International Journal of Computer Science and Applications,

Vol. 7 No. 3, pp. 33 - 44, 2010

RADON TRANSFORMATION AND PRINCIPAL COMPONENT ANALYSIS METHOD APPLIED IN POSTAL ADDRESS RECOGNITION TASK

MIROSŁAW MICIAK

Institute of Telecommunications
University of Technology and Life Sciences in Bydgoszcz
Kaliskiego 7, 85-791 Bydgoszcz, Poland
miciak@utp.edu.pl

In this paper a new method of handwritten characters recognition is introduced. The proposed algorithm is applied to classification of post mails on the basis of zip code information. In connection with this work the research was conducted with numeric characters used in real post code of mail pieces. Moreover article contains basic image processing for instance filtration binarization and normalization of the character. The main objective of this article is to use the Radon Transform and Principal Component Analysis methods to obtain a set of invariant features, on basis of which postal code will be recognized. The reported experiments results prove the effectiveness of the proposed method. Furthermore sources of errors as well as possible improvement of classification results will be discussed.

Keywords: Character recognition; Radon Transform; Principal Component Analysis.

1. Introduction

The today's systems of automatic sorting of the post mails use the OCR (Optical Character Recognition) mechanisms. In the present recognizing of addresses (particularly written by hand) the OCR is insufficient.

The typical system of sorting consists of the image acquisition unit, video coding unit and OCR unit. The image acquisition unit sends the mail piece image to the OCR for interpretation. If the OCR unit is able to provide the sort of information required (this technology has 50 % effectiveness for all mails [Forella (2000)], it sends this data to the sorting system, otherwise the image of the mail pieces is sent to the video coding unit,

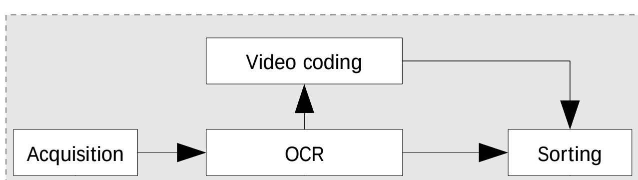

Fig. 1. The automatic sorting system - mail flow.

where the operator writes down the information about mail pieces.

The main problem is that operators of the video coding unit have lower throughput than an OCR and induce higher costs [Forella (2000)]. Therefore the OCR module is improving, particularly in the field of recognition of the characters. Although, these satisfactory results were received for printed writing, the handwriting is still difficult to recognize. Taking into consideration the fact, that manually described mail pieces make 30 percent of the whole mainstream, it is important to improve the possibility of segment recognizing the hand writing. This paper presents the proposal of a system for recognition of handwritten characters, for reading post code from mail pieces.

2. Proposed system overview

The process of character recognition process can be divided into stages: filtration and binaryzation, normalization, Radon Transform calculating, accumulator analysis, Principal Component Analysis, feature vector building, and character recognition stage.

The first step of the image processing is binarization. The colourful image represented by 3 coefficients Red, Green and Blue from the acquisition unit must be converted to the image with 256 levels of grey scale. The next step of processing of the image of mail piece is digital filtration. The filtration is used for improving the quality of the image, emphasizing details and making processing of the image easier. The filtration of digital images is obtained by convolution operation. The new value of point of image is counted on the basis of neighbouring points value. Every value is classified and it has influence on new value of point of the image after filtration.

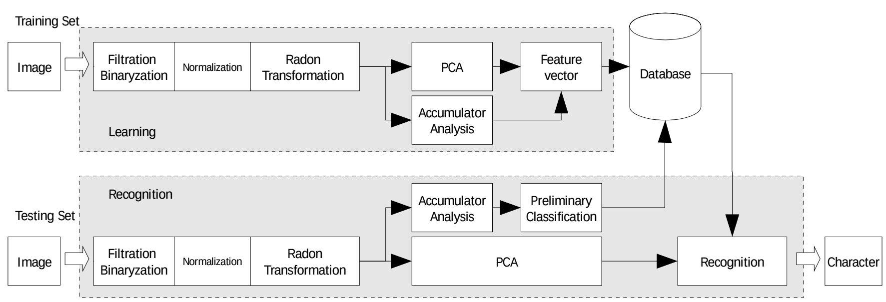

Fig. 2. The proposed method of character recognition.

