

Auto Secured Text Monitor in Natural Scene Images

Sowmya H N¹, Swetha Dareshwar², Sowmya K³, Vijayalakshmi S Katti⁴,
Sushitha S⁵, N.Samanvita⁶
^{1,2,3,4,5,6}MCA, NMIT Bangalore, India

Abstract: Monitoring the text in content based images is an important task. An accurate and robust method to monitor texts in natural scene images. In this paper fast and effective decision tree algorithm (MSERs) is used to extract a character participant by reducing regularized character dissimilarities. Character participant are grouped into text participant by using single-link clustering algorithm. Text participant with high non-text probabilities are removed and texts are establish with a text classifier. Valuable information may be broken in many content-based images, such as web images. We need to detect text before recognizing and retrieving.

Keywords: single-link clustering Algorithm, binarization, Maximally Stable Extremal Regions (MSERs) algorithm

I. Introduction

Text in pictures enclose advantageous information and is misused in numerous substance based picture because of complex foundation and varieties message in characteristic scene pictures must be powerfully recognized before being perceived and recovered. Existing strategies for scene content recognition can generally be sorted into three gatherings: sliding window based routines, associated segment based systems, and crossover techniques. Sliding window based systems, otherwise called area based strategies, utilize a sliding window to scan for conceivable messages in the picture and utilize machine learning methods to recognize content. Joined segment based systems concentrate character from pictures by associated part examination took after by gathering character applicants into content; extra checks may be performed to evacuate false positives.

The main purpose of the system is to reduce the learning parameters by hand. The previous system is complicated to edit or to make changes in text based scene images, because it consumes more time and human effort also it heals more memory .To overcome from previous system we have build an application using advanced algorithm, once the image has been uploaded by the user they have to mention the what changes have to be done in the text box then the service provider will make desired changes and returns the changed document.

To Extract the text from content based image using Maximally Stable Extremal Regions (MSERs) algorithm. Once text based image is uploaded by the user then with the help of provider user will get the changed text in the text format.

II. Related Analysis

The speed and accuracy of the existing systems is not satisfied. The existing systems use rule-based methods and clustering methods which involve tuning parameters by hand, which is time-consuming and error prone. Here several problems remain to be addressed. Most of the redundant components need to be removed before processing. Natural images contain text with different background, there font, size and alignment may be different. So we need to identify them before being retrieved.

III. Current Work

MSER based scene content will find the various leveled structure of MSERs and basic elements, then we use proper decision tree algorithm to check amount of character member to be taken care at extreme points with a high precision. Using self-training distance metric learning algorithm to separate weights and gather edges naturally; then character members are assembled into content member clustering algorithm utilizing the parameters. Finally, utilize a character classifier to gauge the consequent probabilities of content member parallel to non-message and dispose of content member with high non-message probabilities.

IV. System Specification

A. Hardware Requirements

Processor	-	Pentium Dual-Core E5400
Clock speed	-	2.7 GHz
Hard Disk Space	-	2GB RAM

B. Software Requirements

- Technology - Cloud based
- Platform - Windows 2000, XP, NT
- Languages - HTML, Java, JSP
- Databases - MySQL
- Database GUI - Sqlyog
- Supporting Server - Apache Tomcat 5.5

V. System Architecture

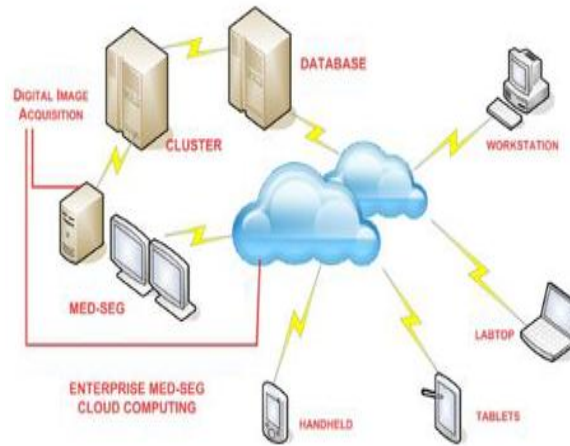


Fig 1: System Architecture

The above diagram (Fig 1) shows the System Architecture.

