OpenABLE: An Open-source Toolbox for Application in Life-Long Visual Localization of Autonomous Vehicles

Roberto Arroyo, Pablo F. Alcantarilla, Luis M. Bergasa and Eduardo Romera

Abstract— **Visual information is a valuable asset in any perception scheme designed for an intelligent transportation system. In this regard, the camera-based recognition of locations provides a higher situational awareness of the environment, which is very useful for varied localization solutions typically needed in long-term autonomous navigation, such as loop closure detection and visual odometry or SLAM correction. In this paper we present OpenABLE, an open-source toolbox contributed to the community with the aim of helping researchers in the application of these kinds of life-long localization algorithms. The implementation follows the philosophy of the topological place recognition method named ABLE, including several new features and improvements. These functionalities allow to match locations using different global image description methods and several configuration options, which enable the users to control varied parameters in order to improve the performance of place recognition depending on their specific problem requisites. The applicability of our toolbox in visual localization purposes for intelligent vehicles is validated in the presented results, jointly with comparisons to the main state-of-the-art methods.**

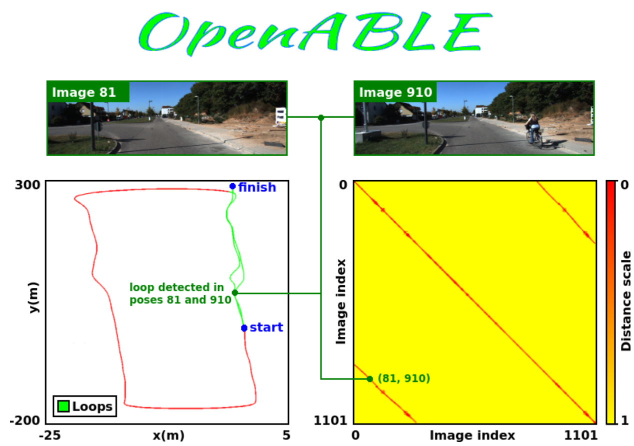


Fig. 1. OpenABLE logo and a general diagram that graphically depicts how the application of our toolbox can help in varied localization problems. The example showed in this case correspond to the recognition of a revisited location in the sequence 06 of the KITTI Odometry dataset.

I. Introduction

Localization is a key aspect to develop a robust automated navigation approach in the context of an intelligent transportation system (ITS). In this sense, visual information is a helpful resource in order to identify locations by means of topological place recognition techniques, which can be complementary to GPS-based solutions, or even supplementary in environments where GPS signal is not completely available or denied [1]. Moreover, the camera-based recognition of locations is commonly used in several localization problems that are typically studied in autonomous vehicles, such as visual odometry [2] or SLAM [3]. In these systems, the visual detection of loop closures also helps to correct the drift appeared over time.

Nowadays, one of the main challenges in visual localization is the recognition of places in a long-term context. This is because of the drastic changes that the appearance of a place progressively suffers due to environmental conditions, as studied in [4]. Probably, the variability in appearance derived from the four seasons of the year is the most critical [5], because it causes important changes in different visual aspects: illumination, weather, vegetation, dynamic elements or field of view, among others. Although some recent works on ITS have analyzed these difficulties in life-long navigation [6] and localization [7], this is still an open problem of great interest in the research community due to the requirement of long-term reliable solutions.

For these reasons, in this paper we present OpenABLE[1], which is an open-source toolbox contributed to the community with the aim of helping researchers in the application of life-long visual localization algorithms, as depicted in Fig. 1. The proposed implementation is based on the philosophy of the topological place recognition method named ABLE, whose main characteristics were introduced in some of our previous research [8], [9], [10]. However, OpenABLE includes several new features and improvements for matching locations using different global description methods and configuration options, which will be extensively described and justified with varied results along this paper.

The motivation of our work is inspired by recent open projects on ITS that have demonstrated to be a great public contribution to the community in their respective areas, such as the KITTI benchmark suite [11] or the LIBVISO libraries for visual odometry [2]. Our goal is that OpenABLE can also give support to ITS applications in the specific area of life-long visual localization for autonomous vehicles.

*This work has been funded in part from the University of Alcalá (UAH) through a FPI grant, the Spanish MINECO through the SmartElderlyCar project (TRA2015-70501-C2-1-R) and the Community of Madrid through the RoboCity2030-III-CM project (S2013/MIT-2748).

