[L3.4] APPROVAL PROCEDURE FOR CONTROLLED ENVIRONMENT TEST EQUIPMENT

Procédure d'homologation des moyens d'essais en environnement contrôlé

Main authors: R. REGNIER (LNE), M. BOUDALI (LNE), A. PIPERNO (UTAC), M. GALLOIS (UTAC), L. PINGOT (UTAC), E. CHATEAUROUX (TRANSPOLIS), Y. MENEROUX (IGN), P. DUTHON (CEREMA)

Keywords: AI based vehicles homologation, testing equipment validation, testing targets, testing driving robots, testing IMU & other sensors.

Abstract. Using the requirements and specifications provided by WP1, the deliverable is to provide methods for ensuring the quality and consistency of test facilities meeting the needs of task 3.2. In particular, it is in charge of providing a best practice guide for the validation of AI-based system test facilities and the implementation of representativeness indicators (retained by WP1, e.g. those relating to the level of realism, the measurement of bias introduced by tests, uncertainty measurements, coverage rate, etc.) for physical tests in controlled environments, drawing for example on data from WP4. The aim is to contribute to recommendations for a future approval procedure for this type of test facility. A particular effort will be made to ensure that these indicators are consistent with those developed in WP2, drawing on the work carried out by the PFA on this subject. Indeed, the outputs of WP3 will also be used as a basis for indicators of the representativeness of simulation tests.

Résumé. En utilisant les exigences et spécifications fournit par le WP1, le livrable doit fournir les méthodes pour s'assurer de la qualité et de la cohérence des moyens d'essais répondants au besoin de la tâche 3.2. IL est notamment en charge de fournir un guide de bonnes pratiques dans le cadre de validation de moyens d'essais des systèmes à base d'IA et de la mise en œuvre des indicateurs de représentativité (retenus par le WP1, par exemple ceux relatifs au niveau de réalisme, à la mesure de biais introduit par les tests, aux mesures d'incertitudes, au taux de couverture, etc.) des tests physiques en environnement contrôlé en s'appuyant par exemple sur des

données venant du WP4. Ceci afin de contribuer à des recommandations pour une future procédure d'homologation de ce type de moyens d'essais. Un effort particulier sera apporté sur la cohérence de ces indicateurs par rapport à ceux développés dans le WP2 en s'appuyant sur les travaux développés par la PFA sur ce sujet. En effet les sorties du WP3 seront aussi utilisées comme support pour les indicateurs de représentativité des essais en simulation.

Table des matières

1 INTRODUCTION : PURPOSE OF THE DOCUMENT

Test Tracks and Closed-Course Testing (using controlled environments such as test tracks or closed courses to assess the vehicle's capabilities in different scenarios like adverse weather, traffic situations, and emergency maneuvers) and HIL/VLl simulation are an important pillar of the evaluation of an automated vehicle and its different components.
However, to be certain that these tests make sense, we must also ensure that the test means we use are of quality. Therefore, the tools and equipment for evaluating AVs also need to be qualified.
In this deliverable, we will see how to put into practice the recommendations of the PRISSMA deliverable 1.6 for the controlled environment testing and it will complement the first chapter of deliverable 2.7 on the validation of simulation test means.

2 QUALIFICATION OF PHYSICAL TEST EQUIPEMENT

2.1 Test track quality

2.1.1. Safety management

All proving ground companies that operate tracks for ADAS and Automated driving system tests shall have a performant safety management system.
Each proving ground has its own functioning and safety management procedures.
The basic rules and requirements are the following:

- A control centre gives the access and controls the activities on each track of the proving ground.
- All vehicles and persons accessing the tracks shall be authorized to do so by the control centre.
- All activities and tests shall have been authorized by the control centre.
- Usually, a prevention plan shall be established by the proving ground with all involved parties. It shall, at least, include:
  - The access and behaviour rules on the tracks,
  - Safety information such as the evacuation plan, the emergency phone numbers and rules…
  - A description of all activities, equipment, and the test plans that will be done on the tracks and workshops,
  - Risk analyses for all these activities,
  - The countermeasures to be implemented,
  - Safety and priority rules if several tests are carried out on the same track,
  - The personal protection equipment required on tracks,
  - The dates of the safety briefings and the names and signatures of all trained and authorized persons.
- According to the complexity and the dangerousness of the tests, the pilot and staff accreditations shall be produced for the proving ground control to authorize the tests.
- Infrastructures and road conditions shall be checked, cleaned and maintained on a regular basis by the proving ground.

2.1.2. Quality management

There is no standard defining the quality and features of a proving ground to test ADS or AI bricks.
Usually, proving grounds have type-approval activities or belongs to vehicle manufacturers which run development, validation and type-approval tests on their tracks.
The Regulation (EU) 2018/858 [1] defines the whole vehicle type approval rules for M, N and O vehicle categories. This regulation also stipulates the rules for the authorities to designate technical services.
Technical services are laboratories in charge of the tests, audits and verifications for the type approvals of vehicles and vehicles components. They are designated and monitored by each Member State in Europe.
[1] Annex III – Appendix 1 stipulates that for activities related to testing for type-approval, two technical services categories exist:

- Category A (tests performed in own facilities): Standard EN ISO/IEC 17025:2005 on the general requirements for the competence of testing and calibration laboratories.
- Category B (supervision of tests, which includes test preparation, where such tests are performed at the manufacturer's facilities or at the facilities of a third party): Standard EN ISO/IEC 17020:2012 on the general criteria for the operation of various types of bodies performing inspection.

The standard EN ISO/IEC 17025:2017 is then the quality standard to apply to show the ability for the proving ground and its staff. The EN ISO/IEC 17025 accreditation shall be passed for each regulation or each protocol.
The track requirements are included in the regulations or protocols such as the Euro NCAP protocols.

2.1.3. Infrastructures

For official protocols and type approval tests, reproducibility is a very important criterion.
It means that the tracks where these tests are carried out shall comply with the regulation or the protocol requirements.
Three track characteristics are usually defined in regulations:

- Geometry / topology,
- Surface condition,
- Grip.

Euro NCAP AEB car-to-car protocol [2] stipulates in paragraph 7 that tests shall be conducted on a dry (no visible moisture on the surface), uniform, solid-paved surface with a consistent slope between level and 1 percent. The UNECE R152 gives the same requirement about the track topology in paragraph 6: The test surface has a consistent slope between level and 1 per cent. This specification applies for the longitudinal and lateral slopes of the track, making this configuration non representative of a real road. Usually, a normal road transversal profile has a slope to evacuate rainwater. The percentage of the cross section depends on road radius and the traffic speed as presented in the SETRA guide [3].
When building a new track, its geometry shall be specified, and validation measurements shall be processed before track commissioning.
Track topology shall be measured by a surveyor. In France, according to the law n° 46-942 of May 7th, 1946, surveyors shall be registered to the surveyor's order to be allowed to practice.

The order imposes to its members to work according to the state of the arts with a qualified staff.

Surface condition is another requirement, Euro NCAP protocols stipulate that the tests shall be conducted on a dry (no visible moisture on the surface), uniform, solid-paved surface, and UN R152 only requires a dry concrete or asphalt surface affording good adhesion. No specification is given about the surface color.

The surface grip can be evaluated by several metrics such as the PBC, the friction coefficient, roughness, etc.
The regulation R152 [4] and Euro NCAP [5] protocols specify a PBC higher than 0.9. The PBC shall be measured by an accredited company. It measures of tire to road surface friction based on the maximum deceleration of a rolling tire, measured using the American Society for Testing and Materials (ASTM) E1136-10 (2010) standard reference test tire, in accordance with ASTM Method E 1337-90 (reapproved 1996), at a speed of 64.4km/h, without water delivery. Alternatively, another method is specified in UNECE R13-H [6] Annex 6 – Appendix 2. This method determines coefficient of adhesion (k).

When building a new track, the proving ground shall stipulate the requirements and validation methods to obtain a surface fit for the regulation or at least representative of usual roads. A French state of the art about road grip is given in a CEREMA document [5].

For other tests that does not require reproducibility, the proving ground company can design its tracks according to its own specifications. Track characteristics shall be documented by the proving ground.

2.1.4. Signs, Road markings and road equipment

Road signs varies across countries. The regulation UE 2021/1985 [6] specifies a list of road signs to be detected by intelligent speed assistance systems for their type-approval. Theses signs only refers to speed limits. Each country, even in the European union, has its own legislation and definition of road signs.
In France, signage is defined by an official document named “instruction interministérielle sur la signalisation routière du 22 octobre 1963” [7]. This document is composed of 9 parts:

- Part 1: Generalities,
- Part 2: Danger signage
- Part 3: Intersections and priorities
- Part 4: Prescriptive signage
- Part 5: Signage for direction, services, and identification
- Part 6: Permanent Traffic Lights
- Part 7: Road markings
- Part 8: Temporary signage
- Part 9: Dynamic signage

Moreover, standards apply to road signs such as NF P98-522 [8], or CE 1826-CPR-04-PAN7 so as to certify these products for the French or the European market.

According to the test objective, the proving ground shall be able to provide signs. Signs can wear with time, and it may be interesting to fine some signs in various states on a track.

Road markings also varies across countries. The UN R130 [9] about the type-approval for lane departure warning systems gives a list of lane marking types in Annex 3. These markings are only longitudinal lines.
In France, road markings are also called horizontal signage. Their shapes and positions are specified in [7] part 7.
Moreover, road marking paint characteristics shall be chosen according to the type of road the track should represent. Paints are certified according to the standard NF P98-691 that defines the following criteria:

- Durability Pr: Expressed in the number of wheel passes on the marking.
    - The durability scale can range from 50,000 to 2 million wheel passes.
- Daytime visibility (Qd): Expressed by the luminance coefficient under diffuse lighting or Qd
    - For yellow products $\geqslant$ 80 mcd/m^2/lx
    - For white retroreflective products $\geqslant$ 100 mcd/m^2/lx
    - For non-retroreflective white products $\geqslant$ 130 mcd/m^2/lx
- Night-time visibility (RL) : Characterized by the level of retroreflection.
    - The minimum value for permanent products is set at 150 mcd/m^2/lx
    - The minimum value for temporary products is set at 200 mcd/m^2/lx
- Adhesion/grip (SRT) Expressed by the coefficient of slip resistance or SRT.
    - The minimum value is set at 0.45 for all roads.
- Temporary durability (T) : Expressed in the number of wheel passes on the marking.
- The durability scale can range from 50,000 to 100,000 wheel passes.
- Night-time visibility in rainy conditions (Rw or Rr) : Characterized by the level of retroreflection in wet conditions (Rw) and in rainy conditions (Rr).
    - The value is set at a minimum of 35 mcd Rw and Rr.

All these parameters were specified with the purpose of adapting the paintings quality and visibility to the uses on the different kinds of roads. Moreover, the markings wear with the weather, sunlight, and wheel passages. The physical conditions of “old” markings shall be documented by the proving ground. However, no metrics are officially defined to measure the wear status of a marking.
In general, the impact of the characteristics of the markings and their wear conditions on the ability of the vehicle sensors to detect the markings is still undefined and will requires further work [10].

The proving grounds shall document the type of paintings and markings installed on its tracks, as well as their wear or conditions.

Common pieces of road equipment shall be found on tracks such as road barriers, bollards, posts, traffic lights or Variable Message Signs. They are not required in official protocols. However, it seams important to validate the ability of automated vehicles to detect or even to classify these elements. So, they should be representative of real items that can be found on real roads and proving grounds shall equip their tracks with components certified according to national or international standards.

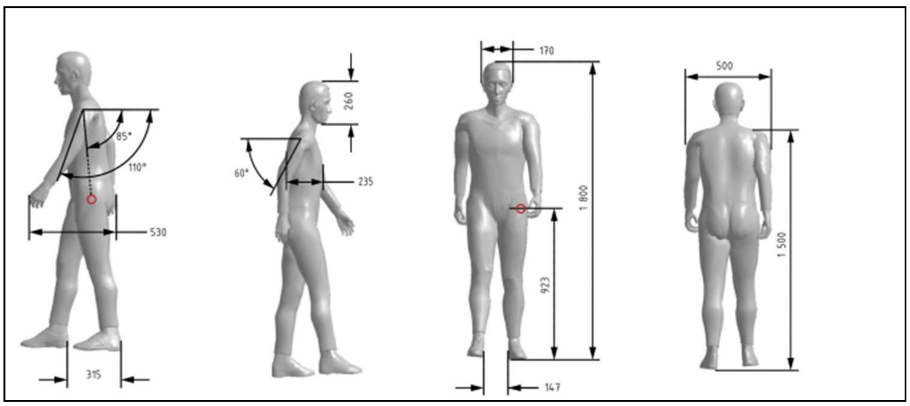
2.2 Test equipment deployed on the track

2.2.1. Road target (VRU, …)

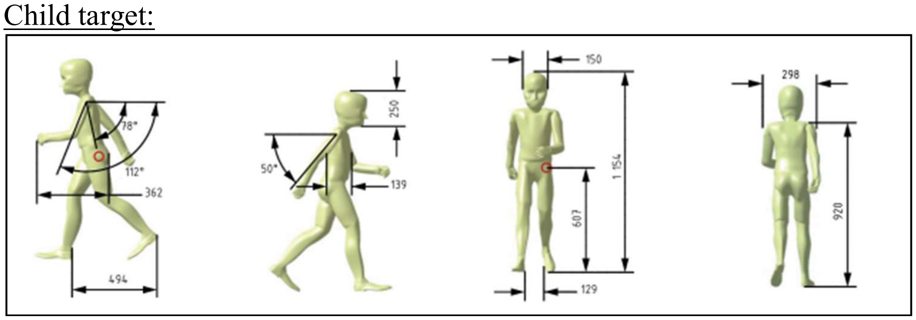
Targets are among most important and expensive equipment for automated and intelligent systems testing. They are many and have much influence on the test result.

The classical used targets for the previous scenario are those defined by the **ISO 19206-2_2018** (Pedestrian) and the **ISO 19206-4_2020** (Bicycle). For the Robustness scenarios, the targets will be adapted.

ISO 19206-2_2018: Adult target:



Child target:

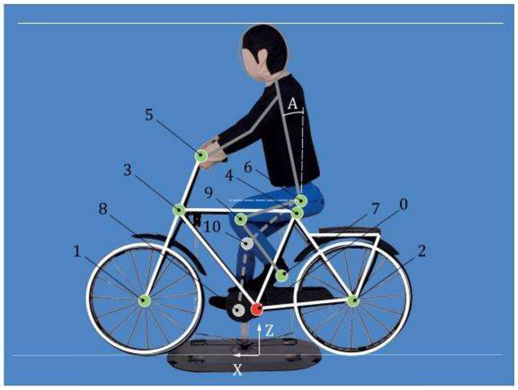


ISO 19206-4_2020: Bicycle target:

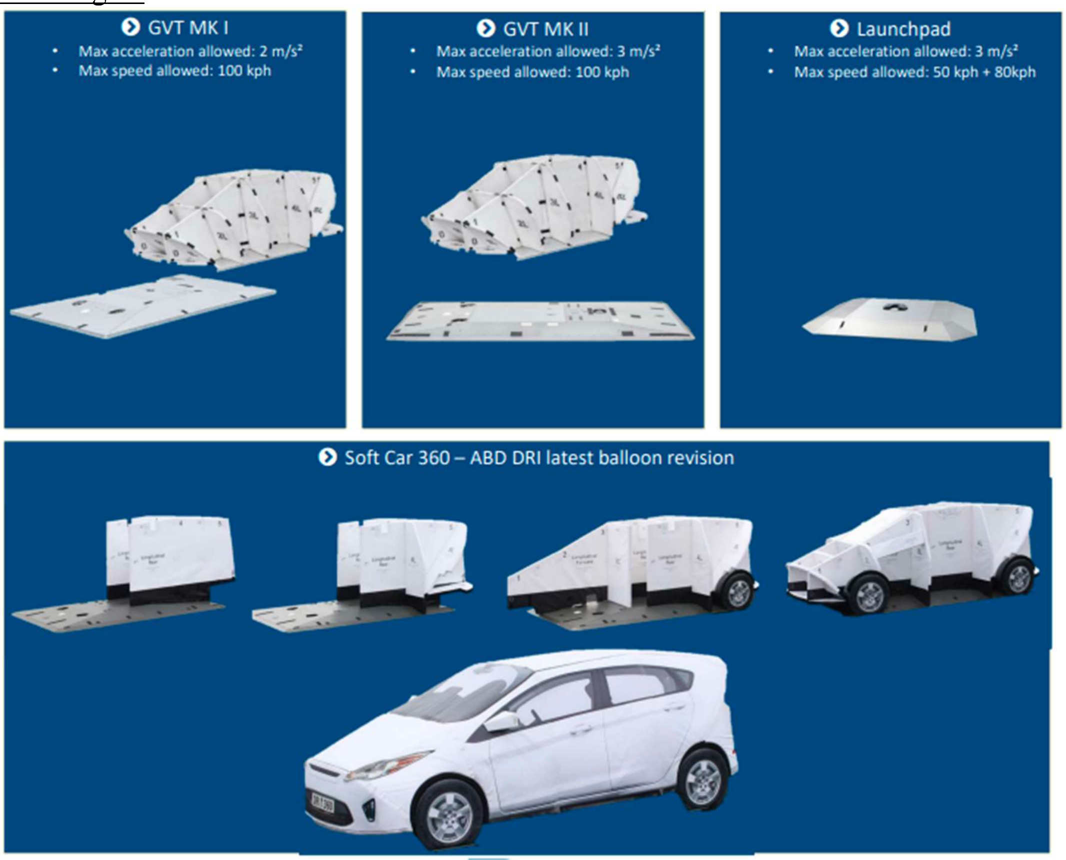
Segment	X	Z	Tolerance	Unit
0 Centre of bottom bracket of BT bicycle	0	280	±10	mm
1 Centre axis front wheel	670	340	±10	mm
2 Centre axis rear wheel	-540	340	±10	mm
3 Front top frame	430	855	±10	mm
4 Rear top frame (upper range sloped top tube)	-215	860	±10	mm
4 Rear top frame (lower range sloped top tube)	-145	460	±10	mm
5 Handlebar	310	1 180	±10	mm
6 Saddle	-235	935	±10	mm
7 Lower edge left foot[a]	105	495	±20	mm
8 Lower edge right foot	80	200	±20	mm
9 Knee point, left [b]	150	860	±20	mm
10 Knee point, right	85	700	±20	mm
Total height (for 10° torso angle)	1 865		±20	mm
Total length	1 890		±20	mm
A Torso angle	10 and 30		±2	°

[a] Lowest point of shoe – centre line tibia.

[b] Knee point: rotation point of knee.


5
A
3
6
4
9
8
10
7
0
1
2
Z
X

Car Target :


GVT MK I
Max acceleration allowed: 2 m/s²
Max speed allowed: 100 kph
GVT MK II
Max acceleration allowed: 3 m/s²
Max speed allowed: 100 kph
Launchpad
Max acceleration allowed: 3 m/s²
Max speed allowed: 50 kph + 80kph
Soft Car 360 – ABD DRI latest balloon revision

PROPULSION SYSTEMS FOR TARGETS

The propulsion systems used are in accordance with the TB029 of ENCAP.

				Euro NCAP Test Targets			
				Global Vehicle Target (GVT)	Euro NCAP Pedestrian Target Adult (EPTa)	Euro NCAP Pedestrian Target Child (EPTc)	Euro NCAP Bicyclist Target (EBT)
			Supplier	ABD	4a	4a	4a
			Product	Soft Car 360	4activePA Adult	4activePA Child	4activeBS
			Version	DRI Rev G Feb 2020	v4v4	v3v3	v5v5
Supplier	**Product**	**Version**					
ABD	GST100	V1.0 (P8503) & (P8328) with car panel					
ABD	GST120	V1.0 (P12218)					
ABD	SPT System	SPT20/SPT20s					
ABD	LaunchPad 50 & Launchpad 60	V 1.0 (P9226) without extension					
		V 1.0 (P9226) with extension					
ABD	LaunchPad 80	V 1.0 (P11000) without extension					
		V 1.0 (P11000) with extension					

Propulsion Systems

A testing center must be an ISO 17025 certified testing Laboratory. At UTAC for example, more than 30 ADAS & AD tests are certified like ABS, ESP, LKA, LCA, LDW, breaking, steering, lighting tests… and all the process for that follow this standard.

This is not mandatory but a warranty of quality for tests, because ISO 17025 requirements for quality are much higher than requirements of the two usual quality standards used in testing labs: ISO TS 16947 and ISO 9001.

Audits of verification and accreditation are regular, every 15 months for verification, every 5 years for new accreditation (made by the COFRAC organism).

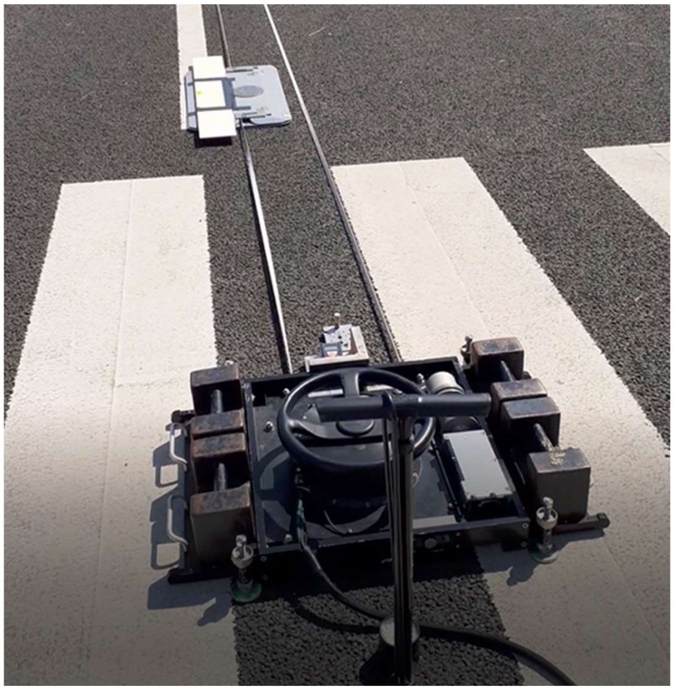
Therefore, all these procedures are confidential so we can only partially describe them and have partial extracts illustrating them.

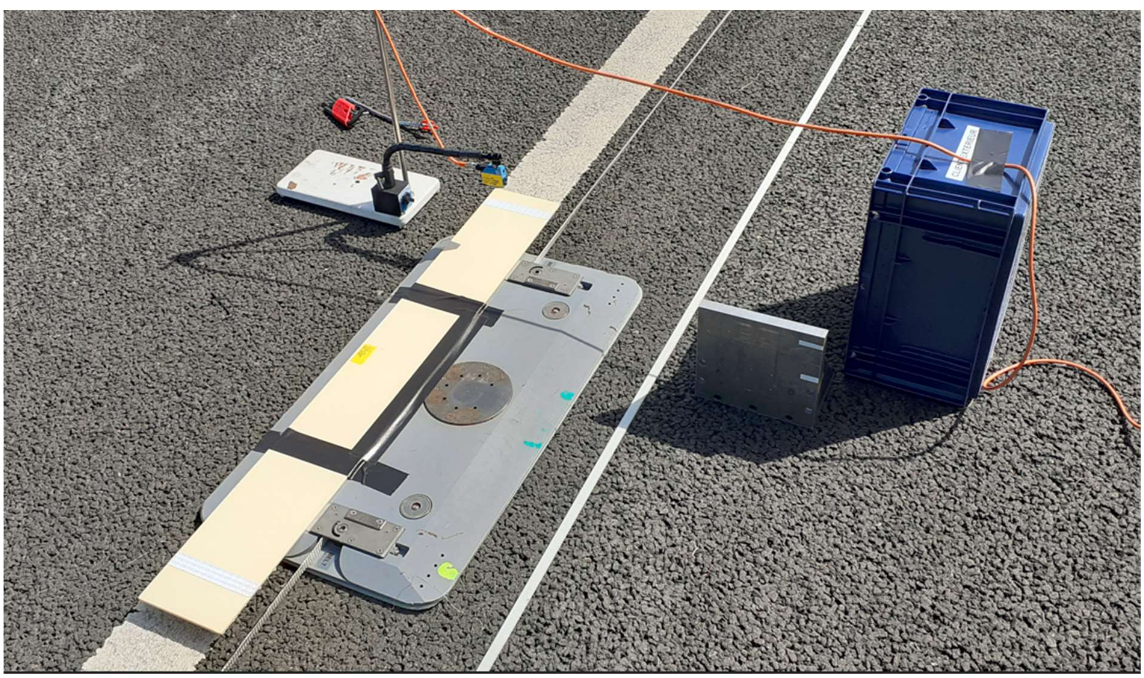
A good practice would be the use of the "5M" method, measuring and maintaining the five items of a test: men, method, material (the vehicle), material (the testing equipment's), and the environment (testing tracks, weather conditions...). Each of these five items is described with a list of variables, which are the most important to measure & calibrate regularly.

It is important an metrology laboratory (intern like in UTAC or extern), for all controls and calibrations.
For example, the UTAC metrology laboratory has more than 100 calibrations procedures and more than 100 pieces of equipment for measure and calibration, partially example in the figure here below:

Grandeurs étendue de mesure	**Etalons mis en œuvre**	**Procédures**	**Exemple d'incertitude d'étalonnage optimum à (k = 2)**
Accélérométrie « 5 g à 500 g 52 à 6000 Hz »	Pot vibrant MET0069 Accéléromètre étalon ACC0868	 Etalonnage des accéléromètres ET.MET.000.051_rev_01	+/- 1.7 %
Electricité – Magnétisme et CEM Uac, Iac, Udc,Idc, R »	Calibrateur multifonction CAL0016 Multimètre étalon MLT0088 Analyseur de réseau – ANA0107	ET.MET.000.019_ Etalonnage de Multimètres ou équivalents_rev_03	+/- 0.01 V +/- 0.02 A
Temps – Fréquence « 1 Hz à 1 MHz »	Fréquencemètre FRQ0004 Chronomètre SEC0024 Générateur de fonctions GEN0020	ET.MET.000.049_Etalonnage des tachymètres_rev_04 ET.MET.000.054_Etalonnage base de temps carte d'acquisition_rev_02	+/- 1 tr/min +/- 0.63 s
Dimensionnelle « 0 – 1 m » « 0 à 360 ° » 	Banc d'étalonnage 1 axe motorisé, résolution 0.001 mm – MET0091 Pied à coulisse DIS0406, DIS0091 Butée micrométrique DIS0164 Inclinomètre NIV0114 Codeur angulaire NIV0115	ET.MET.000.003_ Etalonnage des capteurs de déplacement_rev_04 ET.MET.000.052_ Etalonnage des Capteurs de dépl_angulaire_rev_02 ET.MET.000.004_ Etalonnage des Niveaux et mesures d'angles_rev_03	+/- 0.13 mm +/- 0.46 mm +/- 0.07 °
Vitesse « 0 à 250 km/h »	Fréquencemètre FRQ0004 Drapeau SCU0372	ET.MET.000.043_ Etalonnage de vitesse avec cellule_rev_02 ET.MET.000.053_Etalonnage vitesse banc chute_rev_01	+/- 0.1 m/s

Here below are two photos of pedestrian target propulsion system calibration in UTAC:





2.2.2. V2X equipment

Road side unit (RSU):

This equipment is a component used in C-ITS to provide information to vehicles from the infrastructure. It could be either stand alone to send out DENM messages from road operators. These messages are useful for Safety applications, traffic management and eCall enhancement.

RSUs are also paired with traffic lights to send out information traffic light phase to improve traffic efficiency; It could be also paired with camera to send out information on objects (pedestrian, bike or vehicles) detected.

For example, at UTAC, we have 15 RSUs from Lacroix City installed at 10m height around the tracks to send any messages needed to connected vehicles. Moreover, it is possible to track the position of vehicles that send out C-ITS Messages

On-Board Unit (OBU)

This equipment is a component used in C-ITS to connect a moving vehicle without V2X built-in. The UTAC's OBU (from YoGoKo) is mounted in the vehicle with a GNSS antenna. The functioning of this hardware is identical to the RSUs but smaller to fit vehicles.

Cellular network

On UTAC site, two network providers (Orange & Bouygues Telecom) decided to implement 5G Antenna to cover the tracks with the best connectivity possible, in terms of data rate or latencies.

2.2.3. Sensors on the tracks

This section is composed of three parts: in the first part, the camera-based tracking system designed at IGN is briefly presented, along with its expected objectives, performances and requirements. Then, all the unit tests conducted individually on each module of the system are presented. All these tests are to be conducted again to replicate the experiments with new equipment, and to ensure that the expected accuracy can be reached. Eventually, in a third part, the results of two real-scale experiments designed to assess the accuracy and the capabilities of the fully integrated system, and conducted at Valeo center and UTAC test site respectively, are reported.

2.2.3.1. System overview

2.2.3.1.1. Objectives & requirements

The system is designed to measure an accurate estimation of the trajectory of an autonomous vehicle (AV) driving at moderate speed (15 to 50 km/h), in a small and well-defined area, typically around and next to a roundabout, or any other small-extent road infrastructure (between 50 to 100 m, such as crossing, parking lot or motorway acceleration ramp).

The used approach relies on a camera-based tracking system, with non-permanent targets attached to the vehicle, and ground control points positioned with topometric survey at various locations on the scene to rigorously combine and make the link to the legal reference frame.

This approach has several advantages:
- It is non-intrusive (no permanent equipment on the vehicle) and minimal modification of the environment (which is important for replicating nominal driving conditions for the AV).
- It is completely independent of the AV sensor and navigation systems, which offers an unbiased method for the evaluation or the vehicle capabilities.
- It requires minimal on-site intervention of operators during the data acquisition, which is sometimes of paramount importance for safety reasons.

The expected output is a trajectory estimate, *i.e.* a set of positions and orientations of the vehicle body in a spatial reference frame with a precise absolute timing. A desirable frequency is a few Hz, with an ideal objective of **10 Hz** (corresponding roughly to 1 point every 1.4 m for a vehicle driving at 50 km/h). The required geometric accuracy is about **2 cm** in each axis of a three-dimensional reference frame.

Additional requirements of the system are as follows (see deliverables 1.6. and 3.3. for more details):
- The synchronization of cameras should be performed with a wireless system, in order to make the different poles of cameras physically independent, therefore allowing adapting the system to any required driving scenario (meaning that the geometrical disposition of cameras may be optimized to ensure maximal accuracy for the chosen vehicle path).

- The system should have hard disk capabilities required to register different driving scenarios, for a minimal total of duration of 10 minutes, *i.e.* 6000 full-size raw images per camera.
- The system should be operational under clear weather conditions; despite the fact that nothing prevents its use in rain, or snow conditions (provided that *ad hoc* shelters are set for the cameras and electronic systems), this has not been tested during this project, and in particular, the negative effect of such conditions on optical systems (diffraction, etc.) has not been clearly assessed. Tests have been however successfully performed in very low-luminosity conditions (5 PM on November). For obvious reasons, foggy conditions are prohibited.
- All the operations (topometric survey included) should be realizable with a limited team (4 workers, without staff required to operate the AV under test) in a single day.
- Processing is not required to be done in real time, but should be manageable in a limited amount of time (typically less than one week for a team of 2 people). Considering the large number of images collected (*i.e.* 72 000 images for 12 cameras and only 10 minutes of data acquisition), target detection and global adjustment should be automatized as much as possible.

2.2.3.1.2. Solution

The proposed solution has been thoroughly presented in deliverable 3.3. Its main features are repeated here for convenience.

The system is composed of a (possibly variable) number of poles, each pole being equipped with a maximal number of 3 cameras, connected to a sub-master PC for collecting the images. Each camera is also connected to a common low-cost GNSS receiver, providing the time synchronization (with sub-millisecond capabilities). Cameras are individually calibrated (*i.e.* estimation of focal parameters and distortion map of the lens) 24h to 48h before the experimentation, with *Micmac* IGN homemade photogrammetric software.

A master PC laptop is used to trigger all the cameras, with Wi-Fi connection to keep the benefits of a wireless system.

Since they are autonomous and almost fully independent, the number of poles used for a given experiment may be parametrized accordingly. The nominal configuration is composed of 4 poles, each controlling the maximal allowed number of 3 cameras. Regarding the tests conducted during the project, one test with [1 pole x 3 cameras] and one test with [4 poles x 3 cameras] have been conducted, and results are reported in sections 2.2.3.3.

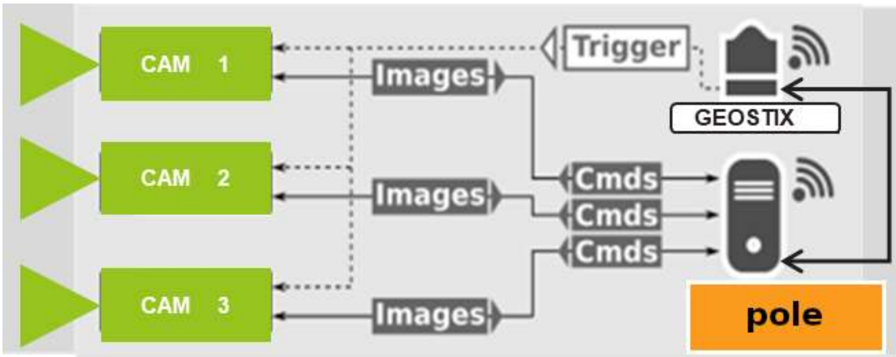


Fig 1. Illustration of 1 acquisition pole with 3 cameras

Once the system of cameras is set up on-site with the chosen configuration, a classical topometric survey is conducted to determine 3D coordinates of a set of points in the scene. These points are of 3 different types:

- Optical center of cameras: it is a non-physical point which can be assumed to be the center on which images are projected to. Its 3D coordinates are determined beforehand, and expressed in a local reference frame, attached to external physical points on the camera sensor.
- Ground control points (GCPs) at various and evenly distributed positions in the scene. GCPs can be further classified into two sub-categories: natural (or opportunistic) points which are already existing in the scene (horizontal and vertical road signs, far antennas, buildings, etc.) and artificial (laser scanner spherical targets, black-and-white targets, cones…) which are specifically positioned in the scene as complementation wherever the former are not sufficient.
- Coded targets on the vehicle body. The nominal number of these targets in the conducted experiments is 15, with a symmetrical disposition (6 on each side of the vehicle, 2 at the rear and 1 on the engine hood. They are designed to allow automatic detection and recognition. The disposition is chosen in order to optimize visibility from cameras (while choosing as much as possible flat surfaces on the vehicle body), to provide redundancy and robustness in the computation steps and to allow for an accurate estimation of the vehicle attitude (especially, heading estimate is better when targets are placed further apart alongside the vehicle main axis).

