# PARTIAL IMAGE RETRIEVAL USING COLOR REGIONS AND SPATIAL RELATIONSHIPS


Yunyun Cao*, Yoshitomo Yaginuma**, Masao Sakauchi*

*Sakauchi Lab, the 3rd Department
*Institute of Industrial Science*
*University of Tokyo*
*4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan*
{cao, sakauchi}@sak.iis.u-tokyo.ac.jp

***National Institute of Multimedia Education*
*2-12 Wakaba, Nihama-ku*
*261-0014, Chiba, Japan*
yaginuma@nime.ac.jp



**Abstract** In this paper, we present a novel approach to retrieve images that contain the query image as a part regardless of the size and position the query image appears. Images in Database are segmented in advance, for each major region obtained, a composite measurement of color, area percentage and position are stored as the feature. While retrieving, the query image is also first segmented, and then the major regions' colors, their area ratios and spatial relationships are generated for narrowing the searching space. The utilization of multiple regions alleviates the influence of inaccurate segmentation, the pre-segmentation of images in database allows indexing of features of color regions and enables the fast retrieving. The experiment shows the advantages and weakness of the proposed method.



**Keywords**: Partial image retrieval, color region, segmentation, spatial relationship


## 1. INTRODUCTION

In recent years, with the dramatic improvements in computer technology and the advent of World-Wide Web, there has been a rapid increase of the


X. Zhou et al. (eds.), *Visual and Multimedia Information Management*

size of digital image collections. However, we can not access to or make use of the information unless it is organized so as to allow efficient browsing, searching and retrieval. Many researchers and institutions are currently involved in providing tools and methods to efficiently manage pictorial digital libraries.

Up to now, a great deal of effort has been made on global measurement of image similarity (Cinque et al., 1999) (Flickner et al., 1995) (Yi et al., 1999) (Smith et al., 1999) (Shih et al., 2000) (Smith et al., 1997) (Li et al., 2000), some of them deal with color regions. However, it is not conducive to building systems that retrieve images which are semantically related to a given query image. Semantic image similarity, which is what the Content-Based Image Retrieval (CBIR) users expect and demand, often follows only from partial match between images as shown in Figure 1, in which the two images are semantically relative since they contain the same announcer.

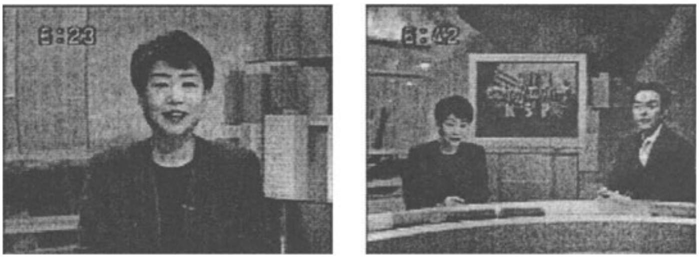


*Figure 1* Semantically related images

What seems to be lacking, however, is the research on partial image retrieval. Only in recently few years, some attentions are paid on it (Cohen, 1999) (Kim et al., 2000) (Moghaddam et al., 1999). However, the current research status is still far away from perfect. Most of them focus on the local feature defining and matching, neglecting the problem of "interactive-time" which is very important while the volume of image database is large.

In this paper, we will present a novel approach for partial image retrieval, which pre-segment the images in database, allowing the indexing of features of color regions and enabling the fast retrieval for query image by utilizing a combination of multiple color regions to calculate the similarity.

The rest of this paper is organized as follows: Section 2 introduces the retrieval method in detail, including the image segmentation, the features stored in database and used for retrieval, and the searching process. In Section 3 experimental results are reported, and advantages and weakness of the proposed method are discussed. Finally Section 4 presents conclusions and our future work.

## 2. PROPOSED METHOD

In this section, we propose a novel partial image retrieval method which utilizes color regions' color, area percentage and their spatial relationships for retrieval. In this method, all images are segmented into regions in advance. For each region, certain features are extracted and indexed for later retrieval. While given a query image, the query image is also segmented first, and the corresponding features of major regions are extracted and combined for local measurements similarity calculation (Figure 2).

