

Hierarchical Segmentation of Falsely Touching Characters from Camera Captured Degraded Document Images

Satadal Saha¹, Subhadip Basu² and Mita Nasipuri³

¹ ECE Department, MCKV Institute of Engg.
Liluah, Howrah, India

² CSE Department, Jadavpur University
Jadavpur, Kolkata, India

³ CSE Department, Jadavpur University
Jadavpur, Kolkata, India

Abstract

An innovative hierarchical image segmentation scheme is reported in this research communication. Unlike static/ spatially divided sub-images, the current innovation concentrates on object level hierarchy for segmentation of gray scale or color images into constituent component/ sub-parts. As for example, a gray scale document image may be segmented (binarized in case of two-level segmentation) into connected foreground components (text/ graphics) and background component by hierarchically applying a gray level threshold selection algorithm in the object-space. In any hierarchy, constituent objects are identified as connected foreground pixels, as classified by the gray scale threshold selection algorithm. To preserve the global information, thresholds for each object in any hierarchy are estimated as a weighted aggregate of the current and previous thresholds relevant to the object. The developed technique may be customized as a general purpose hierarchical information clustering algorithm in the domain of pattern analysis, data mining, bioinformatics etc.

Keywords: *Binarization, Connected component labeling, Hierarchical object, Weighted average threshold.*

1. Introduction

Design of optical character recognition system for camera images is evolving as an emerging field for various industrial applications. In general, in most of the applications due to the poor quality of the camera or due to the unconstrained environment the images contain a substantial amount of noise and shades that affect the output of the OCR as a result of improper segmentation and binarization of the characters. The proposed technique is highly applicable in these types of scenario. Segmentation refers to the process of partitioning a digital image into a number of segments (sets of pixels) in a meaningful way. Segmentation of an image means

isolation of different components of it based on some criteria. If the image is segmented based on the color value or gray value then the outcome of the segmentation process is a set of isolated components having similar colors or gray values respectively. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. The recognition section of an OCR system requires the segmented set of characters as an input. In case of printed documents the characters are normally separated from each other and the segmentation process is more or less simpler. But in case of camera captured degraded document images, some of the adjacent characters are touching in nature because of the gray shades or noise in between. In case of normal segmentation algorithms these touching characters remain in one segment representing a single character. The objective of this technique is to segment the image into a number of components based on the character level texts written therein. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s).

A lot of works has already been reported in the form of research papers during last couple of decades or so. One earlier work [1] has been reported on online handwritten

character recognition of six types of scripts. The method uses vertical projection of the lines and uses the inter-line distance for line segmentation. The words are segmented from the lines using the horizontal projection of the lines and uses the inter-word distance. But the technique may not be suitable for considerably skewed lines. A work [2] is reported on text line extraction from multi-skewed handwritten document images. The work implements a water flow technique for extraction of text lines with high rate of success. But this method cannot dissociate two unintentionally connected neighboring segments. In another work [3] text line segmentation from unconstrained handwritten document images by comparing neighborhood connected components has been reported. The method is applied over four types of scripts with high rate of success for text line extraction. This method also lags in dissociating two unintentionally connected neighboring segments. Three types of clustering algorithms are discussed in a research report [4]. A hybrid model is also discussed. One of the techniques uses bounding box analysis of the adjacent characters using adjacency, aspect ratio and size of the individual character. This method is applicable where a priori knowledge of the position and the size of the characters are known, e.g. localization and recognition of postal code characters. The second method uses splitting of the connected components in under segmentation cases. It finds the segmentation point and the segmentation path by finding the turning point between successive minima of a connected component in anticlockwise direction. The method decides whether the case is of true under segmentation or not by vertically scanning the pixels and counting the number of transitions from black to white pixels. This may not be suitable in case of slanting characters. A third method discussed in the same report uses recognition based segmentation after a preliminary segmentation process and resegmenting the adjacent segments by joining or splitting depending on the recognition result. The method appears to be very time consuming as the recognition and segmentation processes may be repeated a number of times. In another report [5] image gradient is generated and then watershed transform is applied to generate primitive segments. Then a merging algorithm (region adjacency graph) is used to generate finally merged segments. In a work [6] generalized K-mean clustering algorithm is used. Gibbs random field model is used to include the spatial constraints. In the works reported in [4], [5] and [6] a merging algorithm is applied to include any minute segment near the main segment. This may result in losing the isolated minute segments. In one earlier work [7] a region based representation of image and video useful to multimedia services has been represented. It then proposes a partition tree representation of image. Another report [8] uses multiscale Bayesian image segmentation on

multispectral remotely sensed images. For the works [7] and [8] no pixel level object hierarchy is reported in the referred work. No compatible ground truth data are found for comparison. A work [9] is reported in which text is localized and segmented from complex images and videos. The method applies multiresolution approach and uses multimedia text detector and segmentor. It also estimates text color using vector quantization method. But the method assumes horizontal text appearance and the application area is web page images and videos which can definitely be assumed to be fare enough and well designed by the web developer. An effective technique [10] reported uses the colour feature as well as the spatial information of the text region to isolate the potential text region from still and video images. Though the method retrieves the text region from the scene, it retrieves some non-text regions also which bear the edge feature similar to text. Over and above, the method does not isolate the neighbouring characters to generate character level segments as output, though the characters are already isolated in the image as seen visually. A Hough transform based efficient technique is reported in [11] but the method could not dissociate very closely spaced characters. The short review mentioned above indicates that most of the challenges of text segmentation, whether it is line segmentation or word segmentation or character segmentation, have been solved with a satisfied amount of successfulness of the result. In general, most of the previous techniques have been operated upon degraded or non-degraded scanned document images. Now with the technological development of hand held devices, such as mobile phone, digital camera, CCTV camera etc., automatic recognition of the texts from the images captured by the camera of the device has now become a requirement of the IT industry. As obvious, an OCR system is used for the recognition of the text, in which segmentation plays a vital role. Interestingly, the traditional techniques based on document image segmentation do not perform satisfactorily when applied on camera images captured in an unconstrained outdoor environment. The main motivation of the proposed technique is segmentation of text from document images as well as from camera images which are usually degraded by shadows, projection angle of the camera face, environmental pollutions etc.

2. Proposed Technique

2.1 Brief Description of the technique

The designed hierarchical image segmentation scheme is implemented based on image binarization method in the proposed technique. The objective of any binarization

algorithm is to identify a grey level threshold value for a set of pixels in the subimage to convert them into either of the two states, viz. black or white. To do so, some techniques decide the threshold over the complete image, spatially divided subimages or even on the basis of the neighbors of each pixel. The current work, however, attempts to subdivide the image based on logical object hierarchy and compute binarization thresholds for each pixel of the image in each hierarchy. A final decision is then made for each pixel on the basis of the weighted average of such thresholds across different hierarchies. The mathematical derivation of the developed algorithm is discussed below.

Mathematical notations used:

$H(X)$: Original grey level image

$g_{l,k}^{p,q}(x)$: kth grey level object component at level l obtained from qth object at level p

$b_{l,k}^{p,q}(x)$: kth binarized object component at level l obtained from qth object at level p

$t_{l,k}^{p,q}$: local threshold for $g_{l,k}^{p,q}(x)$

$T_{l,k}^{p,q}$: average cumulative threshold for $g_{l,k}^{p,q}(x)$

$n_l^{p,q}$: number of binarized connected components

$b_{l,k}^{p,q}(x)$ at level l

The original grey level image $H(X)$ is used to find the initial binarization threshold $t_{0,0}^{0,0}$ using which the binarized image component $b_{0,0}^{0,0}(x)$ is obtained. A standard binarization algorithm (OTSU) is used for this purpose. This binarized image is passed through a connected component labelling algorithm to generate a set of binarized image objects $\{b_{l,k}^{0,0}(x)\}$ (k ranging from 0 to $n_1^{0,0}$) in the next level of object space hierarchy, The corresponding set of grey level components as obtained

from the original image $H(X)$ is denoted as $\{g_{l,k}^{0,0}(x)\}$

(k ranging from 0 to $n_1^{0,0}$). These grey components are again binarized and the binarized components are then passed through a connected component labelling algorithm to generate the next level objects. The idea is mathematically represented as follows:

$$t_{0,0}^{0,0} | H(X) \rightarrow \{b_{1,0}^{0,0}(x), b_{1,1}^{0,0}(x), b_{1,2}^{0,0}(x), \dots, b_{1,n_1^{0,0}-1}^{0,0}(x)\}$$

which indicates that using the threshold $t_{0,0}^{0,0}$ the 0th level object $H(X)$ is binarized and a CCL algorithm is used to partition the binarized image $b_{0,0}^{0,0}(x)$ into $n_1^{0,0}$ number of 1st level binarized objects $b_{1,0}^{0,0}(x), b_{1,1}^{0,0}(x), b_{1,2}^{0,0}(x), \dots, b_{1,n_1^{0,0}-1}^{0,0}(x)$. Each grey level component object $g_{1,k}^{0,0}(x)$ corresponding to this first level binarized object $b_{1,k}^{0,0}(x)$ is partitioned into further objects by finding average cumulative threshold $T_{1,k}^{0,0}$ for each object k as follows:

$$T_{1,0}^{0,0} | g_{1,0}^{0,0} \rightarrow \{b_{2,0}^{1,0}(x), b_{2,1}^{1,0}(x), b_{2,2}^{1,0}(x), \dots, b_{2,n_2^{1,0}-1}^{1,0}(x)\}$$

