International Journal of Multidisciplinary Research and Development
Online ISSN: 2349-4182 Print ISSN: 2349-5979
www.allsubjectjournal.com
Volume 3; Issue 2; February 2016; Page No. 01-05

An advanced Lab view based OCR system for hand written scripts

[1] Aruna Panda, [2] Prof. Mishra SN, [3] Prof. Das

[1] M.Tech Scholar, Dept. of VLSI & Embedded Systems, KIIT, Bhubaneswar, India.

[2, 3] Dept. of VLSI & Embedded Systems, KIIT, Bhubaneswar, Odisha, India.

Abstract
Optical character recognition usually abbreviated to OCR, involves a computer system designed to translate images of typewritten text (usually captured by a scanner) into machine editable text or to translate pictures of characters into a standard encoding scheme representing them. OCR began as a field of research in artificial intelligence and computational vision. Character recognition is very interesting area of pattern recognition and it deals with offline handwriting recognition. Handwriting Recognition has continued to Persist as a suggests that of communication, collection, recording and sending data in regular life since the centuries even with the arrival of the new technologies. Machine recognition has several sensible applications, reading written communication envelopes, amount written in bank checks, bill processing, government records, industrial forms, signature verification, offline document recognition etc.
In this paper we have implemented OCR for the Odia character recognition using the appearance based recognition technique. Formally, the problem can be stated as follows: given a training dataset x, and an object o, find object x_j, within the dataset, most similar too. Our technique is an appearance based recognition giving satisfactory results

Keywords: Odia Character Recognition, online, offline, pre-processing, classifiers.

1. Introduction
India is a versatile country. In India over 20 official languages are there, Bengali, Malayalam, Hindi, English, Guajarati, Tamil, Kannada, Urdu etc [1]. Odia language is extremely popular language and it's a political candidate language of the Odisha State of India. More than 50 million people speak Odia language [2]. We tend to will notice that work has been done in the field of numerous languages like Chinese, Tamil, Telugu, and South Dravidian etc however terribly less work is done in the field of Odia language recognition.
Pattern recognition has become a very interesting topic for researchers during last few decades. Typed characters will be simply recognized by laptop machine. But handwritten characters don't seem to be recognized efficiently and accurately by computer machine. Many researchers have done work to recognize these characters and many algorithms have been proposed to recognize characters. For over 30 years, researchers have been working on handwritten recognition. Over the few past years, the numbers of firms concerned in analysis on written recognition square measure increasing regularly. Challenges in handwritten characters recognition lie in the variation and distortion of offline handwritten characters since different people may use different style of handwriting.

2. Odia Script
Odia-script used to write the Odia language. The Odia alphabet utilizes overall thirty distinct legitimate and recognized shapes, which mainly includes 59 Characters and 16 diacritics. There are many applications of this type of recognition. Like postal code verification, vehicle range plate recognition, bank cheque processing Assigning ZIP Codes to letter mail, Reading data entered in forms, e.g., tax forms, Automatic accounting procedures used in process utility bills, Verification of account numbers, Automatic accounting of airline passenger tickets, Automatic validation of passports Etc [6].
In particular, machines that can read symbols are very cost effective. A machine that reads banking checks will method several additional checks than a human being in the same time. This type of application saves time and money, and eliminates the necessity that somebody's perform such a repetitive task [1].
Odia digits are having different characteristics. They are having various shapes and it's really difficult to recognize those shapes. Due to varieties in shapes there are some characters that are confusing and possibilities for misclassification are very high [5]. Figure 1 shows the Odia alphabets and digits.

Fig 1: Odia Characters and Odia Digits

3. Application of Character Recognition System
There are number of applications of Character Recognition System [1].

Task-specific Readers
It is basically used for voluminous data processing. It focuses on the specific application like:

- Assigning ZIP codes to letter mail.

- Reading data entered in forms, e.g. tax forms
- Verification of account numbers and courtesy amounts on bank checks
- Automatic accounting procedure used in processing utility bills
- Automatic accounting of airline passenger tickets
- Automatic validation of passports

Address Readers
The address reader in a postal service reads the destination address block on the envelope and also reads the PIN code in the address block. Then using the PIN code it can sort the envelopes according to the area.

