ReliefF Based Feature Selection In Content-Based Image Retrieval

Abdolhossein Sarrafzadeh, Habibollah Agh Atabay, Mir Mosen Pedram, Jamshid Shanbehzadeh

***Abstract*—content based image retrieval systems usually extract low level features to retrieve similar images. But in most cases, selection of suitable features according to their impact on the classification accuracy has been less considered. This paper studies the effects of reducing the number of features and selecting the most effective subset of features in the context of content-based image classification and retrieval of objects. We use Legendre moments to extract features, ReliefF algorithm to select the most relevant and non-redundant features and support vector machine to classify images. The experimental results on Coil-20 image dataset, shows that by selecting much lower number of features when employing ReliefF, we can improve retrieval in terms of speed and accuracy.**

***Index Terms*— Content Based Image Retrieval, Feature Selection, ReliefF Algorithm, Exact Legendre Moments, Support Vector Machine**

I. Introduction

Content based image retrieval or CBIR as a branch of information retrieval is an important research area because of its wide applications as a useful tool for retrieving images from the Internet, medical image archives, pattern recognition and so on. The development of high quality cameras with very low prices and easy distribution of images in Internet and shared storage devices are reasons for increasing demand for applications of CBIR techniques. The main motivation of using CBIR instead of text-based retrieval (annotation) is that manually indexing of large image databases is very tedious and time-consuming.

The basis of most CBIR techniques is to describe images by a set of low level features called feature vector and then retrieving similar images by measuring the similarity of feature vectors between the query image and images of database. There are many types of features that have been employed to construct feature vectors such as color, texture, shape and salient points [1]. Shape features commonly used in shape retrieval, are basic and important group of features used to describe image content specially segmented image regions and specific images such as man-made objects. Moment descriptors are suitable to describe shapes. This is mainly due to their ability to fully describe an image by encoding its contents in a compact way. Orthogonal moment is widely used as descriptors. The main reason is that the orthogonal moments can represent an image with the minimum amount of information redundancy thus the recovery of an image from its geometric moments is accurate. These moments can be classified into continuous and discrete. Examples of continuous orthogonal moments are Legendre, Zernike, Pseudo-Zernike and Fourier-Mellin moments. Examples of discrete ones are Tchebichef, Krawtchouk, Racah and Dual Hahn moments [2]. Orthogonal moment firstly introduced in [3] as a generalization of geometric moments with using Legendre, Zernike and other polynomials as kernel function. Legendre moments are used in many applications such as blurred image recognition [4], tissue classification [5] and pattern recognition and computer vision applications [6].

The improvement of computational efficiency without losing accuracy of CBIR systems can be performed by selecting the best features and reducing length of feature vector. We can divide the feature reduction methods into two groups; feature transform and feature selection [7]. Feature transform methods, such as PCA and ICA, maps the original feature space into the lower dimensional space and construct new feature vectors. The problem of feature transform algorithms is their sensitivity to noise and the resultant features convey no meaning for user. While, feature selection schemes is robust against noise and the resultant features are highly interpretable.

The objective of feature selection is to choose a subset of features to reduce the length of feature vectors with the lowest information loss. Feature selection schemes according to their subset evaluation methods, categorized into two groups: Filter and Wrapper [8]. In filter methods, features evaluated with their intrinsic effect on separating classes while wrapper methods use accuracy of the learning methods to evaluate subset of features.

This paper explores the problem of feature space reduction and selecting effective features in a content-based image classification and retrieval system based on one of the effective feature selection algorithms called ReliefF. To do this, we first construct a feature vector for each image of the Coil-20 [9] gray scale image dataset using moment-based shape features. Then we calculate a weight for each feature using ReliefF algorithm and select the k top features as the effective subset. Then we evaluate this subset using accuracy of SVM classifier. After evaluating subsets and selecting the best one, we calculate the results of retrieval and compare them with results of classification and retrieval when we consider all the features. The results of

Manuscript received December 6, 2011 and accepted December 23, 21011

A. Sarrafzadeh is an Associate Professor and Head of Department of Computing at Unitec, New Zealand. (E-mail: hsarrafzadeh @unitec.ac.nz)
H. Agh Atabay is with Tarbiat Moallem University of Tehran, Tehran, IR. Iran (e-mail: atabay@tmu.ac.ir).
M. M. Pedram is an Assistant Professor with Tarbiat Moallem University of Tehran, Tehran, IR. Iran (e-mail: perdam@tmu.ac.ir).
J. Shanbehzadeh is an Associate Professor with Tarbiat Moallem University of Tehran, Tehran, IR. Iran (e-mail: jamshid@tmu.ac.ir).

