Slant Correction and Detection for Offline Cursive Handwriting using 2D Affine Transform

Fitrianingsih
Gunadarma University
Departement of Informatics
Depok, Indonesia

Sarifuddin Madenda
Gunadarma University
Doctoral Program in Information Technology
Depok, Indonesia

Suryarini Widodo
Gunadarma University
Doctoral Program in Information Technology
Depok, Indonesia

Rodiah
Gunadarma University
Doctoral Program in Information Technology
Depok, Indonesia

***Abstract*—Cursive handwriting recognition is a machine learning field that tries to capitalize the uniqueness characteristics of character to be expressed in a series of objects of observation, then recognized by a system. There are great amounts of shape and style hand writing variation. Each writer has her/his own styles. Thus, it made difficult to perform a cursive segmentation. One of the difficulties is handwriting segmentation with a slope. In this study a slant handwriting method to detect geometry distorted images of handwriting is being proposed. The shifting processes of continued handwriting are carried out by defining the matrix of the letter through 2D affine transformation function. The image of the letter is shifted based on the slope of the words in the direction of the x or y axis. The shifting process are used a certain scaling factor. The experiments are tested using 40 continued offline handwriting. The data are obtained from IAM Database with image slope of -45 to 45 degrees. The results showed that 2D affine transformation methods are successfully detect and correct the slope of the letter and as a result avoid the over-segmentation of the image candidate. It able to improve the accuracy of continued offline handwriting recognition.**

Keywords—2D Affine Transform; Cursive Handwriting; Offline; Slant

I. INTRODUCTION

Handwriting recognition can be implemented in various type of sectors. For example, in education sector an Optical Character Recognition (OCR) is used to check the answer in the form of multiple choice, whereas the documents in the answer form of essay questions was examined using the manual method. OCR itself is a computer system that can read the letters, both from a printer (printer or typewriter) as well as from the handwriting[1]. A cursive handwriting is used as well by the Graphology experts (graphologist) to determine the character of a person's personality[2].

Characteristics of the recognized offline handwriting are influenced by the nature of the individual authors. It also recognized by the process of data acquisition. Offline cursive handwriting recognition is a major challenge[3]. There are many variations in different handwriting, such as the slope of the letter (slant), spaces between words, the size of the letter, the direction of letter, the style of writing, handwritten with a similarity contours of a few letters (for example: *w* and *v*)[4].The similarity of letters can make the segmentation process becoming difficult. The less precise segmentation process can lead to errors in the process of offline cursive handwriting recognition[5].

There are several studies related to the recognition and segmentation of cursive handwriting. Many algorithms try to improve the handwriting recognition problems mainly for slope letter correction. Pervious researchers using vertical separation letter after skeletonized the letters to compute the number of the words pixels that would be the candidates' separation point. The candidates' word will be evaluated on the basis of distance from each other. Candidates for seven adjacent pixels will be united as a point of separation. The implementation is performed using a word taken using a scanner or digital camera. The analysis of contour is done from right side to the left. The reason is as Latin or Roman has at least one contour on the left side. Converting images to RGB allows better visualization of the separation point [6]. Study on offline handwriting try to improve the image offline slanting. Researchers used the long orthography (the right writing) which are projected as a parameter for comparison with the maximum length horizontally. The calculation of the orthogonal image projection started from the value of the last segmented black pixel image P_{Start} and P_{last}. P_l projections are calculated by P_{Start} - P_{last} and will be corrected by δ. There are 500 handwritten image data with a resolution of 800 x 600 pixels that were used in this study. This study was conducted to make the computation of the previous algorithm more efficient and can handle the type of image with the incline degree until 360^o with different styles of writing whether the handwriting is left to right or vice versa. The results of this study were able to recognize that the un-cursive handwriting has a 70% off accuracy and the accuracy for cursive handwriting is 50% [7].

The study was conducted to restore the slant of the letter. Before segmentation process started, a preprocessing process is required in advance to detect and repair the words that contain the slant. The slant is the slope based on the style of writing or font. In previous research, there are several methods used to perform image enhancement containing slant, one of them using Zernike moments [8]. The authors using two methods, among others: the Radon Transform and Hough Transform, where the results of the research was continued by implementing Zernike Moments. Projected

