Google Self-Driving Car

Obstacle Detection and Avoidance Through Image Processing-Review of Autonomous Vehicles Sensing and Classification

Overview

Autonomous Vehicles including autonomous emergency braking (AEB)

Sensors/vision systems data acquisition

- Cameras monocular, or binocular/stereopsis (stereo based/3D)
- Lidar
- Radar

Algorithms data representation

- Segmentation
- Representation and Description
- Object Recognition

Autonomous Vehicles

- Compared to existing car driver:
  - React faster
  - 360º perception
  - DO NOT get distracted, sleepy, or intoxicated
  - 37,423 people died in car accidents in the US in 2008
- Google Self-Driving car
  - 2005- Sebastian Thrun/Stanford wins 2nd Grand Challenge DARPA competition, $2M prize 132 miles in desert.
  - 2007- Cars drive safely in urban environments.
  - 2008- Google car concept No steering wheel or pedals.
  - 2014- Prototype complete Driven 1,000 miles with no intervention and 140,000 miles with occasional intervention. Team of 15 engineers.
  - 2017-2020- Mass production. Project director Chris Urmson.

Google Self-Driving Car

Cost

- $150,000 in equipment including a $70,000 lidar.
- Lidar is made in small quantities by hand. Newer lidars are now appearing at sub-$1,000 price points.

Sensors

- Lidar (scanning/flash laser) Velodyne 64-beam laser range finder that is mounted on the top and creates a 3D map of the environment to estimate the car's location and compare it to GPS. Spins 10 times/s
- Radar sensors 3 sensors in the front and 1 sensor in the back
- Position sensor on the left rear wheel to measure movements & position on the map accurately

