VOL. 10, NO. 8, MAY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

ADVANCED TEXT CLASSIFIER BASED ON THE CONCEPT OF PIPELINING AND CC CLUSTERING

P. S. Ramesh[1] and S. Letitia[2]
[1]Department of Computer Science and Engineering, Arunai Engineering College, Thiruvannamalai, Tamilnadu, India
[2]Department of Electronics and Communication Engineering, TEQIP Directorate of Technical Education, Guindy, Chennai, Tamilnadu, India
E-Mail: psramesh7@gmail.com

ABSTRACT

Image processing is a way to alter an image into digital structure and carry out a few operations on it, in order to obtain an improved image or to take out a little valuable information from it. In this paper, testing has been made to pull out text from remote sensing images. The task is to separate text from the non-text area of an image. The separated text is converted into speech. And the way of segregation of text and translating into speech is supportive in lots of ways, such as guiding visually impaired persons; help to navigate tourists, etc., the citation of text area from a remote sensing image is done by maximally stable extremal region algorithm and the Ada boost classifier. The overall process of text extraction has to be done and achieved through the major concept called pipelining. Varied algorithms have been used in implementing this technique. Traditionally Region-based algorithm was used and it had lots of hindrance mainly the decisions were made locally. Hence connected component has been introduced which reduces the disadvantages of Regional-based algorithm to a certain extent. The main purpose of the experiment is to overcome the disadvantages like photometric degradation, geometric distraction. The result is to extract text area with accuracy and speed and also to overcome the disadvantages of previous techniques.

Keywords: machine learning classifier, connected component (CC)-based approach, MSER algorithm, AdaBoost learning, binarization.

1. INTRODUCTION

A 2D picture which is commonly processed by a computer system is referred to as image processing. In computer graphics, images are manually made from structural properties of objects, and their surroundings, and lightning effects, instead of being acquired from natural scenes, as in most animated. It alludes to focus, or sharpening of image characteristics such boundaries, disparity to make a graphic display more beneficial for display. Noise usually quantified by the percentage of pixels which are degraded. Degraded pixels are either fix to the highest value or have only one bits. There are many ways that noise can be found. The widely used image processing algorithms include noise filtering and several color correction operations. The mixture of algorithms is organized into an image reconstruction pipeline.

Nowadays mobile devices equipped with high-resolution digital cameras are widely available, research activities using these devices in the field of human computer interaction (HCI) have received much attention for the past years. Within that, text deduction and identification in camera captured images have been considered as vital pitfall in computer vision community. The main reason is a text information can be easily recognized by machines and can be used in a variety of applications. Some examples are support for visually impaired people, translators for tourists, and information retrieval systems in indoor and outdoor environments. Although there exist a lot of research activities in this field, scene text detection still remains as a challenging problem. A hierarchical structuring of relations may conclude in providing many classes and a more confused structure to implement. Therefore it is advisable to convert the hierarchical relation structure to a easy structure such as a classical flat one. It is rather direct way to transfer the built hierarchical model into a bipartite, simple model, including of classes on the one hand and flat relations on the other. Simple relations are preferred at the design level for reasons of simplicity and implementation ease. There is no unique or functionality associated with a plane relation.

In proposed system a innovative picture text recognition algorithm depend on two device knowledge classifiers: one allows us to produce candidate word regions and the other removes out non text ones. To be general, the connected components can be extracted (CCs) in images with the help of maximally stable extremal region algorithm. These deduced CCs are divided into clusters an AdaBoost classifier can be train that regulates the neighboring correspondence and cluster CCs by using their pair wise relations. Then candidate word regions can be normalized and intent, whether each region contains text or not. But the level, tilt and shade of each text region can be probable from CCs, a text/non text classifier for normalized images has to be developed. This classifier is based on Kernel discriminate analysis is a method used in statistics and device knowledge to find a straight line mixtures of characteristics which best characterize or separates two or more classes of objects or events. This classifier is based on Kernel discriminate analysis is a method used in statistics and machine learning to find a linear combination of features which best characterizes or separates two or more classes of objects or events.

2. RELATED WORK

Kwang in kim, Keechul Jung and Jin Hyung Kim [1] CC-based methods segment an image into a set of CCs, then classify the final CCs as either text or background by analyzing their geometrical characteristics. Specifically, the system uses a small window to scan the input image, by analyzing its textural properties using an SVM.

Jerod J. Weinman, Erik Learned Miller, Allen R. Hanson [2] the problem of optical character recognition (OCR), or the recognition of text in machine-printed documents. A probabilistic graphical model for STR that brings together bottom-up and top-down information into a single elegant framework.

