INVARIANT FEATURES FOR CHARACTER RECOGNITION

Ryszard S. Choraś
Institute of Telecommunications, University of Technology & Agriculture
Kaliskiego Street 7, 85-796 Bydgoszcz, Poland

Keywords: Feature extraction, character recognition, moment invariants.

Abstract: This paper presents feature extraction method for recognition of isolated characters. Feature extraction is most important factor in achieving high recognition performance. We presented moments invariants as features for pattern recognition. This article analyzes the image feature extraction task on the basis of moments invariants for image recognition problem.

1 INTRODUCTION

Handwritten recognition have been a main research subject in pattern recognition for over thirty years. The application of handwritten character recognition is broad. Typical uses include recognition of handwritten zip codes and reading personal bank checks. The recognition of handwritten characters, like other problems in pattern recognition, consists of two major problems: feature selection and pattern classification. Feature selection is problem-dependent and considered most significant to the final result of a recognition system. Since handwritten characters of the same character class can occur in great variety, it is desirable to generate a representation that is invariant. A feature-based recognition of objects which is independent of their position, size, orientation and other variations has been the goal of much recent research. There have been several kinds of features used for recognition:

visual features (contours, textures),

transform coefficient features,

statistical features (moment invariants).

The objective of this paper is to develop an algorithm for the automatic recognition of handwritten characters. In the algorithm a binary image of the character is obtained and its skeleton is then produced by utilizing a standard thinning algorithm. The classification process incorporates features of the characters such as number of intersections, number of free ends, as well as criteria derived from normalized moments. Moment invariants derived by (Hu, 1999),(Flusser, 1993) are invariant only under translation, rotation and scaling of the object. In this paper we use features which are also invariant under general affine transformations. The block structure of the recognition system is given in Figure 1. The preprocessing stage includes noise filtering, thinning, and character categorization. The feature extraction and classification are described in the next paragraphs.

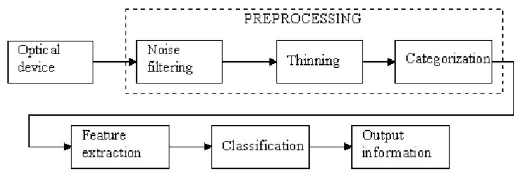

Figure 1: Block diagram of the proposed system.

2 PREPROCESSING

A character pattern is fed to a camera system, and the camera signal is transformed into binary image. The camera must be perpendicular to the character plane and contrast between characters and background must be reasonable. A binary image is regarded as a set of image points (pixels with value 1).

Choraś R. (2004).
INVARIANT FEATURES FOR CHARACTER RECOGNITION.
In *Proceedings of the First International Conference on Informatics in Control, Automation and Robotics*, pages 102-107
DOI: 10.5220/0001139201020107

The pictorial information is represented as a function of two variables (i, j). The image in its digital form is usually stored as an two-dimensional array. If $M = \{1, 2, \ldots, m\}$ and $N = \{1, 2, \ldots, n\}$ are the spatial domains, then $D = M \times N$ is the set of resolutions cells and the digital image P is a function which assigns binary value to each and every resolution cells, i.e. $P : M \times N \longrightarrow B$.

Noise pixels add irregularities to the outer boundary of the characters and may have undesired effects on the recognition system. A smoothing algorithm eliminates small areas and fills little holes. The algorithm modifies each pixel according to its initial value and to those of its neighborhood (Figure 2) according to the following conditions:

$$\text{If } \ p = 1 \ \text{ then } \ p' = \begin{cases} 0 \ \text{ if } \ \sum_{i=1}^{8} p_i \leq T_1 \\ 1 \ \text{ otherwise} \end{cases}$$

$$\text{else } \ p' = \begin{cases} 1 \ \text{ if } \ \sum_{i=1}^{8} p_i \rangle T_2 \\ 0 \ \text{ otherwise} \end{cases} \qquad (1)$$

where p is current pixel value, p' the new pixel value and and T_1 are T_2 the threshold values.

p 4	p 3	p 2
p 6	p	p 1
p 6	p 7	p 8

Figure 2: Pixel notation

2.1 Thinning of the characters

A pixel is considered deletable if it satisfies the following conditions:

1. $1 < B(p) \leq 7$ where $B(p) = \sum_{i=1}^{8} p_i$
2. $A(p) = 1$ is number of 01 transitions in the eight neighbors of pixel p.