R. Arroyo, L. M. Bergasa and E. Romera are with the Department of Electronics, University of Alcalá (UAH), Alcalá de Henares, 28871, Madrid, Spain. E-mails: {roberto.arroyo, bergasa, eduardo.romera}@depeca.uah.es.

P. F. Alcantarilla is with iRobot Corporation, 10 Greycoat Place, Victoria, London, UK. palcantarilla@irobot.com

[1]More information, extra material, videos and source code about the OpenABLE toolbox are available from the website of the project: http://www.robesafe.com/personal/roberto.arroyo/openable.html

II. RELATED WORK

Vision-based methods for identifying locations have been broadly studied in the recent past, especially after the definition of FAB-MAP [12]. This proposal supposed a revolution in the state of the art and became a popularized methodology for topological place recognition and loop closure detection by only analyzing the space of visual appearance. However, FAB-MAP has some disadvantages, such as the need of a prior training stage or its computationally expensive feature extraction and probabilistic inference.

Due to this, other techniques were proposed in order to improve the efficiency and effectiveness of topological localization in a long-term context. In this regard, SeqSLAM [13] can be considered as one of the most successful approaches, which was tested for place recognition in challenging conditions for a same route recorded in a sunny summer day and a stormy winter night. SeqSLAM introduced the idea of using sequences of images to define locations instead of single images in order to enhance the performance in long-term scenarios. Nevertheless, subsequent works demonstrated that SeqSLAM has a dependence on the camera point of view [5]. More recently, methods such as ABLE have ameliorated the behavior for changing field of views in life-long localization across the four seasons, as evaluated in [10].

Some time after their formulation, open implementations of FAB-MAP and SeqSLAM were released for the benefit of the research community: OpenFABMAP [14] and OpenSeqSLAM [5]. Motivated by these contributions, the OpenABLE toolbox presented in this paper also expects to assist future researchers in solving the new challenges associated with the novel trends recently appeared in life-long visual localization for intelligent vehicles [4], [6], [7].

III. ABLE: A BRIEF REVIEW

Before explaining the main properties of our OpenABLE toolbox, it is necessary to introduce a brief review about the basic theory behind the visual topological localization methodology in which is inspired: ABLE (Able for Binary-appearance Loop-closure Evaluation).

The approach originally proposed by ABLE is focused on the application of binary descriptors in visual localization. This is mainly because of their low memory consumption and computational costs in image description and matching, jointly with a satisfactory effectiveness for place recognition. Some of these beneficial properties are due to the simplicity in calculation of these kinds of descriptors, which is initially based on binary tests (τ) computed over smoothed image patches ($\mathbf{p}$) centered in a point of interest, as shown in Eq. 1.

$$\tau\left(\mathbf{p}; f(i), f(j)\right) \;=\; \begin{cases} 1 & f(i) < f(j) \\ 0 & f(i) >= f(j) \end{cases} \qquad (1)$$

In the previous formulation, $f(i)$ is a function that returns an image feature response for a certain cell of pixels ($\mathbf{c_i}$) in $\mathbf{p}$. According to this, ABLE typically uses a binary descriptor named LDB [15], which is based on a $f(i)$ that concatenates binary comparisons over averaged image intensities (I_{avg}) and image gradients (G_x, G_y), as exposed in Eq. 2. Besides, a descriptor named D-LDB is proposed for stereo images in [9], where disparity comparisons (D_{avg}) are also added.

$$f(i) \;=\; \{I_{avg}(\mathbf{c_i}), G_x(\mathbf{c_i}), G_y(\mathbf{c_i})\} \qquad (2)$$

The resulting binary descriptor ($\mathbf{d}$) is calculated as a sequence of n binary tests (see Eq. 3). It must be noted that ABLE efficiently computes LDB as a global descriptor by reducing the processed images to the patch size and taking its center as a keypoint without dominant rotation or scale.

$$\mathbf{d}(\mathbf{p}) = \sum_{1 \leq i \leq n} 2^{i-1} \tau\left(\mathbf{p}; f(i), f(j)\right) \qquad (3)$$

One of the main benefits of applying binary features is that they can be matched using a basic Hamming distance, which is more efficient than using the typical L_2-norm. ABLE takes advantage of this property and matches the computed LDB descriptors by processing the Hamming distance, whose efficiency is due to its simple calculation based on an elementary XOR operation ($\oplus$) followed by a sum of bits (see Eq. 4). After that, the distances between the analyzed images can be stored on a similarity matrix (M).