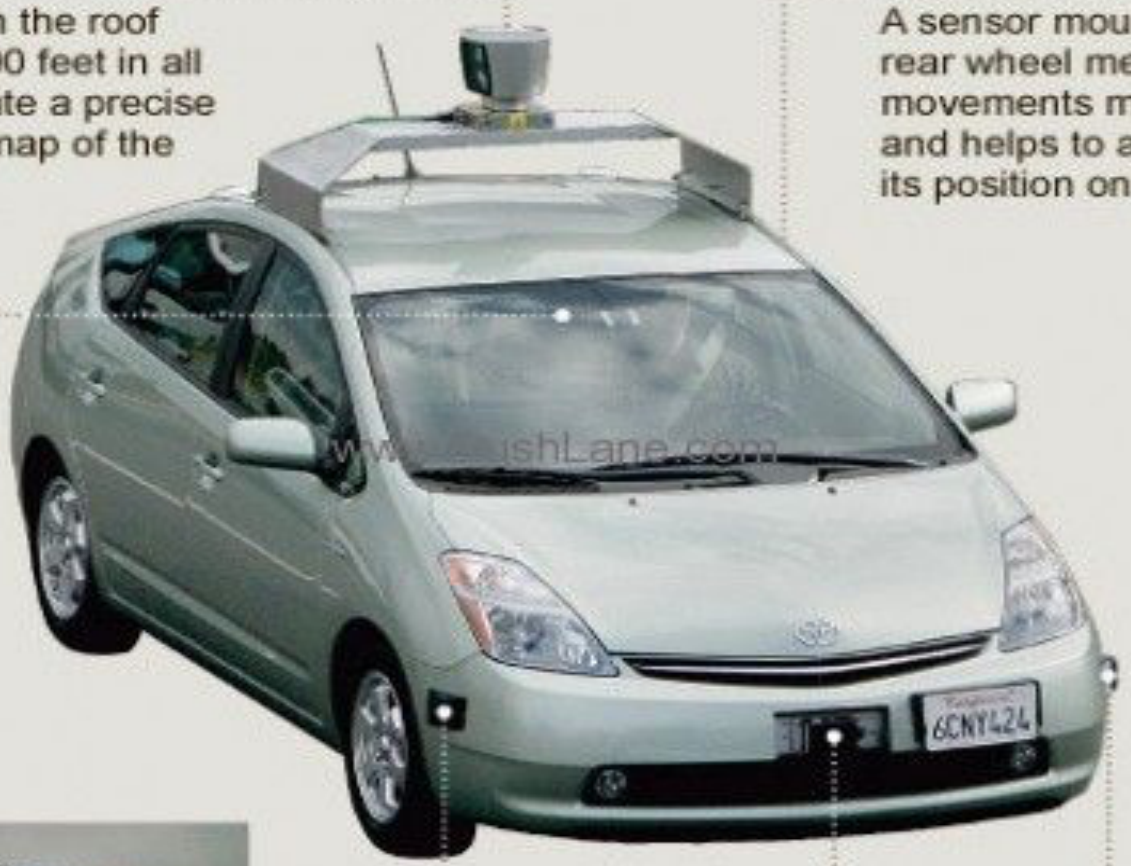
Autonomous Driving

Google's modified Toyota Prius uses an array of sensors to navigate public roads without a human driver. Other components, not shown, include a GPS receiver and an inertial motion sensor.

LIDAR
A rotating sensor on the roof scans more than 200 feet in all directions to generate a precise three-dimensional map of the car's surroundings.

