Journal of Science and Technology
ISSN: 2456-5660, Volume 6, Issue 6 (NOV-DEC 2021)
www.jst.org.in *DOI: https://doi.org/10.46243/jst.2021.v6.i6.pp01-09*

Implementation of Multiple Test Paper Score Cumulating System based on Digital Image Processing

Dr.T.Somassoundaram
Professor, DMI-St John the Baptist University, Mangochi, Malawi.
somsthy2005@gmail.com

To Cite this Article
Dr.T.Somassoundaram, ***"Implementation of Multiple Test Paper Score Cumulating System based on Digital Image Processing"***, *Journal of Science and Technology, Vol.6, Issue 6, NOV-DEC 2021, pp.:01-09.*

Article Info
Received: 18-11-2021 *Revised: 24-11-2021* *Accepted: 01-12-2021* *Published: 07-12-2021*

ABSTRACT

In this paper, through the study of the cumulating system and the final realization process of the accumulation system, the workload of teachers is reduced. This paper studies the implementation of handwritten scores cumulating in test papers. Multiple test paper scores are cumulated based on digital image processing on MATLAB programming platform. The features are extracted and character is identified by analyzing its features and comparing its features that distinguish each character. Besides, the accuracy of the results of the cumulating system is high as the reasonable image processing methods. In the meantime the experimental program is relatively simple with high efficiency under certain requirements.

Key Words: Hough transforms; image preprocessing; character segmentation;

I. INTRODUCTION

In the current situation many software applications and artificial systems are applied to reduce our daily work. For the education industry also we are using the software applications to reduce the complication of teachers. Handwritten test paper score accumulation of students is difficult for teachers and the processing time is long when it is done by manually. After handwritten score accumulation, teachers should verify the final scores is correct or not. Finally it takes lots of time for the score accumulation of test papers. As a result, to overcome the teacher's burden, this paper helps to solve the problem of handwritten manual scores and it is effectively for time management, accuracy and reliability with high speed.

A. BASIC PRINCIPLE OF HOUGH TRANSFORM

Hough transform is a voting principle of the parameter estimation era, specifically used to extract geometric shapes with uniform traits from an original image. It tasks all the points on an immediately line or a curve of the original image to the intersection of all the strains in a parameter space. Within the parameter area, the points are truly collected within the segmented accumulator gadgets, and the peaks of the accumulator devices are searched. The positions in the unique picture corresponding to the accumulator devices in the parameter area in which these peaks locate are the presence of heterosexual lines or curves, wherein instantly strains or curves in the photo may be detected [1]. In truth, the essence of Hough transform is to cluster the pixels with a positive relation inside the authentic picture area, and discover the corresponding parameter area accumulation factors [2] which could companion those pixels with an analytical shape.

B. SEARCHING RED CHARACTERS BASED ON HSV MODEL

The HSV color version is evolved from the CIE 3D color space. Within the HSV color model, every color is represented by hue, saturation, and value. As proven in fig.1, the hexagonal pyramid is the three-dimensional representation of this color model. On this model, the bottom surface of the hexagonal pyramid is placed at the horizontal aircraft of coordinate gadget. The aircraft includes two vital parameters, one of which is hue, and the opposite is saturation. Hue rotates around the vertical coordinate). Six standard colors are marked in fig.1, with adjacent pure colors separated by 60 degrees. Saturation changes along the horizontal path, with the lowest saturation on the origin of the coordinate gadget. Value is located at the vertical axis of the coordinate system, converting usually along the vertical direction [3].

The following nonlinear transformation can trade the RGB color space into HSV color space [4]. It can rise up singular points and unstable fixed factors.

$$h\begin{cases} \text{undefined} & \text{if } \max = \min \\ 60° \times \frac{g-b}{\max-\min} + 0°, & \text{if } \max = r \text{ and } g \geq b \\ 60° \times \frac{g-b}{\max-\min} + 360°, & \text{if } \max = r \text{ and } g < b \\ 60° \times \frac{g-b}{\max-\min} + 120°, & \text{if } \max = g \\ 60° \times \frac{g-b}{\max-\min} + 240° & \text{if } \max = b \end{cases}$$

$$s = \begin{cases} 0, & \text{if } \max = 0 \\ \frac{\max - \min}{\max} = 1 - \frac{\min}{\max}, & \text{otherwise} \end{cases}$$

$v = \max$

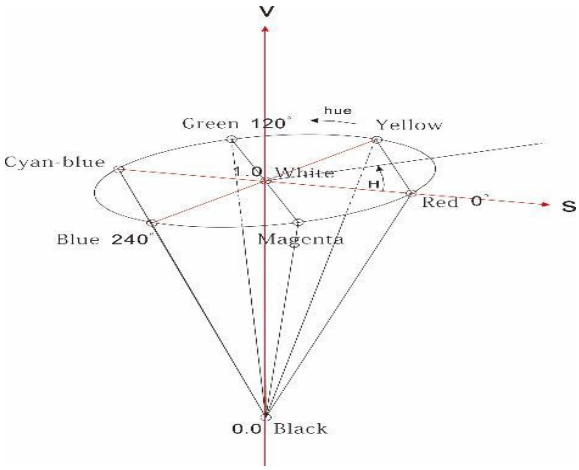

Figure 1: Three-dimensional HSV hexagonal pyramid.

