A Method Using Morphology and Histogram for Object-based Retrieval in Image and Video Databases

Thi Thi Zin and Hiromitsu Hama

Graduate School of Engineering, Osaka City University, Osaka, 558-8585 Japan

Summary
We propose a method for object-based retrieval in image and video databases. In doing so, both color and shape information is taken into account to enrich the querying capabilities for object-based retrieval. The morphological operations are performed on each uniform color region. We use three kinds of histograms: color histogram, distance histogram, and angle histogram. It is shown that the proposed approach is invariant to scaling, rotation, and translation. Finally, performance experiments are carried out by the evaluation of the effectiveness and the robustness of the proposed method.

Key words:
Object-based retrieval, Uniform color region, Image and video databases, Morphology, Color histogram, Distance histogram, Angle histogram

1. Introduction

Advances in multimedia technology have accelerated the amount of digitized information along with the data stored for image and video contents making it difficult to classify, organize, and control in a normal ways. Consequently, there is a dramatic appeal utilizing for efficient and user-friendly multimedia retrieval and indexing techniques.

To deal with a large amount of content we have to provide solutions which enable users to retrieve, edit, and watch selected content. As far as the querying process is concerned in multimedia databases, a set of feature descriptors to represent the visual content is generally taken into consideration instead of the content itself of the image and video. Examples are to recognize the objects of an image and tag the results as metadata or track a target automatically. These techniques are useful for content-based retrieval intended to query the color content of image and video data as a whole. Shape is also another feature that is used for querying the content of image and video.

Recently, many researches have focused on region-based techniques that allow the user to specify a particular region of an image and request that the system retrieve images that contain similar regions. Among many methods for extracting an object, template matching and machine learning are very popular methods. In case of template matching, a pre-defined template of target object is compared with an unknown image. In machine learning, support vector machine [1] and boosting [2] have attracted to significant number of researchers. These methods assume a model of the target and use it for extracting the region of object. So, limitations such as requirements of fixed environment and well defined target are occurred.

There have been proposed many object retrieval algorithms however these are applicable only in a closed domain. To name a few, a method [3] to extract the region including the attention objects based on physical image features and a tracking algorithm [4] to exploit the location data in video databases have been considered. But many open problems are left for the robust retrieval.

In this paper, we propose a new method for object retrieval in image and video databases. Our research focuses on object-based image retrieval, in which it is assumed that every target object is composed of Uniform Color Regions (UCRs). It retrieves images from a database based on the appearance of physical objects. The user designates an object by providing an example to the system, and after that it retrieves all images that contain the query object. Considering the fact that querying by low-level object features is essential in image and video data, an efficient approach for querying and retrieval by shape and color is proposed. After query image is presented, we use morphology and histogram-based approach to enrich the querying capabilities. UCRs of target objects can be extracted using color histograms. For object-based shape queries, morphological dilation is applied to the image. The shape features of the objects are characterized by distance and angle histograms. Then, the integration of color and shape information makes the better retrieval accuracy. Fig.1 shows the overview of our algorithm. In our system, the query and input images are captured from various input devices such as digital camera, mobile phone camera, video camera, TV, online images, photo collection CDs, and so on. This approach is not only invariant to scaling, rotation, and translation, but also valid for low resolution images and also for partial occlusion. The retrieval effectiveness of our proposed method is shown through experiments using a comprehensive set.

Manuscript received September 5, 2007.
Manuscript revised September 20, 2007.

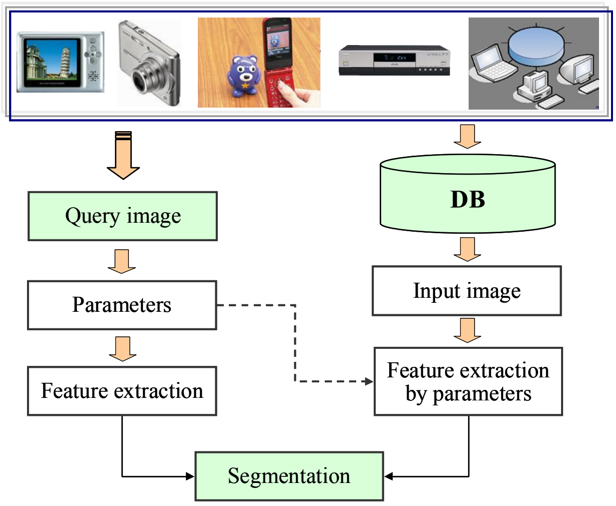


Fig. 1 Overview of our algorithm.

