The Driving Force Behind Autonomous Driving – An Analysis of On-Road Lane Detection Systems

Joseph Schueller
Department of Computer Science
University of Wisconsin – Platteville
Platteville, WI 53818
schuellerjo@uwplatt.edu

Abstract

Lane detection while driving on a road is trivial for a human being but is a complex task for a computer. A human is able to instantly spot patterns and recognize the importance of those patterns. A camera sees the road as a set of pixels. Converting the pixels to actual geometrical patterns includes filtering out visual noise, finding lane markings, determining road boundaries, fitting these patterns to models, and tracking the changes between frames. Lane detection systems must perform quickly, and they have to work on various road conditions, which can range from structured highways to unstructured gravel roads. Autonomous driving is a growing field, and the demand for safety and efficiency drives research in the lane detection field.

Introduction

Lane Detection Systems were originally developed as Lane Departure Warning Systems. These warning systems would warn the driver audibly, visually, or physically if they were veering out of the lane. These systems are widely used in semi-trucks in Europe and North America. If the truck driver veered out of the lane without using a turn signal, the interior speakers of the truck would simulate the sound of a rumble strip. [8]

Lane Departure Warning Systems are becoming more common. They can be found in many newer vehicles today. These systems have upgraded to the point of actually taking control of the steering wheel if it detects the car moving out of the lane. Semi-autonomous cars have the added safety, but most require the driver to maintain control of the car. However there are newer cars that allow "autopilot" unassisted driving under

certain conditions. [4] For future cars to become completely autonomous, Lane Detection Systems must perform reliably and efficiently in all situations.

Active vs Passive Systems

There are two branches in modern lane detection systems: active and passive. Active systems use sensors, such as lasers and sonar, to gather and map the data of the surrounding area. These types of sensors are accurate and fast. They are able to process a lot of data in a short amount of time. The major drawbacks to active systems is that they are prone to interference. Weather and even other cars with similar sensors can interfere with an active systems. Also, equipment for active systems are more expensive than equipment for passive systems. For these reasons, research in passive systems has been a growing topic. Passive systems do not introduce any interference to the surrounding area and only use visual data as input. Most passive systems rely on dash-mounted cameras to retrieve data. Passive systems have the benefit of being reasonably priced and readily available. Also, video frames contain a wealth of information that can be useful in lane detection systems. In this paper, we will focus on passive systems.

Detection

Almost all lane detection systems that operate passively share four common steps: *Preprocess, Feature Detection, Fitting,* and *Tracking*. [5] Please remember that the only input data for a passive system is the video feed from the onboard camera(s). The four steps are essential when designing lane detection system that can perform well in real life situations.

Problems

There are many problems faced when designing lane detection systems. The first is the need for perfect performance in every situation. If a car is to be completely autonomous, lane detection must perform reliably for the safety of the passengers. Mistakes in the system can have disastrous results.

The second problem is the wide variety of road conditions. Lane detections systems perform well on structured roads, that is, roads with clear lane markings and boundaries. However, lane detection systems must also perform well on unstructured road, such as the back country road with no lane markings.

The third problem is the complexity of detecting the road and lanes. There is a high processing cost in applying lane detection systems to each frame of a video feed. Computer systems are limited in computational power. In addition to dealing with complex algorithms, all of these must be carried out in real-time. A car will travel about 74 feet in one second when traveling at 50 miles per hour. Computational speed and efficiency are important topics to bear in mind when considering lane detection systems. This reaction time constraint has pushed researchers to improve computational speed and efficiency.

Preprocess

Preprocessing has two main goals: remove visual noise and prepare the image for the next steps. Noise could include shadows, road discoloration, or other random noise in the image. The image can be shrunk down to only the essential parts in an effort to increase computational efficiency.

Binarization

One way to reduce noise is to use a binarization function to convert the image to a two-toned image. A binarization function will go through each pixel and determine if it should be redrawn as white or black. For example, a grayscale image will contain pixels ranging in intensity from 0 (black) up to 255 (white). The binarization function would have a threshold, say 100. Any pixel at or below 100 would be redrawn as black. Pixels above 100 would be white. Using this technique, the image would be redrawn in two colors. [5] [7] Important features of the image, such as bright lane markings on the road, will stand out and be picked up easier in the Feature Detection step and the other following steps.

