A Rotation Invariant Image Retrieval with Local Features

Hee-Jun You, Dae-Kyu Shin, Dong-Hoon Kim, Hyun-Sool Kim, and Sang-Hui Park

Abstract: Content-based image retrieval is the research of images from database, that are visually similar to given image examples. Gabor functions and Gabor filters are regarded as excellent methods for feature extraction and texture segmentation. However, they have a disadvantage not to perform well in case of a rotated image because of its direction-oriented filter. This paper proposes a method of extracting local texture features from blocks with central interest points detected in an image and a rotation invariant Gabor wavelet filter. We also propose a method of comparing pattern histograms of features classified by VQ (Vector Quantization) among images.

Keywords: Image retrieval, texture, rotation invariant Gabor filter, VQ, interest points.

1. INTRODUCTION

Through the recent emergence of Internet, it is possible to access various data scattered throughout the world. As such, the need for a system that can efficiently search for specific data is on the increase. In particular, a large number of multimedia database and image libraries are being constructed because of the development of large capacity storage equipment used for various purposes. Above all, image data have a particular advantage that the informational capacity is the best.

The image retrieval technique is largely classified into context-based image retrieval and content-based image retrieval. Recently, the efficient and automatic image retrieval by image content such as shape, color or texture etc. is emerging as an important research area with application [1,2].

This paper proposes the use of texture as an image feature for pattern retrieval. These texture features are not extracted from the global features of an image but from the local features calculated at blocks with 17 by 17 pixels whose center is automatically detected as interest points. These features are usually seen as elements of a feature vector. A multiresolution filtering technique based on Gabor filters for texture analysis has been shown to be optimal in the sense of minimizing the joint two-dimensional uncertainty in space and frequency [3]. However, Gabor filter bank has some problems related to image rotation. Therefore, the proposed rotation invariant Gabor filter [4] is used in this paper.

VQ and histogram methods are used for searching image. The feature vectors are classified by VQ and form a histogram of constructed codebook index. Then, image retrievals are performed by comparing features between database images and a query image.

The performance of the proposed algorithm is evaluated through application in images with rotation, small scale change, noise addition, and partial image.

2. FEATURE EXTRACTION

The existing researches have mainly used the global features of an image for image retrieval. The approach cannot represent features such as edges or regions, particularly robustness to partial visibility and high information content [5]. However, the uses of local information calculated at automatically detected interest points, which correspond to the corners, make up for these defects.

A common definition of texture is the repetition of basic texture elements, which have properties such as frequency, direction, phase etc. The Gabor filters are suitable for extracting these features. They have already been used for texture-based image retrieval. However, they are somewhat inappropriate for application to the retrieval of the rotated image because they depend on one direction.

In this paper, we will use local information from blocks whose centers are the detected interest points and the rotation invariant Gabor filter to extract texture features.

2.1. Interest points

Interest points corresponding to the corners known

Manuscript received December 2, 2002; accepted April 23, 2003.

Hee-Jun You and Sang-Hui Park are with the Department Electrical and Electronic Engineering, Yonsei University, Korea (e-mail: { hjy75, psh} @yonsei.ac. kr).

Dae-Kyu Shin is with the Telecommunication R&D Center, Samsung Electronics Co., LTD., Korea (e-mail : daekyu.shim @samsung.com)

Dong-Hoon Kim and Hyun-Sool Kim are with the Mobile Communication Division, Samsung Electronics Co., LTD., Korea (e-mail: {donghoon74. kim, hyunsool}@samsung.com).

as T-Junctions are usually defined as the points where the signal varies multiple-dimensionally, and also the points where the texture changes rapidly. Therefore, these detectors are based on local derivatives [5,6]. Approaches for detecting stable interest points have been widely researched. The detector of Harris and Stephens [6] is known as the most efficient method considering repeatability and information content. This detector uses the auto-correlation function in order to determine locations where the signal changes in two directions and the eigenvectors of a matrix related to the auto-correlation function, which takes into account first derivatives of the signal on a window. The auto-correlation matrix is

$$M = e^{-\frac{x^2+y^2}{2\sigma^2}} \otimes \begin{bmatrix} I_x{}^2 & I_x I_y \\ I_x I_y & I_y{}^2 \end{bmatrix}, \qquad (1)$$

where I_x and I_y are computed by convolution the image Gaussian derivatives. The eigenvalues (λ_1, λ_2) of the matrix M are the two principle curvatures. The interest point is extracted when they are higher than a given threshold.

