Holistic Correlation of Color Models, Color Features and Distance Metrics on Content-Based Image Retrieval

Swapnil Saurav[1], Prajakta Belsare[2], Siddhartha Sarkar [3]

[1]Researcher, Abhidheya Labs and Knowledge Solutions, Maharashtra, India
[2]Researcher, Abhidheya Labs and Knowledge Solutions, Maharashtra, India
[3]Researcher, Abhidheya Labs and Knowledge Solutions, Maharashtra, India

---***---

Abstract - ***An important problem in computer vision is to choose features for image retrieval. To address this problem, we have conducted several experiments, considering combinations of features. The experiment involves the application of various color models and computation (extraction) of the features using different color descriptors like color histogram, color moments and color coherence vectors. The study involves evaluating their performance for Content-Based Image Retrieval. It also presents a holistic representation of comparative effects in correlation with the diversity of parameters chosen. As a means to evaluate these, we have chosen a standard data set of 100 images as the retrieval library. The results are embellished with computed values of precision and recall to give further insights.***

Key Words: *CBIR, color coherence vector, color moments, Euclidean distance, precision, Quadratic distance, recall*

1. INTRODUCTION

Content-based image retrieval is gaining importance with an increase in the volume of the image database. Many researchers are working on CBIR to make this technique fast, accurate and efficient. In order to perform Content-based Image Retrieval color, shape and texture are the most important parameters to be considered. The selection of color model, features to be extracted and distance metrics to be used are crucial parameters of CBIR system.

There have been several systems and techniques developed in both the academic and commercial domains such as the IBM's query by image content (QBIC) system [1], Virage's VIR engine [2], VisualSEEk [3], and PhotoBook [4]. A dynamic user concept pattern learning framework for content-based image retrieval is designed that incorporates multiple instance learning for relevant feedback to discover users concept patterns—especially in the region of greatest user interest [5]. In practice, content-based image retrieval searches the image database by visual content, such as color, texture or shape. The pre-processing of images mainly focuses on extracting features that can represent the relevant visual content. When users perform retrieval, features of the query and database images are compared. Those images that have the most similar features are retrieved. Therefore, feature selection is a key stage in content-based retrieval [6]. This motivated the goal for designing a combined system which will help one to select features that are effective for image retrieval. Without the correct choice of features, the results will become irrelevant and the system efficiency will be affected.

We have tried to bridge the gap between a novice and the vast world of CBIR and also clear the differences between the various parameters selected and results hence obtained.

2. PROPOSED METHODOLOGY

1. Color Space Selection

The results of image retrieval methods are different in different color spaces. It is important to select appropriate color space, which is the most effective regarding image retrieval. In this experiment, we are dealing with 11 color spaces to find out a suitable one. In result section, we have discussed the results obtained by using different color spaces and which color space will be appropriate for retrieval.

2. Feature Extraction Methods

a) Color Moments:

Color Moments can be used to differentiate between images on the basis of their color features. It provides a measurement of similarity between images. On the basis of similarity comparison, image retrieval can be done. The color distribution in an image can be interpreted as a probability distribution. If a color in an image follows a certain probability distribution, the moments of that distribution can be used as a feature to identify the image on the basis of color. Three color moments namely mean, standard deviation and skewness are used in this experiment.

b) Color Histogram:

Color histogram characterizes an image by the representation of the distribution of its color. Basic properties of an image can be obtained from the histogram. The histogram of an image will remain unchanged with the small changes in the camera viewpoint. The shape of the histogram will remain the same for a similar type of objects (even if they are of different color) and different objects will have distinctive histograms. Based on this property of color histograms, it can be used as one of the features for image retrieval.

c) Color Coherence Vector:

Color coherence vector contains more significant information about the spatial distribution of colors. It classifies each pixel as coherent or incoherent. A coherent pixel is part of a large group of pixels of the same color while an incoherent pixel is not. We determine the pixel groups by computing connected components. If the same color group contains more coherent pixels than a threshold value defined beforehand, it belongs to coherence pixels. The rest pixels are incoherent. Color coherent vector gives finer details than color histograms.

3. Distance Metrics

Distance metrics is used to determine the distance between pixels. The distance transforms measures the distance of each object point from the nearest boundary [1]. Different techniques can be used to find distance metrics based on the application. We have used two different techniques for calculating distance as follows:

a) Euclidean Distance
b) Quadratic Distance

4. Classifier Used

The classifier used for image retrieval is the support vector machine (SVM). In image retrieval, images are retrieved from the database based on the relevance of images with the query image. Given a query image, first the features are extracted in specified color space. Then the extracted features are compared with the features of images in the database and most relevant images are retrieved. The support vector machine is used to classify images based on the relevance. The kernel function used for image classification is radial basis function.

