

利用小波變換與顏色字串之類似郵票檢索

Using Wavelet Transform and Color Strings for Stamp Retrieval

林群雄*

Chiun-Hsiun Lin

(收件日期 107 年 6 月 14 日；接受日期 108 年 3 月 11 日)

摘 要

我們的郵票檢索方法可以很有效的檢索出類似的郵票。首先，我們將每張郵票先作二級小波變換的 LL2：粗略近似的圖像取出，再將其尺寸標準化為 60×60 像素，再用 8 條規則（利用 RGB 三原色之相對關係）將每張郵票轉換成顏色字串，並儲存於資料庫。其次，我們利用字串比對進行比較要檢索的郵票的顏色字串與儲存於資料庫的顏色字串。最後，排名前 10 名相似的結果將被檢索出來。我們的系統同時保有：1. 基於內容的郵票檢索系統（基於相似性檢索）和 2. 基於文本的郵票檢索（非常迅速和成熟）兩者的既有優勢。

關鍵詞：小波變換、字串比對、基於內容的郵票檢索、基於內容的圖像檢索 (CBIR)、顏色字串。

*國立臺北大學研究發展處副研究員

Abstract

In this paper, we demonstrate a system for efficient retrieval of stamps images under different conditions. First, we used each stamp to process a two levels wavelet transform, and then we used the LL2: rough approximation of the image. Next we normalized each LL2 image to 60 x 60 pixels. Second, eight rule (using the relationships between the three primary colors of the RGB color model) were used to convert each stamp into a color string, and the color strings were stored in the database. Third, we compare the color strings of stamp images with color strings stored in the database using a color string comparison method. Finally, the ten most 10 similar stamps were shown. Our system maintains the advantages of both content-based (based on similarity searching), and text-based retrieval (rapid and mature) systems.

Key words: CBIR (Content-based Image Retrieval), Color String, Color String Comparison, Stamps Retrieval, Wavelet Transform.

壹、前言

由於數位圖像增加非常迅速，若這些數位圖像不能被有效地檢索，將不足以應付網際網路（全球資訊網）數位圖像之應用與搜尋需求。雖然在過去 30 年有各種用於儲存、瀏覽、搜尋和檢索的圖像系統已被陸續提出，但至今尚無一個令大家都滿意的系統。一個圖像檢索系統是對一個極龐大的資料庫進行瀏覽，搜尋和檢索圖像的計算機系統。不幸的是，大多數的成熟圖像檢索方法目前只有基於文本的方法。現在極龐大的圖像被存儲成數位格式，而通常只能通過網際網路（全球資訊網）上的「關鍵字」進行搜尋 [例如：谷歌 (Google) 或雅虎 (Yahoo) 的「關鍵字」搜尋]，如圖 1。而基於文本的搜尋方法對難以用文字描述之數位圖像就顯得捉襟見肘。

所謂的「基於內容的圖像檢索」[英語：Content-based Image Retrieval (CBIR)；或 Content-based Visual Information Retrieval (CVIR)]，是屬於圖像識別與人工智慧的一個熱門研究領域。基於內容的圖像檢索目的是在給定查詢圖像（可以是一張數位圖像或者是一張手畫的草圖再轉換為一張手畫的數位草圖）的前提下，再依據數位圖像內容信息與指定圖像查詢標準，在極龐大的數位圖像資料庫中搜尋並查找出符合數位圖像查詢條件的相應數位圖像。網際網路上傳統的搜尋引擎，包括 Google、Yahoo 以及 MSN 都已推出相應的數位圖像（片）搜尋功能，但是這些搜尋主要是基於圖像（片）的「文件名建立索引」來實現查詢功能。這種從查詢「文字、文件名」，最後到圖片查詢的機制並不是真正基於內容的圖像檢索。真正基於內容的圖像檢索指的是查詢條件本身就是一個「數位圖像」或「大概內容的數位圖像（如：一張手畫的草圖再轉換為一張手畫的數位草圖）」，它建立索引的方式是通過提取數位圖像內容特徵，然後通過計算比較這些數位圖像內容特徵和查詢條件之間的距離，來決定兩張數位圖像（片）的相似程度。

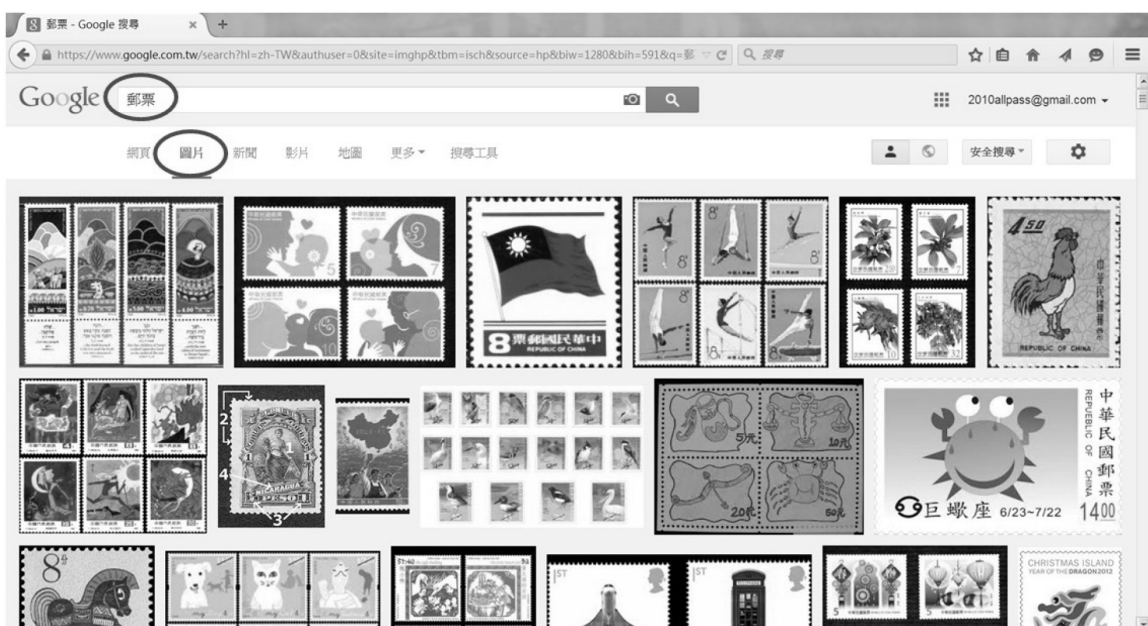


圖 1. 谷歌以關鍵字「郵票」進行搜尋與所得之結果。

谷歌在 2010 年 6 月在谷歌圖片搜尋加入數位圖像搜尋功能 [一個基於內容的數位圖像檢索系統 (CBIR)] 如圖 2。谷歌允許用戶上傳數位圖片檢索相似的圖片。它看起來很有趣，但其性能卻並不夠出色。谷歌的系統嘗試了解圖片的含義並以文字標識當「關鍵字」，然後再使用「關鍵字」以文本檢索方法檢索相似的圖片。而基於內容的檢索方法，應該是實際且直接使用數位圖像的內容，而不應該是像谷歌的系統嘗試了解圖片的含義再使用「關鍵字」以文本檢索方法搜尋，但谷歌圖片搜尋加入數位圖像搜尋功能卻是一個很好的開始。

