An Analytical and Experimental Study of Binary Image Normalization for Scale Invariance with Zernike Moments

Whoi-Yul Kim

Abstract

In order to achieve scale- and rotation-invariance in recognizing unoccluded objects in binary images using Zernike moment features, an image of an object has often been normalized first by its zeroth-order moment (ZOM) or area. With elongated objects such as characters, a stroke width varies with the threshold value used; it becomes one or two pixels wider or thinner. The variations of the total area of the character becomes significant when the character is relatively thin with respect to its overall size, and the resulting normalized moment features are no longer reliable. This dilation/erosion effect is more severe when the object is not focused precisely. In this paper, we analyze the ZOM method and propose as a normalization method, the maximum enclosing circle (MEC) centered at the centroid of the character. We compare both the ZOM and MEC methods in their performance through various experiments.

I. Introduction

Moment-based features from binary images are widely used in many applications for pattern recognition tasks including character recognition[1]. The hardware implementation of computing moments in real-time for images has been realized[2]. An excellent survey on moment-based pattern recognition systems can be found elsewhere[3].

The task of recognizing unoccluded 2-D objects with different sizes randomly scattered through an image requires the system to be invariant to translation, scale and rotation of the pattern. The translation invariance is easily achieved by moving the origin of the coordinate system to the centroid of each object when computing the moment features. For the rotational invariant feature, Hu's 7 moments, circular harmonic expansion, Zernike or pseudo-Zernike moments can be used[4, 5, 6]. For scale invariance, normalization techniques have typically been used, for example, by making the object's height and width equal to standard values [1]. When Zernike moments are used for feature vectors, this normalization is crucial to the problem of object recognition at random orientations. One of the common methods of normalization is to use the zeroth order moment (ZOM) of the object[3, 7, 6, 8].

In most industrial applications where the input image consists of binary patterns, the imaging environment can be controlled precisely so that binary patterns are optimally segmented for inspection or recognition tasks. When the input image is degraded, due to different lighting conditions or off-focused, the system performs poorly. An adaptive thresholding method compensates the contrast of the background, but the method is computationally expensive for real-time applications since the gradient information should be incorporated[9]. Even when objects are segmented with the adaptive method, for example in OCR applications, the stroke width or thickness of the character changes, i.e., the boundary of characters shrinks or expands by few pixels resulting in change in area. This small change becomes non-trivial when the boundary of the object is relatively longer with respect to its area.

In this paper, our aim is to analyze the ZOM-based method for scale invariance. The analysis will be focused on the effect of different threshold values and the blurred effect when Zernike moments are employed as feature vectors. Although it is not limited to characters, different sizes of printed alphabet characters at random orientation are used as input patterns.

The organization of the paper is as follows. Sections 2, 3 and 4 overview the past work relevant to the translation and scale invariance using moments. This section also contains discussions on why normalization is crucial to the scale invariance for the problem of object recognition using Zernike moments. We then analyze in Section 5 the effect of dilation/erosion (D/E) caused by thresholding in-focus and blurred images. In Section 6 we describe the scale invariance using the ZOM method, and compare it with the Maximum Enclosing Circle (MEC) method. We then discuss experimental results of using both methods in Section 7. Finally

Manuscript received September 13, 1997; accepted November 21, 1997.
The Author is with Image Engineering Lab. Department of Electronic Engineering Hanyang University, Seoul, Korea.

section 8 concludes our study.

II. Past Work

Most regular moment-based recognition systems have used either images of similar size or pre-normalized images with single object; an image that contains a single object normalized to a standard size before computing its moments. For example, a single object in each separate image of the size 22×16 are used in[10], 26 character images of size 64×64 are used in[6], library and test images of size 128×128 are used in[11, 12], and 132 small images in[13]. When multiple objects are present in an image, each object has to be first segmented before extracting its features. Cash and Hatamian used a document image of size 1792×2304 digitized using a scanner[1]. The document image consisted of 24 lines of the 62 alphanumerics(1488 characters). Each character has been segmented using the contour tracing algorithm and has been normalized to be filled in a 32×32 array. However, this method is not suitable, when the rotation involved are elongated patterns. For example, the pattern "I" would be too thick when normalized at an angle 90 degree. When Zernike moment is used for feature vectors, one of the common method of normalization is to use ZOM[6, 8]. For example, Khotanzad and Hong have used 314 binary images of size 64×64 with Zernike moments as their feature vectors[6].

The advantage of ZOM method discussed in[6] is that it is easier and simpler to compute directly from each image since there is only single object in it, and parallel implementation can speed up the process. The drawback of the method is that each binary object in the image has to be pre-normalized before computing its features, resulting in extra computational complexity. Furthermore, since Zernike moments are only defined within the unit circle, the size of the normalizing circle or the ratio of the circle with respect to the object should be defined first. The circle should be large enough to include the object completely so that the resulting moments are truly invariant to rotation.

III. Translation and Scale Invariance

The regular moment (or geometric moment) [14] of order $p+q$ for a continuous image function $f(x,y)$ is defined as

$$m_{pq} = \int\int x^p\, y^q\, f(x,y).$$

The translation invariance is easily obtained by moving the origin of the coordinate system to the centroid of the image $f(x,y)$. This moment is called *central moment*, and expressed for a digitized image $f(x,y)$ as

$$m_{pq} = \sum_x\sum_y (x-\bar{x})^p\, (y-\bar{x})^q\, f(x,y),$$

where $\bar{x}$ and $\bar{y}$ are the centroid of the image $f(x,y)$, and are computed from

$$\bar{x} = \frac{m_{01}}{m_{00}} \quad \text{and} \quad \bar{y} = \frac{m_{10}}{m_{00}}.$$

For the scale factors of an image $f(x,y)$, there are in general three parameters to be considered: the scale factors s_x and s_y are respectively for x and y directions of the image, and the scale factor s_f is for the gray scale of the image. The moment of the scaled image is then computed as

$$\begin{aligned} m'_{pq} &= \sum_x\sum_y x^p\, y^q\, f(\frac{x}{s_x}, \frac{y}{s_y})\; dx\; dy \\ &= s_x^{\,p+1}\; s_y^{\,q+1}\; s_f\; m_{pq}. \end{aligned}$$

When different scaling factors s_x and s_y are desired, aspect normalization can be used[11]. With intensity images, this scaling can be considered as resampling by assuming that the gray value $f(x,y)$ does not change with scaling in x and y directions, otherwise contrast invariance is to be used[15]. With range images, $f(x,y)$ denotes the distance from the sensor to the point on the object, and needs to be considered[3]. With binary images, the gray value scaling factor s_f is ignored, and both s_x and s_y is assumed to be s. Then m_{00} or ZOM represents the object area or the total number of pixels in the object.

