Second Balkan IFAC International Symposium
APPLIED AUTOMATIC SYSTEMS AAS'2000
21–23-th September, Ohrid, Republic of MACEDONIA

AAS 7–3

A SIMPLE SYSTEM FOR AUTOMATIC EXAM SCORING USING OPTICAL MARKUP READER

Gorgevic Dejan[1], Grcevski Nikola[2], Mihajlov Dragan[1]

[1]*Faculty of Electrical Engineering, SS. Cyril and Methodiys University,*
P.O.Box 574 Skopje, Republic of Macedonia, dejan@cerera.etf.ukim.edu.mk
[2]*SEMOS, Skopje, Republic of Macedonia*

ABSTRACT: Probably the most efficient way to test somebody's knowledge, about some specific area or to retrieve somebody's opinion is to make a test specially created for that purpose. The test processing is very resourceful and time-consuming operation, what is making automatic test processing procedures present for a long time. The test forms used for today's common test processing machines are generally expensive, have special form and are usually printed in color. These problems lead us to create a simple, efficient and inexpensive program for automatic test processing and exam scoring.

Index terms: optical markup reading, exams, test, scoring

1. INTRODUCTION

Knowledge testing and information retrieval was always important to people. The test as a form of a questionnaire has proven to be one of the most effective ways to retrieve somebody's knowledge. It is especially important to professors, teachers and other professionals when making regular exams, finals or any kind of testing. The test is a special form of questionnaire where the answers are scored according to some rules or regulations. Those questionnaires are to be in some manner suitable for later statistical analysis of the answers.

Test scoring is generally implemented in two basic forms, scoring the correct answers and scoring the correct and incorrect answers. The purpose of this article is to describe the development of a single inexpensive system for test scoring which will be flexible and manageable, easy to use and cost effective.

The OMR (Optical Markup Reader) acronym is used to describe a method of reducing data entry in forms processing environment. OMR is a data collection technology that does not require a recognition engine; it is a technology that detects the absence or presence of a mark, but not the shape of the mark. The general purpose of this kind of programs is to process different types of tests and other informative questionnaires. The advantage over manual test processing is saving of time and human resources, but provides an additional security and reliability because of the absence of human interventions and errors.

The OMR system that we developed is functioning by detecting the presence or absence of a mark by using the lightning and shading areas of the scaned image. The image of the test is scaned by using a standard flatbed scaner with color information discarded option. The color information discardation usually speeds up the process of retrieving test image from the scaning software. The test precision and the size of the answer area for each single answer, imply the resolution to be used for scanning. Mainly, the resolution of 150 dpi meets the requirements for most test forms. The resolution adjustment is very important because high resolution images take more time to scan and more time to process.

The main idea of the scanning is shading detection over some predefined places. If the shading of some place is not high enough the software ignores it and reports an empty answer. The shading parameters that specify the answer must be set carefully because of the different types of writing tools, like pens or pencils. For an example the pencils write-marks are usually scanned with weaker shading (or darkness) which lowers the shading criteria for a found answer. On the other hand using an eraser on some place that was previously marked and then canceled (erased)

still leaves some marks which may be confused for an filled place, and by that missjudging the correct or wrong answer.

## 2. PREPARATION FOR DATA ENTRY FORMS (TEST) PROCESSING

The first step in test analysis and processing is preparation of the system and testing equipment. In order to achieve machine testing system, a standardized testing form must be created like the form shown at the pic. 1 a). After creating the form it should be modeled in the system so that program will know the exact coordinates of the paper, where the data entry fields exist. As asumed the paper size on which the forms are printed is also important for the analysis program.

Further, after entering the page size that will be used and the coordinates of the entry fields that are analysed, the shading parameters must be set in order to prepare the step values for answer activation. The person (or people) who will carry out the testing or information retrieval must set regulations of the writing equipment that will comply the testing process.

The form that will be used must contain at least two markers, so that program can find the eventual angle of rotation and the translation of the test on the

scanning hardware. These markers must be printed on the test form at the upper left corner and the lower left corner, so that both paper dimensions errors (translations and rotation) can be detected. After the position and rotation errors of the test placement in the scanning hardware are detetcted, the program determines if the errors are alovable and do not affect the analysis process. If the errors are in the aloved range, the program adopts it self and corrects the expected data entry fields position. The adoption is carried out for each of the tests that are to be processed. The markers and their position from the upper left corner are entered in to the system, and the shading parameters for marker detecting are also set, because the step values for the markers will usualy differ from the ones of the wriiting equipment (the test forms are usually printed on laser printer with high quality). After the test form is prepared, the system can be tested afterwards for resolution of the images to be scanned (dpi).