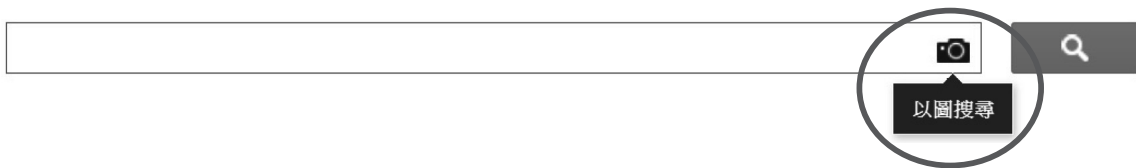
淘寶網幾乎網羅了中國全國的貨品買賣。然而因為淘寶網上的貨品種類繁多，若想要在眾多貨品中找到自己想買的東西並不容易。因而在 2015 年的手機版淘寶 App 的更新中，淘寶網便加入了全新的「拍照購」功能，讓買家可以更方便找到想買的貨品。「拍照購」容許買家透過手機拍攝物件，拍照之後淘寶網便會分析相片中的物件，再嘗試從淘寶網的資料庫中找出相片中相似的產品。為了增加搜尋的準確度，買家在拍攝後還可以在相片中圈出欲搜尋的物件的重點部份。雖然淘寶網「拍照購」功能聽起來十分好用，但可惜的是此功能目前仍未完善，不論是 iOS 還是 Android 版的手機淘寶網 App 中，目前只能搜尋女裝衣服及手提袋。可見假以時日，以圖找圖 (CBIR) 必能在電子商務中佔有一席之地。相信「真正基於內容的檢索方法」將在不久的將來扮演多媒體檢索系統非常重要的一環 Antani et al. (2002)。K. Konstantinidis et al. (2005) 提出了基於 $L^* a^* b^*$ 顏色空間的顏色直方圖的模糊鏈接方法，並提供了僅包含 10 個分類的直方圖。直方圖方法是基於從廣泛多樣化的圖像集中檢索相似圖像所獲得的性能來評估的。他們的方法（直方圖方法）被認為對圖像的各種變化，例如：光照變化、遮擋和雜訊比其他方法較不敏感。Silva et al. (2011) 討論了基於相關反饋的基於內容的圖像檢索 (CBIR) 系統的方法，該方法根據兩種主動學習模式（貪婪和計劃）進行。在貪婪方法中，系統將於每次重複尋找時找出與查詢最相關的圖像 (the most relevant images)。在計劃的方法中，在幾次重複尋找時期間找出最具信息性的圖像 (the most informative images)，並且在最後呈現最相關的圖像。他們提出了一種基於最優路徑森林分類 (OPF: optimum-path forest classification) 的貪婪方法分類器的計劃方法。然而在他們的方法都需要用戶的參與需要用戶提供回應。另外比較著名的系統還包含 Wang et al. (2001) 提出圖像庫的語義敏感整合匹配 (Semantics-sensitive Integrated Matching for Picture Libraries) 之圖像檢索系統，他們使用語義分類方法，並使用基於小波的方法來提取特徵，並基於圖像分割來整合區域匹配。與其他基於區域的檢索系統一樣，其特徵在於顏色，紋理，形狀和位置。系統將圖像分類為語義類別，例如：有紋理、無紋理、圖形、照片。用分類方法以縮小數據庫中的搜尋範圍來增強檢索效率。使用區域匹配方法來整合圖像之間的整體相似度。他們的方法 1. 減少了不準確的圖像分割的不利影響，2. 有助於闡明特定區域的語義，3. 為基於區域的圖像檢索系統啓用簡單的查詢界面。而他們於論文中亦誠實表示該系統有幾個限制：(1) 與其他 CBIR 系統一樣，SIMPLiCity 假設具有相似語義的圖像具有一些相似的特徵。然而這種假設可能並不成立。(2) 形狀匹配過程並不理想。當圖像被分割成許多區域時，應在合併匹配區域之後計算 IRM 距離。(3)

統計語義分類方法不能完美區分不同類別的圖像。此外，圖像可能同時落於幾個不同語義類。(4) 查詢界面不夠強健，系統不允許用戶自由地制定查詢。對於不同的用戶，查詢介面理應提供不同的功能介面。Ma et al. (1997) 介紹一種原型圖像檢索系統，它使用分割圖像區域中的顏色，紋理，形狀和空間位置信息來從圖像庫中搜尋和檢索相似區域。該系統的顯著特徵是它結合了一種強大的自動圖像分割演算法，該演算法允許基於標的物或區域的搜尋 (object- or region-based search)。當圖像包含多個複雜標的物時，圖像分割顯著提高了圖像檢索的質量。在存入圖像庫時將圖像分割成均勻區域，並計算表示這些區域中的每一個的圖像屬性。除了圖像分割之外，系統的其他重要組件還包括：有效的顏色表示，以及用於快速搜尋和檢索的顏色、紋理和形狀特徵的索引。該表示允許用戶撰寫有趣的查詢，例如“檢索包含具有標的物 A 的顏色的區域的所有圖像，標的物 B 的紋理，標的物 C 的形狀。其中各個標的物可以是屬於不同圖像的區域。但他們需要用戶提供的以下參數：1. 要使用的圖像特徵（灰色／顏色，紋理或兩者）。2. 定位所需邊界的首選比例。3. 區域合併算法的圖像區域的數量。然提供這些參數對使用者應是個大問題。Pentland et al. (1996) 是一組用於瀏覽和搜尋圖像和圖像序列的交互式工具。這些查詢工具與標準圖像數據庫中使用的查詢工具的不同之處在於，他們直接使用圖像內容而不是依賴於文本註釋。通過使用保持語義的圖像壓縮使得直接搜尋圖像內容的成為可能，他們將圖像縮小為一小組感知有效係數。他們討論了三種類型的 Photobook 描述：(1) 允許基於外觀 (appearance) 進行搜尋 (2) 使用二維形狀 (2-D shape) (3) 允許基於紋理屬性 (textural properties) 進行搜尋。這些圖像內容描述可以彼此組合併與基於文本的描述相結合，以提供複雜的瀏覽和搜尋功能。但他們的方法都需要用戶的參與需要用戶提供選擇用何種特徵進行搜尋。Smith et al. (1997) 描述了一個以視覺特徵進行搜尋的系統，用戶可繪製有顏色區域的空間排列之圖像來查詢。系統將依據此顏色區域的空間排列之圖像進行搜尋包含類似區域的最相似排列的圖像。首先系統會自動從圖像中提取並索引顯著顏色區域。再通過利用有效的顏色信息，區域大小以及絕對和相對空間位置的索引技術進行搜尋。然而他們的方法需要用戶自行繪製有顏色區域的空間排列之圖像，而不是直接輸入一張圖像。Qian et al. (2013) 則利用機器學習演算法的技術於設計電腦輔助路徑規劃系統 (CAP)，以幫助年輕醫生在外科手術路徑規劃同時維持最高的精確度。可見以圖找圖 (CBIR) 的應用也已經在臨床醫學界中佔有一席之地。

因為郵票有豐富的色彩，更是有實用價值的研究主題，所以在本研究中，我們選擇了圖像檢索的一個子集：「郵票檢索」。我們提出了一個「郵票」的檢索系統，本研究不像大多數 CBIR 的方法嘗試標識一個有字面意義的文字當「關鍵字」，然後使用此「關鍵字」去檢索相似的圖片。我們的方法是直接使用圖片顏色的空間位置去標識一個無字面意義的顏色字串當「關鍵字」，然後使用此「關鍵字」去檢索相似的圖片。我們的系統總覽如圖 3。本系統包含三個階段。一、我們將每張郵票先作二級小波變換，再將其 LL2：粗略近似的圖像取出，最後再將其 (LL2) 尺寸標準化為 60X60 像素，再用 8 條規則（利用 RGB 三原色之相對關係）將每張郵票轉換成顏色字串。二、我們利用字串比對進行比較這些

顏色字串。三、排名前 10 名相似的結果將被檢索出來。本文的其餘部分安排如下。在第 2 節介紹郵票的尺寸標準化與顏色字串的轉換方法。在第 3 節，將利用字串比對進行比較這些顏色字串進行說明。實驗結果和討論在第 4 節。最後，結論和未來的工作在第 5 節說明。

