



International Journal of Advance Research, IJOAR .org

Volume 2, Issue 11, November 2014, Online: ISSN 2320-9186

A NOVEL APPROACH TO A DYNAMIC TEMPLATE GENERATION ALGORITHM FOR MULTIPLE-CHOICE FORMS

Neha Raina, Prachi Sonawane, Mayur Nikam

Author Details (optional)

Neha Raina is currently pursuing bachelor's degree program in Computer Engineering in University of Pune, India. E-mail: nehanishuraina023@gmail.com

KeyWords

Computer Based Testing, Evaluation, MCQ, Paper Based Testing, Template Genration, Tests, etc

ABSTRACT

A dynamic template generation algorithm for MC test forms is used. The basis here is that the first step to take in any OMR machine or software when marking MC test forms is to get the template. The template in this case would be a map of the form, detailing where the answer fields (answer regions) are located. Therefore the final template would consist of only the answer regions, noting their positions and type.

1. INTRODUCTION

Multiple-choice (MC) forms have become the dominant, fast, reliable and the easiest tools of assessment in learning environments such as schools. Present automatic grading systems for such forms, though fast and accurate are either rated as expensive due to materials and equipments they require for their operation or can only operate on a particular set of forms or papers. The empty sheet is first scanned into digital form and then normalized to a given size. To consistently remove non-relevant objects such as lines from the image is proposed here. Finally the true answers are identified and score analysis performed.

Recently, starting from small pre-schools up to the academia, but also in other activities such as market survey, census, and elections. In almost all part of the world, MC tests are being promoted as “objective” and are considered to be the best way of assessing students’ performance in exams. Due to risks of damaging or reprocessing the scripts more than once due to human errors such as mistakes, inexperience, and tiredness manual grading method seem illogical in terms of time, cost, and energy. This in turn impacts heavily on the accuracy. Thereby, many issues are generated in existing mechanism for MC Analysis.

In the world today, the use of multiple-choice (MC) forms has not only dominated the learning environment, starting from small pre-schools up to the academia, but also in other activities such as market survey, census, and elections. Other reliable and efficient means of performing MC tests is by doing it online. The tests are conducted and assessed in real time on a computer system (PC), and in some cases, over the internet. Here, one can conclude that these online systems utterly depend on the availability of a good working PC and, in the latter case, an excellent reliable internet connection. Moreover, the system is not cost effective when dealing with a large number of candidates. In a recent attempt to address the limitations associated with MC test grading systems a camera based MC test grading system was reported. It does not require any special type of paper nor scanner or machine, the answer script need only be captured by a digital camera and then processed. But like other predecessor systems and software mentioned, it does not generate templates from other types of forms and thus will not work on forms produced either by other similar systems or by other means.

The proposed system is a solution to present automatic MC test grading systems in overcoming some of those limitations by adopting the use of a template. This will enable the system to become more flexible and detect the score in an accurate format and thereby facilitate better analysis.

2. LITERATURE SURVEY

S.Bani-Ahmad and A. Audeh [1] presented paper Re-engineering of Multiple-choice Examination form Production Tools: Cost Effective and Quality- Assurance Approach in the year 2010. This study aims at designing and implementing an exam-form generation software tool (named ex-pro) for Multiple-Choice-Based (MCB) exams. The study is motivated by the fact that student number in Jordanian universities is continuously growing of educational resources (instructors, labs, etc). A result of this situation is having large number of students in class-rooms. Consequently, providing and using online-examining system could be intractable and expensive alternatively, paper-based MCB tests can be used. The design and evaluation of ExPro is done by considering a basic set of design principles that are based on a list of Functional Requirements (FRs). Deriving those FRs is made possible by developing ExPro using the iterative and incremental model from software engineering domain. The paper shows that ExPro proves helpful to instructor in preparing multiple-choice tests. Further, ExPro makes archiving previous exams possible and effective to search for in future. The ExPro is available for free over the Internet and has been in use. ExPro users agree upon that ExPro is easy to be learned and used and proves to be a cost-effective alternative to online examining system. Cheating results are having unrealistic exam grades that falsely reflect student's achievement. Consequently, the quality of the education system is declined and overtime cheating corrupts the mind of the students. The advantage of this system is that instructors find it more convenient for themselves and more fair to students to use MCB tests instead of FRB tested. If instructors are well trained to prepare MCB tests and the quality of items is assured, MCB can be an effective assessment technique. It reduces human effort. The confidentiality and personality is maintained and the tests can be of any language. The limitations of this system are PB (pencil based) MCB tests are best choice over CB (computer based) MCB tests when having large no of students setting for exam. In fact, in some circumstances, CB tests can be prohibitively expensive. Another reason why PB tests are preferred over CB tests is that the test takers prior computer experience may affect their performance. Multiple choice questions are more difficult to construct as they are relatively very sensitive to construction mistakes and thus require much of constructor's focus. The most important limitation is that it requires a particular set of paper for multiple choice tests.