In the pre-processing part non-linear filtration was applied. The statistical filter separates the signal from the noise, but it does not destroy useful information. The applied filter is median filter, with mask 3x3.

The image of character received from the acquisition stage have different distortion such as: translation, rotation and scaling. The character normalization is applied for standardization size of the character. Images there are translated, rotated and expanded or decreased.

The typical solutions takes into consideration the normalization coefficients and calculate the new coordinates given by:

$$[x,y,1]=[i,j,1]\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -I & -J & 1\end{bmatrix}\times\begin{bmatrix}m_i & 0 & 0\\ 0 & m_j & 0\\ 0 & 0 & 1\end{bmatrix}\times\begin{bmatrix}\cos\beta & \sin\beta & 0\\ \sin\beta & \cos\beta & 0\\ 0 & 0 & 1\end{bmatrix} \quad (1)$$

where: *I,J* is a center of gravity given by :

$$I=\frac{\sum_i\sum_j if(i,j)}{\sum_i\sum_j f(i,j)} \qquad J=\frac{\sum_i\sum_j jf(i,j)}{\sum_i\sum_j f(i,j)} \quad (2)$$

In the reality we haven't got this parameters starting right now, so we use new coordinate system where center is equals to center of gravity of the character. The value of angle rotation is according to main axes of the image. The value of scale coefficient is calculated by mean value of variation of the character. So the center of gravity of the character is good candidate point of the center of image as a product of normalization stage.

3. Radon Transformation

In recent years the Radon transform have received much attention. This transform is able to transform two dimensional images with lines into a domain of possible line parameters, where each line in the image will give a peak positioned at the corresponding line parameters. This have lead to many line detection applications within image processing, computer vision, and seismic [Peter (1996)]. The Radon Transformation is a fundamental tool which is used in various applications such as radar imaging, geophysical imaging, nondestructive testing and medical imaging [Venturas (2005)]. The Radon transform computes projections of an image matrix along specified directions. A projection of a two-dimensional function *f(x,y)* is a set of line integrals. The Radon function computes the line integrals from multiple sources along parallel paths, or beams, in a certain direction. The beams are spaced 1 pixel unit apart. To represent an image, the

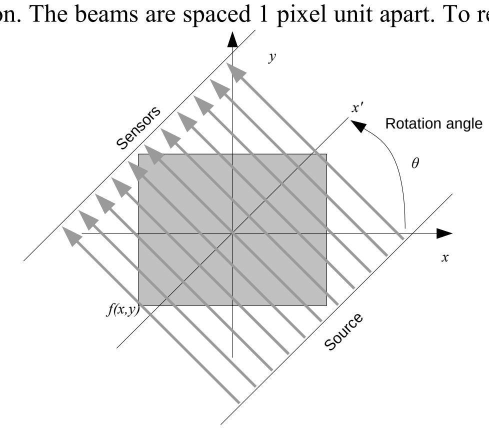

Fig. 3. Single projection at a specified rotation angle.

radon function takes multiple, parallel-beam projections of the image from different angles by rotating the source around the centre of the image.
The "Fig.3" shows a single projection at a specified rotation angle. The Radon transform is the projection of the image intensity along a radial line oriented at a specific angle. The

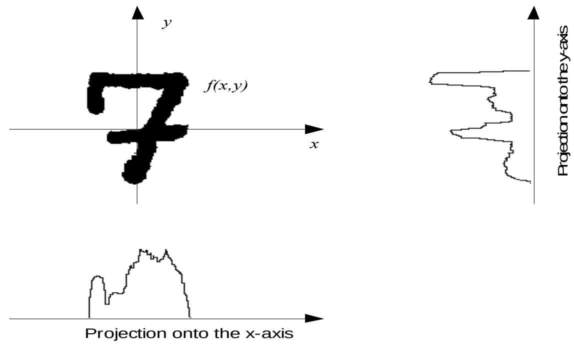

Fig. 4. Horizontal and Vertical Projections of a Simple Function.

