DEWARPING

Arya Mol V.
Computer Science and Engineering
Sree Buddha College Of Engineering
Elavumthitta, Pathanamthitta
aryaullass@gmail.com

Sruthy S.
Computer Science and Engineering
Sree Buddha College Of Engineering
Elavumthitta, Pathanamthitta
sruthy78@gmail.com

Abstract— Digital imaging is a new way for storing images electronically. It has numerous advantages like saves storage space, handles records easily, find documents rapidly and prevents lost records. Dewarping of camera document images has attracted a lot of interest over the last few years since warping not only reduces the document readability but also affects the accuracy of an OCR application. Here, a two-step approach for efficient dewarping of camera document images is proposed.

Keywords—Dewarping, Warping. Transformation model

I. INTRODUCTION

Digital imaging is a new way for storing images electronically. It has numerous benefits like saves storage space, handles records easily, find documents rapidly and prevents lost records. The availability of high performance, low quality equipments have made prolific changes in digital imaging analysis. Document image processing has been extensively studied about past 40 years. In conventional times, document imaging has been done with huge flat bed scanning devices. The images from flat bed scanners give good start up but when they are working on digital cameras they will be giving low clarity images, which cannot be readable by OCR systems. Recently, Portable devices like digital camcorders, digital cameras, PC cams, PDAs, and even cell phone cameras are most commonly used for image capture. They are small, light, portable, easily integrated with various networks, and moreover they are more flexible for many document capturing tasks in less constrained environments. These factors are heading to a natural extension of the document processing community where cameras are used for document image analysis.

Document imaging covers different areas including pre-processing, graphics analysis, writer identification, digital libraries, office automation, and forensics etc. Digital cameras, camcorders, PDAs and phone cameras can supplement the scanners and makes document image processing more flexible. These technical developments leads to advanced researches in the field of document image processing which aims in the video files and also abstraction of images in the text files. The scanner based OCR applications are now being converted into new platforms which are using more flexible image capturing devices.

Document image analysis can be categorized into a number of types: by the techniques used, by the devices equipped, by the intended application. Extraction of document images can be of various types. Document image contains text files with scenes, video frames with captions, etc. The feature difference of the image causes many of the challenges in the extraction of the document. Camera captured images suffer problems like perspective distortion which causes warping along the spline of the book, geometric distortion, low resolution, uneven lighting, complex backgrounds. These types of distorted images cannot be readable by the current OCR systems. Non-linear warping is a major distortion that makes document imaging analysis disgusting. The strongly distorted text in the document image makes the processing more complicated.

Many different approaches have addressed these problems and can be classified into two main categories based on three-dimensional (3D) document shape reconstruction and two dimensional (2D) document processing. Three-dimensional (3D) reconstruction requires specialised hardware like stereo cameras, laser scanners etc. So, it limits the flexibility of camera captured devices. (2D) document image processing uses single camera in an uncontrolled environment.

II. RELATED WORKS

This chapter represents a brief review of several studies and researches related to dewarping of document images. Several papers were studied based on this thesis and among them few papers relevant to this work are reviewed here.

Masalovitch and Mestetskiy [10] proposed a new method for approximation of whole image deformation as combination of single interlinear space deformations. Long continuous branches are used for defining interlinear spaces of the document. They are approximated by cubic Bezier curves so as to estimate the deformation of each interlinear space. After that a whole approximation of document image is built. Here, the initial image should be black and white with black text lines and white background and the initial image should contain one big text block. First the image is binarized and after that discrete binary image is represented as a set of continuous polygonal figures with lowest perimeter. Skeleton of polygonal figure can be represented as a planar graph, where nodes are points on a plane and bones are straight lines that connect the nodes. The main idea of the algorithm is that the image is represented as continuous skeleton system, and then filtering of the skeleton is built such that the unwanted bones are removed. After that extracting long near horizontal branches and then each branch is approximated by cubic Bezier curves and Bezier patches are built based on the obtained curves. One of the steps of this algorithm is the pre-processing step, on which all small garbage branches and branches that can be obviously determined as non-interlinear from the skeleton are deleted. This method has a demerit that it will not give a satisfactory result in the case of vertical borders of image which isn't so accurate.

