

Automatic Detection and Inpainting of Text Images

S.Bhuvaneshwari

Department of Computer Science and Engineering
Annamalai University
Annamalainagar
Tamilnadu, India

T.S.Subashini

Department of Computer Science and Engineering
Annamalai University
Annamalainagar
Tamilnadu, India

ABSTRACT

The proposed system detects text using connect component labelling and a set of selection/ rejection criteria which helps to retain the text region alone. The detected text region is then inpainted using fast marching algorithm which uses the pixel information that is present in the non-text region of the image for inpainting the detected text region. This work is done in two steps. The first step detects the text region from the image without the user manually marking it and in the second step the text is de-occluded from the image using the existing fast marching inpainting algorithm.

Keywords

Niblack's algorithm, CCL, selection criteria, inpainting mask, fast marching algorithm

1. INTRODUCTION

Sometimes an image may contain text embedded on to it, which should be removed for aesthetic and other reasons. Further, detection and recognition of this text can be important in certain applications like OCR etc. In this work, the text in image is detected and fast marching image inpainting [1] is applied to fill the detected text region with the background image data, thereby eliminating the text regions. The term inpainting [2] refers to automatic filling of a specific region in an image called a mask. Normally the mask is marked manually, which will be filled based on the surrounding image region.

There are many applications of image inpainting ranging from restoration of photographs, films, removal of occlusions such as text, subtitle, logos, stamps, scratches, red eye removal etc.

The proposed work is done in two stages.

- 1.Detection of text
- 2.Inpainting



a) Original image



b) Manually selected mask to be inpainted



c) Inpainted image

Fig 1: Inpainting process

The common requirement for all image inpainting algorithms is that the region to be inpainted should be manually selected by the user.

As a first step the user manually selects the portions of the image that will be restored. This is usually done as a separate step and involves the use of other image processing tools. Then image restoration is done automatically, by filling these regions with the information from the surrounding pixels (or) from the whole image. It can be seen from Figure 1 how the lizard in the original image has been inpainted [3] after manually selecting the mask region.

The rest of the paper is organized as follows. Section 2 describes the related work in this area. Section 3 presents the methodology of this work. Section 4 shows the experimental results. Section 5 concludes the paper.

2. RELATED WORK

This section gives an overview of the work that has been carried out in the area of automatic text detection. The text retains an image is of different size, font style, colour, orientation etc., and so, automatic extraction of text is very challenging.

The work in [4] and [5] uses similarity constraints and language statistics for text detection. The authors in [6] compared the edge based, and connected component based approaches and analysed the advantages and disadvantages of these systems. The authors in [7] used edge detection and connected component labelling to extract the text effectively. The work in [8] detects and extracts the text regions from a complex scene, using connected component approach. The authors in [9] propose an algorithm for selecting text regions and for region segmentation using horizontal projection and geometric properties. The work in [10] proposes a new filtering technique in which the image is filtered using connected operator and the selected regions are estimated using inpainting. The authors in [11] compared and analysed the image inpainting techniques such as Exemplar, Poisson, successive elimination and 8-pixel neighbourhood methods. The results showed that Poisson method is the best for filling two dimensional objects.

The work in [12] proposed a method which inpaints the object based on the non-linear diffusion tensor. The authors in [13] proposed an algorithm to inpaint an image using sparse representation based iterative algorithm. The results prove that it recovers different structural components in the image. The authors in [14] used the texture property to identify the text using SVM. The work in [15] introduces new local thresholding technique "NICK". Nick has been inspired from the Niblack's binarization method and exhibits its robustness and effectiveness when evaluated on low quality ancient document images. The work in [16] automatically identifies and recognizes text zones in images taken from a camera and this system can be used on a

mobile environment by blind or visually impaired users. The work in [17] studied the use of level set based Bresson's method, for image binarization and segmentation of Hindi character images. Authors in [18] proposed a graph model using Viterbi algorithm to extract the most probable sequence of characters present in the scene image.

In this work automatic text detection is done for which the image is first segmented using CCL and then a set of selection / rejection criteria are applied to eliminate non-text regions from the image.

3. METHODOLOGY

This work is done in two steps. The first step detects the region of text from the image without the user manually marking it. In the second step the text detected in the previous step is used as a mask for inpainting the image using the existing fast marching inpainting algorithm [19]. The flow of the proposed work is shown in Figure 2.

