

基於主成份分析與模組化半徑基底函數類神經網路於影像語意內容分析

張傳育, 李綺芳

國立雲林科技大學資訊工程研究所

chuanyu@yuntech.edu.tw

摘要

由於數位影像、資料庫和網路技術的快速發展，對於大型資料庫而言，以內容(content)為基礎的影像檢索(image retrieval)已成為一個重要的研究題目。另一方面，隨著數位相機的普及，數位影像的拍攝與取得變得極為容易，現行數位像機中幾乎都是以日期加上編號的方式來命名，造成日後影像檢索與管理上的困難。因此，如何有效的管理大量的數位影像，也是一項值得研究的課題。因此，我們將提出一個有效的影像語意內容分析技術，利用模組化半徑基底函數類神經網路(modular RBF neural networks, MRBF)結合低階的影像特徵分析影像的內容，計算影像中各種高階語意項目的成分，進而達到影像管理的目的。此外，為解決傳統半徑基底函數(RBF)類神經網路的中心點初始化問題，加速 MRBF 的訓練及提升 MRBF 的分類效能，我們將結合自我組織映射(SOM)和向量學習量化(LVQ)網路進行中心點初始值的設定。同時，由於影像特徵的維度眾多，我們也將進一步使用主成份分析(PCA)技術分析影像特徵，從中選取有效及有鑑別度的影像特徵，降低 MRBF 的輸入特徵維度。實驗結果顯示，本論文提出的 MRBF 方法對於語意為基礎的照片分析有高的準確度，而且分析的結果比較接近人類視覺的判斷。

關鍵詞：影像內容分析、模組化半徑基底函數類神經網路。

1. 前言

隨著電腦儲存設備容量與日俱增、網路頻寬的增加，再加上數位相機、掃描器等影像設備的普及，影像取得與傳送更為方便，為我們帶來了大量的影像資料。面對這些龐大的影像資料，影像的檢索與管理已成為一個重要的課題。目前已經有學者提出以內容為基礎的影像檢索系統，例如：QBIC[1]和 Photobook[2]。這些系統嘗試利用一些低階的特徵，例如：影像的顏色分佈、影像的物件形狀及影像的紋理結構等當成影像特徵作為影像檢索時的索引，以便在檢索時進行影像相似度的比對。然而，這類以圖找圖的影像檢索系統無法以高階的語意進行檢索，因為使用者通常不會思考影像的低階特徵，也就是說，使用者思考的是高階語意(semantic)(例如：給我一張日出影像)。然而，上述的影像檢索系統對於此種需求並沒有辦法提供一個很好的解決方式。因此，如何使用影像的低階特徵，將影像分類至適當的高階語意項目中，是一個

相當值得深入探究的題目。一個良好的影像檢索系統，會在比對階段排除不適當的影像，藉此提升系統的檢索性能。在影像分類技術中，已有使用語意標籤的方法[3]被提出使用，希望可以解決上述的問題。然而，對於一張影像而言，影像中可能包含相當多的物件，不同的物件有不同的語意類別，導致影像分類系統很難辨識這麼大量的物件語意。以區域為基礎的檢索系統[4]嘗試解決高階語意與低階特徵的問題，採用物件的影像特徵作為代表，進行影像切割(segmentation)，將一張影像切割成許多小區域，假如這個切割方法是理想的話，每一個小區域代表一個物件區域，此種系統跟人類視覺系統是類似的。

近年來，使用以類神經網路的方法已經被提出並用來解決分類的問題。在眾多架構中，RBF 網路是一個最受歡迎的架構，因為它具有好的學習和近似能力。然而，RBF 的建構重點在於 RBF 的中心點位置，以及參數調整的學習演算法。RBF 對於中心點的初始值相當敏感，為了取得較佳的中心點，它必須找出重要的特徵，提供給 RBF 進一步的分類。另外，RBF 網路也是一種耗時的網路架構，尤其在輸入特徵具有較高維度的時候，此種特性更是明顯。因此，在本論文中，我們採用模組化半徑基底函數類神經網路(modular RBF neural networks, MRBF)網路[5]，並結合 SOM 和 LVQ 網路架構，用以解決 RBF 網路選擇適當的中心點問題和提升分類的的能力。

