

## VIDEO BASED MOBILE MAPPING SYSTEM USING SMARTPHONES

A. Al-Hamad \*, A. Moussa and N. El-Sheimy

Department of Geomatics Engineering, Schulich School of Engineering, University of Calgary, 2500 University Dr.  
NW , Calgary, Alberta, Canada T2N 1N4 - (amalhama, amelsaye and elsheimy)@ucalgary.ca

ISPRS Technical Commission I Symposium 2014

**KEY WORDS:** Close Range Photogrammetry, Mobile Mapping System, Smartphones, MEMS, Epipolar Geometry.

### ABSTRACT:

The last two decades have witnessed a huge growth in the demand for geo-spatial data. This demand has encouraged researchers around the world to develop new algorithms and design new mapping systems in order to obtain reliable sources for geo-spatial data. Mobile Mapping Systems (MMS) are one of the main sources for mapping and Geographic Information Systems (GIS) data. MMS integrate various remote sensing sensors, such as cameras and LiDAR, along with navigation sensors to provide the 3D coordinates of points of interest from moving platform (e.g. cars, air planes, etc.). Although MMS can provide accurate mapping solution for different GIS applications, the cost of these systems is not affordable for many users and only large scale companies and institutions can benefit from MMS systems.

The main objective of this paper is to propose a new low cost MMS with reasonable accuracy using the available sensors in smartphones and its video camera. Using the smartphone video camera, instead of capturing individual images, makes the system easier to be used by non-professional users since the system will automatically extract the highly overlapping frames out of the video without the user intervention. Results of the proposed system are presented which demonstrate the effect of the number of the used images in mapping solution. In addition, the accuracy of the mapping results obtained from capturing a video is compared to the same results obtained from using separate captured images instead of video.

### 1. INTRODUCTION

In the last few years, Micro Electrical Mechanical Systems (MEMS) sensors have witnessed a massive development in terms of the used technologies and manufacturing. The low cost of these sensors encourages various cellular phone manufacturers to use these sensors inside their phones to make it smarter for many applications. In 2012, Yole development has estimated that there are 497M units of smartphones with accelerometers and gyroscopes (Mounier & Développement, 2012). Nowadays, smartphones are becoming more sophisticated with a lot of capabilities and various sensors types. For example, current smartphones are equipped with GPS receivers, high resolution image and video cameras, MEMS inertial sensors and powerful computing processors. All of these developments in smartphones encourage the researches around the world to develop new creative applications and services beyond the traditional voice calls and SMS so that users can exploit its maximum benefits in their daily life activities.

In this paper, smartphones will be used as a platform for mapping applications. Using its GPS receiver, Inertial Measurement Unit (IMU), magnetometers and camera sensors, smartphones are considered ideal platforms which contains all navigation and remote sensing sensors required for any MMS. However, the main challenge of using smartphones for mapping applications is their poor inertial sensors accuracy which needs an external update source to improve its performance.

Video camera will be used in this research work to record a synchronized video with GPS, IMU and magnetometers

measurements inside the smartphone. Current smartphones digital video cameras can be used for various mapping applications. In contrast to a digital image camera, large overlapping area between images can be guaranteed between used images in mapping solution. The paper presents an approach to select a set of images from the captured video with a certain overlapping area between each two consecutive chosen images and use these selected images to estimate the final mapping solution of chosen interest points. The Exterior Orientation Parameters (EOPs) of the selected images will be calculated initially using the different navigation sensors measurements of the smartphone. In addition to a set of matched points between images, epipolar geometry constraints are used to correct the initial EOPs values. These corrected values are used in bundle adjustment software to estimate the final mapping solution.

The rest of this paper is organized as following: section 2 provides a brief literature review about the development of MMS. Sections 3 and 4 discuss the system implementation and methodology to obtain the mapping results using video cameras of smartphones. Results of the developed system are shown and discussed in section 5. Section 6 gives a summary of the paper.

### 2. LITERATUR REVIEW

In the last 20 years, MMS technology has witnessed a huge and rapid development in terms of cost and accuracy which are considered the main two used aspects to assess any mapping system (El-Sheimy, 1996). MMS are composed of two main

---

\* Corresponding author.

types of sensors; Navigation and Imaging (Mapping) sensors. Navigation sensors are the sensors which enable the user to determine the position and the orientation of the imaging sensor at exposure times. Inertial Measurement Units (IMUs) which contains two accelerometers and gyroscopes triads, GPS receiver and magnetometers triad are some examples of the navigation sensors. On other hand, mapping sensors can be passive ones such as video or digital cameras or active ones such as laser scanners (Ellum & El-Sheimy, 2002a).

