該第一物件和一第二物件之間的一距離來決定該

分類的一權重，該第二物件包括用於擷取該圖像的一相機；

基於該權重來更新該分類的一累積權重，其中該累積權重基於來自包括該第一物件的多個圖像的多次偵測；

基於該分類的該累積權重，來決定該第一物件包括由該分類定義的該物件類型的一概度性；

基於該被決定的概度性和與一或多個其他分類相關聯的一或多個概度性，來決定該第一物件包括該物件類型；及

基於由該分類定義的該物件類型，來決定該第一物件的該最小尺寸和該最大尺寸。

【請求項 15】 根據請求項 14 之方法，其中該第一物件是一第一車輛，該第二物件是一第二車輛。

【請求項 16】 根據請求項 15 之方法，其中由該分類定義的該物件類型包括該第一車輛的一車輛類型。

【請求項 17】 根據請求項 1 之方法，亦包括以下步驟：

獲取複數個雷達量測點，其中該複數個雷達量測點基於由該第一物件反射的雷達信號；及

基於該複數個雷達量測點來決定該第一物件的一附加估計尺寸。

【請求項 18】 根據請求項 17 之方法，其中該複數個雷達量測點是經由使用該第二物件上包括的複數個雷達而被獲取的，其中該第二物件包括用於擷取該圖像的一相機。

【請求項 19】根據請求項 17 之方法，亦包括以下步驟：

決定該複數個雷達量測點沿著該第一物件的一朝向方向的一最大擴展，以及

基於該複數個雷達量測點沿著該第一物件的該朝向方向的該最大擴展，來決定該第一物件的該附加估計尺寸。

【請求項 20】根據請求項 17 之方法，亦包括以下步驟：

使用一估計模型來決定該第一物件的一尺寸，其中該估計尺寸和該附加估計尺寸被用作該估計模型的輸入。

【請求項 21】根據請求項 20 之方法，其中該估計模型是一卡爾曼濾波器。

【請求項 22】根據請求項 1 之方法，亦包括以下步驟：

從一成像雷達獲取一雷達圖像；及

基於該雷達圖像決定該第一物件的一附加估計尺寸。

【請求項 23】根據請求項 22 之方法，亦包括以下步驟：

使用一估計模型來決定該第一物件的一尺寸，其中該估計尺寸和該附加估計尺寸被用作該估計模型的輸入。

【請求項 24】一種用於決定一或多個物件的一或多個尺寸的裝置，包括：

一記憶體，被配置為儲存至少一個圖像；及

以電路系統實現的一處理器，被配置為：

獲取辨識在一圖像中偵測到的一第一物件的一邊界區域；

獲取包括複數個地圖點的一地圖，該複數個地圖點

對應於一三維空間中的一或多個參考位置；

將辨識該第一物件的該邊界區域與該地圖中包括的該複數個地圖點中的至少一個地圖點相關聯；及

使用該邊界區域和該至少一個地圖點來決定在該圖像中偵測到的該第一物件的一估計三維位置和一估計尺寸。

【請求項 25】根據請求項 24 之裝置，其中該第一物件是一車輛。

【請求項 26】根據請求項 25 之裝置，其中該一或多個參考位置包括該三維空間中一道路上的車道。

【請求項 27】根據請求項 24 之裝置，其中該處理器被配置為：

將來自該複數個地圖點的一地圖點子集投影到該圖像上；

從該地圖點子集中決定該至少一個地圖點是該一或多個參考位置上離該邊界區域最近的一地圖點；及

基於決定該至少一個地圖點是該一或多個參考位置上離該邊界區域最近的該地圖點，將該邊界區域與該至少一個地圖點相關聯。

【請求項 28】根據請求項 27 之裝置，其中使用校準資料和用於擷取該圖像的一相機的一姿態，將該地圖點子集投影到該圖像上。

【請求項 29】根據請求項 27 之裝置，其中該地圖點子集包括被包括在用於擷取該圖像的一相機的一視野內、並

且處於該相機的一選擇範圍內的地圖點。

【請求項30】根據請求項27之裝置，其中該處理器被配置為：

決定該至少一個地圖點與該邊界區域的一邊緣上的一或多個點相交；

其中基於決定該至少一個地圖點與該邊界區域的一邊緣上的一或多個點相交，該至少一個地圖點被決定為該最近的地圖點。

【請求項31】根據請求項30之裝置，其中該邊界區域的該邊緣包括該邊界區域的一底部邊緣。

【請求項32】根據請求項27之裝置，其中該處理器被配置為：

決定該邊界區域的一邊緣上的一點；及

決定與來自該地圖點子集的其他地圖點相比、該至少一個地圖點最靠近該邊界區域的該邊緣上的該點；

其中基於該至少一個地圖點最靠近該邊界區域的該邊緣上的該點，該至少一個地圖點被決定為該最近的地圖點。

【請求項33】根據請求項32之裝置，其中該邊界區域的該邊緣包括該邊界區域的一底部邊緣。

【請求項34】根據請求項24之裝置，其中為了決定該第一物件的該估計三維位置和該估計尺寸，該處理器被配置為：

獲取該第一物件的一寬度；

決定該第一物件的一朝向在該至少一個地圖點處平行於該地圖的一表面；

基於該第一物件的該寬度和該第一物件的該朝向，決定代表該第一物件的一三維邊界框的一頂點位置，其中該頂點位置對應於該三維邊界框的最靠近用於擷取該圖像的一相機的一拐角；

