MIRROR: An Interactive Content Based Image Retrieval System

Ka-Man Wong, Kwok-Wai Cheung, and Lai-Man Po

Abstract *- A content based image retrieval system, called MPEG-7 Image Retrieval Refinement based On Relevance feedback (MIRROR), is developed for evaluating MPEG-7 visual descriptors and developing new retrieval algorithms. The system core is based on MPEG-7 Experimentation Mode (XM) with web-based user interface for query by image example retrieval. A new Merged Color Palette approach for MPEG-7 dominant color descriptor similarity measure and relevance feedback are also developed in this system. Several MPEG-7 visual descriptors are adopted in MIRROR for performance comparison purpose. The system can be accessed at http://www.ee.cityu.edu.hk/~mirror/.*

Index Terms *- Image Databases, Content Based Image Retrieval, Merged Palette Histogram, Relevance Feedback, MIRROR.*

I. INTRODUCTION

Application of current content based image retrieval (CBIR) is limited and not satisfactory as the intrinsic gap between high level concepts and low level features is not considered in many CBIR systems. Users have to express their requirements in terms of low level image features. MIRROR (MPEG-7 Image Retrieval Refinement based On Relevance feedback) is implemented to develop techniques to address this problem. MIRROR supports various MPEG-7 [1] visual descriptors for representation and extraction of image features. Using MPEG-7 standards also makes interoperating and comparing MIRROR with other MPEG-7 compliant systems possible.

In addition, MIRROR uses relevance feedback (RF) [2] techniques, which takes into account the user's feedback during the retrieval process to effectively capture the user's high level query and concepts. The user's information is used to dynamically update the weights given to low level features. Successful application of relevance feedback technique in CBIR system highly depends on the use of representative image features in the feedback process.

On the other hand, the system is also very suitable for new algorithm development and evaluation. A new merged-color histogram (MCH) approach for similarity measure and relevance feedback feature generation was developed in MIRROR to enhance the retrieval accuracy on using MPEG-7 Dominant Color Descriptor (DCD). In the following sections, the MIRROR system architecture, system features and experimental results on new MCH based techniques for DCD will be presented.

The work described in this paper was substantially supported by a grant from the Research Grants Council of Hong Kong, City University of Hong Kong, Hong Kong SAR, China. [Project No. 9040602 (CityU 1121/01E)].

K. M. Wong and L. M. Po are with City University of Hong Kong, Hong Kong SAR, China
(email: w.carlos@student.cityu.edu.hk, eelmpo@cityu.edu.hk).
K. W. Cheung is with Chu Hai College, Hong Kong SAR, China
(email: kwcheung@chuhai.edu.hk).

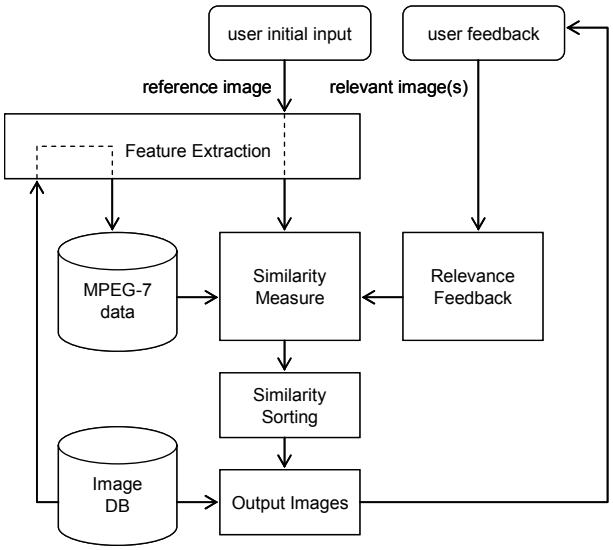

Fig. 1: System Architecture of MIRROR

II. SYSTEM ARCHITECTURE

MIRROR consists of three main modules as shown in Fig.1. Feature Extraction module extracts the descriptor from the original image and converts the information into MPEG-7 data stream. Similarity Measure module uses the user reference image as a query to search relevant images in the database based on the pre-generated MPEG-7 data. Relevance Feedback module receives user feedback on relevant retrieval and generates a new query for Similarity Measure module for a new retrieval.

