NEW SOURCE OF GEOSPATIAL DATA:
CROWDSENSING BY ASSISTED AND AUTONOMOUS VEHICLE TECHNOLOGIES

C. K. Toth[*], Z. Koppanyi, M. G. Lenzano

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

Commission IV, WG IV/4

KEY WORDS: Crowdsensing/Crowdsourcing, Autonomous/Driverless Vehicles, Mobile Mapping, Deep Learning

ABSTRACT:

The ongoing proliferation of remote sensing technologies in the consumer market has been rapidly reshaping the geospatial data acquisition world, and subsequently, the data processing as well as information dissemination processes. Smartphones have clearly established themselves as the primary crowdsourced data generators recently, and provide an incredible volume of remote sensed data with fairly good georeferencing. Besides the potential to map the environment of the smartphone users, they provide information to monitor the dynamic content of the object space. For example, real-time traffic monitoring is one of the most known and widely used real-time crowdsensed application, where the smartphones in vehicles jointly contribute to an unprecedentedly accurate traffic flow estimation. Now we are witnessing another milestone to happen, as driverless vehicle technologies will become another major source of crowdsensed data. Due to safety concerns, the requirements for sensing are higher, as the vehicles should sense other vehicles and the road infrastructure under any condition, not just daylight in favorable weather conditions, and at very fast speed. Furthermore, the sensing is based on using redundant and complementary sensor streams to achieve a robust object space reconstruction, needed to avoid collisions and maintain normal travel patterns. At this point, the remote sensed data in assisted and autonomous vehicles are discarded, or partially recorded for R&D purposes. However, in the long run, as vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication technologies mature, recording data will become a common place, and will provide an excellent source of geospatial information for road mapping, traffic monitoring, etc. This paper reviews the key characteristics of crowdsourced vehicle data based on experimental data, and then the processing aspects, including the Data Science and Deep Learning components.

1. INTRODUCTION

The past decade has seen phenomenal developments in sensor technologies, and by now our environment is continuously observed by an ever growing network of navigation, imaging, mapping and a variety of other sensors. In the developed world, the number of inexpensive sensors outnumbers the population by a large margin, and the trend is still sharply increasing. The general framework is provided by the IoT (Internet of Things), which provides for access and control sensor from virtually anywhere. Smartphones represent the highest sensor integration on any mobile platform, they have 8-10 built-in sensors that make these devices extremely powerful navigation and imaging/mapping tools. Furthermore, these devices provide an easy access to other sensor deployed in our daily life, such as wearable technologies and smart homes.

Most of the sensor data is used locally and not archived currently, but as communication technologies are becoming more affordable along with cloud services, the trend is to archive the data, as it can provide valuable individual and global information for the user, companies and governments. For example, providing location information of smartphones in vehicle creates the best possible data for traffic flow estimation, and these applications are one of the most popular ones of smartphones. In fact, people tend to prefer them compared to dashboard built-in navigation systems due to the currency of the data. Note some new cars are only providing visual interface to the smartphone apps instead of offering a navigation system. Health-related personal data is typically not shared due to privacy concerns, though it has enormous potential for research and disease prevention.

An important aspect of the acquired sensor data is that it typically comes with location information. While this is the primary information source for the smartphone based navigation apps, the use of the spatial context of the sensor data is still not fully exploited. For example, huge volumes of images are acquired with varying georeferencing accuracy, yet current applications don't use it; say, for example, mapping, navigation or object space reconstruction. The trend, however, is that navigation and imaging sensors are increasingly used together.

The Smart City concept is based on fully exploiting the technology potential to use and share information to make the life of people living in big and dense urban areas better by improving all the services provided by companies and governments (Su et al., 2011). One key element of a Smart City is the efficient mobility that considers all the citizens transportation needs, and not just people who are driving. For example, people with disabilities have specific needs to access public transportation from their homes and to get to a doctor's office in a health complex. Recent advances in vehicle technologies have started to offer various levels of autonomy, providing a new dimension to the process of improving mobility in cities.

Autonomous vehicle (AV) technologies, a.k.a driverless car, assisted driving (Advanced Driver-Assistance Systems, ADAS), are rapidly developing, as traditional car manufactures, IT giants, and large numbers of start-up companies have been devoting unprecedented R&D efforts to advance this field. The main disciplines for AV technologies are 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 (Bonnefon et al., 2016; Ibañez-Guzmán et al., 2012).

---

[*] Corresponding author

Most of the early AV technologies have primarily focused on sensing the environment to avoid obstacles, and thus provide for safe driving. But no or limited attention was paid to use the acquired and interpreted data to create or update and existing map. Note that a state-of-the-art AV has comparable sensing capacity a mobile mapping system. Furthermore, it has been also overlooked that using accurate and high resolution map data can improve the process how the vehicles sense and analyze their immediate environment. This paper looks into these aspect of AV technologies, in other words, the potential of crowdsensing to acquire geospatial data along transportation corridors and cities.
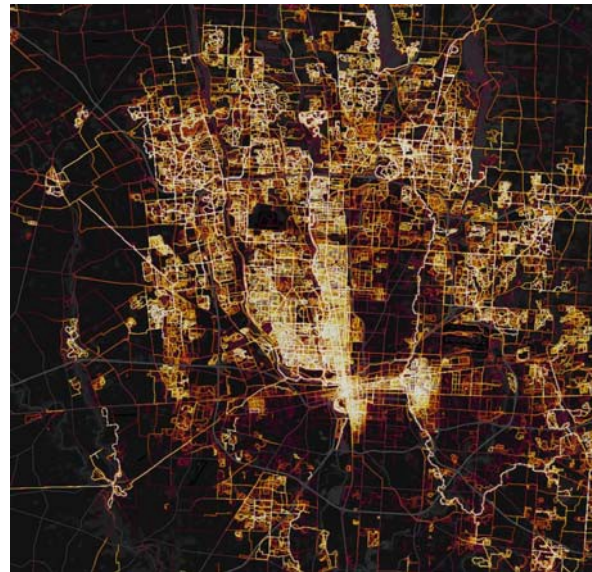
2. CROWDSENSING

Crowdsourcing, created in the Information Technology industry about 10 years ago, originally aimed at combining resources via the internet to solve large tasks. By now, crowdsourcing has been used in a much broader sense than data/computer science. In geospatial practice, crowdsensing is the more adequate terminology, as it is primarily about acquiring data (Heipke, 2010; Toth and Jozkow, 2015). Figure 1 shows an early crowdsourced project, where the movement of the San Francisco taxis were tracked at the SFO airport area (Piorkowski, 2009). Note that at that time, smartphones were less advanced.

Figure 1. Crowdsourced GPS data of taxis from 2009
(CRAWDAD database)

Today most of the smartphone apps attribute location to the logged data streams. For example, fitness apps may log heart rate and other important parameters during exercising, and the entire data stream is stored in the cloud, so the user can access his/her history, compute statistics, etc. In addition, using aggregated data, identity removed, valuable information can be extracted. Figure 2 shows heatmaps based on running/jogging and bicycling activities in the Columbus, OH area, data provided the Strava fitness application. These maps have the traditional location information, for example the bike trails are quite visible, as cyclists prefer them for safety reasons. Running/jogging is less confined to trails, as it requires less distance and thus can be easily done in residential areas. Beside geospatial data, for example, there is socio-economic information in these maps. The density/intensity is much lower in the poorer southern part of the city. People in affluent neighborhoods tend to pay more attention to their health and exercise more, as opposed to economically depressed areas, where fitness is not a high priority for the residents.

