



HAL
open science

Pedestrian Localization: Literature Review and Illustration of a Monocular Vision Based Approach

Wael Elloumi, Rémy Leconge, Sylvie Treuillet

► **To cite this version:**

Wael Elloumi, Rémy Leconge, Sylvie Treuillet. Pedestrian Localization: Literature Review and Illustration of a Monocular Vision Based Approach. International Conference on Signal, Image, Vision and their Applications SIVA'11, Nov 2011, Guelma, Algeria. p. 199-204. hal-00678783

HAL Id: hal-00678783

<https://hal.science/hal-00678783>

Submitted on 13 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pedestrian Localization: Literature Review and Illustration of a Monocular Vision Based Approach

First A. W. Elloumi, Second B. R. Leconge, and Third C. S. Treuillet

Abstract—Locate our position is one of the tasks we naturally do with our vision and our memory. However, ensuring this task in an artificial way is still a challenge for the whole community of computer vision. This paper presents a detailed literature review of pedestrian localization systems and their technologies. In the context of the visually impaired pedestrian navigation assistance, we also present an innovative solution based on the vision localization for both outdoor and indoor navigation. This new solution is established based on the independence of any collective equipment.

Index Terms—Computer Vision, navigation assistance, pedestrian localization, visually impaired.

I. INTRODUCTION

To move is a vital necessity for every person. In former times, man has developed many technical means to facilitate his movements and increase his autonomy: after the map and the compass, navigation assistance systems have experienced a great progress since the beginning of this new century with the advent of GPS (Global Positioning System). If the GPS is the reference system to locate moving vehicles, ships or aircraft, it is still insufficient for pedestrian localization because of its low accuracy. Besides, it is often ineffective in urban areas or in indoor environments.

Other technologies based on networks of communicating tags (WiFi, Zigbee, RFID, etc.) are being tested. But up until now, none is yet operational. The pedestrian localization remains a challenge because the intended accuracy, reactivity and reliability for a continuous guidance are far more important than robotics. In order to be practical for the visually impaired users (VI), a personal assistive navigation system

needs to be wearable, low cost, and free from collective equipment. The most challenging issue is the instantaneous and accurate spatial localization for guiding the VI indoor/outdoor daily walks, supplementary to obstacle avoidance that they can discern on their own using a long cane. To move safely, VI needs a navigation system with sufficient accuracy to keep the walker in a corridor of one meter wide along few memorized paths.

In this paper, we propose a literature review of recent works on this subject and our innovative approach.

The outline of this paper is as follows: the next section presents the specifications of the pedestrian localization against the ones of vehicles or robots. Section 3 details a review of the most important applications of such a pedestrian navigation system. Section 4 shows a state of the art of proposed researches in this field and distinguish between autonomous systems and those based on an existing infrastructure. Finally, we propose our monocular vision based method, our results and conclusion.

II. DIFFERENCE BETWEEN PEDESTRIAN LOCALIZATION AND ROBOT OR VEHICLE LOCALIZATION

The guided robot or vehicle localization is often based on a GPS as they are meant for outdoor use. The GPS receiver can only identify a position but it is unable to know neither roads or wrong ways nor places. That is why the system must use a digital map that includes the road network. The main and secondary roads have been digitalized by specialists with a few meter precision (from 5 to 20m). This map includes the direction of the traffic and the important infrastructure such as car parks, hotels, hospitals, railway stations, airports, etc.. Thanks to this map we can choose the road that we want to take. To have more accuracy, the system often incorporates a gyroscope which enables it to know some motions of the vehicle and so to identify significant changes in direction as the turns. It includes also an odometer to know the distance traveled by the vehicle which may also be obtained by integrating the velocity measured by the vehicle speedometer. The system performs a recalibration, using map-matching algorithms. It compares information from the gyroscope to the one stored the map data. The whole system can achieve an

Manuscript received October 27, 2011.

Wael Elloumi, Author A is with the Laboratory PRISME, Université d'Orléans, Ecole polytechnique, 12 rue de Blois, 45067 Orléans, France (wael.elloumi@univ-orleans.fr).

