2015/2500M-0033

ISSN 1797-3457 (verkkojulkaisu)
ISBN 978-951-25-2755-7 (PDF)

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
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. Also, information of the soldier's context is important for successful operations, e.g. if the soldier is running, crawling or static for a long time. Requirements for the system are stringent; it should function also in indoor environments, lightweight and inexpensive. 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. 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 will implement SLAM using a monocular camera as an input. However, existing algorithms for monocular SLAM do not provide reliable enough results for rescue or military applications.

At present, most functioning indoor localization systems are based on processing short- range radio signals from pre-installed networks and therefore cannot be considered as infrastructure-free. Advances in sensor technology have been rapid during the last several decades. Self-contained Micro-Electro-Mechanical (MEMS) sensors fulfill the size and cost requirements set for an infrastructure-free military and rescue system (Rantakokko et al. 2011). Use of, e.g., inertial sensors provides enough information for propagating a known initial position for the purposes of forming a SLAM solution with a camera. However, the MEMS sensors suffer from biases and drift errors that may decrease the position accuracy substantially (Collin 2006). Therefore, sophisticated error modelling and implementation of integration algorithms are key for providing a viable final result.

Integration of different sensors has been an active research area already for some years, but there does not yet exist an accurate and reliable infrastructure-free indoor positioning system. Our approach is to integrate a monocular camera, multiple Inertial Measurement Units (IMUs), a barometer and a ranging sensor to obtain a solution for SLAM, as well as tactical motion information. This project investigates also some sensors less used for positioning, such as ultrasound, for obtaining more accurate positioning and also resolving the height of the camera, which is needed for the visual processing of the method discussed below. Also, positioning using multiple inertial sensors is studied. One inertial unit is foot mounted and other units will be placed on the helmet and body enabling context recognition, e.g. observing the dynamics of the soldier or if he has been static for a long time and therefore possibly wounded.

Digital TV (DTV) signals penetrate buildings much better than satellite signals and therefore provide absolute position information also indoors. The use of DTV indoors has been research only slightly until now. Intact-project investigates the accuracy of an absolute position solution obtained using DTV signals.

2. Research objectives and accomplishment plan

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. The research investigates also some sensors, previously less used for indoor positioning but suitable for the task, such as ultrasound ranging equipment attached to the soldier's person and digital tv signals, whose antenna may also be attached to soldier's equipment.

At the beginning of the INTACT project an extensive literature review was made for obtaining knowledge of the state-of-the-art. Then, all equipment needed for the research were selected, purchased and studied. Algorithms for obtaining motion measurements (i.e. range, speed, height and heading) from foot-mounted inertial sensor, barometer and sonar were implemented. After evaluating the measurements and performance of each independent sensor a test campaign was done for collecting data from all equipment attached to a test person. Then, development for fusion, machine learning and SLAM algorithms was started. First a Kalman filter was implemented for multi-sensor fusion. As Kalman filter is aimed for fusions with linear models, which motion of an unmounted soldier is definitely not, a Particle filter fusion algorithm was also developed. Simultaneously, development for machine learning algorithms for improved context recognition was carried on. Also, the development of 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. The developed methods were tested and the results analyzed frequently, every time new tasks were accomplished. The developed methods and results are discussed 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

Slam algorithm provides simultaneous localization of the user and a map of the unknown environment. Feasible SLAM solutions have been developed for robots. However, the requirements set for the equipment by the unmounted 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 SLAM algorithm developed by (Civera et al. 2010) will be used as the base for the SLAM development, but in Intact it will be heavily improved by integrating the visual gyroscope and visual odometer (Ruotsalainen 2013) measurements for improved accuracy, by further developing the error detection algorithms and by developing the map properties to be more suitable for tactical applications, namely by computing also the vertical position solution and improving the representation.

3.2 FOOT-MOUNTED PDR

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. This is why pedestrian dead reckoning (PDR) mechanizations are typically used which resort to detecting steps from the inertial sensors' output waveform and applying an external model for the stride length (Beauregard and Haas 2006). Such mechanization avoids the error-prone double integration of inertial measurements. However, human gait differs from person to person, which makes the stride length difficult to predict (Leppäkoski 2015). However, mounting the IMU to the foot of a pedestrian constitutes a special case: Unless the shoe slips, the IMU remains stationary for a short period of time between steps (Foxlin 2005). This makes it possible to compute the step displacement and therefore avoid the need for an external step length model. The method also works when stepping sidewards or backwards, and when walking in stairs

However, the MEMS gyros are quite sensitive to temperature changes and need minutes to stabilize after e.g. entering indoors from cool outdoor conditions before they can provide accurate heading information (Leland 2005). Also, high linear accelerations induce large G-sensitivity errors (Bancroft and Lachapelle 2012). Tactical operations often include both large changes in the temperature and large accelerations. Therefore, foot-mounted IMUs alone do not provide sufficient performance.

