Journal of Engineering and Applied Sciences 12 (Special Issue 12): 9415-9418, 2017
ISSN: 1816-949X

Content Based Image Retrieval System Based on Classifier Ensemble Approach

[1]Vishal B. Padole and [2]V.R. Mankar
[1]Department of Electronics and Telecommunication Engineering,
Sant Gadge Baba Amtavati University, Amravati, India
[2]Department of Electronics and Communication Engineering,
Government Polytechnic Amravati, Amravati, India

Abstract: We formulate the problem of using an individual classifier for searching the images in the huge database by providing image or its rough sketch as a query. In order to used single classifier for image classification it produce numerous drawbacks, so to avoid this problem a systems comprising of a pool of classifier has to be designed which generates the output which is complement of each other. Though this field has acquired much wide popularity, CBIR has not yet reached its mellowness and is not yet being used on a large scale. In this study, various concepts regarding CBIR like objectives while designing a CBIR system, problems associated while designing an efficient CBIR system, analysis of KNN classifier for this system and a new system overcoming the drawback of using existing methods in terms of computational time and correctness have been discussed.

Key words: CBIR, classifier, ensemble, complement, objectives, terms

INTRODUCTION

Providing textual annotations for the images in large data bases posses difficulty to the user which gives idea that why CBIR is of high interest nowadays. CBIR is a system which gives output depends upon the content of a image to be search (Huurnink *et al*., 2012; Swain and Ballard, 1991). While designing any CBIR System the most common objectives for efficient output (correct image) are:

Finding the images (output) that shares precise characteristics the user has in mind. Providing the output images which closely match to the query image sketch. The search engine must permit for fairly unclear outputs and it should not produce the output that contains all the related images and it should not discard all the irrelevant images because we have to pick one of the images as a output from itself. Relevant images provided as output by the search engine should must contain few images out of entire database (Chen *et al*., 2005). Retrieval of images (Query) from the huge database can be divided into following categories:

- Labeling the images by using texts
- Based on visual characteristics of the image content
- Based on the image semantic features

Retrieval of images by labeling using texts includes some drawbacks that it changes the retrieval system into conventional keywords and it also offers too heavy load and the output remains unclear. In second method, accuracy of the output depends solely on the quality of extracted features. In third method, the image retrieval that uses rough sketch of required image as a keyword is more superior than that of conventional text based. But the challenge is difference exists in mathematical feature extraction model on human beings point of view.

Problems associated while designing an efficient CBIR System: On the basis of various techniques related to the designing of CBIR System following problems have been revealed:

- Problem of difference in color visual scence that every human possessed which creates problem while analyzing color images
- If texture chosen of each method is not defining the complete information about image it will creates problem in text based search
- Automatic selection of shape edge creates difficulty in retrieval using shape features

Composition of a CBIR System: CBIR involves the following four parts in system realization:

Corresponding Author: V.R. Mankar, Department of Electronics and Telecommunication Engineering,
Sant Gadge Baba Amtavati University, Amravati, India

Database: It is the set of all images among which we have to search for required image or image in which we are interested.

Feature extraction: It is the key of any CBIR System because whatever the comparison is done between query image and images from database is purely on the basis of features extracted from the images (Wardhani and Thomson, 2004).

Relevance feedback: Relevance is a similarity aspect amongst the images which is shared by other images. Ranking is one of the serious issue while sorting the images and the solution is to arrange the images before presenting to the user in decreasing order of similarity. But for the user it is not possible to choose from two alternatives namely relevant and irrelevant, therefore, images should not classified as a relevant and irrelevant in spite of this search engine should only rank most relevant images first and then irrelevant. Relevance is the process of including user in the CBIR system to increase the preciseness of output.

CBIR System: This is the core of the search engine there are many ways of designing CBIR System on the basis of different feature extracting algorithms in this study, we mostly focused on the study of classifiers and a designing of CBIR System using one or more classifiers which generate effective results in constraints with the computational time and exact image output. In this study, KNN (K Nearest Neighbor) algorithm is discuss along with a new proposed system to overcome the drawbacks of individual classifier. Figure 1 depicts A CBIR System by using classification approach (Shah *et al*., 2004).