$$M_{i,j} = M_{j,i} = \text{bitsum}(\mathbf{d}_i \oplus \mathbf{d}_j) \qquad (4)$$

Apart from the previous comments about image description and matching in ABLE, other applied techniques give an extra robustness to this method in life-long identification of locations. On the one hand, places are considered as sequences of images instead of single images, because it enhances the performance in a long-term context, as introduced in works such as [13]. On the other hand, illumination invariant images are computed to improve the accuracy of ABLE against changes in the lighting of the scene, which is also studied in works for visual perception in intelligent vehicles [16]. All these characteristics are graphically described in the general diagram of ABLE presented in Fig. 2, where the example images recorded in different seasons of the year belong to the Nordland dataset [5].

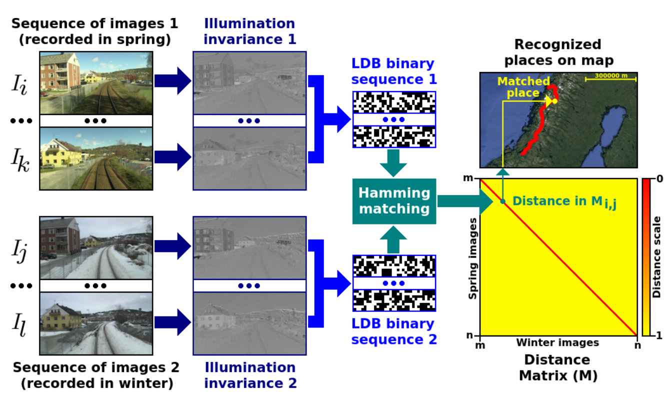


Fig. 2. A graphic representation about how the methodology proposed by ABLE works in life-long visual localization across the seasons of the year.

Although the images showed in the schema of Fig. 2 correspond to a monocular case (ABLE-M [10]), it must be noted that more detailed proposals are also defined for stereo (ABLE-S [9]) and panoramic cameras (ABLE-P [8]).

IV. OpenABLE

OpenABLE is a project that arises from the growing interest in an open-source philosophy in research. Our main objective is to publicly share our toolbox with the ITS, robotics and computer vision communities, because it can be helpful for the common progress of researchers in life-long visual localization topics. Due to this, the code provided in OpenABLE can be freely employed and modified by the users, which can also include our algorithms into a larger system or directly apply them over their localization problems. In this sense, the toolbox contains varied novel configuration options to facilitate its utilization and it is easily adaptable to different applications and specifications.

A. Novelties in OpenABLE

OpenABLE is much more than a simple open implementation of the original ABLE method. The described toolbox contributes several new characteristics and functionalities with the aim of providing a wide range of possibilities to the users:

- Apart from the descriptors originally used by ABLE (LDB and D-LDB), OpenABLE supplies implementations based on BRIEF [17], BRISK [18], ORB [19], FREAK [20], SIFT [21], SURF [22] and HOG [23].
- The typical Hamming matching used for binary features is substituted by a L_2-norm when vector-based descriptors are applied (SIFT, SURF, HOG).
- An image description method based on multiple grids is now implemented, as well as the common global description normally performed by ABLE.
- A thresholding functionality is provided for improving the filtering of loop closures in the similarity matrix.
- Multi-camera approaches are available. OpenABLE exploits the extra image information procured in any case: monocular, stereo or panoramic.

In addition, OpenABLE has some other advantages with respect to similar toolboxes. For instance, OpenFABMAP requires a previous training, while our toolbox is completely training-free. Besides, our method has a better behavior for the changing fields of view typically appeared in camera-based localization for autonomous vehicles, which is one of the weaknesses of OpenSeqSLAM, as evaluated in [5] and in some results presented in this paper in Section V.

B. Main Characteristics of the Toolbox

The toolbox is developed in C++ because it is a standard programming language adaptable to varied system requirements. Although the code has been designed under a Linux operating system, it is easily portable to other platforms such as Windows or Mac OS. OpenCV 3.0[2] is required because some of the computer vision algorithms programmed in OpenABLE apply functionalities of these libraries. A file named *CMakeLists.txt* is provided jointly with the code to facilitate the compilation of the toolbox by using CMake[3].

