Fast Computation of Zernike Moments Using Three Look-up Tables

Sun-Gi Kim, Whoi-Yul Kim, Young-Sum Kim, and Chee-Hang Park

Abstract

Zernike moments have been one of the most commonly used feature vectors for recognizing rotated patterns due to its rotation invariant characteristics. In order to reduce its expensive computational cost, several methods have been proposed to lower the complexity. One of the methods proposed by Mukundan and K. R. Ramakrishnan [1], however, is not rotation invariant. In this paper, we propose another method that not only reduces the computational cost but preserves the rotation invariant characteristics. In the experiment, we compare our method with others, in terms of computing time and the accuracy of moment feature at different rotational angle of an object in image.

Ⅰ. Introduction

Two types of features are often used for rotation invariant pattern recognition systems: Fourier descriptor and moments[3, 6]. The main advantage of the Fourier descriptor based method lies in its speed. However, one of its drawback is that since the descriptor describe an object with the outer boundary only, it becomes very difficult for rotation invariant recognition when the object consists of several boundaries, not to mention the sensitivity to the change of boundary shape as well as the number of boundary curves even preprocessed with morphology based filtering. Moment based methods do not suffer from such drawbacks. In these methods, the whole image is regarded as the sum of two-dimensional polynomials, and the associated coefficients become feature vectors.

Many different types of moment polynomials have been proposed since the introduction of Hu's 7 moment invariants[5]. Among these, Zernike moments are used for its two distinct properties: the rotation invariant characteristics of the feature vectors and their orthogonality[2, 3]. Recently the Zernike moment based method has been successfully employed to discriminate and recognize a trademark from the database that consists of 3,000 trademarks[8]. One of the drawbacks of Zernike moments, however, lies in expensive computational cost because of its repetitiveness. In an effort to reduce the complexity of the computation of Zernike moment basis functions, W.Y. Kim and Po Yuan[2] used two look up tables for basis functions. On the other hand, R. Mukundan and K. R. Ramakrishnan[1] had converted a square image to a circular one, and moments are computed in polar coordinates to reduce the complexity in computation of polynomials. In terms of the number of multiplications, they had $O(N^2)$ complexity. However, since original square image is shaped into a circle by circular transform true sense of rotation invariance can not be obtained.

In our method, we first transform an image in Cartesian coordinates to polar coordinates because radial polynomials of Zernike moment are defined only in terms of radius, so that the pixels at the same radius have the same polynomial values. Then three look up tables were used to avoid the repeated computations. Of course, the amount of memory was increased three times more than in conventional method. However, the speed has been increased by the order of magnitude.

The organization of this paper is as follows. In section 2, we discuss Zernike moment and precious researches on fast computation of Zernike moments(R. Mukundan and W.Y. Kim). In section 3, we introduce Three Look Up Table (TLUT) method. In section 4, we discuss the experimental results and conclusion in section 5.

Ⅱ. Computation of Zernike Moments

1. Zernike Moments

Zernike moments are a set of orthogonal polynomials, which have the rotation invariant characteristics. The orthogonality implies no redundancy or overlap of information between the moments. This property enables the contribution of each moment to be unique and independent of the information in an image. The rotation

Manuscript received September 10, 1997; accepted November 15, 1997.
S. G. Kim and W. Y. Kim are with Image Engineering Lab., Department of Electronic Engineering, Hanyang University, Seoul, Korea.
Y. S. Kim and C. H. Park are with Electronics and Telecommunications Research Institute, Korea.

invariance allows the feature set, the magnitude of the Zernike moments extracted from the image, to be the same at any orientations.

