Lossy compression effect on color and texture based retrieval performance.

Md. Facihul Azam
Department of Information Technology
Tampere University of Technology,
Finland
E-mail: md.azam@student.tut.fi

ABSTRACT

This experiment investigates the compression effects on image retrieval using color and texture features and presents retrieval results of a Content Based Image Retrieval (CBIR) system. The result of system can be evaluated by several performance metrics. Each one is evaluated by different measurement. In this experiment only Two performance metric has been used namely F1 measure and Average Normalized Modified Retrieval Rank (ANMRR) . Here four visual descriptors are used and measure performance individually. Two of the color feature and other two is texture features. The system is tested with two different image database color and gray images with visual feature vectors.

Keywords: CBIR , ANMRR, F1,color feature and texture feature.

1. INTRODUCTION

The amount of images in image database and internet cloud is increasing day by day due to the increasing popularity of smartphones and digital imaging devices such as cameras. This increases many applications like photo library or e-commerce and motivates researchers to develop new algorithms to image indexing and searching from image databases. The image retrieval processes can be done either by textual descriptor or low-level characteristic can be extracted from the image prior to search. Keywords annotation such as textual descriptors were used in traditional image retrieval systems. Keywords annotation method is laborious and not efficient process. Content Based Image Retrieval (CBIR) is developed to overcome key annotation problems. The idea behind CBIR is to retrieve images from the database using extracted visual features such as color, texture ,shape and spatial contents that are existing in images. Many CBIR system has been developed such as GIFT (GNU Image finding tools), *MUVIS* , MIT photobook, QBIC etc[1,3,4].In this paper it was investigated how efficiently image is retrieved from a given image database using color and texture features. The system made with MatLab programming Language.

2. PROCESSING AND EVALUATION

In this section processing of image with CBIR system is described. Here two different image databases are used such as MUVIS real image database where 200 images with 10 different categories and MUVIS texture image database where 200 images with 21 different categories exist. MUVIS real image and texture image database is retrieved after extracting color features and texture features. Here RGB histogram(4,4,4) and Lab histogram(8,4,4) is used as color feature and Gray Level Co-occurrence Matrix GLCM (dx=dy=4) and Local Binary Pattern LBP histogram of bin size 12 is used as texture feature. F1 metric which is calculated from precision and recall value and Averaged Normalized Modified Retrieval Rank(ANMRR) metric is employed to measure performance of retrieval images from the database. These two metrics are discussed below:

F1 is measure by harmonic mean of precision and recall. Where F1 Score reaches 1 gives best result and 0 gives worst result.

$$F1 = \frac{(2 \times precision \times recall)}{(precision + recall)}$$

$$precision = \frac{tp}{(tp + fp)} \qquad recall = \frac{tp}{(tp + fn)}$$

where tp= number of retrieved relevant images
fp = number of retrieved non- relevant images
fn = number of relevant images that are not retrieved.

ANMRR:

$$AVR(q) = \frac{(\sum R(k))}{(N(q))}$$ and

$$W = 2 \times N(q)$$

$$NMRR = \left(\frac{(2 \times AVR(q) - N(q) - 1)}{(2 \times W - N(q) + 1)}\right) \leq 1$$

$$ANMRR = \left(\frac{(\sum NMRR)}{Q}\right) \leq 1$$

Where N(q) is minimum number of relevant images (via ground truth) in a set of Q experiments
W is the size of retrieval window.
R(k) is rank of the retrieved image in **K**th relevant retrieval

First real images were retrieved using color and texture features separately, then ANMRR and F1 is calculated. Then 200 real images are compressed with 50%, 40% , 30% , 20% and 10% quality factor and evaluated retrieval performance respectively . Similar process is done with texture images using only texture features.

3.RESULTS

It has been noticed from table 1, 2 and 3 that

Jpeg Compression	Real Image			
	RGB		Lab	
	AF1	ANMRR	AF1	ANMRR
100%	0.2953	0.7446	0.2863	0.7527
50%	0.2430	0.7994	0.2522	0.7892
40%	0.2437	0.7985	0.2545	0.7880
30%	0.2432	0.7995	0.2525	0.7868
20%	0.2438	0.7989	0.2522	0.7898
10%	0.2452	0.7979	0.2632	0.7777

Table 1 :results of Real Images with colore features

after JPEG compression retrieval performance has been reduced in real images. It is because JPEG is lossy compression technique where information of the image is being reduced. Figure 1 and 2 shows ANMRR and AF1 score against compression ratio. It is clear from the figure that after as compression ratio increases from 50% to 100% performance increases. Below 50% compression is

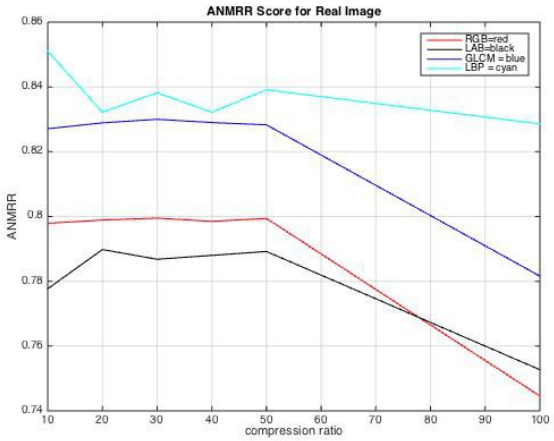

Figure 1: ANMRR score for real Image

is not acceptable because performance of both matric is decreasing. In texture images retrieval performance same as real Images. Performance decrease in accordance with compression ratio. From table 3 has been noticed that ANMRR and AF1 score varies after 50% compression.

