

國立交通大學

電控工程研究所

碩士論文

基於視覺注意與多重特徵的快速物件辨識

Fast Object Recognition based on Visual Attention and
Multiple Features

研究生：簡珮珊

指導教授：周志成 博士

中華民國一百零二年七月

基於視覺注意與多重特徵的快速物件辨識

Fast Object Recognition based on Visual Attention and Multiple Features

研究生：簡珮珊 Student : Pei-Shan Chien

指導教授：周志成 博士 Advisor : Chi-Cheng Jou



國立交通大學

電控工程研究所

碩士論文

A Thesis

Submitted to Department of Electrical and Control Engineering

College of Electrical Engineering

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master

in

Electrical and Control Engineering

July 2013

Hsinchu, Taiwan, Republic of China

中華民國一百零二年七月

基於視覺注意與多重特徵的快速物件辨識

研究生：簡珮珊

指導教授：周志成 博士

國立交通大學電控工程研究所碩士班

摘要

在電腦視覺的領域裡，物件辨識一直是個極具挑戰的問題。物件辨識的困難點在於如何排除外在背景干擾得到圖片中正確的物體，並對物體本身擷取特徵使其成功辨識出物件。常見的特徵抽取方式為對整張圖片做全域的特徵抽取，但全域的特徵抽取易受外在環境干擾，因此本文提出基於視覺注意的快速物件辨識，以醒目性偵測結果輔助物件辨識，藉以降低背景影響，接著以有即時性及可靠性優點的局部二值模式及色彩直方圖做特徵抽取，以達到快速辨識的需求，最後依特徵相似度排序，達成辨識目的。本論文選取 Caltech256 資料庫來對物件辨識問題進行實驗，並且比較辨識結果的準確率，實驗結果證明醒目性偵測的確能有效提升辨識準確率，且使用多種特徵做相似度計算亦可提高準確率。

Fast Object Recognition based on Visual Attention and Multiple Features

Student: Pei-Shan Chien

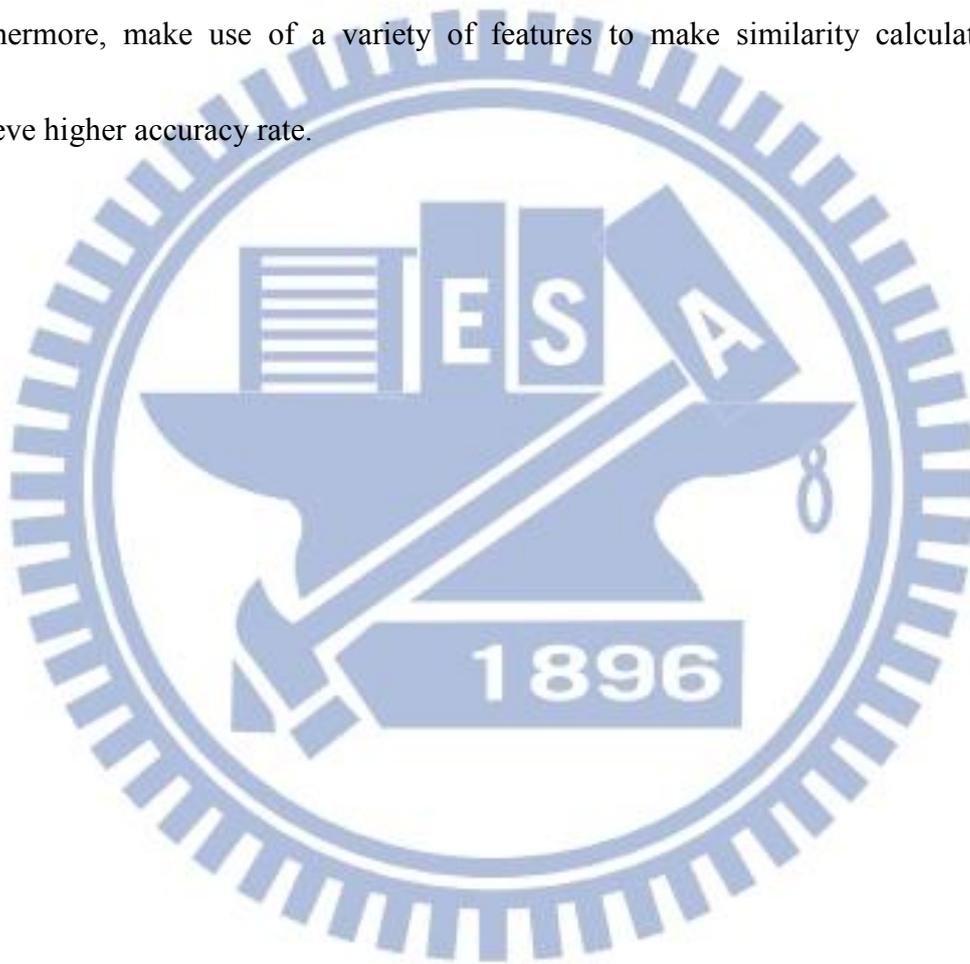
Advisor: Dr. Chi-Cheng Jou

Department of Electrical and Control Engineering
National Chiao Tung University

ABSTRACT

Object recognition has been a challenging issue in the field of computer vision. The difficulty of object recognition is how to exclude external background interference and get the object of the picture, then extract the object features and identify the object successfully. The common approach of feature extraction is extracting global feature for the whole picture. However, the global feature extraction tends to be interfered with outside environment. Therefore, we present a fast object recognition which based on visual attention to reduce the background effect. In order to achieve high-speed requirement, we use fast and robust local binary pattern and color histogram to do feature extraction, and then we classify the object according to feature similarity. After

sorting the similarities, we achieve the purpose of object recognition. In this paper, we utilize Caltech256 database to experiment on object recognition problems, and we calculate the accuracy rate of our method. The experimental result shows that saliency detection indeed can effectively enhance the accuracy rate of object recognition. Furthermore, make use of a variety of features to make similarity calculation can achieve higher accuracy rate.



誌謝

兩年的研究所生涯過得很快，轉眼就要進入下一個階段告別學生生活了。在這兩年的日子裡，不論是專業領域或是在人生的價值觀與方向上，老師都教了我們很多，真的非常感謝老師十分有耐心的教導，老師總能用最淺顯易懂的方式來帶領我們學習，雖然研究的日子並不輕鬆，但相信往後的自己仍會懷念這段與老師單純學習研究的生活，也希望自己在未來仍能保持持續學習的心態。

同時，也十分感謝 胡竹生教授與蔡雅蓉博士在百忙之中親臨指導口試，給予論文許多十分寶貴的意見，讓研究更加完善嚴謹，在此表達致伸謝意。

感謝學長阿勇、駿程在修課、論文上給予許多經驗分享，以及同學姿璿、育豪、彥錚及學妹敏婷這兩年的陪伴，讓研究的日子不孤單，還要感謝一路支持、鼓勵我的宗慶，讓我對自己有信心能夠完成這段研究的日子。

最後，特別感謝總是給予我最多支持及關愛的家人，願與你們分享這份成果與喜悅。

目錄

中文摘要	i
英文摘要	ii
誌謝	iv
目錄	v
圖目錄	vii
表目錄	ix
第一章 序論	1
1.1 研究動機	1
1.2 文獻回顧	5
1.3 論文架構	8
第二章 特徵抽取與醒目性偵測	9
2.1 色彩特徵	10
2.2 紋理特徵	11
2.2.1 局部二值模式基本型	11
2.2.2 多解析度的局部二值模式	12
2.2.3 一致性局部二值模式	14
2.3 偵測顯著物體	15
2.3.1 多尺度對比特徵	16
2.3.2 中心環繞直方圖特徵	17
2.3.3 色彩空間分佈特徵	19
第三章 辨識方法分析	22
3.1 偵測醒目區域	22
3.1.1 背景單一的物體	22

3.1.2 背景複雜的物體	24
3.2 K-最鄰近分類器	26
3.2.1 K-最鄰近分類器基本型.....	26
3.2.2 加權式 K-最鄰近分類器	27
3.2.3 卡方距離.....	27
3.3 以色彩直方圖特徵做辨識.....	28
3.3.1 色彩特徵的特性探討.....	28
3.3.2 結合醒目性偵測與色彩特徵	30
3.4 以局部二值模式特徵做辨識.....	31
3.4.1 局部二值模式參數特性探討	31
3.4.2 結合醒目性偵測與局部二值模式	32
3.5 醒目性偵測結合多種特徵.....	34
第四章 實驗結果	38
4.1 結合不同特徵的卡方距離.....	39
4.2 K-最鄰近分類法分類結果.....	42
4.3 綜合比較.....	43
第五章 結論	47
參考文獻.....	49

圖目錄

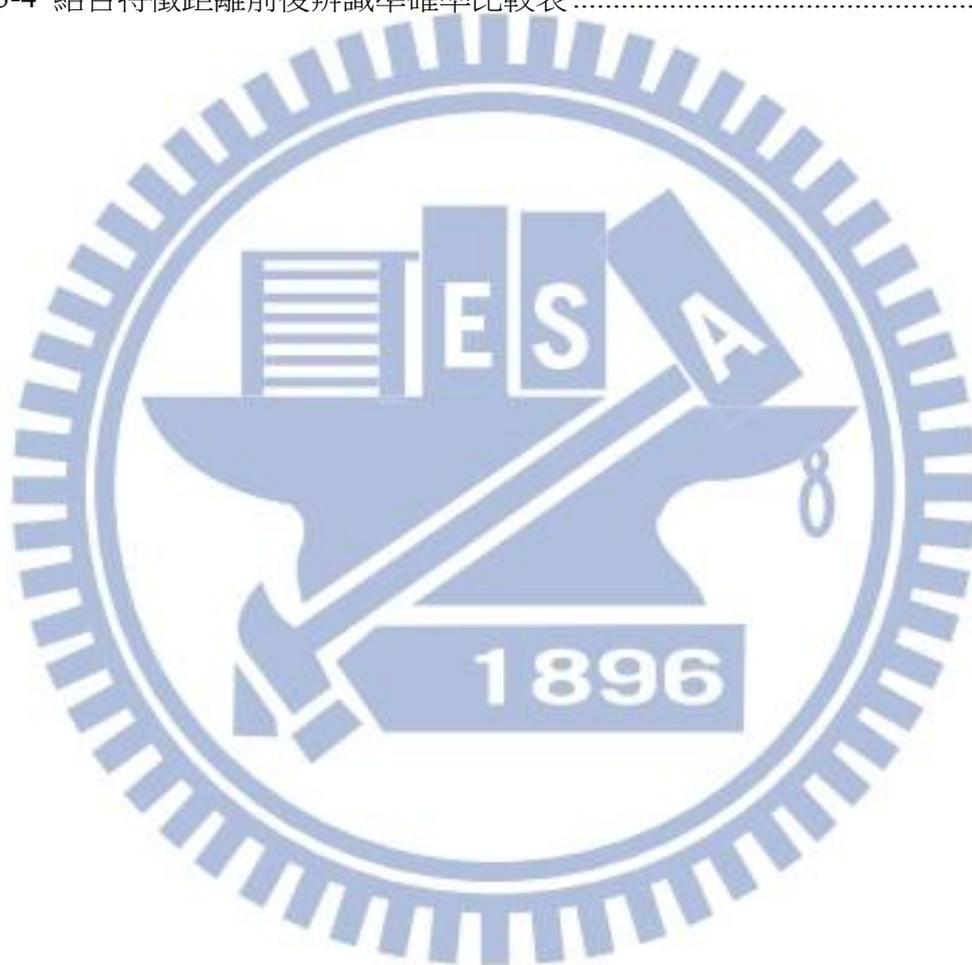
圖 1.1 傳統搜尋方式	2
圖 1.2 Google 做各種圖片查詢實例.....	2
圖 1.3 Google 各圖查詢結果.....	2
圖 1.4 Google 以圖搜圖流程圖.....	3
圖 1.5 雜湊演算法實例	3
圖 1.6 CBIR 流程圖	5
圖 1.7 影像在各種環境下的情形	5
圖 1.8 醒目物體偵測實例	7
圖 1.9 系統流程示意圖	7
圖 2.1 特徵的選擇對辨識的影響	9
圖 2.2 圖形在 RGB 空間的色彩分佈	10
圖 2.3 圖片在不同角度的空間色彩分佈	11
圖 2.4 基本型 LBP 計算方式.....	12
圖 2.5 各種尺寸的局部二值模式	13
圖 2.6 紋理各種一致性情形	14
圖 2.7 偵測醒目物體流程圖	16
圖 2.8 多尺度對比分析圖	17
圖 2.9 中心矩形與周圍矩形色彩直方圖差異	18
圖 2.10 醒目性區域特徵偵測結果	19
圖 2.11 醒目性全域特徵偵測結果.....	20
圖 3.1 背景單調的 10 類圖片	23
圖 3.2 背景單調的 10 類圖片醒目性偵測結果	24
圖 3.3 偵測醒目性效果不佳的圖片	24
圖 3.4 背景複雜的 10 類圖片	25
圖 3.5 背景複雜的 10 類圖片醒目性偵測結果	25

圖 3.6 訓練資料點示意圖	27
圖 3.7 直方圖柱數對辨識準確率的影響	29
圖 3.8 直方圖柱數對時間的影響	29
圖 3.9 實驗流程圖	35
圖 3.10 不同 LBP 卡方距離權重對辨識準確率的影響	37
圖 4.1 物件 70 類代表圖片	39
圖 4.2 隨機抽取資料庫 20 類 10 次的辨識準確率	40
圖 4.3 不同 LBP 卡方距離權重對辨識準確率的影響	41
圖 4.4 不同 LBP 卡方距離權重對辨識準確率(取中位數)的影響	41
圖 4.5 分類器不同 K 值對辨識準確率的影響	43
圖 4.6 各種辨識方法的辨識準確率	44
圖 4.7 各種辨識方法的辨識準確率(取中位數)	45



表目錄

表 3-1 色彩特徵辨識準確率比較表	30
表 3-2 不同 LBP 尺寸辨識準確率表	32
表 3-3 LBP 特徵辨識準確率比較表	33
表 3-4 結合特徵距離前後辨識準確率比較表	36



Chapter 1 序論

1.1 研究動機

隨著科技與網際網路的迅速發展，我們每天從網路接收大量的影像及文字訊息，由於網路世界資料十分龐大，如何從海量資料(Big data)得到有價值的資訊已成為重要的議題。由於攝影的普及，人們對影像查詢的需求也日益增加，許多網站為了使人們可以快速得到想要的影像資訊，提出了全新的查詢方法—以圖搜圖，此方法顛覆了以往的搜尋概念，對於影像或影片的查詢帶來新的視野，使多媒體在現今的應用變得更加廣泛。

