

PAPER

A Comparative Study of Rotation Angle Estimation Methods Based on Complex Moments

Jong-Min LEE[†], *Student Member and* Whoi-Yul KIM^{†a)}, *Nonmember*

SUMMARY Determining the rotation angle between two images is essential when comparing images that may include rotational variation. While there are three representative methods that utilize the phases of Zernike moments (ZMs) to estimate rotation angles, very little work has been done to compare the performances of these methods. In this paper, we compare the performances of these three methods and propose a new, angular radial transform (ART)-based method. Our method extends Revaud et al.'s method [1] and uses the phase of angular radial transform coefficients instead of ZMs. We show that our proposed method outperforms the ZM-based method using the MPEG-7 shape dataset when computation times are compared or in terms of the root mean square error vs. *coverage*. **key words:** rotation angle, phase, Zernike moment

1. Introduction

In pattern recognition and image analysis, sets of Zernike moments (ZMs) are widely used as image descriptors to capture the global features of an image [2]–[13]. ZMs extend geometric movements using orthogonal Zernike polynomials, rather than the conventional transform kernels that geometric moments use. A single ZM is a complex number that encompasses two different values: magnitude and phase.

In order to compare two images that differ by rotational angle, an image descriptor should either be invariant to rotation or able to retrieve the rotation angle. Earlier work using ZMs to describe image patterns most commonly used the magnitude of the ZM because it is invariant to rotation. Recently, several studies have used the ZM phase [1], [14], [15], which is more informative than the magnitude [14]–[16]. Although phase information is not invariant to rotation, the ZM phase can be used to estimate the rotation angle between two image patterns. To calculate the similarity between two images using the ZM phase, this rotation angle must be identified. Apart from comparing two sets of ZM phases, the rotation angle between two images is also used to align images in many computer vision tasks.

A conventional method for estimating the rotation angle between two image patterns is to identify the principle axes, and several methods (i.e., [17] and [18]) that use edge information have been proposed. ZM-based methods, however, provide superior performance in terms of accuracy. Three representative studies have used the ZM phase to estimate the rotation angle between pairs of image patterns [1],

[15], [19]. The method introduced by Kim and Kim [19] calculates a probability density function of the rotation angle using the phase differences of coefficients of two set of ZMs in a discrete angle space. The maximum value of the density function is an estimate for the rotation angle. Revaud et al. presented a method that estimates the rotation angle by minimizing the distance function of two image patterns [1]. The distance is defined using the dissimilarity between reconstructed images using ZMs. Recently, Chen and Sun proposed using a weighted sum of the phase differences in two sets of ZMs [15]. The two phase differences are calculated using two coefficients of consecutive repetition parameters of ZMs.

Although all three methods were proposed for the same purpose and claimed to have the best performance, very little work has been done to compare these methods under equivalent conditions. Only a limited performance comparison between Revaud et al.'s method and, Kim and Kim's method was presented in [1]. Therefore, the question of which estimator is the most accurate and robust with respect to noise remains open.

In this paper, we compare performance in terms of accuracy and noise robustness, and propose a new rotation angle estimation method that uses the phase of coefficients of angular radial transform (ART) [20]. The proposed method is an extension of the work of Revaud et al. [1] and is shown to significantly outperform the three ZM-based methods. For the performance comparison, we employ root mean square (RMS) error vs. *coverage* [15] as criteria. The *coverage* contains information related to the variation of the estimation error. We also compare the computation times of each method, since computational efficiency is one of most important properties of an estimation method. The experimental dataset contains a sufficient number and variety of types of images to facilitate performance comparisons in terms of both accuracy and noise robustness. In order to compare the noise robustness, both photometric and geometric noise images are considered in the comparison.

We briefly review the fundamentals of ZM in Sect. 2 and all three methods in Sect. 3. In Sect. 4, the proposed method is presented with a brief review of the fundamentals of ART. The experimental data and the criteria for our performance comparison are then presented in Sect. 5. The results and analysis of the comparisons are given in Sect. 6, and we conclude the paper in Sect. 7.

Manuscript received September 30, 2011.

[†]The authors are with the Hanyang University, Seoul, 133–791, Korea

a) E-mail: wykim@vision.hanyang.ac.kr

DOI: 10.1587/transinf.E95.D.1485

2. Fundamentals of the Zernike Moment

The ZM coefficients represent complex moments calculated using different orthogonal basis functions. The basis functions are defined on a unit disk in polar coordinates. A ZM of order n with repetition m is defined as

$$Z_{nm} = \frac{n+1}{\pi} \iint_{\rho \leq 1} f(\rho, \theta) V_{nm}^{ZM*}(\rho, \theta) \rho d\rho d\theta, \quad (1)$$

where $f(\rho, \theta)$ is the input image function, $V_{nm}^{ZM}(\rho, \theta)$ is the complex Zernike basis function, and $*$ denotes the complex conjugate. The Zernike basis functions are visualized up to $n = 7$ in Fig. 1.

The Zernike basis functions are defined as

$$V_{nm}^{ZM}(\rho, \theta) = R_{nm}^{ZM}(\rho) \exp(jm\theta), \quad |\rho| \leq 1, \quad (2)$$

where

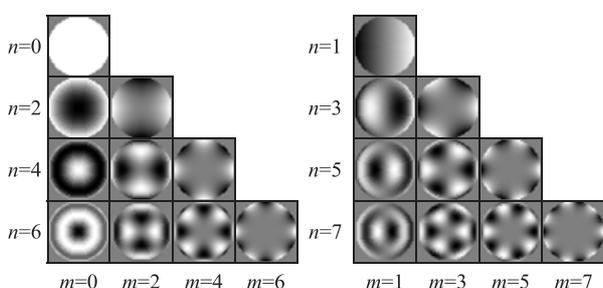
$$R_{nm}^{ZM}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s}. \quad (3)$$

In (3), the order n is a non-negative integer and the repetition m is an integer satisfying $n - |m| = (\text{even})$ and $|m| \leq n$.