The extraction tools of MIRROR are based on MPEG-7 Experimentation Model (XM) [3]. The image features are extracted from the image collection using the extraction tools for indexing and subsequent image retrieval. Image features are described by MPEG-7 Description Definition Language (DDL). Multiple MPEG-7 data streams are required if different descriptors are employed simultaneously. The MPEG-7 data stream is usually generated offline and stored in files for subsequent image retrieval process due to the huge amount of image data involved.

Each image in the database is represented by various descriptors. Given a query image and selected descriptor, similarity retrieval involves searching the database for similar features as the input query by reading the corresponding MPEG-7 data. MIRROR supports MPEG-7 defined color descriptors (Dominant Color - DCD, Color Layout - CLD, Scalable Color - SCD and Color Structure - CSD) and one texture descriptor (Edge Histogram - EHD) for CBIR similarity measures. Various matching tools are defined for different descriptors and description schemes. Particularly, MIRROR improves the retrieval performance by using a new matching tool -

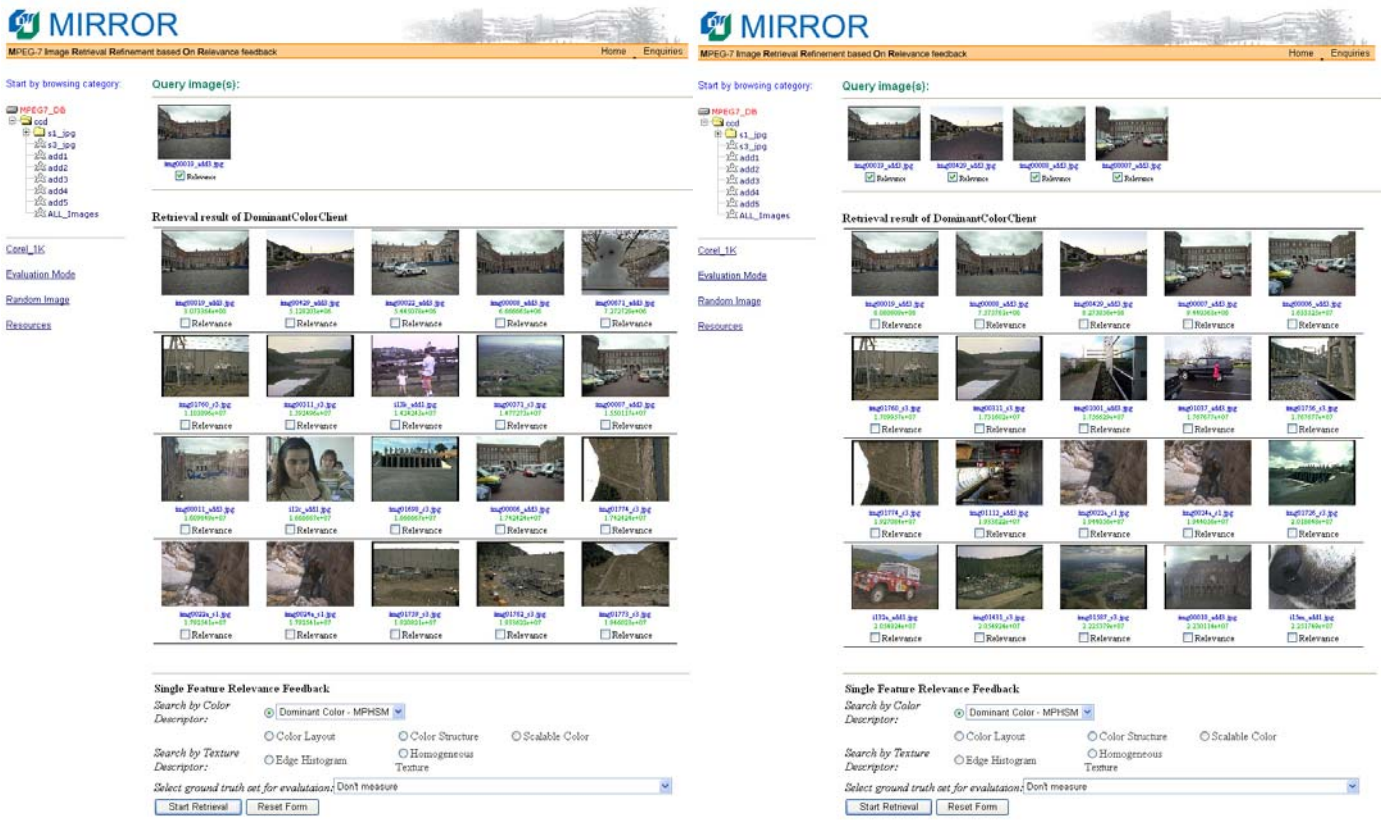

Fig. 2. (a) The reference query image and the initial query results.
(b) Relevance feedback result after 4 more images are selected by the user from the initial retrieval result.

Merged Palette Histogram Similarity Measure for DCD, which will be discussed in section III.

MIRROR employs the query by example paradigm. Fig. 2 shows the web-based user interface of MIRROR. To start the search, the user chooses a reference image from a set of example images. The user may browse the category tree (the top left column of Fig. 2a) for other sets of images. After the user selects the reference image and search method, the system returns the images with the highest ranks (the right column of Fig. 2a). The query image is shown at the top of the screen. Retrieved images are shown at the center of the screen in similarity descending