淡江大學機械與機電工程學系

畢業專題規劃書

(專題題目)

🌕🌕🌕、🌕🌕🌕、🌕🌕🌕(學生姓名)

指導老師：🌕🌕🌕

摘要

本研究計畫將發展視覺輔助的機器手臂系統，執行自動倉儲的貨品挑揀任務。期望發展出有效率的視覺感測裝置與物件偵測演算法，搭配工業機器手臂與夾爪端效器，透過完整的聯網技術，實現機器手臂執行自動挑揀各式貨品的任務。視覺感測與追蹤貨品方面，將使用Kinect感測器，結合多種物件偵測器、色彩索引長條圖、連續自適應均值偏移演算法等，發展視覺偵測與追蹤多個移動貨品的系統。其中物件偵測器偵測包含貨品外形與大小的訊息，色彩索引長條圖則是包含貨品特徵色調分佈的訊息，連續自適應均值偏移演算法則是提供有效率的物件追蹤方法。這些結合將可發揮強健地偵測、描述、與辨識移動貨品的功能，輔助機器手臂執行貨品挑揀的任務。

1. 研究動機與研究問題

1. 在零售業的自動倉儲系統中，使用機器手臂輔助進行挑揀貨品的任務，將是提高工作效率與減少挑錯貨品的重要手段。視覺系統(vision system)可以輔助機器手臂執行自動倉儲的部分關鍵任務，例如在箱櫃中挑揀(bin picking)、包裝(packaging)、貨品堆疊(palletizing)、與搬運(handling)等，這些任務的相關技術之發展是近年智慧型行動機器人的重要研究課題之一。近年積極想實現自動倉儲的最佳範例為全球網購零售巨擘亞馬遜(Amazon)公司，該公司針對倉儲的搬運程序發展行動機器人自動化載運系統以提升效率，成功減少撿貨員在倉儲中大部分的行走路程與找尋貨品的時間[1]。其概念是將供貨倉儲廠房規劃為結構化環境，在地板上佈置大量的數位條碼貼紙。感測與導引裝置行動機器人在倉庫中巡航，每部機器人也具備感測器以避免彼此碰撞。但是，針對困難度較高的挑揀、包裝、貨品堆疊等倉儲任務，目前仍然沒有可以商業化應用的自動化系統，此項技術對機器人研究領域仍然是一項非常困難的挑戰。亞馬遜公司為了促進這些基本科技的發展，在2015年IEEE機器人與自動化國際研討會(ICRA)會場舉辦亞馬遜挑揀競賽(Amazon Picking Challenge, APC)[2]。期望強化產業界與學術界的連結，以及共享非結構自動化(unstructured automation)的研究成果。本計畫即針對視覺系統輔助機器手臂執行挑揀任務的議題，提出研究構想。期望發展出有效率的視覺感測裝置與物件偵測演算法，搭配靈巧活動的手臂機構與可靠抓取貨品的端效器，透過完整的聯網技術，實現機器手臂執行自動挑揀各式貨品的任務。

2. 文獻回顧與探討

使用攝影機輔助偵測不同形式與大小的貨品，關鍵的步驟是貨品形狀與特徵的偵測。必須將這些形狀、特徵、與擺放位置不同的貨品從影像中分辨出來，以便進行物件辨識與追蹤。這是偵測與追蹤具不同形狀與特徵的移動物體之研究議題，在機器視覺與機器人的研究領域仍然是困難的挑戰。文獻[3-7]發展物件偵測器(object detectors)，針對「移動的物體」的特徵、外形、或顏色進行偵測與追蹤，應用在行人、臉部、深度外形、與膚色等物體的偵測與追蹤。在特定環境條件下，可以成功偵測與追蹤移動物體。但是，在一般環境中偵測效果仍然不穩定，而且無法分辨外型相似的物體。使用色彩索引(color indexing)長條圖[8-9]描述物體特徵，除了可以分辨外型相似的物體之外，也可以達到特徵偵測的尺度不變(scale invariant)與方向不變(orientation invariant)的效果，提高物體偵測成功率。本研究將使用物件偵測器搭配色彩索引長條圖，強健地偵測與描述具不同形式與大小且移動中的貨品。另外，均值偏移(mean shift)演算法[10-11]可以應用做為有效率的物件追蹤演算法，其搜尋物件原始位置之最接近且最大的值，並且以機率分佈方式描述[12-14]。均值偏移演算法的改良版本稱為連續自適應均值偏移(continuously adaptive mean shift, CamShift)演算法[15-17]，其追蹤方式是內含均值偏移演算法的程序，透過改變追蹤視窗的大小再次進行疊代運算，直到收斂到設定的門檻值。本計畫將使用連續自適應均值偏移演算法，規劃進行有效率的物件追蹤演算法。