Two-dimensional Zernike moments are defined as:

$$A_{nm} = \frac{n+1}{\pi} \int\int_{x^2+y^2\le 1} f(x,y) V^*_{nm}(x,y)dxdy \quad (1)$$

where, $V^*_{nm}(x,y) = V_{nm}(\rho,\theta) = R_{nm}(\rho)e^{-jm\theta}, |m| \le n$, n-$|m|$=*even*, and radial polynomials are defined as:

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

ρ : radius from origin to (x,y) pixel
θ : angle

In polar coordinates, Zernike moments are defined as:

$$A_{nm} = \frac{n+1}{\pi} \int_0^1 \int_{-\pi}^{\pi} R_{nm}(\rho) f(\rho,\theta) e^{-jm\theta} \rho d\rho d\theta \quad (3)$$

Now, let the image $f(\rho, \theta+\alpha)$ be the rotated image of $f(\rho, \theta+\alpha)$ by α about its origin, then Zernike moments of $f(\rho,\theta+\alpha)$ are given as

$$A'_{nm} = \frac{n+1}{\pi} \int_0^1 \int_{-\pi}^{\pi} R_{nm}(\rho) f(\rho, \theta+\alpha) e^{-jm\theta(\theta+\alpha)} \rho d\rho d\theta$$

and

$$A'_{nm} = A_{nm} \exp(-jm\alpha) \quad (4)$$

In order words, the magnitudes of Zernike moments of an image become rotation invariant because $|A'_{nm}| = |A_{nm}|$.

In a conventional method, the real and imaginary part of the Zernike moments, C_{nm} and S_{nm}, respectively, are computed as follows:

$$C_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) R_{nm}(\rho) \cos(m\theta) \quad (5)$$

$$S_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) R_{nm}(\rho) \cos(m\theta)$$

where ρ and θ are defined as:

$$\rho = \sqrt{\left(\frac{2(x-\bar{x})}{N}\right)^2 + \left(\frac{2(y-\bar{y})}{N}\right)^2},$$

$$\theta = \tan^{-1}\left(\frac{y-\bar{y}}{x-\bar{x}}\right)$$

$$A_{nm} = C_{nm} + jS_{nm}$$

Here, $(\bar{x}, \bar{y})$ is the center of the image.

When the moments are computed in Cartesian coordinates, C_{nm} and S_{nm} have to be recomputed at each pixel location (x,y) in raster scan order. So, in order to compute K moments up to the nth order for an $N \times N$ image, the total computational cost for $R_{nm}(\rho)$ is $K \times N^2 \times O$, where O is the operation for computing single Zernike moment polynomial, and K is determined by n as follows:

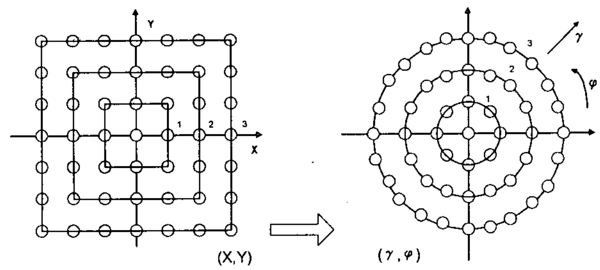


Fig. 1. Schematic of square-to-circle transformation. (Quoted from Mukundan [1])

$$K = \sum_{k=0}^{n} \left[\frac{(k+2)}{2}\right] \quad (6)$$

For example, Eq.(6) yields K=90 when n=17 for a typical application of recognizing a trademark from a database[8].

The basic idea for the fast computation of Zernike moment is as follows: $R_{nm}(\rho)$ is the same for all pixels at the same radius ρ from the origin. In a digital image of $N \times N$, since the number of ρ is limited to N/2, $R_{nm}(\rho)$ needs to be computed only $K \times N/2 \times O$ times in polar coordinates. So when Zernike moments are computed by Eq.(3), the complexity of Zernike moment polynomial reduces from $O(N^2)$ to $O(N)$. In order to apply this concept to a digital image, a transformation from Cartesian coordinates into polar coordinates is necessary and will be discussed in the subsequent sections.

2. *Circular Transform (CT) Method*

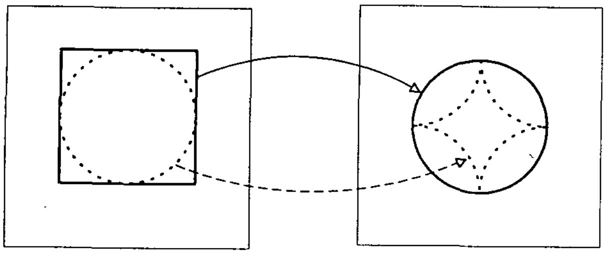
There are three types of distance functions in digital image; they are Euclidean distance(D_e), city-block distance (D_4) and chess-board distance (D_8)[7]. Metrics defined by each distance function from p to q with coordinates (x,y) and (s,t), respectively, are

1) $D_e(p,q) = [(x-s)^2+(y-t)^2]^{1/2}$
2) $D_4(p,q) = |x-s| + |y-t|$
3) $D_8(p,q) = \max(|x-s|, |y-t|)$

The pixels at the same distance will form a circle, diamond and square by these metrics, respectively.