$$T_{1,1}^{0,0} | g_{1,1}^{0,0} \rightarrow \{b_{2,0}^{1,1}(x), b_{2,1}^{1,1}(x), b_{2,2}^{1,1}(x), \dots, b_{2,n_2^{1,1}-1}^{1,1}(x)\}$$

and so on. Similarly, any arbitrary grey level object $g_{1,k}^{1,0}$ is partitioned into further $n_3^{2,2}$ number of binarized objects by finding the average cumulative threshold $T_{2,2}^{1,0}$ and passing the binarized objects through a CCL algorithm as given below:

$$T_{2,2}^{1,0} | g_{1,k}^{1,0} \rightarrow \{b_{3,0}^{2,2}(x), b_{3,1}^{2,2}(x), b_{3,2}^{2,2}(x), \dots, b_{3,n_3^{2,2}-1}^{2,2}(x)\}$$

So in general it can be said that using the threshold $T_{l,k}^{p,q}$ the grey level object component $g_{l,k}^{p,q}(x)$ binarized to generate image component $b_{l,k}^{p,q}(x)$ and it is passed through a CCL algorithm to partition it into $n_{l+1}^{l,k}$ number of new binarized object components as represented below:

$$T_{l,k}^{p,q} | g_{l,k}^{p,q}(x) \rightarrow \{g_{l+1,0}^{l,k}(x), g_{l+1,1}^{l,k}(x), g_{l+1,2}^{l,k}(x), \dots, g_{l+1,n_{l+1}^{l,k}-1}^{l,k}(x)\}$$

The average cumulative threshold $T_{l,k}^{p,q}$ is calculated taking the weighted average of all the local thresholds for all of its predecessor objects in the object hierarchy using the generalised formula:

$$T_{l,k}^{p,q} = \left(w_0 t_{0,0}^{0,0} + \sum_{i=1}^l w_i t_{i,k}^{i-1,q} \right) / (l+1) \quad (1)$$

2.2 Algorithm of the proposed work

1. Start
2. Read Image $H(\mathbf{X})$, where, $\mathbf{X}=[x; y]$.
3. Initialize image hierarchy $p=0, q=0, l=1, k=0, top=0, L=0$.

4. Estimate the binarization threshold $t_{0,0}^{0,0}$ for the input image $H(X)$ generating the 0th binarized image object $b_{0,0}^{0,0}$.
5. Apply connected component analysis algorithm on the object $b_{0,0}^{0,0}$ to generate the first level binarized connected components $\{b_{1,k}^{0,0}\}$ as a set of first level hierarchical images, denoted by the index term k ranging from 0 to $n_1^{0,0}$.
6. The corresponding grey level image components are read from the original image $H(X)$ and are represented as a set $\{g_{1,k}^{0,0}\}$, where the index term k ranges from 0 to $n_1^{0,0}$.
7. If $n_1^{0,0} > 1$ go to step 8 else go to step 10
8. For $k = 0$ to $n_1^{p,q} - 1$ do
 - a. Estimate the local threshold $t_{l,k}^{p,q}$ for binarization of the grey image component $g_{l,k}^{p,q}$.
 - b. Weights are assigned to each level dynamically depending on the current level of the object using the logic that the weight of the previous level is half of that of the current level as implemented below:
 - i. If $l=0$ (i.e. for the starting or 0th level)
 1. $w_0=1.0$
 - ii. else if $l=1$ (i.e. for the 1st level)
 1. $w_0=0.5$
 2. $w_1=0.5$
 - iii. else (i.e. for the rest of the levels)
 1. $w_l=0.5$ (i.e. the current level's weight)
 2. For $i=l-1$ to 2
 - a. $w_i=w_{i+1}/2$ (i.e. the previous level's weight)
 3. End for
 4. $w_1=w_2/2$ (i.e. the 1st level's weight)
 5. $w_0=w_2/2$ (i.e. the 0th level's weight)
 - iv. End if

- c. The cumulative threshold for binarization is calculated as a weighted average of all the local thresholds for binarization as

$$T_{l,k}^{p,q} = \left(w_0 t_{0,0}^{0,0} + \sum_{i=1}^l w_i t_{i,k}^{i-1,q} \right) / (l+1)$$

