

植基粒子群最佳化演算法於語意影像分類

李朱慧
朝陽科技大學
chlee@cyut.edu.tw

許志宏
資訊管理系
s9514639@cyut.edu.tw

摘要

資訊科技的進步使得數位電子產品充斥於日常生活中，尤其以數位相機更為普及，致使影像資料庫的需求大幅成長。於是如何有效率的在龐大的影像資料庫中進行影像分類，便成為一項重要的研究議題。本研究提出一個語意導向的影像分類系統，將影像依照使用者需求區分成數個影像類別，利用粒子群最佳化演算法找出各類物件最佳的觀測點，透過最佳觀測點可以用來推測每類物件佔影像中之比例，並依此進行影像分類。本研究期望能在龐大的影像資料庫中，增進影像的語意分類效率，以滿足使用者的需求。

關鍵詞：粒子群最佳化演算法、語意、影像分類、特徵抽取

1. 前言

由於資訊科技的快速發展，數位產品大量生產，致使社會走向數位化的生活，其中以數位相機更為普及，促使數位影像的快速成長，導致數位影像資料庫日益龐大。如何在龐大的影像資料庫中，有效率的進行影像分類，便成為一項重要的研究議題。在影像資料庫中，分類影像的方法相當多，其中包括使用顏色(color)、關鍵字(keywords)、敘述句(sentences)、圖例(examples)、形狀(shape)、紋理(texture)，甚至是空間結構及位移(spatial structure and motion)。

透過上述的分類影像方法，衍生出許多分類技術，例如：利用區塊特徵向量進行分類的 KNN (K-Nearest Neighbor)[7];利用特徵距離計算任兩 blob 之間距離的決策樹(Decision Tree)[3];找出影像之二元分類的 SVM (Support Vector Machine)[1];藉由將特徵加權後分類的分層級分析法(Analytical Hierarchy Process ,AHP)[2]。

然而良好的影像分類，主要希望能夠自動抽取影像特徵，且能與人類的認知更為相近。

只由影像的標題或鄰近的文章內容獲取某些相關資訊，是可以滿足某些應用上的需求，不過要滿足多數實際應用上複雜的需求，答案是令人懷疑的。一張影像所傳達的內容多半遠超過外觀所呈現的，不同使用者會賦予不同的意義，低階的影像技術要加以涵蓋這些意涵，其技術困難度相當高，也因此在此實際的應用需求上，語意影像的分類技術在未來仍是不可或缺的[5]。

本研究是以一般的消費者影像為對象，消費者影像不同於一般專業影像或定義域專屬的影像，因此顯得更值得注目[6]，我們將分析一張影像中具有特定物件所佔的比例，並依此作為影像語意分類的依據，這些特定物件是依使用者需求定義，研究中利用粒子群最佳化演算法(Particle Swam Optimization algorithm, PSO)找出特定物件最佳的觀測點，從最佳的觀測點可以推測特定物件佔影像中之比例。

本研究的架構如下：在第二節部分，探討粒子群最佳化演算法的發展過程。在第三節中，說明本研究所提出的如何用粒子群最佳化演算法找出特定物件最佳的觀測點。最後在第四節中針對本研究做總結。

2. 相關文獻

粒子群最佳化演算法(Particle Swam Optimization algorithm, PSO)是 Kennedy 和 Eberhart 於 1995 年提出[4]，其主要的概念是源自於群體行為理論，觀察生物群體行動時，能透過個體之間訊息的傳遞方式，使整個群體朝預期較適合的方向或目標移動，此演算法為模仿生物行為以尋求群體最大利益的方法。初期的粒子群最佳化演算法為：先以均勻分佈(uniform distribution)的方式，隨機產生初始粒子群，而每個粒子會參考本身的最佳經驗 P_{best} 和群體的最佳經驗 G_{best} ，以進行移動方向的修正。每個粒子經過不斷地修正後，預期皆會逐漸向最佳解移動。其粒子演算的計算公式如下所示：

$$V_{id} = V_{id} + C_1 \cdot \text{rand}_1 \cdot (P_{best} - X_{id}) + C_2 \cdot \text{rand}_2 \cdot (G_{best} - X_{id}) \quad (1)$$

