SASI: A NEW TEXTURE DESCRIPTOR FOR CONTENT BASED IMAGE RETRIEVAL

Abdurrahman Çarkacıoğlu, Fatoş Yarman-Vural
Middle East Technical University
Department of Computer Engineering
Ankara, TURKEY

ABSTRACT

In this paper, a new texture descriptor, namely, Statistical Analysis of Structural Information (SASI) is introduced as a representation of texture. SASI is based on statistics of clique autocorrelation functions calculated over a set of directional moving windows. SASI defines a set of windows to extract and measure various structural properties of texture by using a spatial multiresolution method. Although it works in spatial domain, it measures the spectral information of a given texture. Experimental results, performed on digitized Brodatz Album, indicate that SASI is very successful in identifying the "similar" textures.

1. INTRODUCTION

In recent years, textural information become important for content based image representation and retrieval problems [1,2]. Although, many texture descriptors exist in the literature, the success of the results heavily depends on the data type and the problem domain. This is basically because of the inherent complexity of texture, which requires laborious mathematical description.

The concept of similarity is defined according to the design of the performance tests. Two basic approaches are used for this purpose. In the first approach, the textured images in a database are divided into subimages each of which are considered to be similar, whereas in the second approach visually similar images are grouped by a user in various classes. Although it may consist of visually dissimilar subimages, the first approach does not require human help, which may create subjectivity in the second approach. The performance of a texture descriptor is, then, measured in terms of retrieval rates of "similar" images in a database.

In recent years, Gabor Filters become very popular for measuring texture similarity. Among many others, the most successful results are reported by Manjunath & Ma [1,3] where the second order statistics of a set of Gabor Filter responses are used as a texture descriptor. The major drawback of Gabor filter descriptor lies behind the selection of the filter parameters, which heavily depend on the characteristics of the textures in the image database. The accurate implementation of a complete Gabor expansion would entail a generally impractical number of filters. Since the Gabor functions are not orthogonal, there is a trade-off between redundancy and completeness in the design of the Gabor Filter Banks. Also, in a digital world, it is not always possible to cope with all sizes of analog Gabor Filters, which may cause problems especially with the textures that consist of sharp corners or small texels.

The SASI descriptor, proposed in this study, is based on second order statistics of clique autocorrelation coefficients, which are the autocorrelation coefficients over a set of directional moving windows. The windows of various sizes and shapes are defined by using the concept of cliques, which describes the characteristics of textures in different granularity. Since the autocorrelation function is a mathematical cousin of Fourier transformation, SASI measures the spectral information, while it works in the spatial domain. This property gives powers of spectral descriptors to SASI. Like Gabor Filter descriptor, SASI employs different orientation and size of the moving windows. This is a spatial multiresolution decomposition property of SASI. However, implementation of SASI is more robust compared to Gabor Filters. SASI can also cope with a broader class of textures, which may consist of sharp corners or small primitives or texels, because of the flexibility in definition of clique windows.

2.STATISTICAL ANALYSIS OF STRUCTURAL INFORMATION (SASI)

2.1. Background

In this study, we assume that texture consists of locally stationary texels with varying granularity and structure, which can be captured by moving windows of varying sizes and shapes. The window size and shape is selected for measuring a stationary component of the image, using autocorrelation function analysis.

SASI combines the concept of clique by the autocorrelation function for defining the clique autocorrelation function and measures the second order statistics of the clique autocorrelation functions defined over the textured image. The following section gives the mathematical framework of SASI.

2.2. Definitions

Definition 1: Base Clique: Given a seed pixel (i,j) in a neighborhood system η , the base clique of ($\mathcal{L}$, η), denoted by a pair of pixels $B_p(ij,kl)$ where p = 1..P , is a subset of the Lattice $\mathcal{L}$ such that [4,5,6]

$$(i,j) \neq (k,l) \;,\; ij \in \eta_{kl}$$

where P indicates the total number of distinct base cliques. Fig. 1.a and 1.b indicate the base clique representation and the corresponding base cliques for second order neighborhood system $\eta 2$, respectively. For this case P=8.

Note that the above definition of base clique is a subset of classical clique definition used in the literature.