One common approach to achieving the scale invariance is by normalizing the area to 1 or to a constant β[3, 6, 8, 12]. To do so, we first compute the area of the object. Using regular moments, the area is computed as

$$m'_{00} = s^2\; m_{00},$$

and the scale factor s is computed as

$$s = \sqrt{\frac{\beta}{m_{00}}}.$$

Therefore, as the area m_{00} changes due to any scale difference, the scale factor s can be easily computed to incorporate changes into the feature vectors. Alternatively $\sqrt{m_{20}+m_{02}}$ can be used[3, 14, 7] in a similar manner to ZOM method.

IV. Rotation Invariance

Zernike moments have been widely studied because of its two distinct properties; the rotation invariance of the feature vectors and their orthogonality[16]. The first property allows the feature set, the magnitude of the Zernike moments extracted from the image, to be the same at any orientation. The second property implies no redundancy or overlap of information between the moments[6, 8]. This property enables the contribution of each moment to the information of the image be unique and independent, and becomes useful for classification of features.

The Zernike moment is defined only inside the unit circle and the radial polynomial $R_{mn}(\rho)$ is defined as

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

Then Zernike moments of order (m, n) in polar coordinate are defined as[6].

$$A_{nm} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho, \theta)\ R_{nm}(\rho)\ \exp(-jm\theta)\rho\ d\rho d\theta.$$

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

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

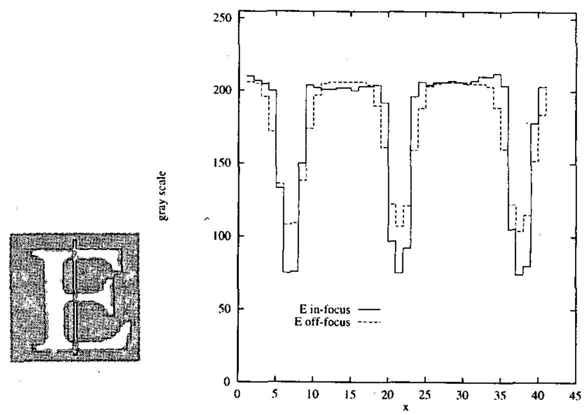
This implies that the rotation of the image has given rise to a phase shift in the moments, and the magnitudes remain the same as before the rotation.

V. Dilation/Erosion (D/E) by Thresholding

In many machine vision applications, the area of the same object or character can be changed in two ways; by scale change and/or by a threshold. When the binary image is used, since ZOM is the same as the area of objects in the image, the diameter of the normalizing circle becomes a function of the area. As shown in Eq. 3, the diameter of the normalizing circle is determined by the scale factor s as a function of the zeroth order moment (ZOM) and m_{00}. The circle is chosen to be large enough to contain the pattern completely.

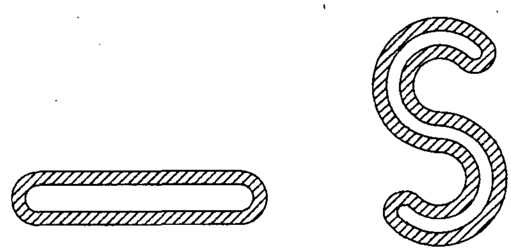
The scale change may be due to the physical change in the imaging system setup such as the change of lenses with different focal length, or different sizes of the same patterns. The other cause is the threshold used to make a binary pattern. That is, the character image may shrink or expand due to the different thresholds used. This phenomenon is severe especially when the image is not precisely focused. As an example, Fig. 5-(a) shows the slightly blurred image of character "E." The size of image is 43×31 and the gray scale along the vertical line of the character is plotted in Fig. 5-(b). The solid line is the profile of the raw image, and the dotted line indicates that of the off-focus image. Notice that the stroke width varies from 1 pixel to 4 in the raw image as the threshold changes. In the off-focus or blurred image, the stroke width ranges from 1 to 6 pixels. The actual boundaries of binary regions of the character in the blurred image are shown in Fig. 3 at the thresholds of 76, 136 and 196 to yield a reasonable binary image for this particular subimage. Here the lowest threshold was manually selected so that all strokes are connected to form the character, and the highest one was selected to exclude just the background region.

When the stroke width is thin compared to the size of the character, and the shape number ($boundary^2/area$) is higher[17], the variations in the area change are more significant. As illustrated in Fig. 2, the quantization and different threshold cause the D/E effect to be more significant when the shape of the



(a) (b)

Fig. 1. Digitized image of "E" with a CCD camera and its cross-sectional profile along the vertical line. The size of the character is 43×31 pixels.



(a) (b)

Fig. 2. Two symbols are modelled here to show the shrinking of characters due to different thresholds.

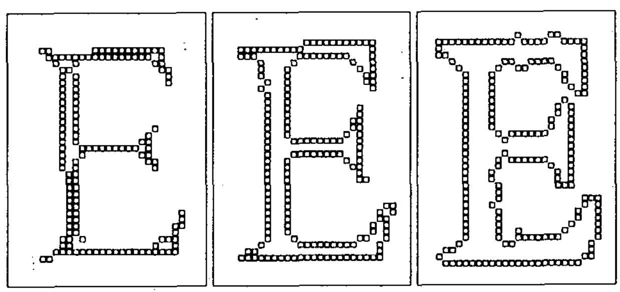


Fig. 3. Boundary points of "E" with different thresholds; 76,136 and 196.

pattern under inspection is elongated and its size is relatively small. In other words, the thickness can increase or decrease by one or more pixels from each side of the stroke. For example, one pixel dilation along the boundary of the bar pattern whose thickness is 2 pixels causes its area to expand almost two fold. Consequently the overall size occupied in the normalized circle reduces by $1/\sqrt{2}$ or 0.7. Similarly, by Eq. 3, with the character shape "S" whose thickness is 1 pixel wide, the area may be expanded to almost three times the pattern. The analysis on this scaling will be thoroughly analyzed in later sections.

