

Warped Document Image Correction Based on Checkboard Pattern and Geometric Transformation

Marian Wagdy¹, Khaild Amin² and Mina Ibrahim³

*Department of Information Technology
Faculty of Computers and Information, Menoufia University, Egypt*

Email: Marian_wagdy_labeeb@yahoo.com¹, kh.amin.0.0@gmail.com², Mina.ibrahim@ci.menofia.edu.eg³

Abstract— Document image warping problem refers to the process of geometrically transforming 2D images. In this work we aim to solve the warped document image problems. The idea of the proposed method is inspired from the concept of image registration to align two images (distorted document and undistorted ground truth checkboard). In this paper, we propose document image dewarping method based on checkboard calibration pattern and geometric transformation. Some pairs of control points are selected between the two images to define the mapping function between them. The dewarping process transforms the warped document image according to the geometric transformation defined by the calculated mapping function. Results on document dewarping dataset CBDAR 2007 demonstrate the success of our strategy. OCR (Optical Character Recognition) error metrics are also used to gauge the success of the suggested approach. Based on the quantitative and qualitative metrics the proposed method outperforms the compared state of the art method.

Keywords— *Warped document image, image registration, polynomial coefficient, least square method, OCR error measure.*

I. INTRODUCTION

In recent years, everyone has his own handheld digital devices such as PDAs and camera phones which use them in capturing any documents such as posters, magazine, and books [1]. This is the simplest way to disseminate and collect information. Scanners are used to transfer the documents into digital format in the past few decades. Efficient storage, management and retrieval of the digitized document images are extremely important in many office automation and digital library applications.

Another scenario when deal with valuable historical manuscripts, the flatbed scanner cannot be used because it is difficult to move the historical document from its place to another one. The historical document by its nature suffers from different distortions, and it has gone through many years, so it became damaged. Also, such these documents can be affected by the heat of the device and physical pressure. In such cases, the standard digital camera is the best solution.

Unfortunately, the captured document images are suffering from a lot of problems such as perspective and geometric distortions. Warping and curled lines are the most common problems appear especially, when deal with thick book and rolled document [1]. Examples of warped document images of bound book are illustrated in Fig. 1.

Document image warping problem refers to the process of geometrically transforming 2D images. Although the word warp may seem to suggest a radial distortion. Document image warping includes the entire range of transformation from complex irregular warps to the simple ones such as rotation or scaling. There are numerous and diverse applications for digital image dewarping [2]. Most of image processing applications have at least one need to translation, scaling or rotation of the image. All these operations are simple examples of warping. One of the most important applications of image dewarping is in remote sensing. Images which are taken from different sensors, satellites, or aircraft suffer from many geometric distortions which are caused by camera lenses or curvature of the earth. Image dewarping is efficiently used to correct these geometric and lens distortions. Recently, image dewarping has much attention focused on a special effect's technique called morphing.

This work focuses on correcting the warping problem of the document images that captured by camera. Optical Characters Recognition (OCR) engines use the captured document image to perform different operations such as, text extractions and character recognition. The accuracy of OCR systems is mainly depending on the dewarping process [3].

The proposed method based on calibration pattern has the following advantages:

- Accurate to recover the shape of the document.
- Recover the shape in the case of complex layout.
- Recover the shape regardless the content of the document.
- No requiring for expensive hardware or software.
- Accurate in the case of bad illumination.

The proposed method based on a well-defined pattern to simulate the distorted and undistorted document images. The proposed method uses the checkboard calibration patterns to get the world and corresponding image points. Mapping function is calculated based on the selected world and image points, regardless the content of the document image. Finally, the dewarping process transforms the warped document image according to the geometric transformation defined by the calculated mapping function.

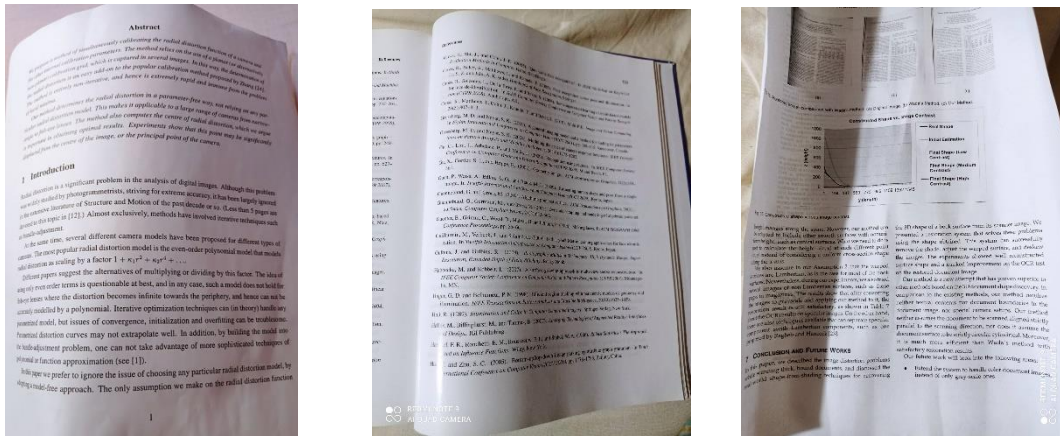


Fig. 1. Examples of warped document images

The remainder portion of this article is structured accordingly. A summary of the relevant document image dewarping techniques is provided in Section II. Section III describes the proposed procedure. Section IV illustrates different examples to explain the usefulness of the procedure proposed. Finally, in Section V, the conclusion is drawn.

II. RELATED WORK

Dewarping methods of paper images have generated a lot of interest in the science community over the past few years and several techniques have been suggested. The accessible techniques are divided into three categories: techniques based on text knowledge, techniques based on recovering of shape, and techniques based on deep learning.

The first type is based on the extraction of features from text characters, words, and lines. A coupled-snakes model based on ridges was used in the approaches in [3,4]. The baseline was measured using a coupled-snakes model, then the starting location of each line of text was established. Eventually, these approaches used a four-point homography approach to eliminate the illusion of perspective. Schneider et al. [5] implemented an approach of adding a vector field from the articulation mesh based on local orientation features. The document is corrected roughly by several linear projections for nonlinear distortion. Baseline detector is used to extract the appropriate local features of orientation. Kim et al. [6] suggested a system based on discrete representation. The features of a text line were extracted. The final aspect is the cost function used by Levenberg-Marquadt approach to resolve the

warping. Few strategies to solve dewarping have been suggested, based on word knowledge. By splitting the text into words, B.Gatos et al. [7] suggested an idea. The baseline and orientation are sensed for each term. Finally, to restore the warping problem, each word is rotated.

A curled pair of baselines were used by Bukhari et al. [8]. Map characters are determined for each respective upper and lower pair of baselines. Zhang et al. [9] have carried out the recovery process based on segmentation and thin plates. The text and vertical stroke boundaries for each line have been set out in Lu et al. [10]. Ulges et al. [11] demonstrated an algorithm covering each quadrilateral cell character and then computed each cell center to correct each cell's path to generate flat text. Generally, text-based techniques are the most common and perform well if the algorithm for segmentation is correct.

Shape recovery is the second category. These methods are based on recovering the shape of document. Researchers are seeking to restore the form of the text based into three methods 2D, 3D and shape from shading-based recovery. Approaches focused on the recovery of the 2D shape used a curved surface projection map to retrieve a rectangular 2D surface as seen in [12]. The physical pattern-based algorithm in [13] directs the uniform parameterization. The approach in [14] divided the text into three vertical sections and projected each part separately. The methods suggested by [15-16] have been used to detect the texts to describe the curved shape by using the neighboring model of transformation.