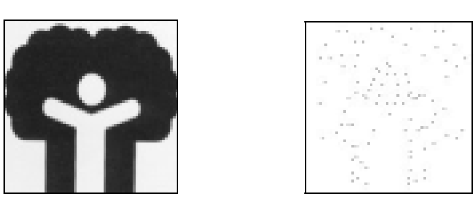
An example of extracting interest points in a trademark image is shown in Fig. 1.

2.2. Gabor filter

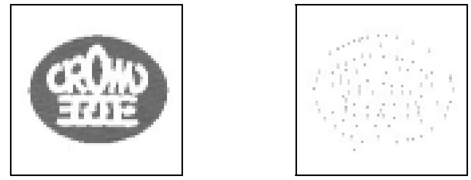
The filters of Gabor filter bank are designed to detect different frequencies and orientations. The Gabor filter takes the form of a two dimension Gaussian modulated complex sinusoidal grating in the spatial domain [7]. It can be written as

$$h(x,y) = g(x^{'}, y^{'}) \cdot e^{-2\pi j(Ux+Vy)}, \qquad (2)$$

where (U,V) defines the position of the filter in the Fourier domain with a centre frequency of $F = \sqrt{U^2 + V^2}$ and an orientation of $\theta = \arctan(\frac{V}{U})$. The term $g(x^{'}, y^{'})$ represents a Gaussian function oriented at an angle θ, where $(x^{'}, y^{'})$ are the rotated co-ordinates given by $x^{'} = x\cos\theta + y\sin\theta$ and $y^{'} = -x\sin\theta + y\cos\theta$. The general form of the Gaussian function is shown in the following.



(a) (b)



(c) (d)

Fig. 1. The interest points extracted from a trademark image: (a) image 1, (b) interest points in (a), (c) image 2, and (d) interest points in (c).

$$g(x,y) = (\frac{1}{2\pi\lambda\sigma^2}) \cdot e^{-\left[\frac{(x/\lambda^2)+y^2}{2\sigma^2}\right]}, \qquad (3)$$

where λ defines the aspect ratio and σ the scale factor. The scale factor is typically determined by the centre frequency of the filter such that high frequency filters are more localized in space.

$$\sigma = \mu / F, \qquad (4)$$

where μ is a constant. A fixed set of filters are usually chosen to generate features for texture classification, centered at the required frequencies and orientations to obtain the optimum coverage of the Fourier domain.

2.3. Rotation invariant gabor filter

The sinusoidal grating of the Gabor filter varies in only one direction thus making it highly orientation specific. Therefore, the filter is very effective at orientation dependent texture analysis, but it is not suitable for rotation invariant texture classification. In order for the filter to be rotation invariant, it is necessary that the sinusoidal grating vary in all directions. Hence, the filter is made circularly symmetric to achieve rotation invariance. This filter is again formed by modulating a complex sinusoidal grating with a Gaussian function. However, both the Gaussian and the grating vary only in a radial direction from the origin, such that the filter is completely circularly symmetric. The circularly symmetric Gabor filter can be written as

$$h(x,y) = g(x,y) \cdot e^{-2\pi jF\sqrt{x^2+y^2}}, \qquad (5)$$

where F is the required centre frequency. In this case, it is unnecessary to rotate the co-ordinates of the Gaussian since it is circularly symmetric. The Gaussian can be written as

$$g(x,y) = \left(\frac{1}{2\pi\sigma^2}\right) \cdot e^{-\left[\frac{x^2+y^2}{2\sigma^2}\right]}. \qquad (6)$$

To extract texture-based features from an image using the circularly symmetric filters, four filters are used, spaced in geometric progression across the Fourier domain to achieve optimum coverage. μ has been chosen such that the filters overlap slightly and Fourier domain is as evenly covered as possible. However, the very low frequencies are not covered since these convey little textural information about the image that has only gradual changes in grey level. An example of a circularly symmetric Gabor filter is shown in Fig. 2.

