JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Content-based Image Retrieval Using Texture Features Extracted from Local Energy and Local Correlation of Gabor Transformed Images

Hee-Hyung Bu*, Nam-Chul Kim*, Bae-Ho Lee**, and Sung-Ho Kim***

Abstract

In this paper, a texture feature extraction method using local energy and local correlation of Gabor transformed images is proposed and applied to an image retrieval system. The Gabor wavelet is known to be similar to the response of the human visual system. The outputs of the Gabor transformation are robust to variants of object size and illumination. Due to such advantages, it has been actively studied in various fields such as image retrieval, classification, analysis, etc. In this paper, in order to fully exploit the superior aspects of Gabor wavelet, local energy and local correlation features are extracted from Gabor transformed images and then applied to an image retrieval system. Some experiments are conducted to compare the performance of the proposed method with those of the conventional Gabor method and the popular rotation-invariant uniform local binary pattern (RULBP) method in terms of precision vs recall. The Mahalanobis distance is used to measure the similarity between a query image and a database (DB) image. Experimental results for Corel DB and VisTex DB show that the proposed method is superior to the conventional Gabor method. The proposed method also yields precision and recall 6.58% and 3.66% higher on average in Corel DB, respectively, and 4.87% and 3.37% higher on average in VisTex DB, respectively, than the popular RULBP method.

1. Introduction

The goal of an image retrieval system is to find images similar to a query image by using representative feature information of images in image databases. The application fields of image retrieval are digital libraries, VOD (video on demand) services, internet shopping malls, image content industry, etc. Image textures have visual patterns in the spatial domain, so they are considered as a significant factor in the areas related to image processing. Typical texture-based methods use brightness to represent characteristics of coarseness, contrast, or directionality and also use frequency transformation to represent energy characteristics.

Texture feature extraction methods are generally based on statistical, structural and spectral approaches. Statistical approaches are used for measuring statistical quantities based on image intensities. These include histograms [1], autocorrelations [2], statistical moments [3], block variance of local correlation coefficients (BVLC) [4], etc. Structural approaches use regularity of pixel patterns of images. These include autocorrelograms [5], gray level co-occurrence matrices (GLCM) [6], etc. Spectral approaches are usually used in the frequency domain. These include Fourier transformation [7], Gabor transformation [8], wavelet transformation [9], local binary patterns (LBP) [10,11], rotation-invariant uniform LBP (RULBP) [12], etc. The approaches are often used in combination with each other.

The Gabor wavelet used in the Gabor transformation was introduced by D. Gabor in 1946 and was proved by Daugman [13] in 1988. Daugman [13] showed that the receptive field profiles of visual cortex simple cells of a cat have 97 percent similarity to the 2-D Gabor wavelet function optimized for each neuron. This means the 2-D Gabor wavelet is similar to the response of the human visual system. The Gabor transformation is also known to be robust to variants of object size and illumination. These characteristics have received huge attention from many researchers.

Early image retrieval systems using the Gabor transformation have not performed as well as expected. In recent years, enhanced Gabor feature extraction methods have emerged and have been actively studied. Han and Ma [14] proposed rotation-invariant and scale-invariant Gabor representation. Fakheri et al. [15] used Gabor wavelets as local descriptors of texture, and gradient vector flow fields as shape information. Deshpande and Tadse [16] proposed the Gabor-Zernike feature extraction method where Gabor filters and Zernike moments are used for texture and shape features, respectively. Patil and Kumar [17] proposed the local gray Gabor pattern (LGGP) which is formed by comparison of pixels in LBP and local Gabor binary pattern (LGBP).

This paper presents an efficient process for combining texture features extracted using local energy and local correlation of Gabor transformed images, after which the extracted features are applied to an image retrieval system. Gabor energy features are often adopted in image processing but Gabor correlation features are not. The LGGP recently proposed in [17] uses the pattern of neighboring pixels, which is different from correlation. Therefore, our contribution is the combination of local energy and local correlation of Gabor transformed images.

The next section is described for the theoretical content of Gabor wavelets and Gabor transformations. The third section describes texture feature extraction methods proposed in this paper. The fourth section shows the experiment and results. Lastly, the fifth section ends with the conclusion.

