International Journal of Electrical and Computer Engineering (IJECE)
Vol.2, No.3, June 2012, pp. 383~394
ISSN: 2088-8708

Standardizing, Segmenting and Tenderizing Letters and Improving the Quality of Envelope Images to Extract Postal Addresses

Milad Shirmohammadi*, Karim Faez, Hassan Khotanlou***
* Departement of Electrical and Computer Engineering, Qazvin Branch, Islamic Azad University
** Departement of Electrical Engineering, Amirkabir University of Technology
* Departement of Computer Engineering, Bu-Ali Sina University

Article Info

Article History:

Received Mar 31st, 2012
Revised May 25th, 2012
Accepted May 30th, 2012

Keyword:

Image rotation
Address extraction
Letter tenderization
Automated postal systems
Handwriting recognition

ABSTRACT

In most mechanized postal systems, envelopes are scanned based on the postal standard using mechanical instruments. In the standard format, the image of envelopes lacks tilts, lines are along the horizontal axis and words are placed in a correct and non-oblique manner. In this article three new algorithms for rotating, segmentation and Tenderizing Letters for standardizing and increasing the quality of an envelope have been presented, which can be used in all text identification systems as three successful pre-processing algorithms. In the algorithm proposed, letters with any forms and tilts during scanning were rotated and standardized by applying a simple two-step algorithm based on what was written on the envelope without requiring the calculation of tilt angle. After standardization, the main regions of the image were specified using the histogram information. Then, in a simple algorithm, the candidate points from the pixels related to the text on the envelope were selected and quality improvement and tenderization were done on the main regions of the image. The advantaged of the proposed algorithm included No need for additional mechanical equipment, less calculation, simplicity and consideration of the structure of words on the envelope in all preprocessing phases.

Corresponding Author:

Milad Shirmohammadi,
Departement of Electrical and Computer Engineering,
Qazvin Branch, Islamic Azad University,
Nokhbegan Blvd, Qazvin, iran.
Email: mshirmohammadi@qiau.ac.ir

1. INTRODUCTION

With the progress of economy, growth of population and increasing number of postal instruments, traditional and manual postal systems are not efficient anymore and it is required to replace new and small automatic systems. Mechanized post was proposed since about 50 years ago in developed countries like Japan, America, Germany and so on and it has made great progress thus far [1]. Current systems use a mechanical separator for extracting postal addressed. This mechanical device removes the probability of angular scan of letters [1]. The biggest problem in developing the mechanical postal systems is massive and expensive instrument and irregular and untidy written words on an envelope. This default is observed in all text detection systems. In this paper, three activities of standardization (determining envelope edges and rotation based on the status of words on the envelope), segmentation and tenderization of letters, which are the main pre-processing phases, for solve this problem were examined. During these three processes, first, the image was scanned without any limitations and was changed to a standard format by applying the proposed algorithm without any need for mechanical equipment. After standardizing the image, the envelope image was mainly divided to three regions of sender address, recipient address and stamp using the image histogram information. Finally, after determining main regions, letter tenderization and quality improvement of these regions were conducted in order to prepare the envelope image for being entered into a letter

identification system. In a normal or standard postal situation, envelope image lacks tilting and the field line related to the addresses written on the envelope is along the horizontal axis. Also, words have a correct form and lack any inversion [2-4]. Standardization, segmentation and tenderization of letters are integral components of all text identification systems [5-7]. The existing algorithms usually detect the boundary of scanned documents, calculate the angle between this boundary and horizontal axis and rotate the image with its negative angle. This process requires many calculations; since rotation is done without considering the way texts are written on the documents, re-rotation is probably required in the following stages. In the proposed algorithm, there was no need for calculating the tilt angel of envelope for rotation and, due to the consideration of addresses written on the envelopes, re-rotation in the text identification phases is not needed while rotating. After this phase, the main regions of the standardized envelope image were extracted and letter tenderization was done on them. Most tenderization algorithms attempt to recognize the composing components of the letter skeletons and then extract the letter skeleton by placing these components beside each other. Two samples of this group of algorithms will be introduced below. In the algorithm of the first sample, first, a series of lines, curves and closed loops is considered from the image of letters and then the skeleton of letters is extracted by placing these forms beside each other [8]. In another sample, an algorithm is presented which defines some standard forms for itself and constitutes the skeleton of all the letters by putting these forms by each other. Therefore, in this algorithm, some fixed forms should be extracted from the image of letters in order to obtain skeleton of letters by placing these images beside each other [9]. There is another group of algorithms which do not search for any special form in the letter image; they remove extra pixels from letters and extract skeleton of letters; the proposed algorithm of this paper was in this group. Letter recognition and tenderization algorithm usually extract structural features [10-13]. The advantaged of the proposed algorithm included No need for additional mechanical equipment, less calculation, simplicity and consideration of the structure of words on the envelope in all preprocessing phases. Moreover, the proposed tenderization algorithm revealed weaker performance for the images with high noise due to the predicted mechanized postal system.
This paper is organized in the following way:

At first, the four corners of the envelope are determined. Then, the kind of field line recognition and the position of word contours in a specific line are expressed. After that, the image standardization and rotation algorithms are described in two phases and image segmentation is expressed using the histogram information of the image. Also, letter tenderization algorithm is explained. After that, the proposed method is evaluated and conclusion is offered based on the proposed algorithm and its applications.

2. RESEARCH METHOD

The general phases of the proposed algorithm for standardizing envelope images are given in Fig. 1.

