

Skew Correction Function of OCR: Stroke-Whitespace based Algorithmic Approach

Mohammad Abu Obaida
Dept. of CSE, DUET
Gazipur, Bangladesh.

Tanay Kumar Roy
CSE, KUET
Khulna, Bangladesh.

Md. Abu Horaira
Software Engineer
DataSoft Systems BD. Ltd
Bangladesh.

Md. Jakir Hossain
CSE, DUET
Bangladesh.

ABSTRACT

As the technology evolves, emergence of Optical Character Recognition (OCR) for both printed and handwritten documents of any language is obvious. In the process of developing an OCR for languages like Hindi, Bengali, Marathi that are among top 15, the most spoken language of the world, the task of skew correction still remains a challenging one as fewer research has been carried out in the field. In this paper, we confront this challenge and describe a stroke-whitespace based algorithmic approach that harnesses horizontal projection technique to correct the skewness of writings precisely for these languages. The paper proposes an easier and effective process named as O_J method that corrects the skewness of images for any degree of rotation. In essence, the paper deals with the images that are rotated by 180° , using Stroke-Whitespace distance method.

General Terms

Optical Character Recognition, Handwritten OCR, Skew Correction, Stroke-Whitespace based text image rotation, Pattern Recognition.

Keywords

OCR, Skew Correction, Text rotation.

1. INTRODUCTION

A field of potential in pattern recognition is Offline character recognition. In essence, documents and data imports along with searching are still intense time consuming for languages other than English in spite of having a huge number of users and various forms of uses. With the view to develop a comprehensive OCR on such languages the skewness of the writing would have to be corrected accurately [6,8,9] to the extent that is accepted without complains. Hough transform [1,2,3,4,5,6], Projection Profile[10,11], Fourier transformation [9] methods are widely used for this purpose. In Hough transform [1,2], the points in the Cartesian coordinate system are described as a summation of sinusoidal distribution: $p = x\cos\theta + y\sin\theta$, the skew angle is calculated on the basis that at the skew angle the density of Transform spaces is maximum. Horizontal projection profile [10, 11] is generally a histogram of the number of dark pixels in horizontal scan lines of a text. Fourier Transformation [9] works on the basic principle that skew angle is the one at which concentration of spectrum is biggest for the document. Clustering [12] is another method, the skew angle for all the connected words in the document is found out and a histogram for the determined skew angle is realized. The maximum clustered skew angle in histogram is the skew angle

of the document. In another method[13], the centers of the nearest neighbors of the connected words in the document are vectorised and later linked to ascertain the skew angle. Cross correlation [14] is quite accurate, the lines of text in a document are considered as vertical lines which are spaced with a uniform distance of d between them. The skewed document vertical lines subtend an angle with the horizontal. So pixels in these vertical parallel lines are translated due to skewing.

The skew corrected image is investigated in several steps to accomplish final character output; it is easily deduced the task of skew correction to be vital. There are various procedures to achieve the final recognition such as, feature extraction. Feature extraction methods have been based mainly on three types of features [15,16,17,18]: a) Statistical Derived from Statistical Distribution of Points b) Structural and c) Transformation-based or Moment-based features. A survey on feature extraction methods is discussed in length in [19]. Moreover, other approaches focus on measuring the similarity/dissimilarity between shapes by mapping one character onto another [20,21]. A choice of Feature Extraction techniques that are mentioned above have been applied with great success to both historical and contemporary document recognition. However, there are also methodologies focused on the unique characteristics of the corresponding historical document they process, such as content and writing style [22,23]. There have been quite a number of successes in determination of invariant features and a wide range of classification methods have been extensively researched.

However, the techniques of skew correction mentioned above fail capriciously while rotating documents in case of skewness very close to 180° , because such images seems skew corrected as it sustains the white spaces and the stroke lines flat. On such stalemate situations we provide O_J method, applicable to stroke based languages (Hindi, Bengali, Marathi etc.) that corrects the skewness of original text for any degree of rotation.

