TECHNISCHE UNIVERSITÄT DRESDEN

Fakultät Informatik, Institut für technische Informatik, Student für Informatik

SEMINARARBEIT

Self-driving car technology introduction

Submitted by: Xi Hu

Matricle: 4755213

Date 03.07.2018

The main symbol comparison table

CPU	central processing units
GPU	graphics processing units
LIDAR	light detection and ranging
MPC	model predictive control
YOLO network	you only look once network
CNN	convolutional neural network
SegNet	segmented network
GPS	global positioning system
x	position in the x-axis direction
y	position in the y-axis direction
ψ	angle between the vehicle body and the x-axis
v	velocity
a	acceleration
δ	steering angle
z	coordinate of predicted position

CONTENTS

1 Motivation

Since 2010, with the development of deep learning technology and the improvement of computing power of Central Processing Units (CPU) and Graphics Processing Units (GPU), the use of algorithms for real-time image processing has also become an realistic thing. Thanks to that, the self-driving car technology based on those technologies is also gaining a high degree of development. This is the possibility of self-driving car technology. At the same time, self-driving is also a major trend in the future: More than 1.2 million people are killed every year because of traffic accidents. These accidents often occur because of the driver's negligence. If cars could be driven by computers instead of people, these accidents would happen less. The amount of accidents will be greatly reduced, which is the necessity of self-driving car technology. In addition, there is a very important point: self-driving technology also makes it possible for people with disabilities to travel. Therefore, an intuitive understanding of this technology will be introduced in this literature research.

2 Abstract

There are two main aspects in this literature research: 1. Road Lane detection 2. Collision Avoidance. One part of road lane detection includes the processing of images, which are captured by cameras placed on the top of the car[1], and another part uses the LIDAR (light detection and ranging) images. Combining the results of both images, the effect of "road lane detection" will be improved[2]. In the collision avoidance part, the information contained in this literature research mainly introduces the theory of model predictive control (MPC)[3][4] and artificial potential field[3]. The purpose of those theories is to plan a path, which can avoid obstacles. These two parts are very important aspects for self-driving technology.

3 Methods for self-driving cars

3.1 Road Lane Detection

a) Image warping process and Polynomial regression[1]

In order to achieve self-driving, it is necessary to detect the lane line in a warped image, and obtain the curvature of the lane line through polynomial fitting. The process of warping the images is handled by changing the perspective of the input images where it tries to approximate the new images' points based on the RoI (Region of Interest) images using polynomial fitting. By getting right and left points, it is possible to make functions to approximate the lane structures by feeding them into the regression method to find the best fit of the road lane detection. The process can be shown as follows:

