

Hand Tracking and Gesture Recognition for Interactive Games

應用於互動遊戲手勢追蹤與辨識之研究

Shih-Pin Chao (趙士賓)

工研院電子與光電研究所

Email: chaoshihpi@itri.org.tw

Cheng-Yuan Tang (唐政元)

華梵大學資訊管理學系

Email: cytang@cc.hfu.edu.tw

Yi-Leh Wu (吳怡樂)

國立台灣科技大學資訊工程學系

Email: ywu@csie.ntust.edu.tw

Maw-Kae Hor (何瑁鎧)

國立政治大學資訊科學系

Email: hor@cs.nccu.edu.tw

Chien-Ming Huang (黃建旻)

華梵大學資訊管理學系

Email: jack11260@gmail.com

摘要—本文主要是開發一套影像為基礎的手部追蹤與手勢辨識系統，目的希望以低成本且能在複雜環境下使用的手勢追蹤與辨識系統，並應用於互動遊戲上。系統主要分為偵測、追蹤與辨識三個部份，利用前景偵測與臉部偵測將背景與畫面中的臉部消去，以減少干擾。追蹤模組則是利用 CAMSHIFT 演算法追蹤手部區域。辨識模組則是將手部區域輪廓利用 PGH (Pairwise Geometric Histogram) 作形狀直方圖統計，並與預先建立的模板作形狀直方圖比對，以判斷使用者的手勢。

在實驗中，實驗了 PGH 對於縮放與旋轉的適應性，以及 PGH 用於模板比對的效果。最後將系統應用於猜拳遊戲上，透過攝影機的擷取手部的移動，判斷使用者是否出拳，並辨識使用者的手勢來出拳。

關鍵詞—手勢辨識、Pairwise Geometric Histogram、互動式遊戲、手部追蹤

Abstract -- In this paper, an interactive system having hand-tracking and gesture recognition has been developed in the complex environment. Our developed interactive system can be applied in an interactive game. Our system consists of three modules, such as the foreground detection module, tracking module and recognition module. The foreground detection module is based on the codebook algorithm. The purpose of the foreground detection module is to remove some portions of the image background. It will reduce the interference caused by the background for tracking and recognition. The CAMSHIFT algorithm utilizing the color of the target object for tracking is used in the tracking module. In the recognition module, the Pairwise Geometric Histogram (PGH) of the hand contour is used to describe shape statistics values into a histogram. Finally, the most similar gesture is

found by comparing the PGH shape histograms which consist of the PGH obtained from the hand contours and the PGH obtained from the gesture template contours.

For PGH, in the experiment, we tested the adaptive scaling and rotation, and the effect of template matching. Our developed system is applied to the interactive game named the finger-guessing game. We use a Webcam to track the movement of hands, determine whether the user selects or not, and recognize the hand gesture of the user who is playing the game.

Keywords: Hand gesture recognition, Pairwise Geometric Histogram, Interactive Game, Hand tracking

一、緒論

隨著科技進步，電腦運算速度也逐年成長，人機互動在近幾年來也開始出現了變化，以往的鍵盤、滑鼠等傳統的輸入設備發展到現在的多點觸控螢幕、動作感應控制、攝影機等，帶給人們更直覺、更容易學習的人機介面。以往使用者必須學習如何操作電腦，使用者無法很自然直接的與電腦互動，發展至今演變為由電腦來配合使用者的使用習慣，使用者能依自己的習慣來操作電腦，能很直覺、簡單的與電腦互動。為了達到這樣的目的，人機介面必須以人的觀點出發，合乎人類行為、更多元、更有效率的方向發展。

為了使人機介面朝向更有效率、更多元化，更合乎人性之方向邁進，縮短使用者與機器間的距離，必須使用人類時常使用的溝通方式來作為與機器溝通的模式，而手勢即是人類的一種相當原始與自然的一種示意方式；在日常生活中，手勢是人與人的互動中常用的溝通方式，對於語言障礙、聽覺障礙人士，也有發展手語系統，取代聲音作為溝通的工具，在將來，手勢輸入將會廣

泛應用在人與電腦之間溝通的介面上。

在遊戲娛樂方面，自從日本任天堂公司推出了Wii遊戲機，其特殊的動作感應設計，使得遊戲玩法產生了很大的變化，與傳統的遊戲相比，動作感應遊戲有著更簡單直覺的遊戲方式，玩家只要照現實中的經驗去作動作揮動，即可與遊戲作互動，使得一般人更能容易上手。這股風潮也使得其競爭對手Microsoft與Sony於2009年相繼發表該公司遊戲平台所對應的動作感應設備，更說明了動作感應遊戲在遊戲市場所帶來的趨勢。