[2]OpenCV is currently available from: http://opencv.org/

[3]CMake is currently available from: https://cmake.org/

The core and main functions of the toolbox are contained in the source code files named *OpenABLE.h* and *OpenABLE.cpp*. An evaluation program (*Test_OpenABLE.h*) is supplied, where an image sequence or a video is processed to return the final distance matrix (M'), which is normalized between 0 and 1, as exposed in Eq. 5. The open code of the LDB descriptor provided by its authors in [15] is also included jointly with our own implementation of D-LDB [9]. In addition, a file named *Config.txt* is incorporated to easily adapt the properties and functionalities of OpenABLE to the different interests of the users. These configuration options are explained in the following sections because of their importance for the application of the toolbox in a varied range of possibilities. Processing times are reported after execution to know the efficiency of an specific configuration.

$$M'_{i,j} = \frac{M_{i,j}}{\max(M)} \tag{5}$$

C. Configuration Parameters for Datasets

Some basic parameters are configurable to select the paths to the input recordings or datasets that contain the images or videos used in the required visual localization tasks.

D. Configuration Parameters for Representation

The configuration options included in this group allow to define the paths where the results generated by OpenABLE will be stored and some other representation features, such as 12 ranges of colors to visualize the distance matrices.

E. Configuration Parameters for Description and Matching

These are the most important parameters for adapting the performance of OpenABLE to the users' priorities and they must be individually described in detail:

*1) **Camera_type**:* This option allows to select the type of camera used for acquiring the images applied in the tests performed with OpenABLE: monocular (ABLE-M), stereo (ABLE-S) and panoramic (ABLE-P).

*2) **Description_type**:* This parameter lets to choose between a computation of features using a global or a grid-based image description. In [9], some results proved that the application of grids can slightly improve the precision, but it also progressively increases the computational cost.

*3) **Patch_size**:* The length of the square defined as patch for the image description process can be adjusted in OpenABLE with this parameter. If a global description is used, the images are also downsampled to this size. In [8], it is tested how a patch of 64x64 can be enough for an effective and efficient visual loop closure detection.

*4) **Grid_x**:* It defines the number of horizontal grids applied over an image if grid-based description is enabled.

*5) **Grid_y**:* It defines the number of vertical grids applied over an image if grid-based description is enabled.

*6) **Panoramas**:* If the toolbox is computing panoramic images, the number of subpanoramas considered in image processing can be selected using this parameter. This concept is explained in detail in [8], where it is exposed how a cross-correlation of subpanoramas can detect places revisited in an opposite direction (bidirectional loop closures).

*7) **Illumination_invariance**:* If this option is enabled, an illumination invariant technique is applied to improve the robustness in changing lighting conditions. Its implementation is based on Eq. 6, where R, G, B are the initial color channels and $\mathcal{I}$ is the final illumination invariant image.

$$\mathcal{I} = \log(G) - \alpha \cdot \log(B) - (1-\alpha) \cdot \log(R) \qquad (6)$$

*8) **Alpha**:* When illumination invariance is used, this parameter represents the value of α shown in Eq. 6. This value can be estimated by following Eq. 7, where λ_R, λ_G, λ_B are the peak spectral responses of each color channel, which are usually available in camera specifications. For example, the camera used in KITTI (PointGrey Flea2) has $\lambda_R = 610nm$, $\lambda_G = 535nm$, $\lambda_B = 470nm$, so in this case $\alpha = 0.47$.

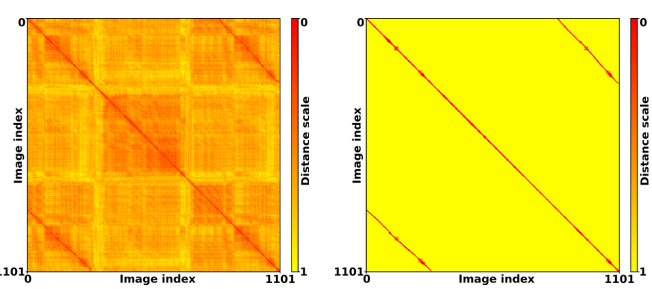
$$\alpha = \frac{\left(\frac{\lambda_B}{\lambda_G} - \frac{\lambda_B}{\lambda_R}\right)}{\left(1 - \frac{\lambda_B}{\lambda_R}\right)} \qquad (7)$$

*9) **Image_descriptor**:* Although ABLE originally applies LDB as core image descriptor (or D-LDB for stereo images), OpenABLE allows to choose other binary features, such as BRIEF, BRISK, ORB and FREAK. In addition, vector-based descriptors are also available, such as SIFT, SURF and HOG. In these cases, a matching based on the L_2-norm is internally computed by default (see Eq. 8), because vector-based features are not compatible with the Hamming distance.

$$M_{i,j} = M_{j,i} = \sqrt{\sum_{k=1}^{n} (\mathbf{d}_{i_k} - \mathbf{d}_{j_k})^2} \qquad (8)$$

*10) **Image_sequences**:* The length of the sequence of images used by OpenABLE is configurable. If the value of this option is 1, single images are employed. A higher length gives better precision in life-long localization. A detailed study of this was provided in [10]. For instance, a length of 300 was suggested as proper for the Nordland dataset.