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

comparisons show that with selecting prominent features, we can reduce the number of features and improve the quality and speed of the classification and retrieval.

The rest of the paper is organized as follows. The next section explains feature vector extraction. Section three and four introduce ReliefF algorithm for feature selection and Support Vector Machine for image classification in brief, respectively. Section five explores the details of our feature selection method and in section six we demonstrate the experimental results of classification and retrieval. Section seven concludes the paper.

II. Feature Extraction

Selecting proper types of extractable visual features for image classification and retrieval in different contexts, is one of the major problems for designing CBIR systems. CBIR systems employ different types of features. Shape features are an important type of features in distinguishing man-made objects in an image. One of the major types of shape features is moment such as Legendre, Zernike, and Tchebichef Moments [10].

Legendre moments are continuous and orthogonal and can be used to represent images with minimum redundancy. Hosney [11] introduced a precise method to calculate this moment named Exact Legendre Moments (ELM). Legendre Moments with order *g=(p+q)* for an image with intensity function *f(x,y)* are defined as:

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

Where $P_p(x)$ is the p^{th} order of Legendre polynomial and defined as:

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

Where x∈[-1,1] and $P_p(x)$ obeys the following rule:

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

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

A set of Legendre polynomials construct a complete set of orthogonal basis in the range [-1, 1] and can be defined as follows [6].

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

where

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

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

$$U_{i+1} = x_i + \frac{\Delta x_i}{2} = -1 + i\Delta x \qquad (7)$$

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

$$V_{j+1} = y_i + \frac{\Delta j}{2} = -1 + j\Delta y \qquad (9)$$

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

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

III. Feature Reduction with ReliefF Algorithm

We can improve the performance of CBIR systems if we consider the weight of features in retrieval rather than treating all the features equal. Because the effect of various features in image classification is may be different. One way of measuring this effect is to use ReliefF algorithm [12]. This algorithm is a variation of Relief algorithm and is used in multi class situations unlike the original Relief. The major drawback of Relief is its sensitivity to noise. ReliefF algorithm has a great capabilities to deal with noisy an unknown data [12].

IV. Classification with SVM

Support Vector Machine or SVM is an unsupervised learning method that has been greatly used in CBIR such as [13]-[16]. SVM classifier uses a hyperplane and a kernel function to nonlinear-classification of two class data. SVM constructs a hyperplane that maximize the margin between two classes. This hyperplane is called the optimal hyperplane. Margin is defined as the distance between closest points called support vectors to the separating hyperplane. To facilitate finding optimal hyperplane, SVM maps the feature space to a high dimensional space using a kernel function. SVM can be used for multi class classification. To do this we have two situations: one-against-one or one-against–all SVMs. In one-against-one implementation of SVM, for each pair of classes one SVM is created and one class is labeled as positive and the other is considered as negative. But in one-against–all SVMs, a SVM is constructed for each class and this class is labeled as positive and the others are considered negative [13]. This paper uses one-against-one SVM.

V. Proposed Method

The objective of this paper is to provide a way to evaluate the impact of features on the quality of semantic classification of specific images and its application in CBIR. We use ReliefF algorithm for this purpose. This algorithm is a feature weighting method and its basis relies on decreasing inter class distance while increasing outer class distance. This algorithm assigns a higher weight to the features that improve this criterion [12].

This paper evaluates ReliefF-based feature selection method by the use of the Coil-20 image dataset that contains 1440 grayscale pictures from 20 classes of objects. These pictures are described by moment-based shape features where these features are extracted using ELM method. Rao et al [16] presented an application of these moments in CBIR where they employed moments of order 4 to 9 of ELM to construct feature vectors and the length of feature vectors in this way is 15, 21, 28, 36, 45 and 55. This paper employs the same feature vector. Major steps of the proposed algorithm are as follows:

