

Research Article

Open Access

Straightening Uniformly Folded Document Image

Patel HN and Modi PR*

Electronics and Communication, Engineering Department, A. D. Patel Institute of technology New Valla Bhai Vidyanagar, Gujarat, India

Abstract

Image processing is very much important area of research nowadays and is being taken interest deeply as the mobile phones and smart cameras are demands of the emerging global markets. Document image processing is the sub branch of it which includes processing the images of the paper may or may not be containing any write up or say textual contents, graphics, tables or any informative thing in form of written or printed stuff. Many a time's we people are habituated to fold the document papers in our hands and put them in pocket or somewhere. If then we want them to be scanned or apply OCR (OPTICAL CHARACTER RECOGNITION) on them then we need to capture an image of such folded documents. For image acquisition there are many sources like digital cameras, mobile phone cameras, printers etc. are available around us. When we need to scan them to convert them into electronic form (to perform OCR on them), then the problems occur because the folded documents cannot be easily scanned as they have folds and ue to such folds there are huge possibilities of shadows in fold parts which can distort the quality of the document as well as this may make the readability of the document poorer. In this paper, the authors have proposed an algorithm to straighten such folded document image to get accurate OCR results in accordance with least distortion in quality of the same. The derived method helps to straighten the uniform folds digitally. So that the processed document can be applied to get improved OCR results with improved in terms of accuracy.

Keywords: Image processing; Corner points; Sub-images; Projective transform

Introduction

When images of folded documents are to be straightened, they need to be pre-processed in order to restore them otherwise scanning such folded documents is not much accurate and the quality of original data in document images get affected due to it, which is not desirable. Also content in document images may distort causing their quality degradation. Folded documents may need to be restored to perform OCR on them or to scan them. It is a very real time application because nowadays all government policies as well as private firms operate digitally. In case such documents were folded by we then that algorithm become helpful when we need to upload the scanned images to the websites suggested by them. For folded document images issues are shadow in fold parts and warping in horizontal as well as vertical direction.

Character and line segmentation techniques were used to know which part in image is warped and then the restored images after straightening them were obtained using Thin Plate Splines (TPS) [1]. Segmentation part provides the warping direction. Line segmentation involve connected component analysis and projection which gives baseline, while character segmentation involve the same along with envelop analysis to provide key points in order to get the destination image from the original warped image. TPS performs global point to point mapping to have restored image finally. Resultant images were obtained by using fine dewarping to straighten the warped images followed by coarse dewarping [2,3]. Coarse dewarping uses projective transform to map curved surface to rectangular area. After that fine dewarping was performed which is based on word detection by its pose normalization. Employs difference between horizontal and vertical projection profile and ends up with concluding that noise was greatly removed with horizontal projection profile [4]. The histogram based technique was used to obtain both the projection profiles. Vertical projection was obtained on binarized image with white background and black data, by summing the black pixels of each column and computing the energy at number of angels and based on angle at which the maximum energy was found, at that angle, image was rotated. Energy was computed as the sum of square of black pixels for each column in document images. The horizontal projection is same but works for rows instead of columns. They have used projective transform to estimate the 3-D shape of an object when a thick bound book was scanned by scanner [5,6]. They also performs shade removal and x, y-axis corrections while scanning. In proposed work, different uniformly folded images were taken as input images [7-11] (Figure 1).

Images of folded documents were captured at different illumination conditions to justify efficiency of the proposed algorithm.

