M.S. Thesis Proposal

A *"Rules Engine"* for

Autonomous Vehicles in Urban Traffic

by

M. Caner Kurtul

B.S. in Computer Engineering, Boğaziçi University, 2000

Submitted to the Institute for Graduate Studies in

Science and Engineering in partial fulfillment of

the requirements for the degree of

Master of Science

in

Computer Engineering

Boğaziçi University

2009

1. INTRODUCTION

Autonomous driving is one of the popular research topics studied by the Artificial Intelligence, Computer Vision and Robotics professionals. The researches are focused either on off-road driving or driving in urban traffic. Thanks to the DARPA Grand Challenge and the DARPA Urban Challenge, significant progress have been made in both domains. Autonomous vehicles equipped with several cameras, sensors, and processors are proved to move sucessfully from a starting point to a predefined destination.

There is remarkable amount of work regarding autonomous driving and its sub-tasks. Most of these researches target the task of moving the car from one point to other, just by eliminating the crashes and taking the most efficient path. This requires efficient path planning and obstacle avoidance algorithms, but no necessity for traffic sign or pedestrian recognition. DARPA Urban Challenge has mandated some specific rules, most importantly the "lane following", but has not covered the traffic rules as a whole. Recognition of traffic lights and signs, and recognition of pedestrians are officially leaved out of scope.

2. MOTIVATION

As mentioned above, most of the current researches focus on the autonomous driving systems. Autonomous vehicles seem to be possible for the future but there seems to be a gap till the vehicles, drivers and roads become appropriate for absolute autonomy. Till then, a transitive solution, which can be applied in near future, is required . This can be a *"Rules Engine"* to evaluate how successful or legal a car is being driven.

Such a *"Rules Engine"* can have various usage domains. It can be used as a means for training atonomous vehicles, in real traffic or other traffic simulators. Regarding the DARPA Urban Challenge vehicles, we can easily say that, they miss a rules engine to evaluate how successful the vehicles navigate in the urban. Our *"Rules Engine"* could have been used as an autonomous referee during the challenge.

Another application area for the *"Rules Engine"* can be the collective transportation vehicles, such as school busses or inter-city coaches. By putting a device on such vehicles, these vehicles can be observed and drivers can be punished more accurately. Such an option would help drivers to avoid traffic rule violations.

3. THESIS PROPOSAL

This work is focused on building a *"Rules Engine"* that measures how successful or legal a car is being driven in urban traffic. It will evaluate the driver, and give a score at the end of the track. This evaulation will be based on the number of challenges and the errors made regarding them. Note that, the challenges are not predefined. They will be the ones simultaneously sensed by the *"Rules Engine"*. As the car navigates on the road, it will encounter unforeseen signs, lane changes and traffic signs.

The work will cover the following sub-tasks:

- Lane Detection: The driver should obey the lane separator lines. If a lane departure occurs when the lane is straight, the driver will be charged for it.
- Traffic Lights: The red/yellow/green lights will be detected by the *"Rules Engine"*. The driver will be evaluated against the red light violations.
- Traffic Signs: The *"Rules Engine"* will be trained for specific traffic signs. The signs will be grouped either as stateful or stateless. For instance a speed limit sign is a stateful one. When the system is running, it will keep track of the most recent speed limit sign. The driver will be evaluated with regard to the most recent signs. On the other hand, left-right turn limitation signs are stateless. These signs will be considerded as long as they are visible. When they move out of the camera vision, their limitations will fade away as well.

