Advances in Intelligent Systems and Computing 855

Raquel Fuentetaja Pizán
Ángel García Olaya
Maria Paz Sesmero Lorente
Jose Antonio Iglesias Martínez
Agapito Ledezma Espino *Editors*

Advances in Physical Agents

Proceedings of the 19th International Workshop of Physical Agents (WAF 2018), November 22–23, 2018, Madrid, Spain

Springer

Advances in Intelligent Systems and Computing

Volume 855

Series editor

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

The series “Advances in Intelligent Systems and Computing” contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within “Advances in Intelligent Systems and Computing” are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

More information about this series at http://www.springer.com/series/11156

Raquel Fuentetaja Pizán
Ángel García Olaya · Maria Paz Sesmero Lorente
Jose Antonio Iglesias Martínez
Agapito Ledezma Espino
Editors

Advances in Physical Agents

Proceedings of the 19th International Workshop of Physical Agents (WAF 2018), November 22–23, 2018, Madrid, Spain

Springer

Editors
Raquel Fuentetaja Pizán
Computer Science Department
Universidad Carlos III de Madrid
Leganés, Madrid, Spain

Jose Antonio Iglesias Martínez
Computer Science Department
Universidad Carlos III de Madrid
Leganés, Madrid, Spain

Ángel García Olaya
Computer Science Department
Universidad Carlos III de Madrid
Leganés, Madrid, Spain

Agapito Ledezma Espino
Computer Science Department
Universidad Carlos III de Madrid
Leganés, Madrid, Spain

Maria Paz Sesmero Lorente
Computer Science Department
Universidad Carlos III de Madrid
Leganés, Madrid, Spain

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-99884-8 ISBN 978-3-319-99885-5 (eBook)
https://doi.org/10.1007/978-3-319-99885-5

Library of Congress Control Number: 2018961590

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Physical agents are likely to change the society in the next years. Despite their inherent risks, many times exaggerated by apocalyptic and catastrophic views, it is expected that they will revolutionize our daily life and will have a positive impact in many different fields, from work to health or leisure. The possibilities for research and application are limitless, and the unstoppable "robotics revolution" will increase well-being and open a myriad of new opportunities and challenges.

The International Conference Workshop of Physical Agents (WAF 2018) is a forum for information and experiences exchange in different areas regarding the concept of *agent* on physical environments, especially applied to the control and coordination of autonomous systems: robots, mobile robots, industrial processes, or complex systems. The authors cover several topics in different areas such as software agents, multiagent systems, robotic manipulators, RoboCup and soccer robots, autonomous and semiautonomous robots, machine learning and robotics, industrial robotics, computer vision and robotics, artificial vision and robotics, and artificial intelligence and robotics.

The nineteenth edition of the conference has been organized at Madrid in November 2018 by the Department of Computer Science of the Universidad Carlos III de Madrid, Spain, and the Spanish Physical Agents Network (*Red de Agentes Físicos*), and technically sponsored by Robotics journal and Springer.

This volume collects 22 papers (73 authors) accepted and presented at the conference. The 73 authors from 5 different countries confirm the international status of the event.

This conference will provide a friendly atmosphere and will be a leading international forum focusing on discussing problems, research, results, and future directions in the area of *physical agents*.

WAF 2018 has been possible thanks to the work of many people. We would like to thank the authors and reviewers. Thanks to the Universidad Carlos III de Madrid for letting us use their facilities for the conference sessions. Thanks for the hard work and dedication of the program and organizing committee members. And special thanks to our editor, Springer, that is in charge of this Conference Proceedings edition. Thank you.

September 2018

Angel Garcia-Olaya
Raquel Fuentetaja
Jose Antonio Iglesias
Agapito Ledezma
M. Paz Sesmero Lorente

Organization

Program Committee

Eugenio Aguirre	University of Granada, Spain
José María Armingol	Universidad Carlos III de Madrid, Spain
Ivan Armuelles Voinov	University of Panama, Panamá
George Azzopardi	University of Groningen, Netherlands
Pilar Bachiller	Universidad de Extremadura, Spain
Antonio Bandera	University of Malaga, Spain
Juan Pedro Bandera Rubio	University of Malaga, Spain
Rafael Barea	University of Alcalá, Spain
Luis M. Bergasa	University of Alcalá, Spain
Alberto Bugarín	University of Santiago de Compostela, Spain
Pablo Bustos	Universidad de Extremadura, Spain
Luis V. Calderita Estévez	Universidad de Extremadura, Spain
Oscar Camacho	Escuela Politécnica Nacional, Ecuador
Miguel Cazorla	Universidad de Alicante, Spain
José María Cañas Plaza	Universidad Rey Juan Carlos, Spain
Juan Carlos Corrales	Universidad del Cauca, Colombia
Fernando Fernandez	Universidad Carlos III de Madrid, Spain
Antonio Fernández-Caballero	Universidad de Castilla-La Mancha, Spain
Miguel Angel Garcia	Universidad Autónoma de Madrid, Spain
Miguel Garcia-Silvente	University of Granada, Spain
Ismael Garcia-Varea	Universidad de Castilla-La Mancha, Spain
Antonio González	University of Granada, Spain
Victor González-Castro	Universidad de León, Spain
Jose Daniel Hernández Sosa	Universidad de las Palmas de Gran Canaria
Roberto Iglesias Rodriguez	Universidad de Santiago de Compostela, Spain
Franziska Kirstein	Blue Ocean Robotics, Denmark
James Little	The University of British Columbia, Canadá