2. Algorithm

The histogram-based approach is adopted to extract color and shape information of images. Both color and shape information is used in most existing systems [5-7] because the combinational usage improves the retrieval effectiveness of the systems. Here, to achieve better retrieval accuracy, we use three kinds of histograms: color, distance, and angle histograms are applied in our system. Furthermore, morphological dilation using ring-shaped and fan-shaped SEs are also used. The detailed flowchart is shown in Fig.2. In the following, these processes are discussed step by step.

2.1 Color histogram for extracting UCRs

Color is the most frequently used feature in querying and retrieval of multimedia data [8]. Not only the images as a whole, but also the color regions that exist in the image, can be queried. Since there may exist more than one UCR in an image, the spatial relations among the UCRs can also be queried. The color histograms [9-11] are used to represent the color distribution in an image or a video frame. The color histogram approach counts the number of occurrences of each UCR on a query image. Since an image is composed of pixels and each pixel has a color, the color histogram of an image can be identified with the size of the areas containing the color as the number of pixels. This method is very simple, easy-to-handle and computationally cheap.

In our approach, color histogram stores Hue-like parameters in *XY* space for all of the pixels belonging to an object. *XY* space is 2D color space and the details and effectiveness of this space has been stated in [12-14]. The aspect of the *XY* space is shown in Fig.3. Each pixel value of an input image is transformed from a point (R, G, B) on the *RGB* space to one point (X, Y) on the *XY* space according to Eq.(1)-(4). The angle (θ) of color on *XY* space is defined by Eq.(5). In Fig.3(c), the color histogram using query image is described. Based on color histogram, we can extract UCRs from the object using the peak regions of the histogram. The aspect of UCR is explained in details in [15]. Here, we use chromatic color only. Achromatic color is defined as in (6).

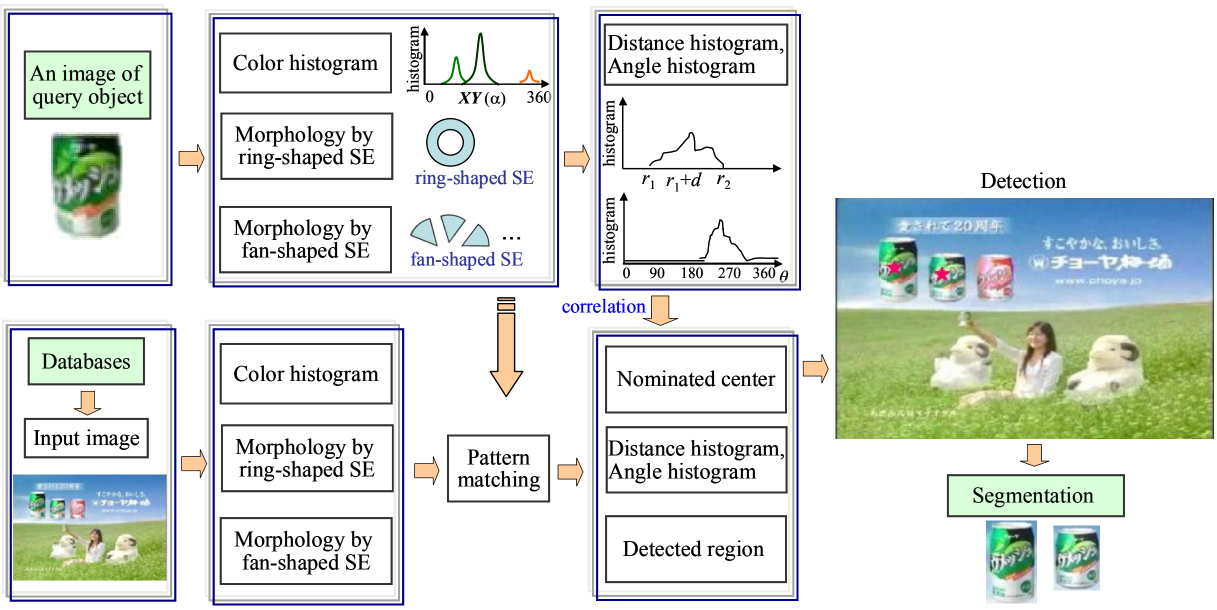


Fig. 2 The detailed flowchart of our proposed system.