VI. Calculation of the Normalizing Circle Diameter

1. Analysis of ZOM Method

Owing to its straightforward nature, the condition given by Eq. 3 sounds simple and this approach to obtain the scale invariance is attractive. The direct implementation as stated, however, is not so because the area information of the largest object in the field is needed. Since the Zernike moment is defined only within a normalizing circle and the size of the pattern is not known in advance, the circle size should be selected large enough to include the largest object. The difficulty has to do with the fact that in the presence of several objects with different shapes and sizes in an image, it may not be easy to determine the size of largest pattern, until the whole image is read and processed. Often *a priori* information of the object in the field can be used. However, when there is large size variance between the patterns, the circle size may be too large for some patterns, and the extracted features may not have any discriminatory power at all for small objects.

Here we take four patterns of three classes as an example for illustration. Figure 4 shows four symbols of "$-$," "$+$," "o" and "$+$" shape. Here, the symbols "$-$", "$+$" and "o" are modeled to have the same area, t_l, of the hatched region. The symbol "$-$" and "$+$" have the same width l. The thickness or stroke width t_b of "$+$" is a little more than the half of t in "$-$" because of the overlapping region at the cross. The thickness of "o" is the same as that of "$-$." Circles surrounding each pattern show the minimum size of circular region where Zernike moment is defined. The patterns "$+$" and "$+$" are similar in shape, and have the same width and height. But the stroke width t_a of "$+$" has been determined to keep the area twice as large as that of "$-$," "$+$" and "o," which is a little longer than t. In order to normalize all of these patterns, we first find the smallest circles for each pattern that includes itself completely, which is similar to the one placed on the symbol "$-$." Here "m_{00}" denotes the ZOM, d_{min} is the minimum diameter to contain the object completely, and s_{zom} is the scale factor computed by ZOM method, and they are computed as shown in the table 1. Here β has been set to tl^3 for

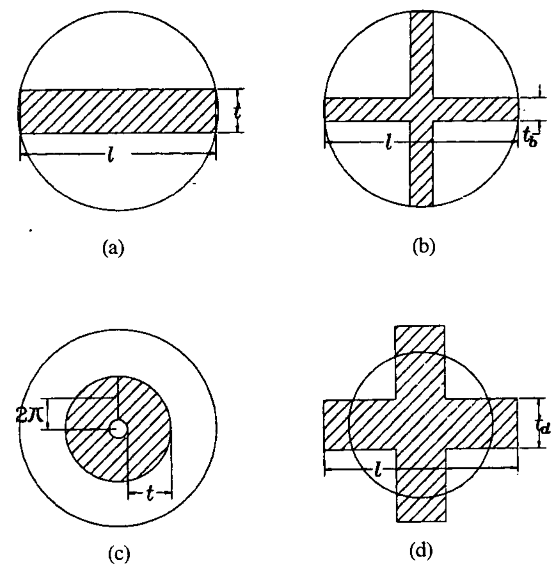


Fig. 4. Three symbols are modelled here to analyze the ZOM method.

Table 1. The properties of symbols are compared in terms of area, minimum radius of the normalizing circle and their ratios.

Case	Shape	m_{00}	d_{min}	s_{zom}	$\frac{d_{min}}{s_{zom}}$
a	-	tl	$\sqrt{l^2+t^2} \approx l$	$\sqrt{\frac{\beta}{tl}} = l$	1
b	+	$\frac{tl}{2} \times 2 - \frac{t^2}{4} \approx tl$	$\sqrt{l^2+t_b^2} \approx l$	$\sqrt{\frac{\beta}{tl}} = l$	1
c	o	tl	$\frac{l}{\pi} + t \approx \frac{l}{\pi}$	$\sqrt{\frac{\beta}{tl}} = l$	$\frac{1}{\pi}$
d	+	$2\ tl - t^2 \approx 2\ lt$	$\sqrt{l^2+t_d^2} \approx l$	$\sqrt{\frac{\beta}{2\ lt}} = \frac{1}{\sqrt{2}} l$	$\sqrt{2}$

simplicity.

The ratio d_{min}/s_{zom} indicates how small or large is the circle determined by the ZOM method with respect to the optimum one. In order to have the object within the circle, this ratio should be less than 1. When the ratio is equal to 1, it becomes the optimum size in a sense that the object is defined at maximum in the circle.

Since the patterns and their sizes are not known in advance, the circle has to be large enough to include the object lest the pixels outside the region should be included. With the ZOM method, the minimum size of the circle should be larger than the largest diameter which is selected as:

$$d_{max} = \max(d_{min,a}, d_{min,b}, d_{min,c}, d_{min,d}) = l$$

When the diameter of the enclosing circle is determined by normalizing its ZOM to a constant so that the patterns (a) and (b) are enclosed by the circle, it would be too small to enclose the pattern (d). On the other hand, this normalizing circle would be too large for the pattern (c) whose diameter is less than all the circumscribing circles and is only l/π. In this case the portion of the pattern that occupies the inside of the normalizing circle is only about one third ($1/\pi=0.32$), i.e., the circle defined by the ZOM method would be π times larger than the smallest enclosing circle. Another difficulty has to do with the D/E effect due to the quantization of thresholding, since different thresholds shrink or expand the binary regions. With dilation, suppose that the pattern (a) expands its boundary by half of its width, $t/2$, so that it becomes $l+2\times t/2$ long and $t+2\times t/2$ wide, the minimum size of circle diameter to enclose the bar by the ZOM method is then

$$d_{\min,dilation} = \sqrt{(l+t)^2+(2t)^2}$$
$$s_{zom,dilation} = \sqrt{\frac{\beta}{(l+t)\ 2t}}\,.$$

