IMPROVEMENT OF ZERNIKE MOMENT DESCRIPTORS ON AFFINE TRANSFORMED SHAPES

Gholamreza Amayeh, Shohreh Kasaei, , George Bebis, Alireza Tavakkoli, Konstantinos Veropoulos

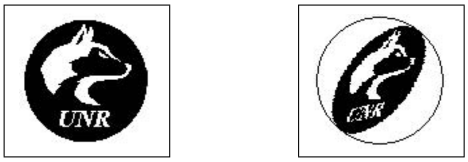
ABSTRACT

In general, Zernike moments are often used efficiently as shape descriptors of image objects, such as logos or trademarks that cannot be defined by a single contour. However, because these moments are defined in a unit disk space and extracted by a polar raster sampling shape, information of skewed and stretched shapes is lost. As a result, they can be inefficient shape descriptors when there is skew and stretch distortion. In this paper, a method is proposed that addresses this issue. More specifically, Zernike moments are obtained from a transformed unit disk space that allows for the extraction of shape descriptors which are invariant to rotation, translation, and scale as well as skew and stretch, thus preserving more shape information for the feature extraction process. The experimental results demonstrate that the proposed algorithm is more accurate in relation to skew and stretch distortions when compared to other available schemes reported in the literature.

1. INTRODUCTION

Shape is a key information to human for distinguishing visual data. To retrieve an image from a large database, we need shape descriptors that have sufficient discriminatory properties and are robust to noise. These descriptors should be invariant to translation, rotation, scale, and perspective of images. Zernike moments have some desirable properties such as rotation invariance, robustness to noise, expression efficiency, and multi-level representation for describing various shapes of patterns [1, 2, 3, 4] but they do not have scale and translation invariance properties. However, with proper preprocessing [5, 1, 6, 7], translation and scale invariance can be achieved. In order to make Zernike moments invariant to translation, the origin of the unit disk is translated to the center of mass of the shape. For scale invariance, *Khotanzad* and *Hong* [6] resampled the shape so that the area of the shape has a constant value. *Bin* and *Jia-xiong* [7] improved the rotation and scale invariance property significantly by fitting the smallest circle on the shape and used this circle as the unit disk. Zernike moments which have been acquired using the latter method were normalized by dividing over the shape area.

Primarily, Zernike moment descriptors are obtained by applying the Zernike moment transform on a shape in the unit disk space. When severe skew or stretching occurs, the shape of the distorted region within the unit circle is changed dramatically. For example, in Figure 1, perceptually, shape 1(a) and shape 1(b) are homogenous. However, due to skew, shape 1(b) only occupies half of the unit circle that shape 1(a) occupies. As a result, less shape information is available for 1(b) compared to 1(a).



(a) (b)

Fig. 1. (a) Original image. (b) Skewed image within a unit circle as defined by shape (a).

Skew and stretch distortions are common among shapes obtained from natural objects. Therefore, it is important to have a descriptor which preserves more shape information and is robust to this type of distortions. Several efforts have been made to improve translation, rotation, and scale invariance properties [5, 1, 8, 9], but not much attention has been given in improving perspective invariance properties [10]. *Zhang* and *Lu* proposed a method which involves rotation and scale normalization. Due to this normalization, loss of some information is unavoidable, especially if the shape is reduced in size.

In this paper, we propose a new method for extracting Zernike moments addressing this issue. The rest of this paper is organized as follows. In Section 2, we briefly provide background information on Zernike moments. In Section 3, the proposed method is described in more detail. In Section 4, through experimental results, we demonstrate that Zernike moments with skew and stretch invariance properties – as extracted by our proposed method – are more efficient shape descriptors and can be used effectively as global shape descriptors in an image, especially in large databases. Finally, Section 5 provides the final thoughts and conclusions of the presented work.

