OPTIMIZED IMAGE FEATURE SELECTION USING PAIRWISE CLASSIFIERS

Mahdi Bazarganigilani
Charles Sturt University
Faculty of Business,
Melbourne, Australia
e-mail: mbazar02@postoffice.csu.edu.au

Abstract

In this paper, we introduce an optimized method to improve the accuracy of content based image retrieval systems (CBIR). CBIR systems classify the images according to low and higher features.In our research, we improve both feature selection and classifier partition of a CBIR system. Results show great performance of our proposed algorithm.

1 Introduction

By increasing amount of multimedia data on information systems, there is a huge demand for more accurate and effective image retrieval systems. The similar semantically images likely have similar vision features. This is the main idea behind CBIR systems [1]. Such systems use positive feedback of the users to improve accuracy of image classification [2].

By improving either classification algorithm or retrieval approaches of the system, the overall accuracy improved. There are researches on classification algorithms [1, 3] and retrieval methodologies [4, 6]. In our research, we focus on choosing efficient features of the images [7]. Consequently, we improve the accuracy of the classifier by using pair wise classifiers.[8]

One of technique to increase the accuracy of the classifier is to use ensemble of classifiers [9]. In this method we use multiple pair wise classifiers instead of one multiple classifier. Pair-wise classifier needs to distinguish between two classes and it has more accuracy toward multiple classifier

2 Feauture Vector of Images

To classify the images, we need to extract the low level feature of images. Low level feature include color and shape properties of the images [10].

In case of low-level image features, a feature vector which describes various visual cues, such as color, shape, or texture, is computed for each image in the database. These include features based on color histograms [11, 12, 13, 14, 15, 16, 17, 18, 19], color moments [18, 20, 21, 22], statistical shape and texture features such as edge direction histograms [23], Tamura texture features [24], and wavelet-based texture features [22, 25]. Given a query image, its feature vector is calculated and those images which are most similar to this query image based on an appropriate distance measure in the feature space are retrieved [26].

Color histograms are generally invariant to translation and rotation of the images and normalizing the histograms leads to scale invariance [27]. However, color histograms do not incorporate spatial adjacency of pixels in the image and may lead to inaccuracies in the retrieval. The image color distribution is represented without any additional information about spatial location or shape of homogeneous colored regions. Images completely different from the perspective of users may have the same color composition. Therefore, retrieval of images based on color histograms are prone to yielding a large number of false hits, i.e., images with completely different content which just happen to have a similar color composition as the query image [2].

To overcome the quantization effects as in color histogram, [20] proposed to use color moments approach. Color moments of an image are a simple yet effective feature for color-based image retrieval [20, 21, 22]. From probability theory we know that a probability distribution is uniquely characterized by its moments. Thus, if we interpret the color distribution of an image as a probability distribution, then the color distribution can be characterized by its moments, as well. Furthermore, as most of the color distribution information can be captured by the low-order moments, using only the first three moments: mean, variance and skewness, it is found that these moments give a good approximation and have been proven to be efficient and effective in representing color distribution of images [20]. These first three moments are defined as:

$$u_i = \frac{1}{N}\sum_{j=1}^{N} p_{ij}$$

$$\sigma_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(p_{ij}-u_i)^2}$$

$$s_i = \left(\frac{1}{N}\sum_{j=1}^{N}(p_{ij}-u_i)^3\right)^{\frac{1}{3}}$$

Where p_{ij} is the value of the ith color channel of the jth image pixel.

Shape is another feature that can be used in content-based image retrieval. In the absence of color information or in the presence of images with similar colors, it becomes imperative to use additional image attributes for an efficient retrieval [26].

Various schemes have been proposed in the literature for shape-based retrieval. Shape representation techniques can be divided into two categories: boundary-based and region-based. The boundary-based uses only the outer boundary of the shape while the latter uses the entire shape region (29). The most successful representatives for these two categories are Fourier descriptors and moment invariants respectively. The main idea of Fourier descriptors is to use the Fourier transformed boundary as the shape feature [30].

