

作為日常物件的輔助認知辨認與教學工具設計與研究---結合視覺辨識與漢米爾敦迴圈

陳侶龍

靜宜大學資訊工程學系
s1002581@gm.pu.edu.tw

楊孟蓓

靜宜大學資訊傳播學系
mcyang2@pu.edu.tw

徐力行

靜宜大學資訊工程學系
llhsu@pu.edu.tw

摘要

一個日常生活物品，如眼鏡的照片對於人的視覺與辨認是有意義的，本研究設計一套運用影像與影像中的輪廓來辨識熟悉物件的方式。這套技術我們希望能提供給因為生病或其他原因無法與人溝通的人，作為溝通與辨認的工具。研究的情境是設想有失智症的老人，因為記憶耗損緣故，無法辨認或說明一些熟悉的物件。透過本研究所設計的一筆畫漢米爾敦迴圈圖工具及相關智能教學方式，可以協助有此困擾的老人記憶回復。

關鍵詞：漢米爾敦迴圈，認知遊戲，影像處理

1 前言

近年來影像辨識是最廣受歡迎的研究之一，影像辨識除了重視電腦辨識以外，近年來也在探討如何與人類視覺做結合應用，這種雙邊視覺能力應用，對於影像辨識來說，也是近年來廣受歡迎研究之一，於本研究中，除了運用影像辨識的技巧外，也會將影像辨識之物體輪廓與漢米爾敦迴圈做結合運用。

在本研究主要分為兩項重要研究工作：(1)由日常生活物件影像，如何尋找到可以產生這一筆畫漢米爾敦迴圈的方式，(2)如何設計透過互動學習，讓使用者可以拉出此一筆畫與連接到相關熟悉物件影像的方式。第一個部分的研究，主要是針對一張物體較為明顯的圖形，設計一套方法，取出物體的輪廓，並取出輪廓的特徵點，利用特徵點可以概括物體的特性，再產生出與輪廓相當的漢米爾敦迴圈。這部份我們透過使用 OpenCV[9]的影像工具，來找尋有效率的方式來完成。在第二部分的研究，我們設計一套方式，讓使用者可以透過學習在平板上畫出漢米爾敦迴圈。並可以依照特徵點、漢米爾敦迴圈理論和輪廓的關係來描述整個物體外觀。經由這樣的方法，可以讓原本對於影像記憶模糊的人，經由學習來重新瞭解影像的意義。

2 研究方法

首先如果有一張背景簡單，日常生活物品影像，如何將物件擷取並且轉換成為本研究所要用到的一筆畫漢米爾敦迴圈？首先，研究會先做在邊

緣檢測前的影像處理，接著會做邊緣檢測，再來會做邊緣的篩選。最後，產生出漢米爾敦迴圈。本研究會設計 GUI 與迴圈合成方式。在下面幾節的說明中，會針對每個小細節做詳細說明，也會在每個說明的同時，會有一些簡單的範例輔助說明。

2.1 影像處理

在輪廓分析前，有許多的前置影像處理方法，這些方法主要是要來確保後續輪廓擷取時，可以得到比較好的輪廓品質，那在前置處理方面，有很多方法，在本研究影像處理部分會有影像大小的調整、灰階處理、平滑處理、門檻值處理和形態學操作，這些處理方法有的是必然經過的方法，有的是要依照輸入圖形而判定使用什麼方法，在這只會對比較重要的方法做詳述，其餘方法會以簡述的方式說明。

2.1.1 影像大小調整

對於輪廓擷取來說，輪廓的定義主要是可以完善表達出該物體的輪廓及可，並且輪廓要能符合連續不斷和封閉性這兩種特性，因此我們針對這樣的需求，進行三個實驗，這三個實驗使用的來源影像皆相同，而且邊緣檢測方法和其餘參數設定也皆相同，不同的只有解析度不同，針對這樣的設定，實驗結果如表 1，輪廓清晰度在於實驗三來說雖然解析度較差，但是其輪廓已經足以表示整個物體特徵，輪廓層次表示輪廓檢測結果是使用多少層輪廓去呈現出整個物體輪廓，因此輪廓層次越小越好，綜合上述來說，實驗三的結果是較符合本研究所需結果，因此在本研究影像辨識處理上，解析度都是以寬度 255 為主，高度則依比例改變，本研究也有針對不同解析度去做其餘參數個別設定，但得到的結果都不理想。