Joaquin Lopez	University of Vigo, Spain
Rebeca Marfil	University of Malaga, Spain
Francisco Martín Rico	Universidad Rey Juan Carlos, Spain
Humberto Martínez Barberá	Universidad de Murcia, Spain
Vicente Matellán	University of León, Spain
Rashid Mehmood	King Abdulaziz University, Saudi Arabia
Manuel Mucientes	Universidade de Santiago de Compostela, Spain
Rafael Muñoz-Salinas	University of Cordoba, Spain
Pedro Núñez	Universidad de Extremadura, Spain
Carlos V. Regueiro	Universidade da Coruña, Spain
Lluís Ribas-Xirgo	Universitat Autònoma de Barcelona, Spain
Humberto Rodríguez	Universidad Tecnológica de Panamá, Panamá
Adrián Romero-Garcés	University of Malaga, Spain
Cristina Romero-González	Universidad de Castilla-La Mancha, Spain
Andrés Rosales	Escuela Politécnica Nacional, Ecuador
Gustavo Scaglia	Universidad Nacional de San Juan, Argentina
Jose Soler	Universitat Politècnica de València, Spain
Nicola Strisciuglio	University of Groningen, Netherlands
Cristina Urdiales	University of Malaga, Spain
Diego Viejo	Universidad de Alicante, Spain
Andrés Vivas	Universidad del Cauca, Colombia
Dimitri Voilmy	Université de Technologie de Troyes, France
Eduardo Zalama	Universidad de Valladolid, Spain

Additional Reviewers

Catalá, Alejandro
Gómez-Donoso, Francisco

Contents

Computer Vision and Robotics

Human Robot Interaction

Mobile Robots

AI and Robotics

Semantic Localization of a Robot in a Real Home

Edmanuel Cruz[1,2(✉)], Zuria Bauer[1], José Carlos Rangel[2], Miguel Cazorla[1], and Francisco Gomez-Donoso[1]

[1] Institute for Computer Research, University of Alicante, P.O. Box 99, 03080 Alicante, Spain
edcruz@dccia.ua.es

[2] RobotSIS, Universidad Tecnológica de Panamá, Santiago, Panama

Abstract. In social robotics, it is important that a mobile robot knows where it is because it provides a starting point for other activities such as moving from one room to another. As a contribution to solving this problem in the field of the semantic location of the mobile robot, we pro- pose to implement a methodology of recognition and scene learning in a real domestic environment. For this purpose, we used images from five different residences to create a dataset with which the base model was trained. The effectiveness of the implemented base model is evaluated in different scenarios. When the accuracy of the site identification decreases, the user provides feedback to the robot so that it can process the information collected from the new environment and re-identify the current location. The results obtained reinforce the need to acquire more knowledge when the environment is not recognizable by the pre-trained model.

Keywords: Robotics · Deep learning · Semantic localization
CNN training · Neural networks

1 Introduction

It is becoming increasingly common to design and create robots with the ability to interact with humans, whether it be caring for the disabled or older adults, or as assistants in shopping malls or receptionists. Social robotics is a research lines that allows these robots to be easily integrated into their environments. In this context, the semantic localization of mobile robots plays a crucial role.

Changing between different environments is currently a challenging task because the systems are unable to adapt to a continually changing environment. The usual scenario is a system that loses accuracy when the changes become more drastic. According to the deployment place of the robot the same semantic categories may also be visually different.

In the same way as a human who first analyzes where he or she is and then decides where to go or what to do, it is important that the robot recognizes where it is in order to perform other activities.

R. Fuentetaja Pizán et al. (Eds.): WAF 2018, AISC 855, pp. 3–15, 2019.
https://doi.org/10.1007/978-3-319-99885-5_1

In this work, we propose to implement a methodology of recognition and scene learning in a real domestic environment. The steps carried out are: (1) taking images from five residences; (2) creating a dataset with images from one residence; (3) training a base model; and (4) experimenting with images from the other residences.

The rest of the paper is organized as follows: Sect. 2 presents the state-of-the-art in the field. Then, Sect. 3 details a description of the proposal. This is followed by Sect. 4, where the procedures for testing the proposed approach are described. Next, Sect. 5 details the results of experiments. Finally, Sect. 6 includes the conclusions and future works.

2 Related Works

In recent decades, many researchers have conducted investigations related to semantic localization. We start with works like [1], where the authors train a neural network to estimate the location of a mobile robot in its environment using the odometry information and ultrasound data. The authors in [2] use a pre-programmed routine to detect doorways from range data. In [3], a system was developed with the ability to learn to use a hybrid methodology based on human demonstrations and user advice. Two years later, the authors in [4] describe a virtual sensor that is able to identify rooms from range data. The same year, in [5], the authors apply different learning algorithms to learn topological maps of the environment.

