A new adaptive distance computation technique for query-by-multiple-example system

Jianqiao Feng, Haifeng Zhao, and Wenhua Jia
Institute of Computer Science & Technology
School of Electronics Engineering and Computer Science, Peking University
Beijing, P.R.CHINA 100871
E-mail: {fengjianqiao, zhaohaifeng, jiawenhua} @ icst.pku.edu.cn

Abstract

Query-By-One-Example (QBOE) is the traditional way of querying in content-based image retrieval (CBIR) system. However, as some recent research points out, QBOE method cannot get accurate result because only one image is not sufficient to express its semantics of the intended query. Therefore, Query-By-multiple-Example (QBME) method is proposed and adopted, in which query images are divided into groups according to relevance to target image class. In order to maximize major features and minimize minor ones, previous researches have introduced adaptive distance computation in QBME. These methods optimize query result compared to QBOE, but still have some defects.

This paper proposes a new adaptive distance computation technique for QBME, which achieves higher performance than previous methods.

1. Introduction

Image retrieval can be divided into two types [1]. The first one is Keyword-based image retrieval, which allows user to search images with high-level semantics and is high-speed. However, it requires all images to be labeled manually. This is an extremely boring task and contains human subjectivity. The other retrieval method is Content-based image retrieval (CBIR) [2], in which low-level features such as color, texture, and structure features are extracted automatically to calculate the distance between query images and images in database.

Generally, the interaction of the CBIR system is Query-By-Example. In most CBIR system user query by one image example (QBOE). However, recent researches [3] show that only one example image cannot form its semantics or concepts of the query. As a result, some systems implement querying by multiple examples (QBME) [1] [4] [5] [6]. Input examples are divided into three groups: positive, negative and neutral. The positive group includes images relevant to query. Negative images are irrelevant to query and neutral images are those between these two groups. In a QBME system named ImageGrouper [1] [4], the covariant between features contained within each group and between groups is used to adjust weight of each feature in distance computation formula. In another QBME system, Data System Group (DSG) [5], an adaptive distance computation technique (ADCT) is devised and shows higher performance than ImageGrouper. DSG uses the range distance [5] of each feature in each group to adjust weight of the feature. The adjustment is based on comparison of range distances between different groups and some minor features are discarded if range distance of negative group is smaller than that of positive group. Zhao et al [6] introduce a modified ADCT (MADCT), which implements a new adaptive fuzzy clustering method (AFC) to divide negative group into small groups and neutral group is removed to reduce input complexity of users. Meanwhile, MADCT reserves a minor feature if its range distance on positive group is relatively small.

MADCT is more effective than ADCT according to experimental results [6]. However, it still has some defects. Firstly, calculation of range distance does not take the number of images in the group into consideration. In most cases, positive group contains more images than negative group, especially when negative group is divided into smaller sub-groups by clustering. As a result, it is possible that one feature is more convergent in positive group although range distance of negative group is smaller, since much less negative images than positive ones. Secondly, distance between positive and negative groups is not taken into account. This distance represents the difference between positive examples and negative ones. While

978-0-7695-3493-0/08 $25.00 © 2008 IEEE
DOI 10.1109/SITIS.2008.20

range distance only shows the convergence of features in different groups. Finally, the range distances are just compared between different groups but not compared among different features in the same group. The range distance of positive group indicates aggregation of query target on each feature, so that should make more contribution to weights adjustment in distance computation metric.

This paper proposes a new adaptive distance computation method for QBME, which called new ADCT (NADCT). In our approach, density of positive group and distance between the centers of positive group and negative groups are calculated to adjust weight of each feature in distance computation. We also adopt AFC method mentioned in [6] to divide negative images into small groups. Experimental results show that the NADCT can optimize result compared to QBOE and achieve higher performance than MADCT.

The article is outlined as follows: Section 2 discusses the feature extraction part of our QBME system. Section 3 introduce NADCT in detail and compares it with MDCT. Section 4 describes experimental results and Section 5 finishes with a conclusion.