$$X_{id} = X_{id} + V_{id} \quad (2)$$

其中， i ：第 i 個粒子。

d ：第 d 個維度。

V_{id} ：第 i 個粒子在第 d 個維度的位移量。

P_{best} ：第 d 個維度的個體最佳解。

G_{best} ：第 d 個維度的群體最佳解。

X_{id} ：第 d 個維度的粒子。

C_1, C_2 ：學習常數。

$\text{rand}_1, \text{rand}_2$ ：0 到 1 之間的隨機變數。

透過公式(1)可以計算出每一個粒子下一步朝向最佳解移動的位移量 V_{id} ，把粒子的原始位置 X_{id} 加上新的位移量 V_{id} ，即為該粒子下一個新位置。將這種具有生物朝向目標移動的特性運用於影像分類，可以依照使用者需求，將影像分類成數個群聚，並可得到每個分群的特徵值。

從圖 1 可以發現每個粒子具有自行朝向目標移動的特性，粒子移動前會先參考自己之前的最佳移動經驗 P_{best} 和群體的最佳移動經驗 G_{best} ，透過這兩種移動經驗，能使粒子朝向群體的目標解移動，最後產生最接近目標解的值。

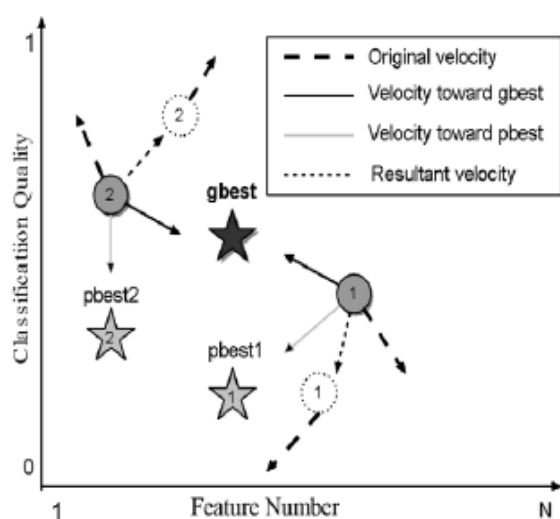


圖 1、粒子移動圖[8]

圖 2 表示粒子群最佳化的演算過程，在演算之前會隨機產生數個粒子，再將這些隨機產生的粒子進行演算，經過粒子群最佳化演算後，最後會取出距離目標最接近的粒子做為最佳解。

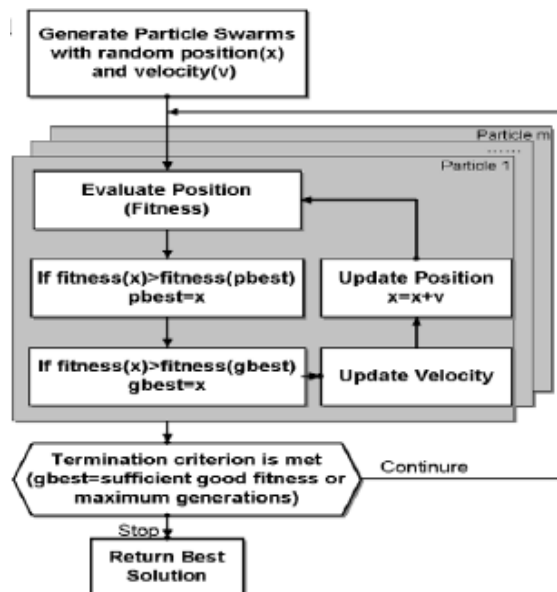


圖 2、粒子群最佳化演算流程圖[8]

3. 語意影像分類

在消費者影像中，常見的物件有人類、天體、植物、建築物、交通工具等，我們希望藉由這些常見物件在影像中出現的面積比例來作為影像分類的依據，例如影像中人類佔了 20%，天體佔了 30% 等，此種分類方式貼近高階的語意，以提供後續階段高階語意搜尋[5]。

不同的物件在顏色、紋路或是形狀上會有其代表性的特徵，一般而言，特徵值可以轉換為多維度向量來表示，如果影像中有出現某物件時，則可由影像的特徵值觀察出來，而且如果某物件在影像中出現的面積越大，則該類對特徵值的影響越大，因此我們希望透過粒子群最佳化演算法，找出某物件的特徵值的最具代表性的表示法，我們稱之為最佳觀測點，以用來預測某物件佔影像中之比例。