在過去，搜尋圖片通常是以輸入圖片關鍵字來做圖片查詢，假設輸入汽車，如圖 1.1，搜尋引擎便根據每張圖片的文字註解來找出含有汽車的圖片提供給使用者，但若使用者不清楚圖片內容物為何，想藉由查詢得知時，此方法便不可行，因此現今網站推出以圖片本身的內容來做查詢，取代過去以文字做查詢的方法，此方法的好處為以圖片的內容做查詢較為客觀，不會受個人主觀想法影響，且此方法可免去人工對圖片做註解的動作，進而降低人力需求。現今常被使用的以圖搜圖搜尋引擎有 Google、TinEye，由於 Google 搜尋引擎的資料庫十分龐大，人們常以此搜尋引擎做各種查詢，因此我們想嘗試測試 Google 的以圖搜圖功能，以了解此網站以圖搜圖的特性。

我們對 Google 搜尋引擎置入了 3 張查詢圖片，如圖 1.2，以圖搜圖的結果則在圖 1.3，觀察圖 1.2 及圖 1.3 可發現查詢的結果與原查詢圖片兩者顏色的位置分布均十分相近，但得到的查詢結果與原圖並不是同類物體，尤其是在查詢的圖片

含有明顯物體類別時更為明顯，如圖 1.2(b)房屋圖及(c)椅子圖的查詢結果，雖然我們無法得知 Google 搜尋引擎是否包含原查詢的房屋和椅子圖片，但就 Google 龐大的資料庫推測應包含此兩類圖片，但查詢結果並無房屋和椅子圖，而是出現了和原查詢圖色彩相似的不同物體，此結果引發我們對 Google 以圖搜圖實作方法的好奇，並探討其方法特性。



圖 1.1 傳統搜尋方式

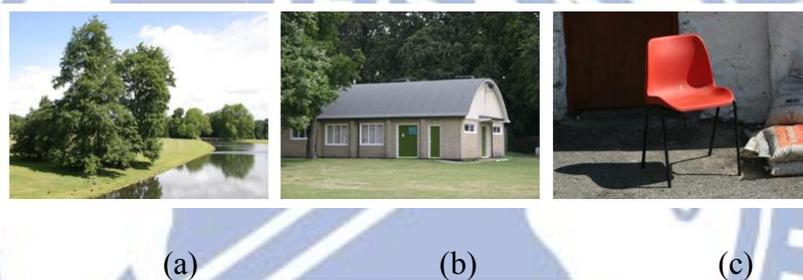


圖 1.2 Google 做各種圖片查詢實例，(a)風景圖，(b)房屋圖，(c)椅子圖

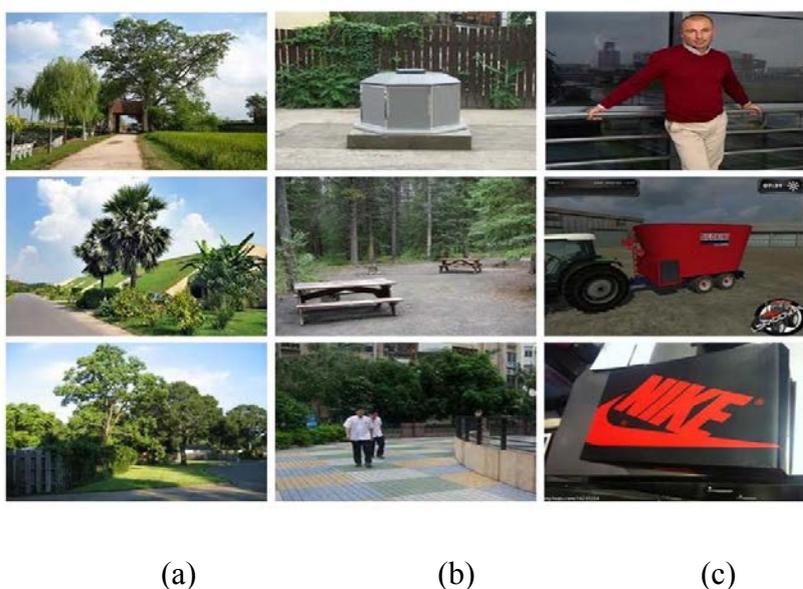


圖 1.3 Google 各圖查詢結果，(a)風景圖結果，(b)建築物圖結果，(c)椅子圖結果

Google 實現以圖搜圖的流程圖如圖 1.4，將圖庫及測試圖片做特徵抽取，再經由比較特徵相似度找出與查詢圖片最相近的圖，最後將相似度比較結果提供給使用者，其特徵抽取方式採用雜湊演算法(Perceptual hash algorithm)，以下將簡單說明演算法步驟，假設查詢圖片為圖 1.5(a)，先將圖片縮小至 8×8 的尺寸，目的為去除圖片的細節只保留原本結構，如圖 1.5(b)，再將縮小後的彩色圖片轉成灰階圖，灰階值範圍在 $[0,63]$ ，用意為保留原圖亮暗分布並捨去彩圖細節，如圖 1.5(c)，接著計算 64 個像素灰階平均值，將每個像素與平均值比較，大於或等於平均值記為 1，小於平均值則記為 0，將此串數字組合在一起得到一組 64 位的整數如圖 1.5(d)，每個查詢的照片與資料庫內的每張圖都會經由上述步驟得到這一串數字，此組數字即為每張圖片的特徵，接著以漢明距離(Hamming distance)做圖片間相似度的比較，若兩張圖片不相同的位數不超過 5，即說明兩張圖片十分相似，若大於 10，即代表這是兩張不同的圖片。

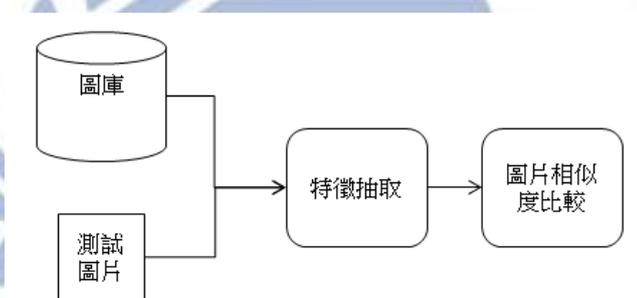


圖 1.4 Google 以圖搜圖流程圖

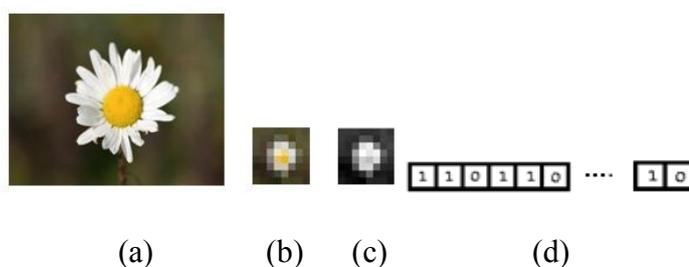


圖 1.5 雜湊演算法圖例

此方法的優點為計算量小、能快速擷取圖片特徵、圖片大小不受限制，但若查詢圖片稍有變動，如旋轉、亮暗、遠近或背景等，便造成查詢結果大不相同，因此適用於找出與原查詢圖片相同的圖，並擷取查詢結果圖片上的文字註解來告知使用者此圖片的相關資訊，但並無任何物件辨識功能包含其中。

基於上述原因，本文欲提出具有物件辨識功能的演算法，即使圖片受到光亮、角度、背景等影響，仍能準確辨識出圖片中的物體，此外，本文欲模擬即時物件辨識系統，讓使用者能快速得知圖片內容物為何，使系統兼具辨識效果及即時傳遞資訊的功能。通常物件辨識被視為分類問題。當使用者輸入影像至系統查詢，系統便開始對影像分析，如偵測影像位置、影像前處理等，並對處理後的影像抽取特徵，藉由對特徵空間做相似度的比較，最後以分類器做分類的動作並告知使用者影像分類結果。例如輸入汽車照片，電腦將查詢的照片與資料庫每類照片比對後，找出汽車類別的照片，代表判斷出此照片為汽車。此種根據影像之間的相似度來判斷兩張照片是否同類的方法，稱為基於內容的圖像檢索 (content-based image retrieval, 簡稱 CBIR)，流程圖為圖 1.6。如何從測試圖片及訓練圖片抽取出正確的特徵即為物件辨識的難題，由於即使是相同圖片，在不同亮度、角度、遠近拍攝下都有著極大變異，如圖 1.7，因此如何使圖片經過特徵抽取後，排除亮度、角度、遠近、背景等外界的干擾使同類變異減小，並將圖片正確分類一直是學界及業界的研究目標。

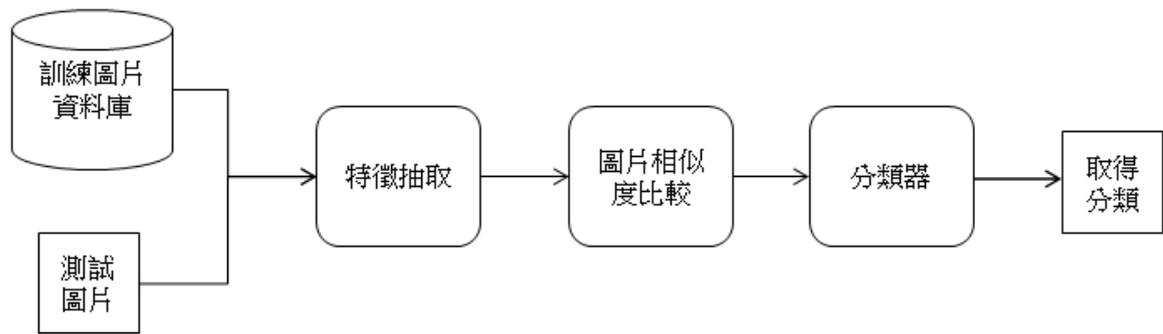


圖 1.6 CBIR 流程圖

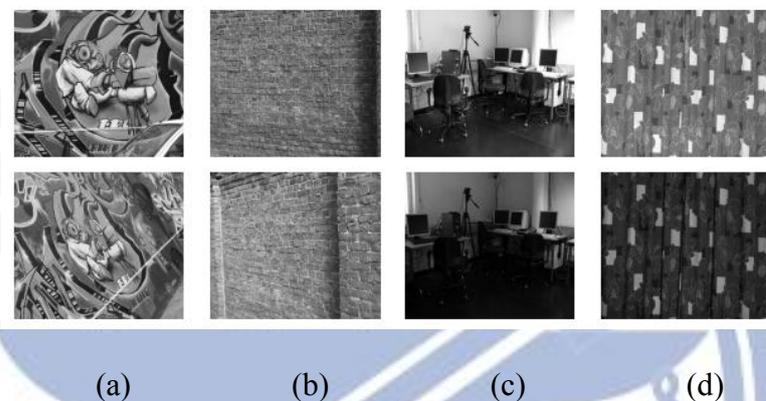


圖 1.7 影像在各種環境下的情形，(a)(b)不同角度，(c)(d)不同亮度

1.2 文獻回顧

早期的物件辨識方法常採用基於外觀的特徵抽取，有學者基於物體外觀的色彩，如 1991 年 J. Swain 提出的色彩檢索(Color Indexing)[1]，藉由抽取圖片色彩特徵作物件辨識，有學者則基於物體外觀的材質，如 2001 年 J. Malik 提出的 3 維空間紋理(3D texton)[2]及 2002 年 T. Ojala 提出的局部二值模式(Local binary pattern)[3]，藉由抽取材質特徵來做分類，有學者則基於物體外觀的形狀，如 2002 年 S. Belongie 等人提出的基於形狀內容的匹配(Shape matching using shape contexts)[4]，以抽取圖片形狀特徵來做分類。基於物件的外觀對其抽取特徵是早期

物件辨識常使用的方法，但由於早期多數的方法只能抵抗小部分外界干擾，在真實情況下，若圖片的背景或光線產生太大變化，這些方法便無法克服，因此對整張圖片抽取外觀特徵的方法逐漸式微。

物件辨識中期提出的方法為抽取圖片局部外觀找到的興趣點當作特徵來做辨識，如 G. Lowe 在 2004 年提出的尺度不變特徵轉換(Scale-invariant transform feature)[5]及 H. Bay 在 2006 年提出的加速穩健特徵(Speeded Up Robust Features)[6]。由於局部不變特徵是基於物體邊緣的興趣點，興趣點不受物體角度、光線或遮蔽物影響，因此對於光線、雜訊、些微角度改變的容忍度相當高，也較接近真實世界的物件辨識，基於上述優點，奠定了局部不變特徵在物件辨識的重要地位。

隨著圖片切割演算法的進步，現今常用的方法則為對圖片做切割，並對不同區域分別抽取全域特徵(global feature)，如 Chunhui Gu 等學者在 2009 年提出的區域辨識(Recognition using regions)[7]、L. Li 等學者提出的場景辨識、註記及切割[8]及 O. Russakovsky 等學者提出的影像分類[9]，或者藉由偵測局部不變特徵位置來做各種特徵提取，藉由綜合不同面向的特徵來做物件辨識。因此早期的特徵抽取方式，藉由加入切割演算法，在現今再度被重視，換言之，此階段的物件辨識並不只由單一方法來辨識物件，因為即使是同類的物體在外觀上也常有極大不同，使得物件辨識的取偽機率相當的高，若只從單一方法來辨識物體則無法得到正確的辨識結果。

在回顧物件辨識文獻後，我們想從現今的切割演算法中挑選適合的演算法來輔助物件辨識，我們選取了在 2011 年由 T. Liu 等學者提出的偵測醒目區域(Learning to detect a salient region)[10]。在一張圖片裡最能吸引人們注意的區域即為醒目區

域，醒目區域的偵測可從圖片的對比度、區域間色彩差異、及顏色的分佈性分析，圖 1.8 為醒目區域偵測的實例，(a)圖為原圖，(b)圖為醒目性偵測結果，中間亮處為偵測出的醒目區域，也是原圖車子的位置。藉由醒目區域的偵測，給予醒目性大的地區較高分數，背景的部分則分數較低，並將偵測結果作為相似度比較的權重，藉此降低圖片中外在背景的干擾進而提升辨識準確度，且由於速度上的考量，我們選取快速的特徵抽取方式來達到系統的即時性，接著將特徵做相似度的比較，最後藉由分類器辨別出物件類別。因此本文的目標為模擬出基於物件辨識的查詢系統，此系統需有一定的物件辨識準確度以及具備即時性的查詢，圖 1.9 為本文系統流程示意圖。