4. LITERATURE SURVEY

Other than the DARPA stuff, there are several important works on the topic. Since the topic is related to every action a human driver can perform, the works are spread over a wide range of research areas, such as:

- Depth Perception with Monocular Vision or Stereo Vision
- Obstacle Avoidance algorithms to avoid collision
- Path Planning to efficiently navigate to the target destination
- Noise Reduction in processed images
- Localization according to predefined objects
- Building systems that can be trained by learning algorithms
- Traffic Sign detection (day and night)
- GPS and Laser Range Finder usage

4.1. Lane Detection

In [6] Jeong *et al.*, have proposed a model that uses inverse perspective transform to get rid of the perspective effect on images. They apply least-mean square method to approximate the lanes. In [7] Aufrere *et al.* uses a monocular camera to track the lanes with a statistical model, which is supported by the localization of the vehicle by tracking the number of lanes. Hiren M. Mandalia and Dario D. Salvucci [8] present an SVM (Support Vector Machines) technique to detect the lane departures. In another research [9], Tolga Birdal and Aytul Ercil have segmented out the image with respect to the gray value difference and several texture analysis methods.

4.1.1. Proposed Approach

The proposed approach in this MS Thesis employs Multiresolution Hough Transform (MHT) for lane detection, followed by two Hidden Markov Models (HMM)for radius and orientation of the candidate lanes.

Multiresolution Hough Transform

The classical Hough transformation approach processes the entire vision data in order to detect the lines. This scenario has two main drawbacks. First, the occluded lines (i.e. another car passing through the line) become noisy since the transformed relative intensity of the line decreases. Second, the relative intensity of the lines also decreases at the curves in the road. The proposed solution divides the road image into partitions, where the sizes of the partitions are inversely proportional to the distance of the partition to the vehicle.

After the image is partitioned, several preprocessing steps are required before applying the Hough transform. These preprocessing steps should be fast because the Hough transform is already computationally expensive for real time applications. Since edge detection techniques are usually computationally expensive for real time applications, each partition is converted to binary images via applying a threshold filter after a color remapping process.

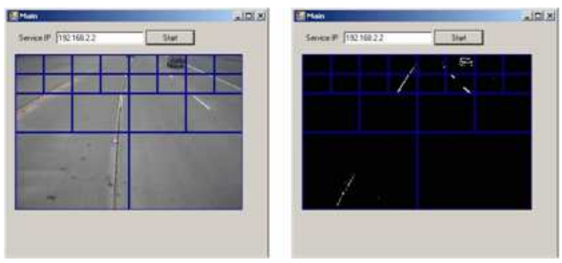

Figure 4.1. (a) Partitioned image, (b) Binary image.

After the image is partitioned, a separate Hough transformation is applied to each single partition. The most intense line in each partition, which is the candidate line segment, is taken into consideration in order to find the global lanes in the image. Since the Hough lines are represented in polar coordinates *(r, θ)* instead of rectangular coordinates *(x, y)*, the candidate lines are grouped according to their slopes and distances to the center of the image as well as their intensities. The center of the frame is chosen as the bottom point.

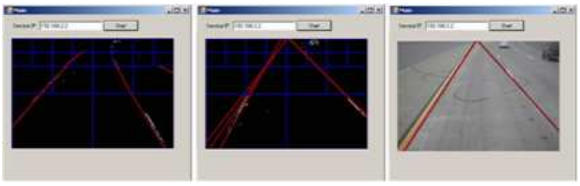

Figure 4.2. (a) Candidate lines, (b) Transformed line, (c) Detected lines.

The transformation of the lines basically changes the center point of the polar coordinates for each transformed line which is achieved by the following translation

$$\begin{aligned} r' &= r + (x - x')\cos(\theta) + (y - y')\sin(\theta) \\ \theta' &= \theta \end{aligned} \qquad (4.1)$$

where *(r', θ')* is the polar coordinates of the transformed Hough line *(r, θ)*. Note that the translation of the center of the Hough transformation is from *(x, y)* to *(x', y')*.

After the lines are grouped, the most intense three clusters are assigned as the lanes. However, there may be less than three lanes if the sum of the intensities of the candidate lines is less than a threshold value.