Form Reader
Form Reader automatically reads the data filled up in the form. It can find the printed and handwritten text in the form and also recognizes the same.

Check Reader
Automatic check reader reads amount and account information from the check image and recognizes the amount as well as the account information.

Signature verifier
Just like check reader there is a signature verifier that also reads the signature image from the check image and recognizes the same.

Bill Processing System
This system is basically used to read payment slips, bills or any value specified in the bill.

Passport Readers
An automatic passport reader can be used for the inspection purpose. That can verify the traveler's information like name, age, passport number and also photograph image which saves time at the airports.

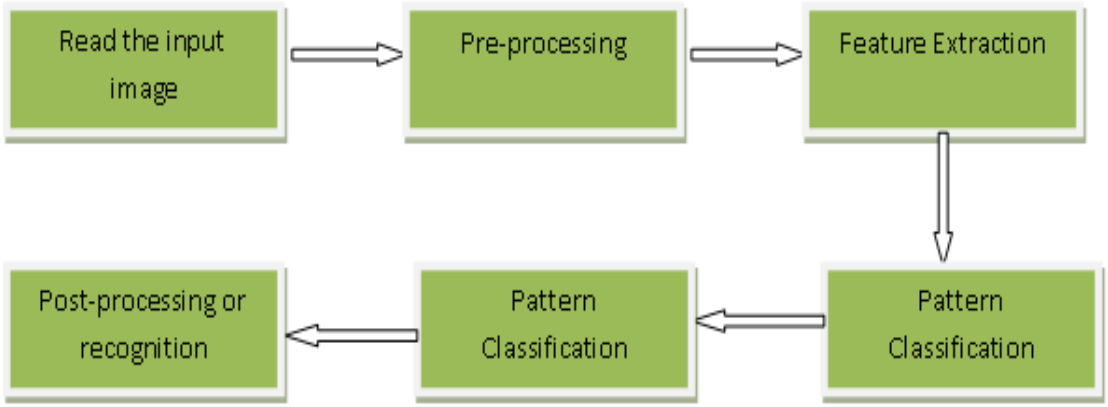

Fig 2: Components of the character recognition system

4. Adopted Character Recognition Scheme
The Pre-processing step aims to improve the image data or the image features that required for the further processing. The pre-processing is a series of operations performed on the scanned input image. It essentially enhances the image. It involves converting an input image to binary image, noise removing, dilation operation, line segmentation and digit segmentation and normalization [9].
Feature Extraction is a very important step for any character recognition system. This step involves the procedures like shape information or style which is very much useful for the classification of the pattern. The feature extraction stage analyses a text segment and selects a set of features that can be used to uniquely identify the text segment [10].
Classification stage uses the features extracted to identify the text segment according to the algorithm. The task is to compare the testing patterns and minimizing the error rate and correct classification of the pattern.
Post-processing involves various approaches dictionary lookup and statistical approach or neural network recognition [11] for the correct recognition.

5. Lab view Based Implementations
Lab view
Lab View is a graphical programming language developed by National Instruments sometime in the mid to late 80's by Jeff Kodosky. A program in Lab View is called a VI, which stands for Virtural Instrument. To create a VI, the programmer uses the Lab View programming environment to make the user interface by dragging and dropping objects, and arranging them as desired. To add functionality to the interface, the diagram, which resembles a flow chart is "wired" with the various structures and functions. So, in most Lab View programs, no lines of code are written, the functionality of the program is provided by the diagram. For this reason Lab View is called a graphical programming language.

