IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 18, Issue 5, Ver. V (Sep. - Oct. 2016), PP 117-120
www.iosrjournals.org

Selection of Legendre Moments for Content Based Image Retrieval Using ACO Based Algorithm

Habibollah Agh Atabay [1]
[1]*(Department of Computer, Gonbad Kavous University, Iran)*

Abstract : *Feature selection is an important step in Content Based Image Retrieval (CBIR) which has a great impact on reducing complexity and increasing efficiency of CBIR frameworks. Swarm Intelligence (SI) methods, as effective optimization techniques, has been used to solve variety of problems. In this paper, a CBIR framework is established to retrieve grayscale shape images of Coil-20 object image dataset where Legendre moments are used as image features, calculated by Exact Legendre Moments (ELM) approach, and Ant Colony Optimization (ACO) algorithm is applied as the feature selection method. Finally, the selected features is feed into Support Vector Machine (SVM) classifier to categorize images. The classification accuracy are compared with the results of the system without feature selection and the results of selection using ReliefF feature weighting method. The results show that, utilizing ACO as a feature selection method can be effective to achieve higher precision in image classification and better reduce the dimensionality of moment based feature vectors.*
Keywords – *Content Based Image Retrieval, Feature Selection, Exact Legendre Moments, Ant Colony Optimization*

I. Introduction

Content based image retrieval (CBIR), is one of the important problems in information retrieval. The algorithms of this domain are used to search and retrieval of images in Internet and archived datasets. Development of imaging technologies and digital cameras, and easy propagation of images through Internet and storage devices, have motivated the increasing applications of CBIR. The basic idea of most methods of image retrieval is describing images using low level visual features, and indexing and retrieval of them using these feature vectors to overcome the tedious and time-consuming process of manual indexing. The features that are used to describe images can be categorized to color, such as color histogram and color layout, texture, such as wavelet transform and Gabor filter, shape, such as aspect ratio and Fourier descriptors and local invariants, like corner points and interest points [1, 2].

Selection of distinguishing features and reducing the feature space dimensionality is one of the most important problems in image retrieval to improve the quality and speed of querying. Dimension reduction methods are categorized into two groups of feature transform and feature selection [3]. Feature transform methods, like PCA and ICA, generate new features by mapping feature space to a low dimensional space. The problem of these methods is the difficulty for users to interpret new features, while in feature selection, the selected features can be easily analyzed. The main purpose in feature selection methods, is the selection of a subset of features among all, such that the information less is negligible. The importance of feature selection methods in pattern recognition problems is because of their effects on increasing speed and performance of the process. Feature selection methods are categorized into two classes, filter and wrapper, according to their method of selecting subsets. Filter methods select features according to intrinsic attributes of them while in wrapper methods, subsets are evaluated and selected according to the quality of a learning algorithm.

In this paper, the problem of reducing the number of features and selecting the most effective ones to use in content based object image retrieval, using a well-known meta-heuristic optimization algorithm called Ant Colony Optimization (ACO) is addressed. ACO is an evolutionary algorithm, proposed by Dorigo [4], and has been applied successfully on many problems such as recent works on robot path planning [5], train routing selection [6], medical image retrieval [7] and applications on engineering domain [8].

The feature selection method is applied on moment features, extracted from shape images of COIL-20 dataset. Different levels of Legendre moments are extracted from each image using Exact Legendre Moments (ELM) method, proposed in [9], to generate feature vectors. The samples are grouped into training, validation and test sets. Then, the ACO-based feature selection method [10] selects a subset of features according to the evaluation result of Support Vector Machine (SVM) classifier. The results of classification and retrieval using the selected subset of features are compared with the results of previous works. The results show that using the proposed framework, the quality and speed of classification and retrieval is improved.

The rest of the paper is organized as follows. In section II, the features and extraction method of them is explored. In section III, the ACO based feature selection method is described. Section IV presents the details of the proposed framework and the results of classification and retrieval and the final section is conclusion.

