COMPRESION EFFECTS ON COLOR AND TEXTURE BASED MULTIMEDIA INDEXING AND RETRIEVAL

Esin Guldogan, Olcay Guldogan, Serkan Kiranyaz, Kerem Caglar[*] and Moncef Gabbouj

Institute of Signal Processing, Tampere University of Technology, Tampere, Finland.
[*]Nokia Mobile Software, Tampere, Finland
E-mail: guldogan@cs.tut.fi, kerem.caglar@nokia.com

ABSTRACT

This paper presents an evaluation of digital compression effects on content-based multimedia retrieval using color and texture attributes. Subjective evaluation tests that are applied on digital image and video databases using different compression and visual feature extraction techniques have been performed and reported. Simulations show that a satisfactory retrieval performance can be obtained from the compressed databases with 10% compression quality (i.e. 97.6% compression ratio in JPEG). Image retrieval based on HSV color histogram performs better than retrieval based on YUV color histogram in the uncompressed domain, and the other way around in the compressed domain. In general, video retrieval based on color histogram in MPEG-4 compressed databases performs better compared to H.263+ compressed databases. However, retrieval performance from H.263+ compressed databases at lower bit rates is more stable, where it drastically decreases in MPEG-4 compressed databases below 128 Kb/s. Retrieval based on texture features produces more robust performance than retrieval based on color. Subjective tests show that 25% compression quality achieves high compression ratio without loosing significant retrieval performance. The results are particularly relevant to applications in which a mobile device is involved in a multimedia retrieval system.

1. INTRODUCTION

During the recent decades, technological hardware and network improvements have caused a rapid increase in the size of the digital visual information. Besides several benefits and usages, such massive digital visual information has brought the problems of storage and management, which have become the main motivation for researchers in digital compression and multimedia indexing and retrieval areas.

In order to reduce the large space requirement for digital multimedia, several studies have been made on digital compression techniques. Various lossy and lossless compression methods have been proposed and applied to digital multimedia items. Successful digital image and video compression standards have been issued by international organizations, such as JPEG, MPEG-1, MPEG-2, MPEG-4 and H.263+ [1]-[5]. So far such efforts have brought a reasonable solution for the storage problem. Thereafter, solving the problem of handling and management of massive digital visual information has become the main motivation to make multimedia indexing and retrieval an active research area. Extensive studies on this area have revealed more reliable indexing and retrieval schemes that are based on the content characteristics of the multimedia items. Consequently, content-based multimedia indexing and retrieval systems have been designed for various educational and commercial applications, such as MUVIS, QBIC and Photobook [6]-[8]. Due to the storage problem of digital multimedia collections, such systems have the tendency to work on compressed domain. Furthermore, indexing schemes may use the information already available within the compressed data, such as motion information in the video bitstream. Therefore, employing content-based indexing and retrieval techniques via digital multimedia compression concepts may provide several benefits in practice. However, lossy compression methods might cause crucial information loss, which may degrade the results of multimedia retrieval. On the other hand, this data loss causes a filtering effect on the visual attributes, and hence the retrieval performance may be enhanced. Such issues have been the main motivation for studying the effects of compression on color and texture based multimedia retrieval.

In practice, JPEG lossy compression is applied on images in the sample image databases, while H263+ and MPEG-4 (simple profile) compression methods are applied on video sequences in the sample video databases. HSV and YUV color histograms [9], and Gabor Wavelet Transform (GWT) [10] and Gray Level Co-occurrence Matrix (GLCM) [11] methods are employed for indexing and retrieval of experimental multimedia databases. MUVIS framework is used to perform experiments upon which the subjective evaluations are based.

2. EXPERIMENTAL CASES AND SUBJECTIVE EVALUATIONS

To run the experiments, several digital images and video sequences were collected, and the following three base databases were created using MUVIS applications:

- Base Image Database: It is a general-purpose database including 1594 uncompressed color images with various sizes (from 100x100 up to 1200x1000 pixels) and color bit depths (8-24 bits) in the bitmap format obtained from several sources.
- Base Video Database: It is a general-purpose database including 300 video clips in AVI format that were captured in real-time from TV broadcast channels using MUVIS *AVDatabase* application. The video clips are CIF-size (352x288 pixels) and are compressed using MPEG-4 simple profile video encoding method with 2 Mbits/sec bit rate. The duration of video clips varies between 40 and 120 seconds.

