Fundamentals and Development of Self-Driving Cars

Yoganandhan A, Subhash SD, Hebinson Jothi

Department of Mechatronics Engineering, Chennai Institute of Technology, India

[praviinvj@gmail.com](mailto:praviinvj@gmail.com) , [subhashsd73@gmail.com](mailto:subhashsd73@gmail.com) , [hebinson007@gmail.com](mailto:hebinson007@gmail.com)

**Abstract** – This paper represents detailed information on the constitutional and development of self-driving cars. For a decade, in the automotive industry, there are lots of malfunctions are performed.  It makes many issues like accidents, driver’s liability and much more.  These kinds of problems arise when the human interfaces with the car.    In real-time Self-driving cars that are driven by digital technologies without using a human interruption. Based on the fundamentals of developing self-driving cars are totally by sensing their environment and automating the tasks.  In our proposed system the Localization, Perception, Prediction, Planning, and Control are to make define and governing the car, certain algorithms are used to control the autonomous system and are used for steering functions. The autonomous car can predict and cruise its path and traffic signs as like as pedestrians. It can minimize accidents, fuel rates, and parking space.

**Keywords** – *Driving assistance system, Fundamentals of self-driving cars, Introduction to Self-driving cars, Localization, Perception, Intelligent vehicles*

**Introduction**

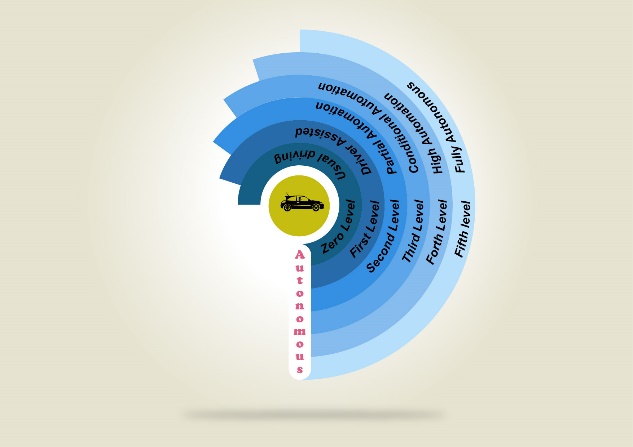
The self-driving car, are also known as Autonomous driving system or driverless cars, a vehicle that is capable of sensing its environment and moving with define lane without human interaction. Self-driving cars technologies mostly involved in the computer system by automating vehicle control parts. These technological parts possess a range of competencies, from forward-collision warning and antilock brakes to lane-keeping and adaptive cruise control, to fully automated driving, Autonomous car combines the variety of sensors, actuators, and cameras. The benefits of automated cars are predicted to increase traffic flow and provide enhanced mobility for all users. The basic fundamentals are in High definition maps, Localization, Perception, Prediction, Planning, and Control of vehicle as follows.

**Origin**

The first era of self-driving cars started in the 1920s. There is a lot of development taken place in creating new technology in the later 1960s. The ALV projects were conducted by the Robotics Institute of Carnegie Mellon University Navlab. By 1994, the double robot vehicles called Vita-2 and VaMP of Daimler-Benz and Ernst Dickman’s demonstrated autonomous driving in free lanes. In 2004 the DARPA (The Defense Advanced Research Projects Agency) conducted the challenge, were self-driving cars complete the course, but no one did that. In 2005 the second challenge was conducted, in that Sebastian Thrun led his team and completed the course. The fully efficient self-driving car was introduced by Toyota Prius modified with Google’s experimental driverless technology was licensed by the Nevada Department of Motor Vehicles in May 2012. The first license issued in the United States for a self-driving car.

**Levels of Automation**

There are five levels of automation systems that enhanced the self-driving system.

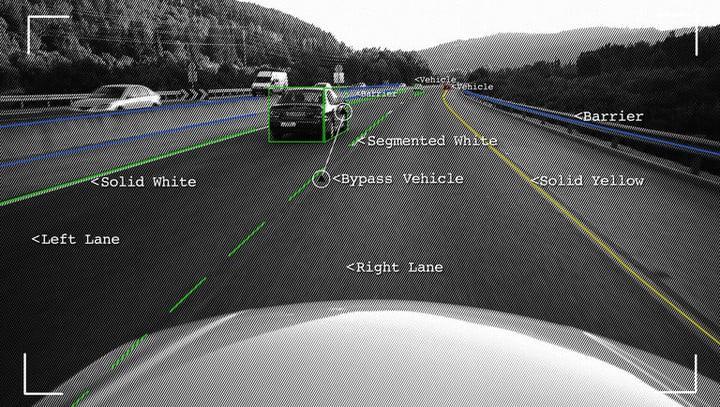
****

*Fig: Five Levels of Automation*

The zero automation is a base level automation system it was a soul decision matter Human driving system. The first level automation is driver assistance, some intelligent aspects were included in it, and the driver was semi-engaged. The second level of automation is partial the automatic cruise control and Automatic lane-keeping system. The third level is conditional automation; Human interface is needed whenever necessary. The fourth level is High-level automation, there is no human interface. The fifth level is a fully autonomous Vehicle that no human and his interface are not needed.

**High-Definition Maps**

The High-Definition maps are not like Normal maps or navigation maps. The HD maps are more important for self-driving cars. They have a higher level of accuracy of objects, lanes, and locations up to 10cm. It contains a huge amount of driver assistance information, Three-dimensional representation of road network, layouts intersection, and location of the signboards. It helps to solve the localization problem, figuring out exactly where the car is around the world. It also recognizes the shape of objects. HD maps are a core of self-driving cars.



The preprocessing and coordinate transformation need to collect data and compare it with HD maps. The uniform coordination system is used in most vehicles. The ***Region of Interest***, the purpose of this section is to build a program that can easily identify the lanes separately in a picture or a video frame from the camera. In that we have to convert this image into grayscale that process a single channel much faster than three channels RGB and less computation intense. Planning with maps makes planners to identify possible routing options. The Maps are also containing information related to the source of data which sensor was used to get the information when the map was last updated.