C.INTRODUCTION TO ANN

ANN is a mathematical model that imitates the shape and questioning methods of brains. ANN may be used to clear up a few tough issues. Its miles a hotspot branch of the artificial intelligence subject. ANN algorithms learn and analyze through facts and features, observe appropriate treatments to grow to be smarter. ANN should have neurons, and these neurons are linked to every other to shape a neural network structure. ANN's getting to know and educating system want to be remembered. Such reminiscence is represented thru the price of synapses which exist among neurons [5]. Compared with other computational models, ANN has as superb benefit information dispensed storage and parallel processing, which greatly improves the performance of statistics processing. Definitely, ANN is a nonlinear dynamic system; the connecting systems of networks have effect on the studying and training competencies in varying tiers. View from the relationship among the diverse variables within the device, the community reminiscence and reputation modes are increase features of the range of neurons [6]. The overall performance of ANN is stricken by four fundamental factors, specifically, the quantity of neurons, connections, community topology and gaining knowledge of strategies.

D.BP ALGORITHM

BP neural community is largely an errors back propagation algorithm of gaining knowledge of method. The studying manner includes two components. One is the forward propagation process, and the alternative is the error backward propagation process. For the forward propagation of information, the neurons within the enter layer in the network model receive the input facts first, and then transmit this statistics to the neurons within the hidden layer, at remaining transfer it to the output layer. Due to influence effect on the neural network overall performance, the very last output can be extraordinary from the anticipated end result. On the same time, the error back propagation procedure begins, which is opposite to the course of the ahead training. The error of the end result propagates from the output layer to the enter layer, and modifies the connection synaptic values between the neurons in every layer. The whole manner follows the mistake gradient descends principle. The two contrary instructions of the propagation method execute alternately,

that allows you to make certain that mistakes maintain to decrease via the adjustment of the synaptic values .the mistake decreases till it reaches an applicable fee, and the training may be stopped [7]. There are varieties of BP neural community fashions, and the maximum typically used are node enter and output version, position feature model, mistakes calculation version and self –mastering version, etc. [8].

II. REALIZING SCORE CUMULATING SYSTEM

The specific steps of implementing the score cumulating system are as follows:

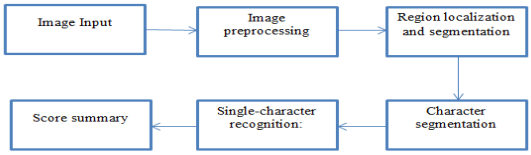

Figure 2: Block diagram of realizing score cumulating system.

Image input: Inputting the images to be processed into computer and then loading the images in the cumulating system.
Image preprocessing: Including image conversion, DE noising, image edge detection, line extract ion and image rotation, etc.
Region localization and segmentation: Positioning the approximate location of the handwritten scores and then dividing the rectangle region
Character segmentation: Separating each question's scores from the rectangle region, and then separating the numbers into several single characters.
Single-character recognition: This step is based on the results obtained in the previous steps, analyzing each character image segmented in previous step, extracting the features, recognizing numbers through the established neural network and finally determining the character category.
Score summary: Calculate the scores of each question, summing all the scores and output the results.

A.IMAGE PREPROCESSING

Image preprocessing is the first step of this experiment, which performs a critical function. The original images no longer most effective have the contour, aspect and color facts of the image, such information may even occupy a certain amount of storage space, which leads to a lower gadget going for walks speed. Consequently, below regular situations, maximum of the time we choose to transform the format of the image with a view to optimize the device and remedy the hassle more quickly and efficaciously. For instance, unique photograph are transformed to Gray scale or binary snap shots. Picture preprocessing is vital Earlier than the photo evaluation, so that the records contained in The photograph may be simplified; simplest on this way are we able to Understand the item characteristics of photographs greater without a doubt and immediately the edges of the image talk to the component wherein the local brightness modifications inside the image is relatively obvious. They could completely replicate the image texture and the shapes of the simple features. In order to facilitate the use of Hough transform step one is image edge detection.

B.SPECIFIC STEPS IN IMAGE PREPROCESSING

- Converting the original images into binary images.
- Using canny operator to detect the edges of the images, and remove small and isolated fake edge segments.
- Using Hough transform to extract lines.
- Adjusting the lines into horizontal direction.

