ICARCV 2014

The 13th International Conference on Control, Automation, Robotics and Vision

December 10 - 12, 2014
Marina Bay Sands, Singapore

IEEE Catalog Number: CFP14532-USB
ISBN: 978-1-4799-5198-7

Introduction

Session Index

Author Index

Search

© 2014 IEEE. Personal use of this material is permitted. However, permission to reprint / republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.


NANYANG TECHNOLOGICAL UNIVERSITY
ICARCV 2014
The 13th International Conference on Control, Automation, Robotics and Vision

Session Fr33
Image based Modeling; Scene Analysis

Fr33.1 Rejecting Mismatches of Visual Words by Contextual Descriptors
Jinliang YAO, Bing YANG and *Qiuming ZHU
Hangzhou Dianzi University
**University of Nebraska at Omaha*

Fr33.2 A New Calibration Method for Vision System using Differential GPS
Chengping YAN, Lincheng SHEN, Dianle ZHOU, Daibing ZHANG and Zhiwei ZHONG
National University of Defense Technology

Fr33.3 A Vision-based System Supports Mapping Services for Visually Impaired People in Indoor Environments
Quoc-Hung NGUYEN, Hai VU, Thanh-Hai TRAN and *Quang-Hoan NGUYEN
International Research Institute MICA
**Hung Yen University of Technology and Education*

Fr33.4 Depth Estimation from a Single Defocused Image using Multi-scale Kernels
Haoqian WANG, Yushi TIAN, Wei WU and Xingzheng WANG
Graduate School at Shenzhen, Tsinghua University

Fr33.5 KTH-3D-TOTAL: A 3D Dataset for Discovering Spatial Structures for Long-Term Autonomous Learning
Akshaya THIPPUR, Rares AMBRUS, *Gaurav AGRAWAL, Adria Gallart DEL BURGO, *Janardhan HARYADI RAMESH, *Mayank Kumar JHA, *Malepati Bala Siva SAI AKHIL and et al.
KTH Royal Institute of Technology
**M.S. Ramaiah Institute of Technology*

Fr33.6 3D Model-based Curve Generation and Manipulation
Chee Kwang QUAH, Xiang XU and Hock Soon SEAH
Nanyang Technological University

2014 13th International Conference on Control, Automation, Robotics & Vision
Marina Bay Sands, Singapore, 10-12th December 2014 (ICARCV 2014)
Fr33.3

A Vision-based System Supports Mapping Services for Visually Impaired People in Indoor Environments

Quoc-Hung Nguyen, Hai Vu, Thanh-Hai Tran
International Research Institute MICA, Hanoi University of Science and Technology
{quoc-hung.nguyen,thanh-hai.tran,hai.vu}@mica.edu.vn

Quang-Hoan Nguyen
Faculty of Information Technology, Hung Yen University of Technology and Education
quanghoanptit@yahoo.com

***Abstract*— This paper describes and extensively evaluates a visual-based system that autonomously operators for both building a map and localization tasks. The proposed system is to assist mapping services to the visually impaired/blind people in small or mid-scale environments such as inside a building or campus of school, hospital. Toward this end, the proposed approaches solely rely on visual data thanks to a self-designed image acquisition system. On one hand, a robust visual odometry method is utilized to create a map of the environments. On the other hand, the proposed approaches utilize FAB-MAP algorithm that is maybe the most successful for learning places in the environments. Map building and learning places in an environment are processed in an off-line phase. Through a matching place procedure, online captured images are continuously positioned on the map. Furthermore, we utilize a Kalman Filter that combines the matching results of current observation and the estimation of robot states based on its kinematic model. We evaluate performances of the proposed system through experimental schemes. The results show that the constructed map coincides with ground truth, and matching image-to-map is high confidence. The evaluations also contain scenarios which the blind pupils move following Robot. The experimental results confirmed that proposed system feasibly navigating blind pupils in indoor environments.**

