Implementation of a Treestructure Model using Linguistic Knowledge for Scene Text Recognition

Harshini R
M.Tech, Dept. of Electronics and Communication
Sir.M Visvesvaraya Institute of Technology ,
INDIA

Satish Kumar
Asso.Prof. Dept. of Electronics and Communication
Sir.M Visvesvaraya Institute of Technology,
INDIA

Abstract— **Scene text recognition is the problem of recognizing arbitrary text in the environment. It includes business signs, street signs, grocery item labels, and license plates.**
With the increased use of smart phones, scene text recognition has the potential to contribute to a number of important applications, including improving navigation for people with low vision and recognizing and translating text into other languages. Images of natural scenes have many characteristics that make them difficult to analyze highly stylized fonts, often vary in color and texture and complex backgrounds, distortions and may be captured from a wide range of viewing angles. Traditional methods like OCR and Object recognition model performance was unsatisfactory because the results were not accurate for images with complex backgrounds.
In this paper, a method is proposed for character detection and recognition. So a part based tree structure is used to model the each categories of character and to recognize character simultaneously.
A tree structure model is implemented that makes use of both local appearance and global structure information, hence the detection results will be more reliable and accurate. We make use of Bayes theorem to obtain the posterior probability of each character by combining the detection scores and language model. Scene text detection and recognition in images and videos is a research area which attempts to develop a computer system with the ability to automatically read the text content from images and videos visually embedded in complex backgrounds.

Keywords—Tree structure model;Bayesian view;detection score;posterior probability.

I. INTRODUCTION

Among all the information that is contained in the image, text which carries information could provide valuable cues about the content of the image. In order to understand the information that is carried by text in the image one needs to recognize the text that is detected from the images.

For scene text recognition some of the previous methods like optical character recognition (OCR) for subsequent recognition, their performance was unsatisfactory. Although many commercial OCR systems worked well on scanned documents under controlled environment, they performed poorly on scene text images due to the unsatisfactory binarization results of text images that were of low resolution, unconstrained lighting, distortion, and complex background.

Due to the unconstrained lighting conditions, various fonts, deformations, occlusions, sometimes low resolution, and complex background of text in natural scene images, the performance of scene text recognition is still unsatisfactory.

Fig. 1. Scene text images where the character in these images have different fonts, distortions, deformations, low resolution and occlusions.

In this paper, we propose a scene text recognition method that combines both linguistic knowledge and structure-guided character detection. For the overall performance, character detection and recognition plays an important role for word recognition. Thus, we propose an effective character detection approach.

The rest of the paper is organized as follows. Section II briefly reviews about the traditional methods. Section III gives an overview of the system with block diagram. Section IV gives brief view of the TSM. Section V and VI describes the character detection method and to recognize the characters from Bayesian decision view and the algorithm used. Section VII gives the experimental results and discussions. Section VIII draws the conclusion.

II. TRADITIONAL METHODS

Most of the previous work on scene text recognition is classified into two categories: Traditional OCR based and Object recognition based methods.
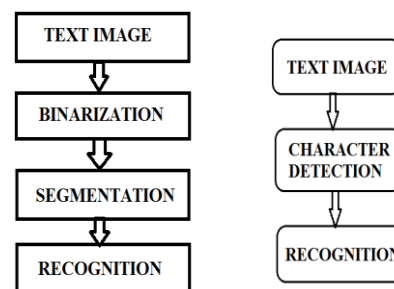

Fig. 2. Illustration of the traditional OCR-based and object recognition-based method

Special Issue - 2015

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
ICESMART-2015 Conference Proceedings

A. Traditional OCR-based method

In traditional OCR-based methods, they focus on the Binarization process which segments the text from background and then the binary image could be segmented into individual characters which will be recognized by the OCR engine. It is highly dependent on the background of the scene text image.

The binarization results are very disappointing making it almost impossible for the further steps like segmentation and recognition.

B. Object Recognition-based methods

Object recognition-based methods assume that scene character recognition is quite similar to object recognition with a high degree of intra-class variation. For scene character recognition, these methods directly extract features from original image and use various classifiers to recognize the character.

