ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240904220222

Preliminary communication
Received: 2024-09-04, Accepted: 2024-10-25

Selection of an Appropriate Extrinsic Camera Calibration Method for Handheld Mobile Mapping Systems

Luka Zalović*, Siniša Mastelić-Ivić, Ante Rončević

Abstract: Mobile mapping systems integrate multiple sensors to collect large volumes of geospatial data in motion. With the growing demand for mapping enclosed and hard-to-reach areas, there has been significant advancement in handheld mobile mapping systems utilizing SLAM (Simultaneous Localization and Mapping) technology. To ensure their data is efficiently usable, these systems should produce oriented images, which are essential for visualization, point cloud colorization, and monoplotting. Achieving so requires precise extrinsic camera calibration. This research provides a comprehensive overview of existing extrinsic camera calibration methods and evaluates their suitability for application in handheld mobile mapping systems. The goal is to identify a method that meets the accuracy and practical needs of these systems, facilitating more effective data processing and utilization in challenging environments.

Keywords: external camera; extrinsic camera calibration; handheld mobile mapping systems; mobile mapping; SLAM

1 INTRODUCTION

Mobile mapping systems (MMS) are measurement systems that integrate multiple sensors to enable the collection of a large amount of data on the move [1]. Usually, these sensors are IMU (Inertial Measurement Unit), LiDAR (Light Detection and Ranging), GNSS (Global Navigation Satellite System) receiver and one or more cameras for generating high-resolution images. Thanks to the integration of many sensors, as well as precise time synchronization and calibration, it is possible to collect georeferenced 2D imagery and 3D laser scans without the need for the system to be stationary. Such an approach is of utmost importance because it has enabled mobile mapping systems to be mounted on moving platforms such as vehicles or aircraft and used for extremely rapid mapping of cities and infrastructure in a non-invasive manner [2]. Despite the remarkable capabilities of such systems, they also have several significant drawbacks. Firstly, the traditional approach to integration entails the use of high-quality, and consequently expensive, measurement sensors, which often poses a barrier to successful implementation. Furthermore, mobile systems are inflexible regarding mounting methods. Ground-based mobile systems are often unable to be mounted on aircraft, just as aerial mobile systems cannot be mounted on ground-based moving platforms. Finally, both aerial and ground-based mobile systems are quite limited in terms of measuring narrow streets, forested areas, underground structures, and indoor spaces. The reason for this lies in the impossibility or impracticality of direct access to these objects. That is why in recent years, increasing attention has been paid to the development of so called PLS (Personal Laser Scanning) mobile mapping systems [3]. These are systems that, for mobility purposes, do not require standard mobile platforms but are light enough for a person to carry on their back or in their hands.

2 HANDHELD MOBILE MAPPING SYSTEMS

Handheld mobile mapping systems are a type of PLS used by operators who carry them by hand through the area being surveyed. For practical reasons, such systems must be as small, light, and mobile as possible, which require certain compromises in terms of sensor integration. Handheld systems typically integrate only a laser scanner and an inertial measurement unit (IMU). Since sensor quality used in such integrations is not at the level of traditional mobile systems, handheld mobile mapping systems utilize advanced and robust SLAM (Simultaneous Localization and Mapping) algorithms. These algorithms enable precise reconstruction of the scanner trajectory using data from the laser scanner and IMU sensor, regardless of their somewhat lower quality [4]. The addition of SLAM algorithms to mobile mapping systems represents a significant technological advancement in the field of mobile mapping, as it enables the realization of measurement systems that are economically viable, lightweight, highly mobile, and ultimately produce precise and reliable output data [5].

2.1 SLAM Algorithms

Simultaneous Localization and Mapping (SLAM) is a fundamental concept in robotics, wherein a robot incrementally constructs a map of an unknown environment while simultaneously determining its position within that map (Fig. 1). This problem is pivotal for autonomous robot navigation, especially when prior knowledge of the environment is unavailable. SLAM can be framed probabilistically, where the objective is to estimate the joint posterior probability distribution of the robot’s pose and the positions of landmarks in the environment, based on the robot’s sensor observations and control inputs [6]. The core mathematical formulation of SLAM involves recursive Bayesian estimation. This can be expressed with the motion model:

$$P(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1}, \boldsymbol{u}_k) \tag{1}$$

which describes the probability P of the robot's current state ($\boldsymbol{x}_k$) given its previous state ($\boldsymbol{x}_{k-1}$) and control input ($\boldsymbol{u}_k$), and the observation model:

$$P(\boldsymbol{z}_k \mid \boldsymbol{x}_k, \boldsymbol{m}) \tag{2}$$

which defines the likelihood of making an observation ($\boldsymbol{z}_k$) given the current state ($\boldsymbol{x}_k$) and map ($\boldsymbol{m}$) [6].