基於該頂點位置，來決定該第一物件的該估計三維位置；及

基於該頂點位置和該邊界區域，來決定代表該第一物件的該三維邊界框的一長度和高度。

【請求項 35】 根據請求項 34 之裝置，其中該處理器被配置為：

獲取該第一物件的一分類；及

基於該分類決定該第一物件的該寬度。

【請求項 36】 根據請求項 24 之裝置，其中該處理器被配置為：

獲取該圖像的該第一物件的一分類，該分類定義物件的一類型；

基於該第一物件的該分類，來決定該第一物件的一最小尺寸和一最大尺寸；及

基於該第一物件的該最小尺寸和該最大尺寸，來決定該第一物件的該估計尺寸。

【請求項 37】 根據請求項 36 之裝置，其中該處理器被配置為：

基於該第一物件和一第二物件之間的一距離來決定該分類的一權重，該第二物件包括用於擷取該圖像的一相機；

基於該權重來更新該分類的一累積權重，該累積權重基於來自包括該第一物件的多個圖像的多次偵測；

基於該分類的該累積權重，來決定該第一物件包括由該分類定義的該物件類型的一概度性；

基於該被決定的概度性和與一或多個其他分類相關聯的一或多個概度性，來決定該第一物件包括該物件類型；及

基於由該分類定義的該物件類型，來決定該第一物件的該最小尺寸和該最大尺寸。

【請求項38】 根據請求項 37 之裝置，其中該第一物件是一第一車輛，該第二物件是一第二車輛。

【請求項39】 根據請求項 38 之裝置，其中由該分類定義的該物件類型包括該第一車輛的一車輛類型。

【請求項40】 根據請求項 24 之裝置，其中該處理器被配置為：

獲取複數個雷達量測點，該複數個雷達量測點基於由該第一物件反射的雷達信號；及

基於該複數個雷達量測點來決定該第一物件的一附加估計尺寸。

【請求項41】 根據請求項 40 之裝置，其中該複數個雷達量測點是經由使用該第二物件上包括的複數個雷達而被

獲取的，其中該第二物件包括用於擷取該圖像的一相機。

【請求項 4 2】 根據請求項 4 0 之裝置，其中該處理器被配置為：

決定該複數個雷達量測點沿著該第一物件的一朝向方向的一最大擴展，以及

基於該複數個雷達量測點沿著該第一物件的該朝向方向的該最大擴展，來決定該第一物件的該附加估計尺寸。

【請求項 4 3】 根據請求項 4 0 之裝置，其中該處理器被配置為：

使用一估計模型來決定該第一物件的一尺寸，該估計尺寸和該附加估計尺寸被用作該估計模型的輸入。

【請求項 4 4】 根據請求項 4 3 之裝置，其中該估計模型是一卡爾曼濾波器。

【請求項 4 5】 根據請求項 2 4 之裝置，其中該處理器被配置為：

從一成像雷達獲取一雷達圖像；及

基於該雷達圖像決定該第一物件的一附加估計尺寸。

【請求項 4 6】 根據請求項 4 5 之裝置，其中該處理器被配置為：

使用一估計模型來決定該第一物件的一尺寸，其中該估計尺寸和該附加估計尺寸被用作該估計模型的輸入。

【請求項 4 7】 根據請求項 2 4 之裝置，其中該裝置是一車

輛。

【請求項48】根據請求項24之裝置，其中該裝置是一機器人。

【發明圖式】

圖像
100

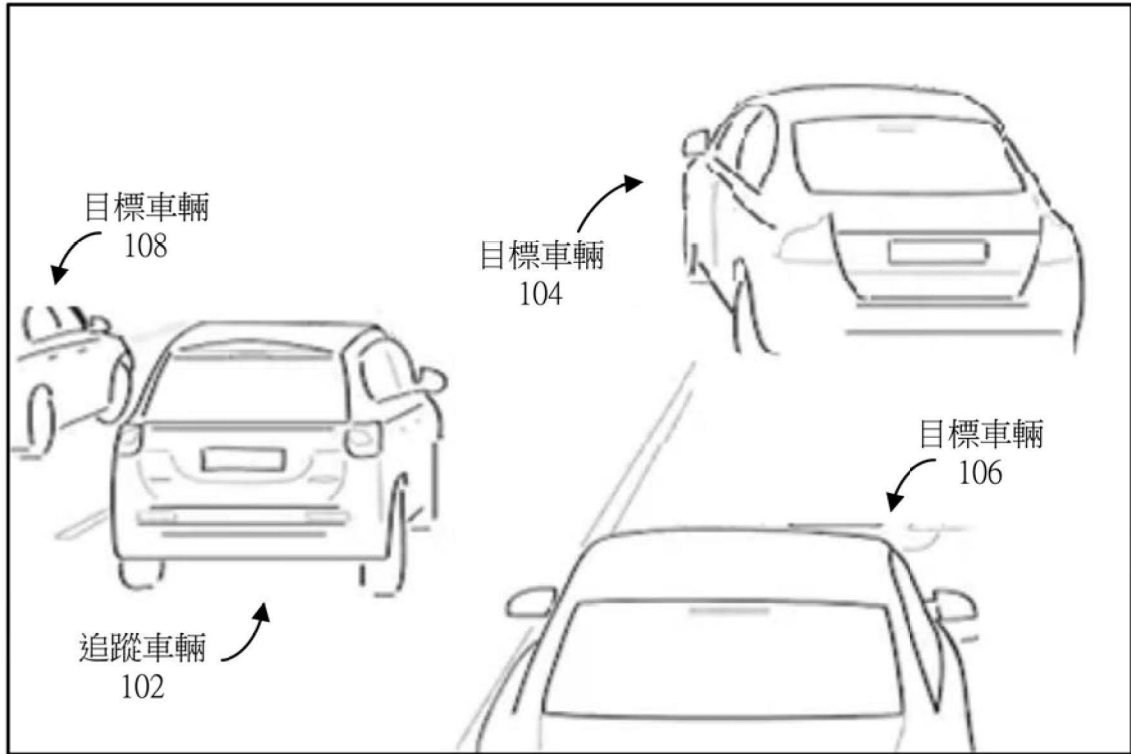


圖1

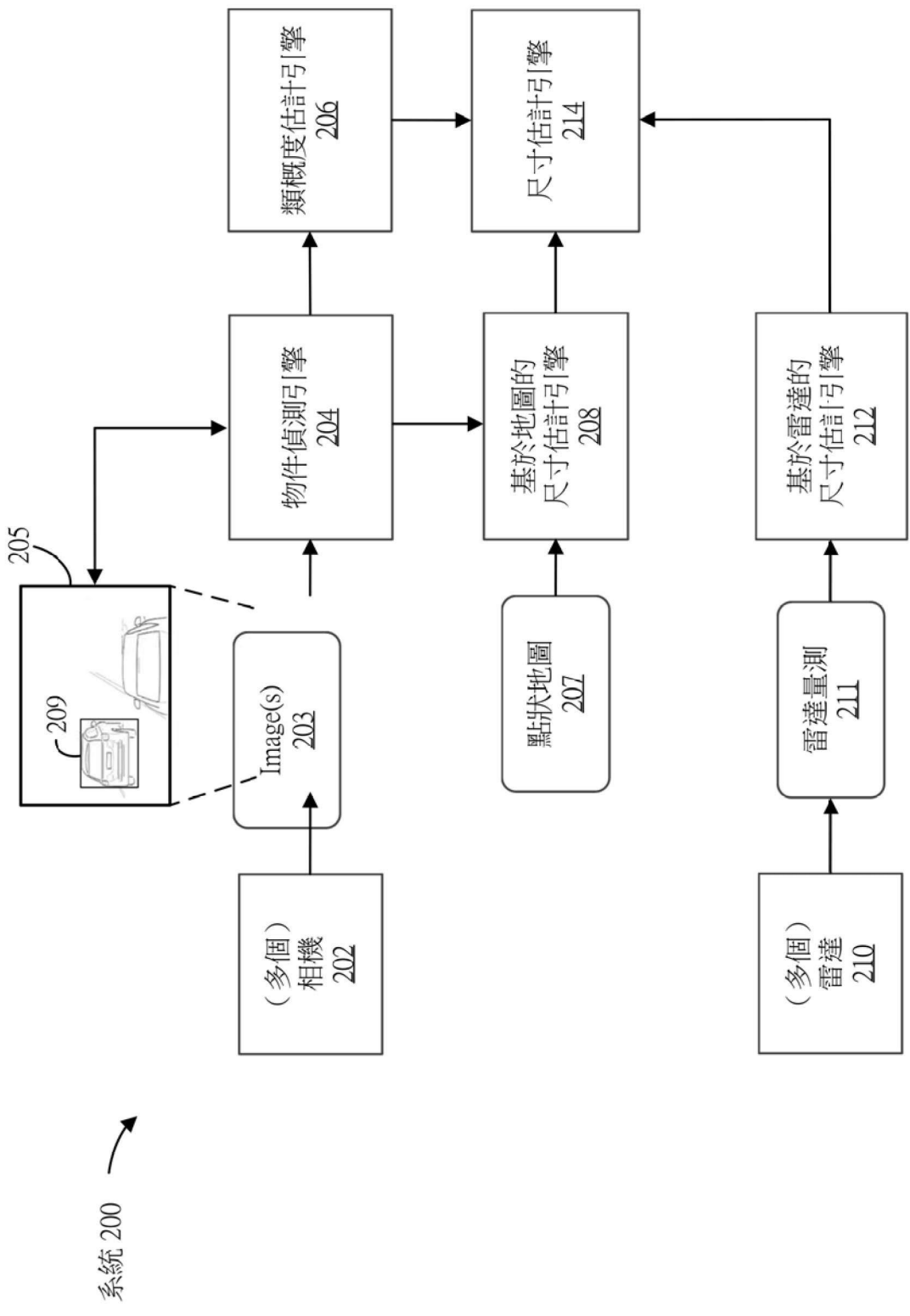
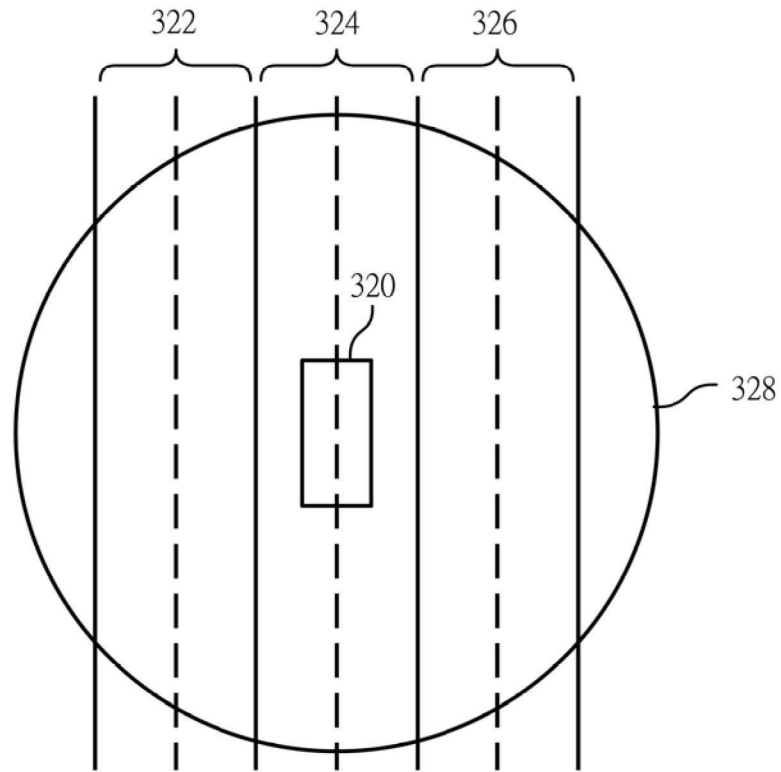