2D-based approaches for shape recovery depend not on any priori information such as camera parameters, scanners, and page curl geometry. These methods often rely only on image content features extracted. The precision of 2D-based techniques, on the other hand, relies on how well the boundaries of the curved structure, which implies constraint, are derived.

Researchers have suggested 3D reconstruction strategies to address the issue of dewarping by transforming the curled text into 3 dimensions (height, width, and depth). By the identification of the depth axis, the dewarping distortion can be corrected, and the curved structure can then be fitted [17- 20]. Special hardware such as laser scans, stereo cameras [21], eye scanner [22], organized light [23], CT scans for objects opaque [24- 25] is required in this process. In order to get a 3D shape of the curved paper, prior information is also needed [26].

In comparison to algorithms based on 2D, most of these approaches are not contents-dependent, so these techniques are more precise than techniques based on 2D to represent the structure. In case of complicated distortions such as folds and rumples, 3D-based approaches are also effective. 3D-based approaches often require special plans to record 3D geometry. More costly technology and hardware is also required. Additional approaches for reconstructing the document's surface shape based on the basis of the shading techniques were suggested. The definition of basis from shading approaches is based on the recovery of shape by using various illuminations called shading. Approaches are provided by Zhang et al. [27-32] using a shading algorithm to create a 3D model.

The last group is approaches focused on deep learning. A new approach based on deep learning was introduced by Saqib et. al. [33]. Their method of solving the dewarping problem is based on the network for deep learning image to image conversion. It transforms the document from one domain to another by taking the damaged and the clean document images. Pix2pixhd network is used in this method [34], which involves high resolution with the aid of GANs. Generative Adversarial Network (CGAN) and Pix2pix networks [35] fail to produce the document of size of 256x256. In the case of a paper that has figures, charts and tables, deep learning-based approaches perform well [36]. However, they need a great deal of data and costly GPUs.

Finally, 3D page shape reconstruction methods produce high quality results to recover the shape of the document regardless the content of the document image. However, specialized, and expensive hardware is required. Furthermore, prior knowledge parameters from camera and scanner are needed. In contrast 2D image processing techniques did not require any parameters or hardware. Unfortunately, recovering the shape in the case of complex layout and bad illuminations is not accurate. Table I presents the summarization of document image dewarping techniques.

In this paper, we can overcome the limitations of the previous dewarping algorithms to recover the shape of the document. The proposed method based on a well-defined pattern to simulate the distorted and undistorted document images. We developed a method which can effectively recover the curved shape regardless the content of the document, and no requiring for expensive hardware or software. The idea of the proposed method is inspired from the concept of image registration to align two images. The first step is using the well-defined checkboard pattern to compute world or fixed points. The second step is using the world or fixed points obtained from the first step to find the corresponding image or moving points from the warped document. The selected pairs of control points (fixed

and moving points) from the previous steps are used to compute the correct mapping function between the warped document image and checkboard pattern as illustrated in the third step. Finally, the dewarping process transforms the warped document image according to the geometric transformation defined by the calculated mapping function.

TABLE I: SUMMARIZATION OF DOCUMENT IMAGE DEWARPING TECHNIQUES.

Category	Techniques	Main idea	Strengthens	drawbacks
Text information-based techniques	Text line-based methods [3- 6]	Divide the document into text lines and use the features extraction to flatten the document.	<ul style="list-style-type: none"> • Most popular one. • Easy to understand. • Easy to implementation. • Accurate if the segmentation algorithm is accurate. 	<ul style="list-style-type: none"> • Sensitive to high degrees of curl and variable line spacing. • Sensitive to the resolution and language. • Time consuming. • Fail in case of complex layout. • Fail in the case of document that contains images and tables.
	Word based methods [7]	Divide the document into words and correct the angle of each word.		
	Character based methods [8- 11]	Divide the document into characters and extract the feature to flatten it.		
Shape recovery-based techniques	2D based shape recovery [12- 16]	Detect the boundary of the warped shape then project it to rectangle shape.	Depend only on features extracted from the image and do not rely on any kind of a priori knowledge.	Not accurate to recover the shape
	3D based shape recovery [17- 26]	Use specialized hardware to recover the 3D shape of the warping document	Accurate representation of physical warping for different structures.	Require specialized and expensive hardware.
	Shape from X based recovery [27- 32]	Use different illuminations to recover the 3D shape of the warping document	Do not Require specialized and expensive hardware.	Fail in bad illuminations.
Deep learning	Deep learning-based methods [33- 36]	These methods depend on image-to-image translation deep learning network	Do not depend on any segmentation or line alignment, also no need for any preprocessing steps for the input document image.	It requires very large amount of data, requires expensive GPUs and hundreds of machines. This increases cost to the users.

III. PROPOSED METHOD

The idea of the proposed method is inspired from the concept of image registration to align two images. Suppose we are given two images for the same document, the original document image, and the warped document image. To observe the changes between these two images, we need to find the correct mapping function between the two. Once the mapping function is determined, the dewarping problem can be accomplished. In this work we use the idea of image alignment to determine the correct mapping function between two images. To get the correct mapping function, we need to find more pairs of corresponding control points. The main problems here are as the following:

1. The original document image without warping is not available; we have only the warped or distorted document image.

2. Which control points to select in one image.
3. How to find the corresponding points in the other image.

What we should to do is to provide sample image of a well-defined pattern (eg, checkboard) to simulate the original document image. Some control points like square corners in checkboard are selected. These corner points represent the world or fixed points. The corresponding points are selected from the warped document image which they represent the image or moving points. The pairs of corresponding control points (image and world) are used to calculate the correct mapping function.

The proposed method consists of six principal steps. The first step includes removing the bad illumination and shading from the document image. In the second step the skew distortion has been detected and corrected. Border noise removal is used in the third step to define the actual text border. Checkboard pattern ground truth is generated in the fourth step to simulate the original document image. The control points are selected in the fifth step. Mapping function is computed based on the selected control points in the sixth step to dewarp the document image. Each step is explained in detail in the following subsections. The general block diagram of the proposed method is shown in Fig. 2.

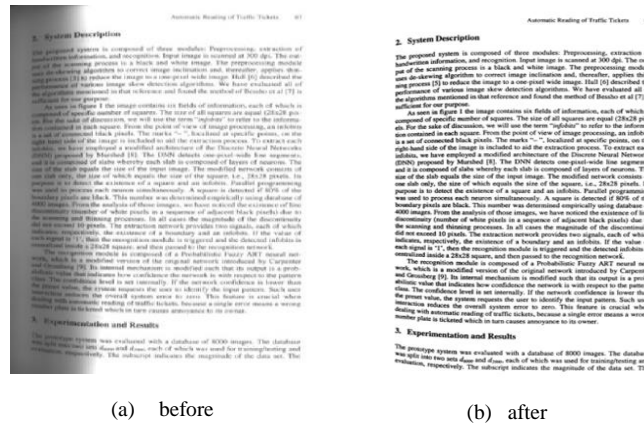


Fig. 3. Document image before and after illumination removal.

