

# 2016 年臺灣國際科學展覽會 優勝作品專輯

作品編號 190004

參展科別 電腦科學與資訊工程科

作品名稱 基於擴充實境之室內影像定位導覽系統的設計與研究

得獎獎項 大會獎：三等獎

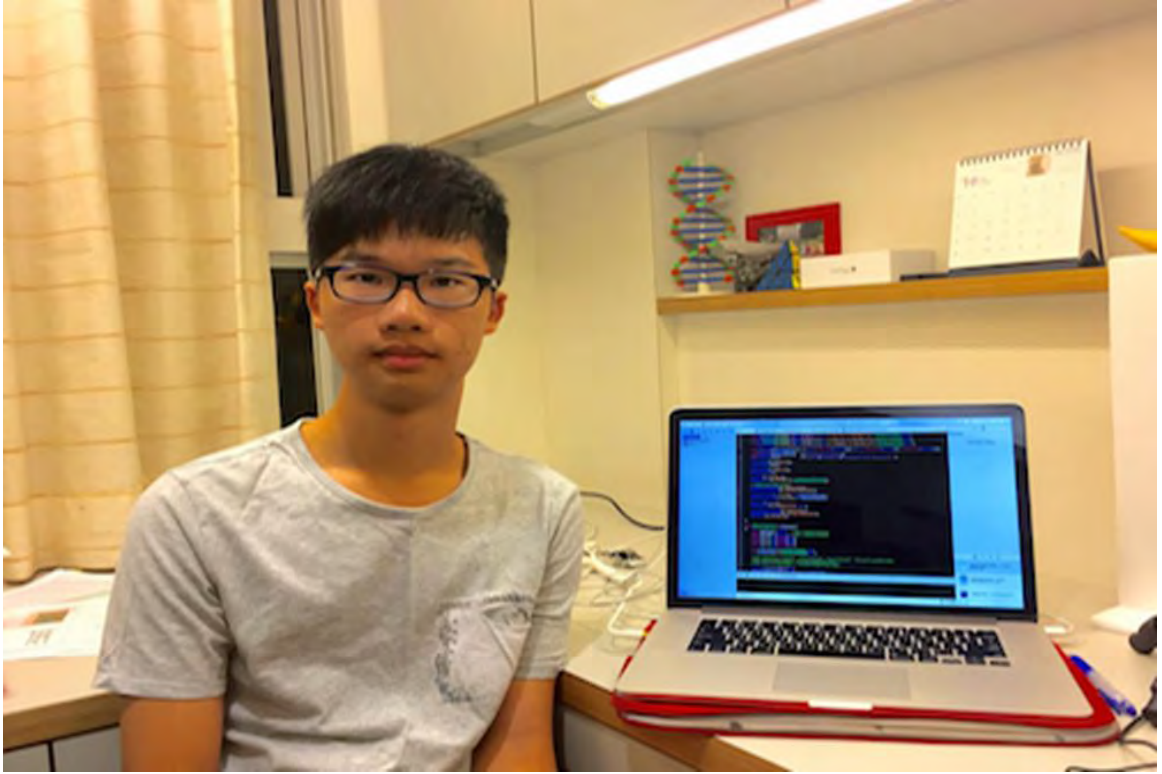
就讀學校 臺北市立建國高級中學

指導教師 王鼎中、莊永裕

作者姓名 陳賢倫

關鍵字 擴充實境、視覺里程計、影像定位系統

## 作者簡介



大家好，我是就讀於建國中學科學班二年級的陳賢倫。在國中就讀數理資優班時與同學一起參加科展，開始了我的研究路。在研究過程中，除了得花費大量時間編寫程式，程式碼中的的錯誤與誤差往往令人感到挫折，但在不斷的努力下，解決了各種困難，自己的能力也在過程中逐漸茁壯。最後，我能完成這一次的研究，要感謝一路上幫助我的各位師長、朋友、以及家人，給予我各方面的支持，在未來，我也將充實自我，在電腦科學的領域為世界貢獻一份心力。

## 摘要

本研究基於擴充實境設置出一套全新的視覺里程計，並以此為基礎建立出一套創新之定位系統。藉由電腦視覺領域的計算方法，本研究藉由計算以參考影像設置的世界座標系與相機座標系之間的轉移矩陣，將其旋轉矩陣與平移向量得出，便可得知相機與參考影像中心再三維世界下  $xyz$  軸的距離差，以及相機所面向的方向，同時利用擴充實境進行虛擬物件的投射，達到提供使用者資訊的效果。此定位系統有別於傳統的定位方法，不需要參考點，僅需得知一影像與其尺寸大小便可建立，無需任何發信器，可被定義為一種全新的定位方式，因此在設置上極為簡易。同時，本研界之定位系統在選用  $30\text{cm} \times 20\text{cm}$  的參考影像時期量測誤差僅於 2 公分以內，每 0.067 秒可完成一張影像的計算，經本研究之評估，可與感測網路、機器人、以及美術館導覽的廣泛領域的未來應用結合，兼具實用性與未來發展性。

## Abstract

This study attempts to create a brand new visual odometry based on the theory of augmented reality and implement this technology to image-based positioning system. The advantages and functions of augmented reality are able to provide the position and pose of an object, while common positioning systems are restricted only to determine the position. By calculating the transformation matrix between world coordinate and camera coordinate, this positioning system obtains the rotation matrix and translation vector between these two coordinates. With these parameters, the xyz distance between camera and the center of pattern image would be acquired. Moreover, as the transformation matrix is obtained, this positioning system would project virtual object to the computer screen by augmented reality to provide additional information to the users for navigation purpose. By implementing this theory, this positioning system is capable of determining the position of objects without any signaling-beacons. In other words, this technology can be defined as a brand new positioning system. After designing the algorithm, this study successfully constructs a program of this image-based positioning system. The advantage of this image-based positioning system is that it requires little setups; achieves an accuracy of the position of camera within 2cm; and calculates each frames in merely 0.067 seconds. The high accuracy and efficiency of this positioning system enable it to be applied to applications in need of an advanced positioning system. This can be a great contribution in navigation of museum and educational exhibition. Furthermore, this image-based positioning system has an overwhelming potential to be combined with various domains such as sensor networking and robotics. In conclusion, this positioning system can determine the position and rotation of an object precisely in real-time and has a feasibility to be combined with diverse domains to dedicate to the world.

# 一、 前言

## (一)、 研究動機

近年來，因電腦視覺領域的發達，各種可以增加生活中視覺多樣性的技術逐漸發展成熟，如擴充實境（Augmented Reality）與虛擬實境（Virtual Reality）等，而在擴充實境技術中，再將虛擬的物件投射到三維坐標系的過程裡的座標系轉換，若將其中的轉換矩陣進行適當的換算，便可得知相機相對於參考物件的三維坐標，而這一部分的技術，有著與透過連續影像已得知相機在三維坐標系中的移動資訊的視覺里程計（Visual Odometry）有著結合的可能性。

然而因科技發展的蓬勃，使定位系統之技術逐漸成長，基於各種不同技術的定位系統也因此產生，如：全球定位系統(GPS)、藍芽(Bluetooth)、磁場等。以上所提出之定位系統被廣泛應用於生活中，將定位所得出之三維空間資料應用於物聯網與導覽系統中。

但是，上述的定位系統在設定上有著繁瑣的步驟，如同需要設置多個參考點等。同時，以上之系統與導覽部分的銜接也相對薄弱，在一般的應用場合，皆需將定位所得出的資料結合另一導覽系統以提供導覽資訊給使用者。因此，這些定位系統皆有著設置上的不便與其他領域之應用結合上的困難。

但若比對現有定位系統的缺點與擴充實境的功能，可以發現擴充實境中，投射虛擬物件的功能，以及計算出使用者的相對座標之可能性，能夠彌補現有定位系統之劣勢，以提供導覽資訊，並建立出可簡易設置的定位系統。基於以上優點，本研究將結合擴充實境之技術，提出一套全新視覺里程計，將其應用於室內定位系統，以改善現有定位系統之缺點，同時將其結合視覺效果的導覽，並評估本系統與其他領域結合之可能性，進而為電腦視覺領域和定位系統方面做出一份貢獻。

## (二)、研究目的

本研究將擴充實境結合視覺里程計，並將其應用於室內的定位導覽系統，試圖透過電腦視覺領域的優勢，改善現有定位系統之不便性，並與多種領域如機器人與物聯網等結合。以下為本研究擬達到之主要目標：

1. 建立出一套結合擴充實境與視覺里程計技術之演算法
2. 實踐此演算法之軟體系統，並將其應用於定位導覽系統
3. 改善此定位系統之準確度與效率，以達到精準的即時計算效果
4. 評估本研究之定位導覽系統與物聯網、機器人等領域的結合可能性

## 二、 研究方法與過程

本研究將依循圖 1 之研究流程進行，首先先對擴充實境與視覺里程計的背景知識與理論進行了解，以進行更深入的研究。對背景知識有了深入的理解後，開始定義須解決的問題並提出實踐程式碼的演算法。最後再根據提出的演算法，進行軟體的實作。

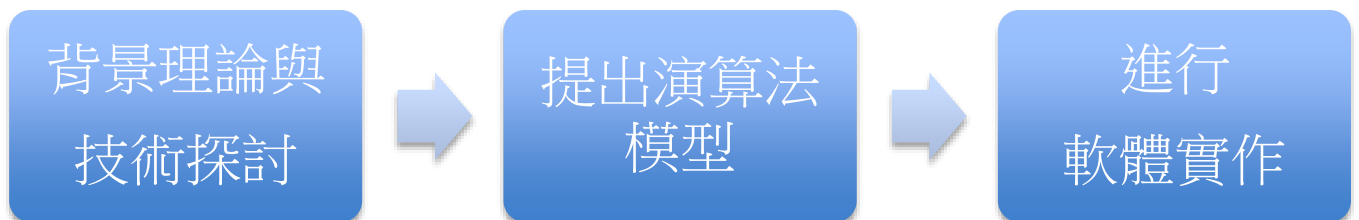


圖 1 研究流程圖

## (一)、背景理論與技術探討

在背景理論與技術探討部分，主要是對擴充實境與視覺里程計之理論與技術進行探討，並將其定為基礎，已進行後續的演算法建立。

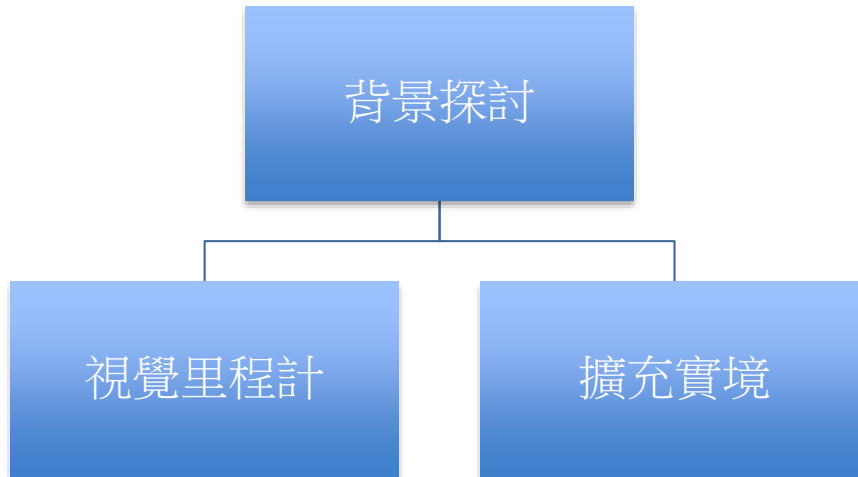


圖 2 背景探討架構圖

### 1. 視覺里程計

視覺里程計 (Visual Odometry, 簡稱 VO) 將透過輸入的連續影像 (即為影片) 進行比對, 以計算出相機本身在三維世界下的旋轉與平移資料。現有的視覺里程計技術已相當成熟, 可即時計算每一時間點下的相機座標的移動。若將每一時間點下的相機三維坐標圖示化並連結, 則可得出相機的移動軌跡。

