

A New Image Classification Framework for Improving Image Retrieval Based on Relevance Feedback

نظام جديد لتصنيف الصور قائم على التغذية الرجعية بهدف تحسين استرجاعها

M. E. ElAlami

Department of Computer Science

Mansoura University

Mansoura Egypt

Moh_EIAlmi@mans.eun.eg

المخلص: تعد عملية استرجاع الصور من التقنيات الواعدة في البحث عن الصور والتي تتم عن طريق سمات الصورة، حيث تتيح التغذية الرجعية للمستخدم التفاعل مع النظام من خلال إضافة تقيمه. يعرض هذا البحث إطار عمل لتصنيف الصور من خلال تحليل سمات الصورة مثل اللون، والشكل، والملمس. يعتمد الإطار الأول على القيمة والاتجاه في الفراغ المتعدد الأبعاد لكل من الهدف والصورة المستعلم عنها، مما يمكن من تصنيف الصور طبقاً لاحتمالات تواجدها بينما يعتمد الإطار الثاني على استرجاع الصور عن طريق استخراج معرفة التصنيف من قواعد بيانات الصور الضخمة في مجال محدد. إن عملية التصنيف تستخدم كمرحلة تدريب يتم فيها تعزيز النظام بشجرة التصنيف لنطاق صور الاسترجاع. وتعد شجرة التصنيف بهذا الأسلوب من أفضل الحلول لتقليل الحجم الضخم للمجموعات التي يمكن إنتاجها من الاحتمالات المختلفة لتواجد تقسيمات قيم خصائص الصور مع بعضها البعض.

Abstract: Content-based image retrieval (CBIR) is a promising technology to assist image finding. CBIR retrieves images by visual features inherent in images. Relevance feedback allows the user to reflect his preference to the system, then the system can reformulate the query according to the positive and/or negative examples responded by the user. This paper presents two efficient frameworks for image classification through the analysis of the visual features (such as color, shape, texture) of an example image. The first framework is based on comparing the norm and the direction of vectors in a multi-dimensional space for both the target and query image vectors, which allows classification according to their probabilities of existence. The second framework depends on acquiring classification knowledge from a large empirical image database in a specific domain and utilizes that knowledge for image classification. The initial classification process is used as a training phase to feed the system with a classification tree for images in the retrieval domain. This tree is the best reduction of dimensionality that would result from all possible combinations of feature divisions.

Keywords: Image classification, Content Based Image Retrieval, Relevance Feedback, Inference Engines, Vector Analysis, Multidimensional Vector Indexing.

Accepted December 30, 2007

1. Introduction

Image capture capabilities are evolving so rapidly that extreme amount of images is produced daily. Great research efforts are needed to handle such quantities of images automatically. Most effort is directed towards content based image retrieval, with the aim to find similar images to given query image [1]. Humans perceive and characterize content using high level concepts, such as the amount of action, romance, comedy or emotional factors, which are not related in a straightforward way to visual features of pixels that compose images [2]. Image classification is the typical unsupervised learning technique for retrieval purpose. It intends to group a set of image data in a way to maximize the similarity within classes and minimize the similarity between different classes. Each resulting class is associated with a class label and images in same class are supposed to be similar to each other [3]. The traditional k-means clustering and its variations are often used for image clustering [4]. Due to the complex distribution of image data (data points are sampled from nonlinear manifold), tradition methods such as k-means clustering often cannot well separate images with different concepts [5]. Probabilistic classification based on Bays theory is among the most powerful clustering tools. The common maximum - a-posteriori (MAP) classifier and its variation maximum-likelihood (ML) classifier have shown great promise for the CBIR problem. However, traditionally it is difficult to apply the classifiers due to the complexity of the MAP similarity function [6].

