논문-00-5-1-03

A Texture Browsing Descriptor based on the Gabor Wavelets

Yanglim Choi*, Hyundoo Shin*, B. S Manjunath**, and Peng Wu**

I. Introduction

In recent years, image texture has emerged as an important visual primitive to search and browse through large collections of similar looking patterns. An image can be considered as a mosaic of textures and texture features associated with the regions can be used to index the image data. To support image retrieval or browsing using texture, an efficient representation is required.

The proposed texture descriptor relates to a perceptual characterization of texture, similar to a human characterization, in terms of regularity and directionality. This representation is useful for browsing type applications and coarse classification of textures. We call this descriptor the Perceptual Browsing Component (PBC). The feature extraction is simple, involving image convolutions with a set of masks. The filters are based on a 2-D Gabor wavelet decomposition. Image convolutions can be efficiently implemented in hardware and software. Using PBC, browsing of image database could be performed (*e.g., show textures that are structured and are oriented at 90 deg.*)

The use of Gabor filters in extracting textured image features is motivated by various factors. The Gabor representation has been shown to be optimal in the sense of minimizing the joint two-dimensional uncertainty in space and frequency[3]. These filters can be considered as orientation and scale tunable edge and line(bar) detectors, and the statistics of these micro features in a given region are often used to characterize the underlying texture information. Gabor features have been used in several image analysis applications including texture classification and segmentation[1],[4], image recognition[5],[7], image registration, and motion tracking[6].

The proposal is organized as following. Section II gives the formation of the texture descriptor. The extraction method is introduced in section III. Experiment results are presented in Section IV.

* DMS Lab, Samsung Electronics
** Department of Electrical and Computer Engineering, University of California at Santa Barbara

II. Proposed Texture Descriptor

The proposed texture descriptor vector is called the "Perceptual Browsing Component" and has the following format:

$$PBC=[v_1 \quad v_2 \quad v_3 \quad v_4 \quad v_5]$$

It's 5-dimensional vector consisting of integers and the semantics of each component is the following.

- $v_1 \in \{1,\ldots,N_v\}$, the larger the value is, the more structured the corresponding the texture is. Thus, provides a confidence measure on the regularity of the texture.
- $v_2, v_3 \in \{1,\ldots,K\}$ give two quantized directions that best capture the regularity. For example, '1' may correspond to 0°, and K correspond to 150° when the direction are quantized to intervals.

- $v_4, v_5 \in \{1, \ldots, S\}$ give two quantized scales that best capture the regularity. For example, S correspond to the number of levels in a wavelet decomposition

The reader is advised to quickly look through the calculated values in that section before jumping into the next section on the extraction of this vector.

III. Extraction Method

1. Multi resolution Decomposition by Using Gabor Wavelet

To perform the multiresolution decomposition, a set of Gabor filters is developed. A two dimensional Gabor function and its Fourier transform[3] can be written as:

$$g(x,y)=\left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right] \quad (1)$$

$$G(u,v)=\exp\left\{-\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right\} \quad (2)$$

where $\sigma_u = 1/2\pi\sigma_x$ and $\sigma_v = 1/2\pi\sigma_y$. Gabor functions form a complete but non-orthogonal basis set. Expanding a signal using this basis provides a localized frequency description. A class of self-similar functions, referred to as *Gabor wavelets* in the following discussion, is now considered. Let $g(x,y)$ be the mother Gabor wavelet, then this self-similar filter dictionary can be obtained by appropriate dilations and rotations of $g(x,y)$ through the generating function:

$$g_{mn}(x,y)=a^{-m}G(x',y'),\ a\rangle 1,\ m,n=\text{ integer}$$
$$y'=a^{-m}(-x\sin\theta+y\cos\theta),\ x'=a^{-m}(x\cos\theta+y\sin\theta) \quad (3)$$

where $\theta = n\pi/K$ and K is the total number of orientations. The scale factor a^{-m} in(3) is meant to ensure that the energy is independent of m. Since the effect of a^{-m} in the x' and y', in computing the energy, will be compensated with the a^{-m} scale factor