radial coordinates are the values along the *x′*-axis, which is oriented at θ degrees counter clockwise from the *x*-axis. The origin of both axes is the center pixel of the image. For example, the line integral of *f(x,y)* in the vertical direction is the projection of *f(x,y)* onto the *x*-axis; the line integral in the horizontal direction is the projection of *f(x,y)* onto the *y*-axis. The "Fig.4" shows horizontal and vertical projections for a simple two-dimensional

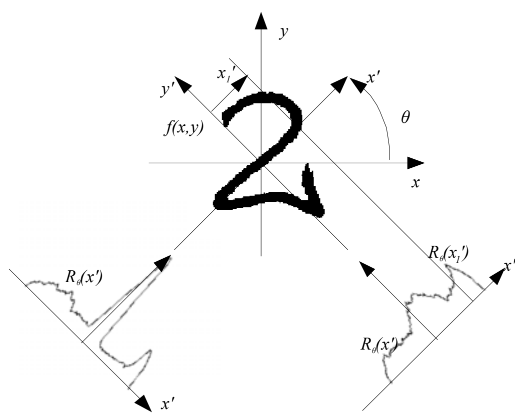

Fig. 5. Geometry of the Radon Transform.

function. Projections can be computed along any angle θ, by use general equation of the Radon transformation [Asano (2002)][Averbuch (2001)][Kupce (2004)]:

$$R_{\Theta}(x') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\delta(x\cos\Theta + y\sin\Theta - x')dxdy \tag{3}$$

where δ(·) is the delta function with value not equal zero only for argument equal 0, and:

$$x' = x\cos\Theta + y\sin\Theta \tag{4}$$

x' is the perpendicular distance of the beam from the origin and θ is the angle of incidence of the beams. The "Fig.5" illustrates the geometry of the Radon Transformation. The very strong property of the Radon transform is the ability to extract lines (curves in general) from very noise images. Radon transform has some interesting properties relating to the application of affine transformations. We can compute the Radon transform of any translated, rotated or scaled image, knowing the Radon transform of the original image and the parameters of the affine transformation applied to it.

This is a very interesting property for symbol representation because it permits to distinguish between transformed objects, but we can also know if two objects are related by an affine transformation by analyzing their Radon transforms [Ramos (2003)]. It is also possible to generalize the Radon transform in order to detect parameterized curves with non-linear behavior [Peter (1996)][Bracewell (1995)][Lim (1990)].

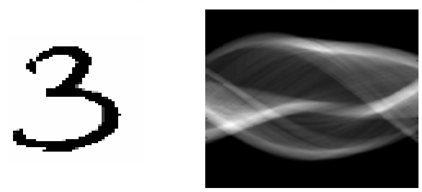

Fig. 6. Sample of accumulator data of Radon Transformation.

4. Principal Components Analysis

PCA is mathematically defined [Jolliffe (2002)][Shlens (2005)][Smith (2002)] as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. PCA is theoretically the optimum transform for a given data in least square terms. The main idea of using PCA for character recognition is to express the large 1-D vector of pixels constructed from 2-D character image into the compact principal components of the feature space [Kim (2003)]. Principal Component Analysis can be used for dimensionality reduction in a data set by retaining those characteristics of the data set that contribute most to its variance, by keeping lower-order principal components and ignoring higher-order ones. Such low-order components often contain the "most important" aspects of the data. For all *r* digital images *f(x,y)* from the normalization stage is creating column vector X_k by the concatenate operation, where *k=(1,....,r)*. For that prepared images we can calculate mean of brightness intensity *M*, difference vector *R* and covariance matrix *Σ*.

$$M_k = \frac{1}{r}\sum_{k=1}^{r} X_k \tag{5}$$

$$R_k = X_k - M_k \tag{6}$$

$$\sum = \frac{1}{r}\sum_{k=1}^{r} R_k R_k^{\ t} \tag{7}$$

Where:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ . \\ X_r \end{bmatrix} \tag{8}$$

$$M = \begin{bmatrix} M_1 \\ M_2 \\ . \\ M_r \end{bmatrix} \tag{9}$$

$$R = \begin{bmatrix} R_1 \\ R_2 \\ . \\ R_r \end{bmatrix} \tag{10}$$

Principal components are calculating from the eigenvectors Φ_l and eigenvalues λ_l of the covariance matrix Σ. The eigenvectors Φ_l are normalized, sorted in order eigenvalue, highest to lowest and transposed, to obtain transformation matrix W, where K is the number of dimensions in the dimensionally reduced subspace calculated by:

$$\frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{l} \lambda_i} \geq p \tag{11}$$

where: p is assumed as threshold [Kim (2003)].The matrix W is given by:

$$W = \begin{bmatrix} \Phi_1^1 & \dots & \Phi_1^K \\ . & \dots & . \\ \Phi_l^1 & \dots & \Phi_l^K \end{bmatrix} \tag{12}$$

