

# A Two-Layer Data Model for Image Retrieval Systems

Been-Chian Chien  
Department of Information Engineering,  
I-Shou University,  
Kaohsiung, Taiwan, R. O. C.  
Email: cbc@isu.edu.tw

Jin-Der Li  
Institute of Information Engineering,  
I-Shou University,  
Kaohsiung, Taiwan, R. O. C.  
Email: m873002m@isu.edu.tw

Din-Yuen Chan  
Department of Information Engineering,  
I-Shou University,  
Kaohsiung, Taiwan, R. O. C.  
Email: dychan@isu.edu.tw

## Abstract

Content-based retrieval is one of the most important research issues in multimedia databases. Since the low-level features are used in content-based image retrieval usually, the gap between low-level features of images and high-level human concept becomes a problem. In this paper, we propose a two-layer image data model including the object layer and the image layer to resolve the problem. The object layer uses low-level image features such as color, shape and texture to describe each object of the image. The image layer represents an image by composing the objects in the object layer. Based on the two-layer image data model, we develop a prototype system for similar image retrieval. Two main algorithms in the system, the object-pair matching algorithm and the similarity matching algorithm, are proposed to measure the similarity between images. A two-layer relevance feedback mechanism is also proposed to update the weights in the two layers according to the user's responses for satisfying human subjective concept. The experiments show that the proposed approach can capture human perception and retrieve relevant images in two or three feedback processes effectively.

## 1. Introduction

Content-based image retrieval has become one of the main issues for image retrieval in the past few years. Content-based image retrieval systems use objective content, such as color, shape and texture, as the image index. These different low-level image features can be extracted by many image processing techniques automatically. However, the features cannot present the high-level semantic meaning of images. The cognitive gaps between low-level features and high-level human concept in an image should become a problem. For solving the problem and improving the efficiency and effectiveness in content-based image retrieval, a lot of techniques have been proposed by researchers. Some representative systems like QBIC [9] developed by IBM Almaden Research Center, VisualSEEK [15] of Columbia University, and MARS [12, 13] of University of Illinois are well-known. These systems adopt different low-level features including color, shape, texture and spatial relationship as their visual features. Although the low-level features describe the objective characteristics of images, it is still far from the human perception and subjectivity. The MARS system [12] proposes a relevance feedback mechanism based on the technique of textual information retrieval to overcome the gap between the low-level features and the user's perspective. The system introduces the concept of interactive retrieval to approximate the user's information need, however, their approach considers only one layer. The relevance feedback mechanism in MARS takes one image as an object, the feedback results will be

improved by refining the weights of various features in similarity metrics. Since a general user is usually attracted by the objects in an image instead of the features of color or texture, a single image layer is not enough for human subjectivity.

In this paper, we proposed a two-layer data model and develop an image retrieval system to retrieve the similar images that are the user's want to. For reducing the gap between human high-level concept and the low-level features of images, the proposed system employs the similarity matching and relevance feedback based on two layers: the object layer and the image layer. In the phase of database creation, we first segment an image into the best number of regions. We take each region as an object, then the features in the object and the relations among different objects are extracted and stored. While the phase of image retrieving, the similarity measure algorithms are given to match object-pairs in the object layer and measure the similarity between images in the image layer. Finally, the relevance feedback mechanism is developed to adjust the weights of measure functions in the object layer and the image layer. Our system is built on the environment of PC/Windows98, the experiments show that the proposed approach can capture human perception effectively and retrieve relevant images in few feedback processes efficiently.

The remainder of this paper is organized as follows: First, in Section 2, we formalize the two-layer image data model for the image retrieval system used in this paper. Section 3 depicts the method of feature extraction including image segmentation and the features used in the system including color, shape and spatial relationship. Then, the processing of similarity image retrieval are presented in Section 4. The relevance feedback technique is discussed in Section 5. The experimental results are shown in Section 6. Finally, concluding remarks are made in the last section.