3.1 物件比例算法

假設資料庫中有 i 類重要的物件，分別為 $X_1, X_2, X_3, \dots, X_i$ ，而第 i 類物件的最佳觀測點為 $(X_{i1}, X_{i2}, X_{i3}, \dots, X_{in})$ 。不同的特徵在不同物件的敏感度是有差異的，例如： $X_{ij} \rightarrow 0$ 即代表第 j 個特徵值在第 i 個物件是越不重要，而 X_{ij} 越大代表第 j 個特徵值越能代表具有第 i 類物件。影像中某物件所佔比例的計算如下：

$$\text{Ratio}(P, X_i) = \frac{\sum_{j=1}^n P_j X_{ij}}{\sqrt{\sum_{j=1}^n P_j^2} \cdot \sqrt{\sum_{j=1}^n X_{ij}^2}} \quad (3)$$

P 代表任意影像的特徵值， X_i 是第 i 類物件的最佳觀測值，藉由上面公式的計算，可以推測該影像具有某類物件的比例，距離愈小即代表該影像具有第 i 類物件的比例越大。

3.2 粒子群最佳化演算法

我們將利用 PSO 將最佳觀測點找出，針對 i 個物件將執行 i 次 PSO 演算，以找出 i 個最佳觀測值，且每次 PSO 的輸入是具有該類物件的代表性影像。首先將每張影像轉換成 n 維向量 $\text{image}(v_1, v_2, v_3, \dots, v_n)$ ，以進行粒子群最佳化演算法。

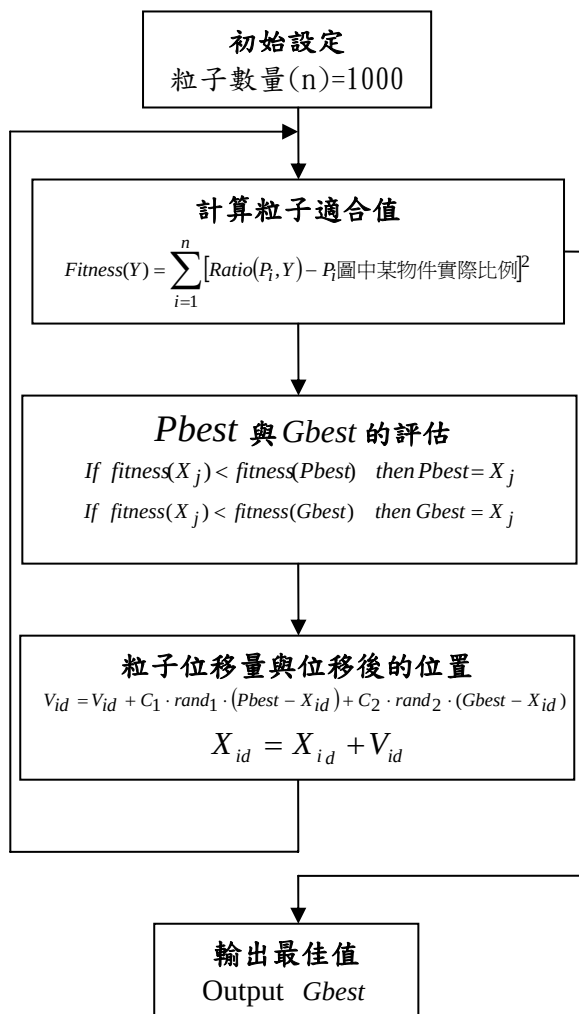


圖 3 粒子群最佳化演算流程圖

本研究之粒子群最佳化演算流程圖如圖

3 所示。在進行粒子群最佳化演算法之前，必需預先設定演算參數，預設的 $Gbest$ 為 $(\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ ，而 $Pbest$ 即為輸入的影像的特徵值，接著計算每一粒子的適合值，該粒子適合值代表目前粒子和 $Gbest$ 之間的距離，粒子適合值愈小，表示該粒子愈接近理想的 $Gbest$ ，也就足以成為該特徵分群的觀測值。粒子適合值的計算：

$$\text{Fitness}(Y) = \sum_{i=1}^n [\text{Ratio}(P_i, Y) - P\text{圖中某物件實際比例}]^2 \quad (4)$$

