A Robust Approach To The Recognition Of Text Based Bangla Road Sign

Bithi Banik
Department of Computer Science & Engineering
University of Chittagong
Chittagong, Bangladesh
Email: bithi_cse@yahoo.com

Fahim Irfan Alam
Department of Computer Science & Engineering
University of Chittagong
Chittagong, Bangladesh
Email: fahim@cu.ac.bd

***Abstract*—Road sign recognition is considered to be one of the most fascinating and interesting field of research in intelligent vehicle and machine learning. Road signs are typically placed either by the roadside or above roads. They provide important information in order to make driving safer and easier. This paper proposes an algorithm that recognizes Bangla road sign with a better percentage. The algorithm starts with capture image from real video scene, text detection from images, character segmentation and recognition of characters through shape matrix. The constructed feature vectors for each individual Bangla road sign are learned into a neural network which later classifies new instance of Bangla road sign. The promising preliminary experimental results indicate a positive potential of our algorithm.**

Keywords: Bangla road sign recognition, character recognition, text detection, bag of words, character segmentation.

I. Introduction

Road signs are installed to guide, warn, and regulate traffic. They supply information to help drivers operate their cars in such a way as to ensure traffic safety. In the real world, when they get tired, drivers may not always notice road signs. At night, drivers are easily affected by headlights of on coming vehicles and may miss road signs. In bad weather, road signs are harder to recognize quickly and correctly. These situations may lead to traffic accidents and serious injuries. A vision-based road sign detection and recognition system is thus desirable to catch the attention of a driver to avoid traffic hazards.

The main objective of an Automatic Road Signs Recognition System is to recognize one or more road signs from complex digital images coming from video camera, mounted on a vehicle moving along the roads and/or the highways. This is a difficult task, considering the complexity of outdoor scenes and the variation of lighting and shadowing conditions. Lighting conditions is a very difficult problem to constrain and regulate.

Recently, many techniques have been developed to detect road signs [1], [2].Most of them are shape based recognition technique.Few [3] concentrates on text attached with road sign.Most deal with single images with simple backgrounds [4], [5]. Lalonde and Li [6] reported a color-indexing approach to identifying road signs. but the computation time increases greatly in complex traffic scenes.Aoyagi and Asakura [7] used genetic algorithms to detection road signs from gray-level video imagery. Unfortunately, because of the limitation of crossover and mutation operators, optimal solutions are not guaranteed.This paper is aimed at developing an inventory system [8], [9] to obtain a complete catalog of all the traffic signs on a given road and gather information about their state. There is very few research work on Bangla road sign recognition.Moushumi and Rahman [10] used sobel edge detection technique in text detection of Bangla road sign.Unfortunately this technique is not very efficient in text detection.But,we have to concentrate more on text detection in road sign recognition system.Besides,there was not any specific methodology for character recognition stage in their [10] algorithm.

In the article, a study on recognition of Bangla Road Sign is presented. Our work is based on a large database of real-life road sign samples. We proposes an algorithm that concentrates on text detection technique.In the recognition system,text detection technique plays an important role in accuracy.Then we segmented lines,words and characters in an efficient way.The feature vector of characters was based on shape matrix which later fed into neural network. The learned neural network later is used to recognize future instances of test data. The experimental result shows a positive feedback of our algorithm.This paper proposes a robust framework in the sense that the feature representation for individual characters are very well descriptive and invariant to transformation changes such as scaling, rotation and translation. One cannot expect to operate and execute an algorithm in an environment free from variations and changes. The algorithm must cope with errors during execution and continue to operate despite abnormalities in inputs. Road sign itself is susceptible to noises and damages. In critical situations, it might be difficult to correctly recognize a particular character which in turn could complicate the entire process. Our aim in the paper was to devise a feature description which can appropriately retrieve the actual representation of characters in worst scenario. The proposed algorithm describes the shape features in a consistent way to differentiate among similar characters and retrieve the actual shape. This robust representation is expected to work with difficult data set in a noisy environment.

II. Bangla Road Sign

Bangla road signs have two recognizing features: a visual template of fixed shape and the text box that contains bangla sentence.Actually,the bangla sentence is nothing but the bangla representation of road or traffic regulation.