## 2. The data model

The system is based on a two-layer image data model. We formalize the data model as two layers: the *object layer* and the *image layer*. Each image is viewed as the collection of a set of objects and each object in the image is represented as a set of low-level visual features. The object layer describing the structures of objects in an image by low-level image features is defined as  $O$ :

$$O = (P, V, F)$$

where  $P$  is the set of pixels inside a region or an object segmented from a raw image.  $V$  is a set of the multidimensional vectors in real number.  $F$  is a set of mapping functions that are used to extract the low-level features associated with the object  $O$ . Assume that  $F = \{f_1, f_2, \dots, f_k\}$ ,  $K$  is the number of extracting functions used by the system. We have  $F: P \rightarrow V$ , for each function  $f_i$  in  $F$ ,  $f_i(P) \rightarrow \langle v_{i1}, v_{i2}, \dots, v_{ii} \rangle$ , where  $i'$  is the length of

the  $i$ th feature vector,  $1 \leq i \leq K$ .

The image layer defines the composition of an image. The image model is a four-tuple:

$$(I, O, R, g)$$

where  $I$  is a raw image data, e.g. GIF or JPEG images.  $O$  is a set of objects segmented from the image  $I$  and each object is modeled by a triple as the object layer.  $R$  is a set of the relations among the objects in  $O$ , e.g. spatial relationships.  $g$  is a mapping function of the relations among the objects in  $I$ ,  $g: O \times O \rightarrow R$ .

### 3. Feature extraction

The first issue in the model is image segmentation. The extraction of features in an image depends on a good segmentation. Many segmentation approaches have been proposed in the past [14]. In our system, we modify Lin's approach [8] to segment color images into the best number of regions. First, the image is partitioned into blocks of size  $X_i \times Y_i$ . Each block is assigned a representative color, which is the mean value of colors for all pixels in the block. We set each block to be an individual region initially. Then each block is merged with its neighboring blocks having the smallest color difference into a new region. Such process is iterative until only one region remains. For determining the best number of segmented regions, the performance index  $Q_k$  [5] is defined as the sum of mean square errors between the image of the  $k$  segmented region and the original image. The best segmentation will be found while  $\rho(k)$  is maximum, where  $k > 0$ , and

$$\rho = \frac{Q_{k-1} - Q_k}{Q_k - Q_{k+1}}$$

After the segmentation, the low-level features in each object will be extracted. The different kinds of image features including color, shape and texture can be considered to describe the content of each object. In the proposed system, we select color and shape features only. For capturing human visual sense, the CIE LUV color space is applied to be the representative color feature of each region. The final color value  $[L, u, v]$  for each region is saved as the color feature. About the feature of shapes, the set of invariant moments proposed by Hu [6],  $\{\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7\}$ , is used to be the shape descriptors. Thus, the representations of the object layer in our system are

$$F = \{f_1, f_2\} \times f_1(O_i) = \langle L_i, u_i, v_i \rangle, \\ f_2(O_i) = \langle \phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7 \rangle$$

where  $O_i$  is the marked  $i$ th object in the image after segmentation.

In the image layer, spatial relationship is used to model the relations among the segmented objects of an image. For example, the set of spatial relationships is  $R$ ,

$$R = \{east, south, west, north, southeast, \\ southwest, northeast, northwest\}$$

and  $g(O_1, O_2) = north$ ; it means that the object  $O_1$  is north of the

object  $O_2$  for an image  $I$  with the set of objects  $O = \{O_1, O_2\}$ . The 2D String [1] is the most famous representation for describing spatial relations. However, the 2D String only supports similarity matching of three types and represents spatial relationships as symbolic strings, it is hard to be combined with the other numerical features like color and shape in our system. To merge various image features as a single ranking function, we modify the vector-based approach[3] as the spatial measuring method in our system. We describe the vector-based spatial matching approach proposed by Chien [3] in the following briefly.

Assume that there are two objects  $O_i$  and  $O_j$  segmented from an image. Let  $(px_i, py_i)$  and  $(qx_i, qy_i)$  denote the bottom-left and top-right coordinates of minimum boundary rectangle (MBR) containing  $O_i$ .  $(Dx_{ij}, Dy_{ij})$  are defined as the vector from  $O_i$  to  $O_j$ .

$$Dx_{ij} = qx_j - px_i \text{ and } Dy_{ij} = py_j - py_i.$$

Let  $O_i$  be the referential object, we define the following parameters:

$$Ix_i = |qx_i - px_i|, Iy_i = |qy_i - py_i|; \\ \alpha_x = \frac{Dx_{ij}}{Ix_i}, \beta_x = \frac{Ix_j}{Ix_i}, \alpha_y = \frac{Dy_{ij}}{Iy_i}, \beta_y = \frac{Iy_j}{Iy_i}, \\ \gamma_x = \frac{\alpha_x + \beta_x}{\beta_x + 1}, \gamma_y = \frac{\alpha_y + \beta_y}{\beta_y + 1}.$$