2. SKEW CORRECTION O_J APPROACH

2.1 Proposition

Horizontal projection method is the fundamental actuation that backs our proposed O_J theory of skew correction, where projection of the document is considered in every angle thereby possessing a perfect projection angle measurement. The theory presumes that, the picture that is entered in for skew correction, contains no white space (white spaces are cropped out) outside the minimum fit rectangle as in Figure 1(b).

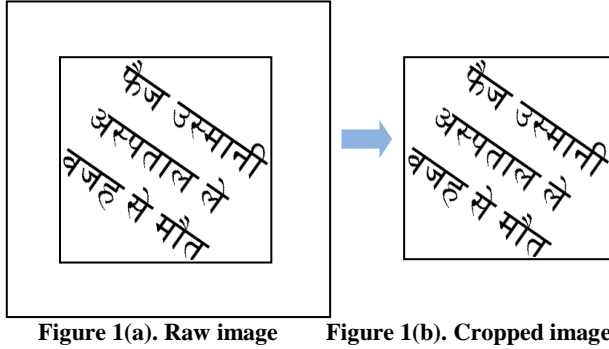


Figure 1(a). Raw image Figure 1(b). Cropped image

A rapid view on the image evident, there is a white line of several pixel widths between text lines (line spacing) that fits the rotation angle. The theory states that finding the angle that counts maximum occurrences of white lines while scanned from top to bottom provides the prospect to correct skewness.

To deal with the problem of 180° rotated images, let us put some light on the stroke-whitespace based 180° skewness correction method. Languages like Hindi, Bengali uses a *stroke zone* (Figure 2(a), Figure2 (b)) or simply a base line to guide the indentation throughout the writing time. We assume each of the letters of these languages is divided in three such regions (Figure 2(a), Figure2 (b)), where each of the characters supports (*Zone above stroke*) < (*Zone below stroke*) in normal cases. We find the distances $d1$ and $d2$ (as depicted in Figure 3) scanning the image from top to bottom (using line spacing thresh-hold). If the distance $d1 > d2$ then we understand the image (Figure 3-R) have to be rotated by 180°.

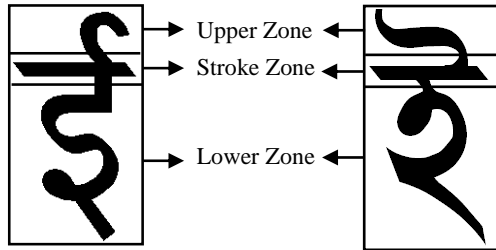


Figure 2(a). A Hindi letter Figure 2 (b). A Bengali letter

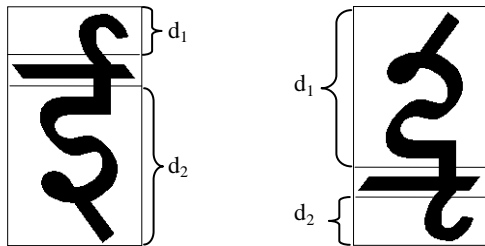


Figure 3 : (L)An all right image{ $d_1 < d_2$ }.
(R) Requires 180° rotation.

2.2 Process Development

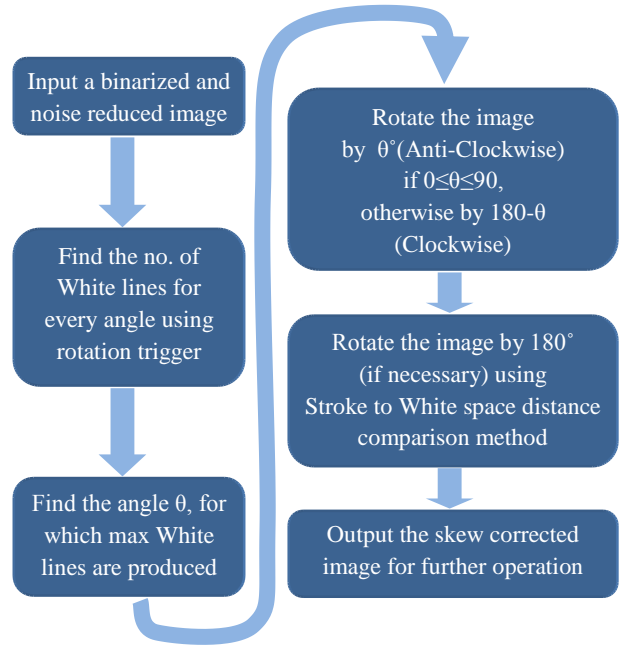


Figure 4. Steps of Skew Correction.

