Real Time Recovery of Text based on FPGA

Baby Sathya S and Rajesh Kumar T

Abstract: For the few decades, the work on digitizing skewed document image is a major challenge for the researchers in the field of OCR. Imaging the old historic or bounded book causes skewness and distortion of the image. The proposed metric recovery method helps in preprocessing, skeletonization, skew detection, skew correction and recognition using horizontal and vertical text line extraction of the image. Text line detection depends on the goal of digitization. Digitization is helpful in museum and library where old historic and damaged manuscripts are imaged. In the real time the image is captured using the camera fitted in FPGA kit (cyclone II). The captured image is preprocessed with the verilog code which is converted from MATLAB code using simulink HDL coder.

Keywords: Optical Character Recognition, Digitization, Skeletonization, Field Programmable Gate Array, Simulink HDL coder.

I. INTRODUCTION

Image captured by flatbed scanner or camera causes more distortion in the image like border noise, skewness and overlapping of the text. To overcome this problem is difficult in the field of OCR. Flatbed scanner images causes skewness of the bounded historical books and document image acquisition using digital camera causes distortion due to curved page and perspective projection of the camera.

For the past days, many different techniques have been proposed to overcome the problem based digitizing the document image. Only the rectification of the document image is proposed. Rectification results only in correcting the slant document image. In our proposed method recognition of the character is included which helps in converting the rectified document image. The output is been displayed in the notepad which helps to identify the character very accurately.

In some previous work rectification is done from coarse to fine rectification method. But each and every time word and text line is to be detected before coarse or fine rectification.

Section II focuses on the general description and the previous rectification method along with some disadvantages. In section III Proposed technique is and conversion of MATLAB code into verilog code is discussed.

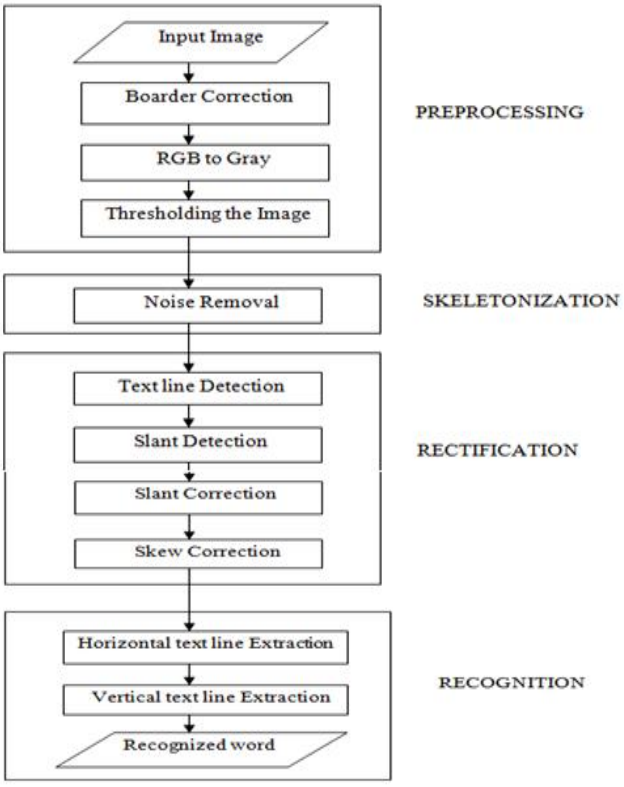

Fig 1. Flow Chart of the proposed method

II. GENERAL RECTIFICATION METHOD

Optical Character Recognition is the mechanical or conversion of image of printed, handwritten or typewritten document into machine - editable form. It has wide application in capturing a document through various representations like scanned image or facsimile multi-page or Ancient scripts on the stones and walls of temples or Palm leaves writings in various languages into high level semantic descriptions of the documents. Thus it involves many interdisciplinary topics including image processing, character recognition, natural image processing, artificial intelligence and database management systems.

The Optical Character Recognition system can be subdivided into five areas as follows. One is System Architecture, which specifies the hardness of the system task leads to modernization of the architecture. Optical Mark Reader or Intelligent Character Recognition (ICR) is used to describe the process of interpreting image data, in particular alphanumeric text which is used in license plate recognition.

Many applications today use direct data entry via keyboard, more and more of these will return to automated data entry. The reasons for this include the increased incidence of operator wrist problems from constant keying and the potential hazards of video display terminal emissions. There are many different things to be noted before capturing the image.

