A Proposal for Print Document Acquisition Using Video Camera Equipped Device

Yousuke Uchimura^{a,*}, Yousuke Katayama^a, Lifeng Zhang^a

^aKyushu Institute of Technology, Sensui-cho, Kitakyushu-city Fukuoka 804-8550, Japan

*Corresponding Author. uchimura@boss.ecs.kyutech.ac.jp

Abstract

In recent years, the smartphone equipped with an imaging device is widely spread. It is very convenient for people to catch information around them. For example, copy one page of a book in a library occasionally, or catch interesting news on the Bulletin board someday. But such a device has a low resolution that cannot give an enough resolution for OCR soft or for human visual recognition if take the full page in one shot. Therefore, naturally, people will partially take character images with a readable size and then expect to stitching them together in efficient way. In this study, we proposed a proposal for print document acquisition using video camera equipped device. It is developed as a part of video camera scan document acquisition system. It uses one dimensional histogram for modification. And the calculation cost is low so that can be installed on a smartphone. The simulation result shows the effect is good.

Keywords : Character image calibration, one dimensional histogram, Image stitching, web camera, OCR

1. Introduction

In recent years, in order to reduce storage and administrative costs of the printed document, the document digitization is growing. Traditionally, using a digital still or a scanner is common method. However, some surface is not suitable for a scanner to work like a wet surface or a bended surface. Moreover, a fragile historical manuscript also cannot be acquired by scanner. As a non-contact scanning method, a digital camera can be used. In this work, a low profile device is used. If taking a picture of the entire page area, the character in such an image surely would have blurred due to the limited resolution. Therefore we proposed a “proposal for print document acquisition using video camera equipped device” and tried scan of printed

document.

This method takes the whole page image partially by several sub-areas, and then connecting these images to get the whole one. In the case of stitching natural images, a slight misalignment at the joint area of the reconstructed image cannot be detected obviously by human eyes. But the character images, especially for the partially acquired images in which the camera state parameter is different, even a slight deviation at the joint part can be noticed immediately. This makes the resulting documents is hard to read and difficult to be recognized by an OCR soft. To reduce such an effect, each video frame is modified independently by the estimating the camera state after the image acquisition so that the position matching can be performed accurately. Therefore, we can perform non-contact scanning independent of the scan targets form and size, with enough resolution.

2. Proposed approach

2.1 System structure

The proposed method is constructed three parts, image acquisition, self-calibration and image connection. Show the flow of system at Figure1.

Fig. 1. Flow of System.

2.2 Image acquisition

If there is a loss area between the adjacent images, character of the area are missed and cannot connect image. Therefore, the next $ImageI_t$ is acquired as overlap part of $ImageI_{t-1}$. To get overlapping images, it is necessary to calculate the cumulative amount of movement of the Web camera. In this study, calculate the cumulative amount of movement of the Web camera using optical flow estimate. Get the video frame when the amount of movement exceeded the threshold and this process is repeated until the entire scan target. Moreover calculate the “acquisition direction” from movement of web camera to connect images. For example, if web camera moved “right” and acquired image, “acquisition direction” is “right”.

2.3 Self-calibration

For a hand-held image acquisition system, the lighting condition and the imaging distance are changing every shot. Moreover the vibrations and view angle are also influence the image quality. But such conditions cannot be improved by user self completely, thus we proposed a calibration method to deal with these stuff. Detail is described the next chapter.

2.4 Image connection

First, search the similar area of the “the image connected” and “the image to connect”. Moreover calculate the “connecting coordinate” from similar area. And then “the image to connect” is connected “the image connected” like overwrite. In this study, perform the template matching for searching the similar area. Therefore set the template area in “the image to connect”. Next, search the template area from “the image connected”. Last, connect to match the template area coordinate and search result coordinate. Figure2 is flow of image connection and Figure3 is concept of image connection.

Fig. 2. Flow of Image Connection.

Fig. 3. Concept of connection.

3. Image calibration

3.1 Flow of Image Calibration

In making the correction, first, perform a threshold processing for the original image, to create a binary image. Then create one-dimensional histogram of binary image for the horizontal direction, to analyze what kinds of distortion it is. Distortion of the acquired image is assumed to occur from the slope of the Web camera, was separated into three different angles of inclination. Pitch angle, roll angle and yaw angle are rotation angles of X-axis, Y-axis and the Z-axis, respectively. (See Figure4) In this work, affine transformation and projective transformation are used to correct these angles.