3. EXPERIMENTS

The images used in this experiment are taken from Wang's database (standard database for CBIR). This database contains 1000 images of 10 categories namely African, Food, Beaches, Dinosaur, Roses, etc. All these images are first classified into three major groups as Manmade Indoor, Manmade Outdoor, and Natural. One image from each ten categories is given as a query image for image retrieval, 100 relevant images are retrieved for each query image. The overall architecture of the image retrieval system is depicted in Fig. 1, 2, and 3. In the following section, each step is explained in detail:

Selection of color spaces:

The color space selection plays a significant role when image retrieval is done on the basis of color features. Choice of appropriate color space is a crucial task. We have used 11 color models *(RGB, I1I2I3, YIQ, HSV, HSI, YUV, LAB, XYZ, CMYk, YCbCr and HMMD).* The RGB image is first converted into one of the color spaces then the feature exaction is done.

Feature Extraction:

Color Moments (CM), Color Coherence Vector (CCV) and Color Histograms are used as features for similarity comparison among query image and database images. To identify the most accurate feature for retrieval, the system is tested for each feature considered separately as well as their combination.

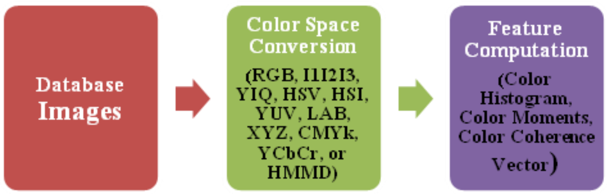


Fig -1: Color Space Selection and Feature Extraction

System Training:

The classifier used for classification purpose is SVM (Support Vector Machine) with Radial Basis kernel. The performance of linear, quadratic and polynomial is also tested. The radial basis kernel provides high performance. The features extracted and targets are fed to the SVM for training and SVM structures are formed according to the major classes mention earlier.

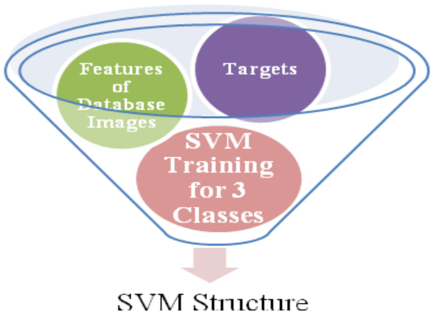


Fig -2: SVM Training of database for three classes

System Testing:

The query image is fed to the system; first the color space conversion is done. The features are extracted and the features are compared to the features of database images. Euclidean distance metrics is used to retrieve the most relevant images. Euclidean distance between the query image and each database image is calculated. The images with less distance are ranked at the top. Top 100 images are retrieved out these ranked images.

Fig -3: Content-Based Image Retrieval

The performance of the system is evaluated on the basis various parameters. The parameters for comparison are: color models *(RGB, I1I2I3, YIQ, HSV, HSI, YUV, LAB, XYZ, CMYk, YCbCr and HMMD)*, feature extraction methods (CCV, CM, CH, CCV-CM, CCV-CH, CH-CM, CCV-CH-CM), and distance metrics (Euclidean distance, Quadratic Distance). The precision and recall is computed for every possible combination and based on that conclusion is drawn. The mean average precision of CCV-CM and CCV-CM-CH are compared for 11 color models.

3. RESULTS

We have evaluated our experiments on Wang's dataset that contains 10 scene categories with manmade, natural, indoor and outdoor environment. It contains 100 images of each scene category, 1000 images in total. We have conducted three sets of experiments (1) comparison of 11 color spaces, (2) comparison of various combinations of three color features i.e. CCV, CH, CM, CCV-CH, CCV-CM, CH-CM, CCV-CH-CM, (3) Euclidean distance and Quadratic distance metrics. SVM classifier is used to classify the images into different categories as Manmade Indoor, Manmade Outdoor, and Natural. One image from each scene category is given as a query to retrieve relevant images. The query is first categorized whether it is Manmade Indoor, Manmade Outdoor or Natural and then relevant images are retrieved accordingly. Very few results are presented in this paper because of space constraints.