GPSVan™ was the first operational land-based MMS, it was developed by the Centre of Mapping at the Ohio State (Ellum & El-Sheimy, 2002a). It integrated a code only GPS receiver, two digital CCD cameras, two colour video cameras and several dead reckoning sensors to obtain a relative accuracy of 10 cm and an absolute accuracy of 1-3 m. All of these components were mounted on the same van. Using VISAT system, a more absolute accuracy has been obtained using a dual frequency carrier phase differential GPS, eight cameras and precise IMU. The main objective for the VISAT project was to develop an accurate MMS for road and GIS data acquisition with an absolute position accuracy and relative accuracy of 0.3 m and 0.1 m respectively at a highway vehicle speed (e.g. 100km/hr) (Schwarz, et al., 1993). Figure 1 shows the VISAT system with all the components mounted on the roof of the used vehicle.



Figure 1. VISAT System

In literature, developed MMS have been used for various purposes. For road mapping applications, several examples can be found in literature such as (Artese, 2007), (Gontran, Skaloud, & Gilliéron, 2007) and (Ishikawa, Takiguchi, Amano, & Hashizume, IEEE 2006). A low cost backpack MMS has been developed at University of Calgary in (Ellum & El-Sheimy, 2002b) which can be used by pedestrians with a 0.2m and 0.3m absolute accuracies in the horizontal and vertical directions and 5 cm relative accuracy. Using smartphones as platform for MMS, only few examples can be found in literature. In (Fuse & Matsumoto, 2014), Smartphone's MEMS sensors and GPS receiver are used to self-localize the camera where measurements from these components are combined using Kalman filter in order to improve the accuracy of the calculated Exterior Orientation Parameters (EOPs) of the used camera at exposure times. These EOPs can be used to estimate the 3D coordinate of interest points from a moving vehicle.

The work in this paper is built on a previous published work in (Al-Hamad & El-Sheimy, 2014). In that research work, promising mapping results have been obtained using captured images from smartphones which shows the ability of using smartphones for different mapping applications. The work in

this paper investigates the efficiency of using captured video from smartphone for mapping applications.

### 3. SYSTEM IMPLEMENTATION

#### 3.1 The Used Device

In this research work, the Samsung Galaxy S4 smartphone has been used as an MMS platform. The resolution of Galaxy S4 video camera is 1920x1080 pixels with a maximum of 30 frames/second recording rate. A specially developed Android application has been used to record a synchronized video with GPS and sensors measurements. In (Al-Hamad & El-Sheimy, 2014) the size of the used images was 4128x3096 pixels which is approximately six times of the size of the images obtained from the recorded video in this work. This degradation of the image size will affect the final mapping solution accuracy as will be discussed in section 5. The types of the GPS receiver and motion sensors inside S4 device are listed in Table 1.

	Type
GNSS	Broadcom BCM47521
Accelerometers	STMicroelectronics LSM330DLC
Gyroscopes	STMicroelectronics LSM330DLC
Magnetometers	AsahiKasei AK8963

Table 1. GPS and motion sensors in Samsung Galaxy S4

#### 3.2 Images Extraction from Video

Using captured images for mapping applications can give more accurate results than using a recorded video especially for pedestrians. In addition to the higher resolution, more stable images can be obtained by capturing it immediately instead of obtaining it from a recorded video. However, recording a video for mapping applications can be easier to use for non-professional users. In addition, MMS using manual captured images can't be convenient in all scenarios. For example, it will be hard for a user to capture images manually from a moving vehicle while it can be done easily using a recorded video from a fixed camera on the roof or the side of the vehicle. In this research work, images obtained from a recorded video are used to estimate the 3D coordinates of the interest points. To extract images from the recorded video, each two consecutive extracted images should overlap with a certain percentage as shown in Figure 2. In this research work, 85% overlapping ratio has been chosen between each two consecutive images. Decreasing the overlapping ratio means obtaining less number of images and therefore less accurate mapping solution. On the other hand, increasing the used overlapping ratio will increase the number of used images for mapping solution and therefore increasing the required processing time.