圖2







-  車道中心線
-  車道邊界線
-  追蹤車輛
-  點子集選擇範圍

圖3

430

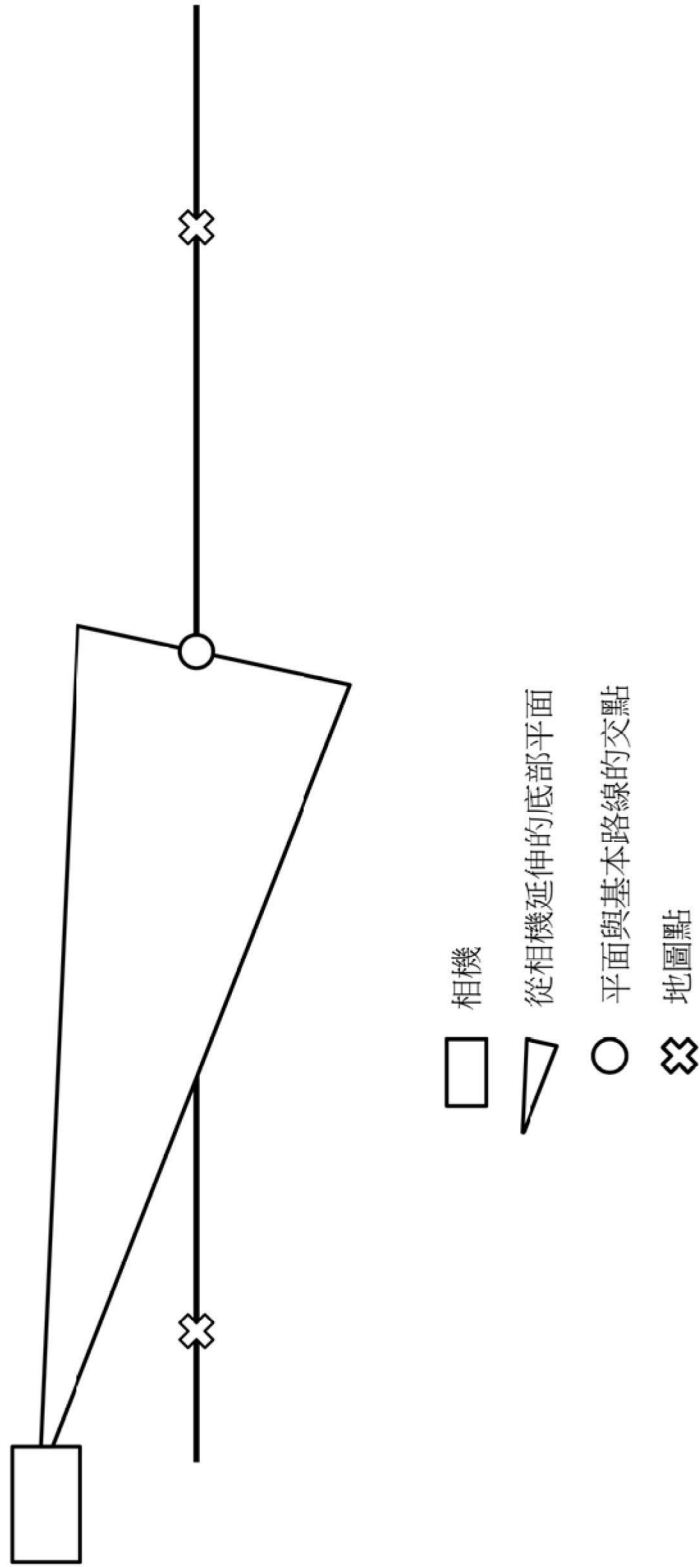


圖4

圖像
505

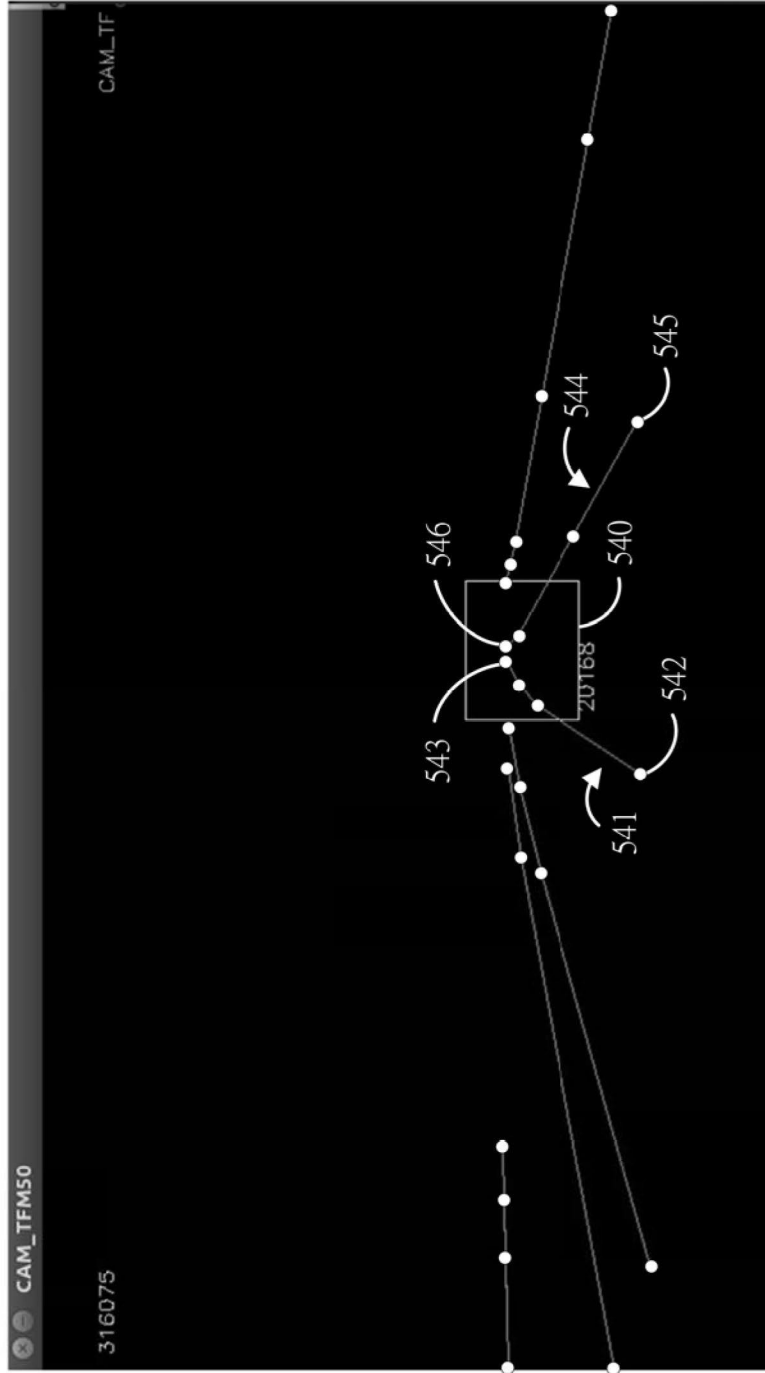


圖5A

圖像
510

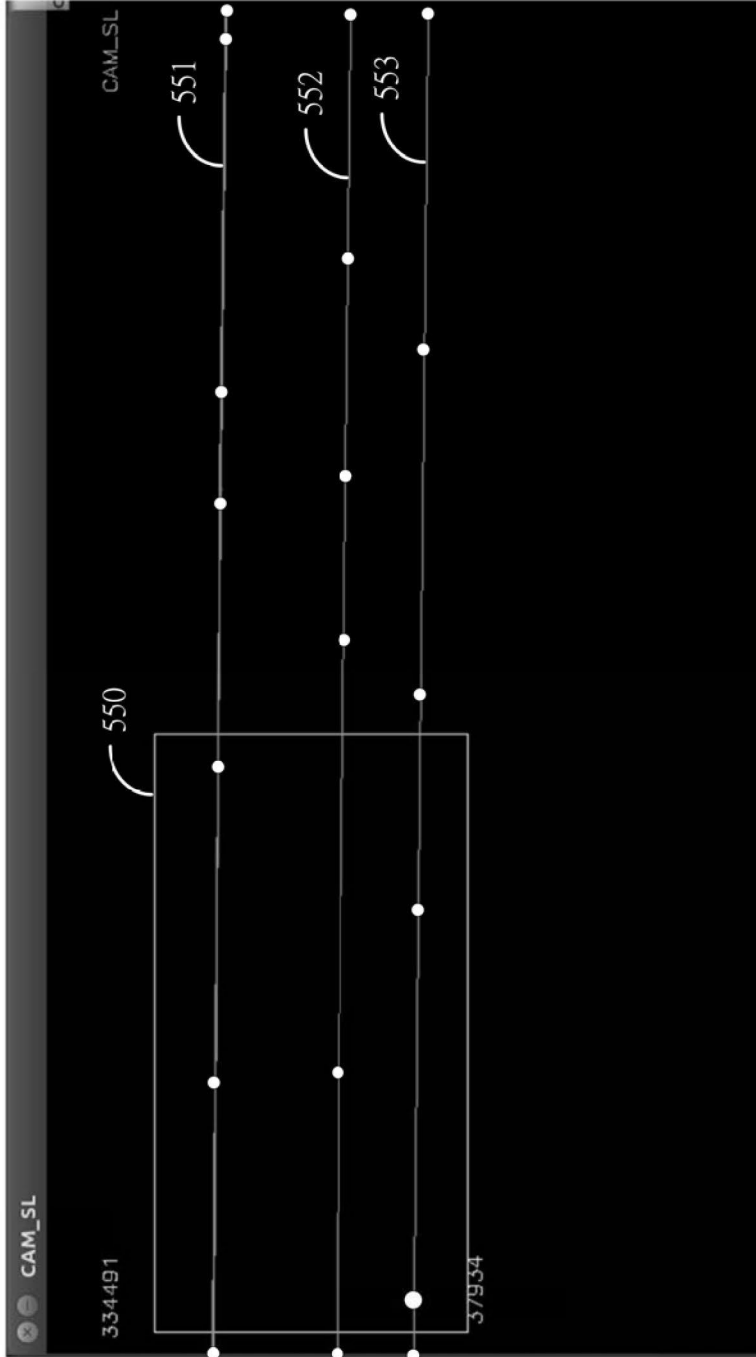
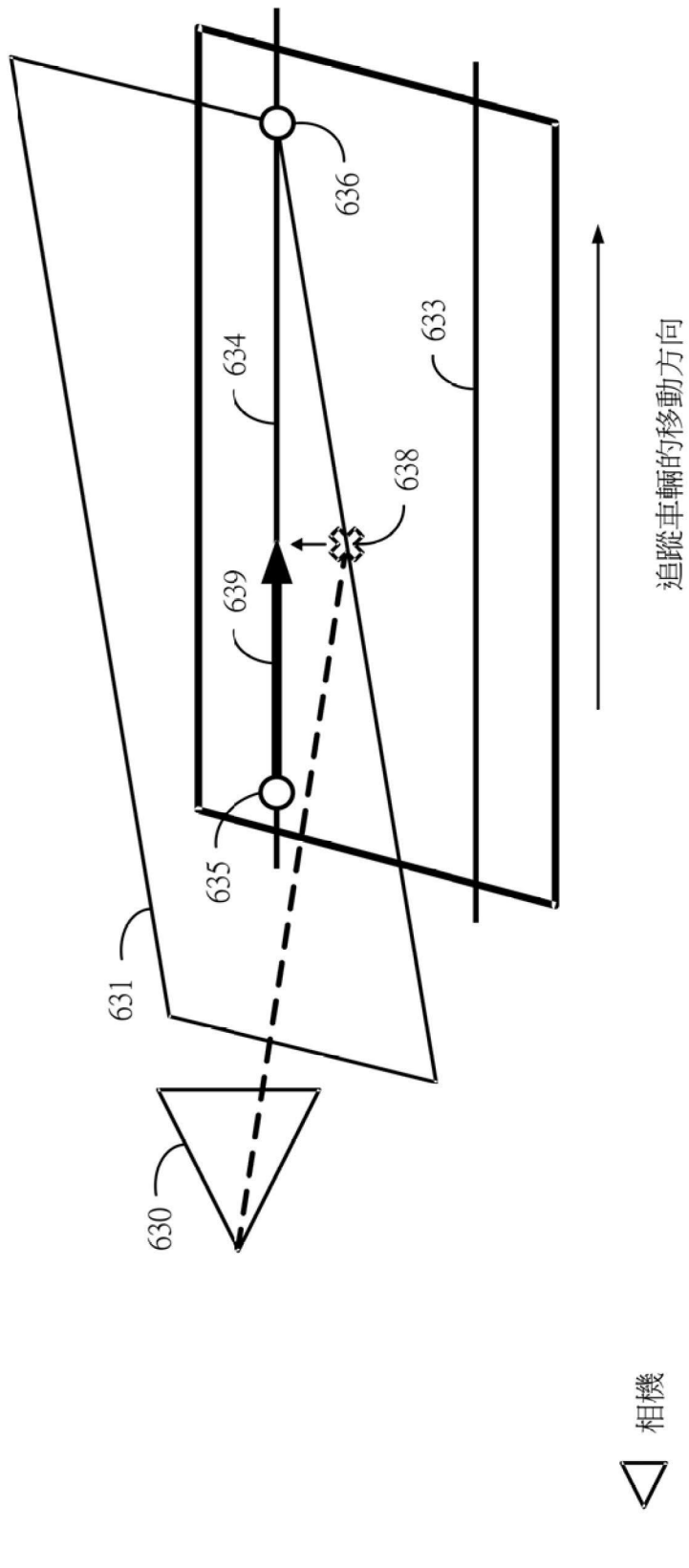