Fig -4: Query Image (Manmade Indoor)

Fig -5: **Category:** Manmade Indoor, **Sub-category:** Food, **Color Space:** HSV, **Features Extracted:** CCV-CH-CM, **Distance Metrics:** Euclidean Distance, **Number of Images Retrieved: 37**

Fig -6: **Category:** Manmade Indoor, **Sub-category:** Food, **Color Space:** HSV, **Features Extracted:** CCV-CH-CM, **Distance Metrics:** Quadratic Distance, **Number of Images Retrieved:** 12

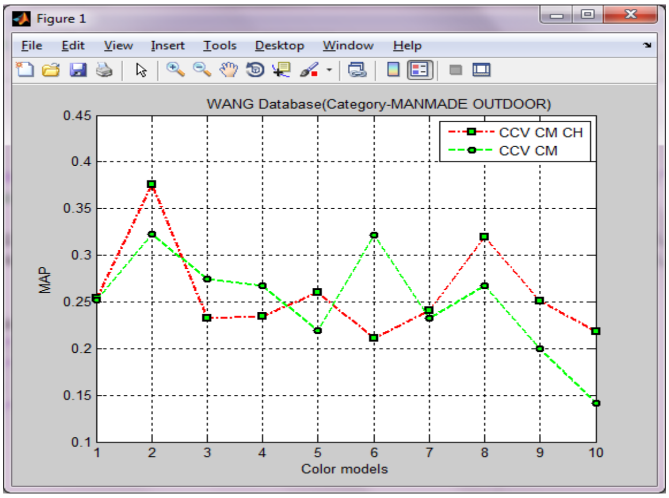
Fig. 5 and Fig. 6 shows the result for a query image of food, with Euclidean distance metrics and Quadratic distance metrics respectively, all other parameters are same for both the retrieval. Using Euclidean distance metrics will retrieve more number of relevant images as compared with Quadratic distance metrics. Similar results are obtained for other query images as well.

Fig -7: **Category:** Manmade Indoor, **Sub-category:** Food, **Color Space:** HSV, **Features Extracted:** CCV, **Distance Metrics:** Euclidean Distance, **Number of Images Retrieved:** 21

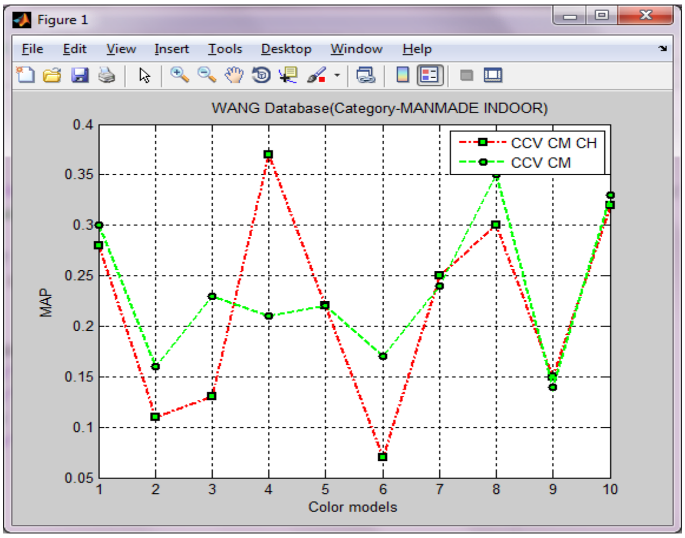
Fig. 5 shows results using CCV-CH-CM color features whereas Fig. 7 shows results for CCV color feature. When CCV-CM-CH color features are used more number of images is retrieved wherein using CCV color feature less number of images is retrieved with few irrelevant images. Use of other combination of color features has led to similar results.