POSITION ESTIMATOR
A sensor mounted on the left rear wheel measures small movements made by the car and helps to accurately locate its position on the map.

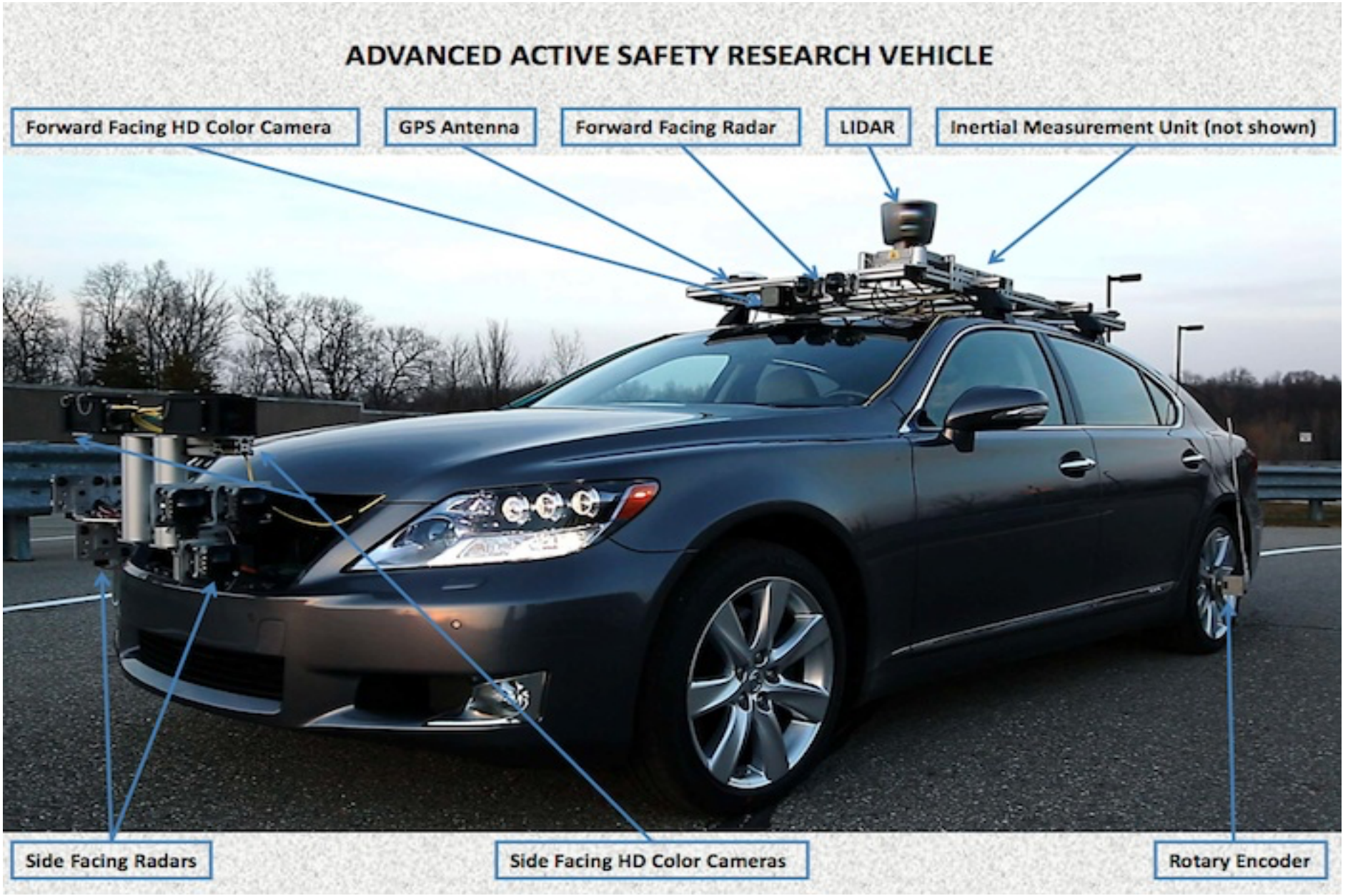
VIDEO CAMERA
A camera mounted near the rear-view mirror detects traffic lights and helps the car's onboard computers recognize moving obstacles like pedestrians and bicyclists.