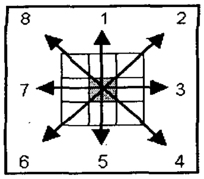

Figure 1.a. Base Clique Representation. Shaded Pixel taken as seed pixel.

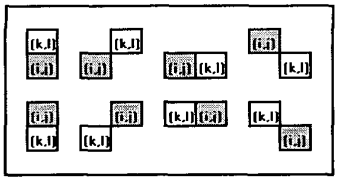

Figure 1.b. 2 neighborhood systems and its Base cliques Bp, P=1,2..8.

Definition 2: Clique Chain: Given pixel (i,j) as a center, the ***clique chain*** $C^\theta_L(i,j)$ **with length L** is the connected chain of pixels with the same base clique. $C^\theta_L(i,j)$ is given by

$$\{B_p(ab,cd)\cup B_p(cd,ef)\cup\ldots\cup B_p(gh,ij)\cup B_p(ij,kl)\cup\ldots\cup B_p(mn,qr)\cup B_p(qr,st)\}$$

where the total number of pixels in $C^\theta_L(i,j)$ is **L** and the total number of distinct base cliques P determines the number of possible θ values.

In η^2, clique chain is the θ degree rotated lines of pixels that has a center on pixel (i,j). Since η^2=8, only 4 direction clique chains can be obtained, as shown in Figure 2.

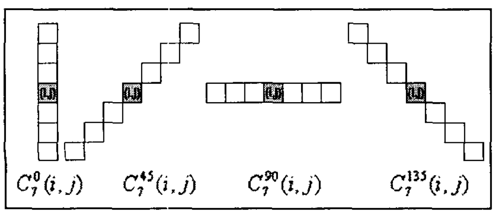

Figure 2. 4 orientations of clique chain with length 7.

Definition 3: Clique Window: $W^\theta_{s,ab}$ is an SxL structuring element which consist of S totally connected clique chains $C^\theta_L(i,j)$.

Practically speaking, in η^2 ,clique window, $W^\theta_{s,ab}$, is an SxL structuring element rotated by θ degrees. Figure 3 illustrates clique window of size L=S=3,7. Note that, if L=S then $W^0_{S,ab}$ is the same as $W^{90}_{S,ab}$. Therefore, with an eight neighborhood system, for a given size S=L, only 3 different types of clique window can be defined.

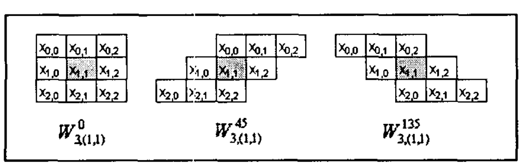

Figure 3.a: 0, 45 and 135 degrees 3x3 clique windows.

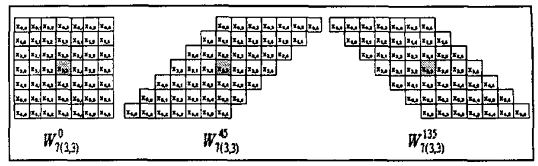

Figure 3.b: 0, 45, 135 degrees 7x7 clique windows.

Definition 4: Clique autocorrelation coefficient at lag vector $\bar{v}(k,l)$ of a given seed pixel for a clique window $W^\theta_{S,ab}$ is given by

$$r(\bar{v})^{W^\theta_{S,ab}} = r(k,l)^{W^\theta_{S,ab}} = \frac{\sum\sum_{\forall(i,j),(i+k,j+l)\in W^\theta_{S,ab}}(x_{i,j}-\bar{x}_{i,j})(x_{i+k,j+l}-\bar{x}_{i+k,j+l})}{\sqrt{\sum_{\forall(i,j)\in W^\theta_{S,ab}}\sum(x_{i,j}-\bar{x}_{i,j})^2\ \sum\sum_{\forall(i+k,j+l)\in W^\theta_{S,ab}}(x_{i+k,j+l}-\bar{x}_{i+k,j+l})^2}}$$

where $x_{i,j}$ represents the gray value of the image at position (i,j), and $\bar{x}_{i,j}$ represents the mean value of $x_{i,j}$ within $W^\theta_{S,ab}$, and $\bar{v}(k,l)$ represents a lag vector is defined between two locations of a clique window. Autocorrelation coefficients of a clique window depend only on the length and direction of the lag vector. Note that $r(k,l)^{W^\theta_{S,ab}} = r(-k,-l)^{W^\theta_{S,ab}}$ and $-1 \le r(k,l)^{W^\theta_{S,ab}} \le 1$.

