Text Localization and Extraction from Still Images using Fast Bounding Box Algorithm

SONAL PALIWAL, RAJESH SHYAM SINGH and H.L. MANDORIA

Department of Information Technology, G. B. Pant University of Agriculture & Technology, Pantnagar, Uttarakhand, India.

ABSTRACT

Text Extraction and Localization from images is a very challenging task because of noise, blurriness and complex color background of the images. Digital Images are subjected to blurring due to many hardware limitations such as atmospheric disturbance, device noise and poor focus quality. In order to remove textual information from images it is necessary to remove blurriness and restore the image for the text extraction. Thus in this paper Fast Bounding Box algorithm is applied for localization and extraction of the text from images in an efficient manner by dividing the image into two halves and then find the dissimilar region i.e. text.

Keywords: Fast Bounding Box, ROI, Bhattacharya coefficient, Precision, Recall and F-measure

INTODUCTION

Digital Image Processing

Digital image processing is rapidly expanding its root in very field and is an important emerging research area. It is developing gradually with the growing applications in Information Technology and computer science engineering. The term digital image processing usually deals with alteration of digital images with the help of digital computer. Signals and systems include image processing as one of its subfield but it particularly focuses on images. It's a technology that applies a wide range of computer algorithm to process the digital image by apply different operations on an image, in order to get an intensified image or to extract some useful information from it. In image processing image is taken as input and output may also be an image or characteristics/features inked with that image.

DIP focuses on two vital tasks

Upgradation of illustrated information for human interpretation. The data inside the image is been processed for transmission, representation and storage for independent machine understanding.

FAST BOUNDING BOX ALGORITHM

Fast Bounding Box (FBB) technique was mainly proposed to quickly detect brain tumor by forming a parallel-axis rectangle throughout the tumor area which essentially finds the most distinct region among the tumorous and non-tumors tissues and bound the tumorous cells in a rectangular box called bounding box. Similarly we are using this algorithm here for text localization in still images. This method automatically locates the text within the image. This is basically used to find most dissimilar region i.e. text and we mention it as Region of Interest (ROI) as it is text localization.

The FBB algorithm works in two consecutive steps. First the input set of images or video frames are processed to find axis-parallel rectangles, i.e. the bounding box on each and every input set of images. In this step the image is divided into two segments left and right and then with reference to other part and the dissimilar part are detected. Secondly, these BB are clustered together to identify the ones that surrounds the text inside the image. FBB algorithm always produces an axis-parallel bounding box for any image slice, even if that image does not contain text that is the dissimilar region that is to be detected.

FBB is used to compute the BB around the object within a given local search region. This process of text detection uses a novel score function that is based on Bhattacharya coefficient is calculated with intensity of the gray level histograms. Bhattacharya coefficient (BC) measures resemblance between two normalized grey level intensity histograms. Therefore, when two normalized histograms are similar, the BC between them is 1 and when two normalized histograms are dissimilar, the corresponding BC value is 0.

Bhattacharya coefficient is calculated by the equation as follows

$$\hat{\rho}(y) \equiv \rho\left[\hat{c}_u(y), \hat{r}_u\right] = \sum_{u=1}^{m} \sqrt{\hat{c}_u(y)\hat{r}_u}$$

...(1)

Where $r^>$ and $c^>$ present the histograms of reference object and the candidate object at location y. The main objective is to find accurate position of the object here text. Thus, main purpose of using Bhattacharya coefficient by FBB algorithm is to explore that how well the peak of the function represents the coordinates of the object which is to be searched. Thus FBB is a novel and fast segmentation approach that uses symmetry to enclose an anomaly by a BB within an image.

In this paper, Fast Bounding Box algorithm is used for localization of text in Still Images. Rest paper is organized as follows. Section II describes Fast Bounding Box algorithm Results are analyzed and discussed in Section III. Finally, Conclusion is summarized in Section IV.

RESULTS DISCUSSION

The algorithm use for text localization and extraction is implemented using MALAB software. For experimental purpose we have created our own database which contains images of different formats Results are discussed by taking three input images of different formats (jpg, png and jpeg) and then following performance metrics given below are considered for analysis:

Fig.1

Precision: Proportion of correct positive classification (True Positive) from the cases that are predicted as positive and calculated by given below formula:

$$Precision = \frac{TP}{TP+FP}$$

...(2)

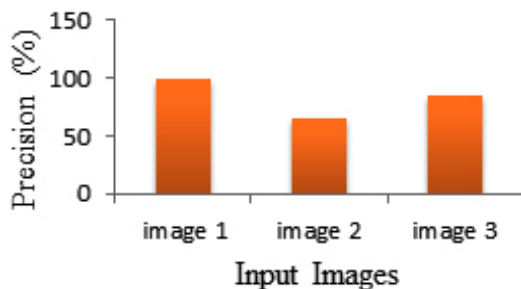

Fig.2

Fig.3

Fig.4

Recall: Proportion of correct positive classification (True Positive) from the cases that are actually positive and calculated by given below formula:

$$Recall = \frac{TP}{TP+FN}$$

...(3)

F-measure: It is a harmonic mean of precision and recall. F-measure is also known as F-factor and calculated by given below formula:

$$F_1 = 2\,\frac{precision \;.\; recall}{precision \;+recall}$$

...(4)

Here, TP means True Positive and FP means False Positive

Input Images taken into consideration are shown below in the figure 1; output after applying FBB and final text localization and detection in the respective images is shown in figure 2 and figure 3 respectively.