$$r = \frac{R}{R+G+B},\ g = \frac{G}{R+G+B},\ b = \frac{B}{R+G+B}. \quad (1)$$

$$\vec{E}_x = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0\right),\ \vec{E}_y = \left(-\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right), \quad (2)$$

$$\vec{a} = (r, g, b) - \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right), \quad (3)$$

$$X = \vec{a} \cdot \vec{E}_x,\ Y = \vec{a} \cdot \vec{E}_y, \quad (4)$$

where $\vec{E}_x$ and $\vec{E}_y$ are perpendicular to each other and unit vectors.

$$\theta = \tan^{-1}\left(\frac{Y}{X}\right). \quad (5)$$

Achromatic region is divided into three categories according to the following:

$$\begin{aligned} &Black: R+G+B \le 25, \\ &White: R+G+B \ge 600, \\ &Gray: 25 < R+G+B < 600 \text{ and} \\ &\quad (r \ge 0.29, g \ge 0.29, b \ge 0.29). \end{aligned} \quad (6)$$

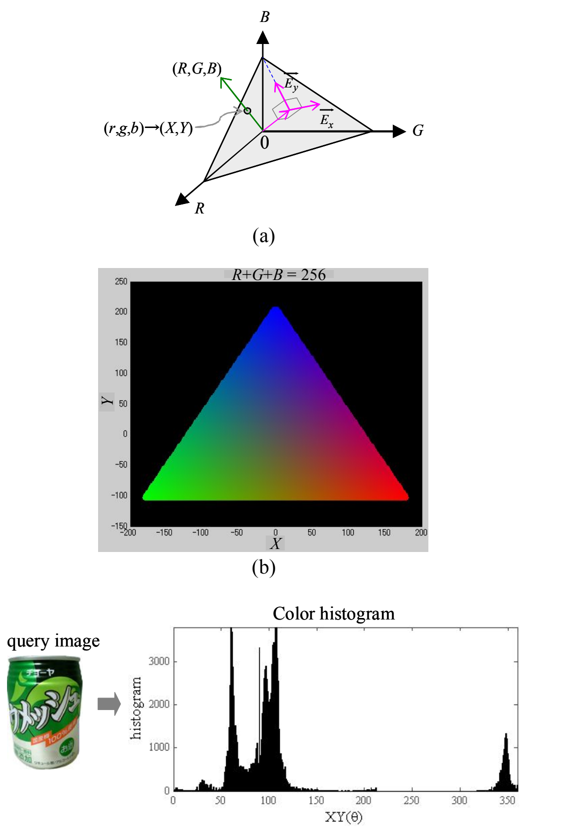


Fig. 3 XY space: (a) relation between *RGB* and *XY* space, (b) *XY* space with colors on *R*+*G*+*B* = 256, and (c) color histogram.

2.2 Morphology for detecting nominated regions

We use morphological dilation for two purposes: i) finding the initial nominated regions for some nominated object's centers (OCs) in the input image and ii) deciding the correct OC and getting the rotation angle of the matched region in the input image with the query object. For the former purpose, we use ring-shaped structuring elements (SEs) on each UCR. Each SE has two radii got from the query object, r_1 and r_2 for the inner ring and outer ring, respectively. This aspect is shown in Fig.4(a). For the latter, we use fan-shaped SEs using the radius R_2 as shown in Fig.4(b). They are successively 45 degree rotated and equal in shape. By using fan-shaped SEs and defining $\boldsymbol{B}_i^{(k)} = Dil(f_i, UCR_k, I)$, for $i = 1, \ldots, 8$, as the binary image of dilation of k^{th} UCR of a query image $\boldsymbol{Q}$ by i^{th} fan-shaped SE f_i, we can obtain the rotation information as follows.