In [6], the authors proposed to use line features to detect corridors and doorways. Furthermore, [7] merges data from local and global features of an image with data from laser sensors. The authors then predict the category of a place by training a Support Vector Machine (SVM) system.

According to [8], many mobile service robots operate in close interaction with humans. They present an approach to people awareness for mobile service robots that utilizes knowledge about the semantics of the environment.

The use of semantic labels was considered in the proposal of [9]. Here, the authors manually insert semantic labels into a 2D image and complement this representation with 3D points.

Before continuing, it is important to mention the tools that have helped the development of semantic localization research. The appearance of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [10] allowed algorithms to be evaluated for large-scale object detection and image classification. Researchers were able to compare progress in detection across a wider variety of objects taking advantage of the expensive labeling effort. The work put forward in [11], introduces a scene-centric database called Places with over 7 million labeled pictures of scenes. The authors propose new methods to compare the density and diversity of image datasets and show that the Places dataset is as dense as other scene datasets and has greater diversity. Using CNN, they learn deep features for scene recognition tasks.

Continuing with our review of previous works, developing a robot with a grounded spatial vocabulary is the proposal in [12]. The authors propose a CNN architecture based on engineered features. Such a vocabulary would allow it to give and follow directions and would give it valuable additional information in aiding localization and navigation.

In [13] the authors describe the problem of location as follows. The problem of semantic localization in social robotics could be defined as the identification of the location of a robot by semantic categories representing a place. The traditional approach for solving semantic localization problems is the utilization of semantic categories such as living room or kitchen, together with the robots perceptions as input data for a supervised classification process.

The survey [14] analyzes the different approaches used to address the localization problem. The problem of recognition of changes in environments is also presented in this work. Robust visual localization under a wide range of viewing conditions is a fundamental problem in computer vision. Dealing with the difficulties of this problem is not only highly challenging but also of significant practical importance, e.g., in the context of life-long localization for augmented reality or autonomous robots.

The authors in [15] propose a novel approach based on a joint 3D geometric and semantic understanding of the world, enabling it to succeed under conditions where previous approaches failed.

In [16], the authors propose a method for semantically parsing the 3D point cloud of an entire building using a hierarchical approach.

The authors in [17] provide an overview of indoor localization technologies, popular models for extracting semantics from location data, approaches for associating semantic information and location data, and applications that may be enabled with location semantics. Environment representation for scene classification could be produced by using different kinds of descriptors such as 3D and 2D.

In [18], the authors propose to develop a study for building robust scene descriptors based on the combination of visual and depth data. The approach was tested for classification problems.

The authors in [19] combine semantic web-mining and situated robot perception to develop a system capable of assigning semantic categories to regions of the space.

In [20], the main idea is to leverage the semantic information provided by the user activities and the accurate metric map created by an assistive robot. In [21], the authors proposal includes the evaluation of several CNN classification models in order to find the one that produces the most accurate classification results.

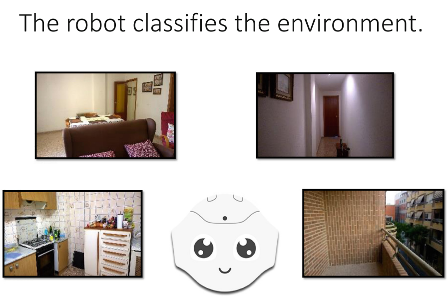
The authors in [22] propose a probabilistic framework that combines human activity sensor data generated by smart wearables with low level localization data generated by robots.

In [23], the authors explore different retraining strategies and experimentation in order to obtain insight about which method provides better precision-training time trade-off. Different settings on the training data are presented, and modifications to different fine-tuning strategies are also explored.

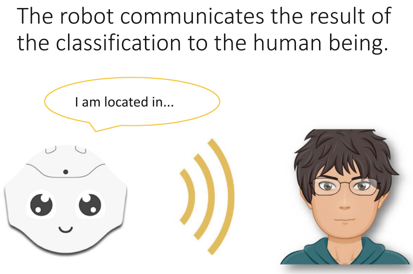
Semantic classification is currently an exciting topic with a great number of published works. In this research, we only focus on current techniques for place categorization that take advantage of DL in order to provide a semantic definition for a place.

3 Proposal

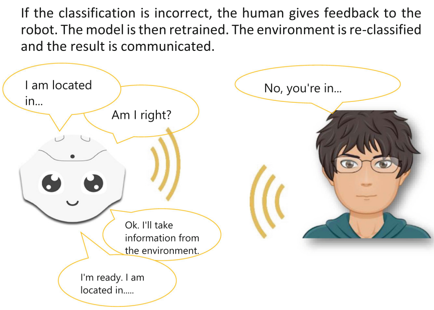
The aim of this study is to implement a methodology to achieve a solid and long-term understanding of the interior scene in changing scenarios. We focus on expanding the work in [23], which presented an optimal methodology for a robot to learn a new environment, from already acquired knowledge. The abovementioned methodology was carried out in a laboratory while the aim of the present work is tested in a real domestic environment.