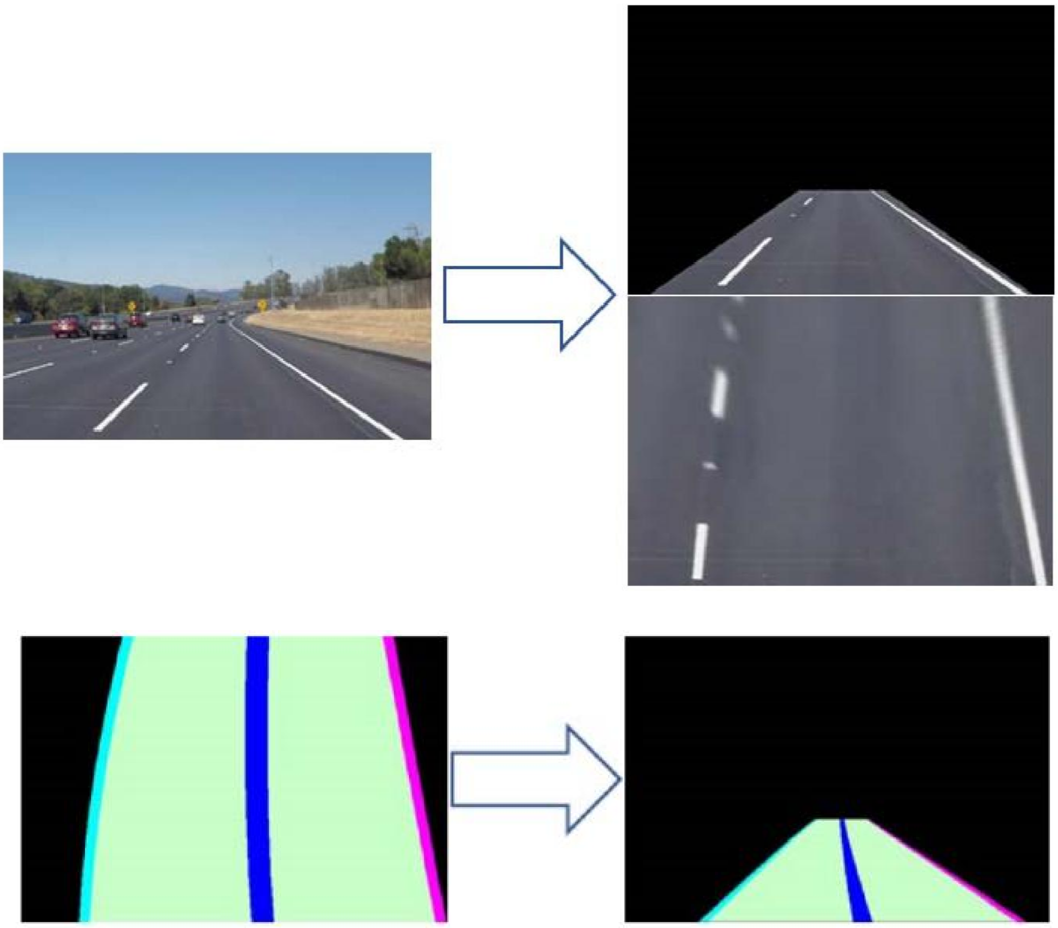

fig1[1].result of road line detection and polynomial regression. With changing perspective, the lane line in bird's-eye view is obtained, which is more suitable to fit lane line with polynomial regression. Then the result of regression can be applied in original image.

b) YOLO network[1]

Besides detecting lane lines, detecting vehicles in images is also a very important task, which can be solved by YOLO (you only look once) network. YOLO is a state-of-the-art, real-time object detection system. Concretely, YOLO implements a deep CNN (convolutional neural network[7]) architecture for its object detection approach. It consists of 24 convolutional layers (using filters to perform convolution on the image) along with 2 fully connected layers, where it is actually inspired by the GoogLeNet[7] model. However, the author of YOLO changes the approach from inception to simply 1x1 reduction layers along with 3x3 convolutional layers. They also have trained a lightweight YOLO for fast object detection where they use less than half convolutional layers along with fewer filters. The architecture of YOLO for the object detection as fig2.

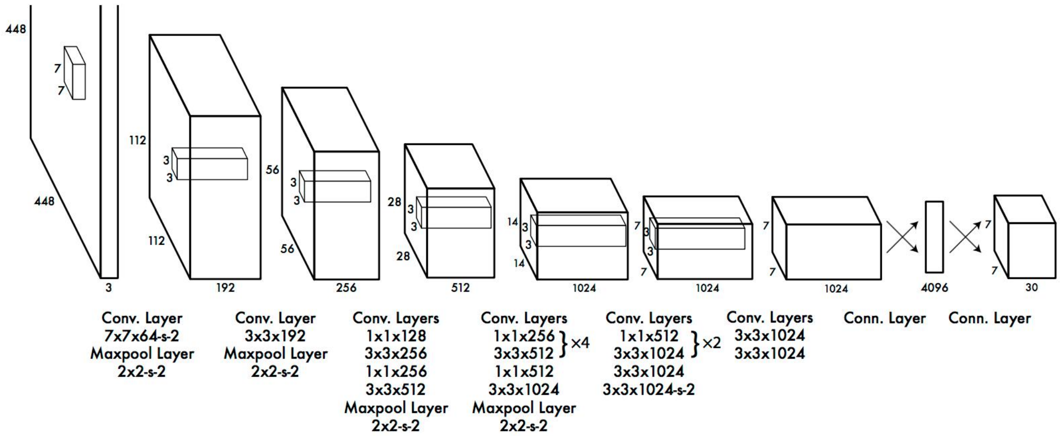

fig2[1].YOLO's original architecture for the object detection

c) SegNet[2]