二、研究方法

本研究發展一個手部追蹤與手勢辨識系統，主要分為偵測模組、追蹤模組與辨識模組。偵測目的在於將畫面中背景部份與使用者的臉部先去除，以減少背景環境與臉部膚色對於追蹤與辨識的干擾。追蹤模組則是利用欲追蹤物體的顏色作為追蹤依據，將畫面中符合顏色區間的區域分離出來，以追蹤手部的區塊。辨識模組則是將手部區域的輪廓描繪出來，並對輪廓做形狀的統計直方圖，並對手勢模板所計算出的形狀直方圖做比較，以計算出最符合的手勢。

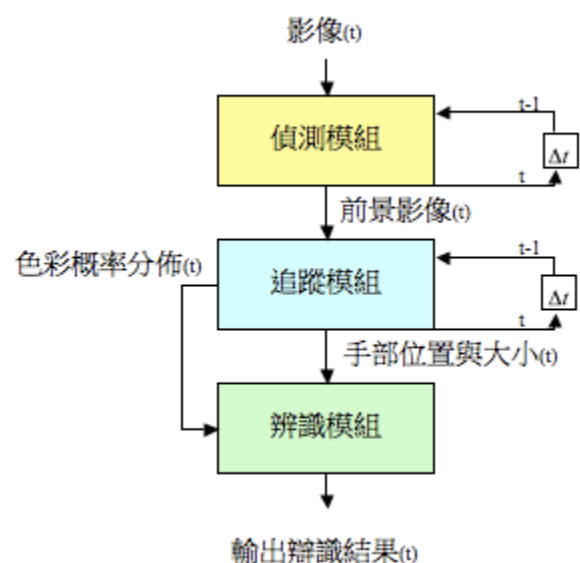


圖 2-1 系統架構流程圖

2.1 偵測模組

在手部追蹤時，手部的移動相對於畫面中其他部份的變化量大得多，使用前景偵測可以將手部與背景分離出來，排除非前景部份以減少環境的干擾。對於畫面中使用者的臉部的影響，則是使用臉部偵測找尋使用者的臉部位置，並對該位置作遮蔽，以消除臉部對於後續追蹤與辨識運算的影響。

前景偵測方法採用以 CodeBook 為基礎的前景偵測演算法，一般 CodeBook 演算法需要事先做訓練，並且在偵測時環境若有改變，就必須重新訓練，而 Tang[2]所提出的方法可以在偵測的狀態下對場景做學習，可以免除事前訓練，並且能克服環境變化，以及對於場景中變化量小的區域作學習，能夠將不需要關注的部份去除。臉部追蹤方法我們則是直接使用 Viola 提出的物件偵測方法，來對使用者的臉部作偵測。

2.2 追蹤模組

追蹤模組主要是使用 CAMSHIFT (Continuously Adaptive Mean Shift) 演算法 [3]，以顏色為追蹤根據，透過物體的顏色來追蹤觀測影像中的物體，為一簡單有效率的追蹤演算法。如圖 3-4 所示，將前景偵測的結果影像輸入，計算前景影像的膚色分佈，再對最大的膚色區塊使用 CAMSHIFT 作追蹤，以判斷手部位置。

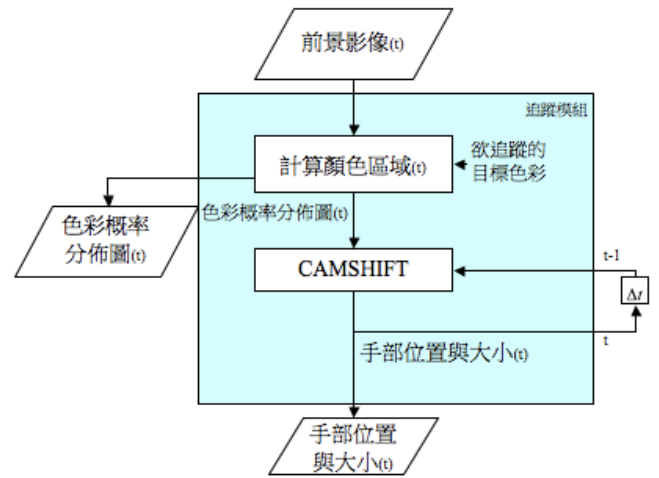


圖 2-2 追蹤模組流程圖

CAMSHIFT 為常見的演算法，網路上也有許多開放的程式碼可供使用，本研究所使用的是 OpenCV (Open Source Computer Vision Library) 中的 CAMSHIFT 函式來實作。