3. 研究方法及步驟

本計畫將結合多種物件偵測器、色彩索引長條圖、連續自適應均值偏移演算法等，應用在攝影機偵測與追蹤多個移動貨品的系統。其中物件偵測器偵測包含貨品外形與大小的訊息，色彩索引長條圖則是包含貨品特徵色調分佈的訊息，連續自適應均值偏移演算法則是提供有效率的物件追蹤方法。這些結合將可發揮強健地偵測、描述、與辨識移動貨品的功能，提供機器手臂偵測與辨識貨架上的貨品，進而輔助機器手臂執行貨品挑揀任務。

視覺輔助機器手臂執行貨品挑揀的系統架構規劃如圖1所示，利用淡江機電系王銀添教授實驗室現有的設備，架構模仿倉儲自動化的測試場景，包括圖2所示的廣明光電SCARA機器手臂與Pioneer P3DX行動機器人。貨品將置放在兩部行動機器人所載運的工作檯上，將使用SCARA機器人做為挑揀貨品的機器手臂，透過微軟Kinect感測器進行物件偵測與辨識。Kinect架設在手臂上，輔助機器手臂抓取物件。

以下將根據所規劃的系統架構，依序說明物件偵測器與色彩索引長條圖、均值偏移與連續自適應均值偏移演算法、RGB-D視覺感測器、貨品位置與方向估測、機器人與視覺聯網等研究議題，闡述本計畫所提出各模組功能的可行方案，以及預期的研究成果。