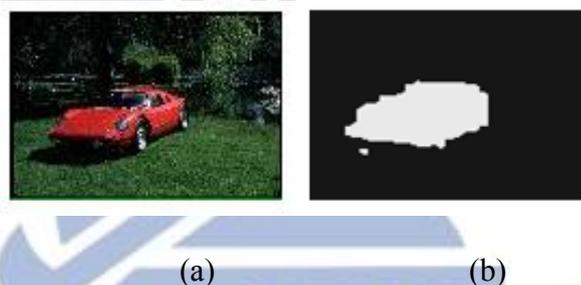


圖 1.8 醒目物體偵測實例，(a)原圖，(b)圖片偵測結果

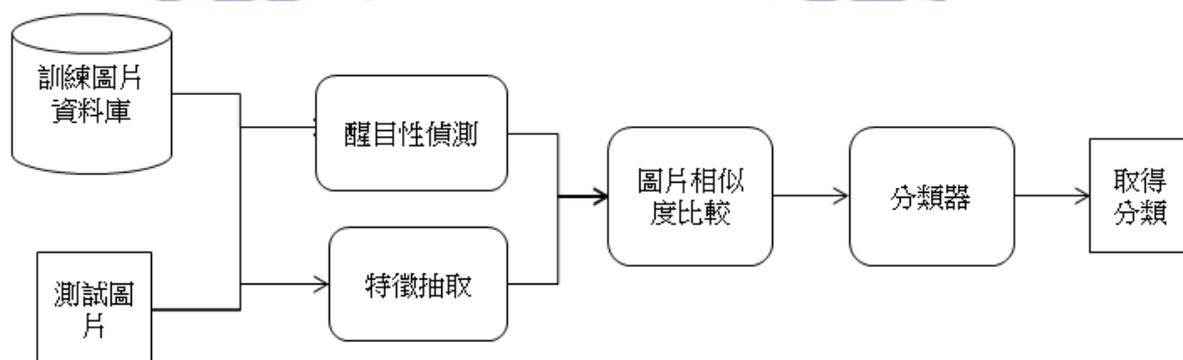
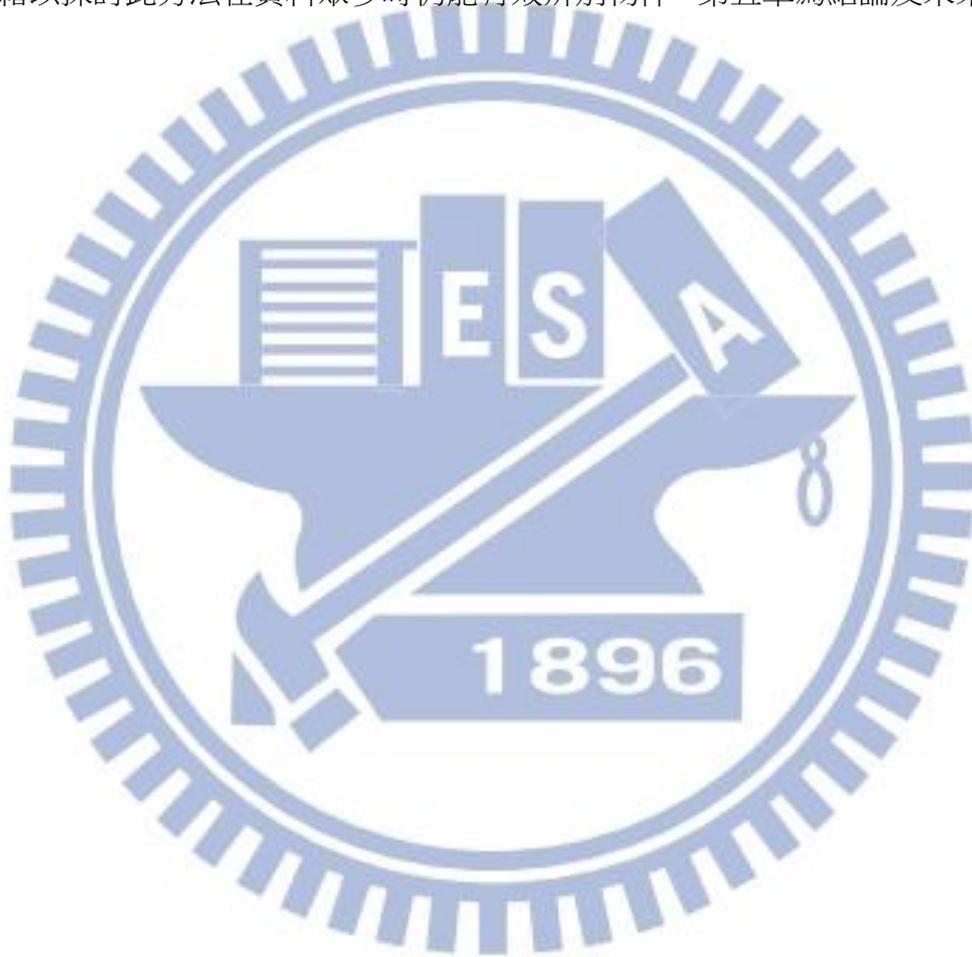


圖 1.9 系統流程示意圖

1.3 論文架構

第一章為問題描述，簡述本文研究動機及研究架構，第二章為探討各種特徵在物件辨識的應用及其優缺點，並介紹醒目性偵測的原理，第三章將探討本論文提出的方法，並以小實驗佐證其可靠度，第四章使用大型資料庫進行物件辨識實驗，藉以探討此方法在資料眾多時仍能有效辨別物件，第五章為結論及未來展望。



Chapter 2 特徵抽取與醒目性偵測

在辨識物體時，人眼會從各方面比較，如顏色、形狀、材質、個人經驗等方面辨別，人眼通常不只以一種物件特性去做判斷，而是綜合上述所列的性質共同判斷，因此如何使電腦擁有人類快速而準確的辨識能力仍是現今物件辨識的重要目標。對電腦而言，辨識影像的方法通常依圖片的特徵去做相似度的比較，然後依相似度結果來做分類。因此選取不同的特徵會得到截然不同的相似度比較結果，以圖 2.1 舉例，若我們以形狀特徵來比較此三樣物品，則第一樣物品跟第二樣物品較為相似，若以色彩特徵來做比較，則第一樣物品和第三樣物品較為相似，因此選取不同的特徵將影響相似度的判別甚鉅。



圖 2.1 特徵的選擇對辨識的影響

除了特徵的選擇會影響辨識結果，相似度的度量方法也可能造成影響，因為即使在相同的特徵空間中，兩個特徵向量的相似度若以不同方式計算也會有不同的結果，例如假設在三維的空間中存在兩組特徵向量 $a = [1 \ 0 \ 0]^T$ ， $b = [0 \ 1 \ 0]^T$ ，如果我們以歐基里德距離計算這兩個特徵向量的相似度為 $\|a - b\| = \sqrt{(1-0)^2 + (0-1)^2 + (0-0)^2} = \sqrt{2}$ ，但若以卡方距離計算這兩個特徵向量，

相似度則為 $\|a-b\| = \frac{1}{2} \left(\frac{(1-0)^2}{1+0} + \frac{(0-1)^2}{1+0} \right) = 1$ 。因此當兩個向量因度量方式不同即可

能得到不同的相似度，所以相似度量的選擇也需列入考量。

由於本文欲模擬一即時物件辨識系統，因此系統的速度和可靠度皆為重要考量，我們考量需要的特性選出色彩特徵及紋理特徵做為我們選取的特徵，以下將詳細介紹這兩種特徵特性。

2.1 色彩特徵

以人類視覺而言，色彩是顯而易見的特徵，也是物件辨識的重要指標之一，最早在 1991 年由 Swain 等學者提出以色彩直方圖來作物件辨識。對電腦而言，圖片中的每個像素都可以用一組數字來表示該像素的色彩，經由統計色彩空間的分佈情形繪成直方圖即得到色彩特徵，圖 2.2 即為圖片在 RGB 空間的例子。

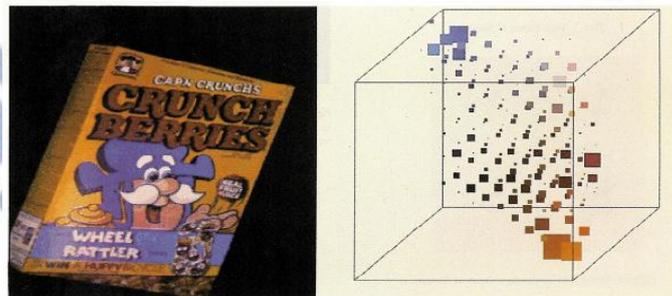


圖 2.2 圖片在 RGB 空間色彩分佈

舉例來說，假設圖片使用的色彩空間為 RGB，色彩範圍為 $[0,255]$ ，將 RGB 三個軸分四等分，第一等分的顏色值為 $[0,63]$ ，第二等分顏色值為 $[64,127]$...等依此類推，由於有 RGB 三個軸，三個軸皆分成四等分，可得到共 4^3 個柱的色彩特徵值方圖。其優點有特徵抽取速度快、不受物體旋轉影響、不受位置改變影響、可容

許物體小部分的遮蔽，以圖 2.3 為例，當物體旋轉、角度偏移、或有小部分的遮蔽，其產生的色彩直方圖仍差異不大，但由於缺少對色彩分佈空間上的資訊，即使完全不同的物體也可能產生相同的色彩直方圖，而即使是同類的物體也可能色彩不同而無法被辨識為同類物體，因此若只單以色彩特徵來做判斷並不周詳，仍需搭配其他特徵共同辨識。

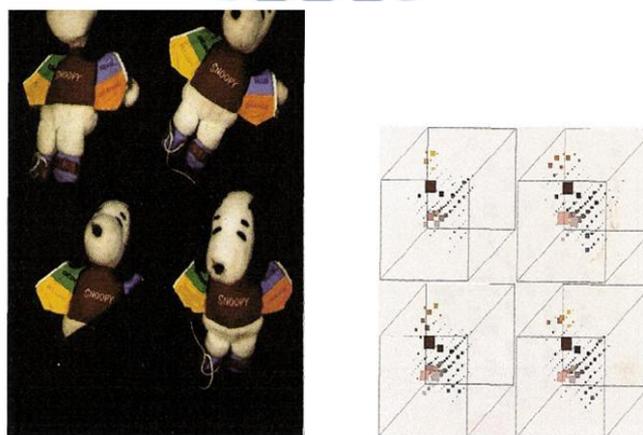


圖 2.3 圖片在不同角度的空間色彩分佈

2.2 紋理特徵

物體的紋理也是常用的辨識特徵，人們可藉由不同紋理來判別物體的種類。紋理是指圖片中重覆出現的空間結構，藉由提取重覆出現的空間結構，即可得到紋理特徵，再依紋理做相似度的比較來做辨識。常見的紋理特徵抽取方式有局部二值模式法、3D 紋理基元、灰階共生矩陣[11]。本文採用的是局部二值模式法，以下詳細說明局部二值模式原理。

2.2.1 局部二值模式基本型

局部二值模式 (Local Binary Pattern, 簡稱 LBP) 是一種用來描述圖片紋理的局部特徵，於 1994 年被提出[12]，現在常見的版本為 2002 年優化後的多解析度的

局部二值模式，由於計算簡單且兼具可靠性，使得它在紋理辨識佔有不可或缺的地位。早期 LBP 演算法原理如下:首先將圖片轉為灰階影像，假設影像內某區塊灰階值如圖 2.4(a)，將區塊內周圍八個像素和中心像素灰階值做比較，若大於或等於中心像素值，則該像素點標記為 1，否則為 0，如圖 2.4(b)，接著將八個位置得到的值依序乘上 2 的次方，如圖 2.4(c)，將 3×3 區塊內的值相加即可得到該中心像素代表的 LBP 值，如圖(d)，此 LBP 值可反應局部區域內的灰階變化。藉由統計圖片中每個像素的 LBP 值得到直方圖，做為圖片的紋理特徵。

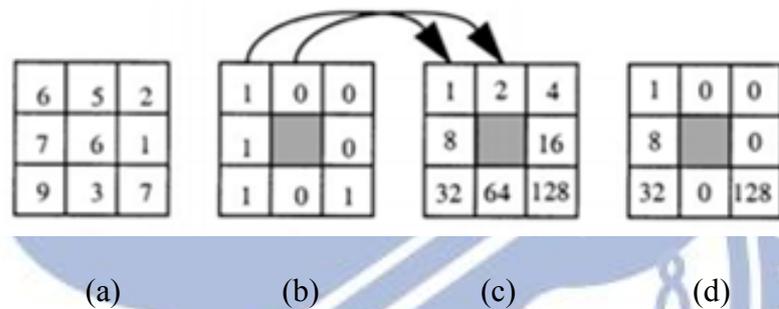


圖 2.4 基本型 LBP 計算方式

2.2.2 多解析度的局部二元模式

由於 LBP 基本型的計算方式侷限於 3×3 的區域無法滿足不同紋理的需要，因此新型的 LBP 因應而生，名為多解析度的局部二值模式 (Multi-resolution Local Binary Pattern)，計算 LBP 的區域從原本的 3×3 矩形擴展為任意區域，並以圓形代替原本矩形來做計算，使用者可依圖片需求自行選擇圓形半徑及周圍採樣點，圖 2.5 為各種尺寸的局部二值模式，其中 P 代表圓形周圍採樣點的各數，R 代表圓的半徑大小。

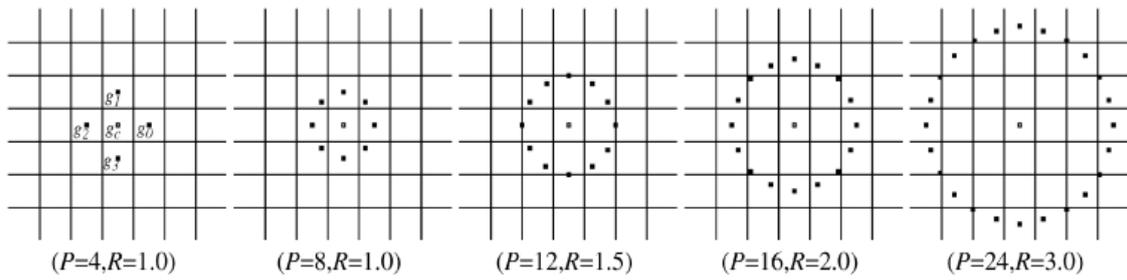


圖 2.5 各種尺寸的局部二值模式

在多解析度的 LBP 中，我們可根據不同紋理挑選適合的半徑及周圍取點數來計算 LBP 特徵，(式 2.1)為多解析度 LBP 計算方式:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (2.1)$$

其中 P 為周圍取點數， R 為圓形半徑， $s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$ ，再藉由統計圖片中每個像素的 LBP 值得到直方圖，做為圖形的特徵。

多解析度的 LBP 計算方式雖解決了尺寸上的問題，但作者 T. Ojala 以此方式辨識紋理卻得到不理想的結果，因此他再度對 LBP 做更深入的研究。經研究後發現，直方圖內的 LBP 值並非隨機分佈，有高達 90% 的 LBP 分佈於某些特定值，此代表圖片中某種空間結構重覆出現的機率相當高，只有少部分的結構是不常出現的，若將此種結構以相同方式一起統計成直方圖，將造成直方圖內的雜訊，在做相似度度量時將帶來干擾。以半徑 $R=1$ ，周圍取點數 $P=8$ 為例，LBP 值將落在 $[0, 255]$ ，但 LBP 值並非隨機分佈於 $[0, 255]$ ，而是明顯的集中於某些 LBP 值，因此 T. Ojala 再度提出新的 LBP 計算方式，稱為一致性局部二值模式。

