Combining Bayesian and AI approaches for Autonomous Driving

Christian Laugier

Combining Bayesian & AI approaches for Autonomous Driving

Christian LAUGIER, PhD & Dr Es Science

Research Director at Inria & Scientific Advisor for Probayes

Inria Chroma team & IRT Nanoelec – christian.laugier@inria.fr

Keynote Speech

Workshop on "Perception & Navigation for Autonomous Robotics in Unstructured and Dynamic Environments"

IROS 2021, Prague, September 27th 2021

The research work presented in this talk includes contributions from several Chroma team members:

- ***Postdocs & Researchers:*** *D. Sierra-Gonzalez, O. Erkent*
- ***Research engineers:*** *L. Rummelhard, A. Negre, N. Turro, J.A. David, J. Lussereau, T. Genevois, A. Paigwar*
- ***Also some former PhD students & Postdocs***

=> I warmly thank each of them

Brief overview of Technologies & Ongoing challenges for Autonomous Driving

- **Due to both Road nuisances** *(pollution, traffic jams, accident…)* **& Current technological advances, Technologies & Services for <u>Human Mobility</u> are evolving rapidly** => *Reduce road Nuisances & Costs* + *Increase Efficiency & Safety*
- **Last decade: A Technological breakthrough** *(Sensors, GPU, DL…)* **& Numerous AV experiments in real traffic conditions**

Commercial Autopilot L2 system
(tested by customers)

- Large fleet & 20 Million miles covered
- **Self-driving taxi service L3**
(1st in US, Phoenix, Dec 2018)

- Similar **Self-driving taxi service L3**
- Experimented in several cities

Autonomous Shuttle service concept

- **Millions of miles have been covered in last decade, but SAFETY is still not guaranteed !** *(as evidenced by several accidents)*
 => *Additional R&D work is still mandatory on 2 main research lines:*
 - ***Improving "Perception & Decision-making" technologies*** *(in particular for Robustness & Scalability issues)*
 - ***Developing "Validation & Certification" approaches for future deployment*** *(in addition to traditional testing approaches)*

 => *Develop a* ***new Framework*** *including* ***Realistic simulators***, ***Augmented reality*** *capabilities, Integration of* ***<u>Formal methods</u>***
 This subject is addressed by ***Inria & IRT Nanoelec*** *(several EU & National projects)* ***… But not presented in this talk !***


Testing setup (Inria & IRT Nano)


Realistic Simulators & Augmented Reality testing systems (Inria & IRT Nano)

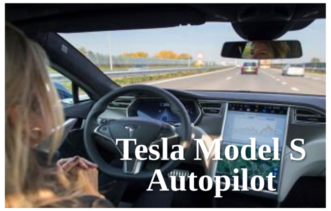

Disappointing evaluation of commercial "emergency braking products"
(Report published by AAA in 2019)

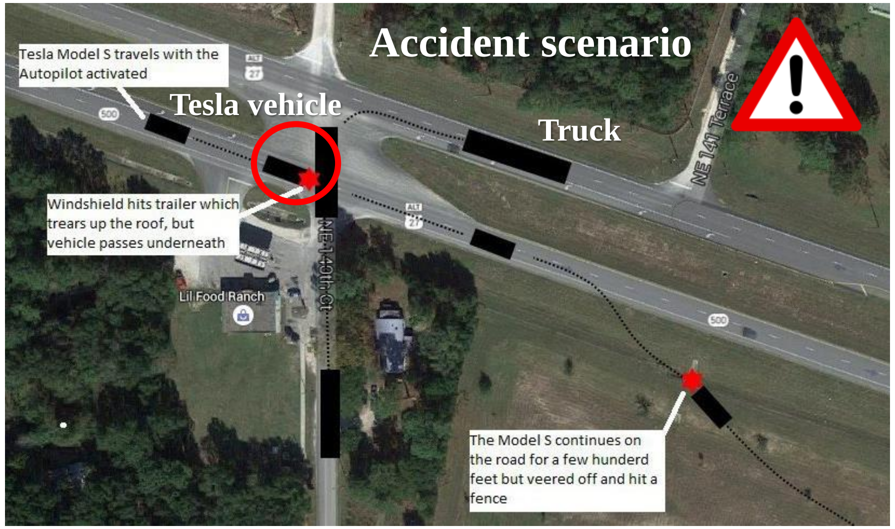
Fatal accidents involving AVs – *Perception failure*

❑ Tesla: Driver was killed in a crash with Autopilot L2 active

May 2016

- ✓ *A truck turned left and blocked the route of the Tesla vehicle*
- ✓ *The Autopilot failed to detect the white moving truck, with a brightly lit sky (Camera Mobileye + Radar)*
- ✓ *The human driver was not vigilant & didn't took over*





❑ Uber: Pedestrian killed by Self-driving L3 vehicle => *First fatal crash involving a pedestrian, see video bellow*

Temple, Arizona, March 2018

- ✓ *Despite the presence of multiple sensors (lidars, cameras …), the perception system failed to detect the pedestrian & didn't disengaged*
- ✓ *In addition, the "Safety Driver" reacted too lately (1s before the crash)*

Improving Safety: AVs have to face two main challenges

1st Challenge : *The need for **Robust**, **Self-diagnosing** & **Explainable Embedded Perception***

Video source: AutoPilot Review @ youtube.com

Video Scenario *(Tesla vehicle with Autopilot active):*

- *The Tesla perception system **failed to detect the barriers** blocking the left side route.*
- *The **human driver has to take over and steer** the vehicle away from the blocked route.*

Improving Safety: AVs have to face two main challenges

2nd Challenge: *The need for **Understandable Driving Decisions** (Sharing the road with human drivers)*

Unfortunately, Human drivers actions are determined by a complex set of interdependent factors which are very hard to model *(e.g. intentions, perception, emotions …)*

$\Rightarrow$ ***Predicting human driver behaviors is inherently uncertain***

$\Rightarrow$ ***AV have to reason about "uncertain intentions" of the surrounding vehicles***

Video source: The Telegraph

Video scenario involving a **Waymo AV &** a **Human Driven Bus.**

The accident scenario was recorded by the dashcam of the Bus (bus not visible in the video)

- ***The Waymo AV** was blocked by an obstacle and **decided to make a left lane change***
- ***The bus driver** misunderstood the AV intention and **didn't yield***
- *The two vehicles collided*

Perception & Decision-making challenges – *Outline*

Dynamic scene analysis & Navigation Decisions

Embedded Perception & Decision-making for Safe Intentional Navigation

Detecting & Reacting in real-time to Unexpected Events

⇒***Scene interpretation*** *using Sensing & Prior knowledge*

⇒***Safe Navigation decisions*** *(using semantics & context)*

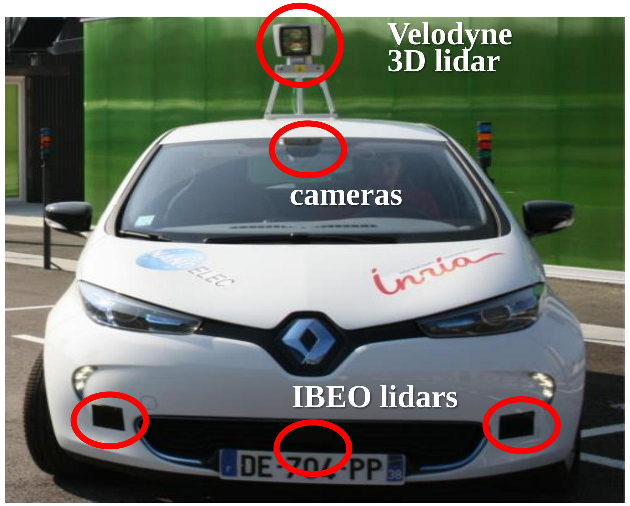
⇒***Detect & Predict*** *imminent collision with* ***"something"*** *(unclassified unexpected event)*