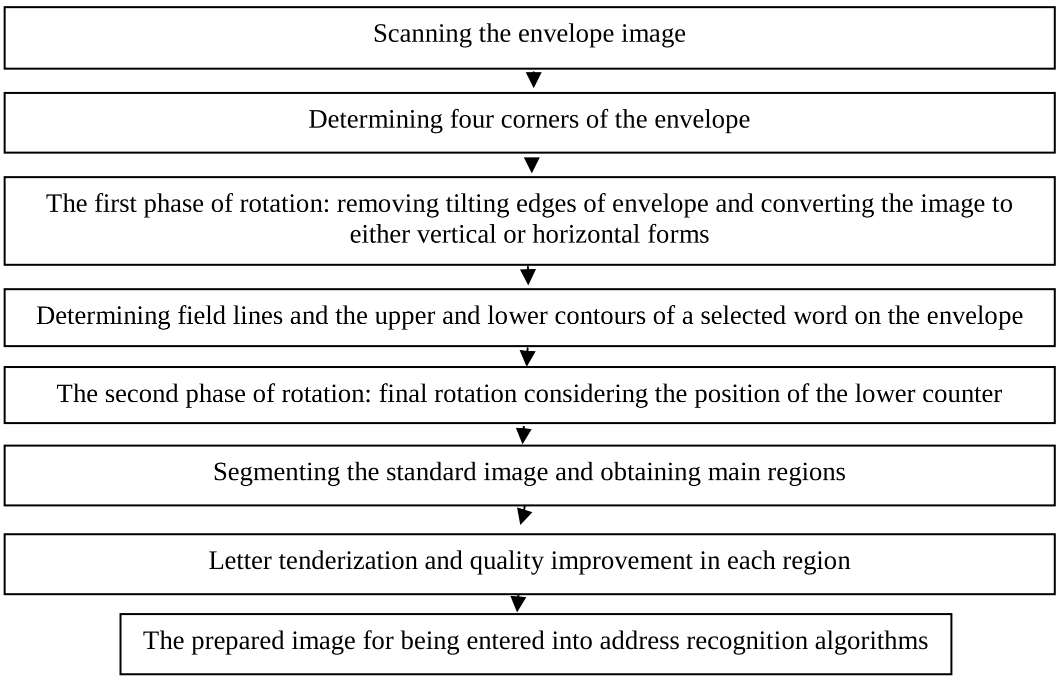


Figure 1. Standardization phases of the envelope image based on the proposed algorithm

2.1. Determining the Envelope Boundary

In most methods, changes in the pixel intensity are used for determining the boundary of the regions in the image [14].

In this paper, the positions of the four corners of the envelope and its boundaries were obtained by changing color intensity of pixels on the image.

If the equation related to the four edges of the envelope were obtained, the positions of four envelope corners could be determined using the intersection points of these lines. This was done according to the three steps shown in Fig.2.

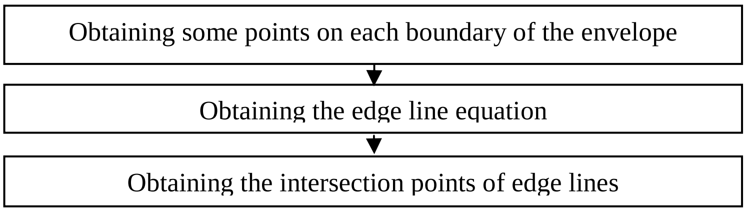


Figure 2. The required steps for obtaining the four corners of the envelope

2.1.1. Selecting points of edges:

The colors of image background and envelope were different from each other. Therefore, some points could be easily selected on the boundary of intensity changes (edges) using the difference of color intensity of pixels. However, due to the dirty or opaque envelope edges, when intensity changes were used for obtaining the position of points on the edges, the obtained points might not be on the same line. Thus, this problem should be solved while obtaining the edge line equation. Considering the dark color of background and white color of envelope, the diagram of intensity changes was used for obtaining two points on the right and left edges of the envelope. If the diagram of intensity changes were obtained for one line of the image, the values on the right and left sides of this diagram would be zero, but they would turn toward 255 at once. Therefore, the position of value changes from zero to 200 and vice versa was selected as the points on the right and left edges on the leftmost and rightmost points of the diagram. In Fig. 3a, a part of an envelope is depicted. For one of the lines of this image, the diagram of intensity changes was obtained, as shown in Fig. 3b. As demonstrated in the figure, the diagram had high intensity changes around the points 259 and 966 and these points were considered the points on the right and left sides of the envelope.

This process was repeated several times for right and left and top and bottom edges in order to obtain several points on each edge.

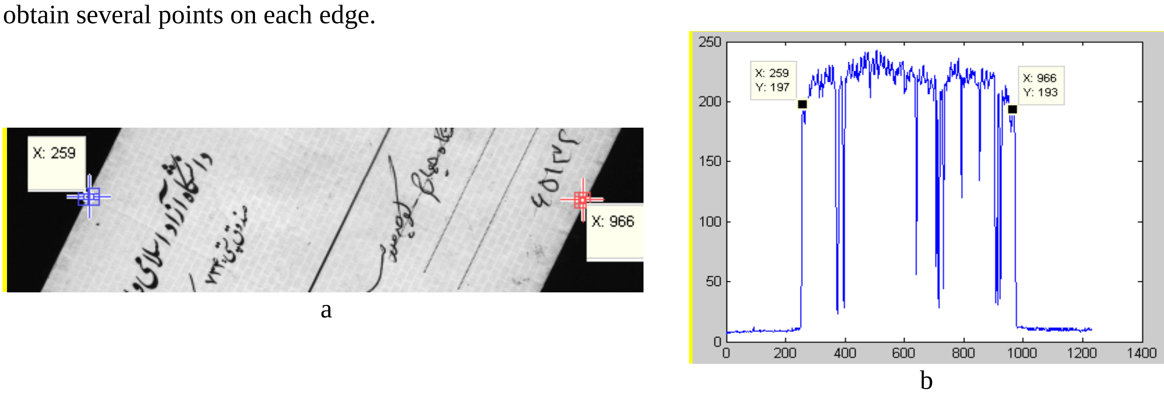


Figure 3. a) A part of envelope image scanned in an oblique way; b) Diagram of intensity changes of the pixels related to one line of envelope image

2.1.2. Obtaining edge line equation:

Using the points on each obtained edge, line equation was made two by two; then, the slope of the resulting equations was averaged using the geometrical average and the equation of average line was obtained.

