



A CONTENT DISCLOSURE AND TEXT IDENTIFICATION BASED ON ROAD-LEVEL SYMBOLISM USING TCH ALGORITHM

Ramya. R^{*1}, Vinothini. K. R², R. Manikandan³

^{*1, 2, 3} Department of Electronics and Communication Engineering, AVC College of Engineering, Annai College of Engineering & Technology, INDIA

*Correspondence Author: ramyarajaraman123@gmail.com

Abstract:

Image enhancement is a process of improving the quality of image by improving its feature. Image contrast enrichment techniques have been largely studied in the past decades. Traffic sign disclosure and realization has been fully prepared for a great past. Traffic board detection and text recognition still leaving a protest in computer view due to its various group and the vast changeability of the data illustrated in them. The essential function can be to make an automated index of the traffic panels placed in a road to holding its care and to aid drivers. Then the figure are defined as a “bag of visual words” and restricted to applying support vector machines. Completely our own text detection and recognition method is tested on those images where a traffic board has been disclosure in system to naturally read and save the data details in the panels. Leading driver aid scheme could also use from text recognition for automated traffic signs and panels description. This visual appearance categorization method is a new approach for traffic panel detection in the state of the art. We propose a language model partly based on a dynamic dictionary for a limited geographical area using a reverse geocoding service. Experimental results on real images from Google Street View prove the efficiency of the proposed method and give way to using street-level images for different applications on ITS.

Keywords:

Traffic Board Detection, Bag of Visual Words, Street-Level Images, ITS, geocoding service.

Cite This Article: Ramya. R, Vinothini. K. R, R. Manikandan, “A Content Disclosure and Text Identification Based on Road-Level Symbolism Using TCH Algorithm.” *International Journal of Research – Granthaalayah*, Vol. 3, No. 2(2015): 38-46.

1. INTRODUCTION

This paper presents a real application to intelligent transportation systems (ITS) of a method to detect and recognize text in images taken from natural scenarios proposed by the same authors in [1]. This text reading algorithm has proved to be robust in many kinds of real-world scenarios, including indoor and outdoor places with a wide variety of text appearance due to different writing styles, fonts, colors, sizes, textures and layouts, as well as the presence of geometrical distortions, partial occlusions, and different shooting angles that may cause deformed text. In this paper, this algorithm is applied, including some modifications and new functionalities, to read the information contained in traffic panels using the images served by Google Street View. The aim of this work is, in the first place, to detect traffic panels and to recognize the information inside them, showing



that the text detection and recognition method proposed in can be generalized to other scenarios, which are completely different to those that have been tested, without needing to retrain the system. In the second place, we want to develop an application that enables the creation of up-to-date inventories of traffic panels of regions or countries that facilitate traffic signposting maintenance and driver assistance. In this paper, we focus on traffic panels in the Spanish territory for two main reasons. First, unlike other countries, the coverage of Street View in Spain is near complete; thus, we can create a huge and diverse data set of images. Second, as far as we know, there is not any official database of all the traffic panels in Spain; thus, there are more possibilities that any government or institution responsible for managing the road network can be interested in having an up-to-date inventory of the traffic panels in Spain with the method here proposed. The reasons for which these organizations may be interested are various. Having a centralized database of all the traffic panels supposes a rapid and economic way of evaluating and analyzing the potential dangerous situations that may arise due to traffic panels that suffer from a bad visibility or show deteriorated or outdated information.

2. SYSTEM ARCHITECTURE

2.1. PROPOSED SYSTEM

The BOVW method stems from text analysis, wherein a document is represented by word frequencies without regard to their order. These frequencies are then used to perform document classification. The BOVW approach to image representation follows the same idea. The visual equivalent of words are local image features. Therefore, the BOVW technique models an image as a sparse vector of occurrence counts of the vocabulary of local image features. In other words, it translates a very large set of high-dimensional local descriptors into a single sparse vector of fixed dimensionality (a histogram) across all images. First, features at some key points are extracted in the train images and converted into feature descriptors, which are high dimensional vectors. In this paper, we compare different descriptors of the state of the art, as it will be explained later in this section. Then, the sampled features are clustered in order to quantize the space into a discrete number of visual words using k-means clustering. The visual words are the cluster centers and can be considered as a representative of several similar local regions. The image can be represented by the histogram of the visual words. The best performance is achieved for words, with the detection and recognition rates above 90% and close to 80%, respectively, when the panel is at short distance. However, the detection rate for numbers and symbols is lower compared with words because our text detection method, as explained in focuses on detecting text lines that has at least three elements; thus, numbers and symbols that may appear isolated in the panel could not be detected. A lower threshold than 3 could be used for this specific application, although the number of false positives may increase; however, what we wanted to show in this paper is that the proposed text detection and recognition algorithm trained with a concrete data set, which is composed of real-world images that are completely different to the Street View images used in



this paper, can be generalized to any other situation similar to the one shown in this work, achieving a reliable performance.