In CT method proposed by Mukundan, the image pixels along the concentric squares were mapped onto concentric circles by transformation as shown Fig. 1.

The index γ along the radial direction takes a values from 1 to $(N/2)$, whereas the index φ along the circumference takes a value from 1 to 8γ. So Zernike moments polynomial $R_{nm}(2\gamma/N)$ needs be computed only once for all pixels mapped on to the same circle.



(a) (b)

Fig. 2. Distance at rotated image in CT method.

The drawback of CT method

In CT method, the pixels from the origin to the pixels in D_8 distance are considered the same in D_e distance after transformation. In other words, pixels along the square will map onto those along the circle (solid line). With this concept, however, a circle (dotted line) in D_e distance will no longer map onto a circle after the transformation as illustrated in Fig. 2(b). Consequently, Zernike moments for a circle will not be rotation invariant unlike those for the square shape except when the angles are multiples of 90 degrees.

3. Basis Look up Table (BLUT) Method.

In Eq. (1), will be the same for any image when the size of image is the same. Therefore, two lookup tables of size $N\times N$ for are prepared off-line, and an input image is normalized to the size, Zernike moments of the image are then computed as follows:

$$C_{nm} = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)\,VR_{nm}(x,y) \qquad (7)$$
$$S_{nm} = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)\,VI_{nm}(x,y)$$
$$A_{nm} = C_{nm}+jS_{nm}$$

where VR_{nm} and VI_{nm} are real and imaginary look up tables of $V^*_{nm}(x,y)$, respectively.

With this method, since there is no need for computation of basis functions on-line, the method requires only one multiplication and one addition per pixel, and the complexity is reduced to $O(KN^2)$, where K and N are the number of moments and the size of image, respectively.

III. Three Look Up Table (TLUT) Method

In our method three look up tables for radius, angle, and S_m are prepared for computing Zernike moments in polar coordinates. To do so, x and y values are first mapped to ρ and θ. Then, the radius ρ, whose range is from 0 to 1, is normalized to R_{max}, and stored in a look up table $T_r(x,y)$ for radius by the following formula:

$$T_r(x,y) = \gamma = [\rho\times R_{max}],\quad R_{max} = [N/2] \qquad (8)$$

where $[x] = [x+0.5]$, which is the nearest integer.

Another table T_a for angle is prepared for the whole $N\times N$ image again as follows:

$$T_a(x,y) = \tan^{-1}\left(\frac{y-\bar{y}}{x-\bar{x}}\right) \qquad (9)$$

The mapping of ρ to $T_r(x,y)$ has an advantage of computing Zernike moments fast to reduce the complexity for R_{nm}. In other words, the computation of R_{nm} can be reduced from $O(N^2)$ to $O(N)$ as follows: suppose we connect the pixels that have the same γ in the radius table, $R_{max}+1$ concentric circles are drawn ranging from 0 to R_{max}. Because the number of different values for γ is $R_{max}+1$, the computation of R_{nm} reduces to $R_{max}+1$ from N^2 for $N\times N$ image when γ is used for ρ. Then Eq. (5) can be rewritten in terms of T_r, T_a as follows:

$$C_{nm} = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)R'_{nm}(\gamma)\cos(m\theta) \qquad (10)$$
$$S_{nm} = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)R'_{nm}(\gamma)\sin(m\theta)$$

where $\gamma=T_r(x,y)$, $\theta=T_a(x,y)$ and $R'_{nm}(\gamma) = R_{nm}\left(\frac{T_r(x,y)}{R_{max}}\right)$