- d. Using the threshold $T_{l,k}^{p,q}$ the image is binarized generating the l^{th} level binarized image object $b_{l,k}^{p,q}$.
- e. Apply connected component analysis algorithm on the object $b_{l,k}^{p,q}$ to generate the $(l+1)^{\text{th}}$ level binarized connected components $\{b_{l+1,r}^{l,k}\}$ as a set of $(l+1)^{\text{th}}$ level hierarchical images, denoted by the index term r ranging from 0 to $n_{l+1}^{l,k}$.
- f. The corresponding grey level image components are read from the original image $H(X)$ and are represented as a set $\{g_{l+1,r}^{l,k}\}$, where the index term r ranges from 0 to $n_{l+1}^{l,k}$.
- g. If $n_{l+1}^{l,k} > 1$ (i.e. if the no. of objects is more than 1 then go for depth first search)
 - i. $k_stack[top]=k$
 - ii. $l_stack[top++]=l$
 - iii. $k=0$
 - iv. $l++$
- h. Else (label the component and trace back one level up)
 - i. Mark $b_{l,k}^{p,q}$ with current label index L
 - ii. $L++$
 - iii. $k=k_stack[top]$
 - iv. $l=l_stack[top--]$
 - v. $p=l-1$
 - vi. $q=k$;
 - i. End if
9. End for
10. End

2.3 Brief explanation of the algorithm

Read Image

The images used are obtained from CCTV camera or mobile camera. The images are 24-bit color images and are stored in bmp format. Reading the header of the image

file the total image is read and is stored in an array $img(i,j)$. The array stores the color value of the image at location (i,j) . The image is then converted to grey level image using standard conversion rule [12] stated below.

$$grey(i, j) = 0.59 \bullet red(i, j) + 0.30 \bullet green(i, j) + 0.11 \bullet blue(i, j)$$

where, $red(i,j)$, $green(i,j)$ and $blue(i,j)$ indicate the red, green and blue components of the color of the pixel at position (i,j) in the image and $grey(i,j)$ is the resultant grey value of the pixel at position (i,j) .

Binarization using OTSU's algorithm

Out of the different methods [13], OTSU's algorithm [14] is selected for the purpose of segmentation as it provides good binarization in case of noisy image also. An automatic threshold selection method for image segmentation is presented by Nobuyuki Otsu who proposed an algorithm for threshold selection.

Let the pixels of a given image be represented in L gray levels $[1, 2, \dots, L]$. The number of pixels in gray level i is denoted by n_i and the total number of pixels can be expressed as $N = n_1 + n_2 + \dots + n_L$. Then suppose that the pixels were dichotomized into two classes C_0 and C_1 , which denotes pixels with levels $[1, \dots, k]$ and $[k+1, \dots, L]$, respectively. This method is based on a discriminant criterion which is the ratio of between-class variance and total variance of gray levels. The optimal threshold of an image maximizes between-class variance which in turn maximizes the separability of the resultant classes in gray levels.

Connected Component Labeling

In this section the connectivity of the foreground pixels are checked. For any black pixel (considering black pixel as foreground pixel) it is found whether it has any black pixel as a neighbor of it. In this section, 8-neighbors are used for the purpose of finding the neighbors. And if one or more neighbor(s) of the center pixel is (are) found as black then for each of them the 8-neighbors are found. This process is done recursively to get all the connected components within the image. Each of the connected components is labeled for future identification.

Object level segmentation

Each of the connected components is termed as segment and all these segments (considered as objects) are isolated from each other. The actual image data corresponding to each segment are picked up from the original image for further processing. Now due to the poor capturing of the image these segments may contain connected objects as a single segment. To solve this problem here segmentation is done in a hierarchical manner. At the first level of hierarchy the whole image is binarized using a single threshold and is segmented using CCL algorithm. If the number of segments at this level is greater than one then it is passed to the second level of hierarchy where each of the segments is taken as input. Each segment is then binarized using average cumulative threshold combining

the local (to the segment) as well as global threshold and the segmentation processes are executed locally to get new sub-segments. If any new sub-segment is generated at this level then all the sub-segments are passed to the third level of hierarchy and the process continues. Intermediately, at any level and for any segment if the number of sub-segments is unity then the segment is labeled with the current label index and the rest of the process continues for other segments at that level and for the remaining levels.

3. Experimental Result Analysis and Conclusion

The segmentation algorithm is applied over a large number of dataset. To test the performance of the algorithm the dataset published in DOBCO-2009 and H-DIBCO-2010 is used as the benchmark. Over and above, to introduce the variety in the dataset, a large number of business card images, printed text document images and handwritten document images all captured from a camera are also used. The technique is applied over 110 printed text document images, 45 handwritten document images, 100 business cards and 67 natural images containing text and the result appears to be quite satisfactory.

Fig. 1. shows some of the test images on which the algorithm is applied upon. The source images constitute of three categories of images: business card images, degraded printed document images and degraded handwritten document images. The images are impaired by shadow or poor printing quality or printing impression from back side of the page.

