International Journal of Electronics Communication and Computer Engineering
Volume 3, Issue 5, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

Application of Radial Basis Function Network for Content Based Image Retrieval

Chandrasekhar G. Patil, Dr.Mahesh T.Kolte, Dr. Prashant N.Chatur, M.M. Sardeshmukh

Abstract **– The performance of active learning of the Classifier and Optimizer decides the accuracy of Content Based Image Retrieval The Classifier is used to reduce the huge amount of the data base and will offer the semantically similar images whereas the optimizer optimizes the semantic difference between the Query Image and the Retrieved Classified images .The active learning plays an important role in this target oriented Image Retrieval strategy. The Image searching will be performed by comparing the similarity of the features with that of stored in the data base. Absolute, Euclidian distances are few of the parameters that can be used for comparing the similarity between the images. The Radial Basis Function Network Neural Network is used in this approach for classifying the Images and the Image retrieval is designed by using the Absolute distance.**

Keywords **– Image Retrieval, Query image, Database, mean, standard deviation, features, semantic, etc.**

I. INTRODUCTION

The Content Based Image Retrieval Process needs the complete characterization of the Image. The Low level Features from the Image like color, texture and shape are used to fully characterize the Image. The Content in the Image depends on the Human Perception. Therefore it is quite necessary to describe the Image by considering the human perception in order to enhance the compatibility of the machine learning capabilities as compared to the human being.

The sizes of the data base and the techniques used for refining it as per the semantic contents from the Images have great impact on the efficiency of the CBIR system [1]. Higher dimensional indexing improves the speed performance; and feature extraction is the key to accuracy performance.

The Low level features are not sufficient to describe the image similar to that of human visual perception (HVP). To enhance the similarity it is necessary to convert these low level features to semantic based High level features. The accuracy of the CBIR system depends on the gap between the Low level and High level features which is known as the Semantic Gap [2].

Computational time is also one of the governing factors that govern the performance of the CBIR. Therefore it is beneficial to transform low-level characteristics to the high level features.

Many efforts have been taken in improving the relevance feedback algorithms such as query refinement [9, 10], reweighting [11], Bayesian learning [12] and kernel-based learning [13]. From the literature it has been observed that optimization in the semantic gap is still a challenge to the researchers. To design a data base with low dimension and high semantically mapping is the requirement of the good CBIR system [3]. Image representation will be more focused towards the semantic contents by transformation of low level to high level features [4].

This paper proposes a simple relevance feedback algorithm based on Radial Basis Function Network (RBFN). Here RBFN is used not only as classifier but also as feature Transformer. It transforms the low level Features of the Image in semantic based high level features

The paper is organized as follows: the section 2 focuses on the Feature Extraction methodology used for this approach. The Radial Based Function Network (RBFN) along with its characteristics and properties is described in the Section 3. Section 4 discusses the application of RBFN as a Classifier for the semantically based classification of the Images. The overall system operation is described in Section 5. Experimental Results can be viewed in Section 6. Section 7 concludes the paper along with future scope.

II. FEATURE EXTRACTION

The feature of pixels decides the characteristics of the Image. Hence in order to characterize the Image Semantically, it is necessary to exploit the visual parameters of the pixel in multidimensional feature vector. The raw image pixels are converted to a collection of feature vectors in a 10-dimensional feature space. Each image is replaced by a collection of feature vectors in a 10-dimensional feature space before classification. It gives some qualitative descriptions to characterize the principal visual properties of images effectively.

To exploit the visual properties each image pixel is characterized by 10-dimensional visual features and 2-dimensional location descriptors (i.e., x and y).

The 10-dimensional pixel-based visual features include 3-dimensional color descriptors (i.e., R, G, and B), 3-dimensional color deviations of R, G, and B calculated by using 5×5 window, 2-dimensional gradients of luminance channel along horizontal and vertical axis, and 2-dimensions of edge and lable matrix that shows the object oriented connectivity of pixels in the Image. Following the feature extraction strategy, each pixel is represented by a 10-dimensional feature vector, and the image as a whole is represented by a collection of feature vectors in the 10-dimensional feature space. A feature vector (FV) is created for each of the Image from the Data base. Once the data base is ready then next important step is to consider the searching methodology that is to be adopted to search a maximum similar Image to that of the Query Image.

From the research it is known that there are two distinct types of the image search with reference to CBIR.

