RESEARCH ARTICLE OPEN ACCESS

Rotation and Scale Invariant Feature Extraction Using Complex Zernike Moments Forfarsiand Arabic Handwriting Character

Mahvashakbari[1], Maryam Khademi[2]
[1]Department of ComputerScience, Science and Research Branch, Islamic Azad University, Saveh, Iran
[2]Applied Mathematics Department, Islamic Azad University, South Tehran Branch, Tehran, Iran

Abstract
Analyzing Farsi and Arabic handwritten documents is one area in image processing whose target is to transform picture documents into symbolic form. This transformation is conducted o make rapid and easy saving, improvements, retrieval, reuse, searching and transferring documents. Analyzing documents is performed in five stages: pre-processing, segmentation representation, recognition and post-processing. In this research, in first stage the required pre-processing is performed to normalize the image. In recognition stage, zernik moments–based method has been introduced to extract feature of Farsi and Arabic handwritten characters. Outputs of zernik moments are put in systematic clustering to decide about characters. The obtained results show that feature extraction using zernik moments is a suitable method that deals with few rotation-independent featuresand this leads to reduce calculations and increase speed of recognition and stability against rotation. Size–independence is obtained using difference and at least distance of image calculations, because of preprocessing stage which has been performed in this algorithm is transfer-independent. Valid rank of zenik moments to extract features during conducted test is 4 – 38. Cluster applying has led to reduce algorithm expenditure tplog(n) and this one advantage of the suggested algorithm .

Keywords: Rotation and Scale Invariant Feature Extraction, zernik complex moments, Farsi and Arabic handwritten characters.

I. Introduction

An image is valued one thousand words is a quotation form Confucius, a Chinese philosopher in 2500 years ago. Todays, essence of this sentence is perceived completely. Visual data plays an important role in the society and this role will be increased and there will be intensive need to process these features in future. Pictures and images are applied in many application environments such as architecture and drawing, journalism, advertisement, amusement, etc.How to search and retrieve considered images quickly against nature and rapid increment of images is a basic consideration that requires and image – retrieval system. Visual properties of images provide a description from their contents. Content–based image retrieval (CBIR) is considered as an efficient and promising tool to retrieve images and search in the large image database. In recent years, CBIR has been student and its meaning is an image retrieval process from a set based on automatic feature extraction. Applicable and effective properties of an image should have some necessary features such as recognition capability,scaleinvarient, transfer and rotation, noise – resistance, statistical independence and reliability.

Generally, image descriptor is a set of numbers obtained to describe image properties. Correctness and accuracy of suitable retrieval needs an image descriptor that can retrieve similar images from a database significantly. Descriptors are often in a diagram form. Image descriptors should satisfy following requirements:

- Descriptors should be complete as much as possible to provide data items.
- Descriptors should be provided and save briefly.
- Calculation of distance between descriptors should be simple.

Optical characters recognition (OCR) techniques are performed for components of the documents image which has been identified as the content by structural analysis techniques and transform document image into a text which is editable by the computer.

Input image → pre-processing → segmentation → representation → recognition
post – processing → output text

Figure 1: block diagram of (OCR) system

It isn't possible to perform recognition methods of Latin texts directly to recognize Farsi texts.

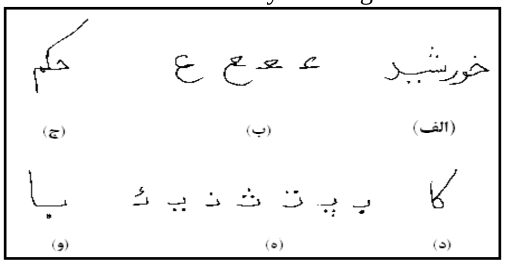


Figure 2: Farsi handwritten character features

In the view of computer–aided processing some Farsi and Arabic text features are as follows:

- ✓ In contract to Latin characters, Farsi and Arabic texts are written from right to left.
- ✓ In Farsi and Arabic words some characters are connected to their neighbor characters from one or both sides and some characters are written separately.
- ✓ Characters inside a word may cover each other. It means the characters can't be separated completely by drawing vertical lines.
- ✓ Some characters have one to three points which may be located above or under the character body.
- ✓ Characters that can be connected to their neighbor character may be written as elongated.

Providing an acceptable method to extract size and rotation invariant features is purpose of this study. This method can be used in digital libraries, for reading addresses, separating envelopes and postal packages based on their addresses, reading car's number plates in traffic control cameras, transforming old paper documents to digital texts to edit and search the content rapidly, deleting or reducing typist role and therefore increasing performance speed in organizations.

