



(19)中華民國智慧財產局

(12)發明說明書公告本

(11)證書號數：TW I571099 B

(45)公告日：中華民國 106 (2017) 年 02 月 11 日

(21)申請案號：104139474

(22)申請日：中華民國 104 (2015) 年 11 月 26 日

(51)Int. Cl. : *H04N13/00 (2006.01)**G06T7/00 (2006.01)*

(30)優先權：2015/10/16 美國

62/242,332

(71)申請人：財團法人工業技術研究院 (中華民國) INDUSTRIAL TECHNOLOGY RESEARCH INSTITUTE (TW)

新竹縣竹東鎮中興路 4 段 195 號

(72)發明人：吳俊德 WU, CHUN TE (TW)

(74)代理人：祁明輝；林素華；涂綺玲

(56)參考文獻：

TW 201241789

審查人員：陳哲賢

申請專利範圍項數：12 項 圖式數：6 共 21 頁

(54)名稱

深度估測裝置及方法

DEVICE AND METHOD FOR DEPTH ESTIMATION

(57)摘要

一種深度估測裝置，包括：立體影像擷取單元、主動式深度感測器以及處理器。立體影像擷取單元提供第一視角影像以及第二視角影像。主動式深度感測器提供深度資料。處理器執行深度感測處理，深度感測處理包括：對深度資料進行空間轉換以產生轉換後深度資料；對第一視角影像與第二視角影像執行特徵搜尋運算，以決定多組特徵組，各特徵組包括分別在第一視角影像與第二視角影像上對應於同一物件之兩特徵點；利用特徵點，與轉換後深度資料執行內插運算，以針對第一視角影像與第二視角影像決定像素位移搜尋範圍；以及在像素位移搜尋範圍內，搜尋第一視角影像與第二視角影像匹配時的像素位移量，以產生視差影像。

A depth estimation device includes a stereo image capture unit, an active depth sensor and a processor. The stereo image capture unit provides a first view image and a second view image. The active depth sensor provides depth data. The processor performs depth estimation process including steps of: performing space transform on the depth data to generate transformed depth data; performing feature search algorithm to determine a plurality of feature pairs, each feature pair including two feature points respectively located on the first and second view images and corresponding to the same object; performing interpolation on the transformed depth data and the feature points to determine a search range for pixel shift; and searching the search range for pixel shift values that the first view image matches the second view image to generate a disparity image.

