ISSN: 2321-7782 (Online)
Volume 2, Issue 12, December 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

Efficient Text and word level recognition using Lexicon Analysis Technique

Dr. K. Swathi[1]
Professor and Head
Department of Computer science & Engineering
Cauvery College of Engineering & Technology
Trichy (TN) – India

S. Uma Maheswari[2]
Scholar
Department of Computer science & Engineering
Cauvery College of Engineering & Technology
Trichy (TN) – India

R. Vijayanathan[3]
Senior Librarian and Head
Department of Library and Information Science
Cauvery College of Engineering & Technology
Trichy (TN) – India

Abstract: Recognition of text in natural scene pictures is changing into a distinguished analysis space owing to the widespread availability of imaging devices in low-priced client product like mobile phones. Detecting text in natural pictures, as hostile scans of written pages, faxes and business cards, is a crucial step for variety of laptop Vision applications, like processed aid for visually impaired and robotic navigation in urban environments. Retrieving texts in each indoor and outdoor environment provides discourse clues for a good kind of vision tasks. During this project, we execute two processes like text detection and text recognition. In text detection, utilize contrast map is then binaries by median filter and combined with Canny's edge map to spot the text stroke edge pixels supported feature extraction. The options extractors are Harris-Corner, Maximal Stable Extremal Regions (MSER), and dense sampling and Histogram of Oriented Gradients (HOG) descriptors. Then implement text recognition. The primary one is coaching a character recognizer to predict the class of a character in a picture patch. The other is coaching a binary character category for every character class to predict the existence of this class in a picture patch. The two schemes are suitable with two promising applications associated with scene text that are text understanding and text retrieval. In additional we tend to extend this idea with word level recognition with lexicon methods with correct results. And additionally recognition text in real time pictures, videos and mobile application pictures.

Keywords: character recognition, text detection, text recognition.

I. INTRODUCTION

Automatic innovation of notice regions is an active investigate area in the design of machine vision systems and is used in many applications such as tourist's assistant systems, mobile robot steering, vehicle license plate detection and recognition. Visions systems are mainly focused on persistently monitoring traffic and observe passing vehicles, take out important features such as vehicle type, color and distinct marks. One feature that separates text from other elements of a scene is its nearly constant stroke width. This can be used to recuperate regions that are likely to contain text. An exhaustive search is then applied to group components into regions and a text level classifier is used for classification of these regions. As an important qualification for text recognition, text detection in natural scene images still remains an open problem due to factors including complex background, low quality images, variation of text content and deformation of text appearance. Recently, Maximally Stable Extremal Regions (MSERs) based text detection has been widely explored. The main advantage of these approaches over other component based approaches is rooted in the effectiveness of using MSERs as character/component candidates. It is based on the observation that text components usually have higher color contrast with their backgrounds and tend to be form homogenous color regions, at least at the character level. The MSER algorithm adaptively detects stable color regions and provides a good solution to restrict the components without explicit binarization. Text detection and recognition in natural scene images has recently received increased attention of the computer vision community. Since text is a pervasive element in many environments, solving this problem has potential for significant impact. Using the SIFT concept to get the text as a voice using the Lexieon analysis.

II. RELATED WORK

In [1] X. Bai, L. J. Latecki, and W.-Y. Liu et al. In this paper, we introduce a new skeleton pruning method based on contour partitioning. Any contour partition can be used, but the partitions obtained by Discrete Curve Evolution (DCE) yield excellent results. The theoretical properties and the experiments presented demonstrate that obtained skeletons are in deal with human visual perception and stable, even in the presence of significant noise and shape variations, and have the same topology as the original skeletons. In particular, we have proven that the proposed approach never produces imitation branches, which are common when using the known skeleton pruning methods. Moreover, the proposed pruning method does not displace the skeleton points. Consequently, all skeleton points are centers of maximal disks. Again, many existing methods displace skeleton points in order to produces pruned skeletons.

In [2] R. Beaufort and C. Mancas-Thillou. With the increasing market of cheap cameras, natural scene text has to be handled in an efficient way. Some works deal with text detection in the image while more recent ones point out the challenge of text extraction and recognition. We propose here an OCR correction system to handle traditional issues of recognizer errors but also the ones due to natural scene images, i.e. cut characters, artistic display, incomplete sentences (present in advertisements) and out- of-vocabulary (OOV) words such as acronyms and so on. The main algorithm bases on finite-state machines (FSMs) to deal with learned OCR confusions, capital/accented letters and lexicon look-up. Moreover, as OCR is not considered as a black box, several outputs are taken into account to intermingle recognition and correction steps. Based on a public database of natural scene words, detailed results are also presented along with future works.

