Buletinul Ştiinţific al Universităţii "Politehnica" din Timişoara

Seria ELECTRONICĂ şi TELECOMUNICAŢII
TRANSACTIONS on ELECTRONICS and COMMUNICATIONS

Tom 49(63), Fascicola 1, 2004

Optical Character Recognition Application using the IMAQ PCI 1411 Acquisition Board

A. Chioreanu[1], R. Carean[1], B. Orza[1], A. Vlaicu[1]

Abstract – The objective of this paper is to propose a solution for extracting the information regarding the identity of a person from a Romanian Id card, to process that information, and to generate a bar code containing the CNP (Numerical Personal Code).
This paper is intended to be a research on the OCR (Optical Character Recognition) and image acquisition, but it is also the first step in the development of a useful product designed for front desk use. The application presented in this paper might be an important part for a more developed platform for University purpose, which would help the accounting department to extract the information from students IDs. The code bar generation would ease very much searching for students in the University databases.
Keywords: OCR, enhancement, code bar

I. INTRODUCTION

The purpose of the text recognition problem is, as its name indicates, to achieve efficiently the recognition of characters from a text source, examples of text sources are books, magazines, letters, ID cards, license plates and so on. The main advantages of the OCR (Optical Character Recognition) are constitute from the fact that a recognize text file is much smaller (about 5%) than an image file with the same text. In an OCR text there is always place for editing, reformatting or information retrieval. Beside all these using OCR the process of typing the information is much faster than regenerating of the document by manually rekeying the text into the computer.

II. ARCHITECTURE OF THE APPLICATION

There is an easy to understand model of an image processing system, which has little changed over the years, being composed of four main blocks: the acquisition block, the processing unit, the display unit and the storage capabilities, as one can see in Fig. 1.

The model fits perfectly for analog and digital processing systems. Talking about digital processing systems, at its most basic level, such a system requires a computer and two special input/output devices: an image digitizer and an image display device. Since computers do not know how to treat images, they work with numerical data rather then pictorial data, an image must be converted to numerical form before processing by computer can begin.

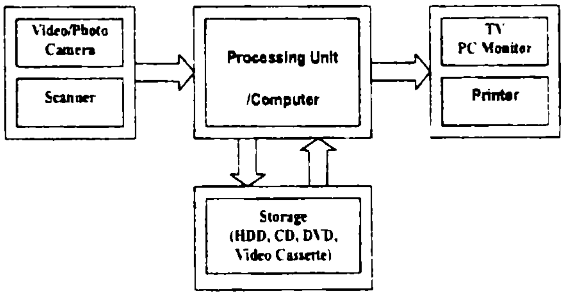


Fig. 1 The architecture of the application

This conversion takes place in the acquisition block and involves two actions: transforming the images into an analog electric signal and then digitizing the electric signal. The first is made by a sensor, sensible for a given domain of the electromagnetic spectrum (the X ray spectrum, the infrared spectrum, the visible spectrum in our case etc.), that gives an analog output signal proportional to the energy of the radiations. The second device, the digitizer, transforms the signal into a signal with discrete values. The numerical data obtain from the acquisition block is then passed in the processing unit. The processing unit can transfer the image data in the memory or directly to the output device or it can make transforms on the numerical data and then store or display the results.

There are many fields that use image acquisition systems, starting from medical investigations (such as topography), to military purposes (satellite image taken from war zone and transmitted to the headquarters), from farming industry (satellite images of crops) to security (iris identifications). If at the beginning of all this there was little difference between the acquisition systems, now there are a big variety of them, this is due to the specific field demands. Some fields need high quality images but

[1] Faculty of Electronics and Telecommunications, Communication department G. Bariţiu Nr. 26-28, Cluj Napoca, e-mail orza@com.utcluj.ro

they don't need a small transfer time (this is the case for Spatial Exploration), but some like in the case of web conferences need high speed (but they can trade this in the detriment of image quality).