表 1. 實驗結果

	實驗一	實驗二	實驗三
影像來源	風扇影像	風扇影像	風扇影像
解析度	3840*2160	1024*576	255*143
輪廓清晰度	佳	普通	差
輪廓層次	104	39	5

2.1.2 灰階處理

灰階處理是由 255 種灰黑色組成，因此在處理順序上會有一些影響，先做灰階處理[2]再做高斯濾波和侵蝕膨脹，或者先做高斯濾波和侵蝕膨脹再做灰階處理，結果會稍微不同[6]。在本研究中，大多都是以先灰階處理，再做高斯濾波和侵蝕膨脹居多，通常這樣都能得到不錯的結果，但少數情況下，會需要先做高斯濾波和侵蝕膨脹，再做灰階處理，反而會得到比較好的結果，通常這種狀況是在於物體於影像上不夠顯著，這時先做高斯濾波和侵蝕膨脹，可以增強它的特徵，再灰階處理會得到比較好的結果。

2.1.3 平滑處理

影像通常會夾雜的一些雜訊，這些雜訊來源可能是電磁波干擾、傳輸過程受到干擾或者取樣時演算法設計不良，都會使得影像夾雜雜訊且不清晰，因此使用濾波處理，可以去除一些雜訊，但也有可能會降低影像的品質，通常要得到比較好的影像品質，都須對影像做過特別處理，花費的時間也會比較多，因此在本研究會選擇一個比較適合大多數影像皆可使用的演算法[6]。

關於影像濾波方法有很多，而濾波的宗旨在於盡量保留影像的完整性的情況下，還可以提高影像的品質，而在濾波方面有線性濾波和非線性濾波，本研究並沒有深入瞭解濾波的實際的算法，僅是為了減少雜訊而使用濾波器，通常最常使用的濾波器是線性濾波器，因此對於非線性濾波就沒有多加探討，在本研究中只研究過平均濾波和高斯濾波，而依照實驗結果，本研究採用高斯濾波為主，其核心設定通常是 3*3。

$$K = \frac{1}{K_{width} \cdot K_{height}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}$$

圖 1. 平均濾波[10]

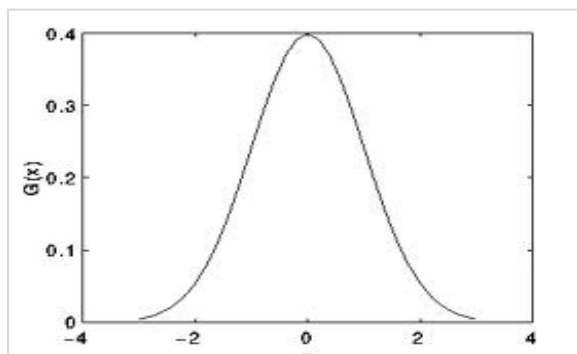


圖 2. 一維高斯函數[10]

2.1.4 門檻值處理

門檻值處理指的是利用一個固定值，然後利用像素去做比較，將顏色歸類，門檻值處理也被稱為二值化處理，通常會歸類成黑白兩色，利用像素高於或低於某個固定值來歸類。門檻值處理最主要是利用像素的顏色，當背景與物體顏色差距越明顯，歸類出來的結果越正確，當背景與物體顏色差距越靠近，歸類出來的結果越不正確，另外背景也是影響物體的最主要因素之一，通常背景越乾淨越單一，物體就會越顯著，背景比較混雜，那麼結果就會比較不理想。除此之外，門檻值處理又有分兩種方式，一種是屬於手動設定門檻值，一種是屬於自適應門檻值，通常對於整體來說，物體已經與背景明顯有不同，那麼就比較適合用手動設定門檻值，而自適應門檻值比較適合用在對於區域來說，物體已經與背景明顯有差別的地方，一個是以整體而言，一個是以區域而言，在本研究中，通常都是手動門檻值居多，自動門檻值較少使用，因為自動門檻值通常雜訊會比較多。