Figure 1: Original Image - Business US Highway 151, Platteville, WI

Figure 2: Binary Image - Business US Highway 151, Platteville, WI

Blurring

Other ways of reducing image noise it to apply a blur to the image, usually done with a Gaussian Blur. Commonly used in image processing and graphic design, it is very effective in removing random "salt and pepper noise" from the image. [5]

Regions of Interest

To increase efficiency, the image may be cut in size to only a portion of the original image. This portion is called a Region of Interest (ROI). A ROI is focused on the road, and cuts out the upper part of the image which usually contains only sky. By cutting down the image size, subsequent steps have less image to process, and therefore cut down on the computational complexity in the future steps. ROIs may be shaped in a trapezoid to fit around the view of the road or multiple ROIs may be shaped around the probable edges and lane markings of a road. [5]
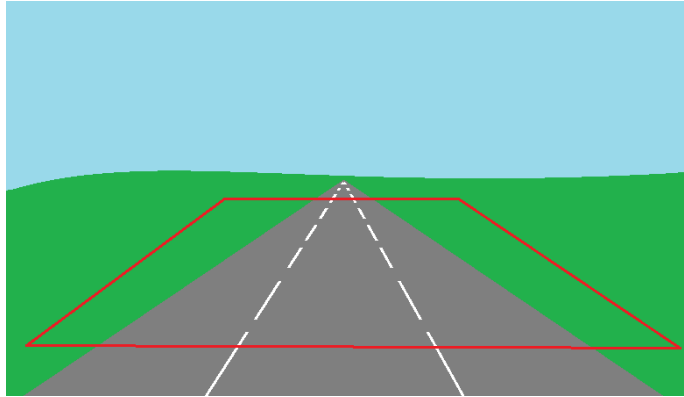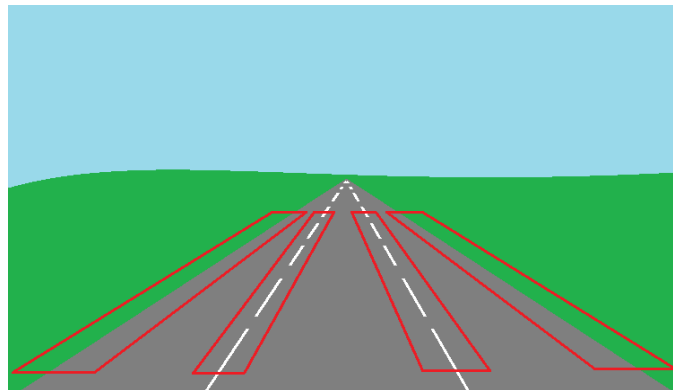
Figure 3: Region of Interest - Trapezoid

Figure 4: Region of Interest - Edges and Lane Markings

Inverse Perspective Mapping

To simplify other steps, Inverse Perspective Mapping (IPM) can be used. IPM maps pixels from a three dimensional view of the road to a two dimensional top-down view of the road. This top-down view is very useful in creating and fitting road models to incoming images. [1] [2]

Feature Detection

The next step, feature detection, is the process of finding and selecting various features from the images. This could include anything of use for autonomous driving, including lane markings, road edges, vanishing points, signs, and other cars.

Feature Selection

Color is rarely used in feature detection because of its high computational costs. Grayscaled images are used because they usually have better resolution, and the same type of information can be gathered from both grayscaled and colored images. To simplify algorithms and increase speed, most designers of lane detection systems prefer to use grayscale images.

A major part of feature detection is edge detection. This is how the system finds the boundaries of the road and lane. Edge detection performs well on structured roads, that is, roads with clear lane markings and edges. However, operating algorithms of edge detection on unstructured roads suffer greatly. The difference between road and nonroad may not totally be clear on unstructured roads. Also there are usually no lane markings on unstructured roads. Another problem many edge detection systems face is picking up false edges. False edges can be anything from bad lighting, shadows, weather, trees, telephone poles, or even mountain slopes. An edge detection algorithm may pick up a false edge and may end up believing that is the boundary of the road. These problems can usually be lessened in the preprocessing stage where visual noise is removed.

Feature Extraction

In feature extraction, important features, such as road areas, road markings, or road boundaries, are extracted from the image using various filters or statistical methods. Feature extraction can be broken down into three main classes: area-based methods, edge-based methods, and area-edge combined methods. All methods attempt to extract the feature from the image, but they each do it in a slightly different way.

Area-based methods attempt to group pixels into different sets of clusters based on either intensity or color. The Iterative Self-Organizing Data Analysis Technique (ISODATA) is one way of breaking an image down into different areas. Based on the k-means algorithm, ISODATA algorithm attempts to partition similar pixels together in the same cluster. The ISODATA algorithm is the same as the k-means algorithm except the ISODATA algorithm allows for a different number of clusters while k-means has a set number of clusters. The algorithm is iterative meaning that the cluster definitions are redefined in each iteration. Each cluster is redefined as the cluster's mean. The pixels are then regrouped under the newly defined clusters. The process of redefining the clusters and regrouping pixels is carried out until no pixels change clusters. The ISODATA algorithm is a great way of defining different areas in a picture. Using the clustering method, similar pixels will be redrawn using the same color. When this is carried out on an image certain areas will stand out from others. Below is the pseudocode for the k-mean algorithm. [3]

. K-Means Pseudocode

1. Define an initial set of clusters C1, C2, C3…Cn. Each cluster has a value M which represents the mean of the data stored in the set.
2. Let D represent all of the input data (pixels).
3. For each input data D, classify D into the most similar cluster (C1, C2, C3…Cn) based on the cluster's M value.
4. For each cluster C1, C2, C3…Cn, set M value to be the mean of the input data classified under that set.
5. For each input data D, classify D into the most similar cluster (C1, C2, C3…Cn) based on the cluster's M value.
6. If any of the input data changed cluster, loop to step 4. Otherwise, success.