In order to more accurately detect vehicles on the road, SegNet[8] is used (segmented network) to segment the objects in images. SegNet is a deep convolutional encoder-decoder architecture for image segmentation. Its architecture uses an encoder-decoder network that is followed by a pixelwise classification layer, where the encoder and decoder networks consists of 13 convolutional layers each. In other words, semantic image segmentation means, that the algorithm automatically separates the object area from the image and identifies the content in this area. It is used to improve visual lane detection. For example, the output of SegNet is shown as fig3.

fig3[2].SegNet's input (left) and output (right). With result of SegNet, the accuracy of lane detection can be improved, because of reducing of noise in the road. (for example: noise causing by different light intensity)

d) LIDAR Technology[2]

LIDAR (Light Detection and Ranging), means implementing detection and ranging with laser. The working principle of LIDAR is the time of flight method, which uses the reentry time after the laser encounters target to calculate the relative distance between the target and the self-driving car. The laser can accurately measure the relative distance between the outline edge of the target and the self-driving car. These information is used to construct a so-called point cloud map and then draw a 3D environment map accurate to centimeters. In application of self-driving car field, the price of camera is lower, but its reliability is also relatively lower due to the influence of ambient light. LIDAR has the advantages of wide detection range, accurate motion estimation of objects and has therefore high reliability, but the price is higher. Comparing LIDAR and cameras, the latter can accomplish tasks such as detecting lane lines, detecting obstacles, and identifying traffic signs; while LIDAR can solve tasks such as edge detection, dynamic and static object recognition, positioning, and map creation. For a dynamic object, the camera can determine whether the object or pedestrian in the two frames is the same object or pedestrian, and LIDAR obtains this information, and then measures the moving speed and the moving displacement of the object or pedestrian.

3.2 Collision Avoidance

a) Method of path planing : Artificial Potential Field[3]

Artificial potential field method is a commonly used method of local path planning. This method assumes that the self-driving car moves under a virtual potential field. The initial point of the object is on a higher "mountain head" and the target point to reach is at the “mountain foot”, which forms a potential field. The object is guided by this potential, avoiding obstacles and reaching the target point. The artificial potential field includes the repulsive field of gravity, where the target point generates gravity on the object and guides the object toward it. Obstacles repel against object (in this case object is the self-driving car) and prevent it from colliding. The resultant force at each point on the path of the object is equal to the sum of all the repulsions and gravitational forces at this point. An example of artificial potential field is shown in fig4.

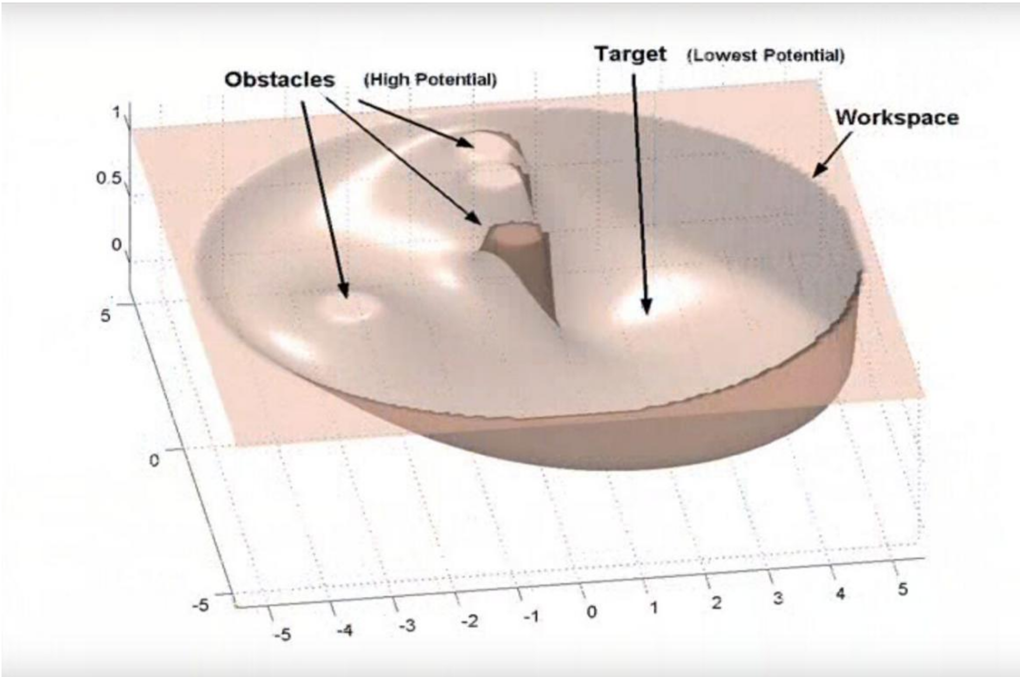

fig4[5].An example of artificial potential field

b) Model Predictive Control[4]

