Retrieving Landscape Images Using Scene Structural Matrix

G. Qiu and S. Sudirman

School of Computer Science, The University of Nottingham
Jubilee Campus, Nottingham NG8 1BB, UK
{qiu,sxs}@cs.nott.ac.uk

Abstract. In this paper, we present Scene Structural Matrix (SSM) and apply it to the retrieval of landscape images. The SSM captures the overall structural characteristics of the scene by indexing the geometric features of the image. A binary image tree (bintree) is used to partition the image and from which we derive multi-resolution geometric structural descriptors of the image. It is shown that SSM is particularly effective in retrieving images with strong structural features, such as landscape photographs. We show that SSM is robust against spatial and spectral distortions thus making it superior to current state of the art techniques such as color correlogram in certain applications. We will also show that images retrieved by the SSM are more relevant than those returned by color correlogram and color histogram.

1 Introduction

Content-based indexing and retrieval [1] have attracted extensive research interests in recent years. Traditional methods use global statistics of local image features, e.g., color histogram [2], color correlogram [3] and their variance as image indices. These methods have been shown to be very successful in retrieving images with similar local feature distributions. However, since these measures do not take into account the locations of the local features, the retrieved results often do not make a lot of sense. For example, using a landscape image with blue sky on top and green countryside at the bottom as query example and trying to retrieve images with similar structures, i.e., blue sky on top and green countryside at the bottom, methods based on global statistics of local features often give very unsatisfactory results. Another scenario is one in which two or more images of the same scene photographed under different imaging conditions, e.g., images of a countryside taken at dusk or dawn under a clear or a cloudy sky. Using one of these images as a query example often fails to retrieve other images of the same scene taken under different time or conditions. In yet another situation maybe one in which a same scene imaged by different, uncalibrated devices. Using one image taken by one device may fail to find the same scene taken by other devices.

In this paper, we present a method which uses a binary image tree [8] to partition an image recursively into hierarchical sub-images and introduce the Scene Structural Matrix (SSM), a 2-dimensional table to summarize the geometric structures of the

H.-Y. Shum, M. Liao, and S.-F. Chang (Eds.): PCM 2001, LNCS 2195, pp. 921–926, 2001.
© Springer-Verlag Berlin Heidelberg 2001

partitioned image. We use the SSM as image indices for content-based image retrieval. Experiments have been performed on an image database consisting of over 7000 high-resolution photographic color images. It is shown that the SSM is particularly effective in retrieving images with strong structural features such as landscape images. It is also shown that the SSM is more robust to spatial and spectral distortions than traditional color histogram and state of the art color correlogram methods and therefore is advantageous in applications such as retrieving images of the same scene imaged at different time and under different imaging conditions.
The organization of the paper is as follows. Section 2 describes the scene structural matrix. Section 3 presents the application of SSM to content based image retrieval. Section 4 presents experimental results and section 5 concludes the paper

2 Scene Structural Matrix

To construct the SSM, we first segment the image using a simple, easy to implement, binary image tree (bintree) adaptive image segmentation method [8], which can be regarded as a much simplified version of binary space partitioning tree image segmentation method [4-6]. An image (assumed rectangular in shape) is first cut into two equal sized sub-images by either a vertical or a horizontal straight line. Each of the resulting two sub-images is again cut into two equal sized sub-images by either a vertical or a horizontal straight line. The process is repeated for each of the subsequent sub-images until a stopping criterion (e.g., when the pixel value variation in the sub-image falls below a preset value) is reached. Figure. 1 illustrates such a scheme.

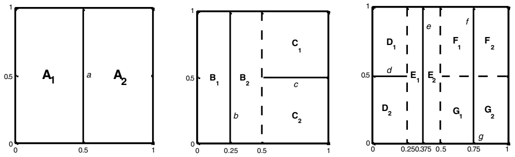


Fig. 1. The bintree adaptive image segmentation scheme. The original image is partitioned by the vertical line a into two equal sized halves (left); Lines b and c partition the resulting two halves (middle); and the resulting four sub-images are partitioned by the lines d, e, f and g (right). The capital letters denote the characteristics of the corresponding region. They might represent the region average color, the distribution of color or even textures.