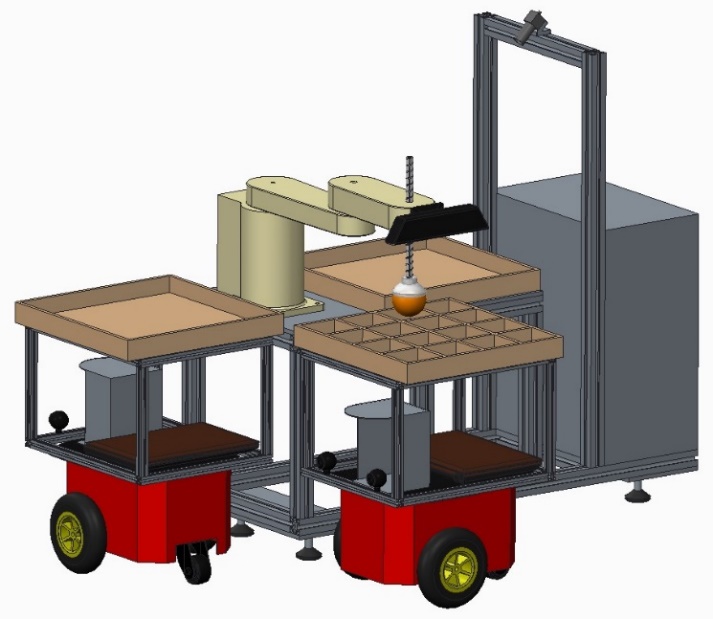


圖1 使用現有設備所規劃的模仿倉儲自動化之測試場景

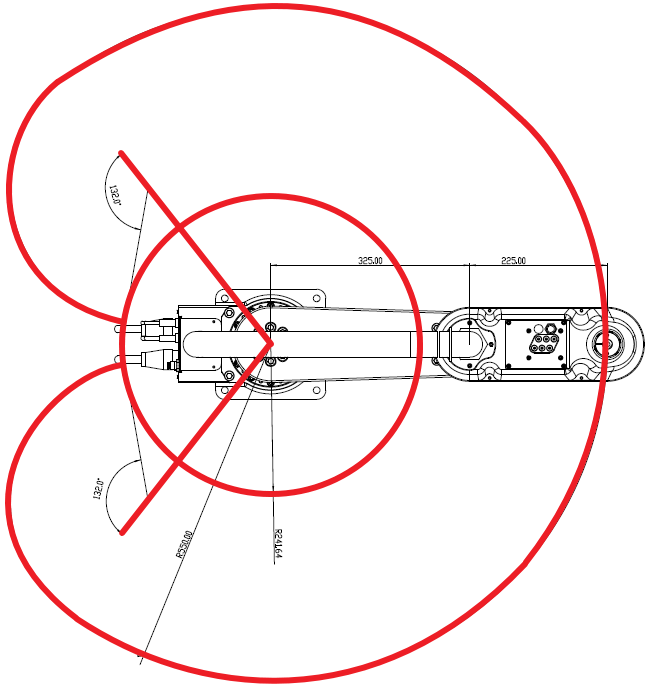
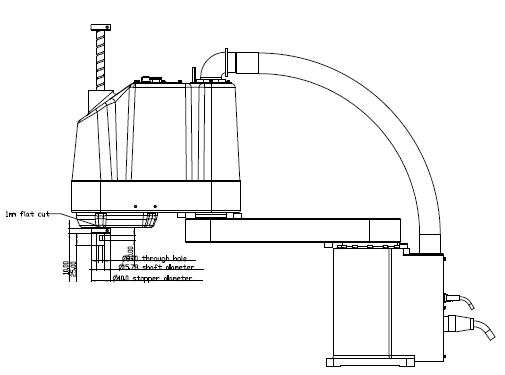
 

圖2 廣明光電SCARA機器手臂(左)與Pioneer P3DX行動機器人(右)

3.1 物件偵測器與色彩索引長條圖

物件偵測器最有名的是Dalal and B. Triggs [4]所提出的行人偵測器，使用具有方向的梯度直方圖(histogram of oriented gradient, HOG)特徵進行行人的偵測，其演算法在靜態影像的行人偵測效能不錯。HOG的概念是將偵測框架中的梯度資訊依照方向做累積，以統計的方式紀錄行人不同部位具有邊緣的資訊。並經由大量樣本的支持向量機(support vector machine, SVM)分類器訓練，使分類器學習複雜環境下行人與非行人的特徵。使用行人偵測器所偵測到行人方框如圖3所示，此方框顯示行人所在的影像範圍。其次，使用尋邊演算法的物件偵測器偵測貨品外框(contour)，如圖4所示為餅乾包裝影像的界線圖，去掉內部的界線即是餅乾包裝貨品的外框。本計畫將使用類似概念的物件偵測器，框出貨品物件的外框，再使用色彩索引長條圖[8-9]描述貨品的色調特徵。

使用色彩索引長條圖描述貨品所呈現的色調特徵，可以強健地描述貨品外貌特徵，也可以達到貨品特徵偵測時的尺度不變與方向不變的效果，提高偵測的成功率。不同貨品具有其獨特的色彩索引長條圖，如圖5(左)所示。色彩索引長條圖偵測特徵時的尺度與方向不變效果，如圖5(右)所示，不同距離與大小的物件，具有幾乎相同的色彩索引長條圖。

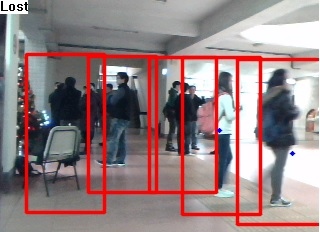


圖3 物件偵測器應用於偵測行人

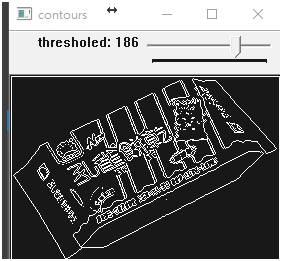
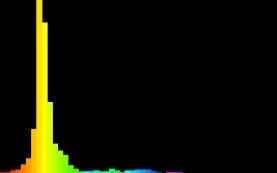
 

圖4 物件偵測器應用於偵測貨品

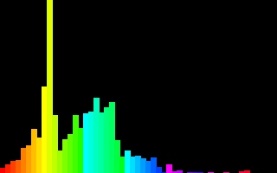
   

圖5 不同貨品的色彩索引長條圖(左)，特徵偵測時的尺度與方向不變(右)

3.2 均值偏移與連續自適應均值偏移演算法

均值偏移(MeanShift)演算法利用一個自訂的固定搜索視窗，計算出物件影像的形心位置。再將原本視窗的形心移動到計算出來的視窗。根據形心計算方程式(1)與高斯核心(Gaussian kernel)方程式(2)，可以計算出位移量*m*。重複運算直到*m*收斂到一個預設的門檻值，

