MPEG-7 Dominant Color Descriptor Based Relevance Feedback Using Merged Palette Histogram

Ka-Man Wong and Lai-Man Po

w.carlos@student.cityu.edu.hk and eelmpo@cityu.edu.hk
Department of Electronic Engineering, City University of Hong Kong,
83 Tat Chee Avenue, Kowloon, Hong Kong SAR of China.

Abstract - This paper proposes techniques to improve the effectiveness of Relevance Feedback (RF) based image retrieval using MPEG-7 Dominant Color Descriptor (DCD). The conventional RF query point moving techniques could not be directly applied on DCD due to the difference of the color spaces used in each DCD histogram. In order to tackle this problem, a new Merged Palette Histogram (MPH) technique is proposed for generating new query DCD histogram from the relevant image set. The new method has been implemented on a MPEG-7 XM with use of an image database containing 1,000 images. The effectiveness of this new MPH-RF technique using different similarity measures is demonstrated by experimental results in the terms of Averaged Normalized Modified Retrieval Rate (ANMRR) and visual retrieved images comparison.

1. Introduction

Digital libraries containing large quantity of video and images are rapidly becoming available. To avoid the expense and limitations of text, there is considerable interest to develop content-based image retrieval (CBIR) techniques based on automatically extractable visual attributes such as colors or textures. A CBIR system facilitates the formulation of visual queries by novice users in which a query can be initiated simply by supplying a sample image. The system extracts the visual features from the query image and matches it against those stored in the database. The result is a set of images that are similar to the query image rather than an exact match. Color is one of the most used image features.

Color is a straight forward visual feature and is widely used in CBIR systems. A suitable color descriptor is one that is accurate, compact and natural. It is also essential to use a uniform color space such as the *CIE Luv* color space that matches user's ability to perceive and differentiate colors in natural images. Conventional color histogram has been widely used in color-based CBIR systems [1,2]. It is attractive because it is very simple and the descriptor storage is very straight forward. However, it is not a very efficient approach compared to new methods developed in recent years. For example, newly developed DCD provides more compact and accurate description on the distribution of color, which only stores the most representative colors. Also the descriptor size is very small, so DCD similarity measure is more computationally efficient.

Although the color-based methods perform surprisingly well, their retrieval performance is still limited. The main reason is because the color representation is low-level. The representation merely captures the dominant colors of an image's contents, but not its semantics. As a result, there are often ambiguities in similarity matching. One way to tackle this problem is to involve the users in identifying relevant images from the list of images retrieved. This information is then used to improve the quality of retrieval in a process known as relevance feedback (RF). [3]

RF has been applied effectively to text-based systems. However, it has not been found to be very effective in color-based systems. The prime reason is because it is difficult to extract meaningful features from color histogram representations of relevant images, especially in ordinary histogram approaches.

MPEG-7 DCD [4] is one of the most commonly used descriptor. It is an effective, compact, and intuitive representation of salient colors in an image region. However, the representative colors of DCD are computed from each image instead of being fixed in the color space. Thus, conventional weight update techniques cannot be directly used in DCD-based relevance feedback image retrieval for improving the retrieval accuracy. A new merged palette approach for DCD will be proposed in this paper.

The new histogram palette merging approach will be very suitable for relevance feedback image retrieval technique because it is based on a common palette generated by merging the dominant color histograms from the selected relevance image. This merged palette formed a new query for similarity measure. Thus the new query represents the dominant color among the selected relevance images.

This paper describes the techniques for image retrieval with RF using merged palette histogram. The rest of the paper is organized as follows. Next section analyzes existing approach in color-based image retrieval and the related color image techniques will be described. Section 3 discusses the proposed merged histogram approach for relevance feedback. The implementation and testing are described in Section 4. Finally Section 5 contains our concluding remarks.

2. Relevance feedback techniques for color based image retrieval

Color has been widely used in content-based image retrieval systems. The problem with using color is that its representation is low level and hence its retrieval effectiveness is limited. Since the selected query image might not be exactly matched with the user's perception. Relevance feedback techniques are developed to refine the query by the user's feedback based on their perception. Although relevance feedback can improve the retrieval effectiveness in a interactive way, it requires suitable feature representation and query update algorithm.