2. Gabor Wavelets and Gabor Transformations

2.1 Gabor Wavelets

The 2-D Gabor wavelet [10] is a Gaussian envelope of a sine plane wave with specific frequency and direction. The function represents well the characteristics of the localized spatial frequency of the specific direction. The 2-D Gabor wavelet used in this paper is defined as follows:

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

where $j = \sqrt{-1}$ and W denotes the modulation frequency. The symbols σ_x and σ_y describe the dilation of the Gaussian function. The symbols x and y mean the spatial location of the filter.

A Gabor wavelet can be used to extract various frequency information using the set of Gabor wavelets with various resolutions and directions. A set of Gabor wavelets can be defined using $g(x,y)$ as follows:

$$g_{s,n}(x,y) = a^{-2 \cdot s} \cdot g(x',y') \tag{2}$$

where $x' = a^{-s} \cdot (x \cdot cos\theta_n + y \cdot sin\theta_n)$, $y' = a^{-s} \cdot (-x \cdot sin\theta_n + y \cdot cos\theta_n)$, $a > 1$ and $\theta_n = n \cdot \pi/K$, for $s = 0, 1, ..., S - 1$ and $n = 0, 1, ..., K - 1$. The symbol S is the number of all scales and K is the number of all directions. The symbols a, σ_x and σ_y of (1) and (2) are given as the following [14]:

$$a = \left(\frac{U_h}{U_l}\right)^{\frac{1}{S-1}}, \sigma_x = \frac{1}{2 \cdot \pi \cdot \sigma_u}, \text{ and } \sigma_y = \frac{1}{2 \cdot \pi \cdot \sigma_v} \tag{3}$$

where U_l and $U_h(= W)$ denote the lower and upper center frequencies, respectively. σ_u and σ_v are the variances in frequency domain corresponding to the spatial variances σ_x and σ_y, which are determined by

$$\sigma_u = \frac{(a-1) \cdot U_h}{(a+1) \cdot \sqrt{2 \cdot ln2}} \text{ and } \sigma_v = tan\left(\frac{\pi}{2 \cdot K}\right) \cdot \left(\frac{U_h^2}{2 \cdot ln2} - \sigma_u{}^2\right)^{\frac{1}{2}} \tag{4}$$

2.2 Gabor Transformations

Spatial frequency components provide important information, which is not gained directly from pixels. The Gabor transformation greatly represents energy characteristics of frequency of local region. It is not affected by variants of object size and illumination.

The 2-D Gabor transformation is defined as the convolution of an input image and the 2-D Gabor wavelet kernel. For the given image $I(x,y)$ and the 2-D Gabor wavelet with scale parameter s and direction parameter n, the general operation is defined by

$$J_{s,n}(x,y) = \sum_{x_1}\sum_{y_1} I(x_1,y_1) \cdot g_{s,n}(x - x_1, y - y_1) \tag{5}$$

where $J_{s,n}$ is the output image made by the convolution of the input image I and the 2-D Gabor wavelet kernel $g_{s,n}$. This operation can generate s × n images of the same size as input image I.

