BEST: International Journal of Management, Information Technology and Engineering (BEST: IJMITE)
ISSN (P): 2348-0513, ISSN (E): 2454-471X,
Vol. 4, Issue 1, Jan 2016, 25-30

A REVIEW ON DIFFERENT APPROACHES TO EFFICIENTLY DETECT AND RECOGNIZE INFORMATION FROM TRAFFIC PANELS

PAMI PRAKASH[1] & ANIL A. R.[2]

[1]M.Tech Scholar, Department of Computer Science and Engineering, Sree Budha College of Engineering, Pattoor, Kerala, India

[2]Head of the Dept., Department of Computer Science and Engineering, Sree Budha College of Engineering, Pattoor, Kerala, India

ABSTRACT

Text detection and recognition has been a major area which is studied for a long time. Traffic signals are assets that are crucial for safe transportation. Computer vision techniques applied to systems used on road maintenance, are playing a major role because of its importance. Reading text in natural images has gained the attention of many researchers during the past few years there is a huge availability of cheap image-capturing devices in low-cost products. The importance of automatic inventory of the traffic panels located in a road to support road maintenance and to assist driver is increased. The detection of the texts with random orientations from traffic panels has become an increasingly important and yet challenging task because of its huge variability. To address this problem, different techniques have been proposed and to classify the related algorithms. This paper provides an explanation of work done for automatic detection of text from images, localization and extraction of text from the traffic panel images having complex backgrounds..

KEYWORDS: Text detection, Computer Vision, Textual Information, Traffic Panel Localization

INTRODUCTION

Text detection and recognition from camera captured images of traffic panels have been considered as very important problem, has been one among the major areas of research over the past few decades. Nowadays, there are many text reading algorithms that had proved to be robust in many circumstances with a huge variability of text appearance due to different viewing angles, writing styles, sizes, colors, fonts, textures etc., as well as the presence of geometrical distortions and partial occlusions that may cause text deformation.

Text detection plays a vital role in content based image retrieval, and content-based visual information retrieval, which is the application of computer vision techniques solve the problems of image retrieval, which means searching for digital images in large databases applications. Texts present in a captured image can give us vital information about that image. This is still a challenging problem due to many reasons. First, there is a large variability of traffic panels as each of them depicts different information, varying in size, color, and shape. Images are of two types, document images and natural images. The image format of the paper document is known as document images.

The images which are taken by camera are natural images. Text information in an image can be broadly classified into scene text and caption text. Scene texts appear in natural images. But the caption texts added manually. Scene texts could be overlapped with the background. Therefore it is not so easy to identify scene texts as compared to that of caption

texts especially in case of natural images. The detection of interest points and local features constitutes the support for many significant computer vision tasks. For example, object recognition, stereo matching, mosaicking, robot navigation etc. rely on the detection of interest points which possess some distinctive, highly invariant and stable properties.

One of the valuable applications of text detection and recognition is helping people with visual impairments. It will be more helpful for them if something can convey them the text information present in an image. As people travel through different places, they find it very difficult to understand the text written on display boards on their way. In such cases, people either look for tour guides or hand held devices that can help them in translating the information written on display boards to the language which they understand. For this, text detection from display boards and its translation is an important part. So an efficient text detection and recognition automated system can help them in the case of lack of coverage of the GPS. License plate recognition in traffic surveillance is one of the important areas where text detection and extraction is used. License plate recognition has an important role in traffic surveillance, at checkpoints.

OCR is a method that can detect the text characters by edges detection. But edges can be recognized by OCR only if characters of the text are having a clear separation from background. Interference from background and degradation of images decreases the efficiency of OCR. So OCR performs in natural scenes is comparatively low. So, it becomes more complicated in the case of detecting texts from traffic panels. Texture analysis, topic based partition etc. are existing methods for text detection. These techniques can work correctly in document images but text string detection from natural image is a difficult task.

LITERATURE SURVEY

The main vision-based detection and recognition techniques for driver assistance systems and automated vehicles can be found in [2]. A recent idea about an intelligent road sign inventory based on image recognition, for traffic signs in place of traffic panels and using images obtained from a vehicle is described in [3]. Also, there is a huge diversity of information contained in traffic panels resulting in difficulty to analyze. To conclude, till now, the research on automatic detection and recognition of the information which is contained in traffic panels is scant. A novel method that extracts only the candidates who have the possibility to become traffic panels is proposed in [5], using a method that detects blue and white regions in the image using the hue and saturation components of the hue–saturation–intensity space. Then, these candidates are classified according to their shapes, for extracting the rectangular blobs. This can be accomplished correlating the radial signature of their fast Fourier transform with a pattern that corresponds to a correct rectangular shape. Then, reorientation of the panel is carried out using a homograph that aligns the four vertices of each blob. After the detection and the reorientation of the panel, segmentation of the foreground objects from the background of the panel is done by examining the chrominance and luminance histograms. Connected components labeling and position clustering is finally done for character arrangement on the traffic panels. This algorithm is invariant to all kind of translations and projective distortion, but it is severely affected by changing lighting conditions. In addition, many parameters and thresholds are adjusted temporarily. Recognition is applied at character level. There is no other information on how and where the images are extracted. Moreover, there is no performance evaluation in the experimental results provided by the authors; hence, it is very difficult to know the effectiveness of their proposal and it is impossible to compare, as they use data sets of their own.