A. Degradation Enhancement

When capture page of thick book, the resulted document image has shading in the spine area of the bound volume as shown in Fig. 3.a. OCR engines and segmentation algorithms face a lot of challenges to recognize the text in the case of bad illuminations. In this work the proposed method removes this limitation by using Retinex theory used in [37]. Mathematically, in Retinex theory the document image is characterized as a product of its reflectance $F(x,y)$ with the illumination $U(x,y)$,

$$I(x,y) = U(x,y) \cdot F(x,y) \tag{1}$$

By using the low-frequency component of the measured signal, the illumination $L(x,y)$ component can be approximated by using smooth filtering technique with a large kernel, it is calculated under the assumption that the lighting variable varies smoothly across the entire image. To obtain a smooth version of the image, the proposed method uses the median filter.

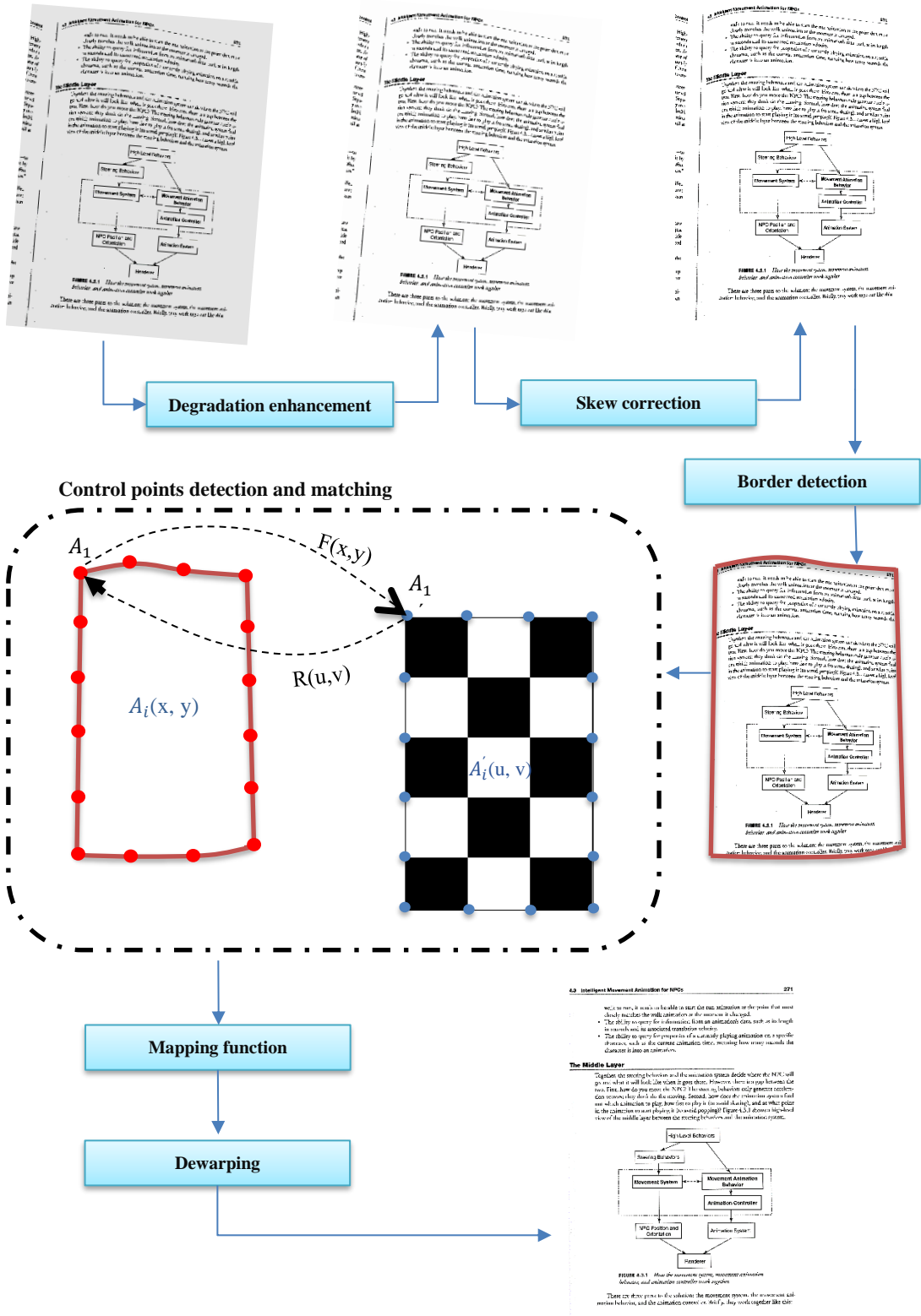


Fig. 2. Block diagram of the proposed method.

Lightness is the estimation of the reflectance $F(x, y)$. It can be calculated as,

$$F(x, y) = \frac{I(x, y)}{U(x, y)} \quad (2)$$

After removing the illumination, it is easy to binarize the document image by using global binarization method used in [38]. Fig. 3 demonstrates an example of document image before and after illumination removal step.

B. Skew Detection and Correction

In this step the skew distortion has been detected and corrected based on ellipse shape as in [39]. The fundamental principle of this method is to measure the inertia ellipse of each related component of the document, then find the orientation of the ellipse based on the following equation,

$$\theta = \begin{cases} \tan^{-1} \frac{(\mu_{rr} - \mu_{cc}) + \sqrt{(\mu_{rr} - \mu_{cc})^2 + 4\mu_{rc}^2}}{-2\mu_{rc}} & \text{if } \mu_{rr} > \mu_{cc} \\ \tan^{-1} \frac{-2\mu_{rc}}{(\mu_{rr} - \mu_{cc}) + \sqrt{(\mu_{rr} - \mu_{cc})^2 + 4\mu_{rc}^2}} & \text{if } \mu_{rr} \leq \mu_{cc} \end{cases} \quad (3)$$

Where θ is the orientation of the connected component. μ_{rr} , μ_{cc} and μ_{rc} are called normalized second row moment, normalized second column moment, and normalized second mixed moment.

$$\mu_{rr} = \frac{1}{A} \sum_{(r,c) \in R} (r - r_0)^2 \quad (4)$$

$$\mu_{cc} = \frac{1}{A} \sum_{(r,c) \in R} (c - c_0)^2 \quad (5)$$

$$\mu_{rc} = \frac{1}{A} \sum_{(r,c) \in R} (r - r_0)(c - c_0) \quad (6)$$

Where A is the area of connected component R , (r, c) is coordinates of each connected component, and (r_0, c_0) is the centroid.

After computing the angles of each connected component (θ_i) from equation 3, the mode is selected of non-zero values to ignore the noise, and then rotate the document by negative value of the mode. Fig. 4 shows an example image of skew detection and correction step.

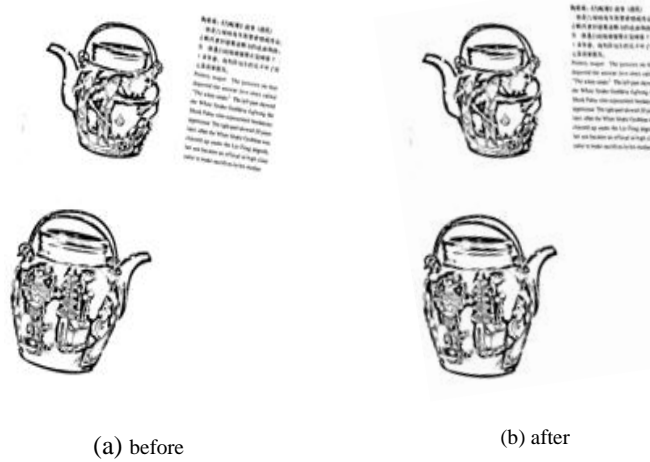


Fig.4. Document image before and after skew detection and correction.