III. ACQUISITION AND PREPROCESSING OF THE IMAGE

The application is made up of several blocks: the image acquisition, the image processing, the OCR, the character set treatment, the presentation (interface), and the barcode generator. The system that make this is compose by a CCD PAL video camera connected to a IMAQ PCI 1411 (provided by National Instruments) acquisition board mounted on an ordinary PC running dedicated recognition software created by us on a Windows platform.

National Instruments, with the help of the NI-IMAQ library, offers support for the IMAQ boards programming on two different environments: VC++ and LabView. We chose to write the code in VC++ because the final product is an executable program that can be used on any machine and because the execution time using LabView is slow. VC++ with MFC (Microsoft Foundation Classes) offer relatively easy use of the interface that can match LabView's interface.

This project is part of a recognition character recognition system. There are a lot of recognition systems: shape/color recognition, character recognition, and biometrics recognition. The main difference between those systems consist in the algorithm they use for recognition and in the way instrument used to acquire the image, this differences arrases from the demands of each specific application which may need to be in real time or not, fast or small, reliable or cheap and easy to use. Basically we can use two ways of taking the picture: by a video camera or by scanning it. The choice of using a camera instead of a scanner is made when we need a greater control over the moment of acquisition and a greater speed. We chose to acquire the image using a scanner when we need a high quality image. The system that we have designed needed an easy to use procedure. For this reason we have chosen to acquire the image using a color video camera then to pass the information on the computer via an acquisition board, so that we can control the acquisition parameter sitting in front of the computer.

The output of the video camera is a video complex signal. This signal is passed to the acquisition board which transforms it into a true color digital image (640*480) that is stored in a buffer in the PC memory. The OCR procedure needs a binary image input so that it can work properly, but till now we only have a true color picture. So the next step of the application is to process the ID card color image. First we will transform the picture from a true color representation in a gray scale representation. This is done with the help of the luminance formula.

$$Y = 0{,}3R + 0{,}59G + 0{,}11B \qquad (1)$$

After we obtain the luminance component of the image, which is a gray scale representation o the image we make the binarization. The binarization function is basically a comparison, so if the value of the luminance of a certain point is under the threshold we will have the output set to zero and if the luminance is greater or equal to the threshold we will have the output set to one. We used for this application a threshold value of 110, this value was determined experimentally, but the threshold can be virtually any value between 1 to 255.

IV. THE OCR (OPTICAL CHARACTER RECOGNITION) ALGORITHM

Nobody can talk about a standardized method for implementing OCR. Different corporations or different programmers choose their own technique in order to achieve the best results for the application the OCR system is intended for. OCR involves four discrete processes subsequent to image capture:

1. **Identification of text and image blocks in the image:** Most software uses white space to try to recognize the text in appropriate order. But complex formatting such as cross-column headings or tables must be manually delineated by "zoning" (identifying and numbering text blocks) prior to OCR. Images interspersed throughout the text will usually be ignored by the OCR software; they will be dropped from simple output formats such as ASCII.
2. **Character recognition:** The most common method of character recognition, called "feature extraction", identifies a character by analyzing its shape and comparing its features against a set of rules that distinguishes each character/font.
3. **Word identification/recognition:** Character strings are then compared with dictionaries appropriate to the language of the original.
4. **Correction:** The OCR output is stored in a proprietary file format specific to the OCR software used (and the image is then usually discarded). The software highlights non-recognized characters or suspicious strings, and an operator inputs corrections.

The flow diagram used in our OCR algorithm is shown in Fig. 2.

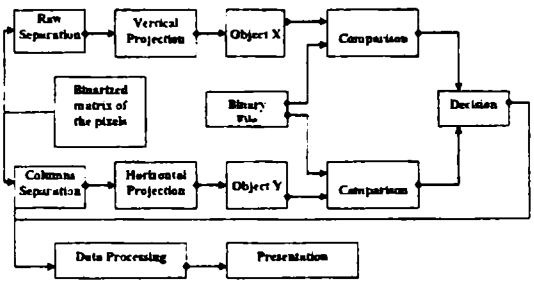


Fig. 2 The flow diagram of the OCR algorithm

The algorithm detection is based on the comparison of the character vertical and horizontal projections, with the horizontal and vertical projections stored in a "data base". Starting from a binary image the horizontal and vertical projection for each character look like in Fig. 3.