Wu et al. [6] proposed a technique to detect text on traffic panels from video. First, k-means clustering algorithm is used for extracting the regions of the same color, and traffic panel candidates are detected by searching for flat regions

perpendicular to the camera axis. Three or more points in two successive frames are used to estimate the orientation of the candidate planes is estimated using; hence, this method needs an accurate tracking method to detect corresponding points in successive frames. Furthermore, on each candidate traffic panel area, a multi-scale text detection algorithm is performed. Edge detection, adaptive searching, color analysis using Gaussian mixture models, and geometry alignment analysis are integrated in text detection method. Every detected text line is covered by a minimum bounding rectangle.

Yao[5] proposed a method to detect text in images with arbitrary orientation using a two level classification technique. Stroke width transform and intrinsic features of text are used in this method. There are four steps in this image. The steps include component extraction, candidate linking, component analysis, and finally chain analysis. In order to reduce parameter tuning manually, a two level classification is used for training. Yi [6] proposed a method for text detection that can detect text with arbitrary orientation. The method comprises of mainly two steps. Firstly, the image is divided based on stability in color and local gradient features of the character component for finding character candidates. Then character candidates are grouped to detect texts based on the structural features of texts like character alignment, neighbouring characters distance, differences in the size of characters. Yi make use of two algorithms assuming that text string in an image may contain a minimum of three characters. They are adjacent character grouping algorithm and text line grouping algorithm. The adjacent character grouping algorithm performs Hough transformation for fixing lines of text with the centroids of texts.

Adjacent grouping method finds similar type of candidate characters as elements of text and merges the intersecting groups of siblings for forming text string. Cortes[8] proposed a new learning machine for solving two-group classification problems. The machine implements the following idea: input vectors are non-linearly mapped to a very high dimension feature space. In this feature space a linear decision surface is constructed. Special properties of the decision surface ensure high generalization ability of the learning machine.

Gllavata[9] proposed a method in high frequency wavelet coefficients obtained by applying a wavelet transform to the image is used for classification of the text area and non-text area which consider the distribution of high. K-means clustering is used for classifying the text area. Then the text localization is done by performing a projection analysis on the text area. Then the binary text image which is segmented is given as the input for the OCR engine. Jung [10] proposed a technique in which they have an inclination to use a text detection rule that extract six options of texts and build use of modest adaboost classifier with a multi consecutive rummage around for classifying text and non-text region. Six options used square measure variance and expectation of X-Y derivatives, native Gabor filter energy, applied mathematics texture live of image bar graph, activity of variance of riffle constant, edge detection and edge interval calculation, connected part analysis.

Kim[11] described a texture based approach for text detection . The proposed method makes use of a combination of both SVM and CAMSHIFT for detecting and extracting the text efficiently. SVM is used for analyzing the texture. Intensity of raw pixels that forms the textural pattern is delivered as the input to SVM. Then extraction of the texture feature is done. After that the text region identification is done by the CAMSHIFT.

Gatos [12] proposed a method for detecting scene text. In this method an image enhancement and binarization is performed first. After that connected component is analyzed for producing a binary image that mainly contains texts. Firstly, grey level and inverted grey level image is produced. After that they produce binary version of these images by applying image enhancement and binarization. After that a decision function is applied that determines text is present in

these images or not. Then, a connected component extraction is performed. Yin[13] described a technique for text detection by extracting MSER which is considered as the candidate region. Character candidates are grouped for forming text candidates is done by using a single-link clustering algorithm. For this, a self-training distance metric learning algorithm which automatically learns clustering thresholds and learning weights is applied. A character classifier is used for estimating the probability of being a text candidate. The text is then identified by a text candidate.

Koo[14] introduced a machine learning approach for text detection. This method uses two classifiers. One is used produces the candidate text regions and the other separates the non-text areas from text areas. MSER algorithm is used for connected component extraction. Then clustering is applied in which extracted connected components are grouped into form various clusters. For clustering on the basis of adjacency relationship, an Adaboost classifier is trained. Then normalization is performed on the candidate regions that contain words. a multilayer perceptron based classifier was finally used for the refinement of non-text area.

Liu [15] described a method which is a hybrid based approach. At first text edge pixels are extracted using color image edge detection algorithm. External and internal contours are detected from these edge pixels by performing connected component analysis on it. Geometric and gradient features of each region are extracted for constructing text region candidates and for separating non text candidates. Then textural features obtained from wavelet domain are used for verifying candidate text areas. Yi [16] describes a robust hybrid approach of text detection. At first a text estimation map is generated using text region detector. This helps in segmenting text components using local binarization. A conditional random field model is used for filtering out non text components. Finally learning based energy minimization method is used for grouping of text components into text lines.

