An Efficient Algorithm of Computerized Checking System for Hard Copy MCQs Test ($H_CM_{CQS}T$)

Tabassam Nawaz[†], *Khurram Ishfaq Qazi*[††], *and* *Muhammad Imran Ashraf*[†††]

[†]*Faculty of Software Engineering, University of Engineering & Technology Taxila, Pakistan*
[††] *Department of Software Engineering, University of Engineering & Technology Taxila, Pakistan*
[†††] *Department of Software Engineering, University of Engineering & Technology Taxila, Pakistan*

Summary:

A noval approach for checking of computerized multiple choice questions test ($H_CM_{CQS}T$) like in NTS (National Testing Service, Pakistan), Engineering Universities and Medical colleges Entry tests without any errors is presented in this world. In this world tests are conducted in two ways; computerized /soft copy and on paper / hard copy. The NTS and Entry tests in Pakistan are conducted on papers i.e. answers that are filled by the contestants are on papers. The contestant has to fill the circles on the paper. The circles are sometime filled properly and sometime not properly. This also contributes in increasing computerized checking with errors. The reasons behind this that machine mostly don't read the incomplete filled circles. When $H_CM_{CQS}T$ are checked computerized, there may be possibility $H_CM_{CQS}T$ paper may be damaged or bend or torn. In this condition checking done may carries some errors due to not proper reading of damaged area by computer. Thus final results don't achieve the 100% accuracy. For this reason, an algorithm is designed that will check the $H_CM_{CQS}T$ answer sheet paper with computer. The efficient algorithm will remove the noise if there is in the $H_CM_{CQS}T$ answer sheet paper, and then $H_CM_{CQS}T$ answer sheet paper will be checked. The results make it more attractive for educational application.

Key words:

Morphing Filter, Region Based Segmentation, Optical Character Recognition, Dilation, Erosion, Closing, Segmentation

1. Introduction:

In this whole world, tests are conducted to judge the abilities of the students and people. There are different categories of test conduction in this world. Mainly they are divided into two categories which are computerized and manual. In computerized systems there are two sub-categories either to conduct online by web or offline conducted on computers in the institutions.

There is no chance of error in this method. In manual systems some tests are conducted on papers and checked by hands manually by the checkers (teachers). The other way is that tests are conducted manually on papers and checked by machines. Entry Tests and NTS Tests in Pakistan are conducted on paper in a pattern which is shown in Fig. 1.

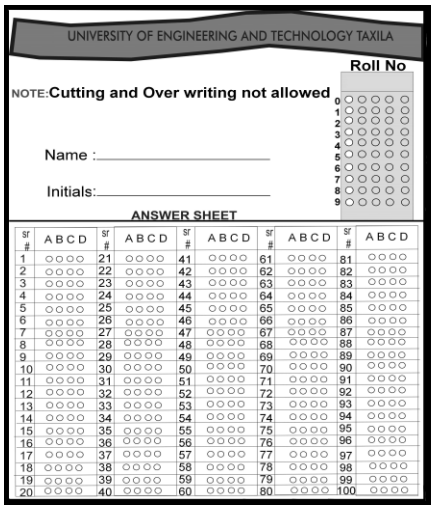


Fig. 1

The Fig. 1 is the $H_CM_{CQS}T$ answer sheet given to students or contestants for filling according to question paper given to them. They have to fill one option in correspondence to a question as shown in Fig. 2.

If a contestant fills more than one option, this will also create confusion for the computer that either to accept correct answer if filled both correct and incorrect options or to consider as wrong answer as shown in Fig.3. It is difficult for computer to first of all check whether more than one option is filled or only one option is filled. The reason is that, some times during computerized checking the option that undergoes checking is correct and other options after that are also filled and computer may skip other options and moves toward next question.

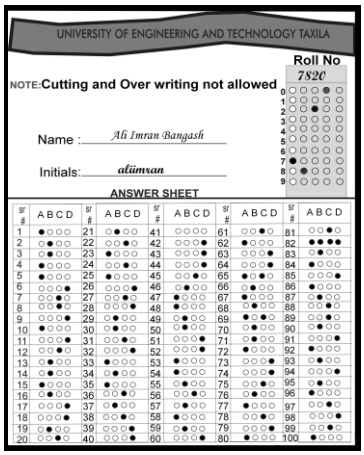


Fig. 2

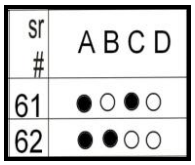


Fig. 3

In this case checking is not done 100% accurate because according to NTS rules double options filled is considered incorrect.

