

## PORTABLE READING MACHINE FOR BLINDS

Mamta Sharma, H.N Bhargaw\*, H.K Sardana  
Central Scientific Instruments Organisation, Chandigarh  
\*Advanced Materials and Processes Research Institute, Bhopal  
mamta\_sharma@csio.res.in

**Abstract:** India today presents an alarming picture with respect to the problems of the education and rehabilitation of the blinds/visually impaired in India. Hitherto existing touch-based methods of Moon and Braille for text cognition by the blind and visually impaired are no longer acceptable technologically. In this work, a portable reading machine for the blind and visually impaired has been developed enabling them to read printed material that has not been transcribed into Braille. In the development, scanning in steps, optical character recognition, and text-to-speech technologies were employed. Portable reading machine is made up of aspheric lens and embedded PC that is suitable for acquiring digital images and speech synthesis in an embedded manner.

### 1. INTRODUCTION

Portable reading machine is an electromechanical device for reading horizontal languages by vertical step scanning a document in parts[1], permitting multiple image-snaps by device with a necessary but nominal step-overlap and its automatic removal to provide continuity of text. It relates to using an array of imaging modules and appropriate lenses placed at a suitable distance, and all fitted in an ergonomically designed chamber as apparatus for convenient handling to suit scanning of document, all the processing would be done with the help of embedded PC to make the device portable like PDA, pocket PC. Since the device is using a pair of imaging modules having negligible but fixed field of view overlap, so first of all the device is calibrated using configuration cum calibration software. Calibration parameters required for the merging of left and right images of the document to get a mosaic image, that are stored in a calibration file for future use.

A certain minimum amount of overlap between two consecutive scans is necessary to achieve the continuity. Continuity between two scans is obtained by using template based search technique which removes the overlap area/text between the two scans and thereby providing the seamless text output. The device operation and the associated reading process is controlled by the keys on the handle and the system readiness prompt, acknowledgment and errors are provided by audible alarms through integrated speaker in embedded PC to facilitate the working by a visually Impaired person. The Device is a standalone unit with embedded PC and can be controlled by audio browsing alerts, e.g.: an audio alert will ask to shutdown the system automatically on a particular key press etc. Application software integrates optical character recognition (OCR) and text-to- speech (TTS) software to give the seamless text output.

Portable Step Scanning device uses two major techniques:

- Automatic Configuration-cum-Calibration of Portable Step scanning device.
- Document layout Analysis Algorithm to obtain the Structural layout of the text for Multi column text Reading.



Fig1: Portable Reading Machine

### 2. AUTOMATIC CALIBRATION OF DEVICE

The device scans a document text with plurality of identical imaging modules and vertical hand-positioning. The plurality of imaging devices allows a wide document width imaged in good resolution in one step at a close working distance, otherwise not possible by a single module. The images from each module need to be merged to form a single mosaic corresponding to the width of the document. Calibration of the device is the preparatory step that is required once since the device involves a pair of imaging modules. Two imaging modules are used to take left and right image of the document. There is a certain minimum amount of overlap between the field of view of the two modules. These two images are then joined together using the calibration

parameters to generate a complete image of a part of document. Once the device is calibrated, using the special software, the hand-held step-scanner can be used to read the document images.

Configuring the device implies identifying the number of imaging modules employed, their identification in the indexed order such as left & right, image sensing parameters such as resolution, exposure, gamma, gain etc. The procedure for calibrating the device with a pair of imaging modules is described here. It is to be conducted only if assembly has been completed or if some part is replaced. Any plane text document can be kept below the device in a manner that the central portion of the pattern forms the overlap area in between the two fields of views corresponding to both modules. The Configuration-cum-Calibration module merges the left image and right image after removing the overlapped area with the help of image processing library known as Open Computer Vision(OpenCV).

This configuration cum calibration algorithm uses correlation based template matching in the two images captured from the cameras, and the assumption that they will have some overlapping area. Algorithm take right most section of the left image as template and search for it in the right image, wherever the maximum correlation will be find out will be use as the points to generate appropriate translations to combine the images. This illustrated as shown in the **Fig 2(a) & (b)**.

The calibration file holds the customization of the device's keypad, identification of the camera, calibration data and color balancing data. The calibration file acts as the drivers for the device.

The invented device requires no calibration if only one module is used, e.g. device for reading books or novels. In such cases, the standard manufacturing and assembly errors are easily tolerated by the technique as are the hand movements during the step-scanning procedure.

