IJCTA
Vol.8, No.1, Jan-June 2015, Pp.326-335

An Improved IHBM using Smoothing Projections

N Gnaneswara Rao[1], V Ramakrishna Sajja[2], S Deva Kumar[3] and M Venkata Rao[4]
[1, 4]Professor, Dept. of CSE, VFSTR University
[2, 3] Assistant Professor, Dept. of CSE, VFSTR University
Mail id: gnani.nitta@gmail.com

ABSTRACT

In the last two decades, the rapid growth of the content based image retrieval has enormously increased the number of image collections available. The accumulation of these image collection including art works, satellite and medical imagery is attracting more and more users in various professional fields. The study of image retrieval, which is concerned with effectively and efficiently accessing desired images from large and varied image collections, has become more interesting and more challenging. Therefore selection of proper similarity measure is an essential consideration for a success of many histogram image retrieval systems. This paper proposed an Improved Integrated Histogram Bin Matching (IHBM) using smoothing projection called neighbor-bank projection which is efficient and effective than existing methods. In this, initially convert the image from RGB into HSV model and then calculate the Ordered Histogram values, which is in turn used in calculating the IHBM similarity measures. The experimental results clearly showed that the proposed Improved IHBM has significantly improved by the smoothing projection technique.

Keywords: Ordinal histograms, Distance functions, Image Retrieval, IHBM, Similarity Measures, Histogram based Image Retrieval.

1. INTRODUCTION

The use of images as a tool to communicate is hardly a new mean of communication in every day's life. However, there has been a rapid growth in the number, availability and usage of images in all walks of life. But this ocean of image information would be useless without the ability to manipulate, classify, archive and access them quickly and selectively. During the past two decades we have seen a rapid increase in the size of digital image collections. As the computational power of both hardware and software have increased, the ability to store more complex data types in databases, such as video, audio and images, has been drastically improved. Large image databases are difficult to browse with traditional text searches. More effective techniques are needed with collections containing millions of images. This is more often referred to as Content Based Image Retrieval (CBIR) [1,16,17,18]. Content Based Image Retrieval relies on the characterization of primitive features such as color [3,4,5,15], shape[9] and texture[10,11] that can be automatically extracted from the images themselves. Queries to CBIR system are most often expressed as visual examples of the type of the image or image attributed being sought. For Example user may submit a sketch, click on the texture pallet, or select a particular shape of interest. This system then identifies those stored images with a high degree of similarity to the requested feature.

To date, image and video storage and retrieval systems have typically relied on human supplied images to provide pattern recognition. The exiting distance functions[4,5,12,13,14,19] for large image and video archives are time consuming to retrieve. They necessitate that each image and video scene is analyzed manually by a domain expert so the contents can be described by their physical attributes. Along with the distance function the smoothing projections can retrieve the similar images very efficiently and effectively.

Currently the most widely used image search engine, the GOOGLE, provides its users with the textual annotation kind of implementation. With lacks of images added to the image database, not many images are annotated with proper description. So many relevant images go unmatched.

The most widely accepted content-based image retrieval techniques use the Quadratic Distance [3,4,5] and the Integrated Region Matching methods[13,14]. The Quadratic Distance method, though yields metric distance, is computationally expensive. The proposed Approach using an improved IHBM method which access the images effectively and efficiently than other existing methods. We also provide an interface where the user can give a query image as an input. The distance is automatically extracted from the query image and is compared to the images in the database retrieving the matching images.

2. RELATED WORK

This section presents some of the popularly existing histogram similarity measures [3,4,5], namely, Histogram Intersection (HI), Histogram Euclidean Distance (HED) and Histogram Quadratic Distance Measures (HQDM).

2.1. Histogram Intersection (HI)

Histogram Intersection [4, 5] is for color image retrieval and to find known objects within images using color histograms,

$$D_{HI}(q,t) = \sum_{i=0}^{M-1} | h_q(i) - h_t(i) |$$

Where $D_{HI}(q,t)$ is the distance between query image q and target image t, and h_q and h_t are the color histograms of query and the target images respectively and m is the number of bins of histogram.