RADAR
Four standard automotive radar sensors, three in front and one in the rear, help determine the positions of distant objects.

Source: Google

THE NEW YORK TIMES; PHOTOGRAPHS BY RAMIN RAHIMIAN FOR THE NEW YORK TIMES


ADVANCED ACTIVE SAFETY RESEARCH VEHICLE
Forward Facing HD Color Camera
GPS Antenna
Forward Facing Radar
LIDAR
Inertial Measurement Unit (not shown)
Side Facing Radars
Side Facing HD Color Cameras
Rotary Encoder

Tesla Model S Autopilot

Tech package

- Hardware added to all new cars in Oct. 2014 for an added cost of $4,250; cannot retrofit older Tesla vehicles.
- Radar in the front
- 12 long range UV sensors
- Forward camera
- Software update that will engage auto-steering or “autopilot”. However, driver should still pay attention.

Features

- Changes lanes when turn signal is turned on
- Warns the driver (option to automatically slow down) when over speed limit by reading speed limit signs

Autonomous Emergency Braking

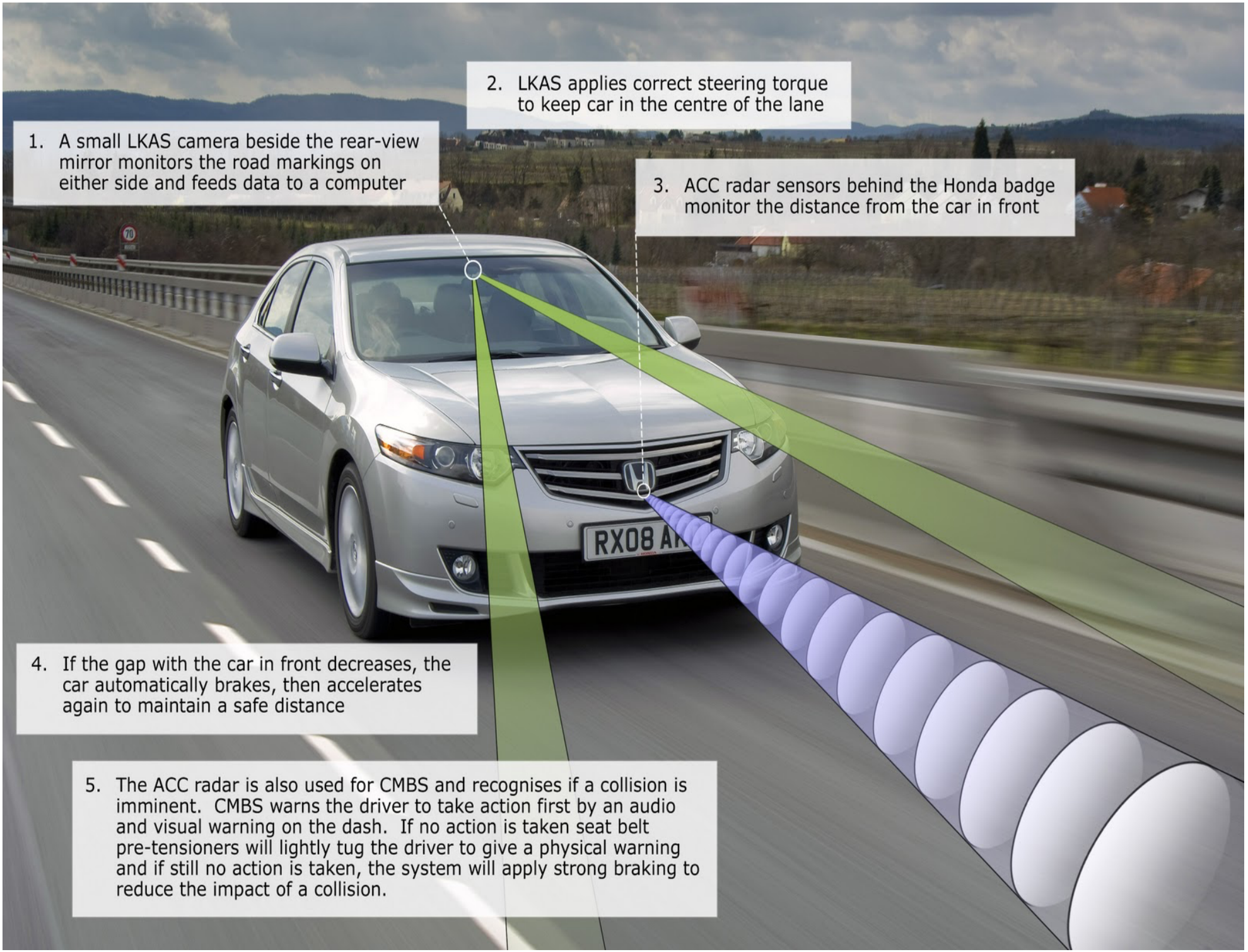
a.k.a Emergency Braking System

- Identifies potentially hazardous targets and applies brakes sooner and harder.

