Investigation of Fuzzy Similarity Measure for Retrieval of Color Images

Poornima N[1], Deivasikamani S M[2], Natrayan M[3], Dineshkumar R[4]

[1]Assistant Professor, Department of Science and Humanities (Mathematics), Sri Ranganathar Institute of Engineering and Technology, Coimbatore, India

[2, 3, 4]Assistant Professor, Department of Mechanical Engineering, Sri Ranganathar Institute of Engineering and Technology Coimbatore, India

Abstract- **Image retrieval has important practical applications in database management, medical, and computer applications etc. For the purpose of effectively retrieving more similar images from the digital image databases, this work uses the color distributions and the mean value to represent the global characteristics of the image. The comparison of image retrieval result between RGB feature vector and hue feature vector in token of image characteristics is taken. The similarity measure is calculated by hue feature vector which is not sensitive to elimination and saturation change and the size of hue vector is reduced on the premise that the retrieval result is not affected. As the experimental results indicated, the proposed technique indeed outperforms other schemes.**

Keywords: Pattern Recognition, Image Retrieval, Similarity Measure.

I. INTRODUCTION

Color image retrieval and classification are very important in the field of image processing. Image retrieval methods are mainly based on color, texture and shape of image. This chapter gives the concept of pattern recognition and its applications in color image and the need for color image retrieval using fuzzy system. Pattern recognition is concerned with the classification of objects into categories, especially by machine. Image analysis deals with the processing and analysis of images. Many pattern recognition systems are designed to classify or analyze images. Pattern recognition is concerned with the automatic detection or classification of objects or events.

A pattern is an entity, vaguely defined, that could be given a name, e.g., fingerprint image, handwritten word, human face, speech signal, DNA sequence [3]. Pattern recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest, make sound and reasonable decisions about the categories of the patterns.

The measurements or properties used to classify the objects are called features and types or categories into which they are classified are called classes. Most pattern recognition tasks are first done by humans and automated later it can be defined as a process of identifying structure in data by comparisons to known structure, the known structure is developed through methods of classification [3]. Two of the main forms of pattern recognition are classification and regression. In classification problems, data are collected and given discrete class labels. In a regression problem, on the other hand, data labels are typically continuous values, not categorical.

II. IMAGE RETREIVAL

Many images on the World Wide Web confront the users with new problems. Images are a fundamental part of our daily communication. The German saying "Ein Bild sagt mehr als tausend Worte" (literally: "A picture says more than a thousand words.") reflects this. The huge amount of pictures digitally available is not manageable by humans any more [2]. A person searching for a picture in a database of 100 images will probably find the search quite fast by just viewing the images or small versions of the images (thumbnails). If a thousand, ten thousand, or even more images are involved, the task becomes boring and interminable. Computers might be able to help here in the same way as they already do for searching text documents. A well-established example for text retrieval is the Internet search engine Google. Entering some keywords often helps finding related documents from the vast amount of documents available on the Internet. Google also offers a possibility to search for images, but the way the search is performed does not always lead to satisfactory results.

A broad variety of applications requires searching for images, in medical applications many images are produced and a physician might search for similar images to learn about treatments of former patients and their outcomes. Image retrieval is the task of searching for images from an image database [2]. The query to the database can be of various types,

Query-by-text: The user gives a textual description of the image he is looking for. Query-by-sketch: The user provides a sketch of the image she is looking for. Query-by- example: The user gives an example image similar to the one he is looking for.

In image retrieval the similarity between two sets of features, extracted from the database image and the query image has been used as a match measure. The match measure has been used to retrieve those regions present in a database of images, which are similar to the query image.