II. Image descriptors

Applied image descriptors techniques are divided in two groups:

Structural descriptors analyze character or word image and generate a property vector between high or low level of properties. On the other hand, image inside patterns and their mathematical properties can be expressed by mathematical descriptors as numbers. Many techniques are applied to compress images.

They try to present an image without explicit pixels determination. Applying fractals is based on this assumption that each image has redundancy degrees and we can find a similarity between their components and determine it[2]. Fourier transformation is based on cosine and sinusoidal waves and analyzes image signals to sinusoidal components in different frequencies. In this condition image is presented in frequency space instead of pixel space. By adding applied frequencies, presented image accuracy in frequency space will be increased. Therefore, information of one image can be used to reconstruct the image by looking at different parts of frequency. Wavelet transformation is in alternative method for Fourier transformation. It follows similar rules but created extra information in addition to frequency which scales amounts.

Reason of this call is that wavelets are small waves with non–zero area and gather in a central small region; therefore, they provide more practical description for generated data in this region [1]. Invariant moments are another group of base transformations. Large range of different moments has been developed which includes legendary, zernik and sub–zernik moments. Final candidates are Zernike moments. Comparative study about efficiency of moments different forms were performed. It has been shown that Zernike moments are preferable to legendary moments because of their noise tolerance and informational contents.[3]

In [8] Broumandnia and Shanbehzadeh used Zernike moments, wavelets and artificial nervous network to recognize Farsi characters independent of rotation and size. The best efficiency with 50 features of input nodes is 98.2%. Shan li and senior member in [9] used Zernike moments to extract image features independent of rotation and size by testing on data base MPEG-7CE-2 test. In this study we have tried to use both features Zernike moments phase and amplitude for independence from size and rotation.

C.Singh, E.Walia and N.Mital in [10] use Zernike moments to recognize grey level face and image of dual characters based on the rotation – independent feature. In this research database ORL has been used and face recognition rate 96.5% and recognition rate of Roman characters dual images 99.7% has been obtained.

III. Zernike moments

Information of this section was obtained from reference [4, 5, 6, 7].

Complex zernike moments are constructed using a set of complex polynomialswhich form a complete orthogonal basis set defined on the unit disc. Vertical feature reduced redundancy and increases accurate of image reconstruction from input moments that leads to advantage for this case.

To conduct moments first all coordinates in the considered region are fitted and polar coordinates will be used instead of Cartesian coordinates. I should be determined whether image is put in unit disc or unit disc covers the image. Farmer selection decreases reconstruction accuracy because empty area in outside of the image generates redundancy, whereas last selection leads to clip out a piece of the image and it may include some important image features. In figure (3) these two methods have been shown.

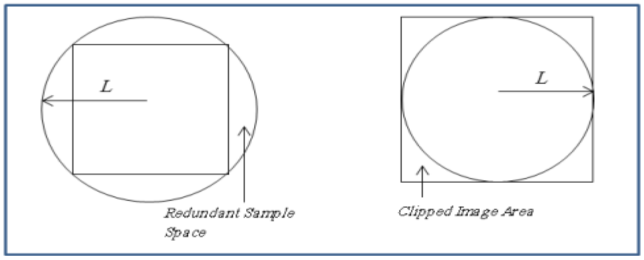


Figure 3: Putting disc inside the image and putting the image inside the disc.

Low order moments provide basic information about picture and high order moments provide more information about their details. Although zernike moment are rotation invariants, they are not characters transfer and size invariants. Calculating formula for an order (multinomial degree) zernik complex moment and angular independence m is as follows:

$$A_{nm} = \frac{n+1}{\pi}\int_x \int_y I(x,y)\bar{V}_{nm}(\rho,\theta)\delta x \delta y$$

In which I(x, y) is pixel intensity function. Its value for dual back and white images is 0 and 1, respectively. Also in above formula, V is basic function ρ , θ are previous coordinates of x and y and following limitations exit:

- $N \geq o$
- $N \geq |m|$
- N - |m| is an even number
- $X^2 + y^2 \leq 1$

Valid orders are obtained by testing differentn. Note that V is the complex conjugate of V.
Discrete form for equation (2) is obtained by transforming integral to sum for all pixels.

$$A = \frac{n+1}{\pi}\sum_{k=0}^{k<maxX}\sum_{j=0}^{j<maxY} I(k,j)\bar{V}_{nm}(\rho,\theta)$$

$V_{nm}(\rho, \theta) = R_{nm}(\rho)e^{im\theta}$

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

If we rotate image calculate moments we find that value A for rotated image is the some pervious transferring stage:
$A'_{nm} = A_{nm}$

It means $|A_{nm}| = |A_{n-m}|$; therefore, moment sizes can be used as a rotation – invariant property for image intensity function.