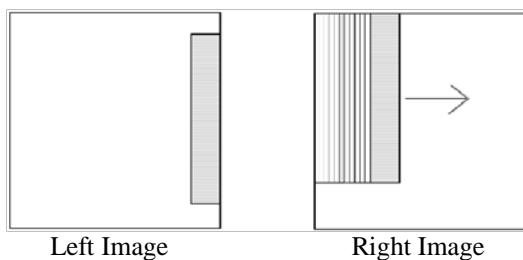
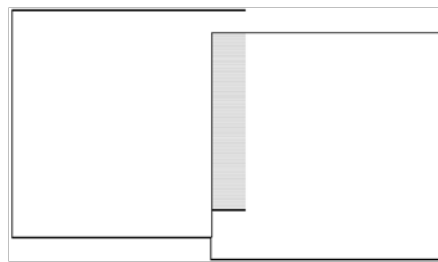


Fig2 (a): Template Matching



Merged Images

Fig2 (b): Result of Template Matching

The basic steps in Calibration program are as follows:

- Selecting the number of cameras present in device. (Single or double camera Device)
- Identification of cameras as left and right camera.
- Image Capture with default parameters.
- Auto-Adjustment of calibration parameters (Gain, saturation, X-trans & Y-trans, FOVs)
- Image Capture with the new properties and parameters.
- Saving the new settings.

### 3. STEP SCANNING TECHNIQUE

Once the device has been calibrated using configuration-cum-calibration utility software it is ready to use for reading documents. Device provides image data from the two imaging modules which independently and simultaneously acquire images from left and right portions of the document. Due to the manufacturing tolerances and assembly misalignments, the spacing between the two imaging modules is so chosen to provide at least a nominal overlap ~1mm at the field of view. This margin allows the tolerances and misalignments to be adjusted with the help of configuration cum calibration utility software. A horizontal merging module for mosaic generation is an essential part of the device driver that works much like the configuration cum calibration utility except that it merges the two images automatically using the pre-stored calibration parameters. Also since the scan length of Device (approximately 2.5" to 3") is smaller than the document length (viz. 11.85" for A4 document) therefore multiple steps are required to scan the full length of the document. A mosaic image is generated at every step-scan in vertical direction. In case of document layout analysis for the whole document information about the complete page is available in single step of analysis. But since step scanning is being used, layout analysis considers every scan as being independent from other. Layout analysis for each scan is done as it is a single page. This will provide complete text in the current scan. Therefore to provide the layout information for whole page all consecutive scans must be analyzed to provide continuity in the text and structure. Only condition

for providing complete information about complete page is to provide a certain minimum amount of overlap between two consecutive scans. A straight application of OCR to the current mosaic image gives the text corresponding to that mosaic. After that a Text Based Template Search technique removes the overlapping text between a pair of consecutive scans after appropriately searching a key-template text already obtained from previous scan.

Once the continuity for the consecutive scans has been established, the algorithm will look for the structural layout of the document, finally seamless text is given as input to TTS which gives enjoyable reading.

#### 4. TEXT-TEMPLATE SEARCH TECHNIQUE

Text Base Template Search Technique works in text domain that integrates a particular language based OCR to remove the overlapping text between a pair of consecutive scans. In this technique, the text in current mosaic image is extracted by the application of OCR analysis upon it. Some characters from the end of this text file are saved as a key template for use in the next consecutive scan. At the same time a previously saved key template from the previous text file has to be searched in this complete current text file.

For multi-column documents, the algorithm would find out the structure of text document (number of column), then it sorts the Text-blocks according to their location (column wise) on document and finally generate a Text-Template for each column. The number of characters that match in two files gives the percentage match between two files and the position of maximum match is used to give the position of overlap. The threshold will decide the continuity between two scans. If the match between two files is above a tolerance limit, taking into consideration the OCR errors, match status is stored as successful. If the match status is successful then overlapping text between two scans is removed from current scan file by using the position of maximum match. The pruned current scan file is then merged with the previous scan file to generate the seamless text output. The threshold test is helpful to decide the beginning of a new page, i.e. when the match between the two current scan files is negligible or down the threshold, it shows the discontinuity or start of new page. White space present at the end of current mosaic image helps to identify the end of page using white break terminators.

For Multicolumn document algorithm will sense the sensibility of reading flow. for example in case of two column document, the algorithm will read out the first column text for the consecutive scans and will hold up or buffered the second column text till the end of page, once the end is detected it will

automatically append the buffered text(second column text) to the TTS.

To find the end of page or end of article or any kind of break the algorithm will use white space information (providing in the end of page) or new article heading.

#### 5. DOCUMENT LAYOUT ANALYSIS

As described earlier at every step scan images from the left and right imaging modules are merged using the pre-stored calibration parameters to generate a mosaic. The mosaic image is given to standard OCR software for layout analysis and to extract the text from it [2]. OCR analyses the image in form of blocks depending upon the structure of the document as shown in Fig3.