2.1.3. Obtaining the intersection of edges:

To obtain the four corners of the envelope, the following stages were followed:

[i] Obtaining top left corner: left boundary line and top boundary line equations were equalized and their intersection was obtained.

[ii] Obtaining top right corner: right boundary line and top boundary line equations were equalized and their intersection was obtained. Two other corners were also obtained by this method.

2.2. The First phase of Rotation: Removing Envelope Tilt

If envelope image is considered as a tilted matrix, it should be rotated, its tilt should be removed and they should be turned to a correct form in this phase. Each line of the envelope image is transferred to a new image; the only difference being that the lines of the new image lack any tilts. The lines of envelope images are transferred from bottom to top to the new image.

According to the governing rules on the equation of a straight line, if the position of two points of a line is present, its equation can be obtained. Envelope images have four corners and their positions were obtained in phase 2.1. Using the address of left corners on the top and bottom, the left boundary was obtained; using the address of right corners on the top and bottom, the right boundary was obtained. Here, there were two lines which specified the edges of the envelope. In Fig. 4, A and B red lines specify right and left boundaries of the envelope. The start was from the bottom of the envelope and, in each step, two points were selected from the left and right boundary determining lines (A and B lines) (Fig. 4, the points shown by "*") and an equation was made by them (Fig. 4, the line that connects two symbols of "*"). The obtained line was one of the rows in the initial image (oblique image or tilted matrix). This row of the initial image (the obtained line) had an amount of slope or tilt, which should be removed in the final image. Therefore, after obtaining the line, its path was scrolled in the initial image and the intensity of pixels on this line was transferred to a row in the new image. These two points made a line equation. After obtaining the line, its path was scrolled on the initial image and the intensity values on this line were transferred to a row of the new image. With each step, a row of the new image was created. This process continued and, in each step, two new points were selected from left and right lines and the obtained line was transferred to a row higher than the new image. The rows which were selected from the initial image had some slope; however, when these rows were transferred to a new image, their slope should be removed. At the end of this phase, the envelope form was vertical or horizontal without any tilts.

It is easy to detect the horizontal or vertical position of the envelope since it is a rectangle. In fact, the histogram produced from the intensity of word pixels can be used for identifying the vertical or horizontal orientation of the field line.

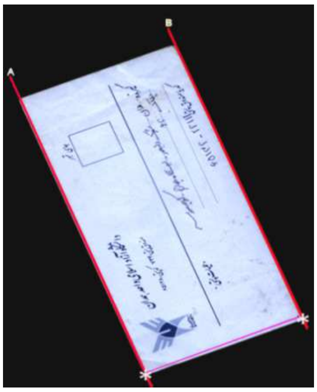


Figure 4. Image of a tilted envelope, the left and right boundaries of which were identified, and two points were selected on these two boundaries using symbol *

2.3. Recognizing Field Line and Words' Upper and Lower Contours

If a histogram is produced using the intensity of word pixels, the components of the total histogram make a curve. If the cure is divided into two parts from the minimum point which shows the position of the field line of the word, the top of the word (top contour) would have a larger range compared with the bottom (bottom contours). This issue in indicated in Fig. 5.

In this phase, a box of the image was selected. This box should contain at least one word. The sender address is usually written in the middle of the envelope. Thus, this is a good place for the position of the selected box. In the selected box, if the orientation of texts is vertical, the intensity of pixels are summed for all the columns of the selected box in order to make a vertical histogram by placing the sum of intensities beside each other; if the orientation of texts is horizontal, the intensity of pixels are summed for the all rows of the selected box in order to make a horizontal histogram by placing the sum of intensities beside each other [2], [15]. If there are several lines in the selected box, several curves are formed in the histogram [3]. The points identified as the minimum points are the field lines related to the lines inside the box.

The existing curve in the histogram was divided into two parts from the minimum point; and the left and right ranges could be used for continuing the process. The smallest point of the histogram was related to the field line of the largest line of the selected box. Smaller and larger ranges showed the positions of word bottom and top, respectively. The output of this section was used in the second phase of rotation (Section 2-4). Fig. 6 demonstrates a box with the width of 200 pixels. This box was selected from the image of a vertical envelope and contained a part with two-line text over the envelope. In this selected box, the text on the right line was thicker than the text on the left line. If the intensity of this box was summed column by column, the histogram in Fig. 4 would be obtained. As shown in Figs. 3 and 4, there were two-line texts in the selected box and the 131st column of the histogram showed the field line of the largest text in the selected box. Two curves were formed in the histogram; if one of them, which had the minimum point, was divided into two parts from point 131, then, the right range would be smaller and the left one larger. So, the contour of the bottom words of this line was on the left side.