2.3 Formulation of Lines

1) $E: ax+by+c=0$.

Any line goes through $(0,0)$ at angle θ is given by,
 $y = x \cdot \tan\theta$ - - - - - (I)

$\Rightarrow x \cdot \tan\theta - y = 0$

So, $a = \tan\theta, b = -1, c = 0$.

2) Equation of pair of parallel lines of (I) at distance d is given by $E_{p1,p2}: ax+by+K=0$. - - - (II)

So, $\pm d = (K-c)/\sqrt{(a^2 + b^2)}$

$\Rightarrow K = c \pm d\{\sqrt{(a^2 + b^2)}\}$ - - - - - (III)

As the value of parameters a, b, c is known from given equation; value of K can easily be calculated by (III).

Using this value of K , a pair of Straight lines parallel to equation (I) is formulated using equation (II).

These straight lines in turns used to find if a line is White (\pm Error Tolerance) or not.

2.4 Algorithm

SKEW_CORRECTION(image, X_{max} , Y_{max} , precision, result_rotation)

result_rotation:=0, count:=0

FOR angle:=0 to 180° STEP precision LOOP

 TMP=COUNT_WLINE(image, X_{max} , Y_{max} , angle,d)

 IF tmp>Count THEN

 Set Count:=tmp

 Set result_rotation:=angle

SKEW_CORRECTION:= result_rotation

COUNT_WLINE(*image*, X_{max} , Y_{max} , *angle*, *d*)
 (1) $WL:=0$
 (2) Let $E: ax+by+c=0$ is the equation that fits the image for given angle.
 (3) IF **ISWHITE**(*image*, X_{max} , Y_{max} , E , Err_Tol , $Flag$) = **TRUE** THEN $WL:=WL+1$
 (4) Now find the equation of parallel lines of E , $ax+by+K=0$ at d distance and assign it to E .
 (5) IF $Flag=TRUE$ goto Step 3.

ISWHITE(*image*, X_{max} , Y_{max} , E , Err_Tol , $Flag$)
 (1) Set $Flag:=TRUE$, $BCount:=WCount:=0$.
 (2) From E we get $y=(-k-ax)/b$ - - - - - (E)
 (3) FOR $x:=0$ to X_{max} STEP 1
 Calculate y from E
 IF $y \geq 0$ AND $y \leq Y_{max}$ THEN
 IF $image_{x,y}=1$ then
 $WCount:=WCount+1$.
 ELSE
 $BCount:=BCount+1$
 (4) IF $Err_Tol \geq BCount/(BCount+WCount)$
 Return **TRUE**
 (5) Set $y_0:=(-k-ax_0)/b$, $y_m:=(-k-aX_{max})/b$
 IF $y_0 \geq Y_{max}$ AND $y_m \leq 0$ THEN
 $Flag=FALSE$
 (6) Return **FALSE**

STROKE_BASED_180REVERSE_ALGORITHM(*image*, X_{max} , Y_{max})
 (1) FOR $y:=0$ to Y_{max} REPEAT Steps 2 to 4
 (2) Set $BPCount:=WPCount:=0$
 (3) FOR $x:=0$ to X_{max}
 IF $image_{x,y}=1$ THEN
 $BPCount:=BPCount+1$
 ELSE
 $WPCount:=WPCount+1$
 (4) IF $Err_Tol \geq BPCount/(BPCount+WPCount)$ THEN
 $tmp=y$, $PixCount[k]=BPCount$, $k:=k+1$; goto Step 5.
 (5) FOR $y:=tmp+1$ to Y_{max}
 $BPCount:=WPCount:=0$
 FOR $x:=0$ to X_{max} Step 1
 IF $image_{x,y}=1$ THEN
 $BPCount:=BPCount+1$
 ELSE