Additional conditions may permit the removal of element if it is on south or east edge or if is on a corner

3. $(p_1 = 0)$ or $(p_7 = 0)$ or $(p_3 = 0$ and $p_5 = 0)$

Table 1: Topological properties of p

THE VALUE OF N_C^4 OR N_C^8	0	1	2	3	4
PROPERTY OF PIXEL p	Internal or isolate	End	Connect	Branch	Cross

The next conditions may permit the removal of element if it is on north or west edge or if it is on a corner

4. $(p_3 = 0)$ or $(p_5 = 0)$ or $(p_1 = 0$ and $p_7 = 0)$

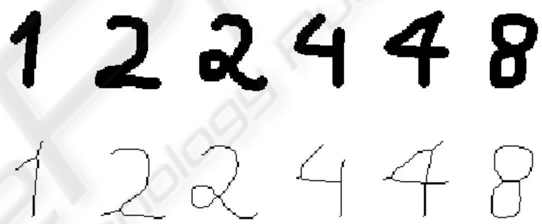

Figure 3: Original a) and thinned b) characters

2.2 Character categorization

A particular difficulty encountered in handwritten characters classification is the large variations in total properties of patterns. For character categorization we used information derived from connected number of point p .When $p = 1$, the connected number N_c of p is defined by the next equation

$$N_C^4 = \sum_{k \in S} (p_k - p_k p_{k+1} p_{k+2}) \qquad (2)$$

$$N_C^8 = \sum_{k \in S} (\bar{p}_k - \bar{p}_k \bar{p}_{k+1} \bar{p}_{k+2}) \qquad (3)$$

where: $S = (1, 3, 5, 7)$ and $\overline{p}$ means $(1 - p)$.

Topological properties of the pixel p are shown in Table 1, and the distribution of characters into the categories is shown in Table 2.

Table 2. Distribution of characters into the categories.

Category code c	Character									
	0	1	2	3	4	5	6	7	8	10
000	0									
200		1 1	2	3	4 4	5	6	7	8	9
400					4 4	5		7 7		
110	0		2		4		6		8	9
310		1 1	2	3	4	5 5	6	7		9
201			2	3	4			7	8	9
401			2		4			7		
220					4		6			9
420		1			4	5		7		

3 MOMENT-BASED CHARACTER CLASSIFIERS

3.1 Geometric Moments

Any character can be represented by the spatial moments of its intensity function

$$m_{pq} = \int\int F_{pq}(i,j)p(i,j)didj \tag{4}$$

where $p(i,j)$ is the intensity function representing the image, the integration is over the entire image and the $F(i,j)$ is the same function of i and j for example $i^p J^q$, or a $sin(ip)$ and $cos(jq)$. In the spatial case

$$m_{pq} = \sum_{i=1}^{m}\sum_{j=1}^{n} i^p j^q p(i,j) \tag{5}$$

The characteristic function of $f(x,y)$ is defined as its conjugate Fourier transform and may be expanded as a power series in u, v , as the following

$$F(u,v) = \int\int e^{-i2\pi(xu+yv)}dxdy = \\ = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(-i2\pi)^{p+q}}{p!q!}m_{pq}u^p v^q \tag{6}$$

The infinite set of moments $m_{pq}, \; p,q = 0,1,2,\ldots$ uniquely determine $f(x,y)$, and vice-versa

$$f(x,y) = \sum_{u}\sum_{v} e^{i2\pi(xu+yv)}[\sum_{p}^{\infty}\sum_{q}^{\infty}\frac{(-i2\pi)^{p+q}}{p!q!}u^p v^q] \tag{7}$$

The central moments are given by

$$m_{pq} = \sum_{i=1}^{m}\sum_{j=1}^{n}(i-I)^p(j-J)^q p(i,j) \tag{8}$$

where (I,J) are

$$I = \frac{m_{10}}{m_{00}} \quad \text{and} \quad J = \frac{m_{01}}{m_{00}} \tag{9}$$

Normalized central moment μ_{pq}

$$\mu_{pq} = \frac{m_{pq}}{(m_{00})^{\alpha}}, \quad \alpha = \frac{p+q}{2}+1 \tag{10}$$

Using nonlinear combinations of the lower order moments, a set of moment invariants (usually called geometric moments), which has the desirable properties of being invariant under translation, scaling and rotation, is derived. Hu (Hu, 1999) employed seven moment invariants, that are invariant under rotation as well as translation and scale change, to recognize characters independently of their position size and orientation.