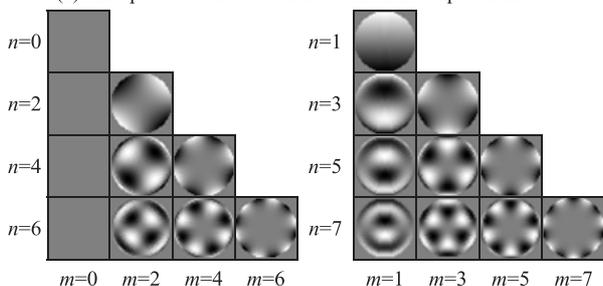
The difference between two ZM phases of an original image Z_{nm} and a rotated image Z_{nm}^r is given by

$$\Theta_{nm} \equiv \arg\left(\frac{Z_{nm}^r}{Z_{nm}}\right) = m\alpha, \quad 0 < \Theta_{nm} \leq 2m\pi, \quad (4)$$

or



(a) Real parts of the Zernike basis functions up to order 7.



(b) Imaginary parts of the Zernike basis functions up to order 7

Fig. 1 The basis functions of the ZM.

$$\begin{aligned} \Phi_{nm} &= (\varphi_{nm}^r - \varphi_{nm}) \bmod (2\pi) \\ &= (m\alpha) \bmod (2\pi), \quad 0 < \Phi_{nm} \leq 2\pi, \end{aligned} \quad (5)$$

where φ is the phase of the ZM. More specifics about the fundamentals and the properties of the magnitude and phase of ZMs are well presented in [9] and [14].

3. Methods for Estimating Rotational Angle Based on the ZM Phase

3.1 Kim's and Kim's Method

In [14], the rotation angle α is defined as

$$\begin{aligned} \Omega_{nm} &= (\varphi_{nm} + 2k_1\pi) - (\varphi_{nm}^r + 2k_2\pi) \\ &= \Phi_{nm} + 2k_{nm}\pi = m\alpha. \end{aligned} \quad (6)$$

here $k_m \in \{0, 1, 2, \dots, m-1\}$. The estimated rotation angle contains an error ε generated by the quantization effect of an image sensor. Hence (5) can be written as

$$\Theta_{nm} \equiv \arg\left(\frac{Z_{nm}^r}{Z_{nm}}\right) = m\alpha + \varepsilon, \quad (7)$$

and the rotation angle for the (n, m) order of the ZM $\hat{\alpha}_{nm}$ is

$$\hat{\alpha}_{nm} = \frac{\Theta_{nm}}{m}; \quad m \neq 0. \quad (8)$$

Since (8) may yield m solutions, a probabilistic model is used as in (9) to find the $\hat{\alpha}$ that minimizes the error $E[(\alpha - \hat{\alpha}_{nm})^2]$, i.e.,

$$P(\hat{\alpha}) = \sum_n \sum_m \xi_{nm} P(\hat{\alpha}|n, m), \quad 0 \leq \hat{\alpha} \leq 2\pi \quad (9)$$

where $P(\hat{\alpha}|n, m)$ is the probability density function of the rotation angle estimated from the ZM of order (n, m) and ξ_{nm} is the corresponding weighting factor calculated using the magnitude of the ZM. The unique solution is determined at the highest peak of the density function.

3.2 Revaud et al.'s Method

Revaud et al. [1] defined the similarity distance between the original and rotated images by using the difference between two images reconstructed by ZM as

$$\begin{aligned} d_{\text{original,rotated}}^2(\alpha) &= \sum_{x^2+y^2 \leq 1} \sum_{(n,m) \in D} \left| \sum_{(n,m) \in D} Z_{nm} \cdot V_{nm}(x, y) \right. \\ &\quad \left. - \sum_{(n,m) \in D} Z_{nm}^r \cdot e^{jm\alpha} \cdot V_{nm}(x, y) \right|^2. \end{aligned} \quad (10)$$

where, $D = \{(n, m) | 0 \leq n \leq \infty, |m| \leq n, |n - m| = \text{even}\}$.

They rewrote (10) into (11) to convert the equation to a function of the rotation angle.

$$\begin{aligned} d_{\text{original,rotated}}^2(\alpha) &= \sum_{x^2+y^2 \leq 1} \sum_{n+1}^{\pi} \left[|Z_{nm}|^2 + |Z_{nm}^r|^2 \right. \\ &\quad \left. - 2|Z_{nm}Z_{nm}^r| \cdot \cos(m\alpha + \varphi_{nm} - \varphi_{nm}^r) \right], \end{aligned} \quad (11)$$

where $-180 \leq \alpha \leq 180$. Then, they roughly searched for a minimum value of $d_{original,rotated}^2(\alpha)$ to find the rotated angle α . Equation (12) can be simplified by removing the constant terms and by aggregating cosine terms with the same frequency:

$$A_1 \cos(m\alpha + B_1) + A_2 \cos(m\alpha + B_2) = |\lambda| \cos(m\alpha + |\lambda|), \tag{12}$$

where λ is a complex that can be expressed as $A_1 e^{iB_1} + A_2 e^{iB_2}$. Then, (12) can be equivalently expressed as a sum of N cosines:

$$f_N(\phi) = \sum_{m=1}^N A_m \cos(m\alpha + B_m), \tag{13}$$

where $A_q \in \mathfrak{R}^+$ and $B_q \in [-\pi, \pi]$.

Equally spread $4N_{N-S}$ points are sampled between $[0, 2\pi]$: $\{x_n = n\pi/2N_{N-S} | 0 \leq n < 4N_{N-S}\}$ where N_{N-S} is the minimum sampling number based on the Nyquist-Shannon sampling theorem [16]. The rotation angle $x_{minimum}$ is calculated using two consecutive points x_n and x_{n+1} by a non-iterative gradient descent algorithm described by

$$x_{minimum} = \frac{x_{n+1}f'_N(x_n) - x_n f'_N(x_{n+1})}{f'_N(x_n) - f'_N(x_{n+1})} = x_n + \frac{\pi}{2N} \frac{f'_N(x_n)}{f'_N(x_n) - f'_N(x_{n+1})}, \tag{14}$$

where $f'_N(x_n) < 0$ and $f'_N(x_{n+1}) > 0$.