*11) **Threshold**:* This parameter allows to apply a threshold (θ) over the distance matrix calculated by OpenABLE. If it is enabled, Eq. 9 is computed to generate the thresholded similarity matrix (M''). The advantage of using this option is that loop closures can be better detected, as shown in Fig. 3.



(a) Before thresholding. (b) After thresholding.

Fig. 3. Comparison between the distance matrix initially obtained by OpenABLE in a standard configuration and after a thresholding ($\theta = 0.2$). The example corresponds to the sequence 06 of the KITTI Odometry dataset.

$$M''_{i,j} = \begin{cases} 0 & \text{if} \quad M'_{i,j} < \theta \\ 1 & \text{otherwise} \end{cases} \qquad (9)$$

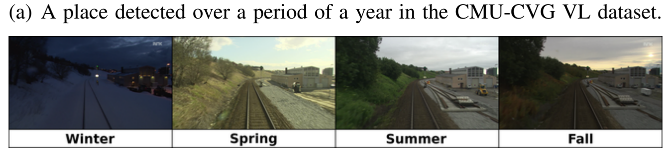
V. Results and Practical Applications

Our toolbox has been widely validated over several datasets and under varied conditions using monocular, stereo and panoramic cameras. Here, we describe the main performance results obtained by OpenABLE in life-long visual topological localization. Additionally, we show examples about the application of the toolbox in different tasks, such as place recognition, loop closure detection or visual odometry and mapping correction. These practical applications could be extended in the future by the users of our open code.

A. Using OpenABLE with Monocular Cameras

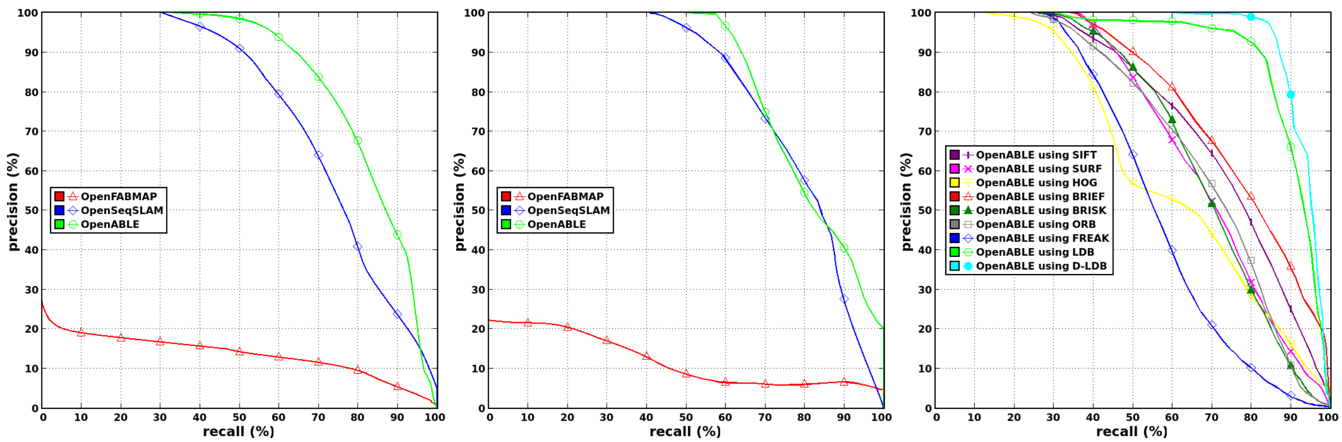
We have performed tests using the CMU-CVG VL [7] and the Nordland datasets [5], which are recorded in long-term conditions using monocular cameras. Precision-recall results are shown for the most difficult case in both datasets: sequences acquired in winter vs spring. These tests are presented in Fig. 5(a) and Fig. 5(b), where our method is compared against other state-of-the-art toolboxes. OpenFABMAP and OpenSeqSLAM are applied using their standard configurations defined in [14] and [5], while OpenABLE uses the original parameters of the ABLE method in this case [10]. The results evidence the better performance of the proposals based on sequences of images (OpenSeqSLAM and OpenABLE). Besides, our toolbox has a more favorable behavior in datasets with changes on the field of view, such as the CMU-CVG VL dataset. Fig. 4 depicts some image examples about the application of OpenABLE in place recognition for all the sequences of the tested datasets.