2. Feature extraction

In our QBME system, color, texture and structure features are extracted for further searching.

For color features, a common method is the color histogram technique [7] [8]. In our system, the HSV color space is used. In order to preserve some spatial information and get better accuracy, we adopt the method mentioned in [6]. An image is divided into 9 sub-images and for each block the average, variance and skewness of H, S and V are calculated. In total, eighty-one color features are computed for every image.

For texture features, Wavelet-based Texture features [9] [10] are used. We adopt Daubechies-4 wavelet to decompose image and analyzed by the method of multi-resolution analysis (MRA) [11]. This process should repeat twice to generate three-level decomposition with ten sub-bands. Average of the wavelet coefficients is calculated for each sub-band as a texture feature. Thus, there are a total of ten texture features for each image.

For structure feature, we adopt the Water-Filling Algorithm [12] [13] to the edge maps of the images. First step is edge detection with Sobel filter [14], which is followed by the thinning step [14] to generate its corresponding edge maps. From the edge map, eighteen features are extracted including the longest edge length, the average and max filling speed of paths, fork number, horizontal cover, vertical cover, path number, loop number and water amount.

Finally, a 105-dimension vector is formed for each image. And all of these features are normalized for further computation.

3. Distance computation

In our QBME system, we use the range distance to evaluate the similarity between query image and images in database. The center of feature vectors in positive group is considered as query target: for feature i, $\overline{Q_i} = average(Q_i)$, where Q is an image vector in positive group.

The distance between Q and image feature vector I is calculated by the formula below:

$$\mathrm{D(Q,I)} = \sum_{i=1}^{N} | Q_i - I_i | \times W_i \tag{1}$$

Where i is the number of features, indexing from 1 to N. And W_i is the weight of feature in distance calculation.

As in [6], we remove the neutral group for reducing complexity of user input. The range distance of image group $d_i = u_i - l_i$, where u_i and l_i are upper and lower bound of the feature. Range distance is the factor used in ADCT and MADCT to determining whether a feature is a major or minor factor in distance calculation. However, we think this is neither sufficient nor accurate enough. One feature should be assigned to a higher weight only if it is more meaningful for distinguish irrelevant images from relevant ones. As for multiple input images, more convergent in positive group and larger distance between positive and negative groups mean the feature is more important for classifying relevant and irrelevant images in database. Thus we introduce two new concepts for weight adjustment and the algorithm based on these two concepts is described in detail as follows.

3.1 Center distance

For feature f_i of all images, center of positive group $cen_{ipos} = average(f_i)$, where f is a positive image vector and similarly, center of negative group $cen_{ineg} = average(f_i)$. Calculate the center distance

$$cd_i = | cen_{ipos} - cen_{ineg} | \tag{2}$$

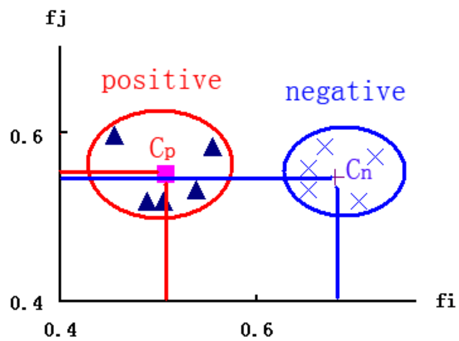

Figure 1.

Larger center distance indicates that it is more easily to differentiate positive images from negative ones on this feature. As a result it is more important for distance computation. An example is shown in Figure 1, in which comparison of range distances between positive and negative groups is similar on f_i and f_j. However, f_i is more important since positive group is far away from negative group on f_i. In Figure 1, C_p and C_n represent the center of positive and negative group respectively.

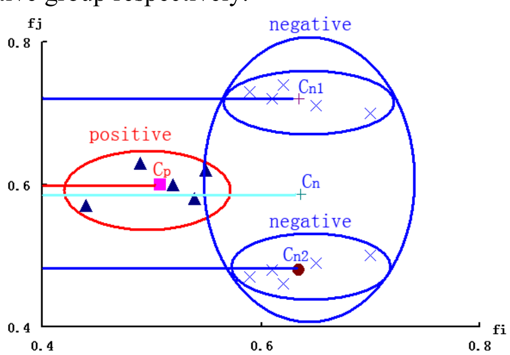

Figure 2.