Material and Methods

The objective of the method is to straighten the document image in the way that the warping is removed and the content in that document image is not distorted after straightening it and to have better OCR



*Corresponding author: Modi PR, Electronics and Communication, Engineering Department, A. D. Patel Institute of Technology New Valla bhai Vidyanagar, Gujarat, India, Tel: 02692 233 680; E-mail: modipayal98@gmail.com

Received June 29, 2016; Accepted July 05, 2016; Published July 30, 2016

Citation: Patel HN, Modi PR (2016) Straightening Uniformly Folded Document Image. J Electr Electron Syst 5: 191. doi: 10.4172/2332-0796.1000191

Copyright: © 2016 Patel HN, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

accuracy [12-16]. The conditions to work the designed method in good way were that the images should be captured with high quality (resolution) camera so that the images were not blurred because if the input image, already is blurred then the resultant straightened image will not produce good OCR accuracy and the corners of the document should not be torn because the proposed method of straightening is based on processing the corners of the document [16-20] (Figures 2-4).

Images of uniformly folded documents were captured. Below are specifications of database prepared for algorithm. Images are in RGB plane and captured by high resolution (5MP to 16 MP) and dpi (72 to







J Electr Electron Syst ISSN: 2332-0796 JEES an open access journal 96 dpi) cameras. Horizontal folded images, vertical folded images and images with both types of folds were captured. Images were taken in different lighting conditions as well as at different times to justify the algorithm. Images should be non-blurred. Corners should not be torn. Pre-processing is a second step in which entire straightening procedure were carried out to make the straightened image which is ready to apply optical character recognition on it. First of all the corners were obtained where there are folds, and based on that input images are divided into sub-images. All these sub-images were processed individually and straightened by applying projective transformation on them. The next pre-processing step was to make sure whether there were characters on fold line by horizontal projection profile for every single sub-image. If yes then the fold line was shifted downward where there were no characters. Sub-images were then resized and merged. All the sub steps in pre-processing are explained in detail as follows.

The folded input images have number of corner points according to type of folds like left up, right up, right down and left down corner locations of folded parts.

For example Figure 5 has horizontal three folds and ten corner points for four sections of the same image. These sections are based on folds. If it is horizontal one fold image it has two such sections, if horizontal two fold image, then three such sections. Further every section has four corners as shown which are upper left corner (LU), upper right corner (RU), lower right corner (RD), lower left corner (LD). Top most sections LD; RD is LU and RU for second section exactly below it respectively. For corner detection edge of folded image needs to be extracted. For that first of all images were converted to binary. When it is done there may be noise or generation of holes due to variation in lightening conditions, which is shown in Figure 6. These holes or noises are harmful because it too would be considered while finding corner points as it's explained in following part and must be removed. Holes in images are filled but still noise won't be totally removed. There is pepper noise in background which will disturb edge detector in its work. So image is filtered by two dimensional order statistic filter which replaces each element in image by the minimum of its north, east, south, and west neighbors resulting in noise-free image in most cases. The resultant images after all these steps are as below.

Where, Filter=
$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Filtered image is eroded twice in order to remove text or any white



Page 2 of 7

dots still remaining outside image in black background because these letters too will have edges which are not required and provide wrong edge information for corner detection [21,22]. For erosion square structuring element is used with all one and size 3*3. Erosion contracts an object in image based on structuring element.

$\operatorname{im}\Theta \operatorname{se} = \left\{ x \in z^2 \mid (\operatorname{se})_x \subseteq \right\}$	im	ı).		(1)
Where,	. 1	1	1]	
Se=Structuring element= 1	. 1	1	1	
im=image.	. 1	1	1	

Finally for image is ready for edge detection. Canny edge detector is used for that as it provides best results than any other methods. The Canny method finds edges by looking for local maxima of computed gradient of image which is calculated by taking the derivative of a Gaussian filter. The canny method uses two thresholds for detection of strong as well as weak edges, and includes the weak edges in the output only if they are connected to strong edges. So this method is not "fooled" by noise, and more likely to detect true weak edges. Thresholds are selected in such a way that ratio of low to high threshold in 0.4 depending on intensity of image. The twice eroded and edge detected images are shown below in Figure 7.

Then edge is detected and corners of edge of an image are to be found. It's having horizontal three fold document image, it is not possible to find all the ten corners at the same time. So first the



Figure 6: (a) Input image (b) Holes filled from an image(c) Image filtered by order statistic filter.



