

ROTATION INVARIANT COLOR RETRIEVAL

Ms. Swapna Borde¹ and Dr. Udhav Bhosle²

¹Vidyavardhini's College of Engineering and Technology, Vasai (W),

Swapnaborde@yahoo.com

²Rajiv Gandhi Institute of Technology, Mumbai

udhavbhosle@gmail.com

ABSTRACT

The new technique for image retrieval using the color features extracted from images based on Log-Histogram is proposed. The proposed technique is compared with Global color histogram and histogram of corners. It has been observed that number of histogram bins used for retrieval comparison of proposed technique (Log Histogram) is less as compared to Global Color Histogram and Histogram of corners. The experimental results on a database of 792 images with 11 classes indicate that proposed method (Log-Histogram) significantly improves Precision/Recall and Complexity of proposed method is less as compared to Global Color Histogram (GCH) and Histogram of Corners (HOC).

KEYWORDS

Content Based Image Retrieval (CBIR), Global Color Histogram (GCH), Histogram of Corners (HOC), Log-Histogram (LH), Histogram Distance (HD)

1. INTRODUCTION

Content Based Image Retrieval (CBIR) retrieves the similar images from the large image database based on visual features such as color, texture and shape. The Content Based Image Retrieval (CBIR) system includes two steps. The first step of CBIR is feature extraction. Set of features extracted from image is called as image signature. Size of image signature is less as compared to original image. Second step of CBIR is similarity measurement. Similarity measurement is used to compare the features of query image with features of images in the database so that top relevant images can be retrieved from the image database [1], [3].

Color is one of the most important image descriptor used in CBIR. Human beings can identify thousands of color shades and intensities as opposed to a few shades of grey. RGB model is a widely used color model. It is composed of three color components red, green and blue. Apart from the RGB color model, there are various other color models that one comes across in literature are NTSC, YCbCr, CMY and HSV. In image retrieval, a color histogram is the most commonly used color feature representation [6].

Texture analysis is one of the most difficult problems in the area of computer vision. Features based on textures can be useful in distinguishing between objects. The two main types used in computer vision to describe the texture of a region are structural and statistical. Structural methods include morphological operator and adjacency graph. Statistical methods include wavelet transform [6].

Shape next to color and texture is considered as important visual feature in describing the objects in images. The two main types used in computer vision to describe the shape of a region are boundary-based descriptor and region-based descriptor. Boundary Based shape descriptor describes the external characteristics of the region. Boundary based shape descriptors include area, perimeter, eccentricity and orientation. Region Based shape descriptor describes the internal characteristics of the region. Region based shape descriptor include statistical moments [2], [4], [5].

2. GLOBAL COLOR HISTOGRAM (GCH)

Histogram of images provides a global description of the appearance of an image. The information obtained from histograms is enormous. Just by looking at the histogram of the image, a great deal of information can be obtained. The main advantage of the histogram is that it can be determined very quickly and it is invariant to rotation [10].

Steps to extract features from a color image by using Global Color Histogram (GCH) are as follows:

1. This method first resizes the image to 128×128 .
2. Compute the histogram of query image as well as images in the database. Figure 2 shows the histogram of the image shown in Figure 1.



Figure 1

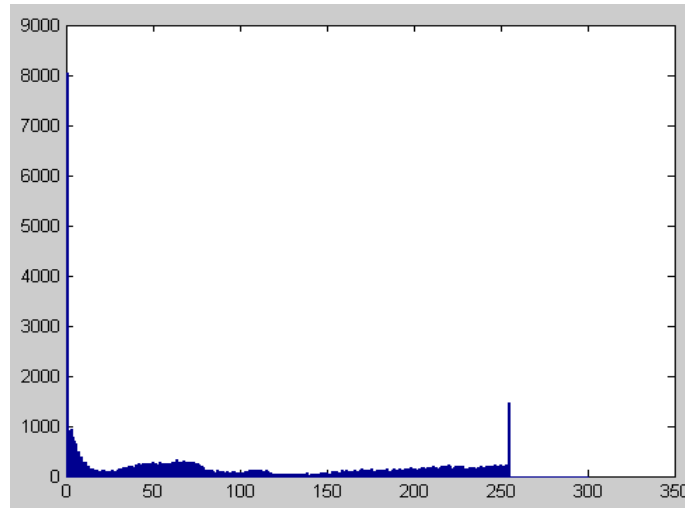


Figure 2

3. Then we use the Histogram Distance (HD) to compute the similarity measure between query image and images from the database.