In[3] X. Chen, J. Yang, J. Zhang, and A. Waibel, et al. In this paper, we present an approach to automatic detection and recognition of signs from natural scenes, and its application to a sign translation task. The proposed approach embeds multiresolution and multiscale edge detection, adaptive searching, color analysis, and affine refinement in a hierarchical framework for sign detection, with different emphases at each phase to handle the text in different sizes, orientations, color distributions and backgrounds. We use affine refinement to recover deformation of the text regions caused by an inappropriate camera view angle. The procedure can significantly improve text detection rate and optical character recognition (OCR) accuracy. Instead of using binary information for OCR, we extract features from an intensity image directly. We propose a local intensity normalization method to effectively handle lighting variations, followed by a Gabor transform to obtain local features, and finally a linear discriminant analysis (LDA) method for feature selection. We have applied the approach in developing a Chinese sign translation system, which can automatically detect and recognize Chinese signs as input from a camera, and translate the recognized text into English.

In [4] A. Coates et al. Reading text from photographs is a challenging problem that has received a significant amount of attention. Two key components of most systems are (i) text detection from images and (ii) character recognition, and many recent methods have been proposed to design better feature representations and models for both. In this paper, we apply methods recently developed in machine learning -- specifically, large-scale algorithms for learning the features automatically from unlabeled data -- and show that they allow us to construct highly effective classifiers for both detection and recognition to be used in a high accuracy end-to-end system.

In [5] N. Dalal and B. Triggs, et al. We study the question of feature sets for robust visual object recognition; adopting linear SVM based human detection as a test case. After reviewing existing edge and gradient based descriptors, we show experimentally that grids of histograms of oriented gradient (HOG) descriptors significantly outperform existing feature sets for human detection. We study the influence of each stage of the computation on performance, concluding that fine-scale gradients, fine orientation binning, relatively coarse spatial binning, and high-quality local contrast normalization in overlapping descriptor blocks are all important for good results. The new approach gives near-perfect separation on the original MIT pedestrian database, so we introduce a more challenging dataset containing over 1800 annotated human images with a large range of pose variations and backgrounds.

III. HISTOGRAM OF ORIENTED GRADIENTS (HOG)

Histogram of Oriented Gradients (HOG) is attribute descriptors utilized in computer vision and image processing for the principle of entity discovery. The method calculates occasion of grade direction in restricted segment of a picture. Local thing facade and form inside a picture can be explained by the allocation of concentration gradients or edge directions. The execution of these descriptors can be attained by separating the picture into tiny related regions, called cells, and for every cell accumulating a histogram of gradient directions or edge orientations for the pixels inside the cell. The grouping of these histograms then symbolizes the descriptor. For enhanced accurateness, the local histograms can be contrast-normalized by manipulative a measure of the passion across a larger region of the picture, called a block, and then utilizing this value to regularize all cells inside the block. These normalization consequences in improved invariance to modify in clarification or shadowing.

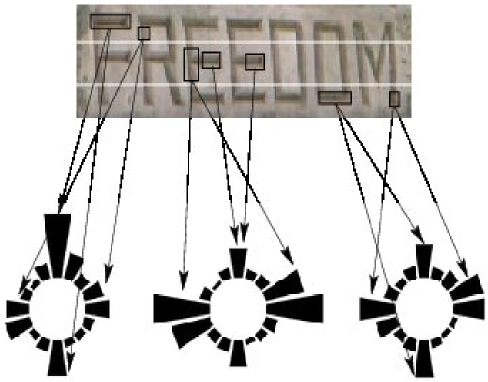

Figure 1: Detect text using hog

Harris Corner Detector

A corner can be definite as the connection of two edges or a point. It is junctions of curves. Normally corner points are extra steady features greater than modify of viewpoint. Corner recognition is extensively utilized in computer vision application such as tracking, image matching and movement discovery. Harris corner detector is utilized to extort the corner points. The Harris corner detector is a well-liked notice point detector. Since there is no result of illumination variation, image noise, scale and rotation on the presentation of Harris corner detector. It is supported ahead the local auto-correlation function of a indication, where the local auto-correlation function actions the local alter of the signal with scrap move by a small amount in dissimilar instructions.