$WPCount:=WPCount+1$
 IF $Err_Tol \geq BPCount/(BPCount+WPCount)$ then
 $PixCount[k]=BPCount$; $k:=k+1$;
 ELSE
 $wlc:=wlc+1$
 IF $wlc \geq Expected_Line_Spacing$ THEN
 goto Step 5.
 (5) FOR $I=1$ to $K-1$
 Set $Diff[I]:=ABS(PixCount[I]-PixCount[i+1])$
 (6) Let I and J be the index of maximum two values in the array $Diff$.
 IF $I > K-J$ THEN
 (Rotate the image by 180)

2.5 Complexity

Worst case running time of **ISWHITE**() is $O(X_{max})$. **ISWHITE**() may be called maximum Y_{max} times in **COUNT_WLINE**(), when the image contains only white lines. Thus its running time is given by $O(X_{max} \times Y_{max})$. Now in **SKEW_CORRECTION**(), the only FOR loop runs in constant time as both 180 and precision are constant. So its running time is $O(X_{max} \times Y_{max})$.

Hence the overall complexity of the algorithm in worst case is $O(X_{max} \times Y_{max})$.

3. EXPERIMENTAL OUTCOMES

3.1 Performance Analysis:

The algorithm put into action in MATHLAB possesses significant improvement correcting skewness than Hough transform, Projection Profile or Center of Gravity method does. This is because the projection is considered in several horizontal angles. Figure 5 demonstrates two skewed Hindi printed text images, whereas figure 6 illustrates the corrected forms of these images.

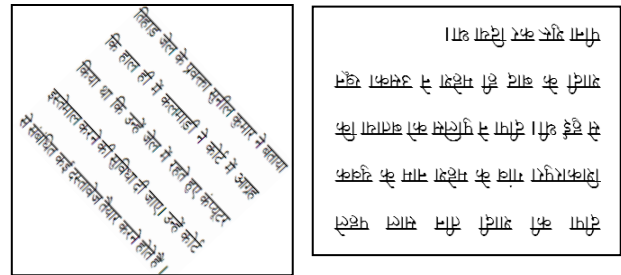


Figure 5: (From Left to Right) subsequently 47°, 180° skewed Hindi printed writings.