The next important step for the testing analysis preparation is the creation of the key test which contains the correct answers and the points for each correct or incorrect answered question. There are several ways of importing the test in to the program, that is by manually entering each field and points for the correct answer, by automatic test retrieval from a test form which is filled with the answers at the correct places, or with the combination of the both methods. The last method is most the common method which provides maximum flexibility. Using

a) sample test form

b) sample key form
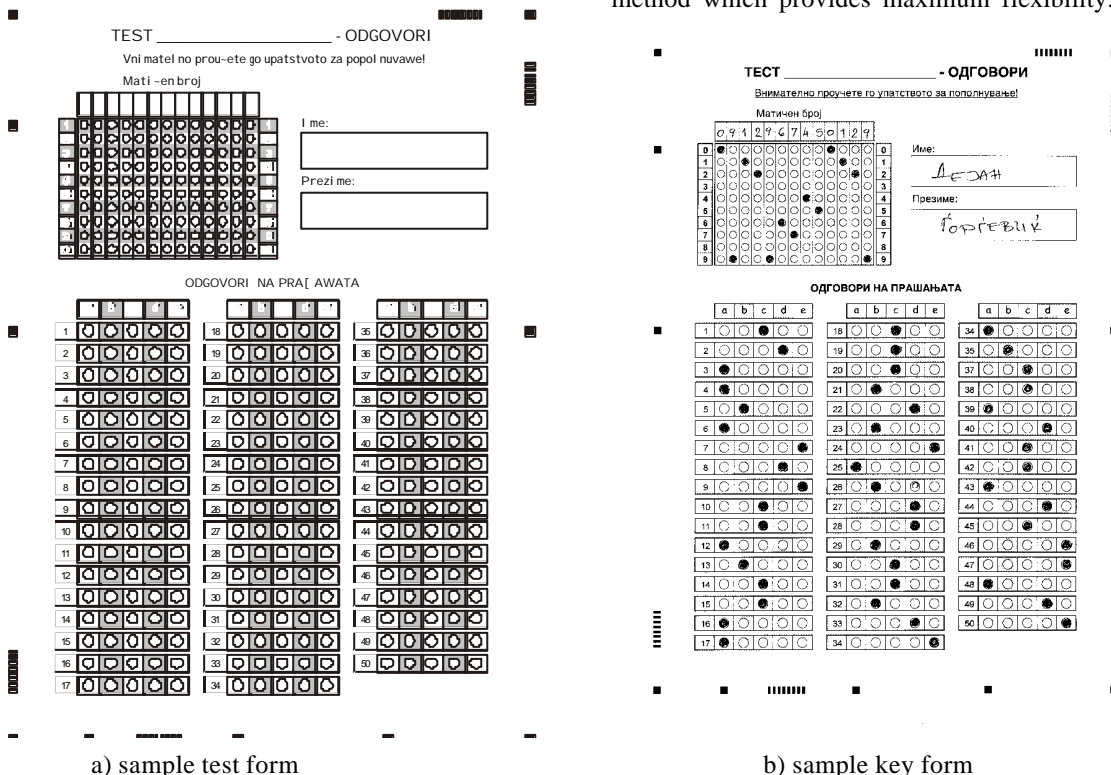
Pic. 1. Sample test forms

the last method, the professor (teacher) filles a single test form with the correct answers, sets the default test point for a correct answer in the program and uses the key option presented at the pic. 2 (at the bottom of the application form). After the test is processed the key database of that test is recorded in the application program and the user can manually alter each of the values for the correct field and points. Also by manual altering, incorrect hits can be assigned negative points which will be substracted after test form validation and analysis. Each imported or entered test is recorded in its own test database, which means that the system can support different kinds of tests.

3. THE TECHNOLOGY USED

The application software was developed for the Microsoft Windows operating system and uses the standard TWAIN protocol which all of the new scanning equipment support. Usage of this protocol enables the program to be hardware independed and to be used in different environments. Unfortunatelly, the software that is started from the TWAIN calls, which is the scanning hardware manufacturer software does not comply with any standards and

regulations. The application software has a special module that is used to describe the manufacturer software and to specify the commands that should be manually executed in order to scan the form. After specifying the commands that user manually executes the application uses Windows API functions to execute the command and to automatically control the scanning software. However, the automatic execution of the command is sometimes immpossible because some of the manufacturers does not use standard Windows forms, menus and buttons, but pictures insted, which are very difficult or eventually imposible to control automatically by using Windows API functions.