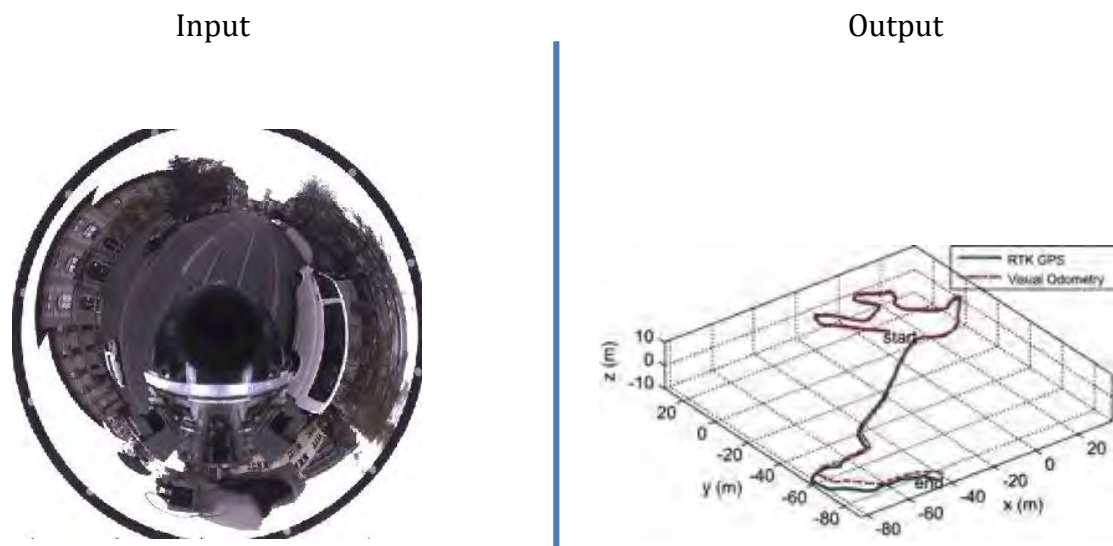


圖 3 視覺里程計示意圖

其中，視覺里程計之所以能計算出不同時間點下拍攝兩不同照片時相機的三維移動資料，主要是透過特徵點計算、相機移動參數計算、以及 Bundle Adjustment 校正，如圖 4 所示。

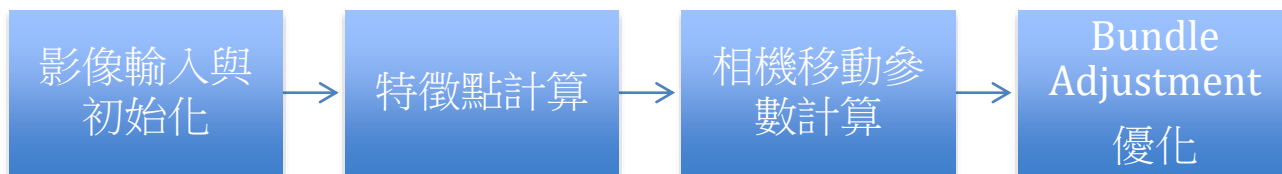


圖 4 視覺里程計流程圖

在視覺里程計的背景討論中，將針對以上流程圖中四個部分進行詳細的說明優化。

### (1) 影像輸入與初始化



在初始化的階段，除了讀入待計算的影像外，還需進行相機內部參數的校正。相機的內部參數包含一個  $3 \times 3$  的內部參數矩陣  $K$  以及一形變參數。校正出以上的參數後，不單是為了校正因相機光學鏡頭所造成的扭曲，其中內部參數矩陣  $K$  將會在計算相機移動參數時被使用到，所以將在相機移動參數部分進行深入探討。以上的校正工作可使用棋盤方格圖配合現有的電腦視覺資料庫進行計算，以校正出內部參數。

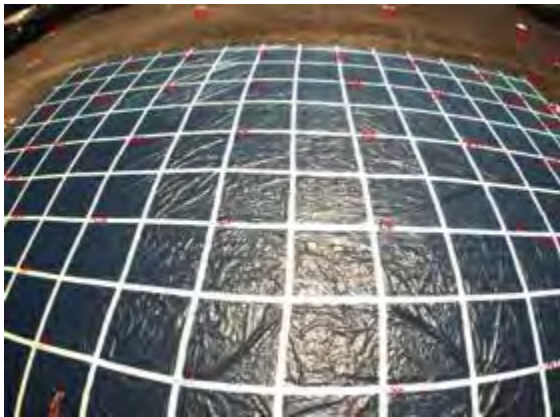


圖 5 照片之光學扭曲

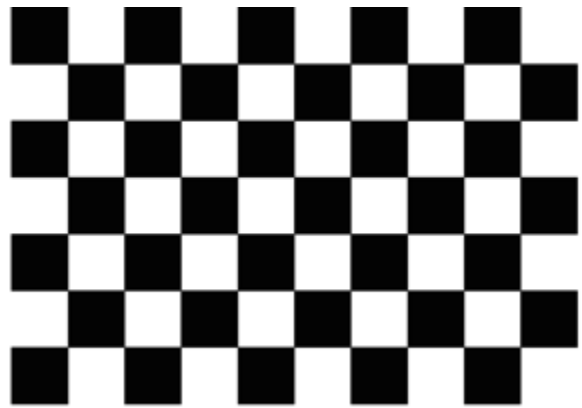


圖 6 棋盤方格圖

## (2) 特徵點計算

特徵檢測為電腦視覺領域中的一個概念。特徵檢測會對一張圖像中的點進行判斷，若為數字圖像中的『特殊』的點，則被演算法判斷為一特徵點。不同的特徵偵測演算法有不同的判斷依據，故無一定的判斷基準，但所有不同的特徵偵測演算法所偵測出的特徵，皆有可重複性這共通點，意味著在不同視角、距離下拍攝的圖像中，可以計算出相同的特徵點，如圖 7 所示。得知了圖像中的特徵點後，便能以圖像中的二維點，推論出相機的三維移動參數。

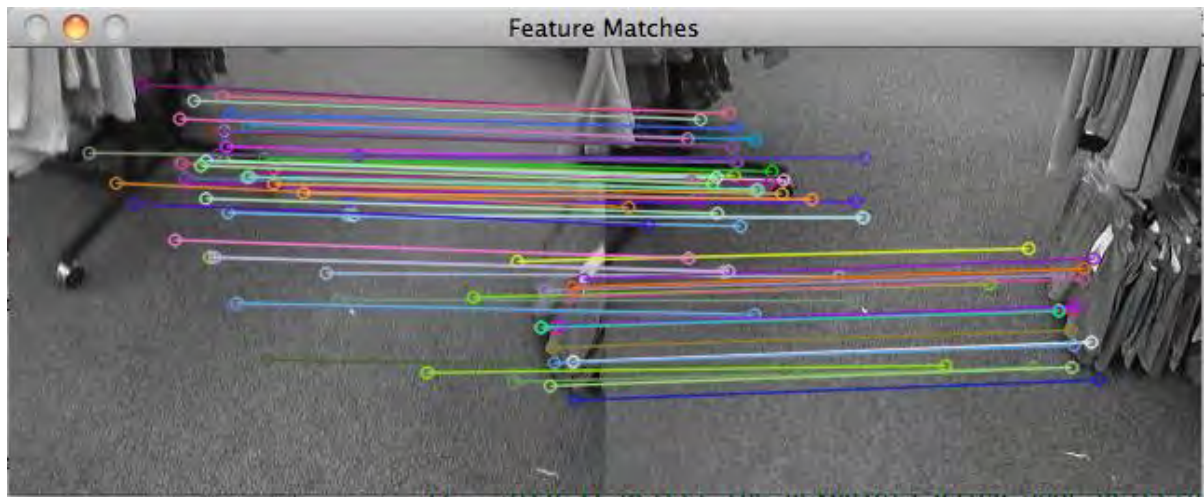


圖 7 特徵偵測與比對示意圖

### (3) 相機移動參數計算

在影像輸入與初始化部分提到，相機的內部參數矩陣 $K$ 在計算相機的移動參數時被使用到。再次對其進行探討。

內部參數矩陣 $K$ 是由焦距 $f$ ，相機中心的  $x$  方向偏移 $x_0$  與相機中心的  $y$  方向偏移 $y_0$ 所組成的  $3 \times 3$  矩陣（公式(1)）。

$$K = \begin{bmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

而在相機針孔模型(Pinhole Camera Model)中(圖 8)，相機中心(Camera Center)到 principal point 的距離即為焦距 $f$ ，同時也是相機中心與影像平面(Image Plane)的距離。

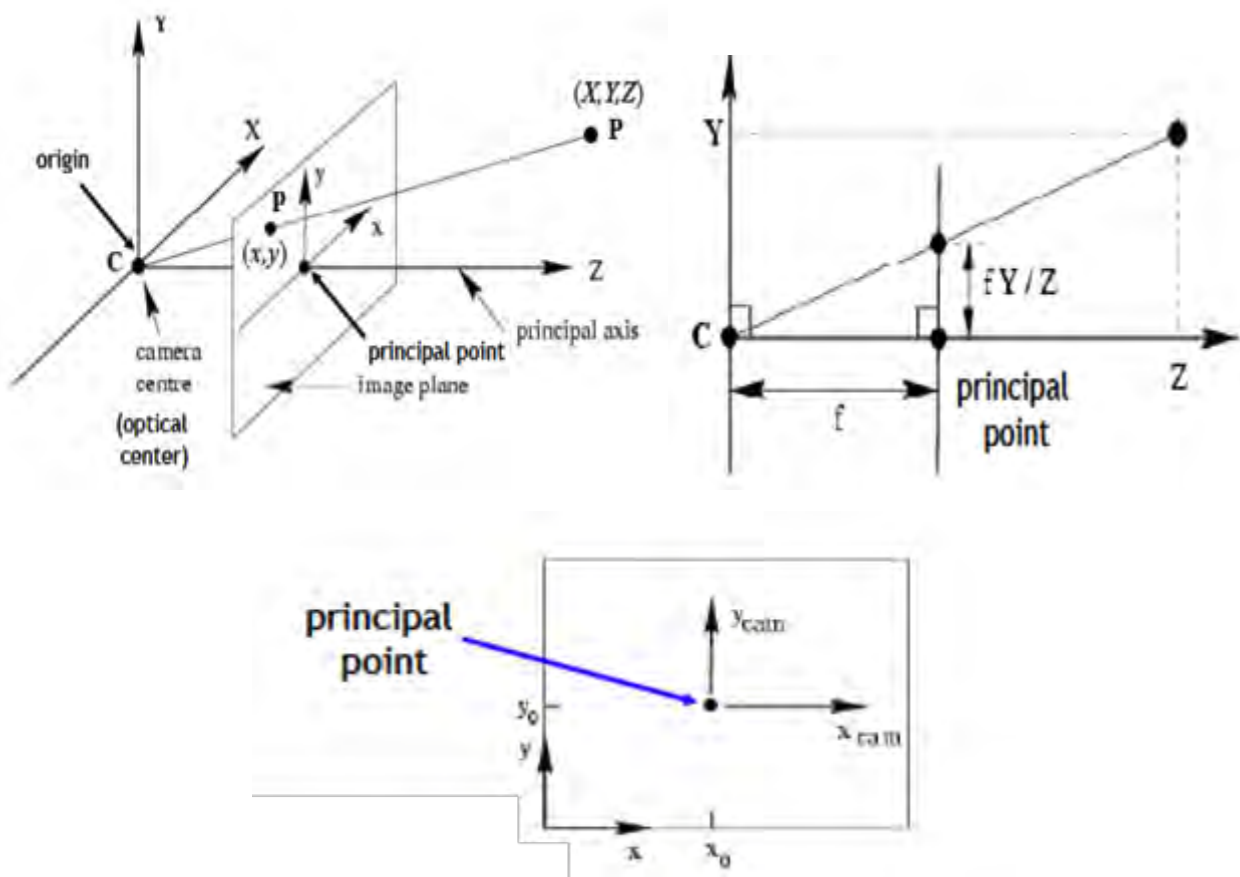


圖 8 相機針孔模型圖

隨後，便可以相似三角形的概念來表示影像平面上的點 $(x,y)$ 與其在三維世界坐標係中的位置 $(X,Y,Z)$ 的關聯，如公式(2)(3)(4)所示。

$$x = \frac{fX}{Z} \quad (2)$$

$$y = \frac{fY}{Z} \quad (3)$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \sim \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & X_0 \\ 0 & f & Y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (4)$$