手動門檻值方法是採用 OpenCV 提供的門檻值設定方式。使用的參數設定變數為 THRESH_BINARY，自動門檻值方法是採用 adaptiveThreshold，使用的參數設定跟手動門檻值一樣，另外自動門檻值還有一個參數是要設定自動門檻值的演算法 adaptive gaussian model[15]。在 OpenCV 的完整函式名稱稱為 ADAPTIVE_THRESH_GAUSSIAN_C，這個演算法會計算每一點 P(x,y)附近的 blockSize * blockSize 加權總數減去 C。由本研究的實驗結果，手動門檻值部分，設定值大概在 80~160 左右，自動門檻值大約落在 23~63 左右。

2.1.5 形態學操作

侵蝕和膨脹處理是形態學操作[7]處理中最基本的操作，形態學操作就是一種基於形狀來做處理的一系列操作，而侵蝕與膨脹通過結構元素來與形狀做運算。藉此加強形狀的特徵，最基本的形態學操作運用於：(1)切割獨立影像元素，(2)連接相鄰影像元素，(3) 消去雜訊，(4)輪廓提取，(5)影像分割。通常形態學操作是用在二值化過後的影像，來將形態的邊緣做細化或粗化等操作，形態學操作主要與結構元素有關，結構元素可以是任何形狀，一般來說以圓形、矩形和交叉形為主，在不指定結構元素的情況下，會以 3*3 矩形為預設值，在指定結構元素下，就要指定明確的形狀以及大小，大小通常以 3*3、5*5 和 7*7 較常見，結構元素大小必須以奇數為主，切割影像元素通常是侵蝕最主要的作用，而連接相鄰影像元素是膨脹最主要的作用，在本研究中，侵蝕與膨脹處理會依照輸入的影像不同，而處理的方式也不同，有時候也會因為輸入的影像已經有較佳的完整性，所以不需要經過侵蝕與膨脹處理，在本研究使用到的

結構元素是矩形，而結構元素大小會依照每個輸入影像的不同，而去做最適當的調整，在本研究設定值大約是 3*3 至 7*7 之間。

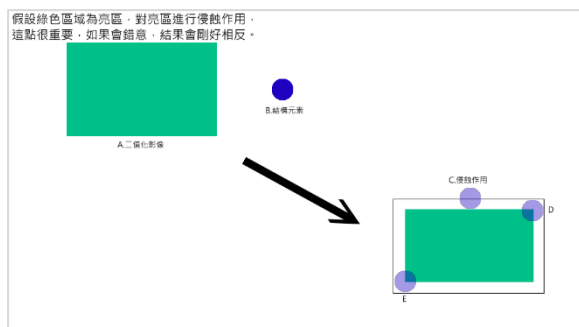


圖 3. 侵蝕作用

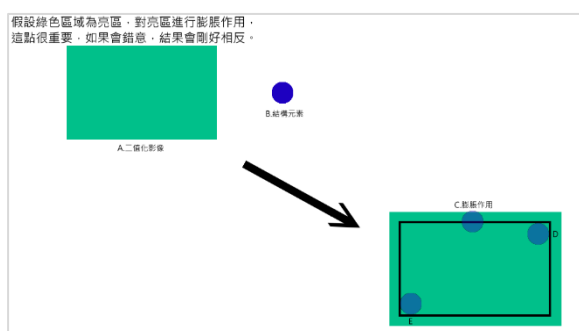


圖 4. 膨脹作用

2.2 邊緣檢測

邊緣對於影像來說，包含重要的資訊，邊緣可以用來檢測物體的大小、形狀和分類物體等等，邊緣通常可以直接表達出影像的訊息，也可以用來減少影像的資料量，邊緣基本上就是像素值在變化上，變化比較明顯的值，這些值可能是深度值的變化、方向值的變化和亮度上的變化等等，這些變化可以視為是影像的特徵，而這些特徵可能會受到環境、背景、角度、亮度和雜訊的影響，雖然是同一個物體，但是在不同狀況下所呈現出來的邊緣會都不一樣，在本研究中，邊緣檢測採用 Canny 算法[1][3]，判斷邊界值控制在 1:3 左右，而 Canny 所檢測出來的邊緣，在本研究盡量是可連續的，封閉的，因為在此的 Canny 檢測結果，在本研究中主要是要做為輪廓使用，所以必須符合輪廓的特性。