Juan YANG Xibin HAN Qian ZHOU [2] presented the paper, The Design and Development of a Semi-auto Computer Generated Testing Paper System a Case Study in the school of continuing education at China University of Geosciences in the year 2011. They proposed that with the rapidly increasing entrance examinations for the School of Continuing education of CUGB, there is an urgent need of the examination papers that have the parallel testing ability for the same academic program. Based on the analysis of parallel factors and the comparison of the manual generated mode and the computer generated mode, the most time-

consuming and experience-consuming steps are found. By providing supports to those steps automatically, a semi-auto computer generated testing paper system was designed and developed for reducing the workload of the experienced faculties to generate testing papers. The papers were generated based on a gradually constructed item-bank. The system offered multiple templates of the testing paper structure plan for assisting semi-auto paper generated procedure and guaranteeing the parallel of the generated papers. In order to construct an item-bank in a short time, an easy-to-use authoring tool with powerful functions was designed and developed. The system supported the spring entrance examination of SCECUGB in 2010 successfully. Faculties gave positive feedback toward the system. Computer generated testing paper system is a subsystem of a Computer Based Testing(CBT) System. Computerized tests are classified into two categories in the literature (e.g. Lilley,2004; López Cuadrado,Pérez,Vadillo and Gutiérrez,2010): computer-based tests (CBTs) and computerized adaptive tests (CATs). The former ones are computerized fixed form tests which are similar to pen-and-pencil tests. No matter based on CBTs or CATs, researches on the computer generated testing paper system usually focus on the algorithm optimization to achieve the fast and precise retrieval. The advantage of this system is that computer generated paper system can solve the problem for its convenient, safe and economic functions . A semi-auto computer generated testing paper system is designed and developed, trying to provide supports assisting paper generation. No plug-ins installation is required in this system. Text-graphics integration is supported in this system and High performance in functions, user-familiar interface are one of the advantages of this system. The limitations of this system are that it is hard to develop, it has poor performance in text-graphics, integration limited functions are there and a plug-in based on MS Word installation is required.

JunhongJi ,Guodong Chen, Lining Sun [3] presented the paper A Novel Hough transform method for Line Detection by Enhancing Accumulator Array in the year April 2011. In this paper, an improved Hough transform (HT) method is proposed to robustly detect line segments in images with complicated backgrounds. The work focuses on detecting line segments of distinct lengths, totally independent of prior knowledge of the original image. Based on the characteristics of accumulation distribution obtained by conventional HT, a local operator is implemented to enhance the difference between the accumulation peaks caused by line segments and noise. Through analysis of the effect of the operator, a global threshold is obtained in the histogram of the enhanced accumulator to detect peaks. Experimental results are provided to demonstrate the efficiency and robustness of the proposed method. Its major advantages include dealing with noise, degradation and partial disconnection and ease of realization. Many improved HT methods have been proposed to extract straight lines, circles, ellipses and natural shapes. Usually, a HT method consists of the following phases: voting, peaks localization, determining the actual parameters and verification. Compared to other methods, there is no parameter tuning needed to adapt to different images. A time-consuming task is avoided and the method can be implemented more autonomously and the effect is less influenced. The limitations of this system are noise and

partial disconnections are there without any prior knowledge. The major difficulty is that the flat part of the enhanced histogram is quite wide.

Aliyu Muhammad Abdu, Musa Mohd Mokji [4] presented the paper A Novel Approach to a Dynamic Template Generation Algorithm for Multiple- Choice Forms in the year Nov 2012. They proposed that multiple choice (MC) have become the dominant, fast, easiest and reliable tools of assessment in learning environments such as schools. Present automatic grading systems for such forms, though fast and accurate are either rated as expensive due to materials and equipment they require for their operation or can only operate on a particular set of forms or papers. Part of the main reasons this is so is due to the inability of such systems to generate a template from an unknown form or another form produced by similar system. As a novel step to creating a more dynamic and efficient automatic MC test grading systems, in this paper a template generation algorithm capable of generating a template from any type of bubble MC test form .The empty sheet is first scanned into digital form and then normalised to a given size .To consistently remove non relevant objects such as lines from the image, a unique combination of Hough Transform and region of interest (ROI) is proposed. Finally, the true answer regions are segmented using region processing by calculating their sizes and positions, this result is final template. The preliminary results showed that the algorithm is dynamic with overall high precision, accuracy even at low resolutions. The system consists of three main stages document registration (pre-processing), template generation and answer extraction. The template generation region is the most crucial part of the system and also the highlights of this paper; it is the stage where a template from the unknown form is generated. The proposed algorithm consists of two main stages: At the first stage a line detection algorithm is utilized to detect any presence of unwanted lines, note their positions and forward it to specially designed it to line removal algorithm for detection . At the second stage, circulatory formulae is used to estimate the type of answer region used. The main advantage of this system is that it is cost effective and the major advantage of this system is that it can operate on any set of papers. The experimental results have demonstrated that the proposed algorithm is dynamic and efficient and is also effective at different resolutions as well.

Mohini and Dr. Amar JeetSingh[5] presented the paper Automated Data Capturing And Recognition Technology To Minimize Human Intervention In University Examination in the year December 2012. System is presented by manual data entry from hand written data is very time consuming, tedious and prone to many errors more so where there is bulk amount of data is involved. This is the area where there is a requirement of automated data capturing and recognition technology. Using a recognition engine to convert text or handwriting from the printed page into computer readable characters saves up to 90 percent of the time it would take to enter the information manually. To fetch the data with minimal human intervention there are three types of recognition engines for this purpose and these are Optical Character Recognition (OCR), Intelligent Character Recognition (ICR) and Optical Mark Reader (OMR). Generally OCR is best for machine print, or type, ICR is

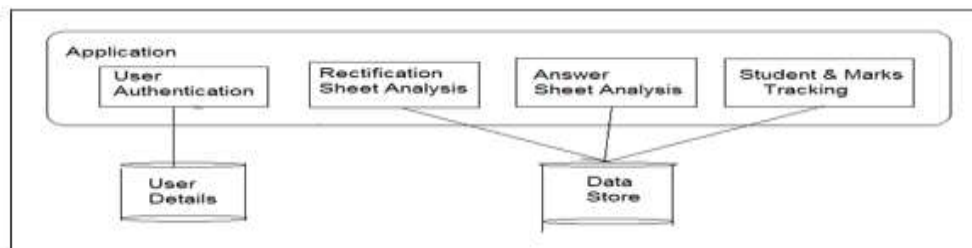
better for converting handwriting data and OMR is best suitable to detect the absence or presence of a mark, but not the shape of the mark. Automated data capturing is rapidly becoming an integral and necessary component in any organization. This not only saves the cost but also increases the speed and accuracy over manually entered data. This paper report on the study of use of these technologies in various processes of examination system especially in fetching awards from OMR technology to get almost hundred percent accuracy. Advantage is that ICR technology is seemingly a good machine facility to human operators to minimize their data entry time, decrease human drudgery and increase overall productivity. Disadvantage is that ICR mode of data fetching is not 100 percent accurate as it is not able to recognize poor handwriting.

Garima Krishna, Hemant Ram Rana, IshuMadan, Kashif, NarendraSahu[6] presented the paper Implementation of OMR Technology with the Help of Ordinary Scanner in the year April 13. In this fast and furious time, people don't want to invest their time in processing. They just want to give an input and take an output immediately. This project is actually an application that we are developing which enables implementation of OMR using an ordinary scanner. It provides tools to the user to design an OMR sheet based on the layout they want. The design of the sheet will be stored as image file format (JPEG). The user can take as many print outs as required, distribute it among others from whom information is desired, and get the filled sheets scanned. The scanned image files will then be provided as input to the software, processing will be done, value of filled fields will be extracted and then the data will be manipulated as instructed by the user. The implementation is done using Java. We are using Net beans IDE so as to easily manage the graphical layout and all the resources. For the GUI java was chosen mainly because of its platform independence and lightweight of the components in java swing. Advantage is to reduce the cost of OMR scanner. Any ordinary scanner will work with this application. The application can work on damaged sheets to a certain level. Damage may be due to inclusion of stains in the sheet or wetting of sheets or degradation due to use of improper scanner. Disadvantage is that to find the corner points we need to design algorithm for those cases when the sheet is not being kept straight during scanning, due to human error. In those cases rotation up to 11 degree is only acceptable.

Shalin A. Chopra, Amit A. Ghadg, Onkar A. Padwal, Karan S. Punjabi, Prof.Gandhali S. Gurjar [7] presented the paper Optical Character Recognition in the year Jan 14 . At the present time, keyboarding remains the most common way of inputting data into computers. This is probably the most time consuming and labour intensive operation. OCR is the machine replication of human reading and has been the subject of intensive research for more than three decades. OCR can be described as Mechanical or electronic conversion of scanned images where images can be handwritten, typewritten or printed text. It is a method of digitizing printed texts so that they can be electronically searched and used in machine processes. It converts the images into machine-encoded text that can be used in machine translation, text-to-speech and text

mining. This paper presents a simple, efficient, and less costly approach to construct OCR for reading any document that has fix font size and style or handwritten style. To achieve efficiency and less computational cost, OCR in this paper uses database to recognize English characters which makes this OCR very simple to manage. Advantage is the computed results are stored in a spread sheet, thus making it easy to understand and interpret the information. The system offers flexibility to the users, which allows designing and printing of forms on simple ordinary sheets, without the aid of any special pre-processing and color constraints. The substitution of special or expensive dedicated high computational machines by ordinary scanner and computer with no stringent requirements on paper quality makes the system cost effective. Disadvantage is if practically any scanner is not perfect; the scanned image may have some noise. This noise may be due to some unnecessary details present in the image.