將兩張不同角度、距離下拍攝的影像進行特徵點偵測後，將得到的特徵點（圖 9 中的藍點）以公式(4)換算，得知這些點的三維坐標（圖 9 中的紅點）後，便可以此為依據，

推算出相機在拍攝這兩張影像之間的轉移資訊矩陣  $P$ 。

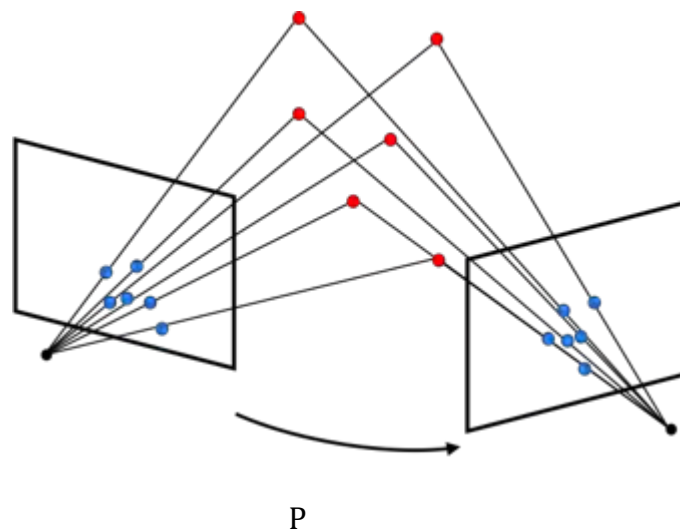


圖 9 相機的移動計算示意圖

其中轉移矩陣  $P$  為一由  $3 \times 3$  的旋轉矩陣  $R$  和  $1 \times 3$  的平移向量  $T$  所組成的  $3 \times 4$  矩陣(公式(5)(6)(7))

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{12} & r_{22} & r_{21} \\ r_{13} & r_{32} & r_{33} \end{bmatrix} \quad (5)$$

$$T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} \quad (6)$$

$$P = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \quad (7)$$

透過上述之推算，便可得知不同時間點下，相機在三維坐標下的旋轉與平移資料。

#### (4) Bundle Adjustment 優化

進行了上述的轉移矩陣計算後，尚需進行優化，避免誤差累積。**Bundle**

**Adjustment(BA)**是一種被廣泛應用於三維重建中的優化方式。

以上為視覺里程計的運作流程，上述的特徵點運算部分到 **Bundle Adjustment** 優化會針對每一時間點下所拍攝的影像與其下一時間點拍攝的影像進行運算，以達到計算整個影片的效果，在一般的程式中，這一部分會以迴圈的方式處理。

但是，如同上述的視覺里程計，普遍上的視覺里程計皆屬於平面與平面之間的比對，因此計算三維空間下的資訊時，會有相當的誤差，為了降低計算誤差，普遍皆使用 **Bundle Adjustment** 等優化方法處理，但進行優化的同時，計算的速度也被影像，導致速度大幅下降。為了克服視覺里程計在速度與準確性之間的權衡，本研究將結合視覺里程計與擴充實境的演算法，提出一套兼具高計算效率與準確性的軟體模型。

## 2. 擴充實境

擴充實境 (**Augmented Reality**, 簡稱 **AR**) 為一即時計算相機影像的位置與角度並在影像內投射虛擬物件的技術，其目標為在電腦螢幕上投射出虛擬三維物件並將其套用在現實世界中，可與使用者進行互動。



圖 10 擴充實境示意圖

如圖 10 所示，擴充實境技術在圖中將虛擬的大樓物件與桌上的虛擬平面圖投射在電腦螢幕中，而當使用者進行移動，旋轉時，擴充實境便會藉由拍攝影像之間的移動以推算出使用者之三維移動資料，再依據其三維移動資料回推虛擬物件的位置，進而使虛擬物件固定在場景中，使虛擬物件有如真實世界中的三維物件一樣，來為使用者提供更多信息。

由於擴充實境可即時並精準的計算使用者在三維世界下的移動與旋轉，故本研究嘗試將其與視覺里程計結合，以彌補其計算效率的缺陷。同時，再搭配擴充實境所投射的虛擬物件，達到使用者互動與導覽的功能。

以下，為擴充實境的計算流程：

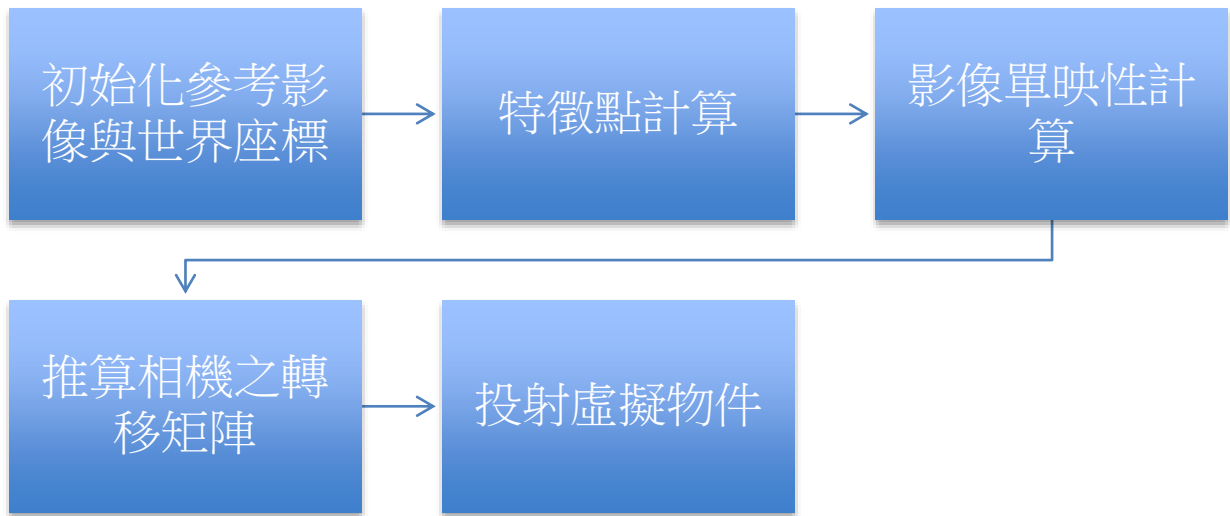


圖 11 擴充實境流程圖

### (1) 初始化參考影像與世界座標

擴充實境的運作原理主要為計算座標系的轉換，分別為計算世界座標系（World Coordinate）與相機座標系(Camera Coordinate)之間的轉換以及相機座標系到影像的平面座標(Image Coordinate)之間的轉換。

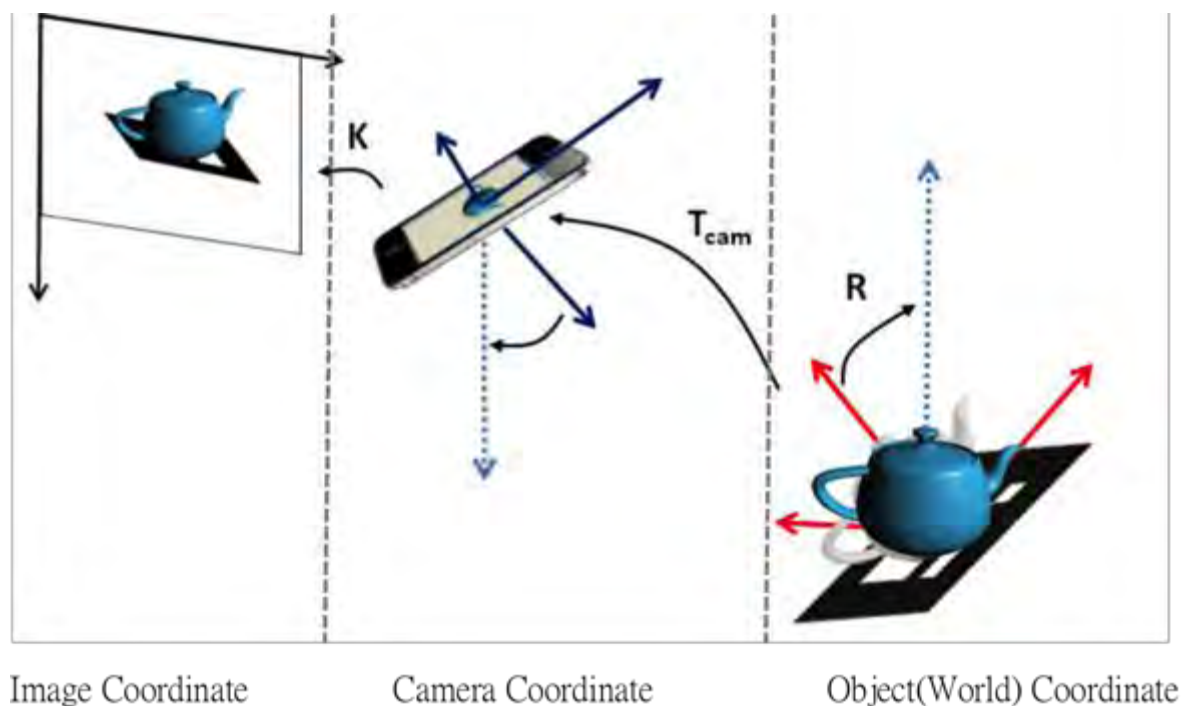


圖 12 擴充實境之座標系轉換

在擴充實境的計算中，會先將場景中一平面影像設定為參考影像（Pattern），並以此影像為基礎，建立一物件座標系，必在此座標系中畫出虛擬物件。再將物件座標系到相機座標系之間的旋轉  $R$  與平移  $T$  計算出來後，便可的之物件座標系與相機座標系之間的轉換。而在相機座標系中，搭配在視覺里程計部分討論的相機內部參數  $K$ ，以公式(4)進行轉換，便可將相機座標系的物件投射到平面影像座標系中，完成螢幕中的投射影像。

