



Automatic Mark Entry System Using Computer Vision-based Approach

Asraf Mohamed Moubark, Mohd Hairi Mohd Zaman, Mohd Asyraf Zulkifley, Sawal Hamid Md Ali
 Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment,
 Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
asrafmohamed@ukm.edu.my, hairizaman@ukm.edu.my, asyraf.zulkifley@ukm.edu.my, sawal@ukm.edu.my

ABSTRACT

Transferring the student marks from the answer scripts to the grade template file, typically constructed using Microsoft Excel, is a challenge to the lecturers when there are too many students in the course. In practice, the lecturers manually enter the student marks the students to the Microsoft Excel grade template file. The manual process not only takes a long time but also allows humans errors to occur. A new system based on called optical marks recognition (OMR) was introduced in this work to ease the burden of lecturers and minimize human errors. With this new system, the matric number of students and marks, shaded in the answer scripts, will be identified and obtained through a low-cost webcam camera processed using Matlab software. Before that, each student will paste an OMR sticker at the top right corner of the answer script, which consists of boxes of matric number and mark, that are given to them during exams. The students only need to shade the matric number boxes. After that, the lecturers need to darken the marks boxes after reviewing the answer scripts. Next, the lecturers placed the marked answer scripts below the webcam to start the scanning process. After Matlab software identified the matric number matrix and marks of particular students through image processing techniques, Matlab will open the Microsoft Excel grade template file and compare the matric number available in the template with the scanned matric number. After founding the matric number, Matlab copied the marks to the Microsoft Excel template. The entire process of scanning and automatic mark entry just takes about 5 seconds for each answer script. As such, the system in this work is not only able to save lecturer time but also can improve the accuracy in inserting the student marks to the grade template file.

Key words : OMR, Image Processing, Marks, Webcam

1. INTRODUCTION

The responsibilities of the lecturers are getting increased every day. Lecturers are not only responsible for teaching students at undergraduate and postgraduate levels, but also have to supervise research mode students. Besides, other requirements assigned to the lecturers is to conduct

world-class research and publish the research outcomes in high impact journals. At the same time, lecturers also become advisers to the students who participated in the competition and managing committee of academic curriculum accreditation, and other management tasks.

One of the ways that are effective for evaluating student understanding is through examinations. In general, there are four stages of implementing controlled or summative exams and tests. The first process is to prepare the examination questions. Then, the second stage is to carry out the examination in a controlled environment. After that, the third stage is to assess and give marks to the answer scripts, and the final step is the mark entry to the grade template file, which is typically built using Microsoft Excel. The first three stages in carrying out the examination processes depend a lot on the effort of lecturers. The stages of preparing questions to assess the student answer scripts are unique and cannot be automated for the time being. These processes require full concentration and commitment from the lecturers. However, the last stage to enter the marks into the grade template file is a repetitive process, and it has a high potential for automation to reduce the burden of the lecturers.

In current practice, the lecturers will write the number of marks on the front page of answer scripts by using a pen and later will enter those marks to the grade template file. Since this is a repetitive process, then the chance for human error happens quite high. Typically, the lecturer will use the 'find' function in Microsoft Excel to find the matric number of each student and then enter the marks in the appeared spaces in Microsoft Excel. The lecturers will repeat this mark entry process for all the students who attend the exam. Human error is most likely to occur if the lecturer mistakenly types the student matric number in the 'find' field in Microsoft Excel. As a consequence, student marks could also be mistaken because of the wrong matric number.

Optical mark recognition (OMR) form has been long used for examination at the school and university levels for multiple choice question type. The current method to identify whether an area in the OMR form has been shaded or not is to use the light emission technique [1]-[3]. In this technique, the system will emit the light on the OMR form and then measures the

amount of reflected light. The area shaded by carbon pencil will absorb more light compared with the unshaded areas. Therefore, student choice can be identified and compared with the available answer template. Another frequently used method in the OMR system is to use a special paper called transoptic paper [4]. The light will emit through the transoptic paper and not reflected like regular OMR paper. Then, the amount of light that is received through the paper will be measured, and if the area was shaded, the light would not go through it. For both techniques, a special machine needs to scan the OMR form. At the same time, the OMR form must be in line with the machine without any deflection. The conventional mechanical scanning of OMR form is also subject to the possibility of errors occurring as much as 2% [5].

To overcome the problem in processing the OMR form, [6-8] used camera and webcam to obtain the information from the OMR form. However, the methods that they adopted is subject to the initial condition or position the OMR form. A special algorithm to correct the distorted OMR form needs to be used before it can be processed[9]-[13]. As such, to minimize human error and reduce lecturer workload, a new OMR system is introduced in this work. The proposed OMR system uses a new algorithm to increase the efficiency in identifying the shaded areas. This proposed method is also not subject to the initial position of OMR form, which must be fixed in advance in other studies, during the scanning process [14]-[17]. The available OMR form in the market has a large size and requires a machine typical to read the information in the OMR form [18]-[20]. This project designed a unique small size OMR form printed on paper stickers for making it easier for the students to paste it anywhere, whether on the standard final exam answer script, mid-semester exams, quiz, or their project reports.