Fig3: Document Analysis

From this layout analyses complete information about the number of blocks generated, type of each block (text, image etc), width and height of the document is obtained. It also gives the coordinates of left most and the rightmost corner, i.e. The diagonal coordinates of each block using which width of the block can be calculated. Now by comparing the position and width of the each block with the width of the current scanned document, document is classified a columned or multi-columned document. Depending on this classification, processing is done to get the seamless text output which is described here. On the first level, if the document contains both text and graphics, these are separated for subsequent processing by different methods [3]. On the second level, segmentation is performed on text by locating columns, paragraphs, words, and characters; and on graphics, segmentation usually includes separating symbol and line components [4]. For instance, in a page containing text and some illustrations similar to the pages of this journal, text and graphics are first

separated. Then the text is separated into its components down to individual characters. The graphics is separated into its components such as rectangles, circles, connecting lines, symbols etc.

Document image analysis refers to algorithms and techniques that are applied to images of documents to obtain a computer-readable description. It is concerned with the global issues involved in recognition of written language in images. A well-known document image analysis product is the Optical Character Recognition (OCR) software that recognizes characters in the image of the document. OCR makes it possible for the user to edit or search the document's contents. Document analysis adds to OCR a superstructure that establishes the organization of the document and applies outside knowledge in interpreting it.

The document layout techniques are also used to join different scans to get the whole document. Text from individual scans is extracted and after detecting continuity with next scans text is saved in reading flow format. The continuity between two scans is detected using text search based processing.

Document image layout has a hierarchical (tree like) representation, with each level encapsulating some unique information that is not present in other levels. The representation contains the complete page in the root node, and the text blocks, images and background form the next layer of the hierarchy. The text blocks can have further detailed representations like text lines, words and components, which form the remaining layers of the representation hierarchy. For many years people claim that, after thousands of years of predominance of paper, it will disappear as medium to save and transport information. Although many products have been developed to reduce the use of paper, for the moment it is not predictable if paper will ever disappear as information storage. As nevertheless more and more processes of everyday life are done electronically, there is a strong need to be able to convert printed documents into electronic ones.

## **6. CONCLUSION AND FUTURE WORK**

Device is used to read seamless text in appropriate reading order independently. As the scanner capture area is small as compared to length of document, so scanning whole page requires multiple steps. While scanning in multiple steps text which is in reading order is passed on to the speech engine. The remaining text is stored and passed on to speech engine only when it is in order. To maintain continuity there must be some minimum amount of overlap between two scans. This overlap is removed using text based search method. Text based approach is a language dependent method and depend upon the languages supported by OCR engine used.

The result shows that text based search technique works well for multi-columned document, up to a maximum of three columns. The technique is

workable under the conditions in which the document has text as well as pictures. In single columned documents all the text recognized in a particular scan is given to speech engine. In double columned document firstly the columns are separated out and then the text which is not in reading order is stored while the other is given to the speech engine. Column endings and page endings which defines the boundaries of documents are detected and are used to purge out the buffered data. Result shows that the technique work well in case of documents having multiple column endings as well as documents having pictures. Similarly three columned documents having column endings and pictures give good results. Reading order is maintained in almost all the cases. So the whole page scanned can be scanned in steps but the result is just like as if whole page has been scanned in single step.

In future work can be carried forward to handle newspaper reading in a seamless manner. In case of newspaper width along with length is more than the field of view of scanner so continuity has to be maintained in both the directions. Inset columns within the main columns can be handled more correctly. It is required to manage the internal memory of the device in such a manner that it can be used by the algorithm to save the scanned text in a readable form for future use (like reading or mailing or editing). Work can be carried out to generate features switched with audio alerts, so the device would be more user friendly and a blind user would be able to interact with the device with the help of audio alerts and input.

## **7. REFERENCES**

- [1] Patent number (WO/2009/115855) Hand held device and step scanning techniques for reading by the blind <http://www.wipo.int/pctdb/en/wo.jsp?WO=2009115855>, accessed on 23 December 2009.
- [2] "Document image analysis: A primer" Rangachar Kasturi1, Lawrence O'gorman2 And Venu Govindaraju, Department of Computer Science & Engineering, The Pennsylvania StateUniversity, University Park, PA 16802, USA, 2Avaya Labs, Room 1B04, 233 Mt. Airy Road, Basking Ridge, NJ 07920, USA, 3CEDAR, State University of New York at Buffalo, Amherst, NY 14228, USA.
- [3] F. M. Wahl, K. Y. Wong and R. G. Casey, "Block Segmentation and Text Extraction in Mixed Text/Image Documents," CGIP 20: 375-390 (1982).
- [4] L. O'Gorman, "The Document Spectrum for Page Layout Analysis," IEEE Trans. on PAMI 5(11): 1162-1173 (1993).