Note that since the vehicle body is moving in the global reference frame, the coordinates of the first two set of points are provided in an absolute reference frame, while the coordinates of the coded targets on the vehicle body only have interest as relative to each other. This set of relative positions (15 x 3 – 3 translation parameters – 3 rotation angles = 39 degrees of freedom) is called rigid block in the present document, and refers to the set of relative constraints, ensuring tight rigid-body, estimation of all the targets. Since coded targets are meant to be removed on the AV, this rigid block is numerically (i.e. with a roto-translation) tied to fix points on the vehicles (e.g. wheels, permanent paintings on the cars, etc.). This tying is necessary to provide final coordinates of a permanent AV reference frame within the absolute reference frame.

Fig 2. Targets on the vehicle

The absolute localization of the topometric survey is ensured with a network of 4 high-precision GNSS receivers, with accurate positions computed with *Bernese* software. To overcome the (relatively) poor orientation of a short-extent survey area, the absolute orientation is reinforced with a set of gyroscopic measurements.

At the end of this step, accurate (5 mm uncertainty) coordinates of this set of points are determined, and they are subsequently used as input to estimate the initial position and orientation of each camera. Theoretically, with accurate knowledge of the optical center, 2 points are enough to determine the camera orientation. For robustness, and highly precise determination, a dozen points (evenly distributed in the image) are used for each camera orientation. Note that these points need not be shared on several cameras since each camera orientation is performed independently. It is however a desirable property during the global adjustment step, thus enabling a common mutual improvement of GCP positions and cameras orientations.

After trigger from the master PC, all cameras are collecting a sequence of images on a synchronized basis, at the maximum rate of 10 Hz, while the AV is moving in the scene. At the end of the experimentation, automatic detection of coded targets is performed on a per-image basis, thus providing a set of 2D pixel coordinates of targets attached to the AV, with their associated identification names. This set of 2D pixels is then converted into 3D bundles $\boldsymbol{b}_{ij}$ (representing the optical ray issued from a given camera i to a given target j). This is possible, since the distortions of cameras have been modeled beforehand. Note that $\boldsymbol{b}_{ij}$ is a unit vector, insofar as it does not provide information on the distance between camera and target – it only states that, at a certain time step t, target j is in the direction pointed by the vector $\boldsymbol{b}_{ij}$ from the optical center of camera i. Note also that because cameras are not necessarily verticalized, this vector $\boldsymbol{b}_{ij}$ is true up to global 3D rotation $\boldsymbol{R}_i$ for all targets $j=1, 2\ldots\ n$ seen from camera I, whose initial value is provided by the orientation step. Besides, to accommodate slight motions of the cameras during the experiment (heat, mechanical vibrations, etc.), this global rotation is to be estimated at each time step: $\boldsymbol{R}_i\,(t)$.

The set of bundle observations (from all cameras, for all time steps, and towards all detected targets) are then input into *Comp3D* topometric software (IGN), along with initial estimates of camera orientations and GCP coordinates, to perform a global adjustment. Intersections from all bundles pointing towards the same target is computed (in a least-squares approach, *i.e.* with a global minimization of squared residuals of angles). These observations are complemented with the prior constraints resulting from the rigid block between targets on the vehicle. When the number of time steps is large (typically more than a few hundreds), or when the number of false detections of targets cannot be kept reasonably small, global adjustment can also be done on a time step basis (*i.e.* each time step is processed individually), thus preventing numerical divergence of solution and easing tracking of gross errors. Experiments conducted on UTAC test site suggested that this surrogate approach does not decrease significantly the accuracy of estimated trajectory points.

The final output is, for each time step (with absolute timing in UTC or GNSS reference time frame), the 3D coordinate of the center of the AV reference frame in the national reference frame RGF93, along with the 3 rotation parameters, all of them associated with estimation standard deviations (in mm for center along each axis, and in degrees for rotations). This set of points can subsequently be interpolated to obtain a proper trajectory at the desired frequency.

2.2.3.2. Optical systems

2.2.3.2.1. General specifications

Please find here the cameras specifications needed to obtain the technical global objectives given in the deliverable 1.6. We precise the camera type and explain the process used to approve the right agreement of the cameras with our project need.

The cameras specifications are:

- Global shutter to avoid image car distortion in the successive images.
- Use GigE interface to grab images from camera placed more than 50m of the acquisition system. This kind of connection offers the possibility to place the camera 100 m so far of the acquisition system. This GigE interface offers the possibility to use many cameras connected to one processing system.
- As the acquisition of most cameras must be driven by external hardware signal, the cameras must have a hardware trigger.
- The acquisition time is depending of time exposition and time delay triggering; the capability to set these parameters is needed. And the best camera is one equipped with sensor with less time to grab image.

As we plan to use two different places for cameras:
Cameras A: cameras placed around the crossroad center.
Cameras B: cameras placed two by two in each side of the roads leading to the center of the intersection.
To be used rightly, the horizontal fields of view of the two cameras type are:

- ✓ **Cameras A** : > 80°
- ✓ **Cameras B** : > 60°

The ground pixel size for each camera type is:

- ✓ **Cameras A** : 1 cm at 20 m
- ✓ **Cameras B** : 1 cm at 40 m

The cameras Frequency Per Second (FPS) triggering is estimated in function of the car speed and the distance between all the trajectory tops. If we plan a trajectory with 1 top/m at 50 km/h, the FPS must be around 13/14 images/second.

For a car speed around 15km/h the FPS must be around 4 images/second. The right choice of the camera is one with FPS greater than 4 Hz and around 14 Hz. As we need both FPS and sufficient resolution, we select the model BFS-PGE-161S7M-C from FLIR. This model offers a good resolution with an FPS around 7 Hz and 12 Hz (in lossless compression mode). This camera is based on a best resolution sensor IMX542 from Sony Corporation, the general specifications are adapted with our needs.

F*ig 3. IMX542 at left and BFS-PGE-161S7 cameras*

Resolution	5328 (H) x 3040 (V)
Mega Pixels	16.19 MP
Supply Voltage	3.3V, 2.9V, (Analog), 1.1V (Digital), 1.8V (Interface)

Package Type	Ceramic, LGA
Chroma	B/W, Bayer
Shutter Type	Global Shutter
Frame Rate	35 to 52 fps
ADC Resolution	8 / 10 / 12 bits
Pixel Size	2.74 µm x 2.74 µm
Type	Back-illuminated Sensor
Sensitivity	2030 to 2571 Digits

Tab 1. Parameters of camera sensors

To choose the camera lens, a main specification of the lenses is a focal distance capability. As we need two different fields of view we need two types of lenses, one with 8 mm of focal and the second model with 12 mm of focal distance.

Camera A	Focal: 8 mm Distance: 20 m	FOV H: > 85° Ground px. Size: 0.69 cm
Camera B	Focal: 12 mm Distance: 40 m	FOV H: > 85° Ground px. Size: 0.69 cm

Tab 2. Parameters of short-length (8 mm) and long-length (12 mm) focal objective lenses

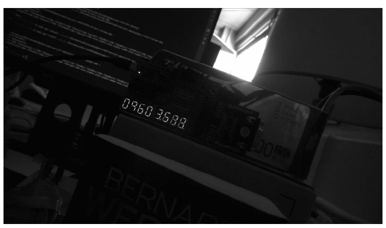
Some manufacturers offer the possibility to use a web selector API to choose the closed model that is adapted to the camera sensor. The selected models are: V0828-MPY2 and V1228-MPY2, the selection process is guided by the selector application offered by the manufacturer (FLIR), but some time is not the best way to select the right Lens for the specific Camera Sensor and for a specific need.

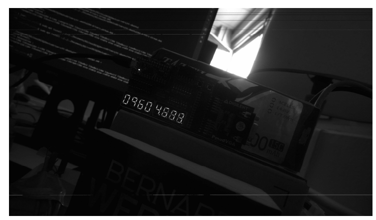
2.2.3.2.2. Camera sensors

From the moment we had established the specifications of the cameras (sensors and lens), a simulation phase allowed us to validate our technical choices on the basis of results consistent with the objectives of the project. To verify that the cameras offer the advertised performance, we planned and carried out tests to check the triggering rate of the cameras, the resolution of the vision system. The experiment is carried out using a pulse generator, a digital-chronometer as well as a VGA screen with a test pattern displayed by triggering. The analysis of the captured images allowed us to verify that the cameras operate at a rate of around 10 Hz, which corresponds to 100ms.

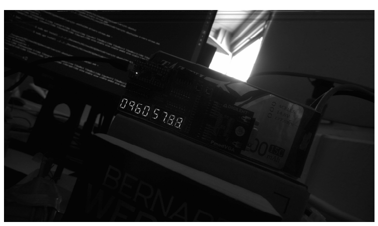
Test with 1 camera

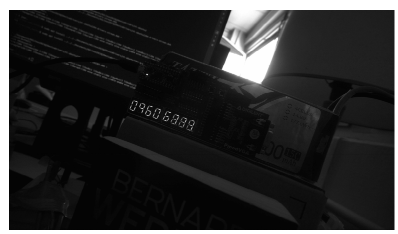
We use a simple method that consists in grabbing images of a digital chronometer triggered at the same frequency as the camera. The displayed sequence of images shows that the delta between the successive images is around ~100ms.





0960,7800-0960,6800 ~ 0,100 s (10 Hz)





0961,0000-0960,8900 ~ 0,100 s (10 Hz)

Fig 4 Sequence of 4 images used to check interval between successive image

c. Camera lenses

Camera resolution

The best way to check the lens resolution specification is to use image pattern or normalized target as the USAF-1951, and for each distance close to the final use we verify the ability to distinguish the smallest element.

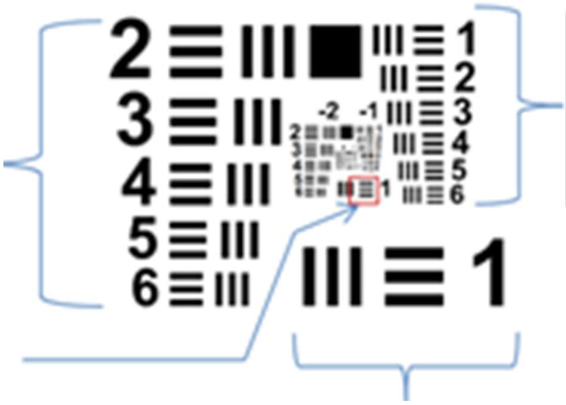


Fig 5. USAF-1951: the lens resolution is better and can resolve 1cm pixel terrain if we are able to see with good definition of the element 1 of the group -2

	• *For machine vision applications* • *8 mm Lens with Fixed focal length* • *12 Megapixel* • *C-Mount* • *Smallest 1.1 "format lens for> 10MP lens in the industry* • *F 2.8* • *3.45µm Pixel size*
	• *For machine vision applications* • *12 mm Lens with Fixed focal length* • *12 Megapixel* • *C-Mount* • *Smallest 1.1 "format lens for> 10MP lens in the industry* • *F 2.8* • *3.45µm Pixel size*

Fig 6. Technical specifications of 8 mm (top) and 12 mm (bottom) lenses.

Calibration of distortions

Calibration is done separately for each pair of sensor/objective, on a textured calibration polygon.

Fig 7. Textured area for calibration (left) and bundle adjustment result (right)

Calibration is performed by taking a set of images with various orientations (to ensure that the calibration is homogenous on the whole image, as illustrated on Fig 8**).**

Typical residual value of calibration is at 0.70 px. This value is highly reproducible between each model of camera (within less than 4%), as shown in Tab 3. Average residual of bundle 3D intersections for calibration**.**

Objective	Sensor	Average residual (px)
12 mm	124	0.72
8 mm	125	0.73
12 mm	126	0.70
12 mm	128	0.70
12 mm	129	0.73
8 mm	130	0.75
12 mm	131	0.71
8 mm	132	0.68
12 mm	133	0.74
12 mm	136	0.69
8 mm	138	0.75
12 mm	149	0.67

Tab 3. Average residual of bundle 3D intersections for calibration

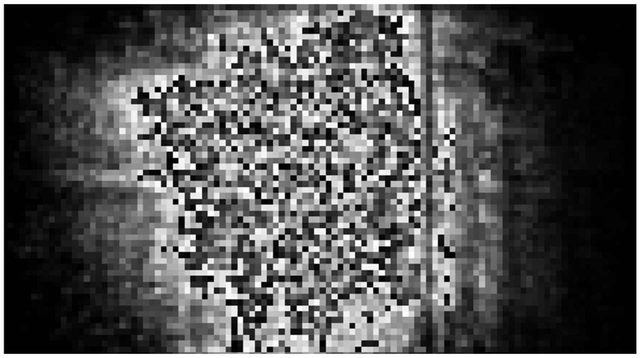


Fig 8. Distribution of tie points in the calibration process. Brighter areas represent higher densities of tie points, therefore better estimation of lens distortions

Tests have been performed to assess the duration of validity of a calibration. This is important, since for practical reasons, calibration of cameras being quite a time-consuming process (~ 12 cameras x 20 min per / camera = 4 hours), it assuredly cannot be done right before the experimentation.

In particular, as illustrated on Fig 9 **and** Fig 10**, calibration seems to be holding for at least a few days, with only a few 1/100th of pixels between on-the-fly calibration and using a pre-calibrated camera.**

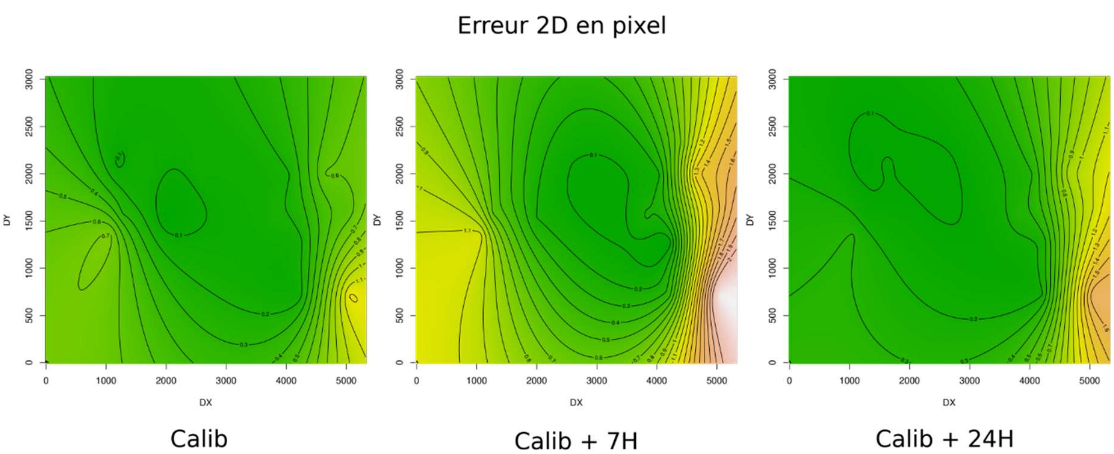


Fig 9. Distortion error just after calibration (left), and a few hours later on the same day (right)

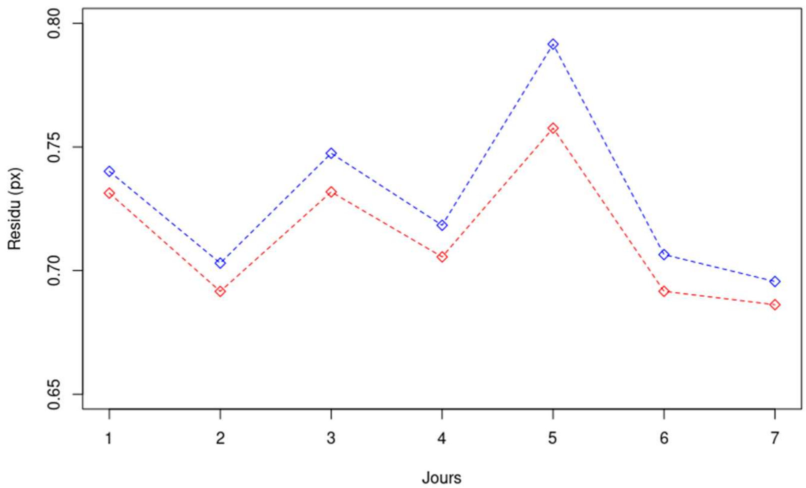


Fig 10. Number of days after calibration vs residual value of bundle adjustment on the same polygon (in px), revealing that calibration of cameras can easily be done up to one week before the experimentation

Fig 10 in particular, shows that the residual error of non-modeled part of lens distortion is not monotonically decreasing with time (at least within one week after calibration), and remains within the typical standard error of all camera models (0.72 +/- 0.02 px).

All these tests lead to conclude that with the chosen model of camera, calibration need not be done on-site right before the experiment POC.

Optical center position estimation

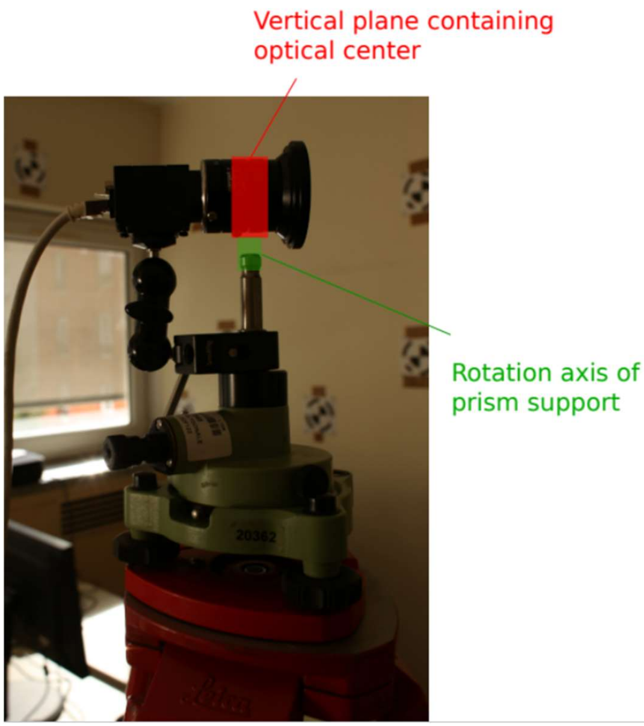


Fig 11. Experimental set up to estimate optical center depth

This is done also for each pair of camera sensor/lens. Optical axes are assumed to be coincidental with the cylindrical symmetry axis of cameras, which leaves only one parameter to be estimated: the depth of optical center position.

Using a calibration polygon composed of 35 coded targets (topometric accuracy ~ 0.1 to 0.2 mm on each axis), disposed on the 4 walls of a square room, camera is placed in the center, and is being rotated, taking picture in stop-and-go. For high accuracy, 100 pictures are captured but experimental comparisons showed that the solution converges with 8 pictures (one in each 45° octant).

Automatic target detection is performed on each frame, and orientation of the camera is performed with photogrammetric spatial resection for each captured image. Of course, this is done with pre-calibrated camera (see section above).

If the optical center is co-located with the device rotation axis, all spatially resected centers coincide in a unique point (up to few 10ths of mm of error). This is of course not the case at first attempt with a new camera. In this case, a circle is fitted on the estimated centers **(**Fig 12**),** and the radius is an estimate of the distance between the optical center and the rotation axis of the device. A second loop of image capture enables to confirm the position of the optical center.

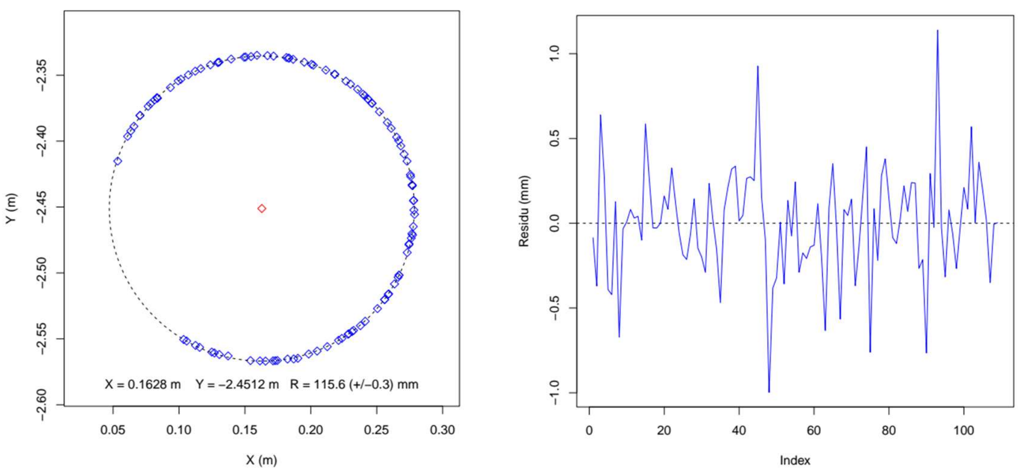


Fig 12. Left: fitted circle on estimated optical center positions (rotation center is depicted in red). Right: residual values along the sequence of 100 images

With this method, an accurate position of the optical center (with about 1 mm standard deviation of error) is estimated in about 5 minutes of experimentation.

Different experimentations revealed that (with the degree of accuracy needed) the position of the optical center does not significantly depend on the sensor aperture (see Fig 13 as an example) but it mainly depends on the focus. Besides, since focus is set for a given aperture, it is important to consider that calibration of optical center is done for a given pair of aperture/focus values.

It was tested also that optical center position does not depend on objective models either (Fig 14). However, another test clearly revealed (as might have been expected) that the depth of optical centers depends on the focus of camera (Fig 13, right).

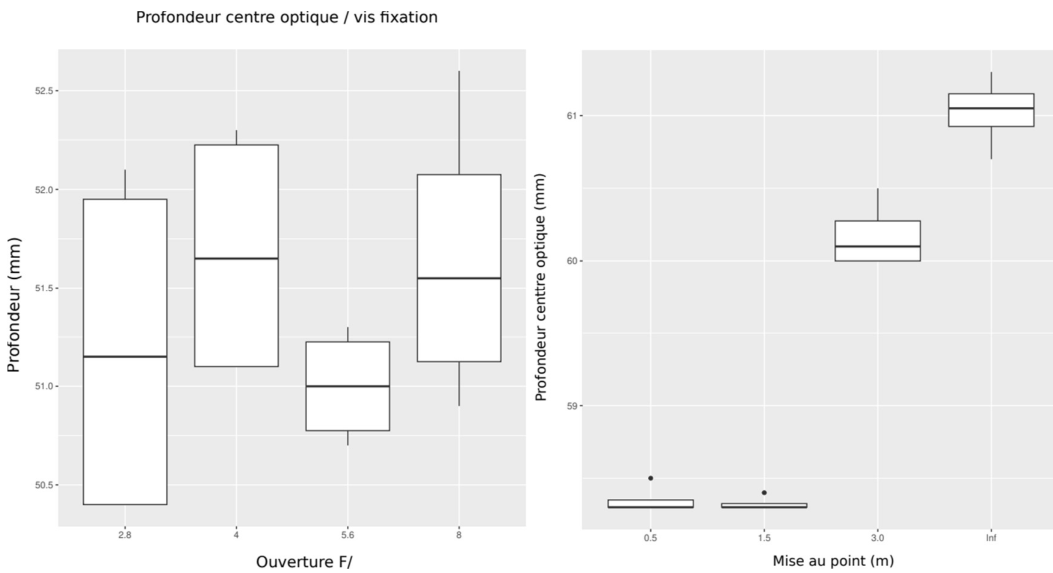


Fig 13. Depth of optical centers (on the optical axis, referenced to an arbitrary point) for different apertures (left) and for different focus (right).

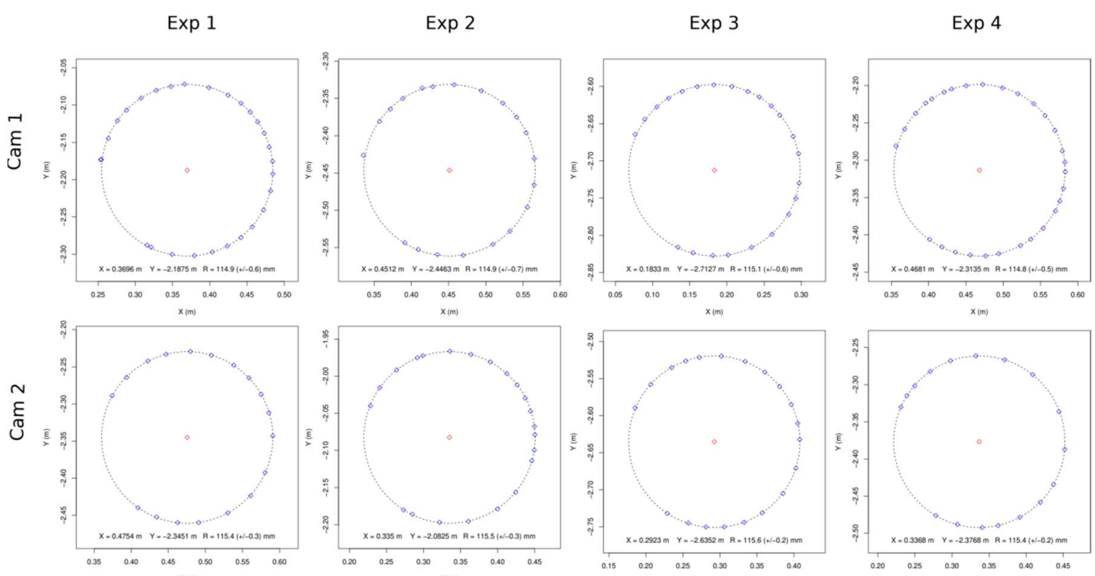


Fig 14. Optical center positions for 2 different cameras and 4 different experiments, revealing that determination is accurate, reproducible, and does not depend on individual camera models.

Procedure

As a summary, the following procedure must be conducted for camera calibration before a mission:

1) Determination of the (maximal) speed at which the vehicle *v* is supposed to move: combining v with the required accuracy σ provides the maximal allowed camera exposition time τ:

$\tau = \sigma / v$

For example, on UTAC test, with a maximal expected speed of 50 km/h (13.9 $m.s^{-1}$), and σ = 2 cm accuracy, the maximal exposition time is 1440 µs.

2) Based on the exposure time τ, and for low (winter) luminosity conditions, aperture of camera is set. It is important to find a trade-off between a large aperture, enabling to get decent exposure of images, even with a short exposition time (exposure is the product of the exposition time by the aperture), and small aperture, enabling to get a long depth of field (the distance between the nearest and the furthest objects that are in acceptably sharp focus). Moreover, with this specific model of camera, tests revealed that apertures below F/8 tend to produce blur artifacts on the borders of images.

Experiments have shown that F/4 is a reasonable value for the aperture. It is important to set this value in pessimistic light conditions. This makes adaptation possible if ever light conditions are better than expected during the test. In that case, adaptation is done by reducing the exposure time τ. Conversely, if the aperture is set in optimistic light conditions, because the calibrations of distortions and optical centers are performed with a fix aperture, adapting to lower light conditions on the field leaves no choice but increasing the exposure time τ, hence introducing undesirable motion blur in the images when the vehicle is close to its maximal speed, which in turns will decrease the target automatic detection capabilities, as well as the accuracy of bundles (each bundle would be an “average” towards the mean position of target on its trajectory during exposure time). In parallel, ISO value of camera is kept constant at a relatively low value. It could be used as a last resort for high-speed driving scenarios (> 50 km/h), or very low light-conditions (see Valeo POC test).

3) Determination of the range of distances at which the vehicle is expected to be seen from the camera (d_{min} and d_{max}). For each camera focal length (8 mm and 12 mm), focus is set in order to have sharp optical definition of all objects at distances ranging between d_{min} and d_{max}. Hyperfocal distance can be used to ease the determination process, in particular when d_{min} and/or d_{max} are not accurately predictable. At the end of this step, rubber tape is used to keep both focus and aperture at a stable position on the objective rings.

4) A distortion model is established for each pair of camera/sensor. This is done by taking a set of 20 images on a textured area, and by performing bundle adjustment as described above.

5) For each (pre-calibrated) pair of camera/sensor, the position of the optical center along the optical axis is determined by taking a set of images of a topometric polygon on a rotational support, as described above. Then the camera/sensor pair is rigidly attached to its support.

At the end of these 5 steps, each camera is ready to be used (within at least one week) for the experiment.

2.2.3.1.2. Synchronization and timing

a. General specifications

Within the context of the PRISSMA project, 12 cameras are used to estimate the trajectory by photogrammetry. These cameras (BlackFly S BFS-PGE-161S7M-C) are equipped with a Sony IMX54216 MP global shutter sensor. We want all the images to be taken in a time interval corresponding to a vehicle movement of less than 1 cm. If the vehicle is moving at 10 m/s (36 km/h), the images would have to be acquired within 1 ms from each other to be assumed to be instantaneous.

b. Trigger delay

In this firs test, the time delay between the capture command (via an electronic pulse on GPIO pin of camera) and the image acquisition is going to be measured.

Device

The test device is based on a Field-Programmable Gate Array (FPGA) board generating a VGA signal at 800 x 600 px resolution and 60 Hz frequency, and displayed on a cathodic screen. An image capture command (synchronized with the VGA output) can be generated with a push button on the device.

Two counters (one on the FPGA *board* and one external) are displaying the time (in 10th of milliseconds) elapsed since the generation of the image capture command. This device does not allow using long exposure time (digit segments become difficult to read as they get superimposed on each other), so VGA output will be mainly used to measure accurate delays.

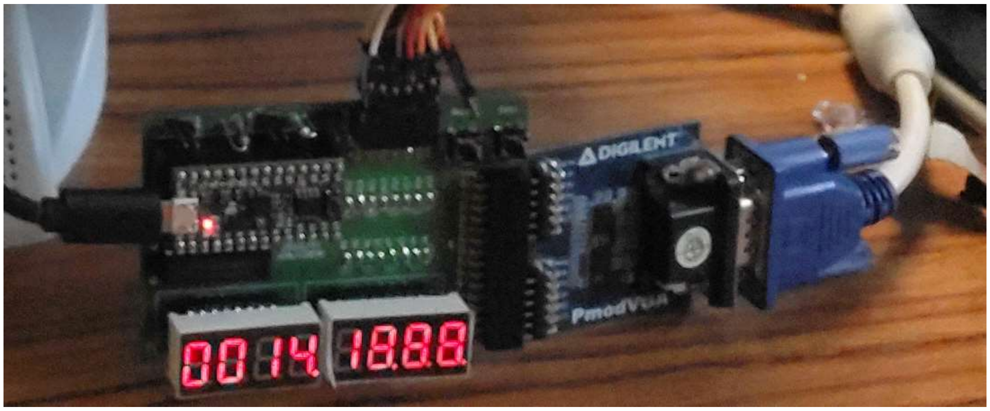


Fig 15. FGPA board (power supply on the left, VGA output on the right and connection to an external digital counter on top.

The FPGA board generates continuous flow of black images, so that the screen gets synchronized with the VGA signal. After a command from the operator, one second delay is introduced, then camera trigger signal is generated along with a special image pattern enabling an easy numbering of lines displayed on screen (see

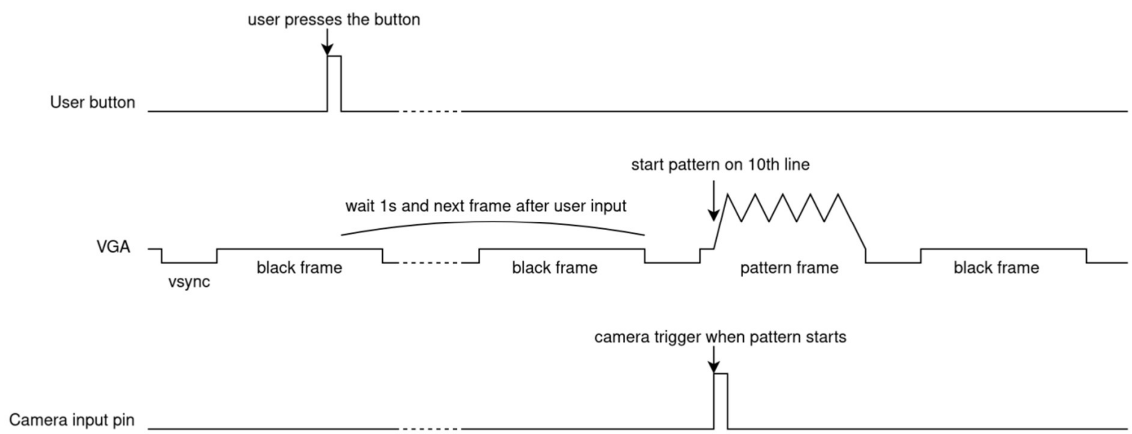


Fig 16).

Camera is set in stand-by mode, waiting for image capture input signal. *SpinView* software (the official industrial software delivered with the camera) is configured as follows:

- Trigger source = Line 3
- Trigger Overlap = Read out
- Trigger Mode = On
- Image capture on falling edge

Displayed pattern

The pattern displayed on screen depicts a binary encoding of line number (most significant bit on right), and is replicated with horizontal symmetry in order to compare the beginning and end of lines (see Fig 1Fig 18).