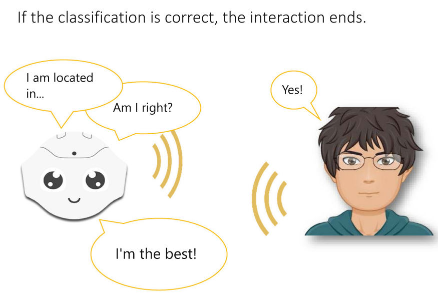
(a) Environment classification



(b) Communication of results



(c) If the classification fails



(d) If the classification is fine

Fig. 1. System interaction

The scenario we wish to present is as follows: first, the robot captures images of the environment and tries to classify them using a previously trained model with images belonging to another place. When the robot is in a new environment, the system is expected to obtain low accuracy due to the differences in the visual features of the new environment. In this case, the user can provide information to the robot to collect data and re-identify the location. If the category provided by the user is not pre-defined, this will be added as a new category, thus allowing the robot to increase its knowledge of the environment. This scenario is shown in Fig. 1.

To complete this goal, we use the neural network architecture shown in Fig. 2. This works as follows: an input image is forwarded to the ResLoc CNN architecture. In this case, we removed the last fully connected layer in order to obtain the visual features descriptor for the input image. As a result, the output of ResLoc CNN is a 2048-dimension feature vector.

The visual features and the respective categories of each image of the dataset are extracted using the ResLoc CNN part of the architecture and inserted into the features database. This feature database is a model that stores the learned data, features of the training samples, and is used during the inference stage.

In the inference stage, the unknown image is forwarded to ResLoc CNN in order to extract the visual feature vector. A K-Nearest Neighbors (KNN) classifier then performs a query on the feature database using the recently computed feature vector. Next, a polling is carried out among the categories of the neighbors, and the most voted category is returned as the final classification of the unknown image.

The performance of the KNN is highly dependent on the k parameter (number of neighbors). Experimentation is carried out using the value of k, which is 3 as it appears in [23]. We used the Annoy[1] implementation of the (approximate) KNN classifier.

4 Experimentation

4.1 Dataset

The experiments described in this work were carried out using our own dataset that provides a semantic category for each RGB image. It is important to clarify that the base model was built with images from only one residence which is identified in the document as House 01. The categories come from the location in which the images were acquired.

Figure 3 shows representative images for the 7 categories available in the dataset.

In order to train the base model we took video sequences from House 01, them randomly shuffled and split them into a 70% training and 30% test ratio. We use only RGB frames.

Table 1 shows the final number of samples per category.

[1] https://github.com/spotify/annoy.

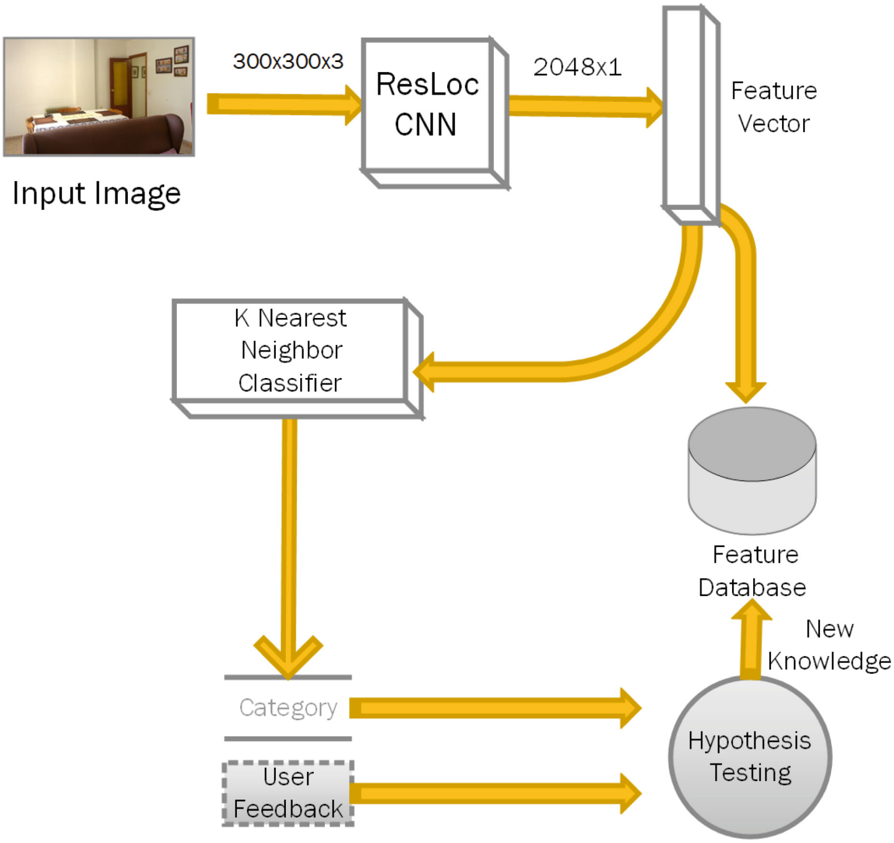


Fig. 2. This Architecture uses the features of a ResLoc CNN with a vector of 2048 features as output. The training samples are forwarded to the ResLoc CNN in order to extract their feature vector. The feature vectors construct the model of a KNN classifier.