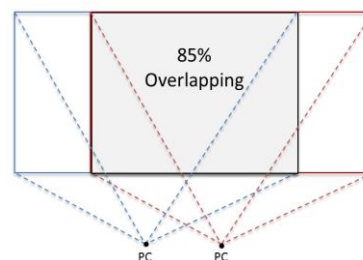


Figure 2. Extract images from a recorded video

### 3.3 Initial EOPs and IOPs Values

Using the GPS receiver and sensors inside the smartphone, EOPs can be initialized. Using equations (1), (2) and (3), the changes of the smartphone's positions between the different selected images in east, north and up directions can be calculated using the change of position derived from GPS.

$$\Delta N = \Delta\phi(R_{earth} + H_1) \quad (1)$$

$$\Delta E = \Delta\lambda(R_{earth} + H_1)\cos(\phi) \quad (2)$$

$$\Delta U = H_1 - H_2 \quad (3)$$

Where:

$\Delta N, \Delta E, \Delta U$  = the changes in the north, east and up directions.

$\phi, \lambda, H$  = the latitude, longitude and height GPS measurements.

$R_{earth}$  = the Earth radius at a given latitude.

On the other hand, smartphone's IMU and magnetometers' measurements are used to initialize the camera roll, pitch and azimuth rotation angles at exposure times of the selected images using equations (4), (5) and (6).

$$roll_{initial} = \sin^{-1}\left(\frac{a_y}{g}\right) \quad (4)$$

$$pitch_{initial} = \sin^{-1}\left(\frac{a_x}{g}\right) \quad (5)$$

$$azimuth_{initial} = \tan^{-1}\left(\frac{mag_y}{mag_x}\right) \quad (6)$$

Where:

$g$  = gravity acceleration value.

$a_x, a_y$  = accelerometers measurements in the x and y axes.

$mag_x, mag_y$  = leveled magnetometers measurements.

In this research work, zero principle point coordinates and distortion parameters and a fixed focal length are used as initial Interior Orientation Parameters (IOPs) values of the used camera. The calculated IOPs and EOPs initial values are used with other epipolar geometry constraints to calculate new refined IOPs and EOPs values which can be used to find a more accurate mapping solution using smartphones as will be shown in the following sections.

## 4. MAPPING USING VIDEO IMAGES

### 4.1 Direct Georeferencing

Georeferencing video images can be defined as the problem of transforming the 3-D coordinate vectors from the image frame to the mapping frame in which the results are required. The strength of MMS is their ability to directly georeference their mapping sensors which means estimating the position and the orientation of these mapping sensors (EOPs) at exposure times without depending on any control points. The relationship between the mapping and the smartphone coordinate frames is shown in Figure 3. The mapping solution of interest point in mapping frame  $r_p^M$  can be calculated using equation (7).

$$r_p^M = r_{SP}^M + \mu R_I^M r_p^I \quad (7)$$

Where  $r_{SP}^M$  is the smartphone position vector in the mapping frame (obtained from the smartphone's GPS receiver),  $r_p^I$  is the position vector of point p in the image frame (I),  $\mu$  and  $R_I^M$  are the scale factor and the rotation matrix between the mapping and the image coordinate systems.  $R_I^M$  can be obtained using the motion sensors inside the smartphone. More information about the georeferencing aspects of the developed system can be found in (Al-Hamad & El-Sheimy, 2014).

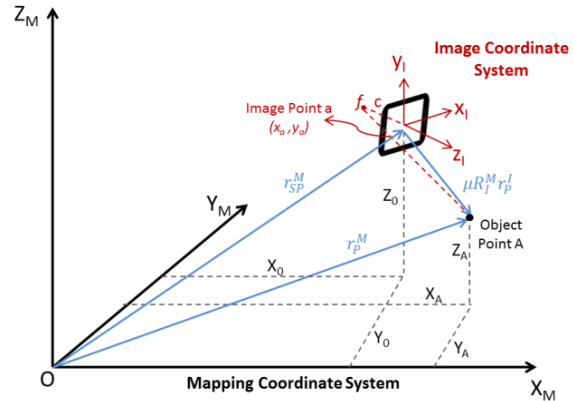


Figure 3. Mapping and Smartphone coordinate systems

### 4.2 EOPs and IOPs Correction

Using the obtained initial IOPs and EOPs from the previous step and a set of matched points between each two consecutive images, a new refined IOPs and EOPs values can be obtained by applying some epipolar geometry constraints. In epipolar geometry, without knowing the scale factor between the image and the mapping coordinate systems, the corresponding point of any point in the first image can be anywhere along the a line in the second image. This line called the epipolar line. The epipolar lines of all matched points intersect in the epipole point which is the image of the perspective centre of the first image on the second image. More information about the epipolar geometry can be found in (Hartley & Zisserman, 2003).

