J. of al-anbar university for pure science : Vol.3: No.1 : 2009

ISSN: 1991-8941

An Automatic System to Grade Multiple Choice Questions paper based exams

Ammar Awny Abbas
Computer Center, University of Baghdad, Baghdad, Iraq

Received:7/4/2008 Accepted:2/12/2008

Abstract: The use of Multiple Choice Question (MCQ) in paper based exams is a very popular choice in the international certificate exams (like TOEFL) because it is very fast to grade and it does not let the student write any unnecessary information. In international tests, a specialized machine is used for grading MCQ paper based exams, this machine is very expensive and it needs a special trained operator to operate the machine correctly and efficiently. This project suggests a method to use a personal computer plus a scanner and a program written in Matlab programming language to grade a specially designed MCQ exam test paper with 15 questions with four choices for each question which the student can choose only one answer per question. The program has been tested to detect the correct answers by comparing each paper with a pre scanned test paper that contains the correct answers, many forms of test papers are used to answer different sets real questions for a real exam that has been conducted in the computer center in the Baghdad University and shown to produce results that matches the results gained from grading the same papers manually. The program is written in such a way that it can tolerate rotating the papers in the scanning operation using the process of image registration, any kind of pencil can be used to make the correct answer no matter its color. The program can detect question with multiple choices and eliminate them from calculations.

Key words: Automatic System , Grade , Multiple Choice Questions , exams

Introduction

A traditional multiple choice questions (MCQs) is one in which a student chooses one answer from a number of choices supplied (normally five choices based on A, B, C, D and E). Basically, MCQs consists of the question (stem), the choices provided after the stem (options), the correct answer in the list of options (key) and distracters which are the incorrect answers in the list of options (1). Some of the main advantages and characteristics of the multiple choice questions are (2):

• Marked quickly, sometimes using automatic scanners.
• Marked by markers with minimal training or preparation.
• Used flexibly in print and electronic forms for assessment (including self-assessment) that provides students and teachers with timely, and sometimes automated, feedback on teaching and learning.
• Highly reliable in that results are consistent from student to student and over time.
• An efficient and effective way of assessing factual knowledge.
• Effectively used for quick perception checks during lectures and for systematic revision.

Problem Definition

The answers to the multiple choice questions (MCQs) are written on special forms with blank squares. These forms are scanned and grading is done by comparing the answers with those inputs by the examiner.

Motivation

Many teachers find correcting answer papers very boring and mundane activity (1). The present paper addresses this issue and discusses the development of an effective method for grading short and simple MCQs.

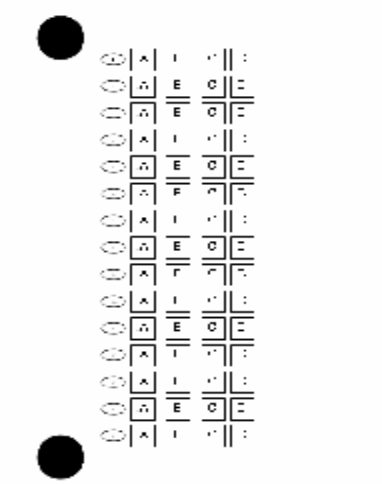


Figure1: A typical MCQs test paper of the proposed method.

Limitations

The student must mark the answer in the square that correspond to the answer the student wants but the answer must be in a special form (only one of the answer choices must be selected) the correct way to answer is shown in figure (2), the square must be completely filled with dark heavy mark so that we cannot see the letter inside the square, light or partial marks may not be read properly by the scanner, we must erase any extra marks completely.

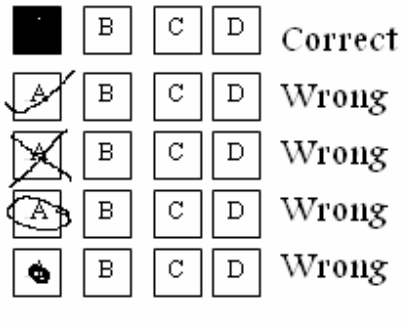


Figure2: The correct way of marking the correct answer.

The mark the student will make must be made by medium-soft (number 2or HB) black lead pencil or any type of ink pencil, this mark will be read by a scanner of any type with a resolution of 75 dot per inch (dpi) which is the lowest resolution that can be selected and will be processed by the Matlab program to determine the correct answer by matching the student answer sheet to the correct answer sheet (called the base sheet) submitted by the person who make the questions.

