Available online at www.sciencedirect.com

ScienceDirect

Image and Vision Computing 25 (2007) 717–726

www.elsevier.com/locate/imavis

Fast Zernike wavelet moments for Farsi character recognition

Ali Broumandnia [a,*], Jamshid Shanbehzadeh [a,b]

[a] Islamic Azad University, Science and Research Branch, Tehran, 14515/755, Iran
[b] Department of Computer Engineering, Tarbiat Moalem University, Tehran, Iran

Received 30 September 2005; received in revised form 7 March 2006; accepted 16 May 2006

Abstract

Farsi character recognition (FCR) systems perform recognition of Farsi documents. This paper presents a novel approach of fast Farsi character recognition based on fast zernike wavelet moments and artificial neural networks. Fast Zernike wavelet moments and artificial neural networks are employed in feature extraction and classification, respectively. A simulation result shows superiority of novel scheme over similar ones in terms of precision 4.37 times in average, and improves recognition speed by about 8.0 times in average.

Keywords: Wavelet moments; Zernike moments; Character recognition; Character segmentation; Neural networks; Fast feature extraction

1. Introduction

The subject of text recognition has been receiving considerable attention in recent years due to increased demand on digital data and the need to transform old paper form data into digital format. Several methods for recognition of Latin, Chinese, and Kanji documents have been proposed. This research area in Farsi and Arabic is still hot and demanding. The Farsi and Arabic alphabets have 28 common characters and the Farsi script has four more characters. The specifications of words in these two languages, such as shape, dots and different fonts, make direct application of methods developed for other languages impossible. The shape of a Persian/Arabic character is a function of its location within a word, where each character can have two to four different forms. This increases the classes to be recognized from 32 alphabet character (28 for Arabic) to 114 (100 for Arabic) as shown in Table 1. As a result Persian character set is divided to isolate and cursive. As shown in Table 1, the number of isolated character set and cursive character set are 32 and 82, respectively.

Researches concerned with Persian/Arabic character recognition have been described in the literatures [1–8]. These techniques were applied for the automatic recognition of printed and handwriting Persian/Arabic characters. Despite so many different approaches proposed, most of them assumed that the character images have the same orientation and font and most of them worked on the isolated character set, only 32 characters. However, this assumption is not realistic for most practical application. For example, if images are obtained from scanning photographs, they are usually subject to certain random skew angles. Also most of Persian/Arabic character set is not isolated.

In this paper, we proposed effective schemes for fast feature extraction and rotation invariant Persian character recognition. These proposed approaches worked for all Persian character set, isolate and cursive, with different rotation, scale and font. The recognition rate and precision are improved versus other methods by proposed approaches.

The outline of this paper is organized as follows: the next section introduces Zernike and wavelet moments theory and proposed fast feature extraction. Sections 3 and 4 introduce two other rotation invariant approaches. Section 5 explores the scheme of the Farsi character segmentation. Section 6 describes Farsi character recognition and the final section presents experiments and results.

* Corresponding author. Tel.: +98 21 22369081.
E-mail address: broumandnia@azad.ac.ir (A. Broumandnia).

0262-8856/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2006.05.014

Table 1
Groups of Farsi letters

Group 1 Isolated Letters				Group 2 Beginning of Word Letters				Group 3 Middle of Word Letters				Group 4 End of Word Letters			
ت	پ	ب	ا	تـ	پـ	بـ		ـتـ	ـپـ	ـبـ		ـت	ـپ	ـب	ـا
ح	چ	ج	ث	حـ	چـ	جـ	ثـ	ـحـ	ـچـ	ـجـ	ـثـ	ـح	ـچ	ـج	ـث
ر	ذ	د	خ				خـ				ـخـ	ـر	ـذ	ـد	ـخ
ش	س	ژ	ز	شـ	سـ			ـشـ	ـسـ			ـش	ـس	ـژ	ـز
ظ	ط	ض	ص	ظـ	ظـ	ضـ	صـ	ـظـ	ـطـ	ـضـ	ـصـ	ـظ	ـط	ـض	ـص
ق	ف	غ	ع	قـ	فـ	غـ	عـ	ـقـ	ـفـ	ـغـ	ـعـ	ـق	ـف	ـغ	ـع
م	ل	گ	ک	مـ	لـ	گـ	کـ	ـمـ	ـلـ	ـگـ	ـکـ	ـم	ـل	ـگ	ـک
ی	و	ه	ن	یـ		هـ	نـ	ـیـ		ـهـ	ـنـ	ـی	ـو	ـه	ـن