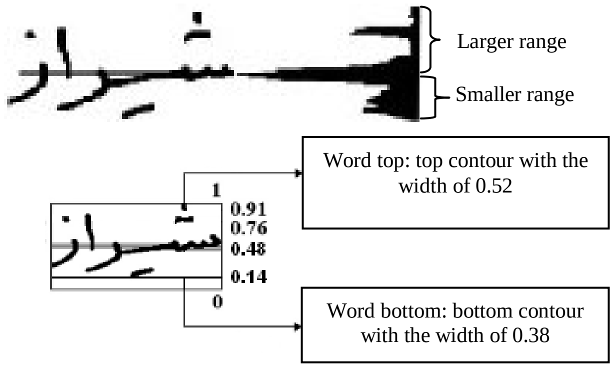


Fig. 5: Histogram structure relative to word status

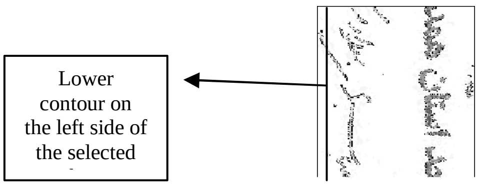


Figure 6. Selecting vertical box from the middle of a sample envelope
(its text was written from top to bottom.)

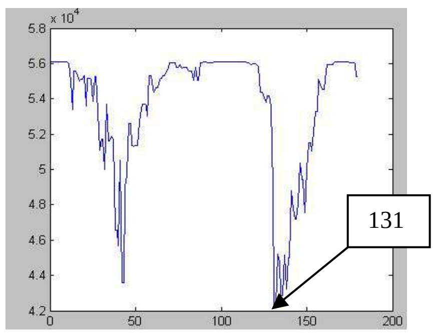


Figure 7. A sample of a histogram for a vertical image

2.4. The Second Phase of Rotation: Final Rotation

Final rotation was done considering the field line and position of contours of words on the envelope. In this stage, the image was rotated so that the written addresses would be in their correct places and the field line would be along with the horizontal axis of the image and the words existing in the address would be in a correct and non-inversion form.

To determine the type of rotation, at first, the algorithm in Section 2-3 was applied to the envelope. By applying this algorithm to the envelope, the position and orientation of the lower contour and upper contour were marked. After determining the placing position of words, the rotation of 90 degrees was made in the vertical envelope toward the lower contour. In the horizontal envelope, if the words were inverted (the bottom contour was at the top of the field line), the envelope image was rotated 180 degrees clockwise; otherwise, there was no need for rotation. It was very simple for 90 and 180 degree rotations and the image was only scrolled row by row once and the intensity of pixels was recorded in the new image. For instance, if the lowest row of the initial image from left to right were recorded in the highest row of the new image from right to left and all the rows were in this manner, then the image would be rotated by 180 degrees. Also, if the first row of the image from right to left were recorded in the first column of the new image from top to bottom and all the rows were in this manner, then the image would be rotated clockwise by 90 degrees. The output of this stage was the standardized envelope image which could be used for segmentation.

2.5. Image Segmentation

The sum of intensity of pixels was obtained column by column or row by row and two histograms were made by placing these intensities beside each other [2], [15]. Maximum points of the horizontal histogram were considered as a candidate for the horizontal boundary of the image regions; around these maximum points, there was a minimum point with the distance less than α. If there were two minimum points before and after the maximum point in the horizontal histogram with the distance of less than α, then, this point would be the boundary between the two points and would not be a component of the boundary of regions (Fig. 8, the maximum point related to the boundary between the lines).

For a maximum point in the horizontal histogram, if there were a minimum point on the one side with the distance less than α and there were no minimum point on the other side with the distance of less than α, then, the current maximum point would be the horizontal boundary of one of the main regions of the envelope. α is the threshold level for the distance between the boundaries, which can be usually estimated based on the distance between minimum points in the diagram.

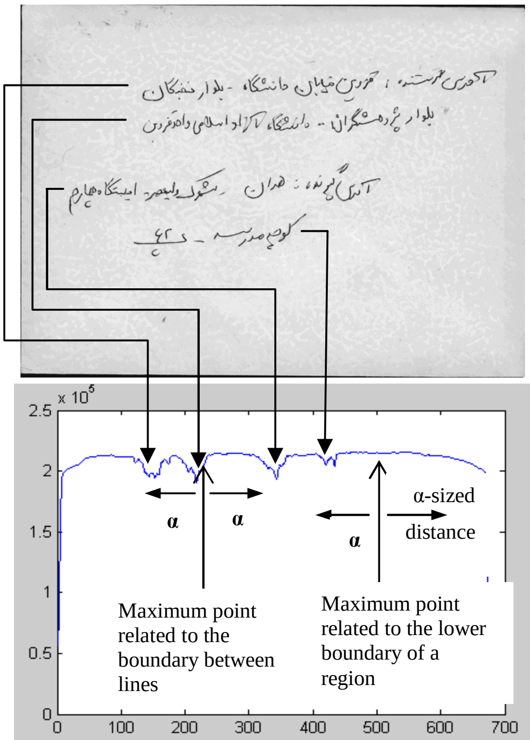


Figure 8. Image of an envelope, the horizontal histogram of which was obtained, and its two maximum points were identified and shown as a sample

After determining the horizontal boundaries of the image, there was a movement toward left or right sides on these boundaries and, when a near-maximum point was reached in the vertical histogram, that point was considered as the vertical boundary of that region.

The proposed segmentation and standardization methods were applied to a non-conventional postal envelope and the results are shown in figure 9 a to d.