2. MATERIAL

This project uses a custom-designed OMR form with a size of 10 cm x 7.5 cm, as shown in Figure 1. The OMR form consists of two crucial information, which is the student matric number and marks given to that particular student.

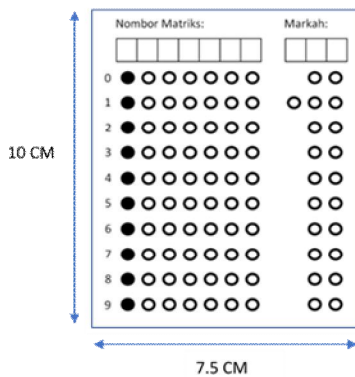
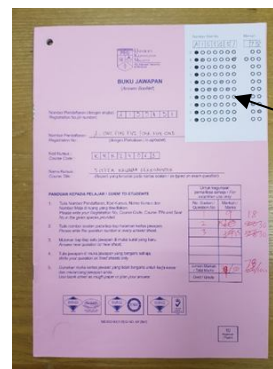


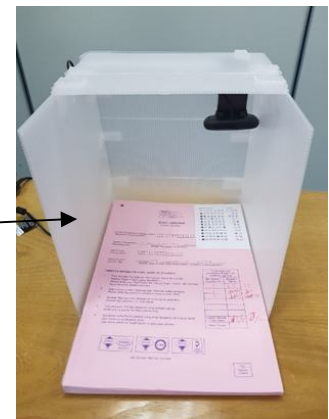
Figure 1: The specially designed OMR form used in this project.

Students are responsible for shading the blank circles that represent each alphabet and numbers associated with their matric number. Meanwhile, the lecturer is only necessary to darken the circle, which represents the total marks obtained by the student. Students and lecturers are encouraged to use 2B grade pencils only for shading the OMR form. Figure 2 (a) shows the OMR forms that have been pasted and shaded.

The process of automatic mark entry to the grade template file in Microsoft Excel format begins when the lecturer opens Matlab software and put the answer scripts under the webcam. Next, the lecturers only need to put the answer script with pasted OMR form one by one into the special box special with the webcam for the scanning purpose. Figure 2 (b) shows the answer script is placed in the custom-designed box with a webcam at the top. Webcam model selected for the project is Logitech with a resolution of 768 x 1024 (length x width). Matlab software version 2019b has been used to process the video acquired using the webcam.



(a)



(b)

Figure 2: a) The specially designed OMR sticker is placed on the answer script. b) After shading the matric number and marks, the OMR form will be placed under the webcam located on top of the box

3. METHOD

A webcam recorded the video of the OMR form and sent the video data to Matlab. Next, the proposed system converted the file format from RGB to grayscale to reduce computational calculations. After that, the algorithm converted the data format from grayscale to binary format to identify the number of circles that exist in the OMR form. The Otsu's adjustment algorithm was adopted to find the threshold so that the grayscale can be classified to 0 and 1 with more efficiency. This is because the light and shadow that fell on the OMR form can affect the results of the grayscale to binary classification. If all circles in the OMR form cannot be identified, the threshold points were continuously calculated until achieving the total amount. Guided by the recognized center and radius of the circle, all circle areas will be calculated, whether it is shaded or not. If the circle is shaded, then the center of the circle was stored, sorted, and classified to identify the matric number and student marks. Figure 3 shows the flowchart of all methods in the automatic mark entry system. Meanwhile, Figure 4 shows the process involved in identifying and extracting the matric number and student marks in the form of images.

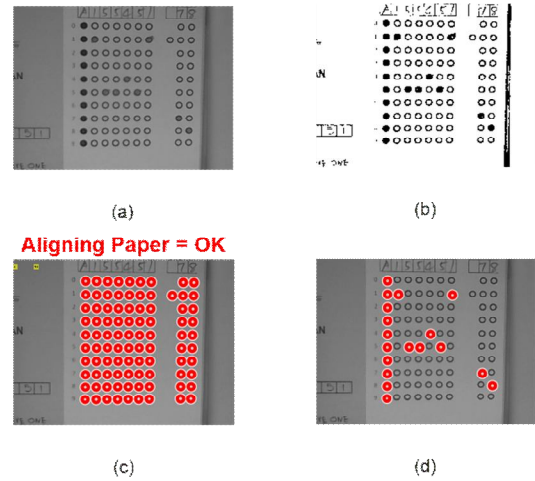


Figure 4: a) Grayscale image of OMR form obtained from the webcam. b) Image binary of OMR form after going through the process of Otsu's adaptation algorithm. c) Each identified circle on the OMR form is mapped on the grayscale image. The message "Aligning paper = OK" will appear if all circles are identified. d) All circles that represent the matric number and student marks are separated and classified