On the other hand, when the bar gets thinned to the single pixel width from its boundary with erosion, we have

$$d_{\min,erosion} = l-2\times t/2+1\approx l$$
$$s_{zom,erosion} = \sqrt{\frac{\beta}{l\times 1}}=\sqrt{tl}$$
$$\frac{d_{\min,erosion}}{s_{\min,erosion}} = \frac{1}{\sqrt{\beta}}\,.$$

As a numerical example, for a bar pattern which is 30 pixels high and 5 pixels wide, the diameter ratio of $d_{min,d}$ to $s_{zom,d}$ is 2.18. This ratio may be too large to be useful. This situation is easily observed in real images when the character is thin and especially when the character is small (say, 30×30), the threshold we choose makes a significant difference when the pattern is normalized. As the threshold changes, it causes the stroke width to shrink or expand by one or more pixels from its own boundary due to quantization. Although the overall size of the character barely changes (to 28×28), the area of the character changes significantly when the stroke width is small.

2. Maximum Extent Circle (MEC) Method

For the reasons stated in the previous section, the normalizing circle should be estimated for each pattern separately based on the boundary of the pattern and should not depend on the area of the pattern itself. As illustrated in Fig. 4, when the maximum diameter is solely defined by the boundary points of the pattern, all different sizes of the pattern will have circles all circumscribing the pattern, yielding true scale invariant characteristics. There are two ways of obtaining such circles:

- o Find the (maximum diameter) circle that circumscribes exactly.
- o Find the maximum diameter from the centroid.

In the first method, the circle is determined by the maximum extent of the points of the boundary. The method is the extension of the very well known, "Convex Hull Finding" algorithm[18]. The computational complexity of this method is $O(n\log n)$ where n is the number of boundary points of the binary pattern. With this method, the center of the circle determined bears no relation to the area of the pattern, and solely depends on the formation of boundary points of the pattern. In the second method, however, from the list of boundary contours the pattern is searched to find the maximum radius from the centroid of mass of the binary region which is the center of the normalizing circle. We have chosen the second method because: 1) its simplicity for implementation because the complexity of the method is only $O(n)$; 2) the center determined by the centroid of mass would be more reliable and less sensitive to noise in the pattern or image.

VII. Experiment

In this section, we compare the two methods discussed in the previous sections by first simulating the D/E effect on the boundary of the character "E" at different thresholds. Then we perform the experiment on the real image captured by a CCD camera. The inverted image is used to compute the moment, i.e., the background 0 and the object region 255.

1. Simulation Results

In order to show the D/E effect more effectively, we amplified the effect by dilating/eroding the character region by 2 pixel steps. In Fig. 5, there are five boundaries overlapped to show the D/E effect. The middle in solid line is the original size to start with. The size of the character is 50 pixels wide and 90 pixels high. Two inner contours show the eroded boundaries of the character by two pixels each. The other two outer contours are for the dilated or expanded boundaries by two pixels each. Fig. 6 shows the relationship between the characters and normalizing circles determined by ZOM method. Fig. 6-(a) shows the original size of character and the circle in which Zernike moment is defined. In this figure, the radius of the surrounding circle is calculated by Eq. 3. The character is scaled accordingly by keeping the circle size fixed for the purpose of illustration. Recall that the scale or radius of the circle increases by the square root of the ratio of increased area to the area of the fixed region, i.e., the size of the character decreases by the same ratio. That is, when the boundary points are expanded by 2 pixels, the relative size of the character with respect to the bounding circle gets smaller in (b) and becomes smallest in (c). Figure 6-(d) shows that the size of character gets larger with respect to the circle, and finally the character is too large for the circle in Fig. 6-(e). For comparison, we have shown the result obtained by MEC method in Fig. 7. Again, the size of the circle is fixed and the object is scaled accordingly to show the effect of the normalized circle plots of "E." This figure shows that the overall size of the

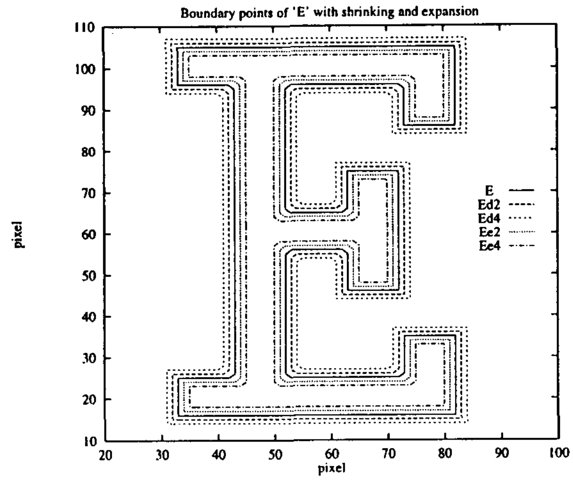


Fig. 5. Overlap of contour plots of the simulated character "E".

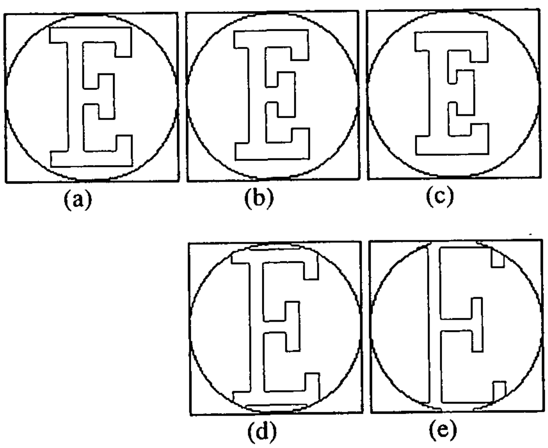


Fig. 6. Normalized plots of bounding circles for "E" computed by ZOM method with shrinking and expansion of the boundary.