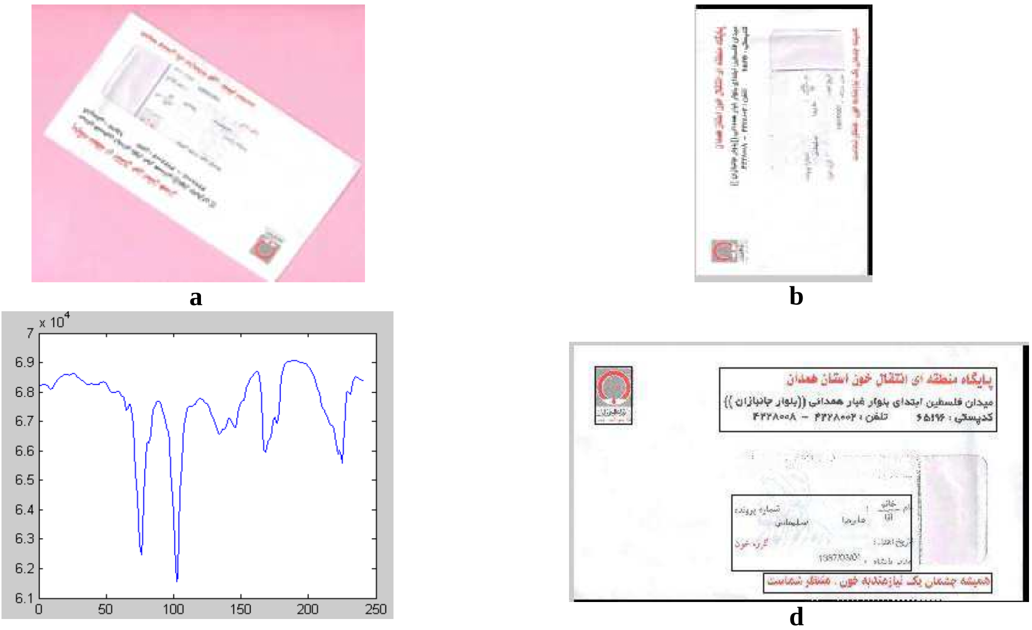


Figure 9 a) Image of a tilted and inverted scanned envelope, b) the image after applying the first phase of rotation, c) the histogram related to a box with the width of 250 pixels from the middle of the image in b, d) the segmented image

2.6. Tenderizing and Quality Improvement

Most of text identification algorithms require a proper tenderizing method. In fact, there are some algorithms which do not use letter tenderization before the text identification process [16]. Different methods have been presented for letter tenderization thus far, most of which are based on letter structure [10], [11]. In most of these methods, letters are analyzed into some main components after tenderization; these components should be connected to each other again. There are some other methods which extract main points on the letters and then connect them to each other [12], [13]. In the previous section, the main regions of the envelope were identified. In this stage, there was an attempt to transform all pixel values in each region to either zero or 255. While doing this activity, letter tenderization was also done, meaning that the thickness of letters was transformed into one pixel.

The pixels of each letter were divided to a number of groups or components. From each group, a number of pixels were selected as representative or candidate points. The intensity of candidate points from each group was equal to zero (black color) and the remaining points of the group were considered equal to 255 (white color). Then, first, a method was presented for determining the font size. Then, to group and select the candidate points, three models were defined. All the groups extracted from the letters were in the form of one of the three models. One model had one-row groups; but, the other two could have groups with more than one row.

Determining Font Size:

To determine font size, several columns were randomly selected in each identified region of the image and the diagram of their intensity was obtained. In the selected columns, the number of neighboring pixels in each diagram which had the intensity difference of 80 (a difference at the one-third level of color intensity range) from the minimum amount (background color intensity) was counted and the number with the highest number of repetition in different diagrams was considered as the font size of that region. Fig. 10a is related to one of the main regions of envelope image, from which several columns were randomly selected. After obtaining the diagram of intensity changes for these columns, the highest number in the diagrams which was the number of repeated neighboring pixels was 5. Therefore, the font size was selected as 5 pixels for this sample. The diagram of intensity changes for the three samples of columns is given in figure 10b.

Grouping Letters and Selecting the Candidate Points of Each Group

To identify the neighboring pixels related to letters, the diagram of intensity changes was used. In this stage, the pixels of each letter were divided to several groups and extra pixels were removed from each group; thus, the width of texts on the envelope became in the size of one pixel. To identify the groups, the image was scrolled row by row; using the diagram of intensity changes, wherever the intensity of several neighboring pixels got closer to zero and severely differed from the color intensity of image background, then those pixels were considered as the first row of one group. Then, based on the following states, group type and probably pixels of the following rows in each group were also identified:

If the number of pixels was less than or equal to the font size, consequently, a one-row group would be composed. Only one pixel from this group would be selected as the candidate point. After determining the candidate point, its intensity would be considered zero (black color) and the remaining intensity of pixels in that group would be assumed equal to 255 or white color. In Fig. 11a, a diagram of intensity changes (Fig. 12a) is drawn for the second row. As can be observed, columns 7 and 8 had an intensity difference of more than 80 from the background color (the minimum level). In this example, since the font size was tripled and there were two neighboring pixels, therefore, these two pixels constituted a one-row group. In another sample, if intensity changes (Fig. 12b) were drawn for row 5, then, three groups of pixels with the conditions mentioned in the previous sample could be extracted. In Fig. 11a, the pixels with green color or symbol "%" are the groups of this model.

The candidate point is usually in the middle of the group. This point should be selected so that the relationship between the candidate point and neighboring candidate points in the preceding and following rows is not stopped. In Fig. 11b, the candidate pixels of each group are shown by green color or symbol "%".

