International Journal of Latest Trends in Engineering and Technology
Special Issue April-2018, pp. 101-105
e-ISSN:2278-621X

SCALE AND ROTATION INVARIANT TECHNIQUES FOR TEXT RECOGNITION IN MEDIA DOCUMENT

Ms. Akila K[1], Dodla Manasa[2], Katrakuti Ramya[3], Ballela Suseela[4]

Abstract- This paper analyzes, compares and contrasts technical challenges, methods and the performance of text detection and recognition research in the color imagery. It summarizes the fundamental problems and enumerates factors that should be considered when addressing these problems. Existing techniques are categorized as either stepwise or integrated and sub-problems are highlighted including text localization, verification, segmentation and recognition. Special issues associated with the enhancement of degraded text and the processing of video text, multi-oriented, perspectively distorted and multi-lingual text are also addressed. The categories and sub-categories of text are illustrated, benchmark datasets are enumerated, and the performance of the most representative approaches is compared. This review provides a fundamental comparission and analysis of the remaining problems in the field.
Keywords – Text detection, Text localization, Text recognition, Survey.

1. INTRODUCTION

Optical Character Recognition (OCR) is a well known process for converting text images to machine editable test format. During the past few decades significant research work is reported in the OCR system. In English there is significant amount of research have been done for languages like Chinese and Japanese. OCR gained so much research interest because of its potential applications in post offices, defense and bank organizations. The first part of this project contains introduction to image segmentation and boundary analysis. We have also tried to elaborate key points which allows to understand quickly the essence of boundary analysis in the rest of the part we have put forth some aspects of theory, algorithms of boundary analysis and also the limitations on the algorithms.

2. PROPOSED ALGORITHM

2.1 Feature Extraction –

At first the centriod of the character image is computed from the foreground pixels. This global centriod is referred as 'Cg'. The character is then partitioned into multiple concentric circles considering the global centriod as the center of each circle.

2.2 Contour Analysis-

A boundary of an object in an image is a set of pixels which separates out the object from background. In Computer System vision coding of a boundary used the code of freeman, two-dimensional coding, polygonal coding are most known.But all these formats of coding are not used ina a contour analysis. However in this project we use the chain code in EmguCV (cross platform/ .net wrapper to OpenCV) for finding contours. It uses correlational algorithms for boundary analysis. The boundary is encoded by the sequence of complex numbers.

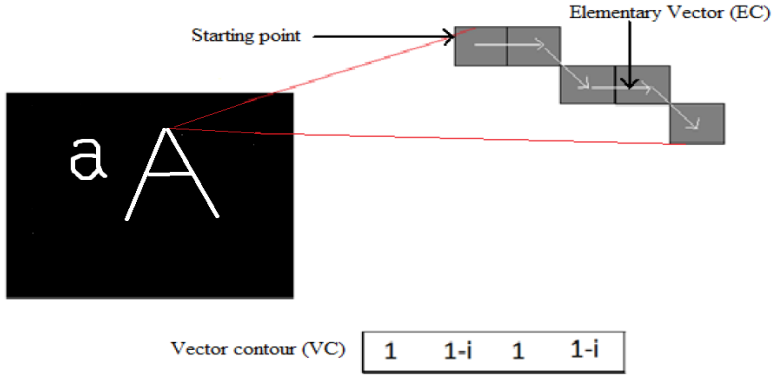

Figure 1. Complex Number Encoding On Contour

[1] Assistant Professor, Computer Science and Engineering, RMK College of Engineering and Technology, Chennai, Tamil nadu, India
[2] Computer Science and Engineering, RMK College of Engineering and Technology, Chennai, Tamil nadu, India
[3] Computer Science and Engineering, RMK College of Engineering and Technology, Chennai, Tamil nadu, India
[4] Computer Science and Engineering, RMK College of Engineering and Technology, Chennai, Tamil nadu, India