An easy way to comply with the conference study formatting requirements is to use this document as a template and simply type your text into it.

K nearest neighbors algorithm: KNN algorithm simply uses the stored information of all existing cases and yields new results based on similarity measures. k-nearest neighbor is falls under the category of supervised learning algorithm in which the result of new image query is classified on the basis of maximum of k-nearest neighbor category. The ultimate aim of KNN algorithm is to categories a new image based on the training examples. With respect to the given query this algorithm finds K number of samples nearest to the input image (Singh and Rajput, 2015). Classification of the images is done by taking into consideration the maximum votes while classifying the k number of objects with the facility of breaking ties randomly. This algorithm use nearby samples categorization as the probable guess value of the new input image:

- K Nearest Neighbors (KNN) is to use database in which the data points are separated into several separate classes to predict the classification of new points
- Input is consisting of k nearest perception examples in the feature space
- Output is class membership
- Maximum number of votes by neighbors will classify the object

As shown in Fig. 2, if we consider green circle as a test sample which we have to classified into either to the class of blue squares or to the red triangle. Solid line circle gives result for K = 3, it assigns to the class of red triangles because they are majority in count and for K = 5 doted line circle assigns to the class of blue squares as they are majority in numbers.

Experimental results for KNN algorithm: For testing KNN algorithms effectiveness while classifying the

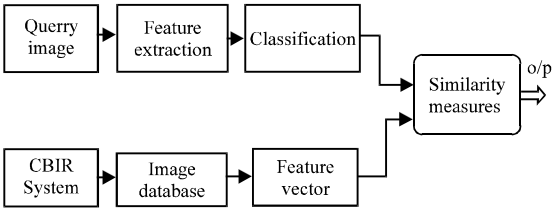

Fig. 1: Generalized block diagram of CBIR system

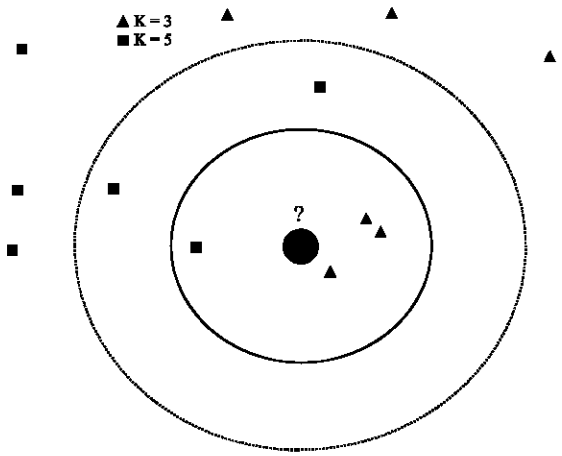

Fig. 2: An example of KNN

images A database containing the images of baby, elephant, dog and flowers is taken and original image is given for testing as a query image it is found that this algorithm searches relevant images in some cases like if dogs image is given as a input it gives similar, that is other dogs image as a output from the combine database but in some cases It also gives irrelevant result. In order to overcome the drawback of KNN or any single classifier the proposed solution is discussed in next study.

MATERIALS AND METHODS

In this research, it is proposed to develop a robust and accurate CBIR system which comprises of fusion of multiple CBIR Systems. In order to overcome the drawbacks and limitations of individual systems, individual systems need to be device and combined in such a way the fused system will be more robust than the individual systems. The proposed approach formulates:

- A probabilistic framework of fusing individual CBIR system and
- It tries to utilize the complementary information of individual systems

CBIR System should be unlike (complementary) from each other that is they complement each other and avoids redundant outputs because redundant images can cause the problem during fusion. The proposed method creates a pool of individual system by extracting various distinct features which captures color, texture and geometry features and by applying different feature matching techniques (Rui *et al*., 1998). The propose reserach tries to apply application orientation fusion of systems, since for different application different specific features would be fused. The selection of systems from the pool of individual systems would be performed based on probabilistic framework of fusion. The different application on which the proposed research is tested upon are: OCR (Optical Character Recognition), logo identification, face detection.