Mean Average Precision Comparison

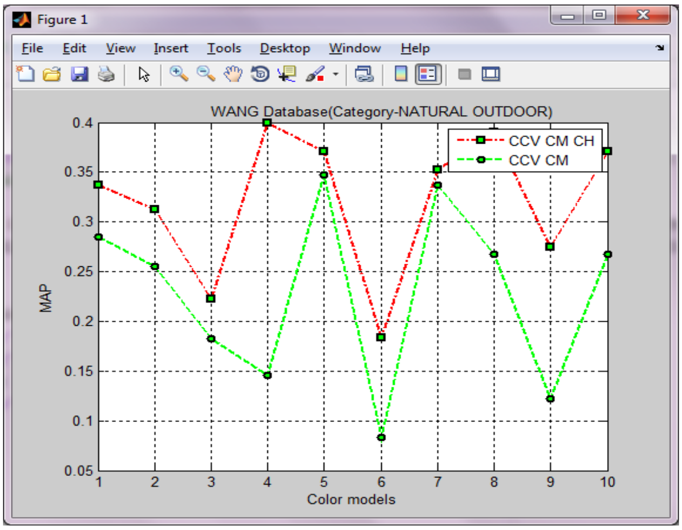
In order to decide which color space gives best results Mean Average Precision (MAP) is calculated for three categories i.e. Manmade Outdoor, Manmade Indoor, and Natural images. Graph 1, 2 and 3 shows a comparison of color features i.e. CCV-CM-CH and CCV-CH based on MAP for each color space. Based on this comparison the best color space for each sub-category is found. Table 1 and Table 2 shows the best color space for each sub-category using CCV-CH-CM and CCV-CH respectively. Table 1 & 2 also describes Maximum Recall and Precision at Maximum Recall value for best color space.



Graph -1: Manmade Outdoor



Graph -2: Manmade Indoor



Graph -3: Natural

Table -1: CCV_CM_CH - Best Color Space for each Category

Category	Best Color Space	Precision at Max Recall	Max Recall
Buildings	HSI	0.954545	0.42
Buses	I1I2I3	0.805194	0.62
African	CMYK	0.884057	0.61
Food	HSV	1	0.37
Beaches	RGB	0.585714	0.41
Dinosaur	CMYK	1	0.99
Elephants	HSI	0.483516	0.44
Roses	XYZ	0.779220	0.6
Horses	HSI	0.556818	0.49
Snowy Mountains	I1I2I3	0.796875	0.51

Table -2: CCV_CM - Best Color Space for each Category

Category	Best Color Space	Precision at Max Recall	Max Recall
Buildings	HSI	0.955556	0.43
Buses	I1I2I3	0.771429	0.54
African	I1I2I3	0.75	0.51
Food	CMYK	1	0.35
Beaches	YUV	1	0.81
Dinosaur	CMYK	1	0.98
Elephants	YUV	1	0.71
Roses	XYZ	0.828947	0.63
Horses	YIQ	0.715789	0.68
Snowy Mountains	I1I2I3	0.647059	0.44

3. CONCLUSIONS

In this paper, we represented a CBIR system that uses a query image and retrieves relevant images. The results show that HSV is the appropriate color space for color feature extraction. It gives good results as compared to other color spaces. As results shows, choice of color features has an impact on image retrieval. When all the three color features are used together, the results are more relevant. Furthermore, the number of images retrieved is also dependent on the distance metrics used. The results of the whole experiment show that choice of color space should be HSV and the distance metrics should be Euclidean. For CCV-CM-CH and CCV-CM features the results almost similar. Since color histogram does not make much difference, CCV-CM-CH and CCV-CM can be used interchangeably. Although these proposed system parameters yield good results but it still can be improved by adding texture features along with shape features. Further work can also be done using color, texture, and shape features for different standard databases. Also, some interesting applications can be developed such as the use of CBIR in the medical field, localization of robot, etc.

REFERENCES

[1] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "*Query by image and video content: The QBIC system,*" IEEE Computer, vol. 28, no. 9, pp. 23–31, Sep. 1995.

[2] The Virage website. [Online]. Available: http://www.virage.com

[3] J. R. Smith and S. F. Chang, *"Visualseek: A fully automated content-based image query system,"* in *Proc. ACM Int. Conf. Multimedia*, Boston, MA, 1996, pp. 87–98.

[4] A. Pentland, R.W. Picard, and S. Sclaroff, *"Solving the multiple-instance learning problem: A lazy learning approach,"* in Proc. Storage and Retrieval for Image and Video Databases II, SPIE-Int. Soc. Opt. Eng., vol. 2185, 1994, pp. 34–47.

[5] S. Chen, S. H. Rubin, M. Shyu, C. Zhang, *"A Dynamic User Concept Pattern Learning Framework for Content-Based Image Retrieval," IEEE Transactions on Systems, Man, And Cybernetics—Part C: Applications and Reviews, Vol. 36, No. 6, November 2006*, pp. 772-783.

[6] M. Jian, J. Dong, R. Tang, *"Combining Color, Texture and Region with Objects of User's Interest for Content-Based Image Retrieval"*, Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2007 IEEE, pp. 764-769.