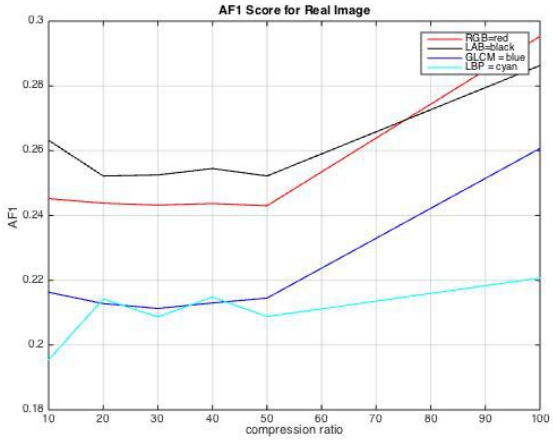

Figure 2 : AF1 score for real Images

Jpeg Compression	Real Image			
	GLCM		LBP	
	AF1	ANMRR	AF1	ANMRR
100%	0.2607	0.7816	0.2207	0.8286
50%	0.2145	0.8283	0.2088	0.8391
40%	0.2130	0.8290	0.2148	0.8322
30%	0.2113	0.8300	0.2087	0.8382
20%	0.2128	0.8289	0.2142	0.8322
10%	0.2163	0.8271	0.1955	0.8510

Table 2 :results of Real Images with texture features

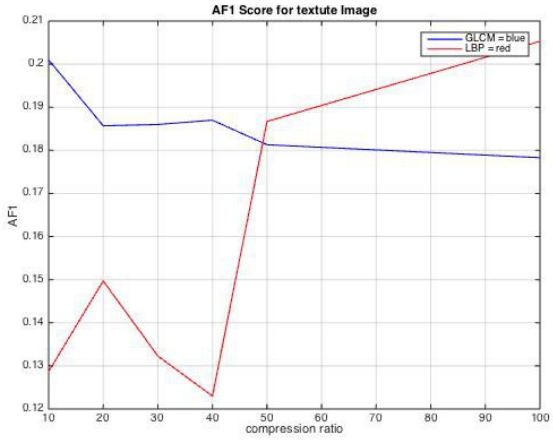

Figure 3: AF1 score for texture Image

From figure 3 and 4 similar performance behavior is noticed like real images. LBP feature fluctuates in between 20% to 30% compression but GLCM is monotonically increasing curve while compression ratio increases.

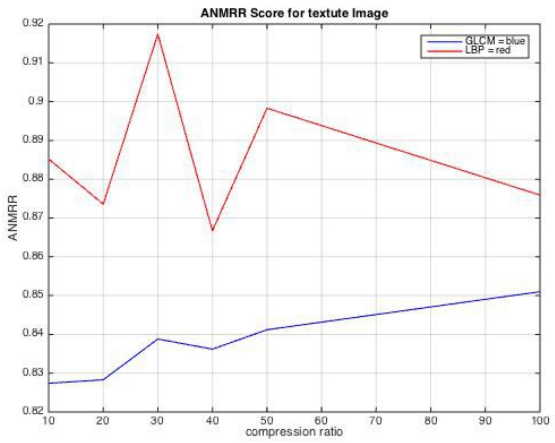

Figure 4: ANMRR score for texture Image

Jpeg Compression	Texture Image			
	GLCM		LBP	
	AF1	ANMRR	AF1	ANMRR
100%	0.1783	0.8510	0.2053	0.8759
50%	0.1813	0.8412	0.1867	0.8983
40%	0.1870	0.8362	0.1230	0.8667
30%	0.1860	0.8388	0.1323	0.9174
20%	0.1857	0.8283	0.1497	0.8736
10%	0.2010	0.8274	0.1287	0.8852

Table 3:results of Texture Images

3. CONCLUSION

Compression and retrieval performance are studied in order to achieve following experimental results

- color feature gives better performance than texture feature.
- Compression reduces retrieval performance in real and texture image databases

Further it has been noticed that ANMRR measure depends on rank of the retrieved images where rank sometimes varies subjectively. It is still not clear whether image database size or the retrieval system effects on ANMRR measure. NMRR evaluating first kth position of retrieved images. Further investigation needed on retrieval using different feature vector.

Appendix

Matlab function used in this project work:
imread,imwrite,strcat,imhist,graycomatrix,rgb2gray,makeform,applycform and sort.

REFERENCES

1. http://www.gnu.org/software/gift
2. "MUVIS",http://www.muvis.cs.tut.fi
3. vismod.media.mit.edu/vismod/demos/photobook