Zhang and Tan [6] divide the document image into shaded and non shaded region. Initially, the shaded region is identified and image is binarized using Niblack's method. They find the text line curves by the connected component analysis and move the components to restore straight horizontal baselines. Images must be always greyscale and have a shaded region. This restoration method uses connected component analysis and regression methods to dewarp the image. A top down scanning method is used to rectify the distortions in clean area and after that alignments are corrected using linear regression method. A bottom down approach is applied to shaded portion and polynomial regression method is used to rectify the distortions. After that, warped text alignments and linear text alignments in both areas are then paired up. The warped text lines are restored by correcting the quadratic curves accordingly based on the corresponding straight text lines. This approach can be applied only for gray scale images.

In Lu and Tan's proposed method for the restoration of camera documents [5], the image is divided into three subsections. They are document partition, the target rectangle construction and document restoration. The document partition step includes two subdivisions. First one is that the distorted document image is divided into X line and base line. Second step converts the identified text line into smaller patches. After that, a target rectangle correspondence is built for each image such that the distorted image is mapped to this rectangle. The target rectangle is constructed based on the number and the aspect ratios of enclosed characters. The character aspect ratios are determined based on character span, character ascender, descender, and character intersection numbers. For each partitioned image patch, a target rectangle correspondence must be constructed within the target image to rectify that partitioned image patch. This method classifies characters to six categories with six different aspect ratios. Characters are classified based on the features including character span, character ascender, descender, and character intersection number. Finally in the image restoration step, rectification homography is applied to dewarp the image. This approach cannot be used when the distortion angle is big.

Zhang and Tian [20] proposed a method for warped document restoration in digital libraries. This method particularly focussed on boundaries to reduce the warping effect and uses Gordon surface model for the text lines of 2d image. Natural cubic splines are used for representing text lines. The image clarity will be less using this method. Here, a document-boundary independent approach to correct arbitrarily warped document images taken using ordinary digital cameras is explained. It is based on the Gordon surface model constructed from a set of text lines extracted from the 2D image. The text lines are represented using Natural Cubic Splines interpolating a set of points extracted from connected component analysis. Most of the images do not have explicit boundary curves for boundary interpolation. However, a ruled surface model based on the text lines is constructed here. This Gordon surface model can be projected to a planar surface without distortion. This 3D Gordon surface model can be applied to the 2D projection image since straight lines are preserved under projection. The iso-parametric lines passing through must also pass through the corresponding 2D projection points. Therefore, the projection of this Gordon surface model can also be parameterized using the projected text lines in the 2D image. This method cannot be used for more distorted image.

Koo, Kim and Cho [7] proposed an algorithm to compose a geometrically dewarped and visually enhanced image from two document images taken by a digital camera at different angles. From the corresponding points in these images, the surface of a book is reconstructed, and then stitches two rectified images for a visually better composite. Initially, a cost function is defined for the correction of geometric distortion, which is related with the geometric transformation of 3-D points. After that image stitching method is used to combine better patches from two images. Graph-Cut Optimization is used as the stitching method. Due to the misalignment of two rectified images and the asymmetry on the amount of information of each image, the simple average of two images is not a good solution to composing an enhanced image. So, better parts from each of the images are found, and then stitch and blend them into a single image. This method fails when distortion angle is big and not suitable for more curled images.

