

使用可適性 YCbCr 模型之猜拳遊戲系統

Roshambo Game System Using Adaptive YCbCr Color Model

黃樹林 張創然 高國淳

Shu-Lin Huang Chuang-Ran Jhang Guo-Chun Gao

摘要

本論文利用多媒體應用設計出一套能適應在不同光源下,使用對應 YCbCr 色彩模型的一套猜拳遊戲系統。首先利用 usb-webcam 讀取手的影像,手勢辨識方法是使用 Holden 提出的。手勢辨識能否成功主要關鍵是能夠利用膚色模型偵測出手的完整輪廓。因為在不同光源下影像色彩是容易受影響,造成使用固定膚色模型會有所誤差。在本研究中主要探討 Holden 的方法在各種光源下的辨識效能,同時參考多位學者在 RGB、HSV 以及 YCbCr 等不同色彩空間的模型,提出一個可適性膚色模型。經過實驗結果發現這個模型在低光源下有良好的效能,在正常光源下也有不錯的結果。在實驗中會分離出背景找出其 Y 值,根據此數值的範圍使用相對應的膚色模型。最後再利用形態學之閉合(膨脹 1 次侵蝕 3 次)做影像邊緣的修補能把效能再提高。使用可適性膚色模型準確率在低光源下從原來的 50% 提升至 78%,其他光源也達到 85%~92%。最後將研究結果實作成一個機器人的猜拳遊戲系統,機器人根據猜拳結果做出相對應的互動,達到一個有趣、有效率的系統。

關鍵詞：手勢辨識、色彩空間、膚色模型

ABSTRACT

This paper designs a roshambo game system using adaptive YCbCr model under various light conditions. First, using an usb-webcam to capture the image of a hand, the hand gesture is identified through the method proposed by Holden [2]. The critical element of successfully identifying a hand gesture is using a skin color model to detect the complete contour of the hand. As image color is easily affected under different light sources, using a static skin color will yield errors. This study mainly explores the performance of Holden's method under various sources of light and references the RGB, HSV, and YCbCr models in different color spaces by several scholars to propose an adaptive skin color model. Experiment results found that this model performs well under low light and decently under regular light. The experiment separates the background to find its' Y value to utilize the proper skin color model using the value's range. Finally the morphology of the closure (dilate once, erode thrice) is used to repair the border of the image to increase effect. Using adaptive skin color models, performance of low light was increased from 50% to 78% while other light sources achieved accurate rates of 85%~92%. The research results were used to create a robotic roshambo game system. The robot performs a relative interaction to roshambo results to achieve a fun and efficient system.

Keywords: Hand gesture recognition, color space, skin color model

一.前言

本論文實作可以和人猜拳的遊戲系統,主要能夠辨識三種手勢-剪刀、石頭、布。現今的手勢辨識是要在一張影像中快速且正確的搜尋到手部區域位置,此技術大致分成:(1)以手套為辨識方式,需要穿戴感測器手套硬體設備來達成;(2)以視訊為基礎之技術,搭配使用軟體影像處理技術來完成。

如 Wagne 等 [1] 提出一種應用於可攜式手勢介面的手勢特徵擷取方法,將手勢區域範圍分成幾個小區塊,統計每個小區塊中的像素後判斷出手勢。此方法在計算上不需太多手部特徵判斷式,但對於手部範圍的擷取需要有較精確的判定。Holden [2] 提出利用雷達圖掃描方式進行手部區域辨識。曹文潔 [3] 利用膚色偵測找尋手部區域,並使用 Holden 所提出的雷達掃描方式而完成一套互動式猜拳遊