在影像資料中，鄰近像素的相關性頗高，因此，我們所擷取出的低階影像特徵可能會存在一些多餘資訊，而且特徵越多不見得分類效果越好，不良的特徵數值會影響有效特徵的數值，可能導致影像分類誤判，也會增加網路的學習時間。因此，我們採用主成份分析(Principal Component Analysis, PCA)[6-7]對所有特徵進行分析，挑選具鑑別度的特徵，以降低特徵維度。

在我們提出的影像語意內容分析系統架構中，包含兩個階段：訓練階段和測試階段，圖一展示影像語意內容分析系統流程。在訓練階段，首先對影像進行分割，並且計算每個區域的各種影像特徵與高階語意的對應，同時，使用 PCA 分析影像特徵，從中選取有效及有鑑別度的影像特徵，做為 MRBF 的訓練樣本。MRBF 結合自我組織映射(SOM)和向量學習量化(LVQ)網路進行初始化中心點的設定，加速 MRBF 訓練的速度及提升 MRBF 的分類能力。在測試階段，類似於訓練階段，先對

測試影像進行分割，計算每個區域的影像特徵，使用 PCA 分析影像特徵，將這些提供給 MRBF 網路進行分類。此外，我們也將使用機率的方式描述影像中物件的高階語意，也就是說，我們將分析照片中包含哪些高階語意的物件，計算每一個高階語意物件占整張影像的機率。

在本論文的其他章節安排如下：在第二章中，簡介影像切割、特徵擷取、PCA 原理及結合 SOM 網路和 LVQ 網路產生適當的 RBF 中心點初始值。我們使用的 MRBF 網路將在第三章中做介紹。第四章介紹實驗結果以及討論。最後，第五章為本論文的結論。



圖 1 影像語意內容分析系統流程

2. 相關理論

2.1 影像切割

由於一張影像可能包含多個物件，影像切割產生的區域資訊對影像分類具有相當大的影響。我們使用 J-Images segmentation (JSEG) [8] 切割方法對影像進行切割。

雖然 JSEG 切割演算法實現一個合理的切割結果，但是此結果對於人類視覺感知是不夠直覺的。因此，經過 JSEG 切割之後的區域，我們透過使用者手動指定區域的語意類別合併成各個物件區域。採用使用者指定語意類別區域的優點在於得到更理想的物件區域。此外，JSEG 演算法可以提供比較好的切割結果供使用者進一步的語意類別指定。因此，JSEG 演算法可以大量減輕使用者在語意類別指定之負擔。圖二展示經由使用者指定區域的語意類別後的範例。

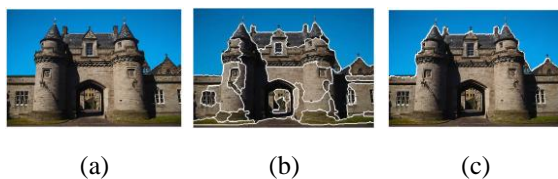


圖 2 經由使用者指定區域的語意類別後的範例 (a) 原始影像 (b) JSEG 切割後的結果 (c) 使用者指定後的結果。

2.2 特徵擷取

我們透過使用者手動指定區域的語意類別後合併成各個物件區域後，對每個物件區域擷取顏色、紋理及形狀特徵。由於顏色、紋理、及顏色特徵與人類視覺感知有高度相關。因此，我們分析這些特徵對影像內容分類的能力。

物件區域的顏色特徵採用 HSV 彩色空間重新量化成 64 個種類，因此，顏色特徵我們使用 64 個彩色特徵作為物件的顏色特徵。使用 5 個灰階統

計參數如方程式(1-5)，配合 4 個不同的角度(0、45、90 和 135 度)，來描述一個物件區域的紋理特徵。由於邊緣方向直方圖 [9] 具有位移和尺度不變的特性，可用來表示物件的形狀特徵。因此，每張影像可表示成一個由顏色特徵、紋理特徵及形狀特徵所組合而成的特徵向量 (feature vector)，維度為 1×108 。圖三展示物件區域的顏色及形狀特徵，建築物的主體與天空的顏色特徵有明顯的差異。

