Development of Vehicle License Plate Recognition System Using K-Means Clustering and Support Vector Machine

Sheela C Loni
Student (PG), Dept. of Telecommunication Engineering
Ramaiah Institute of Technology
Bangalore, India

S.J. Krishna Prasad
Associate Professor, Dept. of Telecommunication Engineering
Ramaiah Institute of Technology
Bangalore, India

***Abstract*— License plate recognition (LPR) technique is a technique which is used to detect, identify and recognize the characters. This work focuses on recognition of Indian cars license plate recognition. This LPR system mainly consists of following sub stages Plate localization, Preprocessing of captured image, grouping of the characters, feature extraction, classification, associated testing and character recognition. In this work, we propose a novel hierarchical character recognition scheme based on supervised K-means and Support Vector Machine (SVM) which not only recognize clear images but also blurred license plate images. Supervised K-means has been used to classify characters into subgroups. Characters of subgroups are further classified by support vector machine based algorithm. Advantage of the proposed approach is to reduce the classes of characters in each subgroup and to further reduce their complexity. This improves accuracy of character recognition. Our algorithms were tested with 100 images both blurred and non blurred image cases captured on the road and parking lot. The algorithm designed exhibits 100% success rate in Indian license plate recognition.**

***Keywords*— Image preprocessing, Morphological approach, K-means clustering, Bounding box and SVM.**

I. INTRODUCTION

In recent years, use of license plate recognition technology is extensively made in traffic monitoring system, vehicle parking and tracking systems. It has attracted lot of attention because it can be used in smart cities to enact on criminal investigation, traffic detection and identification of stolen cars. Therefore, lots of new algorithms have been designed in order to detect the vehicle number plate automatically and to recognize the vehicle numbers in number plate. Despite these algorithms, differences in rules, shape, font and format for the number plate in a certain countries has lead to false detection for quite lot of algorithms. This work mainly aims on Indian vehicles number plates. There are many methods on car number plate detection which have been proposed such as template matching, edge information, morphological operators, optical character recognition, Hough transform and neural networks. License plate recognition systems have been realized on both hardware and software approaches. Main aim of this paper is to focus on the software design aspects. Software design component of our license plate recognition system developed include three steps: License plate detection, applying K-means clustering and SVM.

In one of the work Yang Guang et.al [1] recognized the character using Least Square SVM. In another work Kolour [2] has developed number of license plate detection algorithms. In yet another work Singh M [3] developed an efficient approach on morphological operations. In this work first localization of vehicle plate number in image has been undertaken and then segmentation process of alphanumeric characters are achieved. Recognition is done using the template matching technique. In yet another work, Chittode et.al [4] developed an algorithm on the basis of morphological operations which has been used for number plate recognition. In yet another work Amir et.al [5], developed support vector machine concept in license plate recognition. Support vector machine (SVMs) is the machine learning algorithm used for object classification and regression. Another approach of character recognition is by using SVM and K-means clustering which provides better accuracy [6].

II. PROPOSED METHOD

Fig.1 shows the block diagram for character recognition using SVM and K-means clustering. Here one can capture real time image using named Cyber link You Cam 6 compatible with windows OS and which is interfaced to MATLAB. This is an inbuilt camera which is available in standard laptops. One need to interface this camera with MATLAB from the image acquisition toolbox. The proposed algorithm indicated in Fig. 1uses following steps:

Pre-processing: In this step, we need to convert the RGB image into grey scale image and then send to median filter for noise removal. Once this is done, one needs to perform convolution operation for intensity adjustment and perform thinning operation in order to remove the unwanted portion of

an image that is the removal of horizontal and vertical lines present in license plate image.