Sometimes the contestant fills all the options given on $H_CM_{CQS}T$ answer sheet to a particular question as shown in Fig. 4. The filling of all options also confuse the checker either to consider correct option or to consider it wrong.

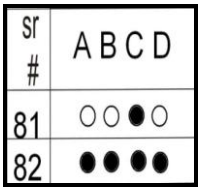


Fig. 4

Sometimes student don't fill any option as shown in Fig. 5

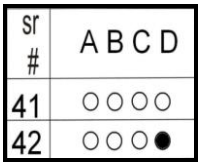


Fig. 5

Contestants sometime may fill the wrong circles while filling the roll number as shown in Fig. 6.

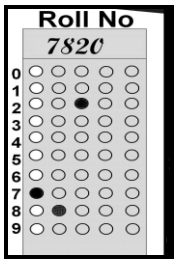


Fig. 6

Observing these conditions an algorithm is designed. This algorithm is designed for $H_CM_{CQS}T$ computerized checking. The $H_CM_{CQS}T$ that contestants completed is scanned. The original filled $H_CM_{CQS}T$ from which tests to be checked also scanned. The algorithm will remove noise from the $H_CM_{CQS}T$ paper if noise is present in the test that can decrease accuracy of the result.

The $H_CM_{CQS}T$ paper will undergo computerized checking. The algorithm used checks test by segmentation **[1]**. The original scanned test undergoes segmentation producing different sub-segments **[2]** which are then compared with the sub-segments of test to be checked.

1.1. *Test Conduction Rules*:

There are some rules for tests that must be followed. They are;

1. Fill the roll number circles carefully.
2. Cutting and overwriting not allowed.
3. Fill the circles properly and completely.
4. If all circles are unfilled, answer is considered wrong.
5. If more than one circle or all circles filled, answer is considered wrong.
6. For double or more than one filled answers, negative marking is done.
7. If only a correct circle is filled, full marks given to that corresponding question.

The hand written and circle filled roll number should be same.

2. *Method's Description:*

Actually there is computerized printing of roll numbers on $H_CM_{CQS}T$ paper as shown in Fig. 6. Before the $H_CM_{CQS}T$ test checking started, it is checked that computerized roll number entered on test sheet and circle filled roll number are same. This is done by Optical Character Recognition **[3].** It is one of best successful applications of hand written recognition. First of all smoothing is applied for better quality of image. The hand written roll number and circled filled area of roll number is extracted by applying region based segmentation [**11**] as shown in Fig. 7

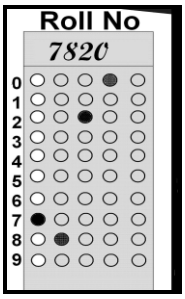


Fig 7

After the normalization [**4**] applied to all numeric digits of hand written roll number. Due to normalization, size of each digit becomes of same size. After the normalization, from hand written roll number, each digit is separated by using image segmentation [**12**] as shown in fig. 8

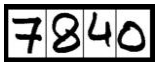


Fig 8

Each digit of hand written roll number is matched with the database of the digits. Actually there is a database of each hand written digit from 0 to 9 with 200 possibilities. Each digit has 20 possibilities of hand written in different styles as shown in fig. 9

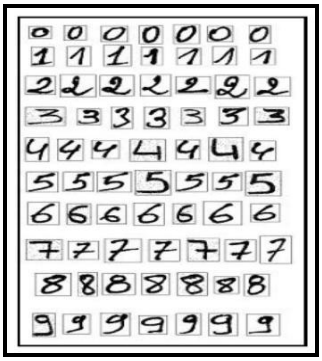


Fig 9

After digit recognition, filled circles are matched. If over writing in hand written roll number, than circle filled roll number is considered correct as shown in Fig.10