然而，這裡的世界座標系是建立在參考影像上，參考影像為一事先拍攝的平面影像(圖 13)，其中將相片的四個頂點設為  $(0,0)$ 、 $(width, 0)$ 、 $(0, height)$ 、 $(width, height)$ ， $width$  為影像的像素寬， $height$  為影像之像素長。以此為依據，開始建立世界座標系，首先將以上四個頂點除以  $height$  和  $width$  之中像素較大者（公式(8)(9)），再將此四頂點設為三維坐標點  $(-W,-H,0)$ 、 $(W,-H,0)$ 、 $(-W,H,0)$ 、 $(W,H,0)$ ，形成一三維世界中的平面。



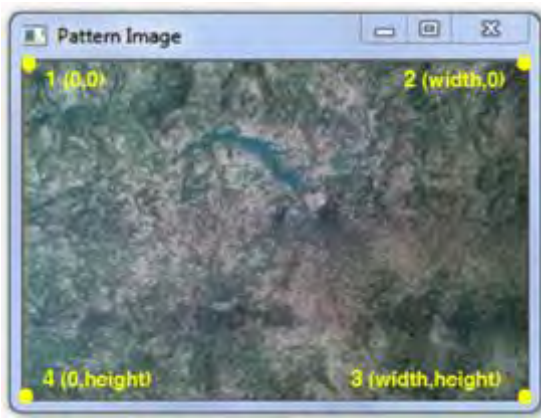


圖 13 參考影像

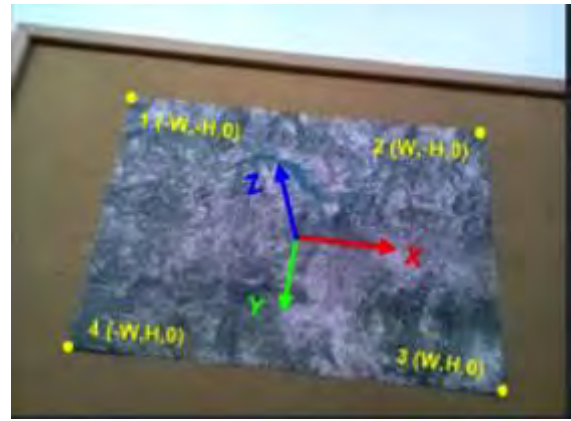


圖 14 物件座標系

$$W = \frac{\text{width}}{\max(\text{width}, \text{height})} \quad (8)$$

$$H = \frac{\text{height}}{\max(\text{width}, \text{height})} \quad (9)$$

以上四個三維點所構成的平面便是世界座標系中的  $xy$  平面，再取其法向量作為  $z$  軸，便形成世界座標系，並完成了擴充實境之初始化。

## (2) 特徵點計算

此部份之特徵點計算與視覺里程計中的特徵點計算雷同，但其進行計算的對象有所不同。在視覺里程計中，是對每一時間點與其下一時間點所拍攝的影像進行比對，但在此擴充實境則是將參考影像與後續拍攝的連續影像進行比對，意味著當參考影像的特徵點偵測初始化後，在後續的工作只需對每一時間點下的影像進行比對，故能有較快的計算速度。

## (3) 影像單映性計算

在完成了參考影像與後續輸入之影片的特徵點計算後，將對其進行影像單映性 (Homography) 的計算。影像單映性為在兩不同角度下拍攝一平面時，其平面在兩相片中的轉換，如圖 15 所示：

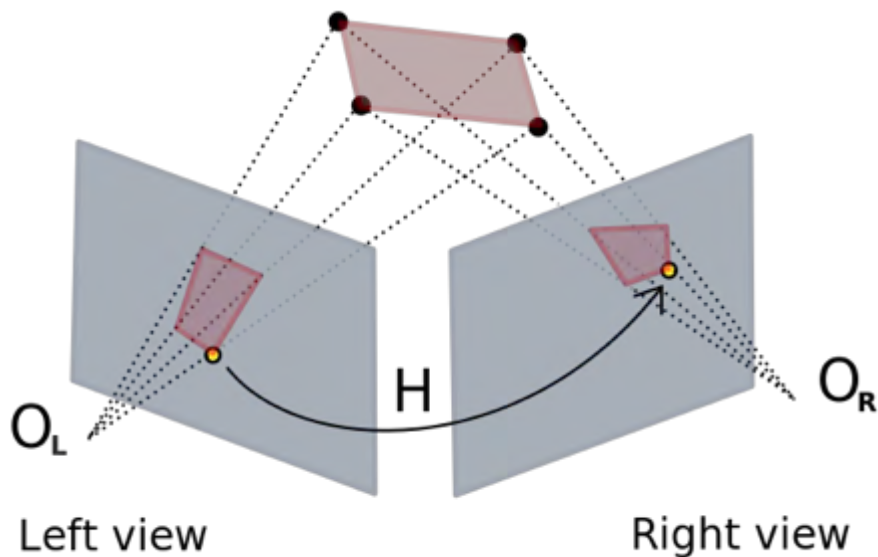


圖 15 影像單映性

其中  $H$  為一  $3 \times 3$  的單映性矩陣，其表示在  $O_L$  與  $O_R$  之間的轉換(公式(10)(11)(12))。

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \quad (10)$$

$$\begin{pmatrix} X'_R \\ Y'_R \\ W \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} X_L \\ Y_L \\ 1 \end{pmatrix} \quad (11)$$

$$\begin{pmatrix} X_R \\ Y_R \\ 1 \end{pmatrix} = \frac{1}{W} \begin{pmatrix} X'_R \\ Y'_R \\ W \end{pmatrix} \quad (12)$$

在擴充實境的應用中，可藉由輸入參考影像與影片場景中的特徵點，計算其之間的

單映性矩陣，便可得知場景中參考影像的四個頂點的位置（圖 16），供下一步驟推算相機之轉移矩陣。

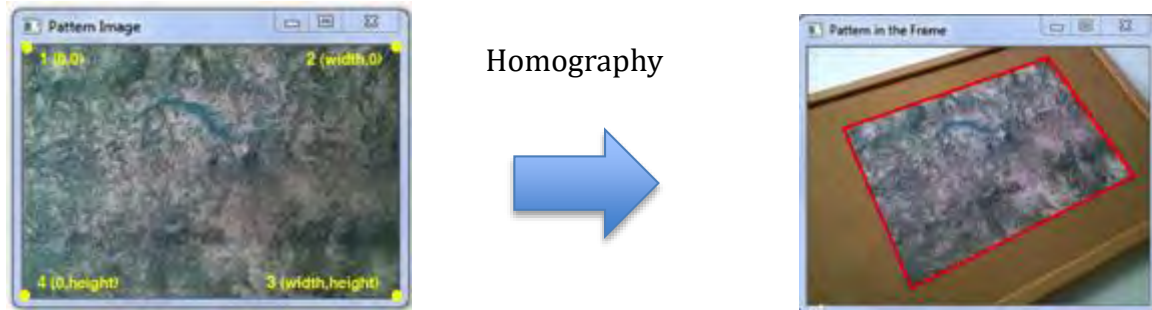


圖 16 單映性轉換

#### (4) 推算相機之轉移矩陣

透過以上的步驟，已可的之參考影像、單映性矩陣、與物件座標系，再加上事先校正出的內部參數矩陣，便可以此推導出相機座標系。由於相機座標系與物件座標系為已知，便可以從其參考影像分別於相機座標系與物件座標系中參考影像至四個頂點來回推兩座標系之間相對的旋轉與平移，而其轉移矩陣便可表示由物件座標系變換到相機座標系的過程（圖 17）。

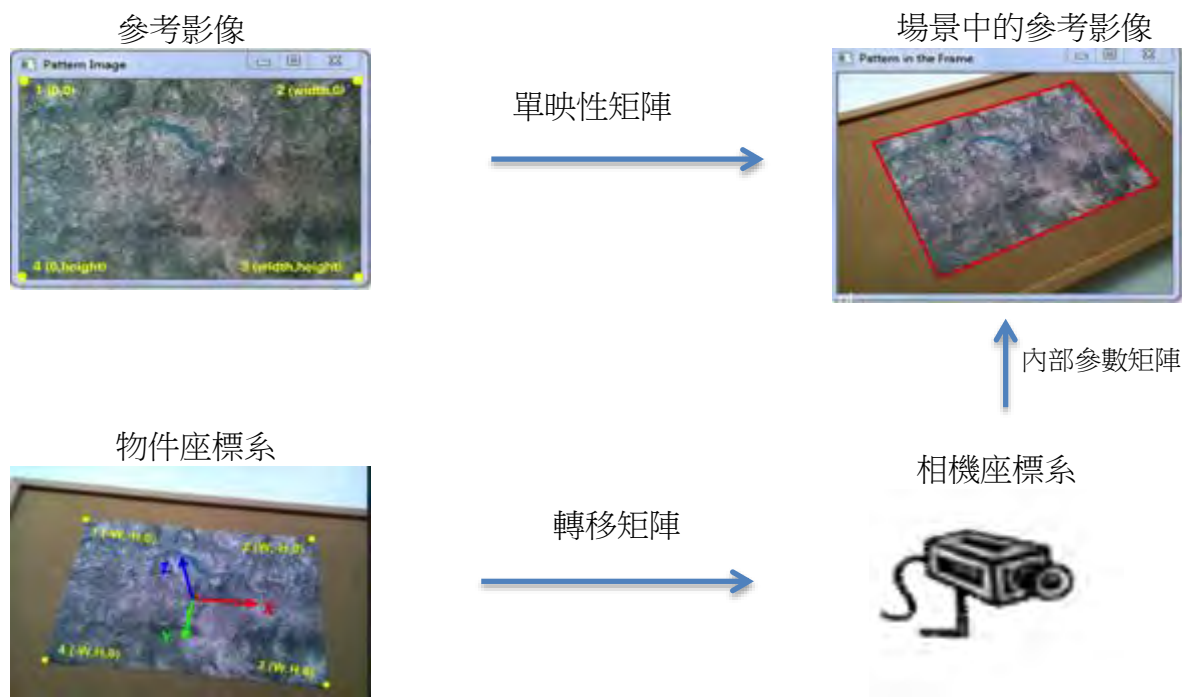


圖 17 轉移矩陣之推導

其中，擴充實境中的轉移矩陣  $P$  與視覺里程計部分之轉移矩陣相同，一樣如同公式 (7) 中表示，是由一  $3 \times 3$  的旋轉矩陣  $R$  和一  $1 \times 3$  的平移向量  $T$  所組成。不過其標示物件座標系到相機座標系之間的轉換，而非相機座標系之間的轉換(公式(13))。

$$\begin{pmatrix} X_{world} \\ Y_{world} \\ Z_{world} \end{pmatrix} = \begin{bmatrix} R_{3 \times 3} & T_{1 \times 3} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} X_{camera} \\ Y_{camera} \\ Z_{camera} \\ 1 \end{pmatrix} \quad (13)$$

### (5) 投射虛擬物件

經由以上步驟，可得出相機座標系與座標系之間的轉換，然而，若在世界座標系中繪出立體圖形，可藉由其轉換到相機座標系，再經由內部參數矩陣投射到平面影像座標，便可在螢幕上呈現虛擬物件，達到擴充實境的功能。