Rémy Leconge, Author B is with the Laboratory PRISME, Université d'Orléans, Ecole polytechnique, 12 rue de Blois, 45067 Orléans, France (remy.leconge@univ-orleans.fr).

Sylvie Treuillet, Author C is with the Laboratory PRISME, Université d'Orléans, Ecole polytechnique, 12 rue de Blois, 45067 Orléans, France (sylvie.treuillet@univ-orleans.fr).

accuracy of 10 meters which is reasonable to guide a vehicle driver in an urban environment.

Unlike vehicle localization, pedestrian localization must overcome many scientific and technological gaps such as the latency and reactivity time which must be reduced and the lack of existing map (unlike the road network). To overcome faster and less predictable motions, it must reach a precision better than those of a guided robot or of a vehicle (motion model). Another tough problem is the human step variation. Unlike a wheel, the humans step can vary from one person to another and even for the same person depending on his state and circumstances. This variation can become a significant source of error because it can increase rapidly. For all these previous reasons, take inspiration and adapt the concepts of car navigation to the pedestrian navigation is still a great challenge.

III. PEDESTRIAN LOCALIZATION APPLICATIONS

The Pedestrian localization has many varied fields of application. This section presents the most important fields of pedestrian localization application.

A. *Emergency intervention*

Entities in charge of emergency services are very interested in navigation systems because the navigation data could greatly improve the security and the exploitation of their agents on the site of intervention. To satisfy the expectations of these emergency organizations, several pedestrian positioning systems have been proposed. [1] propose an indoor pedestrian navigation solution for firefighters' emergency intervention. In [2] a real-time pedestrian localization system has been exposed. It ensures the indoor localization of a policeman.

B. *Visually impaired guidance*

Substitute the traditional tools of guidance (the white cane and the guide dog) is still an ambition for the whole VI community. Several researches and systems have been proposed to help VI in their moves especially in unfamiliar environments [3], [4], [5], [6]. Several European projects have been also proposed as CASBLiP project that led to the definition of a visual-hearing translation system. There are also other projects that are still ongoing as HAPTMAP (Haptic, Audio and Visual Interfaces for Maps and Location Based Services) and NAVIG (Navigation Assistée par VIsion artificielle et Gnss) which try to increase the autonomy of the VI.

Given the different services they offer, navigation and guidance tools can make VI life easier: daily walk assistance, obstacle avoidance, instantaneous position and orientation, etc..

C. *Tourism*

Besides its function of guidance, a pedestrian localization system can provide the carrier with required information about a specific place or point of the itinerary. So we can well imagine that such a system can be useful for guided tours in

cities or museums. Current means of mobile telecommunication allow eventual interactive access to further information about the city (points of interest, public transport, etc.) by the geo-referencing of the path on an existing digital map: everytime the tourist goes close to a touristic area, building, sculpture, etc., the system provides him with interesting and useful information.

D. *Military*

The military sector is a field that gives great importance to the research and development. Indeed, there are many inventions and systems that have emerged thanks to the military researches. Then, they were extended for civilian purposes. For example the GPS is originally a research project of the U.S. Army in the 1960s. The localization of its fleet (vehicles, aircraft, etc.) and each one of its soldiers is strategic and it is of primary importance to all armies. This capacity is already reached for vehicles but it is still scheduled as soon as possible for soldiers.

Each one of the above list of pedestrian localization applications needs some requirements to be respected in terms of accuracy, latency time, reactivity, cost, environment, etc. There are different technologies that can satisfy the needs of these applications: GPS, Bluetooth, Ultra Wide Band (UWB), ultrasound, infrared, wifi, Radio Frequency IDentification (RFID), Microelectromechanical systems (MEMS) and the camera. Most of these technologies were already used in vehicle navigation. Because of lack of space we haven't detailed each one of them. But it is very important to know the advantages and the drawbacks of each one of them to make the best choice.

IV. LITERATURE REVIEW

The diversity of application areas and technologies used in pedestrian localization has led to a multiplicity in systems that have been proposed in the literature. There are mainly two approaches: tag network based systems which are dependant from an existing infrastructure or autonomous systems.