透過求得的粒子適合值，可以評估該粒子是否足以取代目前的 $Pbest$ 和 $Gbest$ 值，如果粒子適合值小於目前的 $Pbest$ ，則將該粒子適合值取代 $Pbest$ ，成目前最佳的 $Pbest$ ，同樣的如果該粒子適合值小於目前的 $Gbest$ ，則取代 $Gbest$ 成為目前最佳的 $Gbest$ 。假設粒子適合值皆小於或等於目前的 $Pbest$ 和 $Gbest$ ，則不進行取代動作。

$Pbest$ 與 $Gbest$ 的取代：

$$\text{If } \text{fitness}(X_j) < \text{fitness}(Pbest) \text{ then } Pbest = X_j \quad (5)$$

$$\text{If } \text{fitness}(X_j) < \text{fitness}(Gbest) \text{ then } Gbest = X_j \quad (6)$$

接著利用目前的 $Pbest$ 和 $Gbest$ 值，計算每一個粒子下一步該往 $Gbest$ 移動的位移量，使得全部的粒子皆朝最佳的 $Gbest$ 值移動，最後可以得到最佳解。

粒子位移量的計算：

$$V_{id} = V_{id} + C_1 \cdot \text{rand}_1 \cdot (Pbest - X_{id}) + C_2 \cdot \text{rand}_2 \cdot (Gbest - X_{id}) \quad (7)$$

再將該粒子目前的位置，加上計算出的粒子位移量，即為該粒子新的目標位置。

粒子新目標位置：

$$X_{id} = X_{id} + V_{id} \quad (8)$$

當粒子群最佳化演算結束後，可以輸出最終的 $Gbest$ ，該 $Gbest$ 值即為影像中各類物件最佳的觀測點。

找出各類最佳的第 i 個觀測點之後，再利用公式(3)，即可估算每張影像中各類物件可能所佔之比例，例如：在一張影像中，天體佔 30%，人類佔 20%，植物佔 20%，這些比例數據可以用來影像的分類。

4. 結論

本研究使用粒子群最佳化演算法，利用其粒子會依照本身過去最佳經驗與群體最佳經驗朝向目標解移動，使得粒子能更接近目標解，最終並產出實際的最佳解。藉由粒子群最

佳化演算法的特性，找出各類物件最佳之觀測點，來估算物件佔影像之比例，最後幫助影像的語意分類。

在後續研究部份，可以藉由使用者輸入欲查詢的影像與每個類別的 *Gbest* 之間距離比例，進行影像相似度的計算，透過影像相似度的計算可以建議使用者輸入的影像比較適合分類至那一個類別，以增進後續影像搜尋的效率並且滿足使用者的需求。

參考文獻

- [1] Zaher Aghbari , Akifumi Makinouchi, "Semantic Approach to Image Database Classification and Retrieval," *NII Journal*, No.7, 2005
- [2] Shyi-Chyi Cheng, T.-C. Chou, C.-L. Yang, and H.-Y. Chang, "A Semantic Learning for Content-Based Image Retrieval Using Analytical Hierarchy Process, " *Expert Systems with Applications*, vol. 28, pp. 495-505, 2005
- [3] Jing Huang, Ravi Kumar, and Ramin Zabih, "An Automatic Hierarchical Image Classification Scheme," *ACM Multimedia*, 1998.
- [4] J. Kennedy and R. C. Eberhart, "Particle swarm optimization", *Proceedings IEEE Int'l. Conf. on Neural Networks*, IV, pp.1942-1948, 1995.
- [5] Chu-Hui Lee, Yi-Ling Yin, "A semantic consumer image classification system based on multi-layer color analysis," *The 19th IPPR Conference on Computer Vision, Graphics and Image Processing*, August 13-15, in Taoyuan, Taiwan, R.O.C..
- [6] J.H. Lim, and J.S. Jin, "Combining Intra-Image And Inter-Class Semantics For Consumer Image Retrieval," *Pattern Recognition*, Vol.38, Issue 6, pp. 847-864, June 2005.
- [7] M. Szummer and R.W. Picard, "Indoor-Outdoor Image Classification", *IEEE International Workshop on Content-based Access of Image and Video Databases*, in conjunction with ICCV'98. Bombay, India, 1998.
- [8] Xiangyang Wang, Jie Yang, Xiaolong Teng, Weijun Xia and Richard Jensen, "Feature selection based on rough sets and particle swarm optimization", *Pattern Recognition Letter*, pp.459-471, 2007