Clustering is necessary for center distance calculation because center distance might be ignored when all negative images are considered as a whole group. A typical example of this case is shown in Figure 2. Only two features f_i and f_j are shown. And the center distance on f_j is small because two sub-groups counteract each other. In Figure 2, C_p is the center of positive group, C_n is the center of negative group, C_{n1} and C_{n2} are centers of sub-groups.

For each negative sub-group after clustering, a center distance is calculated. And for each feature, the average, max and min value of these center distances can be used. We choose the maximum center distance for further computation.

3.2 Group density

For feature f_i, the range distance is defined as $d_i = u_i - l_i$, where u_i and l_i are upper and lower bound of the feature within the group. The group density of f_i in a group is defined as (3)

$$den_i = \mathrm{num} / d_i \tag{3}$$

where num is the number of images in the group.

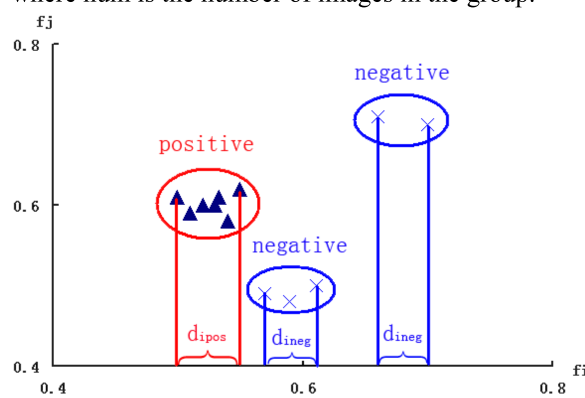

Figure 3.

The range distance is used in [3] and [4] to evaluate degree of convergence in a group. However, range distance is not accurate because it tends to be larger when the group contains more images. Figure 3 shows an example of this case. For feature f_i, the positive group is more convergent than negative group but their range distances are equivalent approximately. But group density is much larger in positive group than in negative one. It is more accurate than range distance for adjusting weights in distance computation.

3.3 Distance computation and matching

The distance metric $\mathrm{D(Q,I)}$ between image (I) and query image (Q) is defined as:

$$\mathrm{D(Q,I)} = \sum_{i=1}^{N} | Q_i - I_i | \times W_i \tag{4}$$

where Q_i is the feature i of query image Q or the center of positive image vectors. I_i is the feature i of image I in database. And W_i, in range [0, 1], is the weight of feature i.

As we have discussed, the maximum center distance and density of positive and negative groups are used to adjust weight for each feature in distance metric.

1. For positive group and each negative sub-group generated by clustering, for each feature (f_i), compute its lower bound l_i and upper bound u_i , and then compute the group density:

$$d_i = num/(u_i - l_i) \quad (5)$$

where num is the number of images in the group.

Thus, for each feature, there are two group densities: d_{ipos} , d_{ineg} for positive and negative groups respectively.

2. For each feature (f_i), for each negative sub-group, the center distance is calculated. Choose the largest center distance as cd_{i_max}.

3. For each feature (f_i), compute the weight:

$$W_i = (d_{ipos})^m \times cd_{i\_max} \quad (6)$$

where m is a factor for adjustment, usually set as 1 or 2.

If for all negative sub-groups, d_{ineg} is larger than d_{ipos} , then $W_i = W_i$ /num, where num is the number of negative sub-groups.

4. Normalize the weight.

$$W_i = 1 - W_i / \sum_{k=1}^{N} W_k \quad (7)$$

Finally, the results are sorted in ascending order, and the K top-rank images are return to user.