C.SCORE REGION POSITIONING AND SEGMENTATION

The localization and segmentation of the test score region is one of the most crucial steps inside the implementation of the cumulating system with the belief of this step, we are able to lock the exact area of the score location and phase the handwritten ratings correctly. The consequences acquired on this procedure are used in the subsequent steps, so the accuracy of the segmentation is directly associated with the very last recognition price of the whole device.

SCORE AREA POSITIONING

In order to ensure the test score positioning and segmentation, we look at the basic characteristics of the image. Teachers use red signature pens while marking test papers. The color property of the score areas is obviously different from that of the surrounding areas, so it is convenient to locate the score area. First off, the unique image is transformed from RGB to HSV, and the 3 components of H, S and V are received. The three basic components of red color are obtained by referring to literatures. By limiting the three components at the same time, the location of score areas can be found. Due to the problems of image quality and clarity, a morphological noise filter consisting of two basic operations which are opening and closing, is applied here to remove small objects that are not related to the object.

SCORE REGION DIVISION

Once a score area has been locked, the x and y coordinate values of the score region are saved, and the minimum and maximum values in both values are recorded, which can be used to reduce aside the rectangle area within which the score are placed. The flow chart of score area positioning and segmentation is shown in Figure 3.

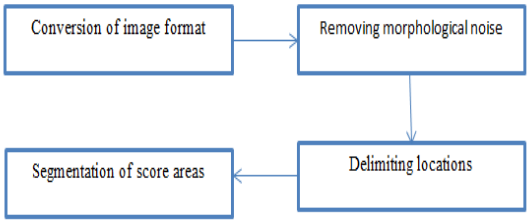

Figure.3 Flow chart of score area positioning and segmentation.

D.CHARACTER SEGMENTATION.

The segmentation of a single character plays a transitional role in the system. The basic idea of this step is to use the matrix projection method, in order to determine whether the X direction exists the score character pixels or not. Figure 4. shows the projection of a matrix of a score region, with each element in the matrix being either 0 or 1. The projection of this matrix is a one-dimensional vector, which deposits both non-zero elements, and 0 elements. The location of the 0

elements means that the corresponding position has no character pixels. Therefore, characters can be split based on 0 and non-0 element positions. For the images in the experiment, the scores of each question are different, with both single-digit scores and two-digit scores. Therefore, we not only need to separate the scores of all questions, but also to separate the ten digit numbers of each question's score. In the experiment, the scores of questions should be separated first, and then the digital numbers should be segmented. In order to distinguish the segmentation of the question score and segmentation of a single number, a 1-dimensional array is defined to record the distance between all the characters presented in the image, and then averages of these values stored in the array. We consider the characters to be in different questions if the distance of adjacent characters is longer than the mean distance, and as for distance values smaller than or equal to the mean, we consider the adjacent characters come from the same question's score. The decision is expressed as follows.

$$\begin{cases} d_i \le \frac{1}{n}\sum_{i-1}^{n} d_i & \text{The i} - \text{th interval is from the same question} \\ d_i > \frac{1}{n}\sum_{i-1}^{n} d_i & \text{The i} - \text{th interval is from different question} \end{cases}$$

Where

- d_i represents the i − th interval in the projected matrix
- n represents the number of total intervals.

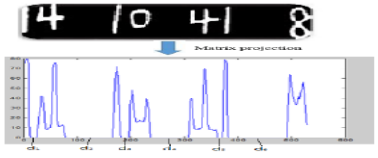

Figure.4 Matrix projection results

E. CHARACTER RECOGNITION

In the implementation of the overall device, character recognition is a very critical element. It determines whether the very last result is accurate or now not. There are many kinds of strategies that we are able to use. There are many kinds of methods that we can use. The most common method is the template matching algorithm and the ANN algorithm. By comprehensively analyzing the accumulation system and considering that the test scores are handwritten, we conclude that the success rate of the template matching algorithm will not be high. In order to ensure the accuracy of the character recognition and the final results, a self-coding neural network is applied in this paper. The self-coding neural network algorithm in the experiment uses a single-character image and its Gabor feature as the input node of the network for single character recognition. We construct each individual character image obtained in the previous steps. A self-coding neural network is used to learn each image. Finally, the correlation between the sample and the true final result are calculated. The maximum correlation value is the result of character recognizing.

F. SCORE SUMMARY

Score summation is the last procedure of the experiment. The main process in this step is to calculate the true score of each question, and finally sum the scores and display the results.