$$\text{Contrast} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i-j)^2 P_{\delta}(i, j) \quad (1)$$

$$\text{Energy} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{\delta}(i, j)^2 \quad (2)$$

$$\text{Entropy} = - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{\delta}(i, j) \log(P_{\delta}(i, j)) \quad (3)$$

$$\text{Homogeneity} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{P_{\delta}(i, j)}{1+(i-j)^2} \quad (4)$$

$$\text{Max probability} = \text{Max}\{P_{\delta}(i, j)\} \quad (5)$$

其中 $P_{\delta}(i, j)$ 是在條件為 $\delta(r, \theta)$ 下的 co-occurrence matrix，而 r 代表像素距離， θ 是兩個像素 (i, j) 的角度。

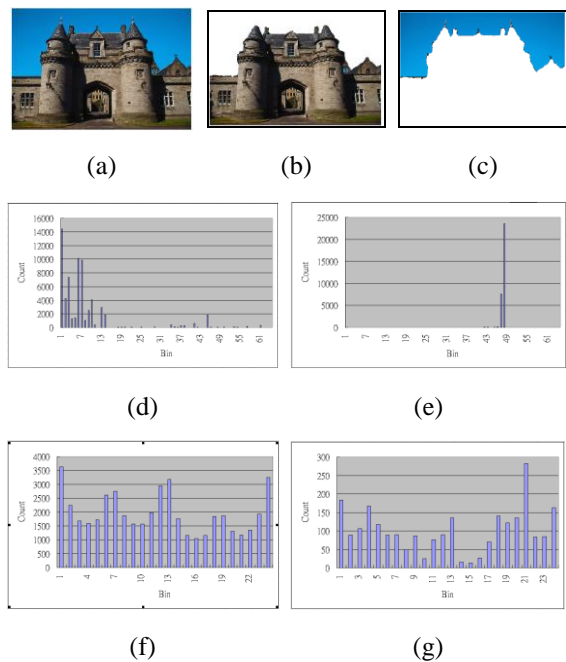


圖 3 物件區域的顏色及形狀特徵 (a) 原始影像 (b) 切割成建築物的區域 (c) 切割成天空的區域 (d) 建築物的顏色特徵 (e) 天空的顏色特徵 (f) 建築物的形狀特徵 (g) 天空的形狀特徵。

2.3 主成份分析

主成份分析 (PCA) 是一種多變量分析的方法，被廣泛地應用在影像處理、資料維度降低等。PCA 是在一個多維空間中找出一條軸線，使得所有的多維向量能夠在這條軸線上產生一個投影點，並使相鄰的投影點彼此擁有最大的變異數。降低維度的過程中，PCA 是將原始資料透過線性組

合產生一組較原始資料少的新變數，這個新變數就被稱為主成份。

假設每張影像擷取的影像特徵有 N 個，形成原始特徵向量 $X = \{x_1, x_2, \dots, x_N\}$ ，首先計算原始特徵向量的共變異矩陣 C_x

$$C_x = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x})^T \quad (6)$$

其中， \bar{x} 為 X 的平均向量。

接著計算共變異矩陣的特徵值 (Eigenvalue)，並依其數值大小排序為： $\lambda_1, \lambda_2, \dots, \lambda_N$ ，並求出相對應的單位特徵向量 (Eigenvector)： E_1, E_2, \dots, E_N 。由於重要的資訊集中在較大特徵值所對應之特徵向量中，利用方程式(7)可以得到第 k 個主成份 y_k ：

$$y_k = E_k^T X \quad k=1,2,3,\dots,N \quad (7)$$

主成份的個數是透過主成份貢獻度 P 來決定，其數學式如下：

$$P = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_M}{\lambda_1 + \lambda_2 + \dots + \lambda_N} \quad \text{其中 } N > M \quad (8)$$

原始的影像特徵參數經過 PCA 轉換後，得到新的特徵參數 $\{y_1, y_2, \dots, y_M\}$ 。這組特徵參數除了維度降低之外，其全域散佈矩陣也是最大，如此便可以讓我們用較少維度的特徵來表示每張影像，並當作 MRBF 網路的輸入。