As time is very important aspect while going for searching an image in the huge database we also tries different strategies to speed-up the search space by applying unsupervised data clustering and by utilizing the relevance feedback. The relevance feedback helps in fast convergence of the query search. Figure 3 shows the block diagram of proposed work.

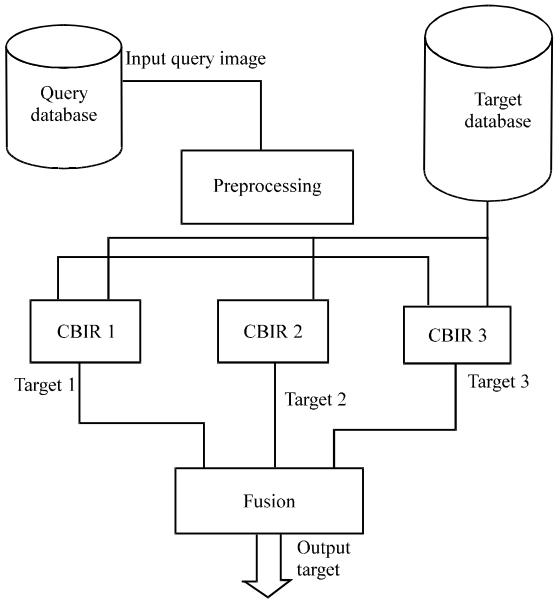

Fig. 3: Block diagram of proposed method showing CBIR System using pool of individual CBIR Systems

RESULTS AND DISCUSSION

A lot of possible applications for CBIR technology has been detected, it will help in:

- Face detection
- Logo identification
- Optical character recognition
- Military applications

The proposed research will reveal better insight into satisfactory interpreter of statistical aspects of image retrieval. Computational time should be considerably reduced due to use of unsupervised clustering and utilization of relevance feedback.

CONCLUSION

Due to the complex features present in the query images and limitations of feature extraction and matching algorithms these methods cannot produce effective results in constraints with the computational time and exact image output. KNN classifier for image classification with its analysis on combine image database is discussed in this study along with the designing of new system by creating a pool of individual system by extracting various distinct features which captures color, texture and geometry features and by applying different feature matching techniques can produce effective results.

REFERENCES

Chen, Y., J.Z. Wang and R. Krovetz, 2005. CLUE: Cluster-based retrieval of images by unsupervised learning. IEEE. Trans. Image Proc., 14: 1187-1201.

Huurnink, B., C.G. Snoek, D.M. Rijke and A.W. Smeulders, 2012. Content-based analysis improves audiovisual archive retrieval. IEEE. Trans. Multimedia, 14: 1166-1178.

Rui, Y., T.S. Huang, M. Ortega and S. Mehrotra, 1998. Relevance feedback: A power tool for interactive content-based image retrieval. IEEE Trans. Circ. Syst. Video Technol., 8: 644-655.

Shah, B., V. Raghavan and P. Dhatric, 2004. Efficient and effective content-based image retrieval using space transformation. Multimedia Proceedings of the 10th International Conference on Modelling, January 5-7, 2004, IEEE, Brisbane, Queensland, Australia, ISBN:0-7695-2084-7, pp: 279-284.

Singh, S. and E.R. Rajput, 2015. Content based image retrieval using SVM, NN and KNN classification. Intl. J. Adv. Res. Comput. Commun. Eng., 4: 549-552.

Swain, M.J. and D.H. Ballard, 1991. Color indexing. Int. J. Comput. Vision, 7: 11-32.

Wardhani, A. and T. Thomson, 2004. Content based image retrieval using category-based indexing. Proceedings of the IEEE International Conference on Multimedia and Expo, ICME'04 Vol. 2, June 27-30, 2004, IEEE, Taipei, Taiwan, ISBN:0-7803-8603-5, pp: 783-786.