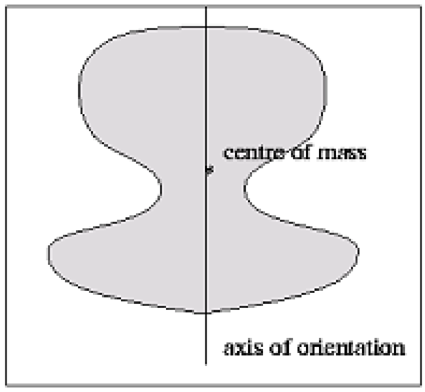

Figure 1. The 0th moment is the area of the object; the 1st moment gives the centre of mass; and the 2nd moments give the axes of orientation.

The main idea of moment invariants is to use region-based moments, which are invariant to transformations, as the shape feature. In [31], the authors represent the shape of an image in terms of seven invariant moments based on the 2nd and 3rd order moments. The idea of using moments in shape recognition gained prominence because of its effectiveness. The central moment of a digitally sampled image in order of $(p+q)$ that has gray function $f(x,y), (x,y=U,...M-1)$ is given by,

$$u_{pq} = \sum_x\sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y)$$

Seven moment invariant based on 2d and 3d order moments gained as follow.

$$\begin{aligned}
&M_1 = (u_{20}, u_{02}),\\
&M_2 = (u_{20}-u_{02})^2 + 4u_{11}^{\;2},\\
&M_3 = (u_{30}-3u_{12})^2 + (3u_{21}-u_{03})^2,\\
&M_4 = (u_{30}+u_{12})^2 + (u_{21}+u_{03})^2,\\
&M_5 = (u_{30}+u_{12})(u_{30}-3u_{12})\left[(u_{30}+u_{12})^2 - 3(u_{21}-u_{03})^2\right]\\
&+(3u_{21}-u_{03})(u_{21}+u_{03})\left[3(u_{21}+u_{03})^2 - (u_{21}-u_{03})^2\right],
\end{aligned}$$

$$M_6 = (u_{20} - u_{02})\left[(u_{30} + u_{12})^2 - (u_{21} + u_{03})^2\right] + 4u_{11}(u_{30} + u_{12})(u_{21} + u_{03}),$$
$$M_7 = (3u_{21} - u_{03})(u_{30} + u_{12})\left[(u_{30} + u_{12})^2 - 3(u_{21} + u_{03})^2\right] - (u_{30} - 3u_{12})(u_{21} + u_{03})\left[(u_{30} + u_{12})^2 - 3(u_{21} - u_{03})^2\right]$$

M_1through M_6 are invariant under rotation and reflection. Scale invariance is achieved through following formula.

$$M'_1 = M_1/n \quad M'_2 = M_2/r^4 \quad M'_3 = M_3/r^6$$
$$M'_4 = M_4/r^6 \quad M'_5 = M_5/r^{12} \quad M'_6 = M_6/r^8$$
$$M'_7 = M_7/r^{12}$$

Where r is the radius of gyration of the object.

$$r = (u_{20} + u_{02})^{1/2}$$

Low-level texture features can provide both global and local information, but are hard to define. Moreover, they are limited in their ability in describing the semantic content in the image. Another limitation of texture features is the high computational complexity of matching based on these features [26].

3 Normalization of Feature Vector

Normalization means putting the real value between two predefined values. Mostly, we map the real values between 0 and 1 [11].To normalize the feature i from vector k,we use the following formula.

$$x'_{i,k} = \frac{x_{i,k}}{\max(x_{i,1}, \ldots x_{i,n})}$$

Normalization is important in learning classifiers. Thus is why, we use sigmoid function in our neural network classifier. Such networks have very small first derivative for the numbers outside the range [-5, 5].

If the network feeded with the values outside the normalized values, the network converges to 0.Thus, the learning is impaired due to non normalized inputs . By normalizing the inputs, we impede such problem [7].