A. Future of Optically-Read Handwriting

Applications exist where it would be desirable to read handwritten entries. Reading handwriting is a very difficult task considering the diversities that exist in ordinary penmanship. However, progress is being made. Early devices, using non-reading inks to define specifically-sized character boxes, read constrained handwritten entries. This resulted in the development of a standard encouraging a certain style of handwriting. The best example of unconstrained handwriting reading was the IBM 3895. This device read the convenience amount entries from checks and then encoded the amount on the check in magnetic E13B characters. It is difficult to design a system to take care of mis-read characters. The 3895 also reads the entries from deposit listings to confirm or to prevent substitutions.

With the advent of image processing systems, this type of recognition is once again being developed. Restrictions on character size and the ability to provide target areas that are outlined in non-read inks will assist the accuracy of recognition. It would be helpful if our school systems could teach the proper manner to write numeric characters to enhance recognition. Once the application has been defined, a decision must be made regarding which paper should be used and what information should be printed prior to distribution.

B. Paper Considerations

The basis weight of the documented image is the first characteristic to be specified and it play a major role in the ultimate cost of the document. In general, most systems in place today prefer the use of a 24 Ib. sheet. (In other words, 500 sheets of paper, 18" x 23", would weigh 25 Ibs.) Paper caliper or thickness is also a consideration.

Other characteristics that may be specified include: tear strength, stiffness, bursting strength, porosity, fold resistance, etc. All of these play a major part in the processing ability of documented image. Other characteristic, such as hardness and smoothness is also specified to provide a better printing surface. In addition, some people specify the cleanliness and undistortion of the paper to avoid spots that will not interfere with scanning.

C. Paper Color

For some applications, it is desirable to use colored stock to help users readily identify images for use in different applications. Color coding is a very simple and satisfactory visual system. However, since OCR depends on the contrast between the background color as well as printed document, some color control is needed.

Color for the most part is restricted by reflectance. Standardized tests are accessible to calculate the virtual reflectance of a sheet in comparison to total white (defined as 100% reflectance). For most systems 65% reflectance is the smallest amount that can be used. The readability of any entity character is resolute by the print difference signal it generates. This is definite by the formula, "PCS = (reflectance of the background - reflectance of the printing) / by the reflectance of the background."

D .Printing Inks

Along with paper color, the color of inks used for pre-printing the document also plays an important role. Some data that is pre-printed includes serial number, roman letters may be OCR read, but other pre-printed data is to be ignored. This creates a new specification for non-reading inks.

E. Related work

The previous rectification work is based on document shape reconstruction and document image processing. The former rectification technique is further subdivided into two categories. First is the detection of distorted text line based on document image and second is the spatial transformation between the warped and dewarped. Techniques of the first subcategory obtain the 3-D shape of the document image using special equipment such as laser scanners [3], stereo cameras [4], [5], or structured light setups [6]. The dependence on special equipment prevents these techniques from being used in an unconstrained environment. On the other hand, techniques of the second subcategory reconstruct the 3-D model from information existing in document images. Cao *et al.* [7] propose a method to rectify warping distortions in document images by constructing a cylinder model. Apart from the cylinder shape assumption, they also have a limitation on the pose that requires the image plane to be parallel to the generatrix of the page cylinder. Liang *et al.* [8] model the page surface by curved developable surfaces to estimate the 3-D shape of the page using texture flow fields.

This method is based on the assumptions that the document is either flat or smoothly curved and the camera is a standard pinhole camera. These techniques require knowledge of lighting, which in most of the cases is unknown. The later technique assumes that the page is smoothly curved or flattened.

III. PROPOSED METHOD

The framework of the project is to retrieve the text from an image captured using flat bed scanner or captured using digital camera. It is possible to rectify the distortions caused by page curl, perspective, border noise due to bounded volume books and their overlapping of text. To detect the text line based on the extraction of the horizontal and vertical text line.

Preprocessing step is applied for the original distorted document image which includes border correction, RGB to gray conversion and thresholding the image**.** Border correction is done for removing some of the noises that are present in the border. The two-dimensional convolution operation is fundamental to the analysis of images. A new value is ascribed to a given pixel based on the evaluation of a weighted average of pixel values in a $k \times k$ neighborhood of the central pixel. Convolution kernel or the filter mask is represented with weights supplied in a square matrix. It is applied to each pixel in an image. Discrete form of the 2D convolution operator is defined by the following relationship between the elements $fi(x, y)$ of the input image, the elements $h(\alpha, \beta)$ of the convolution kernel, and the elements $g(x, y)$ of the output image by the following master formula

$$g(x,y) = \sum_{\alpha=-(k-1)/2}^{(k-1)/2} \sum_{\beta=-(k-1)/2}^{(k-1)/2} f_i(\alpha,\beta)h(x-\alpha, y-\beta),$$

where x, y, α and β are integers [10]. Gray scale conversion is for improving the accuracy of the image.