1. The Objective Search –This searching method uses an absolute/ Euclidian distances between the FVs of the query and images from database. Based on the absolute/ Euclidian distance the Images from database are sorted in increasing order regarding to their distances from a Query Image.

2. The Subjective Search: Here besides the absolute/Euclidian distance the Human Perception is also used for sorting out the most similar Images from the data base. The User is directly involved in the Image Retrieval Process. The user is allowed to refine the sorting process based on his perception. The User decides the Relevancy of the Image that is sorted out from the Objective method and this feedback (Relevance Feedback(RF)) will be considered for further refinement of the final Image so as to optimization in reduction of the semantic gap. [5,6, 10]

The user's relevance feedback is used for improvement in the searching process. The User is exposed to these Images and allowed to select the maximum similar Image and labels subjectively the best-matched candidates. The weights in the Radial Basis Function Network (RBFN) are updated according to the feature vectors of the labeled images. Because of the updating of the weights, the centers and width of the Gaussian function will be modified. The updated RBF model is then used to evaluate subjective similarity in a new search. The process is repeated until the user is satisfied with retrieved results. In practice, several iterative steps are sufficient, as confirmed by intensive simulations.[10,11]

III. The Radial Basis Function Network

Radial basis function (RBF) networks are commonly used for pattern classification. Radial basis function networks typically have two distinct layers as shown in Figure 1. The bottom layer consists of a set of basis functions each of which is a function of the distance between an input pattern and a prototype pattern. A standard choice of basis function is the Gaussian:

$$\text{Ø}_j=\exp\left\{\frac{-\|x-\mu_j\|^2}{2\sigma^2}\right\} \quad \text{......[1]}$$

There are major three techniques for training RBF networks viz

1. Maximum Likelihood,
2. K-means clustering and
3. Gradient descending technique.

First, the structure of standard RBF is presented. The top layer is a simple linear discriminant that outputs a weighted sum of the basic functions. The equation for a single output y_k

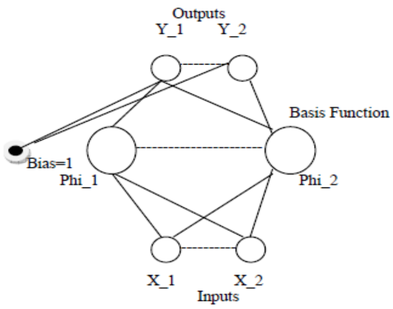


Fig.1. A structure generic RBF network

$$y_k=\sum_{j=1}^{M} w_{kj}\phi_j(x)+w_{k,bias} \quad \text{......[2]}$$

For our networks; we set the weights of the top layer using the least squares error. Rewriting the formula for in matrix notation, equation 2 becomes y(x)=Wϕ. The minimum of the sum squared error

$$E=\frac{1}{2}\sum_n\sum_k\left\{y_k\left(x^n-t_k^n\right)\right\}^2=\phi^T\phi W^T-\phi^T T$$

$$W^T=\phi^\dagger T \quad \text{......[3]}$$

where $\phi^\dagger$ is the pseudo-inverse of $\phi,\phi^\dagger=\left(\phi^T\phi\right)^{-1}\phi^T$.

IV. Radial Basis Function Network as Image Classifier

The architecture of the RBFN is given in Figure 2.It is composed of an input layer, A Gaussian kernel layer and an output layer. The input data will be p-dimensional low-level feature vector. They are connected to the Gaussian kernel layer which is constructed dynamically based on different classes of the Images. The procedure is to cluster samples in each class or categories of the Images.

Let V= {v1, v2,…,vk} be a set of p-dimensional RBF centers. The output for an input vector of an image X is:

$$F(x)=\sum_{i=1}^{k}W_i f(x,v_i,\delta_i)$$
$$=\sum_{i=1}^{k}W_i\exp\left(\frac{-(x-v_i)\Lambda(x-v_i)^T}{2\delta_i^2}\right) \quad \text{......[4]}$$

Wi: Connection weight of the output layer.

V_i, δ_i: center & corresponding width of the ith RBF unit

$$\sigma_i=\gamma.\min_j\|V_i-V_j\|\ j=1,2,.....kj\neq i \quad \text{......[5]}$$

Where Y is a factor which controls the overlapping of different RBF centers. Λ=diag [a1,…,ap] is a diagonal matrix that denotes the relative importance of different feature components determined by positive samples.

