2016/2500M-0044

ISSN 1797-3457 (verkkojulkaisu)
ISBN (PDF) 978-951-25-2840-0

SUMMARY REPORT

Infrastructure-free tactical situational awareness (INTACT)

Laura Ruotsalainen (laura.ruotsalainen@nls.fi, 050 433 6552), Martti Kirkko-Jaakkola, Liang Chen, Simo Gröhn, Robert Guinness, Jose Vallet
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.

Intact will provide means for infrastructure-free tactical situational awareness by developing Particle filtering for multi-sensor fusion, a sophisticated SLAM algorithm and motion recognition means. Our approach is 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.

2. Research objectives and accomplishment plan

Most important research questions are 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 research year the project investigated individual positioning sensors and their performance. Also algorithms for obtaining motion measurements (i.e. range, speed, height and heading) from foot-mounted inertial sensor, barometer and sonar were implemented. Fusion algorithm is the core of multi-sensor positioning and therefore also the basis of a SLAM system. In first year two different fusion algorithms were developed (Kalman and Particle filters) and tested for the purpose. Based on the results Particle filtering was evaluated to be the most feasible method for the tactical applications requiring non-linear user motion observed with sensor suffering from non-Gaussian errors (Arulampalam et al.2002). Simultaneously, the development of machine learning algorithms for improved context recognition as well as a SLAM algorithm based on (Civera et al. 2010) was started, integrating vision-aided algorithms developed earlier at FGI (Ruotsalainen 2013) for improved performance of the SLAM solution.

During the second research year the development of the fusion algorithm, SLAM, and context recognition method were continued and deepened. The errors introduced by different sensors were modelled, and the obtained statistical models will be used in the Particle filtering algorithm for improved positioning performance. The SLAM algorithm was developed further by first improving FGI's visual positioning methods towards more robust performance and then integrating them into the 1-point Ransac SLAM (Civera et al. 2010). Motion context of the soldier is essential for accurate and reliable situational awareness. Command center is able to get information about the soldier's status from the motion context information, the display of the soldier's device may be adjusted based on the situational information drawn from the motion context information, and the accuracy and reliability of the multi-sensor fusion result will be improved by including the motion information into the algorithm. Therefore a thorough study and analysis of the most feasible sensors, features and machine learning algorithms for motion recognition was done and the method was developed further during the second research year. All developed methods were tested and the results analyzed frequently also during the second research year as described in this report.

3. Materials and methods

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

3.1 SLAM

Simultaneous Localization and Mapping (SLAM) means the capability of a user who is placed in an unknown location in an unknown environment to be able to incrementally form a consistent map of this environment and simultaneously determine his location within the map by using probabilistic computation (Durrant-Whyte and Bailey, 2006). Feasible SLAM solutions have been developed for robots. However, the requirements set for the equipment by the dismounted soldiers and rescue personnel, i.e. size and cost, necessitate the development of novel algorithms. Existing methods using a monocular camera and MEMS sensors do not provide sufficient performance yet and therefore new methods are developed in Intact-project. The 1-point Ransac SLAM algorithm developed by (Civera et al. 2010) has been used as the base for the SLAM development, but in Intact it will be extensively improved.

The concept of visual gyroscope and visual odometer (Ruotsalainen 2013) has been developed before at FGI providing measurements that will result in improved SLAM accuracy. During the first research year the methods were implemented into 1-point Ransac. In addition, during the first year we have developed methods to be able to use a wide angle camera, which provides much larger field of view. During the second research year this development has been continued to the point that the SLAM algorithm is now ready to be fused with the multi-sensor fusion algorithm. The methods developed are described below.

The Omnidirectional monocular 1-Point RANSAC Odometer algorithm inputs camera attitudes computed using the visual gyroscope and corresponding image features from consecutive images, and solves the translation using the visual odometer method. The input image feature pairs include many incorrect matches. These erroneous matches are discarded using the RANSAC error processing algorithm. When the SLAM system is initialized using the method described above, the computation is continued using a Perspective-n-Point (PnP) algorithm (Kneip et al. 2014). The benefit of using PnP is that it solves camera's pose with 6 degrees of freedom (i.e. three dimensional position and three dimensional attitude) with respect to provided 3D points, unlike the Omnidirectional monocular 1-Point RANSAC Odometer which only provides XY-translation. However, the map resulting from SLAM in Intact will be presented using only two dimensional point representation, and the three dimensionality will be used to improve the accuracy of the obtained localization. Figure 1 shows the features used for the PnP processing.