Figure 5: score summary for single student

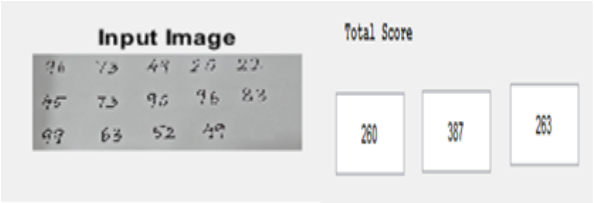

Figure 6: score summary for multiple students

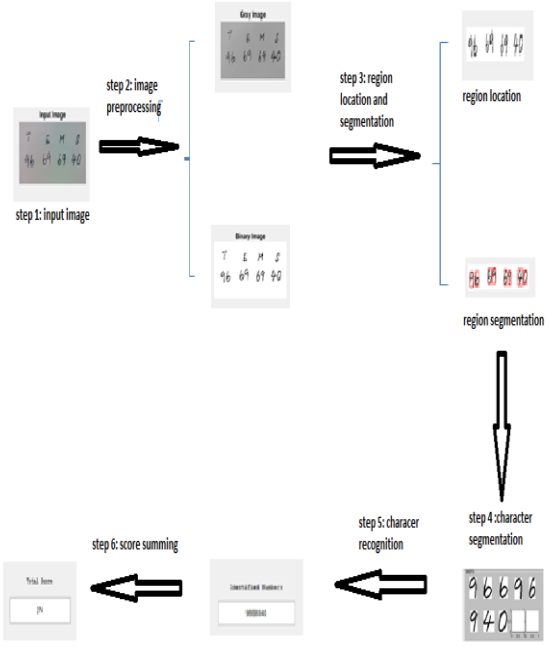

Figure 7: Flow chart for the realization of the system: single student

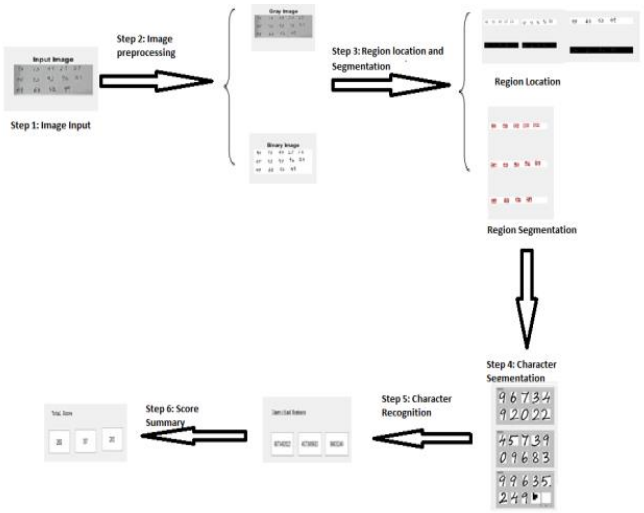

Figure 8: Flow chart for the realization of a system: multiple students.

Figure 7 and 8 describes that :

Step 1: Inputting the marks as image format into the system.

Step 2: Image preprocessing method takes place with the conversion of binary level image into grey level image by sampling theorem.

Step 3: Region localization takes place with the segmentation of appropriate location of handwritten scores which has to be taken into output.

Step 4: Separate the character into several single segments from the region localization.

Step 5: Extracting the features and recognizing the numbers from the previous steps.

Step 6: Finally summing all the scores and output the results.

III. CONCLUSIONS

In this paper, through the study of the cumulating system and the final realization process of the accumulation system for multiple student marks, the workload of teachers is reduced. The accuracy of the outcomes of the cumulating system is high because the reasonable image processing techniques had been used, inclusive of characters division, area normalization strategies, and neural network. At the same time the experimental software is highly simple with excessive-efficiency under certain requirements and time consuming is less. Proposed system of this project is various types of character will be analyzed for output scores.

IV. REFERENCES

[1].[Online].Available:http://wenku.baidu.com/view/e25ddf77e009581b6bd9ebbd.html? From=search (In Chinese)

[2]. Zhou Xiaodong, Li Yajie, Yue Xiaofeng, et al. (2010) Study on the Method of Automobile Instrument Pointer Recognition [J]. *Journal of Changchun University of Technology*, vol. 31, no. 1. pp. 32 ~ 36 (In Chinese)

[3][Online].Available:http://wenku.baidu.com/view/494b9d29bd64783e09122bec.html (In Chinese)

[4][Online].Available:http://wenku.baidu.com/view/3856f3adaef8941ea76e05a1.html (in Chinese)

[5]. Grasberg S. Nonlinear neural networks: principles, mechanisms, and architectures, Neural Network,1

[6]. Chua L.,Yang L. (1988) Cellular neural network: Theory and Applications. *IEEE Transactions on Circuits Systems I*, vol. 35, no. 10. pp. 1273 – 1290

[7]. Jing G, Du W, Guo Y. (2012) Studies on prediction of separation percent in electrodialysis process via BP neural separation and improved BP algorithm. Desalination,vol. 291, no. 2. pp. 78-93.

[8]. MAO Qing-hua. (2003) Application of Artificial Neural Network in Aerospace Experiment Effectiveness Evaluation. Automation of Theory & Applications