(a)) Running/jogging

(b) Bicycling
Figure 2. Crowdsensed exercise activities
in Columbus, OH, USA, 2017

While GPS describes the platform motion at few meter accuracy in general, alone it provides no information of the environment. With the proliferation of imaging sensors, the potential exist that area where the crowdsensing platform travels can be imaged, and thus geospatial data can be acquired. Compared to GPS, there are main differences in the practical use of the sensors. GPS requires no cooperation from the user, once the application has started it logs the data in the background, and no attention is needed from the user. In contrast, imaging sensors should be kept in a position that allows for a reasonable coverage of the area. Furthermore, imaging data by orders larger than GPS, so storing and/or transferring through the network are still a challenge. These problems are less severe on vehicles where there are plenty of resources and sensor mounting is structured. Helmet mounted GoPro and windshield/dash cameras are examples when the platform trajectory area is continuously imaged; for

entertainment and video evidence, respectively. In these cases, the long-term archiving and sharing is not typical.

With the increasing use of AV technologies in the future, there is a tremendous potential to record and aggregate the image sensor data, which then can be used for mapping of the transportation corridors and cities. The real question is how to pass the imagery to the cloud. Vehicle-to-Vehicle (V2V) technology is designed for local communication, and not adequate for handle image sensor data. Vehicle-to-Infrastructure (V2I) technology, however, is for communicating between the vehicles and transportation management and control system, and potentially can handle the task of accepting the image streams. Note that V2X facilitates both V2V and V2I communication through a central unit.

3. STATE-OF-THE-ART IN AV

For the purpose of using AV image data for mapping, the important elements are the number of imaging sensors, their type, and data characteristic, such as spatial resolution, frame rate, accuracy, etc. The environment is generally sensed by cameras, laser sensors, radar and ultrasonic sensors. Clearly, all these sensors represent important sensing characteristics, and ideally should be included on all platforms. However, affordability is a serious concern for stock vehicles, where the cost of the sensors must be limited to keep the vehicle price at an acceptable level. Currently, optical imaging dominates the marker, as these sensors are inexpensive, small and easy to mount on the vehicle, and processing technique are also well developed. Laser is more typical on research and high-value vehicles, such as shuttles.

The Tesla Autopilot system camera configuration is shown in Figure 3; the arrangement is similar on all models. Note that there are three forward looking cameras with different field of views (FOV) to provide comparable resolution imagery over a long range in front of the vehicle.
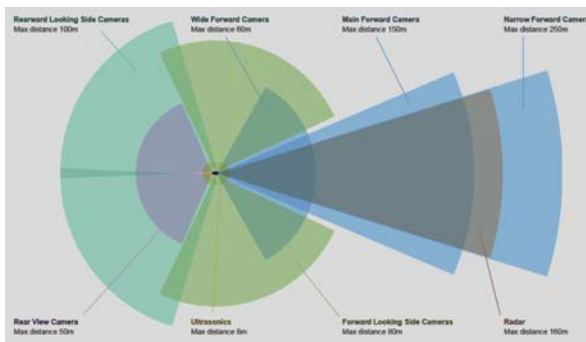

Figure 3. Tesla Autopilot system sensor FOVs; eight cameras, one radar, and 12 ultrasonic sensors (courtesy of Tesla)

The main rival the Cadillac CT6, the Super Cruise, uses only cameras and radar. There are eight cameras installed on the CT6 model, one inside is used for checking the driver's alertness level, and the others are sensing the environment around the vehicle. A unique feature of this system is the high-definition road map that covers 130 miles of freeways in North America and allows the vehicle to achive Level 2 autonomy (SAE, 2018). The map, independently acquired by LiDAR is stated to be accurate about 10 cm.

Waymo, owned by Google, uses laser sensors, a Velodyne mobile scanner, and their systems can be deployed on many stock vehicles. Since the laser sensor provides 360 FOV around the vehicle and the acquired data is 3D, there are less sensors on the vehicle to sense the environment. Figure 4 shows the general sensor arrangement, excluding GPS. A course map with features, such as traffic lights, is needed for the use of this AV technology. Also, there is option for driving on preprogrammed route.
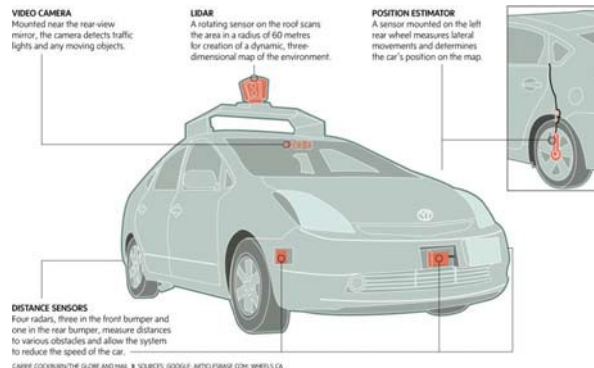