$$\begin{aligned}
\phi_1 &= \mu_{20} + \mu_{02} \\
\phi_2 &= [\mu_{20} - \mu_{02}]^2 + 4\mu_{11}^2 \\
\phi_3 &= [\mu_{30} - 3\mu_{02}]^2 + [3\mu_{21} - \mu_{03}]^2 \\
\phi_4 &= [\mu_{30} + \mu_{12}]^2 + [\mu_{21} + \mu_{03}]^2 \\
\phi_5 &= [\mu_{30} - 3\mu_{12}][\mu_{30} + \mu_{12}]\times \\
&\times[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + \\
&+[3\mu_{21} - \mu_{03}][\mu_{21} + \mu_{03}]\times \\
&\times[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\
\phi_6 &= [\mu_{20} - \mu_{02}][(\mu_{30} + \mu_{12})^2 - (\mu_{21} \\
&+\mu_{03})^2] + 4\mu_{11}[\mu_{30} + \mu_{12}][\mu_{21} + \mu_{03}] \\
\phi_7 &= [3\mu_{21} - \mu_{03}][\mu_{30} + \mu_{12}]\times \\
&\times[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\
&-[\mu_{03} - 3\mu_{12}][\mu_{21} + \mu_{03}]\times \\
&\times[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]
\end{aligned} \tag{11}$$

Any function of moments which is invariant under the general affine transformation

$$\begin{aligned} i' &= a_{11}i + a_{12}j + a_{01} \\ j' &= a_{21}i + a_{22}j + a_{02} \end{aligned} \tag{12}$$

is invariant under simple six transformations
1) $j' = j \quad i' = i + a$ 2) $i' = i \quad j' = j + b$
3) $i' = wi \quad j' = wj$
4) $j' = j \quad i' = di$ 5) $j' = j \quad i' = i + tj$
6) $i' = i \quad j' = j + t'i$

The central moments are invariant under the translation 1 and 2. Normalized central moments are invariant under the scaling 3. We are interested in finding moments invariants also under other transformations e.g. one-axis scaling 4 and skew transformation 5 and 6. Next combinations will provide compound moments that support recognition of the characters

$I_1 = \mu_{20}\mu_{02} - \mu_{11}^2$
$I_2 = (\mu_{30}\mu_{03} - \mu_{21}\mu_{12})^2 - 4(\mu_{30}\mu_{12} - \mu_{21}^2)(\mu_{21}\mu_{03} - \mu_{12}^2)$
$I_3 = \mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2)$
$I_4 = \mu_{30}^2\mu_{02}^3 - 6\mu_{30}\mu_{21}\mu_{11}\mu_{02}^2 + +6\mu_{30}\mu_{12}\mu_{02}(\mu_{11}^2 - \mu_{20}\mu_{02}) + +\mu_{30}\mu_{03}(6\mu_{20}\mu_{11}\mu_{02} - 8\mu_{11}^3) +9\mu_{21}^2\mu_{20}\mu_{02}^2 - 18\mu_{21}\mu_{12}\mu_{20}\mu_{11}\mu_{02} + +6\mu_{21}\mu_{03}\mu_{20}(2\mu_{11}^2 - \mu_{20}\mu_{02}) + 9\mu_{12}^2\mu_{20}^2\mu_{02} -6\mu_{12}\mu_{03}\mu_{11}\mu_{20}^2 + \mu_{03}^2\mu_{20}^3$
$I_5 = \mu_{20} + \mu_{02}$
$I_6 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2$
$I_7 = (\mu_{30} - 3\mu_{12})^2 + (\mu_{03} - 3\mu_{21})^2$

$$I_8 = (\mu_{30} + \mu_{12})^2 + (\mu_{03} - \mu_{21})^2 \quad (13)$$

$I_9 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12}) \times \times[(\mu_{30} + \mu_{12})^2 - 3(\mu_{03} + \mu_{21})^2] + +3(\mu_{21} - \mu_{03})(\mu_{03} + \mu_{21}) \times \times[3(\mu_{30} + \mu_{12})^2 - (\mu_{03} + \mu_{21})^2]$
$I_{10} = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{03} + \mu_{21})^2] + +4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{03} + \mu_{21})$
$I_{11} = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12}) \times \times[(\mu_{30} + \mu_{12})^2 - 3(\mu_{03} + \mu_{21})^2] + +(3\mu_{12} - \mu_{30})(\mu_{03} + \mu_{21}) \times \times[3(\mu_{30} + \mu_{12})^2 - (\mu_{03} + \mu_{21})^2]$
$I_{12} = \mu_{40}\mu_{04} - 4\mu_{31}\mu_{13} + 3\mu_{22}^2$
$I_{13} = \mu_{40}\mu_{22}\mu_{04} - 2\mu_{31}\mu_{22}\mu_{13} - \mu_{40}^2\mu_{13} - \mu_{04}\mu_{31}^2 - \mu_{22}^3$