4.1 Ffrbfn learning algorithm

The parameters of the network are adjusted in an online error correction procedure. The error function is defined as:

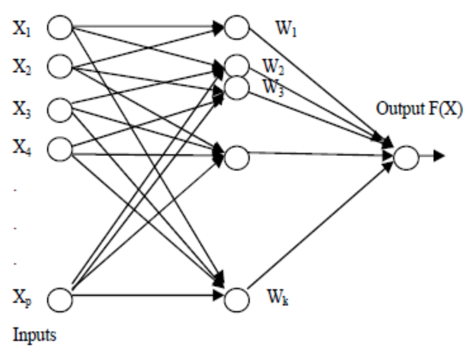


Fig.2. Architecture of Radial Basis Function Network

$$E = \frac{1}{2}\sum_{j=1}^{N} e^2 j = \frac{1}{2}\sum_{j=1}^{N}(Y_j - F(x_j))^2 \quad \text{..........[6]}$$

Where N: Number of total training samples
$F(x_j)$: actual network output for the jth training sample x_j
Yj: desired network output or Target for x_j,.

The Target are set as per the Class of the Images. The Image data base used for the classification is from MIT digital library. Total Five types of Images have been considered for the Classification.1.Coast Images, 2.Forest Images, 3Mounain Images, 4.Open country Images and 5.Tall building Images. For each class of the Images the High Level features have been decided as shown below in Table 1. The general model of the retrieval process is as shown in Fig.3

Sr.No.	Class/Type of the Image	High level Feature output. Target
01	Coast Images	1
02	Forest Images	2
03	Mountain Images	3
04	Open Country Images	4
05	Tall Building Images	5

Table.1. Mapping between the Low level Features and High Level Features of the Images.

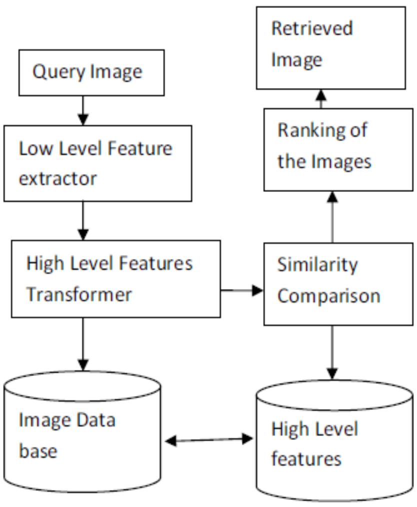


Fig.3. The general model of the Image Retrieval process.

V. System Description

The system has four important phases of the operation viz.

1. The preprocessing operation- Low level Feature Extractor.
2. Low level to high level Feature Transformer/Image Classifier,
3 Retrieval process of the query Image and
4. Displaying the optimistically matched Image to that of the query image.

1. *Preprocessing step:* First step in any CBIR system is finding relevant low-level features j=1,2,…,J (such as color, texture, shape, etc.) describing as best as possible the content of each image i, i=1,2,…,N. Features are expressed by corresponding numerical values, and are grouped into appropriate feature vector **F***i* = [*Fi*1, *Fi*2 ,K, *FiJ*].[11,12] Each coordinate of a vector **F***i* corresponds to particular feature component. FVs were stored in appropriate *feature matrix*, **F**={**F***i*}, of dimension *I*x*J* as described in section 2. The shape is not considered in this research. In general, the terms $F(i,j)$, i=1,2,…,I, j=1,2,…,J, in a feature matrix may have significantly different values since they are obtained in different way. Components with higher variance will be dominant in determining a absolute/Euclidean distance and may produce unfair competition and masking effect. To avoid the dominance of a feature with a large variance and permit the fair influence of all patterns, each term $F(i,j)$ in a feature matrix is column wise rescaled with weighted term $W1j$,

$$W1_j = \frac{1}{mean(F_j)}\log_2(std(\frac{F_j}{mean(F_j)}+2)) \qquad [7]$$

The term *mean* (*Fj*) is the mean value of all elements in *j*-th column of a feature matrix **F** and *std*(.) is the standard deviation. The feature matrix normalized by (8) is then saved and used for calculating the Euclidean distance as a similarity measure between a query and images from database. This weighting strategy, inspired by a text retrieval, was proposed in [5] for image retrieval. A problem of mismatching between an objective measure and subjective sensation of similarity may be resolved by applying weighted Euclidean distances, or by introducing user's relevance feedback (RF). When using weighted distances appropriate weights are associated with feature vector components describing the influence of particular component to a similarity measure. Such a method is applied in systems Photobook [1], QBIC [2], Virage [3], NETRA [4]. Relevance feedback assumes interactive user's assistance in similarity classification. The two main RF strategies in CBIR systems are already applied: a *query shifting* [5] and *distance reweighting* [6, 7].