邊緣檢測結果會被當作輪廓提取的基礎樣板，然後進行輪廓提取，而輪廓提取方法是藉由 OpenCV 所提供的 findContours[13]函式，然而在 findContours 函式中，要取得輪廓的方法有四種，關於這四種輪廓取得方法，可以詳見 OpenCV 參考文件[10]，在於本研究中使用到的輪廓取得方法是 CV_RETR_EXTERNAL，因為本研究中輪廓只需要擷取外層輪廓，不擷取其內層輪廓，擷取輪廓後，要考慮到儲存輪廓的方法，而根據 OpenCV 參考文件說明，儲存輪廓點集的方法共有兩種，而在

本研究中，所使用的儲存輪廓點集的方法是 CV_CHAIN_APPROX_NONE，此方法會儲存所有輪廓點，而另外一種儲存方法是針對水平、垂直或對角線做簡化，例如一條水平線，使用此儲存方法就只會儲存左右兩個端點，在 OpenCV 中，對應到的方法名稱為 CV_CHAIN_APPROX_SIMPLE 方法，在本研究中，其實有實際去嘗試過 CV_CHAIN_APPROX_SIMPLE，並沒有得到很理想的結果，一般來說本研究採用的是日常生活中的影像，因此鮮少影像拍起來會剛好正正方方，只有在非常巧合的情況下才會出現，所以此方法就不適用，在於本研究中，此方法只會增加效能的開銷，並不會得到太大的效益。



圖 5. 原始影像



圖 6. 輪廓提取

2.3 邊緣篩選

在於本研究中，輪廓篩選這部分主要是要將雜質以及不需要的邊緣給過濾掉，在開始探討輪廓篩選之前，我們會藉由前一部份提到的輪廓結果，先對輪廓進一步處理，然後再對輪廓做過濾。

首先會先利用 approxPolyDP[4][14]演算法，將輪廓的點集做逼近算法，讓輪廓的點集裡，不必要的點數去除，留下那些比較重要的角點即可，接著將這些比較重要的點集做面積的過濾，通常面積越靠近自身邊緣檢測結果大小的，就會過濾掉，不然就是面積小於一定值也會被過濾掉，而在本研究中，面積過濾的預設值為 10，當面積過濾完成後，就要接著做角度的計算，角度的計算跟邊緣檢測和輪廓檢測有著極大的關係，原因在於邊緣檢測結果其實可以是非封閉式的，但是輪廓檢測結果一定要是封閉且連續的，因此當邊緣檢測結

果並非為封閉式的結果時，輪廓會在其端點形成接近於 0 度的行走路徑出來，這時候就必須要過濾掉，因為我們要的輪廓是封閉式輪廓，而不是開放式輪廓，圖 7 就是以圖 6 為基礎，做出來的輪廓行走路徑圖。

圖 7 的輪廓行走路徑圖，我們可以看到有三層輪廓，白色部分為原本箭頭的 Canny 邊緣檢測結果，這時候我們利用邊緣檢測結果，做輪廓檢測，就會發現這個箭頭並不是只有一層輪廓組合而成，而是由三層的輪廓組合而成，那麼我來看一下輪廓層次 1，有分為淺黃和深黃色兩條路徑，那麼它的輪廓組成就是淺黃色的路徑過去，然後由深黃色路徑回來，這樣一去一回就組成了箭頭左上部分的輪廓，依此類推，輪廓層次 2 也是如此，淺綠部分過去，深綠部分回來，因此形成箭頭右下角的輪廓部分，輪廓層次 3 也是這樣。