order. This allows users to select relevant images to guide the next search. The system usually retrieves relevant images and false matches. The user checks the images that are more relevant with what he/she is looking for and iterates the process until the desired images are found. During interactive search, the user selects or unselects the example image(s). The user-selected relevant images appear in the upper row of the screen as shown in Fig. 2b.

III. MERGED COLOR PALETTE TECHNIQUES FOR DCD

One of the most significant features of MIRROR is that it supports Merged Palette Histogram Similarity Measure (MPHSM) [4] for enhancing the retrieval accuracy on using MPEG-7 DCD. In this new method, a common palette for two DCD descriptors is generated by merging the selected relevant DCDs' histogram bins using a threshold T_d. Let $F_1 = \{\{c_{1i}, p_{1i}\},\ i = 1,..., N_1\}$ and $F_2 = \{\{c_{2i}, p_{2i}\},\ i = 1,..., N_2\}$ be two DCDs with histograms similar to the ones as shown in the left hand side of Fig. 3. The common palette is generated by searching two closest colors between the two palettes. If this minimum distance is less than or equal to the threshold T_d, the two colors will be merged as

$$c_{m(i,j)} = \frac{p_{1i}c_{1i} + p_{2j}c_{2j}}{p_{1i} + p_{2j}}. \qquad (1)$$

This process continues until the minimum distance is greater than the threshold T_d. A common palette $\{\{c_{mi}\}, i = 1,..., N_m\}$ with N_m colors ($N_m \le N_1 + N_2$) is then generated with the merged colors and unmerged colors from the two palettes as shown in the middle of Fig. 3. This merged palette forms a common color space for the two histograms. The histograms of F_1 and F_2 become

$$F_{1m} = \{\{c_{mi}, p_{1mi}\},\ i = 1,..., N_m\}, \qquad (2)$$

$$F_{2m} = \{\{c_{mi}, p_{2mi}\},\ i = 1,..., N_m\} \qquad (3)$$

where $p_{1mi} = p_{1i}$ if the distance between c_{mi} and c_{1i} is less than or equal to T_d. Otherwise $p_{1mi} = 0$. As the histograms F_{1m} and F_{2m} are based on a common color palette, the histogram intersection method