⇒*Activate appropriate* ***reflexive avoidance strategy***

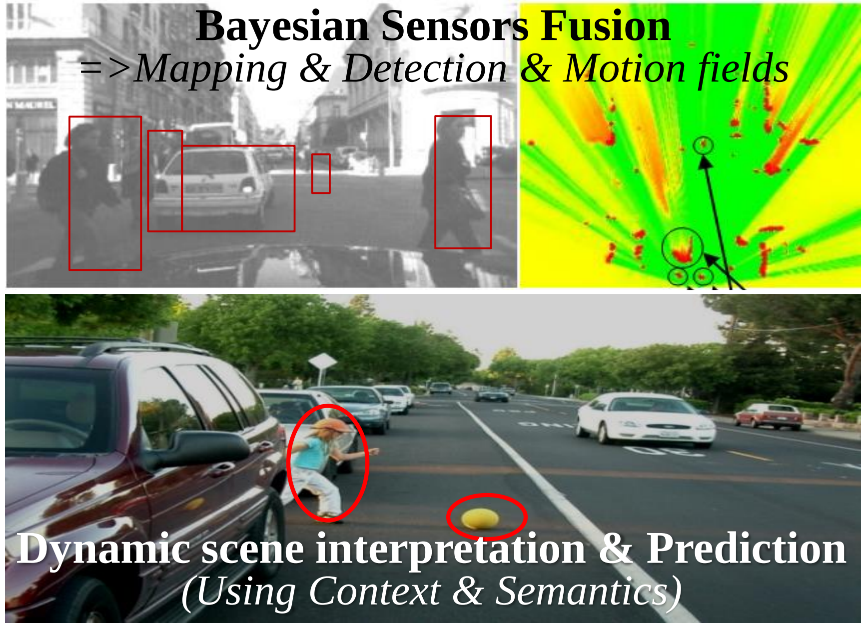
- **Main constraints:** *Dynamic & Open Environments + Incompleteness & Uncertainty + Sensors limitations & Huge amount of data to process*
- **Implementation constraints:** *Real-time processing + Software / Hardware integration*
- **Mixed traffic** (Human in the loop): *A Human-Aware Decision-making process is required*

  => Taking into account Interactions + Behaviors + Social & Traffic rules
- **Our Approach:** ***Combining Probabilistic & IA approaches***

1st Paradigm: Embedded Bayesian Perception



Embedded Multi-Sensors Perception
⇒ Continuous monitoring of the surrounding dynamic environment



Main Objective

1st Leverage dynamic insights to better understand a complex traffic scene !!!

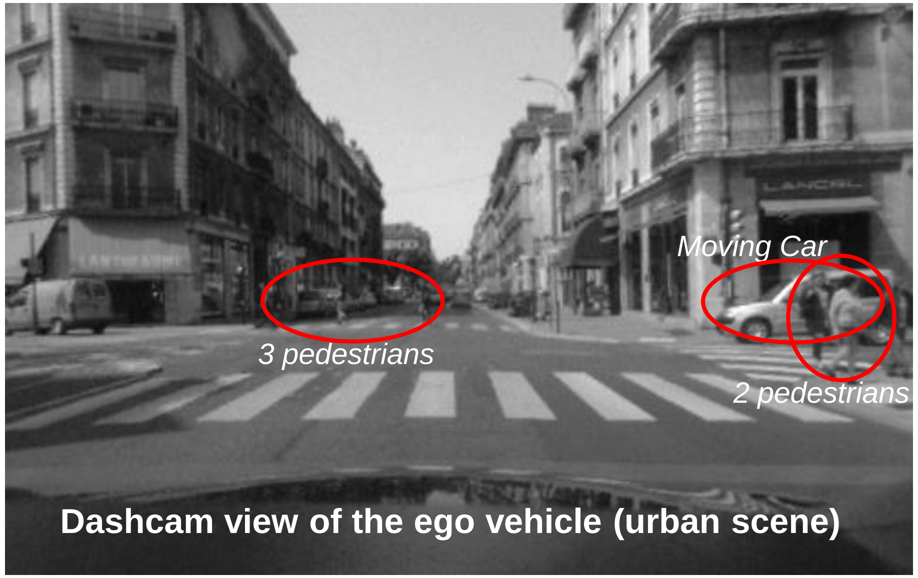
*2nd Process Noisy data, Uncertainty, Dynamics … while respecting strong **Real-time constraints***

Main Features of our approach

- ✓ *Reasoning about **Uncertainty** & a **Time window** => Reasoning on **Past & Future (predicted) events***
- ✓ ***Improving robustness** using **Bayesian Sensors Fusion***
- ✓ ***Interpreting the dynamic scene** using **Contextual & Semantic** information*
- ✓ ***Implementing the system** using **hardware accelerators** (GPU, Multicores, Microcontrollers ...)*

Dynamic Occupancy Grid & Bayesian Filtering

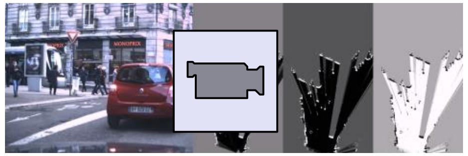
=> Use dynamic information for a better understanding of the observed scene

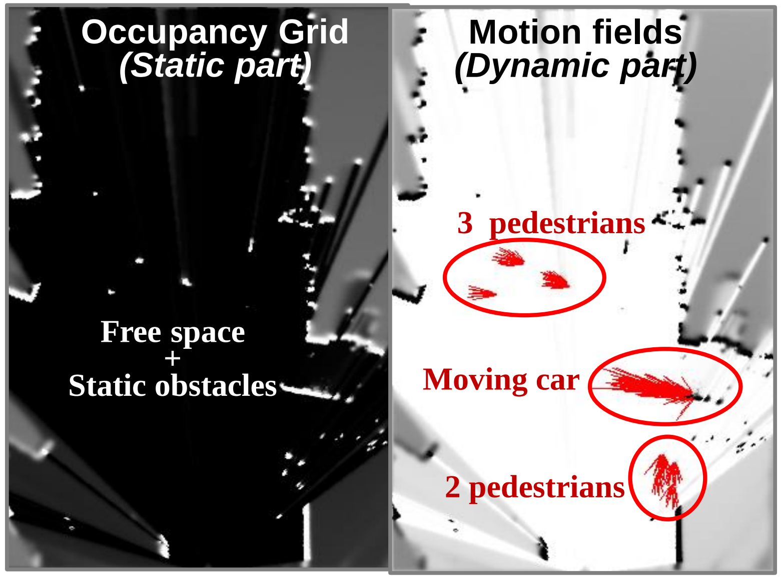


Dashcam view of the ego vehicle (urban scene)

Sensors data fusion
+
Bayesian Filtering
+
Extracted Motion Fields

1st Embedded & Optimized version (HSBOF, patent 2014)