Ni Vision Assistance: Code Generation
This step is simply creating a database of letters for the character recognition of a complete para and converting it into editable text format

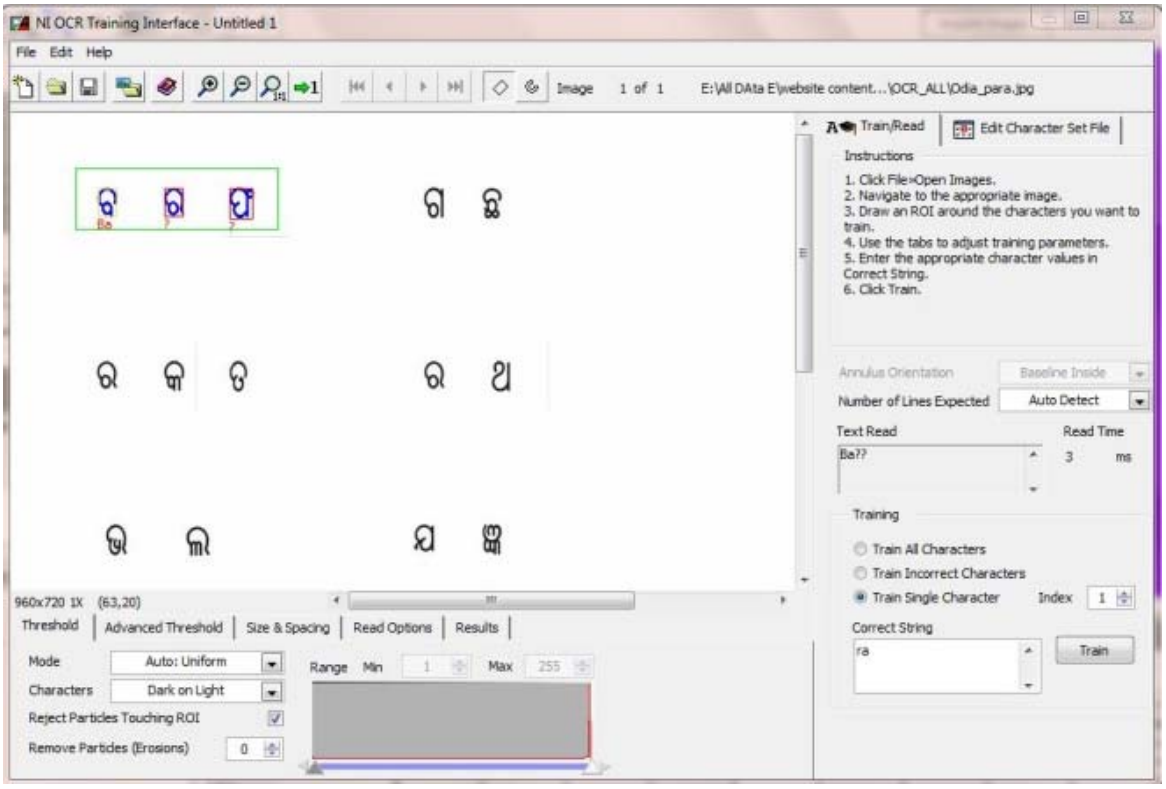

Fig 3: NI Vision Code generation

Ni Lab view based OCR Block

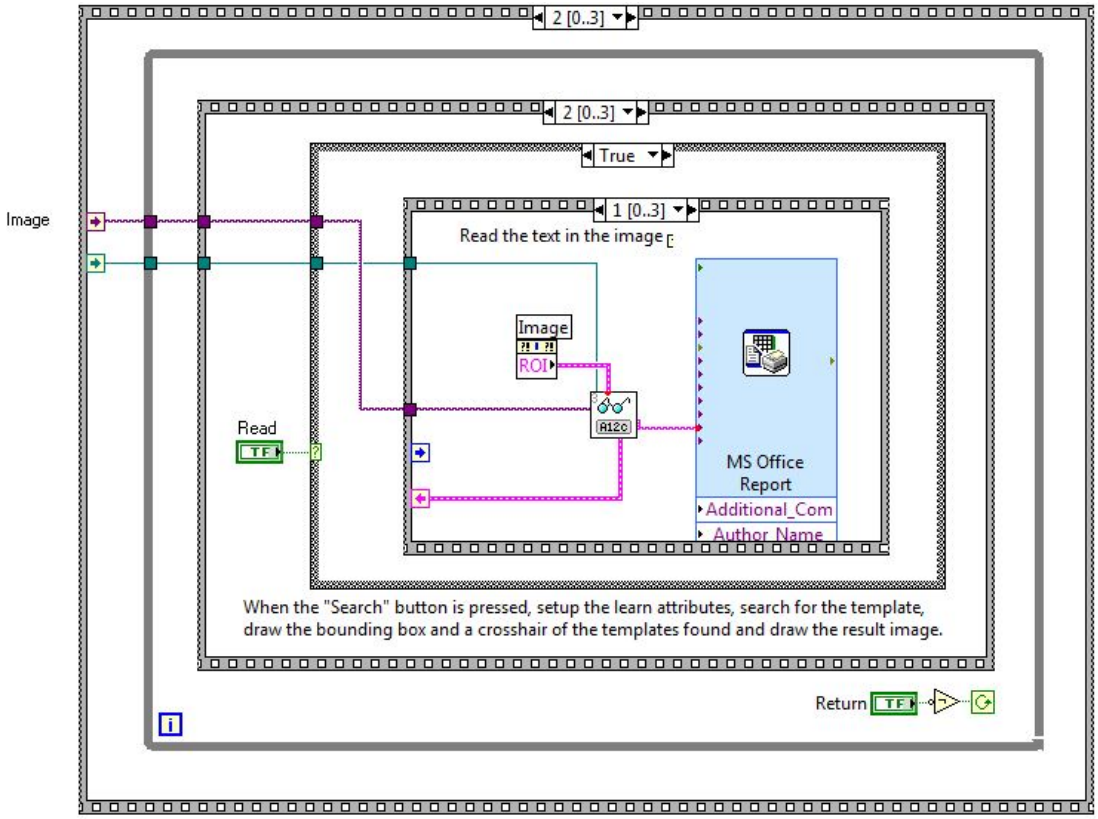

Fig 4: NI Lab view OCR block

Ni Lab view Based OCR Front Panel

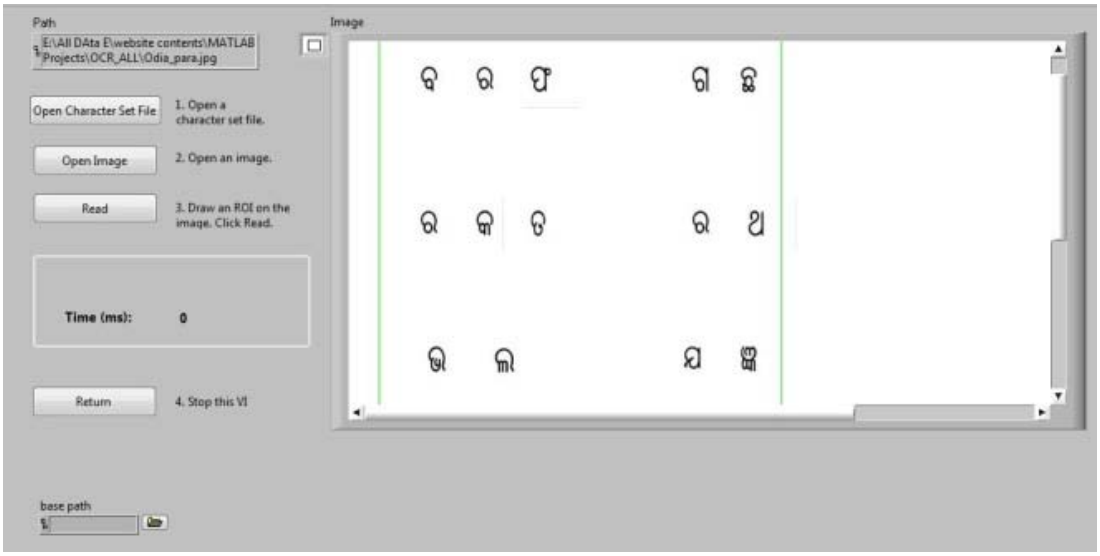

Fig 5: NI Labview OCR front panel

Editable Output for the Odia-OCR

Barapha Gacha
Rakata Rath
Bhala Jagnya
By
Operator :

6. Conclusion

This paper presents the various steps involved in the Odia Optical character recognition system. It also describes the various applications of the character recognition system. Last section presents the proposed Labview based implementations for the Odia language OCR system from Code generation upto getting the editable text output.

