Number Recognition from Captured Images

Randhir J. Patil
Assistant Professor
Sanjay Ghodawat Institutes,
Atigre, India

Akshay G. Bhosale
Assistant Professor
Sanjay Ghodawat Institutes,
Atigre, India

ABSTRACT

Now-a-days because of digitization, it is very important to have all the data in the form of soft copy so that it is easy to store and maintain. Often the handwritten documents are available, but again if we want to use the data written in that document then we need to type that data in any word processing software and then use it. This is very time consuming and tedious job. In this paper an attempt is made to automatically recognize numbers in the document by capturing images of that document. Till now emphasis is given only on recognizing digits or numbers from images. This methodology is implemented in automatic marks filling system. This system captures images of the tables in which marks are written by the teacher in the students' answer book, recognize the numbers or marks from it and fill the same in appropriate database on the computer. The fundamental technique used here is Dynamic Time Warping (DTW). Although this technique is used for speech processing, this paper describes its exclusive use in number recognition from its images. The system is implemented and simulated in MATLAB.

General Terms

Pattern Recognition, Character Matching

Keywords

Image Processing, DTW Technique, Data Acquisition, Data Transformation, Number Recognition, Automatic Marks filling System.

1. INTRODUCTION

Due to digitization it is necessary that whatever data we have that must be available in the form of soft copy. If the data is available in the form of soft copy then it is very easy to maintain it, store it and also process it. But if the data is available in the form of hard copy or handwritten form then it is very time consuming and tedious job to convert the data available in hard format to soft format. It is generally accepted that handwritten characters are hard to recognize. The primary reason is that characters are written in uneven manner. There are some electronic notepads which can be attached to a computer, so whatever we write on that notepad gets converted to digital text. This requires additional software to be installed on the computer. But if we have data already written on a piece of paper then it is difficult to directly convert it to digital format. In such cases this system can work well. There are several technique for isolated template recognition, like Optical Character Recognition (OCR), Artificial Neural Network (ANN), Correlation, etc. Several investigations have been carried out to determine the best technique. It is necessary to develop a system which may have a complex recognition algorithm and may have lower accuracy. But if limited characters are considered like 0 to 9 numbers only, then there is possibility to increase accuracy as well.

The DTW technique that has been described here is most probably used in speech template recognition. It is also applicable to the isolated speech recognition. In speech processing two feature extracted matrices of respective speech templates are compared to find minimum distance. Same technique is used here to develop this system, but two matrices are generated from captured images of characters. The ultimate objective here is to compare two matrices generated from images of the number.

The rest of the paper is organized in different sections as follows. Section II describes the basics of DTW technique. Section III describes detail methodology and the algorithm used in implemented system. After implementing this system in MATLAB, results were obtained and those are discussed in Section IV. Finally, Section V describes the future scope concluding remarks.

2. DTW ALGORITHM

The key idea behind DTW is to align two non-temporal speech sequences and finding minimal DTW score between them. Consider following example shown in figure 1. This figure illustrates the DTW technique used in speech processing. In DTW matrix is computed using MFCC coefficient values. Take all the frames (each frame consist of 13 MFCC coefficient) of stored voice symptom on y-axis and that of testing symptom on x-axis [1][2][3][4].
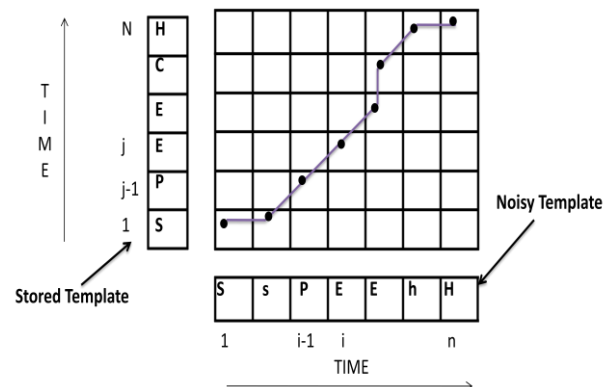