戲。張竣皓 [4] 則先行利用手部的 Haar-Like 特徵，找出雙手位置後進行動態建模，達到可適應性的動態膚色。

上述學者大多使用 Holden 所提出的雷達掃描方式做為辨識的主要方法。搜尋手部區域主要使用膚色偵測，影像膚色判斷最困難的地方，是如何判斷目標物是否受到周圍光線的干擾，而造成判斷上誤判或是目標物無法呈現於影像上。前述學者也都沒有探討辨識物對於環境光源變化造成影像膚色的影響。因為膚色偵測系統的建模，光線的影響是非常大的。所以本研究針對辨識物體在各種不同的光源及不同的色彩空間下進行辨識測試與探討其偵測效能。在本文實驗中分別使用 RGB、HSV 和 YCbCr 做為膚色偵測的色彩空間。膚色模型先根據 Wang [5] 所提出之 RGB 以及 HSV 色彩空間數值，及 Hiremath [6] 所使用 YCbCr 色彩空間數值來進行實驗。也利用環境設計出 7 種不同照度的環境光源進行實驗測試。手勢辨識方法直接利用 Holden 之雷達圖方式進行手指特徵判定。本論文使用開放式影像函式庫 -OpenCV [7] 來實作 Holden 的方法，其函式庫提供大量影像處理函式，能快速實作測試不同的影像處理演算法。

整個研究主要探討 Holden 的方法在各種光源下之辨識效果，找出較佳的環境光源、色彩空間與膚色模型的結合以達到改善偵測效能的目標。從實驗結果發現 Hiremath 的模型在低光源下效果不佳，經過多次實驗發現本研究所提出可適性膚色模型在低光源下能改善 Hiremath 色彩空間的效能，此模型在低光源與正常光源都適用，最後搭配形態學的閉合修補手部影像可讓系統辨識效能可在低光源下達到 80%，正常光源下達到 85%~92%。

本文共分 5 個章節，後面會依序介紹知識背景、實驗方法、實驗結果及結論。

二.背景知識

2.1 Holden 手勢辨識

利用手勢二值影像求出手部重心位置如圖 1(a)，以所求出手部重心為圓心，以雷達掃描畫圓的方式掃描整張影像，半徑由 0 開始逐漸增加，

當圓通過手部點時，則記錄此點的半徑及相對於圓心的角度，圓的最大半徑為手指與圓心最長距離。如此可將手的二值影像(x,y)轉換為一個以圓心為原點的角度距離極座標圖(θ, ρ)來表示如圖 1(b)，橫軸代表角度 θ ，縱軸代表手部重心至手部各點的距離 ρ [2]。

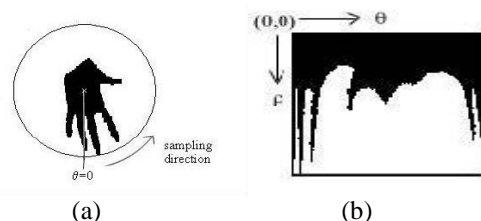


圖 1 手勢二值影像轉換為角度距離極座標圖

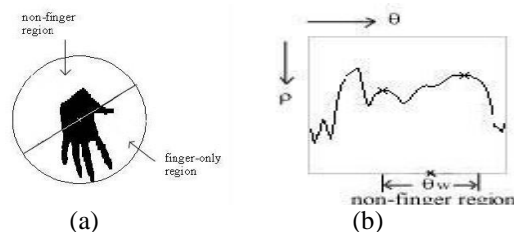


圖 2 找出手勢曲線輪廓，標示手指與非手指區域

如圖 1(b)，可將角度距離極座標圖做邊緣偵測，以獲得如圖 2(b)的曲線圖形，相當於計算由重心點至每個手部輪廓點的距離。接著在手指輪廓曲線圖中找出距離變化較平緩範圍，將此範圍的中心點設為參考點，並假設此點為 θ_w ，根據此點我們可以將手分成手指區域及非手指區域兩部分[2]。圖 1(b)、2(b)中每一個峰值代表一根手指，再由之前所找出的手指區域來判斷峰值數，即可成功辨識出手勢。