Google 圖片



(a) 首先，點選「以圖搜尋」。



(b) 先點選「上傳圖片」與「瀏覽」。再選取「要檢索的的圖片」後點選「開啓舊檔」。



(c) 谷歌「以圖找圖」檢索的結果（相似的圖片）。

圖 2. 谷歌的「以圖找圖」數位圖像檢索系統

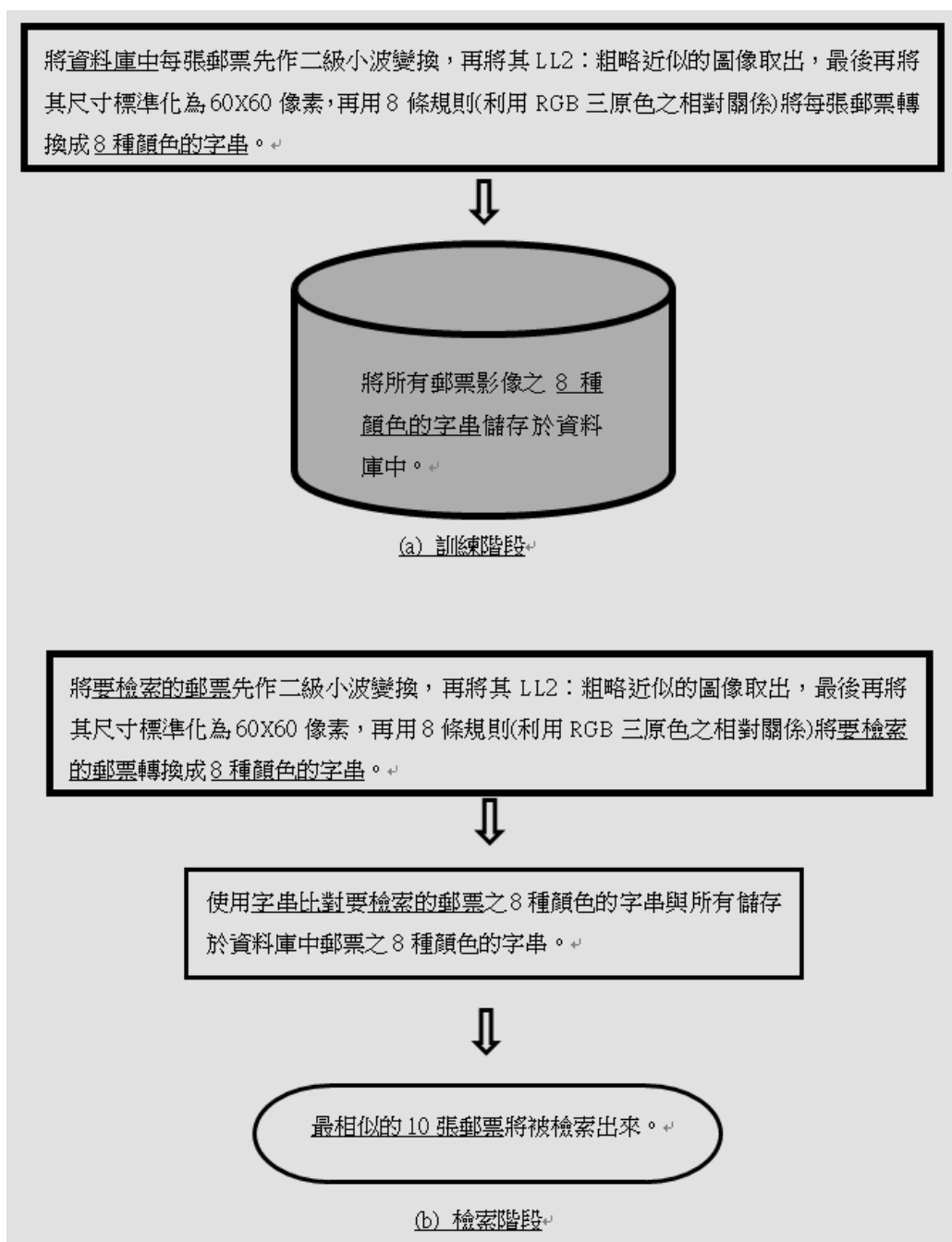


圖 3. 系統總覽

貳、Haar 小波變換

小波轉換最常見的應用是用於影像壓縮。和其他變換一樣，小波變換可以用於原始影像／圖像，然後將變換後的數據編碼，得到有效的壓縮。影像壓縮通常可分為三大步驟，分別是轉換 (Transform)、量化 (Quantization) 和編碼 (Coding)。其中轉換這個步驟是將原始資料轉換成另一種表示法，轉換的目的在於除去訊號取樣的相關性，也就是去除取樣間

的不需要部份（LL2 抽取出的特徵正是我們所需要的低頻部份，我們不需要的高頻部份則被去除掉，且只保留低頻部份也可去除雜訊。）

Haar 小波的母小波函數 $\Psi(t)$ 可以描述為

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ -1 & 1/2 \leq t < 1 \\ 0 & \text{Otherwise} \end{cases}$$

其縮放函數 $\Phi(t)$ 可以描述為

$$\phi(t) = \begin{cases} 1 & 0 \leq t < 1 \\ 0 & \text{Otherwise} \end{cases}$$

Haar 小波可用於去除雜訊，邊緣檢測和壓縮。Haar 小波是最易理解的小波。試著考慮兩個數 a 和 b （一個序列的相鄰樣本），則 a 和 b 有一定的相關性。我們可以用一個簡單的線性變換來代替 a 和 b 的平均值 s 與差值 d 。

$$s = (a + b) / 2$$

$$d = b - a$$

目標是如果 a 和 b 高度相關則可減少 d 所需的位元數，這個計算可以利用逆轉運算來恢復 a 和 b ：

$$a = s - d/2$$

$$b = s + d/2$$

這就是 Haar 小波起始的想法。考慮 $2n$ 個採樣值的信號 S_n 。對於每一對值 (pair)，我們應用這個平均值和差值變換。“對”的數量是 $2n-1$ ，並且這些“對”可產生許多平均值和差值。我們可以將平均值視為原始信號的粗略分辨的版本，並將差值視為更高分辨率的細節。如果原始信號高度相關，則信號的粗略分辨的版本將與原始信號非常接近，且可以高效地表示細節。隨後，我們可以將相同的平均值和差值變換應用於較粗略的信號，將其分解為更粗略的信號和更多細節。憑藉一些奇特的技巧，這種轉換可以在沒有附加緩衝空間的情況下被適用和逆轉，且僅需要 $O(N)$ 時間。對於一維小波變換，我們計算了一個輸入行的所有小波級別，並將結果存儲在具有相同維度的第二行中。對於二維小波變換，我

們對圖像的所有列執行一級變換。具體而言，對於每一列，通過低通和高通濾波器來計算其近似值和細節。結果是兩個矩陣分別是 L 和 H ，每個矩陣都是一半的原始圖像尺寸。隨後，矩陣 (L 和 H) 再次通過低通和高通濾波。結果是被稱為 4 個子帶的四個子矩陣 LL , LH , HL 和 HH ： LL 是全部低通子帶（粗略的近似圖像）， LH 是水平子帶， HL 是垂直子帶，並且 HH 是對角子帶。為了實現二級小波變換，只處理 LL 子矩陣，如圖 4 所示。下一步將只使用低通子帶 ($LL2$ ：粗略近似的圖像)。

