Bengali Longhand Character Recognition using Fourier Transform and Euclidean Distance Metric

Mousumi Hasan Mukti, Quazi Saad-Ul-Mosaheer, and Khalil Ahammad

Abstract—Handwritten Character Recognition (HCR) is widely considered as a benchmark problem for pattern recognition and artificial intelligence. Text matching has become a popular research area in recent days as it plays a great part in pattern recognition. Different techniques for recognizing handwritten letters and digits for different languages have already been implemented throughout the world. This research aims at developing a system for recognizing Bengali handwritten characters i.e. letters and digits using Fourier Transform (FT) and Euclidean distance measurement technique. A dataset with 800 handwritten character texts from different people has been developed for this purpose and these character texts are converted to their equivalent printed version to implement this research. MATLAB has been used as an implementation tool for different preprocessing techniques like cropping, resizing, flood filling, thinning etc. Processed text images are used as input to the system and they are converted to FT. Handwritten character of different person may be of different style and angle. The input dataset is collected from various types of people including age level from 5 to 70 years, from different professions like pre-schooling students, graduate students, doctors, teachers and housewives. So, to match the input image with printed dataset (PDS) each printed data is rotated up to 45° left and right and then their FT is computed. The Euclidean distance among the input image and the rotated 30 images of each printed text are taken as intermediate distance set. The minimum value of Euclidean distance for a character is used to recognize the targeted character from the intermediate set. Wrongly detected texts are not thrown away from the system rather those are stored in the named character or digits file so that those can be used in future for deep learning. By following the proposed methodology, the research has achieved 98.88% recognition accuracy according to the input and PDS.

Index Terms—Euclidean Distance Metric (EDM); Fourier Transform (FT); Fourier Transformed Printed Image (FTPI); Image Euclidean Distance (IMED); Mean Euclidean Distance (MED).

I. INTRODUCTION

The applications of different handwritten character recognition systems are increasing day by day in various fields. Handwritten characters are difficult to recognize as they are not uniform in size, and they vary from person to person, occupation to occupation, age to age. Digital character conversion as for preservation of ancient document, meaning translation, content based image retrieval; keyword spotting and text to speech conversion are some remarkable applications for handwritten character recognition. As an area of pattern recognition process, HCR has been the subject of research interest during the last few decades. Machine simulation of human functions has always been a very challenging task. The ultimate objective of any HCR system is to simulate the human reading capabilities so that the computer can read, understand, edit and do similar activities as human do with the text [1]. Bangla is considered as the fifth most popular language in the world so researchers are trying to computerize the language. Moreover, a good recognition system is needed in every language [2]-[5]. Euclidean Distance Metric (EDM) provides good recognition accuracy. From the transformed images comparing the Euclidean distance that is found from the input dataset and the printed dataset, the minimum distance value that is obtained defines the character or digit that we are comparing to or trying to find out [3].

II. METHODOLOGY

The ability to take input handwriting from sources like, printed document, images and other devices or taking handwriting directly as input from touch screen devices and to interpret them as text defines the term handwritten character recognition [9]. Researchers have already used various methods and techniques to implement text matching system [4]. The methodological steps required to conduct and implement this particular research are presented below.

A. Data Acquisition

Bangla handwritten 50 letters and 10 digits have been collected randomly from 100 people and 7 of them have been sampled using random selection process as the input dataset of this research project. A general range of size of the hand writing is given to the people so that it does not become too big or too small to recognize [2]. These characters are written with pencil, ball- pen, gel pen and of various colors to show that, these factors don’t have any impact on the recognition accuracy. Among these 7 sampled dataset there are handwritten characters of people who write with their left hand and people who write with their right hand. The initial input dataset and the printed dataset are stored indifferent folders and both of them are compared in the recognition process.

B. Cropping and Segmentation

After the collection of various samples of data from different people, the samples are scanned as all the letters and digits are written on paper document. The scanned images are cropped into a size of 24x24 pixels. Any

Published on July 31, 2018
M. H. Mukti, Q. Saad-Ul-Mosaheer and K. Ahammad are with Bangladesh Army International University of Science and Technology, Cumilla Cantonment, Cumilla, Bangladesh. (e-mail: mousumi1612@gmail.com, saad.mosaheer@gmail.com, khalil.ahammad@buiust.edu.bd).

DOI: http://dx.doi.org/10.24018/ejers.2018.3.7.831
handwritten character usually fits within this size. Image size is an important factor in text recognition for this proposed technique [2].

C. Fourier Transform of Image

A Fourier transformed image means an image that is decomposed to its sine and cosine components. The output of the transformation represents the image in frequency domain where the input image is equivalent to the spatial domain. In a Fourier transformed image each point represents a particular frequency contained in the spatial domain. The Discrete Fourier Transform (DFT) is the sampled Fourier Transform containing only a set of samples which is large enough to fully describe the spatial domain image [3].

