International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || Volume 04 - Issue 05 || May 2019 || PP. 127-136

Development of an MCQ checking system using circle detection algorithm

Md. Moazzem Hossain[1], Md. Abu Bakar Siddik[2], Md. Sydur Rahman[3]
[1,3](*Department of Computer Science & Engineering,*
Bangladesh Army University of Science and Technology, Bangladesh)
[2](*Department of Electrical and Electronics Engineering,*
Bangladesh Army University of Science and Technology, Bangladesh)

Abstract: Now a days the use of Multiple Choice Question (MCQ) in paper based exam is a very popular choice in the examination system because it is very fast to grade and it does not let the student to write any unnecessary information. Usually a specialized machine is used for this purpose which is much expensive and needs a special trained operator to operate the machine correctly and efficiently. Moreover there are some vital issues for magnetic interference with the machine for malfunctioning the output result. In this paper we have proposed an alternative approach to examine those MCQ answer papers using a scanner and a computer with a software. To design the software Matlab programing has been used. Here an efficient circle detection technique has been used. Circle detection is the most used technique in the application of computer vision. It is also the most important and most essential task to detect a circle for different applications. Here, our target is to design an automated system for an Optical Mark Recognition (OMR) software which can examine the Multiple Choice Question (MCQ) answer sheets more effectively and efficiently. Circular Hough Transformation (CHT) is a widely used algorithm for circle detection. So CHT has been used to design this system. At first, this software takes the correct answer sheet as input and then identify the correct answers and then take other answer sheets and detect the given answers and then compare with the first one and provide results. By using this software we can check an MCQ sheet and can provide the correct results. Our main goal is to make the software more reliable for examining MCQ sheet. We have done this task roughly. Our software can examine the MCQ sheet and can provide the result in MS Excel format. This is very easy to use. This work mainly emphasizes on some tasks: First of all, it studies image pre-processing, circle detection, comparing two images, and calculation of results and provides output to the user.

Keywords: Automated system, Circular Hough Transformation. Multiple Choice Questions, Optical Mark Recognition,

I. INTRODUCTION

This is a challenging task to make a software by using image processing for MCQ checking system. Now a days we use OMR machine which is very costly and sometimes provide faulty results. And this is not usable for all people because it is very complicated to use. Like a developing country as Bangladesh it is not possible for all organizations to examine their MCQ sheets with this costly OMR machine. Besides by manually examining this sheets is more time consuming and boring task. When we are thinking about digitalization of our country then why we will use this manual process for our education system. So we have tried to make a software to make the whole process digital and efficient also. OMR is the scanning process of paper to detect the presence or absence of a mark in a predetermined position. OMR machine is used to check the answer of MCQ. It is an electric device by which we can examine the MCQ sheets. This machine is used now a days by big organizations. But affordable for small organizations. A traditional MCQ system is one in which a student chooses one answer from a number of choices supplied (normally four choices based on A, B, C and D). Basically, MCQ consists of the question, the choices provided after the question known as options. There is a correct answer in the list of options and others are the incorrect answers in the list of options. There are some of the main advantages and characteristics of the multiple choice questions are:

- 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.

Many of today's OMR applications involve people filling in specialized forms. These forms are optimized for computer scanning, with careful registration in the printing, and careful design so that ambiguity is reduced to the minimum possible. Due to its extremely low error rate, low cost and ease-of-use, OMR is a popular method of tallying votes. OMR marks are also added to items of physical mail so folder inserter equipment can be used.

II. RELATED WORKS

The OMR scanners began to be used in the early 1950's, which used a series of sensing brushes in detecting graphite particles on a document that is passed through the machine to detect the given answer points. Simply the system operation could be distinguished into two modes: learning mode and recognition mode. At first we have to extract data from each area which can be performed based on the horizontal and vertical projections method. We have used the number of black pixels in each answer block is counted, and the difference between those numbers in the input and its corresponding model is used as decision criterion for the purpose of checking answer. It was a transition between punch cards and barcodes. Pegasus Imaging Corporation presented a Software Development Kit for OMR recognition from document images. Recent works include software along with specified scanner for specific design of Document or Form for OMR purpose. Tao and Toan presented some difficult problems of optical mark recognition. There are important problems such as the correcting position of form with pattern and correctly detecting geometrical objects. Hussmann S. et al, 2005 proposed a low cost and high speed system using Field Programmable Gate Array (FPGA), but had constraints on the input of the forms. Pegasus Imaging Corporation presented SDK for OMR recognition from document images, which supported template and free recognition mode. An OMR field is defined as a rectangle area containing a specified number of columns and rows of bubbles to be evaluated. The SDK can scan the region horizontally and then vertically to locate the bubbles apart from the spaces between them. Based on the bubble shape specified, the system scans the discrete zones corresponding to the bubble, counting dark pixels to determine which bubble zones qualify as filled in. The Pegasus, technique can support the plain paper printing and design, but in its application in schools, the multiple choice answer recognition success rate cannot achieve the requirements of the examination. The proposed system on the contrary combines the best features of all the others, and presents a low cost and high speed recognition system, which is flexible and custom configured to user requirements, without the aid of complex hardware. Also the experimental results show the high accuracy rate achieved by the proposed system. Some existing systems are given below:

- OMRHOME
- REMARK OFFICE OMR
- ADMEN GROUP OMR

III. LITERATURE REVIEW

We have studied different types of documents which are needed for our thesis and also studied many papers also. We have studied about OMR machine, its working principle, we have also studied different types of images. Image prepossessing, image segmentation, image enhancement, image histogram equalization, image re-size and reshape, image quality enhancement and so on. We have studied many thesis papers like: Finding edges and lines in images. In this paper we have known about the edge detection techniques using Canny Edge Detection algorithm [1].

Here we have also different types of operator like: Sobel operator, Prewitt operator, Gaussian operator e. t. c. Different circle detection algorithms like Circular Hough Transform. In this paper we have known about the Circular Hough Transformation technique and its advantages and disadvantages. We have also known about the advance CHT for multiple circle detection. This method uses the principle of template matching, which relies on detecting smaller elements matching a template image [2].

We have also studied about fast and robust circular object detection with Probabilistic Pairwise Voting (PPV) method. From this paper we have taught about another efficient algorithm for multiple circle detection. The key idea behind this method was to create a generative model for each place in the database, by learning distributions of word triplets over the placentas training images, and then finding matches between pairs of query features and pairs of database landmarks. These matches then vote for the respective place, with the vote weighted probabilistically [3].

Then we learn about an efficient circle detection scheme in digital images using Ant System Algorithm a completely new, time and memory efficient, multi-agent searching scheme for detecting the intersecting circles in an edge detected grey scale digital image. In this paper they consider the intersecting boundaries as a graph consisting of nodes and branches altogether. All the nodes and branches are found out and the incidence matrix of the graph is constructed. Using this information the closed loops in the graph are found out which are then tested for circles [4].

Circle with maxima minima, Centre detection of circle, multiple circles center's detection. In this section we have studied about Circle detection using maxima minima points and center of the circle.The fast and accurate circle (FACILE) detection algorithm[6], based on gradient-direction-based edge clustering and direct least square fitting. Edges are segmented into sections based on gradient directions, and each section is validated separately; valid arcs are then fitted and further merged to extract more accurate circle information. We implemented the algorithm with the C++ language and compared it with four other algorithms. Testing on simulated data showed FACILE was far superior to the randomized Hough transform, standard Hough transform, and fast circle detection using gradient pair vectors with regard to processing speed and detection reliability. Testing on publicly available standard datasets showed FACILE outperformed robust and precise circular detection, a state-of-art arc detection method, by 35% with regard to recognition rate and is also a significant improvement over the latter in processing speed[6].

Fast algorithms for line and circle detection based on inverted gradient hash maps (IGHM)[7]. Inverted indices are a common technique for storing a map from content of a dataset to its locations in the dataset. Hash maps are typically used to implement associative arrays and reduce search times in large datasets. In this paper, a hash map is used to store an inverted index of image gradient magnitudes and orientations. Algorithms for detecting lines, and circles using IGHMs are presented and shown to be competitive against existing approaches [7].

We have also studied about multiple circle detection technique here. Fast Circle Detection Using Gradient Pair Vectors Detecting lines and circles in an image is a fundamental issue in image processing applications. Extracting circles from digital images has received more attention for several decades because an extracted circle can be used to yield the location of circular object in many industrial applications. So far many circle - extraction methods have been developed.

Hough transform has been the most common method for circle detection exhibiting robustness but adversely demanding a considerable computational load and large storage. Alternative approaches include heuristic methods that employ iterative optimization procedures for detecting multiple circles under the inconvenience that only one circle can be marked at each optimization cycle demanding a longer execution time. In contrast, learning automata (LA) is a heuristic method to solve complex multi-modal optimization problems [8].

Optical Mark Recognition (OMR) is the technology of electronically extracting intended data from marked fields, such as squareand bubbles fields, on printed forms. OMR technology is particularly useful for applications in which large numbers of hand-filled forms need to be processed quickly and with a great degree of accuracy. The technique is particularly popular with schools and universities for the reading in of multiple choice exam papers. This paper proposed OMR based on Modify Multi-Connect Architecture (MMCA) associative memory, its work in two phases: training phase and recognition phase. The proposed method was also able to detect more than one or no selected choice. Among 800 test samples with 8 types of grid answer sheets and total 58000 questions, the system exhibits an accuracy is 99.96% in the recognition of marked, thus making it suitable for real world applications [11].

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 [12].