J Electr Electron Syst

ISSN: 2332-0796 JEES an open access journal

concentration is on finding prime corners i.e. top most left, right and fourth section's bottom left, right corners. For that the steps as below are followed: Construct an array to store all points of edge. Euclidean distance between (1,1) and array; (1,n) and array; (m,n) and array; (m,1) and array is found. For every points where there is a minimum Euclidean distances are the prime corners LU, RU, RD and LD. Size (im)=m*n. Where,

Page 3 of 7

m=total number of rows, n=total number of columns (Figure 8).

Still six corners are undetected. For finding them column difference is useful. Left side: columns of LU up to LD are subtracted from 1 (original image's first column). Right side: column of RU up to RD are subtracted from n (original image's last column) (Figure 9).

Where there are sharp changes in the plot of distance versus row number those points are considered as sub-corners. Peaks suggesting sharp changes are so many as image is not smooth in the plot. If polynomial fitting of order 1(i.e. line y=mx+b: m-slope, b-intercept) is selected then it provides 2 points: first slope and second intercept. The first parameter is useful as where there is a sudden sign change





after many regular signed values those respective rows can be taken as peak. Plots are not smooth so Gaussian filtering is applied with frame size [800,800] and standard deviation 80. Window size for polynomial fitting is selected 50. This method provides true sub-corners rows and with the help of same, respective column too can be found and is accurate in 72.63% folded images of every type when tested on 365 images with different type of folds. Error rate in corners is maximum + or -100 (Figure 10).

In this algorithm, we have separated the folded parts in a document image by using corner points for e.g. if it is horizontal one fold image, then it is separated into two parts, upper part above the fold and lower part below the fold. Two horizontal folded images produces three subimages, three horizontal fold images can be separated into four subimages (Figure 11). The same way it is applicable to all the uniform folds described above. Same way vertical one fold image can be separated into two sub-images, one is left of the fold and other is right of the fold. Image with vertical and horizontal fold can be divided into four subimages. Initially these corner points are obtained manually as upper left, upper right, down right and down left corners. The conditions for validity of this algorithm are that the images should be captured





the way so they are not blurred and the corners are not torn. Here the entire algorithm is explained for horizontal three folds image.

Then we have used the projective transform in order to straighten these individual sub-images obtained in this. Image transform there are parts of folded image or four sub-images and they are needed to straighten individually. For that projective transform is used. In such transformation technique, the spatial transformation structure is required. The logical thing in this entire algorithm is to design the general spatial transformation structure that can straighten all images for our case. Creating spatial transformation structure requires input points and base points which work as control points. For input points (x,y) co-ordinates are set such that all the corner points of horizontal lines are having same y locations and vertical lines are having same x locations. While base points which work as control points are obtained by setting the left, up, down and right parameters according to the mean of them. For e.g. Left is the mean of upper left and lower left's x co-ordinate. Right is the mean of upper right and lower right's x coordinate. Up is the mean of upper left and upper right's y co-ordinate. Down is the mean of lower left and lower right's y co-ordinate as shown in algorithm described below. Now the inter-section of Left and Up gives the final upper left point of resultant image, the inter-section of Up and Right gives the final upper right point of the resultant image and same way the other two points of resultant image are obtained. Finally applying the spatial transformation structure on every individual sub-images of horizontal three fold image we will get the projective transformed straightened images in Figure 12.

All the sub-images are straightened but they are containing black background which is not required. So, the sub-images are cropped such way that the black background is removed and we are left with the subimages without background (Figure 13).

