AUTOMATIC RECTIFICATION OF PERSPECTIVE DISTORTION FROM A SINGLE IMAGE USING PLANE HOMOGRAPHY

Geetha Kiran A[1] and Murali S[2]

[1]Malnad College of Engineering, Hassan, Karnataka, India
[2]Maharaja Institute of Technology, Mysore, Karnataka, India

ABSTRACT

Perspective distortion occurs due to the perspective projection of 3D scene on a 2D surface. Correcting the distortion of a single image without losing any desired information is one of the challenging task in the field of Computer Vision. We consider the problem of estimating perspective distortion from a single still image of an unstructured environment and to make perspective correction which is both quantitatively accurate as well as visually pleasing. Corners are detected based on the orientation of the image. A method based on plane homography and transformation is used to make perspective correction. The algorithm infers frontier information directly from the images, without any reference objects or prior knowledge of the camera parameters. The frontiers are detected using geometric context based segmentation. The goal of this paper is to present a framework providing fully automatic and fast perspective correction.

KEYWORDS

Sobel Operator, Corner detection, Segmentation, Plane homography, Perspective Rectification.

1. INTRODUCTION

Images are perspective in nature. Today's world requires the perspective rectified image for applications like image based rendering and metrology from single view. The applications require parallel view image for photorealistic results. Much work has been done towards the rectification of perspective distortion depicted in document images and also based on multiple images. Recent focus has been on recovery based on single image, a problem that is more challenging than the multiple- view variety and has good potential applications in image based rendering, image mosaicing, machine vision, 3D Reconstruction. Perspective is the projection of actual vision. Many of the applications would need parallel projection because parallel projection facilitates easy extraction and application of the regularities. However they are not parallel in a perspective projection hence their identification and application is of much importance. This leads to rectification of perspective distortion.

The paper is organized as follows. In section 2 a review on the related works is highlighted. In section 3 the overall methodology is viewed through a flowchart. In section 4 the preprocessing steps which further leads to image segmentation are addressed. In section 5 the methodology used to estimate the corners is discussed. In section 6 an overview of the perspective estimation is addressed. In section 7 the details of the perspective rectification is given. Experimental results in section 8 with conclusion in section 9 are to follow.

2. LITERATURE SURVEY

The literature directly relevant to the proposed research, are highlighted here. Researchers have proposed different methods to rectify perspective distortion from a single image. Authors have used perspective cues, typically vanishing points, to predict the depth information and also cues from hough transform[1]. Murali S *et al*.[2] have proposed a method based on perspective transformation and plane homography to rectify the perspective distortion in an image to actual scale with known camera parameters. Jian Liang *et al.* [9] estimates shape from texture flow information obtained directly from the image without requiring additional metric data. Shijian *et al.* [3] extracts character stroke boundaries and tip points based on fuzzy sets and morphological operators. The method needs neither high contrast document boundary nor paragraph formatting information. Yin Fang *et al.* in [4] extracts the endpoints using morphology operators, text baseline is obtained based on the least square method and finally RANSAC method helps in fitting the line corresponding to the vanishing point. Richard Hartley [5] uses methods of projective geometry to determine a pair of 2D projective transformations to be applied to the two images in order to match the epipolar lines. Karfogiannis *et al.* [6] demonstrates the utilization of a fundamental property of distorted coplanar hexagonal lattices in the image rectification framework. Jagannathan and Jawahar [19] have used clues like document boundaries, page layout information, organization of text and graphics components, a priori knowledge of the script or selected symbols for removing the perspective effect and computing the frontal view needed for a typical document image analysis.

3. OVERALL METHODOLOGY

This paper presents a method for automatic perspective correction on distorted images. The flow of the overall methodology used to rectify the distorted image is given in Figure 1.