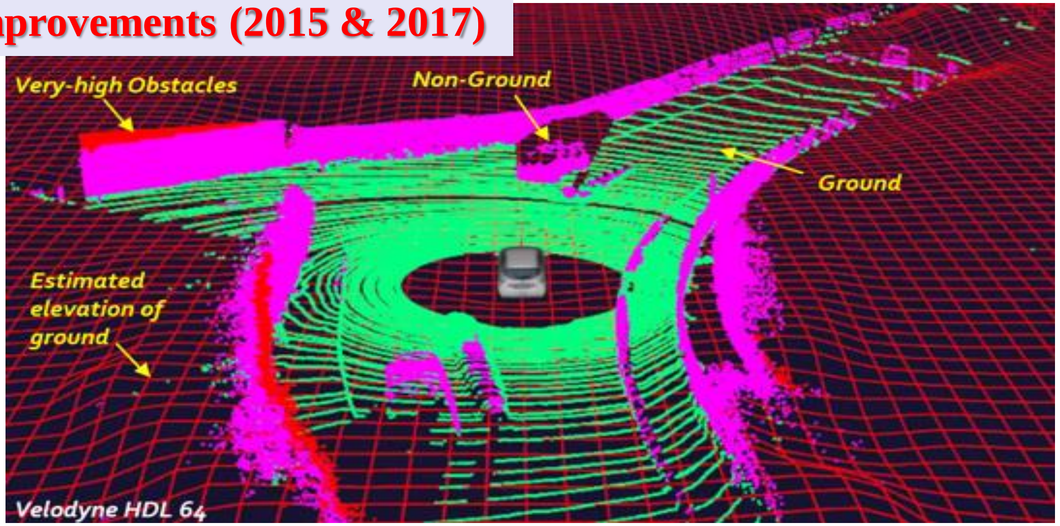




Patented system Improvements (2015 & 2017)

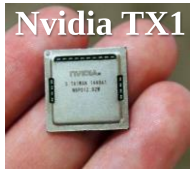
Classification *(using Deep Learning)*

Moving Objects Classification + Detection & Tracking
(CMCDOT 2015 & "Dense Occupancy Tracker")

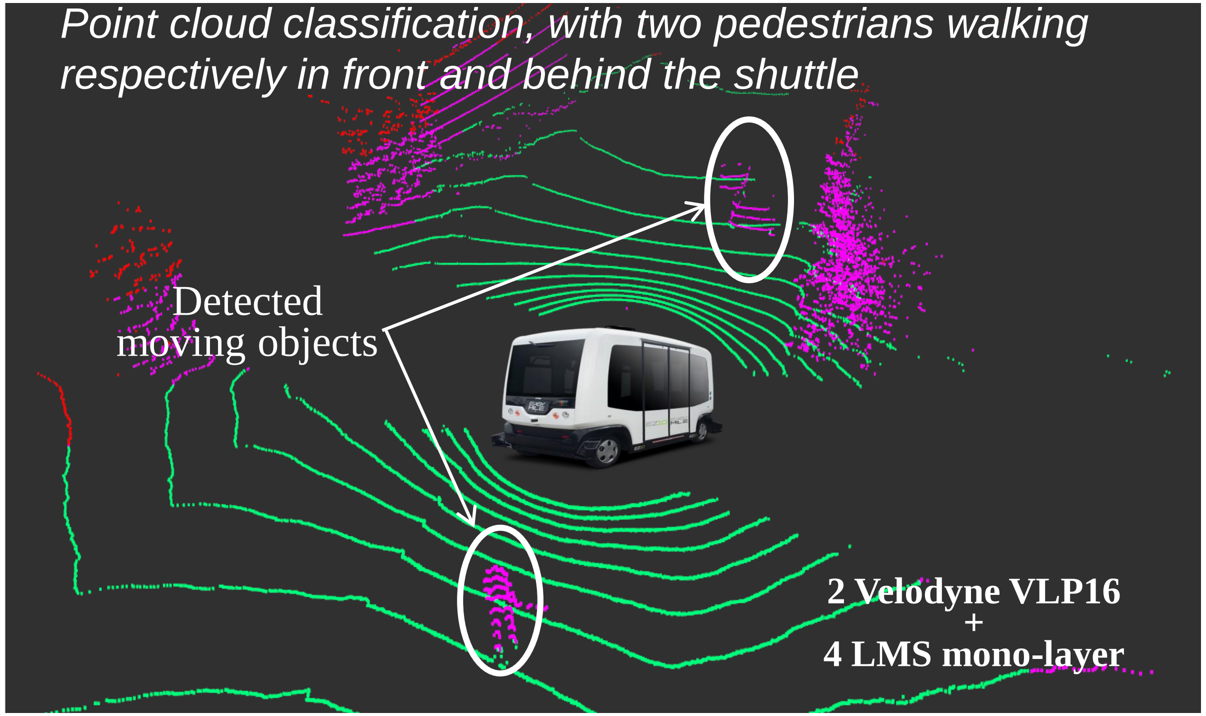


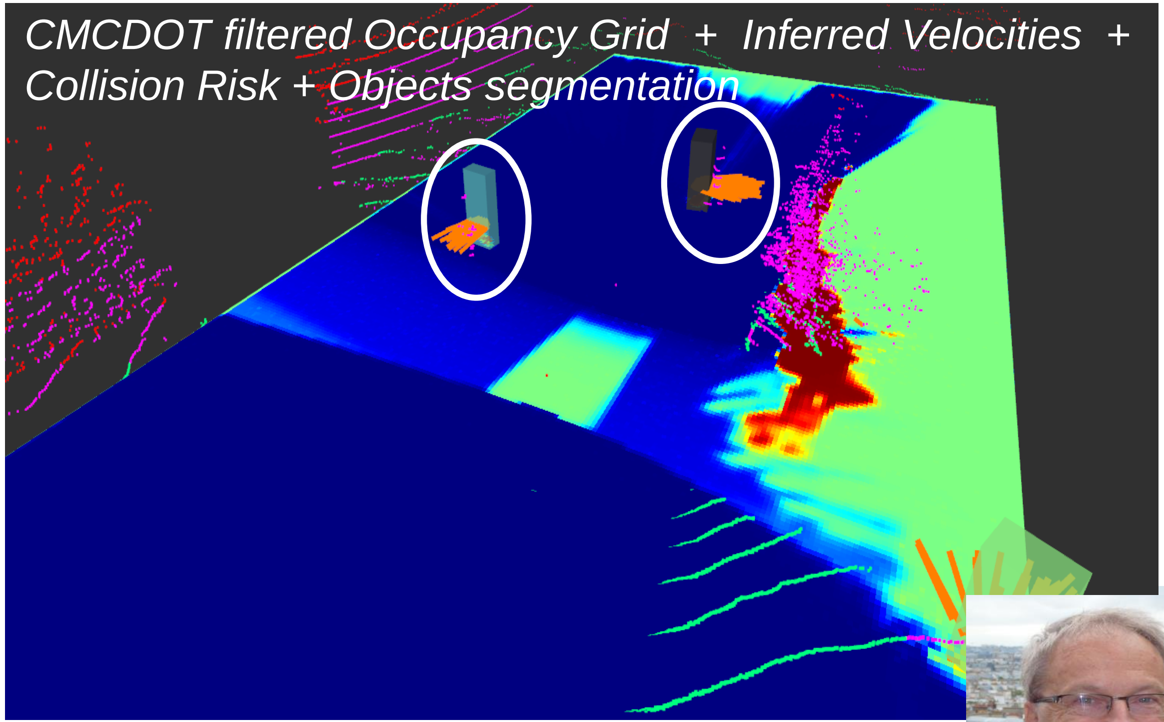
Ground Estimation & Point Cloud Classification
(patent 2017)

System Integration on a Commercial Vehicle



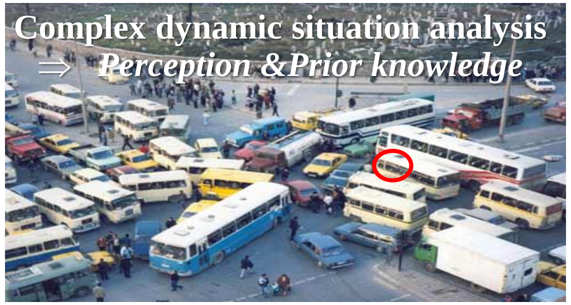
- **POC 2019: Complete system implemented on Nvidia TX1**, and easily connected to the shuttle system network *in a few days* (using ROS)
- **Shuttle sensors data** has been fused and processed in **real-time**, with a successful Detection & Characterization of the **Moving & Static Obstacles**
- **Full integration on a commercial product** still under development with an industrial company (confidential)