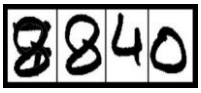


Fig 10

If over writing in circle filling for example two or more circles are filled, than hand written roll number is considered correct as shown in Fig. 11. If both hand written and circle filled roll number contains over-writing than paper is awarded zero marks as shown in Fig. 12

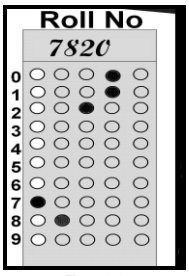


Fig. 11

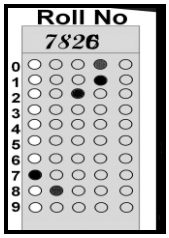


Fig 12

After that it is checked that $H_CM_{CQS}T$ paper is noise free and neither $H_CM_{CQS}T$ paper is bend nor torn for accurate checking. If $H_CM_{CQS}T$ paper is torn and during scanning it gets noise as shown in Fig. 13, the noise is removed by applying noise removal algorithm **[5].**

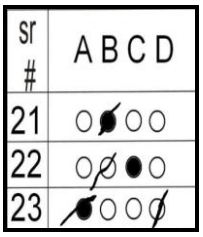


Fig. 13

After noise removal the $H_CM_{CQS}T$ paper becomes clear as shown in Fig. 14

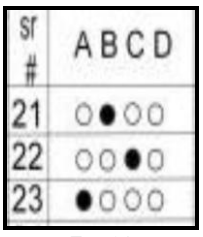


Fig. 14

After noise removal $H_CM_{CQS}T$ paper undergoes checking. The checking is done by Segmentation **[3].** In Segmentation, $H_CM_{CQS}T$ paper undergoes region based image splitting. **[6].** In this $H_CM_{CQS}T$ paper is split into two major regions as shown in Fig. 15 and Fig. 16.

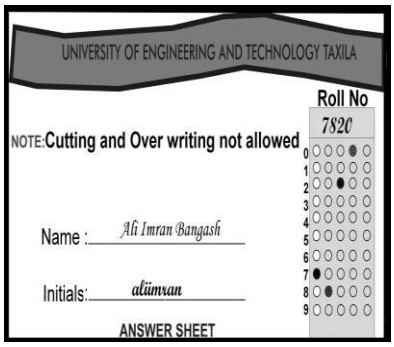


Fig. 15

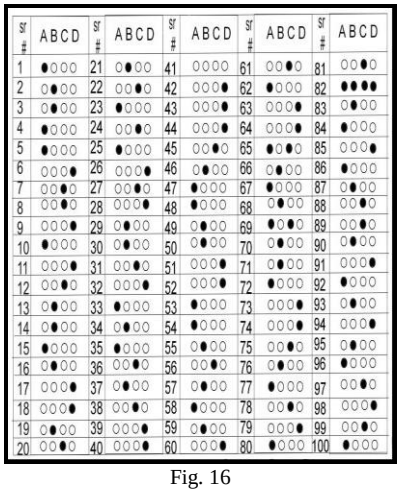


Fig. 16

Each region is given a unique ID. Second region that contains the filled circles region further undergoes region based splitting **[7]** and split into five regions as shown in Fig. 17.

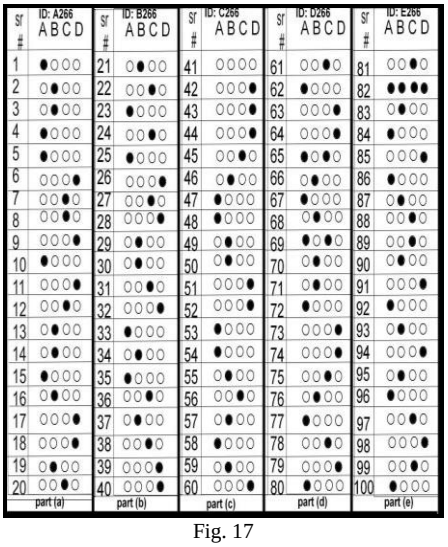


Fig. 17

First splitting was done horizontally. The Second splitting is done vertically column wise. Each column undergoes further region based splitting, splitting each column into 25 sub-regions or segments **[8]**. Now each sub-segment contains all four circle options for one question as shown in Fig. 18.