Outline

The structure of this paper presents the related work in literature in section 3, followed by the basic idea for the problem in section 4, the system layout (system used to solve this problem) is in section 5, and the implementation of the system with a breakdown and description of each of the elements in of that system is in sections 6, the results are provided showing the effectiveness of the proposed method in section 7, the discussions of the results in sections 8, the graphical user interface is in section 9, the future work is found at the end of the document in section 10.

Related Work

The literature on multiple choice tests falls in several categories. Many papers analyze the multiple choice paradigm and provide tests and tools for designing tests that are valid for higher education (3) additionally, much work in psychology and education deal with techniques for designing and phrasing the questions(4).

One reference is particularly relevant to our work. The paper (1) discusses the development of an unconventional, yet an effective method for administering short and simple MCQs as a means of assessing student learning in a focus area of the subject matter. An MCQs test was designed, wherein the five choices (n = 5) of answer to a question is tagged with different numerical numbers based on the multiplication of numeral (n + 1) and not by the alphabets A, B, C, D and E as per normal practice. At the end of the test, the candidate will total the selected choices for questions and the test result can then be obtained by inputting the total choice value in a simple spreadsheet program. The spreadsheet program was written specifically to decode that number and indicate the number of correct answers.

Basic ideas

Base Test Paper Design:

The base test paper is composed of options for 15 questions (four answers for each question) where the examiner must mark the correct answers as shown in figure (2) .The base test paper is shown in figure (3).

In addition to the options for the questions the Base Test Paper and the test paper contain two black circles on the left side, these circles serve as land marks in both base image and answer sheet to align both images at the same coordinate so as to process them correctly in a process called Image Registration (by taking advantage of image processing toolbox of the Matlab program).

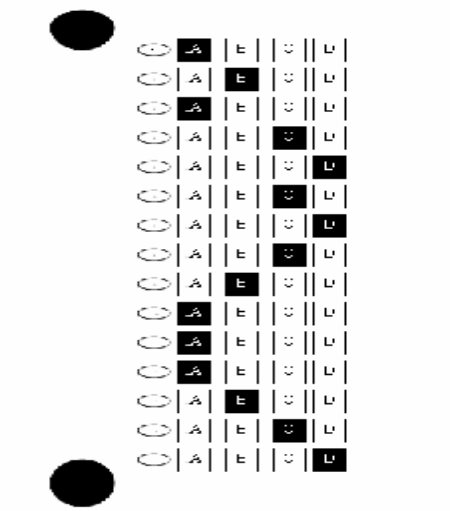


Figure (3): The Base Test Paper

Image Registration

Image registration is the process of aligning two or more images of the same scene. One image called the base image (in this paper we call it the base test paper), this image is considered as the reference to which the other images, called input images (test papers), are compared with. The object of image registration is to bring the test paper into alignment with the base test paper by applying spatial transformation to the test paper (5).

A spatial transformation maps locations in one image to a new location in another image. Determining the parameters of the spatial transformation is needed to bring the base test paper and the test paper into alignment. It is the key to image registration process. The indicators on the left of the base test paper and test paper are used to automatically detect the parameters of the spatial transformation (the simplest registration needs two points detected and verified in the base test paper and test paper images). By detecting the unique size of the circles using Matlab functions, labeling each circle as a unique object, finding the centers of the circles and then passing these centers as the required parameters to make the spatial transformation the registration process is made.

In this research case, the test paper might be rotated in a certain degree due to some human error during scanning but, the program is designed to deal with such cases. So, this program does not need a skillful person to make the scanning (which is the case for dedicated scoring machine).

System Layout

The system model can be broken into three primary stages:

1- Preprocessing.
2- Data Reduction.
3- Feature analysis.

In preprocessing, the operation is performed to make data reduction and analysis easier, the image is converted from a grey image to a binary image. The binary (Black and White) image that contains all the information necessary to discern the object's outline.

In Data Reduction stage the detection of the two indicators on the left of the paper is made by determining their distinctive size from the other objects, then detecting the center of the circles, image registration is performed in this stage. After making the registration, we need to cut only the center of the base test paper and the test paper (where the important information is located) by using the crop function, this process is made to eliminate any unnecessary operations.

In the feature analysis, the features which are extracted by data reduction process are examined and evaluated for their use in the application.

This stage includes ANDing process between the inverted base image and answer image. This will result in an image that contains only the correct answers, where the correct answer may be defined as marking the same square in the base and answer sheets as shown in figure (4).