The research on foot-mounted inertial sensors has been conducted using the OpenShoe platform (http://www.openshoe.org/) as the source of measurements. The OpenShoe inertial measurement unit (IMU) is small in size and weight as can be seen in

Figure 1, allowing non-obtrusive mounting onto the shoe, and its accelerometers and gyroscopes represent the performance level of modern mass-market microelectromechanical (MEMS) inertial sensors.

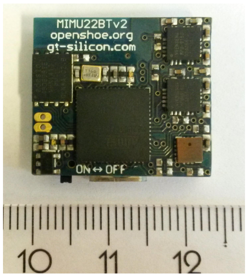

Figure 1 OpenShoe inertial measurement unit (length scale in centimeters)

3.3 ULTRASOUND

Ultrasound has been used for positioning for long already. However, for the moment most of ultrasound positioning systems use transmitters attached to the environment and therefore require an infrastructure, whereas ultrasound ranging equipment contain both the transmitter and receiver. The project investigates how accurate position information may be obtained by doing the measuring continuously to multiple directions using equipment attached e.g. to the soldier. Finally, algorithms are developed for integrating the ultrasound ranging measurements with other sensors and SLAM algorithms and the accuracy of the resulting localization and mapping is analyzed.

3.4 MULTI-SENSOR INTEGRATION

This section gives a short overview of Bayesian estimation methods used for integrating measurements from different sources. Two different integration algorithms, a Kalman filter and a Particle filter, have been developed in Intact in order to find the best performing fusion. The multi-sensor results are discussed in the Results and discussion section.

3.4.1 Recursive Bayesian estimation

Recursive Bayesian estimation algorithms (Gelman et al. 2000, Thrun et al. 2005) are commonly used to estimate the state of a system x_k at the time t_k based on all measurements up to that time. The recursive Bayesian estimation is performed with the following two steps: 1) *Prediction: a priory* probability is calculated from the last *a posteriori* probability using the process model. 2) *Update: a posteriori* probability is updated using the measurement model (9) and the current measurement. Depending on how the probabilities are represented and transformed in the process and measurement models, the recursive Bayesian estimation algorithms are implemented in different ways. For a linear and Gaussian probability density function (p.d.f.) model, the Kalman Filter (KF) is an efficient and optimal solution in the least square sense (Anderson and Moore 1979). Recently, Sequential Monte Carlo (SMC) Filters, such as the particle filter (PF) (Arulampalan et al. 2002) has been applied in the state estimation in the nonlinear and non-Gaussian models, where the probability densities are represented by a set of random state space samples drawn from the corresponding distribution.

3.4.2 Kalman Filter Based Navigation Method

A very simple Kalman filter was developed for obtaining a three-dimensional multi-sensor position solution. The state model of the filter included latitude, longitude, height, heading rate, heading and speed. Filter used data from the visual gyroscope and odometer, XSENS IMU, barometer and sonar as measurements. The filter mechanization, e.g. the covariance matrices, followed the development discussed in (Kuusniemi et al. 2011). The results are discussed more in the Results section.

3.4.3 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.

In Intact, we have implemented a particle filter that computes 3D position, heading and speed, all information needed for an indoor navigation solution. More information of the algorithm may be found from (Ruotsalainen et al. 2016).

3.5 SITUATIONAL AWARENESS

Situational awareness, in the mobile computing community also known as context awareness, is the ability to automatically provide information for seven key questions concerning the user: what, who, where, when, why, in what manner, and by what means (Chen and Guinness 2014). This is done by using machine learning algorithms. In the first year of Intact, preliminary investigation into the capabilities for context recognition using multi-sensor data has been initiated. Using the data from the Xsens IMU, we tested the performance of several different machine learning classification algorithms to recognize several different motion / pose contexts, including: walking, walking slowly, walking very slowly, standing, jogging, ascending stairs, descending stairs, turning around, standing to crouching position, crouching. The results are shown in Section 4.