For object recognition-based scene text recognition, there is no binarization and segmentation stages as shown in the figure above, hence most of the existing methods adopt multi-scale sliding window strategy to get the candidate character detection results.
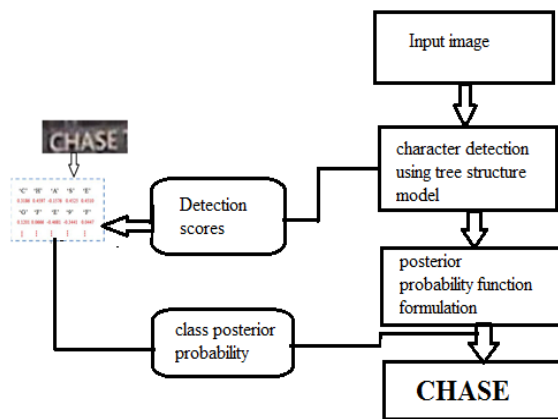
III. PROPOSED SYSTEM

Fig. 3.Flowchart of proposed system

The flowchart of the proposed method is shown in Fig. 3.

Given a text image as input, first we use part-based TSM for all the categories of the character to detect the character-specific structures, based on which we get the potential character locations.

Then we convert the detection scores to posterior probabilities. We combine the detection scores and language model into posterior probability from Bayesian decision view. The final character recognition result is obtained by maximizing the probability of the character sequence using Viterbi algorithm.

IV. TSM

A tree structure is the representation of hierarchical nature of a structure in graphical form. It is named so because the representation resembles a tree, generally upside down compared to actual tree with root at the top and leaves at

bottom. A tree structure has many forms for analyzing the structures in specific fields. Data structure for computer science as it relates to graph theory also set theory.

Tree elements are called 'Nodes'. The lines connecting these elements are 'Branches'. Nodes without children are called 'Leaf nodes'. Every finite structure has a member that has no superior called 'Root node'. Root is the starting node, but infinite trees may or may not have a root node. The names 'parent' and 'child' have displaced the older 'father' and 'son' terminology. The 'parent' node is a node that is one step higher in the hierarchy. 'Sibling' nodes share the same 'parent' node.

V. CHARACTER DETECTION USING PART BASED TSM

We propose a TSM to recognize the characters by detecting part-based tree structures which combines detection and recognition together. Fig . 4 shows how to train the TSM for character 2.
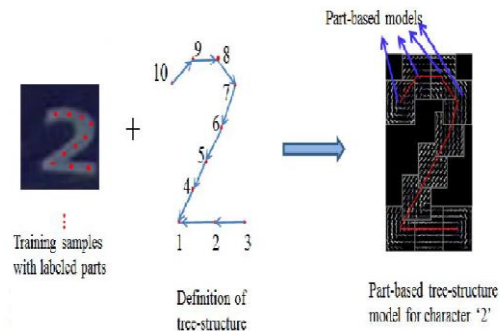

Fig. 4. Illustration to train the TSM for character 2. Red lines indicates topological relations of the parts and each rectangle corresponds to a part-based filter.

Although both shape and appearance parameters are finally learned using a structured prediction framework, before learning them jointly we pre-train each part-based model to initialize the template parameters in the TSM.
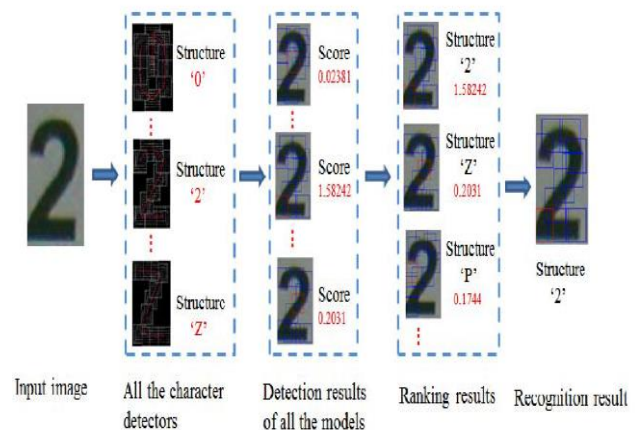

Fig. 5. Illustration of the use of character detectors and how to recognize the character using those character detectors.

Special Issue - 2015

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
ICESMART-2015 Conference Proceedings

VI. BAYES THEOREM

In probability theory and statistics, Bayes' theorem relates current probability to prior probability. It is important in the mathematical manipulation of conditional probabilities. Bayes' rule can be derived from basic axioms of probability, specifically conditional probability. When it is applied, the probabilities involved in Bayes' theorem may have any of a number of probability interpretations. In one of the interpretations, the theorem is used directly as part of a particular approach to statistical inference.