Fig 1: DTW aligning

Calculate local distance between two corresponding frames using Euclidian distance formula sqrt $[(X_i - Y_j)^2]$. Thus cell (n, m) of DTW matrix is result of Euclidian distance between nth frame of test sample and m^{th} frame of reference sample [5][6][7][8]. Then minimum DTW path is constructed through cell (1,1) to cell(r, t) where 'r' and 't' are last frames of reference and test sample respectively. Every next cell is chosen using formula min (celli, j celli-1,j celli,j-1). Thus by calculating DTW path between test signal and all reference signal, nearest neighborhood approach gives recognized word/symptom [9][10]. Here instead of speech features

matrices, image extracted matrices are compared using DTW algorithm.
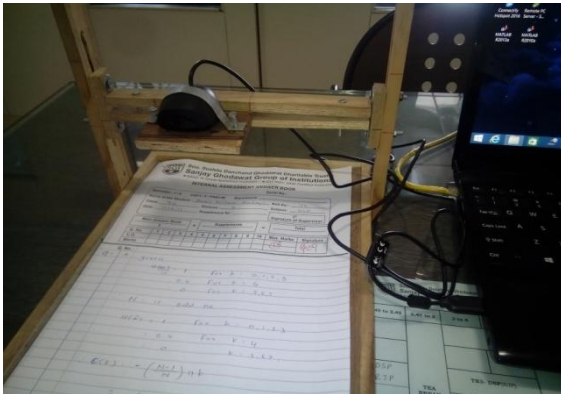
3. HARDWARE SETUP AND METHODOLOGY

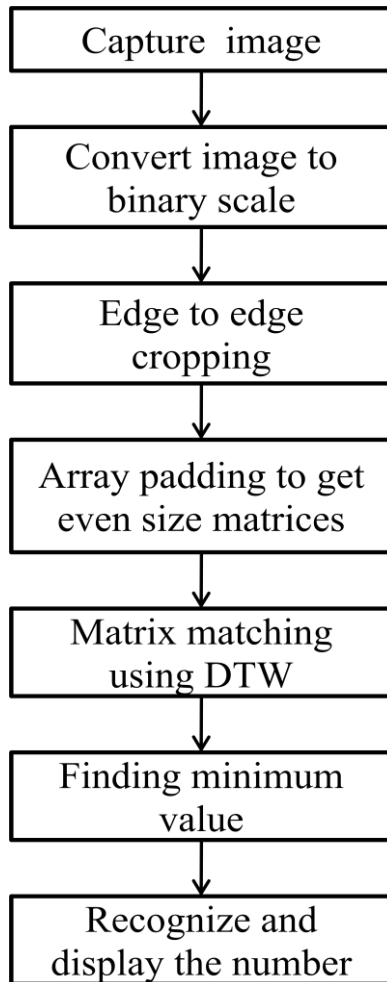**Fig 2: Hardware setup**

Fig 3: Methodology

In the figure 2 hardware setup is shown. Here a camera is mounted on wooden frame with adjustable up, down, right and left movement. The size of wooden frame is same as that of answer sheet. This setup will capture each image from same position. Thus coordinates of marks table will approximately remain same, this avoids change in dimensions of the images, provided that the capturing area is of same size. The system first captures image of the answer sheet. The

written number region is extracted using edge to edge cropping. The resulting data is compared with the data in database so as to recognize a particular number.