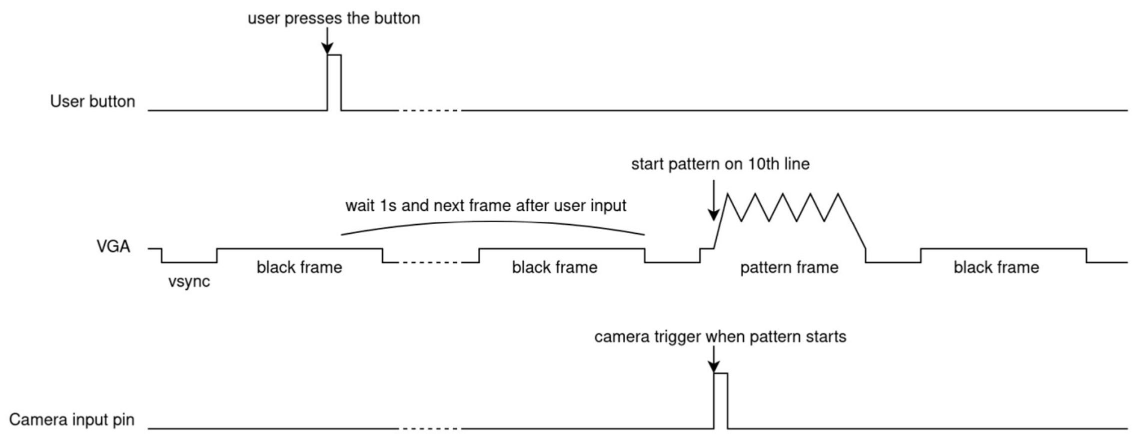


Fig 16. Chronogram of device

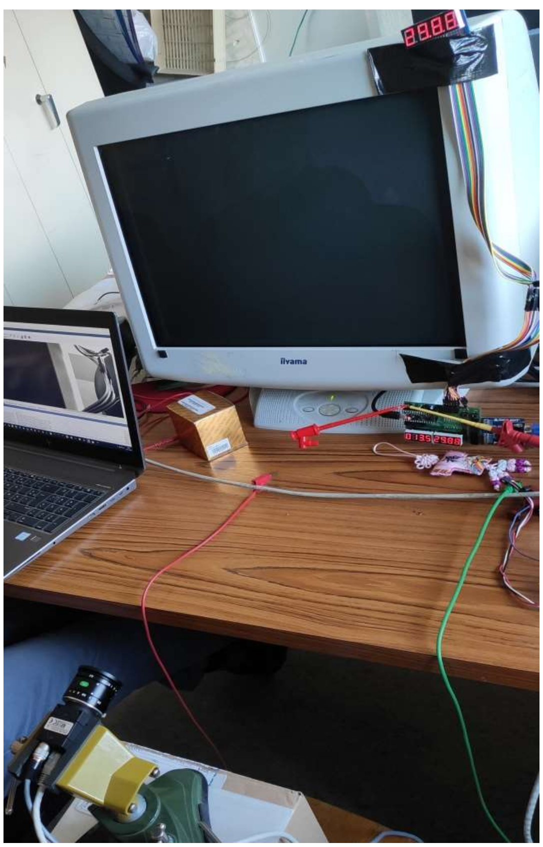


Fig 17. Complete device

The complete device (see Fig 17) is composed of a cathode-ray tube screen, two counters and the camera under test (bottom left). This camera is connected via Ethernet link to a PC (middle left) for configuration and image recording, and to the FPGA board (middle right) via GPIO pins.

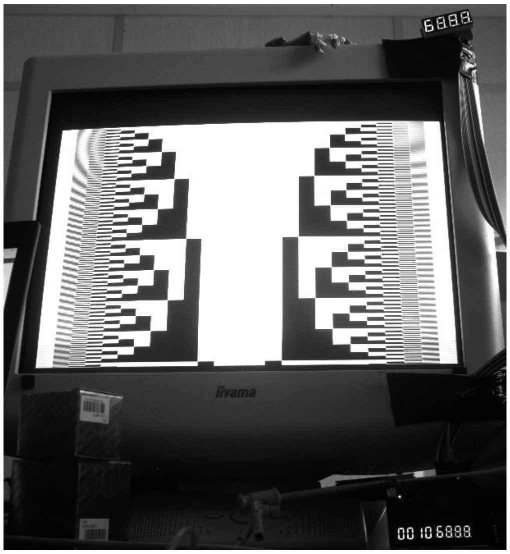


Fig 18. Complete pattern depicted on CRT screen

It is easy to count the rows: on the first 32-pixel-wide column, all the rows are lit (white), then one out of two on the next line, then two out of four, and so on.

The generated image being 800x600 pixel resolution with 60Hz frequency, each line display is lasting 26.4 µs. The full line being 1056 pixel, only 76% of these 26.4 µs are effectively used (http://tinyvga.com/vga-timing/800x600@60Hz). Screen technology makes that there is no noticeable delay between the input VGA signal and the display on monitor. Moreover, screen phosphorescence is short enough to make it easy to differentiate between a currently lit on pixel, and one that was lit on the previous cycle, as illustrated on Fig 19.

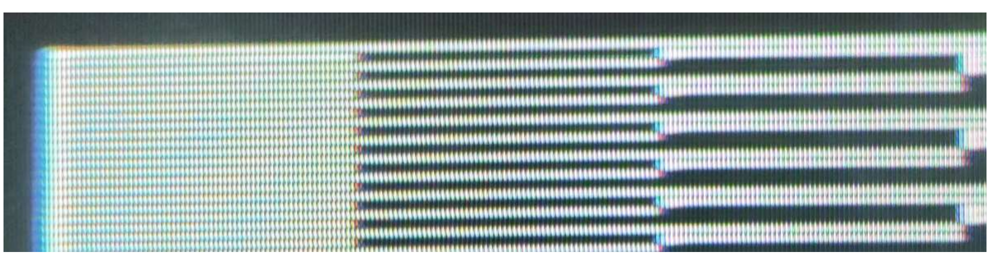


Fig 19. Beginning of pattern

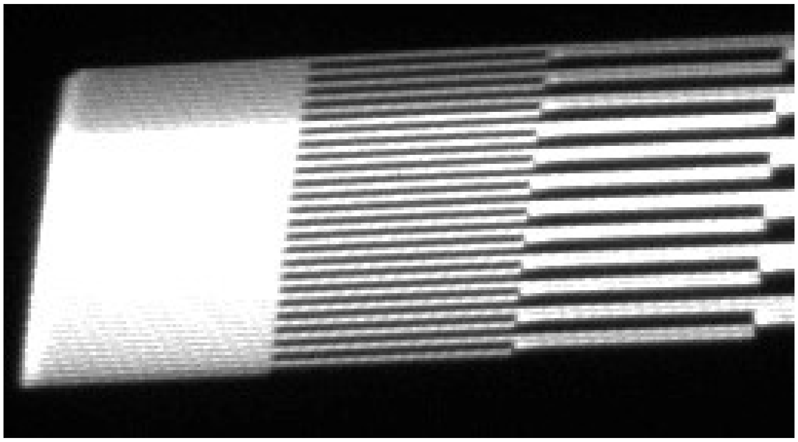


Fig 20. The first 9 rows are darker because they have been lit on before the beginning of the camera exposure. The next 37 rows correspond to the true exposition period. The last rows are darker because they have been lit on during a shorter time interval.

Measurements

For the first test, the camera was set with a 100 µs delay between command signal and acquisition and a 1 ms exposition time.

The images are not perfect: readings with one row accuracy are difficult to achieve. They are depicted on Fig 21.

It is possible to see the beginning of the pattern, with grayer lines at the start, corresponding to the acquisition delay. These lines were illuminated between the instant of the trigger signal and the beginning of the camera exposition, and are then darker.

Fig 21. First experimentation with 100 µs delay and 1 ms exposition time

The reading on this image is about 6.5 rows, which correspond to a time delay of:

$6.5 \times 26.4 + 0.5 \times 0.76 \times 26.4 = 168$ µs

The exposition duration is evaluated at 38 rows, or 1003 µs, which is consistent with the settings of the camera.

A second experiment with the same parameter settings gave 120 µs of delay and 1003 µs of exposition time. Two further tests have been conducted, with a delay estimation varying between 100 and 150 µs. All the exposition times seem very close to 1000 µs.

Tests have been conducted as well for 200 and 400 µs delay command, and results are reported in Tab 4. Control of delaysTab 4.

Parameter delay (µs)	**Exposition time (µs)**	**Screen reading (µs)**
100	1000	125
200	1000	250
400	1000	436

Tab 4. Control of delays

Conclusion

It seems that the delay is about 20 to 50 µs longer than requested. The exposure time is consistent with the demand. There is only a variation of less than 50 µs, which corresponds to a vehicle displacement of 0.5 mm at 36 km/h. This camera model is therefore well suited to the synchronization needs.

c. Synchronization on 1 pole

Methods used to validate the synchronization between the 3 cameras of a station: oscilloscope tests, tests with shots on the VGA monitor, etc. Performances observed “in the lab”.

Triggering three cameras at the same time: We can affirm on the basis of the captured images that the cameras have indeed instantiated the display event of the test pattern on a VGA screen at the same time. The experiment is carried out using a pulse generator, a digital chronometer and a VGA screen of which a test pattern is displayed by triggering.

Fig 22. Display of synchronization lines on CRT screen

Alternative method to check the simultaneity of the triggering for the three cameras of one pole. The grabbed image of the three cameras shows the same number displayed (in seconds) by the digital chronometer (Fig 23).

Fig 23. Synchronization test with digital timer

d. Synchronization between poles

Same methods as above were used to validate synchronization between the 4 poles of the system. Estimation with the oscilloscope, based on the analysis of the signals generated by the 4 GEOSTIX GNSS receiver, the delta is less than 0.1 ms (see Fig 24). The frequency of the signals depends on the chosen FPS, in this case it is 10 Hz (period = 100 ms).

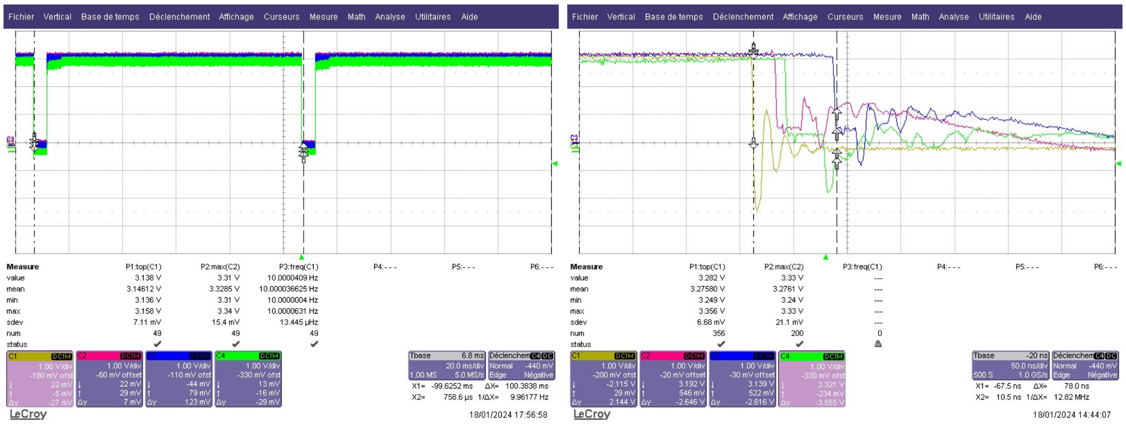


Fig 24. Oscilloscope showing output trigger signal of GEOSTIX GNSS modules. The 4 GEOSTIX triggering signals at the same frequency 10Hz (left) and the jitter between the 4 signals is at most around 78 ns (right)

e. Timing

Each pole of the vision system acquisition is triggered by a signal from a "GEOSTIX", this one is driven as the three others by an RF signal command which gives the GPS time to start the signal triggering and to start the acquisition.

To get the right timestamp of each image we can use two methods: a simple one is based on the image index and based on the stability and precision of the FPS not only the signal but how many time the camera take to get the image. A second method is to use the current PC Time to correct the camera timestamp. This second method need to set the 4 PC system at the same time (GPS Time).

The GEOSTIX receivers provide acquisition trigger pulses, geographic position, GPS time and also a 1PPS pulse of 1s frequency which improves the accuracy of the computer GPS timekeeping. To use GPS and 1PPS pulse we need to install two time synchronization software tools, GPSD and Chrony which is based on Network Time Protocol (NTP).

- **GPSD** (GPS Daemon), install a link between the GNSS receiver (GEOSTIX) and a Linux System. This GPSD needs to identify the serial port to which the GEOSTIX is connected. For this a configuration file "/etc/default/gpsd" will tell it where (serial port) to look for the information and therefore the NMEA frame of the GNSS.
- **Chrony**, adjusts the kernel system clock from external time sources (GNSS) using the NTP protocol. The time sources are mentioned in a configuration file "/etc/chrony/chrony.conf".
- **PPS**, is a signal provided by the GEOSTIX every second (1Hz) that will improve the accuracy of the system's clock. This signal is generated as soon as the date and time messages appear in the frame broadcasted by the GPS. This technique is used to adjust and tune the clocks of a system network. The accuracy achieved is in the order of a micro-second.

To make sure that the GPS is working and GPS time setting is in place on the four systems, you can launch some tools, such as "gpsmon", chronyc sources, chronyc trackings, ...

Image acquisition: as shown in Fig 25, a log file is provided with the acquisition and records indexes, timestamps, and other information…

```
Fichier   Edition   Format   Affichage   Aide
        Frame filename: IGN2_202401301546-00-cam-22348133-1-60551855390809-0
        Frame ID: 1
        Timestamp: 321349859808
        ConvertedTimestamp: 60551855390809
        Exposure time: 661
----------image->info------------------
imagesize: 9324376
width: 5320
height: 3032
 Xoffset : 0
 Yoffset: 0
Pixelformat: 0
        Frame filename: IGN2_202401301546-00-cam-22348133-2-60552055361892-1
        Frame ID: 2
        Timestamp: 321549848360
        ConvertedTimestamp: 60552055361892
        Exposure time: 661
----------image->info------------------
imagesize: 9001656
width: 5320
height: 3032
 Xoffset : 0
 Yoffset: 0
Pixelformat: 0
        Frame filename: IGN2_202401301546-00-cam-22348133-3-60552255361681-2
        Frame ID: 3
        Timestamp: 321749860456
        ConvertedTimestamp: 60552255361681
        Exposure time: 661
----------image->info------------------
imagesize: 8995152
width: 5320
height: 3032
 Xoffset : 0
 Yoffset: 0
Pixelformat: 0
        Frame filename: IGN2_202401301546-00-cam-22348133-4-60552455370648-3
```

Fig 25. Log file of image acquisition

f. Absolute timestamp validation

Absolute timestamp is completely relying on the GNSS time system. Even with poor sky-visibility factor, a positional accuracy of $\sigma = 5$ to 10 m is easily achieved with modern receivers. The receiver internal clock synchronization required to reach such a precision is $\sigma_t = \sigma/c$ where c is the speed of light, which results in time accuracy of a few tens of nano-seconds.

This absolute timestamp error was indirectly validated with the synchronization of 4 poles, each equipped with independent receiver, which revealed a jitter of 78 ns between different poles (note that this delay is not only the synchronization error of the internal clock of the receiver, but it also includes processing time and electronic lag between the receiver and the output signal interface.

The used GNSS module (GEOSTIX) is based on the popular u-blox ZED F9P bi-frequency chip, whose absolute timing capabilities have been thoroughly assessed (see link below, page 4), and are usually estimated at a 30 ns root mean square error.

https://content.u-blox.com/sites/default/files/ZED-F9P-04B_DataSheet_UBX-21044850.pdf

Another interesting and complete analysis of u-blox timing performances can be found in [13].

2.2.3.2.3. Absolute reference

Absolute positioning and orientation of the local survey in a national reference frame is performed jointly with the use of high-precision GNSS receivers and a gyrotheodolite.

The word "receiver" designates here and in the rest of this section, an antenna setup on the point and a receiver where the GNSS observations are recorded.

a. General specifications

Absolute accuracy of points (GCPs and optical center of cameras) is directly transferred to the absolute accuracy of the output trajectory. Therefore a 2 cm absolute accuracy is required for the survey.

b. Position

Though a 3D absolute accuracy of 2 cm can be easily achieved with a single on-field GNSS receiver (whether in *Post Processing Static* mode with several base stations from the GNSS permanent network, or even in *Precise Post Processing* mode for areas without such a network), in order to comply with contexts requiring a much better accuracy, and for the sake of robustness and control, a set of 4 GNSS receivers have been deployed in network.

One GNSS receiver was set in each corner of the surveyed area, and they acquired measurements for all the topometric and driving steps (*i.e.* about 6 to 7 hours of data with a sampling of 15 seconds). This theoretically ensures a precision below 1 cm for each individual receiver. Furthermore, all 4 GNSS observations have been processed with *Bernese* software, enabling to perform mutual control between their individual solutions. *Bernese* is a reference software in GNSS scientific computation, and provides its solution with a detailed report on the standard deviation of absolute reference error: https://www.bernese.unibe.ch/

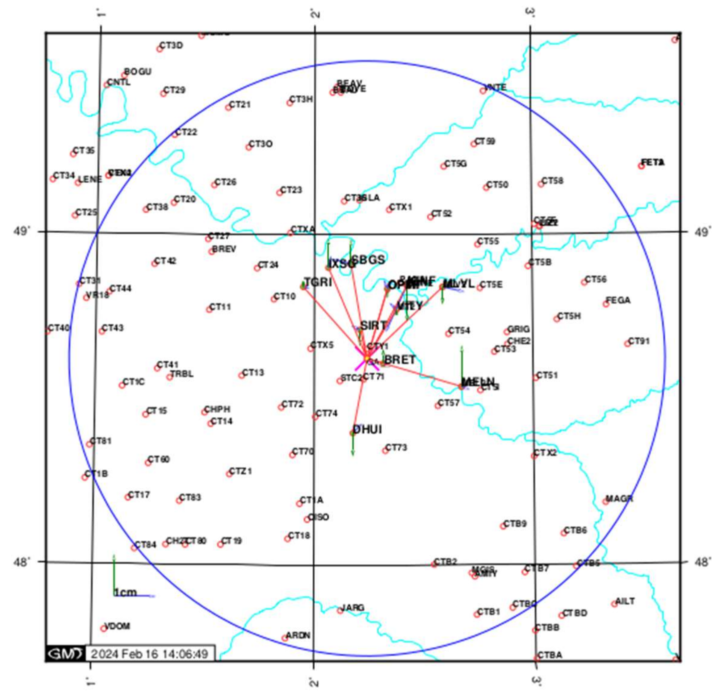


Fig 26. Illustration of baseline computation in Bernese for a GNSS point of the test at UTAC site. Each baseline is a link between a local GNSS receiver and the national permanent GNSS network RGP

Computing solutions in *(Network) Post Processing Static* mode enables to benefit from the absolute reference quality of the RGP (Réseau GNSS Permanent), the permanent GNSS network. For utmost accuracy, solution could be computed with GPS, Glonass and Galileo satellite constellations.

For example, Bernese software provided the following coordinates on UTAC test site for one of the 4 GNSS receivers:

`E = 644070.797 m  N = 6836392.411 m   Alt =  159.029 m`

Planimetric coordinates are expressed in Lambert 93 legal projection of the RGF93 reference frame, and altitude in IGN 69 system converted from ellipsoid height using RAF20 model. The 95% confidence intervals are provided as follows:

`E_N:   9.0 mm     E_E:   8.7 mm     E_H:  18.2 mm`

This makes a 1-sigma uncertainty of about **4.5 mm** in each of the horizontal axis, and **9.1 mm** in vertical axis, which is beyond the required accuracy of 20 mm on each axis. According to *Bernese* report, this uncertainty may be attributed to the baselines measurement inaccuracy (with a theoretical uncertainty of about 2 mm) and to the intrinsic uncertainty of the reference network (about 3 mm)

Processing was done with a combination of 12 base stations (located at distances ranging from 5 to 35 km). The standard deviation of 7-parameters Helmert transformation from the determined coordinates of local GNSS network (including the 12 base stations) to the RGF93 theoretical coordinates has 3 mm standard deviation in horizontal axes and 6 mm in vertical axis.

Processing was also independently carried out with *RTKlib* open source free software, with results differing by less than 10 mm on each processed baselines towards RGP reference stations, which complements to guarantee the validity of the absolute positioning of the survey.

Bernese	`E = 644070.797 m  N = 6836392.411 m`
RTKlib	`E = 644070.806 m  N = 6836392.403 m`

Tab 5. Comparison on one GNSS receiver between Bernese and RTKlib software for validation

Secondarily, a relative network scheme was processed with Leica *Inifnity* software, to improve the accuracy of 3 receivers with respect to the 4^{th} one, set as a reference for the local network.

Note that on UTAC test site, observations from one of the 4 receivers have been discarded, after suspicion of accidental motion of its tripod, based on the bubble plate level indicator tied to the tribrach.

Besides, all receivers are annually checked with long sessions (> 48h) and compared to very accurate reference base, thus enabling to cancel any risk of antenna calibration error or receiver malfunction. This annual checking is part of the ISO-9001 IGN-Metrology department certification, and is necessary to guarantee accurate results.

c. Orientation

To overcome the drawback of GNSS receivers positioned relatively close to each other (e.g. barely above 100 m for the most two extreme receivers), and in the absence of any absolute reference points at far range, the orientation of network is reinforced with a gyrotheodolite observation, which is composed of a theodolite (measuring relative horizontal angles between targets), mounted on a top of a gyroscope (measuring absolute azimuth angle with inertial sensors)

This instrument must be calibrated regularly in order to measure the index difference between the zero graduations respectively of the theodolite and the gyroscope.

The calibration is performed before and after the survey on test site, from a precisely known (~ 5 mm planimetric accuracy) reference base point, sighting a remote (few kilometers) secondary precise reference target point. The observed azimuth is compared with the true azimuth (computed geometrically from the 2D coordinates of reference base and target), and the difference (called e-value) is registered. This index difference is then added on the angles reading to get accurate and unbiased estimates of the azimuths of sighted targets.

This process ensures that usage of gyrotheodolite on the test field provides an absolute orientation of the network with a 0.001° accuracy. This corresponds to 2 mm at a distance of 100 m, thus the orientation provided by this mean is tightly constraining the orientation of the local GNSS network.

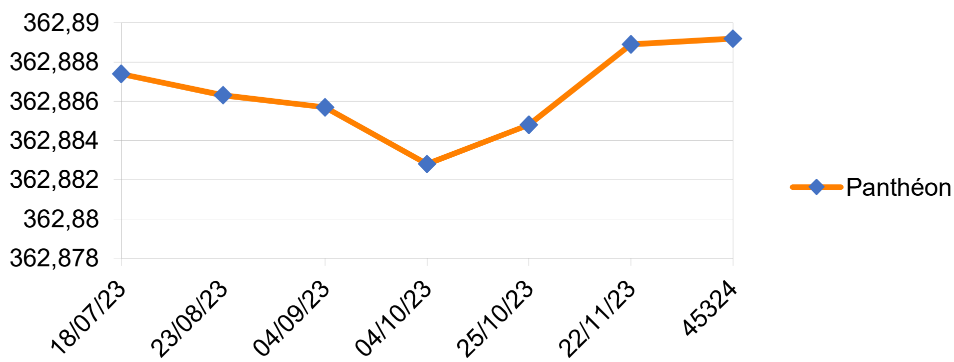


Fig 27. Evolution of the e-value (in gon angular unit: 400 gon = 360°) since summer 2023 until the experimentation at UTAC test site on (01/2023). Variation is stable and continuously monitored.

Fig 28. Gyromat 3000 coupled with a total station Leica TS16

2.2.3.2.4. Topometry

a. General specifications

While a 2 cm accuracy is enough for the uncertainty of the survey global translations, it may not be enough for couple of points separated by short distances. Consider a camera optical

center and its nearest orientation GCP located for example 5 m ahead. A 2 cm uncertainty on this distance turns into a 0.23° standard deviation error on the angle measured from this camera. It is therefore important to enforce a much better relative precision on the survey to ensure an accurate orientation of cameras in 3D space. The problem can be stated as follows: given a test site area of about 100 m extent, and final required accuracy of 2 cm, the expected angle accuracy on each camera orientation is 0.02/100 = 0.01°. Assuming the orientation angle is determined by a unique GCP point located at 5 m distance from the camera optical center, the relative accuracy of points on the scene should be: 0.02/100 x 5 = 0.001 m.
Therefore, about 1 mm relative accuracy is needed between all surveyed points.

b. Calibration

All the surveying equipment used is regularly checked in accordance with a procedure validated by ISO9001:2015 certification, which also ensures the traceability of the equipment used.

c. Preparation

In order to estimate the necessary and sufficient observation redundancy to ensure the expected accuracy, a simulation of the topometric figure is carried out using the Comp3D least-square adjustment software developed at IGN.

Two simulation strategies are readily available in Comp3D:

- Variance propagation method enables to simulate a least-squares computation without observations.
- Monte-Carlo sampling, for a pessimistic “worst-case output”, which unlike the former, can handle more efficiently non-gaussian distributions and extreme values.

Based on this simulation, we can then select the optimal topometric figure allowing to reach the desired accuracy, and estimate the observation time required and the material and human resources needed.

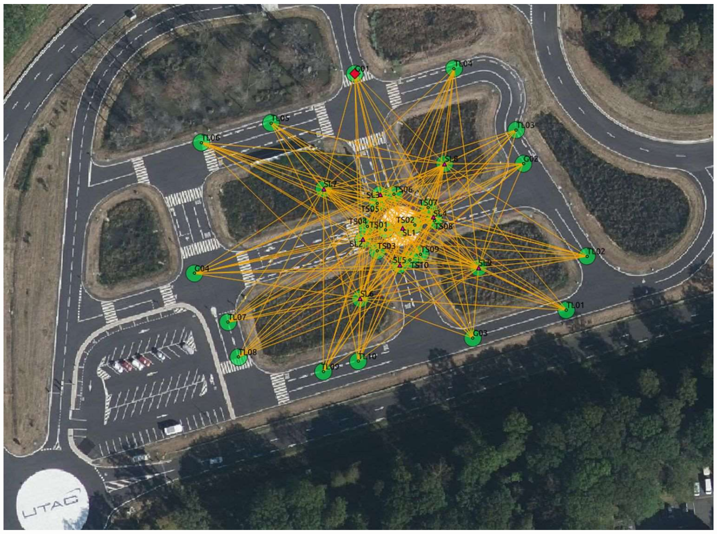


Fig 29. Simulation of topometric observation figure at UTAC

d. Surveying

Horizontal and vertical angle observations and distance measurements are carried out using a precision tacheometer (such as the Leica TM60) combined with high precision reflector prisms. The topometric stations are installed in "free" or "flying" mode - not centered on a material point - so as to avoid centering and height measurement (a source of uncertainty).

For natural ground points, target pointing is done with prism, using a short-length support stick, in order to minimize centering errors.

Measurements were taken in double-turn mode to cancel out collimation errors in the instrument, with verification of the closure of the horizon turns (less than 15 dmg, *i.e.* 0.00135°).

Temperature and atmospheric pressure are measured at each station, in order to correct distances.

e. Processing of observations

The data is recorded in raw form by the total stations, then pre-processed by the *PrepaComp* software, which reduces the horizon circles between right and left circles and makes meteorological corrections. The closures are exported for verification.
All the observations are processed using *Comp3D* adjustment software, developed by IGN and based on a least-squares calculation module:

- An initial calculation is carried out using minimum constraints (one point and one bearing imposed), in order to estimate the internal accuracy of the topometric measurements. This also makes it possible to detect any errors in the measurements before taking GNSS observations into account, and to set the a priori standard deviations of the observations made.

- Finally, a global adjustment of all the observations (topometry, laser, GNSS baseline and gyrotheodolite) is carried out by fixing the coordinates of the GNSS points in order to allow the setting in reference of the whole building site, thus determining the expected coordinates of the GCPs as well as their associated accuracy.

2.2.3.2.6. Target automatic detections

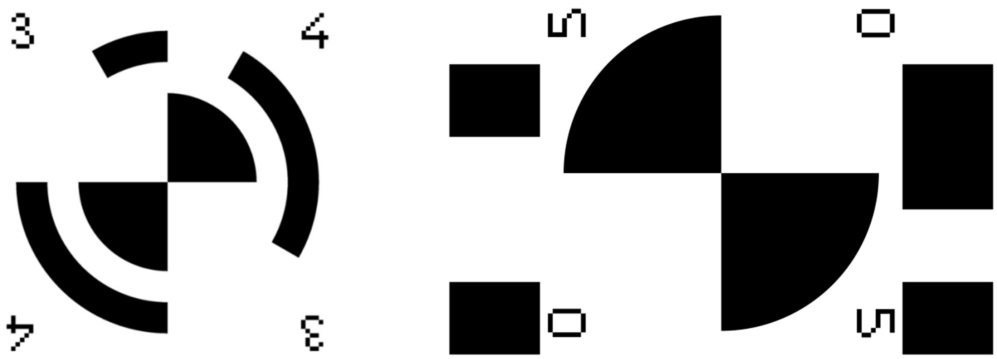


Fig 30. Target pattern

With a few images per second, and multiple cameras aiming at the vehicle, the number of targets to pinpoint on the images may be considerable (up to a thousand per second). Besides, for precise localization of the vehicle on each frame, it is important that the target pointing on images is done with a sub-pixel accuracy. This constraint does not allow for a manual detection of targets in any reasonable amount of time.

For this reason, target design has been conceived in order to enable automatic detection and decoding of targets on a vehicle. This detection must be fast (at most a few seconds per image) and accurate enough to guarantee that bundle from the camera to each target is precise enough to reach a 2 cm accuracy at 50 m range.

To date, detection capabilities have been tested:

- by **simulations**, with a detection ratio of 80 % for targets with pixel size in the image over 15 px (size of the butterfly pattern) and with moderate incidence angle (below 45°) between target plane normal vector and line of sight from camera. This detection rate also includes the correct decoding of target. For the largest target model (XL size: 40 x 60 cm), the butterfly is 32 cm wide: for cameras equipped with 8 mm lens (0.02° angle per pixel), this corresponds to 20 px. Hence this target should be detected and correctly decoded with a success rate of 80 % up to a theoretical maximum distance of **60 m** for the 8 mm lens and a **95 m** for the 12 mm lens. All the other target detection performance can be assessed proportionally given the size of their butterfly pattern:

Center	Dimensions	Number	Distance*
XL	40 x 60 cm	5	60
L	30 x 45 cm	4	40
M	22 x 33 cm	4	35
S	22 x 22 cm	4	35
XS	18 x 24 cm	4	25

*Maximal theoretical distance to get a 80 % rate of detection and correct decoding with the 8 mm camera lens

Tab 6. Model, sizes, number and effective distance of detection/decoding of targets

These distance values are to be considered in relation with the typical ranges on UTAC test track: 10 to 80 m range between cameras and vehicle.

Based on simulation results, the accuracy of target center detection is constant for all targets above 20 pixels. The root mean square error between estimated and true center location is around **0.05 px**. With 8 mm lens camera, this corresponds to an angle of 0.001° (1 mm at 50 m range).

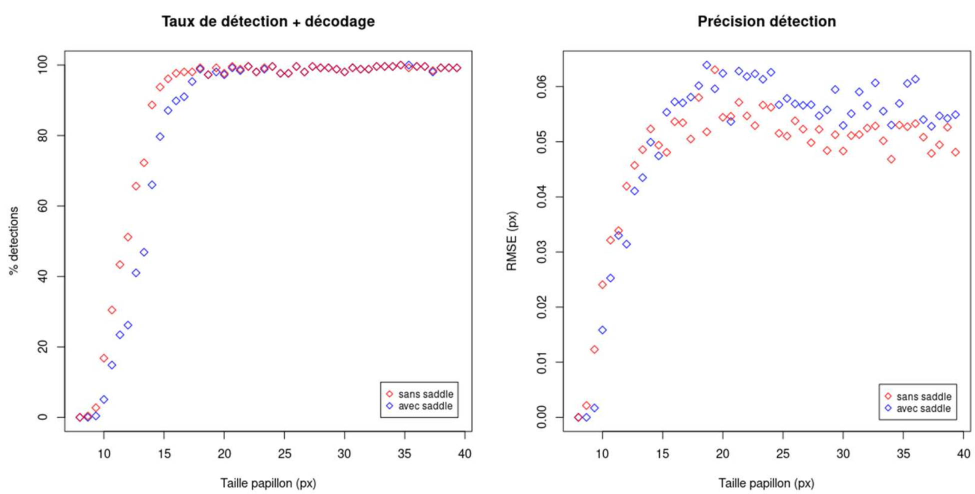


Fig 31. Detection + decoding rate (left) and center localization accuracy (right)

The detection algorithm has been optimized and validated based on simulations (Fig 32). Targets are placed with random positions, sizes and orientations on a typical image, and noise is also introduced to simulate a real camera sensor. This method enables to get a real ground truth for target centers (in pixel coordinates), and to optimize the algorithms on a large number of configurations.

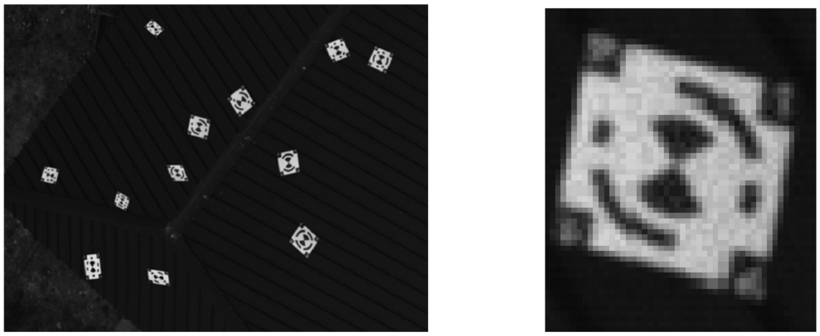


Fig 32. Left: simulation of (circular pattern) targets on an aerial image for target detection and recognition algorithm development, tuning and optimization. Right: simulation of sensor noise.

The relevance of simulation has been confirmed with real experimentation showing approximately similar performances.

- with **real experimentations,** decoding of XL (40 x 60 cm) has been tested with a ratio of 90% within at least a **50 m range**, for moderate angle of incidence (<45°) of targets respectively to the line of sight of the camera. Further test with ground control points showed that the target center localization accuracy is below **0.15 px** (compared to 0.37 px with manual detection). Using automatic detection of targets provided about 5 times more accurate 3D photogrammetric intersections (0.6 mm vs 3 mm error) on an experimentation conducted with ground control points calibration polygon.

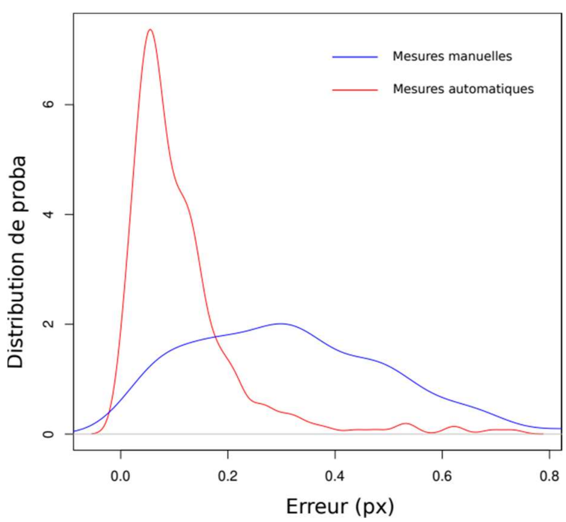


Fig 33. Comparison of distributions of manual (blue) and automatic (red) target detection errors on center localization (in pixels)

Once the detection capabilities of the target extraction algorithm have been assessed, it is important to evaluate how this accuracy transfers to bundle angles. This can be done theoretically, assuming a 0.15 px accuracy on detection, the size of pixel of 8 mm (resp. 12 mm) camera being about 0.018° (resp 0.012°), this results in an angular standard deviation of 0.0018° (resp 0.0012°) for 8 mm (resp 12 mm) cameras.