Now, in Eq. (9), although the total number of R_{nm} is reduced to $N/2$, the total number of multiplication remains to $2\times K\times N^2\times 2$ which is the same as the conventional and that of CT method. This number of multiplication can also be reduced by eliminating the redundancy in the definition of Zernike moments in polar coordinates. In TLUT method, Zernike moments A_{nm} in polar coordinates can be expressed as:

$$A_{nm} = \frac{n+1}{\pi}\int_0^1 R_{nm}(\rho)\int_{-\pi}^{\pi} f(\rho,\theta)e^{-jm\theta}\rho\, d\rho\, d\theta \qquad (11)$$

In Eq. (10), since the 2nd integral term consists of m and ρ, which is the same for any values of n, it is casted into $S_m(\rho)$. Then the Zernike moment A_{nm} is computed by:

$$A_{nm} = \int_0^1 R_{nm}(\rho)S_m(\rho)\,d\rho \qquad (12)$$

Now the equation is expressed by ρ only which is already quantized and mapped to γ. By using γ for ρ in Eq.(11), the real and imaginary parts of A_{nm} are computed by:

$$C_{nm} = \frac{n+1}{\pi}\sum_{r=0}^{R_{max}} R'_{nm}(r)\cdot SC_m(r) \qquad (13)$$
$$S_{nm} = \frac{n+1}{\pi}\sum_{r=0}^{R_{max}} R'_{nm}(r)\cdot SS_m(r)$$

where $SC_m(r)$ and $SS_m(r)$ are cosine and sine terms of

$S'_m(r) = S_m\left(\frac{r}{R_{max}}\right)$, respectively.

In the actual implementation, SC_m and SS_m for each m are computed as follows: since $SC_m(r)$ and $SS_m(r)$ are the functions of distance only, they can be computed by accumulating the result of multiplication between f and sinusoidal term. Furthermore, they are computed in Cartesian coordinates using radius and angle tables which were already prepared. Because the tables have the distance and angle values for each pixel at (x,y), the result can be easily accumulated to SC_m and SS_m terms by indexing the radius table. A pseudo-code for computation of SC_m and SS_m are as follows;

```
For i = 0 to R_max
  SC_m(i): = 0
  SS_m(i): = 0
end

For x : = 0 to N
  For y : = 0 to N
    SC_m(T_r(x,y)) : =
        SC_m(T_r(x,y)) + f(x,y) × cos(m × T_a(x,y))
    SS_m(T_r(x,y)) : =
        SS_m(T_r(x,y)) + f(x,y) × sin(m × T_a(x,y))
  end
end
```

To compare our TLUT method with others in terms of the complexity, the computational costs for multiplication and R_{nm} are listed in table 1, where C_{nm} and S_{nm} are computed up to nth order and the number of moments K is determined by Eq. (6). The actual number of computations is shown in table 2 when the maximum order of the moment is limited to 8, i.e., 25 moments, for 100×100 image.

Table 1. Complexity of each method.

Method	Multiplication	R_{nm}
TLUT	$4 \times N^2 \times (n+1) + 2 \times K \times N/2$	$K \times N/2$
CT	$4 \times N^2 \times K$	$K \times N/2$
BLUT	$4 \times N^2 \times K$	None
Direct	$4 \times N^2 \times K$	$K \times N^2$

TLUT : Proposed Three Lock Up table method
CT : Circular Transform method
BLUT : Basis Look Up Table method
Direct : Direct computation by eq.(5)

Table 2. Actual number of computation.

Method	Multiplication	R_{nm}
TLUT	3.6×10^5	1250
CT	1.0×10^6	1250
BLUT	1.0×10^6	0
Direct	1.0×10^6	2.5×10^5

Table 3. Zernike moments for each order.

Order (n)	Moments (A_{nm})	No. of moments
0	A_{00}	1
1	A_{11}	1
2	A_{20}, A_{22}	2
3	A_{31}, A_{33}	2
4	A_{40}, A_{42}, A_{44}	3
5	A_{51}, A_{53}, A_{55}	3
6	$A_{60}, A_{62}, A_{64}, A_{66}$	4
7	$A_{71}, A_{73}, A_{75}, A_{77}$	4
8	$A_{80}, A_{82}, A_{84}, A_{86}, A_{88}$	5

Table 4. Comparative performance analysis of Zernike moments for various image size. (sec)

