International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 06 | June 2019 www.irjet.net p-ISSN: 2395-0072

Digit Identification in Natural Images

Govardhan Gayake[1], Dr. Arati Dandavate[2]

[1]ME Student, Savitribai Phule University of Pune, Department of Computer Engineering JSPM COE, Hadapsar, Pune
[2]Professor, Savitribai Phule University of Pune, Department of Computer Engineering JSPM COE, Hadapsar, Pune

---

Abstract — Number recognition from images is a complex task to achieve in this digital era. Advanced applications that are beyond Automatic Number Plate Recognition (ANPR) are developed with different algorithms and image processing techniques to recognize numbers from an image with much adverse scenarios. In this paper, we are describing a technique to locate and recognize the numbers in an image. It includes preprocessing for image smoothening, Edge detection to recognize boundaries of all objects. Maximal stable extremal region is used to identify the potential candidate objects. Feature extraction is used to predict the possible number for that object.

Keywords — MSER, Sobel, Feature Extraction

I. INTRODUCTION

Today world has advanced to digital age. People are making out more work with less effort through the help of digitization. The work which is used to take hours is now done within seconds. People prefer digital documents over hard copies as it's handy and portable. The digital documents are digital books, images etc.

In many scenarios, it's required to extract information from these digital documents. This information may include numbers and text.

In this paper, we are discussing process to locate and recognize numbers in the image. Initially image is preprocessed to remove noise. The preprocessing involves gray scaling of RGB image. The preprocessed image undergoes edge detection process. The edge detected image is used to find MSER (Maximally stable extremal regions) regions.

The MSER regions are the regions which has very stable intensity. The MSER regions are the region where numbers lies in the images. Finally feature extraction is used to predict the number.

This paper is divided in 5 sections as: Second section describes pre-processing techniques. The third section explains the working of edge detection technique. The fourth section describes the MSER regions. The fifth section explains feature extraction to predict number and finally conclusion is given.

II. PRE-PROCESSING

The aim of preprocessing is an improvement of image data that suppresses unwanted distortions or enhances some image features important for the image processing. Also, it enhances the visual appearance of images.

A. Gaussian Blur

Gausssian blur is used as pre-processing technique to smoothen the image by noise removal. It is performed by convolving the image using the Gaussian function. The Gaussian distribution in 1-D has the form:

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

Fig 1. Gaussian distribution

Where sigma is the standard deviation of the distribution. Gaussian blurring is commonly used when reducing the size of an image. When down sampling an image, it is common to apply a low-pass filter to the image prior to resampling. This is to ensure that spurious high-frequency information does not appear in the down sampled image.

III. EDGE DETECTION

Edge detection plays important role in image processing. Significant transitions in an image are called as edges. It differentiates between different objects, the outline in the foreground and background object. Different features such as shapes, boundary can be measured between the designs of image

[1]. Edge detection technique works by finding discontinuity in the images [2].

The different edge detection algorithms are as follows:

A. Sobel Operator

Working of Sobel Operator is as follows:

It function like first order derivative, in edge region the difference of pixel intensities is calculated. An image intensity functions approximation of the gradient is computed, the Gaussian smoothing and differentiation is combined. It is a pair of 3x3 convolution kernel that takes first kernel and rotates another kernel by 90 in anticlockwise direction.

-1	0	+1
-2	0	+2
-1	0	+1

Gx

+1	+2	+1
0	0	0
-1	-2	-1

Gy

Fig 2. Kernel matrix

Fig 3. Edge detected image using Sobel operator

The above designed kernels respond to horizontal and vertical running edges. To produce the gradient component in each orientation Gx and Gy, combine these to produce magnitude of Gradient [2] [3].

$$|G| = \sqrt{Gx2 + Gy2} \quad (1)$$

Appoximate magnitude is calculated using equation 2.

$$|G| = |Gx^2| + |Gy^2| \quad (2)$$

The angle of orientation of the edge is calculated by equation 3.

$$\Theta = \arctan(Gy/Gx) \quad (3)$$

B. Canny Edge

The Canny operator works in a multi-stage process. Canny edge algorithm makes use of calculus of variations. The tracking process exhibits hysteresis controlled by two thresholds: *T1* and *T2*, with *T1 > T2*.

Canny Edge algorithm is as follows:

Step 1: Initially, the Gaussian filter is applied to smooth image and noise is removed. **Step 2:** Intensity of images is calculated.

Step 3: To get rid from non-genuine response to edge detection non-maximum suppression is applied. **Step 4:** Potential edges are determined by applying double threshold.

Step 5: Other weak edges are suppressed which are not connected to strong edge [4].

Fig 3 Edge detection using Canny Edge algorithm

C. Prewitts Operator

In Prewitts operator, mask is applied on the image which calculates difference of pixel intensities of the edge region. Horizontal and vertical edges are detected in Prewitts operator. Differences of pixel intensities of corresponding pixel are used to calculate edges.