3.3 Chen's and Sun's Method

Chen and Sun [15] rewrote (6) using the differences between two ZM phases of consecutive m and $m - 1$:

$$\begin{aligned} \alpha &= m\alpha - (m - 1)\alpha \\ &= (\Phi_{nm} - 2\pi k_{nm}) - (\Phi_{n,m-1} + 2\pi k_{n,m-1}) \\ &= (\Phi_{nm} - \Phi_{n,m-1}) \bmod 2\pi, \quad m \neq 0, \quad 0 \leq \hat{\alpha} \leq 2\pi. \end{aligned} \tag{15}$$

They proposed a method that estimates the rotation angle α iteratively using (15). Their method is shown in (16) below.

```

Initialization  $\hat{\alpha}_0 = 0, c_0 = 0$ 
For  $m = 1, 2, \dots, M$ 
  For  $n = m, m + 2, \dots, m + 2 \lfloor \frac{N-m}{2} \rfloor$ 
     $\delta_{nm} = [(\Phi_{nm} - (m - 1)\hat{\alpha}_{m-1}) \bmod 2\pi]$ 
     $w_{nm} = \frac{|Z_{nm}| + |Z'_{nm}|}{2}$ 
  End
   $s_m = \sum_{k=0}^{\lfloor \frac{N-m}{2} \rfloor} \frac{w_{m+2k,m}}{m}; \delta_m = \frac{1}{s_m} \sum_{k=0}^{\lfloor \frac{N-m}{2} \rfloor} \frac{w_{m+2k,m}}{m} \delta_{m+2k,m}$ 
   $\hat{\alpha}_m = \frac{1}{c_{m-1} + s_m} (c_{m-1} \hat{\alpha}_{m-1} + s_m \delta_m); c_m = c_{m-1} + s_m$ 
End
 $\hat{\alpha} = \hat{\alpha}_M$ 
    
```

(16)

The estimated angle $\hat{\alpha}$ is the weighted sum of the phase differences between two ZMs sorted by m .

4. ART and the Proposed Angle Estimation Method

4.1 Angular Radial Transform

ART was proposed by Kim and Kim [20]. ART possesses additional pattern description capability compared to ZM, and takes complexities in both radial and angular directions into account.

$$F_{np} = \int_0^{2\pi} \int_0^1 f(\rho, \theta) V_{np}^{ART*}(\rho, \theta) \rho d\rho d\theta, \tag{17}$$

where, F_{np} is an ART coefficient of order n and repetition p , $V_{np}^{ART}(\rho, \theta)$ is the ART basis function that are separable along the angular and radial directions,

$$V_{np}^{ART}(\rho, \theta) = A_m(\theta) R_p^{ART}(\rho). \tag{18}$$

In order to achieve rotation invariance of the magnitude, an exponential function is used for the angular basis function,

$$A_p(\theta) = \frac{1}{2\pi} \exp(jp\theta). \tag{19}$$

Depending on the types of radial basis functions, two different transforms, ART-C and ART-S can be defined; ART with cosine and sine radial basis functions, respectively.

$$\text{ART-C} : R_n^C(\rho) = \begin{cases} 1 & n = 0 \\ 2 \cos(\pi n \rho) & n \neq 0 \end{cases} \tag{20}$$

$$\text{ART-S} : R_n^S(\rho) = \begin{cases} 1 & n = 0 \\ 2 \sin(\pi n \rho) & n \neq 0 \end{cases}. \tag{21}$$

The magnitudes of the ART coefficients are rotation invariant, similar to ZM. Let the image $f^r(\rho, \theta)$ be the rotated image of original image $f^r(\rho, \theta)$ by α :

$$f^r(\rho, \theta) = f(\rho, \theta + \alpha) \tag{22}$$

The ART coefficients of the rotated image are then given by

$$\begin{aligned} F_{np}^\alpha &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 V_{np}^{ART*}(\rho, \theta) f^r \rho d\rho d\theta \\ &= F_{np} \exp(-jp\alpha) \end{aligned} \tag{23}$$

In this paper, ART-C is used to generate the basis functions. The real and imaginary parts of ART up to order = 3 and repetition = 8 are shown in Fig. 2.

4.2 A Rotation Angle Estimation Method Based on ART

We extend Revaud et al.'s method using ART. This extension yields better estimation accuracy and noise robustness than ZM-based methods. Since the order n and the repetition p of ART are independent of each other, we can construct a set of ART basis function which can describe the

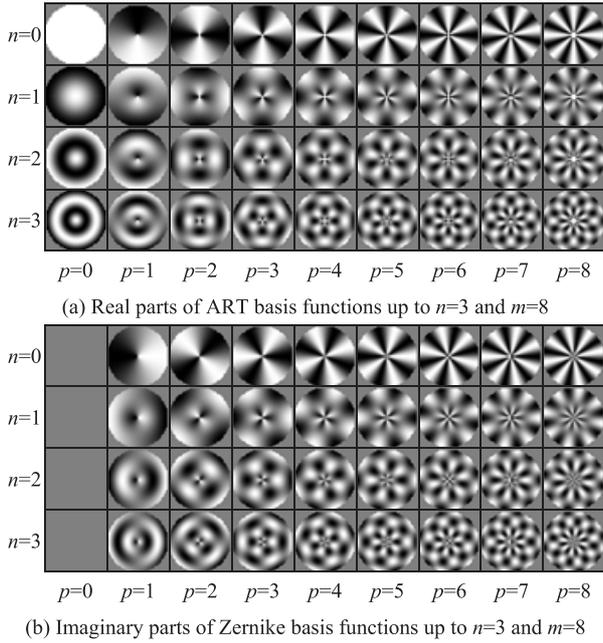


Fig. 2 The basis functions of ART.

angular complexity of an image more intensively than the radial complexity. Therefore, a set of ART coefficients can describe the angular variation of image patterns more intensively than ZMs [20]. In order to retrieve the rotation angle using the phase of the ART coefficients, we rewrote Eq. (11) using ART coefficients instead of ZMs as

$$d_{original,rotated}^2(\alpha) = \sum_{x^2+y^2 \leq 1} \sum_{p \geq 1} \pi \left[|F_{np}|^2 + |F_{np}^r|^2 - 2|F_{np} F_{np}^r| \cdot \cos(p\alpha + \varphi_{np} - \varphi_{np}^r) \right], \quad (24)$$

where $p \geq 1$. The coefficients at repetition $p = 0$ do not have valuable phase information for estimating the rotation angle. Therefore, we exclude those coefficients from the modified distance function. The $\pi/(n+1)$ in (11) represent the inner product values of the basis functions of ZMs; in (24) they are replaced with π since $\langle V_{np}^{ART} \cdot V_{np}^{ART*} \rangle = \pi$. The rest of the estimation of the rotation angle α follows the minimum search scheme of Revaud et al.'s method.