There are different types of road signs in Bangladesh.They are Warning Sign, Information Signs, Route Signs, Supplementary Plates, Obligation Signs etc.Some are shown in figure:1

III. Proposed Framework

Our proposed method consists of two major steps: training and testing.Training phase is very important in recognition task.Training step consists of following stages:preprocessing,Text detection from bangla road sign template,line-word-character segmentation,recognition of individual character through shape matrix feature and atlast confirmation of whole bag of words as bangla road sign.Later in the testing phase,learned network is used to recognize future instances of test data.

A. Preprocessing

In the preprocessing step,images from video sequence containing bangla road sign were collected.It is very common for images to contain noises which results in loss of signicant information from the contents and as a result, the model fails to represent every details of an image.To handle this problem, we used a method described in [11] which is called the adaptive weighted median filter (AWMF).This method introduces the weight coefficients to the well known median filter and by proper adjustments of the weight coefficients, the filter method smoothes each point of the image and as a result, the noises are suppressed with edges and other significant features preserved. As this technique can detect the subtle gray scale variation, it is indeed very a very appropriate method for our recognition problem. Then we normalized each image into 400×300 pixels as to make all images of same size in recognition stage.

B. Text detection

In road sign recognition technique text detection should be performed more accurately.Because,high accuracy of text detection leads to highly accurate road sign recognition.Videos often have low resolution and complex backgrounds and text can be of different sizes, styles and alignments. In addition, scene text is usually affected by lighting conditions and distortions.

There are many Text detection techniques.Some are based on connected components,some are edge based and some are based on texture.The first approach has some disadvantages.It does not work well for all video images. Mousumi et al [10] used the second approach that requires text to have a reasonably high contrast to the background in order to detect the edges.So,these methods often encounter problems with complex backgrounds and produce many false positives. At last, the third approach considers text as a special texture and require extensive training and are computationally expensive for large databases. So,we used an efficient text detection method [12] based on the Laplacian operator.

Text detection method consists of three steps: text detection, boundary refinement and false positive elimination.The detail description is given below

Fig. 1: Sample of Bangla Road Sign

1	1	1
1	-8	1
1	1	1

Fig. 2: The 3×3 Laplacian mask

*1) **Text detection**:* At first,the input image is converted to grayscale and filtered by a 3×3 Laplacian mask to detect the discontinuities in four directions: horizontal,vertical, up-left and up-right (Figure 2).

The Laplacian-filtered image contains both positive and negative values.The transitions between these values correspond to the transitions between text and background. In order to capture the relationship between positive and negative values, we use the maximum gradient difference (MGD), defined as the difference between the maximum and minimum values within a local $1 \times N$ window. The MGD value at pixel(i,j) is computed from the Laplacian-filtered image f as follows.

$$MGD(i,j) = max(f(i,j-t)) - min(f(i,j-t)) \quad (1)$$

where $t \in [-\frac{N-1}{2}, \frac{N-1}{2}]$

Then,we normalize the MGD map to the range [0, 1] and use K-means to classify all the pixels into two clusters, text and non-text, based on the Euclidean distance between MGD values.At last, each connected component in the text cluster we got is a candidate text region (see figure 3(a-d)).

*2) **Boundary refinement**:* In this section,we compute the binary sobel edge map SM of the input image (only for text regions).

The horizontal projection profile is defined as follows.

$$HP(i) = \sum_j SM(i,j) \quad (2)$$

If HP(i) is greater than a certain threshold, row i is part of a text line; otherwise, it is part of the gap between different text lines. From this rule, we can determine the top row i_1 and bottom row i_2 of each text line. The vertical projection profile

