INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 2)
Available online at: www.ijariit.com

Cost effective optical mark recognition software for educational institutions

Vidisha Ware
vidisha.w@gmail.com
Usha Mittal Institute of Technology, Mumbai, Maharashtra

Nithya Menon
nithyamenon3@gmail.com
Usha Mittal Institute of Technology, Mumbai, Maharashtra

Prajakti Varute
prajwarute@gmail.com
Usha Mittal Institute of Technology, Mumbai, Maharashtra

Rachana Dhannawat
rachanadhannawat82@gmail.com
Usha Mittal Institute of Technology, Mumbai, Maharashtra

ABSTRACT

Optical Mark Recognition (OMR) is a technology for effectively extracting data from filled-in fields or bubbles on printed forms. The current systems available for OMR are very expensive and they detect only a marking scheme. Moreover, the image processing techniques used for scanning the OMR sheet also consumes a lot of time and is quite complex, as it includes various restrictions related to the positioning of the sheet. In this paper, a solution to this problem is proposed, where an OMR system is developed using a scanner or a multifunctional printer as an input. The quality of the OMR sheet used in this system is low cost and easily available to any educational institution. The image processing techniques are implemented with the help of PyCharm IDE that not only helps to detect various marking schemes like bubble shape mark and tick mark but also verifies the answers in the sheet and displays the total marks obtained by the student, in a more efficient manner. In order to make the system user-friendly, the GUI of the system is improved and personalized by integrating an online website with the OMR software that displays the results of the individual student.

Keywords— *GUI, Image processing, Low cost, Optical Mark Recognition (OMR), Personalized, PyCharm IDE, User-friendly*

1. INTRODUCTION

Nowadays, the usage of OMR technology is increasing tremendously with the various tests conducted on a large scale, for students. With an increasing need to obtain and manage information from hand-filled forms by the public, it has become necessary to automate large scale processing of data with high accuracy and speed. Currently, the advanced version of this inception finds application in areas including examination forms, automated attendance, feedback forms, questionnaires, ballots and community survey forms. Optical mark recognition technology extracts useful data from marked fields such as fill-in fields and checkboxes very quickly and with great accuracy. The most common use of OMR is in offices, academics and research departments where large numbers of hand-filled documents must be processed such as surveys, questionnaires, exams, reply cards and ballots.

Currently, many OMR technologies are available in the market, but some system has major drawbacks. One of the main drawbacks is that it is very expensive and cannot be used in small scale enterprises like educational institutions, where it has great applications [1]. The current OMR system uses very complex hardware which increases the complexity of the system, thereby, displaying results becomes very expensive. One of the most commonly stated disadvantages is that it can process only a special kind of paper, which is quite expensive. In certain cases, there are chances that the edges are not detected properly, which can affect the accuracy of the system, hence degrading its performance.

In this paper, the OMR system is made using low-cost software and minimal hardware. The use of special paper is avoided by replacing it with a commonly available paper of 70 GSM which is used as the answer sheet. The OMR sheet is processed using various techniques that are mentioned in the paper. The accuracy of the marking scheme is improved with other marking schemes like tick mark and cross mark, whereas, the current system supports only the circular marking scheme [2]. In order to make the system user-friendly, the GUI of the system is improved by introducing an online website which will be used to upload, store and display the evaluated marks of every student.

2. PROPOSED SYSTEM

The proposed technique for OMR neither requires high-quality paper nor high-cost scanning machine. This is image processing-based technique which can be used in any institutes, school and in small scale industries. This system can be implemented by using a camera instead of a scanner. The proposed system is designed in such a manner that it is easy to use. The general steps of the proposed system are illustrated in figure 1.

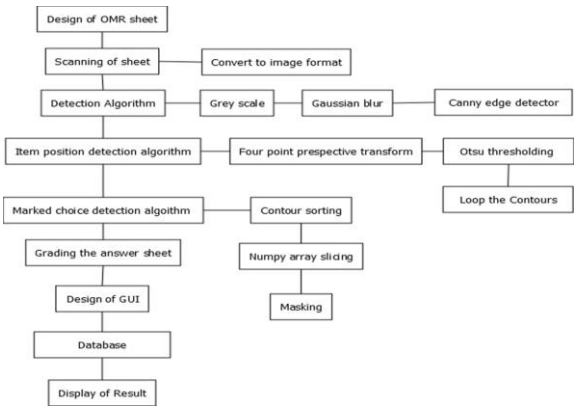

Fig. 1: Proposed system

3. SYSTEM OVERVIEW

The mentioned system consists of an input device, OMR software, an online website integrated with the software and the final output is the result for every student, which will be available online. The input device is a scanner or any multi-functional printer that is easily available in any educational institution. Furthermore, the processing software used for the scanning purposes and calculating the results of the respective student would require a PC computer and the website will be made available to every institution and its students, accordingly. There is no complex hardware required for the system.