Figure 4. Waymo sensor configuration (courtesy of Google)

As the AV car industry continues moving forward from the current autonomy Levels 1 and 2, the amount and quality of the acquired image sensor data is expected to increase. Inexpensive laser sensors are intensely researched, and once became available will improve the potential for directly acquiring geospatial data that could be used to create high-definition maps, such as dense city models. The use of high-definition maps to improve the interpretation of the scene around the AV vehicle is clearly growing.

4. HIGH-DEFINITION MAPPING

The sensor systems developed for AV technology are not designed to acquire highly accurate spatial data. None of the cameras currently used meets the requirement of a metric sensor. However, the observations are highly redundant, as the same sensor will acquire data of an object or area multiple times, and then there are many sensors imaging the same object space. The research question is whether from the highly redundant and moderately accurate data it is feasible to obtain accurate spatial data. A slightly differently posed question is what the optimal sensor configuration is to support safe AV driving as well as provide for accurate mapping. Tests were carried out at the OSU main campus in Columbus, OH, in 2017, to collect data to analyze the performance of object space reconstruction based on using a variety of sensors installed on a test vehicle.

4.1 Platform

A GMC Suburban, customized measurement vehicle, called the GPSVan (Grejner-Brzezinska, 1996), is used as a platform for the data acquisition. The sensors installed on the platform included navigation and imaging/mapping sensors. A light frame structure installed on the top and front of the vehicle provided a rigid platform for the imaging sensors, including LiDARs and different types of cameras. The final sensor configuration consisted 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 (Velodyne family). The location of the sensors on the GPSVan is shown in Figure 55 and Figure 6.
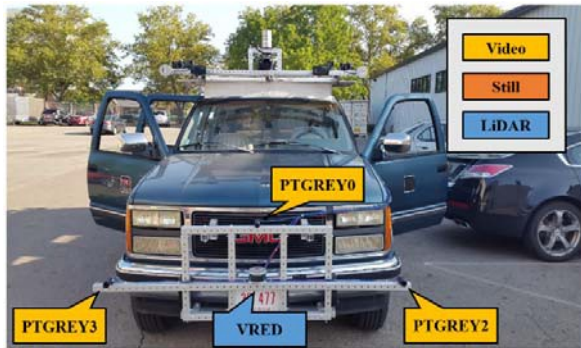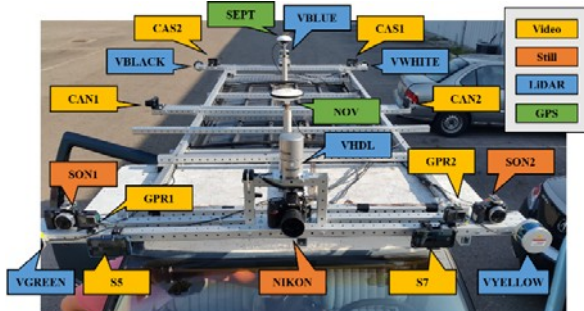
Figure 5. GPSVan front sensor installation

Figure 6. GPSVan top sensor installation

4.2 Test area

Two test sites were selected for the data acquisition, both located at the campus of The Ohio State University. The first route is at west campus and connects two research facilities, and has moderate vehicle and low pedestrian traffic. The second route is on main campus, 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. Due to the present of many moving objects, it clear represent the most challenging scenario for mapping. In addition, this area is a partially GPS/GNSS-denied due to tall buildings located along the route. This dataset contains 15 loops, acquired in about 4 hours, and represents a volume of about 5 TB raw data. A sample of the various imaging data streams is shown in Figure 7. The upper row shows imagery, acquired by three cameras of difference quality, including a low-end GoPro, a medium category Sony, and high-end Nikon. The middle raw shows two side looking cameras, and point cloud, acquired by the main laser scanner. The point cloud of a section of the main campus loop, acquired by the main laser scanner is shown in Figure 8a. Figure 8b shows the same area when all the point clouds of the seven laser scanner are combined.