Then we have choose advanced Circle Hough transform (CHT) [2]as it is one of the best known algorithms and aims to find circular shapes with a given radius r within an image. Usually edge map of the image is calculated then each edge point contributes a circle of radius r to an output accumulator space. For unknown circle radius, the algorithm should be run for all possible radiuses to form a 3-dimensional parameter space, where two dimensions represent the position of the centers, and the third one represents the radius. The

output accumulator space has a peak where these contributed circles overlap at the center of the circle. This algorithm gives more accurate result for multiple circle detection of an image.

IV. METHODOLOGY

The circle Hough Transform (CHT) is a feature extraction technique for detecting circles. It is a specialization of Hough Transform. The purpose of the technique is to find circles in imperfect image inputs. The circle candidates are produced by voting in the Hough parameter space and then select the local maxima in a so-called accumulator matrix. The CHT relies on equations (1) for circle[2]. The equation of the a circle is,

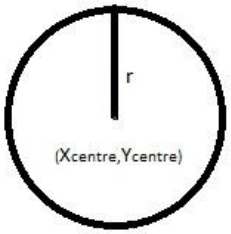


Figure 1. A circle with radius r and centre at (X_{centre}, Y_{centre}).

$$(x - x_{center})^2 + (y - y_{center})^2 = r^2 \qquad (1)$$

where (xcentre ,ycentre) is the centre of the circle, and r is the radius of the circle. From the above equation, we can see we have 3 parameters, so we need a 3D accumulator array for Hough Transformation. The parametric representation of this circle is $x = x_{center} + r\cos\theta$ and $y = y_{center} + r\sin\theta$. In contrast to a linear Hough Transformation, a CHT relies on 3 parameters, which requires a larger computation time and memory for storage, increasing the complexity of extracting information from our image. For simplicity, most CHT programs set the radius to a constant value (hard coded) or provide the user with the option of setting a range (maximum and minimum) prior to running the application. We used an alternative set of transform equations for circles which do not make use of gradient direction information, and therefore avoid trigonometric functions and the inaccuracy of edge gradient directions caused by the edge detector. It is known that a circle can also be represented by equation (2).

$$x^2 + y^2 + 2ax + 2by + c = 0 \ ; \ where \ (a^2 + b^2 > c) \qquad (2)$$

And three non-collinear points can determine a circle and only one circle. Suppose that three points are (x_1, y_1), (x_2, y_2) and (x_3, y_3). After substituting these three points into Eq. (2), the following equations are obtained:

$$x_1^2 + y_1^2 + 2ax_1 + 2by_1 + c = 0$$
$$x_2^2 + y_2^2 + 2ax_2 + 2by_2 + c = 0$$
$$x_3^2 + y_3^2 + 2ax_3 + 2by_3 + c = 0$$

These equations can be solved, using Cramer's rule, we to obtain the coefficients a, b and c from the above equations,

$$a = \frac{D_1}{D}; \quad b = \frac{D_2}{D} \quad ; \quad c = \frac{D_3}{D}$$

Where,

$$D_1 = \begin{vmatrix} -(x_1^2 + y_1^2) & 2y_1 & 1 \\ -(x_2^2 + y_2^2) & 2y_2 & 1 \\ -(x_3^2 + y_3^2) & 2y_3 & 1 \end{vmatrix}$$

$$D_2 = \begin{vmatrix} 2x_1 & -(x_1^2 + y_1^2) & 1 \\ 2x_2 & -(x_2^2 + y_2^2) & 1 \\ 2x_3 & -(x_3^2 + y_3^2) & 1 \end{vmatrix}$$

$$D_3 = \begin{vmatrix} 2x_1 & 2y_1 & -(x_1^2 + y_1^2) \\ 2x_2 & 2y_2 & -(x_2^2 + y_2^2) \\ 2x_3 & 2y_3 & -(x_3^2 + y_3^2) \end{vmatrix}$$

$$D=\begin{vmatrix} 2x_1 & 2y_1 & 1 \\ 2x_2 & 2y_2 & 1 \\ 2x_3 & 2y_3 & 1 \end{vmatrix}$$

Eq. (2) can be rewritten as $(x+a)^2+(y+b)^2=a^2+b^2+c$.Comparing this with Eq. (1), we can see that the circle Centre (x centre, y centre) and radius r are determined by (-a,-b) and $(a^2+b^2-2)^{1/2}$ respectively. Here (x_1, y_1), (x_2, y_2) and (x_3,y_3) are the feature points in the image space which can be obtained by the application of an edge operator. A 2-dimensional accumulator array is used for circle centre, (-a, -b), accumulation and a 1-dimensional histogram for circle radius, $(a^2+b^2-2)^{1/2}$. Every three edge points are selected as a group to calculate one possible centre and radius. The evidences for centre estimates are accumulated in the 2 dimensional parameter space. The radius estimates are stored in a 1 dimensional array. The final radius is determined by the value with maximum frequency in this array.