$$HD(q,t)= \sum_{i=1}^M \text{abs}(h_q(i)-h_t(i)) \text{ --- (1)}$$

Where HD (q, t) is the distance between query image q and images in the database t. h_q and h_t are the color histograms of query and the database images respectively and M is the number of bins of histogram. In this particular experiment, the comparison of query image with images in the database is done on 251 bins (Bin No. 50 to Bin No.300) of histogram.

3. HISTOGRAM OF CORNERS (HOC)

The Harris detector is widely used in object detection and image retrieval system. The number of the corners detected in an image depends on the size and structure of the image [8], [9], [11].

Steps to extract features from a color image by using Histogram of Corners (HOC) are as follows:

1. This method first resizes the image to 128×128 .
2. Compute horizontal and vertical gradient (I_x and I_y) by using prewitt operator.
3. Compute the products of gradients at each pixel

$$Ix2=I_x * I_x, Iy2=I_y * I_y, Ixy=I_x * I_y$$

4. Convolve Gaussian Low Pass Filter with the products of gradients to get the resultant images (S_{x2} , S_{y2} and S_{xy})

5. Define at each pixel (x,y) the matrix

$$C(x,y) = \begin{bmatrix} S_{x2}(x,y) & S_{xy}(x,y) \\ S_{xy}(x,y) & S_{y2}(x,y) \end{bmatrix}$$

6. Compute the response of the detector at each pixel

$$R = \text{Det}(C) - k(\text{Trace}(C))^2$$

7. Corner points of an image can be detected by Comparing R with the threshold value. The Figure 3 shows detected corner points of an image (Figure 1).



Figure 3

8. Compute the histogram around the 16×16 pixel region of corners. Then we use the Histogram Distance (Equation No.1) to compute the similarity measure between query image and images from the database. The comparison of query image with images in the database is done on 251 bins (Bin No. 50 to Bin No.300) of histogram. Figure 4 shows the histogram of the image shown in Figure 3.

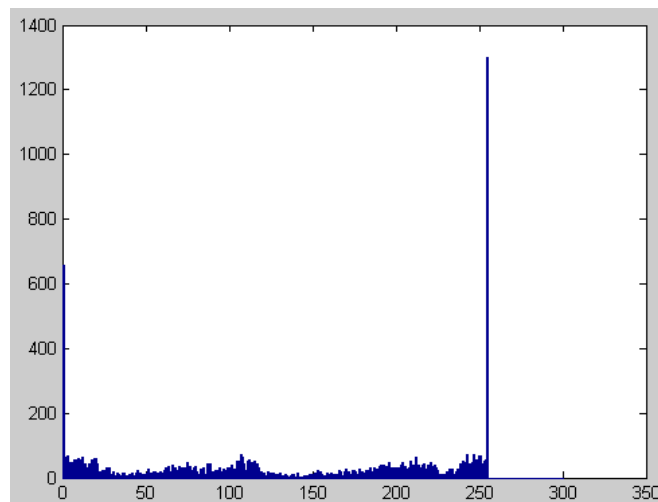


Figure 4

4. Log-Histogram (LH)

The disadvantage of Global Color Histogram (GCH) and Histogram of Corners (HOC) is that large number of bins of histogram (251) is used for comparison. In order to reduce the complexity in searching large image database, the original image is compressed by using a log operator and then histogram of the compressed image is computed. The resulting technique is very efficient in that it uses histogram of only 10 bins of each component for comparison.