2.4. The representation of feature vectors

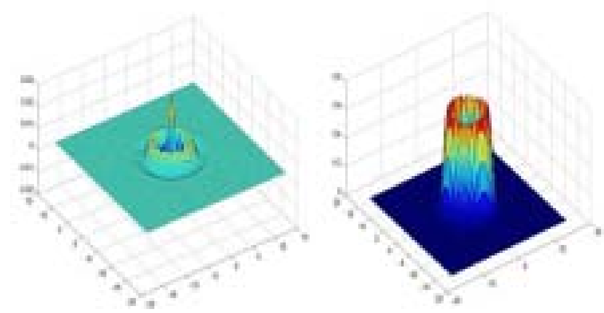
To represent feature vectors, the average (μ_m) of each four-feature vector of pixels ($h_m(x, y)$) in a block with 17 by 17 pixels, which are filtered by four circularly symmetric filters and selected by a criterion that the difference of feature vectors between the interest point and another point in a block is lower than a given threshold, is used. The center of a block is the interest point automatically detected in an image. Therefore, a block consists of all four average feature vectors.

$$\mu_m = \frac{1}{N_x N_y} \sum_x \sum_y |h_m(x, y)|, \qquad (7)$$

where m is the scale of the central frequency. N_x and N_y are the width and height of pixels selected in a block, respectively.

3. SIMILARITY SEARCHING ALGORITHM

In order to be able to compare features of images between database and query, here we, motivate and describe the implementation of a simple VQ-based histogram method. In the preparation stage, the average of the feature vectors is calculated from blocks whose centers are the interest points of images in both query and database. They are then quantized using a LBG algorithm [8] of VQ and represented by the index of the codebook, which efficiently describes all the feature vectors in the database. All images are represented by the histogram describing texture information. Finally, the similarity measure between database and query images is performed by using the histogram intersection method used by Swain in color-based image retrieval [9]. It compares describes all the feature vectors in the database. All



(a) (b)

Fig. 2. A circularly symmetric Gabor filter in the spatial and Fourier domain: (a) spatial domain, (b) fourier domain.

describes all the feature vectors in the database. All images are represented by the histogram describing texture information. Finally, the similarity measure between database and query images is performed by using the histogram intersection method used by Swain in color-based image retrieval [9]. It compares histograms between query image (H_q) and database images (H_d) using the following equation.

$$H(H_q, H_d) = \frac{\sum_{j=1}^{n} \min(H_q(j), H_d(j))}{\sum_{j=1}^{n} H_q(j)}, \qquad (8)$$

where j, which represents the index of the codebook, has ranged from 1 to n. The intersection is normalized by the number of pixels in the histogram of a query image. The larger similarity between two images is the closer value to one this has.

4. EXPERIMENTAL RESULTS

To demonstrate the effectiveness of our proposed method for texture image retrieval, we have used 1100 gray-valued trademark images of 128×128 pixels for the database. The images used for the query images consist of twenty of the images in the database.

In these experiments, all query images, those changed and those not changed, were used to search the similar images stored in the database. To evaluate performance, we computed the average retrieval rank of the desired image as in 9 and compared the performance of the proposed method with those of the existing methods such as Zernike moments (ZM) [10], differential invariants (DI) [5] and Jain's one using gradient [11].

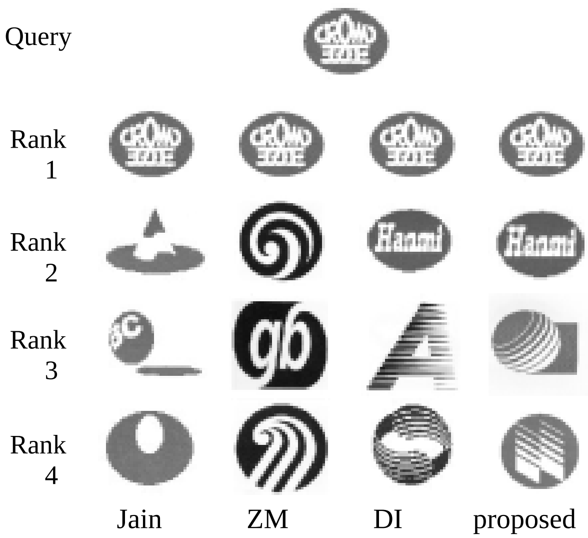


Fig. 3 The result of retrieval test for the original query image.

$$Average \quad retrieval \ rank \quad = \frac{1}{M}\sum_{i=1}^{M} rank[i],$$

$$\% \ rank = \frac{\text{average retrieval rank}}{N} \times 100, \qquad (9)$$

where each M and N is the total number of queries and database images.

Fig. 3 shows the result of retrieval test for query image, which was unchanged. We observe that all the proposed and existing methods correspond to the query image.