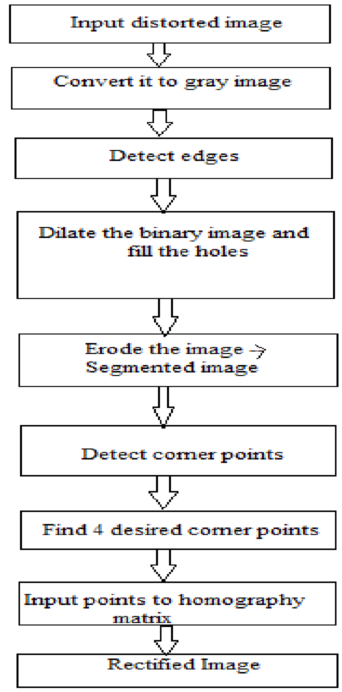


Figure 1. Flow of the overall methodology

4. IMAGE SEGMENTATION

Preprocessing is performed to segment the portion to be corrected. The desired portion is segmented and further used in corner estimation. The steps involved are as follows:

a) The given color image is converted to gray image
b) Apply sobel operator to detect edges.
c) Segment the image by applying the dilation and erosion methods.

The above steps have been explained in subsequent sections.

4.1 Defining Binary Gradient Mask

Image gradients are used to extract information from image. The gradient of an image measures how it is changing. It provides two pieces of information. The magnitude of the gradient tells how quickly the image is changing, while the direction of the gradient tells us the direction in which the image is changing most rapidly. Sobel operator has been used for edge detection. The sobel method finds edges using the sobel approximation to the derivative. It returns edges at those points where the gradient of I is maximum. The operator uses two 3×3 kernels which are convolved with the original image to calculate approximations of the derivatives - one for horizontal changes, and one for vertical. If **A** is the source image, $\mathbf{G}_x$ and $\mathbf{G}_y$ are two images which at each point contain the horizontal and vertical derivative approximations, the computations are as follows:

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * A \quad \text{and} \quad G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * A \tag{1}$$

where * denotes the 2-dimensional convolution operation. They compute the gradient with smoothing. At each point in the image, the resulting gradient approximations can be combined to give the gradient magnitude, using:

$$G = \sqrt{Gx^2 + Gy^2} \tag{2}$$

The gradient's direction is obtained by

$$\Theta = \mathrm{atan2}(G_y, G_x) \tag{3}$$

where Θ is 0 for a vertical edge which is darker on the right side.

4.2. Retrieving the Segmented Image

Dilation and erosion techniques using the structuring elements are used to segment the image. Assuming *E* to be a Euclidean space or an integer grid, *A* a binary image in *E*, and *B* a structuring element.

The dilation of *A* by *B* is given by:

$$A \oplus B = \bigcup_{b \in B} A_b \tag{4}$$

The erosion of *A* by *B* is given by:

$$A \ominus B = \bigcap_{b \in B} A_{-b} \tag{5}$$

Structuring element is used for probing and expanding the shapes contained in the input image.The output of the preprocessing steps is given in Figure 2.