7. References

1. Sanjay D, Savithri TS, Kumar PR. Person follower robotic system. In Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014 International Conference on IEEE, 2014; pp. 1324-1327.
2. Mariappan M, Ramu V, Ganesan T, Khoo B, Vellian K. Virtual Medical Instrument for OTOROB Based on Labview for Acquiring Multiple Medical Instrument LCD Reading Using Optical Character Recognition. In Proc. International conference on Biomedical Engineering and Technology (IPCBEE), IACSIT press, Singapore 2011; 11:70-74.
3. Singla SK, Yadav RK. Optical Character Recognition Based Speech Synthesis System Using LabVIEW. Journal of applied research and technology. 2014; 12(5):919-926.
4. Singh P, Kaur R. Hybrid algorithm for alphanumeric character recognition for vehicles using LabVIEW. In Computational Intelligence and Computing Research (ICCIC), 2014 IEEE International Conference on IEEE, 2014; pp. 1-6.
5. na Chahar R, Bhavsar HG. Automatic License Plate Recoganization System Based on Image Processing Using LabVIEW. International Journal of Advanced Research in Computer Science and Software Engineering. 2014; 4(4).
6. Izák P, Hrianka M. Biomedical Image Analysis by Program" Vision Assistant" and" Labview". Advances in Electrical and Electronic Engineering 2005; 4(4):233.
7. Vaishnavi D, Karthik K, Ramkumar E, Nivethitha I. Language Translator Using OCR In LabVIEW. International Journal for Innovative Research in Science and Technology. 2015; 1(8):185-188.
8. Online and off-line Handwriting Recognition: A comprehensive survey, Rejean Plamondon, Fellow IEEE, sargur Srihari, fellow IEEE, IEEE transactions on pattern analysis and machine intelligence, 2000; 22(1).
9. Indian script character recognition: a survey Pal U, Chaudhuri BB, Pal U, Chaudhuri BB. Pattern Recognition 2004; 37:1887-1899.
10. Offline Typed Odia Character Recognition, Manish Kayasth, Dr. Bankim Patel. 2009; 2(1):0974-3308.
11. CIA/DOE Partnership Program Proposal for FY99 (Sandia National Laboratories Proposal), 1998.
12. Srihari SN. Recognition of Handwritten and Machineprinted Text for Postal Address Interpretation, Pattern Recognition Letters 1993; 14:291-302.
13. Genetic algorithm for feature selection and weighgting for off-line character recognition, Thesis, Faten T. Hussein, Egypt, 1995.
14. Rafael Gonzalez C, Richard woods E, Steven Eddins L. Digital Image Processing using MATLAB, Pearson Education, Dorling Kindersley, South Asia, 2004.
15. Kharma N, ward R. Character recognition systems for the Non-expert, in IEEE Canadian Review 1999; 33:5-8.
16. Rajput KY, Sangeeta Mishra. Recognition and Editing of Devnagari Handwriting Using Neural Network, Proceedings of SPIT-IEEE Colloum and International Conference, Mumbai, India, Vol. 1, 66.

17. Antani S, Agnihotri L. Guajarati character recognition, Proceedings of the Fifth International Conference on Document Analysis and Recognition, 1999; pp.418-421.
18. Diagonal Feature Extraction Based Handwritten Character System Using Neural Network, Pradeep J, Srinivasan E, Himavathi S. International Journal of Computer Applications. 2010; 8(9):0975-8887.
19. Hanmandlu M, Mohan KRM, Kumar H. Neuralbased Handwritten character recognition, in Proceedings of Fifth IEEE International Conference on Document Analysis and Recognition, ICDAR, Bangalore, India 1999; 99:241-244.
20. Comparison of Classifiers for Odia Numeral Recognition Baheti MJ, Kale KV, Jadhav ME. International Journal of Machine Intelligence. Issn: 0975-2927 & E-Issn: 0975-9166, 2011; 3(3):160-163.