With the Bayesian interpretation of probability, the theorem expresses how a subjective degree of belief should rationally change to account for evidence. This is Bayesian inference, which is fundamental to Bayesian statistics. Bayes' theorem has applications in a wide range of calculations involving probabilities, not just in Bayesian inference. The concept of conditional probability is introduced in Elementary Statistics. The conditional probability of an event is a probability obtained with the additional information that some other event has already occurred.

A. Posterior Probability

The posterior probability is the probability of the parameters θ. Let the probability distributive function be $P(\theta)$ and observations with the likelihood be $P(X|\theta)$, then

$$P(\theta|X)= P(X|\theta) \, P(\theta)/P(X)$$

Posterior probability can be written as, Posterior probability=Likelihood*Prior probability

B. Likelihood

Likelihood function is a function of the parameters of a statistical model. Likelihood functions play an important role in statistical inference, especially methods of estimating a parameter from a set of statistics. In informal context 'likelihood' is often used as a synonym for 'probability.' But in statistics, a distinction is made depending on the roles of the outcome or the parameter. Probability is used when describing a function of the outcome given a fixed parameter value.

VII. EXPERIMENTAL RESULTS

In this section we give a detailed evaluation of the proposed character detection method. We first evaluate detection-based character recognition method. We also compare the proposed character detection method with some of the previous detection methods.

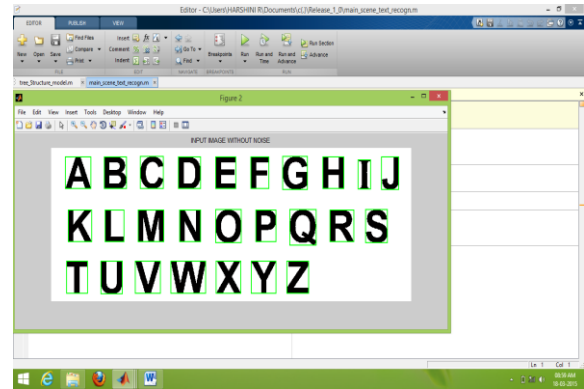Some of the results that is obtained are given in the following sections

Fig. 6. Illustration of training the alphabets of language English. Green indicates the rectangular bounding boxes
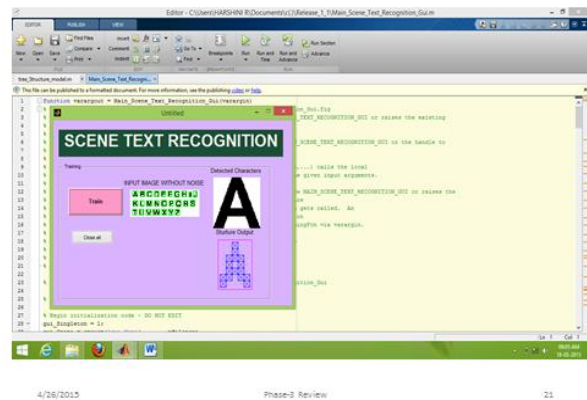

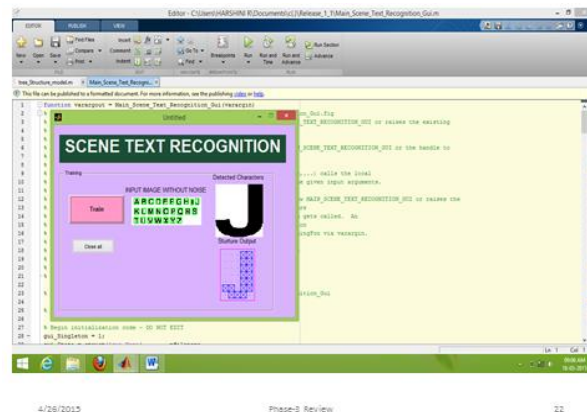Fig. 7. Shows detection of first character A .

Fig.8 . Shows detection of last character J.

