

EXPERIENCES WITH ACQUIRING HIGHLY REDUNDANT SPATIAL DATA TO SUPPORT DRIVERLESS VEHICLE TECHNOLOGIES

Z. Koppanyi^{1,*}, C. K. Toth¹

¹Dept. of Civil, Environmental and Geodetic Engineering, The Ohio State University,
470 Hitchcock Hall, 2070 Neil Ave., Columbus, OH 43210 - (koppanyi.1, toth.2)@osu.edu

Commission II, WG II/10

KEY WORDS: Autonomous/Driverless Vehicles, Mobile Mapping, Sensor Calibration, Benchmark

ABSTRACT:

As vehicle technology is moving towards higher autonomy, the demand for highly accurate geospatial data is rapidly increasing, as accurate maps have a huge potential of increasing safety. In particular, high definition 3D maps, including road topography and infrastructure, as well as city models along the transportation corridors represent the necessary support for driverless vehicles. In this effort, a vehicle equipped with high-, medium- and low-resolution active and passive cameras acquired data in a typical traffic environment, represented here by the OSU campus, where GPS/GNSS data are available along with other navigation sensor data streams. The data streams can be used for two purposes. First, high-definition 3D maps can be created by integrating all the sensory data, and Data Analytics/Big Data methods can be tested for automatic object space reconstruction. Second, the data streams can support algorithmic research for driverless vehicle technologies, including object avoidance, navigation/positioning, detecting pedestrians and bicyclists, etc. Crucial cross-performance analyses on map database resolution and accuracy with respect to sensor performance metrics to achieve economic solution for accurate driverless vehicle positioning can be derived. These, in turn, could provide essential information on optimizing the choice of geospatial map databases and sensors' quality to support driverless vehicle technologies. The paper reviews the data acquisition and primary data processing challenges and performance results.

1. INTRODUCTION

Car manufactures, IT giants, and large numbers of start-up companies have been developing various technologies for autonomous cars at neck braking pace. These R&D efforts are quite interdisciplinary, and include primarily computer science, electrical and mechanical engineering, etc., (Geiger et al., 2012; Ibañez-Guzmán et al., 2012) and then social sciences to address ethical and legal concerns (Bonneton et al., 2016; Ibañez-Guzmán et al., 2012). The mapping community, in particular mobile mapping, is also a contributor to these efforts with bringing in long-term expertise in sensor integration, georeferencing, and mapping of the environment. For instance, currently, autonomous driving could greatly benefit from real-time centimeter level positioning accuracy that cannot be achieved only by GPS/GNSS in urban environment. Using accurate map of the environment and map matching, the required accuracy could be attained. Obviously, the availability of centimeter level geospatial data is important as well as the detailed representation of all objects along the transportation network.

The goals of the data acquisition discussed here are:

- The main objective is to acquire data streams from mobile platforms, including vehicles, bicycles, pedestrians, etc. These data are essential to testing vehicle sensing and maneuvering capabilities, and directly support research and development of autonomous vehicle technologies.
- A second objective is to create a high-definition map of the test area that includes mobile data collection and additional surveying of the area. The availability of such maps has a significant effect on autonomous driving by providing a detailed description of the object space of and around the transportation corridor, greatly improving the reliability of

the vehicle's self-localization and path planning capabilities.

- A third objective is to test the potential of Big Data / Data Analytics technology for map production based on highly redundant multi-sensor data, including crowdsourcing.

This paper describes the data acquisition platform and the first results of primary data processing of the highly redundant sensor arrangement, which is essential to create a benchmark data set. There are several widely used benchmark datasets that support navigation and object space reconstruction research. For instance, the KITTI Vision Benchmark Suite provides data from a Velodyne HDL-64E, stereo cameras, as well as, accurate navigation solution for georeferencing (Geiger et al., 2012). The Cityscape dataset contains stereo video sequences from 50 different cities for semantic analysis (Cordts et al., 2016). Researchers has access to 10 hours annotated images from The Caltech Pedestrian Detection Benchmark (Dollar et al., 2012). The mapping community also has several available data sources from mobile LiDAR sensors or images (Serna et al., 2014) that might be usable for autonomous research.

In contrast to these benchmark datasets, in this study, the focus is on acquiring highly redundant data in a wide range of quality. For example, the sensors cover the similar field of view around the platform, but they differ in image size, lens quality, sampling frequency, etc. This redundant approach allows for assessing to cross-reference the performance of different cameras and LiDAR scanners for specific tasks. As for any benchmark, a key element is to provide accurate georeferencing, and reference sensor calibrations, including the boresight and the hardware specific calibrations, such as lens distortion parameters for the imaging sensors. The paper describes procedures that address a few of the problems presented above and the results including the error budget analysis.

2. DATA ACQUISITION SYSTEM

2.1 Platform

A GMC Suburban, customized measurement vehicle, called GPSVan (Grejner-Brzezinska 1996), is used as a platform for the data acquisition, see Figure 1. The sensors installed on the platform can be categorized as navigation and mapping sensors. The navigation sensors, i.e. GNSS receivers and IMUs are located inside the van. Light frame structure installed on the top and front of the vehicle provides a rigid platform for the imaging sensors, such as LiDAR and different types of cameras. The final sensor configuration consists of two GPS/GNSS receivers, three IMUs, three high-resolution DSLR cameras for acquiring still images, 13 P&S (Point and Shoot) cameras for capturing videos, and seven LiDAR sensors. The location of the sensors on the GPSVan is shown in Figure 1 and the main sensor parameters are listed in Table 1. The four primary purposes of the various sensors are categorized as:

1. **Georeferencing and time synchronization:** GPS/GNSS and IMU sensors provide accurate time as well as position, attitude data of the platform, allowing for time synchronization and sensor georeferencing.
2. **Optical image acquisition:** these sensors are carefully calibrated and synchronized in order to derive accurate geometric data for mapping; for instance, by using stereo, multiple-image photogrammetric and computer vision methods.
3. **Video logging:** these sensors provide a continuous coverage of the environment during the tests. The quality of these sensors does not allow for accurate time synchronization and accurate calibration, which would be comparable to high quality still image sensors. Nevertheless, the moderate geometric accuracy combined with the high image acquisition rate allows for object extraction and tracking, such as traffic signs, road signs,

and obstacles, etc. In addition, dynamic objects, such as vehicles, cyclist, pedestrians, etc., can be tracked.