The experiment described in this paper uses sixteen Gabor wavelets with four scales and four directions of a half circle. In implementation of Gabor kernels, DC biases by finite supports are eliminated. Fig. 1 shows the magnitude parts of the complex Gabor representations for a flower image of 400×400 pixels. The magnitude images with higher resolutions of scales 1 and 2 are delicately displayed in the center details against the images with lower resolution of scales 3 and 4, whereas the latter with lower resolution are smoothly represented in the entire image.
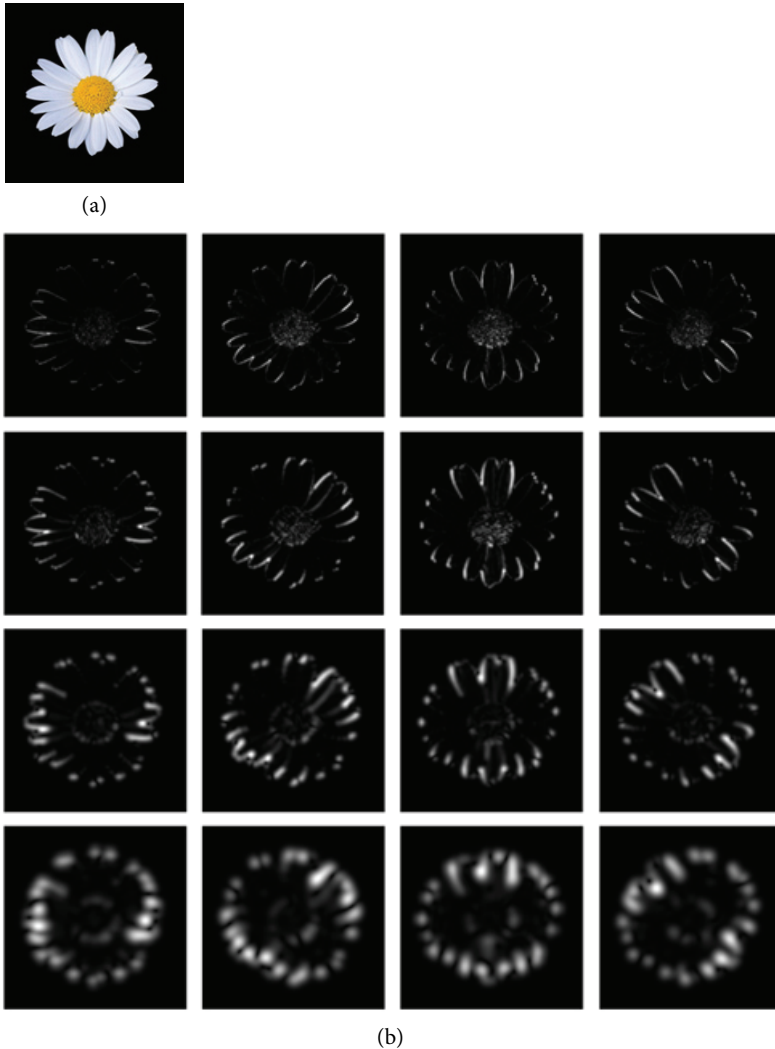
(a)

(b)

Fig. 1. Gabor representations: (a) original flower image of 400×400 pixels and (b) output images by the set of the Gabor wavelets with 4 scales and 4 directions (rows: scales 1, 2, 3 and 4; columns: directions 0°, 45°, 90° and 135°).

3. The Proposed Texture Feature Extraction Method

The block diagram of the proposed method is shown in Fig. 2, which has six steps.

Step 1: transform a query image to a gray image.
Step 2: conduct the Gabor transformation for the gray image.
Step 3: extract local energy and local correlation features from the Gabor transformed image.
Step 4: form the feature vector of the extracted features.
Step 5: measure the similarities between the combined feature vector of the query image and each of the feature vectors stored already in the DB.
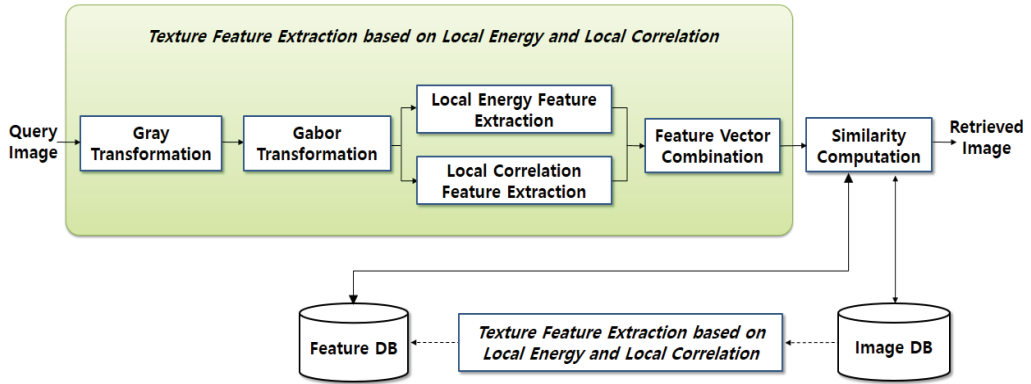Step 6: output the retrieved images from the DB.