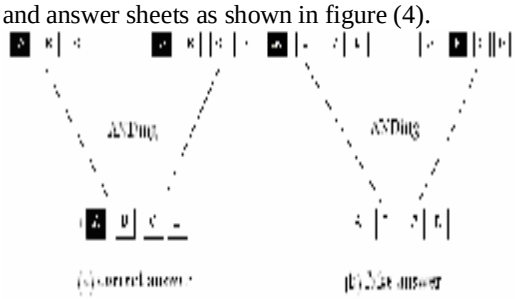


Figure (4): Finding the correct answers by ANDing answers from the Base test paper and test papers.

After the ANDing process we can calculate the squares that results from the process (they also have a distinct size) and these squares are considered as the correct answers. By calculating the percentage of the correct answers to the total answers we can get the final score.

The scanning operation is not included in the model because, it is not part of the computer vision process and it is considered as part of the hardware.

Implementation

The steps of the process are illustrated here with some discussions. We must make sure that the image processing tool box by typing the command (*ver*), which will display information about the version of Matlab that we are using, and the toolboxes that are installed in the program.

Read and |Display the Images

The base test paper and the test paper must be read using (*imread*) function which read the image from graphics file. By using (*imview*) we can display the image. The results are shown in fig (5).

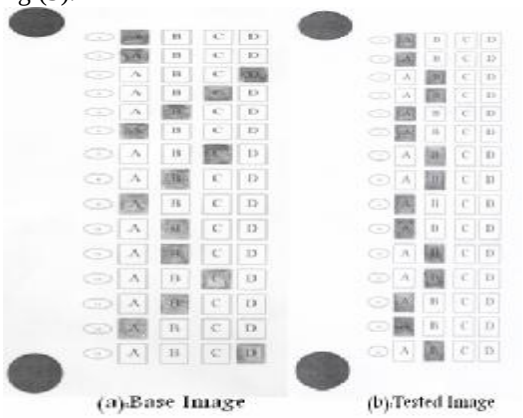


Figure (5): Reading and displaying the base test paper, test paper

Convert the Image to Black and White Image

The base test paper and test paper must be converted to black and white image for further processing. This will reduce the size of the image and eliminate the unnecessary color information using the function (im2bw). The papers were not scanned as black and white in the first place because the large amount of noise that was introduced to the images. The results are shown in fig (6).

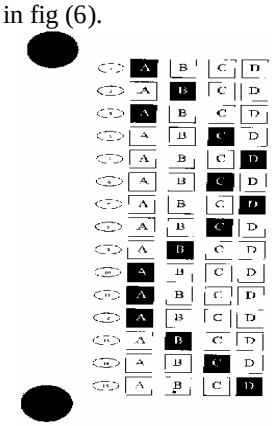


Figure (6): Convert the image to Black and White image

Invert the Colors of the Images

The two black and white images are inverted by making the black color white and the white color black, by doing this the important shapes are converted to white, this step is important because most of the functions processes white shapes only. The results are shown in fig (7).

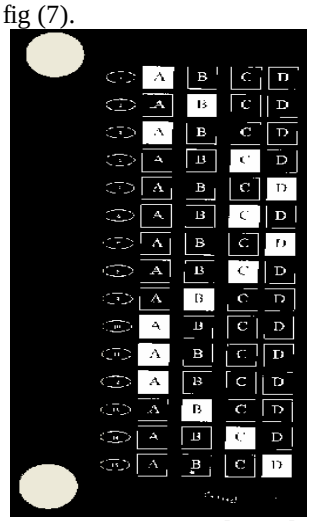


Figure (7): Inverting the Colors.

Eliminating Small Objects

To register the two images we need to specify two points in each image (registration is made to make the two images in the same coordinate system if there is an error in the scanning process). We want to get the centers of the two black circles on the left to pass the centers to the control point selection function (*cpsselect*) to make the registration by eliminating any object that has lesser size 4300pixel(this number is pre-calculated), this will leave only the two circles. The results are shown in fig (8).

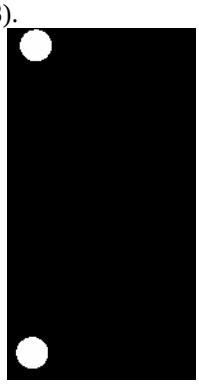


Figure (8): Eliminating Small objects

Label the objects in the Image and Find the Center of the Circles

We must mark all the components in the image, By doing this the number of labeled objects is found in (*numobjectsbase*) parameter, this step is related with another step to measure the basic properties of the image components such as the area of the components and the center of each components.