II. Feature extraction

Choosing the proper type of features for classification of images in various domains, is one of the important problems, especially in designing an image retrieval system. Various types of features have been proposed to describe specific properties of images. One type of features that describe shape properties of images, are moments. In this paper Legendre moments are used to generate feature vectors. Legendre moments are continuous and orthogonal which designed for describing shape of images with minimum amount of information redundancy. A method to exactly calculate this moments was proposed in [9] named Exact Legendre Moments (ELM). Legendre Moments with order $g=(p+q)$ for an image with intensity function $f(x,y)$ are defined as

$$L_{pq}=\frac{(2p+1)(2q+1)}{4}\int_{-1}^{1}\int_{-1}^{1}P_p(x)P_p(y)f(x,y)dxdy\ , \tag{1}$$

where $P_p(x)$ is the p_{th} order of Legendre polynomial and defined as

$$P_p(x)=\sum_{k=0}^{p}a_{k,p}x^k=\frac{1}{2^p p!}\left(\frac{d}{dx}\right)^p\left[\left(x^2-1\right)\right]^p , \tag{2}$$

where $x\in[-1,1]$ and $P_p(x)$ obeys the following rule

$$P_{p+1}(x)=\frac{(2p+1)}{(p+1)}xP_p(x)-\frac{p}{p+1}P_{p-1}(x)\ , \tag{3}$$

with $P_0(x)=1, P_1(x)=x$ and $P>1$.

Hosney [9] introduced a precise method to calculate Legendre moments as ELM. A set of Legendre polynomials construct a complete set of orthogonal basis in the range [-1, 1] and can be defined as follows

$$\tilde{L}_{pq}=\sum_{i=1}^{N}I_p(x_i)Y_{iq}\quad ,Y_{iq}=\sum_{j=1}^{N}I_p(y_j)f(x_i,y_j), \tag{4}$$

where

$$I_P(x_i)=\left(\frac{(2p+1)}{(2p+2)}\right)[xP_p(x)-P_{p-1}(x)]_{U_i}^{U_{i+1}}, \tag{5}$$

$$I_q(y_j)=\left(\frac{(2q+1)}{(2q+2)}\right)[yP_q(y)-P_{q-1}(y)]_{V_i}^{V_{i+1}}, \tag{6}$$

$$U_{i+1}=x_i+\frac{\Delta x_i}{2}=-1+i\,\Delta x\ , \tag{7}$$

$$U_i=x_i-\frac{\Delta x_i}{2}=-1+(i-1)\Delta x\ , \tag{8}$$

$$V_{j+1}=y_i+\frac{\Delta j}{2}=-1+j\,\Delta y\ , \tag{9}$$

$$V_j=y_j-\frac{\Delta y_j}{2}=-1+(j-1)\Delta y\ , \tag{10}$$

In above equations $(U_i,\ V_j)$ is the center of a pixel of any image with coordinates $(x_i,\ y_j)$.

III. Feature selection based on ant colony optimization

Having a set of features with size of *n*, the problem of feature selection is to find a minimum subset with size *s* (*s*<*n*) such that better classification accuracy in comparison with utilizing all features is achieved. In this paper, a feature selection method based on ACO algorithm, presented in [10], called ACOFS, is used. The first step of the method is describing the problem as a graph of routes where artificial ants can move through edges of graph to find the optimal path. In the algorithm, instead of using a complete graph feature selection, the graph of Fig 1 has been used. The reason for that, is that in contrast with problems like routing, where the order of selection is important, in feature selection it is not.