***Keywords*— Visual Odometry, Place Recognition, FAB-MAP algorithms, Navigations.**

I. INTRODUCTION

Autonomous localization and navigation are extreme desirable services of peoples who suffer from visual impairment problems. Most of commercial solutions are based on the Global Positioning System (GPS), Wi-Fi, LIDAR, or fusion of them. However, GPS systems provide the services with stick conditions such as good weather, outdoor environments, no presence of buildings. It is highly cost to setup LIDAR systems in environments where mid-scale areas like campus of school, hospital are. Wi-Fi systems are also not easily installing to cover such environments. They thus are not widely accepted by the users [4]. Recent techniques in the computer vision community and high performances of the smart-phones nowadays offer substantial advantages to address these problems. Consequently, it promises alternative solutions to support the mapping services to the visually impaired or blind people. Our previous works [17] suggests a framework presenting such kind of system that is solely utilizing visual sensor data. Particularly, we toward the new technologies to assist the visually impaired/blind people in small/mid-scale environments. In this paper, we extent framework presented in [17] with two aspects: Firstly, we warp the framework into a mobile robot in order to deploy navigation services assisting blind people, who move following the robot. Secondly, we implement extensive evaluations to confirm feasibilities of the proposed system.

The proposed framework in [17] involves in vision-based methods for understanding and representing environments. It aims to answer two questions. The first question is that "what does the world look like?". This question involves in the map building task. In contrast to this, localization service relates to estimating a pose to a relative position on the created map. In other words, it is to answer the second question "Where am I?". The first question is solved through a learning step which supports us building a map as well as assigning scenes into corresponding positions on it. The major advantage is that it is possible to build incremental map. The second question is a process to match image-to-map. To solve these questions, we simultaneously collect visual data for the off-line process by a self-designed imaging acquisition system. For building a map of the environment, we utilize a robust visual odometry proposed in [15]. This is interesting method because it is successful to build trajectory using only one consumer-grade camera. Furthermore, in order to improve quality of the constructed map, we adapt the algorithms in [15] with contexts of the indoor environments. In order to learn places in the environment, we utilize so-called loop closure detections method [3], [13]. The main idea for learning the visited places is that loop constraints can be found by evaluating visual similarity between the current observation and past images where are captured in one (or several) trials. The second phase is an online process. An agent (such as vehicle, human) is required to mount/wear a mobile device camera. The current observation is matched to the place in the database which is learnt in the off-line phase. This matching procedure is similar to place recognition. Recent approaches like FAB-MAP are aimed at reaching a high recall rate at 100% precision. In this work, we employ a robust FAB-MAP [3] that is reliable to recognize known places through autonomous operation in an intelligent system like a mobile robot. FAB-Map 2.0 has been applied to a 1000 km dataset and achieved a recall of 3.1% at 100% precision (14.3% at 90% precision respectively).

The next sessions of the paper are organized as follows: In Section 2, we briefly survey related works. In Section 3, we present summarize the system navigational aids visually impaired people using vision-based for Robot. In section 4, we focus on evaluating the performance of the proposed system: in term of: Localization based on matching image, accuracy of identifying the starting point; How to control navigation services on Robot using Kalman fiter. Finally, we conclude and give some ideas for future works.

II. Related Works

Developing localization and navigation assistance tools for visually impaired people have been received many intention in the autonomous robotics community [4]. Most of them involve in finding out efficient solutions to the positioning data that come from different sensory modalities such as GPS, laser, Radio Frequency IDentification (RFID), vision or the fusion of several of them. Loomis et al. in [11] surveyed efficiency of GPS-based navigation systems supporting visually impaired people. The GPS-based systems share similar problems: low accuracy in urban-environments (localization accuracy is limited to approximately 20 m), signal loss due to multi-path effect or line-of-sight restrictions due to the presence of buildings or even foliage. Kulyukin et al. [9] proposed a system based on Radio Frequency IDentification (RFID) for aiding the navigation of visually impaired people in indoor environments. The system requires the design of a dense network of location identifiers. Helal et al. [8] proposed a wireless pedestrian navigation system. They integrated several signals such as voice, wireless networks, Geographic Information System (GIS) and GPS to provide the visually impaired people an optimized route. Recent advanced techniques in computer vision offer substantial solutions with respect to localization and navigation services in a known or unknown environments. The vision based approaches are safe navigation and provide a very rich and valuable perception information of the environment. Alcantarilla [6] utilizes well-known techniques such as Simultaneous Localization and Mapping (SLAM) and Structure from Motion (SfM) to create 3-D Map of an indoor environment. He then utilizes means of visual descriptors (such as Gauge Speeded Up Robust Features, G-SURF) to mark local coordinate on the constructed 3-D map. Instead of building a prior 3-D map, Lui et al. [10] utilize a pre-captured reference sequence of the environment. Given a new query sequence, their system desires to find the corresponding set of indices in the reference video.