For a square image of size N×N, the two-dimensional DFT is given by:

\[ F(k,l) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i,j) e^{-i2\pi kl/N} \]

where \( f(i,j) \) is the image in the spatial domain and the exponential term is the origin, corresponding to each point \( F(k,l) \) in the Fourier space. The equation can be interpreted as: the value of each point \( F(k,l) \) is obtained by multiplying the spatial image with the corresponding base function and summing the result.

D. Calculate Image Euclidean Distance

According to the gray level of each pixel image with a fixed size P×Q can be written as a vector, \( x = \{ x_1, x_2, ..., x_{PQ} \} \). The Euclidean distance \( d_E(x_1, x_2) \) between vector images \( x_1 \) and \( x_2 \) is defined as,

\[ d_E^2(x_1, x_2) = \sum_{i=1}^{PQ} (x_1^i - x_2^i)^2 = (x_1 - x_2)^h_{PDS}(x_1 - x_2) \]

Taking different dimensions into account IMED introduces the metric matrix \( G \). The IMED Euclidean distance \( d_E^G(x_1, x_2) \) between images \( x_1 \) and \( x_2 \) is defined as,

\[ d_E^G(x_1, x_2) = \sum_{i=1}^{PQ} \sum_{j=1}^{PQ} g_{ij} (x_1^i - x_1^j)(x_1^j - x_2^j) = (x_1 - x_2)^{h_{PDS}} G (x_1 - x_2) \]

\[ g_{ij} \] indicates the metric coefficient which denotes the spatial relationship between pixels \( p_i \) and \( p_j \).

\[ d_E^2 \] indicates the spatial distance between \( p_i \) and \( p_j \) and the width parameter. For example, if \( p_j \) is at location \( (m, n) \) and \( p_j \) is at location \( (m', n') \) then \( d_E^2 \) is calculated as follows:

\[ d_E^2 = \sqrt{((m - m')^2 + (n - n')^2)} \]

IMED considers the spatial relationship between pixels and it is very sensitive in deforming the pixels [8].

III. IMPLEMENTATION

The proposed methodology has been implemented according to the following flow diagram:

A. Preprocessing PDS

The printed dataset that are written in computer has been scanned in the initial step. The scanned images are then resized and cropped according to predefined specifications. Each cropped and resized images are then saved in a folder named printed dataset.

![Fig. 1. Flow Chart for PDS Processing](image)

B. Preprocessing and Text Detection

Data acquisition and dataset construction is a prerequisite and also vital step to process the IDS before those are compared with printed dataset. The writing style varies from person to person. The font size and styles are also different. To make them comparable with printed dataset, some preprocessing tasks are required. The complete workflow starting from capturing input image to detect it, the following steps are followed:

![Fig. 2. Flow chart for Preprocessing a handwritten character](image)

In the Fig. 2 the entire technique is defined how handwritten characters are recognized by this proposed system implemented using FT and EDM.

DOI: http://dx.doi.org/10.24018/ejers.2018.3.7.831
Step-1  
Step 1 deal with capturing the image of the handwritten character.

Step-2  
In MATLAB tool box, the images are then resized into a size of 24x24 pixel as it is the prescribed size for representing handwritten.

Step-3  
A morphological operation used to remove foreground pixels from binary images more like erosion or opening. Flood filling technique is then implemented on the images to fill the holes in the handwritten digits. Both these operations are applied on the resized images. These operations have a very significant impact on increasing the text detection accuracy level.

Step-4  
In this step, the FT of the resized and preprocessed input images are computed. This step is unavoidable as the reasearch is using the FT of both input and printed text images.

Step-5  
In this step, the Euclidean distance of two fourier transformed images is calculated. To transform the printed text into FT each printed text is fetched from the previously saved PDS folder which will be compared with Fourier Transform of Input Image (FTInI).

Step-6  
Each resized printed image are being rotated max 45° and the interval of rotating the image is 3°. Images are rotated left and right. In each rotation 30 newly rotated images are found and then the Fourier Transformed images are determined. These images are then taken as a set of FTPI.

\[ FTPI = \{FTPI1, FTPI2, FTPI3 \ldots \ldots FTPI30\} \]  

Step-7  
The intermediate set of minimum Euclidean distance is determined in this step. MED is calculated between the FTInI and each FTPI of every set. MED is calculated from every set of FTPI of 60 letters/digits. Thus 60 MED is computed for 60 characters and each MED is obtained according to the following formula:

\[
MED1 = MIN\{(ED1,ED2,\ldots \ldots ED30)\}
\]
\[
MED2 = MIN\{(ED1,ED2,\ldots \ldots ED30)\}
\]
\[
MED3 = MIN\{(ED1,ED2,\ldots \ldots ED30)\}
\]
\[
\ldots
\]
\[
MED60 = MIN\{(ED1,ED2,\ldots \ldots ED30)\}
\]

Step-8  
From the intermediate set of MED only the smallest value along with its reference image is shown, which is the final result of the procedure. The correctly detected images (CDI’s) are stored in the system and are allowed to act as the PDI’s to be compared with the input handwritten images once again.