This paper proposes two efficient classification frameworks that address clusters definitions based on features extracted from image database. In the first framework, the image features are considered as multi-dimensional vector, which has phase angle and scalar norm (magnitude). Comparing the angles and/or the norm of a two vectors leads to different

clustering classes. The similarity measure used is simple and better than the traditional Euclidian distance similarity measures. The early work shows that most of the schemes deal with Euclidean distance, which has number of disadvantages [7]. The second approach has the capability to improve the initial classification based on relevance feedback. Consequently, the second framework uses classification as a learning tool to improve retrieval performance (precision and recall). In addition, cluster prototypes are designed to support quick human detection of relevant images. The second framework starts with extracting n images' features and divides each feature into m intervals. The i^{th} cluster could be generated from the intersection among features' intervals. The number of generated clusters will be m^n . This yields to system overflow due to the curse of dimensionality. Initial classification algorithm is presented to overcome this problem. In this algorithm, the strength of the object oriented and search strategies of the inference engine in the artificial intelligence (AI) language such as CLIPS are utilized [8]. The utilization of the AI language concentrates on defining cluster class and inherits instances of the matched clusters only while dropping the unmatched clusters. Unfortunately, a cluster Interference problem is arisen (i.e a query image may be classified into more than one class). A pruning model is presented to overcome this problem. The proposed framework is expressed as knowledge base (IF-Then rules), which could be used to capture the semantic content and index new untagged images being added to the image database.

2. First Classification Framework

The dot product of two vectors in multi-dimensional space yields two values, the vector norm and the angle. These two values can be used together as a discriminator among the query vector and other vectors in the image database. This

simple and straightforward idea is utilized as the first clustering framework. This framework can be explained as follows:

2.1 Dot product vector for similarity indexing

1- Let an image low level feature vector is:

$$v_i = f_x \underline{i} + f_y \underline{j} + f_z \underline{k} \quad (1)$$

2- Let the query low level feature vector in 3 D is:

$$q_i = q_x \underline{i} + q_y \underline{j} + q_z \underline{k} \quad (2)$$

3- The dot product of the two vectors v_i and q_i is given by:

$$v_i \cdot q_i = \|v_i\| \|q_i\| \cos(\theta) \quad (3)$$

Where;

$$\|v_i\| = \sqrt{f_x^2 + f_y^2 + f_z^2} \quad (4)$$

$$\|q_i\| = \sqrt{q_x^2 + q_y^2 + q_z^2} \quad (5)$$

$$v_i \cdot q_i = f_x \cdot q_x + f_y \cdot q_y + f_z \cdot q_z \quad (6)$$

$$\cos(\theta_{vq}) = \frac{v_i \cdot q_i}{\|v_i\| \|q_i\|} \quad (7)$$

The difference in magnitude between the vector v_i and q_i is given by:

$$\Delta_{vq} = \|v_i\| - \|q_i\| \quad (8)$$

To compare the vector v_i and q_i in both the norm and direction, one can find there probability of the binary existence as shown in truth table 1. So we can get 9 classes. These classes can be represented by the following rules:

- If $(A \cap B \cap C \cap D) = 0$ Then class 1
 If $((A \cap B \cap C) = 0 \ \& \ D = 1)$ Then class 2
 If $((A \cap B \cap D) = 0 \ \& \ C = 1)$ Then class 3
 If $((A \cap C \cap D) = 0 \ \& \ B = 1)$ Then class 4
 If $((A \cap C) = 0 \ \& \ (B \cap D) = 1)$ Then class 5
 If $((A \cap D) = 0 \ \& \ (B \cap C) = 1)$ Then class 6

If $((B \cap C \cap D) = 0 \ \& \ A = 1)$ Then class 7

If $((A \cap D) = 1 \ \& \ (B \cap C) = 0)$ Then class 8

If $((A \cap C) = 1 \ \& \ (B \cap D) = 0)$ Then class 9

Table 1. Comparing the angles and/or the norm of a two vectors.

A	B	C	D	Class No.
Δ_{vq}^+	Δ_{vq}^-	$\delta\theta^+$	$\delta\theta^-$	
0	0	0	0	1
0	0	0	1	2
0	0	1	0	3
0	1	0	0	4
0	1	0	1	5
0	1	1	0	6
1	0	0	0	7
1	0	0	1	8
1	0	1	0	9

2.2 Group arrangement

Since there exist 9 classes and each class may have set of images then one can display the individual set images by sorting them. Finally the performance measures are calculated.

3. Second Classification Framework

A new classification approach is presented in this section. It includes features extraction, features' division, initial classification, relevance feedback and classification pruning. The overall framework is shown in figure 1.