C. Border Detection

Border noise is considered one of the most common problems that appear when capturing or scanning thick book. In this step we aim to remove this noise which comes from the adjacent pages and detect the actual page content area. Border noise removal is used to remove the non-textual and textual noise as illustrated in Fig. 5.a which appears in the borders of the document. Fig. 5 shows an example of document image before and after border noise removal and detection of the actual content area.

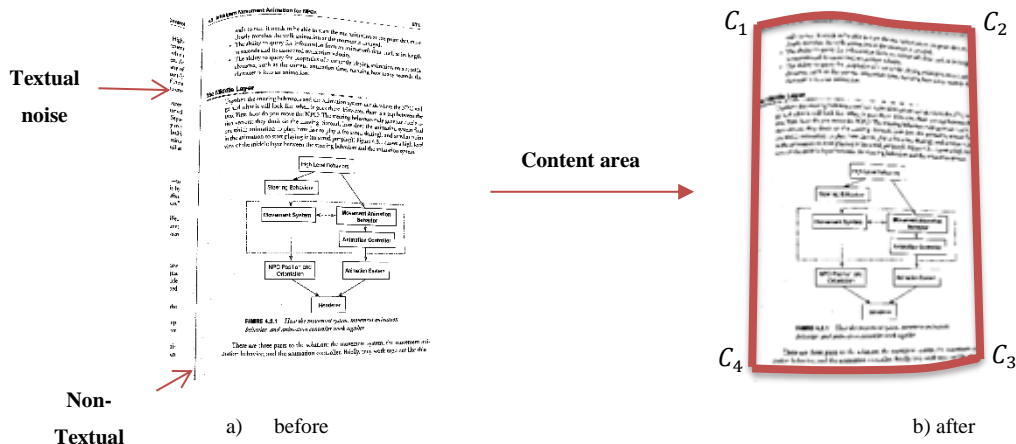


Fig. 5. Example of document image before and after border noise removal and detection of the actual content area.

X-Y cut and filtering techniques are used to remove the non-textual and textual border noise based on morphological operations. Firstly, morphological operations are applied to detect the foreground of the document image. We use line as structured element with width 100 pixel and angle 270° to maximizing the foreground area in vertical direction as shown in Fig. 6.a. The size and type of the structuring element were determined empirically. To remove the textual noise in the left and right parts, X-Y cut algorithm which is applied in [40] is used to cut the border noise. We use X-Y Cut algorithm to cut the border noise by searching for the first column that has the smallest number of black pixel (whitespace) then crop the remaining border. The noise in the upper and lower parts of the document image is non-textual noise (black border, lines, etc.) can be removed by filtering out connected components based on their size and aspect ratio. Fig. 6 shows an example of border noise removal based on morphological operations and X-Y cut algorithm.

Once the non-textual and textual noises are removed, the actual page content can be easily extracted from the warped document. The actual page content boundaries can be extracted by estimating two straight lines that fit the right and left text content along with two curved lines for top and bottom text content. Let $C_1(x, y)$, $C_2(x, y)$, $C_3(x, y)$, and $C_4(x, y)$ denote the left, right, upper and bottom corner points of the actual page content.

For left straight line C_1C_4 that fits the left text content, we detect ten leftmost points at fixed equal distances (x_{L_i}, y_{L_i}) for $i = 1, 2, \dots, 10$. The number of these points was determined empirically. For better line estimation, we ignore the titles, subtitles and empty leftmost points that do not start from the beginning of the actual page content. To eliminate the titles and subtitles leftmost points, median Y_L for all y_{L_i} of the ten leftmost points is calculated. Each leftmost point satisfies the following condition is excluded.

$$|y_{L_i} - Y_L| > 2 \text{ character width} \quad (7)$$

Finally, left straight line C_1C_4 is estimated based on Least Square Estimation (LSE) method to fit all the leftmost points (x_{L_i}, y_{L_i}) that remain after point elimination. Similarly, the rightmost points (x_{R_i}, y_{R_i}) are estimated with the same way to define the straight line C_2C_3 that fits the right text content.

For upper curve boundary C_1C_2 that fits the upper text content, we detect ten uppermost points at fixed equal distances (X_{U_i}, Y_{U_i}) for $i=1, 2, \dots, 10$ according to the following condition,

$$|y_{L_i} - Y_L| < 2 \text{ character width} \text{ and } |y_{R_i} - Y_R| < 2 \text{ character width} \quad (8)$$

By using this condition, we can select the first upper line that guarantees the line is not too small such as title or subtitle line. Coefficients of third-degree polynomial curve are calculated based on Least Square Estimation (LSE) method which fits all the uppermost points (X_{U_i}, Y_{U_i}) in the selected line. Similarly, the bottommost points (X_{B_i}, Y_{B_i}) are estimated with the same way to define the bottom curve C_4C_3 . Fig. 5.b illustrates the actual page content boundaries resulted from the content extraction process.

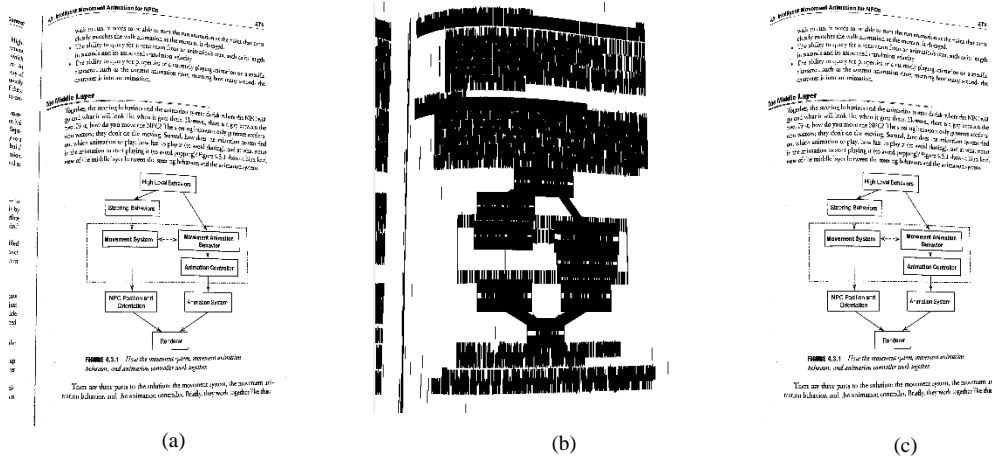


Fig. 6. Example of border noise removal. a) original document image. b) after morphological operations. c) after border noise removal using X-Y cut algorithm.

D. Checkboard Pattern Generation

Checkboard pattern is created in this step to simulate the original document image in real world. The checkboard size is defined by the smallest rectangle that is large enough to contain the extracted actual content area. The selected rectangle is divided into 5 by 3 equal squares to represent the checkboard. The generated checkboard pattern represents the original document image in real world before warping.

E. Control Points Selection

16 World or fixed points are selected from the generated checkboard. The fixed points are selected based on the checkboard corners on the boundary as shown by the blue points in Fig. 2. The fixed points obtained from the checkboard are used to find the corresponding moving points from the warped document image.