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 08, August-2016

images are calculated using the Radon Transform from a certain direction. This direction is taken from a different angle by rotating the projection source around the center of the image. The projection of two-dimensional image $f(x, y)$ is calculated at an angle θ with Radon equation $R_\theta(x')$. This transformation detects a line in an angle. By using this transformation, the maximum value of Radon Transform on all lines that form an angle of $50 \le \theta \le 50$ can be calculated. The angle that produces the maximum $R_\theta(x')$ indicates the slope of a word [8]. This study is exploring the detection and correction of slant handwriting using the core-region detection and the estimation of slanting handwriting. This method consist of three steps, namely: (i) core-region detection, (ii) counting non-horizontal orientation of the fragment by performing word extraction and orientation as well as the fragments height to calculate non-horizontal residual and locations that are related to the core-region and (iii) estimate the word as a whole. The slant word is estimated by considering the orientation and height of each fragment. The experiment in the paper used a dataset of IAM database. The data contain nine different documents with 712 words and a few cases (266 words). It can be found randomly on the page http://www.psychpage.com/learning/library/asse using font ITC Bradley to make as dataset B [9].

The experiment in this paper will implement an algorithms that repair the angle of cursive offline handwriting using Affine Transform 2D. The results of the cursive offline handwriting detection and correction give huge contribution to segmentation process and result that generate cursive offline handwriting recognition that precise and accurate.

I. METHODS AND MATERIAL

A. Image Acquistion

Image acquisition process in this research is done offline in which the input image applied for the segmentation process is obtained from a digital scanner (scanner) that produced a .jpg and .png image. Image acquisition process in general can be seen in Figure 1.

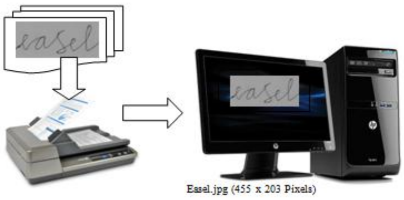


Fig 1. The image acquisition process Handwriting Connect

Image sizes are varied, as seen in the chart Figure 3.1. The research in this paper used the image sized 455 x 203 pixels. Image segmentation process will passed preprocessing with the image cropping. The cropping results are also varied depending on the form of scanned documents. In this study, cursive input image database are used. The experiments used a local dataset of 100 Latin cursive handwriting manifold scanned using a scanner. The image that is employed as training data and test data will not be the same. Secondary data in this study utilized RGB or grayscale image input from IAM database with varied sizes. The images of the word used as test data consist of two to nine small Latin letters. IAM database is used as a comparative test data with local dataset, where the images from database IAM has some characteristics that overlap and the image with a slope that geometrically distorted. This database is needed to see whether the algorithms that were developed runs well on any type of cursive handwriting data. IAM database instance can be seen in Figure 2[10].

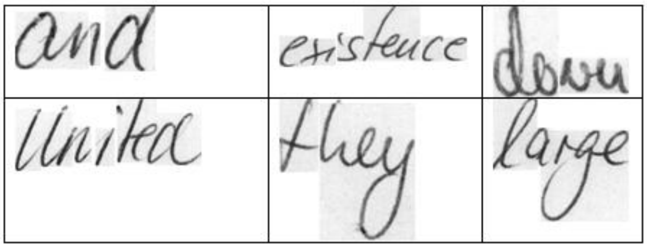


Fig 2. Example of IAM Database[10]

B. Detecting Slant Letter

The slants of the documents often occur in the process of image acquisition. As a result, the written letter in the document in image acquisition process also in a slanted position. Characteristics of cursive handwriting are heavily influenced by the style of individuals, orientation to determine the status and position of the text, whether horizontal, vertical, or form a certain angle. For example alignment is used to format the shape, position, and the position of the text, regardless of its location whether to put on top, middle, or bottom. This geometric distortion can lead to inaccurate results in the following recognition process, thus it necessary to detect and to repair of the slant angle of the image handwriting (image reconstruction).

In general, the process to make the detection of the slanted image handwritten consists of the following steps, among others:

1. Determine the maximum angle of the input image 45°
2. Rotate to transfer the pixel values of the initial writing position with the detected slant from horizontal image of handwriting (corner 0°) by (1) and (2):

$$x_2 = \cos(\theta) \times (x_1 - x_0) - \sin(\theta) \times (y_1 - y_0) + x_0 \qquad (1)$$

$$y_2 = \sin(\theta) \times (x_1 - x_0) + \cos(\theta) \times (y_1 - y_0) + y_0 \qquad (2)$$

where :

x_0, y_0 = Central coordinates of input image

θ = The axis of rotation (Clockwise to horizontal line)

C. Slant Correction Word

Slant letters detection in the image to be processed is done by calculating the projection of the image matrix along the specified direction. The calculation will give the results column of max that are used as a basis for adjusting the slant of the letters in the image. Column of max are used as terms of conditions which will perform 2D affine transformation. Affine transformation is a transformation that maintains the lines and the distance to the four basic operations transformations. The four basic transformations are translation, scaling, rotation and shearing. In cursive slant letter correction, writer uses Shearing operation. Shearing is done by shifting the initial image in the direction of the x or y