The relative distance between the objects  $O_i$  and  $O_j$  on  $x$ -axis and  $y$ -axis are  $f(X_{ij})$  and  $f(Y_{ij})$ , respectively.  $f(X_{ij})$  and  $f(Y_{ij})$  can be used to measure the spatial similarity. They are defined as:

$$f(X_{ij}) = \begin{cases} \frac{1}{(1 - \gamma_x)} & \text{if } \alpha_x + \beta_x < 0; \\ (1 - \frac{1}{\gamma_x}) + 3 & \text{if } \alpha_x > 1; \\ 2\gamma_x + 1 & \text{otherwise;} \end{cases} \\ f(Y_{ij}) = \begin{cases} \frac{1}{(1 - \gamma_y)} & \text{if } \alpha_y + \beta_y < 0; \\ (1 - \frac{1}{\gamma_y}) + 3 & \text{if } \alpha_y > 1; \\ 2\gamma_y + 1 & \text{otherwise.} \end{cases}$$

After extracting the image features, the objects and corresponding features are represented by multi-dimensional attributes as follows:

$$[Image\_id, Object\_id, Feature\_list].$$

$Image\_id$  is the identity of an image,  $Object\_set$  is the set of objects contained in the image, and  $Feature\_list$  is the features of each object in the image. In our system, the features in  $Feature\_list$  are

$$[L, u, v, \phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7, px, py, qx, qy].$$

$[L, u, v]$  and  $[\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7]$  are the features of color and moments as described earlier. The  $(px, py)$  and  $(qx, qy)$  are the bottom-left and top-right locations of object's MBR. The above features will be able to be easily indexed by high-dimensional data

structures such as K-D tree or R-tree to improve the performance of access.

#### 4. Query processing and similarity measure

About the image query phase, the user interface in our system supports two query methods for users:

1. Query by image example: The user gives a sample picture to the system, and the system will find the images which are similar to the sample image.
2. Query by sketch: If the user cannot yield any proper picture as a query picture on his hand, our system allows users to specify the query by drawing the sketch roughly.

While a query image is given, the same processes of feature extraction are performed. The extracted features of the query image are then compared with the features in the feature database by the similarity measure algorithms. Since the proposed model has two layers, the similarity measure includes two main matching algorithms. The first algorithm is to find the best matching of object-pairs between the query image and the target image in the object layer. The second algorithm gives the overall similarity measure of images including the object layer and the image layer.

The goal of the object-pair matching algorithm is to find the best object-pairs whose objects belong to the query image and the target image respectively. Assume that there are  $m$  objects in query image  $Q$  denoted as  $O_1, \dots, O_m$ , and a target image  $T$  in which contains  $n$  objects denoted as  $O'_1, \dots, O'_n$ . The best object-pairs are found by the following steps:

Step 1: Constructing a matrix  $D_{mn}$  of size  $m \times n$ ,  $D_{mn} = [d_{ij}]_{m \times n}$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ , where  $d_{ij}$  is an element in the matrix which stands for the dissimilarity between  $i$ th object  $O_i$  in the query image  $Q$  and the  $j$ th object  $O'_j$  in the target image  $T$ . The  $d_{ij}$  is defined as

$$d_{ij} = \sum_{k=1}^K w_{ik} \Delta_k(O_i, O'_j),$$

where  $\Delta_k(O_i, O'_j)$  is the normalized dissimilarity between  $O_i$  and  $O'_j$  of the  $k$ th feature  $f_k$  in the object layer, and  $w_{ik}$  is the weight of feature  $f_k$  for the object  $O_i$  in the query image.

Step 2: Let  $\min\{d_{ij}\}$  be the minimum element of  $D_{mn}$ . If  $\min\{d_{ij}\}$  is found in  $i = a$  and  $j = b$ , the object  $O_a$  in  $Q$  and the object  $O'_b$  in  $T$  will be grouped into an object-pair.

Step 3: Discarding the  $a$ th row and the  $b$ th column in  $D_{mn}$ , we get a new matrix  $D_{(m-1)(n-1)}$ . The Step 2 is repeated on the new matrix until all objects in either image  $Q$  or image  $T$  have been matched or the remaining values in  $D_{(m-1)(n-1)}$  are too large.