2.4 結合 SOM 網路和 LVQ 網路產生適當的 RBF 中心點初始值

由於 RBF 網路對於中心點的初始化相當敏感，為了找出適當的中心點，我們必須找出重要的特徵，供進一步的網路分類。SOM 網路被證實具有自動分類的能力[10]，為了解決中心點選擇的問題，我們使用 SOM 網路來選擇適當的中心點。為了取得 RBF 網路的中心點初始值，對於一張影像，我們計算影像的多種特徵，組合成一個特徵向量。因此，當每一類有 N 張影像時，我們將會有 N 個特徵向量，經過 PCA 轉換後得到新的 N 個向量提供給 SOM 網路做訓練，用以產生適當的中心點初始值 $v = \{v_1, v_2, \dots, v_M\}$ 。

經由 SOM 網路取得適當中心點之後，相似的中心點可能會結合在一起，為了解決此問題，我們採用 LVQ 演算法做為相似的細微調整。令 $\{v_j\}_{j=1}^M$ 代表使用 SOM 網路取得的中心點，而 $\{x_i\}_{i=1}^N$ 代表要訓練的影像特徵。對於一個輸入影像特徵 $x_i(n)$ ，最相似的第 k 個中心點可以使用如下取得：

$$k = \arg \min_j \{ \|x_i - v_j\| \} \quad (9)$$

令 ℓ_{v_k} 代表經由 SOM 網路取得的中心點，而 ℓ_{x_i} 代表 x_i 的輸入影像特徵，因此，細微的中心點調整可以使用方程式(10)、(11)：

若 $\ell_{v_k} = \ell_{x_i}$ ，則

$$v_k(n+1) = v_k(n) + \eta(n)[x_i - v_k(n)] \quad (10)$$

若 $\ell_{v_k} \neq \ell_{x_i}$ ，則

$$v_k(n+1) = v_k(n) - \eta(n)[x_i - v_k(n)] \quad (11)$$

其中 $\eta(n)$ 是學習速率並隨著時間遞減， $0 \leq \eta(n) \leq 1$ 。

3. 模組化半徑基底函數類神經網路 (MRBF)

在我們使用的 MRBF 中，結合 SOM 網路和 LVQ 網路產生適當的 MRBF 中心點初始值，而且採用模組化的概念，並應用在隱藏層中，提供輸入特徵維度的轉換，因此，每一個模組都是各自做線性加總，也就是說，每一個輸出神經元並沒有互相連結，圖四展示我們使用的網路架構圖，此架構具有 M 個 RBF 和 M 個輸出神經元。這樣的設計，讓每個 RBF 模組獨立運作，可加速網路的訓練及提升個別 RBF 模組的分類能力。我們使用的 MRBF 網路，其參數模型定義如下：

$$f_j(x) = \sum_{i=1}^N w_{ji} \phi_{ji}(\|x - c_{ji}\|) \quad (12)$$

其中 $x \in R^n$ 是一個輸入向量， ϕ_{ji} 是 MRBF 網路的基底函數， w_{ji} 是權重向量， $c_{ji} = (c_{ji1}, c_{ji2}, \dots, c_{jin})^T$ 是第 j 個模組的第 i 個中心節點，符號 $\| \cdot \|$ 代表 Euclidean norm。

MRBF 採用高斯函數當作隱藏層的核心函數，因此，方程式(12)可以被修改成：

$$f_j(x) = \sum_{i=1}^N w_{ji} \phi_{ji}(\|x - c_{ji}\|) = \sum_{i=1}^N w_{ji} \exp\left(-\frac{\|x - c_{ji}\|^2}{\sigma_{ji}^2}\right) \quad (13)$$