The application first computes the horizontal projection of the ROI (Region of Interest). By looking at those projections we can separate the rows (between the rows the horizontal projection has a very low value – usually 0). After the row separation we proceed to the character separation by computing the vertical projection for each row in the ROI. After this by using the vertical projection we can separate the characters (between the characters the value of the projection will go to 0, note - this can be used only if the characters are not bounded together).

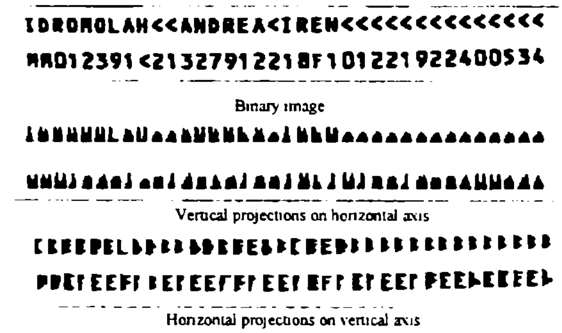
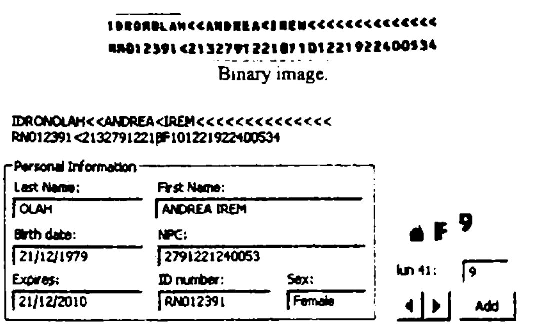


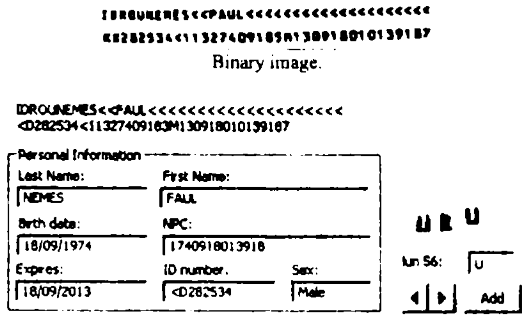
Fig. 3 OCR parameters

After we separated the character we calculate also the horizontal projection for each character, so for every character we will have two sets of projections: an array containing the vertical projections and an array containing the horizontal projection. From experimental results we have determined that this pair of characteristics (projections) is different from one character to another, so we can say that each character has a different horizontal and vertical projection signature. We will compare this pare of projections with ones stored in a database and so we can determine the degree of resemblance between an acquired character and one stored in the data base, and thus perform the OCR.



Print screen taken from the application representing the recognized information from the ID card.

Fig. 4 Experimental result



Print screen taken from the application representing the recognized information from the ID card.

Fig. 5 Experimental result

V. EXPERIMENTAL RESULTS

Using this algorithm we have obtained a 70% recognition rate, and by using a very simple correction algorithm we can increase that ratio for the CNP up to 90%. This correction algorithm is based on the format of the text contained on the Romanian ID card. One can notice that in the first row there are only letter and on the second row all the characters but the first two are digits. The user has the possibility of editing the acquired text and the possibility to replace or to add new charters in the database, so the application is able to learn new characters. The correction algorithm can be improved by using a dictionary lookup for the names. Bellow you can see some experimental results (the smaller image represents the binary input image and the bigger print screens represents the recognized information).

Results obtained using this method can be improved by designing a special acquisition device for the ID card and by improving the processing algorithm (contour enhancement).