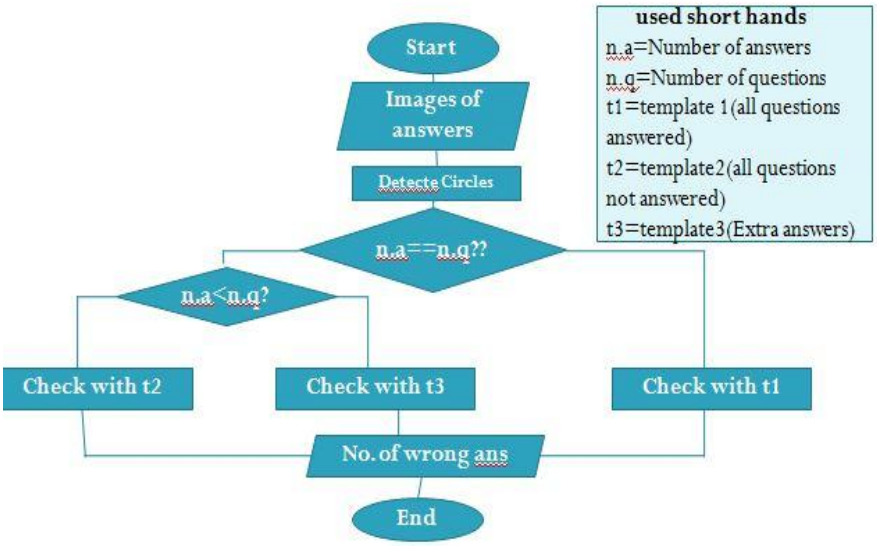


Figure 2. Flow diagram of the system

We used this algorithm to detect circle from the given answer script. First of all we have pre-processed the answer script applying image reconstruction and enhancement process then rotated the image to centre the reference the point. Then image resize applied to make same size like reference image and then circle detection algorithm CHT applied to detect the black circle and corresponding centre location and radius[5]. Using this information the answered options of each questions calculated. Then it was matched with reference correct answered image and marked as green circle for correct answers and red for false answers. Here it also calculated the multiple answered and no answered questions and marked then red as false. Then counted the correct answers using the green circle number.

From the figure 3(b) we can see that here the black circles are the given answers and their centre position we can calculate the actual given answer. Here the first given answer is A because here the Row number of each centre points give the x axis as A, B, C and D on the other hand y axis give the question number. So we see that it's a pretty simple concept if we can get the exact point of the centre from the given answer. According above procedure we can calculate the other answers as Table I.

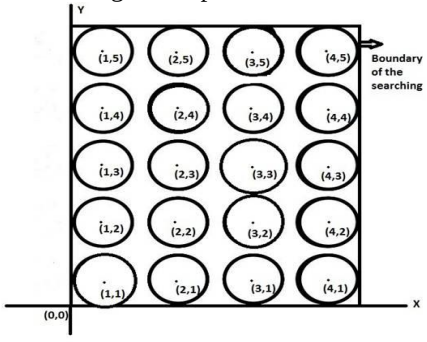


Figure 3(a). A blank MCQ paper with x-y coordinate system.

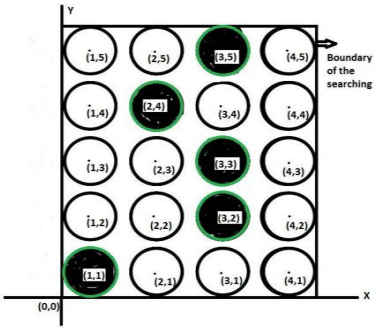
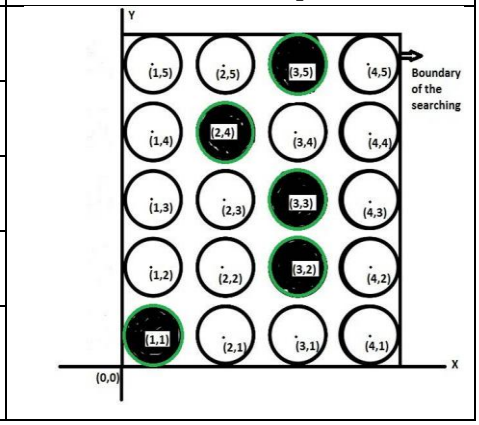
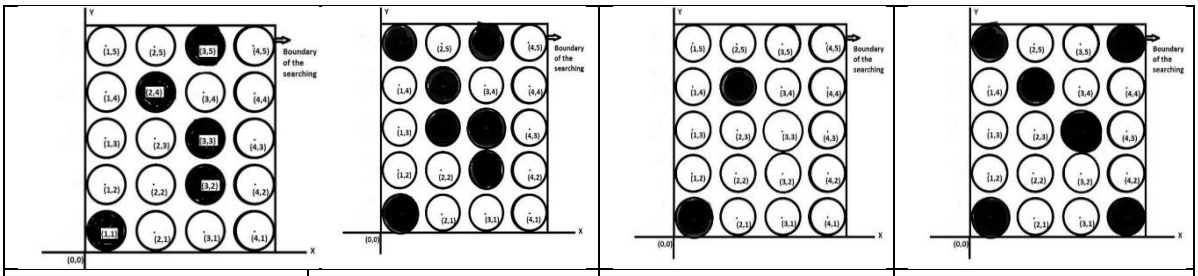


Figure 3(b). A filled answer sheet and the answered positions in the x-y coordinate.