- Base Texture Image Database: It is a constraint database including 1512 uncompressed gray-scale images with 166x166 pixels and 8 bits luminance depth in the bitmap format. Images contain textures of mainly 6 different rock classes as in [11].

Several compressed databases were generated for each experimental case by applying JPEG compression to the base image databases, and MPEG-4 and H.263+ compression to the base video database. The experimental databases and associated features are presented in Table 3. Retrieval experiments were performed on MUVIS *MBrowser* application. Euclidean distance is used as the similarity measurement for retrieval.

A group of people evaluated the retrieval performances from each query in each experimental case. The media item in the first rank is not evaluated at all, since it is the same with the queried media item, and thus it causes a constant up-shift in the final ranking grade. Each person in the evaluation group gave a subjective grade for the remaining 11 retrieved items as shown in Figure 1 and Table 2. The given grades are integer numbers within the range [0-5], which have corresponding subjective meanings as presented in Table 1.

Grade	Subjective Meaning
0	No similarity / Not related
1	Slightly related
2	Related
3	Similar
4	Fairly similar
5	Same / Almost identical

Table 1: The subjective meanings of the evaluation grades

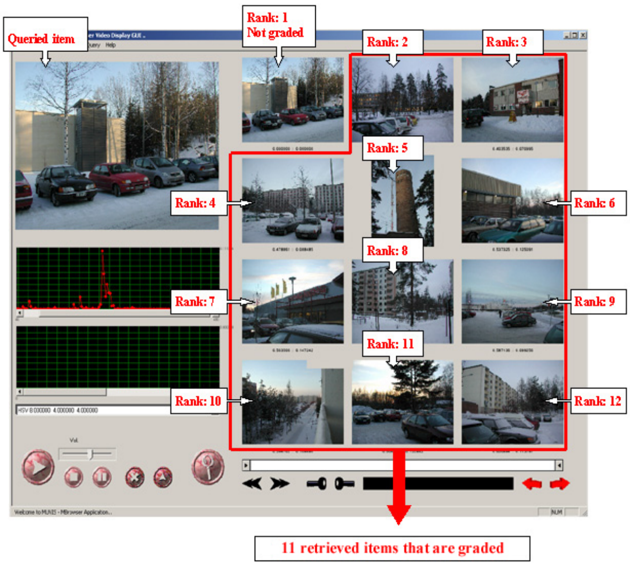

Figure 1: A sample retrieval setup and associated rankings

Rank	2	3	4	5	6	7	8	9	10	11	12
Grade	5	3	4	0	3	2	1	1	1	2	3

Table 2: Grades given by a person for the sample retrieval setup in Figure 1

Each of the 11 graded items has a weight associated with its rank. The *Performance Value (P)* is calculated using (Eq. 1). For instance the *P* value for the sample grading given in Table 2 is 176 over 330. The average performance value within the evaluation group is called the *Query Performance Value (QP)* given in (Eq. 2). The *Overall Query Performance Value (OQP)* for an experimental case is defined as the mean of all *QP* values obtained from the queries of the selected database primitives.

$$P = \sum_{i=2}^{12} G_i \cdot W_i \quad , \qquad (1)$$

where $G_i \in [0,5]$ is the given subjective grade, and $W_i = (13-i)$ is the weight for the item in the rank i.

$$QP = \frac{\sum_{j=1}^{NP} P_j}{NP} \quad \text{and} \quad OQP = \frac{\sum_{k=1}^{N} QP_k}{N} \quad , \qquad (2)$$

where N is the number of queries, and NP is the size of the evaluation group.

3. RETRIEVAL PERFORMANCE RESULTS

The following sections present the retrieval experiments that are performed for subjective evaluations. The size of the evaluation group is 10 (i.e. NP=10).

3.1. Color-Based Image Retrieval Experiments

In the color-based image retrieval experiments, the base image database is used to generate 7 compressed databases whose images are JPEG compressed with a unique compression quality (ratio). So including the base database, a total of 8 image databases are created for retrieval experiments, each of which is used for subjective tests. The properties of the databases and the specific color features used are presented in Table 3. The subjective tests are performed by querying 30 (i.e. N=30) pre-defined images, which are selected from the associated image database where the query is performed, and by calculating *QP* values per image query. Once all subjective tests are completed, the final *OQP* value is calculated using (Eq. 2). Figure 2 illustrates the *OQP* results obtained from the subjective evaluation of the color-based image retrieval experiments.

