Framework for Improving Accuracy in Text Extraction from Natural Image

Roshi Saxena ** Sushil Bansal #

** Student, Department of Computer Science Chitkara University , M.E (2010-2013),
# Assistant Professor & H.O.D , Deptt. Of CSE, Chitkara University

ABSTRACT

Text embedded in natural scene images contain large amount of useful information. Extracting text from natural scene images is a well known problem in image processing area. Data which appear as text in natural scene images may differ from each other in its size, style, font, orientation, contrast, background which makes it an extremely challenging task to extract the information with higher accuracy. Natural scene text extraction with higher accuracy is still a challenging problem. In this paper r we have presented an algorithm, a graphical user interface to extract the text in scene images with higher precision rate and recall rate.

KEYWORDS:

Text, Images, Natural, Accurate

1. INTRODUCTION

Today, most of the useful information is available into the text which is present into the natural images. For eg. Name of the brand embedded into clothes, text written on the nameplates, signboards etc. There should be some mechanism to extract the text from natural images. Recent studies show some methods to extract text from images but the approach didn't worked fine for characters which are small in size. In this paper we have presented an algorithm which will extract the small sized characters and the algorithm works well with the text which is present into the noisy images also.
We have presented a framework which will extract the text from natural images with higher accuracy. Data from ICDAR dataset 2003 is being tested.

2. PREVIOUS WORK

[1] Kim K.C, Byun, H.R ., Song Y.W, Chi, S.Y, Kim, K.K, Chung Y.K presented method that extracts text regions in natural scene images using low-level image features and that verifies the extracted regions through a high-level text stroke feature. Then the two level features are combined hierarchically. The low-level features are color continuity, gray-level variation and color variance. [2] Shivananda V Seeri, Ranjana B Battur, Basavaraj S Sannakashappanavar presented a method to extract characters from natural scene images. Algorithm works well with the medium sized characters. [3] Xiaoqing Liu et al. proposed "Multiscale edge based text extraction from complex images", method which automatically detects and extracts text present in the complex images using the multi scale edge information. This method is robust with respect to the font size, color, orientation and alignment and has good performance of character extraction. [4] Nobuo Ezaki and Marius Bulacu, Lambert

Schomaker presented a text extraction method for blind persons. [5] Xu-Cheng Yin, Xuwang Yin, Kaizhu Huang, Hong-Wei Hao presented robust text detection in natural scene images. A fast and effective pruning algorithm is designed to extract Maximally Stable Extremal Regions (MSERs) as character candidates using the strategy of minimizing regularized variations.[6] Yang, presented the problem of recognizing and translating automatic signatures.

3. PROPOSED ALGORITHM

Text extraction method which is being used in our algorithm is edge based method and reverses edge based method. Our method converts RGB image into HSI plane to extract the exact color of the text from the image. To implement the method we have prepared a graphical user interface which will follow the following steps:

1. Load input image and separate R,G,B Channels
2. Convert RGB image into HSI image
3. Edges are detected by applying Sobel operator on the image in the following ways
   a.) Apply Sobel horizontal Kernel to get the horizontal gradient image.
   b.) Apply Sobel vertical kernel to get vertical gradient image.
   c.) Find magnitude of gradient Image
4. Apply Otsu's thresholding to binarize the image.
5. Apply connected components to find text into the image.
6. Apply filtering to text to remove false objects.
7. Test on the database to calculate the accuracy.

Advantage of our method is that our graphical user interface extracts the small sized characters also with higher accuracy.

3.1 Implementation

3.1.1 Load Input image from the database

First of all, input image is loaded from the database. After loading the image from database, image is separated into R, G,B channels.

3.1.2 Conversion of RGB image into HSI plane

RGB model is used to display the images. No single color can be called as red, green and blue. In RGB Color Model each color appears in its primary spectral components of Red, Green and Blue.

HSI color model is used to process the images. **HUE** is used to extract the dominant color perceived by an observer. **SATURATION** extracts the relative purity of the amount of white mixed with hue. **INTENSITY** is used to check the brightness at different point's i.e. the total amount of light passing through a particular area. Hue, Saturation and Intensity are three important descriptors used in describing colors.

- **Hue** represents the purity of the color. (i.e. pure red, pure yellow, pure green).
- **Saturation** represents the measure of the degree to which a pure color is diluted by white light.
- **Intensity** is the gray level value of the color.