The dissimilarity of an object-pair is measured by their representative colors and shape descriptors. Let  $\langle L, u, v \rangle$  be the vector of feature  $f_1$  and the seven invariant moments  $\langle \phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7 \rangle$  be the vector of feature  $f_2$ . The dissimilarity of  $f_1$  and  $f_2$  are defined by the measure functions  $\delta_1$  and  $\delta_2$ :

$$\delta_1(O_i, O_j) = \sqrt{(L_i - L_j)^2 + (u_i - u_j)^2 + (v_i - v_j)^2},$$

$$\delta_2(O_i, O_j) = \sqrt{\sum_{k=1}^7 (\phi_k^{O_i} - \phi_k^{O_j})^2}.$$

The  $\Delta_1(O_i, O_j)$  and  $\Delta_2(O_i, O_j)$  are the normalized dissimilarity of  $\delta_1(O_i, O_j)$  and  $\delta_2(O_i, O_j)$ , respectively. The detailed algorithm is listed as follows.

**Algorithm:** object-pairs matching algorithm

Input: the color and moment features of pictures  $P_1$  and  $P_2$

Output: the set of object-pair

```

{
  for i = 1 to m
    for j = 1 to n
      compute  $D_{mn}$ ,  $d[i, j] = \sum_{k=1}^K w_{ik} \Delta_k(O_i, O_j)$ ;
  object-pair =  $\emptyset$ ;
  k = 1;
   $d[a, b] = \min\{d[i, j] \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ ;
  While ( $k \leq \min(m, n)$ ) and ( $d[a, b] \leq \text{MIN\_DIS}$ )
    // MIN_DIS is a specified
  {
    // minimum degree of dissimilarity
    object-pair = object-pair  $\cup \{(O_a, O_b)\}$ ;
    for j = 1 to n
       $d[a, j] = \infty$ ;
    for i = 1 to m
       $d[i, b] = \infty$ ;
    k = k + 1;
     $d[a, b] = \min\{d[i, j] \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ ;
  }
}

```

After object-pairs are found, the similarity of two pictures is measured by the similarity matching algorithm. In addition to the image features in the object layer, this algorithm also considers the spatial relationships among the objects in the image layer. The similarity matching algorithm is shown as the following steps.

Step 1: Assume that the object-pair matching algorithm finds  $r$  object-pairs for the given two images  $P_1$  and  $P_2$ . The set of object-pairs is  $\{(O_1^{P_1}, O_1^{P_2}), (O_2^{P_1}, O_2^{P_2}), \dots, (O_r^{P_1}, O_r^{P_2})\}$ .

Step 2: The dissimilarity between pictures  $P_1$  and  $P_2$ , denoted  $Dis(P_1, P_2)$ , is computed by

$$Dis(P_1, P_2) = \sum_{k=1}^K W_k \Psi_k(P_1, P_2) + W_R \Gamma(P_1, P_2),$$

$$\text{where } \Psi_k(P_1, P_2) = \sum_{i=1}^r \Delta_k(O_i^{P_1}, O_i^{P_2})$$

$$\text{and } \Gamma(P_1, P_2) = \sum_{i=1}^{r-1} \sum_{j=i+1}^r \varphi(O_j^{P_1}, O_j^{P_2}).$$

For two distinct pictures  $P_1$  and  $P_2$ , assume that they both consist of two objects  $O_i$  and  $O_j$ . The  $\Delta_k(O_i^{P_1}, O_j^{P_2})$  is defined as above. The  $\varphi(O_j^{P_1}, O_j^{P_2})$  is defined as the normalized dissimilarity of spatial relationship between  $(O_i^{P_1}, O_j^{P_1})$  and  $(O_i^{P_2}, O_j^{P_2})$ . The  $\varphi(O_j^{P_1}, O_j^{P_2})$  is the Euclidean distance of the spatial relative distances  $[f(X_{ij}^{P_1}), f(Y_{ij}^{P_1})]$  and  $[f(X_{ij}^{P_2}), f(Y_{ij}^{P_2})]$ .

$= W_K = W_R = 1/(K+1)$  or  $W_1+W_2+\dots+W_K = W_R = 0.5$ , and the  $\Delta_{ik}^j$  in the update formulas has to be replaced by  $\psi_k^j$  and  $\Gamma$ .