Edge-based feature detection has very good performance works very well, but it only works well on established roads. On unstructured roads, clear lines may not always be visible. The boundaries between road and nonroad are sometime fuzzy. These types of situations hinder feature detection using edge-based methods. Despite these drawbacks, edge-based methods continued to be used for their efficiency and accuracy on established roads.

The main way to detect lines and determine edges is to use the Hough Transformation technique. Suppose the incoming image has three data points marked out along the lane markings of the road as seen in Figure 5. To find a relationship between these points, we first draw all theoretical lines in a 360 degree fashion that run through each point as seen in Figure 6. Ideally all lines are drawn, but for the sake of this example, only 3 lines are drawn for each data point. You will notice that there is a line from each data point that crosses all three data points. Since the equation, $y = mx + b$, for these lines are the same, we draw the resulting line which is seen in Figure 7. [6] The Hough Transformation technique is usually used in finding the vanishing point and is also used as a preliminary stage to the next step, fitting. [5]
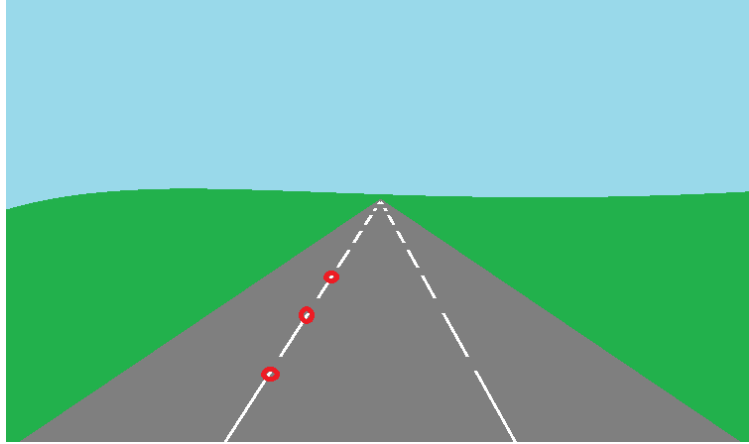
Figure 5: Hough Transformation - Selecting Three Data Points on a Lane Marking
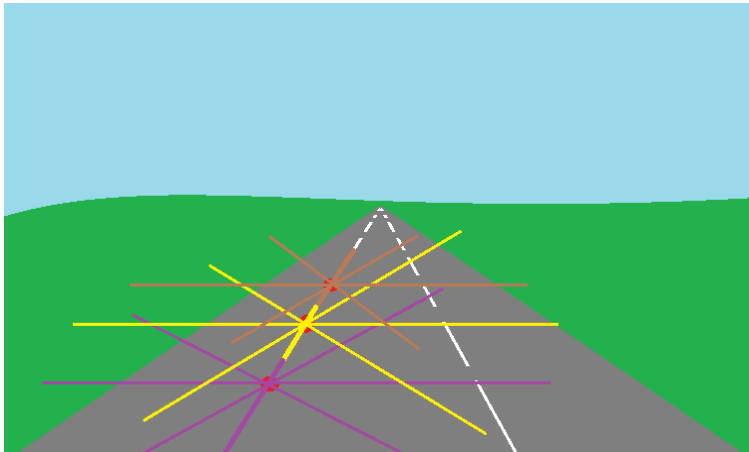

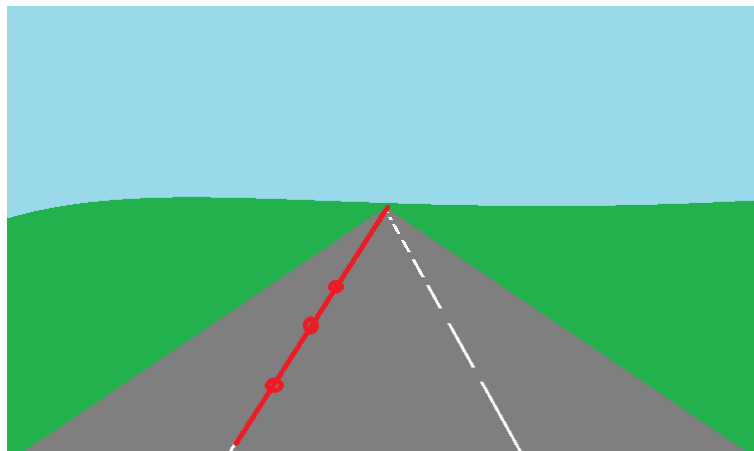Figure 6: Hough Transformation - Drawing All Lines Through Each Point

Figure 7: Hough Transformation - Resulting Line

Area-edge based methods use a combination of both area-based and edge based techniques. The idea is that when one technique falters, the other technique will be able compensate. While this does increase computational complexity, this sort of parallel redundancy is needed to create safe and reliable systems.

