Journal of Engineering and Applied Sciences 10 (9-12): 277-281, 2015
ISSN: 1816-949X

Reconstruction of Scanned Book Pages

Abbas Nasrabadi
Department of Mechatronics Engineering, Hakim Sabzevari University, Sabzevar, Iran

Abstract: In order to retrieve knowledge from the old and voluminous text books its necessary to develop a method to reconstruct the scanned pages texts while defection of these old text pages is very usual which causes two major problems, the shadow of text which is caused by the scattering of ink over the wet paper and the deformation of old paper which naturally occurs as a result of temperature and humidity variations and causes the text to be seen curved and unrecognizable. It's the objective of this present work to represent a method resolving the two inconvenient. First, the border between the text and its shade should be determined utilizing a connected component analysis based noise reduction method. Second, the discovered borders are compared with the previously standardized letters edges and then the current text is identifiable. The method is experienced at several pages and results are displayed.

Key words: Distortions, shade, main septum, Niblack Method, noise reduction

INTRODUCTION

This current job is mainly stimulated by the two immediate work pressures at "national archives of Singapore and NUS digital library". One urgent project was about to store the large amount of old hand written documents and second to store the old thesis and books which were typed but deformed over time. While the hand writings had the problem of scattered ink shades and irretrievable letters the typed books are also difficult to be scanned properly because its not possible to put the pages on flat position due to the physical pressure caused by the glued edge of pages which conforms a solid column. This creates a shadow to be casted on the glued side of the page which is scanned and the deformation of letters over the curved paper surface. This situation is more or less similar to the condition occurs for hand written pages. Before attending to bid a new job on solving these issues we committed a survey on the previous jobs. Baird has presented a six parametered algorithm (Baird, 2007, 1999) but its not clarified about the contributed share of each parameter which are to be presumed at practice when we face the matter of extracting info from various kinds of deformations at a number of different pages. Kanungo *et al*. (1993) has suggested a model to remove the image distortions but this method also relies on some assumptions and parameters which may not be determinable at a real experience. So, at our work here we are not going to utilize any parameter adaptation to the current environment stages.

Here we have benefited from a modified version of Niblack Binary Image Processing Method to remove the shade and then created the connected components to remove the noise and rearrange the bent letters at the shaded area. These operations are executed through four main stages:

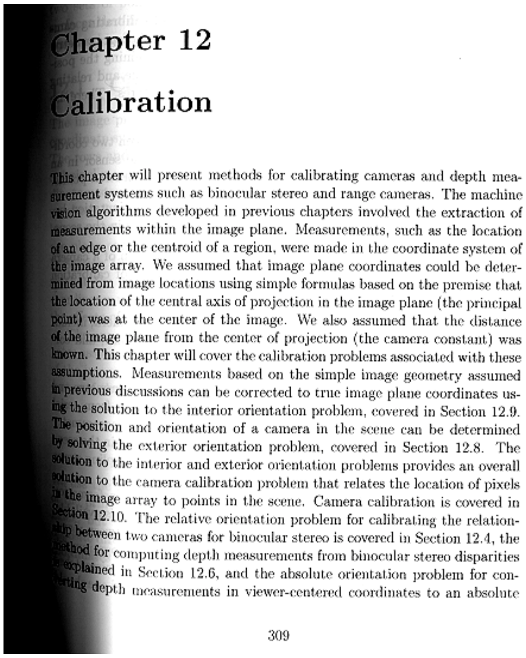
- Indication of shade and text borders
- Converting the image to a binary mode
- Reconstructing and analyzing the connected components
- Final regulation of the image with bent letters

These four stages are explained in next study.

DETERMINING THE SHADE BORDER

Figure 1 depicts an example of scanned gray image of a thick book's page assuming that each picture contains one page it can be observed that only the words at the shadowed hazy area are deformed and bent so we limit the work of letters reformation to the shade vicinity. At this level, we first create a projection profile from the image to determine the location of the projection pike that delegates the location of shade at the original image.