III. NECESSITY OF COLOR IMAGE RETRIEVAL

Color image retrieval and classification are very important in the field of image processing, as a hotspot in image

processing, image retrieval and classification are very important. Image retrieval methods are mainly based on color, texture, shape and semantic-image [8]. Color features are among the most important features used in image database retrieval. Due to its compact representation and low complexity, fuzzy similarity measure is the most commonly used technique in measuring color similarity of images. These features are independent of specific domain and can be used in general systems of retrieval images. The color feature is the first and one of the most widely used visual features in image retrieval and indexing. The most important advantages of color feature are power of representing visual content of images, simple extracting color information of images and high efficiency, relatively power in separating images from each other, relatively robust to background complication and independent of image size and orientation.

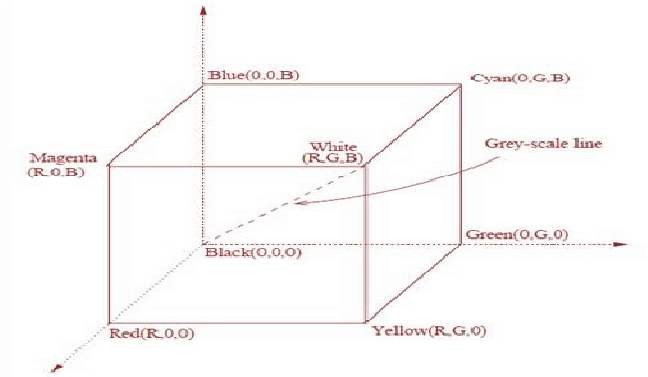
For image matching, features are extracted which may be shape features, texture features, or color features. Thus, a database is formed and for each image in the database, image features are found out. Another image, called the query image, is taken as the input image whose features are also found out. Then, the features of the query image are allowed to match with the features of the images in the database [5, 8]. One important criterion for testing the efficacy of search and retrieval is that the features of the query image should almost be there in the images of the database.

IV. COLOR SPACE

Color is a sensation created in response to excitation of our visual system by electromagnetic radiation known as light. Color is the perceptual result of light in the visible region having wavelength in the region of 400nm to 700nm incident upon the retina of human eye. A color can be specified by a tri-component vector. The set of all colors form a vector space called color space.

RGB Color space

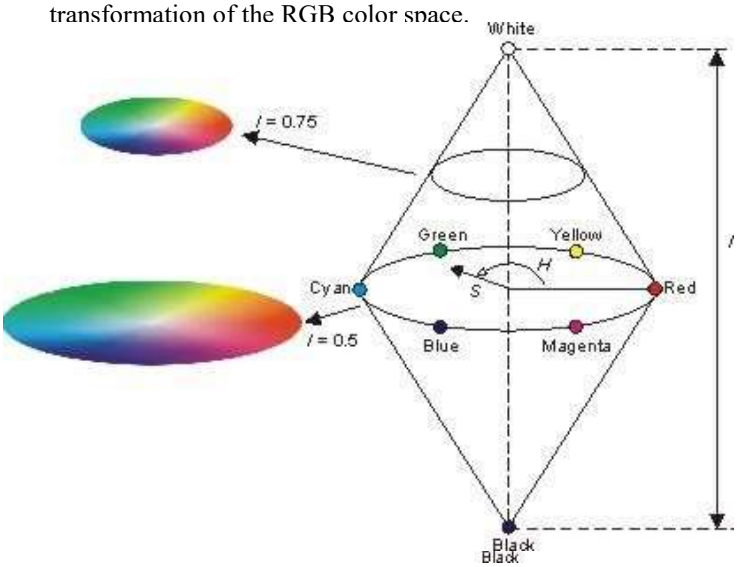
For representing colors several color spaces can be used. A color space is a specification of a coordinate system and a subspace within that system where each color is represented by a single point [6]. Color is perceived by humans as a combination of tristimuli i.e., R(red), G(green) and B(blue) which are usually called three primary colors. In hardware devices like monitors and digital cameras RGB color space is used. It is based on Cartesian coordinate system.



The above Fig represents the RGB color space using the Cartesian coordinate system.