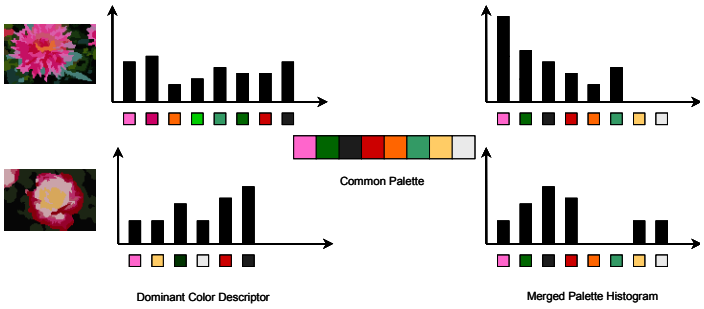

Fig. 3. Merged Palette Histogram using histogram intersection directly.

is used to calculate their similarity. The distance is the non-intersecting area of these two histograms as shown in the right hand side of Fig. 3. This non-intersection histogram area can be represented as

$$D(F_{1m}, F_{2m}) = 1 - \sum_{i=1}^{N_m} \min(p_{1mi}, p_{2mi}) \quad . \qquad (4)$$

The smaller the distance $D(F_{1m}, F_{2m})$, the more similar the two images. The maximum distance is 1. Experimental results of DCD using this new distance measure as compared with the original DCD quadratic histogram distance measure [5] (DCD-QHDM) will be given in section IV.

To further improve the retrieval accuracy, MIRROR also supports relevance feedback. In relevance feedback based CBIR, human and computer interact to refine high level queries to representations based on low level features. MIRROR uses conventional histogram weighting approach [2] for MPEG-7 descriptors of CLD, SCD, CSD and EDH in Relevance Feedback module. However, this conventional method cannot be directly applied on DCD due to the non-fixed color space of DCD. In order to tackle this problem, a merged color palette approach [6, 7] is developed in MIRROR for DCD based Relevance Feedback module. For instance, consider when two or more images were selected by the user as relevant. DCD descriptor will be extracted from the images as the feature data. Different relevant images generate different color palettes. Different color palettes are merged together by Merged Color Histogram method. In general, similar colors among different images will be grouped together. Euclidian distance is used to determine the similar colors. If distance d is less than the threshold T_d, the colors will be considered as similar. Then, a new color entry (c_n) is created by merging the two similar colors and is given by

$$c_n = \frac{p_1 c_1 + p_2 c_2}{p_1 + p_2} \qquad (5)$$

where c_1 and c_2 refer to the colors with color distance less than T_d. Weighting of the new color entry is the sum of the weighting of two similar colors, i.e., $p_n = p_1 + p_2$. After forming common palette from the selected images, the number of color in the palette is checked to see if it exceeds the maximum. If so, unimportant colors, the ones with lowest percentage, will be removed to keep the palette size (number of colors) less than or equal to eight, the maximum number of colors defined in MPEG-7 DCD. Normalization will be performed after generation of each merged color palette. All color histogram bin sizes are added to find the sum p_t. Each color bin will be divided by this sum to complete the normalization. This new DCD's histogram is used to represent all the selected relevant images and used to perform a new retrieval for improving the accuracy. Experimental results will be given in the next section.

IV. SYSTEM PERFORMANCE

Currently, MIRROR includes MPEG-7 Common Color Dataset and Corel_1k dataset. The retrieval metric, Averaged Normalized Modified Retrieval Rank (ANMRR) [8], is used to measure the retrieval accuracy. ANMRR measures the overall performance by averaging the result from each query. Both non-interactive (single image query) and interactive (improvements using relevance feedback) image retrieval performance using various descriptors are summarized in Table I and II. According to the second columns of Table I and II, DCD-QHDM performs worst among all color descriptors. With MPEG-7 Common Color Dataset, ANMRR of DCD-QHDM is 0.0582 worse than a more compact descriptor CLD. Using MPHSM, the difference is reduced to 0.0352. With Corel_1k dataset, MPHSM even outperforms both DCD-QDHM and CLD in ANMRR by 0.1522 and 0.0054, respectively. These results show that MPHSM makes searching with DCD as accurate as using other compact descriptors. Both CLD and DCD are very compact. If they are used to describe a 16Kbytes image, the overhead is smaller than 0.1%. As ANMRR rate of approximately 0.24 may be considered as a good result, both CLD and DCD with MPHSM are good for small storage and fast retrieval.