3.1 OMR sheet template

Generally, the answer sheets of OMR are printed by special printing machine which requires a special kind of paper that is very costly and not easily available, Moreover, the basic image processing techniques applied until now required the ratio of black pixels to white pixels which could only be detected by the three position detection pattern algorithm [3]. In this paper, the answer sheet is more flexible and robust, as no special paper is used. Instead, a normal A4 sized paper with 70GSM thickness, which is easily available in the market and is quite cheap. The template for the answer sheet can be according to the requirements of the institution. A scanner is used to scan the answer sheet and stored in the format of an image without applying any algorithms. This image will further be used for image processing techniques, directly. The design of the OMR sheet is shown in figure 2.

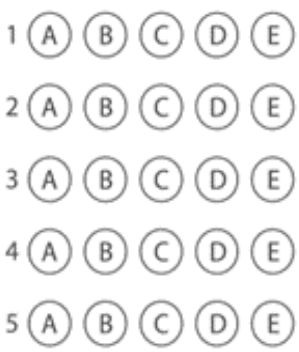

Fig. 2: The layout of the OMR sheet

3.2 The detection algorithm

The initial step is to store and save the answers to the questions in the code in a given format. The answer key provides the integer mapping of the numbers to the index of the correct bubble shown in Fig 3. After the image of the answer sheet is detected, grayscale is applied to the image followed by adding Gaussian blur [4] to the grayscale image. To the above image, the Canny edge detector is applied in order to detect the edges/outlines of the marked choices. In order to obtain the entire outline of the exam in the image, we detect all the contours and sort them to obtain the whole answered part as the main focal point of the image.

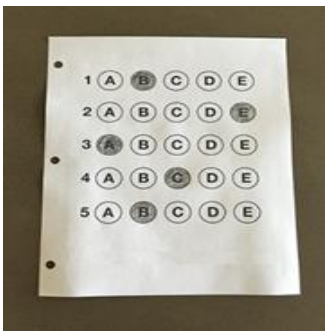

Fig. 3: The scanned image

3.3 The item position detection algorithm

The next step is to apply a Four Point Perspective transform to obtain a top-down, birds-eye-view of the document since the outline has been detected. This is followed by the application of a mixture of Otsu's binary [5] and adaptive thresholding that will separate the foreground of the image from the background of the image.

This binarization will allow us to once again apply contour extraction techniques to find each of the bubbles in the answer sheet [6]. The process is followed by initializing a list of contours, which correspond to every question. In order to detect the images of bubbles, we apply loop over each of the individual contours. The aspect ratio of each contour will help in recognizing whether a contour is a bubble or not [7].

3.4 Marked choice detection

In the next stage, the contours are sorted from top to bottom and then NumPy array slicing is added along with contour sorting to sort the current set of contours from left to right [8]. This helps in identification of the bubbles corresponding to each question. Using the thresh image, a mask is constructed and counting the number of non-zero pixels in the masked region, we obtain which bubble is marked. This enables us to detect any type of marking scheme [9].

Fig. 4: The final output

3.5 Grading answer sheet

Finally, a comparison is made between the bubbled answer of each question and the option mentioned in the answer key

initially in the code. Then the result will be displayed which will show the total marks obtained by the student.

3.6 Design of GUI

A website has been designed to display the results to the students using HTML and CSS. There is a 'Homepage' which contains a navigation bar to the admin and student login pages. The admins must register themselves using the 'Admin Registration Form'. The design of the Admin page is illustrated in figure 7. The teachers will act as the admin and upload the answer key and students answer sheets and the marks will be evaluated by the system. The students also must register using the 'Student Registration Form' shown in figure 5. After logging in, the student can view their individual scores on the website as shown in figure. 6. This will help the students to view their answer sheet from anywhere in the world and complete privacy will be given to everyone in terms of their scores. In case of any issues in their answer sheets, the students can report it on the 'Contact Us Page'.