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 08, August-2016

axis with a certain scale factor as can be seen in Figure 3 Shearing according to the x-axis:

$$(x',y') \rightarrow (x+ky, y) \qquad \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{3}$$

$$(x',y') \rightarrow (x, kx+y) \qquad \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{4}$$

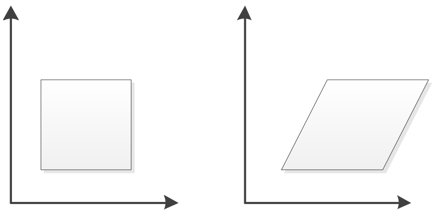


Fig 3. Shearing Process in General

The image shifting can be done by defining the matrix (5). The result of the image slant enhancement can be seen in Figure 4

$$\begin{bmatrix} 1 & sh_y & 0 \\ sh & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

Where :

sh_y = Determining the shear factor along the y-axis

sh_x = Determining the shear factor along the x-axis

(a)

(b)

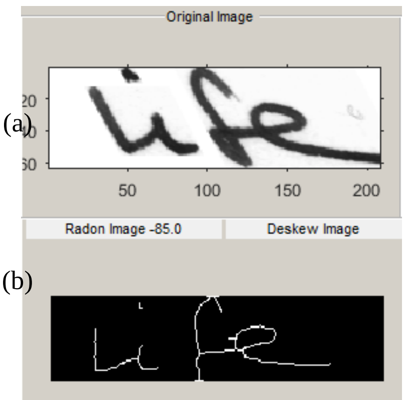


Fig 4 (a) Original Image. (b) Image after slant correction

In general, the slope of the correction process chart said activity diagram can be seen in Figure 3.9. Steps were taken to a slope correction letter in order to get the right image, among others :

1. The input image must be in the format of a white foreground and black background with the following steps:
   1) Determine the value of the threshold T that can be used to distinguish between intensity of objects cursive handwriting and other objects.
   2) Perform binary process to change the value of the grayscale image pixel into the image with a pixel-value logic '1' (white) or '0' (black). Binary process can be performed by using the equation 6. Results of binary image process can be seen in Figure 5.

$$g(x,y) = \begin{bmatrix} 0 \rightarrow & f(x,y) \geq T \\ 1 \rightarrow & f(x,y) \langle T \end{bmatrix} \tag{6}$$

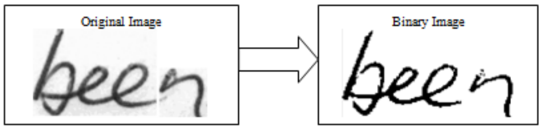


Fig 5. Binary Process of Cursive Handwriting

2. Find max image column using random function *column of max*
3. Create a condition using *column of max* that is obtain from each
4. Make image correction using affine that has shifted threshold on x-axis for each of condition

D. Selecting the Segmentation Candidate

The first stage performed by thinning process. This process is important and carried out before performing the following stages. Thinning removes the outer layer of image handwriting pattern so that only line that contain 1 pixel remain. This line is called a framework (skeleton) of the image. Each iteration of this method consists of two sub successive iteration that applied in contour points of image region, assuming that a pixel with value 1 is the foreground, pixel with value 0 is the background and object pixels has 8-neighbor rules with the following steps :

1. Marked the contour point p to be deleted if conditions 6, 7, 8 and 9 are fulfill by p :

(a) $2 \leq N(p_1) \leq 6;$ (6)

(b) $S(p_1) = 1;$ (7)

(c) $p_2 \cdot p_4 \cdot p_6 = 0;$ (8)

(d) $p_4 \cdot p_6 \cdot p_8 = 0;$ (9)

where $N(p_1)$ is the number of neighbors from p_1 which are not 0:

$$N(p_1) = p_2 + p_3 + \cdots + p_8 + p_9 \tag{10}$$

Table 1
Designations of the nine pixels in a 3 x 3 window [11]

p_9	p_2	p_3
p_8	p_1	p_4
p_7	p_6	p_5

and $S(p_1)$ is the number of transition 0-1 with the order $p_2, p_3, \ldots, p_8, p_9$.

2. Condition (a) and (b) the same as the first stage, while condition (c) and (d) is changed to :

(c) $p_2 \cdot p_4 \cdot p_6 = 0;$ (11)

(d') $p_2 \cdot p_6 \cdot p_8 = 0;$ (12)

In one iteration, the sequence of the algorithm consists of some steps:

1. Run step 1 to mark the border points to be deleted.
2. Remove the points will be marked by change it to 0.
3. Run step 2 in the rest of the border points in step 1 that has not been deleted and that correspond to all the conditions that should be met in step 2 and then marked for deletion.
4. Remove the marked dots by change it to 0.