Next, the image texts should be read at horizontal order right from the location of profile's pike. Each read line should be assigned a break point 'B' which indicates the border between the shade and clear area. This B should be defined with this equation:



Chapter 12

Calibration

This chapter will present methods for calibrating cameras and depth measurement systems such as binocular stereo and range cameras. The machine vision algorithms developed in previous chapters involved the extraction of measurements within the image plane. Measurements, such as the location of an edge or the centroid of a region, were made in the coordinate system of the image array. We assumed that image plane coordinates could be determined from image locations using simple formulas based on the premise that the location of the central axis of projection in the image plane (the principal point) was at the center of the image. We also assumed that the distance of the image plane from the center of projection (the camera constant) was known. This chapter will cover the calibration problems associated with these assumptions. Measurements based on the simple image geometry assumed in previous discussions can be corrected to true image plane coordinates using the solution to the interior orientation problem, covered in Section 12.9. The position and orientation of a camera in the scene can be determined by solving the exterior orientation problem, covered in Section 12.8. The solution to the interior and exterior orientation problems provides an overall solution to the camera calibration problem that relates the location of pixels in the image array to points in the scene. Camera calibration is covered in Section 12.10. The relative orientation problem for calibrating the relationship between two cameras for binocular stereo is covered in Section 12.4, the method for computing depth measurements from binocular stereo disparities is explained in Section 12.6, and the absolute orientation problem for converting depth measurements in viewer-centered coordinates to an absolute

309

Fig. 1: A gray scaled image of scanned book page

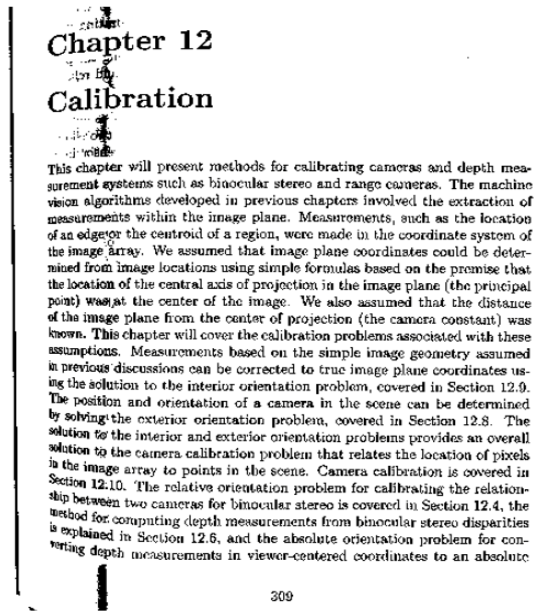
$$B = \{(\beta, i) \mid 0 \leq i < n, \beta = F(i, T)\} \quad (1)$$

Where:

- n = The quantity of pixels at the horizontal line at the image
- T = The predefined border
- F = The return function from the length of the first run from the pixel at ith line which has a larger intensity than T. We use $T = V_{max}/4$ at the current experiment where V_{max} is maximum intensity of the gray scale image here

CONVERSION TO THE BINARY IMAGE

To erase the shade a known version of Niblack algorithm is utilized. The Niblack Method acts with respect to the different thresholds at the image which are calculated from the local average 'm' and the standard deviation 's' which is calculated at a 15×15 neighborhood. So, the threshold is obtained for each pixel as p(x, y) from the equation T(x, y) = m (x, y)ks(x, y) at the center point of the neighborhood. 'k' is a user defined parameter which should be negative. However this rout may not be the optimal choice here because of two matters, one is that the method is sensitive to the amount of K while its not possible to maintain a single value for this parameter which can cover all of the subjected images, second problem is that the high rate of salt-pepper noise is imposed to the shaded area. At our improved method