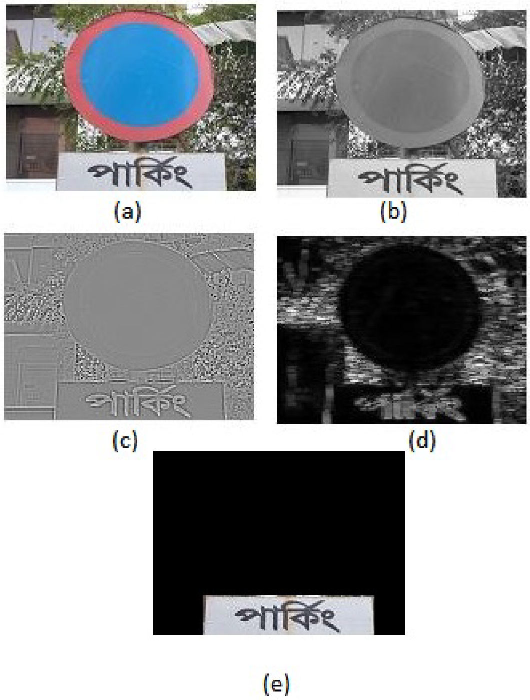


Fig. 3: in the figure,(a)real image,(b)gray image,(c)the laplacian filtered image ,(d) MGD map,(e)text block

is then defined as follows.

$$VP(j) = \sum_{i=i_1}^{i_2} SM(i,j) \tag{3}$$

Similarly, if VP(j) is greater than a certain threshold,column j is part of a text line; otherwise, it is part of the gap between different words.By applying this step recursively, we can determine the accurate boundary of each text block.At last, each detected block is a candidate text block.

*3) **False positive elimination**:* We eliminate false positives based on geometrical properties. Let W, H, A_r, A and E_a be the width, height, aspect ratio, area and edge area of text block B.

$$A_r = \frac{W}{H} \tag{4}$$

$$A = W \times H \tag{5}$$

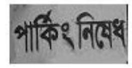
$$E_a = \sum_{i,j \in B} SM(i,j) \tag{6}$$

If $A_r < T1$ or $\frac{E_a}{A} < T2$, the candidate text block is considered as a false positive; otherwise, it is accepted as a text block (see Fig.3-(e)).

C. Line-word-character segmentation

Our text segmentation part is structured into five subsections.They are The Absolute Gradient features,line segmentation,word extraction,detection of the Matra and separation of three different zones of Bangla text line, segmentation of the words into characters.The detail description is given below.

TABLE I: image after AGF and k-means clustering

original grayscale image	normalized gradient image	k-means cluster
		

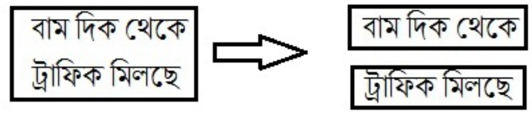


Fig. 4: line segmentation

Fig. 5: division and binarization of word image

*1) **Absolute gradient feature**:* To differentiate low contrast text pixels from background,we are proposing the absolute gradient feature method. For a given text line image,horizontal and vertical gradient information are calculated by using [-1 1]mask to each pixel of the original gray-scale image(see table I).In the terms of mathematics,

$$AGF(I) = g_x(x,y) + g_y(x,y) \tag{7}$$

where, $g_x(x,y) = |I(x+1,y) - I(x,y)|$ and $g_y(x,y) = |I(x,y+1) - I(x,y)|$

After getting this absolute gradient image we use K-Means clustering to binarize the image(see table I).

*2) **Line segmentation** :* K-Means clustering with two classes, gives a text cluster as a binarized edge map for the whole text template. A component analysis is then driven out on the Text cluster to get the average area for the components. Now those components which have area less than two percent of the average area are ignored as noise.Few Bangla road sign has text more than one line.So,first we are to segment those lines(see figure 4). The global horizontal projection method computes sum of all black pixels on every row and constructs corresponding histogram. Based on the peak/valley points of the histogram individual lines are segmented [13].

*3) **Word segmentation** :* After the line segmentation task, we are projecting it vertically to get some continuous zero-frequency columns. Those continuous intervals are marked as word gaps.Based on the Word Gap information, the original grayscale Text line image is now divided into some word images (see figure 5).Each word grayscale image is now binarized using Otsus Thresholding technique [14].