2.2. FLOW DIAGRAM

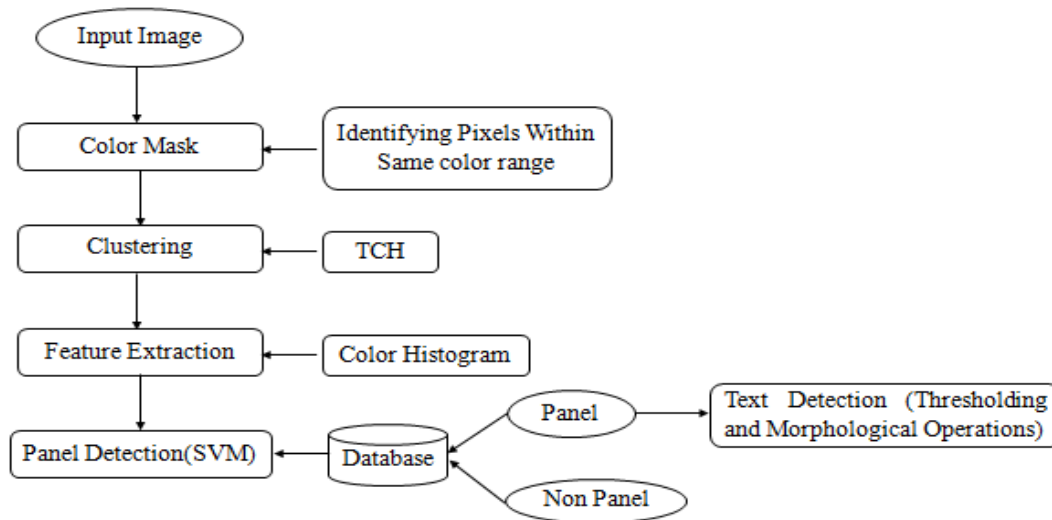


Figure 1: Flow Diagram

3. OVERALL ARCHITECTURE

3.1. MODULES DESCRIPTION

3.1.1. COLOR MASKS

Here SRM is used to create a color map structured objects or semantically meaningful parts of structured objects. In general, refers to a basic capability of the human visual system to derive relevant groupings and structures from an image without prior knowledge of its contents. The ability to impose structural organization on sensory data, which is central to perceptual organization, helps in pruning the set of possible interpretations of the sensory data by emphasizing hypotheses that exhibit structure and organization. We first pick one part and then keep growing the region by trying to group its neighbors with the region. The process stops when none of the region's neighbors can be grouped with the region. To achieve this, we develop a measurement to measure how accurately a region is grouped. The region goodness directly depends on how well the structural relationships of parts contained in the region obey Gestalt laws. The best region corresponds to a single structured object or the semantically meaningful part of the structured object.



3.1.2. FEATURE EXTRACTION

In this module we extract the feature based on the color representation of the images. It will calculate the color distribution of the images. The color histogram can be built for any kind of color space, although the term is more often used for three-dimensional spaces like RGB or HSV. For monochromatic images, the term intensity histogram may be used instead. For multi-spectral images, where each pixel is represented by an arbitrary number of measurements (for example, beyond the three measurements in RGB), the color histogram is N -dimensional, with N being the number of measurements taken. Each measurement has its own wavelength range of the light spectrum, some of which may be outside the visible spectrum. If the set of possible color values is sufficiently small, each of those colors may be placed on a range by itself; then the histogram is merely the count of pixels that have each possible color. Most often, the space is divided into an appropriate number of ranges, often arranged as a regular grid, each containing many similar color values. The color histogram may also be represented and displayed as a smooth function defined over the color space that approximates the pixel counts.

3.2. SVM CLASSIFIER

- Data setup: our dataset contains three classes, each N samples. The data is 2D plot original data for visual inspection
- SVM with linear kernel ($-t 0$). We want to find the best parameter value C using 2-fold cross validation (meaning use $1/2$ data to train, the other $1/2$ to test).
- After finding the best parameter value for C , we train the entire data again using this parameter value
- Plot support vectors
- Plot decision area.