$$\boldsymbol{B}^{(k)} = (\boldsymbol{B}_1^{(k)}, \cdots, \boldsymbol{B}_8^{(k)}), \quad (7)$$

$$\boldsymbol{C}_l^{(k)} \equiv (\boldsymbol{B}_{1\oplus l}^{(k)}, \cdots, \boldsymbol{B}_{8\oplus l}^{(k)}),\ l = 0, 1, \cdots, 7, \quad (8)$$

where $\oplus$ means (sum of mod 8)+1, for example, $\boldsymbol{C}_1^{(k)} = (\boldsymbol{B}_2^{(k)}, \boldsymbol{B}_3^{(k)}, \ldots, \boldsymbol{B}_1^{(k)})$, $\boldsymbol{C}_2^{(k)} = (\boldsymbol{B}_3^{(k)}, \boldsymbol{B}_4^{(k)}, \ldots, \boldsymbol{B}_2^{(k)})$ and so on. Now, $\boldsymbol{B}^{(k)}$ and $\boldsymbol{C}_l^{(k)}$ can be considered as 8 bit-gray images giving the corresponding pixel values at position (x, y) as $\boldsymbol{B}^{(k)}(x, y)$ and $\boldsymbol{C}_l^{(k)}(x, y)$, respectively. Consequently, if we denote the position OC by (x_0, y_0), then the pixel values of query and input images at OC can be expressed as $\boldsymbol{B}^{(k)}(x_0, y_0)$ and $\boldsymbol{I}^{(k)}(x_0, y_0)$, respectively. The set $\boldsymbol{N}$ of nominators of OCs is obtained by:

$$\boldsymbol{N} = \{(x, y) \mid \text{HD}(\boldsymbol{I}^{(k)}(x, y), \boldsymbol{B}^{(k)}(x_0, y_0)) \text{ is equal to zero or one}\}$$

where HD($\boldsymbol{X}_1$, $\boldsymbol{X}_2$) means Hamming distance between $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, for example, HD((0000), (0001))=1, HD((0011), (0110))=2 and so on. If $\boldsymbol{N}$ is not empty, there may be at least one object similar to the query image inside the input image region.

If an input image $\boldsymbol{I}$ is the 45 degree rotated one of the query image $\boldsymbol{Q}$, then HD($\boldsymbol{I}^{(k)}(x_0, y_0)$, $\boldsymbol{C}_0^{(k)}(x_0, y_0)$) is equal to zero. Generally saying, HD($\boldsymbol{I}^{(k)}(x_0, y_0)$, $\boldsymbol{C}_l^{(k)}(x_0, y_0)$), $(l \neq 0)$ is greater than zero. Thus we can know roughly rotation angle. The precise angle may be calculated using angle histogram as described in **2.4**. Since $\boldsymbol{B}_i^{(k)}$ is obtained by fan-shaped SEs which are 45*n (n: integer) degree rotation invariant, so we define $\boldsymbol{N}_0$ and $\boldsymbol{N}_1$ newly as follows:

$$\boldsymbol{N}_0 = \{(x, y) \mid \text{HD}(\boldsymbol{I}^{(k)}(x, y), \boldsymbol{C}_l^{(k)}(x_0, y_0)) = 0\} \quad (9)$$

$$\boldsymbol{N}_1 = \{(x, y) \mid \text{HD}(\boldsymbol{I}^{(k)}(x, y), \boldsymbol{C}_l^{(k)}(x_0, y_0)) = 1\} \quad (10)$$

Nominators for OCs are decided according to the following three cases:

Case 1: If $\boldsymbol{N}_0$ is not empty, then pixels in $\boldsymbol{N}_0$ are nominated as OCs.

Case 2: If $\boldsymbol{N}_0$ is empty and $\boldsymbol{N}_1$ is not empty, then pixels in $\boldsymbol{N}_1$ are nominated as OCs.