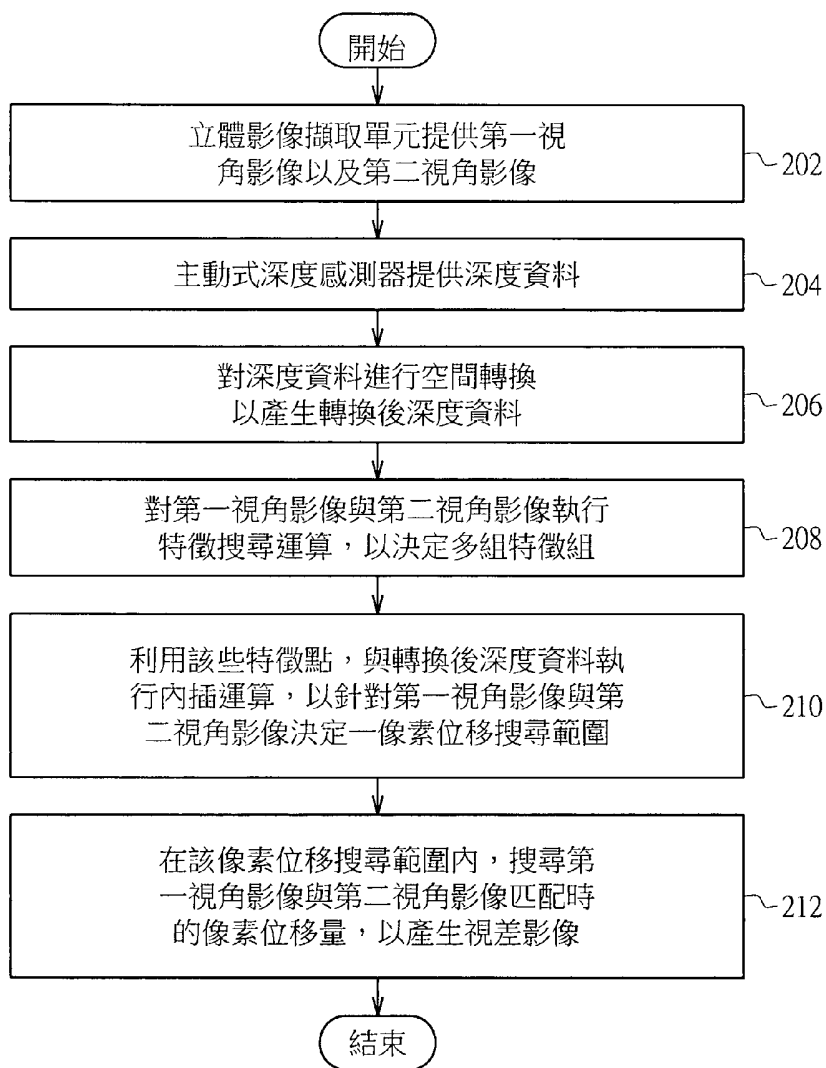
指定代表圖：

符號簡單說明：

202、204、206、

208、210、

212 . . . 步驟



第 2 圖

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

深度估測裝置及方法 /DEVICE AND METHOD FOR DEPTH ESTIMATION

【技術領域】

【0001】 本揭露是有關於一種深度估測裝置及方法。

【先前技術】

【0002】 近年來，深度估測技術被廣泛地應用，諸如車用距離感測、立體成像等等。為了精準地測得與環境物體間的距離(深度)，傳統上會利用光達進行環場式的掃瞄以取得深度資訊。然而，低解析度的光達往往不易偵測到環境中的細節，而高掃描解析度的光達卻又價格不斐，使得商品化並不容易。

【0003】 因此，如何提供一種深度估測裝置及方法，可在提升深度估測效率的同時亦能降低設置成本，乃目前業界所致力的課題之一。

【發明內容】

【0004】 本發明是有關於一種深度估測裝置及方法，可利用雙視角影像之特徵點來提升深度估測效率，並能降低深度感測器的設置成本。

【0005】 根據本揭露一方面，提出一種深度估測裝置，其包括：立體影像擷取單元、主動式深度感測器以及處理器。立體影

像擷取單元提供第一視角影像以及第二視角影像。主動式深度感測器提供深度資料。處理器執行深度感測處理，深度感測處理包括：對深度資料進行空間轉換以產生轉換後深度資料；對第一視角影像與第二視角影像執行特徵搜尋運算，以決定多組特徵組，各特徵組包括分別在第一視角影像與第二視角影像上對應於同一物件之兩特徵點；利用特徵點，與轉換後深度資料執行內插運算，以針對第一視角影像與第二視角影像決定像素位移搜尋範圍；以及在像素位移搜尋範圍內，搜尋第一視角影像與第二視角影像匹配時的像素位移量，以產生視差影像。

【0006】 根據本揭露一方面，提出一種深度估測方法，其包括以下步驟：自立體影像擷取單元取得第一視角影像以及第二視角影像；自主動式深度感測器取得深度資料；對深度資料進行空間轉換以產生轉換後深度資料；對第一視角影像與第二視角影像執行特徵搜尋運算，以決定多組特徵組，各特徵組包括分別在第一視角影像與第二視角影像上對應於同一物件之兩特徵點；利用特徵點，與轉換後深度資料執行內插運算，以針對第一視角影像與第二視角影像決定一像素位移搜尋範圍；以及在像素位移搜尋範圍內，搜尋第一視角影像與第二視角影像匹配時的像素位移量，以產生視差影像。

【0007】 為了對本揭露之上述及其他方面有更佳的瞭解，下文特舉較佳實施例，並配合所附圖式，作詳細說明如下：

【圖式簡單說明】

【0008】

第 1 圖繪示依據本揭露之一實施例之深度估測裝置之簡化方塊圖。

第 2 圖繪示依據本揭露之一實施例之深度估測方法之流程圖。

第 3 圖繪示依據本揭露之一實施例之第一視角影像與第二視角影像之示意圖。

第 4 圖繪示依據特徵組之兩特徵點間在一方向上的像素位移量是否超出一閾值來判斷是否篩除該特徵組之示意圖。

第 5 圖繪示依據本揭露之一實施例之像素位移搜尋範圍之示意圖。

第 6A 圖及第 6B 圖繪示依據本揭露不同實施例之立體影像擷取單元與主動式深度感測器之配置圖。

【實施方式】

【0009】 在本文中，參照所附圖式仔細地描述本揭露的一些實施例，但不是所有實施例都有表示在圖示中。實際上，這些揭露可使用多種不同的變形，且並不限於本文中的實施例。相對的，本揭露提供這些實施例以滿足應用的法定要求。圖式中相同的參考符號用來表示相同或相似的元件。

【0010】 請參考第 1 圖及第 2 圖。第 1 圖繪示依據本揭露之一實施例之深度估測裝置 100 之簡化方塊圖。第 2 圖繪示依據本揭露之一實施例之深度估測方法之流程圖。深度估測裝置 100 主要包括立體影像擷取單元 102、主動式深度感測器 104 以及處理

器 106。立體影像擷取單元 102 例如是具備雙鏡頭的立體相機 (stereo camera)。主動式深度感測器 104 例如是光達 (Lidar)、雷達或其它可透過主動式掃描以獲取深度資料的感測器。