The mask for detection of edges in vertical direction,

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

The mask for detection of edges in horizontal direction,

$$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Above mask will find edges in horizontal direction and it is because of zeros in middle row [7] [9].

IV. MSER

Maximally stable extremal regions (MSER) are used to obtain all possible candidate regions. Candidate regions are the text regions which we are interested in. In MSER we firstly perform the sorting operation on the pixels. After sorting, pixels are marked in the image, and the list of growing and merging connected components and their areas is maintained using the union-find algorithm [8].

There are two types of MSER i.e positive and negative. First type we get by thresholding all intensities of grayscale images.

E.g. for a threshold T = 45, all pixels with intensity <= 45 are assigned black, or foreground, and all pixels > 45 intensity are white or background. While the second type consists of only two intensities 0 and 255.

Compared with other five region detectors which are Harris-affine region detector, Hessian-affine detector, edgebased regions, intensity extremal detector, and salient regions detector MSER provides many advantages.

- Density of Region

MSER offers 2600 regions for texture blur image and 230 regions for light changed and has repeatability of 92% for this tests.

- Size of Region

MSER goes for smaller region detection as compared larger regions.

- Change of Scale

After Hessian-affine detector, MSER comes in second under scale change and in-plane rotation of image.

- Change of light

MSER showed the highest repeatability score for this type of scenario compared to other region detectors.

Fig 4.1 Edge directed image

Fig 4.2 Candidate Region

Above images show MSER region obtained from edge detected image.

V. NUMBER RECOGNITION

Once we get the area where digits are located next step is to recognize the digits. The digits can be recognized using feature extraction technique. The feature extraction is process which involves finding some specific pattern or features which can be easily compared with trained data to recognize the objects in an image. The features can be edges, corners or blob in the image. The shaped based features are nothing but lines, circles. [11][12] [13].

A. Segmentation of Digits

- *Vertical projection*

The vertical projection is used to segment each digit. The vertical projection is the process of calculating sum of pixels in each column of an image. In vertical projection, troughs represent the boundary between two digits [15].

- *Using Contours*

The contours are the boundaries of any object in the image. To find contours in the image, initially image is converted into black and white image. The black and white image will have pixels with intensity 0 and 1 where 0 or white represent background part of image and 1 or black represent pattern [16].

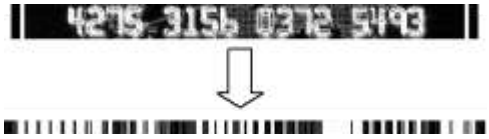


Fig 5. Projection image

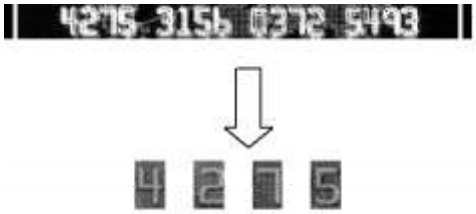


Fig 6. Segmentation of Digits

B. Feature Extraction

After segmentation, bounding box is formed around each segmented digit. For each bonding box features are extracted. Now to recognize each digit, extracted features need to be classified. To classify each digits technique such as KNN neighbor or probabilistic Neural Network can be used. In KNN algorithms, initially samples of digits are used for training classifier. This classifier is used to classify each digit based upon features.

The complete step wise procedure is described below

Fig 7.1. Preprocessing on input image

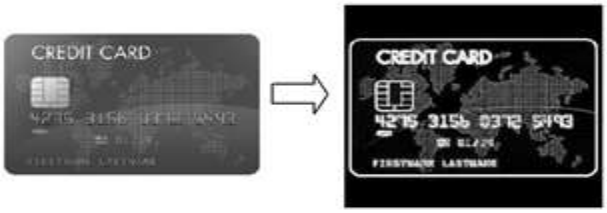


Fig 7.2. Edge detected image using Sobel operator

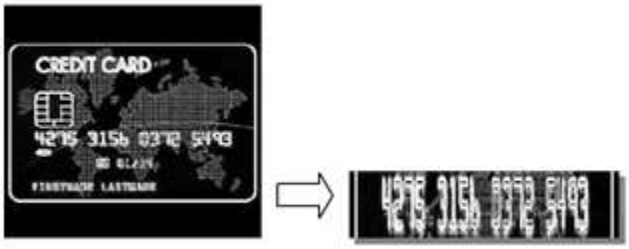
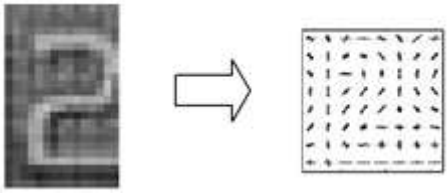


Fig 7.3. Detect MSER Region

Fig 7.4. Segmentation of Digits



7.5. Feature extraction for each digit

Fig 7.6. Classification of each Digit

VI.CONCLUSION

The canny edge detection algorithm requires threshold for detecting edges where as sobel can find all edges using horizontal and vertical gradients. The MSER region provides smaller precise regions than any other region detector. The optical character recognition system uses training data set having hundred of samples which consumes lot of memory space. To recognize numbers using feature extraction, it needs either no training data or smaller training data resulting in saving memory space and computation time.