Case 3: If $\boldsymbol{N}_0$ and $\boldsymbol{N}_1$ are empty, then there is no object similar to the query image inside the input image region, that is, no nominator.

If there is a pixel (x, y) in the nominators such as HD($\boldsymbol{I}^{(k)}(x, y)$, $\boldsymbol{C}_l^{(k)}(x_0, y_0)$) is equal to zero or one, it means possibility that there is an object similar to the $45*l$ degree rotated one of the query image.

2.3 Distance histogram for determining scaling factor

In order to construct a distance histogram, we first denote the minimum and maximum distances between the OC and the pixels within the object by r_1 and r_2, respectively. Then, the interval $|r_1$-$r_2|$ is divided into class intervals of equal length (say d). For each class interval, we draw circles with centre at OC and radius of length r_1, r_1+d, r_1+2d,..., r_2. Finally, we can obtain a distance histogram by using number of pixels of the object located in the annulus form by the circles of radians r_1, r_1+d, r_1+2d, and so on. Alternatively, the distance histogram can be obtained by using lengths of circumscribed rectangles of the corresponding objects located in the annulus form by the circle of radians r_1+d, r_1+2d, and so on. In Fig.5(a), the distance histogram for a particular object is given as an illustration. Here, the correlation between the distance histograms of the nominated region in the input image and the query object is abbreviated as Cor(Dist). By using Cor(Dist), we can obtain the scaling factor in size.

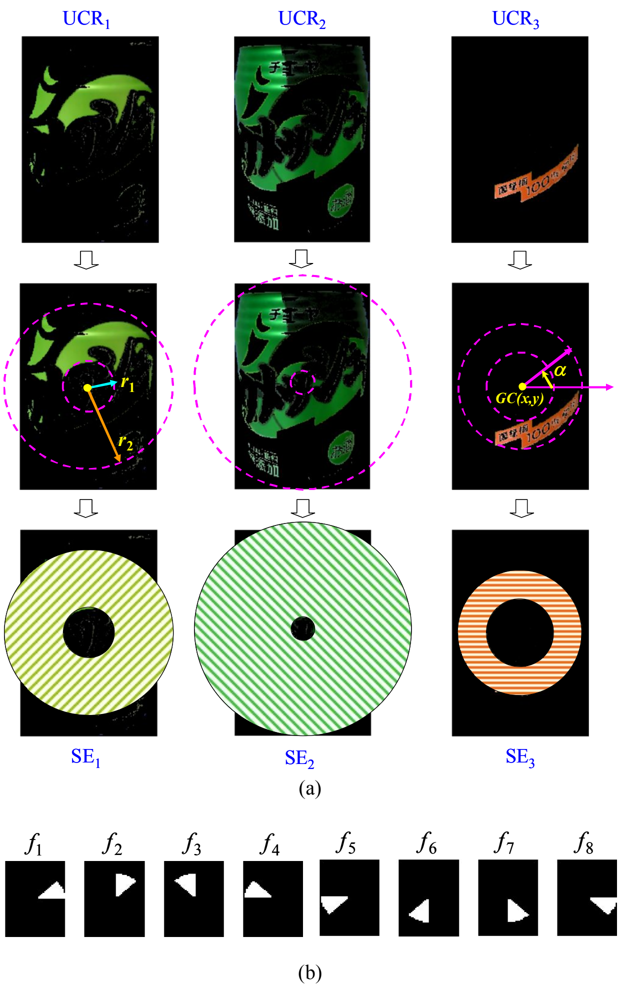


Fig. 4 SEs for morphological operations: (a) ring-shaped SEs and (b) fan-shaped SEs.

