2017/2500M-0063

ISSN 1797-3457 (verkkojulkaisu)
ISBN (PDF) 978-951-663-005-5

SUMMARY REPORT

Infrastructure-free tactical situational awareness (INTACT)

Laura Ruotsalainen (laura.ruotsalainen@nls.fi, 050 433 6552),
Martti Kirkko-Jaakkola, Jesperi Rantanen, Maija Mäkelä
Department of Navigation and positioning, Finnish Geospatial Research Institute FGI

Abstract: The objective of the INTACT project is to analyze and develop methods for infrastructure–free simultaneous localization and mapping (SLAM) and context recognition for tactical situational awareness. Most important research questions are how an accurate and reliable SLAM system may be obtained using a single camera, multiple inertial sensors and ranging equipment, and how good situational awareness the equipment provides. All measurements will be collected using only equipment attached to the user. Indoor environments are selected as a specific research environment, because localization is most challenging in those areas, but all results are well suited also for urban and for some extent for all outdoor environments. The project addresses applications aimed for soldiers, but the methods developed will serve the needs of e.g. police, border guards and rescue personnel as well.

1. Introduction

Tactical situational awareness for military applications should be based on infrastructure-free systems and should be able to form knowledge of the previously unknown environment. Localization of the soldier and formation of a map from the unknown environment are core of forming situational awareness. Also, information of the soldier's motion context is important for successful operations, e.g. if the soldier is running, crawling or static for a long time. The motion information will give the command center invaluable information about the status at the field, the format of the information representation to the soldier may be changed to be less straining base on the situational awareness drawn from the motion context, and the fusion algorithm may be developed to provide more accurate and reliable positioning information when motion context is integrated into it.

The infrastructure-free requirement is motivated by the fact that rescue and military personnel must be able to operate reliably in any environment, regardless of the available infrastructure. Requirements for the system are stringent; it should function also in indoor environments, which is at present the most challenging operation environment for localization methods, it should be lightweight and inexpensive.

Simultaneous Localization and Mapping (SLAM) is a key technology for providing an accurate and reliable infrastructure-free solution for indoor situational awareness (Davison et al. 2007). However, indoor environments and the requirements for the system make the implementation of SLAM using existing algorithms challenging. Most existing algorithms were developed for use in robotics where size and weight requirements are not as stringent. Due to size limitations, we have implement SLAM using methods providing an accurate solution using a monocular camera (Ruotsalainen 2013), a result that has been unavailable before.

Methods developed in INTACT will provide means for infrastructure-free tactical situational awareness. The methods consist of Particle filtering for multi-sensor fusion, a sophisticated SLAM algorithm and motion recognition means. Our approach was to integrate a monocular camera, multiple Self-contained Micro-Electro-Mechanical (MEMS) grade Inertial Measurement Units (IMUs), a barometer and a ranging sensor to obtain a solution for SLAM, as well as tactical motion information. Also, a method using sonar for an improved map formation was developed. This report discusses the research done during the third research year and presents results from two proof-of-concept test campaigns.

2. Research objectives and accomplishment plan

Most important research questions were how an accurate and reliable SLAM solution may be obtained using a single camera, multiple inertial sensors and ranging equipment, and how good situational awareness the equipment provides. During the first two years the research concentrated on the development of fusion algorithm, machine learning algorithms for improved context recognition as well as a SLAM algorithm based on (Civera et al. 2010) integrating vision-aided algorithms developed earlier at FGI (Ruotsalainen 2013) for improved performance of the SLAM solution.

During the third research year the methods were evaluated via two proof-of-concept test campaigns committed in Utti military area by soldiers. The results of the first campaign, done in February, were analyzed and the methods further developed accordingly. The second test campaign was committed in September and its good results give encouraging proof of the usability of the methods for infrastructure-free tactical situational awareness. The project formalized also recommendations for the future system implementation.

3. Materials and methods

This section discusses the materials used in the research and the methods developed in INTACT.

3.1 MULTI-SENSOR FUSION

A Particle filtering algorithm has been developed in INTACT in order to fuse all sensor measurements for an accurate and reliable solution. The multi-sensor fusion results are discussed in the Results and discussion section.

