Multimodal data acquisition prototype for autonomous driving in adverse weather conditions

Gerard DeMAS-GIMÉNEZ(1), Adrià SUBIRANA(1), Pablo GARCÍA-GÓMEZ(2), Eduardo BERNAL(1), Jordi RIU(2), Josep Ramon CASAS(3), Santiago ROYO(1,2)

1. Centre de Desenvolupament de Sensors, Instrumentació i Sistemes (CD6-UPC), Universitat Politècnica de Catalunya, 08222 Terrassa, Spain.
2. Beamagine S.L. Carrer de Bellesguard 16, 08755 Castellbisbal, Spain
3. Image Processing Group, TSC Department, Universitat Politècnica de Catalunya (UPC), 08034 Barcelona, Spain.

Contact: Gerard DeMAS-GIMÉNEZ (gerard.de.mas@upc.edu).

ABSTRACT:

Autonomous driving has made significant progress in recent years, but adverse weather conditions still remain a major challenge for perception and decision-making algorithms. This work presents a multimodal data acquisition prototype designed to enhance autonomous vehicle perception in challenging environments such as fog, heavy rain, and snow. The system is mounted on a Dacia Duster and features a diverse sensor suite including visible, Short-Wave Infra-Red, thermal, and polarimetric cameras with a solid-state LiDAR and automotive radars. The whole system is calibrated with dedicated calibration boards developed on purpose to enable data fusion across all modalities. Initial results demonstrate the effectiveness of our approach, showing that critical environmental features can be detected reliably across various weather conditions. Future work includes the release of a fully annotated multimodal dataset to support further research in adverse weather perception.

Keywords: autonomous driving, multimodal imaging, data fusion, LiDAR, SWIR, thermal imaging, artificial intelligence, data set, adverse weather conditions

1.- Introduction

Autonomous driving has emerged as one of the most transformative areas in modern transportation, promising to redefine mobility and enhance road safety. The development of autonomous vehicles (AVs) relies on advanced sensing and perception systems that allow the vehicle to understand its environment in real-time. As the demand for safer, more efficient transportation grows, the autonomous driving industry has become a highly competitive sector, attracting significant investments and leading to diverse approaches in vehicle design, data acquisition, and perception technologies. Companies across the world are developing different strategies to address the unique challenges of autonomous navigation, from environmental perception to decision-making and control to get an edge over their competitors.

A crucial element in these advancements is the data acquisition systems employed by AVs, as they are in charge to feed the machine learning models that power autonomous driving. While some companies focus on integrating multiple sensor modalities, others have opted for more specialized approaches, focusing on specific data types to optimize performance.

One clear example of a specialized approach is Tesla, Inc.'s vision-centered approach. Their sensor suite consists on having multiple cameras to get a 360º field of view (FOV) around the vehicle. Tesla, Inc.'s "Autopilot" system detects objects and estimates depth with only images and sophisticated neural networks. While this system maintains reduced costs, as visible cameras are generally cheaper than other sensors like Light Detection And Ranging (LiDAR), it compromises safety and robustness in certain environmental conditions where visible cameras might fail, like fog or low light.

In contrast, companies like Waymo, LLC. have embraced a multimodal approach, integrating LiDAR, radar, and visible cameras to enhance perception accuracy and robustness. For instance, low-light scenarios will not present as much problems with this configuration as LiDAR and radars, being active sensors, do not rely on external illumination. However, low-visibility scenarios such as fog, heavy rain, and snow will still compromise the safety of the ride.

While rare in some environments, these adverse weather conditions present a big problem to tackle in order to obtain full autonomous driving, not to mention countries where such conditions are frequent.

This paper explores the development of a multimodal autonomous driving prototype designed to produce a rich dataset for research and development of novel autonomous driving perception algorithms under adverse weather conditions. By integrating unconventional sensors, such as thermal and polarimetric cameras or solid-state LiDAR, alongside more traditional modalities like radar and visible cameras, the prototype aims to create a more diverse and accurate dataset for testing and enhancing autonomous driving systems. Through the careful collection and fusion calibration of such multimodal data, this approach has the potential to offer new insights into the performance and reliability of autonomous systems under any weather condition.