Lane Tracking

For lane tracking, HMM is used to represent the relation between the current frame and its successor. Each line in a specific frame is represented by an individual *(r, θ)* pair. In the succeeding frame, the process will most probably observe the same line at *(r', θ')* which is not very far from the position of the line in the previous frame. The probability of observing *(r', θ')* pair in the next frame is modeled as an HMM problem. In addition, θ and r values are modeled by two different HMM. The θ value is discretized as (0, 1, 2, 3...178, 179) where the r value is discretized at the pixel level. This discretization schema is used in both transmission and emission matrixes. The emission probability matrix shows the probability of observing θ' (or r') in the next frame, having observed θ (or r) in the current frame. In our implementation, the observation and state transition matrix values are derived from two Gaussian distributions with different deviations. The deviation of the transition matrix is assigned to a smaller value than the observation matrix, which means, the state transition matrix aims to preserve the current state where the observation matrix promotes the exploration behavior.

4.2. Traffic Light Detection

Traffic light detection mostly consists of color detection, and the closure of the red, yellow, green lights in the image. For traffic light recognition, Lindner *et al.* [10] have proposed a system that is support by GPS and enhanced digital maps. K. Lu *et al.* [11] have handled the topic in two steps. Firstly, they have detected the light location by region labeling the HSI color-space image. Afterwards classification is done by border detection. Obtained region borders are used as matching features.

4.3. Traffic Sign Detection

Traffic sign detection basically has two major steps: the detection and the classification. Xavier Bar and Jordi Vitri [1] proposes an *"object detection system"* based on rectangular features and radial symmetry. They pass grayscale images through a cascaded classifier. H. Fleyeh [2, 3] presents a fuzzy approach to the sign detection problem. This work uses the color variation cues in HSV color space. Hue and saturation channels are used to generate a binary image of detected traffic signs. A. Soetedjo and K. Yamada [4] proposed a *"ring partitioning"* method for the classification problem. They convert the RGB image to grayscale, and divide the image into several rings starting from the center. The histograms of the rings are used to assign weights to each ring. The method is especially good at handling occluded signs. The research in [5] proposes a successful SVM (Support Vector Machines) approach to the classification of the traffic signs. Their system has yielded good results on the noisy images and even for the damaged, rotated, and translated traffic signs.

4.3.1. Proposed Approach

This research will propose a hybrid approach for traffic sign detection. First, Genetic algorithm (GA) will locate the candidate signs in the video stream, and then a radial symmetry will be applied to prevent false positives. Since the fitness function is the redness of each chromosome, the GA is likely to detect fully red regions as traffic signs. This is where the radial symmetry check comes into the scene.

The initial version of the algorithm is currently implemented. The algorithm runs on RGB images. Initially the chromosomes are distributed randomly over the image. At each iteration the fitness values are calculated according to how red the surrounding of the chromosome is. After 5 iterations half of the chromosomes (most fit ones) are transferred to the next frame. The selected chromosomes are expected to be converged around the circular signs.

Figure 4.3. Detected Traffic Sign

5. TRAINING DOMAINS

5.1. RoboChamps

The *"Rules Engine"* will initially be trained in Microsoft RoboChamps Urban Challenge simulation environment. This environment will primarily help the development of lane detection, and traffic light detection algorithms using image processing techniques. The Microsoft RoboChamps Platform is based on Microsoft Visual C♯ development environment. Hence, the *"Rules Engine"* will be developed in C♯. The simulator is support by KIA Motor Company. It simulator has several sensors: stereo vision, GPS, bumper sensors and laser range finder. Only the vision sensors will be used for the *"Rules Engine"* development.

The disadvantage of RoboChamps environment is that, it provides perfect ligtning conditions, and perfect sensory data. Also lane markers are perfect lines. The algorithms trained with this perfect data may not function well in real-world environment.

5.2. Precaptured Video

For real-world training, precaptured videos are used. The videos are captured from a car moving in the urban traffic with a speed of 50 km/hour. The camera is placed onto the front console of the car. The captured video has a resolution of 512x288 pixels with a frame rate of 29.97.