Since we have only the two circles if we call the (*centroid*) property of the (*regionprops*) function we can obtain the center of the indicators (circles).

Registration Process

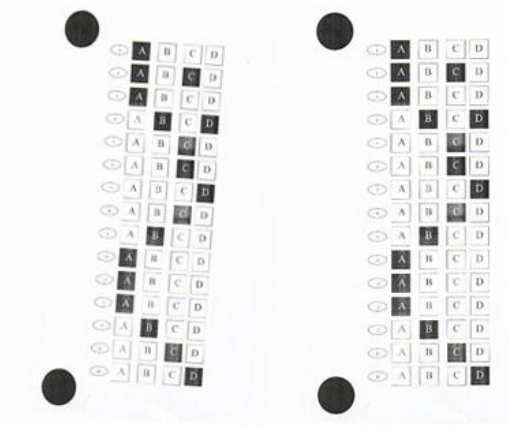
The first operation in the registration process is selecting the control points that we acquired from the last step and pass them to the function (*cpselect*).

Get the Center of the Circles in the Registered Image

In the same way we must get the center of the circles of the registered image, to eliminate any object smaller than 4300 pixel in area. Then we must label the objects and get the properties of the objects in the resulting image. Then we can get the center of the circles by using the centroid property.

Performing the registration

After selecting the control points the registration operation can take place using the function (*imtransform*), the results of the operation is shown in figure



(a) (b)

Figure (8): Registration process, (a): Rotated test paper, (b): Registered image

Information Cropping

This step is made to crop only the Region Of Interest (ROI) from the images. This will reduce the size of the two images to reduce the calculations. The resulted images are shown in figure (9).

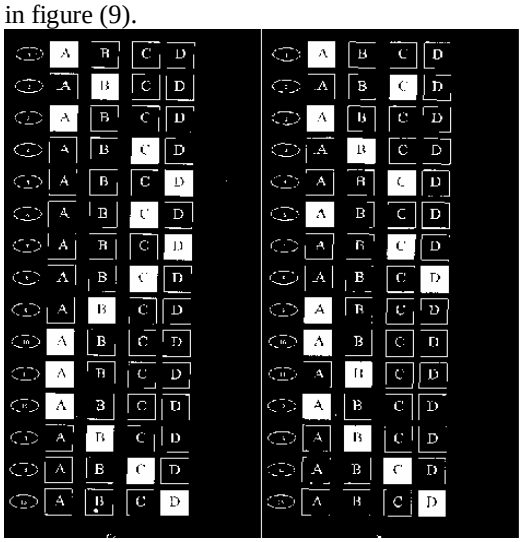


Figure (9): The cropped images (Base test paper on the left and test paper on the right)

Check for Questions that have Multiple Answers and Eliminating them

If the student answers any question more than one answer for a single question this step eliminate that question from the overall calculation (without affecting the results). The original test paper and the test paper after eliminating the question with more than one answer is shown in figure (10)

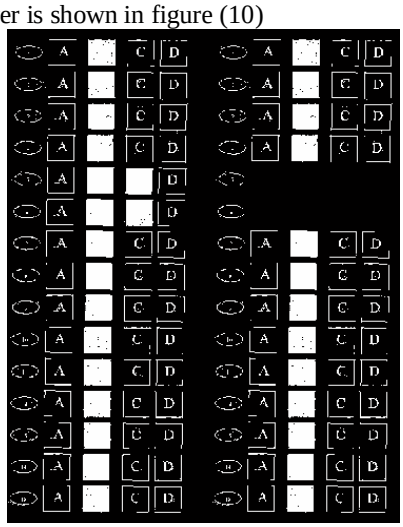


Figure (10):Check for Questions that have Multiple Answers and Eliminating them

Calculate the Correct Answers

An important characteristic of the correct answer is that "it is a white square that is found in the same place in the base test paper

and the test paper, while the false answers is a white square found in an area in the test paper but it is not found in the same area in the base test paper", as shown in figure (4).

Using this property we can multiply the two images and only the correct answer will appear in the resulted image, the correct answer is white and has the value of 1 from multiplying a square of ones from base test paper with a square of ones in the test paper both in the same coordinates in the images, while the false answers that will be in the test paper will disappear because multiplying a square of ones by a square of zeros will result a square of zeros, the resulting image of multiplying a base image with a test paper with seven correct answers is shown in figure(11).