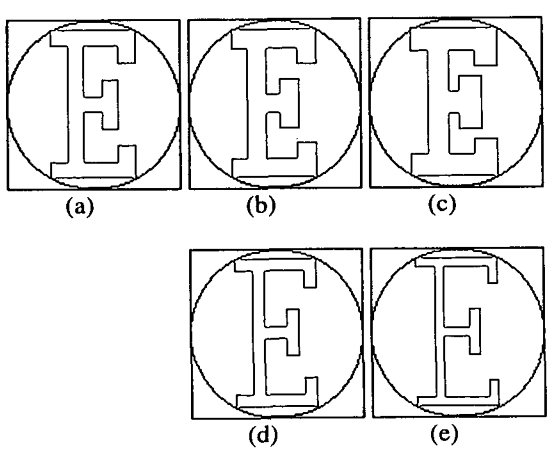


Fig. 7. Normalized plot of bounding circle for "E" computed by MEC method with shrinking and expansion of the boundary.

character barely changes as D/E occurs at different thresholds. It

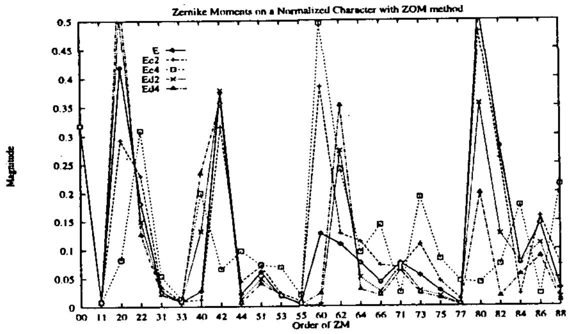


Fig. 8. Zernike moments after normalization to the simulated character "E" using ZOM method.

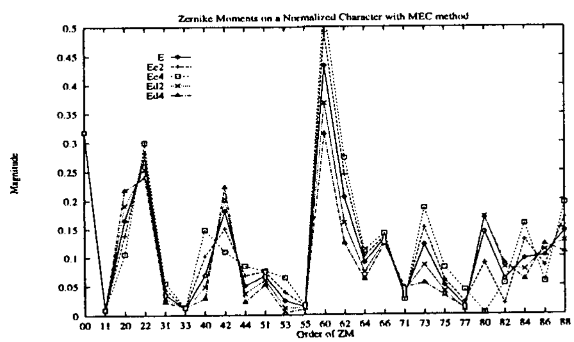


Fig. 9. Zernike moments of the simulated character "E" after normalization using MEC method.

should not matter how large or small the circle is as long as the feature vector that we are interested in is the same. However, this is not the case. Fig. 8 shows the plot of 25 sets of Zernike moments on the image normalized by the ZOM method. The abscissa of the plot shows the order of Zernike moments. Total of 25 moments are shown in the plot. Note that the zeroth order Zernike moments $|A_{00}|$ for each contoured region are the same as the normalized area. The moment $|A_{11}|$ in all plots which should be 0 by definition is not zero because the part of the object region excluded from the circular region caused the centroid to shift. Each curve is annotated by the upper left corner of the plot to show the degree of D/E by numbers. For example, "Ed4" denotes that Zernike moment set computed from the boundary dilated by 4 pixels. In Fig. 8, some degree of correlation is apparent between the moments sets computed from the adjacent contours, however, hardly can any correlation be observed between the contours which are not adjacent. On the other hand, Fig. 9 shows the result obtained using our MEC method. Comparing with the previous plot in Fig. 8 by ZOM method, the feature vectors of Zernike moments of all contours are very similar, and there exists high degree of correlation among the feature vector sets. Because of the thickness change, the overall shape of the character changes radially from the centroid of the character only by couple of pixels, which is only 2 or 3 % error.

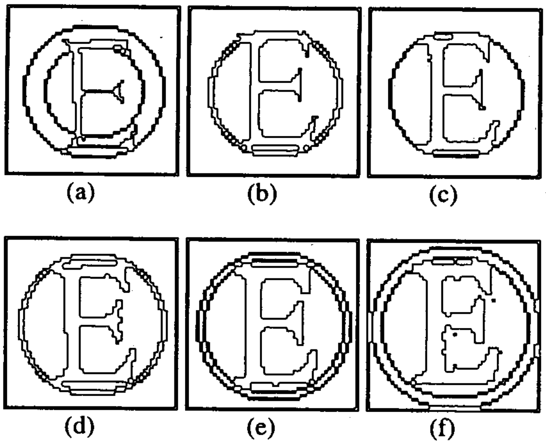


Fig. 10. In these figures, the radius of the normalizing circle determined by the ZOM method varies as a function of area of "E" at different thresholds. The one determined by the MEC method barely changes.