Method	101 by 101	155 by 155	201 by 201
TLUT	0.0978	0.1793	0.2937
CT	0.2830	0.5932	1.0738
BLUT	0.2748	0.5342	0.9334
Conventional	5.184	12.31	20.78

Ⅳ. Experiment and Result

In order to evaluate the performance of our method compared to CT, BLUT and the conventional direct method, the elapsed time in computer simulation and SNR between original and rotated image were measured. Twenty-five moments up to 8th order were computed on IBM PC (Pentium 100MHz CPU) as shown in table 3. Table 4 shows the elapsed time of three methods for various size of image. To verify the rotation invariant characteristics of our proposed method, Zernike moments of a trademark image are computed at different orientations. Fig. 3 shows moments at 0, 30, 60, 90 and 120 degrees, respectively. SNRs between 0, 15, 30, 45, 60, 75 and 90 degrees are depicted in Fig. 4. For more evaluation at different sizes of image, SNRs between 0 and 30 degrees are depicted in Fig. 5. Here, $|A_{nm}|$ are Zernike moments of original image, $|A^r_{nm}|$ are the moments of rotated image and $|\overline{A}|$ is the mean of the moments. SNR is defined as:

$$SNR = 10\log\frac{\sigma^2}{\sigma_e^2} \tag{14}$$

where, $\sigma^2 = E[(|A_{nm}| - |\overline{A}_{nm}|)^2]$, $\sigma_e^2 = E[(|A_{nm}| - |A^r_{nm}|)^2]$

In terms of computation speed, CT, BLUT and TLUT methods yielded much better results than the conventional method as shown in Table 4. Among these, the elapsed time for TLUT method was three times faster than CT and BLUT methods. For the extracted feature vectors, however, the conventional method and TLUT method have the closest rotation invariant characteristics, while CT method fails to do so as depicted in Fig. 3, Fig. 4 and Fig. 5.

