A NEW PALETTE HISTOGRAM SIMILARITY MEASURE FOR MPEG-7 DOMINANT COLOR DESCRIPTOR

Lai-Man Po and Ka-Man Wong

Department of Electronic Engineering, City University of Hong Kong,
Tat Chee Avenue, Kowloon, Hong Kong SAR, China

ABSTRACT

The MPEG-7 dominant color descriptor (DCD) is an effective, compact, and intuitive representation of salient colors in an image region and is very useful for content-based image retrieval and indexing applications. However, the original quadratic histogram distance measure (QHDM) for MPEG-7 DCD similarity matching does not fit the human perception very well. It may cause incorrect ranks between regions with similar salient color distributions. In this paper, a new merged palette histogram similarity measure is proposed to provide an alternative for the QHDM. The new similarity measure is based on a merged color palette as a common color space to define histogram intersection. Experimental results show that this new similarity measure provide better accurate and perceptually relevant image retrieval for MPEG-7 DCD.

Keywords—*MPEG-7, Dominant Color Descriptor, Image Retrieval, Similarity Measure.*

1. INTRODUCTION

In response to the need for visual information retrieval due to a rapid increase in the volume of image and video collections, MPEG-7 specifies content-based descriptors that allow efficient browsing, searching, and retrieval of digital images and video based on their own visual content. Color is perhaps the most expressive of all the visual features and considerable work has been done in designing efficient descriptors for content-based image retrieval applications. In the current version of the MPEG-7 Final Committee Draft, several color descriptors have been approved. They consist of number of histogram descriptors and a dominant color descriptor (DCD) [1,3].

The MPEG-7 DCD provides the distribution of the salient colors in a region of interest or in an image. Its purpose is to provide an effective, compact, and intuitive representation of colors present in a region of interest. This feature descriptor consists of the representative colors and their percentages in the region. In [1], a quadratic histogram distance measure (QHDM) is defined for dissimilarity measure of this descriptor, however, some experimental results show that the QHDM does not match the human perception very well. It may result in incorrect ranks between regions with similar salient color distributions. In this paper, we propose a new similarity measure based on a merged palette generated from the two query dominant color histograms for similarity retrieval in large image database. With use of the merged palette, a common color space is created for the two query histograms, thus conventional histogram intersection measure can be applied to determine the similarity between these two histograms. Our experiments demonstrate that the MPEG-7 DCD with use of this new merged palette histogram similarity measure (MPHSM) could provide accurate and perceptually relevant image retrieval in an image database with 1000 images.

This paper is organized as follows. The second section introduces the original QHDM of MPEG-7 DCD. The third section presents the details of the proposed MPHSM. The fourth section describes the experimental result and some performance evaluations. Some concluding remarks are given in the last section.

2. QUADRATIC HISTOGRAM DISTANCE MEASURE (QHDM)

To generate the dominant color descriptor as defined in MPEG-7, the colors present in a given image or region are first clustered by a modified generalized Lloyd algorithm (GLA) or fast color quantization algorithm with further clusters merging. This could generate a small number of colors and the percentages of these colors in the region of interest. The descriptor could be defined as

$$F = \{\{c_i, p_i\}, i = 1,..., N\} \tag{1}$$

where N is the total number of dominant colors, c_i is a 3-D color vector, p_i is its percentage, and sum of p_i is equal to 1. The QHDM [1] for two dominant color descriptors $F_1 = \{\{c_{1i}, p_{1i}\}, i = 1,..., N_1\}$ and $F_2 = \{\{c_{2i}, p_{2i}\}, i = 1,..., N_2\}$ is defined as

$$D^2(F_1, F_2) = \sum_{i=1}^{N_1} p_{1i}^2 + \sum_{j=1}^{N_2} p_{2j}^2 - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} 2a_{i,j} p_{1i} p_{2j} \tag{2}$$

where $a_{i,j}$ is the similarity coefficient between colors c_{1i} and c_{2j}. The similarity coefficient $a_{i,j}$ is defined as

$$a_{i,j} = \begin{cases} 1 - \|c_{1i} - c_{2j}\| / \alpha T_d, & \|c_{1i} - c_{2j}\| \le T_d \\ 0, & \|c_{1i} - c_{2j}\| > T_d \end{cases} \quad (3)$$

where $\|\cdot\|$ is the Euclidean distance and T_d is the maximum distance for two colors to be considered similar. T_d is between 10 to 20 in the CIE-LUV color space and α is between 1.0 to 1.5. This dissimilarity measure can be shown to be equivalent to the well-known quadratic histogram distance measure [4] that is commonly used in comparing two color histogram descriptors. That is the reason we called this dissimilarity measure as QHDM.