An experimentation was conducted to validate these accuracy predictions. Images have been taken with a pre-calibrated 8 mm camera on a polygon composed of a dozen of targets (whose coordinates have been determined beforehand with conventional topometric methods). Automatic target detection is performed on all images, and 3D bundles from optical center of camera to targets are generated. Global adjustment is performed with Comp3D software. All bundles have been input into the software with a prior standard deviation of 0.002° (according to the theoretical model above). The squared sigma of residuals (σ_0) after adjustment was 1.06, thus indicating that 0.002° is a reasonable standard deviation value for a 8 mm focal-length camera. Using multiple images for these tests, guarantees that σ_0 is not kept artificially low by underweighting of GCP 3D coordinates (assumed to be 0.1 mm standard deviation on each axis, which is already a fairly optimistic estimate, given that targets are printed on regular paper).

Fig 34. Calibration polygon

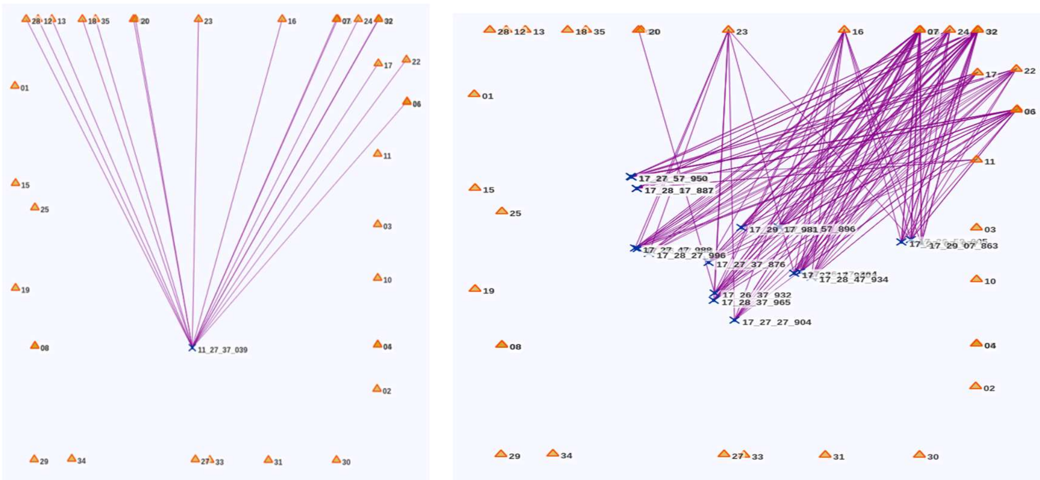


Fig 35. Bundle adjustment computation on a calibration polygon. Bundles (from 8 mm camera) towards automatically detected targets are depicted in purple. Left: assessment of optical center accuracy with 3D spatial resection. Right: assessment of bundle accuracy.

As a conclusion, the following accuracies will be input into all subsequent computations: **0.002°** for 8 mm camera and **0.001°** for 12 mm camera.
Simulations have also been used to test the impact of the target orientation on the detection accuracy.

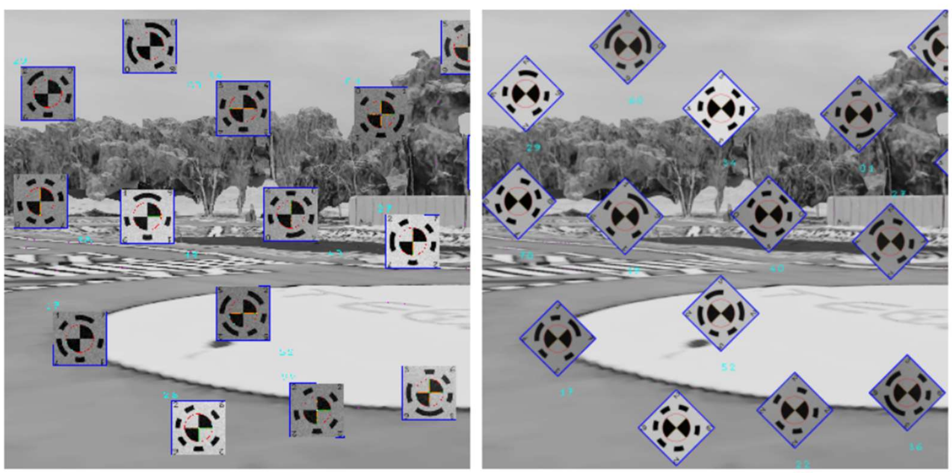


Fig 36. Comparison of detection of targets for 2 different orientations: 0° (left) and 45° (right) angle

According to the theory, localization of target center should be slightly more accurate when the target is at a 45° angle. Doing so enables to get a non-alignment between the target butterfly pattern and the image matrix of pixels, resulting in better correlation. Because the target detection algorithm does not directly rely on correlation, but rather on detection of primitive features (such as lines and ellipses), it needed to be tested whether this orientation of targets at 45° angle is also beneficial here.
Images have been generated (for different sizes of targets) to compare both orientations. Comparison between ground truth and detected target centers are shown on Fig 37.

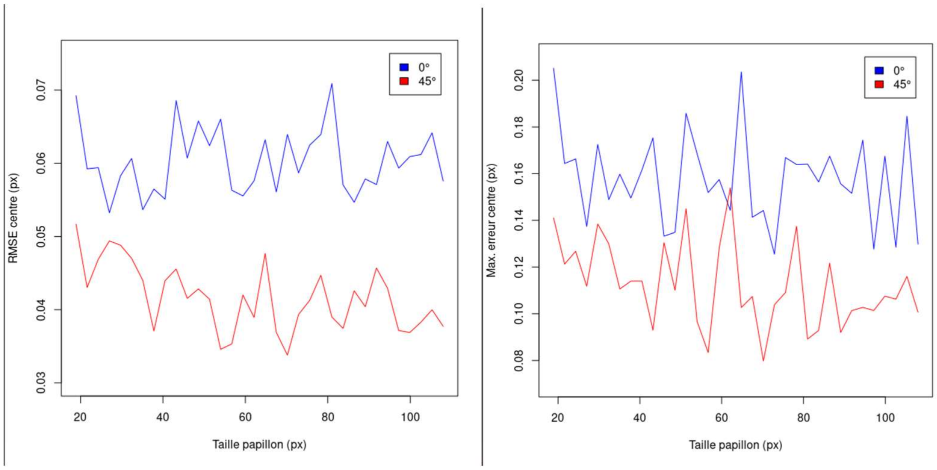


Fig 37. Evaluation of the impact of target orientation on center localization accuracy for different size of target butterfly: 0° angle (blue) is compared to 45° (red) on 2 different performance metrics, root mean square (left) and maximal (right) errors, both being given in pixels

The analysis result demonstrated that with this algorithm, location center is statistically more accurate with 45° angle targets, with **0.04 px** standard deviation of error, against 0.06 px for 0° configuration. The maximal error is also reduced to **0.11 px** compared to 0.16 px. This difference seems to be stable for all sizes of target butterfly pattern (between 20 and 100 px).

However, the difference being not so important, and the magnetic target being already purchased with the 0° orientation angle, it was decided to keep this pattern for real-scale tests during the PRISSMA project.

2.2.3.2.7. Miscellaneous

Various complementary methods are used to ensure that the resulting trajectory has the required accuracy:

- **Laser scanner** is employed during the topometric phase. It enables to record the relative positions of objects at close range (~ below 20 m) with acceptable accuracy (a few mm). Laser scanner is used (1) to get a redundant set of measurements of targets on the vehicle (this will be used to reinforce the rigid-block between targets, and the transformation between target system and the vehicle body reference frame), (2) to register the exact position of cameras mark points (because cameras are aiming at each other, it is possible to use external reference points, provided that theses points are somehow attached numerically to the optical centers of cameras) and (3) as an alternative way to control topometric determinations: though it is less precise than conventional tacheometric instruments, laser scanner measurements have the desirable property of recording automatically a very large number of point targets in a short period of time). Moreover, since observations are processed at the office, it is still possible to get 3D coordinates of non-planned points during the data processing step. In particular, on Valeo test site, the use of laser scanner enabled to detect that one of the camera support was set in reverse direction on its tripod, therefore causing a 10 cm bias of one of the camera position. With redundancy of observations, this error of course was producing statistical test rejection of the global adjustment of observations, without being able to clearly identify the source of error (in such types of networks, parameters and observations are significantly correlated, making it difficult to separate error sources). The use of laser scan is of paramount importance to reduce this risk of gross errors.

- **Auxiliary camera photogrammetry** was also used: images have been taken at various positions on the field with a manual camera, then a global bundle adjustment is computed, mixing these images, with a set of images taken by the (stationary) industrial cameras. This process enables to reinforce the initial position and orientation of cameras.

- **Rubber tape on wheels:** because desynchronization problems may be difficult to identify after the mission (for example a shift of 1 frame of one of the cameras results in a desynchronization of 100 ms; while it is a large error for the global computation, it is difficult to identify merely looking at the images). For this reason, a rubber tape was placed on one radius of each of the 4 wheels of the vehicle. Two cameras (placed on the same side of the vehicle) may be assumed to be reasonably synchronized if the rubber tape is located at the same phase (between 0 and 360°) on each of the wheels seen in common. Using this would have allowed identifying easily and quickly a one-frame shift on one of the cameras on Valeo test site.

- **Coded flash**: for more robust desynchronization detection, it is planned to use an omnidirectional lighting system, producing flash at high frequency (10 kHz or more), with a specific binary code pattern (similar to the coarse-acquisition code used by GNSS receivers, to measure time shift between signal arrival at the receiver from different satellites).

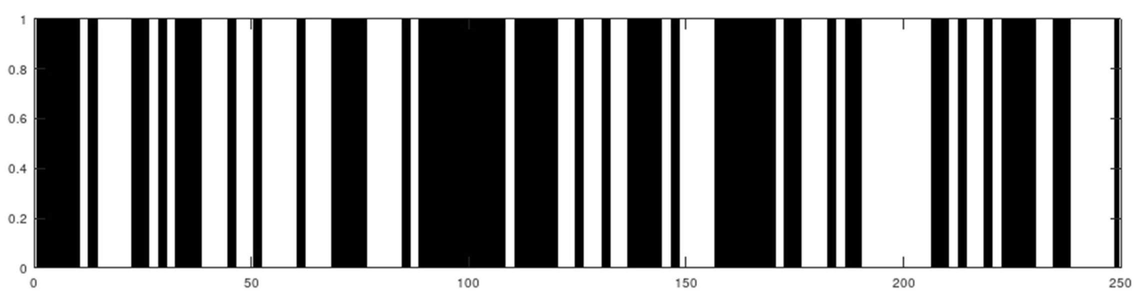


Fig 38. Example of the first 250 bits of the GPS coarse acquisition code. A similar pseudo-random code could be used to ensure synchronization of cameras, with a LED system, placed on the center of the scene, emitting omnidirectional flash pulses (flash is on only when the bit code above is at value '1' or 'HIGH').

For slots of 1 minute, with typical clock drift values of 10 ppm, desynchronization drift should be less than 1 ms, so it may be assumed to be approximately constant. Therefore, cross-correlations between flash signals received at each camera could enable to detect and estimate precisely any potential relative desynchronization of cameras. Moreover, if the flashing system is controlled by a GNSS receiver, absolute desynchronization of cameras can be measured.

This method was not yet evaluated in real conditions.

- **On-Vehicle GNSS receiver**: 2 low-cost GNSS receivers have been attached on the car roof (with magnetic support), one at the front and one at the back. This additional equipment can be used for validation of the absolute timestamping of trajectories. With a centimeter-level absolute precision on position, it can also be used as external control on trajectory. Moreover, using two receivers separated by 2 m, enables to get a rough estimate of the vehicle heading (with

0.3° accuracy), which can also be used as external control for the vehicle attitude determination. Note however that, GNSS positioning accuracy being lower than the expected accuracy of the trajectory estimated by the system, this external reference can only be used for the detection of potential gross errors, but does not enable to assess the full capabilities of the system.

2.2.3.3. Integration tests

In this last section, the results of integration tests – designed to assess the performance and the reliability of the complete system – are reported. A total of 3 main tests were conducted:

1) A simulation: note that simulation was also used in the first place to adapt the design to the UTAC test site (number and placement of cameras, GCPs…). In this section, it is used as a pre-validation tool; the configuration is set to its optimal solution and accuracy of the resulting trajectory assessed.

2) A first test in real conditions, with only 1 pole and 3 cameras, conducted at Valeo center.

3) A second test in real conditions with the complete system, *i.e.* 4 poles, each equipped with the nominal number of 3 cameras. Note that this test was also the final POC of PRISSMA project for IGN contributions. Therefore, after internal validation, the vehicle trajectory issued from this test was provided to Valeo for comparison to the internal AV navigation system trajectory.

2.2.3.3.1. Simulations

Simulations have been performed with *Blender* free software. A 3D model of the vehicle used for the final POC was downloaded and integrated into the simulation. Images of targets are attached to the vehicle at known location points (which enables to generate a ground truth for the evaluation of the results produced by the system).

An orthoimage and digital surface model (DSM) of the UTAC test site are also integrated into the scene, and 12 virtual cameras are placed in the four corners of the road intersection, according to the final configuration retained for the UTAC POC. GCPs are also integrated into the scene.

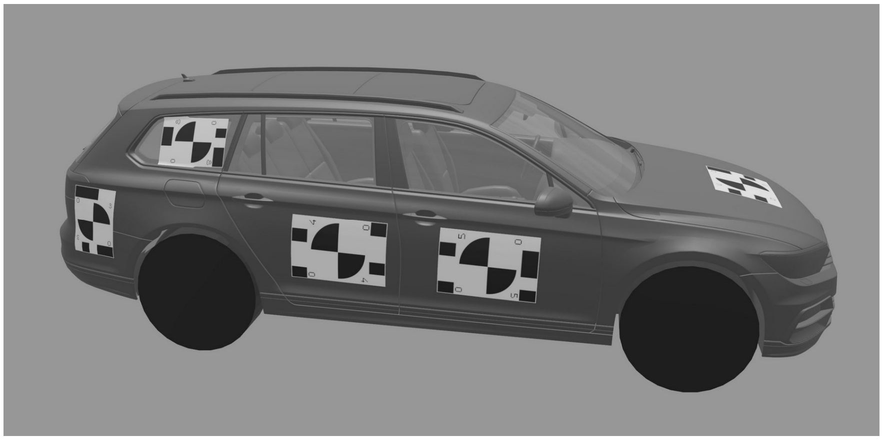


Fig 39. 3D model of the vehicle with the rigid-block of targets. Note that this is not the final configuration of the rigid-block.

Then, a sequence of images is generated at the frequency of 10 Hz, through each virtual camera, with the vehicle driving along a typical path scenario on the round-about. The automatization of the process is done with Python interface to Blender software, which enables to generate a large amount of simulation data at low-cost, for a comprehensive system performance validation in large array of scenarios.

Fig 40. Aerial view of the simulation on UTAC test site. Cameras are simulated with their exact optical parameters, and with approximate positions and orientations (in orange). A sequence of images “as seen” from the cameras will be generated for each position of the vehicle along its trajectory. The color of the vehicle is not important since images will be taken in grey level.

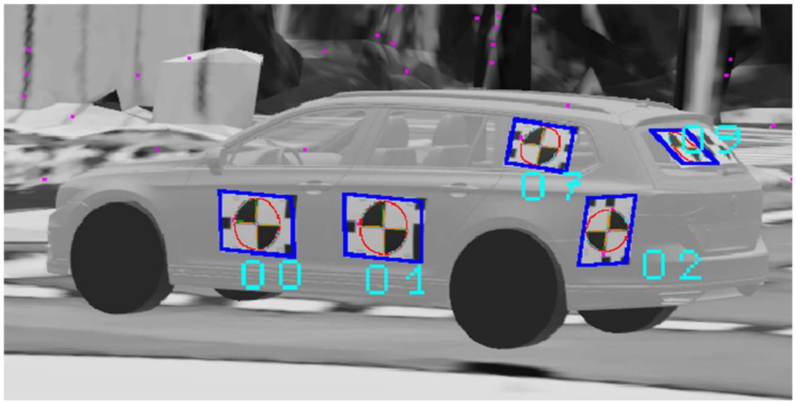


Fig 41. Automatic detection of targets on a simulated image

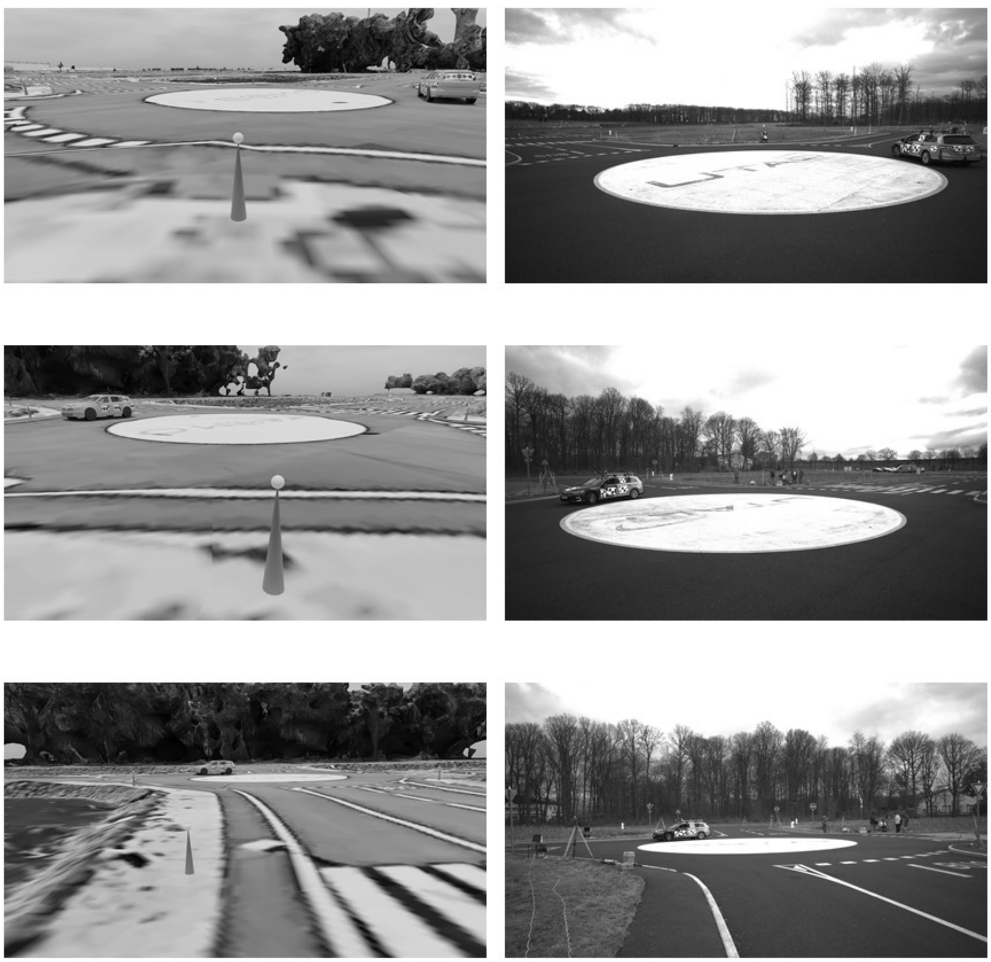


Fig 42. Comparison of simulated images (left) and real images on UTAC test site (right) for 3 different cameras: 8 mm (top and middle rows) and 12 mm (bottom row)

Simulation is very realistic insofar as it takes into account the following elements:

- All cameras are generated at a position very close to the one that would be used at final POC (with 1 or 2 meter uncertainty), and they integrate all their optical parameters (focal length, field of view, sensor resolution and depth of field), allowing to simulate very realistic images

- Background scene is replicated (especially horizontal road signs) which enables to assess the robustness of the target detector to “clutter” noise.

- Motion blur is also integrated into the simulation, to take into account that if luminosity is very low, and for images taken at short distance from the vehicle, a slight motion of the vehicle of a few pixels may be unavoidable during the exposure time. Result analysis demonstrated that this blur does not have any negative impact, neither on the accuracy of the 2D target center detection, nor on the final estimated position of vehicle after global adjustment.

- Targets are attached with curvature on the vehicle body, as it would be in real experiments. This was one of the main drawbacks of the target detection algorithm: its inability to decode significantly non-flat targets. Not knowing very precisely beforehand the exact model of car used for the final test, did not allow to take into account this curvature in the algorithm. Taking into account this curvature during simulation is a real asset of this evaluation method.

- Blender can simulate sun effects, and specular reflections on targets and/or vehicle body, which can be leveraged to assess the robustness of the target detection algorithm to the variability of lighting conditions.

- Potential desynchronization of cameras can be simulated if needed.

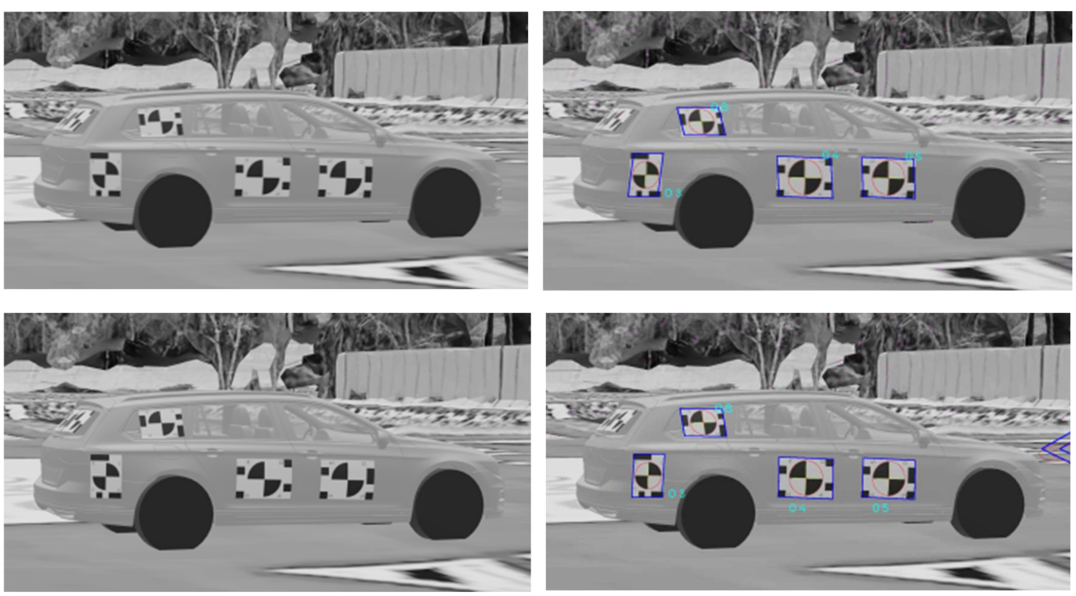


Fig 43. Comparison of detection of targets on (motion) blurry (top) and non-blurry (right) images

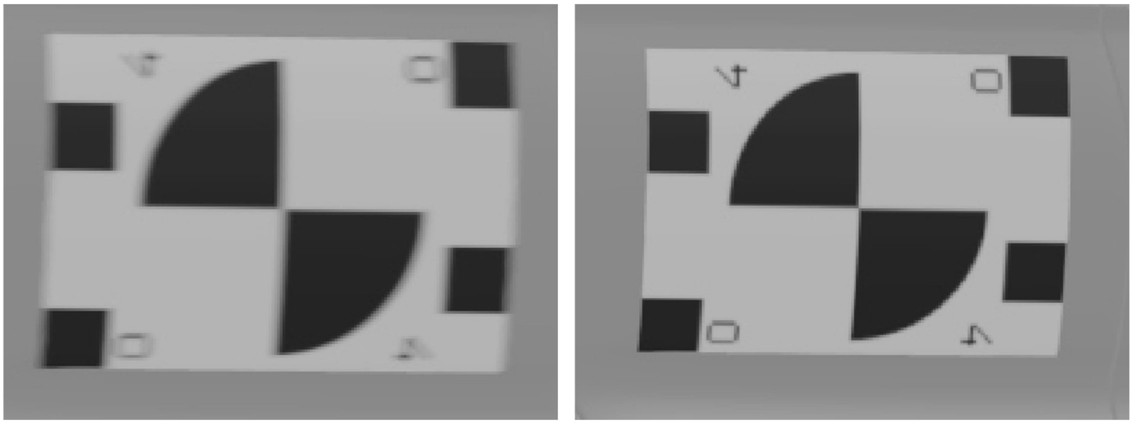


Fig 44. Comparison of a simulated (motion) blurry image (left) and non-blurry image (right)

At each time step, all simulated observations are integrated into Comp3D software, to get the position (and orientation) of the vehicle.

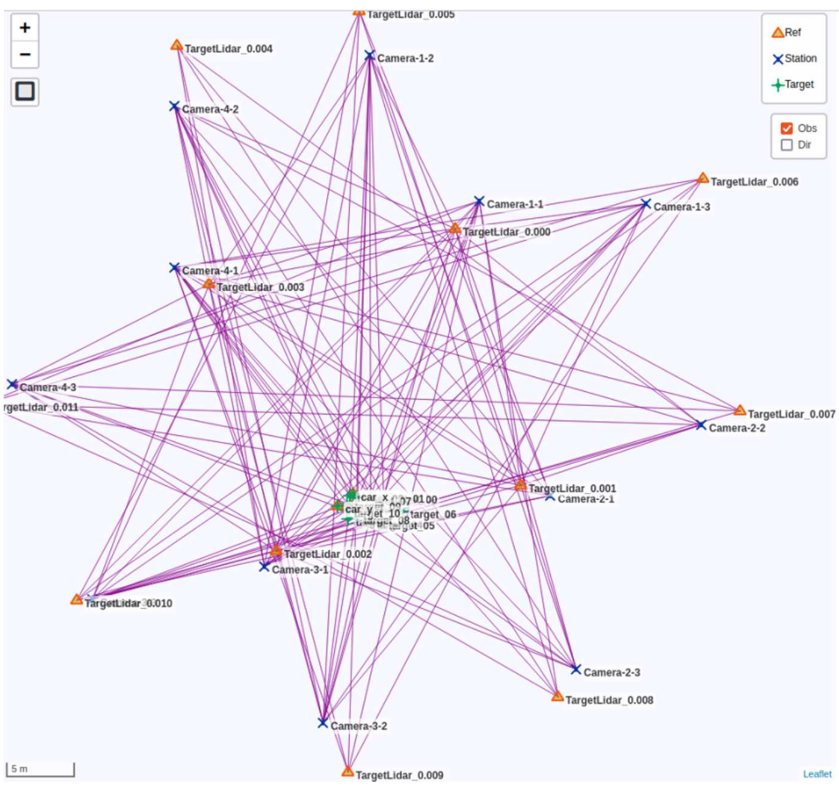


Fig 45. Global adjustment of (simulated) observations at one time step. Targets on vehicle are shown in green and cameras in blue. Bundles towards vehicle targets and GCPs are depicted in purple.

The typical resulting accuracy on targets and vehicle center determination at each time step is **5 mm to 10 mm**.

Results demonstrated that the true error (measured as the vector between the simulated and estimated vehicle center) is always included in the 2-sigma confidence ellipse.

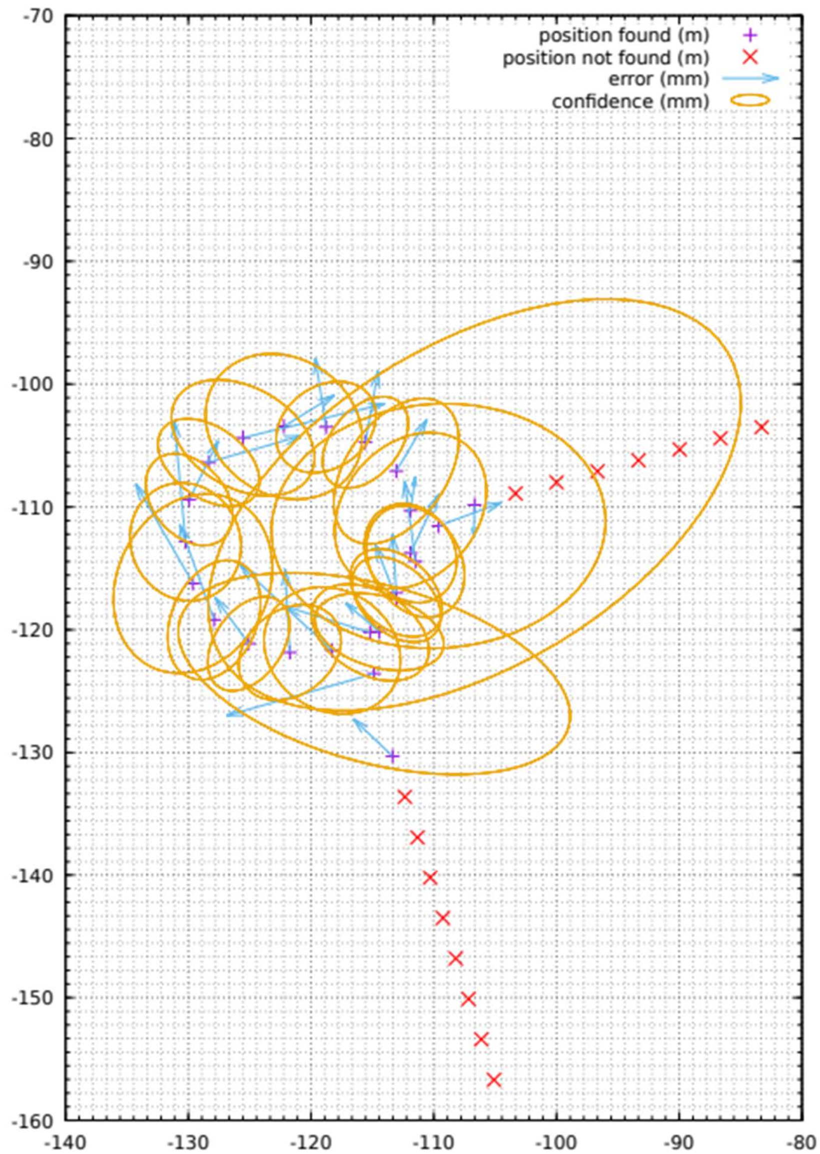


Fig 46. Horizontal uncertainty ellipses of the global adjustment determination of the vehicle center along its trajectory. The true errors between simulated and estimated positions are depicted by blue vectors. Red cross show that vehicle position can hardly be determined out of the area roughly defined by the convex polygon of cameras.

A few limitations of the simulation are however to be mentioned:

- It is difficult to simulate realistic human pinpointing on GCPs, which can still be the main source of uncertainty on the camera orientations, and then on the final trajectory.

- Bundles issued from cameras are weighting according to their experimental standard deviation (which integrates both center localization uncertainty and non-modeled distortions of camera lens). However, it does not take into account the non-stationarity of the residual distortion error within the solid angle seen by the camera: distortion errors of lenses are usually several order of magnitude higher on the edge of images. Because it cannot be modeled easily to a satisfactory degree of precision, distortion errors have been assumed to be stationary. Moreover, it was assumed to be spatially uncorrelated (*i.e.* white gaussian noise) which is not a realistic assumption.

- It does not take into account accidental or non-modeled behaviors, like tripod and camera instabilities, heat deformation, potential movement of GCPs between topometric phase and driving phase…

For these reasons, validation needs to be complemented by real test experiments.

2.2.3.3.2. Real test 1

This first real-scale test was conducted at *Valeo Mobility Tech Center* (Créteil) on a 50 x 25 m test track, on the afternoon of November 28th, 2023.

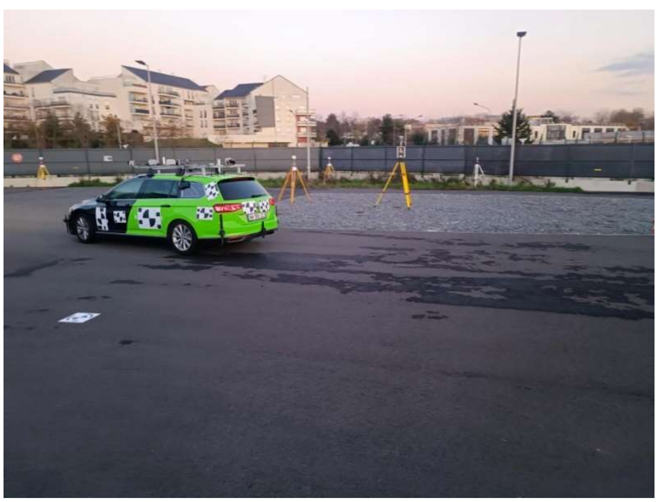


Fig 47. Picture of the test at Valeo center, during the set up. Yellow tripods are GCP targets and a GNSS receiver mounted on top of a topometric reflector.

During this test, only one pole with 3 cameras was tested, therefore all the synchronization was ensured through wires. Cameras have been set up at positions C1, C2 and C3 on Fig 48. Note that, unlike nominal configuration, which involves two 12 mm cameras and one 8 mm camera per pole, here on the opposite, two 8 mm cameras and one 12 mm camera were used. The rationale for that, was to make up for the relatively small scale of this model (compared to the final POC planned at UTAC); using 8 mm cameras enables to get larger field of view (hence more points on the trajectory seen simultaneously by the 3 cameras) while at the same time *simulating* an equivalent situation where vehicle would be seen by a 12 mm camera at a 50% longer distance.