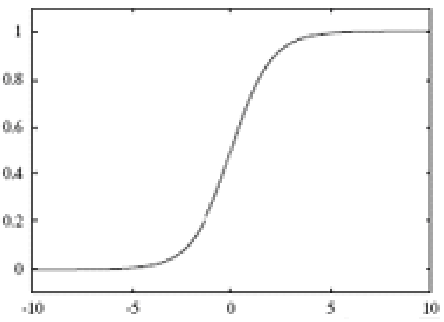

Figure 2. Sigmoid Function.

4 Descretization of Inputs

The main advantage of descretization is to decrease the amount of inputs without losing the accuracy of the classifier [7]. We consider value sets of a feature as $f_k = \{x_1, x_2, ..., x_n\}$.We sort such values, Therefore, we have $x_1 \leq x_2 \leq ... \leq x_n$.We divide the set from the point x_p.The suitable point for x_p dividing feature set of f_k in feature vectors U is the position minimizes the following formula [11].

$$E(f, x_p) = \frac{S_1}{U} Entropy(S_1) + \frac{S_2}{U} Entropy(S_2)$$

$$Entropy(S) = -\sum_{c' \subseteq C} \frac{Sc'}{S} \log_2 \frac{Sc'}{S}$$

In which,

U : The set of feature vectors.

S_1 : The feature vectors which $f_k \leq x_p$.

S_2: The feature vectors which $f_k \succ x_p$.

C: The set of semantic concepts in image inputs.

We continue the division of x_p to obtain the best point for each feature set.

5 Feature Selection

There are many features which is not proper for semantic classification. For example brightness is not a good feature. Since every picture from different classes may have different brightness. The main idea behind feature selection is to identify the most differences among classes [32].

We use Information Gain approach to select best features for classifiers. Information Gain is presents the power of any feature for differentiating the instances [33].Information Gain of feature Ain feature sets of S

$$Gain\left(S,A\right) =$$
$$= Entropy\left(S\right) - \sum_{v \in Value(A)} \frac{|S_v|}{|S|} Entropy\left(S_v\right)$$

In which,

$Value\left(A\right)$: The value set of feature A.

S_v : Subset of S in which feature A has value v.

We select best features and eliminate 20 percent of worst feature sets in this step.

6 Pairwise Classifiers

In this section, we explain using our pair wise classifiers instead of one multiple classifier. In such manner the accuracy of classifier improved. The learning in multiple classifiers is to distinguish one category against other categories.

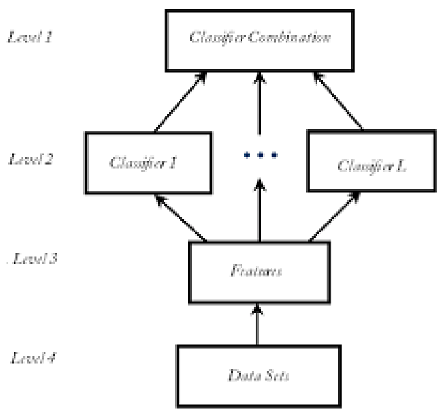

Figure 3. Different levels in making ensemble of classifers.

While in pair-wise one, it is based on the other category. We use a secondary network to combine the results of pair wise classifiers [8].

We feed the feature vectors to our multiple par wise classifiers. In this algorithm, the combination of results obtained by a multilayer perceptron network.

In figure 4, we have shown the accuracy of using pair-wise classifier. In the figure distinguishing the rectangles from other shape like circles is more accurate. On the other hand, discriminating the rectangles among all the shapes leads to less accuracy. In this approach, firstly, the couple classifiers with maximum error rate of classification determined. Such categories need to be feeded to pair wise classifier for improving overall efficiency of the classifier.