圖5B



- ▽ 相機
- - - 來自交點的射線
- ▭ 道路平面
- 車道中心線
- 地圖點
- ⊗ 射線與道路平面的交點
- ↑ 平行距離

圖像
710

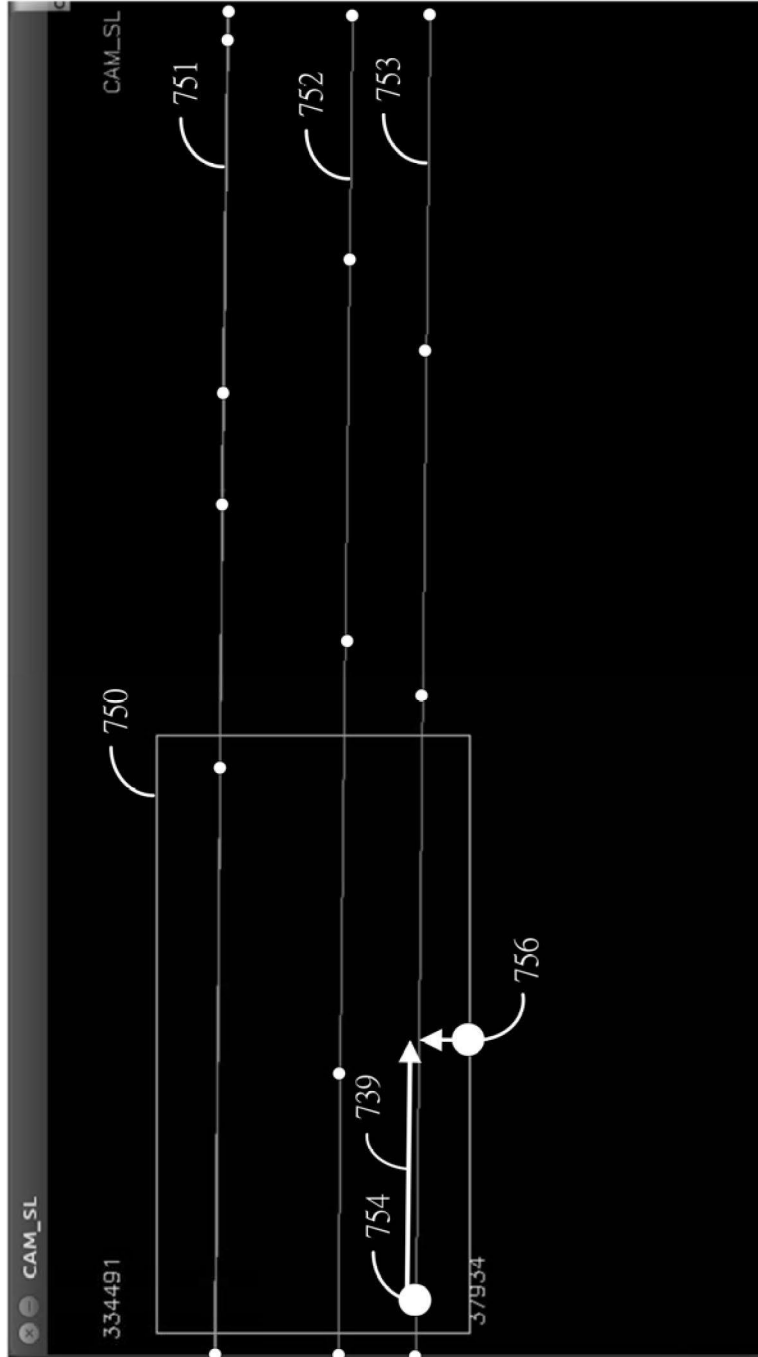


圖7

圖像
810

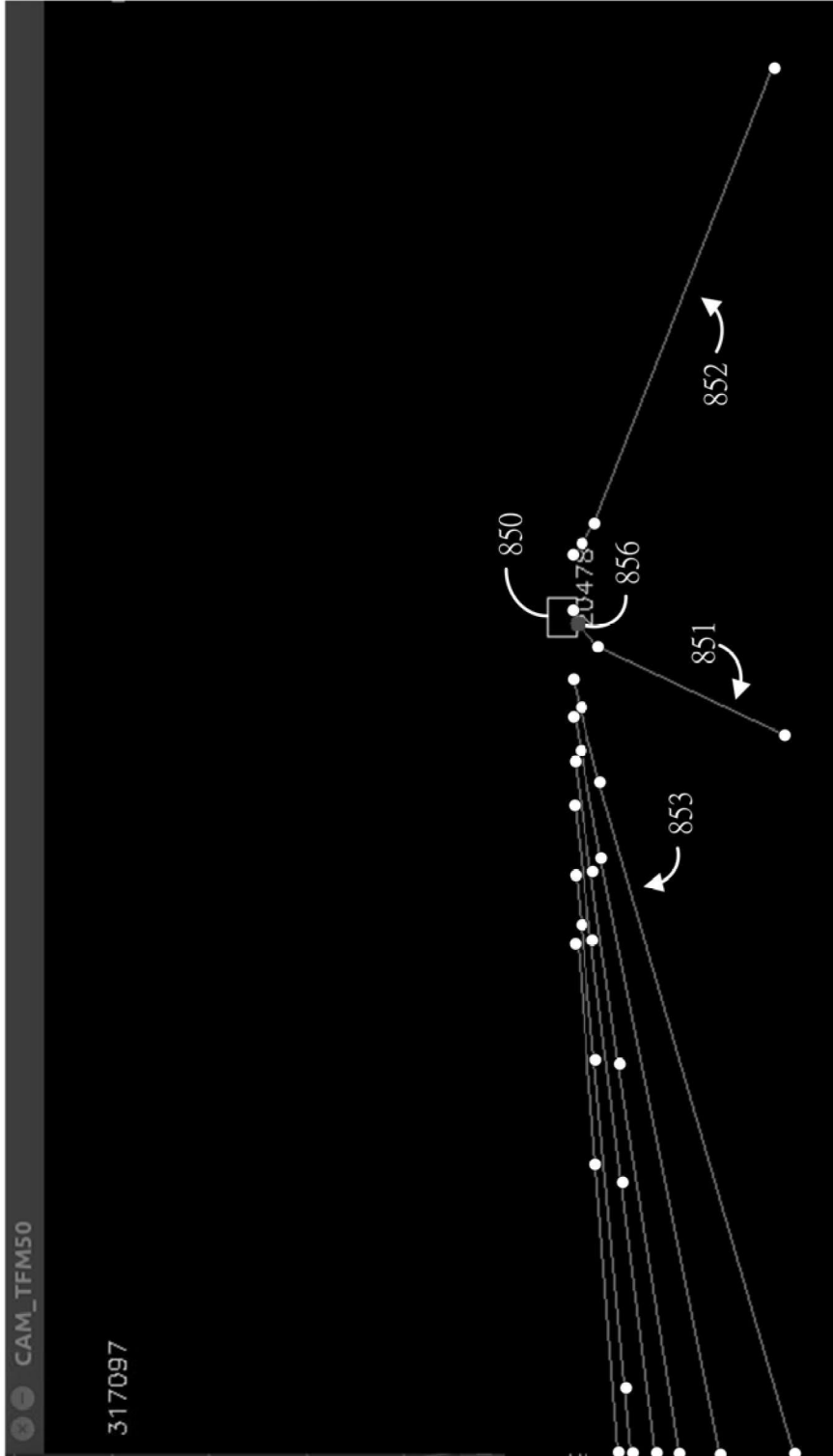