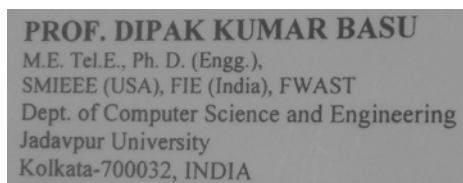
Fig. 2. shows the experimental result when the technique is applied on the test images. First column of the table represents the output after first level of segmentation and the second column of the table represents the final output of segmentation. The visual analysis of the result can summarized as below:

Fig. 2 (a, b): The test image shown in Fig. 1 (a) is a clear printed text business card image. Fig. 2 (a) shows that the discrete characters are all segmented at this level. The "rs" and "ty" of "University" remain connected at this level. Fig. 2 (b) shows the final result of segmentation. The algorithm dissociates the segment "rs" to 'r' and 's' but failed to dissociate the segment "ty" to 't' and 'y'.

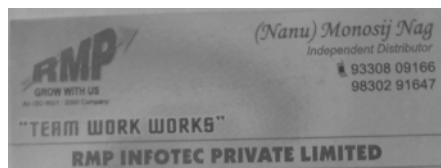
Fig. 2 (c, d): The test image shown in Fig. 1 (b) is a shaded printed text business card image. Fig. 2 (c) shows that most of the discrete characters are segmented at this level. The "OW" of "GROW", the "WP" and "TH" of "WITH", the 'U' and 'S' of "US" remain connected at this level. Fig. 2 (d) shows the final segmentation result. The algorithm could dissociate "WP", "TH" and "US" but could not dissociate "OW" of the segment "GROW".

Fig. 2 (e, f): The test image Fig. 1 (c) is a degraded printed text document image. Fig. 2 (e) shows that the discrete characters are all segmented at this level only with a two exceptions. The “in” of “feiner” and “ft” of the word “ptftel” remain connected at this level. Fig. 2 (f) shows the final result of segmentation. The algorithm could dissociate ‘i’ and ‘n’ of the segment “in” of the word

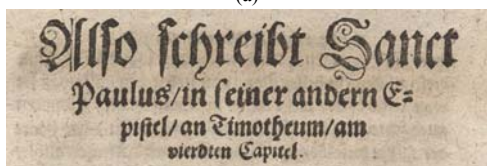
“feiner” properly though the segment “ft” could not be dissociated into ‘f’ and ‘t’. Similarly, the handwritten images Fig. 1 (d-g) are segmented by our algorithm and the results are represented in Fig. 2 (g-n) sequentially in pair representing first and last stage outputs respectively in each row. In most of the cases the algorithm has dissociated the texts successfully.



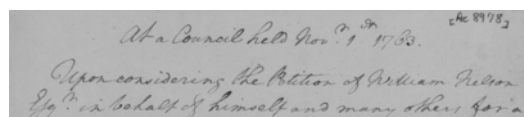
(a)



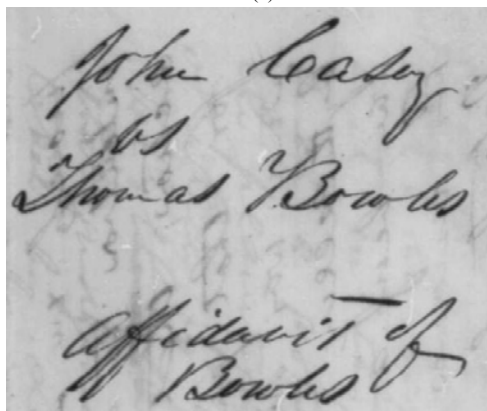
(b)



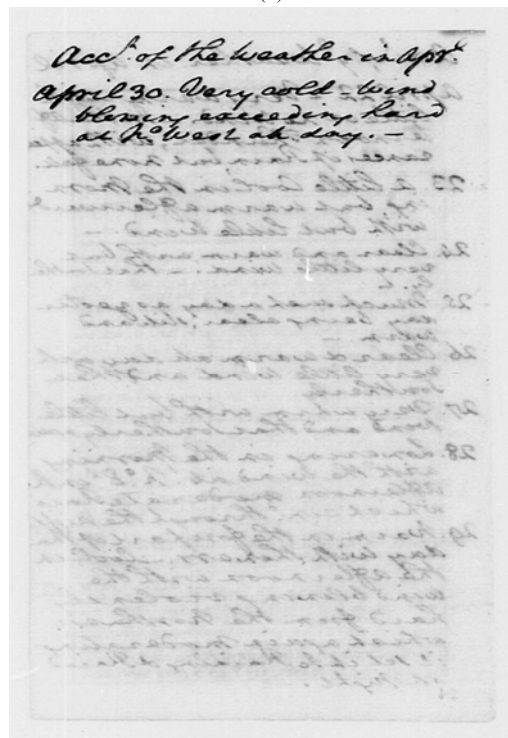
(c)



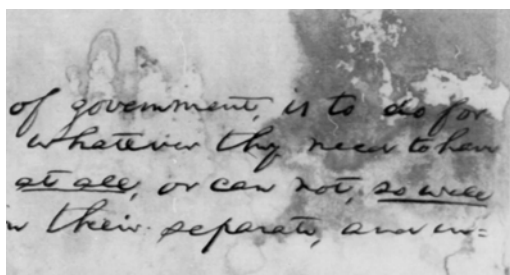
(d)