2.3 辨識模組

辨識模組針對追蹤模組所計算出的位置，在觀測影像上切出一個關注區域 (Region of Interest, ROI)，計算膚色的形狀，進行輪廓的計算，隨後對輪廓使用 PGH (Pairwise Geometric Histogram) [4] 描述其輪廓的形狀，統計出一個 2D Histogram，並與預先建立好的手勢模板所計算出的 PGH 做相似度的計算，計算出觀測影像與哪個手勢模板最為相似，以判斷使用者的手勢。



圖 2-3 觀測影像的輪廓與比對之模板

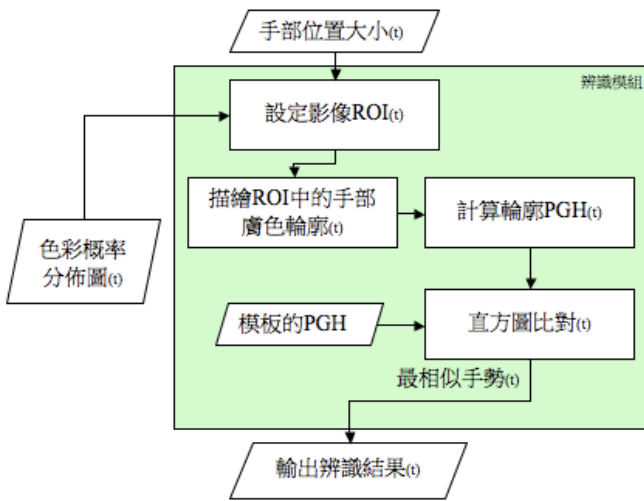


圖 2-4 辨識模組流程圖

2.3.1 成對幾何直方圖

參考吳松航[1]所提出的手部追蹤方法，對於觀測影像旋轉與縮放的問題，以建立多個模板的方法來克服，但多個的模板比對會使得計算時間增加，且會隨著所使用的手勢數量而倍增，並且在不同手勢的狀況下，會出現模板邊緣點數量不同，而對模板比對時以距離轉換計算出的數值不同，進而影響到模板比對的結果。所以本研究對於模板比對，則是選擇可以克服旋轉與縮放且快速的方法。

對於形狀描述，具有良好的旋轉、平移與縮放的適應性為一個非常重要的衡量方式，參考Iivarinen et al.[4]的研究，相較於The Chain Code Histogram (CCH)，依物體的輪廓線的方向，萃取出鏈碼 (chain code)，依據鏈碼統計為直方圖來對形狀做描述，Pairwise Geometric Histogram (PGH) 使用線段間的幾何特性對形狀做描述，包括了線端間的相對角 (Relative Angle) 與垂直距離 (Perpendicular Distance)，藉由這兩種特徵所統計出的二維直方圖來表示所描述的形狀，使得PGH對於不規則的輪廓形狀有著較CCH更精確的描述能力，以及對於旋轉有更好的適應性，但在計算量以及記憶體佔用上則較CCH來的大。

假設由邊緣點 (edgepoint) $(\tilde{x}(t), \tilde{y}(t)) \in R^2$ 定義的一個輪廓，由連續邊緣點定義為多個輪廓中的線段，PGH 的計算步驟如下：取某一線段作為參考線 L_{ref} ，計算其他線段 L_i 與參考線間的相對角 θ ($\theta \in [0, \pi]$) 與垂直距離的最短距離 d_{min} 與最長距離 d_{max} (圖 2.5 (a))，並將結果作二維直方圖統計 (圖 2.5 (b))，重複此步驟直到所有線段皆作為參考線為止。統計結果即為 Pairwise Geometric Histogram。

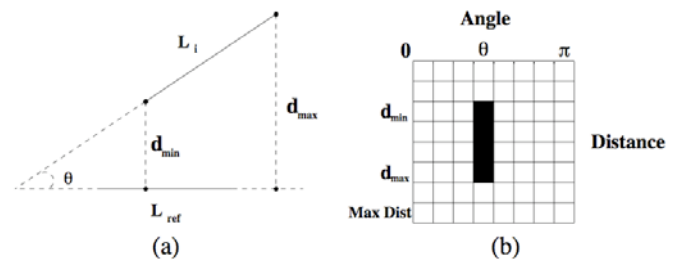


圖 2-5 (a) 兩線段間的相對角度與垂直距離。
(b) Pairwise Geometric Histogram。

計算手部區塊的輪廓，為了讓 PGH 比對有一定的基準，以保持 PGH 直方圖的準確度，我們取輪廓周長的 1% 作為每一個邊緣線段的長度，在不同輪廓形狀中取出相同數量的線段，以確保比對正確性。