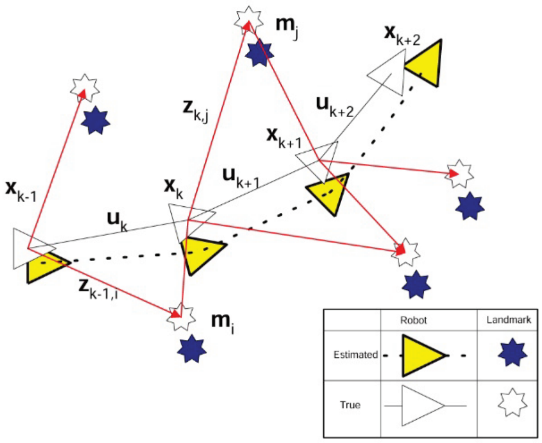

Figure 1 The essential SLAM problem. A simultaneous estimate of both robot and landmark locations is required while true locations are never known or measured directly. Observations are made between true robot and landmark locations.

The recursive nature of this problem allows for continuous updating of the map and pose estimates as new sensor data is acquired. However, the computational complexity of SLAM presents a significant challenge, particularly as the environment becomes more extensive and the number of landmarks increases. In its naive form, the SLAM algorithm scales quadratically with the number of landmarks, leading to prohibitive computational demands that are unsuitable for real-time applications [7]. To address these challenges, various strategies have been developed, such as state augmentation, sparsification, and submapping, which help reduce computational load while maintaining accuracy [8]. Another critical component of SLAM is data association, which involves correctly matching observed landmarks with those already recorded in the map. Errors in data association can lead to catastrophic mapping errors, particularly during loop closure, when the robot revisits previously mapped areas [9]. Loop closure is essential for correcting cumulative errors in the robot's pose estimate but requires accurate data association to avoid significant errors.

Recent advancements in SLAM research have focused on enhancing robustness and accuracy by incorporating richer environmental representations, moving beyond purely geometric models to include appearance-based features, which improve performance in diverse and complex environments. These developments have expanded SLAM's applicability to a wider range of scenarios, including those with dynamic elements and variable lighting conditions. Despite significant progress, challenges remain, particularly in scaling SLAM to large or highly unstructured environments. Ongoing research continues to address these issues, aiming to develop more efficient, scalable, and reliable solutions in the field of SLAM [10].

Today, SLAM is a crucial component of handheld mobile mapping systems, enabling accurate and real-time mapping and localization in diverse environments. SLAM technology has become increasingly vital in fields such as architecture, civil engineering, geodesy, forestry, and speleology, where precise and efficient spatial data collection is crucial.

2.2 Camera Integration

A component that is often missing in handheld mobile mapping systems is the integration of cameras for capturing RGB (Red, Green, Blue) images. Although some applications do not require their use, experience has shown that their availability significantly facilitates and accelerates the utilization of 3D data. Various manufacturers do not integrate cameras into handheld mobile systems because it would increase the mass and decrease mobility, thus reducing the practicality of the system. On the other hand, some opt to integrate small and low-quality imaging sensors, which often result in inadequate photographs. One of the most popular approaches to camera integration in such systems involves using an external camera that can be quickly mounted and removed from the system. All mentioned camera realizations for handheld mobile mapping systems are shown in Fig. 2. Approach with external camera is flexible and gives the user the option to use the camera only when needed. Also, external cameras often have higher-quality optics and image sensors compared to integrated cameras. However, the constant mounting and dismounting of the external camera will lead to changes in positional and angular offsets between the coordinate systems of the camera and the handheld mobile mapping system [11]. Moreover, even the regular usage of system will eventually affect the positional and angular offsets. Such changes can pose a significant problem for determining the exact position and orientation of the images, resulting in difficult manipulation and utilization of the collected data [12]. This issue is generally addressed through extrinsic camera calibration.