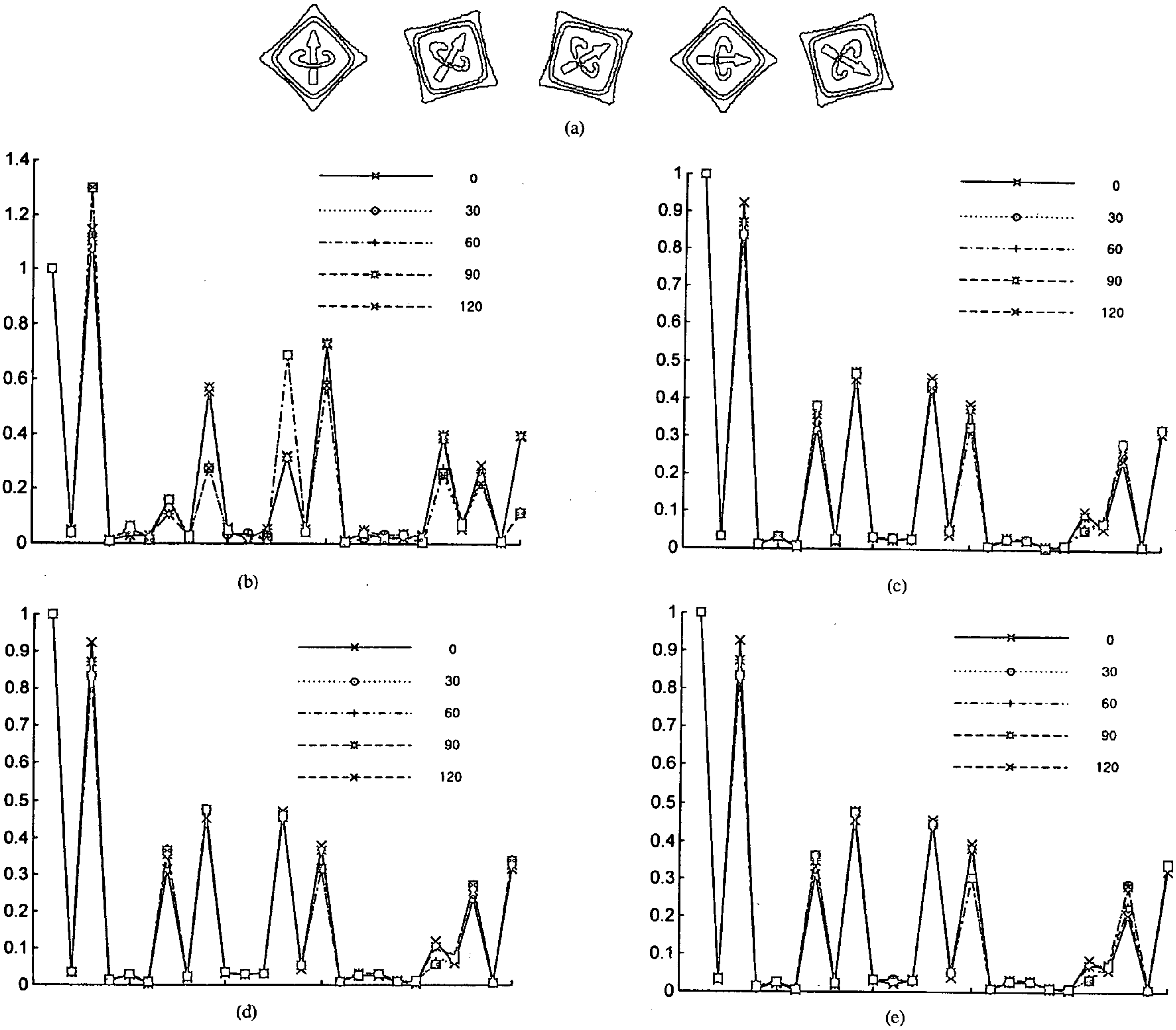


Fig. 3. Zernike moments of the image in (a) at different orientation, by: (b) CT, (c) TLUT, (d) direct and (e) BLUT method.

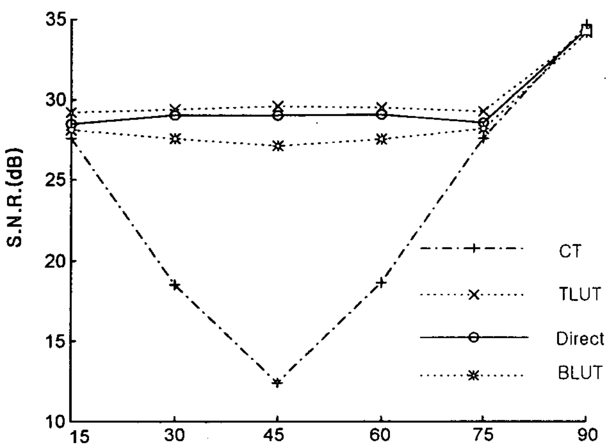


Fig. 4. SNR between original and rotated image for static size of image.

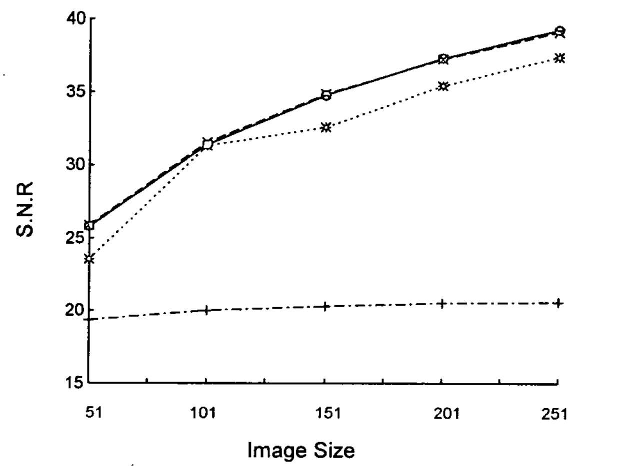


Fig. 5. SNR between original and rotated image for static rotated angle.

V. Conclusion

Zernike moment feature is one of the powerful features that can be used to recognize one out of 3,000 trademarks at any scale and orientations. The computational complexity of Zernike moments, however, has prevented its application to real world problems. In this paper, we have proposed a method to compute Zernike moments fast using three lookup tables. In particular, we have shown that the computational redundancy of radial polynomials can be eliminated when computed in polar coordinates. Although several methods have been proposed to utilize the similar concepts, we used three look up tables instead of one to avoid the repeated computations. At the price of memory increase compared to existing methods, however, the speed was increased by the order of magnitude, resulting in its application possible for Zernike moments under real-time constraints. We have compared their performance in terms of speed and accuracy at arbitrary angles. The results indicate that our method is far superior to any other methods in terms of speed and accuracy.