### 5.1. The update of weights in the object layer

The goal of the object layer is to find the best object-pairs of the query image and the target image for satisfying the user's visual perception. For example, a user prefers an object with a shape of triangle in the image. However, in the initial measure may concern the same importance of color and shape features. The image with triangle objects may not match the objects with the similar shape owing to the domination of a large color difference. To overcome this problem, the weights of features should be refined using the user's relevance feedback information.

We give the same value to the weights of different features initially while the first measuring is made. Assume that there are  $r$  objects in the query image and the number of features used in the object layer is  $K$ . Let  $\omega_k^i$  be the weight of the  $k$ th feature for the  $i$ th object at the  $i$ th feedback time,  $0 \leq \omega_k^i \leq 1$  and  $\omega_1^i + \omega_2^i + \dots + \omega_K^i = 1$ . Thus,

$$\omega_{ik}^0 = 1 / K, \quad \text{for } 1 \leq i \leq r \text{ and } 1 \leq k \leq K.$$

For each feedback processing, the user marks the images as 'yes' if the image is relevant or 'no' if the image is irrelevant. Then, the weights in the object layer will be updated as the following formulas:

$$\sigma_{ik}^{t+1} = \left( \omega_{ik}^t + \frac{\sum_{j=yes} \Delta_{ik}^j - \sum_{j=no} \Delta_{ik}^j}{\sum_{j=yes} \Delta_{ik}^j + \sum_{j=no} \Delta_{ik}^j} \right),$$

$$\omega_{ik}^{t+1} = \left( \sigma_{ik}^{t+1} / \sum_{j=1}^K \sigma_{ij}^{t+1} \right),$$

where  $\Delta_{ik}^{yes}$  and  $\Delta_{ik}^{no}$  denote the dissimilarity of the objects in the images marked 'yes' and 'no' respectively.

### 5.2. The update of weights in the image layer

The same problem in the object layer will occur at the image layer. In an image, the features of objects and the relations among objects should be considered in the image layer simultaneously. The update of weights thus includes the features of objects and the spatial relationship among objects in images. The update method is similar to the object layer except the spatial relationships. The weights of features  $W_k$  and  $W_R$  can be initialized to be  $W_1 = W_2 = \dots$

## 6. Experimental results

For demonstrating the efficiency and effectiveness of the proposed approach, we build a prototype system on PC/Win98. The total number of tested images is about 2500. Types of the images include about 1000 images of photographs with the same size of 256x256 and about 1500 images of trademark with different size. First, the test images are segmented and the features are extracted by the system automatically. Then the features are stored in the feature database as the representation form in Section 3.

We made two different experiments. The first experiment gives a trademark as the query image and evaluates the effectiveness by precision-recall model. Owing to the limitation of the paper length, we only show the evaluation results in Table 1 and Figure 1. There are three retrieving processes totally including two feedback processes. The result of the first feedback retrieval improves the original measuring and the second feedback retrieval improves the first feedback retrieval indeed.

The second experiment gives two sketches as shown in Figure 2 and Figure 5 to be the query images. Figure 2 is the face-like sketch drawn by the user. Figure 3 is the first retrieval result. We then mark the relevant images as 'O' and the irrelevant images as 'X'. After the marking, the result of the first feedback retrieval is shown in Figure 4. The changes of weights in image layer are listed in Table 2.  $W_C$ ,  $W_S$  and  $W_R$  stand for the weights of color, shape and spatial relationship respectively. It depicts that the spatial relationship is more important than others features. For the other image in Figure 5, the relevant retrieval results are shown in Figure 6, Figure 7 and Figure 8. Since we select the images with the similar color and spatial relationships in the feedback processes, the final result shows that the weights of features response the fact and the images with the similar objects can be found effectively.

From the experiments, the system based on the proposed two-layer model can behave more effective than the model with single layer like MARS[12,13]. Furthermore, the number of features used in two-layer model is less than the one-layer model. The processing time of features extraction can be saved, the efficiency is thus improved.

0	Precision(%)	60	52	50	51	48	44	48
	Recall	0.12	0.17	0.20	0.32	0.38	0.51	0.62
1	Precision(%)	80	61	50	48	47	45	47
	Recall	0.19	0.22	0.24	0.33	0.42	0.53	0.61
2	Precision(%)	100	90	84	51	48	45	42
	Recall	0.2	0.4	0.62	0.7	0.74	0.7	0.8

Table 1: Precision and recall for Trademark query.