4. RESULTS AND DISCUSSION

After the process of identifying the matric number and student marks was completed, then the results were displayed in the Matlab command window, as shown in Figure 5. At the same time, the instructions to put the answer scripts into the box was displayed. The answer script that has been scanned and processed before does not need to be taken out. The next answer script to be scanned can be directly placed on top of the previous answer script.

```

Enter the Total Number of Papers : 5
Done
-----
matrix      marks
152875      73
-----
Please Place a New Paper
    
```

Figure 5: Detected matric number matrix and student marks are listed in the Matlab command window for lecturer reference

After completion of the scanning process, Matlab opened the Microsoft Excel template file provided with the student information of the particular course. The student information was obtained from the university website. Matlab compared the matric number in Microsoft Excel with the scanned matric numbers and then enter the marks associated with the identified student using the image processing technique. Figure 6 shows the grade template file with the final mark for each student corresponding to their matric number.

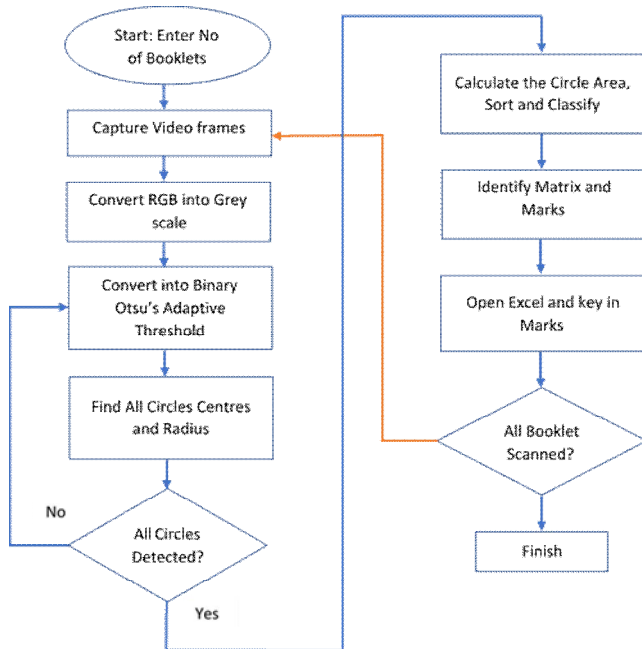


Figure 3: Flowchart of the automatic mark entry system

No	No. matric	Nama pelajar	Sahaja	Prolog	Kul	Prolog	Matkul	Asesmen	Misbah	Nilai
1	A152876	NURUL AINI BINTI MAT PAUZI	4	KZ						73
2	A153780	AHAM AFIF BIN SHARUDIN	4	KZ						60
3	A153872	JEFFRY UNYANG ANAK SELI	4	KZ						51
4	A153957	MUKHAMMAD ZUL HAZID BIN ABDUL KARIM	4	KZ						63
5	A155451	MOHD RIDHWAN BIN ABD GHAFAR	4	KZ						78
PURATA SISIHAN PIAWAI MAKSIMUM MINIMUM BILANGAN DATA YANG DIMPUTKAN										83.00 9.57 78.00 51.00 5

Figure 6: The marks for each student are automatically inserted into the grade template file in Microsoft Excel format after identifying the matric number

A total of five sets of student answer scripts were used as a test material to check the feasibility of the proposed method. All five answer scripts are successfully scanned 50 times without any mistakes where each matric number and its corresponding marks can be accurately and efficiently identified.

5. CONCLUSION

The developed automatic mark entry system, through a vision-based approach, able to reduce the lecturer workload and to minimize human error. Although the system can run smoothly without any issues, the system is still in the initial prototype phase. It still needs thorough testing to identify any potential problems that may occur. This can be achieved when the lecturers began to use the system for a number of courses. In the future, the input from the lecturers will be considered for improving system functionality and effectiveness.

ACKNOWLEDGEMENT

Part of this work was financially supported by University Kebangsaan Malaysia (Grant No.GGPM-2019-007). Funders were not involved in the conduct of the research.

