International Journal of Engineering Research and General Science Volume 3, Issue 4, July-August, 2015
ISSN 2091-2730

Literature Survey On Various Region Extraction Methods

Priya nair, Nithya Susan Abraham

Department of Computer Science and Engineering

Sree Buddha College of Engineering for Women

priyapnair7@gmail.com, nithyasusan55@gmail.com

Abstract— Detecting text in natural images is an important step for a number of Computer Vision applications. Text in various images provides vital information about its content and help in various navigational purpose also. Variations of text due to size, colour, orientation etc. in different backgrounds make the process of text detection complex or challenging. A pre-processing or first step in text detection is candidate region extraction. Various methods are there for region extraction such as morphology based, wavelet transform, MSER, Segmentation, eMSER etc. Each has its own advantages and disadvantages. Among these eMSER has better performance.

Keywords— Morphology; Wavelet Transform; Segmentation; MSER; eMSER; Fuzzy Match-Based Region Extraction; Border Tracing Algorithm; Edge Enhanced Region Extraction

INTRODUCTION

Text detection/localization framework are of two types: Image-analysis-based approaches and Image-analysis and machine learning-based approaches. Image-analysis-based detection depends heavily on image analysis techniques, such as edge detection, connected component, and texture analysis. According to the employed features Text detection/localization methods are classified into two main groups: region-based and texture-based. The region based method is further categorized as the Connected Component (CC) method and the Edge based method. Region-based methods use properties of the color or gray-scale in a text region or their differences with the corresponding properties of the background. They are simple to implement, but they are not very robust for text localization in images with complex backgrounds and usually their performance depends on several threshold values. Texture based methods consider texts as regions with distinct textural properties. They are more accurate when the text is embedded in a difficult background, whereas they suffer in detecting different sizes of text, and are computationally expensive. First, edge detection or CC or texture analysis is performed on the image under consideration. The image's raw features are then grouped according to connectivity, spatial, and geometric correlations to form potential text regions. Second, the potential text regions are further examined using rule-based heuristics, such as the size, aspect ratio, and orientation of the region.

Scene text detection methods are mainly of three types: texture based, region based, and hybrid based. In case of texture based approach various texture properties are extracted using a sliding window and then classifier is used to separate character from non-character. This has one disadvantage that complexity high and extra post processing step needed for text extraction. Region based approaches uses some region based algorithms for region extraction and thereafter some low level cues are used for separating text from non-text. One advantage of this is computational complexity low and no extra post processing step needed for recognition. Hybrid approach is a mixture of both region and texture based approach.

Candidate Region Extraction is the first or preprocessing step of text detection . In this paper various region extraction methods such as morphological method, wavelet transform, segmentation, MSER, eMSER are discussed.

MORPHOLOGY BASED REGION EXTRACTION

Mathematical Morphology is an important method f or analysing and extracting image that are useful in the extraction of various geometric structures , shape representations, boundary detection of objects , their skeletons etc. [1]. These are the basic operations that performed in morphology for region extraction.

Dilation:

Dilation adds pixels to the boundaries of objects in an image. The process may be repeated to create larger effects. The number of elements to be added to the image under processing can be determined by the size and shape of the structuring element.

Erosion:

Erosion removes pixels from the boundaries of objects in an image. The number of elements to be removed from the image under processing can be determined by the size and shape of the structuring element.

Opening:

Opening is an important morphological operator. It is defined as erosion, followed by dilation.

Closing:

Closing is defined as dilation, followed by erosion. In this background regions are preserved that have a similar shape to this Signal &Image.

Structuring Elements

Structuring element are used to carry out erosion and dilation. Every pixel along with its neighbours in the input image is compared with origin of structuring element and translated to each pixel position in the corresponding output image [11].First image is resized , converted into grayscale and filtered then structuring element is used to carry out erosion and dilation operation then its difference is found then it is converted into particular intensity range and convolution is performed to obtain extracted region. To remove the non-text from the resultant image, the size and shape of that region is compared with the text region. Connected components of the non-text whose shapes are obviously different from the text-like connected components and their areas of the text connected component are relatively small comparing to the areas of their non-text. Therefore, they can be easily filled by the above opening, closing and erosion operations. Finally, a fill-hole process is performed to deal with the non-text regions and remove the text regions from a document image.