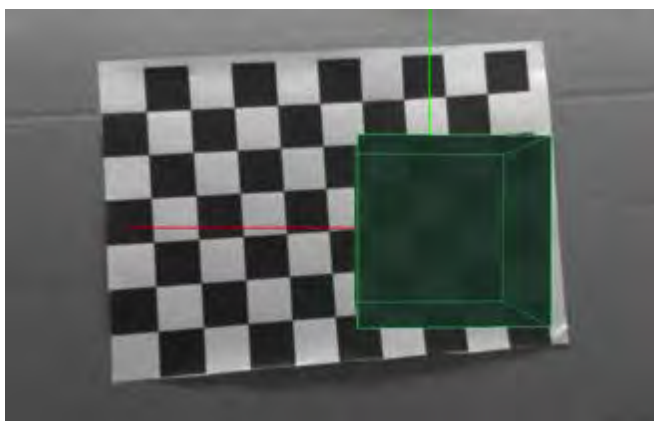


圖 18 擴充實境之投射虛擬物件

## (二)、提出演算法模型

經由背景探討，本研究已對進行實作時所需的知識、理論進行了深入的討論，以方便在後續研究時應用其背景知識。而在視覺里程計與擴充實境的背景探討中，提到了視覺里程計之速率與準確性的缺點，本研究試圖基於擴充實境即時性且準確的精準度，研發出全新的視覺里程計。並將其應用於定位系統，同時利用擴充實境之虛擬物件達到提供使用者資訊的效果。

為了將擴充實境應用於定位系統中，需要達到以下兩點研究方向已使其兩者之間更好的結合：

1. 得出相機相對於參考點的距離
2. 將上一點之距離以公分、公尺等單位表示

下列為本研究之研究方向之詳細說明：

### 1. 得出相機相對於參考點的距離

進行定位時，所需的基本資訊即為一物件相對於參考點之距離。普遍的定位系統中皆採用特定發信器作為參考點，如磁鐵、紅外線發射器、與藍芽等。本研究試圖將擴充

實境中的參考影像作為參考點，計算出相機相對於其的距離，達到定位的效果。

依據公式(13)，擴充實境的以計算出相機座標系與世界座標系之間的轉移矩陣  $P$ ，而轉移矩陣中之平移向量  $T$  可表示兩座標系之件相對位置的移動。然而世界座標系為建立在參考影像上之一座標系，其到相機座標系之間的轉移即可表示為參考影像到相機再三維世界下  $XYZ$  方向之相對距離。因此，在得知了世界座標系到相機座標系之轉移矩陣後，便可將其設計成一視覺里程計，同時以參考影像為準，建立一定位系統之雛形。

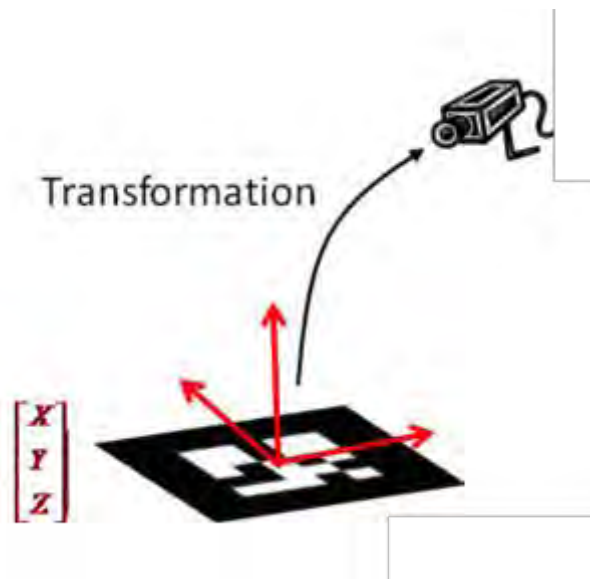


圖 19 計算參考影像與相機之相對距離示意圖

## 2. 將其距離以公分、公尺等標準度量衡表示

在得到了參考影像與相機之間之相對距離後，尚須將其轉換為現實事件中之度量衡。在公式 (4) (8) (9) 中，參考影像之長寬的像素與相機所拍攝的相片皆除以像素，因此相機座標系與世界座標系皆無單位，僅為相對值。同時平移向量  $T$  也無單位，僅為一基於相機座標系與世界座標系之間的相對值。

然而，若得知參考影像之長或寬（以現實世界中之度量衡計），便可得知世界座標系中的單位與公分、公尺之間的轉換，基於單位的一樣，可基於公式(4)與公式(13)推得平移向量  $T$  以公分、公尺計的距離(公式(14)(15)(16))。

$$X_{cm} = \frac{width_{cm}}{W} t_x \quad (14)$$

$$Y_{cm} = \frac{width_{cm}}{W} t_y \quad (15)$$

$$Z_{cm} = \frac{width_{cm}}{W} t_z \quad (16)$$

其中 $X_{cm}Y_{cm}Z_{cm}$ 分別為相機與參考影像之間在  $xyz$  三個方向之相對距離(以公分計),  
 $width_{cm}$ 為參考影像之寬(以公分計)。

基於以上兩點，本研究將結合擴充實境之計算優勢，以視覺里程計的概念將其應用於定位，建立一基於擴充實境之定位系統，以此建構出本研究之演算法流程圖(圖 20)。

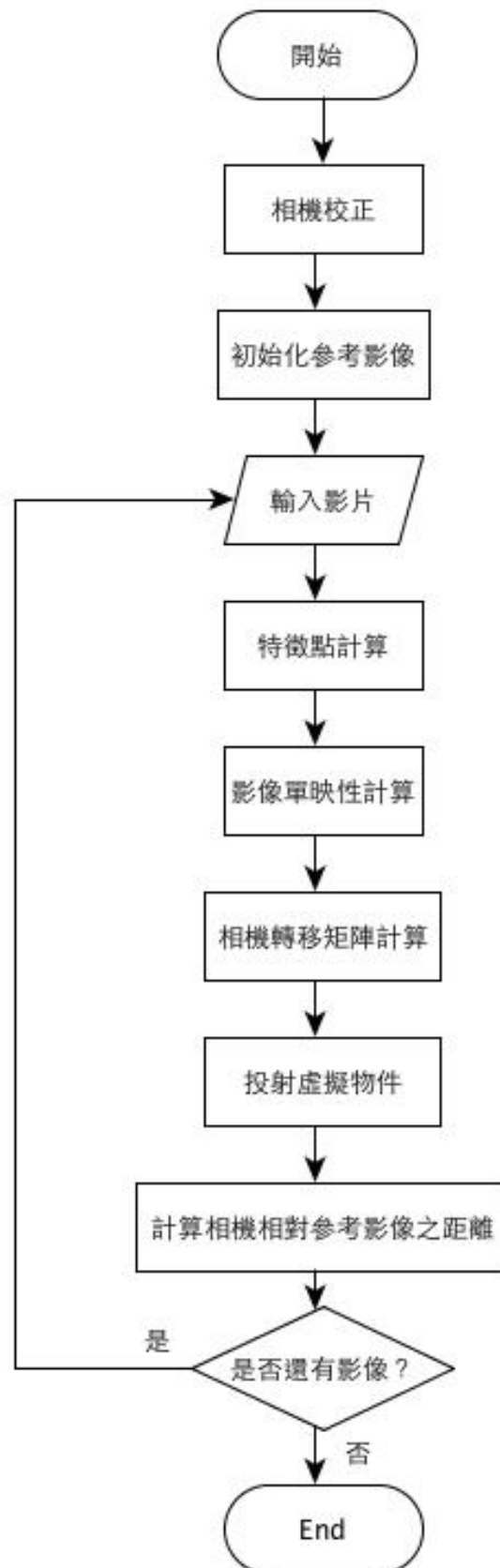


圖 20 本研究之演算法流程圖



### (三)、進行軟體模型實作

在提出了研究目的與建立出軟體模型的演算法後，將對其進行實作，建立出軟體模型，並證實本研究之定位系統之可行性。本研究之軟體模型以 C++ 與 OpenCV、OpenGL 所提供之開放原始碼建立。實作過程先後經由相機校正、以演算法流程圖編寫程式、到設置簡易應用狀況已驗證本研究之定位系統之可行性。以下為實作之流程圖：

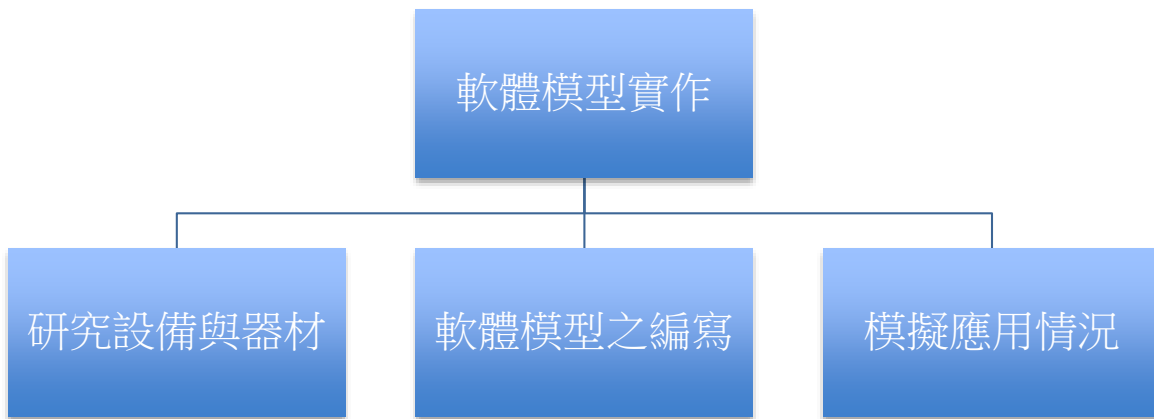


圖 21 實作流程圖

#### 1. 研究設備與器材

下表為本研究所使用之硬體設備與器材：

表 1 硬體設備與器材

實驗設備	型號	數量
網路攝影機	Logitech C615(HD1080P)	1
攝影機腳架	Takara SP-203B	1
電腦	MacBook Pro Retina(Intel Core i7; CPU@2.4GHz; NVIDIA GeForce GT 650M 1024MB)	1
簡易測試模型	由珍珠板切割成的立體支架 與由瓦楞紙製成的方格版,主 要用於測試本軟體模型之精 準度	1

以下圖 22 到圖 24 為本研究之硬體器材：



圖 22 MacBook Pro Retina



圖 23 網路攝影機與腳架



**圖 24 簡易測試模型**

在軟體器材方面，本研究採用蘋果之 Xcode6.4 作為主要開發環境，同時使用 OpenCV 與 OpenGL 開放式原始碼輔助進行程式的建構。

## **2. 軟體模型之編寫**

在此階段，將對於本研究之演算法流程圖進行軟體模型之實作，分別對相機校正、初始化參考影像、特徵點偵測、影像單映性計算、相機移動參數計算、實際距離計算、以及虛擬物件投射進行實作。如以下內容所示：