1. Features are extracted from each image in the dataset and feature vectors are constructed.

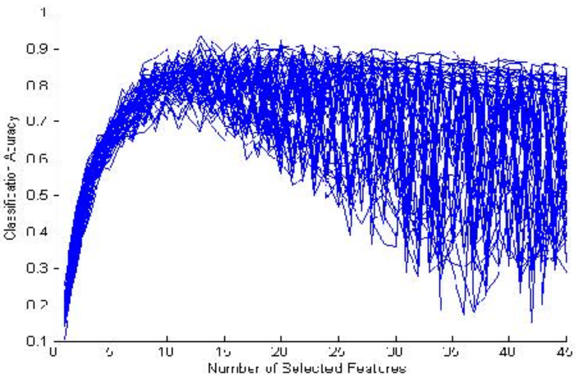

Figure 1: The effects of features on classification accuracy

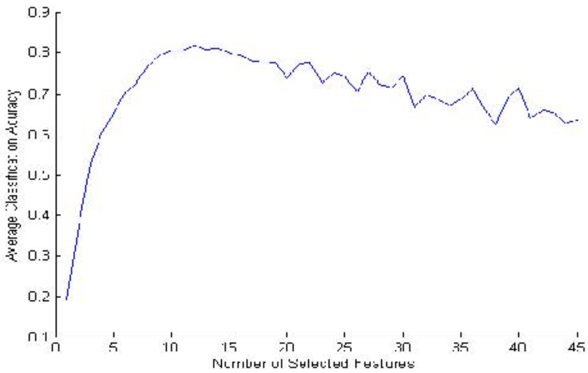

Figure 2: The effects of features on classification accuracy

2. Images in the dataset are divided into Train and Test groups.
3. Weight for each feature is calculated by applying Relief-F algorithm to the train data.
4. The quality of features examined with test data using SVM classifier.
5. One feature that has minimum weight value is discarded.
6. Steps 4 and 5 are repeated until reaching the minimum number of features.
7. Steps 2 and 6 are repeated until reaching maximum iteration.
8. Results for each number of selected features are averaged.
9. Specific number of features that maximize these averages is selected.
10. Finally the best results of classification and retrieval with whole features and with selected number of high weighted features are compared.
11. Then above steps for various orders of ELM as the feature vectors are repeated.

Feature vector for each image is presented as follows where $g=(p+q)$ is the order of ELM:

$$f_i = \{L_{00}L_{01} \dots L_{pq}\} \tag{11}$$

The objective here is to achieve the maximum efficiency by the minimum number of features. Because the results are independent of selecting method of train and test data, the above steps is repeated for a certain number of iterations for randomly selected train and test data and we use the average of the obtained results.

The result of classification has been obtained by calculating the ratio of the number of images correctly classified by the total number of test images. To do this, each feature vectors is given to the CBIR system as query. Then the similarity of that to each of the feature vectors of database are calculated by Canberra distance and certain number of images that has the closest distance to the query image is retrieved. Canberra distance is formulated as follows.

$$d_i = \sum_{i=1}^{M} \frac{|f_q - f_i|}{|f_q| - |f_i|} \tag{12}$$

In (12) f_q is the feature vector of query image is, f_i is any feature vector from database and M is the total number of images.

VI. Experimental Results

This paper used orders 4 to 9 of ELM to construct feature vector for each image in Coil-20 dataset that consists of 1440 grayscale images. Then features database are randomly divided into two groups; 70 percent as train and 30 percent as test data. Then the proposed method is applied on them. This experiment is repeated 50 times for each one of ELM orders. Fig. 1, shows the results for order 8 for 50 examination altogether and Fig. 2 is the average results. The results of other orders are similar to above diagrams. This diagrams show that with a specific number of top-weight features (in the case of order 8 this number is 12) average result of classification using SVM is maximized. We used this feature as the result of feature selection and compared the results of classification and retrieval obtained by this subset with the results obtained by whole features.

Table I, shows the complete results. It observed that the results of classification are improved with the selected subset of features. These improvements, especially when the vector length increases is more visible. The results of retrieval are also improved. It should be noted that the Canberra distance, also treats all the features similarly. So in addition to feature selection, we multiplied weights to distances and show the results in a separate column. In the last column we have used weighted Canberra distance as the similarity measure and the results are also improved.

Table I. The Comparison Of Classification And Retrieval In Case Of Considering And Ignoring Weight Of Features.