【0011】 在步驟 202，立體影像擷取單元 102 提供第一視角影像 F1 以及第二視角影像 F2。第一視角影像 F1 以及第二視角影像 F2 例如是在不同視角下針對一場景所擷取的影像。如第 1 圖所示，立體影像擷取單元 102 包括第一攝像鏡頭 1022 以及第二攝像鏡頭 1024，分別用以拍攝不同視角的第一視角影像 F1 以及第二視角影像 F2。第一視角影像 F1 以及第二視角影像 F2 為經過影像校正 (image rectification) 之影像。在一實施例中，第一攝像鏡頭 1022 以及第二攝像鏡頭 1024 係沿著一水平軸並列設置，此時第一視角影像 F1 以及第二視角影像 F2 相當於一組左、右眼視角影像。然本揭露並不限於此，在一實施例中，立體影像擷取單元 102 可包括多個攝像鏡頭 (例如包括兩個以上的攝像鏡頭)，以分別擷取不同視角的視角影像。

【0012】 在步驟 204，主動式深度感測器 104 提供深度資料 DD。舉例來說，主動式深度感測器 104 可發射掃描波並依據偵測到的回波來估測與被掃描物間的距離，進而產生相應的深度資料 DD。

【0013】 在一實施例中，處理器 106 可依據第一、二視角影像 F1、F2 以及深度資料 DD 執行深度感測處理並產生視差影像 I。所述深度感測處理例如步驟 206~212 所示。

【0014】 在步驟 206，係對深度資料 DD 進行空間轉換(space transform)以產生轉換後深度資料。轉換後深度資料係與第一視角影像 F1 或第二視角影像 F2 的視角一致。空間轉換例如是一座標系轉換，用以將深度資料 DD 所對應的座標系轉換至與第一視角影像 F1 或第二視角影像 F2 相同的座標系。

【0015】 在步驟 208，係對第一視角影像 F1 與第二視角影像 F2 執行特徵搜尋運算，以決定多組特徵組。其中，各特徵組包括分別在第一視角影像 F1 與第二視角影像 F2 上對應於同一物件之兩特徵點。特徵搜尋運算例如包括尺度不變特徵轉換(Scale Invariant Feature Transform, SIFT),加速穩固特徵(Speeded Up Robust Features, SUFT)或 FAST 特徵點檢測等演算法。

【0016】 在步驟 210，係利用該些特徵點，與轉換後深度資料執行內插運算，以針對第一視角影像 F1 與第二視角影像 F2 決定一像素位移搜尋範圍。

【0017】 上述之內插運算例如是近鄰內插(Nearest Neighbor Interpolation)、線性內插(Linear Interpolation)、三次內插(Cubic Interpolation)、多項式內插(Polynomial Interpolation)或其它保留邊緣的演算法(Edge-Aware Data Interpolation)，像是雙邊濾波器(bilateral filter)或雙邊濾波器柵格(bilateral grid)演算法。

【0018】 一般來說，主動式深度感測器 104(如光達)會透過發射多條掃描線進行環場掃描以偵測與環境物體的距離。然該些掃描線所構成的掃描區域係離散的(一掃描線例如對應一特定高度

的水平面)，故仍需透過內差運算以產生完整場景的深度資訊。本揭露實施例之處理器 106 在運算的過程中可將特徵點的資訊對視為已知解再對轉換後深度資料進行內插，故可改善內插運算所造成的誤差，進而降低對主動式深度感測器 104 掃描解析度的要求。

【0019】 在步驟 212，係在該像素位移搜尋範圍內，搜尋第一視角影像 F1 與第二視角影像 F2 匹配時的像素位移量，以產生視差影像 I。由於處理器 106 只需在步驟 210 中所建立的像素位移搜尋範圍內搜尋第一視角影像 F1 與第二視角影像 F2 之間的像素位移量，故可有效降低產生視差影像 I 的運算量。

【0020】 須注意，上述各步驟之執行順序並不以第 2 圖之方塊流程為限。舉例來說，步驟 202、204 之執行順序可以前後互換，亦可同步執行，步驟 206、208 也是一樣。

【0021】 第 3 圖繪示依據本揭露之一實施例之第一視角影像 F1 與第二視角影像 F2 之示意圖。在第 3 圖的例子中，針對物體 OB 在不同視角下進行拍攝可分別獲得第一視角影像 F1 與第二視角影像 F2。由於第一視角影像 F1 與第二視角影像 F2 所對應的視角不同，故兩影像中對應相同物體 OB 的位置會產生偏移，也就是像素位移。一般來說，當物體 OB 距離越遠，兩影像間的像素位移量越小；反之，當物體 OB 距離越近，兩影像間的像素位移量越大。

【0022】 被拍攝的物體 OB 在第一視角影像 F1 中係對應物件 OB1，在第二視角影像 F2 中係對應物件 OB2。透過特徵搜尋運算

(如 SIFT 或 SUFT)，可找出物件 OB1 的特徵點 $P1\sim P3$ 以及物件 OB2 的特徵點 $P1'\sim P3'$ 。其中，對應物體 OB 同一點 A1 的兩特徵點 P1 及 P1' 可視為一特徵組；對應物體 OB 同一點 A2 的兩特徵點 P2 及 P2' 可視為一特徵組；對應物體 OB 同一點 A3 的兩特徵點 P3 及 P3' 可視為一特徵組。相較於平坦/無特徵的畫面(如一面白牆或藍天)，特徵點具有明顯、穩固(robust)的性質，故基於特徵點所進行的深度計算較不易發生錯誤。如第 2 圖所示之步驟 210，可利用特徵點資訊來修正內插結果，以降低對主動式深度感測器 104 掃描解析度的要求，進而降低感測器設置成本。