In another words, clique autocorrelation coefficients can be considered as short-term correlograms over clique windows defined by clique chains. Fine components of the textures are captured in small size clique windows, whereas coarse components are captured in relatively large size clique windows.

In this study, direction of the lag vector $\bar{v}(k,l)$ is set equal to the direction of clique window. Table 1 summaries the selection of lag vectors.

Clique Window	Used Lag Vector (where K<S)			
	We assume $K=\left\lceil \frac{S}{4} \right\rceil+1$			
$W^0_{S,ab}$	v(0,1)	v(0,2)	...	v(0,K)
$W^0_{S,ab}=W^{90}_{S,ab}$	v(1,0)	v(2,0)	...	v(K,0)
$W^{45}_{S,ab}$	v(1,1)	v(2,2)	...	v(K,K)
$W^{135}_{S,ab}$	v(-1,1)	v(-2,2)	...	v(-K,K)

Table 1. Clique Window versus Lag Vectors.

Definition 5: Mean value and **standard deviation** of clique autocorrelation functions with lag vector $\bar{v}(k,l)$ over all clique windows $W^\theta_{S,ab}$ is defined as follows, respectively.

$$\mu^\theta_S(k,l) \approx \frac{1}{N_x}\sum_{\forall(a,b)\in X} r(k,l)^{W^\theta_{S,ab}} \qquad \text{and}$$

$$\sigma_S^{\theta}(k,l) \approx \sqrt{\frac{1}{N_x} \sum_{\forall (a,b) \in X} \left(r(k,l)^{W^{\theta}_{S,ab}} - \mu_S^{\theta}(k,l) \right)^2} \quad ,$$

where X represents an image itself, N_X represents a number of pixel in image X, (k,l) represents a lag vector, θ represents a clique window orientation and S represents the clique window size.

Definition 6: SASI Descriptor: For a given texture **T, SASI** descriptor is defined as a vector with the entries $\mu_S^{\theta}(k,l)$ and $\sigma_S^{\theta}(k,l)$ as

$$D_T = \{ \mu_{S_1}^{\theta_1}(k_1,l_1), \ldots, \mu_{S_Z}^{\theta_Z}(k_Z,l_Z), \sigma_{S_1}^{\theta_1}(k_1,l_1), \ldots, \sigma_{S_Z}^{\theta_Z}(k_Z,l_Z) \}^T$$

where $2 * Z$ is the feature vector size.

Given, $D_T = [f_1, f_2, \ldots, f_{2*Z}]$ then **normalized SASI descriptor,** $D'_T = [f'_1, f'_2, \ldots, f'_{2*Z}]$, is defined by normalizing the entries of D_T as follows.

$$f_i' = \frac{f_i - \mu_{f_i}}{\sigma_{f_i}} \qquad i = 1..2*Z$$

where μ_{f_i} is the mean value and σ_{f_i} is the standard deviation of the features over the entire database.

D'_T measures the structural information by using the second order statistics of local autocorrelation functions for texture T. At this point, any distance measure can be used to measure the mathematical similarity between the textures by using D_T. The size of the descriptor D'_T depends on the image database.

In this study, the mathematical similarity between the textures **T1** and **T2** is measured by the following metric:

$$S(T_1,T_2) = \frac{D'_{T_1} \bullet D'_{T_2}}{D'_{T_1} \bullet D'_{T_1} + D'_{T_2} \bullet D'_{T_2} - D'_{T_1} \bullet D'_{T_2}} \quad ,$$

where $\bullet$ stands for dot product.