3.4. Comparison to MADCT

MADCT is a modified version of ADCT mentioned in [3]. Like ADCT, MADCT is also based on the conception of range distance. For each feature, the smallest range distance d_i is used for weight computation. If d_i is from positive group or the feature is an important feature, that is, compared with other features, range distance of positive group of this feature is relatively small(for example, top 20%), then the weight is calculated as:

$$W_i = 1 - d_i / \sum_{k=1}^{N} d_k \quad (8)$$

where N is the number of features. Else, the W_i will be divided by the number of negative images.

In our algorithm, we use group density instead of range distance to determine whether a feature is important. Only if the density of positive group is smaller than all negative sub-groups, the weight will be divided by the number of negative groups. The density of negative group is just used for judging but not for computation and the center distance between positive and negative groups is introduced. The group density represents the degree of convergence in a group more accurately than range distance. And center distance can be used to represent the degree of separation between negative and positive groups on a certain feature.

4. Experiments and analysis

4.1 Introduction

To compare the performance of our algorithm, NADCT, with previous method, MADCT, we employ the Corel image database [14] with size of 10,000 images, and evaluate the performance of algorithms by NMRR (Normalized Modified Retrieval Rank) and recall rate [15]. For purpose of evaluation, we labeled nine classes of images previously: “landscape”, “butterfly”, “boat”, “map”, “flower”, “car”, “sky”, “sunset”, and “earth”. And we use these labeled images as input. So only images with same label are considered as similar.

4.2 Experiment program

We tested and compared three methods in our CBIR system, QBOE, QBME-MADCT, and QBME-NADCT. To imitate user input, we construct twenty-seven query groups (QG) as positive groups, which contain images with same label. For QBOE, the center of QG is set as the query and query result is used to construct negative group. To imitate user feedback, from first retrieval result set we choose the first fifteen irrelevant images, with different label to query images, as negative group. Then these groups are used as input to retrieval for a second time from database by QBME-MADCT and QBME-NADCT respectively. For each query group, NMRR and recall rate are calculated for all three methods.

Table 1 shows the performance of QBOE, QBME-MADCT and QBME-NADCT methods. Figure 4 and Figure 5 show NMRR and recall rate. These statistic data indicate the performance of three algorithms on all query groups. Figure 6 and Figure 7 exhibit four query examples using QBME-MADCT and QBME-NADCT respectively.

4.3 Result analysis

From Table 1, it proves that QBME is more effective than QBOE through statistic data. Figure 4 and Figure 5 show that, in general, ABME-NADCT achieves lower NMRR and higher recall rate than QBME-MADCT. It proves our method of distance computation presents better performance than previous technique in QBME system.

5. References

[1] Munehiro Nakazato, Ljubomir Manola and Thomas S. Huang, "Group-based User Interface for Content-based Image Retrieval," In proceedings of *Advanced Visual Interfaces 2002 (AVI 2002)*, 2002.

[2] Flickner M, et al. "Query by image and video content: the QBIC system," *IEEE Computer Magazine*, 1995, 28(9):23 ~32

[3] Thomas E. Bjoerge and Edward Y. Chang, "Why one example is not enough for an image query," In Proceedings of *IEEE ICME*, 2004.

[4] Munehiro Nakazato and Thomas S. Huang, "Extending Image Retrieval with Group-Oriented Interface," In Proceedings of *IEEE ICME*, 2002.

[5] Nualsawat Hiransakolwong, Kien A. Hua, Soontharee Koompairojn, Khanh Vu, and Sheau-Dong Lang, "An Adaptive Distance Computation Technique for Image Retrieval Systems," *ACM Symposium on Applied Computing*, 2005.

[6] Haifeng Zhao and Wenhua Jia, "An Adaptive Fuzzy Clustering Method for Query-by-Multiple-Example Image Retrieval", *IEEE SITIS* 2007, 2007.

[7] Ka-Man Wong, Chun-Ho Cheung, Lai-Man Po, "merged-color histogram for color image retrieval", in Proceedings of *International Conference on Image Processing*, 2002.

[8] AK Jain, A Vailaya, "Image Retrieval Using Color And Shape", *Pattern Recognition*, 1996.