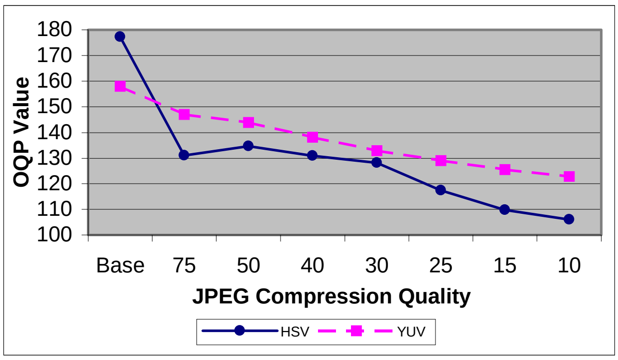

Figure 2: Image retrieval performance based on color histogram

3.2. Color-Based Video Retrieval Experiments

In the color-based video retrieval experiments, the base database is used to generate 2 sets of compressed databases. The first set is compressed by H.263+ encoder and the second set is compressed by MPEG-4 (simple profile) encoder. The subjective tests are performed by querying 15 (i.e. N=15) pre-defined video clips, which are selected from the associated video database where the query is performed, and by calculating *QP* values per video query. After completing all tests, the associated *OQP* value is calculated using (Eq. 2). The *OQP* results from color-based video retrieval are illustrated in Figure 3.

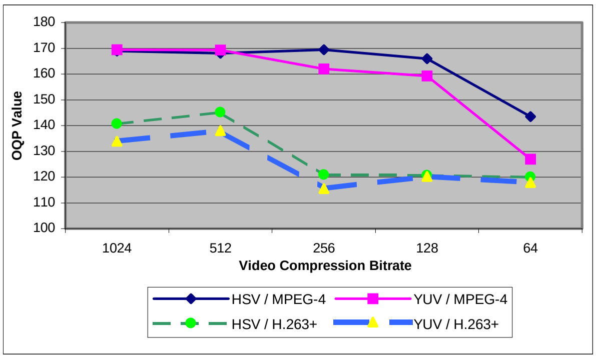

Figure 3: Video retrieval performance based on color histogram

3.3. Texture-Based Image Retrieval Experiments

In the texture-based image retrieval experiments, the base texture image database is used to generate 7 JPEG compressed databases as explained in Section 3.1. The properties of the texture image databases and the texture features used are presented in Table 3. The subjective tests are performed by querying 15 (i.e. N=15) pre-defined images, which are selected from the associated texture image database where the query is performed, and by calculating *QP* values per image query. Once all subjective tests are completed, the final *OQP* value is calculated by (Eq. 2). Figure 4 illustrates the *OQP* results obtained from the subjective evaluation of the texture-based image retrieval experiments.