Yi-Feng Pan, Xin Wen Hou, Cheng- LinLiu [3] Text detection and localization in natural scene images is an important issue for content-based image analysis. To combine unary component properties and binary contextual component relationships, a conditional random field (CRF) model is used.

Chucai Yi, Ying Li Tian [4] Dissimilar from certificate images, from that passage characters are normalized into graceful poses and proper resolutions, sizes, and orientations into complex background. A new framework to detect text strings with arbitrary orientations in complex natural scene images.

Xilin Chen, Jie Yang, Jing Zhang, Alex Waibel [5] the primary challenge lies in the variety of texts: it can vary in character size, direction and situation. Gabor Transform can be used to find local intensity normalization methods to efficiently handle lighting variations.

Chucai Yi,Ying Li Tian [6] Existence of background outliers resembling text characters, such as windows, bricks, and character-like texture. A new proposed algorithms to combine the structural analysis of text stroke with color assignment and filter out background interferences.

Chucai Yi, Ying Li Tian [7] The CC-based methods can extract text efficiently, but have difficulties when text touches itself or other graphical objects. A simplistic solution would be to perform the text detection, pixel by pixel and then verify the following text blocks between consequence pixels.

Yu zhong, Hongjiang zhang and Anil K. Jain [8] the same text may vary its size from frame to frame, due to some special effects. I-frames are compressed using Discrete Cosine Transform (DCT) of local blocks to reduce spatial redundancy.

3. PROPOSED ALGORITHM

Here we present the proposed clustering algorithm. First, we take out the text from images based on MSER algorithm. And then we find a neighbouring association based on the classifier. A cluster component will be produced based on the neighbouring relationship. In this part discusses issues relating to extract text from images and various other implementation issues.

Enhanced argument items and recurrently owned annotations

A. Creation of latest argument items

The dataset consists of natural images annotated with bounding boxes just about each occurrence of a statement. While it contains adequate data in concert assessment we require more information in support of the guidance of our classifiers. We use a classifier that tells us the neighbor relation between CCs in the candidate generation step, and CC-level information is vital for the guidance of such a classifier. Consequently, we improved argument fact to facilitate training set by totaling pixel-level interpretation (binarization results) and text-line information.

B. Text invention

Intended for the candidate generation, CCs in images can be brought out and organize the alienated CCs into clusters, where clustering algorithm is based on a neighbor relationship. From a huge number of CC separation techniques, we have acquired the MSER algorithm for the reason that it shows good concert with small calculation expenditure. To find local binarization results that is stable over a range of thresholds. The MSER algorithm yields CCs that are either darker or brighter than their surroundings. An AdaBoost classifier explains that whether it is adjacent or not. It combines a set of weak classifiers to make a strong classifier. AdaBoost generates and calls a new frail classifier in each of a series of rounds b=1.Applying an Adaboost knowledge algorithm. We use connected factor clustering. Helps to find all neighboring pairs by assessing for all the likely pairs. Based on these neighboring pairs, it is categorized into a set of clusters.

C. Text area normalization

After CC clustering, we have a set of clusters. We normalize corresponding regions for the reliable text/non text classification. For geometric normalization document's skew checking algorithm is used. It depends on the Hough line alteration. Document binarization is a preprocessing task, very useful to document analysis systems. It automatically converts the document images in a bi-level form.

D. Filtering process

The developed text/non text classifier omits non text blocks among normalized images. The poor binarization results should be discarded. Based on the results the text and non text regions are classified into square patches and it should be trained by a kernel discrimination analysis for better classification of text from images. After classification it will be partitioned text and non text. Finally, to integrated the square classification results. The Text from images should be highlighted.

E. Verbal communication creation

A text-to-speech (TTS) method converts regular language text into verbal communication other systems

cause to be symbolic linguistic representations like phonetic transcriptions into speech. Synthesize speech can be produced by concatenating pieces of recorded speech that are stored in a database. Systems alter in the extent of the stored verbal communication units; a system that stores phones or diaphones afford the leading yield range, but may be short of clearness For precise usage domains, the storage space of whole terms or sentences allows for first-rate output. On the other hand a synthesizer can integrate a representation of the vocal zone and other human voice distinctiveness to create a completely "imitation" voice output. The most significant traits of a verbal communication scheme should naturalness and clearness. Naturalness describes how narrowly the productivity sounds like human speech, while clearness is the ease with which the output is understood.