- 45% shorter stopping distance

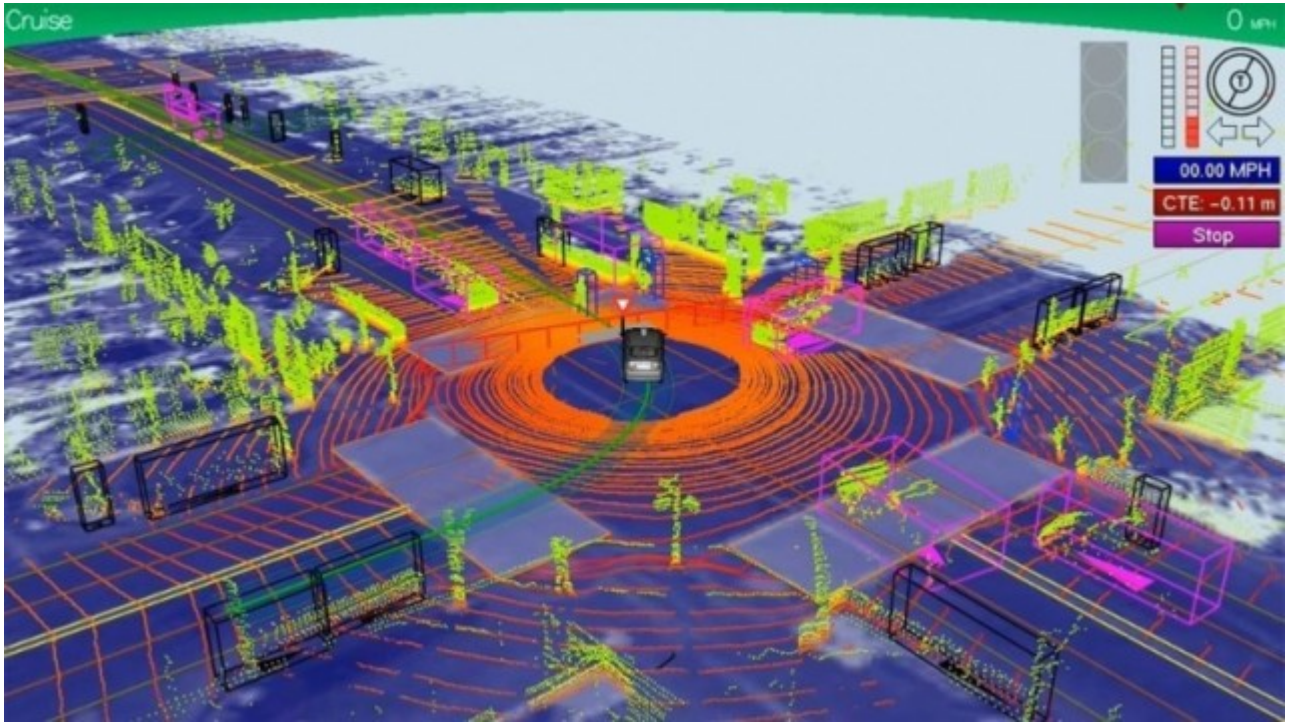
- Estimates say AEB could save 1,100 lives/day
  - 400,000 people under the age of 25 die/yr or >1,000/day.
  - 1.3M worldwide deaths/yr or 3,287 deaths/day.

- Sensors: Radar, Lidar, Ultrasonic, Infrared, or Optical sensors
  - Camera to aid in the classification of an object and radar to detect the relative location of it
- History
- 1996- AEB in Mercedes-Benz S-Class & SL-Class models
- 1998- AEB standard on all Mercedes vehicles
- Acura, Audi, BMW, Infiniti, Land Rover, Rolls Royce, and Volvo have offered their own versions


2. LKAS applies correct steering torque to keep car in the centre of the lane
1. A small LKAS camera beside the rear-view mirror monitors the road markings on either side and feeds data to a computer
3. ACC radar sensors behind the Honda badge monitor the distance from the car in front
RX08 AF
4. If the gap with the car in front decreases, the car automatically brakes, then accelerates again to maintain a safe distance
5. The ACC radar is also used for CMBS and recognises if a collision is imminent. CMBS warns the driver to take action first by an audio and visual warning on the dash. If no action is taken seat belt pre-tensioners will lightly tug the driver to give a physical warning and if still no action is taken, the system will apply strong braking to reduce the impact of a collision.

Sensors

- The localization process starts by using tools like GPS, inertial motion detection and wheel encoders
  - Localization is performed by the sensors to figure out where the car is on the road very accurately
  - Necessary since GPS is not accurate enough to safely drive the car



Cameras vs Lidar

Vision system

- Monocular vs binocular- 3D motion parallax/stereopsis (2 cameras)
- Trinocular (3 cameras) simplifies matching, but 5-10 times too slow
- Light focused onto image chip (i.e. CMOS or CCD) converts light into a digital signal.
- High resolution, reliable, low cost, and low power consumption vs lidar

Lidar/Ladar

- Light based radar. Grayscale images
- Good against interference and high resolution vs radar
- Greater range and less image distortion issues in inclement weather vs camera
- UV, visible, or near infrared light where a narrow laser-beam can map physical features in the environment with 30 cm/px resolution and up to 75m away

Downsides to cameras

- Computer vision is not anywhere nearly good enough today to detect all important features with the reliability necessary for safe driving
- Cameras must deal with lighting variation.
  - Objects are routinely subject to moving shadows and can be lit from any direction, or not lit at all
- Cameras require illumination at night where headlights may not be enough
- Computer vision takes a great deal of CPU or custom chips

Multiple Sensors

Sensor fusion - combining data from multiple sensors improves the quality of the data but is not always reliable.