Table 1. Images distribution per category.

Cat. ID	Category	Training	Test
1	Corridor	3,782	1,622
2	Dining-living room	4,084	1,751
3	Balcony	1,121	481
4	Kitchen	2,664	1,143
5	Laundry	1,714	735
6	Bathroom	3,932	1,686
7	Bedroom	4,868	2,087

Fig. 3. Sample images for each category of our Home Dataset.

4.2 Experimentation Setup

The experiments were carried out as follows:

- Experiments 1 to 2 consist of using a trained model with data from the House 01 in House 02.
- Experiments 3 to 8 consist of using a trained model with data from House 01 in House 03.
- Experiments 9 and 10 consist of using a trained model with data from House 01 in House 04.
- Experiments 11 and 16 consist of using a trained model with data from House 01 in the House 05.

We simulated scenarios in which the robot was incorrectly located in different environments and we obtained feedback from users to correct this knowledge.

For experiments in which new knowledge was included, we used images that were captured in different houses. In these houses, we have the same semantic categories but different visual appearance. The robot then proceeded to capture new information about the environment that the system had failed to identify. Subsequently, the new information was added to the current learned model.

Table 2 shows information on the categories of the different houses. We used this data to validate the efficiency of the model.

Table 2. Images from different houses.

No.	Category	Qty.	House
1	Kitchen	1,113	02
2	Bedroom	1,293	02
3	Bedroom 1	378	03
4	Bedroom 2	330	03
5	Bedroom 3	403	03
6	Bedroom 4	512	03
7	Corridor 1	436	03
8	Corridor 2	529	03
9	Bedroom	359	04
10	Kitchen	212	04
11	Bedroom 1	641	05
12	Bedroom 2	642	05
13	Balcony	571	05
14	Bathroom	427	05
15	Corridor	412	05
16	Kitchen	665	05

5 Results and Discussion

First, we comment on the experiments carried out in the different houses using only the base model, and then we discuss what happened when the system was unsuccessful with the classification and we capture information from the new environment. A summary of the results for the experiments performed can be found in Table 3.

Experiment 1 establishes the baseline we use to compare the following experiments. The total accuracy of the test is 99.98%. This represents the starting line, as no new knowledge was added.

In Experiment 2, a 69.81% success rate was obtained. This was conducted in House 02. This can be considered a considerable success given that we used a completely different environment that the system had never seen before.

Experiment 3 used the bedroom in House 02, obtaining a success rate of 50.27%.

Experiments from 4 to 9 were performed in House 03, obtaining results of (Bedrooms 1 → 48.14%), (Bedrooms 2 → 32.12%), (Bedrooms 3 → 28.53%), (Bedrooms 4 → 60.54%), (Corridor 1 → 85.09%) (Corridor 2 → 75.61%).

As in the kitchen of House 02, the Corridors and Bedroom 4 achieved an acceptable accuracy percentage considering that it was an environment the system had never seen before.

Table 3. Summary of the experiments

No.	Data	House	Acc. without retraining	Acc. with retraining
1	Test	01	99.9	Not applicable
2	Kitchen	02	69.8	100
3	Bedroom	02	50.3	99.5
4	Bedroom 1	03	48.1	100
5	Bedroom 2	03	32.1	100
6	Bedroom 3	03	28.5	100
7	Bedroom 4	03	60.5	100
8	Corridor 1	03	85.1	100
9	Corridor 2	03	75.6	100
10	Bedroom	04	24.5	100
11	Kitchen	04	10.4	100
12	Bedroom 1	05	46.8	100
13	Bedroom 2	05	66.9	100
14	Balcony	05	0.7	100
15	Bathroom	05	48.7	100
16	Corridor	05	44.7	100
17	Kitchen	05	10.5	100

Experiments 10 and 11 were carried out in House 04. Experiment 10 scored 24.51%. This is the lowest success rate obtained in this category. This was mainly due to the visual difference between this category and the images used in the model. Experiment 11 scored 10.37%. As in the previous case, a low percentage was obtained, which was due to the visual difference between this category and the images used in the model.

Experiments 12 to 17 were carried out in House 05, obtaining a result of (Bedrooms 1 → 46.80%), (Bedrooms 2 → 66.97%), (Balcony → 0.70%), (Bathroom → 48.71%), (Corridor → 44.66%) (Kitchen → 10.52%).

It should also be noted that many of the confusions were caused by specific elements appearing in the scenes, such as the case of Experiment 2, which was conducted in a kitchen that had a washing machine, with the system mistaking these images for a laundry (see Fig. 4). When the data was captured, accuracy in all categories increased considerably.

(a) Image from the Laundry category

(b) Image from the Kitchen of House 2

Fig. 4. Image comparison

6 Conclusions and Future Work

As mentioned at the beginning of this study, it is important for the robot to know its location since this is the starting point for other actions it can perform. To provide the robot with the ability to identify its location we used an existing methodology to identify a place. However, this methodology had previously only been developed in the laboratory and had not been tested in a real environment.