Tian and Narasimhan [8] introduced Rectification and 3D Reconstruction of Curved Document Images proposed a method that automatically reconstructs the 3D shape and rectifies a deformed text document from a single image. The regularity in the text pattern is used to constrain the 2D distortion grid. Here, the 2D distortion (warping) grid in an image is estimated by exploiting the line structure and stroke statistics in text documents. This estimation consists of two main steps: 1) text lines are automatically identified and densely traced 2) the text orientation is determined at every location in the image. In most documents 2D image grid can be regarded as a perspective projection of a 3D parallelogram mesh. Here, the process is done by tracing an initial set of text lines, called seed lines, across the document image from randomly selected initial points. These initial points are based on an image self-similarity measure. Then these seed lines are re sampled and refined using dynamic programming. In this work, it is assumed that the camera projection is perspective and each cell of the 2D warping coordinate grid is a parallelogram in 3D space. The second assumption is reasonable because the surface can be assumed to be locally planar or rigid if grid cells are sufficiently small. For most undistorted planar documents, the text lines are parallel and so are local vertical text directions, thus forming a parallelogram grid. But, this method reduces the flexibility of the user because additional hardwares are required.

III. PROPOSED SYSTEM

Image warping is a common problem in the case of scanning or capturing document images from thick volume books or from huge historical manuscripts. Warping causes shade along the spline of the book and also causes curliness on the text lines. This reduces the OCR accuracy and also impairs the readability of the user. Dewarping methods are used for flattening the curled document images and to rectify the distortions in the document image. In this chapter, Coarse to fine dewarping with enclosed box method is explained for dewarping document images. Image binarization (threshold selection) is the first step in this process. It converts a gray scale image into a binary image. Second step concerns with the detection of noisy black border and removal. Third step includes the corner detection of the curled document image. Coarse dewarping which is the fourth step deals with the transformation of the curled document image with a rectangular model. After that each of the letters in the text document is enclosed within rectangular boxes. This is the fifth step. Then fine dewarping is done for better dewarping results. Finally, image enhancement is done with morphological operators to enhance the document image. Fig 1 explains different steps in the dewarping process.