2.2.3 一致性局部二值模式

由於觀察結果發現，多解析度的 LBP 值並非隨機分布，高達 90% 的 LBP 分佈於某些特定值，為了除去不常見的 LBP 值帶來的雜訊，作者 T. Ojala 提出新的 LBP 計算方式，稱為一致性局部二元模式 (Uniform Local Binary Pattern)。一致性局部二元模式的 LBP 計算方式也是採取圓形，使用者一樣可自行決定半徑大小 R 和周圍取點個數 P ，將中心像素和周圍像素比較，周圍像素大於或等於中心像素則標記為 1，小於則記為 0，不同的是，當周圍像素從 0 到 1 的變化次數超過兩次時，則將所有包含此種情況的紋理空間結構視為同一種紋理特徵，如圖 2.6。

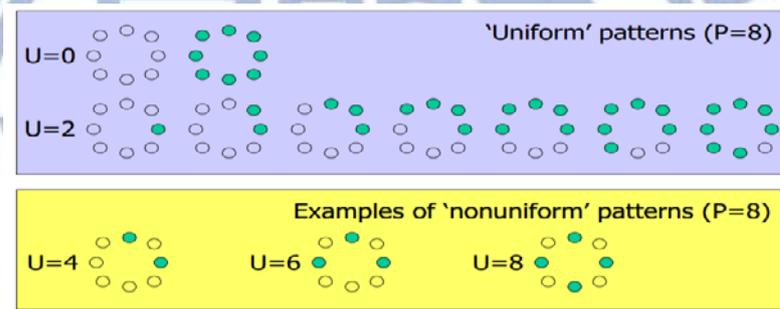


圖 2.6 紋理各種一致性情形

作者將周圍像素的一零變化次數稱為一致性，我們將一致性表示為 U 值，(式 2.2)為一致性計算方式:

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (2.2)$$

其中 P 為周圍取點各數， R 為圓形半徑大小， $s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$ 。

當一致性大於 2 時，將所有一致性大於 2 的像素歸類為相同的 LBP 值，以降低此種 LBP 值對直方圖統計的影響，則(式 2.3)為一致性 LBP 計算方式:

$$LBP_{P,R}^{u2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c)2^p, & \text{if } U(LBP_{P,R}) \leq 2 \\ 2^P, & \text{otherwise} \end{cases} \quad (2.3)$$

藉由將一致性大於 2 列為相同的 LBP 值(2^P)，來去除此種紋理空間架構造成的干擾，再藉由統計每個像素的 LBP 值得到直方圖，做為圖形的紋理特徵，此即為現今常用的 LBP 計算方式。

基於一致性 LBP 有特徵抽取快速能即時分析圖像、抗光亮(當圖片亮度單調遞增或遞減時)、不受物體位置改變影響、相似度比較容易等優點，已在不同領域被廣泛應用，如紋理辨識、紋理切割[13]、以圖找圖[14]、人臉辨識[15][16]等領域。

2.3 偵測顯著物體

心理學家 William James 在 1890 年時對注意力下了這樣的註解“Everyone knows what attention is.”，如 William James 所言，每個人都知道什麼是注意力，藉由注意力的幫助，人類腦部在接收外界大量訊息同時快速且本能的選取出重要的資訊並迅速做出判斷。就視覺注意 (Visual Attention) 而言，某些視覺刺激能夠捕捉注意力，使人眼給予較多視覺刺激的影像關注，這樣的特性在心理學上稱為醒目性 (saliency)。藉由醒目性人眼可以快速做出各種判斷，因此電腦視覺對包含醒目性的影像極富興趣並投入許多研究，目前醒目性已被關注於各種應用，如影像/影片的壓縮、行人偵測、車流偵測、物件辨識等。

本文將利用視覺的醒目性來輔助物件辨識，我們採用 T. Liu 等學者提出的方法來偵測圖片醒目區域，並使用 V. Ordonez 的程式來實現。T. Liu 將偵測醒目區域視為切割問題，目的為將醒目物體和背景切割，為了正確切割醒目區域和背景，作

者提出利用局部的、區域的及全域的特徵來描述醒目的區域，並以條件隨機域 (Condition Random Field, 簡稱 CRF) 結合所有特徵藉以得到最佳的醒目區域描述，並偵測出醒目物體的位置，圖 2.7 為簡單的流程示意圖。以下小節我們將介紹三種醒目特徵抽取方式。

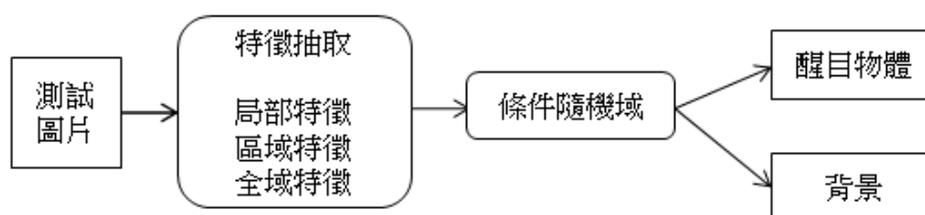


圖 2.7 偵測醒目物體流程圖

以下我們將介紹描述醒目區域的多尺度對比特徵、中心環繞直方圖特徵、色彩空間分佈特徵，此三種特徵代表了圖片局部的、區域的及全域的醒目特徵，如下將詳細描述此三種特徵的抽取方式。

2.3.1 多尺度對比特徵 (Multi-scale contrast feature)

第一個特徵為多尺度對比特徵，由於色彩對比度造成的視覺刺激相當大，因此對比度為視覺注意最常使用的局部特徵，藉由計算圖片各種尺寸的色彩對比度得到多尺度對比特徵，以下描述多尺度對比的實行步驟。

對原影像擷取 6 層高斯金字塔圖，使不同尺寸的影像色彩值皆連續平滑，如圖 2.8(a)，接著對每層金字塔圖做對比度的計算，計算方式為計算 3×3 範圍內中心像素與周圍 8 個像素的差平方，然後將八個值累加即為該像素的對比度，每層的每個像素都依此方式計算，其結果將如圖 2.8(b)，其明顯的描繪出美國國旗色彩對比度，越亮的地方代表對比度越大，最後將每層金字塔圖縮放成原圖大小再做色彩值累加，將每層金字塔圖做累加可使原本對比度大的地方更被關注，並降低圖片

上的不重要的細節，最後將整張圖色彩範圍調整至[0,1]即得到多尺度對比特徵如圖 2.8(c)。

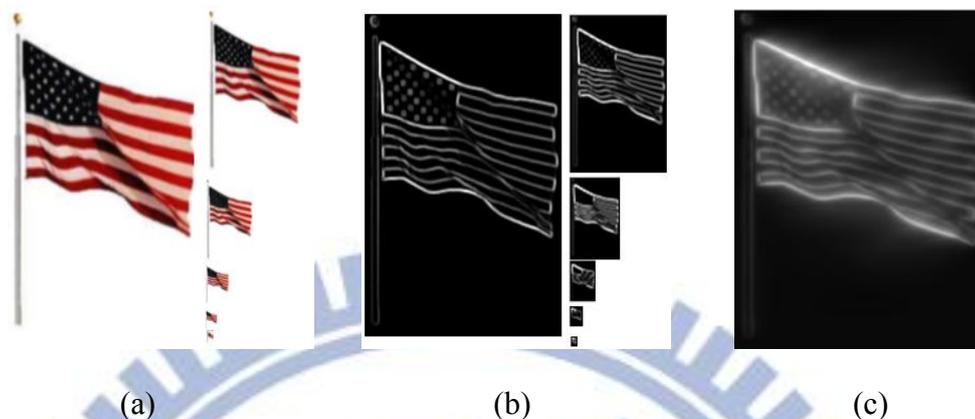


圖 2.8 多尺度對比分析圖

2.3.2 中心環繞直方圖特徵 (Center-surround histogram feature)

第二個特徵為中心環繞直方圖特徵，為區域性的醒目特徵，若圖片中有顯著的物體，我們猜測物體在圖片內佔的面積不會太小，且易於從圖片中區分，因此我們想找出能擷取圖中醒目區域的特徵。先假設圖中的醒目區域可以用一個矩形來框出，我們想找出包含此矩形的醒目區域，首先利用不同尺寸的矩形來偵測圖中各地區，統計中心矩形與其周圍矩形的色彩直方圖，當中心矩形的色彩直方圖與周圍矩形色彩直方圖差異越大，代表此區域越有可能是醒目區域，以下說明詳細實行方法。

假設圖片大小長為 M 寬為 N ，由於無法得知醒目區域在圖中佔的面積，我們假設各種不同大小的矩形藉以描述醒目區域，矩形尺寸為 $k \times \min(M, N)$ ， $k = 0.1 \sim 0.7$ ，長寬比為 $\{0.5, 0.75, 1, 1.5, 2\}$ ，藉由矩形的尺寸大小及長寬比變化，找出造成差異最大的色彩直方圖矩形長寬比，其數學式描述如下：令 x 為圖片的像素， $R(x)$ 為以 x 像素為中心環繞的矩形， $R_s(x)$ 環繞矩形 R 周圍的區域， N 為直

方圖柱數，則最大的色彩直方圖差異公式為(式 2.4)：

$$R^*(x) = \arg \max_{R(x)} \chi^2(R(x), R_S(x)) \quad (2.4)$$

其中

$$\chi^2(R(x), R_S(x)) = \frac{1}{2} \sum_i^N \frac{(R^i - R_S^i)^2}{R^i + R_S^i}$$

由式(2.4)便可找出圖中每個像素位置最大色彩直方圖差異的矩形尺寸，以圖 2.9 為例，圖為不同像素位置得到的矩形尺寸。

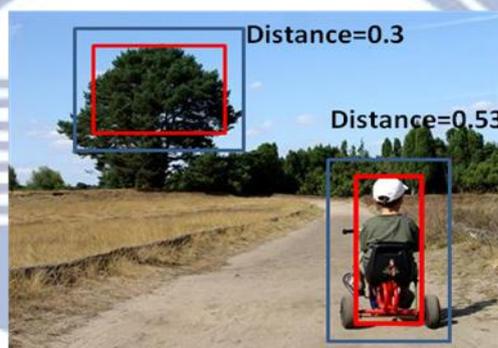


圖 2.9 中心矩形與周圍矩形色彩直方圖差異

接著利用找出的矩形長寬比，累計不同尺寸的矩形造成的醒目區域特徵，令 x 為圖片的像素， I 為圖片， $f_h(x, I)$ 為圖片 I 中 x 像素的醒目性， R^* 為(式 2.4)找出的矩形， R_S^* 為環繞矩形 R^* 的區域， x' 為矩形 R^* 的中心像素，則每個像素的區域醒目性為(式 2.5)：

$$f_h(x, I) = \sum_{\{x' | x \in R^*(x')\}} w_{xx'} \chi^2(R^*(x'), R_S^*(x')) \quad (2.5)$$

其中 $w_{xx'} = \exp(-0.5\sigma_{x'}^{-2} \|x - x'\|^2)$ 。最後藉由調整 f_h 至 $[0, 1]$ 即得到區域性的醒目特徵，圖 2.10 即為區域性特徵實例。

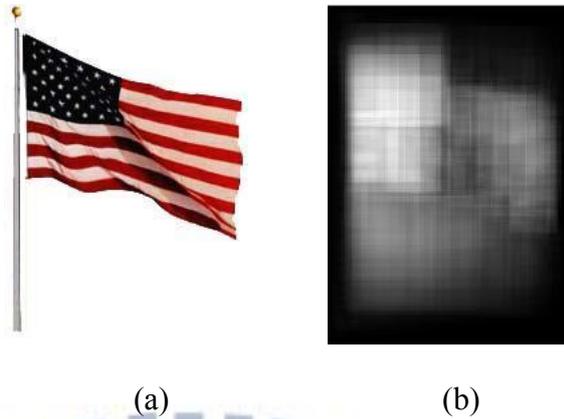


圖 2.10 醒目性區域特徵偵測結果，(a)美國國旗原圖，(b)區域性的醒目特徵結果

2.3.3 色彩空間分佈特徵 (Color spatial-distribution feature)

觀察含有醒目物體的圖片，我們發現當同一顏色集中於圖片部分區域時，人眼會覺得此區域是顯著的，反之，當同一顏色分散圖片各個區域時，人眼會忽略這些部分，因此色彩的空間分佈也影響圖片的顯著性，此即為我們的第三個特徵，詳細計算方法如下。

為了描述圖片的色彩空間分佈，我們先利用高斯混合模型 (Gaussian Mixture Models, 簡稱 GMM) 來計算色彩的空間變異，令 C 為圖片色彩數目， I_x 為圖片中的 x 像素， w_c 為第 c 種顏色的機率， μ_c 為第 c 種顏色的色彩平均值， Σ_c 為第 c 種顏色的變異矩陣，我們可得到每個像素在各種顏色的機率值為(式 2.6)：

$$p(c | I_x) = \frac{w_c N(I_x | \mu_c, \Sigma_c)}{\sum_{c=1}^C w_c N(I_x | \mu_c, \Sigma_c)} \quad (2.6)$$

其中 $N(I_x | \mu_c, \Sigma_c)$ 為多變量密度函數，如下：

$$N(I_x | \mu_c, \Sigma_c) = \frac{1}{(2\pi)^{d/2} |\Sigma_c|^{1/2}} \exp\left[-\frac{1}{2} (I_x - \mu_c)^T \Sigma_c^{-1} (I_x - \mu_c)\right]$$

由(式 2.6)得到每個像素在每種色彩的機率值後，接著計算圖片中每種顏色的水平方向和垂直方向的色彩變異。令 $V(c)$ 為第 c 種顏色在空間中的色彩變異， $V_h(c)$ 為第 c 種顏色在水平方向的色彩變異， $V_v(c)$ 第 c 種顏色在垂直方向的色彩變異， x_h 為像素水平軸座標， x_v 為像素垂直軸座標，我們可得到圖片中各種顏色在圖片中的色彩變異為(式 2.7):

$$V(c) = V_h(c) + V_v(c) \quad (2.7)$$

其中

$$V_h(c) = \frac{1}{|X|_c} \sum_{x=1}^{MN} p(c | I_x) \cdot |x_h - M_h(c)|^2, \quad V_v(c) = \frac{1}{|X|_c} \sum_{x=1}^{MN} p(c | I_x) \cdot |x_v - M_v(c)|^2,$$

$$M_h(c) = \frac{1}{|X|_c} \sum_{x=1}^{MN} p(c | I_x) \cdot x_h, \quad M_v(c) = \frac{1}{|X|_c} \sum_{x=1}^{MN} p(c | I_x) \cdot x_v, \quad |X|_c = \sum_{x=1}^{MN} p(c | I_x)$$

