Digitization of Mechanical Meter Bill Generation using OCR

Pankhuri Tayal
Computer Science Engineering
PES University
Bangalore, India

Pratheeksha P
Computer Science Engineering
PES University
Bangalore, India

Pranamya P Bhat
Computer Science Engineering
PES University
Bangalore, India

Abstract—**Several technologies have been innovated over the years in attempts to digitize every possible process. However, it is much easier to create new technology than to replace the existing ones. There are yet a large number of rural areas where digital meters have not been installed and traditional mechanical meters are still in use. The utility providers have human dependencies to take regular scans of the meters, in person, to generate the bill. To create a hassle-free and accurate billing system, we have come up with the idea of "Digitization of mechanical meter reading using OCR".**

Keeping the under-privileged areas of the world in mind, this solution has been developed. Unlike the advanced metering infrastructure (AMI), this method does not require any advanced complex infrastructure. A camera fixed in front of the mechanical meter takes the image of the meter every month and sends it over to the database. Using the image, the current reading is extracted (by OCR methods). Current reading is subtracted from the previous months' reading and stored back in the database. Based on the obtained monthly reading, tariff calculations are made. For the customer to view the generated bill, they will be required to log in into the web application.

Keywords— Mechanical meter digitization, electricity bill generation, Character recognition, OCR.

I. INTRODUCTION

Government organizations have been employing people since early times to collect meter reading from each house. This process is clearly inefficient, time consuming and inaccurate. The chances of human inaccuracy are high. Misreading of numbers is one of the common mistakes happening across. In some houses, the meters are not placed in easily accessible locations. With the increase in population, the number of houses is also increasing. Hence the organizations are forced to employ an excess number of middle men for this purpose. Therefore, this traditional method can be concluded as erroneous and expensive.

To overcome the mentioned disadvantages and discrepancies, technologies like OCR can be made use of.

"Digitization of analog meter reading "implements the modern application of the OCR technology.

Optical character recognition (OCR) is the new emerging technology in the modern-day innovation. It is well-known for the process of recognizing text from typed content, printed documents and texts. It has its various applications in the world of image processing [1].

Tesseract libraries help in achieving a precision of 0.929. OpenCV renders the required technology for image processing. The proposed approach for designing this system involves storing details in an admin database, monthly capturing the images, processing of the image and tariff calculations.

II. EASE OF USE

A. Service providers' perspective

The current methodology seeks for human intervention for meter reading collection. This method comes with its own disadvantages:

-Cost inefficiency: Service providers have to ensure that it employs a sufficient number of people for meter reading collection. The addition of extra workforce for this purpose increases the costing of the company.

-Inaccuracy: Erroneous reading by humans leads to misleading information and results in incorrect bill amounts[2].

As an addition to the advantages, new users are added into the database by the admin. Access to the database will remain restricted to admins only to avoid trespassing or data theft.

B. Customers' perspective

The existing system involves human discrepancies leading to faulty bill amounts. Process of accessing monthly is made very user friendly. The customer has to just enter his/her user-id to access the monthly bill.
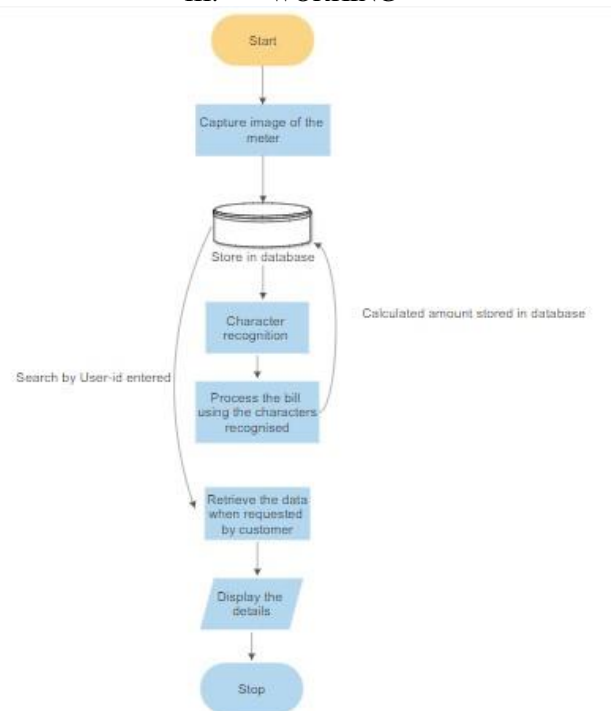
III. WORKING

Fig 2. Bill generation process using OCR.