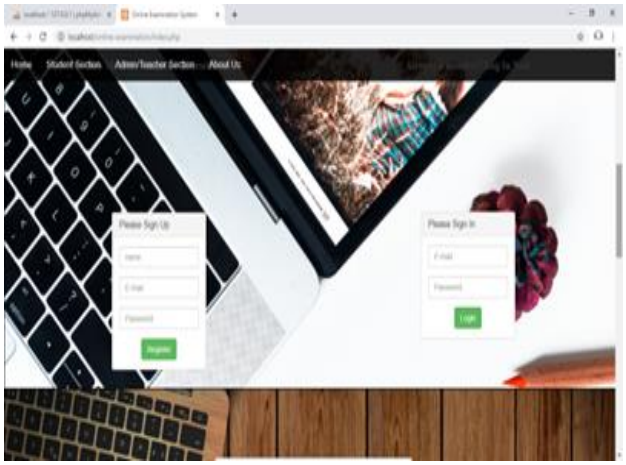

Fig. 5: Student login page

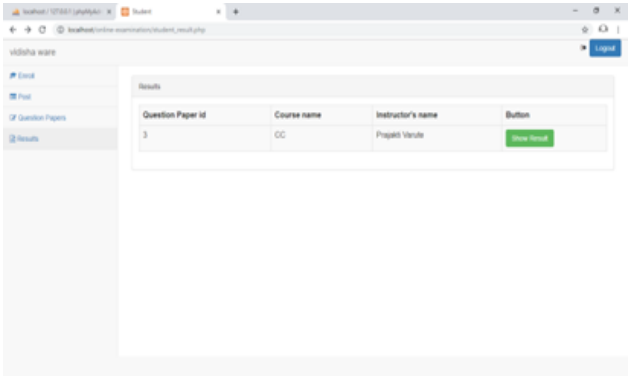

Fig .6: Student result page

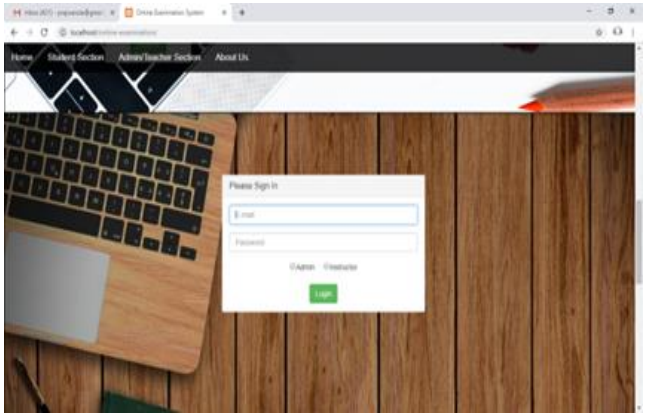

Fig. 7. The admin login page

3.7 Database

In the backend of the website, two separate databases have been created using MySQL: One for the admin and another for students. The admin database is connected with the 'Admin Login Page' using Php. Similarly, the student's information (user id and password) is also stored in the student database and it is integrated with the 'Student Login Page' using Php.

3.8 Integration

The OMR software is prepared using Python programming language and OpenCV [10] library which have been integrated using the PyCharm IDE environment. After the application of image processing techniques, the OMR software is integrated with the website, where the output answer sheet of the software will be uploaded on the website to every student's page.

4. EXPERIMENTAL RESULTS

In order to test the performance of the system, a sample answer sheet was scanned, and the image of the paper is taken as input. Later, the image is processed using the various techniques mentioned and the result is displayed. The resultant image obtained can not only detect circular markings but also other marking methods like tick mark and cross. This has been done using Python programming language and OpenCV library which was integrated into PyCharm.

Another module included a website that is designed to provide a user-friendly environment for storing, uploading and displaying the results of every student belonging to the institution. The admins/faculty uploads the respective scores obtained by each student. The students can view their scores, by logging on to the portal through the username and password provided to them [11].

5. CONCLUSION

In this paper, a cost-effective solution to the OMR software has been proposed. It tries to overcome various drawbacks of the traditional OMR systems. An OMR layout on normal A4 size paper (i.e. 70 GSM thickness) has been designed instead of 120-130 GSM paper which decreases the cost considerably. For scanning purpose, ordinary scanner is used. This system checks a different marking style. Both pen and pencil marks are detected properly. A user-friendly website is created that will save the time of the teachers instead of manually correcting answer sheets. Each student can access their respective answer sheet online. The main limitation is that the questionnaire used is static in nature. In future, the questionnaire can be made dynamic for improved results. Latest technologies such as Cloud Computing and Machine Learning can be integrated and utilized to make the software more adaptive and user-friendly.

6. REFERENCES

[1] Astha Gupta, Sandhya Avasthi (7, July 2016) Image-based low-cost method to the OMR process for surveys and research, International Journal of Scientific Engineering and Applied Science (IJSEAS) – Volume-2: www.ijseas.com.

[2] P. Sanguansat, Robust and low-cost Optical Mark Recognition for automated data entry, 2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTICON), 2015.

[3] R. S, K. Atal, and A. Arora Cost Effective Optical Mark Reader, International Journal of Computer Science and Artificial Intelligence, vol. 3, no. 2, pp. 4449, Jul. 2013.

[4] Y. Sri Sai Saran Reddy, A.Sai Srinivas, G. Mohith Krishna (4 April 2018) OMR Evaluation using Image Processing,

International Journal of Innovations & Advancement in Computer Science IJIACS ISSN 2347 – 8616.

[5] N. Otsu, A Threshold Selection Method from Gray-Level Histograms, IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 6266, 1979.

[6] K. Chinnasarn and Y. Rangsanseri, Applications of Digital Image Processing XXII, 1999.

[7] I. A. Hameed, A Fuzzy System to Automatically Evaluate and Improve Fairness of Multiple-Choice Questions (MCQs) based Exams, Proceedings of the 8th International Conference on Computer Supported Education, 2016.

[8] H. Deng, F. Wang, and B. Liang, A Low-Cost OMR Solution for Educational Applications, 2008 IEEE International Symposium on Parallel and Distributed Processing with Applications, 2008.

[9] E. Greenfield, OMR Scanners: Reflective Technology Makes the Difference, Technological Horizons In Education, vol. 18, 1991.

[10] G. Bradski, The OpenCV Library, Dr Dobb's Journal of Software Tools, 2000.

[11] Ramesh, S, (2005). Exploring the Potential of Multiple Choice Questions in Computer-Based Assessment of Student Learning.pattern recognition letters, VOL.6, NO.7, PP(199-213).