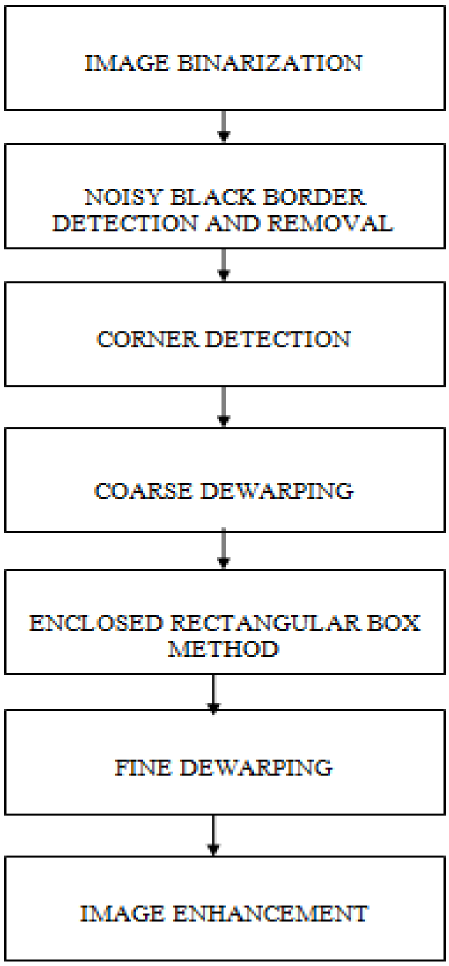


Fig 1: Flow chart of coarse to fine dewarping using enclosed box method

Document image binarization refers to the conversion of a gray-scale image into a binary image. It is the initial step in most of the document imaging analysis. Usually, it differentiates text areas from background areas, so it is used as a text locating technique [22]. Binarization plays a prominent role in document analysis since its performance affects the character segmentation and recognition results. When processing distorted document images, binarization is a hilarious task. Distortions appear quite often and may occur due to several reasons which range from the accession source type to environmental conditions. Examples of degradation influence may include the appearance of variable background intensity, shadows, smudge and low contrast. The binarization scheme consists of five basic steps. The first step includes a dedicated denoising procedure using a low-pass Wiener filter. Here uses an adaptive Wiener method based on local statistics. In the second step, rough estimation of foreground regions is done. Next, as a third step, the background surface of the image is calculated by interpolating neighbouring background intensities into the foreground areas that result from the previous step. In the fourth step, final binarization is done by combining information from the calculated background surface and the original image. Text areas are located if the distance of the original image from the calculated background overshoot a threshold. This threshold adapts to the gray-scale value of the background surface in order to preserve textual information even in very dark background areas. In the last step, we proceed to a post-processing that eliminates noise, improves the quality of text regions and preserves stroke connectivity [23].

Coarse Dewarping

In coarse dewarping step, a transformation model which maps the projection of the curved surface to the 2D rectangular area is applied. The extraction of the curved surface is achieved from the left, right boundaries and top and bottom curled lines. At first the borders are detected [25]. Let NL denotes the number of lines in the curled image and CH denotes the height of each character. Fig 2 shows an example of extracted curved surface.

Rectangular Transformation Model

In rectangular transformation, the projection of the curved surface is transformed with the 2D rectangular area. Let A’ (XI’, Y1’), B’ (X2’, Y2’), C’ (X3’, Y3’), D’ (X4’, Y4’) represents the points on the rectangular area. Let $\widehat{AB}$ be the arc length between points A and B and |AB| represent the Euclidian distance between points A and B. Here, we want to calculate the width W of the rectangle. It is calculated as :

$$W = \widehat{AB}, \widehat{DC}$$

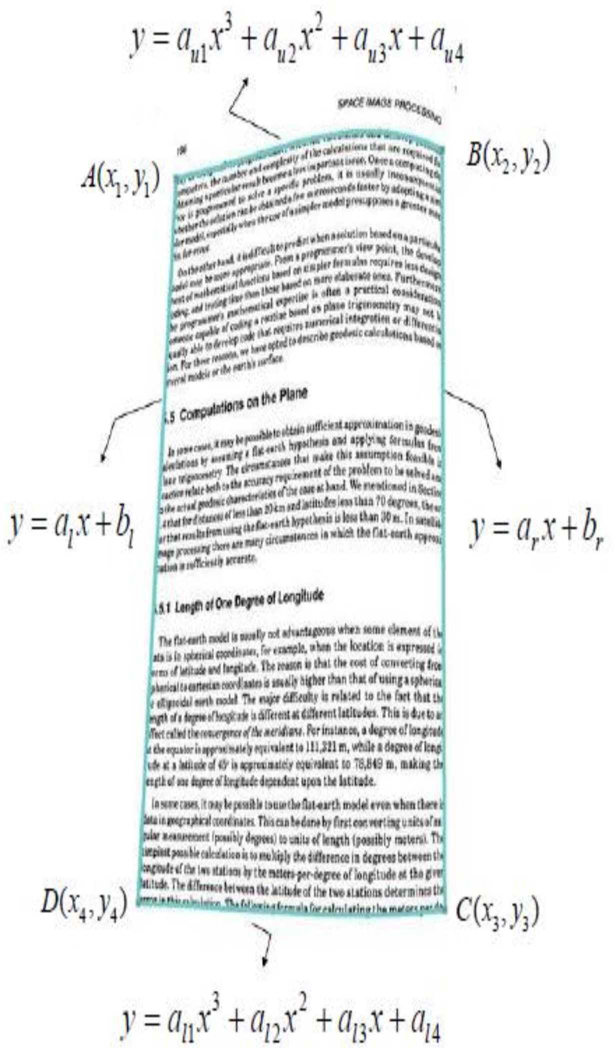


Fig 2: Extraction of curved surface projection

Height of the rectangular area is calculated as:

H = min (|AD|, |BC|)