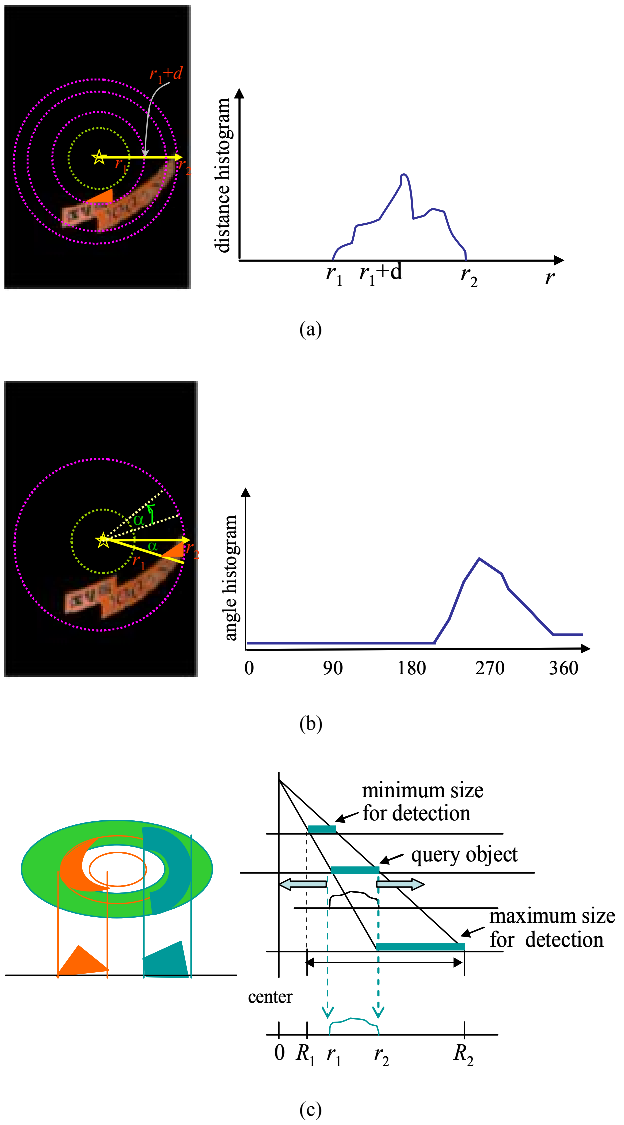


Fig. 5 (a) distance histogram, (b) angle histogram, and (c) scaling in size.

2.4 Angle histogram for obtaining the rotation angle

In this section, we will consider angle histogram. An angle histogram can be obtained by using the angular distribution of the pixels of the object image around OC. Specifically we divide 360° into sectors of α° each. The number of pixels of object located in each sector will be counted down and used as frequency against angles to obtain angle histogram. In Fig.5(b), we illustrate angle histogram of the same object which is used for the distance histogram. By using the correlation of angle histograms as Cor(Ang), we can obtain the angle of rotation for the input object corresponding to the query object.

2.5 Matching process

Pattern matching is performed on the results by morphological dilation to detect OC and rotation angle roughly. In this stage, we can also obtain the estimated scaling factor as shown in Fig.5(c). Using Cor(Dist) and Cor(Ang), we can detect the object region with more accurate rotation angle. Finally, the corresponding region to the query object can be segmented from the input image.

3. Experimental Results

To evaluate the performance of our proposed method, we used various query objects composed of UCRs and 100 input images containing various different objects of interest such as commercial goods, visual signboards, and so on. One example of query image and input image is shown in Fig.6. According to color histogram and morphological operation, we obtained the nominated OCs. So, the query object's regions are detected and then segmented, as shown in Fig.6(a). Using Cor(Dist) of query object and each segmented region, the scaling factors of each region are obtained. In Fig.6(b), the scaling factors of region 1 and region 2 are 1 and 0.8, respectively. This means region 2 is a little bit small. Similarly, we can get the angular position of these segmented regions from Cor(Ang).

Some experimental results are shown in Fig.7. It was seen that the regions got by our method are nearly equal to the manually segmented regions, regardless of wide variety of scale, angle and backgrounds. In this paper, the query object's size is 60*40 pixels and the scaling factor is limited to 0.5 to 2. But it can be expanded easily. In addition, this approach can work well even though the objects under very low resolution and various conditions including partial occlusion and reflection. Moreover, the approach is scalable in the sense that the images having more than one object can be processed without any difficulty. Furthermore, our method allowed the location of object is independent within the whole image. According to our experimental results, the recall rate was 100% based on 100 input images. For the future, we need to extend the huge amount of images in databases.

