International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463
Vol. 3 Issue 10, October-2014, pp: (139-144), Impact Factor: 1.252, Available online at: www.erpublications.com

Automatic Aggregation of Text and Sign on traffic panels using spatial extensions to BOVW

Monali Patil[1], Amit Pimpalkar[2]
[1,2]Department of Computer Science & Engineering, G.H.R.A.E.T, Nagpur University, India

Abstract: Traffic panel detection and recognition is use to support road maintenance and to help drivers. It is use to detect traffic panels and recognize the information present on street-level images. The images of traffic panel are taken by high resolution digital cameras or smartphones.it recognizes text and symbol accurately. To recognize text, system extracts local descriptors after applying green and white color segmentation. Then, images are classified using Naïve Bayes and represented as a "bag of visual words". In images if a traffic panel has been detected then Text detection and recognition method is applied on it to automatically store the information contain on the panels. We propose the system which uses spatial extension to BOVW such as sliding window, branch and bound. To recognize text exactly, we compute the prior probabilities of all the words using unigram language model .The language model completely based on a dynamic dictionary. Various algorithm use which is based on SIFT descriptors to recognize single characters and also on HMMs to recognize whole words.

Keyword: Bag of visual words (BOVW), HMMs, SIFT Descriptors.

INTRODUCTION

Traffic panels provide drivers information about the route by means of different traffic signs and text strings. Various shape of panel is used to show different traffic information such as rectangular panel which is commonly used to show direction of any route and distance. Triangular panel is used to show turning direction, Circular panel use to show only speed limit. There are many panels on the roads that can be confused with traffic panels, like advertisement panels.

Figure 1. (a)Turning direction with speed (b) Stop sign (c) Speed limit sign

Various public organizations and private companies have started to record street-level images. The most well-known service is Street View provided by Google. Text Reading from images is a challenging problem. This is important because there is a huge variety of text appearance because of different writing styles, sizes, textures fonts, colors, and layouts, also the presence of geometrical distortions, partial occlusions, and different shooting angles that may cause deformed text. Images suffer from blurring if it is taken from moving vehicle and its affect the segmentation step. Segmentation method extracts objects i.e. traffic signs from the background using color information. Object localization is an important task to understand the images such as to separate text and sign from the background, or to find the spatial relations between objects in an image. For this functionality Branch and bound is use. The aim of this work is to detect traffic panels and to recognize the information and sign inside them using text detection and recognition method ,for this Multiframe integration is use to improve efficiency of system. For detecting possible rectangular planes color segmentation and shape classification is done on image. A Dictionary is created which contain the words that the system is able to recognize, dataset consist panel that have a blue or a white background. We have taken OCR (Optical Character Recognition) systems as a starting point to identify the misspelled words if the detected traffic panel image is blurred .OCR takes each character as input, then computes its feature vector and using a KNN approach the object is classified into a class.

LITERATURE REVIEW

Various type of the information present on traffic panels automatic visual classification of the information contained on road panels has not been much research. A. Reina, et.al. [1] workedto detect candidates to be traffic panels by using image segmentation for blue and white colors region on image using HSI space. These candidates were classified according to their shape by method correlating the radial signature of their FFT (Fast Fourier Transform) with the pattern corresponding to rectangular shape. A homography between the original plane and the reoriented one, was carried out for an image transformation to correct the angular deviation of the panel in an image. Finally, SVM classifieris applied on the image to classify every symbol and character in grey-scale image. A priori information that can be known from the panels did not take into account because the kind of information present onthe panel depends on the situation over the panel itself.

H. Gómez-Moreno, et.al [2] presented the segmentation methods that can be categorized into edge detection, color-space thresholding and achromatic/chromatic decomposition. The segmentation algorithm consist four stages.1) Segmentation 2) Detection 3) Recognition 4) Tracking. For color segmentation SVMs was use to provide better result but it need some improvements when it applied to achromatic colors. Best methods were Ohtaor RGB Normalized that were normalized with respect to illumination and no need of improvement when it used with(HSI) Hue Saturation Intensity spaces.