When all the sub-images of original folded documents are straightened and background is removed too, then the final step is to merge them to get the entire image which is straightened version of the original image. Before merging images, horizontal projection profile is used to check whether there are characters on fold line on gray scale image. Because if there are characters on fold line then while resizing





Figure 13: Sub-images cropped to remove black background.



before merging these characters will be shifted and not well-aligned resulting in reduced OCR accuracy at that line (fold line) also making it not readable. In projection profile every section of images except top section are processed. Sum of pixels in columns are carried out for each row after converting image to grey scale. If the value of sum is maximum then it means there is white part meaning no character on fold line. So no need to shift fold line. If not then initial 25 lines are checked where value of sum is maximum i.e. where the white line with no characters and first row is shifted upward making white line found by projection profile the first row. Then images are resized. Resizing procedure is done by providing scale in form of number of columns and number of rows of one of the sub-image which computes the number of rows or columns automatically in order to preserve the image aspect ratio. The sub-images are resized first the way that all of them are having same number of columns. Then they can be easily merged which is the straighten version of the input horizontal three folds images in Figures 14 and 15.

The same method can be applied for straightening any uniform folded images as suggested in Figure 1.

OCR helps to convert physical document in electrical or say soft form. When OCR is applied on the resultant images which are straightened after procedure explained above the results were quite promising. For OCR, ABBYY Fine Reader 12 software is used which is trustworthy and reliable. It includes an analysis of OCR results on original folded image, output straightened image without projection profile and with projection profile. When OCR performed on original image accuracy of readability on data in it was 81.25% which is not so bad but time consumed while performing the same was quite high. Whereas the second case is of OCR on straightened image but with characters on fold line and no horizontal projection profile, found resultant accuracy has fallen to 12.50 which is very poor. When the OCR performed on output straightened images with shifted fold line by horizontal projection profile, OCR results rise to 93.75%, which to in less consumption of time compared to both the cases before for same number of images (1.39 sec) (Table 1).

Algorithm

This chapter involves all the steps explained above in summarized form and more discriminative study of the same. The algorithm will give clear picture of what is the goal of the system and how to achieve it.

1) Partition the image into sub-images according to folds by using corner points of original folded image based on concept of Euclidean distance for prime corners and slope by polynomial fitting of column difference for sub-corners as explained above.

2) All the sub-images are needed to be processed individually. Each sub Image has 4 corners: upper left LU, down left LD, upper right RU and downright RD.



Images	Character on fold line(%)	Character on fold line(%)
Folded document images	12.50	93.75

 Table 1: Percentages on character on Fold line.

Citation: Patel HN, Modi PR (2016) Straightening Uniformly Folded Document Image. J Electr Electron Syst 5: 191. doi: 10.4172/2332-0796.1000191

3) The sub-images are straightened using projective Transform. Projective transform is explained briefly in section below. For projective transform, spatial transformation structure needs to be designed using input points and base points. Let the corner points of sub-image be C1=[LU(1) LU(2), RU(1) RU(2), RD(1) RD(2), LD(1) LD(1)]. To get input points C2=[LU1, LD1, RU1, RD1], which are nothing but the documents corner points by not considering black background, for designing spatial transformation structure following mechanism is applied:

LU(1)=upper left corner's column number,

LU(2)=upper left corner's row number,

LD(1)=down left corner's column number,

LD(2)=down left corner's row number,

RU(1)=upper right corner's column number,

RU(2)=upper right corner's row number,

RD(1)=down right corner's column number,

RD(2)=down right corner's row number.

LU1's column value depends on whether LU's column value is bigger or smaller than LD's column value. If it is smaller, then LU1(1) is having 1 making it first column otherwise it is equals to 1 more than the difference between LU and LD's column number to make it smaller. And LU1's row can be obtained the same way but the comparison is made between LU and RU's row. Such type of designing will help the sub-images to make the straightening easier. Do the same for RU1, RD1 and LD1 to get the column and row co-ordinates of them.

4) Corner points detected are now:

C2=[LU1(1) LU1(2), RU1(1) RU1(2), RD1(1) RD1(2), LD1(1) LD1(1)].