If the number of neighboring pixels identified for the first row of the group were more than the font size, the following rows should be also investigated in order to consider the pixels of the following rows with the corresponding position in case they had membership conditions. The number of rows in a group would determine the height of the group. If the height or at least the depth of a group were less than twice of font size, then the whole pixels of the first row of the group would be considered as the candidate point. After determining the candidate pixels of the group, their intensity would be equal to zero. Moreover, the intensity of other pixels of the group would be equal to 255.

Now, an example will clarify the following points. For the third row in Fig. 11a, the diagram of intensity changes (Fig. 12c) was drawn. As can be seen, 9 neighboring pixels had the intensity close to zero; also, these pixels had more than 80 intensity difference from the color of envelope. In this example, the font size was equal to 3 pixels; consequently, the number of neighboring pixels was larger than the font size. In the following row, exactly the same situation was observed in a corresponding position. Considering these two issues, the neighboring pixels in rows 3 and 4 were considered in one group. As a result, a group with the width and height of 9 and 2 pixels were formed, respectively.

In Fig. 11a, the red pixels with the symbol "*" and the pink ones with symbol "+" are the groups of this model. In these groups, the candidate points were selected and the remaining extra pixels were removed; then, Fig. 11b was obtained.

In fact, if the relationship of one group with other neighboring groups is removed due to lack of candidating one part of the pixels of that group, at least one candidate point should be considered for the relationship of that group with the remaining neighboring groups.

For instance, in the example mentioned previously, in Fig. 11a, the pixel in position (4, 5) was not selected as the candidate points; however, the relationship of the two groups was cut by its removal. As a result, it was considered as a candidate point. In Fig. 11b, the pixels with these conditions which were selected as the candidates are shown by yellow color or symbol "#".

Moreover, at first, in this type of group (the left side of the group), one candidate pixel can be removed if it is only connected to its following neighboring pixels (Fig. 13a) and not to the pixels in the same column or its preceding columns (Fig. 13b). In fact, at the end of this type of group (the right side of the group), the opposite conditions can lead to the removal of that pixel. To remove a pixel, the intensity of that pixel is considered equal to 255 or white color. In Fig. 11b, the blue pixels shown by symbol "&" are the samples of these pixels. It should be considered that, at most, one or two pixels from beginning or end of one group can be removed in such a situation.

If the width or height of one group were more than 20 times of the font size, this group would not be probably related to letters; it would be related to the probable lines on the envelope.

If one group were not connected to any other group but some other groups were identified around it (less than 25 pixels), that group might belong to other groups; however, if no other groups were identified around it, that group could be removed.

In Fig. 14 a, b, the stages of segmentation, quality improvement and tenderization are demonstrated.

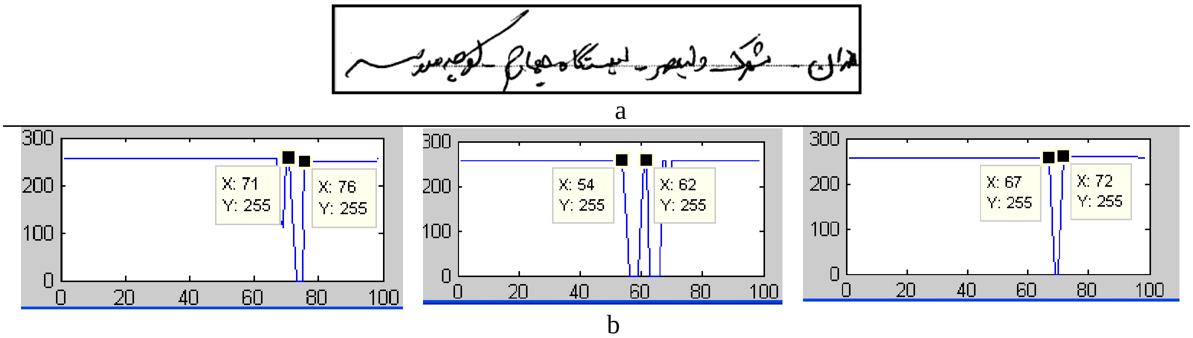


Fig. 10: a) A region of envelope; b) the diagram related to the three columns from image which was randomly selected

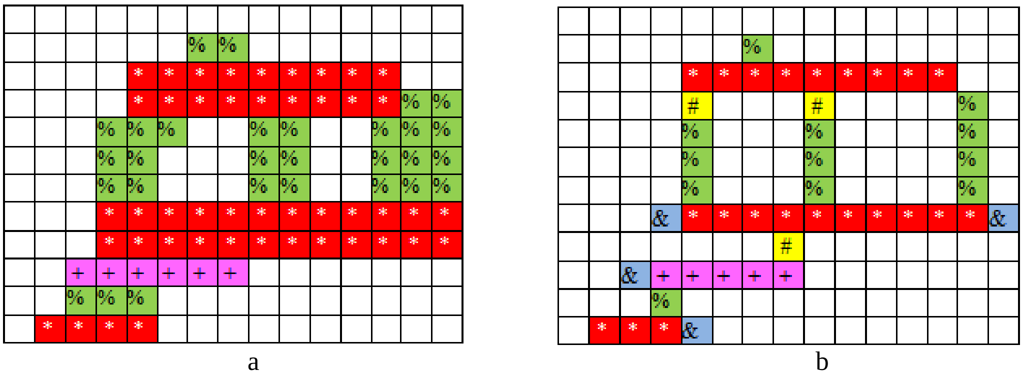


Figure 11. a) The main image of letter (هـ) with the font size of three pixels after applying the grouping algorithm; b) The candidate points selected from each group