In order to experiment with the retrieval of changed images, we converted the queries into the images changed by noise, scale and rotation. For noise addition, we used the queries with approximately 5, 10 and 15 percent of the pixels corrupted by white Gaussian noise. For scale change, the images are changed by 0.9 and 1.1 times. For the queries for image rotation, the images are rotated by 5, 10, 15, 45 and 90 degrees. The codebook size was 100 in VQ procedure. Fig. 4 shows the example of changed images for query.

The results of these experiments were summarized in Table 1.

It indicates that proposed method performs better than existing methods in given situations. In particular, we observe that our proposed method shows a considerable improvement in the rotation change.

Fig. 5 shows the result of the retrieval test for a query image translated to cut off some part of an image. The result indicates that in retrieval of the partial image, the elements of the histogram bin are reduced because of the loss of part of the image, but the global shape of its histogram is maintained and only a few points are necessary to recognize an image.

We applied this proposed method to retrieving natural images. Our database for natural images was composed of 1100 grayscale images with image size of 256-by-256 pixels. Fig. 6(a) and Fig. 7(a) show examples of query images of a tiger and leopard, respectively. The results are shown in Fig. 6(b) and Fig. 7(b).

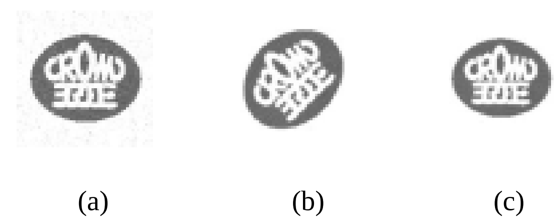


Fig. 4. The example of changed images for query: (a) image corrupted by 15% noise, (b) image rotated by 45°, (c) image changed by 0.9 times.

Table 1. Results of image retrieval tested on trademark images.

Experimental conditions		Jain	ZM	DI	Proposed
Noise	5 %	2.1 (0.19%)	1.3 (0.12%)	3.6 (0.32%)	1.5 (0.14%)
	10 %	2.6 (0.23%)	10.6 (0.95%)	4.3 (0.39%)	1.9 (0.17%)
	15 %	4.2 (0.38%)	41.4 (3.73%)	13.0 (1.17%)	2.2 (0.20%)
Rotation	5 °	Not applicable	3.26 (0.30%)	9.3 (0.85%)	1.3 (0.11%)
	10 °		6.11 (0.56%)	4.26 (0.39%)	1.4 (0.13%)
	15 °		5.67 (0.52%)	7.95 (0.72%)	2.0 (0.18%)
	45 °		13.3 (1.20%)	41.8 (3.77%)	2.1 (0.20%)
	90 °		1.0 (0.09%)	1.2 (0.11%)	1.0 (0.09%)
Scale	0.9	24.7 (2.23%)	22.8 (2.01%)	54.2 (4.89%)	2.3 (0.20%)
	1.1	13.7 (1.23%)	20.4 (1.84%)	57.6 (5.19)	1.4 (0.12%)

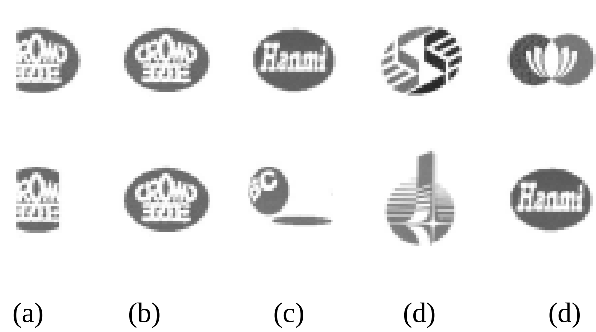


Fig. 5. Retrieval results for a partial image: (a) query, (b) rank 1, (c) rank 2, (d) rank 3, (e) rank 4.

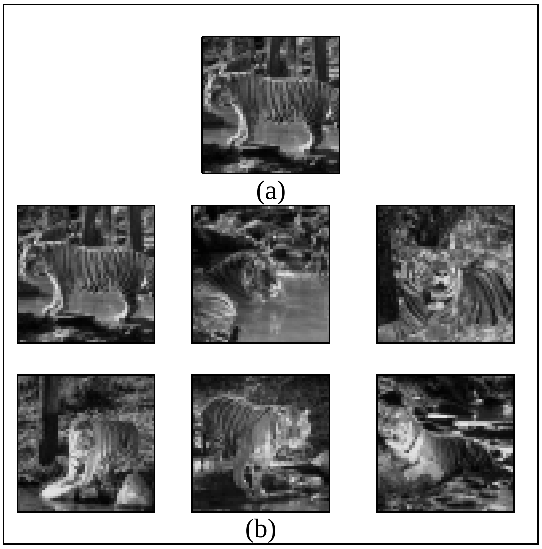


Fig. 6. The results of retrieving nature images: (a) shows the query image of a tiger, (b) shows the best matches ranked in the upper 2%.