2.2. Histogram Euclidean Distance (HED)

The Euclidean distance [4, 5] is, as follows: given histograms h_q and h_t

$$D_{HED}(q,t) = (h_q - h_t)^T(h_q - h_t) = \sum_{i=0}^{M-1}(h_q(i) - h_t(i))^2$$

D_{HED} (q,t) is the distance between query image q and target image t, and h_q and h_t are the color histograms of query and the target images respectively, moreover, M is the number of bins of histogram.

The Fig.1 shown below represents the Minkowski distance measures stated above.

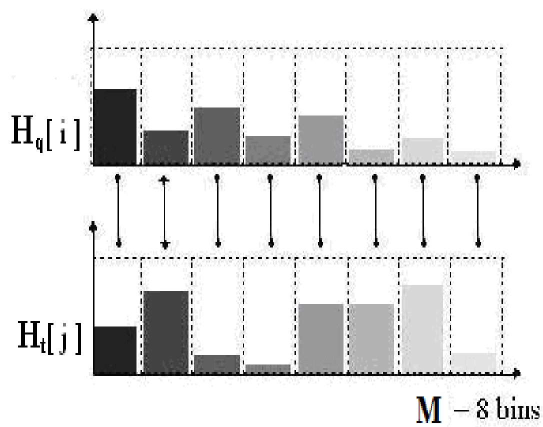

Fig 1. The Minkowski distance measures

2.3. Histogram Quadratic Distance Measures (HQDM)

A Histogram Quadratic Distance Measure is used in IBM QBIC system for color histogram based image retrieval [3, 4, 5]. In [5], it is reported that quadratic distance metric between color histograms provides more desirable results than "like-bins" that are only comparisons between color histograms. The quadratic form distance between histograms $h_q$and h_t given by

$$D_{HQDM}(q,t) = (h_q - h_t)^T A (h_q - h_t)$$

Where D_{HQDM} (q,t) is the distance between query image q and target image t, and h_q and h_t are the color histograms of query and the target images respectively and $A = [a_{ij}]$ and a_{ij} denotes the similarity between image histograms with bins i and j. The Quadratic form metric is a true distance metric when $a_{ij}=a_{ji}$ and $a_{ii}=1$.

The HQDM is computationally more expensive than the Minkowski form metrics since it computes the cross similarity between all histogram bins as shown in Fig.2.

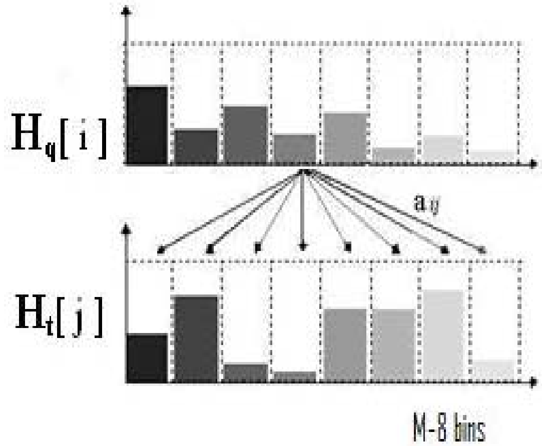

Fig 2. The quadratic distance measure

3. A NOVEL IMPROVED IHBM USING SMOOTHING PROJECTIONS

The proposed approach retrieves the images similar to the Quadratic distance measure by applying smoothing projection efficiently.

The three main steps of the proposed method is given below

1. Conversion of RGB space into HSV space for Quantization.
2. Compute the inter-bin distances matrix HISTd (Q, T) between all pairs of images using smoothing projections.
3. Computation of similarity measure using the IHBM.

3.1. HSV Color Space

The determination of the optimum color space is an open problem, certain color spaces have been found to be well suited for the content-based query-by-color. The proposed method used HSV (Hue, Saturation and Value) Color space [6], because it is natural and is approximately perceptually uniform.

3.2. HSV Quantization

HSV Quantization gives 18 hues, 3 saturations, 3 values, and 4 gray levels, which results 166bins [5, 6] for each image. Then color histogram is computed for 166 bins, and then it is normalized.

