2ECDE62 Computer Vision Report on

"Document Scanning using Camera Calibration"

by

Aum Dhabalia (21BEC027)

Soumik Mohpatra (20BEC123)

to

Bhupendra Fataniya

(Associate Professor)

Abstract – In this report, a document scanning application from camera calibration is performed. Various distorted images were taken as input and algorithms were applied to get required output. Algorithms like SIFT (Scale Invariant Feature Transform), camera calibration estimation from homography, Four-Point transform, edge detection, etc. are used for achieving the document scanning. Distortion in the image was removed using perspective transform since the distortion was tangential in nature. Camera calibration was done to determine calibration parameters and use it for document scanning purpose.

Keywords:- Camera Calibration, SIFT, Four-Point Transform, Distortion, Document Scanning, Edge Detection

I. Introduction

Accurate camera calibration is a prerequisite for many computer vision applications, including 3D reconstruction, object recognition, document scanning and augmented reality. Calibration involves determining the camera's intrinsic parameters such as focal length and lens distortion and extrinsic parameters such as position and orientation in the world. One of the applications of camera calibration is document scanning, which is available as mobile application software.

Document scanning simply captures the document image and then applies image processing and camera calibration matrix to remove any transformations within the image. The transformations include conversion of quadrilateral shape to rectangular shape.

II. Literature Survey

Document scanning using camera calibration matrices has become an enhanced approach in document digitization. Unlike traditional scanning software that relies solely on image enhancement algorithms, calibration techniques involve understanding and compensating for camera related distortions, lens parameters, and geometric transformations.

Key concepts in camera calibration includes the calibration matrix itself, lens distortion and perspective correction. Lens distortion refers to radial and tangential distortion. The distortion observed as curved or bent edges are caused by radial distortion and the tilted image distortion is caused by tangential distortion.

Key concepts of document scanning involve precise edge detection, perspective transformation and improved optical character recognition accuracy. Perspective transformation is done using calibration matrix to any image whether it is in top-down view, flat view or tilted at any angle.

III. Design & Methodology

The methodology used for achieving document scanned image is first calibrating camera using feature point detection using Scale Invariant Feature Transform (SIFT) algorithm and then using the calibration matrix to remove any distortion. After acquiring the undistorted image, edges are detected and bound to the shape of a quadrilateral and then, four-point transform algorithm is used to convert shape of the detected quadrilateral window to rectangular window.

Algorithms used for document scanning are SIFT, Four-Point Transform, calibration estimation from homography and estimating perspective transform matrix. There are six steps of the pseudo-algorithm for document scanning.

a. SIFT (Scale Invariant Feature Transform)

Here, camera is calibrated in real-time meaning by providing live video feed. So, initially using live video feed key-points and descriptors are extracted from current frame and previous frame. Using those key-points and descriptors, image coordinates are determined for current frame and previous frame. Image coordinates are then matched, and perfectly matched coordinates are used to determine homography matrix.

b. Calibration matrix estimation from homography

From the calculated homography, calibration matrix is determined. Since the equation of homography is $H = K[R\ t]$, it is assumed that rotation and translation of the camera is none and so calibration matrix (K) is calculated using singular value decomposition method.

It is also assumed that the camera is positioned at origin and there is skew not present. From this, the calibration matrix determined is almost equivalent to identity matrix and determinant of K is in the range of 0.9 to 1.1

c. Edge detection and contour detection

After camera is calibrated, the image of the document is capture and edges are detected using

Canny edge detection contour detection is done for outlining the shape of the document in the image.

d. Four-Point Transform

Once the contour is detected, the coordinates of the corners of the image are extracted as source points and destination points are considered as corners of the rectangular shape with height and width determined from the height and width of the document from detected the contour.

e. Estimation of perspective Transform

After the destination points are determined, the perspective transformation matrix is calculated from source points and destination points.

f. Applying perspective transform to change the shape of the document scanned image

After acquiring perspective transformation matrix, it is applied on the image and shape is transformed to rectangle of height and width from the original image.

Given Fig. 1 shows the flowchart of the approach for document scanning.