3.1.1 Horizontal localization

The horizontal localization relies on inertial and visual measurements. One of the inertial measurement units is mounted onto the shoe of the user. Although this is somewhat difficult to implement from the instrumentation point of view, the indispensable benefit is that the shoe is known to be periodically at rest during human gait, and these stance phases can be detected from the sensor output. This makes it possible to evaluate the traditional strapdown inertial navigation equations to estimate the movement of the foot at centimeter-level precision: the rapid error accumulation characteristic to low-cost inertial navigation systems can be mitigated by resetting the velocity to zero whenever the foot is detected to be stationary. Foot-mounted inertial navigation does not require user-specific calibration, and it inherently copes with sidestepping and climbing.

The concept of visual gyroscope and visual odometer (Ruotsalainen 2013), providing the user displacement and direction information, were used as another sources of horizontal localization information. In order to fuse the measurements for the best final result, the errors arising from measuring were modelled statistically. Then, inertial and visual measurements were fused using a Particle filter (Liu, 2001) incorporating the correct error models.

3.1.2 Vertical localization

Determining the altitude based on inertial measurements only is prone to drift due to measurement errors and inaccurate knowledge of the local gravitational acceleration. Therefore, we employ a barometer providing the user vertical position component based on the change in the air pressure. However, barometer height measurements suffer from error arising from changes in the air temperature and pressure. Therefore, a sonar pointing downwards and detecting if the change in the height computed by the barometer is really due to the change in height and not the environmental changes is used.

A Kalman filter was developed to fuse the barometer and sonar measurements before entering the height into the Particle filter. The Kalman filter state model consists of true height, height bias (caused by environmental effects) and vertical speed estimated by sonar. Figure 1 compares the fused result of barometric height and sonar observations to the biased height. All measurements were averaged over one second. The bias is the difference between filtered and barometric height and it starts to increase when entering a building, likely because of different ventilation conditions. The method performs somewhat poorer than simple barometric height in stairs (400-480 seconds from start) but better overall.

Largest bias in barometric height is caused by moving between indoor and outdoor environments. The Kalman filter approach able the system to adapt to different environments. The magnitude of bias can be used to estimate the change from one environment to other.

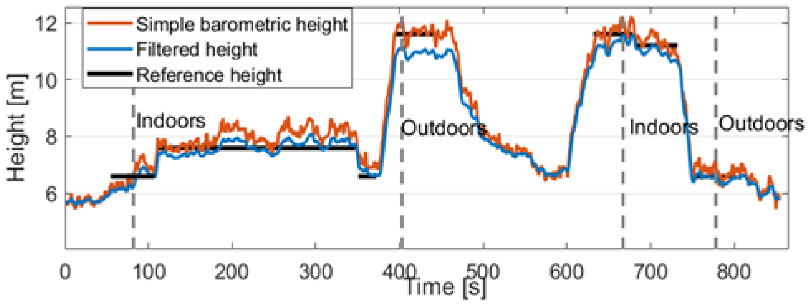

Figure 1: Kalman filtered barometric height and sonar measurements compared to simple barometric height. Grey dashed line mark transition between outdoor and indoor environments.

3.2 SITUATIONAL AWARENESS

The knowledge of the soldier's motion context is important for the localization algorithm, as well as for the tactical purposes. The main goal in the context recognition research in INTACT has been to develop a method that can detect the user's motion independently of the user's personal motion characteristics, and to integrate the obtained information into the fusion algorithm. Therefore we have experimented with various machine learning algorithms and used data from separate persons in order to find best methods for these purposes. The results are shown in Section 4.

Building upon the work done in the previous years of INTACT, during the third year we have studied the different aspects of machine learning. We have studied feature selection from more algorithmic point of view, and also experimented with different pattern recognition methods to see which option would best suit this application.

Previously 14 different motion categories (such as standing, walking, running, crouching, crawling, climbing, turning etc.) were identified. However, some of these motions proved to be difficult to distinguish or were not mutually exclusive. Therefore some of the categories were merged into one category, and some were left out. For instance, categories such as "rising from crouching to standing" or "getting down to crouching" were excluded from the algorithm development, since the duration of the transition from one motion to another is usually very short and can be deduced based on the change in the recognized motion. Thus the movement categories to be detected are walking, running, lying still, standing, climbing up, climbing down and moving forward in a low posture.

In previous work means, variances, dominant frequencies and their amplitudes of the sensor readings were found to be useful features in context recognition. During the third year we used these same features, but studied the feature selection further by algorithmically looking for the best subset of these features. The reduction of the number of features is beneficial in terms of computational costs, but it also helps in avoiding overfitting the chosen classifier.