W. Wu, etal [3] proposed a method to detect text present traffic panels from video. In this, two rules were use. 1) Apply a divide-and-conquer strategy to divide the single task into two subtasks. 2) Detect text from each video frame by combining (2D)two-dimensional image features with the (3-D)three-dimensional geometric structure information of objects which was extracted from video sequence. This rule gave 88.9% text detection rate and9.2% a false hit rate.To extract the region of same colork-means algorithm was used and traffic panel candidates were detected by searching for flat regions perpendicular to the camera axis. A multiscale text detection algorithm was performed on each candidate traffic panel area. Gaussian mixture models and geometry alignment analysis was usedto combines adaptive searching, edge detection and color analysis. For recognition all detected text lines are extracted, but how the recognition is carried out was not mentioned.

A. González, et al [4] described an approach to the VISUAL Inspection of Signs and panel ("VISUALISE") which was an automatic inspection system. It was based on light retro reflection Principle. It used an active infrared illuminator when it comes into contact with the Signs and then panels are reflected. It was capable of classifying a sign or panel into a certain class of material by comparing the luminance measurements with amodel obtained through a prior calibration process. The inspection process can be divided into offline process and online process. In Online process, necessary devices integrated in vehicle with all software applications to recording the input sequences corresponding to actively illuminated roads by the onboard infrared illumination system. In offline process, an image processing device based on a personal computer processed the recorded sequences after processing result contain the retroreflection and the contrast values of every sign and panel.

AshishEmmanuelet.al [5] presented an Optimal Text Recognition and Translation System for Smart phone. Any language can be converted into any other language. Genetic algorithms were used to process the image to maximize the quality for text conversion from image. The size of processed image is reduced and the image segmentation helps in reducing it further thus letting to use multiple asynchronous attempts for translating text.

J. Yebes,et al[6] introduced approach complemented the functionality of a traffic signposting inspection system based on computer vision, which was able to collect data related to the maintenance state of traffic signs and panels automatically which was based on SIFT descriptors to recognize single characters, numbers and symbols and HMMs used to recognize whole word. The text extraction algorithm was divided into three main steps. The first one was the segmentation, which is based on a canny edge detector, Second a geometrical transformation was applied to the image so that the inclination of the panel, due to distortion, is corrected. At last the location of the different elements on the panel is found out. A wrong extraction of the foreground objects was avoided by separating horizontal edges and vertical edges of the image. The advantage of this system was it only works with geometrical features and does not use a method based on colour. so it easily adapted to other traffic signposting regulations.

T.Gevers and C.M. Snoek [7] suggested the properties of color descriptors that never changes using a naming and classifying of invariance with respect to photometric transformations. Image domain and video domain were used to assessed distinctiveness of color descriptors .this paper presented three color descriptors First ,color descriptor based on histograms second color moments and color moment invariants and third was color descriptors based on SIFT. The addition of color descriptors over SIFT improves category recognition. For automated recognition of individual object and scene categories Color descriptor with an appropriate level of invariance was selected.

Christoph H. Lampert [8] proposed Efficient Sub window Search (ESS), a method for object localization. branch and bound scheme was use that allows maximization of quality functions of large class over all possible sub images. The branch-and-bound search quickly identifies an interested region and to find the exact maximum it focuses on its computation in this interested region. Every image was divided into a grid of regions and represents each grid cell by a separate histogram. The proposed method hierarchically split the parameter space into disjoint subsets. For each of the subsets Desired parts of the parameter space are evaluated first andif their upper bound show that they cannot contain the maximum then that parameter space of large parts do not have to be examined further. EXTENSIONS OF ESS has Object localization tasks require the detection of Number of object location in an image.

Bram Alefs et.al.[9] presented a system for detection road sign from panel based on edge orientation histograms. This system can be able to detect 85% of the objects from width of 12pixels and 95% for objects of width of 24 pixels at a low false alarm rate. Edge orientation histogramscan be calculated efficiently using integral images and they are static for small variations in rotation and position.It indicate the use of local indicators in same way as weak classifiers for AdaBoost .The application of this system consists of three step, Image acquisition, image processing and CAN interface.. Third step CAN interface manages the communication with the vehicle network according to proprietary transport protocol. This system performs at 7.5 frames per second on a Laptop PC.

Ravi Shekhar and C.V. Jawahar [10] presented an approach to retrieve efficiently and accurately similar word images from a large database. In Bag of visual Word model, Text or document was represented by an unordered set of words. Text was represented with histogram. To perform document or text classification and retrieval these histograms were used. On four Indian languages they validate this method and report a mean average precision that was more than 0.75. SIFT descriptors were used to quantized visual words of local regions. Limitation of this method was the re-ranking step, which is relatively time consuming.