2.2 膚色偵測與色彩空間

膚色偵測 [8] 是為了能精確找出手掌位置的影像處理技術。由於一般影像使用 RGB 色彩空間模型，分別代表紅色、綠色及藍色，容易受到環境光源對視訊攝影機的影響。HSV 色彩空間模型分別代表色相、飽和度及明度，色相表示顏色不同的數值，飽和度表示顏色的純度，純度愈低表示顏色愈偏向灰色，明度表示顏色的亮度，因此使用 HSV 色彩空間模型在調整明度時，可有效地減低光源影響顏色的因素。RGB 轉換 HSV 計算如公式 1 所示。

$$h = \cos^{-1} \left(\frac{0.5[R-G] + (R-B)}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right),$$

$$H = \begin{cases} \frac{h}{360^\circ}, & \text{if } B \leq G \\ 1 - \frac{h}{360^\circ}, & \text{if } B > G \end{cases},$$

$$S = \begin{cases} 0, & \text{if } \max(R, G, B) = 0 \\ 1 - \frac{\min(R, G, B)}{\max(R, G, B)}, & \text{if } \max(R, G, B) \neq 0 \end{cases},$$

$$V = \frac{\max(R, G, B)}{255} \dots\dots\dots(1)$$

其中， \cos^{-1} 為反餘弦運算， $\max()$ 為最大值函式， $\min()$ 為最小值函式。

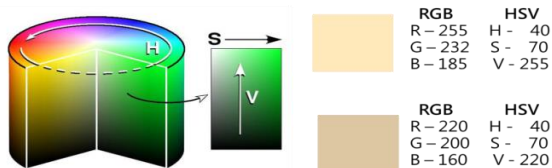


圖 3 HSV 膚色模型

曹文潔認為要找出手部的真實區域，採用 YCbCr 色彩模型有比較好的效益 [3]。這是因為 RGB 色彩影像非常容易受到光線變化的影響，且環境光線變化對膚色判定的影響甚大。因此，將色彩空間由 RGB 轉換到 YCbCr，可以減少顏色對亮度的依賴，其中 Y(luminance) 為亮度元素，Cb(blueness)、Cr (redness)分別為兩個彩度元素，由於對亮度的分離性高，方便與彩度分開操作，很適合於偵測膚色上使用。RGB 轉 YCbCr 色彩空間的轉換關係如公式 2 所示：

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.1687 & 0.33126 & 0.5 \\ 0.5 & 0.41869 & 0.08131 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} \dots(2)$$

三.實驗環境與方法

3.1 系統環境與硬體設備

本實驗中系統環境與硬體設備規格如表 1 所示。實驗方法流程圖如圖 4 所示，首先將圖檔輸入再進行色彩空間轉換取得手部二值影像後，最後執行手勢辨識。

表1 系統環境與硬體規格

硬體方面	軟體方面
CPU : Intel Core 2Duo E4600 2.4GHz	作業系統 : Microsoft Windows XP Service Pack 3
RAM : 2 Giga Byte	開發軟體 : Microsoft Visual Studio 2008
VGA : ASUS 7300GT 256MB DDR2	Open CV 2.2
IPCamera : DM355IPNC-MT5	

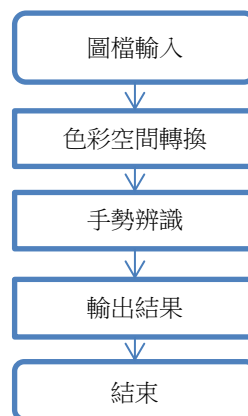


圖 4 系統流程圖

3.2 手部姿勢測試樣本

實驗使用的測試樣本是使用 Microsoft LifeCam HD-5000 去拍攝同學的手，鏡頭距離地面 115cm，距離拍攝手掌 45cm。其參數如表 2 所示。並利用 Minolta CL200 的色彩照度計測量出當下手部測試區域的照度，各自請男、女同學擺出剪刀，石頭與布三種手勢如圖 5，其背景環境為實驗室白板，總共約 800 張來做測試。詳細表格如表 3 所示。

表2 攝影機參數與環境

HD-5000參數	
亮度	70
對比度	5
曝光度	-9
實驗環境	
鏡頭離地面	115cm
背景	白色白板
目標物和鏡頭距離	45cm

表3 測試樣本實驗數據

照度	1330 xl	1055 xl	930 xl	885 xl	721 xl	533 xl	366 xl
人數 (男)	5	13	13	13	13	7	13
人數 (女)	0	7	7	7	7	0	7
手勢總類	3	3	3	3	3	3	3
張數	75	120	120	120	120	105	120