2.3 Algorithm of SASI

1. Determine the clique window sizes S
   1.1. Determine the lag vectors v(k,l)'s to be used
2. For each clique window W
   2.1. For each lag vector v(k,l)
      2.1.1. For each pixel
         2.1.1.1. Define clique window W
         2.1.1.2. Calculate r(k,l) as given in Definition 4.
      2.1.2. Calculate mean value and standard deviation of r(k,l)
3. Construct D_T vector and normalized D'_T vector (as given in Definition 6)

Clique windows sizes and clique autocorrelation coefficients to be used are the parameters of SASI. Most crucial part of the algorithm is selecting the clique window sizes S to be used. Basically, clique windows sizes should be small enough to capture small primitives, and big enough to capture large patterns in images of chosen image database.

3. EXPERIMENTS

Two sets of experiments are fulfilled to test the performance of SASI descriptor, in C programming language. First, the outputs of Gabor Filters are compared to clique autocorrelation functions. Then the retrieval rates of both Gabor and SASI descriptors are compared.

In Figure 4, texture D052 from Brodatz Album is filtered with 4 scales of vertical Gabor Filters. For each orientation and scales different aspects of a texture are captured.

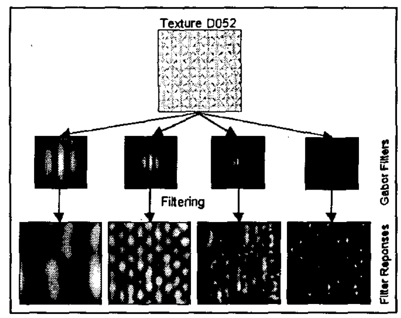

Figure 4. Gabor Filter outputs for vertical direction

In order to see how SASI works as a spatial multiresolution decomposition method, SASI algorithm is applied to the same texture namely D052, shown in Figure 5. Totally 3 vertical clique window is defined which has size 3x3, 5x5, 15x15. Each r(k,l) is scaled to 0 to 255 in order to analyze the outputs in visual environment. Each clique window investigates different aspects of texture like Gabor Filter does. As it can be seen from the figures, SASI captures vertical primitives or texels better than that of Gabor, especially for clique windows of size 3x3 and 5x5.

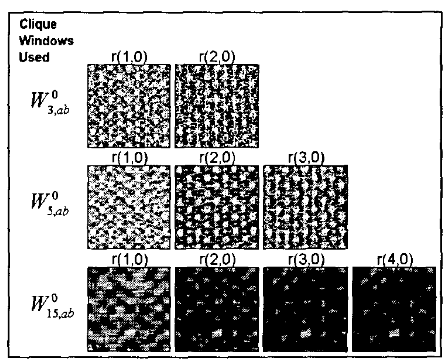

Figure 5: SASI vertical analysis of texture D052.

Retrieval rates of Gabor and SASI descriptors are tested on digitized Brodatz Album, which consists of 112 images (112 Classes) of size 512x512 and 256 gray values. After dividing each image into 16 nonoverlapping subimages, total of 1792 images are obtained. The performance of the proposed descriptor is measured in terms of the average retrieval rate, which is defined as the average percentage number of patterns belonging to the same image as the query pattern in top 15 matches (self matches are excluded) [3]. Another words, for each subimage the most similar 15 subimages are retrieved within 1791 subimages and it is expected that retrieved subimages and the query subimage are part of a same image.

In order to represent any given texture from Brodatz Album, 80 real valued (40 mean value+40 std.dev.s) SASI descriptor is calculated. Selected clique windows and related autocorrelation coefficient is shown in Table 2.

3x3	*5x5*	*15x15*
$r(K_1,0)^{W^{0}_{3,ab}}$	$r(K_2,0)^{W^{0}_{5,ab}}$	$r(K_3,0)^{W^{0}_{15,ab}}$
$r(K_1,K_1)^{W^{45}_{3,ab}}$	$r(K_2,K_2)^{W^{45}_{5,ab}}$	$r(K_3,K_3)^{W^{45}_{15,ab}}$
$r(0,K_1)^{W^{90}_{3,ab}}$	$r(0,K_2)^{W^{90}_{5,ab}}$	$r(0,K_3)^{W^{90}_{15,ab}}$
$r(-K_1,K_1)^{W^{135}_{3,ab}}$	$r(-K_2,K_2)^{W^{135}_{5,ab}}$	$r(-K_3,K_3)^{W^{135}_{15,ab}}$
Where K_1 ***=1,2 and*** K_2 ***=1,2,3 and*** K_3 ***= 1,2,3,4,5***		