A. Image capturing

The new ideology of digitizing analog meters makes use of a camera that will be fixed facing the meter boards. Using a counter module, the camera will capture the image of the meter board on the commencement of a new month. This image captured is sent to the database and stored in the respective tuple.

B. Image processing

a) Image cleaning: Tesseract is the best OCR engine with an average accuracy rate of 98%. Image processing is next to impossible without OpenCV. However, to get the right readings from a raw image, some preprocessing and cleaning is required.

First, the image is converted into grayscale and then into binary. The purpose of this is, grayscale simplifies calculations as the picture is only composed of shades of gray. Conversion to binary, further reduces complexity. Since any pixel in the image can be either black or white, the pixel can be represented with less information. It also highlights the characters to be recognised, which is the region of interest. The binary image is then resized to a standard dimension using inter cubic interpolation [3][4]. A sharpening filter that makes use of a 3-D numpy array is applied to the image to reduce blurriness. The filter enhances the contrast between each pixel and its neighbours thus sharpening the image. Then, erosion and dilation filters are applied. Here, 2 2-D numpy arrays of all ones, where each 1 represents a pixel under the array, and whose number of rows and columns can vary depending on the type of image, are used. Erosion thins down the boundaries whereas, the dilation thickens it. This step ends the preprocessing stage, and the processed image is passed to the pytesseract's image_to_string function that recognises the characters and returns it as a string.

b) Post-processing : The meter reading is composed of only digits. There could still be some errors such as misidentification while reading the characters. For example, 1 is read as l or I, 2 is read as Z, 8 is read as B etc. Basic for loop and if conditions can be used to iterate through the string and fix these kinds of errors and make sure that no alphabets were recognised. Then, the string having only numeric characters is then converted to float and used to calculate the electricity bill.
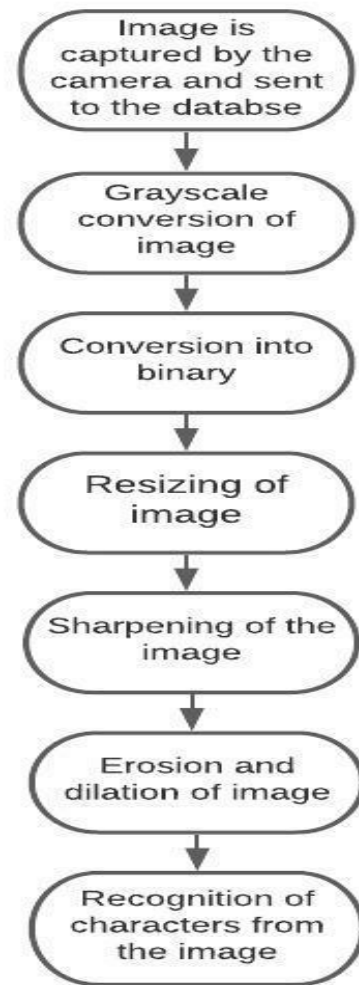

Fig 1. Flow of Image Processing

C. Database management

Access to the database is restricted to the admin. every new user is entered into the records by the admin. This ensures that the data is not mishandled.

User details and other required information are stored in the database using SQLite. The database has the following fields: User-id, image of the meter reading, previous month's reading, current month's reading, units consumed this month and the bill amount for the month.

After the current month's processing is completed, the database is updated for the next month's process. The current month's reading becomes the previous month's reading for next month and so on.

D. Tariff calculations

The resulting output from the character recognition provides us with the cumulative total of units consumed. The total units consumed is obtained after subtracting it from the previous months' reading.

