COLOR AND SHAPE BASED INDEXING USING SCALE-SPACE REPRESENTATIONS

Binnur Kurt, Muhittin Gökmen
Istanbul Technical University
Computer Engineering Department
Maslak 80626, İstanbul
{kurt,gokmen}@ce.itu.edu.tr

Abstract

Performance of an image or video database is directly related to indexing performance. Proposed indexing methods usually rely on low-level features such as color, texture and shape information extracted from an image or a key-frame selected within a shot, which are very sensitive to small changes in intensity or lightning conditions. In this study, we show that one can extract more robust and rich index vectors based on color and shape through using scale-space representation. Color and shape-based index vectors are derived by utilizing their behavior throughout the scale-space. Developed query system allows users to perform color-only, shape-only and hybrid queries. From the experimental results, it is found that the proposed method give more powerful and rich representations compared to the methods that only utilize color-based or shape-based indexes, and could index images with similar color or orientation histograms. Moreover the proposed algorithm yields better and consistent results especially for images with similar content.

Keywords

Multimedia, content-based retrieval, scale-space.

1. INTRODUCTION

Rapid growth in storage capacity, communication bandwidth and computation capability has made multimedia storage, transmission and processing much easier, which leads to high volume multimedia archives. Finding a particular data for processing or browsing purposes from such a multimedia archive becomes a major problem especially when the irregular nature of multimedia data is considered. One solution to the problem is called indexing in which the content of image or video is encoded with bits in such a way that encoded bits point where the actual data resides and they have compact representation compared to the original data. The encoding methods usually involve the extraction of low level features, such as color, shape, texture ([1],[2],[3]). Then the index vectors are formed based on these features. Most frequently used index vector in image or video database is color and edge histogram. Histogram is the frequency distribution of the feature values along the image. In this study, we have showed that one can obtain more rich and consistent index values from color and shape information using scale-space image representation.

Histogram is one of the most frequently used color based index vector. Let $I(p), p \in P = \{(x_i, y_j) | i \in \{1,\ldots,M\}, j \in \{1,\ldots,N\}\}$ denote the image color distribution. In this case, normalized color histogram of the image, $h_k(I), k \in \{0,1,\cdots,L-1\}$, is defined as

$$h_k(I) = \frac{1}{|P|}\sum_{p\in P}\delta(I(p)-k)$$
$$\sum_{k=0}^{L-1} h_k(I) = 1 \qquad (1)$$

Usage of color histogram as an index vector is based on the assumption that similar images have similar color histograms. However for some images having dissimilar content may have similar color histogram distributions or vice-versa. This is due to the fact that color histogram does not take spatial distribution of colors into account as well as their occurrence distribution.

There are several studies on color histogram refinement, modeling spatial distribution as well as color occurrence. Color correlogram method [4] computes a color correlogram matrix which is similar to the co-occurrence matrix defined for texture analysis of images. This matrix counts how many times a color occurs in a distance from the given pixel. Pass and Zabih used another approach [5] called Color Coherence Vector, which partitions pixels based on their spatial coherence. A coherent pixel is a part of contiguous region, while an incoherent pixel is not. The color space is discretized into distinct colors called buckets. A pixel is coherent if the size of its connected component exceeds a fixed value; otherwise, the pixel is incoherent.

Another feature describing the content of an image is shape. Shape provides additional image attributes in the absence of color information or in the presence of images with similar colors. A simple shaped-based index vector is the direction histogram of the edges. Jain has combined color histogram and edge direction histogram and applied it to logo database. The study reports successful results especially for logo images. Another study based on histogram is given in [8]. The effect of noise and shape deformations is reduced by carrying the direction histogram at different detail levels in scale-space.

In this study, color and shape-based indexes are derived through utilizing their behavior through scale-space. Scale-space analysis is introduced in section 2. Derivation of color and shape based indexes are also explained in this section. A content-based query system is introduced in section 3. Experimental results are given in section 4.