2. ZERNIKE DESCRIPTORS

The Zernike moments introduce a set of complex polynomials which form a complete orthogonal set over the interior of a unit circle, *i.e.,* $x^2 + y^2 \leq 1$ [6]. In fact Zernike moments are the projection of the image function on some orthogonal basis functions. Let the set of these basis functions be denoted by $V_{n,m}(x, y)$. These polynomials are defined by [6]

$$V_{n,m}(x,y) = V_{n,m}(\rho,\theta) = R_{n,m}(\rho)e^{jm\rho} \quad (1)$$

where *n* is a non-negative integer, *m* is a non-zero integer subject to the following constrain: $n - |m|$ is even and $|m| < n$. Also, ρ is the length of the vector from origin to the (x,y) pixel, θ is the angle between vector ρ and *x*-axis in a counter-clockwise direction, and $R_{n,m}(\rho)$ is the Zernike radial polynomial. The Zernike radial polynomials, $R_{n,m}(\rho)$, are defined as [6]:

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

Note that $R_{n,m}(\rho) = R_{n,-m}(\rho)$. The Zernike moment of order *n* with repetition *m* for a digital image is [6]

$$Z_{n,m} = \frac{n+1}{\pi}\sum\sum_{x^2+y^2\leq 1} f(x,y)V^*_{n,m}(x,y)\Delta x\Delta y \quad (3)$$

where $V^*_{n,m}(x,y)$ is the complex conjugate of $V_{n,m}(x,y)$. To compute the Zernike moments of a given image, the image center of mass is taken as the origin.

3. PROPOSED ALGORITHM

Our proposed method fits an ellipse instead of a circle to an image object. Note that it is possible to fit different ellipses onto a shape, depending on the angle and ratio between the major axes. The ellipse which is fitted on the shape by the proposed method is unique due to its properties. This ellipse is taken as a transformed unit disk space and the Zernike moments are computed in this space. Finally, the moments are divided by the shape area.

Spatial moments are often used to describe the shape of a region. There are three second order geometric moments for a region. They are denoted by μ_{xx}, μ_{xy}, and μ_{yy} and defined as follows:

$$\mu_{xx} = \sum_x (x-\bar{x})^2 f(x,y). \quad (4)$$

$$\mu_{xy} = \sum_x\sum_y (x-\bar{x})(y-\bar{y}) f(x,y). \quad (5)$$

$$\mu_{yy} = \sum_y (y-\bar{y})^2 f(x,y). \quad (6)$$

These quantities are often used as simple shape descriptors, as they are invariant to translation and scale changes of a 2-D shape.

As a covariance matrix has special value and meaning for any two-dimensional probability distribution, similarly the second spatial moments have special values and meanings for different shapes. If the region is an ellipse, there is an algebraic meaning that can be given to the second spatial moments. If a region R is an ellipse whose center is on the origin, then R can be expressed as[11]:

$$R = \left\{(x,y)|d^2x^2 + 2exy + f^2y^2 \leq 1\right\}. \quad (7)$$

A relationship exists between the coefficients d, e, and f of equation (7) and the second moments μ_{xx}, μ_{xy}, and μ_{yy} given by[11]:

$$\begin{pmatrix} d & e \\ e & f \end{pmatrix} = \frac{1}{4\left(\mu_{xx}\mu_{yy} - \mu_{xy}^2\right)}\begin{pmatrix} \mu_{xx} & \mu_{xy} \\ \mu_{xy} & \mu_{yy} \end{pmatrix}. \quad (8)$$

Equation (8) shows that the second moments μ_{xx}, μ_{xy}, and μ_{yy} determine the orientation and the lengths of the major and minor axes of the ellipse [11]. This ellipse has the same second moments as the query shape. Figure 2 shows the letter “A” and its skewed and stretched images overlayed with the ellipse with the same second moments. Here, $2a$, $2b$ and θ denote the length of major axis, the length of minor axis and the orientation of the ellipse, respectively. To map the coordinate of the unit disk space