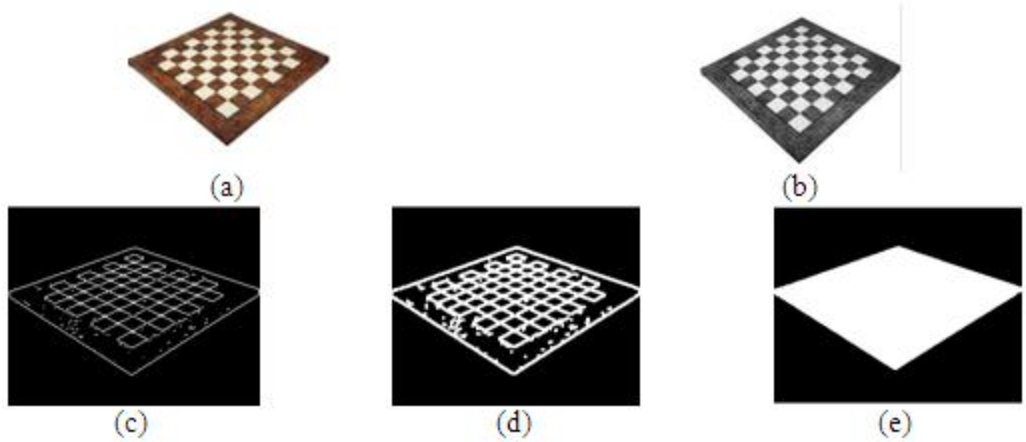


Figure 2. (a) Original image (b) Gray image (c) Edge detection (d) Dilated image (e) Segmented image

The segmented image is the input for corner estimation. In the next section there is detailed description of estimating corners.

5. CORNER ESTIMATION

Harris corner detector is a mathematical operator that finds a variation in the gradient of a image. It is rotation, scale and illumination variation independent. It sweeps a window w(x,y) with displacement u in the x direction and v in the y direction and calculates the variation of intensity.

$$E(u,v) = \sum_{x,y} w(x,y)[I(x+u,y+v) - I(x,y)]^2 \tag{6}$$

where w(x,y) is the window function which is equal to 1 in window and 0 outside , I(x,y) is the intensity at (x,y), I(x+u,y+v) is the intensity at the moved window (x+u,y+v). The large variation in intensity depicts corners. Hence equation (6) has to be maximized, specifically the term:

$$\sum_{x,y}[I(x+u,y+v) - I(x,y)]^2 \tag{7}$$

Let's denote

$$M = \sum_{x,y} w(x,y)\begin{bmatrix} Ix^2 & IxIy \\ IxIy & Iy^2 \end{bmatrix} \tag{8}$$

Now

$$E(u,v) = [u \quad v]\, M \begin{bmatrix} u \\ v \end{bmatrix} \tag{9}$$

A score is calculated for each window, to determine if it can possibly contain a corner:

$$R = \det(M)-k(\mathrm{trace}(M))^2 \tag{10}$$

where $\det(M) = \lambda_1 \lambda_2$; $\mathrm{trace}(M) = \lambda_1 + \lambda_2$; λ_1 and λ_2 are the eigen values of M. R depends on eigenvalues of M. A window with a score R greater than the threshold is selected and points of local maxima of R are considered as a "corner" and are shown in Figure 3(a). Further it is reduced to four using geometric based calculations as desired by the plane homography method.

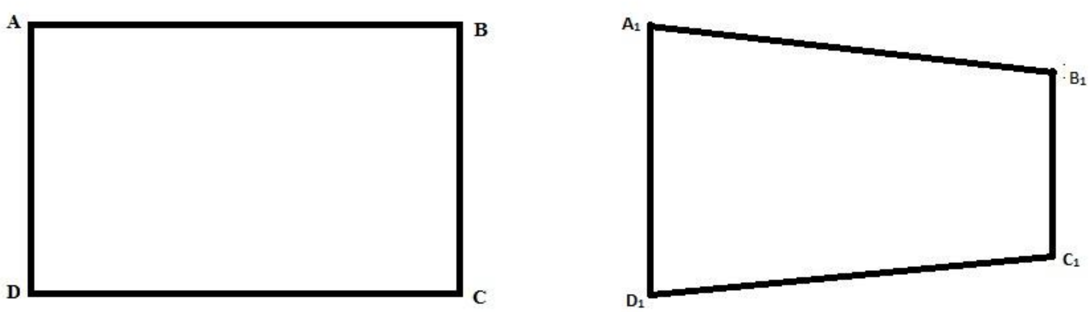
The four points of interest is obtained by selecting the minimum and maximum values in the corner points array as shown in Figure 3(b).