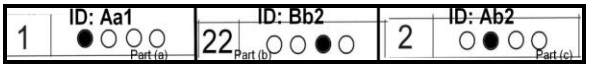


Fig. 18

Same process is applied on the master $H_CM_{CQS}T$ answer sheet from which answers are to be compared. After that, each segment of master $H_CM_{CQS}T$ answer sheet is matched and compared with the segment of answer sheet filled by contestant. The master copy $H_CM_{CQS}T$ is shown in Fig. 19 and test completed by contestant is shown in Fig. 2.

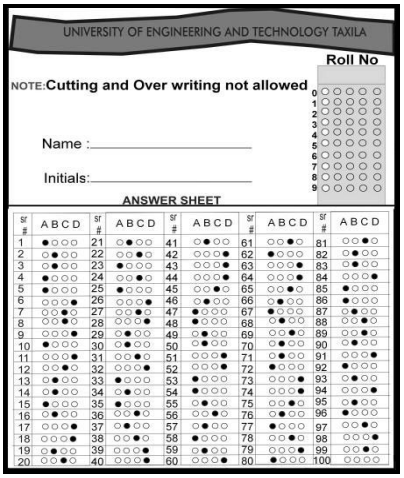


Fig. 19

After region based splitting of both master and contestant filled $H_CM_{CQS}T$ copy, the resultant regions undergo the split test **[9]**. In split test it is tested whether the required region is split or not. After the split test, the comparison is carried out **[10]**. During the comparison, each region of $H_CM_{CQS}T$ master copy is compared with corresponding region of contestant filled $H_CM_{CQS}T$ test copy as shown in Fig 20.

Fig. 20

Actually during the region splitting, each split region is given a unique ID. After comparison if both are matched, ID is given 1 else if all empty ID is given -1 else zero as shown .One (1) means answer is correct and zero means answer is wrong and negative marking for that particular question and -1 means wrong answer with no negative marking as shown in Fig. 21.

Fig. 21

If all options or more than one options are filled than ID is also given zero. The reason is that more than one filled or all options filled are wrong according to NTS rules.
During matching or comparison if more than one circle is filled or wrong circle is filled, negative marking will be done against that particular question.
After processing this, result is calculated. Result is calculated by calculating number of minus 1s (-1), zeros and 1s from the unique IDs. When all calculations had done, result is saved in the database and displayed on the screen.

3. Implementation:

The algorithm describe is implemented. The working of algorithm is shown in flow diagram which shows how whole procedure is carried out as shown in Fig. 24.