•Hue and Saturation represents the color carrying Chrominance (Chromatic) information. •Intensity represents the gray-level Luminance (achromatic) information.

Converting Color from RGB to HSI: To convert RGB image into HSI, following formula is used

$$H = \begin{cases} \theta & if\ B \le G \\ 360 - \theta & if\ B > G \end{cases}$$

$$\theta = \cos^{-1}\left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\}$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R,G,B)]$$

$$I = \frac{1}{3}(R+G+B)$$

Image being converted into HSI

3.4.3 Edge Detection

Edges characterize boundaries in images. Edges in images are areas with strong intensity contrasts – a jump in intensity from one pixel to the next. Detecting an edge in the image reduces the amount of data and filters out useless information, while preserving the important structural properties in an image. The Sobel operator is applied on HSI image to detect the images.

Sobel Operator is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function. The Sobel operator is based on convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction and is therefore relatively inexpensive in terms of computations.

The operator uses two 3×3 kernels which are convolved with the original image to calculate approximations of the derivatives - one for horizontal changes, and one for vertical. If we define **A** as the source image, and $\mathbf{G}_x$ and $\mathbf{G}_y$ are two images which at each point contain the horizontal and vertical derivative approximations, the computations are as follows:

$$\mathbf{G}_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$

Where $*$ here denotes the 2-dimensional convolution operation.

The Sobel operator performs a 2-D spatial gradient measurement on an image.

The magnitude of the gradient is then calculated using the formula:

$$|G| = \sqrt{Gx^2 + Gy^2}$$

Image after edge detection

3.4.4 **Image Binarization**

A binary image is an image that only uses two values for its pixel values (0 and 1). To binarize our image, we have used Otsu's method, which assumes that the image which is to be thresholded contains two classes of pixels i.e foreground and background and then calculates the optimum threshold separating those two classes so that their combined spread is minimal.

Otsu's thresholding method involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels each side of the threshold, i.e. the pixels that either fall in foreground or background. The aim is to find the threshold value where the sum of foreground and background spreads is at its minimum.

Binary Image
HSBC

Image after Binarization

3.4.5 **Dilation and Extraction of connected Components of text**

In order to compute the dilation we will superimpose the structuring element on top of the input image so that the origin of the structuring element coincides with the input pixel position. If at least one pixel in the structuring element coincides with a foreground pixel in the image underneath, then the input pixel is set to the foreground value. In this way we can extract the components which are connected to each other.

Image after dilation
HSBC

Dilated Image

3.4.6 **Filtration by removing Noise**

Image is reversed after dilation and extraction of connected components. OR operation is applied to original image and reversed image to remove the noise.

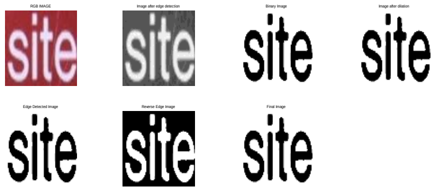
Final Image
HSBC

Final Image

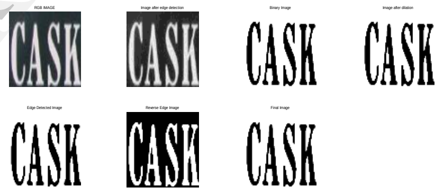
5. EXPERIMENTAL RESULTS

We have conducted the following test on the framework which we have prepared and after conducting the test, accuracy were determined by calculating precision and recall rate.

