Digital Image Analysis- CSL 461

Final Project Report

Number plate recognition using image processing and machine learning

Group members:
Viren Gupta - 2014CSB1080
Vikram Singh - 2014CSB1119

Problem Statement:

We aim to first segment out the number plate from the rear view of the car using image processing techniques. Then we want to extract the exact number from that section using character segmentation followed by learning a model to predict the numbers using machine learning.

Methodology:

Image processing(using MATLAB):

These are the following steps that we have followed.

- RGB to gray and gray to binary
- Remove top ⅓ part of image. We made the assumption that usually the number plate is present in lower ⅔ part of image.
- Segment out the number plate from the image
    - We assume that usually the number plate is a region with highest number of black to white transitions.We applied the following algorithm:
    - Hence we find an upper limit of rectangle and a lower limit of rectangle in which the number of black to white transitions are atleast 20 and maximum 48. These calculations have been done based on the characters pattern in Indian number plate system. This rectangle is the actual number plate region from the original image..
- Smoothing using median filter
- Convolution with 4*4 mask for contrast Enhancements
- Horizontal lines removal using eroding
    - This will reduce the border lines of plate and any irrelevant lines
- filling of holes : This helps us to make a character a single connected component

- ➢ finding connected components using 8 connected neighbourhood.
- ➢ Extract all the connected components into a 20*20 images with each image a single connected component. This is done by using bounding box concept.
- ➢ Applying threshold to remove irrelevant images:
    - We applied a threshold of 10 pixels to remove smaller and irrelevant shapes.

Machine learning(using python)

We have characters as 20*20 pixels images. Now we learn the svm model. Dataset is created by us . The model predicts each character from 1-9 and A-Z. We have taken a threshold of the matching probability to be 0.11. I.e if matching probability of an instance is less than this threshold, we take that character as an irrelevant image and do not consider as a valid character.

Experiments:

- ➢ Used many images of car back view for experiments.
- ➢ If characters in number plate are close enough, then the 2 characters merge into one connected component and cause a problem.
- ➢ Tried different algorithms for number plate extraction using assumption made earlier.
- ➢ Self created Dataset.
- ➢ Assuming that top ⅓ image does not contain number plate works quite well.

Result:

We are able to extract the number plate from the given image and recognize the corresponding character. Cases which may fail are:

- ➔ Less separation between the characters
- ➔ Orientation of plate with rotation.
- ➔ If lot of transitions in background scene are present even after removing top ⅓ part of the image.

We have attached a folder of some sample images on which we have tested our method. There are 9 images out of which 5 work correctly and give the correct number on the plate, 3 give partial correct results and 1 is there which gives incorrect result because of less separation between characters.

Conclusion:

Our algorithm works on some assumptions. In cases where vehicle image is a bit far, we are able to extract the number plate correctly but character recognition fails due to small size of the characters.