圖 5 手部測試樣本

3.3 Microsoft LifeCam HD-5000

本實驗所使用的視訊鏡頭，為 Microsoft 的 HD-5000，詳細規格如表 4 所示。

表 4. HD-5000規格

產品名稱	Microsoft LifeCam HD-5000
感應器	VGA CCD
解析度	1280*720
麥克風	內建麥克風
最高畫框速率	最高每秒可擷取30張畫面
影片解析度	1280*720
相片解析度	最高可達1280× 720

3.4 實驗方法

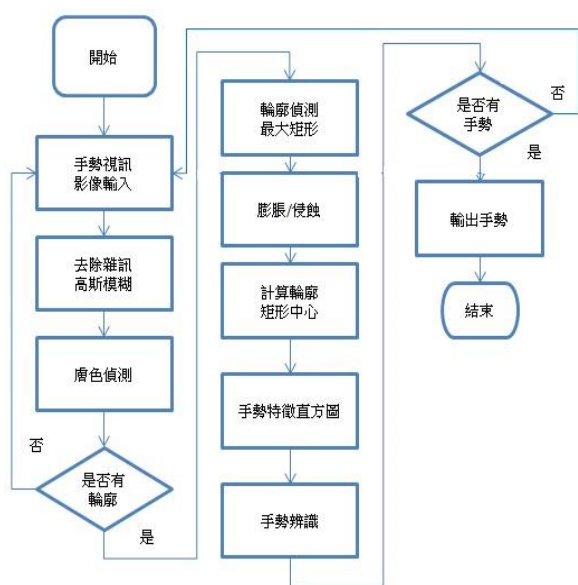


圖 6 手勢辨識流程圖

整個程式是使用 OpenCV 影像處理函式庫來實作完成，其流程如圖 6 所示。先進行影像讀取，再使用 Wang[5]提出 RGB 膚色模型進行膚色偵測。在 RGB 膚色模型範圍中將像素設定為白色，其他則為黑色。接著利用 OpenCV 的 FindContours 函式得到影像中的所有輪廓，輪廓中最大的 area (區域)即為手部區域。對此區域根據實驗需要進行形態學處理如膨脹侵蝕等。最後找出手部區域的中心，進行手勢特徵的分析，依照各手勢之峰值判斷是哪種手勢。在 HSV 及 YCbCr 其他色彩空間的實驗與此流程相似只差在使用色彩空間函式轉換到對應的色彩空間再進行膚色偵測。

四.實驗結果與討論

4.1 手勢辨識評估標準

我們將手勢辨識是否正確的結果區分成三項：第一是完全正確(Correct)，第二是有偵測到膚色但判斷可能受到干擾致使結果錯誤(Wrong)，第三則是膚色偵測錯誤或是因找到的輪廓矩形區域過小而無法偵測手掌輪廓(No Image)。

4.2 RGB 色彩空間的手勢辨識

一般影像皆是使用 RGB 色彩空間，所以先進行實驗作為後面實驗的比較基礎。參考 Wang [5] 在 RGB 色彩空間下針對人類膚色範圍提出之統計值，其膚色範圍分布在 $0.36 < r < 0.465$ 及 $0.28 < g < 0.363$ ，偵測結果如圖 7 所示。在 366x1 的光源下其準確率趨近於零，隨光源的提高，其準確率也隨之增加，可偵測到手部區域(Correct & Wrong)平均約 80%，但其準確率只達到 66%~77%，整體平均約 60%，而無法偵測到手部區域(No Image)高達 25%。從結果看出在低亮度光源時有很大的改善空間。