4. SYSTEM DESIGN

Figure-1 demonstrates the detection of text in image algorithm depends on machine learning techniques. In broad a two cluster partitions are developed: one partition was designed to generate text region using AdaBoost learning and the other categorizer was for the removing of non text candidates using Kernel discriminant analysis. To be certain, it takes out the connected components (CCs) in images with the help of maximally stable extremal region algorithm. These separated CCs are classified into clusters so that it creates text regions. An AdaBoost categorizer was skilled that concludes the neighbour relationship and cluster CCs by using their pairwise relations. Then it normalizes candidate word regions and decides whether each region contains text or not. In view of the fact that the level, slant, and shade of each candidate can be estimated from CCs, we develop a text/non text categorizer for normalized images. This categorizer is based on Kernel discriminant analysis is a technique used in statistics and machine learning to find a linear combination of features which most excellent characterizes or separates two or more classes of objects or events. After extracting the text from images it can be converted into verbal form which is audible. For converting into speech format it should be first consider the area of text region and the way to pronounce the text. Based on these criteria a text from an image can be extracted thoroughly. This shows the overall detail explanation of the working system. Each phase of the deduction of text area can be done thoroughly with the help of pipelining concept. Each step must be overlap with one another. Full explanations of the proposed text extraction algorithm will be presented.

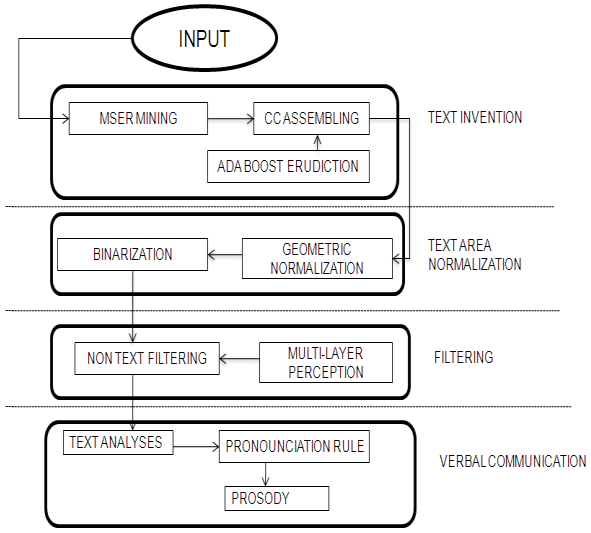

Figure-1. System architecture.

A. Text invention

Intended for the text invention, CCs in images can be carry out and systematize the alienated CCs into clusters, where clustering algorithm is based on an neighbor relationship.

CC Mining

From a huge number of CC separation techniques, we have acquired the MSER algorithm for the reason that it shows good concert with small calculation expenditure. Based on this algorithm a local binarization results can be establish that are steady over a plentiful amount of thresholds. The MSER algorithm produces CCs that are moreover darker or brighter than their surroundings

AdaBoost erudition

Based on the composed samples, an AdaBoost categorizer can be skilled that tells us whether the connected component of the text area is nearby or not. Intended for the samples of the technique, first identify a few narrow properties of CCs. Given a pair of clusters we first find the level distance, level overlap, and perpendicular overlap between two boxes of the each character.

CC assembling

After applying Ada boost learning algorithm. Connected component clustering are used. It can find all adjacent pairs by evaluating for all the possible pairs. Based on these adjacent pair, it is partitioned into a set of clusters. Connected component is mainly used to find the adjacent pixels in a graph. If two pixels are adjacent which should not connected to any other pixels in an image.

B. Text area normalization

After CC clustering, we have a set of clusters. We normalize corresponding regions for the reliable text/non text classification.

Geometric normalization

Intended for geometric normalization we use document's skew checking algorithm, that is depends on Hough line transformation. The algorithm is based on probing for text base lines - black line of text bottoms followed by white line below.

Given word region we first localize its corresponding region. Yet text boxes can practice outlook distortion it can fairly accurate the form of text boxes with parallelograms that's left and right sides are parallel to y-axis. This approximation alleviates difficulties in estimating text boxes having a high degree of freedom (DOF): we only have to find a skew and four boundary supporting points [11]. It avoids all difficulties for evaluating the characters occurred in an image. It calculating all the characters based on the height and width of each and every letter of the word.

To estimate the skew of a given word area, we build two sets from the top-center point and the bottom center point of a bounding box of a text region.

Binarization

Document binarization is a preprocessing task, very useful to document analysis systems. It automatically converts the document images in a bi-level form in such way that the foreground information is represented by black pixels and the background by white ones.