As opposed to the simulated environment, precaptured video sequence provides noisy data with imperfect ligtning conditions. The tests with precaptured video has shown that, lightning conditions will have major effect on the performance of the overall system.

6. SCHEDULE

- Up to now:
  - Literature Survey
  - Implementation of sign detection methods
  - Conference paper on lane detection
  - Ant colony optimization papers studied (not applicable for realtime sign detection)
- 2009 Summer:
  - Implementation for traffic light detection
  - More experiments for algorithms
- 2009 Fall:
  - Implementation for Physical Environment
  - Runs on Physical Environment
  - Collect test data
  - Writing of Thesis Report

7. CONCLUSIONS

This work aims to build a *"Rules Engine"* to evaluate how successful or legal a car is being driven in urban traffic. As of any other real-world applications, urban driving also introduces several challenges. Most striking of them is handling the lightning changes as the car travels. This little tiny problem will have considerable impact on the performance of the overall system. For example, light coming from the front side causes the red traffic signs to look black. Shadowy roads also makes it difficult to distinguish the colors of interest.

Regarding the lane tracking problem, the most noteworthy challenge is the bending roads, which causes the lanes to bend as well. The MHT is fundamentally capable of detecting the straight lines only. Hence the MHT cannot handle the bending lanes correctly. The HMM enhances the results by introducing some statistical correction to the MHT results. But there is still some work to enhance the accuracy of the lane tracking system.

REFERENCES

1. Xavier Bar, and Jordi Vitri, "Fast Traffic Sign Detection on Grayscale Images", *Centre de Visi per Computador+ Dept. Informtica UAB, Barcelona, Catalonia, Spain.*

2. Hasan Fleyeh, "Road and Traffic Sign Color Detection and Segmentation - A Fuzzy Approach", *Department of Computer Engineering, Dalarna University, Borlnge, Sweden.*

3. Hasan Fleyeh, "COLOR DETECTION AND SEGMENTATION FOR ROAD AND TRAFFIC SIGNS", *Department of Computer Engineering, RuEarna University, Sweden.*

4. Aryuanto Soetedjo and Koichi Yamada, "Traffic Sign Classification Using Ring Partitioned Method", *Laboratory for Management and Information Systems Science, Nagaoka University of Technology.*

5. Min Shi and Haifeng Wu and Hasan Fleyeh, "A Robust Model for Traffic Signs Recognition Based on Support Vector Machines", *Image and Signal Processing, Congress on, 4, 516-524.*

6. Seung Gweon Jeong, Chang Sup Kim, Kang Sup Yoon, Jong Nyun Lee, Jong I1 Bae, and Man Hyung Lee, "Real - Time Lane Detection for Autonomous Navigation", *2001 IEEE Intelligent Transportation Systems Conference Proceedings - Oakland (CA), USA - August 25-29, 2001.*

7. Romuald Aufrkre, Roland Chapuis and Frkdkric Chausse, "Real time vision based road lane detection and tracking", *University B. Pascal, FRANCE.*

8. Hiren M. Mandalia and Dario D. Salvucci, "Using support vector machines for lane change detection", *In Proceedings of the Human Factors and Ergonomics Society*

49th Annual Meeting, 2005.

9. T. Birdal, A. Eril, "REAL-TIME AUTOMATED ROAD, LANE and CAR DETECTION for AUTONOMOUS DRIVING", *Faculty of Engineering and Natural Sciences, Sabanci University.*

10. Lindner, F. and Kressel, U. and Kaelberer, "Robust recognition of traffic signals", *Intelligent Vehicles Symposium, 2004 IEEE.*

11. Kuo-Hao Lu, Chao-Ming Wang and Shu-Yuan Chen, "TRAFFIC LIGHT RECOGNITION", *Journal of the Chinese Institute of Engineers, Vol. 31, No. 6, pp. 1069-1075 (2008).*