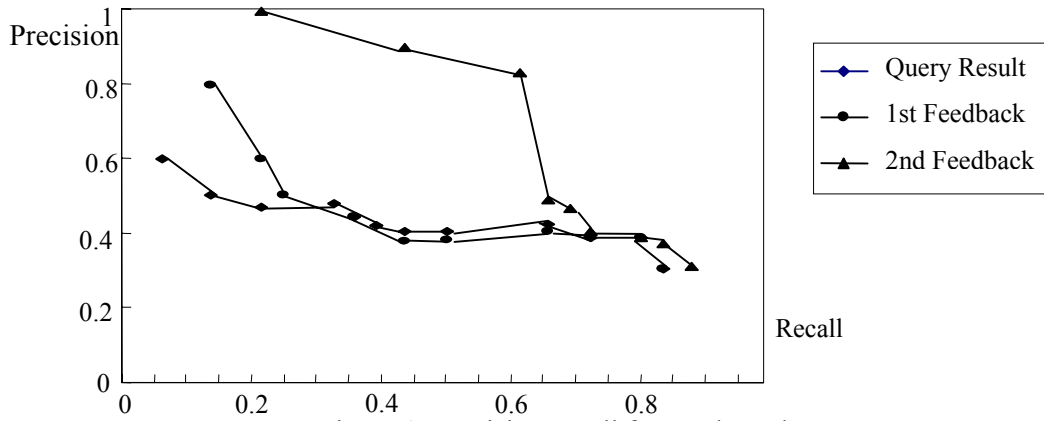


Figure 1: Precision-recall for Trademark query.



Feedback	$W_C$	$W_S$	$W_R$
0	0.3333	0.3333	0.3333
1	0.3545	0.2118	0.4337

Figure 2: The 'face' query image.

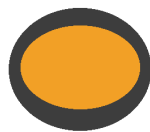
Table 2: The update of weights in 'face' query



Figure 3: The initial retrieval result of 'face' query.



Figure 4: The first feedback result after the marking of Figure 3



Feedback	$W_C$	$W_S$	$W_R$
0	0.3333	0.3333	0.3333
1	0.4146	0.2251	0.3602
2	0.4986	0.2000	0.3014

Figure 5: The 'circle' query image.

Table 3: The update of weights in 'circle' query.

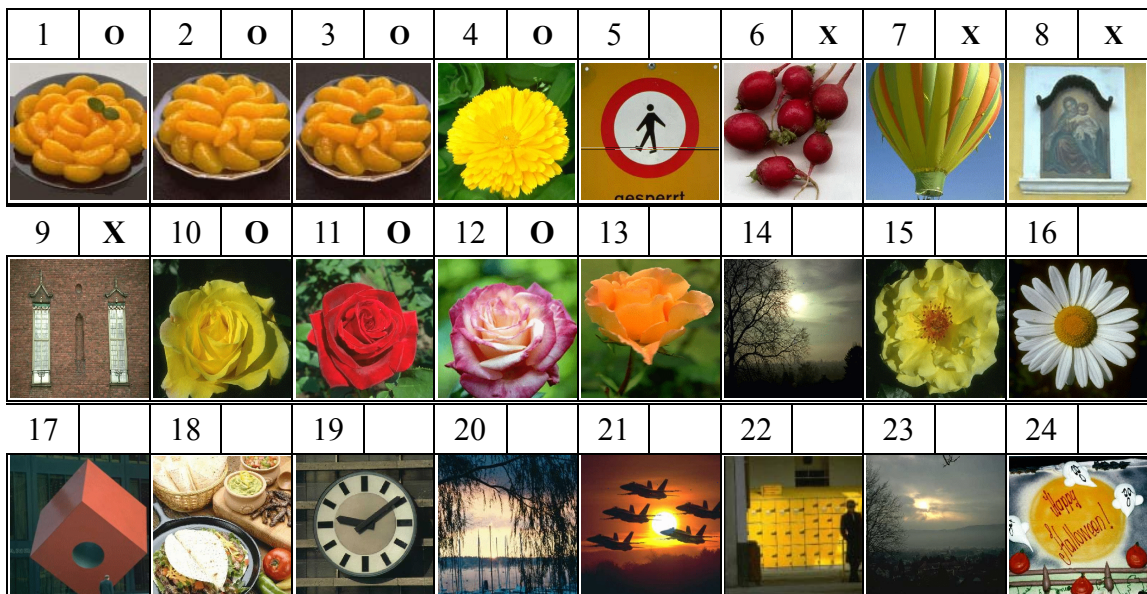


Figure 6: The initial retrieval result of 'circle' query.