4. **3D data acquisition:** Velodyne LiDAR sensors allow for direct 3D data acquisition that can be used for object space reconstruction, and object tracking.

The field of views of the imaging sensors, including the LiDAR and cameras, around the vehicle are shown in Figure 2; note the sensing range is not shown. The sensor arrangement and field of views are designed to acquire highly redundant data that can equally support the high definition mapping of the environment and the algorithmic research related to driverless vehicle technologies. Conventional mobile mapping systems typically utilize high-resolution imaging systems, and use narrower observation field of view around the platform, and the completeness of the acquisition is provided by the details, captured incrementally as the platform is moving. Using accurate georeferencing, the data is merged during processing to obtain a seamless geospatial product. For autonomous driving, the sensor system must be aware of the surroundings at all the time, as the decision-making has to be done in a fraction of a second. Consequently, compared to mapping-orientated data acquisition, the arrangement and alignments of the sensors have to be designed to cover the entire field of view around the vehicle.

Since affordability is a key element of the sensor design for autonomous vehicles, the highly redundant data acquisition provides an ideal data set for performance comparisons; mainly, by providing information of the geospatial navigation and mapping accuracy that can be realistically achieved with simple and inexpensive sensors. For example, what is the performance difference when image sequences from a smartphone, GoPro or a P&S camera are used? Sensor orientation, in particular for scanning sensors, is also important and can be evaluated. Also, point cloud quality can be compared for photogrammetrically derived and LiDAR created data.

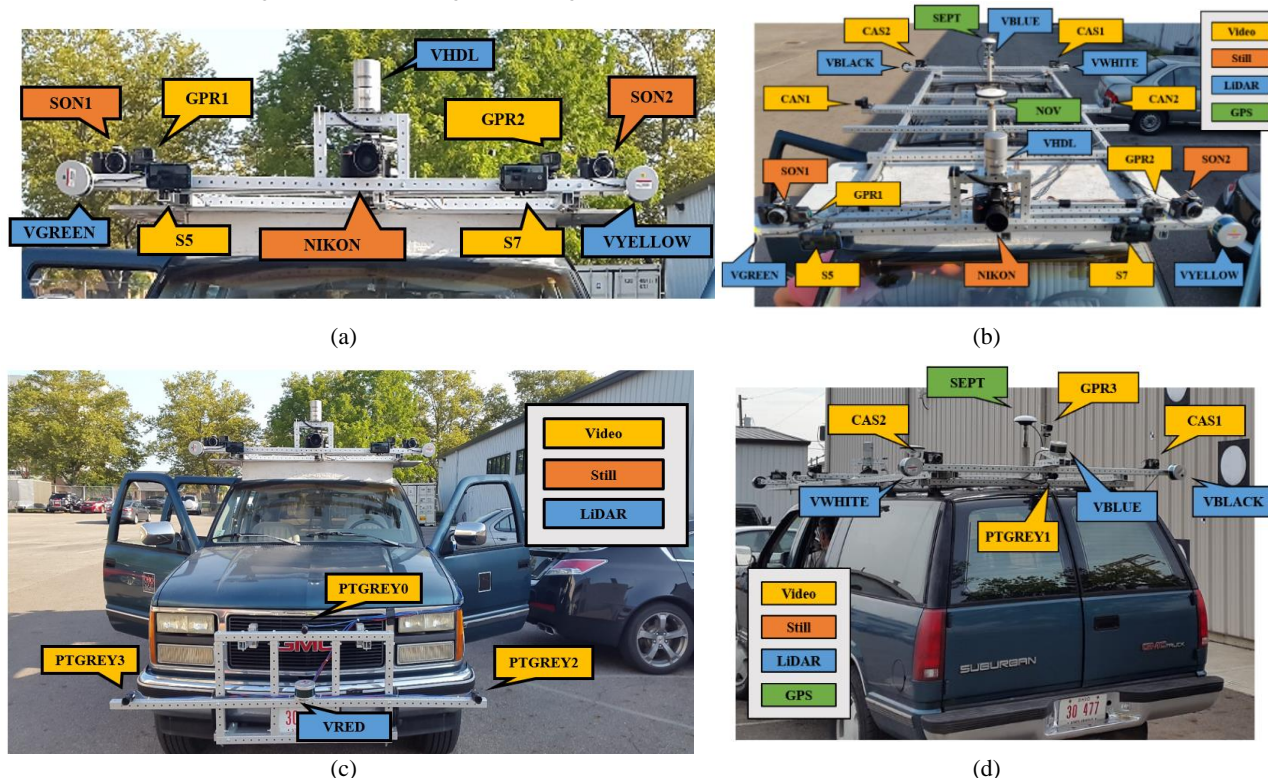


Figure 1. The placement of the imaging sensors: (a) front top view, (b) top view, (c) front view, and (d) rear view.

Table 1. Overview of the sensors; see explanation in the text.