HSI Color space

The HSI color model represents every color with three components Hue (H), Saturation (S), Intensity (I). The HSI color space is very important and attractive color model for image processing applications because it represents color similarly how the human eye senses colors. The Hue component describes the color itself in the form of an angle between [0,360] degrees. 0 degree mean red, 120 means green, 240 means blue, 60 degrees is yellow, 300 degrees is magenta. The Saturation component signals how much the color is polluted with white color. The range of the S component is [0, 1]. The Intensity range is between [0,1] and 0 means black, 1 means white. It is a nonlinear transformation of the RGB color space.



As the above Fig shows, hue is more meaningful when saturation approaches 1 and less meaningful when saturation approaches 0 or when intensity approaches 0 or 1. Intensity also limits the saturation values.

Fuzzy logic starts with and builds on a set of user-supplied human language rules. The fuzzy systems convert these rules to their mathematical equivalents. This simplifies the job of the system designer and the computer, and results in much more accurate representations of the way systems behave in the real world. Additional benefits of fuzzy logic

include its simplicity and its flexibility. Fuzzy logic can handle problems with imprecise and incomplete data, and it can model nonlinear functions of arbitrary complexity.

V. SIMILARITY MEASURE

Similarity measure is a quantity that reflects the strength of the relation between two objects or features. This chapter mainly focuses on similarity measure and the purpose of finding similarity measure and the comparison between the existing measure and proposed measure for pattern data sets.

Similarity is quite difficult to measure. It is a quantity that reflects the strength of the relation between two objects or features [10]. The similarity between two features A and B is usually denoted as S (A,B). It has a normalized range of 0 to 1. Measuring similarity of features endorse to

1) Distinguish one object from another
2) Group them based on their similarity
3) Grouping may also give more efficient organization and ratio of information
4) Predict the behaviour of new object.

The concept of similarity is fundamentally important in almost every scientific field. For example, in mathematics, geometric methods for assessing similarity are used in studies of congruence, as well as in allied fields such as trigonometry. Topological methods are applied in fields such as semantics. Fuzzy set theory has also developed its own measures of similarity, which find application in areas such as management, medicine and meteorology. An important problem in molecular biology is to measure the sequence similarity of pairs of proteins.

Similarity is a core element in achieving an understanding of variables that motivate behavior and mediate affect. It also played a fundamentally important role in psychological experiments and theories. For example, in many experiments people are asked to make direct or indirect judgments about the similarity of pairs of objects [7]. A variety of experimental techniques are used in these studies, but the most common are to ask subjects whether the objects are the same or different, or to ask them to produce a number, between say 1 and 7, that matches their feelings about how similar the objects appear. The concept of similarity also plays a crucial but less direct role in the modeling of many other psychological tasks. This is especially true in theories of the recognition, identification, and categorization of objects, where a common assumption is that the greater the similarity between a pair of objects, the more likely one will be confused with the other. Similarity also plays a key role in the modeling of preference and liking for products or brands, as well as motivations for product consumption.

RGB index

There are three statistical histograms representing one pictures feature. The first is color attribute RED histogram, the second is color attribute GREEN histogram and the last is color attribute BLUE histogram [11]. The statistical index is,

$$\mathbf{U_i} = \{\mathbf{V_{ir}}, \mathbf{V_{ig}}, \mathbf{V_{ib}}\}$$

Where,

$$\mathbf{V_{ir}} = \{\mathbf{x_{ir1}}, \mathbf{x_{ir2}}, \mathbf{x_{ir3}}, \ldots\ldots\ldots\ldots \mathbf{x_{ir128}}\}$$

$$\mathbf{V_{ig}} = \{\mathbf{x_{ig1}}, \mathbf{x_{ig2}}, \mathbf{x_{ig3}}, \ldots\ldots\ldots\ldots \mathbf{x_{ig128}}\}$$

$$\mathbf{V_{ib}} = \{\mathbf{x_{ib1}}, \mathbf{x_{ib2}}, \mathbf{x_{ib3}}, \ldots\ldots\ldots\ldots \mathbf{x_{ib128}}\}$$