5. Dataset and Comparison Criteria

Rotation angle estimation can be performed on binary or gray scale images that contain a character, a symbol or a logo, or a gray image pattern. To compare the performances of the three methods, we compose three datasets, each containing three types of noise: photometric noise, geometric noise and deformation noise. All of the images in our datasets were scaled to 41×41 .

5.1 Experimental Dataset

The photometric noise dataset and the geometric noise

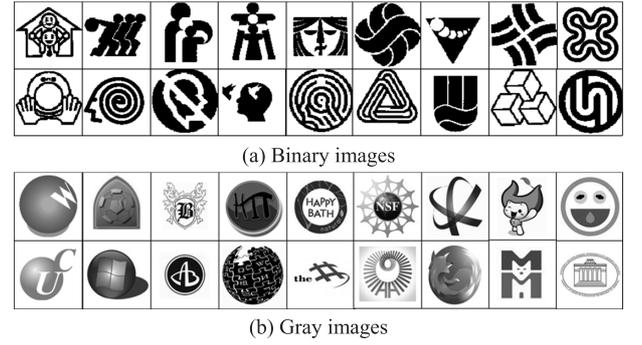


Fig. 3 Examples of original images from the photometric and geometric noise datasets.

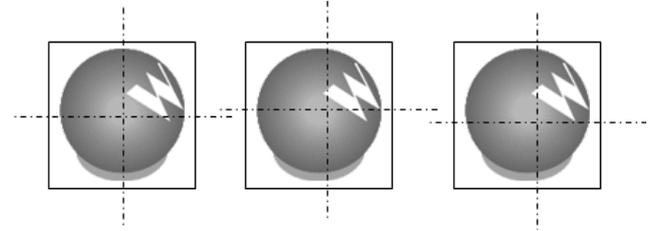


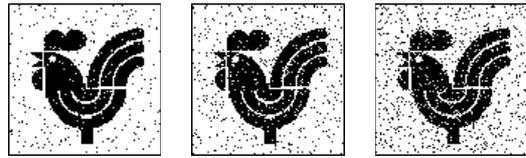
Fig. 4 Examples of images with geometric noise. The crossing point of the two lines represents the origin of the unit circle of the ZM basis functions. The left is the original image and the others are distorted images (level 5).

dataset consist of binary images and gray-scale images. A hundred binary images were randomly selected from the MPEG-7 CE 2 shape dataset and the same number of gray-scale images were gathered from websites and scaled to squares. Examples of the images in the two datasets are shown in Fig. 3.

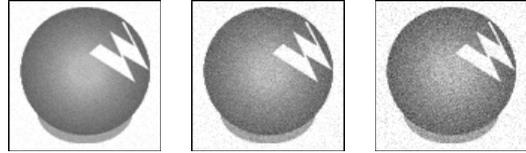
- Geometric noise: the centers of the images were translated from -2 to 2 pixels on the x - and y -axis. Therefore, the distances between the centers of unit disks of the ZM basis and image patterns varied from 1 pixel (level 1) to $2\sqrt{2}$ pixels (level 5). Figure 4 includes examples of images that contain geometric noise.
- Photometric noise: impulse noise was added to the binary images at three levels (3%, 6% and 9%), and Gaussian noise was added to the gray images at three levels (5 dB, 10 dB and 15 dB in terms of SNR). Figure 5 includes examples of images that contain photometric noise.

The deformation noise dataset contains 160 images from the MPEG-7 CE-1 B dataset. The deformation noise dataset consists of 10 groups; each group has one original image and 15 deformed images. Examples of the original and deformed images of the deformation noise dataset are shown in Fig. 6.

In this comparison, we do not consider the noise due to partial occlusion since the methods are not designed to be robust to the distortion caused by partial occlusion.

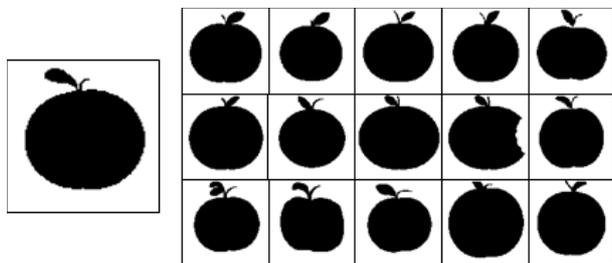


(a) Examples of images with impulse noise (3%, 6%, 9%)

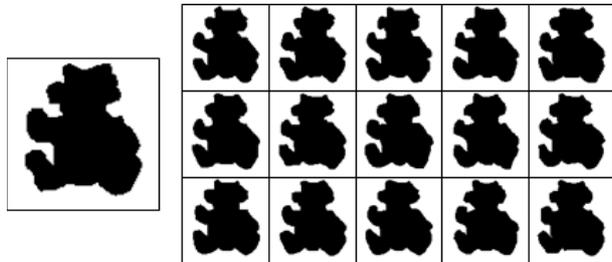


(b) Examples of images with Gaussian noise (5dB, 10dB, 15dB)

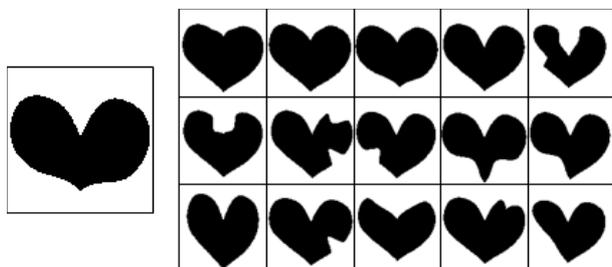
Fig. 5 Examples of images with photometric noise.



(a) Group 1



(b) Group 2



(c) Group 3

Fig. 6 Examples of original images from the deformation noise dataset.

5.2 Evaluation Criteria

The mean error alone is inadequate for assessing the accuracy of estimations because it does not contain any information about how the estimation results vary. Therefore, we used the RMS error E_{RMS} vs. coverage C as a criterion for comparison.