2.3.2 模板比對 (template matching)

計算出觀測影像輪廓的 PGH 後，將其直方圖與預先建立好的模板 PGH 作直方圖的比對，直方圖比對使用的是 Bhattacharyya distance[5][6] 作直方圖的差異度計算，假設 H_1 為觀測影像的直方圖， H_2 為預先建立好的模板直方圖，其計算公式如下：

$$d(H_1, H_2) = \sqrt{1 - \sum_i \frac{\sqrt{H_1(i) \cdot H_2(i)}}{\sqrt{\sum_i H_1(i) \cdot \sum_i H_2(i)}}$$

若是相同的手勢，其直方圖所計算出的 Bhattacharyya distance 會相對的比其他模板所計算出來的結果要來的小，如此一來即可判斷出與當前影像匹配的模板。判斷輸出手勢時，我們連續觀察數個畫面，計算出現最多次數的手勢即判斷為最正確的手勢，如此一來可以增加偵測結果的正確性。

三、 實驗結果

3.1 模板比對

在觀測影像中，我們將 CAMSHIFT 計算出的位置與大小，在膚色的色彩概率分佈圖上切出一個關注區域 (Region of Interest, ROI) 並描繪出該區域的輪廓，但在實際觀測影像時，由色彩概率分佈圖取出的手部膚色通常有些許破碎不完全的部份，對於輪廓描繪與 PGH 的計算會有很大的影響，我們在這先將色彩概率分佈圖利用高斯金字塔 (Gaussian pyramid) 將影像模糊化後，再用設定的門檻將影像值二元化，取得較為完整的膚色區塊。



圖 3-1 由左至右分別為原始色彩概率分佈圖、使用高斯金字塔模糊化、將灰度值二元化的結果

但此方法由於對會對所有的區域作擴展的效果，所以對於雜訊也會相當敏感，對手部形狀產生影響，使得 PGH 計算產生錯誤。所以在前景偵測部份若能有效的消除雜訊，則能減少雜訊所造成的影響。