2nd Paradigm: **Collision Risk Assessment & Avoidance Strategies**

=> Avoiding Pending & Future Collisions



Situation analysis + Prediction & Collision Risk Assessment + Safest Driving Decision

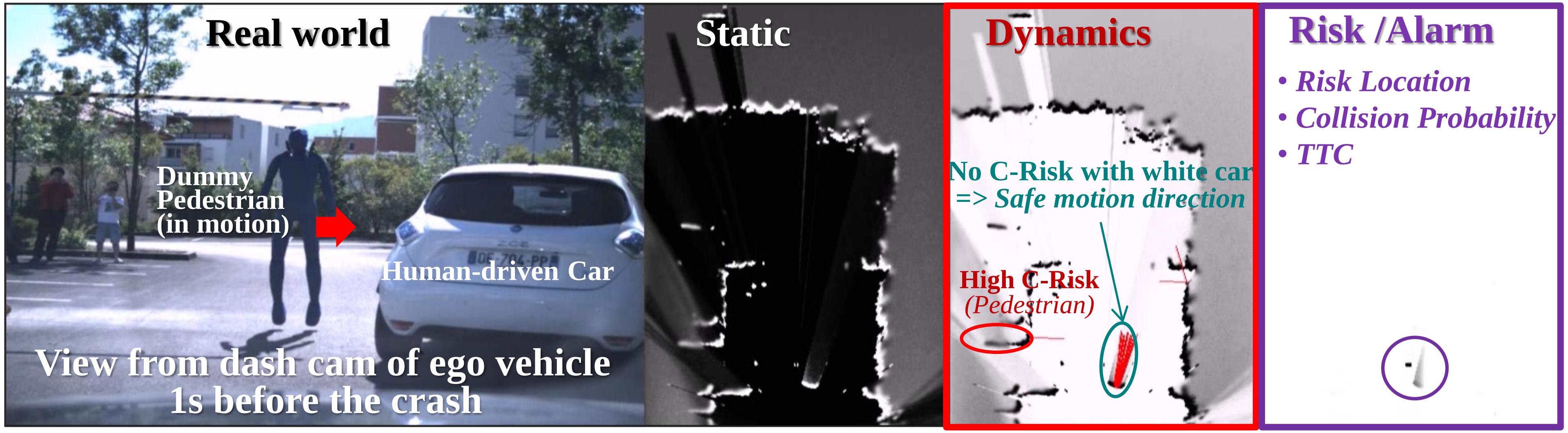
Main Features of our approach:

- ✓ **Predict dynamic environment changes** on a given "**time horizon t+δ**" *(using both History & Motion models)*
- ✓ **Estimate** the **Probabilistic Collision Risk** at horizon **t+*δ*** *(δ = a few seconds ahead)*. ***Two types of Collision-Risk (C-Risk)*** *have to be considered, depending on both the* ***time horizon*** *& the* ***semantic information available***
- ✓ **"Short-term C-Risk" characteristics:** *The **Collision Risk is** estimated at the **Grid-level**, the **Moving obstacles** are not classified, the **time-horizon δ = 3-5s**, and the **Prediction step is based on a "conservative motion hypothesis"***
- ✓ **"Mid-term C-Risk" characteristics:** *The **Collision Risk** is estimated at the **Object level**, the **Involved entities are classified**, and the **Prediction step requires to reason about "Behavior models & Semantic information"***
- ✓ **Driving Decisions** *are taken on the basis of the **Predicted behaviors** of all the observed surrounding traffic participants (cars, cycles, pedestrians…), the **Social &Traffic rules**, and the **expected traffic participants Interactions***

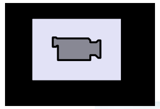
Short-term C-Risk – *Approach & Experimental results*

Experimental results (Urban streets & Test track)

Collision prediction in a crash scenario (in test track)



Video: Short-term C-Risk
- Yellow => *3s before collision*
- Orange => *2s before collision*
- Red => *1s before collision*

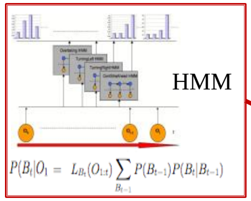


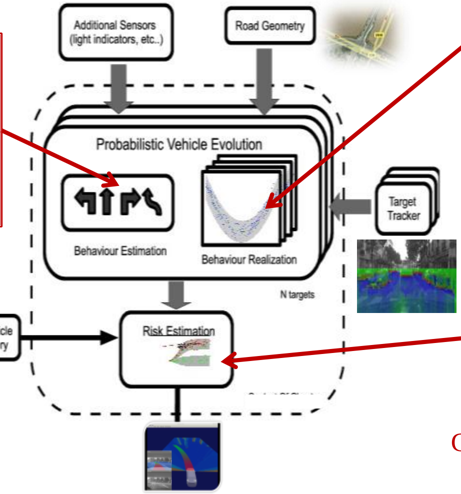
Mid-term C-Risk – *Approach & Experimental results*

=> Increased time horizon & complexity + Reasoning on Behaviors & Interactions

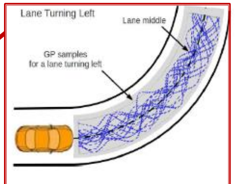
❑ Trajectory prediction & C-Risk assessment *=> Patent 2010 (Inria, Toyota, Probayes)*

Learning & Predicting Behaviors

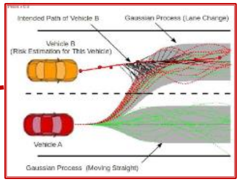




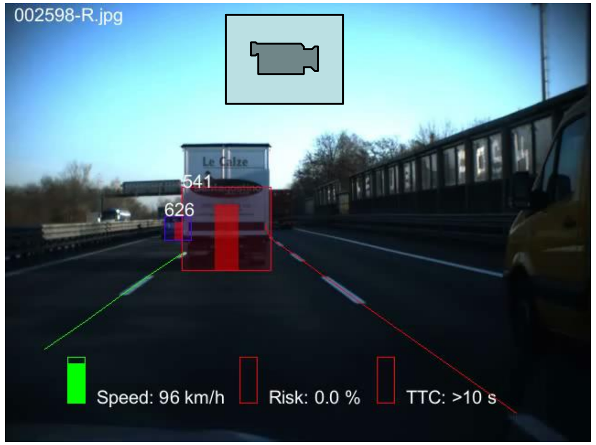
Behaviors to Trajectories



Gaussian Processes



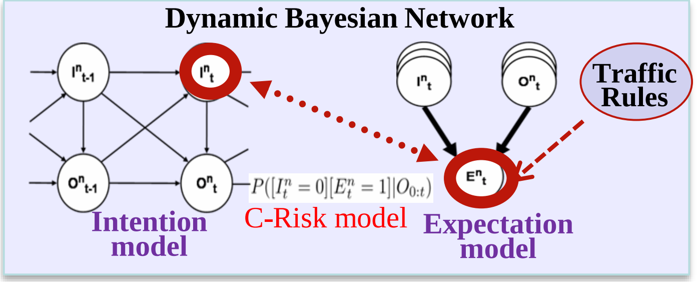
Collision Risk Assessment (MC Simulation)



Courtesy Probayes

Cooperation still on-going (R&D contracts + PhDs)

