AN APPROACH ON RECOGNITION OF HAND-WRITTEN LETTERS

Ahmet ÇINAR , Erdal ÖZBAY

Fırat University Faculty Of Engineering Computer Engineering 23119 Elazig TURKEY

ABSTRACT

In this study, a method for recognizing of hand-written letters is presented. For this purpose, an interactive interface is created to ensure the creation of primarily in hand-written capital font characters. Direction coding method related to structural approach has been used in the developed method. First of all, the characteristic features such as; directional features, gradient of starting and ending points, and the distance of points to letters and each other are extracted. Then, the computed features are compared with the letters' features in the letter table. The drawing is defined according to the most similar letter. Main features are the defined features of that letter. According to specific rules defined in the general case, an average letters recognition rate is about 70%. For this purpose, a software based on Visual Studio C# is implemented.

Keywords: Recognition, Hand Writing, Character Recognition, Initial Teaching Methods.

1 INTRODUCTION

Since people existence, they have tried several ways to get on with each other. Firstly they have learned to speak, emerged languages. After a while, just not enough to talk to people face, they felt the need to leave message. So they have made picture on the cave walls. Pictures in time fixed and turned into an alphabet consisting of special makings, so the foundations of writing were laid. Each statement is written, the language alphabet, numbers of different characters, there are forms and rules. Indicators of the language is aimed at satisfying sound, visual indicators, which is an array of items, or writing letters, arguably humanity's most important is the discovery [1].

Main obstacle of handwriting recognition and control is in the creation of characters lack of a standard rule. In this respect, the letters height, width, angle and other parameters should be measured.But most importantly, to compare with each letter itself, or the like.a and a, b and b in the form of identical or similar letters such as h and k compared these exact size, slope, punctuation marks, letters, spacing, links between letters, words, and the distance between the lines, the page edges of the gaps left in the pen pressure on habits, etc. should be reviewed as well as other factors. Before proceeding to examine the basis of comparison, all documents should be reviewed and determined whether they were similar. Here, the characteristics are determined and the variations (natural and unnatural). [2]

Characteristics are the shapes of letters, letter combinations, basic moves (belt, crown, angle, bump), line quality, compression characteristics, slope (angle), proportion, personal moves, margins, commonly used words, special symbols, special symbols and lines , the speed, the use of paper.

Variations (natural) are letters structuring methods, determination of pen movement, the ink in the lines and spaces, in the proportions of the letters in their own words, in the proportions of the different letters, figures, linked text.

Variations (unnatural) are variations of words, mixed handwriting, signature, paper use, layout, different form letters used by the same author, the importance of variation among authors, spelling and punctuation errors [3].

The purpose of this study is to prepare in an interactive interface restricted area plotted in the form of letters printed on the capital letters in the recognition and demonstration of the online. In this study, the system is designed to recognize 23 different capitals.

2 MATERIAL AND METHOD

In recent years, the studies on writing recognition have been made a great progress. In this way, especially on the floor clean and legible printing product written in a typewriter or computer, writing programs that automatically started to enter into our lives. Recognition of many institutions and companies with a high percentage of these programs are reduced the cost and life quite easier. Alongside these developments, the current problem of handwriting recognition technology is still not fully resolved. Difficulty of handwriting recognition, have

too many different fonts and character as well as letters from person to person varies is also connected to each other due to the write-down. Font-style, state, and used a pencil or a sheet of paper may vary according to the. Persons writing letters, depending on style and speed can be very different shapes and sizes. The letters in the human visual system is not affected by differences in magnitude and direction, but an automated system, they constitute the major problems [4].

As a result of studies on character recognition, characters are created by hand or by printing, many techniques have been developed on how to recognize automatically.

The process of the character recognition is consisted of the scanning the form, handwritten and machine printed elements with the output form of the separation processes. Recognition software, text description languages are used with the help of graphic characters according to match your items. Advanced recognition engines have the capacity to learn by reading text templates over and over again. After you create the final output of the recognition, software allows the user to make changes manually [5].

Character recognition systems was first entered into the system or the screen on the raw data, provides some basic arrangements. These processes contribute to the reliability and accuracy of character recognition system is more advanced stages. In this context, the recognition of characters on the page is very important to pre-treatment phase. Thresholding of pre-treatment, the rate of noise reduction, thinning (thinning), widening (dilation), such as skeleton extraction steps can be mentioned.