3.1 Features extraction

Features extraction is one of the basic processes in content-based image retrieval. In a broad sense, features may include both text-based features (key words, annotations) and visual features (color, texture, shape, faces). However, since there already exists rich literature on text-based feature extraction in the database management systems (DBMS) and information retrieval research communities [9], the techniques of visual feature extraction are stressed here. The features automatically extracted using computer vision techniques, are mostly low-level features (color, texture, shape, spatial layout, etc.) [3].

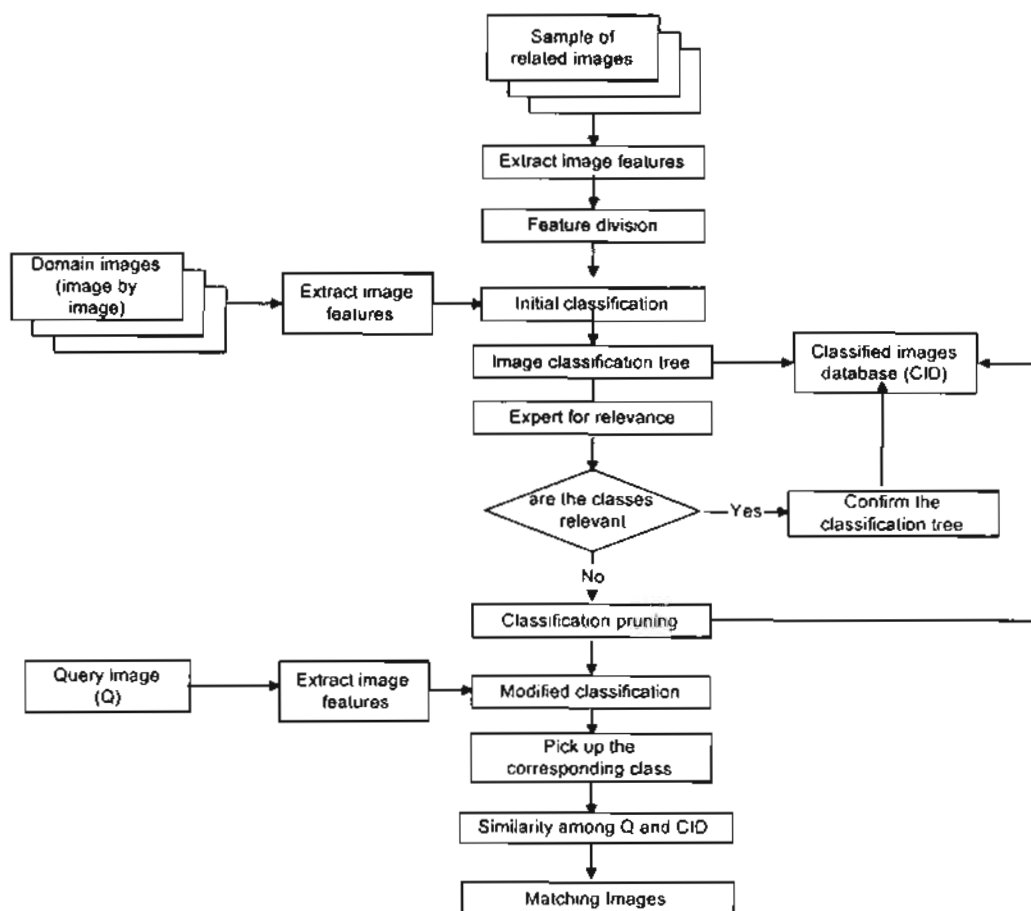


Figure 1. The proposed framework for image retrieval with relevance feedback.

Color feature is one of the most widely used features in image retrieval. Common color features or descriptors in regional based image retrieval systems include, color covariance matrix, color histogram, color moments and color coherence vector [10,11].

Texture provides surface characteristics for the analysis of many types including natural scenes, remotely sensed data, and biomedical modalities [12]. Importance of texture feature is due to its presence in many real as well as synthetic data. As tigers and cheetahs have the same colors but different texture patterns, so using color feature alone cannot clearly distinguish between them. This phenomenon gives clear justification for texture features to be used in CBIR along with color and shape [13].