Type	Sensor Model	Sensor ID	Resolution / Ang. resolution ¹	Location	Sampling Frequency	Field of View ¹	Purpose [*]
GPS	Septentrio PolaRx5 GPS	SEPT	-	Top	10 Hz	-	(1)
GPS	Novatel DL-4 GPS	NOVATEL	-	Top	5 Hz	-	(1)
IMU	MicroStrain 3DM-GX3	MS	-	Inside	200 Hz	-	(1)
IMU	H764G IMU 1	H764G	-	Inside	200 Hz	-	(1)
IMU	H764G IMU 2	H764G	-	Inside	200 Hz	-	(1)
Still image	Sony Alpha 6000 (ILCE)	SON1	6000 x 4000	Front, Right	0.5 Hz	96° x 73°	(2)
Still image	Sony Alpha 6000 (ILCE)	SON2	6000 x 4000	Front, Left	0.5 Hz	96° x 73°	(2)
Still image	Nikon D800	NIKON	7360 x 4912	Front, Center	1 Hz	40° x 40°	(2)
Video	Canon PowerShot SX710	CAN1	1920 x 1080	Side, Right	30 Hz	70° x 55°	(3)
Video	Canon PowerShot G7 X	CAN2	1920 x 1080	Side, Left	30 Hz	74° x 55°	(3)
Video	Casio EX-H20G	CAS1	1280 x 720	Back, Right	30 Hz	72° x 57°	(3)
Video	Casio EX-H20G	CAS2	1280 x 720	Back, Left	30 Hz	72° x 57°	(3)
Video	GoPro HERO5 Black	GPR1	3840 x 2160	Front, Right	30 Hz	102° x 70°	(3)
Video	GoPro HERO5 Black	GPR2	3840 x 2160	Front, Left	30 Hz	102° x 70°	(3)
Video	GoPro HERO3+ Black	GPR3	1920 x 1080	Back, Center	30 Hz	91° x 73°	(3)
Video	Point Grey Flea3 8.8	PTGREY0	3840 x 2880	Front, Bottom	5 Hz	45° x 45°	(3)
Video	Point Grey Flea3 1.3	PTGREY1	1280 x 960	Back, Center	5 Hz	45° x 45°	(3)
Video	Point Grey Flea3 1.3	PTGREY2	1280 x 960	Front, Left	5 Hz	45° x 45°	(3)
Video	Point Grey Flea3 1.3	PTGREY3	1280 x 960	Front, Right	5 Hz	45° x 45°	(3)
Video	Samsung S5	S5	1920 x 1080	Front, Right	30 Hz	45° x 45°	(3)
Video	Samsung S7	S7	1920 x 1080	Front, Left	30 Hz	45° x 45°	(3)
LiDAR	Velodyne HDL-32E	VHDL	H/V: 0.2°/1.33°	Front, Top	20 Hz	H/V: 360°, 40°	(4)
LiDAR	Velodyne VLP-16	VRED	H/V: 1.33°/0.2°	Front, Bottom	20 Hz	H/V: 30°, 360°	(4)
LiDAR	Velodyne VLP-16	VGREEN	H/V: 0.2°/1.33°	Front, Right	20 Hz	H/V: 360°, 30°	(4)
LiDAR	Velodyne VLP-16	VYELLOW	H/V: 0.2°/1.33°	Front, Left	20 Hz	H/V: 360°, 30°	(4)
LiDAR	Velodyne VLP-16	VWHITE	H/V: 0.2°/1.33°	Back, Left	20 Hz	H/V: 360°, 30°	(4)
LiDAR	Velodyne VLP-16	VBLUE	H/V: 1.33°/0.2°	Back, Center	20 Hz	H/V: 30°, 360° ²	(4)
LiDAR	Velodyne VLP-16	VBLACK	H/V: 0.2°/1.33°	Back, Right	20 Hz	H/V: 360°, 30°	(4)

* – see explanations in the text

¹ – angles defined in the platform's coordinate system

² – rotation plane is declined 30° from the horizon

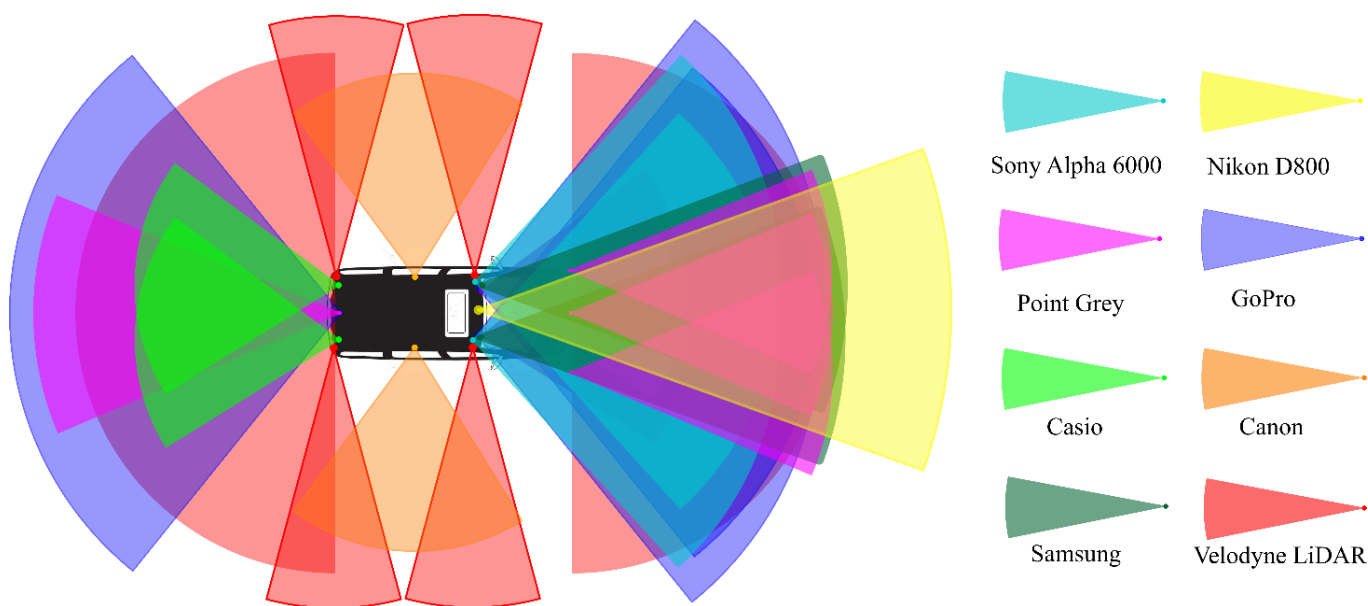


Figure 2. Field of views of the imaging sensors around the GPSVan.