**Localization**

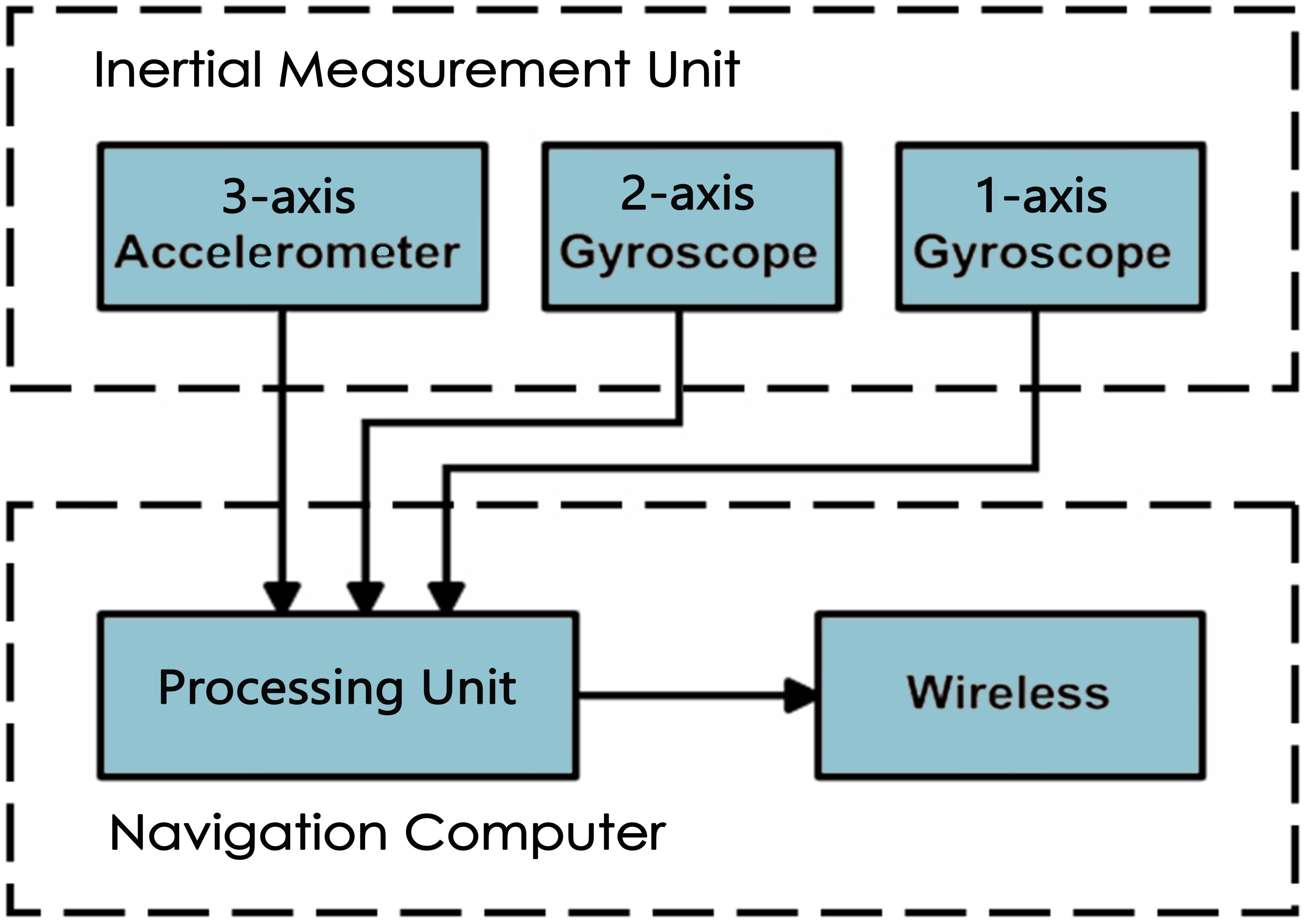
Localization means that shows the exact location of the car. The sensor and the maps are collecting the data to find an exact location. The vehicle coordinate frame and map coordinate frame are vice-versa.

**A) GNSS-RTK**

The GNSS is known as the ***Global Navigation* *Satellite System****,* there are 30 GPS satellites operating in the outer space each Satellite were located on 20,000 kilometers away from the Earth’s surface. The control system is specified around the earth for controlling satellites. The RTK is called a ***Real-Time Kinematic*** positioning system is also a satellite navigation system used to exact position data from a satellite-based positioning system. But the RTK based system was having issues with tall buildings. It was also low-frequency updates like 10MHz. The GPS that equipped in the car can update its location in 0.1 seconds.

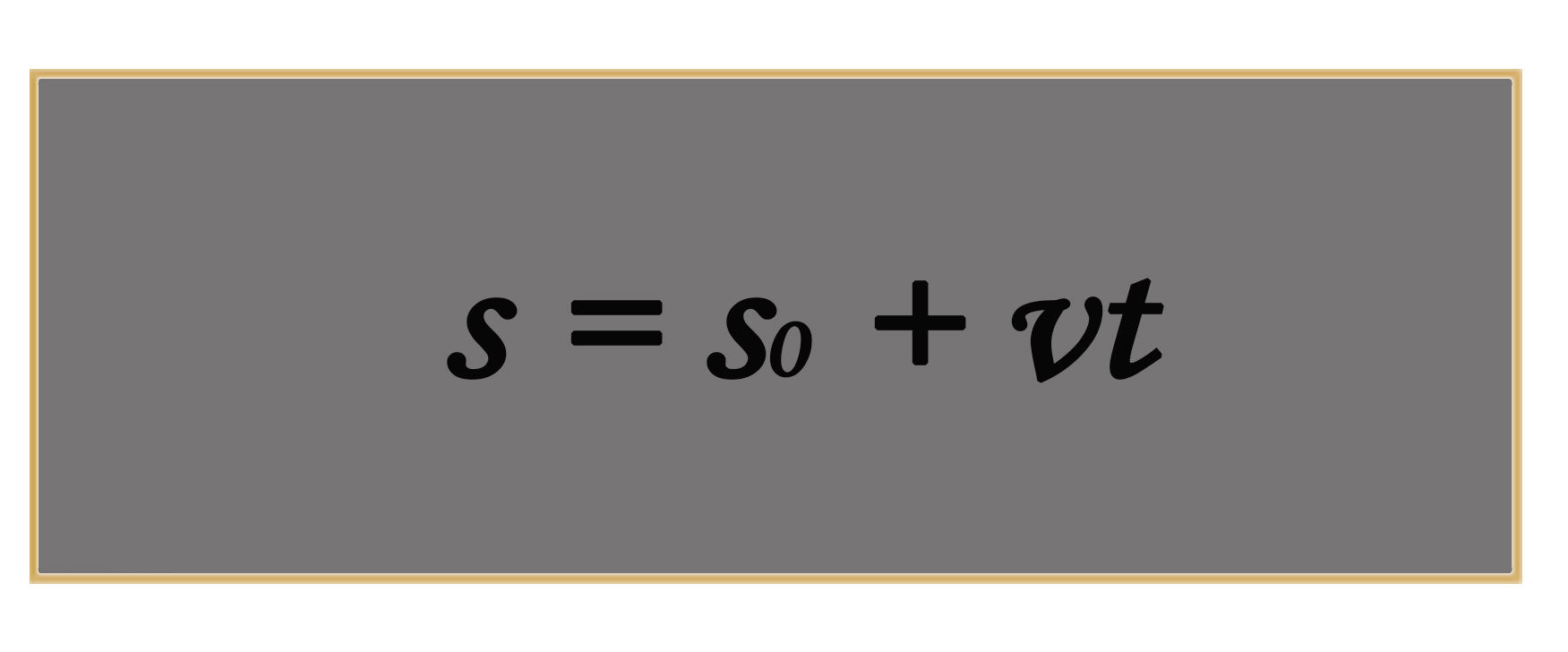
**B) Inertial Navigation**

An Inertial navigation system is a navigational device that uses a computer, motion sensors (Accelerometers) and rotational sensors (Gyroscopes).