2. SCALE-SPACE ANALYSIS

$\lambda\tau$-space representation ([8],[9]) is a scale-space [10] representation. $\lambda\tau$-space representation of an image, $I(p)$, is obtained by convolving it with a filter denoted by $R_{\lambda,\tau}(p)$:

$$I'_{\lambda,\tau}(p) = T(I(p)) = \sum_{q\in P} I(q) R_{\lambda,\tau}(p-q) \qquad (2)$$

The histogram of the image transformed onto $\lambda\tau$-space using the definition by (1) is formulated as

$$h_k(I'_{\lambda,\tau}) = \sum_{p\in P}\delta(I'_{\lambda,\tau}(p)-k)$$

$$= \sum_{p \in P} \delta\left(\sum_{q \in P} I(q) R_{\lambda,\tau}(p-q) - k \right) \tag{3}$$

In the original histogram, same pixel values sum to the same histogram bin while this is not always true for the histogram of the projected image. Let us assume that pixel values are equal at two distinct points ($p_1 \neq p_2$) and $s = I(p_1) = I(p_2)$ is the pixel value at these points. p_1 and p_2 contribute to the same histogram bin (i.e., $h(s)$) for the original image. Projected values at these points need not have to be equal, hence may not contribute to the same histogram bin ($h_k(I'_{\lambda,\tau})$) for the projected image:

$$\begin{aligned} s'_1 &= I'_{\lambda,\tau}(p_1) = T(I(p_1)) \\ s'_2 &= I'_{\lambda,\tau}(p_2) = T(I(p_2)) \end{aligned} \tag{4}$$

s'_1 and s'_2, could be different depending on the local distribution of intensity at the vicinity of the points p_1 and p_2.

2.1. Color-Based Index

We define a new index feature taking the spatial distribution of pixel values into account as well as their frequency:

$$\Phi^c_{\lambda,\tau}(m,n) = \sum_{p \in P} \delta(I(p) - m) \cdot \delta\left(I'_{\lambda,\tau}(p) - n\right) \tag{5}$$

It can be easily verified that one dimensional histogram can be obtained from the two dimensional index vector ($\Phi_{\lambda,\tau}(m,n)$) simply by computing its marginal distributions along m and n as the following

$$\begin{aligned} h_k(I) &= \sum_m \Phi^c_{\lambda,\tau}(k,m) \\ h_k\left(I'_{\lambda,\tau}\right) &= \sum_m \Phi^c_{\lambda,\tau}(m,k) \end{aligned} \tag{6}$$

The properties and the performance of the new feature are investigated in details in [11].

2.2. Shape-Based Index

The approach based on scale-space analysis, introduced above, is also used to derive a shape-based index vector from the edge map and its orientation information. Orientation histogram of edges ($E(p)$) is used as shape-based index. Orientation histogram ($\phi_h(k)$) is defined as

$$\begin{aligned} \phi_h(k) &= \frac{1}{C} \sum_{p \in P} E(p) \delta(\theta(p) - \theta_k) \\ C &= \sum_{p \in P} E(p) \\ \theta_k &\in \left\{ \theta_i \,|\, \theta_i = \frac{2\pi}{A} i; i = 0,1,\ldots,A-1 \right\} \\ E(p) &= \begin{cases} 1 & ; \quad p \text{ is on an edge point} \\ 0 & ; \quad otherwise \end{cases} \end{aligned} \tag{7}$$

where $\theta(p)$ shows the direction of the edges. Let $E_{\lambda,\tau}(p)$ denote the edge map which belongs to the image projected to the λτ-space. In this case, shape-based index, $\Phi^{s}_{\lambda,\tau}(u,v)$, is defined as

$$\Phi^{s}_{\lambda,\tau}(u,v) = \sum_{p\in P} (E(p)\delta(\theta(p)-u)) \cdot \left(E_{\lambda,\tau}(p)\delta\left(\theta_{\lambda,\tau}(p)-v\right)\right) \tag{8}$$

where $\theta(p)$ denotes the direction of edges belonging to the original image and $\theta_{\lambda,\tau}(p)$ denotes the direction of edges belonging to the image projected to the λτ-space.

3. SHAPE AND COLOR-BASED QUERY

In order to retrieve images from the database in a fast and efficient way index vectors should be encoded with as few bits as possible and a distance measure between these index vectors should be defined. In this chapter, a compact representation for index vectors and distance measure are introduced.