WAVELET TRANSFORM

In this candidate text regions are obtained using Dense intensity variety and Contrast between text [3] and its background by assuming that all text regions have dense intensity variety and contrast with its background. The wavelet coefficients around thepixels should have large values when considering intensity variety around the text pixels. Wavelet energy feature is defined by integrating the wavelet coefficients in the three high frequency sub bands. A pixel will be a candidate text pixel in level n if its wavelet energy feature is larger than a dynamic threshold. Pixels whose histogram lies inshadow area be detected as candidate pixels. For the image of low contrast, base threshold is selectedfor Threshold, which ensures that most of the background pixels be excluded .With the increment of image contrast and complexity, threshold is selected and adaptively calculated. And the larger the image contrast is, the bigger threshold should be. Pixels whose contrast is higher than main threshold are selected as candidates. A text region is made of a ‘cluster’ of text pixels. None but ‘dense’ text pixels can construct a text region and the isolated candidate pixels are often noises.

A region growing method isused , in this unlabelled pixels are searched and consider as seed pixel. Then all these types of pixels that are density-connected with Pair labelled with the same region label. Label each found region as a text region. Merge the pixels that are not included in any text region with the background. Thus the region is Extracted.

SEGMENTATION

Candidate Region can also be extracted using Segmentation method [2]. In this image is divided into various segments and is typically used to locate objects and boundaries in images. Each pixel is labeled based on their color, intensity or texture and pixels with same properties are grouped. Output of the image is set of segments or candidate regions.

Segmentation has two objectives. One is to decompose image into various parts and another is to perform a change of representation. Meaningful or efficient pixels in the image are set as higher-level units for further analysis. Image sementation regions must be uniform and homogeneous with respect to various properties such as color, texture etc. Segments must be simple and without many small holes. Different region must have significantly different values with

respect to the characteristic on which they are uniform. Boundaries of each segment should be smooth, not ragged, and should be spatially accurate.

FUZZY- MATCH BASED REGION EXTRACTION

In this [6] the correspondence or correlation between each pair of homologous image points on a given pair of images. Matching is performed by comparing a query object or sub-image region with the test image to find the location of the matched pattern in the scene image. Thus by comparing the gray level properties or other properties such as texture etc, query object region with same properties of the image can be obtained. Query object may or may not be a part of image. It can be applied on both stationary or dynamic image. If an exact part of region is present in the scene image region extraction using segmentation yield better result. Main aim is to obtain spatial boundaries of region with user interest.

Using fuzzy based back propagation algorithm also we can extract regions. It can be applied on both monochrome and color images. It uses fuzzy measures for extracting regions of processed image and use histogram back-propagation algorithm. It help in extracting region in unconstrained image.

BORDER- TRACING ALGORITHM

It is used to extract contours of an object or region [7] **.**Assume that region is binary or previously labeled [9]. Search image from its top-left until a pixel is found, then define a variable 'dir' which shows direction of previous move along the border from previous border element to current border element. If it is 4 connectivity set variable as 0 and for 8 connectivity set variable as 7. Search 3*3 neighborhood of the current pixel in an anti- clockwise direction beginning the neighborhood search at the pixel positioned in the direction. (Variable + 3) mod 4, (Variable + 7) mod 8 if even, (Variable +6) mod 8 if odd. First pixel found with same value as the current pixel is a new boundary element. Update variable. I f current boundary element equals second boundary element and previous boundary element equals stop. Search again for 3*3 neighborhood. Thus border is detected. This algorithm doesn't find region hole boundary. For storage efficient representation chain code [8] is used.

MSER

MSER is Maximally Stable Extremal Region [4]. The input of the MSER algorithm is a grayscale image. It is binarized with a threshold t iterating from 0 to 255. 0 stands for completely black and 1 stands for completely white. White regions are selected as Extremal regions. This can be applied conversely also. It can be found that for how many successive images in the sequence this extremal region stays the same. Select images which are exactly the same in atleast R successive images by selecting a threshold value R. Such regions are called Maximally Stable Extremal Regions.