SVM maps input vectors to a higher dimensional vector space where an optimal hyper plane is constructed. Among the many hyper planes available, there is only one hyper plane that maximizes the distance between itself and the nearest data vectors of each category. This hyper plane which maximizes the margin is called the optimal separating hyper plane and the margin is defined as the sum of distances of the hyper plane to the closest training vectors of each category.

Expression for hyper plane

$$w \cdot x + b = 0$$

x – Set of training vectors

w – vectors perpendicular to the separating hyper plane

b – offset parameter which allows the increase of the margin

Kernel function is used when decision function is not a linear function of the data and the data will be mapped from the input space through a nonlinear transformation rather than fitting non-linear curves to the vector space to separate the data. The classification task is able to scale high dimensional data relatively well, classifier complexity and classification error can be controlled explicitly.



3.3. TCH ALGORITHM

TCH – Transformed Color Histogram algorithm is used for the clustering of the image regions based on color information. Using TCH it has been possible to successfully cluster the descriptors and train the classifier, because convergence was not reached using the other descriptors. TCH transforms RGB into normal distributions of the color channels for the image patches pointed by the detected keypoints. A color histogram is a representation of the distribution of colors in an image. A color histogram represents the number of pixels that have colors in each of a fixed list of color ranges that span the image's color space, the set of all possible colors. The keypoints that are belonging to the similar color regions were identified to be similar clusters. The pixels within the range of the keypoints observed were detected as cluster regions.

3.4. PANEL DETECTION AND TEXT IDENTIFICATION

In this module we identify the panel based on the prediction of the SVM. Which cluster is predicted as a panel that index is used to detecting the area in the traffic panel image. The detected partition portion is merged together to predict the panel in the image. Finally the identified panel is displayed in the images. The thresholding operation identifies the image pixels that are having the pixels similar to the input threshold value. Binary images may contain numerous imperfections. In particular, the binary regions produced by simple thresholding are distorted by noise and texture. Morphological image processing pursues the goals of removing these imperfections by accounting for the form and structure of the image. Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. Morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. The morphological operation eliminates the unwanted pixels and identifies the exact text portions.