[9] Smith, J. R. and Chang S-F. "Transform features for texture classification and discrimination in large image databases," In Proceedings of *IEEE Intl. Conf. on Image Processing*, 1994.

[10] Smith J. R. and Chang S-F. "Quad-Tree Segmentation for Texture-based Image Query." In Proceedings of *ACM 2nd International Conference on Multimedia*, 1994.

[11] S. Mallat, "A theory for multi-resolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989.

[12] Zhou, X. S., Y. Rui, and T. S. Huang, "Water-filling: a novel way for image structural feature extraction", Proc. *IEEE Inter. Conf. on Image Processing*, Kobe, Japan. Oct. 25-29, 1999.

[13] Zhou, X. S. and T. S. Huang, "Edge-based structural features for content-based image retrieval," *Pattern Recognition Letters*, Vol 22/5, Apr. 2001.

[14] Gonzales, R. C. and Woods, R. E., *Digital Image Processing*, Addison-Wesley, 1992.

[15] http://wang.ist.psu.edu/docs/related/

[16] MPEG video group. "Description of Core Experiments for MPEG-7 Color/Texture Descriptors," ISO/MPEGJTC1/SC29/WG11 MPEG98/M2819, July, 1999

Table 1. Performance of QBOE, QBME-MADCT, and QBME-NADCT

Query Group	QBOE		QBME-MADCT		QBME-NADCT	
	NMRR	Recall	NMRR	Recall	NMRR	Recall
Group1(Landscape)	0.194086	0.285714	0.17792	0.326531	0.167112	0.329429
Group2(Landscape)	0.179267	0.193878	0.166912	0.204082	0.178932	0.22449
Group3(Landscape)	0.194503	0.13898	0.18147	0.153061	0.173633	0.173469
Group4(Butterfly)	0.179886	0.41	0.181974	0.47	0.17691	0.5
Group5(Butterfly)	0.146552	0.32	0.200601	0.38	0.194526	0.46
Group6(Butterfly)	0.168269	0.37	0.196137	0.49	0.204549	0.46
Group7(Boat)	0.20003	0.540816	0.192013	0.663265	0.171041	0.704082
Group8(Boat)	0.212215	0.510204	0.203745	0.683673	0.181618	0.734694
Group9(Boat)	0.171667	0.591837	0.167416	0.632653	0.145509	0.734694
Group10(Map)	0.139514	0.642276	0.117324	0.634146	0.166256	0.739837
Group11(Map)	0.13806	0.682927	0.124633	0.666667	0.161648	0.829268
Group12(Map)	0.107126	0.552846	0.160099	0.650407	0.0739639	0.504065
Group13(Flower)	0.185696	0.2	0.185902	0.194737	0.176378	0.231579
Group14(Flower)	0.134022	0.436842	0.121793	0.489474	0.0789355	0.515789
Group15(Flower)	0.200863	0.215789	0.181149	0.268421	0.184159	0.305263
Group16(Car)	0.19285	0.377551	0.217042	0.438776	0.196395	0.520408
Group17(Car)	0.139818	0.428571	0.13815	0.357143	0.113541	0.418367
Group18(Car)	0.1669	0.530612	0.155251	0.581633	0.122747	0.653061
Group19(Sky)	0.204233	0.5	0.172	0.418605	0.154781	0.44186
Group20(Sky)	0.15598	0.476744	0.125798	0.453488	0.148938	0.5
Group21(Sky)	0.215571	0.546512	0.197307	0.488372	0.204543	0.534884
Group22(Sunset)	0.19947	0.309904	0.175404	0.341853	0.168744	0.4377
Group23(Sunset)	0.193214	0.345048	0.236496	0.463259	0.193296	0.428115
Group24(Sunset)	0.215734	0.27476	0.216811	0.383387	0.19005	0.357827
Group25(Earth)	0.184138	0.52381	0.10694	0.547619	0.0807736	0.571429
Group26(Earth)	0.162197	0.571429	0.132618	0.571429	0.09085	0.666667
Group27(Earth)	0.150333	0.619048	0.129368	0.714286	0.103527	0.761905