3.6 DTV POSITIONING

Digital television signals have been recognized to be promising signals of opportunity for wireless positioning and due to their features described below, they are feasible for indoor positioning. DTV signals can be used for computing pseudorange measurements between the transmitter and the receiver like in satellite positioning, and when signals from at least four transmitters are received, a three dimensional position solution may be computed by trilateration. DTV signals have larger transmission power (10 to 15 KW) compared to satellite positioning signals and therefore they are easily transmitted through walls and windows. Also, as the transmitters are on the ground, the signals have shorter path through structures when entering the indoor environment through walls instead of the roof and stores, maintaining more power. DTV signals have also large signal bandwidth (6-8 MHz) which leads to more accurate pseudorange measurements.

In Intact, a self-developed USRP software defined DVB-T receiver is used to sampling the DVB-T2 signals indoors and analyzing the signal spectrum. The central frequency of the sampled signal is set as 177.5 MHz and sampling rate is 10 Msamples/s. The results are discussed in the next section.

4. Results and discussion

This section discusses the results obtained by processing the data collected in a data campaign 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), an Osmium MIMU22BT IMU attached to the foot and a HRUSB-MaxSonar sonar for ranging.

4.1 SLAM

So far our main contribution to current SLAM technologies has been the integration of the visual gyroscope and visual odometer to the 1-point ransac algorithm by (Civera et al. 2010). In addition, we have used a wide angle camera which provides much larger field of view. This caused us to move from more traditional pinhole camera model, radial and tangential lens distortion correction to more flexible camera calibration technique for omnidirectional cameras (Scaramuzza et al. 2006). This improves the accuracy of the visual motion perception, but requires special processing of the images. The initial results of the SLAM localization processing are very promising and have been published in (Ruotsalainen et al. 2015).

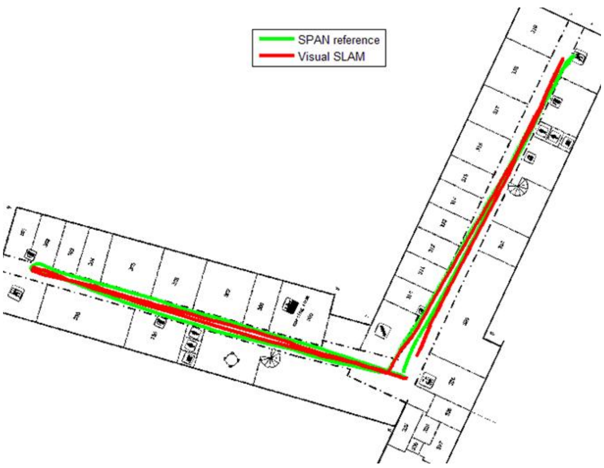

Figure 2. An initial SLAM solution using a monocular camera. The accuracy of the position is 1.8 m, however the turns in the corridor ends are re-initialized by using the reference system.

SLAM –systems using monocular camera provide maps that consists of image feature points. Feature points are extracted from consecutive images and matched. Significant portion of the matches are wrongly paired and this causes errors to the final attitude and translation estimates. To minimize the count of wrong matches, we have used the 1-point Ransac algorithm and further improved its performance as described above.

4.2 FOOT-MOUNTED INERTIAL MEASUREMENT UNIT

Foot-mounted pedestrian dead reckoning was studied in use cases where the user is walking on a level surface, although the foot-mounted inertial navigation mechanization can cope with altitude changes as well; investigating the performance with height changes is left for future work. Raw data were recorded from the IMU at a sampling rate of 125 Hz or more, depending on the test, and all position computations were carried out in Matlab as post-processing; however, the positioning algorithms themselves are not limited to post-processing but can be run in real time.

