Automatic Correction of Multiple-Choice Tests using Digital Cameras and Image Processing

Francisco de Assis Zampirolli, José Artur Quilici Gonzalez, Rogério Perino de Oliveira Neves
Centro de Matemática Computação e Cognição (CMCC) – Universidade Federal do ABC (UFABC)
Av. dos Estados, 5001. Bairro Bangu. Santo André - SP - Brasil
{fzampirolli, jose.gonzalez, rogerio.neves}@ufabc.edu.br

Abstract. *There are currently many commercial solutions for automated scoring of Multiple-Choice Tests, usually composed of a software and a scanner, but the widespread dissemination of laptops, tablets and smartphones with built in cameras offers new possibilities for doing the same job with no need for any extra hardware. This article presents a simple and innovative method to transform captured images of answer sheets into reduced binary matrices containing answers to the questions plus some control elements, using simple morphological operations for segmentation. This methodology is applied to the real problem of automatic correction of Multiple-Choice Tests. Initially, the user positions the answer key sheet in front of the camera in order to save the image to the disk, then the image is gauged to evaluate the test type. Subsequently, student answer sheets can be read using the camera, having the test score displayed on the screen and/or saved to a file.*

Keywords. *Image Processing, Mathematical Morphology, Multiple-Choice tests, Optical Mark Recognition.*

Introduction

Image analysis can profit from strategies where the neighborhood relation between objects in the image, such as regions of answer sheets of multiple choice tests, can be represented simply by matrix. Furthermore, Mathematical Morphology is an elegant form to solve image-processing problems using consistent theoretical bases, that is the theory of sets [1].

The method described was first conceived to address one of the major problems in distance education. To avoid cheating and to verify the identity of the test taker, it is common to require the presence of all students in a test center in a specific date. However, for a large number of students, there might not be enough computer terminals available for an electronic test. In that case, the alternative is to apply printed Multiple-Choice Tests.

Methodology

We used an image database of 674 tests filled by candidates to the graduate program "Specialization in Information Technology", a distance learning modality course offered at UFABC, the test being part of the entry selection process. Each test was printed on one A4 size paper containing both the answer sheet and 24 multiple-choice questions with five alternatives each, only one being the correct answer. The model proposed in this paper supports up to 16 different test sets. We opted to design the answer area in this format for the sake of simplicity, so that any traditional word processor can be used to generate the Multiple-Choice Tests, unlike with traditional OMR forms, Fig. 1.

The developed code was written according to the following steps:

1. read the templates for each set;
2. save a TIFF image containing only the answer sheet area of the template;
3. save a two-dimensional matrix containing the correct answers to each set;
4. read the answer sheets from student tests;
5. save the image of each test to a separate file in TIFF format;
6. add the processed test results to a file in csv format.

To segment the answer sheet area of the test, we used Matlab in a MacBook 2,26 GHz, where each frame was captured and processed using commands: *im2bw*, *imopen*, *imclearborder* and *regionprops*. Using image processing techniques, such as region growing and other filters, image segmentation can be further improved in the algorithm to consider valid answers similar to those presented in Fig. 2.

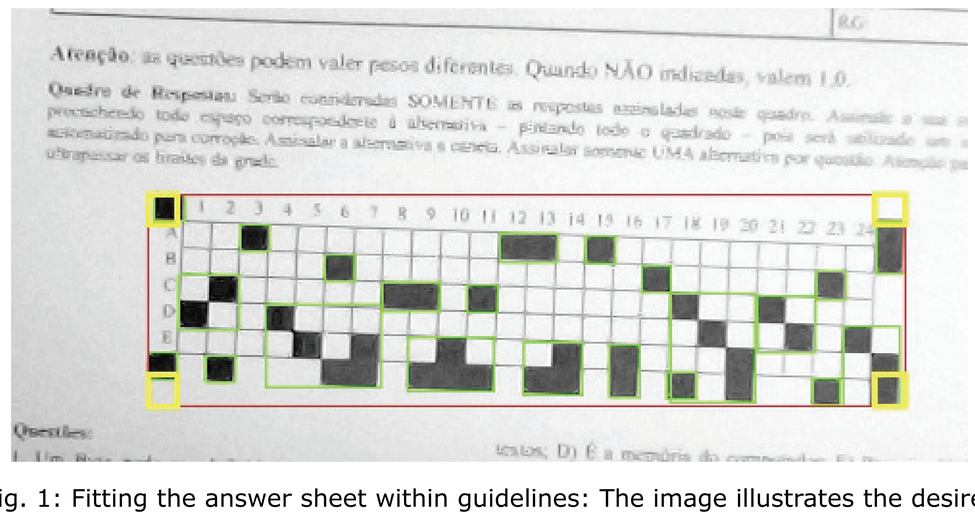

Fig. 1: Fitting the answer sheet within guidelines: The image illustrates the desired position.