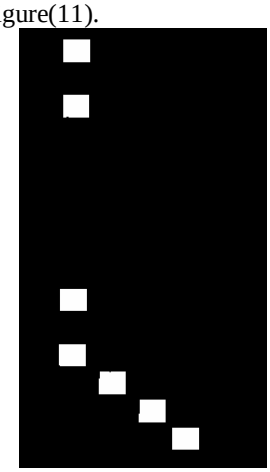


Figure (11): Multiplying a Base test paper with a test paper of 7 correct answers.

When the resulted image appears we can make a simple count for the objects (which in this case the correct answers) in the image, the targeted area is around 1300 pixels (pre-calculated number).

Results

The program is tested with 63 different exams with twenty students for each exam making a total of 1260 test paper; each paper is scanned as an RGB image and with resolution 75 dots per pixel (dpi), which is the lowest resolution that can be selected by Mirascan program.

The images are saved as JPEG image format which is a very small in size compared to other types of images like BMP and PNG and each test paper is compared to the base test paper to calculate the correct answer. The students are asked to fill the test paper by filling the appropriate square that corresponds to the correct answer, the student were free to use any type of pencil of any color to fill the square of the correct answer as long they obey the rules of marking the correct answer that is: filling the square completely; some of the students were asked to fill the correct answer squares incompletely as shown in figure (2), and the program could not recognize the answers. Only when the correct way to fill the squares is used the program works correctly.

The results are calculated almost instantly after the test papers are scanned and fed to the program; and when compared to correct answer that is graded manually the accuracy is 100%.The program shows that it can deal with a major scanning error that is made by the person who performs the scanning operation which is: rotating the paper, the registration process can deal with this problem and correct this error.

Discussions and Conclusions

The present paper has provided an overview of the conventional multiple choice tests and discussed the design of MCQ test paper that is suitable as a means for measuring student learning outcome. The proposed MCQ test papers can be used for any type of exam. Although the scanning process is relatively slow, it was found that the proposed method is applicable and can provide the means to grade papers that is very fast and accurate when compared to manual grading, the proposed method provide a cost effective method with a moderate speed when compared to the specialized machine for grading MCQ paper based exams, And when compared to manual grading it is considered very fast. The MCQ test paper was designed to facilitate rapid marking of the answers. The experiment conducted on a group of 1260 students showed that the MCQs test is popular and achieved considerable student satisfaction for the participants. The following conclusion can be drawn from the research:

- Some of the students complained that the shapes of the answers (squares) are difficult to be filled (especially on the edges) so that it is better to change from the square shape to another shape that does not contain angles (circles or ellipses).
- Regarding the hardware, the use of a personal computer and a scanner to grade multiple choice test papers is an appropriate approach to grade papers, because they are widely available, and it can be widely adapted in high schools and colleges because it is very cheap and relatively fast.
- Regarding the software, the software is

very robust and this procedure does not require any special skills after scanning the operator must give an appropriate name for each test paper and make sure that this name is included in the Matlab work file, in order to be processed properly.

- There is no special preference for the pencil that is used by the students; any type of pencil can be used to grade the paper no matter what color it is.
- The system can tolerate rotating the exam paper in the scanning process.

Graphical User Interface (GUI) for the Program

The Graphical User Interface is shown in figure (12) below:-

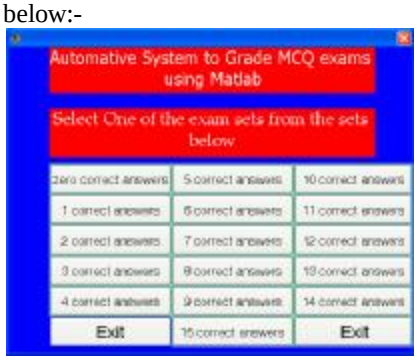


Figure (12): Graphical User Interface

Each one of the buttons when pressed will open a new page similar to the figure (13) below, for instance, if we press button named 7correct answers (which means that all the samples in this paper contains only seven correct answer.