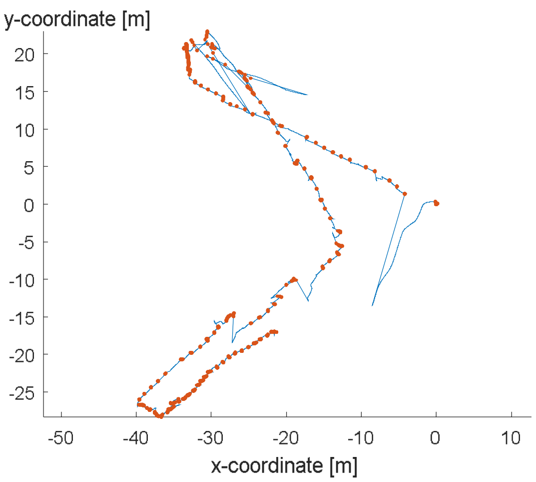

Figure 3 Example foot-mounted pedestrian dead reckoning trajectory estimate

The test results suggest that the foot-mounted pedestrian dead reckoning implementation can yield a few meters relative horizontal position accuracy without external assistance or prior calibration (the only assumption is that the IMU is stationary at start-up, but this period need not last longer than one second). An example pedestrian dead reckoning result is shown in Figure 3 where the red dots indicate instants where the foot was detected to be stationary. It can be seen that when such detections are missing for a long time, the position accuracy degrades rapidly; therefore, it is critical to develop robust foot stance phase detection algorithms that can work in different use cases, such as running and crawling, in addition to regular walking. The reason for the abrupt heading change near position (-30, 20) is unknown, but it is probably related to starting the experiment straight after entering the indoor area or a stance phase detection error.

4.3 INTEGRATION ALGORITHMS

This section discusses the results obtained by the Kalman filter and Particle filter integration algorithms developed for multi-sensor fusion.

4.3.1 HORIZONTAL POSITION SOLUTION

Kalman filtering

The Kalman filter algorithm discussed above was tested by integrating the measurements form different combinations of sensors used from the set of: monocular GPro camera, XSENS IMU and barometer, sonar and Novatel SPAN tactical grade GNSS/IMU system used normally as a reference. The accuracy (mean error) of the horizontal position result obtained varied from 2 meters (SPAN IMU, visual odometer and visual gyroscope) to 5.4 m (XSENS IMU, visual odometer, visual gyroscope). As a conclusion we decided to rely heavily on the measurement of the visual processing, use foot-mounted IMU instead of the body-mounted mechanization and develop a particle filter more suitable for the multi-sensor fusion of the application in question.

Particle filtering

The particle filter developed in the Intact-project was used for fusing measurements from the visual gyroscope and odometer, foot-mounted IMU, a barometer and sonar. Only the intial position and heading were initialized using the reference system, but since all processing was done in an infrastructure-free manner. The mean error of the position obtained in the 160m long path was 3.14m with the standard deviation of 2.8 m. The results, shown in Figure 4, were very promising.

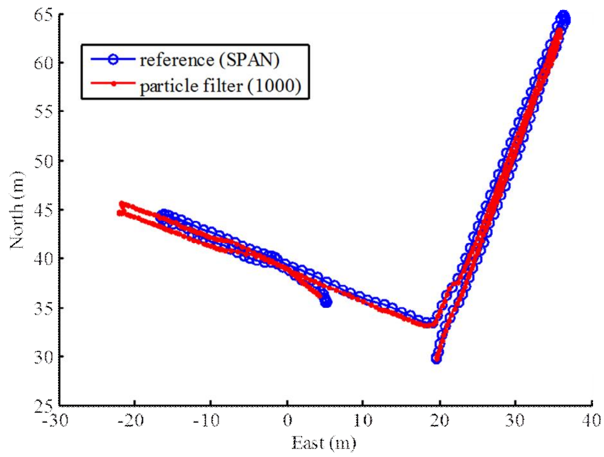

Figure 4. Particle filter fusion results compared with the SPAN reference trajectory in horizontal plane

4.3.2 VERTICAL POSITION SOLUTION

The standard deviation of the height observed in the experiments using only barometer was 0.15m, ultrasonar 0.01m and using the Particle filter developed 0.03m. The results, represented also in Figure 5, show that the integration of barometer and sonar measurements improves the precision and therefore the stability of the height solution. However, the improvement is incremental and the barometer would be sufficient in a favorable environment and situation. The most important benefit from fusing the barometer height and sonar range measurements is that it will provide improved vertical accuracy and reliability in the case of unexpected changes in environment’s pressure and temperature, which distort the height solution heavily.