3.3 MAPPING WITH A SONAR

In addition to the knowledge of soldier's location and motion context, information about the surroundings is important for tactical purposes. For this purpose, we have experimented de-formation of a building plan using sonar ranging device. The map is produced on the move, and develops as the person moves around in the building. Figure 2 displays a suggestion for a representation of an indoor map. The sonar ranging device that was used measures the distance to the nearest object in the direction it is pointing to. In this experiment, the sensor is assumed to point perpendicularly right of the direction of the movement. The green lines with black dots in the end represent the range measurements, and green lines with no dots are measurements larger than the maximum possible measurement, 5 meters. The illustra-tion presented here is from a simplistic experiment.

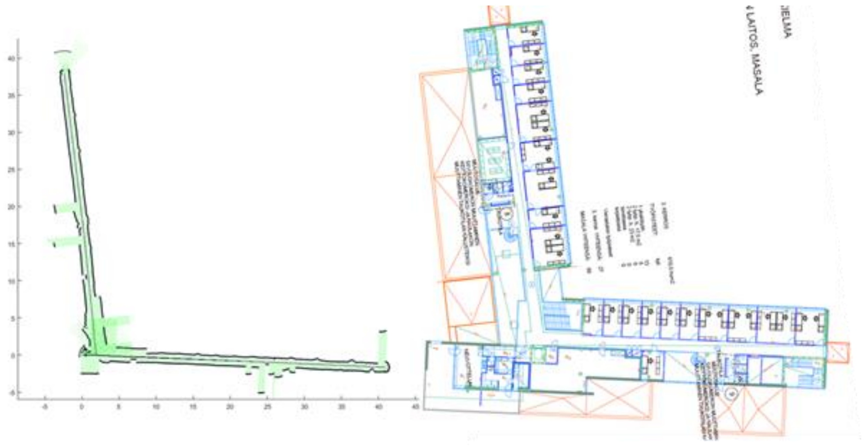

Figure 2: Example of mapping an indoor corridor with sonar ranging device.

4. Results and discussion

During the third research year the methods developed in the project were tested in two proof-of-concept test campaigns in Utti. The first test campaign was carried out in February and best on the results the methods were further improved. The second test campaign was carried out in September and its results are shown below.

4.1 PROOF-OF-CONCEPT

Test campaigns were carried out in Utti military area's exercise hall. The building had confined hallways and small rooms and one larger space approximately 20 by 40 meters in size (Figure 3). Navigation tests consisted of traveling a route that started outdoors, went through all the rooms in the building and around the hall and included some climbing on steps or ladders. Test persons were two soldiers wearing combat equipment to simulate a realistic use situation. The sensors were attached to the combat west, helmet and footwear of the test person. Locations of the sensors are shown in Figure 4.

The test persons walked the same route first at a slow pace and with no sudden movements and the second time faster with some running sprints and more natural movements.

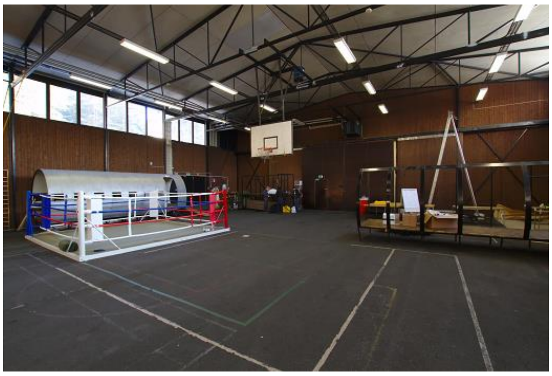

Figure 3: Confined hallway at the test site (on left) and interior of the exercise hall at the test site (on right).

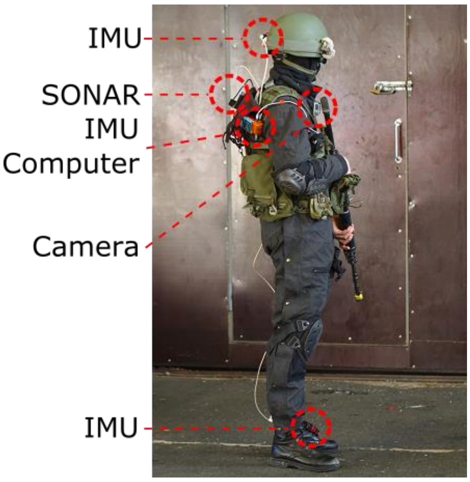

Figure 4: Sensor locations.