1. The image background and foreground information is manipulated to have a dark text region against the light background.
2. The resultant image is resized to 400x400 to reduce unnecessary computation.
3. The resized image is binarized using Niblack algorithm [20] which takes the local mean and standard deviation to find out the threshold for binarizing the image.
4. Connected component labelling (CCL) scans an image and groups its pixels into components based on pixel connectivity, which means all pixels in a connected component share similar pixel intensity values and are in some way connected with each other. Once all groups have been determined, each pixel is labelled with a graylevel or a colour according to the component it was assigned to [1]. So, CCL is applied to obtain the various connected region and the connected regions are bounded using bounding boxes.
5. After applying CCL, the first set of criteria is applied which eliminate non-text macro objects. Here all objects whose area is greater than 3000 pixels and perimeter greater than 2000 pixels are removed. Still smaller objects of non-text regions are left behind.
6. To eliminate those micro non-text regions, CCL is applied to the resultant image once again and the second set of criteria is applied which eliminate non text micro objects. Here all objects whose area is less than 2000 and perimeter less than 1000 are removed. One more criteria namely major axis length is used which is used to retain the text region alone. All objects whose major axis length is between 20 and 95 are considered to be text. This major axis length is found out empirically and this length is almost appropriate for all types of text
7. Then the resultant image is eroded and then dilated to remove the undesired small objects and this image is the binarized text image.
8. Now the image contains only the text regions with the non-text regions fully removed. The original image is now compared with the binarized text image. The pixels locations which are white in colour in the binarized image are replaced with blue colour in the pixel location in the original image which forms the inpainting mask.

9. The image is now inpainted using fast marching algorithm [11], which is a technique for producing distance maps of the points, in a region from the boundary of the region. This method is used to paint the points inside the boundary, according to the increasing distance from the boundary of the region.

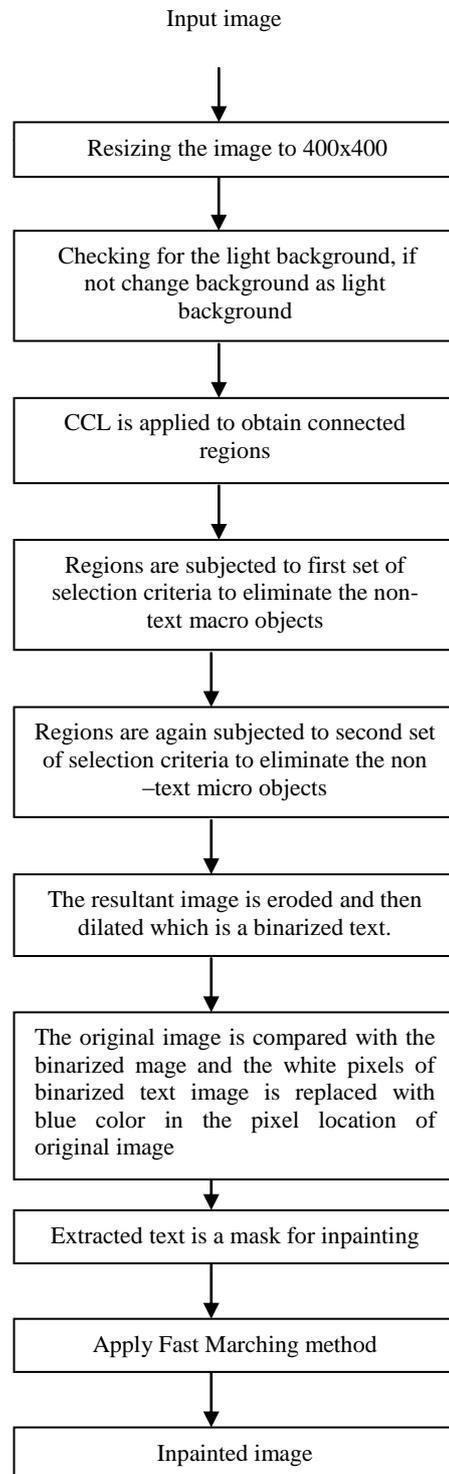


Fig 2: The methodology of the proposed work

4. EXPERIMENTAL RESULTS

The images with simple and complex background are given as input to the proposed system and the outcomes are shown in Figure 3, Figure 4 and Figure 5.