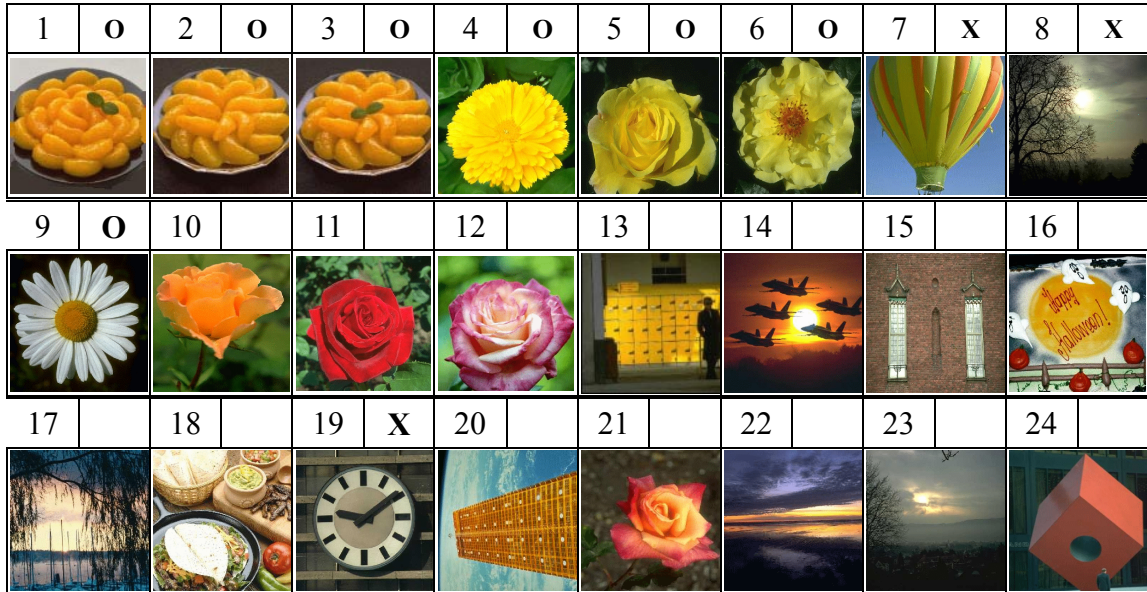


Figure 7: The first feedback result after the marking of Figure 6

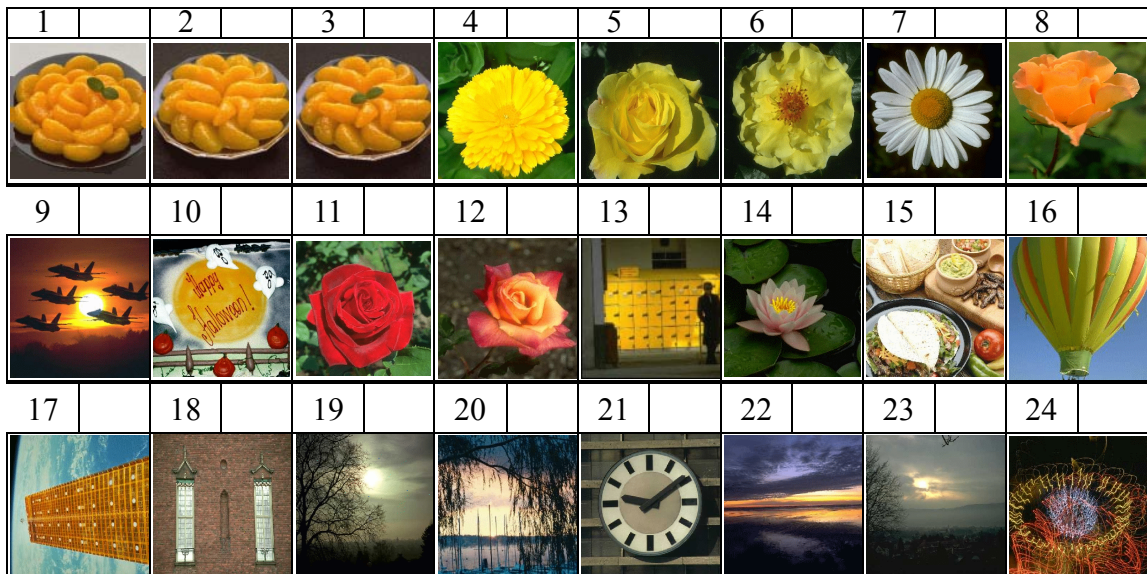


Figure 8: The second feedback result after the marking of Figure 7.