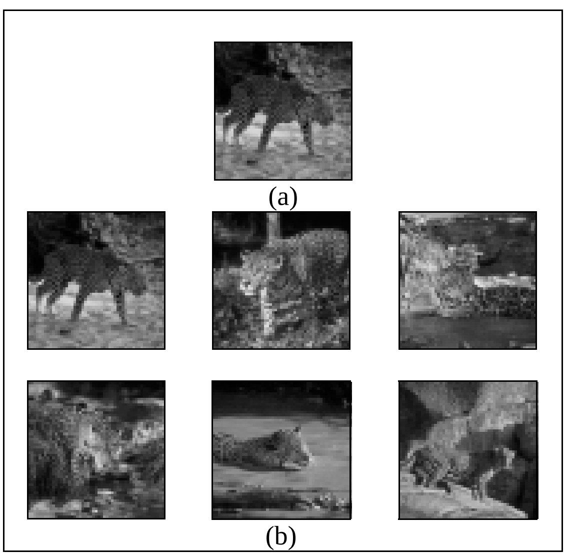


Fig. 7. The results of retrieving nature images: (a) shows the query image of a leopard, (b) shows the best matches ranked in the upper 2%.

Fig. 8(a) and Fig. 9(a) are query images that used a rectangular patch and the results are shown in Fig. 8(b) and Fig. 9(b).

In these experiments, the number of images, which are ranked in the upper 2%, is shown in Table 2.

In the experiment using the original query images, the various backgrounds in the query images but not in the experiment using partial query images affects the retrieved images. The backgrounds in the retrieval images were ignored by the proposed method that used local texture features to handle partial queries.

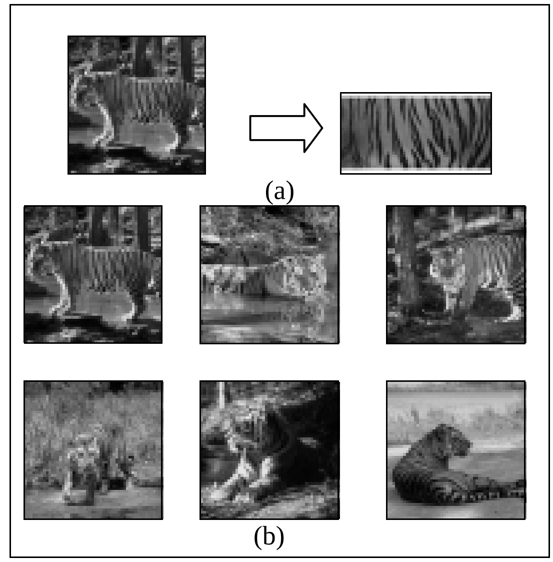


Fig. 8. The results of retrieving partial nature images: (a) shows the partial query image of a tiger, (b) shows the best matches ranked in the upper 2%.

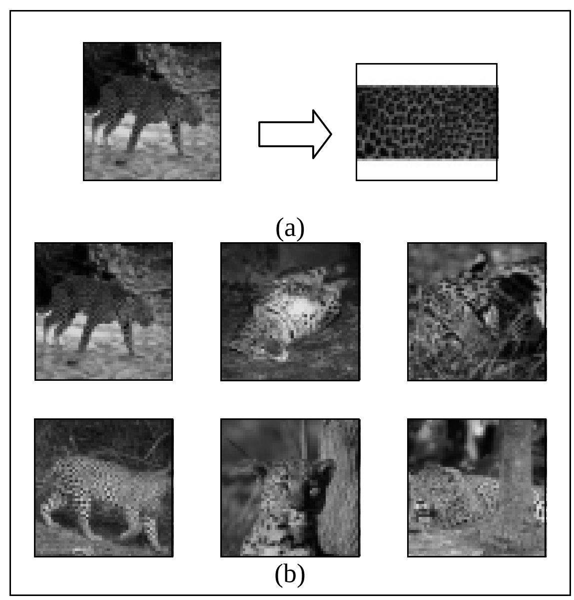


Fig. 9. The results of retrieving partial nature images: (a) shows the partial query images of a leopard, (b) shows the best matches ranked in the upper 2%.