【0023】 在一實施例中，處理器 106 所執行的深度感測處理更包括：判斷各特徵組之兩特徵點間沿著一軸線的一距離差異是否大於一閾值；以及篩除該距離差異大於該閾值的特徵組。

【0024】 舉例來說，若第一視角影像 F1 與第二視角影像 F2 經過影像校正後，係分別對應左眼視角影像與右眼視角影像，原則上左、右眼視角影像的特徵點會在同一水平(或同一軸線)上，此時分別在兩視角影像中的相對應兩特徵點間的像素位移量主要應反映在水平方向，垂直方向的距離差異不應過大。在此情況下，篩除沿垂直方向距離差異過大的特徵組，使之排除於後續的內差運算中，可進一步降低深度計算錯誤的機會。

【0025】 如第 4 圖所示，假設用以提供第一、二視角影像 F1、F2 的第一、二攝像鏡頭 1022、1024 之鏡心連線方向為 x 軸方向，當一特徵組之兩特徵點(如 P1、P1')間沿著 y 軸方向的距離差異大

於一閾值 TH，則此特徵組將被篩除，也就是被排除於後續的內插運算中，以避免採用錯誤的特徵組進行計算。

【0026】 第 5 圖繪示依據本揭露之一實施例之像素位移搜尋範圍 SSR 之示意圖。在本揭露實施例中，處理器 106 可對主動式深度感測器 104 所提供的深度資料 DD 進行空間轉換(space transform)，並利用自第一、二視角影像 F1、F2 取得的特徵點，對轉換後的深度資料(例如視差資料)執行內插運算以產生多筆像素位移參考值，再以此些像素位移參考值為基準，建立像素位移搜尋範圍。

【0027】 如第 5 圖所示，若位在座標 X_0 的像素 P 所對應的像素位移參考值為 SR，可估測其對應像素 P' 係位在另一視角影像中同一 y 軸上的座標 X_0+SR 附近，例如位在座標 $(X_0+SR)-X_1$ 到座標 $(X_0+SR)+X_2$ 的區間內。座標 $(X_0+SR)-X_1$ 到座標 $(X_0+SR)+X_2$ 的區間可定義出像素位移搜尋範圍 SSR。此時，處理器 106 只需針對像素位移搜尋範圍 SSR 進行搜尋即可找出像素 P 與其對應像素 P' 之間的實際像素位移量，而不用逐一比對所有的可能值，故可有效降低深度估測的運算量。在一實施例中，若內插後的轉換後深度資料的深度值變化較大，可適當放大 X_1 與 X_2 的值，以調整像素位移搜尋範圍 SSR 的範圍。也就是說，可依據內插後的轉換後深度資料的深度值變化程度，來調整像素位移搜尋範圍 SSR 的大小。

【0028】 第 6A 圖及第 6B 圖繪示依據本揭露不同實施例之立

體影像擷取單元 102 與主動式深度感測器 104 之配置圖。在第 6A 圖的例子中，立體影像擷取單元 102 之第一、二攝像鏡頭 1022、1024 的鏡心連線 AX 係與主動式深度感測器 104 之掃描線 SL 方向約略平行/實質上平行。

【0029】 在另一實施例中，如第 6B 圖所示，立體影像擷取單元 102 之第一、二攝像鏡頭 1022、1024 的鏡心連線 AX 係與主動式深度感測器 104 之掃描線 SL 方向非平行(例如，約略垂直/實質上垂直)。由於第一、二攝像鏡頭 1022、1024 所提供的第一、二視角影像 F1、F2 係基於鏡心連線 AX 方向(例如水平方向)的像素位移量來判斷對應的視差/深度，故可補足主動式深度感測器 104 基於另一方向(例如垂直方向)掃描之精度。

【0030】 綜上所述，本揭露所提供之深度估測裝置及方法可基於不同視角影像所萃取出之特徵點資訊，對主動式深度感測器提供的深度資料進行內插運算，不僅可改善內插運算的誤差，更能降低對主動式深度感測器掃描解析度的要求。

【0031】 雖然本揭露已以較佳實施例揭露如上，然其並非用以限定本揭露。本揭露所屬技術領域中具有通常知識者，在不脫離本揭露之精神和範圍內，當可作各種之更動與潤飾。因此，本揭露之保護範圍當視後附之申請專利範圍所界定者為準。