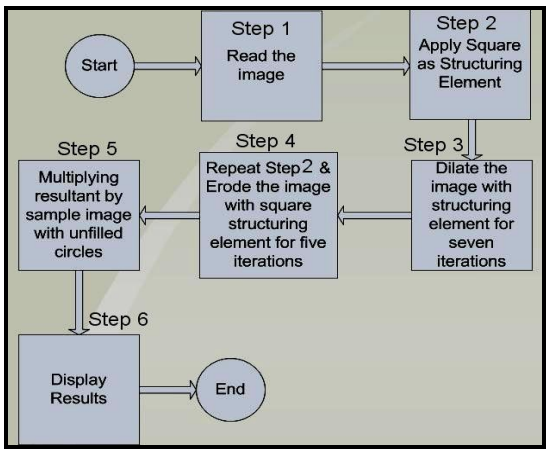


Fig. 22

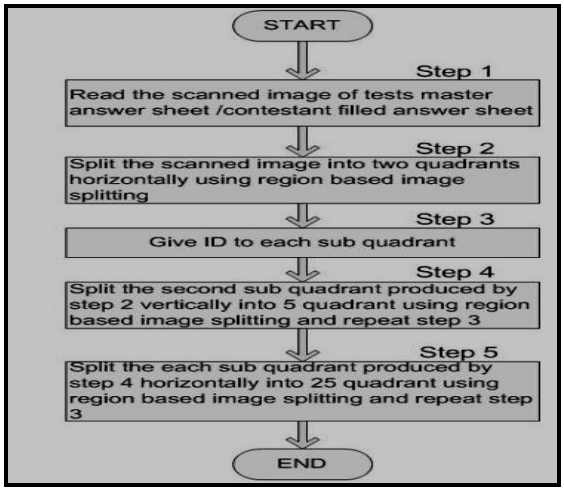


Fig. 23

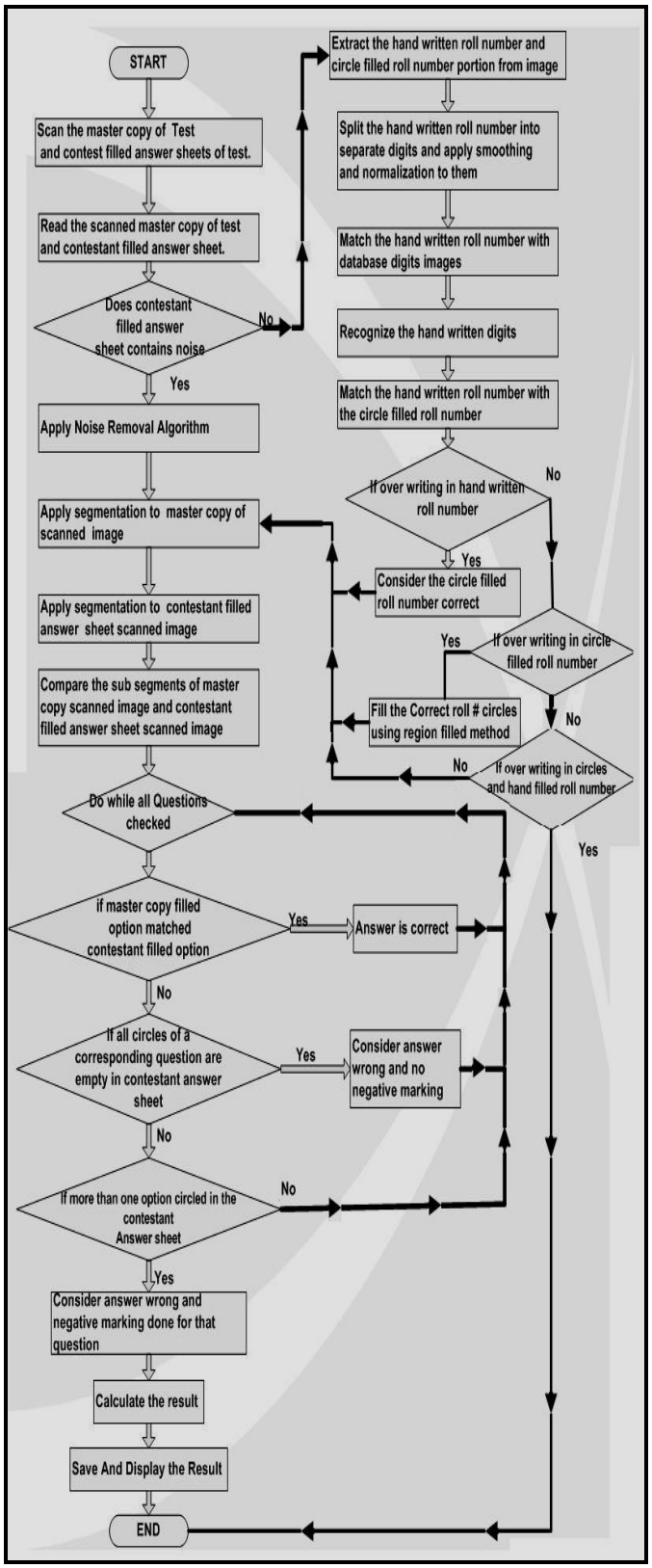


Fig. 24

The Fig. 24 tells that if noise is present in the scanned image due to bending of paper or other external effects, than noise removal algorithm **[5]** is applied. Fig. 22 shows how noise removal algorithm is applied to scanned image. There are steps of contestant and master $H_C M_{CQS} T$ copy scanned image splitting which are shown in Fig. 23.