❑ Intention & Expectation *(Mixed Traffic & Interactions) => Patents 2012 (Inria - Renault) & 2013 (Inria - Berkeley)*



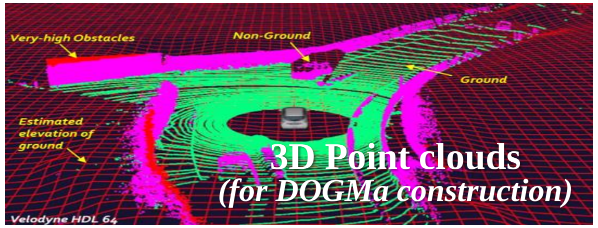
Human-like reasoning

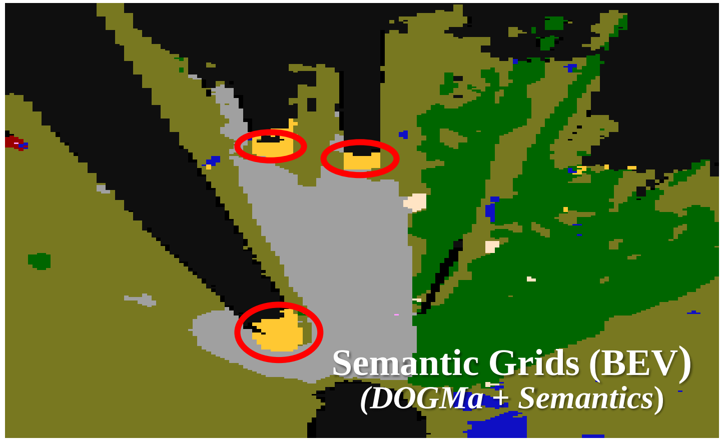
Cooperation still on-going (R&D contracts + PhDs)

3rd Paradigm: Models improvements using Machine Learning

=> *Exploit the complementarities of* ***Bayesian Perception*** *&* ***DL based Computer Vision approaches***
(Richer semantics, Better understanding of 3D dynamic scenes, Mandatory for Decision-making step)

❑ Perception level: *Fusing DOGMa & Semantically segmented RGB images*





- Unknown
- Building
- Sky
- Road
- Vegetation
- Sidewalk
- Car
- Pedestrian
- Cyclist
- Signage
- Fence
- Free
- Static
- Dynamic

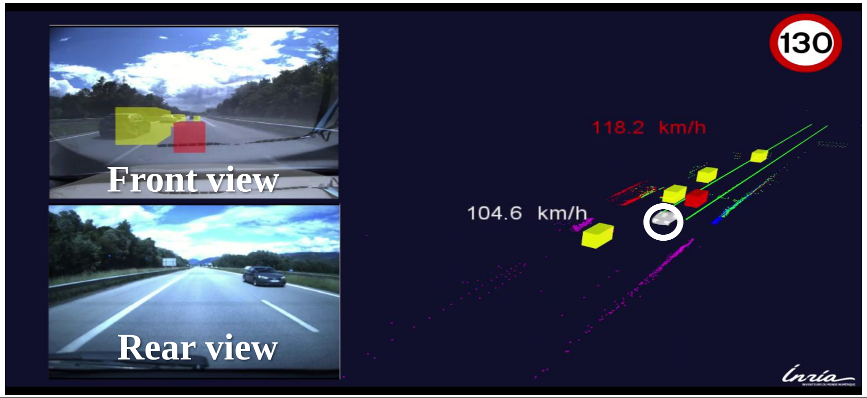
DOGMa: *Dynamic Occupancy Grid Map*

BEV: *Bird's Eye View*

❑ Prediction & Decision-making level: *Learn driving skills for Autonomous Driving*

Step 1: ***Modeling Driver Behavior*** *using Inverse Reinforcement Learning (IRL)*

Step 2: ***Predict motions*** *of surrounding vehicles &* ***Make Driving Decisions*** *for Ego Vehicle*



Perception Level (1): Concept of "Semantic Grid"

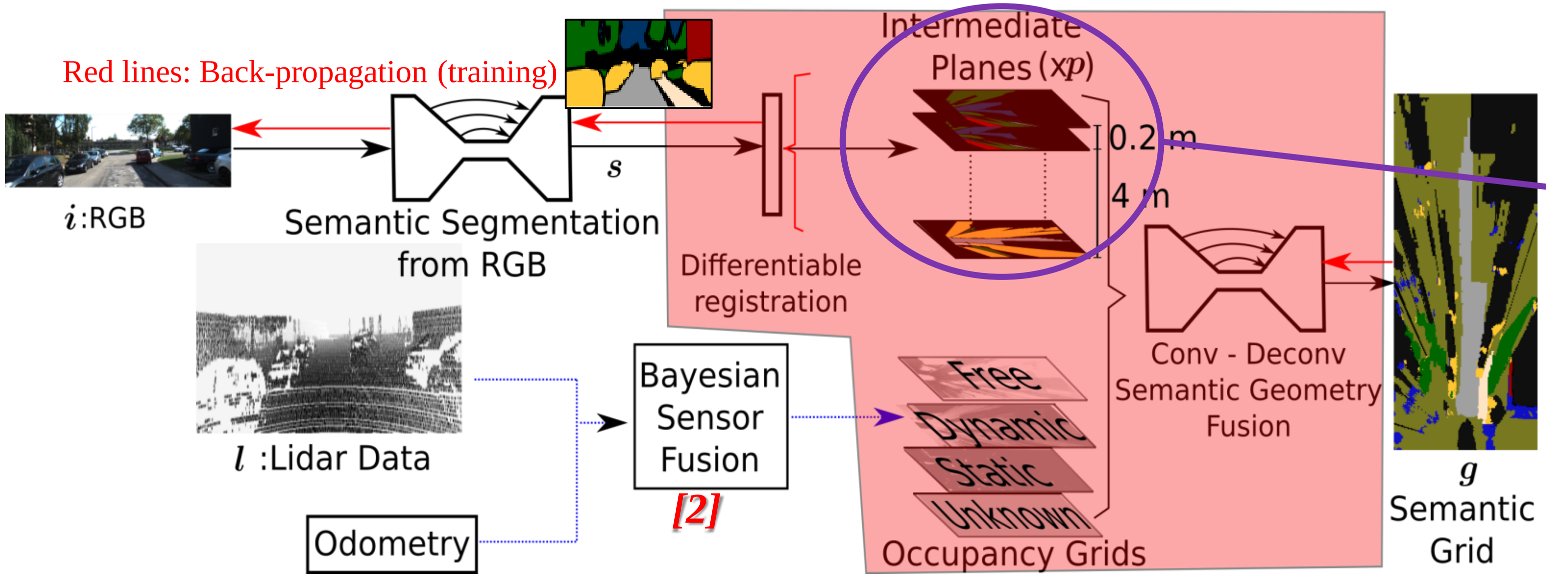
Objective: *Add **Semantic information** (cars, pedestrians, roads, buildings…) in each cell of the Dynamic Occupancy Grid Map, by exploiting **additional RGB inputs***

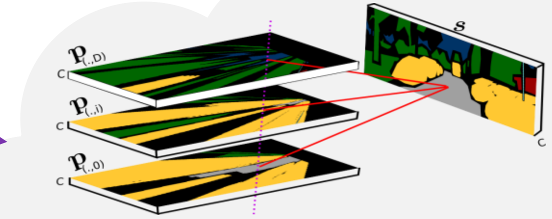
Approach: *A new **"Hybrid Sensor Fusion approach"** combining **DOGMa & Semantically Segmented RGB images (using DL)***