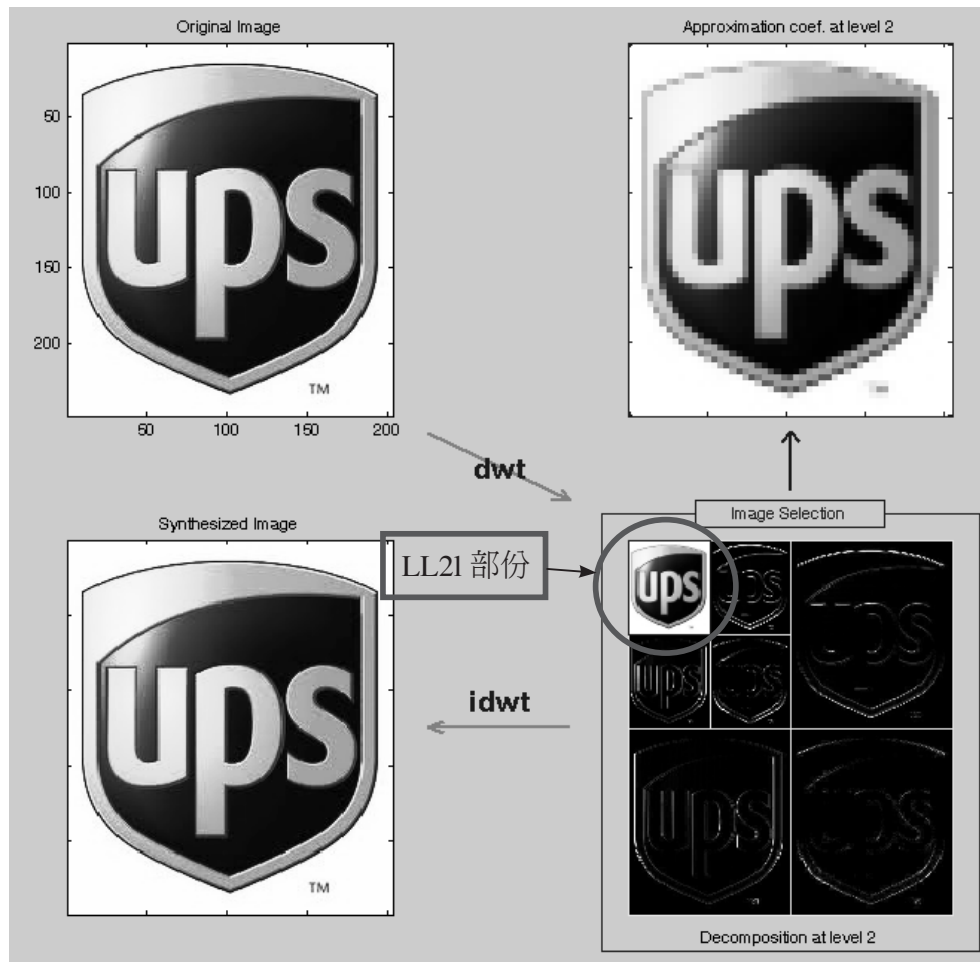


圖 4. $LL2$ 部分 (2 階的 Haar 小波變換) 將用於基於顏色的字串編碼

參、尺寸標準化與顏色字串編碼

由於郵票可能有不同大小，因此，所有的郵票皆先作二級小波變換，再將其 $LL2$ ：粗略近似的圖像取出，最後再將其標準化為標準尺寸 (60×60 像素)，我們使用雙立方插補 (bicubic interpolation) 方法標準化尺寸 Rafael et al. (2007)，它能比佔主導地位的雙線性插補 (Bilinear interpolation) 演算法保留更好的細節質量。雖然標準化尺寸將失去郵票的一些內容信息，但可以得到對不同的大小、模糊和雜訊問題的包容效果，並可加快速度。

三原色光模式 (RGB color model)，又稱 RGB 顏色模型或紅綠藍顏色模型，是一種加色模型，將紅 (Red)、綠 (Green)、藍 (Blue) 三原色的色光以不同的比例相加，以產生多種多樣的色光。因為 RGB 顏色空間是一個三維空間向量，並且每一個像素 $P(i)$ ，是由一個有序三元組 [紅，綠和藍色的坐標， $R(i)$ 、 $G(i)$ 、 $B(i)$]，其分別表示紅色、綠色和藍色光的顏色的強度。我們知道 R 、 G 和 B 的值在改變的光照條件和不同的顏色飽和度的情況下會完全不同。然而， $R(i)$ 、 $G(i)$ 和 $B(i)$ 之間的相對關係卻是變異極小的。此外，RGB 方式只有 6 種排列關係 (即 RGB、RBG、GRB、GBR、BGR 和 BRG)。因此，我們可以利用 $R(i)$ 、 $G(i)$ 、及 $B(i)$ 之間的相對關係 (如： $R(i) > G(i) > B(i)$)，則轉換該像素為 R (第一個系列的「紅色」)。再加入「白色」和「黑色」(總計 8 種字符 (W, K, R, S, G, H, B, C))。我們利用 RGB 三原色之相對關係訂定 8 條規則。例如：有一張郵票影像原本是 600X600 像素。我們將其尺寸標準化為 60×60 像素，換言之，每一個字母是代表原本影像 100 個像素的平均值。因此並不會因位置稍有差異而失敗。[例如：圖 11(a)]

之後再用 8 條規則 (利用 RGB 三原色之相對關係) 將每張郵票轉換成顏色字串以克服於不同光線情況下、不同的色彩飽和度情況下、蓋了不同位置私章的、蓋了不同位置郵戳的相似郵票皆可以被找出來 (這些都是郵票常見的情形)。因為我們是利用 RGB 三原色之相對關係。所以我們可以克服於不同光線情況下、不同的色彩飽和度情況下 (因為我們知道 R 、 G 和 B 的值在改變的光照條件和不同的顏色飽和度的情況下會完全不同。然而， $R(i)$ 、 $G(i)$ 和 $B(i)$ 之間的相對關係卻是變異極小的。) 蓋了不同位置私章的、蓋了不同位置郵戳的、Y 型戳記與像素位移的情況下的相似郵票皆可以被找出來。我們並非試圖辨認出影像中有何物件，而是要找出顏色分布相似的影像。我們所提之方法當屬於相同的照片不足 10 張時，自然會找出顏色分布相似的影像。因為辨認出影像有何物件，似乎只是將影像做了電腦索引，不就等於回到以「關鍵字」找圖了嗎？由圖 11 可證我們所提之方法並不只是能找到的基本上屬於相同的照片。

茲將 8 條規則列示如下：

- (1) if a pixel $235 \leq r(i) \leq 255, 235 \leq g(i) \leq 255, \text{ and } 235 \leq b(i) \leq 255$, then assigns the pixel as 'W'; (White color)
- (2) if a pixel $0 \leq r(i) \leq 20, 0 \leq g(i) \leq 20, \text{ and } 0 \leq b(i) \leq 20$, then assigns the pixel as 'K'; (Black color)
- (3) if a pixel $r(i) > g(i) > b(i)$, then assigns the pixel as 'R'; (the first series of "Red" colors)
- (4) if a pixel $r(i) \geq b(i) > g(i)$, then assigns the pixel as 'S'; (the second series of "Red" colors)
- (5) if a pixel $g(i) > r(i) \geq b(i)$, then assigns the pixel as 'G'; (the first series of "Green" colors)
- (6) if a pixel $g(i) \geq b(i) > r(i)$, then assigns the pixel as 'H'; (the second series of "Green" colors)

(7) if a pixel $b(i) > r(i) \geq g(i)$, then assigns the pixel as 'B'; (the first series of "Blue" colors)

(8) if a pixel $b(i) \geq g(i) > r(i)$, then assigns the pixel as 'C'; (the second series of "Blue" colors)