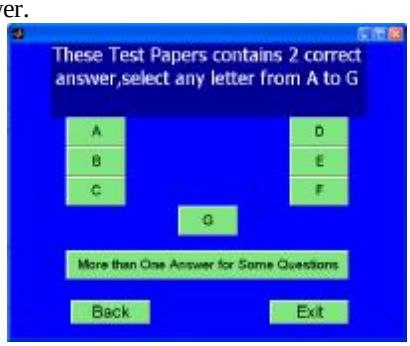


Figure (13): Opening a test paper that contains seven correct answers

After we select one of the samples the (A-G). Images containing the correct and tested paper, final images after processing them and the final image after multiplying the two images and the correct answer appear on the work space. The button named "More than One Answer for Some Questions" when pushed will demonstrate the case when the student answer more than one answer for a single question.

System Drawbacks

The system can tolerate errors in the scanning to a certain degree using registration process, but acute errors in scanning (such as rotating the test paper more than forty five degrees)cannot be corrected by the program.

The program assumes that there is one answer per question. When the student selects more than one choice for a single question the program can eliminate the question from the calculation. The system cannot deal with questions that have many answers for a single question.

Future Work

1- Increase the number of questions in the paper.

2- Increase the number of option for each question in the paper.

3-Change the shape of option boxes from square to circles and ellipses to facilitate the filling of the objects.

4-Change the control point circles to smaller ones.

References

1- 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).

2-Bar-Hillel, M., Budescu, D., & Attali, Y. (2005). Scoring and keying multiple choice tests: A case study in irrationality. Mind and Society, VOL,4,NO.8PP(3-12).

3-D. Traynor and J. P. Gibson. 2005. "Synthesis and analysis of automatic assessment methods in cs1" { generating intelligent mcqs. In SIGCSE. ACM,VOL8,NO.3PP(44-65).

4- M. Bush.2001. "A multiple choice test that rewards partial knowledge". In Journal of Further and Higher Education, VOL25NO.(2).CarfaxPublishing,PP(101).

5- Rafael M.crouch."The Matlab image processing toolbox guide",USA,Adison-Wisley, 2004,PP(116-150).

6- Rafael C. Gonzalez and Richard E.woods,"Digital Image Processing", Reading:Adison-Wesley,USA 2002,PP(20-80).

7-John K. Wally. "The Matlab image processing toolbox guide", 2004.Prentice hall USA,PP(50-70)

8- The official TOEFL website, www.ets.org and the examinee handbook.

J. of al-anbar university for pure science : Vol.3: No.1 : 2009

النظام الالي لتصحيح اوراق اسئلة الاختيارات المتعددة المعتمدة على الاختبار.

عمار عوني عباس
E.mail: scianb@yahoo.com

الخلاصة:

إن استخدام الأسئلة متعددة الاختيارات في الامتحانات العالمية في الامتحانات التي يؤديها الطالب في الورقة شائع جدا في الامتحانات العالمية مثل امتحان التوفل لأنها سريعة التصحيح ولا تترك للطالب إمكانية كتابة معلومات غير ضرورية. يستخدم في الامتحانات العالمية جهاز مخصص لتصحيح الاوراق وهذا الجهاز مكلف جدا ويحتاج إلى مشغل يكون مدربا للعمل على هذا الجهاز بصورة صحيحة ومجدية. يقترح هذا البحث استخدام الحاسب الشخصي مع جهاز الماسح بالإضافة إلى برنامج مكتوب بلغة برنامج Matlab لتصحيح ورقة امتحان مقترحة بها 15 سؤالا لكل سؤال اربعة اجوبة يستطيع ان يختار واحدا منها لكل سؤال. تم اختبار البرنامج لتحديد الاجوبة الصحيحة في اوراق امتحانات حقيقية بواسطة مقارنتها بورقة امتحان للاجوبة الصحيحة مدخلة مسبقاً, تم استخدام اكثر من نموذج للاجوبة المستقاة من امتحانات حقيقة اجريت في مركز الحاسبة في جامعة بغداد واظهرت النتائج تطابقا مع النتائج التي تم الحصول عليها من تصليح الاوراق يدويا.تمت كتابة البرنامج بحيث انه يستطيع تجاوز اخطاء عملية ادخال الصفحة بصورة غير صحيحة بواسطة الماسح الضوئي(تدوير الصفحة بزاوية معينة) باستخدام عملية تسجيل الصورة, يمكن استخدام اي نوع من الاقلام لتاشير الاجوبة الصحيحة مهما كان لون القلم المستخدم. ويمكن للبرنامج ان ان يحدد الاسئلة التي اجاب الطالب بها اكثر من اجابة ويستبعد هذا السؤال من الحسابات.