Fig. 2. The block diagram of the proposed image retrieval system.

3.1 Texture Feature Extraction Method Using Local Energy and Local Correlation of Gabor Transformed Images

A feature extraction method using local energy and local correlation of Garber transformed images is proposed to enhance the feature extraction method using the Gabor transformation. The Gabor transformation has characteristics representing excellently the band-pass frequency components of each scale, which include detailed information of the image in scales. Thereby local energy and local correlation of Gabor transformed images can be effectively used to retrieve similar texture patterns.

1) Correlation

The correlation coefficient is the covariance between two random variables scaled by the product of their standard deviations. It is used to compute the degree of similarity of two variables, which is defined as [18],

$$\rho_{x,y} = \frac{C_{x,y}}{\sigma_x \cdot \sigma_y} \tag{6}$$

$$C_{x,y} = E\big[(X - u_x) \cdot (Y - u_y)\big] \tag{7}$$

where u_x and u_y indicate the means of X and Y, respectively. $C_{x,y}$ indicates the covariance of X and Y. The symbols σ_x and σ_y indicate the standard deviations of X and Y, respectively. The range of the correlation coefficient is $-1 \leq \rho_{x,y} \leq 1$.

This paper will use the formula of correlation coefficients with directional lags.

2) Procedure of texture feature extraction method

The Gabor wavelet chosen in this paper has 4 scales and 4 directions of the half circle, and support regions of ±2.6 sigma range. However, if scale 1 has support regions smaller than 5×5, the regions are enlarged to 5×5.

The detailed procedure of the proposed texture feature extraction method is as follows:

Step 1: Transform a query image to a gray image.

Step 2: Make the Gabor kernel using formula (1) with scale s and direction n.

Step 3: Conduct the Gabor transformation using the Gabor kernel for the gray image derived in step 1. The Gabor transformation is represented in expression (5).

Step 4: Make the magnitude image from the Gabor transformed complex image derived in step 3. In the complex plane, magnitude value r of complex number $z = a + bi$ and the magnitude image are represented as

$$r = |z| = \sqrt{a^2 + b^2} \tag{8}$$

$$R_{s,n}(p) = \left| J_{s,n}(p) \right| \tag{9}$$

where p indicates the pixel position and $|\cdot|$ means the magnitude.

Step 5: Calculate the global average $\mu_{s,n}^R$ and the global standard deviation $\sigma_{s,n}^R$ for the magnitude image using

$$\mu_{s,n}^R = \underset{p \in P}{\text{mean}} \left[R_{s,n}(p) \right] \tag{10}$$

$$\sigma_{s,n}^R = \underset{p \in P}{std} \left[R_{s,n}(p) \right] \tag{11}$$

where the operation std means computing standard deviation and P the set of all pixels.

Step 6: Make the correlation coefficient image from the Gabor transformed complex image derived in step 3. The correlation coefficient image is calculated as follows:

$$\rho_{s,n}(p) = \text{Re}\left\{ \underset{t \in T}{cor} \left[J_{s,n}(p - t), J_{s,n}^*(p - t - \delta\theta_n) \right] \right\} \tag{12}$$

where the operator cor means computing the correlation coefficient. The symbol Re denotes the real part of a complex number. The vector $\delta\theta_n$ is $(cos\theta_n, sin\theta_n)$ and T is the cross region of the 3×3 window. $J_{s,n}$ refers to the Gabor image with scale s and direction n from expression (5). $J_{s,n}^*$ refers to the complex-conjugation of $J_{s,n}$. $\rho_{s,n}(p)$ is the correlation coefficient between the cross region whose center pixel is p and the cross region centered at $p + \delta\theta_n$. Bilinear interpolation is used in the shift by $\delta\theta_n$.

