

Automatic Grading of Scanned Multiple Choice Answer Sheets

Jirapond Muangprathub^{1*}, Oatchima Shichim¹, Yuthaya Jaroensuk¹, Siriwan Kajornkasirat¹

¹Faculty of Science and Industrial Technology, Prince of Songkla University, Suratthani Campus, Suratthani 84000, Thailand

*Corresponding author E-mail: jirapond.m@psu.ac.th

Abstract

The motivations for automatic grading of image-based multiple choice answer sheets include significant time and cost reductions. The proposed method supports any pencils or pens used on thin papers, as well as low-cost gridded paper that is easy to use in a typical test. Fourteen different scenarios pertaining to 560 answer sheets were evaluated with automatic reporting of the final grading and its summary. The result shows that an average accuracy of 100% for the cases with nearly complete pencil or pen markings. In the cases with incomplete markings, such as small markings, overflow, and deleted or unclear markings, the accuracies were 62.42%, 93.16%, 99.57%, respectively. The proposed system operates 2.5 times faster than the conventional manual method.

Keywords: Optical mark reading; OMR; Automatic grading; Correlation coefficient; Multiple choice answer sheets.

1. Introduction

Currently, there are numerous utilization of technologies, whether in the public or private institutions, to shorten the time required for manually marking exams while simultaneously increasing the accuracy and reducing the paper consumption and human errors [1-5]. Printed multiple-choice tests are widely used in education, surveys, or other tests. Consequently, the repetitive task of comparing the answer sheet to the key and score the responses is a common laborious task that should be automated. Nowadays optical mark recognition, also called optical mark reading (OMR), is often employed.

OMR [1-2,4,6-8] is a process for the detection and recognition of marks written on paper that is applicable to the responses to multiple-choice tests. This techniques helps reduce the time an assessor must spend, but requires the use of a formatted answer sheet where the response markings are predictably positioned. However, errors may be incurred when a participant answers with more than one choice or used otherwise ambiguous markings [4,9-11]. These errors are not frequent in the automatic machine marker of this paper, which utilizes template matching technique.

Template matching [12-14] is a technique in digital image processing for finding a target object that matches the given template's pattern. In such technique, the answer image and the answer sheet are compared to verify the right answers by using correlation coefficients. In this paper, different features on both the master solution template and the answer sheet are identified. The reference datum can be implemented by a straight line from left to right or top to bottom, to reduce the tilt error of the answer sheet, and the similarities are then calculated.

This article is organized as follows. Section 2 provides the background in image processing, template matching, and traditional OMR. In Section 3, we present the proposed system. Section 4 shows the experimental results, and Section 6 concludes the article.

2. Background

2.1. Image processing

This section describes image processing of a grid for multiple choice answers or solutions. The answer sheet is divided into two parts, namely identification of the examinee (header part), and the answers part. In this work, the gridded answer sheet design was as shown in Figure 1.

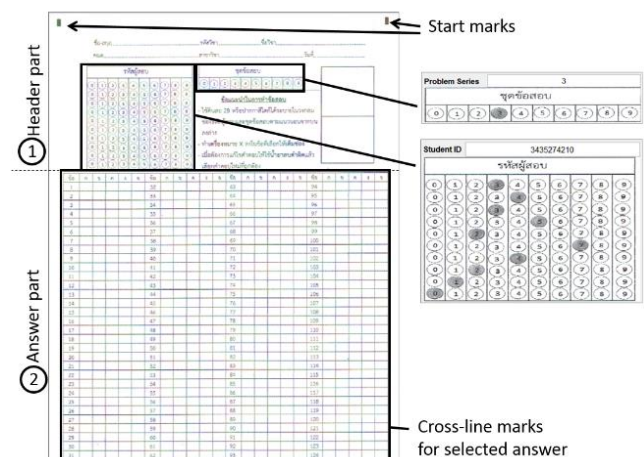


Fig. 1: The gridded answer sheet design

In Figure 1, the designed answer sheet started with 2 black boxes to enable correcting for tilt of the scanned paper (Start marks). The first part contains the identification of examinee including student ID, number of the problem series, and other details to be filled in by the examinee. The examinee needs to fill circle patterns in this part. The second part is the grid in which each square or rectangle can accept a cross marking for the selected alternatives of each multiple choice item.