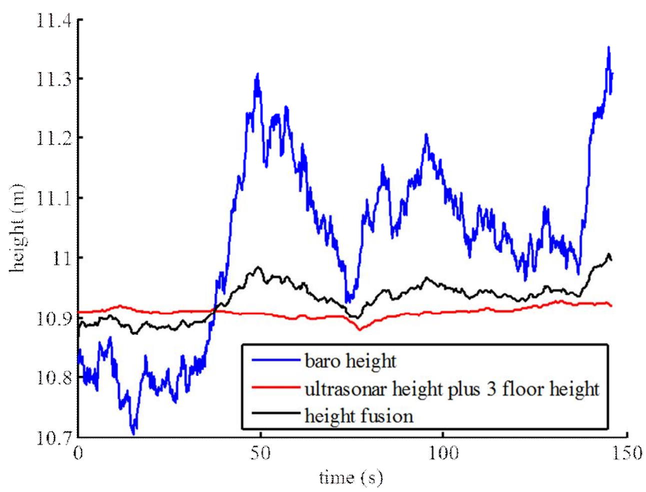

Figure 5. Barometer and sonar height and the fusion results

4.4 Context recognition

Here we present the results from the preliminary investigation into the capabilities for context recognition using multi-sensor data. Using the data from the Xsens IMU, we tested the performance of several different machine learning classification algorithms to recognize several different motion / pose contexts, including: walking, walking slowly, walking very slowly, standing, jogging, ascending stairs, descending stairs, turning around, standing to crouching position, crouching. We found that decision-tree based classifiers performing rather well, achieving correct context recognition about 98.5% of the time for our limited test data. The classification result is shown in Figure 6.

Predicted Label → / Actual Label ↓	walking	ascending stairs	standing	turning around	walking very slowly	walking slowly	descending stairs	jogging	standing to crouch	crouching
walking	**98.96**	0.12	0.45	0.30	0.04	0.02	0.01	0.09	0.00	0.00
ascending stairs	1.27	**98.27**	0.33	0.00	0.00	0.13	0.00	0.00	0.00	0.00
standing	1.10	0.11	**98.36**	0.12	0.17	0.12	0.00	0.01	0.00	0.00
turning around	7.09	0.09	1.00	**91.73**	0.00	0.09	0.00	0.00	0.00	0.00
walking very slowly	0.79	0.00	1.17	0.00	**98.04**	0.00	0.00	0.00	0.00	0.00
walking slowly	0.19	0.00	1.06	0.00	0.25	**98.50**	0.00	0.00	0.00	0.00
descending stairs	0.29	0.00	0.14	0.00	0.14	0.00	**99.43**	0.00	0.00	0.00
jogging	0.68	0.00	0.59	0.14	0.00	0.14	0.00	**98.45**	0.00	0.00
standing to crouch	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**98.33**	1.67
crouching	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.31	**97.22**

ERRORS

ERRORS

(All values are percentage of total in-class samples)

Figure 6: Detailed context analysis using a Confusion matrix algorithm

In the future, we will investigate a wider range of combat-related contexts and integrate multiple sensors into the context recognition routine. We will also in-vestigate different methods for feature generation and feature selection. Lastly, we will expand the number of machine learning classification algorithms to be investigated and include parameter tuning in our performance evaluations.

4.5 DIGITAL TV SIGNALS

This section discusses the indoor DVB-T2 tests carried out in Intact-project. The testing scenario is shown in Figure 7 and the spectrum analysis results are presented in Figure 8. Based on the spectrum analysis. The nominal bandwidth is 7 MHz, and the useful bandwidth shown in the tests is 6.6 MHz, which is consistent with the standard (ETSI 2011). The tests also showed that, the spectrum is not as flat as the theoretical value, which suggests that the DVB-T2 signals should have experienced severe multipath and fading in the indoor scenario.