Units consumed for the month is used to calculate the bill according to the slab designed by the service provider.

E. User interface

Flask framework will be used along with python backend to create a web application via which the user can login in and view their generated bill amount every month.

IV.   RESULTS

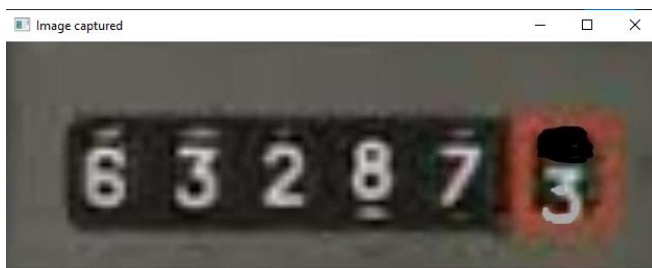Fig 3 (a) Meter board image captured by camera

Fig 3 (b) Image of the meter board after cropping of required part using tesseract

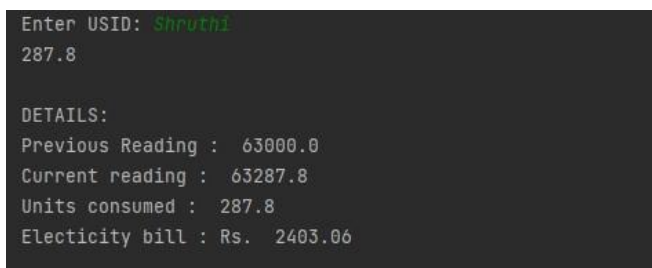Fig 3(c) After image processing

Fig 3 (c) Output obtained upon entering the user-id

From the first result [Fig 3] we observe that it works with perfection and accuracy for even digits detected after the decimal point. Fig 3(c) is the resultant image after cropping, grayscale conversion, binary conversion, erosion and dilation processing.

The database retrieves the data on the basis of the user-id entered. The data extracted helps in the bill calculation. Bill amount is generated using the tariff slabs.

V.   CONCLUSION

The above mentioned approach clearly replaces the conventional manual method for energy consumption calculations. Usage of OCR minimizes the manual data entry work, thereby resulting in better accuracy. By implementation of such an automated method, unscrupulous functions by customers can be avoided. The technique helps in cutting down revenue losses for the government/any service provider. In future, blurred images are sought to be taken care of using newer technologies.

ACKNOWLEDGMENT

REFERENCES

[1] Arpitha Shankar S I , Chaitra T V , Meghana S , Sadhana Prakash Sharma , Shilpa M, 2015, Billing of Electric Meter using GSM, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCRTS – 2015 (Volume 3 – Issue 27)

[2] Kanagarathinam, Karthick & Chitra, S.. (2017). Novel Method for Energy Consumption Billing Using Optical Character Recognition. Energy Engineering. 114. 64-76. 10.1080/01998595.2017.11863765.

[3] C. Patel, A. Patel and D. Patel, "Optical Character Recognition by Open Source OCR Tool Tesseract: A Case Study", International Journal of Computer Applications, vol. 55, no. 10, pp. 0975-8887, October 2012.

[4] M. Brisinello, R. Grbić, D. Stefanovič and R. Pečkai-Kovač, "Optical Character Recognition on images with colorful background," 2018 IEEE 8th International Conference on Consumer Electronics - Berlin (ICCE-Berlin), 2018, pp. 1-6, doi: 10.1109/ICCE-Berlin.2018.8576202.

[5] D. Pejic and V. Vujicic, "Accuracy Limit of High-precision Stochastic Watt-hour Meter," in IEEE Transactions on Instrumentation and Measurement, vol. 49, no. 3, 75 pp. 617-620, Jun 2000. doi: 10.1109/19.850404.

[6] Karthick, K., and S. Chitra. "A Bilingual Numeral OCR System for Creating UniLingual Digitized Numeral Document," Modern Applied Science, Vol 9, No 13 (2015), pp. 148-153. doi:10.5539/mas.v9n13p148.