*Fig 2.1: Inertial Navigation System*

That can continuously calculate the data and the action or process of calculating the position, the orientation and the velocity (direction and speed of movement of vehicles)



**S0 = Initial Location**

**V = Velocity**

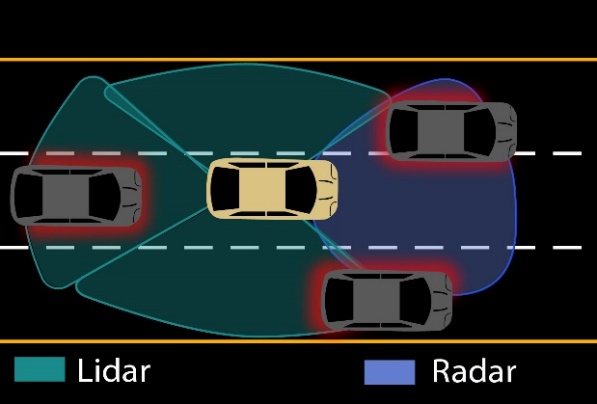
**T = Time Taken**

We are using a 3-axis accelerometer to define the acceleration of the car at any point of the time and also measure the velocity of the current position. Gyroscopes are used to measure the relative position of the spin axis and the three external Gimbals to measure ***Initial Measurement Unit***(IMU).

**C) LIDAR Localization**

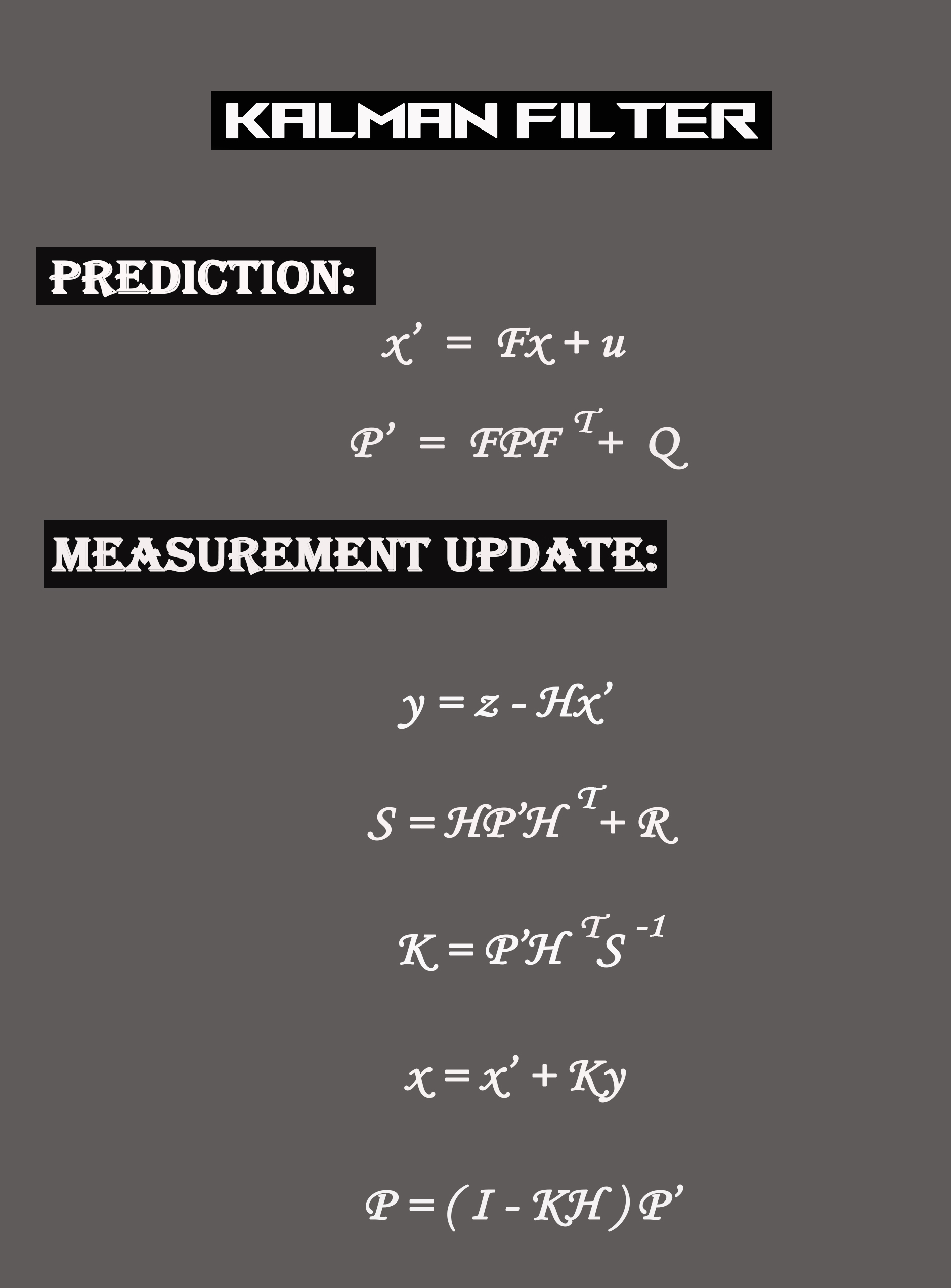
LIDAR (**L**ight **I**ndication **D**etection and **R**anging) means a point cloud matching. This method continuously matches the detected data from the Lidar sensor with HD maps. There are many algorithms are done to match the point of clouds. ***Iterative closet point*** (ICP) is the first approach, filter algorithms are another approach of Lidar localization.

**Kalman Filter** is an algorithm is used to find the assume state which was based on the last state in new sensors measurements.



*Fig 2.2: Lidar and Radar Localization system*

An extended ***Kalman filter*** algorithm is to pose by registering 3D point clouds against ***Gaussian*** mixture multiple solution-maps. This method was proposed on two driverless cars in terrible weather conditions and presented localization estimation errors of about 0.15 m.