Due to the errors of the initial IOPs and EOPs values, the distances between the matched points and their corresponding epipolar lines will not be zero as shown in Figure 4. Using a set of matched points, and therefore a set of distances values, a non-linear least square estimator is used to calculate the best IOPs and EOPs values which minimize the distances between the matched points and their corresponding epipolar lines. A detailed explanation of IOPs/EOPs correction step can be found in (Al-Hamad & El-Sheimy, 2014).

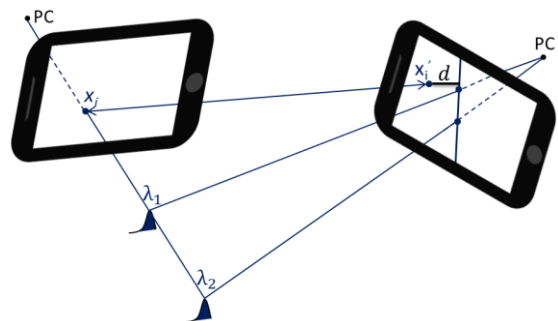


Figure 4. IOPs/ EOPs Correction

### 4.3 Mapping Solution

Using the new obtained IOPs and EOPs values from the previous step, bundle adjustment can be used to estimate the 3D coordinates of the object interest points. Bundle adjustment is a non-linear least square estimation where initial values for the unknown vector are very important to obtain a converged solution. The observation vector in bundle adjustment is the difference between the measured image matched points and the calculated ones using the extended collinearity equations shown in equations (8) and (9).

$$x_a + \delta x_r = x_p - c \frac{r_{11}(X_A - X_0) + r_{12}(Y_A - Y_0) + r_{13}(Z_A - Z_0)}{r_{31}(X_A - X_0) + r_{32}(Y_A - Y_0) + r_{33}(Z_A - Z_0)} \quad (8)$$

$$y_a + \delta y_r = y_p - c \frac{r_{21}(X_A - X_0) + r_{22}(Y_A - Y_0) + r_{23}(Z_A - Z_0)}{r_{31}(X_A - X_0) + r_{32}(Y_A - Y_0) + r_{33}(Z_A - Z_0)} \quad (9)$$

Where  $x_a, y_a$  are the calculated image coordinates of the object point.  $X_0, Y_0$  and  $Z_0$  are camera position at images exposure times.  $r_{ij}$  is the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column element of the rotation matrix ( $R_M^l$ ) between the mapping and image coordinate frames.  $C$  is the perspective distance of the used camera.  $\delta x_r$  and  $\delta y_r$  are the effect of the radial distortion in the x and y directions of the image and  $X_A, Y_A$  and  $Z_A$  are the object point in mapping (ground) coordinates solution.

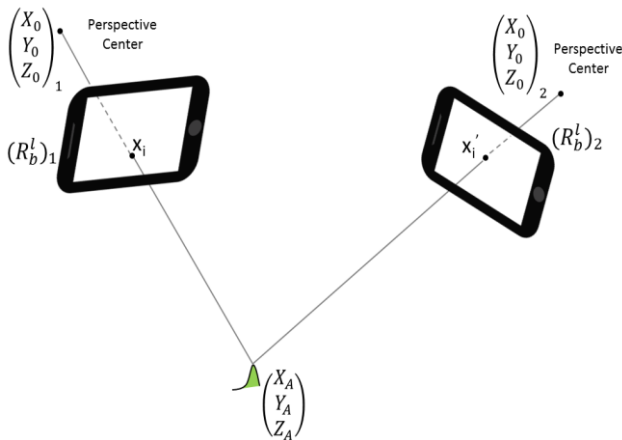


Figure 5. Mapping Solution

## 5. RESULTS AND ANALYSIS

To test the proposed method, a video has been recorded for a test field at University of Calgary with synchronized GPS and sensors measurements. Trying to keep an approximately fixed orientation of the camera will guarantee that a gradual change of the scene will appear in the recorded video. Using an 85% overlapping ratio between each two selected consecutive images, 10 images have been extracted from the video to find the mapping solution of the interest points. Image 6 shows an example of two selected images from the recorded video. The GPS positions of the selected images are shown in image 7. As can be noticed from Figure 7, the motion of the camera during recording the video was straight which will affect the final solution accuracy due to the poor geometry of the chosen images as will be shown later in this paper.



