EXTRACTION AND RECOGNITION OF HANDWRITTEN ALPHANUMERIC CHARACTERS FROM APPLICATION FORMS

(SPONSORED BY K.S.C.S.T., BANGALORE)

A Project Report
Submitted in partial fulfillment of the requirements
for the award of the Degree of Bachelor of
Engineering in Electronics & Communication Engineering
of the Visvesvaraya Technological University, Belgaum.

Submitted by

Shashank Hegde
Shivanand V. Kundral
Raghavendra R.
Rajashekhar S. Udadar

Under the Guidance Of

Prof. Nataraj Vijapur

DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING

K.L.E SOCIETY'S
COLLEGE OF ENGINEERING AND TECHNOLOGY
UDYAMBAG, BELGAUM – 590 008

VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELGAUM

2011 - 2012

ABSTRACT

Character recognition has been the subject of intensive research during the last decades. This is not only because it is a very challenging scientific problem, but also it provides a solution for processing of large volumes of data automatically. Nowadays character recognition even of handwritten script has numerous applications such as Application Form Processing, Address And Zip Code Recognition,Number Plate Recognition & Writer Identification, etc.

In this venture it was thought to develop a reading system which is capable of extracting the handwritten text from application forms and of recognizing the alphanumeric characters. The application forms can be scanned and the handwritten parts can be automatically separated.

Because of the many uncertainties in handwritten character recognition, stochastic modeling is a suitable method for approaching this problem. The character recognition is based on discrete hidden Markov models (HMM). In this direction work was done but due to time limitation, limited progress was achieved. A system was developed which can automatically detect Vehicle license Plate!!!!.

The main objective of our project is to recognize characters. And we have achieved the same using template matching. It's a simpler way of recognition by using the template and correlation function to find out the similarity between the character that needs to be recognized and the template we have in our database, once the closest match is found it is outputted in a text file. Here we need not train the system about the characteristics of the system in contrast to the HMM model, therefore it is less time consuming and easier to implement, and the capacity and the accuracy can be increased by just updating the database with more templates.