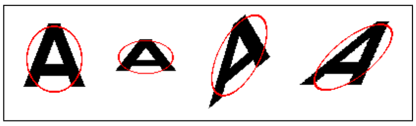
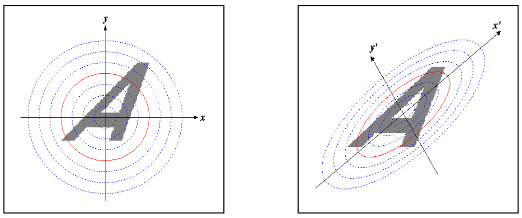


Fig. 2. Some homogenous shapes of letter “A” and the computed ellipses obtained using equation (8).

into the ellipse, first the origin of the unit disk is translated to the shape center and then the coordination system is rotated by θ. Finally, x and y axes are rescaled by factors of $\frac{1}{a}$ and $\frac{1}{b}$, respectively. Now, each object pixel location (x,y) will have a new location (x',y') given by:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{b} \end{bmatrix}\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix}. \quad (9)$$

Figure 3(a) shows the coordinate system (x,y) and Figure 3(b) shows the transformed coordinate system (x',y') using equation (9). To enclose the ellipse on the shape,



(a) (b)

Fig. 3. (a) Original shape in the (x,y) coordinat system. (b) Transformed shape in the (x’,y’) coordinate system, using equation (9). [The red curves indicate unit circle in each coordination.]

the location of each pixel is divided by the maximum distance d_{max} of the boundary points from the origin in the (x',y') coordinate system. Finally, each object pixel location (x',y') will have a new location $(x",y")$ given by:

$$\begin{bmatrix} x" \\ y" \end{bmatrix} = \begin{bmatrix} \frac{1}{d_{max}} & 0 \\ 0 & \frac{1}{d_{max}} \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix}. \quad (10)$$

Figure 4, shows some shapes of the letter “A” with their enclosed ellipse. The final features are normalized by dividing the Zernike moments in the $(x",y")$ coordinate system over the shape area.

4. EXPERIMENTAL RESULTS

In this section, to evaluate the performance of the proposed method, several experiments have been conducted. The experiments mainly indicate that the proposed method posses better skew and starching properties than other available methods. Our test image database consists of 100 binary images of size 100×100 pixels in the TIF format. The first 10 images are shown in Figure 5.

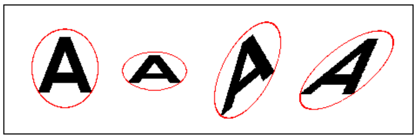


Fig. 4. Results of equation (10) applied on images shown in Figure 2.

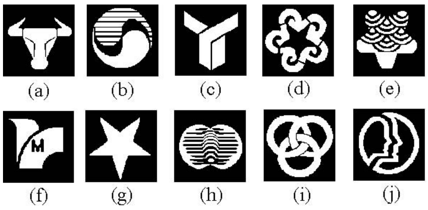


Fig. 5. The first 10 images of 100 images contained in our test database.

Two experiments have been used to evaluate the invariance properties of the proposed method towards rotation, scaling, stretching, and skew. Here, we considered rotation and scale as *Linear* invariance properties, and stretching and skew as *Non-linear* invariance properties. *Bin*'s method [7] shows better linear invariance properties than other reported methods. However, because of circular scanning shapes, this method can not handle non-linear properties as well. Another work by *Zhang* [10], although takes into consideration non-linear invariance properties on affine transformed shapes, still doesn't compare well to *Bin*'s method in terms of linear invariance properties.

In the first experiment, we evaluate linear invariance properties, by comparing the performance of the proposed method to *Bin*'s reported work [7]. In the second experiment, we evaluate non-linear invariance properties by comparing results of the proposed method, to *Zhang*'s reported work [10].
The first experiment can be described in the following steps:

1. Rotation and Scaling transformations have been applied to the test images. Each test image was rotated by 35^o, 75^o, 110^o, and 140^o and scaled by 60%, 90%, 120%, and 230%. As a result. a set of 8 transformed images was created for each test image.

2. Up to 10^{th} order Zernike moments of the test images and their transformed versions were computed by the proposed method and *Bin*'s method.