Several important factors define the working speed of the system, like the scanning device speed parameters, the resolution of the scanned images (usually below 150 dpi), the processing power of the system used for analysis and the automation level that can be achieved with the manufacturers scanning software (the HP Deskscan or Agfa Powerlook software can be fully automatized, while MagicScan of the UMAX is almost impossible to control).
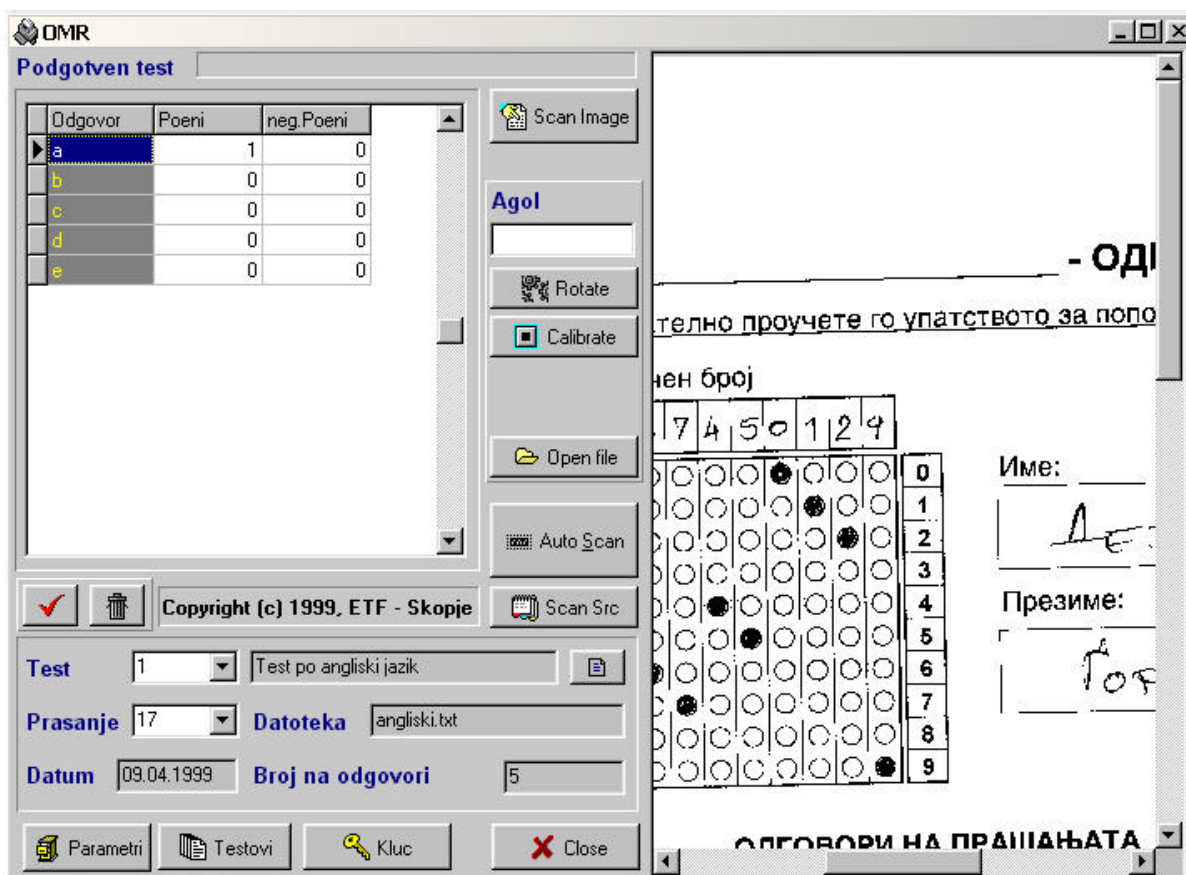

Pic. 2. Application software

4. TEST FORM ANALYSIS

The preparation process is carried out only once for each test type and testing environment conditions. The parameters and data entry fields values are recorded in the application program, and can be changed for different tests by importing and setting new test key. After all of the preparations are carried out the system is ready for data form processing. The test are processed one by one, after each test is analysed its answer values are scored and the recorded in suitable format (comma delimited) which can be easily imported in various database software, for further data analysis and statistics. If the program detects that the test rotation angle and translation in both axis does not fit into predefined range, or if the program can not locate some of the markers it reports to the user to correct the test form placement on the scanning device and rescan.