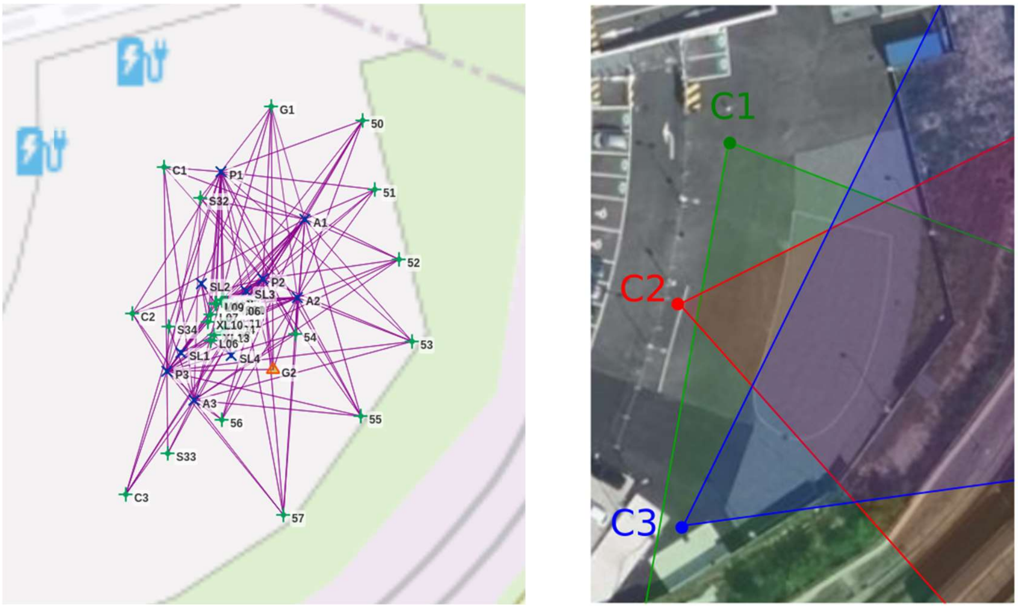


Fig 48. Topometric network of cameras and GCPs on the test track. Cameras are denoted as C1 (8 mm), C2 (8 mm) and C3 (12 mm). Field of views of cameras are depicted on the right plot. The vehicle was mostly moving in the area covered by all the three cameras.

At the end of the topometric phase, all GCP targets and camera optical centers coordinates are known in 3D space with a relative (resp. absolute) accuracy of 0.2 mm (resp. 5 mm). The rigid-block of targets on the AV (here also reinforced by laser scanner) also has a 0.2 mm relative accuracy.

The *base-to-depth* ratio (which is an index measuring how the system can accurately estimate 3D positions by stereoscopy) of the two extreme cameras is similar to the one used at UTAC, *i.e.* about 1.0 when the vehicle is located at the opposite side from the cameras (for comparison, this ratio is expected to range between 0.5 and 1.2 at UTAC, depending on the position of the AV within the round-about).

All the operations have been conducted as they are planned to be conducted at UTAC final POC, except that the number of GCP targets was reduced, gyrotheodolite was not used here, and only 2 (instead of 4) GNSS receivers have been deployed for the absolute georeferencing.

Because the operations were started late in the afternoon, and the topometric phase lasted longer than expected, image acquisition could only be performed from 5 PM, which resulted in very dark images (even though exposition time was set at 5 ms, *i.e.* a value that, theoretically, would not allow the vehicle to move at more than 20 km/h to get non-blurry images).

Images have been corrected by histogram equalization

Fig 49. Original image (top left) and images after histogram equalization on the ranges [0; 90] (top right), [0; 60] (bottom left) and [0; 25] (bottom right).

Without any corrections of images, the contrast is too low and targets on the vehicle cannot be detected automatically. Experiments revealed that the first histogram equalization (which consisted in spreading the interval between [0; 90 bits] to the full 255-bit range) was sufficient to get about optimal level of automatic detections of targets.

After this correction, it can be observed that about 80% of visible targets are detected and correctly identified. This success rate is similar for 8 mm and 12 mm cameras.

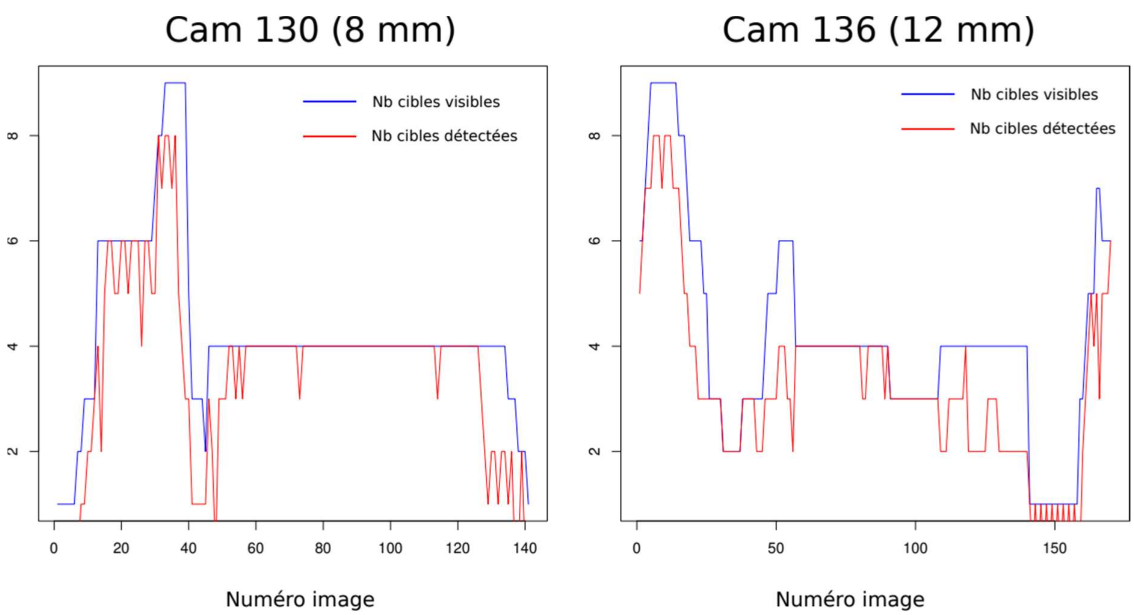


Fig 50. Number of detected targets per image (in red), as compared to the number of visible targets. Here 'visible' means that they could be potentially detected by the algorithm according to its nominal specifications: 20 px minimal size of butterfly pattern, and 60° maximal incidence angle.

On one image from each camera, GCPs are manually pinpointed (with, whenever possible, sub-pixel accuracy). In principle, if camera were rigorously stable, one set of 2D coordinates of GCPs collected on the first image of the sequence would be enough for the entire sequence of images. However, tests performed "at the office" revealed that the camera supports are subject to a residual motion, partly attributable to heat deformation (which could be drastically minimized if camera are powered up 30 to 45 minutes before the acquisition) and a white noise process (wind, cable vibration, etc.). For this reason, stationary targets have been placed right in front of each camera on Valeo test site, in order to assess the residual motion of cameras.

A motion of almost 3 px / minute in horizontal and/or vertical angle was detected on two cameras. The cause of this unexpectedly high motion is still unclear, but it may be assumed to be due to a slight torque applied on the back of cameras by RJ45 and BNC cables connecting them to the pole.

In the subsequent experiments, this problem was almost completely eliminated by using longer cables and tighter camera supports. For the Valeo dataset, the problem was partially offset by implementing image correlation algorithm for tracking GCPs motions in the image. The principle of the method is straightforward: pinpoint of GCP is performed by human operator on one of the images of each camera (it does not need to be the first image on the sequence, on the contrary, a temporally more central image in the sequence might be a better choice). Then a correlation on a small window (ranging from 50 to 200 px depending on the apparent size of the GCP in the image) is computed on all other images of the sequence. The peak of the correlation enables to find sub-pixel translation of the GCP motion in the image (due to the undesirable camera rotation). The updated 2D pixel coordinates of GCP on each image are then input in Comp3D software for the global adjustment.

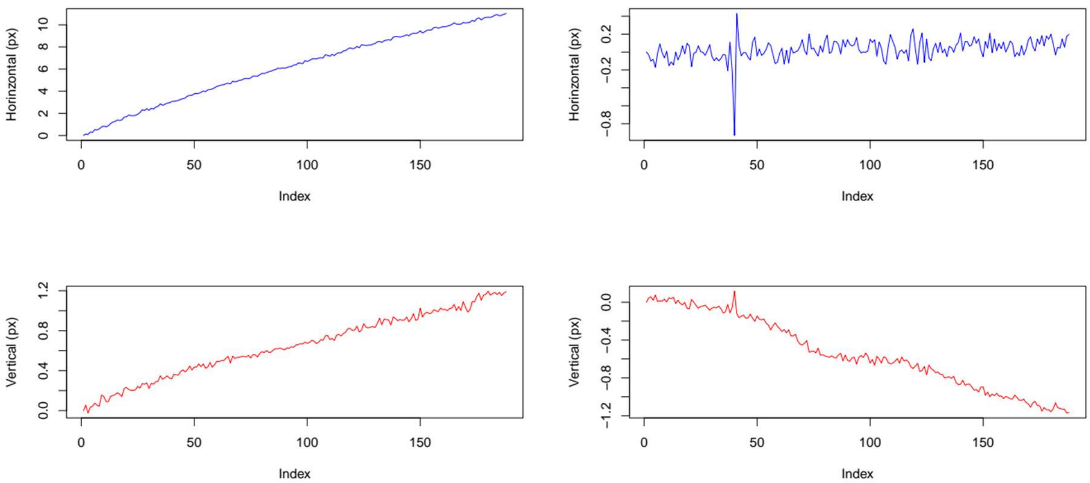


Fig 51. Two examples of (undesirable) camera angular motion at Valeo test site (abscissa in seconds) . Camera on the right is approximately stable in the horizontal angle (except jitter motion at epoch 40) but shows a small drift of 1 px in 3 minutes of acquisition. Camera on the left undergoes an almost linear drift of 10 px in the same time span.

Note however that the accuracy of this automatic tracking cannot be better than the accuracy of the initial pinpointing of the GCP. Therefore, the accuracy of the angle measurements towards a GCP from a camera has two independent components: one depends on the initial localization of center by the human operator – ranging from 0.5 to 2 px depending on the target pattern, and another one is introduced by the tracking process – ranging from 0.05 px to 0.5 px depending on the auto-correlation function of the window image surrounding the GCP. Since the autocorrelation depends on the background window, and not only on the GCP pattern, note that a GCP with a flat auto-correlation function (such as a laser scanner sphere target) may still be tracked accurately if its background provides a peaked auto-correlation.

The computation of the trajectory was done in a batch mode (*i.e.* all angular observations from cameras toward all targets at all epochs being adjusted together).

Being a very time-consuming process (and raising the questions of how observations to the GCP must be weighted if they are artificially replicated at each epoch), this batch processing mode strategy was subsequently discarded at UTAC test site, and all epochs have been processed independently, which did not reveal to be a noticeably sub-optimal solution compared to the ideal batch process, in addition to be more rigorous from the view point of independence of observations towards non-moving GCPs).

Evaluation was done on a piece of trajectory, running from epochs 830 to 849 (20 epochs, or 2.0 sec of motion of the AV). On this section, the vehicle was moving at 14.1 km.h^{-1} and the distances between the vehicle and the cameras were as shown in Tab 7.

Camera	**min. distance (m)**	**Max. distance (m)**
C_1	28.1	31.4

C_2	22.8	23.5
C_3	19.9	26.8

Tab 7. Distances between each camera and the AV.

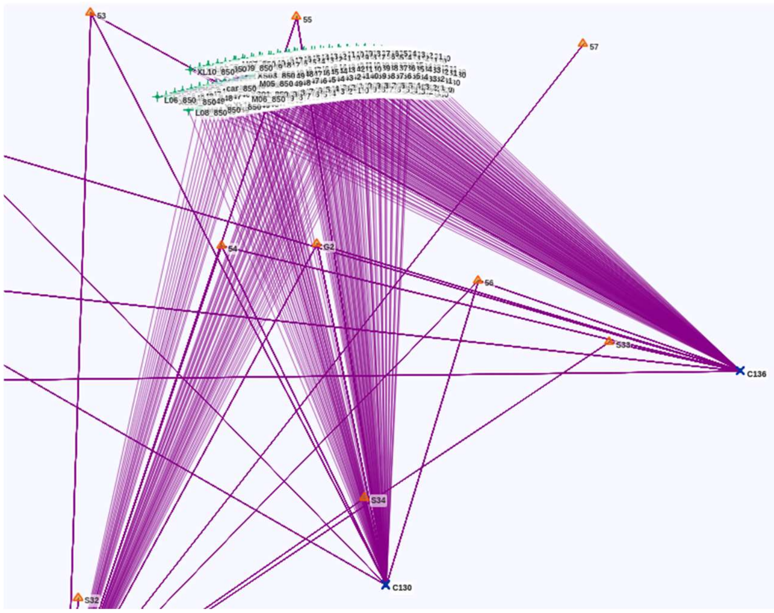


Fig 52. Computation in batch strategy of a piece of trajectory. Batch mode is very consuming in memory and computation time; this simple piece of trajectory required to form a least squares system of 1305 unknown parameters and 2674 observations.

The square root of squared residual mean error after adjustment (σ_0) is 1.12. This is out of the Khi-2 validation test ([0.9522; 1.0524]) meaning that all prior accuracy standard deviations input into the computation are somehow under-estimated by a factor of 1.12. This can be easily corrected, and still contributes to show that all the sensors are roughly behaving as expected (with 10% error in expected accuracies).

Having a set of 3 cameras provides an internal validation of the trajectories. Though the bundles issued from only 2 cameras would always intersect to some degree of satisfaction, at least in the horizontal plane, this is not true in 3D space were a pair of cameras is already providing redundancy because two bundles do not necessarily intersect, and the residual distance between them can be used as a metrics of internal coherence. However, given that the AV and the cameras are essentially coplanar, it is necessary to have a third camera for a truly satisfying internal coherence. If cameras were not oriented correctly, bundles to 3D points in space would not intersect properly, and this would be clearly visible in the value of σ_0. Moreover, if the cameras were not synchronized, bundles to 3D points in space would intersect properly only on non-moving targets; any pair of bundles from 2 cameras would not be consistent with the homologue bundle issued from the third one, towards each target on the vehicle (except when the vehicle is stopped), and the intersection error would be approximately equal to the desynchronization error multiplied by the speed of the AV. Therefore, this usage of a redundant third camera is important to get an internal consistency assessment: comparison of trajectories generated respectively by cameras C_1 and C_2, and cameras C_2 and C_3, separately, resulted in a planimetric error of about 2 cm, which is consistent with the respective uncertainties of trajectories estimated with only two cameras.

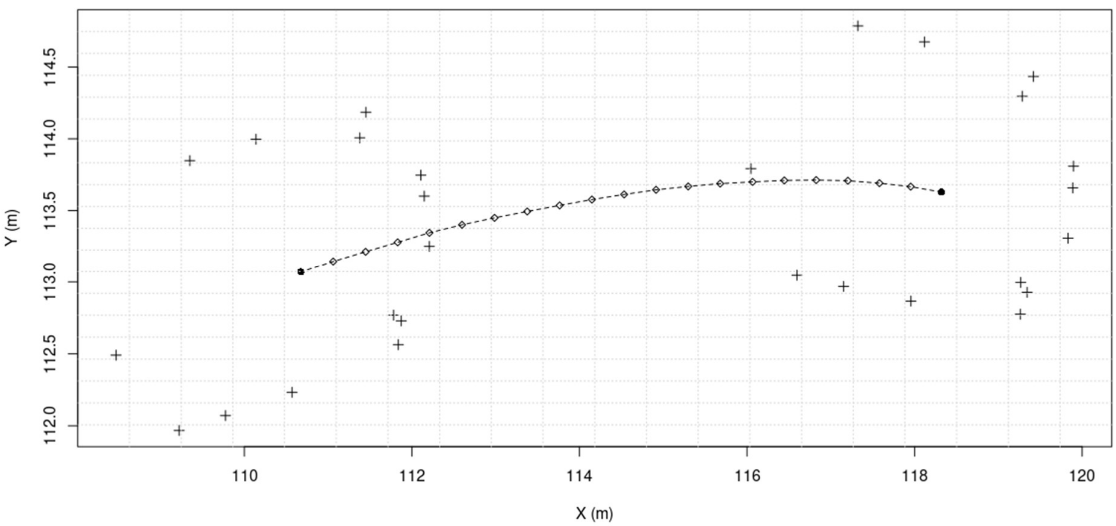


Fig 53. The piece of trajectory used for the evaluation. Two (non-successive) positions of the vehicle are depicted by their rigid-block of targets with cross, and vehicle reference frame center (black dot). Successive positions between them are shown on the dotted line. Coordinates are relative to the test site.

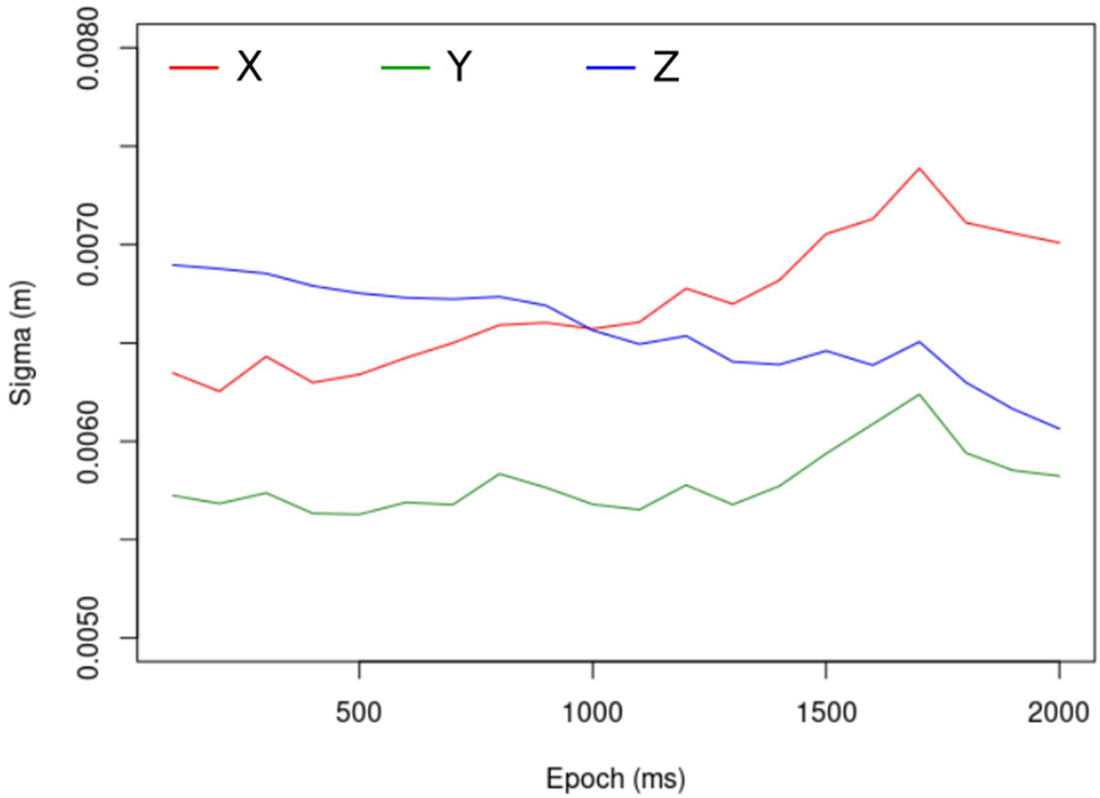


Fig 54. Standard deviations of errors in X, Y and Z directions for the evaluated piece of trajectory

The uncertainty (1σ) along each direction axis is close to 6 mm on average. The average 3D uncertainty is **11 mm**, decomposed as follows: 8.9 mm in horizontal accuracy (σ_h) and 5.6 mm in vertical accuracy (σ_v). Uncertainty is evenly distributed in all the three axes of the reference frame.

Analysis of covariance matrices revealed that vertical and horizontal components are mostly decorrelated. X and Y coordinates, are moderately correlated, and the horizontal ellipse is about 50% in the orthogonal direction with respect to the average line of sight of cameras. As a consequence, the uncertainties in depth are slightly bigger, with 8 mm uncertainty, as compared to the uncertainties along the orthogonal axis, with only 3 mm. This is however not so critical as the ellipse are not so far from being circular, in addition to the fact that final configuration at UTAC test site will involve more cameras and then more viewpoints, which will guarantee more evenly distributed uncertainties, at least in horizontal directions, and more decorrelated errors on coordinate components.

The uncertainty σ_θ on the heading of the vehicle is varying between **0.04°** and **0.07°** along the trajectory.

Having all the epoch being computed in a batch, it is possible to analyze correlation between coordinates of successive positions.

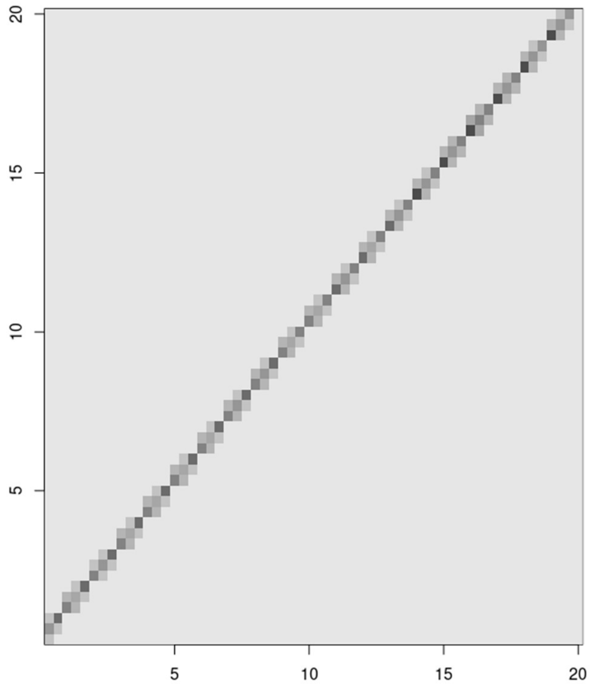


Fig 55. Absolute values of the covariance matrix of the estimation on 20 epochs. Parameters are sorted as follows: (x1, y1, z1, x2, y2, z2, … x20, y20, z20). Darker pixel on coordinate (i, j) means that parameters i and parameters j are more strongly correlated or anti-correlated. Brighter values mean decorrelation of parameters.

Fig 55 shows that parameters are significantly correlated in a 3x3 block scheme, which means that the only significant correlation in the solution is found between the 3 coordinate components of each individual position of the vehicle. No significant correlation is found between epochs. This can be further verified by measuring the uncertainty on the relative distance between positions on the trajectory. A result of 5 mm was found, which is close to the individual uncertainty of points, revealing that the errors on the trajectory may be assumed to be nearly a white noise process.

These results meet the requirements of the system. Final POC at UTAC site will test its capabilities on a 1:1 scale model and at higher speed.

2.2.3.3.3. Real test 2

This second real-scale test was conducted at UTAC test track (Linas-Montlhéry) on a 25-m diameter round-about crossing (TEQMO circuit), on January 30th, 2024.

Weather conditions were ideal, with homogeneous cloud cover, no rain, and almost no wind.

The disposition of cameras, GCPs and GNSS receivers was as depicted on Fig 56.

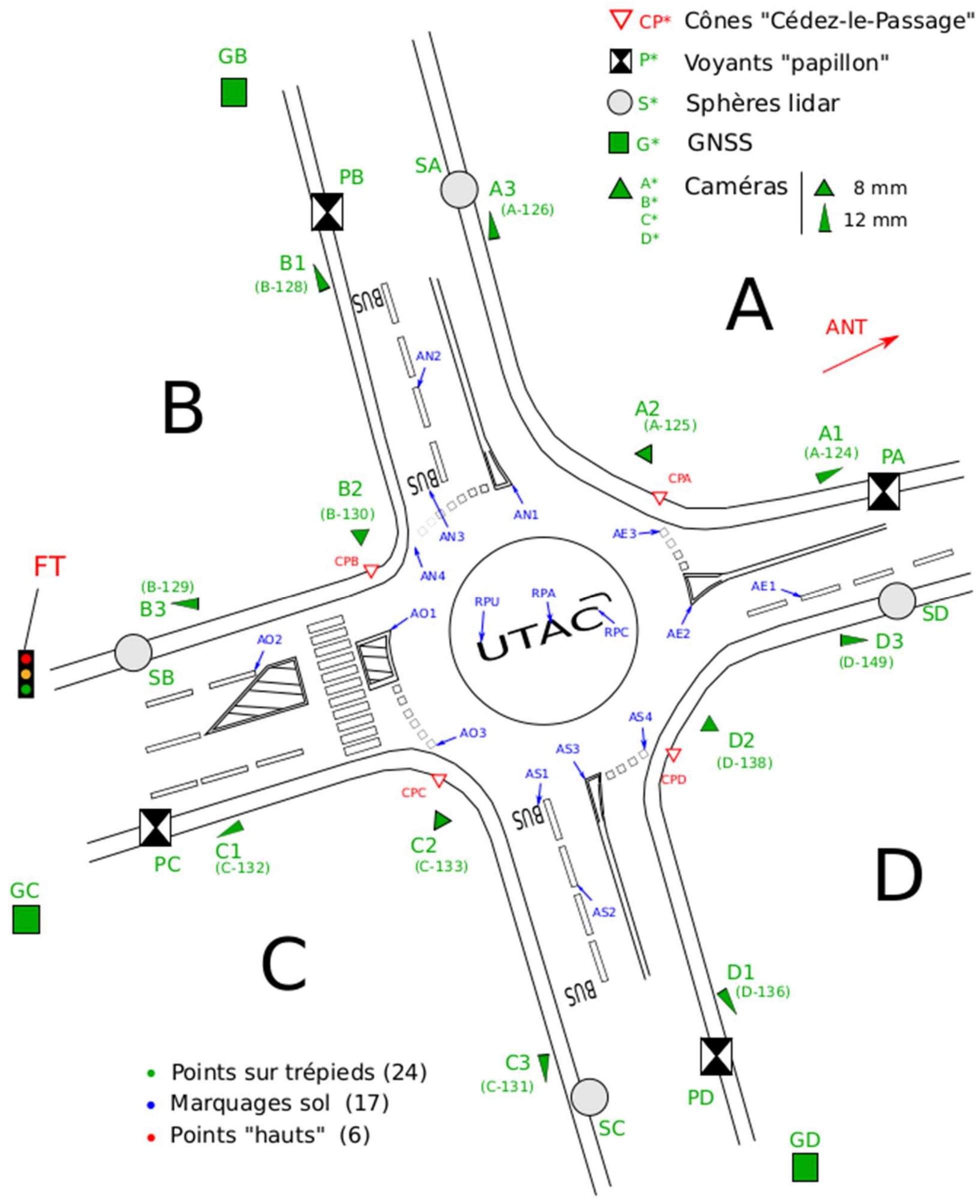


Fig 56. Disposition of GNSS (green squares), GCP targets (blue points, black and white targets and spheres) and cameras.

Fig 57. Set up and topometric phases at UTAC. Vehicle is stationary for measuring relative coordinates of targets (rigid-block). The view point is seen from a position near long-focal distance camera C3 (South-East sector C). From left to right, tripods of cameras B1, A3, A2, A1 and D2 are visible (cameras are not installed yet, as reflectors are used instead to measure the coordinates of points where optical centers of cameras will be placed during image acquisition). GCPs are measured on the ground (such as corners of letters in 'BUS' sign or intersections on pedestrian crossings.

The 4 sectors of the intersection are denoted A, B, C and D, in counter-clockwise, starting from North-East corner. Sector A is composed of:

- 1 pole of 3 cameras: one 8 mm camera (denoted A2) and two 12 mm cameras (denoted A1 and A3) all synchronized by BNC cables to a low-cost GNSS receiver, and connected to a common PC (disposed close to A2)
- 1 laser scanner spherical target SA and 1 black and white target PA, respectively, behind the 12 mm cameras. Spherical targets have the advantage of being equivalently visible from all directions, but it may be difficult to pinpoint accurately their center when placed close to a camera. Therefore, those targets are positioned at a long distance from cameras from which they are susceptible to appear, and are mixed with black and white targets, whose pinpointing is easier.
- 1 conic target on yield way vertical sign, to provide a point on top of images for camera orientations.
- 1 high-precision GNSS receiver GA, for absolute positioning (note that topometric reflector is positioned below the antenna, which can be used as an additional target for camera orientations.

All three other sectors B, C and D are organized similarly. Each sector was powered by an individual electric terminal. About 50 m of BNC and RJ45 cables are connecting each PC to the cameras of its sector. Note that no wire connection is established between sectors, as stated by the project requirements.
The GCP are further completed by 17 horizontal road signs (in blue on Fig 56) and 2 long-range targets (a point on a traffic signal FT and a 50 m-high antenna ANT).

Note that each corner is covered by swamps, making it impossible to set up tripod out of road edges.
Based on the results of Valeo POC, the final rigid-block configuration for the targets on the vehicle is depicted on Fig 58. The total number of 15 targets are used (this number of 15 is chosen in order to optimize the size and number of bits on targets, which is important to guarantee optimal detection and decoding performances) with a symmetrical disposition (6 on each side of the vehicle, 2 at the rear and 1 on the engine hood). 13 targets are magnetic, 2 targets are printed on paper and attached with rubber tape on the rear passenger seat windows.

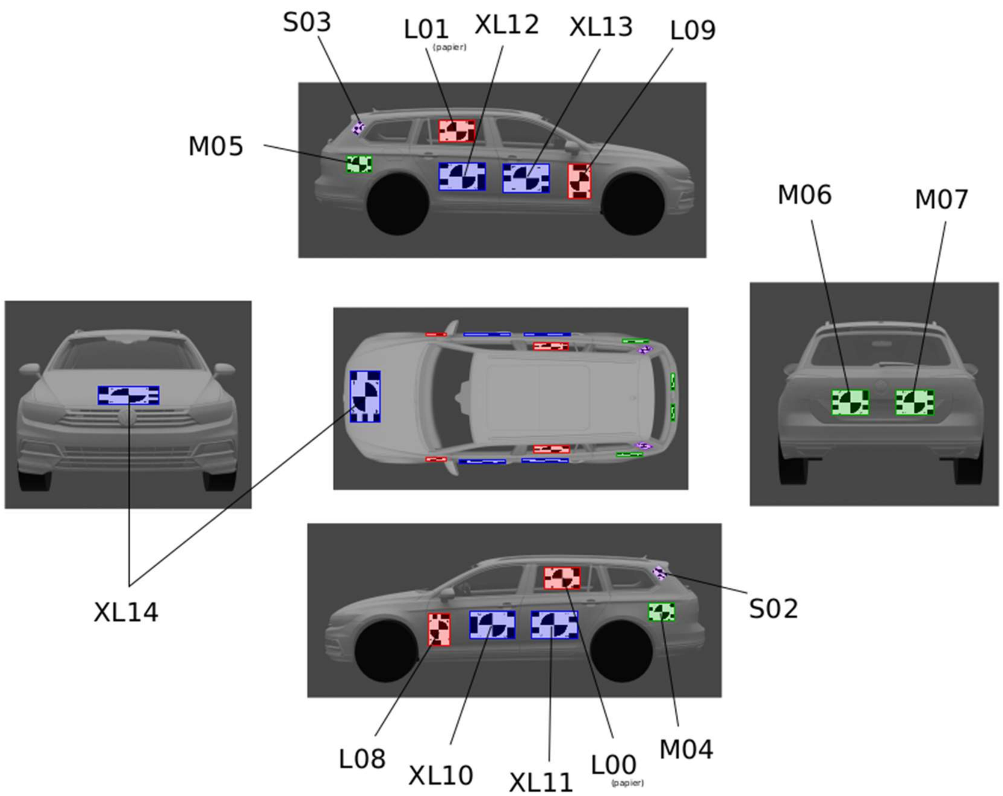


Fig 58. Final configuration of rigid-block of targets on vehicle.

Topometric phase took 5 hours and required 4 operators. 6 hours of GNSS observations were collected. 6 absolute orientation observations have been performed with the gyrotheodolite. At the end of this step, coordinates of GCP and optical centers of cameras are determined with 2 mm accuracy in relative and 4 mm accuracy in absolute position. The targets in the rigid-block have a (relative) accuracy of 0.2 mm. Here also, topometric phase is reinforced by laser scanner.

Acquisition of images could be started at 03:49:12.000 PM.

A few technical problems during the experiments resulted in the following negative consequences:

- a hardware connection problem between the PC of sector B and the two long-focal cameras (B1 and B3) made them unavailable. No images are available from these two cameras, reducing the total usable number of cameras to 10.

- an error in the parameterization of the cameras caused them to discard one image out of two, then capturing image at a 5 Hz frequency instead of the nominal 10 Hz frequency. For a some (still unclear) reason, the cameras of sector A have been capturing images at 6 Hz frequency.

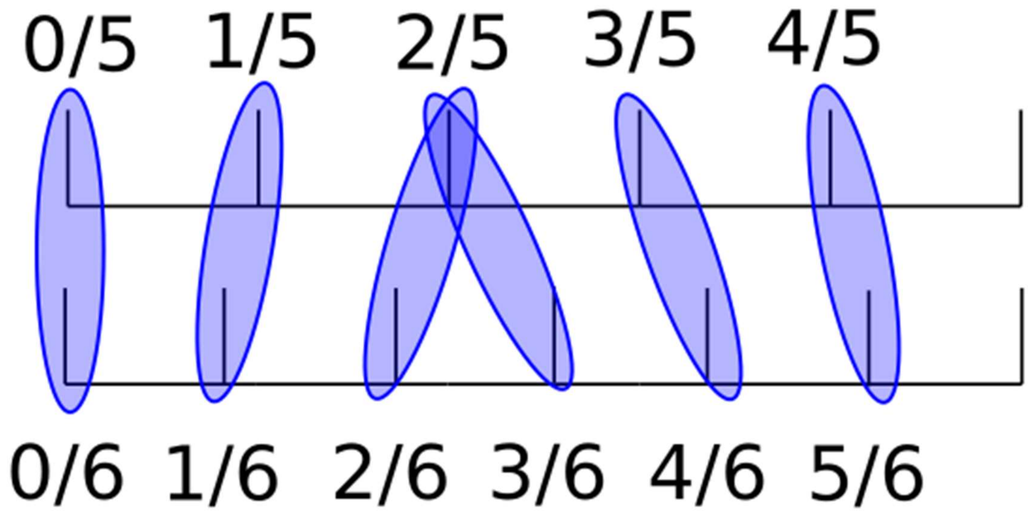


Fig 59. Illustration of desynchronization between cameras of sector A (running at 6 Hz frequency) and cameras of sector B/C/D (5 Hz). On average, 1 out of 6 images of sector A are synchronized with B/C/D, and 1 out of 5 epochs in the 5Hz trajectory can be additionally constrained by measurement issued from sector A. Approximate matching of epochs between A and B/C/D could be done, but at the expense of an unacceptable loss in geometric accuracy (up to 7 ms of error, or 7 cm for an AV moving at 36 km/h).

As a general consequence, 5 out of 6 images are out of sync between cameras of sector A on one hand and B/C/D on the other hand. This result in a total of 7 cameras (from 3 independent poles) is usable for computing the trajectory. The solution could be additionally constrained by $1/6^{th}$ of images issued from sector A, which makes the effective number of usable cameras at 7 + $1/6^{th}$ x 3 = 7.5. However, since it is not particularly interesting to compute a trajectory with non-homogeneous precision (1 out of 5 epochs would be more tightly constrained with the additional measurements provided by bundles issued from cameras of sector A, as illustrated on Fig 59), it was instead decided to proceed as described below.