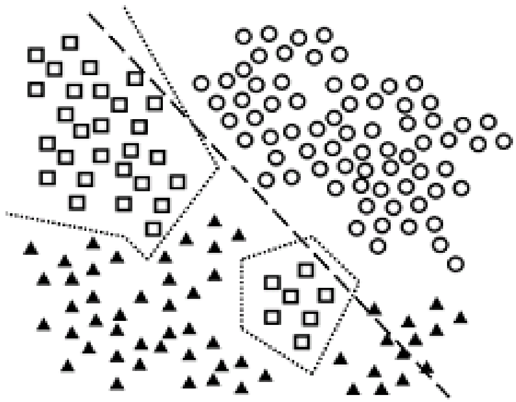

Figure 4. Simplicity of classifier in pairwise boundries.(the boundry of two pices of square with circles has shown with dashed line and with others with dotted line) [8].

7 Determining the Most Erroneous Pairwise Categories

The define which feature should be divided from the multiple learning classifier, we need to define the most erroneous pair-wise categories which have most intersection with each other. Such couple categories feeded to pair-wise classifiers to be categorized more accurately.

We use a multiple neural network classifier to determine error rate of couple categories. We have feeded such networks with the features gained with Information Gain criteria. Consequently, we make a confliction matrix to define the categories need to be classified. The element, $a_{i,j}$ determines the number of incorrect instances from class i classified in class$j$[8].

8 Combination of Pair-wise Classifers

The final step for classifying the feature vectors I to combine the result of primary network and pairwise networks. We use a multilayer perceptron network to combine the results. The output of multi classifier network and pair-wise networks feeded to combination network. The following figure shows a classifier system to both classifying and combining the results [8].

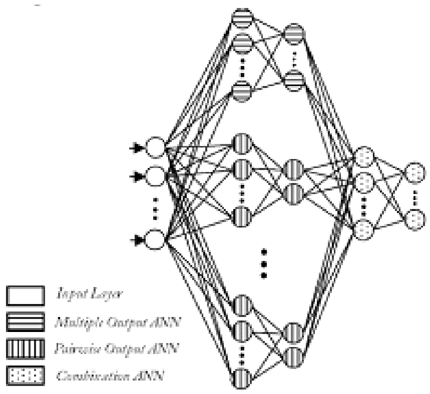

Figure 5. Structure of Artificial Neural network Classifiers.

In figure 5, we have shown the structure of our neural network. In this network the input feature vector feeded to the system. In next layer, there are numbers of pair wise classifier along the multi neural network .the pair wise classifiers created due to couple categories with high error rate of classification. In combination layer all the outputs of pair-wise classifiers and the multi neural networks combined and the output layer consists of the numbers of categories.

9 Experiment Results

To evaluate our algorithm, we choose 300 pictures from LabelMe database [34].Such groups categorized in 10 concept groups including, tree, seashore, sky, sun, hand, cloud, library, window, road and wall. We obtained 13 features of pictures including color and shape properties.

We applied Information Gain in our feature vectors and eliminate 3 worst of them. We feeded the rest features to one multiple Neural Network and determine 4 couple categories with the most intersection and error rate. Therefore, we made a 4 pair wise networks with 10 feature inputs and one multiple network. The combination network has the 18 inputs and ten outputs to determine the concept class of the images. We have shown the results of applying each step in the Table 1.

Algorithm	Number of Feature	Precision
Raw Feature	13	74%
Raw Feature	7	68%
Normalized Features	13	77%
Normalized Features	7	73%
Normalized Descretized Features	13	81%
Normalized Descretized Features	7	75%
Pair-Wise NNs	13	88%
Pair-Wise NNs	7	80%

Table 1. Algorithm Performance

In another experiment, we measure the speed of convergence without considering pair-wise networks. As shown, every step results in faster convergence rate and less numbers of Epochs.

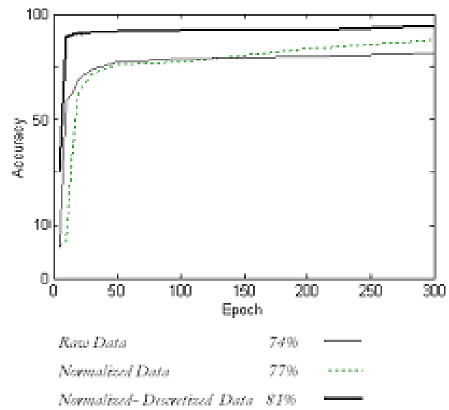

Figure 6. Convergence Speed in Normalized and Descretized Inputs.