(a) (b)

Figure 3. (a) All corners are detected (b) Desired corners obtained

6. PERSPECTIVE ESTIMATION

Images captured in pinhole cameras are perspective in nature. In a perspective image, objects of similar size appear bigger when it is closer to view point and smaller when it is farther. Hence the number of pixels to represent farther objects in the image requires lesser pixels as compared to the number of pixels required to represent the closer objects in the image. Example: when objects are inclined to the viewer, it goes farther away from the viewpoint. Hence, the lesser the pixels means less information for image processing. Perspective images lead to many ambiguous results when we tend to measure the size of the objects in the picture. The rectangular plane ABCD in the Figure 4(a) is seen as $A_1B_1C_1D_1$ in a perspectively distorted case. The perspective view of the rectangle is depicted as in the Figure 4(b). This is how the images of the objects are formed in the image acquisition process.



(a) (b)

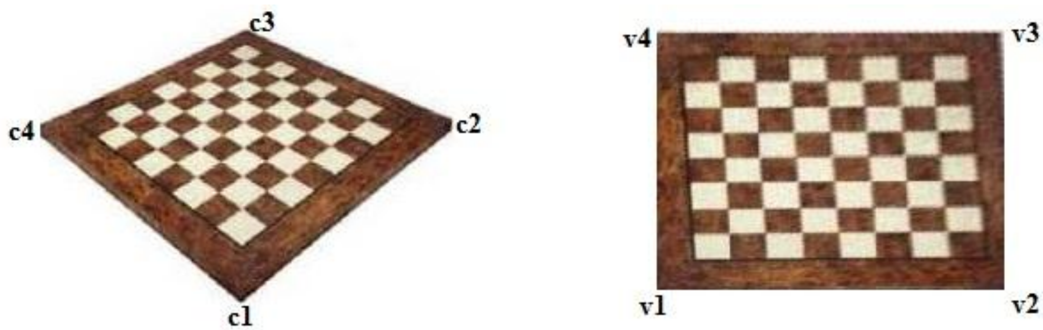
Figure 4. (a) Rectangular Plane (b) Perspective distortion

7. PERSPECTIVE RECTIFICATION

Perspective transformation turns a perspective projection into a parallel projection. Perspective transformation is used for replacing a view volume (prism shaped volume) into a rectangular shape. The parallel projection of transformed primitives is the same as the untransformed image under a perspective projection.

To build a parallel projection of the image from the perspectively distorted image, the plane homography[14] is used. The homography can be computed by knowing the relative positions of the four points on the perspectively distorted image and the positions of the transformed image to be constructed. The four corner points are required to implement the perspective transformation. The corner method detects all the available corner points. The score calculated for each pixel in the corner detector is based on the two eigen values of a matrix. The expression to calculate it, is not arbitrary, but based on observations of how the expression varies with different eigen values. Among the available corner points, the four desired corner points have to be selected. The plane homography is interested with only four points i.e., the intersection point of the edges. It is obtained by taking the minimum and maximum values in the corner points array.

The corner points are further sorted in clockwise direction. The four corner points C_1(X1,Y1), C_2(X2,Y2), C_3(X3,Y3) and C_4(X4,Y4) of the rectangular object (view volume) of the perspective image is used in the plane homography method as in Figure 5(a). The rectified image is obtained having the corners V_1(x1,y1), V_2(x2,y2), V_3(x3,y3) and V_4(x4,y4) where (x1,y1)=(0,0), (x2,y2)=(L,1), (x3,y3)=(L,B) and (x4,y4)=(1,B) as in Figure 5(b).



(a) (b)

Figure 5.

Plane homography is widely used for mapping the perspective boundary to a rectangle boundary which is a necessary step towards perspective transformation. A 2D homography is defined as a 3X3 homogeneous matrix that maps any point p(x,y) on plane π to its corresponding point p′ (x′,y′) on π’ as:

$$p' = H \bullet p \tag{11}$$

The homography is described by a homographic transformatiom when p and p′ are converted to homogeneous coordinates.

$$\begin{pmatrix} wx' \\ wy' \\ 1 \end{pmatrix} = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{12}$$

Compute p′ = H • p to apply a homography H. The homography can be applied in rectifying the perspective distortion of an image by finding the homography H, given p and p′ pairs. The derivation can be attained by first finding the correspondence for a single point and then extending it to four points.