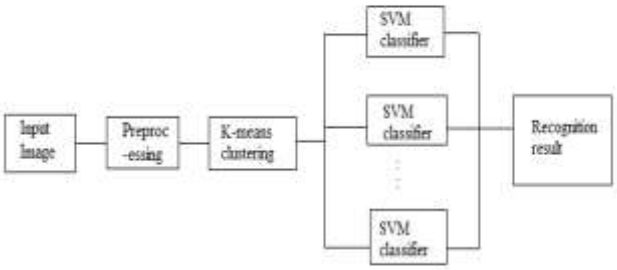

Fig. 1 Block diagram of LPR system

Extraction of characters: In this step, we need to extract all the characters from vehicle number plate. This step is known as localization. Since we are interested only in vehicle license plate number, hence one need to remove all the unwanted portion of an image. Once characters are obtained then, we need to extract the features for all the characters, which are extracting the area of all the characters. This extraction step is inbuilt with preprocessing stage.

K-means clustering: K-means clustering algorithm for grouping the similar characters features is applied in the next step. It helps in differentiating the characters which are having same features and it increases the accuracy in character recognition. The procedure follows a simple and easy way to classify the given data set through maximum 6 clustered groups, such as K clustered group. Each cluster group will have different features incorporated in it. Initially for all the k clusters we need to choose centroids. Centroid is a data point at the center of the clusters. We need to place these centers as much as possible far away from each other. Next step is to take the each point belonging to given data set and associate it with nearest center. Here in this, first we need to extract area of all the characters. One should have a vision of different blurring conditions in the license plate images given as a input to the algorithm. Irrespective of all blurring conditions characters such as medium, less and more blur, grouping of all the characters be done. Grouping is on the basis of the similar features belonging to the specific characters. To perform grouping following formula is been employed.

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} \left(\left\| x_i - v_j \right\| \right)^2 \quad (1)$$

Where $||x_i - v_{i||}$ is the Euclidean distance between xi and vi.

Ci is the number of data points in I^{th} cluster.

C is the number of cluster centers.

J(V) is a variable which is used in the algorithm for grouping.

SVM classification and recognition: This step includes three steps, Feature extraction, Classification and Feature testing. First, features of all the characters are extracted from K-means clustering technique. In K means clustering we algorithmically calculate areas of all the alphanumeric characters. Then perform the classification method for differentiating the similar objects in one subgroup. Here, each character of the input image (Irrespective of the blurring conditions) is sent to its own classifier and then it identifies to which of sub groups this input character belongs to. Comparison is further made with the stored image feature value and input image feature value. This is known as feature testing. If the value matches then it outputs respective character. This method gives the better character recognition as compared to template matching technique irrespective of the blurring conditions.

III. RESULTS

This section illustrates the results obtained from the algorithm developed in MATLAB for the blurred input image of a license plate. The following figures indicates image transformation happening after individual steps execution as indicated in Fig. 1.

