

Data Entries and Location Indication of Product Using Arduino

¹ Kishori Kokate, ² Surekha Thube, ³ Tejashree Kautkar

¹ Electronics and telecommunication, SPPU university, Sandip Foundation, Nashik, India

² Electronics and telecommunication, SPPU university, Sandip Foundation, Nashik, India

³ Electronics and telecommunication, SPPU university, Sandip Foundation, Nashik, India

Abstract - The increasing population and thus the wide expansion of commercial areas have increased need of maintain data in proper form. Data entries are held in different areas like pharmaceutical industries, shopping malls, etc. In this paper literature review for optical character recognition techniques has been done. Using those techniques with help of controller data entries will take place. System objective is to replace manual work and to reduce time. If user has to search particular product location, it can easily search with blinking of LED assigned to that particular block.

Keywords - camera, optical character recognition, template matching, Feature Extraction, controller

1. Introduction

OCR is developed to replace human function by machines like reading. OCR is the successful application for reading by machine in field of Converts the pixel representation of a letter into its equivalent character representation. Image is capture by camera. Segmentation, pre- processing classification, feature extraction takes place on the basis of text matching process [1]. OCR technique is used to extracts numbers from number plate [2]. In bill processing system OCR is used to read payment slips like electrical bills, it will read and recognize amount to be paid and also recognize account number [3]. Manually data entries is most common way for inputting data into computer which is time consuming and labor intensive process. Optical character recognition is the machine replication for the human reading [4]. Vertical character overlapping separation after identification of document text (10,-12). This block only have text search or two lines of white pixel. Separation of overlapping character and touching is done [5]. Pen-Centric Shorthand Handwriting method is used for sub alphabet of upper case of English handwritten. Shorthand method is writing rapidly by substituting characters [6]. Principle component analysis (PCA) is the efficient and simple method for image compression for all images. It is the

orthogonal projection lower dimensional linear space for data. In this paper 'Generalized Hebbian' Algorithm is used [7]. Template Matching is common task for pattern matching. In paper various template matching technique are used like Native Template Matching (NTM), Image Co- relation Matching (ICM) Pattern Co-relation Image (PCI) Grayscale Based Matching (GBM) and Edge Based Matching template matching used for automated reading of data[8]. Optical Character Recognition OCR is developed for machine replication human function like reading. OCR is the successful application for reading by machine in field of pattern recognition. Traditional way of entering data in computer is through the keyboard which is not so efficient. OCR system is required when both humans and machine are able to read information [1]. Feature extraction method is used to classify the text. For classification of Emails dimension reduction techniques are used. For better detection from collection of emails we used reduction technique to keep most instructive and discriminative [9]. In this paper fast method for text feature extraction is discussed. Hash feature requires less memory and less computation. In clustering text feature extraction is done which determines words or terms that occur in each document [11].

2. Existing System

The following figure 1 show the existing system of OCR function. The details of each block are explained below:

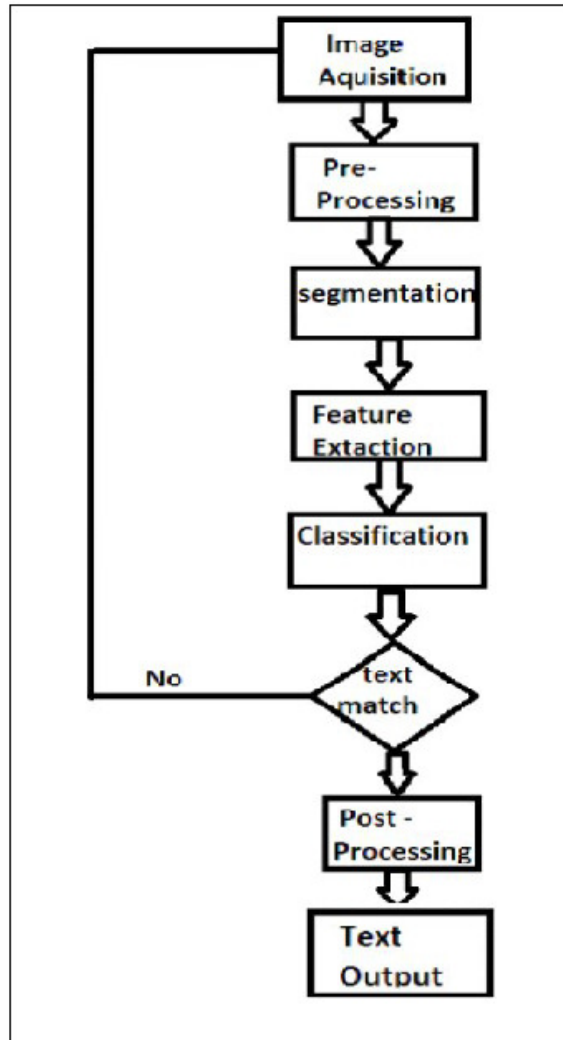


Fig. 1 Flowchart

A. Image Acquisition:

Input image (text image) is captured using electronic devices such as optical (digital/video) mobile phones, camera and scanner. The images will be stored as image file formats like PNG, TIF, JPG/JPEG, BMP etc.

B. Pre-processing:

Pre-processing means avoid interference of noise and improves the image quality.

C. segmentation:

For Optical Character Recognition the blocks are segmented into lines, lines into words and then into individual characters

D. Feature Extraction:

Features are a set of numbers that take the salient characteristics of the segmented image. There are different feature extraction methods for character recognition.

E. Text matching:

Extracted text from the image through the OCR is matched with the database.

F. Post-processing:

In Post-processing string formation of characters is done from the extracted words.

G. Text Output:

Resulting string from the post-processing is generated as a required output which is display in the Notepad. This output string is then used for data entry in tabular form.

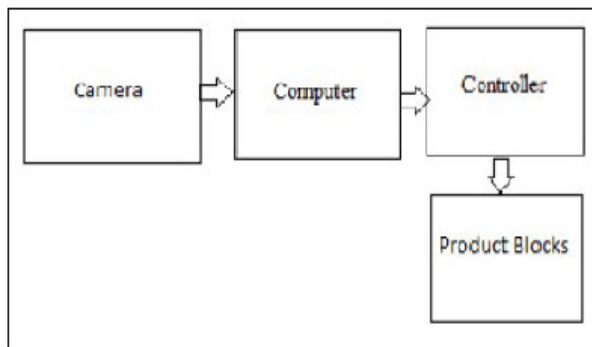
For extraction of data from scanned image we are using OCR techniques. For extraction of feature from the text various techniques can be used.

In [9] this paper proposed work by authors is based on dimension reduction technique. For better detection from collection of emails we used reduction technique to keep most instructive and discriminative .There are two feature selection technique -chi-square and information gain ratio. Feature extraction technique principle component analysis (PCA) is well known technique that can reduce the dimensionality of data by transforming the original features. Second one is latent semantics analysis [LSA] method which is a novel technique in text classification. Another author in his paper [10] proposed document classification techniques. For improving the document classification feature extraction acquires an important subset of feature from data set. In this paper a term frequency (TF) with stemmer based feature Extraction algorithm for high dimensionality

.Text mining is one type of data mining. Text mining extracts only text data from huge volume of data. Feature extraction technique is proposed for Documents classification to improve accuracy and reduce processing time. Feature extraction is used for improve its scalability, accuracy and efficiency. In [11], author describes a fast method for text feature extraction. Hash feature requires less computation and less memory. In clustering text feature extraction is done which determines words or terms that occur in each document. Full text indexing' method is performing feature extraction on large volumes of files. The indexing needs extracts the sequence of words in text feature are integer rather than string.

3. Proposed Work

The hardware layout of the proposed system



is shown in figure

Fig. 2 Block diagram of proposed system

Using optical character recognition and pattern matching technique, specifications of the products are extracted and saved in tabular format. This system is used to find out location of the product. Webcam is connected to laptop or computer. Wi-Fi module is connected to the controller. PC will act as a server whereas controller (Arduino) will act as client. Project module is consisting of number of blocks (e.g. one to ten) in which product or tablets will be placed. LED is assigned to each block. If user entered name of product, that particular location's LED will glow & product will easily find out.