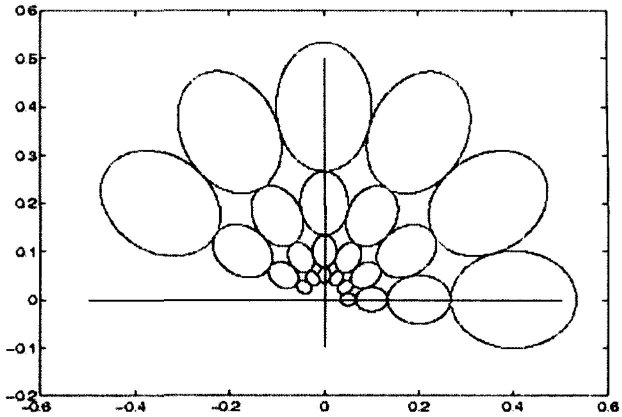


Fig. 1. The contours indicate the half-peak magnitude of the filter responses in the Gabor filter dictionary. The filter parameters used are Uh=0304

in front of $G(x', y')$.

The non-orthogonality of the Gabor wavelets implies that there is redundant information in the filtered images, and the following strategy is used to reduce this redundancy. Let U_l and U_h denote the lower and upper center frequencies of interest. Let K be the number of orientations and S be the number of scales in the multi-resolution decomposition. Then the design strategy is to ensure that the half-peak magnitude supports of the filter responses in the frequency spectrum touch each other as shown in Figure 1. This results in the following formulas for computing the filter parameters σ_u and σ_v(and thus σ_x and σ_y).

$$\sigma_u=\frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}} \quad a=\left(U_h/U_l\right)^{\frac{1}{S-1}},$$
$$\sigma_v=\tan\left(\frac{\pi}{2k}\right)\left[U_h-2\ln 2\left(\frac{\sigma_u^2}{U_h}\right)\right]\left[2\ln 2-\frac{(2\ln 2)^2\sigma_u^2}{U_h^2}\right]^{-\frac{1}{2}} \quad (4)$$

where $W = U_h$and $m = 0, 1, \ldots, S-1$. In order to eliminate sensitivity of the filter response to absolute intensity values, the real (even) components of the 2-D Gabor filters are biased by adding a constant to make them zero mean (This can also be done by setting $g(0,0)$ in (2) to zero). In the experiment, we set $S=4$ and $K=6$. For a given image $I(x,y)$, its decomposed image at scale m and direction n is defined to be .

$$W_{mn}(x,y) = \int I(x,y) g^{*}_{mn}(x - x_1, y - y_1) dx_1 dy_1 \qquad (5)$$

where * indicates the complex conjugate, $W_{mn}(x,y)$ is the filtered image at scale m and direction n.

2. Extraction of PBC

From the multi resolution decomposition, a given image is decomposed into a set of filtered images. Each of them represents the image information at a certain scale and at a certain orientation. The objective of regularity analysis is to get a quantitative measure of regularity from these filtered images. It is based on the following observations.

- Regular textures usually consist of one dominant periodic pattern. The more regular the periodicity is, the stronger the regularity is.
- The repetitive pattern, if it exists, could be captured by the filtered images. This behavior is usually captured in more than one filtered output.
- The dominant scale and orientation information can also be computed from these filtered outputs.

The analysis of regularity is a two step procedure. The first step is the analysis on each filtered image. The objective of this step is to determine the existence of the repetitive pattern. The second step is performed on all the filtered images that are identified as having some kind of periodicity. From the second step, a quantitative measure is derived to characterize the regularity. The information about the dominant scale and the dominant direction, in terms of contributing to the formation of structure, are also derived. The following diagram shows the procedure.

For each filtered image, the projections along horizontal and vertical directions are computed. These are denoted as P_H and P_V, respectively. (if a texture has a dominant orientation other than horizontal and vertical, it is possible compensate for this by taking