(1)

其中*N*(*x*)是鄰近的點，常數函數*K*計算為

(2)

其中*c*為預設常數值。

MeanShift的缺點是其視窗大小是固定的。當物體移動靠近或遠離鏡頭，無法根據物體大小做出適當的放大或縮小。連續自適應均值偏移(CamShift)則利用MeanShift的運算式，再按照比例進行縮放視窗大小。程序是先利用MeanShift進行搜尋，然後再放大*s*倍率

(3)

其中*M*00為原第一張視窗的大小。最後再計算出最匹配的方向。在OpenCV中有提供CamShift的函式，以及利用色彩索引直方圖來統計視窗中的顏色。需修改OpenCV程式，以整合物件偵測器、色彩索引長條圖、CamShift演算法。

3.3 RGB-D視覺感測器

本研究將使用微軟所發展的Xbox 360 Kinect感測器，做為偵測與追蹤貨品的視覺感測器，以及做為量測貨品三維座標的工具。Kinect外觀如圖6所示，規格如表1所列。Kinect感測器具備感測RGB彩色影像與深度D影像的模組，截取的RGB與深度D影像如圖7所示。RGB彩色影像可以利用物件偵測器偵測貨品，再使用色彩索引長條圖描述該貨品，最後使用連續自適應均值偏移(CamShift)演算法追蹤貨品。而RGB彩色影像搭配深度D影像，可以量測貨品在空間中的三維座標。

使用RGB-D感測器量測貨品在空間中的三維座標**m**，可以表示為

 (4)

其中***r***為Kinect攝影機在世界座標系的位置向量，***h***C為貨品在攝影機座標系的視線(ray)向量，為世界座標系轉到攝影機座標系的旋轉矩陣。貨品的三維視線向量表示為

 (5)

貨品在世界座標系的三維座標量測程序為：首先使用Kinect擷取RGB影像與偵測貨品。假設偵測到貨品特徵中心在影像平面上之座標，使用Kinect的影像扭曲模型與修正模型[18]，可以得到與。再利用透視投影關係求算出貨品的攝影機視線向量，成為：

 (6)

其中與為攝影機方向與方向的焦距，單位為像素；為影像平面中心點座標；即是該像素的Kinect深度訊息，可由Kinect的深度D影像取得。因此，當取得貨品之視線向量後，將方程式(6)代入方程式(4)並做整理，即可求算出貨品位於世界座標的三維位置向量：

 (7)

 (8)

 (9)

其中、。、與為攝影機相對於世界座標軸、與的轉動角度，這些係數構成旋轉矩陣的元素[18]。

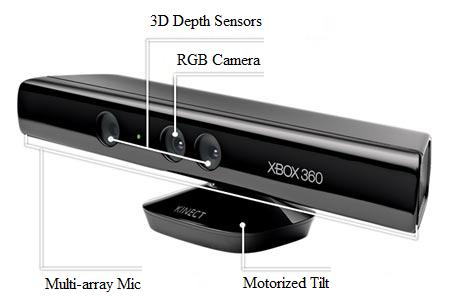


圖6 Xbox 360 Kinect感測器

表1 Kinect規格表

|  |  |
| --- | --- |
| 顏色與深度 | 1.2 ~ 3.6 公尺 |
| 骨架追蹤 | 1.2 ~ 3.6 公尺 |
| 視野角度 | 水平 57 度、垂直 43 度 |
| 底座馬達旋轉 | 左右各 28 度 |
| 每秒畫格 | 30 FPS |
| 深度解析度 | QVGA (320 x 240) |
| 顏色解析度 | VGA (640 x 480) |
| 聲音格式 | 16KHz，16位元 mono pulse code modulation (PCM) |
|  | 四麥克風陣列、24位元類比數位轉 (ADC)、雜音消除 |

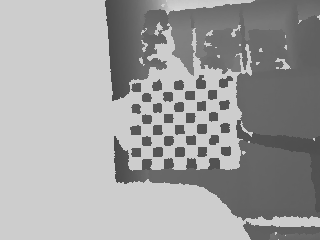
 

圖7 Kinect擷取的RGB影像(左)與深度影像(右)

3.4 貨品位置與方向估測

工作平檯與影像平面皆為平面，可以使用平面投影轉換方法[19]，量測與估測貨品在工作平檯上的位置與方向，此方法也稱為平面之間的單應性量測。而貨品在工作平檯上會有個高度，本研究也規劃高度量測的方法。以下說明如何將單應性矩陣(homography matrix)應用在視覺量測模型中，以及立體物件的位置與方向的估測。

