National Conference on "Recent Trends in Mathematical Modeling, Simulation Methods, Computations, and Physical Sciences"
(MMSCP-2019), Organised by Hislop College, Nagpur, Maharashtra, India
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2019 IJSRCSEIT | Volume 4 | Issue 4 | ISSN : 2456-3307

An Optimized Classification using Color histogram and CNN for Content Based Image Retrieval

Mehta Priyanka Hemantkumar[1], Dr. Ashish Adholiya[2]

[1]Research Scholar, Pacific Academy of higher Education and Research University, Udaipur, Rajasthan, India
[2]Assistant Professor of IT of Marketing, Pacific Academy of higher Education and Research University, Udaipur, Rajasthan, India

#Corresponding Author Email: sbkondawar@yahoo.co.in

*Presenter Author Email : kanchanutanmangate@gmail.com

ABSTRACT

Presently, in the digital globe the use of various hand-held devices, different storage devices, social networking websites and high network bandwidth have made it possible to store large number of images on the web. Since the size of image database is growing gradually, it necessitates the need of a robust image retrieval system. This need has given rise to the researches in the domain of Content Based Image Retrieval System. In this paper, color histogram is used for feature extraction. On this extracted feature, CNN classification technique is applied to retrieve relevant images from the image dataset.

Keywords : Image Retrieval, Color histogram, Classification technique, CNN

I. INTRODUCTION

With the immense use of information technology, images are widely used in majority of the areas of individual's life such as commecre, education, hospital, government, crime prevention, fashion industry, architecture, engineering and so on. As a result large number of images are stored in the image database. Nevertheless only the storage of the large number of images is not adequate since apart from the storage proficient browsing, searching and retrieval of those images should also be possible. For this reason Content Based Image Retrieval has become an interesting research area.

The main objective of Content Based Image Retrieval System is to provide the technique for searching images in the image database and retrieval of similar images very quickly and efficiently. At the very first stage every CBIR system extract visual content from the images. This visual content offers base for CBIR system.

II. Feature Extraction

Digital images are accumulated in the database in the form of 2-dimensional images. Therefore the main problem is involved in extracting meaningful detail from these row data. This process is known as the feature extraction. The feature extraction process extracts the features like color, texture, shape. Here the discussion involves color feature extraction.

1.1 Color feature

The most commonly used visual feature is the color feature. Each and every color is represented by combining the three features red (R), green (g) and blue (B). Color information can be represented in the form of color histograms. A color histogram is one type of bar graph. In this bar graph each bar is representing specified color of the color space which is going to be used. RGB, HSV, YIQ, CMY, YCbCr are some of the color spaces used in CBIR.

1.2 Image Classification

Image classification is used for capturing useful information classes from a given image. Image classification observes the numerical properties of image features of different images and classifies them into various categories. Deep learning neural networks play an important role in image classification. CNN (Convolution Neural Network) is one class of deep learning neural networks.

III. Related Work

Stricker and Orengo [1] introduced an approach using the color moments that overcome the quantization effects that arise in case of color histogram. This approach is able to describe any color distribution by its moments. First moment (mean), second moment (variance) and third moment (skewness) were extracted to represent color feature. They used Weighted Euclidean Distance to calculate the color similarity.

Another approach introduced by H. B. Kekre and Dhirendra Mishra [2] generated feature vector with the help of Walsh transform for content based image retrieval. They multiply Sal functions by j and combine them with Cal functions of the same sequence by visualizing the complex Walsh transform. They calculate angle by taking tan inverse of Sal / Cal between the ranges of 0 to 360 degrees, divided into 4 sectors. The feature vector of 24 components was designed by considering the mean of real and imaginary values of the 4 sectors in all three color planes. The Walsh transform of the color image was calculated in all three R, G and B planes. The experiment was implemented on the database containing 270 images belonging to 11 different classes. For matching similarity, Euclidean distance was used. Average precision and recall is also calculated for the performance evaluation. The overall average of precision and recall is above 50%.

Dr. Sanjay Silakari, Dr. Mahesh Motwani and Manish Maheshwari [3] described splitting of image into Red, Green and Blue components. They calculated average for each component and split every component image to obtain RH, RL, GH, GL, BH and BL images. The RH is obtained by taking only red component of all pixels in the image which are above red average and RL is obtained by taking only red component of all pixels in the image which are below red average. Similarly the GH is obtained by taking only green component of all pixels in the image which are above green average and GL is obtained by taking only green component of all pixels in the image which are below green average and BH is obtained by taking only blue component of all pixels in the image which are above blue average and BL is obtained by taking only blue component of all pixels in the image which are below blue average. Then the Color moment was applied to each divided component i.e. RH, RL, GH, GL, BH and BL and was classified using K-means clustering algorithm.

Ch. Kavitha , M. Babu Rao , Dr. B. Prabhakara Rao , Dr. A. Govardhan,Image [4] anticipated a method that the query image and the images in database were alienated into 6 equal blocks. HSV colour histogram assembled for each block and statistic features such as Energy, Contrast, Entropy and inverse difference acquired from GLCM and builds a combined feature

vector for color and texture. Normalized Euclidean distance was used between the feature vector of query image and target images to attain the similar images from the database. First 20 similar images retrieved with minimum Euclidean distance from their system.