A. *Dependant Localization Systems*

This family of localization systems is composed of systems using sensors or tag networks as the satellite networks (GPS) or the existing local networks (WiFi, GSM) or Radio Frequency IDentification tags (RFID tag). Such a localization system is expensive because it requires the installation of specific infrastructure in all regions where we want to ensure the localization with a sufficient accuracy. That is why most of these methods have been tested only in indoor environments.

Drishti is a pedestrian navigation system proposed by a team of young researchers [3] from Florida University at Gainesville, to guide VI during their travels in the campus or inside its buildings. This device combines the Differential GPS (DGPS) with a Geographic Information System (GIS) for outdoor localization [7]. It uses an ultrasonic positioning system for indoor localization: the receiver is composed of 2 tags attached to the user's shoulders while the transmitters were

composed of 4 ultrasound pilots mounted in the four corners of the building to provide the position of the person. The results show that this system is able to ensure indoor localization with an error inferior to 22cm from the real position.

An infrared localization system was proposed by [8]. It is composed of three main units: a vest that contains a 4x4 micromotors grid which delivers haptic guidance signals on the back of the user, a laptop for itinerary planning and infrared receiver and transmitters to localize the person. For a consistent detection of infrared radiation, the transmitters were mounted so they cover the whole trajectory to be traveled and the IR receiver was held in height so that the IR signals can be easily detected. This system was tested by 12 students. They used to travel 4 different trajectories inside the laboratory. The results showed that each path requires an average of 1.5 minutes of travel and the mean of error number for each individual and for each path varies from 0 to 3 errors.

Other researchers have used RFID to create a communication network and then to ensure the pedestrian localization. An indoor navigation approach for the VI was presented by [4], [9], [10]. This system (entitled —RG||) is inspired from the navigation guide dog. It is composed of a Pioneer 2DX robot platform, a navigation toolkit, a RFID receiver and radio-tags for localization. The robotic platform is attached to a leash as a substitute for a guide dog. This platform has 3 wheels and 16 ultrasonic sonars, 8 in the front and 8 in the back. The navigation toolkit includes a laptop that is connected to the microcontroller of the robot via an USB cable. The laptop is also connected to a RFID receiver to localize the robot guide. This prototype was tested by VI persons. Results show that it suffers from some gaps like its slow speed (0.5 m/s) and its movement jerkiness.

B. Autonomous Localization Systems

Unlike the first family, such a localization system does not require any existing infrastructure and it is often based on a portable pedestrian navigation system (PNS). This PNS may be coupled with a database map of areas or buildings. [11] are based on graph theory to build the building map using a link/node model. Then, they calculate the person's position using a Pedestrian Navigation Module developed in the EPFL laboratory (Ecole Polytechnique Federale de Lausanne) [12]. The PNM is composed of a GPS receiver, a digital magnetic compass, a gyroscope, a barometer and embedded DR algorithms (dead reckoning). All sensors are installed in a small box so that the weight of this module does not exceed 400g. So it can be attached to the user's belt. The combination of position data with those of the map is made by a dedicated matching algorithm. Preliminary results were encouraging. However, its use is limited to indoors.

An indoor pedestrian localization system for VI people has recently been proposed by [5]. This system is composed of a foot-mounted odometer (to measure the speed) and a white cane in which are attached two sensors: a 3-axis gyroscope and a laser scanner to estimate precisely the attitude of the cane. Information from the 3 sensors is fused in two steps to estimate the user's position. The first step uses the inertial

measurements (from the 3-axes gyroscope) and the relative orientation measurement (from the laser scanner) to accurately estimate the attitude of the white cane. The second step estimates the position of the person: it combines linear speed measurements of the odometer, a filtered version of the cane's motion estimation and extracted features (corners detected by the scanner laser).

A team from EPFL [1] has also exposed an almost self-deployable pedestrian navigation solution for emergency intervention. It is based on the hybridization of MEMS sensors and RFID tags, to increase accuracy and robustness of the system.