Table 1: Question Numbers and Corresponding Answers

Question No	Answers	Answer script
5	C	
4	B	
3	C	
2	C	
1	A	

Now there would be many kind of answers provided by users so we have worked with four types of answers at figure 4.



Type-1	Type-2	Type-3	Type-4

Figure 4. Four types of answer sheets

(i) All Questions are answered no double answers.
(ii) All Questions are answered but some are doubled answered.
(iii) All Questions are not answered.
(iv) All Questions are not answered and also some are doubled answered.

And the result table of this four types of answer sheet would be like this Table II.

TABLE II. QUESTION NUMBERS AND CORRESPONDING ANSWER OF FOUR TYPES OF ANSWER SHEETS.

Question No.	Type-1	Type-2	Type-3	Type-4
1	C	X	X	X
2	B	B	B	B
3	C	X	X	C
4	C	C	X	X
5	A	A	A	X

Here we see that for the type-1 user have answered all questions and the answers are recorded and in type-2 user have given some double answers so this doubled answers are recorded as X which indicate wrong answers. And in type-3 here all questions are not answered by user these questions are marked as X which indicate wrong answers. At last type-4 is the mixture of type-2 and type-3 where some answers are double and some are blank. In this case these all are recorded as X which indicate wrong answers.

V. IMPLEMENTATION AND ANALYSIS

Actual working procedure of our software is given below with a Fig. 5. At first we can provide the input images from any scanner like smart phone, normal scanner, camera scanner. Then the images are pre-processed. In pre-processing stage it resizes the image and resolution equalization is also done. After this the image is converted into binary image for processing. Then our software detect the dark circles using the CHT algorithm. Then apply the CHT algorithm for black circles [13]. Then check the no of answer or filled circle

with the numbers of total questions and take the decision.

The pre-processing phase consists in a set of operations that make the scanned image more suitable for the further phases. The first operation performed to the image is the conversion to gray scale; then the image is converted into black and white format using the thresholding method. Next the system does a compensation of rotation effects induced by the scanning operation. The goal of this step is rotate of image answer sheet at a calculated angle to restore it to its normal rectangle. To do that, at first we must calculate the correct angle by using Hough transform method, and then apply bilinear interpolation method with correct angle to rotate all image answer sheet pixels to normal location. Figure 4 shows image answer sheet before and after rotation operation [12].

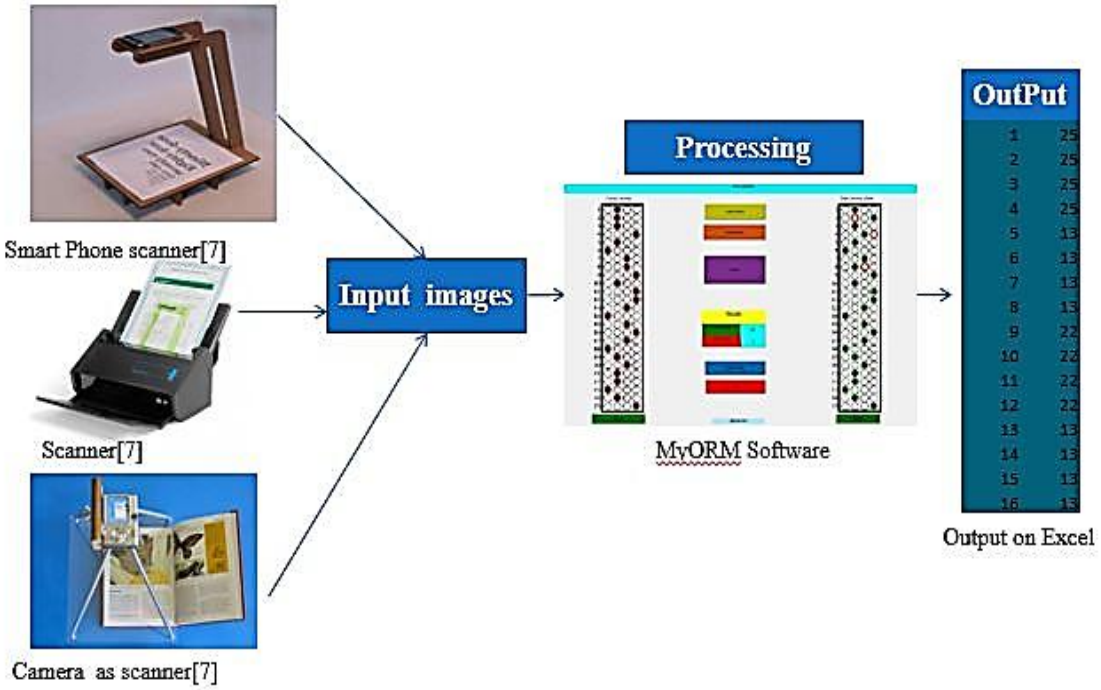


Figure 5. Working procedure of System.