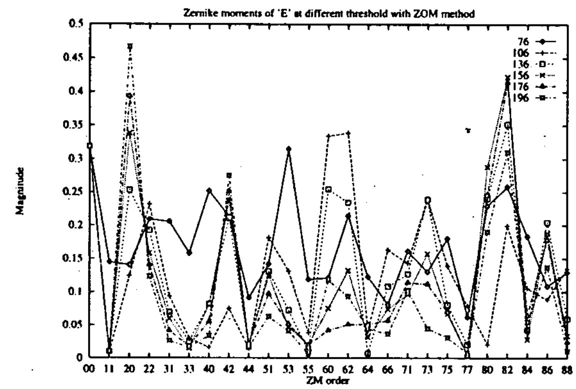
2. Real Image

(A) Single Object Image

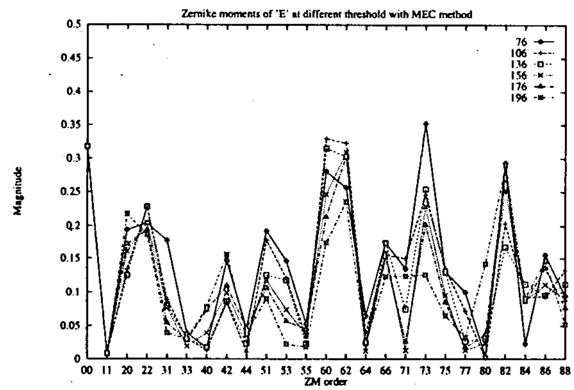
Next experiment is performed on the real thresholded images. An input image consisting of English alphabets is captured by a Panasonic CCD camera with a resolution of 512×480. The raw image as shown in Fig. 5 of the letter "E" is again selected for the experiment. Figure 3 shows the area change due to the different thresholds selected where the image is binarized with thresholds ranging from 76 to 196 to yield different character thicknesses. Normalized circles are placed around the characters in the image to show the extent of the region where Zernike moments are defined. Only the pixels within the character after normalization contribute to the moments. Figures 10-(a) and (f) show the lowest and highest threshold, respectively. Figure 10-(c) shows an optimal threshold selected at the middle of the histogram between two clusters of black and white. The diameters of the normalized circles are determined by both ZOM and MEC methods. From Figs. 10-(c) to (f), the inner circles correspond to the ZOM method and the outer one to the MEC method. As the character stroke becomes thinner, the normalizing circle by the ZOM method becomes smaller, while the circles by MEC method remain almost the same. In Fig. 12, the radii determined by both methods are plotted with respect to threshold. The radii by the MEC method is shown (dotted line) to remain almost constant over the whole span of threshold. Beyond the threshold 178, the radius starts to go up because some of the regions in the background begin to merge as shown in the profile in Fig. 5-(b). Zernike moments as features vectors extracted from these characters are also shown in Fig. 11-(a) and (b), respectively. The results are very similar to that of the simulation.

Table 2. Average number of pixels in the image of the characters as the threshold changes and their standard deviation.

Threshold	130	165	200
Average area	275.34	400.45	562.82
STD	17.74	18.78	21.84



(a)



(b)

Fig. 11. The plot of magnitudes of Zernike moments on the binary image thresholded at 63, 78, 123, 148, 168 and 178, and normalized by ZOM method and MEC method. Notice that all moments are the same at the 0th order because all magnitudes are normalized by the zeroth order magnitude. By the definition, $|A_{11}|$ is zero for all binary images.

(B) Multiple Object Image

To further evaluate the methods employed in this study, the following experiments were performed. A set of approximately 100 characters composed of 24 alphabets were prepared using the Postscript[1)] language. The size of the characters ranges from 12 to 64 pixels with gradual changes while the rotation angle steps by 15 degrees. The largest size of a character is limited only by the resolution of the CCD camera. The image is captured using a

1) Trade mark of Adobe Systems Inc.

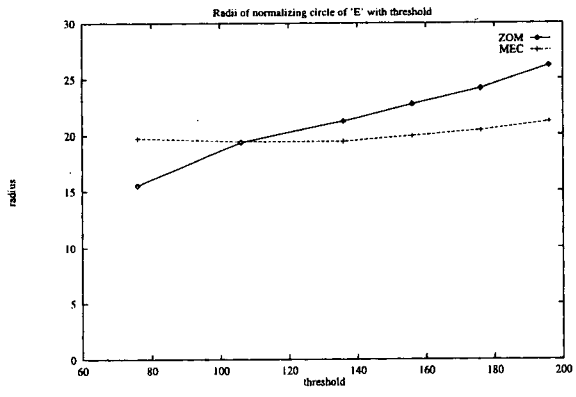


Fig. 12. The radii computed by ZOM method and MEC method.

(a)
(b)
(c)

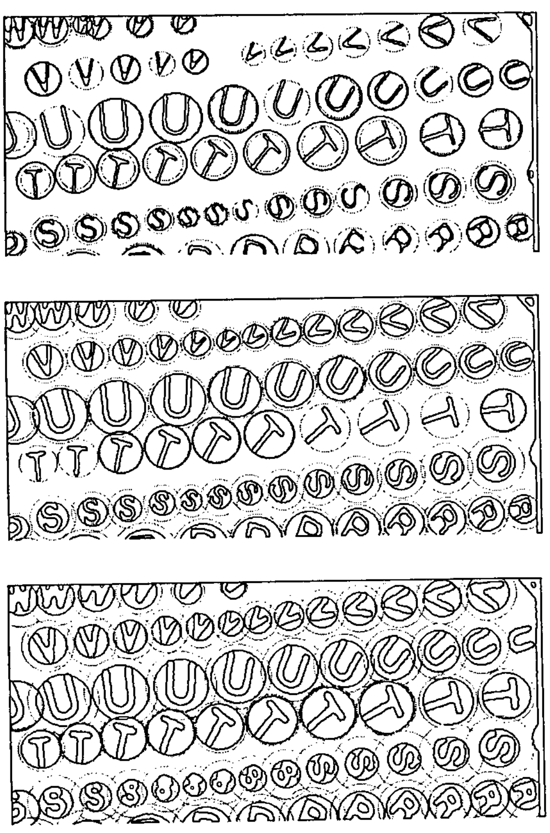


Fig. 13. The figure showes (a)-(c) show the normalizing circles of each character using both ZOM and MEC methods at thresholds 130, 165 and 200, respectively. Circles surrounding each character in dotted line are for ZOM method while in solid lines are for MEC methods.

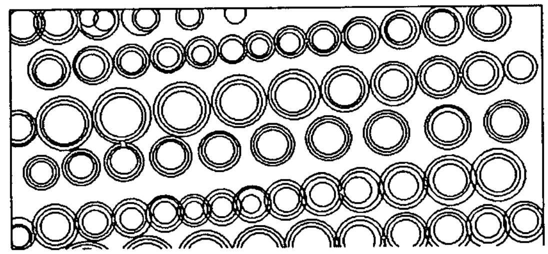


Fig. 14. In this figure, all normalizing circles that circumscribe each characters are overlapped using ZOM method at different thresholds to show the dependency of the threshold used.