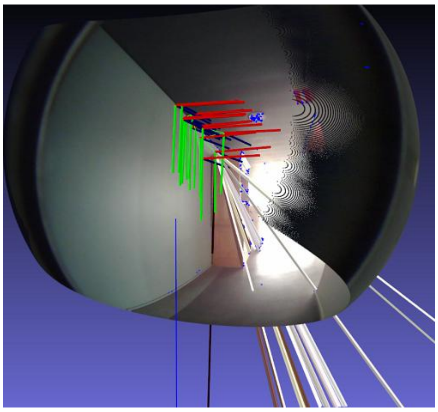

Figure 1: Features used for PnP algorithm

3.2 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 and will be discussed below. The multi-sensor fusion results are discussed in the Results and discussion section.

3.4.1 Measurement models

This section gives a brief overview of the sensors (shown in Figure 2) and measurements used for the multi-sensor fusion.

Self-contained sensors used for positioning provide information about the user motion, mainly translation and heading, or other phenomena useful for positioning purposes. When the initial position and direction are known, a relative position solution may be obtained by propagating the user position with the motion measurements. As mentioned above, the Intact multi-sensor fusion solution consists of measurements from three IMUs, one attached to the foot and used for positioning, one attached to the body and one to the helmet of the user, the two latter being used mainly for motion recognition. In general, the quality of low-cost

MEMS inertial sensors is inadequate for use in the above mechanization except for very short periods of time. However, when the IMU is attached to the foot of the user, whenever the IMU is detected to be at rest, namely during each steps, a zero-velocity update (ZUPT) may be applied to the error-state filter and improved performance obtained.

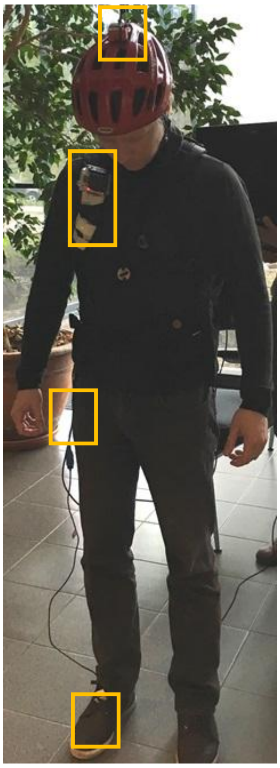

Figure 2: Sensors used for infra-structure-free tactical situational awareness

A camera can be considered as an additional self-contained sensor when integrated to a navigation system via specific mechanization. When the camera is attached to the body (shoulder in Figure 2), motion of features in consecutive images provides enough information for observing translation and heading of the user. Motion of the features may be transformed into heading information in a straight forward manner under favorable conditions and into translation by using a special configuration, namely by attaching the camera into a known height and tilted a bit downwards as discussed in (Ruotsalainen 2013).

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 (on helmet in Figure 2) computing the user vertical position component using the information about the change in the air pressure. However, barometer height measurements suffer from changes in air temperature and air pressure, not caused by the change of height but some other phenomena. Therefore, a sonar (on hip in Figure 2) pointing downwards and detecting if the change in the height computed by the barometer is really due to the change in height and not due to the change in environmental features, is used for making the altitude estimation robust.

3.4.2 Particle Filter Based Navigation Method

Particle filtering is based on the Bayesian statistical theory and Monte Carlo (MC) simulation. Particle filters provide a set of weighted MC samples of the state at each time instant (Liu, 2001). These samples are called particles. Particle filtering estimates the state of the system x_k at the time t_k based on all measurements up to that time (Gelman et al. 2000, Thrun et al. 2005). In Intact we compute the user's three dimensional position, and therefore the system state consists of the position components (horizontal X and Y, vertical Z) as well as the heading (h), namely $x_k = [X \quad Y \quad Z \quad h]$. Traditional fusion algorithms have assumed that the user motion is linear and all measurement errors have Gaussian distribution, which is seldom true in navigation and definitely not true in tactical applications. Therefore, heavy linearization has been done to enable the use of fusion algorithms causing inaccurate position solution. Particle filtering is an estimation method developed to process nonlinear measurements with non-Gaussian error models. Particle filtering algorithm considers the actual measurement error models when sampling the particles and therefore provides improved positioning accuracy.