The detail methodology is explained in figure 3. It consists of different stages such as Image capturing, binary conversion, Image Cropping, Edge to edge cropping, DTW path calculation, Recognition etc.
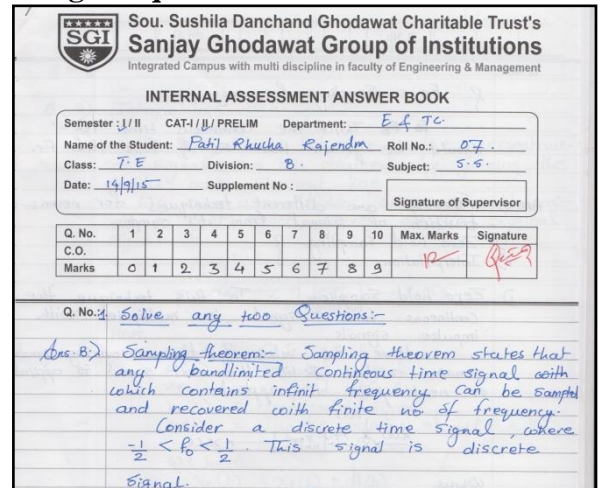
3.1 Image Acquisition

Fig 4: Image captured by camera

The images can be acquired through a digital camera or a webcam. In this system image acquisition is done using 3.1 MP webcam. The system acquires an image of handwritten marks table on the answer sheet. A sample of the captured image is shown in figure 4.

3.2 Database Creation

The pattern of handwritten characters depend on the writing style of an individual. Hence every time new database has to be prepared by the user itself. The first one or two images are considered as reference images in the database. In reference images all the numbers (0 to 9) should be present in the sequence. Thus all the matching of numbers will be done with reference to these database images only.

3.3 Pre-processing

Depending upon the region of interest, the co-ordinates are found and cropping takes place accordingly. These coordinates are based on marks filling area on the answer sheet. Because of uniform paper size and fix capturing point, the coordinates to crop image will always remain same. The captured images are cropped to proper dimensions so as to extract marks table only. The sample cropped image is shown in figure 5.

Fig 5: Cropped image

Image in the figure 5 can be considered as database image since it has all the numbers on it in the sequence. The cropped image is converted into binary image and cropped again to get border free individual number as shown in figure 6. The images shown in figure 6, figure 7, figure 8 and figure 9 are all binary images. For visualization of cropping effect, the background is made grey.

Fig 6: cropped number images.

In the marks table on the answer sheet a number can be written anywhere in the provided space. Hence, further edge to edge cropping is done on all the number images as shown in figure 7.

Fig 7: Edge to edge cropped image

Since the images to be cropped are binary, hence detection of either zero or one is possible. This approach is called as edge processing in which we scan the region of interest from boundaries, which helps to isolate target.

3.4 DTW Implementation

To apply the DTW technique, two matrices are required. Here the matrices are obtained from binary data of the edge to edge cropped numbers image. But, because of different size of handwritten numbers size of their respective matrix is also variable. To overcome this and to improve the accuracy of matching, padding technique is used. In padding, binary matrix of a cropped image is added to some fixed size matrix which contains only zeros. This process is shown in figure 8.

Fig 8: Uniform size Image

After padding, the binary matrix of each image will have uniform size. The DTW algorithm is now applied on sample and reference images. Thus final matrix of each number will be compared to matrices of all the 0 to 9 digits. For the best match, DTW algorithm gives minimum value.