3.1. Compact Representation

The size of the histogram, $\Phi_{\lambda,\tau}(m,n)$, is L^2. Typical value for L is 256 in which case the size is 64kB. In order to reduce the size, we have used binary split vector quantization [12]. Vector quantization is a process to assign pixel values in one of a finite number of vectors. First step in vector quantization is the decomposition of vector set. In binary split tree approach, these vectors are determined in such a way that the quantization error is minimized

$$E = \sum_{n\in Q}\sum_{s\in C_n} \|I_s - q_n\|^2 \tag{9}$$

where $\{q_n\}$ is the set of quantization vectors and C_n is the set of pixels assigned to the vector q_n. Initially all pixels belong to the same class whose vector is the average of the image. Then the class is divided into two sub classes denoted by C_{2n} and C_{2n+1}. The vectors associated with these sub classes are chosen based on the second order statistics within the class

$$\begin{aligned} R_n &= \sum_{s\in C_n} I_s I_s^T \\ m_n &= \sum_{s\in C_n} I_s \\ K_n &= |C_n| \end{aligned} \tag{10}$$

Quantization vector (q_n) is assumed to be equal to the class mean:

$$q_n = \frac{m_n}{K_n} \tag{11}$$

Class covariance matrix is given as

$$\tilde{R}_n = R_n - \frac{1}{K_n} m_n m_n^T \tag{12}$$

Once a class (C_n) is decided to divide into two classes, new class vectors (q_{2n}, q_{2n+1}) are computed. This computation requires the calculation of unit vector e_n which minimizes the expression

$$\sum_{s\in C_n} \left((I_s - q_n)^T e\right)^2 = e^T \tilde{R}_n e \tag{13}$$

This is the largest eigenvalue of $\tilde{R}_n$. Once the unit vectors are obtained, the pixels belonging to the class C_n are assigned to C_{2n} or C_{2n+1} as follows

$$\begin{aligned} C_{2n} &= \left\{ s \in C_n : e_n^T I_s \leq e_n^T q_n \right\} \\ C_{2n+1} &= \left\{ s \in C_n : e_n^T I_s > e_n^T q_n \right\} \end{aligned} \tag{14}$$

Splitting classes stops either when maximum vector number is reached or when the class variance is less than a predefined threshold. If there are S number of vectors, size of the histogram, $\Phi_{\lambda,\tau}(m,n)$, becomes S^2. S is usually chosen as 10.

3.2. Merging Color and Shape Based Indexes

The merged index vector denoted as $\Phi_{\lambda,\tau}$ is formed by concatenating two matrices as follows

$$\Phi_{\lambda,\tau} = \left[\Phi^c_{\lambda,\tau} \middle| \Phi^s_{\lambda,\tau} \right] \tag{9}$$

The distance between joint index vectors of two images (I and J), $D(I,J)$), is computed as the weighted sum of the distances between shape and color based index vectors ($\Phi_{\lambda,\tau}(I)$ and $\Phi_{\lambda,\tau}(J)$) separately

$$\begin{aligned} D(I,J) &= \left\| \Phi_{\lambda,\tau}(I) - \Phi_{\lambda,\tau}(J) \right\| \\ &= w_c \left\| \Phi^c_{\lambda,\tau}(I) - \Phi^c_{\lambda,\tau}(J) \right\| \\ &\quad + w_s \left\| \Phi^c_{\lambda,\tau}(I) - \Phi^c_{\lambda,\tau}(J) \right\| \end{aligned} \tag{10}$$

where $w_c + w_s = 1$.

Shape-based only query is performed by setting w_s to 1, similarly color-based only query is performed by setting w_c to 1.

4. EXPERIMENTAL RESULTS

The retrieval performance of the proposed method is assessed on a database with 82 color image pairs. These images are captured at different distances, from different views, or at different times. The images in the database are indexed for several (λ,τ) values and index vectors are stored in the database (Figure-1). The developed interface allows a user to make color or shape based queries (Figure-2).

Each of the images in the database is searched through shape-only, color-only and hybrid queries. Query results are sorted and listed with respect to the ratings given by the distance measure. In Fig.3 and Fig.4 shape and color-based query results are given. Each image has at least one counterpart image in the database. For each image in the database the ground truth query result is obtained. Only three leading images in the query result are encountered for performance evaluation. If any pair of the queried image is found in the query result list, a corresponding scoring value is assigned to it. This scoring value is 4 if the actual pair is the first image in the query result, 2 if the actual pair is the second image in the query result and finally, 1 if the actual pair is the third image in the query result. Performance is computed as the ratio between the sum and the maximum score. The results are given in Table-1.

Table.1 Query performance.