The difference between picture i and picture j under the RED attribute is

$$r_{ij} = \sum_{k=1}^{128} |x_{irk} - x_{jrk}|$$

The difference between picture i and picture j under the GREEN attribute is

The difference between picture i and picture j under the BLUE attribute is

$$b_{ij} = \sum_{k=1}^{128} |x_{ibk} - x_{jbk}|$$

The similarity measure of two pictures is,

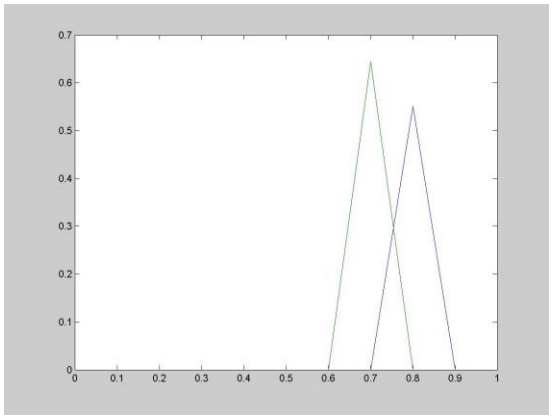
$$\mathbf{ij} = \frac{\mathbf{r_{ij}} + \mathbf{g_{ij}} + \mathbf{b_{ij}}}{\mathbf{3}}$$

The image retrieval method using fuzzy similarity measure is compared with the existing measures [12]. The images are fetched from the database and their degree of similarity is calculated with the query image.

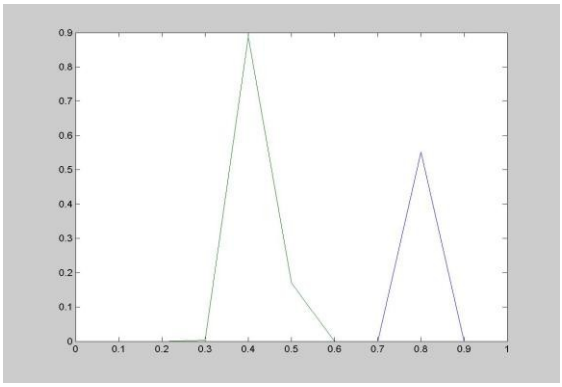
Query image

SNO	Image	Xiaojuan Based Similarity Method	Fuzzy Distance Based Similarity Method
1		0.7093	0.72
2		0.4101	0.1699
3		0.6547	0.5428
4		0.6492	0.5673
5		0.7275	0.72
6		0.7083	0.72
7		0.5379	0.2980
8		**0.8222**	**1**

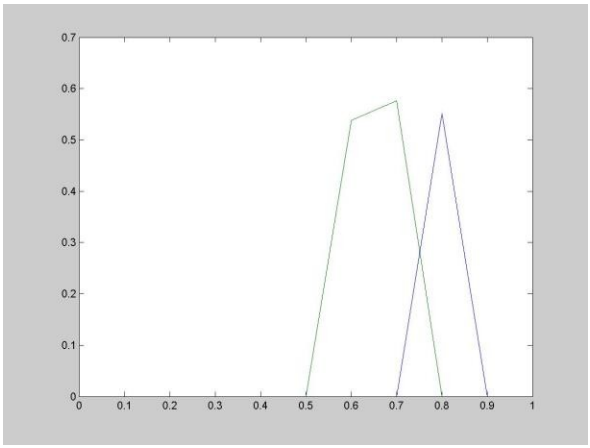
The plot of these images are listed below