2. Zernike and wavelet moments

In a character recognition system, the objects of interest are often represented by a set of numerical features with a goal to remove redundancy in the data and to reduce its dimensions. Moreover, the extracted features are expected to be invariant under translation, scaling and rotation to suit for different viewing directions. Rotation invariant and scaling invariant can be achieved using the Zernike and wavelet moments [9–15]. To get rotation invariant moments, typically the following generalized expression is used:

$$F_{pq} = \int\int f(r,\theta) g_{pq}(r) e^{\hat{j}q\theta} r \, dr \, d\theta, \quad \hat{j} = \sqrt{-1}, \tag{1}$$

where F_{pq} is the pq-order moment, $g_{pq}(r)$ is a function of radial variable r, and p and q are integer parameters. It is easy to prove that the value of $\|F_{pq}\|$ is rotation invariant and the combined moments, such as $||F_{p_1q} \cdot F_{p_2q}||$, are also rotation invariant, where $||x|| = \sqrt{x \cdot x^*}$ and symbol $*$ denotes conjugate of complex number. The proof of the rotation invariant property of $\|F_{pq}\|$ can be briefly given as follows. If an image object $f(r,\theta)$ is rotated by an angle ofβ, its corresponding moment will become $F_{pq}^{\text{Rotated}} = F_{pq}e^{\hat{j}q\beta}$. Since $||F_{pq}^{\text{Rotated}}|| = \sqrt{F_{pq}^{\text{Rotated}} \cdot (F_{pq}^{\text{Rotated}})^*} = ||F_{pq}||$, the rotation invariant property of $\|F_{pq}\|$ is thus proven.

The kernel of Zernike moments is a set of orthogonal Zernike polynomials defined over the polar coordinate space inside a unit circle. The two-dimensional Zernike moments of order p with repetition q of an image intensity function $f(r,\theta)$ are defined as follows [1]:

$$Z_{pq} = \frac{p+1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} V_{pq}^*(r,\theta) f(r,\theta) r \, dr \, d\theta, \quad |r| \leqslant 1, \tag{2}$$

where Zernike polynomials of order p with repetition q, $V_{pq}(r,\theta)$ are defined as follows:

$$V_{pq}(r,\theta) = R_{pq}(r) e^{\hat{j}q\theta} \tag{3}$$

and the real-valued radial polynomial $R_{pq}(r)$ is given as follows:

$$R_{pq}(r) = \sum_{k=0}^{\frac{p-|q|}{2}} (-1)^k \frac{(p-k)!}{k!(\frac{p+|q|}{2}-k)!(\frac{p-|q|}{2}-k)!} r^{p-2k}, \tag{4}$$

where $0 \leqslant q \leqslant p$ and $p - |q|$ is even. The wavelet moments are defined as follows:

$$W_{mnq} = \int_{\theta=0}^{2\pi} \int_{r=0}^{1} f(r,\theta) \psi_{mn}(r) r \, dr \, d\theta, \quad |r| \leqslant 1, \tag{5}$$

where the wavelet function defined along a radial axis in any orientation is denoted by

$$\psi_{mn}(r) = 2^{m/2} \psi(2^m r - 0.5n). \tag{6}$$

The mother wavelet $\psi(r)$ of the cubic B-spline in Gaussian approximation form [2] is defined as follows:

$$\psi(r) = \frac{4a^{n+1}}{\sqrt{2\pi(n+1)}} \sigma_w \cos(2\pi f_0(2r-1)) \exp(-\frac{(2r-1)^2}{2\sigma_w^2(n+1)}). \tag{7}$$

Since wavelet moments and Zernike moments are defined in terms of polar coordinates (r,θ) with $|r| \leqslant 1$, their computation requires a linear transformation of the image coordinates $(i,j), i = 0,1,2,\ldots,R-1, \quad j = 0,1,2,\ldots, C-1$ to a suitable domain $(x,y) \in R^2$ inside a unit circle.