After the thinning process is completed, one way to find the point of letter separation is by using the average of maximum similarity letter. At first it will be used as the input images (resized image) then histogram with large cell diverse are created. Histogram with cell size (for instance: 48 X48). The process to obtained a candidate letter is based on the letter distance can be seen in Figure 6

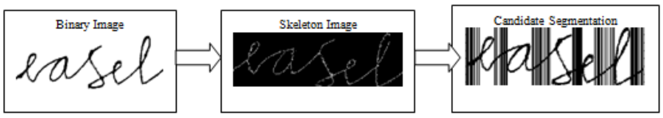


Fig 6. Candidate Selection Process

E. Handwriting Recognition Process

All the candidates of separation point will formed a graph with the similarity weight of the letter. This graph is used to locate the letter separation points. The weighting of the graph utilize the similarity of the letter that was classified with Support Vector Machine (SVM). SVM will produce two results, *i.e.*, positive (+1) or negative (-1) and the similarity probability between 0 and 1. A positive value will be directly used as the similarity of letters, while a negative value will apply 1 - the probability of a letter similarity. Feature letter to be classified by SVM obtained (extracted) using the Histogram of Oriented Gradients (HOG) by compressing the information in the image area where the image validations become cell (worth 1 pixel).

cellSize = [4 4];
hogFeatureSize = 1296;

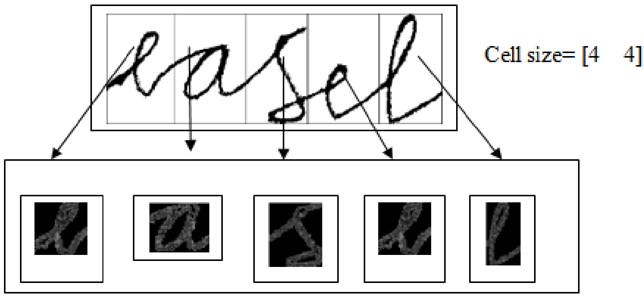


Fig 7. General Schema of Cursive Handwriting Extraction

HOG feature vector obtained by finding the orientation gradient in the image of handwritten continued. The image is divided into several areas of overlap or cell. As can be seen in Figure 7, cell division is done on the size of 4x4 pixels in an image. Each cell is calculated and inserted into the gradient orientation histogram. The entire histogram of the entire cell will be combined into a feature vector HOG. In each histogram consists of a 2x2 cell block or 8 x 8 pixels as shown in Fig 8.

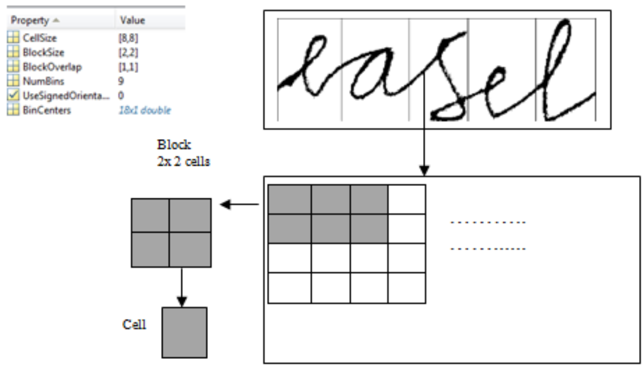
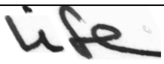
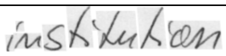
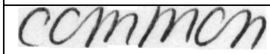
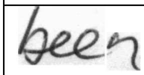
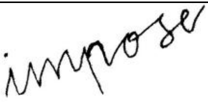
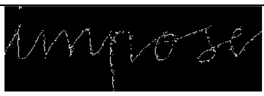
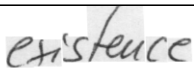
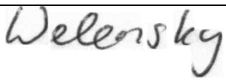


Fig 8. Visualization Features Properties with HOG

II. RESULT AND DISCUSSION

The experiment in this study utilized the letters with varied slant angles (source database IAM). Affine 2D transformation was performed to detect and to repair the slant's angle of offline cursive handwriting. Table I shows some example or slant improvement. Detection and improvement of cursive handwriting is a very important phase to avoid over segmentation of cursive offline handwriting recognition.