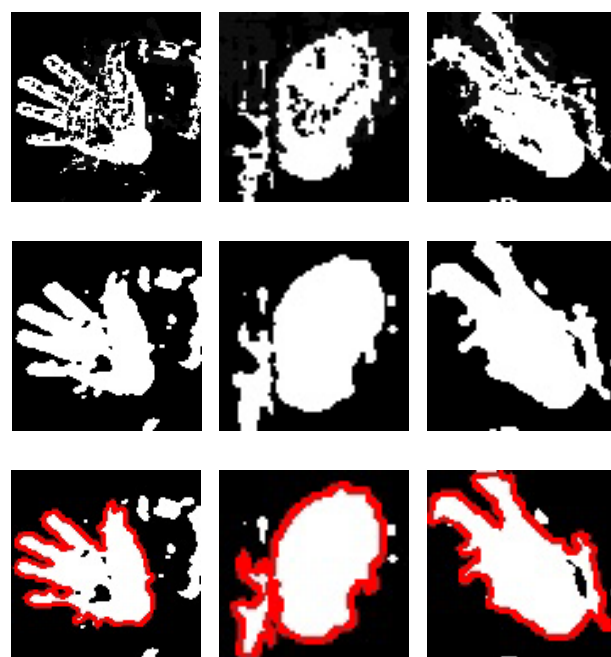


圖 3-2 經過模糊化後，雜訊放大造成的影響與描繪出的錯誤輪廓

隨後描繪手部區域中的區塊輪廓，並保留最大包圍面積的輪廓作為手部的輪廓。對輪廓作 PGH 的計算，並與預先建立好的模板作直方圖比對，結果如下：

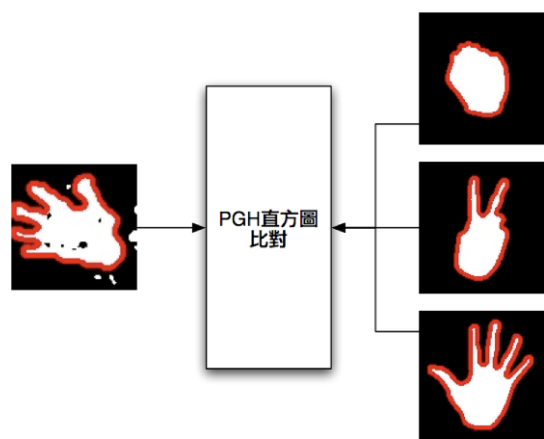


圖 3-3 模板與觀測影像的輪廓 PGH 作直方圖比對

這裡我們分別計算觀測影像與模板的輪廓，並計算成 PGH 後，觀測影像與模板使用 Bhattacharyya distance 作直方圖的差異度計算，

結果如下：

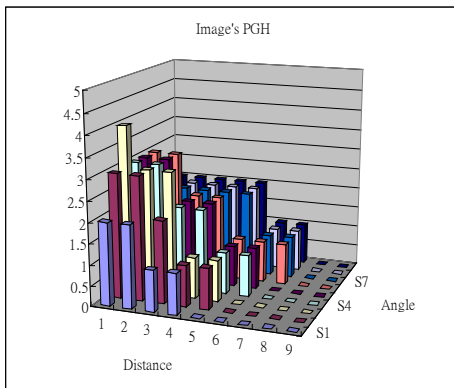


圖 3-4 觀測影像的 PGH

將上圖所示的觀測影像 PGH，與三個手勢的模板 PGH 作直方圖的差異度比對，比對結果如下：

表 3-1 模板與觀測影像直方圖差異度 Score

	Scissor	Rock	Paper
差異度	0.150	0.176	0.06

為了驗證模板的差異性，我們將模板的輪廓描繪出來，互相比對模板的形狀，由於不同模板的輪廓長度不同，取出的線段數量也可能不同，為了統一線段數量以確保模板直方圖比對正確，我們取出輪廓的週長的 1% 作為每個線段的長度，使得每個模板輪廓皆有相同數量的線段來作 PGH 統計。比較結果如下：

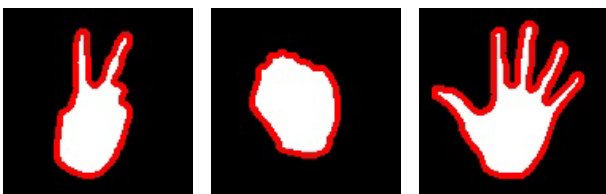


圖 3-5 三個手勢模板與描繪出的輪廓

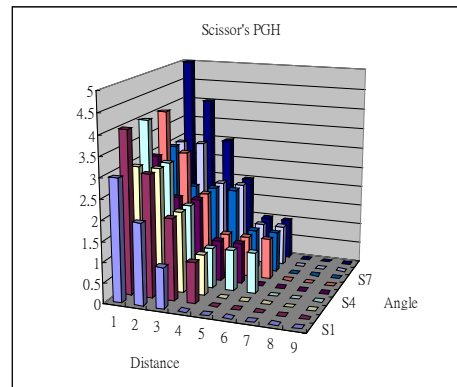


圖 3-6 「剪刀」模板的 PGH

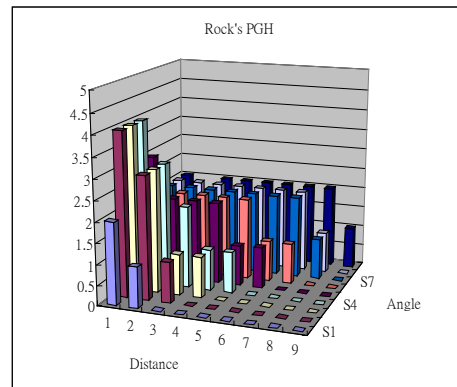


圖 3-7 「石頭」模板的 PGH

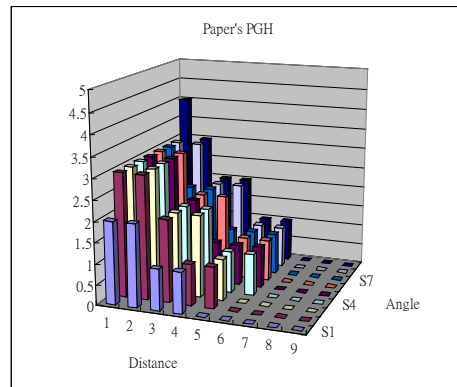


圖 3-8 「布」模板的 PGH

表 3-2 模板間形狀差異度

	Scissor	Rock	Paper
Scissor	0	0.217	0.121
Rock	0.217	0	0.179
Paper	0.121	0.179	0

在模板比對時，我們使用 Pairwise Geometric Histogram 來作形狀的描述，此方法能夠克服旋轉與縮放的問題，不需要建立多個大小、角度不同的模板，即可有效的比對出正確的模板。