After conducting the appropriate evaluations, we conclude that a model will produce better recognition results when tested in an environment similar to that in which it was trained compared to when tested in a different environment.

Our experimental results show that a trained model will obtain more accurate results when re-testing images from the same environment. Furthermore, for images belonging to a different house, the model obtained less accurate results when compared to those for the same house. On the other hand, as was to be expected when adding new knowledge on the environment, the success rate increased considerably.

As future work, we propose to evaluate this approach with different lighting conditions and also introduce more houses in the study. Another feature could be evaluation using 3D information about the places to improve the results using more information.

We also plan to merge the original set of images used in the study with the information generated by user feedback in order to create a full dataset.

Acknowledgements. This work has been supported by the Spanish Government TIN2016-76515R Grant, supported with Feder funds. Edmanuel Cruz is funded by a Panamenian grant for PhD studies IFARHU & SENACYT 270-2016-207. This work has also been supported by a Spanish grant for PhD studies ACIF/2017/243. Thanks to Nvidia also for the generous donation of a Titan Xp and a Quadro P6000.

References

1. Oore, S., Hinton, G., Dudek, G.: A mobile robot that learns its place. Neural Comput. **9**(3), 683–699 (1997)
2. Koenig, S., Simmons, R.: Xavier: a robot navigation architecture based on partially observable Markov decision process models. In: Kortenkamp, D., Bonasso, R., Murphy, R. (eds.) Artificial Intelligence Based Mobile Robotics: Case Studies of Successful Robot Systems, pp. 91–122. MIT-Press, Cambridge (1998)
3. Dillmann, R., Rogalla, O., Ehrenmann, M., Zöliner, R., Bordegoni, M.: Learning robot behaviour and skills based on human demonstration and advice: the machine learning paradigm. In: Hollerbach, J.M., Koditschek, D.E. (eds.) Robotics Research, pp. 229–238. Springer, London (2000)
4. Buschka, P., Saffiotti, A.: A virtual sensor for room detection. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 637–642 (2002)
5. Kuipers, B., Beeson, P.: Bootstrap learning for place recognition. In: Proceedings of the National Conference on Artificial Intelligence (AAAI) (2002)
6. Althaus, P., Christensen, H.: Behaviour coordination in structured environments. Adv. Robot. **17**(7), 657–674 (2003)
7. Pronobis, A., Mozos, O.M., Caputo, B., Jensfelt, P.: Multi-modal semantic place classification. Int. J. Robot. Res. **29**(2–3), 298–320 (2010). https://doi.org/10.1177/0278364909356483
8. Stückler, J., Behnke, S.: Improving people awareness of service robots by semantic scene knowledge. In: Ruiz-del-Solar, J., Chown, E., Plöger, P.G. (eds.) RoboCup 2010: Robot Soccer World Cup XIV, pp. 157–168. Springer, Heidelberg (2011). ISBN: 978-3-642-20217-9

9. Lim, H., Sinha, S.: Towards real-time semantic localization. In: ICRA Workshop on Semantic Perception (2012)
10. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: ImageNet large scale visual recognition challenge. IJCV **115**, 211–252 (2015)
11. Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., Oliva, A.: Learning deep features for scene recognition using places database. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 27, pp. 487–495. Curran Associates, Inc. (2014), http://papers.nips.cc/paper/5349-learning-deep-features-for-scene-recognition-using-places-database.pdf
12. Goeddel, R., Olson, E.: Learning semantic place labels from occupancy grids using CNNs. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3999–4004, October 2016
13. Martinez-Gomez, J., Gimenez, V.M., Cazorla, M., Garcia-Varea, I.: Semantic Localization in the PCL library. Robot. Auton. Syst. **75**(Part B), 641–648 (2016). http://www.sciencedirect.com/science/article/pii/S0921889015001943
14. Lowry, S., Sünderhauf, N., Newman, P., Leonard, J.J., Cox, D., Corke, P., Milford, M.J.: Visual place recognition: a survey. IEEE Trans. Robot. **32**(1), 1–19 (2016). https://doi.org/10.1109/TRO.2015.2496823
15. Schünberger, J.L., Pollefeys, M., Geiger, A., Sattler, T.: Semantic visual localization. CoRR, abs/1712.05773,2017. http://arxiv.org/abs/1712.05773
16. Armeni, I., Sener, O., Zamir, A.R., Jiang, H., Brilakis, I., Fischer, M., Savarese, S.: 3D semantic parsing of large-scale indoor spaces. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016
17. Ma, S., Liu, Q.: Semantic localization. In: Encyclopedia with Semantic Computing and Robotic Intelligence, vol. 01, no. 01 (2017). https://doi.org/10.1142/S242503841630010X
18. Romero-González, C., Martínez-Gómez, J., García-Varea, I., Rodríguez-Ruiz, L.: On robot indoor scene classification based on descriptor quality and efficiency. Expert Syst. Appl. **79**, 181–193 (2017). http://www.sciencedirect.com/science/article/pii/S0957417417301318
19. Young, J., et al.: Making sense of indoor spaces using semantic web mining and situated robot perception. In: Blomqvist, E., Hose, K., Paulheim, H., Lawrynowicz, A., Ciravegna, F., Hartig, O. (eds.) The Semantic Web: ESWC 2017 Satellite Events. ESWC 2014. Lecture Notes in Computer Science, vol. 10577. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-70407-439
20. Rosa, S., Lu, X., Wen, H., Trigoni, N.: Leveraging user activities and mobile robots for semantic mapping and user localization. In: Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI 2017), pp. 267–268. ACM, New York (2017). https://doi.org/10.1145/3029798.3038343
21. Cruz, E., Rangel, J.C., Cazorla, M.: Robot semantic localization through CNN descriptors. In: Ollero, A., Sanfeliu, A., Montano, L., Lau, N., Cardeira, C. (eds.) ROBOT 2017: Third Iberian Robotics Conference, pp. 567–578. Springer, Cham (2018). https://doi.org/10.1145/3215525.3215526