其中 σ_{ji} 是第 j 個模組的第 i 個高斯函數的頻寬，如下所示：

$$\sigma_{ji} = \frac{d_{\max,j}}{\sqrt{k}} \quad (14)$$

其中 $d_{\max,j}$ 是在第 j 個模組中的最大 Euclidean 距離，而 k 是第 j 個模組的中心點個數。

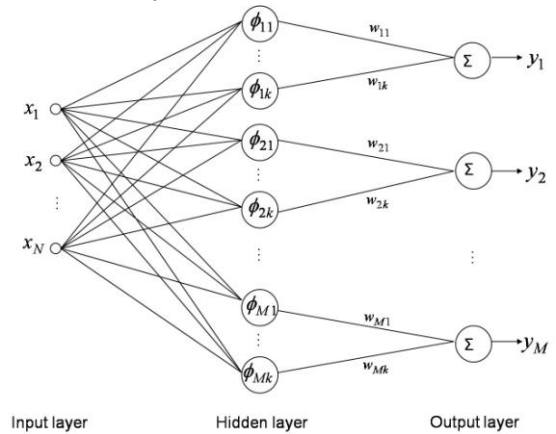


圖 4 MRBF 網路架構

MRBF 網路的錯誤成本函數定義成如下：

$$J(n) = \frac{1}{2} \sum_{j=1}^M \left[y_j^{desired}(n) - \sum_{i=1}^N w_{ji}(n) \exp(-\|\mathbf{x}(n) - \mathbf{c}_{ji}(n)\|^2 / \sigma_{ji}^2(n)) \right]^2 \quad (15)$$

其中 $y_j^{desired}$ 代表第 j 個輸出神經元的希望值。

MRBF 網路的更新方程式定義如下：

$$w_{ji}(n+1) = w_{ji}(n) - \mu_w \frac{\partial}{\partial w_{ji}} J(n) \Big|_{w_{ji}=w_{ji}(n)} \quad (16)$$

$$= w_{ji}(n) + \mu_w e(n) \psi(n)$$

$$\mathbf{c}_{ji}(n+1) = \mathbf{c}_{ji}(n) - \mu_c \frac{\partial}{\partial \mathbf{c}_{ji}} J(n) \Big|_{\mathbf{c}_{ji}=\mathbf{c}_{ji}(n)} \quad (17)$$

$$= \mathbf{c}_{ji}(n) + \mu_c \frac{e(n) w_{ji}(n)}{\sigma_{ji}^2(n)}$$

$$\cdot \exp(-\|\mathbf{x}(n) - \mathbf{c}_{ji}(n)\|^2 / \sigma_{ji}^2(n)) [\mathbf{x}(n) - \mathbf{c}_{ji}(n)]$$

$$\sigma_{ji}(n+1) = \sigma_{ji}(n) - \mu_\sigma \frac{\partial}{\partial \sigma_{ji}} J(n) \Big|_{\sigma_{ji}=\sigma_{ji}(n)} \quad (18)$$

$$= \sigma_{ji}(n) + \mu_\sigma \frac{e(n) w_{ji}(n)}{\sigma_{ji}^3(n)}$$

$$\cdot \exp(-\|\mathbf{x}(n) - \mathbf{c}_{ji}(n)\|^2 / \sigma_{ji}^2(n)) \|\mathbf{x}(n) - \mathbf{c}_{ji}(n)\|^2$$

其中

$$\psi(n) = [\phi_{j1}(\|\mathbf{x}(n) - \mathbf{c}_{j1}(n)\|), \dots, \phi_{jN}(\|\mathbf{x}(n) - \mathbf{c}_{jN}(n)\|)]^T \quad (19)$$

$$\text{和 } e(n) = y_j^{desired}(n) - y_j(n) \quad (20)$$

此外， μ_w 、 μ_c 和 μ_σ 是適當的學習參數。

綜合以上所述，我們可以將 MRBF 網路的訓練程序歸納如下：

- 步驟 1、對輸入的影像特徵進行 PCA 轉換，獲得新的特徵向量 $\{y_1, y_2, \dots, y_M\}$ 。
- 步驟 2、使用 SOM 網路選擇適當的中心點 $v = \{v_1, v_2, \dots, v_M\}$ 。
- 步驟 3、使用 LVQ 演算法進行中心點的微調。
- 步驟 4、根據方程式(14)計算 MRBF 網路的高斯頻寬。
- 步驟 5、使用隨機的方法初始 MRBF 網路的權重，並將步驟 3 的中心點指定 MRBF 的初始中心點。
- 步驟 6、根據方程式(13)，將第 j 類的輸入特徵向量提供給網路做訓練。
- 步驟 7、使用方程式(16)、(17)和(18)更新網路參數。
- 步驟 8、假如網路收斂則停止訓練，若否，則至步驟 6。