Steps to extract features from a color image by using Log-Histogram (LH) are as follows:

- 1 This method first resizes the image to 128×128 and divides the image into R, G and B components. Then apply log operator on each component of the image.
- 2 Compute the histogram of log components of query image as well as images in the database.
- 3 Then Histogram Distance (HD) is used to compute the similarity measure between query image and images from the database.

whereas $HD(q, t)$ (Eq. No.1) is the distance between query image q and images in the database t . h_q and h_t are the histograms of log-R component of query and the database images respectively and M is the number of bins of histogram (10 bins for each component). Similarly compute Histogram Distance (HD) for log-G component and log-B component. Figure 5 shows the histogram of the log-R Component of the image shown in Figure 1.

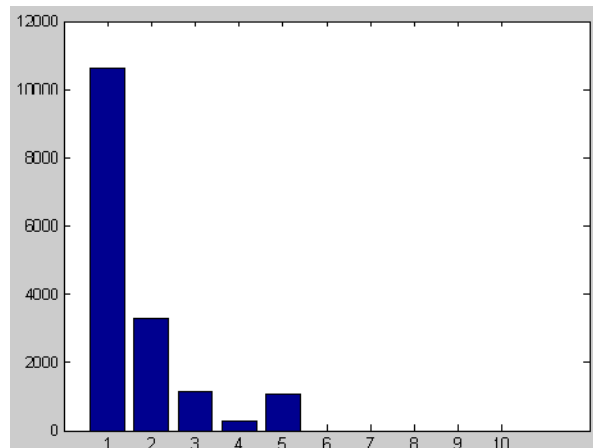


Figure 5

5. EXPERIMENTAL RESULTS

For evaluating the performance of the algorithms, we used Coil-100 Image database [7]. Our image database contains 792 images with 11 different classes (A to K). Some of the sample images which are used as query images are shown in Figure 6.



Figure 6. Sample database of 11 images

Extract the features of query image as well as images in the database by using Global Color Histogram (GCH), Histogram of Corners (HOC) and Log Histogram (LH). Compare the features of query image with features of images in the database by using Histogram Distance (HD) so that top relevant images are retrieved from the database. Performance of the CBIR system is measured by using Precision and Recall [12], [13].

$$\text{Precision} = \frac{\text{Total number of relevant images retrieved from the database}}{\text{Total number of retrieved (relevant+non-relevant) images from the database}}$$

$$\text{Recall} = \frac{\text{Total number of relevant images retrieved from the database}}{\text{Total number of relevant images in the database}}$$

Figure 8 (a), 8 (b) and 8 (c) shows the results of Global Color Histogram (GCH), Histogram of Corners (HOC) and Log-Histogram (LH) using Histogram Distance (HD) for the query image shown in Figure 7.



Figure 7. Query Image (Class 'B')



Figure 8(a). Results of Global Color Histogram (GCH) (Total No. of relevant images retrieved=56, Non-relevant images retrieved=16)