Step 7: Calculate the global average $\mu_{s,n}^\rho$ and the global standard deviation $\sigma_{s,n}^\rho$ for the real part from the complex correlation coefficient image.

$$\mu_{s,n}^\rho = \underset{p \in P}{mean} \left[\rho_{s,n}(p) \right] \tag{13}$$

$$\sigma_{s,n}^\rho = \underset{p \in P}{std} \left[\rho_{s,n}(p) \right] \tag{14}$$

After step 7, constitute the final feature vector by repeating steps 2–7 for all desired scales and directions. The feature vector is described as $f = \left[[\mu_{s,n}^R], [\sigma_{s,n}^R], [\mu_{s,n}^\rho], [\sigma_{s,n}^\rho] \right]$ where $[\mu_{s,n}^R]$ means the vector composed of the global averages for local energy with scale s and direction n. Superscript R and ρ mean local energy and local correlation, respectively.

4. The Experiment and Results

This section demonstrates how the proposed method is superior in image retrieval to the conventional methods. As test DBs, Corel DB [19] and VisTex DB [20] are chosen which are used in most studies of texture image retrieval. Corel DB consists of various images with artificial objects. It contains 11 groups, each of which has 90 images of 192×128 resolution, totaling 990 images. VisTex DB consists of homogeneous pattern images. It includes 75 groups, each of which has 16 images of 128×128 resolution, totaling 1,200 images.

The experiment has three phases: feature extraction, image retrieval, and performance evaluation. The similarity between the feature vector of a query image and the feature vector of a DB image is measured by the Mahalanobis distance expressed as

$$D(f^q, \ f^d) = \left(\sum_{i=1}^{n} \left| \frac{f_i^q - f_i^d}{\sigma_i} \right|^M \right)^{\frac{1}{M}}$$

(15)

where q means a query image and d means a DB image. The symbol n means the feature vector dimension, M the metric order, f_i^q is the ith component of the feature vector of q, and f_i^d is the ith component of the feature vector of d. In the Mahalanobis distance, each component of a feature vector is divided by the standard deviation of the same components of the feature vectors to normalize feature vectors. Thus, the measurement is not affected by specific components [21].

The performance evaluation step utilizes the precision vs recall widely used in image retrieval [22], which is defined as

$$precision = \frac{|A(q) \cap R(q)|}{|A(q)|}, \qquad recall = \frac{|A(q) \cap R(q)|}{|R(q)|}$$

(16)

where $|\cdot|$ means the size of the set, q the query image, $A(q)$ the set of retrieved images for q, and $R(q)$ the set of relevant images for q.

The feature vector dimensions of various retrieval methods are shown in Table 1. As shown, our proposed feature of 64 dimensions is obtained by combining the conventional Gabor energy feature of 32 dimensions and our correlation feature of 32 dimensions.

Table 1. Feature vector dimensions of various retrieval methods

Method	Gabor	Proposed (correlation only)	RULBP	Proposed (Gabor+correlation)
Dimensions	32	32	88	64

In Fig. 3, the left graph shows precision vs recall for Corel DB. Each line indicates the performance at 10, 30, 50, 70, and 89 retrieved images, respectively.

Compared to the Gabor method, the proposed method using correlation only, and the RULBP method, the proposed method improves average precision by 10.6%, 3.5%, and 6.58%, respectively and improves average recall by 5.52%, 1.86% and 3.66%, respectively.

In Fig. 3, the right graph shows precision vs recall for VisTex DB. Each line indicates the performance at 5, 10, and 15 retrieved images, respectively.

Compared to the Gabor method, the proposed method using correlation only, and the RULBP method, the proposed method improves average precision 8.87%, 2.4%, and 4.87%, respectively and improves average recall by 6%, 1.67%, and 3.37%, respectively.