4. RESULTS

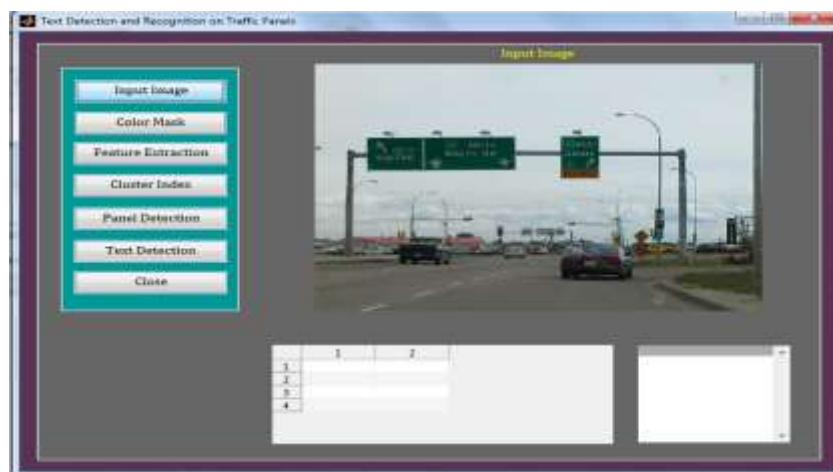


Figure 2: Input Image

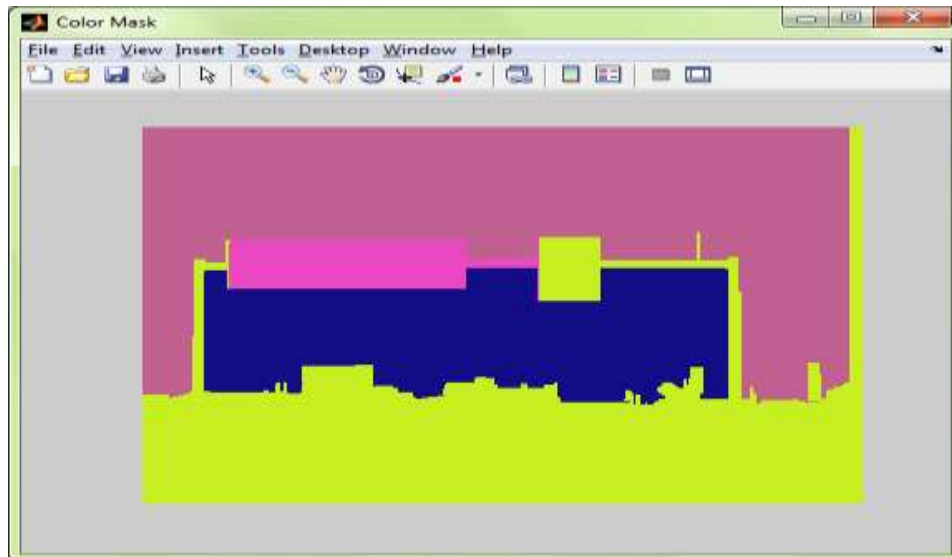


Figure 3: Color Masking



Figure 4: Separated Clusters

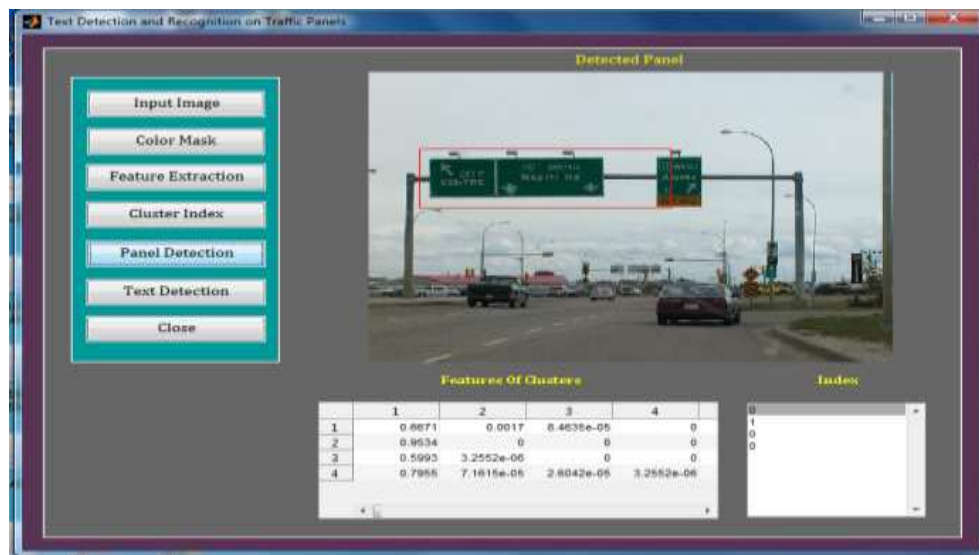


Figure 5: Panel Detection

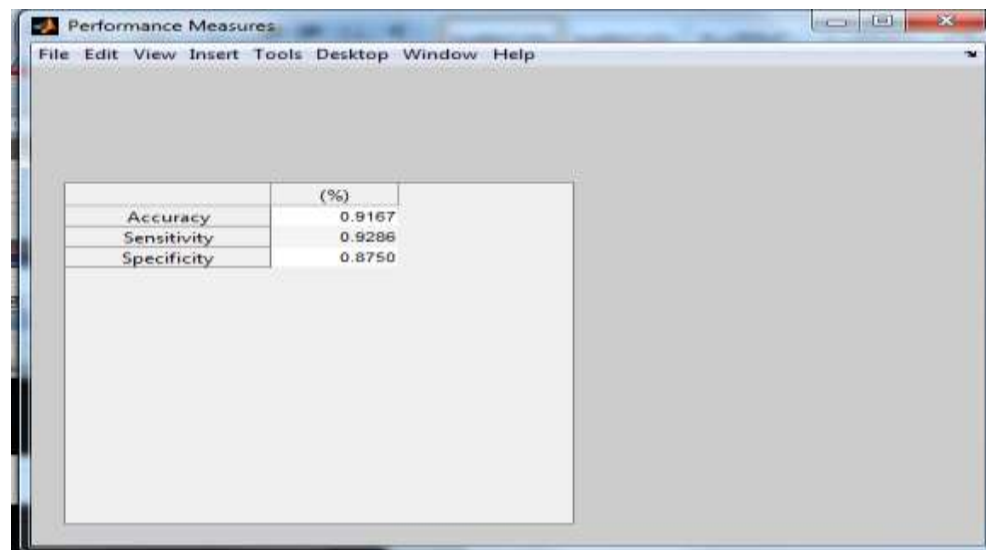


Figure 6: Performance Evaluation

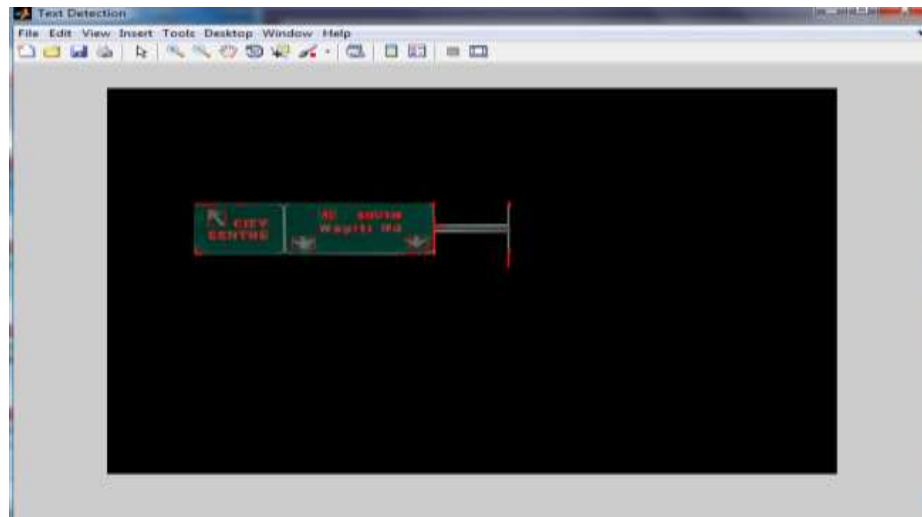


Figure 7: Detected Text

5. CONCLUSION

The traffic panels using panoramic images downloaded from the Google Street View service. The main use of this application is to automatically create up-to-date inventories of traffic panels of whole regions or countries. . This information is very useful for supporting road maintenance and for developing future driver assistance systems. One of the main contributions of this paper is the modeling of traffic panels using a BOVW technique from local descriptors extracted at interest key points, instead of using other features such as edges or geometrical characteristics as it has been done up to now in the literature. This is not an easy task due to the immense variability of the information included in traffic panels. Nevertheless, the experimental results show the effectiveness of the proposed method. An efficient segmentation method based on color masks has been implemented to guide the keypoints searching in the image. In addition, the dimensionality of this descriptor is small (only 37 and 45 elements, respectively); thus, the training time is lower than using other descriptors of higher dimensions. This method is to reduce the size of the dictionary to a limited geographical area using a reverse geocoding service. For the future work, multi frame integration of detected text will be performed.