2.1 Shortcoming of QHDM

In our study, we found that QHDM is not very good measure for two histograms with similar but not exact dominant colors, especially using α=1. For example, we have two perceptually very similar domain color histograms with same number of dominant colors, same color distribution, and the Euclidean distance between two dominant colors are equal to T_d. These conditions can be represented by

$$F_1 = \{\{c_{1i}, p_{1i}\}, i = 1,..., N\} \quad (4)$$

$$F_2 = \{\{c_{2i}, p_{2i}\}, i = 1,..., N\} \quad (5)$$

$$p_{1i} = p_{2i} \quad for \quad i = 1,.., N \quad (6)$$

$$\|c_{1i} - c_{2i}\| = T_d \quad for \quad i = 1,.., N \quad (7)$$

$$\|c_{1i} - c_{1j}\| > T_d, \quad \|c_{2i} - c_{2j}\| > T_d, \quad \|c_{1i} - c_{2j}\| > T_d, \quad for \quad i \ne j \quad (8)$$

The last condition is due to dominant color descriptor generation process, in which the minimum distance between two dominant colors is always greater T_d. With these conditions, the similarity coefficient $a_{i,j}$ and QHDM are given by

$$a_{i,j} = \begin{cases} 1 - 1/\alpha, & i = j \\ 0, & i \ne j \end{cases} \quad (9)$$

$$D^2(F_1, F_2) = \sum_{i=1}^{N} p_{1i}^2 + \sum_{j=1}^{N} p_{2j}^2 - \sum_{i=1}^{N} 2(1 - 1/\alpha) p_{1i} p_{2j} = \frac{2}{\alpha} \sum_{i=1}^{N} p_{1i}^2 \quad (10)$$

The maximum value of this measure is equal to 2 for N=1, p_{11}=p_{21}=1 and α=1, this is not a very small distance. That means for two regions with only one dominant color in each region and the Euclidean distance between these two colors is T_d, their distance is equal 2 in term of QHDM. This is the worth case on using QHDM for dominant color descriptor. In practice, these two regions should be perceptual similar. This example may not be very convening to demonstrate shortcoming of QHDM for images with similar dominant colors. In our second example, we have three images and each image with two dominant colors and their color distributions are equal (p_{1i}=p_{2i}=p_{3i}=0.5 for i=1,2). Their dominant color descriptors can be represented by $F_1 = \{\{c_{11}, 0.5\}, \{c_{12}, 0.5\}\}$, $F_2 = \{\{c_{21}, 0.5\}, \{c_{22}, 0.5\}\}$ and $F_3 = \{\{c_{31}, 0.5\}, \{c_{32}, 0.5\}\}$. In this example, F_1 and F_2 are satisfy the conditions of (4) to (8), such that their QHDM distance is equal to

$$D^2(F_1, F_2) = \frac{2}{\alpha} \sum_{i=1}^{N} 0.5^2 = \frac{1}{\alpha} \quad (11)$$

However, the image 3 with one dominant color c_{31} exactly equal to c_{11} in region 1 ($\|c_{11} - c_{31}\| = 0$) while the second color c_{32} is very different from c_{11} and c_{12} ($\|c_{11} - c_{32}\| >> T_d$ and $\|c_{12} - c_{32}\| >> T_d$). With these conditions, the distance of F_1 and F_3 in term of QHSM is given by

$$D^2(F_1, F_3) = \sum_{i=1}^{2} 0.5^2 + \sum_{j=1}^{2} 0.5^2 - \sum_{i=1}^{2} \sum_{j=1}^{2} 2 a_{i,j} 0.5^2 = 0.5 \quad (12)$$