Shape is a fairly well-defined concept. Shape features of general applicability include aspect ratio, circularity, Fourier descriptors, moment invariants consecutive boundary segments etc.[14].

Besides color, texture and shape, spatial location is also useful in region classification. Region centroid and its minimum bounding rectangle are used to provide spatial location information [15]. A survey of low level feature extraction is described in [3,9,16].

3.2 Features' division

Image classification can actually obtain by the intersection of features' sets (partitions) starting from the first feature interval to the last feature interval. Consequently, one must know the upper and lower limits of each feature then divide the features into segments. This section describes features'

divisions process. Choose some of images in the given domain to get normal distributions for each feature and make the desired divisions as follows:

- Extract the low level features of these images.
- Find the mean, μ_i and standard deviation, σ_i of the i^{th} feature
- Divide each feature into a number of interval, the width of which may be σ . Proper definition of features' intervals will lead to well defined clusters.
- Form the clustering table according to the feature divisions as shown in table 2.

Table 2. The features' intervals of images database

Intervals	Features					
	A	B	C	Z
Interval ₁	A ₁	B ₁	C ₁	Z ₁
Interval ₂	A ₂	B ₂	C ₂	
..
Interval _m	A _m	B _m	C _m	Z _m

3.3 Initial classification

A new image classification approach is presented here. This approach uses features' segments provided by the division process discussed above. It is based on two main actions; pick up the suitable feature interval that reflects specific feature and put up in a tree to successive features in consecutive order to find the matched feature interval. Repeating this processes will guarantee that actual classes are maintained as long as there exist images corresponding to their features. The following steps illustrate how initial classification algorithm is processing through an example of two successive images

1. Get all features of the first image.
2. Consider the value of the first feature, A, belongs to the interval set A₁.
3. Pick up the value of the adjacent feature, B, and suppose it belongs to the interval set B₂. At this level the current class (Cl₁) is characterized by $A_1 \cap B_2 \rightarrow A_1B_2$.

4. Moving to the adjacent feature, C, the picked up value belongs to the interval set C₃. Hence the initial classification algorithm indicates that the cluster Cl₁ will be updated as $A_1 \cap B_2 \cap C_3 \rightarrow A_1B_2C_3$

5. Continuing until reaching the final feature, Z, then the cluster Cl₁ will be defined by:

$$A_1 \cap B_2 \cap C_3 \cap D_2 \dots \cap Z_m \rightarrow A_1B_2C_3D_2 \dots Z_m$$

6. Repeat steps 1-5 for the next image

7. If the next image features belong to Cl₁, then allocate it to Cl₁ else establish new class, Cl₂, which may be defined by:

$$A_1 \cap B_3 \cap C_4 \cap D_1 \dots \cap Z_1 \rightarrow A_1B_3C_4D_1 \dots Z_1$$

The classes are shown in tree representation in figure 2. The tree starts from left and each level represent one feature. The set of blocks in each level correspond to possible segments of that feature.

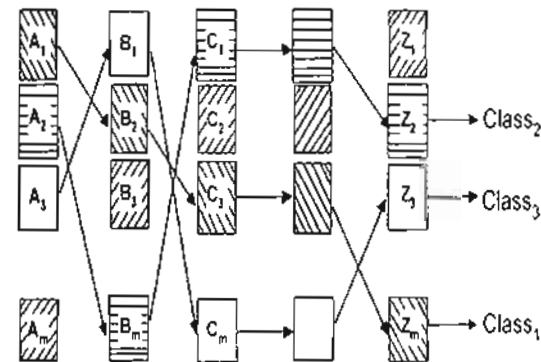


Figure 2. The classification tree.

3.4 Relevance feedback

Usually, relevance feedback schemes are distinguished into two different types of actions: actions that modify the query originally issued by the user (possibly considering multiple query points) and actions that modify the similarity measure used for ranking and retrieving images data in a CBIR system [17]. The relevance feedback results in reallocating images from class to another. This means that these reallocated images will reside in new

classes that are not matching with their feature values. The reallocation process will introduce some changes in classes sets, which as a result changes the features range limits for classes contributing in the images reallocation. The range limits might increase for classes that images move to, or decrease for classes images move from. A global view to the whole classes will show new limits for clusters' features. Interference may be resulted in the values of the same feature in two classes. This problem occurs when two classes, or more, overlap in all features. In this situation, a conflict will be there if retrieval is inquired for an image in the overlapped area of all features. The image could be belonged to both classes.