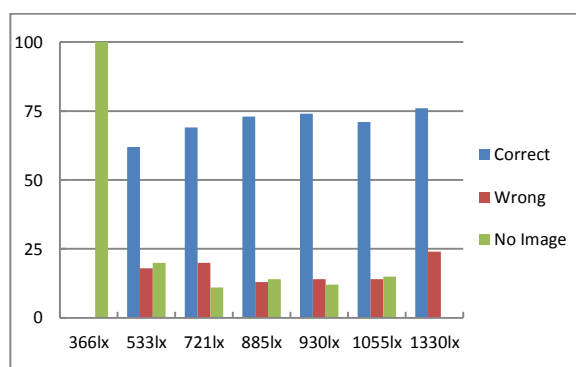


圖 7 RGB 空間偵測結果

4.3 HSV 色彩空間的手勢辨識

4.3.1 Wang's HSV 膚色模型

因 HSV 色彩空間模型在調整明度時，可有效地減低光源影響。本實驗利用 Wang [5]所提出的 HSV 色彩空間針對人類膚色範圍提出之統計值，其膚色範圍分布在 $0 \leq H \leq 50$ ， $0.20 \leq S \leq 0.68$ ，及 $0.35 \leq V \leq 1.0$ ，其偵測結果如圖 8，在 366x1 下其準確率還是趨近於零，但是在其他光源下準確率都有提升 66%~87%，整體平均準確率約 66%。但是在低亮度光源 530x1 下仍然無法改善。

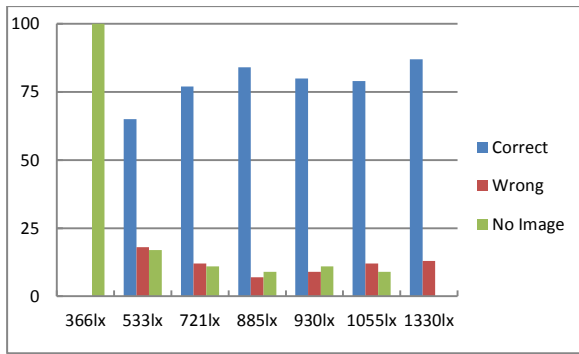


圖 8 Wang's HSV 色彩空間偵測結果

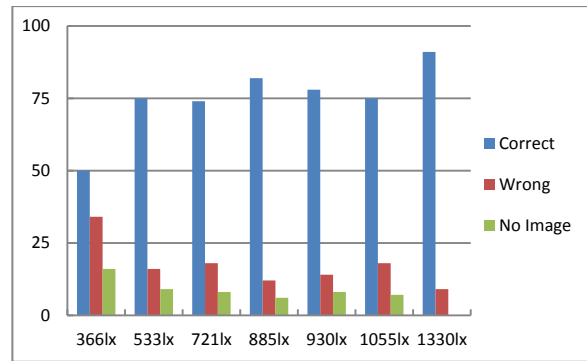


圖 10 YCbCr 色彩空間偵測結果

4.3.2 動態建模(dynamic modeling)

利用人臉偵測找尋到目前玩家的人臉，再選取人臉鼻子的中下面的膚色做為膚色建模的依據此方法我們稱為動態建模。其偵測結果如圖 9 所示。

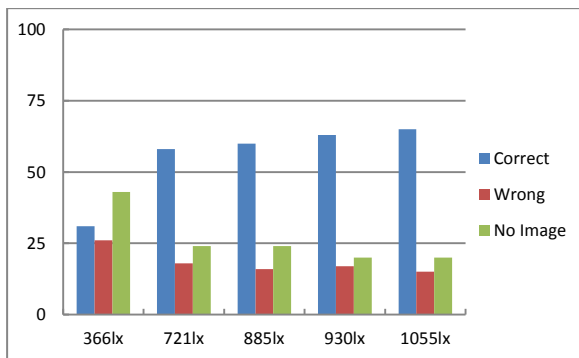


圖 9 動態建模偵測結果

在 366lx 的光源下有 25%的準確率提升，且 No Image 的比例也下降了約 20%，但是也連帶影響到其他光源的偵測效果，大約都下降了 20%的準確率，No Image 比例也提升約 15%。這是因為人臉膚色和手部膚色在相同光源下，還是有些許的差異性。