3.3. Computing distance matrix from Ordered Histograms

This partmainly discusses the computing distance matrix from ordered histograms [2]. An ordered histogram or distribution histogram is one type of histogram where adjacent histogram bins contain all related information. For example, in a gray-level histogram the neighbor dimensions represent pixel intensity values that are almost the same type. So some different natural smooth distributions, such as in color spectrum, the same characteristics are present.

In an ordered histogram the closely situated elements correlate or relate with each other more strongly than elements which are further apart i.e. which are not strongly correlated, a feature vector space can be projected to a smaller number of dimensions without significant loss of information. This kind of smoothing projection together with IHBM distance function induces a Novel Improved IHBM similarity measure. A hypothesis is made that using a smoothing projection, the statistical properties of the samples are more evident and the similarity measure is improved [2,7,8].

The Dimensionality of any histogram is mainly reduced by a linear projection to a subspace, called the neighbor-bank subspace metric. Forming a neighbor-bank subspace is possible only with a set of discrete sampled $\cos^2$ functions. The Histograms are mainly projected on a set of $\cos^2$ functions as shown in Fig 3. Let L be the length of an original histogram $p = (p_{0} \ldots p_{L-1})^T$. N be the number of banks indexed with k from 0 to N-1.

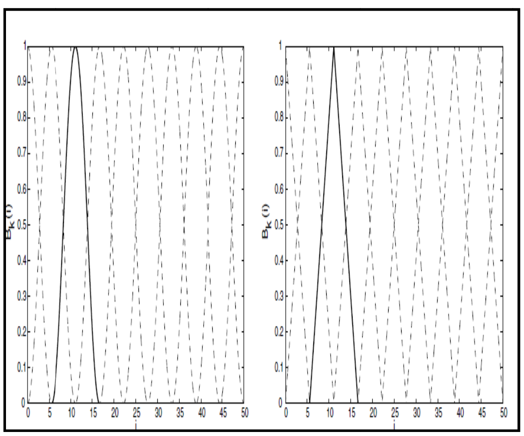

Fig. 3. N = 10 neighbor-banks of $\cos^2$ and triangle functions for discrete histograms of size L = 50

Then for B_K (i) the discrete neighbor-banks of cos ^{2} function can be constructed as

$$B_k(i) = \begin{cases} \cos^2\left(\pi\left(\frac{i}{L}\frac{N-1}{2} + \frac{k}{2}\right)\right) & \text{if } \frac{L}{N-1}(k-1) \leqslant i \leqslant \frac{L}{N-1}(k+1), \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

By constructing a transformation matrix from the above values of neighbor banks of cos ^{2} function, we get the following matrix

$$B = \begin{pmatrix} B_0(0) & B_0(1) & \cdots & B_0(L-1) \\ B_1(0) & B_1(1) & \cdots & B_1(L-1) \\ \vdots & \vdots & \ddots & \vdots \\ B_{N-1}(0) & B_{N-1}(1) & \cdots & B_{N-1}(L-1) \end{pmatrix}, \quad (2)$$

The projection can be performed by matrix multiplication as

$$r = Bp, \quad (3)$$

Where r is the projection.

The distance between a bin pair, HISTd (Q_i, T_j) is the transformation matrix from the above values of neighbor banks of cos ^{2} function computed a priori, independent of the Query image and target images

3.4. Integrated Histogram Bin Matching (IHBM)

IHBM (Integrated Histogram Bin Matching)[12], is a novel metric Similarity measure to compare the color feature of quantized images. The main idea of this, consists of modeling the comparison of color-quantized images as a Transportation problem [12,13,14].

For matching histogram bins of two images, the closest histogram bin pair is considered first. If the bins are of the same size then the two most similar bins are matched otherwise a partial match occurs. This process is repeated until all the histogram bins are matched completely. After matching histogram bins, the similarity measure is computed as a weighted sum of the similarity between histogram bin pairs, with weights determined by the matching scheme.

4. EXPERIMENTAL RESULTS &ANALYSIS

The proposed an Improved IHBM and Histogram Intersection (HI), Histogram Euclidean Distance (HED) and Histogram Quadratic Distance Measures (HQDM) are applied on the following red rose flower image. The retrieval effectiveness is measured, based on the Precision rate of all the four methods considered and they are listed in the tables. Recall measure that indicates the proportion of the relevant images returned is also evaluated for all the images considered, on all the four methods.