2.1. Relevance feedback process

Relevance feedback is an interactive process that starts with normal CBIR that the user input a query, and then the system extracts the image feature and measure the distance with images in the database. An initial retrieval list is then generated. The user reply the system that the retrieved images is relevant or not, and the system refine the query based on the user select relevance images. A new retrieval list is then generated based on the refined query. User can choose the relevance image to further refine the query, and this process can be iterated many times until the user find the desired images.

2.2. Relevance feedback technique for color histograms

While conventional color histogram is used, the color space of images are fixed, relevance feedback query update is normally form by setting weights on the selected relevance images. The new query is then calculated by linear combination on matched dimension in the histogram. The weights will be given by the importance of the relevant images.

One conventional approach is taking the average on the histograms. Consider *k* images selected for relevance feedback with normalized histograms

$H_1 = \sum_i p_{1i}, H_2 = \sum_i p_{2i}, ..., H_k = \sum_i p_{ki}$

and the histogram averaging can be done by :

$$H' = \frac{\sum_i \sum_{j=1}^{k} p_{ji}}{k} \quad (1)$$

The method takes average on same dimension in the histogram. So it is effective only if the color space of both descriptors is fixed.

2.3. Color Descriptor

In MPEG-7 DCD, the colors of DCD are computed from each image instead of being fixed in the color space. It is normal that the images do not have any exactly matched colors. Thus, conventional histogram weight update approach cannot be directly used in DCD-based relevance feedback image retrieval for improving the retrieval accuracy. Consider the following example with 2 selected relevance images with DCD H_1 and H_2:

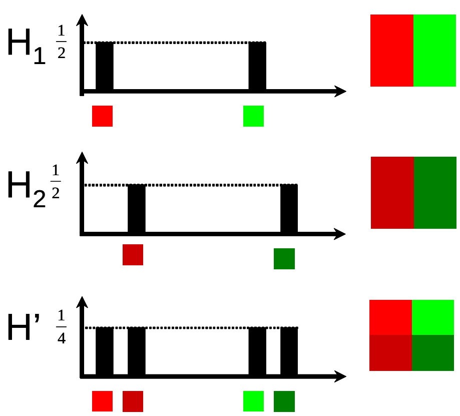

Fig. 1 Using histogram averaging on DCDs

In the example above, the color of selected images are very similar. While using DCD, this approach is equitant to put 2 images together directly, because there is no exactly matched color. By the histogram averaging, the dimension of updated query changed, this may not hold the condition of a valid DCD. (i) the number of color in updated query may exceed the limit of the number of colors defined by MPEG-7. (ii) similar colors considered as different dimension, and cause the problem of ordinary QHDM [4], by definition of DCD, the minimum distance in the palette should larger than a threshold.

These two properties might decrease the performance, since the definitions are used to provide compact description, effective similarity measure, and prevent the problems of ordinary QHDM approaches. [2]

2.4. Color Space

Digital images are normally represented in RGB color space, which is the color space used by CRT monitors. However, RGB color space is perceptually non-uniform. To overcome this problem, perceptually uniform color spaces, such as *CIE Luv* color space, have been developed. *CIE Luv* color space is composed of three components in which *L* defines the luminance, while *u* and *v* define the chrominance. It is intended to yield a perceptually uniform spacing of colors in which any two colors that is equal in distance is perceived as equal in difference by the viewers. It is capable of representing completely any color from the visible frequency domain. *CIE Luv* satisfies the uniformity, completeness, naturalness and compactness properties, and thus is chosen for color representation and retrieval.