4.4 YCbCr 色彩空間的手勢辨識

4.4.1 YCbCr [Hiremath] 色彩空間

有些研究認為，採用 YCbCr 色彩模型對膚色偵測有比較好效益。在此根據 Hiremath [5]所提出統計方法，在 YCbCr 色彩空間膚色範圍分布在 $97.5 \leq Cb \leq 142.5$ 及 $134 \leq Cr \leq 176$ 。由於 Y 為亮度元素，在實驗中為區隔膚色，避免受光線之影響，大都不把 Y 的條件列在膚色判斷式中。實驗結果如圖 10 所示。

在前兩個色彩空間效能都趨近零的 366xl，在此約達到 50%的準確率。成功偵測到手部區域 (Correct & Wrong)約提升到 90%且整體平均準確率提升到 78%。雖然在 366xl 只有 50%的準確率，但比起實驗 4.2 及 4.3.1 兩者的準確率可以說是大幅提升。所以也驗證 YCbCr 在膚色的區隔上優於 RGB 以及 HSV 的色彩空間。

4.4.2 可適性 YCbCr 模型

為進一步改善低亮度光源下的效能，利用背景為黑色的情況下找尋手部區域的 YCbCr 值，用來當作不同光源下其 YCbCr 值的模型，數據如表 5 所示。根據此表提出不同光源的手部可適性膚色模型如下:在 366XL 光源下我們使用的 YCbCr 模型為 $113 \leq Cb \leq 126$ 及 $135 \leq Cr \leq 151$ ，533XL 則為 $109 \leq Cb \leq 121$ 及 $142 \leq Cr \leq 156$ ，而 721XL 以上則因 Cb 及 Cr 值都相近，所以統一使用 $101 \leq Cb \leq 118$ 及 $145 \leq Cr \leq 172$ 當作是 721xl 以上之模型。

表5. 手部可適性YCbCr模型

	Y (min)	Y (max)	Cb (min)	Cb (max)	Cr (min)	Cr (max)
366xl	29	52	113	126	135	151
533xl	51	73	109	121	142	156
721xl	72	107	102	118	145	168
885xl	85	126	102	117	148	171
930xl	89	133	101	117	148	172
1055xl	98	136	101	117	147	171
1330xl	121	162	101	118	148	172

為搭配上上述提出的手部可適性膚色模型使用，於 7 種光源下的背景進行 YCbCr 的 Y 值測試，結果如表 6 所示。目的是用來區隔 366xl、533xl 和 721xl 以上的差異性，好讓我們能判斷目前環境的

Y 值而載入相對應之膚色模型。在 366xl 下 Y 值使用 109~116，530xl 為 186~189，721xl 以上則用 206。

表6 各種光源的背景Y值分布

光源	Y(min)	Y(max)
366xl	109	116
533xl	186	189
721xl	206	206
885xl	209	209
930xl	210	211
1055xl	219	223
1330xl	227	227

圖 11 是使用可適性 YCbCr 膚色模型的結果，在 366xl 和 533xl 明顯有改善其偵測率，366xl 提升 20% 的準確率，533xl 也提升了 5%。大於 721xl 之後的實驗結果和 4.4.1 的結果相近。整體平均準確率可提升接近 80%，成功偵測到手部區域 (Correct & Wrong) 超過 90%。

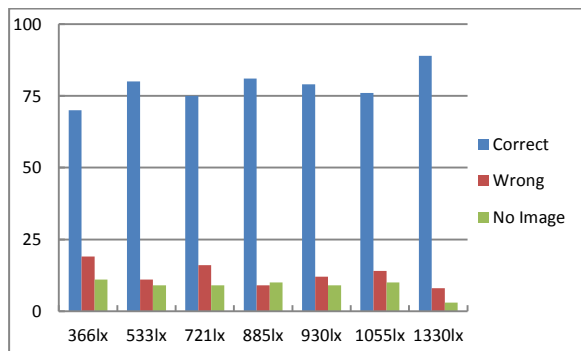


圖 11 可適性 YCbCr 模型測試結果

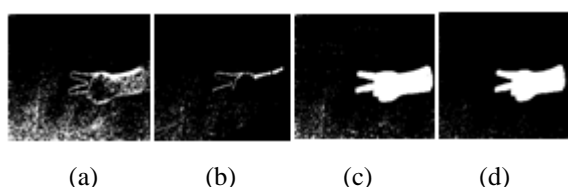


圖 12 (a)RGB(b)HSV(c)YCbCr(d)修訂膚色模型偵測結果

圖 12.分別為 RGB、HSV、YCbCr 以及 YCbCr 修訂膚色模型在 366xl 光源下膚色偵測所顯示之結果。可以很明顯的看出 YCbCr 優於其他二者，而修訂的可適性膚色模型又比 Hiremath 學者所提出的再好一點。