4. RESULTS AND DISCUSSION

The summary of complete process is shown in figure 9.This system is tested for single digit recognition only. Ten different answer sheets were tested. The DTW algorithm provides best possible matchas compared to correlation matching algorithm. The values obtained after DTW processing are summarized in the table 1 and table 2. The accuracy of the system depends on writing style of the user. The system gives 100% accuracy for all the numbers except SEVEN and ONE. In table 1, the comparison of 0 to 4 sample digits is done with 0 to 9 digits available in the database. The orange cells in table 1 show that when a perfect match occurs, then DTW algorithm generates minimum value. For example, the column one shows the comparison of sample digit 0 with all the digits present in the database.In column one, DTW algorithm generates minimum value for database digit 0, this shows that sample digit 0 has been accurately recognized as digit 0. In column 2, the sample digit 1 is compared with database digits.
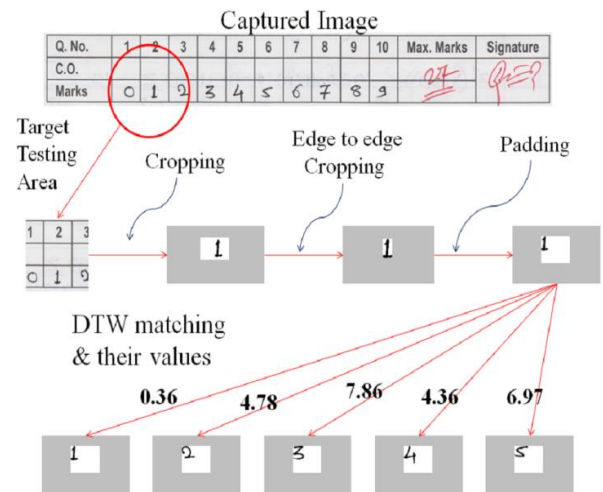

Fig 9: Summary of the process

Here it can be seen that the minimum value in the column 2 is for database digit 7, indicated by grey colored cell, so this indicates that system recognized sample digit 1 as digit 7, which is incorrect.Similarly, the sample digit 7 is recognized as digit 1, which is incorrect. But all other sample digits are perfectly recognized.

Table 1: DTW values for numbers 0 to 4

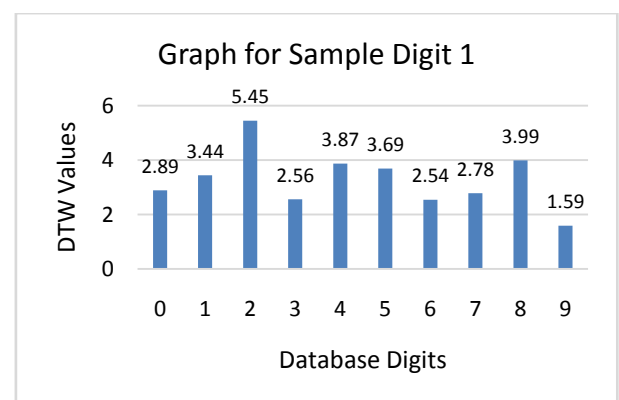| Sample Nos. DatabaseNos.	0	1	2	3	4
0	1.28	2.35	6.66	2.35	3.99
1	5.69	1.69	5.64	2.45	4.85
2	3.25	2.86	0.89	3.48	4.56
3	1.80	3.99	1.49	1.23	2.69
4	2.66	3.44	8.66	6.85	1.50
5	4.56	4.85	2.56	3.69	3.44
6	2.69	2.88	3.87	2.54	5.45
7	8.66	1.40	2.89	2.78	2.54
8	2.45	2.66	3.44	3.86	2.78
9	6.99	8.34	5.45	4.56	2.86

Fig 10: Graph for sample digit 1

In table 2, the comparison of 5 to 9 sample digits is done with 0 to 9 digits available in the database.

Table 2: DTW values for numbers 5 to 9

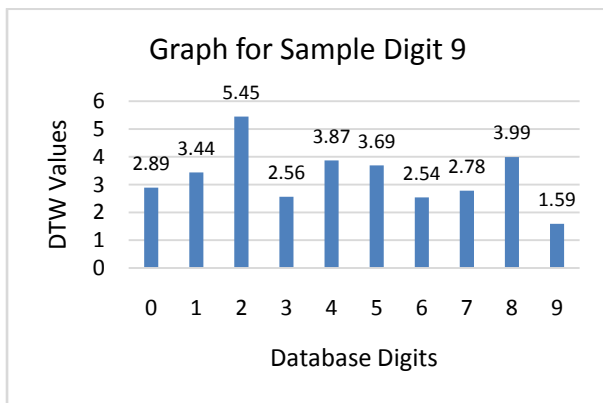| Sample Nos. DatabaseNos.	5	6	7	8	9
0	4.56	4.85	3.56	2.40	2.89
1	2.45	2.86	2.45	3.66	3.44
2	3.48	3.99	6.99	8.34	5.45
3	7.23	3.44	8.66	5.69	2.56
4	6.85	5.88	9.46	6.86	3.87
5	1.85	2.88	3.86	3.99	3.69
6	2.88	1.57	4.56	3.44	2.54
7	6.40	2.66	2.58	4.85	2.78
8	2.66	8.34	5.45	2.22	3.99
9	5.80	3.92	8.46	2.82	1.59