這時候就可以更清楚的解釋角度過濾的意思了，如圖 8 的紫色圓圈標記處，我們可以發現路徑是一去一回，因此在去和回的那個轉角，就是過濾的關鍵，去和回的轉角的角度我們可以發現，它是一個幾乎接近於 0 度的轉角，因此當我們發現有接近於 0 度的轉角，本研究會直接將它認定為是一個錯誤的輪廓，也就是非封閉式輪廓，在本研究中，輪廓擷取部分，只會擷取封閉輪廓，而非封閉輪廓就會被過濾掉，因為非封閉輪廓對於本研究來說，沒有特別的意義，因此在圖 8 的部分，輪廓層次 1、2 和 3 都會被過濾掉，那麼就意味著圖 6 右下角的箭頭，其實是一個開放式輪廓，所以這個箭頭就不是本研究需要的輪廓。

接著是輪廓範圍的過濾，假設找出來的輪廓符合我們的需求，那麼這個輪廓的點分佈就是我們下一步要做處理的地方，關於點的分佈這部分，在本研究有分為內點和外點，而外點主要是用來做輪廓範圍的判斷和輪廓的描述點，內點的主要作用是輪廓的描述點，而不作用於輪廓範圍的判斷，在本研究外點指的是輪廓點集合裡面，被拿來當作 convexHull[12]的角點，而除了角點以外的，剩餘的輪廓點集合就是內點，而這些內點有一個特性，就是這些內點通通會在 convexHull 的範圍裡面，所以在輪廓範圍的過濾部分，其實利用的就是 convexHull，利用 convexHull 找出凸包，然後藉由凸包的角點來判斷輪廓範圍的覆蓋狀況，因為本研究的物體擷取只以單一物體為主，所以輪廓範圍覆蓋狀況可以說明此影像是否由多重物體組合而成，多重物體組合而成的輪廓，並無法代表實際物體的輪廓，因此就必須要去適當的篩選，舉例來說，好比說汽車，汽車除了車體以外，還有車輪，在影像辨識後，車體與車輪會是分開的，但事實上車體與車輪的輪廓組合起來才可以構成汽車本體，因此像這類的多重物體組合成單一物體，就不適合本研究所探討的對象。另外會選擇使用凸包做篩選條件，有三個優點，第一個優點是可以過濾掉凸包範圍與實際輪廓範圍之間的雜訊，第二個優點是影像便於切割，第三優點則是去除由多重物體組合單一物體案例。

以下是一個實際的範例，假設以圖 9 當作來源影像，那麼針對來源影像做影像辨識，辨識出來的結果如圖 10，我們可以發現圖 10 的綠線部分

為 convexHull 的結果，而 convexHull 的角點我們就稱為外點，而在 convexHull 裡面的點我們稱為內點，而輪廓範圍的過濾主要是由外點來過濾，內點則是屬於輪廓的描述點，不做為輪廓範圍的過濾點，藉由內外點的分佈狀況，由圖 10 的輪廓範圍覆蓋狀況可知，此圖有兩個物體存在，而兩物體並沒有多重物體組成的問題，因此這兩個物體都是適合本研究所需要的物體。



圖 7. 輪廓行走路徑圖



圖 8. 輪廓行走路徑圖（角度標記）



圖 9. 原始影像

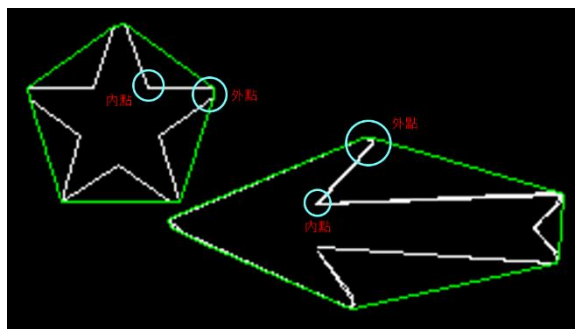


圖 10. 輪廓的凸包

2.4 設計使用介面

本研究的 GUI 介面設計主要是以 Qt 開放式原始碼軟體[11]所設計出來的。而 GUI 主要功能有建立影像、影像辨識、繪圖模式、回上一步和結果判定。除了上面這些功能以外，還有一個繪圖區域，主要是提供影像顯示與互動繪圖區，繪圖區相對於 Z 軸來看，一共有三層，我們由下往上說明的話，第一層是影像底圖，第二層是輪廓點集，第三層是繪圖層，GUI 介面如圖 11，當使用者依照規則完成繪圖時，可以進行結果的判斷，當判斷無誤時，才算正確。介面的設計我們參考使用者經驗研究的相關文獻[5]。遊戲規則是依照輪廓點集來畫線，畫出來的線必須符合物體輪廓，且必須具備漢米爾敦迴圈規則，才算正確。漢米爾敦迴圈的定義是每個點都必須經過且最多經過一次，最後會回到起點，圖 12 是漢米爾敦迴圈[8]的示意圖。