All of the systems listed above have either gaps of precision (GPS, MEMS) or they require expensive installation costs (RFID, IR) or they cannot be used in both indoors outdoors (GPS). To overcome these limitations, the computer vision can be considered as an alternative solution to ensure accurate indoor/outdoor localization using a low-cost autonomous system.

In the literature, there are mainly two approaches to deal with the problem of vision based localization: localization without prior information about the environment (SLAM) or localization based on prior knowledge of the environment.

1) Simultaneous Localization and Mapping (SLAM)

The SLAM or the CML (Concurrent Mapping and Localization) approach was developed for robots or autonomous vehicles navigation in an unfamiliar environment. Progressively, the environment's map is enriched with new features. The localization is based on these features and the position of the former primitives is refined taking into account the new observations.

This approach has been used with sonar [13], [14] or telemeter [15], [16] in indoors and more recently with radars in outdoors [17], [18]. The SLAM approach has also recently been applied to stereovision [19] with SIFT's features for a robot localization. Nevertheless, the development of such an algorithm in real time seems to be difficult. Recently, a real-time monocular SLAM approach, combining a particle filter and an extended Kalman filter (EKF), has been presented in [20], [21] and then improved in [22], [23]. A SLAM algorithm on a camera phone has been also exposed in [24]. These approaches are therefore restricted to limited areas and they cannot be efficient for the pedestrian localization without a reliable motion model. One of the proposed solutions to overcome these drawbacks is the multi-sensor fusion.

2) Pre-recorded Mapping (Visual Memory)

Thanks to this approach, the mapping step, which is the most complex part in terms of computing time, can be processed offline. A partial 3D reconstruction of the environment can be established from a recorded video sequence of the reference path. Once this "visual memory" stored, the localization can be ensured in real time by estimating the camera position (6 DOF). That is why we have

adopted this approach to develop our own pedestrian localization system.

It is also possible to build the 3D map using vision alone or by combining vision with other sensors. For example, [25] use a camera mounted in a rotating platform and a laser telemeter to build a set of panoramic images enriched with depth information. [26] also build a 3D map using a trinocular sensor and an odometer. This map includes the position of vertical lines observed during the learning step. A different mapping approach was proposed by [27]. It uses a camera placed on a vehicle to see the road sides. From these views, a motion's segmentation is performed to differentiate the different buildings facades that are classified according to the distance to the camera.

In this section we have introduced different pedestrian localization systems and methods from the literature. In the next section we will detail our own pedestrian localization approach.

V. MONOCULAR VISION BASED PEDESTRIAN LOCALIZATION METHOD

The approach we have adopted takes advantages of recent advances in artificial vision navigation. Originally, it was proposed in the case of mobile robots: with a body mounted camera and a pre-recorded 3D map, it is possible to obtain instantaneous and accurate localization and orientation of the camera along an itinerary. To establish the environment's 3D map, a video sequence of the reference itinerary is acquired. This video sequence is then automatically off-line processed by a computer. The 3D map uses existing natural singular points in the scene (Harris corner) which are automatically detected in images (angular points on buildings, doors, windows, panels, trees, etc.).

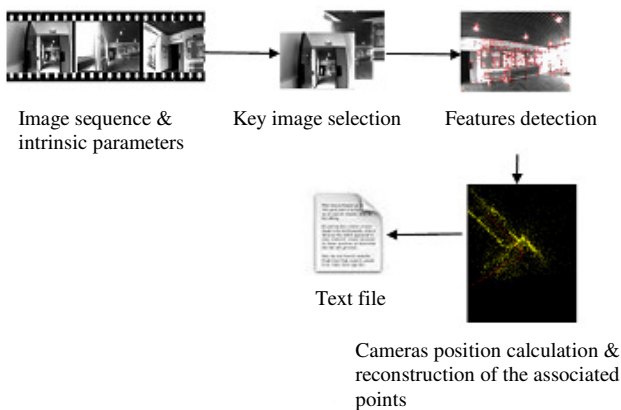


Figure 1: Learning step

This learning step is performed once for each itinerary. The amount of stored data in the visual memory is limited: it includes the positions of some selected key images and the characteristics of 3D features (natural singular points detected in the scene) as shown in Figure 1.