a) Integrated camera   b) External camera   c) No camera

Figure 2 Camera realizations in different handheld mobile mapping systems. Image a) shows system with integrated camera, b) shows system with external camera and c) shows system without camera

3 EXTRINSIC CAMERA CALIBRATION

To efficiently utilize data from handheld mobile mapping systems, it can be argued that 3D laser scan data should be accompanied by oriented imagery. Oriented images are those with known external orientation parameters, specifically 3D coordinates (X, Y, Z) and 3D rotations (angles φ, θ, and ψ) expressed in the mapping coordinate system (m-system) [12]. Such images are an extremely valuable visual tool for identifying and precisely defining the positions of certain objects, which would be difficult or impractical to detect using only point clouds. Furthermore, oriented images combined with laser scans allow for the definition of 2D image coordinates on the photograph, which, thanks to the known internal and external orientation parameters, can be directly projected onto the laser scans, thereby directly obtaining 3D coordinates in the m-system. This approach is known as monoplotting [13] and is highly useful when working with 3D data since it enables the direct definition of 3D points using images instead of laser scans, contributing to faster and more reliable creation of survey products. Finally, oriented images allow for the colorization, i.e., the assignment of RGB values to the collected 3D points, which is important in many applications for obtaining an accurate impression of the real appearance of the terrain and objects [14].

A crucial prerequisite for obtaining oriented images is the successful extrinsic calibration of the mobile system's sensors. Specifically, in the case of mobile mapping systems with fixed-mounted sensors, calibration involves determining external calibration parameters, i.e., positional and angular offsets between the camera coordinate system (c-system) and the mobile system's navigation unit coordinate system (b-system) [11], as shown in Fig. 3.

Extrinsic camera calibration methods for handheld mobile systems are still insufficiently explored. Most of the methods researched so far are suitable for vehicle-based and aerial mobile mapping systems. However, some of these methods, or their combination, could potentially be used for camera calibration purposes in handheld mobile systems. Camera calibration methods for mobile mapping systems can generally be divided into methods based on measurement targets (target-based), methods based on object characteristics (feature based), methods based on 3D alignment (3D alignment-based), methods based on sensor motion prediction (motion-based), and methods based on dependencies (dependence based).

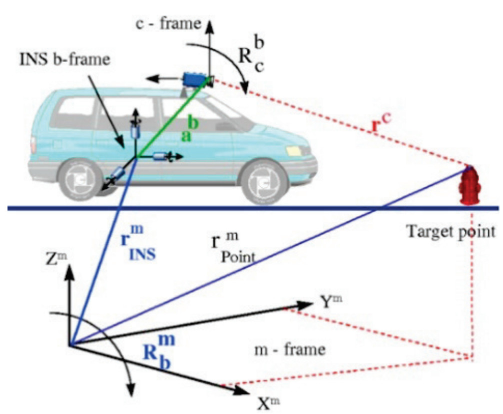

Figure 3 Coordinate systems of the mobile mapping system. External calibration parameters a^b and R_c^b allows for transformation between b-frame and c-frame

3.1 Target-Based Calibration Methods

Target-based methods are certainly the most well-known and thoroughly researched. These methods involve the use of artificial targets of known shape or pattern. The most commonly used targets are two-dimensional black-and-white patterns in the form of crosses, targets, or checkerboards (Fig. 4). Some authors have demonstrated successful implementation of three-dimensional targets such as trihedrals [15] and polygonal plates [16]. Regardless of the specific design of the target, the calibration method relies on finding identical points (in this case, the measurement targets) in the collected images and laser scans. Successfully identifying a sufficient number of identical points in both datasets allows for the mathematical adjustment process and the direct and unambiguous calculation of the external calibration parameters (positional and angular offsets). The use of targets with known shapes and patterns is crucial because images and scans represent two fundamentally different data formats. First and foremost, images are inherently two-dimensional, while laser scans are three-dimensional data. Furthermore, images do not contain information about the scale of the captured object, whereas laser scans do. Finally, the resolution of images is typically higher compared to the resolution of scans [17]. For these reasons, it is very challenging to define a method for the unambiguous and precise selection of identical points on both photographs and scans. Measurement targets with known shapes and patterns significantly simplify, and to a large extent, automate the process of selecting these identical points. The main drawbacks of such methods include the need to select, prepare, and set up the measurement targets, as well as the requirement to use software tools that will automatically (or semi-automatically) detect the targets and find identical points in the photographs and scans. It is important to note that the success of this method depends on the quality and resolution of the sensors used.