4. Conclusion

In this work, we have proposed a method for image-based object retrieval by shape and color features. Experiments showed that our approach is invariant under rotation, translation, and scaling. Moreover, we have confirmed the robustness of our method using images with very low resolution, partial occlusion, and various lighting conditions. The possible applications are video searching, net-window shopping, image keyword searching, online search for sport-team logo and commercial goods on WWW. Moreover, our method is applicable to image zooming of the relating region with the query object.

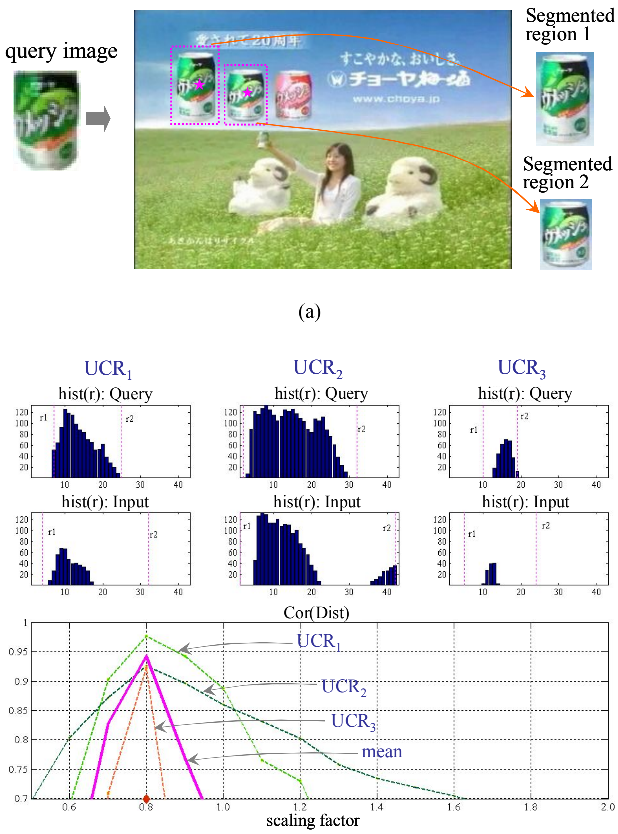
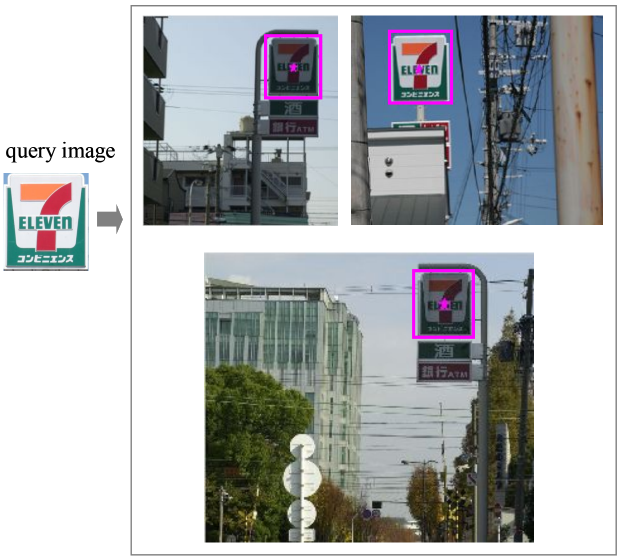


Fig. 6 One example of query image: (a) segmented regions and (b) correlation between distance histograms for segmented region 2.

(a)



(b)

Fig. 7 Experimental results.

Acknowledgment

The research work was supported in part by Research Fellowships of the Japan Society for the Promotion of Science. This work was also supported in part by KAKENHI 19500154 Grand-in-Aid for Scientific Research (C).

References

[1] C.A. Waring and X.Liu, "Face detection using spectral histograms and SVMs", *IEEE Trans. on SMC (B)*, vol. 35, no. 3, pp. 467-476, 2005.

[2] S. Lao and M. Kawade, "Vision-Based face Understanding Technologies and Their Applications", *5th Chinese Conf. on Bimoteric recognition*, SINOBIOMETRICS2004, pp. 339-348, 2004.