### **(1) 相機校正**

此部分將輸入多張不同角度下拍攝的棋盤方格圖已進行相機校正，以得出相機之內部參數與扭曲參數。

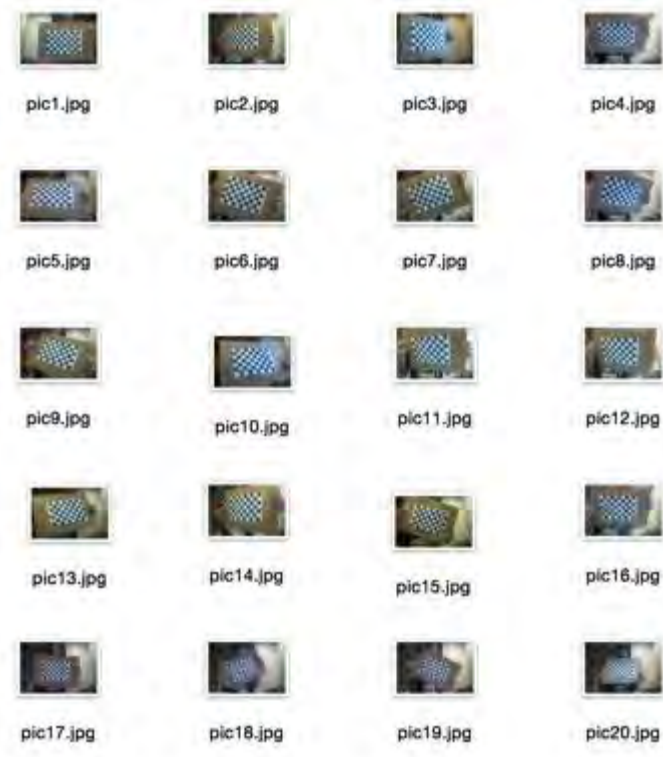


圖 25 相機參數校正

(2) 初始化參考影像



圖 26 參考影像

圖 26 為使用網路攝影機拍攝之參考影像，其實質大小為 30cm x 20cm。以下為初始化參考影像部分之原始碼。

```

void buildpattern(Mat& image, Pattern& pattern)
{
    pattern.size = cv::Size(image.cols, image.rows);
    pattern.frame = image.clone();
    getGray(image, pattern.grayImg);

    pattern.points2d.resize(4);
    pattern.points3d.resize(4);

    const float w = image.cols;
    const float h = image.rows;

    const float maxSize = std::max(w,h);
    const float unitW = w / maxSize;
    const float unitH = h / maxSize;
//設定公式(8)(9)

    pattern.points2d[0] = cv::Point2f(0,0);
    pattern.points2d[1] = cv::Point2f(w,0);
    pattern.points2d[2] = cv::Point2f(w,h);
    pattern.points2d[3] = cv::Point2f(0,h);
//設定參考影像

    pattern.points3d[0] = cv::Point3f(-unitW, -unitH, 0);
    pattern.points3d[1] = cv::Point3f( unitW, -unitH, 0);
    pattern.points3d[2] = cv::Point3f( unitW,  unitH, 0);
    pattern.points3d[3] = cv::Point3f(-unitW,  unitH, 0);
//設定世界座標系
}

```

此部分將輸入之參考影像依據公式(8)(9)進行處理，同時設置世界座標系

### (3) 特徵點偵測

完成了參考影像與世界座標系的設置後，開啟網路攝影機並開始讀入影片，同時對其進行特徵點偵測，以供後續影像單映性與移動參數之計算。然而現有的眾多特徵偵測演算法中，以表 2 列出了適用於本研究之演算法的特徵偵測子，其中雖然 SIFT 與 SURF 可以得出最多的成功匹配特徵，但因速度的考量而放棄使用。BRISK 與 FREAK 則因實驗時誤差過大而捨棄，故採用 ORB 特徵偵測演算子。以下為特徵點偵測部分之原始碼。

表 2 特徵偵測演算法比較表

偵測子	SIFT	SURF	ORB	BRISK	FREAK
計算時間數	15.7233s	4.5235s	0.0235s	0.0642s	0.1562s
成功匹配 對數	5293	2340	492	490	835

```
Ptr<FeatureDetector> featureDetector = new::ORB(5000);
Ptr<DescriptorExtractor> descriptorExtractor = new::ORB(5000);
Ptr<DescriptorMatcher> descriptorMatcher = new FlannBasedMatcher;

vector<KeyPoint> queryKeypoints;

Mat queryDescriptor;

featureDetector->detect(src,queryKeypoints);
descriptorExtractor->compute(src,queryKeypoints,queryDescriptor);
```

以上部分為對輸入的影像已ORB特偵偵測演算法進行計算，採用5000個特徵點。

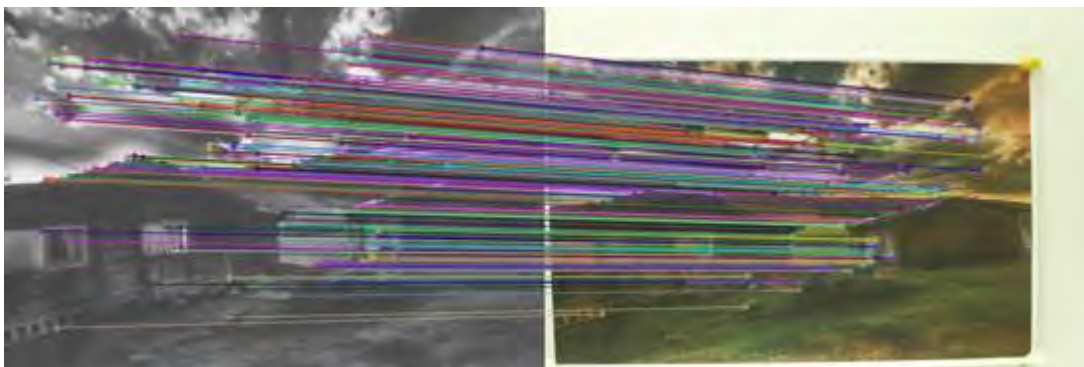


圖 27 ORB 特徵點偵測

#### (4) 影像單映性計算

完成了參考影像與後續輸入之影片的特徵點偵測後，將其輸入 OpenCV 內建知函式，其算出單映性矩陣。

```
homography = cv::findHomography(queryPoints,  
                                trainPoints,  
                                CV_FM_RANSAC,  
                                reprojectionThreshold,  
                                inliersMask);
```

其中 `homography` 為單映性矩陣，`queryPoints` 與 `trainPoints` 分別為參考影像與輸入影片之特徵點。

圖 28 為利用單映性矩陣，將參考影像中的四個頂點根據公式(10)(11)(12)轉換到影片中，即為紅色框框部分。



圖 28 單映性矩陣所推算出的場景中之參考影像

#### (5) 相機移動參數計算

透過以上步驟求得的單映性矩陣、世界座標系、內部參數等資訊，將其作為已知，推算公式(13)之相機移動參數。

```
Mat rvec, tvec;  
Mat r;  
  
solvePnP (pattern.points3d,pattern.points2d,K,distcoeff,rvec,tvec);  
rvec.convertTo(r,CV_32F);  
Mat_<float> rotMat(3,3);
```

```
Rodrigues(r,rotMat);  
raux = rotMat;
```

以上為透過 OpenCV 之內建函數 solvePnP 藉由已知的世界座標系(pattern.points3d)、平面影像座標系(pattern.points2d)、內部參數(K)、以及扭曲參數(distcoeff)，推算出旋轉向量(rvec)與平移向量(tvec)，再以 Rodrigues 公式將旋轉向量轉為 3x3 的旋轉矩陣。

```
Cam Matrix(solvePNP)[0.99961364, -0.0010668441, 0.027773976, -0.12834275;  
0.0020840212, 0.99932706, -0.036620244, -0.0073316661;  
-0.027716218, 0.036663976, 0.99894321, 2.6118462]
```

圖 29 軟體模型所計算出之轉移矩陣

## (6) 實際距離計算

經由上一步驟，本研究之軟體模型計算出了世界座標系與相機座標系之間的轉移矩陣。在這一步驟中，將藉由已知的參考影像之寬(30cm)，基於公式(14)(15)(16)計算出以公分為單位之相機與參考影像中心之距離。

```
const float scale = 15;//cm  
  
x = scale * t.at<float>(0,0);  
y = scale * t.at<float>(0,1);  
z = scale * t.at<float>(0,2);
```

在這裏，透過參考影像的半寬 scale，將平移向量的 xyz 轉換為以公分為單位

## (7) 虛擬物件投射

本步驟將藉由步驟(5)中得出的轉移矩陣與先前校正出的內部參數 K，先於世界座標系匯出三維虛擬物件，之後再投射至平面影像座標。



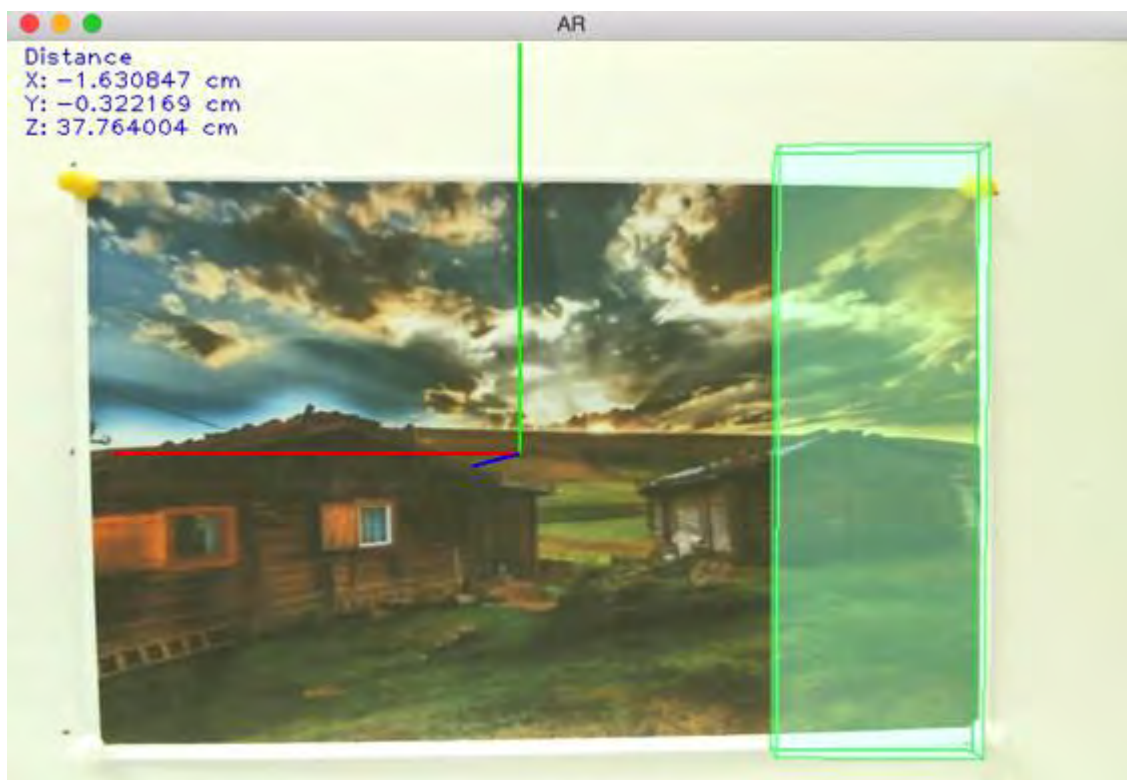


圖 30 擴充實境之虛擬物件投射

### 3. 模擬應用情況

經由上述之軟體模型實作，已可建立出一套完整的定位軟體模型，並完整實踐出演算法中的各個步驟。在此，將使用自製之簡易場景模擬實際使用情況，待配本研究之軟體模型，進行實驗。



圖 31 簡易應用模型

以下圖 32 33 分別為在應用模型中，本研究之軟體模型之輸出影像，其在左上角顯示了攝影機到參考影像中心以公分計的距離，並以擴充實境投射一三維坐標軸與一綠色透明版與影像中。