After that we want to calculate the corner points of the rectangle as follows:

x1' = x1	y1' = y1
x2' = x1' + W	y2'=y1'
x3' = x2	y3' = y2' + H
x4' = x1'	y4' = y3'

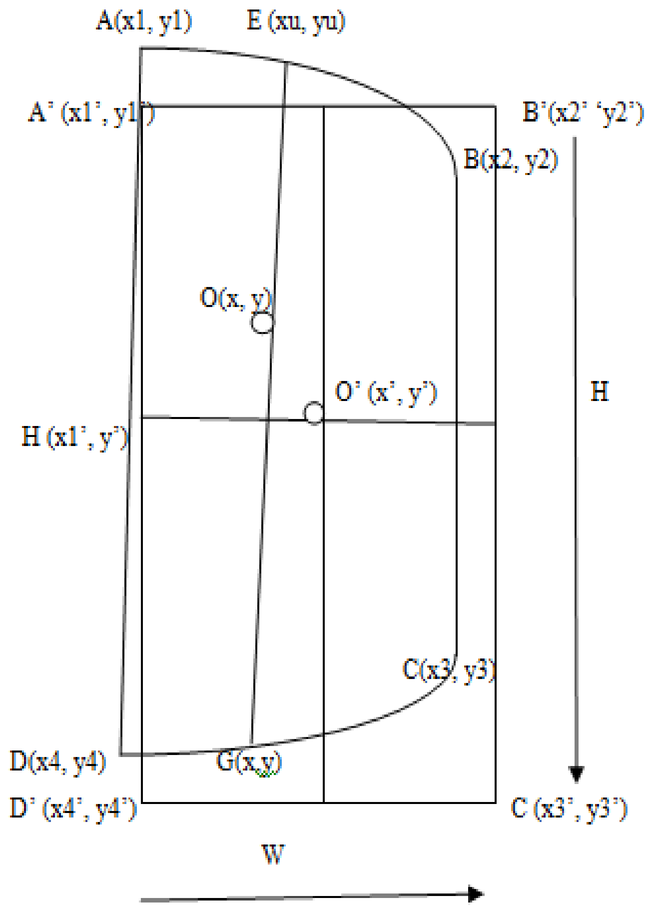


Fig 3:Transformation model

Next step is to represent the points in the curved surface to the points in the rectangular area. Fig 3 shows the transformation model which maps the curved surface area. Consider the point O (x, y). Let us transform this point to the rectangular area by calculating new position O'(x, y) for O(x, y) as follows:

X' = x1' + | A'Z |

Y' = y1' + | A'H |

Where | P'Z |, | P'H | are calculated as follows:

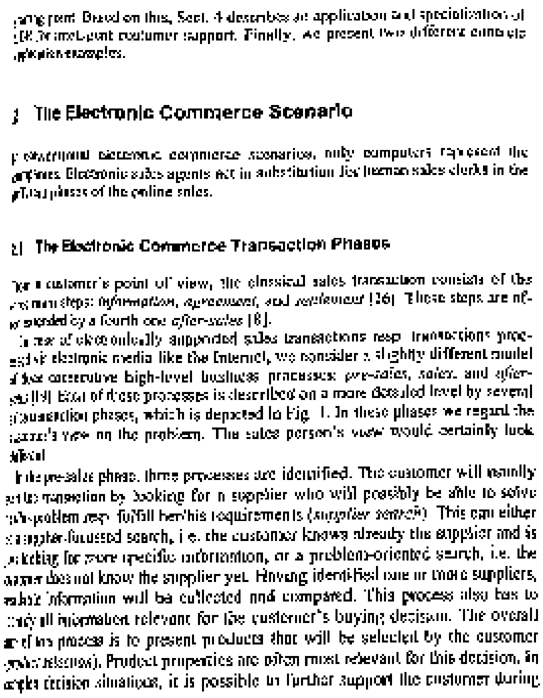
$$\frac{|AQ|}{|AE|} = \frac{W}{|A'Z|}$$

$$= \quad | A'Z | = \frac{W}{|AQ|} * |AE|$$

$$\frac{|EG|}{|EO|} = \frac{H}{|AH|}$$

$$= \quad |A'H| = \frac{H}{|EG|} * |EO|$$

Repeating this terms for all the points in the curved surface area. Finally each and every point on the curved surface area is transformed to new points in the rectangular area. Figure 4 shows the output image after coarse dewarping.