3.5 Classification pruning

A solution to the overlapping problem resulted from relevance feedback is proposed here. This solution is based on avoiding classes' conflict by redistributing images between classes in a way that two classes could not overlap in all features. The model is presented, for two interfered classes. It could be generalized to any number of interfered classes. As stated above, two classes (I and J) are considered to be interfered if and only if all features in class I overlap with their corresponding features in class J. The aim of this model is to distinguish the two interfered classes with the minimal features changes to both classes. So, the algorithm focuses only on one feature to make this distinguish. Moreover, the selected feature is chosen as the most dissimilar feature in both classes, which guarantees the lowest approximation in the features related to that feature after the class refinement process. The selected feature contains the lowest count of candidate images in the overlapped interval. If two (or more) features have the same number of images in the overlapped interval, the feature with the smallest normalized overlapping interval is selected. The algorithm continues with the two classes refinement and then repeated for all

pairs of interfered classes until no classes' causes conflict in the retrieval process. The following steps describe a computational model to redistribute class elements based on adaptation of range limits of selected features in the interfered classes:

1. Select the feature describing least similarity, Feature A:

$\min(\text{count}(i \in \{I_i \cap J_i\}) + \text{count}(j \in \{I_i \cap J_i\}))$, Where, i : element in class I & j : element in class J.

OR

$\min(I_i \cap J_i) / (I_i \cup J_i)$, $i = 1, 2, \dots, n$

Where, n : number of features.

2. Split overlapped and non-overlapped intervals in I_A , temp_{I_A} for class I and J_A , temp_{J_A} for class j

$\text{temp}_{I_A} \leftarrow I_A - J_A$

$\text{temp}_{J_A} \leftarrow J_A - I_A$

$\text{Inter}_A \leftarrow I_A \cup J_A$

3. Move non-overlapped sub-sets from I_A and J_A to temp_{I_A} and temp_{J_A} that matches temp_{I_A} and temp_{J_A} sets boundaries.

move $i \in \{I_A\} \rightarrow \{\text{temp}_{I_A}\}$

move $j \in \{J_A\} \rightarrow \{\text{temp}_{J_A}\}$

4. Calculate centers of overlapped sets in I_A and J_A

$I_c \leftarrow 1/k \cdot \sum\{I_A\}$, k : number of elements in I_A set

$J_c \leftarrow 1/m \cdot \sum\{J_A\}$, m : number of elements in J_A set

5. Define ΔI_A and ΔJ_A sets by redistributing overlapped I_A and J_A sets.

$\text{Inter}_{\min} = \min\{\text{Inter}_A\}$

$\text{Inter}_{\max} = \max\{\text{Inter}_A\}$

If $I_c - \text{Inter}_{\min} < J_c - \text{Inter}_{\min}$

Then

$\Delta I_{\min} = \text{Inter}_{\min}$

$\Delta J_{\max} = \text{Inter}_{\max}$

$\Delta I_{\max} = I_c + (J_c - I_c) \cdot (\text{Inter}_{\max} - I_c) / ((\text{Inter}_{\max} - I_c) + (\text{Inter}_{\max} - J_c))$

$\Delta J_{\min} = J_c - (J_c - I_c) \cdot (\text{Inter}_{\max} - J_c) / ((\text{Inter}_{\max} - I_c) + (\text{Inter}_{\max} - J_c))$

else

$\Delta J_{\min} = \text{Inter}_{\min}$

$\Delta I_{\max} = \text{Inter}_{\max}$

$$\Delta J_{\max} = J_c + (I_c - J_c) \cdot (Inter_{\max} - J_c) / ((Inter_{\max} - I_c) + (Inter_{\max} - J_c))$$

$$\Delta I_{\min} = I_c - (I_c - J_c) \cdot (Inter_{\max} - I_c) / ((Inter_{\max} - I_c) + (Inter_{\max} - J_c))$$

End If

6. Consolidate I_A and J_A sets

$$J_A \leftarrow temp_I_A \Delta I_A$$

$$J_A \leftarrow temp_J_A \Delta J_A$$

The interference region in I_A and J_A is split to both classes on a multiplicity basis. The multiplicities are defined by centers distribution over the overlapped area. Figure 3 shows a schematic layout for this overlap.