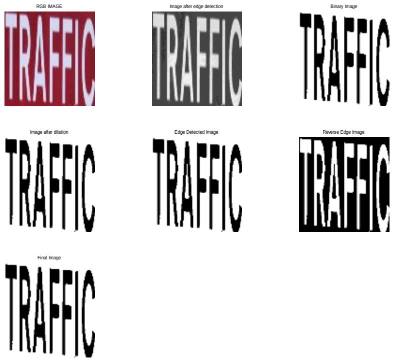
Test 1



Test2

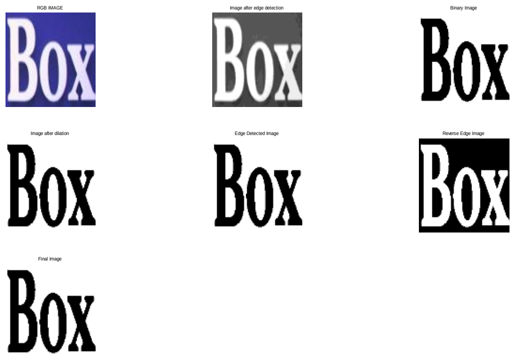


Test 3

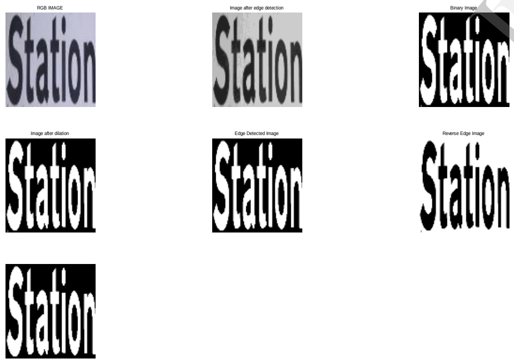


Test4

Test5



Test 6



5.1 Test Results

Test	Precision Rate	Recall Rate
Test 1	100	100
Test 2	98.0769	98.0769
Test 3	100	100
Test 4	100	100
Test 5	100	100
Test 6	98.1818	98.0769
Overall Precision Rate	99.38	
Overall recall Rate	99.34	

5.2 Comparison with other Methods

Method which is proposed into the paper is compared with the existing text extraction Algorithms and the following results were obtained.

Method	Precision Rate (%)	Recall Rate (%)
Proposed Algorithm	99.38	99.34
Shivanand S. Seeri	98.46	97.83
Samarabandu	91.8	96.6
J. Gllavata	83.9	88.7
Wang	89.8	92.1
K.C. Kim	63.7	82.8
J. Yang	84.90	90.0

After Comparing with the other methods, proposed method is better than the existing one and it extracts small sized characters with higher accuracy.

6. CONCLUSION AND FUTURE SCOPE

In this paper we have tried to present an approach which will extract characters from natural scene Images with higher accuracy, precision and recall rate. Algorithm is implemented on small sized characters and it works well with the small sized characters. Limitation of the algorithm is that it did not work well with the character images which are blurred in nature. Future work involves extraction of text characters from blurred images with higher accuracy.

7. REFERENCES

1. Kim K.C, Byun, H.R ., Song Y.W, Chi, S.Y, Kim, K.K, Chung Y.K, " Scene Text extraction in natural scene images using hierarchical feature combining and verification" Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on page 679-682 Volume:2

2. Shivananda V Seeri, Ranjana B Battur, Basavaraj S Sannakashappanavar " Text Extraction from Natural Scene Images", International Journal of Advanced Research in Electronics and Communication Engineering, Volume 1 October 2012

3. Xiaoqing Liu and Jagath Samarabandu, "Multiscale edge-based Text extraction from Complex

   images,"

   IEEE, 2006

4. Nobuo Ezaki, Marius Bulacu, Lambert Schomaker, Text Detection from Natural Scene images : Towards a System for Visually impaired Persons, Proc. of 17th Int. Conf. on Pattern Recognition (ICPR 2004), IEEE Computer Society, 2004, pp. 683-686, vol. II,
   23-26 August, Cambridge, UK

5. Xu-Cheng Yin, Xuwang Yin, Kaizhu Huang, Hong-Wei Hao , " Robust Text Detection in Natural Scene Images" IEEE Explore June 2013

6. J. Yang, J. Gao, Y. Zhang, X. Chen and A. Waibel, "An Automatic Sign Recognition and Translation System", Proceedings of the Workshop on Perceptive User Interfaces (PUI'01), 2001, pp. 1-8.

7. A.K Jain " Fundamentals of Digital Image Processing" Englewood Cliff, NJ: Prentice Hall, 1989, Ch 9

8. R.C. Gonzalez, " Digital Image Processing Using MATLAB"

9. N. Otsu, "A Threshold Selection Method from Gray- Level Histogram", *IEEE Trans. Systems, Man and Cybernetics*, Vol. 9, 1979, pp. 62-69

10. S.M. Lucas, A. Panaretos, L. Sosa, A. Tang, S. Wong, and R. Young, "ICDAR 2003 Robust Reading Competitions", *Proc.of the ICDAR*, 2003, pp. 682-687