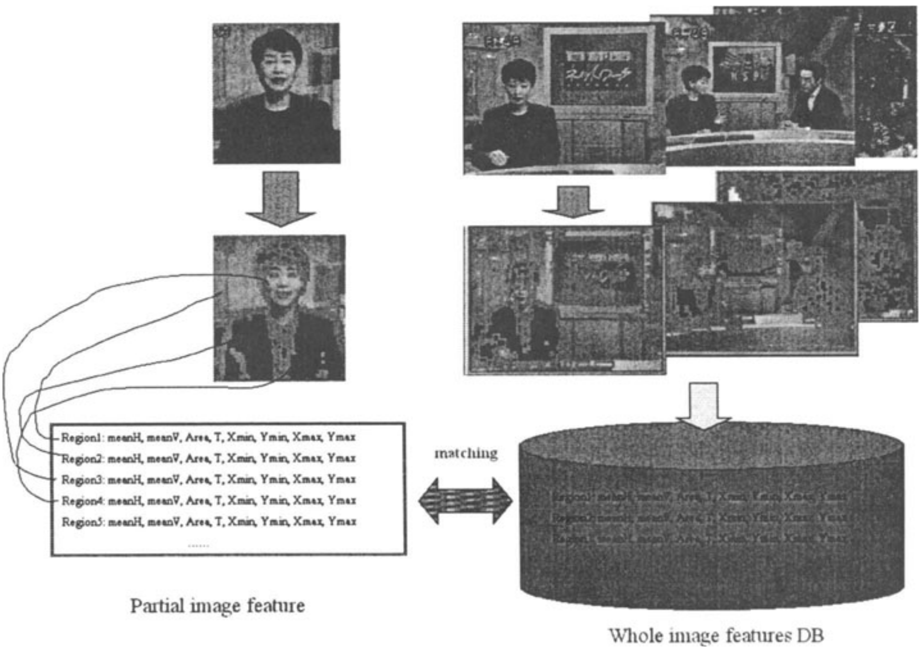


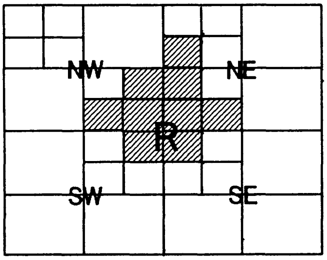

*Figure 2* Partial image retrieval with color regions and spatial relationships

## 2.1 Image segmentation

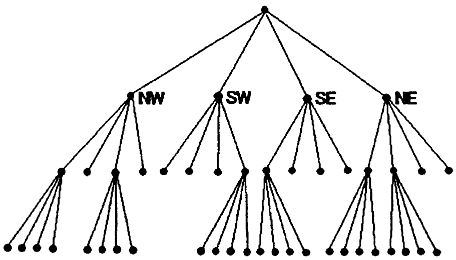
An improved split_and_merge method is developed for image segmentation, which can segment an image into regions with certain colors and texture patterns.

**Split**

The image is at first split into four parts recursively --- NW, NE, SW, SE parts (Figure 3(a)).




(a) Example of splitting




(b) Structure of the quad tree

*Figure 3* Image segmentation with split and merge method

Let $mean_R$ be the average color of part R, $\sigma_R$ be the deviation. The splitting of part R would be stopped in any of the three cases:

1. $Area_R < Threshold_{area}$,

   Which means that when the area of part R is smaller than a certain value, the splitting should be stopped. Since the retrieval which we will discuss later utilizes multiple regions which can reduce the influence of inaccurate segmentation, an appropriate value for $Area_R$ can be taken with a trade-off between the segment accuracy and cost.

2. $\sigma_R < Threshold_\sigma$,

   Which means that if the color in part **R** is almost the same, then this part does not need to be split further.

3. For two consecutive levels of splitting,

$$\sigma_R \approx \sigma_{NW} \approx \sigma_{NE} \approx \sigma_{SE} \approx \sigma_{SW} \text{ and}$$

$$mean_R \approx mean_{NW} \approx mean_{NE} \approx mean_{SE} \approx mean_{SW}$$

Which means that the part **R** may be a region which has a certain color or texture pattern. With this condition, a texture region would not be split into small ones.

The result of the split is represented as a quad tree (Figure 3 (b)), in which the leaf nodes mean the final parts which do not need to be split further.