3.1. Pseudo Code:

```
Main ()
{1
    Scan master copy and all contestant filled copies of test from the scanner.
    Get and Read the master copy and contestant filled copy of test scanned image.
    //Roll number checking
  Read the roll number filled area field by retrieving specific pixels.
  Split the roll circle filled area by region based image splitting.
  Read the hand written roll number area by retrieving specific pixels.
   Split each digit of hand written roll number by region based image splitting and normalize each digit image.
 Read the filled circles for roll number.
  Match the hand written roll number with the circle filled roll number.
    If (over writing in hand written roll number)
{2
Consider circle filled roll number correct
2}
Else
    If (over writing in circle filled roll number)
{3
Consider hand written roll number correct
3}
Else
{4
Award paper zero marks and end the paper processing.
4}
    // noise removal
    Match:
    If (contestant image has noise)
     {5
          {6
        Dilate the image by structuring element square and value 3
        Do seven times
        6}
  If (circles size decreased too much)
        {7
        Erode the image by structuring element square and value 4 for two times
        Erode the image by structuring element square and value 6 for four times
        7}
        Multiply the sample image with resultant image pixel by pixel.
    5}
        // image segmentation of contestant filled test answer sheet
        If (noise removed)
        {8
        Split the contestant filled test copy into two quadrants vertically. (//splitting on the basis of quadrants and this splitting is called region based splitting)
        Give unique ID to each quadrant.
        Identify the middle circle filled quadrant/ region which is second quadrant.
        {9
        Split the resultant second quadrant into 25 sub region horizontally.
        Give unique ID to each sub quadrant.
        9}
        // image segmentation master test copy
        Split the master test copy into two quadrants vertically. (//splitting on the basis of quadrants and this splitting is called region based splitting)
        Give unique ID to each quadrant.
        Identify the middle circle filled quadrant/ region which is second quadrant.
        {10
        Split the resultant second quadrant into 25 sub region horizontally.
        Give unique ID to each sub quadrant.
        10}
        // comparing or matching.
        Do while all questions completed
        {11
        Match the master copy of test answer sheet with the contestant filled answer sheet.
        If (master copy answer sheet region matches with the contestant filled answer sheet region)
        Than
        {12
         Add 1 to CA variable
        //CA stands for Correct Answer and it counts total number of correct answers
        12}
        Else
        If (All circles unfilled in the contestant filled answer sheet and one circle filled in master copy answer sheet)
        Than
        {13
        Add -1 to EC variable
        //EC stands for empty circles; it counts total number of questions whose answers are empty.
        13}
        Else
        If (more than one circle or all circles of contestant answer sheet filled)
        Than
        {14
        Add 1 to MOCF variable
        //MOCF stands for More than One circle Filled
        14}
        Else
        If (wrong circle filled)
        Than
        {15
        Add 1 to WC variable
        //WC stands for Wrong circle filled
        15}
        11}
        8}
```

//Noise Free Image

Else

Do while all questions completed

{16

Split the contestant filled test copy into two quadrants vertically. (//splitting on the basis of quadrants and this splitting is called region based splitting)

Give ID to each quadrant.

Identify the middle circle filled quadrant/ region which is second quadrant.

{17

Split the resultant second quadrant into 25 sub region horizontally.

Give ID to each sub quadrant.

17}

// image segmentation master test copy

Split the master test copy into three quadrants vertically. (//splitting on the basis of quadrants and this splitting is called region based splitting)

Give ID to each quadrant.

Identify the middle circle filled quadrant/ region which is second quadrant.

{18

Split the resultant second quadrant into 25 sub region horizontally.

Give ID to each sub quadrant.

18}

// comparing or matching.

{19

Match the master copy of test answer sheet with the contestant filled answer sheet.

If (master copy answer sheet region matches with the contestant filled answer sheet region)

Than

{20

Add 1 to **CA** variable

//CA stands for Correct Answer and it counts total number of correct answers

20}

Else

If (All circles unfilled in the contestant filled answer sheet and one circle filled in master copy answer sheet)

Than

{21

Add -1 to **EC** variable

//EC stands for empty circles; it counts total number of questions whose answers are empty.

21}

Else

If (more than one circle or all circles of contestant answer sheet filled)

Than

{22

Add 1 to **MOCF** variable

//MOCF stands for More than One circle Filled

22}

Else

If (wrong circle filled)

Than

{23

Add 1 to WC variable

//WC stands for Wrong circle filled

23}

19}

16}

Now calculate number of minus one (-1), zeros and 1s from sub segments IDs.