An advantage of the MSER algorithm is that it is well applicable for finding text character candidates and is affine invariant and can be applied to low quality images and has an efficient implementation. It can be implemented to detect both black and white font text. A set of distinguished regions that are detected in a gray scale image are denoted by MSERs [10]. All of these regions are defined by an extremal property of the intensity function in the region and on its outer boundary. MSERs are stable local detector since it has properties to show superior performance. MSERs are detected at different scales. MSER tracking requires a data structure that can be efficiently built and managed. The component tree is a structure which allows the detection of MSERs within an image and, in addition, constitutes the basis for MSER tracking. The component tree has been recently for efficient implementation of watershed segmentation. The component tree is a rooted, connected tree and can be built for any image with pixel values coming from a totally ordered set. Each node of the tree represents a connected region within the input image. The nodes of the component tree are identified as connected regions within binary threshold images.

EDGE-ENHANCED MSER

MSER is a natural choice for text detection as the intensity contrast of text to its background is typically significant and a uniform intensity or color within every letter can be assumed [12]. MSER has been identified as one of the best region detectors due to its robustness against view point, scale, and lighting changes, it is sensitive to image blur. Thus, small letters cannot be detected or

distinguished in case of motion or defocus blur by applying plain MSER to images of limited resolution. To cope with blurred images the complimentary properties of Canny edges are combined with MSER. The MSER pixels outside the boundary formed by the Canny edges is achieved by pruning the MSER along the gradient directionscomputed from the original gray-scaleimage. The gradient directions can be adapted toguarantee that they point towards the background. Thus individual letters can be identified. This not only improves the performance of geometric filtering but also increases the repeatability of MSER based feature matching under different image blur conditions.

eMSER

eMSER stands for Edge-preserving MSER. In this [5] input is the color image and its required parameters and output are various potential characters.First color image is converted into intensity image and then smooth it using guided filter. Then compute the gradient amplitude map and normalize it to [0,255]. Thus we get a new image, perform MSER algorithm on the new image to extract dark regions on the bright background and vice-versa.MSER algorithm is sensitive to blurred image so gradient amplitude is added to it to produce edge-preserving MSER

ACKNOWLEDGMENT

We acknowledge almighty god.

CONCLUSION

Various region extraction methods are there. Among the above surveyed region extraction methods eMSER has better performance. To avoid sensitiveness of blurred image, gradient amplitude is added in eMSER. Morphological operation such as 'close' operation has a disadvantage that it is often used to connect text pixels into text regions. In theoperation, all of the pixels near to each other will beconnected despite whether they form a 'cluster' of textpixels or not. This problem is not in case of wavelet transform and other above discussed methods.

REFERENCES:

[1] R.Radha and C.P.Sumathi, "A Novel Approach To Extract Text From Licence Plate Of Vehicles": An International Journal (SIPIJ) Vol.3, No.4, August 2012 on Signal & Image Processing

[2] Shapiro and Stockman, "Image Segmentation" Computer Vision: Mar 2000

[3] Qixiang Ye, Qingming Huang, Wen Gao, Debin Zhao, " Fast and robust text detection in images and video frames ", Image and Vision Computing (2005)

[4] Mikhail Zarechensky, "Text detection in natural scenes with multilingual text", Department of Analytical Information Systems, Saint Petersburg State University

[5] Yao Li, WenjingJia, ChunhuaShen, and Anton van den Hengel, " Characterness: An Indicator of Text in the Wild ", IEEE Transactions On Image Processing, Vol. 23, No. 4, April 2014

[6] TamalikaChaira, Ajoy Kumar Ray, "Fuzzy Match-Based Region Extraction" , http://books.google.co.in/books? isbn=1439807094

[7] BorderTracing-Digital Image Processing Lectures,The University of lowa, http://www.icaen.uiowa.edu/~dip/Lecture/Segmentation2.html#tracing

[8] Chain codes- Digital Image Processing Lectures,The University of lowa, http://www.icaen.uiowa.edu/~dip/Lecture/Shape2.html#tracing

[9] G.X. Rtter, J. N. Wilson, "Handbook of Computer vision algorithms In Image Algebra" Second Edition-Chapter 10.4 Chain code Extraction and Correlation, CRC press, New York

[10] M. Donoser and H. Bischof, "Efficient maximally stable extremal region (MSER) tracking," in Proc. IEEE Conf. Comput. Vis. Pattern Recognition, Jun. 2006

[11] Yung - Sheng Chen, "Morphology Based Region Extraction " www.researchgate.net

Huizhong Chen, Sam S. Tsai1, Georg Schroth, David M. Chen, Radek Grzeszczukand Bernd Girod, " Robust Text Detection In Natural Image With Edge-Enhanced Maximally Stable Extremal Regions " in Proc Image Processing