The RMS error is defined as (25),

$$E_{RMS} = \sqrt{\frac{\sum_i (\theta_i - \theta'_i)^2}{\text{Number of pairs}}}, \tag{25}$$

where θ_i and θ'_i are the actual and the estimated rotation angles, respectively.

The coverage is the ratio between the number of estimated angles with errors that are smaller than a specified value ε_i and the total number of image pairs compared,

$$C = \frac{\text{Number of pairs with } \varepsilon < \varepsilon_i}{\text{Number of total pairs}} \times 100, \tag{26}$$

where, ε is $|\theta_i - \theta'_i|$ and ε_i is the specified error boundary. The coverage represents the distribution of errors. A higher C at a lower E_{RMS} represents better performance.

6. Experimental Results

The maximum order n of ZM is fixed at 12 for the comparison; this is the value proposed in [15] for all three methods. For the proposed method, the maximum order of ART is fixed at 3 and the maximum repetition is fixed at 10. Consequently, Kim and Kim’s method, and Chen and Sun’s method both use 42 ZMs while Revaud et al.’s method uses 49 ZMs to estimate the rotation angle. The proposed method uses 40 ART coefficients, which is the smallest among the four methods.

Since the precision of Kim and Kim’s method is 0.1, the total number of the histogram bins of (9) of the probability density function is $360 \times 10 \times 42 = 151,200$. We performed the rotation angle estimation using each of the four methods applied to the original (unrotated, noise-free) image and its variants rotated by intervals of 5° .

6.1 Performance for the Geometric Noise Dataset

The number of geometric noisy images is $360,000 = (1 \text{ (noise-free image)} + 24 \text{ (geometric noisy images)}) \times 72 \text{ (the number of angles)} \times 100 \text{ (number of the images)} \times 2 \text{ (binary and gray images)}$. Table 1 lists the rotation angle estimation results (E_{RMS} and C) for the four methods for the binary images and the gray images under geometric noise. E_{RMS} and C were calculated for three error boundaries: $\varepsilon < 3^\circ$, $\varepsilon < 6^\circ$ and $\varepsilon < 9^\circ$. Figure 7 shows two graphs of the average E_{RMS} vs. average C of the four methods for the geometric noise dataset. The symbols for each line represent the three boundaries of E_{RMS} . Higher C at a lower E_{RMS} represents the better performance in the graphs.

All four methods show good performance for noise-free images. The estimation results for noise-free images are shown in the first column of Table 1; the C of Kim’s and Kim’s method is 100% at $\varepsilon < 3^\circ$, and the C s of the rest are also close to 100%. For binary images, the C of Kim’s and Kim’s method is higher than the proposed method by 1.93% when $\varepsilon < 9^\circ$ in the binary images, however the E_{RMS} of the proposed method is 0.106 lower than Kim and Kim’s

Table 1 Rotation angle estimates for the binary and the gray-scale images under geometric noise (E_{RMS} ($^{\circ}$), C (%)). 1: Chen and Sun’s method, 2: Revaud et al.’s method, 3: Kim and Kim’s method, 4: Proposed method.

Binary Images												
	Noise-free		level 1		level 2		level 3		level 4		level 5	
	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C
Cases of $\epsilon < 3^{\circ}$												
1	0.619	99.83%	0.989	91.73%	1.126	83.93%	1.245	74.97%	1.291	70.67%	1.401	61.36%
2	0.542	98.43%	0.921	93.63%	1.106	89.17%	1.134	79.72%	1.283	76.09%	1.412	67.25%
3	0.681	99.75%	0.964	94.90%	1.073	92.01%	1.118	84.36%	1.214	82.53%	1.330	74.32%
4	0.344	98.17%	0.628	96.49%	0.757	94.75%	0.993	89.67%	1.071	87.81%	1.248	82.37%
Cases of $\epsilon < 6^{\circ}$												
1	0.628	99.94%	1.328	96.51%	1.684	92.78%	1.835	84.49%	1.975	82.22%	2.259	75.81%
2	0.563	98.60%	1.126	96.19%	1.411	94.17%	1.730	89.77%	1.881	87.14%	2.239	82.88%
3	0.706	100.00%	1.177	97.70%	1.315	95.67%	1.491	90.77%	1.622	89.78%	1.933	86.51%
4	0.344	98.17%	0.633	96.55%	0.828	95.45%	1.147	91.74%	1.210	90.07%	1.510	87.26%
Cases of $\epsilon < 9^{\circ}$												
1	0.628	99.94%	1.426	97.07%	1.891	94.28%	2.377	88.43%	2.424	85.64%	2.699	79.52%
2	0.563	98.60%	1.170	96.43%	1.536	94.88%	1.920	91.11%	2.162	89.22%	2.612	86.28%
3	0.706	100.00%	1.230	97.96%	1.479	96.54%	1.725	92.06%	1.777	90.77%	2.111	87.84%
4	0.344	98.17%	0.633	96.55%	0.828	95.45%	1.148	91.74%	1.284	90.44%	1.601	87.63%
Gray-scale Images												
	Noise-free		level 1		level 2		level 3		level 4		level 5	
	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C
Cases of $\epsilon < 3^{\circ}$												
1	0.676	99.00	1.093	93.02	1.260	85.51	1.341	75.28	1.398	71.14	1.587	62.10
2	0.666	97.85	1.067	93.62	1.197	88.63	1.277	81.70	1.368	79.34	1.502	70.39
3	0.606	100.00	0.990	97.61	1.165	95.35	1.219	89.98	1.319	87.50	1.440	80.32
4	0.493	97.68	0.728	96.53	0.913	95.87	1.153	92.98	1.224	91.65	1.395	86.16
Cases of $\epsilon < 6^{\circ}$												
1	0.693	99.22	1.409	98.43	1.832	96.41	2.112	89.59	2.239	86.34	2.407	78.15
2	0.734	98.56	1.301	97.34	1.621	96.34	1.968	94.76	2.100	93.39	2.308	89.03
3	0.606	100.00	1.121	99.92	1.361	99.38	1.602	97.88	1.782	97.44	2.087	95.45
4	0.537	98.07	0.843	97.59	1.024	97.30	1.336	96.56	1.465	96.48	1.815	95.53
Cases of $\epsilon < 9^{\circ}$												
1	0.712	99.28	1.458	98.72	2.004	97.81	2.618	94.28	2.864	92.71	3.125	84.80
2	0.734	98.56	1.335	97.51	1.741	97.18	2.143	96.36	2.408	96.17	2.805	94.22
3	0.606	100.00	1.134	99.98	1.414	99.69	1.710	98.57	1.927	98.56	2.316	97.45
4	0.537	98.07	0.879	97.72	1.097	97.57	1.425	97.01	1.541	96.91	1.876	95.99

method. Thus the ratio ($|$ Kim and Kim’s method – the proposed method $|$ /proposed method) of C is 0.019 however the ratio of E_{RMS} s is 0.128. Therefore, the *coverage* of the proposed method is slightly lower than Kim and Kim’s method, which shows the best performance in terms of *coverage*. Yet the accuracy of the proposed method is significantly higher than that of Kim and Kim’s method.