Many specific applications that also are based on vision sensors are developed to support typical daily activities of the visually impaired people. For example, [2] develops an application, names LocateIt, which supports blind people locate objects in the indoor environments. In [16], ShelfScanner is a real-time grocery detection, that allows online detection of items on a shopping list. With regard to map building and localization services, SLAM has been proven to be quite successful in navigation for autonomous robotic systems [1]. By means of visual SLAM techniques, some wearable applications are proposed. Pradeep et al. [14] presents a head-mounted, stereo-vision for detecting obstacles in the path and warn subjects about their presence. They incorporate visual odometry and feature based metric-topological SLAM. Murali et al. in [12] estimate the users location relative to the crosswalks in the current traffic intersection. They develop a vision-based smart-phone system for providing guidance to blind and visually impaired travelers at traffic intersections. The system of Murali et al. in [12] requires supplemental images from Google Map services, therefore it is suitable with travels at outdoor environments only. With SLAM-based approaches, it is possible to build a map at the same time the location of the people who wears cameras standing/moving in the environment. However, the complexity of the map building task varies in function of environment size. In some case, a map can be acquired from visual sensor, but in other cases, the map is such that it must be constructed from other sensor modalities such as GPS, WIFI [3]. Furthermore, matching a current view to a position on the created map seems to be the hardest problem in many works [1], [7]. In our point of view, an incremental map is able to support us improving matching results. Therefore, different from these systems, we create a rich map as good as possible through many trials. When new observations arrive, these new observations must be locally and globally consistent with the previous construction. These problems are able to solve through the loop closure algorithms [3], [13]. However, major different points from [3] are that our proposed system build solely using visual data for both map and localization services, whereas the works in [3] use GPS data for localizing on the map. With our proposed system, creating a travelling route and learnt scenes are implemented in advance. The blind people/visually impaired people then use a frontal camera of a smart-phone device to capture current view for matching image-to-map. Therefore, it is not able to update the map that is major function in the SLAM systems.

III. Summary of the Proposed System

This section briefly describes a frame-work to build navigation system in indoor environments. Different from conventional navigation system, the proposed system utilizes solely visual data, without requires conventional positioning data such as GPS, WI-FI, LIDAR, so on. Details of the proposed system are described in [17]. In major improvement in this paper is that, we warp the proposed frame work in [17] into a mobile robot, where Kalman filter is an important factor to improve accuracy of robot movements.

The proposed frame work is presented in Fig. 1. According to this framework, its operation consists of two phases:

- **Offline learning phase**: Using the collected visual data, this phase creates trajectories and learns the places along the travels. Because scenes and route images are captured concurrently, the constructed map contains learnt places in corresponding positions of the travel.
- **Online localization phase**: The current view is described using a visual dictionary. A probabilistic function attempts to match this data to the database of labeled places obtained during the offline phase. The current observation can then be matched to a corresponding position on the constructed map.