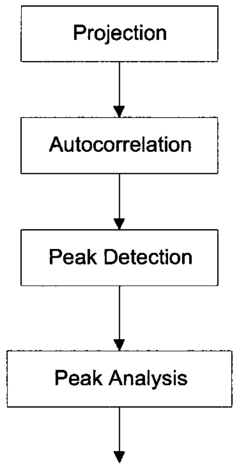


Fig. 2. The diagram of the analysis of each filtered image

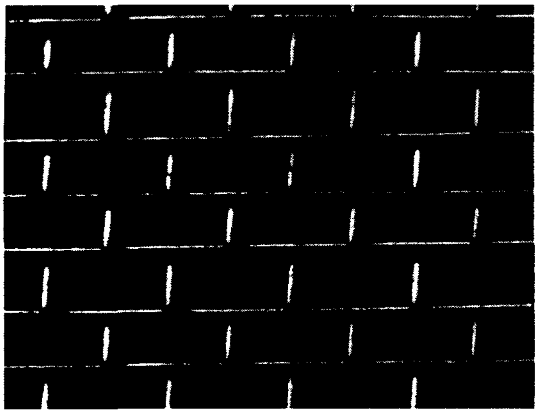
the Radon projections along the dominant direction. For notational simplicity, we consider only the horizontal and vertical projections here.) For each filtered image, its horizontal or vertical projection is a one-dimensional sequence, which can be denoted as $p(l)$, where $l = 1, \ldots, N$. On each of the projections, the normalized auto correlation function (NAC) is

$$NAC(k) = \frac{\sum_{m=k}^{N-1} p(m-k)p(m)}{\sqrt{\sum_{m=k}^{N-1} p^2(m-k) \sum_{m=k}^{N-1} p^2(m)}} \qquad (6)$$

Given the autocorrelation function $NAC(k)$, a peak and a valley are defined in the following manner:

{ $NAC(k)$ is a peak iff $NAC(k-1) \le NAC(k)$ and $NAC(k+1) \le NAC(k)$
$NAC(k)$ is a peak iff $NAC(k-1) \ge NAC(k)$ and $NAC(k+1) \ge NAC(k)$

For the detected peak points and valley points, their position and magnitude are recorded. Let M be the number of peaks and N be the number of valleys, then $p\_posi(i)$, $p\_magn(i)$ are the positions and magnitudes of those peak points, where $i = 1, \ldots M$, $v\_posi(j)$, $p\_magn(j)$ are the positions and magnitudes of the valley points, where $j = 1, \ldots, N$. The



T001.01 *256×256*
Fig. 3(a).

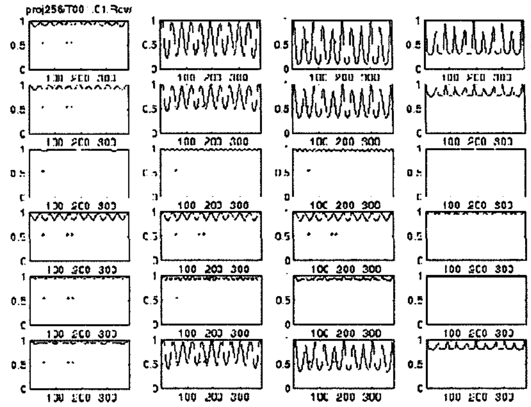


Fig. 3(b). NAC of horizontal projections of all the filtered images from image T001.01.

Fig. 3. An example image and its autocorrelation function NAC of the horizontal projections.

contrast of the projection is defined to be

$$contrast = \frac{1}{M}\sum_{i=1}^{M} p\_magn(i) - \frac{1}{N}\sum_{j=1}^{N} v\_magn(i) \qquad (7)$$

Given a peak sequence $p\_posi(i)$ with length M, two measures are extracted, the average of the distances between the peaks, *dis*, and the standard deviation of distances, *std*. A ratio is defined as following

$$\gamma = \frac{std}{dis} \qquad (8)$$

If this ratio is smaller that a pre-selected threshold T_{p}, the corresponding projection is considered to represent periodic information. The projection would be chosen for further analysis and represented by a vector $p_{mn}(dis, std)$, where m and n denote the scale and direction information associated with the filtered output. These are the potential candidates for further analysis. Among the P_h or P_v the coming from all the filtered images, some of them would be selected as potential candidates. They form a set of vectors $\{p_{mn}(dis, std)\}$. A modified agglomerative clustering is used to select the final candidate projections containing the regularity information. Let us denote the final set of selected candidates as $\{C_{mn}(dis, std)\}(H)$ and $\{C_{mn}(dis, std)\}(V)$ for the horizontal and vertical projections, respectively, which are then used to compute the PBC. Recall that

$$PBC = [v_1 \quad v_2 \quad v_3 \quad v_4 \quad v_5]$$

Once the candidates are selected, four elements of PBC, $PBC.[v_2 \quad v_3]$ and $PBC.[v_4 \quad v_5]$ could be determined. Denoting $C(m^*(H), n^*(H))(H)$ as the candidate projection from P_H that has the maximum contrast and $C(m^*(V), n^*(V))(V)$ as the candidate projection from P_v that has the maximum contrast. Then we have

$$PBC.v_4 = m^*(H) \text{ and } PBC.v_2 = n^*(H)$$
$$PBC.v_5 = m^*(V) \text{ and } PBC.v_3 = n^*(V)$$