The main trajectory (delivered to Valeo for comparison) is computed from images issued from sectors B, C and D, and is qualified with intrinsic consistency estimators (squared residual error, khi-2 test, etc.). Additionally, a separate trajectory is computed from images issued from sector A (with however less accuracy, given that the three cameras A1, A2 and A3 offers a poorer stereoscopic distribution and less redundancy). Both trajectories are then geometrically compared (note that because trajectories are out of sync on 80% of epochs, this comparison requires linear interpolation between points to resample both trajectories at the same frequency). Since they are almost independent (simply relying on the same GCP coordinates, whose quality can be certainly guaranteed by the topometric phase), the comparison provides a good estimation of the system geometric accuracy.

In total 3 minutes of images have been collected, and the system is evaluated on the 5 following (typical) scenarios, for a total of 79.8 seconds (Tab 8):

#	Scenario	Duration (s)	Speed (km/h)
S1	1 turn around round-about	15.8	20.1
S2	Straight-line	3.0	36.3
S3	2 turns around round-about	31.2	20.8
S4	Backward driving	20.0	7.6
S5	Slow speed + stop	9.8	< 6.8

Tab 8. The list of 5 scenarios recorded by the system

Automatic detection of targets is performed with a total of **11 609** detections, which represents between 70 and 80% of all targets visible from camera.

Manual validation of detection on scenario S1 enables to observe that the ratio of false detection is **0.4%**, distributed evenly between false detections on horizontal road signs (38%), vehicle license plate (23%) and other parts of the vehicle (38%). The first two types of false positive detections can be easily mitigated by (1) by covering the license plate and (2) by *pre-training* the detection algorithm on a set of images without vehicle, hence enabling to identify areas in images where false detection are likely to occur, in order to remove them from the images taken with the vehicle. Moreover, outlier detection can be easily removed in second-pass, by comparing mutual position of detected targets and rigid-block coordinates.

No decoding error was reported, meaning that all detected targets have been associated to their correct code.

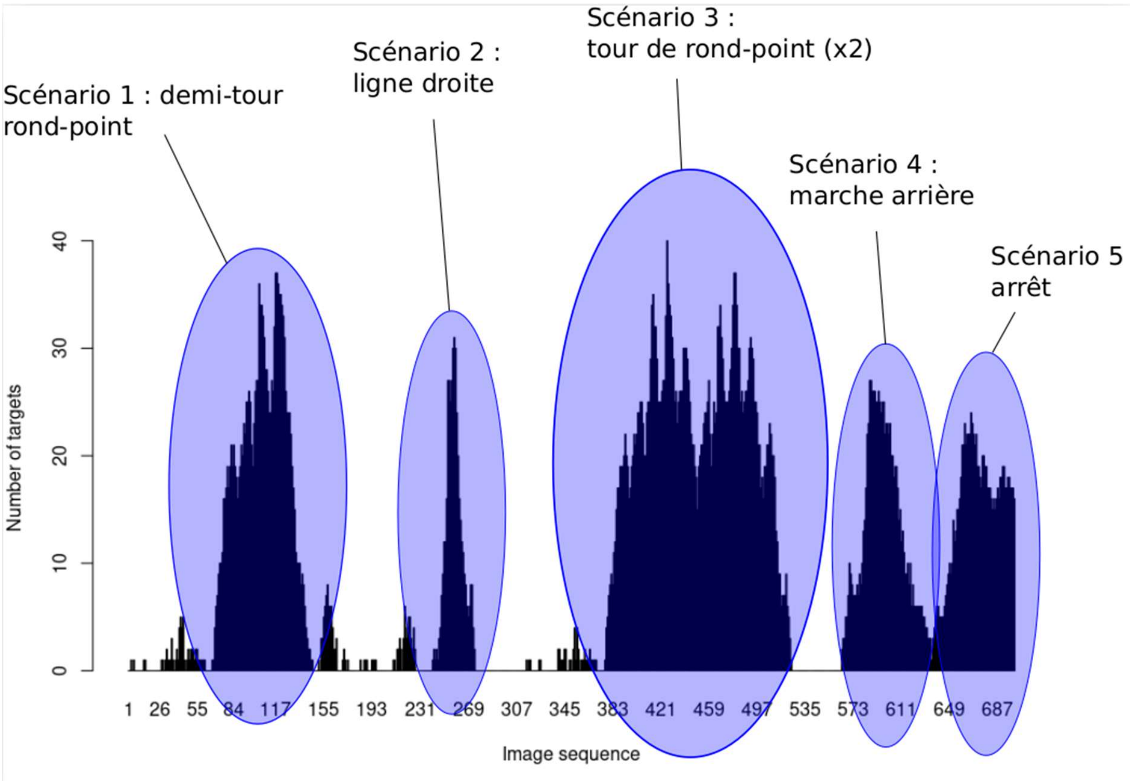


Fig 60. Distribution of target detections on driving scenarios S1 to S5. Up to 40 targets are detected simultaneously from sector B/C/D.

The ratio of detection for each target is given on Fig 61. Though target on the hood can be beneficial for a very accurate estimation of vehicle attitude, it seems to be particularly difficult to detect. This may be partly explained by the high incidence angle and curvature of the target.

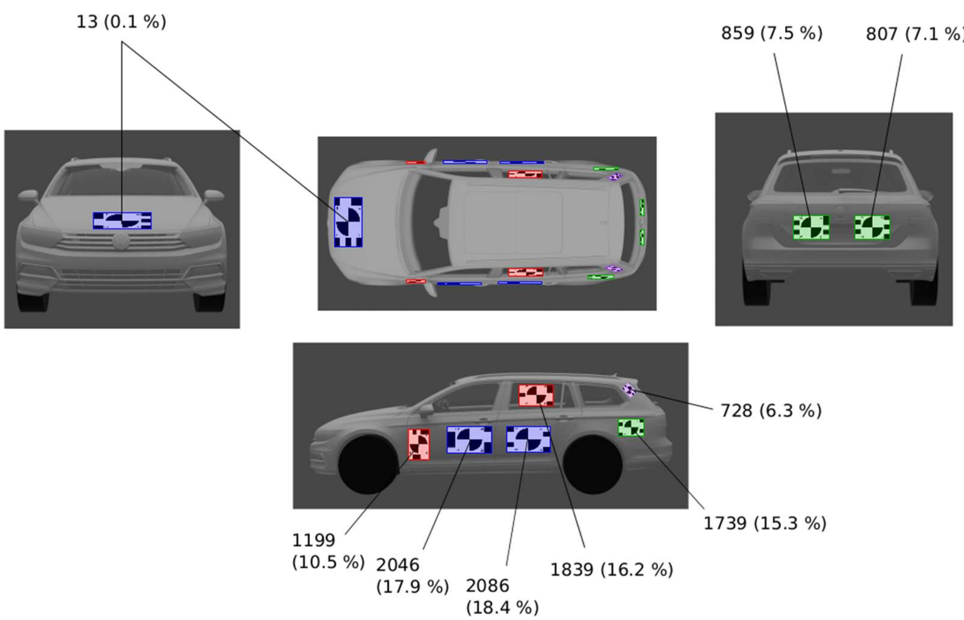


Fig 61. Ratio of detected target: each percentage indicates the ratio of the number of detections of a specific target (left and right sides of car being counted together) out of the total number of detections.

Unlike Valeo test, as shown by the graphics on Fig 62, cameras were stable during the acquisition, hence only measurement towards GCPs on the first image of the sequence could be replicated to compute the orientation of all cameras.

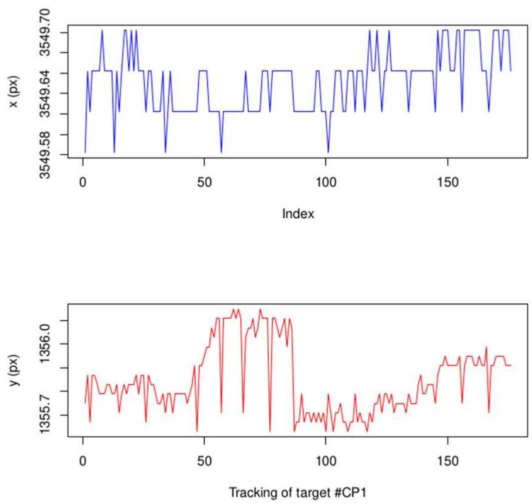


Fig 62. Motion of a GCP target (in px) in both horizontal and vertical angles from one 12 mm camera, showing that the net motion in 3 minutes is below 0.1 px, with a peak-to-peak value of 0.3 px at most. All other cameras have similar stable patterns.

Pixel coordinates are then converted into 3D bundles, corrected from camera distortions, and then input into Comp3D software for global adjustment. As explained in the previous section, computation is done on a per-epoch basis. It required 60 seconds for 508 epochs on an Intel(R) Core(TM) i7-11800H @ 2.30GHz processor with 8 cores and 32GB of RAM, which amounts to about 10 epochs processed per second (as it corresponds to the 10 Hz nominal frequency of images, this opens up possibilities for real time implementation, though target automatic detection of images is still the bottleneck with a few seconds of processing per image…). In total **125 000 equations** are solved to determine **81 000 unknown parameters**.

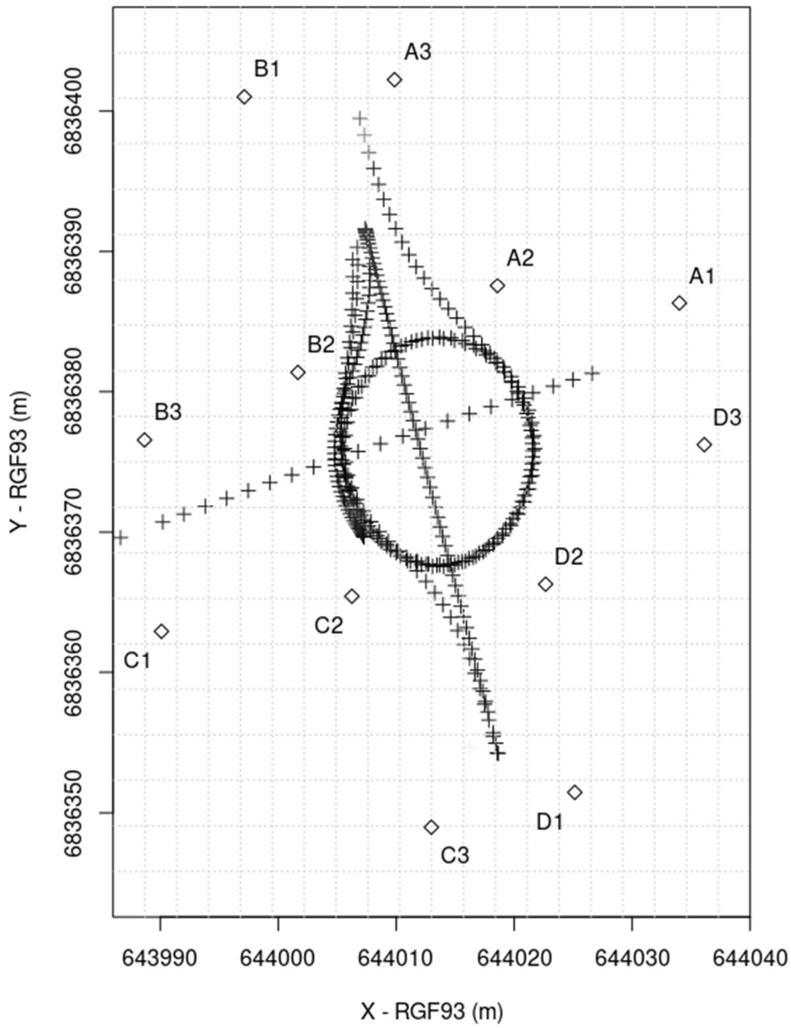


Fig 63. Resulting trajectory after global adjustment. Coordinates are provided in RGF93 Lambert 93 cartographic projection

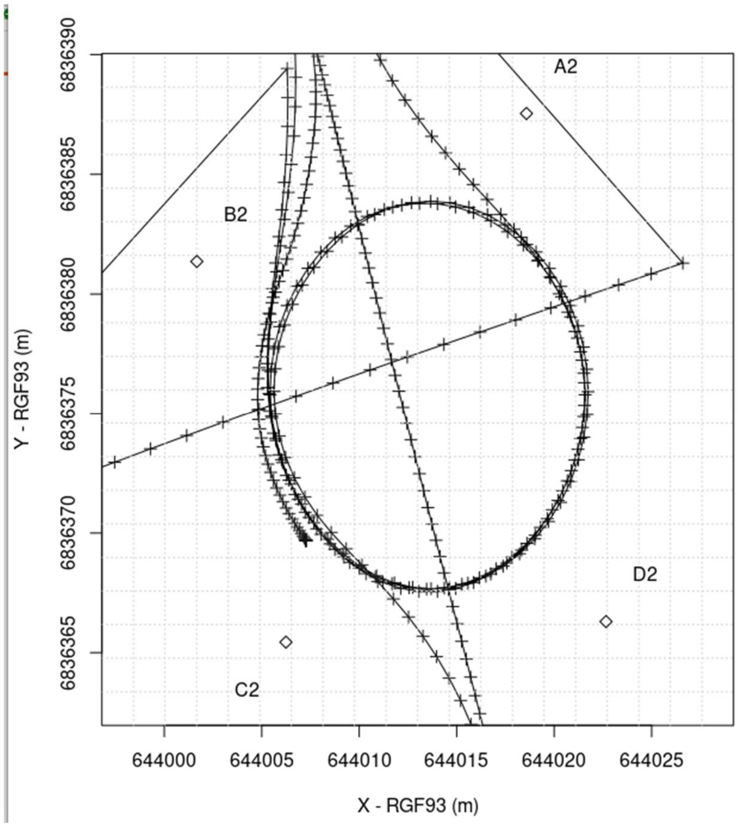


Fig 64. Resulting trajectory after global adjustment: zoom on the central area. Coordinates are provided in RGF93 Lambert 93 cartographic projection

The σ_0 factor is almost always between 1.0 and 2.0, which means that the input uncertainties are relatively well quantified.

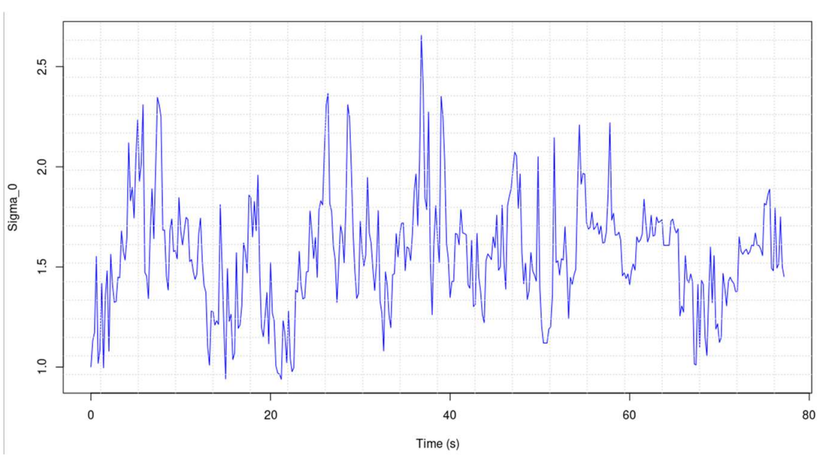


Fig 65. Evolution of σ_0 on the full trajectory

Standard deviations of positions are always systematically below 10 mm, except on sections where the vehicle is about to enter or to exit the round-about.

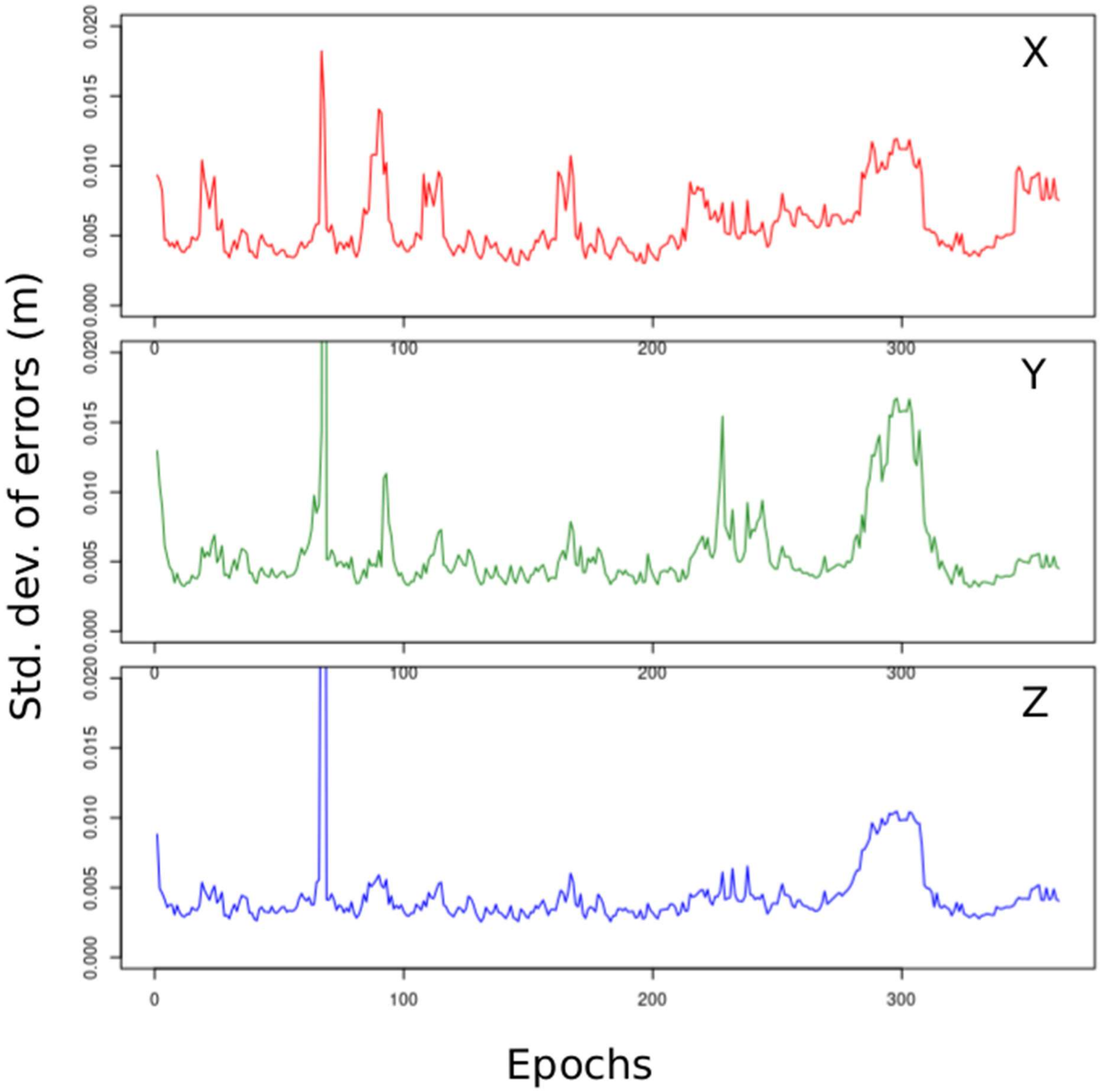


Fig 66. Standard deviations of positions (in m) in X, Y and Z directions **in Lambert 93 cartographic projection**

The mean RMSE of the trajectory is **5.8** mm in both horizontal directions, and **4.6** mm in vertical direction.

The standard deviation error of the heading angle is **0.04°**.

The plot of error ellipses (Fig 71) reveals that the uncertainty is mostly directed in the radial direction, with respect to the roundabout.

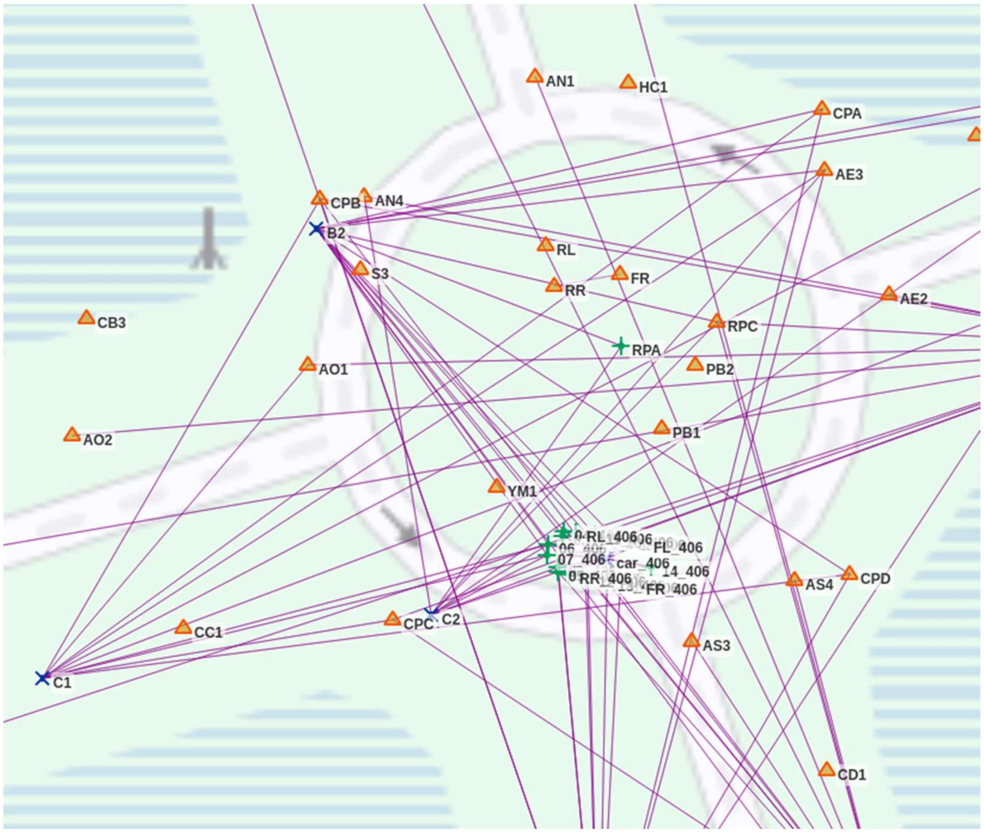
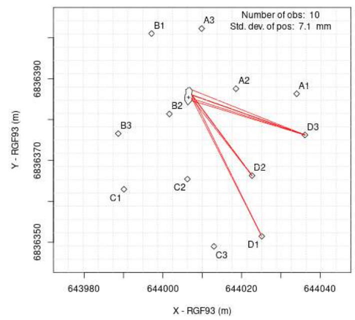
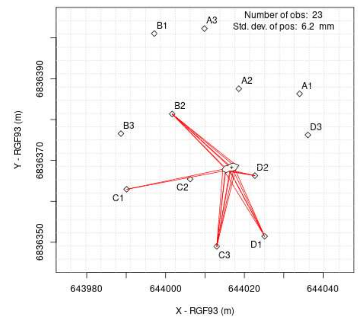


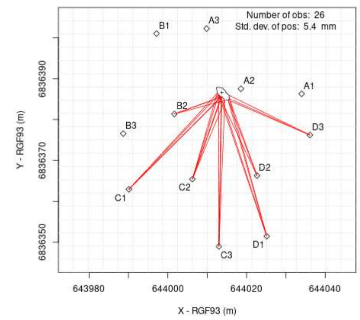
Fig 67. Adjustment of the epoch 406. Vehicle targets are depicted in green, cameras in blue.

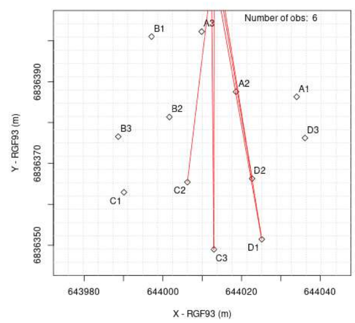
A series of plots here after shows the evolution of the number of observations and the standard error of positioning of the vehicle for different positions in the scene. It can be seen (as confirmed by ellipse plot) that position is slightly less accurately estimated (15 mm uncertainty against a nominal 5 mm on the remaining of the trajectory) in front of sector C. This may be explained by the absence of cameras in sector A.

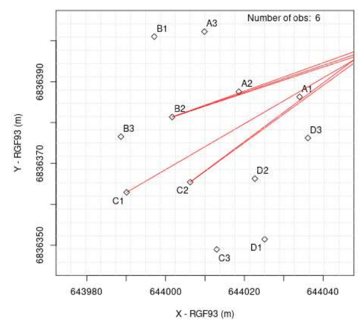
Simulations showed that if all cameras would have been operative, the accuracy of the estimated error would likely be at its lowest value of 5 mm throughout the trajectory.

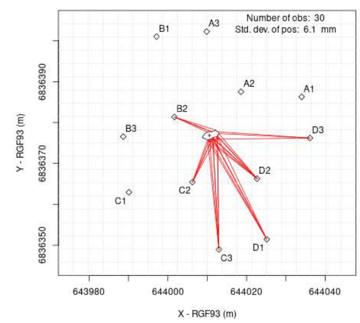

Number of obs: 10
Std. dev. of pos: 7.1 mm
Y - RGF93 (m)
X - RGF93 (m)

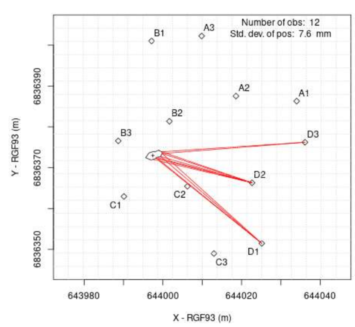

Number of obs: 23
Std. dev. of pos: 6.2 mm
Y - RGF93 (m)
X - RGF93 (m)

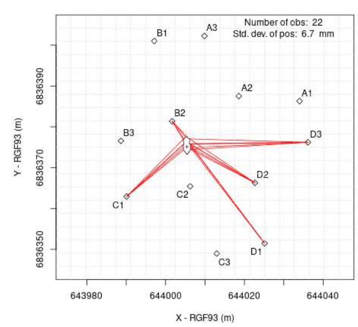

Number of obs: 26
Std. dev. of pos: 5.4 mm
Y - RGF93 (m)
X - RGF93 (m)

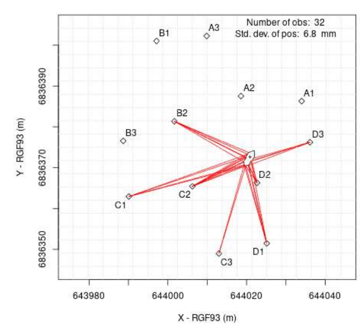

Number of obs: 6
Y - RGF93 (m)
X - RGF93 (m)


Number of obs: 6
Y - RGF93 (m)
X - RGF93 (m)


Number of obs: 30
Std. dev. of pos: 6.1 mm
Y - RGF93 (m)
X - RGF93 (m)


Number of obs: 12
Std. dev. of pos: 7.6 mm
Y - RGF93 (m)
X - RGF93 (m)


Number of obs: 22
Std. dev. of pos: 6.7 mm
Y - RGF93 (m)
X - RGF93 (m)


Number of obs: 32
Std. dev. of pos: 6.8 mm
Y - RGF93 (m)
X - RGF93 (m)

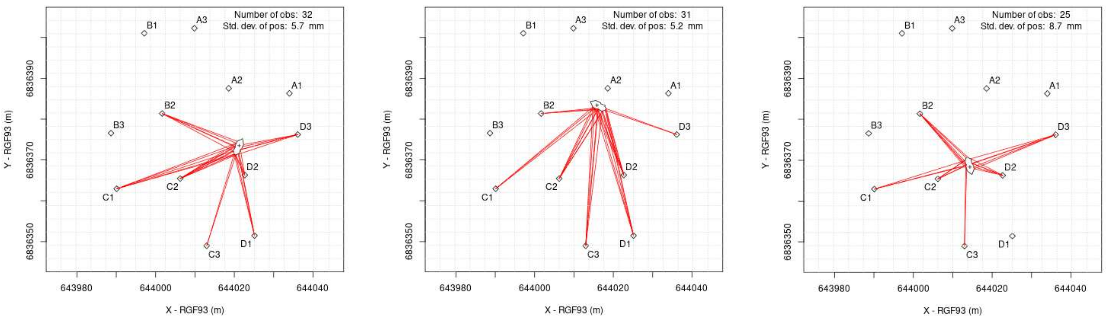


Fig 68. Illustration of bundles from cameras of sectors B, C and D, towards the targets on the vehicle. Standard error of positioning is varying (roughly between 5 and 10 mm) depending on the position of the vehicle in the scene. On average, the vehicle is seen by 20 bundles from cameras

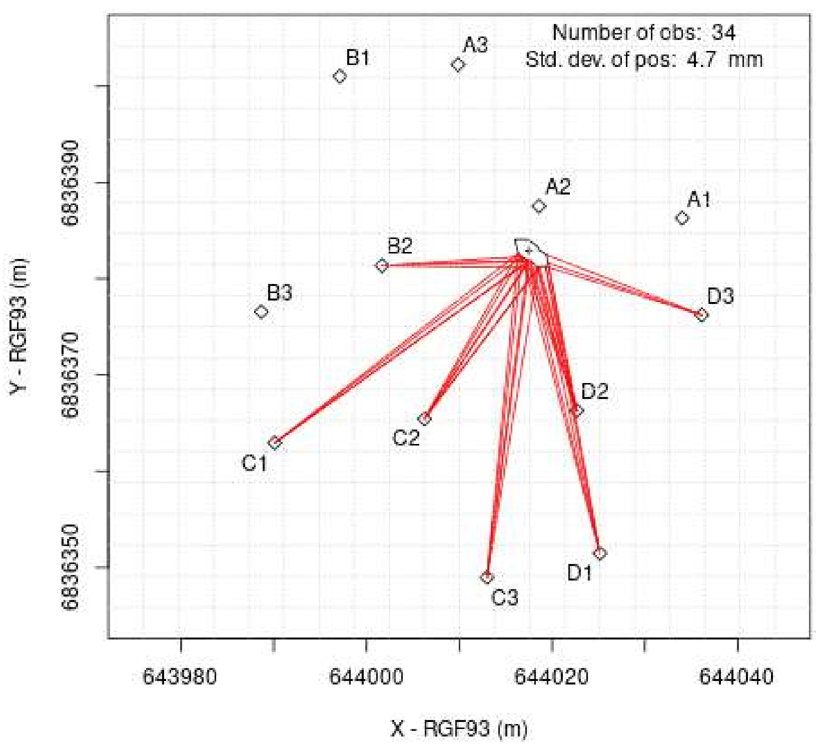


Fig 69. Illustration of bundles from cameras of sectors B, C and D, towards the targets on the vehicle. Standard error of positioning is at 4.7 mm, for a total of 34 observations.