Furthermore, because the camera state is unknown at first, we cannot get the all of the parameters that should be set for a projective transform at one time. The parameters are detected for each angle respectively, and modification is performed step by step. The procedure is as follows.

- Image binarization
- Modification of yaw angle
- Modification of yaw angle
- Modification of yaw angle

Fig. 4. Pitch, Roll and Yaw angle

3.2 Binarization

Normally, the text color and the brightness are changing caused by ambient lighting environment. Therefore, no threshold can be determined uniquely. In this study, we provide an adaptive threshold for each local area in the image for the binarization processing. Figure 5 shows one processing example.

Fig. 5. Image binarization.

3.3 One Dimensional Histogram

In this study we used one dimensional histogram to detect the correction parameters, here is the overview.

For a binarized character image, consider the number of black pixels on the horizontal line, we get the one-dimension histogram as shown in Figure 6. In Figure 6(a), the histogram is appeared regularly like a bar chart. Assume the number of the bar is K , and the most frequency appeared pixel number in each bar area is $n_{a\max}$ ($0 < a < K, a: \text{integer}$), then find out two y positions (R_a, L_a) when the pixel number is less than $n_{a\max}/4$ on the both side of each bar, then we get the width $W_a = R_a - L_a$ of each bar (character line width), and the blank width between two neighbored bar $I_b = L_{a+1} - R_a$ ($0 \leq a \leq K - 1, b: \text{integer}$). These parameters are used to modify the pitch angle describe in following section.

Generally, obtained image has a distortion in some degree. The one dimensional histogram rarely looks like Figure 6(a), in most case the bar is blurred looks like Figure 6(b). This confused us to determine the width of character line and the distance between two neighbor lines. In order to detect the parameters in this case, we developed a local histogram method illustrated in Figure 8. In this study we automatically set three strap areas on the image, then make one dimensional histogram of them. In Figure 7(b) the resulting histograms reflect the condition of character line's distortion even if the overall character line is not horizontal, the modification parameters can be derived from such information. The detail will be described in the following sections.

Fig. 6. One-dimension histogram.

Fig. 7. Local area image binarization.

3.4 Modification of yaw angle (Z-axis)

Calculating the gravity point of each gathered histogram shown in Figure 8(b). The gravity point g_{ck} is calculated by Eq.1

$$g_{ck} = \frac{\sum_{y=L_{ck}}^{r_{ck}} B(y) \cdot y}{\sum_{y=L_{ck}}^{r_{ck}} B(y)} \quad (1)$$

$$\theta = \frac{1}{2K} \sum_{n=0}^{2K} \theta_n \quad (2)$$

Where, k is the gathered histogram number, $B(y)$ is the pixel number when the coordinates of each strap is y . L_{ck} and r_{ck} is boundary coordinates of each gathered

histogram. The resulting image is shown in Figure8, the average θ can be calculated from θ_n . In this study, Eq.2 is though as the yaw angle. Using an affine transform, we got the modified image shown in Figure8(c).

Fig. 8. Yaw angle procedure.

3.5 Modification of roll angle (Y-axis)

After yaw angle modification, we perform local histogram processing again to the previous result again, but this time we set two straps on the image automatically. And finally find four gravity points on it. The gravity point calculation method is same as Eq.1. For an image without roll angle distortion, these four points must be the vertex of a rectangle. Using this characteristic, the roll angle can be modified easily by using projective transformation. The result shows in Figure9(b). The gravity point is shown for comparison.

3.6 Modification of pitch angle (X-axis)

Pitch angle modification parameter is estimated from an overall one dimensional histogram. Difference from previous two, the modification parameters cannot be derived from gravity point. Therefore, we have to estimate the modification parameters through the slightly variation of the width of histogram bar and the distance of blank interval shown in Figure6(a). Here we use the coordinates of gravity points shown in Figure9(b) as initial reference points coordinates

$$(p1_x, p1_y), (p2_x, p2_y), (p3_x, p3_y), (p4_x, p4_y)$$

and the estimation modified points are $(p1'_x, p1'_y), (p1'_x, p1'_y), (p1'_x, p1'_y), (p1'_x, p1'_y)$ respectively. The width of histogram bar W_i , the distance of blank interval S_j , $NumW$ and $NumS$ are numbers of

Fig. 9. Roll angle modification result.