The graphs in Fig. 7 represent the average E_{RMS} vs. the average C of the four methods calculated from the estimation results for images under the five degrees of geometrical noise in Table 1. The performance of the four methods is degraded due to the translation of the location of the image center.

The proposed method outperforms the other three methods, as shown in Fig. 7 (a) and (b). In Fig. 7 (b), similar to the results on noise-free images, the average C of Kim and Kim’s method is slightly higher than the proposed method when $\epsilon < 9^{\circ}$, however its average E_{RMS} is significantly more than that of the proposed method. Further, when compared to Revaud et al.’s method, the proposed method shows sig-

nificantly better performance due to a superior property of ART: ART can describe angular variation more thoroughly than ZMs.

The performance of Chen and Sun’s method is significantly degraded when images contain any noise since that method estimates the rotation angle just from the weighted sum of phase differences, without any consideration for noise. Since Kim and Kim’s method, Revaud et al.’s method and the proposed method performed the estimation by finding the minimum distance and the maximum probability, those three methods are more robust to noise than Chen’s and Sun’s method.

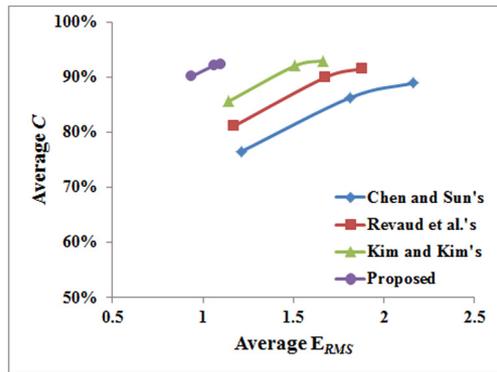
6.2 Performance Using the Photometric Noise Dataset

The number of photometric noisy images is $43,200 = 3$ (photometric noisy images) $\times 72$ (the number of angles) $\times 100$ (number of images) $\times 2$ (binary and gray image). Table 2 lists the rotation angle estimation results (E_{RMS} and C) for each of the four methods for the binary images and the

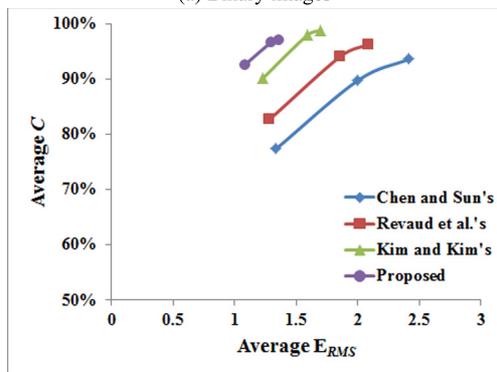
gray images under photometric noise. Figure 8 shows two graphs of the average E_{RMS} vs. the average C of the four methods for the photometric noise dataset.

For the photometric noise dataset, the proposed method yields the best performance, just as for the geometric noise dataset. The average E_{RMS} of the proposed method is signif-

icantly smaller than that of the other methods; the E_{RMS} of the proposed method is almost 0.5 times less than the average E_{RMS} of the other methods. For Kim and Kim's method,

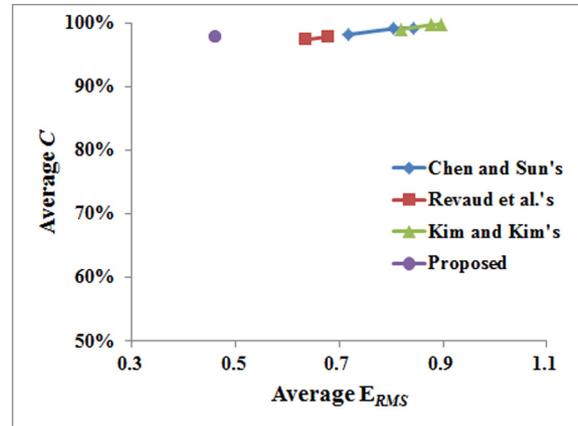


(a) Binary images

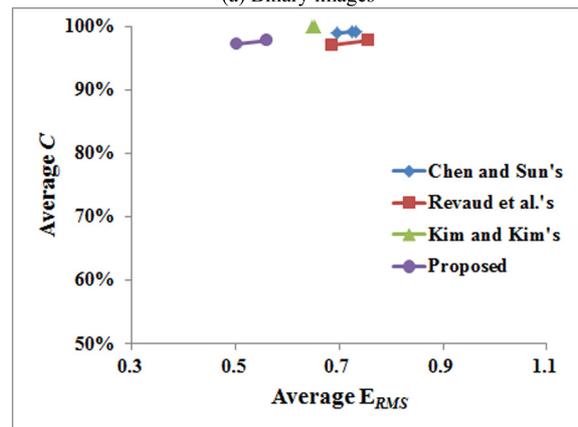


(b) Gray-scale images

Fig. 7 Average C vs. Average E_{RMS} graphs for the geometric noise dataset.



(a) Binary images



(b) Gray-scale images

Fig. 8 Average C vs. Average E_{RMS} graphs for photometric noise dataset.

Table 2 Rotation angle estimates for the binary and the gray-scale images under photometric noise (E_{RMS} ($^{\circ}$), C (%)). 1: Chen and Sun's method, 2: Revaud et al.'s method, 3: Kim and Kim's method, 4: Proposed method.