Prevalent in all areas are used for character recognition applications because there are different techniques have been developed over the years [13]. Drawing of the flow diagram of the letter "F" is as follows.
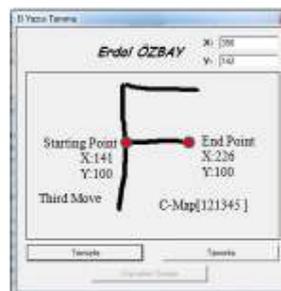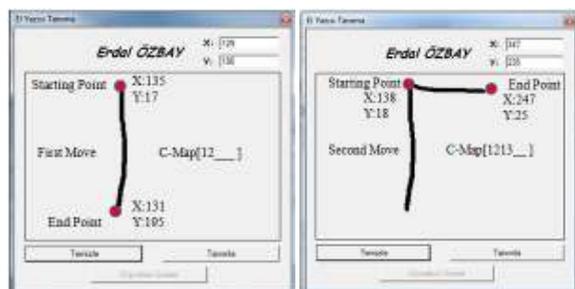
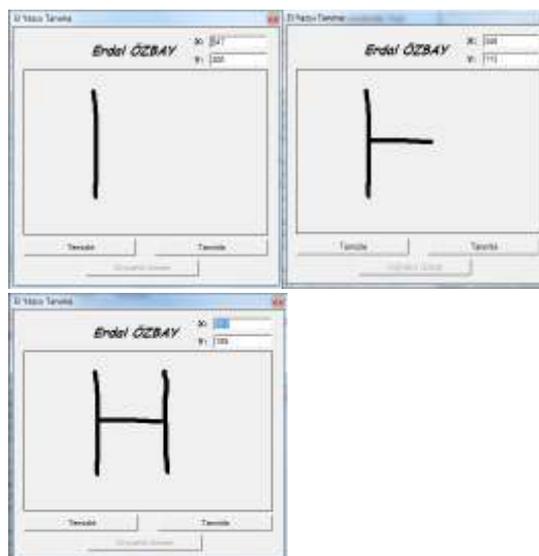**Figure 1:** Drawing of the flow diagram of the letter "F".

Application stage of the study flow diagram can be explained as following. These are;

Step-1: First of all, created an interface to create handwritten letters.

Figure 2: Program interface.

Step-2: Then we want to identify great letter start to write as interactive.

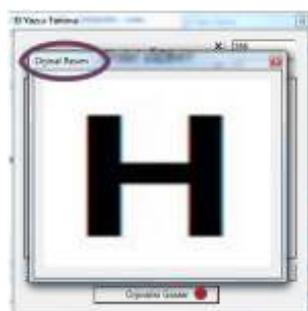**Figure 3:** Interactive drawing of the letter "H".

Step-3: After drawing the letter we want to identify by clicking on the "Tanımla" button, the

letter of the connection map is created.

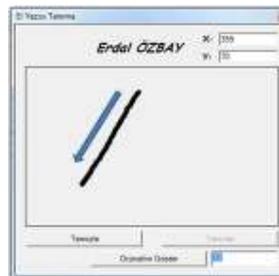**Figure 4:** Connection map of the letter "H".

Step-4: Finally, the map of the letter issued, by the fact that matches, displayed on the screen with the "Orjinalini Göster" button.

**Figure 5:** Accuracy is controlled, "H" letter of the original picture.

In this study, prepared 360x240-pixel limited interface area has been the recognition of some of the interactive drawn capital letters. For this purpose, a software based on Visual Studio C# is implemented. Direction coding method related to structural approach has been used in the developed software. First of all, the characteristic features such as; directional features, gradient of starting and ending points, and the distance of points to letters and each other are extracted. Then, the computed features are compared with the letters' features in the letter table. The drawing is defined according to the most similar letter. As the main features are the defined features of that letter, all the letters which are drawn according to general rules has been recognized.

According to specific characteristics of the letters in recognition of the stage of defining the business logic, one might say the letters "A" and "B".


Figure 6: First move of the "A" letter. Connection map 12.