Query Type	Shape-only	Hybrid	Color-only
Performance	87%	95%	92%

5. CONCLUSION

In this study, color and shape-based index vectors are derived through utilizing their behavior through scale-space. Developed query system allows users to perform color-only, shape-only and hybrid queries. the proposed method give more powerful and rich representations when compared to the methods that only utilize color-based or shape-based indexes and could index images with similar color or orientation histogram. Moreover the proposed algorithm yields better and consistent results especially for images with similar content.

REFERENCES

[1] Y. Rui, T.S. Huang, S. Chang, "Image Retrieval: Past, Present, and Future," Visual Communication and Image Representation, vol.10, pp. 1-23, 1999.

[2] A. Pentland, R.W. Picard, S. Sclaroff, "Photobook: Tools for content-based manupulation of image database," Storage and Retrieval for Image and Video Database II, vol. 2185, pp. 34-47, 1994.

[3] B. S. Manjunath, W. Y. Ma, "Texture features for browsing and retrieval of image data," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18(8), pp. 837-842, 1996.

[4] J. Huang, S.R. Kumar, M. Mitra, W.Zhu, R. Zabih, "Image Indexing Using Color Correlograms," Proceeding of IEEE Conference on Computer Vision and Pattern Recognition, pp. 762-768, San Juan, Puerto Rico, June 1997.

[5] G. Pass, R. Zabih, "Histogram Refinement for Content-Based Image Retrieval," 3rd IEEE Workshop on Applications of Computer Vision (WACV'96), pp. 96-103, 1996.

[6] Anil K. Jain, A. Vailaya, "Image Retrieval using Color and Shape," Pattern Recognition Journal, vol.29, pp. 1233-1244, 1996.

[7] G. Iannizzotto, L. Vita, "A Multiscale Turning Angle Representation of Object Shapes for Image Retrieval," Visual 99, pp. 609-616, The Netherlands, June 1999.

[8] B. Kurt, M. Gökmen, "*Two Dimensional Generalized Edge Detector*," IEEE Proceedings of Int. Conf. on Image Analysis and Proc., ICIAP'99, pp. 302-307, Italy.

[9] B. Kurt, M. Gökmen, "*Yinelemeli ve Uyarlanır Ayrıt Saptayıcı Süzgeçleri*," Conf. on Signal Processing and its App., SIU'2001, pp. 410-415, Northern Cyprus, 2001.

[10] A.P. Witkin, "Scale-space filtering," International Joint Conference on Artificial Intelligence, Vol.2, pp. 1019-1022, August, 1983.

[11] B. Kurt, M. Gökmen, "*İmge ve Videoların İndeksleme Amaçlı Ölçek Uzayı Analizi*," Conf. on Signal Processing and its App., SIU'2002, pp. 335-340, Pamukkale, 2002.

[12] M. Orchard, C. Bouman, "Color Quantization of Images," IEEE Trans. On Signal Processing, vol. 39, no.12, pp.2677-2690, 1991.