Fig. 2 Blurred input image

Fig. 3 Restored blurred image

Fig. 4 Morphological image

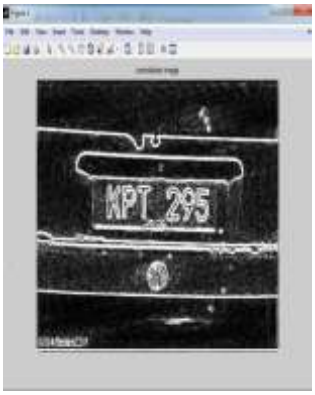

Fig. 5 Convolution image

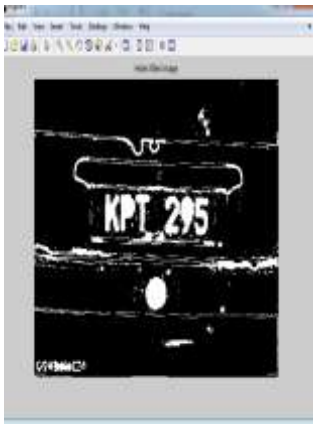

Fig. 6 Hole filled image

Fig. 7 Final preprocessed image

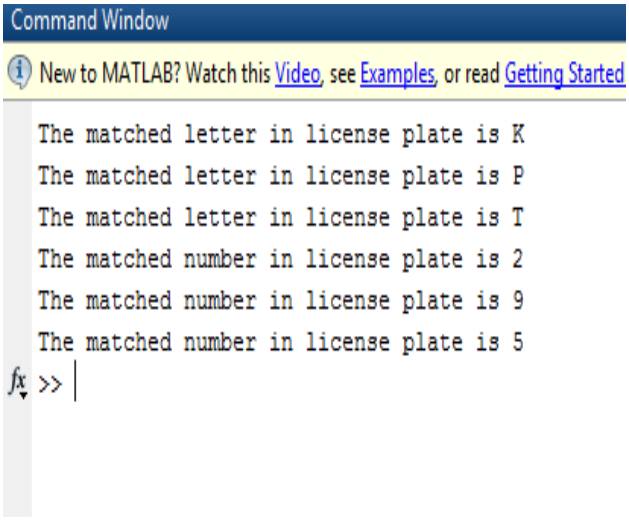

Fig. 8 Output results obtained after recognition

Steps of the algorithm:

- Convert extracted plate image into gray scale image.
- In next step, pass it through wiener filter to restore the blurred image.
- Then further pass it through the median filter for noise removal.
- Apply morphological gradient for edge enhancement. Perform dilation and erosion of an image and then subtract both the images.
- Convolve this subtracted image with [1 1; 1 1] for brightening the edges.
- Erode with horizontal line as structuring element to eliminate the possible horizontal lines from the output image.
- Subtract eroded image from convolved image.
- Fill all the regions of the image.
- Do thinning on the image to ensure character isolation.
- Then extract the features for all the characters like area and group the similar features using K means step. Here we are going to consider the feature value as area. First we need to measure the areas of all the characters. This is known as feature extraction.
- Then we need to perform the SVM classification method where each SVM classifiers will have all the characters extracted area values in it.
- Using the following mat lab functions we can extract the areas of the character.
  1. Region props
  2. Extract field
- The last step is to test the extracted feature value with input image character values and identify the respective character. This is known as feature testing.
- After all the characters are tested, the respective license plate is recognized.

IV.CONCLUSION AND FUTURE WORK

LPR system mainly consists of three phases license plate recognition, character segmentation and character recognition in an individual license plate. This work mainly focuses on the Indian license plates. Algorithm has been tested with 100 images taken from parking lot and image data base. Here we have proposed algorithm on character recognition scheme combined with K-means and SVM for license plate recognition irrespective of blurring conditions. This method also proves for blurred images and hence it provides better accuracy compared to template matching technique. Template matching technique will not be a suitable technique to deal with blurred image. The future vision of this work is that, we can extend this for tilted, hazy images and characters of various language. This can be extended to traffic surveillance, vehicle identification, automated tolling application and intelligent transport systems.

ACKNOWLEDGMENT

This work was supported by MSRIT, Bangalore and Dept. of Telecommunication Engineering.

REFERENCES

[1] Yang Guang, “License Plate Character Recognition Based on Wavelet Kernel LS-SVM”, School of Communication and Information Engineering Xi'an University of Posts and Telecommunications Xi'an, ©2011 IEEE.
[2] Hadi Sharifi Kolour.: An Evaluation of License Plate Recognition Algorithms. In: International Journal of Digital Information and Wireless Communications (IJDIWC) (2011).
[3] Pandya and M Singh,” Morphology based approach to recognize number plates in India,” International Journal of Soft Computing and Engineering,Vol-1, Issue-3, pp 107-113, June2011.
[4] J.S. Chittode and R. Kate, “Number plate recognition using segmentation,” International Journal of Engineering Research & Technology, Vol. 1 Issue 9, November- 2012.

[5] Amir Ebrahimi Ghahnavieh, "Enhancing the License Plates Character Recognition Methods by Means of SVM" in the 22nd Iranian Conference on Electrical Engineering (ICEE 2014), May 20-22, 2014, Shahid Beheshti University.

[6] Wei-Chen Liu et.al, "A Hierarchical License Plate Recognition System Using Supervised K-means and Support Vector Machine" in Proceedings of the 2017 IEEE International Conference on Applied System Innovation IEEE-ICASI 2017.