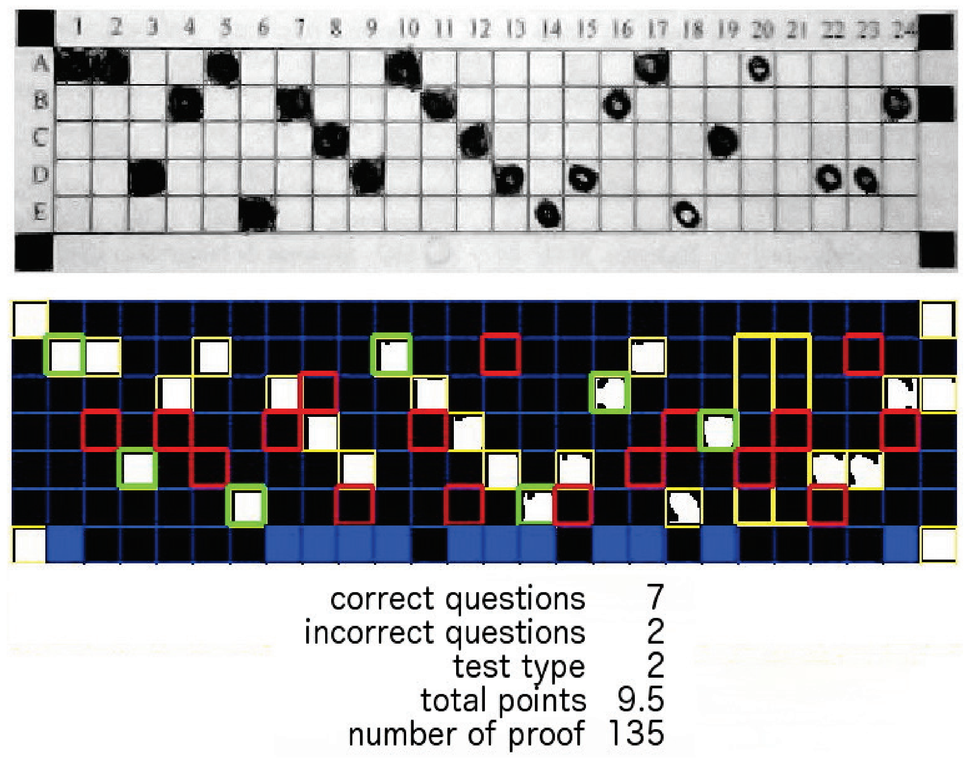

Fig. 2: The gray scale image shows a captured student test. The bottom image shows the test scoring, highlighting the correct answers in green and the wrong answers in yellow. The red squares highlight unmarked correct answers. Control elements are located in the borders, like the outer squares in the first and last column, which contains the four corner guides and the coded test type. The marked squares in the bottom line indicate questions with increased value. Finally, a summary is shown below the image displaying the number of correct answers, invalid answers, the test type, the final score and the test number.

Conclusion

In this work we have presented a system which uses techniques from mathematical morphology, capable of acquiring images from computer cameras (and eventually a myriad of alternative devices) to perform automated scoring of a Multiple-Choice Tests, presenting high accuracy in the results. The steps performed by the algorithm were explained in detail and illustrated by its employment in a real case scenario. The software is currently being ported to smartphones, what will allow teachers to travel to test centers carrying only printed tests with no need for additional resources, apply the tests, capture images and calculate scores in loco, being able to discuss the test results with the students on the same day.

References

[1] HEIJMANS, H. Morphological image operators. Academic Press, Boston, 1994.

[2] ZAMPIROLLI, F. Z., GONZALEZ, J.Q., NEVES, R. P. O. Automatic corrections of Multiple-Choice Tests . http://youtu.be/rhp4gfwrdf0, http://youtu.be/yRPbGS_YSmQ.

http://youtu.be/yRPbGS_YSmQ

Apoio: UFABC CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico