

An approach to Extract Text from Water Meter Images using OpenCV-Python

Shantala¹, Latha H K E², S.V Mohanasundaram³, Gokul Krishnan G³ and Niyathi Shenoy K³

^{1,2}Dept. of E&IE, SIT, Tumakuru, India

gshantala04@gmail.com, lathahke@gmail.com

³Dept. of IOT,C-DAC, Bengaluru, India

mohansv@cdac.in, gokulkg@cdac.in, niyathishenoy@cdac.in

Abstract—The text extraction has been developing rapidly and is an important research field in recognition from meter images. This paper brings out an algorithm to extract the text from a water meter image. The meter image has same scales of the width and height of the target area which makes this algorithm suitable to recognize the numerical characters. The algorithm finds the region of interest (ROI) of water meter using Matplotlib, then all numeric are segmented. The segmented digit, each row value is added and stored as an array values from 0-9. To extract text from a water meter image, first segments the digits and each row value is added and compared with previous stored array values. The comparison is done by finding the Euclidean distance of two arrays and then computed the text information. The algorithm is able to extract text from the specified meter images using OpenCV- Python.

Index Terms—Water meter, OpenCV- Python, Matplotlib, Euclidean Distance.

I. INTRODUCTION

In recent years, the automatic reading of meter system becomes an important symbol of intelligent community gradually. The establishment of remote automatic reading meter system saves the personnel expenditure and it also decrease the need of labors [1]. The traditional way of water meter reading is inconvenient and time consuming where requires huge number of labor operators. Water is an essential resource to fulfill everyone's need in day to day life. According to a survey India and china got the highest stress level of water. In India 70% of water is being utilized for the agricultural purpose and around 19% of water by the industries [2].

There are several text detection and recognition algorithms have been proposed in the past.Camera based images is a very difficult problem because it is not always possible to precisely define the features of text in a colored scene image because of the wide variations in possible formats [3].

A text information in images contains important information in different applications. There are several Automatic detection of text from images has been proposed. These methods aim to detect the characters or numbers based on general properties of character like pixel values [1]. The text extraction technique can broadly classified into four types i.e. Region based, Edge based, Morphological Based and Texture based techniques. A text extraction algorithm for camera based images. The methodology is based on edge

detection technique. The text is extracted using discrete edges around the text boundaries [4].

Text Extraction from Heterogeneous Images using Mathematical Morphology and Thresholding technique is implemented [5]. Reading Values in Electrical Meter Using Image Processing Techniques, where adaptive Thresholding and vertical edge detection algorithm is used [6]. One of the most important reasons is that the algorithm in the automatic reading meter system, which is character recognition algorithms [7]. It has a low recognition score and it affects the normal use of systems.

There are so many algorithms are proposed to extract text from gas, electrical and water meters. It's always a better practice to use a specific algorithm for a specific kind of meter system which increases the flexibility and easy usage of the system. This paper will give a character recognition algorithm suit to a specific kind of water meter.

II. MOTIVATION

According to a research India is in the need of smart water metering system because the water stress level is more here. In India sufficient water resources are available but water is unevenly distributed. It requires an efficient and Improvement in water meter text extraction technique for smart metering systems [2]. Even though there are so many algorithms available for extracting text from meter images which is more complex and time consuming. Here an algorithm is proposed to extract the text information from water meter images using OpenCV-Python.

III. DESCRIPTION OF PROPOSED ALGORITHM

When obtaining the text from water meter for smart metering system, the camera and the meter both are fixed with some distance. The algorithm involves several steps like image digit segmentation, binarization and uses OpenCV Image Processing library which created by Intel and maintained by Willow Garage.

First step is taking the image of meter and crop/resize that, repeat the same procedure to get all 0-9 values for preparing the array. The Red Green Blue (RGB) image is converted to Black White (BW) image, in order to reduce the processing overload. The image then segmented with fixed number of rows and columns. By writing a Python code program to count the number of each row elements and put all the number values in an array. Calculate the Euclidean distance of each element, which gives the minimum Euclidean distance value that is taken as extracted text of that image. In this proposed algorithm an improved text localization and extraction technique from images is obtained. The proposed method is tested with various captured image values.