**Merge**

After the image has been split into a set of sub-regions, the next work is merging the neighboring ones which have similar colors into a larger region.

Figure 4 illustrates some examples of segmentation results.

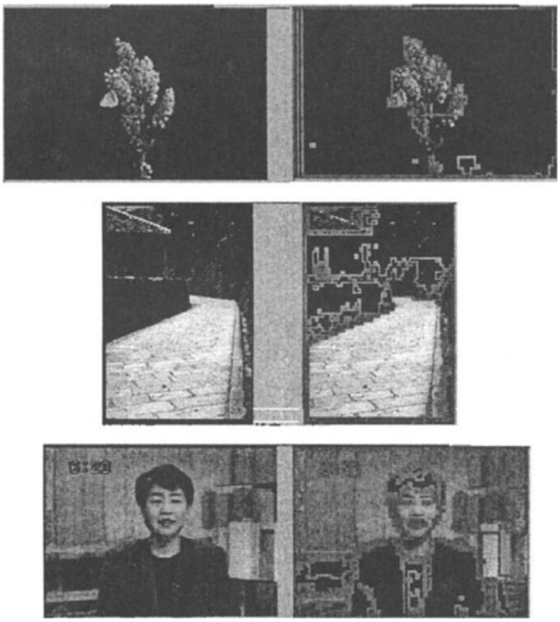


*Figure 4* Examples of image segmentation

## 2.2 Features for retrieval

As we have mentioned at the beginning of Section 2, for each image in Database, the features of its major regions are extracted and stored for later retrieval. Considering the retrieval requirements as well as the costs, we set the number of major regions as 30. For each region, a composite feature of color, area percentage, and position is extracted. The feature vector is described as:

$(imgID, meanH, meanV, meanC, Area, T, Xmin, Ymin, Xmax, Ymax)$

Where *imgID* denotes the ID number of the image which contains the region.
*meanH* denotes the mean value of Hue of this region.
*meanV* denotes the mean value of Value of this region.
*meanC* denotes the mean value of Chroma of this region.
*Area* indicates the percentage of this region's area to the whole image.
*T* denotes this image's position on time axis. It is used in case the image is a frame in a video/image sequence.
*Xmin, Ymin, Xmax, Ymax* indicate the coordinates of the minimum bounding rectangle of this region relative to the top-left vertex of the image, for the whole image they are (0, 0, 1, 1).

For retrieval, the query image is also segmented. However, only main regions (according to the area size) are utilized for retrieval. Suppose that we use the top N regions, the features of them are described as:

$$(meanH_i, meanV_i, meanC_i, Area_i, T_i, Xmin_i, Ymin_i, Xmax_i, Ymax_i), \quad i = 1, \ldots, N$$

Where $Area_i >= Area_{i+1}$

For the query image, the features used for retrieval are described as:

**a. Color feature**

It is described as:

$(meanH_i, meanV_i, meanC_i)$, $i = 1, \ldots, N$

The Euclidean distance is utilized as its distance measurement $D_c$.

**b. Area ratio**

It is obviously that even the size of an object changed, the ratio of the corresponding regions' areas is almost the same as it appears in the original image. Therefore the ratio of region areas can be used as a feature, which is invariant to image size. Since we have obtained a feature element *Area,* which means percentage of a segmented region's area to the area of the whole image, The feature **Area ratio** is described as:

$$\left(\frac{Area_1}{Area_2}, \frac{Area_2}{Area_3}, \ldots, \frac{Area_i}{Area_{i+1}}, \ldots, \frac{Area_{N-1}}{Area_N}\right)$$

**c. Spatial relationships of the major segmented regions**

Some researchers have worked on retrieving images on spatial relationships (Kim et al., 1999) (Venkat, 1998). We considered and briefly investigated various techniques for spatial representation and matching, including the elastic spring models and graph matching. But in the end, considering the spatial information available in our work --- the minimal bounding rectangle of segmented regions, we opted for a much simpler formulation based on the consistency of arrangement of vertex in x and y axis, as illustrated in Figure 5.