The moving points from the dewarped document image are selected by defining six control points from the left extracted border line and six from right extracted border lines. Four points are selected from upper extracted curve and four from bottom curve that corresponding to the checkboard corners. The 16 moving selected points are illustrated by the red points in the actual content boundary as shown in Fig. 2.

F. Mapping Function

The selected pairs of control points (fixed and moving points) from the previous step are used to compute the correct mapping function between the warped document image and checkboard pattern. Let $A_i(x, y)$ represents the control moving points of the dewarped document image, and $A_i(u, v)$ its corresponding fixed points of the checkboard pattern where $i = (1, \dots, n)$, n is number of control points. The objective is to find the correct mapping function between two images based on these pairs of control points. The determination of the mapping function between two images is called the registration problem, where $F(x, y)$ is called forward mapping function and $R(u, v)$ is called the inverse mapping function as in equation 9.

$$A_i(x, y) \begin{array}{c} \xrightarrow{F(x, y)} \\ \xleftarrow{R(u, v)} \end{array} A'_i(u, v) \quad (9)$$

The warping distortions or some time called radial distortions include all types of deformations that can be modeled by polynomial transformations. The mapping function of the warping distortions is approximated mathematically by set of polynomial equations [41]. The functions $F(x, y)$ and $R(u, v)$ are polynomials in x, y and u, v of degree m , whose coefficients are given by coefficient arrays p and c . The mapping function can be represented in more general form as,

$$x = \sum_{i=0}^m \sum_{j=0}^m p_{ij} u^i v^j \quad (10)$$

$$y = \sum_{i=0}^m \sum_{j=0}^m c_{ij} u^i v^j \quad (11)$$

Where m is the order of polynomial, the large numbers of m is used in more complex distortion, $m=3$ is selected in our implementation. So, the equations can be represented as the following,

$$x = p_{00} + p_{01} v + p_{02} v^2 + p_{03} v^3 + p_{10} u + p_{11} uv + p_{12} uv^2 + p_{13} uv^3 + p_{20} u^2 + p_{21} u^2 v + p_{22} u^2 v^2 + p_{30} u^3 + p_{31} u^3 v + p_{32} u^3 v^2 + p_{33} u^3 v^3 \quad (12)$$

$$y = c_{00} + c_{01} v + c_{02} v^2 + c_{03} v^3 + c_{10} u + c_{11} uv + c_{12} uv^2 + c_{13} uv^3 + c_{20} u^2 + c_{21} u^2 v + c_{22} u^2 v^2 + c_{30} u^3 + c_{31} u^3 v + c_{32} u^3 v^2 + c_{33} u^3 v^3 \quad (13)$$

For inverse mapping function,

$$u = \sum_{i=0}^m \sum_{j=0}^m a_{ij} x^i y^j \quad (14)$$

$$v = \sum_{i=0}^m \sum_{j=0}^m b_{ij} x^i y^j \quad (15)$$

$$u = a_{00} + a_{01} y + a_{02} y^2 + a_{03} y^3 + a_{10} x + a_{11} xy + a_{12} xy^2 + a_{13} xy^3 + a_{20} x^2 + a_{21} x^2 y + a_{22} x^2 y^2 + a_{30} x^3 + a_{31} x^3 y + a_{32} x^3 y^2 + a_{33} x^3 y^3 \quad (16)$$

$$v = b_{00} + b_{01} y + b_{02} y^2 + b_{03} y^3 + b_{10} x + b_{11} xy + b_{12} xy^2 + b_{13} xy^3 + b_{20} x^2 + b_{21} x^2 y + b_{22} x^2 y^2 + b_{30} x^3 + b_{31} x^3 y + b_{32} x^3 y^2 + b_{33} x^3 y^3 \quad (17)$$

Least square method is used to determine the parameters of polynomial coefficients (p, c, a, b) .

$$A = [F^T \ F]^{-1} F^T U \quad (18)$$

$$B = [F^T \ F]^{-1} F^T V \quad (19)$$

$$P = [R^T \ R]^{-1} R^T X \quad (20)$$

$$C = [R^T \ R]^{-1} R^T Y \quad (21)$$

Once the correct mapping function between the warped document image and its corresponding checkboard is calculated as presented above. The dewarping process transforms the warped document image according to the geometric transformation defined by the calculated mapping function. The transformation process is performed by scanning the warped document image pixel by pixel and computing the corresponding coordinates in the corrected image. The corresponding coordinates are calculated by evaluating the mapping function and coloring these coordinates in the corrected image with the same color of the warped pixel.

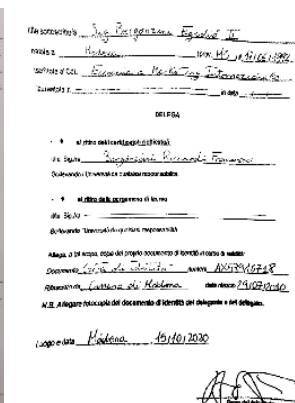
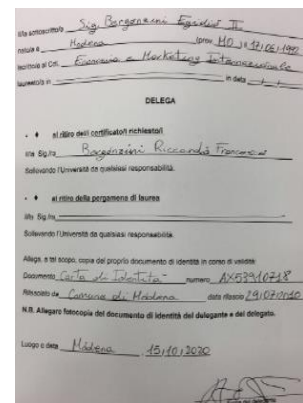
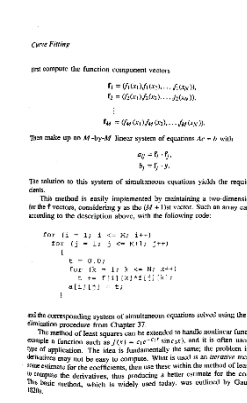
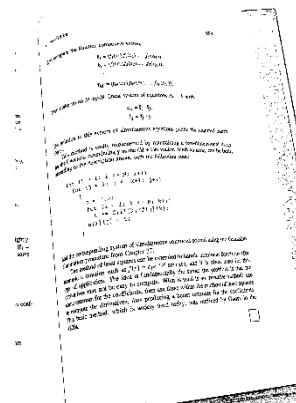
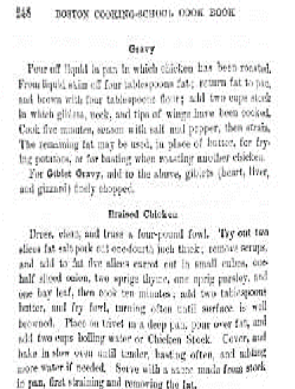
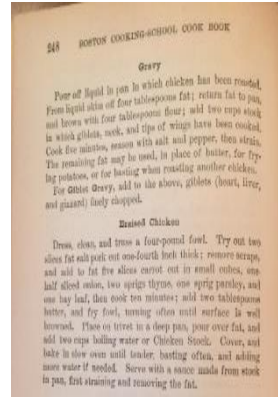
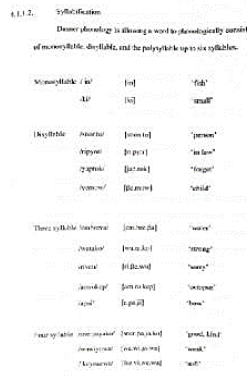
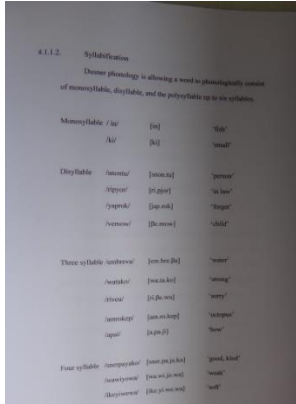
IV. DISCUSSION OF RESULTS

This section assesses the efficiency of the approach suggested. The proposed approach was applied and tested on a Core i5 2.5 GHz CPU PC using MATLAB R2020a. To show the feasibility of the proposed procedure, a number of experiments have been carried out. Warped document images with different levels of complexity are used. Part of these examples was captured by hand phone camera and the other were collected from internet. It was taken into our consideration that the captured and collected images cover all aspects of complexities, for example degradation, handwritten, different layout, border noise and different languages. Fig. 7, 8 presents the results of these samples of document image before and after warping correction. As shown in Fig. 7 and Fig. 8, the proposed method automatically detects the correct mapping function to solve the warping problem for document images that have a

lot of challenges like degradation, border noise and handwritten warped text. The results illustrate the behavior and the efficiency of the proposed method when the document images have skew and warp.