Model predictive control (MPC) is an advanced method of process control that is used to control a process while satisfying a set of constraints. Specifically, in the self-driving field, this theory is used in the following scenarios: After the path is planned, the vehicle parameters (such as acceleration, steering angle, etc.) need to be known to control to make the difference between actual driving path and planning path minimal. The following example shows a simplified vehicle kinematic model :

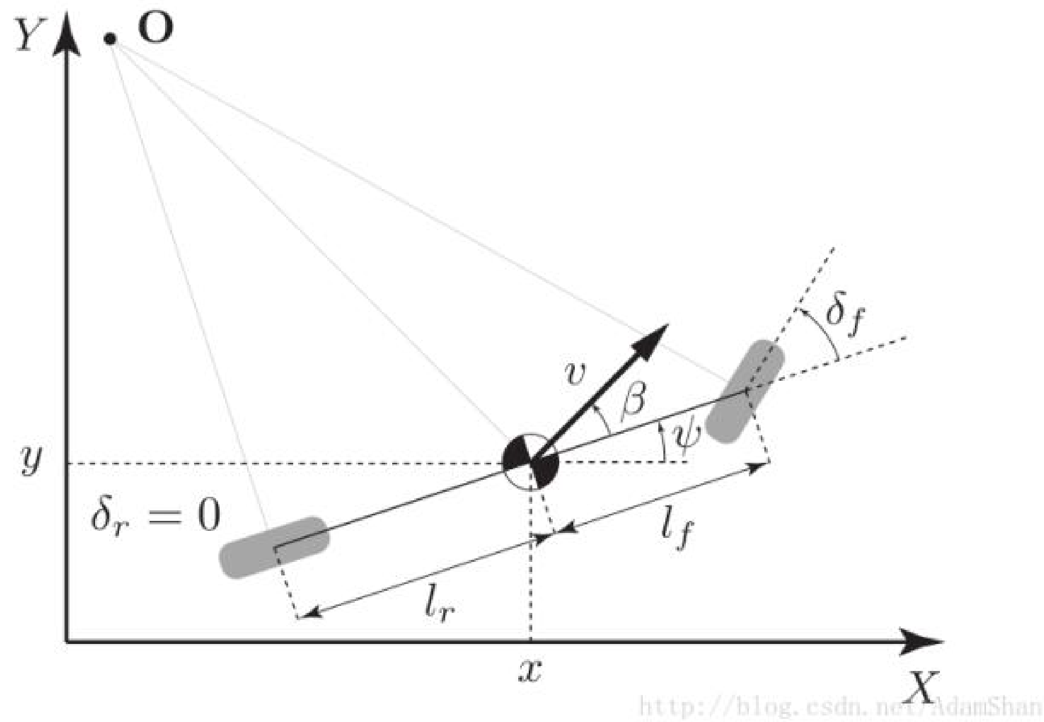

fig5[6]. simplified car kinematics model

In fig5 simplifies the car kinematics to a two-dimensional coordinate system, given the initial state of the vehicle (x, y, ψ, v), and acceleration a (through throttle to control) and steering angle δ (through steering wheel to control), we can predict the status after dt time.

calculated as follows:

$$x_{t+1} = x_t + v_t \cdot \mathbf{cos}(\psi t + \beta) \cdot dt$$

$$y_{t+1} = y_t + v_t \cdot \mathbf{sin}(\psi t + \beta) \cdot dt$$

$$\psi_{t+1} = \psi + \frac{v_t}{l_r} \cdot \mathbf{sin}(\beta) \cdot dt$$

$$v_{t+1} = v_t + a \cdot dt$$

β can be calculated by following: $\beta = \mathbf{tan}^{-1}\left[\frac{l_r}{l_f + l_r} \cdot \mathbf{tan}(\delta_f)\right]$

After obtaining the predicted path, it can be compared with the previously path and measure the gap between them by constructing a loss function as follows:

$$Loss = \sum_{i}^{n} (z_i - z_{ref,i})^2$$

In this loss function, n means model predictions for the next n intervals, and $z - z_{ref}$ is the distance from the prediction point to the actual reference line.