The image processing starts by scanning the answer sheet. The first operation performed to the image is conversion to gray scale, followed by binarization to black and white. Then, the examinee's ID and the number of problem series are detected. The answer part image is extracted for the processing described in Section 3.

2.2 Template matching

The above original designs are transformed individually into their corresponding generalized chains (kinematic chains). The generalized chain will be involved in various types of members (edges) and joints (vertices, or said kinematic pairs) for possible assembly in the following steps.

Template matching [6,12-15] is a Character Recognition technique. It is the process of finding the location of a sub image, called a template, inside a larger image. Once a number of corresponding templates is found, their centers are used as corresponding points. The window that yields the best similarity score is selected as the match location. The two variables are the corresponding pixel values in two images, answer and solution sheet. The scoring is with correlation coefficient

$$Cor = \frac{\sum_{i=0}^{N-1} (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=0}^{N-1} (x_i - \bar{x})^2 \cdot \sum_{i=0}^{N-1} (y_i - \bar{y})^2}}$$

where x is the solution sheet gray level image

\bar{x} is the average grey level in the solution sheet image

y is the answer sheet image section

\bar{y} is the average grey level in the answer sheet image

N is the number of pixels in the section image (size of the answer sheet image)

The value Cor is between -1 and $+1$, with larger values representing stronger relationship between the two images. This formula is applied to calculate matching between answer and solution sheet for each problem.

2.3 The traditional OMR

Image-based OMR has been widely developed and studied. In [2], the authors presented a system which uses techniques from mathematical morphology, capable of acquiring images from computer cameras to perform automated scoring of a multiple-choice test with highly accurate results. The steps performed by the algorithm were explained in detail and illustrated by its employment in a real case scenario. The software is currently being ported to smartphones, which will allow teachers to travel to test centers carrying only printed tests with no need for additional resources, apply the tests, capture images, and calculate scores in loco, being able to discuss the test results with the students on the same day.

Sattayakawee [9] proposes an algorithm for test scoring using non-optical traditional grid answer sheets [16]. The algorithm is based on projection profile and thresholding methods. The main distinctions from other OMRs for test grading are the answer sheet forms which are much harder to process and the detection of 2 or no selected choices. The average accuracy was 99.909%. The actual error came from an abnormal input, rather than the algorithm itself. This algorithm can also be applied in other types of answer sheets in grid form.

Spadaccini et al. presented JECT-OMR, a recognition system for

multiple-choice tests based on the Gamera framework [16]. Using the Gamera framework helped us develop the application in a very short time, and, due to its intuitive use, we could focus on the application domain rather than on programming techniques. The results of the tests show that the application is quite mature, and it has been used for more than two years by a small enterprise whose mission is to prepare students for multiple-choice tests used in the admission tests of numerous clausus University courses. In [10, 17], the author proposed an automatic system to grade multiple choice questions to detect the correct answers by comparing each paper with a pre-scanned test paper that contains the correct answers, as many forms of test papers are used in real exams conducted in the computer center of the Baghdad University. The results were shown to match results from manual grading of the same papers.

3. The proposed method

3.1. Overview of automatic grading system

The motivations for automatic image-based grading of multiple choice answer sheets include significant time and cost reductions. The proposed system applied the correlation coefficient to compare answer sheets and solution sheet images. The proposed method supports any pencils and pens for marking the answers, and also supports thin or low-cost gridded paper as answer sheets, which are easy to use in a general test. An overview of steps in the developed system is shown in Figure 2.