Support vector machines (SVM) [17] and Naïve Bayes [18] can be used for classifying and training images. Initially, the train images are analyzed to extract the features at some keypoints and are converted into feature descriptors [18] which are high dimensional vectors. Good descriptors must be invariant to all image transforms such as intensity, rotation, scale, and affine transformations. Feature extraction, training, and testing are done separately on each detected region. Feature extraction at various keypoints, which are retrieved using the Harris–Laplace salient point detector [19] is performed. It uses a Harris corner detector and subsequently the Laplace operator for scale selection. The method to segment white regions is based on the maximally stable external regions method [20], which is a region detector that allows detecting bright-on-dark regions in the image.

CONCLUSIONS

The survey shows various text detection and extraction techniques. Even though there are large no of algorithms and techniques for text detection and extraction from image, none of them provide a desired output because of deviation in text. I have looked different methods for efficient detection and extraction of text from natural images. We can choose our technique depending on the current requirements. Here provide a detailed summary of each text detection technique and comparative study.

REFERENCES

1. A. González and L. M. Bergasa, “A text reading algorithm for natural images,” Image Vis. Comput., vol. 31, no. 3, pp. 255–274, Mar. 2013.

2. A. Mogelmose, M. Trivedi, and T. Moeslund, "Vision based traffic sign detection and analysis for intelligent driver assistance systems: Perspec- tives and survey," IEEE Trans. Intell. Transp. Syst., vol. 13, no. 4, pp. 1484–1497,Dec,2012.
3. Z. Hu, "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, Feb. 2013.
4. A. González, M. Garrido, D. Llorca, M. Gavilán, J. Fernandez, P. Alcantarilla, I. Parra, F. Herranz, L. M. Bergasa, M. Sotelo, and P. Revenga, "Automatic traffic signs and panels inspection system using computer vision," IEEE Trans. Intell. Transp. Syst., vol. 12, no. 2, pp. 485– 499, Jun. 2011.
5. C. Yao, X. Bai, W. Liu, Y. Ma, and Z. Tu, Detecting texts of arbitrary orientations in natural images, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Page. 1083–1090 Year of Publication: 2012 CVPR.2012.6247787
6. W. Wu, X. Chen, and J. Yang, "Detection of text on road signs fromvideo," IEEE Trans. Intell. Transp. Syst., vol. 6, no. 4, pp. 378–390, Dec. 2005.
7. Yingli Tian and Chucai Yi, Text string detection from natural scenes by structure based partition and grouping, IEEE Transactions on image processing, vol. 20, no. 9, pp. 2594-2605, 2011.
8. C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn.,vol. 20, no. 3, pp. 273–297, Sep. 1995.
9. J. Gllavata, R. Ewerth, and B. Freisleben, Text Detection in Images Based on Unsupervised Classification of High-Frequency Wavelet Coefficients, Proc. of Int'l Conf. on Pattern Recognition, Cambridge, UK, (page 425-428 Year of Publication : 2004 ICPR.2004.1334146).
10. J. Lee, P. Lee, S. Lee, A. L. Yuille, and C. Koch., Adaboost for text detection in natural scene. In Proc.IEEE Int. Conf. Doc. Anal. and Recogn, pages 429–434, 2011.
11. K. I. Kim, K. Jung, and J. H. Kim, Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm, IEEE Trans. PAMI, vol. 25, no. 12, pp. 1631–1639, 2003.
12. Gatos, B.,Pratikakis, I. & Perantonis, S.J. ,Towards text recognition in natural scene Images, in Proceedings of Int. Conf. Automation and Technology, (Page 354-359 Year of Publication 2005)
13. Y. Pan, X. Hou, and C. Liu, "A hybrid approach to detect and localize texts in natural scene images," IEEE Trans. Image Processing,, vol.20, no. 3, pp. 800–813, 2011.
14. Yi-Feng Pan, Xinwen Hou, Cheng-LinLiu(2009), "Text Localization In Natural Scene Images Based On Conditional Random Field," ICDAR,pp 6-10.
15. J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," in British Machine Vision Conference, 2002, vol. 1, pp. 384–393.
16. C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, Sep. 1995.
17. D. D. Lewis, "Naive (Bayes) at forty: The independence assumption in information retrieval," in Proc. ECML, 1998, pp. 4–15.

18. K. Mikolajczyk and C. Schmid, “Indexing based on scale invariant interest points,” in Proc. ICCV, 2001, pp. 525–531.

19. N. Kulkarni, “Color thresholding method for image segmentation of natural images,” Int. J. Image, Graph. Signal process., vol. 4, no. 1, pp. 28–34, Feb. 2012.

20. J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide baseline stereo from maximally stable extremal regions,” in Proc. BMVC, 2002, pp. 384–393.