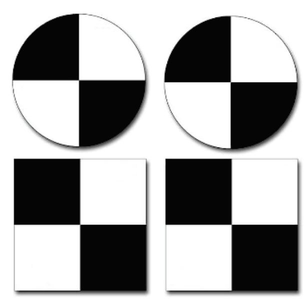

Figure 4 Targets with black-and-white patterns often used for calibration purposes. The difference in measured intensity on black and white segments allows for automatic detection of the target's centre.

3.2 Feature-Based Calibration Methods

Feature-based methods aim to calculate the external calibration parameters by detecting and matching common characteristic points, lines or shapes in 3D scans and images. Unlike target-based methods, the features used in these methods are not artificial targets. The available literature suggests that the most used features are road lines and building planes [18], although some authors have demonstrated successful calibration using car shapes [19], and even the skyline or horizon line [20], as shown in Figure 5. Regardless of the feature type, most of these methods are based on one or a combination of the following approaches [21]:

- Extracting 3D features from laser scans and photogrammetrically reconstructed 3D models
- Extracting 3D features from laser scans and 2D data from images
- Creating a 2D image from laser scans and extracting the same 2D features from both datasets.

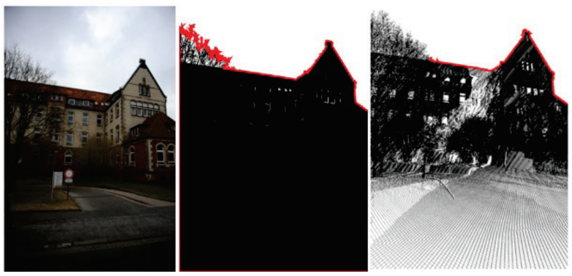

Figure 5 Skyline feature extraction from images and laser scan data. Left photograph is the original image, middle photograph represents threshold image (black pixels represent objects and white the sky) and the right photograph is the laser scan coloured by intensity.

Feature-based methods are well-suited for situations where there is a good distribution and visibility of features used for calibration, such as urban landscapes and built-up areas. On the other hand, these methods often yield poorer results in natural landscapes and areas with significant vegetation. Finally, the accuracy of the calibration itself is highly dependent on the quality of feature extraction in both datasets.

3.3 3D Alignment-Based Calibration Methods

3D alignment-based methods use the so-called "cloud-to-cloud" registration method to align two point clouds in order to determine the external camera calibration parameters [21]. In this process, the first point cloud is obtained from the laser sensor, while the second one needs to be calculated through photogrammetric 3D reconstruction based on the collected images. Point cloud registration involves the application of the well-known ICP (Iterative Closest Point) algorithm (Figure 6), which allows for the precise alignment of the photogrammetric point cloud with the 3D scan of the mobile mapping system, given there is enough overlap between the two datasets [22].

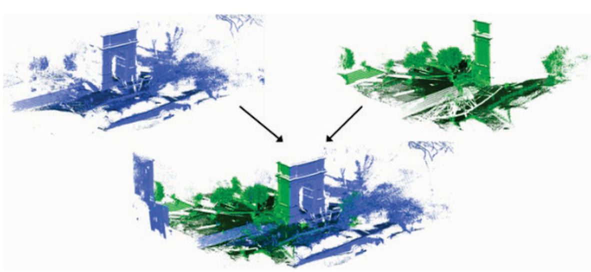

Figure 6 Registration of two point cloud by means of ICP algorithm. To be successful, the algorithm requires an overlap between two datasets.

3D alignment-based methods are accurate and provide reliable data in a wide range of situations; however, they are better suited to aerial mobile systems. This is due to the fact that the reliability of photogrammetric reconstruction heavily depends on the longitudinal and lateral overlaps between images, which are extremely difficult to achieve from the ground but straightforward from the air. Additionally, 3D alignment-based methods require initial rough alignment between the two point clouds and are extremely time-consuming due to the need for dense 3D point cloud reconstruction from the images.

3.4 Motion-Based Calibration Methods

Motion-based methods aim to calculate calibration parameters by utilizing knowledge of the movement of rigidly mounted sensors on a mobile platform [21]. Rigid sensor mounting means that the relative offset between sensors remains constant, regardless of the platform's movement. These methods are closely related to the hand-eye calibration problem (Fig. 7) addressed by the robotic community, where a camera ("eye") is rigidly mounted on a robot gripper ("hand"). The aim of the hand-eye calibration is to estimate the unknown transformation between the camera and the gripper coordinate frames based on the motions undergone by the gripper and camera, with the latter being estimated from captured images [23].

