Look Around and Beyond: Enabling Extended Vision for Future Automobiles

USC University of Southern California

Ming Hsieh Department of Electrical Engineering

USC Viterbi School of Engineering

Hang Qiu, Fawad Ahmad, Ramesh Govindan, Marco Gruteser, Fan Bai, Gorkem Kar

NSL

RUTGERS THE STATE UNIVERSITY OF NEW JERSEY

GM

Abstract

As autonomous cars do today, future vehicles will have rich sensors to map and identify objects in the environment. For example, many autonomous cars today come with line-of-sight depth perception sensors like 3D cameras. These cameras are used for improving vehicular safety in autonomous driving, but have fundamentally limited visibility due to occlusion, sensing range, extreme weather, and lighting conditions. To improve visibility, and therefore safety, not just for autonomous vehicles but for other Advanced Driving Assistance Systems (ADAS), we explore a capability called Augmented Vehicular Reality (AVR). AVR broadens the vehicle's visual horizon by enabling it to share visual information with other nearby vehicles, but requires careful techniques to align coordinate frames of reference, and to detect dynamic objects. Preliminary evaluations hint at the feasibility of AVR and highlight research challenges in achieving AVR's potential to improve autonomous vehicles and ADAS.

Approach

Vehicle Relative Positioning

a) Crowdsourcing a sparse static 3D feature map of the world.

b) Localizing vehicles by detecting, tracking and matching 3D sparse features with the static map.

Static HD feature map

Vehicle Localization using Feature Map

Vehicle Vision Extension

a) Creating the 3D point cloud of vehicle surroundings.

b) Tranforming objects from vehicle perspective to the common world coordinate frame of reference.

c) Sharing objects to other vehicles by sending the point cloud via light weight V2V communication.

ZED Stereo Camera

Point Cloud of Vehicle Surroundings

Challenges

Perspective Transformation

a) Transformation matrix.

$$Tcw = \begin{bmatrix} RotX.x & RotY.x & RotZ.x & Translation.x \\ RotX.y & RotY.y & RotZ.y & Translation.y \\ RotX.z & RotY.z & RotZ.z & Translation.z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

b) Tramsforming voxels to world coordinate frame of reference.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} RotX.x & RotY.x & RotZ.x & Translation.x \\ RotX.y & RotY.y & RotZ.y & Translation.y \\ RotX.z & RotY.z & RotZ.z & Translation.z \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

c) Transforming voxels to another vehicle perspective.

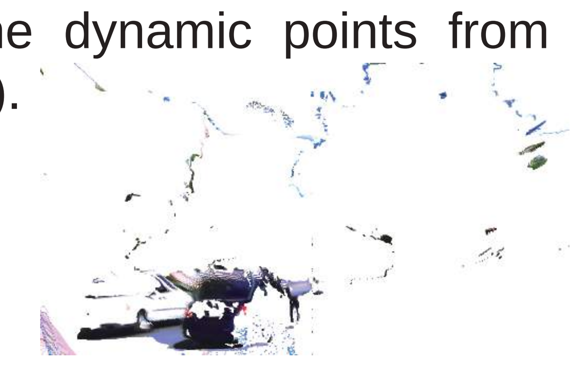
$$V_b = Tbw^{-1} * Taw * V_a$$

Compressing Point Cloud to Meet the Bandwidth

a) Light weight voxel tracking via corresponding pixel matching between consecutive frames.

b) Isolating and transmitting only the dynamic points from moving objects (e.g. cars, pedestrians).