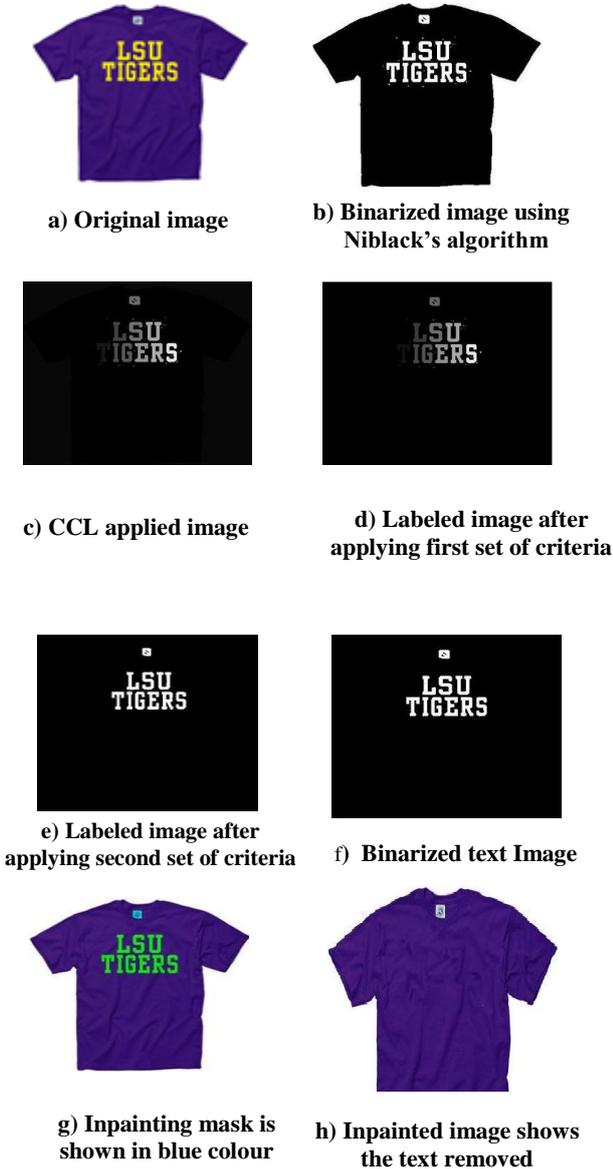


Fig 3: Automatic detection and inpainting of text

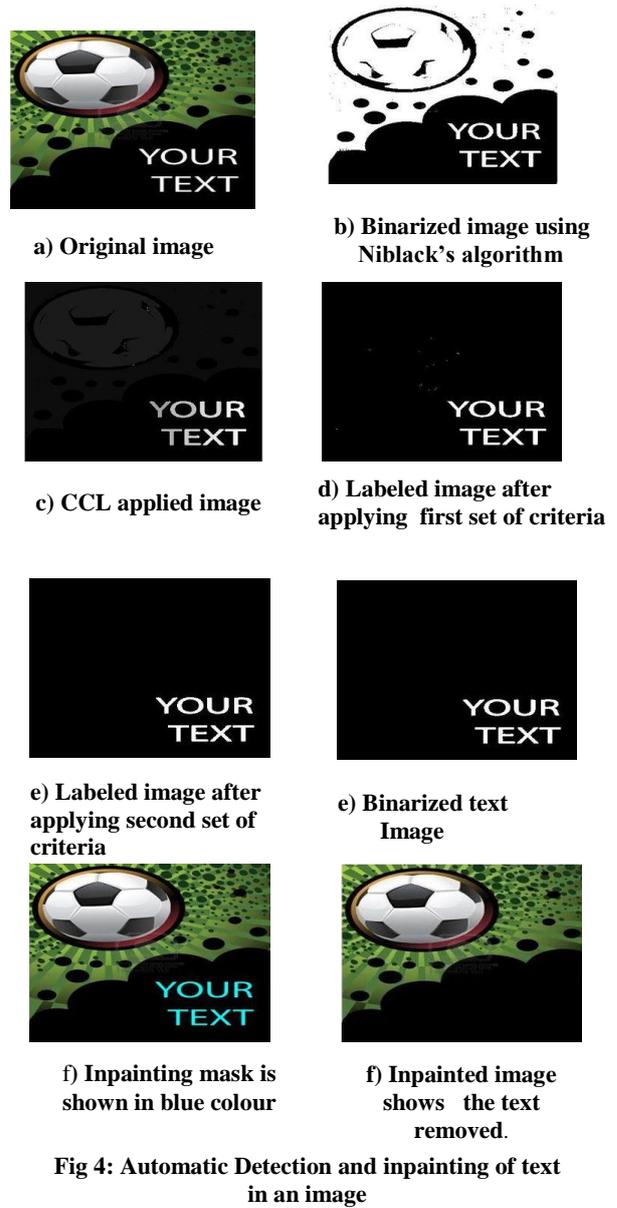


Fig 4: Automatic Detection and inpainting of text in an image

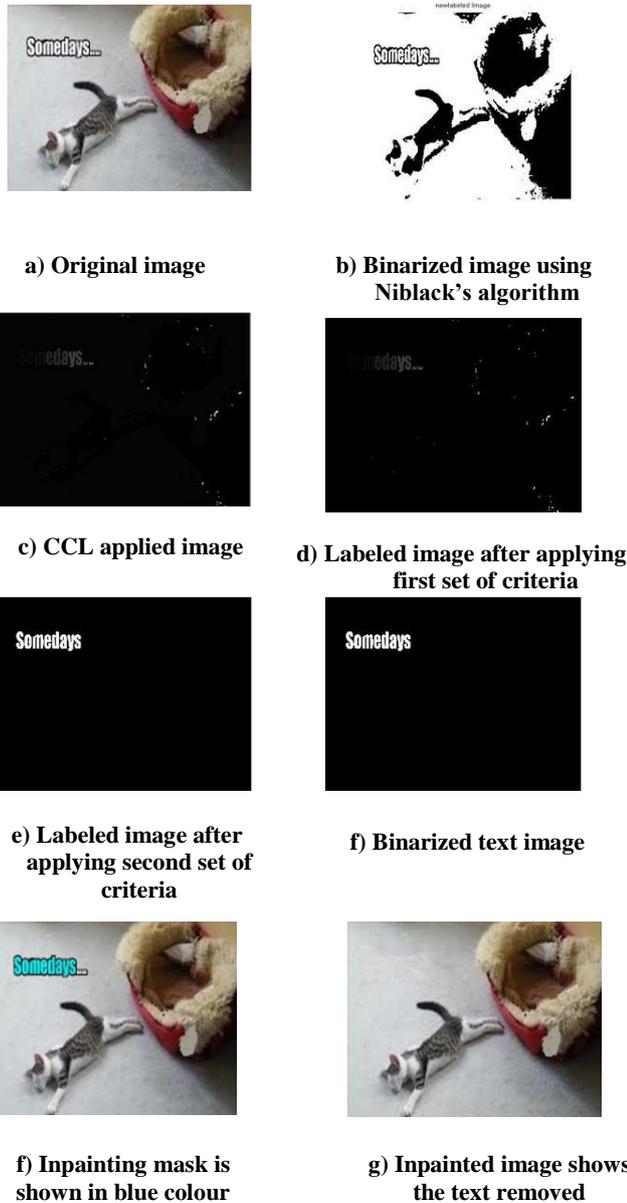


Fig 5: Automatic detection and inpainting of text in an image

Table 1 shows the efficacy of the proposed work in correctly detecting text in images.

Accuracy is given by the number of characters which are correctly detected divided by the total number of text characters present in the image. From Table 1 it can be seen that the proposed method is efficient when font style is almost same. However when the font style is different as seen in the last row of the Table 1, the efficiency is greatly reduced.

5. CONCLUSIONS

In this paper, a method to detect text automatically in colour images is presented. The detected text region is the target region for inpainting.

Niblack's algorithm was applied to binarize the image; CCL was carried out to detect the text. The text detected is used as

Table 1. Experiments and its results

Original Image	Image with mask	Accuracy
		100%
		100%
		100%
		59.09%

a mask and it is de-occluded using fast marching algorithm. Experimental results have demonstrated that the proposed method can be effectively used to detect the text automatically. On the other hand, if the text is of different font style the efficiency comes down. The proposed method works well for both simple and complex images and it can be used to create automatic text masks for the subsequent inpainting process.