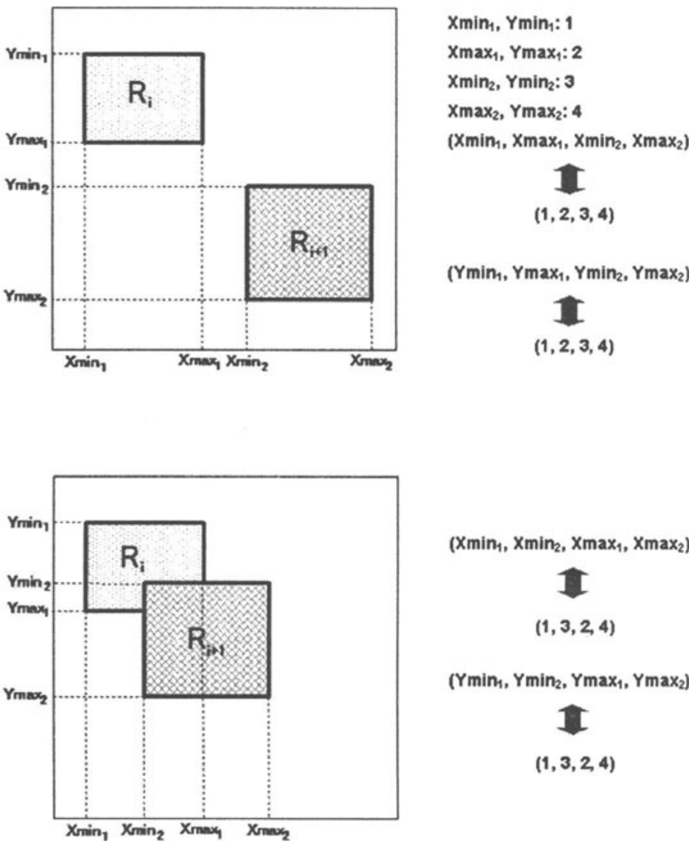



*Figure 5* Representation of spatial relationship between two regions

Let (Xmin, Ymin), (Xmax, Ymax) be the top-left and bottom-right vertex of the minimal bounding rectangle of a region. For two regions $R_i$ and $R_{i+1}$ as shown in Figure 5, assign labels 1, 2, 3, 4 to $Xmin_1$, $Xmax_1$ , $Xmin_2$ , $Xmax_2$ respectively. Sorting the $Xmin_1$, $Xmax_1$ , $Xmin_2$ , $Xmax_2$, the sequence of the corresponding labels becomes the x-axis spatial relationship of $R_i$ and $R_{i+1}$. The y-axis spatial relationships can be obtained similarly.

For example, for Figure 5 (a), the spatial relationship of $R_i$ and $R_{i+1}$ is represented as (1, 2, 3, 4) for x-axis, (1, 2, 3, 4) for y-axis, and for Figure 5 (b), it is represented as (1, 3, 2, 4) for x-axis, (1, 3, 2, 4) for y-axis.

Let $(l_{x1}, l_{x2}, l_{x3}, l_{x4})$, $(l_{y1}, l_{y2}, l_{y3}, l_{y4})$ refer to the x-axis, y-axis spatial relationships between a pair of regions, respectively. The distance between two pairs of regions is calculated with the following equation:

$$D_s = \sqrt{\sum_{i=1}^{4}(l_{xi} - l_{xi}')^2 + \sum_{i=1}^{4}(l_{yi} - l_{yi}')^2}$$

For example, the distance between Figure 5 (a) and Figure 5 (b) is:

$$\sqrt{0+1+1+0+0+1+1+0} = 2$$

Therefore, the spatial relationship of the query image can be described as:

$$\{(l_{xi1}, l_{xi2}, l_{xi3}, l_{xi4}\ l_{yi1}, l_{yi2}, l_{yi3}, l_{yi4}) \mid i=1, \ldots, N-1\}$$

For each $i$, $(l_{xi1}, l_{xi2}, l_{xi3}, l_{xi4}\ l_{yi1}, l_{yi2}, l_{yi3}, l_{yi4})$ denotes the spatial relationship between region $R_i$ and $R_{i+1}$.

## 2.3 Retrieval with combination of multiple color regions

Assuming that the query image is segmented into $N$ major regions: $\{R_1, R_2, \ldots, R_N\}$, from these N regions, three kinds of features discussed in Section 2.2 can be obtained.