Table 2. Selected clique windows and autocorrelation coeff.s

While average retrieval rate of the Gabor descriptor is 74.07%, average retrieval rate of the SASI descriptor is **76.37%** on Brodatz data set.[1]

Depending on the application domain, normalization of a descriptor may not be possible or feasible (e.g. real time image search). In order to show the effect of normalization on SASI and Gabor descriptor, Euclidean distance measure is used without any normalization. This time average retrieval rate of the Gabor descriptor is **62.15%,** whereas average retrieval rate of the SASI descriptor is **73.39%** on Brodatz data set.

Brodatz Album was never intended to give a fully representative sample set of a broad class of textures for testing the full performance of the descriptors. Splitting the images into subimages may sometimes yield visually dissimilar textures. Thus, in order to measure the human consistent performance of a descriptor the similar images can be clustered by the human.

Because it is hard to group each subimage in our test set (112x16=1792 subimage) manually, 112 texture images of Brodatz Album are grouped into 32 different clusters, each of which containing 1-6 similar textures [3]. However, this clustering process cannot fully eliminate the problems of Brodatz, mentioned above.

[1] For more experiments and explanations, see http://www.ceng.metu.edu.tr/~carkaci/sasi.htm

Figure 6 illustrates retrieval rates, based on 32 clusters. Average retrieval rate of SASI descriptor is higher than that of Gabor filter descriptor.

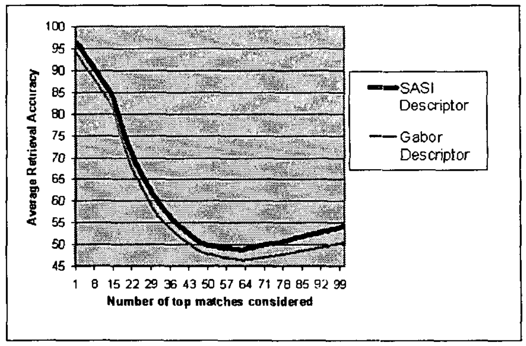

Figure 6. Retrieval Performance After Clustering

The algorithmic complexity of SASI is $O(S * N)$ where S represents clique window size and N represents image size, whereas the algorithm complexity of Gabor Filter is $O(N * \log N)$ when filtering is done in frequency domain.

4. CONLUSION

In this paper, a new descriptor is introduced to measure the similarity of textures. Simulation experiments are done on digitized Brodatz Album. It is observed that SASI descriptor captures the structural property of the texture better then the Gabor filters. This is basically because of the flexibility in designing the clique window and the power of the autocorrelation function defined over the window. The second order statistics of clique autocorrelation functions on a given texture, provides most of the visual information about the appearance of texture. This fact is verified during the performance tests based on average retrieval rates applied on subimages and visually clustered images of Brodatz Album.

5. REFERENCES

[1] B.S.Manjunath, Peng Wu, S.D.Newsam, H.D.Shin, "A Texture Descriptor for Content-based Image Retrieval" Image Communication Journal, Nov 1999.

[2] Yong Rui, Thomas S. Huang, and Shih-Fu Chang, "Image Retrieval: Current Techniques, Promising Directions and Open Issues", J. of Visual Comminication and Image Representation, 10, 1-23(1999)

[3] B.S. Manjunath and W.Y. Ma, "Texture features for browsing and retrieval of image data," IEEE Trans. On PAMI, Vol. 18, No. 8, pp 837-842, Aug. 1996.

[4] Cross and A.K.Jain, "Markov random field texture models", IEEE Trans. On PAMI., vol. 1-5, pp. 25 - 39, 1983.

[5] A.Carkacioglu and F.T. Yarman-Vural, "A Set of Similarity Measures for Binary and Gray Level MRF Textures", IAPR - ICIAP Vol. 2, 127-133, Italy 1997.

[6] A. Carkacioglu and F.T. Yarman-Vural, "Set of Texture Similarity Measures", Electronic Imaging 97, Machine Vision Applications in Industrial Inspection , SPIE Proceedings Vol. 3029, pp. 118-127, San Jose, 1997.