Consider one point correspondence, p(x,y) → p′ (x′,y′).

Then equation (11) becomes

$$\begin{pmatrix} wx' \\ wy' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{13}$$

Solving x' and y': $x' = \frac{ax+by+c}{gx+hy+i}$ and $y' = \frac{dx+ey+f}{gx+hy+i}$. (14)

The two linear equations with respect to unknown coefficients of matrix H is as below:

$$\begin{gathered} ax+by+c-gxx'-hyy'-ix'=0 \\ dx+ey+f-gxy'-hyy'-iy'=0 \end{gathered} \tag{15}$$

By extending the same to four point correspondence,

$$p_i(x_i,y_i) \rightarrow p'_i(x'_i,y'_i) \tag{16}$$

It can be generalized as

$$\begin{gathered} ax_i+by_i+c-gx_ix'_i-hy_iy'_i-ix'_i=0 \\ dx_i+ey_i+f-gx_iy'_i-hy_iy'_i-iy'_i=0 \end{gathered} \tag{17}$$

The above equations can be written in matrix format as **A_i . h=0** for i=1,2,3,4 where h=[a b c d e f g h i] ^{T} is a vector of unknown coefficients in H and A_i is a 2x9 matrix based on known coordinates x_i,y_i,x'_i,y'_i given as

$$\begin{bmatrix} xi & yi & 1 & 0 & 0 & 0 & -x'ixi & -x'iyi & -x'i \\ 0 & 0 & 0 & xi & yi & 1 & -y'ixi & -y'iyi & -y'i \end{bmatrix} \tag{18}$$

Finally equation (11) can be written as

$$P' = H \bullet P \rightarrow A_i \bullet h = 0 \text{ for } i = 1,2,3,4. \tag{19}$$

where A_i is 2x9 matrix, h is a 9x1 matrix and the result is a 2x1 matrix. If all the four matrix equations are put into one equation, it results with the following

$$\begin{pmatrix} A1 \\ A2 \\ A3 \\ A4 \end{pmatrix} . h = 0 \tag{20}$$

where the order of the matrix A_i is (4x2)x9 = 8x9, h is a 9x1 matrix and the result is a 8x1 matrix. The resultant expression can be written as

$$A_i \bullet h = 0 \tag{21}$$

It implies that there are 8 linear equations and 9 unknows. Add constraint ‖h‖=1 which will simplify to A•h=0 subject to ‖h‖=1. The solution is obtained using Singular Value Decomposition for the 4 point correspondence between two planes.

The transform is given by the solution as explained above. The elements $A_{1\text{-}8}$ can be given as

$$\begin{bmatrix} X1 & Y1 & 1 & 0 & 0 & 0 & -X1*x1 & -Y1*x1 \\ 0 & 0 & 0 & X1 & Y1 & 1 & -X1*y1 & -Y1*y1 \\ X2 & y2 & 1 & 0 & 0 & 0 & -X2*x2 & -Y2*x2 \\ 0 & 0 & 0 & X2 & Y2 & 1 & -X2*Y2 & -Y2*y2 \\ X3 & Y3 & 1 & 0 & 0 & 0 & -X3*x3 & -Y3*x3 \\ 0 & 0 & 0 & X3 & Y3 & 1 & -X3*y3 & -Y3*y3 \\ X4 & Y4 & 1 & 0 & 0 & 0 & -X4*x4 & -Y4*x4 \\ 0 & 0 & 0 & X4 & Y4 & 1 & -X4*y4 & -Y4*y4 \end{bmatrix}$$

and the vectors are given as $h_{1\text{-}8} = (a11, a12, a13, a21, a22, a23, a31, a32)^T$, $A_9 = (x1, y1, x2, y2, x3, y3, x4, y4)^T$.