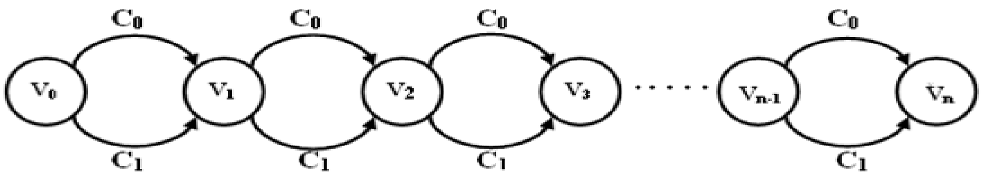

Figure 1. The graph to describe feature selection

In Fig 1, for i_{th} feature$\left(i=1..n\right)$, a node v_i is drawn and v_0 demonstrates the begin point for each artificial ant to route. The ants travel the graph, beginning from v_0 , to reach the node v_n and present the results as a feature selection solution. Two types of edges with labels of C_j^0 and C_j^1 connect two adjacent nodes v_j and v_{j-1}. If an ant in node v_j selects C_j^0 (or C_j^1), the j_{th} feature is (or is not) selected. On each edge C_j^i the artificial value of pheromone τ_j^i is assigned as feedback information of each ant to guide other ants in graph. The matrix of pheromones is initialized with the value of $\tau_j^i=1$ for each $i=1,2,\ldots,n$ and $j=0,1$.

The probability function of selecting a feature or not, is equals to the probability of choosing edge C_j^i in the node v_{j-1} to reach the node v_j calculated by composition of heuristic desirability and pheromone density using the following formula:

$$P_j^i\left(t\right)=\frac{\left[\tau_j^i\left(t\right)\right]^{\alpha}\left(\eta_j^i\right)^{\beta}}{\left[\tau_j^0\left(t\right)\right]^{\alpha}\left(\eta_j^0\right)^{\beta}+\left[\tau_j^1\left(t\right)\right]^{\alpha}\left(\eta_j^1\right)^{\beta}}\ \left(i=0,1\,;j=1,2,\ldots,n\right), \tag{11}$$

Where $\tau_j^i\left(t\right)$ is the pheromone value on the edge C_j^i between the nodes v_j and v_{j-1} in time t which denotes the potential trend of ants to follow the pathC_j^i . η_j^i is the heuristic information of each ant to choose the edge C_j^i . The parameters α and β determine the relative importance of pheromone or heuristic information. The pheromone value of edges must update after each iteration. It is obvious, if an ant chooses the edgeC_j^i , the pheromone value of that edge must be increased and in next iteration the ants will choose that edge with higher probability. Also if ants don't choose an edge, the pheromone value of it, must be decreased. In each iteration, the pheromone values are updated using the following formula:

$$\tau_j^i\left(t+1\right)=\rho.\tau_j^i\left(t\right)+\Delta\tau_j^i\left(t\right)+Q_j^i\left(t\right), \tag{12}$$

where the following relationships are established:

$$\Delta\tau_j^i\left(t\right)=\frac{1}{\left|S_j^i\right|}\sum_{s\in S_j^i}f\ \left(s\right), \tag{13}$$

and

$$Q_j^i\left(t\right)=\begin{Bmatrix}Q & C_j^i\in s_{best}\\ 0 & otherwise\end{Bmatrix}, \tag{14}$$

where S_j^i is the set of solutions which are found in t_{th} iteration and passes from C_j^i , s_{best} is the best solution ever found and Q is a positive constant used for emphasis is on the best route.

IV. Experimental Results

In this paper, an application of shape feature selection on COIL-20 dataset [11] of object images is

Table 1. The comparative results

	Without Feature Selection			**Feature Selection using ReliefF**			**Feature Selection using ACOFS**		
Order of ELM	Number of Features	Classification Accuracy of SVM	Retrieval Result	Number of Features	Classification Accuracy of SVM	Retrieval Result	Number of Features	Classification Accuracy of SVM	Retrieval Result
4	15	0.9815	0.5737	10	0.9899	0.6523	7	0.9752	0.7382
5	21	0.9583	0.5905	11	0.9444	0.6778	8	0.9687	0.7735
6	28	0.9421	0.6103	12	0.9213	0.7121	8	0.9603	0.7895
7	36	0.8981	0.6289	10	0.9815	0.7124	9	0.9812	0.6929
8	45	0.8171	0.6475	12	0.8958	0.7169	10	0.9714	0.7526
9	55	0.7731	0.6586	13	0.9769	0.7201	10	0.9589	0.7834