3. The error between each test image and its transformed image, was computed using following equation:
$$\varepsilon_r = \frac{|Z - Z'|}{\|Z\|} \qquad (11)$$
vectors of the test image and one of its transformed images, respectively.

As shown in figure 6 the mean error values of the Zernike moments obtained by the proposed method (red line) and *Bin*'s method (blue line) for the first 50 test images are close to each other. This indicates that our proposed method is comparable to *Bin*'s method in terms of rotation and scale invariance.

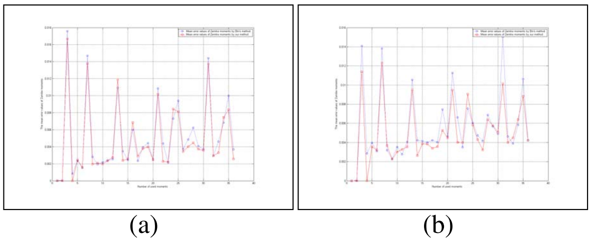


Fig. 6. Mean error values of ZMs in the first experiment for: (a) rotation, (b) scaling.

The second experiment consists of following steps:

1. Stretching and skew have been applied to the test images. Each test image has been stretched by 60%, 90%, 120%, and 230% and skewed by 35^o, 75^o, 110^o, and 140^o in the horizontal and vertical directions, separately. Finally, we applied stretching and skew together on each images in both directions. As a result, a set of 36 transformed images was created for each test image. In Figure 7, 5 of the 36 transformed versions of image 5(a) are shown.

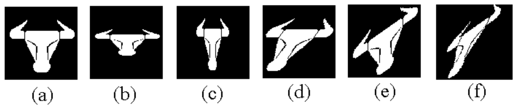


Fig. 7. (a) Original image. Stretched images by 60%: (b) in y axes, (c) in x axes. Rotated images by 35^o: (d) in x axes, (e) in y axes, (f) in both axis.

2. Up to 12^{th} order Zernike moments of the test images and their transformed versions were computed by the proposed method and *Bin*'s method.

3. Zernike moments were computed.

Figures 8 show the Zernike moments of images 5(c) as computed by the proposed method and *Zhang*'s method. Due to the tiny differences between the Zernike moments of the original image and its transformed images, computed by our method, we only show the Zernike moments of the original image and images stretched by 60% and skewed by 35^o in each direction.

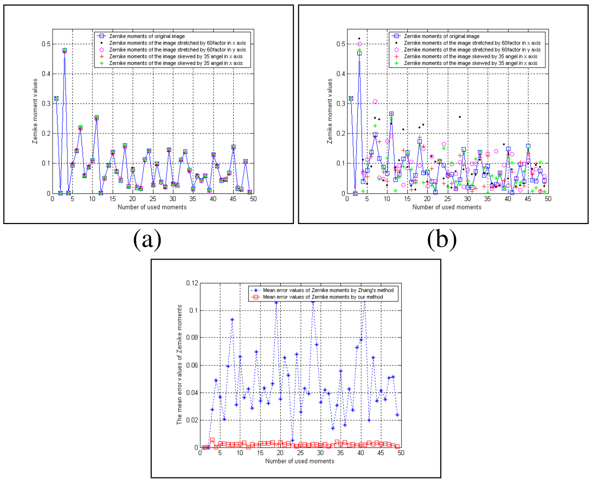
Figure 9 shows the mean error values of the Zernike moments obtained by our proposed method (red line) and

Zhang's method (blue line) for the whole database. As indicated by the comparative graph the mean error of the Zernike moments as obtained by our method is considerably lower than the resulting error of *Zhang*'s method.

5. DISCUSSION & CONCLUSION

Overall, the combined results of two experiments indicate that the proposed method produces more efficient shape descriptors for stretch and skew distortions than other approaches, without loosing any performance on rotation and scale changes.