When the transformation is applied on perspective images, the perspective distortion is rectified as highlighted in Figure 5.

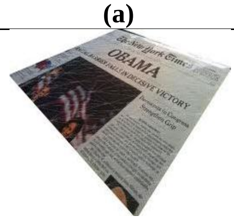
The transformation can be summarized as follows:

a) Input the segmented image and the four corner points.
b) The projective transformation is applied on the coordinates of the input image.
c) The resultant image contains the transformed view volume of the rectified image.

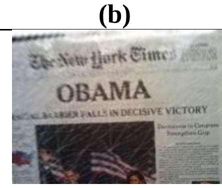
The distorted and the rectified images are given in the Figure 6.

8. EXPERIMENTAL RESULTS

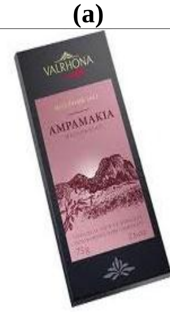
(a)



(b)



(a)



(b)

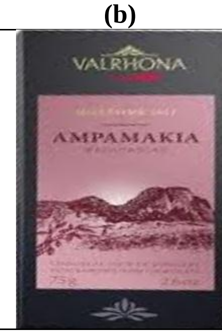


Figure 6. (a) Distorted Images (b) Rectified Images

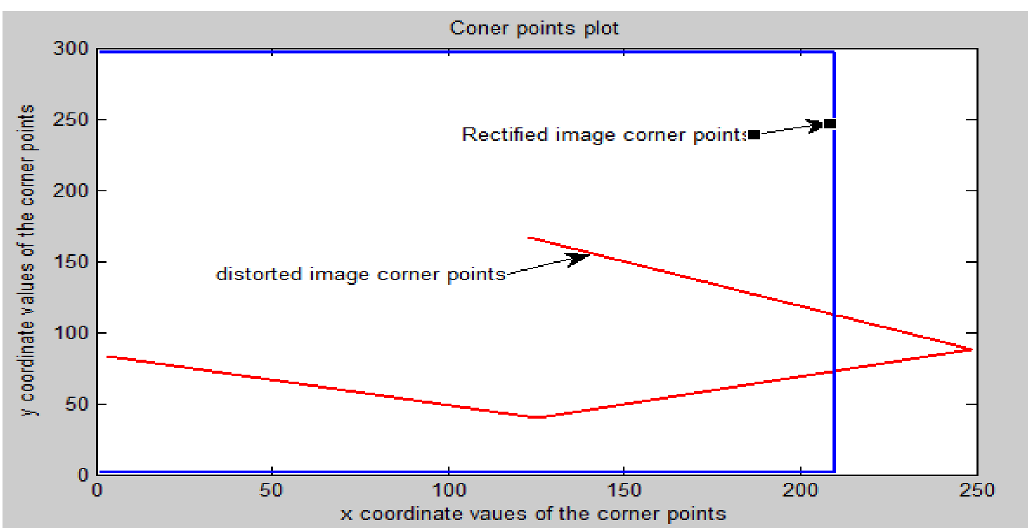


Figure 7.

9. CONCLUSION

An algorithm for rectifying perspective distortion automatically has been presented. Compared to previous approaches, our algorithm rectifies the perspective distortion without any reference objects or prior information about the camera parameters. The rectification may not be to the scale. There is a possibility to rectify to the scale using geometrical manipulation. The scale factor has to be derived before the manipulation using camera calibration technique.