presented. The dataset is consists of 1440 images of 20 class of various objects, captured from different angles. To describe shape features of each image, Legendre moments calculated be ELM algorithm, detailed in section two, are used. An application of these feature was presented in [12] where Legendre moments of orders 4 to 9 were used to generate feature vectors and the classification and retrieval results of COIL-20 dataset was presented. The feature vector lengths in this method are 15, 21, 28, 36, 45 and 55, respectively. The method of ACO based feature selection, detailed in section III, is used. The parameters of this method are set as $\alpha = 1$, $\beta = 1$, $\rho = 0.1$, maximum number of iteration = 50 and number of artificial ants in each iteration = 20. The dataset is randomly split into 70% training, 10% validation and 20% test sets. The accuracy of SVM classifier, on the validation dataset using the selected subset of features, is used for evaluation of the subset. To do this, SVM is trained using training set, described by selected features, and for each sample of the validation dataset a query is proposed, and the average classification accuracy of SVM for all samples, is used as the measure of quality of the selected subset. The F-Score parameter is used similar to [9] as the heuristic function of artificial ants.

The results of the experimented method is compared to the previous work [13], where the ReliefF algorithm was used as feature selection method, and [12] where feature selection method was not used. The results are shown in Table 1. The table show that the classification and retrieval accuracy of feature selection using ACOFS are surpassed the results of retrieval without feature selection and retrieval using ReliefF based feature selection. This improvements are more observable when the feature vectors length is increased.

V. Conclusion

In this paper, an application of ACO based feature selection method (ACOFS), in classification and retrieval of shape images of Coil-20 object dataset, was presented. Legendre moments, extracted by ELM method, were used as image features. The results were compared to other similar works, [12] and [13] where in the former, different orders of Legendre moments have been used without feature selection and in the later, a ReliefF based feature selection algorithm has been applied on similar feature vectors. The results showed that a small subset of feature among moments are exist which have more important role in distinguishing feature vectors to achieve better classification accuracy and thus other feature can be omitted without degrading the retrieval quality. The classification and retrieval speed is also improved by proper feature selection. This results are more obvious when the feature vector length is increased using higher orders of moments.

References

[1]. M. Rehman, M. Iqbal, M. Sharif, M. Raza, *World Applied Sciences Journal, 19(3)*, 2012, 404-412.
[2]. F. Rajam, S. Valli, *Life Science Journal, 10(2)*, 2013, 2475-2487.
[3]. G. Chandrashekar, F. Sahin, *Computers & Electrical Engineering, 40(1)*, 2014, 16-28.
[4]. M. Dorigo, M. Birattari, T. Stutzle, *IEEE Computational Intelligence Magazine, 1(4)*, 2006, 28-39.
[5]. Y. Yao, Q. Ni, Q. Lv, K. Huang, A novel heterogeneous feature ant colony optimization and its application on robot path planning, 2015 IEEE Congress on Evolutionary Computation (CEC), IEEE, 2015, 522-528.
[6]. M. Samà, P. Pellegrini, A. D'Ariano, J. Rodriguez, D. Pacciarelli, *Transportation Research Part B: Methodological, 85(*2016, 89-108.
[7]. V. Jeyakumar, B.R. Kanagaraj, *Journal of Medical Imaging and Health Informatics, 5(7)*, 2015, 1383-1389.
[8]. B.C. Mohan, R. Baskaran, *Expert Systems with Applications, 39(4)*, 2012, 4618-4627.
[9]. K.M. Hosny, *Pattern Recognition, 40(12)*, 2007, 3597-3605.
[10]. L. Chen, B. Chen, Y. Chen, Image feature selection based on ant colony optimization, Australasian Joint Conference on Artificial Intelligence, Springer, 2011, 580-589.
[11]. S.A. Nene, S.K. Nayar, H. Murase, Columbia object image library (COIL-20), Technical report CUCS-005-96, 1996.
[12]. C. Rao, S.S. Kumar, B.C. Mohan, *arXiv preprint arXiv:1005.5437*, 2010.
[13]. A. Sarrafzadeh, H.A. Atabay, M.M. Pedram, J. Shanbehzadeh, ReliefF Based Feature Selection In Content-Based Image Retrieval, Proceedings of the International MultiConference of Engineers and Computer Scientists, 2012.