【符號說明】

【0032】

100：深度估測裝置

102：立體影像擷取單元

1022：第一攝像鏡頭

1024：第二攝像鏡頭

104：主動式深度感測器

106：處理器

F1：第一視角影像

F2：第二視角影像

DD：深度資料

I：視差影像

202、204、206、208、210、212：步驟

OB：物體

OB1、OB2：物件

P1~P3、P1'~P3'：特徵點

A1~A3：點

TH：閥值

P、P'：像素

X_0 、 $(X_0+SR)-X_1$ 、 X_0+SR 、 $(X_0+SR)+X_2$ ：座標

SSR：像素位移搜尋範圍

AX：鏡心連線

SL：掃描線

x、y、z：座標軸

發明摘要



※ 申請案號： 104139474

H04N 13/00 (2006.01)

※ 申請日：104. 11. 28

※IPC 分類：G06T 1/00 (2006.01)

【發明名稱】(中文/英文)

深度估測裝置及方法 /DEVICE AND METHOD FOR DEPTH ESTIMATION

【中文】

一種深度估測裝置，包括：立體影像擷取單元、主動式深度感測器以及處理器。立體影像擷取單元提供第一視角影像以及第二視角影像。主動式深度感測器提供深度資料。處理器執行深度感測處理，深度感測處理包括：對深度資料進行空間轉換以產生轉換後深度資料；對第一視角影像與第二視角影像執行特徵搜尋運算，以決定多組特徵組，各特徵組包括分別在第一視角影像與第二視角影像上對應於同一物件之兩特徵點；利用特徵點，與轉換後深度資料執行內插運算，以針對第一視角影像與第二視角影像決定像素位移搜尋範圍；以及在像素位移搜尋範圍內，搜尋第一視角影像與第二視角影像匹配時的像素位移量，以產生視差影像。

【英文】

A depth estimation device includes a stereo image capture unit,

an active depth sensor and a processor. The stereo image capture unit provides a first view image and a second view image. The active depth sensor provides depth data. The processor performs depth estimation process including steps of: performing space transform on the depth data to generate transformed depth data; performing feature search algorithm to determine a plurality of feature pairs, each feature pair including two feature points respectively located on the first and second view images and corresponding to the same object; performing interpolation on the transformed depth data and the feature points to determine a search range for pixel shift; and searching the search range for pixel shift values that the first view image matches the second view image to generate a disparity image.