圖8

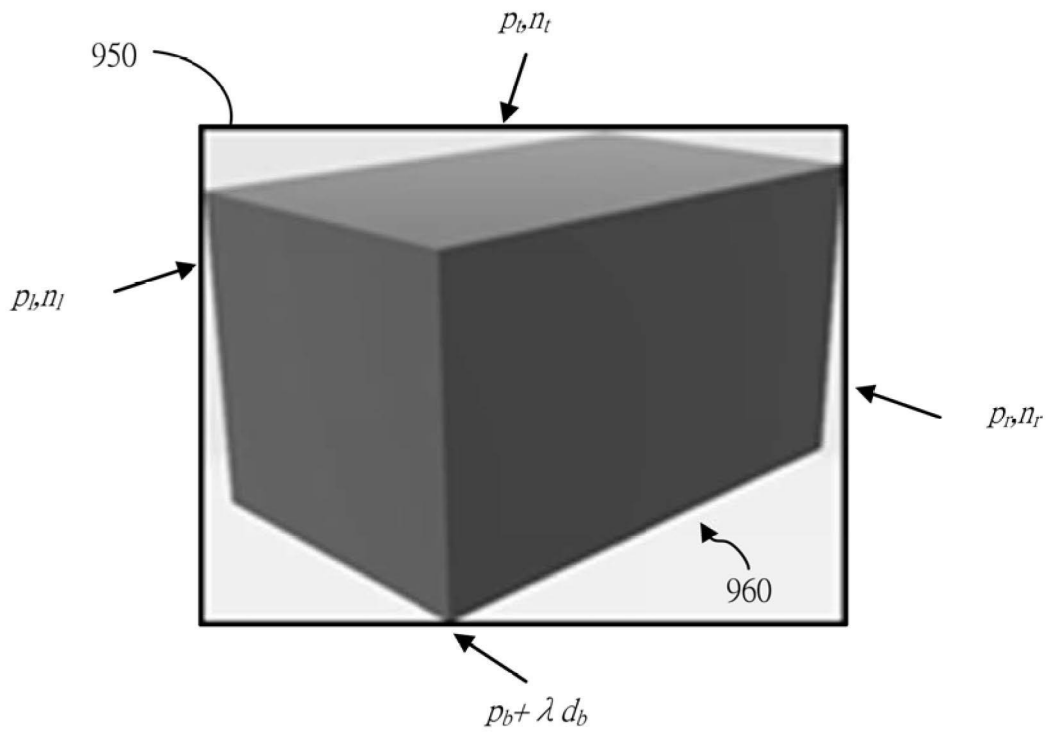


圖9A

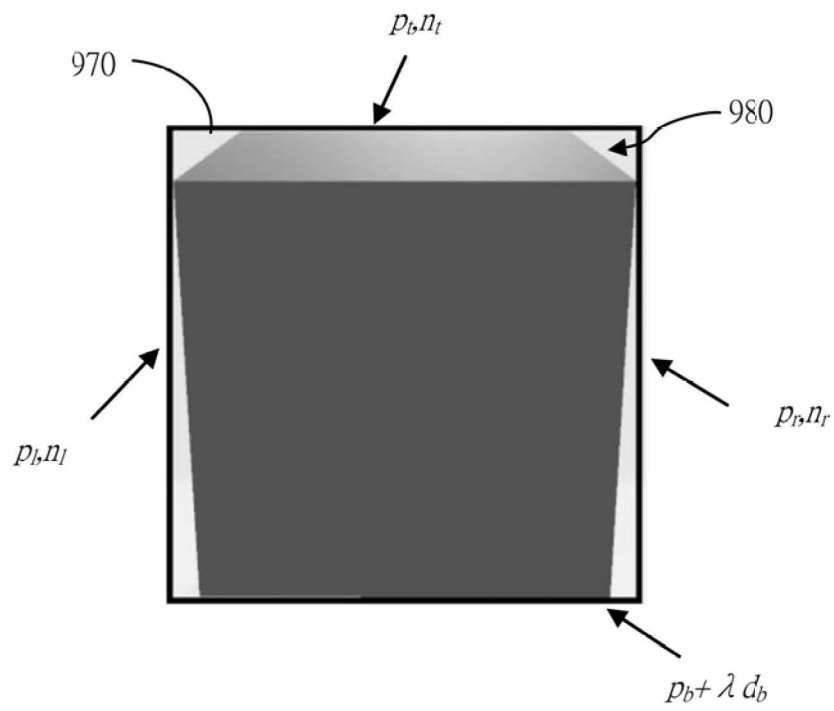


圖9B

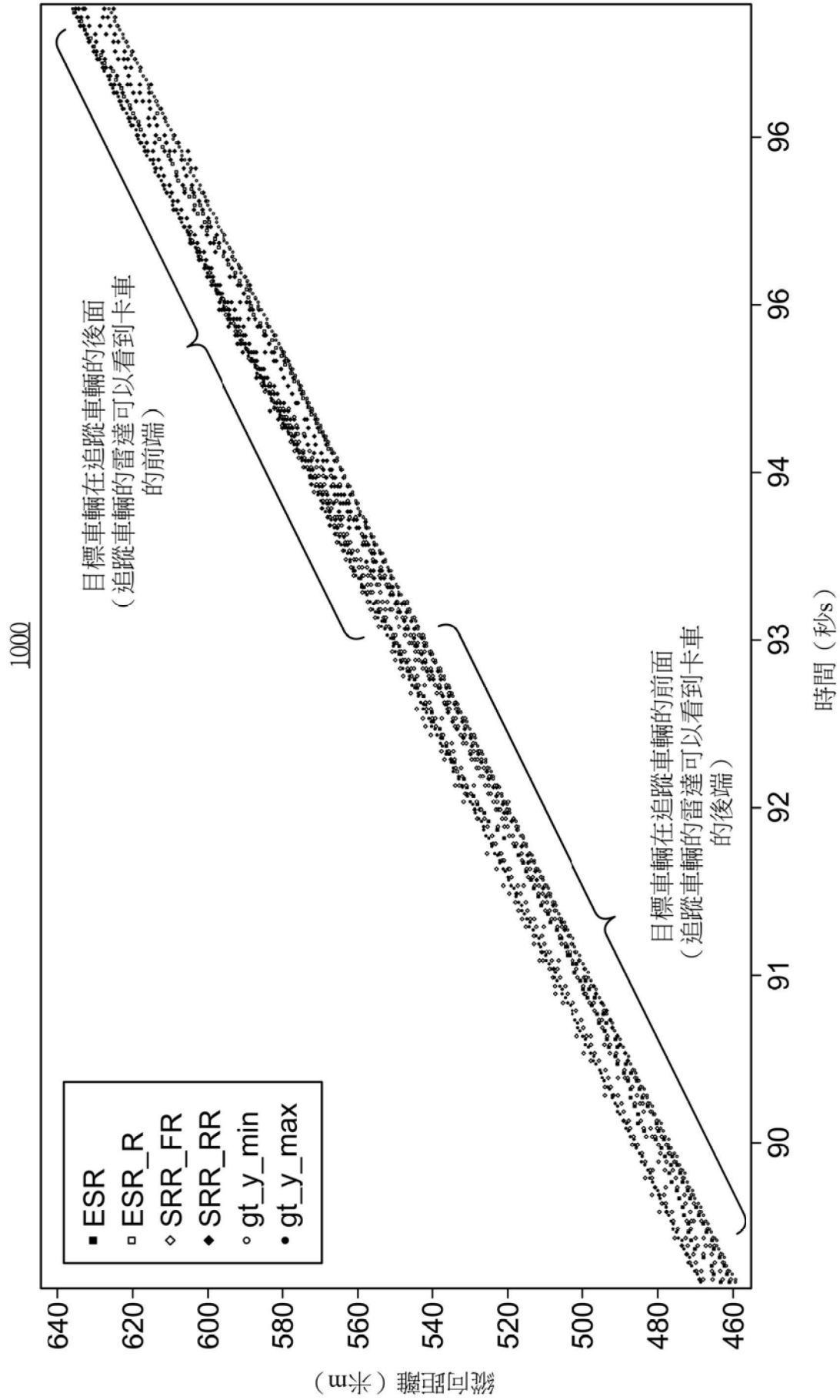


圖10