On the other hand, for each image in the Database, features of top 30 segmented exclusive regions are available. They are represented as:

$$(imgID', meanH'_j, meanV'_j, meanC'_j, Area'_j, T'_j, Xmin'_j, Ymin'_j, Xmax'_j, Ymax'_j),$$
$$j = 1, \ldots, 30$$

Where $Area'_j >= Area'_{j+1}$

An image which contains a region set of $\{R'_1, R'_2, \ldots, R'_N\}$ is considered the final candidate, if it satisfies the following conditions:

**a.** $D_c(R_i, R'_i) < Threshold_c,\ i=1, \ldots, N$

Which means that the colors of corresponding regions should be identical.

**b.** $\left|\frac{Area_i}{Area_{i+1}} - \frac{Area_i}{Area_{i+1}}\right| < Threshold_A,\ i=1, \ldots, N-1$

Which means that the area ratios of the corresponding region pairs should be identical.

c. $D_s((R_i, R_{i+1}), (R'_i, R'_{i+1})) < Threshold_s, \ i=1, \ldots, N-1$

Which means that the spatial relationships of the corresponding region pairs should be identical.

At first, the largest region $R_1$ is searched according to its color feature, a result set of $\{(imgID_i, region_{i1}, D_{ci1})\}$is obtained. then it's the turn of the second largest region $R_2$, it is not necessary to search the whole feature space now. For each $(imgID_i, region_i, D_{ci})$ in the result set, check the 30 regions in image $imgID_i$, if there does not exist a corresponding $R'_2$ which satisfies the three conditions, then delete the $(imgID_i, region_{i1}, D_{ci1})$ from the result set, else the result set becomes $\{(imgID_i, region_{i1}, region_{i2}, D_{ci1}+D_{ci2})\}$. Repeat the process until the last region $R_N$ is processed.

The region-based partial image retrieval is processed serially region by region. At each step, the searching space is narrowed according to the three conditions discussed above.

## 3. EXPERIMENTAL RESULTS

For experiments, we constructed an image database of 11000 images, in which 3300 images with size of 192x128 are from Tesmic System, and 7200 images with size of 162x120 are extracted at the rate of 1 frame per second from two hours videos of NHK news. The remaining 500 images are extracted from a driving shot of urban scene. Although the three types of images have different size and qualities, our method can handle them as the same.

Figure 6 and Figure 7 show two examples of retrieved results.

It can be seen from the figures that the proposed method can retrieve images containing the query object with varying size and locations.

The results show that for an object with special color regions (Figure 6(a)), the method works well, however, for an object with common color regions, e.g., in Figure 7(a), the announcer wears a suit of a very popular color, the approach failed to generate a good precision and recall.

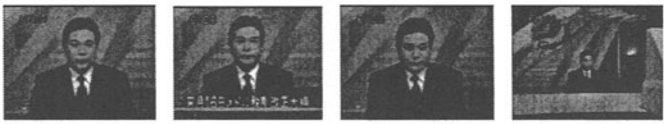
(a) Query image

(b) Examples of true results

(c) Examples of false results

*Figure 6* Results of partial image retrieval (1)

(a)Query image



(b)Examples of true results


(c)Examples of false results

*Figure 7* Results of partial image retrieval (2)