The Query image and retrieved relevant images of red rose images are shown in Fig. 4 and 5 respectively for HI,HED,HQDM,IHBM and Improved IHBM methods considered.

Fig 4. The results of the query image and retrieved relevant images of HI, HED, HQDM and IHBM methods for the image Red rose

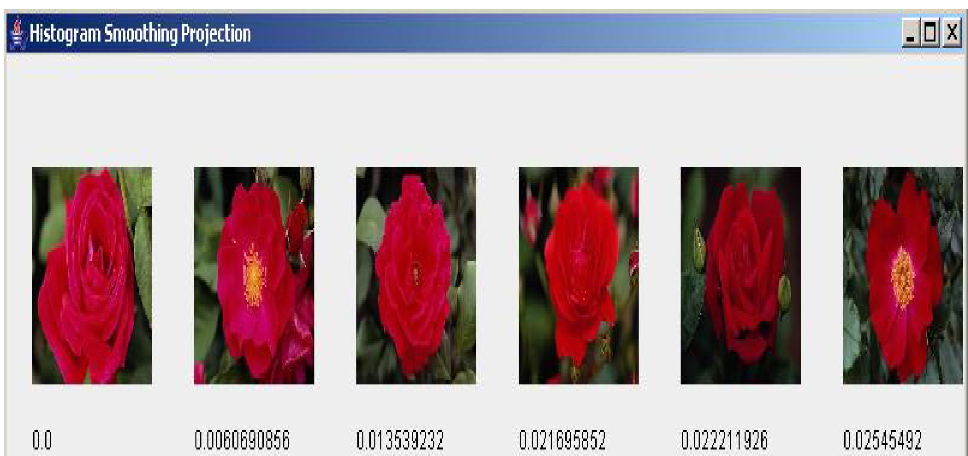

Fig 5. The results of the query image and retrieved relevant images of Improved IHBM using Smoothing Projections

Table-1: Precision for TOP 5, 10 and 20 images

Red rose	HI	HED	HQDM	IHBM	Improved IHBM
TOP 5	5	4	5	5	5
TOP 10	10	9	10	10	10
TOP 20	18	15	18	19	20

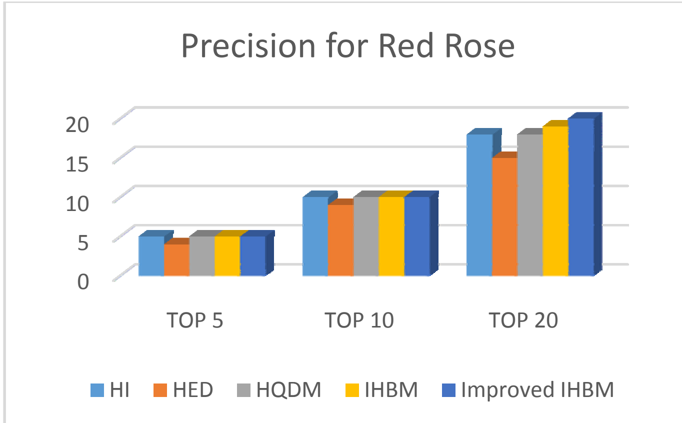

Fig 6. The Precision for Red Rose image for all methods

The Table 1 clearly indicates the Improved IHBM, IHBM, HI and HQDM methods shows the similar results which are superior to HED method for Top 5 and Top 10 red rose images. For TOP 20 red rose images, Improved IHBM outperform the other four methods.

The precision Vs the number of images returned for the red rose image is plotted for the five methods considered in the Fig.6.

5. CONCLUSIONS

In this paper, the properties of various distance metrics were examined primarily in the context of ordered histogram type data. A new smoothing projection, the neighbor-bank projection, IHBM Smoothing were also introduced. The smoothing projection seems to improve the accuracy of some of the studied distance functions and to have advantages when combined with methods utilizing the statistical properties of the data. The new smoothing technique like improved IHBM is experimented on 1000 Corel database color images and the experimental results with the help of table and graph clearly indicate the proposed Improved IHBM is more accurate and efficient than the four existing methods.