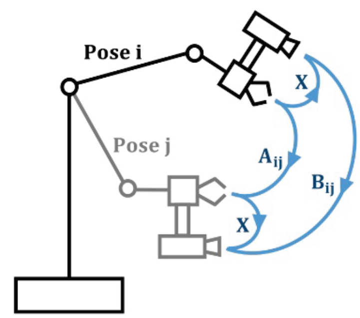

Figure 7 "Hand-eye" calibration problem visualisation. The aim is to find transformation *X* (camera calibration matrix) by using known transformations A_{ij} between different poses (*i*, *j*) and B_{ij} between observed image changes.

In earlier research, the main limitation of these methods was the need to use calibration markers to estimate the camera's motion [24, 25]. Recent studies, however, confirm the feasibility of using visual odometry and Structure from Motion (SfM) algorithms to overcome this limitation [26]. The most common approach in sensor motion prediction-based methods is to estimate the motion of the laser sensor using the ICP (Iterative Closest Point) algorithm, while camera motion prediction is achieved by tracking feature points in the photographs. The advantages of these methods include not requiring overlap between laser and image sensor data and not needing initial values for the calibration parameters. On the other hand, these methods require a significant amount of movement to accurately determine sensor motion, as well as precise temporal synchronization of both sensors.

3.5 Dependence-Based Calibration Methods

Dependence-based methods attempt to estimate the external camera calibration parameters based on the similarity between the signals of the laser scanner and optical images. Typically, this signal is expressed in a two-dimensional space, making it necessary to generate a 2D image from the laser sensor data by interpolating and projecting the laser data onto the image plane [27]. The fundamental assumption of dependence-based methods is that the two signals, one obtained from the laser sensor and the other from the image sensor, are somehow correlated (Figure 8). An example of such a signal is the reflectivity captured by the laser sensors and the grayscale values derived from RGB images. According to the current research and literature, these methods have been tested on sensors used for traditional terrestrial and aerial mobile systems, as well as on smaller 2D laser profilers and simple image sensors. In both cases, these methods have been proven to be extremely fast, reliable, and capable of automation. Unfortunately, they also have some significant drawbacks. Current studies confirm the need for an exceptionally narrow search space for accurate calibration parameters, which means they are limited to optimizing already precisely determined calibration parameters [28]. Finally, factors such as uneven lighting or shadows can significantly affect the correlation between the signals used, potentially resulting in inaccurate calibration parameters.

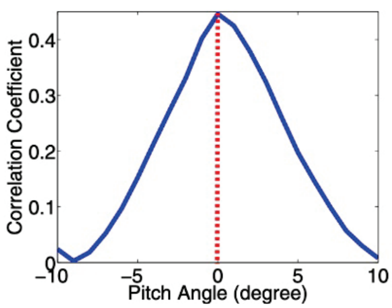

Figure 8 Reflectivity-greyscale Correlation Coefficient rising as it approaches correct calibration parameter (in this case, pitch calibration angle)

4 LIMITATIONS OF EXISTING CALIBRATION METHODS

Despite the advancements and success of the calibration methods discussed above, none of these approaches are fully suited to the specific requirements of handheld mobile mapping systems. The challenges posed by these systems, primarily due to their compact form, lower sensor quality, high degree of movement variability and need for frequent recalibration necessitate a re-evaluation of existing calibration techniques, which are primarily developed for aerial and vehicle-based mobile mapping systems. A brief overview of why each of the described methods faces limitations in the context of handheld mobile mapping systems is given below:

1) Target-based calibration methods require the use of artificial targets that must be carefully selected, prepared, and positioned in the environment. The process is labour-intensive and less practical for handheld systems, where quick deployment and flexibility are critical. Additionally, the lower resolution and quality of sensors in handheld systems can hinder the automatic detection of targets and the accurate identification of identical points across datasets, leading to unreliable calibration.
2) Feature-based calibration methods, while effective in urban environments with clearly defined features, struggle in natural landscapes or areas with sparse or indistinct features. Handheld systems, often used in diverse and uncontrolled environments, may not always have access to the consistent and prominent features required for accurate calibration. Moreover, the dependence on high-quality feature extraction, which can be challenging with the lower-grade sensors typical of handheld devices, further limits the reliability of these methods.
3) 3D alignment-based calibration methods rely heavily on accurate photogrammetric reconstruction and require substantial overlap between images, which is challenging to achieve with handheld systems due to their inherent instability and irregular movement. The need for initial rough alignment and the computationally intensive nature of dense 3D point cloud reconstruction make this approach less feasible for performing camera calibration frequently.
4) While motion-based calibration methods may not require data overlap between the laser and image sensors, they demand precise temporal synchronization and significant sensor movement to determine calibration parameters accurately. Handheld systems, which often experience erratic and inconsistent movement, may not provide the stable conditions needed for these methods to work effectively, rendering them impractical for this application.
5) Dependence-based calibration methods, while fast and potentially automatable, are highly sensitive to variations in lighting and environmental conditions. The correlation between laser and image signals can be disrupted by shadows or uneven lighting, leading to inaccurate calibration. Additionally, the narrow search space required for these methods limits their flexibility, making them less adaptable to the varied and dynamic conditions typical of handheld mobile systems.

All the calibration methods, together with their advantages and disadvantages regarding camera calibration of a handheld mobile mapping system, are summarized in Tab. 1.

Table 1 Existing extrinsic camera calibration methods

	Pros	Cons
Target-based methods	Accurate and computationally fast	Target detection due to sensor quality
Feature-based methods	Effective in urban landscapes and built-up areas	Feature extraction due to sensor quality
3D alignment-based methods	Automatic alignment of point clouds	Image overlap and computational intensity for frequent calibrations
Motion-based methods	No need for data overlap nor initial calibration parameters	Accurate temporal synchronisation and significant amount of motion
Dependence-based methods	Fast, reliable and automated	Narrow search space and influence of shadows on correlation

5 NOVAL CAMERA CALIBRATION APPROACH

Given the limitations of existing calibration methods, it is evident that there is a need for a new method specifically tailored to the unique demands of handheld mobile mapping systems. This new approach should address the challenges posed by lower sensor quality, irregular movement, and diverse operational environments. What is more, it should be straightforward for the end user to implement and conduct frequently, objective, and consider the specific characteristics of handheld mobile systems, which are:

- Unlike aerial or vehicle-based mobile mapping systems, handheld systems can be easily placed over a control point, by utilising so-called "reference plate", as shown in Fig. 9.
- Handheld systems do not require a moving platform or have specific operating limitations, allowing for complete control over the conditions in which the calibration process is carried out.
- Handheld systems often utilise lower resolution external cameras capable of collecting panoramic imagery at a specific time interval. Such images can be efficiently used for photogrammetric purposes by utilising spherical camera model and Structure from Motion algorithms.

Figure 9 Reference plate for handheld mobile mapping systems. The system is placed on a reference point in a way that the cross coincides with the point marker. The positional offset between the centre of the cross and b-system of mobile mapping system needs to be known.

Considering all the characteristics, the new calibration method should be a combination of existing calibration techniques, specifically target-based and motion-based methods. External calibration parameters could be determined through a so-called calibration scan, which would be a short scan (around 1 minute) conducted in a controlled environment suitable for both SLAM and SfM algorithms. Targets with black-and-white patterns should be utilized, as they can be automatically detected in the images. The coordinates of these targets in the mapping system (m-frame) can be extracted by placing the handheld mobile mapping system directly over them using a reference plate. Photogrammetric processing of the collected panoramic images should be performed using optimized SfM algorithms and a spherical camera model, with the target coordinates added as ground control points (GCPs). This process would produce two datasets: images with an initial pose that do not contain positional and angular offsets, and images with an optimized pose that includes these offsets. Finally, the external calibration parameters would be estimated using a least squares adjustment. The described noval approach will be examined in more detail in future research.