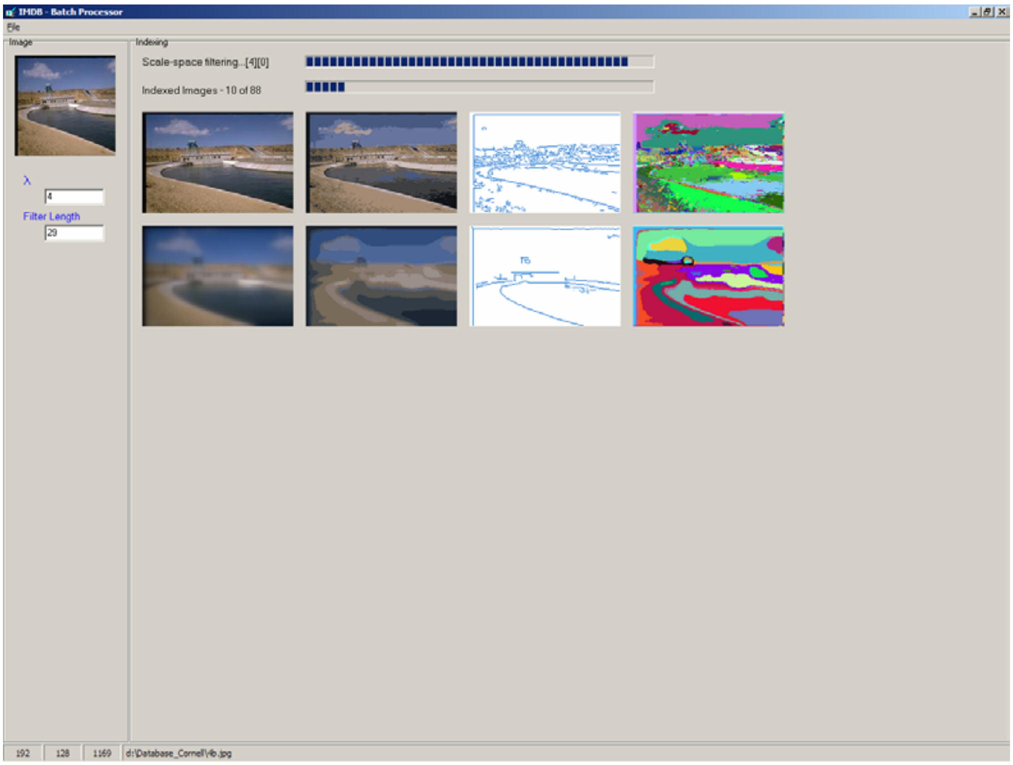


Figure-1 Image indexing system.