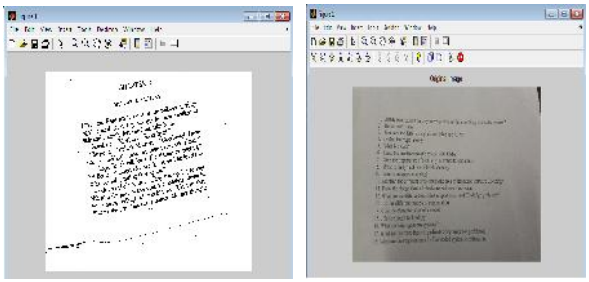

Fig 2 Original images

Thresholding is done by calculating the mean value for the pixels of the image and the values which are above threshold is set to white and for below to black. Thresholding is of two types: Global thresholding and locally adaptive thresholding. Locally adaptive thresholding produces a satisfactory output when compared to global thresholding. To find thresholding, 1) select an initial estimate for T (typically the average grey level in the image).2) Segment the image using T to produce two groups of pixels: G_1 consisting of pixels with grey levels >T and G_2 consisting pixels with grey levels ≤ T. 3) Compute the average grey levels of pixels in G_1 to give μ_1 and G_2 to give μ_2. 4) Compute a new threshold value:

$$T = \frac{\mu_1 + \mu_2}{2}$$

5) Repeat steps 2 – 4 until the difference in T in successive iterations is less than a predefined limit T_∞

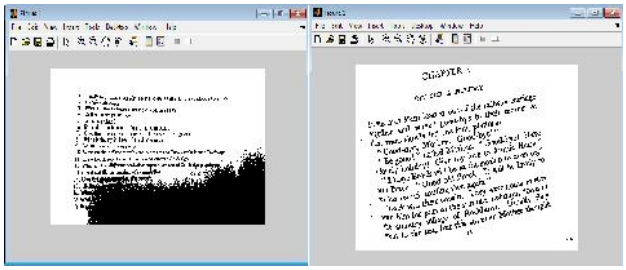

Fig 3 Preprocessed images

Skeletonization is the process of removing the noise that is present in the images. Gaussian Filters are used to remove the noise in skeletonization. The filtering function is given by H(u,v). G(u,v)=H(u,v)*F(u,v) where F(u,v) is the input image. Thinning is the other name for skeletonization.

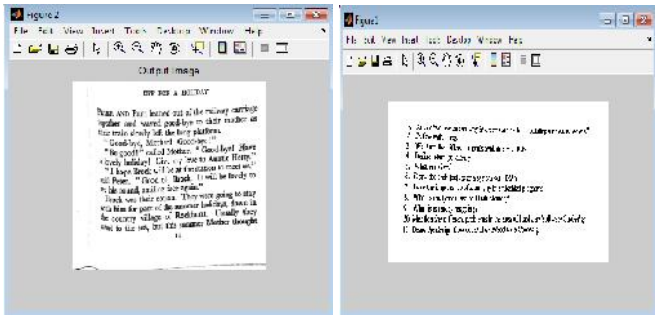

Fig 4 Image after rectification

Text line Detection, Slant detection, Slant Correction and skew correction are the steps involved in Rectification of the images. Recognition of the character is done by horizontal and vertical text line extraction. First the horizontal line extraction is done and then segmentation is made vertically. Thus words are recognized and are displayed in notepad.

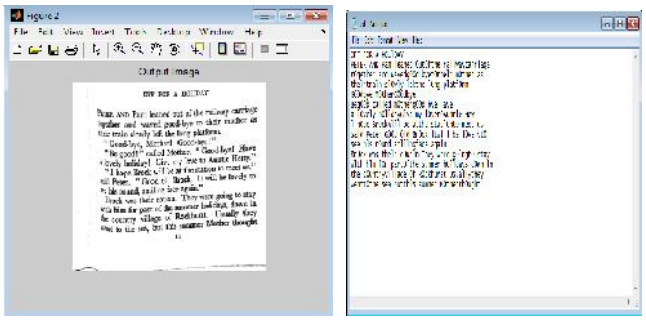

Fig 5 Displaying recognized word in Notepad

HDL Coder used to generate synthesizable HDL code for FPGA and ASIC implementations in a few steps: Model your design using a combination of MATLAB code, Simulink blocks, and Stateflow charts. Optimize models to meet area-speed design objectives. Generate HDL code using the integrated HDL Workflow Advisor for MATLAB and Simulink. Verify generated code using HDL Verifier.

HDL Code Generation from MATLAB

The HDL Workflow Advisor in HDL Coder automatically converts MATLAB code from floating-point to fixed-point and generates synthesizable VHDL and Verilog code. This capability is used to model the algorithm at a high level using abstract MATLAB constructs and System objects while providing options for generating HDL code that is optimized for hardware implementation. HDL Coder provides a library of ready-to-use logic elements, such as counters and timers, which are written in MATLAB.