REFERENCES

1. Gaikwad, Sumitra B. **Image processing based OMR sheet scanning.** *International Journal of Advanced Research in Electronics and Communication Engineering*, vol. 4, no. 3, pp. 519-522, March 2015.
2. Sanguansat, Parinya. **Robust and low-cost optical mark recognition for automated data entry.** *International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, 2015, pp. 1-5.
3. Gupta, Astha, and Sandhya Avasthi. **Image based low cost method to the OMR process for surveys and research.** *International Journal of Scientific Engineering and Applied Science*, Vol. 2, no. 7, pp. 91-95, July 2016.
4. Patel, Niral V., Ghanshyam I. Prajapati, and Bharuch SVMIT, **Various techniques for assessment of OMR sheets through ordinary 2D scanner: a**

5. survey, *International Journal of Engineering Research and technology*, Vol. 4, no. 9, pp. 803-807, Sept. 2015. <https://doi.org/10.17577/IJERTV4IS090675>
6. Ferreira, Fernando JTE, and Fernando JP Lopes. **Webcam-based tachometer for in-field induction motor load estimation.** *IEEE International Conference on Electrical Machines*, 2016, pp. 2380-2388.
7. Hussmann, Stephan, Leona Chan, Celine Fung, and Martin Albrecht. **Low-cost and high-speed optical mark reader based on an intelligent line camera.** *International Society for Optics and Photonics*, vol. 5106, pp. 200-208, Aug. 2003.
8. Nguyen, Tien Dzung, Quyet Hoang Manh, Phuong Bui Minh, Long Nguyen Thanh, and Thang Manh Hoang. **Efficient and reliable camera based multiple-choice test grading system.** *IEEE International Conference on Advanced Technologies for Communications*, 2011, pp. 268-271.
9. Naz, Saeeda, Khizar Hayat, Muhammad Imran Razzak, Muhammad Waqas Anwar, Sajjad A. Madani, and Samee U. Khan. **The optical character recognition of Urdu-like cursive scripts.** *Pattern Recognition*, Vol. 47, no. 3, pp. 1229-1248, 2014.
10. Pérez-Benedito, José Luis, Enrique Querol Aragón, Juan Alonso Alriols, and Ljiljana Medic. **Optical mark recognition in student continuous assessment.** *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, Vol. 9, no. 4, pp. 133-138, Nov. 2014. <https://doi.org/10.1109/RITA.2014.2363005>
11. Mehta, Honey, Sanjay Singla, and Aarti Mahajan. **Optical character recognition (OCR) system for Roman script & English language using Artificial Neural Network (ANN) classifier.** in *2016 International Conference on Research Advances in Integrated Navigation Systems*, 2016, pp. 1-5.
12. Ribeiro, Fabio De Sousa, Liyun Gong, Francesco Calivá, Mark Swainson, Kjartan Gudmundsson, Miao Yu, Georgios Leontidis, Xujiang Ye, and Stefanos Kollias. **An end-to-end deep neural architecture for optical character verification and recognition in retail food packaging.** in *25th IEEE International Conference on Image Processing*, 2018, pp. 2376-2380.
13. Haskins, Bertram. **Contrasting classifiers for software-based OMR responses.** in *Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference*, 2015, pp. 233-238.
14. Ju, Yunxia, Xichang Wang, and Xiangxi Chen. **Research on OMR Recognition Based on Convolutional Neural Network Tensorflow Platform.** in *11th International Conference on Measuring Technology and Mechatronics Automation*, 2019, pp. 688-691.
15. Ali, Asghar, Mark Pickering, and Kamran Shafi. **Urdu natural scene character recognition using convolutional neural networks.** in *IEEE 2nd international workshop on Arabic and derived script analysis and recognition*, 2018, pp. 29-34.

15. Din, Israr Ud, Imran Siddiqi, Shehzad Khalid, and Tahir Azam. **Segmentation-free optical character recognition for printed Urdu text**, *EURASIP Journal on Image and Video Processing* 2017, Vol. 62, no. 1, pp. 1-18, Sep. 2017.
16. Sharma, Vandini, Prakash C. Mathpal, and Akanksha Kaushik. **Automatic license plate recognition using optical character recognition and template matching on yellow color license plate**, *International Journal of Innovative Research in Science, Engineering and Technology*, Vol. 3, no. 5, pp. 12984-12990, May 2014.
17. Cong, L. I. **Artificial Mark Recognition in Augmented Reality Based on FPGA**, *Electronic Science and Technology*, Vol. 27, no. 5, pp. 60-65, May 2014.
18. Bautista, Rey Mark John SA, Vishnu Joshua L. Navata, Aldrich H. Ng, Ma Timothy S. Santos, Justine D. Albao, and Edison A. Roxas. **Recognition of handwritten alphanumeric characters using Projection Histogram and Support Vector Machine**, in *International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management*, 2015, pp. 1-6.
<https://doi.org/10.1109/HNICEM.2015.7393256>
19. Soomro, W.J., Ismaili, I.A. and Shoro, G.M. **Optical Character Recognition System for Sindhi Text: A Survey**. *University of Sindh Journal of Information and Communication Technology*, Vol. 2, no. 2, pp.81-87, April 2018.
20. Kalina, Darina, and Roman Golovanov. **Application of Template Matching for Optical Character Recognition**, in *IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering*, 2019, pp. 2213-2217.