*4) **Matra detection and zone separation** :* Since Bangla text lines can be partitioned into three zones (see Fig.6), it is convenient to distinguish these zones. Character recognition becomes easier if the zones are distinguished because the lower

Fig. 6: three zones of bangla text

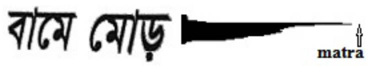


Fig. 7: Matra detection

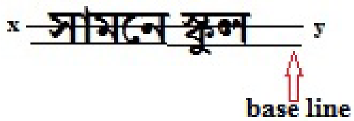


Fig. 8: lower boundary detection

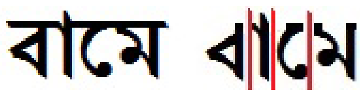


Fig. 9: character segmentation after deletion of matra

zone contains only modifiers, while the upper zone contains modifiers and portions of some basic characters.

Firstly a horizontal projection is applied on the binarized image to find the span of its highest peak, which is referred as the Matra of the text line (shown in Fig.7). From Fig.6 it is clear that the upper zone can be separated from the middle zone of a text line by the Matra, which is already detected. To separate the middle and lower zone consider an imaginary straight line xy that horizontally partitions the text line region into equal halves. Consider only the connected components below xy and connected to xy. The horizontal line, which passes through maximum number of the lowermost pixels of each of these components, is the separator line (Base line) between the middle and lower zones. See Fig.8 for an illustration.

*5) **Character segmentation** :* To segment individual characters of a word we consider only the middle zone. The basic approach here is to ignore the Matra so that the characters get topologically disconnected. To find the demarcation line of the characters a linear scanning in the vertical direction from the Base line is initiated. If during a scan, one can reach the Matra without touching any black pixel then this scan marks a boundary between two characters. For some kerned characters a piecewise linear scanning method [15]has been invoked (see Fig.9).

D. Feature Extraction

After the character segmentation,we get the individual Bangla characters.Moushumi et al [10] used threshold value to extract feature of the characters.But,this method will not be appropriate for variation of angle and rotation of character images.So,to solve this kind of critical cases,we used shape based feature extraction method [16]. For simplification,we

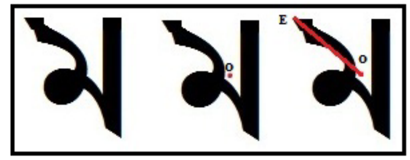


Fig. 10: (a) a character image; (b)center of gravity ;(c)maximum radious

normalized each Bangla character into 32 by 32 pixels. For each binarized and normalized character image I , an n×m shape matrix is extracted. The matrix is formed by polar quantization of the image (see Figure 10(a)). The formation of this matrix is initiated by determining the center of gravity and the maximum radius of character. The center of gravity of I is defined by its pixel coordinates O = (O_x,O_y),

$$O_x = round\frac{1}{N}\sum_{i=1}^{N} x_i \tag{8}$$

$$O_y = round\frac{1}{N}\sum_{i=1}^{N} y_i \tag{9}$$

where N is the number of pixels of I (i.e., cardinality of I), round returns the nearest integer value, and (x_i, y_i)∈ I(see Figure 10(b)). The maximum radius R of I(shown in Figure 10(c)) is defined as

$$R = \max_{(e_x,e_y)} \sqrt{((e_x - O_x)^2 + (e_y - O_y)^2)} \tag{10}$$

Where E=(e_x ,e_y) is a point on the edge of I. Next the matrix representation of I's shape is formed. The matrix is a sampling of I (done by polar quantization of I) where the value of an entry indicates whether or not a pixel is a foreground or background. Thus, the larger the dimensions of the matrix, the more descriptive it is of I's shape. If small values are chosen, then the sampling will not be large enough to give information about all the pixels and therefore, some information in the shape description will be missing.The number of pixels on the maximum radius, i.e., OE, is given by

n=$\lfloor R \rfloor$+1

Similarly, the number of pixels on the circumference of the outermost circle is given by

m=$\lfloor 2 \times \pi \times R \rfloor$+1

To cover all the pixels, the distances between any two adjacent (on the same radius or circumference) sample points have to be smaller or equal to the distance between two adjacent pixels. By keeping this in mind, OE is divided into n-1 equal distances and circles are drawn centered at O and with radii $\frac{R}{(n-1)}, \frac{2R}{(n-1)}, \ldots \frac{(n-1)R}{(n-1)}$ (as shown in Figure-11).