因為 8 種字符 (W,K,R,S,G,H,B,C) 3600 個就代表 8^{3600} ($=2^{10800}$) 的排列方式已是一個天文數目，應該足夠用來判別各種不同的郵票。我們做了一個實例「以 20×20 個像素的 24 位的顏色色板」來說明轉換結果，我們用「紫色 W」代表白色，以便讀者能夠清晰可見 W，如圖 6。我們可以觀察到的 R 和 S 代表兩個系列的紅色的 G 和 H 代表兩個系列的綠色，在 B 和 C 代表兩個系列的藍色中，W 代表白色和 K 代表黑色。

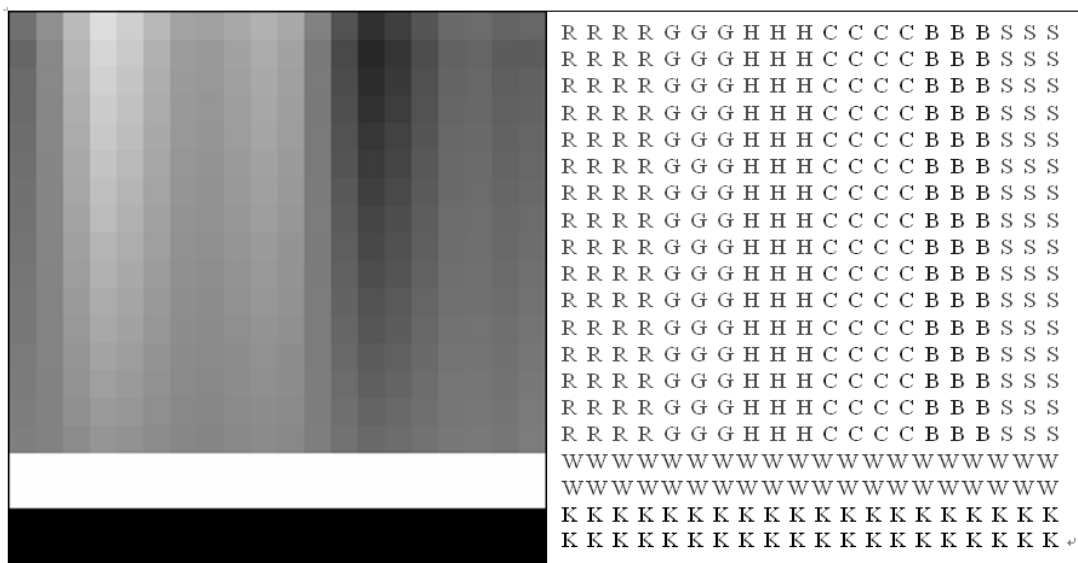


圖 6. 以 20×20 個像素的 24 位的顏色色板說明轉換結果（我們用「紫色 W」代表白色，以便讀者能夠清晰可見 W）。

肆、字串比對

在數學中，字串比對（也稱為相似性的比較）是兩個字串之間的相似性比較。例如：“Research”和“Researcher”的字串可以被認為是相似的。例如：如果兩個郵票圖像是完全相同的，其總比對值就是 3600。如果總比對值為 3600，則距離是 0。愈相似的郵票圖像應具有愈高的總比對值和更小的距離。

因此，欲查詢（檢索）的郵票之顏色字串將和在資料庫中的所有郵票之顏色字串經由字串比對做相似度之計算。由此相似度計算出的「總比對值」則用以確定排名，排名前 10 名相似的結果將被檢索出來。

字串比對之概念如下：

如果「欲查詢郵票之顏色字串」和「在資料庫中的郵票之顏色字串」在相同的位置都

包含相同的顏色字串（例如：兩者都是 R），則總比對值 = 總比對值 + 1。

字串比對之演算法描述如下：

輸入：「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」。

輸出：每張「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」比較後產生之總比對值。

開始

將「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」輸入。

步驟 1) 如果「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」在相同位置的字皆是 R，則總比對值 = 總比對值 + 1。

步驟 2) 如果「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」在相同位置的字皆是 S，則總比對值 = 總比對值 + 1。

步驟 3) 如果「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」在相同位置的字皆是 G，則總比對值 = 總比對值 + 1。

步驟 4) 如果「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」在相同位置的字皆是 H，則總比對值 = 總比對值 + 1。

步驟 5) 如果「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」在相同位置的字皆是 B，則總比對值 = 總比對值 + 1。

步驟 6) 如果「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」在相同位置的字皆是 C，則總比對值 = 總比對值 + 1。

步驟 7) 如果「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」在相同位置的字皆是 W，則總比對值 = 總比對值 + 1。

步驟 8) 如果「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」在相同位置的字皆是 K，則總比對值 = 總比對值 + 1。

步驟 9) 計算每張「欲查詢郵票之顏色字串」和「資料庫中的所有郵票之顏色字串」比較後產生之總比對值。

結束

伍、實驗結果和討論

Y. Mussarat et al. (2013) 表示：「基於內容的圖像檢索的比較需要相同的計算環境和資料庫，因為使用不同的環境和資料庫比較的兩種方法將會導致無意義的結果。」一般而言這是合理的，但在與谷歌的系統比較的情況下是不正確的。由於谷歌的系統資料庫理應比其他資料庫包含更多的圖像（更多的候選圖像理應返回更好的類似圖像，因此，谷歌的系統應該經常返回更好的類似圖像才合理）。否則，我們將永遠失去與谷歌的系統比較之極好機會（谷歌的系統最公開，不能欺詐），因為我們幾乎不可能得到與谷歌的系統相同的計算環境和資料庫。況且實驗的資料庫郵票皆來自網際網路（全球資訊網），總計

有 25,800 張郵票。我們的資料庫包含不同光線情況下、不同的色彩飽和度情況下、蓋了不同位置私章的、蓋了不同位置郵戳的。我們的資料庫部份郵票圖像如圖 7。這些與 K. Konstantinidis 的方法（直方圖方法）被認為對圖像的各種變化，例如：光照變化、遮擋和雜訊比其他方法較不敏感，優點較為相似。因此，我們選擇與谷歌和 K. Konstantinidis 的「以圖找圖」檢索的結果做比較，結果如圖 8~11。我們使用一台 P4 的電腦 CPU 2.5 GHz，記憶體 4G，我們的方法訓練時間是 58.198 秒 /25,800 張郵票和檢索時間是 1.253 秒 / 每次查詢，兩者皆是非常迅速的。當然訓練時間和檢索時間會因為資料庫的擴大與每張圖像的尺寸大小變大而增加。而 K. Konstantinidis 的方法訓練時間是 71.652 秒 /25,800 張郵票和檢索時間是 2.048 秒 / 每次查詢，不論是訓練時間和檢索時間，我們的方法皆優於 K. Konstantinidis 的方法。與 K. Konstantinidis 的比較結果詳如表 1。