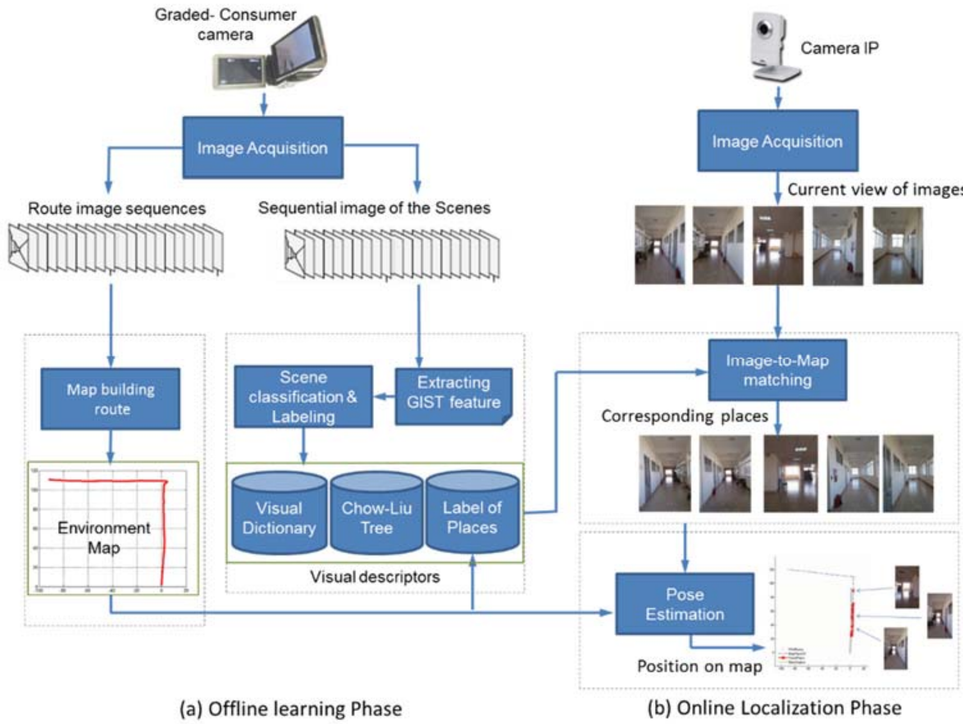

Fig. 1. The framework of the proposed system

The framework utilizes the visual data collected by a self-designed image acquisitions system, as shown in Fig. 2.

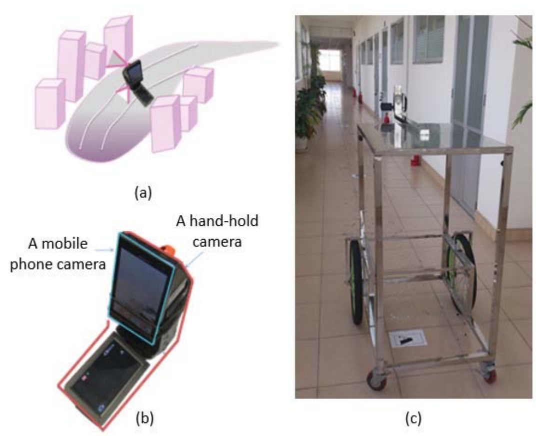

Fig. 2. (a) A schematic view of the visual data collection scheme. (b) The proposed imaging acquisition system in which a mobile phone camera is attached on rear of a hand-hold camera. (c) The image acquisition system attached on a wheel vehicle.

A. *The map building based on visual odometry techniques*

To build route of the travel, we utilize a visual odometry method proposed by Van Hamme et al. [8]. The method is based on the tracking of ground plane features. Particularly, it is designed to take into account the uncertainty on the vehicle motion as well as uncertainty on the extracted features. Well-known issues for visual odometry techniques are that they need to estimate precisely correspondences between the features of consecutive frames. Once the feature correspondences have been established, we can reconstruct the trajectory of the vehicle between the two frames. Due to the floor characteristic of the corridor environment, the number of feature points detected by the original work [15] is quite limited and leads to a very poor reconstruction of the travel. To solve this issue, we manually placed additional markers over the whole journey.

B. *Visual description of the scene*

Visual description of the indoor environments are suffered from repetitive structure and ambiguous scenes. In order to discriminate scenes as well as to find only representative scenes, we filter out similar scenes from collection data. To obtain this, we utilize GIST features [18] and a conventional classifier (e.g., K-nearest neighbor) to extract only representative scenes.

These visual presentations need to be easy implementation and efficient distinguishing scenes. To adapt with these issues, we involve the FAB-MAP algorithms [3] which are recently successful for matching places in routes over long period time. The FAB-MAP is a probabilistic appearance-based approach to place recognition. Each time the image taken, its visual descriptors are detected and extracted. In our system, we utilize SURF extractors and descriptors for creating on a visual vocabulary dictionary. A Chow Liu tree is used to approximate the probability distribution over these visual words and the correlations between them. The FAB-MAP involves co-occur visual word of same subject in the worlds.

C. *Matching image-to-map procedure*

Given a current view, its position on the map is identified through a place recognition procedure. We evaluate the current observation at location L_i on the map by its probability when given all observations up to a location k:

$$\rho(L_i|Z^k) = \frac{\rho(Z_k|L_i)\rho(L_i|Z^{k-1})}{\rho(Z_k|Z^{k-1})} \quad (1)$$

Where Z_k contains visual words appearing in all observations up to *k-1*; and Z^k presents visual words at current location k. These visual words are defined in the learning places phase. A probability $p(Z_k|L_i)$ infers observation likelihood as learnt in the training data. In our system, a L_i is matched at a place $k*$ when $argmax(p(Z_k|L_i))$ is large enough (through a pre-determined threshold T = 0.9).