5) Base points, which are to work as control points in constructing spatial transformation structure, are CR=[L U, R U, R D, L D]. They are obtained by taking mean of respective components of C2. For example L is the mean of LU1 and LD1's column value, U is the mean of LU1 and RU1's row value, R is the mean of RU1 and RD1's column value and D is the mean of RD1 and LD1's row value. Finally C2 and CR work as input and control points respectively to construct spatial transformation structure. Then this projective transformation structure is applied on sub-images obtained. That way inter-section of L and U points, U and R points, R and D points, D and L points will construct the straightened image. It will straighten the sub-images.

6) Resultant straightened sub-images are cropped to remove black background which is not of the interest to get the straightened subimages.

7) Our ultimate goal is to get the image which is straightened version of folded image, not the sub-images. So we need to merge the straightened sub-images obtained in the previous step. But before that the lower sub-images rows are applied horizontal projection profile to check whether there are characters on fold line. If so, then fold line is shifted down wards and at the same value upper sub-images shifted down ward. Also those sub-images are of different sizes. So they need to be resized the way that all the sub-images are having same columns. We have nothing to do with the rows. Finally they are merged at the end to get image which is straightened.

Experimental Results

The same method explained above can straighten any uniformly folded images. But if the images captured are blurred then it will not provide good OCR results. The experimental results for horizontal one fold image, horizontal two folds image, horizontal three folds image, horizontal and vertical fold image, vertical one fold image are shown below. All the images were captured by good quality cameras and all images are JPEG images with colour representation RGB (Figure 16).



J Electr Electron Syst ISSN: 2332-0796 JEES an open access journal Page 6 of 7

Citation: Patel HN, Modi PR (2016) Straightening Uniformly Folded Document Image. J Electr Electron Syst 5: 191. doi: 10.4172/2332-0796.1000191

In Figure 16(a) shows the document image with the one horizontal fold at the centre of the document. The image is captured in brighter lighting conditions and on black surface. After straightening the resultant image is shown.

In Figure 16(b) is an input image with two horizontal folds so the warping in middle portion is seen in image. The image is captured at darker lighting conditions and on black surface. Above the fold-lines shadow would be there. But resultant image shows that it does not affect the resultant straightened image.

In Figure 16(c) is a horizontal three folds image of document taken at brighter lighting conditions than in (b). It is much curvy than above two examples because it has much folds and therefore much warping too. The resultant image is the straightened form of input image. For straightening it needs to be divided into four sub-images by using corner points. Then they are straightened by using projective transformation. So it requires processing four sub-images individually and then merging them as explained in algorithm.

In Figure 16(d) shows the document image with horizontal one fold and vertical one fold. Here to the four sub-images are obtained but the corner points differ then the previous case of horizontal three folds image. Because horizontal fold is partitioned to provide two sub-images: one above the fold and one below the fold. Then after upper and lower both sub-images are again partitioned as left of the vertical fold and right of the vertical fold. Then all sub-images are processed the same way and resultant straightened image is obtained.

In Figure 16(e) is an input image with vertical one fold which is processed in similar way as the horizontal one fold image. The image is partitioned into two sub-images; the only difference is the corner point selection. It has so much warping in horizontal direction which is hard to read the write up in it. But after straightening the warping is removed and it is good to read then.

In Figure 16(f) horizontal one fold document image with coloured background is there with shadow at lower fold portion. The resultant is an output straightened image.

Conclusion

The derived algorithm gives very good quality of content in resultant straightened document image even in the folded parts. It can be proved from the image above. The proposed algorithm has less complexity and good speed. The algorithm has robustness because it can work well with images in all illumination conditions and shadow conditions background independently. OCR results rise to 93.75% for content containing fold line which would be shifted a little when required while straightening than that 12.5% at fold line of input folded images when there were characters on fold line and fold line was not shifted by horizontal projection profile, while folded input image had OCR accuracy of readability 81.25%. So all in all, just by giving the path we can have straight image from the folded document image.

Future Scope

The algorithm can be optimized for straightening of document images with non-uniform folds. Moreover, the algorithm can be optimized so that when straightened the fold is equalized and document image is perfectly straight and fold is not visible in resultant image. If background is white then automatic corner detection is challenging and can be considered as future scope.