[illegible] Based on this, Sect. 4 describes an application and specialization of [illegible] intelligent customer support. Finally, we present two different concrete [illegible] examples.

[illegible] The Electronic Commerce Scenario

[illegible] electronic commerce scenarios, only computers represent the [illegible]. Electronic sales agents act in substitution for human sales clerks in the [illegible] phases of the online sales.

[illegible] The Electronic Commerce Transaction Phases

[illegible] customer's point of view, the classical sales transaction consists of the [illegible] steps: *information*, *agreement*, and *settlement* [26]. These steps are [illegible] by a fourth one *after-sales* [8].

[illegible] of electronically supported sales transactions resp. transactions [illegible] electronic media like the Internet, we consider a slightly different model [illegible] consecutive high-level business processes: *pre-sales*, *sales*, and *after-sales* [9]. Each of these processes is described on a more detailed level by several [illegible] phases, which is depicted in Fig. 1. In these phases we regard the [illegible] view on the problem. The sales person's view would certainly look [illegible]

In the pre-sales phase, three processes are identified. The customer will mainly [illegible] by looking for a supplier who will possibly be able to solve [illegible] problem resp. fulfill her/his requirements (*supplier search*). This can either [illegible] supplier-focussed search, i.e. the customer knows already the supplier and is looking for more specific information, or a problem-oriented search, i.e. the [illegible] does not know the supplier yet. Having identified one or more suppliers, [illegible] information will be collected and compared. This process also has to [illegible] all information relevant for the customer's buying decision. The overall [illegible] of this process is to present products that will be selected by the customer [illegible]. Product properties are often most relevant for this decision. In [illegible] decision situations, it is possible to further support the customer during

Fig 4:image after coarse dewarping

Fine Dewarping

In fine dewarping , word level dewarping is presented. Here, first detect all the text lines and words. For this method, remove all the non text components using the condition:

$$Height > 3 * CH \text{ or } Height < \frac{CH}{4} \text{ or}$$

$$Width < \frac{CH}{4}$$

After that, upper baseline of the word is defined as:

$$y = aij + bij$$

Then all the words are rotated and translated as follows:

$$y^{rs} = y^{r} + d_{ij}$$

$$x^{rs} = x = x^{r}$$

Where

$$y^{r} = (x - x_{min}^{ij}) * \sin(-\emptyset_{ij}) + y * \cos(\emptyset_{ij})$$

$$d^{ij} = \begin{cases} y_{i0}^{ru} - y_{ij}^{ru}, & \text{if } |\emptyset_{ij}^{u}| - |\emptyset_{ij-1}| < |\emptyset_{ij}^{l}| - |\emptyset_{ij-1}| \\ y_{i0}^{rl} - y_{ij}^{rl}, & \text{otherwise} \end{cases}$$

$$y_{ij}^{ru} = (a_{ij}x_{ij} + b_{ij}) * \cos\emptyset_{ij}$$

$$y_{ij}^{rl} = (a_{ij}x_{ij} + b_{ij}) * \cos\emptyset_{ij}$$

Where $\emptyset_{ij}$ is the slope of the word and x_{min}^{ij} is the left side of the enclosed rectangular box. At last all the components that we have been removed are added.

Image Enhancement

Morphological operators are used for image enhancement in curled document images. They affect the structure, layout or shape of an image. The two significant morphological operations are dilation and erosion. Object expansion can be done by dilation. It potentially fills in small holes and connects disjoint objects. Erosion shrinks objects by etching away their boundaries. By the proper selection of structuring element these operations can be customized for an application, which determines exactly how the objects will be dilated or eroded [26].