DOGMa: *Dynamic Occupancy Grid Map*

Ozgur Erkent
Starting Research Position at Inria Chroma team (2017-21). External Collaborator at Inria & Assistant Professor at Ankara Univ (since July 2021)

Semantic Grid Neural Network *[1] [3] [Patent Inria & Toyota 2019]*





~ 20 Intermediate layers

=> Learns the approximate heights of the classes/objects

=> No 3D reconstruction required & Less sensitive to calibration errors (initial camera/3D points calibration)

Implementation / Testing: Segnet Cuda/GPU + Kitti dataset

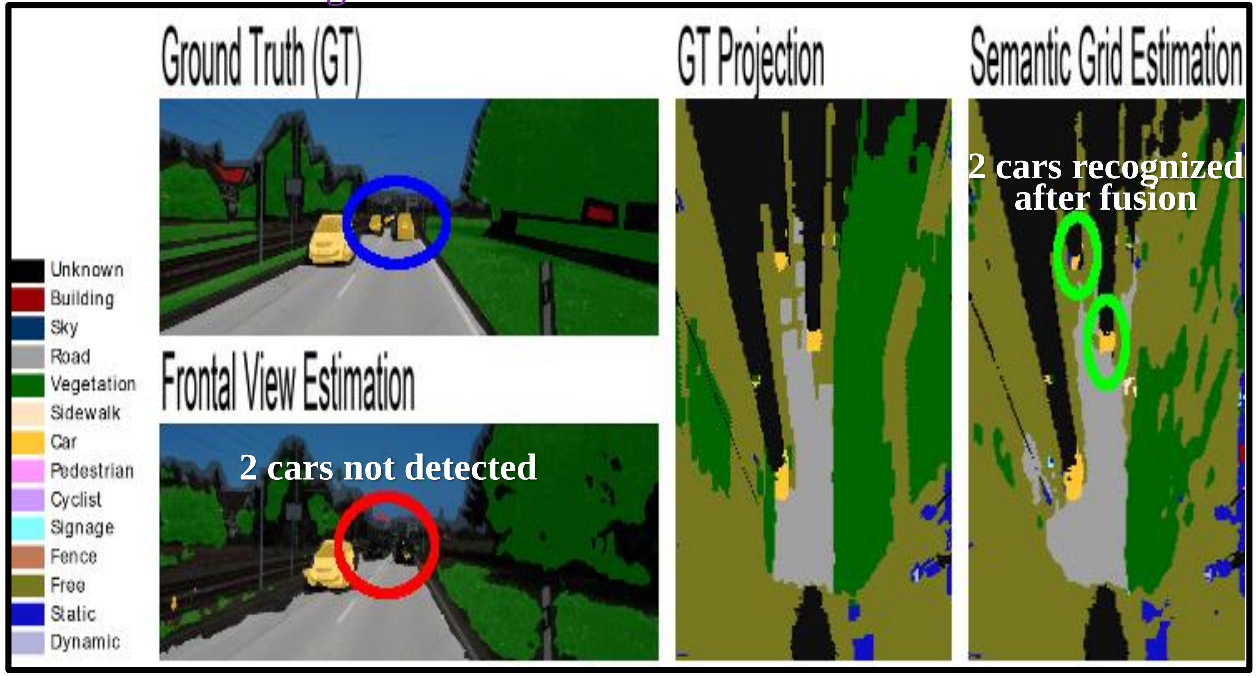
[1] *Semantic grid estimation with a Hybrid Bayesian and Deep Neural Network approach, 0. Erkent et al., IEEE IROS 2018*
[2] *Conditional Monte Carlo Dense Occupancy Tracker, Rummelhard et al., ITSC 2015*
[3] *Segnet: A deep convolutional encoder-decoder architecture for image segmentation, Badrinarayanan et al., IEEE PAMI 39(12) 2017*

Semantic Grids – *Experimental Results [1]*

(Using Kitti dataset)

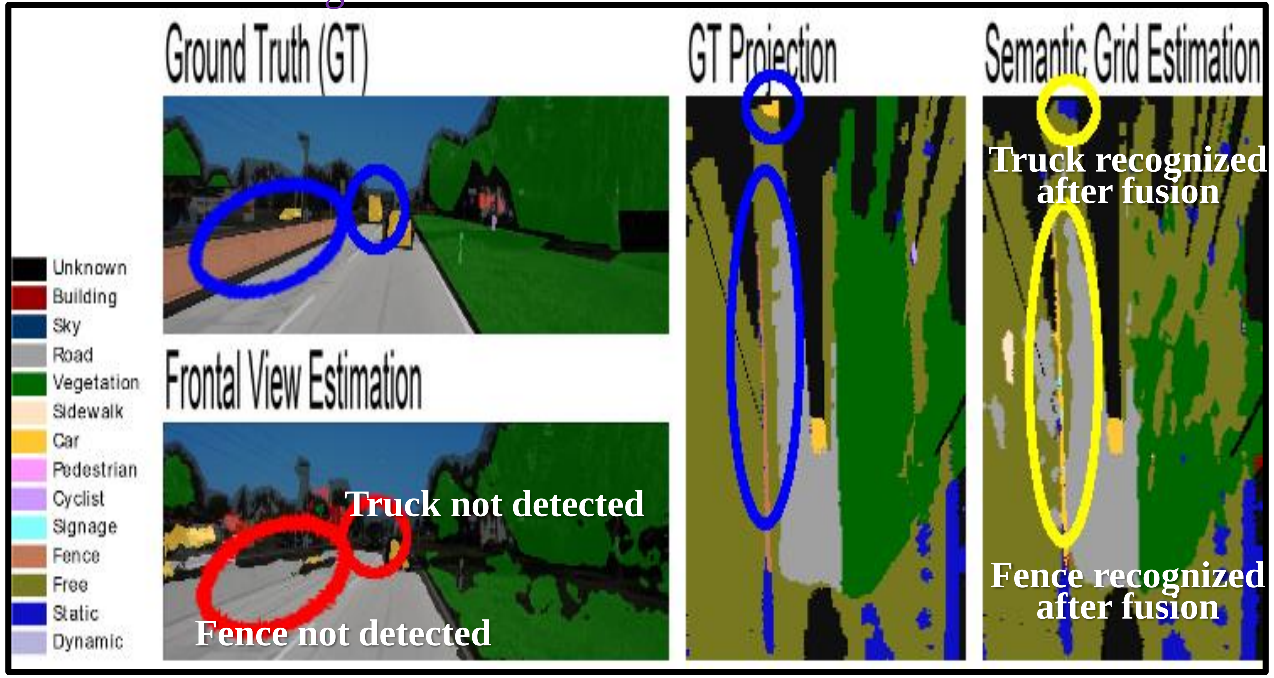
Traditional Semantic Segmentation

Hybrid Sensor Fusion



Traditional Semantic Segmentation

Hybrid Sensor Fusion



=> **2 cars not detected** in frontal view estimation…. but **recognized as an obstacle in Semantic Grid** *(with the help of DOGMa)*

=> **Truck & Fence not detected** in frontal view estimation…. but **recognized as an obstacle in Semantic Grid** *(with the help of DOGMa)*