2.- State-of-the-art Data sets for autonomous driving

2.1 KITTI

The KITTI[1] dataset is one of the most widely known benchmarks in autonomous vehicle research. It is considered the pioneer of modern autonomous driving datasets as it was the first to include 3D laser scanner and 3D annotations in real live scenarios. They equipped a Volkswagen Passat (B6) with: two color and two grayscale visible cameras with a big horizontal FOV, a 64-beam spinning LiDAR with a maximum range of 100m, and a GPS/IMU localization system. The set up has been calibrated to enable data fusion. The data set includes good-weather, urban scenes.

2.2 Waymo Open

Released in 2019, Waymo-Open[2] is Google's approach to autonomous driving. Waymo-Open is presented as the largest and most diverse autonomous driving dataset. Figure 1 shows the sensor layout and coordinate system of Waymo's data acquisition prototype. Their prototype counts with: five rotating LiDAR sensors, five high resolution pinhole cameras. The sensors are precisely calibrated and synchronized to enable data fusion and capture the vehicle's surrounding environment. The data is captured at different times and in different weather conditions, even though sunny scenes are the most common and visibility is always clear, mostly in urban scenes.

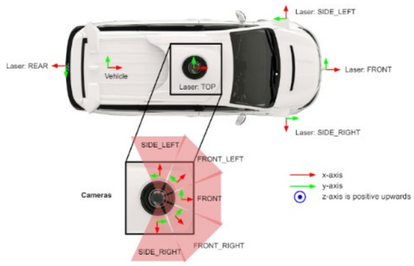

Fig.1: Waymo's data acquisition prototype for Waymo Open[2].

2.3 NuScenes

Introduced by Motional in 2019, nuScenes[3] it is largely considered the reference dataset for autonomous driving. The data was collected in Boston (USA) and Singapore, offering challenging scenes with complex traffic dynamics and dense urban structures. The data is collected at different times of the day (day and night) and weather conditions (sun, light rain, and clouds). However, there is no sign of adverse weather conditions such as heavy rain, fog, or snow. NuScenes uses an extensive sensor suite mounted on a Renault Zoe custom data acquisition prototype as shown in Fig 2. The vehicle is equipped with: six RGB visible cameras that cover 360º of FOV, one 32 beam spinning LiDAR, five frequency modulated continuous wave

(FMCW) radars, and one GPS/IMU system for localization. All sensors are synchronized and calibrated to enable data fusion.

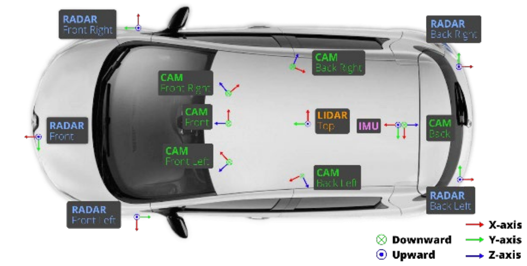

Fig. 2: NuScenes'[3] data acquisition prototype.

2.4 Seeing Through Fog (STF)

Previous data sets only focus in good environmental conditions. Even nuScenes'[3] and Waymo's[2] rainy scenes have fairly good visibility. As a result, existing autonomous systems perform well under normal imaging conditions but fail under adverse weather due to the bias toward clear scenes. Mercedes-Benz alongside Ulm and Princeton universities tackled the issue presenting the first multimodal data set under adverse weather conditions [4]. To overcome the challenges of adverse weather conditions, they presented a sensor suit sensible to the visible, mm-wave, Near Infra-Red (NIR), and Far Infra-Red (FIR) band. They used: two visible stereo cameras, a gated NIR camera, a FMCW radar, a 64-beam spinning LiDAR, a thermal camera, and an IMU. The data set includes urban and highway scenes of northern Europe.

3.- Design and Sensor Suite

As we have just seen, there is no publicly available data set except for STF[4] that has a wide range of sensors and is tested under harsh environmental conditions. Therefore, this work intends to fill the gap in the literature by proposing developing a design for robust multimodal autonomous driving under adverse weather conditions that complements the work done by STF[4]. Our proposed sensor suite can be found in Fig 3. We equipped a Dacia Duster with the following sensors: three RGB high-resolution visible cameras, one short-wave infra-red (SWIR) camera, one thermal camera, one polarimetric camera, two radars, one solid-state LiDAR, and a GNSS/INS system.