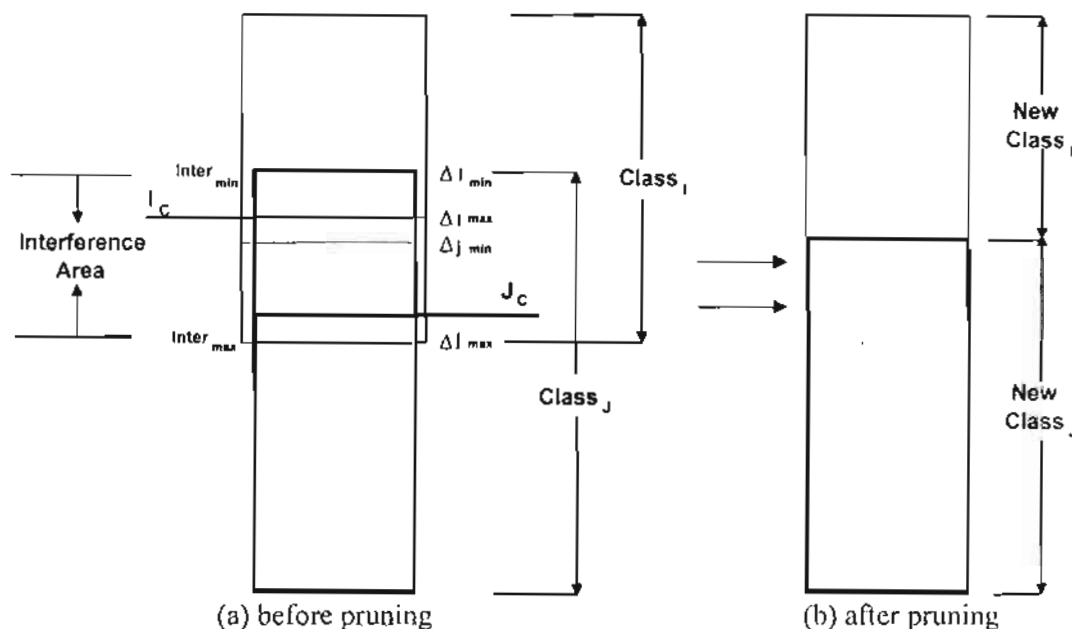


Figure 3. Interference region in most dissimilar feature before and after pruning.

4. Experimental Results

The database used in conducting the experiments is composed of 2250 images for students in a classroom and famous persons downloaded from the Internet. A sample of these images is shown in figure 4. The main goal of these experiments is to compare the system retrieval performance before and after applying each of the proposed frameworks. The retrieval performance was defined by precision, and recall of the retrieved images [18]. Precision (Pr) is defined as the ratio of the number of relevant images retrieved (Nr) to the number of total retrieved images (K), $Pr = Nr/K$.

Recall (Re) is defined as the number of retrieved relevant images (Nr) over the total number of relevant images available in the database (Nt), $Re = Nr/Nt$. The results of image retrieval are explored throughout different stages of the proposed frameworks. Figure 5 shows the retrieval subset images before applying the proposed frameworks. The performance values (precision and recall ratios) are 63.157 % and 72 % respectively. The retrieval subset images after applying the first proposed framework is shown in figure 6. The precision and recall ratios are increased to 68.4211 % and 78 %. The performance measures show high performance values.



Figure 4. Sample of images in the database.