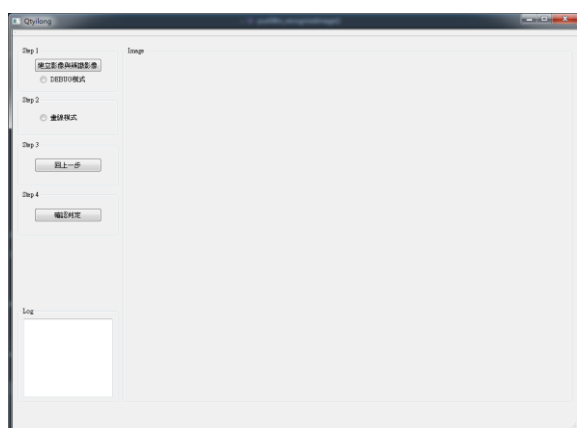


圖 11. GUI 介面

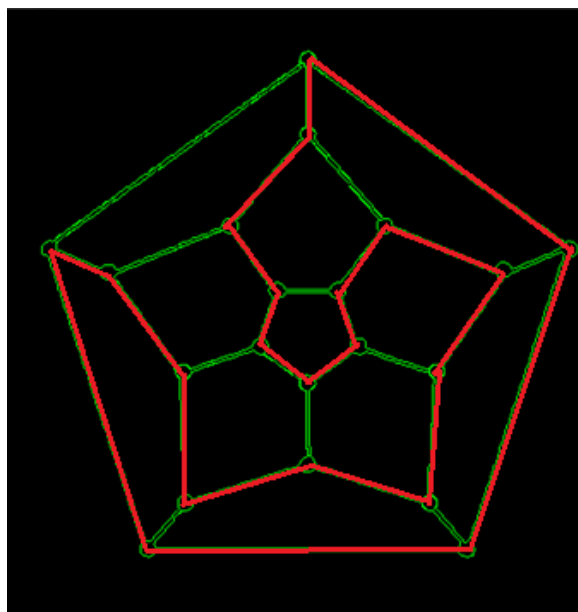


圖 12. 漢米爾敦迴圈

3 研究結果與說明

本章主要說明實驗結果，我們以一個比較簡單的範例來說，在範例中，影像是以日常生活中的鑰匙為主，在於影像辨識方面，此 Canny 門檻值為 20，門檻值過濾為自適應門檻，門檻值參數設定為 27，自動門檻演算法是 adaptive gaussian model，形態學操作是侵蝕，侵蝕結構元素是矩形，大小是 3*3，實驗出來的結果與實際鑰匙輪廓誤差接近 0。

在實驗結果部分，由於受到形態學操作影響，有些影像雖然能順利辨識出來，但是會與實際的物體輪廓稍為有點偏差，但都是在能接受的範圍內。在 GUI 來說，左上角分別有建立影像與辨識影像，然後 debug 模式可以看到整個影像辨識的過程，當影像順利載入與辨識完成，就可以看到右圖的影像區，影像區會看到綠色的點，那些點就是輪廓的實體點集，接著勾選畫線模式，就可以在右邊的影像區對那些點做畫線的動作，然後假設過程中畫線有誤，可以選擇回上一步，而回上一步這部分，最多就回一步而已，當畫線完成後，可點擊確認判定，會確認你的畫線是否符合漢米爾敦迴圈與實體輪廓，在左下角的 Log 區，可以看到辨識完成後的訊息，和畫線的整個路徑過程。