Chapter 12

Calibration

This chapter will present methods for calibrating cameras and depth measurement systems such as binocular stereo and range cameras. The machine vision algorithms developed in previous chapters involved the extraction of measurements within the image plane. Measurements, such as the location of an edge or the centroid of a region, were made in the coordinate system of the image array. We assumed that image plane coordinates could be determined from image locations using simple formulas based on the premise that the location of the central axis of projection in the image plane (the principal point) was at the center of the image. We also assumed that the distance of the image plane from the center of projection (the camera constant) was known. This chapter will cover the calibration problems associated with these assumptions. Measurements based on the simple image geometry assumed in previous discussions can be corrected to true image plane coordinates using the solution to the interior orientation problem, covered in Section 12.9. The position and orientation of a camera in the scene can be determined by solving the exterior orientation problem, covered in Section 12.8. The solution to the interior and exterior orientation problems provides an overall solution to the camera calibration problem that relates the location of pixels in the image array to points in the scene. Camera calibration is covered in Section 12.10. The relative orientation problem for calibrating the relationship between two cameras for binocular stereo is covered in Section 12.4, the method for computing depth measurements from binocular stereo disparities is explained in Section 12.6, and the absolute orientation problem for converting depth measurements in viewer-centered coordinates to an absolute

309

Fig. 2: The result of current conversion algorithm

each s(x, y) is normalized with the dynamic range of derivations 'R'. Also the local average m(x, y) is increased several times instead of increasing the standard derivations. This leads to a lesser amount of pepper noise dots added to the image in response to the decreased derivations. In addition also the amount of sensitivity toward the parameter k is deduced too. Equation 2 shows the improved equation:

$$T(x,y) = m(x,y) \times \left[1 + k \times \left(1 - \frac{s(x,y)}{R}\right)\right] \quad (2)$$

where, s, m are the same with the Niblack Method. We arranged R = 100 and K = 0.1 for the gray scaled objectives. Figure 2 displays the result of Fig. 1's conversion to binary with the reformed method.

RECONSTRUCTION AND ANALYSIS OF THE CONNECTED COMPONENTS

After conversion to the binary mode we may face a sizable amount of noise dots like as seen at Fig. 2. So, we have to distinguish the connected components based on a 8 membered neighborhood. This is also required in order to grab the letter surrounding boxes which are needed to correction of bent letters stage.

Noise filtering and removal: Two filtering strategies are concurrently executed, shape filtering and size filtering. Shape filtering removes elongated and undisciplined areas which are most likely not letters. Size filtering then removes large spots which are bigger than normal letters. The mathematic logic behind the shape filtering is exposed at Eq. 3:

$$F(c)=\begin{cases}\text{true} & \text{if } \sqrt{(x_1-x_0)^2+(y_1-y_0)^2}\leq \sqrt{j\times \text{size\_of}(c)},\ \forall (x_1,y_1)\in c\\ \text{false} & \text{otherwise}\end{cases} \quad (3)$$

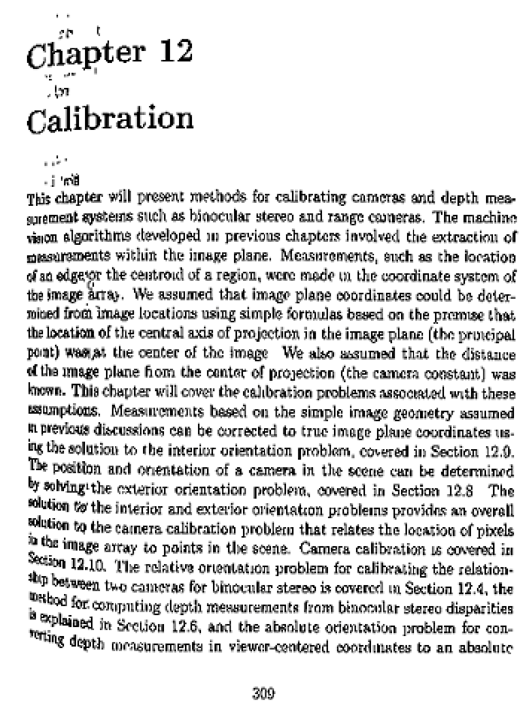
where, c implicates the subjected connected component to be checked for the removal (x_0, y_0) is its centroid coordinates and (x_1, y_1) the coordinates for a optional point at the component. "Size_of" is the return function which determines the number of black pixels at the current c component and j is a default factor which is set at 2.5 here. The result of this filtering approach is shown at Fig. 3.