利用(式 2.6)和(式 2.7)得到的 $p(c | I_x)$ 和 $V(c)$ 即可得到色彩空間分佈特徵 f_s ，計算方式為(式 2.8):

$$f_s(x, I) = \sum_{c=1}^c p(c | I_x) \cdot (1 - V(c)) \quad (2.8)$$

最後將 f_s 調整至範圍 [0,1] 即得到區域性的醒目特徵，圖 2.11 即為全域性特徵實例。



圖 2.11 醒目性全域特徵偵測結果，(a)美國國旗原圖，(b)全域性的醒目偵測結果

藉由以上三種特徵來描述出圖片局部的、區域的、全域的醒目區域，並輔以條件隨機域及訓練資料庫，找出三種特徵的最佳權重來將特徵做線性結合即得到圖片的醒目區域。在 T. Liu 文章中所得到的局部性的特徵權重為 0.24、區域性的特徵權重為 0.54、全域性的特徵權重為 0.22，代表區域性的特徵所觀察到的醒目性最為正確，此也符合人類視覺上的判斷。

下一章節我們將描述本文如何利用此方法得到的醒目性來輔佐物件辨識，使物件辨識基於人眼感知，並藉由醒目性的輔助降低外在背景的干擾，使辨識準確率上升。



Chapter 3 辨識方法分析

由第二章的特徵抽取可發現，直接對整張圖做特徵抽取並直接分類的優點有速度快、不受物體位置改變影響、適合即時分析，但直接採用整張圖片的特徵則易受外界背景干擾，造成物件辨識失敗。因此我們將利用醒目性偵測來找出顯著物體，藉由醒目性偵測結果做為特徵相似度比較的權重，降低外在背景的干擾情形，特徵方面則選取 LBP 特徵及色彩特徵，將兩種特徵的卡方距離得到新的卡方距離，最後將新的距離置入 K-最鄰近分類器做分類，判別物件辨識是否成功。

3.1 偵測醒目區域

為了測試偵測醒目區域演算法的效果，我們採用資料庫 Caltech 256 來實驗是否此演算法能正確找出圖片的醒目物體，此資料庫的圖片共分成 256 類，每個種類的照片至少有 82 張以上，每張圖片尺寸不固定且圖片皆為彩圖，由於每張圖片有著明顯類別，因此常被視為物件辨識資料庫。

此小節我們想藉由背景複雜度探討偵測醒目區域演算法的效果，因此我們從 256 類中隨意選出 10 類，每類選取 10 張，5 張為單調背景，5 張為複雜背景，觀察此演算法是否能正確偵測出醒目物體。

3.1.1 背景單調的物體

原圖包含 10 類物體，此 10 類為：輪胎、車子、卡特曼(卡通人物)、螃蟹、木槿花、袋鼠、椰子樹、向日葵、蝴蝶、番茄、斑馬，每類 5 張，每張圖片的背景

單調，見圖 3.1。圖片中的每個像素在經過醒目性偵測後會得到一個醒目分數(salient score)，分數範圍為 $[0,1]$ ，醒目分數越高表示此位置的像素在圖片中的醒目性越大，我們將顯示醒目分數大於 0.5 的像素，醒目分數小於 0.5 的像素則不被顯示，圖 3.2 則為 10 類圖片醒目性偵測結果。

由圖 3.2 發現，在背景單調的情況下，醒目性的偵測結果相當正確，圖中物體皆被視為醒目區域，背景也幾乎都被去除，僅有少數幾張圖片的部分背景未被完全除去，我們推測當物體本身與相鄰背景顏色相似時，相鄰物體的背景容易被視為醒目的物體，如圖 3.3(a)顯示圖庫內第 21 張的袋鼠和背景顏色相似，造成背景的沙地也被視為顯著區域。形狀不規則的物體也會增加醒目性偵測的難度，使物體難以與背景完全切割，如圖 3.3(b)顯示圖庫內第 26 張到第 30 張的椰子樹，由於葉子邊緣較不規則，無法和背景精確的分隔開來。



圖 3.1 背景單調的 10 類圖片



圖 3.2 背景單調的 10 類圖片醒目性偵測結果

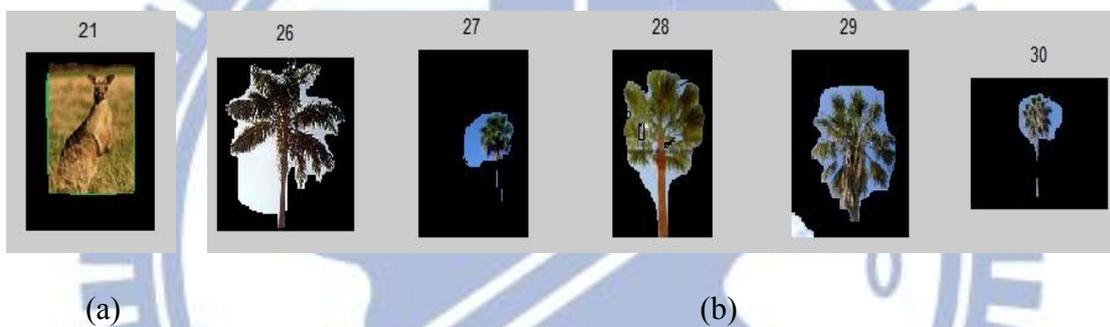


圖 3.3 偵測醒目性效果不佳的圖片，(a)第 21 張袋鼠，(b)26-30 張的椰子樹

3.1.2 複雜背景中的物體

原圖同為 10 類，每類 5 張，每張圖片的背景複雜，見圖 3.4，同樣地，每個像素在經過醒目性偵測後會得到一個醒目分數(salient score)，分數範圍為 $[0,1]$ ，我們將顯示醒目分數大於 0.5 的像素，醒目分數小於 0.5 的像素則不被顯示，圖 3.5 則為 10 類圖片醒目性偵測結果。

由圖 3.5 發現，在背景複雜的狀態下，因為圖片背景的複雜度變大使得醒目性偵測的難度上升許多，以此 10 類別而言，物體偵測的位置大致正確，但受背景混亂的影響，導致上圖許多物體邊界的細部較不完整，尤其是邊緣不規則的物體。

整體而言，醒目性偵測的正確率確實會受背景複雜度影響，複雜背景的圖片較背景簡單的圖片偵測到的物體正確度稍差，但大致能找出物體正確區域，因此本文認為加入醒目性偵測應能對物件辨識有所幫助，而在往後的實驗我們也會加以驗證此項假設。

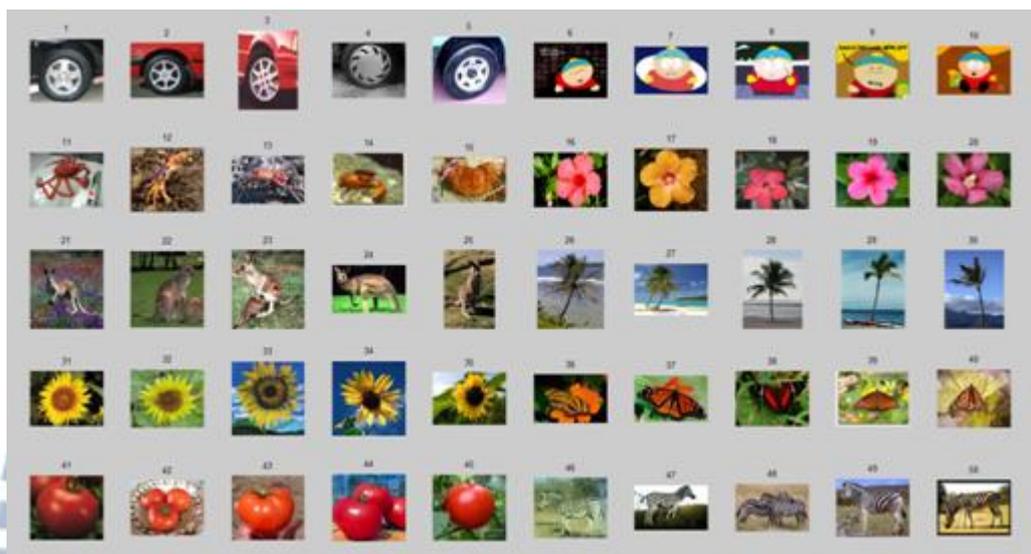


圖 3.4 背景複雜的 10 類圖片



圖 3.5 背景複雜的 10 類圖片醒目性偵測結果

3.2 K-最鄰近分類器

除了不同的特徵抽取方法會影響辨識結果，分類器的不同也會對辨識結果造成影響。在對資料做完特徵抽取後我們將特徵做相似度的比對，然後藉由分類器來進行物件分類的動作。由於物件辨識有同類變異大的特性，因此同類資料也可能並不完全集中，容易三三兩兩分散，因此我們選用 K-最鄰近分類器來做特徵抽取後的分類。

3.2.1 K-最鄰近分類器基本型

K-最鄰近分類法 (K-Nearest Neighbor Classification, 簡稱 KNN) 在 1951 年被提出，是一種簡易的非參數化方法，常應用於物件辨識。假設訓練資料庫有 N 筆資料，每類的資料數為 N_k ，即 $\sum N_k = N$ ，每筆資料各自所屬的類別為 C_k ，當測試資料 x 進入系統時，我們找到與 x 最近的 K 筆資料，若此 K 筆資料內為 C_k 類的資料有 K_k 個，我們可得知各類別佔 K 筆資料內的比例為 K_k / K ，其中擁有最高機率的類別即判別為測試資料的類別，也可以看做 x 為 C_k 類的事後機率 $P(C_k | x) = K_k / K$ 。以圖 3.6 為例，圖中有兩類訓練資料，一類是藍色的正方形，另一類是紅色的三角形，綠色的圓形則是我們待分類的測試資料，當 $K = 3$ 時考慮離綠色圓形最近的 3 筆資料，我們發現紅色三角形個數(2 個)大於藍色正方形(1 個)，因此測試資料判別為紅色三角型的類別。但若當 $K = 4$ ，紅色三角形和藍色正方形個數相同，我們無法判定測試資料是哪一類，為了解決此類問題，下小節介紹加權式 KNN 分類器。

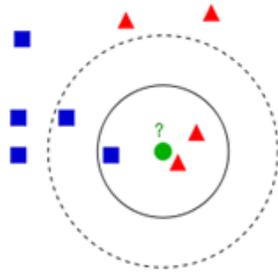


圖 3.6 訓練資料點示意圖

3.2.2 加權式 K-最鄰近分類器

加權式 KNN 分類器 (Distance-weighted KNN, 簡稱 DW-KNN) 的概念為給予離測試資料前 K 筆訓練資料點所投的票賦予權重，離測試資料點越近，給的權重越大，反之越小，我們根據每類所得的票數總和來判定類別。假設現在類別共有 p 類，考慮測試資料最鄰近的 K 的點。這 K 點中每一類別 i 所佔的票數為 K_i ，而這 K 個點距離測試資料的距離分別為 d_{ij} ，其中 $j=1,2,\dots,K_i$ 代表第 i 類中第 j 個點。本論文根據測試資料和訓練資料的距離給予比重 $w_{ij} = 1/d_{ij}$ ，我們可以得到測試資料屬於第 i 類所獲得的分數為 $w_i = \sum w_{ij}$ ，再由各類別分數最高者當選，也就是測試資料屬於這個類別。

3.2.3 卡方距離

將資料置入 KNN 分類器前需先計算資料間的距離，計算距離的方式有許多種，常見的有：歐幾里得距離、卡方距離、餘弦距離、漢明距離...等，本文將採用卡方距離來計算資料間的相似度。卡方距離計算方式如(式 3.1):

$$\chi^2(p, q) = \frac{1}{2} \sum_{i=1}^N \frac{[h_p(i) - h_q(i)]^2}{h_p(i) + h_q(i)} \quad (3.1)$$

其中 p 為訓練資料， q 為查詢資料， h_p 為訓練資料特徵直方圖， h_q 為查詢資料特

徵直方圖， N 為直方圖柱數，此方法取兩直方圖每個柱的差平方，並且在分母做柱的相加，藉以防止當某個柱在整個直方圖佔的比例過高時，卡方距離值受此柱的影響過大。由於計算方便及易於比較兩兩直方圖的曲線趨勢，卡方距離常被應用於直方圖的相似度度量。

3.3 以色彩直方圖特徵作辨識

在此小節我們想以色彩特徵來做物件辨識，色彩對人眼來說也是個顯而易見的特徵，但由於易受背景干擾，因此我們引入上一個小節描述的偵測物體醒目性，先找出圖片中的醒目物體，再對其抽取色彩特徵，而測試的資料為上一小節採用的 Caltech256 資料庫，類別也與上述相同共有 10 類，每類 10 張圖片，5 張為簡單背景，5 張為複雜背景，藉以探討色彩特徵的特性及加入顯著性偵測前偵測後的辨識度。

3.3.1 色彩特徵的特性探討

由於醒目性的偵測不會對色彩特徵的參數造成影響，因此在這個小節我們先不加入醒目性的偵測，直接以原圖抽取色彩特徵做探討。色彩特徵唯一的參數為色彩直方圖的柱數，假設色彩範圍劃成 N 個等分，在 RGB 色彩空間可得到長度為 2^N 的色彩直方圖。由於色彩直方圖的柱數會直接影響到執行時間及辨識準確度，因此我們將以 3.2 節提出的 10 類圖片來探討色彩直方圖柱數的影響，對資料庫內每張圖片做物件辨識，辨識方式為將每張圖與所有圖做相似度的比較，若圖片最後的分類結果正確即代表辨識成功，一次完整的執行時間為計算單張圖片與資料庫所有圖片色彩直方圖相似度所花的時間，相似度度量方式採用卡方距離，分類

方法以加權式 KNN 做分類，K 取 3，圖 3.7 為實驗辨識準確率的結果。

由圖 3.7 可發現，當直方圖柱數太少，色彩範圍間距過大使直方圖的統計失準，導致低辨識率，當柱數上升時，辨識度也急速上升，選取適當色彩直方圖柱數可得到最佳的辨識度，若柱數過多，色彩範圍間距太小也會使得直方圖統計太過瑣碎導致辨識度再度下降。圖 3.8 為在不同直方圖柱數時，每張圖片完成辨識的平均執行時間，由圖可觀察出當 bin 數上升時，執行時間也急遽上升，且柱數過多對辨識準確率並沒有幫助，因此選取適當的直方圖柱數能有效降低執行時間並且達到最佳的辨識準確率。