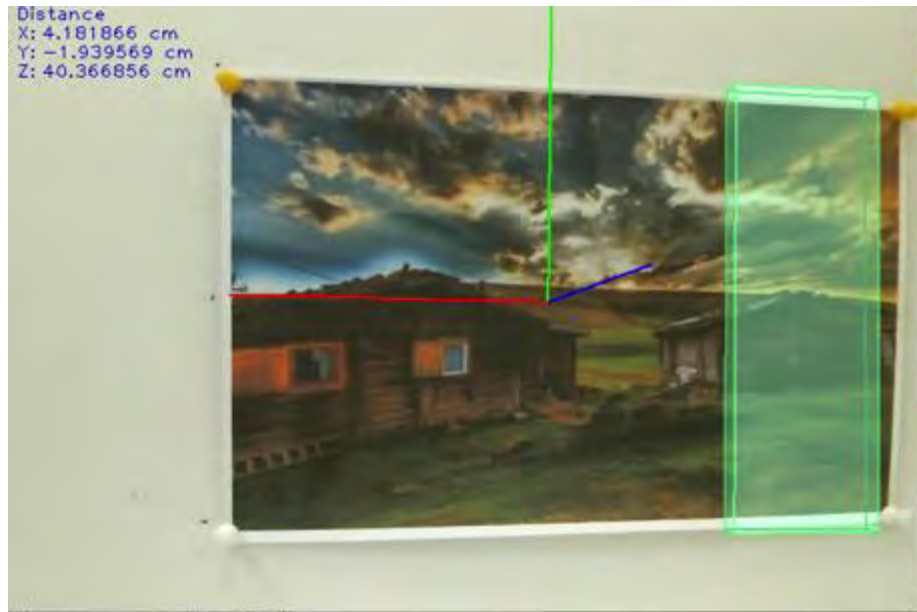


圖 32 輸出結果

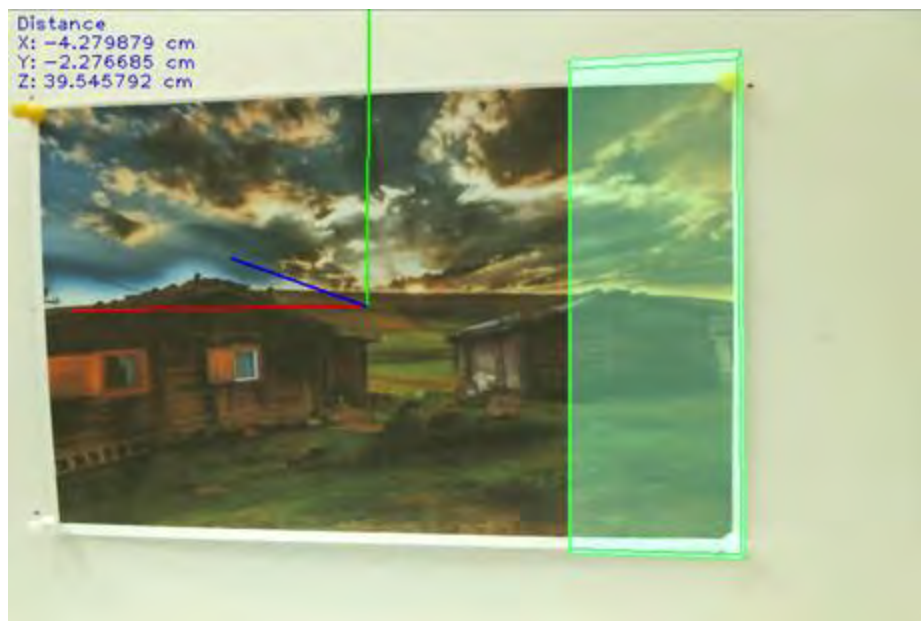


圖 33 輸出結果

### 三、 研究結果與討論

在研究結果與討論部分，主要將針對以下四點研究成果進行討論：

- (一)、 本研究演算法之視覺里程計
- (二)、 本研究演算法之計算精準度與速度
- (三)、 本研究之擴充實境投射物件效果
- (四)、 評估本定位系統與美術館導覽、感測網路、機器人等多領域結合之可能性

#### (一)、 本研究演算法之視覺里程計

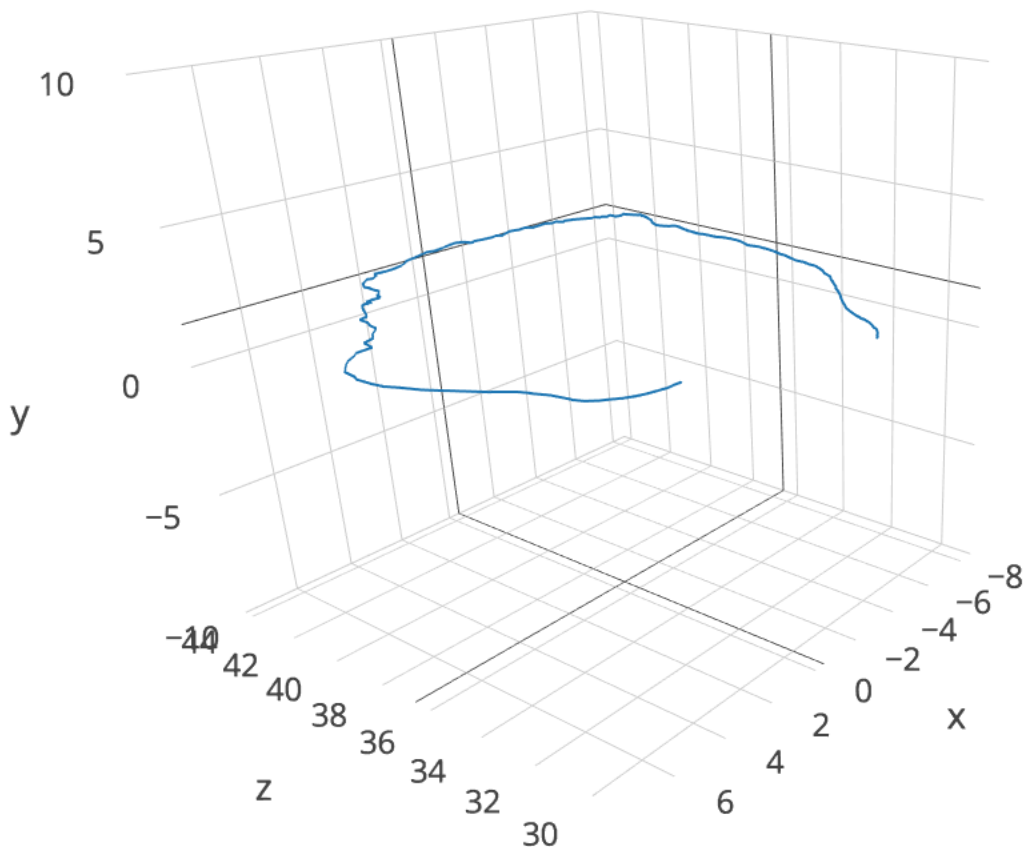


圖 34 視覺里程計之相機移動軌道

由於在擴充實境中計算相機座標系與世界座標系的轉移矩陣時，可藉此轉移矩陣得知相機在三維空間中的移動與旋轉，達到視覺里程計的目的。圖 34 為搜集的相機位置處理後的資料，從此

三維路線圖中可以得知當相機在 XZ 平面中移動時，其 Y 軸的量值並沒有明顯的變動，可驗證本演算法之正確性。並經由此評估，可在未來將本研究應用於視覺里程計或是視覺即時定位與地圖建構 (V-SLAM)。

## (二)、本研究演算法之計算精準度與速度

本研究所建立之軟體模型可成功計算出相機相對於參考影像中心之 xyz 方向的距離，並將其以標準度量衡如公分等表示。在進行距離計算時，本研究之軟體模型可達到即時並精準計算的效果。下表為模擬應用情況中利用自製之簡易模擬模型所量測之數據：

表 3 軟體模型之距離量測資料

	Point 1	Point 2	Point 3	Point 4	Point 5
Actual_X	0	-5	-5	-5	5
Actual_Y	0	0	0	5	5
Actual_Z	35	35	40	40	40
Measured_X	-0.18	-5.17	-5.06	-4.58	5.17
Measured_Y	0.23	0.01	0.08	5.03	5.36
Measured_Z	34.97	35.71	39.42	39.16	39.85
Error_X(cm)	0.18	0.17	0.06	0.42	0.17
Error_Y(cm)	0.23	0.01	0.08	0.03	0.36
Error_Z(cm)	0.03	0.71	0.58	0.84	0.15
Error (cm)	0.29	0.73	0.21	0.94	0.43

在表 2 中，Point 1 到 Point 5 分別代表在不同位置下量測之相機相對於參考影像中心之距離，Actual\_X 到 Actual\_Z 分別代表其實際之三維世界下的距離差（以公分計），Measured\_X 到

Measured\_Z 則為本研究之軟體模型量測之距離差，由表 2 可以看出，其誤差值皆小於 1cm，其平均值更是在 2 公分以內，可十分精準的量測其距離。在本次量測中，採用的參考影像為 30cm x 20cm 大小，在 z 方向之距離差為一公尺以內皆可進行精準的量測，在實際應用場合，將搭配尺寸更大的參考影像（100cm~250cm 寬），可精準的量測至約六公尺之距離，若在同時搭配多張影像，建立一影像資料庫，則可根據定位場所制定出完整涵蓋其範圍之定位系統。

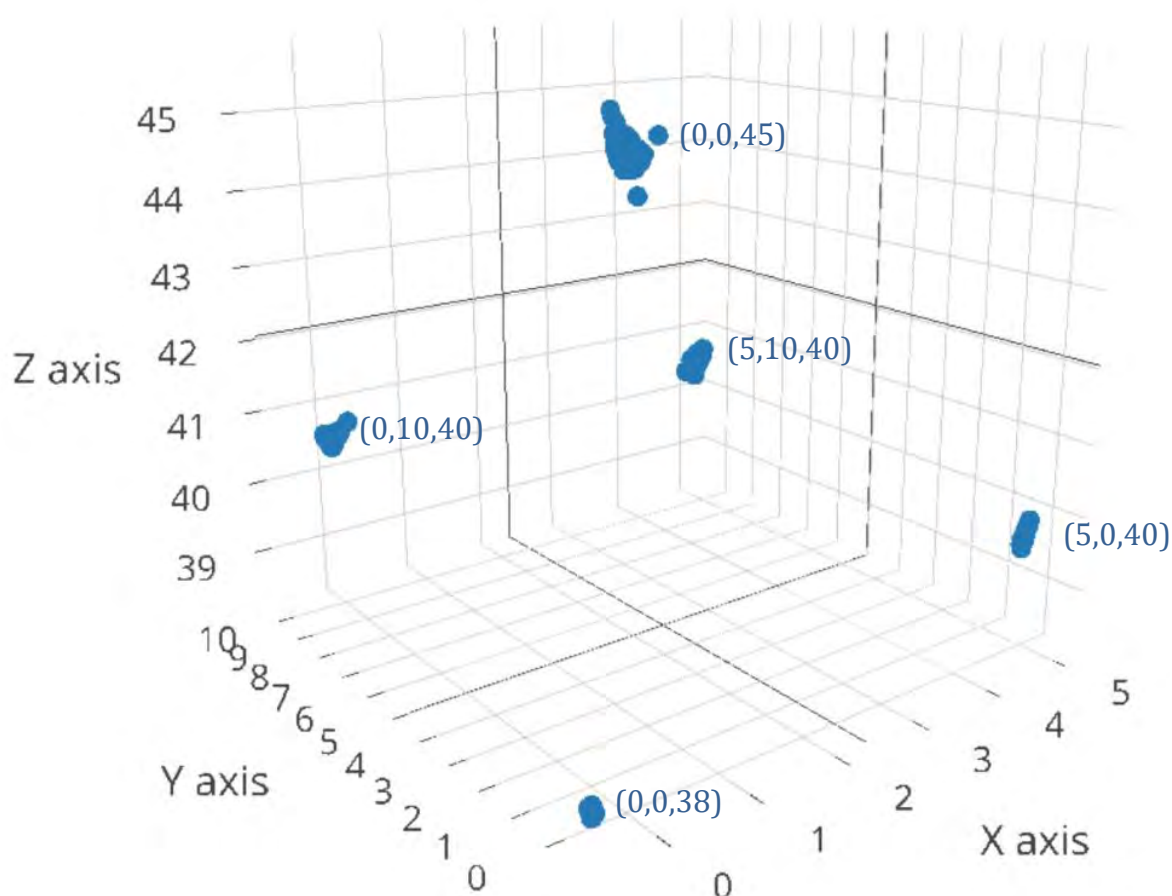


圖 35 穩定性測量

為驗證本研究之定位系統的準確性，上圖為進行穩定性測量的資料，圖 34 中個點為在其位置下 15 秒內所搜集的數據點，並將其進行繪圖。由圖中可得知，其中各個數據點皆相當集中，並不會嚴重的發散，可得知本定位系統之穩定性。