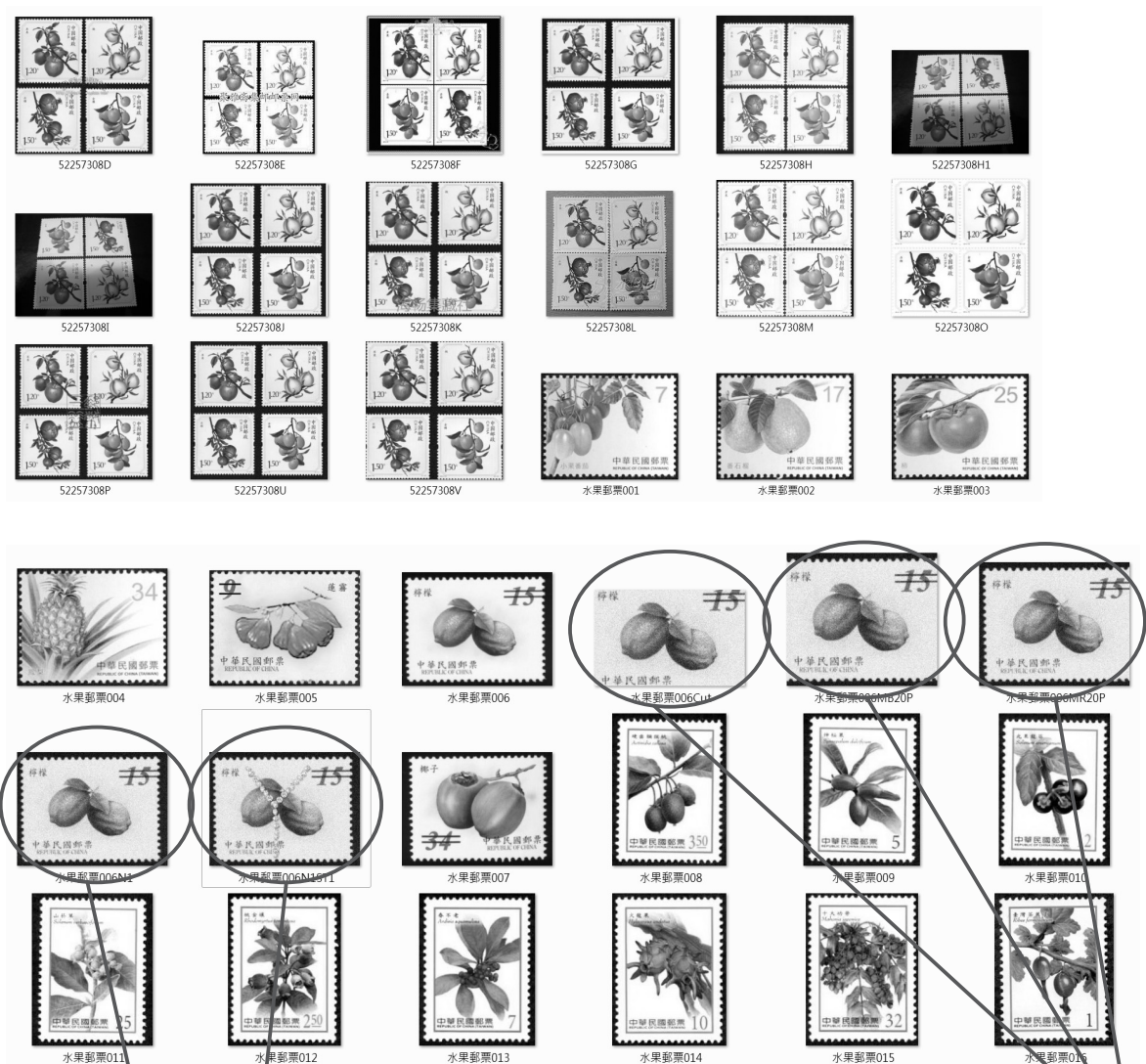
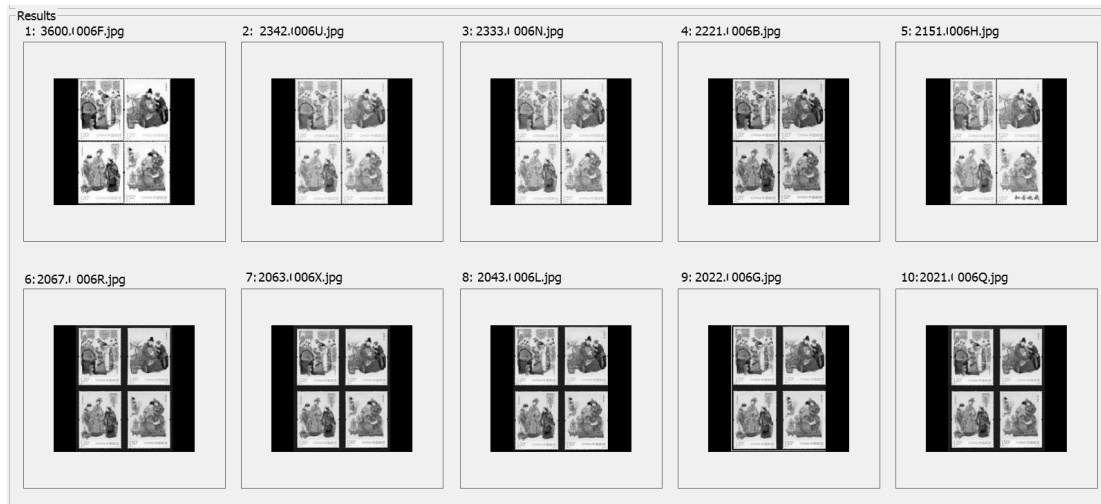


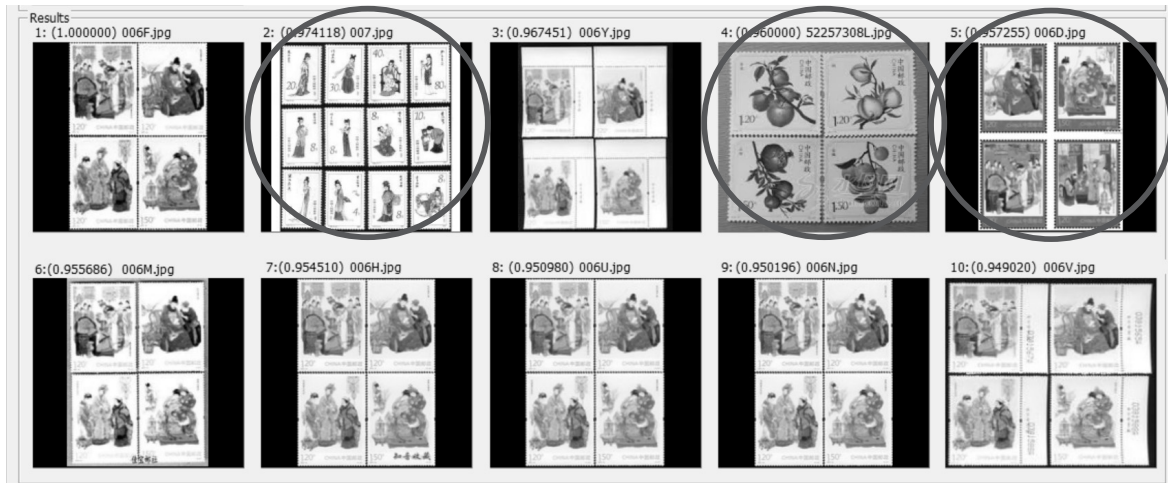
圖 7. 我們的資料庫包含不同光線情況下、不同的色彩飽和度情況下、蓋了不同位置私章、郵戳、像素位移、有雜訊與 Y 型戳記（如橢圓框出的郵票）的情況下。