Whether a sub-image (including the original) is cut by a horizontal or a vertical line depends on the structure of the sub-image (hence the segmentation is adaptive), and there are many criterion can be used. One method is to cut the sub-images based on the predominant edge orientations. In this method, the horizontal and vertical gradients are first calculated and the sub-image is cut based on the magnitudes of the directional gradients. If the vertical gradient dominates, then the sub-image is cut horizontally, otherwise vertically. Some edge detection operators [7] may be used to calculate the directional gradients. Let $G_h(i, j)$ and $G_v(i,j)$, i = 0, 1, ... M-1, $_j$ = 0, 1, ..., N-1, be the horizontal and vertical gradients respectively, of an M x N sub-image to be partitioned. We calculate

$$G_h = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left| G_h(i,j) \right| \quad G_v = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left| G_v(i,j) \right|$$

If $G_h > G_v$, the sub-image will be cut vertically otherwise horizontally. Figure 2 shows an example of the segmentation.

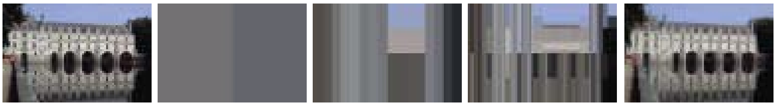


Fig. 2. From left to right: Original landscape image, 1st, 4th, 7th and 12th level partition. Each sub-image is painted by its average color

Based on the bintree segmented image, we construct a table termed scene structural matrix. The rationale for building the SSM is that because the reconstruction are perceptually similar to the original, then the features from the reconstruction can be used to recognize the original. The fact that we recursively cut the image with only horizontal and vertical lines means that there are some very simple geometry structures can be extracted. When a line partitions a sub-image, it intersects with the two borders of the sub-image which are perpendicular to it forming a T shape structure at different resolutions (refer to Figure 1, the dash and solid lines). It is based on this T shape structure we build our scene structural matrix. The SSM is a two dimensional array indexing the T shape structures of the bintree segmented image. There are only two types of T shape structures and their conjugates. Therefore only the + and the - structures need to be indexed since (+ and ¦) and (- and -) will always appear in pairs. It is clear, at different level and depending on how a sub-image is cut, the two arms of the T-shape features will have different length. The SSMs capture this fact by indexing the T-shape features of various sizes. There are two SSMs, SSM$^+$ and SSM$^-$ indexing the two unique T-shape features. Each cell in the matrix corresponds to the T-shape with a certain arm lengths. The values of the cells in SSM$^+$ are the accumulated average color difference across the horizontal arm, and the values of the cells in SSM$^-$ are the accumulated average color differences across the vertical arm. The formal description of the SSM is as follows: Let $h_i = H/2^i$, $v_j = V/2^j$, where i, j = 0, 1, 2, … are two integers, H is the horizontal dimension of the image and V is vertical dimension of the image.

Let h_a and v_a be the lengths of horizontal arm and vertical arm of the T shape features. We have

SSM$^+$ (i , j) = SSM$^+$ ($h_a = h_i$ and $v_a = v_i$) = ACC | C_{top} - C_{bottom} |
SSM$^-$ (i , j) = SSM$^-$ ($h_a = h_i$ and $v_a = v_i$) = ACC | C_{left} - C_{right} |

where ACC denotes accumulation, and C_{top} is the average color of the top half and C_{bottom} is the average color of the bottom half of the partition. Similarly, C_{left} is the average color of the left half and C_{right} is the average color of the right half of the partition. As an example, Table 1 shows the SSMs of the scene partitions of Figure 1.