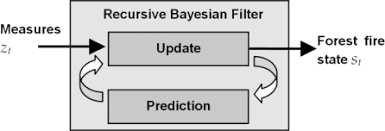


*Fig 2.3: Kalman Filter for 3d Point Clouds*

In a statistical method of the control system, the Kalman filtering is also called as ***Linear Quadratic* *Estimation*** (LQE), this type of algorithm that uses a series of measurements observed over time, and contains other inaccuracies and unknown variables.

**D) Camera-Based Localization**

In camera-based localization self-driving cars are used to estimate the location of the car and relates to the map. A ***Recursive Bayesian*** filter algorithm are used and performs to find inferences in a graph by exploiting its structure and the model of how the car moves, as measured by the visual odometry.



*Fig 2.4: Recursive Bayesian Filer for interferences in graph models*

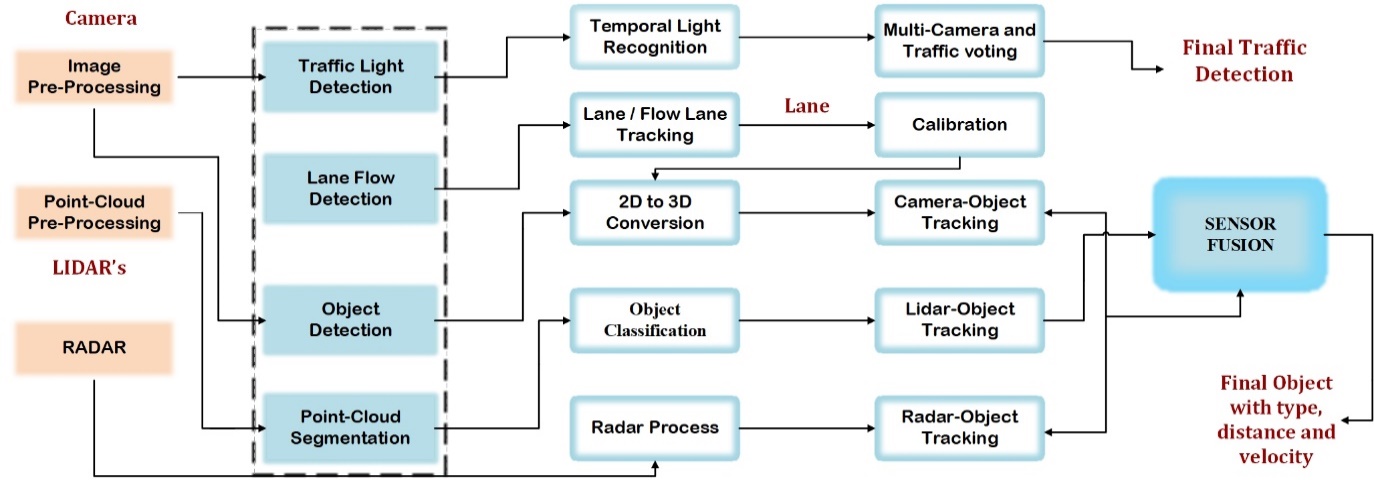
This algorithm is able to point out the car’s position in the graph and increase the probability that the current pose lies in a point graph that is correlated with the latest car movements.

**Perception**

Perception is a tough task in controls of self-driving cars. The perception module has been upgraded completely to handle comprehensive and fusion of sensors.

The perception module incorporated the ability of using Multiple Cameras, Radars, and Lidar to recognize obstacles and fuse to their individual tracks to obtain in final track list from the controller. The obstacle, sub-module detects, classifies and tracks obstacles. The sub-modules are also predicted obstacle motion and position information. For lane-keeping, we had lane instances by post-processing lane parsing pixels and calculate the lane relative location to the vehicle.

The core concepts of self-driving cars are Detection, Classification, Tracking, and Segmentation. ***Detection*** is the means of detecting the object that capture images by cameras or the Lidar inputs. ***Classification*** is a process done by some Neural Network algorithms and classifies in certain manners. The ***Tracking*** is the means of the tracks that the objects from the car like and their velocity, distance, and some other aspects. The ***Segmentation*** in the means of clarifies each pixel form the camera images and semantic category.



*Fig 3.1: Flow Chart of Perception*

**A) Camera Images**

The camera images are the common data; the images are comprised of pixels. Which called small units of color, in every pixel of an image, is just a numerical value, were the values are comprised into an image matrix.

Color Images are more complex. Color images are constructed as Three-Dimensional cubes of values each cube is a Height, Width and the Depth of the value.

**B) LIDAR Images**

The Lidar images are getting from the sensor which creates the point cloud on the environment and defines the objects around it. The Lidar works by the laser coming out of it and getting back with the modified frequency that makes it measure distance.

**C) Machine Learning**

Machine Learning is extremely used to find out the solution to various problems that arise in the manufacturing of self-driving cars. With the inclusion of sensor data processing in an ***Electronic Control Unit***(ECU) in a car, it’s essential to enhance the utilization of machine learning to accomplish new tasks.

The **Supervised Algorithms** make a training dataset to learn and they continue to learn till they get to the level of confidence and they aspire to reduce the probability error. Supervised learning is also sub-categorized into regression, classification, and detection or dimension reduction.

**Unsupervised** **Algorithms** are another set of machine algorithms that fall between unsupervised and supervised. There is a target label in supervised learning; there are no labels in unsupervised learning, the **Reinforcement** **learning** consists of time-delayed and sparse labels for future rewards.

**Regression** is also a kind of algorithm for predicting functions. The Regression Analysis evaluates the relationship between two or more

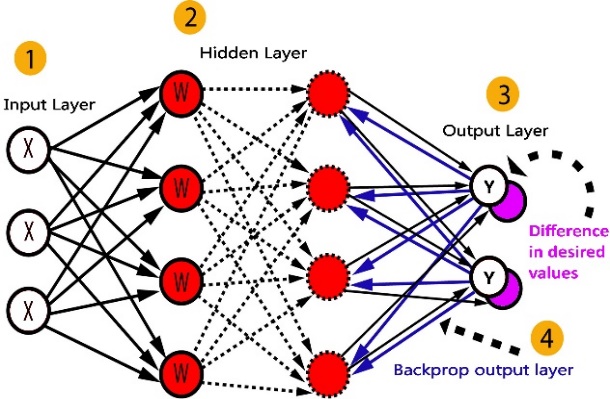
variables to collate the effects of variables on distinct scales and is driven mostly by the metrics.

**Neural Networks**

An Artificial Neural Network is a tool to learn complex patterns of data. Neural Network is comprised of a large number of neurons. For neural networks the most basic representation of the image is “The Pixel value of the Image”.

**A**) **Back Propagation**

The learning is also called Training. It was consisting of the step cycle.



*Fig 3.2: Back Propagation for Error Reduction*

1. **Feed Forward** – Feed each image to Neural Network (n, n) to generate the output value.
2. **Error Management** – The difference between Ground Truth and generated output value.
3. **Back Propagation** – We sent the error to back position through the Neural Network feed to forward on the reverse.