(e)



(g)



(f)


Fig. 1. Some test images used for segmentation using the proposed algorithm (a), (b): business card images, (c) printed document image, (d)-(g) handwritten document images

PROF. DIPAK KUMAR BASU
M.E. M.A.E., Ph. D. (Engg.),
SMIEEE (USA), FIE (India), FAWASIT
Dept. of Computer Science and Engineering
Jadavpur University
Kolkata-700032, INDIA

(a)

PROF. DIPAK KUMAR BASU
M.E. M.A.E., Ph. D. (Engg.),
SMIEEE (USA), FIE (India), FAWASIT
Dept. of Computer Science and Engineering
Jadavpur University
Kolkata-700032, INDIA

(b)

 [New] Monasij [New]
Independent Distributor
93308 91635
93302 91647
"TEAM WORK WORKS"
RMP INFOTEC PRIVATE LIMITED

(c)

 [New] Monasij [New]
Independent Distributor
93308 91635
93302 91647
"TEAM WORK WORKS"
RMP INFOTEC PRIVATE LIMITED

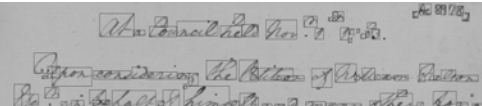
(d)

Also schreibt Sanct
Paulus/in seiner andern E=
pistel/an Timotheum/am
vierten Capitel.

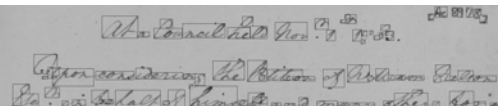
(e)

Also schreibt Sanct
Paulus/in seiner andern E=
pistel/an Timotheum/am
vierten Capitel.

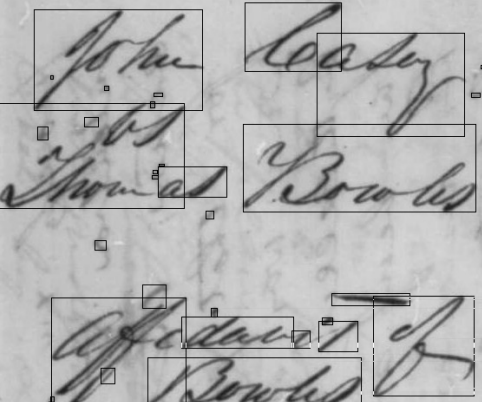
(f)


Upon considering the Petition of William Bowers
for a writ of Habeas Corpus and every other for a

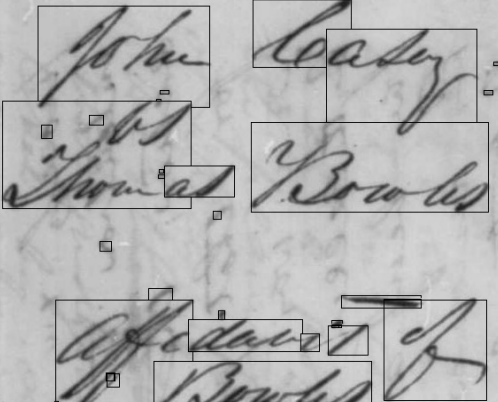
(g)


Upon considering the Petition of William Bowers
for a writ of Habeas Corpus and every other for a

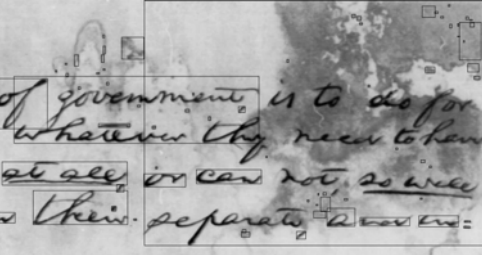
(h)


John Casey
Thomas Bowers
Affidavit of
Bowers

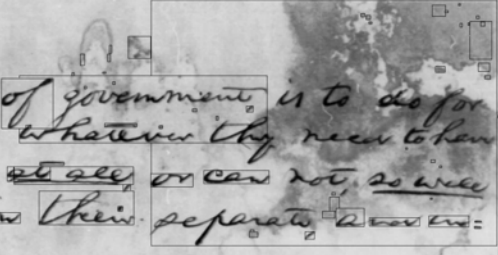
(i)


John Casey
Thomas Bowers
Affidavit of
Bowers

(j)


of government is to do for
whatever they need to have
at all, or can not so well
their separate

(k)


of government is to do for
whatever they need to have
at all, or can not so well
their separate

(l)