Performance Metrics for all the three images are discussed as follows:

Precision

Fig 4 shows Precision for all the three Images respectively. After analyzing and calculating precision for still images it is concluded that precision is in order Image 1,Image 3 and Image 2 where for Image 1 it is best and decreases in the further images.
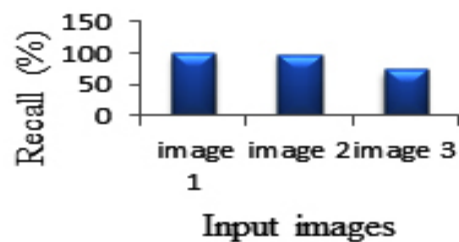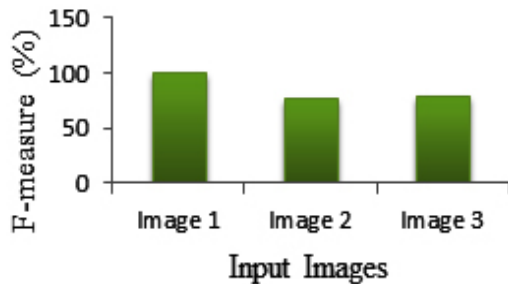

Fig.4

Fig.5

**Table 1: comparison Metrics
For Sample Images**

Test Images	Precision (%)	Recall (%)	F-measure (%)
Image 1	100	100	100
Image 2	65.3	96.9	78.2
Image 3	85.2	78.3	79.11

2 where for Image 1 it is best and decreases in the further images.

Recall

Figure 5 Fig 4 shows Recall for all the three Images respectively. After analyzing and calculating recall for still images it is concluded that recall is in order Image 1, Image 2 and Image 3 where for Image 1 it is best and decreases in the further images.

F-measure

Figure 6 shows F-measure for all the three Images respectively. After analyzing and calculating F-measure for still images it is concluded that F-measure is in order Image 1, Image 3 and Image

CONCLUSION

In this paper fast bounding box algorithm is used for text localization and extraction. Three Images are taken as input of different format and then there precision, recall and F-measure computed. From the obtain results it was clear that the text localized and extracted by using fast bounding box was efficient and accurate. Future work in this method can focus on first remove the noise from the image so that text can be easily localized and extracted and then perform different algorithms based on the quality of the image.

REFERENCES

1.   Amarapur, B. and Patil, N. 2006. Video text extraction from images for character recognition. *Electrical and Computer Engineering Canadian Conference, IEEE,* 198– 201.

2.   Azra, N. and Shobha, G. 2011. Key frame extraction from videos-A survey. *International Journal of Computer Science & Communication Networks. IJCSCN,* 3**(3)**: 194-198.

3   Chen, D., Odobez, J. and Thiran, J. 2004. A localization/verification scheme for finding text   in images and video frames based on contrast independent features and machine learning methods. *Signal Processing, IEEE,* 19: 205–217.

4   Chen, T. 2008. Text localization using DWT fusion algorithm. *International Conference on Communication Technology, IEEE,* 722 – 725.

    Chun, B.T., Bae, Y. and Kim, T. 1999. Text extraction in videos using topographical features of characters. *International Fuzzy Systems Conference Proceedings, IEEE,* 2:1126-1130.

5   Deshmukh, S.P. and Ghongade, R.D. 2015. Detection and segmentation of brain tumor from mri images". *International Journal of Computer Trends and Technology, IJCTT,* 21 :31-33.

6   Krishanamurthy, E.V. and Sen, S.K .2003. Programming in Matlab. *Affiliated East Press Pvt. Ltd.*

7   Huang, X., Ma, H. and Zhang, H. 2009. A new video text extraction approach. *Multimedia and Expo International Conference, IEEE,* 650 – 653.

8   Jain, A.K. and Bin, Y. 1998. Automatic text location in images and video frames. *Pattern Recognition Proceedings of Fourteenth*

International Conference IEEE, 2: 1497-1499.

9   Jain, A.K. 2002. Fundamentas of Digital Image Processing. *Prentice,Hall of India pvt. Ltd.*

10   James, Z.X., Minsoo, S. and Sanjay, R. 1996. Text string location on images. *Proceedings of ICSP, IEEE, 2:* 1354 – 1357.

11   Khodadadi, M. and Behrad, A. 2012. Text localization, extraction and inpainting in color images. *Proceedings of Iranian Conference on Electrical Engineering, ICEE,* 1035-1040.

12.   Kumar, A. 2013. An efficient text extraction algorithm in complex images. *Proceeding of ICE   IEEE,* 6-12.

13   Laurence, L.S., Anahid, H. and Claudie, F. 1995. A hough based algorithm for extracting text lines in handwritten documents. *Document Analysis and Recognition,*

Proceedings of the Third International Conference IEEE, 2: 774 – 777.

14   Lei, S., Xie, G. and Yan, G. 2014. A novel key-frame extraction approach for both video summary and video index. *The Scientific World Journal.* 695168: 09-10.

15   Lempitsky, V., Kohli, P., Rother, C. and Sharp, T. 2009. Image segmentation with a bounding box prior. *Computer Vision 12th International Conference, IEEE,* 277-284.

16   Lienhart, R. and Wernicke, A. 2002. Localizing and segmenting text in images and videos. *Transaction on Circuits and Systems for Video Technology, IEEE,* 12: 256 - 268.

17   Liu, G. and Zhao, J. 2009. Key frame extraction from mpeg video stream. *Proceedings of the Second Symposium International Computer Science and Computational Technology, ISCSCT,* 007-011.