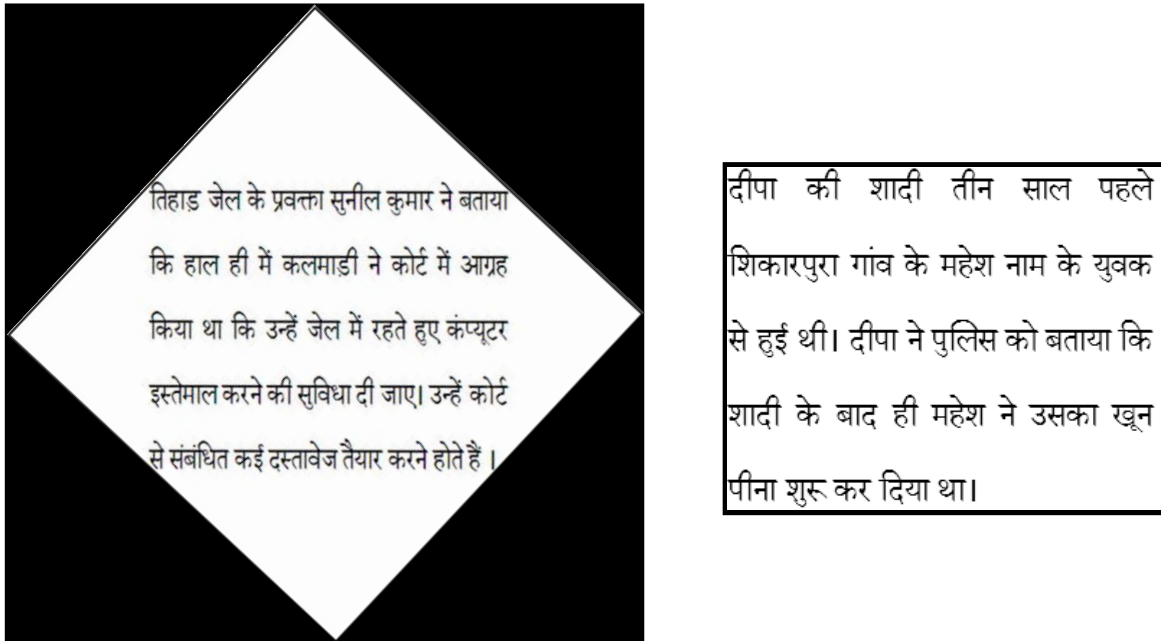


Figure 6:(From Left to Right) Subsequently 47°, 180° skew Corrected Hindi printed writing illustrations.

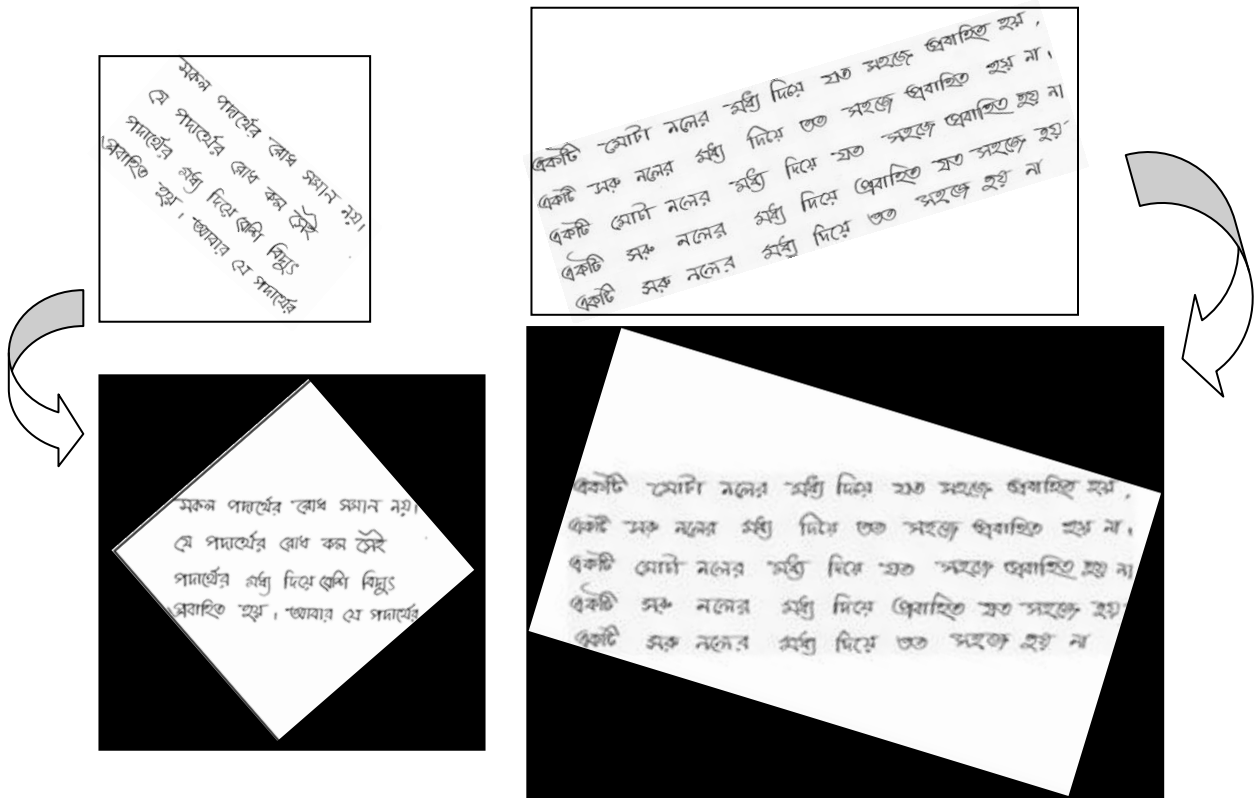


Figure 7: (Top) skewed Bengali Handwritings | (Bottom) Skew Corrected.