During the second research year errors of the methods forming the core of the position solution, namely the foot-mounted IMU, the visual gyroscope and the visual odometer were modelled and will be further incorporated into the fusion algorithm. Results of the error modeling will be presented in the section 4.

3.5 SITUATIONAL AWARENESS

As discussed above, an important part of obtaining tactical situational awareness is the knowledge of the soldier's motion context. The main goal of the context research in Intact is to develop classifiers that are able to detect the user motion independently of the user's motion characteristics, resulting in user independent and therefore scalable method, and to integrate the obtained information into the fusion algorithm.

To achieve the goal, during the second research year we have studied the three main aspects affecting the formation of user independent classifiers for motion recognition. Namely, we have investigated the best set of sensors used for collecting the measurements, keeping in focus the fact that a good balance between the accuracy and the amount of sensors has to be found in order not to disturb the actual tactical operations. We have also studied the best set of features representing the data and used for motion classification, and the best machine learning algorithm for performing the actual classification.

Through discussions with the professionals at the tactical domain, we have identified 14 motion patterns to be detected; standing, walking, running, moving forward in crouching pose, crawling, turning, ascending stairs, descending stairs, getting down to crouching, staying static at crouching pose, rising from crouching to standing, getting down from standing to crawling pose, rising from crawling to standing and jumping.

Selection of a set of features, derived from various sensor measurements that capture information about the various motion contexts is one of the most crucial steps for motion recognition. The used features have to be justified to suit the challenging application area because their usage consumes resources, such as power and memory, and the large amount and obtrusive locations of sensors may disturb the tactical operations. The motions relevant for the applications presented above involve varying physical phenomena, e.g. large horizontal accelerations (crawling), large vertical accelerations (jumping), and large heading changes (turning). However, when the goal is a real-time application, a balance between the number of different features used and computational cost has to be found. In some cases, features may be useful for classification but computationally expensive to generate. We have followed the work of Frank et al. (2010) and Pei et al. (2010) for selecting the features. The features we investigated were means, variances and dominant frequencies of the computed motion measurements.

Machine learning (ML), also known as pattern recognition, is the tool we have chosen for obtaining situational awareness by classifying the different motion states using classifiers learned from training data. During the second research year we have studied the performance of various ML algorithms from the point of view of their classification accuracy and computational demands. The results are shown in Section 4.

4. Results and discussion

This section discusses the results obtained by processing the data collected in various data campaigns and processed by algorithms developed in the Intact-project. The equipment used for data collection were a GoPro camera, XSENS Inertial Navigation Sensor unit (IMU and barometer), two Osmium MIMU22BT IMUs, one attached to the foot and the other one to the body of the user, and a HRUSB-MaxSonar sonar for ranging. All accuracy values presented in the report are obtained by comparing the position solution computed with methods developed in Intact to the ground truth. The ground truth is computed using a Novatel SPAN system containing a dual-frequency GPS receiver and a tactical grade Honeywell's HG1700 IMU (Novatel webpages) and providing a solution with decimeter level accuracy.

4.1 INTEGRATION ALGORITHMS

This section discusses the results of the error modelling and shows the fusion result using the Particle filtering algorithm at its present state.

When the Particle filter integration algorithm is used for multi-sensor fusion with Gaussian error models, the position mean position error during a route with a length of around 200m was 1.88 m with a standard deviation of 3.19 m. The result, shown in Figure 3, is feasible for indoor positioning, however it was obtained during one experiment and relatively smooth motion. Therefore, the next step of the fusion research is to include the obtained more realistic error models into the Particle filtering algorithm.