SSM^+

0	0	0
$\lvert C_1-C_2\rvert$	0	x
$\lvert D_1-D_2\rvert$	x	x

SSM^-

$\lvert A_1-A_2\rvert$	$\lvert B_1-B_2\rvert$	$\lvert E_1-E_2\rvert$
0	$\lvert F_1-F_2\rvert + \lvert G_1-G_2\rvert$	x
0	x	x

Table 1. The Scene Structural Matrices (SSM) of the segmentation of Figure 1. The capital letters denote the average colour vectors of their corresponding sub-images

3 Content-Based Image Indexing and Retrieval Using SSM

We can use the SSMs as image indices for content-based indexing and retrieval in image database application. The SSMs for each image in the database are constructed and image retrieval can be based on comparing these SSMs with that of the query image. It is worth noting that only a few (5 to 8) levels of partition suffice, hence the size of SSM is very small. Let SSMq(i, j) and SSMd(i, j) be the SSMs of the query image and the database image respectively. The similarity of the two images can be calculated by

$$D(q, d) = \min\{ D1(q, d), D2(q, d) \}$$
$$D2(q, d) = \Sigma_i \Sigma_j (W(i,j)(|SSM^+q(i, j)-SSM^+d(i, j)| + |SSM^- q(i,j)-SSM^- d(i,j)|))$$
$$D2(q, d) = \Sigma_i \Sigma_j (W(i,j)(|SSM^+q(i,j)-SSM^- d(i,j)| + |SSM^- q(i,j) SSM^+d(i,j)|))$$

The min {} operator guarantees a zero difference between identical images rotated by 90°. The matrix W(i,j) allows us to use different weights to the contribution of coarse and detail features of the image. Intuitively we would like global features to be more dominant than detail features. A typical weight matrix is

	0	**1**	**2**
0	1	½	1/3
1	½	1/3	x
2	1/3	x	x

Based on the image similarity measures, images that have smaller distances are considered more similar to the query and are returned to the user.

4 Experimental Results

We have implemented the SSM method for image indexing and retrieval using an image database of 7400 high-resolution color photographic images, a subset of the commercially available Corel Photo collection widely used by other research groups. From the database, we randomly chose 50 landscape images. In the experiment, each of these images was subjected to various spatial and spectral processing before being used as query image. The aim was to use the distorted (processed) image as query and retrieve the original image from the database. The processing performed include spectral (color) modification, spatial resolution scaling, and spatial filtering. As a comparison, we have also implemented color histogram (CH) [2] (4096-bin) and color correlogram (CC) [3] (4 distances, 64 colors) methods. For the results presented, the SSMs were built based on 7 level bintree partitions. The cumulative recall rate, i.e., the number of retrieved original images above a certain rank, of various processed images as queries and using different query methods are shown in Table 2.

The results show that SSM method is more robust to color alteration than the other two methods. It is also far more robust to scaling than color correlogram method and relatively stable to spatial smoothing. Furthermore, SSM retrieved far more relevant images than the other two methods. Figure 3 shows some examples. It is clearly seen that CC returns completely irrelevant images while CH despite being able to retrieve the original image the other returned images have little relevance to the query. On the other hand, the SSM method not only successfully found the original image using the distorted version as query, it also returned much more relevant images. These results clearly demonstrate the superiority of SSM method as compared to the others.

5 Concluding Remarks

We have presented an image content descriptor which captures simple geometrical structures within an image and indexes them efficiently into two relatively small matrices. It has been shown that this method has good performance in retrieving the original image when the same image having undergone substantial spatial and spectral processing was used as query. Furthermore, the method has been shown to return much more relevant images than state of the art methods when applied to landscape image retrieval. The concept of SSM is unique in the literature. It is worth pointing out that in the current paper only one of many possible image features was used to fill the contents of SSM.

6 References

[1] Y. Rui et. al., “Image Retrieval: Current Techniques, Promising Directions, and Open Issues”, J. Visual Comm. Image Representation, vol.10, pp.39-62, 1999.

[2] M. J. Swain et. al., “Color Indexing”, Int. J. Computer Vision, Vol. 7, no. 1, pp.11-32, 1991.

[3] J. Huang, et.al., "Image indexing using color correlogram", Proceeding of Computer Vision and Pattern Recognition, pp.762-768, 1997.

[4] H. Radha et. al., "Image compression using binary space partitioning tree", IEEE Trans. on Image Processing, vol.5, pp.1610-1624, 1996.