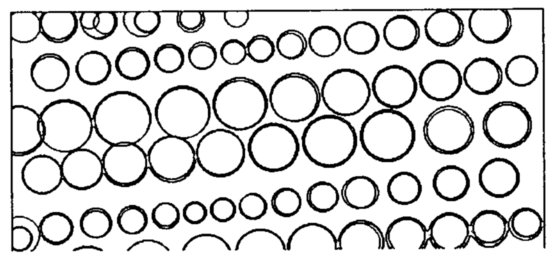


Fig. 15. In this figure, all normalizing circles that circumscribe each characters are overlapped using MEC method to show the invariance of the threshold used.

CCD camera via a frame grabber ITEX 151 imaging system. The binarization process plays an important role especially since the performance of the system depends on the normalization which in turn depends on the binarization. The smallest size of the pattern is 12 pixels, and even a single pixel due to noise can make a significant difference in the systems' overall performance. In order to reduce the noise from the digitizer as well as the artifact of the patterns, one of the simplest method for industrial application is to average a frame of image which can be easily done by the imaging system hardware in real time. Eight frames of image are averaged which takes less than a second.

Figure 13-(a) shows the partial view of our sample sheet. The thresholds here are 130 as the lowest one, 165 being the visually optimum and 200 for the highest not to break the characters into smaller segments. The characters thresholded at 130 look thinner than the optimal one as shown in Fig. 13-(b). The average area of each character at different thresholds are listed in Table 2. The area changes by 31% from optimum at the lowest threshold and 40% at the highest, which accounts for a radius change from 1.2 times to 0.84 of the optimal one. From Table 2, a constant β has been carefully selected for the normalizing circle to cover most of characters in Fig. 13-(b) but not to run over to the adjacent characters in the image. The circles are overlayed to show the difference of the two methods. Dark circles are from the MEC method and gray circles are from the ZOM method. Notice that some "T" characters are not completely contained in the gray circle in Fig. 13-(a). On the other hand, the gray circles are too large for the characters in Fig. 13-(c). The ratio of the smallest and the largest circles is larger as the pattern is elongated or the shape number goes up. As per our estimation, the size difference between circles by both methods are most significant with "S" characters. The character "W," the third one from the top left

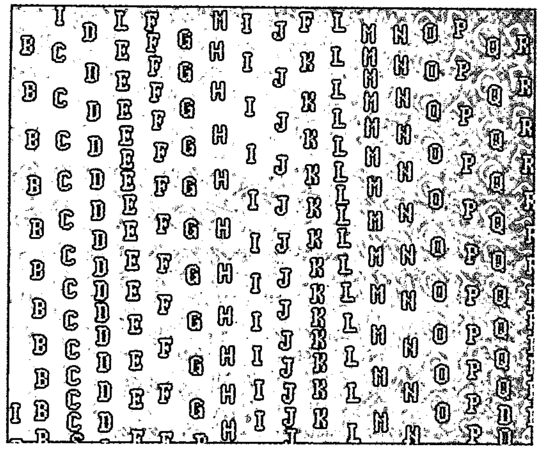


Fig. 16. Experiment with Helvetica character fonts, white fonts illustrate recognition results.

corner is broken into two pieces because of the low threshold and treated as two separated characters. Similarly the 6th character "V" in the second row has been broken into several pieces, treated as noise by the system and deleted from the image. Figure 14 shows the overlap of the circles defined by ZOM at different thresholds. The dependency of ZOM method to threshold value is obvious. On the other hand, in Fig. 15 all circles determined by MEC method changes very little as the threshold varies, and shows the threshold invariance. As shown in Fig. 10, the recognition system using the MEC method with the Zernike moment for rotational invariance has achieved close to 100% recognition rate on the characters in the image except the ones along the image boundary while the ZOM method only yields 85%[19].

To illustrate the threshold invariance, we have run the experiments to recognize the street names labeled along the curve. The image of the map is taken from MAPSCO© using the same CCD camera, and is shown in Fig. 17-(a). Figure 17-(b) shows the results when applied a global threshold of 102 and Fig. 17-(c) shows the results when 162 is applied as a global threshold. In both figures white icons indicate the computer interpretation of the original character. In Fig. 17-(b) some character along the boundary of the image are not recognized. Similarly, no attempt has been made to recognize the characters connected to the road lines in Fig. 17-(c).

Ⅷ. Conclusion

In order to use the scale invariant property using the normalization technique, care should be taken when the zeroth order moment is used. Although the Zernike moment feature is location, scale and rotation invariant, it is no longer invariant when the pattern has been dilated or eroded and if it is treated as a scaled image. In other words, in moment based recognition systems which employ normalization processes as a scale invariant charac-