To evaluate and validate the performance of the suggested approach, we conducted two different types of tests, quantitative analysis based on OCR engine and visual perception based qualitative analysis. Based on visual criteria, our proposed method is efficient and more readable as illustrated in Fig. 7,8 and 9 because warping distortion has been solved efficiently in the case of documents that contain figures, equations, and perspective distortion.

OCR engine is used for quantitative analysis to test the text region of the proposed results against the state-of-the-art methods. In this paper, ABBYY Fine Reader OCR program is used. The ABBYY program is used to compute the number of characters in the document image before and after dewarping technique.



3.4 Applying Goal-Oriented Action Planning to Games 227

[McDermott98] McDermott, Drew, *Planning Domain Description Language*, information about PDDL2.1 available online at www.dur.ac.uk/d.p.long/IPC/pddl.html, 1998.

[Monolith02] *No One Lives Forever 2: A Spy in H.A.R.M.'s Way*, Monolith Productions/Sierra Entertainment, Inc., 2002. Toolkit and SDK available at nolf2.sierra.com/

[Nilsson98] Nilsson, Nils J., *Artificial Intelligence: A New Synthesis*, Morgan Kaufmann, 1998.

[Russell02] Russell, Stuart, and Norvig, Peter, *Artificial Intelligence: A Modern Approach (Second Edition)*, Prentice-Hall, Inc., 2002.

3.4 Applying Goal-Oriented Action Planning to Games 227

[McDermott98] McDermott, Drew, *Planning Domain Description Language*, information about PDDL2.1 available online at www.dur.ac.uk/d.p.long/IPC/pddl.html, 1998.

[Monolith02] *No One Lives Forever 2: A Spy in H.A.R.M.'s Way*, Monolith Productions/Sierra Entertainment, Inc., 2002. Toolkit and SDK available at nolf2.sierra.com/

[Nilsson98] Nilsson, Nils J., *Artificial Intelligence: A New Synthesis*, Morgan Kaufmann, 1998.

[Russell02] Russell, Stuart, and Norvig, Peter, *Artificial Intelligence: A Modern Approach (Second Edition)*, Prentice-Hall, Inc., 2002.

Fig. 7. Examples 1 of document images with different levels of complexity before and after proposed method correction.

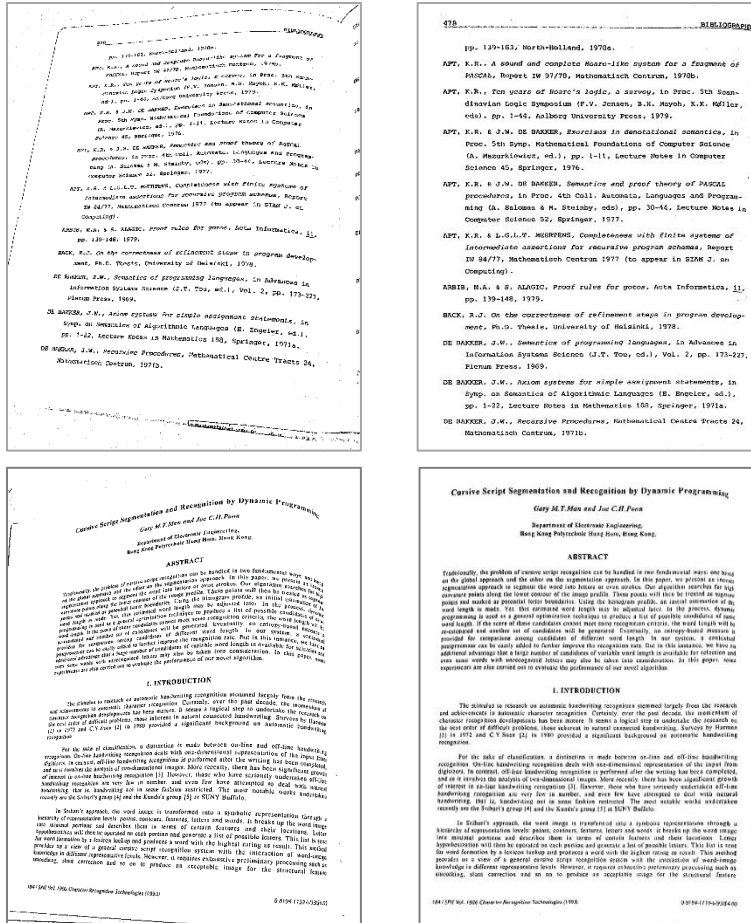


Fig. 8. Examples 2 of document images with different levels of complexity before and after proposed correction.

OCR Error Rate (E_r) metric is the common metric that is used for evaluating the ratio of the number of incorrect characters to the actual number of characters in the illustration of a text is calculated.

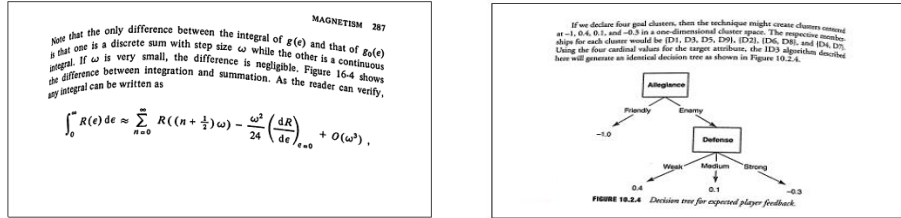
$$E_r = \frac{\text{incorrect character}}{\text{all character}} * 100 \% \quad (22)$$

OCR Edit distance [42, 43] is another OCR error measure which returns the lowest number of changes such as deletions, insertions, and substitutions between the dewarped document image and its ground truth. The mean and median are the normal statistical metrics that calculated from the defined OCR Edit distance to evaluate average changes of the proposed method.

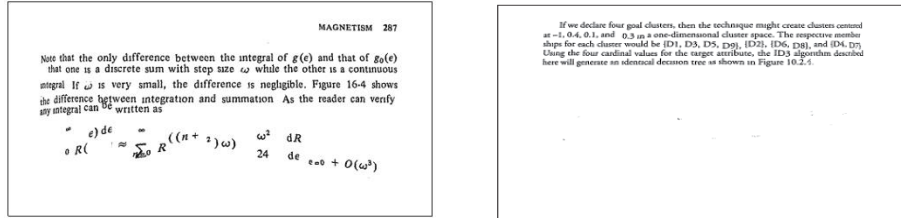
The performance evaluation metrics require ground-truth dewarped images. So, the only available document dewarping dataset (CBDAR 2007) is used to evaluate the proposed method. CBDAR 2007 document dewarping dataset contains 102 binarized document images. These documents were captured from several technical books using an off-the-self hand-held digital camera in a normal office environment. These documents have a lot of border noise (textual and non-textual) with different structure and layout. Furthermore, the ground truth ascii text of each document is also included.