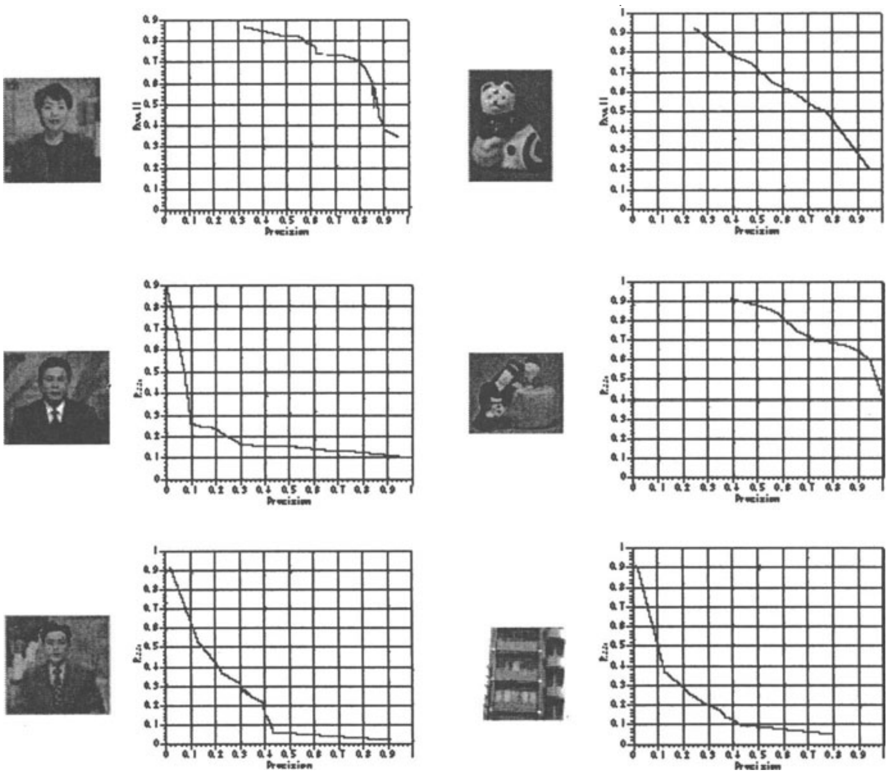


*Figure 8* Precision-recall relations for corresponding objects

Here we use *precision* and *recall* for further evaluation. They are defined as:

$$Precision = \frac{|A \cap B|}{|B|}$$

$$Recall = \frac{|A \cap B|}{|A|}$$

Where $A$ is the set of all *relevant* images, which contains the query image as a part, $B$ is the set of retrieved images resulting from the query.

From the definition we can know that *precision* is the ratio of the number of relevant images retrieved to the total number of images retrieved. Perfect *precision* (100%) means that all retrieved images are relevant. *Recall* is the ratio of the number of relevant images retrieved to the total number of relevant images. Perfect *recall*(100%) can be obtained by retrieving the entire collection, but the *precision* will be poor. Generally there is a trade-off between *precision* and *recall*.

Figure 8 illustrates *precision-recall* relations for retrieving some objects. The horizontal axis indicates the *precision* and the vertical axis indicates the *recall*. From these results, we can also obtain the same conclusion which we observed from Figure 6 and Figure 7 --- this method can work well for objects with special or rich color regions, but is weak for objects which only have popular colors, such as simple black or gray.

The experiment was processed on a PC with CPU of 1G, memory of 384M, and the program was written with Visual C++. The average time of retrieving one image is about 50 seconds, which we thought is an acceptable cost compared with the cost of (Cohen, 1999) --- the time of per query-image comparison is about 0.1 seconds which is obviously a high cost while the volume of database grows to thousands of images.

## 4. CONCLUSIONS AND FUTURE WORK

In this paper, a novel method for partial image retrieval was proposed, which can retrieve images that contain the query image as a part regardless of the size and position. Images in image database were segmented in advance, for each major region obtained, a composite measurement of color, area percentage and position was extracted and stored as the feature. The pre-segmentation and the simple features allows the indexing of features of color regions and enables the fast retrieval. While retrieving, the query image was also first segmented, and then the major regions' colors, their area ratios and spatial relationships were generated for narrowing the searching space. Since the color feature is used as main feature, the proposed method can work well for objects with special or rich color regions, but is weak for objects which only have popular colors, such as simple black or gray, as shown in our experiment.

Our future work includes adding new features such as shape features to enhance the precision, and utilizing a high dimensional index structure to accelerate the retrieving process.