(a) A place detected over a period of a year in the CMU-CVG VL dataset.



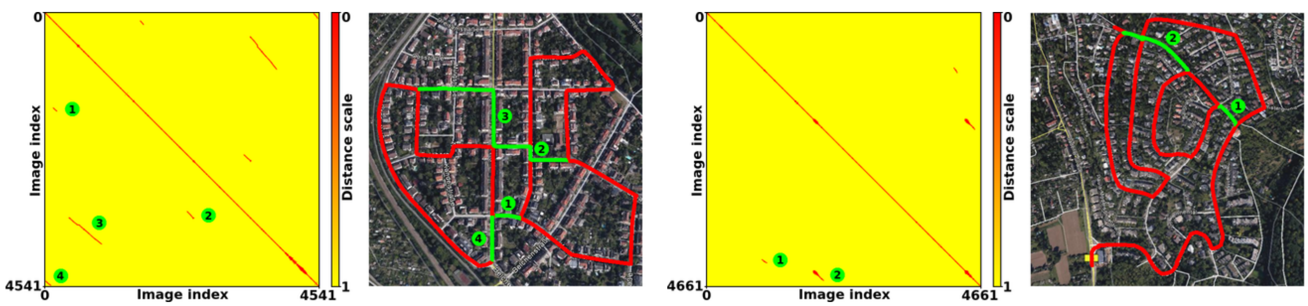
(b) A place detected across the seasons of the year in the Nordland dataset.

Fig. 4. Examples of recognized places using the distance matrices provided by OpenABLE in challenging conditions derived from life-long localization.

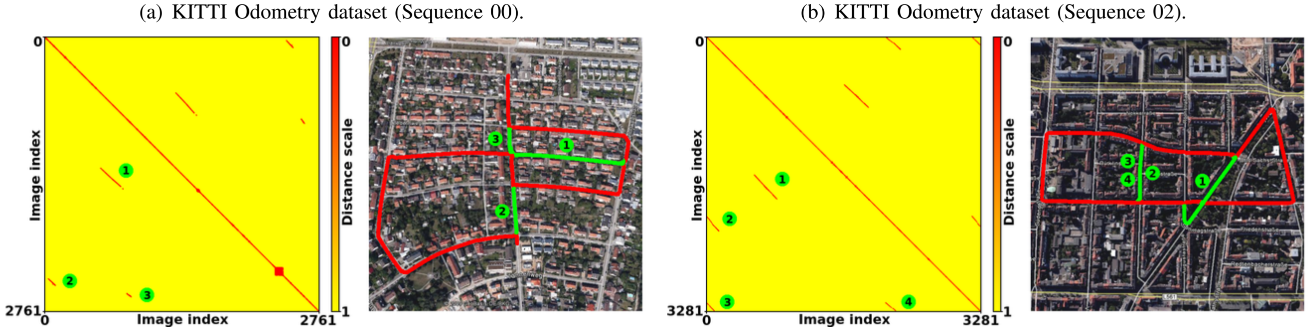


(a) CMU-CVG VL dataset (21/12/10 vs 21/04/11). (b) Nordland dataset (Winter vs Spring). (c) KITTI Odometry dataset (Sequence 06).

Fig. 5. Precision-recall curves for comparing the performance of OpenABLE against other state-of-the-art toolboxes and using different descriptors.



(a) KITTI Odometry dataset (Sequence 00). (b) KITTI Odometry dataset (Sequence 02).



(c) KITTI Odometry dataset (Sequence 05). (d) KITTI Odometry dataset (Sequence 13).

Fig. 6. Some examples testing the application of loop closure detection over the most representative sequences of the KITTI Odometry dataset. The results obtained in the distance matrices calculated by OpenABLE are represented over the metric map, where loop closure zones are depicted in green.