6. REFERENCES

- [1] R.C.Gonzalez and Redwoods, Digital Image Processing, 2nd ed., Pearson Education, 2002.
- [2] A.Criminisi, P.Perez, K.Toyama, "Region Filling and object removal by exemplar based image inpainting", IEEE transactions on image processing, Vol.13, No.9, pp.1-7, 2004.
- [3] S.Bhuvaneshwari, T.S.Subashini, S.Soundharya, V.Ramalingam "A novel and fast exemplar based approach for filling portions in an image", IEEE Proceedings on the International conference on recent trends in information technology (ICRTIT), pp.91-96, 2012.
- [4] D. L. Smith, J. Field, and E. G. Learned-Miller, "Enforcing similarity constraints with integer programming for better scene text recognition", CVPR, pp.73-80, 2011.
- [5] J. J. Weinman, E. G. Learned-Miller, and A. R. Hanson, "Scene text recognition using similarity and a lexicon with sparse belief propagation", PAMI, pp.1-11, 2009.
- [6] J.Sushma, M.Padmaja, "Text Detection in color images", IEEE Proceedings on the International conference on intelligent agent & multi agent systems (IAMA), pp.1-6, 2009.

- [7] Nobuo Ezaki, Marius Bulacu, Lambert Schomaker, “Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons”, IEEE Proceedings on the 17th International conference on ICPR, pp.683-686, 2004.
- [8] J.Ohya,A.Shio,S.Akamatsu, “Recognizing Characters in Scene Images”, IEEE transactions on PAMI, Vol.16, No. 2, pp. 214-224, 1994.
- [9] RodolfoP.DosSantos, S.Gabriela, Clemente, TsangIng Ren, George D.C.Calvalcanti, “Text Line segmentation based on morphology and histogram projection”, IEEE Proceedings on the 10th International conference on document analysis and recognition (ICDAR), pp.651-655, 2009.
- [10] Mariella Dimiccoli, Philippe Salembier, “Perceptual filtering with connected operators and image inpainting”, ISMM Proceedings on the 8th International symposium on mathematical morphology, pp. 227-238, 2007.
- [11] Muthukumar.S, Dr.Krishnan.N, Pasupathi.P.Deepa.S, “ Analysis of Image Inpainting techniques with Exemplar, Poisson, Successive Elimination and 8 pixel Neighborhood methods”, International Journal of Computer Applications(0975 – 8887), Vol. 9, November 2010.
- [12] Faouzi Benzarti, Hamid Amiri, “Repairing and Inpainting Damaged Images using Diffusion Tensor”, International Journal of Computer Science (IJCSI), Vol. 9, July 2012.
- [13] M.J.Fadivi,J.LStarck and F.Murtagh, “Inpainting and Zooming using sparse Representation”, The Computer Journal, pp.64-79, 2007.
- [14] Qixiang Ye, Wen Gao, Wiquiang Wang, WeiZeng, “A robust text detection algorithm in images and video frame”, IEEE Proceedings on the International conference on information, communication and signal processing (ICICS), Vol.2, pp.930-932, 2012.
- [15] Khurram Khurshd, Imran Siddiqi, Claudie Faure, Nicole Vincent, “Comparison of Niblack inspired Binarization methods for ancient documents”, SPIE-ISAT proceedings on the International conference on Document Recognition and Retrieval, 2009.
- [16] Celine Thillou, Silvio Ferreira, Bernard Gosselin, “An Embedded Application for Degraded Text Recognition”, Vol.13, pp.2127-2135, 2005.
- [17] Manjusha.K, Sachin Kumar.S, Jolly Rajendran, K.P.Soman, “Hindi Character Segmentation in Document Images using Level set methods and Non-linear Diffusion”, International Journal of Computer Applications(IJCA), Vol.44-No.16, April 2012.
- [18] Khaoula Elagouni, Christopic Garcia, Franck Mamalet, Pascale Sebillot , “Combining Multiscale Character Recognition and Linguistic Knowledge for natural Scene Text OCR”, IAPR International Workshop on Document Analysis Systems, pp.120-124, 2012.
- [19] Graham Leedham, Chen Yan Kalyan Takru, Joie Hadi Nata Tan and Li Mian, “Comparison of Some Thresholding algorithms for text/Background Segmentation in difficult document images”, IEEE Proceedings on the Seventh international Conference on Document Analysis and Recognition (ICDAR), pp.859-864, 2003.
- [20] Alexandru Telea, “An image Inpainting Technique Based on the Fast Marching Algorithm”, Journal of Graphics tools, Vol.9, No.1, pp.25-38, 2004.