- Zhang Y, Liua C, Dinga X, Zoub Y (2008) Arbitrary Warped Document Image Restoration Based on Segmentation and Thin-Plate Splines. 19th international conference on pattern recognition, ICPR, Tampa, FL.
- Stamatopoulos N, Gatos B, Pratikakis I, Perantonis S (2008) A two-step dewarping of camera document images. Eighth IAPR workshop on document analysis systems, DAS 08: 209-216.
- Stamatopoulos N, Gatos B, Pratikakis I, Perantonis S (2011) Goal-Oriented Rectification of Camera-Based Document Images. IEEE Transactions on image processing 20: 910-920.
- Jain B, Borah MA (2014) AComparison paper on skew detection of scanned document images based on horizontal and vertical projection profile analysis. International journal of scientific and research publications 4: 1-4.
- Tiwari L, Kumar B, Patnaik T (2014) Robust Camera Captured Image Mosaicking for Document Digitization and OCR Processing. International Conference on Information: 100-105.
- Zhang Z, Tan C, Fan L (2004) Estimation of 3D Shape of Warped Document Surface for Image Restoration. 17th international conference on Pattern Recognition 1 Date of Conference: 486-489.
- Negi A (2001) Document Image Processing and Optical Character Recognition Systems for Indic Scripts. School of Computer and Information Sciences University of Hyderabad.
- Zhang L, Tan L (2006) Restoring warped document images using shape-fromshading and surface interpolation. In Proc 18th International Conference on Pattern Recognition 1: 642-645.
- Wada T, Ukida H, Matsuyama T (1997) Shape from Shading with Inter reflections Under a Proximal Light Source: Distortion- Free Copying of an unfolded Book. International Journal of Computer Vision 24: 125-135.
- Brown M, Seales W (2004) Image restoration of arbitrarily warped documents. IEEE Trans On Pattern Analysis and Machine Intelligence 26: 1295-1306.
- Zhang Z, Tan C (2003) Correcting document image warping based on regression of curved text lines. International Conference on Document Analysis and Recognition, Edinburgh, Scotland: 589-593.
- Wahl F, Wong K, Casey R (1982) Block segmentation and text extraction in mixed text/image documents. Comput Graph. Image Process 20: 375-390.
- Okun O, Pietikäinen M, Sauvola J (1999) Document skew estimation without angle range restriction. International Journal on Document Analysis and Recognition 2: 132-144.
- Tang Y, Suen C (1993) Image Transformation Approach to Nonlinear Shape Restoration. IEEE Trans Systems Man and Cybernetics 23: 155-171.
- Lavaille O, Molines X, Angella F, Baylou P (2001) Active Contours Network to Straighten Distorted Text Lines. Proc Int'I Conf Image Processing: 1074-1077.
- Weng Y, Zhu Q (1996) Nonlinear Shape Restoration for Document Images. Proc IEEE CS Conf Computer Vision and Pattern Recognition: 568-573.
- Tsoi YC, Brown MS (2004) Geometric and Shading Correction for Images of Printed Materials-A Unified Approach Using Boundary. Proc IEEE CS Conf Computer Vision and Pattern Recognition 1: 240-246.
- Burger W, Burge M (2008) Digital Image Processing. 1st (edn) An Algorithmic introduction using Java. Springer.
- 19. Gonzalez R, Woods R (2012) Digital Image Processing. $3^{\rm rd}$ (edn) Pearson Education.
- Jayaraman S, Esakkirajan S, Veerakumar T (2011) Digital Image Processing. 1st (ed) Tata McGraw Hill.
- 21. http://www.img.cs.titech.ac.jp/~akbari/pmwiki/uploads/Site/Felipe-rep.pdf
- 22. http://math.stackexchange.com/questions/96662/augmented-realitytransformation-matrix-optimization