where only $a_{1,1}$=1 and other similarity coefficients are equal to 0. Unfortunately, based on (11) and (12), we can find that $D^2(F_1, F_2)$ is always greater than $D^2(F_1, F_3)$ as

$$D^2(F_1, F_2) = \frac{1}{\alpha} > 0.5 = D^2(F_1, F_3) \quad (13)$$

where 1 ≤ α ≤ 1.5. The minimum value of $D^2(F_1, F_2)$ is 2/3 for α=1.5. That means image 1 is more similar to image 3 than image 2 based on the QHDM of their dominant color descriptors. Visually, image 1 is more similar to image 2 as both two colors are very alike. This example also demonstrates that QHSM could not effectively measure the dissimilarity of images with perceptually similar dominant colors. In this paper, we propose an alternative of this dissimilarity measure called merged palette histogram similar measure, which could improve the accuracy in this situation.

3. MERGED PALETTE HISTOGRAM SIMILARITY MEASURE (MPHSM)

The major difference between the dominant color descriptor and the conventional color histogram descriptor is that the representative colors are computed from each image instead of being fixed in the color space, thus conventional histogram similarity measurement such as histogram intersection [2] cannot be directly applied on dominant color histograms matching. However, the format of the dominant color descriptor defined in (1) is very similar to a color palette of a color quantized image stored in GIF format. In order to use the conventional histogram intersection similarity measure for dominant color descriptor, we can generate a common palette for two of these descriptors by merging their color histogram bins using the 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 dominant color descriptors. The common palette is generated by searching the closest two colors between the two palettes and if this minimum distance is less than or equal to the threshold T_d then the two colors will be merged as

$$c_{m(i,j)} = \frac{p_{1i}c_{1i} + p_{2j}c_{2j}}{p_{1i} + p_{2j}} \quad (14)$$

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 use of the merged colors and unmerged colors from the two palettes. This merged palette forms a common color space for the two histograms, thus we can redefine the histograms of F_1 and F_2 with use of this space and each histogram bins is equal to N_m. For example, the redefined F_1 histograms is given by

$$F_{1m} = \{\{c_{mi}, p_{1mi}\}, i = 1,..., N_m\} \quad (15)$$

where $p_{1mi} = p_{1i}$ if the distance between c_{mi} and c_{1i} is less than or equal to T_d; otherwise $p_{1mi} = 0$. Similarly, the redefined F_2 histogram is given by

$$F_{2m} = \{\{c_{mi}, p_{2mi}\}, i = 1,..., N_m\} \quad (16)$$

An example of this merged color palette histogram for two images is shown in Figure 1 with the common palette and the regenerated histograms. As these two histograms F_{1m} and F_{2m} are based on a common color palette, we can directly apply the histogram intersection method to calculate their similarity. The MPHSM is defined as the intersection area of these two histograms, which is given by

$$I(F_{1m}, F_{2m}) = \sum_{i=1}^{N_m} \min(p_{1mi}, p_{2mi}) \quad (17)$$

The larger the value $I(F_{1m}, F_{2m})$ the more similar the two images and the maximum value is 1. To demonstrate the performance of the MPHSM, we could use this new measure to compare the similarity of the images F_2 and F_3 to the image F_1 in second example of section 2. The results are I(F_1, F_2) = 1 and I(F_1, F_3) = 0.5, which mean that all colors of images F_1 and F_2 are similar while only 50% of the colors in images F_1 and F_3 are similar. This result matches our perceptual interpretation. The problem with QHDM does not exist in this new measure. It is because the threshold T_d is defined as the threshold of human perception on similar colors, the colors with distance smaller than T_d are considered as the same color. Experimental results based on MPEG-7 retrieval metric with use of an image database with 1000 images for further evaluation of these two similarity measures are shown in the next section.

4. EXPERIMENTAL RESULTS

In this section, we present the performance of the proposed MPHSM and compare it with the QHDM for image retrieval using dominant color descriptor. 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 images from different categories are also defined for our retrieval experiment. 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) [5] by modifying the dominant color descriptor matching program. The threshold for similar colors is set to T_d = 15 and α = 1.2 for DCD generation from the image database and the similarity measures of QHDM and MPHSM.