首先選取工作平檯與影像平面上四邊形相對應的頂點，如圖8所示。其中工作平檯為標示在攝影機座標系中的平面，影像平面則是所擷取影像的像素平面座標。工作平檯平面上點的座標向量為***X***，相對應之影像平面上點的座標向量為***x***，假設工作平檯與影像平面存在線性的轉換關係為

 (10)

其中***R***稱為單應性矩陣；假設影像平面上某點的座標向量為***x***=(*x*,*y*,1)，其中1為虛擬參數(dummy parameter)，使得座標向量***x***成為3×1的向量；工作平檯上相對應點的座標向量為***X***=(*X*,*Y*,*Z*)，假設工作平檯的*Z*軸座標為定值*Z*=*W*，在單應性轉換關係中設為權重值。將方程式(10)改寫為

 (11)

矩陣乘開後成為

 (12)

 (13)

 (14)

令*a*33=1，再將方程式(14)代入方程式(12)與(13)，整理之後可以得到

 (15)

 (16)

假如工作平檯與影像平面上有四個已知的對應點***X***與***x***，例如使用圖9所示的工作平檯上四個螺孔位置，則利用方程式(15)與(16)可以求算出單應性矩陣***R***中8個未知的元素。將四個已知對應點座標代入方程式(15)與(16)，並將單應性矩陣***R***的8個未知元素提出為向量***a***，成為

 (17)

或表示為

 (18)

假如方程式(18)中矩陣***A***為非奇異矩陣，即其行列式值不為0，則可以求算出單應性矩陣***R***的8個未知元素，。假如矩陣***A***的行列式值為0，則可以使用虛擬反矩陣求算，。

可以根據兩個平面上已知的一對對應的四邊形求出單應性矩陣***R***，進而對於影像平面的每一點，都可以使用單應性矩陣***R***求得其在攝影機座標系中相對應平面上的一個點。

另外，由於立體物件在影像平面上無法判斷出高度，會將物件位置辨識錯誤，故需要求算出物件真實的位置，才可讓機器手臂順利夾取貨品，如圖10所示。藉由方程式(19)可以求算出立體物件真實的視線向量，

 (19)

其中，參數*H*為實際Kinect攝影機高度，參數*h*為貨品高度，參數*L*為根據影像計算出的貨品在工作平檯上的距離，*l*為立體貨品在工作平檯上真實的距離。



圖8 攝影機座標平面與影像平面關係圖



圖9 四個螺孔的相對位置



圖10 影像裝置、目標物與工作平檯高度關係圖

3.5 機器人與視覺聯網

本研究將使用網連技術規劃機器人與視覺聯網，如圖11所示，將IP攝影機、USB攝影機、Kinect等不同形式的視覺感測器，透過不同的網路介面連接起來。再將各個視覺感測器所感測的訊息整合，藉由視覺訊息輔助機器手臂進行偵測、辨識與挑揀貨品。