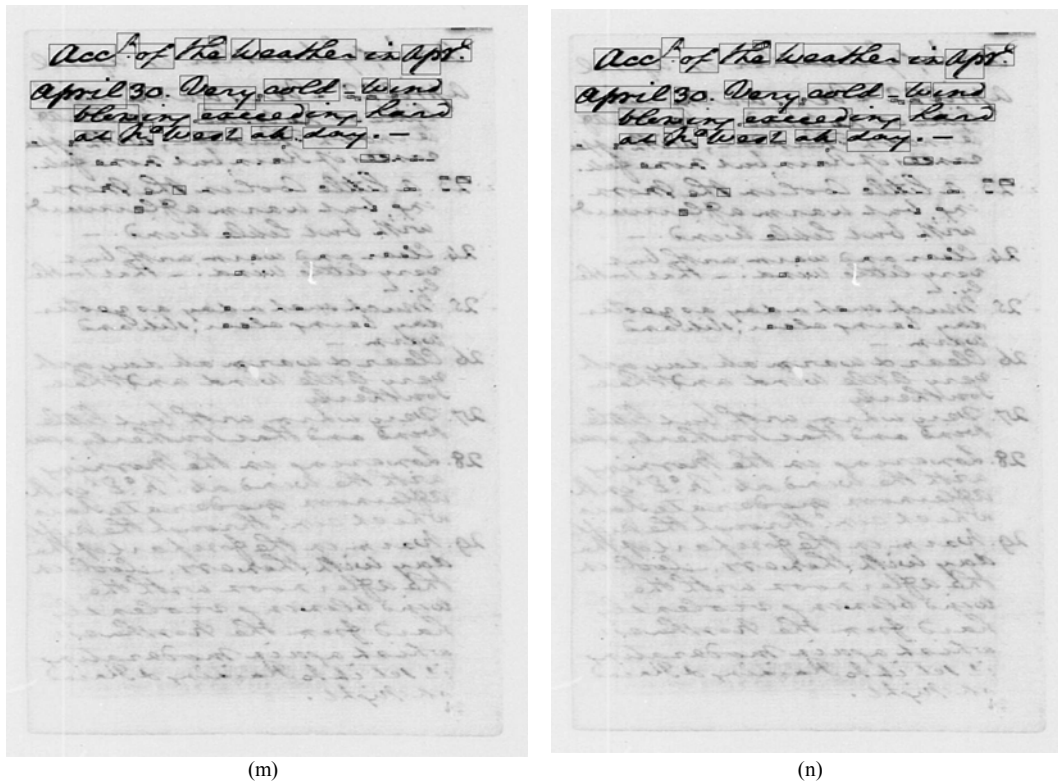


Fig. 2. Segmentation result for business card and document images

3.1 Comparing output segments with ground truth segments

The ground truth segments corresponding to the source image contains the actual or correct segment information. For comparing the result of segmentation the ground truth segments and the resulted segments are counted manually. These two data are compared according to segment-by-segment basis and is used to get the result of matching. The matching-mismatching can be classified into three categories of segments:

- (a) True positive (tp), in which the corresponding segments in both the images are ON (marked).
- (b) False positive (fp), in which the segment in output image is ON (marked) but the corresponding segment in ground truth image is OFF (unmarked).
- (c) False negative (fn), in which the segment in output image is OFF (unmarked) but the corresponding segment in ground truth image is ON (marked).

Both the images are compared and number of different categories of segments, viz. tp , fp and fn , are counted according to the three cases respectively.

3.2 Computing recall, precision and f-measure

For quantitative measure of the segmentation result the following parameters are computed:

Recall (rc) is the fraction of ON segments in the ground truth image that are ON in the output image also.

Precision (pr) is the fraction of ON segments in the output image that are ON also in the ground truth image.

F-measure (fm) is a quantity combining the recall and precision, giving the result of segmentation.

$$rc = \frac{tp}{fn + tp}$$

$$pr = \frac{tp}{fp + tp}$$

$$fm = \frac{2.rc.pr}{rc + pr} \times 100\%$$

The existing segmentation algorithms are based on spatial level variation of color or gray value. These algorithms result in a number of segments with similar color or gray value. As a result two or more objects with similar color or gray value may reside in one segment. In

the proposed approach, the algorithm segments the image at the object level by binarizing it using the hierarchical threshold selection method. Two characters connected by similar but undesired link in between are segmented into separate sub-segments at the second or third level of the proposed segmentation process. The result is shown in Table 1.

Table 1: Result for different categories of images

Type of image	fm	Overall fm
Printed document image	95.6%	91.6%
Handwritten document image	89.2%	
Business card image	92.7%	
Natural image containing text	88.0%	

Acknowledgment

Authors are thankful to the CMATER and the SRUVM project, C.S.E. Department, Jadavpur University, for providing necessary infrastructural facilities during the progress of the work. Mr. S. Saha, is thankful to the authorities of MCKV Institute of Engineering for kindly permitting him to carry on the research work.