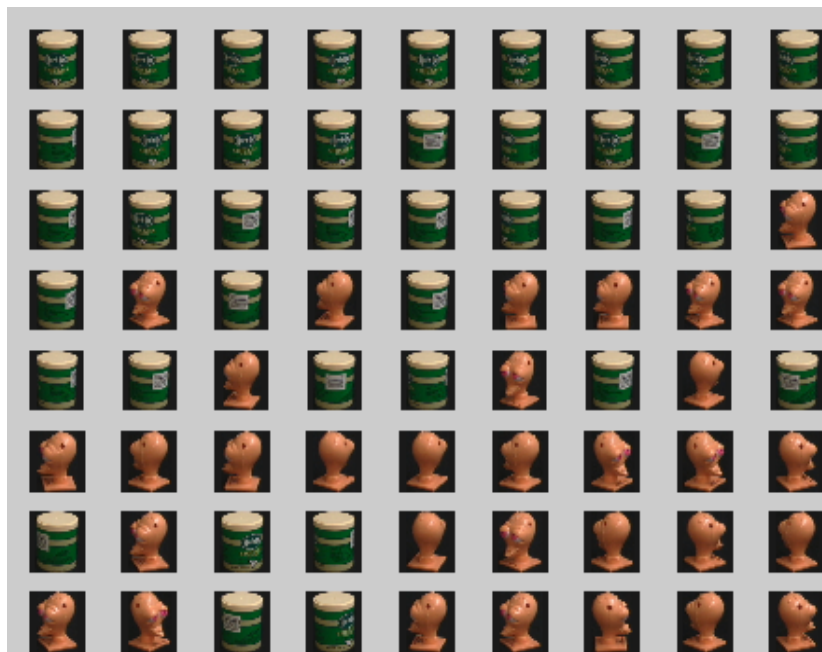


Figure 8(b). Results of Histogram of Corners (HOC) (Total No. of relevant images retrieved=40, Non-relevant images retrieved=32)



Figure 8(c). Results of Log-Histogram (LH) (Total No. of relevant images retrieved=72, Non- relevant images retrieved=00)

Table 1, Table 2 and Table 3 gives % Precision/Recall for all 11 classes using Global Color Histogram (GCH), Histogram of Corners (HOC) and Log-Histogram (LH). From each category randomly five images are chosen as query image and for every query image Precision/Recall values are computed.

Table 1: Average Precision/Recall for 11 classes (A to K) and 5 queries each using GCH

Class	A	B	C	D	E	F	G	H	I	J	K
Query Image 1 %Precision/Recall	100	77.77	100	88.88	100	87.5	100	94.44	91.66	100	91.66
Query Image 2 %Precision/Recall	100	73.61	100	88.88	100	86.11	100	94.44	100	100	100
Query Image 3 %Precision/Recall	100	84.72	100	91.66	100	100	100	94.44	81.94	100	88.88
Query Image 4 %Precision/Recall	100	90.27	100	100	100	88.88	100	94.44	88.88	100	88.88
Query Image 5 %Precision/Recall	100	94.44	100	94.44	100	100	100	95.83	93.05	100	83.33
Average %Precision/Recall	100	84.16	100	92.77	100	92.49	100	94.72	91.11	100	90.55

Table 2: Average Precision/Recall for 11 classes (A to K) and 5 queries each using HOC

Class	A	B	C	D	E	F	G	H	I	J	K
Query Image 1 %Precision/Recall	93.05	55.55	97.22	68.05	95.83	88.88	98.61	77.77	77.77	93.05	84.72
Query Image 2 %Precision/Recall	54.16	69.44	93.05	73.61	94.44	87.5	98.61	95.83	68.05	95.83	91.66
Query Image 3 %Precision/Recall	97.22	68.05	100	61.11	95.83	76.38	94.44	95.83	83.33	68.05	73.61
Query Image 4 %Precision/Recall	95.83	66.66	94.44	75.0	93.05	66.66	98.61	90.27	91.66	91.66	77.77
Query Image 5 %Precision/Recall	86.11	66.66	97.22	83.33	91.66	91.66	98.61	97.22	83.33	95.83	73.61
Average %Precision/Recall	85.27	65.27	96.38	72.22	94.16	82.21	97.77	91.38	80.82	88.88	80.27

Table 3: Average Precision/Recall for 11 classes (A to K) and 5 queries each using Log-Histogram (LH)

Class	A	B	C	D	E	F	G	H	I	J	K
Query Image 1 %Precision/Recall	91.66	100	100	69.44	100	100	94.44	100	100	100	100
Query Image 2 %Precision/Recall	100	100	100	91.66	100	100	100	100	100	100	100
Query Image 3 %Precision/Recall	100	100	100	76.38	100	100	87.5	100	73.61	100	100
Query Image 4 %Precision/Recall	91.66	100	100	70.83	100	100	100	100	81.94	100	100
Query Image 5 %Precision/Recall	95.83	100	100	93.05	100	100	100	100	100	100	100
Average %Precision/Recall	95.83	100	100	80.27	100	100	96.38	100	91.11	100	100

From Table 1, it is clear that average Precision/Recall of Global Color Histogram (GCH) varies from 84.16 to 100%. From Table 2, it is clear that average Precision/Recall of Histogram of Corners (HOC) varies from 65.27 to 97.77%. From Table 3, it is clear that average Precision/Recall of Log-Histogram (LH) varies from 80.27 to 100%

Table 4 gives Histogram Bins Comparison of CBIR Techniques based on Global Color Histogram (GCH), Histogram of Corners (HOC) and Log-Histogram (LH).