【代表圖】

【本案指定代表圖】：第（ 2 ）圖。

【本代表圖之符號簡單說明】：

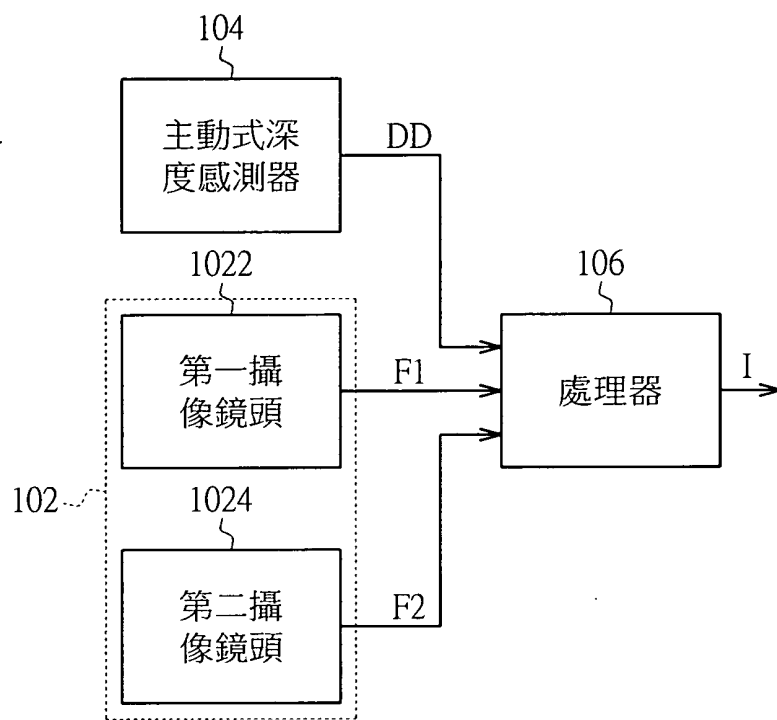
202、204、206、208、210、212：步驟

【本案若有化學式時，請揭示最能顯示揭露特徵的化學式】：

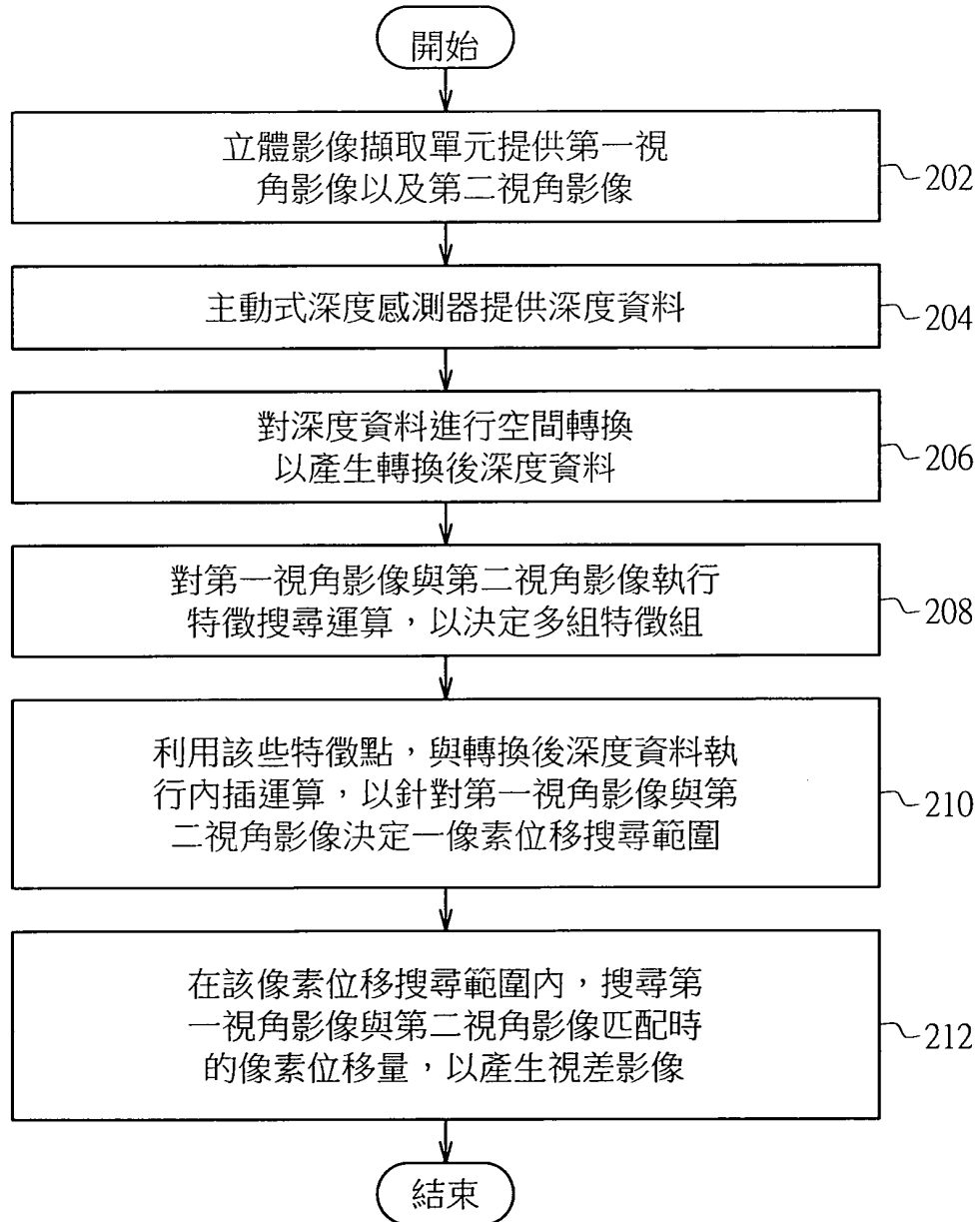
無

圖式

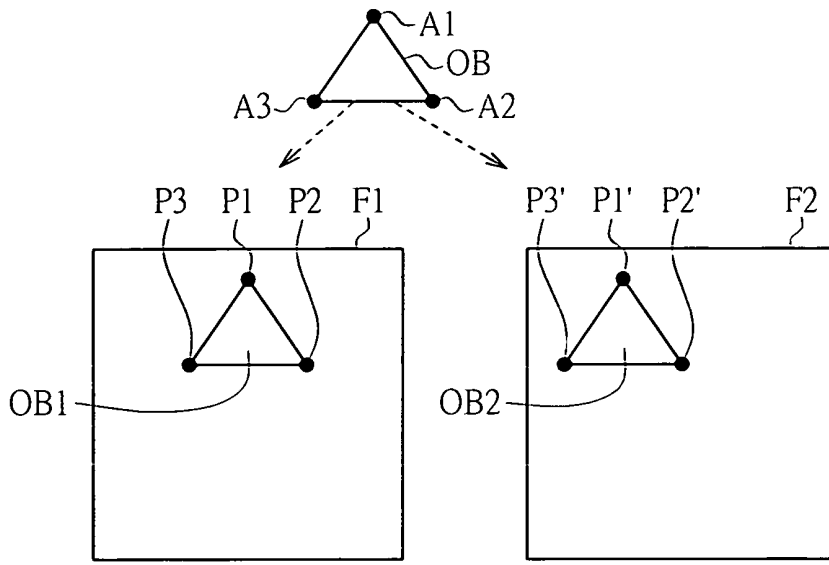
100



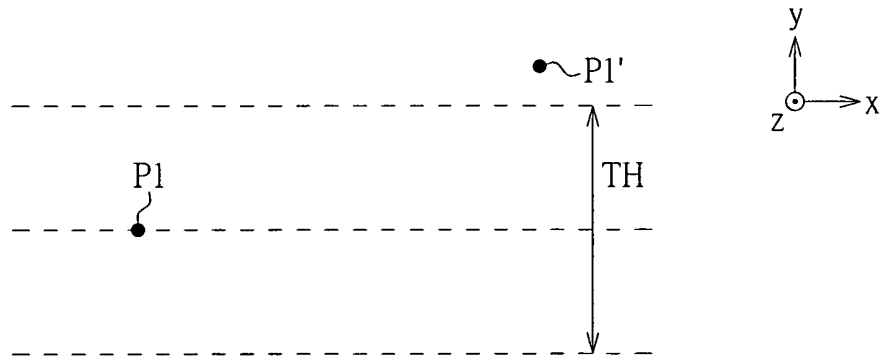
第 1 圖



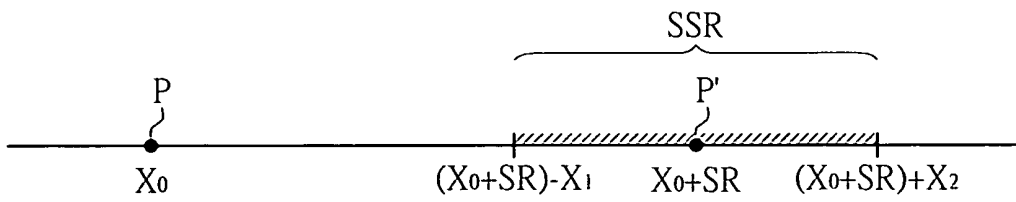
第 2 圖



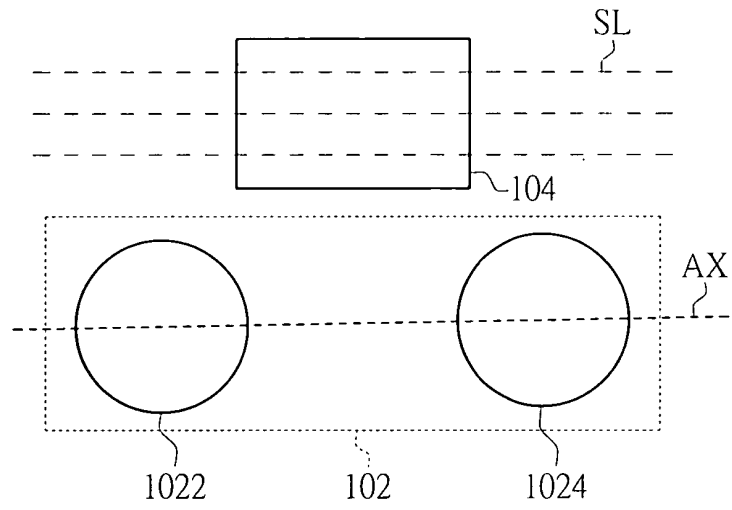
第 3 圖



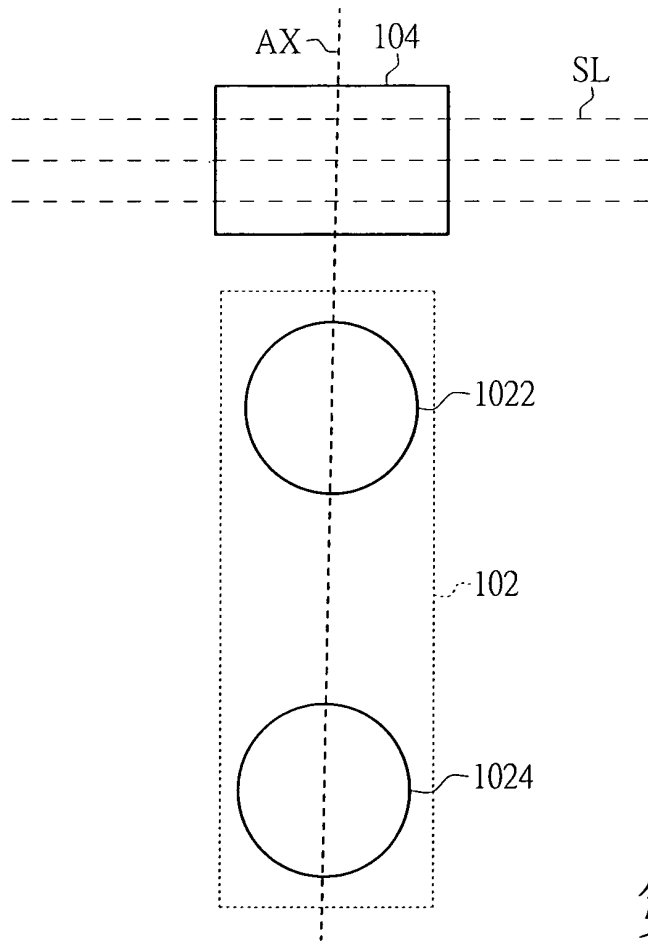
第 4 圖



第 5 圖



第 6A 圖



第 6B 圖

an active depth sensor and a processor. The stereo image capture unit provides a first view image and a second view image. The active depth sensor provides depth data. The processor performs depth estimation process including steps of: performing space transform on the depth data to generate transformed depth data; performing feature search algorithm to determine a plurality of feature pairs, each feature pair including two feature points respectively located on the first and second view images and corresponding to the same object; performing interpolation on the transformed depth data and the feature points to determine a search range for pixel shift; and searching the search range for pixel shift values that the first view image matches the second view image to generate a disparity image.