In Figure 2, we can easily observe that the 3D map (the camera positions are represented by cyan squares and reconstructed

3D point cloud is shown in blue) includes not only geometric shapes of the scene but also the traveled path.

In the localization step, when one moves on the same itinerary, the camera position is automatically calculated by matching the points extracted in each new image with the features of the visual memory (environment's 3D map). The figure below shows an example of localization (the position and the current orientation of the camera are shown in magenta).

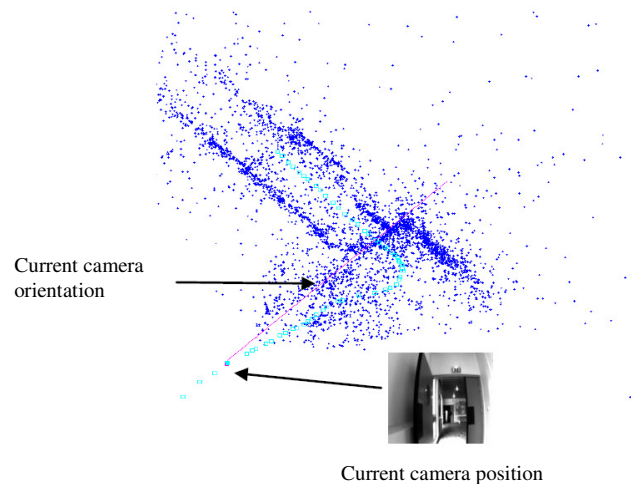


Figure 2: Localization step

This monocular vision based localization seems to have worked in the case of mobile robots. That is why we have decided its transposition to pedestrian navigation. Using an effective guidance interface, this technique should allow the realization of a hand-held and low cost system to significantly increase the mobility and the safety of the VI. The great advantage of such a system is the independence from special infrastructure. It presents serious advantages comparing to current GPS receivers: more accurate, it is effective in indoor/outdoor environments and it is able to calculate the camera position and the movement direction (by calculating the camera position we can retrieve its 6 DOF). Throughout the trip, we have so continuous access to the instantaneous walker position and orientation in relation to the reference itinerary which facilitates its guidance.

VI. RESULTS

Our experimental prototype is composed of a laptop (Intel Core 2 Duo 2.66GHz and 4096MB of RAM) and a 3.5mm lens equipped AVT GUPPY F-033C camera. The camera is previously calibrated to have its intrinsic parameters. The images resolution is 320x240 captured at 25 FPS. Our algorithm is able to process up to 4 fps. This rate is sufficient to ensure the pedestrian localization in real time. Our method has been tested off-line by replaying the same sequence of learning step. The reference itinerary is shown in Figure 2. It is composed of 58 keyframes and 4566 3D points. Off-line localization is perfect (no drift from the reference path). We have also tested the robustness of our algorithm to occlusions

and to scene changes. For this, we have used the "hall_sequence" which is composed of 921 images and captured in the presence of people in the scene. The current camera position and even its orientation are well still onto the reference trajectory (although there is a difference between the two matched images) as shown in Figure 3.

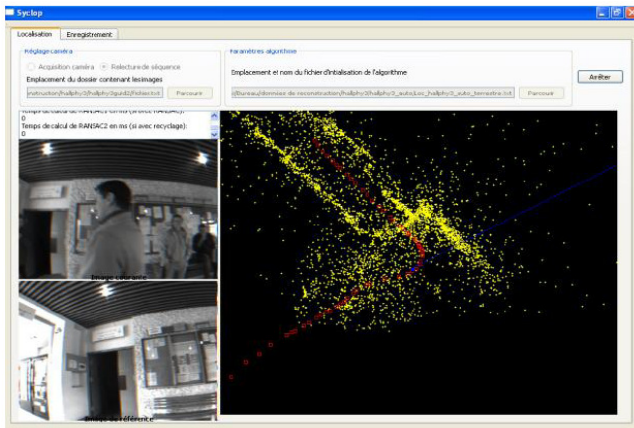


Figure 3: Localization example "hall_sequence"