Firstly, the creation phase of the letters, the letter, which will complete a number of moves, starting and ending points, and breaches of the proximity, slope, and a map was drawn up according to their regular or not. A letter is created, as shown in Figure 6. The first move, this move will breach the start and end points map 12 in the form is created and stored coordinates.
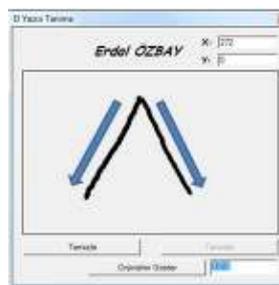
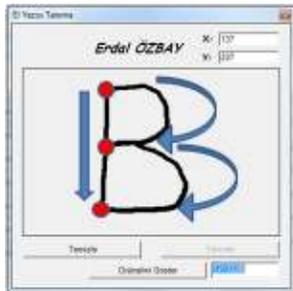**Figure 7:** Second move of the "A" letter. C- map [1213].

By taking consideration with proximity to the starting point of the second move and the starting point of the first move of the "A" letter, C-map take same value 1, while contrary to the others ending point C-map take value 3.
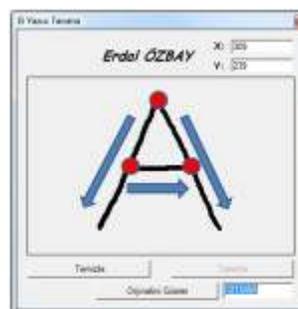
Figure 8: "A" letter after the third move the C-map[121345].

In the same way, recent thrust of A letter due to

violation of the points on the map 45 is added. After this stage, information is controlled by mapping information confirmed case. For example, recent move of the A letter by 45 in the starting and ending points are nearly equal of the middle part of the first and second moves are examined if they could have. As a result, all of a connection to the letter "A" [121345] is mapped in the form of a connection.


Figure 9: The letter "B" after the third move. C-map[121332].

If the letter "B" after the third move has the value of the C- map[121332], this time examined the relationship between connections. For example, end point of the second move and starting point of the third move equal to 3, it must have coincided between starting and end mid-point of the first move. Provided that this requirement added to the map C-map[121332B] as justified.

First of all, the user interactively drawing the letter, according to the movement of the x and y coordinates of pixels for the information kept separately. This pixel data with the four-way limits, the number of points, the points were calculated distances from each other. Characteristic features of the most basic feature are the directional information.

On-line handwriting, aspects of the user goes, the letter moves creating movement and created few steps, the letters between the connection points are able to predict based on compliance.

During the construction of the letters, according to the harmony of the moves in the beginning and ending points for each letter in a unique connection between the map is removed.
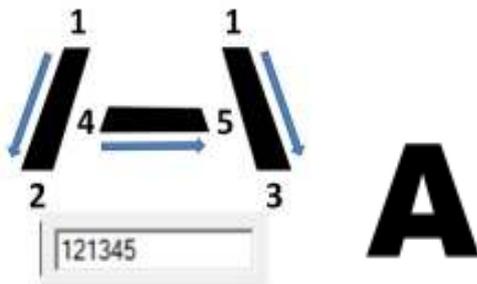
**Figure 10:** Flow diagram by the letter "A" and letter"A".

We have made moves to create a letter, in order of proximity to the start and end points, giving the same or different values. For example, the second step in which with the end point of the 4 for letter K, the starting point of the next move is assuming close to each other, were adopted on the same connection.
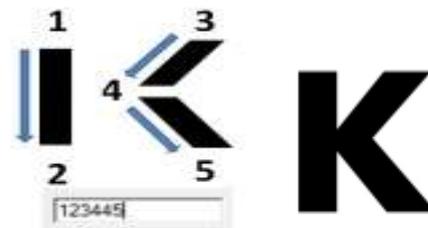
**Figure 11:** Flow diagram by the letter "K" and letter"K".

3 COMPUTING OF CONNECTIONS

Removal of the map, during the creation of the letter, which starts with interactive hits, ending with release of the first moves, and the latter, there is a lack of intimacy is based on the calculation. For example, 121345 connection map, and act with the logic of the letter "A"-like structures are examined, which last moves the center line of the start and end points [Figure12. (4.5)] corresponds to the first and second moves to a need to points close to the middle level.
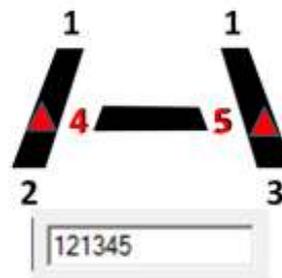
**Figure 12:** "A" letter, control points of connections.