As resulted, the proposed networks results in better accuracy than general multi layer neural network. In this approach , normalization along with descretization leads to more effective input vectors and improved overall accuracy of the structure. This efficient input results in better performance and speed of convergence of the neural network. Furthermore, it is conclude using pair-wise classifiers improved the accuracy of categorization.

Conclusions

In this paper, we introduced an efficient method to optimize feature vectors obtained from images. We firstly, normalized them to avoid impairment of learning algorithms. Consequently, we descretized the inputs to speed up the convergence and improve accuracy of the system. Followingly, we applied Information Gain formula to select the best features to obtain the most accurate classification. Finally, we proposed a novel method to improve the accuracy of neural network classifiers. We used multiple pair-wise networks to classify the worst features in pairs. Results and experiment showed improved performance of our approach in every step.

References

[1] J. Z. Wang, J. Li and G. Wiederhold, *SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries,* IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(9):947–963, 2001.

[2] Ying Liu, Dengsheng Zhang, Guojun Lu and Wei-Ying Ma, *A survey of content-based image retrieval with high-level semantics* , Journal of Pattern Recognition, 40(2007):262–282.

[3] Wei Wang and Aidong Zhang, *Extracting semantic concepts from images: a decisive feature pattern mining approach* , Multimedia Systems, 11(4):352–366, 2006.

[4] Yu Wang, Mingyue Ding, Chengping Zhou and Tianxu Zhang, *A Hybrid Method for Relevance Feedback in Image Retrieval Using Rough Sets and Neural Networks* , international journal of computational cognition, 3(1):78–87, 2005.

[5] Jun-Hua Han, D. S. Huang, Tat-Ming Lok, and Michael R. Lyu, *A novel image retrieval system based on BP neural network* , IJCNN2005, 4(31):2561–2564, 2005.

[6] H. Young Ko and Suk In YOO: *Intelligent Image Retrieval Using Neural Network* , IEICE TRANS. INF. & SYST, E48-D(12):1810–1819, 2001.

[7] A. Tazaree and K. Faez, *Features Selection to Enforce Contents Based Image Retrieval Systems Accuracy,* Proceedings of CSICC2008, CityplaceKish, country-regionIran, March 9–11, 2008

[8] M. Analoyi and M. Meshki, *A sclable Method For Improving Retrival Systems's Accuracy Using Pairwise Classifers* , First Joint Congress on Fuzzy and Intelligent Systems Ferdowsi University of Mashhad, Iran, (In persian). http://confbank.um.ac.ir/modules/conf_display/isfs2007/article/i357.pdf, 2007.

[9] L.I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*, John Wiley & Sons, New Jersey 2004.

[10] M. K. Hu, *Visual Pattern Recognition by Moment Invariants* , IRE Trans. Info. Theory, IT-8:179–187, 1962.

[11] Ioka M., *A method of defining the similarity of images on the basis of color information.* Technical Report RT-0030, IBM Research, placeCityTokyo Research Laboratory, 1989.

[12] Swain M. J. and Ballard D. H., *Similarity of color images.* Journal of Computer Vision, 7 (1):11-32, 1991.

[13] Funt B. V. and Finlayson G. D., *Color constant color indexing.* Technical report 91-09, PlaceTypeSchool of PlaceNameComputing Science, PlaceNameSimon PlaceNameFraser PlaceNameUniversity, CityVancouver, B.C., placecountry-regionCanada, 1991.

[14] D. Lee, R. Barber, W. Niblack, M. Flickner, J. Hafner and D. Petkovic, *Indexing for complex queries on a query-by-content image database.* Proceedings of 12th International Conference on Pattern Recognition (ICPR'94), placeCity-Jerusalem, country-regionIsrael, 142–146, 1994.

[15] J.R. Smith and S.-F. Chang, *Local color and texture extraction and spatial query*, Proceedings of IEEE International Conference on Image Processing, placeCityLausanne, country-regionSwitzerland, 3:1011-1014, 1996.