6 CONCLUSION

This paper has presented a comprehensive review of extrinsic camera calibration methods, emphasizing their suitability for handheld mobile mapping systems. Handheld systems, while offering unique advantages over aerial and vehicle-based platforms, present distinct challenges, especially when it comes to camera integration and calibration. Due to specific characteristics of handheld mobile mapping systems, such as high versatility and mobility, lower mapping and navigation sensor quality and need for frequent camera calibration, traditional extrinsic camera calibration methods are not suitable for them. The calibration process for handheld systems must therefore be designed to be straightforward and quick, allowing users to perform the calibration independently. It is essential that the process remains objective, minimizing user errors while ensuring accuracy. Finally, the calibration method should take advantage of unique capabilities of handheld systems, such as the possibility do be placed over control point, their independence of vehicle-based moving platform and interval-based panoramic image collection which can be efficiently used for photogrammetric purposes. By focusing on the specific requirements and constraints of handheld systems, a novel calibration method could enhance the accuracy and reliability of these systems, enabling more effective and versatile mapping solutions

7 REFERENCES

[1] Løvås, M. (2017). Increasing the Accuracy of Positioning in Mobile Mapping Systems, a Method Supported by Simultaneous Localization and Mapping. *Master thesis*, NTNU, July 2017.

[2] El-Sheimy, N. (2005). An Overview of Mobile Mapping Systems. *FIG Working Week 2005 and GSDI-8*, Cairo, Egypt April 16-21, 2005.

[3] Lauterbach, H., Borrmann, D., Heß, R., Eck, D., Schilling, K. & Nüchter, A. (2015). Evaluation of a backpack-mounted 3D mobile scanning system. *Remote sensing, 7*(10), 13753-13781. https://doi.org/10.3390/rs71013753

[4] Bosse, M., Zlot, R. & Flick, P. (2012). Zebedee: Design of a Spring-Mounted 3-D Range Sensor with Application to Mobile Mapping. *IEEE Transactions on Robotics, 28*(5). https://doi.org/10.1109/TRO.2012.2200990

[5] Durrant-Whyte, H. & Bailey, T. (2006). *Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms*. Australian Centre for Field Robotics (ACFR) J04, The University of Sydney, Sydney NSW 2006, Australia.

[6] Smith, R. & Cheesman, P. (1987). On the representation of spatial uncertainty. *Int. J. Robotics Research, 5*(4), 56-68. https://doi.org/10.1177/027836498600500404

[7] Dissanayake, G., Newman, P., Durrant-Whyte, H. F., Clark, S. & Csobra, M. (2001). A solution to the simultaneous localisation and mapping (SLAM) problem. *IEEE Trans. Robotics and Automation, 17*(3), 229-241. https://doi.org/10.1109/70.938381

[8] Montemerlo, M., Thrun, S., Koller, D. & Wegbreit, B. (2002). Fast SLAM: A factored solution to the simultaneous localization and mapping problem. *AAAI National Conference on Artificial Intelligence*, 593-598.

[9] Neira, J. & Tardos, J.D. (2001). Data association in stochastic mapping using the joint compatibility test. *IEEE Transactions on Robotics and Automation, 17*(6), 890-897. https://doi.org/10.1109/70.976019

[10] Thrun, S., Liu, Y., Koller, D., Ng, A. & Durrant-Whyte, H. (2004). Simultaneous localisation and mapping with sparse extended information filters. *Int. J. Robotics Research, 23*(78), 693-716. https://doi.org/10.1177/0278364904045479

[11] Gopaul, N.-S., Wang, J. & Hu, B. (2016). Camera auto-calibration in GPS/INS/stereo camera integrated kinematic positioning and navigation system. *The Journal of Global Positioning Systems, 14*(3), 1-15. https://doi.org/10.1186/s41445-016-0003-7

[12] Rau, J.-Y., Habib, A.-F., Kersting, A.-P., Chiang, K.-W., Bang, K.-I., Tseng, Y.-H. & Li, Y.-H. (2011). Direct Sensor Orientation of a Land Based Mobile Mapping System. *Sensors, 11*(7), 7243-7261. https://doi.org/10.3390/s110707243

[13] Santos, R. D., Dal Poz, A. P. & Dalmolin Q. (2010). Indirect orientation of images using control points extracted by the means of monoplotting model. *The Photogrammetric Journal of Finland, 22*(1).