Similar conditions are met for all the other letters. In addition to the conditions and limitations to performance of the control condition in case the connection map of the letters that are included with the map.
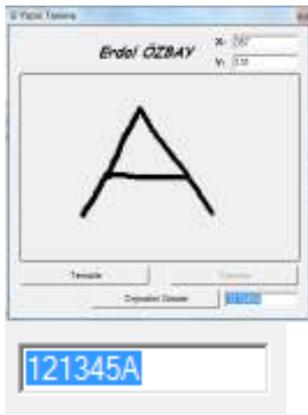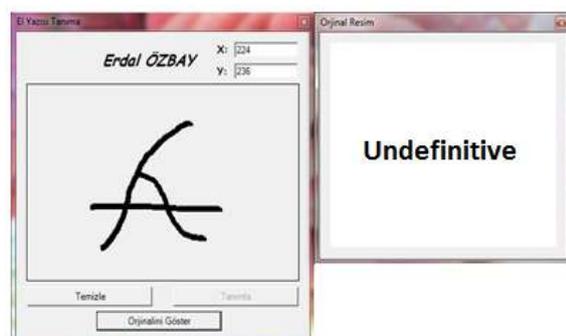
Figure 13: Control is provided "A" letter to be included with the map.

Figure 14: The drawn character, "Tanımla" of the map display after running.
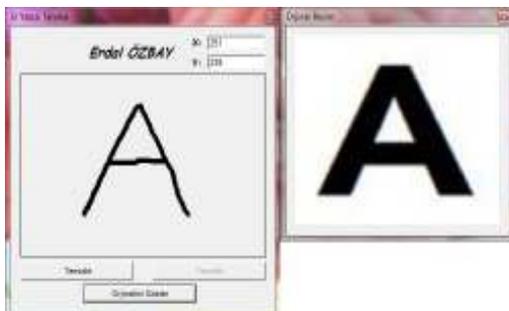
Figure 15: Defined in the original character matching the letter.

Figure 16: The drawn character shown by not to

matching with the original "Undefinitive" statement.

4   DISCUSSION

Handwriting characters recognition by computers has attracted less attention than the optical characters. Hand-written character recognition to increase the performance of two important goal should be to reach. A feature extraction method does not miss significant details and the generalization power of design, but to distinguish between high-low classifier error detection, some special because of the characters in the alphabet (L, U, I, G, T, O, etc.), the recognition of characters, some recognition of the Latin alphabet more difficult problem [6].

The character recognition problem, a variety of security systems, bank checks, postal codes, vehicle plates and barcode reading, form filling, office automation (handwritten or computer printout paper to the electronic environment and again), hand writings of post processing of the screen, such as electronic-components of a common can be applied [7,8].

Facilities of the character recognition technologies, have helped to these technologies are rapidly beginning to be applied in many other business. For example, the addresses of letters above them are automatically recognized and parsed by zip code, automatically checks sent to banks required to recognize and account transactions such as electronic media, the realization of projects carried out by the character recognition technologies.

In recent years, great progress has been writing about recognizing the studies. In this way, especially on the floor clean and legible printing product written in a typewriter or computer, writing programs that automatically started to enter into our lives. Recognition of many institutions and companies with a high percentage of these programs reduce the cost and life quite easier. Alongside these developments, the current problem of handwriting recognition technology is still not fully resolved. Difficulty of handwriting recognition, have too many different fonts and character as well as letters from person to person varies is also connected to each other due to the write-down. Font-style, state, and used a pencil or a sheet of paper may vary according to the. Persons writing letters, depending on style and speed can be very different shapes and sizes. The letters in the human visual system is not affected by differences in magnitude and direction, but an automated system, they constitute the major problems [10,11,12].

The program is simple, and being available to all users in order to address an important advantage.