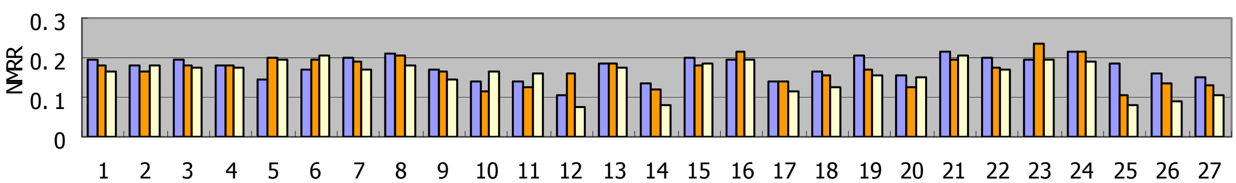

Figure 4. NMRR comparison of QBOE, QBME-MADCT, and QBME-NADCT

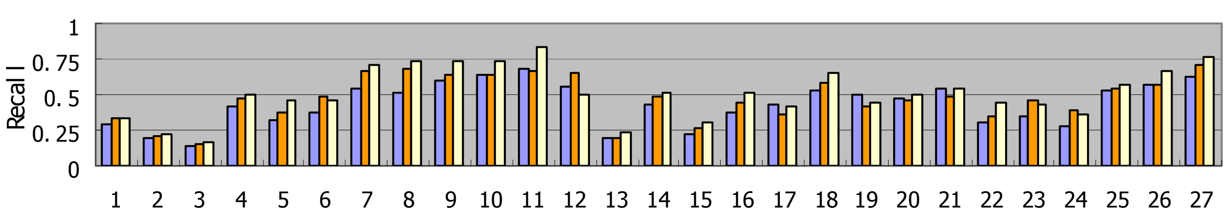

Figure 5. Recall-rate comparison of QBOE, QBME-MADCT, and QBME-NADCT

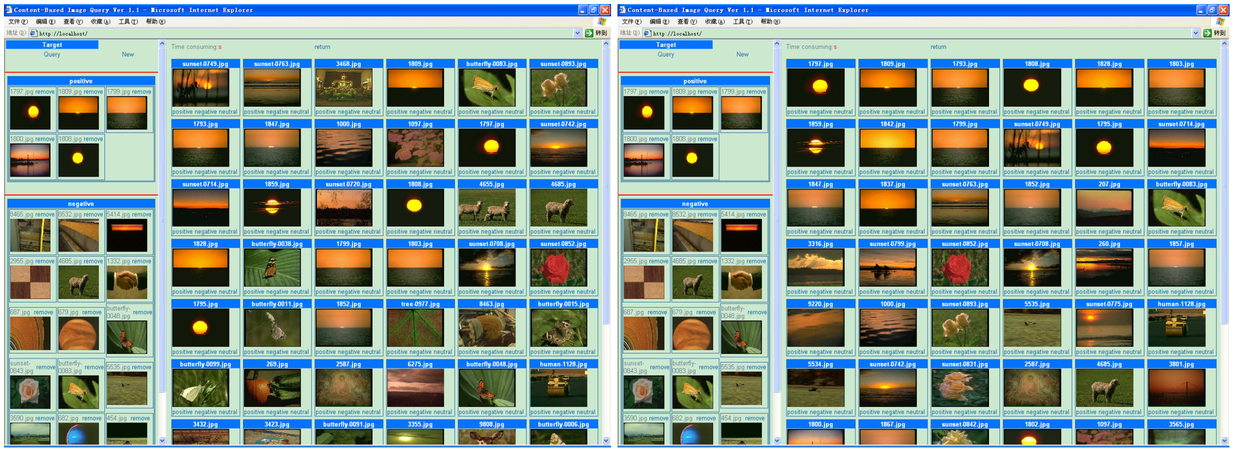

Figure 6. QBME-MADCT method and QBME-NADCT method on "Sunset"

Figure 7. QBME-MADCT method and QBME-NADCT method on "Earth"