$I_{14} = \frac{I_4}{\mu_{00}I_2}$
$I_{15} = \frac{I_1^2}{\mu_{00}I_3}$
$I_{16} = \frac{I_1 I_3}{I_4}$

The algebraic moment invariants are computed from the first *t* central moments and are given as the eigenvalues of predefined matrices, $M[j,k]$, whose elements are scaled factors of the central moments. In contrast to Hu's geometric moment invariants, the algebraic moment invariants can be constructed up to arbitrary order and are invariant to affine transformations. The algebraic moment transform of (Eqn. 5) can be extended to generalized form by replacing the conventional transform kernel $i^p j^q$ with a more general kernel of $P_p(i)P_q(j)$ - the Legendre polynomial or Zernike polynomial respectively. Since both Legendre and Zernike polynomials are complete sets of orthogonal basis, Legendre and Zernike moments are called orthogonal moments. Orthogonal moments allow to accurately reconstruct the described shape. They make optimal utilization of shape information.

3.2 Zernike moments

Zernike moment of order n and repetition m is defined as (Khotanzad, 1990), (Teh, 1988)

$$Z_{nm} = \frac{n+1}{\pi} \iint_{x^2+y^2 \le 1} V_{nm}^*(\rho,\theta) f(x,y)\,dxdy \quad (14)$$

where:
- $f(x,y)$ is the image intensity at (x,y) in Cartesian coordinates,
- $V_{nm}^*(\rho,\theta)$ is a complex conjugate of $V_{nm}(\rho,\theta) = R_{nm}(\rho)e^{jm\theta}$ in polar coordinates (ρ,θ),
- $n \ge 0$, and $n - |m|$ is even positive integer.

The polar coordinates (ρ,θ) in the image domain is related to the Cartesian coordinates (x,y) as $x = \rho cos(\theta)$ and $y = \rho sin(\theta)$.
$R_{nm}(\rho)$ is a radial defined as (Khotanzad, 1990), as follows:

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} \frac{(-1)^s[(n-s)!]\rho^{n-2s}}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!} \quad (15)$$

The first six orthogonal radial polynomials are:

$$\begin{array}{ll} R_{00}(\rho) = 1 & R_{11}(\rho) = \rho \\ R_{20}(\rho) = 2\rho^2 - 1 & R_{22}(\rho) = \rho^2 \\ R_{31}(\rho) = 3\rho^3 - 2\rho & R_{33}(\rho) = \rho^3 \end{array} \quad (16)$$

Zernike moments, which are proven to have very good image feature representation capabilities, are based on the orthogonal Zernike radial polynomials. Zernike moments are defined as continuous integrals over a domain of normalized coordinates. The implementations of such moment functions therefore involve the following sources of errors: (i) the discrete approximation of the continuous integrals, and (ii) the transformation of the image coordinate system into the domain of the orthogonal polynomials.

Moments of order n with repetition m of a discrete image function $f(k,l)$ with spatial dimension $M \times N$ are given by

$$Z_{nm} = \frac{n+1}{\pi} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} f(k,l) R_{nm}(\rho_{k,l}) e^{-jm\theta_{kl}} \quad (17)$$

where the discrete polar coordinates

$$\rho_{kl} = \sqrt{x_l^2 + y_k^2} \quad \theta_{kl} = arctan(\frac{y_k}{x_l}) \quad (18)$$

are transformed by

Table 2: Hu moments invariants for characters with Figure 3.

Hu moments invariants							
1	6.7E-004	3.7E-010	5.6E-015	2.1E-014	1.3E-028	2.5E-019	-1.8E-028
2	6.8E-004	4.9E-010	2.0E-015	7.8E-015	6.0E-030	1.0E-019	3.0E-029
2	6.9E-004	2.2E-010	2.3E-014	1.4E-014	2.4E-028	1.1E-019	-1.1E-028
4	6.8E-004	7.4E-011	8.9E-016	8.2E-015	4.8E-030	-5.7E-020	2.2E-029
4	7.0E-004	3.3E-010	7.8E-015	1.4E-014	-1.4E-028	-1.4E-019	-1.7E-029
8	6.9E-004	2.0E-010	2.2E-015	4.4E-015	6.7E-030	2.9E-020	-1.2E-029

$$x_l = c + \frac{l(d-c)}{N-1} \quad y_k = d - \frac{k(d-c)}{M-1} \tag{19}$$

for $k = 0, \ldots, M-1$ and $l = 0, \ldots, N-1$. c and d are real numbers take values as shown in Figure 4.