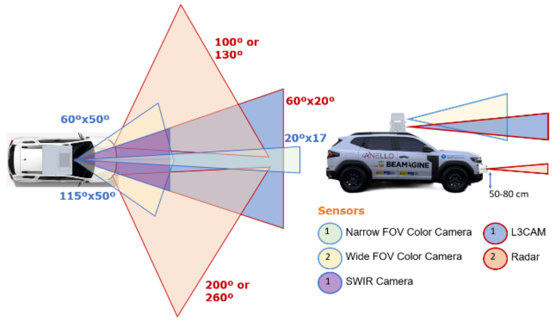

Fig. 3: Design of the developed data acquisition prototype.

Fig. 4: Cameras and LiDAR assembled on the vehicle.

The RGB cameras are all Allied Vision Alvium G1-319c. Two of them have a wide FOV of 60º by 50º that give a combined 115º of horizontal FOV of the vehicle. They have part of their FOVs overlapped to enable stereo imaging, if desired. The other RGB camera has a narrow FOV of 20º by 17º that focuses on small far objects in the frontal direction of the car. The SWIR camera is a Basler Ace 2X a2A640-240gmSWIR sensible in the range from 400nm to 1700nm. This region has a remarkable penetration through turbid media such as fog or smoke. The camera is able to see the lights of cars in dense fog environments before than visible cameras. Both radars are Smartmicro sensor UMRR-96 Type 153 Automotive, they operate in the range of 76-81GHz and have a nominal range of up to 120m with 130º of FOV. This model aims at short to medium range and very wide horizontal angular coverage, providing short, medium- and long-range modes. It is almost unaffected by bad weather. The thermal camera is a Seek Thermal C304SP that operates in the range from 7.8µm to 1400µm with a FOV of 56º x 42º. The polarimetric camera is a LUCID Vision Labs Phoenix™ PHX050S1-QC. This camera gives RGB information as well as linear polarization intensity for 0º, 45º, 90º, and 135º. The last two cameras come together with a solid-state MEMS LiDAR with 60º by 40º of FOV produced by Beamagine S.L. in a multimodal sensor suite called L3CAM. Solid state LiDARs have more resolution and range than their spinning counterparts. However, they cannot have a 360º of horizontal FOV. Thanks to this sensor we can also study how a higher point cloud resolution affects feature extraction and detection in state-of-the-art 3D object detection neural networks. Additionally, the prototype also counts with an Anello EVK, a GNSS unit used for localization and odometry.

Table 1: Failure mode comparison of our prototype.

	Sun	Night	Rain	Snow	Fog
LIDAR	✓	✓	~	~	~
RADAR	✓	✓	✓	✓	✓
RGB	✓	✕	~	~	✕
POL	✓	✕	~	~	~
LWIR	~	✓	~	~	✕
SWIR	✓	~	✓	✓	✓

While our system has considerably more imaging modes than other data acquisition prototypes, our FOV is restricted to just the frontal part of the vehicle. Nevertheless, there is a FOV of 60º by 20º where all modalities overlap. This gives the system a lot of redundancy and robustness against adverse weather conditions. For example, if the prototype is driving in low-light conditions, the visible, polarimetric, and SWIR camera will not produce good data, yet, we can still use the thermal camera combined with radars and the solid-state LiDAR to identify the environment well enough. Table 1 shows a comparison of the performance of each sensor in our sensor suite. It can be seen that there is always at least a 2D sensor and a 3D sensor in good performance regime independently of the weather condition.

4.- Data Fusion

In order to fully exploit the advantages of our multimodal configuration we have to be able to perform data fusion among the imaging modes. Data fusion requires two key components to be implemented: temporal synchronization and spatial registration.

For temporal synchronization, we let each sensor record data at their maximum frame rate and temporally pair the data with our main sensor, the solid-state LiDAR, via best effort synchronization. With that strategy, we achieve temporal differences of less than 50ms between every sensor with respect to the main sensor per frame. The solid-state LiDAR used has a frame rate of 7Hz.

On the other hand, spatial registration has been carried out using [5]. Common reference features have to be found in every mode with the solid-state LiDAR. This is not trivial as each sensor has its own sensitivity range and they may not share it with other sensors in the system. Therefore, we designed two different calibration boards. One for visible, thermal, and LiDAR calibration and another for visible, radar, and LiDAR calibration.