IV. EXPERIMENTAL RESULTS

This work is implemented in MATLAB. In order to verify the validity of this work, hundred images of different resolution are taken. After applying a coarse dewarping which is done by transforming the curled document image with a rectangular model, not a better dewarped result is obtained. So, in order to obtain a better dewarped result, each of the characters in the document image is enclosed within rectangular boxes and a fine dewarping is done on word level.

Table 6.1 Angle measurements of dewarped images

Image	Image 1	Image 2	Image 3
Angle of original image	0.0800	-1.9800	-1.4300
Rotated angle of dewarped image	2.24300	3.1700	3.4300

Images with different warping amount are taken to check the angle variation of document images. The original angle of the warped image and the measured angle of the dewarped images are measuered. This thesis work can correctly rectify the distortions upto angle of -2.

OCR evaluation can be done for checking the accuracy of the image. Hundred images of different resolution before and after dewarping are taken for checking the accuracy of the image. OCR accuracy defined as the ratio of number of correct characters to the total number of characters in the document.

Table 6.2: OCR accuracy

Without dewarping	55.07%
With coarse to fine dewarping	85.56%
With coarse to fine dewarping using enclosed box method	93.98%

This thesis work used data set containing different font sizes and different fonts. The methodology can dewarp document images irrespective of font size and font diversities. This work requires approximately 12 to process one page.

V. CONCLUSION

Document imaging analysis has interest over past few years. Many types of distortion affects document images like warping affect. Warping effect reduces the OCR accuracy and also OCR systems cannot read the curled document images. Many dewarping methods are introduced to straighten the curved document images. In this work, coarse to fine dewarping using enclosed box method is introduced. The curled surface area is projected by a rectangular transformation model to achieve coarse and fine dewarping. In order to enhance dewarping at word level each of the characters in the document image are enclosed within rectangular boxes. Each of the letters is translated and rotated using fine dewarping to improve the dewarping result. Results show that this method can dewarp document images efficiently and improves the OCR accuracy.

REFERENCES

[1] N. Stamatopoulos, B. Gatos, I. Pratikakis and S. J. Perantonis. “A Two-Step Dewarping of Camera Document Images” The Eighth IAPR Workshop on Document Analysis Systems.

[2] J. Liang, D. Doermann, and H. Li. “Camera-based analysis of text and documents: a survey” *InternationalJournal on Document Analysis and Recognition*, 7(2-3), 2005, pp. 84–104.

[3] F. Shafait, T. M. Breuel, “Document Image Dewarping Contest”, *In 2nd Int. Workshop on Camera-Based Document Analysis and Recognition*, Curitiba, Brazil, 2007, pp. 181-188.

[4] A.Yamashita, A.Kawarago, T.Kaneko, K.T.Miura, “Shape Reconstruction and Image Restoration for Non-Flat Surfaces of Documents with a Stereo Vision System”, *International Conference on Pattern Recognition*, vol. 1, Cambridge, UK, 2004, pp. 482-485.

[5] M.S. Brown & W.B. Seales, “Image Restoration of Arbitrarily Warped Documents”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(10), 2004, pp. 1295- 1306.

[6] C.L. Tan, L. Zhang, Z. Zhang & T. Xia, “Restoring Warped Document Images through 3D Shape Modeling”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28(2), 2006, pp. 195-208.

[7] A. Ulges, C. H. Lampert, and T. Breuel, “Document capture using stereo vision” *ACM Symposium on Document Engineering*, Milwaukee, Wisconsin, USA, October 28-30, 2004, pp. 198–200.

[8] N. Gumerov, A. Zandifar, R. Duraiswarni, and L. S. Davis, “Structure of applicable surfaces from single views” *European Conference on Computer Vision*, Prague, 2004, pp. 482–496.