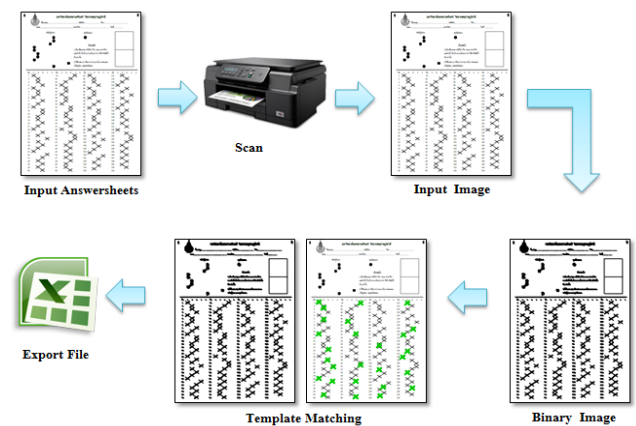


Fig. 2: An overview of the proposed system

As inputs, the answer and the solution sheets are scanned with a scanner device. The images are converted to gray scale and binarized. The header part of an answer sheet is processed while the answer part is extracted for later matching of answers to solutions by using the correlation coefficient. The header part was designed with bubbles to fill, to provide the examinee's ID and the number of problem series. This part is preprocessed with noise reduction and skew detection.

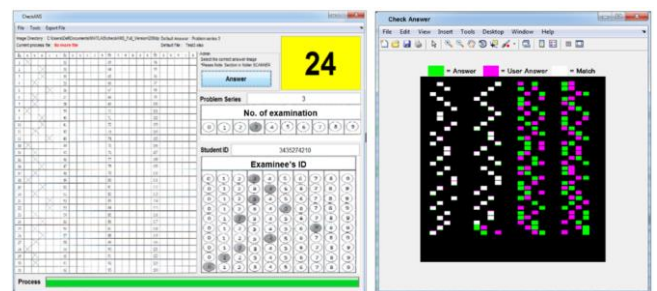


Fig. 3: An example window for automatic grading (left) and re-checked answer sheet (right)

An example of the implemented interface is seen in Figure 3. It shows the answer sheet and the score for the user identified in the left figure. The right figure shows rechecked answer and solution sheets. Moreover, our system can process for all answer sheets to obtain the output scores for all examinees, shown in Figure 4. There is also a statistical summary of the scores, i.e. full score, mean score, minimum and maximum scores, and standard deviation. The results can be exported as an excel file for later use.

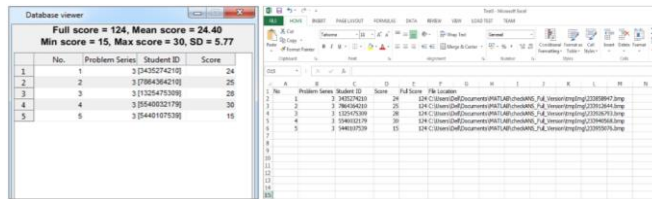


Fig. 4: Summary of scores for several examinees

3.2. The methodology of our system

The main process of the developed system has three parts, i.e., checking the number of problem series, identifying examinee's ID, and checking the answer part. Each part is now described in some detail.

1) Checking the number of problem series

The input answer and solution sheet images are first binarized. The binary image is enhanced for clarity by filling within the bordering of a symbol. Afterwards, the segmentation of the image will be processed to determine the crosshairs. The answer and solution sheet images are compared to check the number of problem series. If the number of problem series images are similar for the answer and the solution sheet, next automatic grading is done. This step follows Algorithm 1. This algorithm provides for getting and checking the number of problem series, where the input data is answer and solution sheet. The answer sheet is first processed to get the selected number of problem series. At the same time, the solution sheet is processed for the number of problem series, to compare with the selected answer sheet. This main function call is *checkNumber()* for image processing.

2) Identifying examinee's ID

This process gets the examinee's ID on the answer sheet by Algorithm 2. First the converted image is slightly improved with filling the lost parts and removing noise. Next, the improved answer image will be cropped for only the region of interest, i.e., the examinee's ID part, and cropped for each row of the examinee's ID part. Each cropped image will be inverted to show only the selected point by eliminating small objects, to enhance image quality. To get the selected point, the centers of the circles are obtained by using their centroid property. This step returns the examinee's ID number.