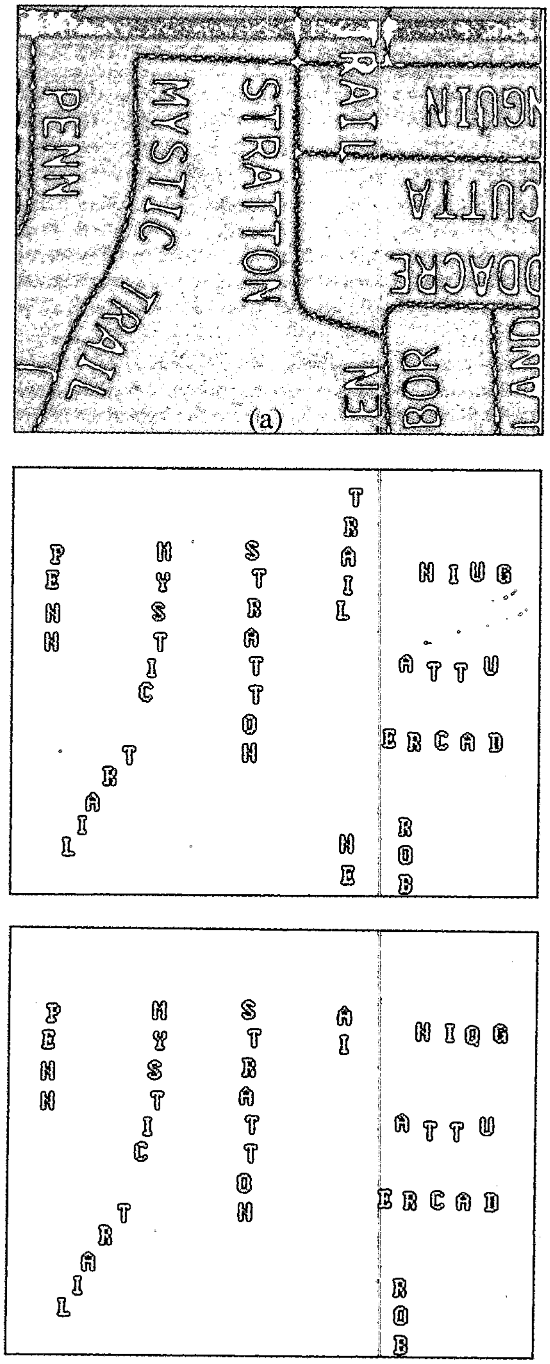


Fig. 17. (a) A map image of size 512×512. (b) Recognition result at the threshold 102. (c) Recognition result at the threshold 162. Some of the characters are not recognized because they are connected to the road lines.

teristic, the feature vectors from D/E patterns are not reliable. When the character thickness reduces due to a high threshold, its area decreases while the width and height of the character changes very little. Therefore, when the zeroth order moment is used to determine the scale factor, the scale reduces by the square root of the ratio of areas resulting in a smaller encompassing

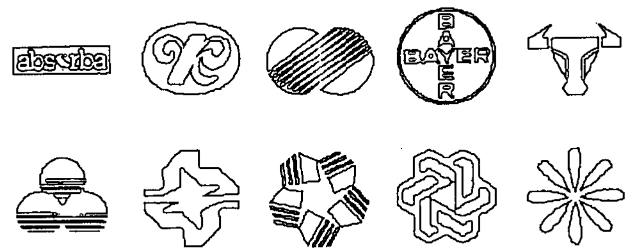


Fig. 18. A set of samples from a database consists of 3,000 trademarks.

circle than the size of the character. On the other hand, when the thickness increases, the circle becomes too large for the character. The resulting set of Zernike moments become drastically different and is not useful as a feature set. With the MEC method, the diameter of the normalized circle centered at the centroid of the character whose diameter can be obtained from the contour tracing algorithm changes very little, yielding a much more reliable feature set for recognition upto 99.5%[19].

This method has been implemented successfully to recognize more complex images than alphabets as illustrated in Fig. 18, where the database consists of 3,000 various shapes of trademarks [20].

References

[1] G. L. Cash and M. Hatamian, "Optical character recognition by the method of moments," *Computer Vision, Graphics and Image Processing*, Vol. 39, pp. 291-310, 1987.

[2] M. Hatamian, "A Real-time two-dimensional moment generating algorithm and its single chip implementation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-34, June 1986.

[3] R. J. Prokop and A. P. Reeves, "A Survey of moment-based techniques for unoccluded object representation and recognition," *CVGIP: Graphical Models and Image Processing*, Vol. 54, pp. 438-460, September 1992.

[4] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, February 1962.

[5] Y. N. Hsu, H. H. Arsenault, and G. April, "Rotation invariant digital pattern recognition using circular harmonic expansion," Appl. Opt., Vol. 21, pp. 4012-4015, 1982.

[6] A. Khotanzad and Y. H. Hong, "Invariant image recognition by zernike moments," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol 12, pp. 489-498, May 1990.

[7] S. O. Belkasim, M. Shiridhar, and M. Ahmadi, "Pattern recognition with moment invariants: A comparative study and new results," *Pattern Recognition*, Vol. 24, No. 12, pp. 1117-1138, 1991.

[8] A. Khotanzad and Y. H. Hong, "Rotation invariant image recognition using features selected via a systematic method," *Pattern Recognition*, Vol. 23, No. 10, pp. 1089-1101, 1990.

[9] J. R. Parker, "Gray level thresholding in badly illuminated images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 13, no. 8, 1991.

[10] P. Gader, B. Forester, M. Ganzberger, A. Gillies, B. Mitchell, M. Whalen, and T. Yocum, "Recognition of handwritten digits using template and model matching," *Pattern Recognition*, Vol. 24, No. 5, pp. 421-431, 1991.

[11] A. P. Reeves, R. Prokop, S. E. Andrews, and F. P. Kuhl, "Three-dimensional shape analysis using moments and fourier descriptors," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. 10, pp. 937-943, November 1988.

[12] A. P. Reeves and R. W. Taylor, "Identification of three-dimensional objects using range information," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. 11, April 1989.

[13] S. A. Dudani, K. J. Breeding, and R. B. McGhee, "Aircraft identification by moment invariant," *IEEE Trans. on Computers*, Vol. C-26, pp. 39-46, January 1977.

[14] M. R. Teague, "Image analysis via the general theory of moments," *J. Pot. Soc. Am.*, Vol. 70, August 1980.

[15] S. Maitra, "Moment invariants," *Proceedings of the IEEE,* Vol. 67, pp. 697-699, April 1979.

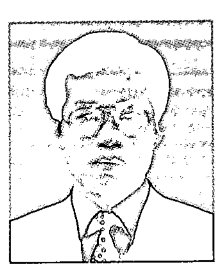
[16] C.-H. Teh and R.T. Chin, "On image analysis by the methods of moments," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. 10, July 1988.

[17] A. Rosenfeld and A. Kak, *Digital Picture Processing.* Academic Press, 1982.

[18] F. P. Preparata and M. I. Shamos, *Computational Geometry.* Springer-Verlag New York Inc., 1985.

[19] W. Y. Kim and P. Yuan, "A Practical Pattern Recognition System for Translation, Scale and Rotation Invariance," *IEEE Conference on Computer Vision & Pattern Recognition*, June 19-24, 1994, pp. 391-396, Seattle, Washington.

[20] Y. S. Kim and W. Y. Kim, "Content-based trademark retrieval system using visually salient feature," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 1117-1138, June 1997.



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.