為了驗證此方法對於縮放、旋轉皆有良好的適應性，實驗中我們建立了，「剪刀」、「石頭」、「布」三個不同手勢的模板，另外建立了作為觀測影像的模板，以每十度為一個單位作順時鐘旋轉，一個手勢共建立 36 個不同角度的影像，並針對這些影像作 PGH 計算，並與先前建立的模板作 PGH 的差異度作計算。

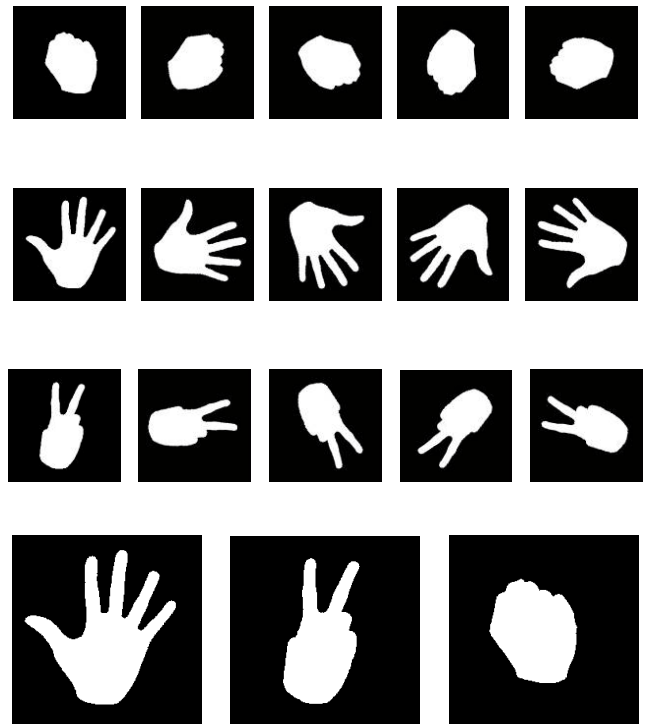


圖 3-9 上三列影像為實驗所使用的觀測影像的一部分，下列為預先建立三個手勢的模板

我們將比對結果統計成圖表，圖表橫軸代表觀測影像所旋轉的角度，縱軸則表示觀測影像與模板間的 PGH 差異度，而線條顏色代表不同的模板所比對結果的分佈。

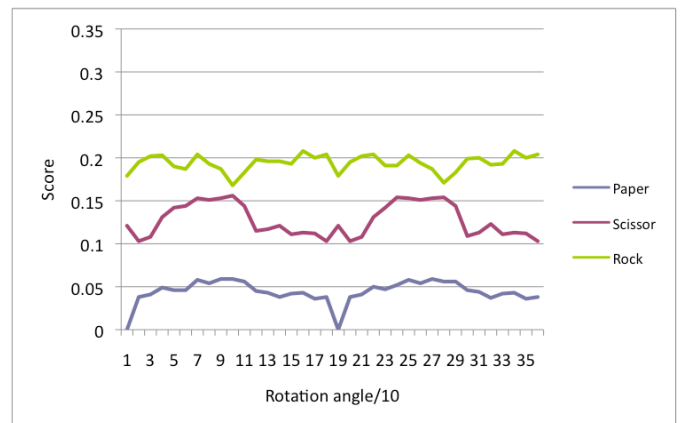


圖 3-10 以「布」作為觀測影像的比對結果

如圖 3-10 所示，以「布」作為觀測影像時，「布」模板 PGH 與觀測影像比對的差異度較其他模板與觀測影像比對的差異度要來得低，而「剪刀」模板的差異度也低於「石頭」模板，表

示其形狀差異度較「石頭」來的小。

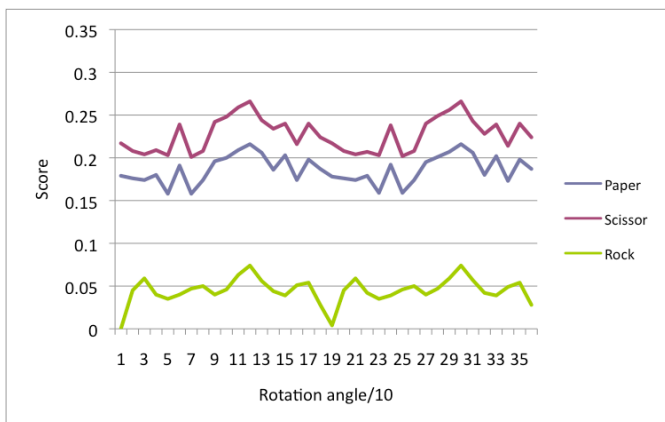


圖 3-11 以「石頭」作為觀測影像

圖 3-11 以「石頭」作為觀測影像時，與「石頭」模板 PGH 比對出來的差異度遠低於其他兩個模板，其他兩個模板的差異度有著較為相似的差異性，我們比較三個手勢所統計出的 PGH，相較於「石頭」模板的形狀來說，「布」與「剪刀」的形狀有著相似的分佈，因此比對結果會有著相似的差異性。