(a) 我們提出的方法檢索的結果。第一個郵票圖像也是查詢圖像。檢索的結果圖像以降序的總比對值方式從左至右並從上到下。

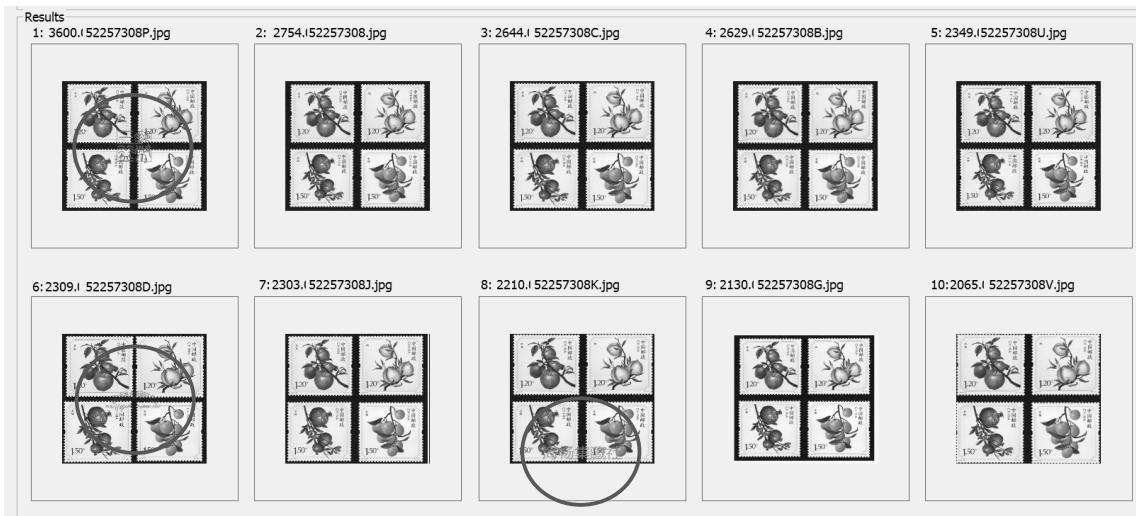


(b) 谷歌「以圖找圖」檢索系統檢索的結果。



(c) K. Konstantinidis 「以圖找圖」 檢索系統檢索的結果，紅色橢圓框出檢索應視為錯誤的郵票。

圖 8. 於不同光線情況下，我們提出的方法與谷歌和 K. Konstantinidis 「以圖找圖」 檢索系統比較的結果。



(a) 我們提出的方法檢索的結果。橢圓框出蓋了私章的位置。第一個郵票圖像也是查詢圖像。檢索的結果圖像以降序的總比對值方式從左至右並從上到下。



約有 113 項結果 (搜尋時間：0.53 秒)



圖片大小：
620 × 605

尋找這個圖片的其他尺寸版本：
所有大小 - 中型

這個圖片最有可能的推測結果：**切手 果物**

特殊切手「野菜とくだものシリーズ 第3集」の発行 - 日本郵便
www.post.japanpost.jp/kitte_hagaki/.../h270223_t.html ▾ 翻譯這個網頁
 種類, 52円郵便切手 (シール式), 82円郵便切手 (シール式) ... から本格的にハウス栽培が始まり、国産の完熟品が出回るようになったことで、注目される果物となりました。

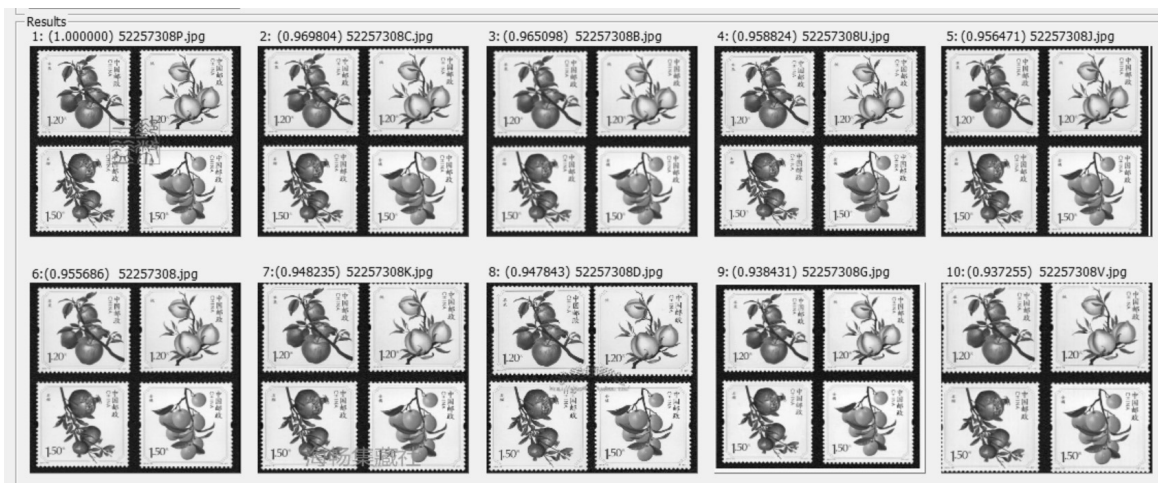
看起來相似的圖片

檢舉圖片



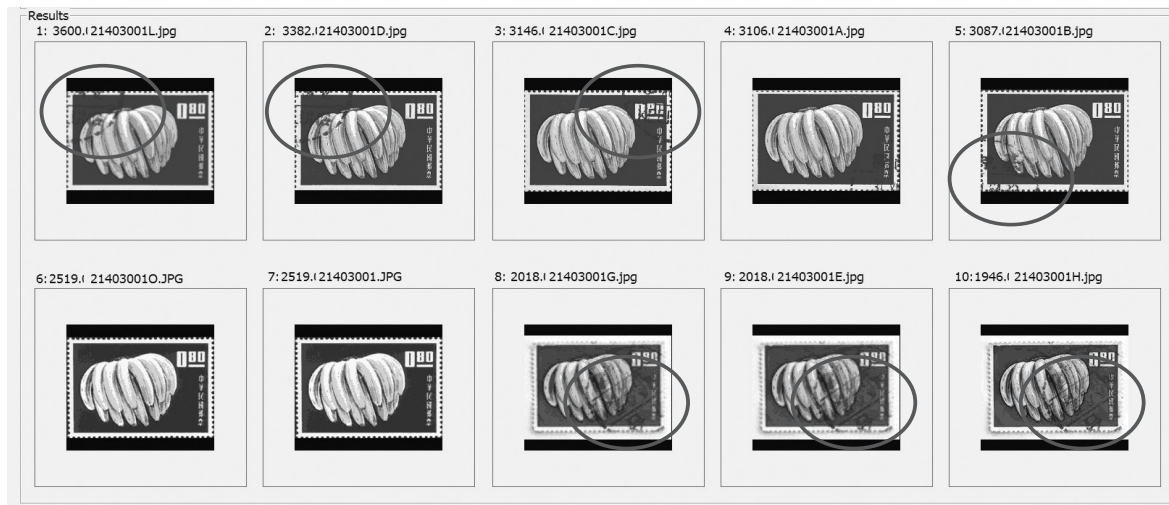
含有相同圖片的網頁

(b) 谷歌「以圖找圖」檢索系統檢索的結果。



(c) K. Konstantinidis 「以圖找圖」檢索系統檢索的結果。

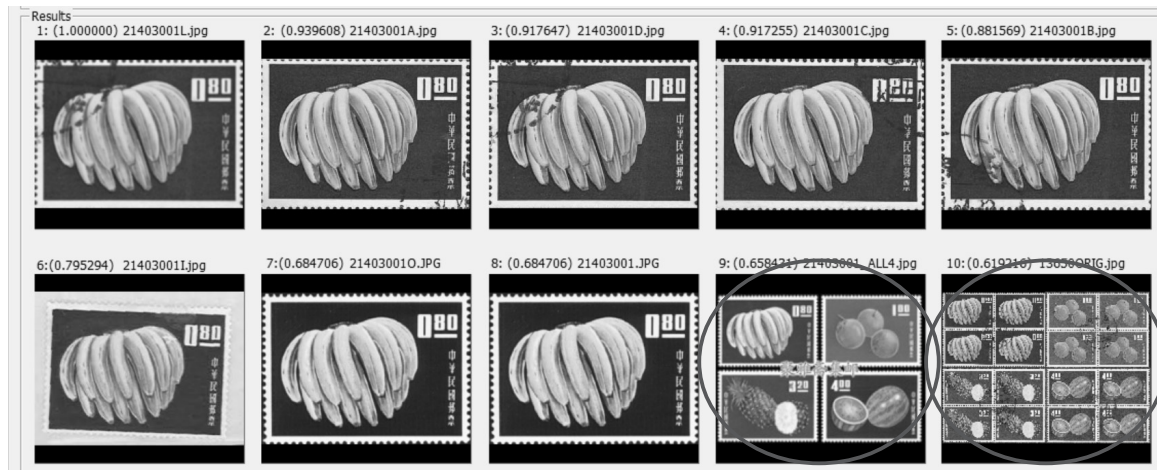
圖 9. 於蓋了私章與不同的色彩飽和度情況下，我們提出的方法與谷歌和 K. Konstantinidis 「以圖找圖」檢索系統比較的結果。