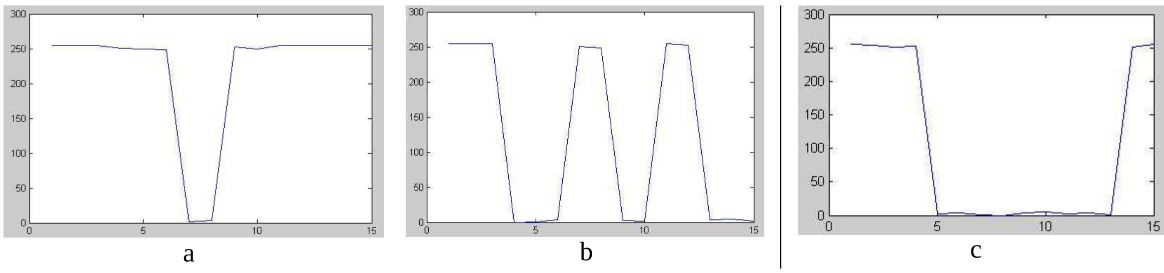


Figure 12. a and b) Diagram of intensity changes from the row belonging to group type a; c) Diagram of intensity changes from the row belonging to group type b

Figure 13. Neighboring pixels of the central candidate pixel

Figure. 14. a) The rotated image; b) the image segmented using the histogram information; c) the image after quality improvement and tenderization

3. RESULTS AND DISCUSSION

The proposed method was tested on 200 envelopes with different dimensions and with the resolution of 300 dpi and bmp format. During scanning, no special limitations were applied to the placing manner of envelopes and they were scanned with different angles. In fact, due to the whiteness of enveloped, the scanner background was considered colorful. In the scanned images, there were official letters (university and Blood Transfusion Organization) in addition to personal letters. In the proposed method, the results presented in Table 1 were generated considering the type of texts written on the envelopes.

Table 1. Accuracy rate of algorithm performance in standardization and quality improvement

Status of text inside the box	Percentage of rotation accuracy	Letter quality improvement and tenderization percentage
Typed texts	100	95
Handwritten texts	99.5	91

The accuracy of the algorithm performance for rotation, quality improvement and tenderization is given in Table 1. Rotation was correctly done in all the envelop samples when the selected box of the envelope included typed letters; therefore, the accuracy level of its rotation was equal to 100%; however, when the selected box included handwritten letters, the accuracy percentage was obtained as 99.5% due to the error in the correct determination of text orientation (correct position of counters) while doing final rotation. Also, since tenderization does not change letter structure and only selects candidate points from among black pixels of a word, therefore, if the proposed method cannot properly tenderize one word, it does not damage the structure of the word and only improves the quality of the word. The proposed method was completely successful in rotating the envelope and improving the quality of its texts. This system has an acceptable function in tenderizing too. Because in Persian language the similar activities have not been done, we are not able to compare the suggested method with other ones.

4. CONCLUSION

In this paper, the important pre-processing stages of the mechanized postal systems were investigated. In the proposed method, envelopes with different forms and tilt levels can be scanned. After applying the rotation algorithm, the envelope was rotated and transformed into a standard format. Then, the segmentation algorithm was applied and the main regions of the envelope were determined; finally, the tenderizing algorithm was applied to prepare the image for entrance into an address extraction system. The rotation algorithm conducted the proper rotation process in two stages with no need for the calculation of tilt angle and consideration of the proper form of words. In this method, only histogram information was used for segmenting. Letter tenderization process was done by selecting a number of candidate points on the words. Novelty in Persian language, less calculation and rotation based on the texts written on the envelope were the advantages of the methods presented in this paper. These algorithms can be used in the mechanized postal systems as a successful and complete pre-processing process. Application of the proposed algorithm in the rotation, segmentation and tenderization processes omits the need for exposing any limitations while scanning the envelopes. As this system is new in Persian language, we won't compare it with other methods.

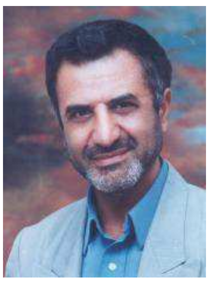
REFERENCES

[1] M. Nahvi, Extract the ZIP code of address printed and manuscripts, a master's degree, Tarbiat Modarres University, Tehran, pp. 20-13, 1997.
[2] R. Azimi, E. Kabir, *"A New Segmentation Technique for Omnifont Farsi Text "*, Pattern Recognition letters 22, pp. 97-104, 2001.
[3] A. Amin, and Masini, "Machine Recognition of Multifont Printed Arabic Texts *", Proc. 8th Int. conf. on Pattern Recognition*. Paris, pp. 392-395, 1989.
[4] V. Marger, "SARAT- A System for the Recognition of Arabic Printed Texts ", *Proc. 11th Int. conf. on Pattern Recognition*. Paris, pp. 561-564, 1992.
[5] N. Koochmshky, M. Fathi, "Real-Time Vehicle Detection and Tracking Algorithm Based on Improved Early Detection of the Line of Symmetry", *An Annual National Conference of Computer Society of Iran*, Amir Kabir University, Iran, 2008.
[6] A. Azimi, E. Kabir, "Designing a Picture Dictionary for the Following Recognition of Printed Words," *Amirkabir Journal*, No. A -59, p. 29-43, 2004.
[7] U. Pal, B.B. Chaudhuri, *"Indian Script Character Recognition: A Survey "*, Pattern Recognition. Indian, pp. 1887-1899, 2004.
[8] M. Soleymani Baghshah, S. Bagheri Shouraki, S. Bagheri Shouraki, "A Novel Fuzzy Classifier Using Fuzzy LVQ to Recognize Online Persian Handwriting", *ICTTA*, pp. 1878 - 1883, 2006.