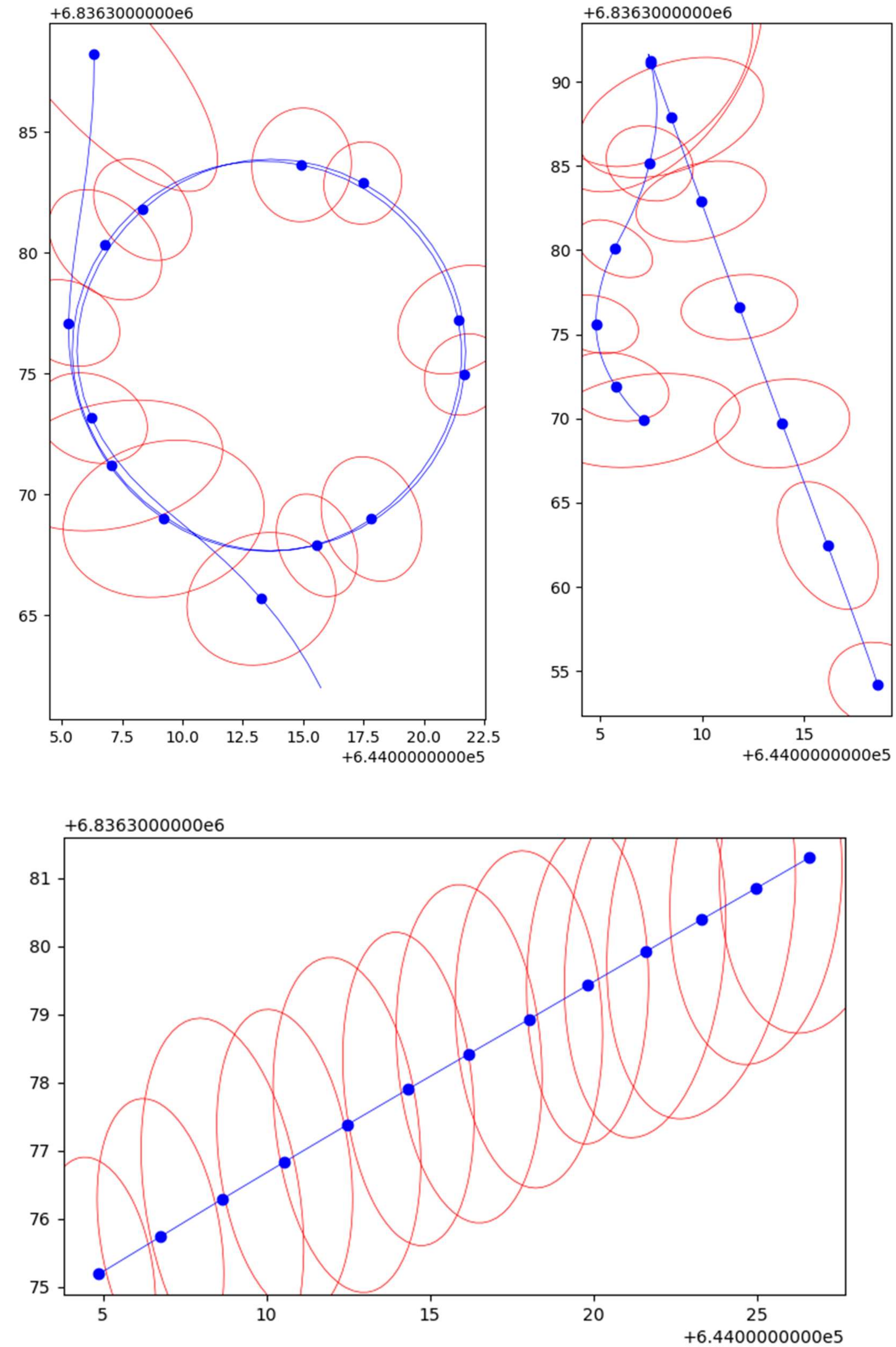


Fig 70. Representation of planimetric uncertainty ellipses (unit scale is x1000, i.e. 1 m of plot axis for 1 mm of uncertainty) for the scenarios S3, S4-S5 (top) and S2 (bottom). To keep the plot readable, one ellipse every 10 epochs is shown for scenarios S3 and S4-S5

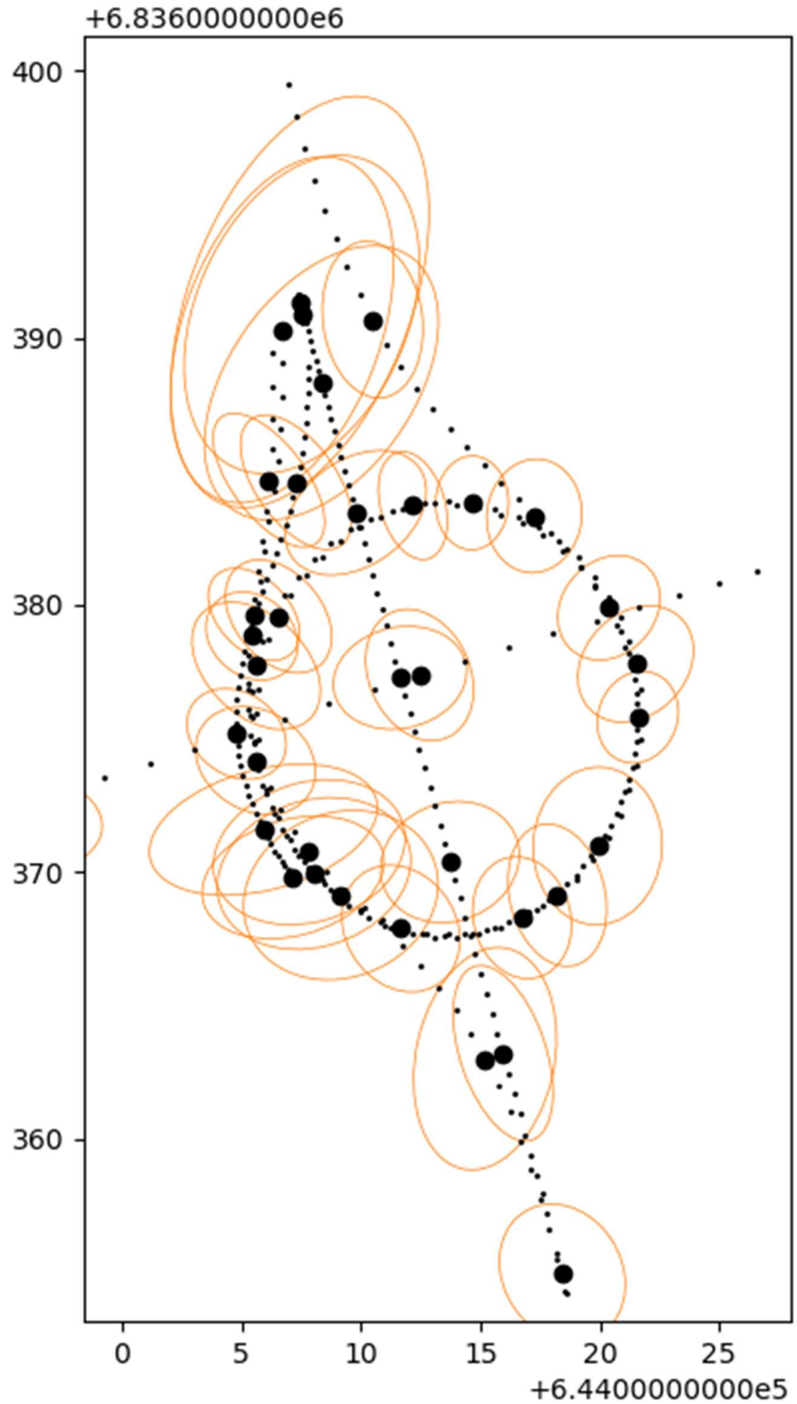


Fig 71. Error ellipse on all scenarios (ellipse scale: 1 m = 1 mm). One ellipse is plotted every 10 trajectory points.

The cumulative distribution function of standard deviations of positions along the trajectory shows that most of points (80%) have a sub-centimetric 3D accuracy. The few points having more than 2 cm uncertainty are located on the entrance/exit lanes of the round-about, where the stereoscopic factor of the system of camera is quite poor.

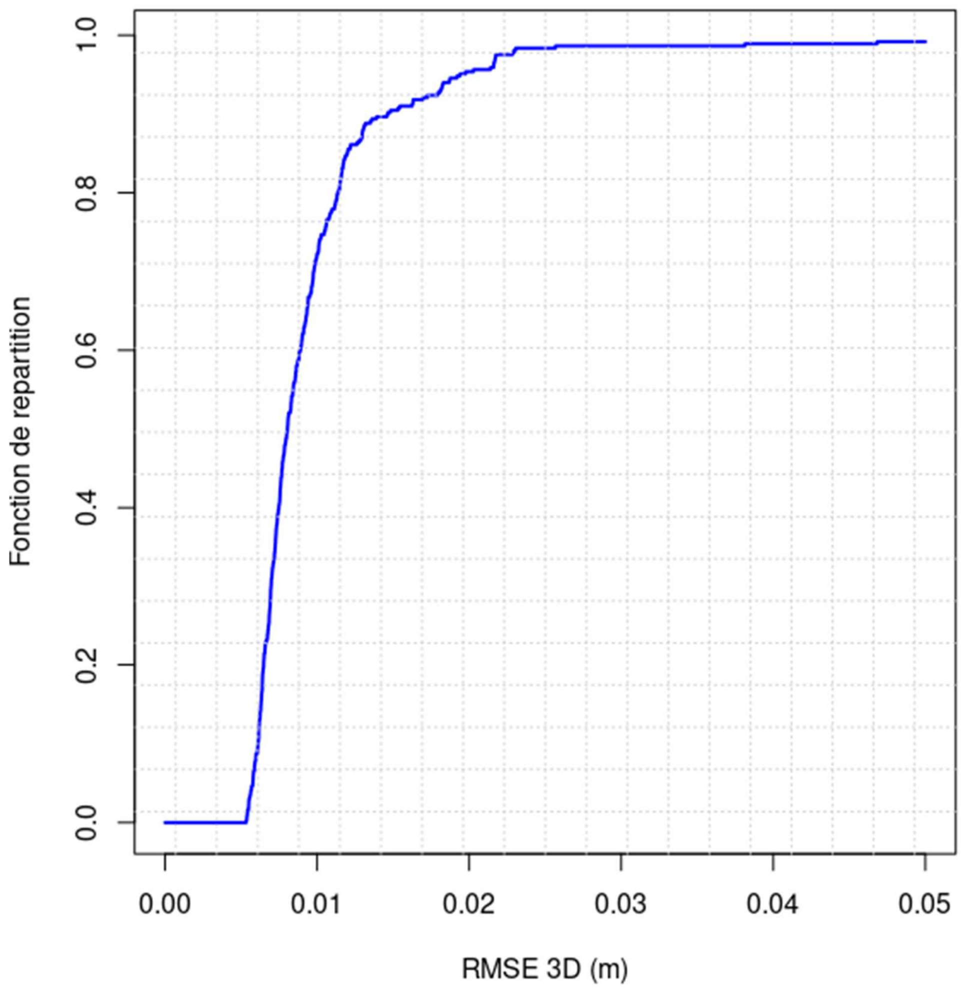


Fig 72. Cumulative distribution function of 3D uncertainty of positions.

Moreover, it can be clearly seen on this plot that all points have at least 5 mm uncertainty, which corresponds to the global georeferencing error.

Though the problem is frankly redundant (therefore allowing an unbiased internal evaluation of the precision of the trajectory estimation), it may be interesting to compare the results with the independent trajectory computed from sector A cameras (recall that cameras from A have not been used so far in the determination process).

Evaluation is performed specifically on the scenario S3.

Since the two trajectories are out of synchronization, comparison between them is purely geometric. More precisely, it is done by matching the two trajectories with a discrete Fréchet distance, as illustrated on Fig 73. Computation was done with *Tracklib* library developed at Lastig (IGN) laboratory.

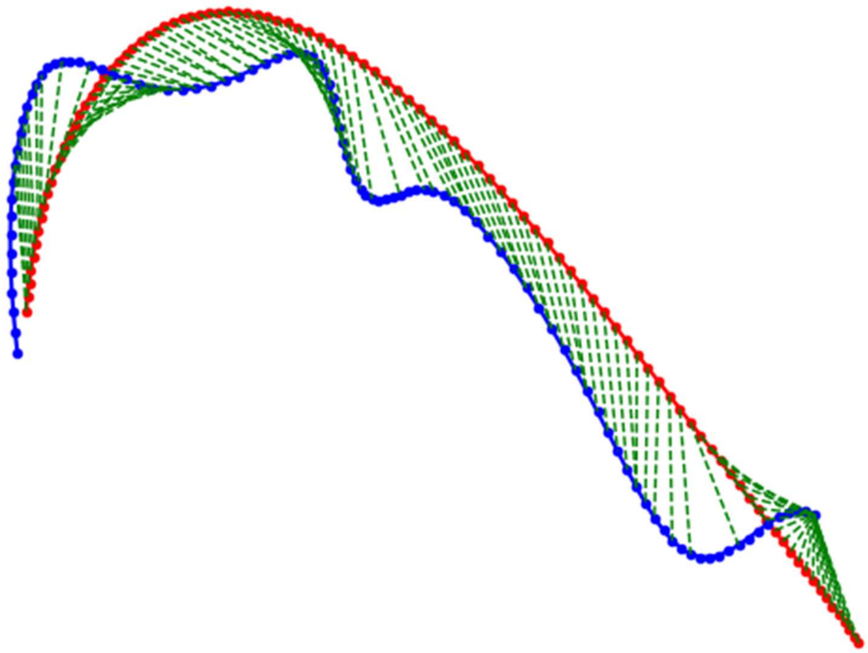


Fig 73. Illustration of the matching of two (non-synchronized) trajectories with discrete Fréchet algorithm: the maximal length of green segments is minimized, while imposing a monotonic travel along both trajectories. This process enables to match complex trajectories (such as the one of scenario S3), and unlike nearest neighbor matching, offers a nearly unbiased estimate of the geometric similarity of the two trajectories being compared.

Comparison of trajectories resulted in the following plot of differences:

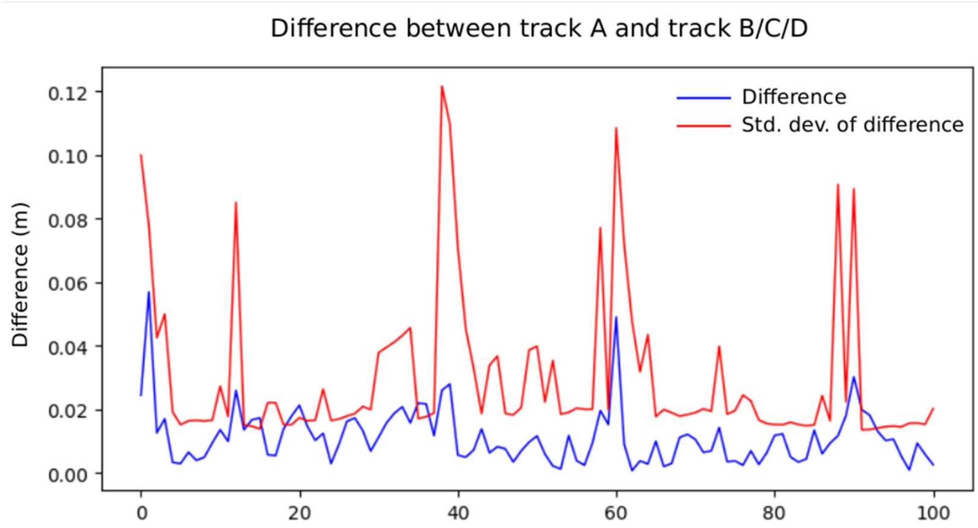


Fig 74. Comparison between trajectories estimated separately and independently, on one hand from sector A and on the other hand from sector B, C and D. The blue curve depicts the 2D distance between trajectories (often between 10 and 25 mm). The red curve is the 2σ confidence band on this difference (derived from individual uncertainties on the two trajectories). Abscissa is the number of epochs.

The comparison shows that the distance between the two trajectories is most of the time below 2 cm, and almost always below the 2σ confidence band describing the intrinsic uncertainties of both individual trajectories. This shows that the difference between the two estimated trajectories is not statistically significant in light of predicted accuracies.

Moreover, since 1 epoch out of 5 in cameras of sector B/C/D is synchronized with an epoch of sector A (Fig 59), two sub-trajectories at 1Hz can be extracted for geometric comparison, thus cancelling uncertainties resulting from the time registration between sector A and B/C/D. Here also, a 10 mm error between the two sub-trajectories has been measured.

This contributes to validate the determination process.

A note on interpolation

Linear interpolation can be performed to oversample the trajectory at the nominal frequency of 10 Hz. However, a first-order Taylor expansion shows that it could potentially result in a 10 cm additional error on the trajectory determination. As a consequence, linear interpolation can be used only if the final required accuracy is lower than 10 cm, otherwise higher-order interpolations schemes must be used.

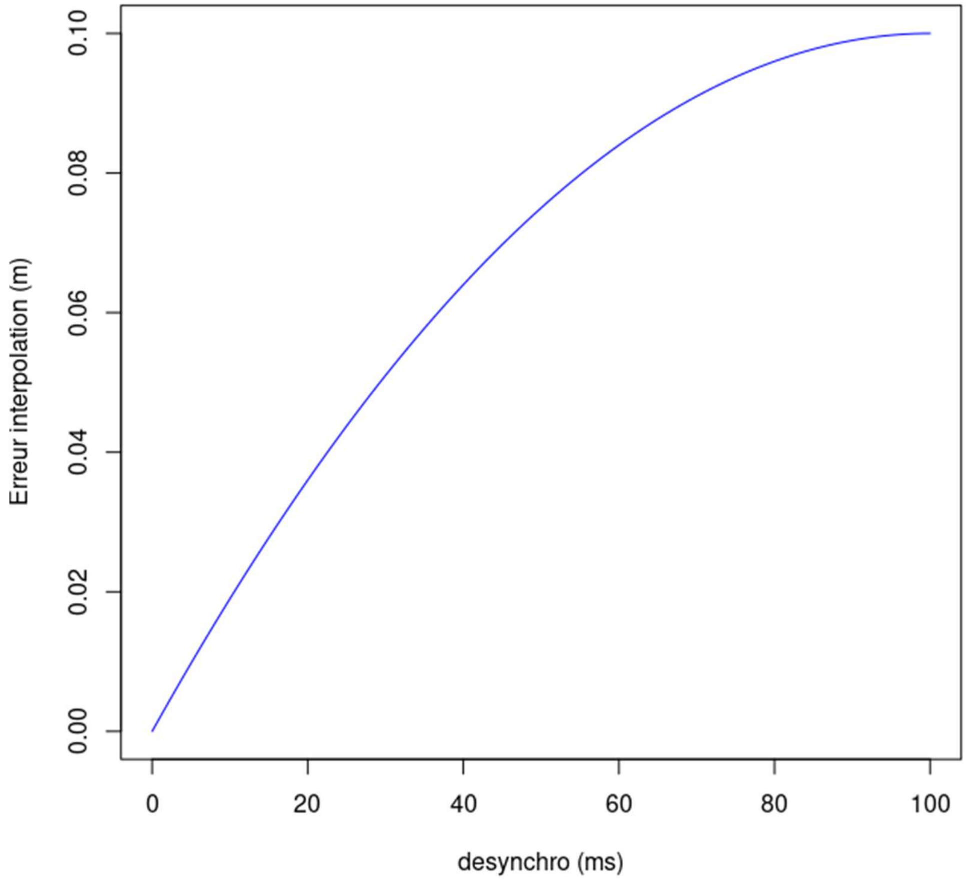


Fig 75. Theoretical error resulting from a linear interpolation of trajectory points. This plot shows that to interpolate at a 10 Hz frequency from the 5 Hz trajectory, it is necessary to interpolate 100 ms out of the observed points, which would produce a 10 cm additional error.

If interpolation is still needed with a high accuracy, input uncertainty of points on the trajectory should be propagated to the interpolated values. This can be done easily with a Kalman filter (KF), assuming constant acceleration model.

Assuming constant acceleration between epochs, the dynamic equation between time steps t and t + 1 can be written in three-dimensional space:

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \\ \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \\ \dot{\mathbf{z}} \\ \ddot{\mathbf{x}} \\ \ddot{\mathbf{y}} \\ \ddot{\mathbf{z}} \end{bmatrix}_{t+1} = \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \\ \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \\ \dot{\mathbf{z}} \\ \ddot{\mathbf{x}} \\ \ddot{\mathbf{y}} \\ \ddot{\mathbf{z}} \end{bmatrix}_{t} + \begin{bmatrix} \frac{\Delta t^3}{6} \cdot \varepsilon_{\mathbf{x}} \\ \frac{\Delta t^2}{2} \cdot \varepsilon_{\mathbf{x}} \\ \Delta t \cdot \varepsilon_{\mathbf{x}} \\ \frac{\Delta t^3}{6} \cdot \varepsilon_{\mathbf{y}} \\ \frac{\Delta t^2}{2} \cdot \varepsilon_{\mathbf{y}} \\ \Delta t \cdot \varepsilon_{\mathbf{y}} \\ \frac{\Delta t^3}{6} \cdot \varepsilon_{\mathbf{z}} \\ \frac{\Delta t^2}{2} \cdot \varepsilon_{\mathbf{z}} \\ \Delta t \cdot \varepsilon_{\mathbf{z}} \end{bmatrix}$$

where Δt is the time interval between a reference epoch, and the interpolated point (for example, 100 ms for interpolation from 5 Hz to 10 Hz), and $\varepsilon_{\mathbf{x}}$, $\varepsilon_{\mathbf{y}}$ and $\varepsilon_{\mathbf{z}}$ are realizations of the noise process on the derivative of acceleration (jerk), respectively in each of the 3 dimensions. This equation can be simplified, noting $\mathbf{X}_t$ the state vector at time step t:

$$\mathbf{X}_{t+1} = \mathbf{A}\mathbf{X}_{t+1} + \mathbf{G}\varepsilon_t$$

where $\varepsilon_t \in \mathbb{R}^3$ is a realization of the jerk process with covariance matrix $\mathbf{Q} \in \mathbb{R}^{3\times 3}$ and:

$$\mathbf{G} = \begin{bmatrix} \frac{\Delta t^3}{6} & 0 & 0 \\ \frac{\Delta t^2}{3} & 0 & 0 \\ \Delta t & 0 & 0 \\ 0 & \frac{\Delta t^3}{6} & 0 \\ 0 & \frac{\Delta t^2}{3} & 0 \\ 0 & \Delta t & 0 \\ 0 & 0 & \frac{\Delta t^3}{6} \\ 0 & 0 & \frac{\Delta t^2}{3} \\ 0 & 0 & \Delta t \end{bmatrix}$$

Then, supposing that $\mathbf{X}_t$ is the solution provided at a particular point (by *Comp3D* for the position variables $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ and by numerical derivations for their 2×3 derivatives), with covariance matrix $\mathbf{\Sigma}_t$, covariance can be propagated to the interpolated value $\mathbf{X}_{t+1}$:

$$\mathbf{\Sigma}_{t+1} = \mathbf{A}\mathbf{\Sigma}_t\mathbf{A}^{\mathbf{T}} + \mathbf{G}\mathbf{Q}\mathbf{G}^{\mathbf{T}}$$

For numerical application, a $\sigma_{\mathbf{x}} = 5$ mm standard deviation on position is assumed on each axis. Standard deviations on speed and acceleration are derived accordingly:

$$\sigma_{\dot{\mathbf{x}}} = \sqrt{2}\frac{\sigma_{\mathbf{x}}}{\Delta t} \qquad \text{and} \qquad \sigma_{\ddot{\mathbf{x}}} = 2\frac{\sigma_{\mathbf{x}}}{\Delta t^2}$$

$$\Sigma_t = \begin{bmatrix} \sigma_{\mathbf{x}}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\mathbf{y}}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\mathbf{z}}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\dot{\mathbf{x}}}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\dot{\mathbf{y}}}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{\dot{\mathbf{z}}}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{\ddot{\mathbf{z}}}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{\ddot{\mathbf{z}}}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{\ddot{\mathbf{z}}}^2 \end{bmatrix}$$

The interval time Δt is set to 100 ms, and the standard deviation on jerk process is set at 15 m.s^{-3}.

Applying the formula above, the propagated standard deviation the interpolated position is close to 10 mm on each of the 3 components.

Note that the same algorithm could be used to merge non-synchronized points of trajectory issued from sector A and sectors B/C/D. After interpolating the trajectory from one point of trajectory A to a given timestamp of trajectory B/C/D, with the prior covariance Σ_{t+1}, Kalman gain is computed then $\mathbf{X}_{t+1}$ and Σ_{t+1} are updated to their new values.

Therefore, Kalman filtering offers and elegant solution in order to (1) interpolate the trajectory at the desired frequency and (2) merge partial trajectories generated by out of synchronization sets of cameras.

The process being not designed to be executed in real time, Kalman filtering can be performed in both forward and backward step before averaging solutions.

With this model, and assuming nearly diagonal covariance matrices of the input trajectory points, simplified calculations show that interpolation can be done up to a 10 Hz frequency, while keeping a sub-centimetric uncertainty on points.

2.2.3.4. Conclusion

This document described the results of an optical tracking system developed at IGN, and designed to estimate the trajectory of an autonomous vehicle, driving at moderate speed (up to 50 km/h) on a small extent area. The system can be deployed in 8 hours, and once automatized, estimation of the trajectory can be computed in a few hours.

Even in case of technical incidents (mainly loss of cameras) the system was proven to be able to estimate the trajectory of the vehicle at a **5 Hz** frequency, with an absolute and relative accuracy of **5 mm** in each of the three-dimensional components, and better than **0.05°** in vehicle heading.

Inter-comparison of observations showed a good consistency, and cross-validation between independent trajectories generated by different subsets of cameras showed a **9 mm** median deviation in planimetric positions, which is in agreement with the predicted uncertainties.

Trajectory can be linearly interpolated to its nominal frequency of 10 Hz. In case utmost accuracy and fine temporal resolution are both required, and under the assumption that the vehicle acceleration can be (at least statistically) bounded, higher-order interpolation schemes or Kalman smoothing techniques can be used to resample the trajectory while keeping a reasonably low standard error on interpolated positions.

The following steps to the development of the system are:

- making it more robust to ease the tasks needed to be done by operators on the field,

- conducting further tests to demonstrate its capability to work properly at high speed (up to 80 km/h?), and with various weather conditions,

-automatizing the search of target outlier detections in the global adjustment (with an iterative process using the prior knowledge on the targets relative positions provided by the rigid-block),

- and eventually optimizing the target detection module to be able to process very large datasets of images and longer trajectories.

2.2.4. Weather conditions

Standards for testing and validating ADAS and AV systems systematically specify the nominal weather conditions to be met. At present, most standards specify that tests must be carried out under clear weather conditions. However, with the advent of AI and the change in the regulatory context, manufacturers will have to specify their ODD. Weather conditions will form part of the ODD and will therefore have to be precisely defined and measured during validation tests. The weather can generally be broken down into four main categories: clear weather, fog, rain and snow. It is therefore necessary to describe the characteristics and physical quantities of each. It is also important to describe the equipment that can be used to carry out the measurement. Another challenge linked to the insertion of AI into driving systems is the need to be able to reproduce numerous use cases. To achieve this, and as has already been explained in the PRISSMA project, it will be essential to carry out tests under simulated conditions. It is therefore important to describe the key characteristics for reproducing weather conditions that are representative of reality. The rest of this section will therefore be broken down into four sections for each of the weather conditions, with the physical definition, measuring instruments and key simulation elements for each one.

Clear Weather

Clear weather conditions are defined by the absence of disturbances such as rain, snow or fog. The description of the following conditions will help to better understand this section. Clear weather means no fog, i.e. a MOR (see definition in fog) above 1000m and no precipitation, i.e. a rainfall rate of 0mm/h (see definition in rain). In clear conditions, it is also important to consider the cloud cover. To do this, the illumination measurement in lux is a good indicator. This measurement certifies that the lighting condition is similar. It also enables us to better define the day/night aspect, or glare conditions. It should be noted that current standards define clear weather with dry ground. In fact, there can be clear weather in the weather sense (atmosphere), but the ground is still wet or snow-covered. These borderline cases are described in the other meteorological aspects. As far as the reproduction of clear weather in simulated environments is concerned, this is a basic case and therefore does not present any specific challenge.

Fog

Fog is made up of water droplets suspended in the air. These droplets absorb and scatter light, which reduces visibility. The Meteorological Optical Range (MOR) is therefore defined as the physical parameter describing density of fog. This is described in detail in the World Meteorological Organisation's (WMO) guide No.8. The MOR corresponds to the distance at which a luminous flux is 95% attenuated in fog. The lower the MOR, the greater the intensity of the fog. The standard NF P 99-320 (AFNOR, 1998) proposes a classification of fog according to the MOR.

	Road visibility class	MOR (m)
Fog (meteorological definition)		< 1000
Fog road class	1	200 - 400
	2	100 - 200
	3	50 - 100
	4	< 50

The sensors used to measure MOR are transmissiometers and diffusometers. These have a head that emits light and a head that receives light. In the case of transmissiometers, the emitter and receiver are aligned and the measurement focuses on the quantity of attenuated flux. In the case of diffusometers, the emitter and receiver are set at 45° and the measurement is of the proportion of light scattered.

When testing under real fog conditions, the only criterion to consider is the MOR. However, when testing under simulated conditions (numerically or physically), it is important to consider drop size. To reproduce a realistic fog, the droplet size must be between 0.5 and 50 microns. This criterion is crucial, particularly when using LIDAR, time-of-flight or thermal camera sensors. The spatial homogeneity of the fog is also important.

Rain

Rain is made up of drops that fall to the ground in aqueous form. Rain has direct and indirect impacts. Direct impacts are caused by the presence of drops that can prevent visibility. Indirect impacts are various such as: wetting the ground (which changes its appearance and its ability to grip), generating spray effects behind vehicles, wetting sensors, etc.. Rain is characterized by its intensity in mm/h. This value, described in detail in the WMO guide no8, corresponds to the height of water falling on 1m2 over one hour. A rain intensity of 0mm/h means that there is no rain, but the rain intensity can then rise to several hundred mm/h during peak thunderstorms. The standard NF P 99-320 (AFNOR, 1998) also proposes a classification of rain according to the rainfall intensity.

Class	**Rainfall intensity (mm/h)**
Very low	0.0 – 0.1
Low	0.1 – 2.5
Moderate	2.5 – 7.5
Heavy	> 7.5

There is a variety of sensors for measuring rainfall intensity. There are bucket rain gauges, disdometers and capacitive impact sensors. The first type of sensor is very accurate, but does not have sufficient temporal resolution test measurements requiring a temporal resolution of a few seconds. It is intended more for meteorological monitoring with a time resolution of a few minutes. The other two types of sensor have finer time resolution, but because the measurement is indirect, they are less accurate. Note that there are also sensors that can be used to estimate the state of the road (wet, dry, etc.) or even measure the height of water on the pavement. This type of sensor can be useful for fine testing on tracks. This is what is done for tire testing, for example.

When testing in real rain conditions, the intensity of the rain is the determining criterion. However, there are other important criteria to consider, the type of pavement on which the test is carried out (micro texture and materials) or the wind can also influence the results. These can vary grip or produce different spray effects for the same intensity of rain.

Consider when creating tests under simulated conditions (numerically or physically), the size and speed of raindrops must also. These are well described in the literature and vary from 0.5mm to a few millimeters, with speeds of around 9m/s. These aspects are important for all sensors, including radar. Spatial homogeneity is also an important criterion to consider.

Snow

Snow is made up of water particles in solid form, in the form of snow crystals.

Like rain, snow has both direct and indirect impacts. Snow can also stick under certain conditions, completely obstructing the visibility of certain sensors (particularly at the bumper). Snow intensity is also characterized by rain intensity in mm/h. This is the rainfall intensity once the snow has melted. The WMO's guide provides a classification of snow:

Snowfall Intensity	**Water equivalent content (mm/h)**
Light	< 1
Moderate	1 - 5
Heavy	> 5

As the magnitude that characterizes snow is the same as for rain, the sensors used to measure it are the same. They are then heated so that they are not caught in the snow, or even to melt the snow and measure it in the form of liquid water. It should be noted that trough-type rain gauges need to know the temperature in order to estimate whether it is rain or snow. As the other sensors take their measurements from the size and speed of the hydrometeors, they are able to

classify rain and snow directly. Again, as with rain, the intensity of snow, although a key criterion during a test in real conditions, is not sufficient to characterize the test: the wind, the type of pavement, the air and pavement temperatures and the condition of the pavement are all criteria that can vary and have an impact on the test. For this, the use of surface condition sensors can be of interest.
For tests under simulated conditions (numerically or physically), once again the microphysical properties of snow (such as distributions for snowflake size and mass) have to be taken into account in addition to the rest.

<u>Other weather-related factors</u>
Other aspects of the weather are present during the tests, but they are secondary. In fact, the current standardization phase involves a gradual transfer from clear weather only to certain deteriorated weather conditions (rain or fog). The following aspects generally have a smaller impact, but are mentioned here:
- Effect of the ground state: the ground can be dry, but also wet or snow-covered. The case of black ice is also important because of the change in the vehicle dynamics.
- Environmental effects: the environment can change according to the season, disruptors such as leaves or spray behind vehicles can be present. All these aspects can influence AI-based perception and decision-making algorithms.
- Wind effects: the wind has not been described here, but it can have an impact on vehicle dynamics or effects on perception by lifting elements from the environment (dust, leaves, etc.).
- Glare: the glare from the headlights of opposing vehicles, at sunset or sunrise or at tunnel exits can have an impact on perception systems. This aspect can therefore also be studied when defining test cases.
- Temperature and humidity: temperature and humidity have no direct impact on vehicle systems. On this subject, the only objective is to check the resistance of the hardware elements to sometimes-extreme variations in temperature and humidity. However, the ageing of vehicles is already systematically verified during specific tests carried out during type approval.

2.3 Test equipment deployed in vehicles

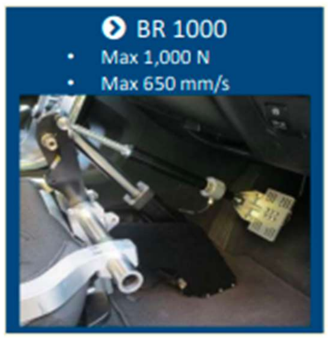
2.3.1. Driving Robot

Driving robots are also an important part of the testing equipment for automated and intelligent vehicles. They are necessary to be compliant to a test protocol, requiring very stable and accurate trajectory and speed for the vehicle (not more than 10 centimeter or 0.1 km/h compared to the nominal specified value). A human driver cannot drive with these requirements!

They are composed of steering & pedals robots:

Steering robots

Pedal robots

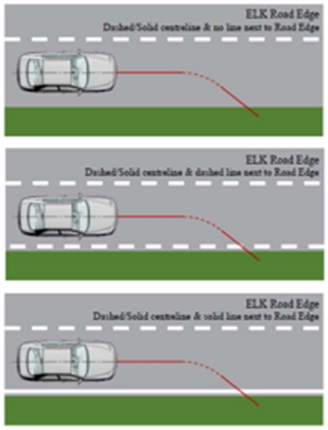


As example, the two figures below are illustrating

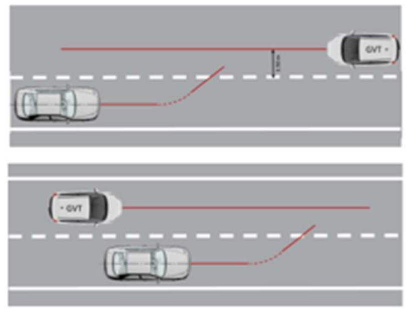
- the Euro NCAP LSS (Lane Support Systems) emergency lane keeping (ELK) testing protocols
- the main variables to monitor (as inputs to make the test and outputs as tests results) :

Euro NCAP LSS test protocol :

Emergency lane keeping tests (ELK):



Road edge tests



Oncoming tests

4.2 Measurements and Variables

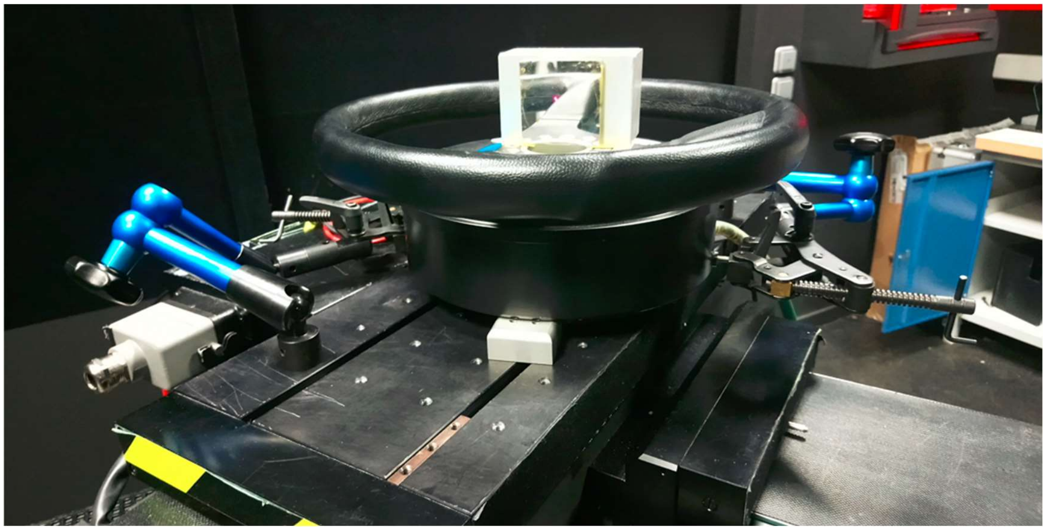
4.2.1	Time	T
	• T_0, time where manoeuvre starts with 2s straight path	T_0
	• T_{LKA}, time where LKA activates (for calibration purposes only if required)	T_{LKA}
	• T_{LDW}, time where LDW activates	T_{LDW}
	• T_{steer}, time where VUT enters in curve segment	T_{steer}
	• $T_{crossing}$, time where VUT crosses the line or road edge	$T_{crossing}$
4.2.2	Position of the VUT during the entire test	X_{VUT}, Y_{VUT}
4.2.3	Position of the GVT during the entire test	X_{GVT}, Y_{GVT}
4.2.4	Speed of the VUT during the entire test	$V_{long,VUT}$ $V_{lat,VUT}$
4.2.5	Speed of the GVT during the entire test	V_{GVT}
4.2.6	Yaw velocity of the VUT during the entire test	$\dot{\Psi}_{VUT}$
4.2.7	Yaw velocity of the GVT during the entire test	$\dot{\Psi}_{GVT}$
4.2.8	Steering wheel velocity of the VUT during the entire test	Ω_{VUT}

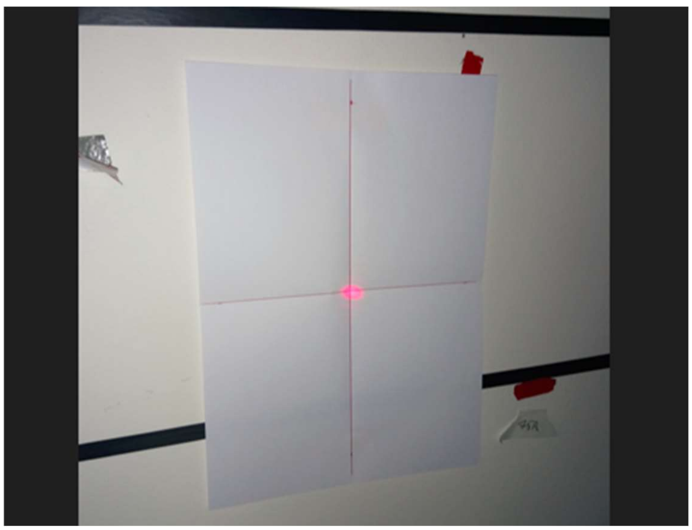
4.3 Measuring Equipment

4.3.1 Equip the VUT with data measurement and acquisition equipment to sample and record data with an accuracy of at least:

- VUT and GVT longitudinal speed to 0.1km/h;
- VUT and GVT lateral and longitudinal position to 0.03m;
- VUT heading angle to 0.1°;
- VUT and GVT yaw rate to 0.1°/s;
- VUT longitudinal acceleration to $0.1 m/s^2$;
- VUT steering wheel velocity to 1.0°/s.