22. Rosa, S., Patané, A., Lu, X., Trigoni, N.: CommonSense: collaborative learning of scene semantics y robots and humans. In: IoPARTS 2018: 1st International Workshop on Internet of People, Assistive Robots and ThingS, Munich, Germany, 10 June 2018, 6 pages. ACM, New York (2018). https://doi.org/10.1145/3215525.3215526
23. Cruz, E., Rangel, J.C., Gómez-Donoso, F., Bauer, Z., Cazorla, M., Garcia-Rodriguez, J.: Finding the place: how to train and use convolutional neural networks for a dynamically learning robot. In: 2018 International Joint Conference on Neural Networks, IJCNN, pp. 3655–3662, July 2018. ISBN: 978-1-5090-6014-6

Positioning System for an Electric Autonomous Vehicle Based on the Fusion of Multi-GNSS RTK and Odometry by Using an Extented Kalman Filter

Miguel Tradacete(✉), Álvaro Sáez(✉), Juan Felipe Arango(✉), Carlos Gómez Huélamo(✉), Pedro Revenga(✉), Rafael Barea(✉), Elena López-Guillén(✉), and Luis Miguel Bergasa(✉)

Robesafe Group, Department of Electronics, University of Alcala, Alcalá de Henares, Spain
tradacete.miguel@gmail.com,
{alvaro.saezc,juanfelipe.arango,carlos.gomezh}@edu.uah.es,
{rafael.barea,elena.lopezg}@uah.es
https://www.robesafe.com/

Abstract. This paper presents a global positioning system for an autonomous electric vehicle based on a Real-Time Kinematic Global Navigation Satellite System (RTK- GNSS), and an incremental-encoder odometry system. Both elements are fused to a single system by an Extended Kalman Filter (EKF), reaching centimeter accuracy. Some varied experiments have been carried out in a real urban environment to compare the performance of this positioning architectures separately and fused together. The achieved aim was to provide autonomous vehicles with centimeter precision on geolocalization to navigate through a real lane net.

Keywords: Autonomous vehicle · Positioning · Odometry
Multi-GNSS · Kalman Filter

1 Introduction

Vehicle positioning and tracking have numerous applications in general transport-related studies including vehicle navigation, fleet monitoring, traffic congestion etc. In the last decade, many works have been focused in studying driving behaviour through examining the vehicle movement trajectory using GNSS signals, mostly GPS [1–3]. These methods have been able to provide both geolocalization and time information to a receiver employing multiple satellite signals while they stay fast, accurate, and cost-efficient. However, their performance has a strong dependence on several system factors such as the number of visible satellites, their positions or the capability of the GPS receiver. In addition, the signal trips through the layers of the atmosphere, and some other

R. Fuentetaja Pizán et al. (Eds.): WAF 2018, AISC 855, pp. 16–30, 2019.
https://doi.org/10.1007/978-3-319-99885-5_2

sources contribute to inaccuracies and errors in the GPS signals by the time they reach a receiver. Thus, the accuracy provided by this methods is low (usually between 1 and 10 m), they need a considerable time (over 30 s) to provide the first position measurement and they do not guarantee a robust service in several situations such as environments with poor signal conditions.

The development of Intelligent Vehicles (IVs) has specially grown during the last years. These systems aim to solve complex issues with specially demanding accuracy requirements (usually decimeter precision) like autonomous driving applications where tasks such as lane maintenance analysis demand centimeter precision [4]. Furthermore, autonomous vehicles also require robust solutions with low latencies and high time availability so standalone GNSS techniques are not adequate for them.

Various solutions are proposed to achieve a better service quality: to deal with the accuracy problem Differential GPS (DGPS) is used to obtain an accuracy enhancement using data from a reference station [5,6] and the more complex Real-time Kinematic (RTK) positioning solution, which uses carrier phase information, has attracted much interest in applications with strict precision requirements due to its centimeter-level accuracy [7]. To approach the robustness issue Multi-GNSS (multiple Global Navigation Satellite Systems) techniques are being widely-used, boosted by the appearance of alternative GNSSs based on different satellite constellations like Russian (GLONASS), European Union (Galileo), Chinese (Beidou) or Japanese (QZSS). Multi-GNSS allows to easily increase the number of tracked satellites to over 10 in good signal conditions and to more than 5 in almost any other situation, even including dense urban areas combining multiple GNSS [4]. Several studies have proven the benefits of these techniques combining GPS and GLONASS [8,9], GPS and Galileo [10] or even four of the available systems (GPS+Galileo+BeiDou+QZSS) [11].