	Binary Image Group						Gray-scale Image Group					
	Impulse noise						Gaussian noise					
	level 1 (3%)		level 2 (5%)		level 3 (7%)		level 1 (5dB)		level 2 (10dB)		level 3 (15dB)	
	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C
Cases of $\epsilon < 3^{\circ}$												
1	0.677	98.71	0.713	98.32	0.764	97.75	0.681	99.01	0.695	98.99	0.708	98.83
2	0.600	97.64	0.636	97.42	0.669	97.25	0.670	97.15	0.685	97.08	0.704	97.15
3	0.769	99.40	0.825	99.00	0.864	98.75	0.616	100.00	0.643	99.93	0.685	99.92
4	0.415	97.90	0.467	97.72	0.500	97.65	0.492	97.39	0.496	97.38	0.518	97.13
Cases of $\epsilon < 6^{\circ}$												
1	0.733	99.33	0.800	99.18	0.878	98.92	0.702	99.25	0.722	99.29	0.754	99.33
2	0.632	97.93	0.681	97.90	0.723	97.85	0.741	97.88	0.755	97.83	0.775	97.88
3	0.812	99.86	0.886	99.78	0.930	99.63	0.616	100.00	0.653	100.00	0.691	100.00
4	0.415	97.90	0.467	97.72	0.503	97.68	0.545	97.88	0.558	97.93	0.577	97.65
Cases of $\epsilon < 9^{\circ}$												
1	0.736	99.35	0.850	99.33	0.942	99.15	0.706	99.26	0.727	99.31	0.766	99.36
2	0.632	97.93	0.681	97.90	0.723	97.85	0.741	97.88	0.755	97.83	0.775	97.88
3	0.828	99.90	0.898	99.82	0.962	99.74	0.616	100.00	0.653	100.00	0.691	100.00
4	0.415	97.90	0.467	97.72	0.503	97.68	0.545	97.88	0.558	97.93	0.577	97.65

Table 3 Rotation angle estimates for the deformed images (E_{RMS} ($^\circ$), C (%)). 1: Chen and Sun's method, 2: Revaud et al.'s method, 3: Kim and Kim's method, 4: Proposed method.

	1		2		3		4	
	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C	E_{RMS}	C
$\varepsilon < 3^\circ$	1.456	2.90	1.014	1.21	1.670	2.32	1.535	2.26
$\varepsilon < 6^\circ$	3.235	5.57	4.065	3.70	3.707	5.03	3.850	5.53
$\varepsilon < 9^\circ$	5.112	8.79	6.192	9.22	5.982	10.00	6.388	13.29
$\varepsilon < 12^\circ$	6.848	11.70	7.338	12.07	7.611	14.61	7.818	19.12
$\varepsilon < 15^\circ$	8.121	13.57	8.072	13.28	8.221	15.98	8.529	21.50

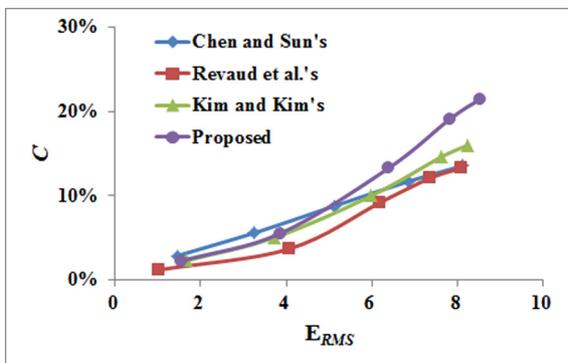


Fig. 9 Average C vs. Average E_{RMS} graphs for deformation noises dataset.

observations similar to those made above for the geometric noise dataset apply; Kim and Kim's method shows seriously degraded performance when a binary image contains impact noise.

6.3 Performance for the Deformation Noise Dataset

The number of deformation noise images is $10,800 = 10$ (image groups) $\times 15$ (deformed images) $\times 72$ (number of angles). Table 3 lists the E_{RMS} and the C of the rotation angle estimation results of the four methods for the binary images and the gray images under deformation noise. The performance of all four methods is significantly degraded for the deformed image dataset in terms of both E_{RMS} and C . Figure 9 shows a graph of E_{RMS} vs. C for the deformation noise dataset.

Since the C s of the four methods are quite low for comparing the performance up to $E_{RMS} < 9^\circ$, we extend the error boundaries to $\varepsilon < 15^\circ$. Because all four methods yield inaccurate estimation results in most cases, the average C of the four methods is only 16.08% for the cases when $\varepsilon < 15^\circ$. As shown in Fig. 9, until the average E_{RMS} is less than about 5° , Chen and Sun's method shows the best performance, closely followed by the proposed method and Kim and Kim's method. However, the proposed method yields the best performance after the average E_{RMS} is larger than $\sim 5^\circ$.

Table 4 ANOVA estimates of the experimental datasets.

Source of variation	Square sum	Degree of freedom	Mean square
Between groups	1.9E+06	3	645432.1
Within groups	1.1E+09	1677304	638.348
Total	1.1E+09	1677307	

Source of variation	F ratio	P-value	F critical value
Between groups	1011.098	0.000	3.782

Table 5 Computation times of the four methods.

	Average computation time [<i>microsec</i>]
Chen's and Sun's method	247.981
Proposed method	1021.185
Revaud et al.'s method	1029.458
Kim's and Kim's method	175218.292

6.4 Statistical Test to Verify Comparison Results

In order to verify the comparison results in this paper, we applied single factor analysis of variance (ANOVA) to the estimation results for each of the four methods. In the ANOVA test, the null hypothesis is that the means of the errors of the four methods are equal and the alternative hypothesis is that the means of the errors of the four methods are not equal. If we obtain evidence to reject the null hypothesis in the test, this indicates that the comparison results in this paper are valid.

The results of the ANOVA test of estimation results of experimental datasets are listed in Table 4. The P-value of 0.000 is less than the significance level 0.01, so we can reject the null hypothesis: the four methods have different accuracies at the 0.01 significance level. The F ratio (1011.098) is also significantly greater than the F critical value (3.782), so again, we can reject the null hypothesis that the mean errors of the four methods are all equal. Therefore, we can reject the null hypothesis and safely conclude that the comparison results are valid at the 0.01 significance level.