From the above results, it is shown that the proposed method using local energy and local correlation of Gabor transformed images yields improved retrieval performance over the conventional Gabor method, the proposed method using correlation only, and the popular RULBP method for the two DBs.
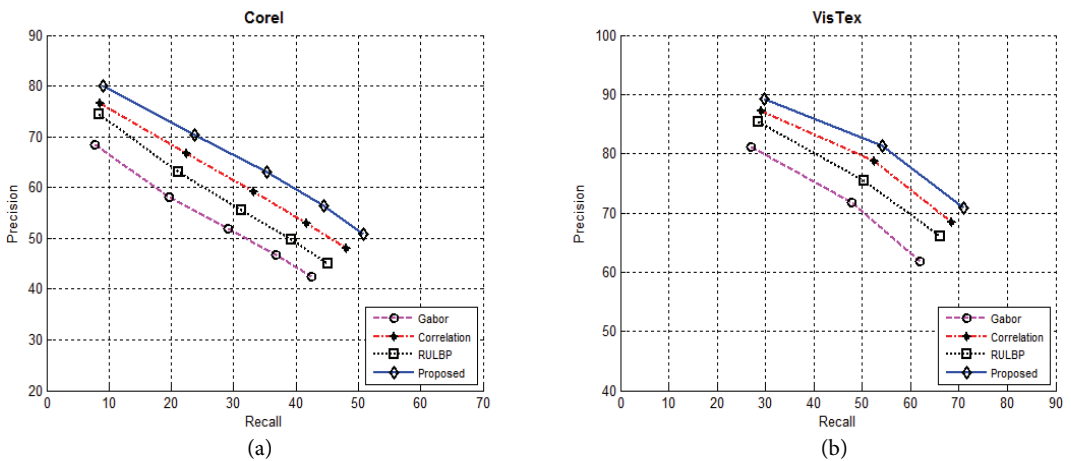

Fig. 3. (a) Precision vs recall for Corel database and (b) precision vs recall for VisTex database.

5. Conclusions

In this paper, the feature extraction method using local energy and local correlation of Gabor transformed images is proposed. The entire process applied to an image retrieval system is also presented. The local correlation to utilize relations between texture patterns of local regions on Gabor transformed images is employed. It is effectively used with the local energy to retrieve similar texture images.

It is confirmed that the proposed method improves results in the aspect of precision vs recall over the conventional Gabor method, the proposed method using correlation only, and the RULBP method in the fourth section. Consequently, the proposed method led to advances in the retrieval texture image system. Future research should explore the further utilization of useful characteristics of Gabor wavelets.

Acknowledgement

References

[1] E. Hadjidemetriou, M. D. Grossberg, and S. K. Nayar, "Multiresolution histograms and their use for texture classification," in *Proceedings of the 3rd International Workshop on Texture Analysis and Synthesis*, Nice, France, 2003, pp. 41-46.

[2] Y. Zhang, K. Qin, C. Zeng, E. B. Zhang, M. X. Yue, and X. Tong, "A data field method for urban remotely sensed imagery classification considering spatial correlation," in *Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, ISPRS Congress*, Prague, Czech Republic, 2016, pp. 431-435.

[3] E. F. Salma, E. H. Mohammed, R. Mohamed, and M. Mohamed, "A hybrid feature extraction for satellite image segmentation using statistical global and local feature," in *Proceedings of the Mediterranean Conference on Information & Communication Technologies, Lecture Notes in Electrical Engineering*. Cham, Switzerland: Springer, 2015, pp. 247-255.

[4] Y. D. Chun, S. Y. Seo, and N. C. Kim, "Image retrieval using BDIP and BVLC moments," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 9, pp. 951-957, 2003.

[5] H. A. Moghaddam, T. T. Khajoie, and A. H. Rouhi, "A new algorithm for image indexing and retrieval using wavelet correlogram," in *Proceedings of the International Conference on Image Processing*, Barcelona, Spain, 2003, pp. 497-500.

[6] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, 1973.

[7] S. Liao and A. C. S. Chung, "A new subspace learning method in Fourier domain for texture classification," in *Proceedings of 17th IEEE International Conference on Image Processing (ICIP)*, Hong Kong, China, 2010, pp. 4589-4592.

[8] N. Jain and S. S. Salankar, "Color & texture feature extraction for content based image retrieval," in *Proceedings of the International Conference on Advances in Engineering & Technology*, Singapore, 2014, pp. 53-58.

[9] J. Agarwal and S. S. Bedi, "Implementation of hybrid image fusion technique for feature enhancement in medical diagnosis," *Human-centric Computing and Information Sciences*, vol. 5, article no. 3, 2015.