(a) 我們提出的方法檢索的結果。紅色橢圓框出郵戳的位置。第一個郵票圖像也是查詢圖像。檢索的結果圖像以降序的總比對值方式從左至右並從上到下。



(b) 谷歌「以圖找圖」檢索系統檢索的結果。



(c) K. Konstantinidis 「以圖找圖」檢索系統檢索的結果，紅色橢圓框出檢索應視為錯誤的郵票。

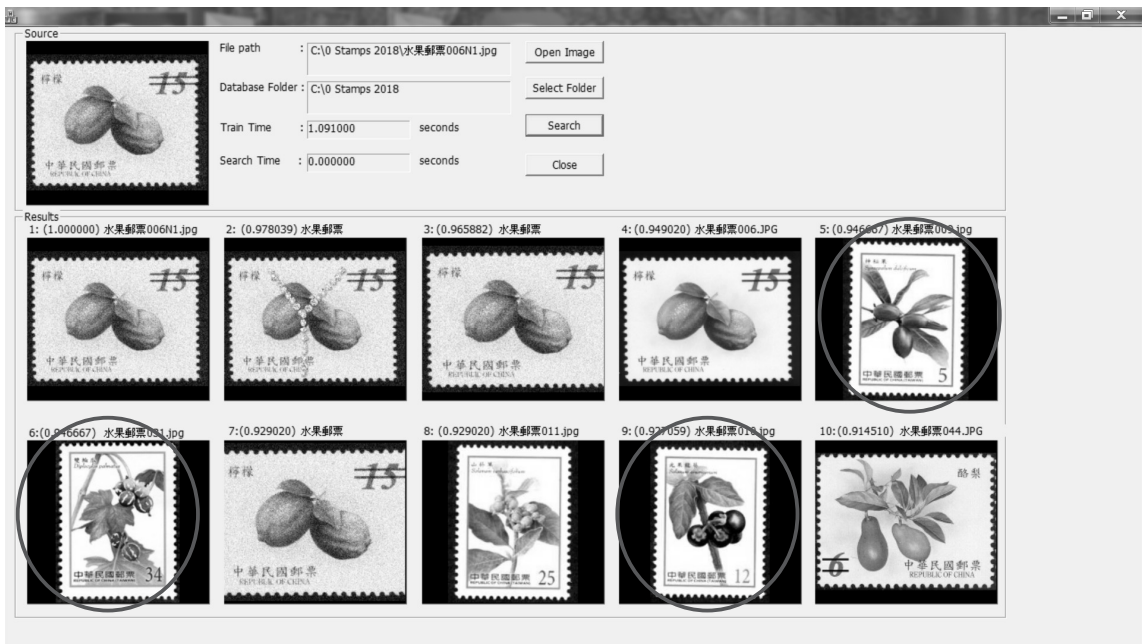
圖 10. 於蓋了郵戳與不同的色彩飽和度情況下，我們提出的方法與谷歌和 K. Konstantinidis 「以圖找圖」檢索系統比較的結果。



(a) 我們提出的方法檢索的結果。於有雜訊、Y型戳記與像素位移（如橢圓框出的郵票）的情況下。第一個郵票圖像也是查詢圖像。檢索的結果圖像以降序的總比對值方式從左至右並從上到下。



(b) 谷歌「以圖找圖」檢索系統檢索的結果。



(c) K. Konstantinidis 「以圖找圖」檢索系統檢索的結果，紅色橢圓框出檢索應視為錯誤的郵票。

圖 11. 於加入雜訊、Y 型戳記與像素位移的情況下，我們提出的方法與谷歌和 K. Konstantinidis 「以圖找圖」檢索系統比較的結果。

表 1. 與 K. Konstantinidis 的比較結果。不論是訓練時間和檢索時間，我們的方法皆優於 K. Konstantinidis 的方法。

執行時間 使用方法	訓練時間 (秒) / 所有郵票	搜尋時間 (秒) / 十個相似郵票
使用我們的方法	58.198/25,800	1.253
K. K. 的方法	71.652/25,800	2.048

陸、結論和未來的工作

因為我們將每張郵票轉換成顏色字串，該郵票檢索系統即成爲一個類似基於文本的檢索系統。因為顏色字串的每個字母皆是包含一系列顏色（例如，白色，黑色，紅色，綠色或藍色系列）。因此，我們的系統可以克服不同的亮度與不同的色彩飽和度的問題，更容許「欲查詢之數位圖像」與「其查詢結果」具有一些差異性（基於相似性的檢索）。因字串比對的計算是極迅速，所以我們的系統執行上是非常迅速的（非常迅速和成熟）。換句話說，我們的系統保持了基於內容的數位圖像檢索系統（基於相似性的檢索）和基於文本的數位圖像檢索系統（非常迅速和成熟）的雙重優點。在未來，我們將改善我們的方法以適用於不同的數位圖像領域，我們希望基於內容的數位圖像搜尋將來能更爲正確、有效與簡單。

參考文獻

英文部分

- Antani, S., R. Kasturi, and R. Jain, (2002) A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video. *Pattern Recognition*, 35, (4), 945-965.
- Konstantinidis, K., A. and I. Andreadis Gasteratos, (2005) Image retrieval based on fuzzy color histogram processing, *Optics Communications*, 248(4 – 6), 375 – 386.
- Ma, W. Y. and B. Manjunath, (1997) Netra: a toolbox for navigating large image databases, *Proceedings of the IEEE International Conference on Image Processing*, Santa Barbara, CA, USA.
- Yasmin Mussarat, Sharif Muhammad, Mohsin Sajjad and Irum Isma, “Content Based Image Retrieval Using Combined Features of Shape, Color and Relevance Feedback,” *KSII Transactions on Internet and Information Systems* 2013; 7(12): 3149-3165.

- Pentland, A., R. W. Picard, and S. Scaroff, (1996) Photobook: content-based manipulation for image databases, HYPERLINK "<https://link.springer.com/journal/11263>" \o "International Journal of Computer Vision" International Journal of Computer Vision, 18(3), 233 - 254.
- Qian, Yu, Rui Hui, Xiaohong Gao, (2013) 3D CBIR with sparse coding for image-guided neurosurgery, Signal Processing, 93(6), 1673 - 1683.
- Rafael, C. Gonzalez, and E. Woods Richard, (2007) Digital Image Processing, (3rd Edition), Prentice Hall. USA.
- Silva, André Tavares da, Falcão, Alexandre Xavier, Magalhães, Léo Pini, (2011) Active learning paradigms for CBIR systems based on optimum-path forest classification. Pattern Recognition, 44 (12), 2971 - 2978.
- Smith, J. R., and S. F. Chang, (1997) VisualSEEK: a fully automated content-based image querysystem, ACM International Conference on Multimedia, Boston, Massachusetts, USA.
- Wang, James Z., Member, IEEE, Li, Jia Member, IEEE, and Wiederhold, Gio Fellow, IEEE (2001) SIMPLIcity: semantics-sensitive integrated matching for picture libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(9), 947 – 963.