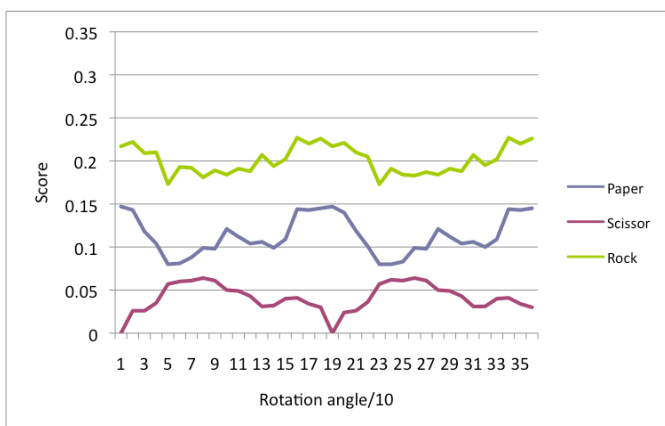


圖 3-12 以「剪刀」作為觀測影像

當以「剪刀」作為觀測影像時，「剪刀」模板的差異度也遠低於其他兩個模板。

以上實驗顯示，PGH 用於模板比對時，觀測影像旋轉與縮放時雖然差異度會有些許的變化，但對在不同形狀間還是能夠有效的作判別，可以認為 PGH 用於模板比對能夠克服旋轉的變

化。

對於縮放的適應性，我們將模板由原先的 240*240 像素分別放大為 480*480 像素和縮小為 120*120 像素，再依序作 PGH 直方圖比對，結果如下表：

表 3-3 原始模板 PGH 與放大後模板 PGH 比對的差異度

	Scissor	Rock	Paper
Scissor_480	0.058	0.174	0.107
Rock_480	0.201	0.05	0.155
Paper_480	0.129	0.166	0.027

表 3-4 原始模板 PGH 與縮小後模板 PGH 比對的差異度

	Scissor	Rock	Paper
Scissor_120	0.043	0.216	0.105
Rock_120	0.227	0.038	0.194
Paper_120	0.123	0.191	0.026

如表 3-3、3-4 所示，模板經過縮放後的 PGH 差異度還是低於一定程度，且還能保有與其他模板的鑑別能力，因此可以認為 PGH 對於形狀的縮放還能夠保有一定的描述能力。

另外，我們實驗了六種未定義的手勢，建立與模板大小相同的以下六種手勢作為觀測用影像，與預先建立好的「剪刀」、「石頭」與「布」

三種手勢模板作比對，以觀察使用者擺出未定義手勢的計算結果，對個別手勢作說明，實驗結果如下：

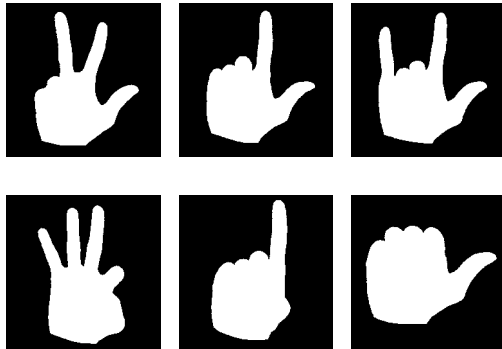

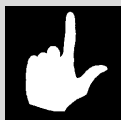






圖 3-13 實驗用的六種未定義手勢

表 3-5 未定義手勢與已定義手勢模板比對結果

	Scissor	Rock	Paper
	0.064	0.209	0.083
	0.11	0.154	0.069
	0.096	0.17	0.053
	0.049	0.214	0.095

	0.096	0.179	0.1
	0.173	0.118	0.110

結果如上表所示，在未定義的手勢部份，模板比對出來後我們選擇與三個模板中最相似的模板，認定為使用者所比出的手勢，若是要判斷使用者所比出的手勢是否為定義過的手勢，則必須觀察每個手勢的模板差異度數值變化，對每個手勢設定一門檻值，才能夠達到認定是否為已定義手勢的效果，但由於門檻並不容易選擇，尤其有大量的手勢模板時，更加難以界定。所以我們未來可以加入 PCA 等方法對手勢作分析，排除掉差異過小的模板，保留手勢差異較大的模板，以維持手勢的差異性。