Next, each circle is divided into m equal arcs (with OE being one of the arcs), each arc being d = 360/m degrees. Then the intersecting points $C_{(c,a)}$ between the circles and the equally spaced arcs where c $\in$ $0, 1, \ldots, n-1$ and a$\in$ $0, 1, \ldots, m-1$ are found (see Figure 11). Notice that c = 0 corresponds to O which is considered to be the first circle.

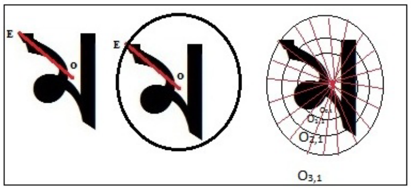


Fig. 11: forming (n-1) circles and m arcs in order to describe its shape with $C'_{(c,a)}$s as the intersecting points between the c^{th} circle and the a^{th} arc.

An n×m shape matrix is then constructed using Algorithm 1.

Algorithm 1 Shape Matrix Formation

Input: an Image *I*,
Output: The resulting shape matrix *M*.

Compute center of gravity O of I using Eq. (8)(9);

Compute *R* of I using Eq. (10), *n*, *m*;

for all $c := 0$ to $n-1$ **do**
  for all $a := 0$ to $m-1$ **do**

    // Find polar coordinates (d, θ) of $C_{c,a}$;

    $d \leftarrow \frac{cR}{n-1}$;

    $\theta \leftarrow a\frac{360}{m}$;

    // Convert (d, θ) to Cartesian coordinates (x, y);

    $x \leftarrow dcos\theta + o_x$;

    $y \leftarrow dsin\theta + o_y$;

    if $pixel(x, y)$ **then**
      $M[c, a] \leftarrow 1$;

    else
      $M[c, a] \leftarrow 0$;
    end if

  end for

end for

The shape matrix of Figure-11 for n = 4 and m = 20 is shown in Figure- 12. The resultant matrix is fed into neural network as feature vector for the segmented characters.

E. Character Recognition

At this point, we have feature representation of the individual characters,each consisting of matrix features.Next, these features need to be learned into a classifier which can classify the future instances of test road sign data. In this paper, we have used the well-known Multilayer Perceptron(MLP), a feedforward articial neural network model that maps sets

$$M = \begin{array}{cc} & \begin{array}{cccccccccccccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Fig. 12: Shape Matrix of Figure 11

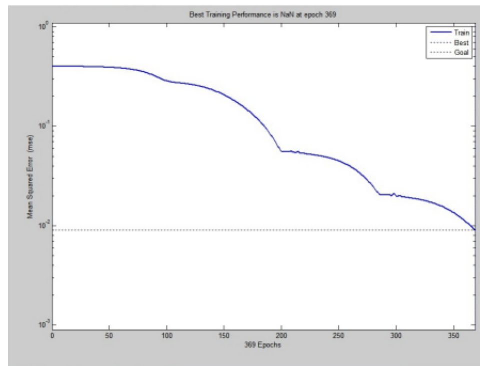


Fig. 13: The variation of MSE with training Epochs

of input data onto a set of appropriate output. For training purposes, the backpropagation algorithm [17] can be used.But the performance of this algorithm critically depends on the choices of several parameters, namely; the learning rate and the momentum factor. For a careful selection of these parameters, we have used a modied version of the algorithm known as adaptive backpropagation algorithm [18] which uses self-adaptive learning rates.The gradient descent back propagation with adaptive learning rate and log-sigmoid transfer functions are used for neural network training.The performance of the training process through MLP classifier is measured by calculating Mean Square Error (MSE) and we record the number of epochs that were needed to achieve our performance goal regarding MSE.Figure 13 illustrates the plot between MSE and number of epochs in our training phase.

F. Recognizing Test Data

When all the bangla characters are recognized,the words containing characters are saved as bag of words.Then the whole bag of words of a particular road sign are used later as learned textual road sign.So,now we have a list of real bangla road sign training data which later would be used for matching with the new testing data.

When a new image of a Bangla road sign is given,it is processed through methodology explained in section 3 and it is feed forwarded into our training model.Then for a particular bangla road sign template,the new recognized bag of words is matched with pre-specified bag of words list to confirm whether it is a Bangla road sign template or not.