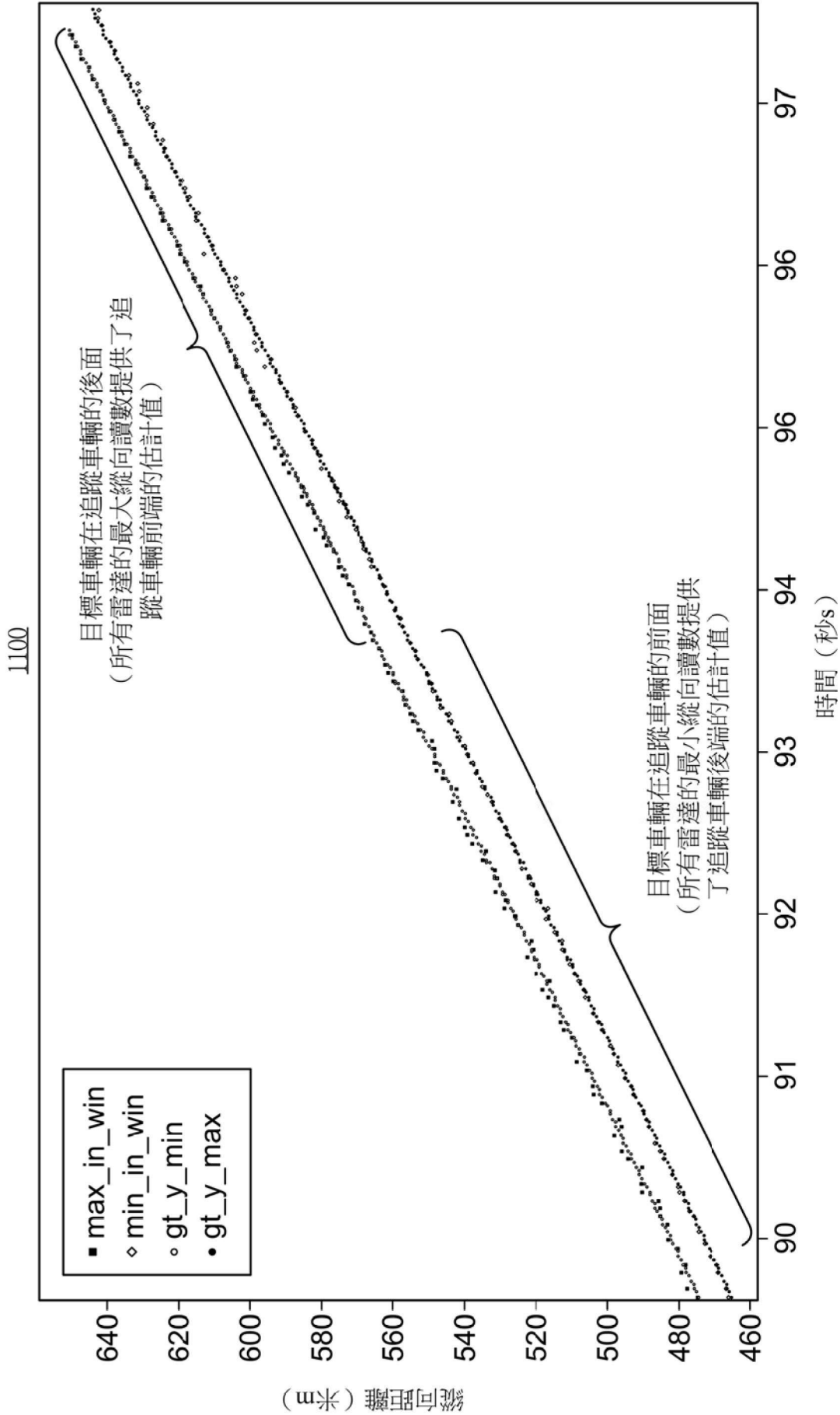


圖11

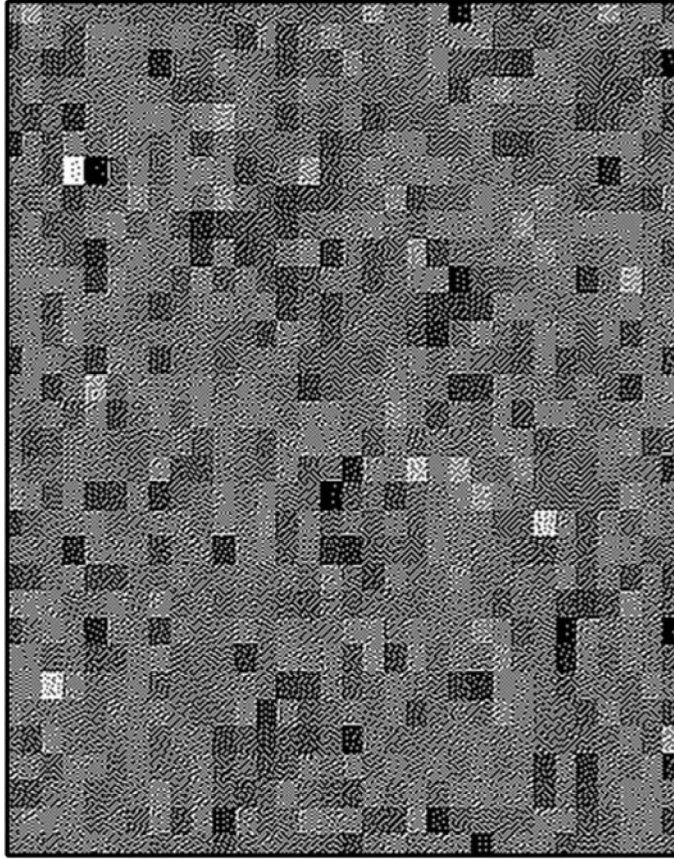


圖12B

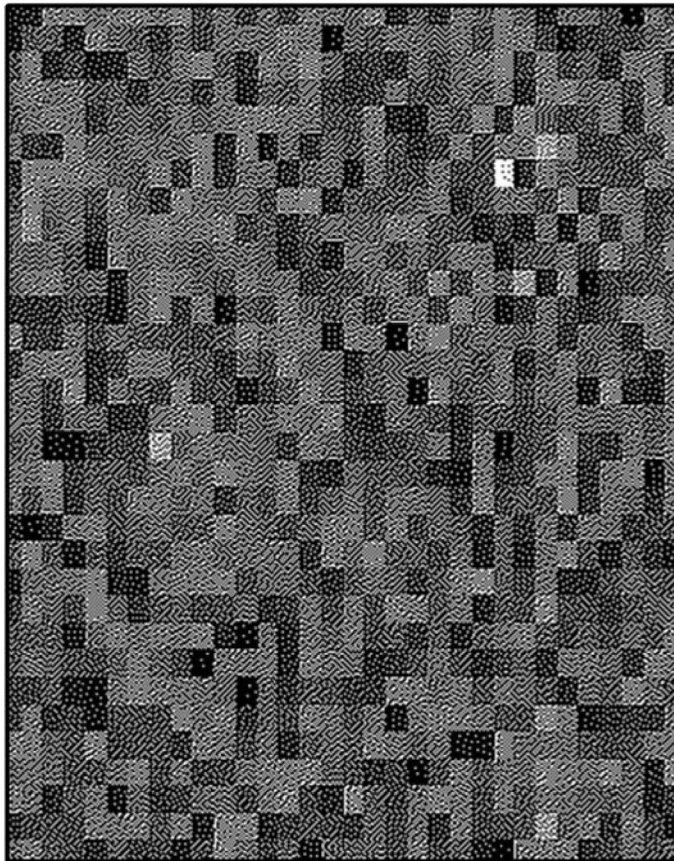


圖12A

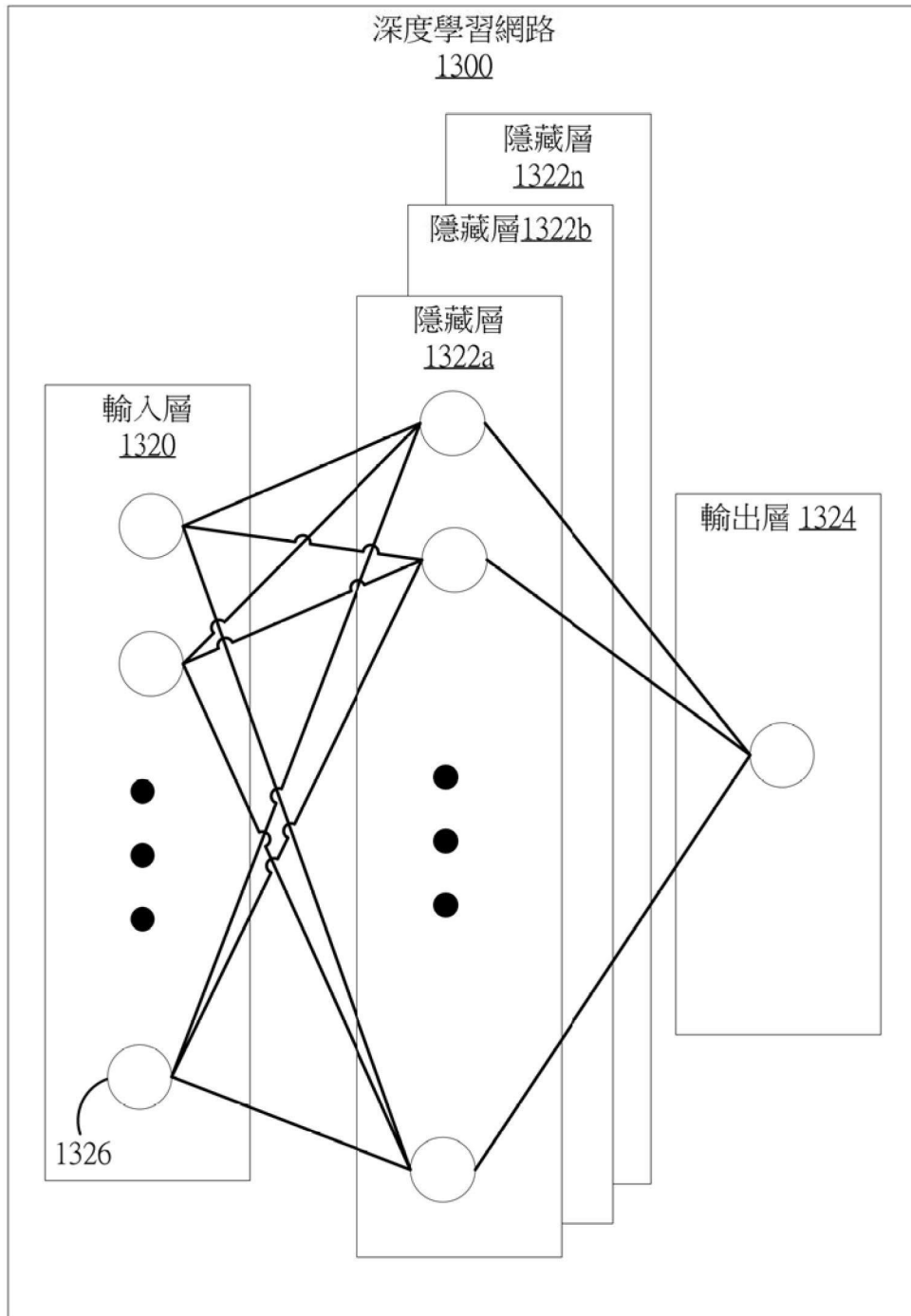


圖13