3. SYSTEM ARCHITECTURE



3.1 SYSTEM FEATURES

Login & Registration: This module facilitates the system to authenticate the user using the system and thereby facilitates only valid users to enter into the system.

3.1.1 Template generation: This module will enable the admin to generate the template (correct answer sheet) and store in the system and thereby, enable identification of the actual answers for validating the answers from various sheets.

3.1.2 Answer Sheet Analysis: This module will facilitate identification of answers specified in the sheet and accordingly, calculation of the final score based on the comparative analysis of the specified sheet with the Correct Sheet. Thereby, this module will utilize certain image processing techniques and thereby filter out the unwanted data and successfully analyse locations of various answer points and decision making of the total score.

3.1.3 Student Identification: This will facilitate identification of the student ID related details from the answer sheet analysis and thereby facilitate its identification and thereby storage of score based details along with the ID in the tracking module.

3.1.4 Marks Tracking & Result Generation: This module facilitates the storage of the total marks analysed by the previous modules and therefore, generation of the final result sheet specifying the overall test result view of the particular answer sheet analysed.

4. PROPOSED ALGORITHMS

4.1.1 Algorithm 1: A line detection algorithm is utilized to detect any presence (removing non-relevant objects), of unwanted lines, note their positions and forward it to a specially designed line removal algorithm for deletion.

4.1.2 Algorithm 2: A circulatory formula is used to estimate the type of answer region used (in this case between circular and rectangular answer regions); the result is then used to segment the regions.

Conclusion

Acknowledgment

The author wishes to thank everyone who was involved in active research.

References

- [1] Izuagba, Angela Chinasa; Afurobi, Ada O; "Quality Education through Innovation: Examples of Tertiary Institutions in Nigeria"; *European Journal of Social Sciences*, vol. 10, Number 4th, 2009, pp. 605 – 613.
- [2] Martin Bush, "Alternative Marking Schemes For On-Line Multiple-choice Tests"; *Proc. Conference on the Teaching of Computing*, Belfast, 1999. I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [3] Bani-Ahmad, S.; Audeh, A., "Re-engineering of multiple choice exam-form production tools: Cost-effective and quality assurance approach"; *Proc. International Conference on Line detection and segmentation accuracy of the proposed algorithm at four (4) different resolutions for the two form categories* 2012 IEEE International Conference on Control System, Computing and Engineering, 23 - 25 Nov. 2012, Penang, Malaysia 220 Educational and Information Technology (ICEIT), 2010, vol. 3, pp 260 264.
- [4] Fang Zhang; Jinhui Li; Hongwei Jang, "Design and implementation of computer examination system based on genetic algorithm"; *Proc. International Conference on Computer Science and Service System (CSSS)*, 2011, pp. 334 –337.
- [5] Aggarwal, N.; Karl, W.C., "Line detection in images through regularized Hough transform"; *IEEE Transactions on Image Processing*, 2006, vol. 15, Issue: 3, pp. 582 – 591.
- [6] Satzoda, R.K.; Sathyanarayana, S.; Srikanthan, T.; Sathyanarayana, S., "Hierarchical Additive Hough Transform for Lane Detection"; *IEEE Transactions on Embedded Systems Letters*, 2010, vol. 2, Issue: 2, pp. 23 – 26.
- [7] Bonci, A.; Leo, T.; Longhi, S., "A Bayesian approach to the Hough transform for line detection"; *IEEE Transactions on Systems, Man and Cybernetics; Part A: Systems and Humans*, vol. 35, Issue: 6, pp. 945 – 955.
- [8] Verdera, J.; Caselles, V.; Bertalmio, M.; Sapiro, G., "Inpainting surface holes", *Proc., International Conference on Image Processing*, 2003, vol. 2, pp. II - 903-906.
- [9] Sunhyoung Han; Vasconcelos, N., "Object-Based Regions of Interest for Image Compression"; *Proc. Conference on Data Compression*, 2008, pp. 132 – 141.
- [10] Hui Liu, "Solving Multiple Choices for English Exams by Web Search"; *Proc. International Conference on Business Computing and Global Informatization (BCGIN)*, 2011, pp. 573 – 576.