\[ Detected_{image} = MIN\{(MED1, MED2, \ldots MED60)\} \]

Step-9  
Within the printed dataset folder there are 60 (as the examined number of Printed digits and characters are 60 in number) other folders for every letter or digit that are correctly detected. Every time when an input character is checked it is compared with the printed input as well as the previously correctly detected letters. When it matches then this newly detected character will be stored into the folder of its name to its related location. So, when the next character comes it will again be compared with the printed character as well as the two other correctly detected characters and this will be continued for each character. Wrongly detected images are not discarded from the system. Rather as it matches with any one of the characters or digits, it is stored gain to enrich system’s PDS. By doing this, the system can detect whether the character is written in the prescribed format or not. For example, the experimented system wrongly detected one of the samples of খ as খ. But after the wrongly detection process, the image was not discarded rather it was stored as খ. Next if a wrongly writing pattern comes in the system, it can be compared with the character and can be detected correctly. This will help in the process of collecting enormous hand writing from people and make a training data set in deep learning.

IV. RESULT AND ANALYSIS  
For the evaluation of the proposed technique of recognizing handwritten characters, 100 samples of data are taken from 100 people and a complete set of 60 characters and digits from 6 people are considered. From the rest other people, random characters are taken as input and are checked by the proposed system. Among these people, there are people of different occupations, different sex and different ages as the handwriting varies with these parameters. Among these 100 people, there are people who write using left hand and right hand.

A. Resultant Tables  
The results of the experiments of randomly picked 6 samples those are correctly detected are presented here along with their actual printed digit/character and MED.
<table>
<thead>
<tr>
<th>Sample</th>
<th>Detect Texts' MED</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0204</td>
</tr>
<tr>
<td>1</td>
<td>0.8594</td>
</tr>
<tr>
<td>2</td>
<td>1.0725</td>
</tr>
<tr>
<td>3</td>
<td>0.5172</td>
</tr>
<tr>
<td>4</td>
<td>0.7178</td>
</tr>
<tr>
<td>5</td>
<td>0.4172</td>
</tr>
<tr>
<td>6</td>
<td>0.6882</td>
</tr>
<tr>
<td>7</td>
<td>0.592</td>
</tr>
<tr>
<td>8</td>
<td>0.492</td>
</tr>
<tr>
<td>9</td>
<td>0.392</td>
</tr>
<tr>
<td>10</td>
<td>1.0057</td>
</tr>
<tr>
<td>11</td>
<td>0.778</td>
</tr>
<tr>
<td>12</td>
<td>0.592</td>
</tr>
<tr>
<td>13</td>
<td>0.778</td>
</tr>
<tr>
<td>14</td>
<td>0.592</td>
</tr>
<tr>
<td>15</td>
<td>0.778</td>
</tr>
<tr>
<td>16</td>
<td>0.592</td>
</tr>
<tr>
<td>17</td>
<td>0.778</td>
</tr>
<tr>
<td>18</td>
<td>0.592</td>
</tr>
<tr>
<td>19</td>
<td>0.778</td>
</tr>
<tr>
<td>20</td>
<td>0.592</td>
</tr>
<tr>
<td>21</td>
<td>0.778</td>
</tr>
<tr>
<td>22</td>
<td>0.592</td>
</tr>
<tr>
<td>23</td>
<td>0.778</td>
</tr>
</tbody>
</table>

**TABLE I: DETECTED TEXTS' MED**
B. Resultant Graphs

Each graph of this section represents the input images and their distance with the printed data presented in Table I. In graph the distance values and the samples of each digits and letters are shown in Y and X axis respectively.
C. Wrongly detected texts

Recognizing handwritten characters can be very difficult when there are large numbers of data in the dataset and there are many characters which are very similar to each other. As a result, all the characters cannot be perfectly matched with the input dataset. In the above graphs presented in section 4.2, the matched letters and digits are shown where the least value Euclidean distance for those characters matches the most with the printed dataset. Because of the similarity of characters, this proposed system cannot match some characters accurately. The eerily detected characters are presented in Table II.

D. Recognition Accuracy

There are 60 considerable characters for Bengali language and this system has taken 7 samples for each of the characters. Thus, there are 420 input characters (60 characters × 7 samples for each character). The dataset also consists 380 randomly collected samples. Total 9 characters were detected wrongly.

So, Precision = (Number of correctly detected text / Number of detected text) [6]. For this system, precision is 0.98875 and 1.125% characters have been detected wrongly. So, the recognition accuracy of the system is 98.875%.

V. CONCLUSION

The main objective of this research is to achieve 100% accuracy in handwritten character recognition. But it is impossible to obtain that much accuracy. Human being themselves cannot sometimes recognize handwritten text properly [7]. Variations of input characters in terms of style, size and angle have been found as the key constraints to achieve 100% recognition accuracy. In addition, some characters are very similar in Bengali language. This proposed system has come up with an excellent accuracy level of above 98% in determining each Bengali handwritten letter and digit.

REFERENCES