Stages of identification and assimilation of some letters in the application are shown below.
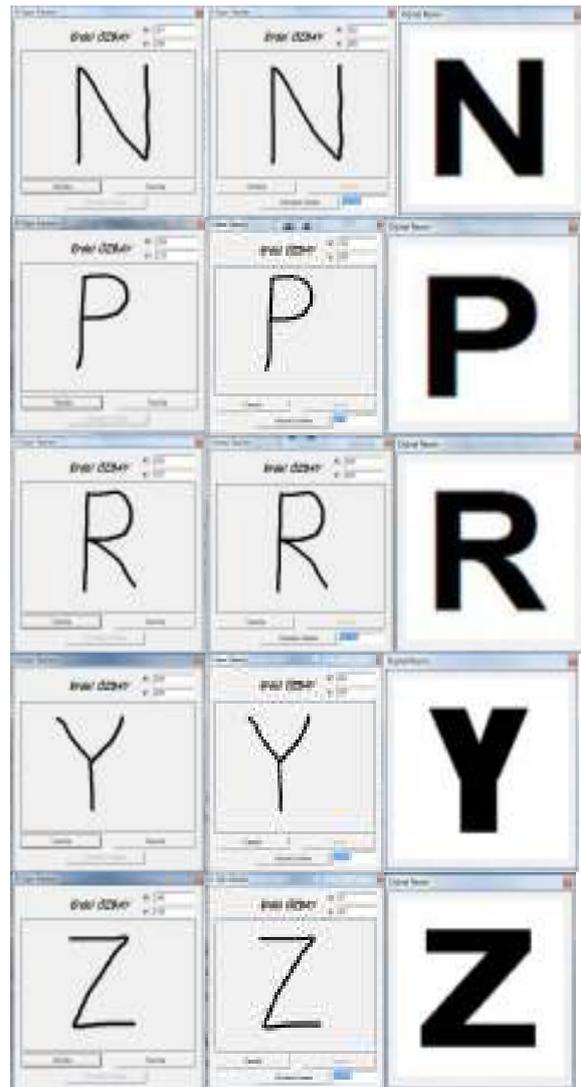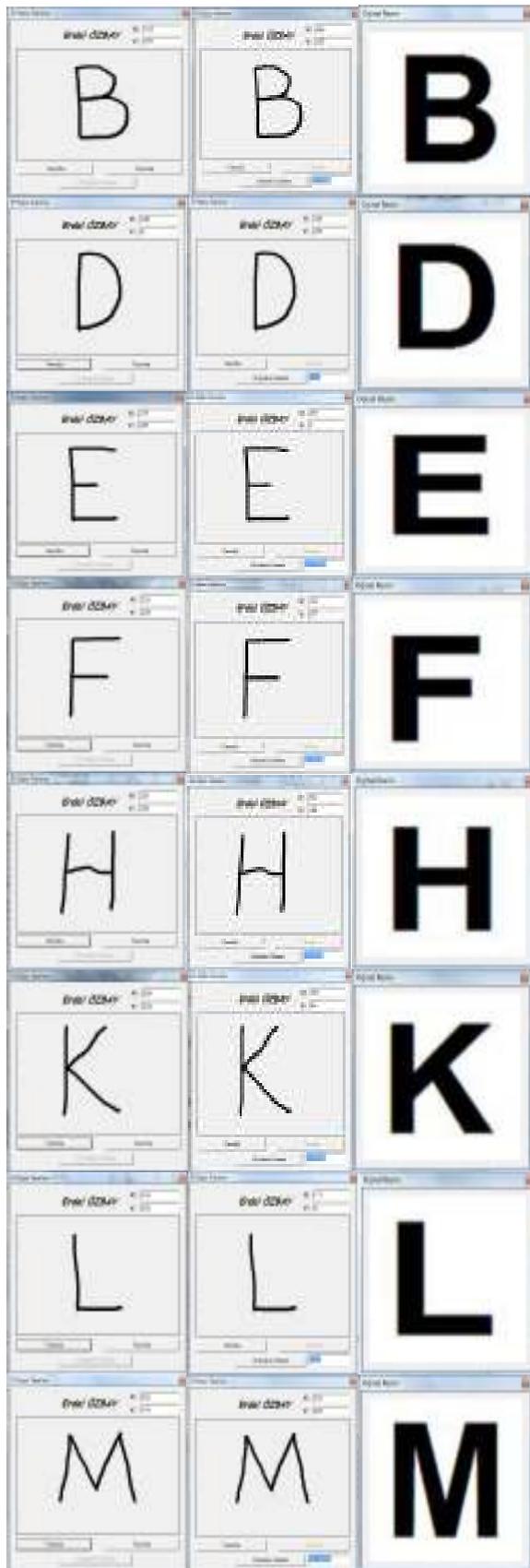
Figure 17: (From left to right) 1-Character creation, 2-Character identification process is approved to perform 3-Accuracy of the real character of the character matched.
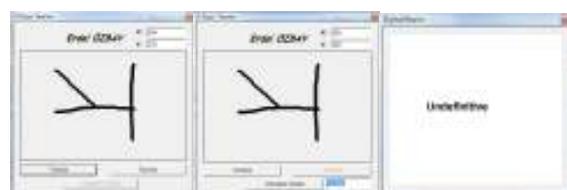
Figure 18: (From left to right) 1-Character creation, 2-Character identification process is approved to perform 3- The accuracy of the character that is not approved, the result does not match the "Tanımsız" expression.