The method of measuring $PBC.v_1$ is based on the following observations on the distribution of candidate vectors.

- For strong structured textures, their periodicity could be captured by multiple projections --- the candidates chosen from the above procedure. These candidates are usually close to each other.
- If the texture is not structured or only weakly structured, the distribution of the candidates, if they exist, usually is sparse and the neighboring relationship can rarely be detected.

If such a consistency in the neighboring projections is detected from among the projections in the candidate set, this would indicate a stronger regularity. The candidate projections are further classified as follows:

- C_1 : For a specific candidate, we can find at least one another candidate at its neighboring scale or orientation.

The value associated with this class is $V_1 = 1.0$.

- C_2 : For a specific candidate, we can find at least one another candidate distributed at the same scale or orientation but no candidate is located at its neighboring scale or orientation. The value associated with this class is $V_2 = 0.5$.
- C_3 : The candidate is the only one distributed at its scale and orientation. The value associated with this class is $V_3 = 0.2$.

And let

$$M = \sum_{i=1}^{3} N_i * V_i \qquad (9)$$

where N_i is the number of candidate projections classified as C_i. M is calculated independetly for the

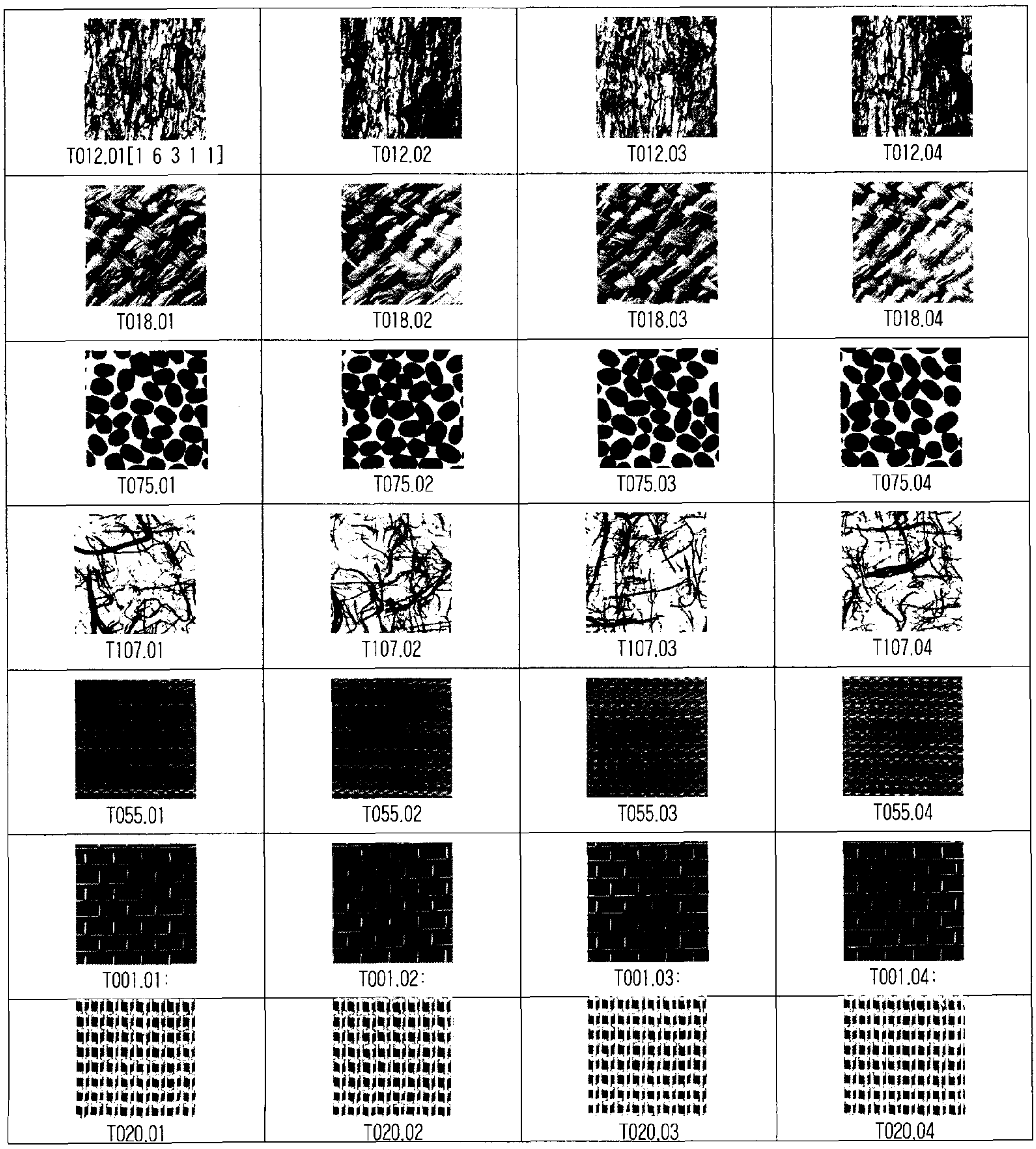


Fig. 4. Example images and their PBC vectors

horizontal (M_H) and vertical (M_v) projections. Let

$$M_{img} = M_h + M_v \tag{10}$$

M_{img} is quantized into N_v bins. The larger the value of M_{img} is, the more structured the corres- ponding texture is. In our implementation, $N_v = 4$. Consequently, each image is associated with a number $B_{img}(B_{img} \in 1, \ldots, N_v)$ to indicate which bin a image belongs to.

$$PBC.v_i = B_{img}$$

IV. Experiment Results