Algorithm 1 : Checking the number of problem series

Input : Answer sheet image (ANSImg)
: Solution sheet image (SolImg)

Output : Next process (if return true) or
Cancel process (if return false)

Method :

1. Read ANSImg and SolImg
2. ansNo=checkNumber (ANSImg)
3. SolNo=checkNumber (SolImg)
4. if (ansNo==SolNo)
5. return true;
6. else

7. return false;

Function checkNumber (Imgsheet)

1. Convert Imgsheet to binary image
2. Fill border of binary image
3. Crop region of interest i.e. problem series part
4. Invert Imgsheet to get only selected point
5. Eliminate small object
6. Improve Imgsheet to distinct and clear
7. Get selected point
8. Check position of selected point
9. Return selected position

Algorithm 2 : Identifying examinee's ID

Input : Answer sheet image (ANSImg)

Output : examinee's ID

Method :

1. Convert Imgsheet to binary image
2. Fill border of binary image
3. Crop region of interest i.e. examinee's ID part
4. While crop row of examinee's ID part
5. Invert Imgsheet to get only selected point
6. Eliminate small object
7. Improve Imgsheet to distinct and clear
8. Get selected point
9. Check position of selected point
10. Concatenate position into ID
11. Return ID

3) Checking the answer part

The most important step is to check the answers on solution sheet and obtain the score for the examinee. Algorithm 3 is used to match the selected answer and solution. This algorithm refers to the *CrossSelect()* function to select the marked point on the solution sheet. Afterwards, the selected point from answer or solution sheet is compared for counting the score.

Algorithm 3 : checking answer correction

Input : Answer sheet image (ANSImg)
: Solution sheet image (SolImg)

Output : Number of choosing answer correction
from examinee (correctNo)

Method :

1. SelectANS[] = CrossSelect (ANSImg)
2. SelectSol[] = CrossSelect (SolImg)
3. correctNo=0
4. For i=0 to |SelectSol[]|
5. For j=0 to |SelectANS[]|
6. If SelectSol[i] == SelectANS[j]
7. correctNo = correctNo+1
8. Return correctNo

Function CrossSelect (Imgsheet)

1. Convert Imgsheet to binary image
2. Crop the grid answer part (ansPart)
3. Crop ansPart into array of colAns[] for each column of answer sheet
4. Crop ansPart into array of rowAns[] for each row (each problem answer)
5. For each colAns[]
6. For each rowAns[]
7. Check position of selected point →
SelectPoint[]
8. Return SelectPoint[]

4. Experimental results

4.1. Data Collection

The proposed system was tested to assess its accuracy in a real examination. In this experiment, the sample was students of the Prince of Songkla University, Surat Thani campus, Thailand. We used as answer sheets 560 test samples with 14 cases, and each case consisted of 40 answer sheets as in Table 1.

Table 1: Detail of data collection in experimentation.

Case Order	Detail of Condition
1	Using pencil less than 2B, 124 questions, examinee is technology food student.
2	Using pencil 2B, 124 questions, examinee is information technology student.
3	Using pencil greater 2B, 124 questions, examinee is economic business student.
4	Using pen color, 124 questions, examinee is chemistry for industrial student.
5	Answer sheets is same the solution, 30 questions, examinee is information technology student.
6	Answer mark more one choice, 30 questions, examinee is economic business student.
7	Answer mark not full in box, 70 questions, examinee is chemistry for industrial student.
8	Answer mark greater the questions in solution, 50 questions, examinee is information technology student.
9	Answer mark less than the questions in solution, 80 questions, examinee is information technology student.
10	Answer mark overflow from the box, 90 questions, examinee is economic business student.
11	Delete mark and rewrite mark again, 100 questions, examinee is information technology student.
12	Answer mark with 60 questions, examinee is economic business student.
13	Answer mark with 70 questions, examinee is public administration student.
14	Answer mark with 110 questions, examinee is information technology student.

4.2. Answer sheet patterns

This section shows answer sheet patterns in the test sample. These answer sheets were marked with crosses as in Figure 5. Moreover, our system can detect the marked cross-line pattern for two answers in one question, as in Figure 6, in which case the test item was assigned score zero.

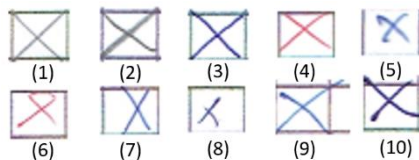


Fig. 1: An example answer sheet patterns



Fig. 2. The examinee select two answer in one question

4.3. Results

We implemented the automatic image-based grading of multiple choice answer sheets, and used the data set in Section 4.1 to test our system. We tested the 15 cases in Table 1. The results for accuracy of our test scoring system are shown in Table 2 and Figure 7.