Figure 6. Two selected images of the recorded video

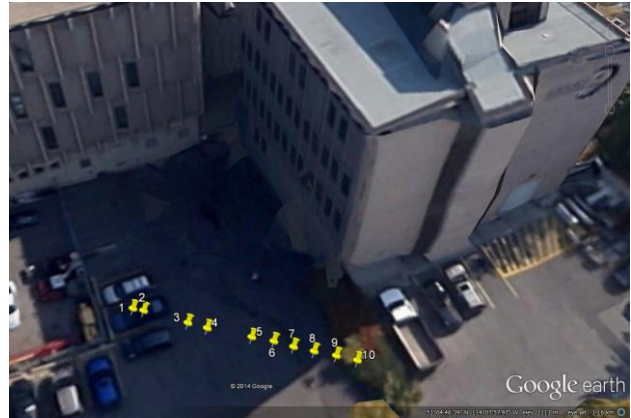


Figure 7. GPS positions of the chosen images

To test the mapping accuracies for different scenarios, two control points have been used to shift and rotate the final mapping solution model. Without using any control point, the final mapping solution will be shifted and rotated due to the position and azimuth errors of the first chosen image which have not been corrected and used as a reference to correct the IOPs and EOPs of the other chosen images.

To investigate the effect of the number of the used images in the final mapping solution, different minimum numbers of images have been used to estimate the 3D coordinates of the interest points. As can be noticed from Table 2, increasing the minimum number to test the mapping solution decrease the accuracy of the final mapping solution. Using 4 images only, the maximum 3D error was about 7 m. Increasing the minimum number of the used images to 5, the maximum 3D error decreased to approximately 2.5 m. On the other hand, the maximum mapping error of using a minimum of 6 images did not exceed 0.8 m. Mapping accuracies using at least six images in east, north and up directions are shown in Table 3.

Minimum Used Images	Number of Points	Mean (m)	Standard deviation(m)	Maximum Error (m)
4 images	24	1.340	2.322	7.316
5 images	21	0.514	0.680	2.586
6 images	18	0.265	0.228	0.778

Table 2. Mapping accuracies using different number of images

	Mean (m)	Standard Deviation(m)	Maximum Error (m)
East	0.224	0.237	0.775
North	0.064	0.069	0.233
Up	0.070	0.053	0.181
3D	0.265	0.228	0.778

Table 3. Mapping solution accuracies using 6 or more images in East, North and Up directions

The final mapping solutions with and without correcting IOPs and EOPs values (explained in section 4.2) are shown in Figures 8 and 9 using at least 6 images for each interest point in both cases. As can be noticed from Table 4, a huge improvement on the final solution has been obtained by applying the IOPs/EOPs correction step.

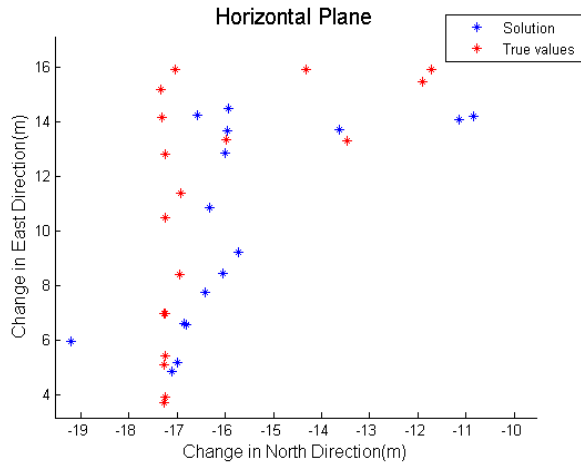


Figure 8. Mapping Solution without using IOPs/EOPs Correction step

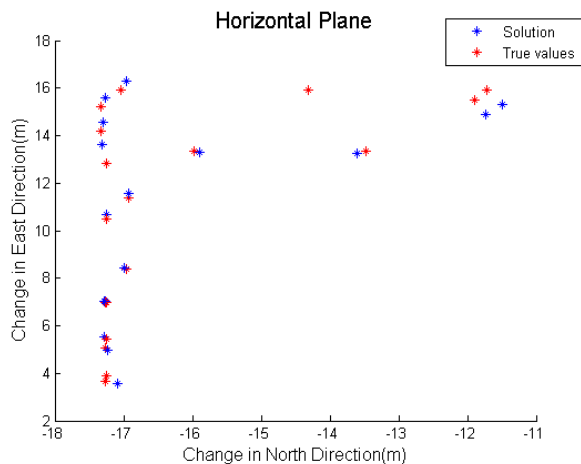


Figure 9. Mapping Solution using IOPs/EOPs Correction step

	Mean (m)	Standard deviation(m)	Maximum Error (m)
No EOPs/IOPs Corrections	3.350	2.220	7.392
With EOPs/IOPs Corrections	0.265	0.228	0.778