We compared our approach with other state-of-the-art approaches in [44], [8], [14]. Fig. 9 presents two examples from the dataset. The selected two images have one of the most challenges that most of the previous methods failed to solve. Fig. 9.a is the original warped document image from dataset. Fig. 9.b shows the resulted document images using method in [44]. As shown in Fig. 9.b, the method in [44] produced bad results in the equation part of the left image and failed when the right document image contains figures. Fig. 9.c shows the resulted document image using method in [8]. The results of the method in [8] are better than the method in [44] in

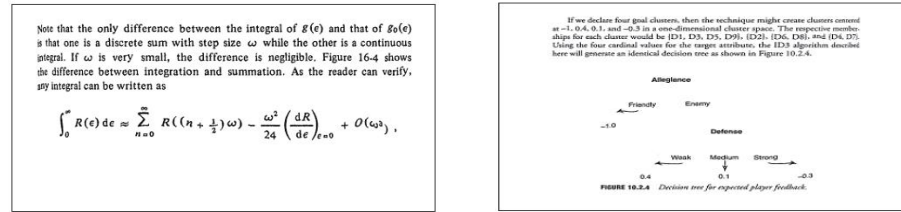
the left image but still failed in the right document image due to the figure inside the document. Fig. 9.d shows the resulted dewarped document images using method in [14]. The results of the recent method in [14] solved the problem of the document that contains figures but removed the title of the figure. Furthermore, it failed to dewarp the equation and text in the left image. Fig. 9.e illustrates the resulted dewarped document images using our method. The proposed method as illustrated produces straight lines in the text part and efficient in the case of documents that contain figures and equations. Table II introduces a comparison of the state of arts methods [8][44][14] with respect to OCR Error Rate (Er) performance.



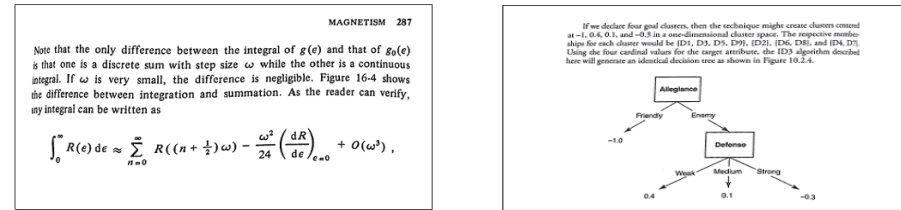
(a) Warped document images in CBDAR dataset.



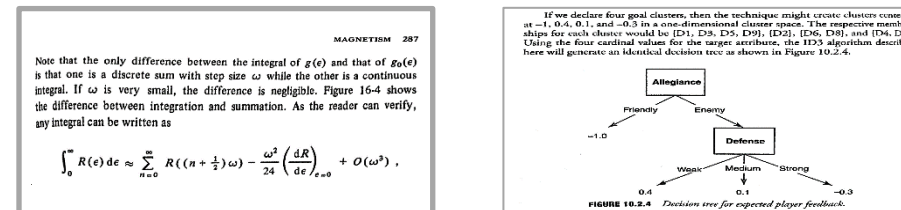
(b) method in [44].



(c) method in [8].



(d) method in [14].



(e) Our proposed method.

Figure 9. The comparison between our approach and other state of arts methods using CBDAR dataset as in [14].

TABLE II: OCR PERFORMANCE (ERROR RATE ER %)

Warped images	42 %
Method in [8]	25 %
Method in [44]	19 %
Method in [14]	12 %
Proposed method	5 %

We compare our proposed method with the available comparison between state of arts methods [7][44][45][4] as in [4] using ABBYY Fine Reader. Table III presents comparison for edit distance between the proposed method and other state of art methods. The results of the proposed method outperform the state of arts method with respect to median and mean edit distance.

TABLE III OCR PERFORMANCE (EDIT DISTANCE %)

Algorithm	Mean edit distance %	Median edit distance%
SEG [7]	4.09	2.22
SKEL [44]	2.15	0.96
CTM [45]	2.12	0.89
CTM2 [45]	1.76	0.83
Vinod et. al. [4]	1.97	0.75
Proposed	0.72	0.61

Based on the experimental results, from tables II, III and Fig. 7,8, and 9 the proposed method outperforms the other state-of-arts methods regardless the content of the document images. It produces high quality results in case of figures, border noise, tables, equations with different levels of complexity and bad illuminations. However, it fails to recover the shape of the document in the case of complex layout. The accuracy of the proposed method depends on the detection of document boundary which is a limitation.

V. CONCLUSION

In this paper we introduce an efficient dewarped document image technique based on checkboard calibration pattern and geometric transform. Pairs of control points are selected from the warped document and its corresponding checkboard pattern. Mapping function between them has been defined to correct the warping problem. Geometric transformation defined by the calculated mapping function is used to dewarp the document image. We have shown the consistency of the proposed approach for a number of warped document images from CBDAR 2007 dataset. The experimental results showed a better quality compared with other state-of-arts methods. The proposed method corrects the warping problem from the document images that have different levels of complexity. The proposed method succeeds to recover the document image which contains figures and tables.

REFERENCES

- [1] M. Wagdy, I. Faye, D. Rohaya, "Border Noise Removal and Clean Up Basedon Retinex Theory", DaEng, Lecture Notes in Electrical Engineering Volume 285, 345-352 (2014).
- [2] M. Wagdy, I. Faye, D. Rohaya, "Fast and Efficient Document Image Clean Up and Binarization Based on Retinex Theory", IEEE 9th International Colloquium on Signal Processing and its Applications, CSPA 58-62 (2013).
- [3] Bukhari, Syed Saqib, Faisal Shafait, and Thomas M. Breuel. "Dewarping of document images using coupled-snakes." Proceedings of Third International Workshop on Camera-Based Document Analysis and Recognition, Barcelona, Spain. (2009).
- [4] Vinod H.C, S.K Niranjana "dewarping of camera captured document images" IEEE International Symposium on Consumer Electronics (ISCE), pp. 13-18, (2017).
- [5] Schneider, D., Marco Block, and Raúl Rojas. "Robust document with interpolated vector fields." Ninth International Conference on Document Analysis and Recognition (ICDAR 2007). Vol. 1. IEEE, (2007).