Categorization of connected components: Its possible to have more than one black areas concluded as a single component like in letter i or j. At such cases we may acquire undesirable results when we try to reform the bent letters to perpendicular shape. So, prior to that we should regroup the excavated connected components based on the closest neighborhood. This operation should be computed through the following levels:

- Level 1: assuming a vacant letter box and assign it to the first unlabeled connected component, we call it as C1
- Level 2: as for any P2 pixel at C2 which is closer to any P1 pixel at C1 than the threshold distance of D is considered as the second component with the same label at the current letter box

If there had been no candidates to be comprised at the first letter box then the first level is imitated for the next component as Ci (Sauvola and Pietikainen, 2000). The result of this approach is depicted at Fig. 4 as its anticipated a suitable threshold distance D will bring separate boxes assigned for each word.

Word based categorization: In addition to the word segmentation results, we also require to confirm the position of the word over the text in order to extract interpretable sentences. At this job, we execute the improved box-hands method (Strouthopoulos *et al*., 1997) to participate lines. At the original basic box-hands method its assumed that the lines are put at nearly perfect horizontal sort which is not a case of practice here. In response to the current situation a chain expansion algorithm is applied on the results of study categorization of connected components. The function behaves as following, the vertical position of the box one is compared to the surrounding nearby boxes at four corners and two opposite pairs which repeat the same variation of height

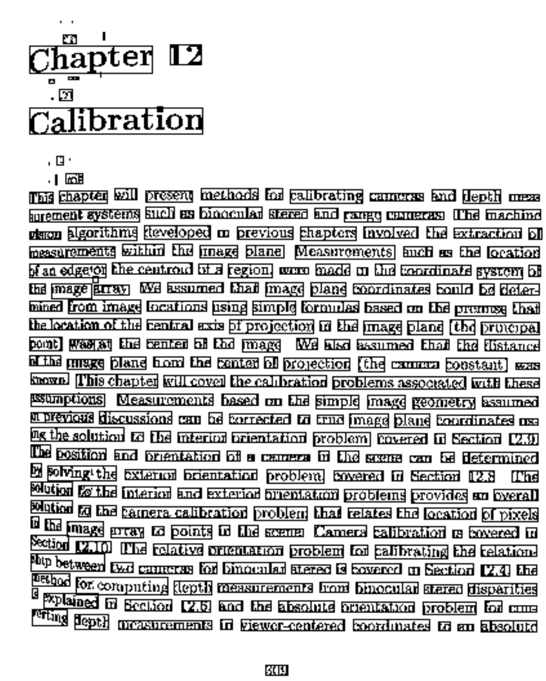

Chapter 12

Calibration

This chapter will present methods for calibrating cameras and depth measurement systems such as binocular stereo and range cameras. The machine vision algorithms developed in previous chapters involved the extraction of measurements within the image plane. Measurements, such as the location of an edge or the centroid of a region, were made in the coordinate system of the image array. We assumed that image plane coordinates could be determined from image locations using simple formulas based on the premise that the location of the central axis of projection in the image plane (the principal point) was at the center of the image. We also assumed that the distance of the image plane from the center of projection (the camera constant) was known. This chapter will cover the calibration problems associated with these assumptions. Measurements based on the simple image geometry assumed in previous discussions can be corrected to true image plane coordinates using the solution to the interior orientation problem, covered in Section 12.9. The position and orientation of a camera in the scene can be determined by solving the exterior orientation problem, covered in Section 12.8. The solution to the interior and exterior orientation problems provides an overall solution to the camera calibration problem that relates the location of pixels in the image array to points in the scene. Camera calibration is covered in Section 12.10. The relative orientation problem for calibrating the relationship between two cameras for binocular stereo is covered in Section 12.4, the method for computing depth measurements from binocular stereo disparities is explained in Section 12.6, and the absolute orientation problem for converting depth measurements in viewer-centered coordinates to an absolute