IV. Experimental Results

We have divided our system in four main phases:text detection,segmentation,character recognition and bag of word confirmation.So,the overall performance of the system directly depends on the performance of the four individual phases.In

TABLE II: Experimental results of text detection

Method	R	P	F	M
Mousumi et al [10]	0.67	0.33	0.44	0.43
Proposed	0.86	0.76	0.81	0.13

TABLE III: Experimental results of segmentation

No.of text	Line Segmentation Rate	Word Segmentation Rate	Character Segmentation Rate
5	96%	97%	94%
10	97%	98.5%	94%
15	98.4%	99%	94.5%
20	99.1%	99.4%	94.8%

the text detection phase we found TDB,FDB,MDB.For each image in the dataset, we also manually count the number of Actual Text Blocks (ATB).

TDB=Truly Detected Block

FDB=Falsely Detected Block

MDB=Text Block with Missing Data

The performance measure are defined as follows.

Recall (R) = TDB/ ATB

Precision (P) = TDB/(TDB + FDB)

F-measure (F) =$2 \times P \times R/(P + R)$

Misdetection Rate(MDR) = MDB/TDB The experimental results are shown in the Table II.

Where,R-Recall (= Detection Rate), P-Precision (= 1-False Positive Rate), F-F-measure, M-Misdetection Rate. Detection method has a great effect on total recognition accuracy.From table II we can easily assume that the detection method we used is more efficient than the detection method in mousumi et al [10]. The performance of the segmentation method we used is very satisfactory.The experiment results are shown in the table III. In the character recognition phase,The recognition rate, rejection rate and substitution error of the technique are estimated using the following decision rule:

1. For all samples collect the condence levels from neural network.

2. For all samples do

If R(x) is a recognized character for sample x and j be the character with highest confidence level, then if this maximum confidence level is greater than a threshold β and j=x, then increment rec. (which is the number of samples recognized), else find the possible substitution by checking the second, third and fourth confidence levels. If any of these is greater than β, increment rej. (number of samples rejected).

Recognition rate(REC)= (number of samples recognized/total number of samples)$\times$100.

Substitution rate(SUB)=(number of samples substituted /total number of samples)$\times$100.

TABLE IV: Experimental results of recognition

parameter	mousumi et al [10]	shape based
REC(learning)	73%	95%
SUB(learning)	23.66%	5%
REJ(learning	3.34%	0.00%
REL(learning)	75.52%	95%
REC(testing)	55.5%	74%
SUB(testing)	34.5%	22.5%
REJ(testing	10%	3.5%
REL(testing)	61.66%	76.68%

TABLE V: Experimental results of confirmation

no.of road sign template	confirmation	efficiency
10	9	90%
20	18	90%
25	25	100%

TABLE VI: Experimental results of overall system with our dataset

module	efficiency of mousumi [10]	efficiency of the proposed
text detection	67.58%	86%
Segmentation	93.89%	96.81%
character recognition	89.66%	95%
connfirmation	91.78%	93.33%

Rejection rate(REJ)=(number of samples rejected/total number of samples)$\times$100.

Reliability(REL)=(Recognition/(Recognition +Substitution))$\times$100.

The results of the recognition are shown in table IV.Our method gives us a satisfactory result for rotated text image.But,mousumi [10] used a method that is unable to recognize rotated text image. Now,look at the experimental result of last part which is bag of word confirmation.The experimental result is given in table V. We made an performance comparison of our overall method with mousumi [10].The details of the comparative study is illustrated in Table VI.

V. Conclusion

In this paper, we proposed an algorithm to support the recognition of Bangla road sign template.Text detection,text segmentation,character recognition and text confirmation were the steps of our algorithm that follows laplacian approach for text detection,global horizontal projection approach for line segmentation,vertical projection for word segmentation,zone separation for character segmentation.Then for character recognition we proposed shape based approach which has a satisfactory experimental result with rotated characters.Then it follows bag of word method for confirmation.The preliminary experiments were done on a dataset created by collecting samples from volunteers.In future,we would like to extend our work including text to speech approach.

REFERENCES

[1] Nick Barnes, Alexander Zelinsky, and Luke Fletcher. Real-time Speed Sign Detection using the Radial Symmetry Detector. *IEEE Transactions on Intelligent Transport Systems*, 2008.