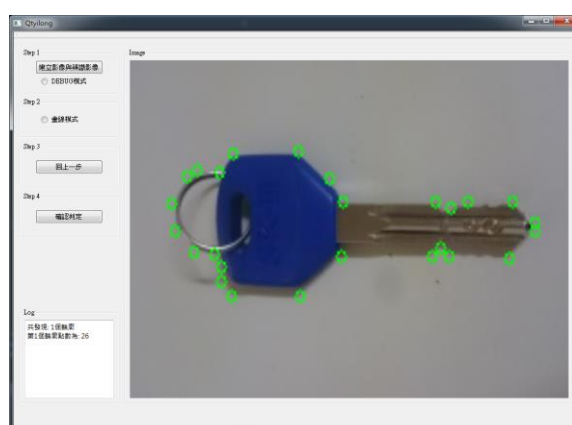


圖 13. 原始影像與輪廓角點

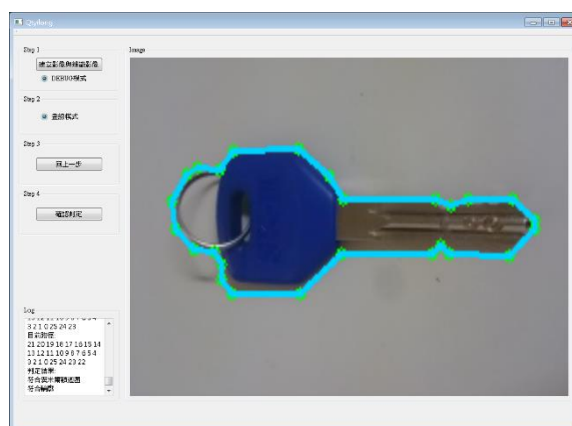


圖 14. 實驗結果

4 結論

我們運用影像處理與圖形理論嘗試設計一套互動式工具。在這個工具將日常生活物品與影像辨識做結合，將找出來的物體輪廓與漢米爾敦迴圈做應用，成為一套運用影像與影像中的輪廓來辨識熟悉物體的方式。這套應用我們希望能提供給因為生病或其他原因無法與人溝通的人，作為溝通與辨認的工具，未來期望可以將此技術放置於平板上，在本研究已經開發出第一個雛形系統，將尋求相關專業領域的研究人員來合作設計研究教案來瞭解其可行性。

參考文獻

- [1] J. Canny (1986), "A Computational Approach to Edge Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.8, no.6, pp.679-698.
- [2] Charles Poynton, Color Space, <http://www.poynton.com/ColorFAQ.html>
- [3] Ding, L., & Goshtasby, A.(2001), "On the Canny edge detector," Pattern Recognition, 34, pp. 721-725.
- [4] Douglas, D. H., Peucker, T. K. (1973). Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature. The Canadian Cartographer, 10(2), 112-122.
- [5] Fleury, Drawing and Acting as User Experience Researcher Tools, ACM APCHI 2012, pp. 269—278, 2012.
- [6] Jain, R., Kasturi, R., & Schunck, B. G. (1995). *Machine vision* (Vol. 5). New York: McGraw-Hill.
- [7] Jean Serra, Image analysis and mathematical morphology, Academic Press, London, 1982
- [8] L-H Hsu, Mathematics in Zoo, Commonwealth publishing, Taipei 2011 (in Chinese)
- [9] OpenCV, The OpenCV library, 2017, <http://opencv.org/>
- [10] OpenCV, The OpenCV documentation, 2014, <http://docs.opencv.org/2.4.9/>
- [11] Qt, Cross-platform software development for embedded & desktop, <https://www.qt.io/>
- [12] Sklansky, J., Finding the Convex Hull of a Simple Polygon. PRL 1 \$number, pp 79-83 , 1982
- [13] Suzuki, S. and Abe, K., Topological Structural Analysis of Digitized Binary Images by Border Following. CVGIP 30 1, pp 32-46, 1985.
- [14] U. Ramer, "An iterative procedure for the polygonal approximation of plane curves," Computer Graphics and Image Processing, vol. 1, no. 3, pp. 244-256, 1972.
- [15] Siggia, A. D., and R. E. Passarelli. "Gaussian model adaptive processing (GMAP) for improved ground clutter cancellation and moment calculation." Proc. ERAD. Vol. 2. 2004.