VI. Flow Chart

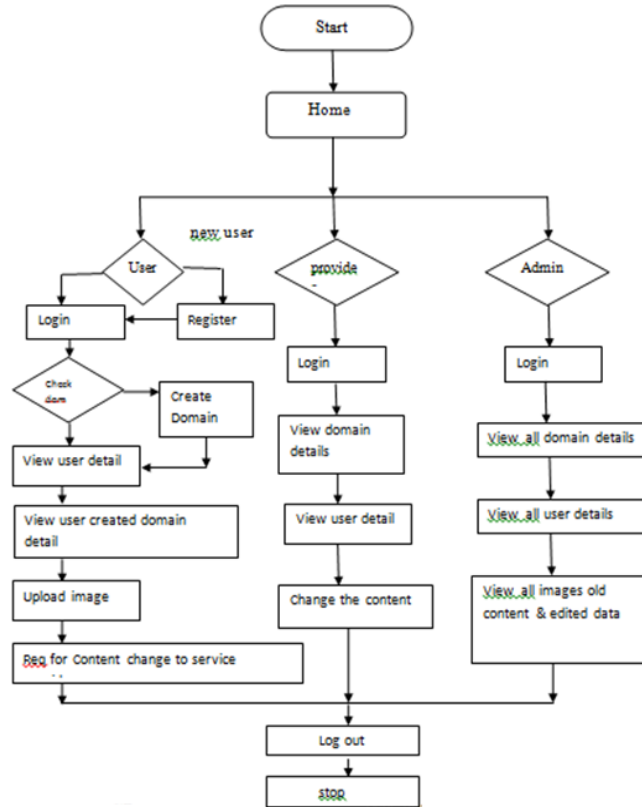


Fig 2: Flow Chart

The above diagram (Fig 2) shows the Flow Chart.

VII. Modules Description

A. Cloud Module

Cloud can view all the User Details & Service Provider Details. View all images which is uploaded by user & modified by service provider .Also can view all the Domain Uploads & Domain Details.

B. Service Provider

Service provider is providing the service to the User according to the priority provided by the Cloud. Service provider can view all domain details, and user details and the main work done by the provider is to view all images and the text which is to be replaced from old text to new text from the different users and make the respective changes without any complications and confusion.

C. Cloud User

User can upload the text containing Digital Images and check the availability of all domains. If the domain is free then user can request for that domain according to his need in terms of memory, life span etc. Upload the text containing digital images in the domain and request for the text change to the provider. Service Provider will identify the image provided by the information “from text” to “text” and make the changes, serves the request , access the edited data by unique id where data is highly secured.

VIII. Results

A. Domain Register Page

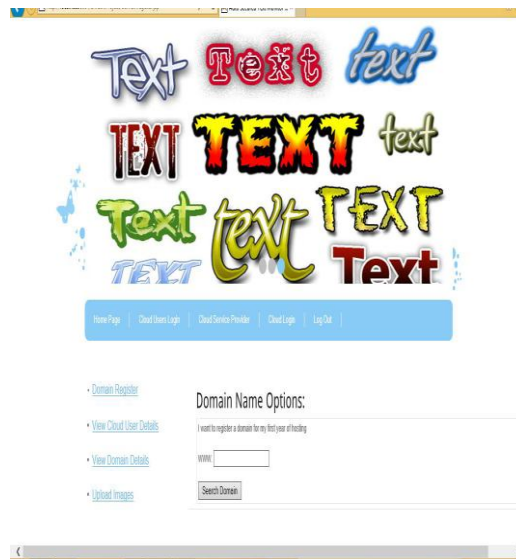


Fig 3: Domain Register Page screenshot

The above diagram (Fig 3) shows the result of Domain Register Page screenshot where it allows the individuals to Register.