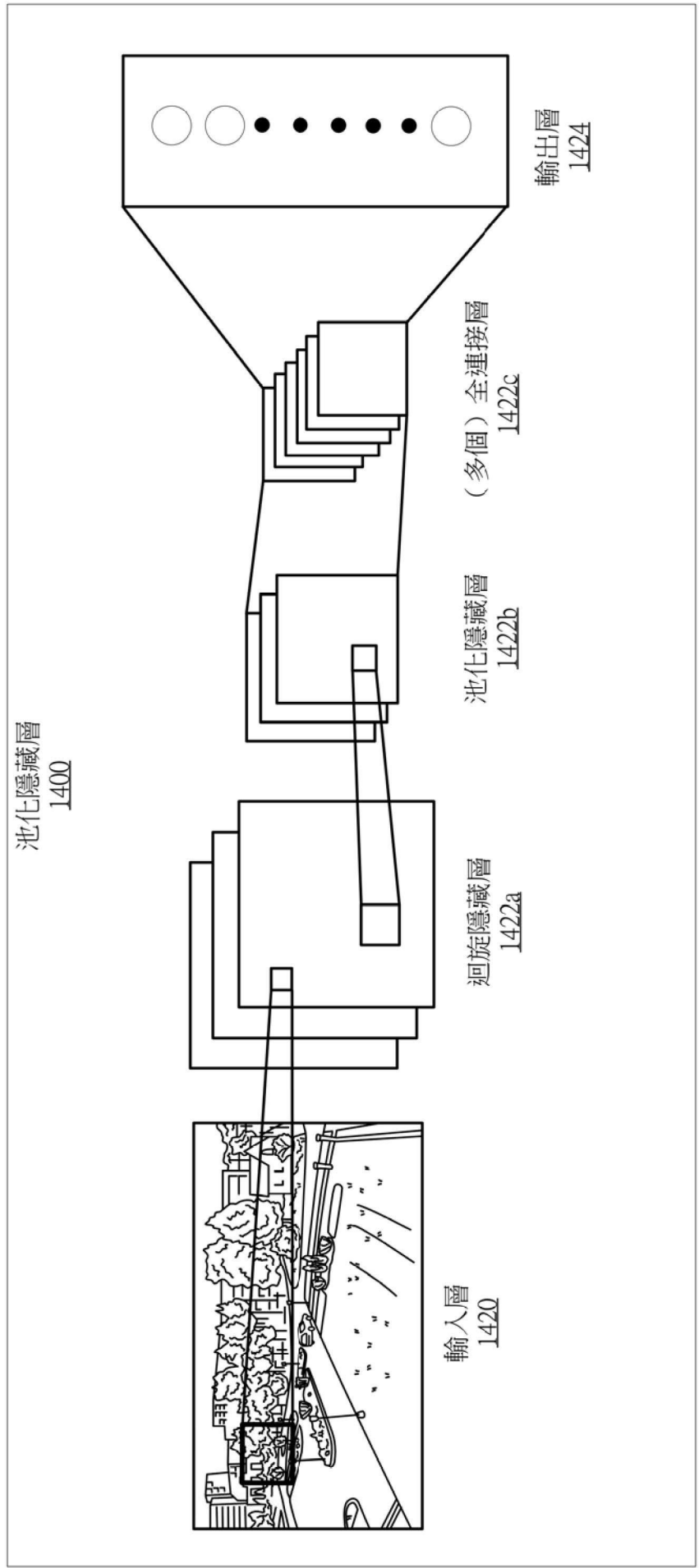


圖14

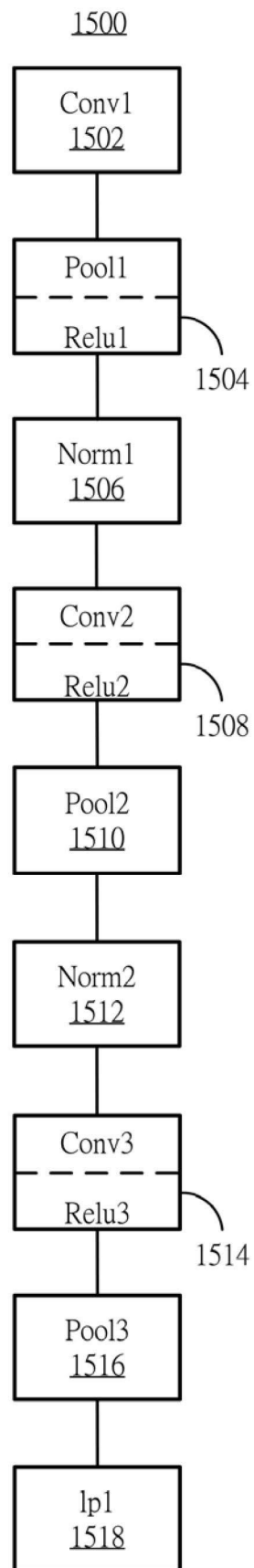
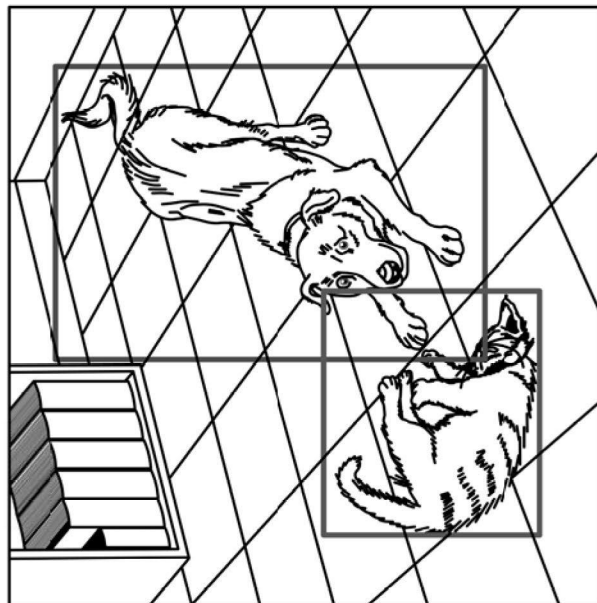
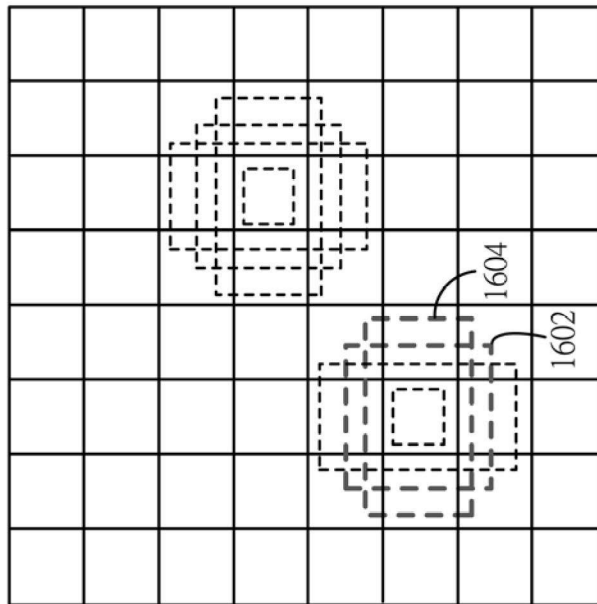