In the process of constructing the loss function, we can also add constraints (for example: the steering angle is always between -90 degrees and 90 degrees). By adjusting the acceleration and steering angle, the loss function is reduced as much as possible, which is the ultimate goal of MPC. By this method a series of control parameters (i.e. acceleration and steering angle) can be obtained, that can automatically deal with environmental noise (for example, due to the difference in slope and roughness of the road surface, the throttle required to obtain the same acceleration is different sometimes) and different constraints.

4 Future work

4.1 Sensor control and fusion

Each type of sensor has its own insurmountable deficiencies. The solution is to synthesize the information collected by different sensors. To achieve perfect sensor fusion, it is necessary to control different sensors to cooperate with each other and use the comprehensive information to more accurately perceive environment, and the results obtained are much better than using different sensors alone. For example, compared with optical sensors, LIDARs are better at detecting distance in bad weather. However, although the optical sensor is limited in bad weather, it can recognize colors (for example, for traffic lights and road signs) and also has an advantage in resolution. The above example proves that different sensors have strong complementary capabilities. In addition, the fusion of different sensors can also provide higher redundancy. Even if a sensor loses effectiveness, it will not affect the safety of the whole vehicle system. Therefore, the control and fusion of sensors is a subject worthy to study further in the future.

4.2 Positioning and navigation

Self-driving cars mainly rely on GPS (global positioning system) to achieve positioning. However, the effect of civil GPS can not meet the requirements of self-driving cars, whose deviation is usually 10 to 100 meters. And frequency of location update of GPS is relatively low , which is usually only 10Hz, when the vehicle is driving fast, it can not provide real-time accurate location information. besides, GPS positioning can be blocked by buildings and trees. For example, in overpasses, tunnels, underground garages, the GPS positioning accuracy is seriously reduced. Therefore, it is still necessary to use other sensors for assisted positioning, besides GPS. After obtaining the exact location of the vehicle, using high-precision maps for path planning is the ultimate goal of positioning and navigation. This is also an essential part of self-driving technology.

5 Summary and Personal Thoughts

As a summary, I would like to present a workflow with respect to the above technologies : Suppose we have a well-designed self-driving car. Firstly it uses the YOLO algorithm to detect the lanes on the road through the visual images collected by the camera, and detects the cars around it through LIDAR, and then proceeds according to the lanes. When encountering obstacles ahead, it uses MPC and artificial potential fields to plan a path to avoid collisions and balances stability and safety when changing directions. The above is the process of how the related technologies involved in the articles' work. Of course, a true self-driving car still has many details that deserve to continue studying. For example, path planning at the crossroads, control of sensors and handling of emergencies, those all require further research.

Reference

[1] Nugraha B T, Su S F. Towards self-driving car using convolutional neural network and road lane detector[C]//Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT), 2017 2nd International Conference on. IEEE, 2017: 65-69.

[2] Lim K L, Drage T, Brä T. Implementation of semantic segmentation for road and lane detection on an autonomous ground vehicle with LIDAR[C]//Multisensor Fusion and Integration for Intelligent Systems (MFI), 2017 IEEE International Conference on. IEEE, 2017: 429-434.

[3] Ji J, Khajepour A, Melek W W, et al. Path planning and tracking for vehicle collision avoidance based on model predictive control with multiconstraints[J]. IEEE Transactions on Vehicular Technology, 2017, 66(2): 952-964.

[4] Funke J, Brown M, Erlien S M, et al. Collision avoidance and stabilization for autonomous vehicles in emergency scenarios[J]. IEEE Transactions on Control Systems Technology, 2017, 25(4): 1204-1216.

[5] Blog of AdamShan : https://blog.csdn.net/adamshan/article/details/79083755

[6] Blog of junshen : https://blog.csdn.net/junshen1314/article/details/50472410

[7] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]. Cvpr, 2015.

[8] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(12): 2481-2495.

.