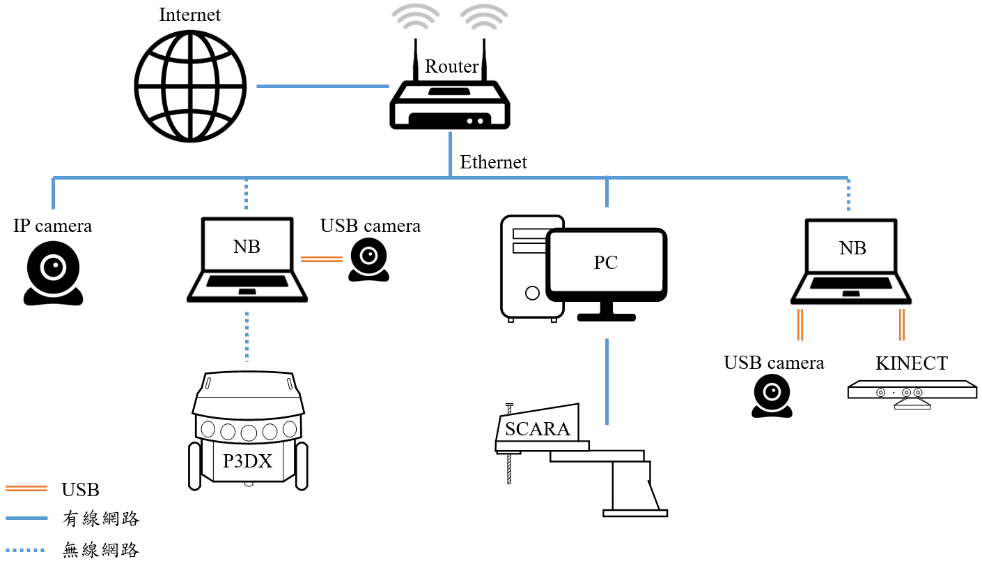


圖11 機器人與視覺聯網(internet of robotics and vision)

4. 預期結果

本研究計畫將結合多種物件偵測器、色彩索引長條圖、連續自適應均值偏移演算法等，發展攝影機偵測與追蹤多個移動貨品的系統，輔助機器手臂執行貨品挑揀的任務。預計完成下列研究項目：

(a) 架設Kinect RGB-D視覺感測器系統與完成感測器校準；

(b) 撰寫影像物件偵測器程式，並且使用色彩索引長條圖描述貨品的影像特徵；

(c) 撰寫與測試均值偏移與連續自適應均值偏移演算法，應用在偵測與追蹤貨品；

(d) 架設機器人與視覺聯網，傳遞視覺訊息與手臂任務規劃訊息；

(e) 執行貨品位置與方向估測，提供機器手臂端效器夾取貨品路徑規劃之用；

(f) 系統整合與測試，實現視覺輔助機器手臂執行貨品挑揀的任務。

5. 參考文獻

* 1. Kiva Robot, http://holeaker.blogspot.tw/2013/12/best-investmet-that-amazon-made.html, December 19, 2013.
  2. Amazon Picking Challenge, http://amazonpickingchallenge.org/ (accessed on December 19, 2015)
  3. W. Choi, C. Pantofaru, and S. Savarese, "A General Framework for Tracking Multiple People from a Moving Camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 35, pp. 1577-1591, 2013.
  4. N. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2005.
  5. V. Ferrari, M. Marin-Jimenez, and A. Zisserman, “Progressive Search Space Reduction for Human Pose Estimation,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2008.
  6. P. Viola and M. Jones, “Robust Real-Time Face Detection,” Int’l J. Computer Vision, vol. 57, no. 2, pp. 137-154, 2003.
  7. M. Jones and J. Rehg, “Statistical Color Models with Application to Skin Detection,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1999.
  8. Swain, M.J. and D.H. Ballard, Indexing via color histograms, Proceedings of Third International Conference on Computer Vision, pp.390-393, 1990.
  9. Swain, M.J. and D.H. Ballard, Color Indexing, International Journal of Computer Vision, vol.7, no.1, pp.11-32, 1991.
  10. Mean shift, https://en.wikipedia.org/wiki/Mean\_shift
  11. Y. Cheng, Mean Shift, Mode Seeking, and Clustering, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.17, pp.790-799, 1995.
  12. D. Comaniciu, P. Meer: Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, pp.603-619, 2002.
  13. D. Comaniciu, V. Ramesh, P. Meer: Kernel-based object tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.25, pp.564-577, 2003.
  14. Y.A. Ghassabeh,"A sufficient condition for the convergence of the mean shift algorithm with Gaussian kernel". Journal of Multivariate Analysis, 135: pp.1-10, 2015.
  15. Bradski, G.R., "Real time face and object tracking as a component of a perceptual user interface," Proceedings of Fourth IEEE Workshop on Applications of Computer Vision, (WACV 1998), pp.214-219, 1998.
  16. Bradski, G.R. Computer Vision Face Tracking For Use in a Perceptual User Interface, Intel Technology Journal, No. Q2. (1998)
  17. E. Emami, "Online failure detection and correction for CAMShift tracking algorithm".Iranian Conference on Machine Vision and Image Processing (MVIP 2013), vol.2, pp.180-183, 2013.
  18. Chi, C.-T., Y.T. Wang, S.-T. Cheng, and C.-A. Shen, Robot Simultaneous Localization and Mapping Using a Calibrated Kinect Sensor, *Sensors and Materials*, Vol.26, no.5, pp.353-364, 2014.
  19. A. Criminisi and I. Reid and A. Zisserman, A plane measuring device, Image and Vision Computing, vol.17, pp.625-634, 1999.