**B) Convolutional Neural Network**

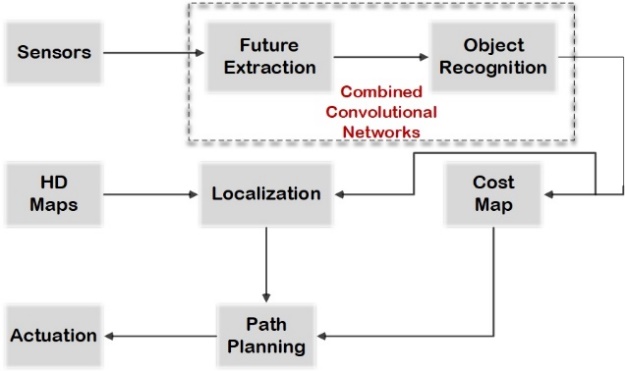
The Convolutional Neural Network is a perfect solution to the Perception problem. The input values for CNN are multi-dimensional values, including two, and three-dimensional shapes that define most of the sensor data.

**C)** **Region-Based Convolutional Network**

The Region-based Convolutional Network (RNN) gets the excellent object detection accuracy by the deep convolutional network to classify the object’s proposals. R-CNN has notable drawbacks.

**1.** **Training is a multi-stage pipeline**

R-CNN is the work to fine-tune the ConvNet on object proposals using log loss. Then, it fits ***Support Vector* *Machines*** (SVM) to ConvNet features. These SVMs are acting as object detectors, replacing the SoftMax classifier learned by fine-tuning. In the third training stage, bounding-box regressors are learned.

****

*Fig 3.3: Combined convolutional Networks*

**2. Training is expensive in space and time**

For SVM the bounding box are the regression training, the feature is extracted from each object proposal in each of the images and written to the output. The very deep neural network, such as the VGG16, in this process it takes 2.5 GPU-days for the 5k images of the VOC07 travel set. These features require hundreds of gigabytes of memory and storage.

**3. Object detection is slow**

At last, test-time that features are extracted from each object proposal in each test image, Detection with VGG16 takes 47s / image on a GPU.

**D) Tracking**

After detecting the object, it is continuously tracked. Detection of every object, frame and identification of each of the object is done with the ***Boundary Box***. If the identity gets the conformation it will match all the objects detected in the previous frame. That object detects in the frame by finding objects with a higher similarity.

**E) Segmentation**

The semantic segmentation involves the classifying of each pixel of the image. ***Fully* *Convolutional Neural Network*** (FNN), in that FNN is replaces the flat layers at the end of a traditional CNN architecture with convolutional layers. The first part of the network is called encoders and fetches on the input image. The second half is a decoder it applies to output.

**F)** **Region of Interest**

The region of interest is based on the object detected by the read-data input from the point-cloud data for certain applications.

**1.** **Single Shot Detector**

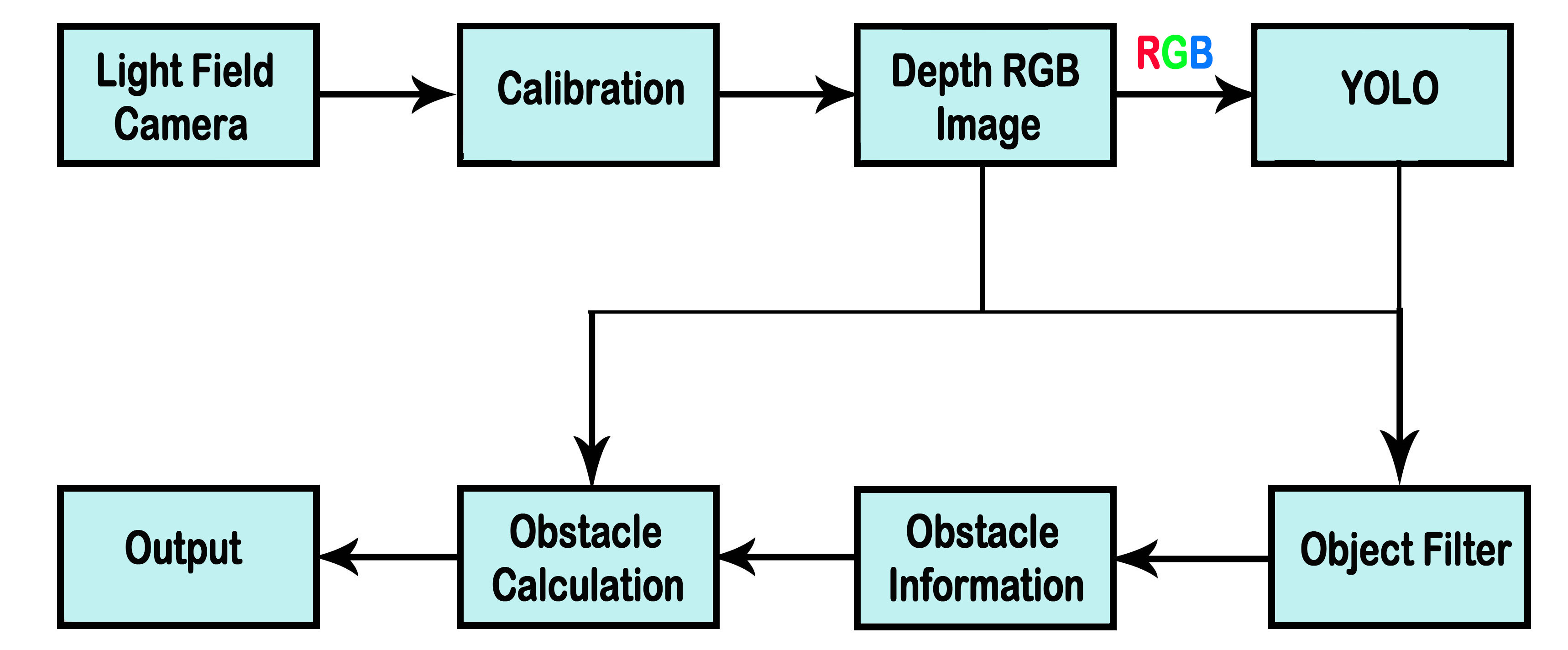
The SSD approach is based on the feed-forward convolutional network that produces a fixed-size collection of ***bounding boxes*** and scores for the presence of the object with the class instances in those boxes, it followed by a non-maximum suppression step to produce the final detections. The Early network layers are based on a standard architecture used for the high-quality image classification technology, which we will call as the base network. We then add auxiliary structure to the network to produce detections with the following structures.

**Multi-scale feature maps for detection,** in that we add convolutional feature layers to the end of the truncated base network.

**Convolutional predictors for detection**, in this approach each added feature layers can produce a fixed set of detection predictions using a set of convolutional filters.