3. PROCESSING

As a standard preprocessing, the sensor orientation must be georeferenced and data streams must be transformed from the sensor coordinate system to a common or global coordinate system prior any further processing. This is a typical problem in mobile mapping system. Here, it is assumed that after the raw sensor stream is preprocessed, the geometric data are available as $\mathbf{x}_s = [\lambda x, \lambda y, \lambda z, \lambda] \in \mathbb{R}^4, \lambda \neq 0$, 3D points with homogenous coordinates defined in the sensor coordinate system. LiDAR sensors directly provide the 3D point coordinates, while, 3D points from cameras can be derived based on stereo or multi-view camera equations. To transform \mathbf{x}_s point defined in the sensor coordinate system to a common global coordinate system is described as:

$$\mathbf{x}_g = T_{g,p} T_{p,s} \mathbf{x}_s. \quad (1)$$

where $T_{p,s}$ is the 4 by 4 time-independent sensor to platform homogenous transformation matrix, also called boresighting, and the $T_{g,p}$ is the time-dependent platform to global transformation, also known as georeferencing.

3.1 Time synchronization

All data streams are synchronized to the GPS time (GPST). However, the accuracy of the time synchronization of the sensors varies. Obviously, the GNSS and IMU sensors are very well synchronized to the GPST. Similarly, Velodyne sensors are accurately synchronized by the 1PPS signal and NMEA messages provided by GPS/GNSS receivers. The time synchronization of the high-resolution cameras was accurately aligned to the GPST. The two Sony cameras were triggered using the 1PPS signals, while the Nikon was using a self-timer for image capture. For all the three cameras, the triggering time of the images were also logged through a connection between the flash output of the cameras and the event input of the GPS/GNSS receivers. There was no accurate time synchronization for the video camera streams. Coarse time tagging was obtained by matching high-resolution images with selected video frames of dynamic content. For example, walking pedestrians or runners are considered as good objects for the matching because similar still images of the same dynamic content can be found in the image sequences across the cameras. In addition, for a few cameras the images were directly logged by a computer, in which case, through the CPU time, the image acquisition time could be estimated with respect to the GPS time. In both cases, time synchronization accuracy is estimated to be in the range of 0.01 – 0.1 s.

3.2 Georeferencing

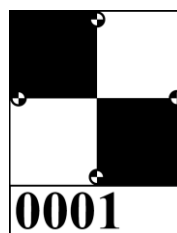
The $T_{g,p}$ platform to global coordinate transformation (platform georeferencing) is a time dependent transformation that provides the position and orientation of the platform. This transformation can be obtained with integrating GPS/GNSS and IMU observations. The results presented in this study are derived based on the data from the Septentrio GNSS receiver and the navigation grade H764G IMU, using the loosely coupled integration model. The GPS/GNSS trajectory solution was derived by PPK (post-processed kinematic) using carrier-phase observations. GPS/GNSS signal was lost several times during the tests due to "urban canyon" effect. The use of navigation grade IMU, however, provided reliable bridging for the gaps in the GPS/GNSS observation.

3.3 Sensor Calibration Range

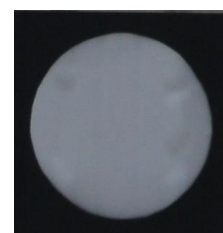
The goal of the sensor calibration is to determine the position and orientation (boresighting) of the sensors in the platform coordinate system, expressed in the $T_{p,s}$ transformation matrix in Equation 1. In addition, other sensor related parameters may be determined, such as the distortion parameters for the cameras. In order to estimate these parameters for the camera and LiDAR sensors, a test range was created at the main facility of the OSU Center for Automatic Research (CAR), see Figure 3. At a corner of the building, 40 optical camera and 5 LiDAR targets were attached to the walls, and imaged at various locations and orientations, see Figure 4. The target locations were measured by total station using triangulation at sub-centimeter accuracy. The points are also tied to the global system using GPS/GNSS measurements.



Figure 3. GPSVan at the calibration range.



(a) Size: 14 cm x 12 cm



(b) Size: 50 cm x 50 cm

Figure 4. Camera (a) and LiDAR (b) targets.

3.4 Camera Calibration and Boresighting

The cameras were first laboratory calibrated, and the obtained parameters were subject to adjustment during the in-situ calibration process (Fraser, 2012). The estimation of the internal and external orientation parameters of the cameras consists of two steps. First, the cameras position and orientation are estimated in the global coordinate system using bundle adjustment (McGlone, 2013), and then the estimated parameters are transformed into the platform coordinate system using the georeferencing solution. During the camera calibration, the targets were manually measured in the images, taken from various platform positions. Since the global coordinates of the targets are known from surveying and the pixel locations of the targets are measured, the rotations ($R_{g,s}$) and positions ($\mathbf{t}_{g,s}$) of the cameras in the global frames are estimated using bundle adjustment without tie points. The linear calibration parameters (i.e., focal length, and principle point coordinates) as well as the three radial and two tangential distortion parameters (k_1, k_2, k_3, p_1, p_2) are also estimated in the bundle adjustment. Figure 5 shows a result for one camera; blue dots depict the targets, blue lines are the rays, and the camera planes are shown in red.