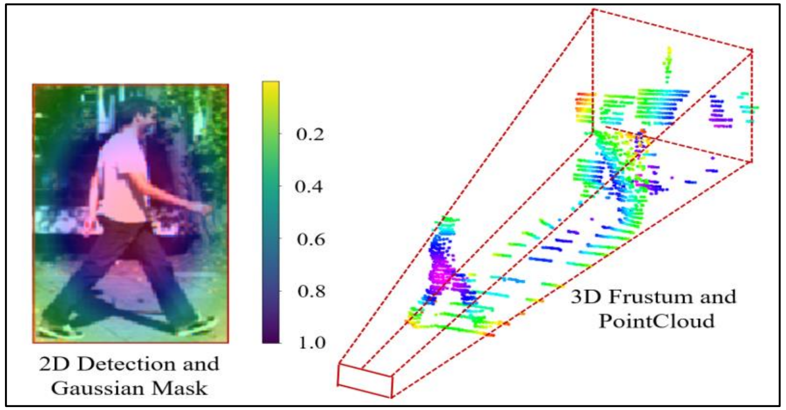
[1] Semantic grid estimation with a Hybrid Bayesian and Deep Neural Network approach, O. Erkent et al., IEEE IROS 2018

Perception Level (2): 3D Object Detection using Lidar & RGB camera
(Frustum PointPillars)

Anshul Paigwar
Research Engineer, Inria

David Sierra-Gonzalez
Starting Research Position, Inria

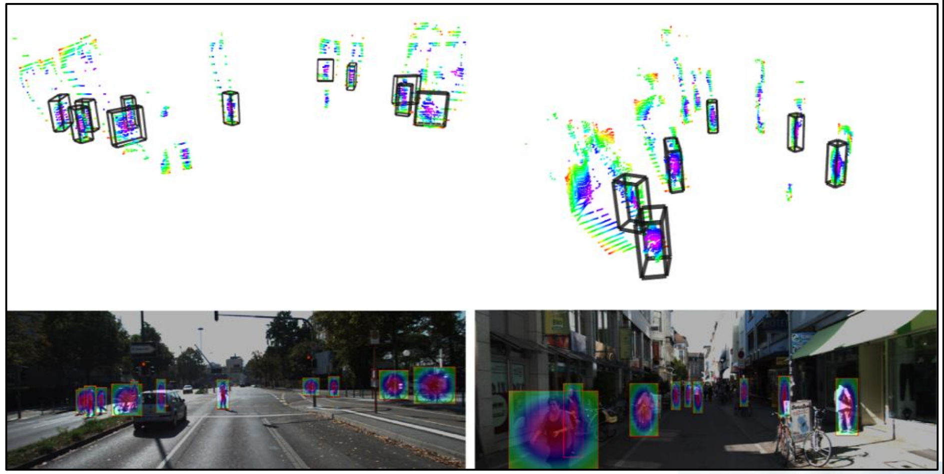
Objective: Realtime accurate detection & localization of <u>small object</u> like pedestrians in *3D point cloud of large-scale scenes*



Challenge: LiDAR only approaches have poor accuracy as pedestrians have *fewer data points and non rigid structure*

Proposed approach: Multi-stage sensor fusion approach *[1]*

- We outperform other multi-stage SOTA approaches for pedestrians BEV detection on the KITTI dataset
- Significantly faster runtime of 14 Hz.



Stage - I	Filtering & Masking of point cloud	Stage - II
2D Detections RGB Camera	→	3D Detection in filtered point cloud

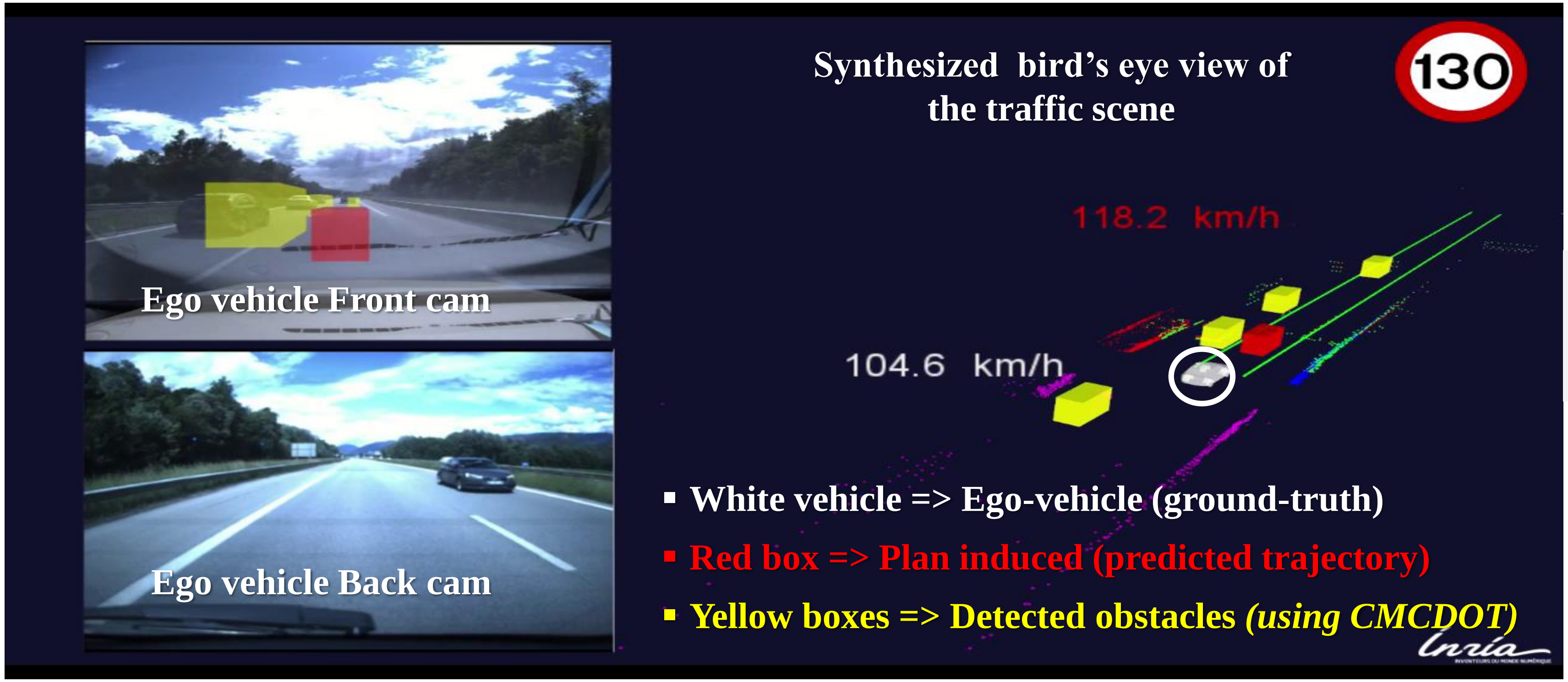
[1] Frustrum-Pointpillars: A Multi-Stage Approach for 3D Object Detection using Lidar and RGB camera, A. Paigwar et al, ICCV 2021 Workshop on "Autonomous Vehicle Vision" (Oct 2021)

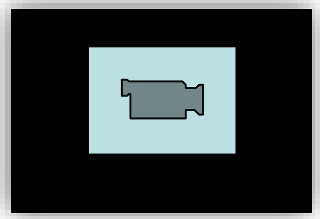
Prediction & Decision-making level: Learning Driving Skills for AD

1st Step: Driver behavior modeling

David Sierra-Gonzalez
Starting Research Position
Inria Chroma team

- **Learn Model parameters** from real driving demonstrations using ***Inverse Reinforcement Learning (IRL)***
- **Driver behaviors** are modelled using a **Cost function & Several hand-craft features** *(e.g. Lane index preferences, Deviation from desired velocity, TTC to frontal targets, Time-gap to rear targets, etc.)*
- A **training set containing *"interesting highway vehicle interactions"*** *has been constructed using our **Lexus vehicle***
- The obtained models can be leverage to both **Predict human-driver behaviors & Generate human-like plans for the ego vehicle in mixed traffic.** *[Sierra-Gonzalez et al, ICRA 2018] [Sierra-Gonzalez PhD thesis 2019]*