Here below are two photos of driving steering robot calibration in UTAC :





2.3.2. IMU and other sensor

Many pieces of equipment are embedded in the vehicle under test to have a precise measure of its position, speed, deceleration, reactions (physical or in the electric CAN):

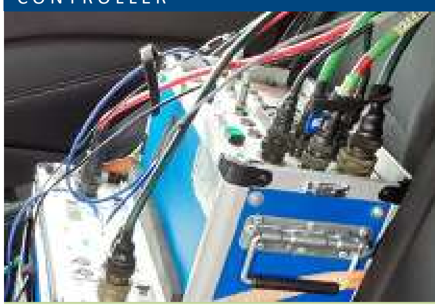
Motion Measurement

MOTION PACK 1

Manufacturer Oxford Technical Solutions (OxTS)	
Unit model TO BE DEFINED	
Sensors Accelerometers (Servo) / Gyros (MEMS)	
Data output rate 100 Hz	*Coupling method* GNSS / INS

Data Recording System

CONTROLLER



Manufacturer Antony Best Dynamics (ABD)	
Unit model XR Omni	
Sampling rate 100 Hz	
Analog input voltage ± 10 V	*A / D conversion* 16 bits

HMI sensors

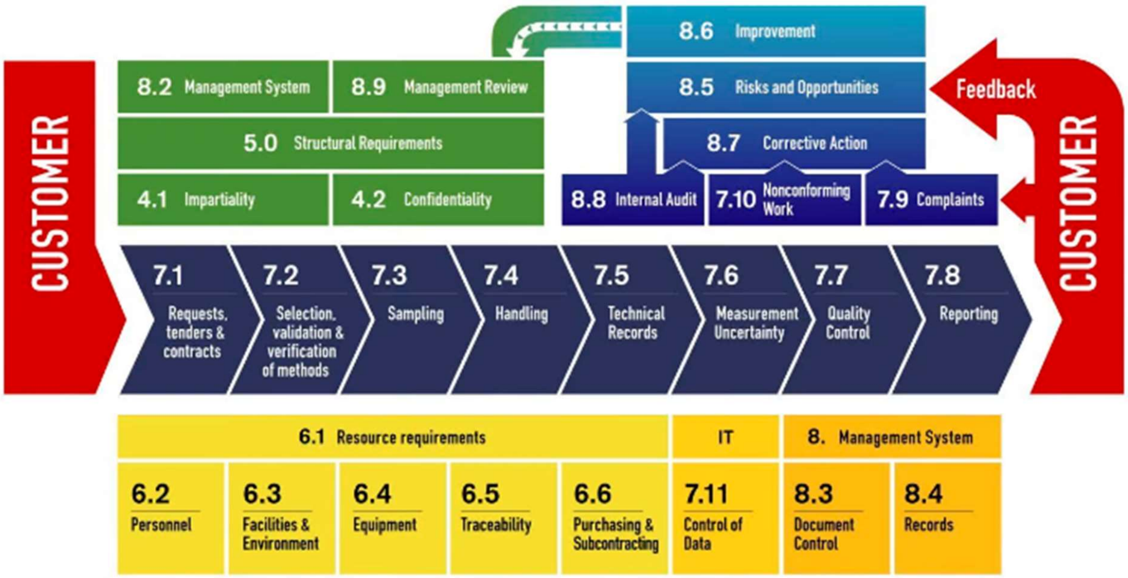
VIDEO VBOX

Manufacturer Racelogic	
Unit model	*Frame rate*

GOPRO

Manufacturer GoPro	
Unit model	*Video resolution*

This testing equipment is also involved in the ISO/IEC 17025 standard, very complete processes with eight steps summarized in the below figure:



As is it very confidential we just give here below illustrations and extracts of UTAC tables listing:

Exigences 17025	Eléments et procédures à mettre en œuvre
Identité de l'équipement et de son logiciel correspondant.	Fiche de vie du moyen et étiquettes Métrologie UTAC. PG.REI.000.001 pour l'aspect logiciel.
Nom du fabricant, l'identification de type, le numéro de série ou autre identification unique.	Fiche de vie du moyen.
Vérifications de la conformité de l'équipement aux spécifications.	cf. § 6.2.1 pour la réception. Fiche de vie (pour l'état, la disponibilité).

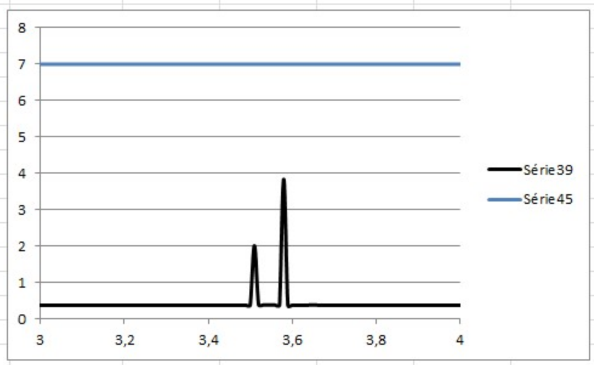
Extract of UTAC procedure and documents for testing equipment life management

Exigences du système	Solutions du service Métrologie
Titre.	- Certificat d'étalonnage. - Constat de vérification.
Nom et adresse du laboratoire ainsi que le lieu si différent du laboratoire.	- Nom et adresse du laboratoire sur la page de garde.
Identification unique du rapport d'opérations métrologiques.	- Numéro unique attribué automatiquement par la Métrologie. - Code : année – n° d'ordre (ex : 2016-00001).
Répétition de l'identification.	- Page de garde et pages des mesures si utile.
Numéro de page et nombre de pages sur chaque page.	- Page N° / Nbr de pages total, sur chaque page du document.
Nom et adresse du client.	- Intitulé du service client sur la page de garde.
Identification de la méthode employée.	- Page 2, référence à la procédure ET + descriptif de la méthode.

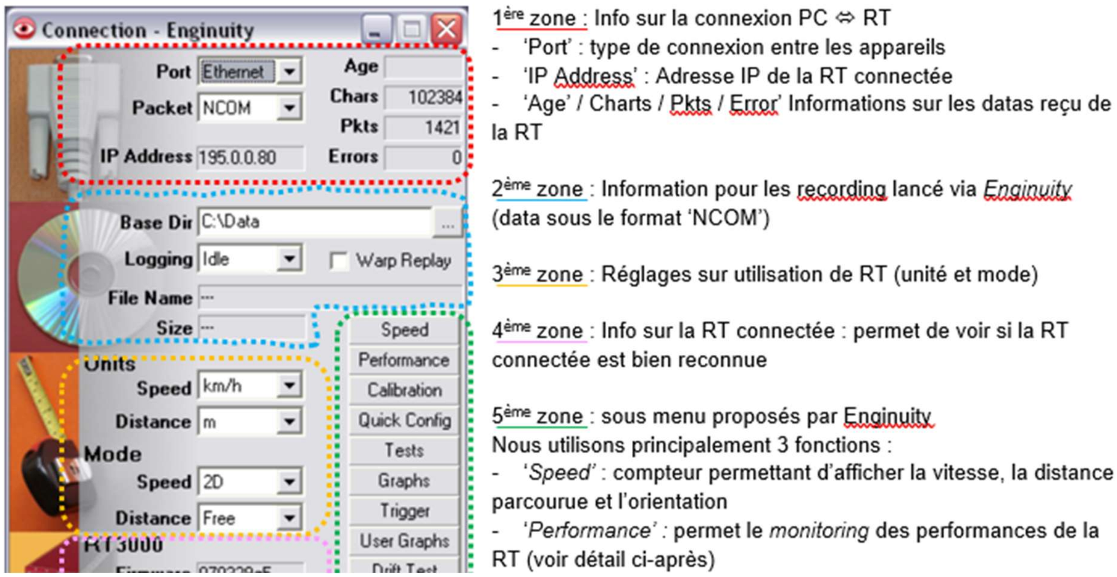
Extract of UTAC certificates list for verification and calibration of testing equipment

We also give here below an extract of the UTAC calibration procedure for a motion measurement equipment or a GPS sensor, to compare average speed measured by motion equipment and calibrated average speed; this is required for all tests and all measured data:

354	349	3,49000001	0,37822387	6,99499989		
355	350	3,5	0,37822387	6,99499989		
356	351	3,50999999	2,01143312	6,99499989	vitesse moy	6,996874928
357	352	3,51999998	0,38952407	6,99499989		
358	353	3,52999997	0,37855622	7		
359	354	3,53999996	0,37888861	6,99499989		
360	355	3,54999995	0,37888861	7		
361	356	3,55999994	0,37888861	7		
362	357	3,56999993	0,37888861	6,99499989		
363	358	3,57999992	3,83541608	6,99499989		
364	359	3,58999991	0,36792076	6,99499989		
365	360	3,5999999	0,37755916	7		
366	361	3,6099999	0,37822387	7		
367	362	3,61999989	0,37822387	6,99499989		
368	363	3,63000011	0,37822387	6,99499989		
369	364	3,6400001	0,37855622	7		
370	365	3,6500001	0,37855622	6,99499989		
371	366	3,66000009	0,37855622	7		
372	367	3,67000008	0,37822387	7		
373	368	3,68000007	0,37822387	6,99499989		



We also give here below an extract of the installation/configuration procedure of RT type motion measurement equipment:



To conclude this part about testing equipment validation for AI-based vehicles:

- Testing equipment have to be fully operational and accurate to validate AI-based vehicles. For example, all UTAC current procedures to validate , use and maintain our testing equipments seem to be absolute necessary for AI-based vehicles validations

- It is quite different to evaluate vehicles parts like sensors and validate the whole vehicle. For an evaluation of a whole, the classical testing equipment (NCAP standard…) seem sufficient for short term.

- Unlike WP2 and its thousands of virtual tests for AI validation, physical tests are long, expensive and only a few dozen can be made on test track.

- Nevertheless, the testing equipment & procedures on track could have to evoluate and address new requirements in the near future due to AI-based vehicles, if regulator or Euro NCAP or manufacturers demand us to evaluate and validate specific dimensions, requiring more specific or accurate equipments. For example if the test tracks evaluator are asked to explain AI-based vehicle behavior inside the whole vehicle, the testing equipment, tools and protocols could have to change because explainabiliy is a new and huge challenge, requiring much more tests, analyses, statistic tools… like done in PRISSMA WP1 and WP2.

- The deliverable L3.3 explained widely the current situation for regulation or Euro NCAP test requirements: only a few tests are required (and reasonably possible) per scenario so we proposed in this L3.3 deliverable only very simple new metric and few new tests to validate AI-based vehicles (KPI metrics). But it could change in the future, requiring to change our vehicle testing equipment & procedures, as preconized by PRISSMA WP1 & WP2.

2.4 Test bench validation

A number of criteria need to be considered in order to validate the test bench. These criteria are in fact intermediate to those of digital twin on the one hand and track testing on the other. They make it possible to obtain much more repeatable and measurable conditions than on a track (lighting, weather, etc.) and also to qualify systems for better integration into digital twin systems (which require a lot of input data).

Here is a list of the test bench validation elements that have been identified in this respect:

- The position of objects must be known with precision. 3D scanning methods can now be used to accurately position objects. This may even enable them to be integrated into digital twins.
- Lighting characterisation is also an important factor. Using test benches means that conditions are much more controlled than on an outdoor track. Lighting characterisation is therefore essential.
- Characterization of all the surfaces: in terms of reflectivity, micro-texture and condition (dry or wet). This is very important for understanding the feedback from the systems tested, but also for integration into digital twins.
- Detailed weather characterisation using dedicated sensors

3 COUPLING WITH SIMULATION OR REAL-LIFE ROAD TESTS

3.1 Digital Twins

This approach was extensively covered in the PRISSMA 2.7 deliverable. In summary from the WP3 point of view, to ensure the representativeness of a digital twin during on-site measurements, several precautions should be taken:

1. Accurate Data Collection: Ensure that all data collected during on-site measurements are accurate and comprehensive. This includes using calibrated instruments and sensors to capture relevant parameters.

2. Validation and Calibration: Validate and calibrate the digital twin against real-world measurements to ensure its accuracy and reliability. Any discrepancies between the digital twin and actual measurements should be addressed through appropriate adjustments.

3. Environmental Factors Consideration: Take into account environmental factors such as temperature, humidity, and pressure, which can affect the performance of both the physical system and the digital twin. Adjustments may be necessary to account for these variations.

4. Sensor Placement and Quality: Place sensors strategically to capture data that accurately reflects the behavior of the physical system. Ensure that sensors are of high quality and properly maintained to minimize errors in data collection.

5. Data Synchronization: Ensure synchronization between the data collected on-site and the corresponding inputs to the digital twin. Any delays or discrepancies in data transmission can

lead to inaccuracies in the digital twin's representation.

6. Continuous Monitoring and Update: Continuously monitor the performance of the digital twin and update it as necessary based on new data collected on-site. This iterative process helps maintain the representativeness of the digital twin over time.

7. Model Validation: Validate the underlying models and algorithms used in the digital twin to ensure they accurately simulate the behavior of the physical system under various conditions.

By following these precautions, you can enhance the representativeness of a digital twin based on on-site measurements, improving its utility for predictive analysis and decision-making.

3.2 Hybrid test between simulation and test bench

Hardware-in-the-loop (HIL) testing represent a crucial step in the development and validation of complex systems across various industries, ensuring reliability, efficiency, and safety. Leveraging 3D simulation in conjunction with HIL tests amplifies the efficacy of the validation process, offering a comprehensive approach towards achieving high-quality outcomes.

Essentially, HIL testing involves integrating physical hardware components into a simulated environment, allowing engineers to evaluate system behavior under realistic conditions. This method not only streamlines testing procedures but also enables early detection of issues, significantly reducing development time and costs.

Integrating 3D simulation adds another layer of sophistication to the validation process. By immersing the hardware components within a virtual environment, engineers can replicate real-world scenarios with remarkable accuracy. This immersive approach facilitates a deeper understanding of system dynamics, enabling thorough analysis and fine-tuning of system parameters.

To achieve high-quality results through HIL tests and 3D simulation, several key considerations must be addressed:

Accurate representation of the physical environment: The virtual environment must closely mimic real-world conditions, including factors such as terrain, weather, and lighting. Precise modeling ensures that the system's response remains faithful to actual operational scenarios.

Realistic interaction dynamics: Interaction between hardware components and the simulated environment should mirror reality. This entails accurately modeling physical interactions, such as collisions, friction, and fluid dynamics, to simulate realistic system behavior.

In our study, we have used different methods, including HIL testing, to compare and evaluate the simulation tools against each other. The HIL test involves using a camera sensor positioned in front of a projected screen to display images generated from two sources: a 3D simulator and the ZED2i camera capturing real scenes. These images are then captured by the camera sensor, as shown in Figure 76.

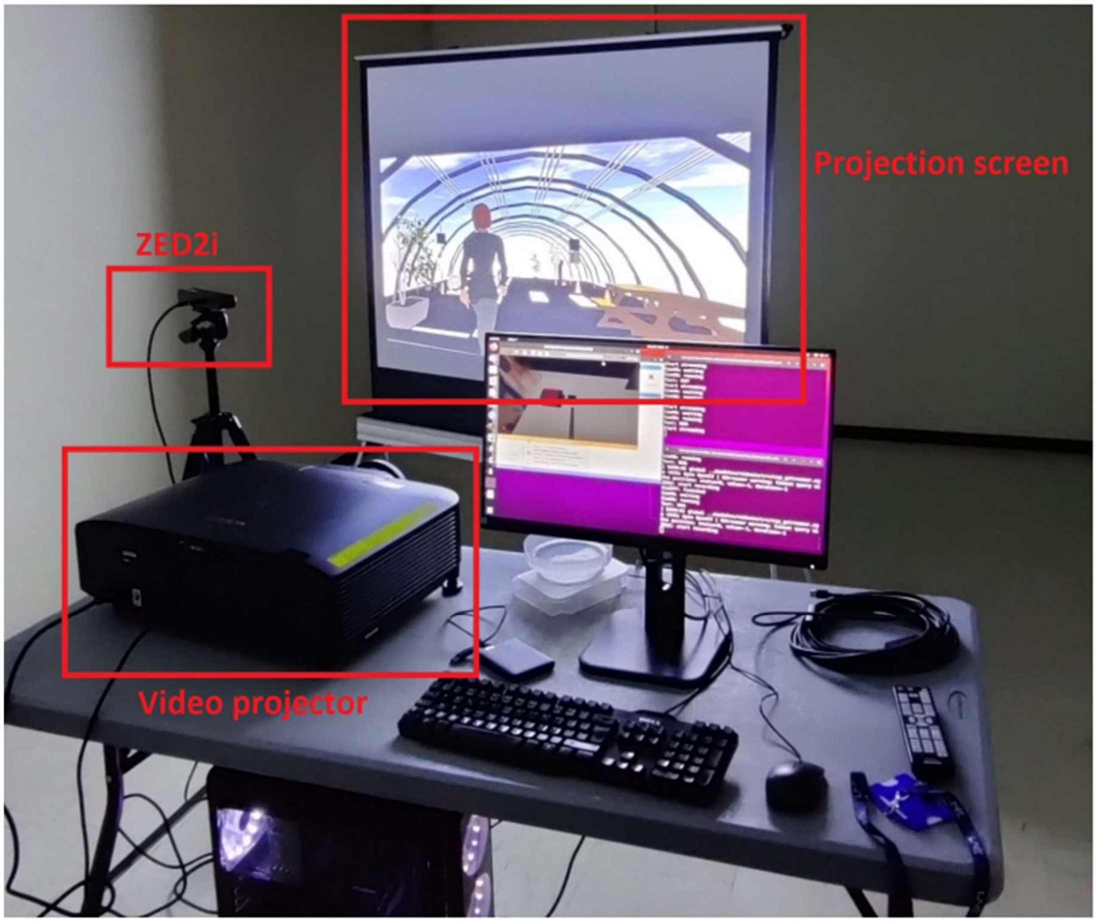


Figure 76: HIL test bench

To ensure high-quality outcomes in this procedure, various critical factors need to be taken into account:

Realistic scene reproduction: One of the fundamental aspects of testing camera sensors using a projected screen is the faithful reproduction of real-world scenes. The 3D simulation environment must accurately replicate various scenarios, including different lighting conditions, weather effects, and object movements. This realism ensures that the camera sensor's response aligns closely with actual operating conditions.

High-fidelity projection system*:* The quality of the projected screen significantly affects the accuracy of sensor testing. A high-fidelity projection system capable of rendering detailed textures, realistic colors, and precise motion is essential. Moreover, the projection setup should facilitate dynamic adjustments to simulate varying distances, angles, and perspectives, enabling comprehensive sensor evaluation.

Calibration and alignment: Precise calibration and alignment of the camera sensor with the projected screen are critical for accurate testing. Calibration ensures that the sensor accurately captures the projected images, while alignment guarantees proper spatial correspondence between the virtual and physical elements. Careful calibration minimizes errors and discrepancies, enhancing the reliability of test results.

Dynamic scene generation: The 3D simulation environment should support dynamic scene generation to simulate realistic scenarios effectively. This includes the ability to generate moving objects, changing environmental conditions, and interactive elements. Dynamic scene generation enables comprehensive testing of the camera sensor's performance under diverse and evolving conditions.

Regarding the realism of the 3D simulator, we have used 4D-Virtualiz simulator and integrated the 3D model of the Cerema platform into it. This model has been meticulously developed by professionals to ensure maximum accuracy in replicating the real platform. However, we encountered some limitations with this simulator, particularly concerning rendering, which was not at a high level because the simulator uses the rendering engine called OGRE which does not have good realism compared to those used in unreal engine or unity.

Additionally, we encountered limitations related to the visual rendering of fog, which was very poor as shown in Figure 77. The fog intensity increases little with the distance and fog haze is practically non-existent. To overcome this problem, the smoke is used in addition to fog in order to enhance the visual rendering. The smoke remained static, with no movement or ripple effects as shown in Figure 78. In this study, two smoke intensities are defined to obtain weak and strong fog. In collaboration with Cerema, we adjusted the smoke intensity values to ensure visual acceptability.

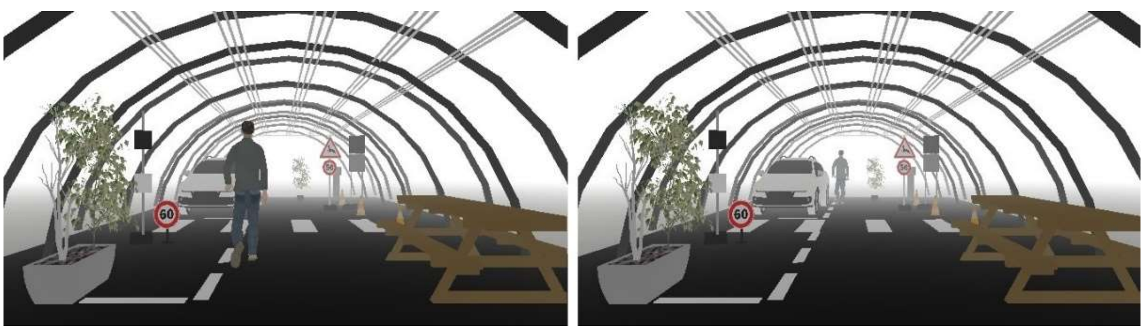


Figure 77: Original fog of 4DV.

Figure 78: Fog using smoke of 4DV.

A high-fidelity projection system **Epson EB-PU1007B** was used in the HIL test. This projection system offers a high definition with 4k enhancement, 7000 lumen for color light output and the ability to reproduce up to 1.07 billion colors. Additionally, it enables adjusting and aligning images with the camera through vertical and horizontal shift adjustments.

To ensure proper alignment of the image and achieve a rectangular shape, we adjusted these shifts to correct the shape of the image retrieved by the camera sensor as shown in Figure 79.

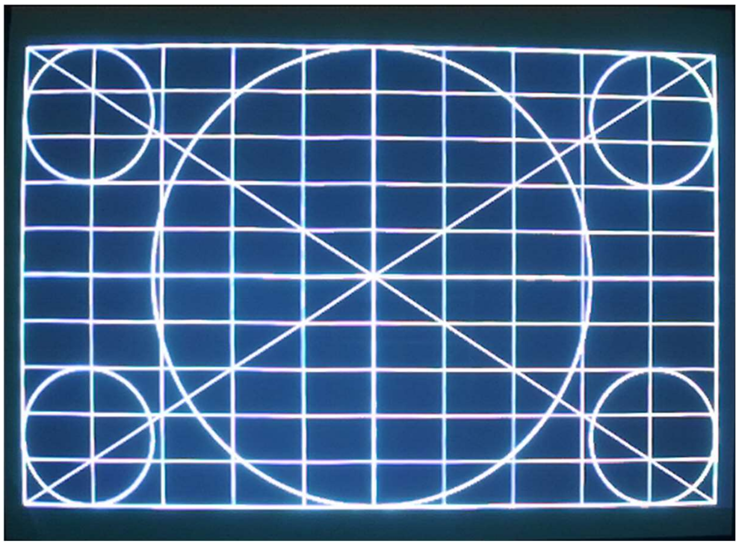


Figure 79: Correcting image shape using horizontal and vertical shifts

Finally, by applying the previous procedure, we can obtain the following images using the 3D simulator and images captured from real scenes.

Figure 80: Images captured from HIL tests: left: image from 3D simulator and right: image from real scene

3.3 Test representativeness

Testing the representativeness of track and bench testing methods for autonomous vehicles involves a meticulous examination of the extent to which these tests accurately mirror real-world scenarios. The evaluation process encompasses a comprehensive analysis of both physical testing on tracks and simulation-based testing on benches.

Physical testing on tracks involves conducting trials in controlled environments that simulate real-world driving conditions. To assess the representativeness of track testing, factors such as track design (markings, infrastructure, curvature, pavement... must comply with the standards of real roads, while still being able to show the fatigue of the latter), environmental conditions, traffic scenarios, and vehicle dynamics must be meticulously considered. Additionally, the diversity of scenarios encountered during testing plays a crucial role in determining representativeness. A comprehensive evaluation involves comparing the range and complexity of scenarios encountered on the track to those encountered in actual driving environments. The problem here is the financial and time cost that these tests can represent, making it difficult to avoid poor scenario coverage. In the same way, we may wonder how the small sample of regulatory targets (or even their appearance) can effectively be representative of the multiplicity of situations on real roads. What was not a problem for tests without AI (or at low levels of automation) where obstacles just had to be detected and reacted to, could become one for highly automated systems where AI will take on a much more crucial role.

Simulation-based testing on benches complements physical testing by offering a controlled environment where various scenarios can be simulated and tested efficiently. Evaluating the representativeness of bench testing involves assessing the fidelity of simulations to real-world conditions and this aspect has been addressed in PRISSMA Deliverable 2.7. This deliverable includes scrutinizing factors such as sensor accuracy, environmental modeling, vehicle dynamics simulation, and the incorporation of complex driving scenarios. The ability of bench testing to replicate real-world scenarios accurately significantly impacts its representativeness.

Furthermore, the integration of physical and simulation-based testing methodologies can enhance the overall representativeness of testing approaches for autonomous vehicles. Combining data from real-world testing with simulated scenarios enables a more comprehensive evaluation of vehicle performance across a broader range of conditions. This hybrid approach leverages the strengths of both physical and simulated testing to achieve a more representative assessment.

To assess representativeness, there is no miracle solution: you need to be able to rely on the complementary nature of the tests (for example, the use of real road tests for track tests, or track tests for simulation, to provide ground truth) and to ensure that the results are consistent according to the different types of test used. In Deliverable 1.6, we have provided an analysis of the strengths and weaknesses of each type of test, particularly with regard to representativeness and how the complementarity and dependency of tests can be used to try to compensate for their weaknesses, which will be looked at in the next section.

3.4 Analysis of test interdependence and complementarity

As explained in Deliverable 1.6, simulation testing (SIL, HIL, VIL), track testing and open road testing are three interdependent components essential for the comprehensive validation of autonomous vehicles. Each testing method offers unique advantages and complements the others, thereby forming a synergistic approach to ensure the safety and reliability of autonomous driving systems. For example, open road testing does not allow for repeatability or reproducibility of tests in an obvious way, is very costly, requires very long tests to be able to cover a minimum of the ODD and does not allow for easy or at least safe testing of the limits of a system (it is preferable to avoid accident-prone situations). However, it provides unrivalled representativeness and can see the emergence of situations not anticipated when the system was created, which is globally the opposite of SIL testing.

Simulation testing serves as a foundational pillar in the validation process by providing a controlled virtual environment where a wide range of scenarios (n particular the quasi accident scenarios) can be efficiently and safely. Simulations allow exploring hypothetical scenarios, fine-tuning algorithms, and assessing vehicle behavior under conditions that may be challenging or dangerous to replicate in real-road or tack testing. But behind this idyllic veil lies a dependence on the quality of virtual tools, a lower level of fidelity (particularly for the SIL), a partial validation of a system of system (particularly for the hardware in the loop) as well as an expensive software life-cycle or test bed (HIL and VIL) et a risk of over-reliance. The only way to remedy the problem of low fidelity is to use ground truth from physical tests, and for VIL tests, they are often based on test tracks. The fundamental need for reference and ground truth was addressed in detail in PRISSMA Deliverable 2.7. HIL and VIL bench testing complements SIL testing by bridging the gap between virtual simulations and real-world performance. Bench testing involves conducting experiments in controlled environment settings using physical prototypes or components of autonomous systems. This method allows researchers to validate sensor accuracy, evaluate hardware reliability (difficult in SIL), and assess system integration under controlled conditions. Bench testing also enables researchers to conduct in-depth analysis and debugging of hardware components, identifying potential issues that may arise during real-world operation.

Track testing represents another stage in the validation process, where autonomous vehicles undergo rigorous evaluation in real-world driving environments but with a more restricted scenario framework. Track testing provides invaluable insights into vehicle performance under dynamic and environmental conditions, including adverse weather conditions or interactions with other road users. By testing on tracks, we can assess their ability to handle real-world challenges effectively and very limited robustness to real conditions.

The interdependence of simulation testing, physical bench testing, and track testing lies in their complementary roles throughout the validation process. Simulation testing lays the foundation by enabling efficient exploration and validation of a wide range of scenarios in a virtual environment but needs ground truth to guarantee a minimum of representativeness. Bench testing complements simulation by validating hardware and software components under controlled environmental conditions, ensuring their reliability and functionality but often relies on digital twins or real tracks. Finally, track testing validates the performance of autonomous vehicles in real-world scenarios, confirming their readiness for deployment on public roads but cannot be exhaustive and needs simulation on this aspect.

In conclusion, the interdependence and complementarity of the different kind of testing are essential for the comprehensive validation of autonomous vehicles. By leveraging the strengths of each testing method and integrating them into a cohesive validation strategy, researchers can ensure the safety, reliability, and effectiveness of autonomous driving systems.

REFERENCES

[1] European Parlement, *Regulation (EU) 2018/858 - the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles,,* https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R0858, 2018.

[2] EuroNCAP, *EUROPEAN NEW CAR ASSESSMENT PROGRAMME - TEST PROTOCOL – AEB Car-to-Car systems,* https://cdn.euroncap.com/media/77302/euro-ncap-aeb-c2c-test-protocol-v42.pdf, 2023.

[3] Sétra, «Comprendre les principaux paramètres de conception géométrique des routes,» janvier 2026. [En ligne]. Available: https://dtrf.cerema.fr/pdf/pj/Dtrf/0004/Dtrf-0004044/DT4044.pdf?openerPage=notice&qid=sdx_q0.

[4] UNITED NATIONS, «Addendum 151 – UN Regulation No. 152 - Uniform provisions concerning the approval of motor vehicles with regard to the Advanced Emergency Braking System (AEBS) for M1 and N1 vehicles,» 2020. [En ligne]. Available: https://unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/2020/R152e.pdf.

[5] EuroNCAP, *EUROPEAN NEW CAR ASSESSMENT PROGRAMME - TEST PROTOCOL – AEB/LSS VRU systems,* https://cdn.euroncap.com/media/77299/euro-ncap-aeb-lss-vru-test-protocol-v44.pdf, 2023.

[6] UNITED NATIONS, «Addendum 12-H: UN Regulation No. 13-H - Uniform provisions concerning the approval of passenger cars with regard to braking,» 2018. [En ligne]. Available: https://unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/2018/R013hr4e.pdf.

[7] CEREMA, «L'adhérence des chaussées - État de l'art et recommandations,» octobre 2015. [En ligne]. Available: file:///C:/Users/e.chateauroux/Downloads/DT6914.pdf.

[8] EUROPEAN COMMISSION, «EU 2021/1958 - specific test procedures and technical requirements for the type-approval of motor vehicles with regard to their intelligent speed assistance systems and for the,» 2021 type-approval of those systems as separate technical

units and amending Annex II to that Regulation. [En ligne]. Available: https://eur-lex.europa.eu/eli/reg_del/2021/1958/oj.

[9] DSR/BSC – DGITM, «INSTRUCTION INTERMINISTÉRIELLE SUR LA SIGNALISATION ROUTIÈRE du 22 octobre 1963,» [En ligne]. Available: https://equipementsdelaroute.cerema.fr/versions-consolidees-des-9-parties-de-l-a528.html.

[10] AFNOR, «Signalisation routière verticale - Décors pour panneaux de signalisation - Méthode d'essai pour la mesure des caractéristiques colorimétriques.,» 1991 NF P98-522. [En ligne].

[11] UNITED NATIONS, «Addendum: 129: Regulation: 130 - Uniform provisions concerning the approval of motor vehicles with,» July 2013. [En ligne]. Available: https://unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/2013/R130e.pdf.

[12] T. Damkjaer, J. Girard, V. Muzet, R. Nuyttens, J. Ritter, K. Sorensen et Z. Huanyu, «Synthesis of bibliographic analysis of WG2 project RMCAD (Road Marking for Connected Automated Driving), task group 100,» 03 2023. [En ligne]. Available: https://nmfv.dk/wp-content/uploads/2023/09/SynthesisWG2_RMCAD_task100bibliography_2023-03-07.pdf.

[13] J. Ackermann, *Timing and LocationPerformance of Recent u-blox GNSS Receiver Modules*, 2020