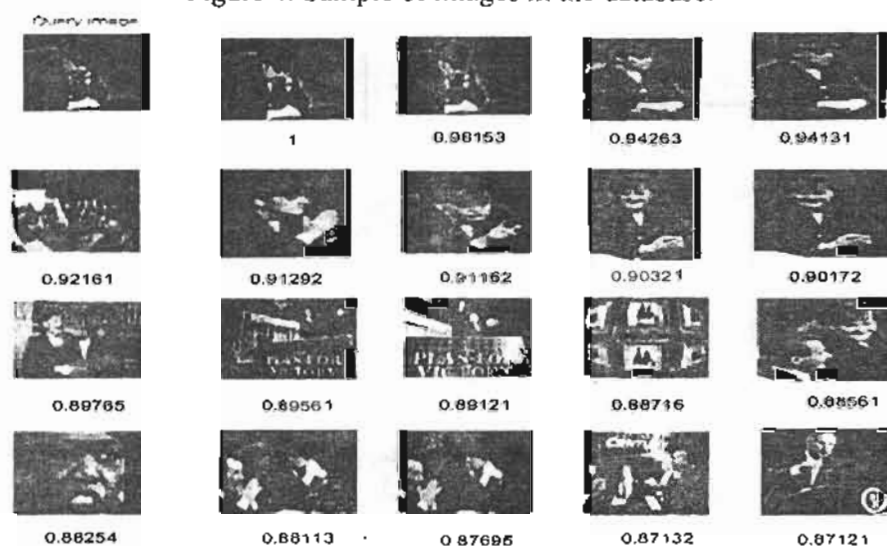


Figure 5. Retrieval subset images before applying the proposed frameworks.

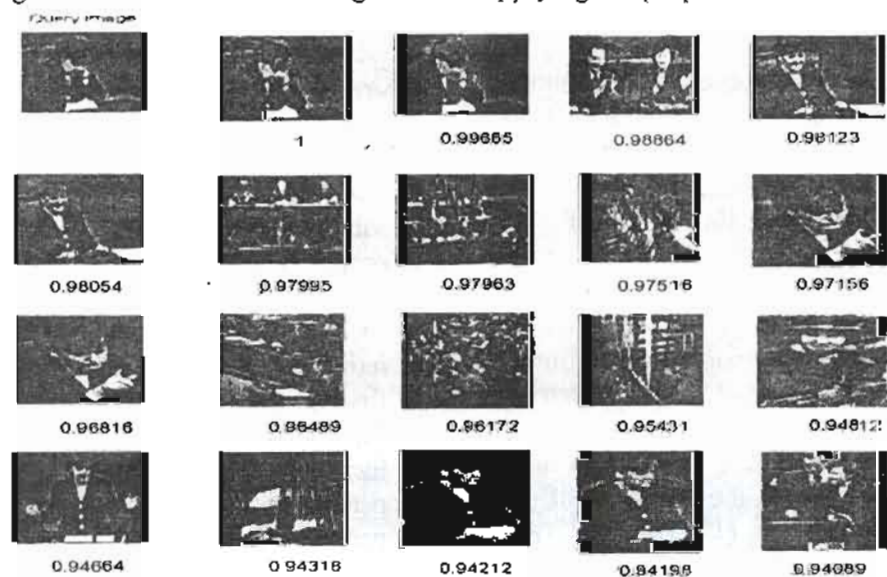


Figure 6. Retrieval subset images after applying the first framework.

For the second framework a same query image is entered; its features are extracted and compared with the classes' limits and the matched class is elected. A similarity measure is used to extract the most similar images within a specific class. This measure is based on the differences between the features of the images being compared. If t and q are the database and the query images respectively having the following feature vectors:

$$I_t = \{f_{t_1}, f_{t_2}, \dots, f_{t_n}\}$$

$$I_q = \{f_{q_1}, f_{q_2}, \dots, f_{q_n}\}$$

The similarity between both images is

measured by the minimum distance, which is:

$$D(I_t, I_q) = \sum_n d_n(I_t, I_q) \quad (9)$$

Where $d_n(I_t, I_q)$ is the distance between the values corresponding to feature n in t and q images. This distance is calculated as:

$$d_n(I_t, I_q) = \sqrt{(f_{t_n} - f_{q_n})^2} \quad (10)$$

Figure 7 shows the image retrieval of the same person after applying the second proposed framework. The precision in this case increased to 74.213 % and the recall is also increased to 84 %.



Figure 7. Retrieval subset images after applying the second framework.

Table 3 shows the precision and recall pairs before and after applying the two proposed frameworks.

Table 3. Performance values before and after applying the proposed frameworks.