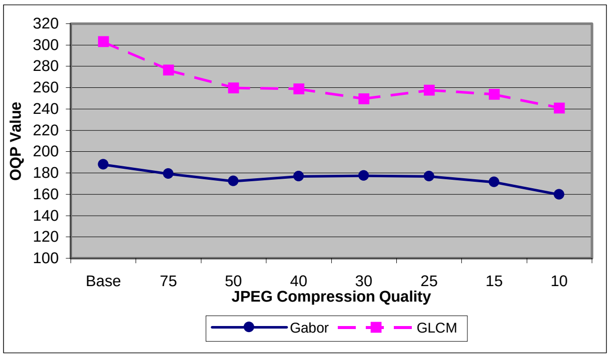

Figure 4: Image retrieval performance based on texture features

4. INTERPRETATIONS, CONCLUSIONS AND FUTURE WORK

Image retrieval experiments using uncompressed and JPEG compressed image databases show that the best retrieval performance based on color histogram can be obtained by querying uncompressed images in uncompressed image databases. This is so far expected since some crucial information might be lost due to lossy JPEG compression, and as a result, retrieval performance is degraded. In both retrieval schemes based on HSV and YUV color histogram, increasing compression ratio gradually decreases the retrieval performance. The subjective tests show that a satisfactory retrieval performance can be obtained from compressed databases with 10% compression quality (97.6% compression ratio). Satisfactory and efficient retrieval performance can be considered as achieving the highest compression ratio without loosing significant retrieval performance. Image retrieval based on HSV color histogram performs better than retrieval based on YUV color histogram in the uncompressed domain, and vice versa in the compressed domain. The reason of this retrieval difference might be explained relying on the design of JPEG encoding. JPEG encoding works in YUV color domain, and transformation into any other color domain such as HSV may increase the degree of degradation. Therefore, HSV color features might yield worse retrieval performance when JPEG compression is applied. From 75% to 30% compression quality, retrieval performance based on HSV color histogram is stable, while it decreases gradually at lower compression qualities. Retrieval performance based on YUV color histogram behaves differently. Considering color attributes in general, 30% quality parameter achieves a satisfactory retrieval performance. Practically, this might be a desirable property in the cases where limited storage and transmission capabilities are encountered, e.g. in applications where a mobile device is involved.

In general, video retrieval based on color histogram in MPEG-4 compressed databases performs better compared to H.263+ compressed databases. However, retrieval performance from H.263+ compressed databases at lower bit rates is more stable, where it drastically decreases in MPEG-4 compressed databases below 128 Kbits/sec bit rate. It might be expected that the video retrieval performance is better in H.263+ compressed databases at very low bit rates (i.e. <64Kbits/sec), since H.263+ encoder used in the experiments is designed for low bit rate applications while MPEG-4 encoder used in the experiments is optimized for better visual quality at high bit rates. Additionally, there are practical differences in the optional MPEG-4 and H.263+ encoder settings, which may change the key-frame selection of the video sequence. Such changes affect the retrieval results, since the indexing scheme of MUVIS is directly based on the key-frames for video feature extraction.

The video retrieval performances based on HSV and YUV color histograms are close to each other in the compressed databases. Retrieval based on HSV color histogram gives slightly better retrieval performance. In general, 512 Kbits/sec can be considered as the optimum compression bit rate for color based video retrieval at high bit rates. For low bit rate applications, 128 Kbits/sec for MPEG-4 compression bit rate and 64 Kbits/sec for H.263+ compression bit rate can be preferable to achieve a satisfactory retrieval performance.

According to the compression effects so far discussed, image retrieval based on texture features gives more robust performance results than retrieval based on color features. The reason of such robustness can be insignificant texture information loss due to degradation in color domain caused by JPEG compression. However, image retrieval experiments based on texture features are performed on constrained databases. Therefore, such results cannot be generalized for overall image retrieval performance based on texture attributes.

Considering the experimental cases, using Gray Level Co-occurrence Matrix (GLCM) feature extraction technique gives higher image retrieval performance than using Gabor Wavelet Transform (GWT) technique with respect to the compression effects. However, retrieval performance based on GWT is more stable. The best image retrieval performance based on texture features can be obtained from the uncompressed database. Subjective tests show that 25% compression quality achieves the highest compression ratio without loosing significant retrieval performance.

For future work, the studies described so far can be extended in order to refine the results and achieve more general and reliable conclusions. There are still several issues that are planned to be the continuation of this study, as listed below:

- Similar experiments will be made to study texture-based image retrieval performance on various multimedia databases.
- Modified performance evaluations such as querying compressed or low-resolution multimedia items in an uncompressed or high-resolution database will be made especially for mobile and Internet applications.
- Experiments can be made for studying compression effects on indexing and retrieval performances based on other features such as audio and shape.

5. REFERENCES

[1] CCITT Recommendation T.81, "Information Technology - Digital Compression and Coding of Continuous-tone Still Images", Sept. 1992.

[2] ISO/IEC IS 11172-2, "Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 2: MPEG-1 Video", 1993.

[3] ISO/IEC DIS 13818-2, "Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Video", Dec. 2000.

[4] ISO/IEC 14496-2:2001, "Information Technology - Coding of Audio-Visual Objects - Part 2: Visual ", July 2001.

[5] ITU-T Recommendation H.263, "Video Coding For Low Bit Rate Communication", Feb. 1998.

[6] M. Gabbouj, S. Kiranyaz, K. Caglar, B. Cramariuc, F. Alaya Cheikh, O. Guldogan, and E. Karaoglu, "MUVIS: A Multimedia Browsing, Indexing and Retrieval System", *Proceedings of the IWDC 2002 Conference on Advanced Methods for Multimedia Signal Processing*, Capri, Italy, Sep. 2002.