2.3 Normalized Scalar Product-.
The Normalized Scalar Product can be calculated as follows:
η =(Γ,N)/(|Γ ||N|)
Where |Γ| and |N| - the norms (length) of boundaries calculated as:

$$|\Gamma| = \left(\sum_{n=0}^{k-1} |\gamma_n|^2\right)^{\frac{1}{2}}$$

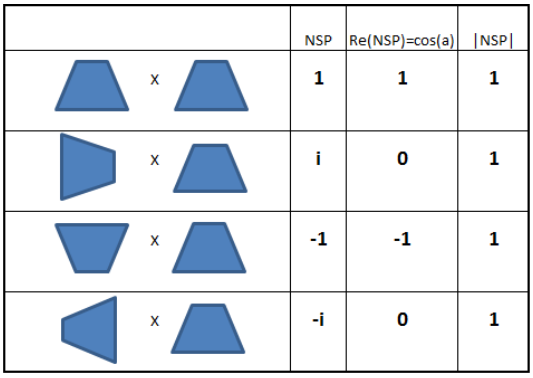

Figure 2. Properties Of Normalized Scalar Product Of Contours

2.4 Correlaation Function of Boundary-
Let's introduce the concept of inter correlation function (ICF) of two boundaries:
τ (m) = (Γ,N(m)), m = 0, … , k-1 (7)
Where N(m) - a boundary received from N by cycle shift by its EV on m of elements.
For an example, if N = (n1, n2, n3, n4), N(1) = (n2, n3, n4, n1), N(2) = (n3, n4, n1, n2) and so on.

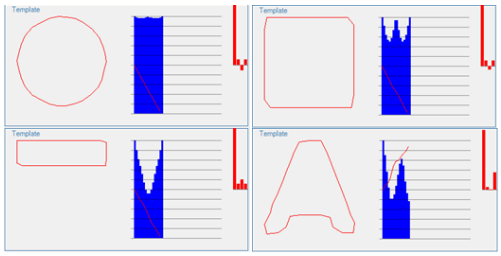

Figure 3. Graphical Representation Of Contours

Let's discover the magnitude having the maximum norm among values an ICF:

$$\tau_{max} = \max\left(\frac{\tau(m)}{|\Gamma||N|}\right), \quad m = 0, \ldots, k\text{-}1$$

2.5 Edge Detection And Filterting-
After processing, the system uses canny edge detection algorithm which performs better than other edge detectors. From smoothed gray scale image, the systems computes the gradient components, estimates the edge strength and orientation of the edge normal respectively. All visited points in the connected contour found are stored for contour analysis. After canny edge detection the contour is generated with unwanted edges and various noises. The system calculates length and area where, a contour length is the total number of the contour pixels and contour area is the maximum (height*width) of the closed contour shape. The system filters out the unwanted edge by intersecting with the minimum contour length (K<200) and minimum contour area ((height*width)<400) from the cammy edge detected image.

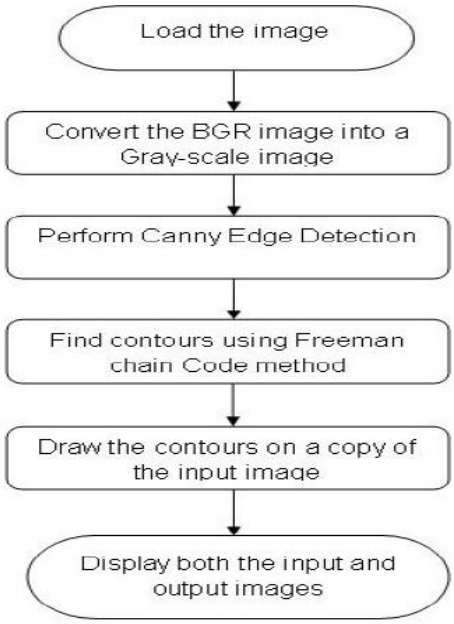

Figure 4. Architecture For Character Recognition

3. EXPERIMENT AND RESULT

3.1 Text In Imagery

Graphic text and scene text are considered as two basic classes of text, where the former refers to machine print text overlaid graphically and the later refers to text on objects, captured in its nature environment.