The accuracy of the point clouds have been checked using building and road features (patches), and the accuracy at the references was 5 cm. The photogrammetrically derived point clouds produced varying and lower accuracy, which is the subject of continuing investigation; an example point cloud of the same area is shown in Figure 9.
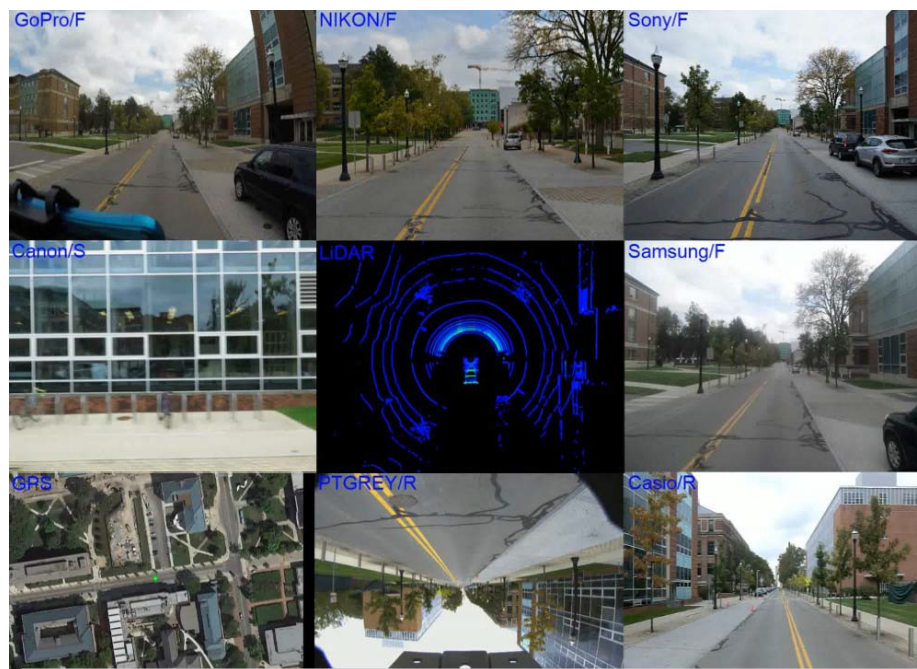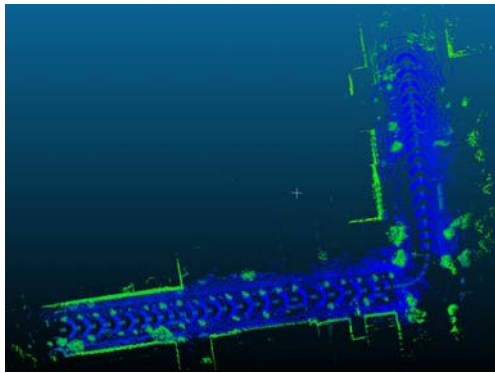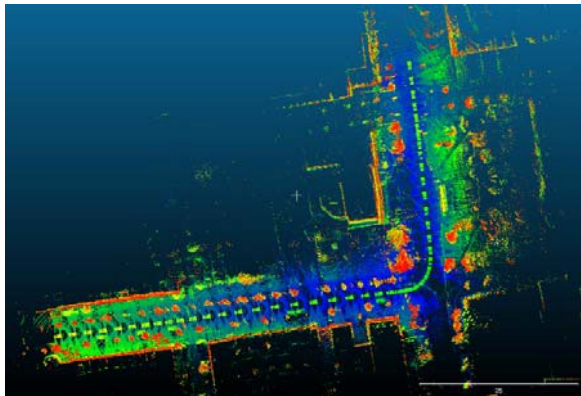