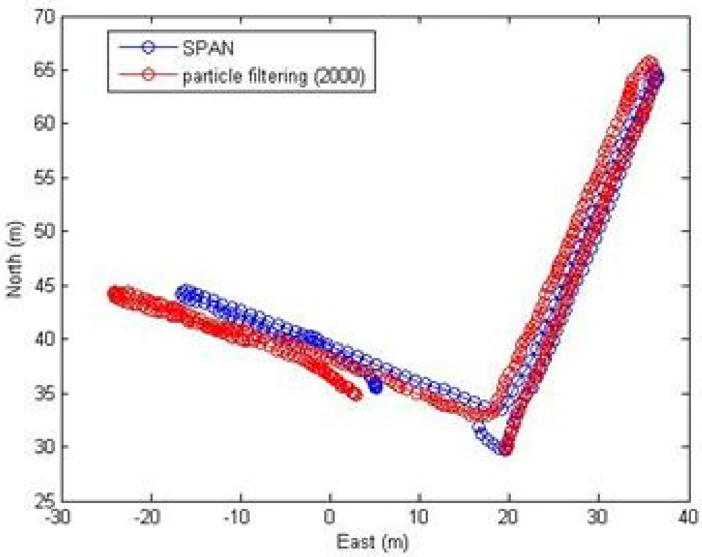

Figure 3. Particle filtering result

The errors arising when using the foot-mounted IMUs and the visual gyroscope and visual odometer were modelled by collecting many set of data during data campaigns and by comparing the obtained results to the ground truth. Figure 4 shows the modelling result for the foot-mounted IMU measurements. Figure on left shows the route walked and measured using the SPAN reference system (with red) and the solution computed using foot-mounted IMU (blue). The obtained position solution is deviated due to the measurement errors. The errors are plotted using a histogram and shown in figure on right. Two probability distribution functions are fitted to the errors, Gaussian and Student-t. As may be seen from the figure, Student-t distribution fits the error distribution better, and the parameters were computed to be mean -1.495 meters and standard deviation of 6.883 meters. Similar computations were made to the visual methods.

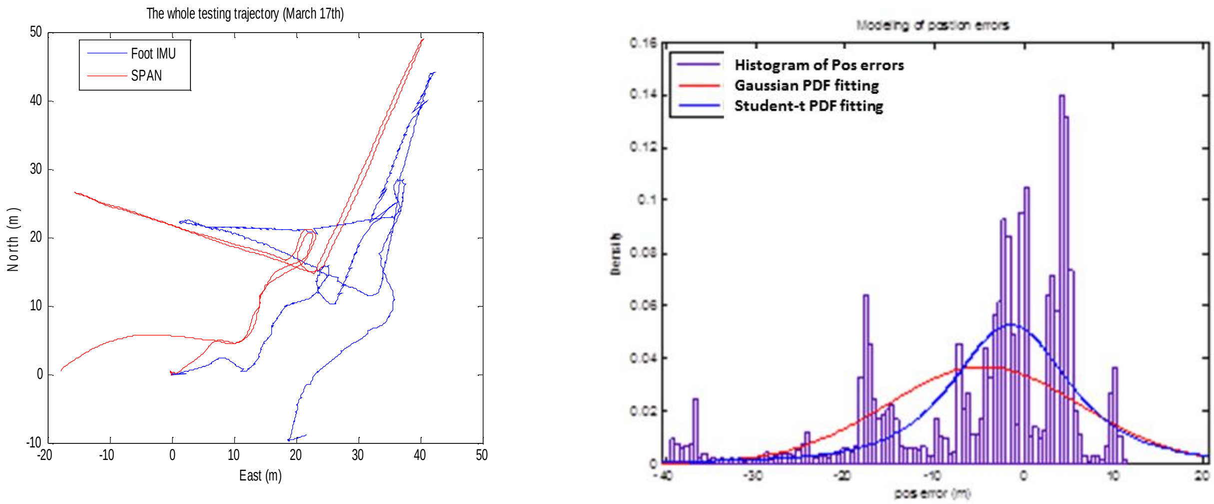

Figure 4. Error modelling result for foot-mounted IMU

4.2 Context recognition

Here we present the results from the extensive motion recognition research done for identifying the best set of sensors and features as well as the best Machine learning algorithm to be used for motion recognition. The classification results are summarized in Figure 5.

The overall accuracy of the classification with this setup was 78.2 % (percentage of correctly classified instances). The main motions, even more unusual ones like crawling, are well detected from the data, namely the detection accuracy for crawling was 100% for the data collected during the tests. However, motions like jumping and getting up from the crawling and crouching poses are often confused with other motions. The reason for this is that the data available for training the classifiers to detect these motion patterns were too few in these data sets. Table 1 shows the appearance of different motions in the data. Also, there are some motions in the data that are not mutually exclusive, like turning that can happen while also walking or running. Classification for these motions resulted in confusion decreasing also the overall accuracy. In the future research, we will join similar transition motions to form one class and therefore contribute better for the classification procedure.