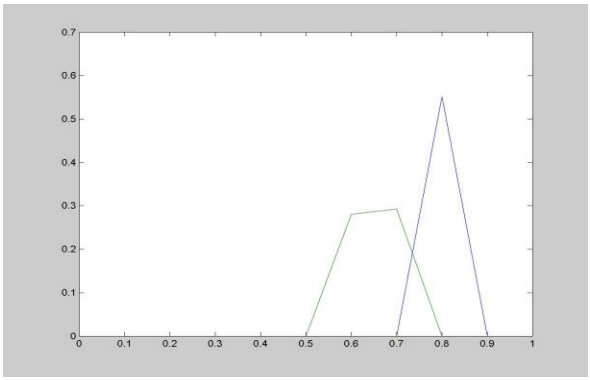
Im 1



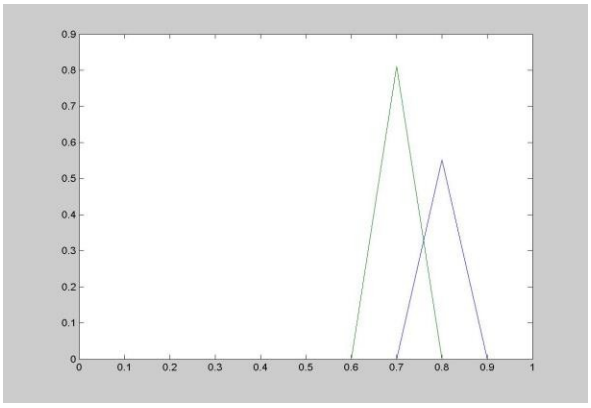
Im 2



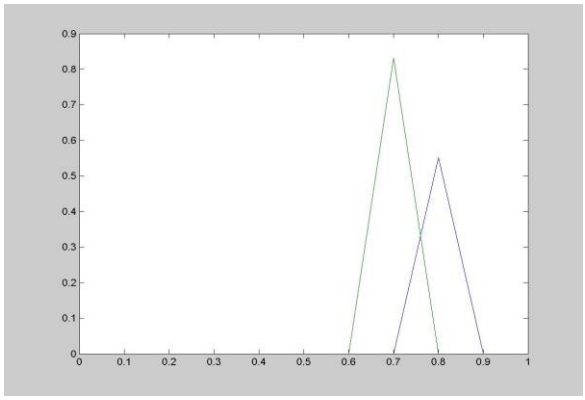
Im 3



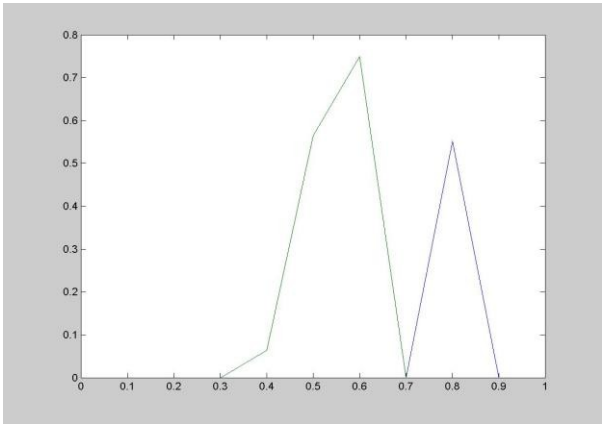
Im 4



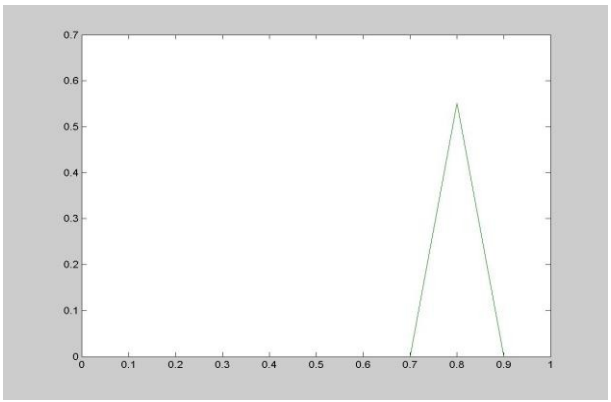
Im 5



Im 6



Im 7



Im 8

VI. COMPARISON

From the Table 5.2, Im 8 is same as that of query image so the degree of similarity should be 1, which is true in the case of fuzzy distance based similarity method but Xiaojuan based similarity method fails to identify the 100% similarity. Also Im 2, is found to be dissimilar with the query image, the degree of similarity is found to be very low in fuzzy distance based similarity method, but the Xiaojuan based similarity method provides a high measure.

Hence it reveals that the fuzzy distance based similarity method provides better similarity when compared with the existing method [9, 13]. The experimental results show that idea which takes the fuzzy similarity measure for the method of image retrieval is found to be reasonable and effective.

VII.CONCLUSION

- Based on the existing research work we have studied the color image retrieval method. This is done using fuzzy similarity measure. The similarity measure is calculated by hue feature vector. The measure greatly reduces the influence of inaccurate measure and provides a very intuitive quantification.
- The results obtained by the proposed method reflect the significance of fuzzy representation rather than the crisp definition.

REFERENCES

[1] Timothy J.Ross, Fuzzy Logic with Engineering Applications, McGraw - Hill, 1997.

[2] William K.Pratt, Digital Image Processing, John Wiley and Sons Publications, Third Edition,2001.

[3] Earl Gose, Richard Johnsonbaugh, Steve Jost, Pattern Recognition and Image Analysis, Prentice Hall of India, Third Edition,2002.