309

Fig. 3: Enhanced version of Fig. 1 after filtering


Chapter 12

Calibration

This chapter will present methods for calibrating cameras and depth measurement systems such as binocular stereo and range cameras. The machine vision algorithms developed in previous chapters involved the extraction of measurements within the image plane. Measurements, such as the location of an edge or the centroid of a region, were made in the coordinate system of the image array. We assumed that image plane coordinates could be determined from image locations using simple formulas based on the premise that the location of the central axis of projection in the image plane (the principal point) was at the center of the image. We also assumed that the distance of the image plane from the center of projection (the camera constant) was known. This chapter will cover the calibration problems associated with these assumptions. Measurements based on the simple image geometry assumed in previous discussions can be corrected to true image plane coordinates using the solution to the interior orientation problem, covered in Section 12.9. The position and orientation of a camera in the scene can be determined by solving the exterior orientation problem, covered in Section 12.8. The solution to the interior and exterior orientation problems provides an overall solution to the camera calibration problem that relates the location of pixels in the image array to points in the scene. Camera calibration is covered in Section 12.10. The relative orientation problem for calibrating the relationship between two cameras for binocular stereo is covered in Section 12.4, the method for computing depth measurements from binocular stereo disparities is explained in Section 12.6, and the absolute orientation problem for converting depth measurements in viewer-centered coordinates to an absolute

309

Fig. 4: Each word is assigned to an individual box

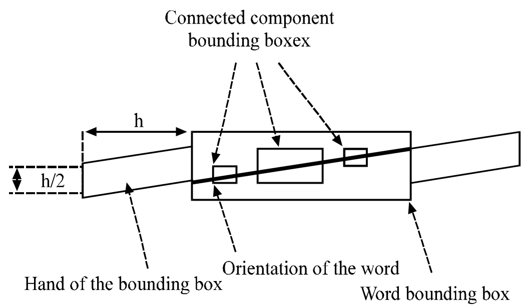


Fig. 5: Chain connected boxes

are presumed as the neighboring word box. Same results can also be obtained if we can find three dots which are positioned at a directed line at the central word box. The direction of the reference line is determined from the direction of central word which is derived from the liner regression (Jain *et al*., 1995). This strategy is illustrated at Fig. 5.

The linear regression is committed through the perception of following equations. The coordinates of each connected component centroid (x_i, y_i) is provided and the line equation of $y = mx+c$ is assembled comparing the two adjacent centroids. Finally, m and c are calculated with the below cited equations:

$$m = \frac{n\sum x_i y_i - \left(\sum x_i\right)\left(\sum y_i\right)}{n\sum x_i^2 - \left(\sum x_i\right)^2} \tag{4}$$

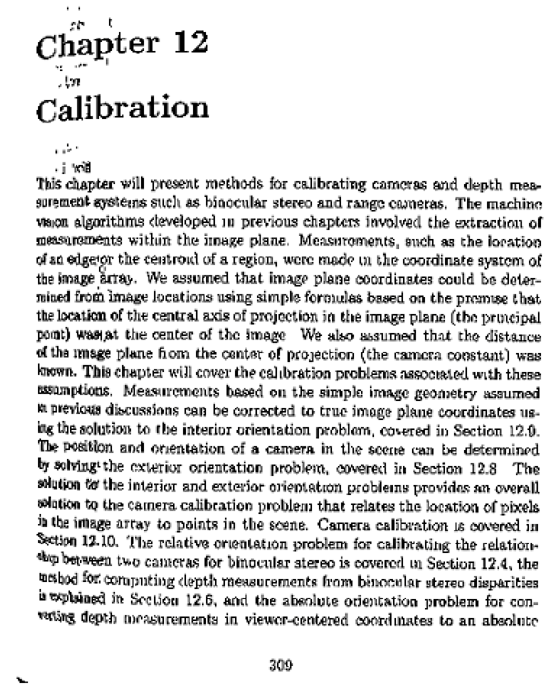
$$c = \frac{\sum y_i - m\sum x_i}{n} \tag{5}$$

Hence, every single text line can be detected by categorization of all the word which are in contact hands and boxes.

In comparison with we are not considering the shared piece between the two points of word's direction and box edges as enrolled. Also instead of Hough transform we utilized linear regression which induced much less computation efforts. This simplifications has soared the effectiveness of the algorithm without reducing the output image quality.