B. Using OpenABLE with Stereo Cameras

The KITTI Odometry dataset [11] is chosen for our experiments with stereo cameras because of its reputation for autonomous driving evaluation. It is composed of 22 sequences recorded in multiple car routes. Some of them contain loop closures, which are specifically defined in a ground-truth described in our previous work [9].

Here, we test the accuracy of OpenABLE using the different available features as core of the description method. Some results about it are exposed in Fig. 5(c), where LDB and D-LDB obtain the most satisfactory performance in this specific case. It is mainly due to the addition of gradient (LDB) and disparity (D-LDB) comparisons in the binary description process, which commonly provides a higher robustness in this sense. However, the other available features could be an alternative option depending on the interests of the users of the toolbox or their particular problems.

Moreover, Fig. 6 shows how the distance matrices generated by OpenABLE are applicable to loop closure detection for tasks related to visual odometry correction, which is studied in works such as [24]. We present several mapping results, where loop closures are annotated while they are identified by OpenABLE. The maps are constructed using KML files for saving the registered poses, which are represented over satellite images by means of Google Earth®[4].

[4]Google Earth is currently available from: https://www.google.com/earth/

C. Using OpenABLE with Panoramic Cameras

Panoramas give a visual perception of the environment in all the possible directions that allows to detect not only unidirectional loop closures, but also the bidirectional ones. OpenABLE exploits this advantage, as demonstrated in the results shown in Fig. 7 for the Oxford New College dataset [25], which is recorded using panoramic cameras. Here, the relative poses provided jointly with the dataset are employed to depict all the associated loop closures identified by our toolbox over the map. The drift perceived in the registered trajectory could be corrected by means of the processed information about loop closures.

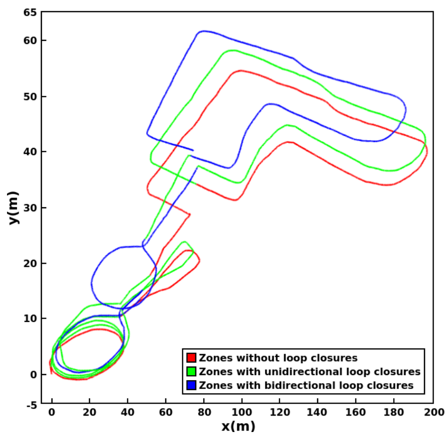


Fig. 7. Unidirectional and bidirectional loops detected by OpenABLE over the uncorrected map provided jointly with the Oxford New College dataset.

VI. Conclusions and Future Works

This paper has formally presented OpenABLE, whose main characteristics and functionalities have been extensively described along this work for facilitating the usage of this open-source toolbox contributed to the research community. A wide set of satisfactory results has validated its performance compared to other open toolboxes in the state of the art, such as OpenFABMAP or OpenSeqSLAM. In addition, several examples have demonstrated the practical application of our toolbox in different life-long visual localization problems for intelligent vehicles, which confirms that OpenABLE can be a useful resource for researchers in this topic.

The source code of OpenABLE is publicly available and it is completely customizable by its users in order to adapt it to their own requirements. This is a clear advantage for the research progress with respect to other localization methods that do not follow this kind of open-source philosophy.

In the future, the code will be maintained and updated when needed in the Github[5] repository of the project. Besides, other future works could be focused on including new powerful descriptors derived from Convolutional Neural Networks (CNNs) and their applicability in novel localization problems [26], such as street-view change detection [27] or other recent similar topics of interest for the ITS community.

[5]The code repository of OpenABLE in GitHub is available from: https://github.com/roberto-arroyo/OpenABLE

References

[1] H. Lategahn and C. Stiller, "Vision-only localization," *IEEE Trans. Intell. Transp. Syst (T-ITS)*, vol. 15, no. 3, pp. 1246–1257, 2014.

[2] B. Kitt, A. Geiger, and H. Lategahn, "Visual odometry based on stereo image sequences with RANSAC-based outlier rejection scheme," in *IEEE IV*, 2010, pp. 486–492.

[3] G. Ros, A. Sappa, D. Ponsa, and A. López, "Visual SLAM for driverless cars: A brief survey," in *Workshops of IEEE IV*, 2012.

[4] A. Bansal, H. Badino, and D. Huber, "Understanding how camera configuration and environmental conditions affect appearance-based localization," in *IEEE IV*, 2014, pp. 800–807.