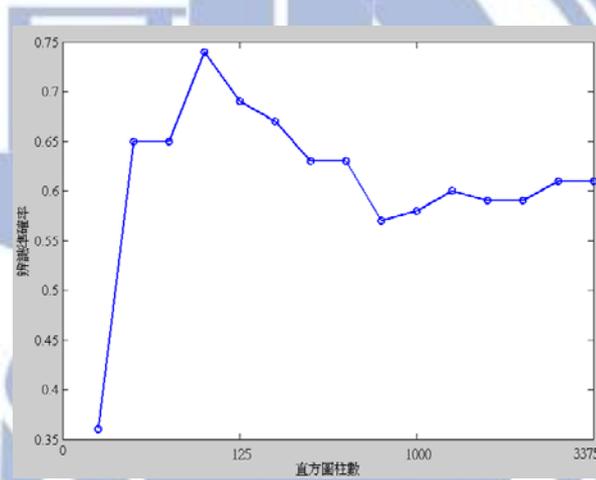


圖 3.7 直方圖柱數對辨識準確率的影響

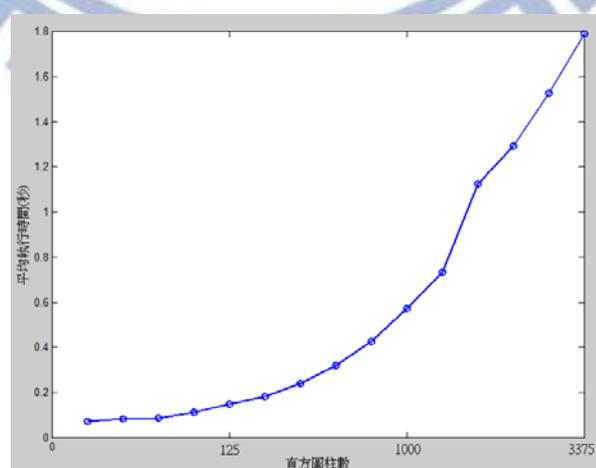


圖 3.8 直方圖柱數對時間的影響

3.3.2 結合醒目性偵測與色彩特徵

在這個小節，我們先對資料庫的所有圖片進行醒目性的偵測，藉以得到圖片內每個像素的醒目性分數，接著對原圖片抽取色彩特徵，在統計色彩特徵直方圖時，將圖片每個像素的醒目性分數做為直方圖的權重，藉由低分的像素將背景的干擾降低，並觀察去除背景干擾後，色彩特徵是否適合作為圖片的特徵。

此實驗我們取色彩直方圖柱數 7^3 ，資料庫同為上小節所述，分類的方法也與上小節相同，以加權式 KNN 做分類，K 取 3，表 3-1 為實驗辨識結果。

表 3-1 色彩特徵辨識準確率比較表

方法	辨識準確率(%)	
	不含醒目性偵測	含醒目性偵測
1 輪胎	70	80
2 Cartman	80	100
3 螃蟹	70	90
4 木槿花	70	60
5 袋鼠	40	80
6 椰子樹	80	70
7 向日葵	60	80
8 蝴蝶	20	30
9 番茄	50	80
10 斑馬	80	80
平均準確率	62	75

由上表可以發現其每類經過醒目性偵測後皆較偵測前辨識結果佳，只有第 4 類辨識率略低於偵測前，觀察第 4 類的圖片，發現由於第 4 類的 10 張圖片背景相似，且同類木槿花的顏色不同，若同類的色彩變異太大，則色彩特徵無法將醒目性偵測後的區域視為同一類，造成第 4 類偵測醒目性後卻略低於偵測前的結果，因此若同類的色彩變異太大，我們可能需輔以其他特徵來做辨識才能使辨識率上升。但就平均準確率來說，含醒目性偵測的準確率(75%)優於不含醒目性偵測的準

確率(62%)，顯示醒目性的偵測在辨識物體上的確有其顯著效果，在此 10 類的測試上，色彩特徵的準確率也尚可接受。

3.4 以局部二值模式特徵作辨識

此小節我們想以局部二值模式來做物件辨識，局部二值模式為目前常用的紋理特徵抽取方式，它有計算量小、不受物體位置影響、容易比較相似度等優點，但由於易受背景干擾，常被應用於圖片變異較小的人臉辨識而較少應用於物件辨識。因此我們想引進醒目性的偵測來降低背景干擾增進辨識準確率，並利用局部二值模式計算快速達到系統的即時效果。測試的資料為上一小節採用的 Caltech256 資料庫，類別也與上述相同，共有 10 類，每類有十張圖片，5 張為簡單背景，5 張為複雜背景，藉以探討局部二值模式參數特性及加入醒目性偵測前偵測後的辨識度。

3.4.1 局部二值模式參數特性探討

局部二值模式的參數有局部空間的半徑、周圍取點數和編碼一致性(uniform)，當描述局部空間的半徑上升，周圍取點數也須跟著上升，對局部空間的描述才夠精確，但周圍取點數的上升會使特徵長度變長，使得執行時間上升，因此適當的空間半徑、取點數才能兼顧辨識度及時間複雜度。

由於醒目性的偵測對局部二值模式的參數並不造成影響，因此我們在這個小節先不加入醒目性的偵測，直接以原圖來做局部二值模式的參數探討，我們取編碼一致性(uniform)為 2，不同的 LBP 尺寸以 $LBP_{P,R}^{u2}$ 表示， $u2$ 代表一致性為 2， P

代表周圍取點數， R 代表空間半徑，藉由改變 R 及 P 觀察辨識準確率和空間半徑及周圍取點數是否有關。資料庫如前小節所述，分類的方法也與上小節相同，以加權式 KNN 做分類， K 取 3，表 3-2 為實驗辨識結果。

表 3-2 不同 LBP 尺寸辨識準確率

	平均辨識準確率 (%)
$LBP_{4,1}^{u^2}$	58
$LBP_{8,1}^{u^2}$	62
$LBP_{8,2}^{u^2}$	55
$LBP_{16,2}^{u^2}$	57

由表發現，半徑的改變對這 10 類圖片沒有太大的影響，推測由於此 10 類圖片的尺寸大小並無太大差異，將半徑改變對辨識率並無太大幫助。把周圍取點數上升也無法有效提高辨識率，且提高周圍取點數將使執行時間上升，因此改變空間半徑及周圍取點數在此 10 類物件辨識上並無幫助，而下小節本文將探討若將圖片輔以醒目性偵測是否能有效幫助物件辨識。

3.4.2 結合醒目性偵測與局部二值模式

在這個小節，我們會先對資料庫所有圖片做醒目性區域的偵測藉以得到圖片內每個像素的醒目性分數，接著對原圖片抽取 LBP 特徵，在統計特徵直方圖時，將圖片每個像素的醒目性分數做為直方圖的權重，藉由低分的像素將背景的干擾降低，並且觀察免去背景干擾後，是否為適合的物件辨識特徵。此實驗我們將 LBP 特徵半徑 R 取 1，周圍取點數取 8，一致性取 2，資料庫同為上小節所述，分類的方法也與上小節相同，以加權式 KNN 做分類， K 取 3，表 3-3 為實驗辨識結果。

表 3-3 LBP 特徵辨識準確率

方法	辨識準確率(%)	
	不含醒目性偵測	含醒目性偵測
1 輪胎	70	80
2 Cartman	90	90
3 螃蟹	70	80
4 木槿花	80	60
5 袋鼠	30	60
6 椰子樹	70	70
7 向日葵	70	60
8 蝴蝶	50	30
9 番茄	50	70
10 斑馬	40	80
平均準確率	62	68

由上表發現，僅 4、7、8 類略低不含醒目性偵測方式，觀察第 4、7、8 類，發現圖同類圖片背景相似，由於相似的背景將導致對圖片抽取全域特徵反而能輔助辨識，因此當醒目性偵測結果不佳時，辨識準確率未必較不含醒目性偵測佳，因此辨識的準確度十分端賴醒目性偵測的效果。但整體來說，含醒目性偵測的圖片平均準確率(68%)仍優於不含醒目性偵測的圖片(62%)，因此以此節的 10 類物件辨識來說，醒目性偵測確實有其效果，單就辨識效果而言，LBP 特徵的辨識度尚可接受。

3.5 醒目性偵測結合多種特徵

在前面的描述我們發現經過醒目性偵測的確能使物件辨識的準確率上升，但由於同類物體的變異極大，單用一種特徵來辨識物體是不夠的，因此我們想利用多種特徵以多個面向來辨識物體以提高辨識準確率。在前兩節中介紹的色彩特徵與局部二值模式特徵各有優缺，由於局部二值模式的特徵抽取法為統計經常出現的局部空間架構，它能描述像素間的大小關係，但沒有描述像素色彩值，當兩張圖片像素色彩值不同，但像素間的大小關係恰好相同時便會造成誤判，而此缺點能以色彩特徵彌補，經由色彩特徵來描述像素的色彩值便能克服此問題。又當同類物體的顏色不同，但若像素局部空間架構相同時，局部二值模式特徵便能降低同類差異，拉近同類但色彩不同的物體距離，因此我們認為色彩特徵和局部二值模式特徵有互補效果，混合使用來做物件辨識應能達到準確率的提升。

本節將綜合 3.1 節的醒目性偵測，搭配 3.3 節的色彩特徵，以及 3.4 節的局部二值模式特徵，提出本論文採用的物件辨識法，流程圖為圖 3.9，以下介紹演算法的步驟：

1. 尋找訓練資料與測試資料的醒目區域。

利用局部的、區域的、全域的特徵來找出圖中的醒目區域，藉以去除圖片背景的干擾。經過偵測後圖片中的每一個像素會得到一個醒目分數，分數越高代表該位置的醒目性越大。

2. 對訓練資料及測試資料抽取色彩特徵，以圖片中每個像素的醒目性做加權用以統計直方圖當作色彩特徵。
3. 將訓練資料及測試資料轉成灰階圖，並抽取局部二值模式特徵，以圖片中每個

像素的醒目性做加權用以統計直方圖當作材質特徵。

4. 求訓練資料及測試資料的色彩特徵與局部二值模式特徵兩兩圖片卡方距離，將色彩距離及 LBP 距離調整至範圍[0,1]。
5. 調整 LBP 距離 D_{LBP} 權重 λ 與色彩距離 D_{color} 權重 $(1-\lambda)$ 得到新的卡方距離 D ，如式(3.2)，其中 λ 介於[0,1]之間。

$$D = \lambda D_{LBP} + (1-\lambda) D_{color} \quad (3.2)$$

6. 以距離 D 做 K 最鄰近分類法完成物體辨識。

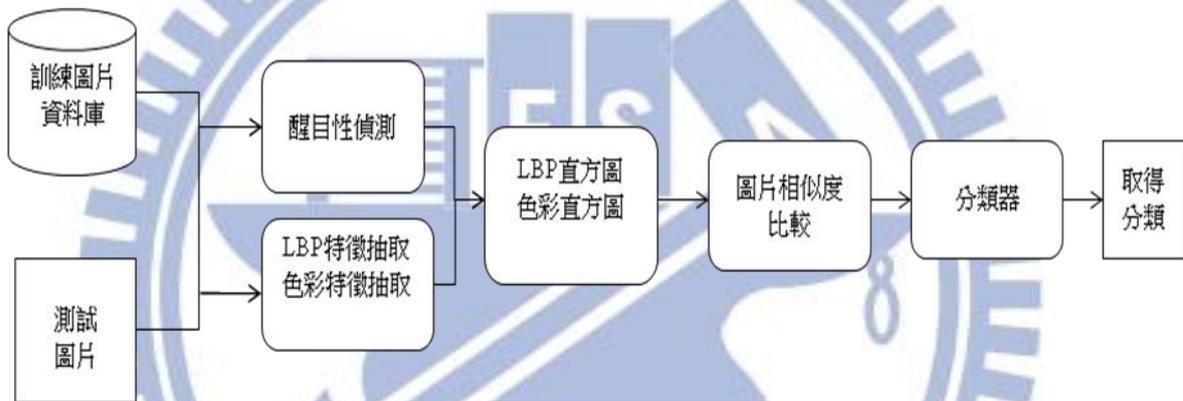


圖 3.9 實驗流程圖

以下為本章前面小節使用的 10 類資料辨識結果，辨識方法採用本文上述提出的演算法，色彩特徵參數 bin 取 7，LBP 特徵參數半徑 R 取 1、周圍點各數 P 取 8 及一致性 U 取 2，LBP 卡方距離權重取 0.9，辨識結果如表 3-4。觀察表 3-4 可發現，將兩種特徵距離結合，取 LBP 距離權重 0.9 得到的辨識結果(78%)皆較只單依色彩特徵距離(75%)或 LBP 卡方距離(68%)佳，此即驗證了 LBP 特徵和色彩特徵的互補性。

表 3-4 結合特徵距離前後辨識準確率比較表

	辨識準確率(%)		
	含醒目性偵測 以色彩距離	含醒目性偵測 以 LBP 距離	含醒目性偵測 結合兩特徵 距離
1 輪胎	80	80	80
2 Cartman	100	90	90
3 螃蟹	90	80	90
4 木槿花	60	60	70
5 袋鼠	80	60	90
6 椰子樹	70	70	70
7 向日葵	80	60	90
8 蝴蝶	30	30	30
9 番茄	80	70	80
10 斑馬	80	80	80
平均準確率	75	68	78

接著我們觀察圖 3.10，x 座標為 LBP 卡方距離所佔的權重，座標最左邊為完全以色彩卡方距離來做最鄰近分類，往右 LBP 權重漸增，最右邊完全以 LBP 卡方距離做最鄰近分類，由圖可觀察出此 10 類圖片以色彩特徵做辨識較 LBP 特徵佳，色彩特徵平均辨識準確率在此 10 類圖片經醒目性偵測後可高達 75%，LBP 特徵的平均辨識準確率也有 68%，若將兩種特徵的卡方距離作權重調整得到新的卡方距離，則能使辨識準確度又更加上升，此實驗結果吻合我們所提出的色彩特徵與 LBP 特徵有其互補性質，不論從原理或實驗結果皆可得到實證。