Table 4. Mapping solution accuracies with and without IOPs/EOPs correction step

To investigate the effect of using captured images or extracted images from a recorded video, results obtained from this paper have been compared to old results obtained from using captured images for mapping solution in Table 5. Mapping results using captured images can be found in (Al-Hamad & El-Sheimy, 2014). As can be noticed from the table, for pedestrians, capturing images for mapping can give more accurate results than using a video. In addition to the higher resolution and

stability of the captured images than a video, a better geometry of using captured images could be obtained since there is no need to guarantee any overlapping ratio between each two consecutive images. However, other images' extraction strategies from a video can be adopted that enable better geometry of the extracted images.

	Standard deviation(m)	Maximum Error (m)
Using Images	0.092	0.376
Using Videos	0.228	0.778

Table 5. Mapping accuracies using images and videos

## 6. CONCLUSION

In this paper, a video based MMS using smartphones has been introduced to overcome the problem of the high cost of the traditional MMS. Although more accurate results could be obtained using captured images for mapping applications by pedestrians, promising results has been obtained which shows the possibility of using videos for mapping to enable casual users to acquire overlapped images for proper mapping. These promising results show that smartphones will indeed be a major source for GIS data in the future.

## ACKNOWLEDGEMENTS

This work was supported in part by research funds from TECTERRA Commercialization and Research Centre, the Canada Research Chairs Program, and the Natural Science and Engineering Research Council of Canada (NSERC).

## REFERENCES

- Al-Hamad, A., & El-Sheimy, N. (2014). SMARTPHONES BASED MOBILE MAPPING SYSTEMS. (pp. 29-34). ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.
- Artese, G. (2007). ORTHOROAD: A low cost Mobile Mapping System for road mapping. In C. V. Tao, & J. Li (Eds.), *Advances in Mobile Mapping Technology - ISPRS Book Series* (pp. 31-41). London, UK: Taylor & Francis Group.
- Ellum, C., & El-Sheimy, N. (2002a). Land-Based Mobile Mapping Systems. *Photogrammetric Engineering and Remote Sensing*, 68(1), 13-28.
- Ellum, C., & El-Sheimy, N. (2002b). Portable Mobile Mapping Systems. *National Technical Meeting ION NTM*. San Diego, California: The US Institute of Navigation.
- El-Sheimy, N. (1996). A Mobile Multi-Sensor For GIS Applications In Urban Centers. *ISPRS, Commission II*. Vienna, Austria.
- Fuse, T., & Matsumoto, K. (2014). DEVELOPMENT OF A SELF-LOCALIZATION METHOD USING SENSORS ON MOBILE DEVICES. *ISPRS Technical Commission V Symposium* (pp. 237-242). Riva del Garda, Italy: ISPRS.

Gontran, H., Skaloud, J., & Gilliéron, P.-Y. (2007). A mobile mapping system for road data capture via a single camera. In C. V. Tao, & J. Li (Eds.), *Advances in Mobile Mapping Technology - ISPRS Book Series* (pp. 43-50). London: Taylor & Francis Group.

Hartley, R., & Zisserman, A. (2003). *Multiple View Geometry in Computer Vision*. Cambridge University Press.

Ishikawa, K., Takiguchi, J.-I., Amano, Y., & Hashizume, T. (IEEE 2006). A Mobile Mapping System for road data capture based on 3D road model. *Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, 2006* (pp. 638-643). IEEE.

Mounier, E., & Développement, Y. (2012). *MEMS Markets & Applications*. 2nd workshop on design, control and software implementation for distributed MEMS.

Schwarz, K., Martell, H., El-Sheimy, N., Li, R., Chapman, M., & Cosandier, D. (1993). VIASAT - A Mobile Highway Survey System. *VNIS-93*. Ottawa: IEEE - IEE Vehicle Navigation & Information Systems Conference.