[4] Rafael C.Gonzalez, Richard E.Woods, Digital Image Processing, Prentice Hall of India, Second Edition,2005.

[5] Nachtegeal.M, Van der Weken.D, De Witte.V, Schulte.S, Melaange.T, Kerre.E.E, "Color Image Retrieval using Fuzzy Similarity Measures and Fuzzy Partitions", Image Processing, ICIP 2007, IEEE International Conference on, Vol-6, pp. VI-13-VI-16, 2007.

[6] Gwangwon Kang, Junguk Beak, Jongan Park, "Features Defined by Median Filtering on RGB Segments for Image retrieval", Computer Modeling and Simulation, EMS '08 Second UKSIM European Symposium on, pp. 436-440, 2008.

[7] Jeong- Yo Ha, Gye- young Kim, Hyang-II Choi, "The Content-Based Image Retrieval Method Using Multiple Features", Networked Computing and Advanced Information management. NCM'08. Fourth International Conference on, Vol-1, pp.652-657,2008.

[8] Yu Xiaohong, Xu Jinhua, "The Related Techniques of Content-Based Image Retrieval", Computer Science and Computational Technology. ISCSCT '08. International Symposium on, vol.1, pp.154-158, 2008.

[9] Zhangyan Xu, Shichao Shang, Wenbin Qian, Wenhao Shu "A method for fuzzy risk analysis based on the new similarity of trapezoidal fuzzy numbers", Expert Systems with Applications, Vol-37, pp. 1920–19270, 2010.

[10] B.Sridevi and R.Nadarajan, "Fuzzy Similarity Measure for Generalised Fuzzy Numbers", Int.J. Open Problems Compt. Math, Vol-2,No-2, June 2009.

[11] Xiaojuan Ban, Xiaolong Lv, and Jie Chen, "Color Image Retrieval And Classification Using Fuzzy Similarity Measure And Fuzzy Clustering Method", IEEE Conference on Decision and Control, 2009.

[12] John Yen, Reza Langari, Fuzzy Logic- Intelligence, Control and Information, Pearson Education, 2011.

[13] Gerasimos Louverdis and Ioannis Andreadis, "Soft Morphological Color Image Processing: A Fuzzy Approach", med.ee.nd.edu/MED11/pdf/papers/t7- 032.pdf.