2. *Low Level to High level features transformer:* This transformer accepts he input vector of low level features F(i,j) and associates with the Target vector T(i,j) as depicted in Table 1. The RBF based neural network is used for training. Once the network is trained, the output will show the five classes of the Images. The multilayer RBF based neural network is found to be suitable for the classification of the Images. These Low level , High level features of the Images are properly stored in the data base.

3. *The Retrieval process of the query Image:* The complete model of the retrieval system is an ordinary CBIR system, where the user inserts the query image Q into the system, and then the low-level features, whose values are used as the input to the RBF neural network, are extracted. When a query image is given to the system, the procedure which took place for the training database images, is also done for the query image. This input vector will be compared with the high-level vectors stored in the database by means of a standard similarity measure like absolute/Euclidean distance. By means of K-nearest algorithm, the class of query image will be determined and that class will be only considered for further processing.

4. *Display of optimistically matched Image:* Once the class of the query image is determined then the search process will be restricted to that class of the Image. By

using standard similarity measure like Euclidean distance, top k relevant images are displayed for feedback from the user.

VI. Simulation Results

An image database from MIT Media Laboratories [17] is used for the simulation. Out of 2000 images, 500 images are selected and manually grouped and properly labeled in to 5 classes. This set of 500 images is used as initial data base. The Images are manually grouped in to the five groups viz. 1.Coast Images, 2.Forest Images, 3Mounain Images, 4.Open country Images and 5.Tall building Images. The advantage of manually aligning the images with known classes lies with the easy analysis of the result. The system performance will also be analyzed easily. The parameters considered for the performance measure are

1. The effectiveness and accuracy of classification and retrieval of the Image. The precision of querying based on the feature vectors extracted; the higher the system's precision, the higher the percentage of matching images returned to a query.
2. The speed of the Classification and Retrieval of the Image .It is desirable to have as lowest as possible the elapsed time for both.

The most favorable condition would be for 100% matching of images, since the idea behind a CBIR system is to retrieve from the database the maximum number of images that match a query image.[13,14] For the evaluation, the following Equation is used for precision.[15,16].

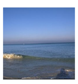
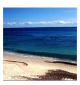
$$\text{Precision} = \frac{\text{Number of Correct Images Returned}}{\text{Number of total Images returned}} \times 100$$

The sample of the retrieval process is shown in Fig.4.The Query Image and the corresponding Images retrieved are shown along with the absolute distance between the Query and retrieved Images.

VII. Conclusion

The Image Classification method, based on RBF neural networks, is used here. It is based on the transformation of Low level Features to High level features. The results show that by using RBF based Neural network, it is possible to accurately retrieve the Query Image from the given data base. But the Response time should also be considered for the further improvement.

Query Image

		
A1-0.00	A2-0.2267	A3-0.3279

Retrieved Images

Fig.4. The sample of the Retrieval Process.

The Precession and the response time are as shown in Table 2

S.No.	Class of the Image	Precision In %	Response time In seconds
01	Coast	100%	23.372030
02	Coast	100%	33.704884
03	Coast	100%	43.355776
04	Forest	100%	53.246911
05	Forest	100%	54.219515
06	Forest	100%	43.295211
07	Mountains	100%	33.191574
08	Mountains	100%	13.247013
09	Mountains	100%	23.358124
10	Open country	100%	33.352995
11	Open country	100%	43.223969
12	Open country	100%	33.180830
13	Tall Buildings	100%	34.231710
14	Tall Buildings	100%	36.218445
15	Tall Buildings	100%	37.232646

Table.2. The Precision and the Response Time of the proposed system.