- Believe every obstacle report to reduce blindness but you potentially add ghost obstacles. Less confusion if each sensor is better at a particular problem and you trust it over that particular regime of examination

Other sensors

- Radar – can see through fog, doppler radar can determine velocity of other cars; radar can bounce off of the road and cars ahead.
- Thermal or long-wave infrared- uses emitted instead of reflected light
- "Hyperspectral" imaging- cameras working in many color bands
- Alternative types
  - Infrared, Ultraviolet, Sonar (sound)

Machine Learning

- Construction and study of algorithms that learn and make predictions off of data
- Ability to adapt to new circumstances and to detect and extrapolate patterns
- Given experience, the performance on tasks improve.
- Predictions are based on known data as opposed to knowledge discovery & data mining (KDD) which is the discovery of unknown properties in the data.
- Makes generalizations based on limited data set performance evaluated on the ability to reproduce known knowledge
- Pattern recognition and computational learning theory shares similar properties with neurons in the inferior temporal cortex that are used for object recognition in primate vision
- Overlaps with computational statistics

Camera Processing

- Image or video sequence
- Machine vision is a simpler, localized analysis vs computer vision which recognizes objects
- Machine vision allows some detection of other cars, pedestrians, edges of the road and road markings
- Machine vision can detect objects independent of scale or rotation

Segmentation and Classification

- Segmentation partitions an image into multiple segments based on two basic properties of intensities: **discontinuity** & **similarity** according to a set of predefined criteria
  - Goal is to simplify or change the representation of an image to something more meaningful and easier to analyze
- Classification locates objects or boundaries such as point, line, & edge detection using derivatives
  - Digital signals/functions instead of derivatives, the discrete nature uses difference methods
- Discontinuity is detected using 1^{st} & 2^{nd} order derivatives using spatial filters or highpass filter

Segmentation

- Methods
  - Thresholding
  - Region growing
  - Region splitting
  - Morphology and motion cues
- Using more than one method is preferred i.e. combine edge detection (discontinuity) with thresholding
- Low level cues/attributes i.e. brightness, color, texture, motion
- Partitioning is inherently hierarchical mainly low-, mid-, and high-level cues must exist
  - Low-level cues must be combined with mid- and high-level cues to either confirm the low-level group or select one for further attention