Figure 7. Data streams from GoPro (video, top left), Nikon (high resolution still images, top center), Sony (high resolution still images, top right), Canon (P&S, video, middle left), Velodyne HDL-32E (LiDAR, middle center), Samsung (mobile phone, built-in camera, middle right), GPS/GNSS (bottom left), PointGrey (video, bottom center), and Casio (P&S, video, bottom right)

(a)

(b)

Figure 8. HDL-32E point cloud (a), and all LiDAR sensors' data combined (b); points with 5 m above the road surface are removed, height is color-coded

Figure 9. Photogrammetrically created point cloud

5. POSITIONING WITH IMAGES: PERFORMANCE

Based on time and location, a basic database was built to provide an easy access to the large volume of data streams acquired in main campus data collection. Besides the accurate imaging sensor georeferencing, features are extracted and stored in the database. As an initial test of using the database for vehicle positioning using images acquired by a camera, a two-step method was evaluated; note that there are many methods available to accomplish this task. The concept implemented here is shown in Figure 10; note that both the database creation and its use for positioning are included.
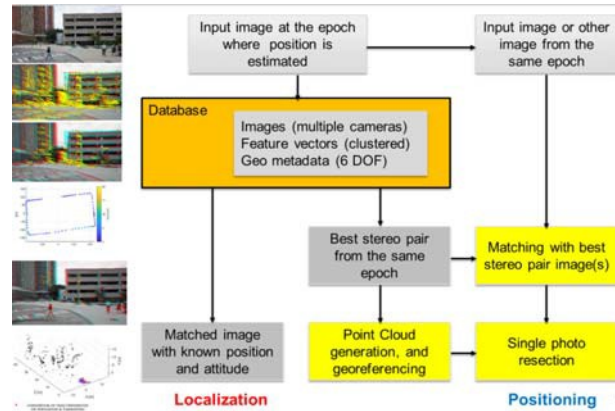

Figure 10. Two-step positioning based on georeferenced image database

In the first step, the vehicle is localized by searching for a close match of an image acquired from the vehicle. The matching is feature-based, using the SIFT feature descriptors. The search is generally accelerated by knowing the approximate location of the vehicle, so no need for an exhaustive global search in general. Different cameras were evaluated, and Figure 11 shows performance results, note that the lower score represents good performance; in other words, the match is unique. The accuracy of this localization is about 4 m, which is sufficient to start the refinement process.
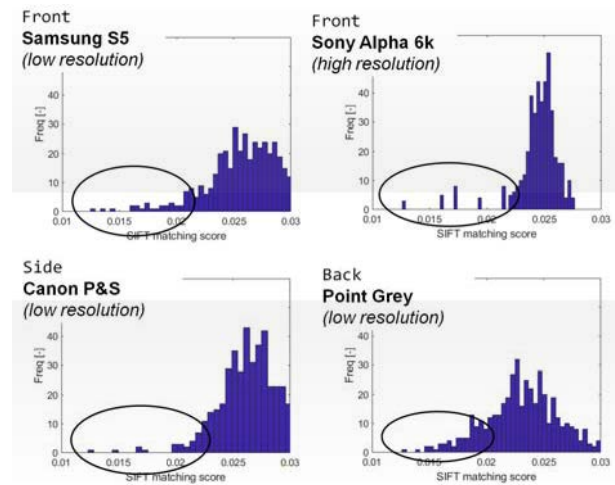

Figure 11. Matching results with various cameras

The second step is based on using 3D data from the database, and classical single photo resection is performed using the matched features. Figure 11 shows a point cloud used to refine the camera position and estimate attitude; green represents initial position, and purple is refined position. The accuracy of this processing depends on the point cloud accuracy, the spatial distribution of the points, and then on the camera quality, expressed in interior orientation characteristics. On average, better than 0.5 m 3D accuracy can be achieved in general; in benign areas with well-calibrated cameras, the accuracy could be below 0.2 m, which is close to the 0.1 m 2D accuracy suggested for AV positioning.