We have also tested our method on-line. The 3D map of the reference trajectory —lab_sequence|| is shown in Figure 4. It is composed of 37 keyframes and 2291 3D points. During the navigation of the reference trajectory online, 142 images of the video stream were processed. All images have been localized onto the reference trajectory. However, 120 images have been properly localized and 22 images have been localized with a bit drift to the reference trajectory. This is because of the abrupt and rapid motions of the camera which greatly affect the step of features matching. It can be related also to the drifts from the actual trajectory to the reference one (we cannot travel exactly the same path of the learning sequence).

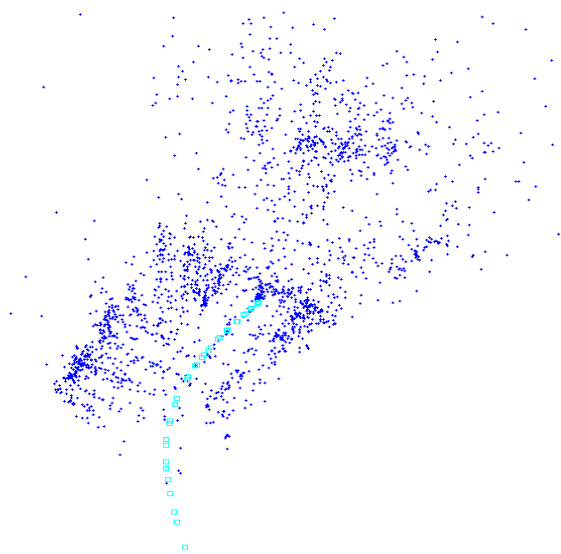


Figure 4: Reference path of the "lab_sequence"

VII. CONCLUSION

We have presented a detailed literature review which has allowed us to highlight the advantages and the drawbacks of the different technologies used for pedestrian localization and hence to select the one that satisfies our purposes.

We have also presented an illustration of our new pedestrian localization prototype. It can be used in indoor/outdoor environments. Moreover, it is interesting since it is independent from any collective equipment. However, the robustness of our system should be improved and this will be investigated in future work.

ACKNOWLEDGMENT

This research program is supported by Heron Technologies (France) and the Conseil Général du Loiret.

REFERENCES

- [1] Renaudin, V., Yalak, O., Tomé, P., and Merminod, B.: Indoor Navigation of Emergency Agents. *European Journal of Navigation*, Vol 5(3), pp. 36–45, (07/2007)
- [2] Gay Bellile, V., Tamaazousti, M., Dupont, R., Naudet Collette, S.: A vision-based hybrid system for real-time accurate localization in an indoor environment. *International Conference on Computer Vision Theory and Applications (VISAPP)*. Angers (2010)
- [3] Ran, L., Helal, S., Moore, S.: "Drishti: An Integrated Indoor/Outdoor Blind Navigation System and Service". *Second IEEE International Conference on Pervasive Computing and Communications (PerCom'04)*, pp. 23. (2004)
- [4] Kulyukin, V., Gharpure, C., Nicholson, J. and Pavithran, S.: —Rfid in robotassisted indoor navigation for the visually impaired. *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Sep. 28-Oct. 2 (2004)
- [5] Hesch, J.A., Roumeliotis, S.I: "An Indoor Localization Aid for the Visually Impaired". *Proc. 2007 IEEE International Conference on Robotics and Automation (ICRA'07)*. (Apr. 2007)
- [6] Treuillet, S., Royer, E.: outdoor/indoor vision-based localization for blind pedestrian navigation assistance. *International Journal of Image and Graphics*. Vol. 10, No 4, pp. 481–496 (2010)
- [7] Helal, A., Moore, S., and Ramachandran, B.: "Drishti: An Integrated Navigation System for Visually Impaired and Disabled". *Proceedings of the 5th International Symposium on Wearable Computer*. October (2001)
- [8] Ertan, S., Lee, C., Willets, A., Tan, H., and Pentland, A.: —A wearable haptic navigation guidance system. *Proc. Second International Symposium on Wearable Computers*, pp. 164—165. (1998).
- [9] Kulyukin, V., Gharpure, C., Sute, P., Graw, N. D. and Nicholson, J.: —A robotic wayfinding system for the visually impaired. *Proceedings of the Sixteenth Innovative Applications of Artificial Intelligence Conference (IAAI-04)*. (2004)
- [10] Kulyukin, V., Sute, P., and Graw, N. D.: —Human- robot interaction in a robotic guide for the visually impaired. *Proc. of the AAAI Spring Symposium on Interaction between Humans and Autonomous Systems over Extended Operation*. March (2004)
- [11] Gilliéron, P.Y., Daniela, Büchel D., Spassov, I., Merminod, B.: Indoor Navigation Performance Analysis. *ENC GNSS*, (2004)
- [12] Ladetto, Q., Merminod, B.: Digital Magnetic Compass and Gyroscope Integration for Pedestrian Navigation. *9th St-Petersburg International Conference on Integrated Navigation Systems*. (2002)
- [13] Leonard, J.J. & Durrant-Whyte, H.F.: Simultaneous map building and localization for an autonomous mobile robot. *IEEE/RSJ International Workshop on Intelligent Robots and Systems*. (1991)
- [14] Wijk, O. & Christensen, H.I.: Triangulation Based Fusion of Sonar Data for Robust Robot Pose Tracking. *IEEE Transactions on Robotics and Automation*, vol. 16, pp 740—752. (2000)