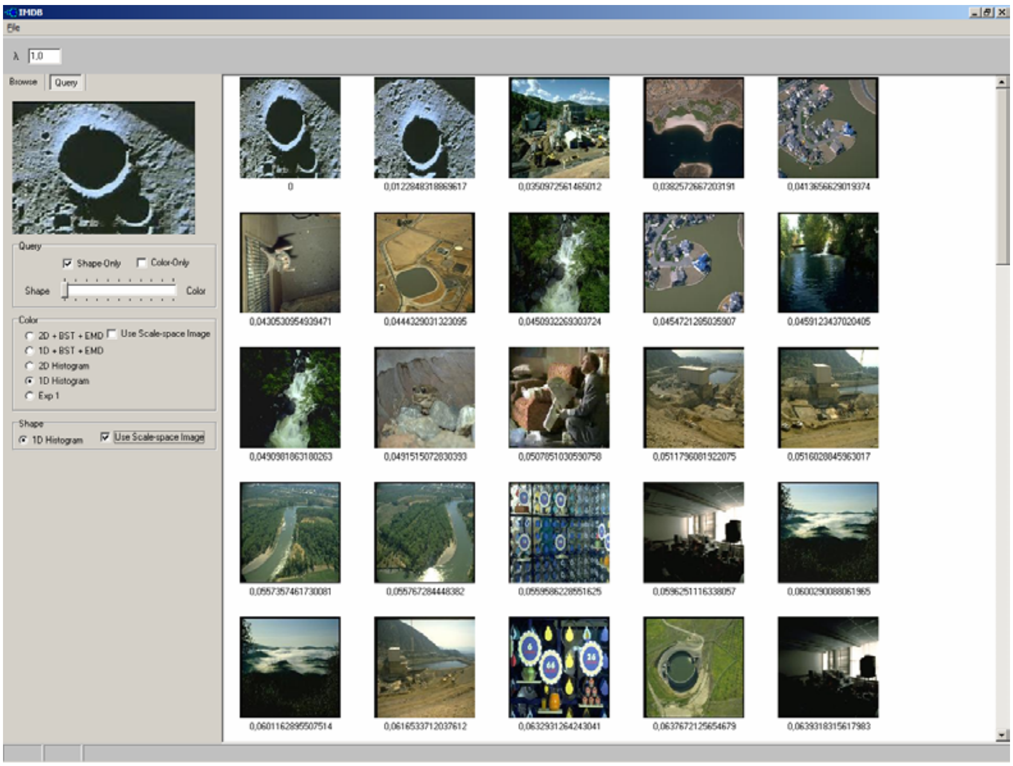
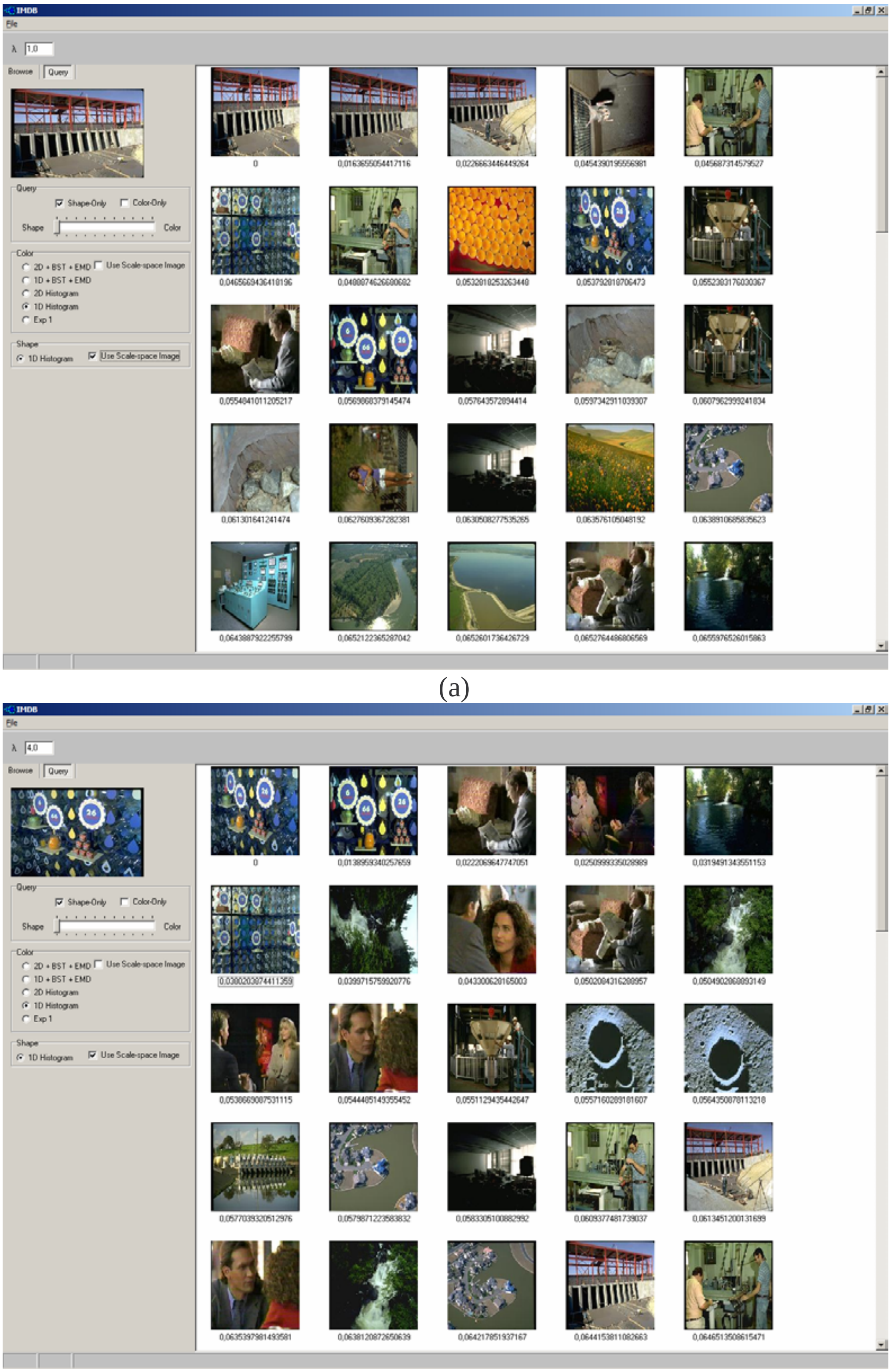
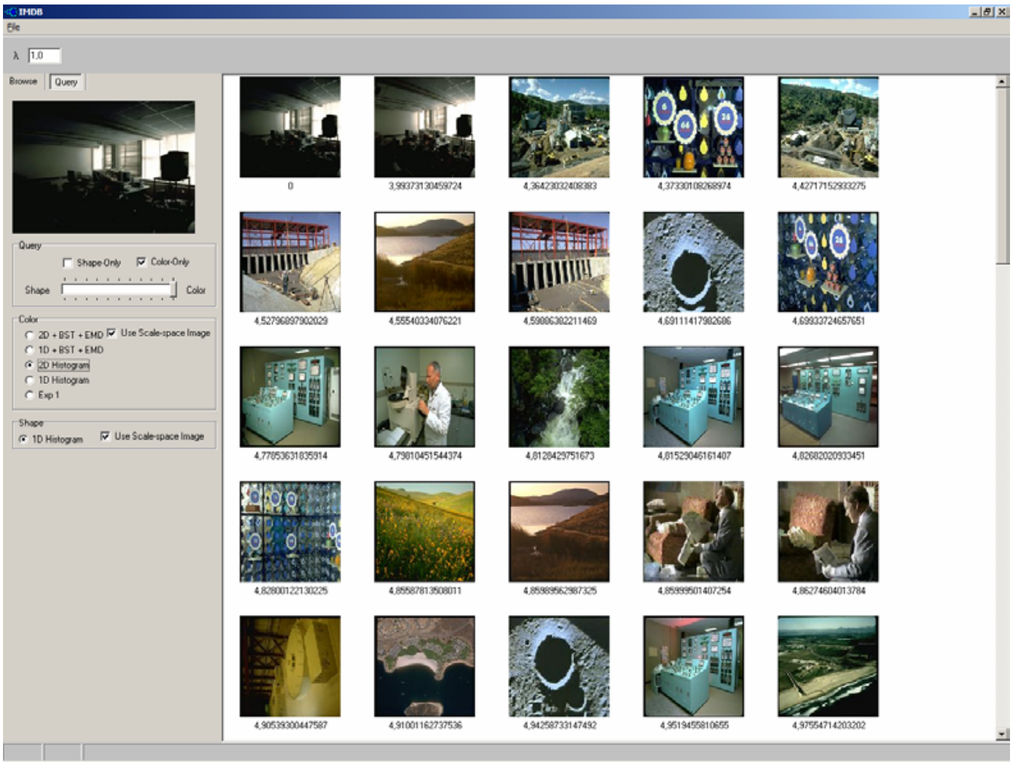