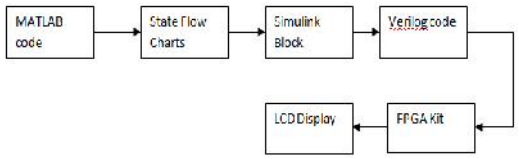

Fig 6 Conversion of MATLAB code to verilog code

HDL Code Generation from Simulink

The HDL Workflow Advisor generates VHDL and Verilog code from Simulink and Stateflow. With Simulink, the algorithm can be converted using a library of more than 200 blocks, including Stateflow charts. This library provides complex functions, such as the Viterbi decoder, FFT, CIC filters, and FIR filters, for modeling signal processing and communications systems and generating HDL code.

IV. CONCLUSION

Metric recovery method is for removing the nonlinear geometric distortions that are present in a document image which is captured using the flat bed scanner. It rectifies the distortions caused by page curl, perspective, and the overlapping of text. The experiments on both synthetic and real-captured images demonstrate the robustness and efficiency of the proposed method. The experimental results show that MRT outperforms these methods in terms of rectification errors and OCR accuracies. Cyclone II FPGA starter development board helps in real time conversion of document images into editable form.

REFERENCES

[1] Nikolaos Stamatopoulos, Basilis Gatos, Ioannis Pratikakis, Member, IEEE, and Stavros J. Perantonis "Goal-Oriented Rectification of Camera-Based Document Images" IEEE transactions on image processing, vol. 20, no. 4, April 2011

[2] J. Liang, D. Doermann, and H. Li, "Camera-based analysis of text and documents: A survey," Int. J.Document Analysis and Recognition, vol.7, no. 2-3, pp. 84–104, 2005.

[3] L. Zhang, Y. Zhang, and C. L. Tan, "An improved physically-based method for geometric restoration of distorted document images," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 4, pp. 728–734, Apr. 2008.

[4] A. Ulges, C. H. Lampert, and T. Breuel, "Document capture using stereo vision," in Proc. ACM Symp. Document Eng., Milwaukee, WI, 2004, pp. 198–200.

[5] A. Yamashita, A. Kawarago, T. Kaneko, and K. T. Miura, "Shape reconstruction and image restoration for non-flat surfaces of document with a stereo vision system," in Proc. 17th Int. Conf. Pattern Recognit.,Cambridge, U.K., 2004, pp.

[6] M. S. Brown and W. B. Seales, "Image restoration of arbitrarily warped documents," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 10, pp. 1295–1306, Oct. 2004.

[7] H. Cao, X. Ding, and C. Liu, "Rectifying the bound document image captured by the camera: A model based approach," in Proc. 7th Int. Conf. Document Anal. Recognition, 2003, pp. 71–75.

[8] J. Liang, D. DeMenthon, and D. Doermann, "Geometric rectification of camera-captured document images," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 4, pp. 591–605, Apr. 2008.

[9] C. L. Tan, L. Zhang, Z. Zhang, and T. Xia, "Restoring warped document images through 3-D shape modeling,"

IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 2, pp. 195–208, Feb. 2006.482–485.

[10] Seul. M, O'Gorman Lawrence, Sammon. Michael, Practical Algorithms for Image Analysis - Description, Examples, and Code, Cambridge University Press, 2000.

Baby Sathya S was born in Tirunenveli, Tamil Nadu. She is currently pursuing the Master degree in Embedded System Technology. She received the Bachelor degree in Computer Science and Engineering from Anna University, Chennai, in 2010. Her research interests include computer vision and machine learning with an emphasis on pattern recognition.

Rajesh kumar T was born in Tirunelveli, Tamil Nadu. . He received his Bachelor degree in Engineering during 1996 from Vinayaka Mission Kirupanandha Variyar Engineering College, Salem under Madras university and Masters degree in Computer Science and Engineering during 2004 from Manonmaniam Sundaranar University, Tirunelveli. He is currently pursuing the Doctoral degree programme under Anna University, Chennai in Faculty of Information and Communication Engineering from July 2011 onwards. His area of interest is Image/Speech Signal Processing and Embedded System. He has thirteen years of teaching experiences in various Engineering Colleges and two years of Industrial experiences. Currently he is working as Assistant Professor (Senior Grade) in Electronics and Communication Engineering Department. His work on authoring a book is in progress for Electronics and Microprocessors and Digital System Design. As far as the Professional memberships are concern, he is a life time member of ISTE and Indian Science Congress, Culcutta. He is an Associate Member in IEEE professional body. Motivating Students with Innovative Ideas in Technical field of Engineering is his axiom.