	Precision	Recall
Without Classification	0.63157	0.72
First Classification Framework	0.68421	0.78
Second Classification Framework	0.74213	0.84

5. Conclusion

In this paper, two new image retrieval frameworks that extract and define similarity knowledge for a specific domain are presented. The first framework is based on the simple idea of the vector dot product, which generates different classes. The application of this framework yields high performance values. The second framework has the ability to generate classes of images dynamically according to the features of the concerned database. This reflected back a great enhancement to CBIR performance by dealing with a class of images rather than capturing all images

in the database and hence yielded to alleviating the dimensionality problem. Two models, initial classification and pruning classification, were introduced to establish images auto-classification and fine-tuning these classes based on relevance feedback. Results showed that retrieval is further enhanced after applying the pruning process, which eliminates any interference between classes.

References

- [1] Da Deng, "Content-based image collection summarization and comparison using self-organizing maps", *Pattern Recognition*, Volume 40, Issue 2, February 2007, Pages 718-727.
- [2] Nuno Vasconcelos, and Andrew Lippman, "Statistical Models of Video Structure for Content Analysis and Characterization" *IEEE Transactions on Image Processing*, Vol. 9, no. 1, January 2000 pp 3-19.
- [3] Ying Liu, Dengsheng Zhang, Guojun Lu and Wei-Ying Ma, "A survey of content based image retrieval with high-level semantics" *Pattern Recognition* Vol. 40, Issue 1, January, 2007 pp262-282.
- [4] Daniela Stan and Ishwar K. Sethi "Mapping low-level image features to semantic concepts" *Proceedings of SPIE: Storage and Retrieval for Media Database 2001*, PP 172-179.
- [5] Xin Zheng, Deng Cai, Xiaofei He, Wei-Ying Ma, and Xueyin Lin "Locality Preserving Clustering for Image Database" *proceedings of the 12th ACM Multimedia*, October 2004.
- [6] Nuno Vasconcelos "On the efficient evaluation of probabilistic similarity functions for image retrieval" *IEEE Transactions on Information Theory*, Vol. 50, No. 7, July 2004.
- [7] Sanjoy K. Saha, Amit K. Das, and Bhabatosh Chanda "Image retrieval based on indexing and relevance feedback" *Pattern Recognition Letters* 28 (2007) 357-366.
- [8] Joseph C. Giarratano "CLIPS User's Guide" version 6.2 March 31st 2002.
- [9] Yong Rui, Thomas S. Huang and Shih-Fu Chang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues" *Journal of Visual Communication and Image Representation* 10 (4) (1999), 39-62.
- [10] Ety Navon, Ofer Miller and Amir Averbuch, "Color image segmentation based on adaptive local thresholds" *Image and Vision Computing* 23 (2005) 69-85.
- [11] N. Jhanwar, S. Chaudhuri, G. Seetharaman and B. Zavidovique, "Content based image retrieval using motif co occurrence matrix" *Image and Vision Computing* 22 (2004) 1211-1220.
- [12] P. W. Huang' and S. K. Dai, "Image Retrieval by Texture Similarity" *Pattern Recognition* Vol. 36, Issue 3, March 2003, PP 665-679.
- [13] Manesh Kokare, B. N. Chatterji and P. K. Biswas, "Cosine - modulated wavelets Based texture features for content Based Image Retrieval" *Pattern Recognition letters* Vol. 25 (2004) pp 391- 398.
- [14] Mao-Hsiung Hung, Chaur-Heh Hsieh and Chung-Ming Kuo, "Similarity retrieval of shape images based on database classification" *Journal of Visual Communication and Image Representation* Vol. 17, Issue 5, October 2006, PP.970-985.
- [15] W. Y. Ma, B. Manjunath and Netra. "A toolbox for navigating large image database", *proceedings of IEEE International Conference on Image Processing*, 1997, pp568-571.
- [16] Mustafa Ozden and Ediz Polat. "A color image segmentation approach for content-based image retrieval", *Pattern Recognition*, Volume 40, Issue 4, April 2007, Pages 1318-1325.
- [17] Chunru Wan, Mingchun Liu. "Content - based audio retrieval with relevance feedback" *Pattern Recognition letters* 27 (2006) 85-92
- [18] Julia Vogel and Bernd Schiele. "Performance evaluation and optimization for content-based image retrieval" *pattern recognition* 39 (2006) pp 897-909.