4. Algorithm

The algorithm contains following steps:

- Σ Give input as image
- Σ Removal of noise
- Σ Segmentation & feature extraction

- Σ Classification of data through Text matching
- Σ Output of image in text format
- Σ Data entries in tabular form.
- Σ Enter the name of product which is to be searched.
- Σ Position/Location of the

corresponding product is shown by blinking LED

5. Conclusion

A data entry of the product using OCR technique eliminates not only human efforts but also provides Error free operation. It is less time consuming compared to traditional way of entering entries using keyboard. This system is automatic so reduce lots of manpower. This system can be implemented at Medical shops, Library, Wear house and at many more places where the no of things to be stored systematically.

References

- [1] Line Eikvil, "OPTICAL CHARACTER RECOGNITION" , in IEEE International [1993].
- [2] Namrata Dave, "Segmentation Method for Hand Written Character Recognition using various segmentation methodologies", in International Journal of Signal Processing, Image Processing and Pattern Recognition [2015].
- [3] Raja Vikramdeep Singh and Navneet Randhawa "Automobile Number Plate Recognition and Extraction Using Optical Character Recognition system", in International Journal of Scientific & Technology Research Volume 3 [2014]
- [4] Dr. Jangala, Sasi Kiran, N. Vijaya Kumar and N. Sashi Prabha , "Literature Survey on Digital Image Processing Techniques in Character Recognition of Indian languages" , in International Journal of Computer Science and Information Technologies [2015].
- [5] N. Papamarkos and T. Koutalios Department of Electrical and Computer Engineering Democritus University of Thrace, "SEPARATION OF OVERLAPPING CHARACTERS", in IEEE Vol. 80, No. 7, pp. 1066-1078[2000]
- [6] Charles C. Tappert and Jean R. Ward Pace University, NY, USA Pen Computing Consultant, MA, USA, "Pen-Centric Shorthand Handwriting Recognition Interfaces" , in 0-7695-2942-9/07 \$25.00 © 2007 IEEE DOI 10.1109/PLT.2007.37[2007].
- [7] Chih-Wen Wang Dept. of Inform. Eng., I-Shou University Kaohsiung, Taiwan Jyh-Horng Jeng Dept. of Inform. Eng., I-Shou University Kaohsiung, Taiwan , "Image compression using PCA with clustering", in IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2012)

- [2012]
- [8] Nazil Perveen, Darshan Kumar and Ishan Bhardwaj Assistant Professor, CSE Department , Chouksey Engineering College, Bilaspur (C.G) India, “ An Overview on Template Matching Methodologies and its Applications ”, in International Journal of Research in Computer and Communication Technology, Vol 2, Issue 10, [OCT 2013]
- [9] Masoumeh Zareapoor Department of Computer Science, Jamia Hamdard, New Delhi, India. Seeja K. R Department of Computer Science & Engineering, Indira Gandhi Delhi Technical University for Women, New Delhi, India, “Feature Extraction or Feature Selection for Text Classification: A Case Study on Phishing Email Detection” I.J. Information Engineering and Electronic Business, 2015, 2, 60-65 Published Online March 2015 in MECS (<http://www.mecs-press.org/>)DOI: 10.5815/ijieeb.2015.02.08 [2015]
- [10] Jebamalar Leavlin Epiphany, Anna University, Tiruchirappalli, India, “Feature Extraction for document classification”
- [11] George Forman Hewlett-Packard Labs Palo Alto, CA, USA ghforman@hpl.hp.com Evan Kirshenbaum Hewlett-Packard Labs Palo Alto, CA, USA evankirshenbaum@hp.com “Extremely Fast Text Feature Extraction for Classification and Indexing” presented at Conference on Information & Knowledge Management, Napa [2008]
- Kishori Kokate** pursuing B.E. Degree in electronics and telecommunication from Sandip Institute of Technology and Research Centre , Nasik ,India.
- Surekha Thube** pursuing B.E. Degree in electronics and telecommunication from Sandip Institute of Technology and Research Centre , Nasik ,India.
- Tejashree Kautkar** pursuing B.E. Degree in electronics and telecommunication from Sandip Institute of Technology and Research Centre , Nasik ,India.