Nevertheless, even the combination of the previous methods might not be enough to cover autonomous vehicles needs in certain environments such as dense urban or concrete places like tunnels. To face this challenging situations, GNSS data needs to be fused with local sensors information when the measurement's quality is degraded. In [12] RTK-DGPS was fused with speed vehicle sensors and steering-wheel position measurements to improve vehicle tracking. Other works like [13] used an Extended Kalman Filter to integrate DGPS with some vehicle sensors like an inertial navigation system (INS) through a kinematic model in order to achieve enough accuracy to enable vehicle cooperative collision warning without the use of ranging sensors.

This paper presents a robust real-time positioning system for autonomous vehicles that reaches centimeter precision. The system uses a GNSS receiver and an incremental-encoder odometry, integrated by an Extended Kalman Filter which leverages quality of the received satellite measurements. As well as, odometry system is calibrated through an automatic process applying a least square adjustment of the position error of a variety of routes. Experiments presented in Sect. 4 show that our system is able to keep the vehicle in the middle of the lane nets even in regions without available differential corrections.

Furthermore, the system is complemented with a reactive navigation module based on vision and Lidar that slightly relaxes the positioning requirements.

This paper is organized as follows: Sect. 2 presents the system's structure together with an analysis of the main modules that compose it and their corresponding standalone performances. Section 3 analyzes the integration of both modules using the Extended Kalman Filter and the following Sect. 4 exposes the results of the performed experiments to test the final system with different configurations. Last Sect. 5 presents the final conclusions and future work lines.

2 System Architecture

The positioning system is integrated in an open-source electric vehicle (TABBY EVO Vehicle 4 seats version) modified and automatized by the University of Alcalá. The system's architecture includes a Real-Time Kinematic Multi-Global Navigation Satellite System based on both GPS and GLONASS with a local base station that broadcasts differential corrections, a GPRS modem, and an incremental-encoder odometry system. Its sensor equipment is composed of a GNSS receiver, a Choke-Ring Antenna for the local base station, and two Kübler 3700 incremental encoders for odometry. All these modules are fused in Robotic Operating System (ROS) using an Extended Kalman Filter. Figure 1 shows the general diagram of the system.

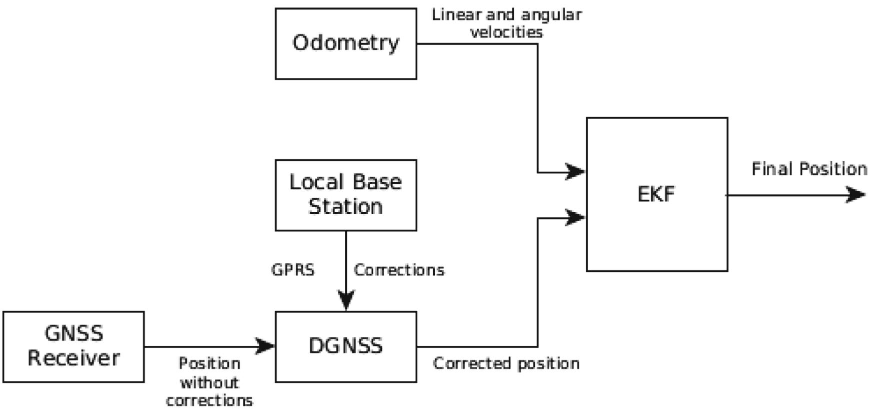


Fig. 1. System architecture diagram

GNSS receiver is set on top of the vehicle to obtain maximum coverage, and Choke-Ring Antenna for a base station is set on the Polytechnic School building's roof. The odometry encoders are assembled in both rear-wheel shafts by 3D-printed pieces. ROS runs on two embedded GPUs looking for the benefits of modularity. These GPUs are a Nvidia Jetson TX2, and a Raspberry Pi 3 for odometry processing. Figure 2 displays the entire vehicle, and Fig. 3(a) (where

Fig. 2. General view of the TABY EVO-OSVehicle

(a) GPS and Lidar (b) Incremental encoder

Fig. 3. Vehicle sensor equipment

Lidar is shown as part of reactive navigation system) and Fig. 3(b) show GNSS receiver and odometry encoder details.

2.1 Multi-GNSS

The main module of the localization system consists of a multi-constellation system (multi-GNSS) with RTK positioning solution. In addition, it also includes two elements: a differential Hiper Pro GPS+ receiver configured as rover, and a local base station to generate differential corrections.

The rover is able to obtain data from both GLONASS and GPS to provide a more robust solution than a standard GPS by increasing the number of visible satellites. It provides positioning information at 10 Hz as autonomous vehicles