6. REFERENCE

- [1] Gonzalez. A and Bergasa.L.M (Mar 2013) “A text reading algorithm for natural images,” *Image Vis. Comput*, vol. 31, no. 3, pp. 255–274.
- [2] Mogelmose.A, Trivedi.M and Moeslund.T (Dec 2012) “Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey,” *IEEE Trans. Intell. Transp. Syst*, vol. 13, no. 4, pp. 1484–1497.



- [3] Hu.Z, (Feb 2013) “Intelligent road sign inventory (IRSI) with image recognition and attribute computation from video log,” *Comput-Aided Civil Infrastruct Eng*, vol. 28, no. 2, pp. 130–145.
- [4] González.A, Garrido.M, Llorca.D, Gavilán.M, Fernandez.J, Alcantarilla.P, Parra.I, Herranz.F, Bergasa.L.M, Sotelo.M, and Revenga.P (Jun. 2011) “Automatic traffic signs and panels inspection system using computer vision,” *IEEE Trans. Intell. Transp Syst.*, vol. 12, no. 2, pp. 485– 499.
- [5] Reina.A.V, Sastre.R.J.L, Arroyo.S.L, and Jiménez.P.G (2006) “Adaptive traffic road sign panels text extraction,” in *Proc. 5th WSEAS ISPRA*, pp. 295–300.
- [6] Wu.W, Chen.X and Yang.J (Dec 2005) “Detection of text on road signs from video,” *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 4, pp. 378–390.
- [7] Csurka.G, Dance.C.R, Fan.L, Willamowski.J and Bray.C (2004) “Visual categorization with bags of keypoints,” in *Proc. Workshop Stat. Learn. Comput. Vis. ECCV*, pp. 1–22.
- [8] Cortes.C and Vapnik.V (Sep 1995) “Support-vector networks,” *Mach. Learn*, vol. 20, no. 3, pp. 273–297.
- [9] Lewis.D.D (1998) “Naive (Bayes) at forty: The independence assumption in info paper No.14136mation retrieval,” in *Proc. ECML*, 1998, pp. 4–15.
- [10] Mikolajczyk.K and Schmid.C (2001) “Indexing based on scale invariant interest points,” in *Proc. ICCV*, pp. 525–531.