D. The Kalman Filter (KF)

Utilizing Kalman Filter is an extension from our original framework in [17]. In the context of assistance service to blind people using mobile robot, Kalman Filter help to combines the matching results of current observation and the estimation of robot states based on its kinematic model. The observations of the robot are images captured over time, which are then converted to coordinates (x, y, z) in a predefined coordinate system using above matching procedure. However, in indoor environment, the scene does not always change significantly. Consecutive scenes could repeat when the robot moves. Therefore, the performance of image matching is not good. Sometimes, a current observation could be matched with a very far forward / backward image that makes incorrect localization of the robot. To overcome this problem, we propose to use a Kalman filter to correct the position of the robot from observation. Effectiveness of using Kalman filter is presented in the experimental results.

IV. Experimental Results

A. Evaluation environments

We evaluate the proposed method in a corridor environment of a building, where is 10^{th} floor of International Research Institute MICA-Hanoi University of Science and Technology (HUST) (Fig. 3).

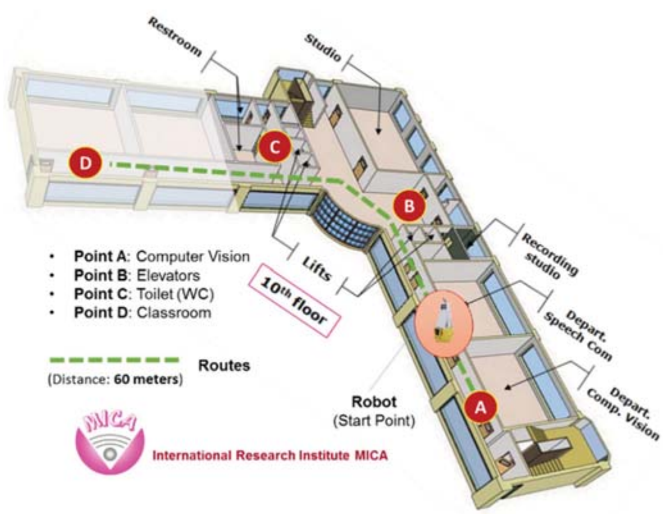

Fig. 3. A 3-D map of the evaluation environment

In this section, we report performance of three experiments, which affect to quality of the navigation services. They are: (1). Performance of the image matching-to-map. This evaluation reports localization accuracy of the proposed method. (2). Performance of the identifying the starting point. This evaluation impacts to navigating blind people services. In scheme of the navigating services, knowing the starting point and ending point, the direction and how to travel are able to predicted (e.g. by shortest path algorithms). Whereas, the ending point is given by user's request, the starting point should be automatically identified by system. (3). Effectiveness of using Kalman filter in controlling robot. This evaluation is to confirm that proposed system is feasible to face practical issues in movements of robot.

B. Experimental results

1) Evaluate image matching-to-map procedure

We collect data in 4 times (each time equals to one trial). To build the visual dictionary in offline phase, we have used images collected from L#1 trial. By experience, we set the size of the dictionary to 1300. We then use the images collected from L#4 trial to learn places along the travel. In total, K = 140 places have been learnt. The visual dictionary and descriptors of these places are stored in XML files. The collected images in L#2 and L#3 travels are utilized for the evaluations.

TABLE I. Result Of The Matching Places (FAB-MAP Algorithms) without and with Scene discriminations

Travels	Without scene discrimination		With scene discrimination	
	Precision	*Recall*	*Precision*	*Recall*
L#2	12%	90%	67%	82%
L#3	36%	85%	**74%**	**88 %**

The TABLE I. shows the precision and recall with L#2 and L#3 travels with/without scene discrimination step. For learning places (using original FAB-MAP, without scene discrimination), the recall of L#3 travel is clearly higher than L#2. The main reason is that some "new" places which were not learnt from L#4 are able to update after L#2 running. Therefore, more "found" places are ensured with L#3 travel. TABLE I. also shows the efficiency of the scene discriminations step, the performances of image-to-map matching obviously increasing and stable for precisions measurement with scene discrimination step, whereas high confidence of the recalls is still consistent.

2) Evaluate the localization of starting point

As mentioned previously, we will use a robot to help blind person navigating in the environment from any position on the map to a predefined destination point. To this aim, we have to localize the robot at the current time. In this section, we will examine if the proposed localization method is good enough to identify starting position of the robot.

Following round-trip along the corridor of total length about 60 m, every 3 m, we take a point and consider it as the starting point of the robot. This procedure is repeated at 3 different times (Testing # 3- morning, Testing # 2- afternoon, Testing #1- evening). For each testing time, the total number of sampled point is 46. To identify the position, we take 100 consecutive images then apply the image matching as presented in section III.C. We determine the most repeatable matched location when its reputation is large enough (larger than 70% in our experiment). The robot is then considered as being localized at that position.