PROPOSED WORK

The propose system is Automatic Aggregation of Text and Sign on a traffic panels Using spatial extensions to BOVW (Bag of Visual Word). This information is use to assist the driver and supporting road maintenance .We will create a Database which contains images of traffic panel these panel detected by using image segmentation for blue, green and white background colors. Segmentation is use to find out the location of text and signs based on their color and thus reduce search space. Color-based segmentation is use which is based on a thresholding of the input image in some color space. Spatial extension such as Branch and Bound is also use to find out the location of text and sign from image. It hierarchically splits the parameter space into different subsets. The parameter space which have maximum bound are explored first, If rest of the parameter space indicates that they cannot contain the maximum upper bound it do not have to be evaluate in future. Text and sign are considered as a feature which is extract from panel by using the Harris Laplace salient point detector. Fast Filter Method is use to remove irrelevant feature. After removing irrelevant feature image is reconstructed and represented as BOVW. We use spatial Extension to BOVW such as branch and bound. Most of the time extracted text is blurred or some character are erased due to environmental maladies so OCR will use to identify the misspelled words or symbols on traffic panels. Unigram probabilistic language model is use to recognize wordand we computes the prior probabilities of all the words to recognize text exactly. We use multiframe integration at each single frame of the recognized information. This refers to estimation and noise reduction technique. There are some panels with uniform white or green backgrounds in the image with letters inside them, so system will consider them as traffic panels generating false positives. Panels are validating using sensitivity and specification defines as:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (2)$$

TP :- True Positive TN:- True Negative

FP: - False Positive FN: - False Negative

The sensitivity shown in equation (1) relates to the system's ability to find out positive samples and the specificity shown in equation (2) relates to the system's ability to find out negative samples, if system detects a panel and that panel is traffic panel then it is true positive. If detected panel is advertisement panel not a traffic panel then it is false positive.If panel is present in frame or image but it is not detected by system then it is True negative.

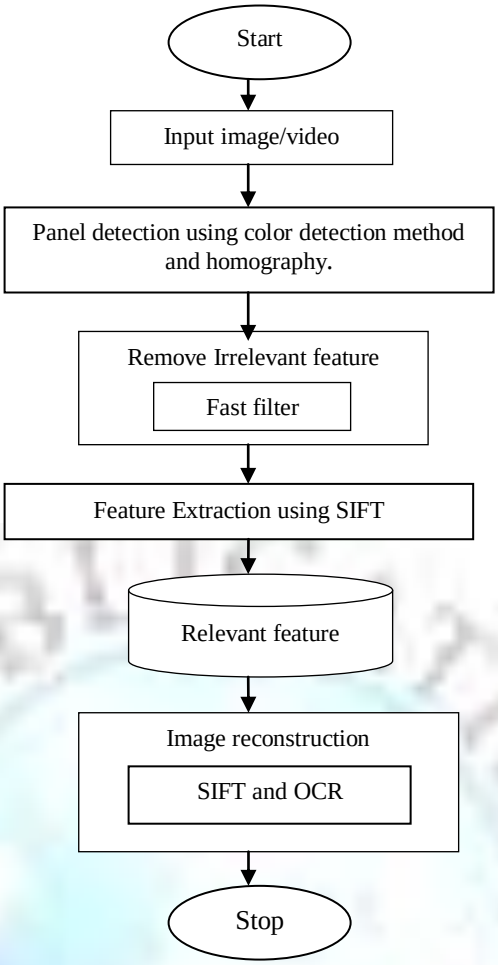

Figure 2: Data flow diagram of proposed system

In Fig 2. Database contains images of traffic panel. Images contain noises so we use Fast Filter Method to remove irrelevant feature. After removing irrelevant feature image is reconstructed and represented as bag of visual word. Panels are green in color and text on it are in black color so color segmentation is performed for that Color descriptor, Scale Invariant Feature Transform (SIFT) and Transformed Color Histogram (TCH) are used to detect green and white traffic panel. To extract the feature Harris-Laplace Salient Point Detector will use. Panels contain text and sign, to classify these objectNaïve Bayes is use. Text reading Algorithm is use to recognize the text and OCR will use to correct the misspelled word. For clustering we will use k-means algorithm which cluster the feature discrete number of visual words.

IMAGE ACQUISITION

The first stage of system is Image Acquisition. Images of traffic panel taken by live camera or mobile camera. Livecamera image will not be in a perfect lighting and background. Image is preprocessed to enhance the image. The color image is converted to a gray scale image and image smoothing, noise removal are performed.

FAST FILTER METHOD

Image taken by camera or mobile contain irrelevant feature that featur**e** are remove by fast filter method. This method has two parts, first how to decide whether a feature is relevant to the class or not. To decide relevant feature method use a user- defends threshold value SU (Symmetrical Uncertainty) values for feature. Second, how to decide whether such a relevant feature is redundant or not when considering it with other relevant features. For this SU values captured F-correlations in order to decide whether a relevant feature is redundant or not. Redundancy is cause if the level of correlation between two features in S0 is high. To avoid redundancy, one of them may be removed from S0. For a feature Fi in S0, the value of $SU_{i,}c$ quantizes the extent to which Fi is correlated to the class C. we will obtain quantized estimations about the extent to which Fi is correlated to the rest relevant features in S0. For measuring correlations between features Symmetry is a desire property symmetrical uncertainty defined as follows:

$$SU(X, Y) = 2 \left[\frac{IG(X|Y)}{H(X)+H(Y)} \right] \quad (3)$$

IG is information Gain .It is the amount by which the entropy of X decreases reflects additional information about X provided by Y and it is given by:

$$IG(X|Y) = H(X) - H(X|Y) \quad (4)$$

IMAGE RECONSTRUCTION

Image reconstruction creates (2D) two dimensional or (3D) three dimensional images from incomplete or scattered data. Image sharpening is use to make an image a readable, usable and useful. But there is problem with image reconstruction if mage contain noise and unwanted data that can interrupt the clarity of an image, so to remove noise or to correct the data OCR is use. In OCR, SIFT descriptor use to extract alphabetic characters, numbers and symbols. A SIFT key point described by a geometric frame of four parameters, the key point center coordinates x and y, its scale and its orientation angle. Thefourth parameter is orientation angle because some characters are symmetrical, like 6and 9 or b. From a set of reference images, SIFT key points of objects are first extracted and the corresponding descriptor vectors are stored in a database. These images consist all uppercase and lowercase letters, all the numbers since 0 to 9 .From new image character or symbol is recognized by comparing each feature vector from the new image to the database and finding candidate matching features based on Euclidean distance of their feature vectors.

COLOR SEGMENTATION

In image segmentation an image is separated into different parts that correspond to something that can easily separate and view as individual objects. The segmentation process is based on various features found in the image. This might be color information that is used to create histograms, or information about the pixels that indicate edges or boundaries or texture information. Some of colors descriptors are SIFT Hue histogram, TCH. We use TCH because it is invariance to scale and shift of intensity also it is not affected by shadow, lightening effect, and viewpoints in the scene. TCH transforms RGB (Red Green Blue) into normal distributions of the color channels for the image patches pointed by the detected key points

NAÏVE BAYES

A naive Bayes classifier is use for classification of objects into category. It assumes that every category has its own distribution on the visual vocabulary and that distributions of each category are observed differently. Let N be the number of visual words. Each and every image represented by $m = [m_1, m_2 \ldots m_n]$ where $m_i$is an N-dimensional vector whose ith component measures the frequency occurrence of visual word in the image. Consider c that represent the category of image. The classification decision is given by

$$cMAP = \arg\max p(c|m) \quad (5)$$

Equation (5) finds the class c so that it maximizes the posterior probability P(c|m).
Applying Bayes' rule to equation (5) can be expressed as

$$cMAP = \arg\max \frac{p(m|c)p(c)}{p(m)} \quad (6)$$

Assuming that distribution on each category is independent equation (6) reduces to

$$cMAP = \arg\max p(c) \prod_{n=1}^{N} p(m|c) \quad (7)$$

P(c) is the prior probability for class c.

CLUSTERING

The clustering referred as finding groups of similar objects in the data. A similarity function is used to measure the similarity between the objects. Clustering in general is an unsupervised process of grouping elements together, so that elements assigned to the same cluster are more similar to each other than to their mining data points.The k-means clustering algorithm uses a set of k representatives around which the clusters are built. These representatives are not necessarily obtained from the original data and are refined somewhat differently than a k-medoids approach. The simplest form of the k-means approach is to start off with a set of k seeds from the original corpus, and assign documents to these seeds on the basis of closest similarity. In the next iteration, the centroid of the assigned points to each seed is used to replace the seed in the last iteration. In other words, the new seed is defined, so that it is a better central point for this cluster. This approach is continued until convergence. One of the advantages of the k-means method is that it requires an extremely small number of iterations in order to converge.

CONCLUSION

The images of traffic panel that is blurred because of the image stitching are correctly detected as panels by this system. Images taken at early morning or sunset suffer from illumination effect, low contrast, and shadows on the panel, these panels are correctly detected due to TCH descriptor. This work improve recognition rate since it use multiframe integration technique. We use prior knowledge concept to reduce the false panel detection rate.

REFERENCES

[1]. Reina, R. Sastre, S. Arroyo , Adaptive traffic road sign panel text extraction, Proceedings of the 5th WSEAS International Conference on Signal Processing, Robotics and Automation, pp. 295-300, 2006.
[2]. S. Maldonado-Bascón,H. Gómez-Moreno, P. Gil-Jiménez, and S. Laurent-Arroyo, "Goal evaluation of segmentation algorithmsfor traffic sign recognition," IEEE Trans. Intel. Transp. Syst., vol. 11,no. 4, pp. 917–930, Dec. 2010
[3]. X. Chen ,J. Yang andW. Wu, "Detection of text on road signs fromvideo," IEEE Trans. Intelligent Transportation System., volume 6, number 4, pp. 378–390, December 2005.
[4]. A. González, M. Garrison, D. Llorca, "Automatic traffic signs and panels inspection System using computer vision," IEEE Trans. Intelligent Transportation System,volume12, number 2, pp. 485–499, June 2011.
[5]. Ashish Emmanuel S, Dr. S.Nithyanandam"An Optimal Text Recognition and Translation System for Smart phones Using Genetic Programming and Cloud"International Journal of Engineering Science and Innovative Technology (IJESIT) Volume 3, Issue 2, March 2014.
[6]. A. González, L. Bergasa, J. Yebes, "Automatic information recognition of traffic panels Using SIFT descriptors and HMMS" in Proceeding ITSC, pp. 1289–1294, 2010
[7]. T.Gevers, and C.M. Snoek, "Evaluating color descriptors for object and scene recognition, IEEE Trans. Pattern Analysis Machine Intelligent, volume32, number 9, pp.1582–1596,September 2010.
[8]. C. Lampert, M. Blaschko, and T. Hoffman, "Efficient subwindowsearch:A branch and bound framework for object localization," IEEE Trans Pattern Analysis Machine Intelligent, volume 31,number 12, pp. 2129–2142, December 2009.
[9]. Bram Alefs, Guy Eschemann, Herbert Ramoser, CsabaBeleznai" Road Sign Detection from Edge Orientation Histograms" Proceedings of the2007 IEEE Intelligent Vehicles Symposium Istanbul, Turkey, June 13-15, 2007.
[10]. Ravi Shekhar and C.V. Jawahar" Word Image Retrieval using Bag of Visual Words" 10th IAPR International Workshop on Document Analysis Systems, October 2012.
[11]. Oi-Mean Foong, SuziahSulaiman and KiingKiuLing "Text signage recognition in android mobile devices" Journal of Computer Science 9 (12): 1793-1802, 2013
[12]. C. Cortes and V. Vapnik, "Support-vector networks" Machine Learning, volume 20, Number 3, pp. 273–297, September 1995.
[13]. S.Lafuente-Arroyo, S. Maldonado-Bascón,"Road sign tracking with a predictive filter solution," in Proceed 32nd IEEE IECON, pp. 3314–331 , Nov. 2006.
[14]. S. Maldonado-Bascón, S. Lafuente-Arroyo "Road-sign detection and recognition based on support vector machines" IEEE Trans. Intelligent Transportation System, volume 8, number 2, pp. 264–278, June 2007.
[15]. A. González and L. M. Bergasa, "A text reading algorithm for natural images" Image Vis. Computer, volume 31, number. 3, pp. 255–274, March 2013.