Another property is $|A_{nm}| = |A_{n-m}|$, it means we can ignore moments with negative m values in generating property vector.
An image has many points or pixels which are considered as elements from image input matrix. In this images with 200 * 200 pixels were used. Therefore, there are 4096 inputs for one image. Using an input with this large size makes clustering hard and volume of calculation increases very much.

Zernik moments are used an alternative method to present the image in smaller dimensions. In mathematics, *Zernike* polynomials are a *sequence* of polynomials that are orthogonalon unit disc and were established by fritz zernik and play an important role in image processing. Zernikepolynomials are either odd or even. Even polynomial is defined as equation (6):

$$Z_n^m(r,\theta) = R_n^m(r)\cos(m\theta)$$

And old polynomial is calculated using equation (7):

$$Z_n^m(r,\theta) = R_n^m(r)\sin(m\theta)$$

In which n, m are non – negative integral numbers and n ≥ m. θ is azimuthal angle and r is radial distance (0 ≤ r ≤ 1). Radial polynomial R_n^m is defined as equation (8):

$$R_n^m(r) = \sum_{k=0}^{(n-m)/2} \frac{(-1)^k}{k!\left(\frac{n+m}{2}-k\right)!\left(\left(n-\frac{m}{2}-k\right)!\right.} r^{n-2k}$$

(n –m) is an add number and these radial polynomial is zero for add (n – m).
Radial multinomial can be written as equation (9):

$$R_n^m(r) = \sum_{k=0}^{(n-m)/2} \frac{(-1)^k (n-k)!}{k!\left((n=m)\backslash 2-k\right)!\left(\frac{n-m}{2}-k\right)!} r^{n-2k}$$

Vertical property of radial components is as equation (10):

$$\int_0^1 \sqrt{2n+2\ R_n^m}(r)\sqrt{2n+2R_{n'}^m}(r)dr = \delta_{n,n'}$$

It is some Dirac delta. Being vertical for δn and n′ angular part is as equations (11) – (13):

$$\int_0^{3\pi} \cos(m\theta)\sin(m'\theta)\, d\theta = \varepsilon_m \pi\, \delta|m||m'|$$

$$\int_0^{2\pi} \sin(m\theta)\sin(m'\theta)\ d\theta = (-1)^{m+m'}\pi\, \delta_{|m\ ||m'|}\ \ , m \neq 0$$

$$\int_0^{2\pi} \cos(m\theta)\sin(m'\theta) d\theta = 0$$

In which ε_m (sometimes is called neuman factor) is 2 for m = 0 and is 1 for m≠ 0. Product of radial and angular components shows that zernik functions are orthogonal to both indexes when they are integrated on unit disc:

$$\int z_n^m\ (r,\theta) z_{n'}^{m'}(r,\theta) d^2 r = \frac{\epsilon_m\ \pi}{2n+n}\delta_{n,n'}.\,\delta_{m,m'}$$

Zernik moments are symmetric and their symmetry is as reflection along x axis as equation (15):

$$Z_n^m(r,\theta) = (-1)^m Z_n^m(r,\theta)$$

Symmetry to coordination center is also as equation (16):

$$Z_n^m(r,\theta) = (-1)^m Z_n^m(r,\theta+\pi)$$

Zernik moments are defined based on zernikpolynomials and are as equation (17):

$$A_{nm} = \frac{n+1}{\pi}\int_0^{2\pi}\int_0^1 V_{nm}\ (r,\theta) f(r,\theta) r\ dr\ d\theta$$

That $n - |m|, |m| \leq n, r \leq 1$ is an even number. Zernikpolynomials V_{nm} (r,θ) is defined based on radial functions:

$$V_{nm}(r,\theta) = r_{nm}\, re^{+jm\theta}$$

IV. Suggested algorithm

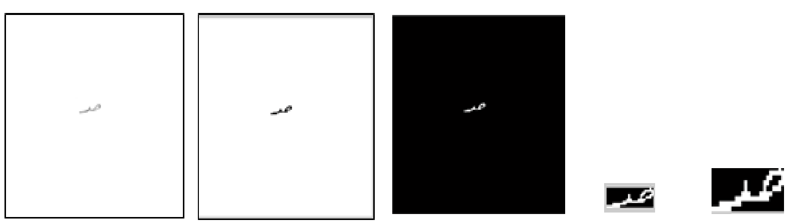
To extract features, first image will be normalized. In this research firs loaded image will be transformed to a grey scale image(if the scanned characters image is colored)scanned images are 200 * 200 pixels. Next stages are as follow:

a) Histogram modulation: if the image is scanned, contrast of the considered image will be low. To improve light intensity distribution histogram modulation will be applied.