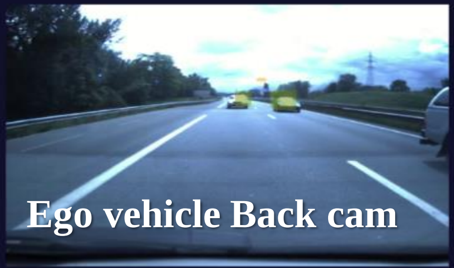
Video: Comparison between **demonstrated behavior** in test set & **behavior induced by the learned model**

Prediction & Decision-making level: **Learning Driving Skills for AD**

2nd Step: Motion Prediction & Driving Decisions

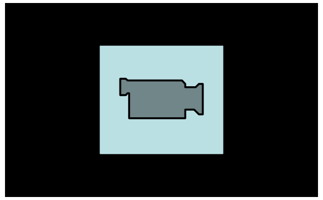
David Sierra-Gonzalez
Starting Research Position
Inria Chroma team

- A **realistic Human-like Driver Model** can be exploited to **Predict the long-term evolution** *(10s and beyond)* of traffic scenes *[Sierra Gonzalez et al., ITSC 2016]*
- For the **short/mid-term**, both the **Driver model** and the **Dynamics of the target** provide useful information to **determine future driving behaviors**
- Our **probabilistic model fuses** ***Model-based Predictions & Dynamic evidence*** to produce robust **lane change intention estimations** in highway scenes. *[Sierra Gonzalez et al., ICRA 2017] [Sierra-Gonzalez PhD thesis 2019]*



Synthesized bird's eye view of the traffic scene
&
Over vehicles expected intentions

- **Orange bar** => *Probability that the target executes a lane change according to the model (given the traffic situation)*
- **Red bar** => *Final lane change intention probability (fusing model-based prediction & dynamic evidence)*



Video: Comparison **between demonstrated behaviors** in test set & **behaviors induced** by the both the *learned model* and the *dynamics evidence*

Ongoing R&D work

- **Panoptic & Instance Segmentation + Domain adaptation** *(PhD student, Toyota & CPS4EU EU projects)*

- **Fusing RGB & Event cameras to improve objects detection in dynamic environments & adverse weather or lighting conditions** *(ES3CAP project)*

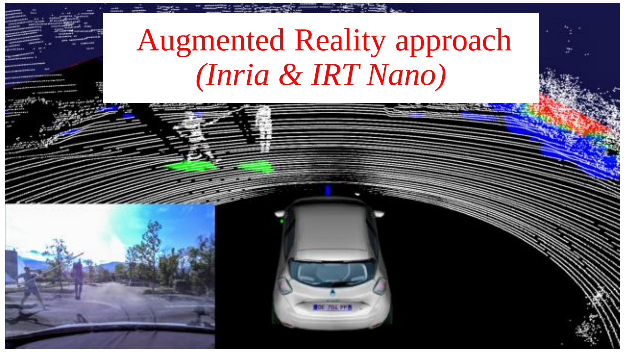
- **Detection & Tracking of Multiple Objects using DOGMa & Semantic Grids** *(Toyota project)*

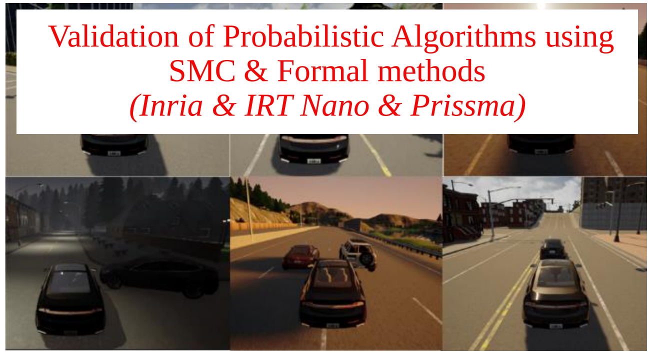
- **Distributed Perception & Situation Awareness using Connected Vehicles & Road-Side-Units** *(PhD student, IRT Nanoelec & International industrial company, confidential)*

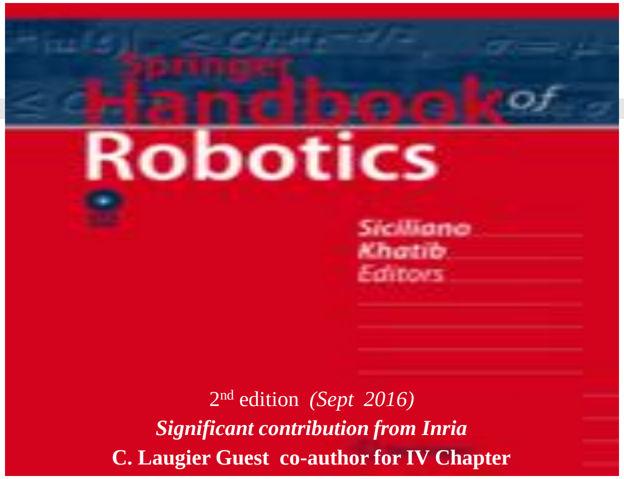
- **Real-time Planning & Control of Collision Avoidance or Mitigation Trajectories** *(PhD student, Renault)*

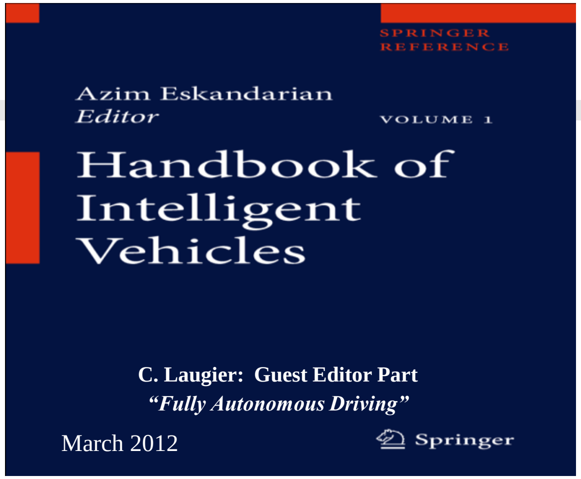
- **Validation & Certification framework combining Real Tests (in test-tracks), Realistic Simulation, Augmented Reality and Formal Methods** *(PhD student, IRT Nanoelec & Prissma project)*

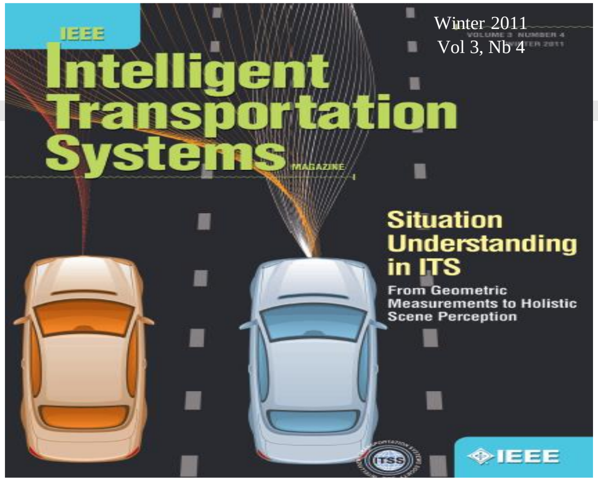

Testing setup *(Inria & IRT Nano)*


Augmented Reality approach *(Inria & IRT Nano)*


Validation of Probabilistic Algorithms using SMC & Formal methods *(Inria & IRT Nano & Prissma)*







Thank You for attending my keynote speech

Any Questions or comments ?

=> Please use IROS 2021 conference system