After image projection into eigenvectors space we do not use all eigenvectors, but these with maximum eigenvalues, this gives the components in order of significance. The eigenvector associated with the largest eigenvalue is one that reflects the greatest variance in the input data. That is, the smallest eigenvalue is associated with the eigenvector that finds the least variance. They decrease in exponential fashion, meaning that the roughly 90% of the total variance is contained in the first 5% to 10% of the dimensions [Kim (2003)].

The projection of *X* into eigenvectors space is given by:

$$Y = W\left(X - M\right) \tag{13}$$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ Y_r \end{bmatrix} \tag{14}$$

The final data set will have less dimensions than the original [Smith (2002)], after all we have *r* column-vector for each input image with *K* values :

$$Y_k = \begin{bmatrix} y_1 \\ y_2 \\ . \\ y_K \end{bmatrix} \tag{15}$$

The Principal Component Analysis module in proposal system generate a set of data, which can be used as a features in building feature vector section. For instance when we use input matrix 8x8 from Radon Transformation stage, as a result we obtained *K=8* values vector, using Cattell's criterion [Ledesma (2007)].

5. Feature vector calculation stage

Two sets of data received from the Principal Component Analysis module and Accumulator Analysis stage are used to create vector of features of character. To obtain one of the first vector's parameter we used the local peaks of the accumulator. Therefore the initial accumulator data reduction is fulfilled by matching of local maximum operation. As a result of this stage is a amount of local peaks limited as to the value of the threshold, that minimizes the intra-class variance and defined as a weighted sum of variances of the two classes, calculated basis on histogram for all accumulator cells. A second restraint is the distance between each of peak, which is determined on basis

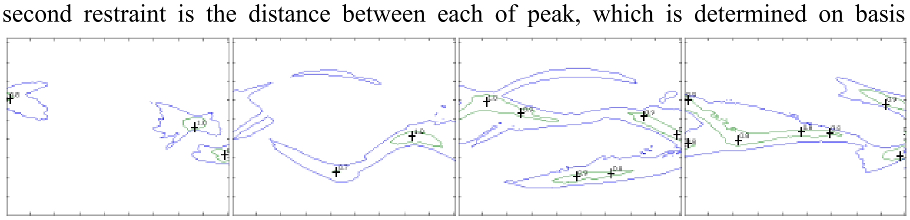

Fig. 7. The accumulator peaks and their localization for few digits from testing set.

double value of mean thickness of the characters (keeping the unitary of scale transformation). Thus, on this stage we can noticed that, a series of characters are divided into some class because of the peaks amount. The "Fig.7" depicts a Radon transformation accumulators local peaks and their localization for few sample digits from testing sets.

Another module: PCA, also uses a Radon transformation results. In this case on basis of normalized and concatenated accumulator values, are received other statistical reference elements of feature vector. The amount of data from Radon transformation depends from the image of character size and numbers of projections. On basis of introduction presented in chapter 3 can be concluded that, the size of Radon transformation accumulator table *ACC* depends mainly from two factors: size of input image (I_{mx}, I_{my}) and quantity of projection intervals θ. This means that approximate height of accumulator table x' it can be calculated by:

$$x' = \sqrt{I_{mx}^{\ 2} + I_{my}^{\ 2}} \tag{16}$$

However the width of the table depends from step θ, which is taken typically for those tasks in range from -90 to 90 degrees, assuming $\Delta\theta = 1$ and range {-89,90}, the width of table equals 180 cells. For example, we use image with size 128x128 pixels and step $\Delta\theta$ equals one degree in range {0,179}. And so, with those parameters we can retrieve accumulator matrix with 180 width and 185 height of array elements. In the our researches we don't use so many elements from the accumulator. One way to reduce of the table data is possible by the resizing operation - when generally size of matrix is most commonly decreased. The most known scaling techniques are [Keys (1981)][Oppenheim (1975)][Avidan (2007)]: method used with pixel art scaling algorithms, bicubic interpolation, bilinear interpolation, Lanczos resampling, spline interpolation, Seam carving, and many others. In our research we make tests with linear, bicubic and bilinear methods. The results with other methods was very similar and do not have influence on the recognition rate of proposed system. As a result of resize operation in our system is a n-by-n matrix, were n has been chosen experimentally on basis of recognition results. Unfortunately disadvantage of this solution is too high computational complexity. Therefore, such parameters should be chosen ($\Delta\theta$ and Δx), to obtain the same size array, as in the case of scaling operation. Practically speaking, the scaling operation will be confined to integral of several data columns array in any of each projections. In our approach, the discrete Radon transformation, is obtained as a successive columns sums, designated for the image rotated by an angle $\Delta\theta$. The obtained vectors are transposed, and formed the matrix with accumulator elements (see on "Fig.8").

To produce distinctive character features we don't use all values from the accumulator. Thus, the next step of vector features preparing is concatenate operation and Principal Component Analysis of matrix from Radon Transformation. There are at least two reasons why PCA is effective in our solution. First, the statistical analysis of the accumulator values allows to specify the array parameters, which are important during the classification process, and secondly, it has a good generalization properties. As a result of Principal Component Analysis module is n element vector $P=\{L_1,L_2,..,L_n\}$ of main values from input data, which will be used to feature vector. Additionally, beside

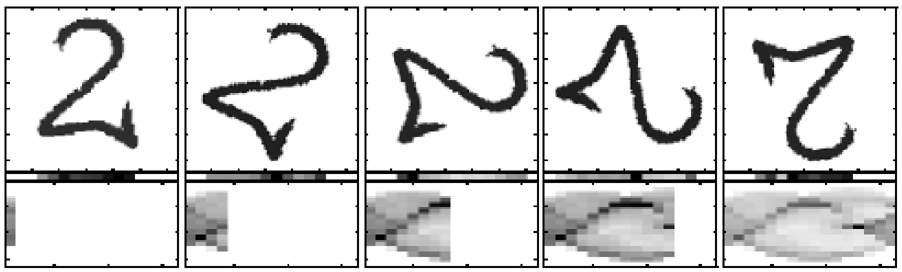

Fig. 8. Radon transformation module: example projections for *1,45,90,135 and 180 degrees.*

the number of peaks *LP* from the Accumulator Analysis stage, vector contains code of known character *ZN* as a Unicode [UTF-8 (2003)]. After all, the feature vector consists of values $FV=\{ZN,LP,P\}$ for each character from training set.

6. Module of preliminary classification

The aim of the preliminary classification is to reduce the number of possible candidates for an unknown character, to a subset of the total character set. For this purpose, the selected domain is categorized into six groups with number of local maximum. It is worth mentioning that, after applying the preliminary classification, the number of wrongly classified characters was decreased. The analysis of the elements belonging to different groups, does not allow to indicate the clear membership rules classes of character, but rather may show their geometrical features. Additionally, pre-classification module can be used to determine rejection of non-digit character. For instance, on the basis of information about number of peaks we can reject unknown objects, which too many peaks. In the near future the author intends to develop a module of preliminary

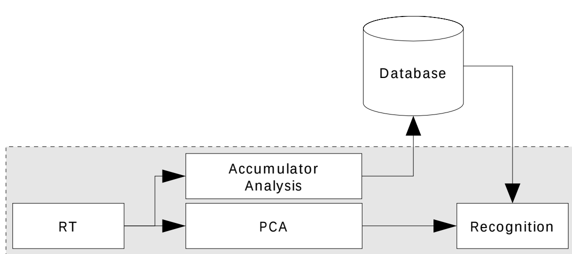

Fig. 9. The Preliminary classification scheme.

classification based on the amount, value and the coordinates of peaks. In summary of this paragraph, current method of preliminary classification based only on the quantity of local maximums (calculated in the Accumulator Analysis module), see on “Fig.7”.