[3] M. Kimura, M.Yamauchi, "A method for extracting region of interest based on attractiveness and study for CE applications", in *Digest of Tech. Papers of the Int. Conf. on Consumer Electronics*, Las Vegas, NV, USA, Jan. 2006, pp. 343-344.

[4] A.H.R. Albers, R.G.J. Wijnhoven, E.G.T. Jaspers, P.H.N. de With, "Smart search & retrieval on video databases", in *Digest of Tech. Papers of the Int. Conf. on Consumer Electronics*, Las Vegas, NV, USA, Jan. 2006, pp. 475-476.

[5] Ediz Şaykol, Uĝur Güdükbay and Özgür Ulusoy, "A histogram-based approach for object-based query-by-shape-and-color in image and video databases", *Image and Vision Computing*, vol. 23, no. 13, pp. 1170-1180, Nov 2005.

[6] E.L. Van den Broek, E.M.Van Rikxoort, and T.E. Schouten, "Human-Centered Object-Based Image Retrieval", *Lecture Notes in Computer Science*, vol. 3687, pp. 492-501, 2005.

[7] Wang Xiangyang, Hu Fengli, Yang Hongying, "A Novel Regions-of-Interest Based Image Retrieval Using Multiple Features", *12th International Multimedia Modeling Conf.*, Beijing, China , pp. 377-380, Jan 2006.

[8] Xiuqi Li, Shu-Ching Chen, Mei-Ling Shyu, Sheng-Tun Li, and Borko Furht, "A Novel Hierarchical Approach to Image Retrieval Using Color and Spatial Information", *Lecture Notes In Computer Science*, Springer-Verlag, vol. 2532, pp. 175 - 182, 2002.

[9] M. Swain and D. Ballard, "Color indexing", International Journal of Computer Vision, vol. 7, no. 1, pp. 11-32, 1991.

[10] J. Hafner, H. Sawhney, W. Equitz, M. Flickner, W. Niblack, "Efficient color histogram indexing for quadric form distance functions", *IEEE Trans. on PAMI*, vol. 17, no. 7, pp. 729-736, 1995.

[11] Sangoh Jeong, Chee Sun Won, and Robert M. Gray, "Image Retrieval Using Color Histograms Generated by Gauss Mixture Vector Quantization", *Computer Vision and Image Understanding*, vol. 94 , no. 1-3, pp. 44-66, 2004.

[12] Thi Thi Zin, Sung Shik Koh, and Hiromitsu Hama, "Optimal Color Space for Relative Color Polygons", *IEICE ELEX*, vol. 4, no. 3, pp. 106-113, Feb. 2007.

[13] Thi Thi Zin, Sung Shik Koh, and Hiromitsu Hama, "Relative Color Polygons for Object Detection and Recognition", Lecture Notes in Artificial Intelligence in LNCS, Springer- Verlag, vol. 4259, pp. 834-843, Nov. 2006.

[14] Thi Thi Zin, Sung Shik Koh, and Hiromitsu Hama, "Two Dimensional Color Space and Color Shifting", *Memoirs of the Faculty of Engineering*, *Osaka City University*, vol. 47, pp. 1-8, Dec. 2006.

[15] Thi Thi Zin, Sung Shik Koh, and Hiromitsu Hama, "Main Peak Finding for Signboard Recognition under Adverse Conditions", *International Journal of Computer Science and Network Security*, vol. 6, no. 12, pp. 28-34.

Thi Thi Zin is a Postdoctoral Research Fellow of Japan Society for the Promotion of Science, in the Graduate School of Engineering, Osaka City University, Japan. Her research interests include mathematical morphology, color image processing, and ITS. She is a student member of IEEE.

Hiromitsu Hama received the Dr. Eng. degrees in Electrical Engineering from Osaka University, Japan. He is currently Dean of Graduate School of Engineering at Osaka City University. His research interests include ITS (Intelligent Transportation Systems), 3D world reconstruction, computer vision, picture processing/understanding. He is a member of IEEE, SPIE, IEICE, IIITE, ASJ, J-FACE and ISCIE.