b) Reversed zernik: to identify characters important parts of the image are determined by black lines and background is white. If standard intensity function is applied instead of characters descriptions with their lines it is obvious principally where these characters don't exist. It is better to provide some kind of zernik moments which show details of black color. Achieve this intensity function will be reversed. In this case white pixels are shown with 0 and 1 respectively. Another advantage of this method is that black pixels of the image are much less than white pixels. In other words, basic function are called just for non-zero intensity function, hence, calculations will be reduced and take less time.

c) We perform two more works to prepare main body image. First whole image redundant parts are removed and only one window which includes characters is remained. This remaining box is fitted on the character main body completely. In next stage remaining character body from previous stage is put in a square image such that it is in center of this body and in middle of new image.By this work whole scanned characters are written in the middle of image and in zernik calculation which is not transfer-invariant calculation errors and comparison mistakes will decrease.

In figure (4) all stages of image preparation are shown for character "ص "



a) 200*200 pixel b) 200*200 pixel c) 200*200 pixel d) 24*9 pixel e) 64*64 pixel

To be size independent,color intensity matrixes difference of the image is used. Experimental image and training images are inputs of this program. This process finds images with minimum difference. Pre-processing stage output will be used as input for zernik moments calculation. In training procedure the clustering is used as retrogression. Inputs of these clusters are zernik moments outputs. In experimental stage pioneer search in clusters is performed.

V. Clustering inputs and their analysis

After performing pre-processing stages the feature vector is sent to calculation zernik momentsorder is considered 4 – 38. Zernik moments output are saved as retrograde clustering in a matrix. To fill clusters we perform as follows: the matrix first column corresponding clustering includes each image indexes. Resultant features vector from zernik moments fills matrix second column. From third column to next two neighbor lines mean in column k – 1 will be put in column k such that each item in the column k – 1 only one time has been involved in making average.

This process will continue until last column just one non-zero item.The number of matrix lines is n and equal training inputs number and number of matrix column is log(n) + 1. Clustering in used because of low expenses of O(logn) in recognition stage. The more number of training inputs, the more recognition speed and expenses in relation to inputs will decrease. In experimental stage, clusters will be investigated as pioneer and suitable case will be selected.

Generally, effective parameters in training stages are as follows:

1. The more number of training patterns, the more training accuracy.
2. Selection range for considered zernik moment order is very effective in recognition accuracy

Various experiments have suggested that if this order increases more than acertain limit (order 38) recognition will not performed correctly.

An example for clustering performance for 16 raining inputs has shown in figure (5):

7.0000	0.0097	0.0129	0.0172	0.0291	0.0461
12.0000	0.0160	0.0216	0.0410	0.0631	0
2.0000	0.0207	0.0345	0.0578	0	0
8.0000	0.0225	0.0476	0.0685	0	0
6.0000	0.0339	0.0542	0	0	0
9.0000	0.0351	0.0614	0	0	0
4.0000	0.0460	0.0662	0	0	0
11.0000	0.0492	0.0708	0	0	0
3.0000	0.0527	0	0	0	0
10.0000	0.0556	0	0	0	0
16.0000	0.0596	0	0	0	0
5.0000	0.0631	0	0	0	0
15.0000	0.0656	0	0	0	0
1.0000	0.0668	0	0	0	0
13.0000	0.0706	0	0	0	0
14.0000	0.0710	0	0	0	0

Figure 5: An example for clustering performance for 16 raining inputs

VI. Statistical analysis of research data and suggested algorithm

In this section, results obtained from statistical analysis by software SPSS is provided. “n” is zernik moments order, elapstime in calculation time for zernik moments and AOH is feature vector for zernik moments. In table(1) as significant level shows value related to “elapstime” and “AOH” have normal distributions. Therefore, relation between "n" and these two variant are states using person correlation coefficient.

Based on information in table (2) relation between two variants is significant and negative at 0.01 level i.e. by increasing “n”and “AOH” value decreases.

Table (1): normality test

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Elapstime	.105	18	.200*	.948	18	.399
AOH	.200	18	.055	.838	18	.006

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Results in table (3) show that there is a positive and significant correlation at 0.01 level. Therefore, we can claim certainly that the more “n” increases, the more “Elapstime” will increase.