However, we perform the binarization separately by estimating text and background colors. It is because (i) the MSER results may miss some character components and/or yield noisy regions (mainly due to the blur) and (ii) we have to store the point information of all CCs for the MSER-based binarization [11].

C. Filtering process

We develop a text/non text categorizer that reject non text blocks among normalized images [11].

Kernel discriminate analysis

Intend for the regulation we require normalized images. Intend for this aspiration we applied our algorithm offered in the preceding sections (i.e., candidate generation and normalization algorithms) to the guidance images. Then, we by hand classified them into text and non text. We eliminated a few images screening reduced binarization results, and collected text block images and non-text chunk images [11]. On the other hand we have set up that further pessimistic samples are required for the dependable refusal of non text mechanism and collected more pessimistic samples by applying the same process to images that do not contain any text. These text/non text images are alienated into squares and we have accomplished a kernel discriminate analysis for the arrangement of square patches.

Incorporation of choice results

Later than organization it will be detachment text and non text. It reflects on only the text. Finally, to incorporated the square classification results. In addition to highlight the text from the scene picture.

5. DISCUSSIONS

We tried to build a text/non text classifier based on normalized gray-scale images (without binarizaton), because gray-scale images appear to be more instructive. To be exact we adopted the AdaBoost erudition scheme and incline characteristics. On the other hand experiments have shown that this approach yielded approximately the similar concert with a extensive total of overhead in guidance. We also tried to use both classifiers in a row (neural network with binary images and Adaboost with gray images), however, we could not find obvious gains.

6. RESULTS

A. Word separation analytics

Even though word separation is not the major concern of the scene text detection trouble, it is vital in the assessment. Hence, we have developed a analytics law that divides deducted text boxes into words. Given a cluster we arrange distances between adjacent CCs in a normalized image in downward order and approximate a (minimum) word spacing distance S. If the distance between adjacent CCs is over S, a being apart between them occurs.

B. Exploitation of multichannel information

The MSER algorithm extorts CCs in a scalar valued image. Consequently our technique occasionally suffers from the missing of information during color lessening. Note that our scheme is indifferent to affine transform of amount and we can apply our technique to the chrominance channel without any alteration. In case that alike (significantly overlapping) text blocks are deducted in more than one channel, we have chosen one text block having the leading textness value. One may think that CC extraction methods in multi-channel images can lessen the same problem, even if they are much slower than single channel methods. However, they may yield over-segmented results.

C. Conversion of text to speech

After extracting the text from images the given text can be converted in speech format by considering the pronunciation and prosody rule.

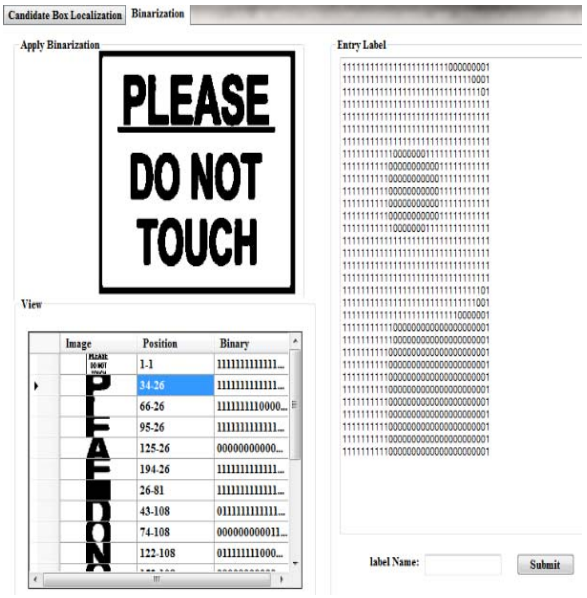

Figure-2. Binarization.

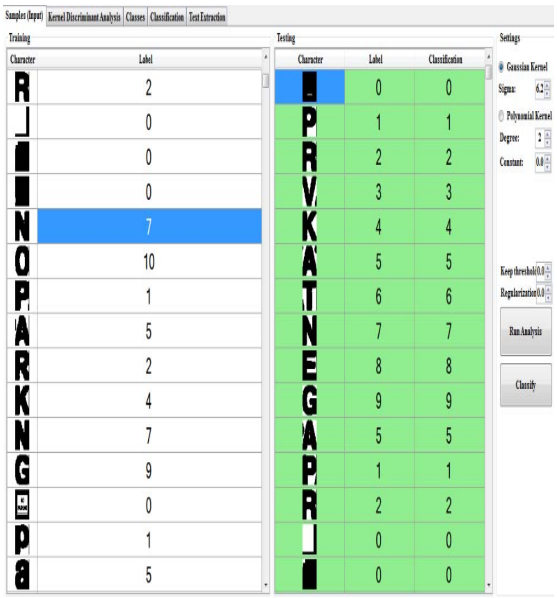

Figure-3. Classifier.