Acknowledgements

This work was supported by the Electronics and Telecommunications Research Institute under grant 97-202.

References

[1] R. Mukundan and K. R. Ramakrishnan, "Fast Computation of Legendre and Zernike moments," *Pattern Recognition*, Vol. 28, No. 9, pp. 1433-1441, 1995.

[2] W. Y. Kim and Po Yuan, "A Practical Pattern Recognition System for Translation, Scale and Roration Invariance," *IEEE, Conference on Computer Vision and Pattern Recognition*, 1994.

[3] A. Khotanzad and Y. Hong, "Rotation Invariant Image Recognition Using Features Selected via a Systematic Method," *Pattern Recognition*, Vol. 23, No. 10, pp. 1089-1101, 1990.

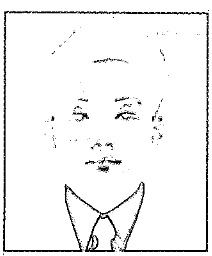
[4] A. Khotanznd and Y. H. Hong, "Invariant Image Recognition by Zernike Moments," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 12, No. 5, pp. 489-497, May 1990.

[5] M. K. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Trans.* Info Theory 8(1), pp. 179-187, 1962.

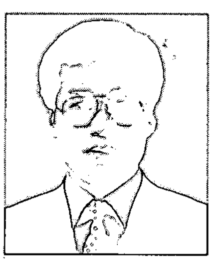
[6] Z. Mingfa, S. Hasani, S. Bhattarai and H. Singh, "Pattern recognition with moment invariants on machine vision system," *Pattern Recognition Letters* 9, pp. 175-180, 1989.

[7] R. C. Gonzalez and R. Woods, *Digital Image Processing*. MA: Addison-Wesley, 1992, pp. 45-47.

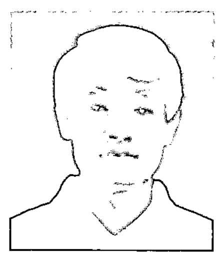
[8] Y. S. Kim and W. Y. Kim, "Conent-Based Trademark Retrieval System Using Visually Salient Feature," *IEEE Conference on Computer Vision and Pattern Recognition*, 1997.



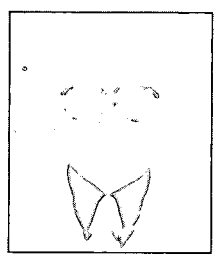
Sun-Gi Kim received the B.S. degree in electronic engineering from Hanyang University, Seoul, Korea, in 1996. He is currently working on his Master's Degree at the same University. His research interests aer pattern recognition and machine vision applications.



Whoi-Yul Kim received the B.S. degree in Electronics Engineering from Hanyang University, Seoul, Korea, in 1980. He received the M.S. degree from Pennsylvania State University, University Park, in 1983, and Ph.D. Degree from Purdue University, West Lafayette, IN, in 1989, both in electrical engineering. From 1989 to 1994, he was with the University of Texas at Dallas. Since 1994, he has been on the faculty of electronic engineering at Hanyang University, Seoul, Korea. He has been doing research in image processing and image understanding, and machine vision application.



Young-Sum Kim received the B.S. and M.S. degree in telecommunications and the Ph.D. degree in computer science from Hanyang University, Korea, in 1989. His research interests include Natural Language Processing Multimedia Information Retrieval, Spoken Language Processing. He is currently senior member of research staff in comupter technology division, ETRI, Korea.



Chee-Hang Park received the B.S. degee in applied physics from Seoul National University, Korea, in 1974, the M.S. degree in computer science from Korea Advanced Institute of Science and Technology, Korea, in 1980, and the Ph.D. degree in computer science from University of Paris 6, France, in 1987. During last 10 years he has been involved as project leader with several large projects such as multi-media computer system development and high speed paralled computer system development. His research interests include multi-media systems, distributed systems, middleware, groupware, network virtual computing, and mobile agent architecture. He is currently vice president of computer technology division of Electronics and Telecommunications Research Institute.