7. Character recognition stage

Information from pre-processing module allows to chose a group of objects which appropriate number of peaks. Then a group of well-known characters is transferred to the classification module. The classification in the recognition module compared features from the pattern to model features sets obtained during the learning process. Based on the feature vector *FV* recognition, the classification attempts to identify the character based on the calculation of Euclidean distance [Turk and Pentland (1991)] between the features of the character and of the character models [Aissaoui (1999)]. The distance function is given by:

$$D(C_i, C_r) = \sum_{j=1}^{N} [R(j) - A(j)]^2 \tag{16}$$

where:
Ci - is the predefined character, *Cr* - is the character to be recognized, *R* - is the feature vector of the character to be recognized, *A* - is the feature vector of the predefined character, *N* - is the number of features.

The minimum distance *D* between unknown character feature and predefined class of the characters is the criterion choice of the character [Aissaoui (1999)].

8. Experiments

In this section we present the results of classification with proposed method. Especially for evaluation experiments, we extracted some digit data from various paper documents from different sources e.g. mail pieces post code, bank cheque etc. The character samples were scanned with 600dpi in color and stored in special data collections [Horstmann (2002)] in form 24 bit RGB and 8 bit grayscale images. It is important that in the case of images with heterogeneous background to perform directional filtering for 0, 45 and 90 degrees.Character image is normalized according to specification of the second paragraph. Based on geometric and central moments, can be achieve center of gravity and main axis angle. For experimental purposes, the character image sizes are ranged from 16x16, 32x32, 64x64, 128x128 to 256x256 pixels. Optimal results have been obtained for 128x128 pixels. Similar scenario was carried out for grayscale levels, where 2,4,8,16,32 and 64 levels were tested. In total, the datasets contain the digit patterns of above 130 writers. In this way were collected 1220 different patterns for training and testing sets.

On this stage each pattern is represented by a radon transformation table. Also for this part are two critical parameters: number of discretization levels and size n-by-n of accumulator table. We use 4,8,16,32 and 64 levels of the accumulator and *n* from 3 to 11. Finally after PCA and AA operation, each pattern can be represented as a feature vector with *n*+2 elements. The results our method for these parameters we can see on "Fig.10". These results obtained for testing 5 sets defined in ratio 32:68 of all samples (testing

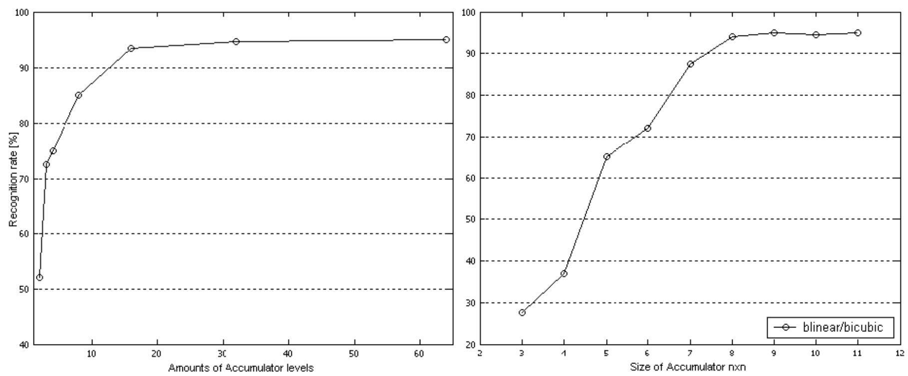

Fig. 10. The Preliminary classification scheme.

set:learning set). For similar character sets from MNIST database, the results were about 10% less (due to a lower resolution).

9. Conclusions and future work

The selecting of the features for character recognition can be problematic. Moreover fact that the mail pieces have different sizes, shapes, layouts etc. this process is more complicated. The paper describes often used the character image processing such as image filtration, normalization and Radon Transformation. The article presents an approach to optical character recognition, specifically used in the recognition of zip code digits. Although, the area is well known and explored, with successful examples of both scientific and commercial implementation, however efficiency of mail sorting systems is imperfect. The author hopes that this solution may be support for other approaches such as [Liu and Sako (2002)] [Kaufmann and Bunke (2000)] [Aissaoui (1999)] task [Bellili and Giloux (2003)]. The most common optical character recognition methods are based on modified quadratic discriminant function, hidden Markov models, normalized Fourier descriptors, MLP-SVM. The task of comparing the results for handwritten character with other researches is a difficult work because are differences in experimental methodology, experimental settings and handwriting databases, therefore, generally they are treated as "lab results". The main advantages of the method are: finding geometric relations of the character by Radon transform, invariance to background noise, low computational complexity, working with grayscale images. Disadvantages: low value of the rejections, unclear data reduction from PCA, need to use preprocessing. Further work will include

Rough Sets theory and Trace Transformation functionals tools to obtain invariant features. Moreover the data set will be upgraded to all alphanumerical signs.