Fitting

In most research, lane detection systems are described as finding road and lanes without prior information about the road. In fitting, domain knowledge is used to apply constraints and fit predefined models to road. Applying domain knowledge to the lane detection system helps improve accuracy.

Domain Constraints

The symmetry constraint assumes that for each road boundary on one side of the image, there exists another road boundary on the other side of the image that is parallel.

The road width constraint uses outside data to estimate the width of the road. The outside data can come from previous systems that have driven the road or from accurate satellite imaging. By having an estimated road width, road edge detection knows generally where to find the edge of the road.

The vanishing point agreement constraint is when lines detected in a system all converge to one point. Any lines that do not converge to the vanishing point may be filtered out. For example, a shadow from a telephone pole may cause edge detection systems to find a horizontal line in the image. Since this line does not converge to a vanishing point, this line will be filtered out.

Predefined models

Models are used to represent the expected road and lanes from an incoming image. Most of the incoming images from a lane detection system are very similar, so it makes sense to use a predefined model that matches the current image. Models are geometrical shapes that closely resemble commonly seen situations on the road. For example, a straight line model uses the parallelism of the road to fit the model. A circular arc model uses the vanishing point as the center of the circle. Road edges are drawn from center of the circle

to the outside of the circle. Lane and road width data is commonly used to fit the incoming image to a model. [5]

Tracking

Lane detection systems can take one frame of video run the analysis on it. However, the use of tracking can reduce the computational cost in the analysis. Tracking uses previous frames and its corresponding data to estimate and help out future frames. The position of road features, such as road signs and lane markings, can be estimated by tracking their movement in previous frames. Instead of searching the entire image over and over for each frame, only certain areas of the image have to be searched to re-find these features.

Systems have used this idea to rethink the way Regions of Interest (ROI) are created. A new system of Lane Bounded Regions of Interest (LBROI) uses previous frames to create an ROI that surrounds the lane and road boundaries. This limits the search space that is to be used in the next frames of video. Lower computational cost and improved accuracy are the driving forces behind tracking.

Conclusion

While there are many methods and techniques that are used in lane detection systems, there does not exist one technique that will perform perfectly in every situation. Methods will have to be combined to produce a safe and reliable system. Also, methods should be able to know their limitations. When the system is falling below a certain tolerance level, the whole system should be switched off. This is for the safety of the driver, passengers, and others on the road. More powerful computers combined with newer, more efficient algorithms will help accelerate the growth of autonomous vehicles.

References

[1] Aly, M. (2008) Real time Detection of Lane Markers in Urban Streets. *IEEE Intelligent Vehicles Symposium*. 1 - 6.
http://www.vision.caltech.edu/malaa/publications/aly08realtime.pdf

[2] Muad, A., Mustaffa, M., Hussain, A., Majlis B. Y., & Samad, S.A. (2004 November). Implementation of Inverse Perspective Mapping Algorithm for the Development of an

Automatic Lane Tracking System. *2004 IEEE Region 10 Conference, Volume A.* 207 - 210. doi:10.1109/TENCON.2004.1414393

[3] Piech, C., (2012) K Means. *Stanford.* Retrieved from http://stanford.edu/~cpiech/cs221/handouts/kmeans.html

[4] Tesla Motor Teams (2014) Dual Motor Model S and Autopilot. Retrieved from http://www.teslamotors.com/blog/dual-motor-model-s-and-autopilot

[5] Yenikaya, S. Yenikaya G, & Düven E. Berndt, T. J. (2013). Keeping the Vehicle on the Road - A Survey on On-Road Detection Systems. *ACM Computing Surveys 46*, *1, Article 2,* 1 - 43. http://dx.doi.org/10.1145/2522968.2522970

[6] Yu, B. & Jain, A. K. (1997) Lane Boundary Detection Using A Multiresolution Hough Transform. *International Conference on Image Processing, Proceedings., Volume 2.* 748 - 751. doi:10.1109/ICIP.1997.638604

[7] (2009 March 23) Canny Edge Detection. Pages 1 - 7.0020http://www.cse.iitd.ernet.in/~pkalra/csl783/canny.pdf

[8] (2008 June 6) Lane Departure Warning System. Retrieved from http://usedsemitrailers.com/lane-departure-warning-system/