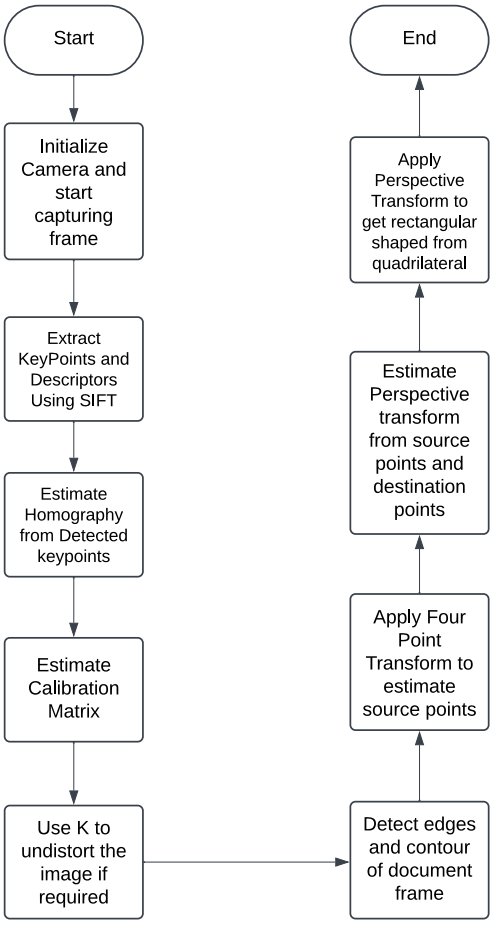

Fig. 1 Document Scanning Flowchart

IV. Outcomes

Three sample images were taken to check the outcome and the results obtained were almost perfect. All the obtained resultant images are rectangular in shape and the lens distortion was observed to be absent. Given figures are the outcomes of the algorithm of document scanner.

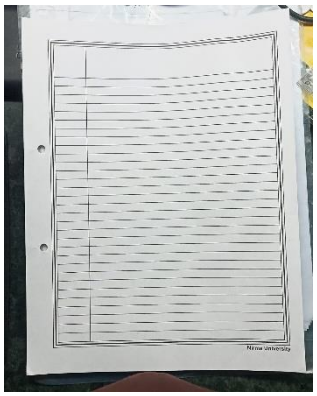

Figure 2a Sample Image 1

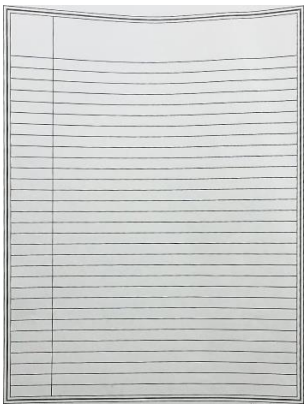

Figure 2b Scanned Document 1

Figure 3a Sample Image 2

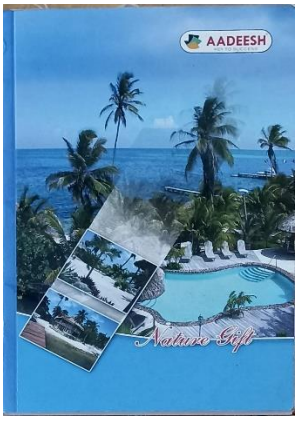

Figure 3b Scanned Document 2

Figure 4a Sample Image 3

Figure 4b Scanned Document 3

V. Conclusion

During the period of performing the project, several learning outcomes were achieved like learning scale invariant feature transform (SIFT) algorithm, four-point transform algorithm to extract corners points of the sample image and calibration matrix estimation using homography from SIFT. The outcomes observed were almost accurate as outcomes achieved from document scanning apps.

VI. References

[1] M. Covell and S. Baluja, "Finding Images and Line-Drawings in Document-Scanning Systems".

[2] R. Maini and H. Aggarwal, "Study and Comparison of Various Image Edge Detection Techniques".

[3] M. Dawood, C. Capelle, M. E. El Najjar, M. Khalil and D. Pomorski, "Harris, SIFT and SURF features comparison for vehicle localization based on virtual 3D model and camera".

[4] H. Jaspers, B. Schauerte and G. A. Fink, "SIFT-based Camera Localization using Reference Objects for Application in Multi-Camera Environments and Robotics".

[5] Z. Zhang and Q. Tang, "Camera Self-Calibration Based on Multiple View Images".