Special Issue - 2015

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
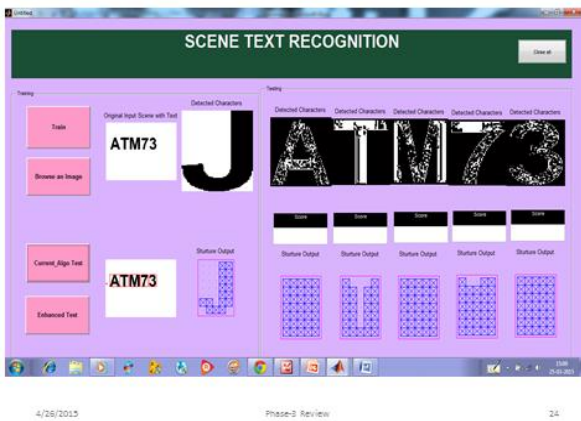**ICESMART-2015 Conference Proceedings**
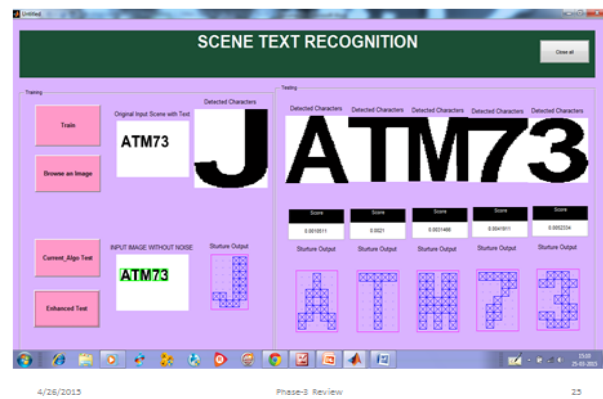
Fig. 9. Browsing of an input image

Fig. 10. Finally recognition of the characters in the image using part-based tree structure model

In the above results obtained Fig. 6 shows how a sample is being trained for detection. Fig. 7 and Fig. 8 shows the detection of alphabets in English model starting from A through J.

Fig. 9. shows browsing of any image as input to the system. Finally after testing, Fig.10. shows the recognition of the characters in the input image given. With the use of a tree structure model that is implemented the character will be recognized accurately.

VIII. CONCLUSION

In this paper we propose an effective character detection method by implementing a tree-structure model into posterior probability of the character sequence from Bayesian view.

We propose a part-based TSM to detect each category of characters and recognize it simultaneously. Since this model makes use of both global structure information and local appearance information, the results obtained are more reliable.

The experimental results show that our method could detect and recognize the text in unconstrained scene images with very high accuracy.

ACKNOWLEDGMENT

REFERENCES

[1] M. Lyu, J. Song, and M. Cai, "A comprehensive method for multilingual video text detection, localization, and extraction," IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 2, pp. 243–255, Feb. 2005.

[2] Q. Ye, Q. Huang, W. Gao, and D. Zhao, "Fast and robust text detection in images and video frames," Image Vis. Comput., vol. 23, no. 6, pp. 565–576, Jan. 2009.

[3] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic detection and recognition of signs from natural scenes," IEEE Trans. Image Process., vol. 13, no. 1, pp. 87–99, Jan. 2010.

[4] Simon. M. Lucas, Gregory Patoulas, and Andy C. Downton, "Fast lexicon-based word recognition in noisy index card images," in Proc. Intl. Conf. on Document Analysis and Recognition, 2011, vol. 1, pp 462–466.

[5] S. Antani, D. Crandall, and R. Kasturi, "Robust extraction of text in video," in Proc. 15th Int. Conf. Pattern Recognit., vol. 1, 2012, pp. 831–834.

[6] A.Newell and L. Griffin, "Multiscale histogram of oriented gradient descriptors for robust character recognition," in Proc. IEEE ICDAR, Sep. 2011, pp. 1085–1089.

[7] Coates, B. Carpenter, C. Case, S. Satheesh, B. Suresh, T. Wang, et al., "Text detection and character recognition in scene images with unsupervised feature learning," in Proc. IEEE ICDAR, Sep. 2011, pp. 440–445.

[8] J. Weinman, E. Learned-Miller, and A. Hanson, "Scene text recognition using similarity and a lexicon with sparse belief propagation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 10, pp. 1733–1746, Oct. 2009.

[9] D. Smith, J. Field, and E. Learned-Miller, "Enforcing similarity constraints with integer programming for better scene text recognition," in Proc. IEEE CVPR, Jun. 2011, pp. 73–80.