**2. You Only Look Once (YOLO)**

In this YOLO, it is a new type of approach detection technology in the self-driving cars. YOLO predicts multiple bounding boxes per grid cell. At training time, we only want one bound in box predictor to be responsible for each product. First, YOLO is extremely fast in object detection technology. Since we frame detection as a regression problem, we don’t need a complex pipeline. We simply run our neural network on an image at a test time to predict detections. Our base network runs at the 45 frames per second with no batch processing on the Titan X GPU a fast version runs more the 150 fps.



*Fig 3.4: YOLO Object Detection from Images*

**Prediction**

The prediction module studies and predicts the behavior of all the obstacles detected by the perception module. Perception receives obstacle data along with basic perception information including positions, headings, velocities, accelerations, and generates predicted trajectories with probabilities for those obstacles.

Prediction needs to be real-time, latency as small as possible accurate; Predictions is also been valued on learning a new behavior of vehicles.

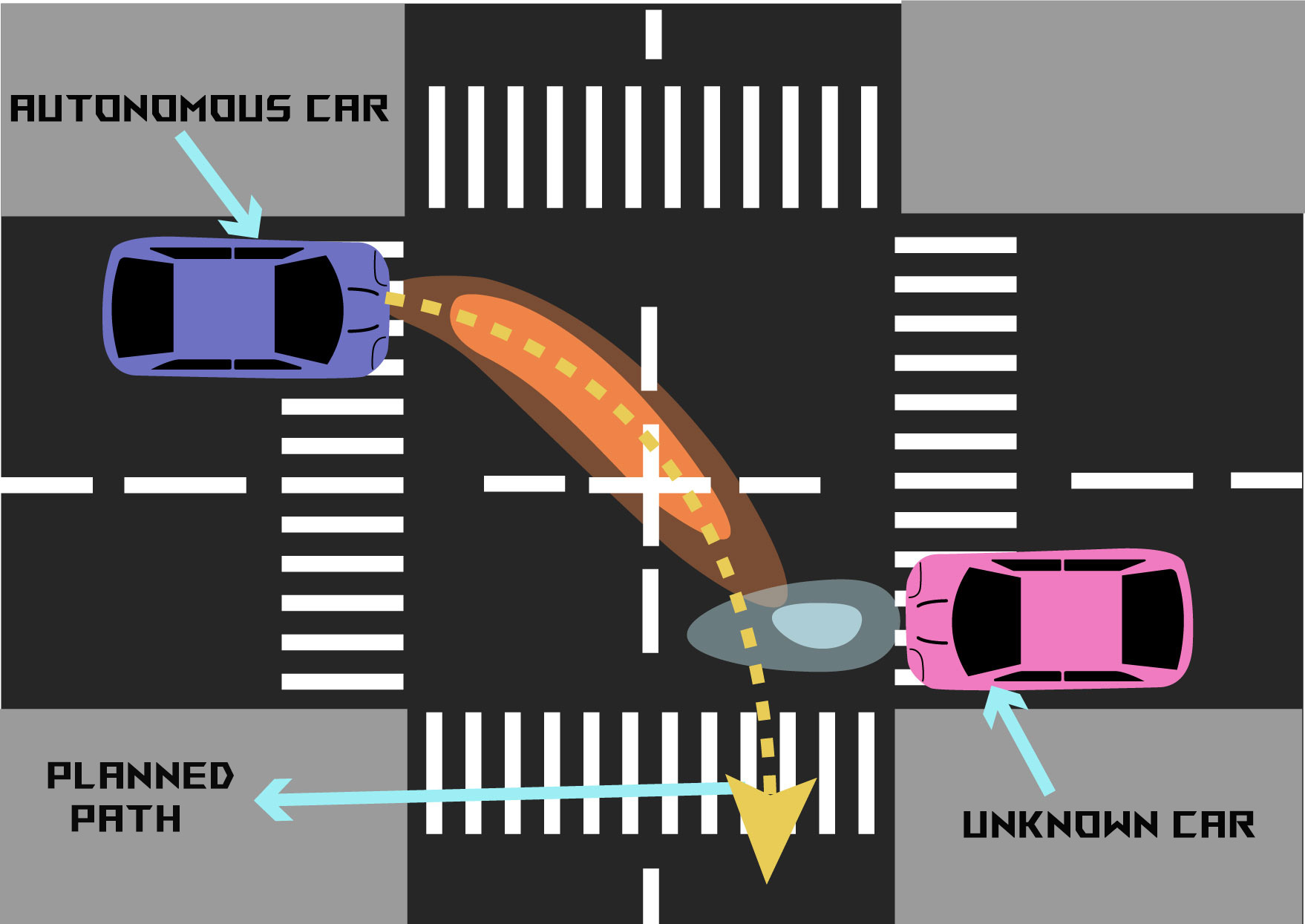
1. **Model-based Prediction**
2. **Driven-based Prediction**

**Model-Based Prediction**, one model describes the moment of a vehicles turning positions. Another model describes the movement of the vehicle whether continuing straight or not.

**Data-Driven Prediction**, it was used by machine learning to train a model based on the observations once the model is trained and able to make predictions in the real world.

**A) Lane Sequence-Based Predictions**

In the lane sequence-based predictions we have to divide the path into multiple segments.



*Fig 4.1: Path Planning Predictions*

Autonomous vehicles are equipped with many advanced sensors that allow them to perceive other vehicles, obstacles, and pedestrians in the environment. If any obstacle status occurs, we have knowledge on predict the state, initially we have to know the state of an object.

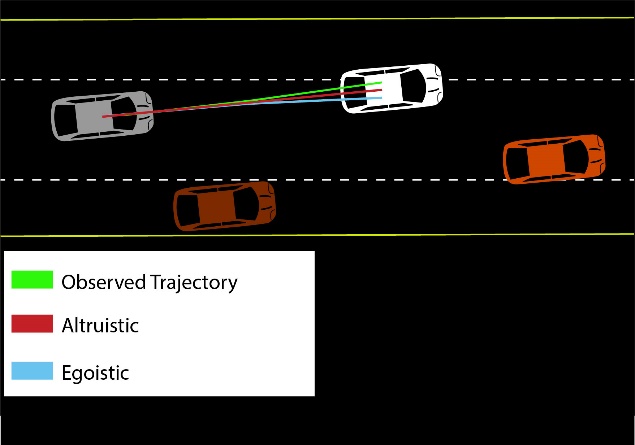
The classical approach of probabilistic graphical models, such as ***factors graph* *spatiotemporal graphs***, and the dynamic Bayesian networks, which bring graphical models into the sequential modeling space, is widely used in self-driving cars community for many reasons including their interpretability and the high-level structure, which can capture various relationships between features to modeling temporal sequences.