References

[1] A. M. Nambodiri, and A. K. Jain, "Online Handwritten Script Recognition", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 26, No. 1, January 2004, pp. 124-130.

[2] S. Basu, C. Chaudhuri, M. Kundu, M. Nasipuri, and D. K. Basu, "Text line extraction from multi-skewed handwritten documents", Pattern Recognition, Vol. 40, No. 6, June 2007, pp. 1825-1839.

[3] Khandelwal et. al., "Text Line Segmentation for Unconstrained Handwritten Document Images Using Neighborhood Connected Component Analysis", S. Chaudhury et al., eds., PReMI 2009, Berlin: Springer-Verlag, pp. 369-374.

[4] R. G. Casey, and E. Reclinet, "A Survey of Methods and Strategies in Character Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 7, July 1996, pp. 690-706.

[5] K. Haris, and S. N. Efstratiadis, "Hybrid Image Segmentation Using Watersheds and Fast Region Merging", IEEE Transactions on Image Processing, Vol. 7, No. 12, December 1998, pp. 1684-1698.

[6] T. N. Pappas, "An Adaptive Clustering Algorithm for Image Segmentation", IEEE Transactions on Signal Processing, Vol. 40, No. 4, Apr 1992, pp. 901-914.

[7] P. Salembier, and F. Marques, "Region-based representations of image and video: Segmentation tools for multimedia services", IEEE Transactions on Circuits and Systems for

Video Technology, Vol. 9, No. 8, December 1999, pp. 1147-1167.

[8] C. A. Bouman, and M. Shapiro, "A Multiscale Random Field Model for Bayesian Image Segmentation", IEEE Transactions on Image Processing, Vol. 3, No. 2, March 1994, pp. 162-177.

[9] R. Lienhart, and A. Wernicke, "Localizing and Segmenting Text in Images and Videos", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No. 5, April 2002, pp. 256-268.

[10] S. P. Chowdhury, S. Dhar, K. Rafferty, A. K. Das, and B. Chanda, "Robust Extraction of Text from Camera Images using Colour and Spatial Information Simultaneously", Journal of Universal Computer Science, Vol. 15, No. 18, 2009, pp. 3325-3342.

[11] S. Saha, S. Basu, M. Nasipuri, and D. K. Basu, "A Hough Transform based Technique for Text Segmentation", Journal of Computing, ISSN: 2151-9617, Vol. 2, Issue 2, February 2010, pp.134-141.

[12] R. C. Gonzalez, and R. E. Woods, Digital Image Processing, Second Edition, Pearson Education Asia, 2002.

[13] M. Nandy (Pal) and S. Saha, "An Analytical Study of Different Document Image Binarization Methods", Proceedings of IEEE National Conference on Computing and Communication Systems (COCOSYS-09), UIT, Burdwan, January 02-04, 2009, pp. 71-76.

[14] N. Otsu, "A threshold selection method from graylevel histogram", IEEE Trans. on System, Man, Cybernetics, Vol. 9, No. 1, 1979, pp. 62-69.

Satadal Saha has done B. Sc. (Physics Hon's), completed B. Tech. in Applied Physics and M. Tech. in Optics and Optoelectronics from University of Calcutta in 1995, 1998 and 2000 respectively. In 2000, he joined as a Project Fellow in the Department of CST in BESU (formerly B. E. College), Howrah. In 2001, he joined as a Lecturer in the Department of Information Technology, Govt. College of Engg. and Textile Technology (formerly known as College of Textile Technology), Serampore. In 2004, he joined as a Lecturer in the Department of Computer Science and Engg., MCKV Institute of Engg, Howrah and he is continuing his service there as an Assistant Professor in the Department of Electronics and Communication. He has authored a book titled Computer Network published by Dhanpat Rai and Co. Ltd., New Delhi in 2008. His research areas of interest are image processing and pattern recognition. Mr. Saha is a member of IEEE, IETE and CSI. He was also a member of the executive committee of IETE, Kolkata for the session 2006-08.

Subhadip Basu received his B.E. degree in Computer Science and Engineering from Kuvempu University, Karnataka, India, in 1999. He received his Ph.D. (Engg.) degree thereafter from Jadavpur University (J.U.) in 2006. He joined J.U. as a senior lecturer in 2006 and he is continuing his service there as an Assistant Professor. Currently, he is the visiting Scientist (Under BOYSCAST Fellowship from Govt. of India) at Department of ECE, The University of Iowa, USA. His areas of current research interest are OCR of handwritten text, gesture recognition, real-time image processing.

Mita Nasipuri is a senior member of the IEEE, U.S.A., Fellow of I.E (India) and W.B.A.S.T, Kolkata, India.