4. 實驗結果及討論

在照片分析實驗中，我們選擇 Corel 影像資料庫中的 132 張影像作為網路訓練影像，並且定義 8 種語意的類別[3]：建築物、雲、天空或水、草地、樹木、花、山和日出，8 種語意的類別影像數量分別有 33 張。8 種語義類別與區塊影像的顏色對照如表 1.所示。

實驗中，每個物件特徵表示成一個由 64 個顏色特徵、20 個紋理特徵及 24 形狀特徵所組合而成的特徵向量(feature vector)，維度為 1×108 。然後，

將每種特徵個別透過 PCA 運算，取出主成份貢獻度 $P \geq 0.95$ 的主成份。最後我們得到每個物件用新的顏色特徵 14 個、新的形狀特徵 11 個及新的紋理特徵 3 個表示成一個組合而成的新特徵向量(feature vector)，維度為 1×28 ，當作 MRBF 網路的輸入。









此外，在 MRBF 網路中，我們使用 8 個 RBF 網路模組，作為 8 個語意類別的分析，每一個 RBF 網路模組都具有 25 個隱藏層神經元。照片分析考慮區域歸屬的語意類別以及面積的大小，因此，對於第 j 類語意類別，其分析機率定義如下：

$$P_j = \sum_{i=0}^k \frac{W_i^j}{S} \quad (21)$$

其中 S 代表影像的大小， k 為影像分割後的區域數目， W_i^j 是第 j 個語意類別的第 i 個區域的區域大小。

圖 5 到圖 8 展示以語意為基礎的照片分析結果。實驗結果中，圖 5(b)內的區塊可被完全正確的歸成日出，因此，日出的類別機率為 100%。圖 6 影像包含四種物件，其中，天空/水有 16.82%的機率、雲有 26.33%的機率、建築物和草地各有 19.31%的機率和 37.52%的機率，雖然有少部分的建築物區塊被誤認為是天空和雲，但分類的結果頗符合人眼視覺感受。基本上，圖 7 影像中包含四種語意類別，但影像右側的低矮建築被誤判為樹木，這是因為影像中右邊的建築物在影像中的較小，而且它的顏色類似於樹幹的顏色。圖 8(a)是一張於樹林中拍攝的影像，雖然仍有些小區域被誤判為日出及花，但是分析的結果樹木仍然占了影像的 87.02%，符合人眼視覺感官的認知。由上述的實驗分析結果可知，即使是影像的分類結果仍存在些微的錯誤，但這些錯誤都是肉眼難以分辨的區域，因此，分析的結果都相當接近人類視覺的判斷。

表 1. 使用在區塊影像的顏色對照表

語意類別	顏色區塊
建築物	
雲	
天空/水	
草地	
樹木	
花	
山	
日出	

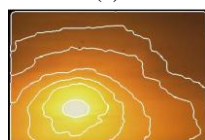
5. 結論

我們提出的影像語意內容分析系統，計算影像中各種高階語意項目的成分，達到影像管理的目

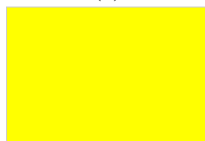
的。我們採用以物件為基礎的特徵擷取方式，而不是整張影像的特徵，並進一步利用 PCA 分析影像特徵，選取有效及具有鑑別度的影像特徵，降低影像特徵維度。因此，可以達到更高的分類準確度及節省 MRBF 的學習時間。此外，為了解決 RBF 的中心點初始化問題，加速 MRBF 的訓練及提升 MRBF 的分類效能，我們結合 SOM 和 LVQ 網路進行中心初始值的設定。MRBF 的設計讓每個 RBF 模組獨立運作，可提升個別 RBF 模組的分類能力，而且模組化的設計容易新增、刪除、及抽換任一 RBF 子網路，而不會影響其他已訓練完成之子網路模組。從實驗結果顯示，分析的結果確實相當接近人類視覺的判斷。