Table 1 Appearance of different motions

Motion	Instances in data [s]
Walking	73
Standing	53
Turning	29
Ascending stairs	26
Descending stairs	21
Running	18
Crawling	16
Staying static at crouching pose	13
Moving forward in crouching pose	10
Rising from crouching to standing	6
Getting down to crouching	5
Rising from crawling to standing	3
Getting down from standing to crawling pose	1
Jumping	1

The results showed also that having three IMUs, namely the first attached to the user body, the second to the foot, and the third to the helmet, will bring more information and therefore accuracy for the classification. We also studied the best combination of other sensors, measurements, and features used for detecting the contexts and found out that removing any of the sensors used in the data collection decreased the classification accuracy. The only feature not contributing to the classification accuracy was the “Mean of heading change”.

We compared the performance of several different ML algorithms for learning the classifiers. Results showed that the decision-tree based algorithm RandomForest outperformed all other tested algorithms using default parameters. In the future the research will continue on forming the relevant motion classes and the classification accuracy is anticipated to improve.

	Asc	Crawl	CroFwd	Crouch	Desc	Run	Stand	Turn	Walk
Asc	92.3							7.7	
Crawl		100.0							
CroFwd			80.0				10.0		10.0
Crouch				53.8			7.7		7.7
Desc					90.5			9.5	
Run						88.9	11.1		
Stand							88.7		11.3
Turn	17.2				24.0		6.9	37.9	17.2
Walk	1.4						6.8	5.5	86.3

Figure 5: Classification accuracy in percent for different motions, horizontally the detected motion and vertically the correct motion pattern. Note that some classes have been left out of this table because of only few occurrences in the data, causing some of the lines not to sum up to 100 %.

5. Conclusions

The results obtained during the second research year assure that the methods selected to be developed for infrastructure-free tactical situational awareness, are the most feasible ones for the purpose. However, there is still a need to develop the fusion of all measurements and methods into one adaptive system, for providing reliable and accurate solutions for all situations and environments.

In order to finally obtain infrastructure-free, accurate and reliable tactical situational awareness, all methods developed so fast should be integrated into one system. Therefore, next steps in the research will be the incorporation of the error models into the fusion algorithm and integration of the Particle filtering and SLAM algorithms. Also the motion recognition classifier will be integrated into the fusion algorithm in order to obtain an adaptive, accurate and reliable system.

Finally, when the performance of the developed system is sufficient, a proof-of-concept testing in a more realistic and challenging operation environment, taking into consideration also e.g. changes in temperature and pressure conditions, should be made. Also constrains and possibilities for implementation of the methods into a real operational system should be discussed in the future.

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 report of the work done during the first research year:

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.

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, submitted.

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) A 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.

References

Arulampalam, M. Maskell, S., Gordon, N. and T. Clapp (2002) “A tutorial on particle fiters for online nonlinear/non-Gaussian Bayesian tracking,” IEEE trans. on Signal Processing, vol.50, no.2, pp.174-188.

Chen, R. and R. Guinness (2014) Geospatial Computing in Mobile Devices, Artech House.

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.

Durrant-Whyte, H. and T. Bailey (2006) “Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms” in IEEE Robotics and Automation Magazine, vol. 2.

Frank, K., Vera Nadalesz, M.J., Robertson, P. and Angermann, M. (2010). “Reliable Real-Time Recognition of Motion Related Human Activities Using MEMS Inertial Sensors” in Proceedings of ION GNSS, September, Portland OR, U.S. Institute of Navigation

Gelman, A.B., Carlin, J. S., Stern, H. S. and D. B. Dubin (2000) Bayesian Data Analysis, 2nd ed. Chapman & Hall.

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

Novatel’s web pages http://www.novatel.com/products/span-gnss-inertial-systems/, last accessed 30.11.2016reference

Pei, L., Liu, J., Guinness, R., Chen, Y., Kuusniemi, H and Ruizhi Chen (2012). ”Using LS-SVM Based Motion Recognition for Smartphone Indoor Wireless Positioning”. Sensors, Vol. 12, pp. 6155-6175.

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

Thrun, S., Burgard, W. and D. Fox (2005) Probabilistic Robotics (Intelligent Robotics and Autonomous Agents). The MIT Press, September.