圖15



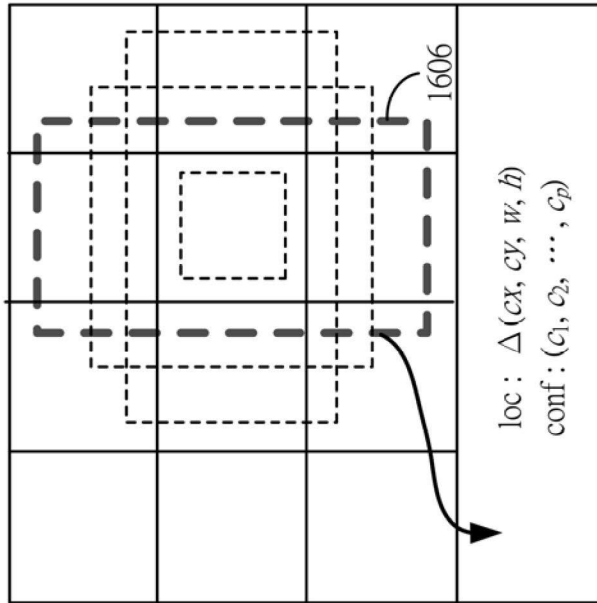
具有GT框的圖像

圖16A



8x8特徵映射

圖16B



4x4特徵映射

圖16C

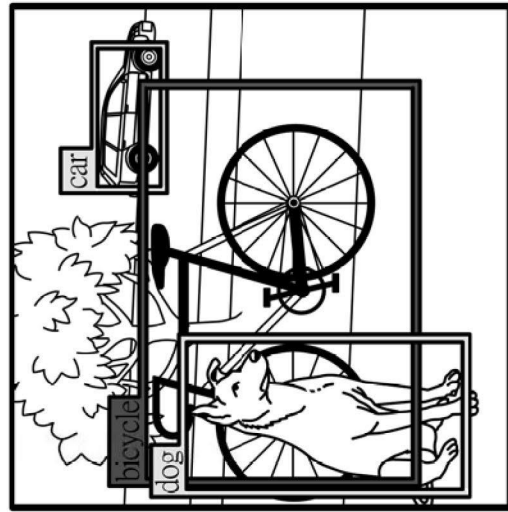


圖17C

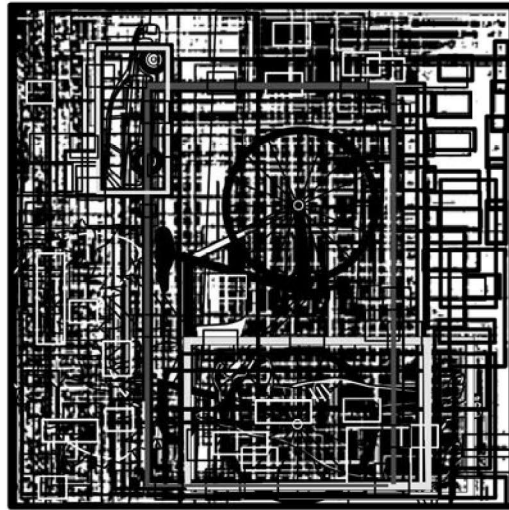


圖17B

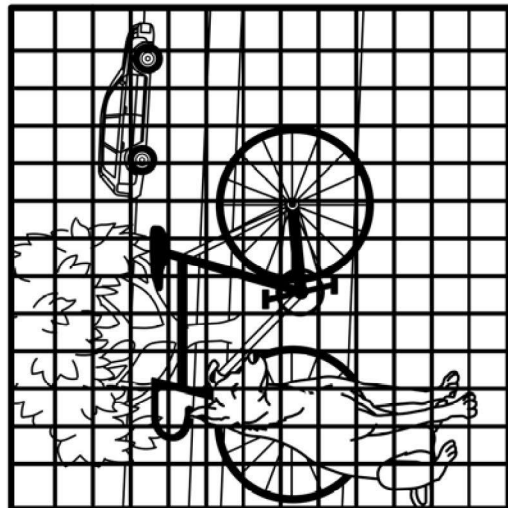


圖17A

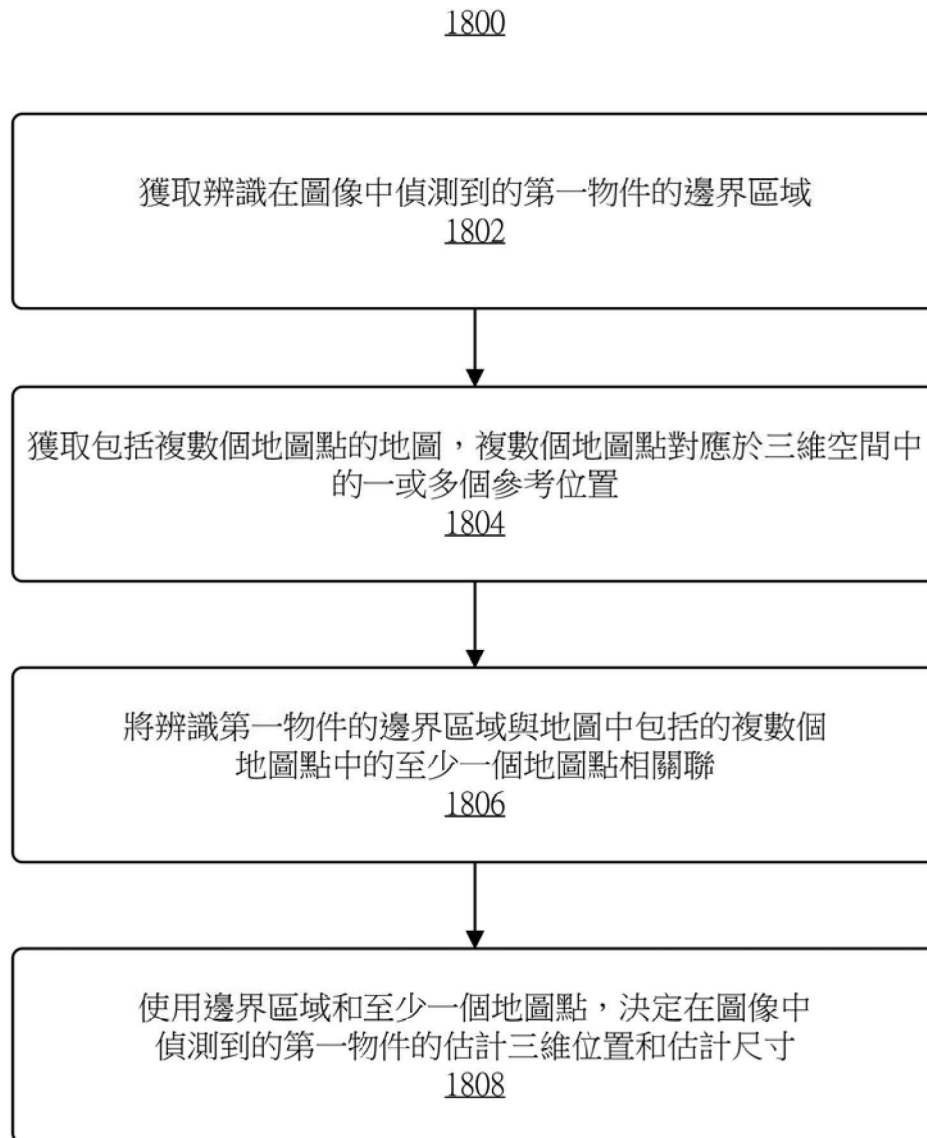


圖18

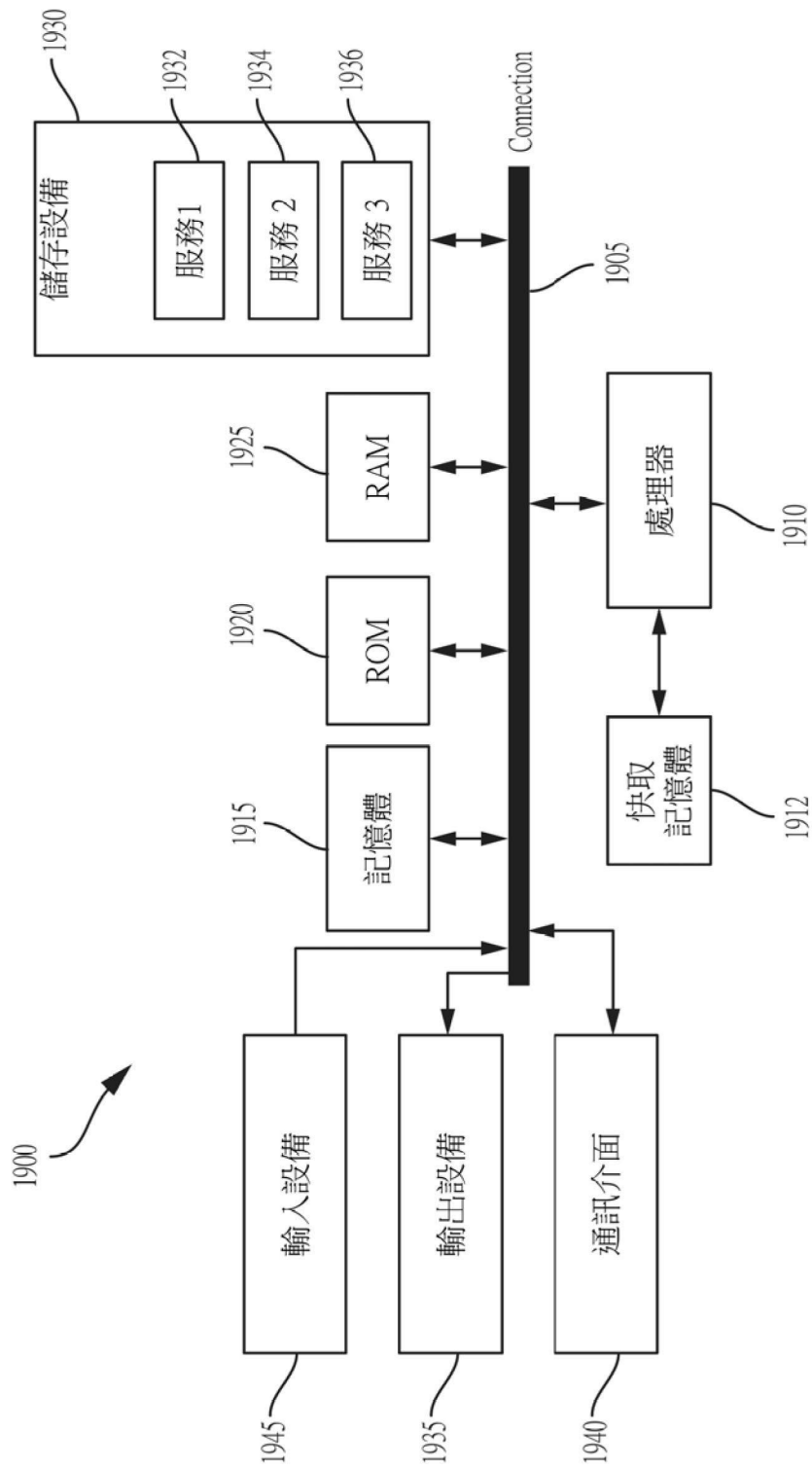


圖19