Figure-2 Query interface.

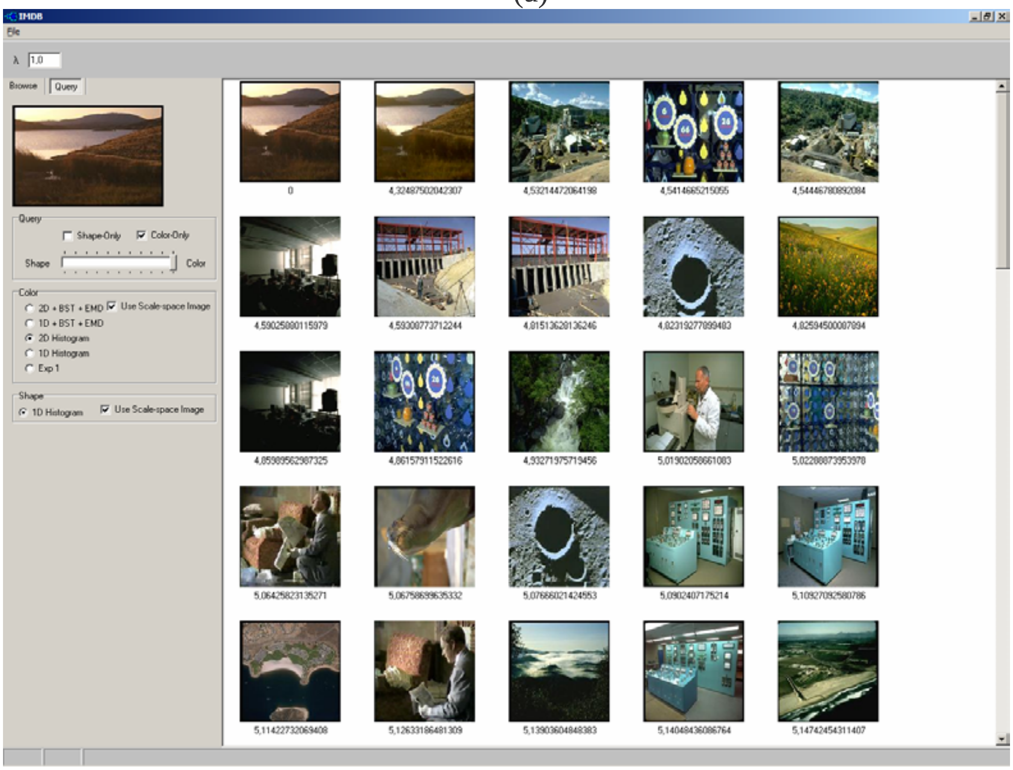


(b)

Figure-3 Shape-only query.



(a)



(b)

Figure-4 Color-only query.