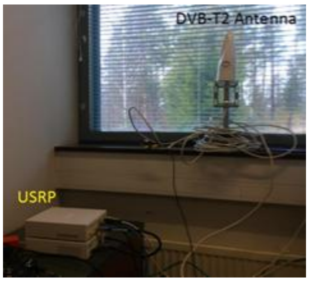

Figure 7. indoor DVB-T2 testing scenario

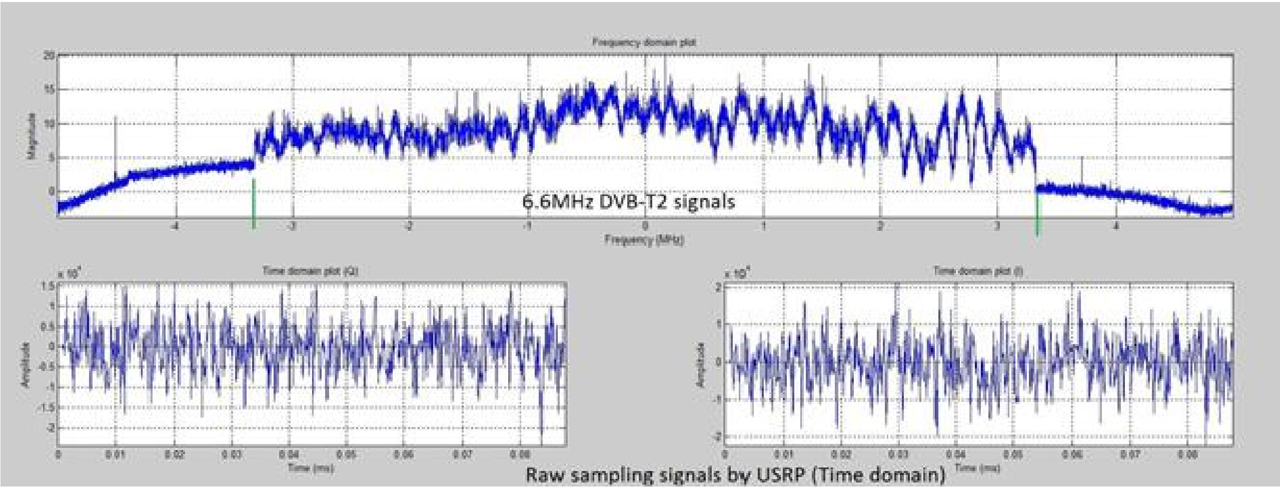

Figure 8. Signal spectrum analysis on DVB-T2 signals acquired indoors

The future work will focus on analyze the signal property of the DVB-T2 and study on the methods for TOA based ranging method.

5. Conclusions

The results obtained are very promising both for the selection of the equipment to be used as for the algorithms developed in the project. However, the accuracy of the position solution should still be slightly improved to be in the scale of 1-2m, usually recognized to be optimal for most positioning needs. Also, so far the algorithms have been tested only in an office corridor, during a few minutes experiment and with a test person mainly walking in a near constant speed. Therefore, the next steps in the research would be to continue the careful error modelling and error detection algorithms for improved integration results, for both improving the accuracy and the scalability of the method. Also, the next steps would be further developing the SLAM algorithm and improving the representation of the map to include more information of the environment. The motion recognition algorithms will be also developed to correctly classify more difficult motion patterns, e.g. crawling. Finally, when the performance of the developed system is sufficient, test campaigns in a more realistic and challenging environment should be made.

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

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, in press.

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.

References

Anderson, B.D.O. and Moore, J.B (1979) Optimal Filtering. Englewood Cliffs, New Jersey: Prentice–Hall.

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.

Collin, J. (2006) Investigations of self-contained sensors for personal navigation, Ph.D. dissertation, Tampere University of Technology, Finland

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.

ETSI (ETSI 2011), "302 755: Digital video broadcasting (DVB), " Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2), 2011.

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

Julier, S. Uhlmann, J. and H. Durrant-Whyte (2000) “A new method for the nonlinear transformation of means and covariances in filters and estimators,” IEEE Transactions on Automatic Control, vol. 45, no. 3, pp. 477–482, Mar

Kuusniemi, H., Chen, L., Ruotsalainen, L., Pei, L., Chen, Y., Chen, R. (2011), Multi-Sensor Multi-Network Seamless Positioning with Visual Aiding, Proceedings of IEEE/ ICL-GNSS 2011, IEEE Xplore, 29-30 June, Tampere, Finland, pp. 146-151.

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

Rantakokko J., Rydell, J., Stromback, P., Handel, P., Callmer, J., Tornqvist, D., Gustafsson, F., Jobs, M. and Gruden, M. (2011), “Accurate and reliable soldier and first responder indoor positioning: multisensor systems and cooperative localization”, in IEEE Wireless Communications, Vol. 18, Issue 2, pp. 10–18.

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

Scaramuzza, D., Martinelli, A. and Siegwart, R., (2006). "A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion", Proceedings of IEEE International Conference of Vision Systems (ICVS'06), New York, January 5-7.

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