6.5 Computation Time

We measured the computation times of the methods throughout the experiment in order to compare their computational efficiencies. The four methods were implemented in Microsoft Visual Studio 2008 on an Intel Pentium Core2 Quad 2.4 GHz. Code was carefully designed to perform without redundancy. The average computation times of the four methods are listed in Table 5.

As listed in Table 5, Chen and Sun's method performs in the shortest time followed by the proposed method; the method is four times faster than the proposed method. The proposed method and Revaud et al.'s method required similar amounts of time to perform the estimation. The average computation time of Kim and Kim's method is about 136

times longer than that of the proposed methods because Kim and Kim's method calculates the probability density of the rotation angle at every order and repetition in the discrete domain.

7. Discussion and Conclusion

In this paper, we presented an experimental comparison of four rotation angle estimation methods that use phases of complex moments for estimation; three ZM-based methods: Kim and Kim's method, Revaud et al.'s method and Chen and Sun's method, and one new ART-based method proposed in this paper. The proposed method is a modification of Revaud et al.'s method. The experimental dataset contains a range of over 410,000 images from simple binary objects to complex gray patterns. In order to assess noise robustness, we performed the methods using images containing three types of noise and compared the estimation results. For quantitative comparison, we employed RMS error vs. *coverage* for the comparison criteria. The graphs of RMS error vs. *coverage* enabled objective comparisons between the performances of the methods. The proposed ART-based method yielded the best performance in terms of average RMS error vs. average *coverage*; the average RMS error of the proposed method is significantly lower than the other three methods for the same *coverage* values in most cases. In terms of computational time, the proposed method required less time than Revaud et al.'s method and, Kim's and Kim's method; Chen and Sun's method took the shortest computational time but yielded the worst performance in terms of the average RMS error vs. average *coverage*. Although the comparison presented here is not exhaustive, the comparison was performed using over 400,000 images and we verified the validity of the comparison by an ANOVA test. Therefore, we conclude that the proposed method is the best tradeoff between accuracy of estimation and efficiency of computation.

References

- [1] J. Revaud, G. Lavoué, and A. Baskurt, "Improving zernike moments comparison for optimal similarity and rotation angle retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.31, no.4, pp.627–637, April 2009.
- [2] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inf. Theory*, vol.8, no.2, pp.179–187, Feb. 1962.
- [3] E. Persson and K. Fu, "Shape discrimination using Fourier descriptors," *IEEE Trans. Syst., Man Cybern.*, vol.SMC-7, no.3, pp.170–179, March 1977.
- [4] W.I. Grosky and R. Mehrotra, "Index-based object recognition in pictorial data management," *Comput. Vis. Graph. Image Process.*, vol.52, no.3, pp.416–436, Dec. 1990.
- [5] Y.Y. Tang and C.Y. Suen, "Extraction of peripheral shape features in Chinese character recognition," *Proc. 12th IAPR Int. Conf. Comput. Vis. Image Process.*, vol.2, pp.377–379, Oct. 1994.
- [6] F. Mokhtarian and A.K. Mackworth, "A theory of multiscale, curvature based shape representation for planar curves," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.14, no.8, pp.789–805, Aug. 1992.
- [7] W.-Y. Kim and Y.-S. Kim, "A region-based shape descriptor using Zernike moments," *Signal Process.: Image Commun.*, vol.16,

pp.95–102, 2000.

- [8] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.4, pp.509–522, April 2002.
- [9] M. Teague, "Image analysis via the general theory of moments," *J. Optical Soc. Am.*, vol.70, no.8, pp.920–930, Aug. 1980.
- [10] S. Liao and M. Pawlak, "On image-analysis by moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.18, no.3, pp.254–266, March 1996.
- [11] A. Khotanzad and Y.H. Hong, "Invariant image recognition by zernike moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.12, no.5, pp.489–497, May 1990.
- [12] S. Paschalakis and P. Lee, "Pattern recognition in grey level images using moment based invariant features," *Proc. Int. Conf. Image Processing and Its Applications*, vol.1, pp.245–249, 1999.
- [13] C.H. Teh and R.T. Chin, "On image analysis by the methods of moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.10, no.4, pp.496–513, April 1988.
- [14] S. Li, M.C. Lee, and C.M. Pun, "Complex zernike moments features for shape-based image retrieval," *IEEE Trans. Syst. Man. Cybern.-A, Syst. Humans*, vol.39, no.1, pp.227–237, Jan. 2009.
- [15] Z. Chen and S.K. Sun, "A Zernike moment phase-based descriptor for local image representation and matching," *IEEE Trans. Image Process.*, vol.19, no.1, pp.205–219, Jan. 2010.
- [16] A.V. Oppenheim and J.S. Lim, "The importance of phase in signals," *Proc. IEEE*, vol.69, no.5, pp.529–550, 1981.
- [17] D.G. Lowe, "Object recognition from local scale-invariant features," *Int. J. Comput. Vis.*, vol.60, no.2, pp.91–110, 2004.
- [18] L. Lin, "A rotation angle calculation method about bearing images based on image correlations," *LNCS*, vol.5226, pp.1047–1055, 2008.
- [19] W.-Y. Kim and Y.-S. Kim, "Robust rotation angle estimator," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.21, no.8, pp.768–773, Aug. 1999.
- [20] W.-Y. Kim and Y.-S. Kim, "A new region-based shape descriptor," *ISO/IEC MPEG99/M5472*, Maui, Hawaii, Dec. 1999.



Jong-Min Lee received M.S. degrees in Information and Communication Engineering from the Graduate School of Communication & Information at Hanyang University, Seoul, Korea in 2001. He is now a doctoral candidate in the Division of Electrical and Computer Engineering at Hanyang University. His research interests include object segmentation, shape representation, and pattern recognition.



Whoi-Yul Kim received the Ph.D. degree in Electronics Engineering from Purdue University, W.L., IN, USA in 1989. From 1989 to 1994, He was with the Erick Johanson School of Engineering and Computer Science at the University of Texas at Dallas. He joined Hanyang University in 1994 where he is now a professor in the Department of Electronics and Computer Engineering. His research interests include visual surveillance, face tracking and identification, motion analysis, face recognition and MPEG-7 applications, and he contributed to the development of the MPEG-7 visual descriptors.