REFERENCES

[1] S.Murali, N.Avinash & Mohan S (2006) "Rectification of perspective distortion using plane homography and clues from Hough transform", Indian conference on Computer Vision, Graphics and Image Processing.
[2] Mohan S, Avinash N & Murali.S (2008) "Rectification of Perspective distortion using camera parameters - A Perspective Geometry Based Approach", ICGST International Journal on Graphics, Vision and Image Processing, GVIP, Vol 8, pp.1-7.
[3] Shijian Lu, Ben M. Chen & C. C. Ko(2005) "Perspective rectification of document images using fuzzy set and morphological operations", Journal on Image and Vision Computing, Vol 23, pp. 541-553.
[4] Yin Fang , Chen Deyun & Wu Rui (2011) " A distortion correction approach on natural scene text image, Strategic Technology ", IFOST, 6th International forum, Vol 2, pp. 1058-1061.
[5] Richard I. Hartley (1999) "Theory and Practice of Projective Rectification", International Journal of Computer Vision, Vol 35, No.2, pp. 115-127.
[6] Koufogiannis, E.T. Sgouros & N.P. Sangriotis, M.S (2012) "Perspective Rectification of Integral Images Produced Using Hexagonal Lens Arrays" , Eighth International Conference on IIH-MSP, pp.75-78.
[7] R.O.Duda, P.E Hart & D.G. Stork (2000), Pattern Classification, 2nd ed. New York,Wiley – Inter science.
[8] Gonzalez R.C. & Woods R.E (2002) Digital Image Processing with Matlab, Prentice Hall.
[9] Jian Liang, Daniel DeMenthon & David Doermann (2008) "Geometric Rectification of Camera-Captured Document Images", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 30, No.4, pp 591-605.
[10] Scharr & Hanno (2000) "Optimal Operators in Digital Image Processing" , Dissertation (in German).
[11] C. Harris & M.J. Stephens (1988) "A combined corner and edge detector", In Alvey Vision Conference, Vol 15, pp 147–152
[12] H. Moravec (1980) "Obstacle avoidance and navigation in the real world by a seeing robot rover", Technical Report CMU-RI-TR-3, Carnegie-Mellon University, Robotics Institute.
[13] C. Schmid, R. Mohr & C. Bauckhage (2000) "Evaluation of interest point detectors", International Journal of Computer Vision, Vol.37, No.2, pp151–172.
[14] Xiang Wang, Reinhard Klette & Bodo Rosenhahn (2006) "Geometric and Photometric Correction of Projected Rectangular Pictures", Image and Vision Computing, New Zealand.
[15] P.F. Felzenszwalb & D.P. Huttenlocher (2004) "Efficient graph-based image segmentation", International Journal of Computer Vision, Vol.59, No. 2, pp. 167-181.
[16] Maurizio Pilu (2001) "Deskewing perspectively distorted documents: An approach based on perceptual organization", HP laboratories, Bristol UK, No. 100.
[17] Barnard S T(1983) "Interpreting perspective images", Artificial Intelligence, Vol 21, pp.435-462.
[18] Intel OpenCV Computer vision Library(C++), http://www.intel.com/research/mrl/research/opencv.

[19] L.Jagannathan & C V Jawahar (2005) “Perspective correction methods for Camera based document Analysis”, Proceedings of the First International Workshop on Camera-based Document Analysis and Recognition (CBDAR), pp 148-154.
[20] Francesco Isgro & Emanuele Trucco (1999) “On Robust Rectification for Uncalibrated Images”, Proceedings of the 10th International Conference on Image Analysis and Processing, pp 297-302.
[21] Chris harris & Mike stephens(1988) “A combined corner and edge detector”, Proceedings of the Fourth Alvey Vision Conference, Manchester, Vol 15, pp 147-151.
[22] Donald Hearn & Baker, Computer Graphics, PHI.
[23] F Isgro E Trucco (1999) “On Projective Rectification”, Seventh International Conference on Image Processing and its Applications,Vol 1, pp.42-46.
[24] C. V. Jawahar (2006) “Homography Estimation from Planar Contours”, Proceedings of the Third International Symposium on 3D Data Processing, Visualization and Transmission, pp.877-884.