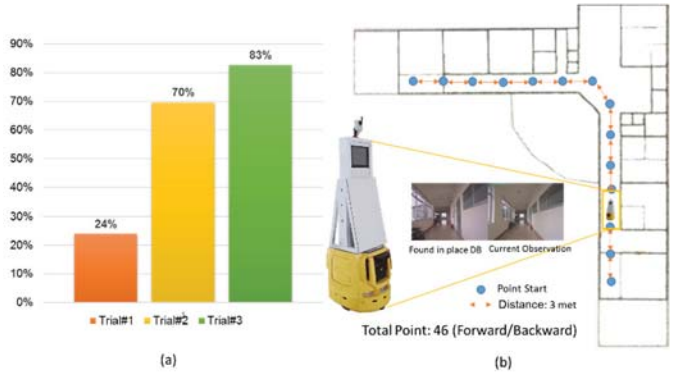

Fig. 4. The performance of localization of starting point of Robot. (a) Evaluation chart prediction starting point. (b) The starting point in the environment.

The Fig. 4 shows the performances with 3 trials in term of recall. The best result among 3 trials is 83%. These results show that environment conditions have a relatively big impact to the experimental results. They also suggest us for further improvement so that the proposed algorithms are solid to the lighting conditions

3) Evaluate the role of Kalman filter

Given the starting point of the robot and the destination point, based on the environment information, we could determine the shortest path to follow. However, when we make the robot moves, dues to the mechanical errors, the robot cannot not attend exactly the positions provided by drive/control module. The Fig. 5 shows the real positions of the robot that drifts away the ground truth ones.

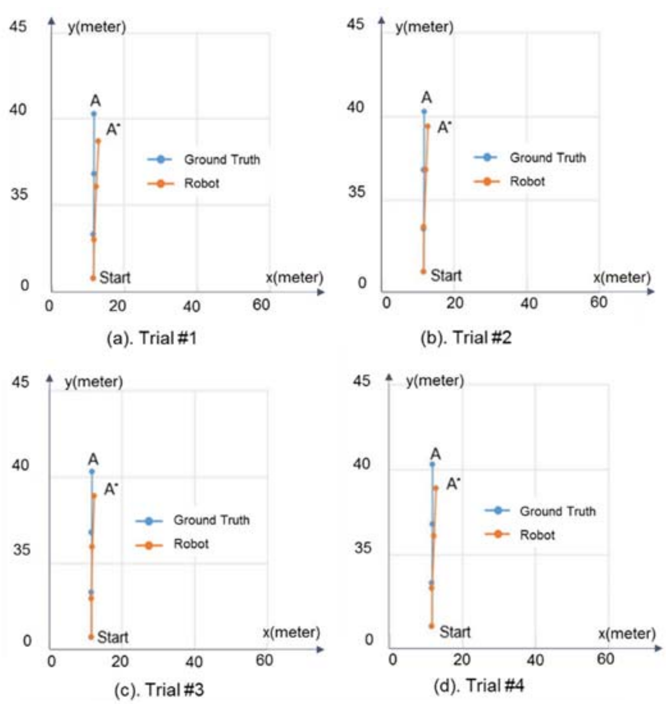

Fig. 5. Comparision of real positions of the robot to the groundtruth ones

Because of these reasons, the robot cannot reach the desired position. We propose to correct the drift with the use of image matching based localization in combination with Kalman filter (see section III.D). We conducted four experiments in each the robot moves follows a straight road of length 30 m at velocity about 200 mm / second in the morning. To show the effectiveness of Kalman filter, Fig. 6 demonstrates navigation data without and with using Kalman filter.

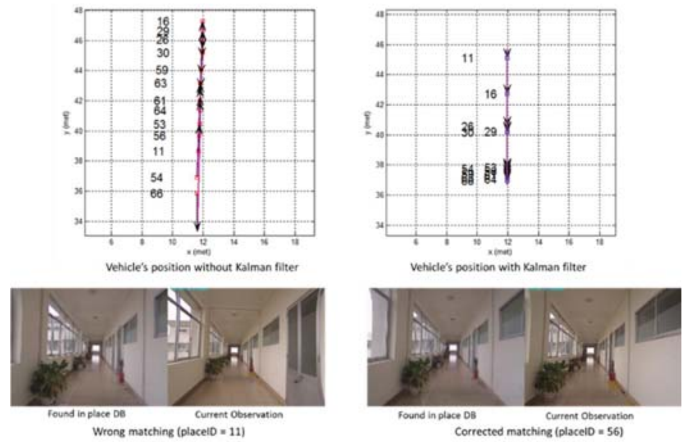

Fig. 6. Vehicle moving without/with Kalman Filter. Top row: Left panel: vehicle positios on the map using only results of the matching image-to-map procedures. The arrows show directions to guide vehicle. Numbers on left of each red box show *placeID* of the current observation. Right panel: positions of the vehicle are updated using Kalman filter. Bottom row: Left panel: This result shows wrong direction to vehicle. Right panel: is a good matching with Kalman filter.