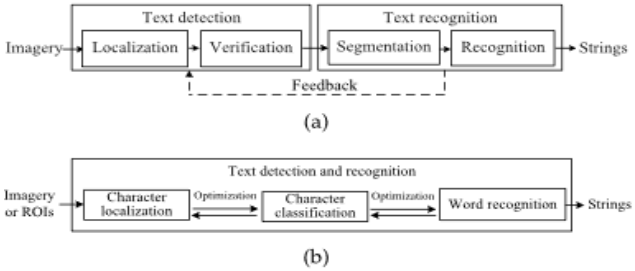

Figure 5. Frameworks of two commonly used text detection and recognition methodologies a)Stepwise Methodology, b)Integrated methodology

a)Stepwise Methodology have four primary steps: Localization, Verification, Segmentation, Recognition.
b)Integrated Methodology have three primary steps: Character Localization, Character Classification, Word Recognition.

3.1.1 Text Localization

The objective of text localization is to localize text components preciseely as well as to group them into candidate text regions. For text localization, connected component analysis (CCA) and sliding window classification are two widely used methods are color,edges, strokes and texture are typically used as features.

3.1.2 Text Verification

The text localization often introduces false positives because a small piece of components/patches may not contain sufficient information for classification. After text localization, holistic features of text regions are available for precise classification and verification. Prior knowledge about color, size and shape consenses, and projectionprofile have been used to perform text verification. Horizontal and Vertical projection profiles are used to verify text candidates. Various features including structure, intensity and shape features, wavelet and HOG texture descriptors where used to perform discrimination of text in verification.

3.1.3 Text Segmentation

Before detected text regions are recognized by an OCR module, set an approaches use binarization, text line segmentation and character segmentation algorithms to obtain the precisely bounded characters.

3.1.4 Text Binarization

Text Binarization operates to extract text pixels and remove the background pixels. Algorithms related to adaptive thresholding, probability models, and clustering have been used in this problem.

3.1.5 Text Line Segmentation

The function of text line segmentation is to convert a region of multiple text lines into multiple sub-regions of a single text lines. A recent advance of text line segmentation comes with the emergencce of skeleton analysis method. Text skeletons are extracted from connected components and a text line is defined as a continuous path on the skeleton from an intersection point to either an end point or an intersection point. The path corresponding to text lines does not include any other points in the middle.

3.1.6 Character Segmentation

Character Segmentation seperates a text region into multiple regions of single characters. It is often difficult to determine an optical projection threshold when degradation or touching characters exist. Adaptive methods including the adaptive morphological operation, clustering and optimization methods have been steadily developed. At two-pass path search algorithm is applied wher the forward search localizes potential cuts and the backward direction removes the false cuts

3.1.7 Text Recognition

Text recognition converts image regions into strings in recent research, word recognition has been central to text recognition because words are well-formulated with statistical models in terms of low-level features and high-level languages priors.

3.1.8 Character Recognition

To recognize characters of a single font, general features and simple classifiers such as linear discriminant analysis (LDA) are often used. Part based implict models have explored for distorted character recognition proposed using deformal part based models and sliding window classification to localize and recognize characters in seen images. Characters are divided into parts, each of which moves in a local domain with panality parameters.

3.1.9 Word Recognition

It is not unusual for a recognition model to assign different labels to identical characters. This is particularly common given distortions are lack of training data for particular fonts. In this case the character segmentation and character recognition can be integrated using optimization methods including Bayesian inference, integer programming and graph models.