If two numbers are equal then check with the template of all question answered. If not then check if the number of given answer is lees then the number of questions then check with the template where all question not answered. If not true then check with the template where extra answers were given. From all this step we get the number of wrong answer after comparing with each other's. Some output of the software is given below:

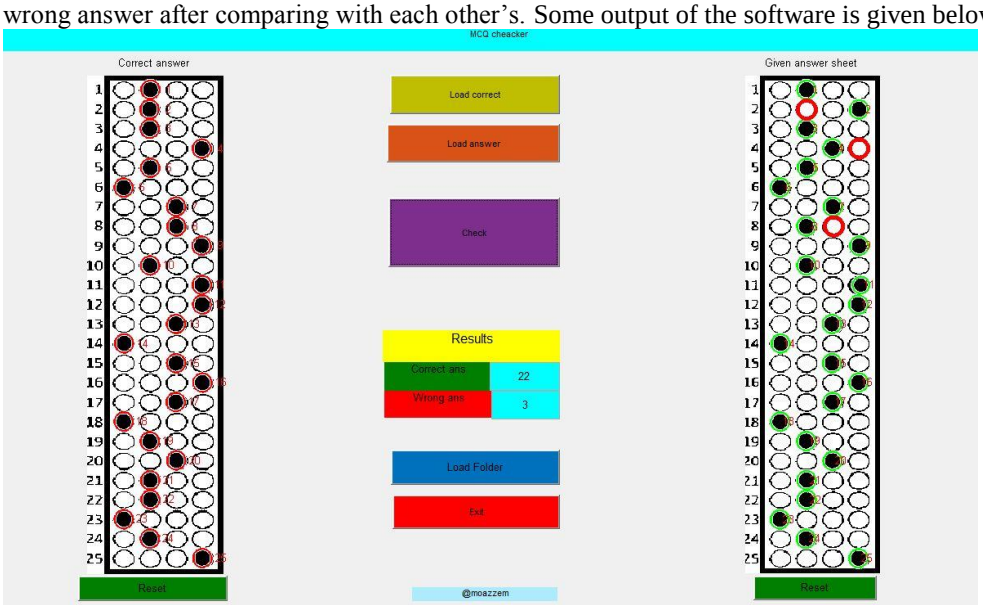


Figure 6. Type-1 real answer script evaluation with result.

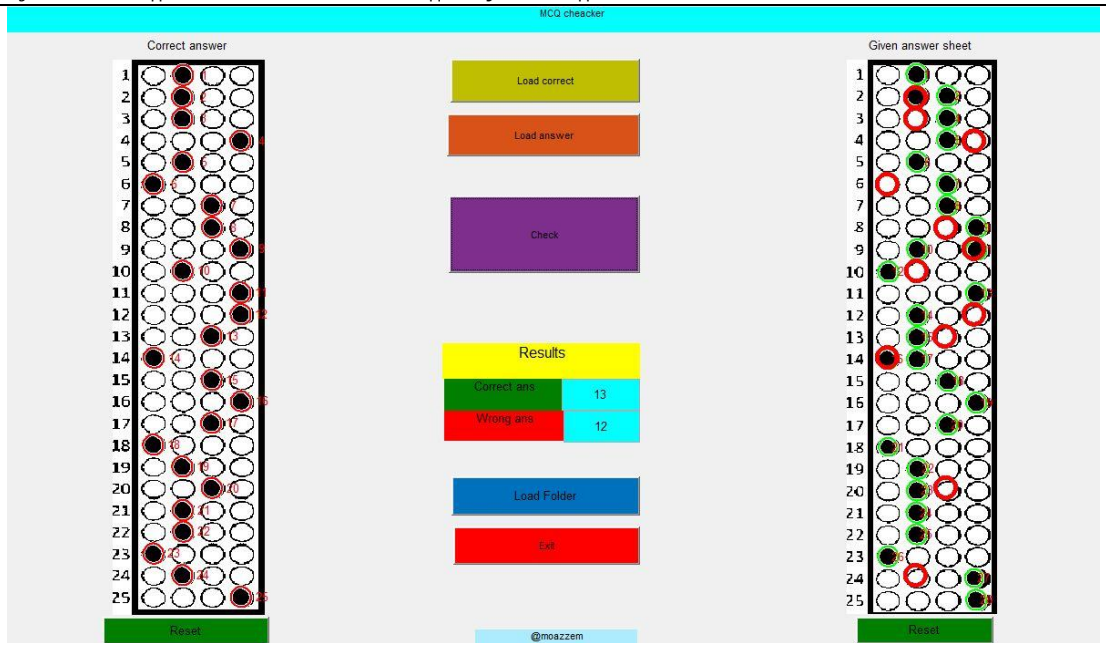


Figure 7. Type-2 real answer script evaluation with result.