A. Image Segmentation and Array Preparation of Digits.

The binarization techniques convert the image to black and white image using Thresholding. According to experiences here to get a good effect that the threshold value is considered in between 80 to 110 when binarizing. The image first converted to gray scale then threshold value is applied. In order to recognize each character from image after binarization processing needs to be segmented to get the interested region of the image and further to segments each character need to be recognized. The character segmentation is based on the several technologies such as gray projection, the character distance determination, determination the ratio of character between width and height, character outline analysis, and so on. The character segmentation is a difficulty in the character recognition of the images. Here after binarization Technique the image is segmented using Matplotlib function in python. Matplotlib is a plotting library environment which is used for analyzing images, plotting the overall accuracy of search systems. For finding the X and Y co-ordinates of the meter image the Matplotlib is imported in the program

The segmentation of digits depends on the X and Y values of the image, where X corresponds to width of the digits and Y corresponds to height of the digits. The values of X and Y for all digits should be same with fixed offset value. After digits separation prepare the array formation of numbers from 0-9 values. The array formation consists of finding the sum of each row values of the digits. The each row values of each digit is added and stored as an array. Each digit contains an array which has different values in it.

Using Python language the array formation can be done but using the instructions provided by the python will be time consuming. So, using built in library functions like Numpy, Scipy which does the faster operations on array numbers. The pre-computed values of these array are used to extract the text from the water meter image. This array formation is done before and stored for further calculations. The Fig.1. Shows the flowchart of the algorithm.

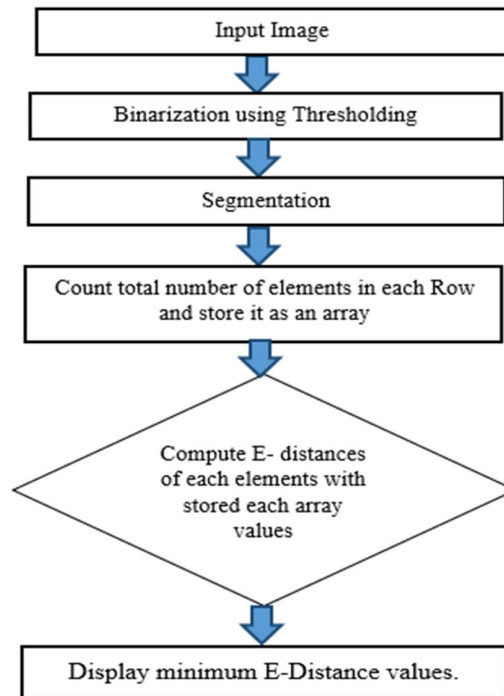


Figure. 1. Flow Chart of proposed algorithm

B. Character Recognition Method Using Pre-Computed Array Values.

It may recognize the characters of meter after determining their positions in the captured images. There are many algorithms can recognize the characters, such as template matching, morphological technique, feature pattern matching, and so on, these methods have the so many good and bad points also [8,9,10]. The previous computed values of arrays are taken, when an image of meter reading is to be extracted first the image is binarized using Thresholding technique.

The binary image is further segmented for all digits. The segmented digits taken and computed each row values are added and stored the 1st segmented digit array. This 1st array is compared with all stored array values from 0-9. The comparison is done by finding the Euclidean Distance of each array values with each other stored array matrix values. The comparison which gives the minimum distance value of the array will be taken as the text which is extracted value from the meter. Follow the steps given in Fig. 1. For the rest of all digits same procedure is followed. The meter consists of six digits so, six times the comparison is done with the stored 0-9 values. The extracted values are displayed on the terminal window and stored for further use of the digits

IV. RESULTS

The Results of the proposed algorithm is as shown in Fig. 2. The water meter image is taken and resized. The meter image is then used to extract the X and Y co-ordinates using Matplotlib function which is a plotting library environment shown in Fig. 3. It is used for analyzing images, plotting the overall accuracy of search systems.

After finding the co-ordinates, the image is converted to BW image using Thresholding technique as shown in Fig. 4.

After binarization the image digits are segmented for further computations, the segmented images shown in Fig. 5.