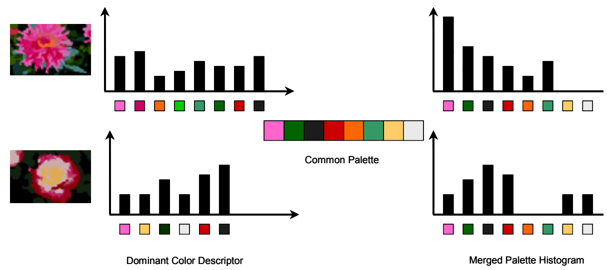

Figure 1: Merged Palette Histogram uses histogram intersection directly.

To measure the effectiveness of these two similarity measures for image retrieval using MPEG-7 DCD, we use the MPEG-7 retrieval metric the Normalized Modified Retrieval Rank (NMRR) [3]. 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

$$NMRR(q) = \frac{\left(\sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)}\right) - 0.5 - \frac{NG(q)}{2}}{K(q) + 0.5 - 0.5 * NG(q)} \quad (18)$$

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 :

$$K(q) = \min(4 * NG(q), 2 * GTM) \quad (19)$$

where GTM is the maximum of $NG(q)$ for 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, and is given by

$$ANMRR = \frac{1}{NQ} \sum_{q=1}^{NQ} NMRR(q) \quad (20)$$

where NQ is number of query images. Table I listed the performance comparison of the dominant color descriptor using MPHSM and QHDM in terms of NMRR and ANMRR for the 20 selected images. It can be seen from the table that the proposed MPHSM achieves better results in terms of retrieval accuracy compared to the QHDM. For some of the testing images the accuracy improvement could be higher than 30% and in average the improvement is about 14%. Figures 1(a) and 1(b) demonstrate the visual differences for the retrieval results of the image #616 using the two similarity measures. The query image is shown in the upper left corner. The upper and the left most images are higher ranked images. These figures also show that the proposed method could achieve better perceptually relevant image retrieval.

5. CONCLUSIONS

In this paper, we proposed a new merged palette histogram similarity measure for dominant color descriptor of MPEG-7. The similarity measure is based on a common palette generated by merging the two query dominant color histograms. This merged palette formed a common color space and used to redefine the two query histograms for conventional histogram intersection measure. Experimental results show that this new MPHSM in average outperforms the QHDM by up to 14% based on the ANMRR and also provide better perceptually relevant image retrieval for MPEG-7 DCD.

ACKNOWLEDGMENT

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

TABLE I

Retrieval Results of the QHDM and Proposed MPHSM.

Query	NG	K	NMRR (QHDM)	NMRR (MPHSM)
#102	17	68	0.644737	0.567337
#197	24	96	0.673256	0.558140
#204	21	84	0.573455	0.606890
#270	20	80	0.571508	0.362011
#309	8	32	0.489437	0.651408
#326	8	32	0.862676	0.742958
#327	27	108	0.481788	0.393327
#400	100	200	0.001053	0.003609
#486	100	200	0.028922	0.035188
#522	19	76	0.537461	0.473065
#582	18	72	0.775707	0.653554
#600	17	68	0.609907	0.267802
#604	11	44	0.521336	0.359926
#614	11	44	0.664193	0.300557
#616	34	136	0.675217	0.376471
#640	26	104	0.746121	0.535490
#703	82	200	0.780956	0.344856
#725	47	188	0.293436	0.095089
#804	54	200	0.585685	0.356221
#826	25	100	0.799643	0.765000
Average (ANMRR)			**0.565825**	**0.422445**

REFERENCES

[1] Y. Deng, B. S. Manjunath, C. Kenney, M. S. Moore and H. Shin, "An Efficient Color Representation for Image Retrieval," *IEEE Transactions on Image Processing*, Vol. 10, No. 1, pp. 140-147, Jan 2001.

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

[3] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan and A. Yamada, "Color and Texture Descriptors," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 6, pp. 703-715, June 2001.

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

[5] *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.\

Figure 1 (a): Retrieval results for the query image #616 using dominant color descriptor with QHDM (NMRR=0.675217).

Figure 1 (b): Retrieval results for the query image #616 using dominant color descriptor with MPHSM (NMRR=0.376471)