【代表圖】

【本案指定代表圖】：第（ 2 ）圖。

【本代表圖之符號簡單說明】：

202、204、206、208、210、212：步驟

【本案若有化學式時，請揭示最能顯示揭露特徵的化學式】：

無

申請專利範圍

1. 一種深度估測裝置，包括：

一立體影像擷取單元，用以提供一第一視角影像以及一第二視角影像，其中該第一視角影像以及該第二視角影像是在不同視角下針對一場景所擷取的影像；

一主動式深度感測器，用以提供一深度資料；以及

一處理器，用以執行一深度感測處理，該深度感測處理包括：

對該深度資料進行一空間轉換以產生一轉換後深度資料；

對該第一視角影像與該第二視角影像執行一特徵搜尋運算，以決定多組特徵組，各該特徵組包括分別在該第一視角影像與該第二視角影像上對應於同一物件之兩特徵點；

利用該些特徵點，與轉換後深度資料執行一內插運算，以針對該第一視角影像與該第二視角影像決定一像素位移搜尋範圍；以及

在該像素位移搜尋範圍內，搜尋該第一視角影像與該第二視角影像匹配時的像素位移量，以產生一視差影像。

2. 如申請專利範圍第 1 項所述之深度估測裝置，其中該轉換後深度資料與該第一視角影像或該第二視角影像的視角一致。

3. 如申請專利範圍第 1 項所述之深度估測裝置，其中該深度感測處理更包括：

利用該些特徵點，與該轉換後深度資料執行該內插運算以產

生複數筆像素位移參考值；以及

以該些像素位移參考值為基準，建立該像素位移搜尋範圍。

4. 如申請專利範圍第 1 項所述之深度估測裝置，其中該深度感測處理更包括：

判斷各該特徵組之該兩特徵點間沿著一軸線的一位置差異是否大於一閾值；以及

篩除該位置差異大於該閾值的特徵組。

5. 如申請專利範圍第 4 項所述之深度估測裝置，其中該立體影像擷取單元包括一第一攝像鏡頭以及一第二攝像鏡頭，該第一攝像鏡頭以及該第二攝像鏡頭分別用以拍攝該第一視角影像以及該第二視角影像，

其中，該第一攝像鏡頭與該第二攝像鏡頭間的鏡心連線係與該軸線垂直。

6. 如申請專利範圍第 1 項所述之深度估測裝置，其中該立體影像擷取單元包括一第一攝像鏡頭以及一第二攝像鏡頭，該第一攝像鏡頭以及該第二攝像鏡頭分別用以拍攝該第一視角影像以及該第二視角影像；

其中該第一攝像鏡頭與該第二攝像鏡頭間的鏡心連線係與該主動式深度感測器之掃描方向非平行。

7. 如申請專利範圍第 1 項所述之深度估測裝置，其中該深度感測處理更包括：

依據內插後的該轉換後深度資料的深度值變化程度，調整該

像素位移搜尋範圍的大小。

8. 一種深度估測方法，包括：

透過一立體影像擷取單元提供一第一視角影像以及一第二視角影像，其中該第一視角影像以及該第二視角影像是在不同視角下針對一場景所擷取的影像；

透過一主動式深度感測器提供一深度資料；

對該深度資料進行一空間轉換以產生一轉換後深度資料；

對該第一視角影像與該第二視角影像執行一特徵搜尋運算，以決定多組特徵組，各該特徵組包括分別在該第一視角影像與該第二視角影像上對應於同一物件之兩特徵點；

利用該些特徵點，對該轉換後深度資料執行一內插運算，以針對該第一視角影像與該第二視角影像決定一像素位移搜尋範圍；以及

在該像素位移搜尋範圍內，搜尋該第一視角影像與該第二視角影像匹配時的像素位移量，以產生一視差影像。

9. 如申請專利範圍第 8 項所述之深度估測方法，其中該轉換後深度資料與該第一視角影像或該第二視角影像的視角一致。

10. 如申請專利範圍第 8 項所述之深度估測方法，更包括：

利用該些特徵點，對該轉換後深度資料執行該內插運算以產生複數筆像素位移參考值；以及

以該些像素位移參考值為基準，建立該像素位移搜尋範圍。

11. 如申請專利範圍第 8 項所述之深度估測方法，更包括：

判斷各該特徵組之該兩特徵點間沿著一軸線的一位置差異是否大於一閾值；以及

篩除該位置差異大於該閾值的特徵組。

12. 如申請專利範圍第 8 項所述之深度估測方法，更包括：

依據內插後的該轉換後深度資料的深度值變化程度，調整該像素位移搜尋範圍的大小。