3. Proposed Merged Palette Histogram for relevance feedback

3.1. Component representation

Assume that an image database *DB* is composed of *n* images, $DB = \{I_1, I_2, ..., I_n\}$. For a query *Q*, the retrieval ranking can be made by measuring the distance $D(Q, I)$ between the query *Q* and images I in the database. Where a query *Q* is obtain by the feature extraction algorithm $Q = F(O)$ of content based image retrieval, and *O* is the user input for the query, this can be a image, sketch, or user defined histogram. After the retrieval process, an initial ranking is generated. User can choose the image that they perceived as relevance. A query update process modifies the query into $Q' = U(Q, I_R)$ where I_R is the user select relevance list. The retrieval process repeat using updated query *Q'* instead of original query *Q*.

3.2. Query update using merged palette histogram

The proposed merged histogram approach is applied in the query update process, since we use Dominant Color Descriptor, we suppose the initial query $Q = \{(c_i, p_i)\}, (i = 1,2,..., N)$

And the user selected *r* image as relevance, the selected list $I_R = \{I_1, I_2, ...I_r\}$. Updated query $Q' = U(Q, I_R)$ can be obtained by the following process:

Step 1: Feature extraction

We perform DCD feature extraction among the selected images I_R, then pseudo query

$$Q'_o = \{Q, Q_1, Q_2, ..., Q_r\}$$
$$= \{\sum_i (c_{qi}, p_{qi}), \sum_i (c_{q1i}, p_{q1i}), ..., \sum_i (c_{qri}, p_{qri})$$

is obtained. Since the colors are different in the images, the number of color is increased much higher than the maximum number of colors defined in DCD, also there would have similar colors. This is still not suitable for an updated query. We use Merged Palette Histogram technique to reduce number of color and form a new query representing the selected relevance images.

Step 2: Palette Merging

We search the most similar color pair with minimum Euclidian distance in *CIE Luv* color space. Then the most similar color pair will be merged together by the following equation.

$$c_m = \frac{p_i c_i + p_j c_j}{p_i + p_j} \quad (2)$$

$$p_m = p_1 + p_2$$

This process iterate until the minimum distance between the color is larger than the threshold T_d which is defined as the maximum distance between colors that human perception consider as similar. In [4], it is suggest that T_d can be 10-15. Fig.2 describes the Palette Merging.

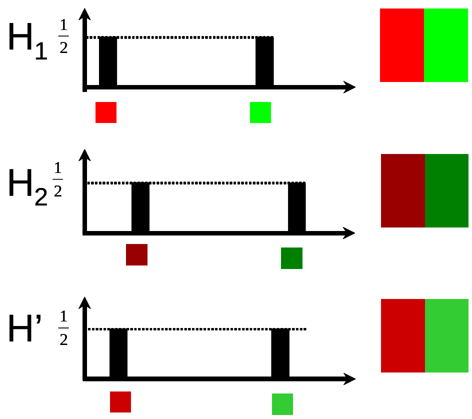

Fig. 2 Using Merged Palette Histogram on DCDs

Where H_1, H_2 are selected relevance feed back, and the color is within the threshold T_d (a pair of red, and a pair or red color), the resulting colors lay between the matched colors. The minimum color distant in the palette is still remains $>T_d$. So it is a more natural description on these selected images than that of conventional histogram averaging approach. So, this method should benefit to all similarity measure metrics such as DCD-QHDM.

Step 3: Approximation

Since the size of palette merged histogram would still larger than maximum number of color defined by DCD (8 colors), approximation is applied to reduce the number of color. This can be done by repeating step 2 that further merge the histogram with desired number of color as the stopping criteria. But it is not very efficient because the number of color will increase by the number of selected images, and the iteration will take longer, also the color will not as accurate as using T_d as the stopping criteria. One possible solution is by sorting the histogram and keeps only the most significant colors as the new query. The sum of p_i in 9^{th} and after dominant colors only has at most 11%, and in most cases each color has fewer than 3%. It will not affect the retrieval.

Step 4: Re-normalization

The merged palette histogram process merge images to form a new query, the histogram size is changed. Re-normalization is done to obtain the new query *Q'*. It can be done by diving the percentages by the histogram sum.