To evaluate our proposed O_j method, we compare both scanned /camera captured, handwritten/printed, Hindi/Bengali writings. Some of the experimental images are listed above as Fig 5 & Fig 7(Left). The approach with high precision can detect skewness with any angle (360°). The experimental result reinforce that the proposed method is effective if compared to well known existing methods. Table 1 & Table 2 evident the claim that proposed O_j method, it exhibits competence compared with other popular methods of image skewness correction.

Table 1. Performance Comparison (Skewness Vs Rotation angle, Calculating Time) of proposed O_j method with other methods for Hindi writings (800 x 600 pixels).

Writing Skewness	Hough Transform[2][4]	Center of Gravity[7]	O_j Method
30°	$29^\circ(1.51s)$	NA	$30.2^\circ(5.3s)$
75° (H*)	$73^\circ(1.94s)$	NA	$73.7^\circ(6.5s)$
90°	$90(1.53s)$	NA	$90^\circ(5.2s)$
135° (H*)	$44^\circ(2.7s)$ / NA	$134^\circ, (0.49s)$	$136^\circ(4.7s)$
180°	$0^\circ(1.09s)$ / NA	NA	$180^\circ(7.8s)$
250° (H*)	$71^\circ(2.65s)$ / NA	NA	$249^\circ(8.9s)$

H*- Refers to Hand writings.

NA- Not Applicable / 180° Rotation problem.

Table 2. Performance Comparison (Skewness Vs Rotation angle, Calculating Time) of proposed O_j method with Hough Transform for Bengali Writings (800 x 600 pixels).

Writing Skewness	Hough Transform	O_j Method
30°	$30^\circ(1.76s)$	$30.2^\circ(6.3s)$
75° (H*)	$74^\circ(1.53s)$	$74.2^\circ(6.7s)$
90°	$90^\circ(1.56s)$	$90^\circ(7.2s)$
135° (H*)	$45^\circ(1.86s)$ / NA	$136^\circ(6.4s)$
180°	$0^\circ(1.14s)$ / NA	$180^\circ(8.8s)$
270°	$90^\circ(1.37s)$ / NA	$270^\circ(9.4s)$

H*- Refers to Hand writings.

NA- 180° Rotation problem.

4. CONCLUSION

To sum up briefly the paper exposes a very successful, however, a bit time consuming method for the validation of correcting skewness for several worlds' most used languages. In our effort we have corrected skewness using stroke-whitespace zoning mechanism. Experiments have been carried out for plain documents (without graphics) and also for noisy images. Experimental results advocate that it works better for all types of documents with or without graphics. Thus we can conclude that proposed O_j method is a successful one, novel and precise for skewed documents. However in case of noisy images performance of proposed method degrades as the noise concentration increases. On the contrary to deficiency in time consumption it is believed that the algorithm out performs; besides further improvements will be suggested shortly to reduce the execution time as low as one/two seconds.