REFERENCES

[1] A. Bhagyalaksluni, V.Vijaya chamundeeswari," A Survey on Content Based Image Retrieval Using Various Operators", 2014 IEEE International Conference on Computer Communication and Systems(ICCCS '14), Chennai, Feb 20-21,2014.

[2] Kamarainen, J.-K., Kyrki, V., K€alvi€ainen, H., June 2001. Similarity measures for ordered histograms. In: Proc. 12th Scand. Conf. Image Anal. SCIA2001, Bergen, Norway, pp. 699–705.

[3] Hafner, J., Sawhney, H.S., Equitz, W., Flickner, M., Niblack,W., 1995. Efficient color histogram indexing forquadratic form distance functions. IEEE Trans. Pattern Anal.Machine Intel. 17 (7), 729–736.Jou, F.D., 2003.

[4] J.R.Smith and S.-F Chang,Integrated spatial and feature image query, Multimedia Syst. 7(2) ,129-140(1999).

[5] J. R. Smith. Integrated Spatial and Feature Image Systems: Retrieval, Analysis and compression . PhD thesis,Columbia University, New York,NY, February 1997

[6] John R. Smith and Shih-Fu Chang,"Tools and Techniques for Color Image Retrieval" Columbia University Department of Electrical Engineering and Center for Telecommunications Research New York, N.Y. 10027.

[7] Aksoy, S., Haralick, R.M., 2001. Feature normalization and likelihood-based similarity measure for image retrieval. Pattern Recognition Lett. 22, 563–582.

[8] Cha, S.-H., Srihari, S.N., 2002. On measuring the distance between histograms. Pattern Recognition 35, 1355–1370.

[9] N. Gnaneswara Rao and V. Vijaya Kumar "Texturebased image indexing and retrieval" VISAPP 2007 International Conference on Computer Vision Theory and Applications, Barcelona, Spain, March,2007, PageNo:177-181.

[10] Vijaya Kumar , N. Gnaneswara Rao, A.L.Narsimha Rao.RTL: Reduced Texture spectrum with Lag value Based Image Retrieval for Medical Images, International Journal of Future Generation Communication and Networking Vol. 2, No. 4, December, 2009.

[11] N Gnaneswara Rao, V Vijaya Kumar, P S V Srinivasa Rao,"Novel approaches of evaluating texture based similarity features for efficient medical image retrieval system"International Journal of Computer Applications (0975 – 8887) ,Volume 20 ,No.7,pp20-26April 2011.

[12] Vijaya Kumar , N. Gnaneswara Rao, A.L.Narsimha Rao ,V.Venkata Krishna, IHBM: Integrated Histogram Bin Matching For Similarity Measures of Color Image Retrieval, International Journal of Signal Processing, Image Processing and Pattern Recognition,IJSIP, Vol. 2, No. 3, September, 2009.

[13] J. Li, J.Z. Wang, G. Wiederhold, IRM: integrated region matching for image retrieval, in:Proceedings of the Eighth ACM Multimedia Conference, 2000, pp. 147–156.

[14] N. Gnaneswara Rao, T. Sravani, Dr.V.Vijaya Kumar "OCRM: Optimal Cost Region Matching Similarity Measure for Region Based Image Retrieval", International Journal of Multimedia & Ubiquitous Engineering; 2014, Vol. 9 Issue 4, p327, June 2014.

[15] V.H. Vu, Q.N. Huu and H.N.T. Thu "Content based Image Retrieval with bin of Color Histogram", published in IEEE, Audio, language and Image Processing, 2012,

[16] Y. Liu, D. Zhang, G. Lu, W.Y. Ma, "A survey of content-based image retrieval with high-level semantics", published by Elesvier, Pattern recognition 262-282, 2007.

[17] N.S. Vassilieva, "Content-based Image Retrieval methods", published in Programmirovanie, vol. 35, No. 3, 2009.

[18] O.A. Vatamanu, M. Frandes, M. Ionescu and S. Apostol "Content-Based Image Retrieval using Local Binary Pattern, Intensity Histogram and Color Coherence Vector", published in IEEE, E-Health and Bioengineering Conference, 2013.

[19] M. J. Swain and D. H. Ballard, "Color Indexing," International. Journal of Compututer Vision., Vol.7, No. 1, pp. 11–32, 1991.