The orientations ($R_{g,s}$) and positions ($\mathbf{t}_{g,s}$) of each camera defined in the global frame are known from the bundle adjustment. In the next step, these parameters have to be transferred to the platform coordinate system. The GPS/IMU integrated navigation solution can be used to obtain the $R_{g,p}$ and $\mathbf{t}_{g,p}$ rotations and translations between the platform and the global coordinate system, described by the following equations:

$$\begin{aligned} R_{p,s} &= R_{g,p}^{-1} R_{g,s}, \\ \mathbf{t}_{p,s} &= \mathbf{t}_{g,p} - \mathbf{t}_{g,s}. \end{aligned} \quad (2)$$

where $R_{p,s}$ and $\mathbf{t}_{p,s}$ are the platform to sensor rotation matrix and the translation vector, respectively. Since these transformations are different for each platform position, the average of the estimated orientation and translation parameters is calculated to obtain the final transformation parameters. In addition, the statistical evaluation of distribution of the camera positions allows for assessing the accuracy of the final solution that includes the errors of the bundle adjustment and the georeferencing.

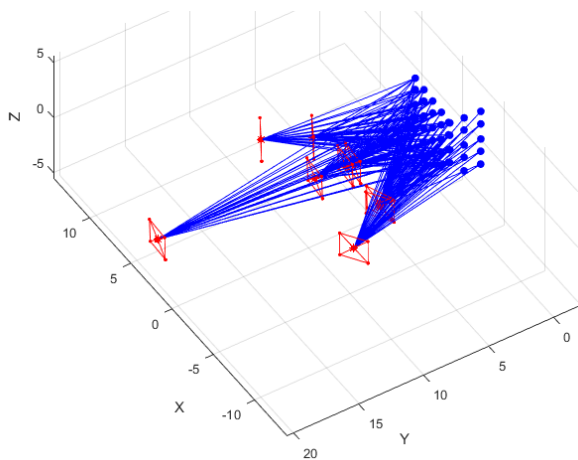


Figure 5. Result of the bundle adjustment: targets (blue dots), ray (blue lines), camera plane (red).

3.5 LiDAR Boresighting

The center of the Velodyne VLP-16 and HDL-32 sensors is well defined, and thus, the positions of these sensors in the platform coordinate system can be easily measured with measuring tape. For the sensor orientations, one can simply manually position the sensor to be in coincidence with the platform's local coordinate system. The accuracy of this direct approach, however, is modest, and it can only provide an approximate or nominal sensor alignment. Indirect methods provide better accuracy, and use object space information for estimating the sensor boresighting, for examples, see Glennie and Lichti, 2010; Atanacio-Jiménez et al., 2011 or Guindel et al., 2017.

The approach used in this study is based on the idea presented by Csanyi and Toth, 2007. The LiDAR specific targets are extracted from the point cloud, and their surveyed coordinates are used to calculate the sensor orientation; the final step of the computation is similar to the photogrammetric indirect georeferencing. Csanyi and Toth applied this concept to airborne LiDAR. The determination of the orientation parameters using this concept for the Velodyne sensors is different due to the limited angular resolution and field of view along the rotation axis.

Consequently, less number of points can be captured from the target; in particular, as the object-sensor distance increases. Nevertheless, several points can be captured from the 50 x 50 cm targets from a range of 3-7 m, and thus used for estimating the boresight parameters. Figure 6 shows a position with three targets seen by the HDL-32E scanner; the targets are highlighted in the figure.

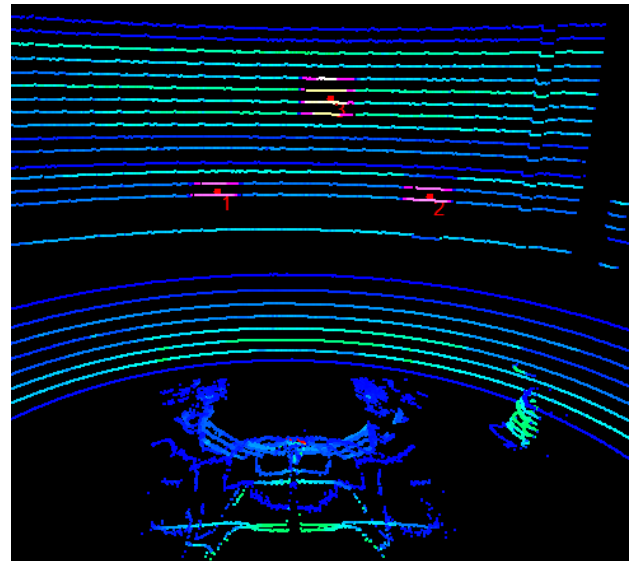


Figure 6. Target points (magenta color) in the point cloud captured by Velodyne HDL-32E scanner.

After segmenting the target points from the background, the next step is to find the center of the targets, based on the captured points. First, the captured target points are transformed to a local plane, defined as the best fitting plane to the points. Figure 7 shows the captured points of a target on this local plane; the colors represent the intensity values of the points. High intensity values (bright colors) are assumed to be captured from the white inner circle of the target, while low intensity indicates that the points are captured around the edges of the target. The center of the white circle can be found by fitting a circle (light blue) with the known radius to the points.

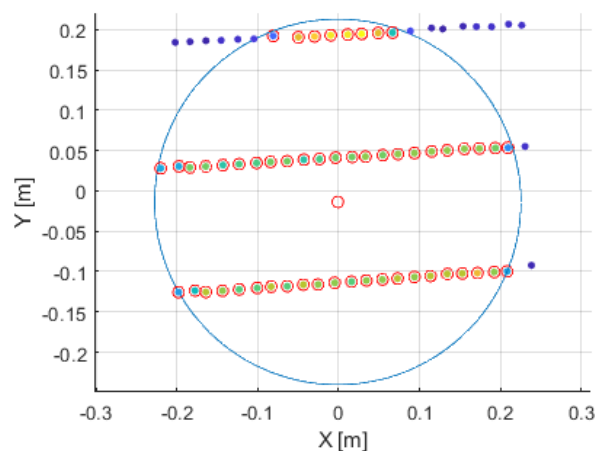


Figure 7. Captured points from a target. The color depicts the intensity: dark: high, and light: low intensity; blue circle represents the best fitting circle around the points.