Maximally Stable Extremal Regions (MSERs)

Lately, Maximally Stable Extremal Regions (MSERs) found text detection has been extensively discovered. The major improvement of these methods in excess of extra part stands methods is entrenched in the efficiency of utilizing MSERs as character/component candidates. It is found on the surveillance that text components typically have higher color contrast with their backgrounds and be predisposed to be type homogenous color area, at slightest at the character level. The MSER algorithm adaptively notice steady color regions and offers a fine explanation to focus the components with no unambiguous binarization. MSER regions are of all-purpose, data-dependent shape, i.e. composite adequate to offer enough restraint to describe affine frames. They are related, randomly shaped, probably nested, and do not cover the whole picture, i.e. they do not form a separation. The formal description of MSERs and a complete explanation of the extraction algorithm are given in. MSER presentation estimation and association to extra detectors can be established in.

Figure 2: Detect text using MSER

IV. PRINCIPLE COMPONENT ANALYSIS (PCA)

Principal Components Analysis is a correlation-based method that select set of delegate dimensions described the principal components based on the degree of difference that they confine from the original data. In exacting, PCA calculates the output points by performing singular value decomposition (SVD) on the document covariance matrix and then multiplies the resulting eigenvectors with their matching eigenvalues. PCA is a technique that decreases data dimensionality performing a covariance examination among factors. The original data will be curved into a fresh coordinate system supported on the difference in the data. PCA relates a mathematical process for changing an amount of interrelated variables into an amount of uncorrelated variables described principal components. The first principal component accounts for as a lot of the inconsistency in the data as probable, and every following component accounts for as a lot of the outstanding unpredictability as possible. PCA is helpful when there is data on a huge number of variables, and there is a number of idleness in those variables. In this case, redundancy means that a number of of the variables are associated with one another. And since of this redundancy, PCA can be utilized to decrease the experiential variables into a lesser amount of principal components that will account for most of the variation in the experimental variables.PCA is suggested as an investigative tool to expose unknown tendency in the data. The method has establish application in fields such as face detection and image density, and is a ordinary method for discovery patterns in data of high dimension.

Linear Discriminant Analysis (LDA)

The Linear Discriminant Analysis (LDA) technique has been useful for decades for dimension decrease of clustered data in pattern identification. It is typically formulate as an optimization difficulty on scatter matrices. A serious drawback of the LDA is that its object function need that the whole scatter matrix be nonsingular. In a lot of modern data mining troubles such as information recovery, facial detection, and microarray data investigation, the total scatter matrix in query can be extraordinary because the data items are from an extremely high-dimensional space, and in all-purpose, the dimension beats the amount of data points. This is recognized as the below sampled or wonder trouble. Linear Discriminant Analysis (LDA) utilizes class data in order to divide well the classes. Newly, the computer visualization community has effectively planned numerous alternatives of LDA that unnaturally pull separately the positive and the negative examples. LDA search for the most excellent division of data from dissimilar classes by reducing the within-class distance and exploit the between-class distance at the same time. LDA execute well in several applications. Due to LDA’s excellent properties and the wants in streaming data mining, incremental LDA draws extra and additional interest. In positions where data come in stream, updating the explanation to LDA with the received data is preferred, because it evades the time-consuming batch-mode recalculation of LDA solution. During the past few years, various incremental LDA algorithms have been urbanized. The majority of them provides estimated solutions and suffers from high computational cost. There are a few learning on incremental LDA utilizing neural networks up till now frequently experience from slow meeting and rigorously challenge the consequence of incremental learning.

Scale-invariant Feature Transform (SIFT)

The images were processed to produce keypoints by mainly utilizing Scale-invariant Feature Transform (SIFT) technique. SIFT (Lowe, 1999) was developed for image feature generation for constructing feature vectors which are invariant to image translation, scaling, rotation, and illumination change. For this approach, robust object recognition can be reached in cluttered partially included images. Extracted features are utilized for solving the problem of recognizes images in various viewpoints, contrasts, and luminance. The SIFT algorithm takes an image and transforms it into a collection of local feature vectors. Every of these feature vectors is invented to be distinctive and invariant to translation, image scaling, rotation and partly invariant to illumination changes and affine or 3D projection. As being widely used in different research (cite), SIFT were proven to be a robust method for object recognition in different image scenes. The SIFT features share a number of properties in common with the responses of neurons in inferior temporal (IT) cortex in primate vision.