References

Aissaoui A. (1999). *Normalised Fourier Coefficients for Cursive Arabic Script recognition*, Universite Mohamed, Morocco.

Asano A. (2002). *Radon transformation and projection theorem*, Topic 5, Lecture notes of subject Pattern information processing, http://kuva.mis.hiroshima-u.ac.jp/~asano/Kougi/02a/PIP/

Averbuch A.; Coifman R.R. (2001). *Fast Slant Stack: A notion of Radon Transform for Data in a Cartesian Grid which is Rapidly Computible*, SIAM J. Scientific Computing.

Avidan S.; Shamir A. (2007). *Seam Carving for Content-Aware Image Resizing*, ACM Transactions on Graphics, Vol 26.

Bellili A.; Giloux M. (2003). *An MLP-SVM combination architecture for handwritten digit recognition*, International Journal on Document Analysis and Recognition, Springer-Verlag.

Bracewell R. N. (1995). *Two-Dimensional Imaging*, Englewood Cliffs, Prentice Hall, pp. 505-537.

Forella G. (2000). *Word perfect, Postal Technology*, UKIP Media & Events Ltd, UK.

Horstmann C. S.; Cornell G. (2002) *Core Java 2-Advanced Features*, Sun micreosystems Press A Prentice Hall Title

Jain A. K. (1989). *Fundamentals of Digital Image Processing*, Prentice Hall.

Jolliffe I. T. (2002). *Principal Component Analysis, Springer Series in Statistics*, 2nd ed., Springer..

Kaufmann G.; Bunke H. (2000). *Automated Reading of Cheque Amounts*, Pattern Analysis & Applications, Springer-Verlag.

Keys R. (1981). *Cubic convolution interpolation for digital image processing*, IEEE Transactions on Signal Processing, Acoustics, Speech, and Signal Processing.

Kim K. (2003). *Face Recognition using Principle Component Analysis*, DCS, University of Maryland.

Kupce E.; Freeman R. (2004). *The Radon Transform: A New Scheme for Fast Multidimensional NMR*, Concepts in Magnetic Resonance, Wiley Periodicals, Vol. 22, pp. 4-11.

Ledesma R. D. (2007). *Determining the Number of Factors to Retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis*, PAR&E, Vol. 12.

Lim J. S. (1990). *Two-Dimensional Signal and Image Processing*, Englewood Cliffs, Prentice Hall, pp. 42-45.

Liu C.; Sako H.(2002). *Performance evaluation of pattern classifiers for handwritten character recognition*, International Journal on Document Analysis and Recognition, Springer-Verlag.

Oppenheim A. V. (1975). *Digital Signal Processing*, Englewood Cliffs, Prentice-Hall.

Peter T. (1996). *The Radon Transform - Theory and Implementation*, PhD thesis, Dept. of Mathematical Modelling Section for DSP of Technical University of Denmark.

Ramos O.; Valveny T. E. (2003). *Radon Transform for Lineal Symbol Representation*, The Seventh International Conference on Document Analysis and Recognition.

Shlens J. (2005). *A Tutorial on Principal Component Analysis*, available at: http://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition_jp.pdf

Smith L. I.. (2002). *A Tutorial on Principal Components Analyzing – Discusion and singular value decomposition*, http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf

The UTF-8 encoding form (2003), *The UTF-8 encoding scheme*. UCS Transformation Format 8, defined in Annex D of ISO/IEC 10646:2003.

Turk M. A.; Pentland A. P. (1991). *Face recognition using eigenfaces*, Computer Society Conference on Computer Vision and Pattern Recognition, pp. 586-591.

Venturas S.; Flaounas I. (2005). *Study of Radon Transformation and Application of its Inverse to NMR*, Algorithms in Molecular Biology.