As opposed to the Csanyi and Toth approach, where they applied a grid-based search to find the center of the circle, here, we solve the following quadratically constrained quadratic program (QCQP) to find the location of the "best" fitting circle that satisfies the condition presented above (Boyd and Lieven, 2004):

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \mathbf{x}^T \mathbf{x} \\ & \text{subject to} && \frac{1}{2} \mathbf{x}^T \mathbf{H}_i \mathbf{x} + \mathbf{k}_i^T \mathbf{x} + d_i \leq 0, i = 1..n \end{aligned} \quad (3)$$

where $\mathbf{x} = [x, y]^T$ is the unknown center of the disc, and n is the number of the captured points. The $\mathbf{H}_i, \mathbf{k}_i, d_i$ elements of the constraints depends on whether the constraint equation is written for a point that is on the white disc or outside. For a (x_i, y_i) point lie inside the disc, $\mathbf{H}_i = \mathbf{I}$, $\mathbf{k}_i = -[x_i, y_i]^T$, and $d_i = \frac{1}{2}(x_i^2 + y_i^2 - r^2)$, and for a (x_j, y_j) point outside of the disc, $\mathbf{H}_j = -\mathbf{I}$, $\mathbf{k}_j = [x_j, y_j]^T$, and $d_j = -\frac{1}{2}(x_j^2 + y_j^2 - r^2)$, where r is the radius of the disc.

After solving Equation 3, for instance, following Mathworks Team, 2018, the circle center is transformed back from the local plane to the sensor coordinate system. An \mathbf{x}_s target center defined in the sensor coordinate system can be transformed to the global coordinate system with applying first a sensor to platform transformation ($T_{p,s}$), and then, a platform to global system transformation ($T_{g,p}$), see Equation 1. The $T_{g,p}$ transformation can be determined from the GPS/IMU integrated navigation solution, and the $T_{p,s}$ homogenous transformation is the LiDAR boresighting, that has to be estimated. The \mathbf{x}_g target centers in the global reference frame are known from surveying, and thus the rigid body rotation parameters of the $T_{p,s}$ transformation matrix can be obtained with least squares:

$$\text{minimize (w.r.t. } T_{p,s}) \|\mathbf{x}_g - T_{g,p} T_{p,s} \mathbf{x}_s\|_2^2 \quad (4)$$

Note that the unknown parameters of the $T_{p,s}$ transformation are only the three rotation parameters, because the position of the sensor is already measured, see above. Thus, three or more targets have to be measured in order for estimating the unknown boresight parameters.

4. DATA ACQUISITION

Two test sites were selected for the pilot data acquisition; both located at the campus of The Ohio State University. The first route is located at west campus and connects two research facilities, see Figure 8a. This route has moderate vehicle and low pedestrian traffic, and data was acquired in completing 9 loops in about one hour. The second route is on main campus, see Figure 8b. This area is heavily used by students and cyclists, and therefore, this dataset can be used for investigating complex scenarios; for example, testing various pedestrian, cyclist or other object detection algorithms, or visual navigation methods with rapidly changing dynamic content. In addition, this area is a partially GPS/GNSS-challenged area due to tall buildings located along the route. This dataset contains 15 loops, acquired in about 4 hours. The two datasets represent about 5 TB raw data.

5. DISCUSSION

The results of the camera calibration for the Nikon, Sony and GoPro cameras showed that the reprojection errors are under 0.6 pixel. The *a posteriori* accuracies are 2-3 cm and 0.2-0.3° for the camera positions and orientations, respectively. These parameters have to be transferred from the global to the platform coordinate system based on Equation 2 using the georeferencing solution. After this transformation, the standard deviation of the camera coordinates is 1-2 cm for X-Y direction, and slightly larger in the Z direction (3-4 cm). The accuracy of the orientation parameters is also slightly larger (0.2-0.3°). Note that only approximate boresight parameters are provided for the rest of the cameras.

The results of the orientation estimation for the LiDAR sensors are listed in Table 2. The table also shows the number of detected targets and platform positions. It is noteworthy that this error is calculated based on the Jacobian matrix, numerically approximated from Equation 4. The results indicate worse orientation estimation than it was for the cameras; the $\hat{\sigma}_R$ *a posteriori* errors are around 1.1°. Here, the main issue is the low point density. The impact of the point density on the boresight accuracy can be seen in the table. The HDL-32E sensor, also tagged as VHDL, is able to capture denser point clouds than the VLP-16, consequently, the reported *a posteriori* error is 0.41° degree as opposed to the VLP-16 error of 1.1°. It is noteworthy that these errors already include the georeferencing error, since the $T_{g,p}$ matrix is used in Equation 4. For example, the VHDL attitude errors results in an 8 cm error, perpendicular to the ray at 10 m distance. The attitude errors for the VLP-16 sensors are larger, and thus, the point clouds obtained from these sensors may need to be aligned to the points captured by the HDL-32E sensor to increase accuracy.

The GPS/IMU solution can be seen in Figure 8a and 8b. Note that GPS/IMU integration provides a seamless trajectory solution in Figure 8b. The positioning accuracy is under 1.5 cm after the filter has converged. The estimated attitude standard deviation is under 0.01° for the roll, pitch and under 0.02° for the yaw components. Figure 9 shows a sample from the synchronized camera streams, LiDAR and GPS/GNSS. Finally, Figure 10a shows the georeferenced HDL-32E LiDAR data for a section of the main campus loop, and Figure 10b shows the same area with all LiDAR sensors' data combined.