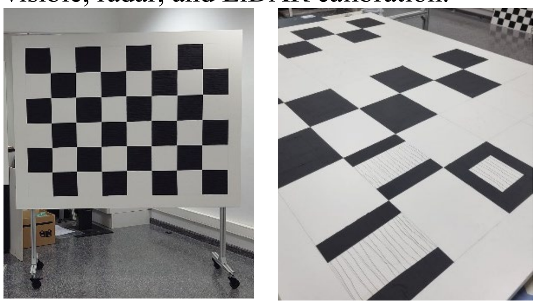

Fig. 5: Visible, thermal, LiDAR calibration board. Left: complete board. Right: nichrome wire under the absorbing tape.

The visible-thermal-LiDAR calibration board (Fig.5) consists of a 5 by 7 squarred checkerboard. Each square is made of a tape that greatly absorbs 1064nm light, so no returns from the LiDAR are obtained from them. Behind the tape, there is a nicrom wire that greatly heats up when current passes through it, making the squares visible for the thermal camera. The intersection between the squares can be obtained through classical image processing algorithms. The later board cannot be used to calibrate the radars as the metallic wire used to detect the pattern with the thermal camera interferes with the radar waves. As a result, another board was designed adapting the previous calibration board ideas of [6]. This time, we generate a similar, smaller pattern with the same absorbing tape to be able to detect the board with both, visible cameras and the LiDAR.

The board is made of Styrofoam, which is invisible for radar waves. Behind the board a metallic retroreflector is placed to force a high-intensity return from the radar. The board can be seen in Fig. 6.

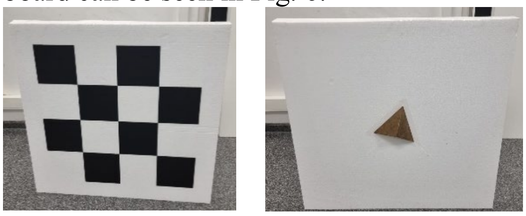

Fig.6: Visible, radar, LiDAR calibration board. Left: frontal part of the board. Right: rear, where the retroreflector is placed.

By taking multiple captures of each board independently we can obtain the intrinsic and extrinsic parameters for every sensor, registering the different imaging modes.

5.- Results

Once the prototype was finally assembled and all sensors were calibrated, we drove it around different environments in search of diverse weather conditions. We also tested how data fusion was consistent even with high dynamic scenes. In Fig. 7 a point cloud projected onto an image can be seen. We can see how the points coincide with the shapes of the objects on the image.

Fig. 7: LiDAR point cloud projected onto an RGB image in a dynamic scene.

Figures 8-10 show captured samples under different weather scenarios. A mosaic of six images is presented. The top row consists of the wide left, narrow, and wide right images respectively. Below, thermal, polarimetric, and SWIR. After that a bird-eye-view of the LiDAR point cloud is presented. Fig. 8 consists on an urban scene with clear vision. Every sensor performs well under these conditions. Differently, Fig. 9 shows a road scene with dense fog. It can be seen how visible cameras are really affected by the condensed water in front of them. The thermal camera also yields poor data as infra-red light is greatly absorbed by water. LiDAR detectivity drops significantly, however, a traffic light visible on the narrow image can still be detected at a high distance. SWIR, on the other hand, still presents a well-defined image with good detectivity of the van's frontal lights. Finally, Fig. 10 presents an urban scene while snowing. Again, visible cameras struggle to produce good results, while the SWIR and the thermal cameras can still operate properly. Especially the thermal camera detecting pedestrians, as they are way hotter than the rest of the snowy environment.

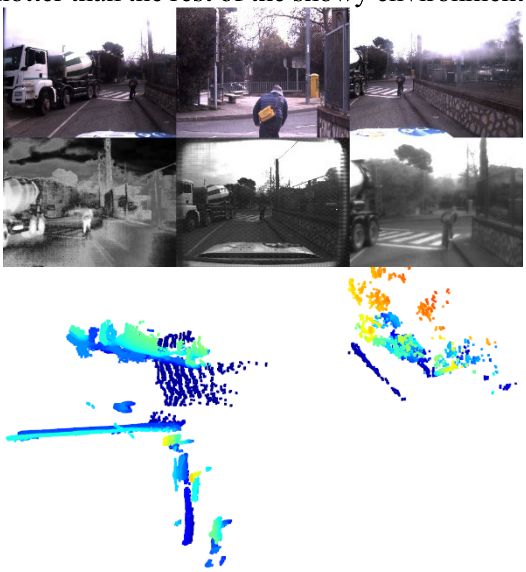

Fig. 8: Urban sample with good visibility.