Using only the place recognition results, the directions supporting navigation services are obviously uncontrolled. Some matching places (show in numbers) are misses and in the wrong order in this case. The main reason is the erroneous matching of some places (e.g., place ID = 11, shown in bottom panel). By using a Kalman Filter, directions supporting navigation services are correctly ordered.

To evaluate the localization accuracy in case of use / nonuse of Kalman filter, we measure median and average error of estimated position with respect to the ground truth one. The evaluation results are presented in the table below:

TABLE II. MEDIAN ERROR (IN METER)

Method	L#1	L#2	L#3	L#4	Average
Vision based localization with Kalman filter	0.6	0.4	0.8	1.3	**0.8**
Only use the shortest path	0.6	0.8	0.9	1.1	**0.9**

TABLE III. AVERAGE ERROR (IN METER)

Step	L#1		L#2		L#3		L#4		Average	
	avg	*std*	*avg*	*std*	*avg*	*std*	*Avg*	*std*	*avg*	*std*
Matching	0.6	0.4	0.6	0.6	1.0	1.0	1.3	0.7	**0.9**	**0.7**
NoMatching	0.6	0.5	0.7	0.4	1.6	2.0	2.3	3.2	**1.4**	**1.6**

The best accuracy is obtained with the 2nd trial ($\Delta \sim 0.4$ meter). We investigate in more detail the deviation at each position on the trajectory. Matching and NoMatching (L#2, L#3) as shown Fig. 7.

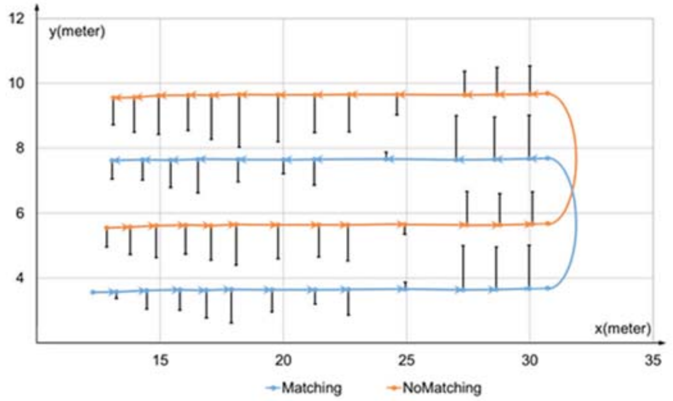

Fig. 7. Compare the results of locate error between Matching and NoMatching

Consequently, if we do not utilize the Kalman Filter and imaging matching, the error rate is Δ ~ 1.2 m high Delta; whereas, by using Kalman filter, the localization error rate significantly reduce to Δ ~ 0.4.

4) Nagivational aids to visually impaired people in Small Pervasive Environments using robot

To show the feasibility of using robot to help blind people navigating in environment, we have asked a blind pupil to follow the robot. The human robot interaction is carried out though the mobile phone.

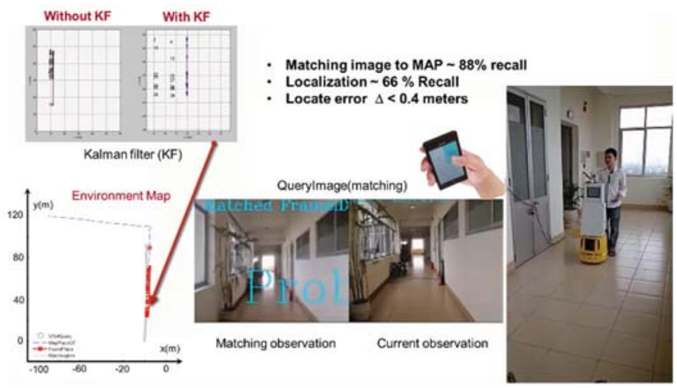

Fig. 8. A blind pupil move following Robot. The places on his travel are marked on the constructed map in red-rectangles

According to guidance by vibration sensor mobile smartphone, he could go whole travel in the corridor environment. We obtain average performances of the Matching ***~88% recall, Localization ~ 66 % Recall***, Locate error Δ ***< 0.4 meters.*** These results are not so far from as shown Fig. 8. In other words, we could match precisely positions the blind people using the constructed map. These results are feasible to deploy automatic navigating system.