Table 2. The number of images retrieved in the upper 2% (The number of images retrieved in the upper 2% / a total of tiger or leopard images in the database).

A query image	Original image	Partial Image
The tiger	9 / 30	17 / 30
The leopard	6 / 27	16 / 27

5. CONCLUSION

In this paper, we have proposed a new image retrieval method robust to rotation using VQ-based local texture information. By using a rotation invariant Gabor wavelet filter at interest points and VQ, the whole image feature has been extracted and quantized. In order to compare similarity, we have used the histogram intersection method.

The experimental results tested on trademark images show that this method performs better than the existing methods in case of noise addition and scale change. In particular, in case of rotation change, our proposed method shows considerable improvement. By using the local texture information even small parts of images can be correctly retrieved. Moreover, this proposed method shows the possibility to retrieve natural images.

REFERENCES

[1] J. Eakins and M. Graham, "Content-based image retrieval," *JISC Technology Applications Report: 39*, October 1999.

[2] C. C. Venters and M. Cooper, "A review of content-based image retrieval systems," *Manchester Visualization Centre*, JISC Technology Application, 2000.

[3] J. G. Daugman, "Complete discrete 2d gabor transforms by neural networks for image analysis and compression," *IEEE Trans. ASSP*, vol. 36, pp. 1, 169-1, 179, July 1988.

[4] R. Porter and N. Canagarajah, "Gabor filter for rotation invariant texture classification," *IEEE International Symposium on Circuits and Systems*, June 9-12 1997.

[5] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530-535, 1997.

[6] C. Harris and M. Stephens, "A combined corner and edge detector," *Proc. 4th Alvey Vision Conference*, pp. 147-151, 1988.

[7] A. C. Bovik, M. Clark, and W. S Geisler, "Multi-channel texture analysis using localized spatial filters," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 55-73, January 1990.

[8] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer deign," *IEEE Trans. on Communications*, vol. COM-28, no. 1, pp. 84-95, January 1980.

[9] M. J. Swain and D. H. Ballad, "Color indexing," *International J. Computer Vision*, vol. 7, no. 1, pp. 11-32, July 1991.

[10] A. Khotanzad and Y. H. Hong, "Invariant image recognition by zernike moments," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 489-497, 1990.

[11] A. K. Jain and A. Vailaya, "Image retrieval using color and shape," *Pattern Recognition*, vol. 29, no. 8, pp. 1233-1244, 1996.

[12] D. K. Shin, H. S. Kim, T. Y. Chung, T. S. Kim, and S. H. Park, "Image retrieval using VQ-based local Gabor feature," *Electronics Letters*, vol. 38, no. 11, pp. 505-507, 23 May 2002.

Hee-Jun You received the B.S degree in Electrical and Electronic Engineering from Myongji University, Korea in 2002. He is currently pursuing the M.S. degree in Electrical and Electronic Engineering at Yonsei University, Korea. His research interests are video compression standards, fast motion estimation and video transcoding.

Hyun-Sool Kim received the B.S., M.S. and Ph.D. degree in Electrical and Electronic Engineering from Yonsei University, Korea. He is currently with Mobile Communication Division, Samsung Electronics Co., LTD., Korea.

Dae-Kyu Shin received the B.S., M.S. and Ph.D. degree in Electrical and Electronic Engineering from Yonsei University, Korea. He is currently with Telecommunication R&D Center, Samsung Electronics Co., LTD., Korea.

Dong-Hoon Kim received B.S. degree in Electronic Engineering from Kyungwon University, Korea and the M.S. degree in Electrical and Electronic Engineering from Yonsei University, Korea, in 2001 and 2003, respectively. He is currently with Mobile Communication Division, Samsung Electronics Co., LTD., Korea.

Sang-Hui Park received the B.S., M.S. and Ph.D. degree in Electrical Engineering from Yonsei University, Korea in 1962, 1964 and 1971, respectively. He was at a Medical College, Washington University, Seattle as a Visiting Professor from 1982 to 1984. During 1982-1984, he served as the President of the Korean Institute of Electrical Engineers (KIEE), Korea. He is currently a Professor in the School of Electrical and Electronic Engineering, Yonsei University, Korea. Since 2002, He has been serving as the Dean of Graduate School of Information, Yonsei University, Korea.