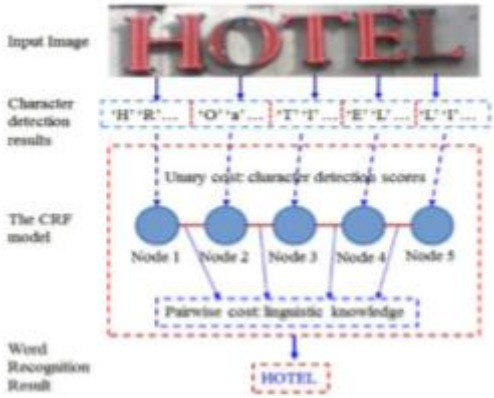

Figure 6. Word Recognition

3.1.10 End-to-End Recognition

Given imagery with complex backgrounds as an input, an end-to-end recognition system embodies the localization, detection and recognition functions to convert all text regions in the imagery into strings. The motivation of words spotting is that " The whole is greater than the sum of parts ", and the tasks looks to match specific words in a give lexicon with image patches using character and word models.

3.2 Photo-Ocr:Reading Text In Uncontrolled Conditions

We describe photo-OCR, a system for text extraction from images. With the goal of text recognition as a user input modality similar to speech recognition. Commercially available OCR performs poorly on this task. We evaluate our system on public bench mark data sets for text extraction and outperform all previously reported results, more than halving the error rate on multiple bench marks. The system is currently in use in many applications at Google, and is available as a user input modality in Google Translate for Android.

3.3 Automatic Text Location In Images

Automatic text location deals with extracting image regions that contain text. The images of these regions can be fed to an optical character recognition module or highlighted for users. This is very useful in a number of applications such as database indexing and converting paper documents to their electronic versions.

3.4 An Adaptive Text Detection Approach In Images

In this paper, an adaptive edge based text detection approach in images is proposed approach can adapt different edge detection methods according to the image background complexity. It mainly consists of four stages-Firstly, images are classifies into different background complexities. Secondly, different edge detectors are applied on the images according to there background complexities. Thirdly, Connected component analysis is adopted on the edge image to obtain text candidates to find the exact position.

4. CONCLUSION

In this paper image segmentation using boundary analysis have been elaborated. Also some unique properties of boundaries of object/text to be detected have been included which helps to extract it from the background in the image. Also correlational algorithms such as ICF and ACF for comparision of two boundaries were used for the feature extraction. This will in turn help to recognize a text as a collection of individual alphabets and digits from the image. With the hope that this paper will simplify the understanding of boundary analysis for image segmentation implementations have been done. In the future the following extensions can be made such as adding augmented reality features to it such that it can be used as real time translator i.e., Translation of image just as you capture it. To do this either a free or licensed API translator could be used.

5. REFERENCES

[1] " Scene Text Detection via Connected Component Clustering and Nontext Filtering". Hyung Il Koo, Member, IEEE, and Duck Hoon Kim, Member, IEEE.

[2] "PhotoOCR: Reading Text in Uncontrolled Conditions". Alessandro Bissacco∗, Mark Cummins∗, Yuval Netzer∗,Hartmut Neven Google Inc.

[3] "Automatic Text Location in Images and Video Frames". Anil K. Jain and Bin Yu Dept. of Computer Science, Michigan State University East Lansing, MI 48824, USA.

[4] Text Localization in Real -world Images using Efficiently Pruned Exhaustive Search ". Lukaˊs Neumann ˇ Centre for Machine Perception , Dept. of Cybernetics Czech Technical University , Prague, Czech Republic neumalu 1@cmp.felk.cvut.cz Jiˇrˊı Matas Centre for Machine Perception , Dept. of Cybernetics Czech Technical University, Prague, Czech Republic matas@cmp.felk.cvut.cz.

[5] N. Ezaki, K. Kiyota, B. T. Minh, M. Bulacu, and L. Schomaker, “Improved text-detection methods for a camera-based text reading system for blind persons,” in Proc. IEEE Int. Conf. Document Anal. Recognit., 2005, pp. 257–261.

[6] U. Pal, and P. P. Roy,"Multi-oriented and curved text lines extraction from Indian documents", Proc. IEEE Trans. on Systems, Man and Cybernetics-Part B, vol. 34, no. 4, pp. 1676-1684, 2004.

[7] P. P. Roy, J. Lladós, and U. Pal,"Multi-Oriented Character recognition from Graphical documents", Proc. Int. Conf. on Cognition and Recognition, pp. 30-35, 2008.