The procedure described here can be considered as a basic feasibility test. Using Big Data methods, a structured frame can be developed for the autonomous reconstruction of the 3D object space, including point cloud representation, feature points,

objects extracted, and even topology of the objects. Using this database, the image based position estimation can be also handled by Data Analytics methods, such as using Deep Learning, for example, CNN (Convolutional Neural Network), to identify objects and interpret scenarios besides retrieving position data (Rawat, 2017; Guo, 2017).
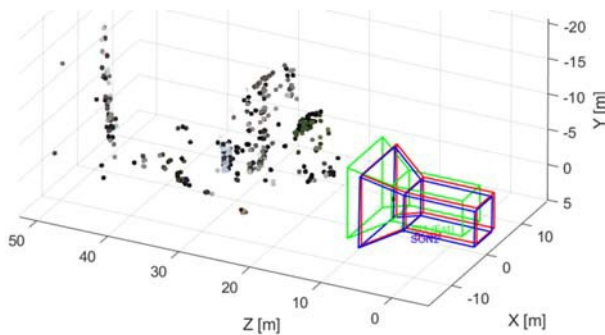

Figure 11. Camera position refinement based on single photo resection

6. SUMMARY AND CONCLUSION

AV technologies continue to rapidly advance, and the sensing capabilities of vehicles are expected to further improve. Inexpensive mobile sensors are still in the development phase, but once introduced, they will be used along cameras, which dominate the currently available AV market. The highly redundant image data, acquired by AV technologies have a big potential to create high-definition maps at good accuracy from crowdsensed data. The limitation of the current technology is that the huge amount of data cannot be easily transferred to the cloud. But, as connectivity improves, such as V2X becomes widely available, the conditions will quickly change.

Creating map data, which include point clouds, images, features, object, semantic information, etc., cannot be accomplished with the existing practice of map production. Frist of all the sheer amount of data represents an insurmountable obstacle. Then, the combination of high redundancy and low/modest sensor quality presents a formidable challenge. Clearly, Big Data technologies are needed to automatically reconstruct the object space around the transportation corridors. The main advantage of crowdsensed map data that it forms a live database, and it automatically adjust as the environment is changing. The crowdsensed map database will equally support AV and the geospatial needs of Smart Cities.

REFERENCES

Bonnefon, J.-F.; Shariff, A.; Iyad R. (2016) The Social Dilemma of Autonomous Vehicles. *Science* 352 (6293) pp. 1573–76. https://doi.org/10.1126/science.aaf2654.

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.

Guo, Y., Yu, L., Georigiou, T., Lew, M. S. (2017) A review of semantic segmentation using deep neural networks, *International Journal of Multimedia Information Retrieval*, pp 1-7. https://doi.org/10.1007/s13735-017-0141-z

Grejner-Brzezinska, D. (1996) Positioning Accuracy of the GPSVan. In Proceedings of the 52nd Annual Meeting of The Institute of Navigation. Cambridge, MA, USA.

Heipke, C. (2010). Crowdsourcing geospatial data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(6), 550-557.

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.

Piorkowski M., Sarafijanovoc-Djukic, N., Grossglauser, M. (2009) A Parsimonious Model of Mobile Partitioned Networks with Clustering, In The First International Conference on COMmunication Systems and NETworkS (COMSNETS), pp. 1-10, Jan. 5-10, 2009

Rawat, W., Wang, Z. (2017) Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review, *Neural Computation* (29), pp. 2352-2449

Su, K., Li, J., Fu, H. (2011) Smarty city and applications, International Conference on Electronics, Communications and Control (ICECC), pp. 1028-1031, 9-11 Sept. 2011

Society of Automotive Engineers (SAE) levels. Web link: https://web.archive.org/web/20170903105244/https://www.sae.org/misc/pdfs/automated_driving.pdf, last accessed: 04/2018

Toth, C., Jozkow, G. (2015) Remote Sensing Platforms and Sensors: A Survey, *ISPRS Journal of Photogrammetry & Remote Sensing*, 115 (2016), pp. 22-36.