Table (2): correlation between “n” and “AOH”

			n	AOH
Spearman's rho	n	Correlation Coefficier	1.000	-.973**
		Sig. (1-tailed)	.	.000
		N	18	18
	AOH	Correlation Coefficier	-.973**	1.000
		Sig. (1-tailed)	.000	.
		N	18	18

**. Correlation is significant at the 0.01 level (1-tailed).

Table (3): correlation between “n” and “Elapstime”

			n	Elapstime
Spearman's rho	n	Correlation Coefficient	1.000	.998**
		Sig. (1-tailed)	.	.000
		N	18	18
	Elapstime	Correlation Coefficient	.998**	1.000
		Sig. (1-tailed)	.000	.
		N	18	18

**. Correlation is significant at the 0.01 level (1-tailed).

VII. Conclusion

Many tests are performed on the system and generally increasing number of experimental database several tens or several hundred folds can increase characters recognition accuracy significantly. Because of pre-processing stage processes algorithm is size, and transfer – independent in addition to rotation – independent.

During conducted tests, valid order for zernik moments to extract feature is 4 – 38. Using clustering in last stage will decrease algorithm expenses to log(n) .

Statistical analysis of result of the considered algorithm performancesuggests that there is a significant and positive correlation between “n” and “Elapstime” at 0.01 level. Therefore, we can claim certainly that the more “n” value increases, the more “Elapstime” value will increase. There is a negative and significant correlation between “n” and “AOH” at 0.01 level i.e. by increasing “n” value “AOH” value will decrease.

VIII. Future works

The authors suggest that togenerate a database with suitable quality and apply more manuscripts with more inclusion to develop this method and also this method can be applied to identify word inside a text.

References

[1] C.S. Wallace, “A Suggestion for a Fast Multiplier”, *IEEE Transaction on Electronic Computers*, (1964), Vol. EC-13 14-17.

[2] L. Dadda, “Some Schemes for Parallel Multipliers”, *Alta Frequenza*, (1965), Vol. 34 349-356.

[3] Z.JMou, F. Jutand, “Overturned-Stairs Adder Trees and Multiplier Design”, *IEEE Transaction on Computers*, (1992), Vol. 41 940-948.

[4] S.Veeramachaneni, K.M. Krishna, L.Avinash, S.R.Puppala, M.B. Srinivas, “Novel Architectures for High-Speed and Low-Power 3-2, 4-2 and 5-2 Compressors”, *IEEE*, 20th *International Conference on VLSI Design,* (2007), 324-329.

[5] D.Baran, M.Aktan, V.G. Oklobdzija, “Energy Efficient Implementation of Parallel CMOS Multipliers with Improved Compressors”, *Proceedings of the 16th ACM/IEEE international Symposium on Low-Power Electronics and Design*, (2010), 147 – 152.

[6] A. Author, B. Author, BOOK TITLE, Publisher, 1990.

[7] http://www.nangate.com/, accessed November 17, 2011. ecognition،Academic Press، 1st edition، January 15، 1999.

[1]. A. Mowlaei.Kt.Faez , a.Haghighat, Feature Extraction with Wavelet Transform for Recognition of Isolated Hamdwritten Farsi / Arabic Characters and Numerals Digital Signal Processing , pp. 923-926,2004.

[2]. D.Neary ,"Fractal methods in Image Analysis and Coding", City University , School of Electronic Engineering , 2000.

[3]. H.s.Kim, H.K. Lee," Invarianat Image Watermark Using Zernik Moments," IEEE Tramsactions On. Circuits And Systens For video Technology ,Vol . 8,pp.766-773,2003.

[4]. S.O.BelkasimM.S"Efficient Algorithm For Fast Computation of Zernike Moments",IEEE 39 th Midwest symposium on Circuits and systems 3 ,pp.1401-1404,1966.

[5]. A.Khotanzad ,Y.H.Hong," Invariant Image Recognition by Zernik Moments," IEE Trans .Pattern Analysis Machine Intelligence , Nr.5 , pp.489-497-1990.

[6]. I.R.Rabiner," A tutorial on hidden markov models and selected applications in speech recoghition", Reading in Speech Recognition , PP.267-296, 1990.

[7]. C.W Chong ,P.Raveendran , R.Mukandan. Translation, 36 , S, 1765-1773 , 2003.

[8] Broumandnia Ali , ShanbehzadehJamshid , Fast Zernike wavelet moments for Farsi character recognition, Elsevier, accepted 16 May 2006,Image and Vision Computing 25 (2007) 717–726

[9] Shan Li, Senior Member, " Complex Zernike Moments Features for Shape-Based Image Retrieval", IEEE, Moon-Chuen Lee, and Chi-Man Pun

[10] C. Singh, E. Walia, N. Mittal, "Rotation invariant complex Zernike moments features and their applications to human face and character recognition", IET Computer Vision. ISSN 1751-9632, April 2011.