The actual operation of the software begins with marker detection on the test form. After the markers are detected they are compared with the predefined values. The difference that is detected is than used to adjust and adopt the software for field detection. After the rotation and translation parameters are detected the system uses field database to retrieve the positions to scan. Using the resolution value and the correction parameters the program determines the actual field location on the image that is scanned. After the position is determined the program uses shading parameters to determine if the place is filled with an answer. After each of the fields are scanned they are compared with the key database and they are

scored according to the correct answers. After the scoring process completes the user is reported to verify the test analysis and the test hits screen is presented to the user.

There are two different test form analysis methods, on fly and batch processing. The most common way, which is actually described in the paragraph above, is the on fly method, which means the tests are scanned and then immediately processed and analyzed. However, the test form images can be scanned all at once at a different computer, with some powerful scanning equipment, than saved in some common format for image exchange (like JPEG, TIFF e.t.c.) and finally analyzed in the program from the image files. The image file importing software supports most of the standard image file formats and also can be run over the network. The last process of test form processing is called batch processing.

The speed of the test processing using the OMR program is generally lower than the scanning performance of the scanning machines, while reducing the costs of equipment purchase, equipment maintenance and test printing.

5. CONCLUSIONS

Machine test form analysis can be implemented in a form of a flexible, cost effective system that can be used in various environments and with various scanning equipment.

The process of test analysis and scoring with such a system can be fully automatic, manageable and suit-
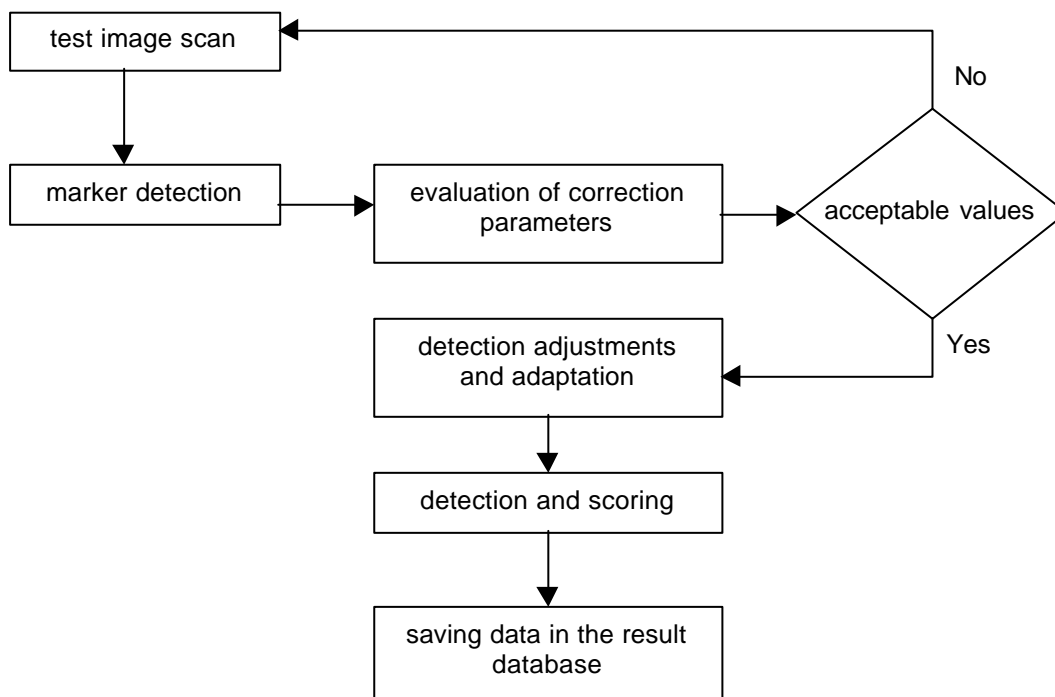

Fig. 1. Workflow diagram of the recognition process

able for data export in various database software for later statistical analysis.

Although, the processing speed of the system is lower than the processing speed of the OMR machines, the systems performance can meet the requirements of the most of the common testing environments.

6. REFERENCES

[1]. Super Scanning Techniques The Hewlett-Packard Guide to Black-And-White Imaging by Jerry B. Day. Paperback (November 1993)

[2]. Optical Pattern Recognition by Francis T.S. Yu(Editor), Suganda Jutamulia (Editor). Hardcover (April 1998)

[3]. Selected Papers on Optical Pattern Recognition (Spie Milestone Series, Vol Ms156) by Francis T. S. Yu(Editor), Shizhuo Yin (Editor). Hardcover (September 1999)

[4]. Computer Analysis of Images and Patterns : 8th International Conference, Caip'99, Ljubliana, Slovenia, September 1-3, 1999 : Proceedings (Lecture note by Franc Solina(Editor), Ales Leonardis (Editor). Paperback (November 1999)