// note zeros represent wrong answers, -1 represent empty and 1s represent correct

Compile the result using EC, WC, MOCF and CA variables values.

Negative marking is done for WC.

Finalize the results.

Save result in the database.

Display the result

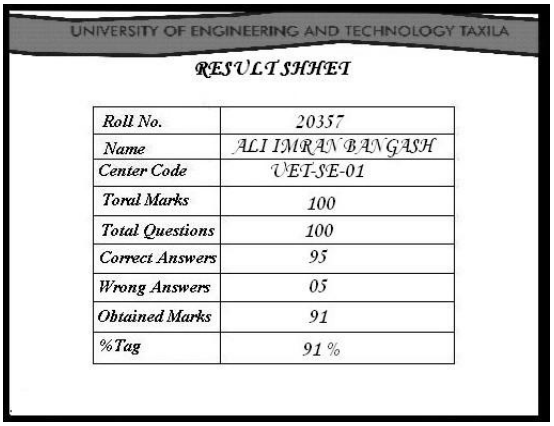
1}

4. Results:

When a contestant filled $H_CM_{CQS}T$ answer sheet as shown in Fig. 2 undergoes the above described algorithm, then it is observed that accurate test checking is done as it should be done manually.

The $H_CM_{CQS}T$ master copy from which contestant filled answer sheet is compared is shown in Fig. 19.

After calculation correct and accurate results were produced by proposed algorithm as shown in Fig 25.


UNIVERSITY OF ENGINEERING AND TECHNOLOGY TAXILA

RESULT SHHET

Roll No.	20357
Name	ALI IMRAN BANGASH
Center Code	UET-SE-01
Toral Marks	100
Total Questions	100
Correct Answers	95
Wrong Answers	05
Obtained Marks	91
%Tag	91 %

Fig 25

The $H_CM_{CQS}T$ filled by contestant has four mistakes. Table 1 shows the types of mistakes and errors checked by above proposed algorithm and method.

Table 1

SR. #	ERROR TYPES
1	All Filled circles.
2	More than one circles filled.
3	Wrong circle filled.
4	No circle filled.
5	Noise Removal from answer sheet

There is one question against which contestant filled all options in $H_CM_{CQS}T$ answer sheet. It is considered wrong according to test conduction rules. Contestant did not answer to one question and over-writing was done against one question in $H_CM_{CQS}T$ answer sheet. Over-writing answer is considered wrong. In the $H_CM_{CQS}T$ answer sheet, answer to one question was wrong. Negative marking done against,

- Two filled options in $H_CM_{CQS}T$ answer sheet against particular question.
- All filled options in $H_CM_{CQS}T$ answer sheet against particular question.
- Over writing and cutting in $H_CM_{CQS}T$ answer sheet against particular question.
- Wrong filled option in $H_CM_{CQS}T$ answer sheet against particular question.

Two marks were deducted for each above discussed four categories while doing negative marking. Whereas one mark deducted for unfilled options against a particular question in $H_CM_{CQS}T$ answer sheet. Fig. 26 shows the result sheet, in which 91 marks given to contestant showing 100% accuracy in $H_CM_{CQS}T$ checking.

Results show that method is fast and error free for the $H_CM_{CQS}T$ of NTS computerized checking. It's efficient and applicable method.

5. Conclusion:

We have presented an efficient, fast and appropriate algorithm for General hard copy MCQs Test with computerized checking. The algorithm not only checks the hard copy by computerized, but also removes noise if noise added to scanned test due to bend or if test torn due to some external effects. After noise removal the algorithm checks for the correct roll number entry, if hand written and circle filled mismatches, it first corrects it and than proceed further.

Fifty papers undergo the above applied algorithm and they computed results with 100% accuracy and efficiently. Observing results it is concluded that this algorithm produces 100% accurate test checking results. This algorithm is particularly attractive for educational applications.