Order of ELM	Number of Features	Result of Classification	Result of Retrieval	Number of Selected Features	Result of Classification	Result of Retrieval	Result of Retrieval with Weights Multiplication
4	15	0.9815	0.5737	10	0.9699	0.6523	0.6628
5	21	0.9583	0.5905	11	0.9444	0.6778	0.6792
6	28	0.9421	0.6103	12	0.9213	0.7120	0.7110
7	36	0.8981	0.6289	10	0.9815	0.7124	0.7148
8	45	0.8171	0.6475	12	0.8958	0.7169	0.7202
9	55	0.7731	0.6586	13	0.9769	0.7201	0.7238

VII. Conclusions

This paper presented the application of ReliefF algorithm as a feature weighting method in selection of useful features for content-based image classification and retrieval using moment-based shape features. Results of these applications were compared with the results in [16] which ignored feature weighting and selection. The simulation results showed that a small subset of features was requires in classification and retrieval of images thus the speed of retrieval was increased subsequently. The effect of feature selection especially by increasing order of ELM and thus increasing feature vector length was more noticeable. This suggested that employing more features had negative effects on the results of CBIR system.

REFERENCES

[1] R. Datta, D. Joshi, J. Li, and J.Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age," ACM Computing Surveys, vol. 40, 2008, pp. 1-60.

[2] [6] HZ. Shu, LM. Luo, and JL. Coatrieux, "Moment-based approaches in imaging . Part 1 , basic features," IEEE Engineering in Medicine and Biology Magazine, 2007, pp. 1-8.

[3] MR Teague. “Image Analysis via the General theory of Moments.”, Journal of Optical Society of America, vol. 70, 1980, pp. 920-930.

[4] H. Zhang, et al., “Blurred image recognition by Legendre moment invariants,” IEEE Trans. Image Process., vol. 19, 2009, pp. 506-611.

[5] V.S. Bharathi and L. Ganesan, "Orthogonal moments based texture analysis of CT liver images," Pattern Recognition Letters, vol. 29, 2008, pp. 1868-1872.

[6] G.A. Papakostas, E.G. Karakasis, and D.E. Koulouriotis, "Accurate and speedy computation of image Legendre moments for computer vision applications," Image and Vision Computing, vol. 28, 2010, pp. 414-423.

[7] M.E. ElAlami, "A novel image retrieval model based on the most relevant features", presented at Knowl.-Based Syst., 2011, pp.23-32.

[8] Chulmin Yun, Donghyuk Shin, Hyunsung Jo, Jihoon Yang, and Saejoon Kim, “An Experimental Study on Feature Subset Selection Methods”, International Conference on Computer and Information Technology IEEE, 2007, pp. 77 - 82

[9]. S. A. Nane, S. K. Nayar, and H. Murase, (1996) “Columbia Object Image Library: COIL-20,” Dept. Comp. Sci., Columbia University, New York, Tech. Rep. CUCS-006-96.

[10] G.A. Papakostas, D.E. Koulouriotis, and E.G. Karakasis, "Computation strategies of orthogonal image moments: A comparative study", presented at Applied Mathematics and Computation, 2010, pp.1-17.

[11] K. M. Hosney, “Exact Legendre moments computation for gray level images”, Pattern Recognition, Vol. 40, 2007, pp. 3597-3605.

[12] I. Kononenko. Estimating attributes: Analysis and extensions of RELIEF. In L. D. Raedt and F. Bergadano, editors, European Conf. on Machine Learning, Catania, Italy, Springer Verlag, New York, 1994. Pp. 171–182

[13] C.Chang , H.Wangb, C.Li, “Semantic analysis of real-world images using support vector machine”, Elsevier, Expert Systems with Applications vol. 36, 2009, pp.10560–10569.

[14] K. K. Seo, An application of one-class support vector machines in content-based image retrieval, Expert Systems with Applications vol. 33, 2007, pp. 491–498.

[15] Wai-Tak Wong , Frank Y. Shih, Jung Liu, “Shape-based image retrieval using support vector machines, Fourier descriptors and self-organizing maps”,Information Sciences, vol. 177, 2007, pp. 1878–1891

[16] Ch.Srinivasa Rao, S.Srinivas Kumar and B.Chandra Mohan, Content Based Image Retrieval using Exact Legendre Moments and Support Vector Machine, The International Journal of Multimedia & Its Application (IJMA), Vol.2, No.2, May 2010

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)