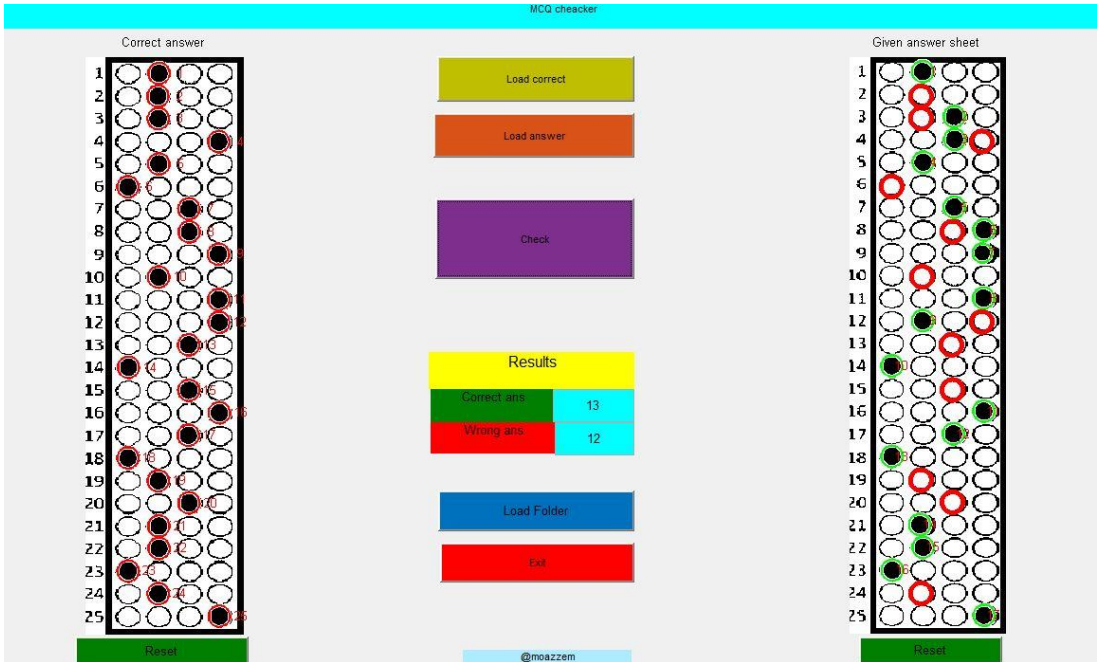


Figure 8: Type-3 real answer script evaluation with result.

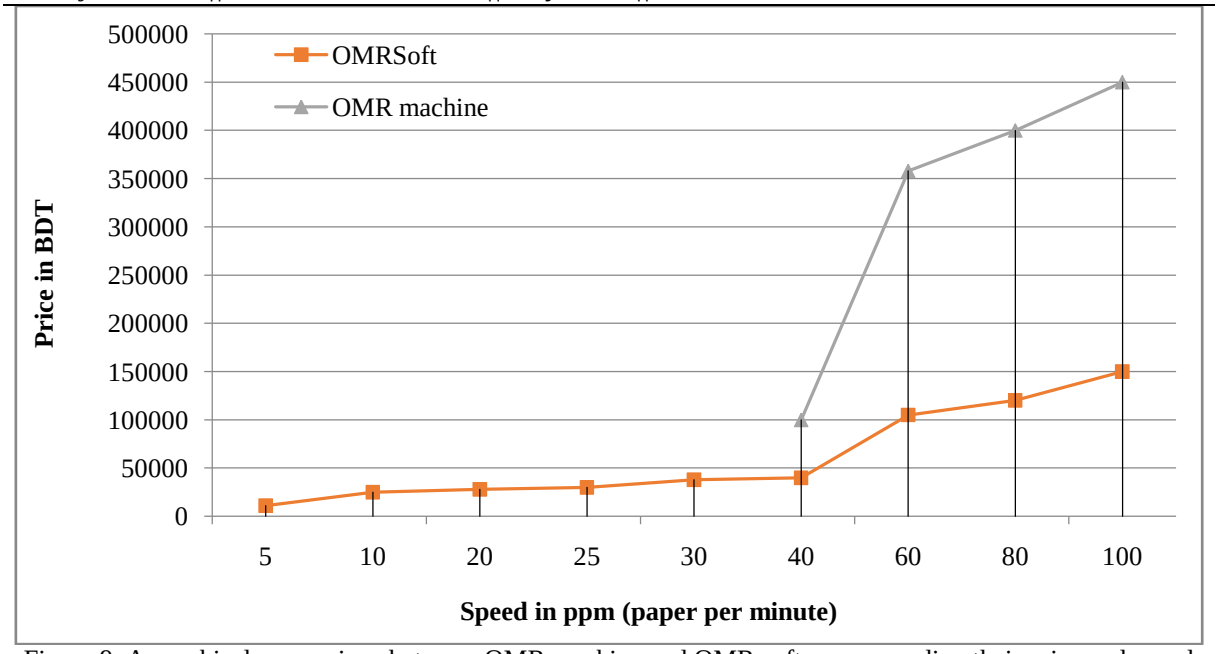


Figure 9. A graphical comparison between OMR machine and OMR software according their price and speed