VII. FUTURE SCOPE

Above techniques performs efficiently when input images has plain background .The accuracy degrades as complexity of background of image increases. The complex background introduces unwanted edges after edge detection due to which unexpected features of images get extracted. So algorithm to remove unwanted edges due to complex background can be included in system to make robust in all condition.

REFERENCES

1. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.
2. B.Pechiammal, Dr.J.Arokia Renjith, “AN EFFICIENT APPROACH FOR AUTOMATIC LICENSE PLATE RECOGNITION SYSTEM”, 2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM), 2017 IEEE
3. Jack Greenhalgh and Majid Mirmehdi, Recognizing Text-Based Traffic Signs, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTA-TION SYSTEMS, VOL. 16, NO. 3, JUNE 2015.
4. Huizhong Chen, Sam S. Tsai, Georg Schroth, David M. Chen, Radek Grzeszczuk and Bernd Girod, Robust Text Detection in Natural Images with egde-enhanced Maximally Stable Extremal Regions, ICIP 2011.

5. Govardhan Gayake, "A Survey on Text Detection from Natural and Scanned Images", International Journal of IT, Engineering and Applied Sciences Research (IJIEASR) ISSN: 2319-4413 Volume 8, No. 1, January 2019
6. Shubham Saini, Bhavesh Kasliwal, Shraey Bhatia, "Comparative Study of Image Edge Detection Algorithms".
7. http://in.mathworks.com/discovery/edgedetection.html?s_tid=gn_loc_dro p
8. https://en.wikipedia.org/wiki/Sobel_operator
9. http://homepages.inf.ed.ac.uk/rbf/HIPR2/sobel.htm
10. https://en.wikipedia.org/wiki/Canny_edge_detector
11. http://www.tutorialspoint.com/dip/prewitt_operator.htm
12. [8] K. Mikolajczyk, T.Tuytelaars, C. Schmid, A. Zisserman, T. Kadir and L. Van Gool, "A Comparison of Affine Region Detectors"; International Journal of Computer Vision, Volume 65, Numbers 1-2 / November, 2005, pp 43-72
13. Jayshree Ghorpade-Aher, Sumeet Gajbhar, Amey Sarode, Govardhan Gayake, Piyush Daund, "Text Retrieval from Natural and Scanned Images" International Journal of Computer Applications (0975 – 8887) Volume 133 – No.8, January 2016
14. Jayshree Ghorpade, Raviraj Palvankar, Ajinkya Patankar and Snehal Rathi, "EXTRACTING TEXT FROM VIDEO, Signal & Image Processing" An International Journal (SIPIJ) Vol.2, No.2, June 2011
15. Sushruth Shastry, Gunasheela G, Thejus Dutt, Vinay D S and Sudhir Rao Rupanagudi, 'i - A novel algorithm for Optical Character Recognition (OCR)', 978 -1-4673-5090-7/13 IEEE, 2013
16. http://docs.opencv.org/3.0beta/doc/py_tutorials/py_feature2d/py_feature s_meaning/py_features_meaning.html#features-meaning [13] Ivind Due Trier, Anil k Jain and Torfinn Taxt ,"Feature Extraction Methods For Character Recognition - A Survey"
17. M. Zahid Hossain, M. Ashraful Amin, Hong Yan, "Rapid Feature Extraction for Optical Character Recognition" 2012
18. http://4imageprocessing.blogspot.in/2009/05/opencv8-find-verticalprojection-of.html
19. http://www.imageprocessingplace.com/downloads_V3/root_downloads/t utorials/contour_tracing_Abeer_George_Ghuneim/intro.html
20. Jack Greenhalgh, Majid Mirmehdi, "Recognizing Text-Based Traffic Signs" IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTA-TION SYSTEMS, VOL. 16, NO. 3, JUNE 2015
21. Huizhong Chen, Sam S. Tsai, Georg Schroth, David M. Chen, Radek Grzeszczuk and Bernd Girod, "Robust Text Detection in Natural Images with egde-enhanced Maximally Stable Extremal Regions", ICIP 2011
22. Binh Quang Long Mai, Tue Huu Huynh, Anh Dong Doan, "A Study about the Reconstruction of Remote, Low Resolution Mobile Captured Text Images for OCR", 978-1-4799-6956-2/14 IEEE, 2014.
23. Faisal Mohammad, Jyoti Anarase, Milan Shingote, Pratik Ghanwat, "Optical Character Recognition Implementation Using Pattern Matching, International Journal of Computer Science and Information Technologies", Vol. 5 (2) , 2014.