## 7. Conclusions

Content-based image retrieval is one of the most important techniques for retrieving multimedia data. It is difficult for researchers to create the so called fully automatic image retrieval to capture human perception. The previous researches proposed the interactive retrieval approach focusing on the different features of images and similarity metrics. In this paper, we motivate a two-layer data model including the object layer and the image layer. We propose an object-based interactive retrieval method to reduce

the gap between human high-level concept and the low-level features of images. The two-layer similarity measure algorithms are given to find the similar images from image databases and a relevance feedback mechanism is developed to adjust the weights of measure functions in the object layer and the image layer. The experimental results also show that the proposed approach is effective and efficient. The future work on expanding the object with visual perception to integrate with human conceptual semantics is under investigated.

## References

- [1] S. K. Chang, Q. Y. Shi and C. W. Yan, "Iconic indexing by 2D-String", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 9, NO 3, pp413-428,1987
- [2] C. C. Chang, "Color Image Retrieval Based on 2D String", *Master thesis, Department of Computer Science, National Tsing Hua University, Taiwan, 1998.*
- [3] B. C. Chien and P. C. Chieng, "A New Approach of Similarity Measurement for Iconic Image Retrieval", in *Proceeding of International Conference on Computational Intelligence and Multimedia Applications, 9-11, Feb. 1998, Churchill, Australia.*
- [4] P. C. Chieng, B. C. Chien, "A New Similarity Metric Based on 2D Vector for Iconic Image Retrieval", in *Proceedings of 1998 International Computer Symposium, Workshop on Software Engineering and Database Systems, 16-19, Dec., 1998, Tainan, Taiwan*, pp.85-92 .
- [5] D., Duta and P. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.
- [6] J. D. Foley, A. V. Dam, S. K. Feiner, and J. F. Hughes, "Computer Graphics: principles and practice," 2nd edition in C, Addison-Wesley, 1996.
- [7] R. C. Gonzalez and R. E. Woods, "Digital Image Processing", third edition, Addison-Wesley, Reading, MA, 1992.
- [8] H. C. Lin, L. L. Wang, and S. N. Yang, "Color image retrieval based on hidden Markov models," *IEEE Transactions on Image Processing*, Vol. 6, No. 2, 1997.
- [9] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, and C. Faloutsos, "The QBIC project: querying images by content using color, texture, and shape," *Research Report*, No. 9203, IBM Almaden Research Center, 1993.
- [10] A. K. Jain and A. Vailaya, "Image Retrieval Using Color and Shape", *Pattern Recognition*, Vol. 29, No. 8, pp. 1233-1244, 1996.
- [11] B. M. Mehtre, M. S. Kankanhalli and W. F. Lee, "Content-Base Image Retrieval Using A Compost Color-Shape Approach", *Information Processing And Management* , Vol. 34, Issue. 1, 1998 , pp. 109-120
- [12] Y. Rui, T. S. Huang, and S. Mehrotra, Relevance Feedback Techniques in Interactive Content-Based Image Retrieval, in *Proceedings of IS&T and SPIE Storage and Retrieval of Image and Video Databases VI* , January 24-30, San Jose, CA, 25-36, 1998.
- [13] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, Relevance Feedback: A Power Tool in Interactive Content-Based Image Retrieval, *IEEE Trans. on Circuits and Systems for Video Technology* , *Special Issue on Segmentation, Description, and Retrieval of Video Content*, Vol. 8(5), 644-655, 1998.'
- [14] Y. Rui, T. S. Huang, and S. F. Chang, Image Retrieval: Current Techniques, Promising Directions and Open Issues, *Journal of Visual Communication and Image Representation*, Vol. 10, 1-23, 1999.
- [15] J. R. Smith and S. F. Chang, Visually Searching the Web for Content, *IEEE Multimedia*, Vol. 3, 12-20.