To demonstrate the performance of PBC, each of 512×512 image from the Brodatz database[2] is divided to 4 256×256 subimages. In this experiment, 15 different image textures are selected which are

class1={T001,T006,T014,T020,T095}
class2={T002,T007,T012,T097,T107}
class3={T009,T018,T055,T067,T075}

These images are selected subjectively so that class 1 includes images that are deemed as strong structured, class 2 contains images that are strong non structured and images in class3 can not get general agreement on its intensity of regularity. The following figure 4 shows the PBC value of the some of the image for each class.

We set $N_v = 4$, $S = 4$, and $K = 6$. So each element of *PBC* vector has the meaning as following,

- $v_1 \in \{1, \ldots, 4\}$, the larger the v_i is, the more regular the texture is. For example, a image with $v_1 = 4$ would be one of the most structured images, like T001.01, etc.
- $v_2, v_3 \in \{1, \ldots, 6\}$ give two quantized orientations that best capture the regularity. The orientation from 0° to 180° is quantized to 6 orientations with a 30° interval. So '1' corresponds to 0°, '2' corresponds to 30° and so on. v_2 and v_3 could be same.
- $v_4, v_5 \in \{1, \ldots, 4\}$ give two quantized scales that best capture the regularity. v_4 and v_5 could be same.

For example, for image T001.01 which has the *PBC* vector as

$$PBC_{T001.01} = [4 \;\; 1 \;\; 4 \;\; 3 \;\; 3]$$

we can tell that it is a strong structured image, its regularity can be best captured from orientation 1 and 4, which are 0° and 90° respectively and scale 3.

From the experiment, we observe the following

- $PCB.v_1$ gives a good description of the regularity that has the agreement with human's perception.
- $PCB.v_1$ also provides a confidence measure of v_i, $i = 2,3,4,5$. The larger the v_1 is, the more trust-ful the v_i, $i = 2,3,4,5$, are.
- For images with large, $PBC.v_1$, $v_i(i = 2,3,4,5)$ provide good estimation about at which orientations and which scales the regularity can be best captured. Note that the inconsistency of *PBC* vectors of T020 is due to the multiple periodicity