Table 2. Results of the LiDAR orientation estimation.

Sensor ID	Sensor type	# of Targets	# of Position	$\hat{\sigma}_R$ [°]
VHDL	HDL-32E	9	3	0.41
VRED	VLP-16	5	3	1.19
VGREEN	VLP-16	4	2	1.68
VWHITE	VLP-16	7	4	1.12
VBLUE	VLP-16	3	3	1.15
VBLACK	VLP-16	4	1	1.42

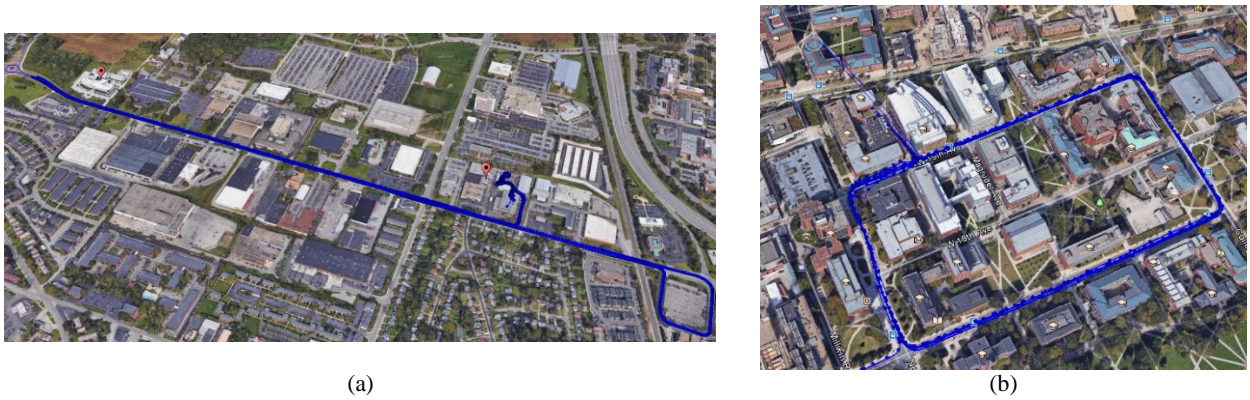


Figure 8. GPS/IMU trajectory solutions for the two test routes: (a) west campus, and (b) main campus loop.

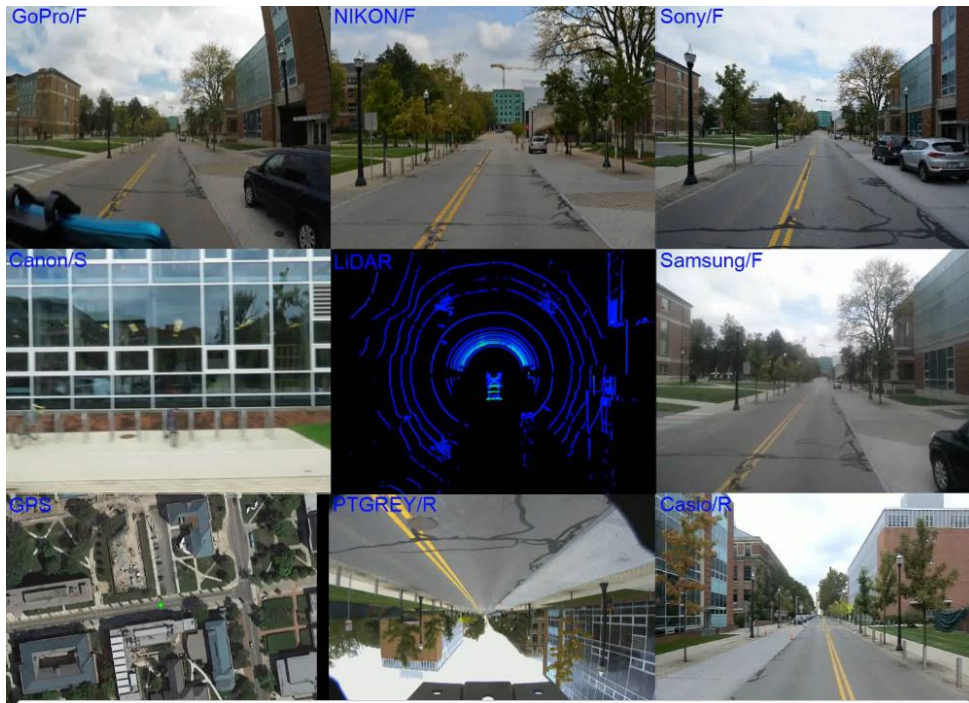


Figure 9. Data streams from GoPro (video, top left), Nikon (high resolution still images, top center), Sony (high resolution still images, top right), Canon (P&S, video, middle left), Velodyne HDL-32E (LiDAR, middle center), Samsung (mobile phone, built-in camera, middle right), GPS/GNSS (bottom left), PointGrey (video, bottom center), and Casio (P&S, video, middle right).