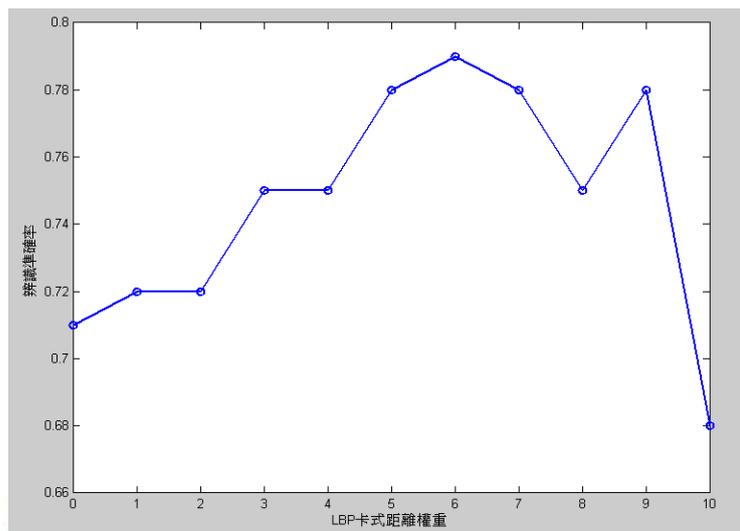


圖 3.10 不同 LBP 卡方距離權重對辨識準確率的影響

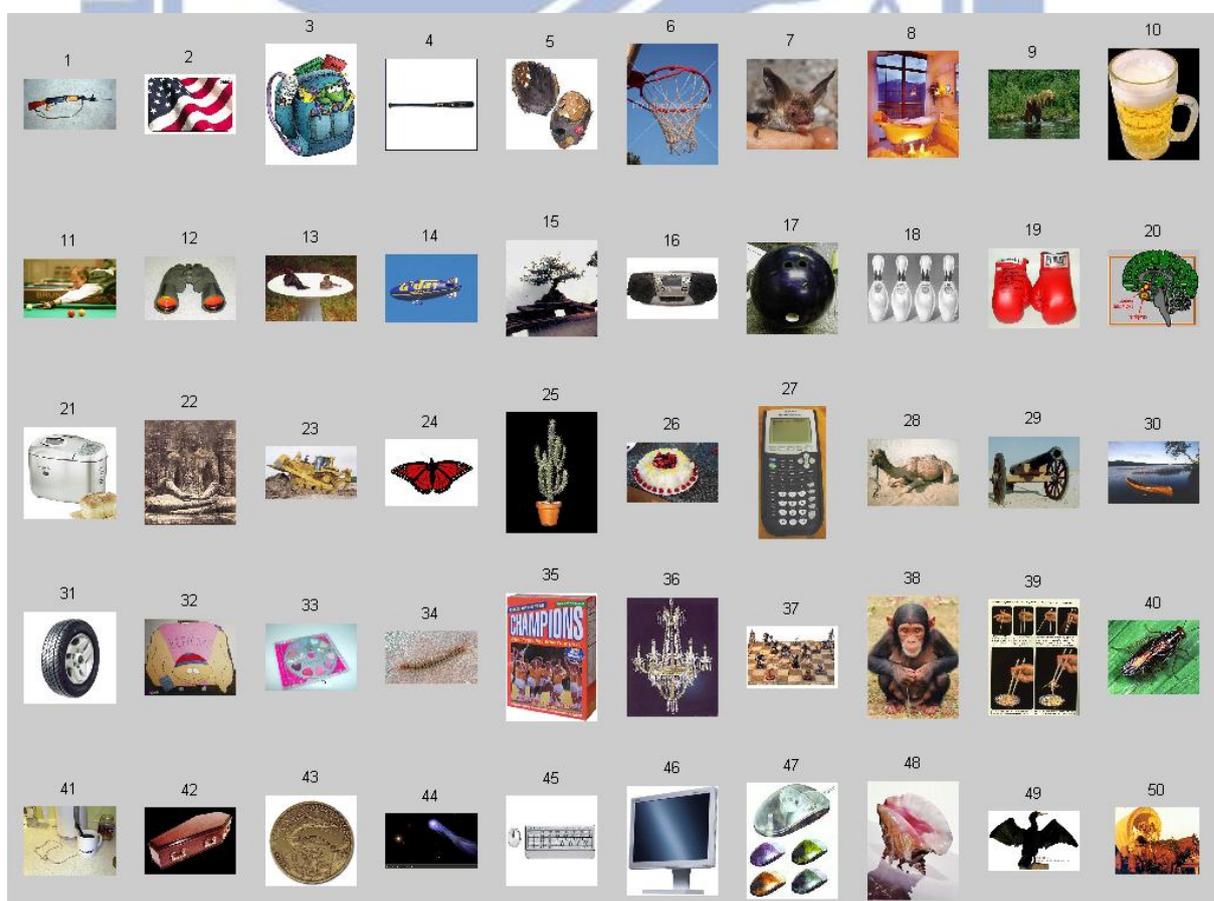
而在下一章我們將把圖片辨識資料庫類別放大，用以驗證在資料量不斷上升的同時，本實驗演算法仍能準確的辨識物體。



Chapter 4 實驗結果

在上個章節我們以色彩特徵及 LBP 特徵來對 10 個類別的資料庫進行物件辨識，並在 3.5 節驗證本文提出的演算法更能使得辨識準確度有效上升。但由於物件辨識的變異性相當大，若更改或置入不同類別的圖片也可能造成物件辨識準確度大幅變動，因此我們擴大原有的資料庫類別，藉以觀察在物件類別數量上升的同時，本文提出的方法是否仍有效使辨識準確率上升。

本章採用 Caltech256 資料庫 1-70 類來做實驗，取每類的編號 1 至編號 20 的圖片來做物件辨識，每張圖片大小不同且皆為彩圖，圖 4.1 為 70 類物件代表圖片，接下來的每小節我們將以此 70 類物件做各種分析探討。



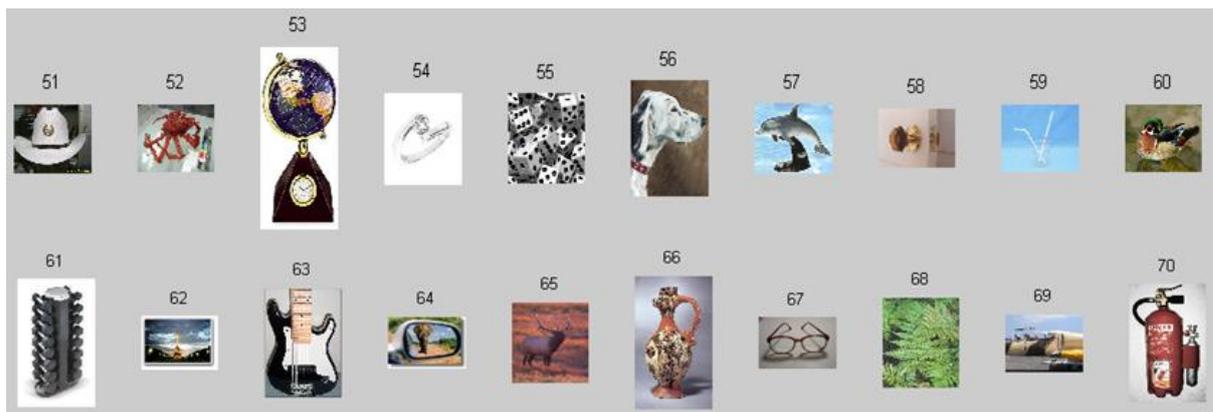


圖 4.1 物件 70 類代表圖片

4.1 結合不同特徵的卡方距離

在第三章我們提出結合 LBP 特徵的卡方距離及色彩特徵的卡方距離可提高物件辨識準確率，並以 10 個物件類別的小實驗證實我們的假設。但物件辨識的組間變異相當大，當物件類別數目上升或類別改變時都會改變辨識準確率，因此我們在此節探討當物件類別數上升或類別改變時，結合此兩種特徵的卡方距離是否有效提升辨識準確率。

我們從資料庫內 70 類圖片隨機挑選出 20 類，每類圖片皆偵測醒目性區域，並對原圖片抽取色彩特徵及 LBP 特徵，色彩特徵直方圖柱數= 8^3 ，LBP 特徵半徑= 1 、周圍點數= 8 、一致性= 2 ，以偵測到的醒目性區域做權重進而統計出 LBP 直方圖及色彩直方圖並計算 LBP 特徵的卡方距離及色彩特徵的卡方距離，最後調整兩卡方距離權重以得到新的卡方距離，以 K-最鄰近分類法完成物體辨識，取 $K=5$ 。圖 4.2 為在不同權種情況下，隨機從資料庫 70 類內抽取 20 類 10 次的辨識準確率圖。

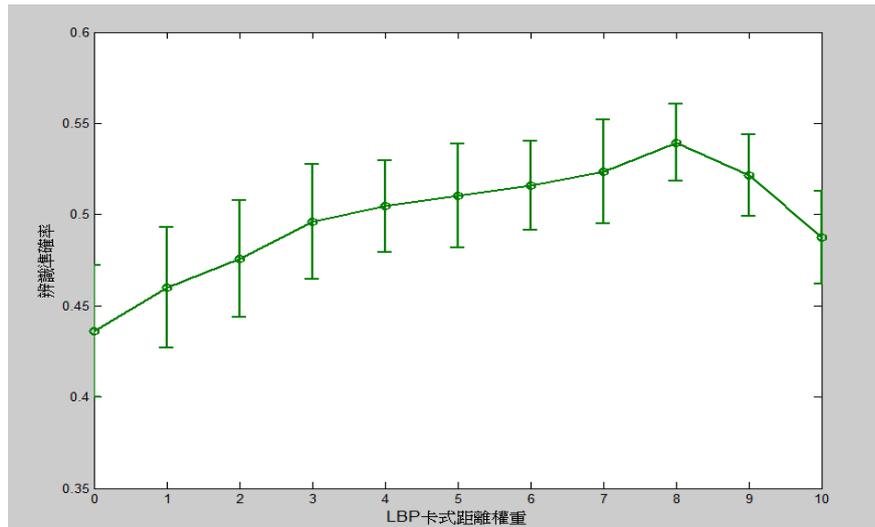


圖 4.2 隨機抽取資料庫 20 類十次的辨識準確率

基於易讀性的考量，我們只在圖 4.2 繪出隨機抽取 10 次的辨識準確率分佈範圍，此章之後的所有實驗仍隨機抽取資料庫類別 10 次，但曲線圖將只繪出隨機抽取 10 次的平均準確率圖，使實驗結果易於觀察。

圖 4.3 為在擴大資料庫類別下，改變 LBP 卡方距離權重造成的辨識準確率變化圖，x 軸為 LBP 卡方距離所佔權重，x 軸最左邊 LBP 卡方距離權重=0，色彩卡方距離權重=1，最右邊為 LBP 卡方距離權重=1，色彩卡方距離權重=0，y 軸為辨識準確率。觀察圖 4.3 可發現單獨以 LBP 卡方距離或色彩卡方距離分類時得到的辨識準確率皆較結合使用時來的低，此及驗證了兩者的互補性，且 LBP 卡方距離辨識結果較色彩卡方距離佳，此和第三章的以 10 類資料庫做辨識的結果不同，我們推測若物件類別大量上升，以色彩特徵來做判斷較容易造成取偽的情形，因為即使是不同類的物體也有可能產生相同的色彩直方圖，同類的物體也可能色彩直方圖完全不同，使取偽的機率在類別上升時大大的提升，因此在物件類別增加時，以 LBP 卡方距離做分類的效果反而比色彩卡方距離好。將兩卡方距離結合的同時，辨識準確率明顯提升，在不同類別數時皆可得到此結果，觀察上圖在類別數

從 20 上升到 60 時，LBP 卡方距離權重=0.8 或 0.9 時可得到最佳的辨識準確率，此也驗證了 LBP 特徵及色彩特徵的互補性。

另一方面，考慮某些圖片可能過於容易辨識或太難辨識，本文將辨識準確率計算方式取類別辨識率的中位數來代表辨識準確率，實驗結果圖為圖 4.4，我們仍可發現將 LBP 特徵及色彩特徵合併的實驗結果較只選取單一特徵做辨識佳，將辨識率取中位數後，辨識準確率的最高值大約位於 LBP 特徵權重 0.6~0.9 之間。

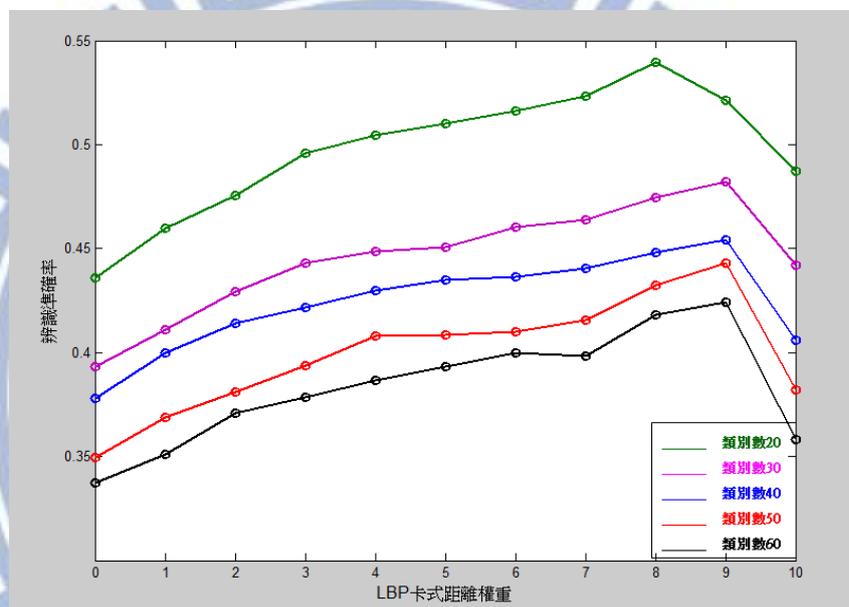


圖 4.3 不同 LBP 卡方距離權重對辨識準確率的影響

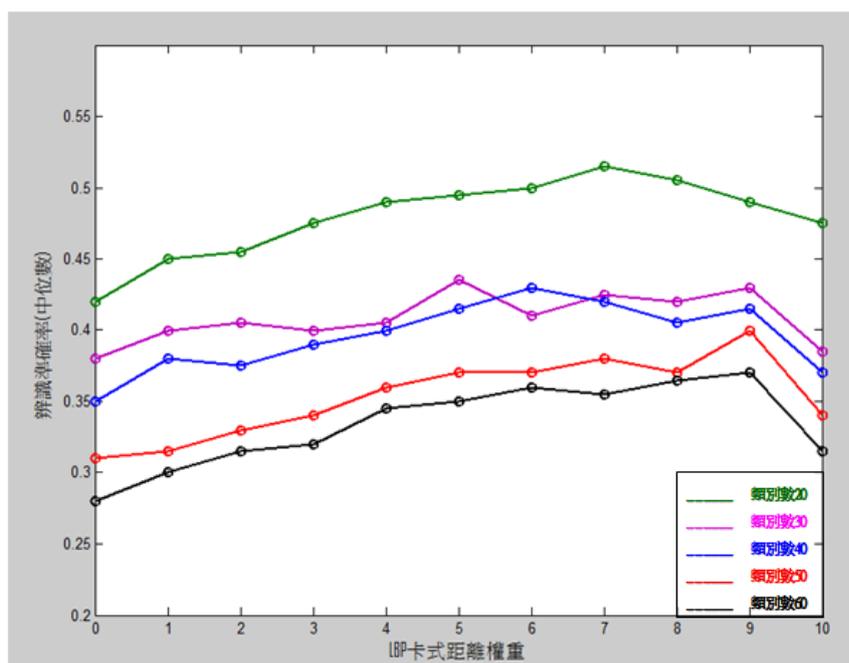


圖 4.4 不同 LBP 卡方距離權重對辨識準確率(取中位數)的影響

4.2 K 最鄰近分類法分類結果

在物件辨識的分類我們採用 K-最鄰近分類法，K-最鄰近分類法做法為比較資料兩兩的特徵距離並計算離本身最近的 K 筆資料，其中佔最多的那一類即為查詢資料的類別，由於只選定相似度的前 K 名來做分類依據，適合在同類資料變異大時使用。

我們採用 3.2.2 節介紹的加權式 KNN 來做分類，此節我們想探討不同的 K 值是否會影響辨識結果。此節用的色彩特徵柱數= 8^3 ，LBP 特徵半徑=1，周圍取點數=8，一致性=2，LBP 卡方距離權重=0.9，圖 4.4 為在擴大資料庫類別下，改變 K-最鄰近分類法造成的辨識準確率變化圖。