V. Conclusions

In this paper, we presented a vision-based system for both autonomously map building and localizing services. We successfully created the map of the indoor environment using the visual odometry and learning places. The results of matching image-to-map are high confidence. Therefore, the proposed system is able to provide us deploying navigating services in the indoor environments. The proposed system directs to support visually impaired peoples in Small Pervasive Environments. Further in-the-loop evaluations with the visually impaired/blind people will direct us to future work.

Acknowledgment

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number **FWO.102.2013.08**

References

[1] Bailey T. and Durrant-Whyte H. (2006), "Simultaneous localization and mapping (SLAM): Part II," *IEEE Robotics & Automation Magazine*, 13, 108-117.

[2] Bigham J. P., Jayant C., Miller A., White B. and Yeh T.(2010), "VizWiz:: LocateIt-enabling blind people to locate objects in their environment," in *Computer Vision and Pattern Recognition Workshops (CVPRW),* 65-72.

[3] Cummins M. and Newman P. (2008), "FAB-MAP: Probabilistic localization and mapping in the space of appearance," *The International Journal of Robotics Research*, 27, 647-665.

[4] Dakopoulos D. and Bourbakis N. G. (2010), "Wearable obstacle avoidance electronic travel aids for blind: a survey," *Systems, Man, and Cybernetics, Part C: Applications and Reviews.*

[5] Everingham M., Van Gool L., Williams C., Winn J. and Zisserman A. (2009), "The PASCAL Visual Object Classes (VOC) challenge,"

[6] Fernández Alcantarilla P., "Vision based localization: from humanoid robots to visually impaired people," (2011), Electronics, University of Alcala, Ph.D. Thesis

[7] Fraundorfer F. and Scaramuzza D. (2012), "Visual odometry: Part II: Matching, robustness, optimization, and applications," *Robotics & Automation Magazine, IEEE*, 19, 78-90.

[8] Helal A., Moore S. E. and Ramachandran B. (2001), "Drishti: An integrated navigation system for visually impaired and disabled," in *Wearable Computers, Proceedings.* 149-156.

[9] Kulyukin V., Gharpure C., Nicholson J. and Pavithran S., "RFID in robot-assisted indoor navigation for the visually impaired," in *Intelligent Robots and Systems, (2004). Proceedings. 2004 IEEE/RSJ International Conference on IROS,* 1979-1984.

[10] Liu J. J., Phillips C. and Daniilidis K., "Video-based localization without 3D mapping for the visually impaired (2010)," in *Computer Vision and Pattern Recognition Workshops (CVPRW),* 23-30.

[11] Loomis J. M., Golledge R. D. and Klatzky R. L. (2001), "GPS-based navigation systems for the visually impaired,"

[12] Murali V. N. and Coughlan J. M. (2013), "Smartphone-based crosswalk detection and localization for visually impaired pedestrians," in *Multimedia and Expo Workshops (ICMEW),* 1-7.

[13] Newman P. and Ho K, (2005), "SLAM-loop closing with visually salient features," in *Robotics and Automation ICRA, Proceedings of IEEE International Conference*, 635-642.

[14] Pradeep V., Medioni G. and Weiland J. (2010), "Robot vision for the visually impaired," in *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 15-22.

[15] Van Hamme D., Veelaert P. and Philips W., (2011), Robust visual odometry using uncertainty models, In: Advances Concepts for Intelligent Vision Systems. Springer, 1-12.

[16] Winlock T., Christiansen E. and Belongie S. (2010), "Toward real-time grocery detection for the visually impaired," in *Computer Vision and Pattern Recognition Workshops (CVPRW),* 49-56.

[17] Quoc-Hung Nguyen et al.(2014), "Mapping Services in Indoor Enviroments based on Image Sequences", in the proceeding of *the 5th International Conference on Communications and Electronics (ICCE), Danang, Vietnam, July 2014*

[18] Aude Oliva, Antonio Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope", IJCV Vol. 42(3): 145-175, 2001