Table 2: Accuracy of test scoring

Test Case	Accuracy of test scoring (%)	Feature of the marked cross-line pattern
1-6	100	Using any pencil, any pen, any question, and selecting two answer in one question
7	62.42	Marking not full in box
8-9	100	Marking answer greater and less than questions
10	93.16	Marking overflow from the box
11	99.57	Delete mark and rewrite mark again
12-14	100	Normally marking for any question

From Table 2 and Figure 7, cases 1-6, 8-9, and 12-14 had 100% accuracy in test scoring. These cases include marked crosses using any pencil, pen, for any question, selecting two answers in one item, not fully marking answers and normally marking for any questions. The feature of cross marks in these cases is a full or nearly full cross in the box (e.g., in Figure 5 numbers 1-4 and Figure 7). Case 7 had 62.42% accuracy in test scoring. This is unsatisfactory because the marked cross did not fill its box (e.g., in Figure 5 numbers 5-6 and 8). Case 10 had 93.16% accuracy in test scoring. It was not fully accurate because of marking overflows from the box (e.g., in Figure 5 numbers 9 and 10). Case 10 had 99.57% accuracy markings had been poorly deleted.

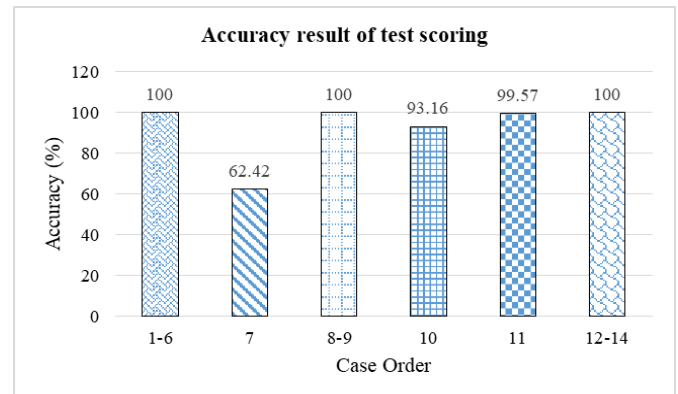


Fig.3: Accuracy of automatic test scoring for each case

In addition, the running time to process with the developed system was compared between our system and manual checking, as shown in Table 3 and Figure 8, showing that the automated system saved time. The time saved was 71.05% of that for manual checking: our system ran about 2.5 times faster than manual checking.

Table 3: Comparison the running time between our system and manual checking

Number of questions	Running time (second)	
	Manual checking	Using our system
20	639	265
30	747	248
40	819	271
50	828	284
60	936	294
70	1320	306
80	1324	360
90	1370	395
100	1584	420
110	1680	526
124	2376	575

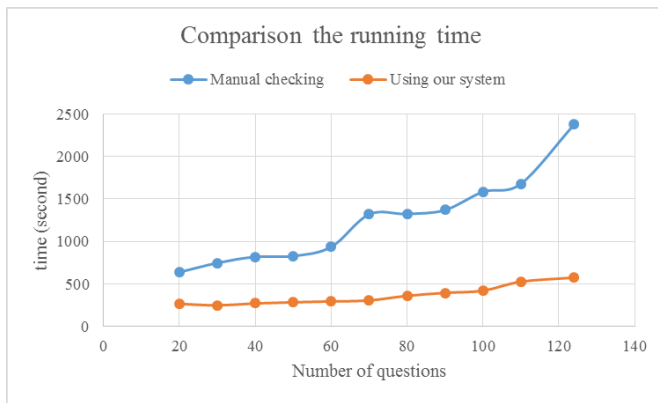


Fig. 4: Comparison of running times between our system and manual checking

5. Conclusion

The automatic image-based grading of multiple choice answer sheets provides significant time and cost reductions. The proposed system applied the correlation coefficient to check answer sheets from solution sheet images. The proposed method supports any pencils and pens used, and also supports thin and low-cost gridded paper used as answer sheets. The proposed method was implemented so it shows the final grading and a statistical summary of scores across all examinees. The developed software was evaluated with 14 cases from 560 answer sheets. The results showed 100% accuracy in detecting cross mark patterns made with any pencil or pen, in any question, and in selecting two answers to one item, not fully marking answers, and normal markings for any questions. The cross pattern in these cases filled or nearly filled its box. The accuracy was 62.42% when the cross did not fill the box. The accuracy was 93.16% with overflow out of the box. Finally, the accuracy was 99.57% with poorly erased markings. In addition, the running time of the developed system was compared with manual checking and showed 71.05% time savings.