- [6] Kim, Beom Su, Hyung Il Koo, and Nam Ik Cho. "Document dewarping via text-line based optimization." *Pattern Recognition* 48.11 3600-3614: (2015).
- [7] Gatos, Basilios, Ioannis Pratikakis, and Konstantinos Ntirogiannis. "Segmentation based recovery of arbitrarily warped document images." *IEEE Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*. Vol. 2, (2007).
- [8] Bukhari, Syed Saqib, Faisal Shafait, and Thomas M. Breuel. "Dewarping of document images using coupled-snakes." *Proceedings of Third International Workshop on Camera-Based Document Analysis and Recognition, Barcelona, Spain*. (2009).
- [9] Zhang, Yu, "Arbitrary warped document image restoration based on segmentation and Thin-Plate Splines." *IEEE 19th International Conference on Pattern Recognition*, (2008).
- [10] Lu, Shijian, Ben M. Chen, and Chi Chung Ko. "A partition approach for the restoration of camera images of planar and curled document." *Image and Vision Computing* 24.8 837-848: (2006).
- [11] Ulges, Adrian, Christoph H. Lampert, and Thomas M. Breuel. "Document image dewarping using robust estimation of curled text lines." *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*. IEEE, (2005).
- [12] Bolelli, Federico. "Indexing of historical document images: Ad hoc dewarping technique for handwritten text." *Italian Research Conference on Digital Libraries*. Springer, Cham, (2017).
- [13] M. Tappen, W.T. Freeman, and E. Adelson. "Recovering intrinsic images from a single image". In *Advances in Neural Information Processing Systems 15 (NIPS)*. MIT Press, (2003).
- [14] Shamqoli, Maryam, and Hossein Khosravi. "Warped document restoration by recovering shape of the surface." *2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP)*. IEEE, (2013).
- [15] Stamatopoulos, Nikolaos, "A two-step dewarping of camera document images." *2008 The Eighth IAPR International Workshop on Document Analysis Systems*. IEEE, (2008).
- [16] Sruthy S, S. Suresh Babu. "Dewarping on Camera Document Images", *International Journal of Pure and Applied Mathematics*, Volume 119 No. 16, (2018).
- [17] Flagg, Christopher, and Ophir Frieder. "Searching Document Repositories using 3D Model Reconstruction." *Proceedings of the ACM Symposium on Document Engineering 2019*. ACM, (2019).
- [18] Y. Lin and W. Seales, "Opaque document imaging: Building images of inaccessible texts," In *Proc. of the Tenth IEEE Int. Conf. on Computer Vision (ICCV05)*, vol. 1, pp. 662–669, (2005).
- [19] Zhang, Li, Yu Zhang, and Chew Tan. "An improved physically-based method for geometric restoration of distorted document images." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30.4 (2008).
- [20] Luis Galarza, Harold Martin, "Integrating low-resolution depth maps to high-resolution images in the development of a book reader design for persons with visual impairment and blindness", *International Journal of Innovative Computing, Information and Control*, Volume 14, Number 3, June (2018).
- [21] Y. Y. Tang and C. Y. Suen, "Image transformation approach to nonlinear shape restoration," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 1, no. 23, pp. 155–171, (1993).
- [22] K. B. Chua, L. Zhang, Y. Zhang, and C. L. Tan, "A fast and stable approach for restoration of warped document images," *IAPR Int. Conf. on Document Analysis and Recognition*, no. 1, pp. 384–388, (2005).
- [23] T. Amano, T. Abe, T. Iyoda, O. Nishikawa, and Y. Sato, "Camera-based document image mosaicing," *Proc. SPIE*, vol. 4669, pp. 250–258, (2002).
- [24] M. Brown and C. Pisula, "Conformal deskewing of non-planar documents," In *Proc. of the 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR05)*, vol. 1, pp. 998–1004, (2005).
- [25] W. Seales and Y. Lin, "Digital restoration using volumetric scanning," *Joint ACM/IEEE Conf. on Digital Libraries*, vol. 1, pp. 117–124, (2004).
- [26] Y. Lin and W. Seales, "Opaque document imaging: Building images of inaccessible texts," In *Proc. of the Tenth IEEE Int. Conf. on Computer Vision (ICCV05)*, vol. 1, pp. 662–669, (2005).
- [27] Zhang, L. and Tan, C. L. "Warped document image restoration using shape from-shading and physically-based modeling". In *8th IEEE Workshop on Applications of Computer Vision (WACV)*, 20-21 February, Austin, Texas, USA, page 29 (2007).
- [28] Zhang, AM Yip, MS Brown, CL Tan "A unified framework for document restoration using inpainting and shape-from-shading." *Pattern Recognition* 42.11 2961-2978: (2009).
- [29] Zhang, Li, Andy M. Yip, and Chew Lim Tan. "Photometric and geometric restoration of document images using inpainting and shape-from-shading." *Proceedings of the National Conference on Artificial Intelligence*. Vol. 22. No. 2. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, (2007).
- [30] Zhang, Li, Andy M. Yip, and Chew Lim Tan. "Removing shading distortions in camera-based document images using inpainting and surface fitting with radial basis functions." *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*. Vol. 2. IEEE, (2007).
- [31] Wada, Toshikazu, Hiroyuki Ukida, and Takashi Matsuyama. "Shape from shading with interreflections under a proximal light source: Distortion-free copying of an unfolded book." *International Journal of Computer Vision* 24.2 (1997): 125-135.
- [32] Z. Zhang and C. L. Tan. "Restoration of images scanned from thick bound documents". In *ICIP '01, Thessaloniki, Greece, Oct (2001)*.
- [33] L. Zhang, A. Yip, C. Tan, "Shape-from-shading based on Lax-Friedrichs fast sweeping and regularization techniques with applications to document image restoration" in: *IEEE Conference on Computer Vision and Pattern Recognition*, (2007).
- [34] Saqib Bukhari, Vijaya Kumar, Bajjer Ramanna, "Document Image Dewarping using Deep Learning", 8th international conference on pattern recognition applications and methods ICPRAM, (2019).
- [35] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. "Image quality assessment: from error visibility to structural similarity". *IEEE transactions on image processing*, 13(4):600–612, (2004).
- [36] Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. "A. Image-to-image translation with conditional adversarial networks". *arXiv preprint* (2017).
- [37] M. Wagdy, I. Faye, D. Rohaya, "Document Image Binarization Based on Retinex Theory", *Electronic Letter on Computer Vision and Image Analysis ELCVIA Vol 14, No 1*, 2015.
- [38] M. Wagdy, I. Faye, D. Rohaya, "Degradation Enhancement for the Captured Document Image Using Retinex Theory", *IEEE Information technology and multimedia ICIMu 2014*, ISBN: 978-1-4799-5423
- [39] Marian Wagdy, Mina Ibrahim, Khaild Amin, "Document Skew Detection and Correction Using Ellipse Shape" *International Journal of Imaging & Robotics*, Volume 19, Issue Number 4, 2019.

- [40] Marian Wagdy, I. Faye, D. Rohaya, "Border Noise Removal from the Document Image Using X-Y Cut and Filtering Technique Based on Morphological Operation", *International Journal of Imaging and Robotics (IJIR)*, Vol. 15, No. 3, 2015.
- [41] Paniphi, Narayan, and Smita Tripathy. "Image Registration using Polynomial Affine Transformation." *Defence Science Journal* 52.3 (2002): 253.
- [42] Qin, J., Xiao, C., Hu, S., Zhang, J., Wang, W., Ishikawa, Y., Tsuda, K. and Sadakane, K. "Efficient query autocompletion with edit distance-based error tolerance." *The VLDB Journal* 1-25, 2019.
- [43] Fuad, Muhammad Marwan Muhammad, and Pierre-François Marteau. "The extended edit distance metric." 2008 International Workshop on Content-Based Multimedia Indexing. IEEE, 2008.
- [44] A.Masalovitch, and L.Mestetskiy, "Usage of continuous skeletal image representation for document images de-warping," In 2nd Int. Workshop on Camera-Based Document Analysis and Recognition, Curitiba, Brazil, Sep. 2007.
- [45] Fu, Bin "A model-based book dewarping method using text line detection." Proc. 2nd Int. Workshop on Camera Based Document Analysis and Recognition, Curitiba, Barazil. 2007.