[2] Andrzej Ruta, Yongmin Li, and Xiaohui Liu. Real-time traffic sign recognition from video by class-specific discriminative features. *Pattern Recogn.*, 43(1):416–430, January 2010.

[3] Wen Wu, Xilin Chen, and Jie Yang. Detection of text on road signs from video. *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS (ITS)*, 6(4):378–390, 2005.

[4] Arturo De La Escalera, Luis E. Moreno, Miguel Angel Salichs, and Jos Mara Armingol. Road traffic sign detection and classification. *IEEE Transactions on Industrial Electronics*, 44:848–859, 1997.

[5] D. Ghica, Si Wei Lu, and Xiaobu Yuan. Recognition of traffic signs by artificial neural network. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 3, pages 1444–1449 vol.3, Nov 1995.

[6] M Lalonde and Y Li. Detection of road signs using color indexing. *Centre de Recherche Informatique de Montreal, Montreal, QC, Canada, CRIM-IT-95/12–49*, 1995.

[7] Y. Aoyagi and Toshiyuki Asakura. A study on traffic sign recognition in scene image using genetic algorithms and neural networks. In *Industrial Electronics, Control, and Instrumentation, 1996., Proceedings of the 1996 IEEE IECON 22nd International Conference on*, volume 3, pages 1838–1843 vol.3, Aug 1996.

[8] Saturnino Maldonado-Bascón, Sergio Lafuente-Arroyo, Pedro Gil-Jiménez, Hilario Gómez-Moreno, and Francisco López-Ferreras. Road-sign detection and recognition based on support vector machines. *IEEE Transactions on Intelligent Transportation Systems*, 8(2):264–278, 2007.

[9] S. Lafuente-Arroyo, S. Maldonado-Bascon, P. Gil-Jimenez, H. Gomez-Moreno, and F. Lopez-Ferreras. Road sign tracking with a predictive filter solution. In *IEEE Industrial Electronics, IECON 2006 - 32nd Annual Conference on*, pages 3314–3319, Nov 2006.

[10] Mohammad Osiur Rahman, Fouzia Asharf Mousumi, Edgar Scavino, A. Hussain, and H. Basri. Real time road sign recognition system using artificial neural networks for bengali textual information box. In *Information Technology, 2008. ITSim 2008. International Symposium on*, volume 2, pages 1–8, Aug 2008.

[11] Yongbum Lee, Noriyuki Takahashi, Du-Yih Tsai, and Kiyoshi Ishii. Adaptive partial median filter for early ct signs of acute cerebral infarction. *Int. J. Computer Assisted Radiology and Surgery*, 2(2):105–115, 2007.

[12] Trung Quy Phan, Palaiahnakote Shivakumara, and Chew Lim Tan. A laplacian method for video text detection. In *Proceedings of the 2009 10th International Conference on Document Analysis and Recognition*, ICDAR '09, pages 66–70, Washington, DC, USA, 2009. IEEE Computer Society.

[13] U. Pal and Sagarika Datta. Segmentation of bangla unconstrained handwritten text. In *ICDAR*, pages 1128–1132. IEEE Computer Society, 2003.

[14] Rainer Lienhart and Axel Wernicke. Localizing and segmenting text in images and videos. *IEEE Trans. Circuits Syst. Video Techn.*, 12(4):256–268, 2002.

[15] B. B. Chaudhuri and U. Pal. A complete printed bangla ocr system. *Pattern Recognition*, 31(5):531–549, 1998.

[16] Ardeshir Goshtasby. Description and discrimination of planar shapes using shape matrices. *IEEE Trans. Pattern Anal. Mach. Intell.*, 7(6):738–743, June 1985.

[17] David E. Rumelhart, James L. McClelland, and CORPORATE PDP Research Group, editors. *Parallel distributed processing: explorations in the microstructure of cognition, vol. 1: foundations*. MIT Press, Cambridge, MA, USA, 1986.

[18] Miguel Moreira and Emile Fiesler. Neural networks with adaptive learning rate and momentum terms. Idiap-RR Idiap-RR-04-1995, IDIAP, Martigny, Switzerland, 10 1995.