5. REFERENCES

- [1] Richard O. Duda, Peter E. Hart, April 1971. "Use of the Hough Transform to detect lines and curves in pictures". *Technical Note 36, AI Center*.
- [2] Srihari, S.N. and V. Govindaraju, 1989. "Analysis of textual images using the Hough transforms". *Machine Vision Applications*, 2: 141-153. DOI: 10.1007/BF01212455.
- [3] Le, D.S., G.R. Thoma and H. Wechsler, 1994. Automatic page orientation and skew angle detection for binary document images. *Pattern Recognition*, 27: 1325-1344.
- [4] Pal, U. and B.B. Chaudhuri, 1996. An improved document skew angle estimation technique. *Pattern Recognition Lett.*, 17: 899-904. DOI: 10.1016/0167-8655(96)00042-6
- [5] Yu, B. and A.K. Jain, 1996. A robust and fast skew detection algorithm for generic documents. *Pattern Recogn.*, 29: 1599-1629. DOI: 10.1016/0031-3203(96)00020-9
- [6] Tian Jipeng, G.Hemantha Kumar, H.K. Chethan : "Skew correction for Chinese character using Hough transform". *International Journal of Advanced Computer Science and Applications (IJACSA), Special Issue on Image Processing and Analysis*.
- [7] Atallah Mahmoud, Al-Shatnawi and Khairuddin Omar: Skew Detection and Correction Technique for Arabic Document Images Based on Centre of Gravity. *Journal of Computer Science* 5 (5): 363-368, 2009, ISSN 1549-3636
- [8] A.F.R. Rahman and M. Kaykobad, A Complete Bengali OCR : A Novel Hybrid Approach to Handwritten Bengali Character Recognition, *Journal of Computing and Information Technology*, Vol. 6(4), 1998, pp. 395-413.
- [9] Omar, K., A. Ramli, R. Mahmod and M. Sulaiman, 2002. Skew detection and correction of jawi images using gradient direction. *Journal of Tech.*, 37: 117-126.
- [10] Hou, H.S., 1983. Digital Document Processing. *Wiley New York*, ISBN: 0471862479.
- [11] Akiyama, T. and N. Hagita, 1990. Automated entry system for printed documents. *Pattern Recognition*, 23: 1141-1158. DOI: 10.1016/0031-3203(90)90112-X
- [12] Hashizume, A., P.S. Yeh and A. Cosenfeld, 1986. A method of detecting the orientation of aligned components. *Pattern Recognition Letters*, 4: 125-132.
- [13] O'Gorman, L., 1993. The document spectrum for page layout analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11: 1162-1173. DOI: 10.1109/34.244677.
- [14] Yan, H., 1993. Skew correction of document images using interline cross correlation. *Computer Vision Graph. Image Process.*, 55: 538-543. DOI: 10.1006/cvip.1993.1041.
- [15] Luiz S. Oliveira, F. Bortolozzi, C.Y.Suen, "Automatic Recognition of Handwritten Numerical Strings: A Recognition and Verification Strategy", *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 2001, Vol. 24, No. 11, pp. 1448-1456.

- [16] K. M. Mohiuddin and J. Mao, "A Comprehensive Study of Different Classifiers for Hand-printed Character Recognition", *Pattern Recognition, Practice IV, 1994*, pp. 437- 448.
- [17] L. A. Koerich, "Unconstrained Handwritten Character Recognition Using Different Classification Strategies", *International Workshop on Artificial Neural Networks in Pattern Recognition (ANNPR), 2003*.
- [18] N. Arica and F. Yarman-Vural, "An Overview of Character Recognition Focused on Off-line Handwriting", *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2001, 31(2)*, pp. 216 - 233.
- [19] O. D. Trier, A. K. Jain, T. Taxt, "Features Extraction Methods for Character Recognition – A Survey ", *Pattern Recognition, 1996, Vol.29, No.4*, pp. 641-662.
- [20] S. Belongie, J. Malik, J. Puzicha, "Shape Matching and Object Recognition Using Shape Contexts", *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24, No. 4*, pp. 509-522, 2002.
- [21] Anil K. Jain, Douglu Zongker, "Representation and Recognition of Handwritten Digits using Deformable Templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, Vol. 19, No. 12*, pp. 1386-1391.
- [22] V.G.Gezerlis and S.Theodoridis, "Optical Character Recognition for the Orthodox Hellenic Byzantine music notation", *Pattern Recognition, 2002, Vol.35*, pp. 895 – 914.
- [23] K. Ntzios, B. Gatos, I. Pratikakis, T. Konidakis and S.J. Perantonis, "An Old Greek Handwritten OCR System based on an Efficient Segmentation-free Approach", *International Journal on Document Analysis and Recognition (IJ DAR), Special Issue on Historical Documents, 2007, vVol. 9, No. 2-4*, pp. 179-192.

6. AUTHORS PROFILE

Mohammad Abu Obaida obtained his Bachelor of Science in Engineering degree from Department of Computer Science and Engineering (CSE), Dhaka University of Engineering & Technology (DUET), Gazipur-1700, Bangladesh. At present he is performing wide-ranging research on networks, cryptography and pattern recognition. His key research interests include Cryptography, Network and Web Security, Wireless networks, Software Architecture, Machine Vision, Artificial Intelligence, Protocol analysis and design.