There is, however, an additional computation cost during pre-processing. Before computation of Zernike moments, we need to compute up to second order geometry moments and transform pixel coordinates by a matrix. Therefore, the preprocessing step for a $N \times N$ image takes $O(N^2)$ time. Computation of Zernike moments up to order M for that image is $O(N^2M^2)$ [12]. This issue can be addressed by using a special hardware accumulation grid architecture, which can allow for efficient computation of any order Zernike moments in real time [13].



(a) (b) (c)

Fig. 8. Invariance of ZMs of image 5(c) by: (a) proposed method, (b) *Zhang*'s method. (c) Mean error of ZMs between original image and its transformed images.

To summarize, we proposed a new method to improve the stretch and skew invariance properties of Zernike moments. We have shown that extracting Zernike moments from a transformed unit disk space not only provide shape descriptors with better stretch and skew invariance properties than other methods, but also retains the rotation and scale invariance properties as well.

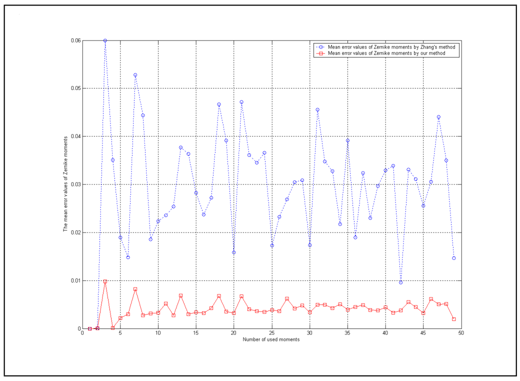


Fig. 9. Mean error of ZMs in the second experiment for all the images in our test database as computed by $Zhang$'s method (blue line) and by the proposed method (red line).

6. REFERENCES

[1] C.H.Teh and R.T.Chin. "on image analysis by the methods of moments.". *IEEE Trans. on PAMI.*, 10:496–513., 1988.

[2] D. S. Zhang and G. Lu. "content-based shape retrieval using different shape descriptors: A comparative study.". *IEEE Int. Conf. on Multimedia and Expo, Tokyo*, pages 317–320, 2001.

[3] H. Kim and J Kim. "region-based shape descriptor invariant to rotation, scale and translation.". *Signal Proc.: Image Communication*, 16:87–93, 2000.

[4] S. Belkasim, E. Hassan, and T. Obeidi. "radial zernike moment invariants". *The Fourth Int. Conf. on Computer and Information Tech.*, 1:790–795, 2004.

[5] E.M. Arvacheh and H.R. Tizhoosh. "pattern analysis using zernike moments.". *Instrumentation and Measurement Tech. Conf.*, 2:1574–1578, 2005.

[6] A.Khotanzad and Y.H. Hong. "invariant image recognition by zernike moments". *IEEE Trans. on PAMI.*, 12:489–498., 1990.

[7] Y. Bin and P. Jia-xiong. "improvment and invariance analysis of zernike moments". *IEEE Int. Conf. on Communications, Circuits and Systems and West Sino Expositions*, 2:963–967., 2002.

[8] A. Choksuriwong, H. Laurent, and B. Emile. "comparison of invariant descriptors for object recognition". *IEEE Int. Conf. on Image Processing*, 1:377–380, 2005.

[9] X. Yongqing, S. Liao, and M. Pawlak. "on the improvement of rotational invariance of zernike moments". *IEEE Int. Conf. on Image Processing*, 2:842–845, 2005.

[10] D. Zhang and G. Lu. "improving retrieval performance of zernike moment descriptor on affined shapes". *IEEE Int. Conf. on Multimedia and Expo*, 1:205–208., 2002.

[11] Haralick and Shapiro. *"Computer and Robot Vision."*, volume 1. Addison-Wesley 1992.

[12] G. Amayeh, A. Erol, G. Bebis, and M. Nicolescu. Accurate and efficient computation of high order zernike moments. *First Int. Sym. on Vision and Computation, NV, USA*, pages 462–469, 2005.

[13] L. Kotoulas and I. Andreadis. "real-time computation of zernike moments". *IEEE Trans. on Circuits and Sys. for Video Tech.*, 15:801–809, 2005.