在計算速度方面，可於平均 **0.067 秒**內完成一影片之計算，包含演算法流程圖中之特徵點偵測到投射虛擬物件部分。此速度可完美的應用於 10 fps(Frame Per Second)與 15 fps 的影片中，在

30 fps 之影片中，可挑整為每兩張照片進行一次計算，依舊可達到視覺暫留以呈現出影片之效果，並不影像定位部分之計算。因此本研究之軟體模型整體上有高計算準確性與高計算速度之優勢，可被應用於實際場合。此外，由於本研究僅需影像來作為參考，並不需任何發信器作為參考點，因此可被視為一全新的定位系統。

### (三)、本研究之擴充實境投射物件效果

下列圖 34 到 38 為在不同角度與距離下擴充實境投射虛擬物件之情況：



圖 36 擴充實境之截圖

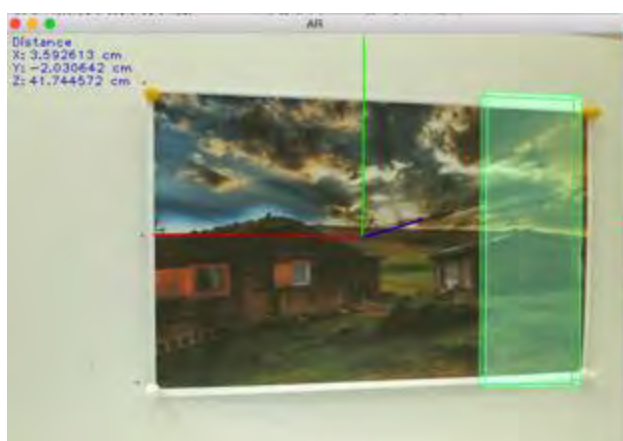


圖 37 擴充實境之截圖

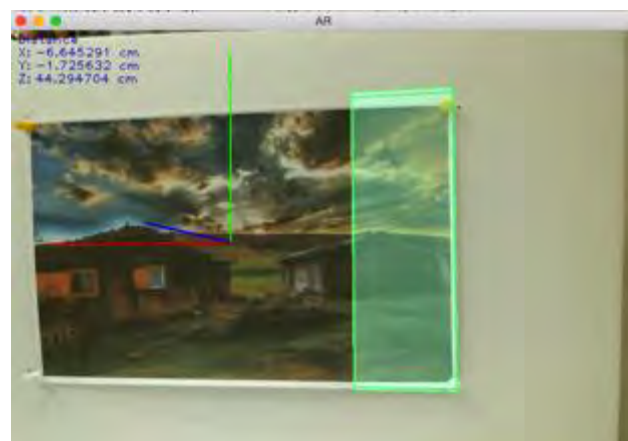


圖 38 擴充實境之截圖

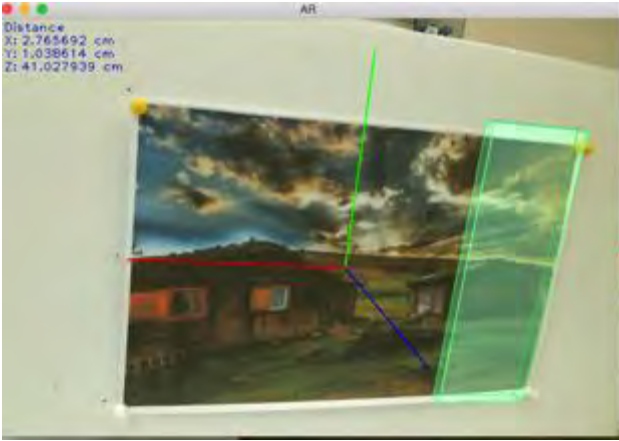


圖 39 擴充實境之截圖

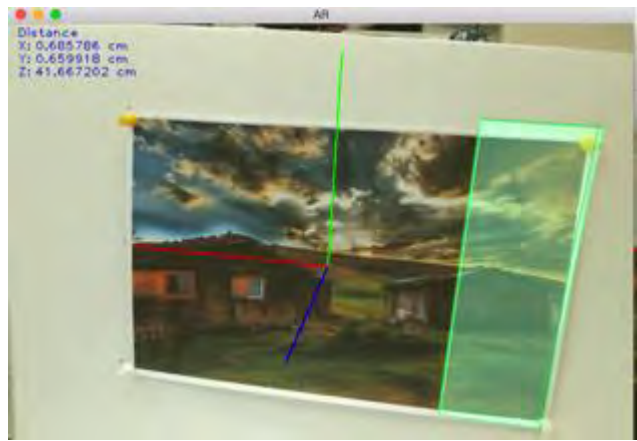


圖 40 擴充實境之截圖

由以上圖片可以看出，無論是在不同距離或是不同角度下拍攝，擴充實境之投射物件皆可『黏』在參考影像上，宛如真實的三維物件一般，達到逼真的效果。而在不同角度下拍攝時，世界座標系與相機座標系之間的旋轉矩陣也隨著改變，所以須做到精準的測量才能使投射的虛擬物件逼真的呈現在螢幕中，從以上的輸出圖看來，其旋轉矩陣的計算十分成功。在未來的應用中，可將多種圖片與文字以 OpenGL 匯成物件，投射至擴充實境中，達到對使用者提供導覽資訊的效果。此方法與 QRCode 相比，QRCode 僅為辨識出一參考圖像後顯示出二維的平面資料如網址等，然而本研究之擴充實境方法所投射之物件可依據使用者相機的移動旋轉，使其物件移動，可達到更加逼真的效果，同時，本研究之軟體模型是藉由參考影像作為辨識，比起 QRCode 的辨識條碼更加美觀，因此可提供更佳的使用者體驗。

#### (四)、評估本定位系統與美術館導覽、感測網路、機器人等多領域結合之可能性

由於本研究使用之擴充實境的定位方法屬於影像定位的一種，因此於美術館、博物館等文物展覽場合，可使用其展覽畫作作為參考影像，進行定位，讓使用者在胸前別上小型攝影機，同時搭配虛擬投射物件，進行使用者導覽的功能。若搭配機器人，可於自走車或是四軸飛行器上安裝攝影機取向，除了於螢幕上的資訊提供為，可藉由機器人進行物理意義上的引導。在感測網路方面，由於本研究之定位方法出了三維空間下的位置，也可藉由擴充實境中計算出之旋轉矩陣判斷使用者面向的方向。因此本研究之定位系統可成為感測網路中的感測器之一，可被應用的範圍廣

大，舉服飾店為例，在服飾店中，可將大型的模特兒看板設為參考影像，進行定位，在得出使用者之位置與面向方向後，將此數據統整，藉由其位置與面向方向判斷客人對哪些位置的服飾較有興趣，同時以擴充實境提供服飾資料等。

由此章的討論，可看出本研究之研究成果同時具備高計算精準度與速率，並可成功的投射虛擬物件提供資訊，也具備與其他領域，如感測網路、機器人、或是美術館導覽等應用結合。具有多方面的好處，可對多種層面提供非凡的貢獻。



## 四、 結論與應用

本研究成功的建立一套以擴充實境為基礎之視覺里程計，以此為基礎，建立出一全新的零點定位系統，在定位的功能外，可藉由擴充實境提供使用者額外的資訊。同時根據實作中的數據與軟體模型之輸出圖片驗證其效能。以下為歸納出之結論：

- (一)、 建立一基於擴充實境之視覺里程計，並利用其世界座標系與相機座標系之間的轉換，得知相機與參考影像之街的距離。
- (二)、 透過參考影像之以公分計的長或寬，將上一點之距離單位轉換為公分。
- (三)、 此定位系統在使用 30 cm 寬之參考影像時，其距離量測誤差僅為 **0.52cm**，且可根據尺寸更大的參考影像的出更大的定位範圍，因此不受參考點的限制。
- (四)、 本研究之軟體模型可高效率的進行計算，平均每次計算僅需 **0.067 秒**，可被應用於多種規格之影片，若搭配規格更高的繪圖卡，則可更進一步的提升其速度。
- (五)、 本研究之研究成果除了進行定位，有別於一般的定位系統，可配合擴充實境之功能達到提供使用者導覽訊息的作用。
- (六)、 本研究之定位系統僅需參考影像與其尺寸大小的資料便可建立，**不需發信器**，可被定義為一種**全新的定位系統**，在設置的方便性上，有著壓倒性的優勢。

本研究之軟體模型在需要精準定位與導覽的場合如美術館、博物館、教育機構等可提供極大的貢獻，同時可與多種領域進行跨領域結合，本軟體模型之高效率、提供多元化資訊、以及與其他領域結合之可能性有極大之發展潛力。本研究將往一下目標進行努力：

- (一)、 驗證本研究於大尺寸影像與多張參考影像場合之計算精準度
- (二)、 改善演算法，提升其計算精準度與速度
- (三)、 將本研究之研究成果與上述之應用場合結合
- (四)、 將本研究與感測網路、機器人等技術結合，提供多種層面的應用

經由以上所歸納出之結論，可以看出本研究不止於定位方面有著極大的貢獻，同時具有與多種領域進行跨領域結合之未來發展，具有十分之潛力。本此研究僅為這部分的一個開端，未來，也將持續努力，已讓此成果對人類的發展貢獻出一份心力。

## 五、 參考文獻

- [1] David Scaramuzza and Friedrich Fraundorfer(2011). Visual Odometry Part1: The First 30 Years and Fundamentals. IEEE Robotics and Automation Magazine, Vol. 18, issue 4, 2011.
- [2] Daniel Lelis Baggio. Mastering OpenCV with Practical Computer Vision Projects. PACKT open source publishing, 93-127, 2012.
- [3] Friedrich Fraundorfer and David Scaramuzza (2011). Visual Odometry Part2: Matching, Robustness, Optimization, and Applications. IEEE Robotics and Automation Magazine, Vol. 19, issue 2, 2012.
- [4] OpenCV. "Camera Calibration and 3D Reconstruction". OpenCV 2.4.11.0 documentation(OpenCV API Reference)
- [5] OpenCV. "Feature Detection and Description". OpenCV 2.4.11.0 documentation(OpenCV API Reference)
- [6] S.Benhimane & E. Malis(2004). Homography-based 2D Visual Tracking and Servoing. The International Journal of Robotics Research, Vol. 26, No.7, July 2007, pp.661-676.
- [7] Sanni Siltanen(2012). Theory and Application of Marker based Augmented Reality. VTT
- [8] 莊永裕(2013)。數位視覺效果。取自 <http://www.csie.ntu.edu.tw/~cyy/courses/vfx>

## 【評語】 190004

1. 研究主題具應用價值，且系統完成度高，值得肯定。
2. 研究實驗可多加強。尤其可針對各個步驟的重要性/準確性/辨識率進行探討。