4. Experimental Results

In this section, we present the performance of the proposed MPH relevance feedback query update for both QHDM and MPHSM similarity measure. In our simulations, a database with 1,000 color images from COREL without region segmentation is used to test these similarity measures. These images are divided into 10 categories based on semantic concepts and ground truth sets of 20 sample query images from different categories are also defined. The ground truth sizes among these selected images are about 8~100. In addition, this experiment is implemented on a MPEG-7 Experimental Model (XM) [6] by modifying the dominant color descriptor client application program. The threshold for similar colors is set to T_d = 15 and α = 1.2 for dominant color descriptor generation from the image database and the similarity measures of DCD-QHDM and MPHSM.

In the retrieval process, a sample image is selected and used as the initial query for each query, and we assume the user won't make a wrong feedback. In each feedback, user can browse the first 20 retrieved images, relevance images will be chosen if the image is in the ground truth set of the initial query image. And the relevance set is accumulated in each feedback.

The result is then measured by Averaged Normalized Modified Retrieval Rate (ANMRR) for 3 passes. Also we simulate the result for both MPHSM[5] and QHDM, which show improvement of the relevance feedback on different similarity measure.

To measure the effectiveness of the improvement on MPH relevance feedback, we use the MPEG-7 retrieval metric the Normalized Modified Retrieval Rank (NMRR) [4]. NMRR not only indicates how many of the correct items are retrieved, but also how highly they are ranked among the retrieved items. NMRR is defined by

$$NMRR(q) = \frac{\left(\sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)}\right) - 0.5 - \frac{NG(q)}{2}}{K(q) + 0.5 - 0.5 * NG(q)} \quad (3)$$

where $NG(q)$ is the size of the ground truth set for a query image q, $Rank(k)$ is the ranking of the ground truth images by the retrieval algorithm and $K(q)$ specifies the "relevance rank" for each query. As the size of the ground truth set is normally unequal, a suitable K(q) is determined by :

$$K(q) = \min\left(4 * NG(q), 2 * GTM\right) \quad (4)$$

where GTM is the maximum of $NG(q)$ for the all queries. The NMRR is in the range of [0 1] and smaller values represent a better retrieval performance. ANMRR is defined as the average NMRR over a range of queries, which is given by

$$ANMRR = \frac{1}{NQ} \sum_{q=1}^{NQ} NMRR(q) \quad (5)$$

where NQ is number of query images. Table I,II listed the performance improvement of the relevance feedback using MPHSM and DCD-QHDM in terms of NMRR and ANMRR for the 20 selected images.

It can be seen from the table that the proposed MPH relevance feedback query update can improve both DCD-QHDM and MPHSM. For some of the testing images the accuracy improvement could be higher than 20% and in average the improvement is about 4.6% for QHDM, 3% for MPHSM. Since the initial retrieval result of some queries is already very good (entries grayed), it is very hard to further improve the performance. While those queries not count in ANMRR, we can see that improvements more clearly. Where the MPH improves QHDM by 5.1% and improves the MPHSM by 6.5% Fig. 3 demonstrates the visual differences for the retrieval results of the images and #270 using the MPH-RF. The query image is shown in fig.3(a), in the retrieval results, the top left corner is the highest ranked image and sorted row by row.

TABLE I
Result of MPH relevance feedback with QHDM

QHDM	0RF	1RF	2RF	3RF
#102	0.4559	0.4368	0.4377	0.4503
#113	0.6754	0.5586	0.6225	0.5974
#204	0.4181	0.4526	0.4548	0.4366
#270	0.5683	0.4472	0.6028	0.4683
#309	0.4826	0.5486	0.5417	0.5382
#326	0.6778	0.6711	0.6778	0.6778
#327	0.4798	0.5588	0.4225	0.4094
#400	0.0011	0.0011	0.0011	0.0011
#486	0.0289	0.0289	0.0289	0.0289
#522	0.5343	0.4438	0.4826	0.5005
#582	0.6839	0.6211	0.5572	0.5356
#600	0.6059	0.5959	0.6036	0.5995
#604	0.6777	0.6336	0.6042	0.6042
#614	0.6575	0.6722	0.7089	0.7089
#616	0.7251	0.6861	0.7105	0.7061
#640	0.7429	0.7594	0.7544	0.7544
#703	0.7791	0.5459	0.3556	0.3532
#725	0.2927	0.2297	0.2395	0.2395
#804	0.5885	0.6480	0.6286	0.6286
#826	0.7897	0.7132	0.6933	0.7004
ANMRR	0.5433	0.5126	0.5064	0.4969
ANMRR'	0.6020	0.5679	0.5610	0.5505