由圖 4.5 可看出當 K 值越大辨識準確率越低，代表當我們取的 K 值越大，越容易產生取偽的情形，我們猜測由於 Caltech256 資料庫內同類物件在外型上也可

能有很大的差異，所以資料在抽取特徵後同類並不會全部聚集在一起，仍可能會三三兩兩散佈，因此若取太大的 K 值會使辨識準確率下降，在類別數上升時的情況也相同， K 值上升依然造成辨識準確率的下降，但 K 值下降的幅度卻有趨緩的情形。上圖我們還可發現當類別數越大，物件辨識的複雜度越高，辨識準確率也越來越低，但類別數上升時，辨識準確率的下降值有減小的趨勢。

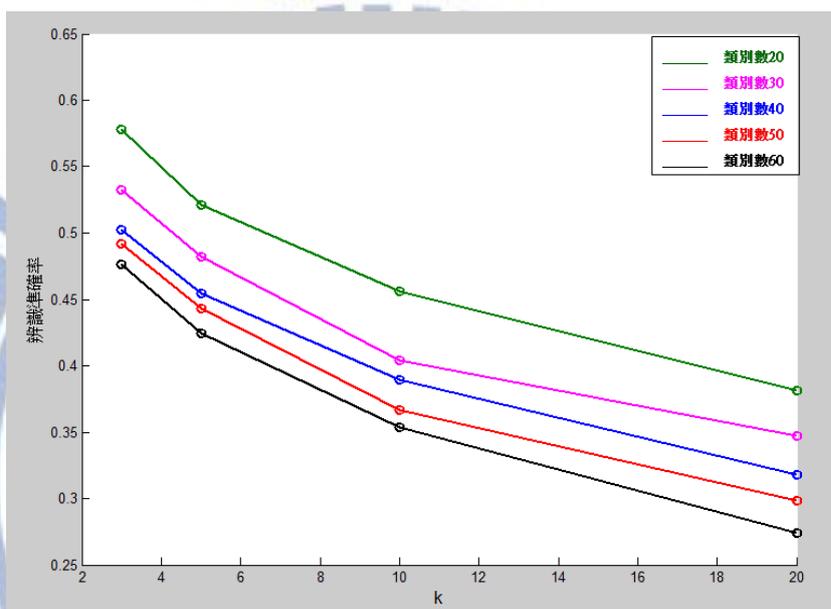


圖 4.5 分類器不同 K 值對辨識準確率的影響

4.3 綜合比較

在此節我們想比較各種不同方法的辨識結果，藉以觀察本文提出的物件辨識演算法是否明顯提升辨識準確率，我們將比較加入醒目性偵測前後的差異及單獨使用一種特徵，或結合不同特徵來辨識圖片，以及增加類別數目時各種方法的辨識準確率變化，藉以分析本演算法的優缺點。

以下我們將列出在資料庫 70 類中隨機抽取 20 類、30 類、40 類...至 60 類 10

次，各種方法的辨識結果，我們所選用的參數為色彩特徵直方圖柱數=8³，LPB 特徵的半徑=1，周圍取點數=8，一致性=2，LBP 卡方距離權重=0.9 及選取 K=5 來做最鄰近分類法，圖 4.6 為我們隨機抽取不同類別數 10 次所得到的平均準確率值。最上方藍色曲線為本文提出的演算法辨識結果，紅色曲線為對圖片做偵測醒目性且抽取 LBP 特徵的辨識結果，綠色曲線為不對圖片做醒目性偵測直接抽取 LBP 特徵的辨識結果，黑色曲線為對圖片偵測醒目性並抽取色彩特徵的辨識結果，紫色曲線為不對圖片做醒目性偵測直接抽取色彩特徵的辨識結果。

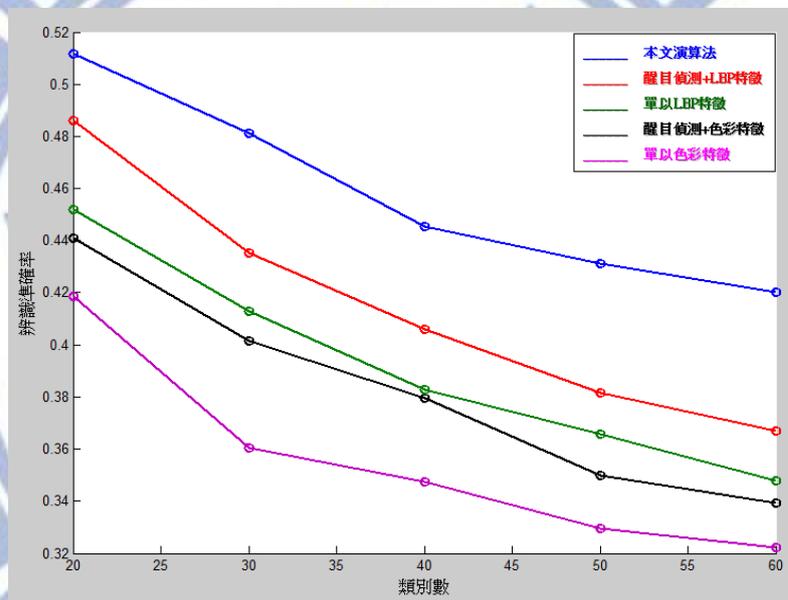


圖 4.6 各種辨識方法的辨識準確率

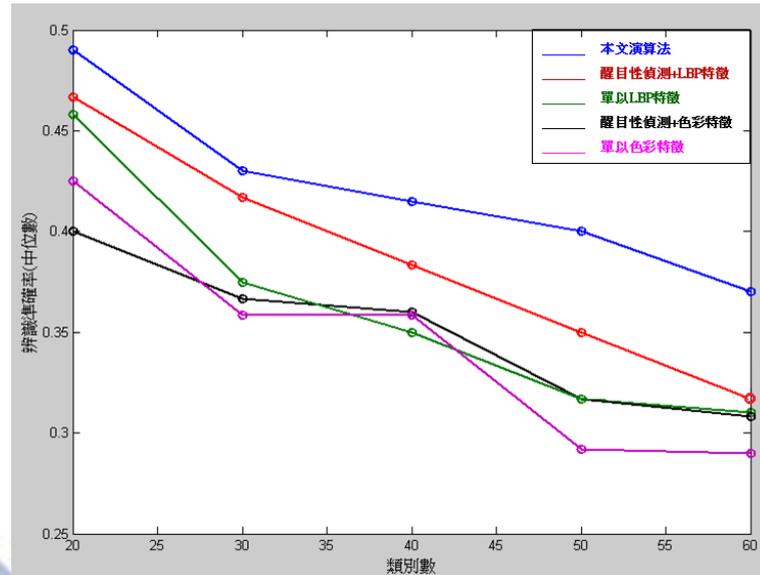


圖 4.7 各種辨識方法的辨識準確率(取中位數)

由圖 4.6 可發現本文提出的演算法(藍色曲線)明顯改善辨識準確率，在類別數=20 時辨識準確率為 53%，隨著類別數上升至 60 辨識準確率剩下 46%，顯示當類別數上升時資料分類難度變高辨識準確率隨之下降，但觀察藍色曲線其下降幅度卻有趨緩的情形，且此情形較其他曲線來的更明顯，推測由於醒目性的偵測及卡方距離的結合，使得同類圖片的相似度上升，即使類別數增大，同類間的相似度夠大以致不受新加入的類別影響，因此造成類別數上升時，辨識準確率下降的幅度趨緩，也可明顯的觀察出本文提出的演算法與其他方法準確率在類別上升時差距更大。

再來觀察加入醒目性偵測前後的結果差異，加入醒目性偵測並以 LBP 特徵分類(紅色曲線)的辨識準確率比未偵測醒目性的綠色曲線、紫色曲線結果佳，但加入醒目性偵測並以色彩特徵做分類的黑色曲線準確率卻比未偵測醒目性的綠色曲線準確率差，由於綠色曲線抽取的是 LBP 特徵，黑色曲線抽取的是色彩特徵，由前幾小節的推論得知，當類別數上升時，色彩特徵的辨識結果比 LBP 特徵不理想，

這是由於當類別數上升時，只提取色彩特徵當圖片的特徵向量不夠可靠，當類別數漸增即使不是同類物體也可能有色彩相同的情形，此會使資料點的組內變異大組間變異小，因此即使有做偵測醒目性的處理，在只以色彩特徵做分類的狀況下，辨識準確率仍不盡理想。

由於色彩特徵在類別數上升或同類物體色彩變異太大時並不可靠，在單獨使用的情況皆不理想，且由於本文認為色彩特徵和 LBP 特徵有其互補的效果，並在前面的實驗皆獲得驗證，因此以 LBP 特徵計算的卡方距離權重=0.9，色彩特徵計算的卡方距離權重=0.1 來當作圖片與圖片間的距離。由圖 4.6 觀察結果，當類別數目不斷上升時，結合兩種特徵卡方距離的方式(藍色曲線)明顯比只單用一種特徵去做辨識的效果佳，此再度證實了色彩特徵與 LBP 特徵的互補特性，即使在資料庫類別不斷增大的情況下，仍能得到相同結果。

再來我們改變辨識準確率的計算方法，將每類的辨識結果取中位數做為辨識準確率，得到圖 4.7 的實驗結果，觀察結果可發現將辨識準確率取平均值得到的效果比取中位數來的佳，這是由於某些圖片可能太過容易辨識或難以辨識，因此取中位數做為辨識準確率或許更加具有可靠性，且在醒目性偵測及多重特徵的互補之下辨識準確率仍明顯上升

因此本文的演算法在此章驗證了在大資料庫的情況下，偵測圖片的醒目性的確能幫助物件辨識，而且使用色彩特徵及 LBP 特徵除了符合即時運算的功能，將兩種特徵的卡方距離結合再度增進物件辨識的準確率，且在 LBP 特徵的卡方距離權重遠大於色彩特徵的卡方距離權重情況下，物件辨識可達到較佳的效果。

Chapter 5 結論

在現實生活中影像常受到外在環境大量干擾，造成物件辨識的不容易，因此本文提出藉由醒目性區域偵測來對輔助物件辨識，使外在環境的干擾下降，其方法為利用醒目性的偵測結果做為相似度比較的權重，藉由醒目性物體可獲得較大的權重，不重要的背景的權重較低來降低外界環境的干擾，在物體相似度的比較上，則選取 LBP 特徵及色彩特徵做卡方距離的結合來代表圖片之間的相似程度，選取 LBP 特徵及色彩特徵的用意為由於此兩種特徵有著互補效果，結合兩者使用能讓辨識效果更加顯著，且由於本文的訴求為模擬出一即時物件辨識系統，特徵抽取的速度為一大考量，過於精細煩雜的特徵抽取方式並不適合，LBP 特徵及色彩特徵一大優點即為特徵抽取快速，此和本文需求不謀而合，因此選取此兩種特徵做為本文特徵抽取方式。

本文以 Caltech256 資料庫做物件辨識實驗，在醒目性偵測方面，利用多種醒目特徵來做醒目區域偵測能正確找出物件在圖片的區域，藉由醒目性偵測的結果來做相似度的權重也的確能有效降低圖片背景的干擾，我們在觀察加入醒目性偵測前後的實驗結果後，也證實醒目性偵測確實能改善辨識準確率。再者，利用色彩特徵與 LBP 特徵的互補性將兩特徵做結合也提高了辨識準確率，使物件辨識在可靠度及速度上都得到上升。

本論文雖然藉由醒目性偵測及特徵的結合使得同類圖片間的相似度提高，但仍有許多值得改善的空間，得以做更多的探討：

1. 當醒目性偵測的結果不佳，便無法降低外在環境的干擾，甚至會造成同類圖

片的相似度降低，且若同類圖片的背景十分相似，再加入醒目性偵測的結果不一定優於不加入醒目性偵測。

2. 對一張圖片做醒目性偵測需花費 30 秒左右，若要對訓練資料庫的每張圖片做醒目性偵測將十分費時，因此如何有效降低醒目性偵測執行時間仍是此方法的重點。
3. 色彩直方圖只能描述色彩分佈的情形，但無法描述其色彩在圖片中的位置情形，只靠色彩分佈的描述十分容易造成取偽的現象。
4. LBP 特徵在描述同一材質時十分可靠，但一個物體可能由各種材質構成，因此以 LBP 特徵來做相似度比較也具有一定風險，若同類物體的變異太大將導致辨識結果不佳。



參考文獻

- [1] J. Swain and H. Ballard, "Color Indexing," *International Journal of Computer Vision*, 7:1, pp.11 - 32 , 1991.
- [2] T. Leung and J. Malik, "Representing and Recognizing the Visual Appearance of Materials using Three-dimensional Textons" *International Journal of Computer Vision* 43(1), pp.29 - 44, 2001.
- [3] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns" *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, 2002.
- [4] S. Belongie, J. Malik, and J. Puzicha, "Shape Matching and Object Recognition Using Shape Contexts" *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 24, 2002.
- [5] G. Lowe, "Distinctive Image Features from Scale-Invariant keypoints," *International Journal of Computer Vision* 60(2), pp. 91 - 110, 2004.
- [6] H. Bay, Andreas Ess, T. Tuytelaars, L.V. Gool, "SURF: Speeded Up Robust Features," *Computer Vision and Image Understanding (CVIU)*, vol. 110, no. 3, pp.346 - 359, 2008.
- [7] C. Gu, J.J. Lim, P. Arbelaez and J. Malik, "Recognition using Regions," *Computer Vision and Pattern Recognition*, 2009.
- [8] L.-J. Li, R. Socher and L.-F. Fei, "Towards Total Scene Understanding : Classification, Annotation and Segmentation in an Automatic Framework," *Computer Vision and Pattern Recognition (CVPR)*, 2009.
- [9] O. Russakovsky, Y. Lin, K. Yu and L.-F. Fei, "Object-centric spatial for image classification," *European Conference on Computer Vision (ECCV)*, 2012.

- [10] T. Liu, Z. Yuan, J. Sun, J. Wang, and N. Zheng, "Learning to Detect a Salient Objects," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 33, no. 2, 2011.
- [11] M. Haralick, K. Shamugam, and I. Dinstein, "Textural Features for Image Classification," IEEE Trans. on Systems, Man, and Cybernetics. SMC-3 (6), pp.610 – 621, 1973.
- [12] T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions", Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR), vol. 1, pp. 582 – 585, 1994.
- [13] T. Ojala and M. Pietikainen, "Unsupervised texture segmentation using feature distributions," Pattern Recognition Society, vol. 3, pp. 477 – 486, 1999.
- [14] V. Takala, T. Ahonen, and M. Pietikainen, "Block-based methods for image retrieval using Local Binary Patterns," In Proc. 14th Scandinavian Conference on Image Analysis (SCIA), pp.882 – 891, 2005
- [15] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, 2006
- [16] A. Hadid, M. Pietikainen, and T. Ahonen, "A Discriminative Feature Space for Detecting and Recognizing Faces," Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition .