International Journal of Computer & Information Technologies (IJOCIT)
www.ijocit.org & www.ijocit.ir
ISSN = 2345-3877

Robust Optical Character Recognition under Geometrical Transformations

Mohammad Sadegh Aliakbarian[1], Fatemeh Sadat Saleh[2], Fahimeh Sadat Saleh[3], Fatemeh Aliakbarian[4]
Isfahan, University of Tech.[1], Sharif University of Tech.[2], Alzahra University[3], Amirkabir University of Tech.[4]
ms.aliakbarian@ec.iut.ac.ir[1], saleh@ce.sharif.edu[2], f.saleh@alzahra.ac.ir[3], f.aliakbarian@ce.aut.ac.ir[4]

Abstract: Optical character recognition (OCR) is a very active field for research and development, and has become one of the most successful applications of automatic pattern recognition. Dealing with scaled, translated and rotated characters are some challenging problems nowadays. On the other hand, another important issue is the dealing with high dimension local features of a character. In this paper, a geometrical transform invariant feature extraction is proposed. After this feature extraction, the dimensionality of extracted features is reduced to a very lower dimension space. Employed supervised dimensionality reduction method not only maximizes the between-class distances and minimizes within-class distances simultaneously, but also makes no loss in class separability power. Experimental results show that the accuracy of classification on extracted features is strongly high for translated, scaled and rotated characters. Another experiment result is that a reduction in feature space dimension to M-1, which M is the number of classes, makes no loss in class separability power.

Keywords: Optical Character Recognition, Geometrical Transform, Feature Extraction, Dimensionality Reduction.

1. Introduction

Optical character recognition (OCR) is a very active field for research and development, and has become one of the most successful applications of automatic pattern recognition.

International Journal of Computer & Information Technologies (IJOCIT)
Corresponding Author: *Mohammad Sadegh Aliakbarian*
February, 2014 Volume 2, Issue 1

Optical character recognition, is the task of conversion of scanned images of handwritten, typewritten or printed text into machine-encoded text. There are many researches around simple OCR [1, 2 and 3]. Simple OCR is widely used as a form of data entry from some sort of original paper data source, whether documents, sales receipts, mail, automatic number plate recognition, importing business card information into a contact list or any number of printed records. It is a common method of digitizing printed texts so that they can be electronically searched, stored more compactly, displayed on-line, and used in machine processes such as machine translation, text-to-speech and text mining. But in dealing with rotated characters, scaled ones or random translated characters, simple optical character recognition methods may fail. So this is a challenging problem to deal with geometrical transformed characters. Converting handwriting in real time to control a computer, defeating CAPTCHA anti-bot systems, converting the textual to machine readable text or converting rotated, scaled and translated characters in an image to machine readable text are some of the examples of geometrical transformed character recognition. There are a few researches about rotated, scaled or translated character recognition [4,5,6]. This paper presents a robust and fast method for this kind of OCR.

In proposed optical character recognition, each character is first separate from other characters in text image. Then, geometrical and photometrical normalization will be applied on each segment. In simple words, a geometrical normalization step, resize, crop and use proper white padding if needed to make all images in a specific size. Photometrical normalization is the task of normalizing the image in term of brightness, gray level of the black and white image and binarization. Then an accurate morphological edge detecting method extracts the edge and boundary of each character. After all these operations, local features (boundary coordinates) are extracted. The next remaining steps which we

© 2014, IJOCIT All Rights Reserved

International Journal of Computer & Information Technologies (IJOCIT)
Corresponding Author: *Mohammad Sadegh Aliakbarian*
February, 2014 Volume 2, Issue 1

focused mostly on them in this paper are extracting geometrical transforms (rotation, scaling and translation) invariant features using Discrete Fourier Transform, robust supervised dimensionality reduction and finally classification. In extracting geometrical transforms invariant features, the problem of sampling origin is also solved by generating sampling origin invariant features. Employed supervised dimensionality reduction method in this paper, not only maximizes the between class distances and minimizes within class distances, but also makes no loss in class separability power. And finally for simplicity, Nearest Neighbor classifier is used to determine the closest training character feature vector for each test character vector.

The remainder of this paper is organized as follows. The proposed algorithm is described in Section 2 precisely. Section 3 presents some experimental results, and finally some conclusions are given in Section 4.

2. Proposed Methodology

As a quick view of our proposed method we can say that: each character is first separate from other characters in text image. Then, geometrical and photometrical normalization will be applied on each segment. Then a morphological edge detecting method extract the boundary of each character. After all these operations local features (boundary coordinates) are extracted. The next steps are extracting geometrical invariant features using Discrete Fourier Transform, dimensionality reduction and finally classification which this article mainly focuses on these three latter steps.

Figure. 1: Geometerical Transformed Characters

The rest of this section talks about our proposed methodology more precisely.

2.1. Character separation

In this step, the goal is to separate each character from others in a text image. This separation is done using basic segmentation methods. Then

© 2014, IJOCIT All Rights Reserved

each separated character goes for normalization step.

2.2. Character Normalization

After character separation, each character is ready to be normalized. In this article, two types of normalization are applied on each character: geometrical normalization and photometrical normalization. In geometrical normalization, each character must have a specific constant size. To this end, proper cropping and resizing is done for each character. White padding is added if needed. In photometrical normalization, the brightness and gray level of each character will setting up. To this end, after converting to black and white, we define a gray level threshold. If the intensity of the pixel is smaller than defined threshold, the pixel value is change to 0 and otherwise the pixel value is change to 1. This is called binarization. Result of this step is illustrated in Figure. 2.

Figure. 2: Separated Normalized Characters

2.3. Morphological Edge Detection

Edge detection is an important task involves in many object recognition problems and also is an important pre-processing step in image segmentation. In this step, mathematical morphology has been employed for edge detection [7]. Mathematical Morphology is a powerful tool for solving various problem in image processing. There are four main operators in morphology: erosion, dilation, opening and closing. With combination of these operators and defining a proper structure element, we can perform an edge detection on character images. In the following we introduce two main morphological operators that are used in edge detection and then we define morphological edge detector.

Let I denote a gray-scale 2-dimentional image and B denote structuring element.

Dilation of an image $I(x,y)$ by a structuring element $B(s,t)$ is defined as bellow,

$$(I \oplus B)(x,y) = \max\{I(x-s,y-t) + B(s,t)\} \qquad (1)$$

© 2014, IJOCIT All Rights Reserved

International Journal of Computer & Information Technologies (IJOCIT)
Corresponding Author: *Mohammad Sadegh Aliakbarian*
February, 2014 Volume 2, Issue 1

Erosion of an image $I(x,y)$ by a structuring element $B(s,t)$ is defined as bellow,

$$(I \ominus B)(x,y) = \min\{I(x+s,y+t) - B(s,t)\} \qquad (2)$$

We denote edge of image I by $E_d(I)$ that is defined as the difference set of the dilation domain of I and the domain of I. This is also named as dilation residue edge detector.

$$E_d(I) = (I \oplus B) - I. \qquad (3)$$

The edge of image I also can be defined as the difference set of the domain of I and the erosion domain of I. This is also named as erosion residue edge detector:

$$E_e(I) = I - (I \ominus B). \qquad (4)$$

The size and shape of the structuring element is an important factor in the final result of detected edges. Figure.4 shows the effects of different Filter sizes. In this paper we use a 3×3 crisscross structure element that shown in Figure. 3 because of its complete edge detection and no extra information.

0	1	0
1	1	1
0	1	0

Figure. 3: 3×3 crisscross structure element

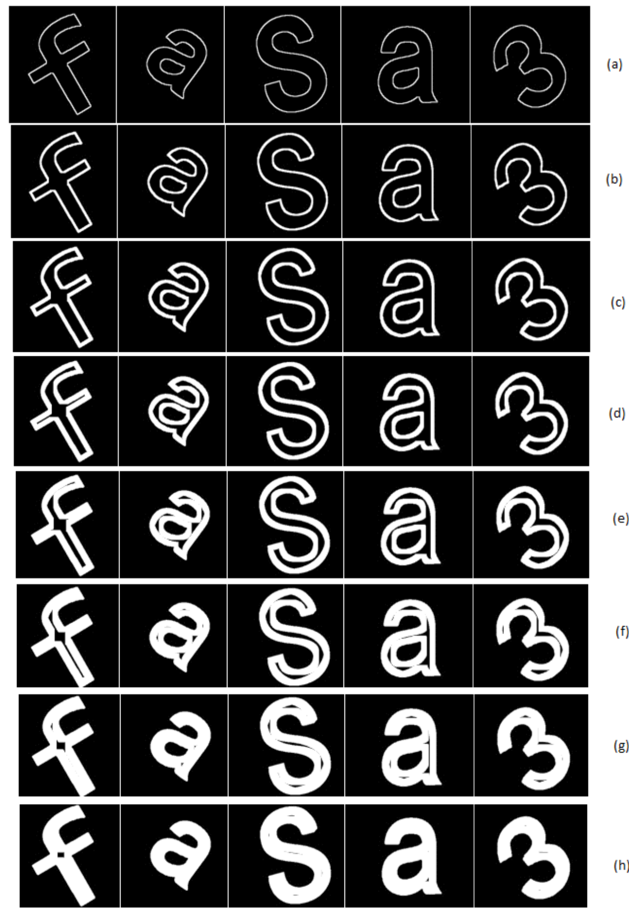

Figure. 4: Mathematical Morphological Edge Detection with Different Filter Sizes: From Top to Bottom 3, 5, 7, 11, 13, 15 and 17

© 2014, IJOCIT All Rights Reserved

2.4. Local Feature Extraction

After passing previous steps, there is no texture on characters any more. All the property of characters are boundary points. All the things in this step is to extract local features. Local features are coordinates of boundary points in pair (x, y).

2.5. Geometrical Transform Invariant Feature Generation

The features which are used in this paper, is totally robust and invariant to geometrical transformations such as translation, rotation and scaling. These features are also invariant to the sampling point origin. The rest of this section talks about these robust features.

2.6. Fourier Coefficients as Robust Features

Consider the pair (x, y) be one of the N feature pairs from boundary of a normalized separated character. For each boundary coordinate pair (x_k, y_k), an imaginary number can be defined as shown in Equation (6).

$$u_k = x_k + jy_k \tag{6}$$

Where the DFT (Discrete Fourier Transform) can be calculated for each boundary descriptors u_k as shown in Equation (7).

$$f_l = \sum_{k=0}^{N-1} u_k \, exp\left(-j \, {}^{2\pi}/_{N} \, lk\right), l = 0,1,2, \dots, N-1 \tag{7}$$

In this paper, Fourier coefficients [8] f_l are used with a modification as geometrical invariant features. In order to reconstruct the boundary of each character, u_k is computed by calculated f_l. However, in this article, the goal is not character reconstruction, so there is no need to have all Fourier coefficients. The rest of this subsection talks about geometrical transform invariant feature extraction and generation. These geometrical transforms are translation, rotation and finally scaling.

In translation of a character form one position to another one in a screen, the coordinate of each point will be translated. In other words, a specified real number will be added to the pair (x, y). New coordinate can be defined as shown in Equation (8).

$$(x'_k, y'_k) = (x_k + \Delta x, y_k + \Delta y) \tag{8}$$

© 2014, IJOCIT All Rights Reserved

So, the corresponding imaginary number will be constructed using Equation (9).

$$u'_k = u_k + (\Delta x + j\Delta y) \equiv u_k + \Delta u' \qquad (9)$$

Therefore, the Fourier coefficients of translated point will be modified as shown in Equation (10).

$$f'_l = \sum_{k=0}^{N-1}(u_k + \Delta u')\, exp\left(-j\, {}^{2\pi}/_{N}\, lk\right) = \sum_{k=0}^{N-1}(u_k)\, exp\left(-j\, {}^{2\pi}/_{N}\, lk\right) + \sum_{k=0}^{N-1}(\Delta u')\, exp\left(-j\, {}^{2\pi}/_{N}\, lk\right) = f_l + \Delta u' \delta(l) \qquad (10)$$

If we look more detailed at above equation, we can find that Fourier coefficients of translated character is different from original Fourier coefficients in just zero point (zero's Fourier coefficient). In other words, translation transform affects only f'_0 coefficient. So, by discarding the 0^{th} Fourier coefficient of translated features, this feature will be invariant to translation transform. It means

$$f_l = f'_l \,,\, l \neq 0 \qquad (11)$$

In rotation of a character, the event is just a change in angle. All the point of character boundary will be rotated by a specified angle θ. The corresponding imaginary number of rotated point can be constructed as shown in Equation (12).

$$u'_k = u_k\, exp(j\theta) \qquad (12)$$

As it is obvious from Equation (12), rotation transform affects only in imaginary part by a constant change in angle. Now, Equation (13) can be derived by simplification of modified Fourier coefficients.

$$f'_l = \sum_{k=0}^{N-1}(u_k) exp(j\theta)\, exp\left(-j\, {}^{2\pi}/_{N}\, lk\right) = \sum_{k=0}^{N-1}(u_k)\, exp\left(-j\, {}^{2\pi}/_{N}\, lk + j\theta\right) \qquad (13)$$

As it is shown in Equation (14), rotation transform can change just the phase of Fourier coefficients and the amplitudes of rotated Fourier coefficient are completely equal to original ones. So, by extracting the real part of Fourier coefficient and discarding the imaginary part, a rotation invariant feature will be generated. In other words,

© 2014, IJOCIT All Rights Reserved

$$|f_l| = |f_l'| \quad (14)$$

Finally, in scaling of a character, each point of character boundary is multiplied to a constant real number. If scaling goal is larger character, the constant real number is greater than 1 and otherwise it is smaller than 1. Multiplication of a constant number to the coordinate of the boundary just can cause a simple change in corresponding imaginary number which is shown in Equation (15).

$$u_k' = a u_k \quad (15)$$

By simplification of Fourier coefficients after scaling transform, it can be concluded that it is possible to generate scaling invariant features. The scaled character Fourier coefficients is shown below.

$$f_l' = \sum_{k=0}^{N-1} a(u_k)\, exp\left(-j\, {}^{2\pi}/_{N}\, lk\right) = a \sum_{k=0}^{N-1} (u_k)\, exp\left(-j\, {}^{2\pi}/_{N}\, lk\right) \quad (16)$$

As we can find from Equation (16), the only modification is the multiplication of scaling constant number to the Fourier coefficients. If we let the new features to be the ratio of two subsequent Fourier coefficients, this new feature will be invariant to scaling transform. In other words,

$$f_j'/_{f_i'} = f_j/_{f_i} \quad (17)$$

2.7. Dealing with Sampling Point Origin:

The first decision to be taken, prior to the computation of the Fourier coefficients, is to define the first sampling point (x_0, y_0) on the boundary. In practice, the choice of this point for each character has a degree of randomness. The choice of a different sampling origin corresponds to a relative translation of, say, $k_0 < N$ samples (since the boundary is a closed curve, the relative translation will always be $(k - k_0)$ modulo $N < N$). As it is bright, if we start sampling boundary of character from multiple different points, different feature vectors (with local features and boundary coordinate pairs) will be generated for one specific character and image. So, a significant reduction in classification accuracy will be observed. In order to overcome this problem, we have to generate feature vectors which are also invariant to sampling origin. In

© 2014, IJOCIT All Rights Reserved

International Journal of Computer & Information Technologies (IJOCIT)
Corresponding Author: *Mohammad Sadegh Aliakbarian*
February, 2014 Volume 2, Issue 1

other words, we have to generate features which produce one specific feature vector for one specific character while sampling is started in any point. The sampling origin point can be defined as,

$$u'_k = u_{k-k0} \tag{18}$$

In order to have sampling origin invariant features for optical character recognition, we can write

$$f'_l = \sum_{k=0}^{N-1} (u_{k-k0}) \exp\left(-j \,{}^{2\pi}/_{N}\, lk\right) \rightarrow f'_l = f_l \exp\left(-j \,{}^{2\pi}/_{N}\, lk0\right) \tag{19}$$

As we can see, a changing in sampling starting point just affects phase or imaginary part of Fourier coefficients and the amplitude or real part is remained without any changes. So, by extracting the real part of Fourier coefficient and discarding the imaginary part, sampling origin invariant features will be generated. So,

$$|f_l| = |f'_l| \tag{20}$$

2.8. Dimensionality Reduction

After generating geometrical transform invariant features, it is time to reduce the dimensionality of these features. Boundary of a normalized character has many local features (coordinate pairs (x, y)). After generating those robust features using Fourier coefficient in latter subsection, each coordinate pair will be converted to one number. It means the data gets smaller. But, these new features are still high dimension and can reduce the speed of recognition execution. Therefore, it is needed to reduce the dimensionality of feature space. Proposed reduction method not only reduce dimensionality to the minimum dimension space, but also this reduction makes no loss in class separability power. In short, beside dimensionality reduction, this algorithm tries to maximize between class distances and minimize within class distances simultaneously. Rest of this subsection talks about this algorithm in detail.

2.8.1. Scatter Matrices

First of all, we define two matrices which play important roles in dimensionality reduction process. These matrices specify the way feature vector samples are scattered in the *l*-dimensional

© 2014, IJOCIT All Rights Reserved

space. The first matrix, *within-class* scatter matrix is defined in Equation (21).

$$S_w = \sum_{i=1}^{M} P(\omega_i) E\{(x-\mu_i)(x-\mu_i)^T\} \tag{21}$$

And second matrix, *between-class* scatter matrix is defined in Equation (22).

$$S_b = \sum_{i=1}^{M} P(\omega_i)(\mu_i - \sum_{i=1}^{M} P(\omega_i)\mu_i)(\mu_i - \sum_{i=1}^{M} P(\omega_i)\mu_i)^T \tag{22}$$

Where ω_i is the i^{th} class, x is the sample, μ_i is the mean of i^{th} class and finally $P(\omega_i)$ is the probability of i^{th} class which can be calculated using i^{th} class samples and all samples. According to definition of scatter matrices, an optimization criteria can be defined. Maximizing this optimization criteria is the goal of dimensionality reduction. This criteria is defined in Equation (23).

$$Criteria = Trace\{S_w^{-1} S_b\} \tag{23}$$

According to Equation (23), the smaller S_w, the closer samples and consequently smaller variance. On the other hand, the greater S_b, the greater between-class distances.

2.8.2. Minimum Dimension Feature Extraction

In order to project current dimension space to a lower dimension space which satisfy our optimization criteria, we have to find a transform matrix to maximize the $Criteria$. Let y be the feature vector in reduced dimension space. y can be defined as shown in Equation (24).

$$y = A^T x \tag{24}$$

where A is transform matrix. The size of y is *l*l*. The goal is to find a transform matrix A which maximizes the optimization ciretia as follows.

$$Criteria(A) = Trace\{S_w^{-1} S_b\} = Trace\{(A^T S_{xw} A)^{-1}(A^T S_{xb} A)\} = Tr \tag{25}$$

In order to maximize Equation (25), we can take the derivative of this function and make the result equal to zero. By solving and simplifying the result of derivation, Equation (26) can be derived which is very similar to the relation of eigenvectors and eigenvalues of a matrix.

$$(S_{xw}^{-1} S_{xb})A = A(S_{yw}^{-1} S_{yb}) \tag{26}$$

It was proved [9] that there is a linear combination, S_{yw} and S_{yb} convert to diagonal matrices. This diagonalization is done using

© 2014, IJOCIT All Rights Reserved

matrix B as shown in Equation (27) and Equation (28).

$$I = B^T S_{yw} B \tag{27}$$

$$D = B^T S_{yb} B \tag{28}$$

From another perspective, the goal is to find $\hat{y}$ which have the minimum error according to y. We define $\hat{y}$ as an estimation of y.

$$\hat{y} = B^T y = B^T A^T x \tag{29}$$

It was also proved [9] that in projection of y to $\hat{y}$ there is no change in the value of optimization criteria.

$$a(\hat{y}) = Trace\{S_{\hat{y}w}^{-1} S_{\hat{y}b}\} = Trace\left\{\left(B^T S_{yw} B\right)^T \left(B^T S_{yb} B\right)\right\} = Trace\{B^{-1} S_{yw}^{-1} S_{yb} B\} = Trace\{S_{yw}^{-1} S_{yb} B B^{-1}\} = Criteria(y) \tag{30}$$

Combination of Equation (27), Equation (28) and Equation (29) derive a new equation which is very similar to the relation of eigenvalues and eigenvectors. In Equation (31) the size of $S_{xw}^{-1} S_{xb}$ is $m*m$ and $C=AB$ with the size of $m*l$.

$$(S_{xw}^{-1} S_{xb}) C = CD \tag{31}$$

Matrix *C* is a diagonal matrix containing eigenvalues of $S_{xw}^{-1} S_{xb}$ on its main diagonal values. Matrix *D* is containing the eigenvectors which are corresponding to those eigenvalues in *C*. Now the goal is to select *l* eigenvectors form all *m* eigenvectors which can cause the best results in term of criteria optimization. It can be proved that selecting *l* eigenvectors with highest eigenvalues helps to reach our goal [9]. Another property of this dimensionality algorithm is its minimum MSE (Mean Square Error). Feature space can be reconstruct using Equation (32).

$$y = A^T x \rightarrow Ay = x \tag{32}$$

In addition to maximization of between class distances and minimization of within class distances, this reduction minimize the projection error [9].

$$MSE = E\left\{\|x - \hat{x}\|^2\right\} = E\left\{\left\|\sum_{i=0}^{m-1} y(i)\, a_i - \sum_{i=0}^{M-2} y(i)\, a_i\right\|^2\right\} = E\left\{\left\|\sum_{i=M-1}^{m-1} y(i)\, a_i\right\|^2\right\} = \sum_{i=M-1}^{m-1} \lambda_i \tag{33}$$

Where x is original feature space and $\hat{x}$ is the estimation of reduced feature space. $y(i)$ is the value of projected sample on a_i . λ_i is the i th eigenvalue. Equation (33) shows that discarding eigenvectors which have lower eigenvalues minimize the projection error and result higher accuracy. According to the concept of *matrix rank* [9] the best dimension for new space (the value of *l*) which makes no loss in class

© 2014, IJOCIT All Rights Reserved

separability power is equal to the number of class labels minus one [9]. In our character recognition, there are 28 classes for letters and 10 classes for digits, so the samples are in new 38 dimension feature space.

2.9. Classification

In order to determine the closest training character feature vector for each test vector we perform nearest neighbor classification using Euclidean and Mahalanobis distance metrics on reduced dimension space. Vector $(a_1, \ldots, a_d)$ is the output of proposed optical character recognition estimation algorithm. Given a set of training characters, we obtain a corresponding set of vectors $(a_1^{id}, \ldots, a_d^{id})$, where id is an index over the set of training characters. Similarly, given a test character, we obtain a vector $(a_1, \ldots, a_d)$ of reduced dimension vector of that character. To complete the optical character recognition algorithm, we need an algorithm that classifies $(a_1, \ldots, a_d)$ with the index id, which is the most likely match. For simplicity, we use the nearest-neighbor algorithm that classifies the vector $(a_1, \ldots, a_d)$ with the index.

$$\mathrm{argmin}_{id}\ dist\left((a_1, \ldots, a_d), \left(a_1^{id}, \ldots, a_d^{id}\right)\right) = \mathrm{argmin}_{id} \sum_{i=1}^{d} (a_i - a_i^{id})^2 \qquad (34)$$

All of the results mentioned in this article use the Euclidean distance. Alternative distance functions, such as the Mahalanobis distance, could be used instead if so desired.

3. Experimental results

3.1. Experiment conditions

In this paper a dataset with 5000 characters with 20 different fonts is generated. Generated dataset contains 1000 normal characters, 1000 rotated characters, 1000 scaled characters, 1000 translated characters and finally 1000 characters with combination of all geometrical transforms. All transformed characters are generated randomly using bootstrap functions with random angels for rotation, random scale parameter for scaling and random translation distance for translation transform. Then training and testing sets are generated. In each group, 800 random

© 2014, IJOCIT All Rights Reserved

International Journal of Computer & Information Technologies (IJOCIT)
Corresponding Author: *Mohammad Sadegh Aliakbarian*
February, 2014 Volume 2, Issue 1

characters are selected as training set and 200 remaining characters construct test set.

3.2. Experiment results

All input character images are normalized to size *32*32* after binarization and morphological edge detection. Then coordinates of boundary pixels are extracted as character local features. Then these local features are employed for next step, geometrical transform invariant feature extraction. The remaining of this section shows and compares different experimental conditions results.

TABLE 1 shows the effect of dimensionality reduction on Fourier coefficient features. This experiment is done for 6 different feature vector sizes and three different states for K-Nearest Neighbor classifier with K=1, K=2 and K=3 with the average accuracy of normal character, rotated character, scaled characters, translated characters and finally combination of these condition classification. As we can see, this experiment also confirms the fact that reducing the dimensionality to M-1 which M is the number of classes makes no loss in class separability power. In this paper the number of classes is 38. TABLE I shows that there is no significant loss in classification accuracy for feature vector size 100, 80, 60, 40 and 37. But when the dimension is less than 37, the accuracy of the classification decreases significantly.

TABLE 2, shows a complete comparison of many different experiment conditions. Four character conditions, three different feature sets and three different K-Nearest Neighbors with K=1, K=2 and K=3 for each pair of character and feature set condition.

Table 1: Result of Dimensionality Reduction, Best Feature Vector Size is 37, One Less Than Class Numbers

Dimension Size	K-Nearest Neighbors Classification Accuracy		
	K=1	K=2	K=3
100	97.68	97.71	97.65
80	98.73	98.22	97.91
60	97.66	98.30	98.12
40	98.66	98.60	98.74
37	***98.71***	***98.73***	***98.93***
30	94.21	95.33	95.09

© 2014, IJOCIT All Rights Reserved

Table 2: Complete Comparison between Proposed Methods and Pixel Value Features

Condition	K-Nearest Neighbor Classification Accuracy								
	Pixel Value Features			Geometrical Transform Invariant Features			Geometrical Transform Invariant Features with Reduced Dimension		
	K=1	K=2	K=3	K=1	K=2	K=3	K=1	K=2	K=3
Normal Characters	94.87	97.45	96.98	97.78	98.88	98.88	98.30	98.33	97.77
Translated Characters	71.00	74.50	69.22	97.34	95.66	98.90	97.81	95.77	98.31
Rotated Characters	29.50	24.66	31.14	9674	96.53	97.33	95.97	94.22	97.99
Scaled Characters	40.00	43.63	43.78	96.98	98.88	98.53	98.12	97.90	99.25
Combination of Transformations	28.72	30.21	28.00	96.63	98.52	98.25	98.71	98.73	98.93

This experiment strongly confirms the vantage and superiority of proposed optical character recognition under rotation, scaling and translation transforms.

4. Conclusion

The most challenging problem in the scope of optical character recognition is dealing with scaled, translated and rotated characters. On the other hand, another important issue is the dealing with high dimension local features of a character. In this paper, after passing segmentation, normalization and morphological edge detection steps, a geometrical transforms invariant feature extraction is presented and applied on local features of characters boundary. Our robust features are not only invariant to scaling, translation and rotation transforms of a character, but also is invariant to the sampling origin point. In other words, using these features, sampling from different boundary points generate just one single feature vector for that character. After these step, the dimensionality of generated feature space is reduced to a lower dimension space. Employed supervised dimensionality reduction method not only maximizes the between class distances (the distance between

© 2014, IJOCIT All Rights Reserved

different characters samples) and minimizes within class distances (the distance between the feature vectors of a specific character), but also makes no loss in class separability power. The mathematical equations and facts in proposed methodology section prove the strength of generated features under geometrical transforms and reduced dimension features in term of classification accuracy. Experimental results also show that classification accuracy for translated, scaled and rotated characters are still high and also show that reduction in feature space dimension to M-1, which M is the number of classes, makes no loss in class separability power.

References

[1] P. Hidayatullah, N. Syakrani, I. Suhartini and W. Muhlis, W. "Optical Character Recognition Improvement for License Plate Recognition in Indonesia," European Symposium on Computer Modeling and Simulation (EMS), pp. 249-254 2012

[2] X. Jianhong, "Optical Character Recognition Based on Least Square Support Vector Machine," Third International Symposium on Intelligent Information Technology Application, 2009. IITA, pp. 626-629, 2009

[3] J. Leimer, "Design factors in the development of an optical character recognition machine," IRE Transactions on Information Theory, Vol. 8, Issue 2, pp. 167-171, February 1962

[4] T. Lijing, Z. Yan, Z. Huiqun, "A Warped Document Image Mosaicing Method Based on Registration and TRS Transform," 10th International Conference on Computer and Information Science (ICIS), 2011 IEEE/ACIS, 16-18, pp. 179 – 183, May 2011

[5] C. Singh, R. Upneja, "Accuracy and numerical stability of high-order polar harmonic transforms," Image Processing, IET, Vol. 6, Issue: 6, pp. 617 - 626, August 2012

[6] M. Miciak, "Character recognition using Radon Transformation and Principal Component Analysis in postal applications," International Multiconference on Computer Science and Information Technology, pp. 495-500, 2008. IMCSIT 2008

[7] Z. Yuqian, G. Weihua and C. Zhencheng, "Edge detection based on multi-structure elements morphology," Proceedings of the sixth world congress on intelligent control and automation, Dalian, China, June 21-23, 2006

[8] E. Persoon and F. King-Sun, "Shape Discrimination Using Fourier Descriptors," IEEE Transactions on Systems, Man and Cybernetics, Vol. 7, Issue 3, pp. 170-179, March 1977

[9] S. Aliakbarian, A. Fanian, F. Saleh and T. A. Gulliver, "Optimal Supervised Feature Extraction in Internet Traffic Classification," presented at the 21th Iranian Conference on Electrical Engineering (ICEE'13), Mashhad, Iran, 2013

[10] K. Fukunaga, Narendra and M. Patrenahalli, "A Branch and Bound Algorithm for Computing k-Nearest Neighbors," IEEE Transactions on

© 2014, IJOCIT All Rights Reserved

International Journal of Computer & Information Technologies (IJOCIT)
Corresponding Author: *Mohammad Sadegh Aliakbarian*
February, 2014 Volume 2, Issue 1

Computers, Vol.C-24, Issue 7, pp. 750-753, July 1975

Authors Profile:

Mohammad Sadegh Aliakbarian is currently a Researcher and Developer at Electronic Research Institute (ERI) working on Computer Vision, Image and Video Analysis at Sharif University of Technology, Tehran, Iran. Before that, he received his B.Sc. in computer engineering at Isfahan University of Technology, Isfahan, Iran. His research interests are statistical pattern recognition, machine learning, image and video processing.

Fatemeh Sadat Saleh is currently a Researcher and Developer at Electronic Research Institute (ERI) working on Computer Vision, Image and Video Analysis at Sharif University of Technology, Tehran, Iran. Before that, she received her M.Sc. in artificial intelligence at Sharif University of Technology, Tehran, Iran and B.Sc. in computer engineering at Isfahan University of Technology, Isfahan, Iran. Her research interests are statistical pattern recognition, machine learning, speech processing, image and video processing.

Fahimeh Sadat Saleh is currently a M.Sc. student in artificial intelligence at Alzahra University, Tehran, Iran. Previous to that, she received her B.Sc. in computer engineering at University of Isfahan, Isfahan, Iran. Her area of researches are medical image processing and machine learning.

Fatemeh Aliakbarian received her M.Sc. in

artificial intelligence at Amirkabir University of Technology, Tehran, Iran. Previous to that, she received her B.Sc. in computer engineering at University of Tehran, Tehran, Iran. Her areas of researches are data mining and machine learning.

© 2014, IJOCIT All Rights Reserved