TABLE II
Result of MPH Relevance Feedback with MPHSM

MPHSM	0RF	1RF	2RF	3RF
#102	0.4071	0.3162	0.2574	0.3405
#113	0.5116	0.4962	0.3546	0.4122
#204	0.5248	0.4487	0.4188	0.4014
#270	0.3439	0.2778	0.2322	0.2483
#309	0.6458	0.5451	0.5313	0.5035
#326	0.5978	0.6156	0.6178	0.6933
#327	0.3804	0.4332	0.3548	0.4231
#400	0.0169	0.0169	0.0169	0.0169
#486	0.0391	0.0391	0.0391	0.0391
#522	0.5485	0.5614	0.5479	0.5651
#582	0.5217	0.5728	0.4706	0.4272
#600	0.1799	0.2530	0.2576	0.2453
#604	0.2039	0.3398	0.3085	0.3232
#614	0.2847	0.4059	0.3967	0.3967
#616	0.4368	0.2692	0.3318	0.3055
#640	0.5605	0.5618	0.5490	0.3300
#703	0.3003	0.1264	0.1264	0.1264
#725	0.0727	0.0727	0.0727	0.0727
#804	0.3543	0.3551	0.4401	0.4510
#826	0.7431	0.7292	0.7147	0.7410
ANMRR	**0.3837**	**0.3718**	**0.3519**	**0.3531**
ANMRR'	**0.4912**	**0.4506**	**0.4248**	**0.4263**

5. Conclusion

In this paper, we proposed a new merged palette histogram for relevance feedback on dominant color descriptor of MPEG-7. The query update is based on a common palette generated by merging the two dominant color histograms from selected relevance images. This merged palette formed a common color space and used to redefine the new query histograms for DCD similarity measure. Experimental results show that the proposed MPH-RF in average improve the dominant color descriptor's QHDM by 4.6% and MPHSM by 3% based on the ANMRR and also provide better perceptually relevant image retrieval.

fig. 3a The relevance image #270

fig. 3b Initial retrieval– 15 ground truth images found, ANMRR = 0.3439

fig. 3c First relevance feedback - 17 ground truth images found, ANMRR = 0.2778

fig. 3d Second relevance feedback - 18 ground truth images found, ANMRR = 0.2322

Acknowledgments

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)]

References

[1] M. Stricker and M. Orengo, "Similarity of color images," SPIE Storage and Retrieval for Image and Video Databases, 1995.

[2] M. Swain and D. Ballard, "Color indexing," International Journal of Computer Vision, vol., 7 no. 1, 1991.

[3] Y. Rui, T.S. Huang & S. Mehrotra, "Content-Based Image Retrieval with Relevance Feedback in MARS." Proceedings of IEEE International Conference on Image Processing, Santa Barbara, California, Oct, 2003, pp. 815-818.

[4] B. S. Manjunath, J.-R. Ohm, V.V. Vasudevan, A. Yamada, "Color and texture descriptors," IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 6, pp. 703 -715, June 2001.

[5] K. M. Wong, C. H. Cheung, T.S. Liu, L.M. Po, "Dominant color image retrieval using merged histogram", Proceedings of the International Symposium on Circuits and Systems, vol. 2 , May 2003 pp. 908 -911

[6] T. Ojala, M. Aittola and E. Matinmikko. "Empirical evaluation of MPEG-7 XM color descriptors in content-based retrieval of semantic image categories." Proc. 16th International Conference on Pattern Recognition, Quebec, Canada, vol.2, pp. 1021-1024. 2002