OMR scanner are big machines manufactured outside Bangladesh, spares are difficult to get and vendors charge as per their desire as no manufacturer is from Bangladesh. OMR scanner reads standard OMR form. A decent OMR scanner with 2,500 scan/hour can cost somewhere around Tk. 4-5 million. Maintenance and support is a pain. No image is made for reference of the scanned form in most of OMR scanner. Up gradation is also difficult. OMR software can use normal document scanner which is made by renowned brands like Canon, Kodak, Fujitsu etc. OMR software reads standard OMR form and plain paper form[9]. An OMR software and document scanner that scan 2,500 per hour can cost around Tk. 60,000 Maintenance and support can be taken easily as software is manufactured in Bangladesh. An image is made of the OMR[10] form, with facility of image cropping is also available. Up gradation is easy in our software. A graphical view of the comparison of OMR machine and OMR software is given below in figure 9.Here we see that the price of OMR machine is so high in accordance with the same speed of OMR software. Now our world going to be digitalized so all our system should be software oriented. So we think, it is a good approach for this goal.

VI. CONCLUSION

We have proposed a simple and cost-effective software tool to evaluate OMR forms with high accuracy and execution speed. This system can easily replace current day heavy machinery which consists of expensive dedicated scanners that achieve the same objective with multiple dependencies and strict constraints. The proposed system uses normal printer and scanner with no added cost, and allows the forms to be printed on normal paper, without any constraints on quality or colour [9].User is provided with an easy interface, which helps him to design his own forms or modify existing ones, and evaluate and store the filled forms in an Excel spreadsheet. The system is designed to cater the need of vast population, also it is easy to install and use. Experimental results clearly depict the robustness and correctness of the proposed system, proven to be invariant of colour, brightness and affine transformations. Another dimension to the utility of the software is that the program is made to execute in parallel using multiple threads and implements pipelined architecture, thus improving performance and processing capability. The proposed system finds its application in fields such as institute examinations, voting, lottery, community surveys, etc. The idea can also be extended to implement the system through a website and make this utility available over the internet, where authenticated users can login, design forms and get the filled.

REFERENCES

[1] J. F. Canny, "Finding edges and lines in images." DTIC Document, Tech. Rep., 1983.
[2] S. J. K. Pedersen, "Circular Hough Transform," *Aalborg University, Vision, Graphics, and Interactive Systems*, 2007.
[3] L. Pan, W.-S. Chu, J. M. Saragih, F. De la Torre, and M. Xie, "Fast and robust circular object detection with probabilistic pairwise voting," *Signal Processing Letters, IEEE*, vol. 18, no. 11,pp. 639–642, 2011.
[4] K. Chattopadhyay, J. Basu, and A. Konar, "An efficient circle detection scheme in digital images using ant system algorithm," *arXiv preprint arXiv:1106.0962*, 2011.
[5] J. H. Han, L. T. Koczy, et al(1993) "Fuzzy Hough transform," [Proceedings 1993] Second IEEE International Conference on Fuzzy Systems, San Francisco, CA, USA, 1993, pp. 803-808 vol.2..
[6] J. Wu, K. Chen, and X. Gao, "Fast and accurate circle detection using gradient-direction based segmentation," *JOSA A*, vol. 30, no. 6, pp. 1184–1192, 2013.
[7] R. Gonzalez, "Fast line and circle detection using inverted gradient hash maps," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on. IEEE*, 2015, pp. 1354–1358.
[8] E. Cuevas, F. Wario, V. Osuna-Enciso, D. Zaldivar, and M. Perez-Cisneros, "Fast algorithm for multiple-circle detection on images using learning automata," *IET Image Processing*, vol. 6, no. 8, pp. 1124–1135, 2012.
[9] S. Rakesh, K. Atal, and A. Arora, "Cost effective optical mark reader," *International Journal of Computer Science and Artificial Intelligence*, vol. 3, no. 2, p. 44, 2013.
[10] Cheong Loke, Seng, et al, (2018). A new method of mark detection for software-based optical mark recognition. PLOS ONE. 13. e0206420. 10.1371/journal.pone.0206420.
[11] Erdal Tümer, Abdullah,et al,. An Image Processing Oriented Optical Mark Recognition and Evaluation System. International Journal of Applied Mathematics, Electronics and Computers. 6. 59-64. 10.18100/ijamec.2018447788, 2018
[12] Baghdadi, Ammar. An Automatic System to Grade Multiple Choice Questions paper based exams. Journal of university of Anbar for Pure science. 3. 174-181, 2009.
[13] Yadav, Virendra. Approach to Accurate Circle Detection: Circular Hough Transform and Local Maxima Concept, 2013.