4.5 膨脹侵蝕

經過上面的實驗發現，不論是使用學者或自行提出的可適性膚色模型都還約有 5~10% 的樣本無法做手勢辨識的處理 (No Image)。因為這些少數的樣本邊緣訊號較不完整而造成無法做後續的辨識。

所以利用形態學的閉合來做修補。經測試發現需先做膨脹再侵蝕，膨脹次數最少一次，且侵蝕次數要超過膨脹次數才能達到較好的效果。最後於 4.4.2 的輸出使用膨脹 1 次侵蝕 3 次的影像處理，其辨識結果如圖 13 所示。從實驗結果可發現，準確率除 366xl 可達到 78% 外，其餘也都達到 80%~92%。

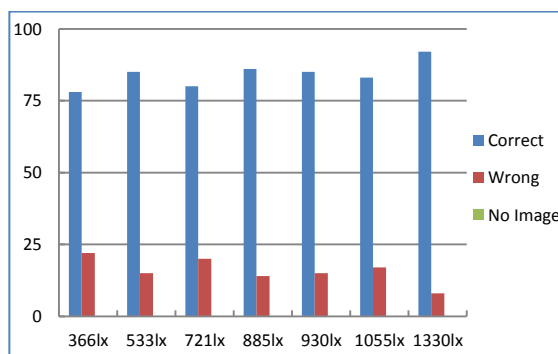


圖 13 膨脹侵蝕測試結果

五.結論

在上面實驗中發現在 RGB、HSV 及 YCbCr 色彩空間中，以 YCbCr 對於膚色的區隔是最好的，且我們提出的可適性 YCbCr 模型在低光源下有好的效能提升，在正常光源下也有不錯的結果。本論文提出一個可適性 YCbCr 膚色模型，其利用影像背景 Y 值計算估算目前環境亮度，再依此判斷應使用之對應膚色模型。

進一步利用形態學的閉合(膨脹 1 次侵蝕 3 次)提升系統的效能。辨識結果在低光源(366xl、533xl)下達到 81% 的準確率。正常光源(721xl 以上)下則達到 85%~92% 較高的準確率。最後將研究結果實作成一個機器人的猜拳遊戲系統，畫面如圖 14 所示。機器人根據猜拳結果做出相對應互動，機器人贏就旋轉前進，輸就做類似擺頭的動作往後退，達到一個有趣、有效率系統。



圖 14 猜拳遊戲系統介面

參考文獻

1. S. Wagne, B. Alefs and C. Picus (2006), Framework for a Portable Gesture Interface, 7th International Conference on Automatic Face and Gesture Recognition, pp. 275 – 280.
2. E. J. Holden and R. Owens (2003), Recognizing Moving Hand Shapes, in Proc. on 12th International Conference on Image Analysis and Processing, pp.14-19.
3. 曹文潔 (2007), 猜拳機, 國立中央大學電機工程研究所碩士論文。
4. 張竣皓 (2010), 基於可適性膚色切割之雙手手勢辨識系統, 國立台北科技大學資訊工程系研究所碩士論文。
5. Y. Wang and B. Yuan (2001), Novel Approach for Human Face Detection from Color Images under Complex Background? Pattern Recognition vol. 34, pp. 1983-1992.
6. P.S Hiremath and A. Danti (2006), etection of Multiple Faces in an Image Using Skin ColorInformation and Lines-of-Separability Face Model, International Journal of Pattern Recognition and Artificial Intelligence, vol. 20, no.1, pp.39-61.
7. <http://opencv.willowgarage.com/>
8. 程煌瑞 (2001), 生活照片之人物分類系統研究, 國立交通大學電機與控制學系碩士論文。