- [15] Moutarlier, P. & Chatila, R.: Incremental free-space modeling from uncertain data by an autonomous mobile robot. *International Workshop on Intelligent Robots and Systems*, pp. 1052—1058. Nov. (1991)
- [16] Thrun, S., Burgard, W. & Fox, D.: A Real-Time Algorithm for Mobile Robot With Applications to Multi-Robot and 3D Mapping. *IEEE International Conference on Robotics and Automation*. (2000)
- [17] Guivant, J. & Nebot, E.: Optimization of the Simultaneous Localization and map Building Algorithm for Real Time Implementation. *IEEE Transactions on Robotics and Automation*, Vol. 17, pp. 242—257. June (2001)
- [18] Dissanayake, G., Newman, P. M., Durrant-Whyte, H-F., Clark, S. & Csorba, M.: A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotic and Automation*, Vol. 17, pp. 229--241, May (2001)
- [19] Se, S., Lowe, D. & Little, J.: Local and Global Localization for Mobile Robots using Visual Landmarks. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 414--420, October (2001)
- [20] Davison, A. J.: Real-Time Simultaneous Localisation and Mapping with a Single Camera, In *9th International Conference on Computer Vision*. pp. 1403--1410, October (2003)
- [21] Davison, A. J., Molton, N. D., Reid, I., and Stasse., O.: MonoSLAM:Real-time single camera SLAM, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 29(6), pp. 1052--1067. (2007)
- [22] Strasdat, H., M. Montiel, J. M., and Davison, A. J.: Real-time monocular SLAM: Why filter?. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. (2010)
- [23] Strasdat, Hauke, J. Montiel, M., and J. Davison, Andrew: Scale Drift-Aware Large Scale Monocular SLAM. *Robotics: Science and Systems*. (2010)
- [24] Klein, G., and Murray, D. W.: Parallel tracking and mapping on a camera phone. *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR)*. (2009)
- [25] Cobzas, D., Zhang, H., and Jagersand, M.: Image-based localization with depth-enhanced image map. *International Conference on Robotics and Automation*. (2003)
- [26] Ohya, A., Miyazaki, Y., and Yuta, S.: Autonomous navigation of mobile robot based on teaching and playback using trinocular vision. *IEEE Industrial Electronics Conference*. (2001)
- [27] Li, S. and Tsuji, S.: Qualitative representation of scenes along a route. *Image and Vision Computing*, Vol. 17, pp. 685—700. -(1999)