References

[1] A. Pentland, R. Picard, and S. Sclaroff, "Photobook: Tools for content-based manipulation of image databases", in *Proc SPIE Conf. Storage and Retrieval of Still Image and Video Databases II*, pp 34-47, San Jose, CA, USA, 1994

[2] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, and B. Dom, "Query by image and video content: The QBIC system", *IEEE Computer*, vol. 28, no. 9, pp. 23-32, Sept. 1995

[3] J. R. Back, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Jain, and C. F. Shu, "The virage image search engine: An open framework for image management", in *Proc. SPIE Conf. Storage and Retrieval for Still Image and Video Databases IV*, pp. 76-87, San Jose, CA,USA, February 1996.

[4] W. Y. Ma and B. S. Manjunath, "NETRA: A toolbox for navigating large image databases", in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, 1997, pp. 568–571.

[5] Y. Rui, T. Huang, and S. Mehrotra. "Content-based image retrieval with relevance feedback in MARS", in *Proc. IEEE Int. Conf. on Image Processing*, pp 815–818, 1997.

[6] Y. Rui and T. Huang, "Relevance feedback: A power tool for interactive content-based image retrieval", *IEEE Trans. On Circuits and Systems for Video Technology*, vol. 8, no. 5, pp 644– 655, 1998.

[7] J. Peng, B. Bhanu, and S. Qing, "Probabilistic feature relevance learning for content-based image retrieval", *Computer Vision and Image Understanding*, vol. 75, no. 1/2, pp 150–164, 1999.

[8] R. Brunelli, and O. Mich, "Image retrieval by examples", *IEEE Trans. Multimedia*, vol. 2, no. 3, pp. 164-171, Sept. 2000

[9] D. Zhang, A. Wong, M. Indrawan, and G. Lu, "Content-based image retrieval using Gabor texture features", in *Proc. IEEE Pacific-Rim Conference on Multimedia* (PCM'00), pp.1139-1142, Sydney, Australia, December 13-15 , 2000

[10] P. Muneesawang, and L. Guan, "An interactive approach for CBIR using a network of radial basis functions", *IEEE Trans Multimedia*, vol. 6, no. 5, pp. 703-716, Oct. 2004

[11] K.-M. Lee, and W. N. Street, "Cluster-driven refinement for content based digital image retrieval", *IEEE Trans. Multimedia*, vol. 6, no.6, pp. 817-827, Dec. 2004

[12] G. Liu, "Techniques and data structures for efficient multimedia retrieval based on similarity", *IEEE Trans. Multimedia*, vol. 4, no. 3, pp. 372-384, Sept. 2002.

[13] J. Calic, N. Campbell, A. Calway, M. Mirmehdi, T. Burghardt, S .Hannuna, C. Kong, S. Porter, N. Canagarajah, and D. Bull, “Towards intelligent content based retrieval of wildlife videos”, in *Proc. 6th Int. Workshop WIAMIS-05*, EFPL, Montreux, Switzerland, April 2005.

[14] M. Clausen, and F. Kurth, “A unified approach to contentbased and fault-tolerant music recognition”, *IEEE Trans. Multimedia*, vol. 6, no. 5, pp. 717-731, Oct. 2000

[15] V. Đorđević, N. Reljin, and I. Reljin, “Identifying and retrieving of audio sequences by using wavelet descriptors and neural network with user’s assistance”, in *Proc. Conf. EUROCON- 2005*, Vol. 1, pp. 167-170, Belgrade, 21-24 Nov. 2005

[16] G. Zajić, N. Kojić, V. Radosavljević, S. Čabarkapa, B. Reljin,” Feature vector reduction in CBIR system with relevance feedback”, in *Proc. 13th Inter. Conf. IWSSIP-2006*, pp. 479-482, Budapest, Hungary, 21-23, Sept. 2006.

[17] http://vismod.media.mit.edu/pub/VisTex/

[18] Corel Gallery Magic 65000, http://www.corel.com

AUTHOR’S PROFILE

Mr. Chandrashekhar G. Patil

is currently pursuing his Ph.D.. (Electronics Engineering) in the Department of Electronics and Telecommunication Engineering, Government College of Engineering, Amravati and Sant Gadgebaba Amravati University, Amravati . His areas of interest are Digital Image Processing, Content Based Image Retrieval, Digital and Wireless Communications.

Dr. Mahesh T. Kolte

is currently working as Professor and Head of the Department ,Department of Electronics and Telecommunications Engineering, Maharashtra Institute of Technology, Pune , Maharashtra His areas of interest are Digital Signal Processing Digital Image Processing, Speech Processing.