B. Domain Detail Page



Fig 4: Domain Detail Page screenshot

The above diagram (Fig 4) shows the result of Domain Detail Page screenshot where it depicts the details of all the domains.

C. Image Upload Page

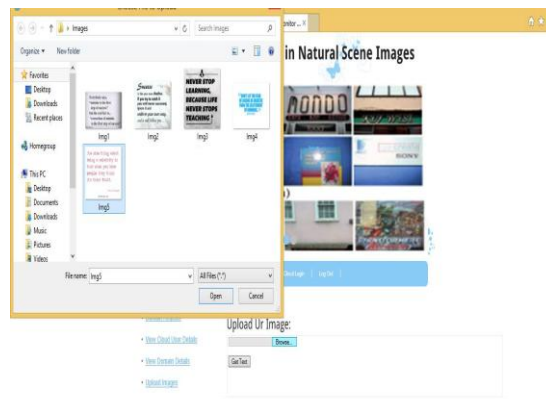


Fig 5: Image Upload Page screenshot

The above diagram (Fig 5) shows the result of Image Upload Page screenshot where it depicts the details of uploading the images.

D. Content Page

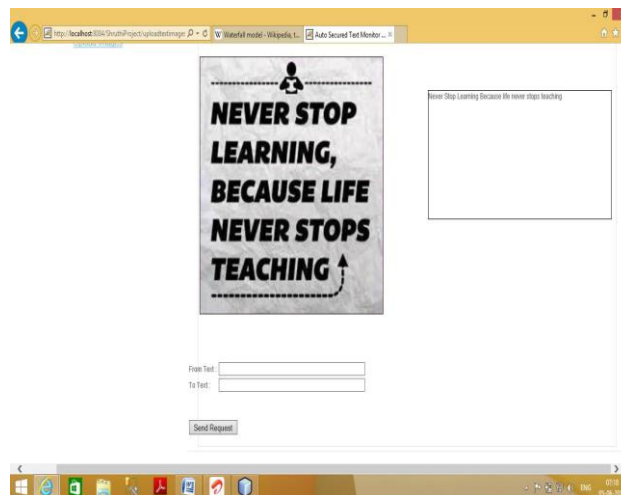


Fig 6: Content Page screenshot

The above diagram (Fig 6) shows the result of Content Page screenshot where it depicts the details of contents.

IX. Conclusion

This paper shows routines for mining content information with the utilization of side-data. Numerous types of content databases contain a lot of side data or meta-data, which may be utilized to enhance the grouping procedure. We also used an iterative dividing method as a estimation process which processes the significance of various types of meta-data. This methodology is utilized to outline both bunching and order calculations. Results on genuine information sets define the capability of our methodology. The outcomes demonstrate that the utilization of meta-data can incredibly upgrade the nature of content grouping and order, while keeping up an abnormal state of productivity.

Future Enhancements

New modules can be added to the existing system with less effort and amount of complications; in future we can add color correction, textual effects, image comparison, and 3D images.

References

- [1]. C.C.Aggarwal and P.S.Yu, (2012) "On text clustering with side information," in Proc. ICDE Conf, Washington, DC, USA,
- [2]. J. Chang and D. Blei, "Relational topic models for document networks," in Proc. AISTASIS, Clearwater, FL, USA, 2009, pp. 81–88.
- [3]. P. Domingos and M. J. Mazzini, "On the optimality of the simple Bayesian classifier under zero-one loss," *Mach. Learn.*, vol. 29, no. 2–3, 1999 pp. 103–130.
- [4]. A. Banerjee and S. Basu, "Topic models over text streams: A study of batch and online unsupervised learning," in Proc. SDM Conf., 2009, pp. 437–442.
- [5]. D. Cutting, D. Karger, J. Pedersen, and J. Tukey, "Scatter/Gather: A cluster-based approach to browsing large document collections," in Proc. ACM SIGIR Conf., New York, NY, USA, 1992, pp. 318–329.
- [6].