Table I
The Result of the Application[10]

Original Image	**Slant Correction**
	
	
	
	
	
	
	
	

The experiment implemented in this study use 40 images with different type and slant [10]. The slant of the image ranged between -45 to 45 degrees. The attempts are performed to obtain the validation accuracy of the separation points of the cursive words. For example, 8 images sample from 40 samples of experiments on Table 1. In the image sample with words ***impose*** (Image with the pixel size 560 x 269), the test are done with the following variables:

- max_angle = 45
- resolution = 0.5
- threshold = length (rounded) x image after *thinning* / 25 = 18
- CellSize = [4 4]

The experiment of slant correction image as well as the result of segmented image validation with the precise separation point can be seen in Figure 9.

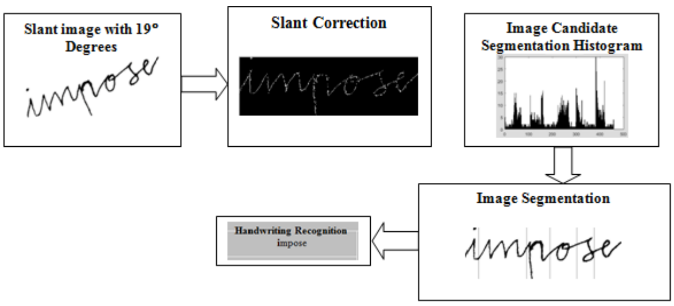


Fig 9. Word Learning Process

III. CONCLUSION

The implementation of affine transformation 2D method on the image of cursive slant offline handwritten which geometrically distorted has successfully corrected. The segmentation of the candidate slant images obtained the accurate letter separation points. The exploration slant angles are obtained using affine transformation 2D method. This method can rotate the image's angle from -45 to 45 degrees and then select the best angle. The angles that consider the best is the one that has histogram with the highest peak. The candidate of letter separation point is the column image that has a number of pixel less than equal to 1 ($\leq$ 1). The histograms are made from number of vertical pixel in each image column. The column with the small pixel values might be the candidate for a separation point. The result of low image pixel shows precise image recognition. Over segmentation typically occur due to the geometric distortion of the image. It also appears on some overlap letters. As a result, there are errors on separation and letter recognition process.

ACKNOWLEDGMENT

The authors would like to thank to Gunadarma University and Doctoral Program in Information Technology of Gunadarma University

REFERENCES

[1] Aparna.A, Prof I Muthumani. *Optical Character Recognition for Handwritten Cursive English Characters*. International Journal of Computer Science and Information Technologies. Vol 5(1). 847-848. ISSN : 0975-9646. 2014

[2] Handwritten Cursive English Characters. International Journal of Computer Science and Information Technologies. Vol 5(1). 847-848. ISSN : 0975-9646. 2014

[3] Prof.Seema Kedar,Ms.Vaishnavi, Ms. Shweta Kulkarni. Personality Identification through Handwriting Analysis : A Review. International Journal of Advanced Research in Computer Science and Software engineering. Volume 5, Issue 1. ISSN: 2277 128x. 2015

[4] Park, Jaehwa, Venu Govindaraju, and Sargur N. Srihari. OCR in a hierarchical feature space. Pattern Analysis and Machine Intelligence, IEEE Transactions on 22.4 (2000): 400-407.

[5] Alessandro Winciarelli. A Survey On Off-Line Cursive Word Recognition. Pattern Recognition. Vol 35, No.7. pp 1433-1446. 2002

[6] Plamondon, R and Sargur N. Srihari. 2000. On-line and Off-line Handwriting Recognition: A Comprehensive Survey. IEEE Transaction on Patern Analysis and Machine Intelligence. Vol 22 No. 1.

[7] Neeta Nain Subhash Panwar. Handwritten Text Recognition System Based on Neural Network. Journal of Computer and Information Technology Vol 2, No.2. 2012

[8] Amit Choudhary and Rahul Rishi. Improving The Character Recognition Efficiency Of Feed Forward Bp Neural Network , International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 1. 2011

[9] Ishaan Agrawal, Alok Vashishtha and Rajiv Kumar, Slant Angle Estimation in Handwritten Documents, Int. Journal of Computer Science and Management Studies. Vol. 14, Issue 5, pp. 1-5, 2014.

[10] A. Papandreou, B. Gatos. Slant Estimation and Core-Region Detection for Handwritten Latin Words. Pattern Recognition Letters 35. 16-22. 2014.

[11] http://www.fki.inf.unibe.ch/databases/iam-handwriting-database/

[12] T.Y.Zhang and C.Y.Suen. A Fast Parallel Algorithm for Thinning Digital Patterns. Image Processing and Computer Vision. Volume 27, Number 3.