4.1.1 Localization, UTTI 2

Herein, horizontal localization result for the second UTTI test campaigns, and therefore the final result of the project, is shown in Figure 5 (left). Figure 5 shows also a rough true path of the test, drawn manually (on right). Due to the lack of a reference solution, the accuracy of the result is evaluated by measuring the error in the loop-closure, i.e. the deviation of the solution end point from the correct end point. The error was 2.5 m after the first round and 4.6 m after the second round committed by running. Also, the error in the second round was caused mainly from the error in the starting heading that could be corrected by a mean providing absolute heading information.

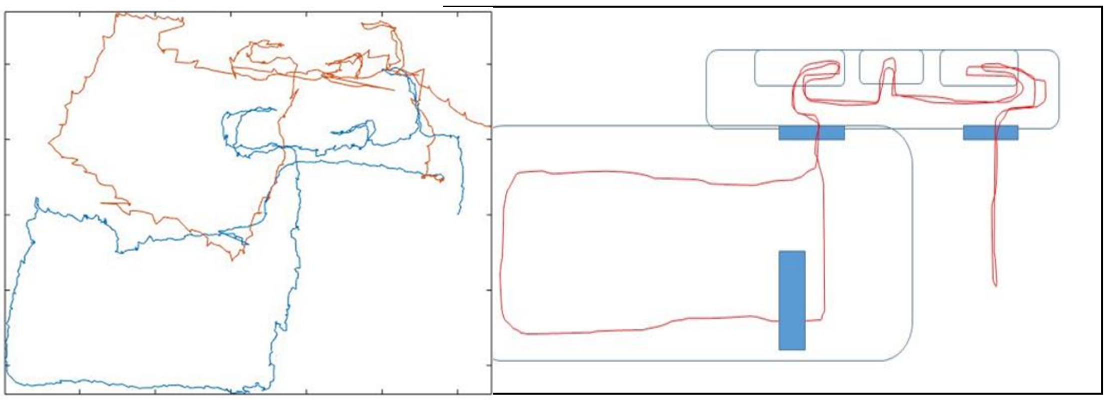

Figure 5: Horizontal localization result (on left) from UTTI test campaign, blue path done by walking and red by running. On right a rough sketch of the true path.

4.1.2 Effect of a pressure shock to the height solution

Effect of a pressure shock to sensors was tested by firing three practice cartridges. The effect is noticeable but does not cause a permanent bias in height measurement. Figure 6 shows that the effect disappears quickly. The first cartridge was fired at 220 seconds of the test, the next one two second after, and the third one at 240 seconds. An extremely small effect was also observed in accelerometers but this had practically no effect on horizontal localization.

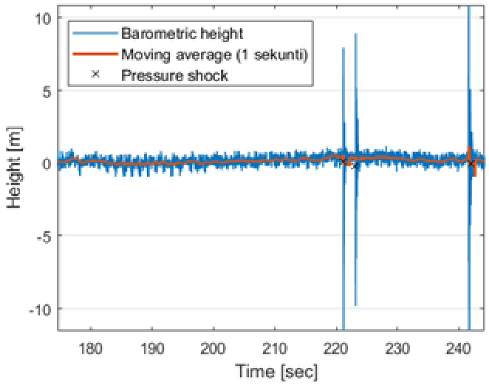

Figure 6: Barometric height during the test when pressure shock causes disturbances to sensors.

4.1.3 Motion recognition

The data used in these tests was gathered in Utti in February 2017. Amounts of the gathered training and test samples are presented in Table 1. We tested motion recognition using training data from one person, and test data from two different persons in order to test generalizability of the classification algorithms from person to person. We have experimented extensively with different classification and feature selection algorithms. Here we present results of two different classifiers; RandomForest, which has been used for motion recognition in INTACT before, and Naïve Multinomial Bayesian. These two classifiers produced the best outcome of the experimented ones. The tests were done using Weka, an open source data mining software.

Motion category	**Training data**	**Test data, person 1**	**Test data, person 2**
Walking	53	35	32
Running	39	39	65
Crawling	37	-	33
Crouching	40	35	22
Lying down	34	14	25
Standing	79	2	15
Climbing up	20	4	14
Climbing down	17	4	14

Table 1: Number of training and test samples for context recognition.