Acknowledgements

The authors are deeply grateful to the Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani campus, Thailand. This research was financially supported by Prince of Songkla University, Surat Thani Campus, 2016. The authors also gratefully acknowledge the helpful comments and proving language in this paper of Assoc. Prof. Dr. Seppo Karrila.

References

- [1] S. Kumar, "A Study on Optical Mark Reader", International Interdisciplinary Research Journal, vol. 3(11), pp. 40-44, 2015.
- [2] H. Deng, F. Wang, B. Liang, "A Low-Cost OMR Solution for Educational Applications", In Parallel and Distributed Processing with Applications, ISPA'08. International Symposium on IEEE, pp. 967-970, 2008.
- [3] N.V. Patel, G. I. Prajapati, "Various Techniques for Assessment of OMR Sheets through Ordinary 2D Scanner: A Survey", International Journal of Engineering Research & Technology, vol.4(9), pp.803-807, 2015.
- [4] S. Gupta, G. Singla, P.S. Sandhu, "A Generalized Approach To Optical Mark Recognition", In International Conference on Computer and Communication Technologies, pp. 26-27, 2012.
- [5] N.S. Ratchapaiboon, D. Jetpipattanapong, "Handwritten Digits OCR for Identifying Examinee Number on Objective Test Answer Sheet", Naresuan University Journal: Science and Technology, vol.22(2), pp.94-105, 2014.
- [6] S. Chatree, M. Phayung, S. Sunantha, "Scanner-based optical mark recognition", Information Technology Journal, vol. 5(1), pp.69-73, 2009.

- [7] K. Chinnasam, Y. Rangsanteri, "An Image-processing Oriented Optical Mark Reader", International Symposium on Optical Science, Engineering, and Instrumentation, International Society for Optics and Photonics, pp. 702-708, 1999.
- [8] J.A. Levi, Y.A. Solewicz, Y. Dvir, Y. Steinberg, "Method of Verifying Declared Identity in Optical Answer Sheets", Soft Computing, vol.15(3), pp.461-468, 2011.
- [9] N. Sattayakawee, "Test Scoring for Non-Optical Grid Answer Sheet Based on Projection Profile Method," International Journal of Information and Education Technology, vol. 3(2), pp. 273, 2013.
- [10] F. A. Zampirolli, J. A. Q. Gonzalez and R. P. O. Neves, "Automatic correction of multiple-choice tests using digital cameras and image processing," In the Proceedings of IX Workshop de Visão Computacional, pp. 1-6, 2013.
- [11] H.M. Parul, M. Kaur, "A novel optical mark recognition technique based on biogeography based optimization", International Journal of Information Technology, vol. 5(2), pp.331-333, 2012.
- [12] M. Gharavi-Alkhansari, "A Fast Globally Optimal Algorithm for Template Matching Using Low-Resolution Pruning", IEEE Transactions on Image Processing, vol. 10(4), pp. 526-533, 2001.
- [13] N.R. Pai, S.K. Vijaykumar, "Design and Implementation of Optical Character Recognition using Template Matching for Multi Fonts/Size". International Journal of Research in Engineering and Technology, vol. 4(2), pp.398-400, 2015.
- [14] J. P. Lewis, "Fast Template Matching", In Vision interface, vol. 95(120123), pp. 15-19, 1995.
- [15] J.A. Fisteus, A. Pardo, N.F. Garcia, "Grading Multiple Choice Exams with Low-Cost and Portable Computer-Vision Techniques", Journal of Science Education and Technology, vol.22(4), pp.560-571, 2013.
- [16] A. Spadaccini, V. Rizzo, "A Multiple-Choice Test Recognition System based on the Gamera Framework," arXiv preprint arXiv:1105.3834, 2011.
- [17] A. A. Abbas, "An Automatic System to Grade Multiple Choice Questions paper based exams", Iraqi Academic Scientific Journals, vol. 3(1), pp. 1-8, 2009.