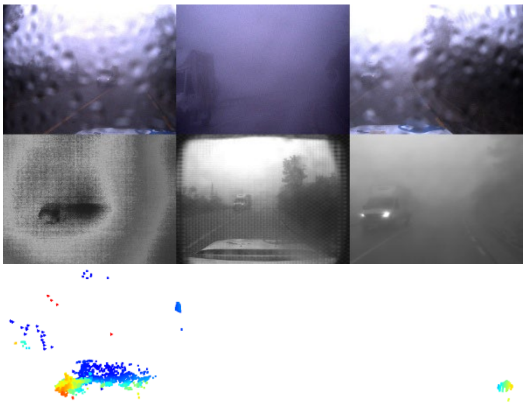

Fig. 9: Road sample with dense fog.

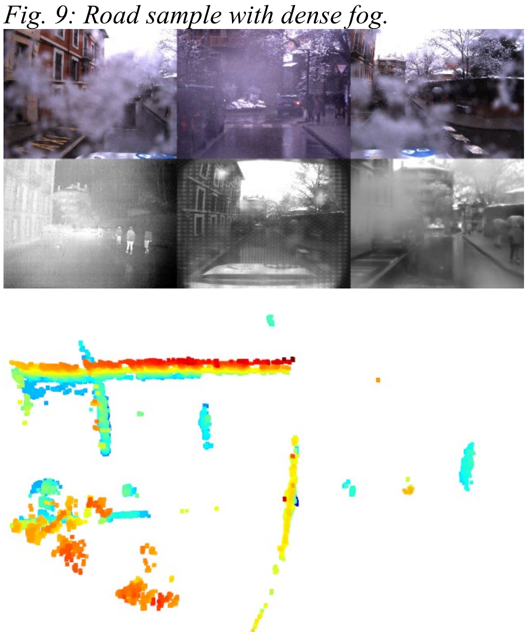

Fig. 10: Urban sample while snowing.

6.- Conclusions and future work

A multimodal data acquisition prototype specially designed to be able to explore perception under adverse weather conditions has been developed and described. Furthermore, two novel multimodal extrinsic calibration boards have been developed to enable data fusion across all sensors. Results show that relevant features from the environment can be obtained regardless of the weather conditions, and that each sensor presents failure modes. As future work, a multimodal annotated data set is currently under development and is expected to be publicly published later this year.

Acknowledgements: This work has been sponsored by the Spanish Ministry of Science through projects TED2021-132338B-I00 and PID2023-152427OB-I00 and by the European commission through project GA-101139941. This work is part of a Ph.D. thesis with grant number 2023 FI-1 00229 co-funded by the European Union.

References

[1] Andreas GEIGER, Phillip LENZ, Raquel URTASUN, *"Are we ready for autonomous driving? The KITTI vision benchmark suite"*, in Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 3354-3361, 2012.

[2] Pei SUN, Henrik KRETZSCHMAR, Xerxes DOTIWALLA, Aurélien CHOUARD, Vijaysai PATNAIK, Paul TSUI, *"Scalability in Perception for Autonomous Driving: Waymo Open Dataset"* in Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2443-2451, 2020.

[3] Holger CAESAR, Varun BANKITI, Alex H. LANG, Sourabh VORA, Venice Erin LIONG, Qiang XU, *"nuScenes: A Multimodal Dataset for Autonomous Driving"*, in Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11618-11628, 2020.

[4] Mario BIJELIC, Tobias GRUBER, Fahim MANNAN, Florian KRAUS, Werner RITTER, Klaus DIETMAYER, *"Seeing Through Fog Without Seeing Fog: Deep Multimodal Sensor Fusion in Unseen Adverse Weather"* in Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11679-11689, 2020.

[5] Pablo GARCÍA-GÓMEZ, Santiago ROYO, Noel RODRIGO, Josep R. CASAS, "*Geometric Model and Calibration Method for a Solid-State LiDAR*", Sensors, 20, 10, 2898, 2020.

[6] Joris DOMHOF, Julian F.P. KOOIJ, Dariu M. GAVRILA, "*An Extrinsic Calibration Tool for Radar, Camera, and LiDAR*" in Proceedings of 2019 International Conference on Robotics and Automation pp. 8107-8113, 2019.