239 different features were computed for each sample of motion. However, as the number of training instances (319) is not much higher than the number of features, the number of features should be reduced as a one option in avoiding overfitting of the classifier. Overfitting usually leads to poor generalization of the classification algorithm. In RandomForest classifier this reduction comes naturally, as for each Random tree contributing to the final outcome a certain number of features is randomly chosen from all of the available features. For the Bayesian classifier we chose 75 most useful features based on each features worth in OneR classifier. These selected features contained readings from all sensors except sonar, thus verifying the previously obtained result that all sensors used for positioning contribute also to the motion recognition accuracy. The lack of sonar in these selected features may be due to the subideal direction of the sonar, resulting to measured distances larger than the capability of the sensor for most of the time.

The total classification accuracy for RandomForest, when both the training and test data are produced by the same person, is 93.23 %. This supports the previously obtained result of RandomForest classifiers good performance in motion recognition. However, when the test data was produced by different person than the training data, the overall accuracy reduced to 67.27 %. Crawling was often mixed with Crouching, and Lying down with Crouching and Standing. Therefore it can be deduced that the personal movement styles affect especially the recognition of these motion categories.

To improve the motion recognition in case of training data from different person than the final user, we experimented different classifiers and resulted in Naïve Bayesian Multinomial classifier. First, we replaced the possibly missing values with mean value of corresponding feature in the training data. Then, we scaled all of the feature values to range from 0 to 1 in order to prevent features with higher numerical values from dominating the outcome. After that we performed the previously described feature selection process in order to obtain 75 most relevant features. We also merged the Crawling and Crouching classes into "Moving forward in low posture", as the preliminary tests had shown that these categories were still difficult to distinguish from each other. After these preprocessing steps we trained the Bayesian classifier. The results of the classification are displayed in Figure 7. The total classification accuracy is 85%, which is a significant improvement compared to the performance of the RandomForest, when training and test data were from different persons. When not trying to distinguish Crawling and Crouching from each other, Moving forward in low posture is recognized from the test data with 100% accuracy. Climbing up and Climbing down are still often mixed with each other, Standing and Walking. This may be due to the insufficient sonar measurements in this dataset.

%	Walk	Run	Forward down	Lying down	Standing	Climbing up	Climbing down	Actual class
Walk	**84,38**	4,62	0	0	13,33	14,29	14,29	
Run	3,13	**90,77**	0	4	6,66	0	0	
Forward down	3,13	4,62	**100**	28	0	0	0	
Lying down	0	0	0	**68**	0	0	0	
Standing	6,25	0	0	0	**80**	7,14	21,43	
Climbing up	0	0	0	0	0	**64,29**	7,14	
Climbing down	3,13	0	0	0	0	14,29	**57,14**	
Classifier output							Total accuracy	85

Figure 7: Context recognition results using Naive Bayesian Multinomial classifier and test data produced by different person than the training data was produced by.

4.2 RECOMMENDATIONS FOR THE FUTURE IMPLEMENTATION

4.2.1 Strengths, Weaknesses, Opportunities, and Threats

The main strengths of the infrastructure-free tactical situational awareness system is its wide applicability without prior preparation of the operating environment. However, the relative nature of the observations cause the performance to degrade over time, and the absolute heading (with respect to North) is difficult to determine. However, the system could be integrated with other sources of information, and ideally the various sensors would be integrated to existing equipment carried by the soldiers. The most significant threat is expected to be the harsh operating conditions, posing strict requirements on the impact protection of the equipment.

Strengths:	**Weaknesses:**
• Applicability in variety of environments • Rapid deployment, no onsite preparation • Resilient to jamming and spoofing • Demonstrated with low-cost hardware • Can accommodate information from other sensors/sources as well	• Performance degrades over time in absence of absolute position information • Many sensors mounted at various parts of the body are needed • Difficult to determine the absolute heading when indoors • Co-operative radio ranging requires a high transmit power indoors
Opportunities: • Integration with other systems • Smart clothing and related technology to ease installation • Advances in sensor technology yielding better performance • Co-operative (networked) positioning can implement communications	**Threats:** • Users may ignore possible requirements on calibration • Extreme use conditions causing physical damage to components

4.2.2 Communication requirements

The transfer of information between soldiers is most convenient using a wireless radio signal. Despite the signal (WLAN, Bluetooth, Ultra-Wideband), to avoid congestion in communications one should only transmit the necessary information, especially when the size of the network is large. Therefore, the situational awareness solution should be computed locally and only the end result transmitted with adaptive transmission pace depending on the respective need. The channel access method, such as TDMA, ALOHA or some other, should be chosen carefully to meet the needs of the application. The wireless communications should also be encrypted to avoid leaking information to a hostile party.