**B) Recurrent Neural Networks**

An approach that takes special advantage of time-series data (Back Propagation) apart from its standalone utility is recurrent neural network. Input monolithic and relativity simple SSD model provides a useful building block for larger systems that employ an object detection component. A promising future direction is to explore its use as part of a system using Recurrent Neural Networks to detect and track objects in Video simultaneously.

**C) Trajectory Generation**

Trajectory planning was a final step of the prediction process. We can be getting constrains that will eliminate most of the candidate trajectories. We make an assumption that the car will align with the center to the target lane.

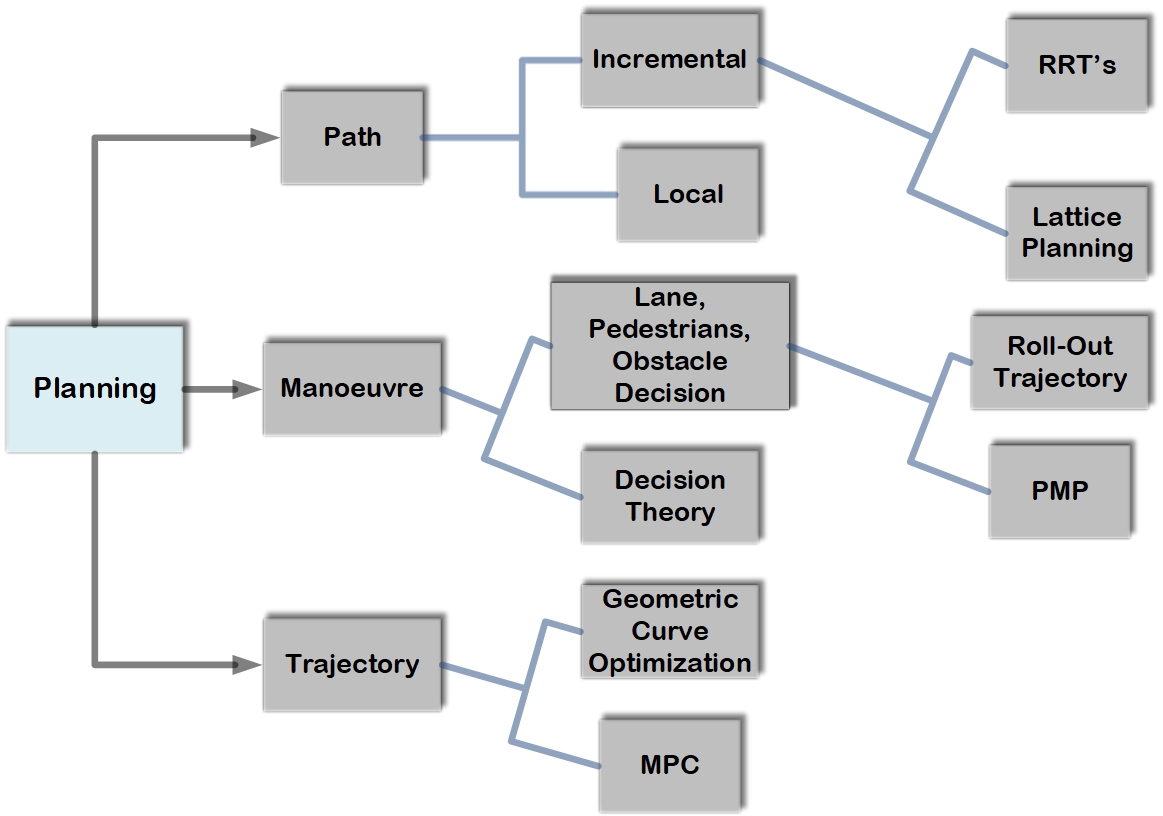


*Fig 4.2: Trajectory Planning to attain correct path*

In that above figure, Path Planning for autonomous vehicles becomes possible after technology considers the urban environment in a way that enables it to search for a path. Put, simply thread-life physical environment is transformed into a digital configuration or a state space. Path planning technology searches for and detects the space and corridors in which a vehicle can drive.

**Planning**

Planning is a base of Routing. The routing takes the map data as input and output a navigable path.

****

*Fig 5.1: Flow Chart of Planning*

**A) Routing**

Routing was planning to go from starting point to destination. It needs three inputs

* **Map**
* **Current position on map**
* **Destination**

**Route Module,** Trajectory planning how we make accurate decisions to avoid obstacles and create a smooth ride for the passengers.

**B) Graph Analysis**

The graph is not the state-space graph, in fact unlike the state-space graph in which a plan is a path through the graph. The planning graph is essentially a flow in the network flow sense. Planning Graph is closer in spirit to the ***Problem Space Graphs*** (PSG).

**Nodes** – Section of Road

**Edges** – Connection between on those sections

**Constraints,** In the real world it was plenty of constraints, which was a major use of trajectory to a collision-free, obstacle-free passengers make to feel more comfortable.

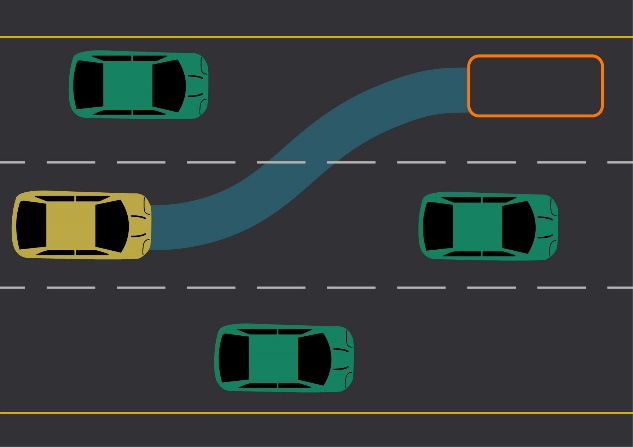
**Frenet coordinates,** it helped us to describe the position of cars with respect to the road.

**Trajectory Planning,** it was the most crucial moment of planning of the car the ***Path-Velocity*** decoupled planning.

* **Path-Planning**
* **Speed-Planning**

**C) Path Generation and Selection**

The path Generation and Selection is the next process after it defines all constraints; it was based on the position of the car.



*Fig 5.2: Path Generation to avoid Collison*

**D)** **Lattice Planning**

The trajectory was an implement in 3d representation longitudinal dimension, lateral dimension, time dimension. There are two kinds,

* **SL - Trajectory**
* **ST - Trajectory**

**Controls**

Control is the main strategy of actuating the vehicle to move it towards the road. The control inputs are Steering, Acceleration, and Brake. It is especially for safety as planning and control the smoothness of driving, it is the main option to control.