5   CONCLUSION

In this paper, we designed an interface limited interactive space. The program interface is prepared on   visual programming language using Visual Studio C #. Encoding method is used in the program written in the direction of the structural approach. First, the user will draw on the interface of pixels is limited to 360x240. The user interface is interactive while simultaneously drawing in which directions were calculated. For this purpose, the following by the user were transferred to an array of x and y coordinates. In order to find the direction of changes in the values of x and y values are compared to the pixel coordinates were continuously monitored. This value was recorded in the knowledge of the direction of the letter. Other points in the direction of the case for the recognition of the character's knowledge of the location, such as number, letter to each other and the distance from the calculated values are computed. This study identifies an average recognition rate of the each letters according to certain rules to the letter was found to be about %70.

The prepared program on personal computers (PC), are available on Electronic boards, Tablet PCs. Mobile phones and handheld computers (PDAs) can be used when the necessary adaptations. As a result, increasing the accuracy of the attribute to extract the characteristics of the letter and the pixels are drawn as an interactive computer system is accelerated to increase the rate of getting a higher recognition rates can be obtained. System in different programming languages can also be used to speed up.

6   REFERENCES

1. Büyükbayrak H. , 2005. Online Handwritten Mathematical Expression Recognition, Master Thesis, Sabanci University
2. On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey Rejean Plamondon Sargur N. Srihari JANUARY 2000
3. Characteristics of Students in the First Grade of Primary School and Initial Teaching of Reading andWriting Aynur BİLİR Ankara University, Journal of
   Faculty of Educational Sciences, year: 2005, vol: 38, no: 1, 87-100
4. FLink, Gernot A and Plotz, Thomas, 'Unsupervised Estimation of Writing Style Models for Improved Unconstrained Off-line Handwriting Recognition',
   International Workshop on Frontiers in Handwriting Recognition, Suvisoft, France, 2006
5. Gunter, Simon and Bunke, Horst, 'Off-line Cursive Handwriting Recognition - On the Influence of Training Set and Vocabulary Size in Multiple
   Classifier Systems' Elsevier Volume 43, Issues 3-5, March-May 2005, Pages 437-454
6. Araokar, Shashank, 'Visual Character Recognition using Artificial Neural Networks', CoRR, Vol abs/cs/0505016, 2005
7. Hewavitharana, S.; Fernando, H. C. and Kodikara, N.D., 'Off-line Sinhala Handwriting Recognition using Hidden Markov Models', Indian Conference
   on Computer Vision, Graphics and Image Processing, India, 2002
8. Shi, Daming; Shu, Wenhao and Liu, Haitao, 'Feature Selection for Handwritten Chinese Character Recognition Based on Genetic Algorithms',
   International Conference on System, Man, and Cybernetics, 1998
9. A. Amin, "Recognition of Hand-Printed Characters Based on Structural Description and Inductive Logic Programming," Pattern Recognition Letters, vol.
   24, pp. 3187-3196, 2003
10. S. Mozaffari, K. Faez, and M. Ziaratban, "Structural Decomposition and Statistical Description of Farsi/Arabic Handwritten Numeric Characters," Proc.
   Int'l Conf. Document Analysis and Recognition, pp. 237-241, 2005.
11. H. Bunke, S. Bengio, and A. Vinciarelli, "Off-Line Recognition of Unconstrained Handwritten Texts Using HMMS and Statistical Language Models,"
   IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, pp. 709-720, 2004.
12. Chevalier, S., Geoffrois, E., and Preteux, F., 2003, A 2D Dynamic Programming Approach for Markov Random Field-based Handwritten Character
   Recognition, Proceedings, pp. 617 630, IAPR International Conference on Image and Signal Processing (ICISP'2003), Agadir, Morocco.
13. Liu, C-L. and Fujisava, H., 2005, Classification and Learning for Character Recognition: Comparison of Methods and Remaining Problems, pp. 1-7,
   First IAPR TC3 NNLDAR Workshop, Seoul, Korea,