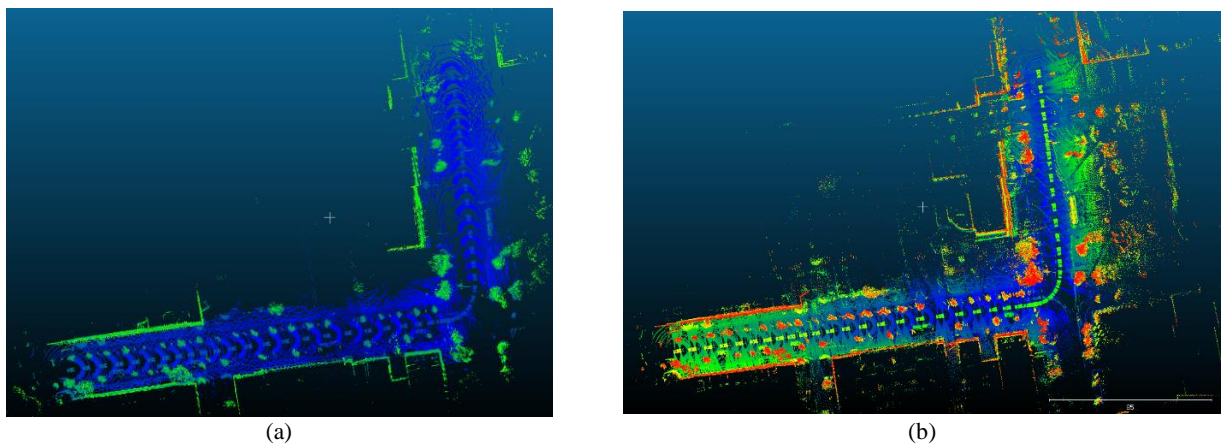


Figure 10. A sample from the georeferenced HDL-32E data (a), and all LiDAR sensors' data combined (b); points with 5 m above the road surface are removed, height is color-coded.

6. SUMMARY AND CONCLUSION

The paper describes a data acquisition effort to collect highly redundant geospatial data to support driverless vehicle technology research. A measurement vehicle was equipped with a high performance georeferencing system, multiple still and video cameras, and several LiDAR sensors. Data was acquired in multiple sessions at two locations at The Ohio State University campus.

The acquired data can be used to obtain high definition map of the test areas, and to serve as benchmark data for algorithmic research to support autonomous driving. The sensors simultaneously acquired data around the vehicle with high redundancy, meaning that multiple sensors collected data from the same and/or overlapping areas around the platform. This dataset allows for not just comparing the performance of sensors with different imaging qualities/capabilities, but also to evaluate the algorithms in terms of robustness against various types of data streams. Selected aspects of the platform georeferencing and calibration of the camera and LiDAR sensors are discussed in the paper. The achieved georeferencing and sensor modeling accuracy provides a good reference solution for algorithmic evaluation.

Currently, a software tool is developed to provide a simple interface for visualizing, editing, combining and exporting the various data streams. The ultimate goal is to share the collected data along with the developed tools and documentations through a webpage, and thus, to make it available as benchmark dataset.

ACKNOWLEDGEMENTS

Support from CAR, OSU and Teradata is greatly appreciated.

REFERENCES

- Atanacio-Jiménez, G.; González-Barbosa, J.-J.; Hurtado-Ramos, J. B.; Ornelas-Rodríguez, F. J.; Jiménez-Hernández, H.; García-Ramirez, T.; González-Barbosa, R., 2011. LIDAR Velodyne HDL-64E Calibration Using Pattern Planes. *International Journal of Advanced Robotic Systems* 8(5) pp. 70–82.
- Bonnefon, J.-F.; Shariff, A.; Iyad Rahwan., 2016. The Social Dilemma of Autonomous Vehicles. *Science* 352 (6293) pp. 1573–76. <https://doi.org/10.1126/science.aaf2654>.
- Boyd, S.; Lieven, V., 2004. *Convex Optimization*. Cambridge University Press, Cambridge, ISBN 978-0-521-83378-3, pp. 152-167.
- Cordts, M.; Omran, M., Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B., 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. ArXiv:1604.01685 <http://arxiv.org/abs/1604.01685>.
- Csanyi, N.; Toth, C. K., 2007. Improvement of Lidar Data Accuracy Using Lidar-Specific Ground Targets. *Photogrammetric Engineering and Remote Sensing* 73(4) pp. 385-396.
- Dollar, P.; Wojek, C.; Schiele, B.; Perona, P., 2012. Pedestrian Detection: An Evaluation of the State of the Art. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34 (4):743–61. <https://doi.org/10.1109/TPAMI.2011.155>.
- Fraser, C. S., 2012. Automatic Camera Calibration in Close-Range Photogrammetry. *Photogrammetric Engineering and Remote Sensing* 79(4):381-388
- Geiger, A.; Lenz, P.; Urtasun, R., 2012. Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite. In 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 3354–61. <https://doi.org/10.1109/CVPR.2012.6248074>.
- Glennie, C.; Lichti, D. D., 2010. Static Calibration and Analysis of the Velodyne HDL-64E S2 for High Accuracy Mobile Scanning. *Remote Sensing* 2:1610–24.
- Guindel, C.; Beltran J.; David, M.; Garcia, F., 2017. Automatic Extrinsic Calibration for Lidar-Stereo Vehicle Sensor Setups. In *IEEE International Conference on Intelligent Transportation Systems*, 2017
- Grejner-Brzezinska, D., 1996. Positioning Accuracy of the GPSVan. In Proceedings of the 52nd Annual Meeting of The Institute of Navigation. Cambridge, MA, USA.
- Ibañez-Guzmán, J.; Laugier, C.; Yoder, J.-D.; Thrun, S., 2012. Autonomous Driving: Context and State-of-the-Art. (Book Chapter) *Handbook of Intelligent Vehicles*, 1271–1310. Springer-Verlag London Ltd.
- Mathworks Team, 2018. Linear or Quadratic Objective with Quadratic Constraint, <https://www.mathworks.com/help/optim/ug/linear-or-quadratic-problem-with-quadratic-constraints.html?requestedDomain=true> (12 January 2018)
- McGlone, J. Chris., 2013. *Manual of Photogrammetry*. American Society of Photogrammetry and Remote Sensing, 6th Edition, Bethesda, Maryland, USA.
- Serna, A.; Marcotegui, B.; Goulette, F.; Deschaud, J.-E., 2014. Paris-Rue-Madame Database: A 3D Mobile Laser Scanner Dataset for Benchmarking Urban Detection, Segmentation and Classification Methods. In *International Conference on Pattern Recognition Applications and Methods*, Angers, Loire Valley, 6-France, March 8, 2014.