Homography from Consecutive Frames

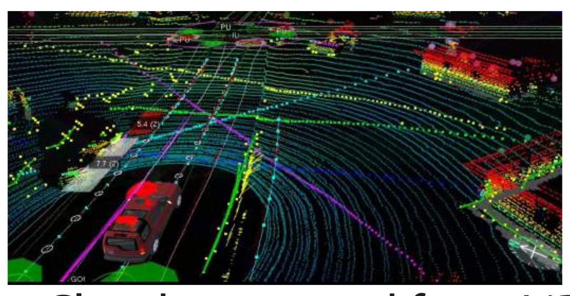


Dynamic Points Isolation

Motivation

A crucial component of future vehicles is a 3D sensing capability that provides depth perception of a car's surroundings. State-of-the-art technologies include LiDar, Radar, (stereo) camera and Infrared sensors, most of which only provide line-of-sight perception so that obstacles can often block a vehicle's sensing range. Moreover, the effective sensing range of these sensors is often limited by different weather conditions (e.g. fog, rain, snow, etc.) or lighting conditions. In this paper, we explore the feasibility of a simple idea: extending the visual range of vehicles through communication. Extended vision can help in many settings, such as buses that occlude children crossing at a crosswalk, or trucks that occlude a left-turning vehicle's view of approaching cars, etc.

LiDar Sensor



Point Cloud generated from LiDAR

photos credit: Velodyne & http://www.hizook.com/blog/2009/01/04/velodyne-hdl-64e-laser-rangefinder-lidar-pseudo-disassembled

Results

Experimental Setup:

Each experiment car is equipped with ZED stereo camera mounted on the top of the front windshield and a smartphone attached to it. The stereo camera records the video stream and computes the 3D point cloud as well as depth information, while the mobile phone records GPS and all motion sensors, i.e., gyroscope, accelerometer, magnetometer, etc..

Stereo Camera Setup

Extended Vision:

Leader View

Follower View

Extended View

Quantifying the Bandwidth:

a) Full: everything it sees.

b) Dynamic: moving voxels only.

c) Object: point clouds belonging to the objects detected.

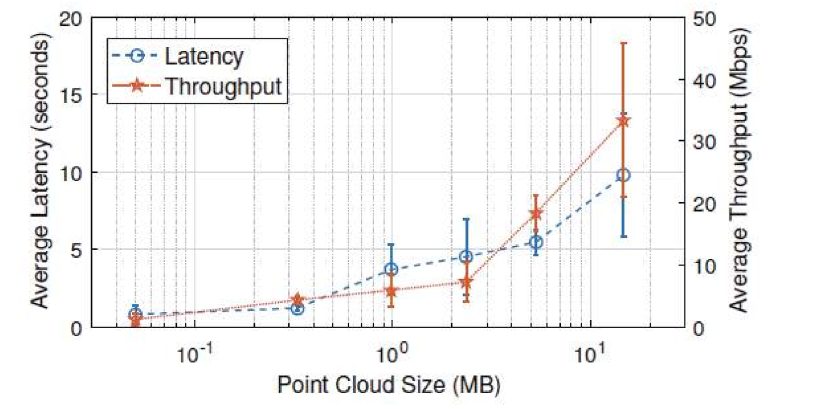
d) Labels: 3D bounding box and the label of the object.

	VGA (640 x 480)	720P (1280 x 720)	1080P (1920 x 1080)
Full	4.91 MB	14.75 MB	33.18 MB
Dynamic	0.79 MB	2.36 MB	5.30 MB
Object	0.33 MB	0.98 MB	2.21 MB
Labels	0.05 MB	0.05 MB	0.05 MB

Point Cloud Data Size per Frame

Throughput and Latency:

Currently, no existing wireless system can sustain AVR bandwidth requirement. We establish a WiFi P2P network to explore how much is the gap. The highest throughput experienced is 60Mbps. Handshaking causes too mcuh overhead for AVR to run in real time. We plan to work around 802.11p to improve the performance.



Transmission Throughput and Latency of Different Point Cloud Data Size

Conclusion

In this paper, we present the concept of augmented vehicular reality and the design and implementation of AVR system. AVR is not only beneficial to human drivers as a driving assistant system, but also, more importantly, enables extended vision for autonomous driving cars to make better and safer decisions. With extended vision, we have demonstrated that AVR can effectively remove sensing range limitations and line-of-sight occlusions. Future work on AVR will focus on improving the throughput of processing objects, addressing the bandwidth constraints, and reducing latency.