3.2 應用於互動遊戲

本研究將發展出的手勢追蹤與辨識系統應用於互動遊戲上，範例中將系統應用於網頁的 Flash 所製作的猜拳遊戲上，系統辨識使用者手部的軌跡方向，如同一般人猜拳的習慣，若手部向下揮動，就判斷為使用者出拳，並判斷使用者所比出的手勢，視為使用者所要出的拳，並為使用者選擇，以達到簡單的遊戲互動。

下圖我們是用 ENIUM 公司所發行的 Flash 猜拳遊戲，作為我們的互動遊戲範例，此遊戲規則為玩家與電腦猜拳，勝利者會攻擊失敗者，若先取得三次勝利，則會將對手打入河流中，贏得最後的勝利。

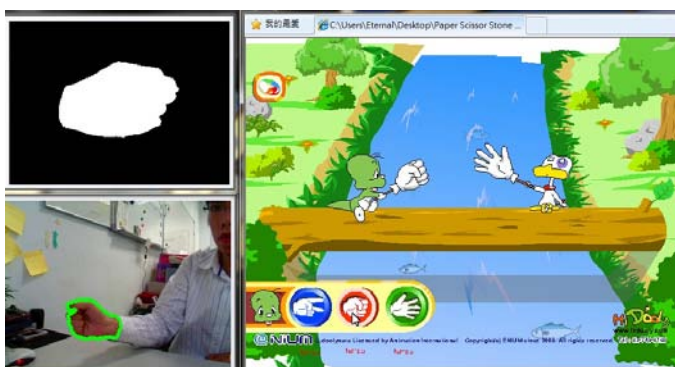


圖 3-14 手部追蹤與手勢辨識系統應用於互動遊戲

四、結論與未來方向

4.1 結論

CAMSHIFT 計算，以找出手部的大略位置。再利用手部的輪廓計算出手部的形狀，並與預先建立好的手勢模板進行比對，達到偵測手掌的位置與手勢的辨識。

對於 CAMSHIFT 演算法，背景與臉部膚色的影響會導致追蹤錯誤的情形，可整合光流法等其他追蹤演算法以物體追蹤的效果與正確性。

找尋手掌輪廓部份，僅利用膚色部份找尋手部形狀輪廓稍有不足，會因光線陰影等影響造成膚色區域的不完整破碎，或是受到背景顏色的干擾而產生計算誤差，雖然本研究利用影像模糊化的方式填補細小的區域破碎，但對較大面積的破碎則無法有效的消除，並且會使得雜訊被放大。未來可使用邊緣等形狀資訊，加強對手部形狀的完整性。

4.2 未來工作

1. 希望能解決手部遮蔽的問題，透過其他追蹤器的整合達到更有效率的手部追蹤。
2. 希望能結合到 3D 手的模型。

3. 目前只能對一隻手進行追蹤，未來可以應用到更多雙手追蹤等應用。
4. 對於大量手勢模板，使用 PCA 等方法對手勢分析，排除差異過小的手勢，以維持手勢的差異性。

五、參考文獻

- [1] 吳松航「在複雜場景使用階層式貝式濾波器作手追蹤」，華梵大學資訊管理學系碩士學位論文，民 97。
- [2] C.-Y.Tang, Y.-L Wu, S.-P.Chao, W.-C. Chen, P.-L. Chen, "Anomaly Foreground Detection through Background Learning in Video Surveillance," KES International Symposium IDT, pp. 427-435, 2009.
- [3] Gary R. Bradski, "Computer Vision Face Tracking For Use in a Perceptual User Interface," Intel Technology Journal Q2 '98
- [4] J. Iivarinen, M. Peura, J. Särelä, and A. Visa, "Comparison of Combined Shape Descriptors for Irregular Objects," In Proceedings of the 8th British Machine Vision Conference, volume 2, pages 430-439, University of Essex, UK, Sept. 8-11, 1997.
- [5] A.P.Ashbrook, N.A.Thacker, P.I.Rockett, and C.I.Brown, "Robust recognition of scaled shapes using pairwise geometric histograms," *BMVC*, 503-512, 1995.
- [6] G.R. Bradski, A.Kaehler, "Learning OpenCV," O'REILLY, 201-203, 2008.
- [7] P. Viola, M.J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," *IEEE Conference on Computer Vision and Pattern Recognition*, 2001