(a)



(b)



(c)

影像包含的類別機率

建築物	0 %
雲	0 %
花	0 %
草地	0 %
山	0 %
天空/水	0 %
日出	100 %
樹木	0 %

(d)

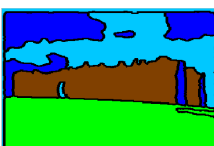
圖 5 語意為基礎的照片分析結果 (a)原始影像(b)切割後的圖片(c)分類後所對應的區塊影像(d)分析結果。



(a)



(b)



(c)

影像包含的類別機率

建築物	19.31%
雲	26.33%
花	0 %
草地	37.52%
山	0 %
天空/水	16.82%
日出	0 %
樹木	0 %

(d)

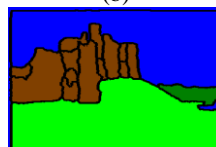
圖 6 語意為基礎的照片分析結果 (a)原始影像(b)切割後的圖片(c)分類後所對應的區塊影像(d)分析結果。



(a)



(b)



(c)

影像包含的類別機率

建築物	17.6 %
雲	0 %
花	0 %
草地	40.42 %
山	0 %
天空/水	39.59 %
日出	0 %
樹木	2.38 %

(d)

圖 7 語意為基礎的照片分析結果 (a)原始影像(b)切割後的圖片(c)分類後所對應的區塊影像(d)分析結果。



(a)



(b)



(c)

影像包含的類別機率

建築物	3.57 %
雲	2.83 %
花	0.58 %
草地	1.61 %
山	0.69 %
天空/水	3.37 %
日出	0.32 %
樹木	87.02 %

(d)

圖 8 語意為基礎的照片分析結果 (a)原始影像(b)切割後的圖片(c)分類後所對應的區塊影像(d)分析結果。

誌謝

本論文的研究經費由國科會提供，NSC95-2221-E224-059。

參考文獻

- [1] C. Faloutsos, R. Barder, M. Flickner, J. Hafner, W. Niblack, D. Petkovic and W. Equitz, "Efficient and effective querying by image content," *J. Intell. Inform. Systems*, vol. 3, pp. 231-262, 1994.
- [2] A. Pentland, R.W. Picard and S. Sclaroff, "Photobook: content-based manipulation of image database," *SPIE*

Storage and retrieval for image and video database II,
vol. 2185, pp. 34-47, 1994.

- [3] Yixin Chen, James Z. Wang and Robert Krovetz, "CLUE: Cluster-based retrieval of image by unsupervised learning," *IEEE Trans. Image Processing*, vol. 14, pp. 1187-1201, August 2005.
- [4] Y. Liu, D. Zhang, G. Lu and W.Y. Ma, "Region-based image retrieval with high-level semantic color names," *IEEE Proc. Multimedia Modeling Conference*, pp. 180-187, 2005.
- [5] Chuan-Yu Chang and Shih-Yu Fu, "Image Classification using a Module RBF Neural Network," *Proc. of IEEE International Conference on Innovative Computing, Information and Control*, vol. 2, pp.270-273, 2006.
- [6] J. Latin, D.Carroll and P.E.Green, *Analyzing Multivariate Data*, Duxbury, 2003.
- [7] A.C.Rencher, *Multivariate Statistical Inference and Applications*, John Wiley and Sons,1998.
- [8] Y. Deng, B.S. Manjunath and H. Shin, "Unsupervised segmentation of color-texture regions in image and video," *IEEE Tran. Pattern Analysis and Machine Intelligence*, vol. 23, no. 8, pp. 800-810, Aug. 2001.
- [9] A.K. Jain and A. Vailaya, "Image retrieval using color and shape," *Pattern Recognit.*, vol. 29, pp. 1233-1244, 1996.
- [10] Paisarn Munesseawang and Ling Guan, "A neural network approach for learning image similarity in adaptive CBIR," *IEEE Proc. Workshop on Multimedia Signal Processing*, pp. 257-262, 2001