[7] W. Niblack, R. Barber and et al., "The QBIC Project: Querying Images by Content Using Color, Textures and Shape", *Proc. SPIE Storage and Retrieval for Image and Video Databases*, pp. 124-128, Feb. 1996.

[8] A. Pentland, R. W. Picard, S. Sclaroff, "Photobook: Content-Based Manipulation of Image Databases", *International Journal of Computer Vision*, 1996.

[9] M. J. Swain, D. H. Ballard, "Color Indexing", *International Journal of Computer Vision*, Vol.7, No. 1, 1991.

[10] W. Y. Ma, B. Manjunath, "Texture Features for Browsing and Retrieval of Image Data", *IEEE Trans. PAMI, vol. 18, pp 837-842*, Aug. 1996.

[11] M. Partio, B. Cramariuc, M. Gabbouj, A. Visa, "Rock Texture Retrieval Using Gray Level Co-occurrence Matrix", *Proc. of 5th Nordic Signal Processing Symposium*, Oct. 2002.

Feature Ext. Method	Feature Ext. Param.	Dbs	Comp. Method	Comp. Param. Quality, Comp. Rate/ Brate, Size, Frame Rate
HSV Color Hist.	Hist. BINS 8 - 4 - 4 8 - 8 - 8 16 - 1 - 1 1 - 1 - 16 1 - 16 - 1	Base Image	Uncompressed	
			Lossy JPEG	75% Quality, 87% comp. 50% Quality, 92% comp. 40% Quality, 93.2% comp. 30% Quality, 94.4% comp. 25% Quality, 95.2% comp. 15% Quality, 96.7% comp. 10% Quality, 97.6% comp.
		Base Video	MPEG-4	1024 Kbits/sec, CIF, 25 fps 512 Kbits/sec, CIF, 25 fps 256 Kbits/sec, QCIF, 25 fps 128 Kbits/sec, QCIF, 15 fps 64 Kbits/sec, Sub-QCIF, 10 fps
			H.263+	1024 Kbits/sec, CIF, 25 fps 512 Kbits/sec, CIF, 25 fps 256 Kbits/sec, QCIF, 25 fps 128 Kbits/sec, QCIF, 20 fps 64 Kbits/sec, QCIF, 20 fps
YUV Color Hist.	Hist. BINS 8 - 4 - 4 8 - 8 - 8 16 - 1 - 1 1 - 1 - 16 1 - 16 - 1	Base Image	Uncompressed	
			Lossy JPEG	75% Quality, 87% comp. 50% Quality, 92% comp. 40% Quality, 93.2% comp. 30% Quality, 94.4% comp. 25% Quality, 95.2% comp. 15% Quality, 96.7% comp. 10% Quality, 97.6% comp.
		Base Video	MPEG-4	1024 Kbits/sec, CIF, 25 fps 512 Kbits/sec, CIF, 25 fps 256 Kbits/sec, QCIF, 25 fps 128 Kbits/sec, QCIF, 15 fps 64 Kbits/sec, Sub-QCIF, 10 fps
			H.263+	1024 Kbits/sec, CIF, 25 fps 512 Kbits/sec, CIF, 25 fps 256 Kbits/sec, QCIF, 25 fps 128 Kbits/sec, QCIF, 20 fps 64 Kbits/sec, QCIF, 20 fps
GWT	Orient.: 4 Scale: 3 Side: 45	Base Text. Image	Uncompressed	
			Lossy JPEG	75% Quality, 81% comp. 50% Quality, 86.2% comp. 40% Quality, 88% comp. 30% Quality, 90% comp. 25% Quality, 91% comp. 15% Quality, 92.5% comp. 10% Quality, 93.8% comp.
GLCM	Distances 1 - 1 5 - 5 12 - 12 24 - 24	Base Text. Image	Uncompressed	
			Lossy JPEG	75% Quality, 81% comp. 50% Quality, 86.2% comp. 40% Quality, 88% comp. 30% Quality, 90% comp. 25% Quality, 91% comp. 15% Quality, 92.5% comp. 10% Quality, 93.8% comp.

Table 3: Experimental Cases