Hidden Markov Models (HMM)

MSER regions are of general, data-dependent shape, i.e. complex enough to provide sufficient constraints to define affine frames. They are connected, arbitrarily shaped, possibly nested, and do not cover the entire image, i.e. they do not form a partitioning. The formal definition of MSERs and a detailed description of the extraction algorithm are given. MSER performance evaluation and comparison to other detectors can be found. Template comparison methods of speech recognition directly compare the unknown utterance to known examples. Instead HMM creates stochastic models from known utterances and compares the probability that the unknown utterance was generated by each model. HMMs are a broad class of doubly stochastic models for a non-stationary signals that can be inserted into other stochastic models to incorporate information from several hierarchical knowledge sources. Since we do not know how to choose the form of this model automatically but, once given a form, have efficient automatic methods of estimating its parameters, we must instead choose the form according to our knowledge of the application domain and train the parameters from known data. In voice recognition HMMs are used to model a non-stationary signal. However they have been used in a variety of fields such as language, financial, and biological modeling.

V. Conclusion

Text detection in natural scene images remains a challenging problem due to complex background, low image quality and/or variation of text appearance. In proposed presented a technique of scene text recognition from identify text regions, which is well-matched with mobile applications. It identifies text regions from image or video and distinguishes text information from the identify text regions. In scene text detection, describe analysis of color disintegration and horizontal alignment is performed to search for image regions of text strings. In scene text recognition, two methods, text understanding and text retrieval, are correspondingly proposed to take out text information from surrounding location. Using the SIFT technique to get the text from the image and then convert it voice using the HMM technique to the text as voice.

References

1. X. Bai, L. J. Latecki, and W.-Y. Liu, "Skeleton pruning by contour partitioning with discrete curve evolution," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 3, pp. 449–462, Mar. 2007.
2. R. Beaufort and C. Mancas-Thillou, "A weighted finite-state framework for correcting errors in natural scene OCR," in Proc. 9th Int. Conf. Document Anal. Recognit., Sep. 2007, pp. 889–893.
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. 2004.
4. A. Coates et al., "Text detection and character recognition in scene images with unsupervised feature learning," in Proc. ICDAR, Sep. 2011, pp. 440–445.
5. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2005, pp. 886–893.

6. T. de Campos, B. Babu, and M. Varma, "Character recognition in natural images," in Proc. VISAPP, 2009.
7. B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in Proc. CVPR, Jun. 2010, pp. 2963–2970.
8. P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan,"Object detection with discriminatively trained part-based models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, pp. 1627–1645, Sep. 2010.
9. T. Jiang, F. Jurie, and C. Schmid, "Learning shape prior models for object matching," in Proc. CVPR, Jun. 2009, pp. 848–855.
10. S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Johsi, "Text extraction and document image segmentation using matched wavelets and MRF model," IEEE Trans. Image Process., vol. 16, no. 8, pp. 2117–2128, Aug. 2007.

AUTHOR(S) PROFILE

Dr.K.Swathi obtained her under-graduation in B.E., (Computer Science & Engineering) from Bharathidasan University, Trichy in 1999. She obtained her M.E. degree in Computer and Communication Engineering from Anna University, Chennai in 2004.She obtained Ph.D degree in Faculty of Information & Communication Engineering from Anna University, Chennai in 2014. Presently, she is working as Associate Professor in Computer Science & Engineering, Cauvery College of Engineering and Technology, Trichy in 2008. She has 14 years teaching experience and also she had attended many workshops, seminars and conferences on Research issues in Image processing. She has published papers in international journals and presented papers in various Conferences. She is the life member of Indian Society for Technical Education (ISTE). Her areas of interest include Image processing, Data mining, network security and Software engineering.

S. Uma Maheswari was born in Trichy, Tamilnadu, India, in 1992. She received the B.E. degree in Computer Science and Engineering from Saranathan engineering college, Trichy affiliated to Anna University Trichirapalli, India, in 2013, and Currently she is pursuing her M.E. degree in computer science and engineering at Cauvery Engineering college, Trichy affiliated to Anna university, Chennai., India, Her area of interest is Image Processing.

Vijayanathan. R. received his Master of philosophy Library and Information Science from Annamalai University in 1999 Also he obtained his post-graduate degree in Master of Economics in Bharathidasan University in 1996 and PG Diploma in Computer Application ,PG Diploma Tourism Management and PG Diploma in co- operative management from Annamalai Unversity .He is working as a Sr. Librarian, Department of library and Information Science in Cauvery College of Engineering and technology, Trichy from 2009. He has guided many M.Phil scholars produced and member in various Universities in Tamil Nadu. He served more than 14 years as Senior Librarian in reputed Engineering College and also he had attended many workshops, seminars and conferences on Research issues in Bibliometric ananalysis.He has published so many research papers in National and International journals. His areas of interested are networking; Cloud computing, IC Technologies, Environmental study, Library Automation, Webometric study, Scientometric study, Bibliometric analysis, and citation study.