histogram bar and blank interval respectively. Because we use an iterative algorithm for parameter estimation, first we define a evaluation function as

$$E = \frac{(comp_W - center_W) + (comp_S - center_S)}{2} \quad (3)$$

where,

$$center_W = \frac{\sum_{i=1}^{NumW} W_i \times i}{\sum_{i=1}^{NumW} W_i}$$

$$center_S = \frac{\sum_{j=1}^{NumS} S_j \times j}{\sum_{j=1}^{NumS} S_j}$$

$$comp_W = \frac{\sum_{i=1}^{NumW} i}{NumW}$$

$$comp_S = \frac{\sum_{j=1}^{NumS} j}{NumS}$$

If $comp_W < center_W \cap comp_S < center_S$,

$$p1'_x = p1_x + \frac{F_w + F_s}{2}, p1'_y = p1_y$$

$$p2'_x = p2_x, p2'_y = p2_y$$

$$p3'_x = p3_x + \frac{F_w + F_s}{2}, p3'_y = p3_y$$

$$p4'_x = p4_x, p4'_y = p4_y$$

Otherwise,

$$\begin{aligned}
p1'_x &= p1_x, p1'_y = p1_y \\
p2'_x &= p2_x + \frac{F_w + F_s}{2}, p2'_y = p2_y \\
p3'_x &= p3_x, p3'_y = p3_y \\
p4'_x &= p4_x + \frac{F_w + F_s}{2}, p4'_y = p4_y
\end{aligned}$$

Where

$$\begin{aligned}
F_w &= (comp_W - center_W) \times (comp_W)^2 \\
F_s &= (comp_S - center_S) \times (comp_S)^2
\end{aligned}$$

Perform the projective transformation iteratively using the estimated coordinates, and then evaluate the accuracy of the correction by the evaluation function E . Terminate the iteration when $E \leq T$ (T is threshold). Otherwise remake a one dimensional histogram and repeat this procedure again. Finally, we get the completely modified result.

3.7 After Processing

After modification, there are needless pixels in the calibrated image. In order to remove these pixels, we performed two after processing.(Figure10)

Moreover, processed image are initialized to a same character scale.

(a) Pitch angle modified result (b) Remove needless pixels

(c) Remove break string

Fig. 10. After Processing

4. Simulation

Figure11(a) are Original images captured by a web camera(logicool Webcam Pro 9000, frame size 320×240), Figure11(b) are calibrated images of Figure11(a), the distortion was modified. And Figure12 is connected image.

(a) Original acquired partial document images

(b) Calibrated images of above

Fig. 11. Simulation result.

Fig. 12. Connected image.

5. Conclusion

In this research, we focused on the image reconstruction and the calibration of a printed document obtained partially by a Web camera. The simulation results show that the distortion is repaired well and our method is able to recognize character of connected image.

However, there is still a drawback that the image is

blurred after processing. This is mostly comes from the iterative calculation of pitch angle estimation, and we need to improve the algorithm to reduce the cumulative error.

In near future, we would improve the accuracy of image calibration and process of image connection.

Acknowledgment

This work is partially supported by Global Security Design, Inc. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

- (1) Jian Liang : “Geometric Rectification of Camera-Captured Document Images”, IEEE Transactions On Pattern Analysis and Machine Intelligence, pp.591-605, 2008
- (2) Jian Liang, David Doermann, Huiping Li : “Camera-based analysis of text and documents: a survey”, International Journal on Document Analysis and Recognition, doi:10.1007/s10032-004-0138-z, 2005
- (3) Nikolaos Stamatopoulos, Basilis Gatos, Ioannis Pratikakis, and Stavros J. Perantonis : “Goal-Oriented Rectification of Camera-Based Document Images”, IEEE Transactions On Image Processing, Vol. 20, No. 4, pp.910-920, April 2011
- (4) Petr Somol and Michal Haindl : “Novel path search algorithm for image stitching and advanced texture tiling”, The 13th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, pp.155-162, 2005
- (5) Mehmet Sezgin and Bulent Sankur : “Survey over image thresholding techniques and quantitative performance evaluation”, Journal of Electronic Imaging 13(1), pp.146-165, January 2004.
- (6) Y.Suematu, H.Yamada : “Image Processing Engineering”, Corona Publishing Co., Ltd., pp.114-115, 2000.
- (7) Nara Institute of Science and Technology : “OpenCV programming book” , Mynavi Corporation, pp.1-440, 2008