The extraction of the text from the images is done using OpenCV-Python. The segmented images are stored in an array by considering non-zero elements of each row. All the extracted present image array values and

previously stored array values are compared to find the minimum Euclidean Distances. The minimum value of each digit is taken as the extracted text from the image. The meter consists of 6 digits and 6 digits are displayed on the terminal window as shown in Fig. 6.

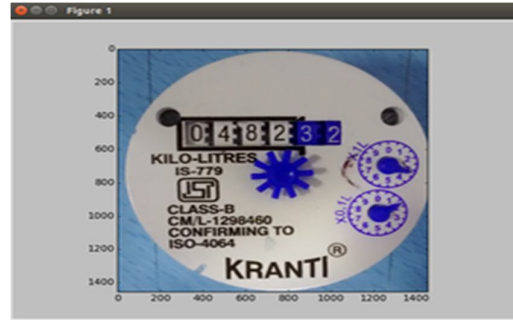


Figure. 2. Water meter image. Figure. 3. Water meter Image using MatPlotLib function for finding the co-ordinates.



Fig. 4. Black and White image of the meter after binarization

0 4 8 2 3 2

Fig.5 Segmented text from image

```

intern@intern-ThinkCentre-M81: ~/Desktop/paper
intern@intern-ThinkCentre-M81:~/Desktop/paper$ python prg1.py
0
4
8
2
3
3
2
intern@intern-ThinkCentre-M81:~/Desktop/paper$

```

Fig. 6. The extracted text result on terminal window

V. CONCLUSION

In this paper the proposed algorithm is used to extract text from specified water meter images using OpenCV-Python. It takes less than 1 minute time for extracting the text information. In future the algorithm can be used to extract the several number plate, electrical meter, gas meter and for other application. The performance can be improved by considering the features of the text and can be made to increase in the accuracy of the overall system.

ACKNOWLEDGMENT

I would like to thank Dr. Hari Babu, Dept. of IOT C-DAC Bangalore for his help in carrying out the experiments. The authors also thank the Executive Director Dr. Sarat Chandra Babu N C-DAC for their encouragement and support to carry out this work.

REFERENCES

- [1]. Ashna K, Sudhish N Gorgre, "GSM based automatic energy meter reading system", IEEE International conference held at Kottayam, pp. 65-72, 2013.
- [2]. Aditya Gupta, Sudhir Mishra, Neeraj Bokde and Kishore Kulat.: "Need of Smart Water Systems In India", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 4 (2016) pp 2216-2223 © Research India Publications.
- [3]. Manisha V Shinde, Pradip W Kulkarni, "New Development In Energy Meter Reading System", International Journal of Scientific Research and Management Studies (IJSRMS) ISSN: 2349-3771 Volume 1 Issue 4, pg. 124-129.
- [4]. S P Chowdhury, S Dhar, A K Das, B Chanda and K McMenemy, "Robust Extraction of Text from Camera Images", 2009 10th International Conference on Document Analysis and Recognition.
- [5]. G.Rama mohan babu, P.Srimaiyee, A.Srikrishna, "Text Extraction from Hetrogenous Images Using Mathematical Morphology", Journal of Theoretical and Applied Information Technology.
- [6]. Parthiban K, Palanisamy A.M, "Reading Values In Electrical Meter Using Image Processing Techniques", IEEE International Conference on Intelligent Interactive Systems and Assistive Technologies, August 2-3, 2013, Coimbatore, INDIA.
- [7]. H.G.Rodney Tan, C. H. Lee and V. H. Mork, "Automatic power meter reading system using GSM network", IEEE, 8th International Power Engineering Conference, pp. 465-469, 2007.
- [8]. H. I. Koo and D. H. Kim, "Scene text detection via connected component clustering and non-text filtering," IEEE Transactions on ImageProcessing, vol. 22, no. 6, pp. 2296-2305, 2013.
- [9]. P. Ramasubramanian, R. Jerlin Emiliya, R. Janaki, B. Gifston Daniel and C. Anand, "Number Plate Recognition and Character Segmentation using Eight Neighbors and Hybrid Binarization Techniques" International Conference on Communication and Signal processing, April 2014.
- [10]. G.Rama mohan babu, P.Srimaiyee, A. Srikrishna. "Text Extraction From Hetrogenous Images Using Mathematical Morphology", Journal of Theoretical and Applied Information Technology