5. Conclusions

The final results obtained during the two proof-of-concept tests assure that the methods selected and developed in the project for infrastructure-free tactical situational awareness are feasible for the implementation of a functional system. Recommendations for future implementation were given in the report.

6. Scientific publishing and other reports produced by the research project

1) A paper describing the development of a monocular SLAM algorithm encompassing novel methods for observing the heading and translation of the user from images:

Ruotsalainen L., Gröhn, S., Kirkko-Jaakkola M., Chen L., Guinness, R. and H. Kuusniemi (2015) "Monocular Visual SLAM for Tactical Situational Awareness", In Proceedings of the IPIN, 13-16 October, Banff, Canada, 10.1109/IPIN.2015.7346957.

2) Matine reports of the work done during the first two research years:

Ruotsalainen L., Kirkko-Jaakkola M., Chen L., Gröhn, S., and Guinness, R. (2015) Infrastructure-free tactical situational awareness (INTACT). MATINE Summary Report, ISBN 978-951-25-2755-7.

Ruotsalainen L., Kirkko-Jaakkola M., Chen L., Gröhn, S., Guinness, R. and Vallet J. (2016) Infrastructure-free tactical situational awareness (INTACT). MATINE Summary Report, ISBN 978-951-25-2840-0.

3) A paper describing the development of a Particle filter algorithm for fusing measurements from a foot-mounted IMU, camera, barometer and sonar for an accurate 3D localization:

Ruotsalainen L., Kirkko-Jaakkola M., Chen L., Gröhn, S., Guinness, R. and H. Kuusniemi (2016) "Multi-Sensor SLAM for Tactical Situational Awareness", In Proceedings of the ION ITM, 26-28 January, Monterey, California.

4) A paper discussing the research done on situational awareness, mainly on motion recognition:

Ruotsalainen L., Guinness R., Gröhn S., Chen L., Kirkko-Jaakkola M., and Kuusniemi H. (2016) Situational Awareness for Tactical Applications. In Proceedings of the ION GNSS+, 12-16 September, Portland, Oregon.

5) journal paper discussing the goals and results of INTACT done during the first 1.5 research years:

Ruotsalainen, L., Chen, L., Kirkko-Jaakkola, M., Gröhn, S., and H. Kuusniemi (2016). INTACT – Towards infrastructure-free tactical situational awareness. European Journal of Navigation, Vol. 14, No. 4: 33-38. ISSN 1571-473-X.

6) A journal paper discussing the future implementation of a method fusing the intact localization with Hyperspectral Lidar mapping and target recognition

Kaasalainen S., Ruotsalainen L., Kirkko-Jaakkola M., Nevalainen O., and Hakala T (2017). Towards Multispectral, Multi-Sensor Indoor Positioning and Target Identification, IET Electronics Letters, DOI: 10.1049/el.2017.1473.

Four journal papers will be submitted still during year 2017 and will therefore be added to the project's publication list.

7. Comment from the project steering group

"INTACT project addressed the user needs for creating and maintaining independent and infrastructure-free situational awareness. The project introduced the status and suitability of the existing commercial technology for positioning and obtaining situational awareness. The development done during the project and results obtained using the developed methods were promising. The developed methods and the chosen implementations create a good basis for follow-up actions on the way towards providing situational awareness in situations where e.g. the use of satellite positioning is denied."

Maj. Mika Nuutinen, Finnish Defence Forces, Army Command

References

Civera, O. Grasa, A. Davison and J. Montiel, "1-Point RANSAC for EKF Filtering: Application to Real-Time Structure from Motion and Visual Odometry," Journal of Field Robotics - Visual

Mapping and Navigation Outdoors, vol. 27, no. 5, pp. 609-631, 2010.

Davison, A., Reid, I., Molton, N. and O. Stasse (2007) "MonoSLAM: Real-Time Single Camera SLAM," in Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, Issue 6, pp. 1052-1067.

Liu, J. S. (2001), Monte Carlo Strategies in Scientific Computing. Springer.

Ruotsalainen, L. (2013) Vision-Aided Pedestrian Navigation for Challenging GNSS Environments, vol. 151, Doctoral Dissertation. Publications of the Finnish Geodetic Institute.