Fig 11: Graph for sample digit 9

Figure 11 shows the graph for sample digit 9. From the graph, it is clear that the DTW value for the database digit 9 is smallest. This indicates that the sample digit 9 matches with database digit 9 and hence system recognizes the digit correctly.

5. CONCLUSION

The automatic marks filling systemuses a series of image processing techniques for identifying the handwritten marks from the captured image of marks table from the answer sheet. The system uses the DTW algorithm for comparison of image matrices. This gives better results than correlation matching algorithm. The system is implemented using MATLAB. The appropriate GUI (Graphical User Interface) is designed to bundle the executable application. The simulation results show that the system robustly recognizes the marks and fillsthem automatically to excel sheet or any database, and can be implemented in school or any educational organization for mass data entry.More accuracy could be achieved by applying more advanced algorithms. This idea could be further extended to detect two digit numbers and also for handwritten letters in an alphabet. Also some algorithms could be thought of to improve the accuracy of character recognition.

6. ACKNOWLEDGEMENT

7. REFERENCES

[1] ChadawanIttichaichareon,SiwatSuksri, and ThaweesakYingthawornusuk, " Speech Recognition Using MFCC," International Conference on Computer Graphics, Simulation and Modeling(ICGSM'2012),pp.135-138, July 28-29, 2012.

[2] Fang Zheng, Guoliang Zhang and Zhanjiang Song, "COMPARISON OF DIFFERENT IMPLEMENTATIONS OF MFCC," Journal Computer science & Technology, 16(6): 582-589, Sept. 2001.

[3] Yan-Sheng Lin and Chang-PengJi, "Research on Improved Algorithm of DTW in Speech Recognition," International Conference on Computer Application and System Modeling (ICCASM 2010), pp.418-421, 2010 IEEE.

[4] ShivankerDevDhingra, GeetaNijhawan, and PoonamPandit, "ISOLATED SPEECH RECOGNITION USING MFCC AND DTW," International Journal of advanced Research in Electrical,Electronics and Instrumentation Engineering, Vol.2, Issue 8, August 2013.

[5] Li Yang, Le Jing, Yang Yuxiang, and Wang Jian, "Improvement Algorithm of DTW on Isolated-Word Recognition", International Conference 2011 IEEE.

[6] Shicheng Li, and HaipengRen, "An isolated word recognition system based on DSP and improved dynamic time warping algorithm,"pp.136-139, International Conference 2010 IEEE.

[7] LindasalwaMuda, MumtajBegam and I.Elamvazuthi, "Voice Recognition Algorithm using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time warping (DTW)Technique", JOURNAL OF COMPUTING.VOLUME 2.ISSUE 3,MARCH 2010,ISSN 2151-9617.

[8] Ran Yaniv and David Burshtein, "An Enhanced Dynamic Time Warping model for improved estimation of DTW Parameters", IEEE TRANSACTION ON SPEECH AND AUDIO PROCESSING, VOL.11, NO.3, MAY 2003.

[9] Zhang Jing and Zhang Min, "Speech Recognition system Based Improved DTW Algorithm," International Conference on Computer, Mechatronics, Control and Electronics Engineering (CMCE), 2010 IEEE.

[10] Talal Bin Amin and IftekharMahmood, "Speech recognition Using Dynamic Time Warping,"2nd International Conference on Advances in Space Technology, pp.74-79,29th -30th November, 2008 IEEE