References:

[**1**] Te-Won Lee And Michael S. Lewicki **"**Unsupervised Image Classification, Segmentation And Enhancement Using Ica Mixture Models**"** Ieee Transactions On Image Processing, Vol. 11, No. 3, March 2002

[**2**] Hua Li, Abderrahim Elmoataz Jaral Fadili Su Ruan **"**An Improved Image Segmentation Approach Based On Level Set And Mathematical Morphology**"** Third International Symposium On Multispectral Image Processing And Pattern Recognition

[**3**] Ilya Pollak, Alan S. Willsky, Fellow, And Hamid Krim, Fellow,"Image Segmentation And Edge Enhancement With Stabilized Inverse Diffusion Equations" Ieee Transactions On Image Processing, Vol. 9, No. 2, February 2000

[**4**] Salim Ouchtati, Mohamed Redjimi, Mouldi Bedda And Faouzi Bouchareb "A New Off Line System For Hand Written Digits Recognition " Asian Journal Of Technology, 2006

[**5**] Tabassam Nawaz, Khurram Ashfaq Qazi, M Imran Ashraf "Efficient Noise Removal Algorithm From A Torn Paper" 12th International Conference On Computer Graphics And Artificial Intelligence 3IA, 2009 Athens, Greece

[**6**] Fabiana Zama "Image Denoising And Segmentation Using Comsol Multiphysics" Excerpt From The Proceedings Of The Comsol Users Conference 2007 Grenoble

[**7**] Gabriele Peters And Jochen Kerdels "Image Segmentation Based On Height Maps" Proceedings Of The 12th International Conference On Computer Analysis Of Images And Patterns, Vienna, Austria, August 27-29, 2007.

[**8**] Mario A. Nascimento, Veena Sridhar, Xiaobo Li "Effective And Efficient Region-Based Image Retrieval" Journal Of Visual Languages And Computing

[**9**] Lifeng Liu And Stan Sclaroff_ "Region Segmentation Via Deformable Model-Guided Split And Merge" Boston University Computer Science Technical Report No. 2000-24, Dec. 2000

[**10**] Sinisa Todorovic And Narendra Ahuja "Scale-Invariant Region-Based Hierarchical Image Matching"19th International Conference On Pattern Recognition (Icpr), Tampa, Fl, December 2008

[**11**] Lifeng Liu And Stan Sclaroff "Region Segmentation Via Deformable Model-Guided Split And Merge" Boston University Computer Science Technical Report No. 2000-24, Dec. 2000

[**12**] H C Sateesh Kumar 1, K B Raja2, Venugopal K R2 And L M Patnaik "Automatic Image Segmentation Using Wavelets" Ijcsns International Journal Of Computer Science And Network Security, Vol.9 No.2, February 2009

[**13**] Mario A. Nascimento*, Veena Sridhar, Xiaobo Li "Effective And Efficient Region-Based Image Retrieval" Journal Of Visual Languages And Computing 14 (2003) 151–179

[**14**]Hakan C ¸ Aglar, Sinan Gu ¨Ntu ¨Rk, Bu ¨Lent Sankur, And Emin Anarim "Vq-Adaptive Block Transform Coding Of Images" Ieee Transactions On Image Processing, Vol. 7, No. 1, January 1998

[**15**] Rudolf Mester And Uwe Frank "Statistical Model Based Image Segmentation Using Region Growing, Contour Relaxation And Classification " Proc. Spie Cambridge Symposium On Visual Communication And Image Processing 88 (T.R.Hsing Ed.) Cambridge/Mass, November 1988

Tabassam Nawaz has PhD degree from University of Engineering & Technology Taxila, Pakistan in 2008. He got Masters Degree in Computer Sciences and in Computer Engineering in the years 2001 and 2005 respectively. His interest lies in the development of computer vision algorithms for vision based human machine interaction. He developed algorithms for gaze direction estimation and face recognition which were prototyped in human computer interaction system for the patients. He holds the position of Associate Professor at the Faculty of Telecommunication & Information Engineering, UET Taxila, Pakistan.

Khurram Ashfaq Qazi is in final semester of B.Sc Software Engineering at University of Engineering and Technology, Taxila, Pakistan. He has done research work in noise removal from scanned images. In 2008 he stayed in National Database and Registration Center to Study finger print or thumb impression image processing.

Muhammad Imran Ashraf is in final semester of B.Sc Software Engineering at University of Engineering and Technology, Taxila, Pakistan. He has done research work in noise removal from scanned images.