[5] G. Qiu and S. Sudirman, “Representation and Retrieval of Color Image Using Binary Space Partitioning Tree”, Proceeding of 8th Color Image Conference, pp.195 - 201, 2000.

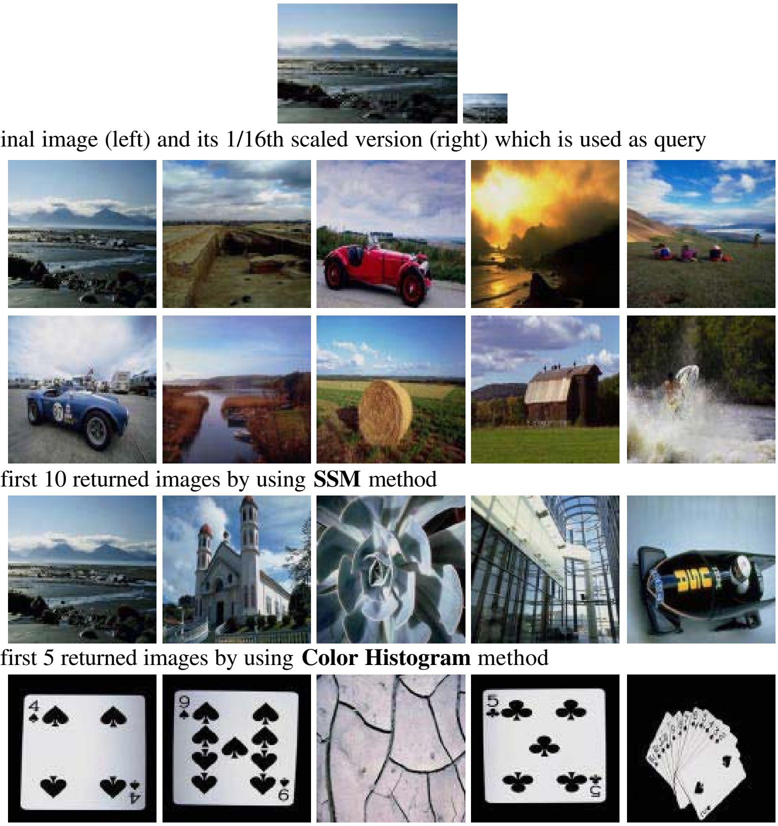
[6] X. Wu, “Image Coding by Adaptive Tree-Structured Segmentation”, ", IEEE Trans. on Image Processing, vol.38, pp.1755-1767, 1992.

[7] R. C. Gonzalez et. al., Digital Image Processing, Addison-Wesley, 1992

[8] H. Samet, Applications of spatial data structures: computer Graphics, image processing, and GIS, Addison Wesley, 1989

		Ranks											
		1			< =10			< = 20			< =50		
Distortion		C	S	F	C	S	F	C	S	F	C	S	F
Method	SSM	2	37	47	23	46	50	38	49	50	46	49	50
	CC	0	0	49	0	1	50	0	2	50	0	3	50
	CH	0	50	50	0	50	50	0	50	50	1	50	50

Table 2. Cumulative Recall Rate (out of 50) when the query images were subjected to various distortions. Method codes: SSM-Scene Structural Matrix; CC-Color Correlogram; CH-Color Histogram. Distortion codes: C-color modification (-80% of overall hue and saturation, and +30% of overall intensity); S-spatial scaling (reduced to 1/16 in size); F-Filtering processing (7x7-neighborhood averaging). The table should be interperated as for example: when the query images were subjected to colour distortion (C), SSM retrieved 2 (out of 50) target images in the 1st rank, CC and CH retrieved none in the 1st rank etc.



Original image (left) and its 1/16th scaled version (right) which is used as query

The first 10 returned images by using **SSM** method

The first 5 returned images by using **Color Histogram** method

The first 5 returned images by using **Color Correlogram** method

Fig. 3. The first few returned images by using three retrieval methods when the target image is scaled down to 1/16th of its original size. The images are displayed as square for convenience