## REFERENCES

Cinque L., Levialdi S., et al. (1999) Color-Based Image Retrieval Using Spatial-Chromatic Histograms, *Proc. of Int. Conf. on Multimedia Computing and Systems Volume II*

Cohen S. (1999) Finding Color and Shape Patterns in Images, *Stanford University Technical Report STAN- CS-TR-99-1620*

Flickner M., Sawhney H., et al. (1995) Query by Image and Video Content: The QBIC System, *IEEE Computer, Vol. 28, No. 9*

Kim B. and Um K. (1999) 2D+ String: A Spatial Metadata to Reason Topological and Directional Relationships, *Proc. of the 11th International Conference on Scientific and Statistical Database Management*

Kim T-Y and H. Han J. H. (2000) Partial Image Matching by Measures From Connected Color Regions, *Proc. of IEEE International Conference on Multimedia and Expo, pp. 367-370*

Li J., Wang J. Z. and Wiederhold G. (2000) IRM: Integrated Region Matching for Image Retrieval, *Proc. of the 8th ACM International Conference on Multimedia*

Moghaddam B., Biermann H. and Margaritis D. (1999) Defining Image Content with Multiple Regions-of-Interest, *Proc. of IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL'99)*

Shih T. K., Wang C-S., Chang A. Y., and Kao C-H (2000) Indexing and Retrieval Scheme of the Image Database Based on Color and Spatial Relations, *Proc. of IEEE International Conference on Multimedia and Expo, pp. 129-132*

Smith J. R. and Chang S. F. (1999) Single Color Extraction and Image Query, *Proc. of IEEE International Conference on Image Processing*

Smith J. R. and Chang S. F. (1997) Querying By Color Regions Using the VisualSEEK Content-Based Visual Query System, *Intelligent Multimedia Information Retrieval. AAAI Press, Menlo Park, CA , pp. 23-41*

Venkat N. G. (1998) $\Theta\Re$ String: A Geometry-Based Representation for Efficient and Effective Retrieval of Images by Spatial Similarity, *IEEE Trans. on Knowledge and Data Engineering, Vol. 10, No. 3, pp. 504-509*

Yi T. and William I. G. (1999) Spatial Color Indexing: A Novel Approach for Content-Based Image Retrieval, *Proc. of the IEEE International Conference on Multimedia Computing and Systems Volume I*


# BIOGRAPHIES

**Yunyun Cao** received the B. Eng. Degree in Computer Application in 1987 and the M. Eng. Degree in Computer Organization and Architecture in 1990 from Tianjin University, and Ph.D degree in Information and Communication Engineering from the University of Tokyo in 2001. She is now an associate professor, School of Science, Tianjin University and a poster doctor, Institute of Industrial Science, University of Tokyo. She is a member of IEEE, ACM and IEICE(Japan). Her research interests include multimedia database, multimedia systems, image processing and understanding.

**Yoshitomo Yaginuma** received the B.Eng. degree in Phys. Instr. in 1988, and the M. Sci. degree in Physics in 1990, both from the University of Tokyo. He is now an associate professor, National Institute of Media Education. He is a member of Information Processing Society of Japan. His research interests include data structure, multimedia database and database acquisition

**Masao Sakauchi** is the Director general and a professor, Institute of Industrial Science, Univ. of Tokyo. He received the B.Sci. degree in Electrical Engineering from the Univ. of Tokyo in 1969 and the M S and Ph D degrees in Electronics Engineering from the Univ. of Tokyo in 1971 and 1975, respectively. He has acted as, General chairman of 4 International

Conferences and Workshops, Including the IEEE International Workshop on Machine Vision and Machine Intelligence (1987), IAPR and IEEE International Conference on Document Analysis and Recognition (ICDAR '93) (1993), International Symposium on Multimedia Mediation Systems(2000), Program Chairman of 3 International Conferences, Including IEEE International Conference of Multimedia Processing and Systems (IEEE Multimedia 96) (1996), ITSC 99 (1999), and Organizing and Program committee member of Many International conferences. He also served as Chairman of the technical committee on Machine Vision and Machine Intelligence, IEEE IE society (1985-1992), Associate editor of IEEE Multimedia Magazine, (1993-99) IEEE trans. on Industrial Electronics. (1984-1992), Editor in chief of Transactions on Information and Systems of Institute of Electronics, Information and Communication Engineers (IEICE) in Japan. (1989-1991), Chairman of three technical committees on Image Engineering of IEICE of Japan, and, Institute of Television Engineers of Japan (1987-1993). He wrote over 340 refereed papers in the research fields of multimedia databases, multimedia systems, image processing and understanding, spatial data structures, and geographical information systems and fault tolerant computing. He is now the project leader of two big research project on Multimedia Mediation Systems and ITS.