[9] Eric L'Homer, “Extraction of Strokes in Handwritten Characters”, Pattern Recognition, *France*, vol. 33, pp.1147-1160, 2000.
[10] H. Almuallim and S. Yamaguchi, “A Method of Recognition of Arabic Cursive Handwriting”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 9, pp. 715-722, 1987.
[11] M. Ziaratban, K. Faez, F. Allahveiradi, “Novel Statistical Description for the Structure of Isolated Farsi/Arabic Handwritten Characters”, ICFHR, *Canada*, 2008.
[12] G. Baptista and K. M. Kulkarni, “A High Accuracy Algorithm for Recognition of Handwritten Characters”, *Pattern Recognition*, vol. 21, pp. 287-291, 1988.
[13] J. Hu and H. Yan, “Structural Primitive Extraction and Recognition for Handwritten Character Recognition”, *Pattern Recognition*, pp.493-509, 1998.
[14] S. Saburi, F. Rzazy, A. Teimuri, "The Composition of the Vertebrate Class for Precise Detection of Natural Faces in Static Color Images", *Computer Society of Iran's Annual National Conference*, Amir Kabir University, Iran, 2008.
[15] A. Vinciarelli, “A Survey on off-Line Cursive Word Recognition”, *Pattern Recognition*, Switzerland, pp. 1433-1446, 2002.
[16] W.M. Pan, T.D. Bui, C.Y. Suen , “Isolated Handwritten Farsi Numerals Recognition Using Sparse and Over-Complete Representations”, *10th International Conference on Document Analysis and Recognition*, 2009.
[17] H. Almuallim and S. Yamaguchi, “A Method of Recognition of Arabic Cursive Handwriting”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 9, pp. 715-722, 1987.
[18] M. Ziaratban, K. Faez, F. Allahveiradi, “Novel Statistical Description for the Structure of Isolated Farsi/Arabic Handwritten Characters”, ICFHR, 2008.
[19] G. Baptista and K. M. Kulkarni, “A High Accuracy Algorithm for Recognition of Handwritten Characters”, *Pattern Recognition*, vol. 21, pp.287-291, 1988.
[20] J. Hu and H. Yan, “Structural Primitive Extraction and Recognition for Handwritten Character Recognition”, *Pattern Recognition*, pp. 493-509, 1998.
[21] S. Saburi, F. Rzazy, A. Teimuri, "The Composition of the Vertebrate Class for Precise Detection of Natural Faces in Static Color Images", *Computer Society of Iran's Annual National Conference*, Amir Kabir University, Iran, 2008.
[22] A. Vinciarelli, “A Survey on off-Line Cursive Word Recognition”, *Pattern Recognition*, Switzerland, pp. 1433-1446, 2002.
[23] W.M. Pan, T.D. Bui, C.Y. Suen, “Isolated Handwritten Farsi numerals recognition Using Sparse and Over-Complete Representations”, *10th International Conference on Document Analysis and Recognition*, 2009.

BIOGRAPHY OF AUTHORS

Milad Shirmohammadi was born in Hamedan, Iran. He received his Technician degree in Computer Software from Hamedan Branch, Technical and Vocational University in january 2005, and his BSc. degrees in Computer Engineering from Hamedan Branch, Islamic Azad University in July 2008, and is now graduate student in artificial intelligence Engineering at Azad University of Qazvin.
Emails: mshirmohammadi@qiau.ac.ir, mshirmohammadi@gmail.com



Karim Faez was born in Semnan, Iran. He received his BSc. degree in Electrical Engineering from Tehran Polytechnic University as the first rank in June 1973, and his MSc. and Ph.D. degrees in Computer Science from University of California at Los Angeles (UCLA) in 1977 and 1980 respectively.
Professor Faez was with Iran Telecommunication Research Center (1981-1983) before Joining Amirkabir University of Technology (Tehran Polytechnic) in Iran in March 1983, where he holds the rank of Professor in the Electrical Engineering Department.
He was the founder of the Computer Engineering Department of Amirkabir University in 1989 and he has served as the first chairman during April 1989-Sept. 1992.
Professor Faez was the chairman of planning committee for Computer Engineering and Computer Science of Ministry of Science, Research and Technology (during 1988-1996). His research interests are in Biometrics Recognition and authentication, Pattern Recognition, Image Processing, Neural Networks, Signal Processing, Farsi Handwritten Processing, Earthquake Signal Processing, Fault Tolerance System Design, Computer Networks, and Hardware Design.
Dr. Faez coauthored a book in Logic Circuits published by Amirkabir University Press.
He also coauthored a chapter in the book: *Recent Advances in Simulated Evolution and Learning*, Advances in Natural Computation, Vol. 2, Aug. 2004, World Scientific. He published about 300 articles in the above area. He is a member of IEEE, IEICE, and ACM, a member of

Editorial Committee of Journal of Iranian Association of Electrical and Electronics Engineers, and International Journal of Communication Engineering. Emails: kfaez@aut.ac.ir, kfaez@ieee.org, kfaez@m.ieice.org.

Hassan Khotanlou was born in Hamedan, Iran. He received his BSc. degree in computer Engineering from IUST University in 1995, and his MSc. degrees in artificial intelligence Engineering from Shiraz University in 1997, his Ph.D. degrees in artificial intelligence Engineering from Pierre & Marie Curie University (Paris VI – TELECO) in 2007, and is now Bu-Ali Sina University faculty.
Email: hkh@basu.ac.ir