S.Arivazhagan, L.Ganesan, S.Selvanidhyananthan [5] distinguish the edge points of the image by using canny operator. The shape of the image was mark out by scanning the edge image and to keep away from discontinuities in representing the shape, resampling was made. The resultant image was removed line by line. To distinguish the number of adjacent points and to obtain the shape features, the neighboring pixels were discovered.

Hiremath and Pujari [6] proposed CBIR system based on the color, texture and shape features by partitioning the image into tiles. The features computed on tiles serve as local descriptors of color and texture features. The color and texture analysis are analyzed by using two level grid frameworks and the shape feature is used by using Gradient Vector Flow. The comparison of experimental result of proposed method with other system [7]-[10] found that, their proposed retrieval system gives better performance than the others.

Gudivada (2010) [11] discussed an approach to improve retrieval effectiveness via relevance feedback in text retrieval systems. He also showed how these relevance feedback techniques have been adopted to CBIR context and their effect on retrieval effectiveness. The need for test collections in advancing CBIR research is also discussed in his work.

IV. The Proposed Approach

This section describes the proposed algorithm for image retrieval which is based on color histogram and CNN.

Optimized classification using color histogram and CNN

Step 1: Upload dataset and divide into N categories.
Step 2: Perform image pre-processing like filtering, smoothing.
Step 3: Extract image features using color histogram.
Step 4: Generate normalized image.
Step 5: Optimize the image with SDG.
Step 6: Build CNN module, train module and label each image of dataset.
Step 7: Probability matrix is calculated. On the bases of the probability the fully connected
node of the CNN is classified.
Step 8: Classify the object.
Step 9: Display images of particular category.

V. Experimental Results

The system is tested on a general-purpose set of Animal Image Dataset which is downloaded from https://www.kaggle.com/ashishsaxena2209/animal-image-datasetdog-cat-and-panda. The database contains 3000 images grouped into 3 different categories Dog, Cat and Panda where each category contains 1000 images. Figure 6.1 exhibit cat as query image and the similar images retrieved from the dataset.

Query Image

Retrieved Images

Figure 6.1: results of query image cat

Table 6.1 demonstrates performance of the system with the help of different parameters such as Precision, Recall and F_1-score.

Category	Precision(%)	Recall(%)	F_1-score(%)
Cat	100	92	96
Dog	87	100	93
Panda	87	100	93
Others	100	73	84
Average	93	91	92

Table 6.1 Average Precision / Recall for different categories

VI. CONCLUSION

In this paper, an optimized classification using color histogram and CNN method has been proposed. This proposed method provides 92% accuracy as depicted by F1-score. The accuracy provided by this method sounds better as compared to the other techniques reviewed before.

VII. REFERENCES

[1] M. Stricker, and M. Orengo, "Similarity of color images", SPIE Storage and Retrieval for Image and Video Databases III, vol. 2185, pp.381-392, Feb. 1995.

[2] H. B. Kekre, Dhirendra Mishra, ," Four Walsh Transform Sectors Feature Vectors for Image Retrieval from Image Databases", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 1 (2) , pp.33-37, 2010.

[3] Dr. Sanjay Silakari, Dr. Mahesh Motwani and Manish Maheshwari, "Color Image Clustering using Block Truncation Algorithm", IJCSI International Journal of Computer Science Issues, Vol. 4, No. 2, 2009.

[4] Ch. Kavitha , M. Babu Rao , Dr. B. Prabhakara Rao , Dr. A. Govardhan,Image, "Retrieval based on Local Histogram and Texture Features", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (2) , p.741-746, 2011.

[5] S.Arivazhagan, L.Ganesan, S. Selvanidhyananthan , "Image Retrieval using Shape Feature", International Journal of Imaging Science and Engineering (IJISE), Vol.1,No.3, July 2007.

[6] P. S. Hiremath and J. Pujari, "Content Based Image Retrieval based on Color, Texture and Shape features using Image and its complement", 15th International Conference on Advance Computing and Communications. IEEE. 2007.

[7] Y. Chen and J. Z. Wang, "A Region-Based Fuzzy Feature Matching Approach to Content Based Image Retrieval", IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 24, No.9, pp. 1252-1267, 2002.

[8] J. Li, J.Z. Wang and G. Wiederhold, "IRM: Integrated Region Matching for Image Retrieval", InProceeding of the 8th ACM International Conference on Multimedia, pp-147-156, Oct. 2000.

[9] M. Banerjee, M. K. Kundu and P. K. Das, "Image Retrieval with Visually Prominent

Features usingFuzzy set theoretic Evaluation", ICVGIP, 2004.

[10] Y. Rubner, L. J. Guibas and C. Tomasi, "The earth mover's distance, multidimensional scaling, and color-based image retrieval", Proceedings of DARPA Image understanding Workshop. Pp- 661-668,1997.

[11] Gudivada, V.N., 2010. "Relevance feedback in content based image retrieval". Marshall University.