REGULATION OF THE BENT WORDS

Rearrangement of leaned words stage is consisted of two sub levels of regulating the location and sorting the direction. Consider a previously detected text line L extracted at 3-3. We use Hough transform at the center of all connected components belonged to L and enclose them to a straight line S. We simply relocate all the words


Chapter 12

Calibration

This chapter will present methods for calibrating cameras and depth measurement systems such as binocular stereo and range cameras. The machine vision algorithms developed in previous chapters involved the extraction of measurements within the image plane. Measurements, such as the location of an edge or the centroid of a region, were made in the coordinate system of the image array. We assumed that image plane coordinates could be determined from image locations using simple formulas based on the premise that the location of the central axis of projection in the image plane (the principal point) was at the center of the image. We also assumed that the distance of the image plane from the center of projection (the camera constant) was known. This chapter will cover the calibration problems associated with these assumptions. Measurements based on the simple image geometry assumed in previous discussions can be corrected to true image plane coordinates using the solution to the interior orientation problem, covered in Section 12.9. The position and orientation of a camera in the scene can be determined by solving the exterior orientation problem, covered in Section 12.8. The solution to the interior and exterior orientation problems provides an overall solution to the camera calibration problem that relates the location of pixels in the image array to points in the scene. Camera calibration is covered in Section 12.10. The relative orientation problem for calibrating the relationship between two cameras for binocular stereo is covered in Section 12.4, the method for computing depth measurements from binocular stereo disparities is explained in Section 12.6, and the absolute orientation problem for converting depth measurements in viewer-centered coordinates to an absolute

309

Fig. 6: The final result of enhancing the Fig. 1

of L in a manner that their centroids accord to the this most likely imposes on the words which are placed at the shade border (stage 2).

Now it's time to reform the curved letters. While a letter is leaned to a side its main diameter is not perpendicular to the straight horizontal line S and may has conformed an angle A with it. Now its only necessary to rotate the letter about its centroid with the quantity of A degrees. The final crop of this work is displayed at Fig. 6 which in respect to Fig. 3 or 1 is highly evolved and improved.

CONCLUSION

At this study, we first presented the problem of scanned image corruption and then suggested a correction-enhancement method to solve it. The algorithm initially utilized Niblack Thresholding Method to convert the image into the binary mode. Then, we assumed the removal of some specific connected components in order to reduce the noise consequences. Further analysis executed over the connected components to construct letter and words boxes and these were on the next turn modified and used to measure and extract the text lines. After these Regulating the line texts implies the rearrangement of each letter's location and direction.

One of the appealing aspects of further research is to work over the repositioning of connected components utilizing the optical flow methods which may even repair the highly damaged letters. First a method to determine the level of injury should be developed with respect to the gray level data of each letter or connected component this data should then be used to estimate the direction and amount of leaning of the surface paper so an algorithm can be composed to fix the letters and words. Further work can be accomplished over the data extraction from more complicated pictures and documents containing even graphical data (Zhang *et al*., 2004).

REFERENCES

Baird, H.S., 1999. Document image quality: Making fine discriminations. Proceedings of the Fifth International Conference on Document Analysis and Recognition, 1999. September 20-22, 1999, IEEE, Bangalore, pp: 459-462.

Baird, H.S., 2007. The State of the Art of Document Image Degradation Modelling. In: Digital Document Processing. Chaudhuri, B.B. (Ed.). Springer Berlin, London, England, pp: 261-279.

Jain, R., R. Kasturi and B.G. Schunck, 1995. Machine Vision. 1st Edn., McGraw-Hill, USA., ISBN: 0070320187, pp: 549.

Kanungo, T., R.M. Haralick and I. Phillips, 1993. Global and local document degradation models. Proceedings of the Second International Conference on Document Analysis and Recognition, October 20-22, 1993, IEEE, Tsukuba Science City, pp: 730-734.

Sauvola, J. and M. Pietikainen, 2000. Adaptive document image binarization. Pattern Recognit., 33: 225-236.

Strouthopoulos, C., N. Papamarkos and C. Chamzas, 1997. Identification of text-only areas in mixed-type documents. Eng. Appl. Artif. Intell., 10: 387-401.

Zhang, Z., C.L. Lim and L. Fan, 2004. Estimation of 3D shape of warped document surface for image restoration. Proceedings of the 17th International Conference on Pattern Recognition, August 23-26, 2004, IEEE, pp: 486-489.