[14] Vechersky, P., Cox, M., Borges, P. & Lowe, T. (2018). Colourising Point Clouds Using Independent Cameras. *IEEE Robotics and Automation Letters, 3*(4). https://doi.org/10.1109/LRA.2018.2854290

[15] Gong, X., Lin, Y. & Liu, J. (2013). 3D LIDAR-Camera Extrinsic Calibration Using an Arbitrary Trihedron. *Sensors, 13*(2), 1902-1918. https://doi.org/10.3390/s130201902

[16] Park, Y., Yun, S., Won, C.-S., Cho, K., Um, K. & Sim, S. (2014). Calibration between Color Camera and 3D LIDAR Instruments with a Polygonal Planar Board. *Sensors, 14*, 5333-5353. https://doi.org/10.3390/s140305333

[17] Mishra, R.-K. & Zhang, Y. (2012). A Review of Optical Imagery and Airborne LiDAR Data Registration Methods. *The Open Remote Sensing Journal, 5*, 54-63. https://doi.org/10.2174/1875413901205010054

[18] Cui, T., Ji, S., Shan, J., Gong, J. & Liu, K. (2017). Line-Based Registration of Panoramic Images and LiDAR Point Clouds for Mobile Mapping. *Sensors, 17*, 70, 1-20. https://doi.org/10.3390/s17010070

[19] Li, J., Yang, B., Chen, C., Huang, R., Dong, Z. & Xiao, W. (2017). Automatic registration of panoramic image sequence and mobile laser scanning data using semantic features. *ISPRS Journal of Photogrammetry and Remote Sensing, 136*, 41-57. https://doi.org/10.1016/j.isprsjprs.2017.12.005

[20] Hofmann, S., Eggert, D. & Brenner, C. (2014). Skyline matching based camera orientation from images and mobile mapping point clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, II-5*, 181-188. https://doi.org/10.5194/isprsannals-II-5-181-2014

[21] Pentek, Q., Kennel, P., Allouis, T., Fiorio, C. & Strauss, O. (2020). A Flexible Targetless LiDAR–GNSS/INS–Camera Calibration Method for UAV Platforms. *ISPRS Journal of Photogrammetry and Remote Sensing, 166*, 294-307. https://doi.org/10.1016/j.isprsjprs.2020.05.014

[22] Abayowa, B.-O., Yilmaz, A. & Hardie, R.-C. (2015): Automatic registration of optical aerial imagery to a LiDAR point cloud for generation of city models. *ISPRS Journal of Photogrammetry and Remote Sensing, 106*, 68-81. https://doi.org/10.1016/j.isprsjprs.2015.05.006

[23] Tsai, R. Y. & Lenz, R. K. (1989). A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. *IEEE Transactions on Robotics and Automation, 5*(3), 345-358. https://doi.org/10.1109/70.34770

[24] Shiu, J.-C. & Ahmad, S. (1989). Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form *AX=XB*. *IEEE Transactions on Robotics and Automation, 5*(1), 16-29. https://doi.org/10.1109/70.88014

[25] Wang, C. C. (1992). Extrinsic calibration of a vision sensor mounted on a robot. *IEEE Transactions on Robotics and Automation, 8*(2), 161-175. https://doi.org/10.1109/70.134271

[26] Schmidt, J., Vogt, F. & Niemann, H. (2005). Calibration–free hand–eye calibration: a structure–from–motion approach. *Joint Pattern Recognition Symposium*, Springer, 67-74. https://doi.org/10.1007/11550518_9

[27] Wang, R., Ferrie, F.-P. & Macfarlane, J. (2012). Automatic Registration of Mobile LiDAR and Spherical Panoramas. *Computer Vision and Pattern Recognition Workshops (CVPRW2012), IEEE Computer Society Conference*. https://doi.org/10.1109/CVPRW.2012.6238912

[28] Pandey, G., McBride, J.-R., Savarese, S. & Eustice, R. M. (2015). Automatic Extrinsic Calibration of Vision and Lidar by Maximizing Mutual Information. *Journal of Field Robotics, 32*(5), 696-722. https://doi.org/10.1002/rob.21542

Authors' contacts:

Luka Zalović, mag. ing. geod. et geoinf.
(Corresponding author)
University of Zagreb, Faculty of Geodesy,
Fra Andrije Kačića Miošića 26, 10000 Zagreb, Croatia
lzalovic@geof.hr

Siniša Mastelić-Ivić, prof. dr. sc.
University of Zagreb, Faculty of Geodesy,
Fra Andrije Kačića Miošića 26, 10000 Zagreb, Croatia
sinisa.mastelic.ivic@geof.unizg.hr

Ante Rončević, prof. dr. sc.
University North, Business Economics,
Jurja Križanića 31b, 42000 Vataždin, Croatia
ante.roncevic@unin.hr