[9] H. Cao, X. Ding and C. Liu, “A cylindrical surface model to rectify the bound document image” *International Conference on Computer Vision*, vol. 1, Nice, France, 2003, pp. 228.

[10] A. Masalovitch and L. Mestetskiy, “Usage of continuous skeletal image representation for document images dewarping” *In 2nd Int. Workshop on Camera-Based Document Analysis and Recognition*, Curitiba, Brazil, 2007, pp. 45-53.

[11] O. Lavaille, X. Molines, F. Angella and P. Baylou, “Active Contours Network to Straighten Distorted Text Lines” *International Conference on Image Processing*, Thessaloniki, Greece, 2001, pp. 1074-1077.

[12] Z. Zhang & C. L. Tan, “Correcting document image warping based on regression of curved text lines” *International Conference on Document Analysis and Recognition*, Edinburgh, Scotland, 2003, pp. 589–593.

[13] H. Ezaki, S. Uchida, A. Asano & H. Sakoe, “Dewarping of document image by global optimization” *International Conference on Document Analysis and Recognition*, Seoul, Korea, 2005, pp. 500-506.

[14] C. Wu & G. Agam, “Document image De-warping for text/graphics recognition” *Joint IAPR International Workshop on Structural, Syntactic, and Statistical Pattern Recognition*, Windsor, Canada, 2002, pp. 348-357.

[15] A. Ulges, C.H. Lampert & T.M. Breuel, “Document image dewarping using robust estimation of curled text lines” *International Conference on Document Analysis and Recognition*, Seoul, Korea, 2005, pp. 1001- 1005.

[16] Shijian Lu, Chew Lim Tan, “The Restoration of Camera Documents through Image Segmentation”, *Workshop on Document Analysis Systems VIII*, Nelson, New Zealand, 2006, pp. 484-495.

[17] B. Gatos, I. Pratikakis and I. Ntirogiannis, “Segmentation Based Recovery of Arbitrarily Warped Document Images” *International Conference on Document Analysis and Recognition*, Curitiba, Brazil, 2007, pp. 989-993.

[18] J. Liang, D. DeMenthon, and D. Doermann, “Flattening curved documents in images” *Conference on Computer Vision and Pattern Recognition*, San Diego, USA, 2005, pp. 338–345.

[19] Y. C. Tsoi and M. S. Brown, “Geometric and shading correction for images of printed materials–A unified approach using boundary” *Conference on Computer Vision and Pattern Recognition*, Washington, DC, 2004, pp. 240-246

[20] Z. Zhang & C. L. Tan, “Warped image restoration with applications to digital libraries”, *International Conference on Document Analysis and Recognition*, Seoul, Korea, 2005, pp. 192-196

[21] M. Wu, R. Li, B. Fu, W. Li and Z. Xu, “A Page Content Independent Book Dewarping Method to Handle 2D Images Captured by a Digital Camera*”, International Conference on Image Analysis and Recognition*, Montreal, Canada, 2007, pp. 1242-1253.

[22] B. Gatos, I. Pratikakis & S.J. Perantonis, “Adaptive Degraded Document Image Binarization”, *Pattern Recognition*, 39, 2006, pp. 317-327.

[23] N. Stamatopoulos, B. Gatos, and A. Kesidis, “Automatic Borders Detection of Camera Document Images” *In 2nd Int. Workshop on Camera-Based Document Analysis and Recognition*, Curitiba, Brazil, 2007, pp. 71-78.

[24] U. V. Marti and H. Bunke, “Using a statistical language model to improve the performance of an HMM-based cursive handwriting recognition system”, *International journal of Pattern Recognition and Artificial intelligence*, 15(1), 2001, pp. 65–90.

[25] http://quito.informatik.uni-kl.de/dewarp/dewarp.php

[26] ABBYY FineReader OCR: http://finereader.abbyy.com/