The performance improvement using relevance feedback can be measured in MIRROR system and the results are given in the third and fourth columns of Table I and II. In Table I, MPH-RF (Merged Palette Histogram with Relevance Feedback) gives ANMRR improvement of 0.0717 to 0.0852 after 3 iterations. The retrieval performance using DCD with MPHSM and MPH-RF is improved up to 0.1752 in ANMRR.

The result outperforms CLD with or without RF. Also the result is only slightly interior to SCD without RF by 0.0107. Histogram weighting technique improves the retrieval results of different descriptors by 0.0284 to 0.0806. Fig. 4 visually shows the effectiveness of using MIRROR's relevance feedback retrieval. Fig. 4a is the query image and Fig. 4b is the ground truth set of the query image. Fig. 4c to 4e are the initial retrieval result, the result after one user's feedback, and the result after two user's feedback, respectively. These results show that the number of images belonging to the ground truth set of the initial query image increases in the retrieval result after user's feedback. The measurements can be easily performed in MIRROR with use of the evaluation mode.

V. CONCLUSION

MIRROR is an interactive CBIR system using relevance feedback algorithm with novel similarity measure to increase the usability and retrieval accuracy. MIRROR uses search by example paradigm. The system is designed with flexibility for CBIR development and experimentation. Comparisons of performance among all methods used in the system are performed show that the combination of the MPHSM similarity measure and the MPH-RF relevance feedback

TABLE I
SUMMARY OF ANMRR RESULTS USING MPEG-7 CCD DATASET

Descriptor	Initial Search ANMRR	ANMRR after 3 RF	RF Improvement
DCD-MPHSM	0.2604	0.1752	0.0852
DCD-QHDM	0.2834	0.2117	0.0717
CLD	0.2252	0.1814	0.0438
CSD	0.0399	0.0115	0.0284
SCD	0.1645	0.1019	0.0626
EHD	0.3217	0.2411	0.0806

TABLE II
SUMMARY OF ANMRR RESULTS USING COREL_1K DATASET

Descriptor	Initial Search ANMRR	ANMRR after 3 RF	RF Improvement
DCD-MPHSM	0.3946	0.3298	0.0648
DCD-QHDM	0.5468	0.4900	0.0568
CLD	0.4000	0.3571	0.0429
CSD	0.3246	0.2366	0.0880
SCD	0.3552	0.3276	0.0276
EHD	0.5089	0.4374	0.0715

can increase the performance of MPEG-7 Dominant Color Descriptor.

REFERENCES

[1] Special issue on MPEG-7, IEEE Trans. CSVT, vol. 11, no. 6, June 2001.

[2] Y. Rui, T. S. Huang, M. Ortega and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," IEEE Trans. CSVT, vol. 8, no. 5, pp. 644-655, September 1998.

[3] *MPEG-7 Visual Part of Experimentation Model Version 8.0*, A. Yamada, M. Pickering, S. Jeannin, L. Cieplinski, J.-R. Ohm, and M. Editors, Eds., Oct. 2000, ISO/IEC JTC1/SC29/WG11/N3673.

[4] L. M. Po and K. M. Wong, "A new palette histogram similarity measure for MPEG-7 dominant color descriptor," accepted for publication in Proceedings of 2004 IEEE International Conference in Image Processing, Oct. 2004, Singapore.

[5] J. Hafner, H. S. Sawhney, W. Equitz, M. Flickner, and W. Niblack, "Efficient color histogram indexing for quadratic form distance functions," IEEE Trans. Patt. Anal. Mach. Intell., vol. 17, pp. 729-736, July 1995.

[6] K. M. Wong and L. M. Po, "MPEG-7 Dominant Color Descriptor based relevance feedback using Merged Palette Histogram," in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing 2004. vol. 3, pp. 433-436. 17-21 May 2004. Montreal, Quebec, Canada.