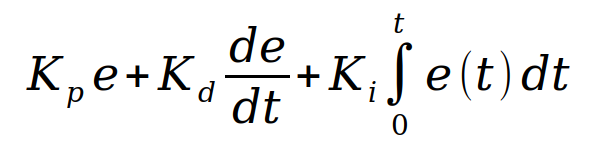
1. **PID – Proportional Integral Derivative**
2. **LQR – Linear Quadratic Regulator**
3. **MPC – Model Predictive Control**

**Control Pipeline**

Two input aspects are ***Target*** ***Trajectory*** and ***Vehicle* *state***. The Target Trajectory comes from the planning module. Each point of trajectory as designates position (x, y), velocity (***v****)* and the acceleration of the car (***a***). The vehicle state determines the position of the vehicle by using the localization module. This gets data from the sensor in the steering, acceleration, and brake.

**PID Control**

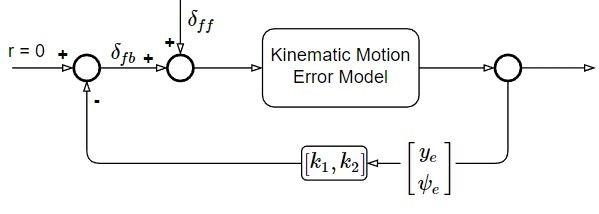
The most important characteristic of autonomous vehicles is their safety and their ability to adapt to various situations and road conditions. We are comparatively comprised of three implementations of such controlling methods, a proportional-derivative (PD) controller built with in accordance with the sensors in steering, a PID controller as an extension of the steering control, a controller designed via the most versatile evolutionary computing methods.



In the above equation that defines the PID controller, where **Kp, Ki, Kd**refers to ***proportional***, ***Integral*** and ***Derivative*** gains constants respectively. For implementation in discrete form, the controller equation is modified by using the backward Euler method for numerical integration. The term which represents the sampling time **ts** is simply eliminated because of multiplication with constants values the PID gains **Ki**and **Kd**

**Linear Quadractive Regulator**

In Linear Quadractive Regulator method the constant and time-varying Vehicle Speeds. The latter is implemented by using a simple gain scheduling method at the grid of the operating points.



*Fig 6.1: Linear Quadratic Control System*

In the design of stabilizing LQR state space control coefficients (***k1, K2***) for the given linear system and obtain the time-varying controllers are different vehicle speed as (***k1*(*V*(t)), *k2*(*V*(t))**) that can be obtained through the Matlab.

**Model Predictive Control**

The MPC is an advanced method of process control that is used to control a process while satisfying a set of constraints. It can obtain further speedup by solving the planning problem approximately it can also fix barrier parameters and limit the total number of Net won steps. It can also run on **Kilohertz** rates.

**Conclusion**

This paper presents detail information about the fundamentals and development and major aspects of self-driving cars.  The framework formulates the problems that arrive from the automotive industries that makes major losses. The errors from the industry are non-predictable.  Multiple solutions from the different scenarios make that problem to clear. One of the main aspects is the era of self-driving cars is to make and the successful experimental results show utility to this approach.  In recent days many companies involved in research and manufacturing of the self-driving cars in efficient methods to solve various problems on certain aspects. In our work, we detailed the real-time self-driving cars and also about every circumstance clearly and also approach it in the right way.  In future, the different technologies will be implemented on further problems to reduce difficulties.

# References

[**1**] Aswanth Suresh, D. G. (2019). "Design and Development of Intelligent Self-driving cars using ROS and machine vision system". *Springer* .

[**2**] Girshick, R. (2019). "Fast R-Cnn" . *IEEE International Conference on Computer Vision .*

[**3**] Girshick, R., Donahue, J., Darell, T., & Malik, J. (2014). "Rich Feauture hierarchues for accurate object detection using RNN-based Deep Models". *IEEE International Conference on Computer Vision.*

[**4**] He, K., Girshick, R., & sun, J. (20156). "Faster R-CNN: Towards Real-Time object detection and Region Proposal Networks". *IEEE International Conference on Computer Vision.*

[**5**] Kumar, A. K., Scheufule, B., & Beucer, D. (n.d.). "Indoor Localization of Vehicles using Deep Learning".

[**6**] Lu, W., Erhas, D. A., & szesedy, C. "SSD: Single Shot Multibox Detector".

[**7**] Mahale, T., Chen, C., & Zhang, W. (2018). "End to End Video Segmentation for Driving: Lane Detection for Autonomous Car".

[**8**] Molla, T. (2018). "Self Driving Car". *Research Gate* .

[**9**] Pannu, G. S., Mohammed, & Gupth, P. (2015). "Design and Implementation of Autonomous Car using Raspberry Pi". *International Journal of Computer Applications* .

[**10**] Patel, s., Griffin, B., & Kusano, K. (2019). "Predicting Future Lane changes of Highway vehicles using RNN-based Deep Models".

[**11**] R. G., Carneio, R. V., Azevedo, R., & Cardoso, V. B. (2019). "Self-Driving car: A Survey". *IEEE Explore* .

[**12**] Salman, Y. D., Ku-Mahamud, K. R., & Kamiouku, E. (2017). "Distance Measurement for

Self-Driving cars using stereo camera". *Proceedings of the 6th International Conference.*

[**13**] Sha, A., Roy, D. D., Alam, T., & Deb, K. (2012). "Automated Road Lane Detection for Intelligent Vehicles" . *Global Journals Inc.*

[**14**] Wang, C.-C., Huang, S.-S., & Fu, L.-C. (2006). "Driver Assistance System for Lane Detections and Vehicle Recognisition with Night Vision". *IEEE Explore.*

[**15**] Lassa, Todd (2013). "The Beginning of the End of Driving". *Motor Trend*

[**16**] Umar Zakir Abdul, Hamid; et al. (2016). "Current Collision Mitigation Technologies for Advanced Driver Assistance Systems - A Survey". *PERINTIS eJournal*.

[**17**] Gibson, David K. (2016). "Can we banish the phantom traffic jam?" *BBC*

[**18**] "Driver licensing system for older drivers in New South Wales", Australia. *NSW Government* (2016).

[**19**] Miller, John (2014). "Self-Driving Car Technologies Benefit, Potentials Risks and solutions". *theenergycollective.com*

[**20**] Whitwam, Ryan (2014). "How Google's self-driving cars detect and avoid obstacles". *ExtremeTech.*

[**21**]Henn, Steve (2015)."Remembering Driverless Elevators Drew Skepticism".

[**22**] Adhikari, Richard (2016). " Feds Puts AI in the Driver's Seat". *Technewsworld*

[**23**] E.h. Tseng et al. (2019). "The development of vehicle stability control at Ford". IEEE/ ASME Transactions on Mechatronics 4, no.3, pp.223-234

[**24**] Rajesh Rajamani (2005). "Vehicle Dynamics and control", 1st ed. *Springer*