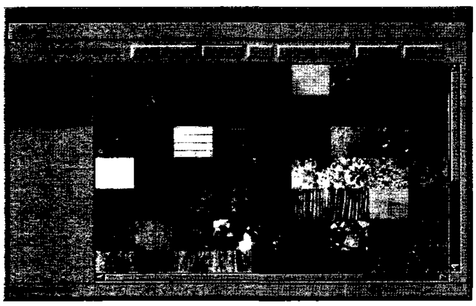


Fig. 5(a). The high lighted image is used as query image for browsing

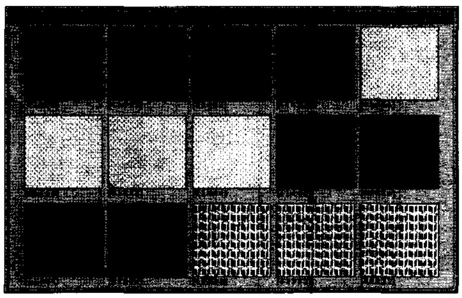


Fig. 5(b). The images that have the similar regularity with the highlighted image in (a) contained in the image texture.

In the following, an example of using *PBC* in an image browsing application is presented. Fig 5(a) gives the interface for image browsing. Fig 5(b) provides an example to browse images that are as structured as the image highlighted in Fig 5(a).

V. Conclusion

A simple 5-dimensional vector called "Perceptual Browsing Vector" is devised as a descriptor for a texture image. It quantifies the regularity, the directionality and the scale information of a given texture. The extraction process of such vector is explained. It is based on using the Gabor filter banks to filter an image and the analysis of the projection images of such filtered images. The performance evaluation of such vector is carried out on the Brodatz database. Subjective testing shows that it does achieve what it's supposed to do, i.e. the characterization of the browsing information of the given texture. In the future, the extraction process needs to be improved in terms of precision and speed and more experiments on various kinds of images should be done.

References

[1] C. Bovic, M. Clark, W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 12, pp. 55-73, Jan. 1990.

[2] P. Brodatz, "Textures: A photographic album for artists & designers," *Dover, NY*, 1966.

[3] J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Trans. ASSP*, vol. 36, pp. 1169-1179, Jul. 1988.

[4] B. S. Manjunath and R. Chellappa, "A Feature Based Approach to Face Recognition", *Proc. IEEE Conf. CVPR'92*, pp. 373-378, Champaign, Ill, Jun. 1992.

[5] B. S. Manjunath, C. Shekhar, and R. Chellappa, "A New Approach to Image Feature Detection with Applications", *Pattern Recognition*, Apr. 1996.

[6] J. G. Daugman, "High Confidence Visual Recognition of Persons by a Test of Statistical Independence", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1,148-1,161, Nov. 1993

[7] C. S. Li, J. R. Smith, V. Castelli and L. Bergman, "Comparing Texture Feature Set for Retrieving Core Images in Petroleum Applications", *Proceedings of SPIE Storage and Retrieval for Image and Video Databases VII, San Jose, CA*, Jan. 26-29, 1999, SPIE vol. 3656, pp.2-11.

저 자 소 개

Yanglim Choi

Yanglim Choi has received B.S. in 199 from the Korea Advanced Institute of the Science and Technology and Ph.D. in 1997 from the California Institute of Technology, both in Pure Mathematics. He joined Samsung Digital Media System Lab at 1998 and since then has been involved in the MPEG-7 Activity. His area of interest is the application of geometric topology into the area of image processing.

저 자 소 개

Hyundoo Shin

Hyundoo Shin has received Ph.D. from the Brown University in the area of Applied Mathematics. Currently he is a senior manager of the Digital Media System Lab at Samsung. His work for the past 2 years mainly has been following the MPEG-7 activity. His area of interest is the watermarking and the data hiding techniques.

Peng Wu

B.S. Manjunath is currently a professor at the department of Electronic and Computer Engineering of the University of California as Santa Barbara. He received his Ph.D. in 1991 from the University of Southern California. He joined the MPEG-7 activity 2 years ago and now serves as the chair for the Color-Texture CE adhog group. His area of interests are various image feature descriptor, face recognition techniques, image segmentation scheme, etc.

B.S Manjunath

Peng Wu is currently a graduate student at the University of California at Santa Barbara pursuing the degree of Ph.D. He worked with Professor Manjunath on developing the perceptual texture descriptor. His current research of interest in the efficient indexing scheme of the high dimensional data.