[7] K. M. Wong, C. H. Cheung, T. S. Liu and L. M. Po, "Dominant Color Image Retrieval Using Merged Histogram," in Proc. IEEE ISCAS, vol. 2, pp. 908-911. 25-28 May, 2003. Bangkok, Thailand.

[8] Subjective Evaluation of the MPEG-7 Retrieval Accuracy Measure (ANMRR), ISO/WG11, Doc. M6029, May 2000.

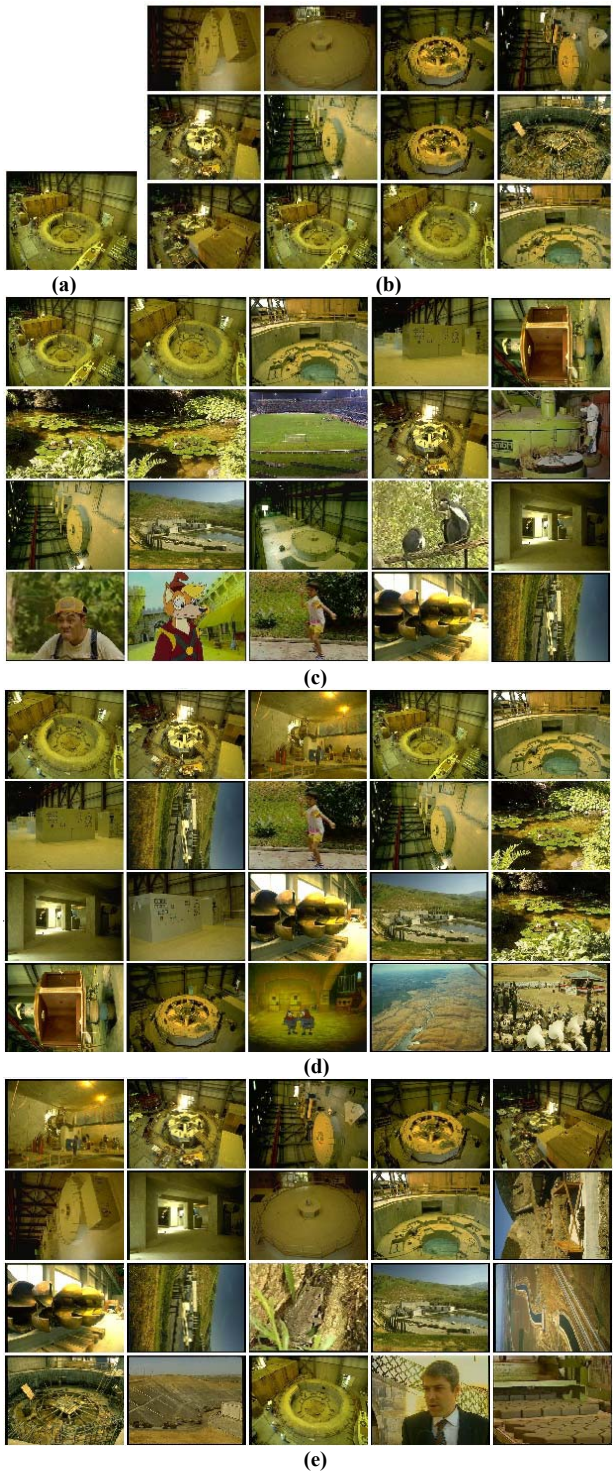

Fig. 4. (a) The reference image of MPEG-7 CCD query #24. (b) Ground truth images of MPEG-7 CCD query #24. (c) Initial retrieval result of MPEG-7 CCD query #24 using QHDM with MPH-RF – 5 of 12 ground truth images found in first 20 retrieval (NMRR = 0.5125). (d) First relevance feedback result of MPEG-7 CCD query #50 using QHDM with MPH-RF– 6 of 12 ground truth images found in first 20 retrieval (NMRR=0.3738). (e) Second relevance feedback result of MPEG-7 CCD query #50 using QHDM with MPH-RF– 9 of 12 ground truth images found in first 20 retrieval (NMRR=0.1963).