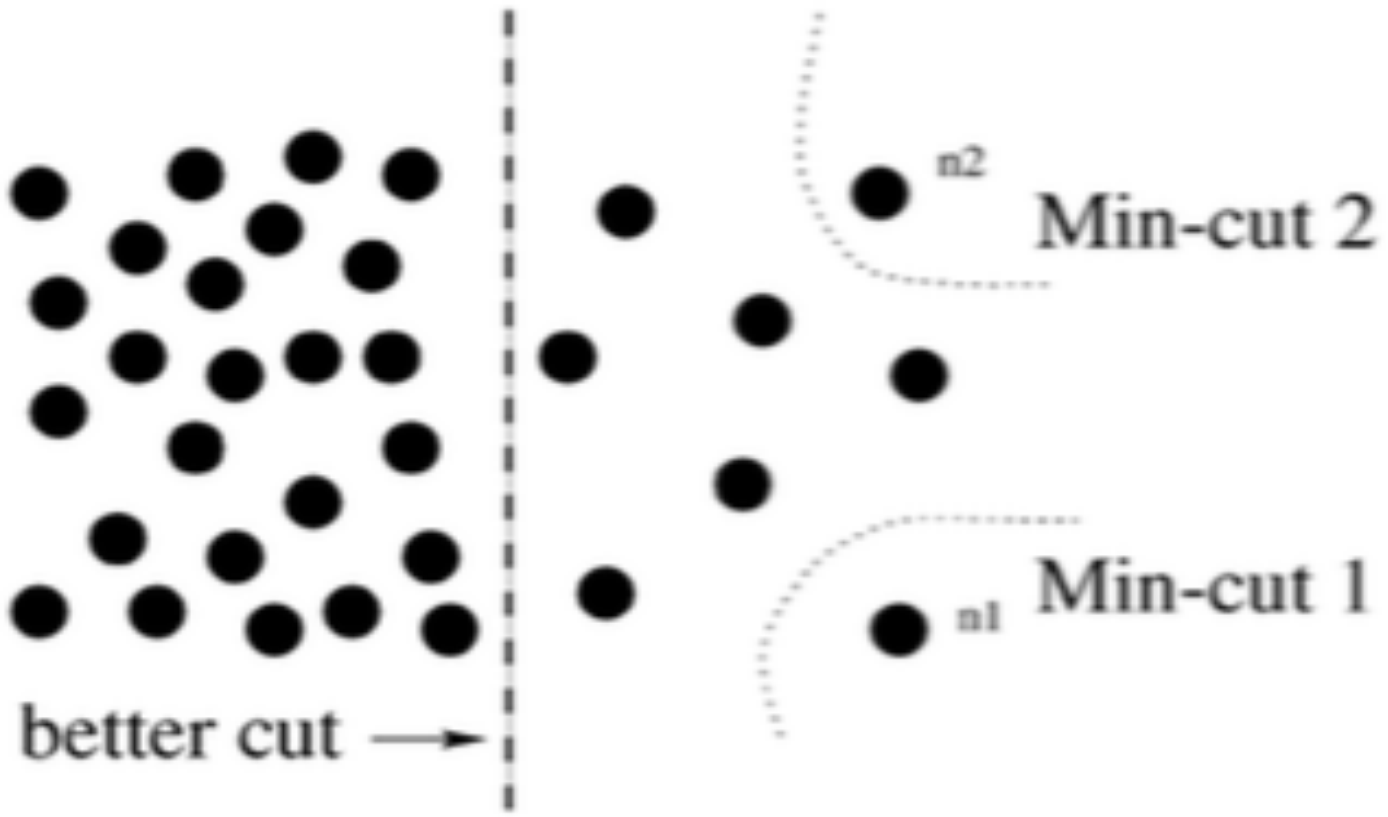


Fig. 1. A case where minimum cut gives a bad partition.

Algorithms

- Optimization Minimization of some loss function on a training set of examples
- Approaches
  - Decision tree maps observations about an item to conclusions about the item's targeted value
  - Association Learning rule discovering interesting relationships between variables in a large database
  - Clustering assignment of a set of observations into subsets or clusters based on similarity
  - Bayesian networks update the probability of a hypothesis as more evidence is acquired

$$P(H \mid E) = \frac{P(E \mid H) \cdot P(H)}{P(E)}$$

Representation

After segmenting the image into regions/superpixels, representing a region involves two choices:

1) Represent the regions in terms of their boundary or external characteristics

2) Represent the regions in terms of their internal characteristics or the pixels comprising the region

Description

Next, is describing the region based on the previously chosen representation

Features/descriptors by which to describe boundary by

- Length
- Orientation of the straight line joining its extreme points
- # of concavities in the boundary

Object Recognition

- Regions now referred to as objects or patterns
- Pattern recognition approaches:

  1) Decision-theoretic quantitative descriptors i.e. length, area, texture

  2) Structural relational descriptors recursive relationship involving primitive elements
- Goal of Object Detection
  - NOT to understand how humans perceive/detect objects with such ease rather to create computer methods for automatic object detection

References

1. Shi, J., Malik, J.. “Normalized Cuts and Image Segmentation.”
2. http://www.cvc.uab.es/~asappa/publications/C__LNCS_4739_1105.pdf
3. http://www.cs.colostate.edu/~draper/papers/badal_wacv94.pdf
4. http://www.templetons.com/brad/robocars/cameras-lasers.html