[10] L. K. Rao and D. V. Rao, "Local quantized extrema patterns for content-based natural and texture image retrieval," *Human-centric Computing and Information Sciences*, vol. 5, article no. 26, 2015.

[11] S. K. Vipparthi and S. K. Nagar, "Color directional local quinary patterns for content based indexing and retrieval," *Human-centric Computing and Information Sciences*, vol. 4, article no. 6, 2014.

[12] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.

[13] J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 7, pp. 1169-1179, 1988.

[14] J. Han and K. K. Ma, "Rotation-invariant and scale-invariant Gabor features for texture image retrieval," *Image and Vision Computing*, vol. 25, no. 9, pp. 1474-1481, 2007.

[15] M. Fakheri, M. C. Amirani, and T. Sedghi, "Gabor wavelets and GVF functions for feature extraction in efficient content based colour and texture images retrieval," in *Proceedings of the 7th Iranian Conference on Machine Vision and Image Processing (MVIP)*, Tehran, Iran, 2011, pp.1-5.

[16] A. Deshpande and S. K. Tadse, "Design approach for content-based image retrieval using Gabor-Zernike features," *International Archive of Applied Sciences and Technology*, vol. 3, no. 2, pp. 42-46, 2012.

[17] J. K. Patil and R. Kumar, "Comparative analysis of content based image retrieval using texture features for plant leaf diseases," *International Journal of Applied Engineering Research*, vol. 11, no. 9, pp. 6244-6249, 2016.

[18] Correlation and Dependence [Online]. Available: https://en.wikipedia.org/wiki/Correlation_and_dependence.

[19] Corel Corporation, *The Corel Stock Photo Library*. Ottawa, Canada: Corel Corp., 1994.

[20] R. W. Picard, C. Graszyk, S. Mann, J. Wachman, L. Picard, and L. Campbell, *VisTex Benchmark Database of Color Textured Images*. Cambridge, MA: MIT Media Lab, 1995.

[21] W. Y. Ma and B. S. Manjunath, "A comparison of wavelet transform features for texture image annotation," in *Proceedings of the International Conference on Image Processing*, Washington, DC, 1995, pp. 256-259.

[22] D. Comaniciu, P. Meer, K. Xu, and D. Tyler, "Retrieval performance improvement through low rank corrections," in *Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries, 1999 (CBAIVL '99)*, Fort Collins, CO, 1999, pp. 50-54.

Hee-Hyung Bu

She received B.S., M.S. and Ph.D. degrees in Computer Engineering from Mokpo National University, Jeonnam, Chonnam National University, Gwangju and Kyungpook National University, Daegu, Korea, in 2004, 2006 and 2013, respectively. She was a research professor in the Department of Computer Engineering at Mokpo National University in 2013. Her research interests include image retrieval, video compression and image processing.

Nam-Chul Kim

He received B.S. degree in Electronic Engineering from Seoul National University, in 1978, and M.S. and Ph.D. degrees in Electrical Engineering from Korea Advanced Institute of Science and Technology (KAIST), Seoul, Korea, in 1980 and 1984, respectively. Since March 1984, he has been with the School of Electronics Engineering at Kyungpook National University, Daegu, Korea, where he is currently a full professor. During 1991–1992, he was a visiting scholar in the Department of Electrical and Computer Engineering, Syracuse University, Syracuse, NY. His research interests are image processing and computer vision, biomedical image processing, and image and video coding.

Bae-Ho Lee

He received his Ph.D. degree in 1993 from the University of Missouri in Columbia, MO, USA. Since 1993, he has been a professor in electronics and computer engineering department at Chonnam National University, Gwangju, Korea. His research includes image processing, computer vision and artificial intelligence.

Sung-Ho Kim

He received his B.S. degree in Electronics from Kyungpook National University, Korea in 1981, and his M.S. and Ph.D. degrees in Computer Science from Korea Advanced Institute of Science and Technology (KAIST), Korea in 1983 and 1994, respectively. He has been a faculty member of the School of Computer Science & Engineering at Kyungpook National University since 1986. His research interests include real-time image processing and telecommunication, multi-media systems, etc.