Table 4: Histogram Bin Comparison of CBIR Techniques

Technique	No. of Bins used for Comparison
GCH	251 Bins of 128*128 Image
HOC	251 Bins of 16*16 pixel region of corners
LH	10 Bins of each component

6. CONCLUSION

In this paper, Rotation Invariant Log Histogram (LH) is proposed to retrieve similar images from the large image database. The Global Color Histogram (GCH), Histogram of Corners (HOC) and Log-Histogram (LH) Based CBIR techniques are tested on the image database with 792 images spread across 11 classes. From experimental results, it is clear that Performance of CBIR is improved using all the techniques such as Global Color Histogram (GCH), Histogram of Corners (HOC) and Log-Histogram (LH) considering 60% average Precision/Recall as acceptable norms. But Complexity of Log-Histogram (LH) is less as compared to Global Color Histogram (GCH) and Histogram of Corners (HOC). Hence the proposed technique (Log-Histogram) is very efficient in that it uses histogram of only 10 bins of each component for comparison.

7. REFERENCES

- [1] Guoping Qiu, "Color Image Indexing Using BTC," IEEE Transactions on Image Processing, VOL.12, NO.1, pp.93-101, January 2003.
- [2] B.G.Prasad, K.K. Biswas, and S. K.Gupta, "Region -based image retrieval using integrated color, shape, and location index," computer vision and image understanding, October 2003.
- [3] Minh N. Do, Member, IEEE, and Martin Vetterli, Fellow, IEEE, "Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback-Leibler Distance," IEEE Transactions On Image Processing, VOL.11, NO.2, February 2002.
- [4] Dr. Fuhui Long, Dr. Hongjiang Zhang and Prof. David Dagan Feng, "Fundamentals of Content-Based Image Retrieval,"
- [5] Michael Eziashi Osadebey, "Integrated content -based image retrieval using texture, shape and spatial information," Master Thesis Report in Media Signal Processing, Department of Applied Physics and Electronics, Umea University, Umea Sweden.
- [6] Rajashekhar, "Novel Image Retrieval Techniques domain specific approaches," Ph.D. Thesis Department of Electrical Engineering Indian Institute of Technology - Bombay, 2006.
- [7] Sameer A. Nene, Shree K. Nayar and Hiroshi Murase, "Columbia Object Image Library(COIL-100)", Technical Report
- [8] K. Velmurugan, Lt. Dr. S. Santosh Baboo, "Image Retrieval Using Harris Corners and Histogram of Oriented Gradients," International Journal of Computer Applications (0975-8887) Volume 24, No. 7, June 2011
- [9] Junqiu Wang and Hongbin Zha, Roberto Cipolla, "Combining Interest Points and Edges for Content-based Image Retrieval," IEEE Journal, June 8, 2010.
- [10] Neetu Sharma, Paresh Rawat and Jaikaran Singh, "Efficient CBIR Using Color Histogram Processing," Signal & Image Processing: An International Journal (SIPIJ) Vol.2, No.1, March 2011.
- [11] Minakshi Banerjee, Malay K. Kundua, Pradipta Maji, "Content-based image retrieval using visually significant point features," Elsevier, Fuzzy Sets and Systems 160 (2009) 3323-3341
- [12] Swapna Borde, Dr. Udhav Bhosle, "Image Retrieval Using Contourlet Transform," International Journal of Computer Applications (0975-8887), Volume 34-No.5, November 2011.
- [13] Swapna Borde, Dr. Udhav Bhosle, "Image Retrieval Using Steerable Pyramid," International Journal of Computer Applications (0975-8887), Volume 38-No.7, January 2012.