[5] N. Sünderhauf, P. Neubert, and P. Protzel, "Are we there yet? Challenging SeqSLAM on a 3000 km journey across all four seasons," in *Workshops of IEEE ICRA*, 2013.

[6] W. Churchill and P. Newman, "Continually improving large scale long term visual navigation of a vehicle in dynamic urban environments," in *IEEE ITSC*, 2012, pp. 1371–1376.

[7] H. Badino, D. F. Huber, and T. Kanade, "Visual topometric localization," in *IEEE IV*, 2011, pp. 794–799.

[8] R. Arroyo, P. F. Alcantarilla, L. M. Bergasa, J. J. Yebes, and S. Gámez, "Bidirectional loop closure detection on panoramas for visual navigation," in *IEEE IV*, 2014, pp. 1378–1383.

[9] R. Arroyo, P. F. Alcantarilla, L. M. Bergasa, J. J. Yebes, and S. Bronte, "Fast and effective visual place recognition using binary codes and disparity information," in *IEEE/RSJ IROS*, 2014, pp. 3089–3094.

[10] R. Arroyo, P. F. Alcantarilla, L. M. Bergasa, and E. Romera, "Towards life-long visual localization using an efficient matching of binary sequences from images," in *IEEE ICRA*, 2015, pp. 6328–6335.

[11] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *IEEE CVPR*, 2012, pp. 3354–3361.

[12] M. Cummins and P. Newman, "FAB-MAP: Probabilistic localization and mapping in the space of appearance," *Int. J. Robot. Res. (IJRR)*, vol. 27, no. 6, pp. 647–665, 2008.

[13] M. Milford and G. F. Wyeth, "SeqSLAM: Visual route-based navigation for sunny summer days and stormy winter nights," in *IEEE ICRA*, 2012, pp. 1643–1649.

[14] A. J. Glover, W. Maddern, M. Warren, S. Reid, M. Milford, and G. F. Wyeth, "OpenFABMAP: An open source toolbox for appearance-based loop closure detection," in *IEEE ICRA*, 2012, pp. 4730–4735.

[15] X. Yang and K. T. Cheng, "Local difference binary for ultrafast and distinctive feature description," *IEEE Trans. Pattern Anal. Mach. Intell. (T-PAMI)*, vol. 36, no. 1, pp. 188–194, 2014.

[16] W. Maddern, A. Stewart, and P. Newman, "LAPS-II: 6-DoF day and night visual localisation with prior 3D structure for autonomous road vehicles," in *IEEE IV*, 2014, pp. 330–337.

[17] M. Calonder, V. Lepetit, M. Özuysal, T. Trzcinski, C. Strecha, and P. Fua, "BRIEF: Computing a local binary descriptor very fast," *IEEE Trans. Pattern Anal. Mach. Intell. (T-PAMI)*, vol. 34, no. 7, pp. 1281–1298, 2012.

[18] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *IEEE ICCV*, 2011, pp. 2548–2555.

[19] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *IEEE ICCV*, 2011, pp. 2564–2571.

[20] A. Alahi, R. Ortiz, and P. Vandergheynst, "FREAK: Fast retina keypoint," in *IEEE CVPR*, vol. 2, 2012, pp. 510–517.

[21] D. G. Lowe, "Object recognition from local scale-invariant features," in *IEEE ICCV*, 1999, pp. 1150–1157.

[22] H. Bay, T. Tuytelaars, and L. van Gool, "SURF: Speeded up robust features," in *ECCV*, vol. 3951, 2006, pp. 404–417.

[23] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE CVPR*, vol. 2, 2005, pp. 886–893.

[24] L. Caramazana, R. Arroyo, and L. M. Bergasa, "Visual odometry correction based on loop closure detection," in *Open Conference on Future Trends in Robotics (RoboCity16)*, 2016, pp. 97–104.

[25] M. Smith, I. Baldwin, W. Churchill, R. Paul, and P. Newman, "The New College vision and laser data set," *Int. J. Robot. Res. (IJRR)*, vol. 28, no. 5, pp. 595–599, 2009.

[26] R. Arroyo, P. F. Alcantarilla, L. M. Bergasa, and E. Romera, "Fusion and binarization of CNN features for robust topological localization across seasons," in *IEEE/RSJ IROS*, 2016.

[27] P. F. Alcantarilla, S. Stent, G. Ros, R. Arroyo, and R. Gherardi, "Street-view change detection with deconvolutional networks," in *Robotics: Science and Systems Conference (RSS)*, 2016.