VI.CONCLUSIONS

As a long time investment an OCR (Optical Character Recognition) system as this one, gives us a lot of benefits at a low maintenance cost. For instance the job of a secretary can be eased and fastened, thus she can allocate the time given for copying information from a document (in our case an id card), to other office activities.

For a variety of reasons, current OCR devices cannot even read as well as a seven-year old child. In order to extend the scope of prospective applications to more difficult material, researchers in many countries are hard at work to develop better methods for reading printed text. We speculate on what might be in the wings, but try to keep an open mind about the nature of the solutions that might emerge. There are four important potential sources of improvement:

- Improved image processing;
- Adaptation of the classifier to the current document;
- Multi-character recognition;
- Increased use of context.

BUPT

The first method is based on more faithful modeling of the printing, copying and scanning processes. The second specialized character classification to the typeface of the current document, while the third exploits style consistency in typeset text. The fourth method depends only on linguistic properties, and varies from language to language.

While OCR reliability is continually improving, and prices for all ranges of OCR software are decreasing, it is important to evaluate materials for OCR case by case. Each material type, and each instance of a given material type, has its own set of idiosyncratic problems for OCR. A single factor, such as a tight binding, can potentially eliminate scanning/OCR as a digitization option.

On the other hand, the effect of many of the OCR-hampering factors cited above can be minimized by ensuring appropriate system set-up, such as loading appropriate language dictionaries, adjusting brightness settings, dpi, etc., and by choosing appropriate originals and employing the hardware necessary to ensure optimum throughput (e.g., document feeders). Recent in-house experimentation suggested that OCR can likely be performed cost-effectively on clear English, French or bilingual print documents, including some material presented in columns, employing a variety of standard fonts, or printed on poor-quality paper. Nevertheless, documents that present any one of the factors noted above should be carefully analyzed to ensure that OCR is the appropriate and cost-effective digitization solution.

The application described in this paper meets most of the requirements for an OCR system to work properly with maximum accuracy and efficiency. The capture device represents the reason why this application does not meet all these qualities. The camera delivers very noisy images to the PC. That is the fact why after the preprocessing the characters shapes are very unlinear and a considerable difference is present between the same characters. The use of a linear scanner or an image from a good digital camera would increase for sure the results of the applications.

The reason why this camera was used is for the fact that we wanted to design a more automated system, which would be very good for office work. The use of a scanner or a camera would slow down the process.

This algorithm can be improved to obtain better results even using this capture device. Some of the improvements that can be implemented can be:

- the use of a dictionary look-up method
- the combination with another algorithm for OCR
- the use of letter frequencies
- the use of a method for contour improvement for characters.

In order for the application to run optimally the ID card should be place neatly and all the time in the same position. A good idea would be to implement a method that would allow the user to place the ID within limits (skew or slanted) and not to bother arrange it carefully.

REFERENCES

[1] Aurel Vlaicu – „Prelucrarea digitala a imaginilor" (Second Editions) Microinformatica Cluj-Napoca 1997

[2] Kenneth R. Castleman – „Digital Image Processing" Prentice Hall, Englewood Cliffs, New Jersey.

[3] „The Measurement and Automatition Catalog 2004" – National Instruments 2004

[4] „321326A - NI-IMAQ Function Reference Manual" – PDF National Instruments 2002

[5] 321326B - NI-IMAQ™ User Manual Image Acquisition Software"" – PDF National Instruments 2002

[6] „321326G – Release Notes" – PDF National Instruments 2002

[7] „322157A - IMAQ PCI/PXI™-1411 User Manual Single-Channel Color Image Acquisition Board for PCI, PXI, and Compact PCI Chassis" – PDF National Instruments May 1998

[8] „322157B - IMAQ PCI/PXI™-1411 User Manual Single-Channel Color Image Acquisition Board for PCI, PXI, and Compact PCI Chassis" – PDF National Instruments October 1999

[9] http://searchcio.techtarget.com/sDefinition/0,,sid19_gci 213536,00.html

[10] Jeff Prosise – „Programming Windows with MFC" (Second Edition) Microsoft Press 1999