Table-1. Kernel discrimination analysis.

Component	Eigen value	Proportion	Cummulative proportion
0	7,679,337,723..	0.10918	0.10918
1	7,412,364,086..	0.10539	0.21457
2	6,992,699,926..	0.09942	0.31399
3	6,576,703,561..	0.09351	0.40750
4	6,483,136,951..	0.0918	0.49968
5	6,011,734,688..	0.08547	0.58515
6	4,461,670,515..	0.06344	0.64589

Table-1 shows all the component values of kernel discrimination analysis to find the cumulative proportion of each component. Based on the data a component proportion chat can be drawn which is shown in Figure-4.

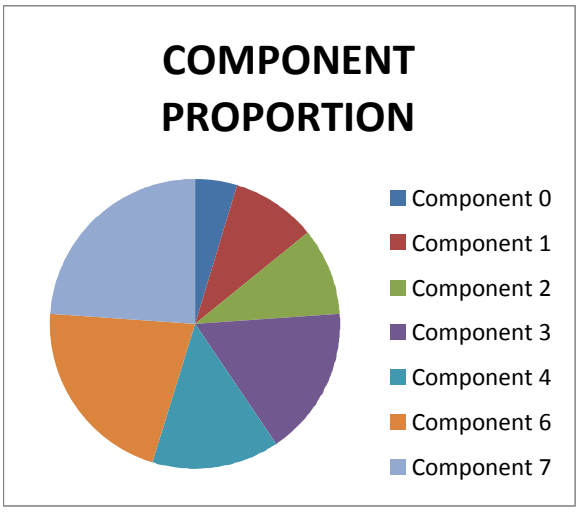

Figure-4. Component proportion chart.

7. CONCLUSIONS

In this paper a new scene text detection algorithm based on two machine learning classifiers: one allows us to generate candidate word regions and the other filters out non text ones has been proposed. To be specific, it extracts connected components (CCs) in images by using the maximally stable external region algorithm. These deductions of CCs are classified into clusters so that we can produce candidate regions. And then an AdaBoost classifier can be trained that determines the adjacency relationship and cluster CCs by using their pair wise relations. Then it normalizes candidate word regions and determines whether each region contains text or not.

It Overcome the problem of photometric degradation and geometrical distortions. Speed should be increased by concept of pipelining. Non-text region can be omitted perfectly with this algorithm. Parallelism and performance also increased.

REFERENCES

[1] K. Jung, “Text information extraction in images and video: A survey,”Pattern Recognit. vol. 37, no. 5, pp. 977-997, May 2004.

[2] J. Liang, D. Doermann, and H. Li, “Camera-based analysis of text and documents: A survey,” Int. J. Document Anal. Recognit. vol. 7, nos. 2-3, pp. 84-104, 2005.

[3] J. Zhang and R. Kasturi, “Extraction of text objects in video documents: Recent progress,” in Proc. 8th IAPR Int. Workshop Document Anal. Syst. Sep. 2008, pp. 5-17.

[4] S. Lucas, A. Panaretos, L. Sosa, A. Tang, S. Wong, and R. Young, “ICDAR 2003 robust reading competitions,” in Proc. Int. Conf. Document Anal. Recognit. 2003, pp. 682-687.

[5] S. Lucas, “Icdar 2005 text locating competition results,” in Proc. Int. Conf. Document Anal. Recognit. 2005, pp. 80-84.

[6] A. Shahab, F. Shafait, and A. Dengel, “ICDAR 2011 robust reading competition challenge 2: Reading text in scene images,” in Proc. Int. Conf. Document Anal. Recognit. 2011, pp. 1491-1496.

[7] X. Chen and A. Yuille, “Detecting and reading text in natural scenes,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2004, pp. 366-373.

[8] X. Chen and A. Yuille, “A time-efficient cascade for real-time object detection: With applications for the visually impaired,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Workshops, June 2005, pp. 1-8.

[9] C. Yi and Y. Tian, “Text string detection from natural scenes by structure-based partition and grouping,” IEEE Trans. Image Process. vol. 20, no. 9, pp. 2594-2605, September 2011.

[10] 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. Jun. 2012, pp. 1083-1090.

[11] Hyung II Koo and Duck Hoon Kim“Scene Text Detection via Connected Component Clustering and Nontext Filtering”. vol. 22, no 6, June 2013.