To calculate the Zernike moments of an image $f(x, y)$, the image is first mapped to the unit disk using polar coordinates, where the centre of the image is the origin of the unit disk. Those pixels falling outside the unit disk are not used in the calculation.

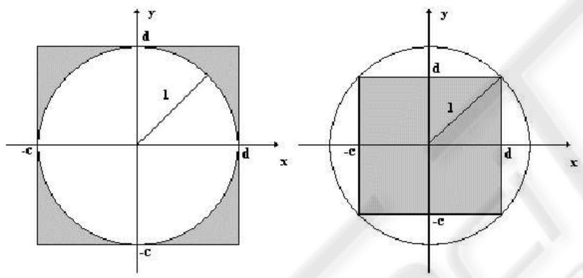

Figure 4: Mapping of a discrete image function a)$c = -1$, $d = 1$ and b)$c = \frac{-1}{\sqrt{2}}$, $d = \frac{1}{\sqrt{2}}$

Normalize the Zernike moments

$$Z_{mn} = \frac{Z'_{nm}}{m_{00}} \tag{20}$$

where, Z_{mn} is the Zernike moments.

Because Z_{mn} is complex, we often use the Zernike moments modules $|Z_{mn}|$ as the features of shape in the recognition of pattern.

The magnitude of Zernike moments has rotational invariant property. An image can be better described by a small set of its Zernike moments than any other types of moments such as geometric moments, Legendre moments, rotational moments, and complex moments in terms of mean-square error. Zernike moments do not have the properties of translation invariance and scaling invariance. The way to achieve such invariance is images translation and image normalization before calculation of Zernike moments.

4 SIMILARITY MEASURE

Recognition is made by associating a feature vector calculated for unknown character with a set of feature vectors for a character obtained with a similar training set. The moment invariant approach to character identification attempts to represent the pattern by a set of K moments invariant-features, thus as a point in K-dimensional feature space. Points corresponding to patterns of the same class are assumed to be close together, not close to those of different classes. The similarity distance (d_m) between two feature vectors M^X and M^Y for a pair of character images X and Y (C(ategory) code image X is identical to C(ategory) code image Y) is computed as the Euclidean distance as follows

$$d_m(M^X, M^Y) = \sqrt{\sum_{k=1}^{K} (m_k^X - m_k^Y)^2} \tag{21}$$

The value of d_m is zero or small for identical or similar characters and high for different characters.

5 CONCLUSION

The main contribution of this paper is presentation of character recognition using set of orthogonal moments. In particular, we have constructed feature vector by applying the normalized various moments. This vector used in conjunction with a simple classification measure such as the Euclidean distance, is capable of achieving satisfactory performance levels.

In our experiment we used our own database which provides handwritten numerals from a hundred writers. Each numeral has 10 samples. Since the size of a sample image varies, we first normalized each image into the size of pixels. If the Hu moments invariants were used, the recognition rate of over 93.2% could be obtained. Five samples of each characters were used as training sets. The recognition rate of using all combinations of moments Eq.13 was found to be significantly better - 98.7%. For real time applications, we prefer only first four moment invariants e.g. with recognition rate 96.8%. When more training sets were used, the recognition rate was found to be higher.

REFERENCES

Hu, M. (1999). Pattern recognition by moment invariants. *Proc. IRE*.vol. 49, pp.1428. .

Flusser, J., Suk, T.(1993) Pattern recognition by affine moment invariants. *Pattern Recognition*, vol. 26, pp. 167-174.

Haralick, R., Shanmugam, K., Dinstein, I. (1973) Textural features for image classification. *IEEE Trans. on Systems, Man, and Cybernetics*. SMC-3(6), pp.610-621, 1973.

Iivarinen, I., Peura, M., Sarela, J., Visa, A. (1997) Comparision of combined shape descriptors for irregular objects. Proc. 8th British Machine Vision Conf.. vol.2, pp.430-439, 1997.

Suen, C. Y., Nadal, C., Mai, T.A., Legault, R., Lam, L. (1992) Computer recognition of unconstrained handwritten numerals. Proc. IEEE, vol. 80 (7), pp. 1162-1189, 1992.

Teh, C.H., Chin. R.T., (1988) On image analysis by the methods of moments. *IEEE Trans. Pattern Anal. Machine Intell.*, 10 (4), 496-513, July 1988.

Khotanzad, A., Hong, Y.H. (1990) Invariant image recognition by Zernike moments. *IEEE Trans. Pattern Anal. Machine Intell.*, 12 (5) , 489-498, May 1990.