[16] G. Pass, and R. Zabih, *Comparing images using joint histograms. ACM Journal of Multimedia Systems,* 7 (3):234–240, 1999.

[17] A. Vailaya, M. Figueiredo, A. Jain, and H. J. Zhang, *A Bayesian framework for semantic classification of outdoor vacation images.* Proceedings of SPIE: Storage and Retrieval for Image and Video Databases VII, placeCitySan Jose, StateCA, 3656:415–426, 1999.

[18] A. Vailaya, M. Figueiredo, A. Jain, and H.J. Zhang, *Content-based hierarchical classification of vacation images.* Proceedings of IEEE Multimedia Systems'99, placeCityFlorence, country-regionItaly, 1:518–523, 1999.

[19] R. O. Stehling, M. A. Nascimento and A. X. Falcao, *On Shapes of colors for content-based image retrieval.* Proceedings of ACM International Conference on Multimedia, placeCityLos Angeles, StateCA, country-regionUSA, 171–174, 2000.

[20] M. Stricker and M. Orengo *Similarity of color images. Proceedings of SPIE Conference on Storage and Retrieval for Image and Video Databases III*, San Jose, CA, 2420:381–392, 1995.

[21] S. Ghosal and R. Mehrota, *A moment-based unified approach to image feature detection.* IEEE Transactions on Image Processing, 6 (6):781-793,1997.

[22] M. K. Mandal, *Wavelet based coding and indexing of images and video.* Ph.D. thesis, CityplaceUniversity of Ottawa, country-regionCanada, 1998.

[23] A. Vailaya, A. Jain and H. J. Zhang, *On image classification: city images vs. landscapes*, Journal of Pattern Recognition, 31(12):1921–1935, 1998.

[24] Q. Li, J. Yang and Z. Zhuang, *Web-based multimedia retrieval: balancing out between common knowledge and personalized views.* Proceedings of 2^{nd} International Conference on Web Information System Engineering, Kyoto, Japan, 92–101, 2001.

[25] J. Li, J. Z. Wang, and G. Wiederhold, *Classification of textured and nontextured images using region segmentation.* Proceedings of Seventh International Conference on Image Processing, CityplaceVancouver, StateBC, country-regionCanada, 2000.

[26] J. Lakshma, *Enhancing retrieval of images on the web through effective use of associated text and semantics from low-level image features.* Ph.D. thesis, School of Computing and Mathematics; College of Health and Science; University of Western Sydney, 2006.

[27] A. Vailaya, Y. Zhong and A. K. Jain, *A Hierarchical system for efficient image retrieval.* Proceedings of 13th International Conference on Pattern Recognition, placeCityVienna, country-regionAustria, C356–360, 1996.

[28] A. Vailaya, *Semantic classification in image databases.* Ph.D. thesis, placeCityMichigan State University, country-regionUSA, 2000.

[29] Y. Rui, A. C. She and T. S. Huang, *Modified Fourier descriptors for shape representation – a practical approach.* Proceedings of First International Workshop on Image Databases and Multimedia Search, CityAmsterdam, The placecountry-regionNetherlands, 1996.

[30] M. Bouet, A. Khenchaf and H. Briand, *Shape representation for image retrieval.* Journal of ACM Multimedia '99, 1–4, 1999.

[31] A. K. Jain and A. Vailaya, *Shape-based retrieval: a case study with trademark image databases.* Pattern Recognition, 31 (9):1369–1390, 1998.

[32] Witten I.H. and Frank E., D*ata Mining: Practical Machine Learning Tools and Techniques with JAVA Implementations*, Morgan Kaufman, CityplaceSan Francisco, 2005.

[33] Tom. Mitchell, Machine Learning, McGraw-Hill, 1997.

[34] B.C. Russell, A.Torralba, K.P. Murphy and W.T.Freeman, *LabelMe: a database and web-based tool for image annotation* , MIT AI Lab Memo AIM-2005-025, 2005. www.labelme.csail.mit.edu